6-3 Real Living Experiments with Conversational Robots at Ubiquitous-Home

UEDA Hirotada, MINOH Michihiko, CHIKAMA Masaki, SATAKE Junji, KOBAYASHI Akihiro, MIYAWAKI Kenzaburo, MATSUMOTO Naoko, and KIDODE Masatsugu

The situation recognition ability of the robot is enhanced by connecting a home ubiquitous network and conversational robots and context-aware services can be provided suitably. By taking such an approach, the situation explanation ability of the robot is also enhanced by acquiring information through the network. The new development of the human robot interaction can be expected in total. In this paper we describe a prototype system that is developed in the Ubiquitous-Home based on such a concept. Then, the study of the actual proof experimental life in the house is discussed.

Keywords
Home, Ubiquitous network, Conversational robots, Context-aware services, Ubiquitous-Home, Actual proof experimental life

1 Introduction

We have seen an increasing number of prototypes of rooms equipped with numerous sensors embedded in the living space. Such studies date back to the Xanadu House[1] in the US, and more recently, active R&D efforts are underway at the Georgia Institute of Technology[2], the University of Tokyo[3], and the Digital Human Research Center of the National Institute of Advanced Industrial Science and Technology (AIST)[4]. However, when viewed from the perspective of our actual daily lives, none of the results to date have been able to present any obvious or significant benefits for inhabitants, and it may be said that no one yet has a clear outlook on what sort of “killer app” to expect. A new concept and framework is required in order to dispel these uncertainties for the future.

In the UKARI Project, a new concept was proposed in which new developments in human-robot interactions would be promoted by enhancing the context recognition capacities of robots, by connecting conversational robots to a home ubiquitous network, and also by enhancing the ability of these robots to explain, with reference to information they could acquire through the network. Based on this concept, a system offering various services has been designed using a conversational robot as the interface in the Ubiquitous Home. Long-term real-life validation experiments have been performed to evaluate the system; here we will give an overview of the prototype and report on the results and findings of the experiment.

2 Basic concept

A system with integrated sensors and appliances may be regarded as an invisible type of robot. The basic concept for realizing services inside a Ubiquitous Home lies in offering previously unavailable services by creating a new form of interaction between humans and robots — specifically through cooperation between the invisible-type robot and the visible conversational robot. The information collected by the sensor groups deployed over the entire house will enhance the ability of the robot to recognize contexts, and this enhanced ability will feed back to a system that can provide context-aware services that make full use of the results of such context recognition. Further, the conversational robot's capacity for explanation will help establish a new working relationship between humans and robots.

In the UKARI Project, a mother-child metaphor has been proposed as an integral framework for the entire system of the Ubiquitous Home (Fig. 1)[5]. The house, which is an invisible type of robot, will autonomously control the appliances connected to the network based on the information obtained through the sensors. These functions have been designed along the lines of the metaphorical mother, who will warmly watch over the family and who will casually extend a helping hand when needed. The functions of a visible conversational robot having an independent body are designed according to the child metaphor, in which services are provided through conversations with the inhabitants and cooperation with its mother.

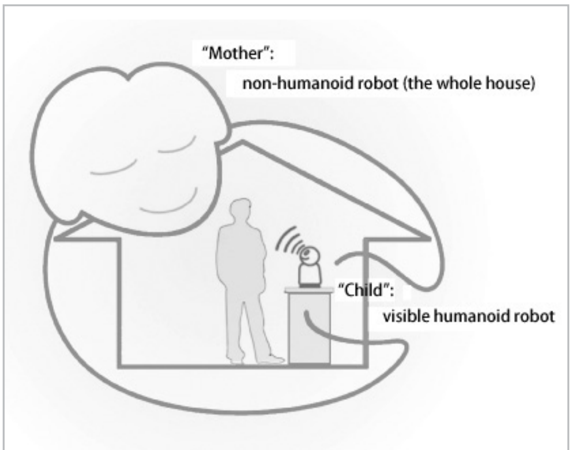


Fig. 1 Mother-child metaphor

The adoption of this architecture will lead to an interface that engages in natural-language conversation, in turn enabling the provision of high-quality, context-aware services. The conversation between humans and the robots will not simply be limited to confirmation of the user's intent before executing a service, but will also involve the introduction of services which the user is unaware of through casual discussions, and responses to questions from the user after execution of services. These explanation functions may be expanded to create a mechanism by which the robot begins to learn the preferences of the user.

The Ubiquitous Home is an R&D facility as well as the site of validation testing. In order to shorten the construction schedule significantly, the facility was constructed inside a pre-existing building and not built separately as a stand-alone house. However, since weight was given to the actual proof experimental lives under extended-stay conditions, it featured two bedrooms, a living room, a dining room, a kitchen, a bathroom, and a lavatory. The entrance hall through which the residents enter and leave the Ubiquitous Home was also specially designed to allow the participants in the experiment to enter or leave the Home freely regardless of the time of day, in order to simulate as closely as possible daily life in a normal apartment, including commuting to work and attending school.

3 Overview of the Ubiquitous Home

The Ubiquitous Home is equipped with a multitude of sensors and appliances. Pressure sensors (resolution of 180 mm × 180 mm) are installed below all floors, and it is possible to track the position and path of any given person. The infrared human-detection sensors installed above the doors to each room and near the floors of the hallways and kitchen provide powerful assistance in detecting the presence of people. Excluding the bedrooms, TV cameras and microphones are installed in the four corners of each room to monitor and simultaneously record the behavior and actions of the residents. Open/close sensors are installed in all doors, windows, cupboard doors, and drawers, and the bookshelf in the study is equipped with a photoelectric sensor to detect the placement and removal of objects. The sofas are installed with a seat sensor. Two types of RFID tag systems have been adopted: an area-type that detects the presence of tags in each room using a receiver installed in the ceiling, and a gate-type that detects the tags as they pass through the door to each room. A conversational robot has been placed in each room and in the entrance hall.

Figure 2 shows the arrangement of the appliances and sensors in the living room as an example, and Fig. 3 shows a scene from the kitchen while a resident is cooking. The kitchen counter is inlaid with glass covered with a special film on which images may be projected to enjoy cooking assistance from a conversational robot. All data collected by the sensors are relayed to the computer room adjacent to the living space and stored in the database. The video images captured by the ceiling camera may be retained for up to two months in the database at temporal resolution of five frames per second. The continuous voice and audio information collected by the ceiling and robot microphones may also be stored for two months in the database. Time stamp information containing the time of data collection is attached to all sensor data so that the required data may be quickly searched, displayed, and reproduced.

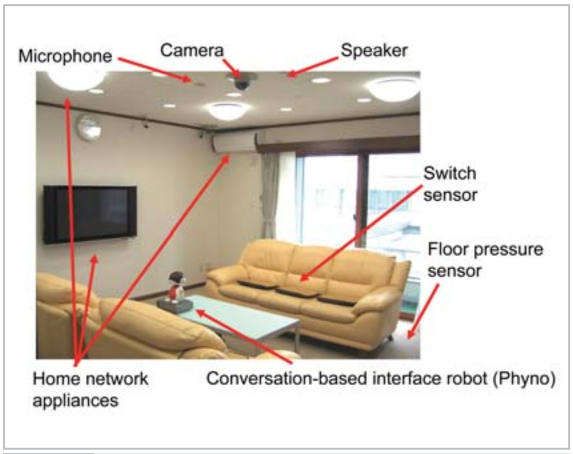


Fig.2 Positions of appliances and sensors in the living room

Fig.3 Scene from kitchen while resident is cooking (photographed by the ceiling camera)

The data storage function has completely eliminated the need for direct or indirect contact between the participants of the real-life experiment and the researchers, and has guaranteed conditions under which the participants can conduct their daily lives naturally in the Ubiquitous Home during the extended real-life validation experiment. Researchers have also conducted interviews of the participants after the end of the experiment to review the required data.

4 Conversational interface robot

The conversational interface robot (referred to as “Phyno”), which plays the role of the child in the mother-child metaphor, was designed based on the following concepts[6].

- Incapable of bipedal walking but placed in various positions instead
- Small enough to fit anywhere
- Body language will reproduce the cuteness of a child
- Voice recognition will take place using a built-in microphone

Since the robot should be of a size that will allow it to be placed anywhere in the house, the total height was limited to 25 cm, and the freedom of motion was set to 6, 3 for the head, 1 for each arm and 1 for the body. A full image of the Phyno prototype robot is shown in Fig. 4. A servo system for radio-controlled models and USB cameras were used to create a compact and lightweight body, as well as to reduce the costs of development.

The CCD camera, microphone, and speaker were selected as devices to interact with the user. We believed that having users wear a microphone or placing microphones for voice recognition separately from the robot would force the users to unnaturally alter their patterns of daily life. Thus, the voice recognition function had to be enabled using only the built-in microphone of the robot. Although it was not easy to ensure successful voice recognition using only the built-in microphone inside a room (where sounds from the TV or the kitchen are likely to interfere), we made it possible for the robot to accurately recognize a voice prompt spoken at a distance of approximately 50 cm, through fine-tuning of the audio model[7].

In the design of the human interface, it was important to present affordances such that the user was able to create and maintain a favorable mental model[8]. In the case of Phyno, the present conversational robot, it was feared that deficiencies in the present technologies for voice recognition and natural language communication would disappoint the user. According to the literature[9], reasoning concerning physical phenomena such as objects falling down a slope becomes possible at 3 to 4 years of age. It has also been reported based on the results of observation that by that time, a child is capable of understanding the concept of category, at least to a certain degree. Thus, we have opted to incorporate these findings and set the age of Phyno, a full-bodied entity of the child metaphor, at 3 years old. In a conversation between the user and Phyno (who will continually refer to the database that has recorded all sensing data inside the Ubiquitous Home, in chronological order), the reasoning and categorical data-handling abilities of a 3- to 4-year old will be combined and expressed using the body language of a child. We hoped that this would create a mental model of the robot as a “clever child” who is highly knowledgeable of the sensor network and appliances in the minds of the user.

Fig.4 Phyno, the conversational robot

5 Context-aware services

The Ubiquitous Home has been packaged with context-aware services for its residents under the framework presented in Fig 5. Some services are of the “invisible and autonomous type”, such as the light switch service, which will turn off the lights when a user leaves the room; other services are executed through voice interaction with Phyno. Some examples of the latter are given below.

(1) Home appliance control

All remote-controlled functions of home electronic appliances such as light fixtures, air conditioners, and TV are included. (When deemed possible, user-dependent or context-dependent controls will also be provided.)

(2) Forgotten items check

When the system judges that a user is leaving the house based on the results of floor sensor tracking, it will search the database to obtain information on the potential destination based on the schedule of the individual identified by the RFID tag, searches for the list of things taken by the user to the corresponding destination based on archived records, and

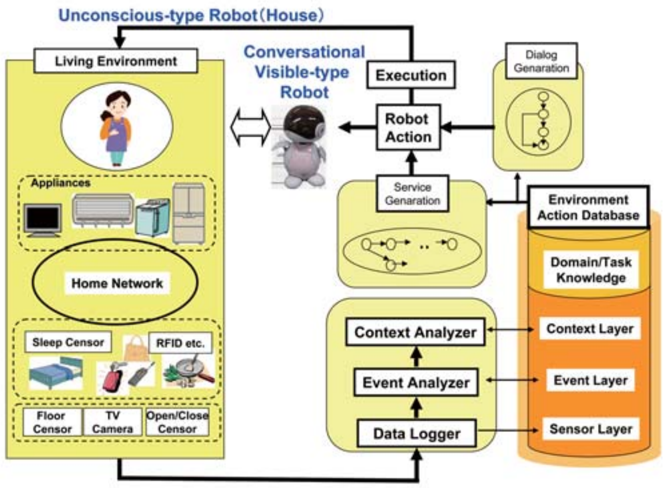


Fig.5 *Framework of context-aware service execution*

checks for items that may have been forgotten by sensing the tags attached to the belongings.

(3) TV program recommendation function

In response to requests made in conversation or when a user approaches the TV, the robot recommends a currently broadcast or recorded program using the electronic TV program listing obtained through the network according to the past records of TV programs selected by the user or his or her entered preferences[10].

(4) Cooking recipe search

Using a mechanism called the word-association conversation[11], menus that may not have been conceived by the user alone will be selected from 1,000 recipes stored in the database.

(5) Washing machine cycle completion notice

When the washing machine has completed its cycle, the location of the user who must be notified is sought using the sensors, and the nearest functional element (speaker or screen) is used to provide the notification.

6 The robot's conversational strategy

A child sometimes interrupts a conversation between adults (which he cannot totally understand), and when he does, he applies his association and reasoning abilities and his knowledge to the fullest extent, interacting via speech while observing the reactions of the adults. When he judges that the reactions of the adults respond positively to the content of the speech he previously provided, the child will apply his association and reasoning abilities to try to continue that subject of conversation. Such a "playing-ball" conversation based on association is referred to as a word-association conversation and was incorporated into the robot's conversational system[12]. This style of conversation allows for the introduction of services to users that they were previously unaware of. The subject of conversation may veer off track due to incorrectly heard words, but it was believed that even misunderstandings could provide the user with helpful hints.

Another strategy is seen in the robot's explanation function. In a Ubiquitous Home, context-aware services are provided by sensing; service information and trigger events from one moment to the next are all collected into the distributed environment behavioral database (See Fig. 5). This log may be used to answer questions posed by the user. The explanation function was created using this log in combination with the user adaptive function. By adopting this conversational strategy, it becomes possible to create a framework in which a solution to a problem faced

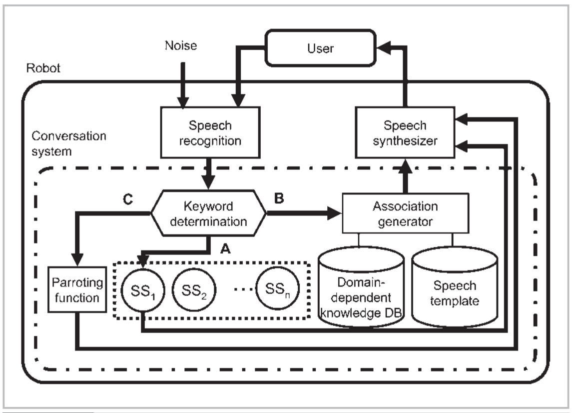


Fig.6 Architecture of the voice interaction software

by the user is found by the robot by searching for information, based on the operational instructions from the network based on the context of the problem, the robot then provides the appropriate guidance to the user.

Figure 6 presents the architecture of the prototype voice interaction software. The user's speech is recognized by a speech recognition device with a large vocabulary, and input into a keyword assessment unit as a character string, where the speech will be categorized into one of three types and sent forward to the corresponding processor units. When a keyword is included in the service script (SS), the service executer will carry out the corresponding service. When the keyword is included in the domain-dependent knowledge, it is handed over to the association generator and a word- association conversation is exchanged with the user. If a corresponding keyword is not found, the robot parrots the user's words as feedback to the user that the word has been recognized, but that the robot cannot judge how to process the word, in an attempt to gain clarification from the user.

7 Actual proof experimental life and results

In an actual proof experimental life, four families with different family structures and age-group compositions were asked to spend a little over two weeks in the prototype Ubiquitous Home. At the start of their stay, they received a briefing on the sensors installed and the various services available. They were instructed to follow their regular routines in daily life, but were asked to refrain from using the remote controls of the home appliances and to operate them by voice interaction with the conversational robot instead. They were also asked to keep notes on whatever they noticed during their stay, and at the end of the two-week experiment, they were asked to return for a three-hour interview on another day.

The participants were selected from groups of families not related to the project who responded to a public announcement. The experiments were carried out after receiving approval from the biometrics-study ethics panel formed by NICT, after careful examination of the overview of research, the experimental facility and the instruments used, the psychological and mental stress on the participants, the method for obtaining written consent from the participants, and a method of data handling that would not violate the privacy of the participants.

The following are the major results of the experiment.

All participants replied that while they did not mind being observed by the built-in camera of the robot, they were uncomfortable with the ceiling camera and felt as though they were being watched. Also, all participants said that they felt some tension in the first three days and would constantly try to sit up straight and act normally, but after four days they relaxed and did not mind the ceiling camera.

As for the control of the TV, air conditioners, and lights by voice interaction with the robot, it took only about three days for the participants to become used to the style of speech that would increase the rate of successful voice recognition. At the start of the experiment, the recognition rate (calculated by the number of repeated requests) was 20 — 30%, while on the fourth day, the rate increased to 80 — 90%. As a result, the participants praised the service, saying that they were relieved of the effort of searching for the remote control, and that they would like to see all remote-controlled functions to be integrated into the interactive voice control of the robot.

Some participants responded that they found the forgotten-items check service to be useful. However, in the present state, the sensitivity of the RFID reader is insufficient, and the bag containing the items must be held up against the tag reader for successful identification. The identification process also requires several tens of seconds. Improvements in RFID reader performance are thus eagerly awaited.

The evaluation of the washing machine cycle completion notice was divided. There were two housewives who answered that it was significantly helpful when they were busy.

The father of a three-year old stated that his child seemed to recognize that Phyno was the entity that had the control over the TV channels, and that he could use the real-bodied Phyno as a frightening character in disciplining the child.

A couple in their sixties who participated in the experiment commented that they were eager to return to the house from shopping as they were worried about Phyno. Some also said that they always watched the entire program when Phyno had recommended it. A young housewife noticed that the Phyno in the kitchen seemed to follow her instructions well, while the Phyno in the living room seemed to be more loyal to her husband. This is believed to be the result of humans learning to voice their requests appropriately with the robot in the respective rooms, but the report seems to be expressed in a manner suggesting that it was the robot that learned. Such examples were often observed in the present real-life experiment, and as reported in various references[13], this phenomenon stands as an important theme for the future, one that should be analyzed from the perspectives of whether robots are sentient or not, or whether humans may become emotionally attached to the robots that share in their lives.

8 Conclusions

Through the UKARI Project, the authors have repeatedly asked what “killer apps” may be offered by such a system. We believe that an approach in which the daily patterns of each household and the preferences and habits of each user are quickly identified for daily and non-daily situations and services are adapted to each household or user is a more valid approach than trying to design a service which will be universally useful to all households. Thus, we have proposed a concept that will allow for the realization of previously unavailable services by creating a framework for cooperation between a home ubiquitous network and a conversational robot.

The present paper reported on a prototype conversational robot system and examples of context-aware services, in addition to the findings of a real-life validation experiment in which inhabitants actually spent two weeks inside a house equipped with such a system and services. The effectiveness and the potential of the present system have been confirmed through the prototype and experiments. These results lead us to conclude that a day will come when robots will truly be able to coexist with people and to understand their intentions. From this point forward, it is important that we continue to carry out experiments and deepen our understanding in this area, in addition to enhanced efforts at interdisciplinary research.

References

1 R. Mason, et al., "Xanadu: The Computerized Home of Tomorrow and How It Can Be Yours Today!", Acropolis Books, 1983.

2 C. D. Kidd, et al., "The Aware Home: A Living Laboratory for Ubiquitous Computing Research", In the Proceedings of the Second International Workshop on Cooperative Buildings - CoBuild'99. Position paper, 1999.

3 Taketoshi Mori, Hiroshi Noguchi, and Tomomasa Sato, "Sensing Room — Room-Type Behavior Measurement and Accumulation Environment —", JRSJ Vol.23, No.6, pp.25-29, 2005. (in Japanese)

4 Yoshifumi Nishida, Yoichi Motomura, and Tatsuhiro Yamanaka, "Sensing and Modeling Everyday Life Behavior for Child Injury Prevention", 2006 IEICE General Conference, No.HK-1-5, pp.SS-9/SS-10, 2006. (in Japanese)

5 Hirotada Ueda, Jun Sato, Masaki Chikama, and Masatsugu Kidode, "Visible/Unconscious Robot Cooperation Mechanism — Mother-Child Metaphor —", The 28th Human Interface Society, Correspondences on Human Interface, Vol.6, No.3, pp.57-64, 2004.

6 Hirotada Ueda, Masaki Chikama, Junji Satake, Jun Sato, and Masatsugu Kidode, "A Prototyping of a Dialog Interface Robot at the Ubiquitous-Home", IPSJ Ubiquitous Computing Systems (UBI), No.2005-UBI-7(34), pp.239-246, 2005.

7 Ryuichi Nishimura, Masaki Chikama, Akihiro Kobayashi, Junji Satake, and Hirotada Ueda, "Speech and Noise Recognition for Dialogue Robot in Ubiquitous Home Network", IEICE Pattern Recognition and Media Understanding (PRMU), No.PRMU2005-59, pp.47-52, 2005.

8 D. A. Norman, "The Psychology of Everyday Things", Basic Books Inc., U.S., 1988.

9 Inui, Toshiro (eds.), "Ninchi hattatsu to shinka", Tokyo: Iwanami Shoten, pp.88-93, 2001. (in Japanese)

10 Seiji Tsuchiya, Junji Satake, Masaki Chikama, Hirotada Ueda, Kazumi Ohkura, Hiroshi Kano, and Masashi Yasuda, "A Study of Effectiveness of the TV Program Recommendation System", IPSJ Human Interface (HI), Vol.2006-HI-117, No.14, pp.95-102, 2006. (in Japanese)

11 Hirotada Ueda, Akihiro Kobayashi, Junji Satake, Masaki Chikama, Jun Sato, and Masatsugu Kidode, "Dialog Strategy for Interactive Robot Interface in Home Network", Trans. Of IPSJ, Vol.47, No.1, pp.87-97, 2006. (in Japanese)

12 Akihiro Kobayashi, Hirotada Ueda, Junji Satake, Masaki Chikama, and Masatsugu Kidode, "Real-life Experiments of the Dialogue Robot in the Home Ubiquitous Environment and an Evaluation of the Dialogue Strategy", Trans. Of IPSJ, Vol.48, No.5, pp.2023-2031, 2007. (in Japanese)

13 N. Matsumoto, et al., "The cognitive characteristics of communication with artificial agents", International Symposium on Advanced Intelligent Systems, pp.1269-1273, 2006.

UEDA Hirotada, *Ph.D.*

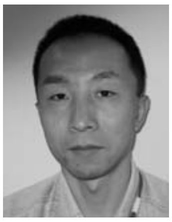
Professor, Faculty of Engineering, Kyoto Sangyo University (Former: Research Fellow, Distributed and Cooperate Media Group, Keihanna Human Info-Communication Research Center, Information and Network Systems Division)

Computer Vision, Human Interface, Intelligent Media

MINOH Michihiko, *Dr. Eng.*

Professor, Academic Center for Computing and Media Studies, Kyoto University (Former: Group Leader, Distribute and Cooperative Media Group, Keihanna Human Info-Communication Research Center, Information and Network Systems Division)

Three dimensional Modeling, Environmental Media, Communication Support System



CHIKAMA Masaki

NICT Technical Staff (Former: NICT Technical Staff, Distrubuted and Cooperative Media Group)

Dialog System

SATAKE Junji, *Ph.D.*

Expert Resercher, Universal City Group (Former: Expert Resercher, Distributed and Cooperative Media Group, Keihanna Human Info-Communication Research Center, Information and Network Systems Division)

Computer Vision, Human Interface

KOBAYASHI Akihiro, *Ph.D.*

Expert Resercher, Universal City Group (Former: Expert Resercher, Distributed and Cooperative Media Group, Keihanna Human Info-Communication Research Center, Information and Network Systems Division)

Computer Vision, Human-Robot Interaction

MIYAWAKI Kenzaburo

Faculty of Information Science and Technology, Osaka Institute of Technology (Former: Trainee, Universal City Group)

Life Support Applications in Ubiquitous Environment

MATSUMOTO Naoko, *Ph.D.*

Graduate School of Information Science and Electrical Engineering, Kyushu University (Former : 21COE Program (Framework for Systematization and Application of Large-Scale Knowledge Resources), Tokyo Institute of Technology)

Psychology, Human Interface, Pragmatics

KIDODE Masatsugu, *Ph.D.*

Professor, Graduate School of Information Science, Nara Institute of Science and Technology

Intelligent Media Understanding, Robot, and Interface