Multisensors-based Approach for Intention Reading with Soft Computing Techniques

Zeungnam Bien, Dae-Jin Kim and Hyong-Euk Lee
zbien@ee.kaist.ac.kr, {djkim, helee}@mail.kaist.ac.kr
Div. of EE, Dept. of EECS, KAIST
373-1 Guseong-dong, Yuseong-gu, Daejeon 305-701, Korea

Kwang-Hyun Park
akaii@robotian.net
Human-friendly Welfare Robotic System Engineering Research Center, KAIST
373-1 Guseong-dong, Yuseong-gu, Daejeon 305-701, Korea

Haiying She, Christian Martens and Axel Gräser
{she, cmartens, ag}@iat.uni-bremen.de
Institute of Automation, University of Bremen
Otto-Hahn Allee, NW1
D-28359 Bremen, Germany

***Abstract*— Human's intention plays a key role in human-machine interaction as in the case of a robot serving for a handicapped person. The quality of a service robot will be much enhanced if the robot can infer the human's intension during the interaction process. In this paper, we propose a soft computing-based technique to read a user's intention using some multisensors-based approach. We have tested the technique by a scenario of 'serving a drink to the user'. With such force/torque or vision sensor, the robot can effectively infer the user's intention to drink the beverage or not to drink. As an application, this intention technique is employed for building a rehabilitation robot, called KARES II, to perform various human-friendly human-robot interaction.**

***Index Terms*—Soft Computing, Intention Reading, Contact Detection, Force/Torque Sensor, Mouth Openness.**

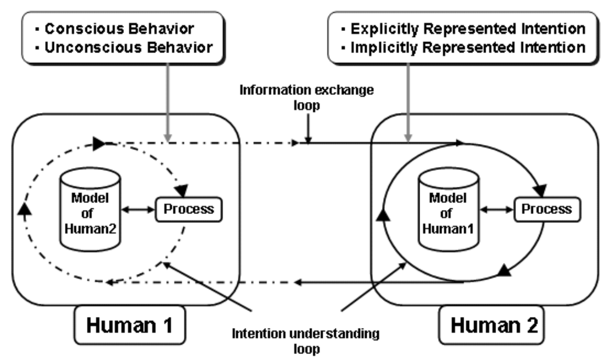


Fig. 1. Intention reading loop in human-human communication [2]

I. Introduction

INTENTION refers to *one's determination to act, in a certain way, what one intends to do, or a concept considered as the product of attention directed to an object of knowledge* [1]. In case of a human being, one's intention is usually represented by means of two forms such as 'conscious behavior' and 'unconscious behavior'. Human beings understand each other's intention by interpreting information received during communication. Thus, information exchange with the other person and interpretation of the received information are essential functions to understand human intention.

When a human being understands a partner's intention through communication, two loops of information processing realize the above functions as shown in Fig. 1. Human beings exchange information with each other through an information exchange loop. Information exchange is realized by means of language and behavior. On the other hand, a human being understands the other partner's intention in an intention understanding loop. Intention understanding is accomplished on the basis of knowledge concerning the partner (see Fig. 1)[2].

In this paper, we propose a novel intention reading scheme based on soft computing techniques for a multisensors-configured robotic system. Specifically, we have used two types of sensors which are a force/torque sensor and a vision sensor. We consider the situation such as 'serving a drink to the user' by robotic arm with the sensors. Soft computing-based approach proves to be effective to deal with uncertain and ambiguous information from the sensors.

This paper is organized as follows. Brief introduction to our rehabilitation robot, KARES II and a scenario for human-robot interaction are mentioned in Section II. In Section III, soft computing-based intention reading is introduced. Here, we show detailed scheme of our approach for intention reading using two sensors. In Section IV, The real experiment is described

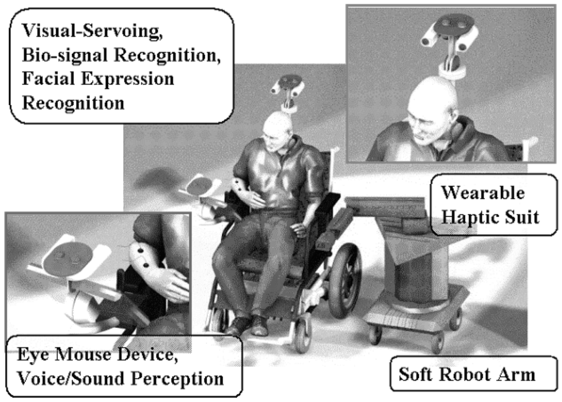


Fig. 2. The wheelchair-based robotic arm and its human-robot interaction technologies

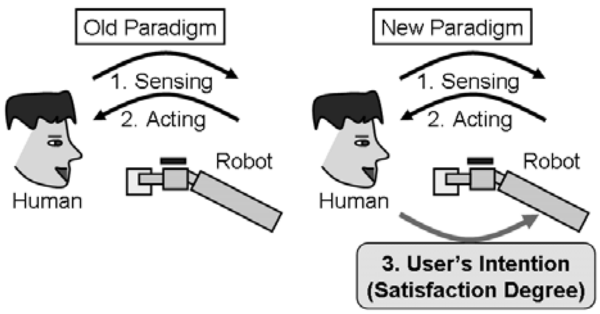


Fig. 3. Intelligent visual servoing scheme

for the rehabilitation robot performing the task of 'serving a drink to the user'. Finally, concluding remarks are given in Section V.

II. Intention Reading in Rehabilitation Robot, KARES II

A. Rehabilitation Robot: KARES II

Rehabilitation robotics is concerned with application of the robotic technology to the rehabilitative needs of people with disabilities as well as growing population of the elderly [3]. Typically, it is known not only as a kind of service robot based on robotic technology but also as a kind of assistive devices based on rehabilitation engineering. Rehabilitation robots aim to solve daily living problems of individual activities. Hence, the primary role of rehabilitation robots is to endow more independence with improvement of life quality.

In this paper, we consider a wheelchair-based rehabilitation robotic system, KARES (KAIST Rehabilitation Engineering Service system) II, which we are developing as a service robotic system for the disabled and the elderly (see Fig. 2). KARES II has many human-robot interaction subsystems for human-friendly usage [4]. Among these, we consider vision sensor and force/torque sensor on the robotic hand.

B. Human-friendly Interaction using Multisensors

To implement human-friendly interaction using multisensors, an 'intelligent visual servoing' scheme is proposed by W.-K. Song [5] as shown in Fig. 3. On the contrary to old paradigm, the new paradigm proposes usage of user's satisfaction degree as a new feedback signal from the user by using various sensor systems. Using this scheme, the robot acts more intellignet way to provide safety and comfortability during services. Facial expression is well-known as a very natural and comfortable way for interaction process. Note, however, that, when the robotic hand is very close to the human's face, it is almost impossible to get enough information from the user's face and to act accordingly. Thus, in this case, we need another sensor to infer the user's intention during drinking. Furthermore, if we cannot utilize such kind of information, the robotic hand may spill a drink on the user's body.

Fig. 4 shows a procedure of the task 'serving a drink to the user' based on the proposed multisensors-based approach. As shown in Fig. 4, the robot can serve as an intelligent servant according to the user's intention which can be read from multisensors. Force/torque sensor is dominant when the robotic hand is closely located near the user's mouth. On the other hand, the vision sensor is effective when the distance between the user's mouth and the robotic hand is far away.

III. Soft computing-based Intention Reading

For a typical robotic system, reading of the user's intention is considered as a very ambiguous and difficult task in view of a conventional mathematical approach. However, humans do such intention reading without much burden. Because of its similar information processing strategy, soft computing is widely used as one of the solutions for these kinds of high-level problems.

As mentioned in the previous section, two sensors are used in a mixed way for intention reading for the 'serving a drink to the user' task of KARES II. (see Fig. 4).

More specifically, in our environment, we have combined and fused information from 2 sensors according to the level of dominance as shown in Fig. 5. Thus, the weighted sum of two information is acquired during human-robot interaction. In the following, we describe intention reading for each type.

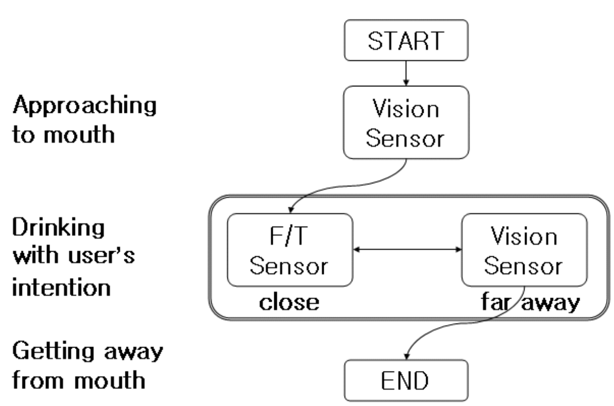


Fig. 4. 'Serving a drink' procedure using multisensors-based approach

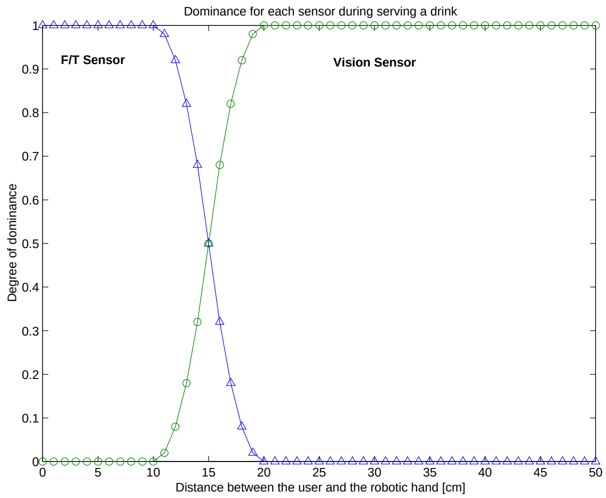


Fig. 5. Dominance of each sensor during the interaction

A. Force/Torque Sensor-based Approach for intention reading: Contact/Non-contact Detection

For a successful task of 'serving a drink to the user', the robot should detect whether the user touches the glass with his/her mouth or not. Thus, the force/torque sensor can be mainly used for contact/non-contact detection.

During drinking a beverage, the force/torque sensor responds very differently due to the various motion of the user's mouth, as shown in Fig. 6.

Although there are reported many similar cases [6], the results are not satisfactory because of their usage of simple threshold-based decision for contact/non-contact detection. However, humans can easily detect the point where the real contact point is with given force signal in Fig. 6.

Hence, a soft computing-based approach (especially, a fuzzy logic-based approach) comes into consideration naturally. We have defined three input linguistic variables such as 'F' (absolute value for force signal from force/torque sensor), 'Fv' (force velocity: absolute value for first derivative of force signal) and 'Fa' (force acceleration: absolute value for second derivative of force signal). For each input linguistic variable, we considered two linguistic values 'L' (Large) and 'S' (small). For output linguistic variable, 'ST' (state) is considered with two linguistic values 'C' (contact state) and 'NC' (non-contact state). After analyzing the signals, the decision rules are chosen as follows:

Rule 1: If F is L and Fv is L then ST is C.
Rule 2: If F is L and Fa is L then ST is C.
Rule 3: If Fv is S and Fa is S then ST is NC.

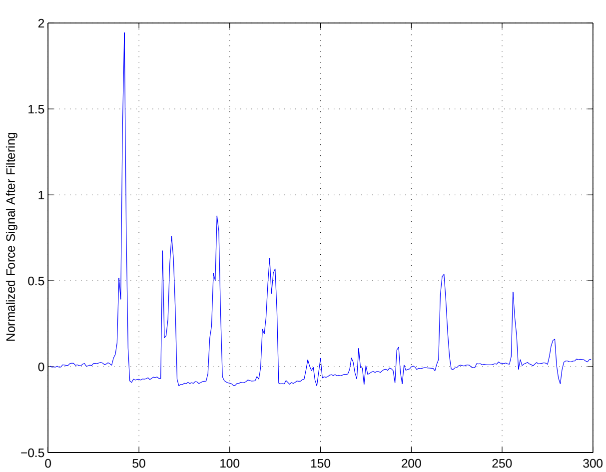


Fig. 6. Force signal during contact of ten times

TABLE I
SIMPLE SCENARIO FOR INTENTION READING

Intention	Behavior
Positive intention	When the user wants to drink, the user opens his/her mouth.
Negative intention	When the user does not want to drink, the user closes his/her mouth.
	In some cases, the user can show his/her negative intention by shaking head.

B. Vision-based Approach for intention reading: Mouth openness and head shaking

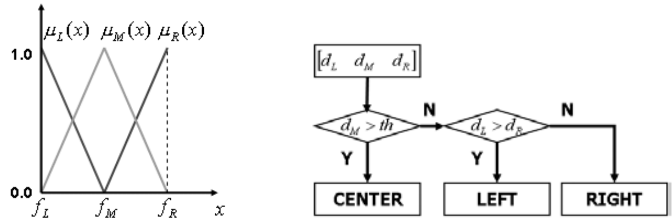
During the task of 'serving a drink to the user', we assume that user's intention can be read from the degree of mouth openness to drink or not during the service by the robotic arm (see Fig. 7). Furthermore, the user can represent his/her strong negative intention by shaking head. Thus, we can make a simple scenario for intention reading as shown in Table I.

Thus, based on this scenario, we use a vision-based approach for intention reading with mouth openness and head shaking [7]. To acquire degree of mouth openness and degree of head shaking, we propose two image features through severe considerations.

1) Degree of mouth openness: To get the degree of mouth openness, we need some special features to roughly determine user's intention. Here, we propose a Gabor-Gaussian (GG) feature [8]. GG feature consists of gabor filtered coefficients, its projection and gaussian weighted form. GG feature f_G is defined as Eqn. (1).

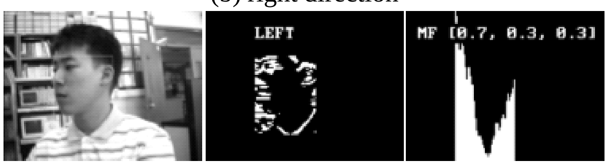
$$f_G = \frac{\sum_{j=1}^{H-1} w_G(j) dy_{proj}(j)}{\sum_{j=1}^{H-1} w_G(j)}, \quad (1)$$

Fig. 7. Positive/negative intention from user's mouth openness



(a) membership functions and decision logic

(b) right direction

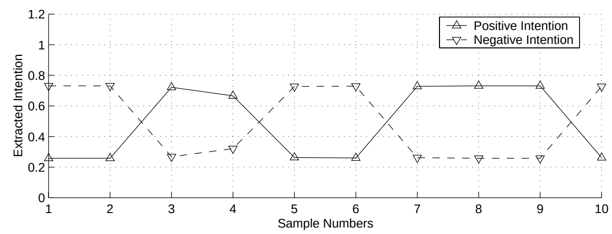


(c) left direction

Fig. 8. Membership function-based approach for shaking head

Fig. 9. Visual Servoing with intention reading

(a) Ten sequential images of the user with varying mouth openness



(b) Intention through ten sequential images

Fig. 10. Intention reading of the user based on mouth openness

where, $w_G(j)$ means the Gaussian weights, $dy_{proj}(j)$ is the absolute values of the derivative of the projected value $y_{proj}(j)$, H is the height of Gabor-filtered image.

2) Degree of head shaking: To acquire the information about shaking head, membership function-based features are proposed such as Eqn. (2)-(4). These features are calculated using horizontal edge image. By simple decision logic, we can effectively acquire the directional information of the user's head.

$$d_L = \sum_{x=f_L}^{f_M} I_{Hproj}(x) \cdot \mu_L(x) \quad (2)$$

$$d_M = \sum_{x=f_L}^{f_R} I_{Hproj}(x) \cdot \mu_M(x) \quad (3)$$

$$d_R = \sum_{x=f_M}^{f_R} I_{Hproj}(x) \cdot \mu_R(x) \quad (4)$$

where, d_L, d_M and d_R denote degree of each facial direction, respectively. I_{Hproj} is a horizontally projected image of edge coefficients. Other parameters are depicted in Fig. 8(a).

In Fig. 8(b)-(c), some examples for direction estimation are given with three feature values. Using these procedure, we can effectively know the status of the user's head whether it is shaking or not. Among three available states, the degree of mouth openness can be extracted only for frontal face (in case of 'center' direction).

IV. Results

For the experiment, we used 6-DOF PUMA type configured robotic arm, POWERCUBE (from Amtec), micro-type color CCD cameras CV-M2350 (from JAI) as vision sensor and JR3 6-DOF force/torque sensor (from Nitta) to acquire the user's intention.

Specially, we also adopted our small-sized/light-weighted cable-driven stereo camera head in eye-in-hand configuration [9]. Fig. 9 shows the intermediate process of our experiment during serving a drink by the robotic arm.

Fig. 10 shows a result for intention reading using ten sequencial facial images. Using proposed mouth extraction method and soft computing-based intention reading scheme, we can clearly extract the intention information as shown in Fig. 10(b). Here, solid line means the degree of positive intention and dashed line means the degree of negative intention.

Fig. 11 shows a result of human-robot interaction based on vision sensor-based approach including head shaking information. As shown in Fig. 11, when the user shakes his/her head, the robot hand moves backward. As we already mentioned, the degree of mouth openness is estimated only for 'center' directed face (see Fig. 11, located in the center, from top to bottom). The robot hand approaches to the user's mouth according to the degree of mouth openness. When the degree is high, the robot hand approaches fastly. However, when the user tightly closes his/her mouth, the robot hand stops immediately. By this way, the robot hand provides human-friendly human-robot interaction in natural way.

Additionally, when the distance between the user's face and the robotic hand is close, force/torque sensor-based intention reading is effectively performed. Fig. 12 shows the result of contact/non-contact detection with given fuzzy logic rules for force signal in Fig. 6.

Here, in Fig. 12, the solid lines designate the detected information of contact/non-contact based on the proposed deci-

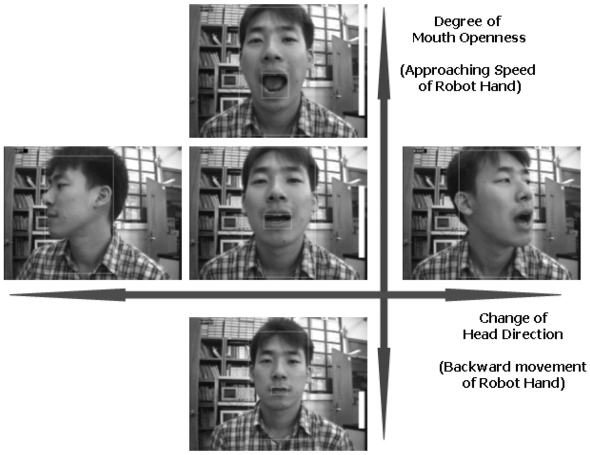


Fig. 11. Human-robot interaction using vision sensor-based approach

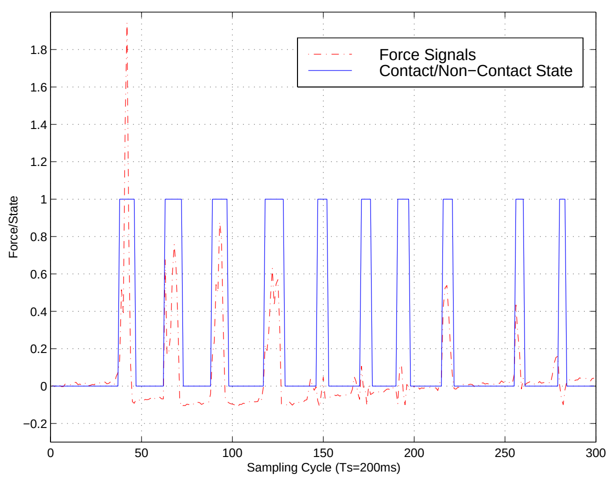


Fig. 12. Detection results using fuzzy-logic

sion rules. The dashed lines refer to the force signal. It can be seen that, though the force signal for each contact is very different, the fuzzy logic-based approach is quite successful to detect contact/non-contact situation.

V. Concluding Remarks

In this paper, we have introduced soft computing-based intention reading through multisensors-based approach. Force/torque sensor and vision sensor are used according to their dominance, depending on the distance between the user's mouth and the robotic hand. Soft computing is essential to deal with many types of ambiguous information during human-robot interaction. As a real example, we have applied our method for the human-robot interaction with the rehabiliation robot, KARES II. The result seems a very promising. We believe that this kind of multisensors-based approach is very useful and essential for real-world applications involving human-machine interactions.

As further works, we are studying ways of incorporating learning capability with more complex situations with multi-level decision values and its application to other human-robot (or human-computer) interaction process.

Acknowledgments

This work is supported by the Ministry of Science and Technology of Korea as a part of Critical Technology 21 Program on "Development of Intelligent Human-Robot Interaction Technology". And this work is done by cooperative work between IAT, University of Bremen in Germany and HWRS-ERC, KAIST in Korea which is partially supported by the DAAD as a part of the HOST program.

References

[1] "Merriam-Webster's Collegiate Dictionary," *http://www.m-w.com/cgi-bin/dictionary*.

[2] D.-J. Kim, W.-K. Song, J.-S. Han and Z. Bien, "Effective Intention Reading Technique as a means of Human-robot Interaction for Human Centered Systems," in IEEE International Conference on Fuzzy Systems, S303, 2001.

[3] M. Hillman, "Introduction to the special issue on rehabilitation robotics," Robotica, vol. 16, p. 485, 1998.

[4] Z. Bien, W. Song, D. Kwon, M. Chung, P. Chang, H. Park, D. Kim, J. Kim and K. Lee, "A Wheelchair Robot System and its Various Interface Methods for the Disabled Persons," in 1st IARP/IEEE-RAS Joint Workshop on Technical Challange for Dependable Robots in Human Environments, VI-4, 2001.

[5] W.-K. Song, "Intelligent Visual Servoing using Space Variant Vision with Application for Rehabilitation Robots," Doctor's thesis, Dept. of EECS, KAIST, 2001.

[6] W. Friedrich, P. Lim and H. Nicholls, "Versatile end-effector control for stationary and mobile manipulator platforms". Proceedings of FSR'99, 1999.

[7] D.-J. Kim and Z. Bien, "Soft computing-based Intention Reading through the user's mouth for Human-friendly Human-Robot Interaction". Proceedings of SCIS&ISIS2002, in press.

[8] W.-K. Song, D.-J. Kim, J.-S. Kim and Z. Bien, "Visual servoing for a user's mouth with effective intention reading in a wheelchair-based robotic arm," in Proc. IEEE ICRA 2001, pp. 3662–3667, 2001.

[9] J.-Y. Choi, "Design of a Behavior-based Controller using a Novel Camera Head and its Application to Service Robots," Master's thesis, Dept. of EECS, KAIST, 2000. (in Korean)