Service-Oriented Software Design Model for Communication Robot

T. Kobayashi, K. Arai, T. Imai
Graduate School of Engineering,
Nagasaki University
Nagasaki, Japan
toru@cis.nagasaki-u.ac.jp

A. Kanai
Faculty of Science and Engineering
Hosei University
Tokyo, Japan
yoikana@hosei.ac.jp

S. Tanimoto
Faculty of Social Systems Science
Chiba Institute of Technology
Chiba, Japan
shigeaki.tanimoto@it-chiba.ac.jp

T. Miyazaki, A. Tsujino
Department of Neurology & Strokology
Nagasaki University Hospital
Nagasaki, Japan

H. Sato
Information Technology Center
The University of Tokyo
Tokyo, Japan
schuko@satolab.itc.u-tokyo.ac.jp

***Abstract*— A communication robot like an AI speaker has become popular as one of consumer services in recent years. As communication robots are widely used, various new demands arise. Therefore, system development is required in order to reflect these demands. Meanwhile, a Service-Oriented Architecture (SOA) or Service-Oriented Software Engineering attracts attentions for productivity improvement. However, there are few examples that apply a SOA or Service-Oriented Software Engineering to consumer services based on communication robots. We have developed communication robots that enable users to receive communication services via Social Networking Service (SNS) regardless of information literacy skills or detect sings of cranial nerve diseases. In this paper, we introduce these examples and discuss a framework common to them. Based on this framework, we propose a human-friendly and service-oriented software design model for communication robots.**

Keywords*— *SOA, Service-oriented software engineering, Communication robot, Internet of Things, Artificial Intelligence

I. Introduction

A communication robot typified by an AI speaker has become popular as a one of consumer services. As communication robots are widely used, various new demands arise. Therefore, system development is required in order to reflect these demands. Particularly in Japan where the population keeps decreasing and aging, quick development of services for the elderly or persons with disabilities are desired. In addition, as the number of an elderly person living alone increases [1], to detect sings of cranial nerve diseases such as dementia or a stroke at early stages is desired. Meanwhile, a Service-Oriented Architecture (SOA) or Service-Oriented Software Engineering attract attention for productivity improvement. However, there are few examples that apply a SOA or Service-Oriented Software Engineering to consumer services based on communication robots.

We have developed the Social Networking Service (SNS) Agency Robot that enables a person who cannot use a smartphone to have message communication by conversation with this robot. This system uses LINE that is the most popular message communication application in Japan to allow message communication. It uses an external cloud Application Programming Interfaces (API) for voice and face recognition and translation. We implemented this system on several robots having different characteristics so that the system can be used by anyone according to their information literacy skills or accessibilities. Moreover, we have made it possible to detect sings of dementia from ordinary conversation by adding the function to detect signs of dementia. Furthermore, the addition of the function to detect sings of a stroke enables detection of stroke sings that is fatal for an elderly person living alone.

In this paper, we discuss a common framework from the point of view of a SOA or Service-Oriented Software Engineering after we introduce six examples of our developped communication robots. Based on this framework, we finally propose a service-oriented software design model for human-friendly services using communication robots.

II. Examples of Communication Robots

A. SNS Agency Robot (PaPeRo)

As Japanese aging rate increases every year, it is predicted that it continues. Under such situation, many watching services for the elderly are offered. However, most of the services are designed to grasp, check condition, detect abnormality of the elderly by a sensor, and inform to their families. For example, ZOJIRUSHI offers a service called “MIMAMORI Hotline” [2].

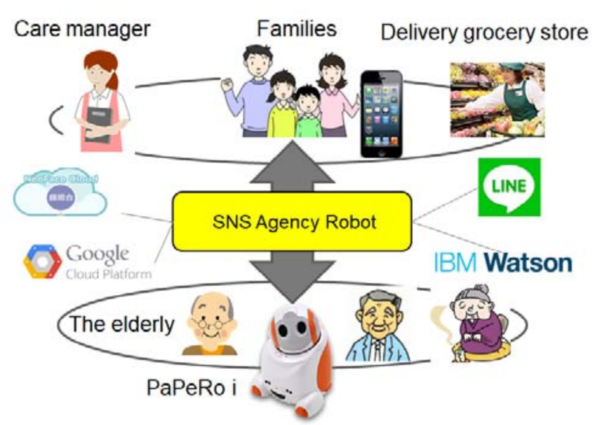

Fig. 1. Outline of SNS Agency Robot

978-1-7281-6972-9/20/$31.00 ©2020 IEEE
DOI 10.1109/SOSE49046.2020.00011

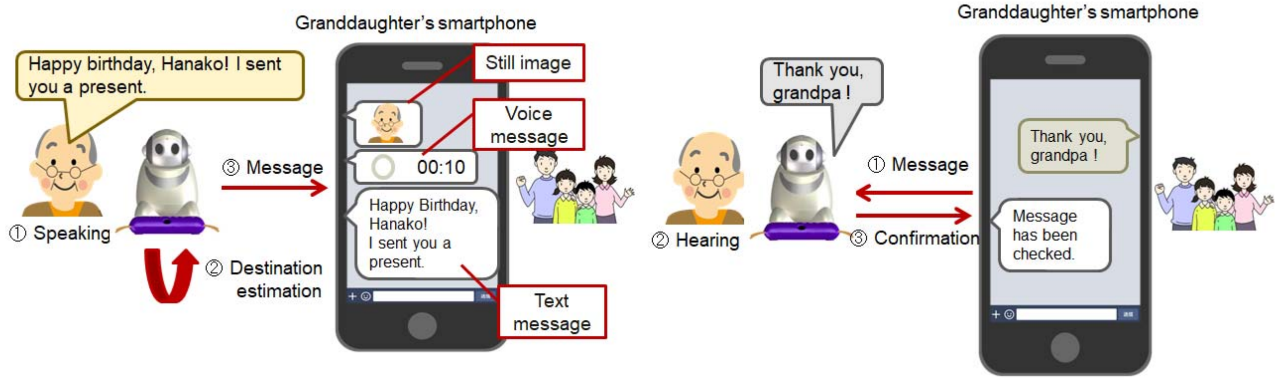

Fig. 2. Use case of SNS Agency Robot (PaPeRo)

This service transmits information via email to an elderly person's families who live apart when the elderly person uses an electric pot. Meanwhile, SNS such as Twitter became popular among the young people. However, the elderly face big hurdle in mastering smartphones to use SNS. Therefore, we developed the SNS Agency Robot which enables the elderly to communicate interactively via Twitter or LINE, even though they are unable to use smartphones [3] (Fig. 1). For example, an elderly person talks to the robot "Happy birthday, Hanako (his granddaughter), I sent you a present." Then his still picture, his voice message, and its converted text message are shown on LINE of Hanako's smartphone. As Hanako replies "Thank you, grandpa" in a text, the robot delivers the message with voice synthesis (Fig. 2). It is also possible to send still pictures and videos to the elderly person from LINE.

Its principle feature is that the appropriate address is automatically set depending on its context just by talking to the robot, since the Internet of Things (IoT) technology and artificial intelligence are in cooperation. Specifically, this is realized by making Watson Natural Language Classifier (NLC) [4] learn the message exchange history with communication partners. For learning of artificial intelligence, the correct data is to be given. In this case, the only person who knows the correct data, "a destination address" is the elderly person. However, it is impossible to make the elderly person input the correct data. Therefore, we implement "message exchange learning type address estimation system" based on reverse thinking that makes it learn the response from the young to the message from the elderly. We use Google Cloud Platform as the voice recognition function which converts the elderly's voice to the text, NEC's PaPeRo i [5] as the robot at the elderly's that the face recognition function is mounted on, and NEC's NeoFace KAOATO as the face recognition for personal identification. The main control function is constructed on Sakura VPS, a commercial cloud service (Fig. 3). This is the second feature that this system configuration makes the best possible use of external open innovation and commercial cloud services which are inexpensive and highly maintainable.

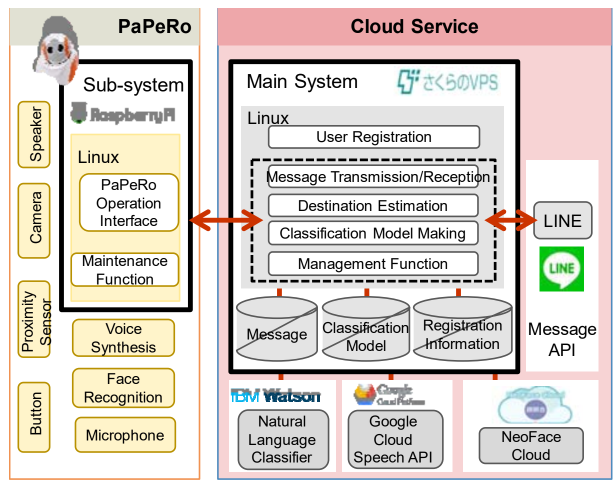

Fig. 3. System configuration of SNS Agency Robot (PaPeRo)

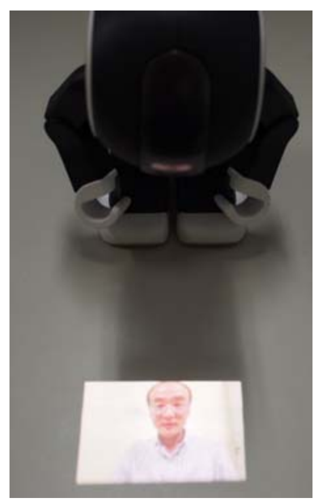

Fig. 4. Built-in projector of RoBoHoN

B. *SNS Agency Robot (RoBoHoN)*

For use of PaPeRo i, it is required to connect another monitor by HDMI to itself in order to check still pictures and videos sent via LINE. It is also necessary to press a button for its basic operation. Therefore, we developed the SNS Agency Robot by using SHARP's RoBoHoN which makes it possible to check still pictures and play videos by a built-in projector function (Fig. 4) [6]. It can be operated by a voice. The

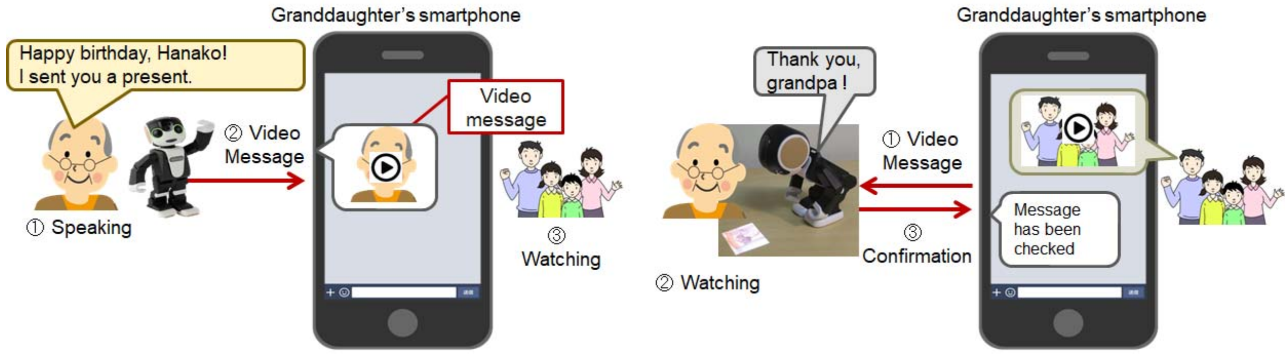

Fig. 5. Use case of SNS Agency Robot (RoBoHoN)

difference in a use case is that the message the elderly can transmit is only limited to video messages. Therefore we do not implement the function of address estimation (Fig. 5).

Fig.6 shows the system configuration of the SNS Agency Robot (RoBoHoN). We assigned the main system of the SNS Agency Robot to Amazon Web Services, also the sub-system to RoBoHoN as an Android application. Sharp offers a software development kit for a RoBoHoN Android application. The main system and sub-system are communicating with each other to execute all functionalities. The main system provides mainly a LINE messages (texts and videos) exchange function and a system management function. The sub-system offers mainly an interface function between the main system and RoBoHoN, which includes the LINE messages (texts and videos) exchange function and a RoBoHoN application interface such as a robot movement, an embedded projector, and facial recognition.

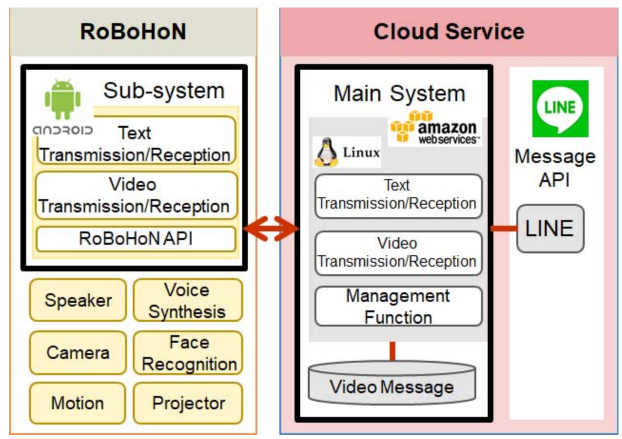

Fig. 6. System configuration of SNS Agency Robot (RoBoHoN)

C. SNS Agency Robot (Original robot)

Fig.7 shows the appearance of the SNS Agency Robot as our original robot. There is a push button switch on the top of the robot. When the switch is pressed down for a short time (about two seconds), the robot transmits LINE messages via a built-in Wi-Fi router after 20 seconds recording by an incorporated camera and a microphone. When the switch is depressed for more than three seconds, the received message is reproduced. The camera for recording is embedded in around the mouth of the robot. Below the mouth, there are two kinds of LEDs. When a LINE message is received, the red LED on the left side is put on. The blue LED keeps lightning during twenty seconds' recording when the function of short-pressing of the switch on the head is in operation. The video recorded by the camera and the voice by microphone that is equipped beneath LEDs are synthesized in an mp4 file. A human sensor for sensing the movement of a human body is mounted. When a new message is received and the human sensor senses a person in the state that the red LED is lighted, the robot starts to read the received message.

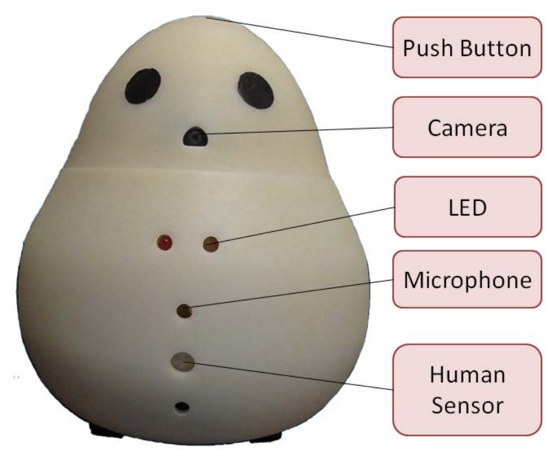

Fig. 7. SNS Agency Robot (Original robot)

Fig.8 shows the inside of the original robot. Raspberry Pi3, a single board computer, is embedded. Behind this, an USB type Wi-Fi router is connected, which realizes Wi-Fi connection environment. Fig.9 shows a hardware configuration of the original robot. Since Raspberry Pi 3 has standard input and output devices for computers so that it can be connected with various devices and easily recognize devices by a plug and play function. Voice and videos are output from display monitors, external devices by HDMI of Raspberry Pi 3. Fig.10 shows the system configuration of the original robot and a cloud service. It is basically the same as the system configuration of PaPeRo i shown on Fig.3, however it is specialized only for a conversation function.

D. Bilingual SNS Agency Robot (RoBoHoN)

The spread of smartphones and SNS enables communication among people beyond time and place. However, there are a certain number of people who are not able to use these communication methods due to physical disabilities. Although the disabilities are caused by various reasons, some have with their hands, some with their eyes due to intractable diseases. Some of them cannot share their anxieties or worries because there are only a few who suffer from the same diseases or medical specialists within their countries. Therefore, we propose the Bilingual SNS Agency Robot for those who have difficulty operating smartphones or tablets due to physical disabilities [7]. This voice user interface embedded robot enables them to communicate with others who suffer from the same disease or medical specialists in their mother tongues.

Fig. 11 shows the use case of an interactive communication between a person with a disability and the other who suffers from the same disease or a medical specialist (a communication person). The communication person transmits a text message written in English by SNS. After the message is translated into Japanese, the robot notifies the person with a disability by voice. His/her response in Japanese is recorded in a video and transmitted to the communication person. The video and the text, the contents of the response, are translated into English and displayed on the smartphone of the communication person. This use case is based on our developed SNS Agency Robot [6].

Fig.12 shows the system configuration of this system, which centers on SAKURA VPS (Intel® Xeon® CPU,CentOS 6.8), a cloud service. We used the RoBoHoN made by SHARP with a displaying function of images and videos. It is operable by voice recognition. Android (Android5.0) is mounted on RoBoHoN. Functions in a robot side, sub-system, is implemented as an Android application. Language Translator, a translation API provided by IBM, is used as a translation service, Cloud Speech-to-Text API by Google as a voice recognition service, and LINE Messaging API as a messaging service.

- In case a message is transmitted from a smartphone to RoBoHon, the entered text message on LINE is translated by the main system and transmitted to RoBoHoN. The voice synthesis function of RoBoHoN enables voice notification.

- In case a video is transmitted from RoBoHoN to a smartphone, the recorded video file by RoBoHoN is transmitted to the main system by HTTPS. A voice file which only the flac formatted voice is separated from is created. After the voice is converted to a text by voice recognition, the translation API translates it into English.

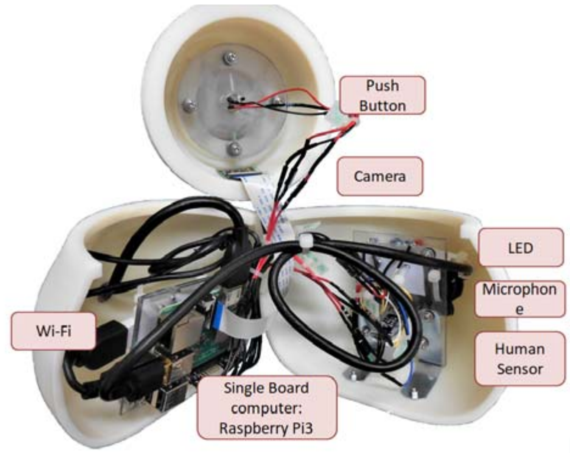

Fig. 8. Inside of the original robot

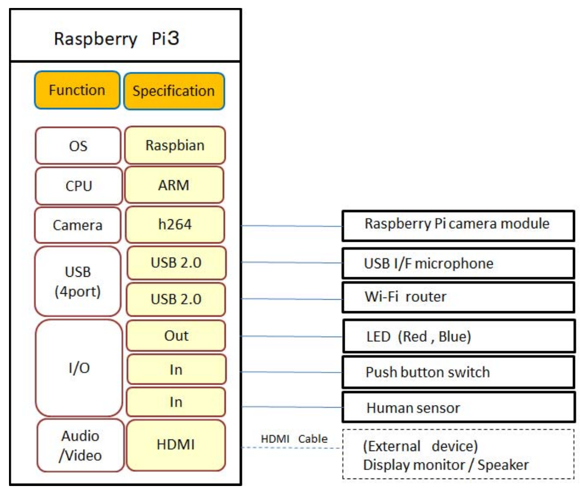

Fig. 9. Hardware configuration of the original robot

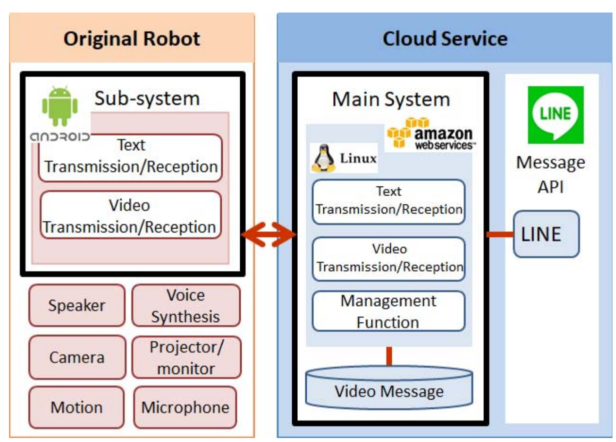

Fig. 10. System configuration of SNS Agency Robot (Original robot)

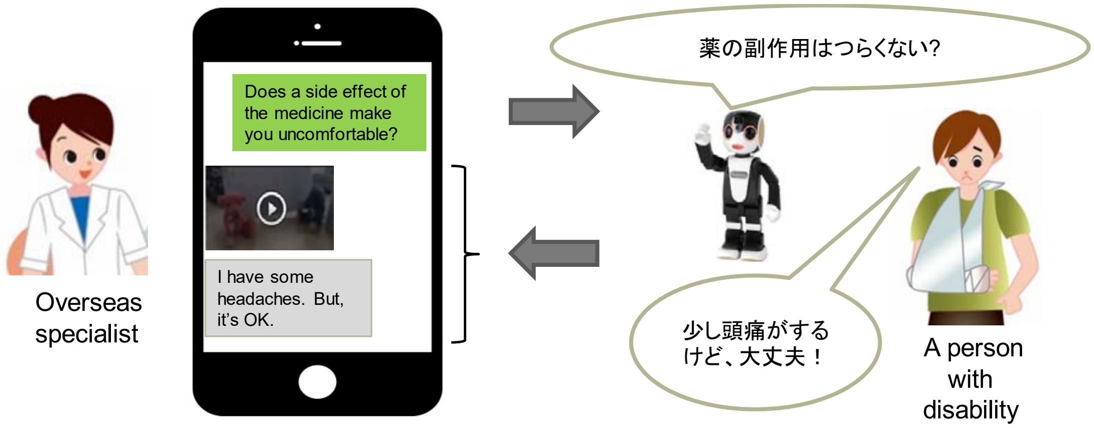

Fig. 11. Use case of Bilingual SNS Agency Robot (RoBoHoN)

E. Dementia Signs Detection Robot (PaPeRo)

As the aging society advances, the percentage of the elderly population living alone to total population of the elderly 65 years old and over tends to increase every year, men account for 13.3%, women for 21.1% in 2015 [1]. Since the elderly living alone have less contact points with society and tend to lead reclusive lives, there is a risk that dementia progresses without even knowing it. According to the study by Keio University, the social cost on medical and nursing-care of dementia reached 14.5 trillion yen in 2014, and estimated 24.3 trillion yen in 2060 [8]. Meanwhile, it is said that medicines at an early stage can delay the progress of dementia, which makes them live a healthy life for longer. If the early signs of dementia are detected and the progress is prevented, it leads to extend "mental health expectancy" of the elderly, which may largely reduce medical and nursing-care expenses in the whole society.

In order to be diagnosed with dementia, the elderly are required to see medical professionals or operate the application on tablet PC. That means the elderly go all the way to visit them to operate the devices and feel a lot of pressure to take tests, which have psychologically and physically imposed strain on them. Moreover, the questioning, rating, preparations of machines, and feedback to the elderly take a lot of time of medical professionals, who conduct the tests of dementia. For these reasons, missing an opportunity to take a test leads to the progress of dementia without realizing it. In order to detect and prevent dementia at an early stage, reducing the burden on those around them as well as the elderly is a key point.

Therefore, we developed the system that a communication robot objectively evaluates dementia based on natural conversation with the elderly living alone and inform suspected dementia to their families through social media [9][10]. This is implemented by adding dementia signs detection function to the SNS Agency Robot (Fig. 13). The detection function of "impairment of cognitive function" and "disability of functioning" is implemented as objective evaluation to detect signs of dementia. At present, the demonstration experiment of "impairment of cognitive function" has been hold at Nagasaki University Hospital. A newly added function to the SNS

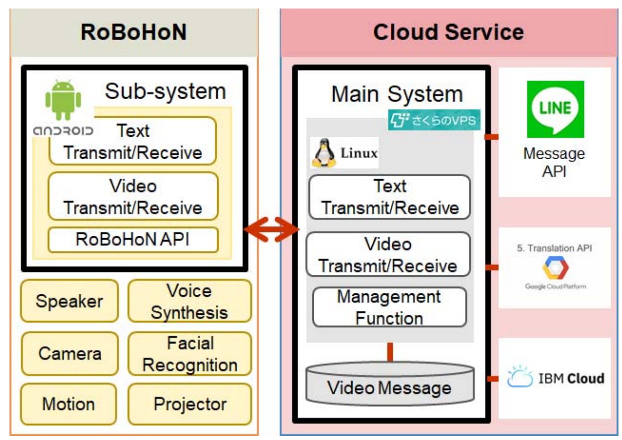

Fig. 12. System configuration of Bilingual SNS Agency Robot (RoBoHoN)

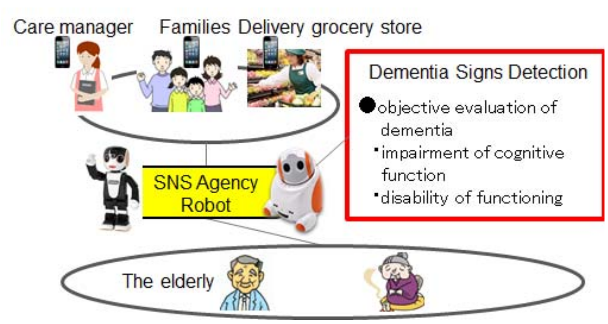

Fig. 13. Built-in projector of RoBoHoN

Agency Robot is an interactive dementia diagnosis function. The results of questioning to the elderly evaluate "impairment of cognitive function" objectively by linking this function to evaluation/scoring function of PaPeRo i (Fig. 14). For example, it calculates scores whether the elderly's answers to nine questions, such as "today's date" are correct or not. The elderly's speech is converted to a text by Google Cloud Speech

API and then processed by IBM Watson/Assistant. Specifically, as contents of the elderly's speech are analyzed, the subsequent suitable questions are chosen depend on answers of the elderly. Google Cloud Speech API is a function to convert a voice to a text by machine learning and able to recognize over 110 languages and dialects.

We performed two evaluation experiments with 10 subjects (male: 6, female: 4) from 60 years old to 78 years old who are patients of Nagasaki University Hospital. They do not develop dementia.

- (1)Cognitive function score automatic calculation using the prototype
- (2)Cognitive function score manual calculation based on medical doctor interview according to the conventional way

We performed both experiments at Nagasaki University Hospital from November 14, 2018 to January 31, 2019. First, we conducted the experiment of (1), then (2) one week later. Each experiment has been executed at the dispensary room of Hospital (Fig. 15). After the experiments, we evaluated a precision of (1) comparing to the results of (2) as correct answers (Fig.16).

The average precision of the prototype system was 85% comparing to the conventional method. The precision of the prototype system is inferior to the conventional method from these experimental results. The main reason was the voice recognition accuracy. Therefore, by improving the accuracy of the voice recognition, we can expect performance enhancement of the prototype system. We confirmed that the prototype system and the conventional method had a high correlation of 0.93 as Spearman's rank correlation coefficient. We confirm that the purpose of the dementia screening system can be accomplished even if an absolute difference with the conventional method was not buried. That is why we could suggest to the elderly to go to see a medical doctor, if the cognitive function score informed by the Dementia Signs Detection Robot would relatively worsen.

F. *Stroke Signs Detection Robot(RoBoHoN)*

Fig. 5 shows the use case of the SNS Agency Robot between a girl and her grandfather. By adding a function to detect signs of a stroke to this robot, we aimed to realize a

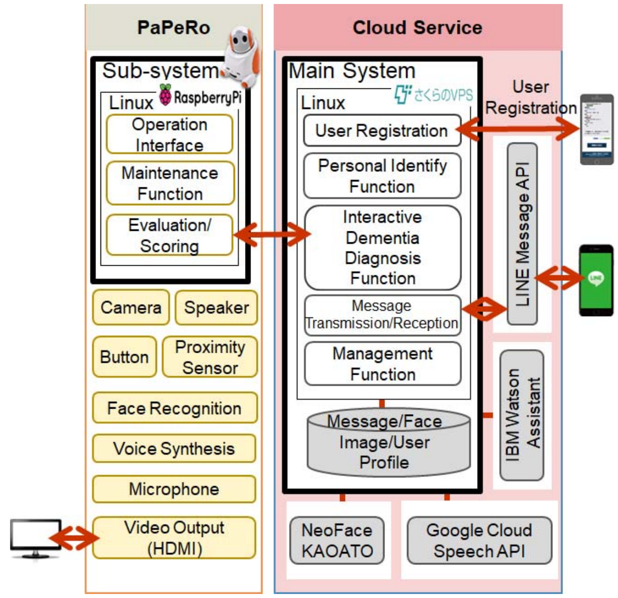

Fig. 14. System configuration of Dementia Signs Detection Robot (PaPeRo)

Fig. 15. Demonstration experiment at Nagasaki University Hospital)

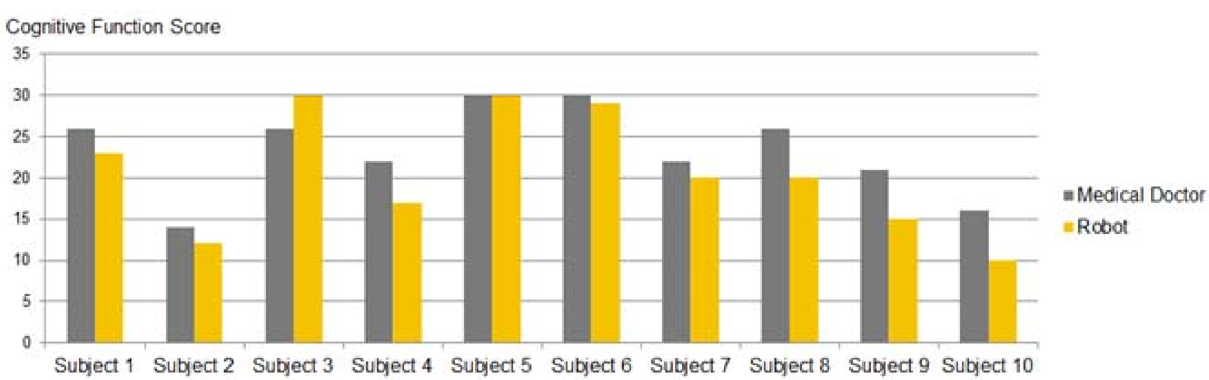

Fig. 16. Evaluation result of Dementia Signs Detection Robot (PaPeRo)

system that informs dangerous signs, which is often unconscious, to those around the elderly living alone. Specifically based on the Cincinnati Prehospital Stroke Scale (CPSS), it is realized by linking AI on a cloud with the robot at the elderly's (Fig.17). The CPSS tests whether both arms move equally, the patient uses correct words with no slurring, and the both sides of face move equally (Fig. 18). If any one of the three tests shows abnormal findings, the patient has a 72% probability of a stroke. Our system implemented two of the three CPSS scales.

We focused on two methods, "arm drift" and "slurring of speech" of the CPSS, which are implemented in the communication robot [11]. We used RoBoHoN made by sharp as a communication robot, since it can make voice conversation, record a video, and express how to drift arms by gestures. As the elderly person talks to RoBoHoN, it records the video of his/her movements after identifying an individual by face recognition. The recorded video is transmitted to a cloud service, "arm drift" and "slurring of speech" are checked. If it shows any abnormal findings, the result is transmitted to the elderly's families, who live apart.

This system is composed of sub-system on RoBoHoN (Android 5.0) and main system on the cloud service (CentOS 6.9) (Fig. 19). On sub-system, a gesture motion function to recognize "arm drift" and "slurring of speech" is newly added to the existing ones of the SNS Agency Robot. The specific functions are as below.

- RoBoHoN records the video of the elderly person as it holds his arms straight out in front according to the voice instruction.
- RoBoHoN records the video of the elderly person after instructing him/her to repeat a particular phrase in voice instruction.

In addition to the existing functions of the SNS Agency Robot, functions to recognize an individual (personal identification), capture a still image/voice from the video (still image capture), and detect signs (signs detection) are new in on the main system. NeoFace Cloud by NEC is used for personal identification. The captured still image automatically detects the position of the elderly's wrist by using OpenPose [12], which has a function to visualize the human pose by deep learning. The detected position information is stored in a Wrist Position Information DB. The function to detect signs recognizes the abnormalities by collating with the stored wrist position information. Google Cloud Speech API, implemented by neural network, converts the captured voice to a text. The existence of abnormalities is detected by whether a particular phrase is correctly pronounced or not as the function to detect signs. If any abnormalities in "arm drift" or "slurring of speech" are detected, an alarm is transmitted as a text message with a still picture via SNS, which is the existing function of the SNS Agency Robot. LINE, a major instant messaging exchange application in Japan, is used for SNS.

We evaluated the proposed system by the experiment that assumed a university student a subject. Regarding "arm drift", the still image is captured from the recorded video by RoBoHoN and OpenPose enables to detect the wrist position online, as Fig. 20 shows. In terms with the accuracy, we

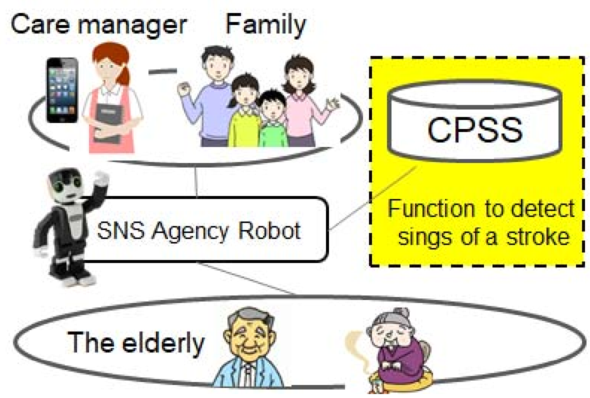

Fig. 17. Implementation of Stroke Sings Detection System

Fig. 18. Cincinnati Prehospital Stroke Scale

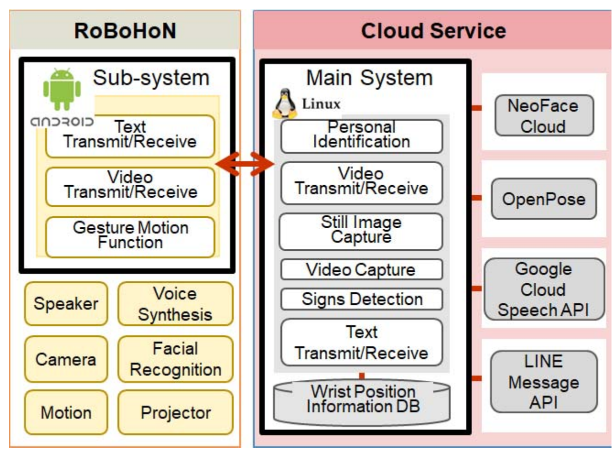

Fig. 19. System configuration of Stroke Sings Detection System

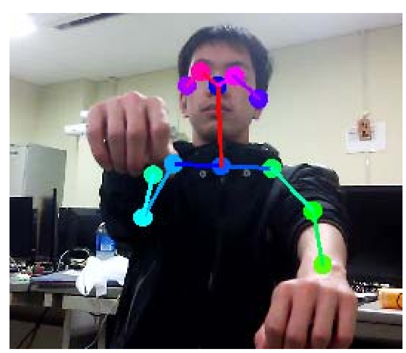

Fig. 20. Result by OpenPose

compared the angle from a neck to a wrist (0° ~20°) between an actual value and an estimated result by the proposed system (Table I). The difference was only less than 0.6%. As for "slurring of speech", the accuracy of the proposed system was 83 %. This showed that the system can check whether the repeated phrase is the same as RoBoHoN instructed or not. The experiment proved that the alarm with still pictures was transmitted via LINE in case any abnormalities are detected (Fig. 21).

TABLE I ESTIMATED ANGLE ACCURACY

Angle from neck to wrist	0°	10°	20°
Proposed system (cos value)	0.9995	0.9860	0.9341
Actual value (cos value)	1	0.9848	0.9397
Difference (%)	0.05	0.12	0.59

Fig. 21. LINE notification image

III. SERVICE-ORIENTED SOFTWARE DESIGN MODEL

Fig. 22 shows the extracted common framework after system configurations of our developed examples mentioned in Chapter II are abstracted. It is characterized by realizing the process with the use of several external cloud services or different kinds of robots for the main system of the SNS Agency Robot. Fig. 23 shows the refinement of this common framework based on a model-View-Controller (MVC) framework. Each robot corresponds to View. Operating each device such as a camera, a speaker, a microphone, and a human sensor realizes a user interface for the elderly or persons with disabilities. The main system of the SNS Agency Robot consists of Controller and Model. Controller is in charge of a control part and controls the entire functions by ordering to View or Model. Model has two functions of an internal process and an external API. An internal process controls each function that a robot system should provide. For example, it is the function to assess the degree of dementia in detecting the signs,

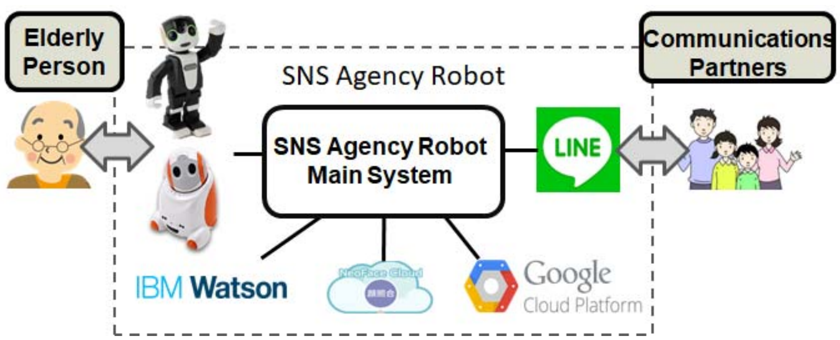

Fig. 22. Common framework for Communication Robot

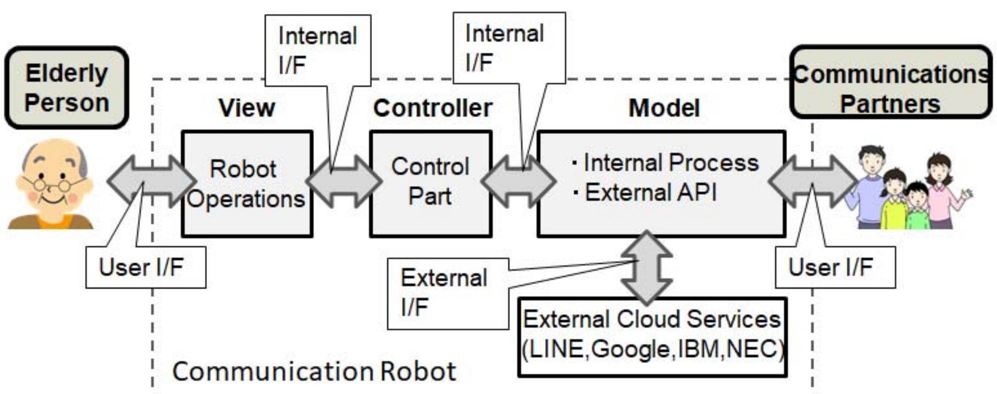

Fig. 23. Service-oriented software design model for communication robots

or detect abnormal arm drift in detecting signs of a stroke. An external API plays a function in using external cloud services. For example, it is an API for message exchange via LINE, voice recognition of Google, or IBM Watson API for support of natural conversation. An external API defines an external interface with external cloud services. This internal processes or an external API also provides a user interface for communication partners who communicate with the elderly or persons with disabilities. Moreover, an internal interface is defined between View and Controller, Controller and Model.

The independency of each component is improved by defining an interface between each internal process or between an internal process and an external cloud service. The following four merits can be expected by this defining.

- flexibility: an external cloud service including a robot is selectable according to a purpose.
- productivity: a testing part can be localized if there are not any changes required for adding functions in interfaces among each components.
- extensibility: new functions according to a purpose can be added by addition of an internal process.
- maintainability: even if an API specification is changed, only modification of an external API will suffice.

IV. Conclusion

In this paper, we proposed a service-oriented software design model for human-friendly communication robots by analyzing the common framework from examples of our developed communication robots. The characteristic of this model is that an interface is defined between each component based on a MVC framework. This enables us to efficiently develop communication robots having excellent flexibility, productivity, extensibility, and maintainability to meet various purposes.

We also referred to usefulness to detect sings of dementia or a stroke as the application service example of the communication robot for the elderly. In future work, we work on a method to detect sings of lifestyle-related diseases such as diabetes by a communication robot using this software design model. We aim not only to be human-friendly, but also to enhance the quality of human life.

Acknowledgment

This work was partially supported by Strategic Information and Communications R&D Promotion Programme (SCOPE).

References

[1] Cabinet Office Japan: Annual Report on the Ageing Society [Summary] FY 2019, https://www8.cao.go.jp/kourei/english/annualreport/2018/2018pdf_e.html

[2] ZOJIRUSHI, http://www.mimamori.net/

[3] Toru Kobayashi, Taishi Miyazaki, Rinsuke Uchida, Hiroe Tanaka, Kenichi Arai, "Social Media Agency Robot for Elderly People," Journal of Information Processing Vol.26, pp.736-746, Oct. 2018.

[4] " Introducing the IBM Watson Natural Language Classifier". https://developer.ibm.com/watson/blog/2015/07/10/the-ibm-watson-natural-language-classifier/

[5] "PaPeRo". https://en.wikipedia.org/wiki/PaPeRo "Office help and training". https://support.office.com/en-us/

[6] Toru Kobayashi, Koushin Kuriyama, and Kenichi Arai, "SNS Agency Robot for Elderly People Realizing Rich Media Communication," 2018 6th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering, Apr. 2018.

[7] Toru Kobayashi, Nobuyuki Yonaga, Tetsuo Imai, Kenichi Arai, "Bilingual SNS Agency Robot for Person with Disability," 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), Oct 15-18, 2019.

[8] Ryo Shikimoto, Mitsuhiro Sado, Masaru Mimura, "The social costs of dementia in Japan: Focusing on the informal care cost," Brain and Nerve, Vol.68, No.8, pp. 939-944, 2016.

[9] Toru Kobayashi, Taishi Miyazaki and Kenichi Arai, "Dementia Diagnosis System Using Social Media Agency Robot," The 11th International Conference on Mobile Computing and Ubiquitous Networking,October 5-8, 2018.

[10] Toru Kobayashi, Naohiro Sameshima, Testuo Imai, Kenichi Arai, Tomoki Watanabe, Teiichiro Miyazaki and Akira Tujino, "Dementia Screening System based on SNS Agency Robot," 2019 IEEE 9th International Conference on Consumer Electronics (ICCE-Berlin), Berlin, Germany, Sep 8-11, 2019.

[11] Toru Kobayashi, Kenichi Arai, Tetsuo. Imai, Koushin. Kuriyama, Satoshi. Ichimaru, and Tomoki. Watanabe, "Stroke Signs Detection Sys tem by SNS Agency Robot," 2020 IEEE International Conference on Consumer Electronics (ICCE), 2020.

[12] Sen Qiao, Yilin Wang, Jian Li, "Real-time human gesture grading based on OpenPose," 2017 10th Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Oct. 2017.