Proceedings of the 3rd IUI Workshop on Interacting with Smart Objects

Telecooperation Lab

Dirk Schnelle-Walka, Stefan Radomski, Jochen Huber, Oliver Brdiczka, Kris Luyten, Max Mühlhäuser (eds.)

Fachbereich Informatik
Telekooperation
Prof. Dr. Max Mühlhäuser

TECHNISCHE UNIVERSITÄT DARMSTADT

Table of Contents

Preface

You are now reading the Proceedings of the 3nd IUI Workshop on Interacting with Smart Objects.
The workshop was held in conjunction with the IUI 2014 at February 24, 2014 in Haifa, Israel [1].

Objects that we use in our everyday life are ever-expanding their interaction capabilities and provide functionalities that go far beyond their original functionality. They feature computing capabilities and are thus able to capture, process and store information and interact with their environments, turning them into smart objects [2]. Their wide range was covered by the submissions to this workshop. In his keynote, "The (Dumb) Internet of (Smart) Things" Carlos Duarte mentioned the following aspects about smart objects. They know something about their users and thus allow for natural interaction. Natural Interaction in contrast does not imply smartness. Smartness requires interaction with users and provides help. There are already commercialized products available that expose their properties and interaction capabilities. In order to enrich their potential and to lower affordances, they need to talk to each other. Here, making sense out of the available data is still an open research question. The overall goal should be to build an interactive ecosystem, that (i) seamlessy discovers, connects and talks to its environment, (ii) is ubiquitous and (iii) which allows the user to be in control.

The workshop examined these issues with regard to the following aspects:

- Voice-based interaction
- Support for people with disabilities
- Smart avatars
- Tangible Interaction
- Use of context Information

Hofmann et al. Came up with a definition of proactivity in voice-based interaction in cars. Their definition is derived from the following characteristics in work psychology (i) anticipatory, (ii) change-oriented and (iii) self-initiated. Tintarev et al. Introduced SasSy, a system that made decisions transparent with argumentation and natural language generation. Their systems enables smart objects to describe their behavior with the help of a rule-based knowledge source.
Karsenty et al. focused on input aspects for deaf people. They developed a glove that can be used to translate sign language to spoken text. Another support aid, especially for the blind, was presented by Grosse-Puppendahl et al.. Their goal was to ease integration of blind users into professional life. Therefore, they equipped compartments in a shelf with RFID tags that helped blind users in localizing compartments by acoustic notifications.
Amores et al. suggested smart avatars to design new narratives between technology and users with the help of avatars that are displayed in augmented reality. They also note that objects do not need electronics to become smart. Augmented reality can also be used to improve smartness of objects.
Nunes et al. discuss how the interaction space above multi-touch tables can be augmented as an interaction space. They provided TACTIC[1] an open source solution and

[1] http://accessible-serv.lasige.di.fc.ul.pt/~tactic/

suggest a setup for tangibles and tabletops to support continuous interaction. Other aspects of tangible interaction were introduced by Sellitsch with a context aware music player. The music player automatically seletcs a set of music that fits the current situation based on Thayer's mood model. Lozano et al. applied tangible interfaces to cognitive therapies, mainly to simplify and enhance interaction capabilities for people with special needs. Their system is implemented as a set of games to interactivley stimulate cognitive skills of children with intellectual disabilities. Finally, Hirsch et al. proposed an electroluminescent flexible screen that can be combined e.g. with touch interfaces where the surface size and shape are the main requirements.
Contextual information was exploited with regard to several aspects. Auferbauer et al. suggested a system that allows for a supervision of groups of children. They integrated sensors in children's cloths and provided an Arduino based device to a supervisor. Thereby, they provided quick access to environmental context under stress. Another important aspect of their work was to preserve privacy. In this light Vanderhulst et al. introduced geo fencing to address privacy issues when people use their smart phones to connect to smart devices in a smart city. Their research is focused around the claim that possession is insufficient to claim identity.

These proceedings contain the keynote from Carlos Duarte and twelve submissions around the different aspects of interacting with smart objects.

The 3rd Workshop on Interacting with Smart Objects was an interesting experience where participants with all their different backgrounds had lively discussions about interacting with smart objects. If you contributed to it in any way, we are grateful for your involvement. If you wish to participate in the future, please come and find out more information about the next workshop http://www.smart-objects.org and join the community. We wish that these proceedings are a valuable source of information in your efforts. We hope that you will enjoy reading the following pages.

May 2014

The program chairs,
Dirk Schnelle-Walka, Stefan Radomski, Jochen Huber, Oliver Brdiczka, Kris Luyten and Max Mühlhäuser

REFERENCES

[1] Schnelle-Walka. et al. 2014. 3rd IUI Workshop on interacting with smart objects. In *Proc. IUI '14*, 481–482.

[2] Kortuem, G. et al. 2010. Smart objects as building blocks for the Internet of things. *IEEE Internet Computing*. 14, 1 (Jan. 2010), 44–51.

Thanks

We would like to send out our thanks to everybody who has helped us to organize this event:

- the authors, who have written and submitted their papers to the workshop
- the program committee, for their time and effort to write substantial and constructive review reports

Program chairs

- Dirk Schnelle-Walka, Technische Universität Darmstadt, Germany
- Stefan Radomski, Technische Universität Darmstadt, Germany
- Jochen Huber, Singapore University of Technology & MIT Media lab, USA
- Oliver Brdiczka, Palo Alto Research Center, USA
- Kris Luyten, Hasselt University -tUL- iMinds Expertise Centre for Digital Media. Belgium
- Max Mühlhäuser, Technische Universität Darmstadt, Germany

Program Committee

- Bo Begole, Samsung, USA
- Marco Blumendorf, yetu AG, Germany
- Jingyuan Cheng, DFKI, Germany
- Aba-Sah Dadzie, University of Birmingham, UK
- Fahim Kawsar, Bell Labs, BE
- Alexander Kröner, Georg Simon Ohm Fachhochschule Nürnberg, DE
- Germán Montoro, UAM Madrid, ES
- Patrick Reignier, Grenoble INP, FR
- Boris de Ruyter, Philips Research, NL
- Geert Vanderhulst, Alcatal-Lucent Bell Laboratories, Belgium
- Massim Zancanaro, FBK, IT

The (dumb) Internet of (smart) Things

Carlos Duarte
LaSIGE, University of Lisbon
Portugal
cad@di.fc.ul.pt

ABSTRACT

With the proliferation of connected sensors, appliances and applications, the Internet of Things is finally breaking out of the lab and reaching out into the consumer homes. Smart sensors and smart appliances are now capable of controlling their operation based on environmental data. Owners can be aware of the status of their homes while away, through mobile applications. However, a truly useful Internet of Things is still not here. First, most sensors or appliances live in their own cloud, not knowing nor communicating with other sensors or appliances. Second, we haven't really understand how to take advantage of the potential that a connected environment of smart devices can bring to its inhabitants.

This talk addresses the problem of designing the Internet of Things from the peoples' perspective, not from the Things' perspective. While in certain contexts we expect an intelligent environment to be able to operate without requiring input from its users, this will not be true in every context; furthermore, most of the actions from this intelligent environment will impact its inhabitants. Consequently, we need to understand how people react to intelligent interactive environments, something which they are not accustomed to, in order to be able to properly design these environments and seamlessly integrate the Internet of Things within our lives.

Author Keywords

SmartObjects; Internet of Things; Interaction design; Affordances.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

General Terms

Design.

THE PROBLEMS

The number of so called Smart Objects has increased significantly in recent years. This trend is expected to continue, and even to accelerate in the coming years, with Gartner predicting that the total economic value add for the Internet of Things will be $1.9 trillion dollars in 2020[1].

While these objects offer an increasing array of features, and are able to sense a variety of information, their applicability is still limited because they do not share the sensed information with other objects. The true potential of Smart Objects will be revealed when information exchange between them is realized, improving their performance in specific tasks, and widening their range of applicability, leading to new areas and scenarios of use.

THE CHALLENGES

In order to make the above a reality, three challenges must be overcome:

- **Ensure all Smart Objects can communicate between themselves and with services offered on the Web** – Current efforts from ZigBee Alliance and ZWave Alliance strive to find a protocol to enable communication.
- **Make sense of all the available data** – Research efforts on semantic web, collective intelligence, information fusion, among others, will enable the exploration of all the data made available by the vast amount of SmartObjects.
- **Design an interactive ecosystem characterized by objects with disappearing affordances** – Users in this new interactive ecosystem will face the challenges of discovering with what to interact, when to interact, and how to interact. At the same time, they will expect ubiquitous, seamless interaction, with the right amount of intrusiveness, while remaining in control. Design principles, like reciprocity – where there is output, let there be input – will be paramount to overcome this challenge.

CONCLUSION

The future environments will be characterized by an array of *embedded* SmartObjects, contributing to an increasingly efficient *context awareness*, supporting both *personalized* and *adaptive* interaction, leading to what can be termed *anticipatory* interaction. To be able to successfully design these environments we will have to be able to fully characterize the context where activities will take place, using methodologies like *contextual design*. We will have to explore the boundaries of *plastic interfaces* and *distributed interaction*. And we must consider, above all, the overall *user experience*, giving the due importance to usability, hedonic and goodness attributes.

[1] http://www.gartner.com/newsroom/id/2621015

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel.

Proactivity in Spoken Dialog Systems

Hansjörg Hofmann, Ute Ehrlich, André Berton
Daimler AG
Ulm, Germany
{hansjoerg.hofmann, ute.ehrlich, andre.berton}@daimler.com

ABSTRACT
Proactive speech interfaces have been a hot research topic for many years. However, until today, no precise definition of proactive behavior in spoken dialog systems (SDSs) and its influencing factors has been made. Therefore, this paper aims at defining the characteristics of proactivity with the focus on SDSs. The definitions are derived from other research fields and then transferred to SDSs.

A general proactivity system model, which describes the relevant system components and their interaction is described. A proactive system receives information from a knowledge source and notifies the user about an incoming event without a user request. The system has to act user-friendly and take the current user state and the environment into account. Thus, the proactive behavior can be identified as anticipatory, change-oriented and self-initiated. A proactive human-machine speech dialog can be structured in 3 stages. First, the user has to be notified about an incoming event, then the problem solving process has to be started. Finally, the new task has to be completed and possibly paused tasks have to be resumed.

Author Keywords
proactivity; spoken dialog systems; situation-awareness

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Today, smartphones are considered as people's companions and are used in various daily situations. People use smartphones to browse the Web, buy things online or communicate via social media, Email and other (instant) messaging applications. Especially, the need to stay "always connected" has increased enormously within the last years and people exchange more and more over-the-top content (OTT) messages[1]. According to Informa Telecoms & Media each OTT messaging user sends an average of 32.6 OTT messages every day [3]. In online communication users take the initiative to interact but also the smartphone triggers the interaction and notifies user about new incoming events. This so-called "proactive" behavior will increasingly demand users as more and more messages are exchanged per day.

In some situation where the manual use of smartphones is not in focus, for instance when driving your car, this increased mental demand can be distractive and dangerous. In so-called dual-task scenarios people perform a secondary task (e.g. reading an email) in parallel to a primary task (e.g. driving), which requires the attention to several sources of information simultaneously. Wickens proposes that performing dual-tasks in parallel is achieved best, when the required user workload is distributed on several resources [15]. As many primary tasks in dual task scenarios tap haptic input and visual output channels speech interfaces are a good means to assist users in a comfortable and safe way [13]. Therefore, spoken dialog systems (SDSs) should be used to perform secondary tasks in a dual-task scenario and to notify users about new information.

Research about proactive SDSs has emerged in the last 10-15 years. However, up to today, no definition of proactivity has been introduced, yet. Related research investigates and characterizes some aspects of proactive behavior of speech interfaces but does not define proactivity and its requirements concretely (e.g., [10, 14, 4]). Therefore, this paper aims at defining the characteristics of proactivity in human-machine interaction (HMI) with the focus on SDSs. In the next Section, a general proactivity system model is presented, which describes the relevant components and influencing factors, which have an impact on proactive systems. In Section 3, a general definition of proactive behavior derived from literature and different research fields is presented, which can also be applied to the field of HMI. Subsequently, the definition of proactivity is transferred to SDSs and an overview about existing proactive SDSs is given. In the final Section, the findings are summarized.

PROACTIVITY SYSTEM MODEL
Imagine, two communication partners (*A* and *B*) would like to exchange information via any system. Therefore, *A* sends the information carrier to the system. When the system has received the information it has two possibilities to interact to handle the information. Either, the system stores the information and waits until *B* requests the information, or it proactively sends the information to *B*. The latter activity would

[1]OTT-messaging applications are downloadable smartphone applications, which enable users to send (instant) text messages for free, using mobile Internet access [3].

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel

resemble proactive behavior. Figure 1 illustrates a simple and generalized system model of such a proactive acting system in which *A* would be represented by the source and *B* by the user. The system which manages the information exchange could be a postal service, a smartphone or an in-car SDS. For example, imagine, *A* sends a letter to person *B* via any postal service. The postal service receives the letter from *A*. Instead of waiting until the recipient picks up the letter from the post office, the postman delivers the letter proactively to the addressee [8].

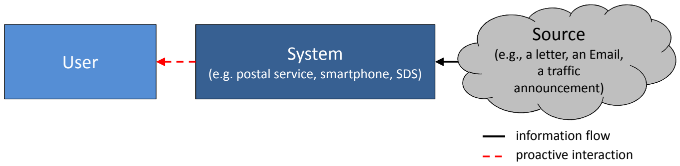


Figure 1. General proactivity system model.

Successful proactive behavior can only be achieved if the proactive system observes the environment in order to act in advance on a future situation and to deliver the content at the right point in time. Therefore, the system model needs to be extended by a context component, which is illustrated in Figure 2 [8]. Context-awareness can relate to the current location, time or situation, knowledge about user preferences, etc. The context knowledge can be gained by observing the environment and from the user himself. For example, if an addressee has changed his residence recently the postal service needs to be aware of the new address. If person *A* would like to send a letter to *B* without knowing the new address the postal service has to take care that the letter arrives at the correct address.

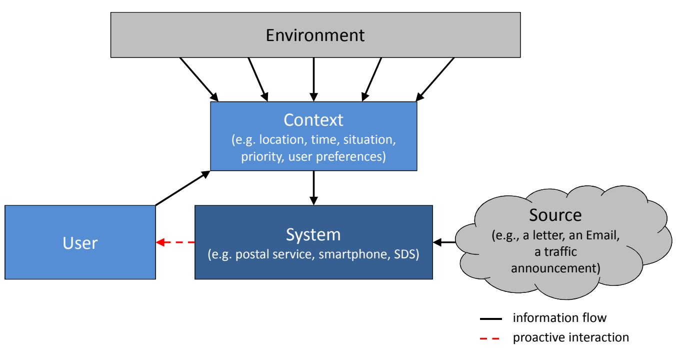


Figure 2. Extended proactivity system model.

If the user is performing a primary task in parallel, such as driving a car the system model needs to be extended (see Figure 3). Depending on the nature of the primary task the task can occupy several input and output channels simultaneously. For instance, by steering and keeping one's eyes on the road driving a car demands a person visually and manually. The state of the primary task needs to be included in the context knowledge. Thereby, the system might know if the user is currently able to process the information, which the system tries to deliver. E.g., if the user is very busy performing the primary task the system should not proactively interact with the user in order to not interfere the primary task.

Proactivity is a relatively new field in HMI and therefore, proactive behavior has not precisely been characterized, yet. Apart from HMI, research in human-human communication

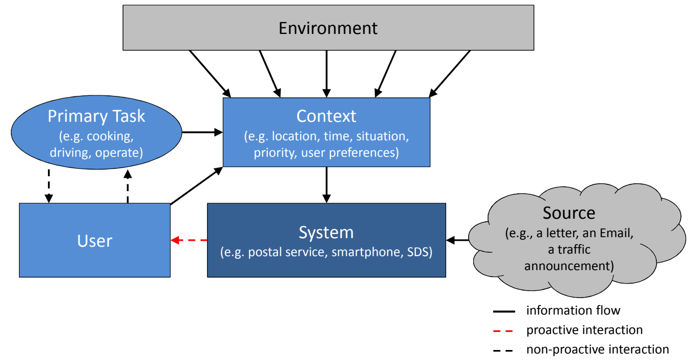


Figure 3. Extended proactivity system model in a dual-task scenario.

and industrial/organizational psychology investigate proactive behavior. The next Section presents a characterization of proactive behavior derived from literature and other research fields, which is applied on SDSs afterwards.

CHARACTERISTICS OF PROACTIVE BEHAVIOR

Dictionary definitions [1, 9, 16] typically contain two key features of proactivity. First, an anticipatory element is emphasized, which involves acting in advance of a future situation, such as "acting in anticipation of future problems, needs or changes" [9]. Second, these definitions highlight taking control and causing change, for example: "controlling a situation by causing something to happen rather than waiting to respond to it after it happens" [16]. These two elements - anticipation and taking control - can be found in most conceptualizations of general proactive behavior. E.g., [12, p. 636] define proactive behavior as "self-initiated anticipatory action that aims to change and improve the situation". In addition, definitions of proactive behavior often emphasize its self-initiative nature, which addresses the attempt to solve problems, which have not yet occurred [5]. Summarizing these definition, proactivity can be described by three key features (as defined in [11]): proactive behavior is

1. **anticipatory** - instead of reacting it involves scanning the environment and acting in advance to a further situation;
2. **change-oriented** - instead of passively adapting to the situation or waiting for something to happen being proactive means to take control or cause something to happen;
3. **self-initiated** - the control is taken on a self-initiative base without being requested to do so.

Most of the definitions are applied in work psychology and used to describe proactive behavior of employees in order to improve individual and organizational effectiveness. For example, a nurse, who is waiting for the doctor sees a patient and prepares the equipment and data the doctor might need. Thereby, the doctor can do his work more effectively. The nurse acts anticipatory by thinking ahead and anticipating the doctor's needs. Instead of waiting for the doctor to come she becomes active and prepares the equipment. The initiative to do so is taken all by herself without being requested by the doctor [11].

PROACTIVITY IN SPOKEN DIALOG SYSTEMS

As the definitions for proactive behavior are formulated in a general manner they can be transferred to other research areas, such as human-human communication or HMI, too. The focus of this research work is on SDSs. Therefore, the characteristics of proactive behavior and the system model theory are transferred to speech interfaces in the following. Furthermore, the process of a proactive speech interaction is explained in detail.

Proactive Behavior in Spoken Dialog Systems

The proactive behavior of an SDS can also be characterized by the 3 key features proposed by Parker et al. [11]. A proactive SDS should act *anticipatory*, *change-oriented* and *self-initiated*. As the proposed proactivity system model applies for all kind of proactive systems it can be applied to SDSs, too (see Figure 4).

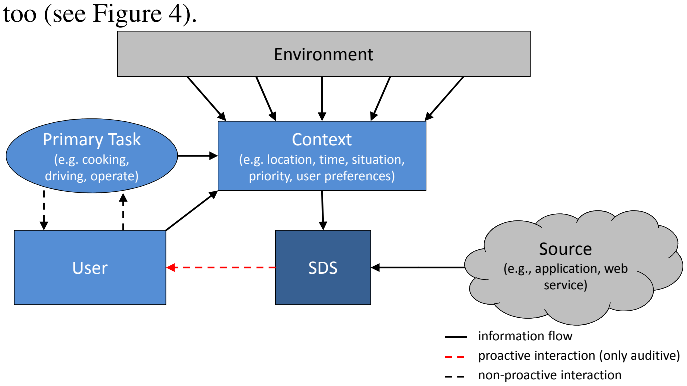


Figure 4. SDS proactivity system model in a dual-task scenario.

A proactive SDS receives the information from an application or a Web Service which is linked to the system. The SDS has to capture the spacial, temporal and user specific context of an interaction in order to take control anticipatory in advance to a further situation, possibly even before the users have become aware of the problem. Furthermore, the system needs to understand the user's current psychological situation, intention and actions and has to keep track of the dialog history. Then, it is able to assist the user in a meaningful way. A proactive SDS initiates the speech interaction itself and not only upon the user's request [10].

Imagine an in-vehicle navigation system, which observes the traffic density on the previously configured route while driving. As the system detects a traffic jam, which would prolong the length of the trip the system speaks up to the driver and suggests to take a different route. Here, the system acts anticipatory by observing the traffic density ahead and preventing the user from a possible traffic jam. Instead of ignoring the pending problem the system suggests to change the route to bypass the traffic jam. The system initiates the dialog itself without a request by the user. The only difference to the proactive behavior of the employee above is that the system only makes suggestions to the driver and does not decide about the new route itself. The driver has the control of changing the route himself.

Proactive Speech Interaction Process

The proactive speech interaction process can be structured in several stages. The dialog flow is illustrated in Figure 5 and the different dialog steps are described in the following.

Before the SDS delivers some new incoming information or informs the user about an upcoming problem two different scenarios are conceivable: Either the user is idle or the user interacts with the system already. When the SDS initiates the speech interaction, the different speech dialog steps have to be walked through:

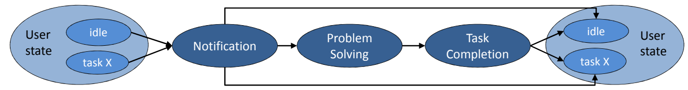


Figure 5. Proactive speech dialog flow.

1. **Notification**: First, the system has to grab the user's attention to tell him that there is some new information. The manner an SDS interrupts the ongoing dialog or initiates the interaction should be situation-sensitive and user-friendly [4]. If the situation does not allow a speech interaction at the moment, the system should not address the user. E.g., if the user is dictating a sensitive Email he should not be interrupted by the system. When the SDS decides to notify the user, in order to appear user-friendly, the SDS could allow the user to decide if he wants to enter into the new dialog or to reject talking about the newly introduced topic.
2. **Problem Solving**: In the course of the speech dialog the user interacts with the SDS on a regular basis. Depending on the dialog modeling and the competences of the SDS the speech interaction can appear more or less conversational.
3. **Task Completion**: When the problem has been solved or the new information has been delivered, the new task is completed. Depending on the initial state of the user the previous task should be resumed or the SDS should disappear again. Again, by negotiating the desired process the system could leave the decision to the user in a user-friendly manner.

In order to design a proactive SDS the different stages of the presented speech dialog flow should be taken into consideration. Several approaches to proactive speech interaction have been made within the last years. The most advanced projects and products are described in the following.

Overview on Proactive Spoken Dialog Systems

There are only few research projects who incorporate proactive behavior in SDS. The DARPA Communicator program[2] (2000-2001) focused on the improvement of SDSs, which allow for performing complex tasks by using speech as sole input modality. The DARPA projects helped to gain knowledge about proactive dialogue management. In the SmartKom project[3] (1999-2003) complex multimodal dialogs are aspired in which the user as well as the system can initiate interactions. Kwaku, as part of the Neem project [2] was a virtual meeting partner, which performed organisational tasks, such as monitoring the time spent on certain agenda points and reminds participants proactively to go to the next item, if necessary. Strauß et al. [14] envisaged an SDS which listens to multiparty conversations and assists the users proactively in a restaurant search and investigate the user behavior in a Wizard-Of-Oz study.

Today, there already exist products, which notify the user proactively. E.g. most of today's navigation systems employ data retrieved from the Traffic Message Channel (TMC)

[2]Websites offline.

[3]http://www.smartkom.org/

for routing taking into account real-time traffic situations [6]. The speech-enabled navigation systems prompt the driver proactively or play a warning sound if, for example a traffic jam on the route appears. Smartphones alert users about incoming emails or instant messages, upcoming appointments by playing sounds. Location-based information, gained knowledge about the user from the smartphone use and real-time data gathered from the Internet set the basis for successful context-awareness. The smartphone app Google Now[4] uses this context-knowledge to notify the user proactively about relevant information by presenting the information on the screen. E.g., when the user enters a subway platform he can see the schedule of the next trains leaving the station on his smartphone. Another example of a proactive situation-aware system is the Warning and Informationmanagement (WIM) system by Heisterkamp et al. [7], which ranks messages and warnings, which can occur while driving a car. Those warnings and messages are communicated to the user only in appropriate situations.

The overview of proactive systems shows that few projects or products exist, which address the employment of proactive SDSs in different environments. However, up to today, no SDS has been developed, which covers all stages of the presented speech interaction process and which satisfies the demands of proactive behavior in a user-friendly and situation-sensitive way. Research should investigate the desired interaction style of the different stages and the appropriate timing of the proactive interference.

CONCLUSIONS

This paper defined proactive behavior in SDSs and its influencing factors. The definitions are derived from other research fields and then transferred to SDSs. First, a general proactivity system model, which describes the relevant system components and their interaction has been described. A proactive system receives information from a knowledge source and notifies the user about an incoming event without a user request. The system has to act user-friendly and take the current user state and the environment into account. Thus, the proactive behavior can be identified as anticipatory, change-oriented and self-initiated. In a proactive human-machine speech dialog, first, the user has to be notified about an incoming event, then the problem solving process has to be started. Finally, the new task has to be completed and possibly paused tasks have to be resumed.

Future research should focus on investigating user-friendly speech interaction strategies for the 3 different stages of the proactive speech dialog flow. These investigations should take the influencing factors into account and consider different strategies for different situations.

REFERENCES

1. American Heritage Dictionaries Editors. *American Heritage Dictionary of the english language, fifth edition*. Houghton Mifflin Harcourt, 2011.
2. Barthelmess, P., and Ellis, C. A. The neem platform: An evolvable framework for perceptual collaborative applications. *Intelligent Information Systems 25*, 2 (2005), 207–240.
3. Clark-Dickson, P., Talmesio, D., and Sims, G. VoIP and IP messaging: Operator strategies to combat are threat from ott players (revised and updated). Tech. rep., Informa Telecoms & Media, 2013.
4. Edlund, J., C., O., and Gustafson, J. Investigating negotiation for load-time in the gethomesafe project. In *Proc. of IAST Workshop* (2012).
5. Frese, M., and Fay, D. Personal initiative (PI): An active performance concept for work in the 21st century. *Research in Organizational Behavior 23* (2001), 133–187.
6. Hamerich, S. Towards advanced speech driven navigation systems for cars. In *Proc. of IE* (2007), 247–250.
7. Heisterkamp, P., and Rothe, S. Daimlerchrysler mobile research environments layout, integration & paths to productization. Augmented Cognition Concept Validation Demonstration and PI Meeting, 2004.
8. Hermanutz, M. Entwicklung und evaluation von sprachdialogkonzepten im fahrzeug bei proaktivem systemverhalten. Master's thesis, Ulm University, 2013.
9. Merriam-Websters online dictionary. Proactive, 2013. Retrieved October, 15 2013, from http://www.merriam-webster.com/dictionary/proactive.
10. Minker, W., Lopez-Cozar, R., and McTear, M. F. The role of spoken language dialogue interaction in intelligent environments. *Journal of Ambient Intelligence and Smart Environments 1* (2009), 31–36.
11. Parker, S. K., and Collins, C. G. Taking stock: Integrating and differentiating multiple proactive behaviors. *Journal of Management 36*, 3 (2010), 633–662.
12. Parker, S. K., Williams, H. M., and Turner, N. Modeling the antecedents of proactive behavior at work. *Journal of Applied Psychology 91*, 3 (2006), 636–652.
13. Peissner, M., Doebler, V., and Metze, F. Can voice interaction help reducing the level of distraction and prevent accidents? meta-study on driver distraction and voice interaction. Tech. rep., Fraunhofer-Institute for Industrial Engineering (IAO) and Carnegie Mellon University, 2011.
14. Strauss, P.-M., and Minker, W. *Proactive Spoken Dialogue Interaction in Multi-Party Environments*, 1st ed. Springer Publishing Company, Incorporated, 2010.
15. Wickens, C. D., and Hollands, J. G. *Engineering Psychology and Human Performance*, 3rd ed. Prentice Hall, Upper Saddle River, NJ, 2000.
16. WordNet 3.1 Princeton University. Proactive, 2013. Retrieved October, 15 2013, from http://wordnetweb.princeton.edu/perl/webwn?s=proactive.

[4] http://www.google.com/landing/now/

Combining multi-touch surfaces and tangible interaction towards a continuous interaction space

Rafael Nunes
LASIGE, Universidade de Lisboa
rlnunes@di.fc.ul.pt

Nikolay Stanchenko
LASIGE, Universidade de Lisboa
stanch@lasige.di.fc.ul.pt

Carlos Duarte
LASIGE, Universidade de Lisboa
cad@di.fc.ul.pt

ABSTRACT

Multi-touch interaction is usually limited to one surface, even when combined with tangibles. Traditional scenarios where people interact with physical objects on and above tables or other surfaces have failed to be fully translated into existing technologies, such as multi-touch set-ups, which don't support natural interactions by combining the surface and the area above it into one continuous interaction space.

We aim to build and explore a set-up that allows users to benefit from a continuous interaction space on and above the table with multi-touch and tangible support. We expect to find and solve problems that can arise in various scenarios, both individual and collaborative.

A set of different existing technologies will be integrated to handle user interactions and an accompanying API will be developed and presented to serve as a tool for building future applications that draw from this kind of set-up.

Author Keywords

Tangible, Multi-touch, Continuous interaction space

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

General Terms

Human Factors; Design; Measurement; Experimentation.

INTRODUCTION

Multi-touch surfaces are emerging in an ever growing list of everyday scenarios. Traditionally, this type of set-up allows interaction with elements on the projected surface through touch input. This paradigm has been studied in many works ([12], [16]), and it became clear that it could benefit from some augmentations adding to the interaction experience. This is shown e.g. in [17], where a small room installation is proposed to explore a variety of interactions related to interactive displays and the space they inhabit; and in [10] which proposes an interaction technique and hardware sensor for detecting handoffs that use physical touch above the table.

But even these new scenarios aren't fully capable of supporting natural user behaviour with physical objects around a surface, such as object manipulation, exchange, moving and other actions, which can be present both on and above the surface. We wish to address the existing problem of combining digitally controlled interaction capabilities with real actions that are natural to the user in non-technological set-ups.

The area above the multi-touch surface can pave the way for new interactions when combined with the surface itself, functioning as a continuous interaction space. Gesture recognition adds natural user interaction *above* the surface, while touch recognition allows it *on* the surface. Physical objects also come natural to any user in everyday scenarios with traditional tables, and we wish to contribute towards a multi-touch augmented set-up that supports natural user interaction through touch, gestures and tangibles on and above the surface, by combining different technologies. To achieve this we have built a prototype of such a set-up.

An API is being developed to aid in building client applications for this type of set-up by blending different technologies into one form of input while providing a set of automatic tools for expected interactions and system behaviour.

In this paper we present our proposed set-up and the employed technology. The proposed API is explained in detail with a list of events and properties. Finally we describe our expectations for future work on our set-up and API along with conclusions we have drawn from this work.

RELATED WORK

Extensive research has been done on similar set-ups, but it has mainly focused in either *on the surface* ([7], [9]) or *Above the surface* ([14], [13]) interactions. *On the surface* interactions are traditionally actions like selecting, grabbing, throwing, rotating and moving while *Above the surface* interactions are typically point and select and accessing areas not reachable by the user's direct touch [11].

In [11], the continuous interaction space is defined as being composed of the direct touch surface and the space above (see figure 1). This allows new ways of interaction such as extended continuous gestures, which are described as gestures that start through direct touch on the surface and continue in the space above it. Mirrored gestures allow users to perform the same action via a gesture either directly on the surface

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel

or the space above it, supporting the user's natural behaviour by allowing to choose either a gesture on or above the surface to interact with the digital content. It is also possible to pick and drop digital objects through physics-based interaction by grabbing a digital object on the surface and pulling it upwards. These gestures take advantage of the continuous interaction space by removing traditional restrictions and allowing both modalities to work together. However, this work does not support manipulation of actual physical objects in the continuous interaction space which is a necessary step towards natural user behaviour.

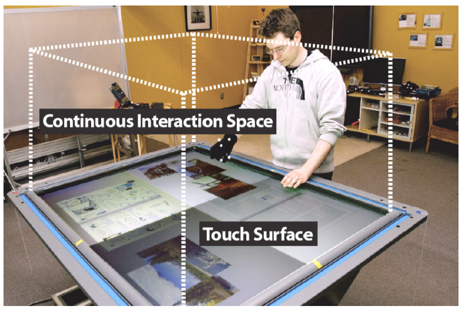


Figure 1. Continuous interaction space, as depicted in [11]

Although there is an extensive amount of research in multi-touch set-ups, few works have focused on user identification for this continuous interaction space. User identification enables new interaction possibilities by making the interface aware of which user is interacting with it. In [6] a proximity-aware multi-touch table is presented. Using 138 proximity sensors it detects a user's presence and location, determine body and arm locations, as well as distinguishing between right and left arms and map touch points to specific users and hands. In [15], a technique is proposed for user identification on interactive surfaces that enables users to access personal data on a shared surface, associate objects with their identity and customize appearance, content or functionality of the user interface. This type of research is relevant when considering collaborative scenarios, where multiple users interact with the same system at once. In [8] research was made on collaborative workspaces in which multiple users work on the same data set. Users would sit on special chairs that attribute different signals to each user, allowing the surface to know where it was being touched and by whom. This was further explored in [10], where *electroTouch* was used to detect interactions between users above the surface by detecting current interruptions.

SET-UP COMPONENTS

Different technologies allow our set-up to detect user interactions throughout the continuous interaction space. The following sections describe how these technologies were deployed and what they offer towards our continuous interaction space.

Different technologies allow our set-up to detect user interactions throughout the continuous interaction space. The following sections describe how these technologies were deployed and what they "bring to the table".

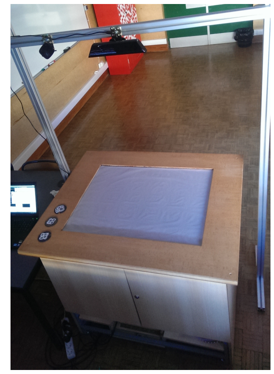


Figure 2. Our multi-touch set-up. Two cameras are placed inside and above the surface

On the surface

We use the Frustrated Total Internal Reflection (FTIR) [1] method for touch interactions on the surface, which allows us to detect touch input through an array of infra-red light. A camera inside the set-up captures this light, which is then interpreted and translated into TUIO [5] protocol messages for client applications.

We decided to use *reacTIVision* [4] to achieve object tracking on the surface by tracking fiducial markers that are placed on physical objects. Since the previous camera inside our set-up had an infra-red filter, it could not detect visible light, which was a necessary feature for the fiducial tracking. As such, a second camera was introduced to capture the fiducial markers and translate them into TUIO [5] protocol messages for client applications.

Above the surface

We used ThreeGear [3] to detect precise finger and hand tracking above the surface using one Kinect camera, as seen in Figure 2. This system allows detection of small gestures, like pinching and wrist movements instead of traditional arm detection. It is coupled with a corresponding API that allows writing software applications based on its technology.

Tangible interactions above the table are inferred through gesture recognition. After an object is placed on the surface, it is registered in the system as having hand *A* holding it, and can be tracked even after being lifted by tracking the hand holding it. Given that we cannot be always sure that the object is still being held by the same hand, we decided to complement this process by a camera above the surface that allows fiducial tracking. This allows confirmation of which hand is holding the object, instead of requiring this information to be inferred.

Although the camera above removes the need to register tangibles on the surface before using them above, one cannot rely solely on fiducial tracking from above, since the marker can be easily lost due to other objects or hands moving over it. This is especially problematic in collaborative scenarios. Constant management of inferred and confirmed data from these sources is needed to achieve the best results, while removing limitations from the user, thus allowing natural interaction.

Bridging the components

Having so many different technologies is quite cumbersome when considering communication between client applications and the various sources of input. We deployed RabbitMQ [2] messaging middleware in our system architecture (see figure 3), to allow seamless communication between any components. The API manages all the information from RabbitMQ and TUIO to deliver events to client applications.

Any set-up component can use RabbitMQ to publish and subscribe to events, thus all the different technologies and languages are easily bridged, making our system highly modular, since it is easy to add new components without making changes to previous configurations.

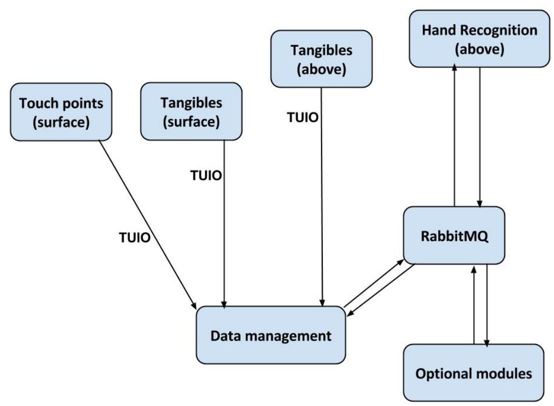


Figure 3. System architecture

Solving Occlusion

Since fiducial tracking is supported by a camera that captures visible light, surface projections can get in the way and cause missing fiducial markers in a tracking scenario. To solve this problem, we searched for a background color that would allow easy and full fiducial tracking on the surface and applied it to an *object follower*. An *object follower* is a circle that surrounds the fiducial when it is tracked for the first time and constantly moves below it while maintaining itself above any projection. Optionally, tangibles can be rotated to increase or decrease the radius of *object followers*. This way we can be sure that the color below the fiducial will always be the desired color for tracking and significantly reduce the probability of miss-tracking a fiducial marker.

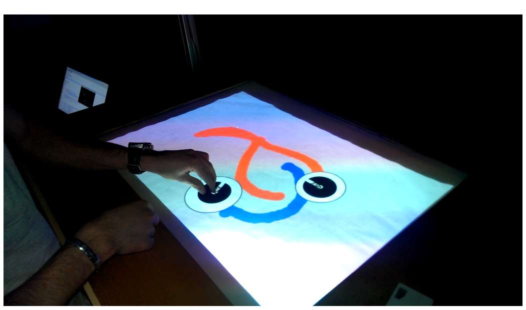


Figure 4. API controlling *object followers* for tangibles

In Figure 4 tangibles are being used as painting brushes with different colors. This scenario can rapidly cause color confusion in the projection, damaging the system's tracking abilities. Thanks to the *object follower* method, objects are easily tracked even when moving above painted zones.

BUILDING AN API

Event handling

Our system uses an extensive amount of technologies, so when considering the future of this work, we felt that an API would serve us greatly towards the development of applications. The API is fully implemented in JavaScript to support building HTML client applications. We have, so far, introduced some basic events and features to this API. These will be explained below:

API Events	
Events triggered whenever an object **enters**, **moves** or **leaves** the surface	**object_added** **object_updated** **object_removed**
Events triggered whenever an object **enters**, **moves** or **leaves** an element that is expecting this event	**object.added** **object.updated** **object.removed**
Event triggered whenever a **touch** is tracked **inside** an element that is expecting this event	**touch.press**
Event triggered whenever a **touch** already being tracked **moves** inside an element that is expecting this event	**touch.update**
Event triggered whenever a **touch** is **no longer** tracked inside an element that is expecting this event	**touch.release**

Element Properties

Some properties can be easily attached to HTML elements by adding the respective CSS class to them. This way the API saves the user the trouble of making extra calculations. Next we detail a set of classes that can be added to elements and the properties they receive:

- **movable**

  A movable element is automatically moved by the API whenever a touch is registered inside it and movement follows. When the touch is released the element stays in the new position.

- **touchable**

  A touchable element receives events related to touch inside the area that corresponds to it, and can then respond to those events in whatever way the user wishes.

  Events received: **touch.press**; **touch.update**; **touch.release**

- **object-aware**

  An object-aware element receives events related to object tracking inside the area that corresponds to it and can then respond to those events in whatever way the user wishes.

  Events received: **object.added**; **object.updated**; **object.removed**

FUTURE WORK

We want to find new ways to achieve user identification in our set-up, while making it easier to use in collaborative scenarios, reducing confusion between multiple user interactions. Hand recognition above the table can be expanded to provide multiple hand pair recognition, thus allowing identification of user interactions by the corresponding set of hands.

The API will be improved to add more events that we consider necessary for application developers. The algorithm for constant swapping of information between inferred and confirmed data above the table is going to be improved to achieve better results and reduce errors as much as possible. We will run tests to gather data on development of applications for this type of set-up with our API.

A set of applications will also be developed to test our set-up and validate our findings towards natural user interactions within the continuous interaction space.

CONCLUSIONS

We proposed an augmented multi-touch set-up that supports multi-touch and tangible interactions *on* and *above* the surface merging both into one continuous interaction space. By combining a set of technologies this set-up is capable of supporting natural user interactions with various modalities without interruptions when transiting from the surface to the area above it, and back.

An accompanying API was developed to support future advancement for this type of set-up by integrating all of these technologies into one single form of input. By providing a set of events and tools, the API allows to build client applications with ease.

ACKNOWLEDGMENTS

This work was supported by FCT through funding of SiTECOSE project, ref. PTDC/EIA-EIA/117058/2010, and LaSIGE Strategic Project, ref. PEst-OE/EEI/UI0408/2014

REFERENCES

1. Frustrated total internal reflection. `http://wiki.nuigroup.com/FTIR#Links:`, Dec 2013.
2. Rabbitmq. `http://www.rabbitmq.com`, Dec 2013.
3. Threegear systems. `http://www.threegear.com/technology.html`, Dec 2013.
4. A toolkit for tangible multi-touch surfaces. `http://reactivision.sourceforge.net/`, Dec 2013.
5. Tuio 1.1 protocol specification. `http://www.tuio.org/?specification`, Dec 2013.
6. Annett, M., Grossman, T., Wigdor, D., and Fitzmaurice, G. Medusa: a proximity-aware multi-touch tabletop. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST ’11, ACM (New York, NY, USA, 2011), 337–346.
7. Cao, X., Wilson, A., Balakrishnan, R., Hinckley, K., and Hudson, S. Shapetouch: Leveraging contact shape on interactive surfaces. In *Horizontal Interactive Human Computer Systems, 2008. TABLETOP 2008. 3rd IEEE International Workshop on* (2008), 129–136.
8. Dietz, P., and Leigh, D. Diamondtouch: a multi-user touch technology. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, UIST ’01, ACM (New York, NY, USA, 2001), 219–226.
9. Epps, J., Lichman, S., and Wu, M. A study of hand shape use in tabletop gesture interaction. In *CHI ’06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’06, ACM (New York, NY, USA, 2006), 748–753.
10. Jun, L., Pinelle, D., Gutwin, C., and Subramanian, S. Improving digital handoff in shared tabletop workspaces. In *3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems (TABLETOP 2008)* (2008), 9–16.
11. Marquardt, N., Jota, R., Greenberg, S., and Jorge, J. A. The continuous interaction space: interaction techniques unifying touch and gesture on and above a digital surface. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part III*, INTERACT’11, Springer-Verlag (Berlin, Heidelberg, 2011), 461–476.
12. Park, S. Implementation of new gestures on the multi-touch table. In *ICT Convergence (ICTC), 2012 International Conference on* (2012), 168–169.
13. Parker, J., Mandryk, R., and Inkpen, K. Integrating point and touch for interaction with digital tabletop displays. *Computer Graphics and Applications, IEEE 26*, 5 (2006), 28–35.
14. Parker, J. K., Mandryk, R. L., and Inkpen, K. M. Tractorbeam: Seamless integration of local and remote pointing for tabletop displays. In *Proceedings of Graphics Interface 2005*, GI ’05, Canadian Human-Computer Communications Society (School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005), 33–40.
15. Schmidt, D., Chong, M. K., and Gellersen, H. Handsdown: hand-contour-based user identification for interactive surfaces. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, NordiCHI ’10, ACM (New York, NY, USA, 2010), 432–441.
16. Schöning, J., Hook, J., Bartindale, T., Schmidt, D., Oliver, P., Echtler, F., Motamedi, N., Brandl, P., and Zadow, U. Building interactive multi-touch surfaces. *Tabletops-Horizontal Interactive Displays* (2010), 27–49.
17. Wilson, A. D., and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, UIST ’10, ACM (New York, NY, USA, 2010), 273–282.

Preserving Privacy in Social City Networks via Small Cells

Geert Vanderhulst
Alcatel-Lucent Bell Labs
Copernicuslaan 50
2018 Antwerp, Belgium
geert.vanderhulst@alcatel-lucent.com

Fahim Kawsar
Alcatel-Lucent Bell Labs
Copernicuslaan 50
2018 Antwerp, Belgium
fahim.kawsar@alcatel-lucent.com

ABSTRACT

An increasingly large amount of small cells – e.g. WiFi hotspots – is being deployed in residential areas to connect a plethora of smart devices to the Internet. In this paper, we present a social city network leveraging small cells for sharing content geographically and temporarily whilst preserving the privacy of its users. Unlike a social network built around friends, we propose a social city network addressing geographically co-located people and smart objects, e.g. residing in a street, on a square, around a building, etc. Our goal is to facilitate interaction with smart cities by easily sharing short-lived data fragments with others in a given area and for a limited time span. To this end, we designed an architecture in which small cells deliver location proofs that grant access to location-restricted content.

INTRODUCTION

Social networks have radically changed the way content is shared on the Web. By following someone on Twitter or becoming one's friend on Facebook, relationships between people – family, friends, colleagues – are explicitly composed. Also location data retrieved by social networking applications is often used to associate shared information with a point on the map, e.g. checking in at a place on Foursquare or Google Places. The impact of such networks has grown beyond sharing one's social status to marketing a business (e.g. collecting 'likes') and engaging in a dialogue with customers (e.g. after sale service). Yet, within the scope of a city, present social networks still lack intuitive means to disseminate content that is constrained in place and time. Rather than addressing known individuals, we envision a location-based social network [8] in which content is shared with a street, a square, a part of a city, . . . and only for a limited period. Such a social network also fits the vision of the Internet of Things where smart objects autonomously exchange social facts derived from sensor data with other objects in the direct vicinity or across city boundaries [1]. The notion of friendly objects and associated configuration work for setting up proper sharing policies could vanish if users can be assured that shared data is short-lived and constrained within geographical boundaries rather than being publicly available on the Internet.

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel.

To preserve a user's privacy when using a city's social network, shared content thus must be contained within the indicated spatial and temporal boundaries. However, present location-based services typically rely on an implicit trust relationship between content provider and consumer: it is assumed that a user indeed resides at the location he claims to be at. As location coordinates can easily be falsified, a more robust solution is needed to prevent users from lying about their whereabouts and hence obtain unauthorized access to shared content. In this paper, we leverage small cells as witnesses that can testify a user's presence in a given area at a given time. We contribute to the state of art with an adapted algorithm for generating secure location proofs that preserve the privacy of the prover. With a location proof, a user can authenticate herself to the social city network and access location-restricted content.

RELATED WORK

GeoLife [12] is a location-based social network which enables users to share life experiences and build connections among each other using human location history. Similarity matching of location trajectories is applied to generate friend and travel recommendations. In our system no explicit relationships between people or objects are established. Everyone visiting an area within a period that content is shared, can interact with it.

Previous works have studied the use of WiFi hotspots to offload traffic from cellular networks [2, 9] and propose social participation to propagate data through opportunistic networks [5]. In these cases, local delivery nodes are exploited to increase network efficiency and overcome overloaded cellular networks. Likewise, access to shared content can be regulated via services running in a hotspot's local network [11] which requires a user to be connected with a specific wireless node to download a piece of data. As opposed to this approach, we exploit small cells as witnesses that can testify the presence of a user at a certain place and time.

Several systems have also been proposed to give users the ability to prove that they were in a particular place at a particular time [10, 6, 13]. Some systems rely on computing an upper bound of the user's distance, e.g. by measuring the round-trip time of a wireless signal [4]. Other approaches which do not depend on dedicated hardware, obtain proofs from wireless access points (e.g. VeriPlace [6]) and Bluetooth devices (e.g. APPLAUS [13]). Our privacy system is based on a simplified version of the APPLAUS architecture, yet adapted to hotspots as trusted witnesses instead of Bluetooth devices. In many usage scenarios, location proofs are used as evidence for later, e.g. to proof to a teacher that all classes were attended or to detect loyal customers. Instead, we leverage them as on-site as authentication tokens to gain access to location-restricted content. IP-to-Geo schemes [7] are often too inaccurate for this matter and can easily be tricked using proxies. On the one hand, we want to protect the privacy of the content provider by restricting the place and time where and when content can be accessed, but on the other hand we also need to preserve the privacy of the content consumer by ensuring that the latter's identity and current location are obfuscated in a proof. As pointed out in [13] and [6], this can be achieved by distributing trust among multiple parties involved in the provisioning and verification of location proofs.

SOCIAL CITY NETWORK

In this section we elaborate on a design of a social city network service in which messages and media can be posted, similar to Twitter. The social city network sets itself apart from other social networks as (i) content is shared within specific geographic areas rather than with a predefined set of people and can only be accessed by users physically residing within this area; and (ii) content is only made available for a limited duration and then disappears again – hence it does not circulate in the network until it is explicitly deleted. Consider the following motivational use cases:

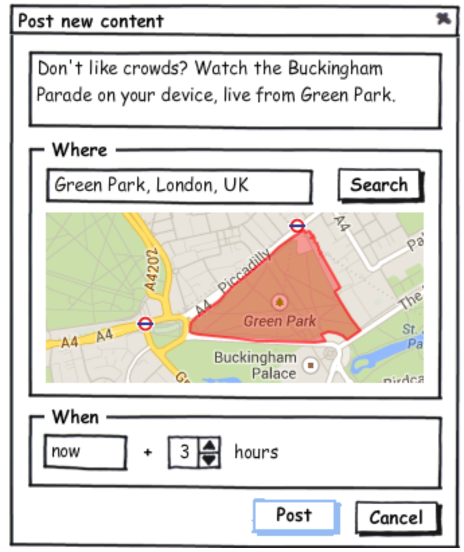
1. The water company announces construction works in Arlington Road and indicates an expected cut of the water supply tomorrow between 8 am and 12 am. Live updates on the works are propagated via the social city network to the residents of the street.

2. A restaurant owner interacts with the city's social network to advertise daily lunch specials to people in the neighborhood. This information is shared between 11 am and 1 pm within a 10 km radius of the restaurant.

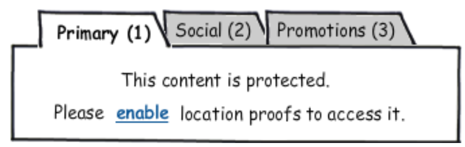
3. During a summer festival in Regent's Park, camera feeds capturing the stage from different angles are made available to the local audience. Hence people at the festival who have no clear view on the podium can still watch the performance on their mobile devices.

4. A smart car automatically broadcasts a flat tire to approaching vehicles, i.e. shares a situation with a particular road segment.

Figure 1 shows a user interface prototype for a mobile application by which users can engage with the social city network and is further elaborated on in the remainder of this section.



(a) Posting content.



(b) News feed

Figure 1. User interface mock-up of an application for interacting with the social city network.

Posting Content

When posting new content to the social city network using the dialog depicted in figure 1, users need to specify *where* and *when* a fragment is shared. For the *where* part, addresses or place descriptions can be specified or an area can be indicated on the map using selection tools. We use OpenStreetMap[1] data to translate place names into geometric shapes composed of location coordinates, e.g. the geographic boundaries of a park in the figure. This spatial data is stored along with the content in the social network as it determines the location from where the content can be accessed. For the *when* part, users can specify a start time and a duration after which the content is deleted. By focusing on short-lived data, we anticipate privacy concerns and promote the news value of published content. Moreover, a desired privacy and security policy can be enforced in our system: trust a user's provided location or require a location proof – the latter being selected as default.

News Feed

The news feed mimics the layout of Google Mail where social and promotional messages are separated from a primary feed as illustrated in figure 1. These feeds are

[1] `http://www.openstreetmap.org`

populated with content that is revealed based on the user's location. Hence, when moving through the city, the news feed will continuously update. Note however that our approach is different from e.g. Foursquare where information about a place (e.g. restaurant reviews) is made available on-site. Displayed content is not necessarily linked with the user's current location, but it has rather been shared within the area the user currently happens to be in. To retrieve protected content, users need to authenticate via a location proof which is retrieved from small cells in the vicinity as discussed in the next section.

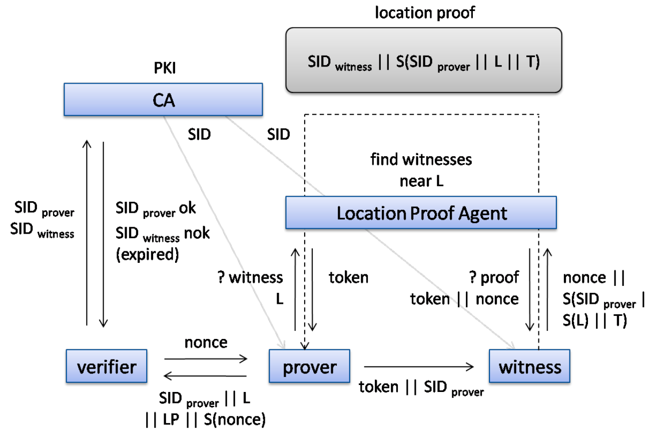
LOCATION PROOFS

Since location coordinates can be spoofed easily, we need a mechanism to verify that a user indeed resides at a claimed position. Even if access to shared content is restricted to a number of small cells, there is no guarantee that an individual is within the range of e.g. a WiFi hotspot as a local proxy can route network traffic from anywhere in the world to the hotspot (e.g. wormhole attack). To protect against this, we modified an algorithm presented in [13] for handing out personalized location proofs and leverage it for instant location-based authentication. The location proof authorizes a user to access content that resides on the Internet (i.e. in the cloud) but that was only shared within a particular city. These proofs are handed out by small cells which act as unique trusted witnesses (i.e. when managed by a trusted provider) that confirm the presence of a user.

In the next sections, we discuss our location proof architecture supported by figure 2, and identify the role of the different entities in the context of a social city network.

Algorithm and Architecture

As illustrated in figure 2(a), several entities are involved in the provisioning and verification of location proofs to preserve the privacy of its users. No single entity is aware of both the identity and the location of a user at any moment. Provers and witnesses communicate with each other and a Location Proof Agent (LPA) using secret identities which also serve as public encryption keys. To this end, provers and witnesses first identify themselves to a Certified Authority (CA) that provides them with secret identifiers (public keys). To obtain a Location Proof (LP), a prover queries the LPA for nearby witnesses that can confirm its location. The LPA responds with a list of witnesses (e.g. WiFi hotspot identifiers) and notifies nearby witnesses that a prover wants to acquire a location proof using the shared token as a reference. The prover then contacts a witness within range using the shared token and its secret identifier – i.e. connects to a hotspot with a given SSID and interacts with a service in the hotspot's network. The witness generates an intermediate LP that consists of the location of the witness (L), the current time (T) and the prover's secret identifier (SID_{prover}). This data is signed with the witness' private key ($S(...)$) and forwarded to the LPA. The



(a) Different parties involved in the provisioning and verification of proofs.

CA	LPA	Prover	Witness	Verifier
ID prover	L prover	SID witness	SID prover	SID prover
ID witness	L witness	L witness		SID witness
SID prover	SID prover			L prover
SID witness	SID witness			L witness

(b) Preserving privacy: who knows what? (ID = identity, SID = secret reference to identity, L = location)

Figure 2. Location proof architecture.

LPA verifies the identity of the witness using its public key ($SID_{witness}$) and composes a final LP which is passed on to the prover. This location proof is then sent by the prover to the verifier which checks with the CA whether the secret identifiers of prover and witness are not expired and if the LP is valid, i.e. signed by the witness. Note from figure 2(b) that the LPA is only aware of the location of a user and the CA only knows the identity of a user. By distributing this information amongst different parties run by different organizations, the user's privacy is protected. To discourage cheating, we let location proofs expire (similar to session cookies) and limit the number of proofs per IP address. To increase trust, a content provider may also demand for multiple location proofs handed out by different witnesses.

Applied to the Social City Network

At the heart of the social city network, a verifier process regulates access to shared content. When a valid location proof is received, location-restricted content is unlocked for the corresponding user for a predefined time span. After that, the user's session expires and a new location proof must be provided as the user might have moved to a different place from where the content can no longer be accessed. To guarantee uninterrupted access to location-restricted content, a mobile application – running a prover process – can pro-actively collect location proofs and (re)authenticate to the social city network. Given the continuously growing network of WiFi hotspots, we believe that WiFi hotspots in particular are suitable candidates for generating these proofs. WiFi nodes are readily being used as location beacons due to their stationary nature and the limited range of their radio signals. Many residential and public hotspots are also managed by telecom providers which can fulfill the role of Location Proof Agent. Note that the underlying authentication mechanism is completely transparent for the end-user of our service: a mobile application (prover) retrieves a location proof from a WiFi hotspot (witness) and passes it to the social city network (verifier) – no manual interaction is required.

CONCLUSIONS AND FUTURE WORK

A social city network built around *where* and *when* content is shared enables several new use cases that are hard to realize with present social networks. In our approach, content is shared within the boundaries of a street, a park or a custom geographic area and automatically expires. Instead of composing static relationships between people or objects, we enable users to reach out to crowds of people that are connected by the places they visit. Even so in an emerging world of connected objects that generate massive amounts of data, spatio-temporal sharing can assist in delivering the right facts to the right place at the right time. In this work, we have explored the technical requirements for safeguarding the privacy of the users of a social city network and propose location as an unobtrusive authentication mechanism. Although we have mainly focused on public WiFi hotspots as delivery vehicles of location proofs, the presented techniques are also applicable to other wireless nodes with limited coverage like Femtocells [3].

Further effort needs to be spent in combining our application prototypes and conceptual architecture into a real-world proof of concept implementation. Other directions for future work include quality control and ownership management of shared content. Possible pitfalls of a location-based social network service are inappropriate content postings and users pretending to be someone else. The former can be addressed via a crowd-sourced voting system (down-voting inappropriate content) or a cost model that attributes a fee to advertising messages based on the target area and the lifetime of the content. While content might be retrieved anonymously via location proofs, further research is needed to prevent identity abuse when posting content.

REFERENCES

1. Atzori, L., Iera, A., Morabito, G., and Nitti, M. The Social Internet of Things (SIoT) - When Social Networks Meet the Internet of Things: Concept, Architecture and Network Characterization. *Computer Networks 56*, 16 (Nov. 2012), 3594–3608.
2. Balasubramanian, A., Mahajan, R., and Venkataramani, A. Augmenting Mobile 3G using WiFi. In *International Conference on Mobile Systems, Applications, and Services (MobiSys'10)*, ACM (2010), 209–222.
3. Chandrasekhar, V., Andrews, J. G., and Gatherer, A. Femtocell Networks: A Survey. *IEEE Communications Magazine 46*, 9 (2008), 59–67.
4. Ferreres, A. I. G. T., Alvarez, B. R., and Garnacho, A. R. Guaranteeing the Authenticity of Location Information. *IEEE Pervasive Computing 7*, 3 (2008), 72–80.
5. Han, B., Hui, P., Kumar, V. S. A., Marathe, M. V., Shao, J., and Srinivasan, A. Mobile Data Offloading through Opportunistic Communications and Social Participation. *IEEE Transactions on Mobile Computing 11*, 5 (2012), 821–834.
6. Luo, W., and Hengartner, U. VeriPlace: a Privacy-Aware Location Proof Architecture. In *International Conference on Advances in Geographic Information Systems (GIS'10)*, ACM (2010), 23–32.
7. MaxMind. GeoIP Database. http://www.maxmind.com/en/ip-location/.
8. Microsoft Research. Location-based Social Networks. `http://research.microsoft.com/en-us/projects/lbsn/`.
9. Pitkänen, M. J., Kärkkäinen, T., and Ott, J. Opportunistic Web Access via WLAN Hotspots. In *International Conference on Pervasive Computing and Communications (PerCom'10)*, IEEE (2010), 20–30.
10. Saroiu, S., and Wolman, A. Enabling New Mobile Applications with Location Proofs. In *Workshop on Mobile Computing Systems and Applications (HotMobile'09)*, ACM (2009), 3:1–3:6.
11. Vanderhulst, G., and Trappeniers, L. Public WiFi Hotspots at Your Service. In *International Conference on Pervasive Computing and Communications (PerCom'12 Workshops)*, IEEE (2012), 411–414.
12. Zheng, Y., Xie, X., and Ma, W.-Y. GeoLife: A Collaborative Social Networking Service among User, Location and Trajectory. *IEEE Data Eng. Bull. 33*, 2 (2010), 32–39.
13. Zhu, Z., and Cao, G. APPLAUS: A Privacy-Preserving Location Proof Updating System for Location-based Services. In *International Conference on Computer Communications (INFOCOM'11)*, IEEE (2011), 1889–1897.

A Sign-to-Speech Glove

Olga Katzenelson
Solange Karsenty
Hadassah Academic College
HaNeviim 37, Jerusalem, Israel
olga.katzenelson@gmail.com
solange@hadassah.ac.il

ABSTRACT

In this paper, we describe a smart glove – JhaneGlove - that turns sign language gestures into vocalized speech via a computer to help deaf people communicate easily with people who do not understand the sign language. We have developed a handmade device: a glove with sensors, and the software that will transform signs into text. Text is converted into speech using standard software. A neural network based agent can learn the gestures interactively and allows users to define new signs. As a result, each user can have a custom sign language independent from other users. We present the system and early experimental results.

Author Keywords

Sign language, neural networks, hand-shape recognition, gesture recognition, glove

ACM Classification Keywords

K.4.2 [Computers and Society]: Social Issues—Assistive technologies for persons with disabilities; H.5.2 [Information Interfaces and Presentation]: User interfaces—Input devices and strategies (e.g., mouse, touchscreen)

INTRODUCTION

Our environment is continuously evolving with new computerized and electronic systems that are changing the way we live and interact with others. Electronics become smart, and sensors are everywhere, thereby enabling the development of devices that can analyze, recognize and interpret our actions. Communication is perhaps one of the major concerns when developing such systems, and it is a major concern at the Hadassah Academic College where projects of the Computer Science department can be initiated in direct collaboration with the Communication Disorders department.

While most deaf individuals with various degrees of deafness can lip-read and understand someone speaking, they almost always communicate with sign language. This language is usually unknown by people around them. A typical situation is a person (student) who wants to use sign language to give a presentation to other students who do not understand the sign language. To enable this type of communication, we have built an experimental system for the automatic translation of sign language into text and speech.

PREVIOUS WORK

A sign is a language that uses manual communication and body language to convey meaning. This can involve simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to fluidly express thoughts. Some systems attempt to recognize as a whole the body language, for example using the kinect [1] that combines cameras and microphone to capture 3D motion and voice, enabling a wide range of applications. Other systems, based on more portable devices such as a glove, focus on the hands only.

The hands are used in sign language to code all letters of the alphabet (Figure 1). This is the primary means of communication for the deaf. We are focusing on this type of sign language, aiming at recognizing all letters to enable the constructions of words and sentences. Furthermore, we generalize the concept of a letter to represent a word or sentence as well: the signs can be mapped to a letter, a word or even a sentence.

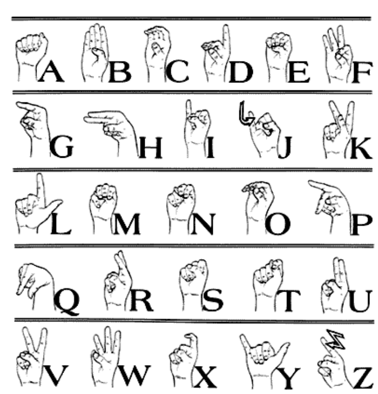


Figure 1. The Sign Alphabet

There have been previous attempts in solving the problem of sign language translation with a glove. A team of Ukrainian students (in the Software Design Competition of the 2012 Microsoft Imagine Cup) developed a glove to turn signs into text [3]. Their system relies on many sensors, including sensors between the fingers allowing very precise gesture recognition. It runs on an Android device, at relatively low speed, and the gestures are translated into text.

Similarly, we have built a custom glove with fewer sensors, but with an accelerometer and gyroscope, and greater flexibility since we can reprogram and run pre-processing computations on the glove itself. Also, the accelerometer and gyroscope helps us detect hand orientation, therefore enabling recognition of rotated hand signs.

Calibration of the glove is an essential initial process to ensure precise gesture recognition. It can be a time consuming operation, particularly for deaf users who need clarifications on about how they should be positioning or moving their hands.

Another glove system aiming at accurate calibration [4] uses a Cyberglove 22 flex sensor for gesture recognition. The system is built on 22 predefined gestures and prompts the user to learn how to operate the glove using a simulator.

Our system is different from both prior attempts in two respects. First, in both prior systems the set of gestures is predefined (after some initial calibration). Our goal is to provide a generic system that enables the user to define any other gesture or sign language. Furthermore, each user's set of signs is stored and enabled after authentication.

Second, the calibration process for the Cyberglove is rather cumbersome and requires the user to repeat 44 gestures. Our system on the other hand has very simple calibration process from a user perspective. The user only needs to move his hand in all directions, and bend and unbend the fingers.

Note that there are a number of systems for sign language translation that have been developed for the kinect [1,2]. The kinect offers very accurate sensors. Systems using the kinect require to be set up in permanent areas such as classrooms, while our vision is to build a wearable device that can easily be carried around. Therefore we focus on hands only, and we rely on fewer sensors provided with a custom made glove.

THE JHANEGLOVE

The human hand has 27 degrees of freedom: four in each finger, three for extension and flexion and one for abduction and adduction. The thumb is more complex: it has five degrees of freedom, and six degrees of freedom for wrist rotation and translation.

The JhaneGlove (figure 2) was built using three types of sensors: 5 flex sensors, 8 contact pads sensor, and an accelerometer and gyroscope. Each flex sensor has three states. Each sensor can be turned on or off. It enables 20 degrees of freedom, allowing more than a hundred different gestures. The flex sensor is a device that changes its resistance proportional to its form (figure 3), therefore allowing the detection of a bending movement when placed on the finger. The contact pads allow identifying the contact between fingers and the hand palm, such as the "Y" sign.

Finally the gyroscope allows identifying hand motion and thereby the recognition of gestures based on hand orientation.

This last feature is unique and makes our system more flexible in terms of recognition, as opposed to others. We can track hand movements and rotation as part of the gesture. For example the A sign can be trained and recognized in different orientations.

THE SIGN LANGUAGE

What makes our glove different from other systems is its ability to easily create custom sign languages. Most systems focus on implementing the conventional sign language. Not only we can support this language, but we can also train and store one different language per user.

As a trade-off between the recognition precision and time needed to train the system, our system currently allows the definition of 30 signs that can be processed by the system's recognizer. In addition to these signs, we have the *space* sign that enables marking the end of a word, and the *dot* sign that marks the end of a sentence. These are necessary for the recognizer to form the words. Once the words are captured, the text-to-speech agent can convert them into speech.

If the system is used to enter the 26 letters of the English alphabet, then four more signs remain that can be assigned to frequent occurring words or even full sentences. If the system is used to define a domain-specific language, then one can assign full words or sentences to the 30 signs and create a custom sign language.

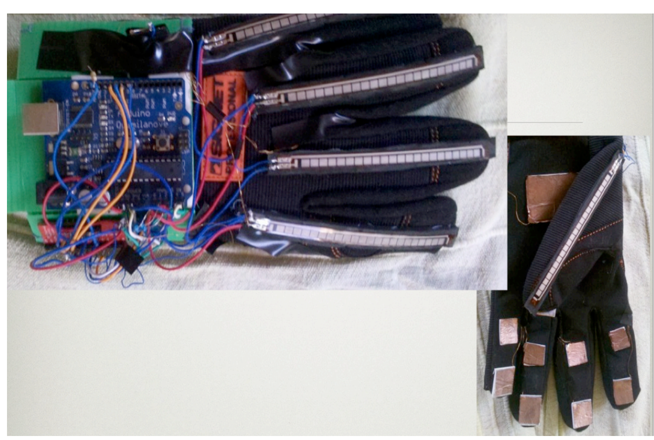


Figure 2. The JhaneGLove

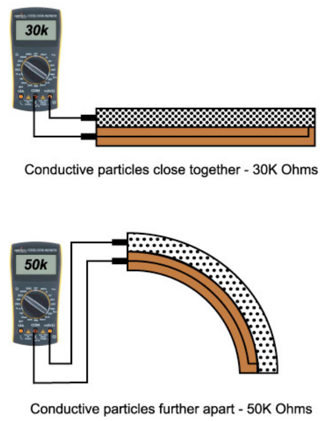


Figure 3. Flex sensors

CALIBRATION AND TRAINING

The glove is a device that continuously sends data from its sensor. In order to pre-process the data to be sent to our main agent (a server side component), we have used an electronic board, the Arduino, to develop the pre-processing software. The board is attached directly to the glove.

Arduino is an open-source electronics prototyping platform based on flexible, easy-to-use hardware and software. It can be used in various ways to create interactive systems. The Arduino board is shown in Figure 2. It can be used with a variety of sensors to control for example lights or motors. The microcontroller on the board can be programmed using a custom programming language. This enabled us to implement pre-processing computations on the glove itself. Arduino projects can be stand-alone or they can communicate with software running on a computer (e.g. Flash, Processing). We have used this ability to develop a computer client that handles communication between the glove and the recognizer. It also contains everything needed to support the microcontroller; we connect it to the computer with a USB cable or power it with a AC-to-DC adapter or battery to get started.

Calibration

In order to use the system, the user must first calibrate the glove. The first stage is to record the min/max values that the user can emit with the glove. This process requires the user to bend and unbend all fingers, as well as move the hand in all directions. The calibration is very short and non tedious for the user. The min/max values are stored on the Arduino board. The pre-processing software embedded in the Arduino board also performs normalization (0-128) on the values that range between (-200,200) for flex sensors and (4000-8000) for the gyroscope. From that point, all data emitted from the glove is sent to our server in a normalized form and can be processed by the gesture recognition agent.

From a user point of view, the software provides a simple user interface. In order to calibrate the glove, the user presses the "Start calibration" button. At that moment, the system stores the timestamp of startup. The glove keeps the calibration signal and starts to send raw data to the server until the "Stop calibration" event generated by the stop button. On generation of the "Stop calibration" event, the system stores the timestamp of shutdown, loading all the raw data from the database which was received from the user during the start-stop time range. Then it computes the minimum and maximum value per sensor. These values are then sent back to the glove.

When the glove receives the last bit of calibration data, it send a "Ready" message to the server and continues to send the calibrated sensor's data to the server by using the new set of minimum/maximum values.

Training

After calibration, the user can train the system to recognize his signs. The training process is relatively intuitive, with a simple user interface where each sign must be entered four times (Figure 4). The user controls the training progress: he can take a break when necessary and can complete the training at a later time.

The problem we are solving a multiple class classification problem. The input dimension corresponds to the sensors input: 20 dimensions with different input ranges for each sensor, multiplied by the number of continuous values recorded for each gesture. The more the user gesture is precise, the better the recognition engine will perform. In addition, the more gestures there are, the more there is a probability of similarity (close distance between gestures in our 20 dimension space). This is why the recognition agent can function with more than 30 signs in practice, but it actually performs more accurately with fewer signs (see experiments below).

Different users use the glove: we implemented an authentication system that guarantees a personal set for each user. After successful authentication, the system loads the user specific training set, and the user can start using the glove with the signs he previously entered.

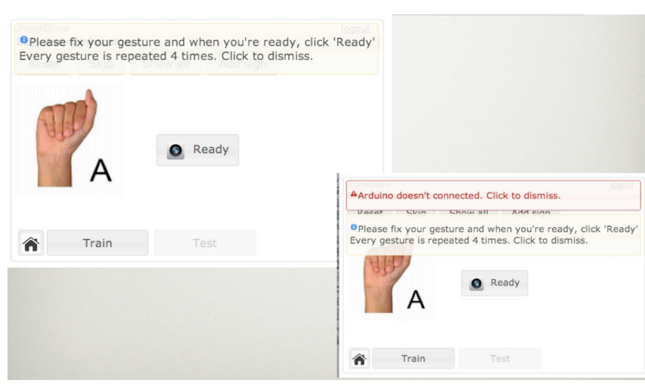


Figure 4. Training

THE SOFTWARE ARCHITECTURE

Our system is built in a client-server architecture where the server side component includes the gesture recognition agent, the database that collects the training and calibration data, and each user's personal language data (the neural network serialization).

The glove with the Arduino is the "wearable" device. On the Arduino board we have the processing software that stores calibration data. The Arduino is connected to a local computer (the client) that transfers the glove data to the server side. The whole setup at this stage is not as mobile as we would hope that the operational system would be. Nevertheless it is sufficient to allow us to experiment and test the feasibility and usefulness of such system. Note that, in this project, we are facing both computing, algorithmic, as usability challenges.

The recognition agent produces text. Then we convert the text into voice using the third-party available text-to-speech software.

THE GESTURE RECOGNITION AGENT

The gesture recognition agent is implemented as a back propagation neural network [5]. The back propagation algorithm is a common learning algorithm. It is used for training multilayer networks by means of error propagation. It aims at minimizing the sum of squared approximation errors by adjusting the network parameters. The neural network uses data that was stored in the database during the training process. This data is user specific so each user benefits from a different customized recognizer. The system is initialized with a predefined dataset including the alphabet, allowing unregistered users to be trained on this set.

The neural network is composed of an input linear layer of size 20, a middle sigmoid layer of size 13, and an output layer of size 1. Data is continuously fed into the neural network and the recognition of a sign is validated after five consecutive recognitions of the same sign. A sign generally last for less than half a second up to one second: slow as well as fast signs can be recognized.

EXPERIMENTS

The recognition accuracy of the device was tested under various conditions: users were asked to enter a full sentence using the alphabet, then they were asked to add a new sign to the dataset and use this sign within a sentence. The results are summarized in Table 1: the percentage represents the amount of correctly recognized symbols. It clearly shows a very good percentage for 15 signs but the recognition accuracy drops and it is too low for a full set of 30 signs. This can be explained both from a software and hardware aspect. The handmade glove requires a lot more sensors to provide more accurate data, and the neural network configuration could be improved.

Natural sign language is certainly much faster without the glove and also much faster than typing. We measured the time required for the user to make the sign and for its recognition. We found that a user sign lasts for about half a second (that includes collecting data and storing it in the database). The recognition is achieved in about half a second (that includes retrieving the data from the database and activating recognition engine). So about a second for a sharp gesture or more for a shaky gesture, is required for each sign. That gives us a maximum rate of 60 signs per minute.

The transition to from one gesture/letter to another works well, but we have so far been unable to identify the repetition of a gesture, for example where a letter appears twice in a word. Recall that we are not using specific signs to signal the end of a sign, although we could. Many systems do use a start/end signal sign to facilitate the parsing. Adding such feature makes the user experience a lot slower (twice as many signs for the same word) and tedious. This is why we did not implement such functionality.

# of signs	**accuracy**
15	92%
20	88%
30	80%

Table 1: Recognition accuracy

CONCLUSION

We have described early results of a homemade glove that allows defining and recognizing the conventional and also custom sign languages. The core system functions were completed and tested. The system is open for extensions both on the hardware and software sides. As a first experiment, we identified several aspects that could benefits from significant improvements. First we could use more sensors, better quality sensors. Two of five flex sensors have a very small range and the gyroscope sometimes shows inconsistent values. We could also add flex sensors between fingers to enhance the precision of gestures. Second we could rely on a wireless Bluetooth connection between the glove and the computer to free completely the user from the physical computer. On the long term, we would aim for a smartphone application where the system would really become wearable. The smartphone would contain the application front-end (user interface of the system), while the server side would perform the computations of the recognition agent and hold the database.

Our current focus is on improving the gesture recognition agent since we still have not reached a satisfying recognition of the full set of signs. We will investigate different neural network configurations with less input in the middle layer for instance, while keeping an acceptable training time. After we will have improved the recognition agent, we plan on performing user tests to evaluate the system from a user perspective. For example we are not sure about using separator signs for experienced users. A simple long pause might be enough instead of a *space* sign. Finally we are exploring other directions such as an auto-complete ability to enhance the user experience and shorten the numbers of signs required to express a word or sentence.

REFERENCES

1. Zahoor Zafrulla, Helene Brashear, Thad Starner, Harley Hamilton, and Peter Presti. 2011. American sign language recognition with the kinect. In Proceedings of the 13th international conference on multimodal interfaces(ICMI '11). ACM, New York, NY, USA, 279-286.
2. EnableTalk http://enabletalk.com
3. Simon Lang, Marco Block, and Raúl Rojas. 2012. Sign language recognition using kinect. In Proceedings of the 11th international conference on Artificial Intelligence and Soft Computing - Volume Part I (ICAISC'12)
4. Matt Huenerfauth and Pengfei Lu, Accurate and Accessible Motion-Capture Glove Calibration for Sign Language Data Collection, ACM Transactions on Accessible Computing, 3, 1, Article 2 (2010).
5. P. Sibi, S.Allwyn Jones and P.Siddarth, Analysis of Different Activation Functions Using Back Propagation Neural Networks, *Journal of Theoretical and Applied Information Technology*, (2013). Vol. 47 No.3

SAsSy - Making Decisions Transparent with Argumentation and Natural Language Generation

Nava Tintarev
University of Aberdeen
Aberdeen
n.tintarev@abdn.ac.uk

Roman Kutlak
University of Aberdeen
Aberdeen
r.kutlak@abdn.ac.uk

ABSTRACT

An autonomous system consists of one or more physical or virtual agents that can perform tasks without continuous human guidance. In order to realise their promise, techniques for making such autonomous systems scrutable and transparent are therefore required. To address this issue the Scrutable Autonomous Systems (SAsSy) demo shows how argumentation and natural language can be combined to generate a human understandable dialog explaining the operation of an autonomous system. On the one hand argumentation theory is used to simulate human understandable reasoning mechanisms. On the other, natural language generation tools are used to translate logical statements into simple plain English. The idea is to generate a dialog that enables the user to understand and question the reasoning present in autonomous systems.

Author Keywords

Explanations, Argumentation, Natural Language, Agents

ACM Classification Keywords

H.5.2 User Interfaces: Natural Language, Interaction styles, Graphical user interfaces (GUI)

BACKGROUND AND APPLICATION CONTEXT

An *autonomous system* consists of physical or virtual entities, or *agents*, that can perform tasks without continuous human guidance. Autonomous systems are becoming increasingly ubiquitous, ranging from unmanned vehicles, to robotic surgery devices. Such systems can potentially replace humans in tasks which can be dangerous (such as refuelling a nuclear reactor), mundane (such as crop picking), or require superhuman precision (as in robotic surgery). While increasing reasoning capacity can enable an autonomous system to handle a wider range of situations, modelling and verifying the operation of such systems becomes increasingly difficult.

The increasing amount of independent reasoning that takes place within an autonomous systems means that humans struggle to establish why a system chose to behave as it did, to identify what alternative actions the system considered, and to determine why these alternatives were not selected for execution by the system. In other words, such systems are *opaque*. Such opacity is exacerbated by the formal models typically used to drive the reasoning behaviour in such systems — a human (and particularly a non-expert) often struggles to understand what is going on in the system. This lack of understanding can lead to unrealistic expectations of an autonomous system, or alternatively to a lack of trust in it, causing inefficiencies at best, and leading to dangerous outcomes in the worst cases. When things go smoothly transparency may not be that important. However, it is equally vital that a user can identify undesired actions before they are carried out, and interfere appropriately if need be: even if the user understands the system they need to be able to cancel actions or suggest alternatives with relative ease in a timely manner.

The SAsSy project[1] has for the last year been investigating computational mechanisms for providing transparency to humans regarding the internal workings of an autonomous system. We use formal argumentation in combination with natural language to offer explanations to a human. The system explains which sequence or actions, or what plan, have been chosen for execution by the system. More specifically it explains *why* a certain plan has been selected. That is, the user should be able to follow a chain of reasoning with arguments and counter arguments. The system also allows users to provide additional information which can be used to modify the arguments, and subsequently, the plan.

The SAsSy team is working with industrial partners in the hydrocarbon exploration and unmanned vehicle domains to identify users' explanatory needs in these domains. However, the system architecture is flexible enough to model new domains and in this demo we present a simplified scenario built around delivery logistics. The components that are domain specific are: a workflow, domain rules, and a lexicon.

The scenario is based on a delivery driver in Scotland who is delivering a package from Edinburgh to Inverness and driving back. In this case, the plan is a choice of route with a sequence of driving actions to a number of intermediate locations. Each location serves as a potential choice point from which several other locations may be possible. Some routes yield shorter distances between points, but given other factors such as traffic and road conditions the shortest route is not always the best option. The driver executes the plan step by

[1] `http://scrutable-systems.org`

step, and questions the system as required while they execute the plan.

SYSTEM DESCRIPTION

Explanations have frequently been a component of intelligent systems (IS) such as expert systems [1, 4, 8] and recommender systems [5, 7]. The explanation capabilities in expert systems have often been evaluated with users in terms of whether they increase acceptance of an intelligent system or acceptance of decisions. However, there are other reasons why explanations may be introduced to an IS including transparency and scrutability [7] – helping users understand how decisions were made (transparency), as well as allowing users to tell the system that it is wrong (scrutability).

The demo shows the interaction between two core technologies: human understandable reasoning mechanism (represented through argumentation theory) and natural language generation tools to translate logical statements into plain (natural) English. Our system is developed in Python and is available under the BSD licence[2].

According to the recent classification of explanations in IS by [4], the explanations in our system can be classified as justification, and to certain extent, trace-based explanations. Justification type explanations describe 'why' a decision was taken and supply the descriptive knowledge used to reach that decision. Trace explanations supply information about 'how' the decisions are made, e.g. which rules were applied. In our system they are exposed to the user thanks to argumentation. Our system does not present the user with the full reasoning trace, but rather allows the user to discuss or argue with our system about each subsequent fact that is discussed. Our system also allows alteration to the rule base. Since all knowledge in our system is captured through rules, the information can be updated by the user through adding rules, removing them, or adding exceptions to existing rules. This in turn affects subsequent decisions, since the rules in turn affect the arguments and their outcomes.

Reasoning mechanism. Our system uses two kinds of rules: *defeasible* and *strict*. *Strict rules* capture traditional implication – whenever the literal or literals on the left-hand side are true, so is the literal on the right-hand side. An excerpt of the knowledge base can be seen in Figure 1. The main difference between these two kinds of rules are that defeasible rules can be affected by exceptions, while strict rules cannot.

An example of a strict rule might be: `flood_road --> closed_road`, which might be read as "*if a road is flooded, it is closed.*"

Defeasible rules capture defeasible implication – whenever the literal(s) on the left-hand side are true the right-hand side is **usually** also true. An example of such rule might be: `snow_road ==> closed_road`, which might be read as "*if a road is covered by snow, it is usually closed.*"

The previous defeasible rule can be changed to the following rule: `snow_road = (-plough_road) => closed_road`, which might be read as "*if a road is covered by snow, it is usually closed unless it has been ploughed.*"

Rules without pre-conditions on the left-hand side are used to represent asserted or assumed information. For example, `==> accident_on_bridge` means that "*there was an accident on the bridge.*" We use the terms *rule* and *information* interchangeably.

Part of the knowledge base represents legal actions in terms of locations that can be reached from each point – we illustrate this in Figure 2. The reasoning mechanism in SAsSy is based on argumentation, where the arguments are derived from rules in the knowledge base. Arguments are stored as a directed graph, in which the nodes constitute arguments and arcs between nodes symbolise attacks between arguments. The advantage of using argumentation as opposed to traditional rule-based reasoning, is that it can reason even in the face of contradicting statements. The use of argumentation also enables two important features: non-monotonicity and

```
forecast_high_wind =(-forecast_old)=> edinburgh_bridge_closed
edinburgh_bridge_closed --> edinburgh_perth_not_possible
edinburgh_bridge_closed --> perth_edinburgh_not_possible

forecast_high_snow ==> inverness_perth_not_possible
forecast_high_snow ==> perth_inverness_not_possible

accident ==> traffic_slow
accident_on_bridge ==> traffic_very_slow

traffic_very_slow --> traffic_slow

R1: ==> -stirling_shorter
R2: ==> -kincardine_shorter
R3: traffic_slow =(-dont_care_slow)=> kincardine_shorter
R4: traffic_very_slow =(-dont_care_slow)=> stirling_shorter

R1, R2 < R3 < R4

#
--> kincardie_bridge_10
A1: ==> vehicle_weight_15
```

Figure 1. Excerpt of the knowledge base used in our system.

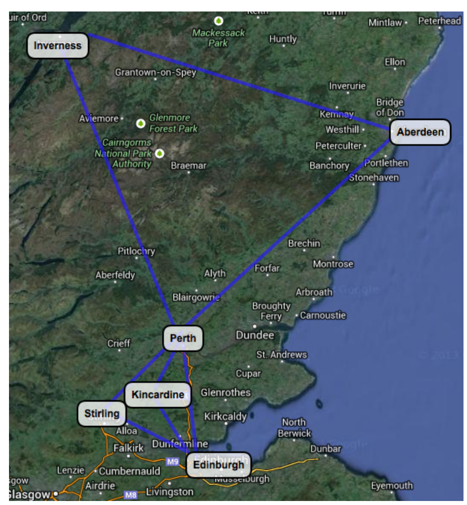


Figure 2. This map is included as an illustration of the different geographic locations used in the example. In our system, routes are represented in the knowledge base as rules of what kinds of actions are allowed. Locations on the map are slightly shifted to increase legibility.

[2] **https://bitbucket.org/rkutlak/sassy**

alternative options. Non-monotonicity means that (when the user) adds new information, it may invalidate past conclusions and justify new ones. Alternative conclusions are represented in the argumentation graph which makes them available for us to both explain what alternatives were considered by our system as well as justify why they were not chosen.

The explanations build on the instantiated grounded game presented in [3]. Our system allows users to step through the argument graph in a sequences of discussion moves presenting arguments and counter arguments. Intuitively, such discussion can be seen as a debate, or reasoning for and against a course of action. The psychology of human reasoning as validation for argumentation semantics is a largely unexplored area [6], but a recent strand of work takes its inspiration from human dialogue to find intelligible explanations of an argument's status (i.e., whether to accept or reject the conclusion of an argument) [2].

The Natural Language Generation. *Natural Language Generation* (NLG) is the study of computer algorithms which produce understandable and appropriate texts in English or other human languages, from some underlying non-linguistic representation of information. In our case, the non-linguistic information are the rules capturing the "knowledge" in our system.

As we have already seen, the rules are formed from literals (e.g., `snow_road`), which can be ambigious or difficult to understand. We use NLG techniques to convert the literals to more natural text as well as to improve the presentation by removing unnecessary information. For example, since defeasible rules capture implications that are usually true, we do not present the exceptions to such rules to the user.

Presenting the available information in concise and unambiguous language could be crucial in e.g., the unmanned vehicle domain, where an operator has limited time to comprehend our system's reasoning and potentially change the course of action. We plan to include a summary of the presented plan as well as to use other NLG techniques such as aggregation (combining simple sentences together for better presentation) and referring expression generation (e.g., using pronouns when referring to past entities) to improve the presentation of the information. Indeed, in order to effectively be able to communicate 'why' certain decisions are preferable, the user needs to be able to understand *what* the recommended plan is first.

DEMO

Our demonstrator shows how argumentation can be used to emulated dialog, and how natural language can be used to present reasoning rules in a way that is familiar to users. The possible actions are represented as a workflow which is visualized as a graph, which is accompanied by natural language descriptions. In addition, the user can contribute to the dialog by asking questions. Figure 3 demonstrates a screenshot of the SAsSy demo.

The system allows the user to ask three different types of questions:

1. *"What can I do next?"*: The left hand pane suggests a next possible location to drive to, e.g. 'Next task: Go to Inverness'. The workflow at the top shows the alternative options, e.g. Stirling1, Kincardine1 and Perth1.

2. *"Why does the system NOT say that I should do Y?"*: likewise, the user can ask why a certain options is rejected: "why out Perth1?". To form a reply the system derives the relevant rules and translates them into natural language in a form such as "Going through Stirling is faster because the traffic is very slow.".

3. *"Why does the system say that a certain thing is true?"*: The user can type a question such as "why traffic_very_slow" in the dialog field and receive an answer such as "The traffic is very slow because of an accident on the bridge.".

For the time being the user input is limited to controlled symbolic constants such as 'Perth1' or 'traffic_very_slow'. Later versions of the system will make explicit the vocabulary available to a user regardless of whether it is controlled or processed from natural language.

The user may choose to question the system several times. For example, when the system tells the user that they cannot take a certain sub-route the user can query what knowledge was used to derive this argument, such as the case with traffic_very_slow in the screenshot and example above.

A user can choose to manually override or alter the knowledge base. A manual override means they simply proceed as they wish, for example by typing "next Perth1". Conceptually, altering the knowledge base is important if a similar process is likely to be repeated – so that users do not have to manually override repeatedly. For example, the situation may have changed and the user may want to tell the system that the accident is no longer causing issues:
Driver> why traffic_very_slow
SYSTEM> The traffic is very slow because an accident on the bridge.
Driver> retract accident_on_bridge

The system deletes the information of an accident on a bridge, which affects the argument against using the bridge. The system now no longer has a reason to recommend the route though Stirling, and returns to the default of recommending going straight to Perth.The system also represents a preference ordering among rules in case of conflicts to mitigate deadlocks, or situations where the system is undecided. So, if two arguments attack each other, only the attack of the argument with the more preferred rule holds true.

Apart from out of date information, the user and the system may have different information. For example, the system may be telling the user that they cannot go through Kincardine because of the weight of their van, but has gotten the weight wrong:
Driver> why out Kincardine1
SYSTEM> you cannot go from Edinburgh to Kincardine because the maximum allowed weight on the bridge is 10 tons and the weight of the vehicle is 15 tons.

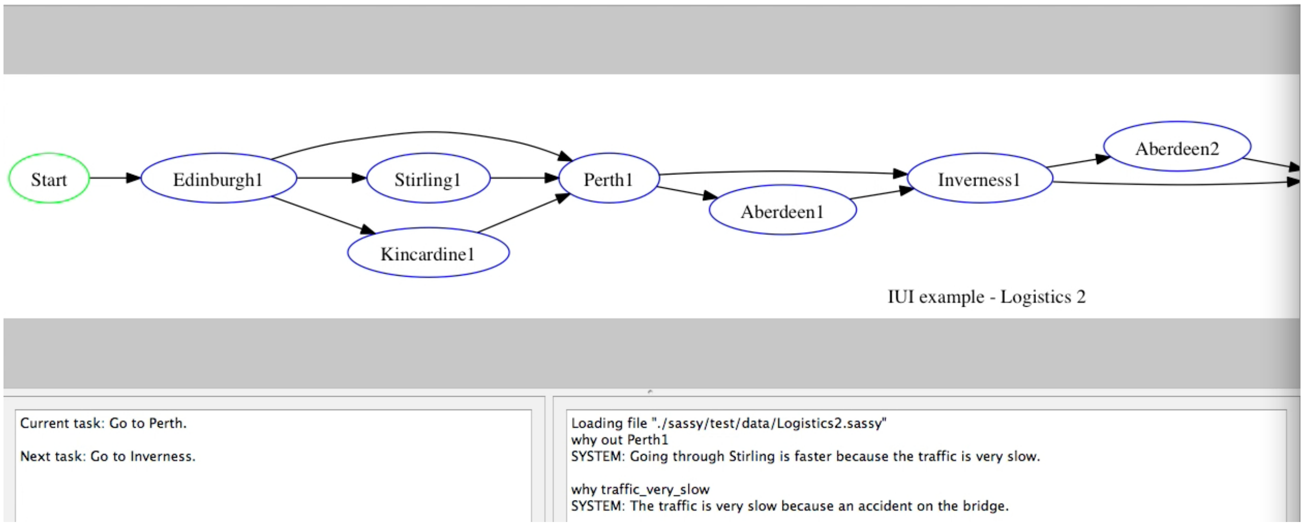


Figure 3. Screenshot of the SAsSy system. *Top:* Workflow with all the possible options, represented as a graph. The edges are actions such as 'drive from a to b', and the edges represent states equivalent to reaching a certain location. Perth1 is the state reached in the outbound journey leaving from Aberdeen, whereas Perth2 is that state of the return journey from Inverness. *Bottom left:* Status of system with current and next steps. Note that some steps could be triggered by the system, in which case the change of status and action taken are also represented here. *Bottom right:* Dialog between user and system. This is where the user asks questions and the system responds. The user asks why they cannot go through Perth, and the user supplies a justification (traffic is slow). However, the user may not be willing to accept this justification. They query why traffic is slow and receive an answer for this as well (there was an accident on the bridge).

Driver> retract vehicle_weight_15
SYSTEM> deleted

In addition, users can introduce new information to the system such as a new weight for the vehicle:
Driver>assert vehicle_weight_5

The explanation capability in SAsSy is undergoing continued development, with another 2 years remaining on the project. Development is guided by ongoing user-centered experimentation. We are currently planning experiments on presenting information differently depending on the areas of responsibility of a user. This would allow the system to tailor the information presentation to, for example, someone who is responsible for deliveries at an airport versus someone who in ensuring the delivery of a specific package.

ACKNOWLEDGMENTS

Many thanks to our esteemed colleagues for their valuable comments. This research has been carried out within the project "Scrutable Autonomous Systems" (SAsSY), funded by the Engineering and Physical Sciences Research Council (EPSRC, UK), grant ref. EP/J012084/1.

REFERENCES

1. Buchanan, B. G., and Shortliffe, E. H., Eds. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison Wesley, 1984.
2. Caminada, M., and Podlaszewski, M. Grounded semantics as persuasion dialogue. In *COMMA*, B. Verheij, S. Szeider, and S. Woltran, Eds., vol. 245 of *Frontiers in Artificial Intelligence and Applications*, IOS Press (2012), 478–485.
3. Caminada, M., Podlaszewski, M., and Green, M. Explaining the outcome of knowledge-based systems; a discussion-based approach. In *AISB* (2013).
4. Darlington, K. Aspects of intelligent systems explanation. *Universal Journal of Control and Automation 1* (2013), 40–51.
5. McSherry, D. Explanation in recommender systems. *Artificial Intelligence Review 24(2)* (2005), 179 – 197.
6. Rahwan, I., Madakkatel, M. I., Bonnefon, J.-F., Awan, R. N., and Abdallah, S. Behavioural experiments for assessing the abstract argumentation semantics for reinstatement. In *Cognitive Science* (2010).
7. Tintarev, N., and Masthoff, J. Evaluating the effectiveness of explanations for recommender systems: Methodological issues and empirical studies on the impact of personalization. *User Modeling and User-Adapted Interaction 22* (2012), 399–439.
8. Ye, L., Johnson, P., Ye, L. R., and Johnson, P. E. The impact of explanation facilities on user acceptance of expert systems advice. *MIS Quarterly 19*, 2 (1995), 157–172.

AvatAR: Tangible interaction and augmented reality in character animation

Adso Fernández-Baena
Seamless Interaction, GTM
La Salle RD, La Salle - URL
08022 Barcelona, Spain
adso@salleurl.edu

David Miralles
Seamless Interaction, GTM
La Salle RD, La Salle - URL
08022 Barcelona, Spain
davidme@salleurl.edu

ABSTRACT

In this paper, we present a novel interaction system, which combines tangible interaction and augmented reality for controlling a virtual avatar. By physically interacting with a cube, it is possible to drive avatars motion that occurs in the real world. The cube acts as a motion controller and as an AR marker reaching input and rendering purposes. The cube facilitates users the avatar positioning and motion customization, providing a fine control for both. In this first version, the avatar is able to stand and move. The current motion state is picked rotating the cube over the same plane where the avatar lies. We have implemented two scenarios in our prototype: a sketch-based controller and an interactive controller. The first one enables users to draw paths on the floor that the avatars follow; on the contrary, the second allows drive avatars position during all the time. The idea of using tangible objects in augmented reality environments for controlling avatars strengthens the link between the user and the avatar providing a better sense of control and immersion.

Author Keywords

tangible interaction; augmented reality; character animation

ACM Classification Keywords

H.5.2. Information Interfaces and Presentation (e.g. HCI): Input devices and strategies

General Terms

Human Factors; Design.

INTRODUCTION

User interfaces have been evolved a great deal from command-line interfaces, through graphical interfaces to last generation. This last generation brings together tangible and natural interfaces. Tangible user interfaces (TUI) [3] permit to interact with digital content through a physical environment. TUI allows to design interactions using everyday objects, breaking the traditional approach of interacting with computers and giving a strong sense of immediacy to the user. These objects which are involved in TUI applications are regarded as kind of smart objects.

In computer animation, intuitively controlling the motion of a virtual character is considered as a difficult task. One reason for this is that a virtual character usually used in the field has a high degree-of-freedom (DOF) for controlling the position and orientation of all body joints, thus difficulting the design of an intuitive interface to manage the individual body joints. One intuitive approach to solve it is using an instrumented puppet in order to retarget puppet pose to a virtual character [5]. Along the same lines, Oshita et al. [6] propose to use hand manipulation based on traditional puppet mechanism to control a character. Both approaches obtain a virtual character moving, but their move is restricted by the input. Even though the efficiency of motion retargeting is high, it makes the generation of real motion very difficult for the user. So, to obtain more natural complex motions (which involves a coordinated movement of different body parts) it is necessary to add an intermediate layer to select proper motion capture data. This is what Lockwood et al. [4] done. They use a touch-sensitive tabletop for generating full-body animations, where two fingers are used to pantomime leg movements. In this work, we would go a step further and control motion of a virtual character using objects (see Figure 1), but even providing a higher layer to reproduce plausible whole body motions like locomotions.

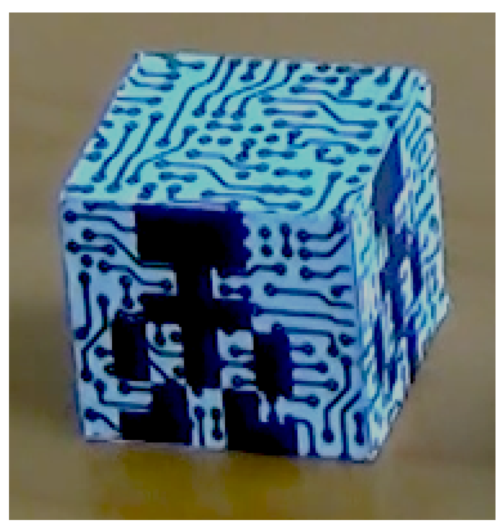


Figure 1. Cube for controlling avatAR.

IUI 2014, Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel.

Turning to another issue, augmented reality is a growing area of interest because it provides an enhancement of physical information, which is presented jointly facilitating the user understand correspondences. Moreover, interaction in augmented reality environments is lifelong investigating in character animation although some advances have been achieved in other topics like Internet of Things (IoT). Recently, Heun et al. [2] propose smarter objects, which among other things, it consists on associate a virtual object with physical object to support an easy means of modifying the behavior of that physical object. It uses augmented reality to match physical and virtual objects and supports both tangible and graphic interaction (through a tablet). Furthermore, Amores et al. [1] present SmartAvatars. SmartAvatars is a concept that also relies in a augmented reality environment and uses virtual characters to display system feedback. They propose a prototype called Flexo where virtual characters interact with objects but their behavior is constraint for the system reaction. However, they do not have neither tangible and direct control on avatars motions, which is precisely what we do in the interaction system described below.

Thus, we propose a tangible interface for controlling avatars in a augmented reality environment. We use a physical cube for driving virtual character motion, which in turn, it is used as augmented reality target. In this manner, the cube is mixed in the virtual scene providing more awareness in spatial perception, enhancing the character animation controllability of the user. To our best knowledge, there is no other character controller than mixes tangible interaction and augmented reality.

INTERACTION SYSTEM

AvatAR interaction system (see Fig. 2) is formed by a cube, a tablet and a plane surface. The tablet is placed on a stand in landscape position, looking in at the surface and in turn defining the interaction area. So, the interaction area is located behind the screen. The cube is our character controller and at the same time our AR marker. So, a virtual character appears inside the cube when such cube is detected. In this manner, augmented reality becomes tangible. From this moment the virtual character is mixed with the reality viewed by the camera of the tablet.

Our character is able to move on a horizontal plane that in this prototype coincides with the tablet orientation (due its pose). For this purpose, we only consider idle motion and locomotion. We indicate through the cube the position where the avatar has to move on. Indicating positions is straigthforward due the cube lies in the scene jointly with the avatar. The local orientation of the cube defines the type of motion providing a range that covers from walking, through running to sprinting. Stablishing the analogy that likens the cube movement to a potentiometer. If we rotate the cube by 360 degrees, the motions displayed are: idle, walking, running, walking, idle. So, the avatar step displacement varies jointly with the type of locomotion.

Up to this point, we can get locations and a variety of motion

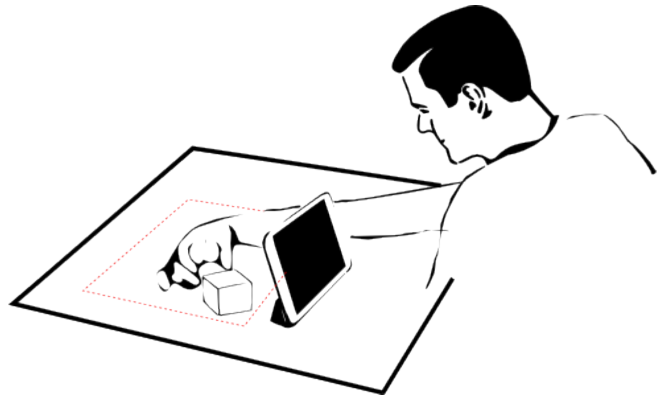


Figure 2. Interaction system

of our avatar by moving and rotating the cube. Nevertheless it has to be decided how to use locations in order to design an easy-to-use character controller. To get started, we design two scenarios.

Scenario 1: Sketch control

This scenario allows to define paths where the virtual character has to follow. So, we propose two steps. First, the path is defined by the translation of the cube across the scene. In order to guide the user in the path creation, we draw intermediate points (see Figure 3). Once we have the path defined (now we have restricted to eight points), the virtual character traces the path in loop mode. While it is moving, the user can control its behavior by rotating the cube. In this manner, it is possible to observe how the avatar can covered the same path performing different motions. Moreover, the user can interrupt avatars movement by assigning idle motion which leads the avatar to remain in a static position.

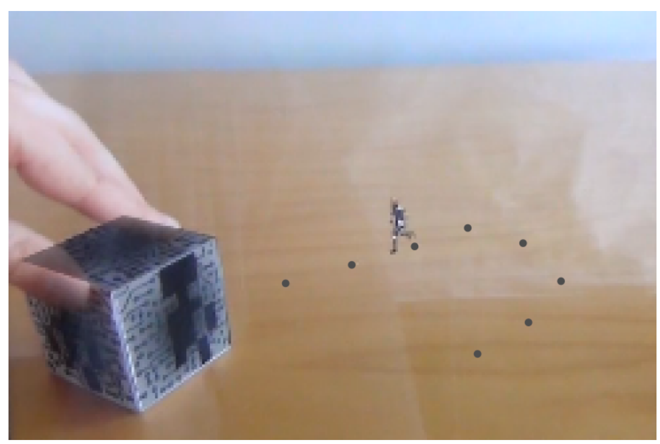


Figure 3. The avatar walk along the path.

Scenario 2: Interactive control

One of the most common uses of virtual characters is in videogames. In this context, virtual characters are driven by players allowing positioning and behavior performing at all times. In the same way, we implement this scenario.

In this case, the virtual character have always defined its target position, which is the cube position. So, the virtual character follows the cube. How longs it takes to reach the cube depends on the type of locomotion. Obviously, if we set running locomotion, the character achieve faster the cubes, in case of setting walking locomotion, the character comes later.

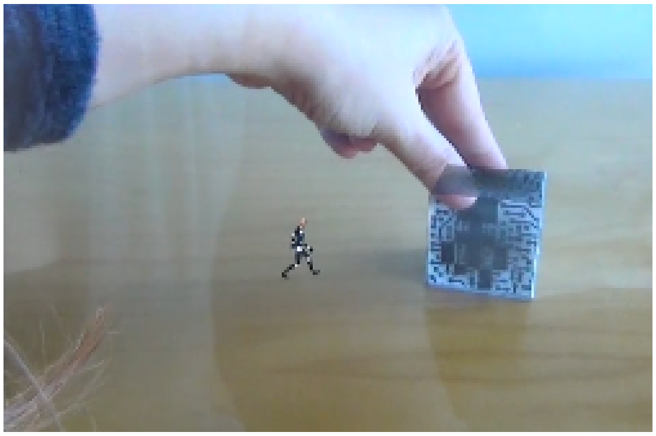


Figure 4. The avatar is walking to the cube.

IMPLEMENTATION

In order to make our first prototype we use Unity Game Engine [8]. Unity is a game development ecosystem which includes Mecanim [9], a powerful and flexible animation system which we use for create our animation controller. For augmented reality purposes, we use the software platform Vuforia [7] developed by Qualcomm. The following is a more detailed explanation of the animation controller, the motion control interface and the cube target we use.

Animation controller

Mecanim enable easily contruct and edit complex state machines and blend tree for complete control of how the virtual characters move. We want that our virtual character can perform different motion clips which are idle, walking and running. For this purpose, we have constructed a locomotion controller composed by two states: Idle and WalkRun. Idle state is an idle loop motion, and WalkRun is a blend tree which we will deeply explain later. To decide the current motion state we use "Speed" motion control parameter. Speed parameter balance between Idle and WalkRun states, assuming that if speed is greater than 0.5 the avatar is moving, and so, walkRun must to be the current state. On the contrary, the avatar is in Idle state.

As we have mentioned, WalkRun state (see Figure 5) is a blend tree which in turn is formed by Walks and Run blend trees. Blend trees provide variations of motion clips by blending similar phases of the input motions, so, motions have to be previously aligned. For both Walk and Run, we implement a classic blend tree which is composed by turn left, straight and turn right motions. The "Angular" motion control parameter for the turn key goes between -90 and +90, controlling which animation is being played. In case of walk blend tree, it is formed by short, medium and wide strides; and for run blend tree medium and wide. Moreover, "Speed" motion control parameter is also used to discriminate between Walk and Run blend trees, where its value goes from 0.5 to 5, remaining Walk state until 3.

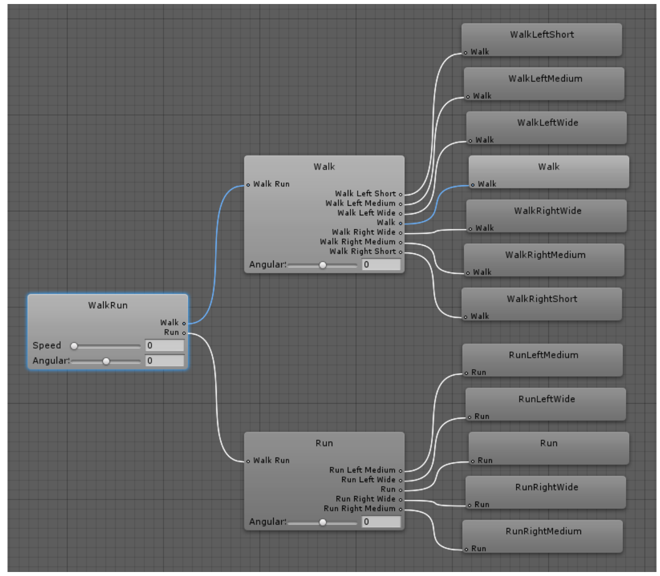


Figure 5. WalkRun state blend tree

Cube target

We design a cube with a set of image targets in its faces. Each face is illustrated by an icon that denotes one of the motion clips. We use 3 icons: idle, walking and running. The idea is to drive the virtual character by the cube faces. How to do this is addressed below. Moreover, we have texture the background with a chip pattern. In this manner, we improve the proper detection of the marker. In Figure 6 it is illustrated the graphic design of the marker.

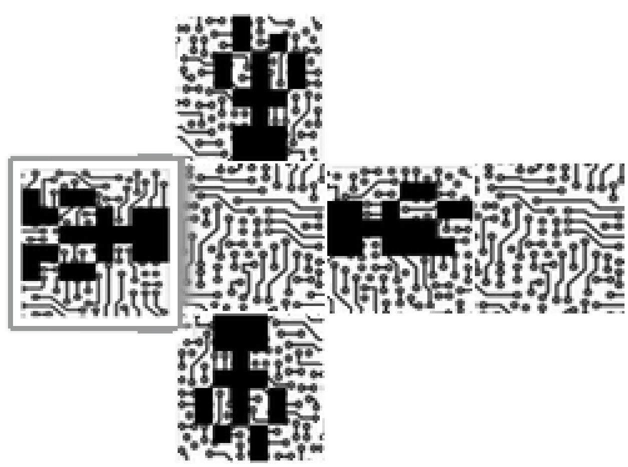


Figure 6. Graphic design of the target cube

Motion control interface

As we have mentioned, there are two motion control parameters in our animation controller which are "Speed" and "Angular". The way we infer "Speed" value is by computing the cube orientation respect to the camera orientation, taking into account the images on the cube faces. So, we compute the angle between the forward vector of the camera and the forward vector of the cube. Then, we transform this value depending on the section. If the obtained angle is between 0 and 45, "Speed" value is directly 0. In other cases, "Speed" value is obtained by linearly transforming the range of the value from 0 to 180 to 1.55 to 5.55. As for the "Angular" motion control parameter, it is extracted from the difference between the current direction of the movement and the previous one. This depends on the interaction scenario.

DISCUSSION AND FUTURE WORK

This work comes up with a proposal to improve the intuitive control of virtual characters and its inclusion of them in daily environment. We use a cube as a tangible interface that permits positioning and to select motion clips for latter drive a virtual character. The graphic design of the cube and the interaction model allows to easily understand how it works. This is because motion control parameters have been taken into account in the design process.

Although the proposed interaction already fulfill the animation system requirements, it is limited by the cube rotation. The cube rotation only allows a degree of freedom, and therefore manage one motion control parameter. In our prototype, it is to only support locomotion control, which in terms of videogames community, it is denoted as a gameplay. Change the gameplay by performing a gesture (i.e. shake the cube) and controlling it with rotation could be a possibility to increment the avatar capabilities, and to improve the character controller. As a consequence, the graphic design of the marker should be changed for not confusing the users since the cube faces denote the active motion.

On the other hand, the fact of using augmented reality to display the character performances produces several benefits. Firstly, the user can feel and see the digital content at same time. This way provides more visual and spatial awareness to the user. Secondly, it allows the scene and the interaction area to move to different locations This would allow to drive the virtual characters in different places through using mobile devices. However, some changes may to be introduced in the current prototype for computing the appropriate poses of the virtual characters.

ACKNOWLEDGMENTS

We thank Daniella Raffo for her work on making the cube prototype. The authors would like to thank Àngels Aragonès for her help of recording the accompanied video (https://vimeo.com/82277562) of this paper. Also thank Francesc Fernández for help in the sketch drawings.

REFERENCES

1. Amores, J., Benavides, X., Comín, M., Fusté, A., Pla, P., and Miralles, D. Smart avatars. In *IUI 2014 Workshop on Interacting with Smart Objects (submitted)* (Haifa, Israel, February 2014).
2. Heun, V., Kasahara, S., and Maes, P. Smarter objects: Using ar technology to program physical objects and their interactions. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, ACM (New York, NY, USA, 2013), 961–966.
3. Ishii, H., and Ullmer, B. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, ACM (New York, NY, USA, 1997), 234–241.
4. Lockwood, N., and Singh, K. Finger walking: Motion editing with contact-based hand performance. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '12, Eurographics Association (Aire-la-Ville, Switzerland, Switzerland, 2012), 43–52.
5. Numaguchi, N., Nakazawa, A., Shiratori, T., and Hodgins, J. K. A puppet interface for retrieval of motion capture data. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '11, ACM (New York, NY, USA, 2011), 157–166.
6. Oshita, M., Senju, Y., and Morishige, S. Character motion control interface with hand manipulation inspired by puppet mechanism. In *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '13, ACM (New York, NY, USA, 2013), 131–138.
7. Qualcomm Connected Experiences, Inc. Unity Extension - Vuforia v2.8, Dec. 2013.
8. Unity Technologies. Unity - Game Engine, Dec. 2013.
9. Unity Technologies. Unity - Mecanim - Simple and powerful animation technology, Dec. 2013.

Smart Avatars: using avatars to interact with objects

Judith Amores
Seamless Interaction, GTM
La Salle RD, La Salle - URL
Barcelona, Spain

Xavier Benavides
Seamless Interaction, GTM
La Salle RD, La Salle - URL
Barcelona, Spain

Michel Comín
Seamless Interaction, GTM
La Salle RD, La Salle - URL
Barcelona, Spain

Anna Fusté
Seamless Interaction, GTM
La Salle RD, La Salle - URL
Barcelona, Spain

Pol Pla
Fluid Interfaces Group, MIT
Media Laboratory
Boston, USA

David Miralles
Seamless Interaction, GTM
La Salle RD, La Salle - URL
Barcelona, Spain

ABSTRACT

This paper focuses on a new interaction system named Smart Avatars, which is based on a mixed reality environment containing virtual avatars as a medium for the interaction between the user and the smart objects of the environment. As a prototype, we introduce Flexo. With Flexo we investigate the virtues and obstacles that occur when we use an augmented reality (AR) avatar to interact with a simple smart object. Then, we conclude with the advantages and disadvantages of this kind of interaction in order to create richer interactions in future prototypes.

Author Keywords

Mixed Reality, Augmented Reality, Internet of things

General Terms

Human Factors; Design.

INTRODUCTION

Internet connection provides us a new way to access and to interact with real and tangible objects. This technology has renamed these objects as smart objects. Taking advantage of this technology we present an application that allows these smart objects to self-explain their functions. In this paper we show a way to access smart objects through a new augmented reality (AR) interface called smart avatar.

When we look for relevant projects that use AR as an interface to control Smart Objects, we have to take a look to CRISTAL [11], a project that goes a step further in this kind of interaction. In this project a collaborative interactive tabletop system is presented as a way to control all kinds of digital devices through a virtually augmented image of the surrounding environment. In this project the video stream becomes a form of interaction between the virtual devices that are connected to the network.

IUI 2014, Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel.

Another project that is related to the intelligence of the physical objects is Augmenting Sticky Notes [9]. In this project, the authors explored the way to enrich the experience of using the ubiquitous physical sticky-notes by augmenting its information and linking this to the digital world.

An improvement in the ubiquitous computing is the introduction of AR as an interface to interact with digital devices. A significant contribution is done in [7] where a single controller that uses an augmented reality (AR) interface is explored as a way to interact with several different smart objects connected to the home network. Further work would be evaluating the benefits of having real-time feedback on AR applications [8].

Finally, parallel to this investigation, Huen has been working on how to create smarter objects [6]. Through an AR application, the user can amplify the interaction of the surrounding smart objects. Also, using this AR application, the user is able to connect different smart objects. As an example, a speaker is shown with a simple interaction: go to previous or next track, and raise or reduce the volume. When the user points with the mobile device to the speaker and the AR is displayed, then an enhanced control layer appears and the user is able to add the tracks that he wants to listen to. In addition, if the user drags and drops this speaker in the display to the nearest speaker, this one will reproduce the same track that the first speaker is playing. In further steps, Huen has implemented Reality Editor [5], a system that it is able to edit the behavior and interfaces of smart objects, using an AR interface over this.

In this work we introduce an interaction system that could go further in helping to advance in the connection between tangible and digital objects. Our goal is to explore how digital information can modify the real world (and vice versa) in order to design interfaces that create new narratives between the user and technology. Here we present a digital interface which is strongly linked to a real object. The visual representation of this interface is an avatar and it only appears through a tablet device when the real object is present. Furthermore, the avatar is capable of changing the state of the object. Obviously this object must be a smart object and it needs to be connected to the tablet. This avatar is the essential part of

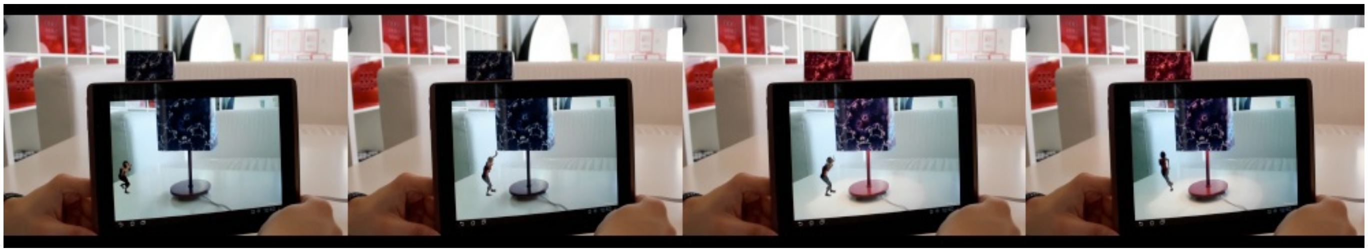


Figure 1. Sequence where a smart avatar comes close to the lamp and turns it on.

a new interface between users and real objects called Smart Avatars.

INTERACTION SYSTEM

Smart Avatars is a concept based on a mixed reality environment created using augmented reality. From the emergence of augmented reality as a new interface to interact with our environment, experts have been made great efforts to find new ways of introducing virtual data into the real world as a way of increasing the available information of tangible elements. Smart Avatars use augmented reality to create a new way of intuitive communication between the user and these tangible elements. With the mixed reality virtual objects are merged with physical objects generating a transversal environment where the virtual and the real coexist in the same layer. With this new type of interface, the advantages of the two worlds can be exploited to design improved interactions.

This interface allows the user to interact with the smart object through an augmented reality dynamic interface and can configure the functionalities of the object at each moment based on his own needs. The aim of Smart Avatars is to redefine the interfaces for these new smart objects that can be considered as robotic entities with a new approach which aims to bring down the barriers between people and complex technology. There is no longer the need to have hundreds of input devices to interact with them because every augmented reality interface will appear when needed.

Figure 2. Flexo prototype.

The interface contains human-like avatars that create a closer relation between the user and the smart object. Different research [2, 1] has shown that humans can solve problems easily when they are driven by emotions. Moreover, other researches have shown the positive value of the use of emotionally expressive avatars in virtual environments [3]. With Smart Avatars, this new narrative included in the interaction increases the implication of the user with the system, as it is generating a direct influence into user emotions.

In addition, the use of a mobile device to display this augmented reality becomes another key aspect for the hardware performance. All the operations are computed in the mobile device; hence there is no need for a high computational cost hardware. This is the fact that makes Smart Avatars a wholly multiplatform concept. Besides, being connected to the cloud at any time allows the smart object to send data and receive data whenever it is needed.

As a first example, we have developed a prototype named Flexo. Flexo applies the concept of Smart Avatars to a basic smart object, a lamp (Figure 2). Using a tablet, the user can turn on and off the lamp through the augmented reality displayed on the device. When the user performs an action (Figure 1), a virtual avatar appears next to the lamp and acts as the medium to execute that action. In this case, the avatar approaches the lamp and turns it on. The avatar becomes the interface between the user and the smart object. The user gets more involved in the interaction as the expressive avatar used to turn on the light creates this new narrative that captures the users attention. Flexo uses an extremely basic smart object with only two states. However, the idea is to use this concept into more complex smart objects where the interaction will be more intuitive and more natural than using a simple graphic layout using buttons. Moreover an interactive control of the avatar [4] could be implemented in order to interact with different smart objects.

IMPLEMENTATION

The Flexo prototype has passed through three different development stages. The first stage was to hack the lamp so that it would be recognized by a smartphone since the device would display augmented reality content. To do so, we designed a squared lampshade that had an image target in each face (Figure 3). To recognize and track the features of the image targets we used Vuforia [10].

The second stage was to give a special character to the augmented reality content. To do so, we created a 3D animated avatar. To animate the avatar in the most real way, we recorded the movement of a person using Motion Capture. We mapped the animation to the 3D model so the model performs the same actions as the actor. Throughout Unity [12] we unified the power of the rendering of a game engine for 3D content with Vuforia augmented reality.

The third stage was to make the lamp smart enough to communicate with the smartphone or the tablet device. One of our goals was to provide an easy wireless connectivity to our object. After evaluating different technologies, we decided that the Wi-Fi was the best solution - we can have a direct connection to Internet without using an intermediate device such a Bluetooth or Zigbee. Wi-fi is a versatile technology because it allows the object to connect to different devices. In order to have the wireless control of turning on/off the light, we designed an electronic circuit that is connected to a microcontroller. Through the microcontroller we can deactivate a 220 V relay that cutts off the power supply, and as a consequence, the light turns off. To achieve the control of that system and provide reliable transmission, we have created a TCP/IP communication socket to the mobile device in charge to send the state of the light. In this way, we are displaying augmented reality content according to the information we are receiving.

DISCUSSION AND FUTURE WORK

Sometimes the connection between smart objects and users must be enhanced in order to achieve a reliable experience. In this work we have developed a new digital interface that permits the interaction of smart objects through an avatar. Although part of this interaction is displayed on a mobile device a deep spatially awareness is perceived. Linking Internet of things and AR Smart Avatars go beyond the mobile display and the graphical layout of buttons.

With this new approach, the connection between the user and the smart object becomes deeper and more affective. The user gets emotionally involved in the interaction, as the medium for it is a human-like avatar. As a result, a better user experience is achieved.

This case study focuses on a lamp as the smart object to interact with. This one has been chosen as a simple object to make the concept more understandable. However, this idea has been thought out to be applied to more complex objects, where the user needs some kind of help to understand their functionalities.

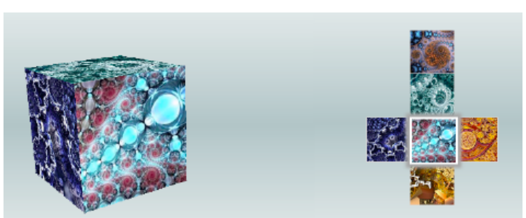


Figure 3. 3D marker used as lampshade.

Some aspects about tangibility could be improved in this kind of interface. Custom-made prototypes including some sensors within them can provide to Smart Avatars a tangible component. In this way users get more involved with smart objects.

ACKNOWLEDGMENTS

The authors would like to thank to Marc Expósito for recording the video that accompanies this paper. (https://vimeo.com/82336816)

REFERENCES

1. Cooper, B., Brna, P., and Martins, A. Effective affective in intelligent systems building on evidence of empathy in teaching and learning. In *Affective Interactions*, A. Paiva, Ed., vol. 1814 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2000, 21–34.
2. Damasio, A. *Descartes' Error: Emotion, Reason and the Human Brain*. Random House, 2008.
3. Fabri, M., and Moore, D. The use of emotionally expressive avatars in collaborative virtual environments. In *Joint Symposium on Virtual Social Agents* (2005), 88–94.
4. Fernandez-Baena, A., and Miralles, D. Avatar: Tangible interaction and augmented reality in character animation. In *IUI 2014 Workshop on Interacting with Smart Objects (submitted)* (Haifa, Israel, February 2014).
5. Heun, V., Hobin, J., and Maes, P. Reality editor: Programming smarter objects. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, ACM (New York, NY, USA, 2013), 307–310.
6. Heun, V., Kasahara, S., and Maes, P. Smarter objects: Using ar technology to program physical objects and their interactions. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, ACM (New York, NY, USA, 2013), 961–966.
7. Lee, J., Kim, J.-I., Kim, J., and Kwak, J.-Y. A unified remote console based on augmented reality in a home network environment. In *Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on* (2007), 1–2.
8. Liu, C., Huot, S., Diehl, J., Mackay, W., and Beaudouin-Lafon, M. Evaluating the benefits of real-time feedback in mobile augmented reality with hand-held devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, ACM (New York, NY, USA, 2012), 2973–2976.
9. Mistry, P. Quickies: The future of sticky notes. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, ACM (New York, NY, USA, 2009), 3529–3530.
10. Qualcomm Connected Experiences, Inc. Unity Extension - Vuforia v2.8, Dec. 2013.
11. Seifried, T., Haller, M., Scott, S. D., Perteneder, F., Rendl, C., Sakamoto, D., and Inami, M. Cristal: A collaborative home media and device controller based on a multi-touch display. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, ACM (New York, NY, USA, 2009), 33–40.
12. Unity Technologies. Unity - Game Engine, Dec. 2013.

Tangible User Interfaces applied to Cognitive Therapies

Elena de la Guía
Computer Science Research Institute
University of Castilla-La Mancha
Albacete, Spain
MariaElena.Guia@uclm.es

María D. Lozano
Computer Systems Department
University of Castilla-La Mancha
Albacete, Spain
maria.lozano@uclm.es

Victor M. R. Penichet
Computer Systems Department
University of Castilla-La Mancha
Albacete, Spain
Victor.Penichet@uclm.es

ABSTRACT
Interactive games to support cognitive training are increasingly becoming an indispensable resource in cognitive therapies. At the same time, technological advances are definitely causing the appearance of new paradigms and different styles of interaction. In this paper, we take advantage of real physical objects and the benefits that new technologies offer us, in order to design a new way to interact with interactive games in cognitive therapies. The system is based on physical objects that integrate NFC technology and allow the final user to interact with Distributed User Interfaces. We analyze the effects of interacting with smart objects in Multi-Device Environments developed for people with intellectual disabilities.

Author Keywords
Tangible interaction, NFC technology, Distributed User Interfaces, Collaboration

ACM Classification Keywords
H.5.2. Information interfaces and presentation: User Interfaces. – Graphical user interfaces

General Terms
Design, Human Factors, Experimentation.

INTRODUCTION

Nowadays three per cent of the world's population has some type of intellectual disability [29]. This disability limits their life and the life of those around them. Recent research on brain plasticity emphasizes that conducting a systematized practice and repetition makes the brain favourably modify its structure and operation, offering the possibility to optimize the performance and cognitive abilities [23].

Technology is a crucial tool in cognitive therapies, as it offers an interactive way to improve, stimulate and develop the user´s cognitive abilities [24]. People with special needs experience difficulty using traditional computers. For this reason engaging and enjoyable software is necessary to encourage them to use it. In this way videogames are elements with a high level of motivation. They can have a positive effect on people by helping them focus their attention and enhance their interest in the tasks they are performing [28].

Nowadays a number of games for people with special needs or intellectual disability have been developed. These games offer many advantages, nevertheless interaction based on using mouse and keyboard and virtual reality devices can be an obstacle for people with limitations.

This paper presents a system, called TraInAb, based on a set of interactive games that uses Tangible and Distributed User Interfaces to offer an amusing environment so as to improve and stimulate cognitive capabilities of people with special needs. The main goal is to create an amusing scenario especially designed for people with special needs overcoming the technological barrier between people and technology.

Tangible User Interfaces (TUI) refers to user interfaces which give physical form to digital information, making the parts directly malleable and perceptible [8]. Tangible User Interfaces are based on smart objects, and provide a natural and easy style of interaction that proves intuitive and motivating for non-experts in technology and people with special needs [10, 15]

According to Niklas Elmqvist in [6], Distributed User Interfaces (DUI) can be defined as a user interface which components can be distributed through one or more dimensions. These dimensions are: input, output, platform space and time. Although some definitions of the term have been proposed, at present there is no formal definition. Some approaches to a definition of this term can be found in [20].

The rest of the paper is organized as follows: Section 2 describes some related works, introducing types of games for intellectual disabilities and styles of interaction in the new scenarios. In section 3 we present the system we have developed. Section 4 shows the experimental data obtained after evaluating the system with real users. Finally, section 5 presents conclusions and final remarks.

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel

RELATED WORKS

This section includes some important concepts with respect to the proposal we present in this paper and a brief overview of several related works to highlight the differences with respect to the system proposed.

A video game is a software program created for entertainment and learning purposes in general. It is based on the interaction between one or more persons and an electronic device that executes the game.

"Serious Games" are games that simulate real situations for people with disabilities, such as shopping in the supermarket. The main objective is to develop the skills that can help them in their daily activities [16].

It is not easy to determine which game is more adequate for intellectually disabled players. The barriers that people with cognitive disabilities may find during the activities are complex and varied as described in [4] and [9]. These studies highlight that the key element in the games must be simplicity.

On the other hand, in Virtual Reality applications using helmets, gloves and other simulators, the user may feel more immersed in the game, and it is very engaging and motivating, but the problem is the high cost of devices, and the difficulty of using certain devices. In addition, an expert is required to control the players and devices [17, 18].

The advantages offered by computer games in general are numerous. They enhance positive attitudes in users while being appealing and encouraging. However these systems present the following disadvantages:

-The user needs a minimum knowledge about computer use. Not everybody can use a computer and some devices, like the mouse or the keyboard are not intuitive for people with cognitive disabilities. They need someone to help them.

-These systems sometimes require highly specialized hardware/software which can be expensive (simulators, virtual reality).

- In some games, impaired users may have difficulty finding specific information.

The rapid evolution of technology has changed the way in which we can interact with interactive systems. New scenarios have appeared such as Multi-Device Environments (MDE). These scenarios implicitly support Distributed User Interfaces. However people with disabilities have significant limitations with new technologies; for this reason it is necessary to provide an intuitive and simple style of interaction so that this is not an obstacle to use the system.

Among the existing interaction techniques in the new scenarios next we highlight those that are more adequate for people with cognitive limitations.

-Touching. This technique involves touching an object, either with a finger or with a mobile device, to perform a task. Some examples of projects using this technique can be found in [1] [7] [21].

-Scanning. The mobile device or any other device is capable of scanning information and interact with the system to provide a service to the user [22].

- Approach&Remove [14]. It is a style of interaction that allows the user to interact with distributed user interfaces by approaching a mobile device to digitized objects.

The system we present in this paper is based on this last style of interaction. Some of the projects that use this type of interaction are the following: Interactive EcoPanel [19] is a system that uses RFID technology to support interactive panels to provide a collaborative application where users can share their opinions and ideas about an environmental issue just by using natural and intuitive gestures.

A map has been digitized in [14], with the aim of allowing users to find and watch on their mobile devices some information related to a specific area. These systems enable users to interact with the digitalized objects, which are used as physical interface.

The main difference with respect to our proposal is that all of them use mobile devices to interact with objects, whether posters, walls, or in the case of the "Touching" technique, touch devices. However, in our proposal not all players need a mobile device. Just one device is required to allow that multiple-users can interact with the system. Thanks to digitized objects, interacting with the system is simple and intuitive for all users. This new style of interaction helps users overcome their fear to interact with new technologies. In this way, people with special needs feel motivated and self-confident when they use the system

TRAINAB SYSTEM

TraInAb (Training Intellectual Abilities) is an interactive and collaborative set of games designed to stimulate people with intellectual disabilities.

The system integrates a new style of interaction. The user can interact with the system through everyday objects such as cards, toys, coins, etc.

The collaborative system is based on the distribution of interfaces and device mobility; it offers the possibility to be used individually or by multiple users.

The functionality of the system is as follows. The game interface is projected onto the wall. Users interact with the game through digitized tangible objects integrating NFC tags inside. The user has to bring these objects closer to the mobile device that incorporates an NFC reader so that the mobile device can identify the object. For example, when the game is waiting for the user to select a concrete object, the user only has to bring the corresponding object closer to the mobile device, and then the system recognizes it and displays the outcome of the game showing whether it was the correct one or not.

Design guidelines

The system has been designed according to some key guidelines described next.

The game interface is executed on a computer and projected onto a wall, what does not require the use of other peripherals like joystick or mouse. This requisite ensures that the game can be set in any environment, provided the existence of the following devices, correctly configured: a projector, the smart objects (tangibles user interfaces) and a mobile device with NFC reader. Smart objects are the main elements to interact with the games.

According to Brundy in [2], there is a need to engage and involve the user in the game. To ensure this requisite, the game sends out ludic cues, as invitations to play. This ludic cues corresponds to animated images and sounds. Ludic cues include verbal messages to facilitate the use of the game and get the attention, generate enthusiasm, and maintain the concentration of users, who are guided by the game interfaces.

Another important requirement is avoiding frustration due to failure. Our system shows positive and encouraging messages when the user fails. When the user gives a right answer, the system shows immediate reinforcement based on encouraging messages and points, motivating the progress toward goals and skills development.

Besides, the Distributed User Interfaces have been designed taken into account visual impairments. In our system, the game interface is executed on large-print displays, alternative colours on the screen, and voice output to compensate for some reading and attention problems.

Finally, flexibility of performance and continuous gratification is another requirement included in the system. The game automatically adapts to the skill of the user while in performance. This allows a greater control over the challenge proposed on the different skill levels of the end users.

Distributed User Interfaces

The interface is the means by which the user interacts with the system, therefore it is very important that they are easy to understand, simple and intuitive. TraInAb is implemented in a distributed user interface setting composed of t three types of interfaces:

-MainUI (Main User Interface). It is the main interface of the system. It graphically displays the game information, including animations, texts and sounds. At all times, it shows the progress and course of the gameplay (See Figures 3, 4, 5).

-MobileUI (Mobile User Interface). It corresponds to the graphical interface of the mobile device that incorporates the NFC reader. Its function is to recognise the object chosen by the user when they bring it close to the mobile device (approach & remove). It also shows the instructions of the game at the beginning (See Figure 1).

-TangibleUI (Tangible User Interfaces). These are common physical objects used as interaction resources (IR) to interact with the game. In this case we have three different types of objects: cards with pictures of the different games, coins and notes (See Figure 1.b-c-d). These objects have an NFC tag integrated inside what allows that the system can identify them through the mobile device (See Figure 2).

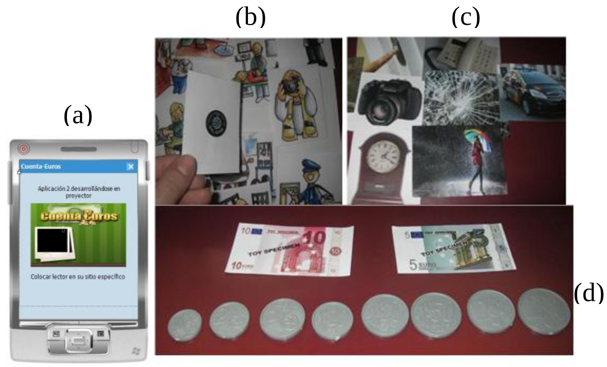


Figure 1. (a) Mobile device interface. (b,c,d) Tangible user interfaces for Game 1(b), Game 2 (d) and Game 3 (c)

Interaction Style

The style of interaction of the user with the system simulates the usual style of the user in their environment. Therefore no prior knowledge is necessary, as using the system is easy and intuitive (see Figure 3). The user only has to bring the chosen object closer to the mobile device. The actions that result from this are transparent for the user.

The game interface displayed onto the wall shows the game executed in that moment. Depending on the game, it may show different objects and the user will have to interact with the system by choosing the correct object and bringing it closer to the mobile device. From that moment all processes are run implicitly. The game interface will display the pictures, texts and sounds according to the final result of the user action. In case of failure, the game will keep the game at the same skill level and in case of success, it will move on to the next level of difficulty.

Next we describe the interaction components and their function in the system:

IR (Interaction Resources). The digitised physical objects. These resources support the interaction with the system. In this case study, we have used everyday objects such as coins or cards to interact with the system. ID (Interaction Devices). The interaction devices correspond to mobile devices. They allow the communication between the digitized objects and the system through NFC and wireless technology.

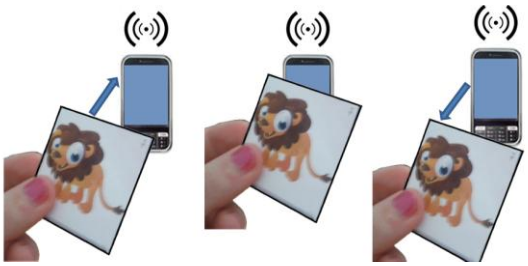


Figure 2. Style of Interaction. The user brings the digitized object (Tangible User Interface) closer to the mobile device (interaction resource) that incorporates the NFC Reader.

Games

TraInAb integrates a set of games focused on improving abilities for the integration of the user in a city. The system has two modules: The first one corresponds to the Management and Administration by the therapists and the second one corresponds to the gameplay management.

Management and Administration Module

This part is responsible for controlling the games and users' statistics. This module has been designed to be used by the therapist, teacher, mentor or psychologist. This person is in charge of administering the system. This module allows the user to choose the game they want to run, as well as to change and save it.

Gameplay Management

The system consists of three different types of game, each aimed at stimulating a different cognitive ability such as memory and attention, calculation, and auditory discrimination.

All the games have been designed with three different levels of difficulty to adapt to the user's skills.. If the user fails, they lose a life and if the user wins they move on to the next level. The feedback messages are motivating for the user to feel encouraged to continue playing. Besides, the system shows this feedback information in a different way depending on the level of difficulty.

In addition, the game shows the status and game results at any moment.

Game 1: Memory

The first game is aimed at memory training. It trains the ability to temporarily retain in memory some information which is later used to produce a specific result. This is a special component of other higher order cognitive processes and includes other cognitive skills such as attention, concentration, mental control and reasoning.

The game presents three cards that the user has to memorise. After a few seconds, the cards disappear and only two of the initial three cards appear again. The goal is that the user remembers the missing card. Then, the user selects the correct tangible objects corresponding to the missing card and brings it closer to the mobile device. The result is displayed on the game interface with positive and encouraging messages regardless of the outcome of the game. The sequence of this gameplay is depicted in Figure 3.

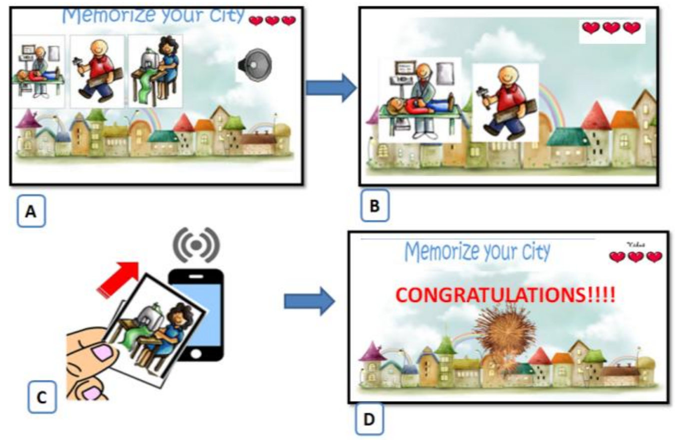


Figure 3. Game 1 Sequence. (a) Main interface of the game. Firstly, users have to memorize the cards (b) One card disappears (c) Bringing the selected tangible interface closer to the mobile device (d) Game results

Game 2: Counting Euros

This game has been designed to improve cognitive abilities related to calculation and logic. The game shows the price of different items such as food, clothes, drinks, and so on. The user has to calculate the price using the coins and notes available as tangible objects. The main goal is to improve concentration as well as the ability to make calculations and handle money (See Figure 4).

Figure 4. Main interface of the game designed to stimulate calculation skills (Game 2)

Game 3. Auditory discrimination

To train hearing skills and increase the user concentration, this game reproduces common sounds in the city such as cars, phones, etc. and the user has to concentrate in order to associate the sound with the corresponding card (See Figure 5).

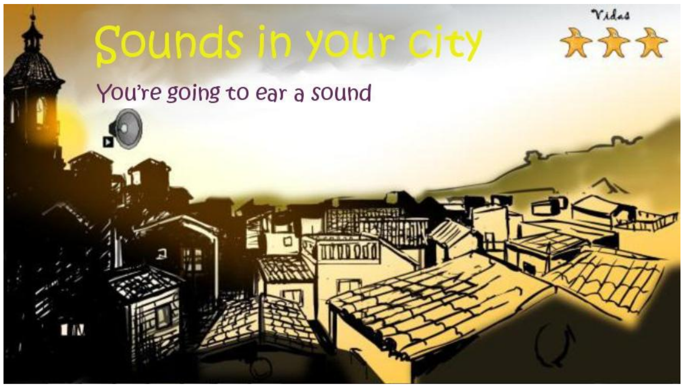


Figure 5. Main interface of the game designed to improve listening skills, attention and concentration (Game 3).

System Advantages

The main advantages of the system can be summarized as follows:

Reduction of cognitive load. The system has been designed with simple and easily understandable graphics and tangible objects are easily recognized.

Flexibility. It refers to the multiple ways in which the user and the system can exchange information. The information exchanged is displayed as text, voice, cheerful sounds or using graphics. The goal is to adapt the system to any user, regardless of the disability or limitation they may have. Besides, it refers also to the flexibility in the number of users. This is a multi-player game. It allows sharing and exchanging experiences with other users. The situation of each user may be complex and variable and, for this reason, the game can also be used by only one player. Flexibility in terms of space. Players can be moving around while playing, the only requirement is that the mobile device is connected to the server.

Cheap to deploy. Passive NFC tags are quite inexpensive. Besides, only one mobile device with NFC reader is required.

Expandable. It offers the possibility to extend the games. The topics can be easily changed. The only requirement is that the NFC tags must be integrated inside the tangible objects.

Interaction with the system is simple and intuitive. Common items are familiar and can be easily assimilated by users, making it more predictable to use. They do not need prior knowledge of the system or device to use it.

Cognitive stimulation may enhance mental abilities such as perception, attention, reasoning, abstraction, memory, language, orientation processes, while optimizing the users' performance. These games can be an objective therapy for cognitive deficit.

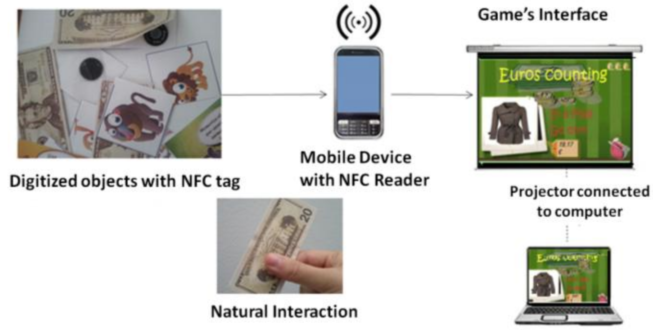


Figure 6. Digitized objects with NFC tags that communicate with the game interface when users bring them closer to the mobile device.

System Architecture

The TraInAb system is a client-server system designed as follows. The client system runs in the mobile device. It is connected to the server application through a wireless network and it is communicated with objects via NFC when the user brings the object closer to the mobile device. A tag (or more) is integrated inside the object or card depending on the size of the object; each tag describes a unique identifier. When the tangible object is brought closer to the NFC reader in the mobile device, the NFC tag inside the object is excited by electromagnetic waves sent by the NFC reader, and then the component controller sends the identifier to the server. The server maps this information in the database and executes the steps necessary to return the information to the mobile device. The games which are running in the PC are displayed onto the wall through the projector (See Figure 6).

EXPERIMENTAL STUDY

This section describes an evaluation of the interactive system performed by people with intellectual disability. The main goals of carrying out the experimental study were:

- To test the strengths and weaknesses of the system.
- To investigate if people´s cognitive abilities improved and whether they enjoyed playing with the games.
- To test the effect of new user interaction based on tangible objects.

Participants

Twelve people with intellectual disability participated in the evaluation (3 female and 9 male). They were recruited from the Center for Assistance to Disabled Persons and Families in Albacete, Spain. Participants ranged from 7 to 62 years (mean=34,83; sd=2,86). Participants had no prior experience with the system.

Apparatus

The hardware used in the evaluation consisted of a Smartphone Samsung Google Nexus S, (16 GB de memory, 512 MG de RAM, processor ARM Cortex A8 1GHZ which incorporates a NFC reader and a private wireless ad-hoc network. A laptop (Processor Intel Core2 CPU t7500 2.20 Ghz with 2,0 GB de RAM) was used to execute the games. It was connected to a projector that showed the game interface onto the wall. The two devices (Smartphone and PC) were communicated through a wireless network. The projector was connected to the PC with a VGA wire. The host application and device application were both developed using C#. The server application and game application are hosted on the same computer. The server application received incoming connections from the mobile device and computer.

Procedure

In order to perform the experiments, we conducted three sessions with the same participants. The first session was divided into three phases: pre-test, test and post-test.

The pre-test phase aimed to obtain information about the participants' profile. In order to find out who had previous experience with technology and stimulation activities we asked the following questions: (1) Have you ever played a video game? (2) How many times a week do you play? (3) Do you like video games (4) What kind of devices have you used to perform the games?: computer, board games or other devices.

The second phase consisted in playing with the interactive system "TraInAb". Firstly, we explained the games and the new way to interact with the system. Then we introduced them the game focused on memory skills (games 1), the game focused on improving calculations and logic abilities (game 2) and the game focused on improving auditory discrimination (game 3).

In the third phase, called post-test, we distributed Smileyometer tests and then we asked them some questions such as: (1) Have you enjoyed playing?, (2) Would you play again? (3) What did you like most about the game?.

While performing the tasks, a video camera was recording the complete session, two evaluators wrote down the playing time and the number of errors occurred, while a psychologist was responsible for assessing the nonverbal messages of users.

The second evaluation session was conducted a week after the first session and the third one two weeks after the first session. In these session, the participants played games 1, 2 and 3. The goal was to analyze whether the participants improved their performance after using the system.

Methods

During the evaluation sessions, we followed the steps defined in [11]. We also used the Smileyometer test that is a method which allows the user to choose from among five pictorial representations, ranging from awful to brilliant, to express their opinion [13] (See Figure 7).

Another method used was Direct Observation [12]. The purpose of this kind of evaluation is to enable designers to better understand how the users make use of the systems in their natural environments. Data is collected in informal and natural ways, with the aim of causing as little disturbance as possible to users. In order to obtain data on the emotions and feelings about the game (playability), we analysed nonverbal messages, gestures [5] and facial expressions [3] shown by the participants.

Figure 7. The Smileyometer test: awful, not very good, good, really good, brilliant

Results

The results of the first evaluation session were as follows: 80% of the participants had played a video game before and they often played for an average of 3.8 times a week, and 72,3% of them liked these games. 31% of the participants used board games compared to the 69% who used computers. None of them had used another technological device such as Kinect or Nintendo Wii. We obtained the following results from non-verbal messages. The first users' reaction was as follows: 44,7 % of the participants were disappointed, 21,3% were unmotivated and only 31,5% were happy and calm. When they interacted with the game, we found that 12,5% of the participants were surprised by the way they interacted with the system. The remaining 20% were still disappointed, and they couldn´t begin to relax until they began to play and win. While they were playing the game, 90% expressed motivation and interest, but 10% expressed indifference to the tasks. 11,1 % of the participants were disappointed, 7.63 % were unmotivated and 65.1 % were happy and calm. Figure 8 shows how the participants felt while playing with the games. The mood of the participants changed, increasing the feeling of calm and relax. That is, the users stated enjoying while playing, decreased fear and users who were unmotivated while playing was just 6.2%.

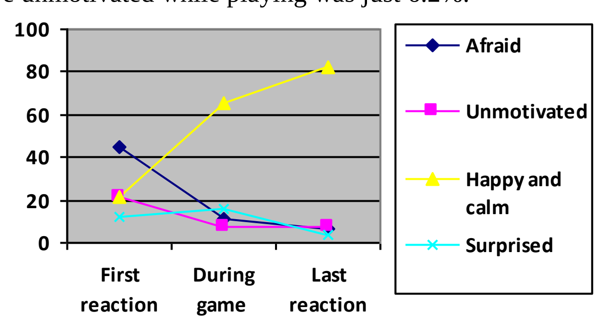


Figure 8. Participants' emotions while using the TraInAb system

In the post-test phase, the following data was collected. In response to the questions stated before, 87% of the participants replied that they perceived the game as innovative and exciting, as opposed to the 13% who were not impressed. 90% would like to play again because they had a good time and were quite entertained. What 70% of the participant liked most was interacting with the system through physical objects, because they were engaged and felt curiosity and interest in the functionality of the system.

The Smileyometer test results were as follows: 7 out of 12 participant thought that the games was "Brilliant," versus 3 out of 12 participant who thought it was "Really good", 1 out of 12 thought it was "good " and 1 out of 12 thought it was "not very good." None of them thought that the games were awful.

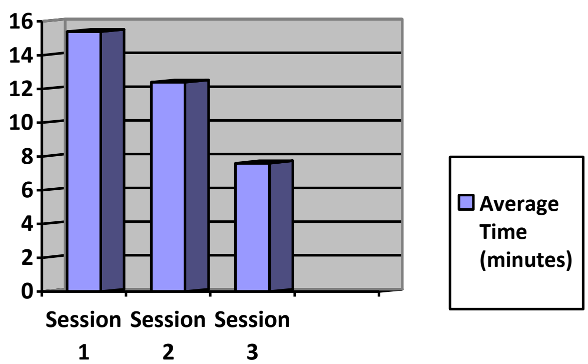


Figure 9. Mean time obtained from the three evaluation sessions performed

To check if the participants had improved their cognitive capabilities, we studied the average execution time of certain tasks. Then, we analysed the data obtained. The average execution time was the dependent variable of the experiment and the independent variable was the session (session 1, session 2, session 3). To analyse if the independent variable influenced the dependent variable we proposed the following hypothesis: the average execution time of the games would be the same from one session to the other. We repeatedly performed the procedure ANOVA on the mean error rate data and found a significant main effect with target width ($F(9, 87) = 27.9$, $p < 0.001$).

Contrary to our hypothesis, we found that the averages were not equal, since the time decreased in each session (See Figure 9). Therefore, the repetition of the game affects the performance and improved cognitive abilities such as memory, calculations and attention.

3. CONCLUSIONS

TraInAb (Traning Intellectual Abilities) is an interactive and collaborative set of games based on distributed and tangible user interfaces developed with emerging technologies such as NFC. The main objective is to stimulate and improve cognitive abilities of people with intellectual disabilities. In order to interact with the system users have to bring tangible objects integrating NFC tags inside closer to the mobile device that incorporates the NFC reader, and then the results and other relevant information is projected onto the wall. This style of interaction is simple and intuitive; its purpose is to eliminate the technological barrier for people with intellectual disabilities. In order to evaluate the beneficial effects of the games and the new style of interaction in the user, we have performed an evaluation based on "Direct Observation". The results have been very positive. Users with Intellectual disabilities were highly motivated and interested in the system. They found it easy to use and enjoyed it. The evaluation was performed in different sessions some weeks apart, in order to check whether cognitive capabilities like memory, calculation and attention improved after playing the game. The results were very optimistic, each session was better than the previous one and the users needed less time to play and concentrated more easily.

ACKNOWLEDGMENTS

This research has been partially supported by the Spanish research project TIN2011-27767-C02-01, from the Ministry of Economy and Competitiveness, and the regional projects with reference PPII10-0300-4174 and PII2C09-0185-1030, funded by Junta de Comunidades de Castilla-La Mancha. We would also like to give special thanks to Erika González and Yolanda Cotillas for their collaboration in the project, without whose support all this work would not have been possible.

REFERENCES

1. Broll ,G. Graebsch ,R., Holleis ,P. , Wagner,M. Touch to play: mobile gaming with dynamic, NFC-based physical user interfaces, Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, September 07-10, 2010, Lisbon, Portugal
2. Bundy, A.C., Play and Playfulness: what to look for. In D. Parham & L. Fazio, orgs. Play in Occupational Therapy for Children. St. Louis, MO: Mosby, 52-66. 1997.
3. Cohen, D. (2007). Body Language: What you need to know. Sheldom Press.
4. Dawe, M.: Desperately seeking simplicity: how young adults with cognitive disabilities and their families adopt assistive technologies. In: CHI '06: Proceedings of the

SIGCHI Conference on Human Factors in Computing Systems, pp. 1143–1152. ACM, New York (2006)

5. Doherty-Sneddon, G. (2009). El lenguaje no verbal de los niños. Lumen, Buenos Aires, Argentina

6. Elmqvist, N. Distributed User Intefaces: State of the Art.Workshop on Distributed User Interfaces2011 (DUI) at the 29th ACM CHI Conference on Human Factors in Computing Systems 2011, ISBN: 978-84-693-9829-6, Vancouver, Canadá, May 7-12, 2011.

7. Hardy, R., Rukzio, E. Touch & interact: touch-based interaction of mobile phones with displays, Proceedings of the 10th international conference on Human computer interaction with mobile devices and services, September 02-05, 2008, Amsterdam, The Netherlands

8. Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. In CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 234–241, New York, NY, USA. ACM Press. 19, 20, 27, 36, 73, 96, 187

9. Lewis, C.: HCI and cognitive disabilities. Interactions 13(3), 14–15 (2006)

10. Marshall, P. "Do Tangible Interfaces Enhance Learning?" Proc. First Int'l Conf. Tangible and Embedded Interaction (TEI '07), pp. 163-170, 2007.

11. MacKenzie, S., Castellucci, S. "Empirical Research Methods for Human-Computer Interaction". Course 28. May 9, 2012 Austin, Texas CHI 2012

12. Preece, J. et al. "Human Computer Interaction" ISBN: 0-201-62769-8. Addison Wesley.

13. Read, J. C., MacFarlane, S. J., & Casey, C. (2002). Endurability, engagement and expectations: Measuring children's fun. Paper presented at the Interaction Design and Children, Germany.

14. Romero, S., Tesoriero, R., González, P. , Gallud, J. A., Penichet, V. M. R.: Interactive System based on NFC to manage Georeferenced Documents. Interacción 2009, X Congreso Internacional de Interacción Persona-Ordenador. Barcelona. Septiembre 2009. ISBN-13:978-84-692-5005-1

15. Schneider, B., Jermann, P., Zufferey, G. and Dillenbourg, P. "Benefits of a Tangible Interface for Collaborative Learning and Interaction," IEEE Trans. Learning Technologies, vol. 4, no. 3, pp. 222-232, July-Sept. 2011.

16. Sik Lanyi, C. and Brown, D. J. Design of Serious Games for Students with Intellectual Disability. India HCI 2010/ Interaction Design & International Development 2010 20 - 24 March 2010 .Indian Institute of Technology, Bombay, Mumbai, India

17. Standen, P.J. and Brown, D.J.. Virtual Reality in the Rehabilitation of People with Intellectual Disabilities, Review, CyberPsychology & Behaviour, vol. 8, no. 3, pp. 272-282, 2005.

18. Takacs, B. Special Education & Rehabilitation: Teaching and Healing with Interactive Graphics, Special Issue on Computer Graphics in Education IEEE Computer Graphics and Applications, pp. 40-48, September/October 2005.

19. Tesoriero, R., Tébar, R., Gallud, J. A., Penichet, V. M. R., Lozano, M.: Interactive EcoPanels: Ecological Interactive Panels based on NFC . Proceedings of the IX Congreso Internacional de Interacción Persona-Ordenador Interacción 2008. ISBN: 978-84-691-3871-7; pp 155-165.

20. Vanderdonckt, J. Distributed user interfaces: How to distribute user interface elements across users, platforms, and environments. In Proceedings of the International Conference on Interaccion, 2010.

21. Vandervelpen, Ch., Vanderhulst, K., and Coninx, K. Light-weight Distributed Web Interfaces: Preparing the Web for Heterogeneous Environments. Proc. of 5th Int. Conf. on Web Engineering ICWE'2005 (Sydney, July 25--29, 2005). Lecture Notes in Computer Science. Springer-Verlag, Berlin, 2005.

22. Want, R., Fishkin, K.P., Gujar, A., Harrison, B.L.: Bridging physical and virtual worlds with electronic tags. In: Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, ACM Press, Pittsburgh, Pennsylvania, United States, 1999.

23. Frank Loprestia, Edmund; Mihailidisb, Alex; Kirschc Ned. Assistive technology for cognitive rehabilitation: State of the art. Neuropsychological Rehabilitation: An International Journal. Special Issue: Technology in Cognitive Rehabilitation. Volume 14, Issue 1-2, pp. 5-39. Taylor & Francis Group, 2004.

24. Griffiths, Mark. The educational benefits of Videogames. Education and Health, Vol. 20 No.3, pp. 2002

25. World Health Organization (WHO). World report on disability; 2011.

Electroluminescent based Flexible Screen for Interaction with Smart Objects and Environment

Orkhan Amiraslanov
German Research Center for artificial Intelligence
Trippstadter str. 122, D-67663, Kaiserslautern, Germany
Orkhan.Amiraslanov@dfki.de

Jingyuan Cheng
German Research Center for artificial Intelligence
Trippstadter str. 122, D-67663, Kaiserslautern, Germany
jingyuan.cheng@dfki.de

Peter Chabrecek
SEFAR AG
Töberstr. 4, CH-9425, Thal, Switzerland
Peter.Chabrecek@sefar.ch

Paul Lukowicz
German Research Center for artificial Intelligence
Trippstadter str. 122, D-67663, Kaiserslautern, Germany
paul.lukowicz@dfki.de

ABSTRACT
In this paper we propose an adjustable structure for flexible screens, based on electroluminescent phenomenon. The final product is thin, flat, flexible, long-lasting, easy to modify, reproduce and install. When combined with pressure matrix, it could become a touchscreen. Changeable pixel number and pixel size, plus the flatness and flexibility, make this structure ideal for interaction interface prototype for smart objects, where the surface size, shape and flatness are the main requirements. As demonstration we show this flexible screen on a window, on a bottle and on a gymnastics mat.

Author Keywords
Flexible screen, electroluminescent, interaction with smart objects.

ACM Classification Keywords
B.1.1 Hardware: Control structures and microprogramming

Control Design Styles: Hardwired Control

General Terms
Design, Verification.

INTRODUCTION
Visual interface demonstrates a very important role in human-machine and human-environment interactions. Nowadays, most of the screens however, are square-shaped and rigid, which limits their applicability in wearable and ubiquitous screen applications. When a soft surface with irregular contours, shapes and even thickness are the case (e.g. a pant or a spherical surface), then current commercially available screens are not able to fulfill the requirements. In this paper we focus on providing an easy prototyping method to equip unusually shaped surfaces with the flexible screen which fits perfectly in such applications. Although either used hardware structure (column and row scanning, details in section "System Structure") and the material (electroluminescent or light emitting capacitor) are not completely new, the combination of them makes the following contributions:

1. Easy prototyping: the screen is made by cutting the raw material and sticking uncoated electrodes along the rows. The control circuit design was implemented on a breadboard for ease of prototyping and design is ready to be implemented on a printed circuit board(PCB). Both pixel number and pixel size also the shape can manually be picked and chosen, while no special equipment other than soldering iron is required.

2. More natural surface, foldable and easy to install: the screen is flat and flexible in 2 dimensions. It can be rolled up before installing (e.g. for the wall out side of a building, the whole screen can be dropped from ceiling or window and spread out naturally through gravity, without any external scaffolding as when installing rigid display) or rolled up and carried easily before/after usage (e.g. gymnastic mat). The material is durable enough to withstand severe weather conditions.

3. Possible to become a touch-screen: combined with another layer of flexible resistive matrix [9], as a further implementation, we are planning to make a single layer of flexible touch screen.

IUI 2014 Workshop on Interacting with Smart Objects, Feb. 24, 2014, Haifa,Israel.

We would like to emphasize that, neither we plan to compete with commercial devices, nor we want to develop new electroluminescent material from scratch. The novelty of this work lies in combination, that we provide a method for quick prototyping a flexible screen in laboratory. The surface size can be either small or big with small or big pixel sizes, can be set to any shape, thus making it useful for human-environment and human-smart object interaction (e.g. a room where walls, cabinets, windows, tables and chairs, cups and etc. are all covered seamlessly with this screen. Examples are shown in Fig. 2, 3 and 4).

Related work

Looking into flexible screen solutions in the market, the most anticipated invention could be Youm: a rollable display for Samsung mobile [8], which will be released shortly. Such screen counts back to 2006, when Philips showed an OLED prototype. The application of these screens varies from small area as wrist display for smart watch, curved smartphone screen to big area as TV. We probably can expect buying such OEM screens with selectable size in the near future. However, having it in user-designed shapes (e.g. a spherical) and sizes are still infeasible. The other solution is LED displays, both as commercial available device for out-door advertisement or as small grids on chest [6] or as wrist worn display [4]. The benefit of LED display lies in its low cost and scalability. However, for large area, the amount of LEDs must be increased or an external layer has to be laid above to spread light from on a single LED into a bigger area. For small area, the pixel number is then limited by the smallest LED size ($\sim$ mm level). The height of single LED also makes a flat surface not straightforward (still possible, when embedded in a layer of transparent plastic). Additional work has been done on developing new material for flexible, elastic display [3], [1]. There are enough work on driving matrix structured displays, e.g. [5], [7], just to name a few.

SYSTEM STRUCTURE

The hardware structure of our system is given in Fig. 1. The whole display is composed of $n \times m$ electroluminescent plates plus n vertical drive lines and m horizontal drive lines. When one vertical line and one horizontal line are enabled and the rest are disabled, the electroluminescent plate at the crossing point is selected and lit up. By scanning trough all the vertical and horizontal lines at a speed quick enough (in this case, higher than 10Hz for human eyes), we form up an image on the screen. The whole circuit decoupled into two parts: high voltage AC part (current flow that provides energy for lighting up pixels) and low voltage DC part (driving circuitry with microcontroller). With a DC-AC inverter the required high AC voltage to drive the flexible display can come from a low-voltage battery pack, which means the whole system is portable (no need for getting energy from commercial power line).
The Electroluminescent (a.k.a. a Light-Emitting Capacitor, or LEC) panels are widely used in the industry to make background lights for LCD displays and as night lights. The panel is a capacitor where phosphor plays a role as dielectric and its sandwiched between the conductive plates. The light emission occurs only when high frequency($\sim 1KHz$) alternating

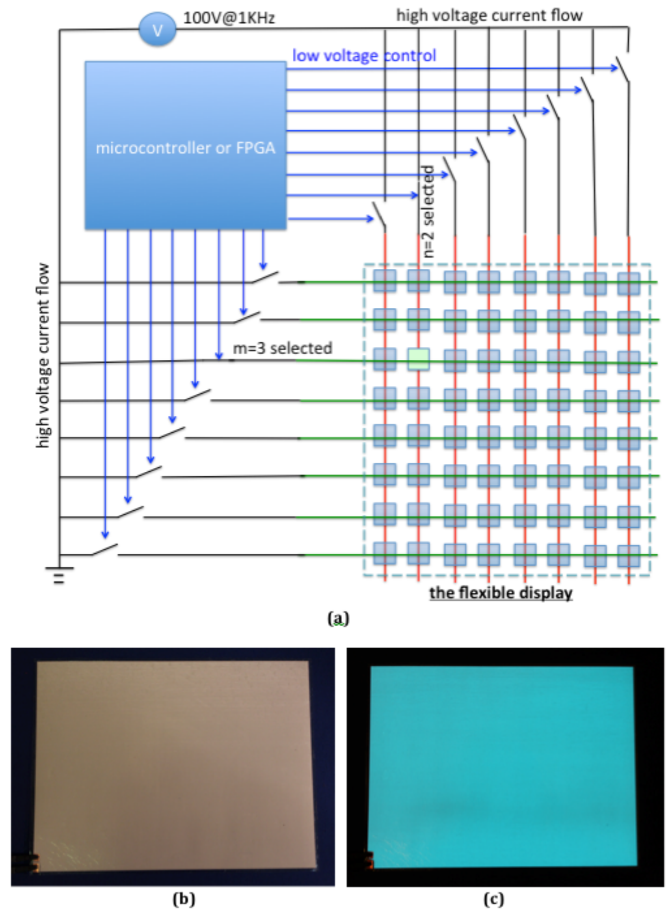


Figure 1. System structure: (a) blog diagram where pixel(2,3) is selected, (b) unlighted raw material, (c) lighted raw material

current is applied. The material we are currently using is designed by SEFAR AG, which should be driven by AC voltage of $100 \sim 110\ V, 400 \sim 1500\ Hz$. Overall lifetime of the material varies between $10.000 - 15.000$ hours. Depending on the brightness level and operating frequency. Moreover, it features very low power consumption ($0.1\,W$ for $1\,cm^2$ area), and can be easily bent and twisted without any damage. The brightness of the material can be adjusted either by changing the scan time or by adjusting input voltage of the inverter.

The structure we currently deployed is scalable, the control circuit grows with $n + m$ while the pixel number grows with $n \times m$. Pixel size for the complete matrix can vary depending on the required resolution. Practically, driving circuit can provide up to $150\ mA$ current for each pixel. In our case, the smallest pixel size that is manually possible to implement is $1 \times 1 mm$, and it can be increased up to $30 \times 20 cm$ depending on the resolution requirements.

In this paper we just demonstrate the idea with prototyping and we set the pixel number to be 8×8, which is already efficient to display a single letter, a digit or simple signs and the pixel size to $9 \times 9\ mm$. The limit on pixel number lies more in the electroluminescent material, which is lighted up only under certain voltage and frequency range. Also there exists ”light leakage” from lighted pixel to the adjacent pixels. The current screens are all made manually, that is to say, the pixels are cut out by hand and the connected and wired by hand. Manufacturing the screen by machine lies in the capability of SEFAR AG and is our next research step. This could enable higher pixel numbers and better performance. Trying out other materials, e.g. FLATLITE from E-Lite Technologies [2], is also included in our future plans.

APPLICATION SCENARIOS

To demonstrate variety of possible applications of the flexible screen, below we show 3 basic scenarios:

1. On a bottle, where the surface is curved.
2. On a window, where weather information is provided while keeping see-through function still.
3. On sport mat, which can be used to give support trainees during an exercise and later easily be rolled up together with the mat without and damage.

On a curved surface

In Fig. 2 we display digits on a curved surface. For simplicity, natural digits $(0 - 9)$ are chosen. The purpose of this experiment was to print temperature of liquid inside of the bottle directly on its surface. To better demonstrate the possibility of equipping irregular surface, we light up the first roll at bottle neck, too. Further interesting information can include: liquid’s chemical composition, brand name, important information on how to use the bottle’s content etc. All the enumerated options are extra and can easily be done by upgrading the circuitry and adding special sensors. A video is available at `http://youtu.be/SVHoxsPJVKo` where we display digits. This scenario can be expanded to other containers with complex shapes (e.g. cup, bowl, kettle, flowerpot), complex surface (sofa chair, pillar, vacuum cleaner) and of course regular surfaces (e.g. desktop, cabinet, washing machine).

Weather on a window

In Fig. 3 we attach the screen to a glass window facing to street. We print symbols for two different weather information on it: cloudy or clear weather. Pixels are a bit shifted, so that user can read the screen plus have a transparent partial view of the outside. Some basic background information (e.g. outside temperature, ambient light level, humidity) can also be added easily.

Sport signs on a mat

With this application we want to support people when they are training on a sport mat, both the newbies who need technical direction on training routine or guideline on how to perform the activity correctly, and by continuously monitoring the trainees, training counts, goals and burned calories etc. data can be printed on the mat directly. Since the display is flat and foldable, it can be integrated into normal sport mat and later just be folded or rolled up. As a simple demo (Fig. 4) we print the signs of a person seated with arm raised up, standing straight and doing push-ups. A Further step would be to combine resistive pressure matrix and our flexible screen to make a smart training mat which recognizes the sport activities and gives feedback to the user at the same time without any external disturbance.

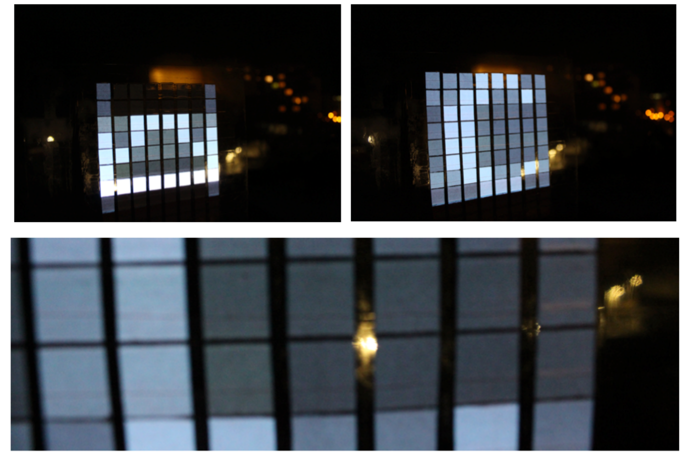


Figure 3. Weather information on a window: cloudy (cloud is displayed) and clear night (a crescent), and the streetlight visible between pixels

Figure 4. A sport mat with 3 posture signs: sitting, standing, doing push-ups

Other possible scenarios

The advantage of our screen against others are the combination of flatness, flexibility, the area it can cover and cost of the material. Possible applications can be more:

1) Outside of the buildings: at night the wall of most of the buildings are dark or just partly lit up by some big screens. To display a big picture, which is visible from far-away, requires big area while the pixel resolution doesn’t have to be very high. To give an example, all buildings around a square where a outdoor new year party is held, can be quickly covered up with such screens and uncover the next day. The rolled up screens can then be transported to somewhere else and re-used again.

2) As a window blind: thickness of the EL material allows to make window blinds which can print some information, such

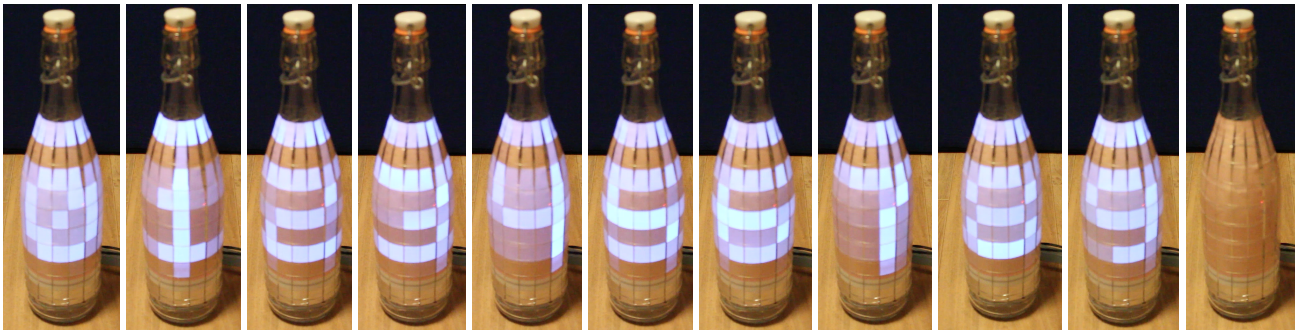


Figure 2. Digits $0 - 9$ **are displayed on a bottle (plus off-state)**

as clock, room temperature, humidity, etc. in forms of text and symbols.

3) For fans: nowadays fans of football or music stars carry with them glow stick, scarf, or portable LED displays to show their love and share their feelings. Our screen can be folded, carried in the pocket and later spread out, providing fans the possibility to express more. Similar application can be foreseen by parade, police evacuating people or demonstrations and so on.

CONCLUSION

We propose a combination of matrix driving circuit and electroluminescent material based screen, which enables easy prototyping of scalable and flexible display for interaction with smart environment and objects. We describe the hardware structure and discuss working principle of the screen comprehensibly. Three basic application scenarios are carried out to prove the idea of flexibility and to demonstrate the possibility of covering different areas of human surroundings with it, which hopefully will raise new interaction methods.

Acknowledgments

This work was partially supported by the collaborative project SimpleSkin under contract with the European Commission (#323849) in the FP7 FET Open framework. The support is gratefully acknowledged.

REFERENCES

1. Chwang, A. B., Rothman, M. A., Mao, S. Y., Hewitt, R. H., Weaver, M. S., Silvernail, J. A., Rajan, K., Hack, M., Brown, J. J., Chu, X., Moro, L., Krajewski, T., and Rutherford, N. Thin film encapsulated flexible organic electroluminescent displays. *Applied Physics Letters 83*, 3 (2003), 413–415.
2. FLATLITE, E-lite Technologies Inc. - official website. `http://www.e-lite.com/catalog/`.
3. Fukuda, Y., Watanabe, T., Wakimoto, T., Miyaguchi, S., and Tsuchida, M. An organic led display exhibiting pure rgb colors. *Synthetic metals 111* (2000), 1–6.
4. Kanis, M., Winters, N., Agamanolis, S., Gavin, A., and Cullinan, C. Toward wearable social networking with iband. In *CHI'05 extended abstracts on Human factors in computing systems*, ACM (2005), 1521–1524.
5. Kazlas, P., Au, J., Geramita, K., Gates, H., Steiner, M., Honeyman, C., Drzaic, P., Schleupen, K., Wisnieff, B., Horton, R., et al. 12.1: 12.1" svga microencapsulated electrophoretic active matrix display for information appliances. In *SID Symposium Digest of Technical Papers*, vol. 32, Wiley Online Library (2001), 152–155.
6. Laibowitz, M., and Paradiso, J. A. The uber-badge, a versatile platform at the juncture between wearable and social computing. *Advances in Pervasive Computing, Oesterreichische Computer Gesellschaft 363* (2004), 368–373.
7. Morokawa, S. Method of driving liquid crystal matrix display, July 7 1987. US Patent 4,679,043.
8. Youm Display - official website. `http://www.youmdisplay.com/`.
9. Zhou, B., Cheng, J., Sundholm, M., and Lukowicz, P. From smart clothing to smart table cloth: Design and implementation of a large scale, textile pressure matrix sensor. In *Springer Lecture Notes on Computer Science (LNCS) series (accepted)*, ARCS 2014 - International Conference on Architecture of Computing Systems (2014).

Smart Objects in Accessible Warehouses for the Visually Impaired

Tobias Grosse-Puppendahl[1], Justus Weiss[1], Pia Weiss[1], Sebastian Herber[1], and Hans-Jörg Lienert[2]
[1]Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany
{tobias.grosse-puppendahl, justus.weiss, pia.weiss, sebastian.herber}@igd.fraunhofer.de
[2]Dräger & Lienert Informationsmangement, Marburg, Germany
hlienert@dlinfo.de

ABSTRACT
The inclusion of persons with handicaps in working life is of increasing importance as equality for both disabled and non-disabled persons is a long-term goal in modern societies. Regardless of physical disabilities, the pursuit of a career suited to personal aptitude is an important concern of every individual. The demand for inclusive design and accessible workplaces is thus great, especially in consideration of the aging European society. In this work-in-progress paper, we investigate the use of smart objects for workplaces of visually handicapped persons. Based on our observations with a commissioning system in production use, we developed a concept for an accessible warehouse by integrating wireless sensor nodes into compartments, boxes and objects. We intend to equip these smart objects with various sensory and actuating technologies, which leverage a time-efficient and easily accessible commissioning process in a warehouse.

Author Keywords
Smart objects, Low-Energy Wireless Communications, Capacitive Sensing

ACM Classification Keywords
H.5.2 User Interfaces: *User-centered design*

General Terms
Human Factors; Design;

INTRODUCTION
With a total of five percent of the German population being visually handicapped or blind [1], we face the challenge of enabling them to be part of the workforce without greater impact on their efficiency in their respective fields of work. While most modern workplaces are equipped with helpful electronic devices, people with low vision are faced with using drawers, file cabinets and paperwork. Nowadays, pervasive technology helps and enables the visually handicapped to use office equipment just as efficiently as people without disabilities. These components have to support different target groups, ranging from kids to elderly people. One has to keep in mind the variety of causes for visual impairment as well - reaching from genetic defects or infection in utero to external influences like injuries. Moreover, in consideration to the age structure of our society, intuitive access and ergonomic design should be key guidelines in designing the workplace of the future. Therefore, the design of such accessible technologies is a crucial and highly demanding task.

In this work, we investigate the use of smart objects in a workplace of a visually impaired person. In particular, we present a concept to equip a warehouse with smart compartments, which provide acoustic feedback for easy and time-efficient localization. In the following, sensing and actuating techniques within the compartments provide naturally accessible functionality for managing object quantities and inserting new items into the warehouse.

BACKGROUND & RELATED WORK
Due to recent advancements in wireless communications, many smart objects and devices pervaded our everyday life. For example, the item finder “Chipolo” uses Bluetooth technology to locate items reliably without requiring a lot of space or high-level hardware [2]. The project employs low-energy Bluetooth receivers to locate items based on their signal strength, using a mobile phone. In consideration to our target group of vision-impaired persons, tangible objects provide an easy and intuitive way of symbolically accessing computer functions [3], [4]. In order to communicate the state of a system, acoustic [5] and haptic [6], [7] feedback techniques are applied. These use two other human senses with both high information density and low reaction time with easy implementation on a technical level.

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel

Figure 1. In this commissioning system, each compartment is equipped with an RFID tag. A text-to-speech program provides hints on finding the right compartment with the scanner.

CURRENT RFID-BASED IMPLEMENTATION

The idea for a new commissioning system based on smart objects was developed during a brainstorming session with a company developing accessible workplaces for the visually handicapped. The company employs two experts in interaction design, both having a low vision. In this section, we will outline the company's current experiences with a warehouse system based on RFID and a scanner. Figure 1 shows an example of the current commissioning system. Each compartment and object within the compartments is equipped with a small RFID tag. An accessible user interface runs on the PC which connects the Tags to a database, containing information on the current compartment contents, the position of the compartment, and the quantity of items inside one. When the user wants to locate an item within the warehouse, the person is able to search the computer for the desired item. Once the item is found, an iterative finding process is initiated.

The finding process is depicted in Figure 2. After a successful request, a cabled or wireless RFID reader is used to locate the right compartment. The starting point is an arbitrary compartment, which is scanned. Subsequently, the PC outputs voice commands for a refined search (e.g. "move down four compartments"). The search is continued until the desired compartment is located, as depicted in step 4 of Figure 2. The user is able to remove the desired objects from the compartment and feed the PC with the updated object quantity.

ENVISIONED CONCEPT

Requirements

Based on experiences with the RFID system in production use, we aim at achieving a number of enhancements. Specifically, our goal is to make the commissioning process less time-consuming for the user. Moreover, it is desirable to simplify the process, with peripheral devices like an RFID reader no longer required. The sources of error are to be minimized as well.

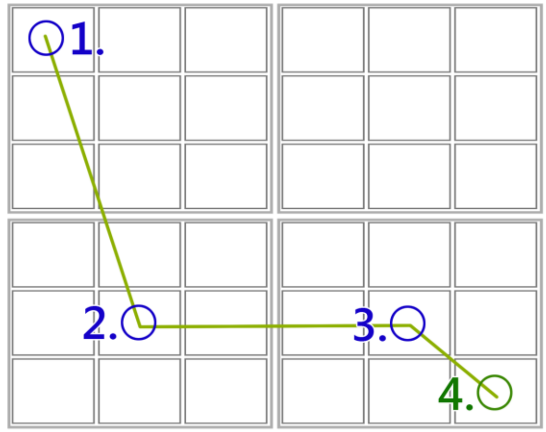


Figure 2. Locating the right compartment is an iterative process. First, the user starts with an arbitrary compartment, then the search is refined by voice commands from the user's PC.

Moreover, the following factors were taken into account when designing the system based on smart objects [8]:

- Natural Mappings: As far as possible, interaction itself shall be initiated by the objects that are naturally involved in the commissioning process. In this case, these objects are the compartments and objects located within.
- Design for Error: Errors are likely to occur, therefore, it is always important to acknowledge the fact that users will make mistakes. Therefore, it might occur that a user selects the wrong compartment or the wrong object. Moreover, he or she might not update the quantity of objects added or removed to the compartment.
- Providing Feedback with Multiple Senses: In order to help users finding the corresponding object, we can rely on other human senses such as touch (for haptic feedback) and hearing. For example, it makes sense for smart objects to send acoustic feedback.

Workflow Design

In order to allow for an efficient commissioning process, we designed a number of different workflows for (1) finding an item and updating its quantity, (2) inserting new items into a compartment, (3) maintenance and changing the batteries.

1. Finding an Item: One of the most obvious use-cases is the search for items within compartments and updating object quantities in the stock database. This begins with a user search for an item within the database. Once the user has found the desired item, the compartment containing the object sends out an acoustic signal which will guide the user towards its location. When the user is within reach of the compartment (approximately 15 cm), it becomes difficult to differentiate the compartment's exact location. Therefore, the compartment will change its acoustic output frequency according to hand distance: With acoustic feedback, selection of the right item is effortless. When an item is removed, it is necessary to update the changed number of items in the database. There are two possible options: (1) When the box is empty, the user may signal this by turning it over or (2) if there are still items left in the box, the user would lightly tap on the box's surface and thus indicate how many items were removed. This workflow uses acoustic feedback extensively during selection, but can also alert the user if a wrong item is selected.

2. Inserting New Items: Here, a user may add new items into arbitrary compartments. In this case, the person first uses the PC and manually enters the name of the item. After, the user selects an arbitrary compartment and taps on it multiple times, depending on the number of items added. Again, acoustic feedback is used to signal how many items are added and if the process was successful.

3. Energy Management: Even though we are not delving into the technical implementation yet, it is obvious that the compartments need a source of energy, e.g. batteries. Depending on how frequently it is used, the power consumption will vary strongly. However, it is also necessary to alert the user when battery power is low. Therefore, the compartments will send heartbeat messages in regular long intervals about their current energy status. When a low battery status is detected, the user will be informed of it. Consequently, the user may locate the corresponding compartment through acoustic clues, similar to the first workflow presented.

Technical Realization

The technical realization depends on a number of key constraints presented in the following. Figure 3 shows an overview of the technical components transforming the compartment into a smart object which can identify usage and is able to communicate with a PC. An essential factor for the successful implementation of the system is the battery runtime of the smart objects. An average battery runtime of one year for normal usage is an important milestone for the project.

A requirement is wireless communication between smart compartments and the PC. Different technologies were evaluated, each with its own advantages and drawbacks. Energy consumption needed for wireless communications represents a substantial amount of the total power consumption. The power consumption also depends on the communication distance to be achieved. In our scenario, this communication distance is between 5-20m within a building. However, this distance depends on the size of the warehouse and if more than one room is involved. We took into account many technologies, such as Bluetooth Low-Energy, ZigBee and proprietary communication methods in the Sub-GHz-domain.

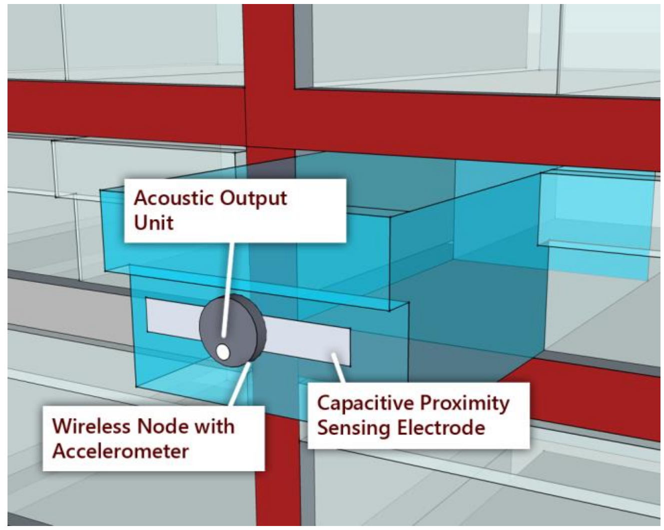


Figure 3. The smart compartment will be equipped with a wireless node which is connected to an accelerometer and a capacitive proximity sensing electrode. It may also use an acoustic output unit to generate polyphone sounds.

We came to the conclusion that a single-chip solution, combining microcontroller and RF-functionality is most feasible for our use-case. For this, we selected an ultra-low power micro-controller with RF functionality by Texas Instruments[1] (CC430F5137). The RF functionality also supports sophisticated Wake-on-Request handling which will be an important feature for energy saving. Therefore, when a compartment communicates with the PC, it is woken up by a request message from it. The compartment itself will only be listening to the incoming message for a few milliseconds and then go back to hibernation. These stand-by times will be adjusted to the frequency of usage of the system and can range from 5 seconds to a few minutes. However, this feature will decrease the reaction time, and a suitable trade-off must be found in the upcoming user studies.

In order to detect object manipulations, such as taps on the compartment's surface and the proximity to a user's hand, two different sensors will be integrated. An acceleration sensor (ADXL345)[2] will measure the compartment's accelerations in three axes. Moreover, it can automatically detect taps and double taps. The micro-controller can also be woken up by the accelerometer when activity is detected.

Furthermore, we integrated a capacitive proximity sensor from the OpenCapSense project [9]. The proximity sensor conducts a self-capacitance measurement by loading and unloading the capacitance between the environment and the

[1] http://www.ti.com/product/cc430f5137
(date accessed: 13-Dec-2013)

[2] http://www.analog.com/static/imported-files/data_sheets/ADXL345.pdf
(date accessed: 13-Dec-2013)

sensing electrode. Depending on the sensing electrode's size, the sensor may recognize the proximity to a human hand for distances up to 35 cm. Currently, we are aiming at recognition distances of approximately 15-20 cm, which will be a more reasonable goal to achieve.

We have not yet made a decision on the acoustic output unit, aiming at maximum energy-efficiency and easy controls. Therefore, different frequencies have to be supported – e.g. by providing a pulse-width-modulated output signal on the microcontroller.

Current Work-in-Progress

As described in the sections above, the workflows have been specified working closely with future system users. Moreover, suitable components were already selected, which are required for realization of the smart compartments.

In the technical development process we use CC430 evaluation boards to develop a communication and broadcasting mechanism. Simultaneously, we implemented bindings for the capacitive sensor and the accelerometer. We plan on conducting the first experiments on energy savings once the wireless communication concept has been successfully implemented. Hereafter, we will move to the final PCB design. The size of the final node and the housing will also greatly depend on the battery required for the desired runtime of one year.

The next steps will include creating interfaces for the warehouse's database with an accessible front-end. Due to this, we plan to integrate an additional component into the current RFID-based commissioning system.

CONCLUSION

In this paper, we presented an approach for the realization of an accessible warehouse for vision-impaired persons using smart objects. Specifically, we presented the concept of smart compartments, aware of physical manipulation and with wireless communication abilities. We introduced a concept which applies various means of natural interaction to make the commissioning process easier and more time-efficient.

Currently, we are at a vital stage of the process. It is not yet resolved if the implementation of the desired features with the specified requirements will comply with our goals, among them are low energy consumption and low maintenance effort. Therefore, it might be necessary to change certain components or revise the concept. Moreover, it is also necessary to find a suitable trade-off between the system's reaction time whilst locating a compartment. Thus, we plan on developing a predictive algorithm which schedules the microcontroller's sleep times intelligently.

After completion of the system, extensive user studies will reveal if our ambitious goals on simplifying the commissioning process can be met. However, our solution may also be transferred onto other use-cases with which smart objects can support handicapped persons in their personal or work life.

ACKNOWLEDGMENTS

We would like to thank the students Oskar Bechtold, Sebastian Schurig, Lukas Strassel, Stefan Wegener (TU Darmstadt), and Frauke Taplik (Hochschule Offenbach) for their effort spent on the system's interaction concept and software development.

REFERENCES

[1] Statistisches Bundesamt, "7,3 Millionen schwerbehinderte Menschen leben in Deutschland (In German)," *www.destatis.de*, no. September, 2012.

[2] Chipolo, "Chipolo: Nothing Is Lost," 2013. [Online]. Available: http://www.chipolo.net/. [Accessed: 13-Dec-2013].

[3] E. Riedenklau, T. Hermann, and H. Ritter, "Tangible Active Objects and Interactive Sonification as a Scatter Plot Alternative for the Visually Impaired," in *International Conference on Auditory Display*, 2010.

[4] D. McGookin, E. Robertson, and S. Brewster, "Clutching at straws: using tangible interaction to provide non-visual access to graphs," in *CHI '10*, 2010, pp. 1715–1724.

[5] Y. Liu, J. Bacon, and R. Wilson-Hinds, "On Smart-Care Services: Studies of Visually Impaired Users in Living Contexts," *First Int. Conf. Digit. Soc.*, pp. 32–32, 2007.

[6] C. Avizzano and S. Marcheschi, "A multi-finger haptic interface for visually impaired people," *Robot Hum. …*, pp. 165–170, 2003.

[7] K. MacLean, "Putting haptics into the ambience," *Haptics, IEEE Trans.*, vol. 2, no. 3, pp. 123–135, 2009.

[8] D. A. Norman, *The Design of Everyday Things*. 2002.

[9] T. Grosse-Puppendahl, Y. Berghoefer, A. Braun, R. Wimmer, and A. Kuijper, "OpenCapSense : A Rapid Prototyping Toolkit for Pervasive Interaction Using Capacitive Sensing," in *IEEE International Conference on Pervasive Computing and Communications*, 2013.

On Aiding Supervision of Groups in the Mobile Context

Daniel Auferbauer
Vienna University of Technology
Favoritenstrasse 9-11/187, 1040 Vienna Austria
daniel.auferbauer@gmail.com

Hilda Tellioğlu
Vienna University of Technology
Favoritenstrasse 9-11/187, 1040 Vienna Austria
hilda.tellioglu@tuwien.ac.at

ABSTRACT
In this work we introduce and examine the possibility of aiding in the oversight of mobile groups by assisting the supervisor in his or her awareness of the physical presence of members. The subject matter of this paper is to find out whether or not that is viable, and why. Our approach is thus: we have lead interviews with users representative of the target audience in order to gather information on group supervision and define requirements. Secondly we have assessed five wireless technologies for use in an actual implementation. As a third step we have engineered an actual prototype based on the information gathered thus far. Lastly, this device was evaluated both under laboratory conditions and in the field. We find high acceptance and demand among prospect users and conclude from evaluation that there are strong indications to the viability of reducing the workload of supervising mobile groups by assisting the person in charge with awareness of physical presence of members.

Author Keywords
Assistance; group; supervision; mobile; presence awareness; assisted supervision

ACM Classification Keywords
H.5.2 user interfaces: user-centered design

General Terms
Experimentation; human factors; verification

INTRODUCTION
It is easily observable that managing a group of people is difficult, especially in the mobile context. The person in charge has to ensure that all members transfer safely to the destination, without anyone getting lost or being hurt. Instantiations of this scenario are, for example, a school class on a field trip or a tourist group on its way through a foreign city. The supervisor of such a group usually has to do some heavy multitasking to assure that all of his or her charges stay out of harms way while navigating, e.g., an urban area. That being established, the idea is that we may aid a group supervisor by providing electronic assistance to the task of checking for absentees, thus reducing the number of things he or she has to keep in mind simultaneously. We seek to investigate this possibility of improving both working conditions and safety. The challenges, as far as human-computer-interaction is concerned, of providing such electronic assistance are to design a mobile device that provides an interface that is usable even in highly stressful situations. It also provides exactly the information needed to assist with group supervision. The high cognitive load and fragmented nature of the user's attention span in such demanding environments is discussed in e.g. [8]. In order to evaluate the prospect of assisted group supervision, we work together with a group of people deemed representative of the target audience: primary school teachers. They provide know-how on best-practices as well as known problems and serve as test users for the prototype's evaluation.

The goal of this paper is to investigate the viability of assisting group supervisors by aiding their awareness of physical presence. "Viability" in this context refers to both technological feasibility (meaning whether it is possible to implement at all) as well as usefulness in the intended setting (meaning whether it indeed has potential to support the user). Herein, we provide some primary research on this. To the best of our knowledge, we are first to explore this avenue of assisting group supervisors; at this point we know of no prior publications in academia on that topic.

Our approach to examining the subject matter involved four steps: Leading interviews with users experienced in supervising groups, doing an assessment of available wireless technologies, implementing a prototype, and finally the evaluation thereof. Each of these steps and its results will be described in turn before we discuss the conclusions drawn from the collected data at the end of this paper. In the conclusion, we will focus on drawing up design guidelines for systems such as the one proposed in this article.

PRELIMINARY INTERVIEWS
"The first step in the usability process is to study the intended users and use of the product" [7, p. 73]. Therefore, prior to designing and implementing a prototype, we went out and talked to users who have previous experience with supervising groups of people on the move: primary school teachers. As wresting control over large groups of sometimes unruly individuals is their source of daily income we expected them to provide valuable feedback based on experience. There were several goals to these interviews: firstly we wanted to gauge reactions to the idea of electronic assistance to group supervision; we considered outright rejection to be possible and wanted to know the reasons why, if this was indeed the case. Secondly, we wanted to learn about typical problems encoun-

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel.

tered when supervising groups, which situations aggravate those and what best practices currently are employed to mitigate this. Thirdly, we tried to find out which features our electronic assistance would have to provide to make it useful. In total there were six participants ranging in age from their early twenties to fifties, the average being in the mid-thirties. The interviews were done at the participants' work spaces, i.e., the rooms of their respective classes, to provide a familiar setting and relaxed atmosphere. The talks were held in a semi-structured manner: a small number of key points with relation to the topics mentioned above was always addressed, but free talking was encouraged at all times.

Initial reception of the idea was very favourable with remarkably little reservation against an "intrusion" by electronic assistance. We identified a core set of requirements, which were mentioned by all participants. These are rather basic and consist of: detecting complete absence of any group member, displaying names of absentees and providing an easily recognisable, auditive alarm signal. Desired features mentioned by at least three of the participants were: setting up and managing a list of members currently with the group prior to venturing out, being able to chose the maximum distance, which the group member are allowed to venture from their supervisor, and sounding an alarm not only if a group member is completely separated from the group but also if they leave a certain "safe zone" (meaning, effectively, an early warning system). Furthermore, exact positioning of all group members was called for, put not considered vital.

We were able to classify three different types of situations during which group supervision is difficult, with an example of each of these being mentioned by at least three of the participants. Firstly, there are such scenarios as require moving past a certain point under time constraints; named examples include boarding or exiting a public transport vehicle or crossing a street. Secondly, there are situations where the group remains in an area with wide boundaries for a prolonged period, such that the supervisor may lose visual contact with members; an example for this is staying on a playground. Lastly, there was mention of crossing highly frequented areas, where the flow of people may disrupt the group's structure; think of, e.g., moving through a subway station.

ASSESSMENT OF TECHNOLOGIES

To find a wireless technology that fits the needs of a prototype system we have compared five wireless technologies currently available: Wi-Fi, Bluetooth, GPS, ZigBee, and Ultra-Wideband. The criteria for this assessment were maximum range of communication, reliability (resilience to interference), energy consumption as well as availability (accessibility on the consumer market). With these in mind, *ZigBee* was chosen for use in the prototype. It is a standard that builds additional layers on top of the IEEE 802.15.4 specifications and is concerned with the setup and routing of multi-hop networks [2]. ZigBee defines three different device types, which enable the formation of various topologies; most importantly and relevant for this work it is able to form mesh networks. Put simply, every node of a certain type is in theory able to expand the network. As an in-depth description of ZigBee is

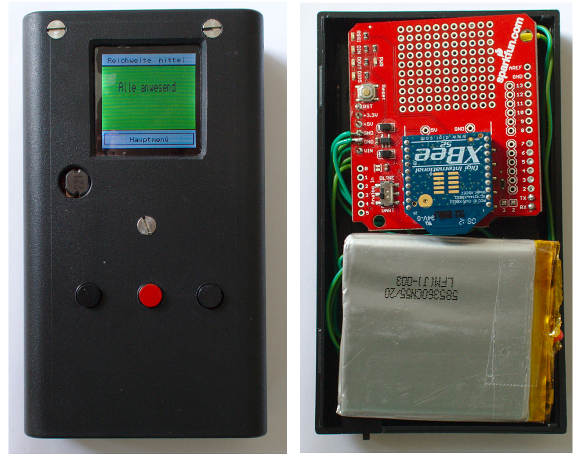


Figure 1. Supervisor's prototype (top-down view).

beyond this paper, please refer to other literature like e.g. [2, 3]. Mesh networking is one of the reasons why we chose to use ZigBee. It allows to mitigate range restrictions imposed by a star topology and provides robustness due to multiple available paths [1, 4]. Also working in favor of this technology is its low energy consumption as compared to other wireless technologies [5, 6].

It should be noted that Ultra-Wideband does, on paper at least, provide better transmission characteristics than ZigBee for our purposes, featuring good range and robustness paired with low energy consumption. Our reason for not using it was the lack of availability in prototyping platforms or other commercial products rather than any obvious technological shortcomings. Bluetooth and Wi-Fi on the other hand differ too much in their intended application profile (high data throughput vs. energy consumption and mobility) to recommend themselves for this use case. GPS was discarded due to its also high energy consumption in combination with outages when indoors or even in urban canyons [9].

PROTOTYPE

To evaluate the prospect of assisting with group supervision we have constructed a prototype based on the information gathered during the preliminary interviews. It consists of one main device intended for the group supervisor (Figure 1) as well as five different wearable devices for group members. The latter include high visibility vests, a pendant, and two kinds of headwear. The more successful designs (refer to next Section) may be seen in Figures 2 and 3.

Figure 2. Pendant item for group members.

Figure 3. High visibility vest for group members.

The prototype is based on the Arduino platform, with its LilyPad variant being used for the wearable group member devices. They communicate with each other using Arduino's XBee wireless modules, which provide an implementation of the ZigBee standard, including mesh networking capabilities. The way it works is that the main device does periodic checks on the difference between the current time and the last time a device has reported in. If this delta crosses a certain threshold the supervisor is alarmed of that specific group member being missing. This is enough to implement the core functionality as identified during the preliminary interviews. Further discussion of the exact workings and configuration of the prototype are beyond this paper.

The prototype's interface is kept rather simple on purpose. The main screen is visible in Figure 1. It shows a green background and a text saying that everyone is present (if this is indeed the case). Should contact be lost with one of the other devices, the screen turns red and displays the name of the person in question. In addition, an audible alarm is sounded which may be silenced by the press of a button. No other interaction is required during regular operations. There is a menu structure, accessible from the main screen, which allows adding and removing of group members and provides a few settings (e.g., for switching off sound); this menu structure does not necessarily have to be accessed while the group is moving. Care has been taken to always display important information on the main screen, where it is quickly accessible. All interactions are carried out via three hardware buttons located under the display. The left and right button are used when navigating menus to move the current selection up and down, respectively. The center button (red) is used for confirmations such as acknowledging an alarm or entering a menu.

Due to the highly stressful environment that the user will be operating in frequently, we recommend to keep a focus on prominently and concisely displaying important information where the user may access it easily; assume that he or she will not have more than a few moments time to interact with the device. We have deliberately decided to forgo the use of touch-based input in favor of hardware buttons. The decision was in part due to remarks from participants during the preliminary interviews, where they stated to prefer hardware button for their robustness and resilience to dirt and water; this is a view that we agree with. Independently of these opinions we have preferred hardware buttons because they provide tactile feedback and do not require visual contact to be maintained with the screen during interaction, both of which are desirable characteristics for situations with many distractions. Ultimately, omitting touch-based input has served us well in this specific use case and was received positively.

EVALUATION

During the last phase of our work, the prototype was evaluated both under laboratory conditions and in the field. For test users we remained with the primary school teachers that we had already talked to during the preliminary interviews. We hoped that the users being familiar with our work would make them feel more tied into the endeavor and thus improve the quality of feedback.

Before venturing out into the field, the prototype carried by the group supervisor was first evaluated in a controlled environment. The "controlled environment" in this case was again the class rooms of the participating teachers, same as during the preliminary interviews. Four of the six teachers already involved through the earlier interviews were chosen as test users. Their age was deliberately kept diverse with the youngest participant being in her early twenties and the oldest just turning fifty. These tests revealed only minor flaws in the structure of the device's menus, but indicated its readiness for field testing otherwise. We would like to concentrate on the latter in this paper, as it is more relevant to answering the question of viability for practical use.

Field Testing

Having assured that the device intended for the supervisor is ready for field testing as far as functionality and usability are concerned, we conducted evaluation in a realistic environment. To this end we have accompanied a teacher and her class during field trips on two occasions. They would use the public transport system to get to a nearby park and playground for a stay of two to three hours before returning to school. We considered this a good context for evaluation, as it would include all three of the problematic situations we identified earlier (see section "Preliminary Interviews"). The whole class consisted of 25 children, five of which had equipped wearable prototypes during the trip; you can see them being worn on Figure 4. Having only five items available, we were not able outfit all of the class with prototypes (refer to "Caveat" at the end of this article); instead they were rotated between different pupils on a regular basis so we could observe a good number of wearers. The teacher served as test user for the main device; she was the oldest person of the demographic participating in evaluation (being 51 years of age) and already familiar with the project, having taken part in the earlier interviews and usability tests.

The field tests have provided additional insights into the challenges of supervising groups in the mobile context as well as useful observations regarding the prototype and its usage. Specific to the use case of school classes (or groups of children in general) is the insight that there was a remarkably high acceptance among them as far as the prototypes were concerned. Even after the novelty factor had worn off, the behavior of the children wearing one of the items ranged from unconcerned to being proud of their adornments. Their usual behavior on the playgrounds and in transit did not seem to be affected at all, an observation that was confirmed by the teachers; we consider the minimal impact of our devices to be a positive characteristic. Also mostly (yet not exclusively) applicable to the context of younger target audiences is the finding that of the different types of items being tested, the high visibility vests proved to be the most practical. They were preferred by the majority of children and also liked by the teachers due to their high visibility. They may also be worn as the top layer of clothing independent of other garments and outside temperatures, which makes them the most versatile.

Observing the teacher using the prototype showed how important the use of concise presentation of information is. Between being riddled with questions by some of the children, keeping an eye on the others and making sure the group catches the next bus or train, she barely had time to glance at the display once an alarm went off before having her attention drawn elsewhere again; this short attention span is in accordance with the findings of [8]. For this reason, our prototype displays the current group status prominently and easily accessible on the starting screen. In the concluding talks the teacher stated using the device did not encumber her in the usual routines and activities related to group supervision. This is as much a positive indication for the device's usability as we can hope for with regards to the current state of the prototype – please refer to section "Caveat" for further remarks on this.

As for the technological aspects of the prototype, the performance of Arduino's own implementation of ZigBee was acceptable. Due to the rather low-power transmissions usually employed by ZigBee radios, some false alarms were triggered in environments that were rich with interferences, such as crowded trains and buses. Such problems were encountered in three out of four rides with public transports. These issues were not encountered in outside areas; the system worked reliable both while the group was mobile or staying within a designated area (e.g. a playground). However, false alarms are a serious problem as they undermine the user's trust in the system and lead to unnecessary commotion within the group as the teacher tries to determine where the "lost" members are. We suspect the number of devices may have been too low or the routing too inflexible to achieve positive effects from mesh networking in such environments and with regards to a rapidly shifting network topology. An implementation on a larger scale with more devices should be able to handle such situations better through the availability of additional routes to the main device. Further investigation on this topic is required to draw significant conclusions.

Figure 4. Children wearing the prototypes: cap, high visibility vest (twice), cap, and pendant, left to right.

CONCLUSIONS

In this paper we have presented our approach to assisted group management in the mobile context by aiding the supervisor in his or her awareness of absentees. We attempted this by providing a prototype that informs of the absence of group members. Acceptance among users representative of the target audience (teachers) was high and the idea was received favorably by all participants. Evaluation of a prototype has provided us with strong indications of the viability of this approach. Technologically the device has held up to expectations under most circumstances, providing a stable network where interferences are not too prominent; it is our opinion that a more stable performance can be achieved in the future. As far as the usefulness of the device is concerned, we have received positive feedback from our target audience after evaluation. We feel, thus, that strong indications are given for the viability of the avenue presented in this paper.

We have found the following guidelines to designing a mobile assistance system for assisting in group management: first and foremost, the system should be kept simple both in its user interface and functionality. In case of the interface, simplicity is necessitated by the environment the user operates in. People in charge of mobile groups are bound to be under stress and accordingly will have little time to spare for interactions. It is therefore important to always display relevant information prominently, concisely and easily accessible on the screen; furthermore interactions should be kept to a minimum and, if unavoidable, short in duration while the group is mobile. For these reasons, the prototype always displays the group status on the home screen and demands only a single interaction after setup (which is acknowledging an alarm by pressing any button). As was mentioned earlier, hardware buttons should be preferred for user input as they provide tactile cues and do not require visual contact to be maintained with the screen, as do touch-based input variants. Apart from the user interface, functionality should also be kept to a manageable extent in order for an assistance system to remain useful. The following functionality, provided by the prototype, was shown to be sufficient: group setup (i.e. declaring who is in the group) prior to moving out, sounding an alarm in case of missing group members and displaying their names. We urge to weigh the addition of more functionality against the risk of unnecessarily bloating the system, thereby making it more cumbersome to use.

Lastly we would like to point out the importance of keeping the presence awareness aid local. What we mean by that is to not utilize technologies which facilitate positioning of individual group members on a global scale, such as geolocation through GPS or GSM cells. Apart from the technological shortcoming of not being able to operate without certain infrastructure being present, these technologies are also likely to diminish acceptance of the system. Our society currently experiences a sensitization towards the issues of surveillance and tracking. The children themselves did not seem to mind the prospect, but their parents are likely to [10]. Likewise, we suspect skepticism towards these technologies from adult group members in other use cases. Indeed, multiple participants have also issued concerns on this topic during preliminary interviews and evaluation. Their fears were alleviated once we assured them that the prototype only tracks within a certain range around the supervisor, not globally. We are well aware of the added possibility that the usage of, e.g., GPS would bring, yet we urge to refrain from global tracking for the sake of acceptance among the prototype's target audience.

Caveat

There is currently no full-scale implementation of the prototype – meaning that there are only five child devices and one for the supervisor. Consequentially, evaluation in the field could not be done to an optimal extent. The test user did not have the option to rely on it as they were intended to with a more complete system. A complete system would feature at least thirty devices to be of use in the context of school outings. But with only a subset of pupils under assisted supervision, the teachers had to do their regular routines of counting heads and checking for absentees in addition to operating the prototype. It was thus hardly feasible to conduct measurements of cognitive load and stress or the reduction thereof. Ultimately this means that, while we feel strong indications are given for the usefulness of the prototype by the favorable responses of our test users, we cannot prove it with further data points yet.

REFERENCES

1. Baker, N. Zigbee and bluetooth strengths and weaknesses for industrial applications. *Computing Control Engineering Journal 16*, 2 (april-may 2005), 20–25.
2. Baronti, P., Pillai, P., Chook, V., Chessa, S., Gotta, A., and Fun Hu, Y. Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards. *Computer Communications 30*, 7 (2007), 1655–61695.
3. Farahani, S. *Zigbee wireless networks and transceivers*, 1st ed. Elsevier Ltd., 2008.
4. Gill, K., Yang, S., Yao, F., and Lu, X. A zigbee-based home automation system. *Consumer Electronics, IEEE Transactions on 55*, 2 (may 2009), 422–430.
5. Guo, W., Healy, W., and Zhou, M. Zigbee-wireless mesh networks for building automation and control. In *Networking, Sensing and Control (ICNSC), 2010 International Conference on* (april 2010), 731–736.
6. Lee, J., Su, Y., and Shen, C. A comparative study of wireless protocols: Bluetooth, uwb, zigbee, and wi-fi. In *Industrial Electronics Society, 2007. IECON 2007. 33rd Annual Conference of the IEEE* (nov. 2007), 46–51.
7. Nielsen, J. *Usability Engineering*. Interactive Technologies. Elsevier Science, 1993.
8. Oulasvirta, A., Tamminen, S., Roto, V., and Kuorelahti, J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, ACM (New York, NY, USA, 2005), 919–928.
9. Retscher, G., and Thienelt, M. NAVIO A Navigation and Guidance Service for Pedestrians. *Journal of Global Positioning Systems 3*, 1–2 (2004), 208–217.
10. Vasalou, A., Oostveen, A., and Joinson, A. A case study of non-adoption: the values of location tracking in the family. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW '12, ACM (New York, NY, USA, 2012), 779–788.

A Context Aware Music Player: A Tangible Approach

David Sellitsch
Vienna University of Technology
Favoritenstrasse 9-11/187, A-1040 Vienna
david.sellitsch@gmail.com

Hilda Tellioğlu
Vienna University of Technology
Favoritenstrasse 9-11/187, A-1040 Vienna
hilda.tellioglu@tuwien.ac.at

ABSTRACT
In this paper we explore new ways of interacting to configure and personalize music selection in ambient environments. We propose a prototype for a context aware music player and a novel interaction concept to deal with it. Context information refers in this work both to the user, especially the mood situation, and the environment. The interaction concept lets the user express information about the mood and the current activity in a subtle way and customize the system to music preferences. This is achieved by using sensors to capture the environment data and a tangible user interface to enter and modify the context information related to the user. In usability tests of the prototype and in analysis of the interviews we conducted with test users we found out that customization options and making autonomous decisions transparent are two key factors to enhance user experience in context aware music systems.

Author Keywords
Context awareness; user-centred design; tangible user interface; sensor-based; mood.

ACM Classification Keywords
H.5.2 User Interfaces: Input devices and strategies, prototyping, user-centered design.

INTRODUCTION
Music is often consumed in the background, especially in ambient environments. People listen to music, while they read, workout, or while they are having dinner. Choosing the right music can be time consuming and distract users from what they wanted to do in the first place. Context aware music systems aim to free the user from choosing music and creating playlists. This is achieved by sensing the environment and gathering information about the user. Based on this information, music that fits the current situation is selected. The findings and figures in this paper are based on the master's thesis of the first author [12].

In the past, some research had the focus on specific parts of context aware music systems, such as music categorization [5] or reasoning [7, 9]. Most of the resulting prototypes are rather systems evaluated in lab environments and not ready to be applied by real users outside the research environment. They neglect design questions about the actual use of such systems.

In this paper we explore new ways of interacting to configure and personalize music selection in ambient environments. First we discuss the role of context information to capture certain data about the use environment. Then we briefly present mood models we used in our work before we introduce our approach and the resulting prototype. After the description of the tangible front end interaction interface of the system, we explain how we implemented the configuration and the automatic rating used in music selection. We conclude our paper and point out some future work.

CONTEXT AWARENESS
In real settings context information is very valuable to capture the current situation and circumstances of use. Context information can be gathered implicitly, by using sensors or explicitly by simply asking users. While implicitly captured information seems to work well for context information about the environment including light, temperature, and noise in the room, sensing the users' mood implicitly is a difficult task. Some of the existing research projects try to avoid user input at all cost. This leads to constructs where they even try to determine mood based on information like a users' stock portfolio [4].

The problem every context aware music system faces is music selection. This can be done based on audio-analysis and classification of songs. One problem that comes with automatic classification is that it neglects personal preferences. For example, one would think that uptempo music is the obvious choice for people while they workout. According to a 2012 survey [2] college students in the US prefer Hip Hop music, which typically has a rather low number of beats per minute, while they exercise. This shows that personal preferences play an essential role for music selection, and we therefore suggest that users need configuration options, which let them customize the system to their preferences.

Unlike Cunningham et al. [3], we believe that the users' mood cannot be determined based only on the information about the environment. It should also be seen as an additional contextual factor for music selection. Even if we try to sense a user's mood by making them wear a cumbersome brainwave detection sensor [8], we would still only have information about the current mood. Since people use music also for

IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel

mood transition, we would still need to know the users' desired mood.

MOOD

We believe that mood models for context aware music systems should be based on models and insights from the field of psychology. There are many different approaches and models, when it comes to moods and emotions. Some of them are even complex 3-dimensional models [10]. For our context aware music player we needed a model with moods that makes sense for music selection. Most people would find it hard to think about music that fits to the feeling of surprise or disgust. Han et. al [5] used a 2-dimensional mood model based on Thayer's insight about human emotion [13]. Since we wanted that users are able to easily make a connection between music mood descriptions, we decided to further simplify the model for our system, by keeping only extreme moods (see Figure1).

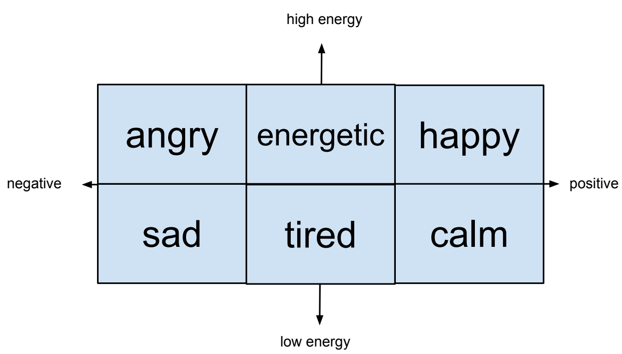


Figure 1. Simplified 2-dimensional mood model based on the work of Han et. al and Thayer.

APPROACH

We approached this research project with a user centered design process. After initial literature review, we started creating ad-hoc personas [1]. We then developed a first low-fidelity version of the prototype and evaluated it in a usability test with three participants, who resembled our identified personas. This version of the prototype featured the look and feel of the final product, and then we used the Wizard of Oz technique to mock its functionality.

Using the information we gathered from the evaluation, we implemented a second prototype. This second version is a fully functional integration prototype. We invited the three participants from the first test and three new participants for a second more detailed usability test. In this test users got specific tasks to solve, covering all features (including configuration). There was also a more realistic second part of the test, where users had to perform activities like cooking or reading a book while using the system. After the test we asked them some questions about the prototype in a semi-structured interview.

The usability test and the interview were video-recorded, in order to analyze them later in detail. The aim of these tests was to get detailed information about how participants use such a system, and how their thought process looks like while they use the system. This is why we used the thinking aloud technique during the tasks.

PROTOTYPE

The resulting prototype is a system that consists of a USB-device that functions as a tangible user interface, has a light sensor and a physical next-button attached to it, 9 single-sided activity cubes, 6 double-sided mood cubes, and a software application that runs on a Mac, gets weather information via a web-service, does all the reasoning, and plays back the selected music (see Figure 2). We used two Arduino Uno micro-controllers, equipped with two NFC-shields for identifying cubes and handling sensor input.

Figure 2. Setup of the system on a table.

Users can also access a graphical user interface via the computer, where they can customize the system. The idea is that once the system has been configured, there is no need to touch the computer again, unless the user wants to reconfigure it. As a source for music we used genre specific Internet radio channels, since they provide a wide variety of music, ranging from ambient music to death metal. The source for music can be easily substituted by local or online playlists, as long as the tracks featured in it have similar characteristics.

When the application gets launched, music starts playing immediately. When there is no user input, music gets selected purely based on environmental information. Users may then place up to two cubes on the tangible user interface.

There are two types of cubes. Activity cubes, which represent an activity, and show a symbol on just one side (white cubes in Figure 3). Mood cubes are double sided. They have a white side, which represents the current mood, and a black side, which represents the desired mood. Figure 3 shows both sides of mood cubes in the bottom right section.

The tangible user interface is a token and constraint interface [14]. The two cavities on the box are not only a constraint, but also cover exactly half of a cube. Thus, mood-cubes placed in one of the cavities appear as either black or white. Double sided mood cubes enable users to express mood transitions. For example, a transition from angry to calm may be expressed by placing the angry cube with the white side on top and the calm cube with the black side on top. When a new cube is placed on the device (cf., Figure 4) or an environment

Figure 3. Cubes used for expressing moods and activities.

factor (e.g., weather) changes, new ratings are calculated and usually new music starts playing.

We also implemented audio feedback for user input. When a cube is placed on the device a text-to-speech voice reads the input out loud. For the above given example this would mean that the system says "I am angry and I want to be calm". By repeating the user input, a dialogue between the system and the user is created. Users can thereby verify if what they wanted to express, was understood by the system. This feature has been added after the first usability test. It did not only help users prevent input errors, but also let new users identify symbols and learn how to use the system without even reading about the concept of current and desired mood on the provided info-graphic.

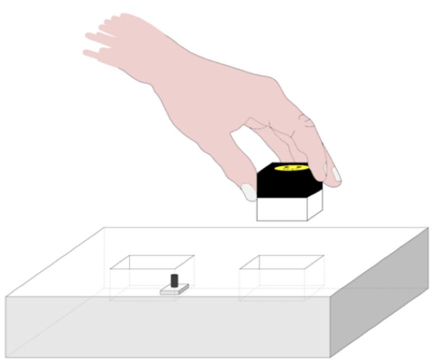


Figure 4. Expressing mood by interacting with double-sided cubes.

Users can always skip music they don't like, by either pressing the skip-button, which is the only physical button on the device, or by clapping twice. We experimented with clap detection software [6], because we wanted to give users the opportunity to skip music remotely in situations where they are not standing near the device, such as cooking or exercising. Users generally liked this feature, but the clap-detection did not always work properly. The clap detection software needs to be configured specifically for every user and one may also need to reconfigure it when using it in a different environment, since the acoustics change when the device is relocated.

CONFIGURATION

People describe their music preferences in sentences like "on a rainy evening, I like to read a book while listening to classical music" or "when I'm tired, house music gives me the energy to continue my workout". Our configuration approach tries to enable users to express it in a similar way by using tags. We translated sensor information into pre-defined tags, which users can assign to radio stations by simply clicking a button. Figure 5 shows a section of this user interface with activated context-tags for a lounge radio station.

Figure 5. Screenshot of a part of the configuration interface

Users can further customize the system by adding new activity cubes, new radio stations, and by changing the weighting of categories.

MUSIC SELECTION

In our prototype system music gets selected using a priority list based on a simple rating function (1). Basically ratings for radio stations ($rating_x$) score points for every context-tag that matches the current situation ($NMC_{x,n}$). In order to prevent music with a higher number of assigned tags from being ranked high in almost every situation, we subtract points for every non-matching context tag that is assigned to it ($-(TNC_{x,n} - NMC_{x,n}$). The complete rating is calculated by multiplying the score of each category with a weighting factor for that category and summing up all category scores. The weighting factors of the seven categories: activity, current mood, desired mood, weather, temperature, time, and light can also be changed by users using the configuration-GUI. This decision has proved to be right during the interviews, since every user stated that moods and activities are more important than other categories for their music selection.

$$rating_x = \sum_{n=1}^{7} WF_n \cdot (NMC_{x,n} \cdot rf - (TNC_{x,n} - NMC_{x,n}) \cdot pf) \quad (1)$$

Symbol	Name	Default-Value
rf	reward factor	5
pf	penalty factor	1
WF_n	weighting factor of category n	

Table 1. Constants

Sym.	Explanation
$NMC_{x,n}$	number of category-n matches with radio station x
$TNC_{x,n}$	total number of characteristics in category n assigned to x

Table 2. Variables

SOME FINDINGS

The prototype as well as the interaction concept proposed in this paper were accepted very well by the users. During the usability tests, we found out that when there is a lack of feedback, users tend use the system in a way it was not intended to by the designer. This is due to mismatching mental models of how the system works. When there was no audio feedback (as described above), people used mood cubes not to express their mood, but as a feedback mechanism, when they did not like the music that was playing (e.g., with a sad cube). One of the users stated that he wants an option to mute the audio-feedback, because he does not want everybody else in the room know, when he is in a bad mood.

One other interesting finding was that users who said that they felt in control of the system, also stated that they had a great experience using the system. Some users said that the tangible user interface gave them back some of the haptic experience they liked about CDs, but lost since they started listening to MP3s at home.

We tried to tackle the problem of awareness mismatch described by Schmidt [11] by making the music selection process transparent, by highlighting context tags of the presently playing radio station (on the computer screen) that are matching the current situation's context. Unfortunately users had problems recognizing changing highlighted tags, since changes usually happened while they interacted with the external USB-device, and therefore did not look at the screen.

CONCLUSION

In this paper we briefly presented a novel tangible user interface for context aware music player that integrates a customizable configuration interface for automatic music selection. Based on our observations, we conclude that for context aware music players, two key factors that influence user experience are customization options and making autonomous decisions transparent.

In the future we aim to further improve the visualization of the decision making. We also want to include the computer as well as some speakers into the box that is currently an external USB-device. We want to create a single device, which can be easily carried between rooms. The configuration GUI could then potentially be accessed via a web-interface from any device with a browser.

REFERENCES

1. Adlin, T., and Pruitt, J. *The Essential Persona Lifecycle: Your Guide to Building and Using Personas*, 1st ed. Morgan Kaufmann, Burlington, MA, USA, May 2010.
2. Barney, D., Gust, A., and Liguori, G. College students' usage of personal music players (pmp) during exercise. *The ICHPER-SD Journal of Research in Health, Physical Education, Recreation, Sport & Dance* (2012), 23–26.
3. Cunningham, S., Caulder, S., and Grout, V. Saturday night or fever? context-aware music playlists. In *Proceedings of the 3rd International Audio Mostly conference on Sound in Motion (AM 2008)*, Interactive Institute Sonic Studio (Piteå, Oct. 2008).
4. Guan, D., Li, Q., Lee, S., and Lee, Y. A context-aware music recommendation agent in smart office. In *Proceedings of the Third international conference on Fuzzy Systems and Knowledge Discovery*, FSKD'06, Springer-Verlag (Berlin, Heidelberg, 2006), 1201–1204.
5. Han, B. J., Rho, S., Jun, S., and Hwang, E. Music emotion classification and context-based music recommendation. *Multimedia Tools Appl. 47*, 3 (May 2010), 433–460.
6. iClapper. Website: `http://iclapperapp.com`; visited on October 23rd 2013.
7. Lee, J., and Lee, J. Music for my mood: A music recommendation system based on context reasoning. In *Smart Sensing and Context*, P. Havinga, M. Lijding, N. Meratnia, and M. Wegdam, Eds., vol. 4272. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, ch. 14, 190–203.
8. neurowear. mico. Website: `http://neurowear.com/projects_detail/mico.html`; visited on November 5th 2013, 2013.
9. Park, H.-S., Yoo, J.-O., and Cho, S.-B. A context-aware music recommendation system using fuzzy bayesian networks with utility theory. In *Fuzzy Systems and Knowledge Discovery*, L. Wang, L. Jiao, G. Shi, X. Li, and J. Liu, Eds., vol. 4223 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, Berlin, Heidelberg, 2006, ch. 121, 970–979.
10. Plutchik, R. The nature of emotions. *American Scientist 89*, 4 (2001), 344.
11. Schmidt, A. *Context-Aware Computing: Context-Awareness, Context-Aware User Interfaces, and Implicit Interaction*. The Interaction Design Foundation, Aarhus, Denmark, 2013.
12. Sellitsch, D. *Master's Thesis: A Context Aware Music Player*. Vienna University of Technology, Vienna, Austria, 2013.
13. Thayer, R. *The Biopsychology of Mood and Arousal*. Oxford University Press, New York, 1989.
14. Ullmer, B., Ishii, H., and Jacob, R. J. K. Token+constraint systems for tangible interaction with digital information. *ACM Trans. Comput.-Hum. Interact. 12*, 1 (Mar. 2005), 81–118.