ACOUSMA: UBIQUITOUS & INTELLIGENT AUDITORY DISPLAYS

Andreas Michelakis¹, Emmanouil Zidianakis¹, Asterios Leonidis¹, Stavroula Ntoa¹, Margherita Antona¹ and Constantine Stephanidis¹,²

¹ Institute of Computer Science, Foundation for Research and Technology - Hellas, Heraklion, Greece
² University of Crete, Heraklion, Greece

ABSTRACT
ACOUSMA is a platform based on micro-services architecture developed to allow effortless enhancement of Ambient Intelligence (AmI) environments with personalized auditory displays. ACOUSMA consists of two system packages designed for interaction with three groups of users: auditory display experts, AmI application developers and AmI environment inhabitants. Auditory display experts can generate, upload and share their designed auditory representations using an intuitive Content Management System (CMS). AmI application developers with no prior knowledge on the field can integrate personalized auditory displays in any AmI environment with minimum effort using a simple application programming interface. An intelligent module was developed to adapt and grow to user preferences in real time, enabling conversion of semantic information into meaningful, user-tailored auditory displays. Finally, AmI environment inhabitants experience auditory displays that occur in the right place at the right time and are adjustable to their preferences. Usability evaluation from both experts and final users has shown that the target user groups can benefit from using ACOUSMA as a platform to store and share auditory representations, as well as to provide personalized, meaningful and context-sensitive auditory displays. Moreover, users could successfully and efficiently communicate their needs and adjust auditory displays using the intelligent module.

KEYWORDS
Auditory Display, Ambient Intelligence Environment, Recommendation System

1. INTRODUCTION

Interaction multimodality is often claimed to be a fundamental feature of Ambient Intelligence environments and related applications (Aarts & Wichert, Ambient intelligence, 2009). However, in practice, for the majority of applications, the focus is on the visual channel to display information. Due to such focus on visual displays results the auditory channel is often not efficiently utilized by AmI applications, as developers have limited specific domain knowledge and lack tools to provide meaningful, personalized auditory displays. To address this issue, this paper introduces ACOUSMA, a complete framework and suite of tools designed to aid AmI developers and Auditory Display experts in enhancing AmI environment applications with information-rich auditory displays that are personalized to inhabitant’s preferences and adapt according to environment and inhabitant context. The system provides: (i) A standardized way of storing and sharing created auditory representations, (ii) A standardized way to deploy and control auditory display playback, (iii) Flexibility in use and monitoring of hardware (output devices), (iv) Auditory displays that are in harmony with user and environment context (occurring at the right place, at the right time), (v) Auditory displays that can be adjusted in real time to user preferences, (vi) Dynamic evaluation of environment context before and during playback of an auditory display, (vii) Meaningful and personalized auditory displays for any given AmI environment data or event.
2. BACKGROUND & RELATED WORK

2.1 Background

ACOUSMA is built upon the AmI paradigm and the background theory of auditory representations, auditory displays and sonification. The AmI paradigm’s realization is a result of microprocessors decreasing in size, the high computing power to user ratio and the successful user experience with mobile devices and technological artifacts (Adelstein, Gupta, & Richard, 2015). An AmI environment is characterized by five primary features (Aarts & Encarnação, 2006) (Schmidt, Kranz, & Holleis, 2005): Embeddable - It is built upon technological advancements of computer networks, sensors, actuators and distributed computing. Context-aware - Utilizes contextual information regarding users, creating profiles, recognizing user activities, situation and location. Personalizable - It is customizable according to user preferences. Adaptive - It adapts dynamically to change in user preferences and context. Anticipatory - Uses artificial intelligence to predict user intentions. ACOUSMA as a platform for AmI environments inherits those characteristics. Auditory representation is defined as a technique or design approach that aims to the creation of a sound that can be used to communicate and represent information or a meaning (Csapó & Wersényi, 2013) (Serafin, et al., 2011). In literature, auditory representations are summarized into three categories, basic, speech based and sonification based (Csapó & Wersényi, 2013). Basic auditory representations include auditory icons (Gaver, 1986) and Earcons (Blattner, Sumikawa, & Greenberg, 1989). Speech based representations include Spearcons (Walker, Nunce, & Lindsay, 2006). Spindices (Jeon & Walker, 2009), Auditory Emoticons (Froehlich & Hammer, 2004) and Sp emoticons (Nemeth, Olasz, & Csapo, 2011). Sonification includes Auditory Warnings (Patterson, 1990), Morphocons (Parseihiyan & Katz, 2011) and Musicons (McGee-Lennon, Wolters, McLachlan, Brewster, & Hall, 2011). ACOUSMA integrates in its interfaces the majority of auditory representation types. Auditory Displays, as stated in (Walker & Kramer, Ecolological Psychoacoustics and Auditory Displays, 2004) is the field that examines the use of sound to depict information in computing environments. Four approaches for creating an auditory display are established in literature (Walker & Nees, Theory of sonification, 2011) (de Campo, 2007). Those are, Audification (Kramer, 1994), Event-Based – Parameter mapping (Grond & Berger, 2011), Continuous and Model Based (de Campo, 2007). Another, rather new approach is that of Blended Sonification (Tünnemann, Hammerschmidt, & Herrmann, 2013) and has been used for ambient, calm auditory displays. ACOUSMA is designed to support playback of auditory displays generated with any approach.

2.2 Related Work

Only a limited amount of research work exists in the field of Auditory Display in the context of AmI environments. To the best of our knowledge, no platform or framework for the management and distribution of personalized, context sensitive auditory displays in AmI environments is currently available. The majority of Auditory Displays do not consider environment context and user preferences before deployment or during playback of auditory displays. Examined here are data sonification tools, programming interfaces and existing AmI auditory display applications.

Data sonification tools are used for generating auditory analogies to datasets and are mainly used for data analysis and exploration. Interaction with those systems is performed via a Graphical User Interface (GUI) and there is no requirement for programming skills. For inputting data, a tabular document is needed (text, csv), as the tools offer no database support. Sonification Sandbox (Davison & Walker, 2007) and xSonify (Candey, Schertenleib, & Merced, 2006) use the parameter-mapping approach for sonification and MIDI for sound output. Another promising sonification tool, Sonifyer (Dombois, Brodwoj, Friedli, Rennert, & Koenig, 2008), uses the Audification approach and a limited version of parameter-mapping (frequency modulation synthesis). A recent work, the Sonification Workstation (Phillips & Cabrera, 2019) is an open-source, multi-platform application designed for general sonification tasks. Functions covered by the tool are data exploration and art, while approaches available are mainly parameter mapping and Audification. The audible output of the tool can only be heard with the tool running and there are no options for audio file export.
Programming frameworks have been developed to allow sonification of data. Differently from sonification tools, such frameworks require an adequate knowledge of sonification techniques and programming. Interaction with these frameworks happens in a text-based environment, a source code editor. SuperCollider (Wilson, Cottle, & Collins, 2011) is an open-source platform for audio synthesis and algorithmic composition, used by musicians, artists, and researchers working with sound. SoniPy (Worrall, Bylstra, Barrass, & Dean, 2007) is a publicly available framework based on the Python programming language. It uses components of Python for data acquisition, storage and analysis and adds modules for perceptual mappings and sound synthesis. Csound (Lazzarini, et al., 2016) is also a publicly available sound and music computing system.

Auditory display applications designed for intelligent environments draw real-time data produced by environment’s sensors to generate meaningful auditory feedback. Systems use a variety of hardware and software tools (some already discussed here) in order to provide meaningful auditory displays. UPSTAIRS (Tünnermann, Leichsenring, Bovermann, & Hermann, 2015) and Knock ‘Knock (Tünnermann, Hammerschmidt, & Hermann, 2013) are prototype systems that use the Blended Sonification approach with contact microphones placed in the environment to capture data and the SuperCollider engine to transform those data into meaningful auditory displays. In Powerchord (Lockton, Bowden, Brass, & Gheerawo, 2014) methods were examined to raise awareness for energy consumption. In the augmentation of a kitchen described in (Groß-Vogt, et al., 2018), the application used two loud speakers for playback and a laptop to receive data emerging from environment sensors.

Currently, there is no standardized hardware or software framework to deploy and control auditory displays in AmI environments. Context-awareness and personalization approaches are also missing. There is no flexibility in the usage of output devices (meaning that applications are bound to the hardware used). In AmI environments where an abundance of information arises from sensors and services, it is crucial that standardized methods exist for harnessing information as well as software and hardware infrastructure to support deployment of applications that create memorable user experiences. The rationale behind the design & development of ACOUSMA is precisely the gap found in literature and current applications. In detail, there is no system to allow management of auditory displays in AmI environments as well as no standardized way for providing auditory displays by AmI applications and non-expert users (application developers). Aligning with the AmI paradigm’s major characteristics, ACOUSMA provides context-aware, personalizable and adaptive auditory displays. Moreover, ACOUSMA uses available network resources (devices and speakers) to direct and control auditory displays anywhere in the intelligent environment.

3. DESIGN PROCESS & USER REQUIREMENTS

A user center design process was followed to elicit functional and non-functional requirements for ACOUSMA. The platform is targeted to three target user groups: Auditory Display experts, AmI application developers and AmI environment inhabitants. Auditory Display experts can use sophisticated, domain specific tools to design, develop and deploy auditory displays. AmI application developers are involved in the creation of applications for multiple domains. They develop multimodal interfaces that use the environment’s infrastructure (sensors, actuators & services) to provide information-rich user experiences. Finally, AmI environment inhabitants are people that live in and experience AmI environments, ranging from residents of a smart home to visitors of a museum. These users have extremely diverse needs and preferences, requiring multimodal interfaces capable to adapt in a natural way to their profile.

In the process of eliciting user requirements for ACOUSMA, Brainstorming, Group Discussion, Direct Observation & Scenarios were used. Experienced researchers in the field of HCI and AmI as well as UX/UI (user experience / user interface) experts and AmI application developers participated in two brainstorming sessions and group discussions. Multiple Direct Observation sessions with user representatives were also executed. As a main outcome, using ACOUSMA’s interfaces, Auditory Display Experts should be able to store & share auditory representations, and quickly and easily test stored designs in an AmI environment; AmI application developers should be able to integrate and deploy personalized auditory displays, and allow the system to evaluate the right context to initiate and auditory display; AmI environment inhabitants should be able to experience personalized and adjustable auditory displays.
4. IMPLEMENTATION

The system consists of two distinct packages, the Speaker Director and the Auditory Display Director. Both packages combined aim to enhance Ambient Intelligence environments with personalized auditory displays. Auditory displays are deployed based on user, environment and application context, as well as user specific preferences regarding auditory representations and sound categories.

4.1 Speaker Director Package

In order to present an auditory display, the most basic requirement is an output device, a speaker. Therefore, output devices in an AmI environment are required to be ubiquitous, enabling the distribution of auditory displays anywhere and in every condition. The Speakers Director package was developed as a solution to support this requirement consisting of the following modules.

**An AmI Audio Client** can be a physical device (e.g., smartphone, TV) with one or more audio output devices, running software enabling playback control of these devices over network. A cross-platform software was developed enabling easy integration and over network playback control for any device usually found in AmI environments. Consequently, any device with audio output capabilities can be enhanced with the software and transformed into an AmI Audio client. In essence, through a network of AmI Audio Clients the seamless distribution of auditory displays in AmI environments is achieved. Every AmI Audio Client establishes a connection to the Speaker Server and sends real-time updates regarding output devices playback status, as well as the client’s location (e.g. “kitchen” in a Smart Home).

**The Speaker server** is a RESTful API, developed for easy direction of a playback request to an AmI Audio Client. As mentioned above, the Speaker Server keeps track of each connected AmI Audio Client’s output devices status, as well as the client’s location, allowing an auditory display request to be directed to any of the available output devices.

**AmI Speakers Explorer** is a web application/user interface for testing auditory displays and monitoring the status of AmI Audio Clients. It was developed for Auditory Display experts and AmI application developers willing to quickly view available locations and output devices and deploy tests while developing an application for an AmI environment (see Figure 1, left). The information is made available in real-time by the Speakers Server. In case a developer or a technician would like to add a new AmI Audio Client, they need to run the provided software at a compatible device (see Figure 1, right).

![Figure 1. AmI Speakers Explorer (left) & AmI Audio Client (right) user interface](image)

4.2 Auditory Display Director

The Auditory Display Director package’s components provide an extensive suite of interfaces for AmI application developers and Auditory Display experts to deploy meaningful and personalized auditory displays to AmI environment inhabitants. The Auditory Display Director package consists of the following components:

**The AmI Audio Library** was developed as a cloud-based database structure to store the complete variety of auditory representations found in auditory display design. The structure in which database items are stored enables retrieval in the form of an audio stream, including information regarding the representation, sound category and information type, as well as a list of descriptors/tags describing the semantic content of the representation. Items stored in the database are retrieved and used from various other modules of ACOUSMA.
**AmI Audio Library Explorer**, a content management system (CMS), was developed for the AmI Audio Library to enable searching, listening, generating and uploading auditory representations. It is a responsive and easy to use web application for all devices (see Figure 2, left), intended to be used by auditory display experts. The application offers three main functions: **Searching**. The user can easily search for a specific auditory representation filtering database item by representation type, sound category, information type or tag. **Adding a new representation**. The user can either choose to upload a created representation to the database or generate a new auditory representation. Currently, only the generation of Spearcons is supported. The auditory display expert can upload a new auditory representation to the database and add additional tags or remove the extracted ones before uploading. **Bookmarking**. The user can choose to bookmark auditory representations. Apart from the ability to listen to bookmarked auditory representations, the Auditory Display expert can view and copy their unique IDs to be used for testing with AmI Speakers explorer.

**Figure 2.** AmI Audio Library Explorer user interface (left), Simulation space infrastructure (right)

**The Auditory Display Server** is a RESTful API which exposes functions that application developers can query to: (i) Start and control a specific auditory display playback, (ii) Request and control a recommended auditory display playback (for a specific user). After each query, a specific workflow is followed. In the case of (i), a specific representation is given with a playback action, location or user and urgency level. All these parameters are passed to the queue manager in order to be sent to the Speaker Director for playback. In the case of (ii), specific input for the recommender is given, as well as the desired playback action and urgency level. The recommender then produces a recommended representation which is passed to the queue manager in order to be then sent to the Speaker Director for playback. In order to better understand the two workflow directions, the individual components that take part in the process will be analyzed.

An **API Client Library** was established in order for application developers to easily query and use exposed functions of the Auditory Display Server API. Developers with applications in different programming languages can add language specific modules that ease communication with the RESTful API. Currently, modules for C# and Typescript programming languages have been designed and developed for developers ranging from zero to expert knowledge in auditory displays. Practically, experts in auditory displays may want to have better control using the specific auditory display playback functions, while developers with no knowledge of the field can use the recommender for producing auditory displays.

**The Auditory Display Recommender** is a knowledge-based recommendation system designed using current state-of-the-art approaches in auditory representations and displays as well as this type of recommender and technique (Aggarwal, 2016). The aim of the recommender is to return a single recommended auditory display for a given query-target while keeping the recommended item personalized to user preferences. The recommender runs a similarity function for each feature found in the expanded target (T), comparing it to each feature found in every item in the auditory representation database (X). The score of each similarity function is regulated by a specific weight configured by auditory display experts. Each function has a feature specific asymmetric reward configured by user’s preferences. The refinement of the case-based recommender results works by using the technique of critiquing. Using different modalities provided by each AmI application, the user can critique the recommended auditory display. The critique is then used by the recommender to adjust the user’s profile and therefore the produced auditory display. Usually, a list of recommended auditory display is presented to the user for further critiquing. If a user likes a recommended auditory display, then the respective case in the case history is adjusted.
The Queue manager component evaluates queues and directs playback requests to the Speaker Director. For every location in an AmI environment, a queue to hold incoming requests is initiated. Queues are evaluated according to requests’ urgency. Effectively, that means that displays with higher urgency will precede those of lower urgency in the order of play. In the case of the emergency – urgency level, all current active displays will be superseded. Since multiple AmI applications may request different auditory displays for different users and locations at a single point in time, the Queue Manager ensures that those are handled appropriately and sent to the Speaker Director in an urgency-specific order. This, in combination with Speaker Director’s evaluation of contextual information, prevents the scenario of a user getting overwhelmed and confused by auditory displays. Therefore, auditory displays are presented according to applications demands and according to users’ activities and busyness of the auditory channel.

5. EVALUATION

Both a heuristic and a user-based usability evaluation were carried out for all modules of ACOUSMA. The ten well established heuristics by (Nielsen & Molich, 1990) were used and six experts in the field of HCI were invited to participate. The outcome of the evaluation indicated a total of 62 usability issues for the AmI Audio Library Explorer and 12 for the AmI Speakers Explorer interface. The major usability problems had to do with Visibility of system status, Flexibility and efficiency of use, and Error prevention. Those along with all minor issues were solved by adjusting or adding new functionality to both interfaces.

A weather application for an AmI environment was prepared for the user-based evaluation. A total of 17 people participated (7 people were assigned to the role of the auditory display expert, 6 to the role of the AmI environment inhabitant and 4 to the role of AmI application developer). The scenario was the following: “A construction company composed by architects, designers and programmers, builds apartments based on AmI technologies. With each purchased apartment, the company offers an interactive platform through which residents can install AmI applications. Those applications utilize shared resources and services provided by the apartment and display useful information in a variety of methods (audio, video) facilitating and improving the day-to-day lives of users. The company is currently utilizing a team of Auditory Display experts and AmI application developers to create a weather forecast application. The application, using audio, will promptly inform a user regarding the weather conditions at the location of a scheduled event on their calendar. For this purpose, the company uses the ACOUSMA system. Moreover, the company has a simulation space (simulating apartment rooms) to test applications before release. When complete, the application will be available for use through the apartments’ interactive platform”. Three use cases derive from this scenario: a) an auditory display expert using ACOUSMA to generate and test auditory representations for various weather conditions, b) an AmI application developer using the API client library module to initiate user-tailored auditory displays for various weather conditions, and c) an AmI environment inhabitant (apartment resident), experiencing the auditory result of the weather forecast application with the ability to critique. Each participant had to complete a number of tasks based on their role. In order to reenact the process of working on a weather forecast application, an interactive 3D space (see Figure 2, right) was utilized (Zidianakis, et al., 2021).

User-based evaluation for the role of Auditory Display Expert followed the think-aloud method (Someren, 1994) and through the completion of 11 tasks aimed at examining the usability of ACOUSMA. At the end, participants were asked to fill out the System Usability Scale (SUS) questionnaire (Brooke j., 1996) (Brooke J., 2013) and a number of additional questions that aimed at gathering statements regarding user’s perception of both interfaces functions. Results were positive with a high average SUS score of 94.17 and with all users having a positive perception of AmI Audio Library Explorer functionality and 83% of them being positive about the functionality of the AmI Speakers explorer.

User-based evaluation for the role of AmI Environment Inhabitant was designed to measure the effectiveness of the Auditory Display Recommender module as well as the perceived usability of an application enhanced with ACOUSMA’s features. To that end a blank preference profile was initiated letting the participant dynamically shape their profile through interaction with the weather forecast application and ACOUSMA. Participants followed 5 everyday scenarios that could take place in the smart home. According to each scenario, the participant would trigger events at the simulation space, initiating characteristic soundscapes, which were created to simulate real-life acoustic conditions. The weather forecast application would inform the participant about weather conditions according to each scenario. Moreover, the evaluator would collect comments regarding the understandability of the display. Also, the participant, using a tablet device, would critique the provided recommendation. At the end, participants were asked to fill the Usability Metric for User Experience (UMUX) (Finstad, 2010). The questionnaire was chosen as it has a high correlation
to the SUS (Berkman & Karahoca, 2016), making it ideal to capture perceived usability for this user-group, as the evaluation was burdened with extensive questions and comments. The application scored an average UMUX score of 89.3/100. The poorest score indicated the need for improvements in the effectiveness and efficiency of the ACOUSMA’s recommender. Additionally, 71.4% of participants could precisely understand the information being transmitted with the recommender having no prior knowledge of the user’s preferences, while the percentage increased to 85.7% after participants used critiquing.

Finally, for the role of the AmI application developer, after the user-based evaluation, a number of comments regarding the documentation and overall usability of the API were gathered. Participants were given a template project of a web-application and were asked to complete certain tasks. Moreover, participants had to set up an AmI Audio Client. Afterwards, they were asked to fill in the SUS questionnaire and were interviewed individually. Overall, three out of four participants commented that the API was well documented. Also, only one participant had trouble finding the right function to use.

6. CONCLUSIONS

This paper has presented the design, implementation and usability evaluation of ACOUSMA, a platform that encapsulates all functions necessary to standardize the integration of context-aware, personalized auditory displays in Ambient Intelligence environments. Using ACOUSMA’s interfaces and modules, auditory display experts can generate, upload, store and share auditory representations; AmI application developers can efficiently integrate meaningful auditory displays that are personalized and context-aware; AmI environment inhabitants can use the intelligent mechanism to critique and adjust auditory displays to their preferences at any time. The user-based evaluation of ACOUSMA yielded positive evidence that user requirements were successfully met. It is of course evident that, in order to ensure the fulfilment of all user requirements and usability of ACOUSMA, AmI applications must be developed to integrate its features. Results of the user-based evaluation have shown that ACOUSMA is ready to be used by AmI applications to enhance any AmI environment with auditory displays. The next step of this work is therefore to further evaluate the effectiveness of all ACOUSMA’s modules through the usage of ACOUSMA in AmI applications of various domains.

REFERENCES


