Cyber Physical Multimedia Systems: A Pervasive Virtual Audio Community

Markus Duchon, Corina Schindhelm, Christoph Niedermeier
Siemens AG, Corporate Technology CT T DE IT 1, Otto-Hahn-Ring 6, 80200 Munich
{*markus.duchon.ext, corina.schindhelm, christoph.niedermeier*}*@siemens.com*

***Abstract*—In recent years, pervasive systems have gained importance in the context of home automation. Social online communities are becoming more and more popular, and indoor positioning techniques have made considerable progress. Thus, creation of virtual environments integrated with the physical world that enhance the user's perception and cognition, becomes feasible. In this paper, we propose a system combining these concepts into something we call a cyber physical multimedia system. This technology leverages pervasive audio communities that facilitate social activities for people with limited mobility, such as the elderly or handicapped.**

***Keywords*-Social Networks, Multimedia, Pervasive Computing, Indoor Positioning, Cyber Physical Systems;**

I. INTRODUCTION

Nowadays being virtually connected with friends, business partners, or items of interests has already become an inherent part of our lives. Via Twitter, Facebook, Ebay and Amazon, we follow news feeds of our favorite musicians or political happenings, we keep in contact and chat with friends and we go shopping while sitting on the couch, respectively. The web has opened a new world of interconnection and communication; however, only the young and computer savvy generation is able to grasp its benefits. People with limited mobility, such as the elderly or handicapped, would benefit greatly from social communities due to their naturally tends to suffer from isolation, yet they have only few possibilities because of their lack of computer knowledge and/or lack of special in-/output assistance for those features. For Germany, the demographic forecast shows that by the year 2035 more than two-thirds will be over 60 years old. Many projects (e.g. Smart Senior) are based on this fact and its resultant problems, such as rising health care costs. The main goal is to create comprehensive cyber physical systems that allow elderly people to stay in their home environment as long as possible, for example, by enabling remote healthcare or community platforms for friends and family to prevent isolation. Cyber physical systems (CPS) [1] are physical and engineered systems whose operations are monitored, coordinated, controlled and integrated by a computing and communication core. We will extend this by the term multimedia reinforcing the utilization of audio communication to a Cyber Physical Multimedia System (CPMS).

In this paper, we will focus on the design of a CPMS, which enables handicapped and elderly people to communicate in an easy way. The remainder is organized as follows. In Section II, we will introduce a scenario and detail basic requirements we consider indispensable for a suitable CPMS enabling a pervasive virtual audio community. Our system architecture is presented in Section III. Section IV discusses related work before Section V concludes the paper.

II. REQUIREMENTS

Before discussing requirements of a CPMS, we will outline an application scenario. Imagine three different apartments like illustrated in Figure 1 of Peter, Mary and Paul. They agreed on a coffee party within their own kitchens. Peter and Mary are already in their kitchens and have a

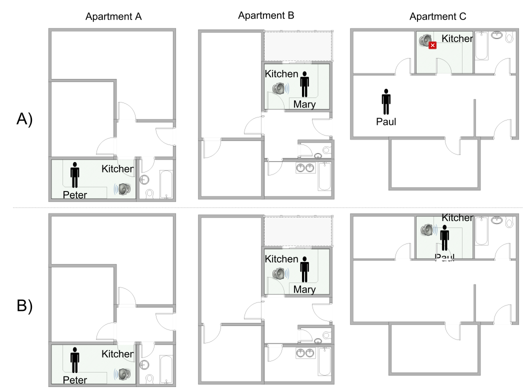

Figure 1. Three different apartments forming a virtual audio community

conversation about the latest news, whereas Paul resides in another room at that time (A). As soon as he enters his own kitchen, it is recognized by the system and an audio channel is opened to a virtual kitchen where he participates in the conversation of Peter and Mary (B). Seeking for a CPMS to enable a distributed, virtual audio community, we must consider some basic requirements a proper solution has to fulfill.

A. System Configuration

The system should provide different levels of configuration mechanisms. This is necessary because such a complex system as the proposed virtual audio community, in some cases, requires fine-tuning involving a considerable amount of user interaction whereas during normal operation it should be possible to configure basic behavior in a simple and intuitive way. The system must provide a detailed configuration interface that could be accessed via a Web browser. This kind of interface may be used for fine-tuning, in particular during initial configuration of the system for a user that has just joined the community or wants to perform major changes in his configuration. For instance, the web-based configuration interface may be used for specification of a mapping function for a physical actor to be integrated into the virtual environment. The system must provide a basic configuration interface that can be controlled via voice commands. This kind of interface may be used to change settings during normal operation of the system. For instance, a user could temporarily map a corner of his living room to the virtual kitchen by issuing a respective voice command.

B. Physical Actors

The system may integrate with the physical environment of the real rooms which are mapped to virtual rooms of the pervasive audio community wherefore some requirements are essential. Integration of physical actors into the virtual environment must be restricted to certain contexts. The virtual actor should only have an effect on real actors belonging to a community user if the respective user has authorized and activated a corresponding context. For instance, lighting control should only occur in a particular room if the user has actually entered the corresponding virtual room. Integration of physical actors requires a mapping between states of the physical actors and the virtual actor associated with them. For a simple actor that just has an ON and an OFF state, this is trivial. But, if a whole spectrum of states (e.g. light intensity or volume of background music) is to be addressed, a mapping function is required. When a user leaves a virtual room, the actors in the real room must be restored to their original states (i.e. the states that they had before the user entered the virtual room) unless the user explicitly decides otherwise.

C. Audio System

Since we want to convey the impression that physically separated people are together in an audio community the audio system should support surround sound capabilities for each room. Since the system should not be restricted to a single room but rather provide coverage for several rooms or even a complete apartment, several input and output devices are necessary. These resources need to be connected and individually responsive to an audio management component. A dynamic change of audio components in terms of volume adjustment and device handover is required to provide an accurate impression when moving around. Also, the exchange of audio communication data among remote participants has to be achieved especially for conference-like community communication.

D. Maps and Indoor Positioning

The system must be able to map physical rooms onto virtual rooms and also be able to find the current position of the user. Therefore, two aspects have to be considered:
Maps of the home environment must be available in digital form and context must be added to these maps. The maps serve as a basis for the positioning system, and the additional context is necessary to identify a room e.g., as living room etc. making a mapping to virtual rooms possible.
A positioning system must be installed. The higher the precision of this system, the better the features developed for it will be. Room level accuracy is the minimum accuracy the system requires. WLAN positioning systems offer a precision around 1-2 meters, which would be sufficient to determine with high probability in which room the user resides. If high precision positioning is available, for example with ultra-wideband (UWB), special effects are possible. As an example, we would like to describe sound fading. Imagine a user group in the virtual room ”kitchen” standing around the stove. One user physically moves away from this stove but doesnt leave the room. The voices of the remaining users would become softer out of the speakers, which creates a more realistic experience.

III. Architecture

Driven by today’s ambient assisted living approaches our system utilizes several mature technologies and combines them to a CPMS enabling a `Pervasive Virtual Audio Community` for elderly, handicapped or disabled people to enhance their participation in social life including a realistic experience. An overview of our system architecture

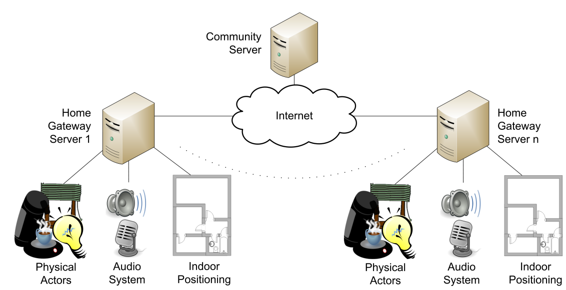

Figure 2. System architecture for a pervasive virtual audio community

is illustrated in Figure 2. Because our goal is to facilitate a community among physically separated users, a platform represented by a `Community Server` (CS) that provides virtual rooms and offers community capabilities is used. Connected through broadband internet, the platform can be accessed by several `Home Gateway Servers` (HGS), which represent the basis on the user side and are able to control and sense activities within flats. To ensure the user’s privacy the proposed system has to be activated and deactivated explicitly by the user. Several physical actors are connected to the gateway and can be controlled regarding the community context - imagine light, window shutters or coffee makers. The required hardware, speakers and

 ISBN: 978-1-61208-129-8

microphones, for audio system is deployed and also linked to the control entity. Since the system heavily depends on the current location of the users moving around in their homes, a fine-grained indoor positioning system is utilized.

A. Community Server

The CS provides virtual rooms where participants can exchange audio signals in the first step. The main idea is sketched in Figure 3, whereby the kitchens of three different flats are unified in a virtual kitchen. These virtual rooms represent certain areas of a habitation where multiple users can participate similar to a telephone conference. Thereby the CS needs to provide certain group functionalities in order to provide multiple virtual kitchens for several independent circles of friends as well as user registration capabilities.

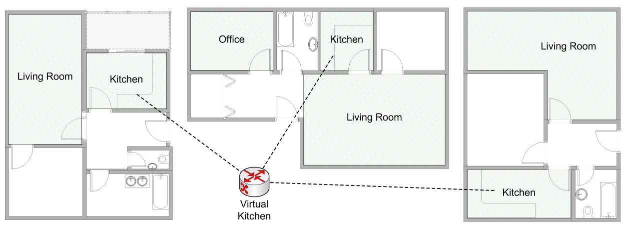

Figure 3. The CS provides a virtual room (kitchen) to which the real kitchens in several flats are interconnected through their HGSs.

B. Home Gateway Server

Within our system the HGS has to fulfill several tasks, which will be explained in turn.

Interaction and Communication: Considering our target group the handling during operation should be kept as simple as possible wherefore our system utilizes voice commands. These commands can be easily defined and extended following a simple syntax: `System: <command>`. For example, the user can use the command `System: start/stop` to start or stop the community system. This command structure can also be used for simple system configuration tasks. For example, the command `System: map to living room` can be used to set the current room the user is in, to the living room. After the recognition of a command the gateway maps the corresponding orders into machine readable commands by applying a ruled based approach. Audio communication within the community is crucial, since we want to achieve interaction with other users. We decided to utilize audio communication in the first step wherefore a VoIP system is sufficient for the desired functionality. The user's command `System: call kitchen` initiates a connection to the CS to join the virtual kitchen of that group the user is associated with. Once the registration process at the CS is completed an audio channel is set up to the specified virtual room. In addition to the transmission of audio signals, notification messages among remote participants need to be exchanged which include information that will be utilized by the connected HGS to adapt the local environment. For these messages we utilize the SUBSCRIBE request dialog [2], [3], which can be used for status notifications to remote users and therefore can be interpreted by the remote HGSs.

Physical Actors, Audio- and Positioning System: The HGS is also responsible for managing the user's environment by the interpretation of notification messages. According to the content of these messages a command is generated and passed to the corresponding home appliance. These messages are created according to the status of the real home appliances which are connected to their virtual counterpart. This will enhance the user's perception and cognition when participating in a pervasive virtual audio community. For instance, community users have attached their room lighting to a virtual lighting control system which than can be used to dim the lighting in the virtual room. As a consequence, the lighting of the real rooms mapped to the virtual room is dimmed accordingly. Another example is a coffee party that has agreed to meet in a virtual kitchen at a certain time. A community user starts all coffee machines mapped to a virtual coffee machine by starting his own machine so that coffee is ready when the other users show up in the virtual kitchen. Besides the adaptation of the physical environment we want our users to be informed about the communication status. Therefore, controllable status lights (i.e. small LEDs) are deployed in appropriate rooms. These lights indicate the remote activity in a specific virtual room by changing color and intensity.

The audio system mainly consists of several microphones and speakers deployed in the household, which represent different sources and sinks in terms of acoustic signals and can be accessed and controlled individually by the HGS. It should be emphasized that an actual voice transmission only takes place when the user is nearby. A user who leaves the kitchen physically, also leaves the virtual kitchen and therefore does not participate in a conversation anymore which takes place there. However, when a user enters the (virtual) living room, where another group resides, the HGS needs to switch to the corresponding room at the CS as well as change the microphone source and the speaker sinks to the user's new location. Furthermore, by using a more precise positioning system it is possible to adjust the surround volume at the remote sides considering the user's position in the real room in terms of a cardinal direction. As highlighted, the audio system needs to be highly coupled to a positioning system. Whenever the user changes his location, the HGS takes action: a) to inform remote gateways about the activity change (status lights, position update) and b) switching the communication in-/output in accordance with the user's physical location. The coupling is performed by the HGS, which holds and manages an indoor map, an overlay map of the defined virtual rooms, and an overlay map for the audio components and their corresponding volumes. The better the precision of the positioning system, the more fine-grained the audio control performance is, the more realistic a virtual conversation will be.

IV. Related Work

In the following, we will outline the state of the art of technologies and subcomponents that are part of virtual audio communities.

A. Physical interaction

Regarding physical control, so called home automation servers are used for processes like adjusting the heating according to the current temperature or adjusting the lighting system as well as the background music either by voice commands or proactively by context recognition. Lots of attention has been put on voice controlled systems [4]–[6] to facilitate their utilization. Also several communication standards and network protocols are widely used. An overview of relevant standards is given in [7].

B. Audio Communities

Prominent representatives like Skype or, GoogleTalk are also capable of conference calls, but are inconvenient to operate, require direct computer interaction and lack of spatial perception of human interaction. Healy et al. [8] present a prototype of audio spatial augmentation headphones in order to take advantage of the innate psycho acoustical perception of sound source locations. Hyder et al. [9] outline the difficulties to identify the speaker of a teleconference and present a solution which adds a virtual acoustic room simulation. Kim et a. [10] propose a 3-dimensional VoIP system for two user groups, whereby the participants can hear the voice of remote users as if each remote user speaks at his or her corresponding position.

C. Positioning Systems

One key information is the position of the user within his home environment. When there is only one person living in a household, the identification of the person is not actually necessary, and as a result, positioning could be accomplished via smart home equipment, such as motion sensors in each room or contact sensors on doors. However, when more than one person lives within a household, the identification of the person using the system is essential. A vast number of positioning technologies and systems exist which support different levels of accuracy. An overview is provided in [11]. If radio-based technologies are used, the user needs to wear or carry a tag. For room level accuracy, WLAN positioning is sufficient and can be enhanced by adding motion sensors. If a higher level of accuracy is needed, UWB systems enable positioning in the range of centimeters.

V. Conclusion

As for now, most of the systems dealing with pervasive computing and home automation put the focus on an easy interaction with the user, by utilizing gesture control, voice control, or on fully automated approaches which analyze the user's context in order to support him in his daily life. As a result, only one facility is addressed which is inhabited by one or more users. Our approach enhances pervasive systems by adding the community aspect. This is achieved by adopting a mixed-reality approach that associates several real and a single virtual environment. At the current stage of implementation we identified future research directions and possible enhancements. The system could be extended by integrating augmented reality aspects both acoustically and optically. This could further enhance the user experience and intensify the illusion of actually spending time together rather than communicating via the internet. This could mean a significant improvement to the lives of people not able to leave their homes and thus being subject to social isolation.

References

[1] E. A. Lee, "Cyber Physical Systems: Design Challenges," in *ISORC*, 2008.

[2] G. Camarillo, A. Roach, and O. Levin, "Subscriptions to Request-Contained Resource Lists in the Session Initiation Protocol (SIP)," RFC 5367 (Proposed Standard), IETF, October 2008.

[3] A. B. Roach, "Session Initiation Protocol (SIP)-Specific Event Notification," RFC 3265, IETF, June 2002.

[4] I. Mporas, T. Ganchev, T. Kostoulas, K. Kermanidis, and N. Fakotakis, "Automatic Speech Recognition System for Home Appliances Control," in *PCI*, 2009, pp. 114–117.

[5] J. Zhu, X. Gao, Y. Yang, H. Li, Z. Ai, and X. Cui, "Developing a voice control system for ZigBee-based home automation networks," in *IC-NIDC*, 2010, pp. 737–741.

[6] A. Gárate, N. Herrasti, and A. López, "GENIO: an ambient intelligence application in home automation and entertainment environment," ser. EUSAI, 2005, pp. 241–245.

[7] W. Kastner, G. Neugschwandtner, S. Soucek, and H. M. Newmann, "Communication systems for building automation and control," vol. 93, no. 6, pp. 1178–1203, 2005.

[8] G. Healy and A. F. Smeaton, "Spatially augmented audio delivery: Applications of spatial sound awareness in sensor-equipped indoor environments," in *MDM*, 2009, pp. 704–708.

[9] M. Hyder, M. Haun, and C. Hoene, "Placing the participants of a spatial audio conference call," in *CCNC*, 2010, pp. 1–7.

[10] C. Kim, S. C. Ahn, I.-J. Kim, and H.-G. Kim, "3d voice communication system for two user groups," in *ICACT*, 2005, pp. 100–105.

[11] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *Communications Surveys & Tutorials, IEEE*, pp. 13–32, 2009.