Detection and Transmission of "Tsumori": an Archetype of Behavioral Intention in Controlling a Humanoid Robot

Masataka Niwa[1], Shinya Okada[1], Shota Sakaguchi[2], Kentaro Azuma[2],
Hiroyuki Iizuka[1], Hideyuki Ando[1], and Taro Maeda[1]

[1]Graduate School of Information Science and Technology, Osaka University
[2]School of Engineering, Osaka University

ABSTRACT

We propose novel architecture to realize a low-degree of freedom (DOF) Tsumori controller to manipulate a high-DOF robot. We believe that humans can immerse themselves to manipulate a robot if the robot reflects user intention. In our Tsumori controller, the user intuitively operates a joystick, and the robot moves semi-autonomously and reflects the user's intention that is extracted from intuitive input sequences. In this paper, we explain how to extract Tsumori, which is an archetype of human behavioral intention, and how to realize a Tsumori controller. We experimentally evaluated our proposed Tsumori controller by investigating how well the robot movements reflect subject intentions. We show that the robot can move with about a 60% increase in accuracy and that the context of its motions affects the intuitive input sequences. To reduce the miss-detection of intentions due to context, we investigated the context dependencies of the intuitive input sequences and found that at most a single behavior segment of the robot affects the intuitive input sequences.

KEYWORDS: Robot manipulation, behavioral intention, telexistence

INDEX TERMS: B.1.1 [Control Structures and Microprogramming]: Control Design Styles; H.2.8 [Database Management]: Database applications

1 INTRODUCTION

In a conventional robot control interface, we need to learn the preset control commands in advance to control a robot. If the intended robot motion becomes complex, the command combinations or sequences become too complex to remember. In this case, such a control interface is not intuitive because users need to convert their intentions into symbols that can be understood by the robot. In contrast, telexistence [1, 2] allows intuitive control without learning by exploiting the complete physical association between a controller and a robot. The disadvantage of this technology is that the system requires high accuracy and a large space, which is expensive.

On the other hand, science fiction books and movies often contain scenes where people freely control a robot with a simple small controller. For example, in a famous Japanese cartoon, the main character controls a humanoid robot by just holding two sticks. Children watching the TV version can put themselves into the main character and move toy controls as if they are controlling the robot. In these cases, the controlling actions of the toy controller and the movements of the animation robot do not match in terms of time and space. However, the behavioral intention (Tsumori) of controlling a robot and the actual robot movements correspond.

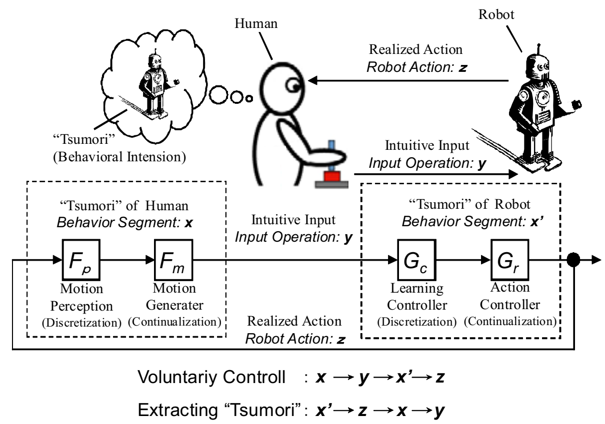

Fig. 1 Flow of controlling robot and extracting behavioral intention: Tsumori

If we can extract user intention from input sequences, we can develop a novel interface to control a robot. The interface maintains the association between the intentions and robot motions. The behavioral intentions appear in the intuitive input sequences, as in the cartoon. Based on this idea, we propose the Tsumori controller to manipulate humanoid robots more intuitively using behavioral intention. Tsumori is a Japanese word that means intention to represent the mental state where one feels like one is performing an action something (actually one is not doing). In the case of robot controlling animation, the children are actually not manipulating the robot but they feel as if they are doing so while watching TV. Thus, we call our novel interface the Tsumori controller.

2 METHOD TO EXTRACT HUMAN BEHAVIORAL INTENTION AND CONTROL THE ROBOT

Consider controlling a robot by the input device shown in Fig. 1. The segmented behavioral intention [3] of human x is converted to controller input operation y by motion generator Fm. y is converted to the behavior segment of robot x' by learning controller Gc. x' is converted to the realized action of robot z by action generator Gr. The human operator (Fig. 1) observes z, which is converted to behavior segment x. The operator compares

[1,2]2-1 Yamada-oka, Suita, Osaka, Japan.
[1]{niwa, okada.shinya, iizuka, hide, t_maeda}@ist.osaka-u.ac.jp
[2]{sakaguchi.shota, azuma.kentaro}@ise.eng.osaka-u.ac.jp

20th International Conference on Artificial Reality and Telexistence (ICAT2010), 1-3 December 2010, Adelaide, Australia
ISBN: 978-4-904490-03-7 C3450

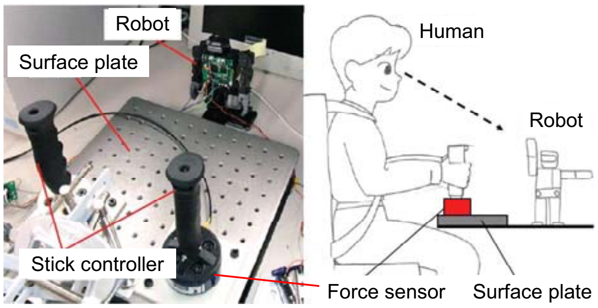

Fig. 2 Experimental equipment

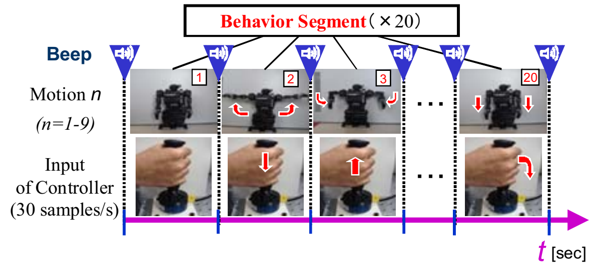

Fig. 3 Time sequences of extraction of human behavioral intentions

this x to the first x. Based on this information, the operator determines the next behavioral intention.

For robot control by the prearranged commands, users have to know the correspondence between robot motions z and controller inputs y that produce segmented motion x'. The users convert behavioral intention x into controller input y by motion generator Fm. In this way, x and x' can be mapped exactly. On the other hand, for the Tsumori controller, the interface has to know relations Gc between y and x' to map x and x'.

Thus, in the Tsumori controller, experimentally acquiring the Gc function is a problem. Therefore, we use the idea of animation where a child remembers all the robot motions and moves the toy joystick controller based on them as if he is actually controlling the robot. The procedure is as follows. First, the robot motions are segmented (x'), and a subject watches a series of segmented robot motions as a continuous sequence that consists of behavior segmented motions z. Next, the subject converts the displayed motion into segmented behavior intention x that corresponds to the movements and operates the controller while watching the robot motions, feeling as if he is actually controlling the robots. Input sequences y must reflect the subject behavioral intentions. Gc is created from the association database of inputs y and robot segmented motions x'.

Such Tsumori control can be realized by mapping the human and robot behavior segments.

3 EXTRACTING HUMAN BEHAVIOR INTENTION

Using the procedure described in the previous section, a person watches robot motion z, which is memorized and freely operates controller y. A database, which is created from relationship y and x' (gotten from z), denotes Learning Controller Gc.

3.1 Experimental system

The experimental setup is shown in Fig. 2 (left). The humanoid robot (TOMY COMPANY, LTD., i-sobot) has 17 degrees of freedom. The robot is controlled all of its 17 servo motors by a PC through a microcomputer (Microchip, PIC18F2525). The robot, which was controlled at 30 Hz, autonomously generates predetermined motion sequences by combining the robot behavior segments. One robot behavior segment is designated as a motion from an absolute pose to another pose. The time of a single behavior segment is 0.8 or 1.6 sec: a quarter or a half breathing cycle.

The two stick controllers were made from grips attached to a 6-axis force sensor (Nitta Corporation, IFS-67M25A) and fixed to the surface plate. The force sensors output the direction of translational force Fx, Fy, Fz and moment force Mx, My, Mz at 30 Hz. The robot, stick controller, and subject positions are shown in Fig. 2 (right).

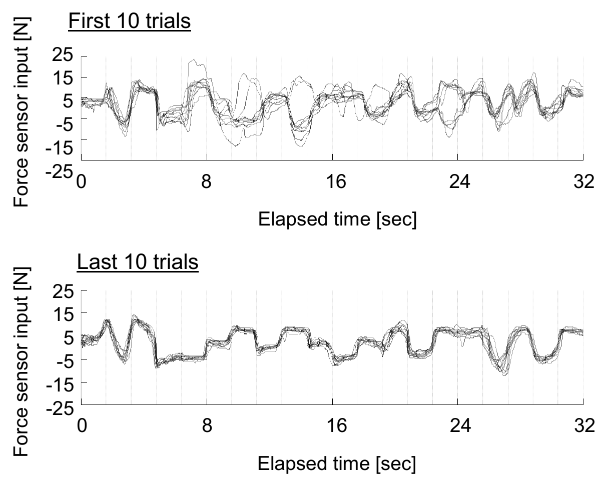

Fig. 4 Example of right stick controller input sequences of Fx. First and last ten trials are shown at top and bottom.

3.2 Procedure of extracting human behavioral intention

We prepared 120 robot behavior segments and created nine robot motion sequences. A single robot sequence consists of 20 robot behavior segments.

The subjects freely input with the stick controller while watching the motion sequence. At this time, it is difficult to detect what subject input corresponds to which part of the robot motion. So the subjects listen to a beep at the behavior segment to match one robot behavior segment and one human behavior segment (Fig. 3). Each motion sequence was presented five times × 12 trials, and each trial was presented in random order (nine motions × five times × 12 trial: 540 trials). We checked the last ten motions. All subjects had to remember the robot motion, and the force inputs to the stick controller were stable enough (Fig. 4). We recruited five people in their twenties as subjects (A, B, D, and E: men, C: woman). Subjects A, B, and C performed the experiments under the condition where the robot behavior segments length was 1.6 sec, and subjects A, D, and E were tested with 0.8 sec segments. Subject A was involved in both.

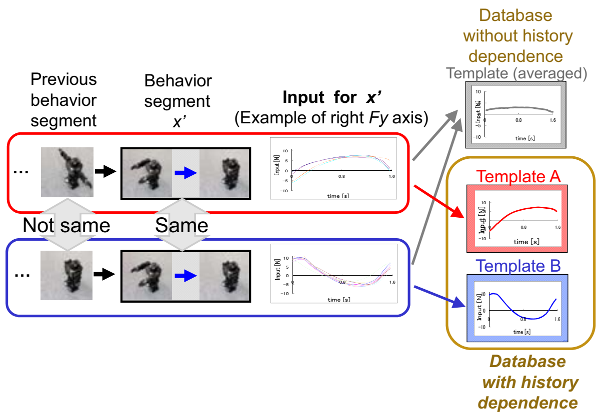

Fig. 5 When previous robot behavior segment is different, template is divided even if robot behavior segment is equal.

3.3 Building and evaluating a database of learning controller *Gc*

Based on the preceding section, a database for each subject was constructed from the relationship between stick controller inputs y and robot motions z. Each bit of data in the database has robot behavior segments and 12-axis (6-axis × 2 (left and right)) force inputs that were averaged over the last ten trials of input waveforms (template).

To evaluate the database obtained in the previous section, the subjects performed the same experiments as the novel robot motion sequences to investigate how accurately the database can generate correct robot motions from intuitive input sequences.

Then a new robot motion sequence was created with 12 robot behavior segments from the database. The subjects operated the stick controller in the same way as in Section 3.2. The last ten of 60 trials were used because by that time the subjects completely remembered the robot motion. Therefore, we assumed they controlled the robot.

The acquired input data and each template were matched using a regularized cross-correlation function, as shown in expression (1). When I_{NCC} is highest, the data that match the template are chosen as a result. This piece of chosen data is assumed to estimate the robot behavior segment.

$$I_{NCC} = \sum_t \left\{ \sum_t^{12} \left(\frac{f_i(i) - \bar{f}_i}{\sqrt{\sum_t \left(f_i(t) - \bar{f}_i\right)^2}} \cdot \frac{g_i(t) - \bar{g}_i}{\sqrt{\sum_t \left(g_i(t) - \bar{g}_i\right)^2}} \right) \right\} \quad \ldots\ldots \quad (1)$$

$F\left[f_{rFx}, f_{rFy}, \ldots, f_{lMy}, f_{lMz}\right]$: Data of force sensor inputs (12-axis)

$G\left[g_{rFx}, g_{rFy}, \ldots, g_{lMy}, g_{lMz}\right]$: Data of database (12-axis)

We investigated how much the estimated robot behavior segments matched the behavioral intention of the subjects. The percentage of correct matches was about 60% in all subjects. This value exceeds chance, suggesting the realizability of our proposed robot control system.

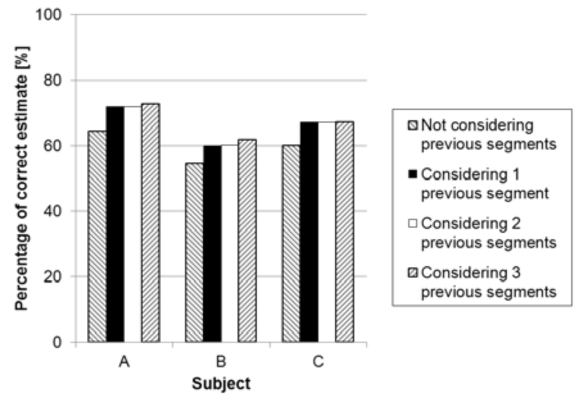

Fig. 6 Matching result when length of robot behavior segment is 1.6 sec

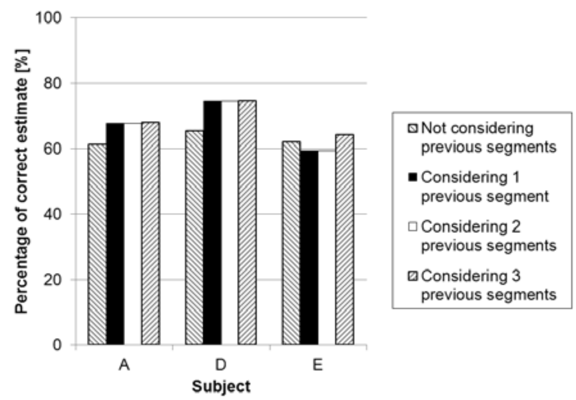

Fig. 7 Matching result when length of robot behavior segment is 0.8 sec.

3.4 Building a database that considers previous robot behavior segments

We found that stick controller input is not always the same even if the robot behavior segment is identical. The input sequence seems to depend on the previous robot behavior segment. So we modified the database to adapt the context dependencies by storing the templates separately if the previous robot behavior segments are different (Fig. 5). This database was evaluated using the input data and the same technique obtained in Section 3.3.

The result is shown in Figs. 6 and 7. As the result of five subjects, the percentage of correct matches increased when the built database considered the previous robot behavior segments. This result means that human behavior intention was influenced by the previous behavior intention. Moreover, regardless of the time length of the segment behavior, it influenced one previous behavior intention. Human behavior intention was influenced by the unit of the human behavior segment. In addition, no difference was found between the results of 1, 2, and 3 previous behavior segments. Just one final behavior segment of the robot affected the intuitive input sequences.

4 Discussion

For the results of subject E in Fig. 7, since the percentage of correct matches did not increase when the built database considered the previous robot behavior segments, we examined the input data. As shown in Figs. 8 and 9, the inputs and templates

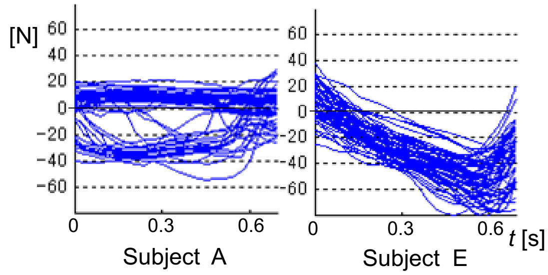

Fig. 8 Examples of subjects input when one robot behavior segment that previous robot behavior segments are different (5 patterns).

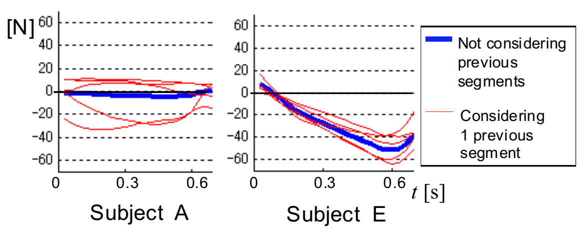

Fig. 9 Template of database made based on input of Fig. 8

of subject A are divided into several groups. However, the inputs and templates of subject E have similar trends. So we could confirm that the input of subject E was not influenced by the previous behavior intention. In the future, we will increase the number of subjects and investigate how many show this trend.

5 Conclusion

We proposed novel architecture to realize a low-degree of freedom Tsumori controller to manipulate a high-degree of freedom robot. We extracted Tsumori, which is an archetype of human behavioral intention, built, and evaluated a Learning Controller database. We showed that the robot moved with about 60% accuracy and that the context of its motions affected the intuitive input sequences. To reduce the miss-detection of intentions due to the context, we investigated the context dependencies of the intuitive input sequences and found that at most a single last behavior segment of the robot affects the intuitive input sequences.

In the future, we will investigate a matching algorithm for our proposed system.

Acknowledgments

This work was partially supported by grant-in aid (No. 22240008) and also by a grant-in-aid from the Global COE program: Founding Ambient Information Society Infrastructure of the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

References

[1] S. Tachi, K. Tanie, K. Komoriya, and M. Kaneko. Tele-existence (I): Design and Evaluation of a Visual Display with Sensation of Presence. Proceedings of RoManSy'84, The Fifth CISM-IFToMM Symposium, pp. 245-254, 1984.

[2] S. Tachi and H. Arai. Study on Tele-existence (II)-Three Dimensional Color Display with Sensation of Presence-, Proceedings of the '85 ICAR (International Conference on Advanced Robotics), pp. 345-352, 1985.

[3] M. Ito and J. Tani: On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system. Adaptive Behavior, Vol. 12, No. 2, pp. 93-115, 2004