Proceedings of the 32nd ISR(International Symposium on Robotics), 19-21 April 2001

Sensory-Motor Fusion Architecture Based on High-Speed Sensory Feedback and Its Application to Grasping and Manipulation

Akio Namiki　　Masatoshi Ishikawa
JST/Univ. of Tokyo　　Univ. of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
URL : http://www.k2.t.u-tokyo.ac.jp/
E-mail : namik@k2.t.u-tokyo.ac.jp

Abstract

In this paper we propose a new hierarchical parallel processing model based on high-speed realtime sensory feedback to realize manipulation in a dynamically changing environment. In this model high-speed sensory feedback is used at hierarchical three stages: (1) servo control, (2) motion planning, and (3) adaptation of the motion planner and the motion controller. Because the structure of the motion planner and the motion controller is changed according to an external environment at a high rate, responsiveness and flexibility to dynamic changes of the environment is achieved. And grasping and handling of a dynamically moving object is achieved based on this model. Experimental results are shown with a dextrous hand-arm and a high-speed active vision.

1. Introduction

A human being has an excellent ability and a difficult task is achieved such as grasping of a flying object. To realize such an ability to grasp in a dynamically changing environment, realtime sensory feedback (vision, tactile sensor, and force sensor) is one of the most important element [1].

Under this background, mainly using visual feedback, several applications of realtime sensory feedback to a grasping system have been studied. Allen et al. realized grasping of a model train using visual feedback based on the optic flow field [2]. Hong et al. realized grasping of a flying ball using visual feedback by stereo vision [3]. Namiki et al. realized high-speed grasping of a dynamic moving object using visual feedback at the cycle time of 1[ms] [4].

In these systems realtime sensory feedback is used mainly for servo control. But because external sensors such as vision are used to observe the state of an environment directly, realtime sensory feedback can be used not only for servo control but also for various uses. And a hierarchical parallel processing architecture is appropriate for designing such a multipurpose system[5]. Conventionally, several models of hierarchical processing were proposed, for example, sensory-motor fusion[6], and multiple control, and hierarchical sensory feedback[7]. However, most of these models were conceptual and how to use realtime sensory feedback was not clearly shown.

To solve this problem, in this paper, a new simple hierarchical parallel processing model using high-speed realtime sensory feedback. A brief outline of the paper is as follows. In section 2 we discuss functions of realtime sensory feedback in a grasping task, and based on the result we propose a simple hierarchical parallel processing model. In section 3 a grasping algorithm for one example problem is derived from the proposed model. At last in section 4 an experimental result is presented using a 6-axis manipulator with a multifingered hand and an active vision.

2. Hierarchical parallel realtime sensory feedback model

2.1. Use of realtime sensory feedback

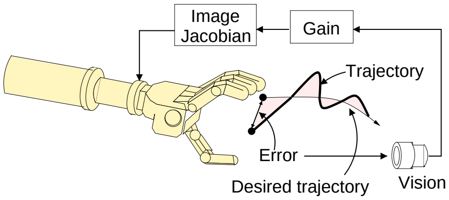
In a grasping task realtime sensor information can be used for various uses and main uses are as follows.

(1) Feedback for servo control
Sensory feedback is used for servo control. Fig.1 (a) is an example in which a hand is controlled by observing the positional error with a vision. In this example a vision is used in the same way as internal sensors which is used as merely a measurement equipment. In this case, to achieve stable and high responsive control, processing frequency in sensory feedback should be more than 1[kHz] [8].

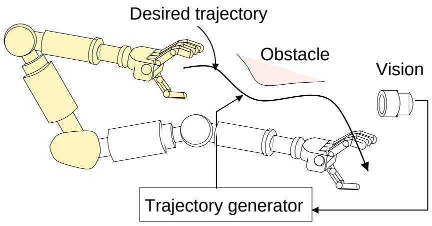
(2) Desired trajectory generation
Sensory feedback is used for desired trajectory generation. Fig.1 (b) is an example in which using visual feedback an optimal arm trajectory is computed by real-time against a moving obstacle. To generate a trajectory in a dynamical changing environment, high-speed realtime sensory feedback is needed.

(3) Task switching according to the state of an environment
Sensory feedback is used for switching tasks. Fig.1 (c) shows the example in which three subtasks is switched according to the state of the environment; grasping, object handling, and collision avoidance. To select an optimal subtask in a dynamical changing environment, high-speed realtime sensory feedback is needed.

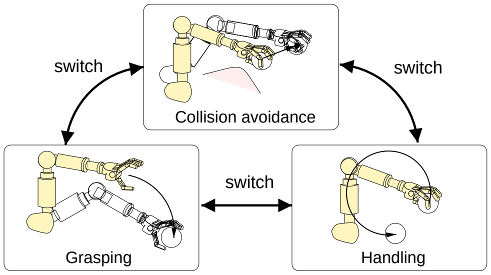
We propose a new hierarchical parallel processing model on the basis of this classification.



(a) Servo control



(b) Trajectory generation



(c) Task switching

Figure 1: Use of realtime sensory feedback

2.2. Hierarchical parallel realtime sensory feedback model

To simplify the problem, it is assumed that the state of a system is uniquely described by a state variable $\boldsymbol{z} \in \boldsymbol{R}^{m_z}$, and $\boldsymbol{\theta} = \boldsymbol{f}(\boldsymbol{z})$ is satisfied by a motion parameter $\boldsymbol{\theta} \in \boldsymbol{R}^{m_\theta}$. On these assumptions a hierarchical parallel processing model which consists of three layers is constructed as shown in Fig.2.

(1) Control layer
Corresponding to the use **(1)**, servo control is executed using high-speed realtime sensory feedback. A feedback signal is calculated by fusing some kinds of sensor information. Concretely, if the projection of an observed value to the state space is $\boldsymbol{z}_i = \boldsymbol{z}_i(\boldsymbol{s}) \in \boldsymbol{R}^{m_z}$ $(i = 1, 2, \cdots, M_z)$, the fused value is given as a linear combination of them.

(2) Planning layer
Corresponding to the use **(2)**, computation of a desired trajectory which is given to the control layer is executed. To deal with various types of a task, a desired trajectory is given as an integrated value of several trajectories of subtasks. Concretely, if the projection of a desired trajectory of a subtask to the state space is $\boldsymbol{z}_{di} = \boldsymbol{z}_{di}(\boldsymbol{s}) \in \boldsymbol{R}^{m_z}$ $(i = 1, 2, \cdots, M_{dz})$, the integrated value is given as a linear combination of them.

(3) Adaptation layer
Corresponding to the use **(3)**, the structure of integration of subtasks in the plan layer and the structure of sensor fusion in the control layer are changed according to a working environment. This is realized by changing positive definition coefficient matrixes $U_i(\boldsymbol{s}) \in \boldsymbol{R}^{m_z \times m_z}$ $(i = 1, 2, \cdots, M_{dz})$ and $S_i(\boldsymbol{s}) \in \boldsymbol{R}^{m_z \times m_z}$ $(i = 1, 2, \cdots, M_z)$ which are weighting matrices in the linear combination.

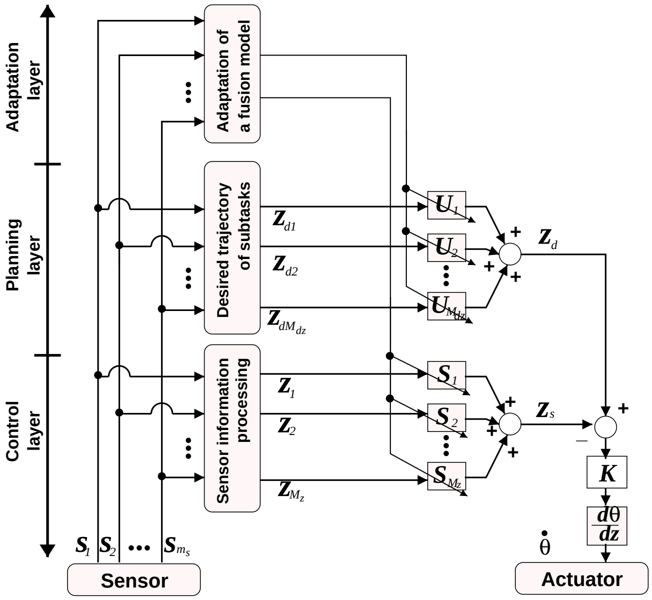


Figure 2: Hierarchical parallel realtime sensory feedback model

As a result the hierarchical parallel processing model is described as,

$$\dot{\boldsymbol{\theta}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{z}} \left(\sum_{i=1}^{M_{dz}} U_i(\boldsymbol{s}) \boldsymbol{z}_{di}(\boldsymbol{s}) - \sum_{j=1}^{M_z} S_j(\boldsymbol{s}) \boldsymbol{z}_j(\boldsymbol{s}) \right), \quad (1)$$

where $\sum_{i=1}^{M_{dz}} U_i(\boldsymbol{s}) = \sum_{j=1}^{M_z} S_j(\boldsymbol{s}) = I_{m_z}$ is satisfied. The matrix I_{m_z} is an unit matrix.

Because processing in each layer is computed by high-speed realtime sensory feedback in this model, processing corresponding to dynamic changes of an environment is realized. Moreover, it has an advantage that a design of a system such as addition and deletion of realtime sensory feedback is easy because processing in each layer is specifically separated.

3. High-speed grasping algorithm

3.1. Assumption of a target system

We assume that a system has a 6-axis manipulator with a multifingered dextrous hand and one monocular vision. And each joint has a force sensor and contact condition can be recognized using them.

As a goal task, a series of a general manipulation process is assumed, namely, grasping of an object, and handling of an object, and interaction between a handling object and another object.

We adopt following subtasks to realize such a process: (1) tracking to keep an optimal position for grasping by canceling the relative positional error between the hand and the object (Fig.3 (a)) , (2) reaching to stretch the arm to the object (Fig.3 (b)), (3) object handling to a desired position (Fig.3 (c)), (4) collision avoidance of a manipulated object (Fig.3 (d)). In Fig.3 these subtasks are shown. Other subtasks should be needed for general-purpose manipulation, but in this study we adopt only above subtasks to simplify the problem.

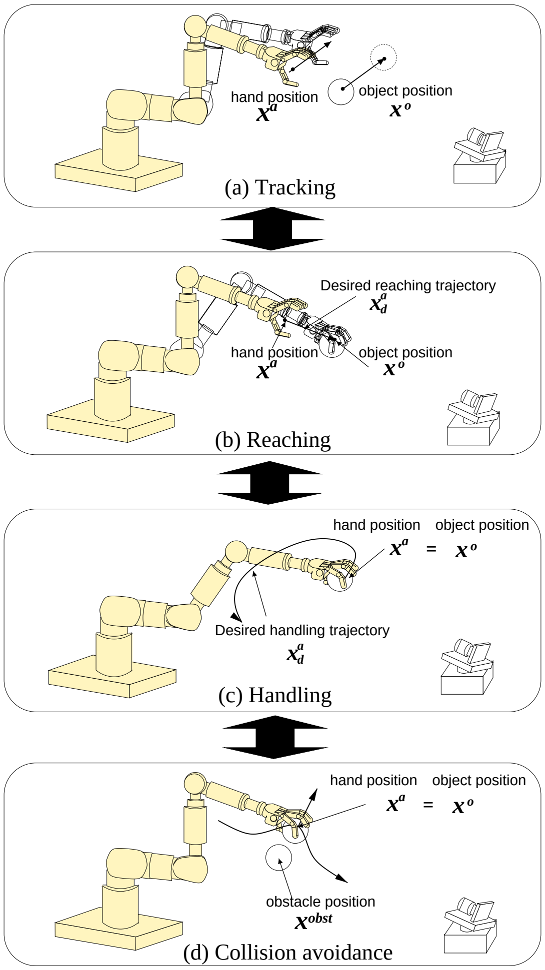


Figure 3: Desired subtasks

3.2. High-speed grasping algorithm

If the joint angle vector of a manipulator is described as $\boldsymbol{\theta}^a \in \boldsymbol{R}^6$, and the position and the orientation of the hand derived from $\boldsymbol{\theta}^a$ is described as $\boldsymbol{x}^a \in \boldsymbol{R}^6$, and the position and the orientation of the object observed by vision is described as $\boldsymbol{x}^o \in \boldsymbol{R}^6$, and force and moment observed by the force/torque sensor is described as $\boldsymbol{F}^a \in \boldsymbol{R}^6$, the desired joint angle velocity $\boldsymbol{v}_d^a \in \boldsymbol{R}^6$ is computed as follows,

$$\boldsymbol{v}_d^a = {J^a}^{-1} K^a \left(\boldsymbol{x}_d - \boldsymbol{x}_s\right) - K^{av} \dot{\boldsymbol{\theta}}^a + {J^a}^T K^{af} \boldsymbol{F}^a, \quad (2)$$

where $J^a \equiv \frac{\partial \boldsymbol{x}^a}{\partial \boldsymbol{\theta}^a}$ is a Jacobian matrix and K^a, K^{av}, and K^{af} are a positive definition diagonal matrix. In this equation the first item means position control based on realtime sensory feedback, and the second item means velocity feedback, and the third item means force feedback using the force/torque sensors.

The variables in the first item is calculated as follows,

$$\begin{aligned} \boldsymbol{x}_d \equiv\ & (I_6 - G^a)(I_6 - G^m)(I_6 - G^s)\, \boldsymbol{x}^o \\ & + (I_6 - G^a)(I_6 - G^m)\, G^s\, \boldsymbol{x}_d^a, \\ & + (I_6 - G^a)\, G^m \boldsymbol{x}_d^o \\ & + G^a\, \boldsymbol{x}_d^c, \quad (3) \\ \boldsymbol{x}_s \equiv\ & (I_6 - C^m)\, \boldsymbol{x}^a + C^m\, \boldsymbol{x}^o, \quad (4) \end{aligned}$$

where $\boldsymbol{x}_d$ is the desired trajectory by integrating desired trajectories of subtasks. This is derived from the desired tracking trajectory $\boldsymbol{x}^o$, the desired handling trajectory $\boldsymbol{x}_d^o \in \boldsymbol{R}^6$, the desired avoidance trajectory $\boldsymbol{x}_d^c \in \boldsymbol{R}^6$, the desired reaching trajectory $\boldsymbol{x}_d^a \in \boldsymbol{R}^6$. And $\boldsymbol{x}_s$ shows the state of the system observed by fusing visual information $\boldsymbol{x}^o$ with internal sensor information $\boldsymbol{x}^a$ (the fingertip position computed by the joint angle $\boldsymbol{\theta}^a$). The matrix $I_6 \in \boldsymbol{R}^{6\times 6}$ is an unit matrix.

Coefficient matrices are calculated as follows,

$$G^s = \mathrm{diag}\,(1,\ 0,\ 0,\ 0,\ 1,\ 1)\,, \quad (5)$$

$$G^m = \mathrm{diag}\,(\,g_i^m\,)\,, \quad i = 1, 2, \cdots, 6, \quad (6)$$

$$g_i^m = \begin{cases} 1 & (1 < \bar{g}_i^m) \\ \bar{g}_i^m & (0 \leq \bar{g}_i^m \leq 1) \\ 0 & (\bar{g}_i^m < 0) \end{cases}, \quad (7)$$

$$\bar{g}_i^m = \int_{t'=0}^{t} \gamma_i^m \mathrm{sgn}(\tau^h(t') - \tau_o^h)\, dt', \quad (8)$$

$$\tau^h = \sqrt{\sum_{i=1}^{4} {\boldsymbol{\tau}_i^h}^T H_i \boldsymbol{\tau}_i^h}, \quad (9)$$

$$G^a = \mathrm{diag}\,(\,g_i^a\,)\,, \quad i = 1, 2, \cdots, 6, \quad (10)$$

$$g_i^a = \frac{1}{1 + \exp(\gamma_i^a (l - l_o))}, \quad (11)$$

$$C^m = C^\alpha\ G^m, \quad (12)$$

where τ^h shows the size of the weighted average of the joint torque of the hand $\boldsymbol{\tau}_i^h \in \boldsymbol{R}^3$, $\tau_o^h \in \boldsymbol{R}$ is a threshold, $l \in \boldsymbol{R}$ shows the distance between the manipulated object and the obstacle, and $l_o \in \boldsymbol{R}$ is a threshold for adjusting the distance in which avoidance motion starts. The matrices C^α, H_i and the constant numbers γ_i^m, γ_i^a are an appropriate coefficient.

In each phase Eqn(3) is written as followings.

(a) Tracking and Reaching
Because $G^m = O_6$ and $G^a = O_6$ in the reaching subtask and the tracking subtask, Eqn.(2) is described as follows,

$$\begin{aligned} \boldsymbol{v}_d^a = {J^a}^{-1} K^a \left\{ \left(I_6 - G^s \right) \left(\boldsymbol{x}^o - \boldsymbol{x}^a \right) + G^s \left(\boldsymbol{x}_d^a - \boldsymbol{x}^a \right) \right\} \\ - K^{av} \dot{\boldsymbol{\theta}}^a + {J^a}^T K^{af} \boldsymbol{F}^a, \end{aligned} \tag{13}$$

where because tracking motion $\left(I_6 - G^s \right) \left(\boldsymbol{x}^o - \boldsymbol{x}^a \right)$ is orthogonal to reaching motion $G^s \left(\boldsymbol{x}_d^a - \boldsymbol{x}^a \right)$, there is no interference between them.

(b) Handling
Because $G^m = I_6$ and $G^a = O_6$ in the object handling subtask, Eqn.(2) is described as follows,

$$\boldsymbol{v}_d^a = {J^a}^{-1} K^a \left(\boldsymbol{x}_d^o - \boldsymbol{x}_s^o \right) - K^{av} \dot{\boldsymbol{\theta}}^a + {J^a}^T K^{af} \boldsymbol{F}^a, \tag{14}$$

where $\boldsymbol{x}_s^o \equiv (I_6 - C^\alpha) \boldsymbol{x}^a + C^\alpha \boldsymbol{x}^o$ is the value by fusing visual information with internal sensor information. In this equation servo control $\left(\boldsymbol{x}_d^o - \boldsymbol{x}_s^o \right)$ is executed.

(c) Collision avoidance
Because $G^m = I_6$ and $G^a = I_6$ in the collision avoidance subtask, Eqn.(2) is described as follows,

$$\boldsymbol{v}_d^a = {J^a}^{-1} K^a \left(\boldsymbol{x}_d^c - \boldsymbol{x}_s^o \right) - K^{av} \dot{\boldsymbol{\theta}}^a + {J^a}^T K^{af} \boldsymbol{F}^a, \tag{15}$$

where an avoidance control $\left(\boldsymbol{x}_d^c - \boldsymbol{x}_s^o \right)$ by fusing visual information with the internal sensor information is executed. As for a desired trajectory for collision avoidance, for example, it is computed as follows,

$$\boldsymbol{x}_d^c \equiv \begin{cases} \boldsymbol{x}_s^o + \begin{bmatrix} \boldsymbol{n}^l \left(l_o - l \right) \\ \boldsymbol{0} \end{bmatrix} & \text{if } l_o - l > 0, \\ \boldsymbol{x}_s^o & \text{otherwise,} \end{cases} \tag{16}$$

where $\boldsymbol{n}^l \in \boldsymbol{R}^3$ is an unit vector in the direction from the obstacle to the manipulated object.

Remark 1
In this algorithm, switching from reaching to object handling is executed using the grasping power $\boldsymbol{\tau}^h$ and switching from object handling to collision avoidance is executed using the distance l observed by a vision.

Remark 2
The structure of sensor fusion changes based on hand joint torque $\boldsymbol{\tau}^h$. Therefore the object is recognized by only visual information before a hand grasps an object, and it is recognized by fusing visual information with internal sensor information after it.

Remark 3
In this method it is designed that tracking and reaching are orthogonal. But generally it is necessary to change the condition of separation according to the state of the environment.

Remark 4
As for calculation in the adaptation layer, by adopting appropriate coefficients except Eqn.(6) and Eqn.(10), it is possible to realize the similar operation. But the global stability of the whole system should be satisfied. These problems are future works.

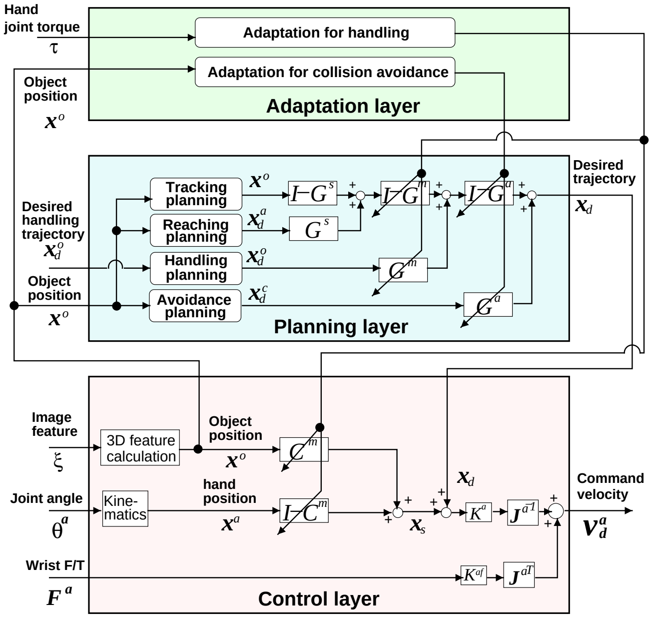


Figure 4: Block diagram

4. Experiment

4.1. Experimental system

The experimental system based on 1ms sensory-motor fusion system [9] is shown in Fig.5. This system consists of a DSP hierarchical parallel processing system, an active vision, a 7-axis axis arm, and a 4-fingered hand.

The DSP hierarchical parallel processing system is a multiprocessor system in which eleven general-purpose floating-point DSPs (TMS320C40, Texus Instrument Ltd.) are connected with each other. This is the main part to execute high-speed sensor information processing by the cycle time of 1[ms]. On the other hand, the active vision is a 2-DOF actuator (tilt and pan) and equips with high-speed vision chip system SPE-256 (resolution:16 × 16 pixel). And the arm is a 7-DOF actuator which loads the wrist with the 6-axis force/torque sensor. The hand is a 4-fingered 14-DOF actuator which equips with force sensors to each joint.

To avoid occlusion problems, the hand-arm and the active vision are arranged opposite to each other and a manipulated object is placed in the middle position. And two planes, the beginning plane and the ending plane for reaching, is settled between the hand and the active vision, and reaching is planned so that the hand is transfered from the beginning plane to the ending plane.

To make a trajectory which is difficult to estimate, a manipulated object is optionally moved by a human hand on the constraint plane.

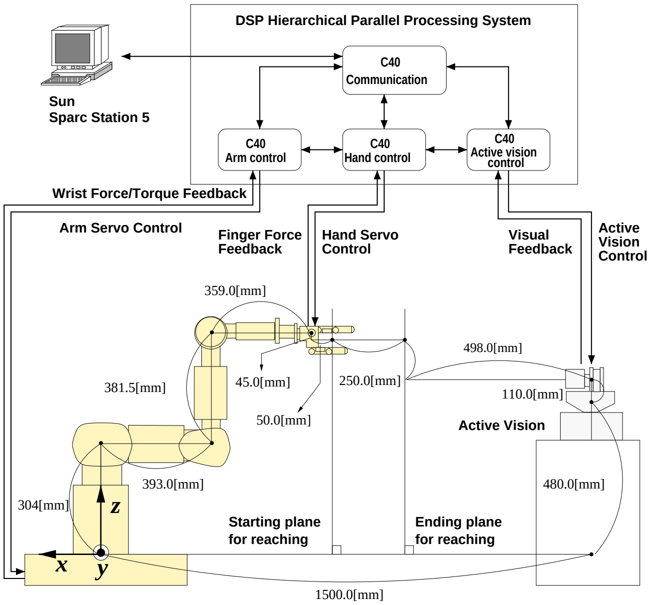


Figure 5: Experimental system

The cycle time of whole system control is 1.5[ms], and the cycle time of active vision control is 1.0[ms] to speed up visual processing.

4.2. Experimental result

To simplify the problem, it is assumed that in the object handling subtask the goal is to make the object stand still at a specified position.

In Fig.6(a)(b)(c), the time response of main parameters, the time response of an element of coefficients for adaptation G^a and G^m, and the time response of the grasping power τ^h are shown. In this figure the tracking subtask and the reaching subtask are realized from 0.0[s] to 4.0[s], and task switching to the object handling subtask is executed after 4.0[s]. And in the object handling subtask collision avoidance is executed when an obstacle approaches to the manipulated object.

Task switching from the reaching subtask to the object handling subtask is realized by the change of the coefficient matrix G^m from a zero matrix to an unit matrix according to the grasping power τ^h. And task switching from the object handling subtask to the collision avoidance subtask is realized by the change of the coefficient matrix G^a from a zero matrix to an unit matrix according to the approach of the obstacle. Because these task switchings are executed by realtime sensory feedback, they are executed at high-speed according to changes of an environment.

In Fig.7 a total manipulation process is shown as a continuous sequence of pictures every 0.3[s]. In this picture from 0.0[s] to 1.5[s] tracking motion is executed, and from 1.8[s] to 3.6[s] object handling (standing still in this case) is executed, and from 3.9[s] to 4.5[s] collision avoidance is realized at high-speed.

5. Conclusion

In this paper, a hierarchical parallel processing model based on high-speed realtime sensory feedback is proposed. And based on the model grasping and manipulation for a multifingered hand is realized which is high responsive and flexible to dynamic changes of an environment.

In the proposed grasping algorithm high level processing, such as estimation and learning, is not used. But from the experimental results it is shown that manipulation adaptable to dynamic changes of an environment can be realized using the realtime sensory feedback algorithm. This result also shows that high-speed realtime sensory feedback is effective not only to realize responsive and flexible manipulation, but also to achieve reliability by preventing complicated processing.

At present several applications of a hierarchical parallel processing model has been developed to various manipulation tasks.

References

[1] L. Weiss, A. Sanderson, and C.P. Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE J. of Robotics and Automation*, RA-3(5), 1987.

[2] P.K. Allen, B. Yoshimi, and A. Timucenko. Real-time visual servoing. *Proc. IEEE Int. Conf. Robotics and Automation*, pages 2376–2384, 1991.

[3] W. Hong and J.E. Slotine. Experiments in hand-eye coordination using active vision. *Proc. 4th Int. Symp. on Experimental Robotics (ISER'95)*, 1995.

[4] A. Namiki, Y. Nakabo, I. Ishii, and M. Ishikawa. High speed grasping using visual and force feedback. *Proc. IEEE Int. Conf. on Robotics and Automation*, 1999.

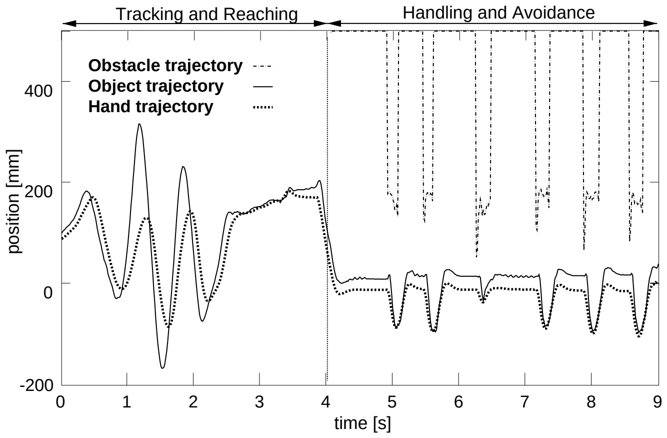
[5] M. Ishikawa. Sensor fusion : The state of the art. *J. of Robotics and Mechatronics*, 2–4:235–244, 1991.

[6] J.S. Albus. Outline for a theory of intelligence. *IEEE Trans. on Systems, Man, and Cybernetics*, 21(3):473–509, 1991.

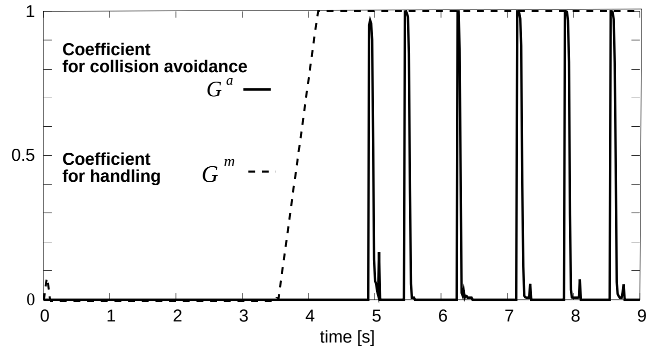
[7] R.A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.

[8] I. Ishii, Y. Nakabo, and M. Ishikawa. Target tracking algorithm for 1ms visual feedback system using massively parallel processing vision. *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 2309–2314, 1996.

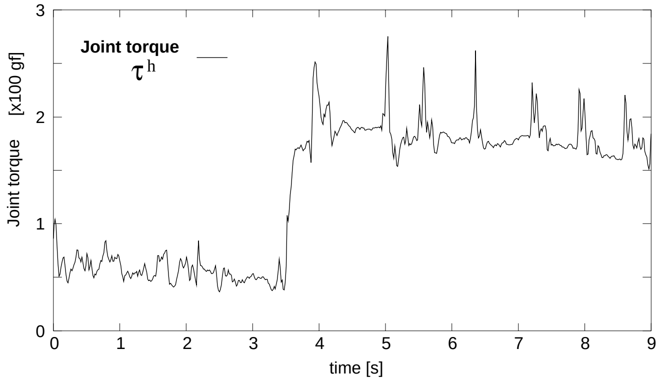
[9] A. Namiki, Y. Nakabo, I. Ishii, and M. Ishikawa. 1ms sensory-motor fusion system. *IEEE/ASME Trans. on Mechatronics*, 5(3):244–252, 2000.



(a) Trajectory



(b) Coefficient



(c) Joint torque

Figure 6: Time response: task switching

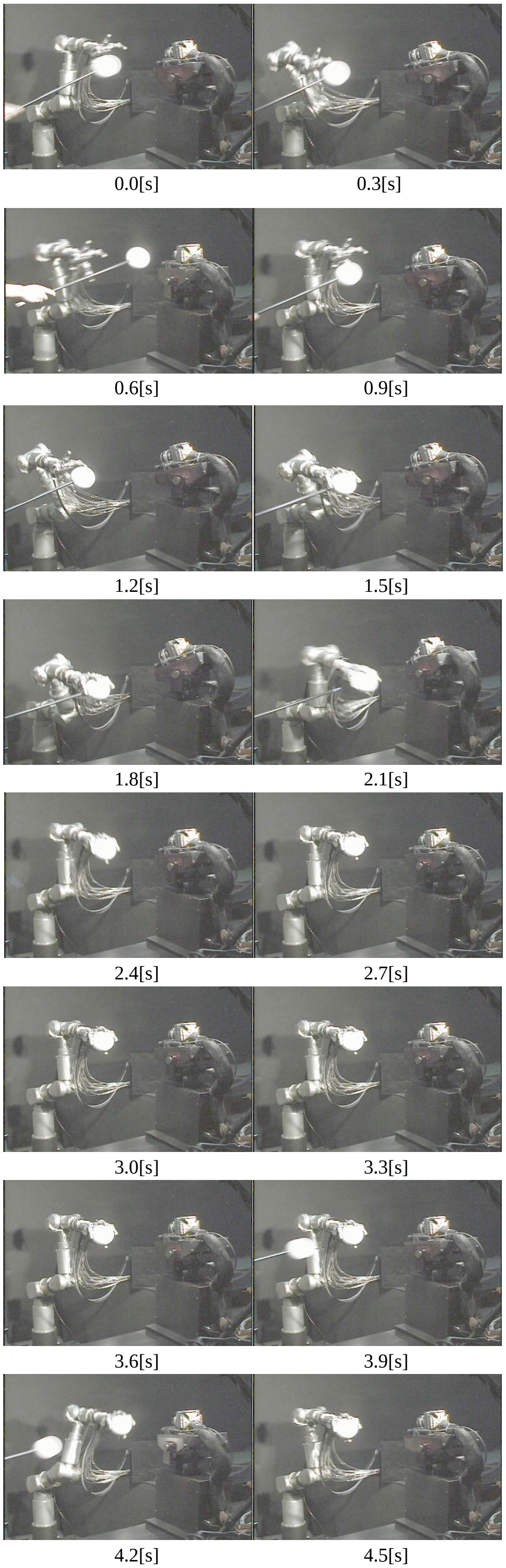


Figure 7: Experimental result: collision avoidance