An Active Sensing Method Using Estimated Errors for Multisensor Fusion Systems

Toshiharu Mukai and Masatoshi Ishikawa

Faculty of Engineering
University of Tokyo

Introduction

Intelligent robots in unknown environment should have

Multiple sensors,

Active sensing method.

How to move sensors This presentation

How to integrate data from the sensors

Parallel Kalman Filter (H.R. Hashemipour *et al.*)

Contents

1. To propose a method to select sensor positions to minimize estimated errors,
2. The method is expanded to also achieve association correctly,
3. Application 1: Target tracking using 2 CCD cameras,
4. Application 2: Visual and tactile data fusion.

Definition of the object state and observation

$$\boldsymbol{x}(t+1) = F(t)\boldsymbol{x}(t) + G(t)\boldsymbol{w}(t)$$
$$\boldsymbol{z}(t) = \boldsymbol{h}(\boldsymbol{x}(t), t) + \boldsymbol{v}(t),$$

t: time,

$\boldsymbol{x}(t)$: state vector (we want to estimate),

$\boldsymbol{z}(t)$: observation vector,

$\boldsymbol{h}(\cdot)$: nonlinear function,

$\boldsymbol{w}(t), \boldsymbol{v}(t)$: zero mean noises with variance $Q(k)$ and $R(k)$, respectively.

Extended Kalman filter

Prediction:

$$\hat{\boldsymbol{x}}(t+1|t) = F(t)\hat{\boldsymbol{x}}(t|t)$$

$$P(t+1|t) = F(t)P(t|t)F^T(t) + G(t)Q(t)G^T(t)$$

Update:

$$\hat{\boldsymbol{x}}(t+1|t+1) = \hat{\boldsymbol{x}}(t+1|t) + W(t+1)[\boldsymbol{z}(t+1) - \boldsymbol{h}(\hat{\boldsymbol{x}}(t+1|t), t+1)]$$

$$P^{-1}(t+1|t+1) = P^{-1}(t+1|t) + J_h^T(t+1)R^{-1}(t+1)J_h(t+1)$$

$$\text{where, } W(t+1) = P(t+1|t+1)J_h^T(t+1)R^{-1}(t+1)$$

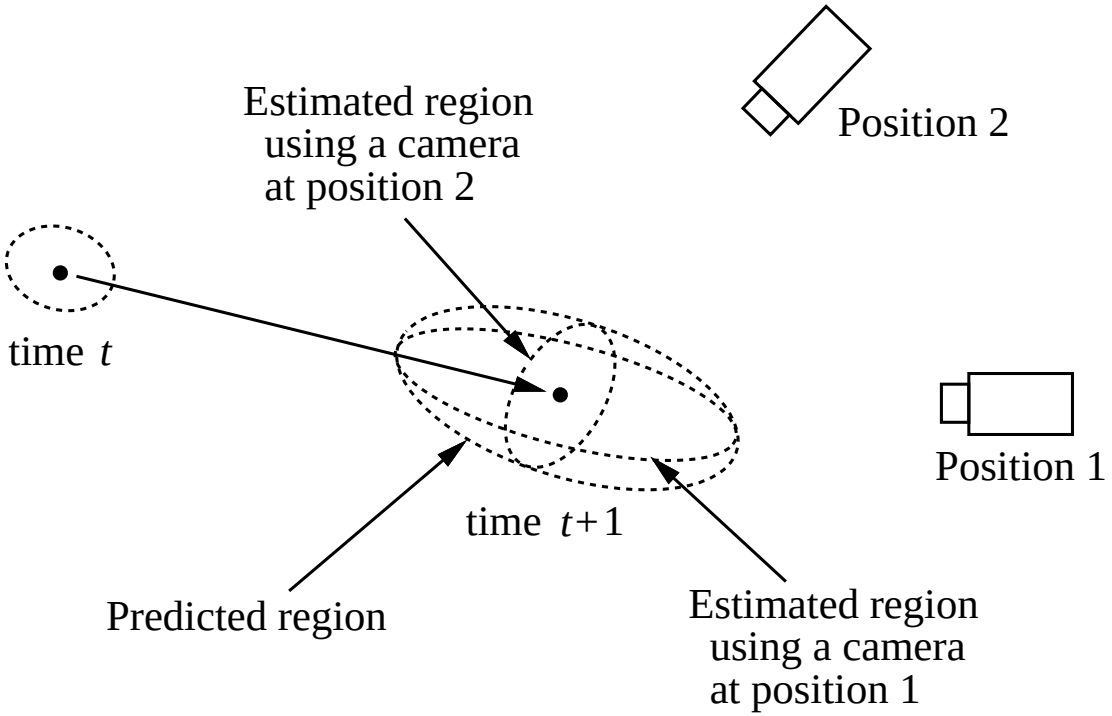
$$\text{where } J_h(t+1) = \frac{\partial \boldsymbol{h}(\boldsymbol{x}, t+1)}{\partial \boldsymbol{x}}|_{\hat{\boldsymbol{x}}(t+1|t)}$$

$\hat{\boldsymbol{x}}(t|t)$: the estimate of $\boldsymbol{x}(t)$

$P(t|t) = E[(\hat{\boldsymbol{x}} - \boldsymbol{x})(\hat{\boldsymbol{x}} - \boldsymbol{x})^T]$

You can calculate $P(t+1|t+1)$ without observing $\boldsymbol{z}(t+1)$.

Concept for selecting sensor position



Target tracking

Minimizing the error of the estimate

Moving sensor $\rightarrow$ $\boldsymbol{h}(\cdot)$ changes $\rightarrow$ P also changes

$$\Downarrow$$

$\boldsymbol{h}(\cdot, \boldsymbol{x}_s), P(\cdot, \boldsymbol{x}_s)$ is used, where $\boldsymbol{x}_s$ is the sensor position

At time t, select sensor position $\boldsymbol{x}_s$ to minimize

$$\mathcal{R}(\boldsymbol{x}_s) = \det P(t+1|t+1; \boldsymbol{x}_s).$$

$\mathcal{R}(\boldsymbol{x}_s)$ is the product of all the eigen values of P

$\Longrightarrow \mathcal{R}(\boldsymbol{x}_s)$ corresponds to the error of the estimate.

Multiple sensors and objects

Subscript i ($i = 1, \ldots N$): Object number

Subscript j ($j = 1, \ldots M$): Sensor number

$\boldsymbol{x}_s = [\boldsymbol{x}_{s1}^T, \boldsymbol{x}_{s2}^T, \ldots, \boldsymbol{x}_{sM}^T]^T$

$$\boldsymbol{x}_i(t+1) = F_i(t)\boldsymbol{x}_i(t) + G_i(t)\boldsymbol{w}_i(t),$$
$$\boldsymbol{z}_{j;i}(t) = \boldsymbol{h}_j(\boldsymbol{x}_i(t), t) + \boldsymbol{v}_{j;i}(t).$$

Select sensor positions to minimize each

$$\mathcal{R}_i(\boldsymbol{x}_s) = \det P_i(t+1|t+1; \boldsymbol{x}_s) \qquad (i = 1, \ldots N),$$

where $P_i(t+1|t+1)$ is calculated like the case of only one sensor and one object.

Nearest neighbor data association algorithm

Associate sensor j's output $\tilde{\boldsymbol{z}}_{j;k}(t+1)$ with $\boldsymbol{z}_{j;i}(t+1)$ which minimize the Mahalanobis generalized distance between $\tilde{\boldsymbol{z}}_{j;k}(t+1)$ and $\hat{\boldsymbol{z}}_{j;i}(t+1)$:

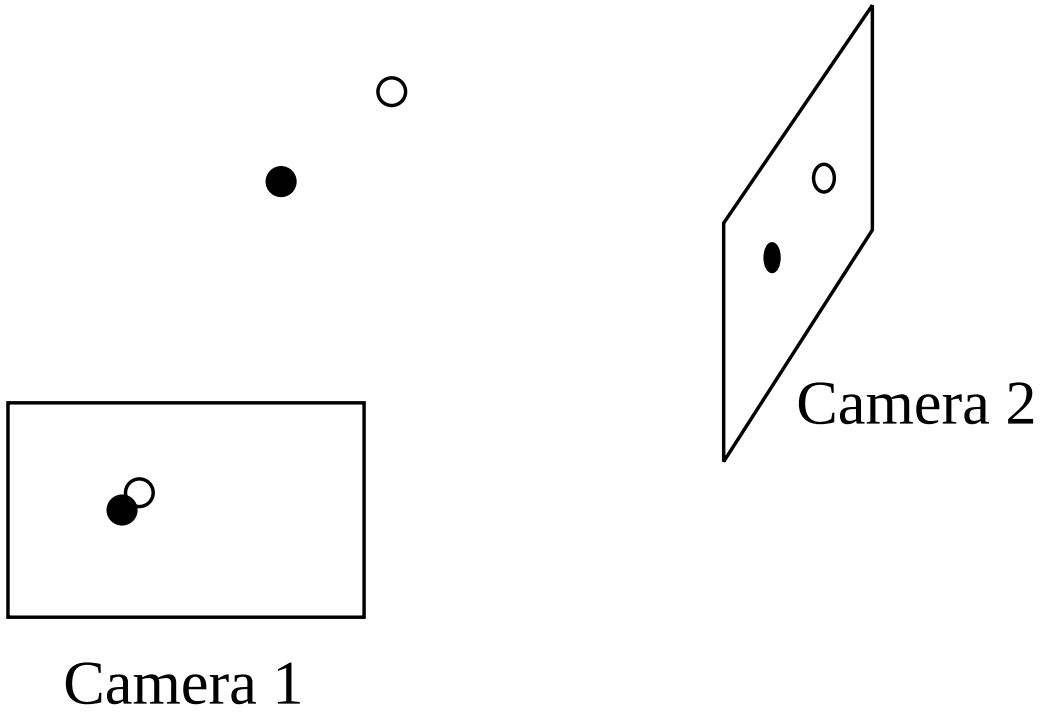
$$\mu_{j;ik} \equiv (\hat{\boldsymbol{z}}_{j;i}(t+1|t) - \tilde{\boldsymbol{z}}_{j;k}(t+1))^T S_{j;i}(t+1|t)^{-1} \cdot(\hat{\boldsymbol{z}}_{j;i}(t+1|t) - \tilde{\boldsymbol{z}}_{j;k}(t+1)),$$

where

$$\hat{\boldsymbol{z}}_{j;i}(t+1|t) = \boldsymbol{h}_j(\hat{\boldsymbol{x}}_i(t+1|t)),$$

$$S_{j;i}(t+1|t) = \frac{\partial \boldsymbol{h}_j}{\partial \boldsymbol{x}}|_{\hat{\boldsymbol{x}}_i(t+1|t)} P_i(t+1|t) \left(\frac{\partial \boldsymbol{h}_j}{\partial \boldsymbol{x}}|_{\hat{\boldsymbol{x}}_i(t+1|t)} \right)^T.$$

Concept for selecting sensor position to achieve association correctly



Target tracking of two objects

Making good association possible

Select $\boldsymbol{x}_s$ to maximize the Mahalanobis generalized distance of each pair of sensor outputs:

$$\begin{aligned} D_{j;i_\alpha,i_\beta}(\boldsymbol{x}_s) \equiv & (\hat{\boldsymbol{z}}_{j;i_\alpha}(t+1|t) - \hat{\boldsymbol{z}}_{j;i_\beta}(t+1|t))^T \\ & \cdot (S_{j;i_\alpha} + S_{j;i_\beta})^{-1} \\ & \cdot (\hat{\boldsymbol{z}}_{j;i_\alpha}(t+1|t) - \hat{\boldsymbol{z}}_{j;i_\beta}(t+1|t)) \\ & (i_\alpha, i_\beta = 1, \ldots, N; i_\alpha \neq i_\beta). \end{aligned}$$

Definition of evaluation function

To achieve both good estimation and association, maximize

$$\mathcal{I}(\boldsymbol{x}_s) \equiv \mathcal{E}^p_{\mathrm{All}}(\boldsymbol{x}_s) + \lambda \mathcal{E}^a(\boldsymbol{x}_s),$$

where

$$\mathcal{E}^p_{\mathrm{All}}(\boldsymbol{x}_s) \equiv \frac{1}{\mathcal{R}_{\mathrm{All}}(\boldsymbol{x}_s)} = \frac{1}{\sum_{i=1}^{N} \mathcal{R}_i(\boldsymbol{x}_s)},$$

$$\mathcal{E}^a(\boldsymbol{x}_s) \equiv -\sum_{j=1}^{M} \frac{1}{D_j(\boldsymbol{x}_s)} \qquad \text{with } D_j(\boldsymbol{x}_s) \equiv \min_{i_\alpha, i_\beta} D_{j;i_\alpha,i_\beta}(\boldsymbol{x}_s)$$

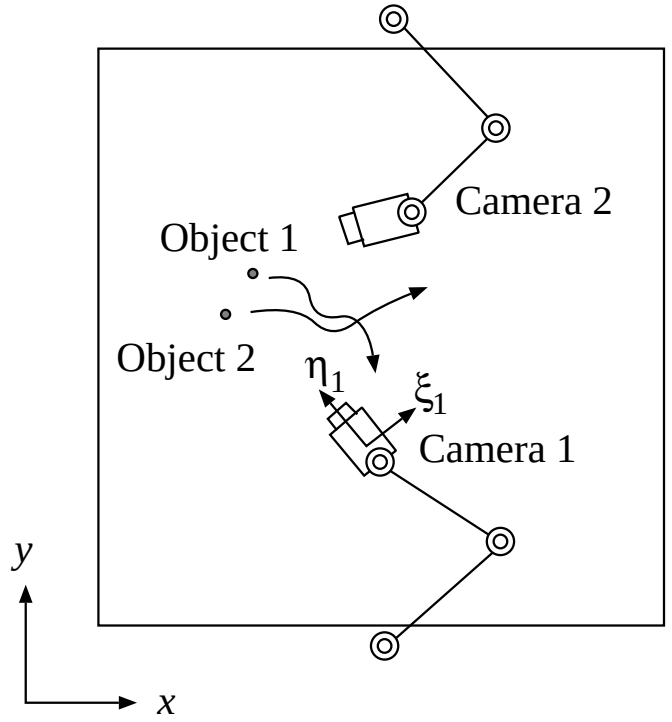
Integration of outputs of multiple sensors

Parallel Kalman filter

$$\hat{\boldsymbol{x}}_i(t+1|t+1)$$
$$= P_i(t+1|t+1)[P_i^{-1}(t+1|t)\hat{\boldsymbol{x}}_i(t+1|t)$$
$$+\sum_{j=1}^{N}\left(\frac{\partial \boldsymbol{h}_j}{\partial \boldsymbol{x}}|_{\hat{\boldsymbol{x}}_i(t+1|t)}\right)^T R_{j;i}^{-1}(t+1)\{\boldsymbol{z}_{j;i}(t+1)$$
$$-\boldsymbol{h}_j(\hat{\boldsymbol{x}}_i(t+1|t)) + \frac{\partial \boldsymbol{h}_j}{\partial \boldsymbol{x}}|_{\hat{\boldsymbol{x}}_i(t+1|t)}\hat{\boldsymbol{x}}_i(t+1|t)\}]$$

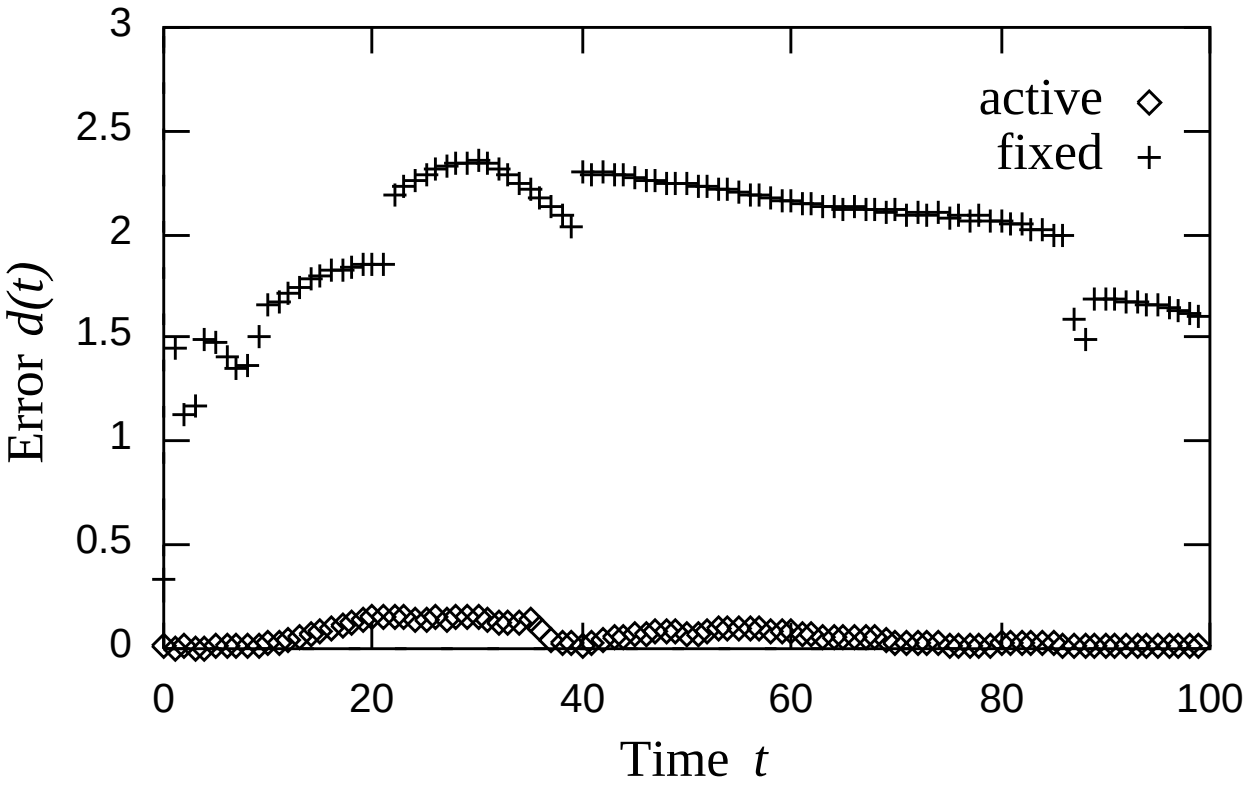
is used.

Application 1: Target tracking



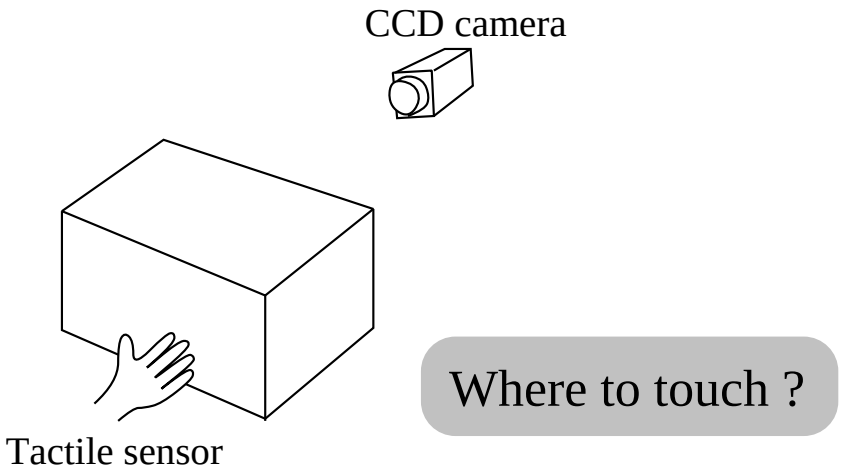
Hand-eye camera tracking system

Simulation results of the tracking



$$d(t) \equiv \sum_{i=1}^{2} \{(x_i(t) - \hat{x}_i(t|t))^2 + (y_i(t) - \hat{y}_i(t|t))^2\}^{\frac{1}{2}}$$

Application 2: Visual and tactile data fusion

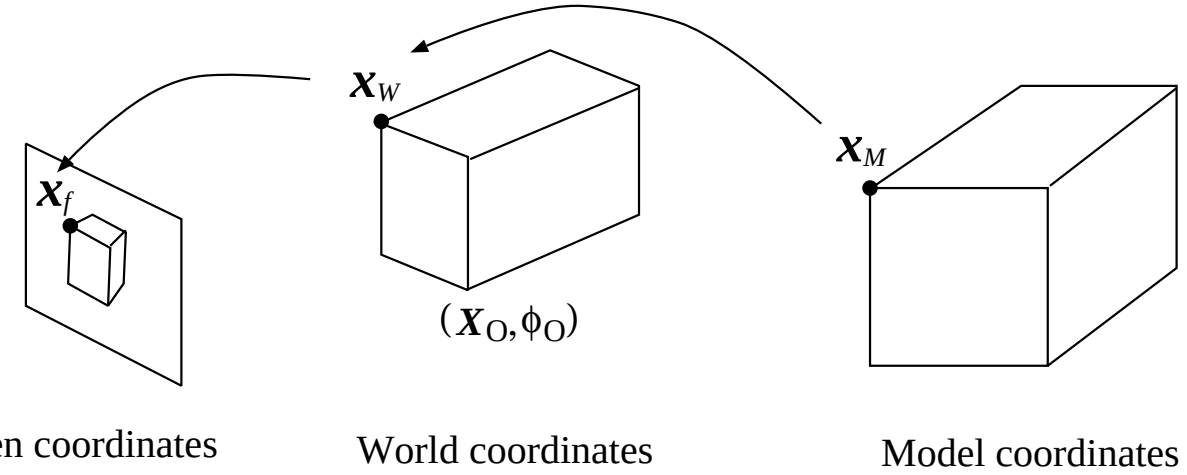


- By using the CCD camera, rough position of the object is obtained. Using the information, how to explore with the tactile sensor is determined.
- The tactile sensor is assumed to obtain local information of the 3D position of contact points and edges.
- 3D geometrical model of the object is utilized.

How to obtain the position X and orientation ϕ of a 3D object

The case of using characteristic points on a CCD image:

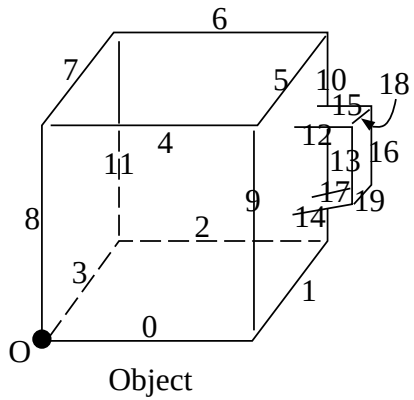
$$\boldsymbol{x}_W(\boldsymbol{X}_O, \boldsymbol{\phi}_O; \boldsymbol{x}_M) = R(\boldsymbol{\phi}_O)\boldsymbol{x}_M + \boldsymbol{X}_O$$

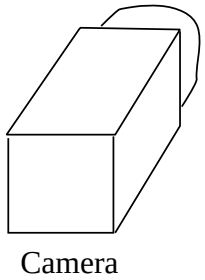


$$\boldsymbol{x}_f = \boldsymbol{h}_c(R(\boldsymbol{\phi}_O)\boldsymbol{x}_M + \boldsymbol{X}_O) + \boldsymbol{v}_c \equiv \tilde{\boldsymbol{h}}_c(\boldsymbol{X}_O, \boldsymbol{\phi}_O; \boldsymbol{x}_M) + \boldsymbol{v}_c$$

can be used as an observation equation of the proposed method.

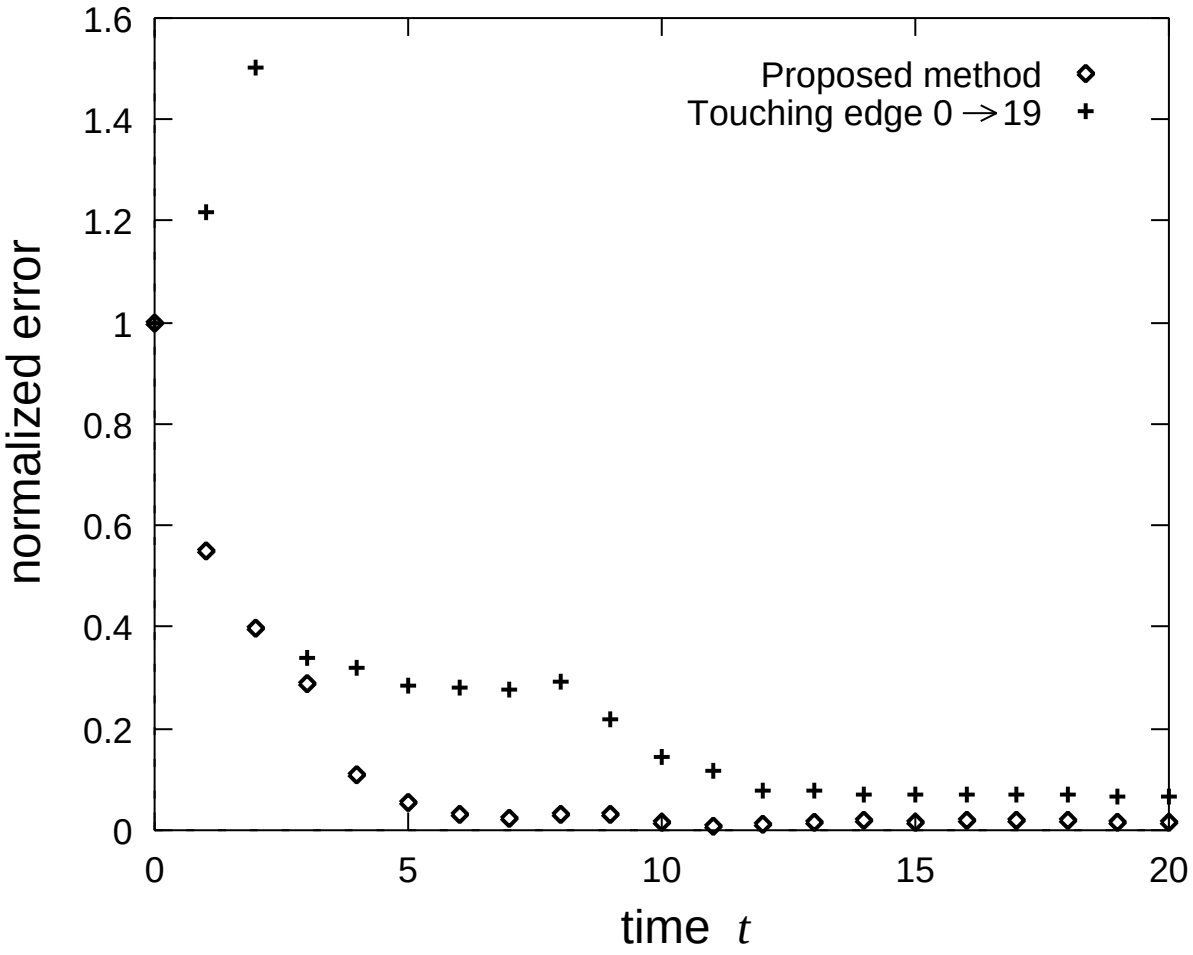
Situation of the simulation





1. Look once by the CCD camera.
2. Matching all vertices with the 3D model, obtain rough position of the object.
3. Using this information, determine edges to explore with the tactile sensor.

Simulation results



t	Selected edge
1	3
2	0
3	1
4	2
5	3
6	17
7	3

An example of selected edges

Conclusion

1. A new method to select sensor positions to minimize estimated errors has been proposed,
2. The method has been expanded to also achieve association correctly,
3. As an application, target tracking by 2 CCD cameras has been illustrated.
4. Visual and tactile data fusion has also been explained as an application. In this system, 3D geometrical model was used to select good sensor position.

Future work

Implementation of this method to a system with a CCD camera and a manipulator with a force/torque sensor.