Sensor Dependent Task Definition: Object Manipulation by Fingers with Uncalibrated Vision

Koh Hosoda, Takuya Hisano, and Minoru Asada
Dept. of Adaptive Machine Systems, Osaka University
hosoda@ams.eng.osaka-u.ac.jp

Abstract. In this paper, a method is proposed to control a robot hand with uncalibrated camera based on *sensor dependent task definition*. To achieve a robust and quick response without precise calibration, adaptive visual servoing is applied. On the other hand, internal force control also has to be applied to maintain grasping. To integrate these controllers, we utilize the principle of virtual work, and eliminate the interference. Experimental results are shown to demonstrate the validity of the proposed method.

1 Introduction

Recently, technology is mature enough to build up a complicated robot system consisting of various sensors and actuators, such as a humanoid. The control scheme for such a robot system becomes more complicated and the effort to program all the degrees grows enormous. A popular method to deal with many actuators and sensors is (1) to build a world model based on *a priori* knowledge, (2) to fuse data from various sensors in the model, (3) to derive outputs for actuators to achieve the task in the model, which is based on so called sensor fusion technique. The task corresponds to a behavior in the model, and hopefully corresponds to the one in the real world, which is not always true since the real world contains a lot of uncertainties that cannot be modeled beforehand. To apply this technique, all the external sensors must be calibrated with respect to the model, which needs tedious calibration process and the resultant system becomes weak to changes of parameters and modeling errors.

To construct an adaptive robot system that can work even in an unknown environment, it seems to be a better idea that tasks for the robot are defined in the robot's own sensor spaces, not in the world model. The behavior of the system is not directly defined, but is a consequence of interaction between the environment and tasks defined in the sensor spaces. It does not need any *a priori* knowledge about the world model since this method does not have to fuse the sensory data in it. The merit of this idea, *sensor dependent task definition*, is that we can derive one control module for each sensor and do not have to deal with all the degrees of freedom and all the sensors at a time. This makes it easy to construct the whole control architecture for a robot that has many actuators and sensors. Based on this idea, we have proposed a vision/force hybrid controller for a robot manipulator[1] and a walking controller for a legged robot[2].

A robot hand is one of such robot systems that have many actuators and various sensors. In the previous work on a robot hand, as we have mentioned above, they used a world model to visually control the robot hand, which needed *a priori* knowledge on the robot and on the environment, and therefore needed tedious calibration [3, 4, 5].

In this paper, we focus on controlling a robot hand based on the sensor dependent task definition. A problem to realize a robot system based on the definition is how to integrate the independently designed controllers for tasks. As for the case of a robot hand with uncalibrated vision, how to integrate the controller to maintain the stable grasp and the one to manipulate the object based on uncalibrated vision. We utilize the principle of virtual work to integrate these controllers.

The remainder of this article is organized as follows. First, a robot hand system with uncalibrated vision is introduced and the problem is described. The robot is controlled by two tasks defined in each sensor space independently: (1) to control internal force, and (2) to align the manipulated object in the image plane. We derive internal force control and adaptive visual servoing control to achieve these tasks. To integrate these controllers, we utilize the principle of virtual work. Finally, we demonstrate that robot hand can manipulate an object with uncalibrated vision as a consequence of the combination of these tasks by experimental results.

2 A Robot Hand with Uncalibrated Vision

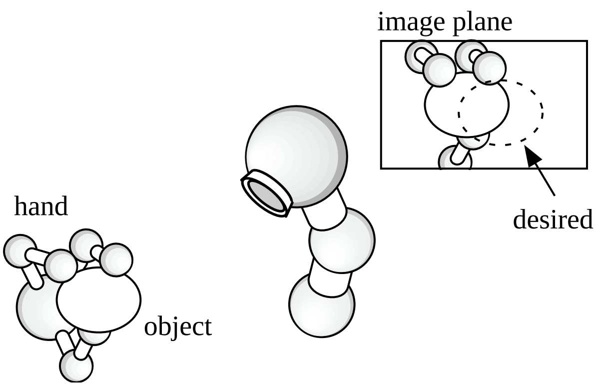

Figure 1: A robot hand with uncalibrated vision: The tasks for the system are to make the image features converge to the desired and to maintain the grasp.

In Figure 1, we depict a robot hand system with an uncalibrated camera. Each finger has a force sensor at the tip. The manipulated object is observed by the camera. The tasks for the system are (1) to make the internal force converge to given desired one, and (2) to make the image features on the manipulated object in the image plane converge to given desired ones. Note that these tasks are defined not in the world model but in each sensor subspace independently, that is, in a sensor dependent task definition manner.

Once a robot hand is manufactured, its kinematic parameters usually do not change. Therefore, we assume that the kinematic parameters of the hand are known. However, relative position and orientation between the camera and the hand are constantly changing according to the movement of the camera/hand. The controller has to estimate them.

3 Internal Force Control

Stable grasp is achieved by applying appropriate internal force. In this section, we derive an internal force controller for stable grasping. Note that this task is achieved in the force sensor subspace, and is not concerning to vision. We focus on 3-finger grasping here, but we can easily extend it to more-finger cases.

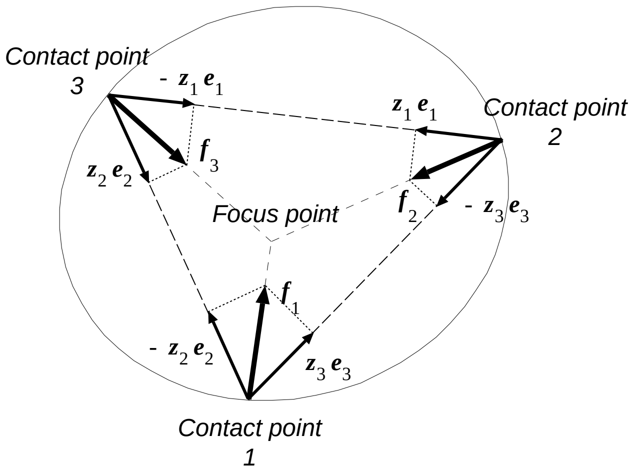

Figure 2: Representation of internal force: the cross section of manipulated object including three contact point

In Figure 2, the cross section of manipulated object including three contact points is shown. Let $\boldsymbol{e}_i$ and $\boldsymbol{z}$ be a unit vector from a contact point to another and a vector representing magnitude of internal force, respectively. Finger tip force vector $\boldsymbol{f}$ is

$$\boldsymbol{f} = \boldsymbol{E}\boldsymbol{z}, \tag{1}$$

where

$$\boldsymbol{E} = \begin{bmatrix} \mathbf{o} & -\boldsymbol{e}_2 & \boldsymbol{e}_3 \\ \boldsymbol{e}_1 & \mathbf{o} & -\boldsymbol{e}_3 \\ -\boldsymbol{e}_1 & \boldsymbol{e}_2 & \mathbf{o} \end{bmatrix}.$$

Applying feedforward and PI control, we can derive finger tip force $\boldsymbol{f}_g$ to accomplish internal force control

$$\boldsymbol{f}_g = \boldsymbol{E}\{\boldsymbol{z}_d + \boldsymbol{K}_p(\boldsymbol{z}_d - \boldsymbol{E}^+\boldsymbol{f}) + \boldsymbol{K}_i \int_0^t (\boldsymbol{z}_d - \boldsymbol{E}^+\boldsymbol{f})dt'\} + (\boldsymbol{I} - \boldsymbol{E}\boldsymbol{E}^+)\boldsymbol{k}, \tag{2}$$

where $\boldsymbol{k}$ denotes an arbitrary vector that describes redundancy of the hand with respect to the internal force task, therefore we can utilize it for visual servoing.

4 Visual Servoing Control

4.1 Relation between finger force and image features

Although it is popular to give the velocity command for the visual servoing control, we have to obtain force command instead to integrate it with internal force control. Therefore, we have to derive the relation between the finger force and image features. We utilize the principle of virtual work to derive this relation.

There is a relation between the image feature vector $\boldsymbol{x}$ observed in the image plane and the finger position $\boldsymbol{p}$,

$$\delta\boldsymbol{x} = \boldsymbol{J}_{xp}\delta\boldsymbol{p}, \tag{3}$$

where $\boldsymbol{J}_{xp}$ denotes a Jacobian matrix that describes the relation between them. Note that not all variation of $\delta\boldsymbol{p}$ appears in object movement in the image plane. Let $\delta\boldsymbol{p}$ be sum of $\delta\boldsymbol{p}_x$ that appears in the image plane and $\delta\bar{\boldsymbol{p}}_x$ that does not appear,

$$\delta\boldsymbol{p} = \delta\boldsymbol{p}_x + \delta\bar{\boldsymbol{p}}_x, \tag{4}$$

$$\boldsymbol{J}_{xp}\delta\bar{\boldsymbol{p}}_x = \mathbf{o}, \tag{5}$$

therefore, we can derive,

$$\delta\boldsymbol{p}_x = {\boldsymbol{J}_{xp}}^{+}\boldsymbol{J}_{xp}\delta\boldsymbol{p}. \tag{6}$$

According to the principle of virtual work, finger tip force $\boldsymbol{f}$ and virtual force in the image plane $\boldsymbol{\xi}$ has a relation

$$(\delta\boldsymbol{x})^T\boldsymbol{\xi} = (\delta\boldsymbol{p}_x)^T\boldsymbol{f}. \tag{7}$$

Substituting eqs.(3) and (6) into eq.(7), we can get finger force $\boldsymbol{f}_v$ to realize $\boldsymbol{\xi}$ in the image plane,

$$\boldsymbol{f}_v = ({\boldsymbol{J}_{xp}}^{+}\boldsymbol{J}_{xp})^{+}{\boldsymbol{J}_{xp}}^{T}\boldsymbol{\xi} + \{\boldsymbol{I}_9 - ({\boldsymbol{J}_{xp}}^{+}\boldsymbol{J}_{xp})^{+}{\boldsymbol{J}_{xp}}^{+}\boldsymbol{J}_{xp}\}\boldsymbol{k}', \tag{8}$$

where $\boldsymbol{k}'$ is an arbitrary vector that describes the redundancy. To determine virtual force $\boldsymbol{\xi}$, we simply apply proportional control to the image error

$$\boldsymbol{\xi} = \boldsymbol{K}_x(\boldsymbol{x} - \boldsymbol{x}_d), \tag{9}$$

where $\boldsymbol{x}_d$ and $\boldsymbol{K}_x$ denote the desired image feature vector and a feedback gain matrix, respectively.

4.2 On-line Jacobian estimator

Because the camera is not calibrated with respect to fingers, the Jacobian matrix $\boldsymbol{J}_{xp}$ is unknown. Therefore, we apply an on-line estimation technique to obtain it at k-th step:

$$\widehat{\boldsymbol{J}}_{xp}(k) = \widehat{\boldsymbol{J}}_{xp}(k-1) + \frac{\{\delta\boldsymbol{x}(k) - \widehat{\boldsymbol{J}}_{xp}(k-1)\delta\boldsymbol{p}(k)\}\delta\boldsymbol{p}(k)^T\boldsymbol{P}(k-1)}{\rho + \delta\boldsymbol{p}(k)^T\boldsymbol{P}(k-1)\delta\boldsymbol{p}(k)}, \tag{10}$$

where $(0 < \rho < 1)$ is a forgetting factor, and $\boldsymbol{P}(k)$ is a covariance matrix calculated as

$$\boldsymbol{P}(k) = \frac{1}{\rho}\left\{\boldsymbol{P}(k-1) - \frac{\boldsymbol{P}(k-1)\delta\boldsymbol{p}(k)\delta\boldsymbol{p}(k)^T\boldsymbol{P}(k-1)}{\rho + \delta\boldsymbol{p}(k)^T\boldsymbol{P}(k-1)\delta\boldsymbol{p}(k)}\right\}. \tag{11}$$

We have already demonstrated that many kinds of robots can be controlled with this estimator without any *a priori* knowledge on the robot[6]. Utilizing estimated $\widehat{\boldsymbol{J}}_{xp}$, the finger force is

$$\boldsymbol{f}_v = ({\widehat{\boldsymbol{J}}_{xp}}^{+}\widehat{\boldsymbol{J}}_{xp})^{+}{\widehat{\boldsymbol{J}}_{xp}}^{T}\boldsymbol{\xi} + \{\boldsymbol{I}_9 - ({\widehat{\boldsymbol{J}}_{xp}}^{+}\widehat{\boldsymbol{J}}_{xp})^{+}{\widehat{\boldsymbol{J}}_{xp}}^{+}\widehat{\boldsymbol{J}}_{xp}\}\boldsymbol{k}'. \tag{12}$$

From eqs.(2), (9), and (12), we can derive a controller for a hand to manipulate an object with uncalibrated camera:

$$\begin{aligned}\boldsymbol{f} &= \boldsymbol{E}\{\boldsymbol{z}_d + \boldsymbol{K}_p(\boldsymbol{z}_d - \boldsymbol{E}^{+}\boldsymbol{f}) + \boldsymbol{K}_i\int_0^T(\boldsymbol{z}_d - \boldsymbol{E}^{+}\boldsymbol{f})dt\} \\ &\quad +(\boldsymbol{I}_9 - \boldsymbol{E}\boldsymbol{E}^{+})({\widehat{\boldsymbol{J}}_{xp}}^{+}\widehat{\boldsymbol{J}}_{xp})^{+}{\widehat{\boldsymbol{J}}_{xp}}^{T}\boldsymbol{K}_x(\boldsymbol{x}_d - \boldsymbol{x}) \\ &\quad +(\boldsymbol{I}_9 - \boldsymbol{E}\boldsymbol{E}^{+})\{\boldsymbol{I}_9 - ({\widehat{\boldsymbol{J}}_{xp}}^{+}\widehat{\boldsymbol{J}}_{xp})^{+}{\widehat{\boldsymbol{J}}_{xp}}^{+}\widehat{\boldsymbol{J}}_{xp}\}\boldsymbol{k}''.\end{aligned} \tag{13}$$

The last term denotes the redundancy of the hand with respect to two tasks.

5 Experiments

5.1 Experimental equipment

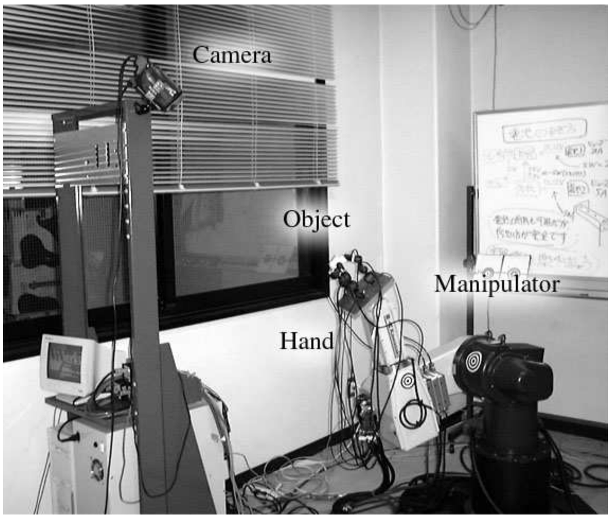

Figure 3: An experimental robot system consists of a multi-fingered hand, an arm and a camera.

We show several experimental results to demonstrate how a robot hand behaves according to the proposed method. In Figure 3, the overview of the experimental robot system is shown. In Figure 4, equipment used for the experiment is depicted. The hand has three fingers, and each finger has 3 degrees of freedom. Adaptive visual servoing and internal force control are programmed as parallel processes on VxWorks(R) 5.3.1, and the sampling rates are 33 [ms] and 2[ms], respectively.

Four image features are captured by the camera (see Figure5), and $\boldsymbol{x} \in \Re^8$ consists of their coordinates in the image plane.

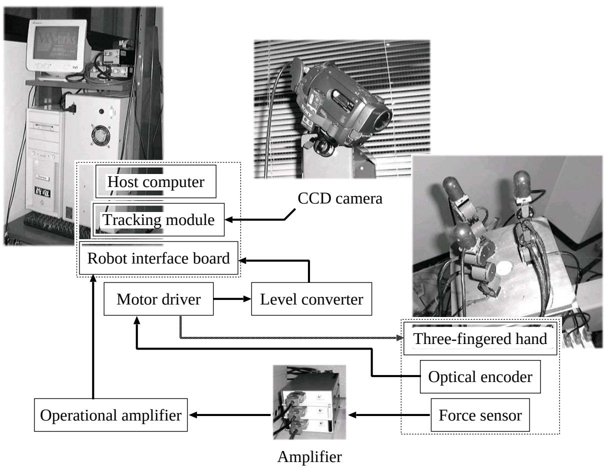

Figure 4: Experimental equipment used for the experiment

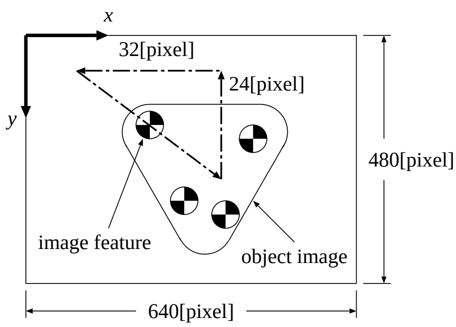

Figure 5: Desired trajectory

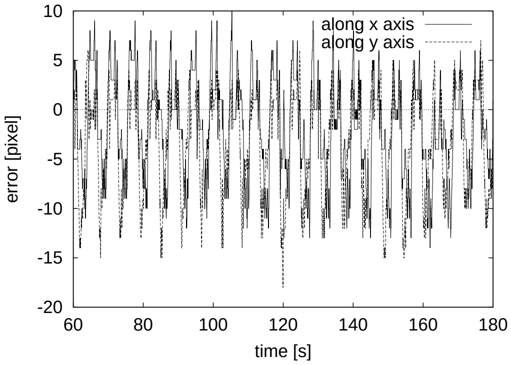

Figure 6: Error between one image feature coordinate and the desired in the image plane during manipulation

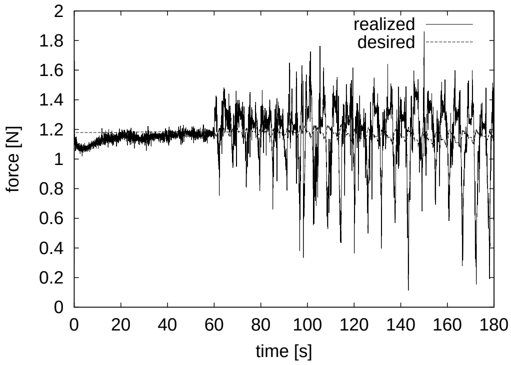

Figure 7: Internal force z_1 during manipulation

5.2 Experiment 1: trajectory tracking

First, we gave manipulating trajectories and saw how the proposed method worked. The desired trajectories of image features were given in triangle shapes (Figure 5) whose velocities are 20 [pixel/s]. We applied only internal force control for 60 [s] to stabilize the grasp, and then began manipulation by applying adaptive visual servoing along the trajectories. The desired value for the internal force was calculated so as to make $\|\boldsymbol{z}_d\|$ be 2.94[N].

Since we assumed that the fingers are not calibrated beforehand, we gave an initial estimated Jacobian matrix $\widehat{\boldsymbol{J}}_{xp}(0)$ as

$$\widehat{\boldsymbol{J}}_{xp}(0) = \begin{bmatrix} 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}.$$

The robot hand successfully manipulated the object without dropping it. We show error between an image feature coordinate and the desired in the image plane in Figure 6 and internal force z_1 in Figure 7. Although the estimated initial Jacobian matrix was given arbitrarily, the proposed method could manipulate the object within 10 [pixel] error. The reason we had still 10 [pixel] error might be that we could not increase the feedback gain since the initial Jacobian matrix was arbitrary. We may be able to eliminate the error by gradually increase the gain.

We can see from Figure 7 that the internal force converges to desired value before the manipulation. However, during manipulation, the force does not converge to the desired very well. The reason might be there was rolling contact between the fingers and the object during manipulation.

5.3 Experiment 2: Moving the arm

In this experiment, we fixed the desired image features as initial ones, and moved the arm on which the hand was fixed. The pose transition of the hand by the arm movement

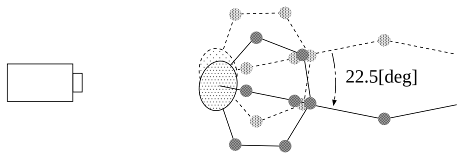

Figure 8: Pose transition of the hand by the arm

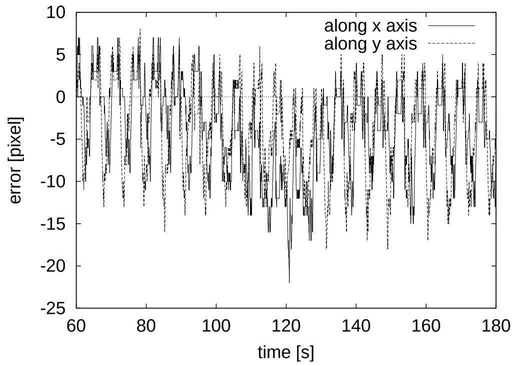

Figure 9: Error between one image feature coordinate and the desired in the image plane during movement of the arm

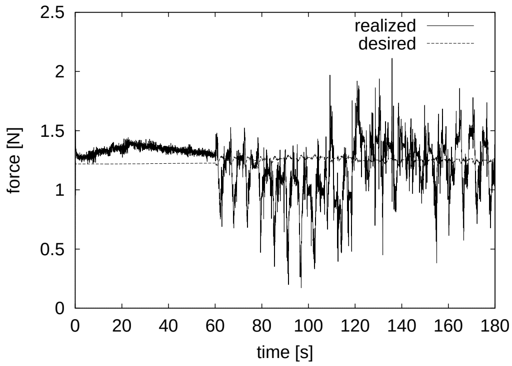

Figure 10: Internal force z_1 during movement of the arm

is shown in Figure 8. We applied internal force control for 60 [s] to stabilize the grasp, and then began moving the arm 22.5 [deg] until 80[s].

We show error between an image feature coordinate and the desired in the image plane in Figure 9 and internal force z_1 in Figure 10. We also show the movement of fingers during this experiment in Figure 11.

In this experiment, the robot hand successfully manipulated the object without dropping it, too. The proposed method can manipulate the object within around 10 [pixel] error. Around 120 [sec], however, error in the image plane increases. The reason may be that one of fingers is near to the singular configuration at the time (see Figure 11, the joint indicated by the dotted line).

Although the hand did not know the movement of the arm, it can manipulate the object successfully. This fact suggest us that the proposed method will be very powerful tool to achieve modular type control structure.

6 Conclusion and Discussion

In this paper, we have proposed to control a robot hand with uncalibrated vision based on sensor dependent task definition. To integrate the controllers, we have introduced the principle of virtual work. Experimental results have been shown to demonstrate the validity of the proposed method.

When a robot has many degrees of freedom and a variety of sensors, the conventional way of controlling it based on the world model is no more effective. For such robots, sensor dependent task definition is the one of most powerful way of controlling it.

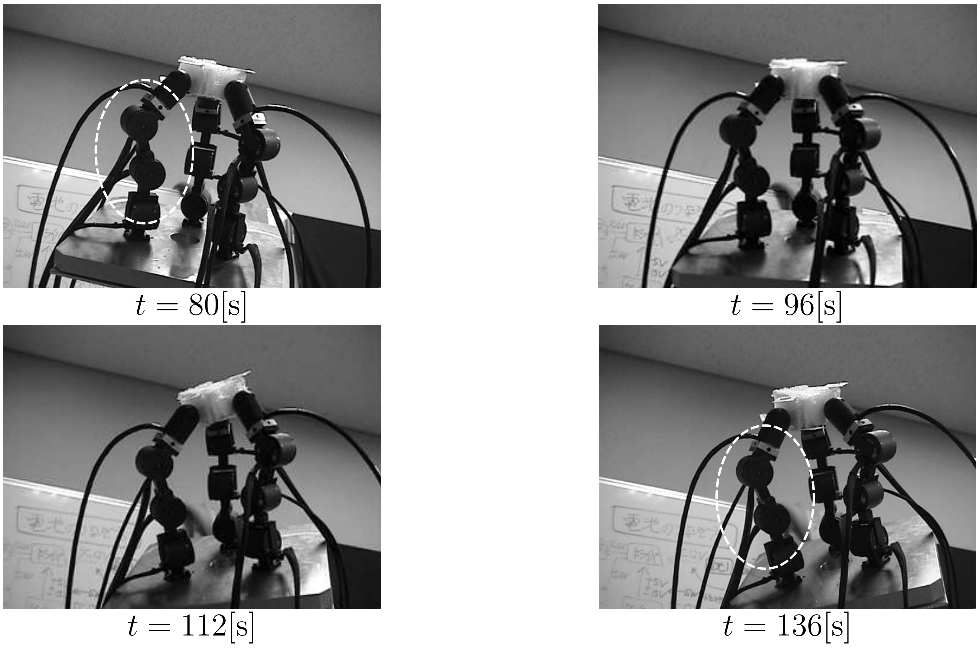

Figure 11: Movement of the fingers

References

[1] K. Hosoda, K. Igarashi, and M. Asada. Adaptive hybrid control for visual servoing and force servoing in an unknown environment. *IEEE Robotics and Automation Magazine*, 5(4):39–43, 1998.

[2] T. Miyashita, K. Hosoda, and M. Asada. Reflective walk based on lifted leg control and vision-cued swaying control. In *Proc. of 1998 BSMEE International Symposium on Climbing and Walking Robots (CLAWAR'98)*, pages 349–354, 1998.

[3] P. K. Allen, A. T. Miller, P. Y. Oh, and B. S. Leibowitz. Using tactile and visual sensing with a robotic hand. In *Proceedings of the 1997 IEEE International Conference on Robotics and Automation*, pages 676–681, 1997.

[4] T. Matsuoka, T. Hasegawa, and K. Honda. A dexterous manipulation system with error detection and recovery by a multi-fingered robotic hand. In *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 418–423, 1999.

[5] Y. Yokokohji, M. Sakamoto, and T. Yoshikawa. Vision-aided object manipulation by a multifingered hand with soft fingertips. In *Proceedings of the 1999 IEEE International Conference on Robotics and Automation*, pages 3201–3208, 1999.

[6] K. Hosoda and M. Asada. Adaptive visual servoing for various kinds fo robot systems. In A. Casals and A. T. de Almeida, editors, *Experimental Robotics V*, pages 547–558. Springer, 1998.