

A Smart Home Control System based on Context and Human Speech

Yoonseon Han*, Jonghwan Hyun[†], Taeyeol Jeong[†], Jae-Hyoung Yoo[†], James Won-Ki Hong[†]

* Division of IT Convergence Engineering, POSTECH, Pohang, South Korea
Email: {seon054}@postech.ac.kr

[†] Department of Computer Science and Engineering, POSTECH, Pohang, South Korea
Email: {noraki, dreamerty, styoo, jwkhong}@postech.ac.kr

Abstract—Traditionally, home appliances are controlled by switches or remote controllers manually. In order to control the appliances, a user have to be near the switches or carry remote controllers all the time. To alleviate this inconvenience, this paper proposes a control system for home appliances using human speech and context information. We present an architecture of the control system and a process to control home appliances composed of three steps: speech detection, speech recognition, and command execution. To validate our system, we show two useful case studies by testing it at POSTECH smart home and UPnP-OSGi simulator.

Keywords—Context Information, Home appliance control system, Ontology, Smart Home, Speech Recognition

I. INTRODUCTION

Nowadays, most home appliances such as TV, light bulbs, and air-conditioners are controlled by switches or remote controllers manually. In order to control switch based appliances, a user has to be near the switches. To overcome this inconvenience, there have been many researches to use human speech to control home appliances [1]–[4].

Some inherent problems of human speech originated from human perceptions and social customs make it difficult to realize a control system for home appliances using human speech. A representative problem is missing of necessary parameters, which is needed for a computer system to understand the meaning or to identify objects. These problems are not an important issue for human communication because all participants generally allow these phenomena by some tacit agreements. This ambiguity in human speech prevents a control system using human speech to understand the meaning. To overcome the inherent problems of human speech, we have employed context information to know the situation of the user and the environment. The context information provides clues to interpret human speech as a home appliance control command.

In this paper, we propose a control system for home appliances using human speech and context information which is monitored from sensors around a user at home or inferred from previously monitored context information. The process of our home appliance control system is composed of three steps: speech detection, speech recognition, and command execution. The first step is speech detection from acoustic sound such as

noise and music by monitoring the environment using several microphones. In the second step, the recorded sound is passed for the speech recognition to analyze an exact meaning of the speech and to extract necessary parameters for controlling home appliances. The last step, command execution, identifies the requested command from the speech and controls home appliances. If the request extracted from human speech is not enough to commit a command to control a home appliance, our control system integrates context information with the uncompleted command to infer unspecified parameters such as location of the requested appliance. To validate our system, we have deployed the system into two test environments: POSTECH smart home and UPnP-OSGi simulator. It shows that our proposed system enhanced the user convenience to control home appliances compared to use only human speech.

II. RELATED WORK

In this section, we present previous work on developing a control system for home appliances using human speech. The first step of all approaches is recognizing human speech or voices from home environmental situations [1], [2]. The main aim of those works is to efficiently distinguish human speech dialogs from environmental acoustic sound. The next step is extracting a user command from ordinary dialogs by analyzing human speech grammatically and semantically [3], [4].

Context information is defined as any information that can be used to characterize the situation of an entity [5]. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application. The fundamental idea of context-awareness in home environments is to integrate the natural human environment with a computer system to improve the quality of services and to provide new types of services. The researches on context-awareness includes developing models to express context information, seamless service convergence, and efficient resource management [6]–[8].

III. HUMAN SPEECH DETECTION

The speech detection module is the starting point of the proposed system. This module monitors the sound of all microphones in order to detect the presence of speech, and

records the sound when it detected human speech. Human speech can be detected by analyzing time and frequency domain features and silence ratio [9]. The recorded sound is passed to the speech recognition module to extract the structure and meaning of the speech. While passing the recorded sound, we also transfer the index number of a microphone to detect user location for the command executor.

IV. HUMAN SPEECH RECOGNITION

Human speech is a complex phenomenon, a dynamic process without clearly distinguished parts. All modern descriptions of speech are probabilistic at some degree [10]. Human speech can be extracted from a continuous audio stream recorded by a microphone. To build up a human speech recognition module, we have used CMU Sphinx-4 framework [11]. Sphinx-4 is a flexible, modular and pluggable framework to develop applications for speech recognition written in Java. To improve the recognition accuracy and the utilization of human speech for controlling home appliance, we defined language model in Java Speech Grammar Format (JSGF) as described in Fig. 1. The role of this language model is to confirm that human speech is well structured as our system allows, and the speech is constructed with allowed vocabularies. By limiting the structure and vocabularies of human speech, recognition accuracy is improved. Moreover, our system easily utilizes human speech as a home appliance control command. The output of Sphinx-4 is imperative sentences like “turn on the light in the kitchen”. The output sentence is parsed into several tokens as we defined. We have three sets of tokens: command, object, and location token set. Table I shows possible words for each set of tokens. The output of speech recognition module is a tuple of tokens, <command, object, location>, and command and location token could be omitted.

```
public <controls1> =
(<lighting1> | <clock> | <volume> | <music>
| <air-conditioner> | <refrigerator>
| <washing-machine>) [<adv_location>];
public <controls2> = <lighting2>;

<location> =
kitchen | bedroom | dining room | bathroom;
<adv_location> =
( in the kitchen | in the bedroom |
in the dining room | in the bathroom);

<lighting1> =
(turn) (on | off) [the] (light | bulb);
<lighting2> =
(turn) (on | off) [the] [<location>] (light | bulb);
<clock> =
(read | tell) [me] [the] (hour);
<volume> =
(turn) (down | up) [the] [<location>] (tv | music player) (volume);
<music> =
(start | stop | pause | resume) [the] (music player);
```

Fig. 1. A part of the language model for speech recognition

TABLE I
THE SET OF WORD TOKEN FOR HOME APPLIANCE CONTROL

Command tokens	turn on, turn off, turn up, turn down, increase, decrease, play, start, stop, pause, resume, check, read, change, move next, move previous
Object tokens	light, light bulb, TV, music player, air, air conditioner, fridge, refrigerator, washing, washing machine, hour, clock
Location tokens	kitchen, bedroom, living room, sitting room, bathroom, dining room

V. USER COMMAND EXECUTION

The last step is to execute requested command obtained from human speech. To control home appliances with a proper action, we have to figure out exactly what the user’s intention is, and consider the situation which the user and home appliances faced. The role of command executor is to determine which action is the most appropriate by considering human speech and the situation.

Fig. 2 shows the internal architecture of the command executor which consists of four sub-modules. Environment Monitoring Interface is a communication interface with sensors and appliances. Using this module, the sensor values, which are used to determine context information and status of appliances, are collected. Statement Interface receives the result of speech recognition module. When the result is arrived, the module starts a inference process to execute a user command. Ontology Framework allows to use an ontology model and inference rules which is essential for expressing and storing context information and for inference process to find missing parameters of human speech. As ontology framework, JENA framework, a Java framework for building Semantic Web applications [12], is used. It provides a programmatic environment for RDF, RDFS, OWL and SPARQL. It also includes a rule-based inference engine. Command Generator generates a device-specific command for home appliances. The specific command varies depending on the type or vendor of appliances. To address diversity of appliances, this module translates a command tuple into a device-specific command.

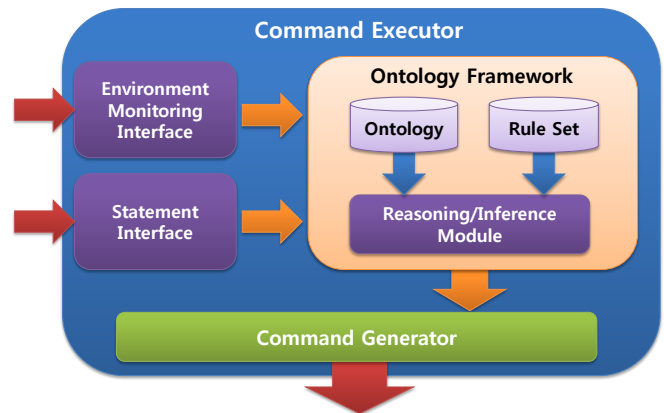


Fig. 2. The architecture of command executor

A. Designing Ontology Model

The concept of ontology is introduced in a semantic web area. It aims to make web resources more readily accessible to automated processes by adding information. Usually, the term “ontology” is defined as the specification of conceptualizations, used for helping programs and human to share knowledge. In our system, an ontology model is employed to express context information and the status of home appliances. Fig. 3 shows a part of the ontology model to control home appliances. To design an ontology model, we used OWL DL as language using protégé - OWL tool.

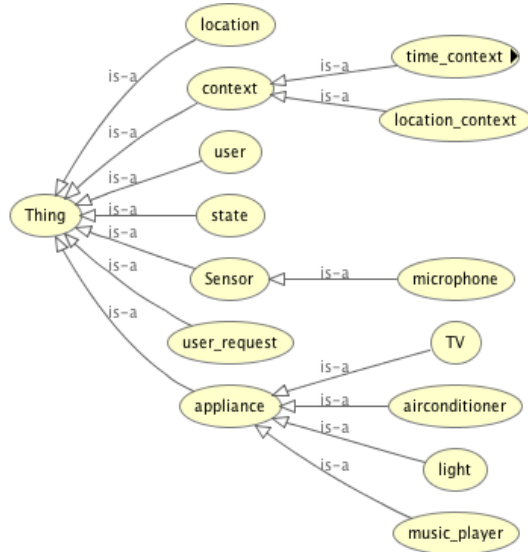


Fig. 3. A part of the ontology model for home appliance control

B. Specifying Inference Rules

A set of inference rules describes how to use the knowledge to infer new information. Actually, the inference rule is first-order logic which can derive more concrete properties. An inference rule makes possible to determine unknown factors based on known factors. To infer new knowledge based on ontology, we have to specify how to use existing knowledge. The language used for specifying inference rules is a JENA general inference rule. Fig. 4 shows an example of inference rule to identify the requested home appliance from a output of speech recognition module.

C. Inferencing missing parameters

The output of speech recognition module is a tuple, <command, object, location>, and command and location tokens could be omitted depending on situation. The role of ontology framework is to infer this missing parameters by utilizing context information.

- <none, object, location>: toggles requested appliances states if the appliance has two mutually exclusive states such as on/off and start/stop.

```
@prefix smart_home:
<http://www.semanticweb.org/ontologies/2011/11/smart_home.owl#>.
@include <RDFS>.

[objectIdentification:
(?req rdf:type smart_home:user_request),
(?req smart_home:has_request_object_string ?rstr),
(?voca smart_home:has_name ?rstr),
(?obj smart_home:has_voca ?voca),
(?req smart_home:has_stated_location ?loc),
notBNode(?loc),
(?obj smart_home:located_in ?loc)
-> (?req smart_home:has_request_object ?obj),
print(the requested device ?obj .)]
```

Fig. 4. An example of inference rule

- <command, object, none>: the current user location is applied for a missing appliance's location to identify a unique appliance.
- <none, object, none>: toggles the state of the requested appliance in same location with the user.

All these strategies to infer missing parameters are expressed by a set of inference rules.

VI. CASE STUDY

We developed the proposed system in a Java and deployed the system into two test environments: POSTECH smart home and UPnP-OSGi simulator. In this section, we describe details of each test environment and show that our system works correctly and improves user convenience for controlling home appliances.

A. Case Study 1: POSTECH Smart Home

POSTECH had set up a smart home environment in the campus to support research activities. In the smart home, many sensors and device controller are installed to monitor and control the environment. Further, there exist many appliances such as TV, air conditioners, and light bulbs. To show the feasibility of our system, we integrate the system with a Power Line Communication (PLC) based switch box manufactured by PlanetINT. The PLC based switch box provides many functionalities such as switch on/off for appliances, metering energy consumption, and checking the status of connected appliances. By integrating our system with PLC based switch box, we realized a control system based on human speech.

We have installed the proposed system into the smart home with PLC based switch box. This switch box communicates with our system using serial port. Fig. 5 shows POSTECH smart home environment. By generating commands for the PLC based switch box, we showed that our system can control real home appliances through human speech.

B. Case Study 2: UPnP and OSGi Simulator

OSGi-UPnP simulation environment allows deployment of various types of home appliances with a rich set of commands. Universal Plug and Play (UPnP) is a protocol stack that aims at self-care environments like residential and small

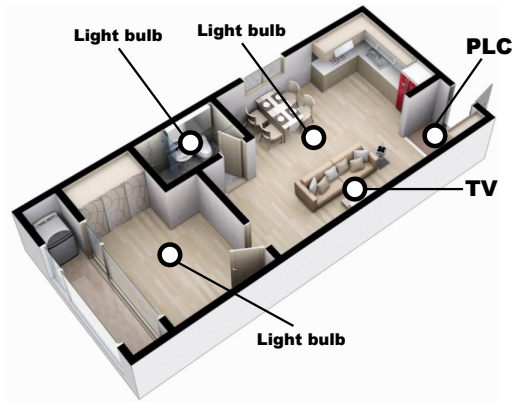


Fig. 5. The POSTECH smart home environment

office networks. Open Services Gateway Initiative (OSGi) is a well-known Java module system and service platform. It exposes UPnP devices as the services that can be monitored and controlled in home network. We selected OSGi-UPnP for simulation to take advantage of industry standards conformance, ease of simulated devices development, and eventing support by nature. By modifying control points and developing virtual UPnP devices, we can quickly establish a useful simulation environment for smart home. A control point delegates commands from our system to corresponding appliances. Using this simulation environment, we showed that our system successfully generates device-specific commands from human speech. Fig. 6 shows the appliances deployed in UPnP-OSGi simulation environment.



Fig. 6. UPnP-OSGi simulation environment

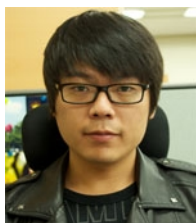
VII. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a system to control home appliances using human speech and context information. The system is composed of three modules: speech detection, speech recognition, and command executor. Moreover, we introduced ontology model to deliver proper services by considering context information. To show the feasibility, we have developed the system and deployed it into the POSTECH smart home and the UPnP-OSGi simulator. As a result, we have shown that our system has worked correctly at both test environments, and improved user convenience by supporting context-awareness.

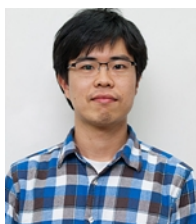
The speech recognition accuracy is still quite low to extract all user requests from ordinary human speech dialogs. As future work, we will improve the accuracy of speech recognition. Another important issue is the heterogeneity of home appliances that the providing services and command set is different by types and manufactures. To address this heterogeneity, we will develop a more general ontology model and an advanced translation mechanism.

REFERENCES

- [1] B. Lecouteux, M. Vacher, and F. Portet, "Distant speech recognition for home automation: Preliminary experimental results in a smart home," in *Speech Technology and Human-Computer Dialogue (SpeD)*, 2011 6th Conference on, may 2011, pp. 1–10.
- [2] M. Vacher, N. Guirand, J.-F. Serignat, A. Fleury, and N. Noury, "Speech recognition in a smart home: Some experiments for telemonitoring," in *Speech Technology and Human-Computer Dialogue, 2009. SpeD '09. Proceedings of the 5-th Conference on*, june 2009, pp. 1–10.
- [3] M. Denkowski, C. Hannon, and A. Sanchez, "Spoken commands in a smart home: an iterative approach to the sphinx algorithm," in *Proceedings of the artificial intelligence 6th Mexican international conference on Advances in artificial intelligence*, ser. MICAI'07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 1025–1034.
- [4] I. Mporas, T. Ganchev, T. Kostoulas, K. Kermanidis, and N. Fakotakis, "Automatic speech recognition system for home appliances control," in *Informatics, 2009. PCI '09. 13th Panhellenic Conference on*, sept. 2009, pp. 114–117.
- [5] A. K. Dey, "Understanding and using context," *Personal Ubiquitous Comput.*, vol. 5, no. 1, pp. 4–7, Jan. 2001.
- [6] A. Meliones, D. Economou, I. Grammatikakis, A. Kameas, and C. Goumopoulos, "A context aware connected home platform for pervasive applications," in *Self-Adaptive and Self-Organizing Systems Workshops, 2008. SASOW 2008. Second IEEE International Conference on*, oct. 2008, pp. 120–125.
- [7] J. Zhao and L. Lu, "Smart home model based on ontology and context-awareness," in *Granular Computing, 2008. GrC 2008. IEEE International Conference on*, aug. 2008, pp. 830–833.
- [8] M.-Y. Weng, C.-L. Wu, C.-H. Lu, H.-W. Yeh, and L.-C. Fu, "Context-aware home energy saving based on energy-prone context," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, oct. 2012, pp. 5233–5238.
- [9] G. Lu and T. Hankinson, "A technique towards automatic audio classification and retrieval," in *Signal Processing Proceedings, 1998. ICSP '98, 1998 Fourth International Conference on*, vol. 2, 1998, pp. 1142–1145 vol.2.
- [10] B. H. Juang and L. R. Rabiner, "Automatic speech recognition - A brief history of the technology development," *Elsevier Encyclopedia of Language and Linguistics*, 2005.
- [11] W. Walker, P. Lamere, P. Kwok, B. Raj, R. Singh, E. Gouvea, P. Wolf, and J. Woelfel, "Sphinx-4: A flexible open source framework for speech recognition," *Tech. Rep.*, 2004.
- [12] B. McBride, "Jena: a semantic web toolkit," *IEEE Internet Computing*, vol. 6, no. 6, pp. 55–59, 2002.



Yoonseon Han is a Ph.D. student in the Division of IT Convergence Engineering at POSTECH. He received his B.S. in the Department of Computer Science & Engineering at POSTECH in August 2009. His research interests are knowledge modeling and autonomic network management.



Jonghwan Hyun is a Master & Ph.D combined student in the Department of Computer Science and Engineering, POSTECH. He also received his B.S. degree in the Department of Computer Science and Engineering, POSTECH. His research interests include network traffic monitoring and analysis and Internet of Things (IoT).



Taeyeol Jeong is a Master & Ph.D combined student in the Department of Computer Science and Engineering, POSTECH. He received his B.S. in the Department of Electrical and Computer Engineering from Sungkyunkwan University in 2011. His research interests include Software-Defined Network and Network Function Virtualization.



Jae-Hyoung Yoo is a Research Associate Professor in the Department of Computer Science and Engineering at Pohang University of Science and Technology. He received his BSc and MSc degrees in Electrical Engineering from Yonsei University, Korea, in 1983 and 1985, respectively, and the PhD degree in Computer Engineering from Yonsei University, Korea, in 1999. He had worked for KT (Korea Telecom), the largest telecommunications company in Korea from 1986 to 2012, where he was responsible for developing management system for various network systems, such as PSTN, ATM, and Internet. He has been an active volunteer in various committees in IEEE NOMS, APNOMS and KICS KNOM. His research interests include network management and security, software-defined networking (SDN), and OpenFlow.



James Won-Ki Hong is a Professor in the Department of Computer Science and Engineering at Pohang University of Science and Technology. He had worked as the Chief Technology Officer and Senior Executive Vice President for KT (Korea Telecom), the largest telecommunications company in Korea from March 2012 to February 2014, where he was responsible for leading the R&D effort of KT and its subsidiary companies. He was Chairman of National Intelligence Communication Enterprise Association and Chairman of ICT Standardization Committee in Korea. He co-founded and is currently Executive Director of SDN/NFV Forum in Korea. His research interests include network innovation, such as software-defined networking (SDN) and network function virtualization(NFV), cloud computing, mobile services, IPTV, ICT convergence technologies (e.g., Smart Home, Smart Energy, and Healthcare), and Internet of Things (IoT). James had served as the Head of the Department of Computer Science and Engineering, Dean of Graduate School for Information Technology, Director of POSTECH Information Research Labs (PIRL), and Head of the Division of IT Convergence Engineering at POSTECH. He was also co-founder and CTO of Netstech, a Silicon Valley-based startup developing network integrated ultradense, blade servers from 2000 to 2002. Over the past 20 years, James has been an active volunteer in various committees in IEEE, ComSoc, and KICS. He has served as Steering Committee Chair of IEEE NOMS, IM, and APNOMS, as well as Chair of IEEE CNOM (ComSoc Committee on Network Operations and Management) and KNOM. He has also been serving as EiC of Wileys International Journal of Network Management (IJNM), an editorial member of the IEEE Transactions on Network and Service Management (TNSM), Elseviers Journal of Network and Systems Management (JNSM), and Journal of Communications and Networks (JCN). James received his HBSc and MSc degrees in Computer Science from the University of Western Ontario, Canada, in 1983 and 1985, respectively, and the PhD degree in Computer Science from the University of Waterloo, Canada, in 1991.