CAAI Transactions on Intelligence Technology

Special Issue: Recent Advances in Intelligent Robotic Systems

Proposal of initiative service model for service robot

ISSN 2468-2322
Received on 13th September 2017
Revised on 6th November 2017
Accepted on 15th November 2017
doi: 10.1049/trit.2017.0021
www.ietdl.org

Man Hao[1,2], Weihua Cao[1,2], Min Wu[1,2], Zhentao Liu[1,2] ✉, Jinhua She[1,2,3], Luefeng Chen[1,2], Ri Zhang[1,2]

[1]*School of Automation, China University of Geosciences, Wuhan 430074, People's Republic of China*
[2]*Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, People's Republic of China*
[3]*School of Engineering, Tokyo University of Technology, Tokyo 192-0981, Japan*
✉ *E-mail: liuzhentao@cug.edu.cn*

Abstract: Service robots are becoming a part of today's life. To realise friendly human–robot interaction, an initiative service model for service robot is proposed, which is composed of three layers, i.e. information acquisition layer, in which communication information and environmental information are collected by sensors; decision layer, in which personal intention and emotional states are obtained by multi-modal emotion recognition and intention understanding, and users' demands are obtained by demand analysis model; initiative service execution layer, in which the task execution and emotion adaption of robot are carried out. Initiative service can sense motive of people, by which a humanised and real-time service could be realised. Drinking service experiments are performed, including passive service and initiative service of robots. Experimental results show that the drinking time of passive service lags behind the initiative service 30 min on average, and the initiative service renders drinking service 2 times more than passive service on average, besides, the water quantity provided by initiative service is more reasonable, which demonstrates the feasibility of the proposal. In prospect, initiative service model could be applied to many occasions in people's daily life, e.g. family service, caring for the elderly, and medical rehabilitation.

1 Introduction

Service robot is an important part of robot industry, and it is a key entry point to improve the way of human life. As the service robot stepping into all aspects of people's daily lives, e.g. guest service [1], family service [2], disabled pension [3–5], medical rehabilitation [6] etc., the simple and mechanical service of robots has been unable to meet the needs of people. It is hope that robots can be endowed with the human-like capabilities of observation, interpretation, and even emotion expression [7], i.e. robots have the ability to interact with people emotionally, including react appropriately to human expectations and behaviour. Hence intelligent service research has attracted more attention.

In recent years, there are many researches on service robot. The ambient assisted living service system based on scale-invariant feature transform (SIFT) algorithm that proposed in [8] can locate the objects who have service demands for elderly and disabled. A service architecture is proposed to solve the problem of heterogeneity in multi-robot cooperation [9]. A mobile robot is able to perceive the surrounding environment and interact with people in [10]. However, function of cognitive and understanding object's intention are still not enough to meet humans' requirements. The Sacario service robot system that is used in the hotel service can realise the human–robot interaction (HRI), including speech recognition, face location, and gesture recognition [11]. This system, however, lacks the perception of emotion in speech recognition, thus it cannot correctly understand the intention of users. The robot behaviour adaption and HRI researches have been done in [12, 13]. Although these robots are able to execute corresponding service tasks through instructions mentioned above, the service pattern is passive, which causes robots cannot integrate into people's life. If a robot can serve users initiatively, i.e. robots provide initiative service for human, a natural and harmonious HRI could be turned into reality.

To realise initiative service in HRI, communication information of human is necessary, which is divided into two parts, i.e. surface communication information and deep cognitive information [14, 15]. Surface communication information refers to the information that can be obtained directly by sensors or the combination of knowledge base, e.g. speech, facial expression, and gesture. Deep cognitive information is based on human cognitive ability and inference ability to understand the intricate physiological or psychological changes in human mind, e.g. emotional states, personal intention etc. Most of the service robots mentioned above serve for people based on surface communition information. They execute tasks mechanically according to the instruction, without understanding the thoughts and feelings of humans, (i.e. deep cognitive information). Deep cognitive information reflects people's inner thinking, which is an important basis for initiative service if we can make good use of them.

An initiative service model for service robot is proposed to solve the above problems, in which surface communication information and deep cognitive information are both taken into account. The initiative service model is composed of three layers. From bottom to top, the first one is information acquisition layer, in which surface communication information (e.g. speech, expression and gesture) and other influence factors (i.e. personal information and environment information) are collected by sensors. The second layer is decision layer, where the deep cognitive information (i.e. emotional state and personal intention) is obtained through emotion recognition and intention understanding. Furthermore, a demand analysis model is established to obtain a fuzzy relationship between deep cognitive information and users' demands, by which users' demands are obtained. The last one is initiative service execution layer, in which the task execution and emotion adaption is accomplished according to users' demands. To validate the proposal, drinking service experiments are performed, in which three men and two women within 20–25 years old are invited as the experimenters. At the same time, the drinking time and water quantity of experimenters at each time by initiative service and passive service

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)

are recorded, respectively. The experimental results validate that initiative service is real time and reasonable.

The remainder of this paper is organised as follows. The architecture of initiative service model is proposed in Section 2. Initiative service method is introduced in Section 3. Drinking service experiments using initiative service model are given in Section 4.

2 Initiative service architecture

In recent years, service pattern of service robot is mainly passive service pattern, in which robot carries out tasks when an instruction is received. This pattern is too mechanised to meet humanised demands. To solve the problem, an architecture of initiative service is proposed to provide a humanised and real-time service.

2.1 Introduction of passive service and initiative service

For passive service, service robots need to receive commands (e.g. semanteme) from the users, and then perform the corresponding tasks. For instance, the user tells the robot that he/she wants to drink water specifically, and then the robot brings him/her the water. The schematic of the passive service is shown in Fig. 1.

For initiative service, emotional states and personal intention are obtained by analysing the surface communication information (e.g. physiological signals and speech), which are used to estimate users' demand. According to this demand, the robot provides initiative service including the corresponding task execution and emotion adaption. For example, the robot obtains users' drinking demands through physical condition information and environmental information if the user is short of water, thereby providing water to the user initiatively and generating emotion adapting to the emotional state of the user. The schematic of the initiative service is shown in Fig. 2.

Initiative service is the spontaneous execution of the robot via a series of decisions, and it does not need to wait for users' instructions, which can make the service more humanised and effective.

2.2 Architecture of initiative service

Users' demands are essential to implement initiative service of robots. Usually, emotional states and personal intention are closely related to users' demands; therefore, emotion recognition and personal intention understanding are indispensable for initiative service, which could be obtained by surface communication information (e.g. facial expression, speech), personal information, and environmental information. Based on the above analysis, a three-layer framework for initiative service model is proposed, which is composed of three parts, i.e. information acquisition layer, decision layer, and initiative service execution layer, as shown in Fig. 3.

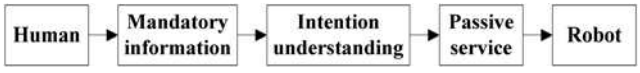


Fig. 1 *Schematic of passive service*

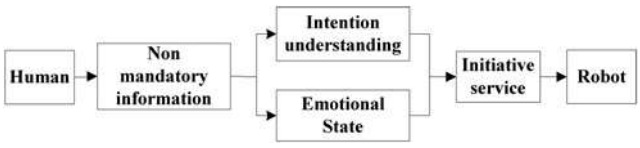


Fig. 2 *Schematic of initiative service*

In information acquisition layer, the information that is related to users' emotional states and personal intention is collected. In consideration of common sense of daily life, surface communication information (e.g. facial expression, speech, physiological signal), personal information, and environment factors which are highly associated with personal intention and emotion are obtained via sensors.

In decision layer, emotional states of the user are obtained through the identification of multi-modal information, and personal intention is inferred based on emotional states, personal information, and environmental information. In the demand analysis model, machine learning, fuzzy inference etc., could be used to get the relationship between deep cognitive information (i.e. personal intention and emotional states) and users' demands.

In initiative service execution layer, there are two kinds of services, i.e. task execution, in which appropriate task according to users' demands are executed; emotion adaption, in which appropriate emotion is generated adapting to the users' emotional state.

3 Demand analysis based on deep cognitive information

The premise of realising initiative service is to obtain users' demand, which is related to user's emotional state and personal intention. Firstly, emotional state is obtained by emotion recognition, including speech emotion recognition, facial expression recognition etc. Secondly, personal intention is inferred by intention understanding. Also then a demand analysis model is established, in which fuzzy inference is adopted to acquire users' demand.

3.1 Multimodal emotion recognition

Humans' emotion is usually divided into seven basic categories, i.e. neutral, happy, anger, surprise, fear, disgust and sadness [16]. In the communication between humans and robots, surface communication information (e.g. speech, facial expressions) can reflect humans' emotional state that is obtained by emotion recognition.

3.1.1 Speech recognition: As fast and easy understand way for communication, speech signal is considered to identify emotions

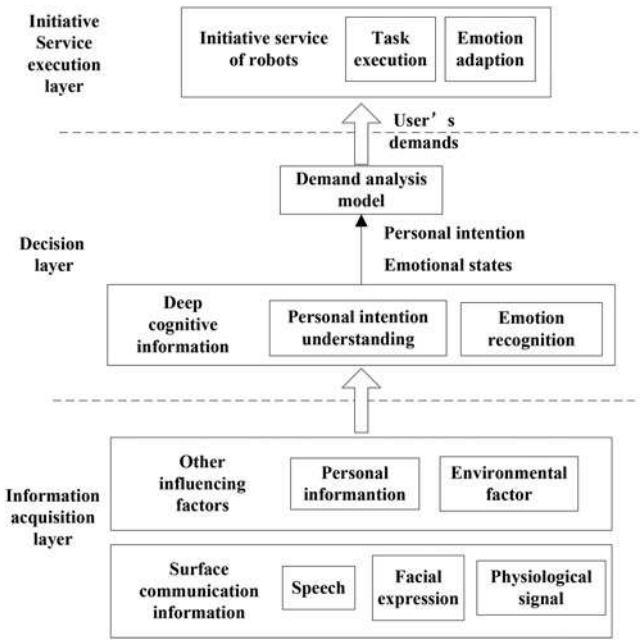


Fig. 3 *Initiative service architecture*

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)

[17]. It is believed that speech conveys not only syntactic and semantic contents of the linguistic sentences but also the emotional states of humans [18]. Speech signal can be collected by audio sensors, and the pre-processing of speech signal is adopted at the same time. Firstly, the emotion features (e.g. fundamental frequency, Mel-frequency Cepstral Coefficients and other statistic features) are extracted according to our previous works [19]. After obtaining the initially extracted features from pre-processed speech signal samples based on the analysis and extraction of speech emotion features, the correlation analysis is used to gain the lower redundancy features and reduce the dimension of selected feature set. Next, these selected features are put into pre-trained extreme learning machine (ELM) model, by which the speech emotion is recognised.

3.1.2 Facial expression recognition: Facial expression is the most expressive way to convey emotions, and facial expression contributes 55% to the effect of the speakers' message [20]. User's emotional state can be estimated through facial expression recognition. The procedure of facial expression recognition mainly consists of three steps [21], i.e. face detection and segmentation, in which Adaboost learning is adopted for detecting the face region and Haar classifiers are used to further detecting and segmenting salient facial expression patches which are closely related to facial expressions; feature extraction, in which 2D Gabor and local binary pattern (LBP) are combined to extract discriminative features of facial expressions, and principal component analysis is used for dimensionality reduction; facial expression classification, in which ELM is adopted for facial expression classification using the extracted features.

3.1.3 Fusion of multi-modal emotion recognition: To improve the accuracy of recognition results, the multi-modal emotion information is taken into account. There are two fusion layers for multi-modal emotion recognition [21]. The first one is feature level fusion, where the features from all modalities are transformed into one feature vector for emotion recognition. An ELM classifier is designed for the feature vector, and the final emotion result is obtained. The other is the decision level fusion, where Naive Bayes classifier is designed for every modality. The probability of each basic emotion estimated by single modality is multiplied to obtain the final probability. By the product rule, the emotion with the biggest probability is chosen as the final decision result of multi-modal fusion.

3.1.4 Emotional state represented in a 3D emotion space: A 3D emotion space named affinity-pleasure-arousal (APA) space is employed to represent emotional state of humans [22], in which emotional state and its dynamic changes can be presented intuitively, as shown in Fig. 4, being thought over the similar relationship between the three axes of emotion space, namely, 'Affinity', 'Pleasure–Displeasure' and 'Arousal–Sleep'. For instance, 'Affinity' relates to the mutual relationship among individuals, 'Pleasure–Displeasure' relates to favor, and 'Arousal–Sleep' relates to liveliness in communication. Moreover, by using APA emotion space, it is not only possible to represent fixed mentality states but also possible to take into account rapid variations in emotional states.

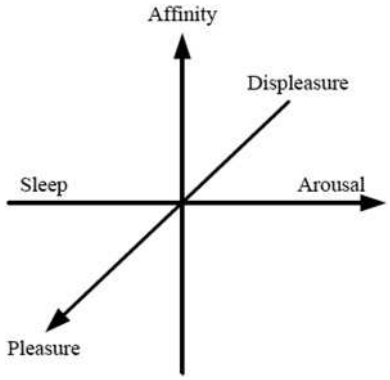


Fig. 4 *Affinity–Pleasure–Arousal emotion space*

3.2 Personal intention understanding

Intention understanding is mainly obtained by emotion, and intention is understood by collecting and analysing of multimodal information (e.g. speech, expression and gesture). In addition, personal intentions can also be affected by personal information and surrounding environment information. Personal information (e.g. the gender and exercise) are obtained through the sensor detection and the knowledge base. Moreover the surrounding environment information (e.g. temperature, humidity etc.) are obtained directly through the sensor.

An intention understanding model based on two layers fuzzy support vector regression is adopted to comprehend humans' inner thoughts in our previous work [14], which includes two parts, i.e. local learning layer and global learning layer. In the local learning layer, fuzzy C-means (FCM) clustering is used to classify the input data, and then the training data set are split into several subsets using the FCM algorithm. Therefore, the centre and spread width is obtained. After the local learning, the global output of the proposed two-layer fuzzy support vector regression is based on the fusion model, which is calculated using fuzzy weighted average algorithm.

3.3 Task execution and emotion adaption

It is difficult to describe users' demands with accurate mathematical model, which belongs to the empirical model. Therefore, fuzzy production rule is adopted in this paper. Firstly, personal intention and emotional state of humans are fuzzified. Secondly, the relationship between deep cognitive information (i.e. emotional states and personal intentions) and users' demands are established. In addition users' demands are obtained by fuzzy inference. Thirdly, robots' tasks and emotion are defuzzified, and execute the corresponding task execution and emotion adaption. The model of fuzzy production rules is shown in Fig. 5.

3.3.1 Fuzzification: There are many kinds of intention of users when robot service for people. When the intention is estimated, the intention level (IL) of user's desire is divide into four levels, i.e. urgent (U), normal (N), slight (S), and needless (NL).

Emotional state of humans is usually divided into seven categories, i.e. happiness, surprise, fear, neutral, disgust, sadness, and anger. For each categories, the emotional state level EL can be divided into five levels, i.e. very strong (VS), strong (S), normal (N), weak (W), very weak (VW).

The level of robot's tasks and emotional states are fuzzified as IL (i.e. U, N, S, NL) and EL (i.e. VS, S, N, W, VW), respectively, which are the same as humans in order to achieve smooth communication between humans and robots.

3.3.2 Fuzzy inference and defuzzication: The intention and emotional state of robots can be obtained by using IF–THEN rules

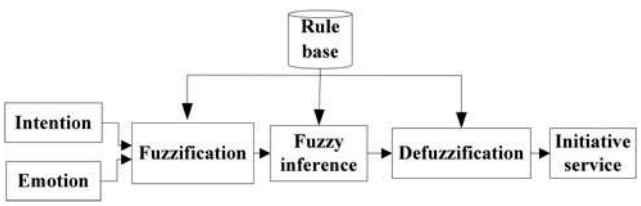


Fig. 5 *Fuzzy production rules model*

CAAI Trans. Intell. Technol., 2017, Vol. 2, Iss. 4, pp. 148–153
This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)

shown in (1) on the basis of fuzzification

$$\begin{aligned} &\text{If Humans} = \{\text{IL and EL}\} \\ &\text{Then Robots} = \{\text{IL and EL}\} \end{aligned} \tag{1}$$

The initiative service of robots includes two kinds of service, i.e. task execution and emotion adaption.

For task execution, it will generate the corresponding intention to perform a task when it receives users' demands, and according to the users' desire level of demands, the corresponding task execution (TE) and the rules are expressed in (2). For instance, TE can be expressed as the water quantity in drinking service is large amount, normal, small amount and none

$$\begin{aligned} &\text{If Robot} = \{\text{IL}\} \\ &\text{Then TE} \end{aligned} \tag{2}$$

For emotion adaption, the three-dimensional emotion space is used to generating the robot's emotion to maintain consistency between robots and humans. It can generate its own emotional state to adapt to users according to the emotional state of users. The emotion adaption of robots (EA) is expressed through the screen, mechanical arm, and other executive agencies, in which each emotional state is divided as very strong (VS), strong (S), normal (N), weak (W), very weak (VW). Also the rules are expressed as

$$\begin{aligned} &\text{If Robot} = \{\text{EL}\} \\ &\text{Then EA} \end{aligned} \tag{3}$$

4 Experiments on initiative service model for drinking service

In our daily life, there is a significant relationship between people's drinking habits and the health of the body. In this paper, experiments on drinking service are performed to validate the proposal. In the scenario of drinking service, how often should people drink is an important problem. The goal of the initiative service model is to ensure that people can drink water at the right time and supplement enough water.

4.1 Drinking water demand analysis

People often forget to drink water because of tedious daily works, and they only drink water when they feel thirsty. However, the body has been in a state of lack of water when they feel thirsty. People lose fluid so rapidly that the brain cannot respond in time. Therefore, it is hoped that the robot can provide water initiatively according to the demands of users to ensure the health of people.

In addition to drinking time, water quantity is another important thing in drinking habits. When you drink too much water, you may develop a condition known as hyponatremia, in which the excess water floods your body's cells, causing them to swell up. Lack of water, or dehydration inversely, reduces the amount of blood in your body, forcing your heart to pump harder in order to deliver oxygen-bearing cells to your muscles. The amount of daily drinking water of an adult is 1500–2000 ml according to a survey. Also the quantity of each time is ~100–400 ml. Furthermore, factors such as speech, physiological signal, temperature, humidity, and gender affect people's drinking water requirements [23].

Usually, talking makes human short of water, and we want to drink water when talking for more than 20 min continuously ordinarily. Thus, it hints that users need to drink water when they speak a lot or the voice becomes hoarse. The length of speech can be obtained by recording equipment and timers. Firstly, a fixed threshold is set, and then two sound acquisition units are set up. Recording equipment starts recording when the intensity ratio of two sound signals is larger than the threshold, and stops when the ratio is less than the threshold. At the same time, the length of speech is acquired by timer during recording. In addition, the frequency of sound decreases when voice becomes hoarse. The collected speech set is processed to realise frequency spectrum analysis, in which the change of frequency is obtained to determine the change of tone.

The body will consume a lot of water after exercise; therefore, exercise is an important factor affecting people's drinking water demand. The common measures of exercise intensity are oxygen intakes and heart rate. Heart rate is easy to measure with a small instrument, while the measurement of oxygen intake is not easy to get since it needs expertise and a large apparatus [24]. Thus, heart rate is adopted to calculate users' exercise intensity in this paper, which is obtained by sphygmomanometer.

In addition, environmental factors also affect people's drinking water demand. For example, the amount of water will increase in demand when the temperature rises or humidity decreases. For

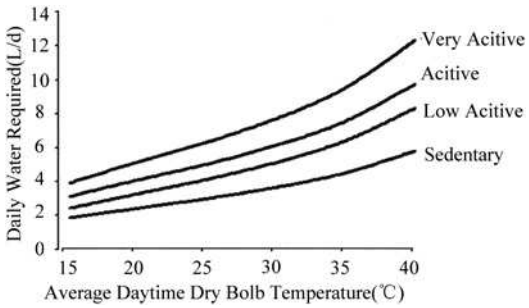


Fig. 6 *Relationship between temperature and drinking need [25]*

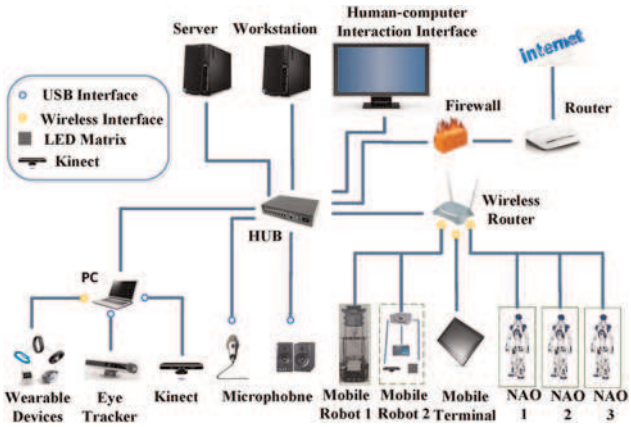


Fig. 7 *Topology structure of MEC–HRI system*



Fig. 8 *Scenario of 'drinking service'*

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)

Table 1 Activities of experimenters

Time	8:30	9:40	10:40–14:00	14:40	17:30	19:45	22:00
activities/day	start	discussion	temperature increment	normal activity	exercise	normal activity	finish

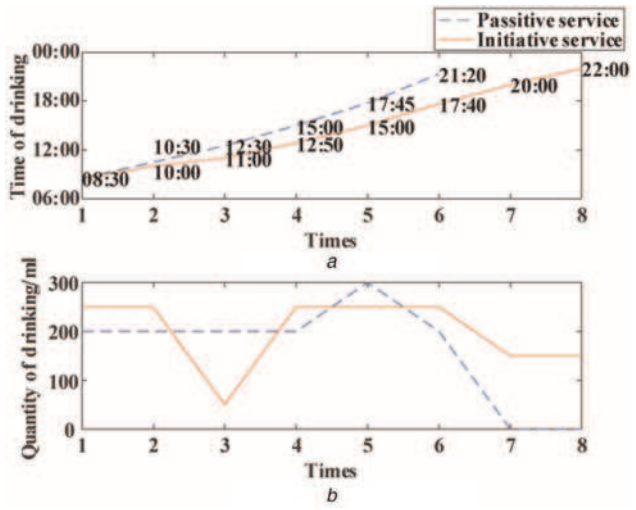


Fig. 9 *Experimental results of drinking service*

a Time of drinking
b Quantity of drinking

example, daily water requirements for any given energy expenditure in temperate climates (20°C) can triple in very hot weather (40°C), as shown in Fig. 6 [25]. In addition to air temperature, other environmental factors also modify sweat losses, including relative humidity, air motion, solar load, and choice of clothing for protection against sweat losses. Moreover, the personal information (e.g. gender and age) affects drinking demands as well. A study from Maastricht University in the Netherlands found that women lose more water during exercise than men [25].

4.2 Experiments on drinking service

Experiments are performed using the multi-modal emotional communication based human–robot interaction (MEC–HRI) system [21], which consists of three NAO robots, two mobile robots, Kinect, eye tracker, two high-performance computers (i.e. a server and a workstation), and other wearable sensing devices as well as data transmission and network-connected devices. The topology structure of MEC–HRI system is shown in Fig. 7. Also we selected three men and two women who are within 20–25 years old as the experimenters. In the experiment, some foundational data, such as humidity, temperature, and physiological data are collected under our supervision. If there is a big mistake in acquiring information such as physical condition and environment, we will intervene in the data collection until it returns to normal. Thus, the reliability of the collected data is guaranteed by our supervision. The scenario of drinking service is shown in Fig. 8.

People often forget to drink water because of busy work, and they may drink water when they feel thirsty. However, when people feel thirsty, it may have been a long time since the body has been lack of water. To solve the above problems and ensure human health, firstly, the factors that influence humans' drinking water demands are selected after in-depth analysis, including physical activity, physical condition, and environment. Secondly, drinking time and water quantity are inferred according to correlation analysis between these factors and water demand. Thus, robots can provide water to users timely and moderately even if users are unaware of their lack of water or forget to drink water. Consequently, initiative service can ensure a real-time and rationality drinking service.

User's information (e.g. speech, physiological, signal, and gender) and environmental information (e.g. temperature and humidity) are captured by the underlying hardware including Kinect, eye tracker, microphone, and other wearable devices, respectively. Then emotional state and personal intention are obtained by emotion recognition and intention understanding, respectively, and users' water demand is obtained by the demand analysis model.

When the intention of drinking is estimated, the intention level can be divided as urgent (U), normal (N), slight (S), and needless (NL) and the emotion state level of human is divided as very strong (VS), strong (S), normal (N), weak (W), very weak (VW). According to Section 3.3, the water quantity provided by the robot is divided into four levels correspondingly, i.e. large amount (400 ml), normal (300 ml), small amount (100 ml) and none (0 ml). The emotional state level of robot is divided as VS, S, N, W, and VW.

4.3 Experimental results and discussion

Experimental results are recorded, including activities from five experimenters, drinking time and water quantity at each time, as shown in Table 1. The curve of drinking time that robots serve the user initiatively is shown as the solid line in Fig. 9*a*. The time when experimenter feels thirsty and orders robot to provide water is shown as the dotted line in Fig. 9*a*. The curves of water quantity supplied initiatively to the user each time are shown as the solid line in Fig. 9*b*. The water quantity that experimenter orders robot to provide is shown as the dotted line in Fig. 9*b*.

In the experiments, two experimenters began to discuss at 9:40, and after a period of discussion, robot recorded this discussion time has reached 20 min, then the robot renders a drinking service at 10:00. However, experimenters were busy discussing until they felt thirsty and drank water at 10:30. As time goes by, temperature rose after 11:00, robot provides normal amount of water to experimenters after the temperature reached 30°C. Air-condition is used in our laboratory due to the heat in summer, which makes air dry, thus robot provided water to experimenters at 12:50 and 14:00, while experimenters felt thirsty and drank a lot of water at 12:00 and 14:00 in fact. Then experimenters ran for 30 min during 19:00–19:30, and heart rate of the experimenter was transmitted to robot, from which robots judged that the experimenters have a strong exercise and served experimenters drink a large amount of water at 19:30, since their body had lost a lot of water due to sweat losses.

From the experimental results, the time of passive service lags behind the initiative service 30 min on average; the initiative service renders two times water supply more than passive service on average; in initiative service, the water quantity for each time is decided according to the loss of water in the body, which is more reasonable.

It can be seen that the drinking time of initiative service and passive service is consistent when the water loss of the body is serious, for example, the drinking time and water quantity provided by initiative service is the same as passive service after exercise, but it is not be noticed by experimenters when the body drains a small amount of water. Experimenters do not drink water timely when they do not realise that they are short of water, which lead to the difference of water supplying time between initiative service and passive service become bigger and bigger.

CAAI Trans. Intell. Technol., 2017, Vol. 2, Iss. 4, pp. 148–153
This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)

5 Conclusion

An initiative service model for service robot is proposed. Firstly, multi-modal information (e.g. speech, expression and gesture) is collected by the underlying hardware (e.g. Kinect, eye tracker, and microphone). Also deep cognitive information (i.e. emotional state and personal intention) is acquired by using emotion recognition and intention understanding. Users' demands are obtained by a demand analysis model using deep cognitive information. Finally, robot executes corresponding task and realises emotion adaptation. By using this model, robots could understand actual demands of people in such an initiative and humanised way and provide corresponding service correctly.

Drinking service experiments are performed by three men and two women aged from 20 to 25 years old. For initiative service, robots can render drinking service for the user in a timely manner. Experimental results show that the time of passive service lags behind the initiative service 30 min on average and the water supply in initiative service is two times more than passive service on average, which avoids the situation that the body is short of water. In addition, by initiative service, water quantity for each time is more reasonable, which ensures that the user can supplement enough water.

Initiative service of service robot is a new concept, which promotes the realisation of intelligence and humanisation. Initiative service model can be applied to many scenarios in our daily life, e.g. family service, caring for the elderly, medical rehabilitation.

6 Acknowledgment

This work was supported by the National Natural Science Foundation of China under Grants nos. 61403422 and 61273102, the Hubei Provincial Natural Science Foundation of China under Grant no. 2015CFA010, the Wuhan Science and Technology Project under Grant no. 2017010201010133, the 111 project under Grant B17040, and the Fundamental Research Funds for National University, China University of Geosciences (Wuhan).

7 References

[1] Nuovo, A.D., Acampora, G., Schlesinger, M.: 'Guest editorial cognitive agents and robots for human-centered systems', *IEEE Trans. Cog. Dev. Syst.*, 2017, **9**, (1), pp. 1–4

[2] Zhou, F., Wang, G., Tian, G., *et al.*: 'A fast navigation method for service robots in the family environment', *J. Inf. Comput. Sci.*, 2012, **9**, (12), pp. 3501–3508

[3] Bedaf, S., Gelderblom, G.J., Syrdal, D.S., *et al.*: 'Which activities threaten independent living of elderly when becoming problematic: inspiration for meaningful service robot functionality', *Disabil. Rehabil. Assist. Technol.*, 2014, **9**, (6), pp. 445–452

[4] Jiang, J., Dawei, T.U.: 'Human-robot collaboration navigation of service robots for the elderly and disabled in an intelligent space', *CAAI Trans. Intell. Syst.*, 2014, **9**, (5), pp. 560–568

[5] Xu, S., Tu, D., He, Y., *et al.*: 'ACT-R-typed human–robot collaboration mechanism for elderly and disabled assistance', *Robotica.*, 2014, **32**, (5), pp. 711–721

[6] Cheng, X., Zhou, Z., Zuo, C., *et al.*: 'Design of an upper limb rehabilitation robot based on medical theory', *Procedia Eng.*, 2011, **15**, (1), pp. 688–692

[7] Iengo, S., Origlia, A., Staffa, M., *et al.*: 'Attentional and emotional regulation in human–robot interaction'. Proc. of 21st IEEE Int. Symp. on Robot and Human Interactive Communication, Paris, France, 2012, pp. 1135–1140

[8] Hendrich, N., Bistry, H., Zhang, J.: 'Architecture and software design for a service robot in an elderly-care scenario', *Engineering*, 2015, **1**, (1), pp. 27–35

[9] Cai, Y., Tang, Z., Ding, Y., *et al.*: 'Theory and application of multi-robot service-oriented architecture', *IEEE/CAA J. Autom. Sinica*, 2016, **3**, (1), pp. 15–25

[10] Hermann, A., Sun, J., Xue, Z., *et al.*: 'Hardware and software architecture of the bimanual mobile manipulation robot HoLLiE and its actuated upper body'. Proc. of IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics, San Diego, 2013, pp. 286–292

[11] Pinillos, R., Marcos, S., Feliz, R., *et al.*: 'Long-term assessment of a service robot in a hotel environment', *Robot. Auton. Syst.*, 2016, **79**, pp. 40–57

[12] Chen, L.F., Liu, Z.T., Wu, M., *et al.*: 'Multi-robot behavior adaptation to humans' intention in human–robot interaction using information-driven fuzzy friend-Q learning', *J. Adv. Comput. Intell. Intell. Inform.*, 2015, **19**, pp. 173–184

[13] Wang, X., Yuan, C.: 'Recent advances on human-computer dialogues', *CAAI Trans. Intell. Technol.*, 2016, **1**, (4), pp. 303–312

[14] Chen, L.F., Liu, Z.T., Wu, M., *et al.*: 'Emotion-age-gender-nationality based intention understanding in human–robot interaction using two-layer fuzzy support vector regression', *Int. J. Soc. Robot.*, 2015, **7**, (5), pp. 709–729

[15] Schuller, B., Batliner, A., Steidl, S., *et al.*: 'Recognizing realistic emotions and affect in speech: state of the art and lessons learnt from the first challenge', *Speech Commun.*, 2011, **53**, (9), pp. 1062–1087

[16] Sandbach, G., Zafeiriou, S., Pantic, M., *et al.*: 'Static and dynamic 3D facial expression recognition: a comprehensive survey', *Image Vis. Comput.*, 2012, **30**, (10), pp. 683–697

[17] Ayadi, M.E., Kamel, M.S., Karray, F.: 'Survey on speech emotion recognition: features, classification schemes, and databases', *Elsevier Sci. Inc*, 2011, **44**, (3), pp. 572–587

[18] Anagnostopoulos, C.N., Iliou, T., Giannoukos, I.: 'Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011', *Artif. Intell. Rev.*, 2015, **43**, (2), pp. 155–177

[19] Liu, Z.T., Li, K., Li, D.Y., *et al.*: 'Emotional feature selection of speaker-independent speech based on correlation analysis and fisher'. Proc. of the 34th Chinese Control Conf., Hangzhou, China, 2015, pp. 3780–3784

[20] Mehrabian, A.: 'Communication without words', *Psychology Today*, 1968, **2**, (4), pp. 53–56

[21] Liu, Z.T., Pan, F.F., Wu, M., *et al.*: 'A multimodal emotional communication based humans-robots interaction system'. Proc. of the 35th Chinese Control Conf., Chengdu, China, 2016, pp. 6363–6368

[22] Hirota, K.: 'Fuzzy inference based mentality expression for eye robot in affinity pleasure-arousal space', *J. Adv. Comput. Intell. Intell. Inform.*, 2008, **12**, (3), pp. 304–313

[23] Hao, M., Cao, W.H., Wu, M., *et al.*: 'An initiative service method based on intention understanding for drinking service robot'. IECON, Beijing, China, 2017

[24] She, J., Nakamura, H., Makino, K.: 'Selection of suitable maximum-heart-rate formulas for use with karvonen formula to calculate exercise intensity', *Int. J. Autom. Comput.*, 2015, **12**, (1), pp. 62–69

[25] Sawka, M.N., Samuel, N., Cheuvront, P.R.: 'Robert carter III PhD, MPH, human water needs', *Nutr. Rev.*, 2005, **63**, (s1), pp. s30–s39

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)