

# Estimation of User's Request for Attentive Deskwork Support System

Yusuke Tamura, Masao Sugi, Tamio Arai and Jun Ota  
*The University of Tokyo Japan*

## 1. Introduction

Since the late 1990s, several studies have been conducted on intelligent systems that support daily life in the home or office environments (Sato et al., 1996; Pentland, 1996; Brooks, 1997). In daily life, people spend a significant amount of time at desks to operate computers, read and write documents and books, eat, and assemble objects, among other activities. Therefore it can be said that supporting deskwork by intelligent systems is of extreme importance. Many kinds of intelligent systems have been proposed to provide desktop support. In particular, augmented desk interface systems have been eagerly studied. DigitalDesk is one of the earliest augmented desk interface systems (Wellner, 1993). It requires a CCD camera and a video projector to integrate physical paper documents and electronic documents. Koike et al. proposed EnhancedDesk, which uses an infrared camera instead of a CCD camera to improve sensitivity to changes in lighting conditions and a complex background (Koike et al., 2001). In addition, Leibe et al. proposed one called Perceptive Workbench, which requires both a CCD and an infrared camera (Leibe et al., 2000), and Rekimoto proposed SmartSkin, which is based on capacitive sensing without cameras (Rekimoto, 2002).

Raghavan et al. proposed a system that requires a head-mounted display to show how to assemble products (Raghavan et al., 1999). These systems have been limited to show some information to the user. Ishii & Ullmer proposed an idea referred to as "tangible bits (Ishii & Ullmer, 1997)," which seeks to realize a seamless interface among humans, digital information, and the physical environment by using manipulable objects. Based on this idea, they proposed metaDESK (Ullmer & Ishii, 1997).

Pangaro et al. proposed a system called Actuated Workbench (Pangaro et al., 2002), and Noma et al. proposed one called Proactive Desk (Noma et al., 2004). Both systems convey only information to the user through movement of physical objects. They do not support the user from physical aspects.

On the other hand, especially in rehabilitation robotics, several studies have been conducted on supporting humans working at desks from a physical aspect (Harwin et al., 1995; Dallaway et al., 1995). Dallaway & Jackson proposed RAID (Robot for Assisting the Integration of Disabled people) workstation (Dallaway & Jackson, 1994). In RAID, a user selects an object through a GUI, and a manipulator carries it to the user. Ishii et al. proposed a meal-assistance robot for disabled individuals (Ishii et al., 1995). The system user points a laser attached to his head to operate a manipulator. Topping proposed a system, Handy 1,

which assists severely disabled people with tasks such as eating, drinking, washing, and shaving (Topping, 2002). In these systems, every time a user wants to be supported, the user is required to consciously and explicitly instruct their intention to the systems. Such systems are not really helpful.

Moreover, a few studies have focused on the physical act of passing an object from a human to a manipulator, or vice versa (Kajikawa et al., 1995; Agah & Tanie, 1997). These studies focused on the realization of human-like motion of the manipulators. When a user needs to be supported, on the other hand, the systems are required to support the user as fast as possible. The studies did not consider the requirement.

In this study, we propose a robotic deskwork support system that delivers objects properly and quickly to a user who is working at a desk. The intended applications of the proposed system are assembly, repair, simple experiment, etc. In such applications, the system often cannot know a sequence of used objects by workers in advance. To achieve the objectives, the system fulfils two primary functions: It estimates the user's intention, and it delivers objects to the user.

Intelligent systems are used by ordinary people; therefore, it is important that the systems be intuitive and simple to use. One of the most intuitive ways to control such systems is using gestures, especially pointing (Bolt, 1980; Cipolla & Hollinghurst, 1996; Mori et al., 1998; Sato & Sakane, 2000; Tamura et al., 2004; Sugiyama et al., 2005). Although pointing is intuitive, it is bothersome for a user to explicitly instruct the systems every time he/she wants to get objects. Furthermore, as pointing direction can be determined only when the user's hand and finger remain stationary, the recognition process takes long time. In the approach proposed here, the system estimates a user's intention inherent in his action without explicit instructions. In fact, the system 1) detects a user's act of reaching, 2) predicts the target object required by the user by measuring continuous movement of his body parts, especially hands and eyes, and finally 3) delivers the object to a user (Figure 1).

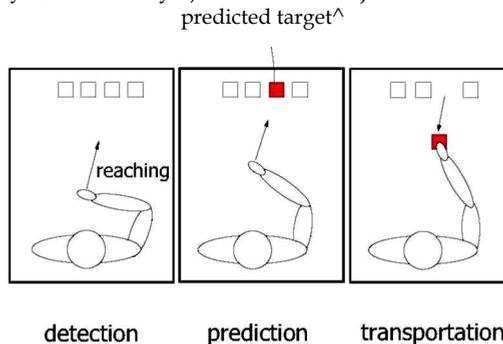


Fig. 1. Concept image of the proposed system

In this chapter, the first two items, involving detection and prediction, are mainly described and discussed.

For the third problem, it is unreasonable to use manipulators for carrying objects. Using manipulators for delivering objects has the following difficulties:

- Weight capacities of manipulators are generally low for their size.
- As manipulators move three-dimensionally, there is a tremendous danger in their high-speed movements.

- Because of the large size of manipulators, many manipulators cannot be operated simultaneously at a desk. Therefore, a manipulator can deliver a target object only after it grasps the object.

As a result, a system using manipulators cannot quickly and safely support a user. Moreover, small wheeled mobile robots present problems relative to speed and accuracy of movement.

One solution for the quick and accurate delivery of multiple objects to a user is to use movable trays driven with a Sawyer-type 2-DOF stepping motor (Sawyer, 1969). The motors are small and have high speed, positioning accuracy, and thrust.

The movable tray has high weight capacity, and moves only on a desk plane. Furthermore, because multiple trays can be placed simultaneously on a desk, multiple objects can be loaded on the trays. Therefore, a system using the movable trays can quickly and safely support a user.

In this chapter, we assume that our deskwork support system uses such movable trays and objects are loaded onto the trays. Assumed size of each tray is 130 x 135 x 25 (mm). In this study, we assume a normal size desk for the system. The width of a normal desk is at most 1200 (mm). According to this, the number of trays lined up in one row sideways is less than nine. In order to quickly deliver objects, a straight route is preferable for each tray. Even if the arrangement of the trays is schemed, the possible number of trays on a desk will be at most ten. We also assume that the distance between the trays and a user is greater than the user's reach. This assumption is for not obstructing a user's work.

In order to quickly deliver objects to a user, the trays are required not only to move fast but also to start early. Considering the speed of the user's hand and the movable trays, the preparation time for carrying objects (detection of the user's reach and prediction of the target object) should be less than a half of an average duration of reaching movements. According to a preliminary experiment, the average duration is about 0.8 (s) without any help. Therefore, the preparation time should be less than 0.4 (s).

In section 2, an algorithm used to detect reaching movement of a user is presented. A method used to predict a target object among multiple objects is described in section 3. In section 4, experiments for verifying the proposed method are described and discussed. In the experiments, the movable trays are not used. Experiments using the movable trays are presented in section 5. We conclude this chapter and refer to the future research in section 6.

## 2. Detection of human reaching movements

To deliver an object to a user, it is necessary that the system determine whether the user is performing an unrelated task or reaching for the object in question. When an individual reaches for an object, his hand and eyes move almost simultaneously toward the object. It has been reported that saccadic eye movement occurs before the onset of a reaching movement (Prablanc et al., 1979; Biguer et al., 1982; Abrams et al., 1990) and the saccade is followed about 100 (ms) later by a hand movement (Prablanc et al., 1979). In this study, therefore, a user's hand movements are measured to detect his reaching movements. When individuals perform tasks at desks, their hand movements are limited to a specific area, and their hands turn around frequently. When reaching for objects, on the other hand, individuals move their hands toward the outside of the working area at a high speed. The trajectories of hand movements are known to be relatively straight and smooth (Morasso, 1981). In addition to these characteristics of hand movements, eyes move toward a target

object to localize the position of the object for guiding hand movements (Abrams et al., 1990). Based on the facts reported above, in this study, the deskwork support system interprets a hand movement as a reaching movement if the following conditions are satisfied:

- The speed of a hand movement is rapid,
- The trajectory of a hand movement is relatively smooth and straight, and
- The directions of the gaze and hand (see Figure 2) are close, and the hand and gaze point are far from the head position.

We define a hand movement as the trajectory of the center of a user's hand. To measure hand movements, we use a color CCD camera attached to a ceiling. The RGB video data is first converted to the hue, saturation, and value (HSV) space. These values are then thresholded to acquire binarized hand images. After that, we apply a morphological erosion operator to the obtained hand region until it becomes smaller than a predetermined threshold value, and the center of a user's hand is given as the resulting region's center of mass. This procedure makes the hand's center insensitive to changes of the shape of the hand image due to a closing or opening motion of the hand (Oka et al., 2002). A tracking system that requires no physical contact is used to measure head and eye movements.

The parameters are defined in Figure 2.

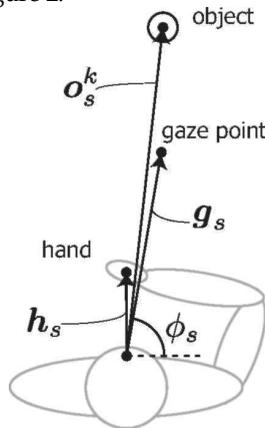


Fig. 2. Definition of parameters

Head and eye positions are measured three-dimensionally. However, in what follows, all positions are projected within a desk plane and are considered to be two-dimensional. Thus, all vectors are also two-dimensional.  $h_s$  is a vector from the user's head to user's hand at time  $s$ ;  $g_s$  is a vector from the user's head to a gaze point; and a vector from the user's head to object  $k$  is denoted by  $o_s$ . In this study,  $v_s$  the speed of a hand movement, is defined using the following equation:

$$v_s = \frac{\|h_s\| - \|h_{s-1}\|}{\Delta t} \quad (1)$$

where  $\Delta t$  is the sampling time of the camera.

To enable the system to determine whether a hand movement is a reach or some unrelated movement is difficult. Failures to detect the target movement can be eliminated by

integrating multiple criteria. Therefore, probabilities are established for three criteria, speed of hand movement, curvature of hand trajectory, and the relationship between the hand position and gaze point, which are used to detect the act of reaching.

**2.1 Speed of a hand movement**

The speed of a hand movement during reaching is much greater than that when performing tasks that occur close to the trunk of the body.

Therefore, we assume that the faster the relative speed of a user's hand to his head is, the higher the probability that the hand movement is an act of reaching will be. Here, we adopt a function whose output ranges between 0 and 1 and increases monotonically with its input as a probability function. Following this policy, we define  $R_v$ , the estimated probability from a hand speed at time  $s$ , as the following equation:

$$R_v = \frac{1}{1 + \exp\{-a(v_s - \beta)\}} \tag{2}$$

where  $a$  and  $\beta$  are parameters representing the motion characteristics of each user.

**2.2 Curvature of a hand trajectory**

In this study, the curvature of a user's hand trajectory is used as a criterion to indicate straightness and smoothness.

We regard the curvature of the circle passing through points  $h_{s-2}$ ,  $h_{s-1}$ , and  $h_s$  as the curvature of the hand trajectory at time  $s$  (Figure 3).

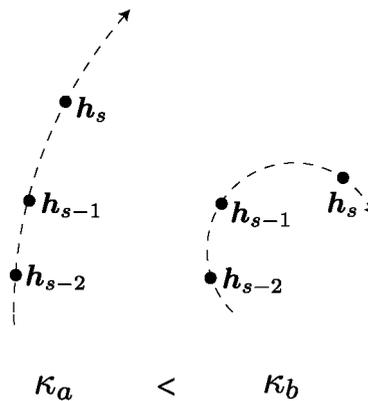


Fig. 3. Definition of the curvature of a hand trajectory at time

$\kappa_s$  is the curvature of the hand trajectory at time  $s$  calculated by the following equation:

$$\kappa_s = \frac{2|h_{s-1} \times h_{s-2} + h_{s-1} + h_{s-2} \times h_s|}{\sqrt{\|h_{s-1} - h_{s-2}\| \|h_s - h_{s-1}\| \|h_{s-2} - h_s\|}} \tag{3}$$

As reported earlier, reaching movements are generally straight and smooth. Therefore, the smaller the curvature of the hand trajectory, the greater the probability that the movement is

an act of reaching. Based on this, we define  $\hat{r}_s$ , the estimated probability from a hand trajectory at time  $s$ , with the following equation:

$$R_s = \gamma^{-K_s} \quad (4)$$

where  $\gamma$  is a parameter representing the motion characteristics of each user.

### 2.3 Relationship between the hand position and gaze point

When an individual reaches for an object, he first locates the object and then reaches for it. To map the location of the target object, a saccadic eye movement occurs about 100 (ms) before the reaching motion begins (Prablanc et al., 1979), as reported above. Because the trajectories of reaching movements are relatively straight (Morasso, 1981), the gaze direction and the direction from head to hand are supposed to be almost equal during the act of reaching. Furthermore, in the act of reaching, the hand position and gaze point are farther from an individual's body (Figure 4-a) than during other unrelated tasks (Figure 4-b).

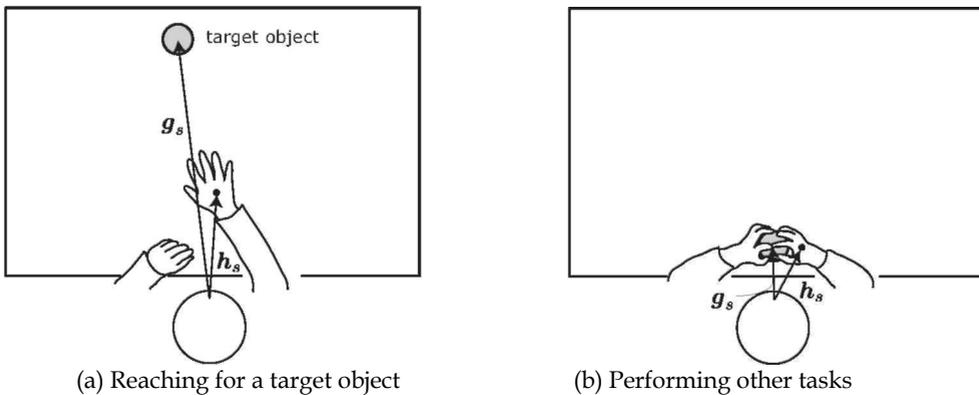


Fig. 4. Relationship between hand position and gaze

Based on these facts, we use the inner products of  $h_s$  and  $g_s$  at time  $s$  to detect acts of reaching.

$$I_s = h_s \cdot g_s \quad (5)$$

The large values of  $I_s$  suggest that the directions of the hand and gaze are close and the hand position and gaze point are far from the person's head.  $R_i$ , the estimated probability from the relationship between the hand position and gaze point at time  $s$ , is defined as follows:

$$R_i = \frac{1}{1 + \exp\{-\delta(I_s - \zeta)\}} \quad (6)$$

where  $\delta$  and  $\zeta$  are parameters representing the motion characteristics of each user.

### 2.4 Parameter determination

Because individuals differ in the motion characteristics of their hands and eyes, the five parameters reported above ( $\alpha, \beta, \gamma, \delta, \text{ and } \zeta$ ) are required to determine a specific individual's characteristics.

In this study, we determine the parameters through the following sequence. In this chapter, we take two parameters of  $R_v$  ( $\alpha$  and  $\beta$ ) as an example.

1. A hand movement while a user is performing some tasks and occasionally reaches for objects is measured.
2. Acquired velocity data  $v_s$  are discretized into several ranges, and times that a hand movement in a certain range of velocity is a reaching movement and those that the movement is not a reaching movement are counted respectively.
3.  $p(\text{reaching} | v_{B_i} \leq v_s < v_{B_{i+1}})$ , the likelihood that a hand movement in a certain range of velocity is a reaching movement, is plotted as Figure 5, where  $v_{B_i}$  is the boundary value between range  $i-1$  and range  $i$ . Here, because we assume that the system cannot determine a prior probability  $p(\text{reaching})$ , we normalize the observed data.
4. A curve represented as (2) is fitted to the plotted data points by the Levenberg-Marquardt method, which is a non-linear least-squares method, to acquire  $a$  and  $i_5$ . Here, we adopt the inverse of the probabilities

$p(\text{reaching} | v_{B_i} \leq v_s < v_{B_{i+1}})$  as weight factors for fitting. An example of fitting  $R_v$  curve

to the observed data points is shown in Figure 5.

The other parameters are determined in the same way.

## 2.5 Detection of reaching movements

The proposed system detects a hand movement as a reaching movement when  $R$ , the integrated probability at time  $s$ , exceeds the predetermined threshold value.  $R$  is defined as follows:

$$R = RvRsRi \quad (7)$$

To reduce missed detections, a large threshold value  $rj$  was used to indicate the distance between the user's head and hand, which acts as a safety net. Even if  $R$  does not exceed the predetermined threshold value, the hand movement is detected as an act of reaching when  $|| \mathbf{h}_s ||$  is larger than  $n$ .  $n$  is a sufficiently large value which is used to prevent missed detections, and it is empirically set to 400 (mm).

Moreover,  $R$  is set to 0 for a given length of time after detection, where the length of time is empirically set to 1.0 (s). This rule is used to prevent false detections.

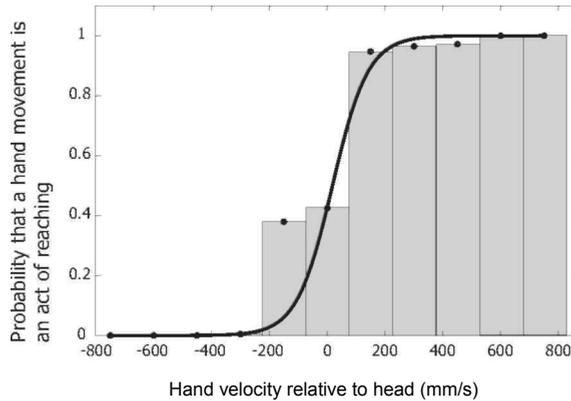


Fig. 5. Fitting a curve of the estimated  $R_v$  to the observed probabilities that a hand movement is a reaching movement

### 3. Prediction of the target object among multiple objects

People usually use multiple objects at a desk, and the sequence in which the objects are used is not predetermined. To deliver a required object to a user, a system must be able to correctly interpret a user's act of reaching at an early stage. It is also necessary that the system predict the required object as soon as possible.

Accurately predicting the user's hand trajectory seems to offer the right way to predict the target object. Studies of mathematical models for human arm trajectory planning have attracted considerable attention; such models include the minimum jerk model (Flash & Hogan, 1985), the minimum torque change model (Uno et al., 1989), and the minimum variance model (Harris & Wolpert, 1998). As these models do not consider human trunk movements and some of them require musculoskeletal parameters that are not easily acquired, it is difficult to apply them here.

In this study, knowledge of precise trajectories is not necessary; however it is necessary to identify the target object. There have been several researches on prediction of the target icon based on movements of a cursor in graphical user interfaces (Balakrishnan, 2004; Asano et al., 2005), however it cannot be applied to our situation because movements of hands and cursors are different. The following two assumptions are made to predict the target object.

- The approach rate of human hand to the target object is higher than to any other objects in the presence of multiple objects, and
- When an individual reaches for an object, his gaze directions are distributed around the direction of the object.

Based on these assumptions, certainty values from hand movements and eye movements are calculated and integrated probabilistically for each object.

#### 3.1 Inference from hand movements

The certainty that object  $O^i$  is a target object given the trajectory of the user's hand after starting a reaching movement  $H_s = [h_0, h_1, \dots, h_s]$  is defined as follows:

$$p(O^* = O^k | H_s) = \frac{f(o_s^k)}{\sum_{i=1}^N f(o_s^i)} \tag{8}$$

where  $O^*$  represents the target object and  $o_s^i$  is a vector from the user's head to the object  $O^i$  ( $i=1, 2, \dots, N$ ) at time  $s$ . Furthermore,  $f(o_s^k)$  is calculated with the following equations:

$$f(o_s^k) = \begin{cases} g(o_s^k)^2 & (g(o_s^k) < 0) \\ 0 & (g(o_s^k) \geq 0) \end{cases}, \tag{9}$$

$$g(o_s^k) = \sum_{j=1}^s \frac{\|o_j^k - h_j\| - \|o_{j-1}^k - h_{j-1}\|}{\|o_0^k - h_0\|}. \tag{10}$$

Equation (10) is transformed into the following equation:

$$g(o_s^k) = \frac{\|o_s^k - h_s\| - \|o_0^k - h_0\|}{\|o_0^k - h_0\|} \tag{11}$$

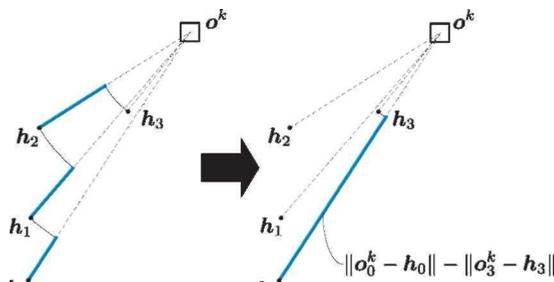


Fig. 6. Prediction of the target object on the basis of the ratio of the reduction in the distance between the hand and the object

The above equation yields the ratio of the reduction in the distance between the hand and the object  $O^k$  from the time the reaching movement is detected (Figure 6).

### 3.2 Inference from eye movements

In this study, we assume that the user's gaze direction  $\arg(g_s)$  follows a wrapped normal distribution (Gumbel et al., 1953), which is for circular data, with mean  $\arg(o_s)$ , the direction of the target object. Thus, we can represent the conditional probability density function for  $\arg(g_s) = \varnothing_s$  given target  $O^* = O^k$  as the following equation:

$$p(\arg(g_s) = \varnothing_s | O^* = O^k) = \frac{1}{\sqrt{2\pi\sigma}} \left\{ -\frac{(\varnothing_s - \arg(o_s^k) + 2\pi i)^2}{2\sigma^2} \right\}, \tag{12}$$

where  $\sigma$  is estimated using previous data for each user.

### 3.3 Integration of information from hand and eye movements

The system integrates the probabilities from the user's hand and eye movements based on Bayes' rule and predicts  $O$ , the target object, using the following equation:

$$\begin{aligned} \delta^* &= \arg \max p(O^* = O^k | H_s, \arg(g_s) = \emptyset_s) \\ &= \arg \max_{O^k} p(O^* = O^k | H_s) p(O^* = O^k | \arg(g_s) = \emptyset_s). \end{aligned} \quad (13)$$

## 4. Evaluation of the methods

To examine the usefulness of the methods, experiments were conducted. Generally, human motion in experiments is not reproducible. If the experiments are conducted in different conditions, therefore, it is impossible to fairly compare the methods. To tackle this problem, in this section, we conducted the estimation using three methods for the same human motion to fairly compare the methods. In the next section, we demonstrated the usefulness of the proposed method in the system that the movable trays can move.

### 4.1 Subjects

A total of 11 volunteers (10 males and 1 female, aged 21-42 years old) participated in the experiments. All subjects were right-handed, and three of them wore eyeglasses.

### 4.2 Experimental apparatus

An overhead digital color CCD camera (VCC-8350CL, CIS) measured the subjects' hand movements two-dimensionally.

For image processing, we used a Windows PC (Intel Xeon 3.0GHz Dual) with an image-processing board (GINGA digital CL-2, Linx) and image-processing software (HALCON7.0, MVTec).

To measure the subjects' head and eye movements, a head and eye tracking system (faceLAB4.2, Seeing Machines) that requires no physical contact and a Windows PC (Intel Pentium4 3.8GHz) were used. The frame rates for the measurement of the hand and eye movements are 30 (fps) and 60 (fps), respectively. Acquired three-dimensional data were projected to a desk plane and transformed into two-dimensional data.

### 4.3 Experimental procedure

In the experiments, subjects assembled a plastic model of a car from five types of subassemblies (Figure 7) five times, and the movements of each subject's hand, head, and eyes were recorded.

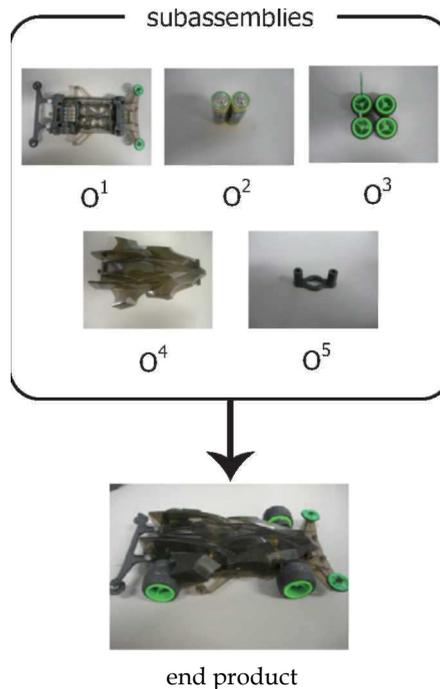


Fig. 7. Assembly of a plastic model from five types of subassemblies

The subjects reached with their dominant hand (right hand) for the subassemblies. In what follows, "hand" indicates the dominant hand (right hand). The subjects were asked to grasp only one type of subassembly at a time.

The arrangement of a subject, a desk, and the subassemblies are shown in Figure 8. During the experiments, the subjects were asked to sit at the desk and assembled a plastic model from five types of subassemblies in no particular order.

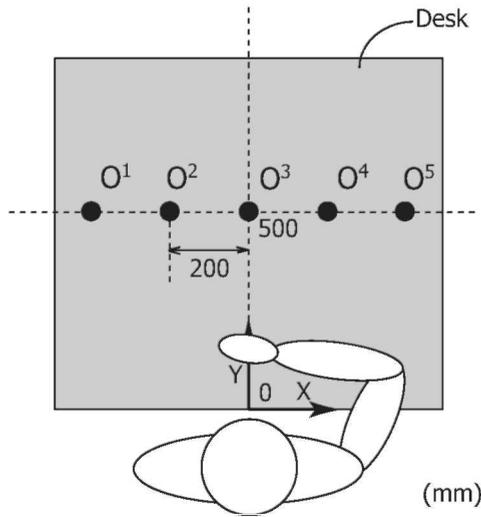


Fig. 8. Arrangement of an experimental subject, a desk, and five types of subassemblies ( $O^1$ ,  $O^2$ , ... ,  $O^5$ ). The initial position of a subject's head is around  $(0, -200)$  (mm).

**4.4 Experimental results**

An example of the observed trajectories of a hand, head, and gaze point is shown in Figure 9.

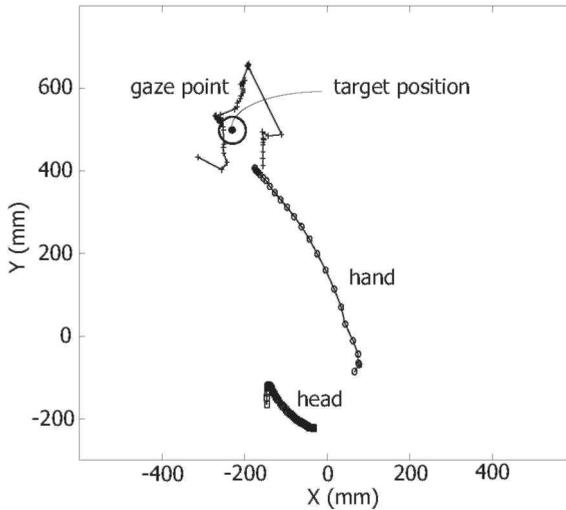


Fig. 9. Example of the observed trajectories of hand, head, and gaze point while a subject reached for  $O^2$

The observed trajectories of the hand are relatively straight and smooth as presented in the literature (Morasso, 1981), and the gaze points are distributed around the target object. As evident in Figure 9, gaze directions are distributed around the target direction. In the proposed estimation method, detection and prediction processes are independent each other. Thus we independently examined the usefulness of the detection method and prediction one.

**Detection of reaching movements**

We compared the detection performances of the proposed method with those of MD and MS. MD represents a detection method that uses only a threshold value of distance. MS represents a detection method uses only a threshold value of hand speed. In MD, we set a threshold value for the distance between the user's head and hand. When the distance  $| | h_s | |$  exceeds the predetermined threshold value, the hand movement is detected as a reaching movement. In MS, we set a threshold value for speed of a hand movement. When the speed exceeds the predetermined threshold value, the hand movement is detected as a reaching movement.

We used two metrics, required detection time and detection accuracy, to evaluate the methods. Because there is a tradeoff between false detections and missed detections, we defined the detection accuracy using a following equation:

$$DA = \frac{N_{correct}}{N_{correct} + N_{false} + N_{missed}}, \tag{14}$$

where  $N_{correct}$ ,  $N_{false}$ , and  $N_{missed}$  mean the number of correct detections, false detections, and missed detections, respectively. The threshold values of MD and MS were determined for each experimental subject to maximize DA. The required time for detection and calculated values of DA are shown in Table 1 and Table 2, respectively.

Subject	MD	MS	Proposed method
CR	0.363	0.187	0.173
GJ	0.348	0.126	0.196
HH	0.488	0.215	0.296
KO	0.407	0.215	0.297
MI	0.385	0.123	0.314
MK	0.462	0.230	0.234
MT	0.555	0.295	0.369
OJ	0.475	0.296	0.312
SH	0.264	0.097	0.185
SM	0.519	0.184	0.230
YY	0.374	0.258	0.212
Average	0.422	0.202	0.256

Table 1. Comparison of the required time (s) to detect an act of reaching

The average required time to detect an act of reaching is 0.256 (s). This is about 0.16 (s) shorter than that of MD and 0.05 (s) longer than that of MS. A total of five acts of reaching went undetected when the proposed method applied. The reason for the occurrence of the

missed detections is that false detections occurred just before the act of reaching and  $R$  was set to 0 as a result of the false detections.

Subject	MD				MS				Proposed method			
	C	M	F	DA	C	M	F	DA	C	M	F	DA
CR	23	2	19	0.52	22	3	43	0.32	25	0	2	0.93
GJ	24	4	14	0.57	23	5	15	0.54	28	0	4	0.88
HH	26	2	21	0.53	21	7	41	0.30	28	0	1	0.97
KO	29	8	7	0.66	20	17	37	0.27	37	0	2	0.95
MI	22	4	19	0.49	16	10	28	0.30	25	1	10	0.69
MK	38	2	13	0.72	33	7	76	0.28	40	0	7	0.85
MT	45	0	8	0.85	36	9	63	0.33	45	0	4	0.92
OJ	27	3	39	0.39	22	8	68	0.22	29	1	8	0.76
SH	37	1	7	0.82	32	6	24	0.52	37	1	1	0.95
SM	44	1	32	0.57	38	7	62	0.36	43	2	7	0.83
YY	25	0	15	0.63	15	10	63	0.17	25	0	0	1.0
Average	30.9	2.5	17.6	0.61	25.3	8.1	47.3	0.31	32.9	0.45	4.2	0.88

Table 2. Comparison of the detection accuracy (C: Correct detection, M: Missed detection, F: False detection)

However, the number of missed detections of the proposed method was less than 20% of MD and less than 6% of MS. Furthermore, the number of false detections of the proposed method was less than 25% of MD and less than 10% of MS. As a consequence, the average DA of the proposed method was much larger than that of MD and MS.

According to the results, MD is inferior to the proposed method in both accuracy and required time. Although MS can detect an act of reaching a little sooner than the proposed method, its accuracy of detection is much lower than the proposed method. Therefore, the proposed detection method is more useful than MD and MS.

#### Prediction of target objects

We compared the prediction performances of the proposed method with those of a method using only hand movements (MH) and one using only eye movements (ME). Figure 9 shows the relationship between the timing of the prediction and the rate of correct target predictions.

As shown in Figure 10, the rates of the proposed method and MH increased with time. On the other hand, the rates of ME decreased slightly with time. This could be attributed to the fact that the subjects sometimes did not see the target object in the final stage of the reaching movements.

Differences in the prediction performances of the three methods were tested using the paired t-test with a significance level of 5%. The difference in the average rates of correct prediction between the proposed method and MH was statistically significant from 0 (s) to 0.1 (s) ( $p < 0.05$ ), and the difference between the proposed method and ME was also significant after 0.067 (s) ( $p < 0.05$ ).

As reported before, it takes about 0.26 (s) to detect a user's reaching movements. Therefore, it is necessary for the system to predict the target object within 0.14 (s) in order to meet the time requirement described in section 1. At 0.133 (s) after detection, while the rates of correct

predictions of MH and ME were 84% and 62%, respectively, the proposed method's rate of correct predictions was 92%.

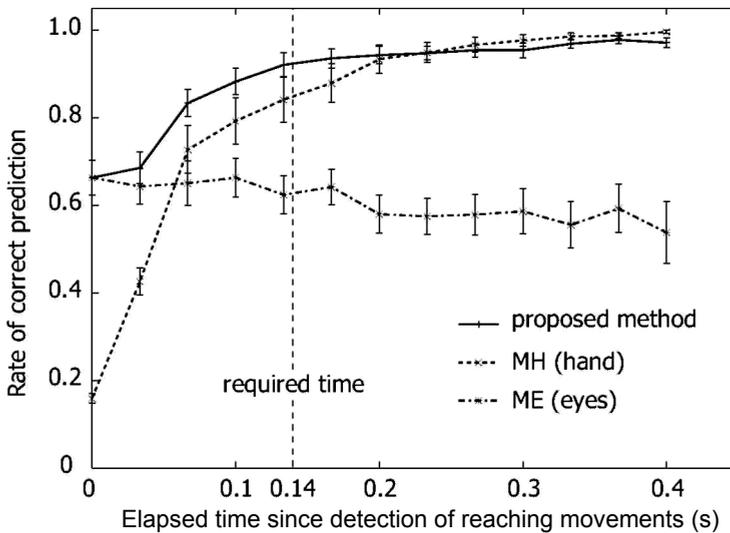


Fig. 10. Average rate of correct prediction of the targets with the standard error of the mean, N=11

The difference in the rates of correct predictions between MH and the proposed method at the time was not significant, but a statistical trend in which the proposed method exceeds MH ( $p=0.069$ ) was observed.

### 5. Application possibility to the deskwork support system

To verify the possibility of applying the proposed methods to our deskwork support system, experiments using movable trays were conducted.

#### 5.1 Subjects

A total of 4 male volunteers (aged 24-33 years old) participated in the experiments. All subjects were right-handed, and they had also participated in the experiments stated in section 4.

#### 5.2 Experimental apparatus

Our deskwork support system is shown in Figure 10.

To measure movements of the subjects' hands, heads, and eyes, we used the same apparatus with section 4.

To deliver objects to the subjects, five self-moving trays, which are driven by a Sawyer-type 2-DOF stepping motor, were used. The trays move on a motor *platen* (i.e. iron plate) with its

size being 1200 x 1000 (mm). The size of each tray is 130 x 135 x 25 (mm). The trays can move 1000 (mm/s) at maximum for each axis.

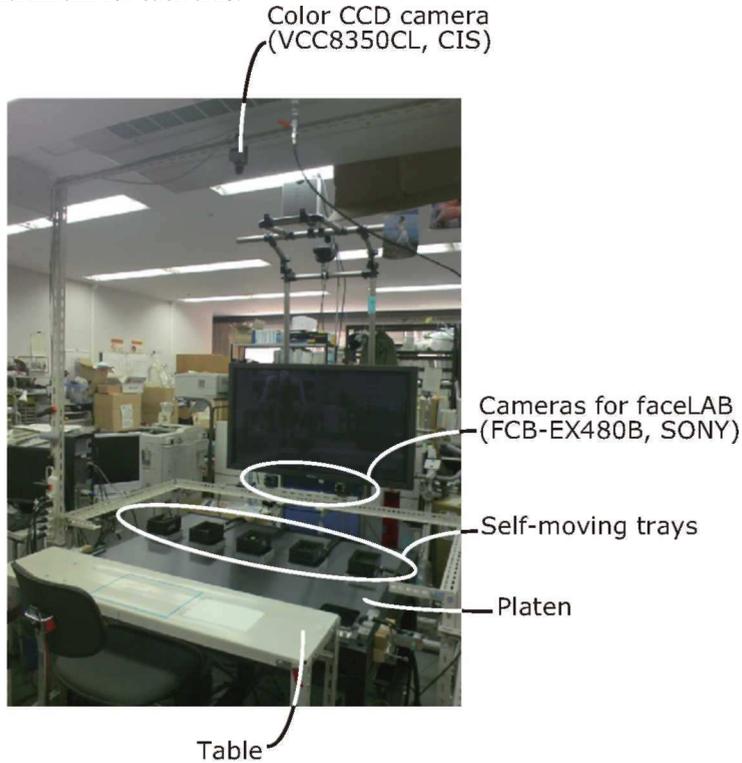


Fig. 11. An overview of the implemented deskwork support system

### 5.3 Experimental procedure

In the experiments, the subjects assembled a plastic model of a car from five types of subassemblies four times as well as the experiments described in section 4. The subassemblies (Figure 7) were loaded on each tray, and when a subject reaches for an object, the deskwork support system detects an act of reaching, predicts a target object, and delivers the object using a self-moving tray.

In the experiments, we set the duration for prediction to 0.133 (s), and we use the same parameters for the prediction process as in section 4. In other words, the parameters were obtained in the reachable target condition.

The initial arrangement of a subject and the subassemblies on the trays are shown in Figure 12. In this experiment, the subassemblies are located farther from the position of the subjects than the experiment in section 4.

During the experiments, the subjects were asked to sit at the desk and assembled a plastic model in no particular order, and place the end product on tray 3.

**5.4 Experimental results**

An example of the experimental scene where the proposed method was applied is shown in Figure 12.

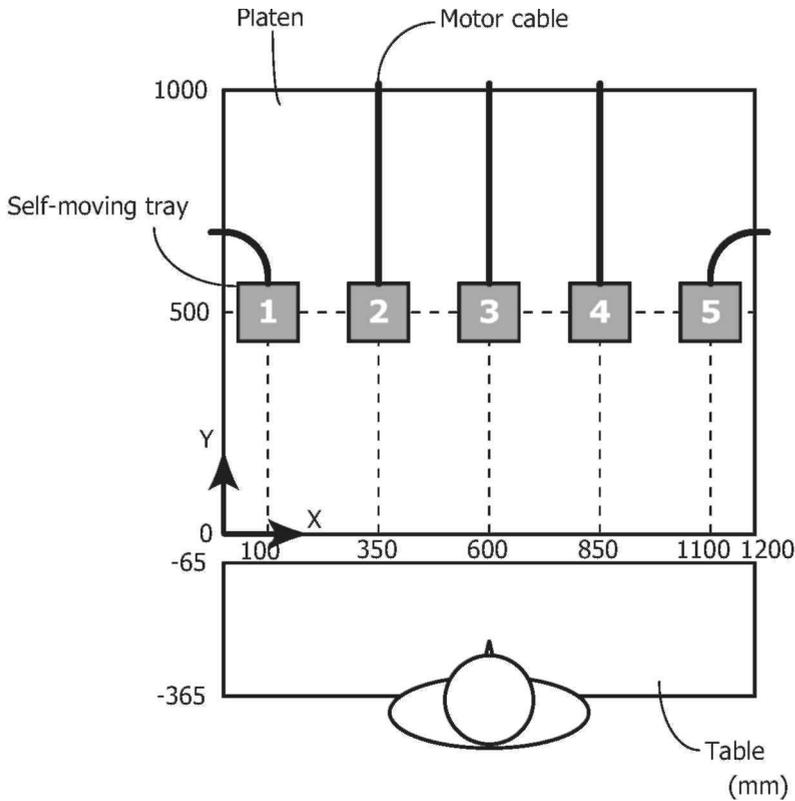


Fig. 12. Arrangement of an experimental subject and five trays (O is on tray 1) The calculated values of DA are shown in Table 3.

Subject	Correct	Missed	False	DA
CR	40	9	0	0.82
MI	40	4	0	0.91
MK	40	4	0	0.91
SM	40	3	0	0.93
Average	40	5	0	0.89

Table 3. Detection accuracy when applied to the deskwork support system

The average value of DA was 0.89. This was almost the same as the result in section 4. The proposed detection method can be applied to the deskwork support system in which the self-moving trays deliver the objects.

We also test the validity of the proposed prediction method. The calculated rates of correct prediction of the targets are shown in Table 4.



Fig. 13. An example of the experimental overview. The subject was reaching for tray 3, and the tray was moving towards the subject.

Subject	Correct	False	Correct rate
CR	40	2	0.95
MI	40	0	1.0
MK	40	1	0.98
SM	40	2	0.95
Average	40	1.25	0.97

Table 4. Rate of correct prediction of the targets when applied to the deskwork support system

The average rate of correct predictions was 97%. This was higher than the result in section 4. This result confirms that the proposed prediction method can be applied to the deskwork support system.

Based on the above experimental results, the proposed method is useful in the deskwork support system, and the parameters acquired from the reachable target condition can be applied to the unreachable one.

## 6. Conclusion

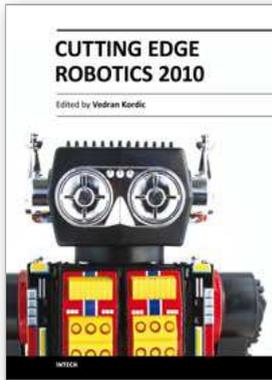
We have presented methods to detect an act of reaching among other hand movements and to predict target objects based on measurements of a user's hand and eye movements. In the detection method, we adopted speed, the smoothness and straightness of a user's hand movements, and the relationship between hand and eye movements. The usefulness of the proposed method was experimentally demonstrated.

In the future, an error recovery algorithm should be developed for more reliable deskwork support system.

## 7. References

- Abrams, R.A.; Meyer, D.E. & Kornblum, S. (1990). Eye-hand coordination: Oculomotor control in rapid aimed limb movements. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 16, No. 2, 248-267
- Agah, A. & Tanie, K. (1997). Human interaction with a service robot: Mobile-manipulator handing over an object to a human. *Proceedings of the 1997 IEEE International Conference on Robotics and Automation*, pp. 575-580
- Asano, T.; Sharlin, E.; Kitamura, Y.; Takashima, K. & Kishino, F. (2005). Predictive interaction using the Delphian desktop. *Proceedings of the 18<sup>th</sup> Annual ACM Symposium on User Interface Software and Technology*, pp. 133-141
- Balakrishnan, R. (2004). "Beating" Fitts' law: Virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies*, Vol. 61, No. 6, 857-874
- Biguer, B.; Jeannerod, M. & Prablanc, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. *Experimental Brain Research*, Vol. 46, 301-304
- Bolt, R.A. (1980). "Put-that-there": Voice and gesture at the graphics interface. *Proceedings of the 7<sup>th</sup> Annual Conference on Computer Graphics and Interactive Techniques*, pp. 262-270
- Brooks, R.A. (1997). The intelligent room project. *Proceedings of the 2<sup>nd</sup> International Cognitive Technology Conference*, pp. 271-278
- Cipolla, R. & Hollinghurst, N.J. (1996). Human-robot interface by pointing with uncalibrated stereo vision. *Image and Vision Computing*, Vol. 14, 171-178
- Dallaway, J.L. & Jackson, R.D. (1994). The user interface for interactive robotic workstations. *Proceedings of the 1994 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1682-1686
- Dallaway, J.L.; Jackson, R.D. & Timmers, P.H.A. (1995). Rehabilitation robotics in Europe. *IEEE Transactions on Rehabilitation Engineering*, Vol. 3, No. 1, pp. 35-45
- Flash, H. & Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *Journal of Neuroscience*, Vol. 5, No. 7, 1688-1703
- Gumbel, E.; Greenwood, J. & Durand, D. (1953). The circular normal distribution: Theory and tables. *Journal of the American Statistical Association*, Vol. 48, No. 261, 131-152
- Harris, C.M. & Wolpert, D.M. (1998). Signal-dependent noise determines motor planning. *Nature*, Vol. 394, 780-784
- Harwin, W.S.; Rahman, T. & Foulds, R.A. (1995). A review of design issues in rehabilitation robotics with reference to North American research. *IEEE Transactions on Rehabilitation Engineering*, Vol. 3, No. 1, 3-13
- Ishii, H. & Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits, and atoms. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 234-241
- Ishii, S.; Tanaka, S. & Hiramatsu, F. (1995). Meal assistance robot for severely handicapped people. *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1308-1313
- Kajikawa, S.; Okino, T.; Ohba, K. & Inooka, H. (1995). Motion planning for hand-over between human and robot. *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 193-199
- Koike, H.; Sato, Y. & Kobayashi, Y. (2001). Integrating paper and digital information on EnhancedDesk: A method for realtime finger tracking on an augmented desk system. *ACM Transactions on Computer-Human Interaction*, Vol. 8, No. 4, 307-322

- Leibe, B.; Starner, T.; Ribarsky; Wartell, Z.; Krum, D.; Singletary, B. & Hodges, L. (2000). The perceptive workbench: Toward spontaneous and natural interaction in semi-immersive virtual environments. *Proceedings of the IEEE Virtual Reality 2000*, pp. 13-20
- Morasso, P. (1981). Spatial control of arm movements. *Experimental Brain Research*, Vol.42, 223-227
- Mori, T.; Yokokawa, T. & Sato, T. (1998). Recognition of human pointing action based on color extraction and stereo tracking. *Intelligent Autonomous Systems 5*, Kakazu, Y.; Wada, M. & Sato, T. (Eds.), pp. 93-100, IOS Press
- Noma, H.; Yoshida, S.; Yanagida, Y. & Tetsutani, N. (2004). The proactive desk: A new haptic display system for a digital desk using a 2-DOF linear induction motor. *Presence*, Vol. 13, No. 2, 146-153
- Oka, K.; Sato, Y. & Koike, H. (2002). Real-time fingertip tracking and gesture recognition. *IEEE Computer Graphics and Applications*, Vol. 22, No. 6, 64-71
- Pangaro, G.; Aminzade, D.M. & Ishii, H. (2002). The actuated workbench: Computer-controlled actuation in tabletop tangible interfaces. *Proceedings of the 15<sup>th</sup> Annual ACM Symposium on User Interface Software and Technology*, pp. 181-190
- Pentland, A. (1996). Smart rooms. *Scientific American*, Vol. 274, No. 4, 54-62
- Prablanc, C.; Echallier, J.F.; Komilis, E. & Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target - I. Spatio-temporal characteristics of eye and hand movements and their relationships when varying the amount of visual information. *Biological Cybernetics*, Vol. 35, 113-124
- Raghavan, V.; Molineros, J. & Sharma, R. (1999). Interactive evaluation of assembly sequences using augmented reality. *IEEE Transactions on Robotics and Automation*, Vol. 15, No. 3, 435-449
- Rekimoto, J. (2002). SmartSkin: An infrastructure for freehand manipulation on interactive surfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 113-120
- Sato, S. & Sakane, S. (2000). A human-robot interface using an interactive hand pointer that projects a mark in the real work space. *Proceedings of the 2000 IEEE International Conference on Robotics and Automation*, pp. 589-595
- Sato, T.; Nishida, Y. & Mizoguchi, H. (1996). Robotic room: Symbiosis with human through behavior media. *Robotics and Autonomous Systems*, Vol. 18, 185-194 Sawyer, B.A. (1969). Magnetic positioning device. *US patent*, 3,457,482
- Sugiyama, O.; Kanda, T.; Imai, M.; Ishiguro, H. & Hagita, N. (2005). Three-layered draw-attention model for humanoid robots with gestures and verbal cues. *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2140-2145
- Tamura, Y.; Sugi, M.; Ota, J. & Arai, T. (2004). Deskwor support system based on the estimation of human intentions. *Proceedings of the 13<sup>th</sup> IEEE International Workshop on Robot and Human Interactive Communication*, pp. 413-418 Topping, M. (2002). An overview of the development of Handy 1, a rehabilitation robot to assist the severely disabled. *Journal of Intelligent and Robotic Systems*, Vol. 34, 253-263
- Ullmer, B. & Ishii, H. (1997). The metaDESK: Models and prototypes for tangible user interfaces. *Proceedings of the 10<sup>th</sup> Annual ACM Symposium on User Interface Software and Technology*, pp. 223-232
- Uno, Y.; Kawato, M. & Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics*, Vol. 61, 89-101 Wellner, P. (1993). Interacting with paper on the DigitalDesk. *Communications of the ACM*, Vol. 36, No. 7, 87-96



## **Cutting Edge Robotics 2010**

Edited by Vedran Kordic

ISBN 978-953-307-062-9

Hard cover, 440 pages

**Publisher** InTech

**Published online** 01, September, 2010

**Published in print edition** September, 2010

Robotics research, especially mobile robotics is a young field. Its roots include many engineering and scientific disciplines from mechanical, electrical and electronics engineering to computer, cognitive and social sciences. Each of this parent fields is exciting in its own way and has its share in different books. This book is a result of inspirations and contributions from many researchers worldwide. It presents a collection of a wide range of research results in robotics scientific community. We hope you will enjoy reading the book as much as we have enjoyed bringing it together for you.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yusuke Tamura, Masao Sugi, Tamio Arai and Jun Ota (2010). Estimation of Users Request for Attentive Deskwork Support System, Cutting Edge Robotics 2010, Vedran Kordic (Ed.), ISBN: 978-953-307-062-9, InTech, Available from: <http://www.intechopen.com/books/cutting-edge-robotics-2010/estimation-of-users-request-for-attentive-deskwork-support-system>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.