Intelligent Interactive Robot System for Agricultural Knowledge Popularity and Achievements Display

Xiaowen Sun[12], Ran Zhao[2], Abdul Mateen Khattak[2,4] ,Kaite Shi[12], Yanzhao Ren[1,3], Wanlin Gao[1,2*], Minjuan Wang[1,2*]
1.Key Laboratory of Agricultural Informatization Standardization, Ministry of Agriculture and Rural Affairs
Beijing, China
2. College of Information and Electrical Engineering, China Agricultural University
3.College of science, China Agricultural University
4. Department of Horticulture, The University of Agriculture, Peshawar, Pakistan
xiaowen0505@126.com, ran.zhao@cau.edu.cn, mateen@aup.edu.pk, skate@cau.edu.cn, xiaozhaochina@163.com,
wanlin_cau@163.com, minjuan@cau.edu.cn

Abstract—**The development of a guide robot with explanatory features has always been a subject of concern. Most urban populations, especially the youth, have little knowledge about agricultural activities. To address this issue, many exhibition centers featuring the theme of local agriculture were built. These agricultural exhibitions provide opportunities to extend agricultural knowledge to urban communities. However, the exhibition centers also need efficient ways of communication for introducing visitors to the exhibitions' achievements. In this regard, we designed an intelligent interactive robot system. For this purpose, we first collected agricultural knowledge to develop custom database questions and answers and made corresponding images. Then, designed a new interactive way through combining the questions and answers with the image display. Finally, two interactive modules were designed for people of two ages (youth and adults) attending the exhibition. The developed robot system was used in "The first china international smart agriculture achievements exhibition" for the first time. The feedback results showed that the robot system helped the visitors learn more agricultural knowledge. Moreover, this robot system provided a more comprehensive way of understanding the exhibition for the visitors.**

Keywords—voice interaction; human-robot interaction (HRI); interaction design; voice recognition; agricultural knowledge; explanation of the exhibition hall

I. INTRODUCTION

The colossal urbanization during recent decades has diverted people from farming to other professions. People living in cities have little knowledge about agriculture. They do not know exactly where the food, that they consume every day, comes from and how is it grown? In this regard, a need was felt to create awareness among the city folks concerning the different techniques and procedures involved in agriculture and food production. For this purpose, many exhibition halls with the theme of local agricultural features were built. These exhibition centers display and promote different existing and new innovative achievements in field of agricultural alongside others. In this connection, the use of information technology is an effective way to display and promote these important commercial agricultural achievements.

Since the past decade, various human-computer interactive robots have been studied. Some of these include shopping mall welcoming robots, intelligent sound robots [1] and educational robots [2]. They greatly facilitate people's lives. Several studies suggest that human-computer interactive robots have the potential to help people gain knowledge [3]. Based on Nao [4], a guide robot with Chinese and English voice recognition and expression was developed by Hu et al. [5]. This robot interacts with visitors and guides them in the exhibition hall. Further, a secondary development based on Nao was commenced to explore whether social robots could assist the instructors' work [6]. Pandey et al. [7] took a social robot named Pepper to a shopping mall for one day to see the different initial responses it could acquire from people. Karat et al. [8] assessed the use of humanoid robots in services based on human interaction. However, the high prices of both the Nao or Pepper make them difficult for use in agricultural exhibition centers. Besides, Community-based Question and Answering (CQA) systems have also become popular knowledge-sharing platforms, where users get answers for their questions [9]. These have received great attention both in the industry as well as academia. An agriculture researcher analyses the question routing problem in a CQA system named Farm-Doctor that is exclusively for agricultural knowledge [10]. It can answer farmers' specific professional questions, but is not suitable for the popularization of agricultural knowledge. So far there have been no reports on the use of robotics to solve such problems. This intensified the need to conduct further research on developing a robot system that could provide agricultural services.

In this article, such an intelligent new interactive robot system is proposed. The system mainly uses sound source positioning, voice interaction, and image recognition technology. For this purpose, custom question and answer databases were developed, because the existing interactive systems could not answer the agricultural knowledge problem very well. An image display database corresponding to the question and answer was also produced. Here, a combination of voice broadcasting and image display continuously attracts participatory attention. From the economic point of view, the proposed system would be more suitable for agricultural exhibition halls.

The remainder of this paper is organized as follows: second part introduces the architecture and components of the agricultural robot system; the third part presents the visitors' feedback from "The 10th China International Modern Agriculture Exhibition" regarding this agricultural robot system. The last part describes the results and discussion, conclusions, and suggestions for future work.

II. MATERIALS AND METHODS

A. Hardware composition and structure

A humanoid shell robot was developed (architecture is shown in Fig. 1). The robot comprised of LED light, monocular camera, capacitive touch screen, Baidu audio, STM32 controlled mobile chassis, IFLYTEK microphone array and Intel's industrial computer.

The mentioned hardware performed the following functions:

LED light: Indicator as robot's working state;

Monocular camera: Acquired user image information;

Capacitive touch screen: Displayed custom image and got user interactive information;

Baidu audio: Voice broadcast;

STM32 controlled mobile chassis: Controlled the movement of the robot;

IFLYTEK microphone array: Six wheat XFM10621 for voice information entry and sound source positioning;

Intel's industrial computer: Total control platform of the robot, controlled the light on and off; processed voice information from the microphone array and image information from the monocular camera; while controlling the robot chassis according to sound source positioning information obtained by the microphone array.

B. System Architecture Overview

Because of the large number of people in the exhibition hall, both sound positioning and face recognition methods are used to determine whether to enter the question and answer module. At the same time, two interactive interfaces were designed according to the age of the visitors: professional interactive interface for adults and cartoon interactive interface for young. The designed system included modules such as sound source positioning, mobile platform, age recognition technology, and interactive interface (as shown in Fig. 2). The robot functions through a process of the following adjacent interconnected steps. First, the microphone array positions the questioner's distance and angle after receiving the awake word. Second, the camera captures the questioner's image information as the robot moves in front of the questioner. Third, the robot identifies the questioner's age based on the sound and image information. Finally, judging which module the robot enters to work on according to the identification results. In the next subsections, the technologies used in this system are described in detail.

The system is based on a Model-View-Controller design, with the core system programed in windows with a custom C#/C++ application in Visual Studio 2015.

1) Sound source localization

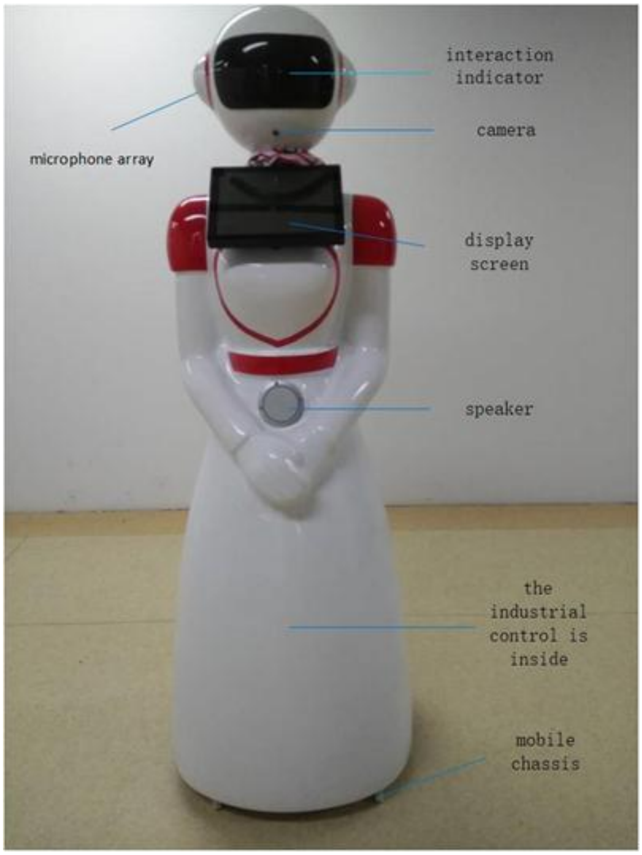


Fig. 1. Hardware composition and structure of the humanoid robot

The exhibition hall received many visitors on daily basis leading to noise in the venue. Sometimes the questioner is not necessarily standing in front of the robot. So, voice data and image information are doubled to achieve identification. So, it is especially important to move the robot towards the questioner. In this study, sound source localization technology was used to determine the location of the questioner.

The sound source positioning algorithm based on the microphone array is generally divided into three categories. The first is based on beamforming [11], the second is based on high-resolution spectral estimation, and the third is based on time delay of arrival (TDOA) method [12]. In the present work, the microphone [13] array with six wheat XFM10621 model was used. The model utilized 6 microphone arrays for 360 degrees of voice signal acquisition through the sound source position to determine the direction of the target speaker. This study only processed the angle information into four directions ({front}, {back}, {left}, {right}), and transferred controlled information to the STM32 under-bit machine.

2) Mobile platform

STM32 under-bit machine would receive the direction of motion from the main control machine to control the steering wheel to move the robot. By positioning the sound source multiple times, the robot would then move eventually to the front of the user. At this point, the monocular camera would get the user's facial information.

3) Age recognition technology

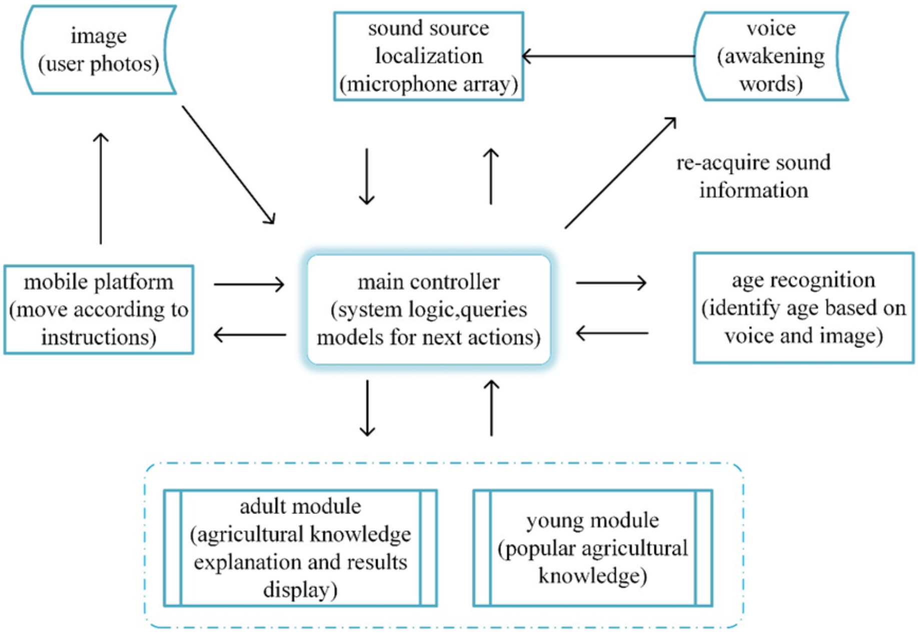


Fig. 2. System architecture diagram

Through the coordination of the mobile platform and sound source positioning, the robot was finally moved to the front of the user. At this point, the monocular camera got the user's facial information. The age of the user was determined by age recognition of voice data and image information to decide which interaction module would be selected.

In this section, we mainly introduce the key technologies of the system's age recognition. This scheme involved voice age recognition and image age recognition. The age judging logic in the system is shown in Fig. 3. If the voice and image recognition results indicated young, the young interaction module was entered. If the results showed all adults, the adult interaction module would be followed. Otherwise, the system would not enter an interactive module. At this point, the system needed to re-acquire the sound information.

a) Voice age recognition

Application Programming Interface (API) was provided by the IFLYTEK service platform AIUI database to recognize the speaker age. During the operation, the recognition module would send audio information captured by a microphone to the AIUI database. AIUI database recognized the age and gender attributes of voice information. The database returned data for gender (male, female) and age range (young, middle-aged, senior). Only the age range recognition results were used in the study, combining middle-aged and seniors into adults. Because the system had only two modules of interaction between young and adults.

b) Image age recognition

In this research, the BOW (bag-of-words) model [14] was used to classify the age of the questioners. The detail described as following:

Step 1: Using speeded up robust features (SURF) for feature point extraction;

Step 2: Using the K-means [15] algorithm for the clustering of feature points;

Step 3: Constructing a bag of words for the picture of the training set;

Step 4: Finally, using the SVM [16] algorithm to train a classifier of a second classifier.

Above this achieves the classification of the camera's image.

4) Interaction design

Normally, most visitors of the agricultural exhibitions are adults who come to visit and study themselves. However, some visitors also bring their children to increase their knowledge about agriculture. The adult visitors need professional explanations. They want quick answers to their questions and expect that the robots would help them understand the agricultural achievements on display in the pavilion. On the other hand, most of the young visitors, who grew up in cities, have no idea about agricultural production, because they have never been able to visit any orchards or farms. Given this situation, the database of agricultural knowledge popularization was designed. Moreover, a cartoon interactive interface for young people was designed. This cartoon interface was utilized to attract the attention of young generation.

The structure of user interaction is shown in Fig. 4. In this procedure, the robot would first identify user problems, then pass the result to the AIUI database. At this point, the AIUI database would perform fuzzy matching according to the problem. The match would return as audio and JavaScript Object Notation (JSON) text. The local database would match the local picture information according to the text information and display to the user via a capacitive touch screen. The local database would match the local picture according to the text information. This way, the picture was displayed to the user via a capacitive touch screen. Thus, the user experience and sense

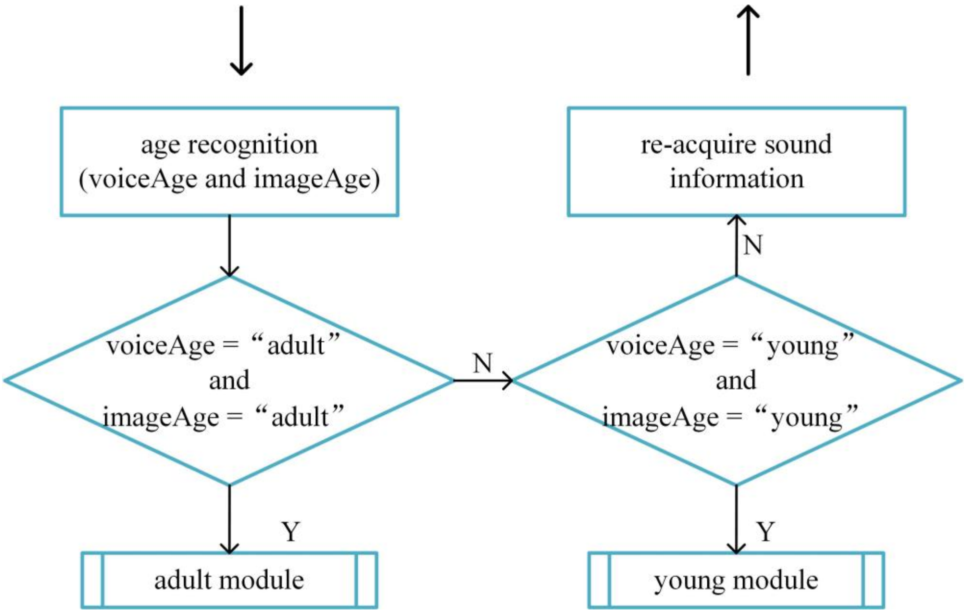


Fig. 3. Age judging logic

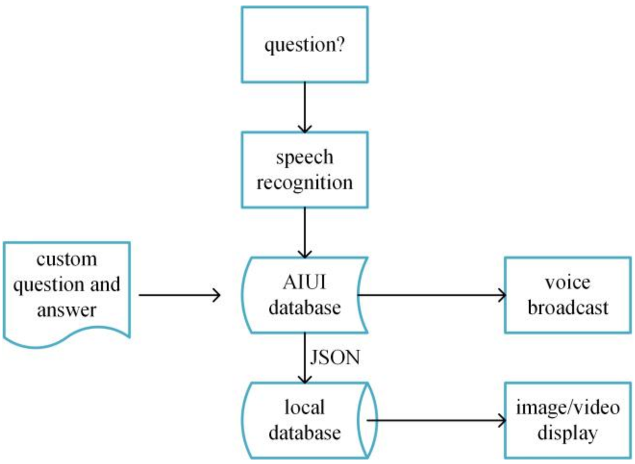


Fig. 4. Interactive structure diagram

of participation was improved through visual and auditory interaction at the same time.

a) Adult module

The adult module used a more professional interface with two main parts. One part collected and produced 100 agricultural knowledge and corresponding display pictures. Fig. 5 (a) shows the first five of the custom databases, while Fig. 5 (b) depicts the first corresponding picture display information. The other part collected information from the manufacturers in the exhibition hall and produced the corresponding display pictures.

b) Young module

A special cartoon interactive interface was designed for the young. For this, collected pictures of key points at the cultivation, growth, and transportation stages of fruits and vegetables in daily life. A total of 36 pictures of fruit or vegetables were produced, as showed in Fig. 6

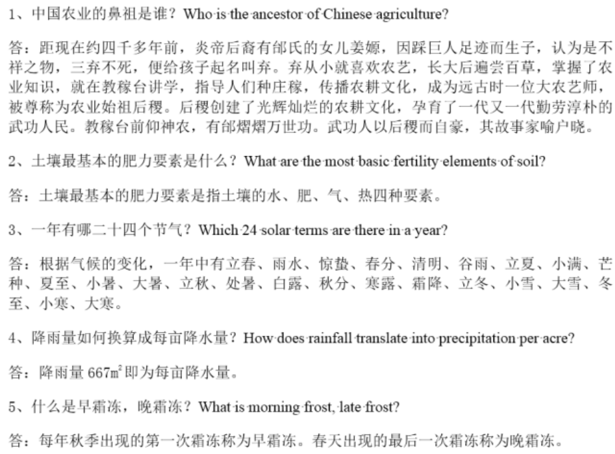
III. Results and Discussion

This robot system designed in this study was tested for the first time in “The first china international smart agriculture achievements exhibition”. According to the achievements presented at the exhibition, corresponding images and custom question-and-answer databases were made. To evaluate the effectiveness of this system more efficiently, 10 questions were formulated and enquired as shown in Table I. The feedback results showed that the robot system helped visitors learned agricultural knowledge in a better and effective way. Moreover, the visitors’ understanding of the exhibition became more comprehensive with use of this robot system.

A. Participants of the exhibition

The exhibition participants registration figures show that the proportion of young people attending the event was 16%. The proportion of middle-aged people was 81% and that of senior adults was 3% as shown in Fig. 7 (a). This study considered people under the age of 16 as young, people between 16 and 60 as middle-aged, and those over 60 as senior adults. Although they were categorized into two modules of interactions, i.e. young and adults, people over 60 were considered separate based on their better knowledge of agriculture.

Nearly 200 visitors interacted with the robots during the three-day exhibition. Their age composition is shown in Fig. 7 (b). Among them, 19 percent were young, 70 percent middle-aged and 11 percent senior adults. Out of those, 15 young, 30 middle-aged, and 5 senior adults were randomly selected. Thus, a total of 50 participants filled out the questionnaire provided in Table I.



1、中国农业的鼻祖是谁？Who is the ancestor of Chinese agriculture?

答：距现在约四千多年前，炎帝后裔有邰氏的女儿姜嫄，因踩巨人足迹而生子，认为是不祥之物，三弃不死，便给孩子起名叫弃。弃从小就喜欢农艺，长大后遍尝百草，掌握了农业知识，就在教稼台讲学，指导人们种庄稼，传播农耕文化，成为远古时一位大农艺师，被尊称为农业始祖后稷。后稷创建了光辉灿烂的农耕文化，孕育了一代又一代勤劳淳朴的武功人民。教稼台前仰神农，有邰熠熠万世功。武功人以后稷而自豪，其故事家喻户晓。

2、土壤最基本的肥力要素是什么？What are the most basic fertility elements of soil?

答：土壤最基本的肥力要素是指土壤的水、肥、气、热四种要素。

3、一年有哪二十四个节气？Which 24 solar terms are there in a year?

答：根据气候的变化，一年中有立春、雨水、惊蛰、春分、清明、谷雨、立夏、小满、芒种、夏至、小暑、大暑、立秋、处暑、白露、秋分、寒露、霜降、立冬、小雪、大雪、冬至、小寒、大寒。

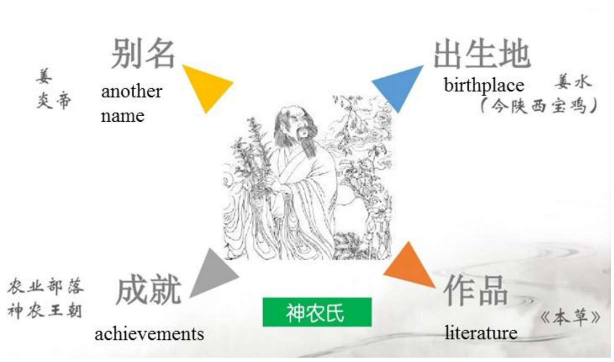
4、降雨量如何换算成每亩降水量？How does rainfall translate into precipitation per acre?

答：降雨量 667㎡ 即为每亩降水量。

5、什么是早霜冻，晚霜冻？What is morning frost, late frost?

答：每年秋季出现的第一次霜冻称为早霜冻。春天出现的最后一次霜冻称为晚霜冻。

(a)



(b)

Fig. 5. (a) Part of the question and answer content of the custom database; (b) Display image information corresponding to the first question

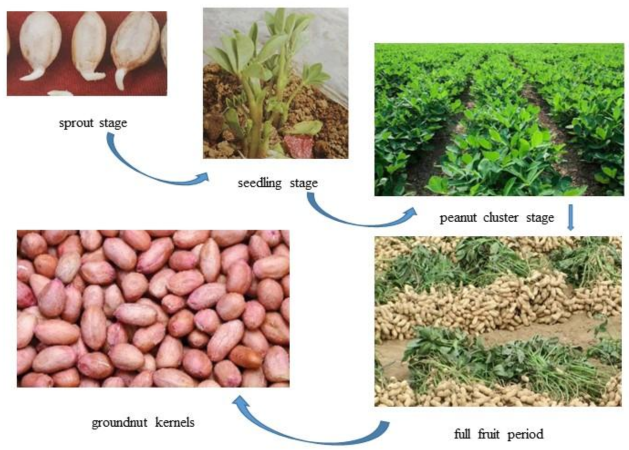
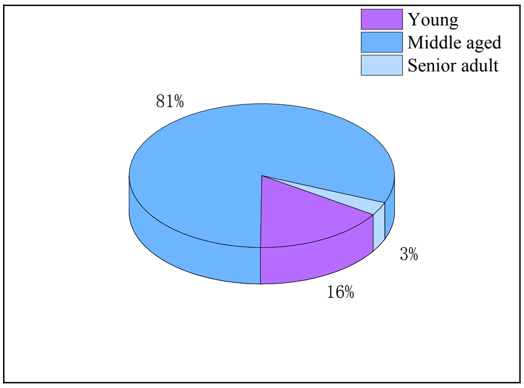


Fig. 6. An example of a peanut growth cycle picture in our database

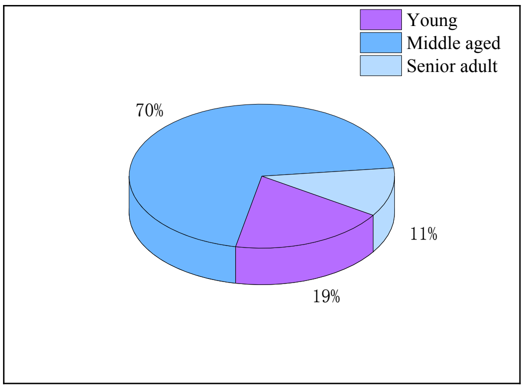
B. Evaluation of Responses

Questions A1 and A2 were designed to observe the participants knowledge of agriculture (see table I). The results of questions A1 and A2 of this experiment are shown in Fig. 8. Regarding question A1 (experience with robots), only 13 percent young participants had experienced similar robots before. In case of the middle-aged participants, 33 percent middle aged had encountered similar robots. Whereas, only 20 percent senior participants had seen similar robots. Concerning their exposure to farm activities (question A2), 80% of the young participants had never been to farmlands or orchards. Among the middle-aged participants, 77% had visited farmland or orchards, while all the senior participants had been to farmland or orchards.

On the first day of the exhibition, many people had no relevant interactive experience, especially young people with little knowledge of agriculture. In the beginning, they would not know how and what to interact with. So, on the first day the



(a)



(b)

Fig. 7. (a) The age composition of the people who attended the exhibition;(b) The age composition of the participants of this study

TABLE I. A SAMPLE OF THE QUESTIONNAIRE FILLED BY ALL THE PARTICIPANTS OF THE EXPERIMENT (1= STRONGLY DISAGREE, 2 = DISAGREE, 3 = NEUTRAL, 4 = AGREE, 5 = STRONGLY AGREE)

Serial Number	Statement	Response
A1	Have you ever touched a similar robot before	☐ Yes ☐ No
A2	Have you ever been to an orchard or farmland before	☐ Yes ☐ No
A3	The facilitation of the robot was of high quality	1 2 3 4 5
A4	I felt comfortable with the appearance of the robot	1 2 3 4 5
A5	I understood the robot	1 2 3 4 5
A6	I trusted the information given by the robot	1 2 3 4 5
A7	The robot was friendly	1 2 3 4 5
A8	It was fun to interact with the robot	1 2 3 4 5
A9	I gained more agricultural knowledge by interacting with robots	1 2 3 4 5
A10	Through interaction with robots, I have a more comprehensive understanding of the exhibition	1 2 3 4 5

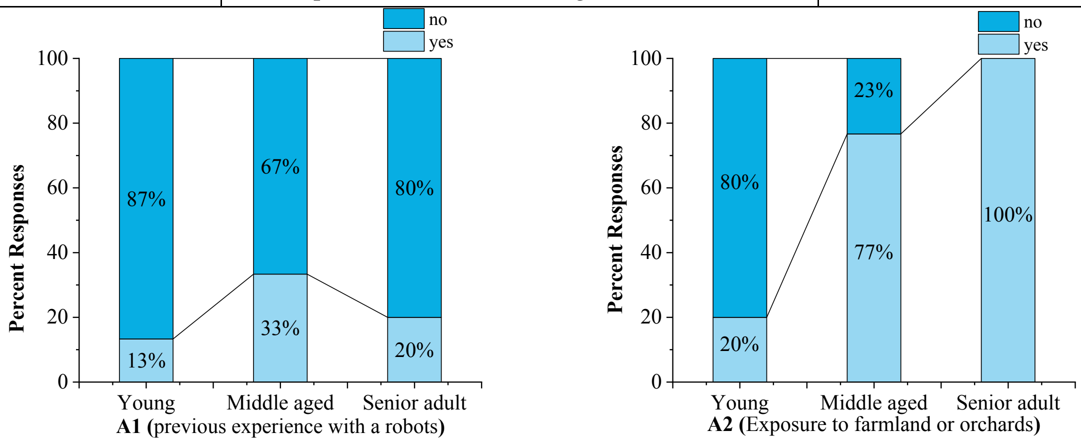


Fig. 8. Participants experience in relation to robots (A1) and farmland or orchard (A2) exposure

participants were dissatisfied or indifferent to our robot. However, during the next two days of the exhibition, the custom questions in the database were printed out. Thus, the participants collected information they were interested in order to interact with the robot. With some external guidance, the participants and robots interacted smoothly and efficiently. The results of questions A9 and A10 (table I) are presented in Fig. 9. The first day results are with external guidance (before) and the last two days results are with external guidance (after). Significant data changes can be seen as obvious from Fig. 9, where these changes are more pronounced in young and older people.

Through learning enormously about agriculture by interacting with robots, young people showed the biggest change, with the average increasing from 2.67 to 4.6. The seniors average change increased from 3.2 to 4.2. The response results of question A10 followed a similar pattern as those of question A9. Young people showed maximum change with their average increasing from 2.87 to 4.47. This was followed by the change that the external guidance brought in seniors' perceptions where it increased from 3.2 to 4.4. However, the external guidance had little influence on the middle-aged participants.

Data regarding the complete interaction time and the number of pause times are presented in Fig. 10. The young participants interacted with robots for the longest time (6 minutes). They also had the highest number of forced pauses (7 times). The time of middle-aged participants interaction with the robot system was 5 minutes. In their case, 3 times the speech playback was forced to pause. The seniors took the shortest time (3 minutes) to interact with the robot system. They also came up with the fewest forced pauses (1 time) in speech playback.

Under the right external guidance, all participants scored more than 3 points, meaning that they were well satisfied with the provided facilities. The response score of questions A3 to A10 (table I) is shown in Fig. 11. All participants responded that the facilitation of the robot was of high quality (A3). However, they seemed less satisfied with the appearance of the robot (A4). Here the middle-aged were the least satisfied (3.1 score) compared with either young (3.53) or senior (3.8) participants. In terms of understanding and ease of communication with the robot (A5), the young were the least satisfied with a score of 3.87. However, the middle-aged and senior participants were at par scoring 4.17 and 4 on robot understanding respectively. Regarding their trust in the lowest trust in robots (3.8 score). The reason might be that the seniors

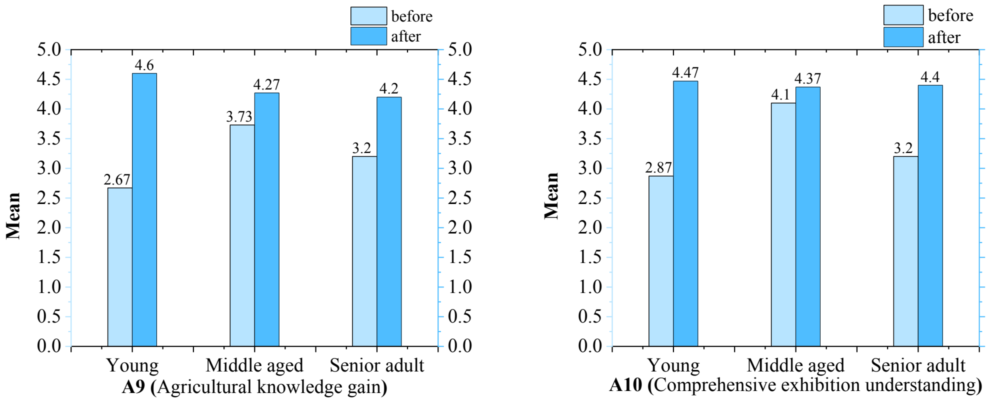


Fig. 9. Participants experience in relation to robots (A1) and farmland or orchard (A2) exposure

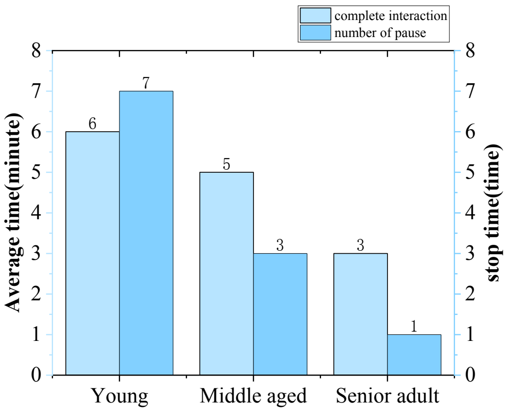


Fig. 10. Average time of complete interaction and the number of pause sound playback

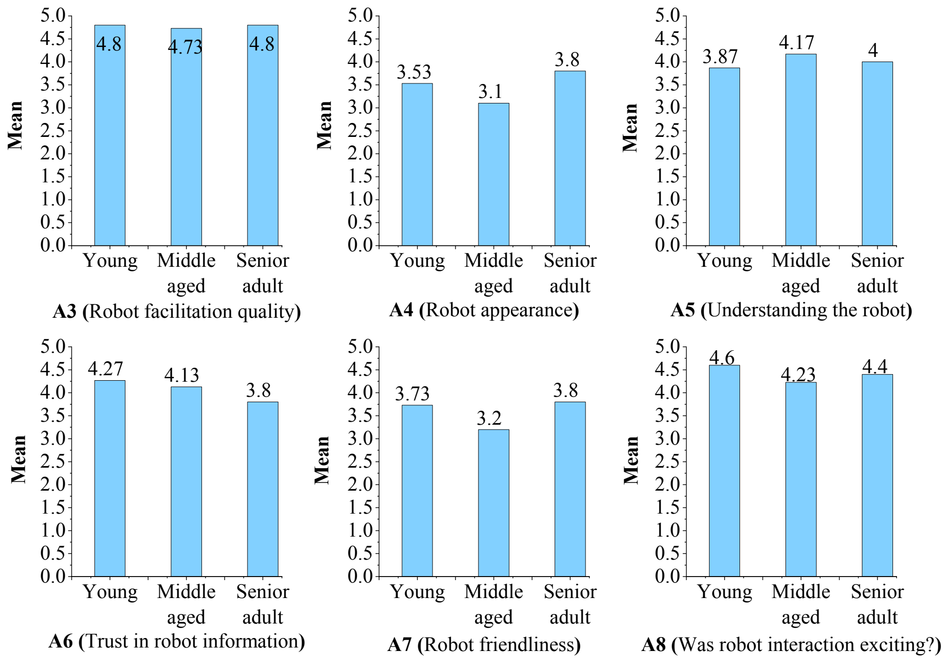


Fig. 11. Participants response regarding their interaction with robots in the agricultural exhibition

had extensive agricultural experience. And in their opinion, there was still room for improvement in the interactive information. The senior and young participants had a similar response (3.73 and 3.8 respectively) to the question whether robots were friendly (A7). However, middle-aged participants response was the lowest with a score of 3.2. In terms of whether it was fun in interacting with robots (A8), the average response score for all participants was greater than 4, showing their satisfaction and interest in robots. Here again the young participants were the most enthusiastic with the highest score of 4.6.

C. Conclusions of the study

An intelligent interactive robot system based on question and answer (Q&A) and image display was developed to introduce agricultural knowledge and achievements to exhibition visitors. The study showed that Q&A combined with the image display was a feasible and beneficial function. In addition, the experiment showed that our system was attractive to people of all ages. Here, the learning efficiency was improved by keeping visitors' interest through asking, listening and watching (i.e. audio-visual participatory approach). This system contributes in introducing a new interpretation mode for agricultural exhibitions.

The responses from participants of three different ages were thoroughly analyzed and evaluated. This brought us to a few conclusions. First, the young and senior participants needed external guidance to interact and make use of the robot effectively. Second, the designed robot system attracted the youth attention more than the middle-aged and senior citizens. So, it is an effective system for creating awareness concerning agricultural knowledge, especially among the young folks. Last, through the combination of voice broadcast and image display, the participants could understand the agricultural knowledge and the exhibition achievements more intuitively.

Moreover, we designed a robot system for agriculture exhibition, similar to that of Hu, et al. [5]. However, we built our robot according to the functional requirements. From the cost and economic perspective, our robot is more suitable for mass production and use in agricultural exhibitions.

D. Suggestions for future work

The interactive robot system is still in the early stage of development. This study introduced the primary functions of this interactive model. However, it opens a wide field for exploration. There is still much future work to be done to address all the fundamental challenges arising from technical difficulties. These include voice tracking, operational sensitivity, and image recognition accuracy, among others. The sound source positioning effect of our system was not ideal and also needs further improvement or direct removal.

To make the system more knowledgeable, the database needs to be expanded. Similarly, improved path planning and autonomous navigation can be used to enhance the intelligence of the robot.

ACKNOWLEDGMENT

This work was supported by the Shandong Key Research and Development Program under Grant 2019GNC106091.

REFERENCES

[1] Y.K. Cho, D. Yook, S. Chang and H. Kim, "Sound source localization for robot auditory systems," IEEE Transactions on Consumer Electronics, vol. 55, no. 3, pp. 1663-1668, August 2009.

[2] Q. Wu, S. Wang, J. Cao, B. He, C. Yu and J. Zheng, "Object Recognition-Based Second Language Learning Educational Robot System for Chinese Preschool Children," IEEE Access, vol. 7, pp. 7301-7312, January 2019.

[3] G. Gordon, S. Spaulding, J.K. Westlund, J.J. Lee, L. Plummer, M. Martinez et al. "Affective personalization of a social robot tutor for children's second language skills[C]," Thirtieth AAAI Conference on Artificial Intelligence. 2016.

[4] D. Gouaillier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourc et al. "Mechatronic design of NAO humanoid," IEEE International Conference on Robotics and Automation, pp. 769-774, May 2009.

[5] R. Hu ; X. Yu, B. He ; Z. Li, "Implementation of voice interaction based on questions and answers for guide robot in exhibition hall," 2nd International Conference on Robotics and Automation Engineering (ICRAE), pp. 105-109, 2017.

[6] R.R Kima, Y. Koren, M. Yachini, G. Gordon, "Human-Robot-Collaboration (HRC): Social Robots as Teaching Assistants for Training Activities in Small Groups," 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 522-523, 2019.

[7] A.K. Pandey, R. Gelin, "A mass-produced sociable humanoid robot: pepper: the first machine of its kind[J]". IEEE Robotics & Automation Magazine, vol. 25, no. 3, pp. 40-48, 2018.

[8] Z. Al barakeh, S. alkork, A.S Karar, S. Said, T. Beyrouthy, "Pepper Humanoid Robot as a Service Robot: a Customer Approach," 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART), pp. 1-4, 2019.

[9] H. Fang, F. Wu, Z. Zhao, X. Duan, Y. Zhuang, M. Ester, "Community-based question answering via heterogeneous social network learning," Thirtieth AAAI Conference on Artificial Intelligence, 2016

[10] X. Shen, L. Gu, A.L. Jia, "Cultivating Online: Question Routing in a Question and Answering Community for Agriculture," arXiv preprint arXiv, 2019.

[11] D. Pavlidi, A. Griffin, M. Puigt, A. Mouchtaris, "Real-time multiple sound source localization and counting using a circular microphone array[J]," IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 10, pp. 2193-2206, July 2013.

[12] J.M. Valin, F. Michaud, J. Rouat, D. Letourneau, "Robust sound source localization using a microphone array on a mobile robot," Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1228-1233, 2003.

[13] M. Jia, J. Sun, C. Bao, "Real-time multiple sound source localization and counting using a soundfield microphone," vol. 8, no. 6, pp. 829-844, November 2017.

[14] L. Fei_Fei, P. Perona, "bayesian hierarchical model for learning natural scene categories," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 524-531, July 2005.

[15] K. Alsabti, S. Ranka, V. Singh, An efficient k-means clustering algorithm[J]. 1997.

[16] S.V.M. Vishwanathan, M.N. Murty, "SSVM: a simple SVM algorithm," vol. 3, pp. 2393-2398, May 2002.