Teaching Assistant System used Eye Tracking Device Based on Gaze Estimation by Neural Network and Intention Recognition by Fuzzy Inference

INFORMATION
Volume 17, Number 9(B), pp.4661-4676

ISSN 1343-4500

Teaching Assistant System used Eye Tracking Device Based on Gaze Estimation by Neural Network and Intention Recognition by Fuzzy Inference

Mao Wang* , Yoichiro Maeda** and Yasutake Takahashi*
*Graduate School of Engineering, University of Fukui, Fukui, 910-8507 Japan
**Faculty of Engineering, Osaka Institute of Technology, Osaka, 535-8585 Japan
*{mwang,yasutake}@ir.his.u-fukui.ac.jp
**maeda@bme.oit.ac.jp

Abstract

Intention recognition can use multiple factors as inputs such as gestures, face images and eye gaze position. On the other hand, eye tracking technology, with its special advantages of applying to Human-Computer Interaction (HCI), can be utilized to develop assistant systems for people with mobility difficulties. In this paper, we propose gaze estimation position information as input of fuzzy inference to achieve intention recognition based on object recognition and construct an assistant system by using humanoid robot. Our approach is divided into three parts: user's gaze estimation, intention recognition and behavior execution. In gaze estimation part, differing from the previous studies, neural network has been used as the decision making unit, and then gaze position on computer screen is estimated. In intention recognition part, user intention is recognized by using gaze frequency and continuous gaze staying time as input of fuzzy inference after an initial intention region set has been found. At last, by using an autonomous humanoid robot, experiments are performed based on the result of intention recognition. After confirmed by user, the robot was controlled with an assistant task for user precisely.
Keywords : Eye Tracking, Neural Network, Intention Recognition, Fuzzy Inference, Humanoid Robot

1 Introduction

Recently, intention recognition, recognizing intention of a user or an agent by analyzing their actions or changes of state, is becoming an important issue in various research fields. This is because intention recognition can make the Human-Computer Interaction (HCI) more convenient. So far, many intention recognition approaches have been proposed. Much of early works were in the context of speech understanding and response automatically[1]. For example, Pynadath and Wellman achieved plan recognition on a problem in traffic monitoring through the context exploited by using a general Bayesian framework[2]. More recently, Pereira et al described an approach to tackle the intention recognition by combining dynamically configurable and situation-sensitive Causal Bayes Networks plus plan generation techniques[3]. Mao and Gratch have presented a utility-based approach to solve recognition of intention, which is realized by incrementally using plan knowledge and observations to change state probabilities[4]. In their researches, the probability is a main factor which used to infer the human intention.

Eye tracking technology has special advantages in application in area such as amyotrophic lateral sclerosis (ALS) and is useful in aiding interactions with computer and others. Unlike traditional methods, interaction through eye tracking may feel convenient and direct, especially for users needing to interact with computers but unable to manage a keyboard or mouse[5]. For instance, ALS victims, who ultimately lose any ability to initiate and control voluntary movement, are able to use eye movement in eye-tracking-based interaction.

Eye tracking technology has been usually used as an interaction tool between user and computer by researchers mainly. For example, it has been used to control a home operating system[6]. Also, it has proven to be a useful method of navigating a robot[7]. Most researchers and research groups focus on research related to improve eye tracking accuracy[8]. But it is always acting as a controller between human and computer in all the applications mentioned above. Different from the usual application, this paper focuses on the intention recognition of human based on soft computing methods.

We designed an eye tracking system with the features of low cost, easy to use and high accuracy at first. And based on this device, we proposed an approach to infer user's intention region recognition by using fuzzy theory. One of the main issues in the intention recognition process was finding an initial set of possible intentions[9]. Therefore, in this research, we found the possible user's intention region based on object recognition by image processing. Then based on data of frequency and time of user's gaze appeared by eye tacking, intention region was estimated by fuzzy inference.

The overall procedural flow of the system proposed is summarized in Fig.1. First, we detected pupil center and performed gaze estimation by using eye tracking device. Second, an initial set of intention was constructed by processing image from robot vision. Third, user's intention was recognized by fuzzy inference.

2 Eye Tracking System

Before intention recognition, user's gaze position must be detected and estimated accurately. Thus an eye tracking system with high accuracy is necessary for this research. In this section, our prototype eye tracking system is explained about hardware device and software program.

2.1 Prototype Eye Tracking Device

Eye tracking technology is divided into two modes: remote camera based mode and wearable device based mode. The remote camera-based mode has the ad-

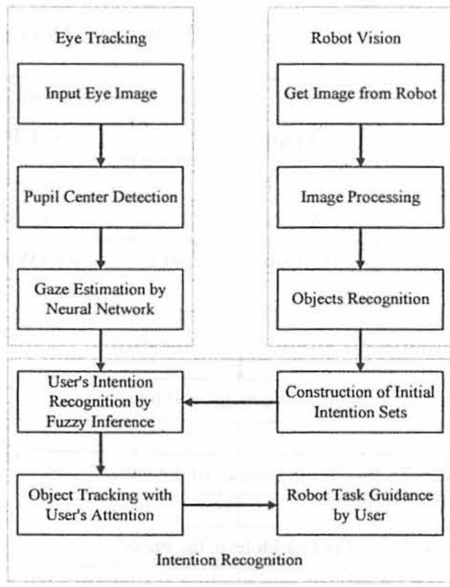


Figure 1: Overview Process of Proposed System

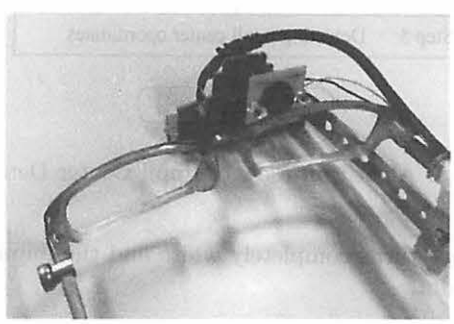


Figure 2: Prototype Eye Tracking Device

vantage of being non-intrusive, convenient and applicable to different computer applications. It has the disadvantages however, of requiring a high-resolution camera due to how far the camera is from the user's eye and more than one camera pan-tilt device must be need. All of this increases system complexity and cost. The wearable device based mode estimates user's gaze through a camera with near infrared (NIR) light illuminators attached to a glass frame or a helmet. These eye tracking systems developed for commercial use are based on the remote camera such as Tobii TX300 eye tracker manufactured by Tobii technology. These eye tracking systems have the disadvantages of being very expensive. So development of a cheaper and simple eye tracking system is needed.

As shown in Fig.2 and detailed in Table 1, the eye tracking hardware we fabricated is wearable device including an eye capture camera attached with NIR LED. To make eye tracking easier, we illuminate the eye with IR light and observe it through an IR sensitive camera with a visible light filter. After doing

Table 1: Specifications of Prototype Device

CCD Camera	Spatial resolution	640×480 pixels
	Frame rate	60 FPS
	Lens focus	fixed
NIR LED	Wavelength	940nm
	Luminous intensity	40mW/Sr

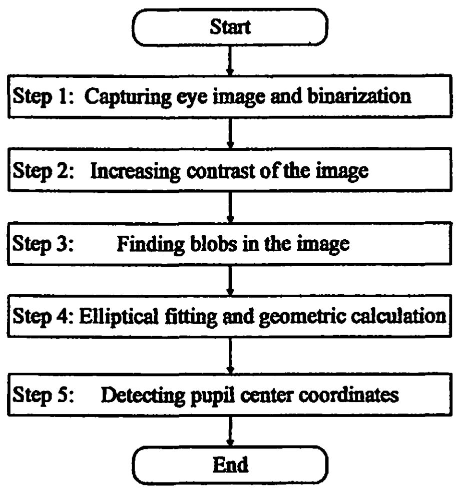


Figure 3: Algorithm Flow of Pupil Center Detection

this the iris of the eye turns completely white and the pupil standing out as a high-contrast black dot.

2.2 Pupil Center Detection

As showed in Fig.1, the process of eye tracking consists the following three steps:

1) Input eye image: the image obtained by using the CCD camera mentioned in Table 1 illuminated by IR light;

2) Pupil center detection: a process by image binarization, contours detection and convex hull calculation;

3) Gaze estimation: a process estimating user's gaze by using neural network;

Pupil center detection is the first part of an eye tracking system, the most important part at the same time[10]. In this research, we detect user's pupil center through eye image processing. The schematic diagram of process flow of pupil center detection is shown in Fig.3.

We first captured eye image in step 1 using a CCD camera and process the binary image, as shown in Fig.4. In step 2, image contrast is increased to make detection process easier. Although mathematical transformation is a large change, it is generally not apparent in image. Program searched for any blobs existing in the image and recorded feature points of the optimal one after

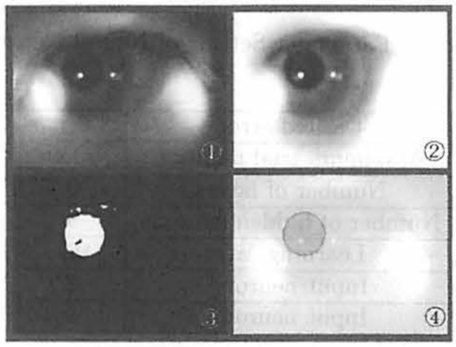


Figure 4: Eye Image Capturing and Pupil Center Detection

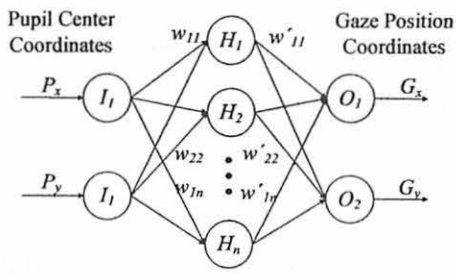


Figure 5: Neural Network for Gaze Estimation

filtering in step 3. In step 4, utilizing a Sklansky algorithm[11], convex shape of feature points is calculated. Finally, pupil center coordinates were obtained by calculating the geometric center after elliptical fitting in step 5. Pupil center coordinates are obtained as shown in 4 of Fig.4.

2.3 Gaze Estimation

The primary task in eye tracking system is estimating user's gaze, which is also the foundation of interaction between the human user and the computer in this method. Gaze estimation is achieved using a neural network to improve system robustness and adaptability. The calibration process based on pupil center coordinates obtained by using the two-input and two-output neural network with a standard back propagation algorithm as shown in Fig.5.

Where input P_x, P_y and output G_x, G_y are pupil center coordinates on the 2D camera image plane and the user's gaze position coordinates on the computer screen. w_{1i} is the weight between input node (I_1) and hidden node H_i.We use a sigmoid function as the transmission function. The parameters of NN are shown in Table 2.

In Table 2, desired error 0.001 means error pixels per 1366 pixels. Because in our program, the outputs of NN must be in the range of 0 to 1, but actually the data source are screen positions which in our experiments are in the range of 0 to 1366 pixels. Fig.6 shows the learning result for the mean square errors (MSE) with various choice of NN's hidden neuron number according to the learning process of NN training. In this figure we can find that when hidden

Table 2: Parameter Setting of Neural Network

Desired error	≤0.1%
Maximum trial number	3000
Number of layers	3
Number of hidden neurons	9
Learning rate	0.7
Input neurons	2
Input neurons	2

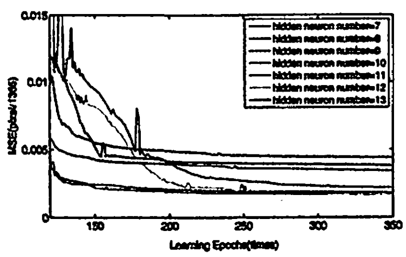


Figure 6: Learning Result of NN According to Different Hidden Neuron Nodes Number

neuron number is nine, training process has the best convergence rate. Thus, in this research, we selected nine as NN's hidden neuron number.

Output value G_x of the neural network is calculated as follows:

$$G_x = (1 + exp(-\sum_{i=1}^{n}(1 + exp(\sum_{j=1}^{2} I_j w_{ji}))^{-1} w_{ij}))^{-1} \tag{1}$$

Also the output value Y is calculated by using the same method.

In calibration process, researchers usually use some designed points such as the calibration points[12]. Sometimes this results in mistaken calibration guessing that experiments be repeated. Specifically, users performing experiments several times move their gaze to the next prospective point before previous calibration finishes. To give calibration greater universality and reduce the possibility of users anticipating subsequent calibration process, our experiments used random calibration, i.e., positions of individual points are given randomly, to eliminate user anticipation. The set consisting of all points must cover the whole screen, as shown in Fig.7.

3 Intention Recognition

Intention recognition is a task of recognizing the intentions of a human or an agent. The task usually achieved by analyzing some or all the actions or changes

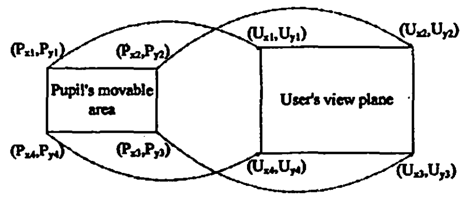


Figure 7: Coordinate Mapping between User View Plane and Pupil Move Area

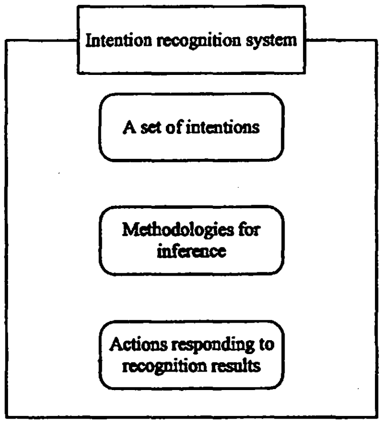


Figure 8: Structure of Intention Recognition System

of some features. The basic structure of an intention recognition system is usually composed of three,parts, as shown in Fig.8.

Specifically, it is realized by two steps: founding a set of features or actions and inferring based on a theory. In this research, for the first step, we found the initial set of intentions by object recognition in the image of robot vision and chose human's gaze as actions. For the second step, we inferred human intention based on fuzzy theory.

3.1 Object Recognition

Before intention processing, initial set of possible intentions must be founded. The set also should depend on situation at hand. We found the set by recognizing objects in image showing to user. Usually, objects are divided into two types, known objects and unknown ones. To obtain the intention set, one or some known objects are organized, which also means that the objects have been learning by system.

We used an autonomous humanoid robot NAO as the mainly component of object recognition and execution unit. It is developed by a French company, Aldebaran Robotics and it's a humanoid robot with many features and capabilities. To keep in proportion with the rest of its body, the NAO only has short arms. Its hands are controlled by one motor and therefore the fingers cannot be operated independently. NAO has 2 CMOS cameras each one provides an image of 640x480 and a frame rate of 30 fps. They are located with offset of 40 degrees

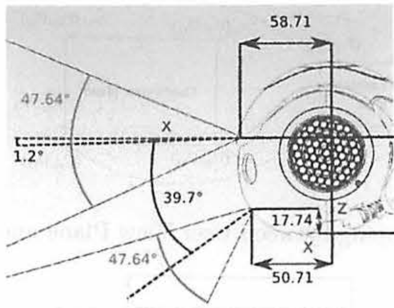


Figure 9: Field View of NAO[13]

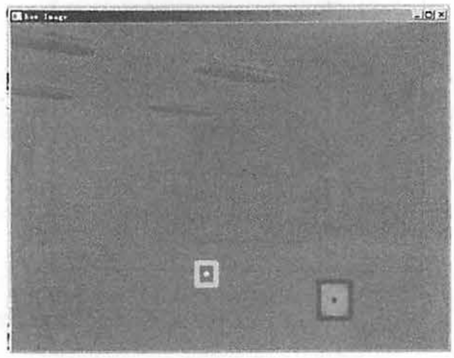


Figure 10: Example of Object Recognition

in the vertical axis. Before recognition of objects, NAO must learn images with object one by one at first. Here we use the top camera to capture image and store it in database in NAO by Choregraphe, graphical programming software, to create program without writing code. The field view of NAO is shown in Fig.9.

After learning object recognition[13], when an object appeared in the view field of NAO, position of object in NAO's coordinate system can be obtained. Feature points of object's contour can also be obtained at the same time. This process was achieved by using the ALVisionRecognition module from NAO SDK, which is a vision module in which NAO tries to recognize different pictures, objects sides or even locations learned previously. This module is based on the recognition of visual key points and is only intended to recognize specific objects that have been learned. An example of object recognition is shown in Fig.10. In this figure, there are two balls on the table, pink one and yellow one. When balls appeared in view field, the rectangle figure out the region and position information.

3.2 Intention Recognition by Fuzzy Inference

After object recognition, regions where objects exist are considered as intention region of user and ready for using by intention inference. In this research, initial

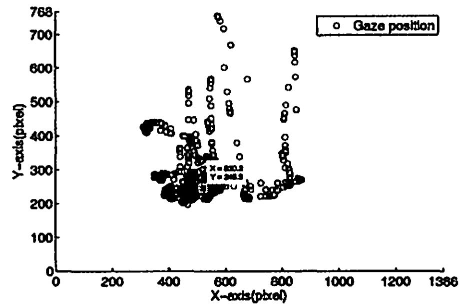


Figure 11: Example of Gaze Position Distribution

intention set is all regions' position where the objects recognized in last section.

By using the eye tracking device mentioned above, we analyzed gaze positions on computer screen in small period. And distribution of these points which stand for gaze positions is shown in Fig.11, where axis X and axis Y is 0~1366 and 0~768 respectively, correspond with the screen resolution. In Fig.11, we can see that at user's intention region, whose center is marked in the figure at (520.2, 245.5), the points' distribution density is high.

In fact, user's gaze will have a longer time staying and higher frequency appearing in the intention region than in not noticed region. Based on this, we used frequency and continuous staying time as factors for intention inference.

So far, researches on intention recognition are mainly divided into three classes according to the second component in Fig.8, which are logic-based, case-based and probabilistic approaches[14].

According to uncertainty of human's conceptual judgment and reasoning way of thinking, by using fuzzy sets and fuzzy rules in reasoning and making a fuzzy comprehensive judgment, we can solve complicated problems which are difficult for normal methods such as intention recognition. Thus, we applied fuzzy inference as method to recognize human's intention. Fuzzy inference is based on fuzzy logic and resembles human reasoning in its use of approximate information and uncertainty to generate decisions[12].

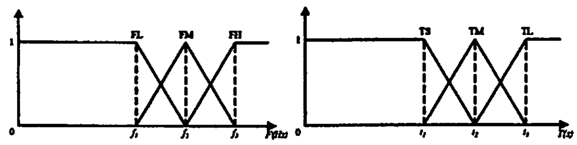
We used fuzzy rules to describe relationship between user's gaze and his/her intention. Fuzzy rule map is shown in Table 3. Fig.12 shows the membership functions and singletons. We used gaze appearing frequency F and gaze continuously staying time T in the regions got by object recognition part as inputs of fuzzy rules. Input space was divided into three subsets- TS, TM, TL, and FL, FM, FH- representing gaze continuously staying time short, intermediate, long and frequency low, intermediate, high, respectively. Probability values of individual object compose output space. Output space was divided into six subsets- PVL, PL, PM, PLH, PH and PVH- representing very low, low, intermediate, slightly high, high, and very high.

4 Experimental Results

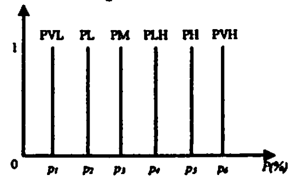
Experiments of proposed method were conducted on a notebook computer with Intel Core i3-380M CPU, 2 GB RAM and Microsoft Windows 7 operating system. Program was developed in Code::Blocks which is an open source IDE. Part

Table 3: Fuzzy Rule Map Used in Intention Recognition

T \ F	FL	FM	FH
TS	PVL	PL	PLH
TM	PL	PM	PH
TL	PLH	PH	PVH



(a) Membership Functions in IF Part



(b) Singletons in THEN Part

Figure 12: Membership Functions and Singletons

of pupil center detection is achieved by using OpenCV, openFrameworks, and an open source C++ toolkit. NAO SDK was also used for getting images from robot's eye camera and controlling it by an external computer.

4.1 Eye Tracking Results

In experiments, we used 16 points as reference points and user stares at each point for 1.5 seconds. In this process, traditional method is that 16 points appeared in turn in a 4×4 grid which is so-called calibration points and user is asked to look at the points when they appear, as shown in Fig.13. Coordinates of user's eye gaze estimation position are recorded and used as mapping data together with the appearing position of calibration points. Order of points appears is according to an "s" type.

We found that a subject's behavior easily became set, however, after several experiments and caused a poor calibration result, as mentioned in Section 3.2.2. We therefore use a method by making standard points appear randomly and replacing the previous one. Similarly, when standard points appeared randomly, we recorded both standard and estimation position coordinates. Image acquisition speed was 60 FPS, so 1440 (1.5×60×16) points are used as neural network inputs and for output. Because the neural network here is a back propagation (BP) neural network, teaching signals also used the above coordinates. After 1000 trails in the calibration process, actual error matches desired error, 0.001, which is set in Table 2.

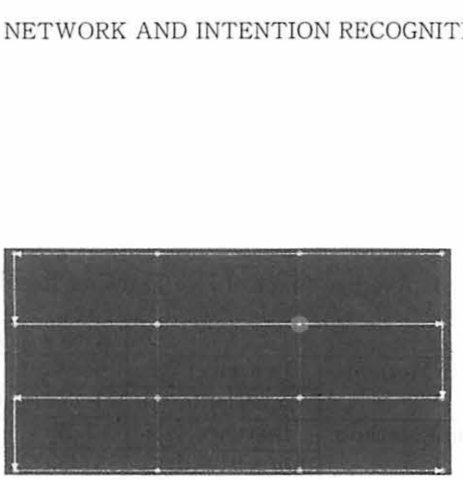


Figure 13: Position of Calibration Points

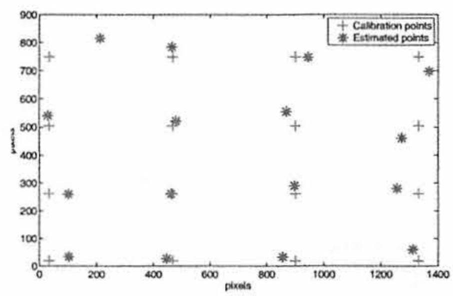


Figure 14: Results of Eye Tracking Verifying Experiment

After calibration, we carried out experiments to validate the calibration results. In experiments, 16 points were giving at first as reference points. Next, user was asked to look at each point in sequence. Position data of reference points and gaze average coordinates at each position would be recorded at the same time. Each estimated point and reference point are shown in Fig.14. Average error of results is shown in Table 4. The error of eye tracking results by traditional geometrical method is shown in Table 4. From the results in both two tables we can see that the method proposed by us worked effectively.

In experiments, distance between user's eye and computer screen is 45cm and resolution of screen is 1366×768 pixels. At the same time width and height of screen are 28.4cm and 21.3cm, respectively. The schematic diagram is shown in Fig.15. Thus, one pixel on computer screen stand for about 0.02cm. Assuming average error in one pixel is t, the degree error are calculated as follow equation.

$$\theta \approx \arctan(0.02t/45) \tag{2}$$

4.2 Intention Recognition Results

In this research, after a calibration process, user was asked to look at the scene image from NAO's camera and controlled NAO's head rotate by gaze. When an object or some objects are found, gaze frequency and continuous time in objects existing regions in a certain period were used as inputs of the fuzzy inference system which mentioned above.

Then after an intention recognition result is given, robot would get the object by using its hands. It is worth to note that there are three types of coordinate

Table 4: Average Error of Eye Tracking Results

		X-axis	Y-axis
Proposed Method	Distance(%)	3.24	3.68
	Direction(deg)	1.172	0.998
Traditional Method	Distance(%)	3.42	3.81
	Direction(deg)	1.253	1.074

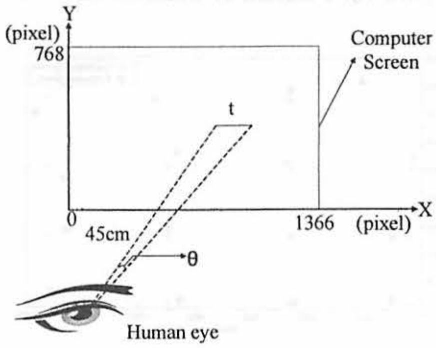
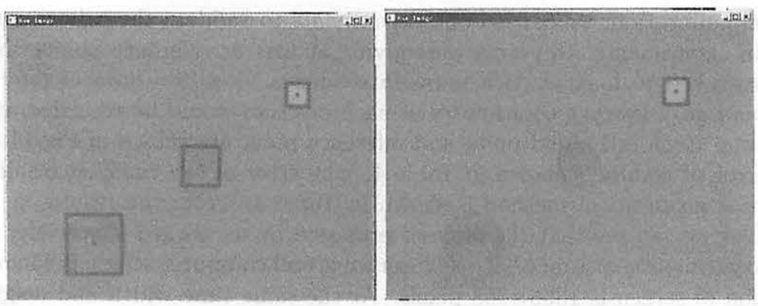


Figure 15: Calculate Method of Degree Error



(a) Object Recognition without Fuzzy Inference

(b) Intention Recognition with Fuzzy Inference

Figure 16: Intention Recognition by Fuzzy Inference

systems for NAO; FRAME_TORSO, FRAME_WORLD and FRAME_ROBOT. When creating a command for NAO, much attention needs to be placed on the space used to define the command. In this process the position of object gotten from NAO is according to FRAME_TORSO, also the same for operating hands to grasp an object. Here FRAME_TORSO is one of the 3 spatial references used by NAO's motion components.

Fig.16.(a) shows the various object recognition results in experiments. We can see that there are three objects have been recognized before intention inference. And Fig.17 shows the distribution of user's gaze positions when looking

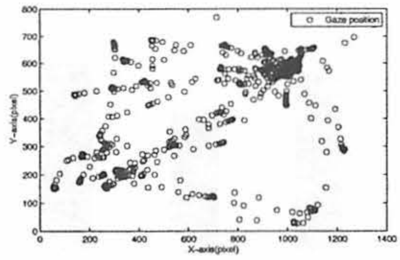


Figure 17: Distribution of User Gaze Position

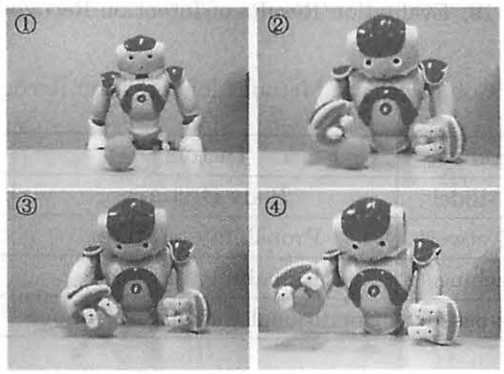


Figure 18: Experiment Scenes of Getting Object by NAO

at this scene. Output result of intention recognition by fuzzy inference is shown in Fig.16.(b). According to Fig.16.(b), user's intention is at the region where marked by a rectangle.

After intention recognition, experiments of getting object for user were conducted. Experiments scenes are shown in Fig.18.

We performed experiments aiming to verify the assistant system based on gaze estimation by NN and intention recognition by fuzzy inference work well with human. We also tried to find out whether subject can have a good impression in experiments. First, we asked subject to put on eye tracking device and carry on calibration process. Then scene image from NAO's camera was shown to subject on screen. According to the scene image, subject controlled NAO's head by gaze and searched in view field. When subject's intention is recognized, NAO would locate the object and get it for subject. At last, after the experiments, subject was asked to express his/her perceptions about the intention recognition result.

Evaluation results of intention recognition obtained is shown in Fig.19. Same experiments were executed by five subjects about twenty years old. After each time of experiments, subject was asked to evaluate at the result from three respects: accuracy of intention position, fitness of intention region contour and wait time before an object be found. We called the third factor as real time. And each factor of recognition has five ranks and represented from 1 (worst) to

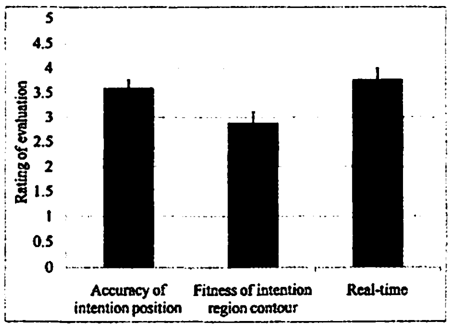


Figure 19: Evaluation Results of Intention Recognition

Table 5: Comparison of Methods for Intention Recognition

No.	Method	Factors	Reference
1	Markov Model	Body Gestures	[16], [17]
2	Bayesian Network	Probability of States	[18], [19], [20]
3	State Machines	Probability of States	[21], [22]
4	Ontology-based	Conditional Entropy of Actions	[23], [24], [25]

5 (best). Then we calculated the average value and standard deviation of the accuracy of intention position, fitness of intention region contour and real-time, respectively, based on the survey results.

From Fig.19, we can see that the system can achieve user's intention recognition basically, but the accuracy of recognition result is not always precise. This is because according to Chen's research by using structural equation modeling (SEM)[15], subjects' intention, especially behavior intention, is affected by destination image, trip quality, perceived value, satisfaction, behavioral intentions and so on. But in this research, we can consider that the gaze factor reflected behavioral intention and the object recognition reflected destination image factors. The two factors, but not the only ones can affect user's intention.

As the last of this research, we make a comparison between the typical intention recognition methods by using robot. The results are shown in Table 5. And from Table 5 we can see that related researches on intention recognition can be divided into two modes. One is based on conducting of user's states while the other realized by recognizing actions. For both two modes, users may have different intentions even they in a same state or doing the same action. But in the novel method proposed by us, we considered user's gaze position as a important factor for intention recognition, which is more directly and can minimizing the effect of environment.

5 Conclusion

We have designed a low-cost, wearable eye tracking system and have proposed new calibration using random calibration points and a neural network to replace conventional spatial mapping in the calibration process. In experiments, we have confirmed that after 1,000 learning trials using enough reference points, better calibration results have been obtained. We have also proposed the method establishing an initial set of intention based on object recognition results by gaze estimation used neural network from robot camera image. Then user's intention was inferred by fuzzy inference by using gaze frequency and continuous gaze staying time as input. At last, a humanoid robot could extract the object in the environment images and calculate the position in its coordinate system, finally get the object for user.

For the future research, We plan to work on enriching factors which affecting intention, such as user's interests, relationship between objects and saliency map of scene image to achieve a better intention recognition result and a more intelligent assistant system.

References

[1]Sadri F., Logic-Based Approaches to Intention Recognition, *Handbook of Research on Ambient Intelligence: Trends and Perspectives*, 2010.

[2]Pynadath D. V. and Wellman M. P., Accounting for Contextin Plan Recognition, with Application to Traffic Monitoring, *Proceedings of the Eleventh International Conference on Uncertainty in Artificial Intelligence*, 472-481, 1995.

[3]Pereira L. M. and Anh H. T., Intention Recognition via Causal Bayes Networks Plus Plan Generation, *Progress in Artificial Intelligence*, 138-149, 2009.

[4]Mao W. and Gratch J., A Utility-Based Approach to Intention Recognition, *Proceedings of the AAMAS 2004 Workshop on Agent Tracking: Modeling Other Agents from Observations*, 2004.

[5]Nickel K. and Karlsruhe U., Pointing gesture recognition based on 3D-tracking of face, hands and head orientation, *Proc. of the 5th International Conference on Multimodal Interfaces*, 140-146, 2003.

[6]Bonino D., Castellina E., Corno F. and Russis L. D., DOGeye: Controlling your home with eye interaction, *Interacting with Computers*, 29(2011), 484-498.

[7]Mitsugami I., Ukita N. and Kidode M., Robot Navigation by Eye Pointing, *Proc. Entertainment Computing*, 256-267, 2005.

[8]Akashi T., Wakasa Y., Tanaka K. and Fukumi M., Robust eye tracking for practical use, *Information*, 11(2008), 377-386.

[9]Ko Y. J., Lee E. C. and Park K. R., A robust gaze detection method by compensating for facial movements based on corneal specularities, *Pattern Recognition Letters*, 29(2008), 1474-1485.

[10]Tahboub K. A., Intelligent Human-Machine Interaction Based on Dynamic Bayesian Networks Probabilistic Intention Recognition, *Journal of Intelligent and Robotic Systems*, 45(2006), 31-52.

[11] Hansen D. W. and Ji Q., In the Eye of the Beholder: A Survey of Models for Eyes and Gaze, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(2008), 1474-1485.

[12] Harris J. W. and Stocker H., Handbook of Mathematics and Computational Science, *Springer-Verlag New York*, 1998.

[13] Documentation of Aldebaran Nao from Aldebaran Robotics -Documentation Version, 1.14.3.

[14] Lee W. O., Lee J.W., Park K. R., Lee E. C. and Whang M., Object recognition and selection method by gaze tracking and SURF algorithm, *2011 International Conference on Multimedia and Signal Processing*, 261-265, 2011.

[15] Chen C. F. and Tsai D. C., How destination image and evaluative factors affect behavioral intentions?, *Tourism Management*, 28(2007), 1115-1122.

[16] Zhu, C. Cheng Q. and Sheng W. H., Human intention recognition in Smart Assisted Living Systems using a Hierarchical Hidden Markov Model, *Automation Science and Engineering*, 253-258, 2008.

[17]Kelley R., Nicolescu M., Tavakkoli A., King C., and Bebis G., Understanding human intentions via hidden markov models in autonomous mobile robots, *The 3rd ACM/IEEE international conference on Human robot interaction*, 367-374, 2008.

[18]Gao J. X., Juan Z. C. and Gao L. J., Modeling and Simulation of Travel Behavior Forecasting System Based on Bayesian Network, *Information*, 15(2012), 5515-5522.

[19]Karim A. T., Intelligent Human-Machine Interaction Based on Dynamic Bayesian Networks Probabilistic Intention Recognition, *Journal of Intelligent and Robotic Systems*, 45(2006), 31-52.

[20]Luis M. P. and Han T. A., Intention Recognition via Causal Bayes Networks Plus Plan Generation, *Progress in Artificial Intelligence*, 138-149, 2009.

[21]Awais M., and Henrich D., Human-robot collaboration by intention recognition using probabilistic state machines, *Robotics in Alpe-Adria-Danube Region*, 75-80, 2010.

[22]Turaga P., Chellappa R., Subrahmanian V. S. and Udrea O., Machine Recognition of Human Activities: A Survey, *Circuits and Systems for Video Technology*, 18(2008), 1473-1488.

[23]Hochul J., Taehwan K. and Joongmin C., Ontology-Based User Intention Recognition for Proactive Planning of Intelligent Robot Behavior, *Multimedia and Ubiquitous Engineering*, 244-248, 2008.

[24]Yamaguchi T., Sato E. and Takama, Y., Intelligent space and human centered robotics , *Industrial Electronics*, 50(2003), 881-889.

[25]Craig S., Anthony P., Zeid K., Stephen B. and Sebti F., Ontology-based state representations for intention recognition in humanCrobot collaborative environments, *Ubiquitous Robotics*, 61(2013), 1224-1234.