Smart Gesture Sticker: Smart Hand Gestures Profiles for Daily Objects Interaction

Ayman Atia, Shin Takahashi and Jiro Tanaka
Department of computer science
University of Tsukuba
Tsukuba, Japan
ayman@iplab.cs.tsukuba.ac.jp, shin@cs.tsukuba.ac.jp, jiro@cs.tsukuba.ac.jp

***Abstract*—In this research, we present a system that can support object gestures in ubiquitous environment. We propose various techniques to use any object around users for devices and appliances interactions. These techniques make user interactions more natural and intuitive. We propose technique to use any object around users for devices and appliances interactions. These techniques make user interactions in a more natural and intuitive. The experiment results lead us in designing an architecture that assists users to interact and customize gesture profiles for objects existing in their ubiquitous environment. We call it "Smart Gesture Sticker (SGS)". In this paper, we present the detailed architecture of SGS. Furthermore, we implement applications with SGS. We evaluate SGS performance by applying different scenarios and situations for users using it in everyday life. Our results show that SGS supports users to interact with their environment and allow users to discover new usage for objects around them.**

***Keywords*-Object gestures, Hand gestures, Interaction in ubiquitous environment.**

I. Introduction

Interaction with everyday objects is one of the challenging research areas in smart ubiquitous environments. Recently, smart applications that depend on user's activity recognition directly use everyday objects. Objects have a direct impact on the interaction in smart environments according to their shapes [1]. There are some gestures that done by users naturally while they hold objects unintentionally. Object gestures are defined as moving an object in the spatial space. In several cases, using hands to interact with devices is not applicable such as to check email while driving a car. The user is involved in another activity and his/her hands might not be free for interaction. Hand gestures might have some drawbacks such as the social acceptance or rejection in public space. It is not appropriate for a user to perform a big circle hand gesture in public areas. It will seems to others as someone is threating them. Gesture by itself is a production of users imaginations, so there are almost no standard gestures that could match the satisfaction of many users. In this paper, our target users, those who are interacting with various objects to perform actions in different situations. We consider different factors for situations where the user can exist in ubiquitous environments. A situation is defined as a combination of user's context parameters. The main context parameters can be shown in Figure 1.

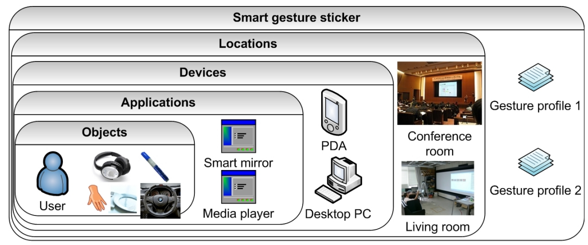

Figure 1. SGS main context parameters

In this paper, we propose a technique to use any object around users for devices and appliances interactions. The technique make user interactions more natural and intuitive. We developed an architecture called "Smart Gesture Sticker" (SGS) to facilitate user interaction with everyday objects in ubiquitous environment. Our architecture is designed to allow users to customize gestures for objects according to their context parameters. SGS packs the user's gestures and provides a customized interface in his/her environment. Gestures of users can be captured with many techniques such as using a camera [2]. The use of camera issue the privacy concerns from users about being captured. Consequently, we use a small attachable coin size wireless 3Daccelerometer sensor. Figure 2 shows the 3D accelerometer sensor used in this study. In order to develop a system capable of capturing gestures of users in ubiquitous environments, we have conducted a detailed study about object gestures.

II. Object gestures

SGS supports users in different situations by providing gestures for the user anywhere. The use of different objects inspires users for more intuitive interaction with the environment. We conducted a pre-experiment study that focuses on how users expect object gesture shapes. We perform an analysis that measure user satisfaction about their gestures when they create their gestures from scratch.

In this study, we choose objects with different shapes including human body parts so we can study the most intuitive and suitable objects that can make users interact with applications fast and accurate. We asked five subjects to do this pre-experiment. All subjects are aged from 25

Figure 2. 3D accelerometer sensor.

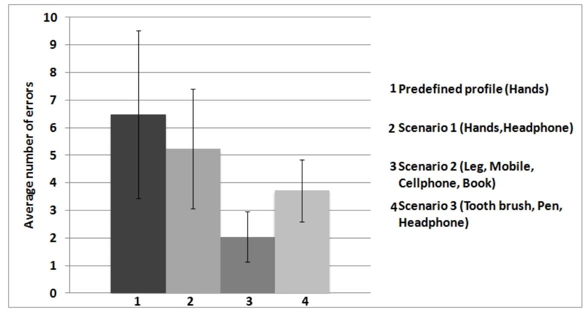

Figure 3. Pre-experiment results

to 33. At the beginning of the experiments, we show the users different 10 objects and they choose three to use in the experiments. Objects are (user's hands, user's legs, stylus board pen, bottle, book, cellular phone, tooth brush, umbrella, handy fan and wireless headphones).

A total of three scenarios to be done, two scenarios to be done in the living room while watching movies using media player and using two different objects. The other scenario is implemented in the bath room while interaction with smart mirror to control media player. The media player has nine functions that can be customized (Play, Pause, Next, Previous, Increase volume, Decrease volume, Mute, Stop and Close). In the first two scenarios, the subjects switch between the two objects and determine the preferred scenario to play. In the third scenario, one profile will be downloaded automatically in their context. First we calculate the number of conflict gestures done by users while they customize their gestures. Second we measured the accuracy of interaction with predefined gestures and the three customized gestures scenarios. Third we get some feedback from users about appropriate usage of objects for gestures. If the user enters a gesture that is not adequate. He/she should repeat the gesture up to eight trials.

A. *Preliminary results*

The results show that the majority of users 76% create their hand gesture without having any conflicts, and 16% had at least one conflict gesture. The average time to create object gesture was nine seconds. Some subjects show more consumed time up to 86 seconds because they preform very similar pattern gesture shapes that leads to conflicts.

The results show that when users get familiar with the system, they can generate gesture ideas such as using ear headphones to play and pause media player. They choose the gesture of taking off the headphone to pause the media player and putting up headphone to play action. They can create a smart headphone appliance through recognizing their current activities and command tasks towards those activities.

The results show that when users have a chance to create their own gestures from scratch or even tune up some existing gestures, they show more accurate and fast interaction with the interface. Figure 3 shows that the three scenarios done by users in different situations are all more accurate than the usage of predefined hand gesture profiles.

Comments and verbal feedback from the users about SGS system was also recorded. One user tried to use his leg for interaction and defining the eight functions. Since the gestures that can be done using the leg is limited, as the space for gesturing by leg is small. It was hard for the user to customize all the functions for the media player using his leg. Another user says that hitting and tapping gestures was easy to memorize. The results analysis show that the strength of doing the gesture with the object can affect the level of doing the command such as increasing or decreasing volume. Users says that gestures need some time to be learned. However, after learning, the usage of gestures was very similar to real-life situations. Customizing gestures were intuitive to use specially the scenario of using the cellular phone was very easy to understand.

We noticed that some users could make gestures like throwing in the air, which might have some social threaten if used in public space area. Some users uses a headphone object on his/her head, trying to do similar gestures pattern. They fail as there is a limited space to do gestures with headphones fixed on head.

We notice four main design principles for gesture customization and interaction interface based on our primary experiments and users feedback. First, the gestures should be fitting all situations of people implicitly. People can change their location, device, application or object they interact with. The ubiquitous environment should maintain well understanding of the gestures. Second, is the intuitiveness and natural use of the gestures. People should interact with devices using gestures as they interact with everyday objects. Third, people do not prefer to consume much time to learn and get trained for the gestures. This can cause frustration and make them bored from using the system. Hence, there is a need for an easy way to customize interface that can depend on minimal number of learning gestures and can support the ability to learn the gestures easily. Forth, the adjustment of a predefined profile is much easier for a beginner user than creating a profile from scratch.

III. SGS Overview

SGS system is functioned by tracking users basic context parameters: location, devices, applications, objects and users in different ubiquitous environments based on Ubi-Gesture infrastructure [3]. A subscribed user to the system enters some location and selects one of the devices he/she wants to interact with. User sticks the sensor to one of the objects by binding it with double side tape or bounding piece of cloth. SGS shows an interface on the device that allows the user to choose the applications he is permitted to run on this device. After the user selects application, SGS will loads the appropriate gesture profile for user in this ubiquitous environment. Users create new profiles or they can adjust gestures they performed in similar ubiquitous context.

In this research, we used a coin size small sensor that can be attached easily to various objects. It has a built-in 3D accelerometer, angular rotation rate sensor and temperature sensor. The sensor is a 20 g weight with dimensions of 39mm (w) x 44mm (h) and 12mm (d). The sensor sends data as a pattern of signals that represent the tilt of the sensor. The output of the accelerometer is a strip sequence of 3D points denoted by G, such that a point can be represented as $h_t\left\{a_x, a_y, a_z\right\}$. h_t is time stamp generated automatically by the sensor and a_x, a_y, a_y are the readings recorded from the accelerometer at time t.

A. Gesture recognition

We used k-mean clustering algorithm with three clusters to compare the distance between the stored object gestures. After the entered gestured is classified into one of the clusters, SGS applies DP-matching algorithm [4] to get similar gestures. The cost function for DP-Matching has been calculated as the Euclidean distance between the two 3D vectors. SGS system calculates the minimum value between the user's object gesture and all stored template gestures within the cluster. Each object has a normal functionality rather than being used for interaction. The user has to decide when he wants to use object for interaction and when he wants to use it naturally. The user must makes four successive shake gestures to start using the object for interactions and once again to stop capturing it gestures.

B. Selecting appropriate profile

SGS keeps tracks of a shared network log repository that can be accessed in ubiquitous environment. SGS load users context parameters through this log. Any user may create a profile with respect to his/her context. Next, SGS lists all objects existing in the user location and available applications to associate objects with. SGS searches for the most appropriate profile for the application according to the following arguments. Each context parameter has some attributes defining it. An example, objects exist in the same location might have some common attributes. On the desk in the office room the bottle and the cup might be similar in their usage. The authors of this paper put primary attribute values for those context parameters based on observations from the primary study. Table I shows the context parameters primary attributes. SGS will match the profiles with maximum similar attributes for context parameters by giving priorities for objects, device, location and user parameters respectively. We give users parameter lowest order for selecting profile to preserve the privacy of users. However, each object gestures have default gestures for applications to be selected at the initialization of the system.

Table I
SGS CONTEXT PARAMETERS ATTRIBUTES

Parameter	Sample
Objects	Long, Short, Big, Small, Fixed on hand, Fixed on head, Sphere shape, Stick shape and Cone shape.
Application	Media players, Presentation viewers, Browsers,Text editors.
Device	Portable, Fixed, Small display, Big display, Large projected display, Speakers exist and Speakers not exist.
Location	Office, Car, Indoor general, Outdoor general, Restaurant, Mall and Train station.
User	Beginner, Moderate, Expert (to use gestures), Computer science specialist, Other specialty.

Those attributes are changeable and each profile item can be classified by many attributes. A gesture profile example classified as (User is expert, Object attached to hands, Location is in public space).

IV. Smart Appliances

The SGS users create their own smart appliances by adjusting object gestures for context. We have three main applications to be run by SGS. The first application is media player application. We used Microsoft windows media player. Second, is photo viewer application that we developed to browse and apply effect on pictures. It has 15 functions (Start PTZ camera, Show image, Close application, Zoom out all pics, Zoom in all pics, Send picture to my home, Navigate right, Navigate left, Navigate down, Navigate up, Save captured image and close, Black white, Crop image, Take picture, Rotate image). Third, is PTZ camera protection application that allows user to control the camera directions and make some reactions in case a person felt being watched. The PTZ camera protector application has nine functions (Tilt up, Pan right, Tilt down, Zoom out, Pan left, Redirect to board, Zoom in, Close, Redirect to desktop).

Figure 4. (a) Smart headphone (b) Smart phone (c) Smart hands (d) Smart book

A. *User appliances*

We ask users to interact with the SGS and ask them to create their own gesture imaginations. Hence, we have classifeid the interaction with objects and applications into 4 smart scenarios.

- Smart headphone: The user created intuitive gestures for controlling their appliance. Putting headphones on the head mapped to play music file. Putting off the headphone on the desk pause the music file. Tilting head to the right and tapping headphone increase volume. Figure 4 (a) shows a user using headphone and tilt and tap gesture to control media player.
- Smart mobile phone: The user created in train station. The user map the gesture of play and pause media files as moving his mobile phone from right hand to left hand and vice versa. The user assigns the stop command as to put his mobile phone in his pocket. Figure 4 (b) shows a user using mobile phone in simulated train station seat.
- Smart hands: The user maps gestures done by his hands to control photo viewer application. The user mapped the rotate image command as a capital 'R' shape gesture, and the black and white filter into 'B' shape gestures. Figure 4 (c) shows a user interacting with his hands.
- Smart book: The user used a book on the table to control PTZ camera protector application. The user maps the waving gesture as a blocking command for PTZ camera and redirecting it to a white board. Circle clockwise and anticlockwise to zoom in and out respectively. Figure 4 (d) shows a sensor attached to book to control camera protector application.

V. Evaluation

We conducted an experiment to compare between the SGS selected profile and the adjusted profile by users in means of speed and accuracy. In addition, we studied which parameters affecting the object gesture profiles and its effect on the accuracy of interaction. We asked six subjects to do the experiment. All subjects are aged from 25 to 32. We configure three different situations. The first scenario is in living room while watching photo picture viewer application. The second scenario is in the office room while sitting on the desk and using a PTZ camera protection application. The third scenario is in the outdoor location like train station and using media player.

In this experiment, each user creates a new profile using the customization interface. The user first selects one of the objects to use for his/her scenario, and then the user asked to select an application to run. SGS will copy the appropriate profile for the user. The user uses the SGS profile to interact with the application and train for all functions one time. The user evaluates the SGS profile without any adjustment for gestures twice. Then we ask the user to adjust some of the gestures that he/she did not like then evaluate the profile again three times. In the experiment, each time a message appears to the user asking him to enter specific gesture and shows a recorded video of the gesture. If the user enters different gesture pattern, then it is counted as an error. Time is measured to complete the whole session and enter all the functions of the application.

A. *Results*

Initially users take time to learn about the SGS created profile gestures specially for the first scenario. Figure 5 shows the comparison between the three scenarios in means of speed and accuracy. After subjects adjusted their gestures in scenario 1 (Photo viewer), number of errors was reduced relatively. It was observed that when the number of functions for application increased as the case of the photo viewer application, the number of errors was much higher compared to other two scenarios. In scenario 2 and 3 we observed that the selected SGS profile speed figure 5 (a) and accuracy figure 5 (b) remains constant even after user adjusted some of gesture shapes. Thus, we think that SGS can support users with appropriate gesture profiles up to nine gesture shapes. Moreover when users get trained to their gestures a learning effect could be seen for speed and accuracy in scenario 1 and 2. Subjects in scenario 3 tried using headphones, however the tapping gesture was new for users so they took time to adjust and understand how the tapping gesture works. Thats why in session 3 we got slightly increase in error rates.

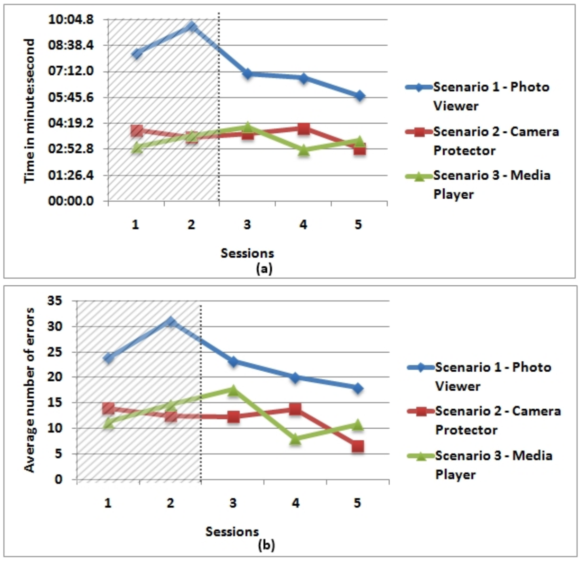

Figure 5. Average time to finish each session (a), Average number of errors per session (b), Shaded area for SGS object profile

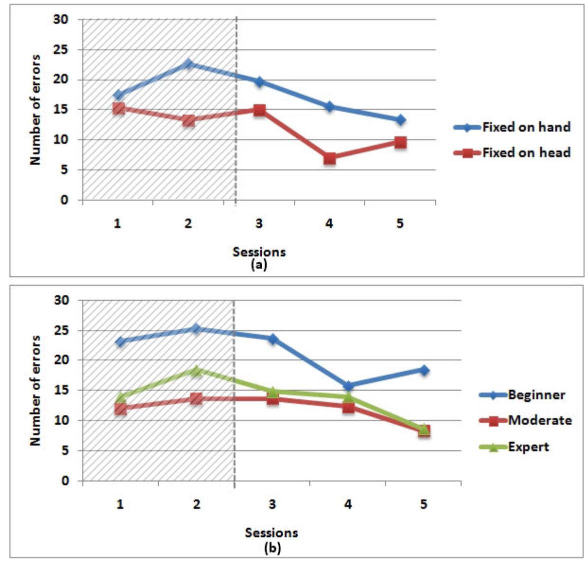

Figure 6. Total average error for object attachment parameter (a), User preference skill level (b), Shaded area for SGS object profile

We analyzed two of the parameters for selecting profiles objects attachment and users skills in more details. We found that objects that were fixed on the head like the headphone have more accurate results. We think that head has a fixed initial position for all the subjects. The movement of head has a limited spatial for doing gestures so users depend on tilting and tapping gestures to perform their gestures. The tapping gesture can be easily achieved by most of the subjects because of its fixed pattern. Figure 6 (a) shows that when users used the objects that fixed on head can get less number of errors for both SGS profile or adjusted profiles.

We classify the subjects according to their previous experience in using gestures for interactions. It can be shown in figure 6 (b) that beginner users have the highest error counts; however they show enhanced in their interaction after adjusting their gestures. Expert users too showed that they were not comfortable of the provided SGS profile, hence when they adjust some gestures their accuracy enhanced. The reduction of the errors was due to the innovative gestures that were performed by the expert subjects. However, moderate users shows a constant error level for for SGS provided profile and their adjusted profiles. Thus we expect that SGS can support normal users with appropriate profiles and can support beginner and expert users to adjust and use their own customized object gestures.

VI. Related work

Hyper-objects was discussed from designing point of view and their effect on lifestyle patterns in [5]. They discussed the potential benefit of hyper-objects on the daily life of people. Kameas et al. [6] design an architecture that aims to provide a conceptual and technological framework for engaging and assisting ordinary people in configuring everyday objects, which are able to communicate, using wireless networks. However, the tangible objects were limited to activity recognition of users. They did not provide a way to profile the activities per the users and the object locations was fixed. Kawsar et al. [7] used objects augmented with sensors to provide value added services in context aware environment. They show three types of user applications and how can objects be used to get context of the user. In our research, we try to allow users to use any object around in different ubiquitous contexts with ability to customize his gestures.

Hand gestures and object gesture are similar in their usage as both are a movement in the spatial space. However, some object gestures have some limitation for their movements as they have less spatial space such as objects attached to heads, neck and legs. Some users prefer hand tangible devices like smart glove [8] or using a handled device like the magic wand device [9] to control devices around them. We try to generalize the case of using objects and allow users to use any object to control devices in their context. Prekopcsak et al. [10] showed in their study hand gesture interface design principles like ubiquity, unobtrusiveness, adaptively and simplicity. We tried to take care of those design principles in addition to adding new parameter of situation aware gesture profiling. Ronkainen et al. [11] studied the usability of hand gestures in different ubiquitous environments. They conduct a survey on the social implications of hand gestures in public spaces. A tap gesture for interacting with mobile devices as a type of socially acceptable hand gesture was presented in their study. They point out that there are gestures that are perceived as being threatening in public spaces. Kela et al. [12] showed from their study on user gesture types for a VCR controlling tasks that users can have different gestures for different tasks which lead to the importance of personalizing the gestures. The study also concludes that gesture commands can assist users in natural interaction especially for commands of spatial association.

Context tracking and location aware services that can support the user in his/her daily activities had a lot of techniques and ideas that were discussed in [13] [14]. Situation dependent user profiles in context aware environments have been discussed in the literature from different aspects and points of view. Sutterer et al. [15] present arguments for structuring the user profile into situation-dependent sub-profiles. They extend their work in [16] to focus on the problem of profiling from the user point of view and his/her context by adding ontology reasoning to select the appropriate user profile.

VII. Conclusion and Future Work

In this research, we have presented SGS system, which is capable of understanding users object gestures in ubiquitous environment. We conducted a primary experiment to understand how users think about object gestures, how they prefer to customize their gestures and appropriate objects for interaction. In general we found that predefined gestures are not always sufficient for users interactions. Also the results guide us in designing "Smart Gesture Sticker (SGS)". SGS was presented and used to implement smart applications with daily life objects. SGS was evaluated by applying different scenarios and situations for users. Results show that SGS can assist and support users to interact with their environment and provide sufficient gesture profiles that can allow users to discover new usage for objects around them.

We think that vast applications that run with objects can get use of SGS, Smart door key, smart meeting rooms and others. We need to conduct further study about different parameters that can affect the selection of appropriate SGS object gesture profile.

References

[1] A. Atia, S. Takahashi, and J. Tanaka, "Coin size wireless sensor interface for interaction with remote displays," in *The 12th International Conference on Human-Computer Interaction (HCI International 2007*, 2007, pp. 733–742.

[2] S. Nagao, S. Takahashi, and J. Tanaka, "Mirror appliance: Recommendation of clothes coordination in daily life," in *Human Factors in Telecommunication (HFT2008)*, 2008, pp. 367–374.

[3] A. Atia, S. Takahashi, K. Misue, and J. Tanaka, "Ubigesture: Customizing and profiling hand gestures in ubiquitous environment," in *Human-Computer Interaction. Novel Interaction Methods and Techniques,13th International Conference, HCI International 2009*, 2009, pp. 141–151.

[4] R. A. Wagner and M. J. Fischer, "The string-to-string correction problem," *J. ACM*, vol. 21, no. 1, pp. 168–173, 1974.

[5] I. Mavrommati and A. Kameas, "The evolution of objects into hyper-objects: will it be mostly harmless?" *Personal Ubiquitous Comput.*, vol. 7, no. 3-4, pp. 176–181, 2003.

[6] A. Kameas, S. Bellis, I. Mavrommati, K. Delaney, M. Colley, and A. Pounds-Cornish, "An architecture that treats everyday objects as communicating tangible components," in *PERCOM '03: Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*. Washington, DC, USA: IEEE Computer Society, 2003, pp. 115–x2.

[7] F. Kawsar, K. Fujinami, and T. Nakajima, "Augmenting everyday life with sentient artefacts," in *sOc-EUSAI '05: Proceedings of the 2005 joint conference on smart objects and ambient intelligence*. New York, NY, USA: ACM, 2005, pp. 141–146.

[8] E. Farella, O. Cafini, L. Benini, and B. Ricco, "A smart wireless glove for gesture interaction," in *SIGGRAPH '08: ACM SIGGRAPH 2008 posters*. New York, NY, USA: ACM, 2008, pp. 1–1.

[9] X. Cao and R. Balakrishnan, "Visionwand: interaction techniques for large displays using a passive wand tracked in 3d," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 729–729, 2004.

[10] Z. Prekopcsak, P. Halacsy, and C. Gaspar-Papanek, "Design and development of an everyday hand gesture interface," in *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. New York, NY, USA: ACM, 2008, pp. 479–480.

[11] S. Ronkainen, J. J. Hakkila, S. Kaleva, A. Colley, and J. Linjama, "Tap input as an embedded interaction method for mobile devices," in *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*. New York, NY, USA: ACM, 2007, pp. 263–270.

[12] J. Kela, P. Korpipaa, J. Mantyjarvi, S. Kallio, G. Savino, L. Jozzo, and D. Marca, "Accelerometer-based gesture control for a design environment," *Personal Ubiquitous Comput.*, vol. 10, no. 5, pp. 285–299, 2006.

[13] A. Schmidt, M. Beigl, and H. w. Gellersen, "There is more to context than location," *Computers and Graphics*, vol. 23, pp. 893–901, 1998.

[14] T. Strang and C. Linnhoff-Popien, "A context modeling survey," in *In: Workshop on Advanced Context Modelling, Reasoning and Management, UbiComp 2004 - The Sixth International Conference on Ubiquitous Computing, Nottingham/England*, 2004.

[15] M. Sutterer, O. Droegehorn, and K. David, "Making a case for situation-dependent user profiles in context-aware environments," in *MNCNA '07: Proceedings of the 2007 Workshop on Middleware for next-generation converged networks and applications*. New York, NY, USA: ACM, 2007, pp. 1–6.

[16] S. Michael, D. Olaf, and D. Klaus, "User profile selection by means of ontology reasoning," in *AICT '08: Proceedings of the 2008 Fourth Advanced International Conference on Telecommunications*. Washington, DC, USA: IEEE Computer Society, 2008, pp. 299–304.