# Multimodal Interface for Mobile Cloud Computing

Hoon Jeong[1], Sungjin Kim[1,] Hana Do[1], Euiin Choi[1],Yeojin Jeong[2], Yongho Kang[2]

[1]Dept of Computer Engineering, Hannam University, Daejeon, Korea
{hjeong, kimsungjin04,dohana77, eichoi}@dblab.hannam.ac.kr
[2] R2soft Co.,LTD, Dunsan-Dong, Seo-Gu, Daejeon, Korea
{vatang88, kang}@r2soft.co.kr

***Abstract***:Mobile Cloud Computing has become as a new IT paradigm because of the growth of mobile device like smartphone and appearance of Cloud Computing environment. This mobile cloud environment provides various services and IT resources according to users' requests, so an effective providing of service and IT resources is required.Multi-modal is a technique which can convert information to suitable format for the optimum communication between the system and the user. We should represent as modalities to contexts with the user and the environment for improved interface in mobile environment. Also, we need to multi-modal technology for understanding contexts and providing suitable services. User's actions then are used as main information of context-awareness and multi-modal service. In order to constrain user's actions as basic ones and provide a good service by grasping each user's situation through context model, there needs a multi-modal interface and it makes for an agent with this to provide a multi-modal interaction service. Therefore, this paper presents a more improved interface using a multi-modal interaction.

***Keywords***: mobile cloud computing, Multimodal, Interface, Context-awareness, Monitoring-control system**.**

## I. INTRODUCTION

The market of mobile recently has been evolving rapidly and cloud computing is spreading into mobile as well. That is why mobile cloud computing is becoming a new issue today. Cloud computing is the computing that provides virtualized IT resources as a service by using Internet technology. In cloud computing, a user lends IT resources (software, storage, server, network) as needed, uses them, get a support of real-time scalability according to service load, and pays as he/she goes. Especially the cloud computing environment distributes IT resources and allocates according to user's request, so there should be a study on technology that manages these resources and deals with effectively. Multi-modal is a technology that looks for a variety of ways of utility, converts information into suitable formats for the system to let itself communicate optimally with user and transmits. It helps humans and computers interact by using 2 or more of input/output modality including a few of modes, aspects, and senses[1,2].

The direction for study of multi-modal can be divided into 3 big parts; first, it has an direction to replace natural movements like speaking, gazing, and acting to the way of input of equipment operation so as to use the study as a method that heightens the measurement of input method for each; voices, steady gaze of eyes, and movements, or overlapped ones and the precision of grasp of intention. Second, it is the study of the optimum interface of output equipment. While products are combined centered space and service purpose in the ubiquitous environment, existing services provide product-centered interfaces such that audio let people hear the music through speaker and TV transmits contents with images and sound at the same time.For instance, it is the study to develop integrated media service, not TV.Third, it aims to embody a multi-modal that knows the situation rightly, confronts appropriately, provide a visual, sound or which complicated interface and how much information[3].

There are various results from the first directed study among them, however, the results from the second and third whose technologies are able to integrated to products and applied to services are not good enough yet. So, we in this paper would like to discuss the monitoring-control system that manages and controls surroundings to fit user's requirements through analysis of user's modality in the side of service offering.

It is not good enough for the existing monitoring-control products to be utilized in ubiquitous computing environment that IT field pursues today.There are a lot of inconveniences such that it has a problem of limited space, a problem that it is ok to control in certain places, and a problem of supplement of personnel because a manager or a clerk related to needs to be involved.Thus, this paper proposes a improved monitoring-control system compared to existing ones by using a multi-modal interaction agent which applies suitable multi-

modal interaction services by catching user's situations through embodiment of context model.

## II. Related works

### 1.Multimodal interface

*A. Multimodal Framework*

Multi-modal interface means that the user can communicate easily and conveniently with the computer using compounded modalities such as speech, visual, tactile, etc. And multi-modal interface which is different with current uni-modal based interfaces should need to synchronize between each interface. Hence, W3C(World wide web consortium) defined Multimodal Interaction Framework which is composed of interaction manager, service session element, system environment element and application service element.The information which is inputted by the user using input interfaces such as speech, pen touch, keyboard, etc is sent to interaction manager which can generate semantic information, and then multimodal input data that is called EMMA(Extended Multimodal Annotation) is generated[4].EMMA can represent semantic information which is described by XML as well as all characteristics of input data. Thus, EMMA data which is generated by each modalitiesis integrated by interaction manager.Interaction manager provide to send data to application server in back-end through compositing each input EMMA data, and the result which is converted to the specialized language for applications is sent to the user[5].

*B. SCXML(State Chart XML)*

SCXML is a state machine that made by combining Harel State Tables with a event based state machine, and is represented to XML[6].Harel State Tables which is also included UML are state machine presentation model, and can represent to semantics about parallel states.Figure 1 shows SCXML which has a core role in interaction manager and multi-modal runtime which compose of languages based on XML.

Graphic interface is represented by HTML and speech interface is represented by VoiceXML.In case of implementing speech interface using telephone, CCXML(Call Control XML) is optionally used for a detailed control each other.SCXML is loosely coupled with interface and has state transition based on events, so has a structure that can easily add new various modalities instead of graphic and speech.

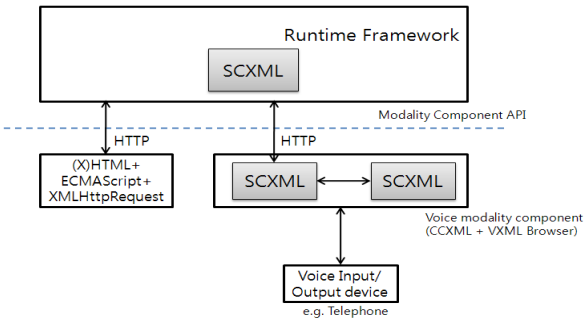


Fig. 1. Multi-modal Runtime Diagram

### 2.Context-awareness

*A. Gaia project*

Gaia project is a typical case among systems applying ontology and semantic web technology in ubiquitous computing environment[7].CORBA, traditional middleware, is used for a way of transmission between distributed objects in ubiquitous environment based on Gaia.Middleware based on CORBA, Java RMI and SOAP can be possible to communicate between a variety of objects, but they do not provide a method for semantic interoperability.In this project, machine study based reasoning is possible using ontology which is describe a CORBA middleware based context as a predicate.When each agent interacts with other agents in ubiquitous environment, they can share context information by aplying semantic technology.

*B. SOCAM(Service Oriented Context-Aware Middleware)*

SOCAM is proposed as a middleware for easy development of context-aware service in mobile environment[8]. Ontology is used in SOCAM as a method of various context information modeling, and context information model using ontology can provide semantic representation, context reasoning, context sharing, etc.Also, service-oriented middleware has been developed for supporting acquisition, detection,interpretation of context information between context-aware systems.Through context information modeling using ontology language, the system can share

the knowledge about context between other objects and can reason a high-level context from a low-level context.

## III. Composition of interface using user interaction agent

### 1.User Interaction

The user activity is used as core information of multi-modal and context-aware services in ubiquitous system.But, we need to grasp interaction level of the user activity which we can understand the user requirement for connecting services and activities.The user interaction of purpose and state level has a large scope of activity, so embodiment and personalization of services are difficult.However, basic activity is defined as the least unit of activity that has definite and concrete purpose among the user activities, and purpose of basic activity is possible to grasp by current technology and has proper interaction level for ubiquitous service. Hence, we have applied basic activities level to user interaction.

### 2.Interface

Context and context-awareness are explained through embodiment of each user's Context Cube and multi-modal interaction for ubiquitous service will be unfolded built on a basis of this information.Information of user's action, location, used equipment, and used time will be situations.It is also possible to know what purpose of action user is doing through the context bar composed by the schedule of the time. And it is able to provide multi-modal interaction service by inferring suitable service method by grasping other related features of action and environment.Figure 2 shows the composition of multi-modal interface.

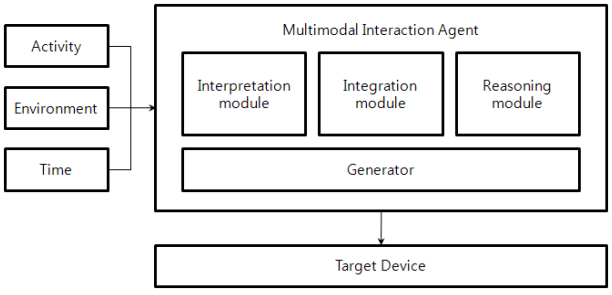


Fig. 2.Composition of Multi-modal interface

For example, let's suppose the situation that a user is sitting on a sofa in the living room at 8pm and an entertainment channel is on TV at the same time.Environment information of context bar at the same time tells you that the user follows the normal schedule by the schedule and action information is for leisure. As the user goes to the kitchen at this time, the volume of TV gets higher to help the user hear it and as the user still stays there after certain time, it searches an equipment that makes the user watch TV in the kitchen, chooses web pad, and provides an interaction service which telecasts TV broadcasting without cutting. In the case of passing information to user's device, it converts interface and provides information which is suitable for the device.

## IV. CONCLUSION

We in this paper designed components of interface which observes situations "at anytime at anywhere" using context-awareness technique and multi-modal technique and enables to offer multi-modal interaction service by inferring suitable service method. Context-awareness technique enables to judge the given monitoring-control environment or situation correctly and objectively and helps provide a service type suitable for the situation judged through context-awareness reflecting user's movements and modality of surroundings.

### ACKNOWLEDGMENT

This research was financially supported by the Ministry of Education, Science Technology (MEST) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation.

## REFERENCES

[1] Michael H. Coen (1998), Design principles for intelligent environment. In Proc. The Fifteenth National Conference on Artificial Intelligence(AAAI'98):547-554
[2] Laurence Nigay, Joelle Cutaz (1993), A design space for multimodal systems:concurrent processing and data fusion. In Proceedings of INTERCHI'93, ACM Press:172-178
[3] Hyun-Jeong Kim, Hyun-Jin Lee (2005), Ubiquitous Context-aware Modeling and Multi-Modal Interaction Design Framework. Korean Society of Design Science:273-282
[4] EMMA(Extended Multimodal Annotation) (2007), http://www.w3.org/TR/emma/,W3CCandidate

Recommendation
[5] W3C Emotion Markup Language Incubator Group, http://www.w3.org/2005/Incubator/emotion/.
[6] SCXML(State Chart XML) (2007), http://www.w3.org/TR/scxml/, W3C Working Draft
[7] M. Roman, C. K. hess, R. Cerqueira and A. Ranganathan et al (2002). Gaia: A middleware infrastructure to Enable Active Spaces. IEEE Pervasive Computing:74-83
[8] T. Gu, H. K. Pung, and D. Q. Zhang (2004), A Middleware for Building Context-Aware Mobile Services. In Proceedings of IEEE Vehicular Technology Conference:2656-2660 Vol.5