3D Feature Based Tracking using SVM

Sehoon Kim , Seungjoon Choi , Sungjin Kim and Sangchul Won

Department of Electronic and Electrical Engineering, POSTECH, Pohang 790-784, Kyungbuk, Korea
(Tel: +82-54-279-5576; Fax: +82-54-279-8998; Email:robotics@postech.ac.kr)

Abstract: Tracking is one of the most important pre-required task for many application such as human-computer interaction through gesture and face recognition, motion analysis, visual servoing, augment reality, industrial assembly and robot obstacle avoidance. Recently, 3D information of object is required in realtime for many aforementioned applications. 3D tracking is difficult problem to solve because during the image formation process of the camera, explicit 3D information about objects in the scene is lost. Recently, many vision system use stereo camera especially for 3D tracking. The 3D feature based tracking(3DFBT) which is on of the 3D tracking system using stereo vision have many advantage compare to other tracking methods.
If we assumed the correspondence problem which is one of the subproblem of 3DFBT is solved, the accuracy of tracking depends on the accuracy of camera calibration. However, The existing calibration method based on accurate camera model so that modelling error and weakness to lens distortion are embedded. Therefore, this thesis proposes 3D feature based tracking method using SVM which is used to solve reconstruction problem.

Keywords: SVM, stereo calibration, 3D feature based tracking

1. Introduction

Computer vision is a very broad field of research that is intended for helping computers and robots to see. With the advances of digital camera and imaging technology, computer vision is playing an increasingly important role in automating tasks that involve visual sensory input.
Tracking is one of the most important pre-required task for many vision application such as human-computer interaction through gesture and face recognition, motion analysis, visual servoing, augment reality, industrial assembly and robot obstacle avoidance. Recently, 3D information of object is required in realtime (3D tracking) for many aforementioned applications. The three dimensional tracking is difficult problem to solve because during the image formation process of the camera, explicit 3D information about the scene or objects in the scene is lost. It is because that imaging process is basically projection of 3D scene onto 2D discrete image plane(in pixel). Even though it is possible to perform 3D tracking using a monocular image sequence [20]. These method have many weakness such that they have many constraints and hypothesis and are very sensitive to noise and so on. Therefore, many researchers who worked on monocular tracking before have now shifted to work on stereo tracking. It is because it is easier to "track" 3D objects in 3D space than in 2D image space. Especially, 3D feature based tracking(3DFBT) [18] [19] using stereo vision have advantage over other methods. The 3DFBT is able to run in real-time, it is because it use simple 2D feature tracking with both camera independently and the 2D feature tracking is more easy rather than 3D feature tracking. Then tracked 2D feature of both cameras reconstruct 3D information of object.
There are two computational subproblems associated with 3D tracking of target using stereo pair feature points on image. One is feature correspondence problem , another is camera calibration problem for reconstruction. It is difficult and important to solve the correspondence problem(temporal[15][14] and stereo[21][22][3]).
If we assumed the correspondence problem is solved, the difficulty of 3DFBT problem depends on the amount of a priori information of camera(s) available. If the intrinsic and extrinsic parameters of the camera(s) are determined by camera calibration method[1]-[7], then reconstruction in absolute coordinates is possible. However, the accuracy of the reconstructed structure is sensitive to the accuracy of these obtained parameters. However, the existing method based on accurate camera model so that modelling error and weakness to lens distortion are embedded. Therefore, this paper focused on stereo calibration method using Support Vector Machine(SVM)[8][9][12][13] to solve reconstruction problem which is a subproblem of 3D feature based tracking using stereo vision.And we also present real time 3D feature based tracking using proposed stereo camera calibration method using SVM. We propose a method that track the 3D position and pose of object from tracking 2D stereo corresponding features in realtime(3DFBT). The remainder of this paper is organized as follows: - section 2 provides a preliminary that will be used throughout this paper. - The proposed stereo calibration method is described in section 3, - Simulation and experiment results are shown in section 4 and then we conclude this paper.

2. Preliminary

2.1. Camera calibration and stereo reconstruction

The problem of camera calibration is to obtain the relation between image coordinate and world coordinate as (1).

$$\begin{bmatrix} sI_u \\ sI_v \\ s \end{bmatrix} = \begin{bmatrix} P_1 & P_2 & P_3 & P_4 \\ P_5 & P_6 & P_7 & P_8 \\ P_9 & P_{10} & P_{11} & P_{12} \end{bmatrix} \begin{bmatrix} W_x \\ W_y \\ W_z \\ 1 \end{bmatrix} \qquad (1)$$

where $\mathbf{I}$ means image coordinate and $\mathbf{P}$ is transform or relation matrix and $\mathbf{W}$ means world coordinate. This matrix $\mathbf{P}$ is typically consists of extrinsic and intrinsic parameters. Extrinsic parameters represent the position and pose of camera

with respect to specific world coordinate. In multi-camera system, extrinsic parameters also describe the relationship between two cameras. Intrinsic parameters represent inherent properties of the camera optics, including the focal length, the image center, the image scaling factor and lens distortion coefficients. Generally, the objective of the camera calibration procedure is to determine values for these parameters(extrinsic and intrinsic) based on image observations of known 3-D target points[1]-[3][4]-[6]. And of course, the camera calibration can be thought as determining transformation matrix or relation matrix **P**.
Reconstruction problem[1]-[3] is how to reconstruct 3D position (x_i, y_i, z_i) of a point from corresponding 2D image point (u_i, v_i). Methods for reconstruction can divide into two classes: using stereopsis and using motion. In this section, simple method using stereopsis is described briefly.

The projection matrix **P** is known if camera calibration

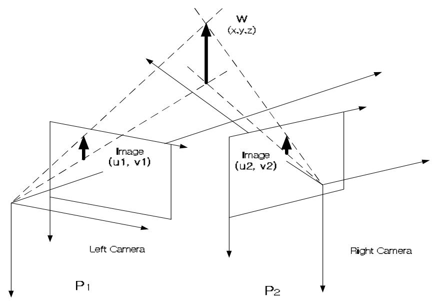

Fig. 1. Stereopsis

process is performed before. If stereo camera are used, **P** matrix for each can be used to determine 3D position of target object as shown at Figure 1 using Linear Triangulation methods [2]. The Linear Triangulation method is following. The equation (2) can be derived from two calibrated camera

$$I_1 = \mathbf{P}_1 \cdot W \quad , \quad I_2 = \mathbf{P}_2 \cdot W \tag{2}$$

where I_1, P_1, I_2, P_2 are image coordinate data (matched) and projection matrix of camera 1 and camera 2 respectively, and $I_1 = (u_1, v_1)$, $I_2 = (u_2, v_2)$, and W is corresponding world coordinate 3D data point, $W = (X, Y, Z, 1)$. The objective of reconstruction is to find world coordinate point from obtained stereo pair image point. The solution can be determined by solving the equation (3) for W.

$$\begin{bmatrix} u_1\mathbf{P}_1^3 & - & \mathbf{P}_1^1 \\ v_1\mathbf{P}_1^3 & - & \mathbf{P}_1^2 \\ u_2\mathbf{P}_2^3 & - & \mathbf{P}_2^1 \\ v_2\mathbf{P}_2^3 & - & \mathbf{P}_2^2 \end{bmatrix} \cdot W = 0 \tag{3}$$

where $\mathbf{P}_i^j$ are the i-th row of $\mathbf{P}_i$.

2.2. Optical flow

A critical but difficult problem for tracking, obviously, is constructing correspondences. In 2D tracking, to find temporal correspondences means tracking itself. In 3D tracking, to find temporal and stereo(between two cameras) correspondences make the 3D tracking possible. The optical flow[1][14][15] indicates that an image point moves from here to there(correspondence) temporally or spatially(stereo). We denote an image by I(x,y,t), and the velocity of an image pixel p=[x,y]^T is

$$\boldsymbol{v}_p = \dot{p} = [v_x, v_y]^T = [dx/dt, dy/dt]^T$$

Assuming the intensity of p keeps the same during dt, i.e.,

$$I(x + v_x dt, y + v_y dt, t + dt) = I(x, y, t)$$

If the brightness changes smoothly with x,y and t, we expend the left-hand-side by Taylor series:

$$I(x, y, t) + \frac{\partial I}{\partial x} v_x dt + \frac{\partial I}{\partial y} v_y dt + \frac{\partial I}{\partial t} dt + O(dt^2) = I(x, y, t)$$

If we ignore high order term $O(dt^2)$, we have (4).

$$\begin{aligned} \frac{\partial I}{\partial x} v_x dt + \frac{\partial I}{\partial y} v_y dt + \frac{\partial I}{\partial t} dt = 0 \\ i.e. \quad \nabla I \cdot \boldsymbol{v}_p + \frac{\partial I}{\partial t} dt = 0 \end{aligned} \tag{4}$$

where $\nabla I = [\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}]^T$ is image gradient at pixel p, which can be obtained from images. And also can be obtained from images easily, We call this equation *optical flow constraint equation*.
Apparently, for each pixel, we have only one constraint equation, but we need to solve two unknowns, i.e., v_x and v_y, which means that we can't determine optical flow uniquely only from such optical flow constraint equation. Given this constrain equation, we can only determine the normal flow, i.e., the flow along the direction of image gradient, but we can not determine those flow on the tangent direction of isointensity contour, i.e., the direction perpendicular to the image gradient. This problem is so called aperture problem.

2.3. Support vector machine

Support Vector Machines(SVM)[8][9] proposed by Vapnik early in 90's have become a subject of intensive study in statistical learning theory. The main idea of support vector machine is to find the hyperplane which have maximum margin.SVMs are a system for efficiently training the linear learning machines in the kernel-induced feature space. Initially, it was designed to solve pattern recognition problems [11], where in order to find a decision rule with good generalization capability and a small subset of the training data(called the support vectors). Recently, SVM has also been applied to various fields successfully such as object recognition [10], pattern recognition[[11]], function regression[[12]][[13]]. When SVM is employed to tackle the problems of function approximation and regression estimation, the approaches are often referred to as the support vector regression(SVR) [12][13]. The function approximation using SVR is very effective, especially for the case of having a high-dimensional input space because the SVM is a universal approach for solving the problem of multidimensional function estimation.
In a regression problem, the goal is to find a fit to the input-output data points, $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)\}$,$(\mathbf{x} \in$

$R^m, y \in R$), sampled from unknown function. SVM can also be applied to regression problem by the introduction of an alternative loss function[12]. These loss function are shown in Figure 2. A support vector machine is an approximation

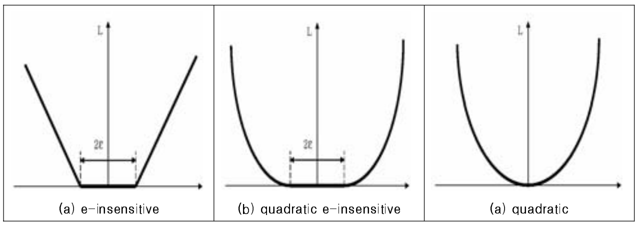

Fig. 2. Loss functions

of the form as (5).

$$\boldsymbol{f}(\mathbf{x}) = \sum_{i=1}^{n} w_i \phi_i(\mathbf{x}) + b \tag{5}$$

where $\phi : \mathbf{X} \rightarrow \mathbf{F}$ is a nonlinear map from the input space to some feature space. Then the non-linear SVR solution, using an ϵ-insensitive loss function, is given by,

$$\max_{\alpha,\alpha^*} W(\alpha, \alpha^*) = \max_{\alpha_i,\alpha_i^*} \sum_{i=1}^{n} \alpha_i^*(y_i - \epsilon) - \alpha_i(y_i + \epsilon)$$

$$+ \sum_{i=1}^{n}\sum_{j=1}^{n} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(\mathbf{x}_i, \mathbf{x}_j) \tag{6}$$

$$subject\ \ to \quad \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0, 0 \leq \alpha_i, \alpha_i^* \leq C \tag{7}$$

where K(·) is a kernel function which is inner product of two feature space, i.e. $<\!\phi(x_i), \phi(x_j)\!>$. Kernels that usually used are represented at Table 1 . Solving equation (6) with con-

Table 1. Kernels

	Linear	polynomial	RBF
$\mathbf{K}(x, x')$	$<x, x'>$	$(<x, x'> +c)^p$	$exp(-\frac{\|x-x'\|^2}{2\delta^2})$
	ERBF	Tangential	Fourier
$\mathbf{K}(x, x')$	$exp(-\frac{\|x-x'\|}{2\delta^2})$	$tanh(\rho <x, x'> +\varrho)$	$\frac{sin(N+\frac{1}{2})(x-x')}{sin(\frac{1}{2}(x-x'))}$

straints equation (7) determines the Lagrange multipliers, α_i, α_i^*, and the regression function is given by (8),

$$\boldsymbol{f}(\mathbf{x}) = \sum_{SVs} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b \tag{8}$$

It should be known that '*Training an SVM*' can be expressed as solving a quadratic program.

3. 3D feature based tracking using SVM

3.1. Stereo camera calibration using SVM

In this paper, stereo calibration using SVM for reconstruction is proposed. It is proposed to solve the reconstruction problem which is one of the subproblem of 3DFBT. We use stereo pair image as input and corresponding world coordinate point as output to train SVM. Three SVMs are used for learning the relation(or nonlinear function) between stereo pair image and world coordinate data. The proposed stereo camera calibration method for 3D reconstruction using SVM is shown in Figure 3. where $(\mathbf{u}_{1i}, \mathbf{v}_{1i})$ and

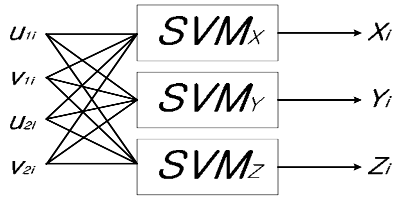

Fig. 3. Proposed stereo calibration system using SVM

$(\mathbf{u}_{2i}, \mathbf{v}_{2i})$ are image coordinate data (in pixel) which are generated by the same world coordinate point $(\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i)$. And $(\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i)$ are actual world coordinate data which are mapped as $(\mathbf{u}_{1i}, \mathbf{v}_{1i})$ and $(\mathbf{u}_{2i}, \mathbf{v}_{2i})$ on two images. If we have N known input-output data,$\{(\mathbf{u}_{11}, \mathbf{v}_{11}, \mathbf{u}_{21}, \mathbf{v}_{21}, \mathbf{X}_1, \mathbf{Y}_1, \mathbf{Z}_1), \ldots, (\mathbf{u}_{1N}, \mathbf{v}_{1N}, \mathbf{u}_{2N}, \mathbf{v}_{2N}, \mathbf{X}_N, \mathbf{Y}_N, \mathbf{Z}_N)\}$, we can train SVMs with these N input-output data. After three SVMs are trained(or solve the quadratic program) with these input-output data, we can determine *SVs(support vectors)* and $\alpha - \alpha^*$ of equation (8) for each SVM. Then we can reconstruct 3 Dimensional position with equation (9) as follows.

$$\boldsymbol{X} = \sum_{SVx} \beta_{xi} K(\mathbf{s}_{xi}, \mathbf{x}) + b \;\;, \quad \boldsymbol{Y} = \sum_{SVy} \beta_{yi} K(\mathbf{s}_{yi}, \mathbf{x}) + b$$

$$\boldsymbol{Z} = \sum_{SVz} \beta_{zi} K(\mathbf{s}_{zi}, \mathbf{x}) + b \tag{9}$$

where β is $(\alpha - \alpha^*)$, $\mathbf{x}$ is (u_1, v_1, u_2, v_2) which is image coordinate data of both cameras corresponding world coordinate data *(X,Y,Z)*, $\mathbf{s}$ is support vectors, subscript x, y, z represent that it is for (SVM_x, SVM_y, SVM_z) of Figure 3. After we train the SVM on a range of interest, the SVM can determine the world coordinate point for any matched pair of image points on a range of interest from (9).

If correspondence problem is solved, i.e. it is apparent that (u_1, v_1, u_2, v_2) are projection of same world coordinate point *(X,Y,Z)*, the accuracy of 3DFBT depends on the accuracy of camera calibration. However, The existing method based on accurate camera model so that modelling error and weakness to lens distortion are embedded. Therefore some learning method using neural network have been developed[7]. Similarly to neural network method, it should be noted that this approach is different form conventional camera calibration techniques in the sense that *no extrinsic and intrinsic camera parameters are found*. Instead, the system is trained so that it can determine 3D position of objects w.r.t world coordinate directly. And this method also learns internal characteristics of camera such as lens distortions and cell size, so there are no need to correct distortions and to know any camera parameters. Simply speaking, trained SVM represent stereo camera itself.

And it shows some advantages compare to neural network. One is that it's robust property against noise(good generalization property : prevent to over-fit) due to its maximal margin property and that the number of free parameters in the function approximation scheme is equal to the number of support vectors. Another is that the SVM approach are convex hence have no local minima. Finally, the capacity of a system is controlled by parameters does not depend on the dimensionality of the space.

3.2. Real-time 3D-FBT using SVM

Our system uses stereo views of interest point features of object for the 3D object tracking in realtime. Overall system is shown in Figure 4. The 3D object tracking system operates

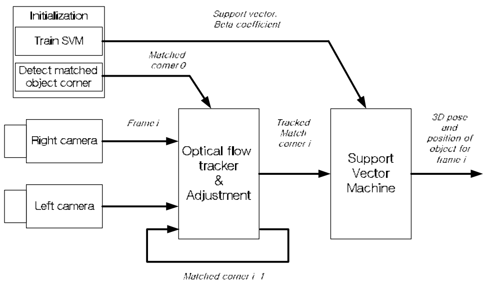

Fig. 4. Proposed real time 3D feature based tracker

in three phases: an initialization phase, a 2D feature tracking phase and reconstruction phase.

Initial phase is made up of three components: initial feature selection, determining initial correspondences and training SVM. The feature extraction is not a focus of this paper so we would like to avoid discussing the details. Target features are selected manually. We use a publicly available implementation of the work by Tomasi and Kanade [17][16] to extract corner features in our experiments. Then initial correspondence of selected features are determined by preknown structure of target. SVM is trained for stereo camera to learn the relation between 2D image pair data and 3D world coordinate data with chess board pattern as shown in Figure 5.

The second phase consist of two components: 2D feature

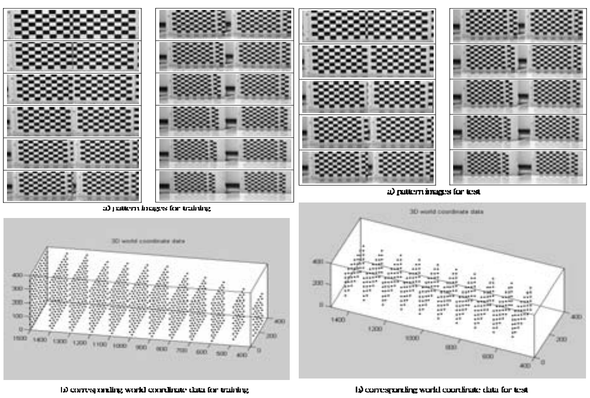

Fig. 5. Training data for SVM

tracking and adjustment of tracked feature. Feature points are then tracked in each camera image independently using a pyramidal implementation of the Lucas-Kanade optical flow algorithm[1][14]. Feature tracking is performed by computing optical flow from frame *i-1* to frame *i*. 2D feature tracking techniques can much more reliably establish feature matching(tracking) between image frames than stereo matching. After computing optical flow, i.e. feature points of frame i are tracked from frame *i-1*, feature points are adjusted to strong corner by Tomasi and Kanade [17][16] so that it can prevent accumulation of error. If tracking is performed successfully it maintain feature correspondences. These features are feed back for frame *i+1*.

In third phase, the tracked features from phase 2 are used as input for trained SVM to reconstruct 3D information of 3D object. From the initialization phase, SVs and coefficients β are delivered for each support vector machine. Using these parameters, it can compute corresponding 3D positions of matched feature pairs easily from equation 9. Due to simplicity of equation (9), proposed 3 Dimensional feature based tracking using SVM can run in real time. The real test data for testing reconstruction error is shown in Figure 5. And the results of simulation and experiment shows that it can reconstruct more accurate 3D object feature position. So proposed method is more accurate real time 3D feature based tracking method than others. The real time 3D feature tacking result and comparison of reconstruction error will be shown in next section.

4. Results and conclusion

4.1. Simulation results of stereo camera calibration

For the simulation, a simple pinhole camera model with radial and tangential lens distortion is used as virtual cameras. Intrinsic and extrinsic parameters of virtual camera are shown in Table 2, Figure 6 respectively. The distance

Table 2. Intrinsic parameters of left and right virtual camera

f	$\mathbf{D}_u$	$\mathbf{D}_v$	**a**	$[u_0, v_0]$
25[mm]	136.7143	142.8565	0.9570	[650,515]

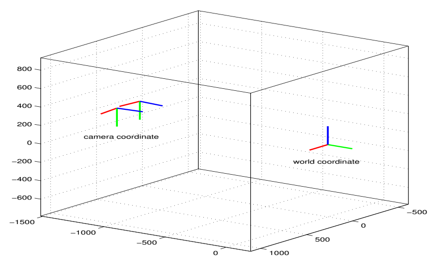

Fig. 6. Pose and position of camera w.r.t world coordinate

between two cameras is 250[mm]. The RBF kernel is used. The training and test data for output, i.e world coordinate data, are shown at Figure 7.

The simulation procedure is as follows. First of all, world coordinate data(3 dimensional location of each corners of Figure 7) are projected to two virtual cameras. Then these projected data are quantized to integer(due to pixel unit), and an gaussian noise and lens distortion are added. And known

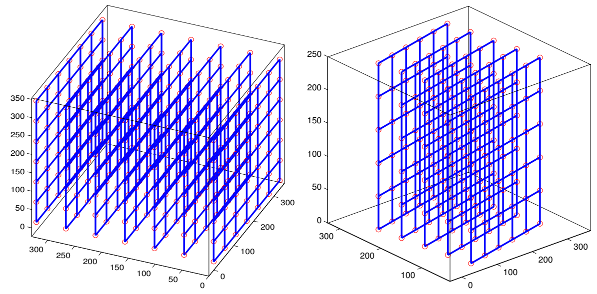

Fig. 7. Training and Test world coordinate data

3D world coordinate data(output) shown in Figure 7 and corresponding stereo pair data(virtual camera image) are used to train three SVMs(Figure 3). After training, test stereo pair data are obtained through same procedure of training. If we use test image data obtained above as input of trained SVMs, then the output of SVMs represents 3D reconstruction. To measure the reconstruction error Root Mean Square Error (10) is used.

$$Err_{reconst} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(X_i - \hat{X}_i)^2 + (Y_i - \hat{Y}_i)^2 + (Z_i - \hat{Z}_i)^2} \quad (10)$$

where i-th test world coordinate is (X_i, Y_i, Z_i), i-th output of SVM is $(\hat{X}_i, \hat{Y}_i, \hat{Z}_i)$, N is number of test data points.
The simulation are performed at varying variance of gaussian noise with 5 types of lens distortion as shown at Table 3. The nonlinear calibration technique(minimizing cost

Table 3. 5 types of lens distortion for simulation

	k_1	k_2	p_1	p_2
Type 1	0	0	0	0
Type 2	-2.8e-3	4.5e-5	-9.0e-4	2.0e-5
Type 3	-5.7e-3	6.3e-5	0	0
Type 4	-4.1e-3	5.3e-5	-3.5e-3	4.0e-5
Type 5	-6.3e-3	6.1e-5	-7.5e-3	5.0e-5

function (10) using result of standard DLT as initial guess [4][6](DLT & Optimisation)), and multi-step calibration method(estimate camera parameters and lens distortion by LM algorithm and then correct the distortion [6](Heikkila)) and Tsai's Method [5](Tsai) are used to compare reconstruction error. Simulation results are shown at Figure 8. From results of simulation, If there is no lens distortion traditional calibration method shows good result. But it is impossible and there is no model to represent the camera exactly and no camera without lens distortion. And Heikkila method shows good result for all types of simulation , but it requires pre-known camera parameters. As shown results of simulation, proposed method shows good result without regard to lens distortion.

4.2. Experiment Results of Real-time 3D Tracking using SVM

For experiment two SONY XC-75CE monochrome camera are used. The RMSE of reconstruction with real imageis shown at Table 4. The realtime tracking results are shown at Figure (9)(10).

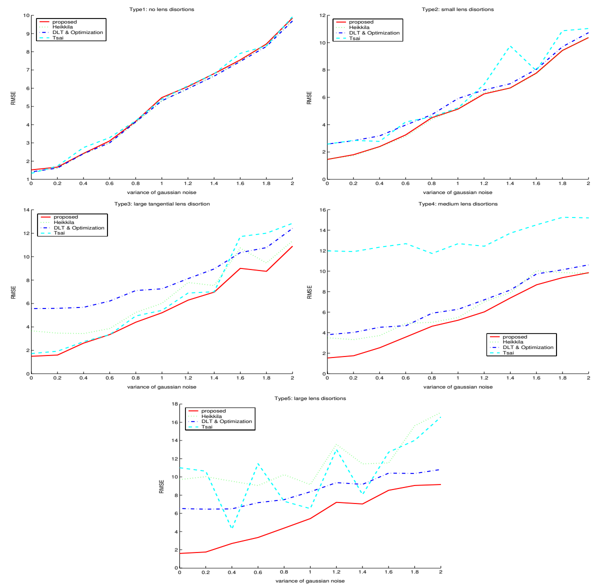

Fig. 8. Simulation result and comparison

Table 4. RMSE for reconstruction with real data

	Proposed	Heikkila	DLT&Opt.	Tsai
RMSE	3.00318295	3.0358397	3.211661	3.2427883

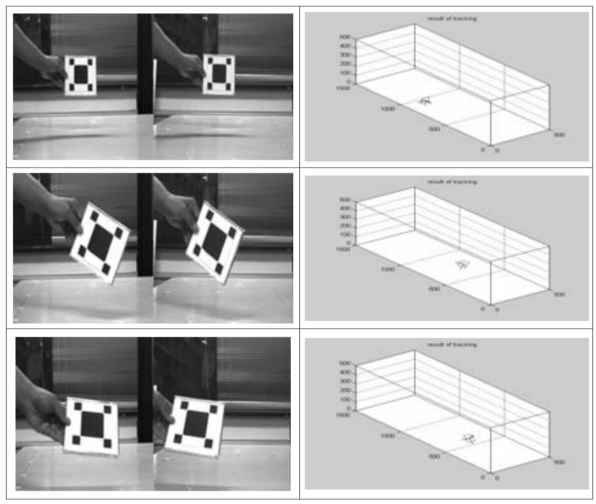

Fig. 9. Result of experiment

4.3. conclusion

A new 3 Dimensional Feature Based Tracking method using SVM was presented. The 3DFBT is a 3D tracking method that uses 2D feature tracking and stereopsis. This kind of method for 3D tracking has many advantages over other tracking method. Because 2D feature tracking is much more easier than 3D tracking. This method has two main subproblems: correspondences and reconstruction. If we assumed correspondences problem is solved, the accuracy of the 3DFBT is depend on the accuracy of the camera calibration. However, existing camera calibration methods have embedded weakness because these method based on exact camera model and because there is no camera model to represent camera exactly. So, a new stereo calibration method us-

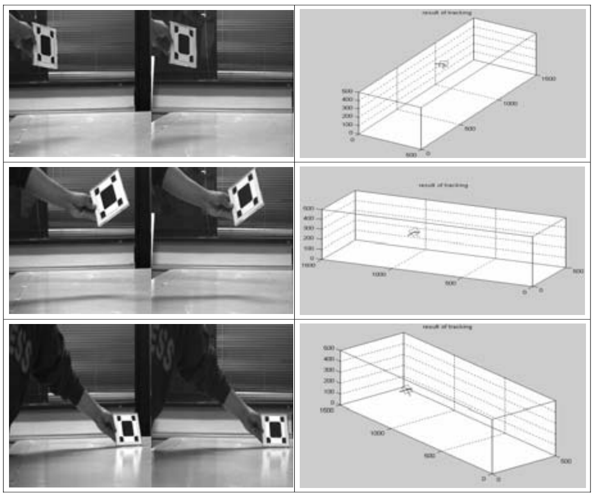

Fig. 10. Result of experiment[cont.]

ing SVM was presented in this paper. We demonstrated how to make the SVM learn the relationship or nonlinear function between matched stereo pair image and corresponding world coordinate. The proposed method is like neural network method (intrinsic camera calibration method). But, the proposed method has several advantage against traditional camera calibration methods and neural network methods. a) The SVM has robust property against noise ,so this method are robust against noise(generalization problem). b) This method has no local minimum solution because training SVM is to solve quadratic problem. c) There are no need pre-known camera parameters and no need to estimate lens distortion and to correct them. d)It is more accurate without regard to lens distortion because trained SVM imitate stereo camera itself. The reconstruction using SVM tested with simulation and experiment with real image. The result in last section show that proposed method more accurate without regard to lens distortion over other methods. The experiment for propose 3D Feature Based Tracking using SVM was tested and shown in last section. The result of experiment shows that proposed method run in real time and is more accurate other methods.

References

[1] O. Faugeras , *"Three dimensional computer vision : a geometric viewpoint"*, MIT Press, 1993.

[2] R. Hartley, A. Zisserman, *"Multiple view geometry in computer vision"*, Cambridge University Press, 2000.

[3] D. A. Forsyth , J. Ponce, *"Computer Vision: A Modern Approach"*, Prentice Hall, 2002.

[4] Abdel-Aziz and Karara, *"Direct linear transformation into object space coordiantes in close-range photo grammetry"*, Proc. Symposium on Close-Range Photo grammetry, Urbana, Illinois, p.1-18, 1971.

[5] R.Y. Tsai, *"A versatile camera calibration technique for high accuracy 3d machin vision metrology using off-the-shelf TV camera and Lenses"*, IEEE Journal of Robotics and Automation, vol. RA-3, NO.4, August, 1987.

[6] J. Heikkila and O. silven, *"A four-step camera calibration procedure with implicit image correction"*, Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 1997.

[7] Q. Memon and S. Khan, *"Camera calibration and three-dimensional world reconstruction of stereo-vision using neural networks"*, International Journal of System Science, 2001, Vol. 32, NO. 9, p. 1155-1159, 2001.

[8] C. Cortes and V. Vapnik, *"Support vector networks"*, Machine Learning, Vol. 20, p. 273-297, 1995.

[9] N. Cristianini and J. Shawe-Taylor, *"An introduction to support vector machines (and other kernel-based learning methods"*, Cambridge University Press, 2000.

[10] M. Pontil and A. Verri, *"Support vector machines for 3D object recognition"*, IEEE Transactions on PAMI, Vol. 20, NO. 6, June, 1998.

[11] C. Burges, *"A tutorial on support vector machines for pattern recognition"*, Data Mining Knowledge Discovery, Vol. 2, NO. 2, 1996.

[12] A.J. Smola, B. Scholkopf, *"A tutorial on support vector regression"*, Royal Holloway College, London, U.K., NeuroCOLT Tech. Rep. TR-1998-030, 1998.

[13] Steve Gunn, *"Support vector machiens for classification and regression"*, ISIS Technical Report ISIS-1-98, Image Speech and Intelligent Systems Research Group, University of Southapton, May, 1998.

[14] B.D. Lucas, T. Kanade, *"An iterative image registration technique with an application to stereo vision"*, In Proc. Int. Joint Conf. on AI, pages 674-679, 1981.

[15] B. K. P. Horn, B. G. Schunck, *"Determining Optical Flow"*, Artificial Intelligence, Vol. 17, pp. 185-204, 1981.

[16] J. Shi, C. Tomasi, *"Good Features to track"*, In Proceedings of the IEEE Conference on CVPR, pages 593-600, Seattle, WA, June 1994.

[17] C. Tomasi, T. Kanade, *"Detection and tracking of point features"*, Technical report. CMU-CS-91-132, Carnegie Mellon University, April, 1991.

[18] T. K. Cheng, L. Kitchen, *"Multi-agent real-time 3D tracking in robot vision"*, In Proceedings of ACSC'95, Eighteenth Australasian ComputerScience Conference, pages 80–89, Adelaide, Australia, February 1995.

[19] C. Y. Tang, Y. P. Hung, *"A 3D Feature-Based tracker for Multiple Object Tracking"*, In Proceedings of National Science Council, Republic of China, Vol. 23, No. 1, pp. 151-168. 1999.

[20] D. Koller, K. Daniilidis, *"Model-based object tracking in monocular image sequences fo road traffic secnes"*, IJCV, Vol. 10, No. 3, pp. 257-281. 1993.

[21] Y. Ohta, T. Kanade, *"Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming"*, IEEE Transactions on PAMI, Vol. 7, NO. 2, pp 139-154, 1985.

[22] T. Kanade, M. Okutomi, *"A stereo matching algorithm with an adaptive window: theory and experiment"*, IEEE Transactions on PAMI, Vol. 16, NO. 9, pp 920-932, September, 1994.