Amigo Context Management Service with Applications in Ambient Communication Scenarios

Joerg Schmalenstroeer[1], Volker Leutnant[2], Reinhold Haeb-Umbach[3]

University of Paderborn, Department of Communications Engineering, Germany
[1]schmalen@nt.uni-paderborn.de,[2]leutnant@nt.uni-paderborn.de,
[3]haeb@nt.uni-paderborn.de

Abstract. The Amigo Context Management Service (CMS) provides an open infrastructure for the exchange of contextual information between context sources and context clients. Whereas context sources supply context information, retrieved from sensors or services within the networked home environment, context clients utilize those information to become context-aware.
An ambient communication scenario realizing follow-me elements is used to showcase how applications benefit from the combination of acoustics-based context sources and positioning sensors.

1 Introduction

Ambient intelligence (Ami) describes the vision of disappearing computer or other electric hardware while their functionality is still readily available to support users in their daily work [1]. One of the important building blocks of such a system is a communication infrastructure, a network, which connects heterogeneous devices. Currently, a major obstacle to realizing the vision is the lack of interoperability between devices of different manufacturers. The Amigo project has been set forth to overcome this by developing an open, standardized and interoperable middleware [2]. Further, demonstrators are being set up to showcase the benefits of a networked home to end-users.

Services can now be developed on top of the middleware. They can be perceived by users as "intelligent" if they adapt their behaviour according to the current context, where the term context summarizes all environmental or usage information which may be of relevance for the service's appropriate behaviour. The collection and provision of context information is thus of paramount importance for the overall system. This paper describes the Amigo project's structure to this, the Context Management Service (CMS). To demonstrate some of the potential of CMS, so called "ambient communication" has been realized, a hands-free communication system, where the call follows the user as he moves from one room to another, utilizing i/o devices (microphones, loudspeakers, screens) most appropriate with respect to the current user's location.

2 Scenario

Maria recently moved out of her father's home for a job in another city. However, they want to keep their close relationship, despite the physical distance, by using an ambient communication system. When Maria comes home the system recognizes her and, according to her preferences, an audio connection to her father's system is established. As she walks through the home the audio connection follows her automatically. John hears Maria doing her homework and can talk to her when ever he wants. If cameras and displays are available, a video communication may also be established.

The aforementioned scenario places a number of requirements on the home network, and in the paper we will present the solutions proposed by the Amigo system. In the next section we will introduce the Amigo Context Management Service (CMS) and its concepts for distributing context information. Section 4 is about acoustic scene analysis as a means of gathering context information for services. In section 5 we discuss the realization of ambient communication and how it makes use of the Amigo CMS, before we finish with some conclusions drawn in section 6.

3 Amigo Context Management Service

The major task of the Amigo Context Management Service is the collection and distribution of context information for services and applications [3]. Fig. 1 gives a brief overview of the CMS architecture and its components.

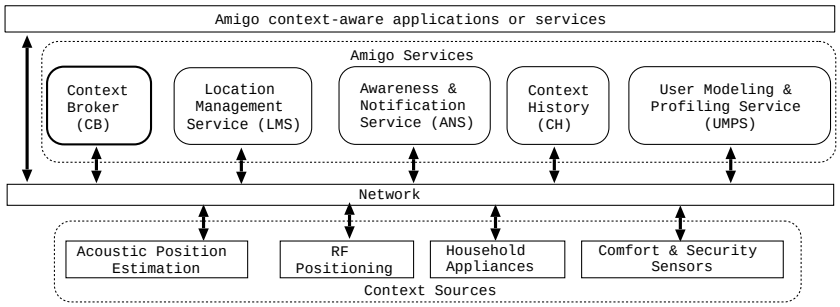

Fig. 1. Amigo CMS architecture

3.1 Amigo CMS services

The Context Broker (CB) is the central component of the Amigo CMS services and responsible for the look up of context sources. Additionally the CMS offers the following services:

– *Awareness and Notification (ANS)*: Service for notifying applications or services about specific events.

– *Context History (CH)*: Service for recording context information and thus enabling services to exploit past context conditions.
– *User Modeling and Profiling (UMPS)*: Service for storing user preferences, based on stereotypes and configurations.
– *Location Management (LMS)*: Service providing information about user and device positions.

Each service is specialized on a certain task and provides basic data processing on context information for applications. The LMS for example, assembles a database about position information of users and devices from context sources offering location information.

3.2 Context sources

A context source can be a sensor or a service which provides information about its current status, measurement data, database entries or usage conditions. Examples of context sources and the information they provide are:

– Sensors: temperature, air humidity, brightness, water leakage, fire detector
– Household appliances: goods available in the refrigerator
– Consumer electronics and mobile devices: usage status, login data
– Applications, Services, Databases: status
– Audio, Wlan, Bluetooth, Radio Frequency: user or device location

Each context source must implement the IContextSource interface, encapsulating physical sensors and devices, to full-fill the interoperability requirements.

There are two basic modes of data exchange between a context source and a context client (application): asynchronous and synchronous. In the asynchronous mode the application registers to context events and waits for notifications, i.e. a change of the context. Whereas in the synchronous mode the application immediately requests context information. All context information is delivered in a RDF/XML [4] description format.

The Amigo CMS reflects the dynamic nature of a networked home, where new context sources and services may be introduced in the home and others being removed. So context sources register at a central service (Context Broker) to announce their capabilities and to store a webservice reference to them. Timeout regulations guarantee that inactive context sources are removed from the list of available context sources.

3.3 Context-aware applications

Applications searching for context information query the CB with a RDF description of their needs for context sources and thereupon retrieve a list of references to matching sources. SPARQL [5] queries are used to request context information directly or to subscribe to events matching the SPARQL query. Applications deploying the asynchronous data exchange have to implement a notify method according to the IContextSource standard, which is called by the subscribed context source for notification.

4 Acoustic Scene Analysis

Acoustic scene analysis refers to the retrieval of all kinds of information inherent in audio signals captured by either single microphones or microphone arrays. This can be information about persons (e.g. age, gender, position, identity, mood), the environment (e.g. distortions, music) or conversations (e.g. participants, topics). In this paper we will focus on the position estimation and speaker identification tasks, since they are valuable context information in ambient communication scenarios.

4.1 Acoustic Position Estimation

Estimating the position of a speaker requires multi-channel audio signals recorded by spatially distributed microphones, either by using beamforming techniques [6] or generalized cross correlation (GCC) approaches [7].

Compared to GCC methods, beamforming has the advantage of providing an enhanced audio signal with improved Signal-to-Noise Ratio (SNR), in addition to position information. We use linearly arranged microphone arrays and Filter-Sum-Beamformer (FSB), which perform a Principal Component Analysis (PCA) and thus blindly adapt towards the loudest speaker. Each FSB delivers a Direction-of-Arrival (DoA) estimate of the speech source. As the microphone positions are known, a system of equations can be set up, whose solution delivers the speaker's location [6].

An interesting application for the acoustic positioning technique is automatic camera control in communication scenarios, where the estimate of the speaker location in combination with a face detection algorithm is utilized to steer a camera towards the speaker. Using acoustic cues in addition to visual cues allows directing the camera towards the speaker, even if the speaker is currently not in the field of view.

4.2 Speaker diarization

A user's location is a valueable context information for the networked home environment. Among the various location estimation techniques acoustic position estimation has the advantage of being usually more precise than RF-based techniques. Further it does not require the speaker to carry a special device or wear a tag. However it demands the identification of the speaker to complement the context information. This task is referred to as speaker diarization and addresses the following question: "Who is speaking, when and where?".

Fig. 2 gives an overview of the used information sources and the chosen architecture. A Viterbi decoder is utilized to simultaneously segment and annotate the audio data, instead of performing the steps sequentially. This allows short delays and online processing of audio streams [8].

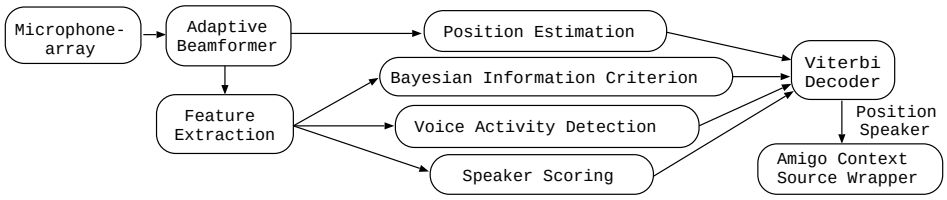

Fig. 2. Speaker diarization system architecture

5 Ambient Communication

Ambient Communication (AmCom) is the concept of virtually connecting spatially separated locations. Humans living in such environments should experience the real distance between the communication partners to a lesser extent, by developing a feeling of participating in the life of the distant partner. The communication itself moves to the background of the user's conscious attention as each partner follows his daily live.

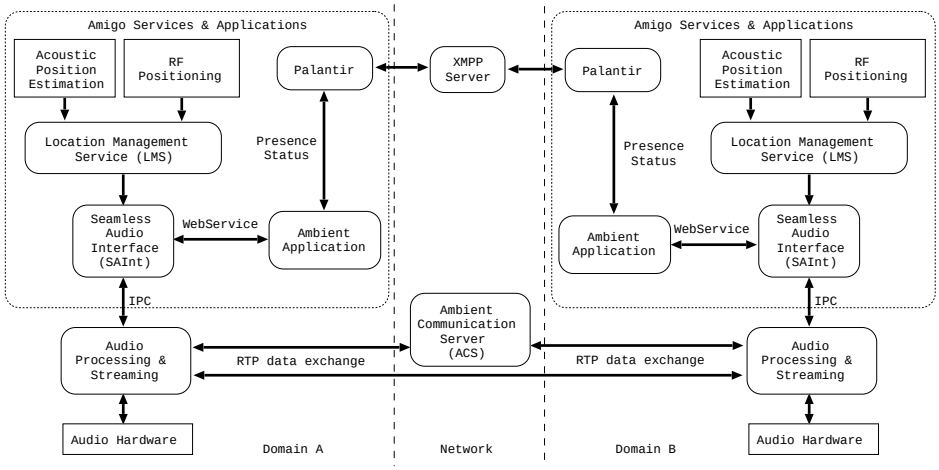

Fig. 3. System architecture for ambient communication

Fig. 3 sketches the building blocks of the AmCom system developed within Amigo. It is divided in blocks responsible for audio processing and in blocks realizing the control layer, both connected via an interprocess communication (IPC). Sophisticated signal processing algorithms have been employed to realize high-quality hands-free communication, such as wideband speech coding, multichannel echo cancellation, acoustic beamforming, noise reduction and real-time streaming. The control layer uses the Amigo middleware and services for gathering context information (e.g. location, awareness status of users) and offering services like the Seamless Audio Interface (SAInt) via Webservices to applications. The Palantir service exchanges context information about the users "availability and willingness to communicate" via an XMPP server and thus enables applications to control the ambient communication according to the user's current status.

One of the AmCom requirements is to be independent of terminals and to allow the user to move freely within his home. The communication follows the user ("Follow-Me") seamlessly throughout the home, configures itself according to the available hardware (audio and/or video) and does so by adhearing to privacy rules.

6 Discussion

Ambient Communication is an attractive service for the networked home environment, but it demands a certain amount of context sources and a service infrastructure providing high-level context information about the user.

The presented Amigo Context Management Service copes with these requirements by offering standard interfaces, discovery mechanisms for context sources and procedures for exchanging contextual information. Additional services like the Location Management Service (LMS) increase the range of contextual information as they encapsulate several context sources to gain improved location information. We introduced the acoustic scene analysis as a context source for the LMS and briefly discussed the technique. In contrast to common RF systems it offers position estimates without forcing the user to carry any devices. Applications, such as ambient communication, can utilize a combination of these context sources for location or proximity based services.

7 Acknowledgment

This work has been supported by the European union project "Amigo - Ambient intelligence for the networked home environment" [IST 4182]. For more information visit "www.amigo-project.org"

References

1. Marzano, S., Aarts E.: The New Everyday View on Ambient Intelligence, 2004
2. Amigo Project website: http://www.amigo-project.org, 2007
3. Ramparany, F., Poortinga, R., Stikic, M., Schmalenstroeer, J., Prante, T.: An open Context Information Management Infrastructure - the IST-Amigo Project, Conference on Intelligent Environments, Ulm, Germany, 2007
4. Resource Description Format (RDF) Specifications: http://www.w3.org/RDF/
5. SPARQL Protocol and RDF Query Language: http://www.w3.org/TR/rdf-sparql-query/
6. Warsitz, E., Haeb-Umbach, R.: Acoustic Filter-and-Sum Beamforming by Adaptive Principal Component Analysis, ICASSP05, Philadelphia, USA, 2005
7. Knapp, C. and Carter, G.: The generalized correlation method for estimation of time delay, IEEE Trans. ASSP, vol. ASSP-24, pp. 320-327, Aug. 1976
8. Schmalenstroeer, J., Haeb-Umbach, R.: Joint Speaker Segmentation, Localization and Identification for Streaming Audio, Interspeech 2007, Antwerp, Belgium, 2007