Advanced Robotics, Vol. 8, No. 3, pp. 263–283 (1994)

Paper

Haptic sensing system with active perception

YUTAKA SAKAGUCHI

Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo, 7–3–1, Hongo, Bunkyo-ku, Tokyo 113, Japan

Received for *AR* 10 April 1993; accepted 13 November 1993

Abstract—Human haptic perception is not caused by simple mechanical stimulation to the skin; it is achieved by integrating sensory information from various receptors (such as mechano-receptors and thermo-receptors) and by observing objects in various ways. The author constructed sensing systems which simulate human haptic processes in order to clarify the mechanism of such sensory integration and active perception. In this article, first the author formalizes these processes from a theoretical point of view and constructs an 'intentional observation' algorithm on the basis of that formalization. This algorithm is to select appropriate sensors from many sensors based on information criteria to recognize objects more accurately by fewer observations. Second, the author describes two kinds of haptic sensing systems which recognize objects' materials and surface textures utilizing actively several sensor devices, and shows that the proposed algorithm is effective in these actual systems. Some related problems are also discussed.

1. INTRODUCTION

We human beings realize an effective sensory information processing by integrating various sensory information and by collecting selectively significant information for accomplishing the task. Moreover, we deal with a vast amount of information from the external world by discarding much redundant information. As many psychologists have pointed out [1, 2], these are remarkable characteristics of human information processing and play significant roles in human perception.

The purpose of this study is to clarify the mechanism of such sensory integration and active perception.

Here, the author would like to note that 'active perception' differs from 'active sensing': it does not mean moving sensors or emitting energy in sensing processes; it means perceiving objects through various observations according to what we want to know.

In light of such characteristics of human brains, engineering researchers have studied 'multi-sensor integration' [3], 'sensor fusion' or 'intentional sensing' [4]. Especially in robotics, object recognition, autonomous locomotion, visuo-motor co-ordination, visuo-tactile fusion, etc., have been discussed from these points of view.

In order to reveal the mechanism, the author proposed neural network models for space perception and visual recognition [5–7]. One is a model for a figure recognition process which observes local features selectively, forms a global internal image from the observed features using neural dynamics and recognizes the figure based on the internal image [7]. This model obtains useful local features actively and integrates them into internal representation in recognizing a figure.

In the present article, the author concentrates on haptic perception processes.

When perceiving objects by touch, we human beings utilize not only deformation of the skin, but also information from various sensory organs including thermo-receptors and proprioceptors. Moreover, we make various observation behaviors, such as rubbing and pushing, according to what property we want to know. Therefore, the haptic perception process is an active perception process (this is why the author does not use the word 'tactile' which implies a passive sense, but the word 'haptic') and is a suitable subject for the author's purpose.

Touch sensors are developed mainly in the robotics field. However, many of them are made just for detecting 'contact', for example, noticing that the robot body collides with some obstacles and examining whether the robot hand catches an object. They are not made from a viewpoint of haptics. Not to lose this viewpoint, the author has tried to construct haptic sensing systems which perceive objects by actively utilizing various sensor devices. Since these systems treat feeling of touch, in addition, they will provide useful suggestions to make a standard for representing human haptic feeling objectively.

In the first part of this article, the author formalizes an iterative sensory integration process from a theoretical point of view (Section 2) and constructs an algorithm to select iteratively useful sensors for accomplishing the recognition task (Section 3), where the active perception process is described as a sort of iterative experimental design.

In the second part of this article, the author introduces two kinds of haptic perception systems (Section 4). They observe object properties with several sensor devices and perceive object materials and surface textures using the active perception algorithm. It is shown that the proposed algorithm is effective in the actual sensing systems: they can discriminate objects more accurately by fewer observations with the algorithm. Finally, some problems related to sensory integration systems are discussed (Section 5).

2. A THEORETIC SCHEME OF SENSORY INTEGRATION AND ACTIVE PERCEPTION

As mentioned in the previous section, sensory integration and active perception are remarkable characteristics of human brains. However, since we human beings realize them based on extensive knowledge and intelligence which have developed since our infancy, it is too difficult to model the whole of the mechanisms. Here, the author extracts their essence and formalizes them from a viewpoint of information theory.

Figure 1 shows schematically the process of sensory integration and active perception [8]. The system observes an object with various sensors. Each sensor receives signals which undergo some pre-processing and sends them to the recognition center. The recognition center guesses the object through integrating information from the local sensors.

Here, it is assumed that the recognition center cannot use more than one sensor at the same time. Then, the system needs to select the sensors one by one and to collect

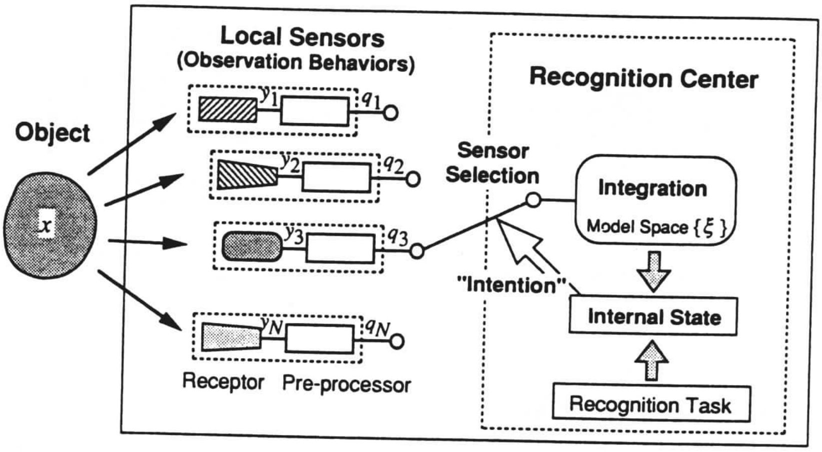

Figure 1. Schematic model of sensory integration and active perception processes. The system observes an object by various sensors. Each local sensor receives signals, puts them through some pre-processing, and sends them to the recognition center. The recognition center guesses the objects by integrating information from the sensors. The system selects one sensor at a time and collects the information iteratively until it gets a recognition result.

information iteratively until it gets a recognition result. It is also assumed that the object state is not changed by observation.

To simplify the following discussion, it is supposed that sensor signals are quantized into discrete values at the local sensors. Besides, when a sensor has controllable parameters, it is treated as a set of separate sensors (Fig. 2). For instance, a situation that a single camera observes an object from various positions can be considered as many cameras observing the object from their own respective positions. When the system utilizes one sensor in various manners, the author's scheme regards each manner as one 'sensor'. Accordingly, the local sensors in the scheme represent not only physical sensors but observation behaviors.

Now, let x, y_i $(i = 1, \ldots, N)$, and q_i $(i = 1, \ldots, N)$ denote the presented object, the signal detected by the i-th sensor and quantized value of y_i, respectively, where N means the number of the sensors, and ξ means the model represented in the recognition center and corresponds to object x in the external world. We call $X = \{x\}$, $Y_i = \{y_i\}$, $Q_i = \{q_i\}$, and $\Xi = \{\xi\}$ object space, sensory signal space, quantized sensory signal space, and model space, respectively. The numbers of the objects and models are both M and that of quanta in Q_i is K_i. In the following, the author equates the object space and the model space ($x \equiv \xi$ and $X \equiv \Xi$) and, for simplicity's sake, calls a quantized sensory signal space just a sensory space: the (raw) sensory signal will never be reffered to below.

It may be noted that the observation process at each local sensor can be regarded as a process of 'encoding' the object state into sensory signals. However, it encodes only restricted information that can be detected by itself.

On the other hand, the guessing process at the recognition center is regarded as 'decoding'. Since the message from a single local sensor contains only partial information, the recognition center needs to collect signals from various sensors and to integrate

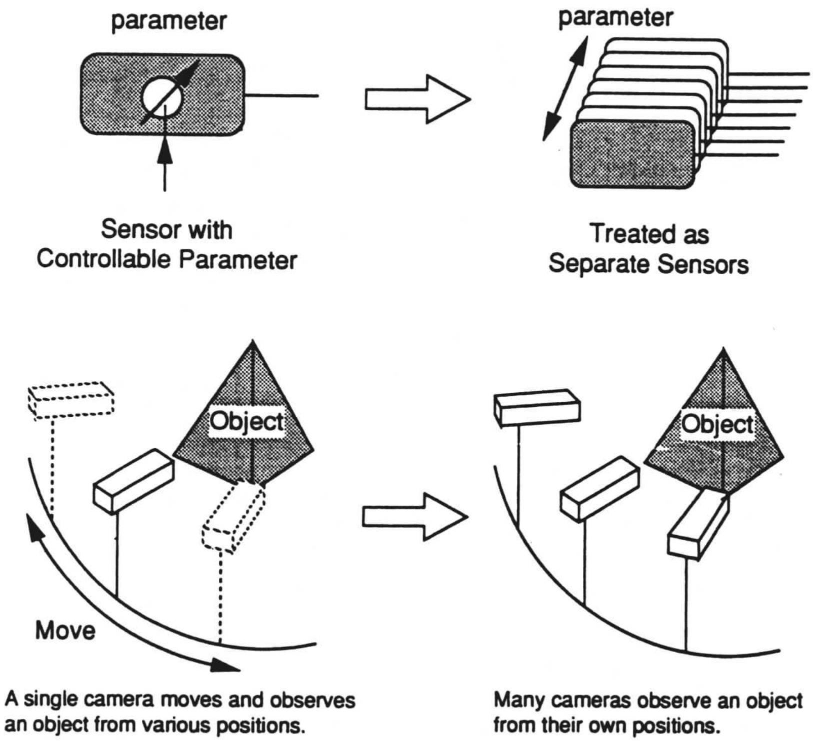

Figure 2. Sensor with controllable parameter. If a system utilizes a sensor in various manners, each observation manner is regarded as a 'local sensor'. For example, a camera which moves and observes an object from various positions is treated as multiple sensors, each of which observes the object from its own position.

them in order to know the object state. This 'integration' process is one of the most significant points of the system.

On the assumption that the system can use only one sensor at a time, how to select sensors from the available local sensors is an important problem because its performance is much affected by whether it selects appropriate sensors or not. Such sensor selection is the second problem. It becomes rather important when the system must pass judgement in a short time. Also, it is closely related to the channel selection problem in communication systems and to the feature selection problem in pattern recognition systems.

In order to integrate the sensor signals, moreover, it is necessary to understand the relation between the objects and sensory signals and to make reliable 'decoders'. In other words, the system should build an appropriate internal model of the external world. This is another significant point of sensory integration systems.

Accordingly, the problems are summarized in (1) how to integrate information from local sensors and to estimate an object state, (2) how to select local sensors, and (3) how to construct an appropriate internal model. The present article focuses on the former two problems, especially on the second problem. The third problem will be mentioned in Section 5.

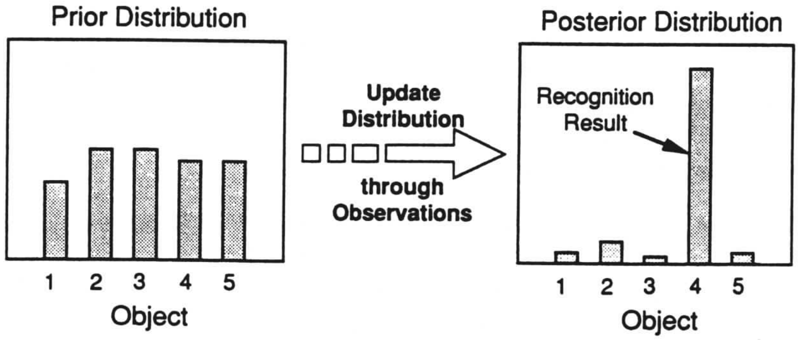

Figure 3. Update of probability distribution of model space. The system updates the probability distribution of the model space using Bayes' rule. As the system repeats the observation, the entropy of the model space becomes smaller and smaller. When the entropy becomes smaller than a certain threshold, the system puts out the model whose probability is the largest as the recognition result.

The author uses the Bayesian method as a solution to the first problem. Concretely, the system holds the probability of the model space $\pi(\xi)$ which represents the recognition state. Starting from some initial distribution $\pi_0(\xi)$, the system updates the probability distribution using Bayes' rule after every observation (see Fig. 3).

To measure the ambiguity of the recognition, the author considers the entropy of the model space. The system repeats observations until the entropy becomes smaller than a threshold value. When this condition is satisfied, the system regards the model whose probability is the most at the time as the recognition result.

Based on this formulation, the author sets about the second problem.

What sensor should be selected depends on what the system want to know. This means that observation behavior varies according to the recognition task. Supposing the system has the 'intention' to recognize objects, observation behaviors are affected by the intention. In this sense, active perception can be thought of as an intentional process, and hence the author calls the active observation strategy 'intentional observation'. The concrete algorithm is described in the next section.

3. AN ALGORITHM FOR INTENTIONAL OBSERVATION

3.1. Sensor selection criterion

In order to measure how much each sensor contributes to recognition, the author uses an information criterion, that is, mutual information or averaged symmetric divergence. Hutchinson and Kak dealt with a similar problem [9]; they used the Dempster–Shafer theory to combine the 'evidence' obtained by each local sensor and to measure the ambiguity of the sensory information.

Mutual information between the model space and the i-th sensory space shows how much the entropy of the model space is reduced on average when the system observes the object using the i-th sensor. By selecting the sensor which has the most mutual information, it is expected that the entropy of the model space will be reduced the most. On the other hand, symmetric divergence indicates 'distance' between two probability distributions and the system selects a sensor whose averaged divergence is the largest. It is known that these two criteria are related to each other [10]. Although it is also useful to use a min-max estimation instead of these average estimations, a min-max method is not discussed in this article.

3.2. Sensor selection algorithm

First, the author deals with the case in which the sensor signals have no mutual interaction.

Let $\pi_t(\xi)$ and L_t denote the probability distribution of the model space and the entropy of the model space at time t, respectively. L_t is written as

$$L_t = \sum_{\xi=1}^{M} -\pi_t(\xi)\log \pi_t(\xi). \tag{1}$$

If the system observes the object with the i-th sensor and gets a signal q_i, then the entropy will become

$$L_{t+1}(q_i; i) = \sum_{\xi=1}^{M} -\pi_{t+1}(\xi, q_i)\log \pi_{t+1}(\xi, q_i), \tag{2}$$

where

$$\pi_{t+1}(\xi, q_i) = \frac{p_i(q_i \mid \xi)}{P(q_i)}\pi_t(\xi),$$

$$P(q_i) = \sum_{\xi=1}^{M} p_i(q_i \mid \xi)\pi_t(\xi)$$

and $p_i(q_i \mid \xi)$ is the conditional probability of signal q_i conditioned on object ξ. The system cannot know the true value of q_i before it actually observes the object. Then, the system can only estimate its expected value,

$$\langle L_{t+1}(i)\rangle = \sum_{q_i=1}^{K_i} P(q_i)L_{t+1}(q_i; i). \tag{3}$$

Mutual information $I_t(i)$ between the model space and the i-th sensory space is the expected entropy reduction of the model space if the system observes the object with the i-th sensor and is written as

$$\begin{aligned} I_t(i) &= L_t - \langle L_{t+1}(i)\rangle \\ &= \sum_{\xi=1}^{M} \pi_t(\xi) \sum_{q_i=1}^{K_i} p_i(q_i \mid \xi)\log\frac{p_i(q_i \mid \xi)}{\sum_{\xi=1}^{M}\pi_t(\xi)p_i(q_i \mid \xi)}. \end{aligned} \tag{4}$$

On the other hand, symmetric divergence $g_i(\xi_1, \xi_2)$ $(\xi_1, \xi_2 \in \Xi)$ between $p_i(q_i \mid \xi_1)$ and $p_i(q_i \mid \xi_2)$ is given as

$$g_i(\xi_1, \xi_2) = \sum_{q_i=1}^{K_i} (p_i(q_i \mid \xi_1) - p_i(q_i \mid \xi_2)) \log \frac{p_i(q_i \mid \xi_1)}{p_i(q_i \mid \xi_2)}. \tag{5}$$

$g_i(\xi_1, \xi_2) \geqslant 0$ always holds and $g_i(\xi_1, \xi_2)$ is equal to 0 if and only if

$$p_i(q_i \mid \xi_1) \equiv p_i(q_i \mid \xi_2).$$

However, it is not 'distance' in a strict sense because it does not satisfy triangle inequality. Averaging $g_i(\xi_1, \xi_2)$ over all pairs of the models (ξ_1, ξ_2), we get the averaged symmetric divergence,

$$G_t(i) = \sum_{\xi_1, \xi_2}^{M} \pi_t(\xi_1)\pi_t(\xi_2) g_i(\xi_1, \xi_2). \tag{6}$$

Note that $g_i(\xi_1, \xi_2)$ is independent of $\pi_t(\xi)$: the system can calculate $g_i(\xi_1, \xi_2)$ beforehand as long as the conditional probabilities $p_i(q_i \mid \xi)$ are fixed.

The system calculates $I_t(i)$ (or $G_t(i)$) for every i and selects one which gives the largest value of $I_t(i)$ (or $G_t(i)$). Let $\bar{i}$ denote the index of the selected sensor.

After getting sensor data $q_{\bar{i}}$ actually by the selected sensor $\bar{i}$, the system calculates posterior distribution $\pi_{t+1}(\xi)$ using Bayes' rule as

$$\pi_{t+1}(\xi) = \frac{p_{\bar{i}}(q_{\bar{i}} \mid \xi)}{\sum_{\eta=1}^{M} p_{\bar{i}}(q_{\bar{i}} \mid \eta)\pi_t(\eta)} \pi_t(\xi), \tag{7}$$

and updates the entropy using $\pi_{t+1}(\xi)$:

$$L_{t+1} = \sum_{\xi=1}^{M} -\pi_{t+1}(\xi) \log \pi_{t+1}(\xi). \tag{8}$$

If L_{t+1} is smaller than threshold θ, the system selects the model whose probability is the most as the recognition result. If not, it continues the observation process.

In generally, we may set the initial distribution $\pi_0(\xi)$ as a uniform distribution, that is, $\pi_0(\xi) = 1/M$. Given some prior knowledge of the object, the system can determine the initial distribution according to the knowledge.

In this article, the author uses discrete distributions as the probability distributions because the signals are quantized into discrete values. If the sensing signals are subject to Gaussian distributions, the system can estimate and update the distributions in a parametric manner and hence reduce the amount of computation.

The sensor selection process is interpreted intuitively as follows. Using prior distribution $\pi_t(\xi)$ and conditional probabilities $p_i(q_i \mid \xi)$, the system calculates the probability distribution $P(q_i)$ $(\equiv \sum_\xi p_i(q_i \mid \xi)\pi_t(\xi))$ which predicts what signal tends to be detected by the i-th sensor. If the distribution concentrates on some specific values, the signals detected by the sensor would be almost the same for all the objects and accordingly the sensor has little meaning to use. If the probability is broadly distributed, to the contrary, it is expected that the system can restrict the candidates in the model space using the sensor. The proposed algorithm selects such a sensor.

Next, suppose that two sensor signals are rather correlated to each other. Assume that the system uses one of them and updates the distribution of the model space. Then, the posterior distribution of the other sensory space will be narrower and accordingly mutual information will be reduced. It means that the system does not select the sensor whose signal is correlated to the signals already used.

3.3. Sensor selection through feature space

In the previous section, we have dealt with the case in which multiple sensor signals have no mutual interactions. However, it is often the case that the signals make sense only if they are given at the same time ($p_{ij}(q_i, q_j \mid \xi) \neq p_i(q_i \mid \xi)p_j(q_j \mid \xi)$). Consider a simple case that the recognition problem has an exclusive-OR structure, that is,

$$p_{12}(1,1 \mid 1) = p_{12}(2,2 \mid 1) = p_{12}(1,2 \mid 2) = p_{12}(2,1 \mid 2) = \frac{1}{2}$$

and

$$p_{12}(1,1 \mid 2) = p_{12}(2,2 \mid 2) = p_{12}(1,2 \mid 1) = p_{12}(2,1 \mid 1) = 0$$

hold (where $\Xi = \{1,2\}$, $Q_1 = Q_2 = \{1,2\}$, and $N = 2$). In this case, the system obtains no information if it knows a value of either q_1 or q_2: the probability distribution of the model space does not change when the system updates the distribution because $p_i(q_i \mid \xi) = 1/2$ for all i and ξ. In order to treat such a case, the author considers a product space of the interacting signal spaces and calls it an 'AND feature space' or a 'conjunctive feature space'.

When two sensor signals are strongly correlated to each other, on the other hand, the recognition center can get enough information even if it selects only one of them. It is also useful to prepare a feature space which unifies such correlated signals. The author calls this unification space an 'OR feature space' or a 'disjunctive feature space'. Thus, the system constructs an AND feature space for interacting signals and an OR feature space for correlated signals. The former feature is determined when all of the related signals are given, and the latter feature is determined when one of the signals is given (see Fig. 4).

Let $F^{(k)}_{i_1 i_2 \ldots i_{n(k)}}$ and c_k denote the feature space generated from sensory spaces Q_{i_1}, $Q_{i_2}, \ldots, Q_{i_{n(k)}}$ and the required observation times to get the feature, respectively. Following the idea described in the previous section, the system selects a feature or a sensor which is expected to bring much information. In this case, the sensor selection algorithm is inevitably complicated because the system must take account of the observation cost c_k and because one sensor (or observation) is related to multiple features. Anyway, the system needs to estimate the expected entropy reduction after several observations,

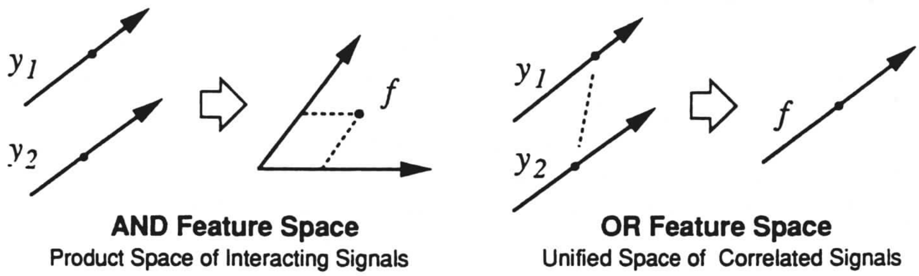

Figure 4. Feature space. The system constructs an AND feature space for interacting signals and an OR feature space for correlated signals.

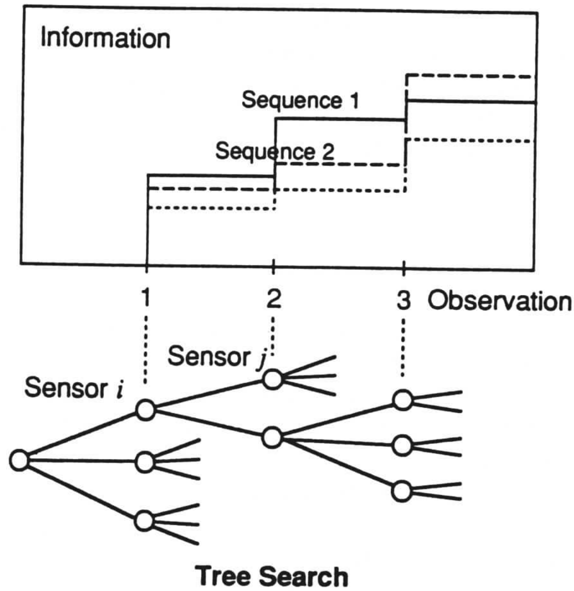

Figure 5. Selection of observation sequence. When selecting observation sequences, the system estimates the expected entropy reduction after every observation sequence, assigns the value to the corresponding node of the tree, and selects the best sequence using a tree-searching method.

that is, the system selects not a single observation but an observation sequence consisting of multiple observations. Here, a tree-searching method will be useful (Fig. 5). The system estimates the entropy reduction for every observation sequence, assigns the values to a node of the tree and then selects the best sequence using a tree-searching method.

Since the required calculation increases rapidly as the searching depth increases, it is not hopeful to search to too much depth. (It would be nonsense if the cost of sensor selection became higher than that of observation!) Actually, the system should determine the appropriate depth according to a trade-off between the calculation cost and the observation cost.

3.4. Attribute recognition

We often want to know only the object's attributes, such as texture and shape, instead of the object itself. In such a case, by regarding 'attribute' space as the model space, the system can operate in just the same way as in the previous sections.

In this situation, the system has a sort of semantic network and assigns attribute spaces to the nodes of the network. The system can update the probability distributions using Bayes' rule so long as the network does not have recurrent structure.

4. HAPTIC SENSING SYSTEM WITH ACTIVE PERCEPTION

When perceiving objects by touch, we human beings do not simply detect the deformation of the skin, but utilize information from various receptors. Moreover, we take various observation behaviors according to what property we want to know.

In light of these characteristics of the human haptic perception, the author has constructed some sensing systems which perceive objects utilizing various sensors and actuators. Here, the authors shows two systems: one treats mainly object material and the other treats mainly surface feeling.

4.1. System 1

4.1.1. Structure. The first system is to discriminate object materials by utilizing four sensor devices [8, 11].

Figure 6 shows the schematic structure of the system and available sensor devices. It is a robotic arm equipped with a thermo-sensor, a pressure sensor, a vibration sensor, and a piezoelectric sensor. Since the sensors are not built into one unit, the author collects experimental data with each sensor separately and afterward performs recognition experiments on a computer.

The thermo-sensor is constructed by burying three thermistors into artificial skin made of silicon rubber, whose temperature is kept constant by a heat controller. It senses the transient change of its temperature when it touches the object.

The pressure sensor is to measure the object's deformation when imposing a certain strength of pressure on the object. Operating this sensor corresponds to the situation in which human beings 'push' an object with the hands to know its elasticity. The piezoelectric sensor measures the frequency shift when it touches the object: it has an oscillator circuit whose oscillation frequency is determined by the resonance of the piezoelectric device and changes according to the object materials.

The vibration sensor is made by covering a small microphone with silicon rubber. This sensor measures the vibration of the rubber while the system rubs the object with this sensor and gets information about surface roughness.

4.1.2. Experimental results. The signals obtained by the sensors are pre-processed and are quantized in a proper manner. After calculating the conditional probabilities, the system discriminates object materials with the intentional observation algorithm. In the experiments, the author dealt with six kinds of materials shown in Table 1. For each material, three different sizes of samples were prepared.

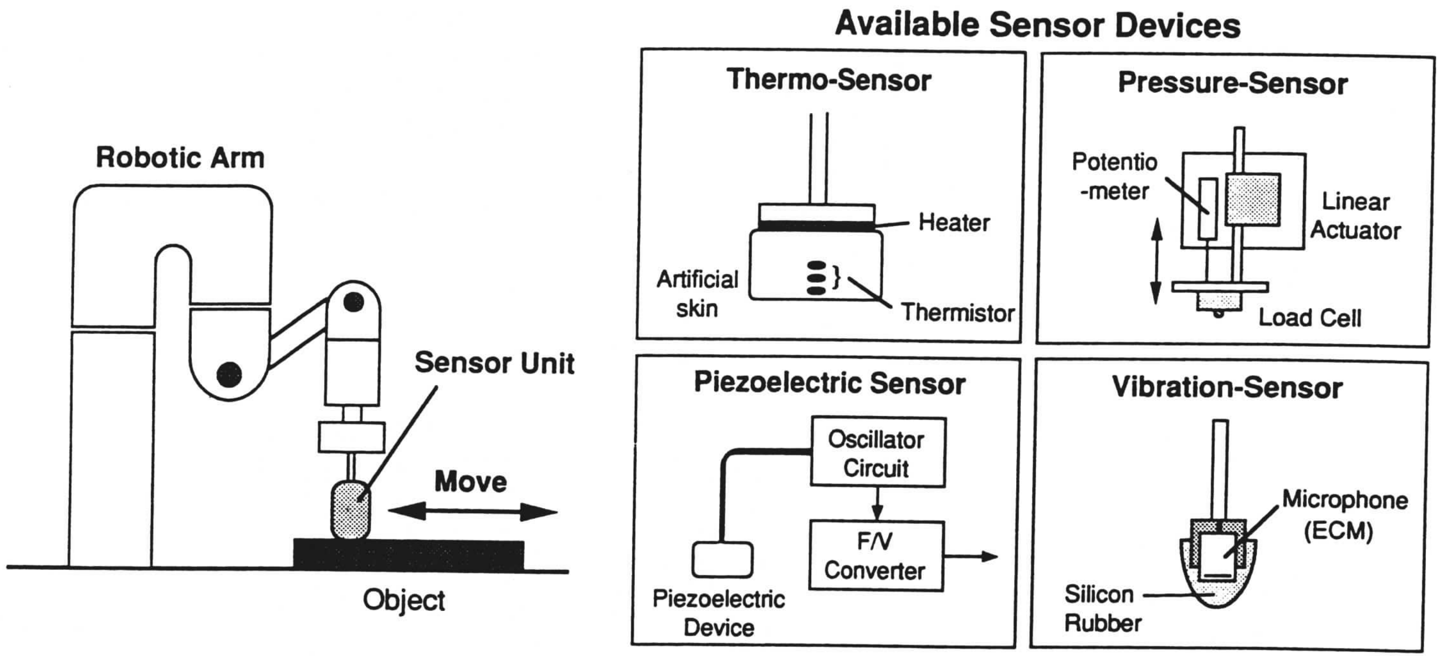

Figure 6. Schematic structure of system 1. System 1 is a robotic arm equipped with four sensor devices, that is, a thermo-sensor, a pressure sensor, a vibration sensor, and a piezoelectric sensor. The system selects one sensor at a time and touches the object with the selected sensor. The sensor data integration is performed on a computer.

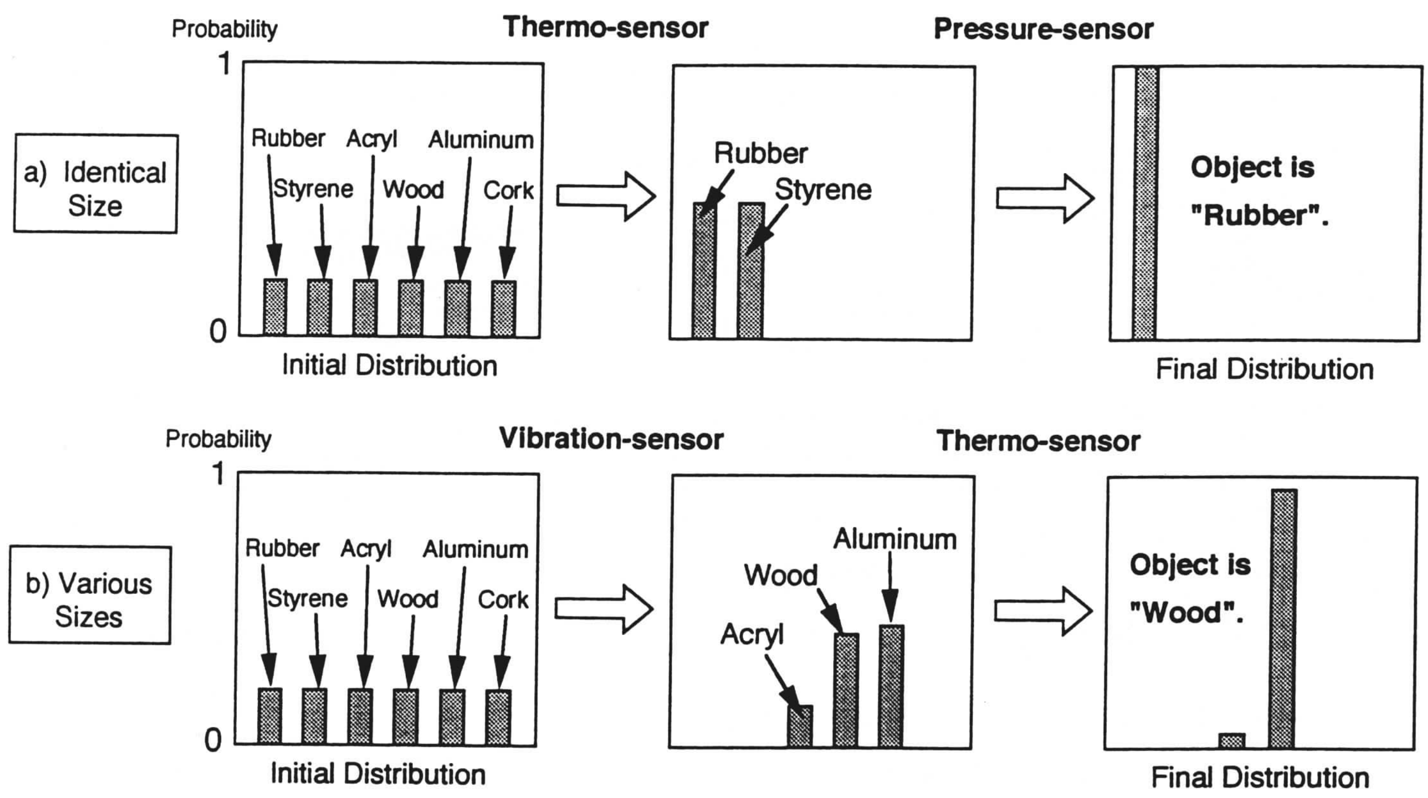

Figure 7. Examples of observation sequences: (a) and (b) show the results for the case where the sizes of the objects are identical and for the case where their sizes differ, respectively.

Table 1.
Objects used for experiments on system 1

Object materials	Object sizes
Aluminium	5.0 cm×5.0 cm
Wood	3.0 cm×3.0 cm
Cork	1.0 cm×1.0 cm
Acrylic resin	
Polystyrene	
Rubber	

Six kinds of materials were dealt with in the experiments. For each material, three sizes of sample objects were prepared.

Two examples of sensor selection sequences are shown in Fig. 7.

Figure 7(a) is a result for the case where the sample objects are identical in size. In this case, the thermo-sensor seems to be most useful at the initial state. As a result of observing the object with the thermo-sensor, the candidates were restricted to 'rubber' and 'polystyrene'. Then using the pressure sensor, the system judged that the object was 'rubber'.

Figure 7(b) is a result for the case where the objects are of various sizes. In this case, the thermo-sensor is not useful because the amount of thermal diffusion is different according to object sizes even if object materials are the same. Instead, the vibration sensor seems to be the most useful. Using the vibration sensor and the thermo-sensor iteratively, the system recognized that the object was 'wood'.

Figure 8 shows recognition accuracy and observation times for the cases of intentional observation and of random observation where the system selects the sensors at random (however, one sensor is selected no more than once). This results shows that recognition accuracy is improved and observation times are reduced by the intentional observation.

When the system takes the random observation strategy, the accuracy is not high in spite of using almost all the sensors. This is because past experience has an influence

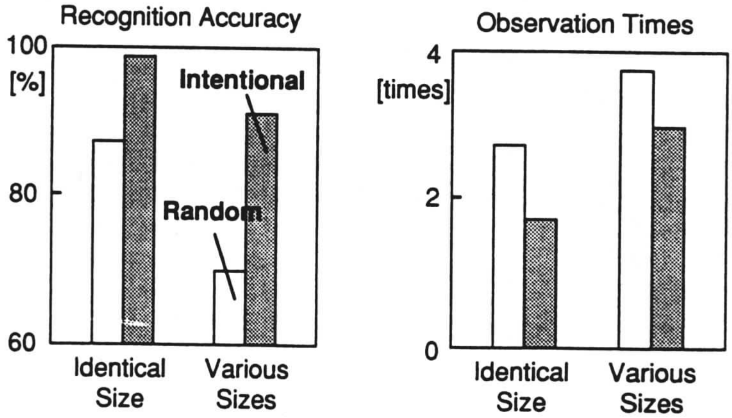

Figure 8. Recognition performance of system 1. The figure describes recognition accuracy and observation times for the cases of intentional observation and of random observation. It shows that recognition accuracy is improved and observation times are reduced by the intentional observation algorithm.

on future judgement in an iterative estimation process: if the system happens to use a sensor which gets unreliable information and updates the internal state based on the information, the system may form a wrong judgement. Anyway, this result is interesting because it illustrates that wrong preconception leads to bad results.

4.2. System 2

4.2.1. Structure. The second system deals with the texture and friction feeling of an object surface [12].

Figure 9 shows the structure of the system. The system consists of two parts: one is a 'sensor head', which rubs and observes an object surface, and the other is a 'stage' which moves up and down with an object on it. When the system observes an object, the stage is lifted up until the object surface comes in touch with the tip of the sensor head. Then, the sensor head moves horizontally and rubs the object surface.

The structure of the sensor head is illustrated in Fig. 10. The tip is equipped with a small microphone covered by silicon rubber. It is almost the same as the vibration-sensor used in system 1: the microphone catches the vibration of silicon rubber while the sensor head moves along the object. The speed of the movement is controlled by a computer. In addition, the head is connected to the mount by parallel links and leaf springs: it rotates around the mount but suffers a force from the springs which brings it back to the center position. Measuring the rotation angle, the system knows the total amount of friction between the sensor tip and the object surface.

On the other hand, the stage is equipped with a load cell to measure the force from the sensor tip upon the object surface. Utilizing this cell, the system adjusts the touching force to the object. Changing the two parameters of the moving speed and the touching force, the system 'touches' the object surface in various manners.

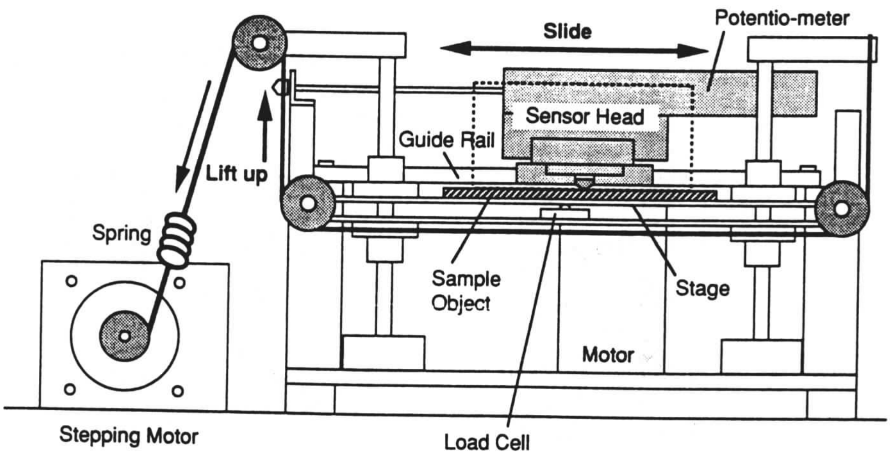

Figure 9. Structure of system 2. The system consists of a 'sensor head' and a 'stage'. When the system observes an object, the stage is lifted up until the object surface touches the tip of the sensor head. Then, the sensor head moves horizontally and observes the object surface.

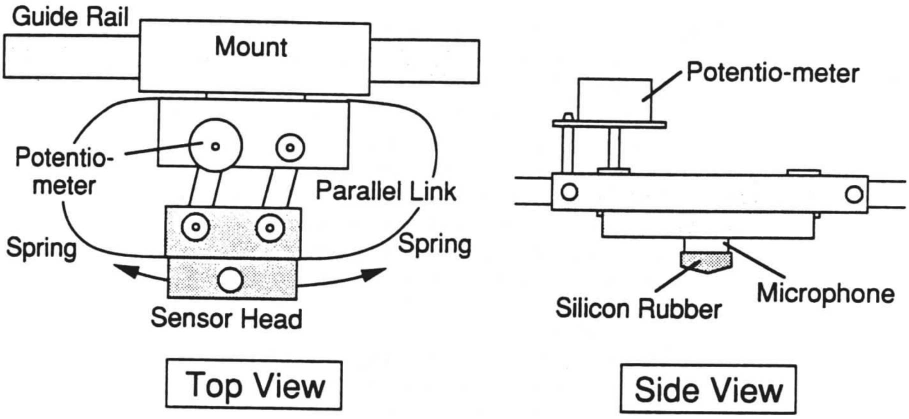

Figure 10. Structure of sensor head. The sensor head is equipped with an small microphone covered by silicon rubber, which measures vibration of the rubber while it moves along the object surface. Also, the head is connected to the mount by parallel links and leaf springs. Using this mechanism, the system knows the total amount of friction between the sensor tip and the object surface.

In the experiments, the author chose three forces and three speeds, and measured the rotation angle and the signal from the microphone for every pair of a pressure and a speed. These signals were processed and transformed to some object characteristics.

4.2.2. Experimental results. The author prepared eight basic objects which gave us apparently distinctive haptic feeling (see Table 2), and chose six characteristics which showed large difference for the objects. Every object was represented as a set of the six characteristics in the system.

The selected characteristics are shown in Table 3. Maximum amplitude of the microphone output contains information on surface ruggedness. The number of zero crossings

Table 2.
Objects used for experiments on system 2

Basic objects	Test objects
(1) Rubber 1	(9) Rubber 3
(2) Rubber 2	(10) Rubber 4
(3) Suede	(11) Wood 2
(4) Chrome leather	(12) Sand paper 2
(5) Ceramic tile	(13) Paper 1
(6) Wood 1	(14) Paper 2
(7) Cork	(15) Cloth 1
(8) Sand paper 1	(16) Felt
	(17) Velour

Eight objects were prepared as basic objects which gave us distinctive haptic feeling. Besides, nine test objects were provided for examining the relation between the system's feeling and human feeling.

Table 3.
Characteristics selected for object representation

	Characteristics
a	Maximum amplitude of microphone output (rubbing speed 48 mm/s, touching force 0.29 N)
b	number of zero crossing of microphone output (rubbing speed 48 mm/s, touching force 0.29 N)
c	power distribution ratio of microphone output (rubbing speed 48 mm/s, touching force 0.29 N)
d	maximum rotation angle of sensor head (rubbing speed 48 mm/s, touching force 0.29 N)
e	maximum rotation angle of sensor head (rubbing speed 27 mm/s, touching force 0.39 N)
f	maximum rotation angle of sensor head (rubbing speed 48 mm/s, touching force 0.15 N)

The author chose six characteristics which showed large difference, for the basic objects in Table 2.

of the microphone output can be a rough estimate of the vibration frequency of the silicon rubber. The power distribution ratio means the power ratio of high frequency component to all the components, which is calculated from a power spectrum of the microphone output. These two characteristics represent surface roughness. Finally, the rotation angle of the sensor head reflects the total friction between the sensor tip and the object surface.

First, the author shows how the basic objects are represented in the system. Figure 11

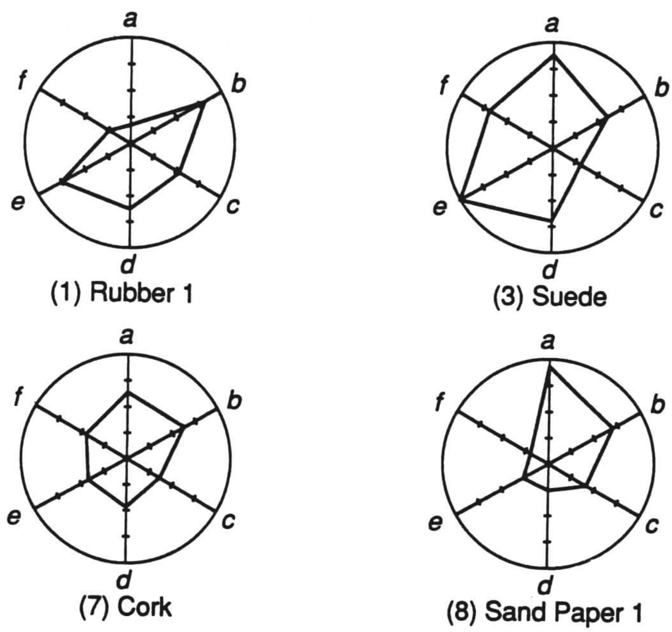

Figure 11. Representation of haptic feeling of basic objects. The system's representations of four basic objects are indicated with radar charts. The details are explained in the text.

describes the results for four basic objects with radar charts whose axes correspond to the six characteristics. As shown in the figure, each basic object has a specific representation. The result suggests the following conclusions:

(1) Objects which feel rough such as 'suede' (3) and 'sand paper' (8) mark large values on axis a.

(2) Though both 'suede' (3) and 'cork' (7) feel rough, the former feels sticky and the latter does not. This difference appears on axis e.

(3) 'Plain rubber' (1) feels smooth when we rub it with a slight pressure, but feels sticky when we rub it with a certain amount of pressure. This feature appears in axes a, d, and e.

Therefore, the system catches some features of haptic perception.

The author applied the intentional observation algorithm also to the second system. As a result, the algorithm proved to be effective in the system. The details of the results are omitted.

4.2.3. Comparison between system's feeling and human feeling. Another experiment was performed to investigate the relation between representation in the system and in human perception. The author prepared nine test objects in addition to the eight basic objects (Table 2) and, for each of them, asked the system and three human subjects which basic object felt the most similar to the given object. The system's answer was determined based on the normalized distance in the six-dimensional characteristic space.

The results are shown in Fig. 12. The figure describes the percentages that the subjects' answers agreed to what the system judged the most similar, the second most similar, and the third most similar. The upper and lower bars show the results when the subjects were taught nothing particular and after they were taught the system's judgement, respectively. It should be noted that the answers of the subjects did not always agree with one another.

The experimental results are well illustrated in the following typical example. When 'felt' (16) was given as an object, the system judged that 'wood 1' (6) was the closest. On the other hand, all subjects answered that 'chrome leather' (4) was the most

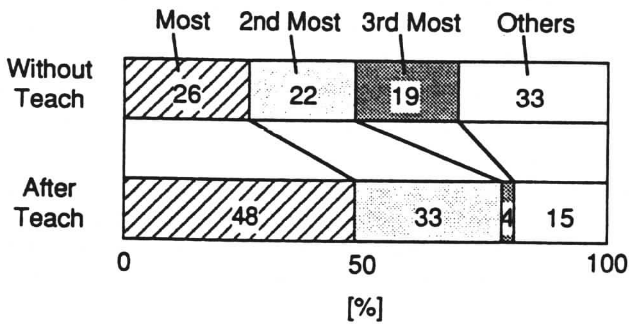

Figure 12. Comparison between system's feeling and human feeling. The figure indicates how much the system's judgement agreed with the subjects' answers: it shows the percentages that their answers agreed to what the system judged the most similar, the second most similar, and the third most similar. The upper and lower bars show the results when the subjects were taught nothing and after they were taught the system's judgement, respectively.

similar. When the experimenter taught them the system's judgement, they changed their judgements to the same as the system's one, and said "If I concentrate my attention upon surface texture, (6) is the closest. But it did not feel similar because it is different in the thermal feelings". This shows that the subjects perceived the objects based not only on the surface texture but also on thermal sensation: their answers agreed well to the system's answers when they judged based only on surface texture.

As mentioned in Section 1, human haptics is based on various sensory information: several kinds of mechanical receptors and thermo-receptors are buried in the skin and supply information about skin vibration, skin deformation, and thermal diffusion when we touch an object [17, 18]. Moreover, muscle receptors detect the force between the body and the object. Such characteristics of human haptic processes have been pointed out in the field of psychophysics [1, 2, 19].

The above results suggests that the system realizes well the human perception of surface roughness and friction, and that thermal information is no less significant in human haptic perception than tactual information. The author would like to improve the system by adding other appropriate sensors and to make a more faithful representation of human haptic feeling.

5. DISCUSSION

In this section, the author would like to discuss some problems related to sensory integration systems.

5.1. Learning probability distribution

The system needs to know conditional probabilities $p_i(q_i \mid \xi)$ in order to use the proposed algorithm. The author has discussed only the recognition procedure on the assumption that they had already been obtained. However, if the system does not know them in advance or if they change in time, the system must estimate them together with recognizing objects. In this case, the system sacrifices observation efficiency and inserts 'testing' observation to identify the distributions.

It would be useful if the system could accomplish these two contradicting goals of efficient observation and efficient learning at the same time. Unfortunately, this problem is one of the difficult problems known as 'dual control' or 'TAB (two-armed bandit) problem' [13, 14]. It is realistic to treat them as separate problems, that is, to pursue only observation efficiency on the assumption that the learning has already been finished.

To construct internal representation is a more significant problem. It is regarded as a process of selecting an appropriate model to represent the external world structure. Considering that 'model selection' is one of the main interests in the field of statistics, fruitful discussion is expected by treating this problem from a viewpoint of statistics.

5.2. Attention to unexpected signals

The author has treated the sensor selection from a viewpoint of 'top-down', that is, based on the system's intention to discriminate objects. Here, the author considers the case where the sensor catches 'attention' of the system in a 'bottom-up' manner.

Assume that each local sensor always receives signals from the object and the recognition center fetches its information when the system selects it. As mentioned in Section 3.1, the system can calculate the probability $P(q_i)$ ($= \sum_\xi \pi(\xi)p_i(q_i \mid \xi)$) which predicts what signal tends to be detected by the i-th sensor. Now, suppose that the i-th sensor detects a signal whose predicted probability is extremely small (that is, $P(q_i) \approx 0$). This situation suggests that the probability distribution of the model space may be wrong. In other words, when the 'detected signal' differs far from the 'expected signal', the system needs to pay 'attention' to the signal and to examine it by observing the object again with the same sensor or with other sensors whose signals are correlated to the signal concerned. Getting the same signal by the second observation, the recognition center should change its internal state. If the signals obtained by two observations are contradictory to each other, it is plausible that the sensor is out of order.

To realize this, every local sensor calculates the distribution, compares it to the detected signals and makes an alert when the corresponding probability is less than a certain threshold. Since the calculation and comparison can be performed independently at each sensor, the task of the recognition center will not increase. By developing this idea, a system can be made which proceeds its task without observation so long as the local sensors do not make any alert.

5.3. *Learning quantization at local sensors*

When treating signals as discrete values, the system needs to quantize sensory signals into a certain number of discrete values. We usually use such quantization algorithms as preserve as much information in the signal as possible, or as maximize the entropy of the quantized signal space [15]. It is sure that this strategy is good in the sense that the system represents the input signal faithfully. In recognition systems, however, it is also helpful that the quantization mechanism should reflect the recognition task.

Such a 'task-dependent' learning quantization is realized by a simple mechanism. For instance, using the hill-climbing method, the system can change quantization mechanism so as to increase the mutual information between the model space and the quantized signal space. Another skilled algorithm using Bhattacharyya distance (one of the distances between two probability density functions) is proposed by Longo *et al.* [16].

If the signal space is divided into too many regions, a vast amount of memory and testing observation will be required. One method to avoid this is to set the initial quantization rough and to make it precise gradually according to demand. As a result, the signals contributing much to recognition are precisely quantized and, conversely, those contributing little are roughly quantized.

6. CONCLUSION

The author formalized the sensory integration and active perception process from a viewpoint of information theory and constructed an algorithm for intentional observation as an iterative experimental design. The algorithm was applied to actual sensing systems which recognized object materials and surface textures in a similar manner to human haptic perception processes, and realized higher recognition accuracy by fewer observations than the random observation algorithm.

The active perception process is found not only in haptic perception but also in visual perception and in auditory perception. Its validity appears rather remarkably in these processes because the system inevitably deals with a large amount of information in visual and auditory processing. Moreover, it plays an essential role in the processes of the 'cocktail party effect' and 'selective attention' in human brains. The author would like to develop the theory and to clarify the mechanism of these interesting functions.

Acknowledgement

This research is partly supported by special coordination funds of the Science and Technology Agency of the Japanese government. The author would like to express his great thanks to Prof. K. Nakano for his helpful advice on this research and to Mr. T. Kurokawa and Mr. K. Sumitomo for their cooperation in constructing the sensing systems.

REFERENCES

1. J. J. Gibson, *The Senses Considered as Perceptual Systems*, Boston: Houghton Mifflin, 1966.
2. J. J. Gibson, "Observation on active touch," *Psychological Review*, vol. 69, pp. 477–490, 1962.
3. R. Luo and M. Kay, "Multisensor integration and fusion in intelligent systems," *IEEE Trans. Systems, Man and Cybernetics*, vol. SMC-19, pp. 901–931, 1990.
4. H. Yamasaki and M. Ishikawa, eds, *Sensor Fusion*, Tokyo: Corona Publishing, 1992 (in Japanese).
5. Y. Sakaguchi, "Information representation in neural networks," *Computrol*, vol. 24, pp. 70–81, 1988 (in Japanese).
6. Y. Sakaguchi, "Topographic organization of nerve field with teacher signal," *Neural Networks*, vol. 3, pp. 411–421, 1990.
7. Y. Sakaguchi and K. Nakano, "Figure recognition with intentional observation," in *Proc. IJCNN-92-Baltimore*, pp. III–792–797, Baltimore, MD, 1992.
8. Y. Sakaguchi and K. Nakano, "Active perception with intentional observation," in *Proc. ISMCR-92*, pp. 241–248, Tsukuba, Japan, 1992.
9. S. Hutchinson and A. Kak, "Planning sensing strategies in a robot works cell with multi-sensor capabilities," *IEEE Trans. Robotics and Automation*, vol. RA-3, pp. 765–783, 1989.
10. R. Takiyama, "Evaluation of pattern recognition systems by the concept of distance," *Tech. Rep. Kyushu Univ.*, vol. 37, pp. 318–324, 1964 (in Japanese).
11. T. Kurokawa, Y. Sakaguchi and K. Nakano, "Intentional haptic sensing system," in *Prod. 31th SICE Annual Conf.*, pp. 671–672, Kumamoto, Japan, 1992 (in Japanese).
12. K. Sumitomo, Y. Sakaguchi and K. Nakano, "A system simuilating human haptic perception," in *Proc. 32th SICE Annual Conf.*, pp. 33–34, Kanazawa, Japan, 1993 (in Japanese).
13. A. Fel'dbaum, "Dual-control theory I–IV," *Automation and Remote Control*, vol. 21, pp. 874–880, 1033–1039; vol. 22, pp. 1–12, 109–121, 1962.
14. I. Witten, "The apperent conflict between estimation and control — a survey of the two-armed bandit problem," *J. Franklin Institute*, vol. 301, pp. 161–189, 1976.
15. D. Meeserschmitt, "Quantizing for maximum output entropy," *IEEE Trans. Information Theory*, vol. IT-17, pp. 612, 1971.
16. M. Longo, T. Lookabaugh and R. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Information Theory*, vol. IT-36, pp. 241–255, 1990.

17. Y. Iwamura, "Structure and function of human tactile receptors," *J. Robotics Soc. Jpn*, vol. 2, no. 5, pp. 54–60, 1984 (in Japanese).
18. R. F. Schmidt, "Somatovisceral sensibility", in R. F. Schmidt, ed., *Fundamentals of Sensory Physiology*, pp. 81–125, New York: Springer-Verlag, 1981.
19. R. L. Kratzky, S. J. Lederman and V. A. Metzger, "Identifying objects by touch: An 'expert system'," *Perception and Psychophysics*, vol. 37, no. 4, pp. 299–302, 1985.

ABOUT THE AUTHOR

Yutaka Sakaguchi was born in 1963. He received the B.E., M.E., and Ph.O. degrees from University of Tokyo in 1986, 1988, and 1994, respectively. Since 1988, he has been with the Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo, where he is currently a Research Associate. His research interests are in human sensory information processing, especially in sensory integration, active perception, attentional mechanisms, and information representation. He is a member of the Institute of Electronics, Information and Communication Engineers of Japan, the Society of Instrument and Control Engineers, International Neural Network Society, Japanese Neural Network Society, and IEEE.