

Super-Flexible Skin Sensors Embedded on the Whole Body, Self-Organizing Based on Haptic Interactions

Tomoyuki Noda*, Takahiro Miyashita†, Hiroshi Ishiguro*†‡, and Norihiro Hagita†

* AMS, Graduate school of Engineering, Osaka Univ., 2-1 Yamada-oka, Suita, Osaka, 565-0871, Japan.

Email: noda@ed.ams.eng.osaka-u.ac.jp

†ATR Intelligent Robotics and Communication Laboratories, Kyoto, 619-0288, Japan

‡Asada Synergistic Intelligence Project, ERATO, JST, Osaka Univ. Osaka, 565-0871, Japan

Abstract—As robots become more ubiquitous in our daily lives, humans and robots are working in ever-closer physical proximity to each other. These close physical distances change the nature of human robot interaction considerably. First, it becomes more important to consider safety, in case robots accidentally touch (or hit) the humans. Second, touch (or *haptic*) feedback from humans can be a useful additional channel for communication, and is a particularly natural one for humans to utilize. Covering the whole robot body with malleable tactile sensors can help to address the safety issues while providing a new communication interface. First, soft, compliant surfaces are less dangerous in the event of accidental human contact. Second, flexible sensors are capable of distinguishing many different types of touch (e.g., hard v.s. gentle stroking). Since soft skin on a robot tends to invite humans to engage in even more touch interactions, it is doubly important that the robot can process haptic feedback from humans. In this paper, we discuss attempts to solve some of the difficult new technical and information processing challenges presented by flexible touch sensitive skin. Our approach is based on a method for sensors to self-organize into sensor banks for classification of touch interactions. This is useful for distributed processing and helps to reduce the maintenance problems of manually configuring large numbers of sensors. We found that using sparse sensor banks containing as little as 15% of the full sensor set it is possible to classify interaction scenarios with accuracy up to 80% in a 15-way forced choice task. Visualization of the learned subspaces shows that, for many categories of touch interactions, the learned sensor banks are composed mainly of physically local sensor groups. These results are promising and suggest that our proposed method can be effectively used for automatic analysis of touch behaviors in more complex tasks.

I. INTRODUCTION

Robots are becoming more ubiquitous in our daily lives [1][2][3][4], and humans and robots are working in ever-closer physical proximity to each other. Due to this proximity, there is increased potential for robots to inadvertently harm users. Physical nearness also increases the need for robots to be able to interpret the meaning of touch (or *haptic*) feedback from humans. Covering the whole robot body with malleable tactile sensors allows us to address both of these concerns. First, soft, compliant surfaces are less dangerous in the event of accidental human contact[5]. Second, flexible sensors are capable of distinguishing many different types of touch (e.g., hard v.s. gentle stroking). This is important, as soft skin actually invites

more natural types of touch interaction from humans, so it is critical that the soft surfaces of robots be touch sensitive.

To extract information about humans' physical contact with robots, the distribution density of tactile sensor elements, sampling rate, and resolution of kinesthetic sense all must be high [3], resulting in a high volume of tactile information that must be processed. To do so, the following three problems must be solved. First, there is the problem of reduced system robustness due to an increased number of possible failing components. The second is the high cost of data processing. The third is the administration of the sensors' configuration.

The previous study of [6] proposed highly dense distributed skin sensor processing based on interconnecting a self-organized sensor network. Spatiotemporal calculation in each node with spatially seamless tactile information gathered from adjacent nodes enabled haptic interaction features to be extracted, solving the first and second challenges. For instance, an edge detection method is applied to extract features of haptic interaction within the local sensor, yielding an efficient data compression. This type of distributed processing requires that the configuration of tactile sensor position is described in the distributed programs of network nodes, which is the remaining third challenge. Manually describing 3-dimensional tactile sensor positions, changing with robot's postures, is very labor intensive and error prone. Moreover, distributed processing typically requires predefined sensor banks, defining which tactile sensors are used in distributed processing for each network node, also a labor intensive task.

In [7], we found that interaction scenarios could be successfully classified simple k-nearest neighbors (KNN) using a novel feature space based on cross-correlation between tactile sensors, achieving performance of 60% in a 13-way forced choice task. We also found that many categories of touch interactions can be easily visualized by arranging sensors into a "Somatosensory Map" using MultiDimensional Scaling (MDS)[8] applied to this feature space as a similarity measure. These promising results suggest that this feature space can be effectively used for automatic analysis of touch behaviors in more complex tasks.

In this paper, we propose a method for learning "self-

organizing tactile sensors” using the feature space from [7] to solve the remaining third challenge. In the proposed method, a classifier is constructed using CLAss-Featuring Information Compression (CLAFIC) [9], a type of a subspace method, applied to a data set consisting of the full cross-correlation based feature space of [7]. Instead of directly using a learned subspace as the input for a classifier, we select the sensor pairs that are most “useful”, i.e., have the highest relevance for the classifier output, to form a more compact sensor bank to be used as input to the classifier. Since now different sensor nodes are involved in different classifiers, it is possible to distribute processing around the body, which can be implemented as in-network processing on the self-organizing sensor network [6]. We call each learned sensor bank a “self-organizing tactile sensor”. We found that classifiers based on self-organizing tactile sensors could classify interaction scenarios with an accuracy of up to 80% in a 15-way forced choice task, a significant improvement over prior work. The learned subspaces can also be visualized in a “Somatosensory Map”, showing sensor point distribution in a 2D plane.

The rest of this paper is organized as follows. In Section we II describe related work, and contrast our study with others. Section III describes the basic idea, and then details the proposed method for self-organizing tactile sensors. Section IV describes experiments dealing with human-robot haptic interaction used to construct a haptic interaction database. In section V, using the database, the performance of the proposed method is shown. Also, the learned subspaces are visualized in the Somatosensory 2D Map. Section VI discusses the results, and Section VII concludes.

II. BACKGROUND AND COMPARISONS

Prior work on studies of robots with tactile sensors has tended to focus on the development of the physical sensors and transmitting sensor data. For instance, [10] proposes a Large-Scale Integration (LSI) technique for processing data from tactile sensors. Iwata et al.[4] demonstrated physical interaction with users via a skin equipped with 6-axis-kinesthetic sensors. Pan et al. [11] and Inaba et al.[2] described tactile sensors using electrically conductive fabric and strings as a whole-body distributed tactile sensor for humanoid robots. And Shinoda et al.[12] proposed a wireless system for transmitting tactile information by burying wireless sensors under the robot “skin.” Thus so far, this research is mainly limited to the problems of collecting tactile information, and solving the necessary wiring and physical implementation problems.

Compared to other sensory modalities such as vision and audio, relatively little prior work has been done on processing haptic interaction from incoming tactile sensor signals. Miyashita et al. [1] estimated user position and posture in interaction using whole-body distributed tactile sensors. Naya et al. [3] classified haptic user interaction based on output from tactile sensors covering a robot pet. Francois et al. [13] also classify different two touch styles, namely “strong” and “gentle”. Though the above research classifies human-robot interaction using tactile sensors, these and other prior studies

have not to our knowledge been successful in classifying several haptic interactions while robots are interacting with users.

Pierce and Kuipers [14] proposed self-organizing techniques for building a “cognitive map”, which represents knowledge of the body corresponding to physical position of sensors. This map shows the position of each sensor installed on the surface of a robot. However, this method will not work out for a robot having high degree of freedom and soft skin because the positions of the tactile sensors in 3-d “world” coordinates dynamically change during an interaction. Rather than construct spatial maps to acquire physical sensor positions our objective is to use, interpret and visualize underlying haptic interaction features.

Kuniyoshi et al. [15] proposed a method for learning a “Somatosensory Map,” showing the topographic relationship of correlations between incoming signals from tactile sensors distributed on the whole body surface of a simulated baby. In their somatosensory map highly correlated sensor points are plotted on a 2D plane close to each other. As the result, the map showed the structure of robot body parts rather than the physical sensor positions as in Pierce and Kuipers[14].

In this paper, our goal is similar to that of [15], so that highly correlated sensor points will be located close to each other, and thus we keep the name “Somatosensory Map[15].” However, we use a different technique to acquire the map, and use real-world human-robot interaction rather than a simulated baby. Moreover, we attempt to classify haptic interactions using correlations between incoming signals from tactile sensors distributed on the whole body surface.

III. SELF-ORGANIZING TACTILE SENSOR METHOD TO DECIDE SENSOR BOUNDARIES

A. Basic idea

Suppose that N tactile sensors are implemented on a robot, and that i -th tactile sensor stream during one human-robot interaction is called S_i ($i = 1, \dots, N$) where S_i is a vector of the n time-step sampling result of the sensor outputs ($S_i = \{(S_i)_1, \dots, (S_i)_n\}$, $(S_i)_t \in \mathbf{R}$). Features need to be extracted from this time series of the data stream. However, less work has been done on processing tactile features than on audio or vision. In conventional works[3][13], since the data are high dimensional, summary statistics, such as mean, standard error, minimum, max, and coefficients of fast Fourier transformation, are computed from one sensor or all of the sensors to be used as features. A feature space defined from one sensor will not be enough when several sensor are activated by touches, e.g., distinguishing a finger tap from a hand tap or a tickle. On the other hand, the feature space computed from combining all the sensors is less robust, since the features could be drastically changed if e.g. one sensor is broken. Feature space defined from several sensors, at least from two sensors, could be an effective happy medium.

We proposed a feature space using cross-correlations computed from sensor pairs, satisfying the above condition, in [7]. The cross-correlation is one important statistics in human

tactile system – Dince et al.[16] reported that discrimination ability of the two point stimulus is improved when correlated stimulus is added continually to two close separated point of a human finger. In fact, the visualization results of our feature space, i.e., the Somatosensory Map, show characteristics of many categories of touch interactions[7] by arranging sensor point. Fig. 7 is the Somatosensory Map made from a 2 minute interaction between human and a robot. This presentation simplifies the haptic interaction; e.g., distinctive sensor point cluster of both arms’ sensors are to the result of subject touching the robot’s arms at the same time.

The cross-correlation based feature space has another advantage that the choice of sensors to be included in a sensor bank for a certain computation can be decided via the elements used in the classifier’s selected subspace. Often the useful subspace is composed of combination of sensors located in close spatial proximity, since distant sensors usually have low correlated signals and thus have less mutual information than adjacent sensors.

B. Feature space

A feature vector \mathbf{a} is computed from the cross-correlation matrix defined by equation (2) as follows:

$$\mathbf{a} = \left(\underbrace{R_{(1,2)}, \dots, R_{(1,N)}}_{N-1}, \underbrace{R_{(2,3)}, \dots, R_{(2,N)}}_{N-2}, \dots, \underbrace{R_{(N-1,N)}}_1 \right)^t \quad (1)$$

where $R_{(i,j)}$ is cross-correlation matrix element at (i, j) between N sensors, i.e.,

$$R_{ij}(S_i, S_j) = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \quad (-1 \leq R_{ij} \leq 1) \quad (2)$$

where

$$C_{ij} = \sum_{t=1}^n ((S_i)_t - \bar{S}_i) ((S_j)_t - \bar{S}_j) \quad (3)$$

is cross variation of (i, j) , and \bar{S}_i is average of time series data S_i .

C. Overview

We construct a classifier for detecting haptic interaction between robot and human uses the CLAFIC method[9], a type of subspace method. This method represents each class as eigenpairs computed from a training data set. The subspace method starts from an idea of Watanabe et al.[9] that, as the feature space grows, the data set will converge to a limited small subspace. The CLAFIC method approximates this subspace with eigenpairs. Since the feature space defined in III-B is also a high dimensional space of $O(N^2)$, feature vectors of a data set should also be restricted mostly to a limited feature subspace.

In our proposed method, a dimension reduction of the extracted subspace is additionally applied by selecting base vectors having large inner product values. The classifier output is then calculated from a selected subspace consisting of only a few base vectors chosen from cross-correlation elements of the coefficient R_{ij} computed from all sensor pairs. Hence,

this subspace and feature selection results in useful subsets of sensors that can be used to distribute processing. Since these subsets are found automatically, we call this method finding “self-organizing tactile sensors”.

D. CLAFIC method

Fig. 1 shows the overview of classification. At first, a data set \mathbf{X}_k is prepared from p_k feature vectors (\mathbf{a}_k) computed from a time series of sensor streams labeled as a class k ($\stackrel{def}{=} \omega_k$), where p_k is the number of training data set. Thus

$$\mathbf{X}_k = \{\mathbf{a}_{k1}, \dots, \mathbf{a}_{kp_k}\} \quad (4)$$

The approximated subspace of CLAFIC method begins by performing singular value decomposition (SVD):

$$\mathbf{X}_k = \mathbf{U}D_\lambda\mathbf{V}^t \quad (5)$$

where the columns of \mathbf{U} is left singular vectors; D_λ has singular values and is diagonal; and \mathbf{V}^t has rows that are the right singular vectors. We perform a forward feature selection of singular vectors to perform classification using a low dimensional subspace. First, we arrange the vectors obtained from the SVD performed only on data from class k in decreasing order of the $((d_k)_1, (d_k)_2, \dots, (d_k)_{smallest})$ singular values, and compute a discrimination function DF_k of class k , which takes input an unknown vector \mathbf{x} , and computes

$$DF_k(\mathbf{x}) = \sum_{j=1}^m (\mathbf{x}^t \mathbf{u}_{kj})^2 \quad (6)$$

using only the first m vectors, where \mathbf{u}_{kj} is a left singular vector derived only from data in class k . We choose m by starting at $m = 1$ and increasing m until the cumulative contribution ratio in eq. 6 exceeds a threshold value C_1 . The output of the discrimination function corresponds to the square of the length of an unknown vector orthographically projected onto the low dimensional subspace. The classifier outputs the class name which has maximum DF output.

$$\max_{k=1, \dots, c} \{DF_k(\mathbf{x})\} = DF_l(\mathbf{x}) \quad \Rightarrow \quad \mathbf{x} \in \omega_l \quad (7)$$

Using the fidelity value τ as Watanabe et al. proposes[9], the unknown is vector classified as “unknown” class if the maximum DF is not significantly difference from the second maximum DF , i.e., if

$$\frac{DF_l(\mathbf{x})}{\max_{k \neq l} (DF_k(\mathbf{x}))} > 1/\tau \quad (8)$$

evaluates to “false”, the classification is “unknown”.

E. Learning Sensor Banks

A sensor bank for a classification task is decided by selecting a useful subspace that has high relevance for the classifier. As equation (6) shows, the output of the classifier is composed of inner products. Considering the \mathbf{u}_{ij} elements are weights for the unknown vector elements, if a p -th element $\{\mathbf{u}_{ij}\}_p$ is close to 0, the element $\{\mathbf{x}\}_p$ could be ignored. In

computer vision, Ishiguro et al.[17] has proposed this type of idea, describing it as a form of “attention control”.

Let $\tilde{\mathbf{u}}_{ij}$ be an approximated subspace ignoring the all elements close to 0, where all $\{\mathbf{u}_{ij}\}_p$ smaller than C_2 are simply replaced by 0. Now we construct approximated discrimination function $DF_k^*(\mathbf{x})$ as follows.

$$DF_k^*(\mathbf{x}) = \sum_{j=1}^m (\mathbf{x}^t \mathbf{u}_{kj}^*)^2 \quad (9)$$

where \mathbf{u}_{kj}^* are assumed to be nearly orthogonal, and are normalized ($\mathbf{u}_{kj}^* = \tilde{\mathbf{u}}_{kj} / \|\tilde{\mathbf{u}}_{kj}\|$).

The approximated discrimination function $DF_k^*(\mathbf{x})$ is also composed of inner products of \mathbf{u}_{kj}^* and \mathbf{x} , however now we need to know only the q -th elements ($\{\mathbf{x}\}_q$), where

$$\{q : \{\mathbf{u}_{kj}^*\}_q \neq 0\} \quad (10)$$

From the definition of feature space in equation (1), only elements $\mathbf{R}_{(r_q, s_q)}$ need to be computed, where (r_q, s_q) is a sensor pair needed to compute q -th element $\{\mathbf{x}\}_q = \mathbf{R}_{(r_q, s_q)}$. These sensor pairs define whether the sensor is used or not used in the sensor bank, facilitating distributed processing, since each classifier only needs a subset of sensors.

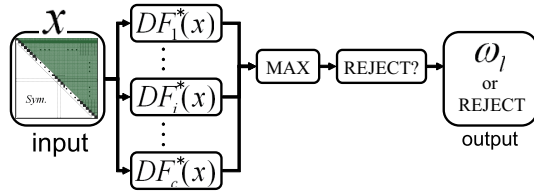


Fig. 1. Overview of the classification

F. Somatosensory Map

To visualize feature vectors and sensor banks, we define dissimilarity as d_{ij} converted from the coefficient of the cross-correlation matrix R_{ij} with the following equation (11),

$$d_{ij}(R_{ij}) = -\log(|R_{ij}|). \quad (0 \leq d_{ij} \leq \infty) \quad (11)$$

This dissimilarity definition defines a “distance” between (i, j) sensors, i.e., higher correlated (or negatively correlated) sensor pairs have smaller dissimilarity. (Note that it does not satisfy all properties of a true distance notion.) Since the self-correlation coefficient is always 1, the dissimilarity with itself is always 0, i.e.,

$$d_{ii} = 0. \quad (S_i \neq \text{constant}) \quad (12)$$

In the Somatosensory Map, the N sensor points are arranged into a 2D map using MDS[8] based on the dissimilarity definition of the equation (11). This 2D map can be used to visualize a vector of the cross-correlation feature space, e.g., fig. 7 is the Somatosensory Map plotted using a feature vector during a human-robot interaction[7]. In section V we apply this method to \mathbf{u}_{ij} to interpret experimental results.

TABLE I

SCENARIO BASED DEFINITION OF THE INTERACTION CLASSES

Class name (step)	Approaching a person	m ($C_1 = 0.15$)
class1	“Hello.”	4
class2	“Let’s shake hands.”	4
class3	“Nice to meet you.”	5
class4	“What’s your name?”	5
class5	“Where are you from?”	5
class6	“let’s play!”	5
class7	“Do you think I’m cute?”	4
class8	“I wish you’d pat me on the head”	3
class9	“Whee!”	6
class10	“I want to play more.”	5
class11	“Tickle me.”	7
class12	“That tickles!”	3
class13	“Thanks”	4
class14	“Give me a hug.”	7
class15	“Bye-bye!”	4

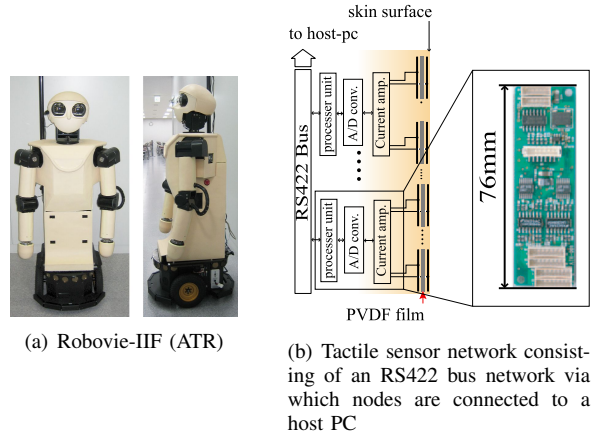


Fig. 2. Overview of our tactile sensor system

IV. EXPERIMENTS

A. Hardware

The hardware on which we are testing our proposed technique is detailed below. Fig. 2 shows an outline of the hardware for the experiments described in this section. Fig. 2 (a) shows the communication robot Robovie-III-F[20], provided with high-density soft tactile sensors and a sensor network consisting of a RS422 bus network via which nodes connected to a host PC (Fig. 2(b)). Fig. 3 shows the structure and materials of the skin sensors installed on the Robovie-III-F surface, and Fig. 4 is the location of embedded piezofilms.

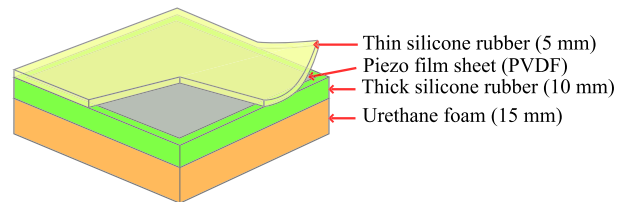


Fig. 3. Architecture of skin sensor devices

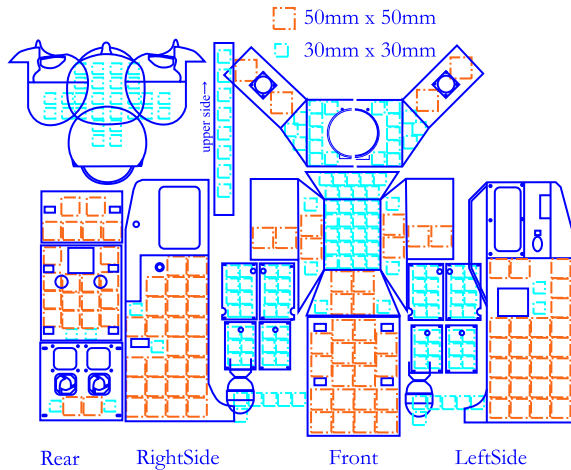


Fig. 4. Position of the skin sensors (PVDF films) on the deployed surface of the Robovie-IIF



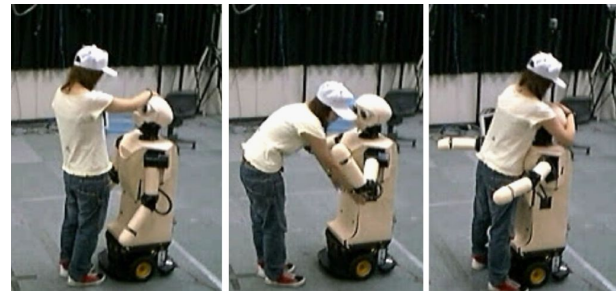
(a) pat the head (b) Hug (c) Touch the robot

Fig. 5. Observed haptic interactions between the robot and the visitors at Osaka Science Museum May. 2005

274 Piezofilms ($3\text{ cm} \times 3\text{ cm}$, or $5\text{ cm} \times 5\text{ cm}$) are embedded in soft silicone rubber. Sampling time for the 16 bit A/D converter is set to 100 Hz, and tactile sensor outputs are read to the host PC of Robovie-IIF at every sampling. Using this hardware, we conducted two experiments detailed below and shown in Fig. 5 and Fig. 6.

B. Field experiment

Our research group conducted several events in the Osaka Science Museum with socially interactive robots[19]. The first experiment is a field experiment during the event named “Let’s play with Robovie”. In the event, the Robovie-IIF was displayed in the Osaka Science Museum during May 2005. We asked visitors to play with the Robovie-IIF and with the goal of investigating what kind of haptic interaction can be realized between humans and the Robovie-IIF. Figs. 5 (a) to (c) show the three haptic interactions such as (a) Patting on the head, (b) Hug and (c) Touch the robot body, observed in this experiment. From these observations we designed a “haptic interaction scenario” for a second experiment to encourage human-robot interaction in which we expected subjects to touch Robovie-IIF, detailed below. Table I shows the interaction scenario stages. Each stage in the scenario consists of the control rule



(a) “I wish you’d pat me on the head.” (b) “Tickle me.” (c) “Give me a hug.”

Fig. 6. Observed subject’s behaviors in each step of the scenario during the experiment for database construction

that the Robovie-IIF tries to sustain interest of a subject to keep interaction going and proceeds to the next stage after finishing each interaction. (See “approaching a person” of Table I.)

C. Construction of the haptic interaction database

In a second experiment, using a “Wizard of OZ” method [18] based on the scenario-based rules described in Section IV-B, Robovie-IIF is controlled by an experimenter with several monitoring displays. We expect the robot to be touched by subjects during these interactions. We constructed a scenario based interaction database which includes the monitoring videos, all of the tactile sensor signals, the command signals sent to control the robot, with all data time stamped using a common clock.

Each subject was asked to interact with Robovie-IIF in 3 trials separated by 10 minutes each. There were a total of 48 subjects, 24 males and 24 females, all of them college students. Each trial took around 5 minutes and was set up with the same condition except for the subject’s position at the start of the scenario. These positions were each 2m away from Robovie-IIF, at 45 degrees to the right, 45 degrees to the left, and 0 degrees (where 0 degree is defined as in front of the robot). We asked the subjects to simply play with Robovie-IIF (which has a child-like voice and uses other cues to encourage humans to treat it as a child), and explained to them before each trial the following rules: (1) The subjects can touch the whole body of Robovie-IIF, (2) The subjects are required to listen carefully to what Robovie-IIF is saying, and (3) The subjects are required to be close to the robot in order to turn on the robot by touching it at the start of each trial.

Fig. 6(a) to (c) show the observed haptic interactions in the experiment, such as (a) “I wish you’d pat me on the head.” of *class8*, (b) “Tickle me.” of *class11* and (c) “Give me a hug.” of *class14*. Excluding approximately 24 cases in which there were technical difficulties, approximately 120 cases of data were acquired to form a “haptic interaction database” of data collected from real interaction scenarios. Segments of the tactile sensor data are automatically clipped and labeled using the time stamps for when each scenario stage (as defined in Table I) begins and ends. Thus, unlike previous work in which interaction segments were hand-labeled by an experimenter,

we do not perform any manual coding of the data.

V. RESULTS

To emphasize the tactile sensor data when subjects touch the robot, we prepared \hat{S}_i from the output of tactile sensors as shown in the equation (13) since our database includes lots of information caused by the robot movements.

$$\hat{S}_i = \left\{ \begin{array}{l} S_i, \left(\left| S_i - \bar{S}_i \right| > \sigma_i \right) \\ \bar{S}_i, \left(\left| S_i - \bar{S}_i \right| \leq \sigma_i \right) \end{array} \right\} \quad (13)$$

where S_i is the i -th sensor output, average of time series data of S_i is \bar{S}_i , and standard deviation of S_i is σ_i . If the absolute difference between S_i and \bar{S}_i is smaller than standard variation σ_i , S_i is replaced by the average \bar{S}_i . In this section, all of the results are from \hat{S}_i .

Fig. 7 shows the 2D Somatosensory Map obtained from a cross-correlation matrix of each tactile sensor during an interaction in a field experiment between the Robovie-III and a subject (Fig. 5 (c)). Figs. 8 and 9 show the results of Leave-One-Out cross validation tests for evaluation of the classifier using the K-nearest neighbor (KNN) method ($k = 3$) and the currently proposed method, respectively, on the whole dataset. Figs. 9 through 12 show the results of choosing different values of C_1 and C_2 using the proposed method. In these figures, *class* k ($k = 1, \dots, 15$) corresponds to the classes defined in Table I. For example, the data set labeled *class2* consists of tactile sensor data collected between the time when the start command of *class2* (“Let’s shake hands”) was sent to the Robovie-III and the time when the start command of *class3* (“Nice to meet you”) was sent. (The data sets include some cases in which subjects did not deliver expected interaction.) As can be seen in Fig 8, the KNN method achieved classification of over 60% for many haptic interactions such as *class2*, *class6*, *class7*, *class8*, *class11*, and *class15*, using only the correlation patterns of all tactile sensors.

Fig. 9 shows the correct recognition rates for the proposed method, while figure Fig. 10 shows the “false alarm” rate, computed for each class as the number of times an example was incorrectly classified as belonging to that class, divided by the number of examples that actually belong to that class (thus these numbers can be greater than 1). For these experiments, the fidelity value τ was experimentally fixed to be 0.95, which did not change the recognition rate but improved the false alarm rate. Each figure has 6 conditions that are in the set $\{(C_1, C_2) : C_1 = 1, 0.15 \ C_2 = 0, 1, 2\}$. The number of orthogonal base vectors (that are left singular vectors, \mathbf{u}_{ij}) is decided by the parameter C_1 , shown in Table. I. The parameter C_2 determines the reduction of size of the feature space, e.g., the reduced feature space in case $C_2 = 0, 1, 2$ were 0%, around 80 – 85%, and around 94 – 96%, respectively.

The proposed method improved classification to 80% for most haptic interactions, including *class1*, *class2*, *class6*, *class7*, *class9*, *class11*, *class14*, and *class15* when using $(C_1, C_2) = (0.15, 0)$. This performance was almost the same as for $(C_1, C_2) = (0.15, 1)$, which used only 15% of the

feature space. When the feature space size is reduced to 5% in the condition $(C_1, C_2) = (0.15, 2)$, performance is still as high as 60% for haptic interactions of *class1*, *class8*, *class11*, and *class15* (note that random, “by chance” performance is less than 7%).

Fig. 11 shows the result of the feature space reduction. Elements of vector \mathbf{u}_{81} are arrayed onto a matrix of the same size as a cross-correlation matrix, and large weighted cross-correlation elements are visualized with darker (more black) colors. Adjacent tactile sensors usually have closer numbers, and are displayed as square line boxes corresponding to their part names. As expected, the useful feature spaces are composed mostly of adjacent sensor pairs. The boxes shown, which include several highly weighted elements, are the result of self-organizing results corresponding to the boundaries of tactile sensors. These self-organizing results are also shown in Fig. 12 (a) and (b), which shows the arrangement of sensors in a 2D Somatosensory Map using $\mathbf{u}_{(8)1}$ and $\mathbf{u}_{(11)1}$. Closer sensor pairs have larger weight in their cross-correlation element. Fig. 12 (a) shows that head sensors, probably touched in the “pat me” interaction, are clustered apart from other sensors. Fig. 12 (b) also clustering of the front side of the body (F-body) and of the left and right side of the body (LSide-body and RSide-body), which are often touched together in the “tickle” interaction. Note that these results do not make use of any knowledge about the spatial position of sensors but only using sensor streams from the whole robot body.

VI. DISCUSSION

Comparing our Somatosensory Map with previous work [14][15], we found that haptic interactions form clusters in the map that often can be grouped by body part. Using this representation for human-robot interaction we achieved good classification results for those interaction categories in which there was some human touching. In previous work for classification of haptic interaction [3][13], the data sets for learning classifiers were hand labeled by the experimenter. In our case, the database is self-labeled during scenario based interactions. The label of each data point is based on what the current designed scenario is, rather than given *post-hoc* by the experimenter asking subjects to touch the robot. We assume this is a more natural and practical database construction.

The Somatosensory Map shows large weighted elements mainly between spatially-localized sensors. This is consistent with the idea of Watanabe et al. that, as the feature space grows, data sets converge to limited subspaces. Additionally, the learned subspace was composed mainly of adjacent sensor pairs in the tactile system, as seen in Fig. 11. Thus, the CLAFIC method is able to achieve higher classification performance even when using smaller subspaces of 15% size. In fact, performance of 80% classification was achieved, improving over the KNN method of 60%, despite a much more challenging evaluation than in previous work. Instead of using static objects consisting of tactile sensors, we constructed the database from real human robot interactions. Since the robot has malleable tactile sensors embedded under soft skin,

and the robot is moving during the experiment, it is possible that the results come from the classification of self-sensations provided by self-movements. However, it seems like more classification is provided from subject's touches, because the recognition rate of the classes that we don't expect to see subject's touches had low recognition rates. Nevertheless, this problem is unavoidable while the robot moving during the touch from other, so in the future we probably will also need to use proprioception in the tactile system.

VII. CONCLUSIONS

In conclusion, the proposed method was found to be efficient with the classification of real human-robot interactions, and was able to be implemented as distributed in-network processing.

In this paper, we describe a haptic interaction classification method using cross-correlation matrix features, and propose a self-organizing technique to define a bank of sensors to be used in distributed processing of each class. The cross validation rests results in recognition of 80% for those interactions in which we expect subjects to touch the robot, using only 15% of the feature subspace. The Somatosensory Map visualization shows that the selected feature space was composed mainly of spatially-adjacent sensor pairs. These promising results suggest that our proposed method may be useful for automatic analysis of touch behaviors in more complex future tasks.

ACKNOWLEDGMENT

Valuable comments from Dr. Ian Fasel improved the presentation quality of this paper. The research in Section IV was supported in part by Japan's Ministry of Internal Affairs and Communications. The analysis of the experimental data in Section V was conducted as part of the JST ERATO Asada Project.

REFERENCES

- [1] Takahiro Miyashita, Taichi Tajika, Kazuhiko Shinozawa, Hiroshi Ishiguro, Kiyoshi Kogure and Norihiro Hagita, "Human Position and Posture Detection based on Tactile Information of the Whole Body", In Proc. of IEEE/RSJ 2004 International Conference on Intelligent Robots and Systems Workshop (IROS'04 WS), Sep. 2004.
- [2] M. Inaba, Y. Hoshino, K. Nagasaka, T. Ninomiya, S. Kagami and H. Inoue, "A Full-Body Tactile Sensor Suit Using Electrically Conductive Fabric and Strings", Proc. 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 96), Vol. 2, pp. 450-457, 1996.
- [3] Naya, F., Yamato, J. and Shinozawa, K., "Recognizing Human Touching Behaviors using a Haptic Interface for a Pet-robot", Proc. 1999 IEEE International Conference on Systems, Man, and Cybernetics (SMC'99), pp. II-1030-1034.
- [4] H. Iwata, H. Hoshino, T. Morita, and S. Sugano, "Force Detectable Surface Covers for Humanoid Robots", Proc. 2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM'01), pp. 1205-1210, 2001.
- [5] H. Iwata, S. Sugano, "Whole-body covering tactile interface for human robot coordination" Proceedings - IEEE International Conference on Robotics and Automation. Vol. 4, pp. 3818-3824. 2002.
- [6] Tomoyuki Noda, Takahiro Miyashita, Hiroshi Ishiguro, Kiyoshi Kogure, Norihiro Hagita, "Detecting Feature of Haptic Interaction Based on Distributed Tactile Sensor Network on Whole Body", Journal of Robotics and Mechatronics, Vol.19, No. 1, pp. 42-51, 2007.

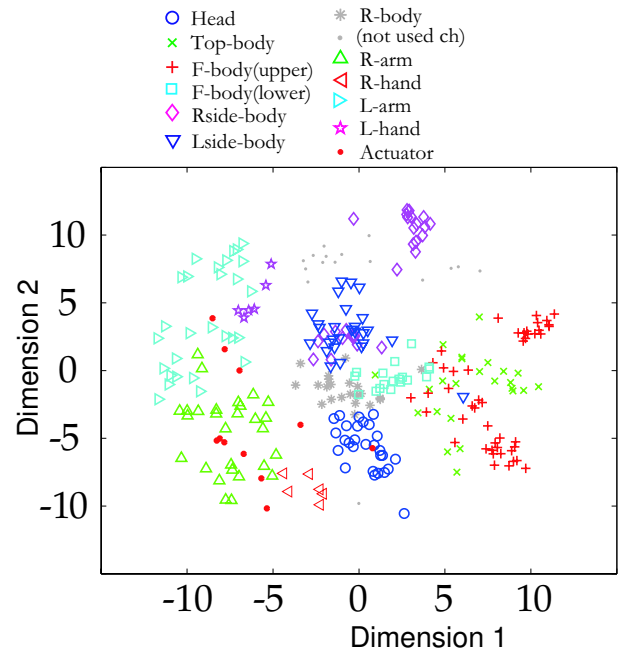


Fig. 7. This map[7] is the 2D Somatosensory Map obtained from the haptic interaction shown in fig. 5(c) in the field experiment. The time window used to compute the distance matrix, $\{d_{ij}\}$, is 2 minute of the whole interaction between the robot and the subject. The two cluster of the "F-body (upper)" sensors' distribution, around the locations of $(Dimension1, Dimension2) = (7, -6)$ and $(11, 4)$, can be interpreted as a result of that she touched the two area of the upper body with her left/right hands. This result suggests that cross-correlations contain touch features.

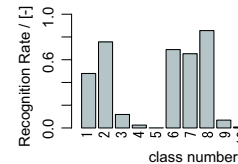
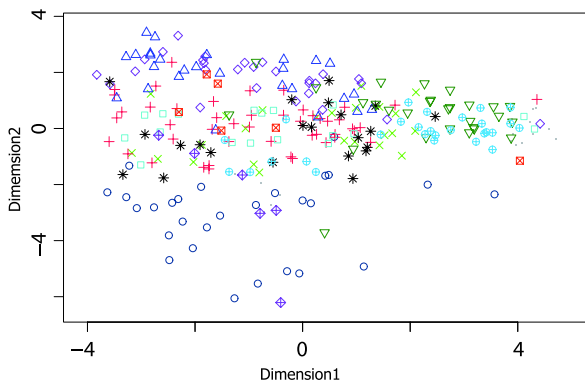
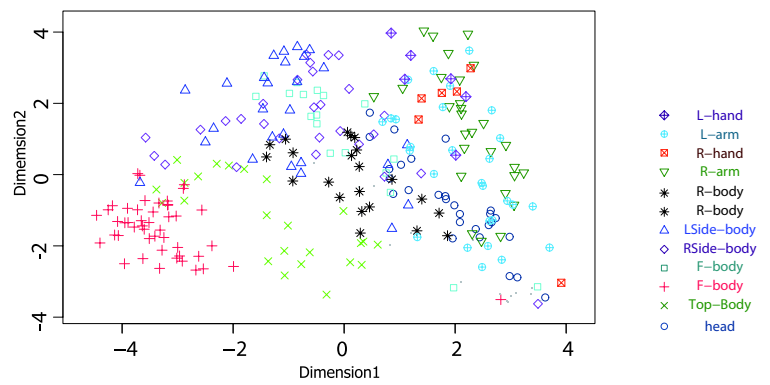


Fig. 8. Recognition Rate of the classifier using the K-Nearest Neighbor method in 15-way forced choice task ($k = 3$ is the same condition as [7], but 2 classes added to the choice. [7] was in 13-way forced choice task.)

- [7] Tomoyuki Noda, Takahiro Miyashita, Hiroshi Ishiguro, Norihiro Hagita, "Map Acquisition and Classification of Haptic Interaction Using Cross Correlation between Distributed Tactile Sensors on the Whole Body Surface", Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 07), 2007.
- [8] Torgerson, W., "Multidimensional scaling : I. Theory and method.", Psychometrika. 17, pp.401-419, 1952.
- [9] S. Watanabe and N. Pakvasa, "Subspace method in pattern recognition," Proc. 1st Int. J. Conf on Pattern Recognition, Washington DC, pp. 2-32, Feb. 1973.
- [10] Atsushi Iwashita, Makoto Shimojo, "Development of a Mixed Signal LSI for Tactile Data Processing", in Proc. of IEEE Int. Conf. on Systems, Man and Cybernetics 2004 (SMC2004), 2004.
- [11] Z. Pan, H. Cui, and Z. Zhu, "A Flexible Full-body Tactile Sensor of Low Cost and Minimal Connections", Proc. 2003 IEEE International Conference on Systems, Man, and Cybernetics (SMC'03), Vol. 3, pp. 2368-2373, 2003.
- [12] H. Shinoda, N. Asamura, T. Yuasa, M. Hakozaki, X. Wang, H. Itai, Y. Makino, and A. Okada, "Two-Dimensional Communication Technology Inspired by Robot Skin", Proc. IEEE TExCRA 2004 (Technical Exhibition Based Conf. on Robotics and Automation), pp.99-100, 2004.
- [13] D. Francois, D. Polani, K. Dautenhahn "On-line Behaviour Classifica-



(a) Somatosensory map made from $\mathbf{u}_{(8)1}$ (class 8)



(b) Somatosensory map made from $\mathbf{u}_{(11)1}$ (class 11)

Fig. 12. Somatosensory map can also visualize a \mathbf{u}_{kj} as sensor point distribution in a 2D plane, since \mathbf{u}_{kj} has the same dimension as the feature vector \mathbf{a} has. In these two maps, closer sensor pairs have larger weight in their cross-correlation element.

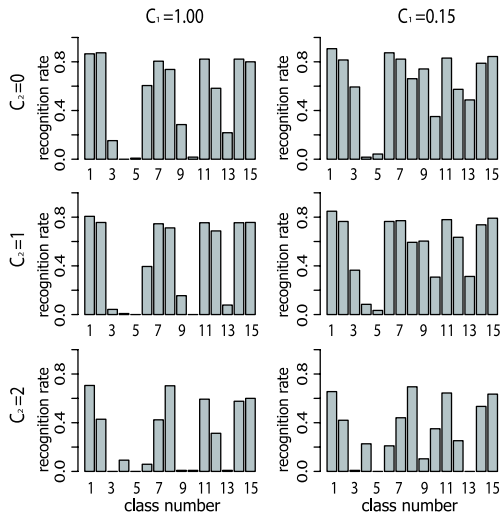


Fig. 9. Recognition rate

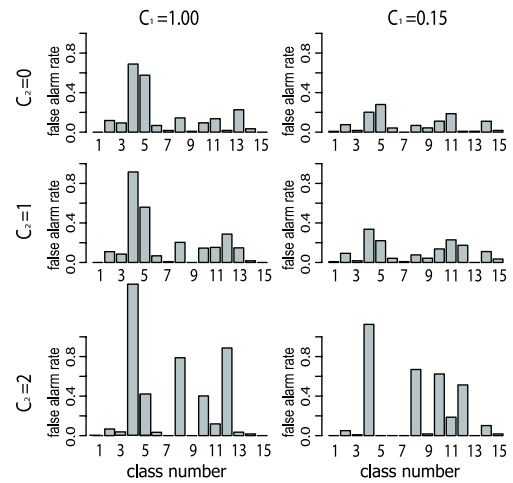


Fig. 10. False alarm rate

- tion and Adaptation to Human-Robot Interaction Styles”, poster session of HRI(2007), HRI, 2007
- [14] D. Pierce and B. Kuipers, “Map learning with uninterpreted sensors and effectors”, Artificial Intelligence, vol. 92, 1997.
- [15] Yasuo Kuniyoshi, Yasuaki Yorozu, Yoshiyuki Ohmura, Koji Terada, Takuya Otani, Akihiko Nagakubo, Tomoyuki Yamamoto, “From Humanoid Embodiment to Theory of Mind”, July 7-11, 2003, Revised Papers, pp.202–218.
- [16] Hubert R. Dince et. al., “Improving Human Haptic Performance in Normal and Impaired Human Populations through Unattended Activation-Based Learning”, ACM Transactions on Applied Perception, Vol. 2, No. 2, pp.71-88., April 2005,
- [17] Hiroshi Ishiguro, Masatoshi Kamiharako and Toru Ishida, “State Space Construction by Attention Control”, International Joint Conference on Artificial Intelligence (IJCAI-99), pp. 1131-1137, 1999.
- [18] Dahlback, N., Jonsson, A. and Ahrenberg, L., “Wizard of Oz studies - Why and How”, Proc. of the international workshop on Intelligent user interfaces, pp.193-200, 1993.
- [19] Masahiro Shiomi, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita, “Interactive Humanoid Robots for a Science Museum”, IEEE Intelligent Systems, vol. 22, no. 2, pp. 25-32, Mar/Apr, 2007
- [20] Takahiro Miyashita, Taichi Tajika, Hiroshi Ishiguro, Kiyoshi Kogure, and Norihiro Hagita, “Haptic Communication between Humans and Robots”, in Proc. of 12th International Symposium of Robotics Research, CD-ROM, San Francisco, CA, USA, Oct. 2005.

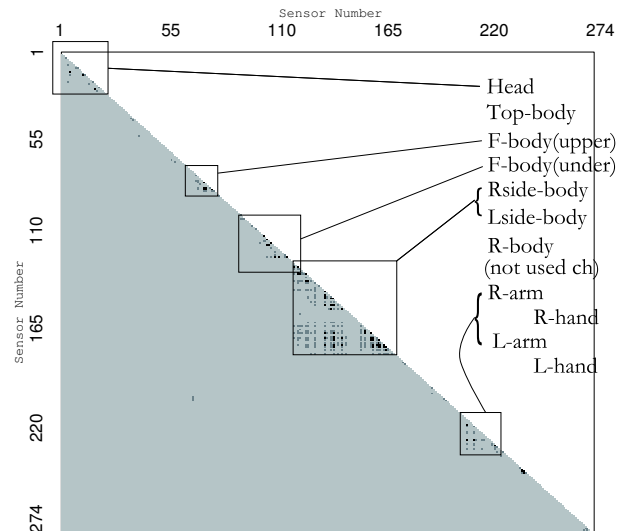


Fig. 11. Color image matrix reconstructed from orthogonal vector \mathbf{u}_{81} , highlighting the large weighted elements of the cross-correlation