Proceedings of the 2nd Workshop on

Pervasive Advertising

in conjuction with Informatik 2009 in Lübeck

Jörg Müller, Paul Holleis, Albrecht Schmidt,
Michael May (Editors)

October 2nd, 2009
Lübeck, Germany

http://www.pervasiveadvertising.org

Title page image created with http://www.wordle.net/ based on the content of the entire proceedings.

Table of Contents

Pervasive Advertising

Jörg Müller[1], Paul Holleis[2], Albrecht Schmidt[3], Michael May[4]

[1]University of Münster
[2]NTT DoCoMo EuroLabs
[3]University of Duisburg-Essen
[4]Fraunhofer IAIS

Abstract: Pervasive Advertising (pervasiveadvertising.org) explores the application of pervasive computing technologies to the domain of advertising. Over the last decade many pervasive computing technologies have emerged, most notably mobile devices, ubiquitous digital displays, and a variety of location system and tracking tools. As these technologies become widespread new opportunities and challenges for advertising arise. In this second workshop on pervasive advertising we hope to get a wider overview of current research efforts and application opportunities as well as a chance to map out a roadmap for further research and deployment challenges in this domain.

1 Background

Electronic displays have become ubiquitous and replace traditional posters and billboards. Mobile phones come with high speed internet access and location sensing built-in. Newspapers and books are being replaced by their digital counterparts. Social networks including detailed profiles and interests of users are placed on the internet and shared with friends. These technological advances, besides others, change the opportunities and challenges for advertising radically. In turn, advertising is a major driving factor paying for the real-world deployment of these innovations. Paid by advertising, pervasive services can be provided to people free of cost, considerably accelerating their adoption. The combination of pervasive computing and advertising is holding considerable economic potential, and therefore the potential to impact everyday life of the worlds urban population over the next years.

2 Motivation

Many people only know advertising from an audience perspective. Because ultimately advertisers pay for advertising, when badly done advertising often respects the advertisers interests much more than the audiences interests. Many people have experiences where they feel offended or annoyed by blunt, aggressive ads. Naively, this could lead people to assume that a world without any advertising may be best. People who run their own business on the other hand often know from their own experience how badly they need to communicate their messages and advertise their products and services to potential consumers to keep the business running. Those people know how important a healthy advertising ecosystem really is. And indeed, interests of advertisers and audiences can be balanced, as is shown by street furniture (e.g. bus stops, public toilets), web sites or newspapers paid by advertising. Therefore, we consider it less responsible to simply ignore the development of pervasive advertising and let its go its way, but consider it very important to investigate and shape this urban future.

3 Topics

The topics proposed for this Workshop were:

- Digital Signage Advertising
- Mobile Advertising
- Location and context-based advertising
- Tracking technologies for advertising (GPS, RFID, GSM, Bluetooth, sensors, vision)
- Techniques for assessing sensor data
- Advertising in electronic news papers and e-books
- Assessing the effectiveness of advertisements in the real world (e.g. counting impressions, audience measurement, see-to-buy ratio, scan-to-buy ratio etc.)
- Mixing content with advertisement
- New pervasive computing technologies that are applicable to advertising
- Dealing with limited attention (SPAM prevention in the real world)
- Integration of social networks with pervasive advertising
- Privacy and pervasive advertisements

Particularly interesting topics and challenges for pervasive advertising are:

Technologies Technologies that are considered important for pervasive advertising are digital signage, mobile devices, electronic book or newspapers and also robots. Interesting applications emerge from the combinations of these media, for example a user profile may be stored on a mobile device, but ads may be shown on digital signs.

Players The question of which players or stakeholders are important for pervasive advertising is not yet solved. As with any medium, there will be an audience and content providers, of which advertisers are a special case (who pay for their content being shown). Whether moderators are necessary, or which other stakeholders are important (such as device owners or 'pervasive service providers'), is an issue of discussion.

Business Models The construction of viable business models for pervasive advertising is not yet solved. How services could be prized and how the resulting revenue stream could be distributed among the players is still evolving. Serendipitous ads, which address needs the audience has not yet been aware of, have been proposed as beneficial for pervasive advertising. Also, 'pervasive service providers', who act similar to mobile service providers, have been proposed.

Targeting Targeting is a very active area of discussion. The possibility to deliver the right ad to the right people in the right moment, using the best medium, format and presentation, is unsurpassed in pervasive computing environments. Different approaches have been proposed how to determine the interests of the audience, either by having them construct explicit profiles, monitoring their behaviour and response to previous ads, or by harvesting information sources like social network sites.

Audience Measurement Audience measurement, the determination of the audience reaction to ads presented, is another area where pervasive advertising excels. Approaches to measure the audience reaction via face detection (e.g. for digital signage with cameras installed on top of the signs), interaction data or coupons, have been presented.

The tug-of-war Interests of audiences and content provides conflict in certain situations. In particular, in the attention economy, advertisers want to attract the audiences attention, while audiences want to protect their attention from ads that are not interesting to them. Because audience attention is limited, these interests must be balanced to avoid an impression of SPAM by the audience while still maintaining a healthy advertising ecosystem.
Similarly, in order to better target ads, advertisers seek to collect as much information about the audience as possible. Seeking to protect their privacy, audiences may want to keep certain information for themselves. Also these interests must be balanced.

Aesthetics and Audience Experience As pervasive adverts may include the entire surroundings of the audience, aesthetics and audience experience are discussed as important issues for pervasive advertising. In particular, the impact of different designs are being investigated and approaches to automate these factors are being proposed.

Implementation Pervasive advertising is in need of complex middleware to manage the huge numbers of different devices and different players. Different technological approaches are currently being discussed and tested.

Interactivity and Dialog One of the great virtues of pervasive advertising is to go beyond mass-media unidirectional communication (as applied with billboards, tv or radio) and open a true dialog between the advertiser and the audience.

4 Objectives

The aim of this workshop is to bring together researchers, developers and practitioners from academia and industry who are concerned with envisioning, creating and implementing future advertising systems. By fostering discussion on this emerging topic, we hope that science can actually shape this emerging reality instead of only analyzing it after the fact. We hope that the workshop will foster a community that has an interest in understanding the future of advertising in the light of pervasive computing.

5 Visions

Taking a positive view we can envision advertisements that precisely match a persons interests and fit the current situation so well that we enjoy receiving them and that we do see advertising as a pleasant distraction. On the contrary taking a negative view one could image a world where we cannot escape from advertisement, where we are continuously tracked and where advertisements reduce the quality of life. Both views, even though they are extreme, are worthwhile a further discussion and we hope the workshop can provide a venue for this.

6 Program Committee

We would like to thank all members of the program committee for Pervasive Advertising: Alois Ferscha, Antonio Krüger, Florian Michahelles, Felix Graf von Reischach, Mira Spassova, Philipp Lehmann, Oliver Paczkowski, Juliane Exeler, Florian Alt, Keith Cheverst, Martin Strohbach and Lucia Terrenghi.

Engaging Passers-by with Interactive Screens – A Marketing Perspective

Daniel Michelis[1], Hendrik Send[2]

[1]Anhalt University of Applied Science
Strenzfelder Allee 28
D-06406 Bernburg
dm@e-acht.de

[2] Institute of Electronic Business
Affiliated Institute to the University of the Arts Berlin
Hardenbergstraße 9A
D-10623 Berlin
send@ieb.net

Abstract: This position paper approaches pervasive advertising from a marketing perspective. It provides a short introduction into the shift from an economy of attention to an economy of engagement. In the second part it analyses seminal research on interactive screens and shows how passers-by can actively engage with public displays.

1 From attention to engagement

For a long time, information was considered to be one of the most important resource of the information age but with the spread of digital technology, information is no longer a scarce commodity but became ubiquitous. As Richard A. Lanham puts it: "If economics is about the allocation of resources, then what is the most precious resource in our new information age? Certainly not information, for we are drowning in it. No, what we are short of is the attention to make sense of that information." [1] Thus, for an advertising company, not the spread of information was the main difficulty, but the attention of their customers. The attention economy served as a guiding paradigm for traditional advertisers and e.g. led to the established AIDA model, that is used to describe a basic buying process that includes Attention, Interest, Desire and Action. As attention is the first step in the buying process it plays a very central role in successfully selling products [2].

Today, we see an important shift in consumer behavior. With the advent of large participatory platforms like Wikipedia, YouTube, and Facebook, usage of the Internet is not limited to passive reception. Passive consumers become actively engaged. Given the growth of user generated content, businesses and advertising in particular explore possible competitive advantages and try to best benefit from the development [3].

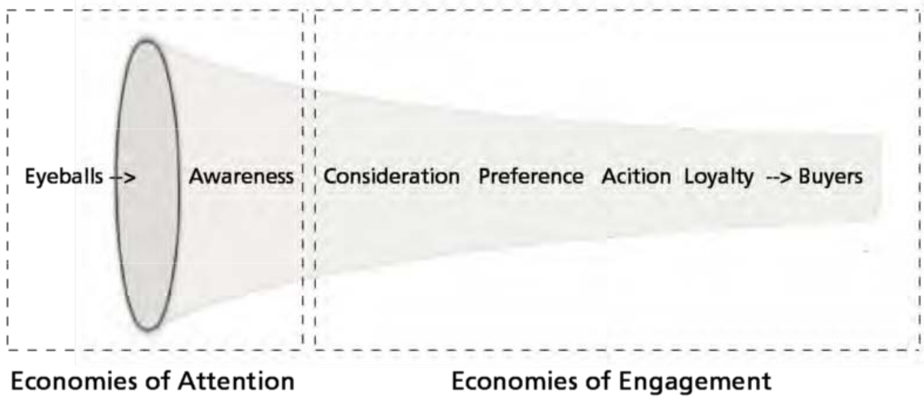


Figure 1: Influence of Attention and Engagement [according to 4]

In their analysis of the altered relation between companies and consumers, Li and Bernoff use the marketing funnel in figure 1 to describe the transition from an economy of attention to an economy of engagement: "In tradition marketing theory, consumers are driven into the big end through awareness activities like advertising. They proceed through stages – including consideration, preference, and action – to become buyers. Marketers have little control over what happens in the middle stages" [4] but the influence of the economy of engagement seems to be the strongest there.

The authors describe active engagement as when customers start to lead the conversation about a product and when they support each other: „With so many products trying to get people's attention, shouting at them isn't nearly as effective as it used to be. […] Marketers no longer dictate the path people take, nor do they lead the dialogue. Once people are aware of you product, a new dynamic kicks in: people are learning from each other. Social technologies have revved up that word-of-mouth dynamic, increasing the influence of regular people while diluting the value of traditional marketing. […] Costumers in the middle of the funnel are engaged in conversations on blogs, in discussion forums, and in social networks. Your company can participate in these places, but shouting doesn't work. Conversations do." [4]

To advertise their products, companies still need to get the attention of their customers. But understanding the dynamics of an economy of engagement, marketers should connect attention with engagement. In the process of communication they should first get the attention of their customers and second enable engagement further down the marketing funnel. In enabling engagement they can strengthen the relationship with the consumer and increase their involvement.

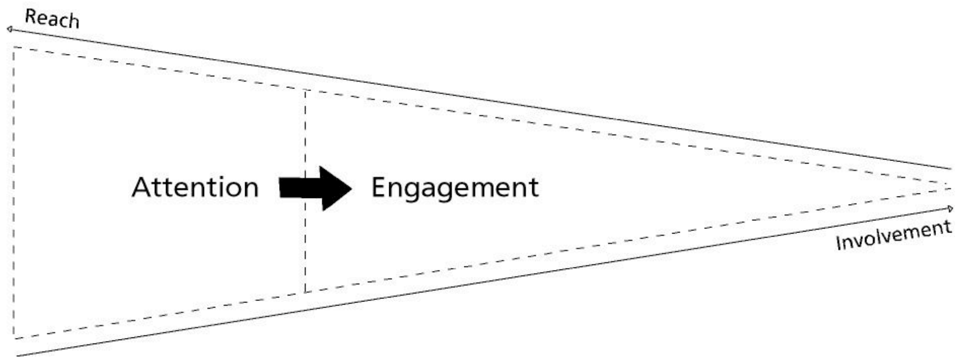


Figure 2: Connecting Attention and Engagement

In the following, this paper will analyze engagement possibilities with interactive screens. While in the past the ability to actively engage with public displays was not an option, this now seems to be changing.

2. Attention and engagement with interactive screens

Initial analysis of public interactive displays has already provided valuable insights the distinction between attention and engagement. While Brignull and Rogers developed a "two-thresholds framework" with peripheral and focused activities, they point out the possibility of engagement that they call direct interaction [5]. Vogel and Balakrishnan differentiate between ambient display, implicit, subtle, and direct interaction. In this "four-phase framework" engagement with the viewer is only possible in the fourth phase that they call the phase of personal interaction [7]. Last but not least, Streitz et al.'s "three-phase framework" differentiates between ambient, notification, and interactive zone. Here, active engagement is only possible in the latter [6]. To understand how engagement with passers-by can be realized, I will present a short overview of how users are actively engaging in the different frameworks.

2.1. Two-Thresholds Framework

Brignull and Rogers make a distinction between and divide the process of interaction into three general phases. These phases range from perception to direct interaction with the featured content, as shown in figure 1, and include: peripheral awareness activities, focal awareness activities, and direct interaction. Direct interaction activities concern the active engagement with the interaction system: “In this activity space, an individual (or a group acting cooperatively) type in their opinion to the display. In their analysis Brignull and Rogers not only look at these three kinds of activities but also at the transition zones between them. Their analysis revealed that the transition zones between different types of activities represent a key bottleneck in public interaction behavior. [5]

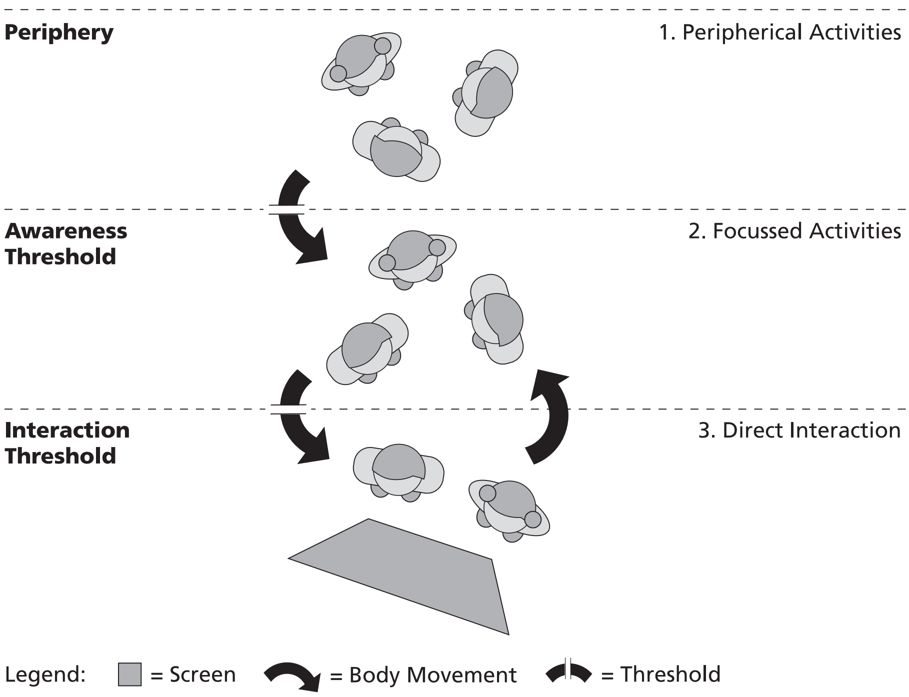


Figure 3: User-engagement through direct interaction [8 according to 5]

2.2. Three-Phase Framework

By dividing the space into ambient, notification, and interactive zones, as Streitz et al.'s "Three-Phase Framework for Interaction Phases" in figure 2 shows, content can be adapted to each phase [6]. Their model focuses on the design of the system from the ambient to the interactive zone, rather than the behavior of the user. As a viewer approaches the surface, or once her presence is registered, interactive content is displayed that can provide personal, detailed information:

"The third zone is active, once the person is very close to the GossipWall [display]. In this case, the person can approach the GossipWall and interact with each single cell (= independent interactive pixel). This is able to store and communicate information in parallel in combination with mobile devices."

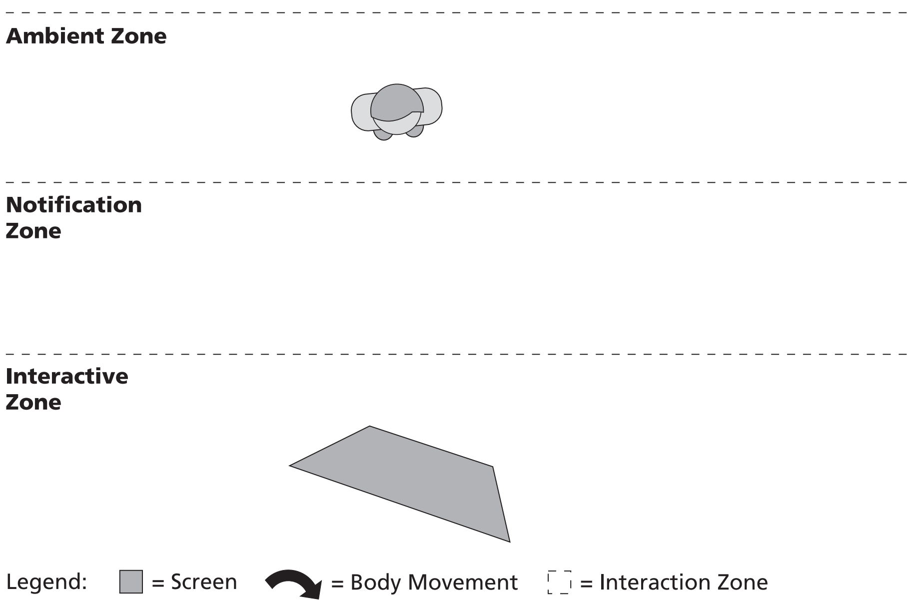


Figure 4: User-engagement within the interactive zone
[8 according to 6]

As figure 2 illustrates, the Three-Phase Framework focuses on the differentiation between specific zones of interaction. Details about the user's activities are not given.

2.3. Four-Phase Framework

In their "four-phase framework" Vogel and Balakrishnan differentiate between ambient display, implicit, subtle, and direct interaction. The framework differs from Streitz et al.'s three-phase model in that it emphasizes the interaction between the system and the user and the "fluid transitions between phases, and [...] supports sharing by several users each within their own interaction phase". By dividing Streitz et al's 'interaction zone' into 'subtle' and 'personal' interaction phases, and by generalizing the notion of a 'notification zone' into an 'implicit' interaction phase, [the] framework suggests a wider range of implicit and explicit interaction techniques." [7]

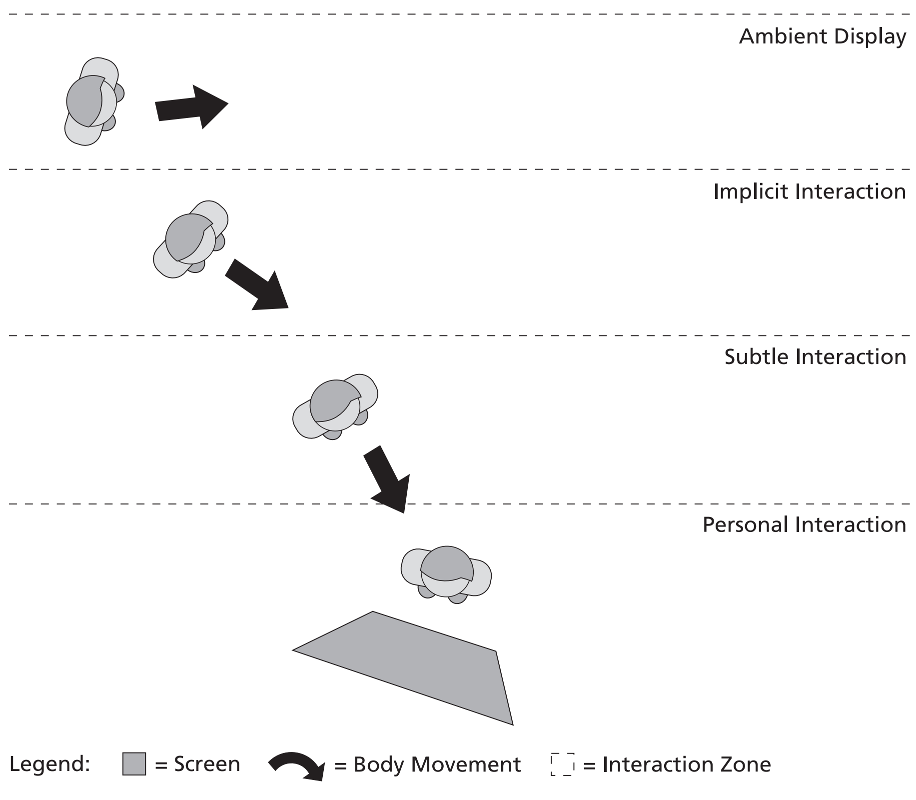


Figure 5: User-engagement through personal interaction [8 acc. to 7]

The framework's four continuous phases cover a range of activities from distant implicit public interaction to personal interaction. It also looks at fluid inter-phase transitions between every phase. The subtle interaction phase merely provides a description of use behavior. However it does indicate that “to this point, the user has only interacted implicitly.” In the final phase, the user is close to the display. Since she stands close to the surface “the user should be able to move closer to the screen and touch information items for more details, including personal information. […] Since the user is close to the display, their body can help occlude the view of their personal information from others.”

3. Conclusion

By approaching pervasive advertising from a marketing perspective this paper gave a short overview on the shift from an economy of attention to an economy of engagement. In the second part it presented seminal work that analyzed the process of interaction with public displays to search for user-engagement possibilities within the area of pervasive advertising. The author of this paper hopes that pervasive advertising will not start to shout out loud to get attention of customers, but enable a new way of calm advertising where customers collaborate and freely and quietly engage with public screens based on their current needs.

References

[1] Lanham, Richard (2006), The Economics of Attention: Style and Substance in the Age of Information, University of Chicago Press

[2] Marketing: Meffert, H. Burmann, C., Kirchgeorg, (2007), Gabler Verlag

[3] H. Send, D. Michelis: Contributing and socialization - biaxial segmentation for users generating content, 9th International Conference on Innovative Internet Community Systems

[4] Li, C. / Bernoff, J. (2008), Groundswell: Winning in a World Transformed by Social Technologies, Harvard Business School Press

[5] Brignull, H., Rogers, Y. (2003), Enticing People to Interact with Large Public Displays in Public Spaces, INTERACT'03

[6] Streitz, N. A., Röcker, C., Prante, Th., Stenzel, R., van Alphen, D. (2003), Situated Interaction with Ambient Information: Facilitating Awareness and Communication in Ubiquitous Work Environments. In: Tenth International Conference on Human-Computer Interaction (HCI International 2003), June 22-27, 2003.

[7] Vogel, D., Balakrishnan, R. (2004), Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In Proceedings of the 17th Annual ACM Symposium on User interface Software and Technology (Santa Fe, NM, USA, October 24 - 27, 2004). UIST '04. ACM, New York, NY, 137-146

[8] Michelis, D., Interaktive Großbildschirme im öffentlichen Raum: Nutzungsmotive und Gestaltungsregeln, Gabler, 2009

Traditional and Digital Signage

Jörg Müller

joerg.mueller@uni-muenster.de

Abstract: In this paper we explain how the content of traditional and digital signs heavily depends on the context in which the sign is seen. This is followed by a comparison of the properties of traditional and digital signs, in which we show that digital signs enable us to exploit many proporties that signs have, in particular context adaptivity and indexicality.

1 Introduction

Signage has influenced urban life since ancient times. Today, we can observe that some of the traditional, static signs are being replaced by digital signs (see Figure 1). While it may seem obvious that there are differences between traditional signs and digital signs, there is also a good number of common properties, which should not be forgotten.

Figure 1: Traditional and digital signage live side by side.

2 Signage

The study of the meaning and use of signs and symbols is called Semiotics [KP06]. Semiotics emphasizes that information is a representation of some object. The Semiotic Triangle (see Figure 2) depicts the relationship between a representation (e.g. on a sign), the object (that is depicted) and the interpretation (of the audience).

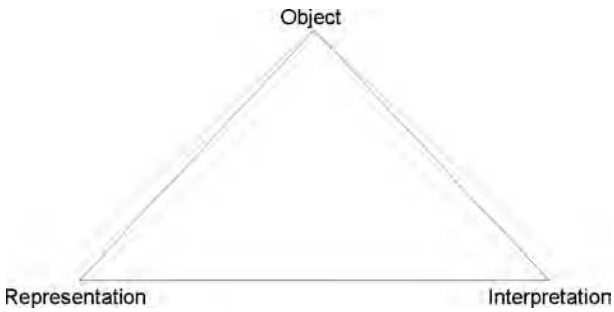


Figure 2: The Semiotic Triangle

Three types of representation are possible. Symbolic representations refer to an object via convention (e.g. script). Iconic representations use similarity to a depicted object (e.g. images). Finally, indexical representations refer to an object in the immediate context (e.g. via an arrow). Symbols and icons may work independent of context, i.e. in a "one size fits all" manner. Indices in contrast are deeply embedded in the context of the audience. Indices can refer to something spatially (e.g. 200m left) or temporally (e.g. tomorrow). Potentially they could also refer to social context (e.g. "the guy with the red hair"). The applicability of indexicality to context-aware computing systems has been investigated in [KP06].

Not only for indices, but also for symbols and icons, the interpretation of the audience may widely differ depending on the context in which the representations are delivered. Marshall McLuhan with his famous quote "The medium is the message"[ML64] emphasized that the medium itself, e.g. a sign, has a certain effect on the audience independent from and interacting with the content shown. William Mitchell captures this effect in his book "Placing Words" [Mit05], page 9, as follows:

> "But the introduction of technologies for inscribing physical objects with text, and the associated practices of writing, distribution, and reading, created a new sort of urban information overlay. Literary theorists sometimes speak of text as if it were disembodied, but of course it isn't; it always shows up attached to particular physical objects, in particular spatial contexts, and those contexts — like the contexts of speech — furnish essential components of the meaning."

This effect of different interpretations in different contexts is related to the psychological effect of *priming*[And04]. People interpret new facts in the contexts of stimuli they have been presented before (e.g. when being told about old women before, they are more likely to recognize an old woman in an ambiguous picture). Mitchell captures this effect as follows:

> "Shouting "fire" in a crowded theater produces a dramatically different effect from barking the same word to a squad of soldiers with guns. Writing it on a hydrant yields yet another result. The meaning of a message depends not only upon the information that it contains, but also upon the sort of local ignorance or uncertainty that it reduces — in other words, upon what the message's recipients require information *about*." [Mit05], page 3.

This underlines the importance that time and location have both for understanding and designing digital signage. In order to implement context adaptive digital signage, the signs need to be equipped with models of their spatial and temporal context. Signs can be seen from a variety of distances and potentially refer and react to their context in a variety of scales. Therefore, both space and time need to be understood in different scales — which is a classical topic of GIScience[1]. Montello[Mon93] highlights that to understand scales of space, it is important how the space is perceived, e.g. from an airplane a country may be perceived differently than from the ground. He introduces the scales of figural, vista, environmental and geographical space. Figural space is projectively smaller than the body, e.g. a table or distant landmark. Vista space is projectively as large or larger than the body but can be perceived without moving around, e.g. a room or the horizon. Environmental space surrounds the body, but it can be directly perceived by moving around, e.g. a city. Geographical space is too large to be directly perceived, it can only be perceived by representations, e.g. maps that map it to figural space. This indicates that, while a single sign will probably always be within vista space, whole display ecologies can fill environmental or even geographical space. Signs can also index to space in any of these scales.

For digital signage, it is not only important how the physical space surrounding the signs is structured, but also how people experience it and behave in it. The unreflected use of "space" to refer to locations where humans interact with other humans or technology has come under critique. Harrison and Dourish[HD96, Dou06] draw on their experiences with Media Spaces to introduce the distinction between space and place. Space is the mere spatial environment and configuration ("the opportunity"), which, in the case of built up environments, certainly has evolved through social processes. Place, in contrast, is the social meaning of a location ("the understood reality"), which evolves through social behavior of people, in coordination with mores and norms. It is emphasized that placeness can not be designed into systems, it can only be designed for. The evolution of placeness also heavily depends on the feeling of ownership people experience. In the case of the Xerox PARC Media Space for example, cheap equipment could be reconfigured and appropriated by users, which strongly influenced their feeling of ownership. In the Bellcore Media Space on the other hand, expensive equipment was used to create a life-like video conference experience, but users could not appropriate the technology. As Harrison and Dourish state: "It wasn't theirs, and they could not *make* it theirs."

[1] Geographic Information Science

Table 1: Comparison Traditional and Digital Signage (iterative improvements in *italic*, completely novel in **bold).**

Property	Traditional Signage	Digital Signage
Content	Text, Graphics	**Animation, Video, Reactive, Interactive**
Content Source	Mostly big business, sometimes smaller, user-generated	*Anybody*, **Life sensor data (weather), life from the Web (News, Train table)**
Multiple Content	A few, cumbersome	*Any content available*
Updates	Expensive, about 10 days	*Cheap, within milliseconds*
Content Depth	What fits the sign	*Any depth with interaction*
Indexical Content	To location, to time and audience within limits	**Any measurable context: e.g. time, history, audience**
Context Adaptivity	Location, time (put out Open sign), audience by location	*Any measurable context: Time in milliseconds, audience up to individual*
Audience Measurement	Manual counting or interviews	*Automatic*
Learning	Manually by media planner	*Automatic*
Interactivity	Sometimes mobile phone	**Touch, Gesture, mobile phone, Web etc.**

Further insights on the behavior of people in space are given by Goffman[Gof59]. In his analysis of social life interpreted as a scene play, Goffman uses a theatrical metaphor for place. “Front stage” and “backstage” distinguish different modes of behavior in interaction. A salesperson may for example treat a customer more formally in the shop room than in the storage area. A digital sign can be interpreted as a stage, where people can perform *on it* by submitting certain content as well as perform *in front of* it by behaving accordingly (e.g. users performing in front of the Dynamo displays [Bri05]). As users constantly seek to maintain a certain role, they may for example want to know whether a photo of them appears distorted on the sign.

3 Properties of Digital Signage

Digitial signs have many properties that differ from their paper and wood counterparts. Originally, signs were painted manually on the presentation surface. Later, printing and computerized printing technologies simplified reproduction and inclusion of other artwork, e.g. photos. Light installations made simple animations possible (as you still can see in Las Vegas). Scrollable billboards enabled to show a few different content pieces in a row. Digital signage goes beyond the possibilities of these technologies in a number of ways (see Table 1). While traditional signs supported text, graphics and very simple animations, digital signs enable full animations, video, content that reacts to the audience, and interactive content. For traditional signs, big business was the most important source of content. Some smaller businesses could afford some signs, and (illegal) graffiti have been an instantiation of user-generated content in major cities around the world.

Often, content providers need to visit the location of the sign themselves. Digital signs enable anybody to submit content, regardless of his location. This opens whole new possibilities for harnessing the 'long tail' of numerous small content providers as well as for user-generated content, despite the unsolved issue of content filtering. Additionally, life sensor data (like weather) and life Web data (like news or train delays) can be presented. While scrollable billboards enabled a limited number of different content items, on digital signs unlimited different content can be presented. On traditional signs, updates are cumbersome, and many advertising signs are only updated every 10 days or so. On digital signs, updates are cheap and can be performed within milliseconds. On traditional signs, the depth of the content is confined to what fits on the sign. In order to provide further information, signs need to refer to other resources (e.g. Web addresses). With a "would you like to know more?" functionality, interactive digital signs can harness the power of Hypertext and provide unlimited content depth in-place.

For presenting indexical content, traditional signs can refer to the location (e.g. "200m left"), to time (e.g. "party tomorrow"), or to the audience (e.g. "I want YOU for US Army"). Because of the high effort of updates, with traditional signs it is often difficult to get indices to dynamic factors right (e.g. the sign still being there one week after the party, or "YOU" referring to nobody when there is no audience in front of the sign). Digital signs can flexibly index to time, history ("the film you watched here yesterday") and audience ("buy the same jeans as the guy with the red hair").

Regarding context adaptivity, traditional signs have been adapted to the location (and the expected audience at that location) for a long time. There have also been limited possibilities of adapting to the time (e.g. putting an "Open" sign on a shop door). Digital signs in contrast can adapt to any measurable context within milliseconds, and also adapt to the audience up to the individual (and his facial expression). Audience measurement with traditional signage has been performed manually, by counting, observing and interviewing the audience. Although traditional signage can also be equipped with automatic audience measurement, this is more common with digital signage (e.g. for interactive digital signage, interaction logging is easy). In order to learn models which content works well in which context, traditionally media planners created statistics manually. Digital signage opens the possibility to accomplish this learning automatically with Machine Learning mechanisms[Mit97] (although this can in principle also be applied to traditional signage).

Recent developments enabled partially interactive traditional signage. Being equipped with dedicated hardware, such signs could send data, e.g. ring tones, to mobile phones. Digital signage offers broad possibilities for interaction: Touch interfaces, gestures, Web interfaces and interaction via mobile phones.

4 Conclusion

As we have seen, the effect of signs depends heavily on the context. This includes the behavior of the audience in front of the signs, but also the behavior of content providers who want to know their content appears in context and if this fits their social role. Additionally, signs can refer to their environment via indexicality. While digital signs are in many ways similar to traditional signs, they offer a number of novel possibilities especially regarding context adaptivity and indexicality.

References

[And04] John R. Anderson. *Cognitive Psychology and its implications (6th edition)*. Palgrave Macmillan, 2004.

[Bri05] Harry Brignull. *Understanding and Designing for the Voluntary Adoption of Community Displays*. Dissertation, University of Sussex, 2005.

[Dou06] Paul Dourish. Re-space-ing place: ”place” and ”space” ten years on. In *CSCW ’06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, Seiten 299–308, New York, NY, USA, 2006. ACM.

[Gof59] Erving Goffman. *The Presentation of Self in Everyday Life*. Penguin, New York, 1959.

[HD96] Steve Harrison und Paul Dourish. Re-place-ing space: the roles of place and space in collaborative systems. In *CSCW ’96: Proceedings of the 1996 ACM conference on Computer supported cooperative work*, Seiten 67–76, New York, NY, USA, 1996. ACM.

[KP06] Jesper Kjeldskov und Jeni Paay. Indexical interaction design for context-aware mobile computer systems. In *OZCHI ’06: Proceedings of the 20th conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction: design: activities, artefacts and environments*, Seiten 71–78, New York, NY, USA, 2006. ACM.

[Mit97] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Science/Engineering/Math, March 1997.

[Mit05] William J. Mitchell. *Placing Words. Symbols, Space, and the City*. MIT Press, 2005.

[ML64] Marshall McLuhan und Lewis H. Lapham. *Understanding Media: The extensions of man*. Routledge, 1964.

[Mon93] Daniel R. Montello. Scale and Multiple Psychologies of Space. In *COSIT ’93: Proceedings of the 1993 Conference on Spatial information theory: A theoretical basis for GIS*, Seiten 312–321, Berlin, 1993. Springer.

Meaningful advertising: pervasive advertising in the experience economy

Peter van Waart[1], Ingrid Mulder[2]

[1]School of Communication, Media & Information Technology
Rotterdam University
Pieter de Hoochweg 129
NL-3024 BG Rotterdam
p.van.waart@hro.nl

[2]ID-StudioLab, Faculty of Industrial Design Engineering
Delft University of Technology
Landbergstraat 15
NL-2628 CE Delft
i.j.mulder@tudelft.nl

Abstract In traditional marketing advertising is part of product promotion, one of the four P's of marketing. In the 20th century media channels such as papers and magazines, and later on radio and television, were used to distribute and broadcast advertisements in order to create product and brand awareness and preference amongst consumers, in order to raise product sales. In the last two decades however, economical developments resulted in the experience economy: a new era of marketing and branding, in which traditional advertising is becoming less effective and meaningful experience branding is key. Meaningful (brand) experiences can make a difference to consumers, because of the personal relevance for people with regard to their personal values and beliefs that are the basis of people’s consumer behaviour and brand preferences. Pervasive advertising should seek for applications with which people can interact with brands in an intellectual, an emotional and a physical way.

Introduction

Where in pre-industrial societies commodities delivered economical value, industrialisation made manufacturers compete on product quality, and modern society produced more and more value from delivering services (see Figure 1). For companies and brands however, it is getting harder to differentiate from each other with products and services because of an endless choice that promises personal freedom, but at the end results in a confused consumer that is not able to value the worth of his choice [Sc04]. Pine and Gilmore [PG99] described the economical development towards an experience economy. In this experience economy meaningful (brand) experiences can make a difference to consumers, because of the personal relevance for people with regard to their personal values and beliefs which are the basis of people’s consumer behaviour and brand preferences [Ro73; Ba09] and identity and lifestyle [Bo84].

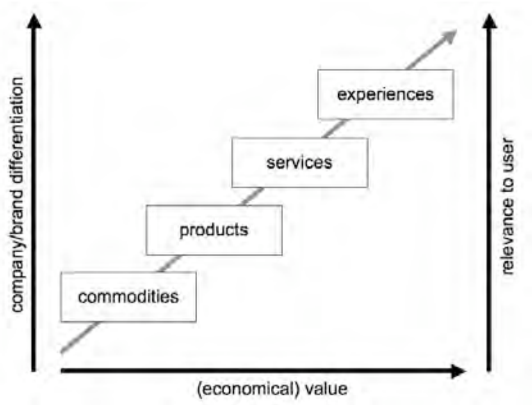


Figure 1. Towards an experience economy [PG99].

This economical development shows the transformation of product- and sales-driven organisations, for which cost-reduction and pricing became less and less competing factors, into experience brand-based organisations. The experience brand strategy differs fundamentally from the traditional marketing strategies. Traditional marketing focused on the product or corporate profit, resulting in an advertising industry that Godin [Go03] calls the TV-industrial complex: buying advertisement space in the media to promote products, sell the products and than use the profit to buy even more advertisement space. At the end this resulted in an advertising rat-race where the companies with the largest budgets end up as the winners of the competition. But what about the relevance for people? Goldhaber [Go97] proposed that the so-called information economy is actually an attention economy where attention of consumers became scarce due to the enormous amount of information [DB01]. Attention can be seen as a scarce resource, which drives economical value. Experience branding differs from attention economy in the way that it put human values as drivers of economical value. From that point of view, attention is triggered when messages are relevant to people, that is, when they adapt to human values of people.

In the current experience brand strategy the consumer is more than ever in the centre of the brand experience more than ever. In fact, the emphasis is on ‘humans’, rather than the consumer. Experience brands aim to provide ‘meaningful experiences’ to people for which a real understanding of human values is required, because at the end something can only be of meaning to someone when it comes to what one values in life [PG99; SDR06].

In the field of experience branding, or meaningful branding, the traditional concept of advertising has been reconsidered. When the product or corporation is no longer in the center of the marketing strategy, advertisements that solely communicate product features and benefits fail to lead to a meaningful human experience. Without taking into account the personal beliefs and former history of (life) experiences, a true experience will not appear because an experience has to be constituted by an individual even more than by the experience (brand) provider [PG99]. Brand and person really have to interact with each other. Brand experience has a behavioural impact; it affects consumer satisfaction and loyalty [BSZ09].

In experience branding the concept of 'touch points' is key. A touch point is any contact or moment in which one is physically or virtually interacting with a brand. All touch points together influence one's brand experience. It is not so that every touch point tells the same story, but that every touch point contributes to the entire expression of brand meaning to a person [Ma08; SDR06; Wh06].

Consequently, traditional advertising no longer holds in the experience economy. Instead of pushed and persuasive advertisements, experience brands need to develop a consistent architecture of touch points that is naturally intertwined within the context of the user and that addresses his values.

2 Technological opportunities facilitating meaningful brand experiences

Current technological developments such as pervasive and ubiquitous computing changed the media landscape dramatically through the shift from broadcasting to personal and interactive media. Today's audience is able to interact with the sender or, even more important, is able to interpersonally exchange information immediately with thousands or billions of other consumers. Additionally, these emerging media provide access to information on production and sales strategies and processes resulting into a transparency that now give powers to consumers that is beyond control of advertisers. Brands cannot have the entire control of the brand experience and learn to share the control with the users, the people. At best, a brand finds itself in a co-creation with people [PG99] (See Figure 2.)

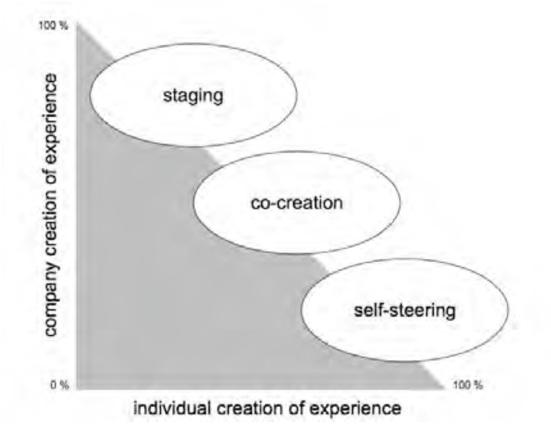


Figure 2. Co-creation as shared control between brand and people [PG99].

An interesting example of co-creation and the use of pervasive technology in experience branding is Nike+. Nike joins efforts with Apple's iPod to come up with a service that enriches people lives in what people value: health. Nike+ is of meaning to people by enriching their experience of keeping in physical shape by running. The Nike + iPod Sports Kit uses a sensor in one's running shoe that keeps track of your running. The running data is stored onto the iPod (or iPhone) and also the iPod serves you appropriate music to inspire your running. Back home, the running data can be uploaded to nikeplus.com where you can monitor all your runs and share motivations with other runners. This example shows that the brand experience is facilitated by the brands Nike and iPod combined with a pervasive application designed for this experiences. However, the actual experience is created by individuals themselves. Together these individuals co-create a branded running universe in which they share their data, bond with each other in training programs or compete with each other in for instance man versus women competitions.

Pervasive and interactive media offer immense opportunities to become interactive touch points and act as ‘meaningful’ media [HY08]. For centuries, myth/story, arts and rituals, respectively addressing the intellect, emotions and physical experience, proved to be powerful ‘media’ to constitute meaningful experiences. Pervasive and interactive media enable people to constitute brand experiences, individually as well as socially. Differently put designers and practitioners cannot design *an* experience, they just can design *for* an experience [MW07). Those pervasive systems and applications that enhance social capital as collective goods involving shared goals and values and social norms of reciprocity, are most usable to be of meaning over time [CM08]. What people experience with YouTube or Flickr or in social networks like Facebook or Twitter, is not so much defined by ‘product features’ or by ‘brand values’. The perceived value is for the larger part constituted by people themselves based on their personal values. People use these brands not only because the are useful, but foremost because they are meaningful: people don’t share their movie clips, photo’s and stories, they enjoy sharing experiences in their social network!

3 Pervasive advertising in the era of meaningful branding

In the era of experience economy, persuasion of consumers might become obsolete. It is not anymore about the effort a manufacturer has to put in creating a high volume of consumers buying his product through persuasive advertisements. Meaningful brands can only express themselves to people in a way that they respect and endorse people's choices and lifestyle as expression of their personal values and beliefs. This is in keeping with the vision of pervasive computing, which foresees novel scenarios of highly interactive environments in which communication takes place between single users and devices, between devices and devices, and between users and users. Moreover, technology becomes increasingly invisible and personal. Who fears technology that is of personal meaning? The fear for annoying or even harassing people is, in a way, superfluous in the paradigm of meaningful branding. In similar vein: has there been any truly believing Christian being annoyed by the view of a church?

In this way, brands can make use of pervasive technology in various touch points facilitating meaningful brand experiences (co-)created by people themselves. Therefore, pervasive advertising should seek for the embedding of brands into the natural living environment of people, in which people can interact with brands in an intellectual (symbolic) way with which meaning can be transferred between brand users, an emotional way by which users will hold a sustainable memory of the experience, as well as a physical (experiential) way in which the immediate conscious and unconscious impact takes place through the interaction with the applied technology. The era of advertisement not to persuade, has just begun.

References

[Ba09] Batey, M. (2009). Brand Meaning. New York: Routledge.

[Bo84] Bourdieu, P. Distinction: A Social Critique of the Judgement of Taste. London, Routledge, 1984.

[BSZ09] Brakus, J.J. Schmitt, B.H. & Zarantonello, L. Brand Experience: What Is It? How Is It Measured? Does It Affect Loyalty? In: Journal of Marketing. Vol. 73, May 2009, p. 52–68.

[CM08] Carroll, J.M. & Mentis, H.M. The useful interface experience: the role and transformation of usability. In (Schifferstein, H.N.J. & Hekkert, P. Eds.): Product Experience. Amsterdam, Elsevier, 2008.

[Go03] Godin, S. Purple Cow: Transform Your Business by Being Remarkable. Portfolio, 2003.

[Go97] Goldhaber, M. H. The Attention Economy and the Net. 1997. First Monday 2 (4), http://www.firstmonday.org/issues/issue2_4/goldhaber

[DB01] Davenport, T.H., Beck, J. C. Attention Economy. Harvard Business School press, Boston, MA, 2001.

[HY08] Highland, M. & Yu, G. Communicating Inner Experience with Video Game Technology, In: Heidelberg Journal Religions on the Internet, 2008 Vol. 03.1, 2008, p. 267-289.

[Ma08] Mager, B. Service Design. In (Erlhoff, M. & Marshall, T. Eds.): Design Dictionary. Basel, Birkhauser Verlag AG., 2008.

[MW07] McCarthy, J. & Wright, P. Technology As Experience. MIT Press Ltd., 2007.

[PG99] Pine II, B.J. & Gilmore, J.H. The Experience Economy: Work Is Theater & Every Business a Stage. Boston, Academic Service, 1999.
[Ro73] Rokeach, M. The Nature of Human Values. New York, Free Press, 1973
[SDR06] Shedroff, N. Diller, S. & Rhea, D. Making Meaning: How Successful Businesses Deliver Meaningful Customer Experiences. New Riders Publishing, 2006.
[Sc04] Schwartz, B. The Paradox of Choice: Why More Is Less. New York, HarperCollins Publishers Inc., 2004.
[Wh06] Wheeler, A. Designing Brand Identity: A Complete Guide to Creating, Building, and Maintaining Strong Brands. Hoboken, Wiley, 2006.

Scent Marketing: Subliminal Advertising Messages

Bernadette Emsenhuber

bernadette.emsenhuber@jku.at
Johannes Kepler University Linz, Department of Pervasive Computing

Abstract: Store chains and service providers beguile customers with a pleasant shopping atmosphere often realized by installing scent diffusers to evaporate overwhelming fragrances. Such systems named as olfactory technology are becoming a standard interior of commercial locations as well as public places and are more and more gain importance in human-computer interaction. This paper delivers insight into current air design technology, the psychological background of scent marketing and gives a basis of discussing the relevance of olfactory communication for pervasive advertising and human-computer interaction.

1 Introduction

Like other media-based business advertising industry is increasingly confronted with the problem of information overload. Penetrating this *bead of information* and reaching costumers with advertising messages appreciably becomes difficult. An important issue for advertisers is finding new methods to persuade consumers of their goods and services. First impressions coming up with the word *advertising* are posters, newspaper ads, commercials on TV or radio, etc., actually visual and auditory media. But the advertising industry also uses consumption raising instruments like scents which can manipulate consumer behavior by unconsciously raising emotions and consequently manipulating purchase decisions. Experts defined this type of advertising as *scent marketing* or *air design*. The olfactory perception is an interface which allows a very subliminal communication between human beings and their environment because of a direct connection of the olfactory system to our emotional center. In the mid nineties psychologist examined the effect of scents on purchase behavior and confirmed that perfumed sales rooms would contribute to increase sales [Stö98]. Therefore marketing experts did not hesitate to use the olfactory channel as medium for subliminal messages.

2 Psychological background of olfactory perception

Olfactory perception takes an exceptional position in the neurological processing of sensory stimuli. The sense of smell differs from other forms of perception in the direct connection between the olfactory and the limbic system, our emotional center. The perception of odors nearly always causes emotional reactions - there is no way to avoid smelling odors or filtering olfactory information. As a result the olfactory channel offers high potential to send advertising messages to the consumer without any loss of data. Nevertheless, the interpretation of odors is subjective and related to individual emotions. The recognition of odors is based on a learning process which starts in the embryonic phase and is composed by perceiving odors and saving them in combination with memories, incidents and emotions forever in the long-term memory. Perceiving a known odor is comparable with an odor-combined memory recall, a process which is known as *proust phenomena*, according to the French writer Marcel Proust, who worked with descriptions of sensory experiences [CD00]. According to this, the sense of smell should be a perfect interface to the human brain to eternally deposit information and recall it by re-perception.

An additional effect to the proust phenomena are color-odor associations, which are constant companions in our everyday life. The food industry is reliant on these associations informing the consumer about taste and smell of products. Heinrich Frieling, a color psychologist and expert for color associations, explains how colors influence consumerism and why different sensory stimuli can complement another [Fri05]. Table 1 shows colors and the connoted odors or tastes as they are used for food packaging and advertising. The tasting process is actually a smelling process because the olfactory epithelium determines flavors through the internal channel to the pharynx. Tasting and smelling are nearly the same perception process, an important fact for food producers. Color-taste and color-odor associations are derived from natural experiences as they are evolutionary fixed in our genes. Therefore it would be contraproductive to produce e.g. blue gummy bears with orange flavor. But comparing different cultures such connotation can diverge, what forces producers to adapt their products for various markets.

Color	Fragrance
PINK	sweet, mild
LAVENDER	sweet, unerotic
MAGENTA	heavy, narcotic, charmingly, sweet
INDIGO	scentless
BLUE	scentless
MINT	juciy, fresh to salty
GREEN	fresh, fragrant, parfume with green fragrance
OLIVE	musty
LIME GREEN	sour, dry, fresh, bitter
YELLOW	parfume, flower
ORANGE	hearty
RED	sweet hefty, hot
GOLD	sweet, good, stunning
OCHER	sourly, neutral
BROWN	aroma, musty
WHITE	scentless
GREY	bad

Table 1: Color-odor associations by Heinrich Frieling [Fri05]

Another exceptional quality of olfactory perception is a direct connection between our smelling system and the hypothalamus, a neurologic area which controls amongst others the autonomic nervous system. This connection allows influencing the viscera by perceiving odors and is responsible that aromatherapy works. Air designers are geared to color-odor associations and the aromatherapeutic effect of fragrances when they are developing perfumes for scent marketing purpose.

For biologists olfactory communication stands for the exchange of pheromones within one species. Pheromones are volatile components causing instinctive reactions within one particular species. They are detected by the *veromonasal* or *jacobson organ*, a part of the nasal septum [Sch06]. The pheromonal communication was only identifiable at animals, but since 2006 it has been proved that also humans communicate via pheromones. For instance, women can identify if a man genetically correlates to the own DNA by checking his body odor including pheromonal information. Pheromones also plays a role when mothers recognize their babys through smelling the baby's body odor [Wat01]. But it is also known that pheromones of other species can influence humans. Parfume components like cibet or musk are animal scents which contain sexual pheromones of the male cibet cat and musk deer. Women find a little bit of such animal scents charming but do not instinctively react on them [Wat01]. Using human pheromones for scent marketing purpose would maybe allow controlling consumers' instincts – a method of persuasive advertising, which isn't without controversy.

3 Scent Marketing

Scent marketing relies on the neuropsychological processing of olfactory stimuli in the human brain. The area of research which analyses neuropsychological effects of advertising and commercial activities on the consumer is called *neuroeconomics*. The term *scent marketing* came up in 2002, defines a subarea of the neuroeconomic research and describes the usage of scents for marketing purpose [Bar08].

The main target of scent marketing is the creation of a pleasant atmosphere for clients. They should stay in stores as long as possible and should enjoy the ambience to accordingly buy more products or raise consumption [MKC95] [Hir09]. In former days bakeries, coffee houses and restaurants often unintentionally worked with scents as attractant. Their chimneys and ventilation systems released enough food aromas that people's mouth watered. Today such shops systematically works with synthetic fragrances to effect similar reactions.

Furthermore costumers should treasure these shopping scenarios as pleasant and relaxed which is due to the already mentioned Proust phenomena. This effect gives store, hotel and service chains an idea of using fragrances as part of their corporate identity. Every time a client perceives the unique parfume of the chain, he or she should recall the shop, the situation and this pleasant atmosphere [MS00]. In 2006 hotel chains like Westin, Sheraton, Omni, Four Points or Hyatt incorporated special fragrances as part of their brand image [Hig06].

Another form of scent marketing is the improvement of the ambience in negatively afflicted locations like hospitals or dental practices. It is a social impact of the western world that people get scared when they are smelling chlorine-camphor and phenol, which give disinfection agents the typical smell. To cope with this issue, dentists tried to improve the atmosphere of their practice by gilding this "doctor's fragrance" to put the patients at ease [LML^{+}05].

The scent of food and goods plays also a relevant role in selling products. Not only food aroma have to meet one's expectations, also other products are liable to associations and connotations. As an example new cars do not smell like plastic and metal. They are sprayed with a oil or leather fragrance to let drivers feel more familiar with them. As the RAC Foundation found out this can lead to overestimation of driver's own capabilities and accordingly to a higher accident risk [NFD05]. But some producers experiment with such associations and perfume their products with unexpected aromas getting a unique selling proposition or jollying their costumers along. Products like scented writing utensils, socks, CDs, USB sticks, papers, etc. have been introduced to market (www.everythingsmells.com, Figure 1). EPAMEDIA, an Austrian public space advertising company has also a controversial approaches and uses unusual advertising forms like smelling posters and illuminated panels [Epa08].

(a) (b) (c) (d)

Figure 1: (a) AromaSocks®, (b) Cocos® flavored USB sticks, (c) Smencils®, (d) Sony Ericsson SO701i

Marketing strategies often includes promotion events as advertising platform and presentation of provider's individuality. Event managers use the latest entertainment forms to compete for visitors and publicity. So scents also became interesting for event managements and nobody wonders that there is a new occupation called *scent DJ* lifting the audience's spirit with fragrance compositions [Emo08]. Affecting ones mood by evaporating special fragrances is also used during social interchange like meetings or class room situations [Hig06]. Researchers noted that some fragrances could contribute to improved cogitation and retentivity [HSB04].

4 Olfactory Technology and Interaction within Pervasive Environments

Information overload is a well known problem of the advertising industry and leads to find other modalities of transmitting information to various target groups. Odors are a new possibility to persuasively present consumers products if they like them. The very subjective interpretation of olfactory stimuli require an adaption of olfactory messages to individual preferences. Therefore scent marketing systems needs to become “smart”. General scent diffusers like scent candles, fragrance lights or other kinds of air fresheners work mechanically, without any electricity. But in the last view years more and more electric scent gadgets were developed and trend to replace their analog pendant (Figure 2).

(a) (b) (c) (d)

Figure 2: (a) Kinlan USB fragrance oil burner, (b) Shenzhen scented USB bracelet, (c) AromaUSB® fragrance diffuser, (d) AIR WICK® FreshMatic

Digital olfactory devices evaporating odors or detecting volatile components can be defined as *olfactory technology* and act as *olfactory displays* or odor sensors. Such technology increasingly implements more interfaces to allow integrating it in and controlling it by pervasive environments. So olfactory technology is becoming “intelligent” and dynamic adjustable for individual purpose.

4.1 Sending Olfactory Information

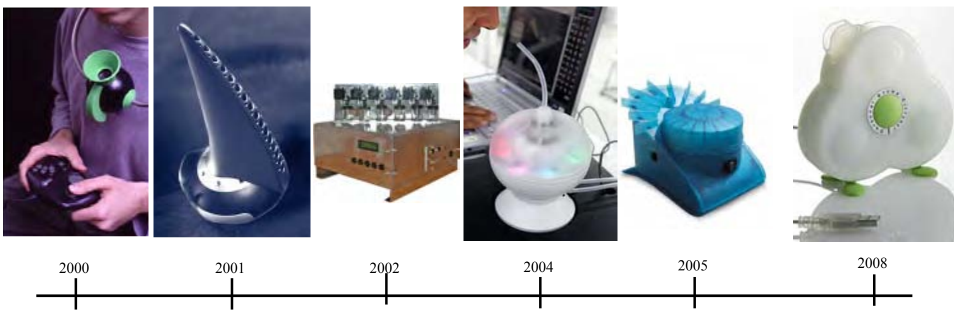
Most scent systems in salesrooms are integrated in the air condition, but there are also heat-based standalone devices. First commercially available air design systems could not adjust running time, fragrance or scent volume, they continuously run and lead to a flood of fragrancy. Modern systems offer an adjustment of time and volume as well as changing between various fragrances (Figure 3).

Not only advertising takes advantage of odors, olfactory communication receives more and more attention in human-computer interaction. Today digitally controlled odor diffusers are not only applied for advertising purpose, they are increasingly used as ambient indicator like an olfactory display in HCI-systems (Figure 4). For instance, Keye reported on an ambient olfactory reminder system [Key01]. An integration of an augmented reality

(a) (b) (c) (d) (e) (f) (g)

Figure 3: (a)(b)(c)(d) heat-based systems of Aromea, (e) ventilation-based aroma cubes of Sensarama, (f)(g) EnviroScent

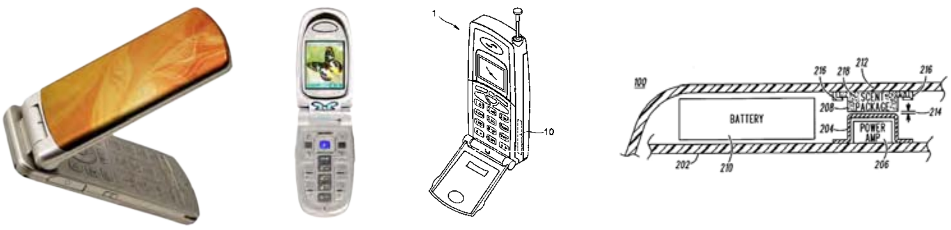
application with an odor machine to improve on the the reality experience is presented in [Ems06]. NTT Communications have developed a smell machine called *Aroma Geur*, laying the path to the first olfactory emails in 2004 [Com07]. This device was also used to create an ambient smell when listening to Tokio FM. In 2005 TriSenx launched their *ScentDome* to enable websites emitting scents [Tri05]. In the meantime, telecommunication industries have also found the olfactory information channel as useful medium and market the first scenting mobile phones [Sof07][Mot07]. The special smoothness of olfactory interaction spaces was the central subject of the *Space-of-Scent*-project realized by Haque Design & Research in 2002.



(a) (b) (c) (d) (e) (f)

Figure 4: (a) AromaJet Pinoke, (b) DigiScent iSmell, (c) FH Hagenberg SmellBox, (d) NTT Com Aroma Geur, (e) TriSenx Scent Dome, (f) Osmooze Personal Diffuser

For the telecommunication industry smell has been successfully introduced as new sensory modality for interactions between human and mobile devies. The first "smelling" mobile phones were placed on market in 2008 (Figure 5). The Sony Ericsson SO701i is scented with an aroma therapy fragrance to support relaxing during stressful phone calls. To satisfy different preferences the mobile phone is available with 8 different fragrances, which can also be useful for advertising purpose and tagging personal things like mobile phones with corporate scents. The Hyunday MP280 integrates an individual refillable scent diffuser which acts as “smelling tone”. Samsung and Motorola holds also patents for smell phones [Sam06] [Mot07]. German inventors have already patented a mobile phone with a smell chip which allows sending and receiving *smell messages* [IH08]. These mobile devices could be the future of mobile advertising – they are offering a new method to send not only informative but also emotional advertising messages.



(a) (b) (c) (d)

Figure 5: (a) Sony Ericsson SO701i, (b) Hyundai MP280, (c) Samsung, (d) Motorola Smell-o-Phone

Mentioned odor diffusing facilities provide using odors for static as well as mobile smart installations. Current scent marketing systems treat each costumer like the others and don’t care about their individual preferences. Making such systems intelligent allows to respond to these preferences as well as the consumers’s psychological state like aroma therapist do. The combination of emotion recognition, referring to *emotional computing* [Pic97], and the usage of smart olfactory technology can be a powerful instrument for pervasive advertising.

4.2 Sensing Olfactory Information

Not only the output of olfactory information is an increasing subject for information technology research, but also using volatile substances as input for digital communication increasingly becomes useful. Gas sensor arrays and electronic noses are especially used in forensic investigation for the detection of explosives and medical science to diagnose diseases like cancer. Today they are increasingly used to control digital systems. For instance the Japanese *Hanahana*-installation allows manipulating flower-animations by ten different perfumes. In 2005 Wyszynski, Yamanaka and Nakamoto already sent an recorded odor by email and reproduced it at the receiver [WYN05].

Researchers of the Austrian Konrad-Lorenz-Institute are currently developing a system to recognize individuals by their body odor which represents the individual DNA like a volatile fingerprint [Pea07]. Body odor is the volatile state of sweat whose components are genetically influenced. Emotions like fear can also manipulate the sweat composition and contribute to the production of cold sweat [CKL06]. According to this, body odor has the potential to become a new data source for intelligent systems to recognize individuals as well as their emotions.

Recognizing and manipulating emotions are basic instruments of advertising. The olfactory information channel allows both detecting emotions by body odor analysis and causing emotions by aroma therapeutic scents. Maybe in the near future smart advertising systems will be able to react to individual emotional states and manipulate them at the same time. Such bidirectional communication via odor acts within unconventional zones of interaction. Each odor-emitting human or thing is surrounded by a definite waft of scent like an olfactory aura defined as *Olfactory Interaction Zone (OIZ)* in [EF09]. This zone can be dynamically extended from square meter size to spare kilometer scale. Consumers often are able to smell aromas of near bakeries or restaurants over many meters leading them from the street in their salesroom like an *olfactory direction sign* [Ems09].

Sensing odors could be a useful instrument for advertisers to find out more information about their clients, especially emotional states. Another valuable data are individual odor preferences which could be examined by identify their body odor or personal perfume. Furthermore such preferences would refer to color preferences which could be useful for e.g. fashion style suggestions.

5 Conclusion and Discussion

Evolutionary we are used to communicate via odor, but digital communication still does not allow this. With scent marketing advertisers have initiated the integration of olfactory interaction in modern communication facilities. Today scent advertising is still a simple form of perfuming rooms and flavoring goods to make them more pleasant for costumers, but odor is more than that. It is a medium for unconscious messages, manipulates mood and behavior and recalls often forgotten memories. But the disability of avoid smelling can rapidly lead to olfactory information overload. Therefore current issues of air designers and also computer scientists include the improvement of evaporation and intensity control, advancing psychological investigation of the olfactory perception and being responsive to consumer's individual preferences.

The detection of body odors will enable identifying and tracking of human beings as well as recognizing their emotional state by one input data stream. But computer scientists still avoid using the olfactory communication channel, probably because olfactory technology is not sophisticated yet. But people suffer from visual and auditory information overload, so the olfactory communication channel gets more and more important for interaction between human and human as well as human and things.

6 Acknowledgments

This work is supported under the FFG Research Studios Austria programme under grant agreement No. 818652 DISPLAYS (Pervasive Display Systems), and under grant FACT (Flexible Autonomic Context Technologies) of Siemens AG, CT-SE 2, Munich.

References

[Bar08] Franziska Bartzos. Duftmarketing - Eine spezielle Form des Neuromarketings. Diploma thesis, University of Vienna, Vienna, October 2008.

[CD00] Simon Chu and John J. Downes. Odour-evoked Autobiographical Memories: Psychological Investigations of Proustian Phenomena. *Chemical Senses*, 25:111–116, 2000.

[CKL06] Denise Chen, Ameeta Katdare, and Nadia Lucas. Chemosignals of Fear Enhance Cognitive Performance in Humans. *Chem. Senses*, 31:415–423, March 2006.

[Com07] NTT Com. Movie Enhanced with Internet-based Fragrance System. `http://www.ntt.com/`, 2007. 2008-09-02.

[EF09] Bernadette Emsenhuber and Alois Ferscha. Olfactory Interaction Zones. In *Pervasive 2009 Adjunct Proceedings*, Nara, 2009. Pervasive 2009.

[Emo08] Emotion. Ein dufter Job. *Emotion*, pages 94–99, June 2008.

[Ems06] Bernadette Emsenhuber. Integration of Olfactory Media and Information in Pervasive Environments. Technical report, Univ. of Munich, Dep. of Computer Science, Media Informatics Group, March 2006.

[Ems09] Bernadette Emsenhuber. *Das olfaktorische Medium - Die Integration olfaktorischer Information in die Mensch-Maschine-Kommunikation*. PhD thesis, Universität für künstlerische und industrielle Gestaltung Linz, 2009. *to appear*.

[Epa08] Epamedia. Duftendes City Light für Donna Karan. `http://www.epamedia.at/`, October 2008. 2009-05-28.

[Fri05] Heinrich Frieling. *Farbe hilft verkaufen*. Hansen Muster-Schmidt, Northeim, 2005.

[Hig06] Michelle Higgins. Figs? Coconut Sunscreen? Hotels Choose Their Scents. *The New York Times*, April 2006.

[Hir09] Rainer Hirt. Multisensory.de - Der Blog zum Thema Sensory-Branding. `http://multisensory-de.blogspot.com/`, May 2009. 2009-05-27.

[HSB04] Rachel S. Herz, Corrente Schankler, and Sophia Beland. Olfaction, Emotion and Associative Learning, Effects on Motivated Behavior. *Motivation and Emotion*, 28:363–383, December 2004.

[IH08] Inside-Handy. Deutsche Erfinder entwickeln Dufthandy-Anwendung. `http://www.inside-handy.de/news/11777.html`, May 2008. 2008-09-02.

[Key01] Joseph N. Keye. Symbolic Olfactory Display. Master's thesis, MIT, Massachusetts, May 2001.

[LML+05] J. Lehrner, G. Marwinski, S. Lehr, Johren P., and L. Deecke. Ambient odors of orange and lavender reduce anxiety and improve mood in a dental office. *Physiology & Behavior*, 86:92–95, 2005.

[MKC95] D. J. Michell, B. E. Kahn, and Knasko S. C. There's something in the air: Effects of congruent or incongruent ambient odor on consumer decision making. *Journal of Consumer Research*, 22:229–238, 1995.

[Mot07] Motorola. Communication device having a scent release feature and method thereof. `http://www.uspto.gov/`, April 2007. 200-09-02.

[MS00] M. Morrin and Ratneshwar S. The impact of ambient scent on evaluation, attention and memory for familiar and unfamiliar brands. *Journal of Business Research*, 49:157–165, 2000.

[NFD05] Sue Nicholson, Sarah Forrow, and Kevin Delaney. The Scent of Danger. `http://www.racfoundation.org/`, June 2005. 2008-09-02.

[Pea07] D. J. Penn and et al. Individual and gender fingerprints in human body odour. *Jour. of the Royal Society Interface*, 4(13):331–340, April 2007.

[Pic97] Rosalind W. Picard. *Affective Computing*. MIT Press, Cambridge, London, 1997.

[Sam06] Samsung. Mobile phone having perfume spraying apparatus. `http://www.uspto.gov/`, March 2006. 2008-09-02.

[Sch06] J.W. Schumann. Pheromone, Körpergeruch und Partnerwahl. *Der Gynäkologe*, 9:731–740, 2006.

[Sof07] Softpedia. Aromatic Personality with Sony Ericsson SO703i. `http://news.softpedia.com/`, July 2007. 2008-01-26.

[Stö98] Anja Stöhr. *Air-Design als Erfolgsfaktor im Handel*. Deutscher Universitäts-Verlag, 1998.

[Tri05] TriSenx. ScentDome. `http://www.trisenx.com/`, 2005. 2009-01-13.

[Wat01] Lyall Watson. *Der Duft der Verführung*. Fischer Verlag, Frankfurt am Main, 2001.

[WYN05] Bartosz Wyszynski, Takao Yamanaka, and Takamichi Nakamoto. Recording and reproducing citrus flavors using odor recorder. *Sensors and Actuators B: Chemical*, 106(1):388–393, April 2005.

eMir: Digital Signs that react to Audience Emotion

Juliane Exeler, Markus Buzeck, Jörg Müller

{j.exeler, m.buzeck, joerg.mueller}@uni-muenster.de

Abstract: In this paper we present eMir, digital signage (public electronic displays) that show human faces which react to audience emotion. Using a camera installed at the sign, the system observes the audience and detects whether someone watches the display via face detection software. The face detection is able to classify facial expressions and determine gender. This information is used to let a human character on the screen react accordingly and encourage interaction with the face/sign. The system has been deployed for one month on a digital sign in a university building. We present experiences with the system, our findings from the collected interaction data and results from interviews with eight users.

1 Introduction

Public electronic displays become more and more important as one can conclude from the increasing number of displays installed in public places and buildings. Digital signage is a cheap and easy way to present customizable information. Shopkeepers are increasingly installing mid- to large-scale displays in order to attract the customers' attention and advertise new products or special offers. But there are also places where displays are installed for non-commercial purposes such as information or entertainment. These places are for example waiting areas, e.g. at the doctors or at the citizen bureau. Studies have shown that public displays are often not able to attract the passerbys' attention except some displays installed in waiting areas[1]. Although there are style guides for display size, height, color composition of the content and the content itself[2] to make displays more perceivable many digital signs are still not noticed. We started exploring how to make displays more perceivable and enjoyable that people get attracted by it. In a field study we presented content which was rated by the view time of the users collected with face detection software[3]. This was a kind of unconscious, indirect interaction because the user didn't know that he was influencing the content to be shown. In our current study we try to find out, how videos of human faces that mirror the users' emotion and gender can attract the users attention and lead to direct interaction. As a first step, we want to examine how users react to our electronic mirror system (eMir) and if the people try to interact with it. We think that in future interaction with digital signs is one way to make digital signage more attractive and perceivable.

2 Related Work

Experiences from our previous work with digital signage have shown two major shortcomings of public displays this project is supposed to overcome. First we observed the effect of Display Blindness[1]. People tend to ignore public displays intentionally or subconciously, which is influenced by negative expectations towards contents. Second we often experience that people react very reserved towards digital signs. In particular when digital signage is equiped with sensors measuring context and audience attention, people often do not feel comfortable[3]. This is partly due to privacy concerns, but often there are still negative feelings when the rational concerns are dispelled and people trust the person who installed the system.

Against this background we propose that eMir can have positive effects on user acceptance in several ways. Evidence from fields such as cognitive, evolutionary, and developmental psychology, as well as cognitive neuroscience, has pointed to the special nature of human perception of faces. Faces begin to be seen as a separate class of objects within the first six months of life[4]. Christian and Avery presented the Smart Digital Kiosk Project which is an electronic information display showing a human face that looked at the users and talked to them[5]. This project received enthusiastic feedback from the users. Faces tend to attract our attention. So we propose that a public display showing a life-sized human face with natural movement and expression will catch the people's eye in the first step. In the next step we propose that the presence of a calm and friendly face smiling back at the user in an unobstrusive way might contribute to a comfortable feeling. And finally there is the aspect of interaction with the face on the sign, that might appeal to peoples curiosity. Shawney, Wheeler and Schmandt[6] suggest to use unobstrusive means for sensing user intention and activity patters for interaction in transitional spaces. They investigated proximity and movement as an interface to maintain a casual and natural interaction. By means of simulating non-verbal interpersonal communication through facial expression we want to explore a similar approach. Advanced face detection technology allows us to make use of an even more subtle form of interaction.

3 A Concept for interactive Digital Signage

According to the capabilities of the face detection software, we define five emotional states of users the system differentiates: a neutral state, happiness, sadness, angriness and surprise. For each of these categories we also consider the gender as being either undefined, male or female. So we end up with 15 different scenarios that might occur with a neutral state and undefined gender as the default state of the system. The current state is determined by the modes of both emotion and gender detected in the camera frames within a certain time slot. According to the current state a small video sequence is selected that fits the situation. Each interaction, which we consider as a variation from the default state, is initiated by the user. The digital sign reacts passively by reflecting the same emotion through the facial expression of a human of the same gender. Happiness shall be encouraged by a happy face. A sad face shall receive empathy by another sad face. An angry face on the screen shall make angry users laugh. The same applies to surprised faces. A user with a neutral facial expression will not be involved into any interaction by the sign. The face on the sign will remain in a default neutral state, looking around the hallway quite bored.

4 Implementation

The prototype consists of a digital sign installed in a university institute with approx. 60 employees. It is located in the hallway and is frequently passed by students and employees. Before being used for this project the sign was used for the university information system iDisplays[7] which is installed on several other signs in the building. Audience emotion is measured via camera installed on top of the sign. We use the Real Time Face Detector from Fraunhofer IIS[8] to analyze the video stream. This software is able to detect multiple faces within the camera image during runtime. It also determines gender and emotion from the facial expression. Depending on the emotion and the gender that has recently been detected most frequently, a small video sequence is selected that shows a face with similar attributes. Therefore we produced videos each showing a person in a portrait close-up. We tried to position the persons and cut the videos in a manner that allows us to play the videos in loops and one after the other without breaks looking as natural as possible (see figure 1). As these videos need to have a minimum length to reveal a certain emotion naturally,

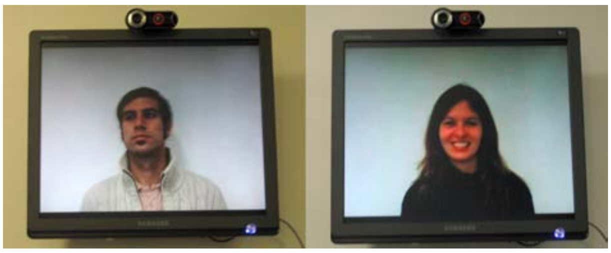


Figure 1: Faces presented on the eMir display.

we do not interrupt a video once it has started. As a consequence the system may react with a delay with maximum length equal to the length of the current video. Whenever a new video is started, either a default video or an interaction video is triggered by the face detection, the attributes of the video are stored in a database in order to be able to analyze user behavior and interaction.

5 Data Collected

During the deployment phase of one month (24/7) about 111000 videos had been played. For each displayed video an item containing an id, gender and emotion information and the timestamp was inserted into a database. We can neglect most of the database entries for our analysis because in these cases default videos have been played. Default videos have been displayed whenever the face detection software could not identify any human face in front of the sign. Then a video with the face of a male person and a neutral emotion was shown but was inserted into the database with the attributes gender is undefined and emotion is neutral. The collected data allows us to investigate how passerbys react to the eMir system and if there is any interaction between the digital sign and a passerby who can then be considered a user. This interaction is recorded as sequences of database items representing non default videos.

In 569 cases a non default video was played. We grouped the collected data into 125 sequences of videos that indicate some interaction has taken place. An interaction sequence ends with a default video shown, a new interaction sequence starts when a default video is succeeded by a non default video. Thus as a first result we can state that during the deployment phase every day (including the weekends when there is no activity in the building) there were five interaction sequences in average. This is more than we expected from our last study[3], where we got some negative feedback by users concerning display contents and privacy concerns. Having a closer look at the duration of each interaction sequence we found out that in average every interaction takes about 30 seconds. Hence we can say that the shown faces attract attention and actually stimulate interaction. Depending on the duration of the individual displayed videos such an interaction comprises two to four videos. Classifying the 125 interaction sequences into four subsets according to the number of videos we see that in most cases (51 times) two, three or four videos build one interaction sequence. For 44 times a sequence of interaction only consists of one item. Here we have to note that the face detection software also detects a face which is just turned towards the digital sign, regardless of the fact that it may not look at it. In 19 cases an interaction sequence comprehends five up to ten video items. For interactions sequences with more than ten video items eleven blocks were identified (see table 1).

Table 1: 569 items of non default videos could be assigned to 125 sequences of interaction

Interaction sequences with...			
1 item	2 - 4 items	5 - 10 item	more than 10 items
44	51	19	11
$\sim$ 30 seconds each interaction			

The analysis of the collected data shows that there has been a trend in the frequency of interaction sequences. Splitting the deployment phase into two parts of equal length the result is that in the first phase there were 33 (75%) interaction sequences with one video item, 30 (58%) with two up to four items, 15 (78%) with five up to ten items and six (55%) with even more than ten items. This indicates that the interest for the eMir system decreases with time. One reason might be that the default video is always the same. Different default videos with alternating faces might also attract the passerbys' attention in the long run.

6 Interviews

After running eMir for one month, we conducted semi-structured interviews with 8 employees and regular visitors of the institute. The interviews were partially transcribed and evaluated using affinity analysis[9]. Perception and acceptance of the system was mixed. The presentation of a human face on the public display was appreciated by seven users. These stated that they found it funny and amusing. This was mainly due to the fact that the presented faces were familiar to the users. Five users said it made a big difference if they knew the face or it was a foreign face. ”The first time it was appealing to see him there. That made me stop at the sign to see what happens.” (U2) ”It would be curious to see more familiar faces.” (U7) Two interviewees stated that any human face attracted attention. We also wanted to investigate if the presence of a human face on a screen compared to textual inforation conveys a different atmosphere. Five interviewees said that the system had a positive effect on the atmosphere. ”It enlightens the mood.” (U4) Drawing a comparison to other systems which were previously installed on the same device, one user said: ”There is an increase in the attraction of user attention from the iDisplay information system, over video presentation to faces being presented. It is appealing on an emotional level.” (U2) ”It is a warmer, more cordial atmosphere, as if there was just information running on the screen.” (U5) When it comes to interaction, four out of eight interviewees stated they had intentionally interacted with the system. ”It is funny to stop for some minutes in front of the display and to try it. It is exciting and nice playing with it.” (U4) ”I tried it. I was curious to see more images.” (U7) Two users did not recognize it was an interactive system. They were not stimulated to watch the display long enough to evoke any reaction from the system. Another two users deliberately did not interact. One of these two gave her negative attitude towards cameras as a reason. ”I don’t like standing in front of cameras. I feel like being observed.” (U3) Moreover six out of eight interviewed users said to have witnessed or have personally been involved in situations in which several people were interacting with the system and with each other. There were discussions of people altogether exploring the functionality of the system. ”We were standing there several times and tried to imagine how it works. We were talking about it, because we found it really funny.” (U2) Figure 2 illustrates a similar situation.

7 Conclusion

In this paper we have presented eMir, a digital signage system that shows human characters reacting to audience emotion. The system was deployed for one month and evaluated through analysis of logging data and interviews with users. We wanted to investigate how such a system is accepted by the user, whether it stimulates interaction and whether the mere presence of a human face on a screen influences the atmosphere. The analysis shows that most users react with curiosity to the display and try to interact. This effect is even stronger when the presented face is familiar to the user. Interaction sequences took place every day and many users spent some time to trigger the different emotions to be shown. However there are also some users who reject such a system because of distrust towards

Figure 2: People observing the face on the eMir display.

the technology. In particular the camera being visible and obviously running is a big issue, although the sign suggests for what it is used. To those users who do not have any privacy concerns, the face on the screen looking around the hallway is funny and effects their mood in a positive way. At least compared to pure information being presented on the sign the system appeals to the users in a more emotional way and thus more effectively. Perhaps this could be an approach to make digital signage more perceivable in general.

References

[1] J. Müller, D. Wilmsmann, J. Exeler, M. Buzeck, A. Schmidt, T. Jay, and A. Krüger. Display Blindness: The Effect of Expectations on Attention towards Digital Signage. In *Proc. of Pervasive 2009*

[2] E. Huang, A. Koster, and J. Borchers. Overcoming assumptions and uncovering practices: When does the public really look at public displays? In *Proc. of Pervasive 2008*

[3] J. Müller, J. Exeler, M. Buzeck, and A. Krüger. ReflectiveSigns: Digital Signs that Adapt to Audience Attention. In *Proc. of Pervasive 2009*

[4] C. A. Nelson. The development and neural bases of face recognition. In *Infant and Child Development*, Volume 10, Issue 1-2, pages 3–18, 2001.

[5] Andrew D. Christian and Brian L. Avery. Digital smart kiosk project, In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems, 1998*

[6] N. Shawney, S. Wheeler, and C. Schmandt. Community Aware Portals: Shared Information Appliances for Transitional Spaces. In *Personal and Ubiquitous Computing*, 5 (1):66-70, 2001.

[7] J. Müller, O. Paczkowski, and A. Krüger. Situated public news and reminder displays. In *European Conference/Symposium on Ambient Intelligence*, pages 248–265, 2007.

[8] C. Küblbeck and A. Ernst. Face detection and tracking in video sequences using the modified census transformation. In *Image and Vision Computing*, 24 (6):564–572, 2006.

[9] H. Beyer and K. Holtzblatt. Contextual Design : A Customer-Centered Approach to Systems Designs, 1997

En passant Coupon Collection

A. Ferscha, W. Swoboda and Ch. Wimberger
Johannes Kepler University Linz, Department of Pervasive Computing
Altenberger Straße 69, 4040 Linz, Austria
{ferscha, swoboda, wimberger}@pervasive.jku.at

Abstract: Spontaneous interaction in public places has evolved as crucial concern in interaction design, particularly in the domain of public advertising. Implicit interaction is a mode of spontaneous interaction in which the user is not attentively or explicitly expressing input to a system, but the system from observing the user, and considering all the available information that describes the user situation, is used to generate input - and behaves accordingly. We have investigated implicit interaction as a potentially effective means to design for spontaneous interaction in the public. An *en passant* coupon collection system, CCS, has been developed to implement an implicit interaction based pervasive advertisement process. CCS is based on scanning and identifying Wi-Fi hotspots by their BSSID, and has been implemented for internet enabled smart phones. Scenarios demonstrate how the system collects bonus points, rewards, sales coupons, reward coins, etc. by users just walking by points of interest like stores, bus stops, schools or offices. CCS neither not being reliant to a preconfiguration of the Wi-Fi infrastructure, nor its use for wireless data communication, encourages for a parasitic use of existing Wi-Fi networks (already available in public places like shopping malls, airports, train stops, stations or city centers) as a pervasive advertisement system.

1 Implicit Interaction and Public Displays

Pervasive Computing has developed a vision where the "computer" is no longer associated with the concept of a single device or a network of devices, but rather the entirety of situative services originating in a digital world, which are perceived through the physical world. It is expected that services with explicit user input and output will be replaced by a computing landscape sensing the physical world via a huge variety of sensors, and controlling it via a plethora of actuators. The nature and appearance of these services will change to be hidden *"in the fabric of everyday life"*, invisibly networked, and omnipresent. They will greatly be based on the notions of context and knowledge, and will have to cope with highly dynamic environments and changing resources.

Interaction with such computing landscapes will presumably be more implicit, at the periphery of human attention, rather than explicit, i.e. at the focus of attention: *"we will be able to create (mobile) devices that can see, hear and feel. Based on their perception, these devices will be able to act and react according to the situational context in which they are used"* [Sch00]. Accordingly, as from the users point of view, *"implicit human computer interaction is an action performed by the user that is not primarily aimed to in-*

teract with a computerized system but which such a system understands as input." [Sch00]. Implicit interaction is hence based on two main concepts: perception and interpretation. Perception is information gathering about the environment and situations, usually involving (technological) sensors. This information is generally provided implicitly to the system and displayed naturally to the user. Interpretation is the mechanism to understand the sensed data. Conceptually, perception and interpretation when combined are described as situational context. Clearly, *pervasive advertisement*, like e.g. advertisement via *public displays*, is a case demanding for principles of implicit interaction.

Public display systems (PDS) [FV02], often referred to as *Out-of-Home Media Systems*, by gaining from technological advances in display technology, but also wireless communication technologies, have emerged as such a landscape, provisioning and controlling content related to *context*. "Context" refers to any information describing the situation of an entity, like a person, a thing or a place, so the content being displayed in a PDS (in e.g. shopping malls, bus and train stations, airports, offices, schools, public places, etc.) can be adapted wrt. to all these types of information.

While user interaction with PDS was not intended over their first generations of existence, the upcoming electronically operated and digitally controlled DSPs have encouraged a variety of solutions for *explicit interaction* [SD08], [BRS05], [BBRS06], [VB04] mostly being based on mobile phones as the interaction device [HR08], [WFGH99], [TCF+07], the primary types of interaction being navigation and browsing [RYL+07] [WK06].

Interaction with such PDS landscapes will presumably be more implicit, at the periphery of human attention, rather than explicit, i.e. at the focus of attention.

With this paper therefore we address *implicit interaction* with PDSs, addressing the issues of a seamless, non-disrupting, yet situation aware modes of interaction in public advertisement. We have developed a Coupon Collection System (CCS) based on (even non-cooperation) scanning for public Wi-Fi hotspots, and the mapping of unique hotspot IDs to geo-referenced advertisement services. The Coupon Collection System is developed in Section 2, and the roles of stakeholders engaging in a pervasive advertisement service system are related in Section 3. The corresponding software architecture and implementation details are outlined in Section 4. Section 5 exemplifies the use of the CCS system with smart phones as context aware CCS clients. Conclusions are drawn in Section 6.

2 The CCS Coupon Collection System

CCS is divided into three major components. There are stationary components, which are installed on places where a user should be able to collect coins. The users carry the mobile components, which detect the stationary components automatically, without any explicit interaction. Finally, a server is needed, to manage the content which will be delivered to the mobile components.

2.1 Stationary Component - WI-FI Access Points

The major task of a stationary component is to be detected by the mobile component in a specific radius around its geographical location. One option to do that, without provoking an explicit interaction by the user, is to broadcast a radio signal with limited range Fig. 1. Of course, if there are two or more stationary components with overlapping radius, the interference between that signals has to be considered. To distinguish between different stationary components, each stationary component also needs a unique identifier.

There is an existing and also very common technology which meets the requirements, the 802.11 wireless LAN access point. A functionality provided by access points is the beacon broadcast - also known as SSID-Broadcast-, a specific package, sent at certain intervals. Usually the beacon broadcast is meant to make the access point known to wireless client devices in vicinity, therefore the beacon packages are not encrypted. Such a package contains, among other information, the MAC address of the access point, the name of the network, the time interval between the broadcasts and the encryption of the network [IEE07]. Using the beacon package, a client device is also able to evaluate the signal strength. Important for this system is mainly the MAC address of the access point for the use as a unique identifier. Usually, the interval between the beacons is set to 100 milliseconds. That leads to a very fast possible reaction time for the client. Other technologies are much slower, for example Bluetooth (with query scan interval set to default, the reaction can take up to 2.56 seconds [SIG09]). Also the frequency bands which are used by 802.11 Wi-Fi networks (2,4 and 5,18-5,7 GHz) are usable without a license.

Figure 1: access points with SNR visibility: (from left to right) (i) one AP, (ii) multiple APs, (iii) multiple APs with overlapping SNR visibility, (iv) Possible Distribution of Wi-Fi Access Points in public space

The interference management is provided by the 802.11 standard and supported by the hardware of access points [Ohr03]. That fact is important for the scalability of the system, because thereby overlapping domains are no longer a problem. The effective range of wireless LAN depends on factors as the transmission power, the used antenna and of course the consistency of the surroundings. As a reference value, 802.11b standard access points can reach a radius of up to 100 meters. To limit the radius of the access points it's either possible to decrease the transmission power, which is does not work on every access point, or to filter signals under specific signal strength on the side of the mobile component [Ohr03]. If the signal strength of an access point is under certain value, the mobile component will just ignore it in first place or doesn't redeem an BSSID for a coin when connected to the server. This feature doesn't exist in the prototype by now, but it's a likely target for further development. As an disadvantage, specification of a collection area for a specific virtual coin is not very accurate. It's not possible to define precise boundaries which separate one collecting area from a non-collecting area or an area of another coin. For operating the CCS one has to be aware of this disadvantage and compensate it by clever distribution of access points and mapping of coins.

In areas where there is already a wireless LAN infrastructure , the operator can call on existing access points. In fact, every access point with enabled beacon broadcast can be used to act as a stationary component in this system, even if it doesn't belong to the operator of the system.

2.2 Mobile Component - Smart Phone

The requirements a mobile component has to meet to work in this system are matched by many of the current and future generation smart phones. They provide the required connectivity like wireless LAN for collecting BSSIDs and UMTS for connection to a server on the web, more than enough computing power and storage, as well as high resolution displays and keys or even touch-screens for interaction with the software. However, not every smart phone is equally well qualified for the use in this system, even if it meets all hardware requirements.

The mobile component of the system is implemented as software on a smartphone. If the user wants to start collecting, he has to activate the software. From that point of time, until he switches it off again, the software collects coins.

An important point which has to be considered is the increased power consumption. Without a program running, a smart phones switch into an energy saving mode, where all but the most basic functions are deactivated. While the software searches for access points, the phone will not be able to do that, which leads to a significantly reduced operation time.

The prototype for the mobile component of the system was developed for the Apple iPhone, a quite innovative and widespread piece of technology.

2.3 Server

After the mobile component collected for a while, its database has stored some MAC addresses. It's not known yet what MAC address belongs to what coin. So at certain points in the program the mobile phone has to connect to a database, where a mapping between MAC addresses and coins happens. Keeping this database on the phone itself leads to various problems. How to keep the database up to date? What can be done to keep it save from malicious manipulation? In addition there might be a lack of storage space on the device. Therefore the database is not located on the phone itself but on a web-server. As a result the mobile device needs internet connectivity, which should be no problem in times of UMTS and wireless LAN hotspots. As the prototype exists now, it's not possible to regard the time when a token was collected. But it's an idea for future versions to store timestamps with collected BSSIDs and make the type of collected coin not only dependent on the BSSID but also on the time when it was collected.

3 Stakeholders and the CCS Service Architecture

For a pervasive advertising scenario, the individual stakeholders engaging in CCS services and their roles are sketched in Figure 2(left). The first stakeholder is the operator of the system. It's the one who sets the system up and maintains it. The operator also contracts with the businesses the system advertises for. Those businesses are the second stakeholder. The third party are the customers of the businesses, the ones the advertisement is addressed to. This third stakeholder is in possession of the platforms for the mobile components of the system, the smart phones. The operator owns the right on the software. The property situation between operator and businesses - who owns the access points and who runs the server - can be different from scenario to scenario.

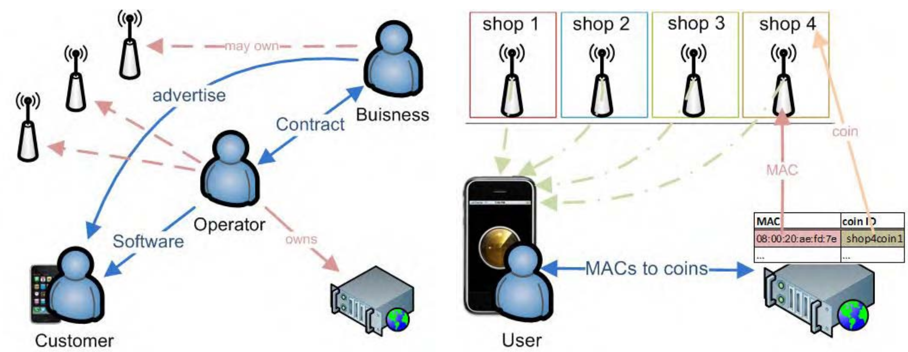


Figure 2: left: Role and Interrelationship of Stakeholders; right: System Architecture and Mapping

3.1 Setting up a CC System

A good way to describe the architecture of this system is to go through a setup process. Therefore a short scenario has to be explained.

In this scenario there is a section of a street with a number of shops or other businesses. Each of this business wants to participate in the system. For that reason they enter into a contract with the operator of the system. The later user of the system, the customer, should earn coins by walking down this street. When he gets near a shop he gets a coin of this shop.

For that to work, the operator of the system has to install one wireless access point, one of the stationary components, inside or near each shop. He doesn't have to add this access points to a network, he just has to switch them on and configure them. However, the operator might want to provide additional services like internet connectivity over those access points, either for the shops or for customers. Doing that wouldn't interfere with the functionality of the system described here. As mentioned before he can also use existing access points, for example if the shopping street already possesses a wireless LAN infrastructure (Fig. 2 right).

After the access points are set up and running they have to be mapped to coins. That happens by storing their MAC address and a reference to a coin to the database on the server. Storing additional information about the access points, like the GPS coordinates, the model or the range, might also prove valuable. The operator can also map one MAC address to more coins and vice versa. By that the user can get some different coins out of one single access point.

The only thing missing now are the mobile components. It's the operator's task to distribute the software, which converts the user's smart phones into mobile components of the system. After the software is installed, the user just has to run it to collect coins. The system is set up and running at this point. It can be easily scaled by adding new mobile and stationary components.

3.2 Interaction between Components

The interaction between the components is quite simple. The software on the smart phones just scans for the beacon broadcasts and stores the MAC Addresses of the access points. As soon as the user wants to check which and how many coins he collected so far, the software connects to the server and gets the proper coin for each MAC address. If the user collected enough coins of a kind to redeem them, for example at a shop, the software connects to the server again to get the information needed to cash the benefit. That could be, for instance, a picture of a bar code [Luf].

4 The CCS Software Architecture

In order to achieve the best possible user acceptance, a clear and straight-forward user interface has to be presented to the person using the software on the smart phone [STH07]. In addition, the software has to be easily operated on the go, because the handling of the coupons after collecting them should be at least as simple as it is with traditional coupons - even for less tech-savvy people. In order for this, we first introduced the metaphor of a "coin" to represent a coupon. A coin in its appearance should offer all the used modalities of handling it, like keeping, saving, or droping it. to manage coins in the user interface of a smart phone, the direct manipulation of visualizations of coins appeared inevitably necessary. Therefore, the Apple iPhone was chosen to purposely serve as the platform for the prototype client. The programming language used is Objective-C as it is the main language for developing on the iPhone platform.

For optimal scalability, even with thousands of collected records, SQLite is used for storage on the client. SQLite is a lightweight and server-less library that implements a transactional SQL database engine. After installing the client software, no additional setup is required to start collecting coins. The entire collecting and redeeming process is completely anonymous.

The server software was developed in PHP with a MySQL database backend. It basically understands three different commands from the client and responds accordingly after issuing some SQL commands to the database. A web frontend can be easily developed to allow third parties to add their coins and access point addresses.

4.1 Scanning for Wi-Fi 802.11 Beacon Broadcasts

Immediately after touching the collector icon on the home screen, the client starts and finds itself in a discovery mode listening for Wi-Fi 802.11 beacon broadcasts. All that is visible to the user at that time is a nearly empty display with an icon displaying a radar screen in the center. In order to keep the system utilization as low as possible, animations are disabled during collecting. The act of finding nearby access points happens in a completely passive way (without sending any Wi-Fi 802.11 packets) thus battery life can be kept at a maximum.

The client continuously keeps an updated list of available access points at the current location. As soon as a new access point appears, it is added to the list and the MAC-address, along with a timestamp, is stored into the database. When an access point in the list cannot be reached anymore, it is removed from the list.

At this point the client doesn't know if the collected MAC-addresses will be exchanged to collected coins in the future. In reality most of the collected addresses will be ignored by the server and only a few will transform into coins. Hence the list of found addresses can potentially be very long, especially if the user keeps collecting over a long period of time without checking for collected coins.

The process of collecting can be interrupted by the user at any time by closing the application and can be continued by simply restarting the application again. By touching the radar screen with his fingers, the user can check if any coins have been collected for him. At this point an internet connection is established and the server is contacted.

4.2 Client-Server Communication

Only at a few certain moments, the server has to be contacted. On one hand when collected MAC-addresses should be exchanged for coins and on the other hand when these coins should be exchanged for the final coupon. To assure the best possible interoperability, the entire communication is based on the standardized Hypertext Transfer Protocol (HTTP). So even in environments with firewalls and proxy servers there should be no problem communicating with the server through the internet. If no internet connected Wi-Fi 802.11 hotspot is around, the iPhone connects to the internet through your mobile provider.

When communicating, the connection is always established by the client and then a HTTP-GET query is issued. Depending on the query, the server responds with a plain text answer or with an image. This answer is then parsed by the client.

There are basically three commands which can be sent by the client.

- getCoins: The client sends his device id and a list of BSSID and timestamp pairs. The server responds with a plain text answer that is composed of multiple text lines. Each line can be of one of three types. A coininfo record containing meta information regarding the collected coin, a bonus record describing which bonus is available for a type of coin and a coin record itself containing a server generated coin id and the type of the collected coin. When a coin type is received for the first time, the coin image has to be downloaded from the server with the following command.
- getCoinImage: The client sends the id of the coin type for which the image is needed and the server directly responds with the image data encoded in the PNG file format. This image has a 80 x 80 pixels dimension.
- getBonus: The client sends his device id, the id of the coin type it is trying to redeem, the amount of coins it is sending and a list of previously collected coin ids. The server responds with the image data of the bonus encoded in the JPEG format. This image has a 320 x 480 pixels dimension to fit exactly the screen of the device. This image can be dynamically generated on the server and a barcode can be included for easier redeeming.

Example getCoins query:

```
http://collector.server.com/collector/api.php?action=
getCoins&did=B1D35D01-32DA-590B-9518-BFDB8F5BC4B8&
data=0:1b:63:2b:4a:b7-1243326354;
```

Answer from the server (fields in the lines are tab separated):

```
I euNCXjTOOJ Plus City http://www.plus-city.at/
B euNCXjTOOJ 10 5 Euro Gutschein
C dn2ccssrXp euNCXjTOOJ
```

```
I f1GRgsU3z7 Müller http://www.mueller-drogerie.at/
B f1GRgsU3z7 10 -10 % auf alles
C d66nrs5cyT f1GRgsU3z7
I skb2a95j2G Cosmos http://www.cosmos.at/
B skb2a95j2G 10 -5 % auf alles
C 9QIKl7O5DD skb2a95j2G
```

4.3 The Coin Management Interface

After touching the radar icon in the collecting screen a new screen for managing the collected coins appears. After the server responded to the getCoins command issued by the client, all newly collected coins appear at the top border of the screen. These coins are branded with the store's logo and can be dragged with the fingers to two drop zones at the lower border of the screen. The left drop zone represents the already collected stack of coins. By dropping a coin onto it the coin is added to the stack and is also collected in the future. The drop zone on the right side symbolizes a horizontal slot to throw newly collected coins away. By throwing coins away they are added to an ignore list and won't be collected anymore. So the client learns which coins are of interest for the user and which ones are not.

For easier decision making, the user can touch any coin and is directed to a page with additional information about that coin, the issuer and the available bonuses for this kind of coin. The additional information can be plain text, a web page or even a video describing the benefits of collecting this coin.

By touching the throw away slot a list of ignored coins is displayed. A coin can be removed from the ignore list as well, but previously ignored coins have to be recollected again - they are not saved.

To get a detail view on the collected coins the user simply needs to touch the stack of coins on the bottom left corner. He then gets to a detail screen where the collected coins are displayed together with the available bonuses. Each type of coin has its own page and by swiping over the screen the user can switch between these pages. With this feature the user is able to get a quick overview of collected coins and how many still have to be collected to receive a coupon. Of course the user can also go to the coin description from this page.

4.4 Screenshots of the User Interface

From left to right Fig. 3 shows the client in collecting mode, the coin organizing screen, the detail collection view and the final coupon.

Figure 3: Screenshots of the collecting and redeeming chain

4.5 Security/Abuse Considerations

To not let the user show his coupon to more than one cashier, the screen displaying it is discarded after a limited time. In addition, a little animating element as pictured in the rightmost screen in Fig. 3 stops the user from taking screenshots from his screen and using this screenshot afterwards.

To ensure that nobody tries to ask the server for a bonus with faked coins, every virtual coin issued by the server has a globally unique, random id. This id is also stored on the server and compared to the ids coming from the client. Even in the case of smartphone backups where parts of it's file system can be restored to a previous state, the coins are marked as redeemed on the server and cannot be redeemed again. To ensure that no fake requests for a high amount of coins are accepted, the server limits the amount of coins issued per client per access point in a certain time span. A scenario with a user living near an access point and collecting the same coins every day is not considered as fraud. The user simply doesn't get a bonus for every coin. Patterns of misuse can be filtered on the server and bad clients could easily be banned for some time.

5 Scenarios

While the CCS was developed for a very wide scope of applications, the following scenarios are among the most obvious ones.

5.1 En passant Coupon Collection while Shopping

The user's activity in this scenario is shopping. It can take place in a mall, a shopping street, or even a whole city. The dimension of the system in this scenario is very well scalable.

The operator of the system has contracts with the businesses in the shopping areas. That can be, for example shops, cafes, restaurants or kiosks. One thing those businesses want is advertisement. Another benefit they might get is information about the customers. Information about the routes they take through a shopping area, how fast they walk through, at which shops they stay for a longer time or which offers they are interested in might prove valueable. The operator is responsible for maintaining the system, distributing the software needed for die client devices, and promotion of the system.

Of course, there has to be some benefit for the customers as well or they would never take action to get and install the software on their smart phones. Therefore they get bonuses from collecting the virtual coins. In this scenario the bonus would be, for instance, a coupon to get discount at a certain store.

So the operator installs access points at various locations, at advertisement walls, entrances and of course the shops. He maps the access point to the coins. For instance, in exchange for passing an advertising wall where Shops A, B and C are advertised, the customer should get coins for shops A, B and C. When passing shop D, he should get three coins for shop D. To make that work the operator maps the MAC of the one access point near the advertising wall to shop A-, B- and C-coins, the access point inside shop inside shop D with three shop D-coins. Some of the bigger stores have their own coins, some of the smaller ones, the cafes in the area for example, have one joint coin together, so that the customer can choose in which one he wants to redeem his coupon.

So, the customer walks through the shopping area. He gets to a shop, shop A (Fig. 4), but he is not very interested in what shop A has to offer, so he walks by without entering. Nevertheless the software on his phone collects the shop A coin. He continues ambling through the area, peeks in shop windows, enters some of the shops to take a look at the offers. By that he collects coins for many of the stores around, also some more shop A-coins.

Figure 4: collecting a coin

Figure 5: redeeming a coupon

Now, aware of the fact that he should have collected some coins by now, he takes a look at his smart phone. He browses through the coins and looks up what bonus or coupon he gets for what type of coin and what quantity he needs to redeem them. He deletes all coins he is not interested in, and keeps the ones which he intends to use in future. The coupon he gets from collecting shop A-coins attracts him and he has already collected enough shop-A coins to redeem it, so he goes back to shop A to buy the article. Now, at the cashier, he wants to redeem the minus twenty percent discount he gets on the article he wants to buy. Therefore he switches to the redeeming screen on his phone and shows it to the cashier. The redeeming screen can show, for instance, a barcode or a simple number code Figure5 [Luf]. After he redeemed the coins, the coins are gone and can't be redeemed again until he collected new ones.

After a long day of shopping, both, the customer and the businesses are satisfied. The customer, because he got some things below the ordinary retail price. The businesses achieved that the customer stood longer in the shopping area than he has intended initially and maybe bought some things he didn't intend to buy in first place. In addition, the businesses can analyze the customer's behavior by taking a look at the order in which the coins were collected, the time between collection of coins, the time the customer stood inside the area and so on. This data might help improving the attractiveness and therefore the profit of the shopping area.

5.2 Coupon Collection while Traveling

Yet another scenario is traveling. There are several possibilities where the system can be used by a traveler. It mostly depends on how he travels, be it by car, by public transportation or by plane.

This scenario might be a bit more difficult for the operator. He is responsible to distribute the access points. In the shopping scenario, even if the shopping area is big, it's manageable. For the traveling scenario we need access points almost everywhere, at train stations, airports, gas stations, advertisements beside roads and in the ideal case everywhere along every root a traveler can take. This, of course, is hardly possible, but there is a solution for an approximation to this. As mentioned before, the operator doesn't have to own the access points used by the system. He can use existing access points. That can be public access points, like the internet hotspots in cities, private access points which are placed in apartments or access points for the internal network of companies. The MAC addresses have to be mapped to coins, which is a bit of an effort, but much cheaper than setting up own access points.

Naturally using foreign access points has disadvantages as well. The reliability is lower and the maintaining costs are higher, because private access points might be removed or exchanged. Some of them are highly reliable, though. Public hotspots, for instance. Anyway, on places where reliability is crucial the operator might want to install his own access points.

Now let's assume the traveler travels by car. He has to drive a long distance between two cities. He travels very frequently and is aware of the benefits the usage of the system can provide. So, at the beginning of the trip, he activates the software on the phone. After a while of driving he needs to refuel. Now he could pick any gas station, but, while driving, he might have collected a number of coins of a certain gas company. That can be, for example, at advertisements beside the road or gas stations of the same company. So he checks his smart phone and realizes that he has enough coins to redeem a minus ten percent coupon at gas company B. He makes his choice of a gas station conditional on the amount and type of coins he collected.

Sadly, after arriving in the destiny city his car breaks down beyond repair. So, in order to get home after he finished his business in the city, he uses public transports. He is used to go by public transports and is also aware of the advantages the coupon collecting system can provide in this situation. A bus company has contracted with the systems operator. In previous visits he has collected some of the company's coins. Now he is able to redeem a coupon to get a short distance with the bus for free.

Later he arrived at the train station. He is a frequent train user and has some train coins on his stack. He collected them by doing some shopping in the train station, passing by train company advertisements and of course by using the train. So he gets his train ticket cheaper by redeeming the train company coupon. If he had chosen to go by plane, some similar benefit would have been possible, like some kind of frequent flyer's miles.

The benefit for the contractors with the system's operator is at the one hand direction of the customer. With clever usage of the system they can make a customer to choose their business over other businesses. On the other hand they can reinforce the relationship to existing customers by giving them benefits for frequent usage of their offerings. Again, it's possible to analyze the data the system extracts out of the user behavior.

6 Conclusions

Pervasive advertisement systems in the future will be crucially reliant to modes of spontaneous, implicit interaction with users: implicit interaction is a mode of spontaneous interaction in which the user is not attentive to a system or the environment. Also, the user is not, an cannot be expected to give explicit response to an interaction system in public space. After investigating on implicit interaction as a potentially effective means to design for spontaneous interaction in the public, we have developed an *en passant* coupon collection system, CCS, to support implicit interaction based pervasive advertisement processes on a wide scale of applications. CCS is based on a scanning for Wi-Fi 802.11 beacon broadcasts, and has been implemented for internet enabled smart phones, here for the iPhone. With scenarios we can demonstrate how the system collects bonus points, rewards, sales coupons, reward coins, etc. by users just walking by points of interest like stores, train- or bus stops, or just in public places of towns or cities. CCS is neither reliant to a preconfiguration of the Wi-Fi infrastructure, nor does it use for wireless data communication. By that, CCS is a very universal approach of implementing geo-referenced, context-aware, personalized services in pervasive advertisement.

7 Acknowledgements

This work is supported under the FFG Research Studios Austria programme under grant agreement No. 818652 DISPLAYS (Pervasive Display Systems), and under grant FACT (Flexible Autonomic Context Technologies) of Siemens AG, CT-SE 2, Munich.

References

[BBRS06] Rafael Ballagas, Jan Borchers, Michael Rohs, and Jennifer G. Sheridan. The Smart Phone: A Ubiquitous Input Device. *IEEE Pervasive Computing*, 5(1):70, 2006.

[BRS05] Rafael Ballagas, Michael Rohs, and Jennifer G. Sheridan. Sweep and point and shoot: phonecam-based interactions for large public displays. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1200–1203, New York, NY, USA, 2005. ACM.

[FV02] Alois Ferscha and Simon Vogl. The WebWall. In *Proceedings of the Ubicomp 2002 Workshop on Collaboration with Interactive Walls and Tables*, Göteborg, Sweden, September 2002.

[HR08] Robert Hardy and Enrico Rukzio. Touch & interact: touch-based interaction of mobile phones with displays. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 245–254, New York, NY, USA, 2008. ACM.

[IEE07] IEEE. *IEEE Standard paper 802.11 - 2007*, 2007.

[Luf] Lufthansa. Lufthansa boarding passes using barcode on mobile phone displays.

[Ohr03] Roeder Ohrtman. *WI-FI Handbook 2003*. McGraw-Hill, 2003.

[RYL+07] Han-Sol Ryu, Yeo-Jin Yoon, Myeong-Eun Lim, Chan-Yong Park, Soo-Jun Park, and Soo-Mi Choi. Picture navigation using an ambient display and implicit interactions. In *OZCHI '07: Proceedings of the 19th Australasian conference on Computer-Human Interaction*, pages 223–226, New York, NY, USA, 2007. ACM.

[Sch00] Albrecht Schmidt. Implicit human computer interaction through context. *Personal and Ubiquitous Computing*, 4(2):191–199, June 2000.

[SD08] Corina Sas and Alan Dix. Designing and evaluating mobile phone-based interaction with public displays. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3941–3944, New York, NY, USA, 2008. ACM.

[SIG09] Bluetooth SIG. *Bluetooth core specification v3.0*, 2009.

[STH07] Albrecht Schmidt, Lucia Terrenghi, and Paul Holleis. Methods and guidelines for the design and development of domestic ubiquitous computing applications. *Pervasive Mob. Comput.*, 3(6):721–738, 2007.

[TCF+07] Nick Taylor, Keith Cheverst, Dan Fitton, Nicholas J. P. Race, Mark Rouncefield, and Connor Graham. Probing communities: study of a village photo display. In *OZCHI '07: Proceedings of the 19th Australasian conference on Computer-Human Interaction*, pages 17–24, New York, NY, USA, 2007. ACM.

[VB04] Daniel Vogel and Ravin Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 137–146, New York, NY, USA, 2004. ACM.

[WFGH99] Roy Want, Kenneth P. Fishkin, Anuj Gujar, and Beverly L. Harrison. Bridging physical and virtual worlds with electronic tags. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 370–377, New York, NY, USA, 1999. ACM.

[WK06] Hendrik Witt and Holger Kenn. Towards implicit interaction by using wearable interaction device sensors for more than one task. In *Mobility '06: Proceedings of the 3rd international conference on Mobile technology, applications & systems*, page 20, New York, NY, USA, 2006. ACM.

Selling the Aether –
A New Billing Schema for Mobile Advertising

Matthias Böhmer and Gernot Bauer
Münster University of Applied Sciences
{matthias.boehmer,gernot.bauer}fh-muenster.de

Abstract: The ubiquity and personalization of mobile phones makes them a promising platform for advertising. Currently, practices known from online advertisement can be adapted, e.g. content-aware ads. However, location as a special feature requires further refinements for mobile advertising. In this paper, we present a conceptual proposal of a billing schema for location-aware advertising. The introduced model describes the pricing of mobile ads related to the spatial distribution of their advertising strength.

1 Introduction

Mobile computing has become an omnipresent issue in nowadays information technology. The processing power of handheld devices is still steadily increasing and mobile connectivity is becoming faster, cheaper and more reliable. In addition, the penetration of devices with embedded sensing capabilities is rising, especially for geographic positioning. Our everyday environment is already pervaded with a high density of these devices. Such mobile and personalized always-on devices provide a new platform for advertising.

Today, a conventional ad is an audio-visual representation of an advertising message placed at a certain location on a specific medium (e.g. billboard, TV spot, radio spot). In this paper – in contrast – we propose a pervasive ad to be a digital audio-visual representation of an advertising message bound to a certain context. This defines the situation the ad should appear in, i.e. the location of the ad, the time span for advertising and the current activity of a customer.

In conventional advertising there is a trade off between the costs for an ad and the number of people who can be reached. Therefore the aim is to minimize the costs and maximize the revenue by focusing on people who will possibly buy the good. An advantage of pervasive advertising – in contrast to conventional advertising – is the possibility of a more precise targeting. In a pervasive computing environment a large knowledge base about people is available, either explicitly by user input or implicitly by sensing and reasoning. Therefore a deeper targeting can be done by profile-based and context-aware filtering. For instance, an ad for fishing equipment would certainly reach people not interested in fishing, although the advertisement was localized near to the coast by choosing a local newspaper or radio channel. However, whereas in conventional advertising the problem of flooding is naturally regulated by a billing schema that evolved over time, in digital advertising the number of appearances of an ad does not relate to costs on the advertiser's side a priori.

In this paper we focus on the spatial distribution of mobile advertising and design an appropriate billing schema. We introduce an advertising space inside a digital layer that covers the earth's surface - conceivable as the aether. The contribution of this paper is a model for pricing pieces of this space for mobile advertising. An advertiser pays for placing his ad on the mobile displays within a certain piece of the aether.

2 Designing a Space for Mobile Advertising

2.1 Idea and Model of a Distribution Volume

In conventional advertising media, e.g. billboards or TV, the space for advertising has a physical and therefore very clear nature: it is available in the form of surface or time units of a resource. As a rule of the thumb, the more units of the resource an advertiser buys, the higher is the probability that people recognize the ad.

Mobile advertising is a form of digital advertising. Therefore, billing schemas need to be based more on artificial regulations than on physical constraints. This can for instance be a scheduling algorithm, mostly according to the business model of the operator. Within such regulations advertisers usually pay for the probability that their ad will be seen, e.g. paying more for pushing an ad to the foremost position on a famous website. This can be referred to as the strength of an advertisement. A similar model is required for mobile advertising.

Mobile computing has the user's location as key function for advertising. Advertisement canvassers can place ads at locations or set boundaries for a better targeting. However, by simply selling spatial regions the notion of advertising strength gets lost (an advantage of digital advertising). Therefore, we propose to add the strength as a new dimension to the geospatial distribution of advertisements.

Figure 1 illustrates our approach. We introduce a virtual layer on the earth's surface. An advertiser can place his ad in a subspace of this sphere. He binds his ad to a three dimensional solid within this layer. While the width and the length of this solid represent the spatial distribution, the advertising strength of the ad can be modeled and visualized as the height of the solid. The latter is not limited and can change within the subspace depending on the location.

This strength can be interpreted and realized in different ways, e.g. as an ordering for the appearance of ads in a list or proportional to display time. Also an observation of the users' click-through rates allows a scheduling of the ads in a priority according to their bought strength. However, the property of strength still needs to relate to the likelihood of an ad to be seen by a user.

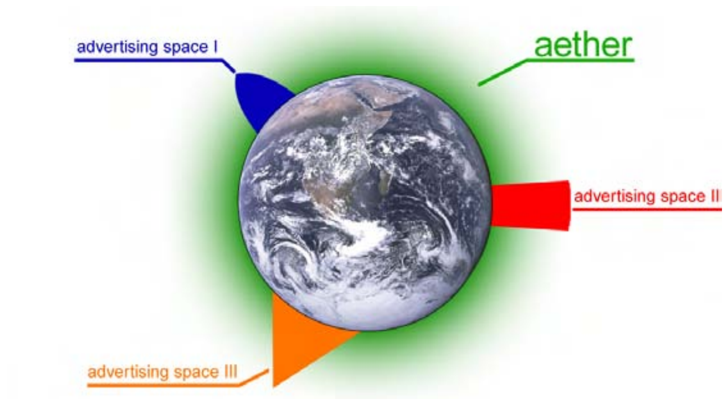


Figure 1: A digital layer covering the earth's surface as space for mobile advertising.

2.2 Pricing the Distribution Volume

In conventional advertising the advertisers pay for the space that an ad occupies (e.g. in newspapers, in magazines) or the time span within an ad is active (e.g. on television, on air) following estimations how many people have been be reached.

Our billing schema for mobile advertising has an analogy to the advertising space on the Internet[1]. Here, advertisers can bid for the likelihood of their ads to appear and be recognized in a certain content range, e.g. given by the content of a search query or a website, in form of an auction. For mobile advertising, we also want to enable advertisers to tune the likelihood of their ad to be seen and to influence the range by means of spatial distribution. Therefore, in our mobile advertising billing schema, the product of the spatial distribution and the strength represents the volume of the solid as advertising space and relates to the price. If an advertiser wants to reach a certain strength for his ad in a certain area, he has to bid a specific price which results in a certain volume.

A simple way to monetize the volume would be a direct conversion of volume units to price units, e.g. payed once or periodically. More specific and dynamic billing units can be adopted from Internet marketing, e.g. by applying methods like pay per click or pay per lead, while the payment itself still correlates to the volume of the advertising space.

2.3 Different Solids for Customizing Advertising Campaigns

The billing for an ad is estimated according to the volume of the solid the ad is bound to within the advertising layer. For a given amount of marketing budget, an advertiser can decide whether he prefers to cover a larger spatial area, or to have a higher advertising strength at a certain location. Concerning this, he can form the shape of his advertising

[1]Most prominent example is Google AdWords, https://adwords.google.com/

space. From a geometrical point of view, the only constraint is that the floor of the solid is planar.

Based on the introduced model, advertisers can become creative in spreading their ads into the aether. They can form different spheres for running mobile advertisement campaigns. Especially, they are able to adapt their campaign to the geographic characteristics of a region, e.g. their point of sale.

Figure 2 shows the use of primitive geometric solids for running mobile advertisements. Cuboids and cylinders can be used for running widespread campaigns for covering large areas as well as to place precise and focused ads. In contrast to solids with a constant height, solids with falling off edges also seem to be useful, e.g. cones, pyramids and hemispheres. They allow to place ads in a certain area with a peak at the center and decreasing advertising strength to the edges. Also the superposition of basic solids can be interesting for modeling campaigns, e.g. to widespread an ad to a large area (e.g. using a flat cuboid) and lay spots at some places (e.g. using high cones).

Figure 2: Basic geometric solids for building advertising spaces.

The use of polygons even provides a higher degree of freedom for advertisers. Figure 3 shows a scenario, where the advertiser is able to model his advertising space according to his geographical circumstances. The sketched iceman of course does not want to pay for advertising above the water, because there is no potential customer. Those are at the beach, wherefore the iceman focuses on this are. However, he also knows that his catchment area is limited – even if people from far away saw his ad, they would not walk too far for getting an ice. So the iceman does not invest very much to promote his ice far beyond the beach across the street.

Pieces of the space in the advertising layer can be taken several times. Especially when the owners of the ads are competitors in business this might lead to an interesting phenomenon. In conventional advertising, this is somehow similar e.g. to over-pasting posters on bill-boards. In our mobile advertising model, this corresponds to the height of the solids as the strength of the ads. For example, it is imaginable that pizza baker Alfredo sets up a mobile advertisement in his hometown bound to a cone. Ads for his restaurant will be shown on devices of people within this space. However, when Pizza baker Beppo places another ad

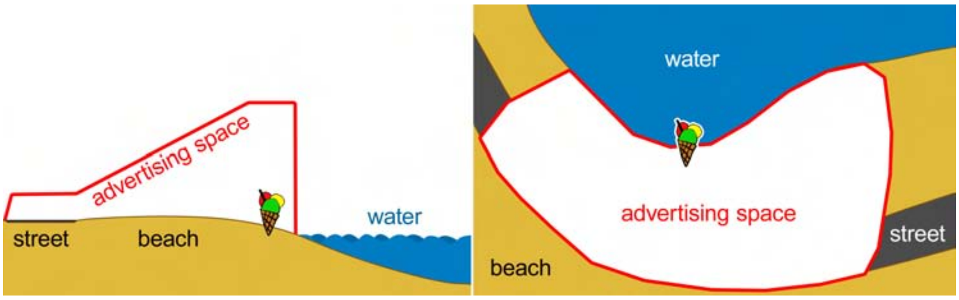


Figure 3: Profile (left) and top view (right) of the mobile advertising space for ice cream at a beach.

bound to a bigger, especially higher solid at the same location, people will more likely see his ad. This would lead to some kind of competition with the advertising budgets as upper barrier. Figure 4 shows this phenomenon as intersection of solids within the advertising layer. Of course, the forms of feasible solids have a large impact on advertising campaigns, not only because of the possibility of intersections.

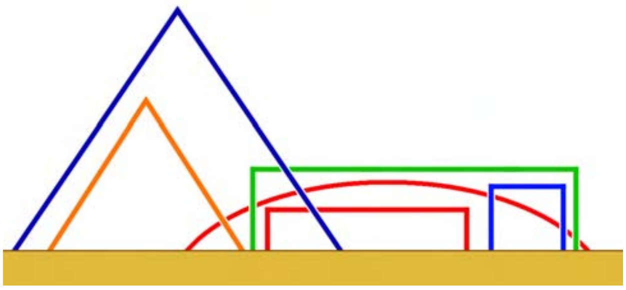


Figure 4: Profile view of the advertising layer showing competition as a result of reselling the space.

We assume, that this form of competition will also lead to different prices for different areas in the world. Places which are worth much more than others for mobile advertising and which guarantee a higher revenue will accommodate higher advertising solids. At places where the usage of mobile devices is sparse and therefore advertising not very worthwhile, presumably no peaks will emerge. This competition will further lead to economic price levels for different areas, which will only be limited by the cost-benefit calculation that each advertiser estimates on his own.

3 Mobile Devices as Advertising Medium

Mobile advertising uses mobile devices as medium for displaying ads. The described model was especially designed for mobile phones. Figure 5 shows different opportunities

Figure 5: Today's ads on a mobile phone: within a website [ZEI09] (left), within an application [myM09] (middle), and as a standalone application [Poi09] (right).

for presenting ads on such devices. Adapted from classical advertising on the Internet, ads can be embedded in mobile websites. Small banners, text ads or videos can also be placed inside applications. These two forms of mobile advertising are already common on today's devices, also in a location-aware manner. A third way are standalone applications, which are shown by an icon in the main menu of a device. Such marketing applications also already exist, but are not pushed onto devices depending on the user's location. Regarding the presented billing schema for location-aware advertising, the presented ads on a mobile device will change when the user is on the move.

Another, hitherto not explored way of mobile marketing could use the idle screens of mobile phones as ad space in form of a screen saver. Private phones lying around e.g. on tables in a café or at desktops are not only visible to their owners, but also to people sitting nearby or passing. The proposed model might also work for billing advertisements on other mobile displays like for example on cars [AES09] or clothes.

4 Conclusion

There is a strong consensus in the research community, that location aware advertising is reasonable and beneficial. However, from an advertiser's point of view, why should he fence his ad in specific location areas, when there is no reason given a priori. Introducing a price per area is a first approach for designing a billing schema. In this paper, we presented an extended model for billing space for mobile advertising. Combining the spatial distribution of an ad with a measure for the likelihood of an ad to be seen by customers we set up a three dimensional space covering the earth surface. This space forms a digital layer as advertising space. Within this layer, advertisers can buy subspaces for running mobile campaigns with a price predetermined by the advertisers depending on the volume of the solid. By competition for the best advertising spaces economic prices will emerge.

A first extension of the presented proposal includes an additional dimension for modeling the strength of ads at certain times. This is imaginable as pulsing advertising spaces or waves, with an interpolation of the shape between given timestamps. Methods like pay per click allows an easy billing, because they are discrete and every ad has a price at a certain location and time.

Having designed a billing schema for mobile advertising, one questions remains open: who should earn the money? It would be possible for network providers or device manufacturers to implement an appropriate platform. However, as long as there is no standard for connecting these, there obviously will be a very dispersive and heterogeneous marketing scenery. Instead, the proposed digital aether should be a common infrastructure, similar for example to the domain name system of the Internet. Real posters can also only be put up in our unique and common environment.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable suggestions. This work was funded by the German Federal Ministry of Education and Research under grant 1748X08.

References

[AES09] Florian Alt, Christoph Evers, and Albrecht Schmidt. Users' View on Context-Sensitive Car Advertisements. In *Proc. Pervasive '09*, pages 9–16, 2009.

[myM09] myMobileApps Inc. iChess. Screenshot of iPhone application available on the iPhone AppStore, 2009.

[Poi09] PointAbout Inc. BURGER KING NOW: PHONE ORDERING. Application available on the iPhone AppStore, 2009.

[ZEI09] ZEIT Online GmbH. ZEIT ONLINE. http://mobile.zeit.de, 2009.

TaxiMedia: An Interactive Context-Aware Entertainment and Advertising System

Florian Alt[1], Alireza Sahami Shirazi[1], Max Pfeiffer[1], Paul Holleis[2], Albrecht Schmidt[1]

[1]Pervasive Computing and User Interface Engineering Group
University of Duisburg Essen
Schuetzenbahn 70, 45117 Essen
[2]DOCOMO Euro Labs
Smart and Secure Service
Landsberger Str. 312, 80687 Munich
{florian.alt, alireza.sahami, albrecht.schmidt}@uni-due.de,
max.pfeiffer@stud.uni-due.de,
holleis@docomolab-euro.com

Abstract: The use of public transport vehicles, such as trams, buses, and taxis as an advertising space is increasing since several years. However mainly the outside of the vehicles is used to show advertisements using paintings, foil or roof-mounted displays. Nowadays, with advances in display technologies, small high-resolution displays can be easily embedded in vehicles and be used for entertainment or advertising purposes. In this paper we introduce an interactive context-ware advertising system designed for cabs, which is targeted to offer context-aware information such as advertisements, points of interest, events, etc. during a cab ride. Additionally it is possible for advertisers to upload their contents and define areas where their advertisements should be shown.

1 Introduction

Since the 1990s, public transportation vehicles have been used as an advertising space. The first example was a Pepsi advertising campaign started in 1993, where the Pepsi logo was painted on urban buses in the Seattle area. This type of advertising is mainly static in that the advertisements do not adapt to their contexts. One of the first approaches towards dynamic advertising content involved mounting electronic displays on cars where the ads could have been updated based on the location, e.g., on taxis in the Boston area. However, all these types of advertising are deployed on the outside of the vehicles. Nowadays small, high-resolution displays are embedded inside busses or trains mainly for advertising, news, or the visualization of location-based information such as details on the current/next train station. In the last few years, also displays in private cars became more common. Such displays are either integrated in the front seat headrests or the central console of the car. These displays are mainly used for entertainment purposes such as watching movies or playing games.

In-car displays are well suited for entertaining the customer by showing him news, information on close points of interest, short videos or even letting him play interactive games. There are already systems commercially available for advertising in taxis such as Tapinto[1], an international taxi-TV supplier, using radio frequency for transferring data.

To explore the potential of in-car displays we implemented the TaxiMedia system, an interactive context-aware system for taxis. Our system is able to show different types of information based on the context of the taxi and can be used to entertain the passengers during the trip. We mainly use our system to visualize location-based advertisements. Those advertisements can be selected based on the location a passenger is picked up at as well as the current location of the taxi, but also further context such as time, weather, or interest of the passenger may be considered.

In the following we show how context information can be used, to more precisely target advertisements to the interests of the user. For example, if a passenger takes a cab from the airport on a week day morning, we may assume that the passenger is currently on a business trip. Hence, advertising hotels in the business district, as well as restaurants, cinemas, or special events in the evening may be interesting for the passenger. In contrast, weekend passengers are more likely to be tourists. Hence advertisements on hotels close to the major tourist sites as well as on interesting upcoming events may be shown. In contrast to available commercial systems, our system is working also offline without any means of connectivity to the server.

The rest of this paper is structured as follows: in chapter 2 we present our concept of a location-based entertainment and advertising system for taxis and how different types of context can be used to enhance the impact of advertisements. In chapter 3 we provide an overview about the architecture of the system and its components. After that we provide a brief overview on related work. Finally we provide a conclusion and talk about future work.

2 Concept

The overall idea as outlined in the previous section suggests itself as a rather simple and intriguing business case. The availability of information about the person looking at an advertisement and the close contact of this person with the digital information can considerably enhance the impact of advertising. In the following, we briefly go into some details about the advantages and requirements for the proposed solution.

2.1 Location

Location is one important type of context for any provision of adapted information. In the event of someone using a taxi, this can be exploited very easily and thoroughly. In contrast to many other settings, three types of location are known at any point in time: the start of the journey, the current location throughout the journey, and the destination.

[1] www.tapinto.info (last access 30.05.2009)

This data can be exploited in various ways:

Starting point: depending on the type of location and the current time and date, some conclusions can already be derived about the purpose and type of customer. Someone arriving at the airport in the evening or being picked up at a business office will most probably be on the way home or looking for a hotel to stay overnight. Someone arriving with a train in the morning can be expected to stay in town most of the day and potentially overnight. This information can be used to present targeted advertisements of, e.g., hotels or theatre tickets.

Current location: the current location during a journey is important to be able to display advertisements that fit a particular region or place that is passed. Targeted advertisements can be shown such as exhibition tickets, offers for sightseeing, or particular events.

Destination: the knowledge of the planned destination of a passenger can be used in a similar way as the knowledge of the initial pick up point. In addition, one can predict that the person will be spending some time at the destination and generate targeted advertisements, e.g. for lunch offers, evening entertainment, and local hotels.

However, also other types of context may be of interest:

Weather: the weather forecast for the destination of the passenger may be used to adjust the advertisements shown on the screen: whereas sunny forecast might result in a suggestion to visit one of the city's nice beer gardens, for rainy weather a restaurant next to a cinema or theatre might be proposed.

Passenger: nowadays, reliable algorithms exist when it comes to sensing gender and age group of a person in front of a display. This information can be used to show gender-specific advertisements as well as to target advertisements towards the most likely age group of the passenger.

2.2 Requirements

The system needs to support the requirements of three main user groups. First, for *advertisers*, an easy and remote opportunity is required to add advertising content and to configure the conditions under which they will be displayed such as the location, time of day, and the frequency. Further it is necessary to agree upon a pricing model and to think about how the impact of the advertisements can be measured (e.g. by providing customers a coupon which they can redeem later in a store). Finally, advertisers need to be provided with feedback about the actual use of the system.

Second, drivers and taxi companies can also benefit from the use of such a system. Whereas it first has to be assured that the hardware can seamlessly be integrated into the car, feedback for the driver as to how customers interact with the system and especially advertisements may be useful. If the driver can also see the currently displayed advertisements, e.g., on a museum, he might be able to provide additional information to the customer. The driver could also use information about the currently displayed advertisement or information on the user interaction (Which advertisements are closed by the user? Which advertisements generate additional queries by the user?) to adjust the route (e.g., if the user wants to buy tickets for a concert at night and the ticket office is just a few meters off the route). However, a business model has to be generated upon which also the driver or taxi company could be compensated.

Third, also the customer has to be taken into account. It has to be assured that advertisements are displayed unobtrusively and that the advertisements generate some added value for the passenger, e.g. by fitting them to the context or by directly providing an opportunity to get more information about the displayed advertisement.

One aspect that is reflected in this list of requirements but only rarely incorporated in realized systems is to give people the opportunity to directly interact with the data on the display. Showing an advertisement about a product to a potential customer is in vain if the person forgets the contents. Providing a direct link to more information or, potentially, even to a shop where the product can be bought can be essential for the success of advertisements. Existing approaches include trying to give the product a catching name or URL, indicating where (in the vicinity) the product can be purchased, or providing a visual barcode that links to the product's web page. The first two solutions are not applicable in the setting within a car. Taking a picture of a visual code and browsing to a web page with a mobile phone is already relatively advanced and has several disadvantages (difficult in moving and less lit vehicles, other passengers are cut off the action, can incur costs for the customer, etc.). We tackle these issues by providing a touch screen that can directly and concurrently be used to display advertisements and let the passenger engage and retrieve appropriate additional information.

In addition to the mentioned features, the proposed infrastructure and concepts can be extended to include further potential in this area.

2.3 Challenges

In the following we present several challenges one might be faced with when deploying such a system.

Appropriateness for specific types of customers: a critical issue is that particular advertisements might be inappropriate for specific customers. This includes problematic cases such as the use of foreign languages, adult contents, and content that is culturally differently interpreted. Other areas such as advertisements focused on male / female customers or elderly people might not be problematic but will reduce the impact if not appropriately displayed. Some aspects can be automatically detected, e.g. the current number of passengers.

Targeting at the particular customer: the concept mentioned in the last paragraph can be brought further to personalized services. People have specific opinions, knowledge and preferences that could be incorporated into the design of targeted advertising. Such information can be automatically derived only with difficulties (e.g. age through image recognition). Two feasible solutions include entering information manually by the driver or the customers themselves or by using personal preferences stored in the customers' mobile phones.

Access to data: in addition to directly accessing additional information as described above, an important fact often neglected in deployed systems is that the retrieved information is lost as soon as one leaves the location (in this case the car). One simple approach (besides having a printer installed locally) would be to push information to the customers' phones.

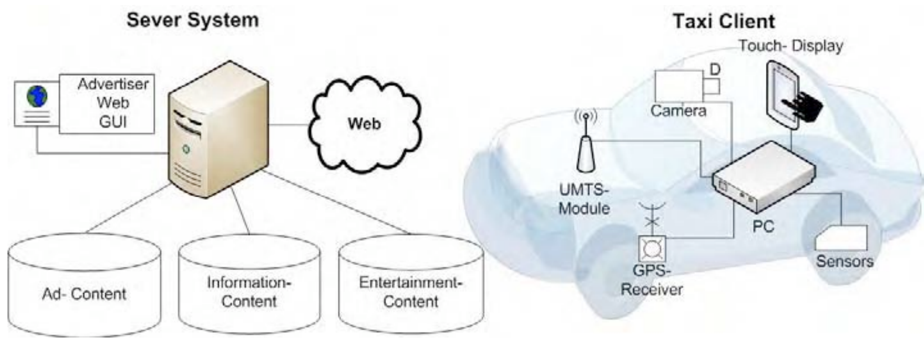


Figure 1: The Taxi Media system architecture.
The system includes two parts: Server System and Taxi Client.

3 System Architecture

The system we implemented includes two parts, the server system and the taxi client shown in . The server system is responsible to mange the information and has a database for the advertisements and other information such as events, news, and points of interest. Additionally we implemented a web-based application which lets advertisers upload their advertisements and define regions where thy want their advertisements to be shown.

The second part, the client system, is a Python-based application running on an *Eee PC*[2] computer that is connected to a touch display embedded into the headrest (see Figure 2). The *Eee PC* has a GPS receiver to track positions and Internet connectivity using a UMTS stick to communicate with the server system. The client system is scheduled to automatically update its contents daily and save all necessary information locally. The update mechanism includes news, advertisements, and entertainment contents. This approach has the advantage that the system does not need any permanent connectivity and works even if the cab is in a tunnel and does not have any coverage. On the other hand the menu layouts of the user interfaces is generated automatically based on the current folder structure and provides the associated content to the costumer. That means that if new contents are added to the system later the system can easily update the user interface without requiring any update in the system core. Some contents such as advertisements or hotspots are classified based on the location and shown at the appropriate positions.

Since the display is touchable, the passenger can interact with the user interface. In the idle mode the display shows different location-based advertisements. The passenger can start interacting with the system by touching the display. Then the menus are visualized and the passenger can select the contents. The contents provided are the available hotspots, restaurants and bars, news, and entertainments based on the current location. Additionally other entertainment contents such as games or videos are provided. Further, a small part of the screen is always reserved for advertising while the content is shown. Selecting the advertisements is done based on matching predefined location of each advertisement and the current location.

Figure 2: A touch display is embedded in the headrest.

[2] http://eeepc.asus.com/ (last access 30.5.2009)

4 Related Work

Location-based services have been explored in various researches. Kölmel *et. al.* [6] implemented a middleware, which provides a host of functional software modules to enable straight-forward deployment procedures of location-based services. SMMART, explained in [4], is a context-aware application running on smart phones. Ad-me is a context-sensitive advertising system, which aims to deliver more palatable, less intrusive and personalized advertisements integrated with a mobile tourist guide [2]. On the other hand, context-sensitive car advertisements are explored in different researches. Based on the result of a survey reported in [1] more than half of the participants were interested in displaying advertising on their cars. Tester et al presented CommuterNews [8], a prototype aiming at engaging the user into active interaction with an in-car entertainment system.

These researches reveal that context-based advertising is an important aspect, which can be considered for providing more related advertisements in appropriated contexts and cars can be a space for providing location-based information and advertisements. During this project we investigated on how to use vehicles as an advertising space and looked into advertising inside vehicles, especially cabs.

As described in the previous chapter, we use a local folder structure to store and make data available offline. Our approach inspired by the Code and Odyssey Systems [2][6][7] – however we were not able to evaluate performance and scalability of the approach in a real-world environment.

5 Conclusion and Future Work

Nowadays, vehicles are more and more used as advertising spaces. Most of these advertisements are static and only deployed on the outside of the vehicles. However, with new technology being available small high-resolution displays are embedded inside the vehicles and trains. These displays are mostly used for entertainment purposes or advertising. To explore the advantage of these types of display we implemented the TaxiMedia system, an interactive context-aware advertising and entertaining system for cabs. Our system includes a touch display embedded in the headrest and interactive user interfaces. The main purpose of this system is to show location-based information and advertisements. In contrast to other available systems, our system works with a local data basis, which also works without any Internet connectivity. Additionally, the system offers a web-based application for the advertisers to upload their advertisements to the system and to define context such as certain areas where their advertisements should be shown. So the advertisers can easily manage their advertisements.

For our future work, we plan to deploy and evaluate the system in a real environment. We cooperate with a local taxi company, which is interested in setting up the system in one of their vehicles. Hence we could use this high-fidelity prototype as a basis for presentations and further discussion with taxi drivers and advertisers. This would give us further insight as to which ideas are applicable (e.g., that a driver provides additional information or adjusts the route) and which are not. Further, we plan to refine the set of requirements based on qualitative interviews with passengers. We also think of incorporating further sensors such as a video camera into the system setup in order to test additional types of interaction such as a video chat function.

Finally we plan to look into how we can encourage passengers to interact with our system. One option would be to integrate personalized information. Hence data from social networks, friend finders, or web services such as dopplr could be used to show passengers, which colleagues, friends, etc. are in town at the same moment.

Literature

1. Alt, F., Evers, C., Schmidt, A.: User's View on Context-Sensitive Car Advertisements. In Proceeding of Pervasive 2009, Nara, Japan, May 2009.

2. Howard, J. H., Kazar, M. L., Menees, S. G., Nichols, D. A., Satyanarayanan, M., Sidebotham, R. N., and West, M. J. 1988. Scale and performance in a distributed file system. ACM Trans. Comput. Syst. 6, 1. 1988. 51-81.

3. Hristova N., & G.M.P. O'Hare, "Ad-me: Intelligent Context-Sensitive Advertising within a Mobile Tourist Guide", In Proc. of 12th Irish Conference on Artificial Intelligence and Cognitive Science (AICS), National University of Ireland, 2001.

4. Kölmel, B., and Alexakis, S. "Location Based Advertising," in Proceedings of the First International Conference on Mobile Business, Athens, Greece, July 2002.

5. Kurkovsky, S. Using Principles of Pervasive Computing to Design M-commerce Applications. In proceedings of ITCC-05, Las Vegas, United States, Apr. 2005.

6. Noble, B. D. and Satyanarayanan, M. 1994. An empirical study of a highly available file system. In Proceedings of the 1994 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems, 1994. Nashville, United States, 1994. B. D. Gaither, Ed. SIGMETRICS '94. 138-149.

7. Satyanarayanan, M. 1996. Fundamental challenges in mobile computing. In Proceedings of PODC 1996, Philadelphia, United States, 1996. 1-7.

8. Tester, J., Fogg, B., and Maile, M. 2000. CommuterNews: a prototype of persuasive in-car entertainment. In CHI '00 Extended Abstracts on Human Factors in Computing Systems, The Hague, The Netherlands. 2000. 24-25.

A Modular Framework to Detect and Analyze Faces for Audience Measurement Systems

Andreas Ernst, Tobias Ruf, Christian Kueblbeck
Fraunhofer Institute for Integrated Circuits IIS
Department Electronic Imaging
Am Wolfsmantel 33, 91058 Erlangen, Germany
{andreas.ernst, tobias.ruf, christian.kueblbeck}@iis.fraunhofer.de

Abstract: In this paper we describe an approach that enables the detection, tracking and fine analysis (classification of gender and facial expression) of faces using a single web camera. One focus of the paper lies in the description of the concept of a framework that was designed in order to create a flexible environment for varying detection tasks. We describe the functionality, the setup of the framework and we also give a coarse overview about the algorithms we are using for the classification tasks. Benchmark results are given on standard and publicly available data sets. Although the framework is designed for general object recognition tasks, our focus so far has been in the field of face detection and fine analysis. With the capabilities provided by the system we see one main application field in the area of digital signage where providers are enabled to gather information, measure audience focus or even create interactive advertisement solutions. The varying applications require different software features and the provided framework approach allows to easily create different recognition setups in order to fullfill these specific needs. We show in a demo application for still images and movies that our framework can be well used for these purposes. This can be downloaded from `http://www.iis.fraunhofer.de/EN/bf/bv/kognitiv/biom/dd.jsp`.

1 Introduction

Technology that enables the detection of faces is developed and improved especially with respect to their performance, error rates and computational effort. The same is valid for systems that analyse faces in order to extract information like time of focus, gender, age or type of reaction. We think that there are other important features that such technology should focus on:

- possibility to efficiently use the software in dedicated applications, so that developers using the software can use the technology with little integration effort,
- configuration capabilities that on the one hand make the software easy to use and on the other hand offer extensive and manifold configuration possibilities.

We present a method for fast and robust face (and object) detection and analysis. The algorithms are embedded in a library called SHORE, an acronym for Sophisticated High-speed Object Recognition Engine. Using different setups the engine can be used either for simple face detection or for more complex tasks like gender classification, tracking of faces, the analysis of facial features, classification whether the eyes are opened and closed and so on.

In this paper we at first describe the functionality and the capabilities of the system. Then we explain the design of our framework in order to fullfill the two features mentioned above. Afterwards we describe the algorithms that we are using and finally we show the results of benchmark tests that we have carried out on standard datasets.

2 Functionality

Our software framework enables the detection, analysis and identification of objects. Up to now we have focused on faces as special objects, but it is also possible to handle other objects that consist of a typical structure. Thus we have already made some promising experiments with the detection of hands or cars.

In the training phase a certain amount of annotated sample images is needed for building a model. *Annotated* means that in advance a certain set of specific points is defined and that these points are marked in each image.

We have created a system that is able to rapidly detect faces in arbitrary images and to further analyze them. So we can also detect the positions of the eyes, the nose and the mouth. In addition we can analyze whether the eyes and the mouth are open or closed. Furthermore we have implemented the classification of gender and the estimation of certain facial expressions (so far happy, angry, sad, surprised).

3 The Framework

Software frameworks are reusable design solutions for a domain of applications or problems [GHJV94]. The presented framework addresses the problems of object detection, fine analysis and recognition (identification). We provide a so called black-box framework [Rie00] to the client, who only needs to know how to configure and use it. The internal details are hidden by the framework interfaces.

Figure 1 shows an overview with the fundamental parts of our framework including three applications (on the top), which use the framework in different use cases. The components used at runtime are located on the left and the middle of figure 1. The right side contains the parts employed for the training of models and annotation of images. Figure 1 gives an outline of our design efforts based on the following objectives:

- coverage of the functionality described in section 2,
- easy and flexible configuration and simple usage,

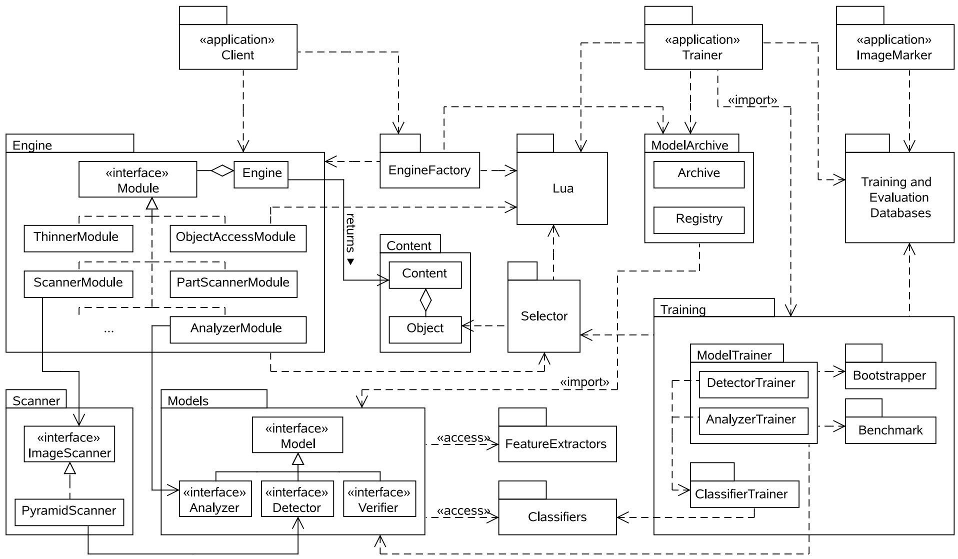


Figure 1: Overview with the fundamental components of our framework as an UML package diagram.

- generic object representation,
- reusability and extensibility,
- loosely coupled training,
- generic annotation of image databases.

To cover the functionality of object detection, fine analysis and recognition we introduced the concept of *Models* (left/middle bottom in figure 1). The three interfaces `Detector`, `Analyzer` and `Verifier` are specialized models that provide these functionalities. Models in general can be seen as an abstract representation of object characteristics, that are used for classification.

Easy configuration capability is achieved by the packages `Engine` and `EngineFactory`. A client can easily set up an `Engine` with the `EngineFactory` and the help of the Lua scripting language, see section 3.4.

Object detection requires an internal representation of the found objects. With a generic representation we can handle different objects, for example faces and cars, with one construct. An `Object`, middle of figure 1, holds its type, a region, name value pairs for ratings, attributes, marker points and other objects as parts. Different objects can be represented by these class attributes. The `Engine` returns a `Content`, that contains all detected and analyzed objects in an image.

Reusability and extensibility are common design issues. Important parts within our framework concering these issues are the `Module` and `Model` interfaces, `Classifiers` and `FeatureExtractors`. Functionality can be easy extended with new modules or models. Further feature extractors and classifiers can help to improve the overall performance.

The following three sections describe the three major use cases for the framework and show the basic design ideas. Thereafter section 3.4 gives some insights how we use the Lua scripting language within the framework.

3.1 Annotating Images

Figure 1 shows the `ImageMarker` application in the top right corner as a part of the framework. The ImageMarker was developed as a tool for flexible manual annotation of objects in images. It is possible to annotate the type and the region of objects in the image. In addition predefined and named marker points and key value pairs, so called attributes, can be annotated for each object. Which kinds of objects and how the objects are annotated is configured in the application settings. New object types can be added any time simply by defining them in a preference file.

A frontal face in the image for example is annotated by marking both eyes, the tip of the nose and both mouth corners. Each marker has a key that describes the marker, e.g. `"LeftEye"`. In addition different attributes can be assigned to the face. In the frontal face example the gender or facial expression can be annotated using the tool, e.g. `Gender:="Male"` or `Expression:="Happy"`. A screenshot of this sample annotation using the ImageMarker is given in figure 2.

The output of the annotation process is a file containing the information about all the objects in the image in xml. The annotated image databases are used for the training and benchmark of models within the framework as shown in figure 1. During the training process the objects that are used as training data can be exactly selected and cut out by using the image annotation. This is described in more detail in the following section.

3.2 Training Models

The training concept in our framework has two important properties. First the training parts are decoupled from the remaining framework, that allows us to provide only the components needed at runtime to the user of the library. Secondly the whole training is scriptable. For the training we use Lua-based training scripts, which define the training and related tasks. The `Trainer` application in figure 1 can parse these scripts with the help of Lua and uses the `Training` package to carry out the training. Details on the use of the Lua scripting language can be found in section 3.4.

A typical training script comprises the following workflow:

- select training images from annotated databases,

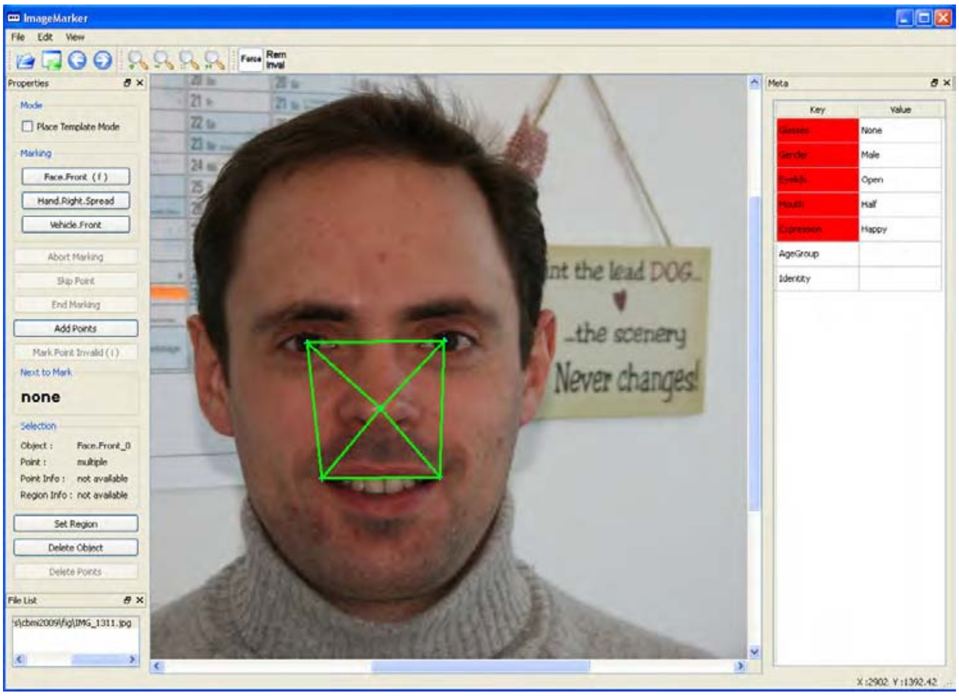


Figure 2: Screenshot of the ImageMarker with a sample annotation.

- choose feature extractors (package `FeatureExtractors`),
- choose trainers for the classifiers (package `ClassifierTrainers`),
- start training with a concrete model trainer (package `ModelTrainer`),
- optimize the model, e.g. with the `Bootstrapper` package,
- benchmark the trained model (package `Benchmark`),
- save the model with the `ModelArchive`.

The selection of the training images is done by the `Selector` package. Using selectors we can exactly define the subsets of images, that match the designated classes for the classifier training, see section 3.4.1. Depending on a concrete model we need a suitable model trainer, which knows how to train its model and handle training related tasks.

3.3 Using the Engine

The left side of figure 1 shows the usage of the framework from the viewpoint of the client application. Essentially the client only uses the `EngineFactory` and the `Engine` itself. The engine factory is able to create a custom tailored engine by parsing a Lua-based setup script that defines the configuration of the engine. An example for an engine setup script is depicted in section 3.4.3.

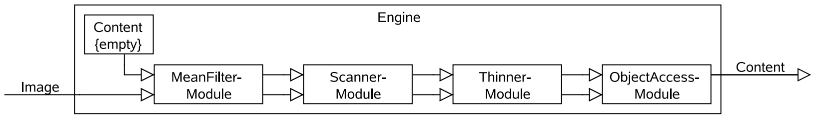


Figure 3: An engine setup configured for a simple face detection task.

Depending on the setup, a concrete engine can do just primitive image processing jobs or even quite complex object detection and analyzing tasks. A simple engine for example could just detect faces in images, whereas a more complex one makes an additional facial feature fine search and a facial expression or gender analysis.

The engine itself encapsulates a concatenation of different *Modules* as shown in figure 3 for a simple configuration example. When using the engine the image and a `Content` instance is provided to the modules one after each other. Each module is allowed to use or modify the image data and/or access the objects in the content.

A client may use the framework in the following way: First it provides the engine setup script to the engine factory. The factory parses the script to create the engine and returns the engine back to the client application. Thereafter the engine is ready to process images provided by the application. For each image the engine returns the content that contains all the objects and the information gathered by the engine. The application can read out the content for each image and may do further processing depending on the information in the content. When the engine is not needed any more it can be destroyed by the application again.

3.4 The Lua Scripting Language as an Instrument for Flexible Configuration

Lua is a very powerful, light-weight and fast scripting language. It can provide scripting capabilities to all kinds of applications. Looking at `http://www.lua.org` gives a good impression of its properties and advantages compared to other scripting languages.

In our framework we use it as an *extension* and *extensible* language [lua03]. These views summarize how Lua and C interact, in which Lua is implemented. The use of Lua as an extension language means that C code can be extended with the scripting facilities of Lua. As an extensible language Lua offers the facility to extend it with new functionality written in C or another language [lua03]. Within our framework we use both interaction modes for the `Selector`, `ObjectAccessModule` (extension), the `EngineFactory` and the `Trainer` application (extensible).

3.4.1 The Use of Lua as an Extension Language

A `Selector` exactly defines or selects a set of objects that fulfill the selector expression. Lets take a simple example. A Lua-based selector that selects all objects of type `"Face"` and where the attribute called `"Gender"` equals `"Male"` looks like this:

```
Object.Type                  == "Face" and
Object.Attribute["Gender"] == "Male"
```

The selector is realized by pushing all the object properties first in the so called Lua state. This makes them available as variables within Lua. After that the selector expression is parsed by Lua. The result of parsing the selector is a boolean that tells us whether the object belongs to the selected subset or not.

Selectors are widely used during training and runtime. Within the model training selectors exactly define which objects in a training image database belong to the predefined classes. This way it is for example very easy to define the positive and negative training datasets for a classifier that distinguishes between male and female faces. During runtime the selectors are used in modules of the engine that access or manipulate objects. Let's take for example an engine that detects faces as well as hands. In this case it can easily be defined by selectors which modules in the engine take care of or manipulate the hand objects and which ones take care of the face objects.

The `ObjectAccessModule` uses Lua to provide a very flexible way to access and manipulate objects within the engine. This module is realized in a way similiar to the `Selector`. All the object properties are pushed in the Lua state first. After that a given Lua expression is executed that grants access to all the object properties. Thereby the object can be manipulated within the Lua state. At the end the object is read out again from the Lua state and replaces the original one.

A simple example how to use this module is shown at the end of the engine setup in listing 1. It is described in more detail in section 3.4.3.

3.4.2 The Use of Lua as an Extensible Language

The `EngineFactory` and the `Trainer` application extend Lua to use the framework in a flexible and comfortable way. Both register functions and types (classes) in the Lua environment [lua03] to make them available for scripting. For the registration of external parts their identifiers and the name for the Lua access must be pushed to the Lua state. For technical details see [lua03, Part IV] or [IdFC06].

Since the `Trainer` is a stand-alone application, we extended the default Lua interpreter with our requirements. All components accessible from a training script are registered in the Lua state of the interpreter. The registration by hand would be impractical and inefficient, because we have lots of training related components. One easy and flexible way to make C and C++ accessible in Lua is to employ SWIG (Simplified Wrapper and Interface Generator), see `http://www.swig.org`.

Unlike the `Trainer` the `EngineFactory` is not an application. But it is also able to parse scripts with configuration information to setup an engine. The factory holds its own Lua state, which is extended with the `Engine` class and setup functions. An extract of our registered functions for the setup can be found in listing 1.

3.4.3 Configuration Example

A sample engine configured for a simple face detection task is shown in figure 3. It starts with a `MeanFilterModule` that just applies a 3×3 mean filter on the input image to reduce noise. The second module scans the image for faces using a face model. The `ThinnerModule` removes multiple detections by applying geometric constraints. And finally the `ObjectAccessModule` is able to modify the found objects in a very flexible way by using Lua.

The corresponding setup script for this configuration is given in listing 1. A primitive example for the flexibility of manipulating objects with the `ObjectAccessModule` is shown at the end of the script. In this case it calculates the eye distance and adds it to the object as an attribute with key `'EyeDistance'`. The provided selector makes sure that only faces are accessed by the module. The `EngineFactory` is able to parse this script to create the `Engine`.

The `Engine` package in figure 1 shows five different sample modules. In our current framework we implemented over 25 different modules for various kinds of image analysis tasks but also for object and content manipulation that can be used within the engine.

Further more complex engine sample configurations are realized and shown in our demo application (for a screenshot see figure 4).

4 The Algorithms

In this section we explain in short all four parts that make up our (and normally every) pattern classification system: preprocessing, feature extraction, classification and training. A more detailed description of the workflow with respect to certain speedups of the algorithms is presented in [KE06].

As a preprocessing step we simply use a mean filter which has turned out to be an easy and fast way to smooth noise artifacts. Further preprocessing is not carried out.

For feature classification we have decided to use features that are at least somewhat independent with respect to illumination conditions. From these features we are creating a *feature pool*. In the training phase it is selected which features are actually going to be used.

In our feature pool we have implemented census features (also known as local binary

```
function CreateFaceEngine()
   -- Create an empty engine instance
   engine = NewEngine()

   -- Add a 3x3 mean filter to the engine
   engine:AddMeanFilter()

   -- Define the model used for detection
   model = "FaceFront_24x24_2008_08_29_161712_7"

   -- Add the face scanner
   engine:AddImageScanner(
      0, 0, 1, 1,   -- Search region in image
      0, 1,         -- Min/max face size
      engine:GetPyramidScanner(
         engine:GetGridScanner(
            model,         -- Defined above
            2,             -- Search step x
            2              -- Search step y
         ),
         1.24,     -- Scale step
         2         -- Thread count
      )
   )

   -- Define a selector that selects
   -- only objects of type "Face"
   selector = 'Object.Type == "Face.Front"'

   -- Remove multiple detections
   engine:AddObjectThinner(
      selector,  -- Only frontal faces
      0.2,       -- Min overlap
      "Score"    -- Rating key
   )

   -- Calculate the eye distance and add it as
   -- an attribute with key 'EyeDistance'
   engine:AddObjectAccess(
      selector,  -- Access only faces
      [[
         le = Object.Marker.LeftEye
         re = Object.Marker.RightEye

         dx = le.X - re.X
         dy = le.Y - re.Y

         ed = math.sqrt( dx*dx + dy*dy )

         Object.Attribute["EyeDistance"] = ed
      ]]
   )

   return engine
end
```

Listing 1: The engine setup script for the sample configuration of figure 3.

patterns) [ZW96]. These features are defined as structure kernels of size 3×3 which summarize the local spatial image structure. Within the kernel structure information is coded as binary information $\{0, 1\}$ and the resulting binary patterns can represent oriented edges, line segments, junctions, ridges, saddle points, etc. In addition we are using resized

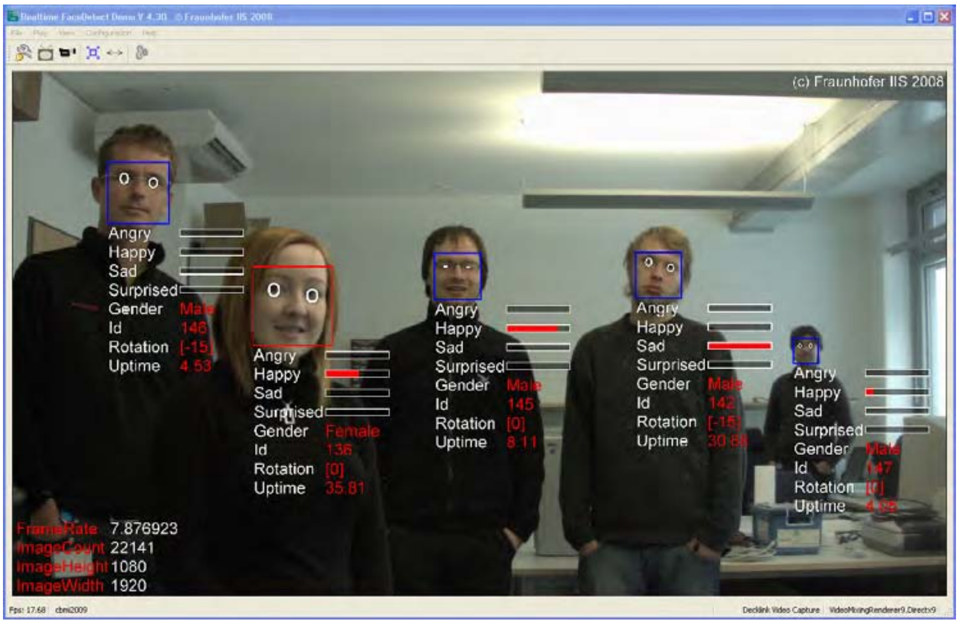


Figure 4: Screenshot of our demo application.

versions of these features which means that the features are not only calculated on simple 3×3 patterns but also on $3n \times 3m$ multiples of these windows.

Other features in the pool are edge orientation features. They can be calculated for each feature as $atan2(sx, sy)$ whereas sx and sy mean the result of the sobel filter applied to the respective pixel in x- and y- direction similar to the description in [FK02].

Doing so, starting from a 24×24 window we yield a total of approximately 30000 features, building the set $\mathcal{F}$ of available features.

During the training phase these features are selected and weighted using the AdaBoost method. By using boosting a number of weak classifiers are combined to form a final strong classifier. The goodness of a weak classifier is measured by its error ϵ on the training set. Our classifier then consists of a set of look-up tables $\{h_{\mathbf{x}}; \mathbf{x} \in \mathcal{F}\}$ for the features $\mathbf{x}$ chosen by the algorithm. Each look-up table holds a weight for each feature value. This method is described in more detail in [KE06, FS99].

One important advantage of using the AdaBoost method is that we gain a very fast classifier that uses look-up tables for the classification process. The detection can be further improved by using multiple stages as described in [VJ01] and [KE06]. Further speed improvements are made by carrying out a coarse-to-fine grid search as explained in [FK02].

5 Benchmark Results

The performance of our system was tested on several image data sets. None of the datasets that were used for benchmarking are part of the training sets, all sets of data are disjoint. We show rates for face detection, gender classification and the classification of happy faces. The face detection rates were calculated on the CMU+MIT (consisting of 130 images with 507 faces) and the BioID data base (consisting of 1521 images with 1522 faces[1]). The outcomes are shown in figure 5. When examining the detection rate for certain false positives we compare to competing technologies as shown in table 1. The numbers from Viola and Delakis are picked from the paper [RV08], where also other methods are compared.

false positives	10	31	65
This work	91.5	93.3	93.9
Viola and Jones[VJ01]	83.2	88.4	92.0
Garcia and Delakis[GD04]	90.5	91.5	92.3

Table 1: Comparison of selected algorithms with respect to detection rate on the CMU+MIT dataset. All numbers are given in percent.

The following measurements were made on images where we used pre-annotated eye-positions. The given recognition rates are based on the optimal classification thresholds choosen separately for each test data set.

The performance of our gender classification module was tested on the BioID data set and on the Feret fafb data set [PWHR98]. We receive a recognition rate of 94.3% on the BioID data base and 92.4% on the FERET fafb data base.

Our happiness analyzer was evaluated on the JAFFE data base [LAKG98]. It consists of 213 images of Japanese women with 31 annotated as happy. Here the recognition rate is 95.3%.

In table 2 we show the results of speed measurements on our system. The measurements were carried out on an Intel Core 2 Duo 6420 CPU. Although the library supports multi-threading we only used one core in this test. For evaluation we used the full BioID data set containing images of size 384×286 and we determined the average calculation time per image.

face detection	×	×	×	×
eye fine search		×	×	×
gender classification			×	×
analysis of 4 expressions				×
time [ms]	9.4	19.3	19.9	22.0

Table 2: Assessment of computation time for different engine setups on a single core of a Intel Core 2 Duo 6420 CPU for an image of size $\mathbf{384} \times \mathbf{286}$.

[1] In the BioID data set there is one image (number 1140) showing two faces.

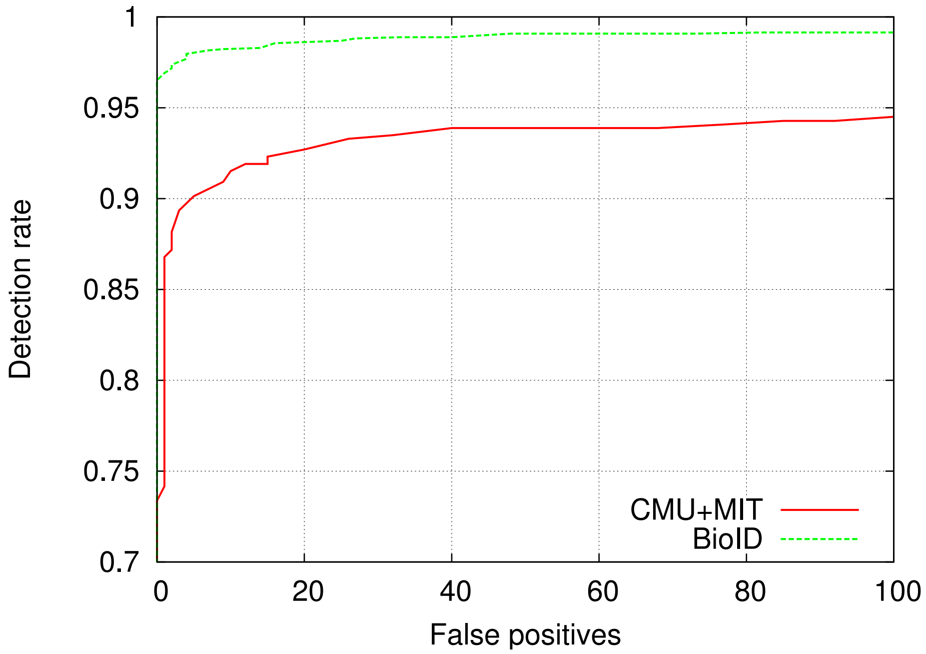


Figure 5: ROC-curve for the two data bases BioID [Fri] and CMU+MIT [Wan].

6 Conclusions

In this paper we have shown an approach to a modular framework to detect and analyze faces. The system features flexibility and modularity so that the integration to specific tasks can be carried out efficiently. We have shown a variety of different recognition tasks that can be combined by setting up the recognition engine in different ways. In addition it is easy to use the script Language Lua to add additional functionality like script logging or a XML-based TCP/IP-Server. We think that due to these properties the system is well suited for running in a real-time environment for audience measurement.

Everyone who is interested trying out the system can download and evaluate our demo software from `http://www.iis.fraunhofer.de/EN/bf/bv/kognitiv/biom/dd.jsp`. At the moment we are involved in adding functionality like age estimation and pose estimation. In addition we intend to train further classifiers for example for robust hand detection. Furthermore the different components like feature extractors and classifiers will probably be extended in the future.

This research work was partly supported by the European Commission under contract: FP6-2005-IST-5 (IMAGINATION), `http://www.imagination-project.org`.

References

[FK02] Bernhard Fröba und Christian Küblbeck. Robust Face Detection at Video Frame Rate based on Edge Orientation Features. In *International Conference on Automatic Face and Gesture Recognition (FG '02)*, Seiten 342–347, Washington D.C., May 2002.

[Fri] Robert Frischholz. The BioID Face Database. Website. Available online at `http://www.bioid.com/downloads/facedb/index.php`; visited on January 7th 2009.

[FS99] Yoav Freund und Robert E. Shapire. A Short Introduction to Boosting. In *Journal of Japanese Society for Artificial Intelligence*, number 14, Seiten 771–780, September 1999.

[GD04] Christophe Garcia und Manolis Delakis. Convolutional Face Finder: A Neural Architecture for Fast and Robust Face Detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(11):1408–1423, 2004.

[GHJV94] Erich Gamma, Richard Helm, Ralph Johnson und John Vlissides. *Design Patterns - Elements of Reusable Object-Oriented Software*. Addison-Wesley, 1994.

[IdFC06] Roberto Ierusalimschy, Luiz Henrique de Figueiredo und Waldemar Celes. *Lua 5.1 Reference Manual*, August 2006. Available online at `http://www.lua.org/manual/5.1`; visited on January 7th 2009.

[KE06] Christian Kueblbeck und Andreas Ernst. Face Detection and tracking in video sequences using the modified census transformation. *Image Vision Computing*, 24(6):564–572, June 2006.

[LAKG98] M. Lyons, S. Akamatsu, M. Kamachi und J. Gyoba. Coding facial expressions with Gabor wavelets. *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, Seiten 200–205, Apr 1998.

[lua03] *Programming in Lua*. Roberto Ierusalimschy, first. Auflage, 2003. Available online at `http://www.lua.org/pil`; visited on January 7th 2009.

[PWHR98] P. Jonathon Phillips, Harry Wechsler, Jeffery Huang und Patrick J. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5):295 – 306, 1998.

[Rie00] Dirk Riehle. *Framework Design - A Role Modeling Approach*. Dissertation, Swiss Federal Institute of Technology Zurich, May 2000. Available online at `http://www.riehle.org/computer-science/research/dissertation/diss-a4.pdf`, visited on January 7th 2009.

[RV08] Mauricio Correa Rodrigo Verschae, Javier Ruiz-del-Solar. A unified learning framework for object detection and classification using nested cascades of boosted classifiers. *Machine Vision and Applications (2008) 19:85103*, 2008.

[VJ01] Paul Viola und Michael Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition*, 2001.

[Wan] Chieh Chih Wang. CMU Image Data Base: face. Website. Available online at `http://vasc.ri.cmu.edu/idb/html/face/frontal_images/`; visited on January 7th 2009.

[ZW96] Ramin Zabih und John Woodfill. A Non-Parametric Approach to Visual Correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996.

Sensor Web and Geoprocessing Services for Pervasive Advertising

Theodor Foerster , Arne Bröring , Simon Jirka , Jörg Müller

International Institute for Geoinformation Science and Earth Observation, ITC, Enschede, the Netherlands
foerster@itc.nl

Institute for Geoinformatics, University of Muenster, Muenster, Germany
{arne.broering,simon.jirka,joerg.mueller}@uni-muenster.de

Abstract: Pervasive advertising attracts attention in research and industry. Sensor information in this context is considered to improve the content communication of Pervasive Environments. This paper describes an architecture for integrating sensor information into Pervasive Environments. The sensor information is accessible through an abstraction layer, the Sensor Web, which is based on Web Service technology. The Sensor Web provides access to any deployed sensor for any compliant infrastructure, such as a Pervasive Environment. It thereby does not only access sensors that are deployed specifically for this system, but any sensor in the world that is available through the Sensor Web. In order to extract specific sensor information from the available sensor data, Geoprocessing Services are deployed as an intermediate component in the proposed architecture.

1 Introduction

Advertising currently introduces novel pervasive technologies installed in urban spaces (e.g. digital signage, mobile phones). These new media enable a) to show live data about the environment collected by sensors (e.g. weather forecast) and b) to adapt information presentation to the context, as determined by sensors (e.g. ads for ice cream, when the sun is shining). Currently, local and individual sensors are deployed at the devices that show the adverts. For example, cameras are installed on digital signs and GPS devices are installed in cell phones.

At the same time sensor information and geoprocesses become available on the web through Web Service technology. This trend is summarized by the term of the Geospatial Web [LA07]. The Geospatial Web evolves due to the advancements in processing and network capabilities as well as in standardization. The Sensor Web provides an abstraction layer on top of the individual sensor hardware and allows applications (e.g. advertising services) to access whole sensor networks independently of individual, often proprietary sensor hardware interfaces. Such sensor environments are already deployed in public spaces and provide measurements of real world phenomena (such as temperature or air pollution). To finally provide sensor information customized for individual applications, geoprocesses are a valid means to transform data (as for instance provided by the Sensor Web) into information. To enable the vision of such web-based sensor information Geoprocessing Services evolve currently.

Integrating Sensor Web and Geoprocessing technologies for pervasive advertising has not been achieved yet. However, this is promising to release the potentials of sensor networks for pervasive advertising and to open up the Sensor Web for a promising new type of application. Such enhanced pervasive advertising applications will provide more adequate and up-to-date information to the individual user and thereby increase the impact of the displayed content.

This paper proposes an architecture based on current standards of the Geospatial Webby integrating sensor information a) to represent it on the display and b) to adapt content presented on the display by measuring the impact of the displayed content through a feedback loop. Section 2 introduces the related concepts of pervasive advertising, Sensor Web and geoprocessing. Based on the introduced concepts Section 3 describes an architecture integrating Sensor Web and geoprocessing for pervasive advertisement applications. Section 4 discusses the architecture and provides an outlook.

2 Related Concepts & Literature

This section reviews the current activities in the context of pervasive advertising and the Geospatial Web. The Geospatial Web in a nutshell describes distributed services providing data and information. Two aspects of the Geospatial Web are the Sensor Web and Geoprocessing Services. In the Geospatial Web, the communication between the services is ensured through standardized data formats and service interfaces. In this context the Open Geospatial Consortium (OGC) plays a major role in setting standards and best practice approaches.

2.1 Pervasive Advertising

Recent advances in computing technology lead to vastly increased numbers of computers embedded into everyday items and environments, a paradigm called pervasive computing or ubiquitous computing [WE91]. For example, in the year 2007, 78,4 % of German citizens owned a mobile phone [AL07], and advances in display technology like LED, LCD, OLED and e-paper displays lead to increasing numbers of digital displays installed in public spaces. These new media are creating a pervasive information environment, where every citizen in urban spaces can be reached at almost every time via a variety of different channels. Advertising is expected to be an important business model in these pervasive information environments [MU09, RA02, KU06]. Because these media are dynamic, they present new challenges for sensor technology, both to deliver content to be presented in the media, and to determine the context of the user in order to adapt content presentation to this context [MU09a]. Because the users are often mobile, location is one of the most important context factors, such that it makes sense to use space and time as a main framework in which to interpret sensor data.

2.2 Sensor Web

Several elements of the architecture presented in this paper rely on standards published by the OGC and especially the OGC Sensor Web Enablement (SWE) suite of standards.

The OGC is an international body for standardization comprising more than 380 members from industry, research, education and governmental agencies. Its activities are centered on realizing the vision of the so called Geospatial Web. This term refers to the aim of integrating all kinds of geospatial data sources as well as processing functionalities into a common framework so that these resources become accessible through the Web. Among the different topics addressed by the OGC for the system presented in this paper, the integration of sensors and sensor data into the Geospatial Web is of special interest. This aspect is the core focus of the OGC Sensor Web Enablement Working Group [BO07].

The work performed within the SWE Working Group has lead in the last years to the creation of a set of standards that aim at fulfilling the following objectives of the Sensor Web:

- Access to sensor data (real time and historic observations)
- Controlling sensors (i.e. setting sensor parameters)
- Alerting based on user-defined alert conditions and sensor measurements (e.g. if a certain threshold is exceeded)
- Accessing sensor parameters (e.g. information about the sensor configuration)
- Provision of sensor descriptions (sensor metadata)

- Discovery of sensors and sensor data.

As a result an architecture formed by a framework of standards has been developed. This framework can be divided into two parts that address the two main aspects which have to be standardized: the SWE Information Model comprising data formats for sensor data as well as sensor metadata and the SWE Service Model consisting of (web service) interface standards offering different types of sensor related functionality.

In Figure 1 the different standards forming the two sub models of the SWE framework are shown. In the next paragraphs these standards will be introduced in more detail in order to allow a better understanding of the system architecture presented in this paper.

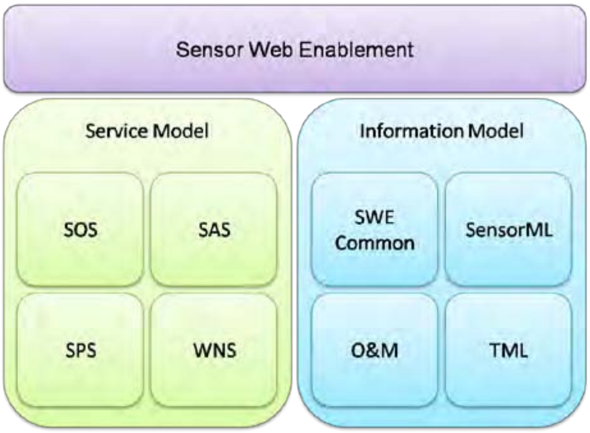


Figure 1: Overview of the SWE framework.

As outlined before the standards for encoding sensor data and metadata are contained in the SWE Information model. Thus, these standards ensure a common data model for exchanging sensor related data between SWE services and client components. There are four different standards which form the SWE Information model:

- SWE Common [BO07a]: In order to avoid overlap between the different SWE standards (i.e. the re-definition of certain simple data types within each standard), SWE Common provides several basic building blocks that are common to all SWE standards.

- Observations and Measurements (O&M) [CO07, CO08]: O&M offers a common encoding for data captured by sensors.

- Sensor Model Language (SensorML) [BO07a]: In order to correctly interpret sensor data or for allowing discovery of sensors, there is a need for describing sensor metadata. This issue is solved by SensorML which offers a standardized encoding for sensor metadata.

- Transducer Markup Language (TML) [HA07]: TML is a data format which allows the encoding of sensor data as well as sensor metadata. However compared to O&M and SensorML, TML has been optimized for supporting data streaming.

Complementary to the SWE Information Model, the SWE Service Model is formed by four different standards that define interfaces for accessing sensor related functionality. In detail, the following four standards are contained in the SWE Service Model:

- Sensor Observation Service (SOS) [NA06]: The SOS is intended for accessing sensor data and metadata. Its operations allow requesting sensor data based on spatial, temporal and thematic filter criteria as well as retrieving metadata for sensors. Furthermore the SOS offers operations for inserting observations and sensors. Typically, the sensor data is returned as O&M, metadata are usually returned as SensorML documents.

- Sensor Alert Service (SAS) [SI06]: The SAS offers a complementary approach for accessing sensor data. Whereas the SOS follows a pull-based communication model, the SAS is capable of pushing sensor data to subscribers. These subscribers usually define (i.e. through alert conditions) in which data they are interested. As a consequence the SAS is especially optimized for generating alerts if user defined conditions occur.

- Sensor Planning Service (SPS) [SI07]: The SPS can be used for controlling and tasking sensors. The most common use case is setting sensor parameters like the sampling rate. Furthermore the SPS offers a broad range of operations for managing sensor tasks (e.g. deleting, updating and cancelling tasks).

- Web Notification Service (WNS) [SI06a]: The WNS does not offer directly sensor related functionality. Instead it has to be understood as a kind of helper service enabling asynchronous communication within the SWE architecture. One important use case is the transmission of alerts generated by the SAS to users via communication means like SMS or email.

2.3 Geoprocessing Services

A geoprocess is considered to transform geodata into geoinformation. An example of a geoprocess can be a simple buffer calculation on specific objects or more complex calculations of climate change models. Complex geoprocesses are modeled as workflows, which consist of multiple processing steps.

As mentioned before, Geoprocessing Services gain currently attention from research and industry, as they enable the shift from web-based geodata to web-based geoinformation. They are an integral part of the vision of the Geospatial Web. Brauner et al. [BR09] describe the current challenges in the context of web-based geoprocessing: improving performance, enhancing service orchestration and semantic process descriptions.

One of the major attempts to standardize geoprocesses on the web is the OGC Web Processing Service interface specification (WPS) [OG07]. It describes a standardized set of operations to publish and execute any type of geoprocess on the web. According to the WPS interface specification, a process is defined as an algorithm, calculation or model that operates on spatially referenced data.

In detail, the WPS specification describes three operations, which are all handled in a stateless manner: *GetCapabilities*, *DescribeProcess* and *Execute*. GetCapabilities is common to any type of OGC Web Service and returns service metadata. In case of a WPS it also returns a brief description of the processes offered by the specific WPS instance. To get further information about the hosted processes, the WPS is able to return the process metadata through the *DescribeProcess* operation. This operation provides the description of all parameters, which are required to run the process. Based on this information the client can perform the *Execute* operation upon the designated process. As any OGC Web Service, the WPS communicates through HTTP-GET and HTTP-POST based on an OGC-specific XML-encoding.

WPS implementations have already been successfully applied in several projects ranging from groundwater vulnerability analysis [KI06] and map generalization [FO06]. WPS is also used for workflow modeling [SC08] and integration into Google Earth™ [FO09]. Additionally, an extensive discussion about the applicability of the WPS and its current drawbacks can be found in [FR07].

3 A Geospatial Web Architecture for Pervasive Advertising

The architecture applies Geospatial Web technologies to pervasive advertising. The pervasive advertising application accesses the Geospatial Web to provide relevant information to the user and to sense the context of the user. To provide the sensor data as information to the user, specific processing is required. This processing is performed by single or chained Geoprocessing Services.

The conceptual architecture is depicted in Figure 2. The pervasive environment retrieves all the information through Geo Processing and Sensor Web facilities. In particular, the Sensor Web measures the user context and the global context. The user context is provided for instance by cameras or profile settings of mobile phones. The global context is measured by for instance weather stations and satellites. We foresee that the distinction between user context and global context will disappear in the future technology-wise, as technology will advance and provide information about global and user context using any sensor. Though, the Sensor Web provides a coherent infrastructure and a common interface to all this information. Geo Processing facilities extract the relevant information. Based on this information the pervasive environment can decide what to display and measure the reaction of the user as a feedback loop.

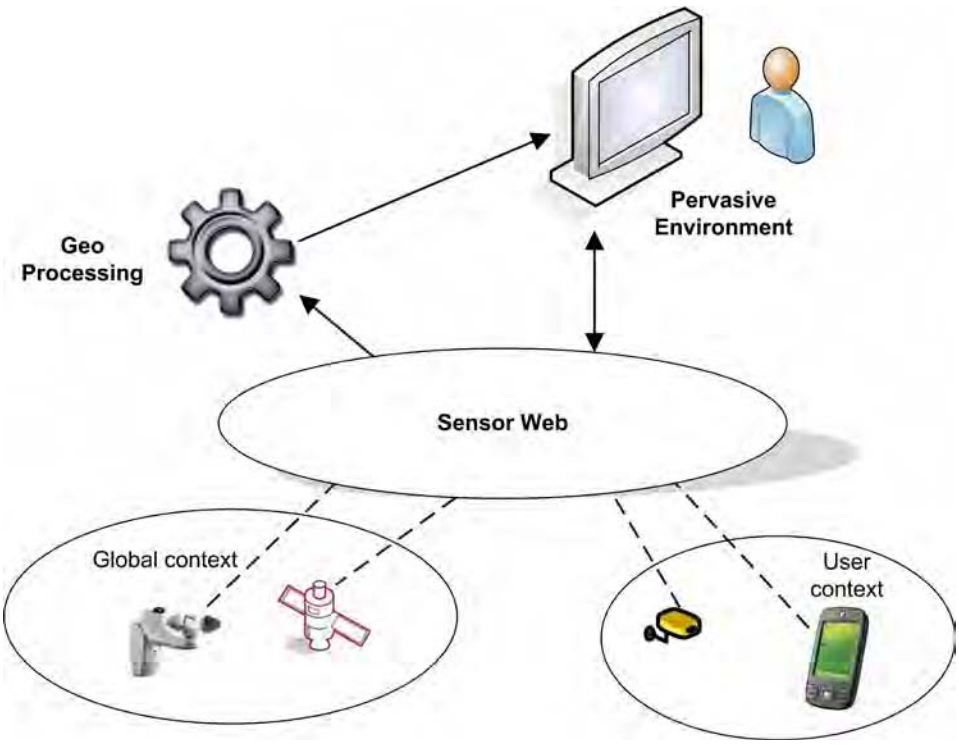


Figure 2: Conceptual architecture.

Figure 3 depicts the mapping of the conceptual architecture to the different service interfaces defined by the Geospatial Web.

In particular, three modes of interaction can be defined in the architecture, as explained in the following.

1. Pervasive Environment displays sensor data.

Sensors can be made available on the Sensor Web by registering them at one of the web services such as the SOS. To be able to do that a metadata description of the sensor in form of a schema-valid SensorML document has to be supplied. Sensors currently accessible on the Sensor Web range from simple thermometers attached to weather stations up to standard web cameras or even satellites. The data gathered by those sensors is ingested and subsequently provided by SOS instances. To access and use this sensor data the Pervasive Environment can on the one hand directly invoke the SOS to request the data in the O&M format. Since applications from the Pervasive Environment domain are often not capable of handling this format, we are proposing to pass the data firstly to a particular WPS instance to transform the O&M data to a more common format (e.g. KML, JPEG or MPEG). The data returned by the WPS is then easily displayable on the user interface.

2. Pervasive Environment reacts on events/alerts occurring in the Sensor Web

Besides a simple querying and subsequent visualization of sensor data the Sensor Web technology can also be used to react on events triggered by sensors or sensor networks. Therefore a sensor has to be registered as a data publisher at an SAS instance. This service is constantly listening for incoming data and forwards it to interested parties (this can be other web services such as an SOS) and computes alerts if for instance certain thresholds are exceeded. If such an alert occurs the SAS notifies applications and users by means of a WNS. An application, for example of the Pervasive Environment, can register at the WNS to receive notification messages via a previously specified delivery method such as SMS, FAX, e-mail or even phone calls. Once the application is notified about the occurrence of an alert or event it can adjust its display by querying a new data visualization from the WPS or can react in other ways.

3. Pervasive Environment influences sensors.

Also, it is possible for the Pervasive Environment to directly influence the behavior and state of the Sensor Web. Sensors which are registered at a SPS can be tasked. The service offers a standardized way to control sensors of arbitrary type. For instance an application of the Pervasive Environment could change the zoom level or angle of a web camera or modify the sampling rate of thermometer.

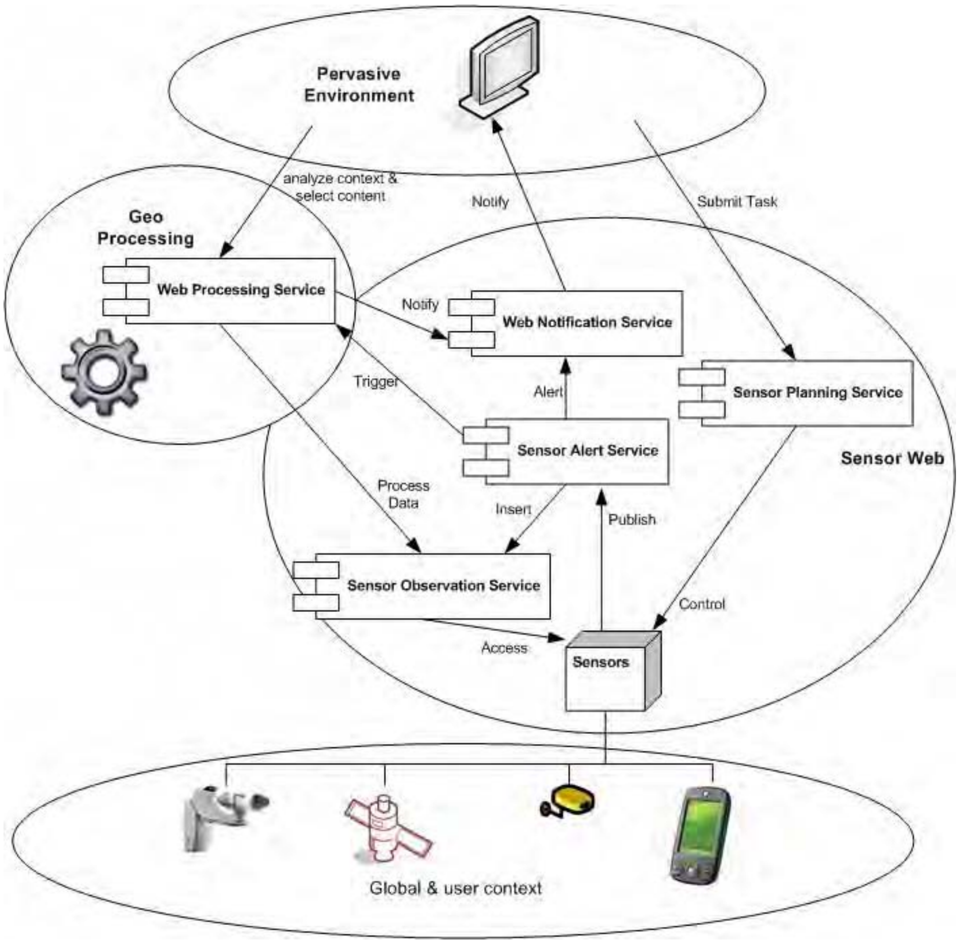


Figure 3: Implementation architecture accessing the Geospatial Web for pervasive applications.

A conjunction of the Geospatial Web and the Pervasive Environment, as outlined in the architecture above, allows the display to portray live data from global and user context as well as to adapt information presentation to these contexts. Data from the Sensor Web can be retrieved via standardized web service interfaces. The ability to send notifications in case of occurring events allows a live display of relevant information. The Sensor Web technology can be used to adapt the Pervasive Environment and the presented information on-the-fly to the context, as determined by sensors. An application, e.g. a public display showing advertisements, can adjust itself once it receives notifications or alerts from the Sensor Web. If for instance a thermometer registered on the Sensor Web starts to measure high temperatures the public display application could adjust itself by switching to an ice cream advertisement.

4 Outlook and Conclusion

Our work presents an architecture based on the Geospatial Web to improve the content communication of pervasive advertising applications. The pervasive environment such as a pervasive advertising application accesses the sensors which observe the global and user context. Based on the gathered information from sensors and processed through Geoprocessing Services the pervasive environment can adapt the displayed information. Due to the standardized interfaces defined by the Sensor Web, pervasive advertising systems can use any kind of available sensor such as weather stations or web cameras. Geoprocessing Services provide a standardized way to derive higher-level knowledge from the retrieved sensor data. The components relating to the Geospatial Web, which are part of the presented architecture, are already implemented and available as open source software solutions (see http//www.52north.org). However, a future challenge will be to create a proof-of-concept prototype for a pervasive advertisement application and to show the feasibility and utility of the proposed approach.

References

[AL07] Allensbacher Markt- und Werbeträger-Analyse: Handybesitz. http://de.statista.org/statistik/diagramm/studie/12740/umfrage/handybesitz/, 2007.

[BO07] Botts, M.; Percivall, G.; Reed, C.; Davidson, J.: OGC Sensor Web Enablement: Overview And High Level Architecture, Wayland, MA, USA: OGC, 2007.

[BO07a] Botts M.; Robin, A.: OpenGIS Sensor Model Language (SensorML) Implementation Specification 1.0.0, Wayland, MA, USA: OGC, 2007.

[BR09] Brauner, J.; Foerster, T.; Schaeffer, B.; Baranski, B.: Towards a Research Agenda for Geoprocessing Services. In *12th AGILE International Conference on Geographic Information Science*, edited by Jan Haunert, B. Kieler, and J. Milde. Hanover, Germany: IKG, Leibniz University of Hanover, 2009. http://www.ikg.uni-hannover.de/agile/fileadmin/agile/paper/124.pdf.

[CO07] Cox, S.: Observations and Measurements - Part 1 - Observation schema 1.0., Wayland, MA, USA: OGC, 2007.

[CO08] Cox, S.: Observations and Measurements - Part 2 - Sampling Features 1.0, Wayland, MA, USA: OGC, 2008.

[FO06] Foerster, T.; Stoter, J.E.: Establishing an OGC Web Processing Service for generalization processes. *ICA workshop on Generalization and Multiple Representation*, Portland, Oregon, USA: 2006.

[FO09] Foerster, T.; Schaeffer, B.; Brauner, J.; Jirka, S.: Integrating OGC Web Processing Services into Geospatial Mass-Market Applications. In *International Conference on Advanced Geographic Information Systems & Web Services, 2009.*, 98-103. Cancun, Mexiko: IEEE, 2009.

[FR07] Friis-Christensen, A.; Ostlander, N.; Lutz, M.; Bernard, L.: Designing Service Architectures for Distributed Geoprocessing: Challenges and Future Directions. *Transactions in GIS*, vol. 11, Dec. 2007, p. 799—818.

[HA07] Havens, S.: OpenGIS Transducer Markup Language (TML) Implementation Specification 1.0.0, Wayland, MA, USA: OGC, 2007.

[KI06] Kiehle, C.: Business logic for geoprocessing of distributed data. *Computers & Geosciences*, vol. 32, 2006, pp. 1746-1757.

[KU06] Kurovsky et al. Using ubiquitous computing in interactive mobile marketing, Personal and Ubiquitous Computing 2006 10, Springer, Berlin (2006).

[LA07] Lake, R.; Farley, J.: Infrastructure for the Geospatial Web. In *The Geospatial Web*, edited by Arno Scharl and Klaus Tochtermann, 15-26. Advanced Information and Knowledge Processing Series. London, UK: Springer, 2007.

[MU09] Müller, J.; Schmidt, A.; Quigley, A.; Begole B.: Proceedings of the 1st Workshop on Pervasive Advertising, http://pervasiveadvertising.org, International Conference on Pervasive Advertising, Nara, 2009.

[MU09a] Müller, J.: Context Adaptive Digital Signage in Transitional Spaces, PhD thesis, University of Münster, 2009.

[NA06] Na, A.; Priest, M.: Sensor Observation Service 1.0, Wayland, MA, USA: OGC, 2006.

[OG07] OGC: *OpenGIS Web Processing Service*. OGC implementation specification. Open Geospatial Consortium, 2007. http://www.opengeospatial.org/standards/wps.

[RA02] Ranganathan, A.; Campbell, R.H.: Advertising in a Pervasive Computing Environment, Proceedings of the 2nd international workshop on Mobile commerce, Atlanta, 2002.

[SC08] Schaeffer, B.; Foerster, T: A Client for Distributed Geo-Processing and Workflow Design. *Journal for Location Based Services* 2, no. 3 (September 2008): 194-210.

[SI06] Simonis, I.: OGC Sensor Alert Service Candidate Implementation Specification 0.9, Wayland, MA, USA: OGC, 2006.

[SI06a] Simonis, I.; Echterhoff, J.: Draft OpenGIS Web Notification Service Implementation Specification 0.0.9, Wayland, MA, USA: OGC, 2006.

[SI07] Simonis I.; Dibner, P.C.: OpenGIS Sensor Planning Service Implementation Specification 1.0, Wayland, MA, USA: OGC, 2007.

[WE91] Weiser, M: The Computer for the 21st Century. Scientific American, 1991, 265 (3), pp. 94-104.

Tokenized Interaction Architecture

Boxian Dong, Bernhard Wally, Alois Ferscha

{boxian.dong, bernhard.wally}@researchstudio.at, ferscha@pervasive.jku.at

Abstract: Out-of-Home (OOH) advertising is currently advancing into a new era: "pervasive advertising" is one of the buzz words describing a soft change from the traditional push-only media to an interactive and bi-directional advertisement experience. While ambient media solutions try to conquer each and every piece of potential advertisement surfaces by various sorts of print media, another trend can be observed as well: OOH advertising is becoming smarter (at least from an electronics point of view). Digital signage products, SMS, Bluetooth interaction in city lights and talking billboards that react on the feedback of motion sensors—pervasive advertising is on its way. However, the development of a pervasive advertisement platform is beyond the scope of any single engineering or art discipline; instead it is a multidisciplinary effort including electronic engineering, computer science, art, cognitive psychology and even social sciences. Therefore, we introduce a *tokenized interaction architecture*; a foundation on top of low level issues that simplifies sub sequent implementation efforts by addressing interaction, context awareness and hardware heterogeneity, so to release designers of pervasive advertising campaigns from computer related challenges.

1 Introduction

New wireless communication technologies, the unceasing evolution of CPU power and efficiency as well as improvements in signal processing systems make the Weiser's (1993) vision more and more realistic. Many implementations show the feasibility of systems where numerous computing devices are widely distributed throughout our natural environment but stay invisible to the human eye, operating in the background to support natural human behavior and interaction. At the same time, the advertisement industry is driving and expanding their market by permanently looking for and introducing new ways of advertising. Pervasive computing is a powerful force to advance and support the advertisement industry, because of its capabilities like personalization and ubiquity.

In our vision, it is necessary to introduce a relatively specific architecture to be commercially implemented, with the following key aspects: integrability into existing advertising platforms, hardware abstraction (release designers from technological burden) and a clear focus on innovative forms of advertisement. This paper first analyzes the features of pervasive advertisement systems and their implementations in Sec. 2 to sub sequentially abstract the requirements of such systems. Sec. 3 will first conclude the related work in pervasive platform research, and then introduce a tokenized interaction architecture, optimized for interactive advertising. Sec. 4 presents a proof-of-concept demonstration of the system, followed by conclusions and an outlook in Sec. 5.

2 Related Work

2.1 Advertising Technologies

Advertisements that adopt pervasive computing technologies have been discussed and researched for years—several pervasive systems have been built to deliver context aware ads at the right time, to the right person at the right place [RC02]. [Gra02] and [BK01] use mobile phones as the main channel to display advertisements and other information, in [SPD06] a system that collects audience statistics by exploiting Bluetooth to detect people's presence is presented. [SRA09] shows how to track customer information by profiles generated by users' Facebook accounts through Bluetooth and RFID, while [BLSC01] and [BL97] use the interactive digital home as platform for advertisement delivery. [BK01] and [MKHS08] generate advertisement with capabilities to let people interact implicitly, [SRA09] and [BL97] focus on explicit interaction. Besides those applications, there are also some public display based implementations, such as those from [MAMI07], [BPS+08] and [BFS+08], where the usage of public displays also inherently provides potential for advertisement applications.

Previous work indicates how methods and technologies of pervasive computing can be used in the advertisement market as well as in public information systems. Based on these findings the remainder of this section will first analyze the challenges of and possible solutions for pervasive advertisement and then conclude with the requirements of a universal architecture to be capable of supporting future advertisement implementations.

2.2 Context Frameworks

[SBHM02] solved the communication decoupling by using context-based message exchange (context-matching) and unifying the messages produced by the input/output devices (common interface). However, [SBHM02] doesn't decouple the encapsulating software from the device, and runs the interface support tools and the processing engine on the physical device. This increases the computing efforts on the physical device and therefore limits the choice of the technology of the end device. [VSdI06] and [CFJB05] release the end device from computational expensive processes by introducing the context-broker that produces user-agents that run on a server on behalf of the end devices. This approach makes it possible to use cheap devices in the system like simple sensors or wirelessly controlled power switches by leaving the computing work to the user-agent in the network. [VSdI06] also improved the system capabilities presented in [CFJB05]'s by decentralizing the user-agent: through the introduction of the orchestrator which scans, matches and controls the distributed user-agents. However, [VSdI06] did not distribute the orchestrator itself and implemented both the intelligence function and the matching function in the Orchestrator, limiting the scalability and maintainability of the system. [BDO+04] divided the ambient intelligence into three layers: environmental, reactive and deliberative. The environmental layer abstracts all low-level interaction (sensors, actuators), the reactive layer delivers quick responses (and forwards more complex situations to the deliberative layer) and the deliberative layer handles all the complex tasks. This approach reduced the complexity of the implementation by simplifying the functions of each module, and increased the maintainability of the overall system. [BCA$^+$08] also decouples the communication from context-matching. The system uses a center arbitrator Adaptation Decision Engine (ADE) to manage the communications, and uses a distributed Adaptive Engine (AE) to manage the behavior and intelligence of the end devices. This approach masks the communication and underlying technology details to a maximum, and enables the higher level engineers to concentrate on their applications. However, the center arbitrator Adaptation Decision Engine depicts a bottle neck of the system. Another framework for "the internet of things" has been presented in [RDG$^+$08], involving OSGi with extensions such as R-OSGi over Bluetooth or IEEE 802.15.4. The system proposed modularizes the system in that it maps each hardware device as a software service into a distributed platform with support for streaming data such as continuously measured sensor data. From a technical point of view, the so called "software fabric" is closely related to the system presented in this document, especially when it come to the mapping of hardware devices to software service instances. However, while the system presented in [RDG$^+$08] is basically a distributed framework for transparent access to various devices over a unified Java platform, the focus of our platform is on the ad-hoc communication of mobile devices ("tokens") with the environment by exchanging a small set of information (a "profile"). In our architecture the (remote) transparent invocation of methods is not a basic requirement nor is it the core component to be implemented—we require a platform that is able to react on spontaneous implicit user interaction by adapting service behavior to the profile data read out. The system presented in [SK05] focuses on (tangible) user interaction and methods for multi modal, wearable, augmented reality (AR) based user interfaces. With a clear focus on explicit and immediate user interaction, the system is a framework for handling various input methods in a consistent manner to allow presenting AR scenes accordingly. Our envisioned use case, implicit user interaction in OOH advertisement settings, has no connection to AR technology or tangible user interaction at all. [DAS01] is an early example of a context aware computing platform developed for a smart conference setting. It aggregates various context data including location, user interest level, presentation content, etc. into a context-aware application that can among others support attendees of a conference in selecting workshops in a multi-track schedule based on the interests declared at the registration. Especially the people interests part is of relevance to our own architecture, as it gives an example of how interests (and their weighting) can be used to build an adaptive system.

To solve the challenge of versatility, we adopt the AE concept from [BCA$^+$08], and the context-broker concepts from [VSdI06] and [CFJB05] to minimize the computation requirements at the end device. To solve the challenges of communication decoupling and scalability, we introduce a distributed context-based routing (similar to the context management in [SBHM02], or the AES in [BCA$^+$08], but without intelligence).

3 Tokenized Interaction Architecture

The tokenized interaction architecture is designed to abstract technology related challenges like hardware heterogeneity, methods of communication and context matching in an interactive advertisement environment. We use the term “tokenized” because our concept is based on “tokens”, the core functional units in our architecture. For the user a token could be a mobile phone, smart key fob, etc., while environmental tokens include sensor nodes, display adapters, etc (cf. Fig. 1).

Figure 1: Examples of possible user tokens (not to scale): (left and middle) a custom designed ZigBee platform and its housing (an actual car key), (right) a mobile phone communicating over Bluetooth or Wi-Fi.

3.1 Challenges to Advertising

Advertisements are all about influencing people, e.g. trying to make them purchase specific products or making them believe in the messages transferred by the ads. There are three important challenges to efficiently convince people: information delivery, information attractiveness and feedback, as concluded from [RC02]:

Information Delivery Traditional advertisements are distributed through mass media like TV, radio, billboards and newspapers in a unidirectional, broadcasting way. This works well to deliver information to the people who are interested, but at the cost of lavishing resources, because in order to guarantee the exposure the advertisements need to be delivered over and over again, and (as the main side effect of mass media) even to people that have no interest for certain topics. Pervasive advertising technologies can help to increase the efficiency by knowing the environment advertisements are delivered at [SRA09, BK01, BFS$^+$08].

Information Attractiveness Traditionally, there are two ways to make advertisements more attractive: either through adding emotional elements such as humor or consternation or through a focused adaption to an individual’s personal interests. The first method has the advantage of universality (at the cost of impact) while the second method can only be applied to a certain group of people. In pervasive environments we have the chance to reach people in individual ways that rely on people's interests and can even introduce interactivity thus enabling more enchanting advertisements [BL97, BK01, MAMI07, MKHS08, BPS+08, BFS+08, SRA09].

Feedback The last very important but all too often neglected challenge of advertisement is user feedback. Feedback is not only the source for the advertiser to collect his revenues, but also for the advertisement and media agencies who design and realize the advertisement campaigns to name and number the results of their implementations in order to refine their strategies for future campaigns [SPD06].

3.2 Requirements for a Tokenized Interaction Architecture

Based on the discussion above, it is clear that a true pervasive advertisement architecture includes some kind of a context framework to understand and work with contextual data: when and where, who is around, what are their interests and what are they doing. Content needs to be presented with respect to the environment it is shown in: what kind of content, how to interact with the surrounding (choosing the right modality) and what kind of feedback can be read from the environment. To achieve this, we need an architecture which includes ambient awareness by sensors or "user profiles" present in the environment, which has the capability to choose the best output device based on the information it got from the environment and can record the environment changes for long term analysis and realtime user feedback. Therefore, the basic requirement of the platform is the capability of wiring different kinds of input, output and storage components together.

3.3 General Architecture

The architecture is mainly constructed by three components: one physical component (tokens) and two virtual components (token shadows and a virtual ecosystem). Tokens are the basic functional units in the system, including displays tokens, audio tokens, user tokens, sensor tokens, etc. For each token, there is a virtual representation residing on the servers which we call the "token shadow". Token shadows represent and abstract the heterogeneity of the physical tokens and their communication channels. The sum of the token shadows together with some core and extending services running on distributed servers form the virtual ecosystem. The computational heavy tasks such as supporting a unified interface for different hardware, context matching, communication decoupling and artificial intelligence can be realized in this ecosystem. Therefore, the token itself can be a very low-power device realized in any kind of technology. The ecosystem serves to decouple the communication, so that there is logically only communication inside the virtual ecosystem, while the shadows act as gateways to their physical representations if required (cf. Fig. 2). "Residential servers" act as hubs for ecosystem-side tokens, such as display or sensor tokens and are thus the physical representation of the gateways from the physical tokens to the ecosystem.

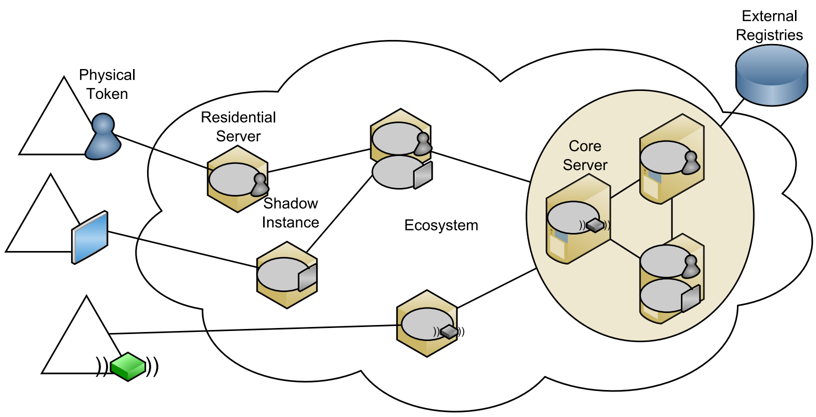


Figure 2: The tokenized interaction architecture: physical tokens and their specific technologies are abstracted by their respective token shadows that exist in a virtual ecosystem as instances of token shadows (those might be distributed within the ecosystem).

Fig. 2 demonstrates the general concept of the architecture. There are three tokens in this scenario: each a display, a user and a sensor token.

- The user token is typically mobile and connects to the system from various locations, while its shadow remains in the virtual ecosystem, allowing shadow instances to follow the user's token. The main purpose of the user token is to provide a user profile to the ecosystem which can be used to infer context information.
- The sensor token resembles an instance of an environmental sensor node that is able to capture environmental information (e.g. the amount of people around, lighting conditions, weather information).
- The display token uses contextual information it can gather from the virtual ecosystem to decide on the content to display. That includes information provided by sensor tokens and profiles injected into the ecosystem by user profiles.
- External registries depict content servers or other databases and computing systems that can be queried in order to get additional context information or to distribute content to display tokens.

As shown in Fig. 2, the ecosystem is to be implemented as a distributed system with a certain hierarchical approach, based on the underlying network topology: "core servers" act as the central entities that can be queried for the ecosystem's state. This mechanism is e.g. used by display tokens to receive information about their environment. Such calls are possibly routed through several residential servers that can act as a cache for the core servers, i.e. if they know the response to a query initiated by a shadow, they can respond accordingly without forwarding the request to the core servers. This mechanism is especially designed for shopping mall scenarios, where the various residential servers might be interconnected over a wireless communication technology like ZigBee, without direct access to the Internet. A special residential server would then act as the gateway of the sensor network to the Internet, thus each query that is targeted to the core servers passes this gateway. By implementing a cache of the core servers on this gateway, multiple requests for the same information could be streamlined to only one request that is sub sequentially delivered to the requesting shadows.

3.4 Software Stacks

The architecture has three software components: the software on the physical token (firmware), the software on the core and residential serves that makes up the virtual ecosystem and the software depicting the token shadows. The latter item is stated separately and not as part of the ecosystem due to the way we look at token shadows: in our vision, the software comprising the token shadows is injected onto the core and residential servers either from authorized entities outside the ecosystem or it might be injected to residential servers by the core servers. The token shadow instances are thus the organisms "living" in the virtual ecosystem.

3.4.1 Physical Token Software

In our model, the physical tokens only have minimal functions to allow small embedded devices made of components with very limited hardware resources. Depending on the actual hardware platform, a token could still be able to execute certain processes, which could possibly help (depending on the situation) reducing communication overheads and saving power.

3.4.2 Virtual Ecosystem

The virtual ecosystem is responsible for constructing the respective token shadow for the physical token and make sure that the shadow is able to communicate with the token.

Fig. 3 shows the stack of the virtual ecosystem's components (residential servers in Fig. 2). The bottom two layers are currently covered by the operation system and the Java VM. The server function layer and thus the token shadow layer are implemented in Java. The challenge of the virtual ecosystem is to

- hide the communication details for components of the ecosystem to talk to each other (cf. Fig. 4) and for the communication between the ecosystem and the physical tokens,
- construct the shadows for newly connected tokens and
- update the shadow data for tokens whose state has changed.

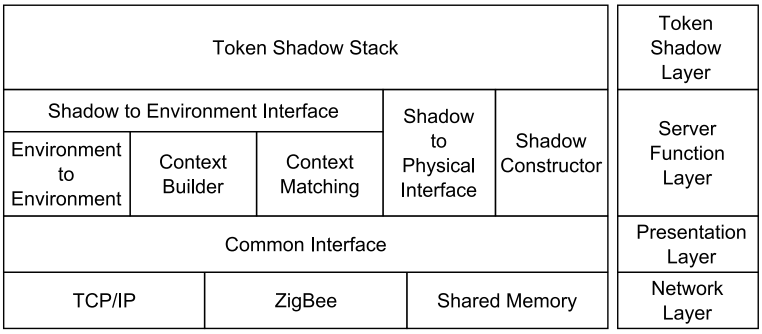


Figure 3: The software stack of the components constructing the virtual ecosystem (residential servers).

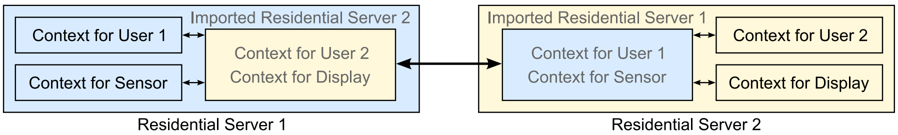


Figure 4: Residential servers talking to each other by importing remote data over the transparent server function layer.

The shadow to shadow communication is abstracted by the server function layer that is responsible for

- handling the context builder, which will register the context of the shadow (what data is supplied for the display tokens, what data is gathered by the sensor tokens),
- matching context information, i.e. forwarding sensed data to shadows that require the respective sensor information or polling sensor data whenever required by another shadow (based on formal context descriptions) and
- handling the communication between the residential servers by querying the core servers for the connection settings for remote residential servers and instantiating the respective communication channels.

3.4.3 Token Shadow Instances

The token shadow instances are the glue between various residential and core servers of the virtual ecosystem. While the residential servers provide the required infrastructure for further information and data exchange, the token shadow instances define what data is provided or required and when and where it is needed (cf. Fig. 5).

The token shadows mainly have two layers, the driver layer, which integrates with the ecosystem and provides a uniform interface to the application layers and the application

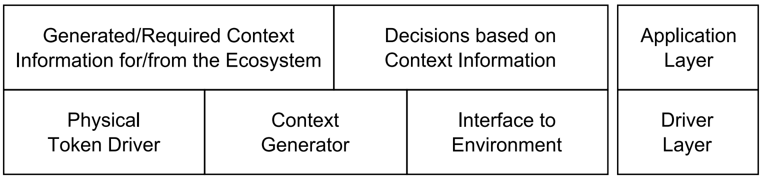


Figure 5: Software stack of the token shadow.

layer, which implement the functionality of the shadow including reasoning, art design or e.g. advertisement related decisions.

4 A ZigBee based Tokenized Interaction Platform

As a proof-of-concept and first demonstration of the platform, we are developing a ZigBee based tokenized interaction platform for implicit interaction in urban areas, especially for shopping mall scenarios, as those are typically crowded public places with the advantages of indoor settings (no heavy rain showers, sun or wind; electricity to run the components required for this platform).

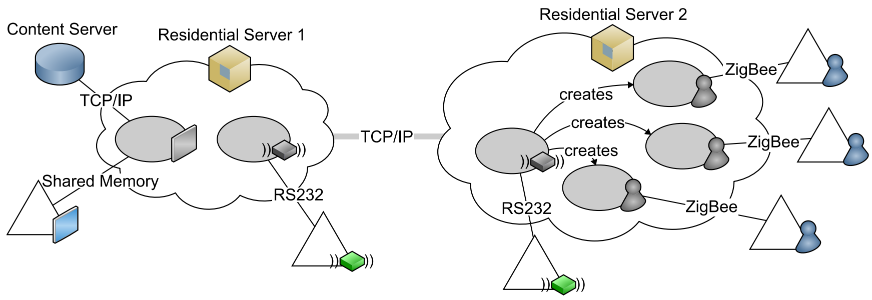


Figure 6: ZigBee based tokenized interaction platform.

The system, as depicted in Fig. 6, is mainly constructed by two residential servers. Server 1 runs two token shadows (one for the sensor token connected via RS232 and one for the display token running natively on the residential server, i.e. the display is connected to the graphics adapter of the residential server). Server 2 incorporates a ZigBee module, connected to the server via RS232, that is able to sense the presence of user tokens over its ZigBee radio interface. It reads out the contextual information stored on the user tokens (user profiles) and instantiates corresponding user token shadow instances on the residential server. The ZigBee communication was implemented using a ZigBee development kit by Texas Instruments, namely the CC2431ZDK. The sensor token on running on server 2 is a TI CC2430 evaluation board, while the user tokens are TI CC2430 demonstration boards, both running our firmware. The sensor token on server 1 is an Arduino Duemilanove which is interfacing ultra sonic sensors to gather environmental information (distances of people nearby). The residential server software stack is implemented in Java SE on a PC platform. While the content to be displayed on the display is planned to be stored in a remote location, it is currently stored on a hard disk inside the residential server 1. The two servers communicate with each other over a socket based communication link using a stream of seralized objects.

The basic idea behind this hardware and software development effort is a shopping mall scenario in which a digital display that is part of a digital signage system can select the content to display by evaluating the current context it is situated in. For this specific scenario this includes information from the ultra sonic sensors (how close are passers-by) and user profile data provided by user tokens. The user profiles depict interests and some basic information about the user and can thus help the digital signage system to display matching content. Of course, it could also just help users find their shops of interest or show the way to the closest money teller or toilet, if required. A person visiting a shopping mall could define his/her personal profile prior to leaving for the shopping mall; once there, the digital displays around would help him/her to navigate through the mall, present special offers and even give discount in the target shops for using the tokenized interaction system. For the shops in the mall this system allows to promote their products and offers target group specific with a probably higher hit rate than a typical mall signage video loop would allow. This, in turn, could lead to a new billing system on the mall signage operator's side: Pay-per-Visualization or Pay-per-Trusted-Contact. The former meaning a payment model linked to the air-time of the displayed content, the latter describing a payment model linked to the number of profile matches found during corresponding content replay.

With the current setup, the content server needs to have its content annotated with meta data to let a context matcher find good matches for a certain user profile. The data gathered from the distance sensors could then be used to choose from a number of different visualization options that have been prepared for various viewing distances. That way a commercial shown on the digital screen could even smoothly transform from a far distance still image to a close distance video visualization while a user would come closer to the display. Please note, that such a functionality is currently not implemented but an option for further development tasks.

5 Conclusion

In this paper we have proposed a tokenized interaction architecture and presented the early stage of a ZigBee based implementation thereof. We are aware of the fact that user profiles and their usage could conflict with privacy concerns, however at this stage of development we didn't want to discuss this in detail but just keep in mind that this could be an issue in real deployment scenarios. However, we think that a user token, following the user on its way through urban areas, could become some sort of a bonus card, totally accepted by users: a user would collect bonus points by passing by displays that can read out the personal profile. When buying a product in a store supporting this tokenized interaction platform, the user would receive a discount e.g. linked to the number of contacts with displays.

We are currently developing both the user token hardware/firmware and the residential server software stack further to have a running prototype with the functions described in this paper within the next few months. Since digital signage products are already out in the field that solve the distribution of content to the screens, the tokenized interaction architecture also provides an interface for applications outside the ecosystem: a content provider could query the system through one of the core servers to get contextual information and use the response for deciding on the kind of content to display. That way, existing digital signage infrastructure could be used to display advertisements with respect to current and maybe even future situations.

6 Acknowledgments

This work is supported under the FFG Research Studios Austria program under grant agreement No. 818652 DISPLAYS (Pervasive Display Systems) and under grant FACT (Flexible Autonomic Context Technologies) of Siemens AG, CT-SE 2, Munich.

References

[BCA+08] V. Barbosa, A. Carreras, H.K. Arachchi, S. Dogan, M.T. Andrade, J. Delgado, and A.M. Kondoz. A scalable platform for context-aware and DRM-enabled adaptation of multimedia content. *Submitted to ICT Mobile and Wireless Communications Summit*, 2008.

[BDO+04] J.F. Bradley, B.R. Duffy, G.M.P. OHare, A.N. Martin, and B. Schön. Virtual personal assistants in a pervasive computing world. In *Proceedings of IEEE Systems, Man and Cybernetics, UK-RI 3rd Workshop on Intelligent Cybernetic Systems, ICS04*, pages 7–8, 2004.

[BFS+08] D. Byrne, J. Freyne, B. Smyth, A.F. Smeaton, and G.J.F. Jones. Classifying public display systems: an input/output channel perspective. In *Workshop Proceedings of Designing and Evaluating Mobile Phone-Based Interaction with Public Displays (co-located with CHI 2008 Conference on Human Factors in Computing Systems)*, April 2008.

[BK01] John Barton and Tim Kindberg. The Cooltown User Experience. Technical Report HPL-2001-22, Internet and Mobile Systems Laboratory, HP Laboratories Palo Alto, February 2001.

[BL97] P. Baudisch and D. Leopold. User-configurable advertising profiles applied to Web page banners. In *Proceedings of the first Berlin Economics Workshop*, pages 24–25, 1997.

[BLSC01] T. Bozios, G. Lekakos, V. Skoularidou, and K. Chorianopoulos. Advanced techniques for personalized advertising in a digital TV environment: The imedia system. In *Proceedings of the eBusiness and eWork Conference*, 2001.

[BPS+08] Timothy W. Bickmore, Laura Pfeifer, Daniel Schulman, Sepalika Perera, Chaamari Senanayake, and Ishraque Nazmi. Public displays of affect: deploying relational agents in public spaces. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3297–3302, New York, NY, USA, 2008. ACM.

[CFJB05] H. Chen, T. Finin, A. Joshi, and MARYLAND UNIV BALTIMORE. *Semantic web in a pervasive context-aware architecture*. Defense Technical Information Center, 2005.

[DAS01] Anind K. Dey, Gregory D. Abowd, and Daniel Salber. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-Computer Interaction*, 16(2):97–166, December 2001.

[Gra02] E. Gratton. M-commerce: The Notion of Consumer Consent in Receiving Location-Based Advertising. *Canadian Journal of Law and Technology*, 1(2):59–77, 2002.

[MAMI07] A.H. Morales-Aranda and O. Mayora-Ibarra. A Context Sensitive Public Display for Adaptive Multi-User Information Visualization. In *Proceedings of the Third International Conference on Autonomic and Autonomous Systems*. IEEE Computer Society Washington, DC, USA, 2007.

[MKHS08] Hema Mahato, Dagmar Kern, Paul Holleis, and Albrecht Schmidt. Implicit personalization of public environments using bluetooth. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3093–3098, New York, NY, USA, 2008. ACM.

[RC02] Anand Ranganathan and Roy H. Campbell. Advertising in a pervasive computing environment. In *WMC '02: Proceedings of the 2nd international workshop on Mobile commerce*, pages 10–14, New York, NY, USA, 2002. ACM.

[RDG+08] Jan S. Rellermeyer, Michael Duller, Ken Gilmer, Damianos Maragkos, Dimitrios Papageorgiou, and Gustavo Alonso. The Software Fabric for the Internet of Things. In C. Floerkemeier, M. Langheinrich, E. Fleisch, F. Mattern, and S.E. Sarma, editors, *Proceedings of the International Conference on the Internet of Things, Zurich, Switzerland*, volume 4952 of *Lecture Notes in Computer Science*, March 2008.

[SBHM02] P. Sutton, M. Brereton, C. Heyer, and I. MacColl. Ambient Interaction Framework–Software Infrastructure for the Rapid Development of Pervasive Computing Environments. In *Inaugural Asia Pacific Forum on Pervasive Computing*, page 1. Australian Computer Society, 2002.

[SK05] Christian Sandor and Gudrun Klinker. A rapid prototyping software infrastructure for user interfaces in ubiquitous augmented reality. *Personal Ubiquitous Computing*, 9(3):169–185, 2005.

[SPD06] Matthew Sharifi, Terry Payne, and Esther David. Public Display Advertising Based on Bluetooth Device Presence. In *Mobile Interaction with the Real World (MIRW 2006) in conjunction with the 8th International Conference on Human Computer Interaction with Mobile Devices and Services*, September 2006.

[SRA09] Stabeler M. Shannon R. and Quigley A. Profiling and targeting opportunities in pervasive advertising. In *Pervasive 2009 workshop on Pervasive Advertising*, Nara, Japan, 2009.

[VSdI06] J.I. Vazquez, I. Sedano, and D.L. de Ipina. Evaluation of orchestrated reactivity of smart objects in pervasive semantic web scenarios. In *Proceedings of The Second International Workshop on Semantic Web Technology For Ubiquitous and Mobile Applications (SWUMA06) at the 17th European Conference of Artificial Inteligence (ECAI 2006)*, 2006.

Presence Sensing Billboards

Bernhard Wally, Alois Ferscha, Markus Lenger

bernhard.wally@researchstudio.at, {ferscha, markus.lenger}@pervasive.jku.at

Abstract: The Out-of-Home (OOH) media markets currently rely on a plethora of methods and technologies for measuring their potential or real success. From nationwide interviews over statistical analyses to device based mobility records, each of the OOH markets (in different regions or countries) has chosen its own mixture of measurement efforts in order to provide more or less clear numbers to their clients. In this contribution, we are presenting the *Presence Scanner*, a new and inexpensive technological approach to detecting the presence of people in front of OOH media sites. Based on an array of ultra sonic range finders and equipped with wireless communication capabilities, the Presence Scanner is on its way to resemble a flexible and scalable solution to plug and play activity sensing in (at least) indoor scenarios.

1 Towards Aware Billboards

Out-of-Home (OOH) and ambient media are still representing a substantial part of our daily advertising dosage, with about 10% and 8% of gross contacts, respectively, as of 2007. For instance, City-Light-Posters enjoy increasing popularity within the young and mobile audience. The attentiveness towards OOH media is lower than for most other media, except for radio, but is still rated around 2.8 (on a scale from 1 (low attentiveness) to 6 (high attentiveness) [DFIS08].

However, the OOH market and especially its main shareholders—OOH advertising companies such as EPAMEDIA and JCDecaux—are facing rough times, as they are still struggling with the definition and measurement of *contacts*. Contacts however (or better: the *opportunity to see*, OTS [Kos09a]), are the currency of advertisement and in times of the Internet where views and clicks of advertisement banners are rather easy to count, the measurement methods used in classical advertisement markets such as TV, Radio and OOH have to be reconsidered. The OTS is closely related to the "frequency" of a specific OOH media site: how many OTS/passers-by are there in a certain amount of time (e.g. per hour)?

Pedestrians, cars and public transport users are all among the group of individuals to be considered in OOH media situations. While cars are usually tracked by official authorities and the usage of public transport is known by the public transport companies, pedestrians are harder to grasp (in numbers), as they roam around rather freely. Statistical analyses from data gained in interviews or questionnaires can serve as a powerful and accurate source for media planning, however "real" data, measured at the place of occurrence has a clear potential for higher accuracy and it can even serve for realtime decisions if required.

In this paper we will thus first give an overview on currently used contact definitions for OOH media, motivate our approach for non-cooperative people counting, present different technologies for presence detection in public spaces and provide a detailed view on the Presence Scanner (a cheap and easy to use presence detection device suitable for outdoor usage) and its system architecture. We will also present two application scenarios of the Presence Scanner for the OOH market that are currently under development.

2 Frequency Counting

International OOH research uses mainly three methods for calculating individual related billboard contacts [Sch05]: questionnaires, observations and the usage of existing traffic counting data. Since questionnaires tend to be rather inaccurate, more and more observation based methods are implemented and used for the derivation of contacts [Sch05].

2.1 Plakatwertung Österreich

The Plakatwertung Österreich (PWÖ) is the Austrian implementation of a billboard contacts definition, introduced to the market in 2000 and revised in 2004. It consists of the following criteria [Kos09b]: (1) *site* different sites must be at least 25m apart; multiple billboards are considered a part of the same site if closer than that, (2) *visibility and distance* a billboard situated in parallel to the flow of traffic can be viewed from up to 80 meters distance in an angle of at least 30°, (3) *road distance* the distance is measured from the middle of the road, (4) *billboard direction* with respect to the viewer, from 0° to 90°, in 15° steps, (5) *frequency* based on statistics of the Kuratorium für Verkehrssicherheit (KfV); for roads without available statistics, additional measurements have been conducted, (6) *velocity* based on data of the KfV, including traffic light slowdown and (7) *density* the more billboards on a single site, the higher the devaluation of the site.

The PWÖ classifies the 61,563 billboards in Austria into six classes (A-F; Superstar, Star Plus, Star, Top, Selekt, Standard), based on the Plakatwert (PW) calculated from the criteria above. As of the end of 2008, the EPAMEDIA owns 51.9% of all registered billboards, the Gewista 31,6% and the remainder OOH companies sum up to 16,5% [EPA09].

2.2 Buitenreclame

The Buitenreclame (Netherlands) is based on the national traffic counting (supported by annually 120,000 respondents [Kos09a]), which allows finding out start and end points of the registered routes. Using questionnaires and calculations it is possible to infer the actual paths and thus the contact chance [Sch05]. In order to get better contact numbers, this model is extended with a rating for each billboard site, regarding properties such as the distance to the road and the angle of vision [Kos09a].

2.3 ma Plakat

The ma Plakat is the method of choice in Germany and replaced the Plakat-Media-Analyse (PMA) in 2004 [Arbb]. In 2007 the ma Plakat comprised a computer assisted telephone interview (CATI) part (ma 2007 Plakat CATI) with 21,000 samples and a GPS-based mobility study (ma 2007 Plakat GPS) with 8,600 samples conducted in 24 cities [Arba, HK08]. The CATI part was realized by five institutes in 2007 and cost about EUR 77.- to EUR 84.- per interview in a first phase of the project [Rit07]. The GPS part was realized by test persons who carried a GPS receiver for a period of seven days which stored GPS coordinates every second [HK08].

After injecting (1) *CATI and GPS data* into the required model, additional information is added [HK08]: (2) *weighted data set* demographic information, such as the Mikrozensus, etc. (provides a basis for sub sequential target group evaluations), (3) *frequency atlas* road-based traffic flows/frequencies (allows the calculation of passage values for each individual site), (4) *polygon system* redistributing the empirically recorded mobility data from the 2007 ma Outdoor CATI (form finely designed mobility spaces for the entire area of Germany) and (5) *k-value* site-specific visibility criteria (different types of advertising media and their individual sites to be regarded in a differentiated manner). Finally, *exposure and passage probabilities* are calculated using a statistical method, the event analysis procedure (Kaplan-Meier method), resulting in the visibility-weighted passage exposure per poster site, which is the reporting unit of the ma Plakat.

2.4 Poster Audience Research

The Poster Audience Research (Postar) is the OOH measurement method developed and applied in the UK since 1996. It builds upon the older Outdoor Site Classification and Audience Research (OSCAR) model from 1981 [Kos09c] and breaks up the measurement into four distinct research projects [Posa]: (1) *vehicular & pedestrian traffic counts* vehicle counts for approximately 25,000 locations and data for pedestrians are put through a neural network program to provide up-to-date traffic estimates [Posc], (2) *visibility study* realized using infrared eye tracking hardware on users that don't know they are part of a study concerning posters [Pose], (3) *travel survey* individuals are interviewed using a computer aided personal interview (CAPI) technique [Posd][1] and (4) *panel audit* each poster site is analyzed (photographs and up to 254 separate measurements are taken) to determine its particular characteristics [Posb].

[1]The overall opinion of the project partners is that GPS could be a good technology for travel data collection, but some problems need to be addressed: drops in the satellite fix and data conversion. Another idea would be not to rely solely on GPS but use other recording methods as a secondary source of information [Bea07]. Also, it was harder (read: impossible) to find respondents for the GPS based survey, even though the incentive was increased from the usual GBP 20.- to 40.-, so a panel was used to find respondents [Kal07].

2.5 Swiss Poster Research Plus

The Swiss Poster Research Plus AG (SPR+) uses GPS traces of people to calculate billboard contact opportunities. The following steps are taken in the process from raw data to a media planning software for potential customers: GPS data collection → paths and tracks → data validation (GIS data) → individualization (site specific data: buildings, visibility) → passages and contacts → modeling and extrapolation (site specific data, mobility data, infrastructure data) → planning tool → customers [Hof08]. The final OTS is calculated by taking some site specific factors into account. In order to get reliable data for e.g. the Zürich area, 1,800 measurement weeks are required.

2.6 Mediawatch

The Mediawatch is a Single-Source-Multimedia-Measurement-Device in the form of a digital wrist watch, currently in its fifth generation [Hac08]. It can be used to measure radio, TV, print and OOH consumption and was developed in cooperation with over 16 companies, institutions and four universities and federal institutes of technology [GfK09]. The functional principle is very simple: an embedded microphones is activated up to six times a minute for four seconds and records audio, which sums up to approx. 12,000 Bytes. Using a data reduction algorithm, the captured audio is converted to a 100 Bytes audio footprint that is stored on the watch. In the data reduction step it is also guaranteed that the original audio sample cannot be restored, which is important regarding data protection and privacy aspects [GfK09]. The maximum operating time is four weeks [Hac08]. In order to infer media contacts, the recorded footprints are matched against a central data server that collected continuous footprints of all relevant media [GfK09, Hac08]. The Mediawatch can even receive radio frequency signals thus it can be used for measuring OOH media OTSs as well as cinema and shopping mall visits [Hac08]. However, this requires the billboards or other points of interests to be equipped with small transmitters, that need to be powered [Sch05].

3 Presence Detection in Public Urban Areas

Using interviews to generate traffic information about pedestrians is a rather tedious and costly task if conducted on a regular base and with a large sample. Technologies such as the Mediawatch or the SPR+ GPS tracking method provide a cheaper (in the long run) and more reliable (in interviews there is no guarantee that the answers given by the test persons are correct) way of getting the required mobility information. However, it is required that the users carry some kind of technology (a "user token") with them, leading to two issues: (1) only parts of the citizens are considered and (2) participants have to be found. In this classification, also Bluetooth MAC address sniffing falls into the category of user token based approaches, it is therefore not considered any further (see below).

We believe that such user token based techniques are one answer to gather information about billboard passers-by (and many additional scenarios are possible by relying on these user tokens); still, techniques that don't require the passers-by to actively participate in the counting process unleash another potential: that of the big masses. This "non-cooperative" approach is not not only able to respect privacy issues (as long as it's not based on camera systems), because passers-by don't have a known ID in the system, but could even be easier to setup and maintain if the right usage scenarios and technology decisions are made: only selected OOH media sites need to be equipped with some kind of technology instead of lots of probands. Additionally, the user-token-less counting sensors could be an enabling technology for device free interactive installations (be it implicit or explicit). This leads to another advantage of non-cooperative solutions: realtime capability. Statistical approaches (such as user token based approaches, traffic counting, etc.) can only give an estimate of the current situation with a certain probability, while devices mounted on-site have a chance to provide "real" realtime data to be used either locally (implicit advertisement adaption, user interaction) or more globally for realtime media planning decisions, as well as realtime billing.

Based on the different methods described in Sec. 2, we believe that such detailed data could even help making better media planning decisions now, but it will be a necessary part of tomorrows OOH advertisements. Currently, static OOH media is exchanged every 10 to 14 days, degrading detailed frequency information to non-significant supplemental information. But with the advent of digital signage products, OOH media start being able to react on situations in the blink of an eye. Detailed data, such as activity and presence patterns in front of billboards and screens, could be the key information for delivering messages to the target groups with much better efficiency.

A number of technologies exists for detecting the presence of people in a certain area, ranging from very cheap to extraordinarily expensive, from extremely small to huge. The following list gives an overview of different technologies and their main features. Thereafter an informative ranking of the different technologies with respect to the purpose of detecting and counting people in front of public displays is given.

Laser Ranger Finders: Laser range finders provide extremely detailed distance measurement resolution ranging from sub-inch to several hundred meters distance. Miniaturized sensors fit into handheld devices and can measure distances up to approx. 100m with a power consumption of 1mW (for the laser module) [Min, TEM]. In public settings, laser range finders might not be a good choice because their usage of laser light could lead to eye injuries, thus lowering citizen acceptance. Multiple sensors need to be mounted in order to gather detailed information about the movement of people, except when using scanning laser range finders that can observe a certain field of view. Such sensors are unfortunately more bulky and more expensive than the fixed directional sensors, with dimensions around 80x80x80mm^3 [Acr].

Optical Range Finders: Optical range finders emit light that is than detected using a photo sensor; based on the time-of-flight for each pixel of the sensor a 2-dimensional depth image can be generated. The quality of this kind of sensor is very high (e.g. featuring a resolution of 176x144 pixels and a horizontal field of view of 43.6°), the package size is around 65x65x68mm^3. Sensors are usually designed for indoor usage and support a scanning distance of up to 5m [MES09].

Passive Infrared: Passive infrared (PIR) sensors can be used to detect the presence of people nearby at a rather cheap price with a small footprint. The detection range for miniaturized modules can be as far as 10m, with a field of view of 360°[IR-06]. Motion detectors are binary sensors, i.e. they can say, that there is a human, but they cannot say where and in what distance he/she is.

Radar: Radar allows detecting the distance and movement speed of objects up to several kilometers away. In the short range sector, improvements with regards to the size of radar modules have been accomplished, enabling modules in the size of 60x45x20mm^3 detecting objects up to 30m distance [KGH$^+$05].

Ultra Sonic Range Finders: Ultra sonic range finders work similar to radar range finders, but make use of acoustic waves instead of electromagnetic waves. Due to the difference in speed (speed of light vs. speed of sound) no highly precise electronic components need to be deployed, resulting in cheaper circuits. Ultra sonic sensors can cover distances up to 18m [Fut].

Computer Vision: Camera based presence detection and counting systems provide very flexible and accurate readings at the cost of high data rates and consequently a larger amount of computing power. For the purpose of public space observation, cameras face legitimately harsh opposition from privacy advocates, which is the main reason why we ruled out this technology for our purposes. Micro cameras can be as small as 30x30x30mm^3 while providing high quality images at frame rates as high as 80fps.

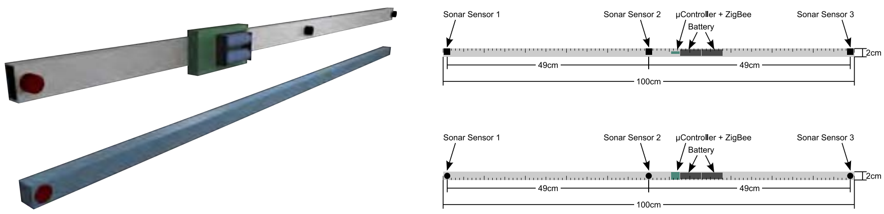
Photo Electric Sensors: Photo electric relays ("light barriers") are a very cheap, reliable and technically robust solution, but they require a slightly more complex setup: light source and detector or reflector need to be aligned correctly, if there is a spot to mount them at all. In outdoor/urban scenarios there is very often no "opposite side" or if there is one it will most probably be owned by a third party, complicating the setup (and maintenance) process. Miniaturized light barriers can cover distances up to 14m [Dis09a, Dis09b].

For our purpose of building a cheap, small and outdoor capable sensor system for presence detection and people counting we ruled out camera based systems for reasons of user acceptance and privacy concerns, even though good progress has been made in detecting and counting people using camera based systems [HHD99, SGMR01, YnG03, RF03, SBGPO08]. With regards to the material cost radar and optical range finders were not a choice. Laser range finders were excluded because of possible eye injuries with laser light (especially in public/crowded areas). Light barriers are not practicable in terms of installation, as we envisioned a system that is capable of true plug and play, but light barriers need the installation of a reflector or the detector opposite to the light source—this introduces additional complexity and might prove impossible in many outdoor scenarios.

That left PIR and ultra sonic range finders in the shortlist. Since pricing and module size is about the same for both technologies we decided to use ultra sonic, because it provides more detailed information (distance of objects) allowing finer grained application response as well as better people counting capabilities through path estimation and plausibility checks (e.g. walking speeds).

4 Presence Scanner

With the Presence Scanner we started implementing a self-contained, scalable, wireless presence detection device, prepared for plug and play usage. One Presence Scanner comprises three ultra sonic range finders, arranged in line with a distance of 49cm from the outer sensors to the center one, resulting in a bar with 1m length. The first prototype, realized based on a cable duct, has a cross section of 16x30mm^2 and a crude ZigBee enabled radio platform of another 100x50x30mm^3 attached to it (cf. Fig. 1a.top). We are currently planning a second generation made of a custom designed plastic housing (1000x20x20mm^3) incorporating all required electronics, running on two exchangeable AA batteries (cf. Fig. 1a.bottom and Fig. 1b). The sensor beam pattern of the Presence Scanner is depicted in Fig. 2a, the raw distance measurements over a period of three minutes are shown in Fig. 2b.

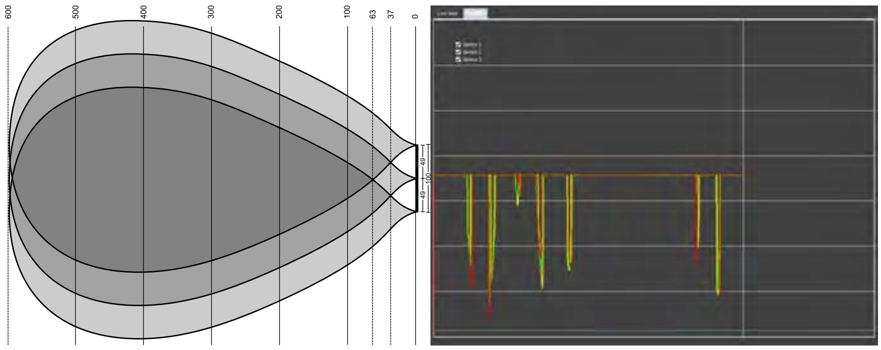


(a) (b)

Figure 1: (a) an early prototype (top) and the envisioned Presence Scanner design (bottom). Each of them is 1m wide and features three ultra sonic sensors with 49cm distance to each other as well as an integrated microcontroller and wireless communication chip. (b) detailed view on the envisioned indoor design (including a metering rule texture for "analogue" usage).

The ultra sonic range finders in use are MaxBotix LV-EZ1. In order to avoid interferences between the three senors, they are enabled sequentially. The built-in support for daisy chaining these sensors didn't work flawlessly, so the daisy chaining was realized in software on the microcontroller. The delay between the activation of each of the sensors depends on the application scenario: the time increases with larger distances to measure (because ultra sonic range finders are based on time of flight measurements) and with the amount of resonance of the probing audio signal, i.e. the properties of the acoustic environment.

We empirically identified the following values as a robust setting for indoor usage (opposite wall distance >8m, consisting of glass; glossy tiled floor with lots of reflections): (1) the delay between sensor activation was set to 160ms, (2) after all three sensors we added



(a) (b)

Figure 2: (a) the approximate sensor beam pattern of the Presence Scanner (adapted from [Max07], numbers in cm) and (b) raw sample measurements of the three sensors of the Presence Scanner over a period of three minutes in an indoor setting (maximum distance 358cm). The distance between two white lines equals 100cm and reaches from 0cm (bottom) to 700cm (top). The seven clearly visible peaks depict bypassing from different directions in different angles and distances.

another 20ms delay to make one cycle take 500ms for reading all three ultra sonic sensors sequentially (cf. Fig. 3.bottom). In another indoor setting, with the opposite wall being closer than 2m, the inter-sensor delay was set to 50ms and no “final calm down” delay was required for getting stable readings from the Presence Scanner (cf. Fig. 3.top).

4.1 Modularity and Scalability

Our decision to go wireless with the Presence Scanner was partly driven by improving the scalability of the system. Multiple Presence Scanners could be mounted close to each other, extending the observation area or taking into account different observation angles or heights. The main issue in deploying multiple Presence Scanners in close distance is the potential interferences with other Presence Scanners, thus the ultra sonic sensor activation needs to be synchronized. The simplest solution to this issue is activating each Presence Scanner sequentially, as it is done within a single Presence Scanner (cf. Fig. 4). This daisy chaining of the Presence Scanners results in a rather poor performance, with a sampling rate of 0.69Hz when using three Presence Scanners. Higher sampling rates can be achieved by allowing distant sensors to be activated simultaneously, either on a Presence Scanner granularity (less performance) or on a single sensor granularity (better performance).

But how to decide if two sensors are far enough from each other? The main factor is the maximum distance to be measured by the sensors: two sensors must be placed in a distance of at least double the maximum expected distance, because in this case the audio signal of one sensor will arrive at the other sensor later than the signal sent out by the other sensor

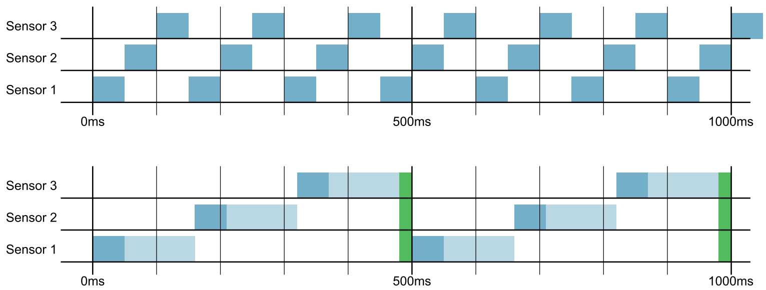


Figure 3: Timing issues in the Presence Scanner: (top) indoor setting with low maximum distance (<2m), (bottom) indoor setting for echoic environments. The darker blue resembles the minimum delay (for re-activating the same sensor again) as of the sensor specification, the lighter blue depicts the safety delay before activating a new sensor, while the green bar represents the artificial “synchronization” delay at the end of one cycle (to add up to 500ms total cycle time).

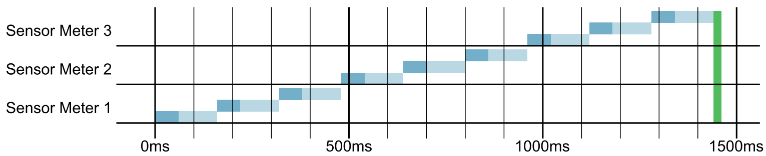


Figure 4: Daisy chaining multiple Presence Scanners to avoid interferences. After one complete cycle 20ms of calm down delay are inserted. The delay between activation of each sensor is taken from the indoor scenario with the higher amount of echoes (cf. Fig. 3.bottom).

itself. In the scenario depicted in Fig. 5 the maximum distance to be measured is 2.4m resulting in a minimum distance of 4.8m between two sensors to avoid erroneous readings. Two consecutive readings of different sensors are allowed every 80ms (this value should be sufficient to let the echoes pass off). The result of this setup in which 2 pairs of sensors are far enough from each other (out of a total of 8 Presence Scanners, corresponding to 24 sensors) is a very poor (and almost neglectable) sampling rate improvement from 0.52Hz to 0.57Hz per sensor.

4.2 Counting Passers-by

Two counting algorithms have been implemented so far, both of them being a work in progress: (1) detect on leave and (2) live evaluation. Currently, we cannot deliver any detailed data on the accuracy of the algorithms, as we are still exploring these and other ways to reliably count passers-by.

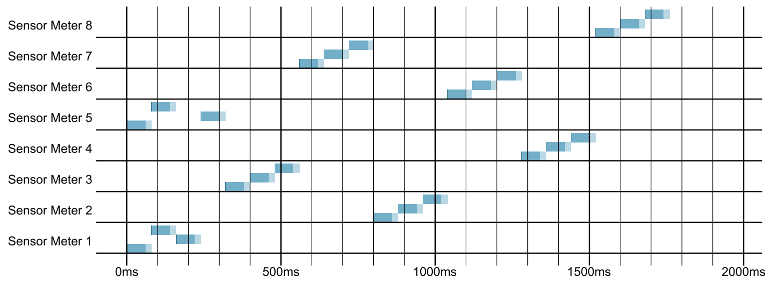


Figure 5: In this setting, sensors 1 and 3 of Presence Scanners 1 and 5 are far enough from each other in order to be operated simultaneously. All other sensors need to be operated sequentially. This corresponds to a sampling rate of 0.57Hz compared to 0.52Hz in pure daisy chaining mode.

4.2.1 Detect on Leave

Useful for detecting people passing by the Presence Scanner in parallel. The algorithm is triggered on a rising edge of one of the outer sensors—such an event can be interpreted as “somebody just left the sensing area of a sensor”. In that case, the history of the other two sensors is analyzed to find out whether a person could just have walked by them, sequentially. Assuming a walking speed of 0.5-1.7m/s (≙1.8-6.12km/h) and regarding the distance of the sensors of 49cm, the history of the last 0.29-0.98s is taken into account for the center sensor and the last 0.58-1.96s for the other outer sensor. If there is a matching measurement (regarding both the time it was read and the distance measured) on both sensors, the people counter is increased by one. Additionally, the direction the person was walking in (left or right) is known. By evaluating the distances measured by the three sensors an estimation of the angle in which a person was passing by can be inferred. This algorithm is a single user algorithm with no support for multiple people passing by or stopping in front of the Presence Scanner.

4.2.2 Live Evaluation

The live evaluation algorithm is a more sophisticated approach to the detection and counting of passers-by. For each sensor reading, a comparison to the state at the last reading is conducted (considering the readings of all three sensors of a Presence Scanner). That way a virtual passer-by (VPB) is created for each sensor reading that couldn’t be a follow-up of another VPB (regarding the same constraints like in the previously discussed approach). Existing VPBs are repositioned using the sensor readings and a possible path and speed is calculated. This approach is much more flexible, especially when taking into account several Presence Scanners and combining their readings for VPB creation and tracing. Unfortunately, the current state of implementation is far from delivering reliable results, but we believe that it has enough potential for good results and even—to some extent—multi user support.

5 Conclusion

We presented a new and highly scalable plug and play device for outdoor people counting, the Presence Scanner, that can be mounted on any billboard and provides the number of people passing by as well as real time interaction possibilities (see Sec. 5.2). Currently, a storage and processing unit needs to be attached to the Presence Scanner(s), as the on-chip memory is too low for long-term storage. We have yet to improve the people counting algorithms and inter-device communication, synchronization and negotiation, but first steps have been taken. The solution so far is low cost (less than EUR 200.- for one Presence Scanner when relying on the currently used components) and can be made cheaper when changing to other components or producing in higher volumes. With the Presence Scanner, interactive OOH advertising is gaining another tool for simple installations, especially for short range implicit interaction (cf. Fig. 6). At the same time, the OOH market receives a new tool for evaluating billboard sites that requires no explicit user participation and respects user privacy concerns by sensing distance values only—no sensible data that could raise privacy concerns is sensed or inferred.

We believe our system is suitable for indoor and outdoor usage. Indoor usage could be validated by our lab experiments: as depicted in Fig. 2b, the measured data is relatively stable and allows algorithms to rely on this low noise data. Outdoor usage is just assumed, as we haven't extensively tested the Presence Scanner in outdoor scenarios, yet. There might be problems with the stability of the sensor data due to environmental noise introduced by machines, cars (parking assistant systems), etc.

5.1 Drawbacks

Compared to some of the methods presented in Sec. 2, our approach has considerable drawbacks: (1) it is not able to identify people, i.e. it is impossible to generate accurate traces of people through the city or find out that a person would pass by a certain site every working day at a certain time. (2) due to the technology used and the shape of the Presence Scanner, it can only detect objects up to a distance of 7m. (3) if an algorithm for counting cars would be implemented it would not be possible to count the number of passengers residing in the cars. (4) the same holds true for public transport.

Many more drawbacks of our approach could be found, however when thinking of the Presence Scanner as "one of the tools" that make up the overall view on frequency counting, we believe it could play a considerable role, because it would even enable value added services such as *interactive advertising* to the frequency counting functionality (cf. Fig. 6). Also, the Presence Scanner was not designed to deliver an answer to the questions described in any of the previously stated drawbacks. Our goal was to implement a system that would be able to count pedestrian passers-by. When we decided to use ultra sonic range finders, we knew that this would introduce drawbacks such as restricted distance scanning, occlusions, non-identification, etc. For our purpose, this was acceptable.

5.2 Outlook

We are currently implementing and incrementally enhancing the counting algorithm in order to get reliable data, while at the same time a truly embedded system is designed and a wireless negotiation protocol is implemented. Our goal is an embedded system to be presumably bought in a hardware store to be used in a plug and play manner.

The performance of the counting algorithm will be tested in a field study that is currently in preparation: a sensor bar is mounted on a real outdoor advertising billboard or "City Light" and Presence Scanner counts are compared to manual counts as well as other techniques such as Bluetooth MAC address sniffing.

In another implementation we are envisioning an interactive ad for a City Light, based on a Presence Scanner embedded therein and some of the lighting coupled to the sensors, such that only parts of the ad will be illuminated, based on the sensed distances of the three sensors (cf. Fig. 6). For this installation, a different layout of the sensors (other than lined up in parallel) is intended, with all three sensors placed close to each other and pointing in different directions—the angle between two sensor center axis is 50°, resulting in an observation area of almost 180°. Our intention with this installation is to find out if there is any improvement in recall values in comparison to "normal" ads and to test the user acceptance of this technology. This scenario is based on observations of previously implemented creative city light ads that would light up various areas of the poster using a constant loop. Multiple users would simply light up more areas at the same time, therefore no negative impact on the user experience or presentation of the product is expected.

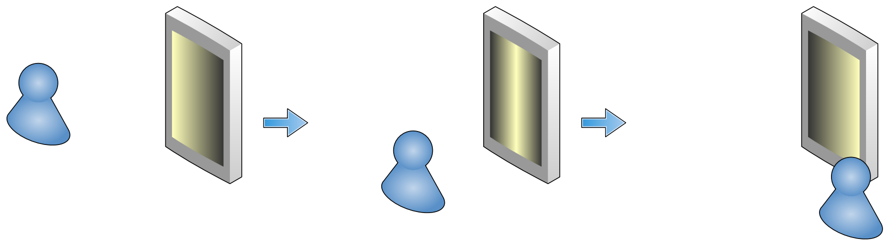


Figure 6: Future development: only parts of the advertisement are illuminated, depending on the location of the viewer. The distance of the viewer could be used to control various light properties.

6 Acknowledgments

This work is supported under the FFG Research Studios Austria program under grant agreement No. 818652 DISPLAYS (Pervasive Display Systems) and under grant FACT (Flexible Autonomic Context Technologies) of Siemens AG, CT-SE 2, Munich.

References

[Acr] Acroname Inc. Hokuyo UHG-08LX Laser. Internet: http://www.acroname.com/robotics/parts/R311-HOKUYO-LASER2.html. Last Visited: April 26, 2009.

[Arba] Arbeitsgemeinschaft Media-Analyse e.V. Erhebungsmethode. Internet: http://agma-mmc.de/03_forschung/plakat/erhebung_methode/erhebungsme thode.asp. Last Visited: April 26, 2009.

[Arbb] Arbeitsgemeinschaft Media-Analyse e.V. Plakat. Internet: http://agma-mmc.de/03_forschung/plakat.asp. Last Visited: April 26, 2009.

[Bea07] Beacon Dodsworth Limited. POSTAR GPS Project. Technical report, Beacon Dodsworth Limited, November 2007.

[DFIS08] Anna Maria Deisenberg, Christian Franzen, Jan Isenbart, and Udo Schendel. *Focus-Jahrbuch 2008*, chapter MindSet - Näher dran, pages 321–332. Focus Magazin Verlag, 2008.

[Dis09a] Distrelec. Lichtschranken und Lichttaster. Internet: https://www.distrelec.at/ishopWebFront/catalog/product.do/para/keywords/is/Licht schranken_und_Lichttaster/and/language/is/de/and/shop/is/AT/and/series/is/1/and/id/is /01/and/node/is/DC-53480.html, April 2009. Last visited: April 24, 2009.

[Dis09b] Distrelec. Lichtschranken und Lichttaster. Internet: https://www.distrelec.at/ishopWebFront/catalog/product.do/para/keywords/is/Licht schranken_und_Lichttaster/and/language/is/de/and/shop/is/AT/and/series/is/1/and/id/is /01/and/node/is/DC-61839.html, April 2009. Last visited: April 24, 2009.

[EPA09] EPAMEDIA. Stellenbewertung PWÖ. Internet, 2009.

[Fut] Futurlec. Ultra Sonic Sensors. Internet: http://www.futurlec.com/Ultrasonic_Sen sors.shtml. Last Visited: April 26, 2009.

[GfK09] GfK Telecontrol AG. Mediawatch. Internet: http://www.telecontrol.ch/typo3/uploads/media/Product_Brochure_Mediawatch.pdf, March 2009. Product Brochure.

[Hac08] Tanja Hackenbruch. *Focus-Jahrbuch 2008*, chapter Elektronische Publikumsforschung mit Radiocontrol/Mediawatch, pages 457–475. Focus Magazin Verlag, 2008.

[HHD99] Ismail Haritaoglu, David Harwood, and Larry S. Davis. Hydra: Multiple People Detection and Tracking Using Silhouettes. In *Proceedings of the International Conference on Image Analysis and Processing*, Los Alamitos, CA, USA, 1999. IEEE Computer Society.

[HK08] Lothar Hannen and Christiane Korch. Proudly announced: The german ma Outdoor has been published. In *Proceedings of the Annual European Media Research Organisations (EMRO) Congress 2008*, May 2008.

[Hof08] Urs Hofmann. Plakatforschung 'SPR+'. Ein Einblick in eine der weltweit größten GPS-Mobilitätserhebungen, May 2008.

[IR-06] IR-TEC International Ltd. *MS-360LP Low Power PIR Motion Sensor Module*, July 2006. Technical Datasheet.

[Kal07] Kaleidoscope Research and Consultancy Limited. EXPERIMENTAL TRAVEL PILOT KALEIDOSCOPE REVIEW FOR POSTAR. Technical report, Kaleidoscope Research and Consultancy Limited, October 2007.

[KGH+05] R. Kulke, C. Günner, S. Holzwarth, J. Kassner, A. Lauer, M. Rittweger, P. Uhlig, and P. Weigand. 24 GHz Radar Sensor integrates Patch Antenna and Frontend Module in single Multilayer LTCC Substrate. In *Proceedings of the EMPC 2005*, June 2005.

[Kos09a] Wolfgang J. Koschnick. Plakatforschung. In Wolfgang J. Koschnick, editor, *FOCUS-Lexikon, Werbeplanung - Mediaplanung, Marktforschung - Kommunikationsforschung - Mediaforschung*. FOCUS Magazin Verlag GmbH, 2009.

[Kos09b] Wolfgang J. Koschnick. Plakatwertung Österreich (PWÖ). In Wolfgang J. Koschnick, editor, *FOCUS-Lexikon, Werbeplanung - Mediaplanung, Marktforschung - Kommunikationsforschung - Mediaforschung*. FOCUS Magazin Verlag GmbH, 2009.

[Kos09c] Wolfgang J. Koschnick. POSTAR. In Wolfgang J. Koschnick, editor, *FOCUS-Lexikon, Werbeplanung - Mediaplanung, Marktforschung - Kommunikationsforschung - Mediaforschung*. FOCUS Magazin Verlag GmbH, 2009.

[Max07] Maxbotix Inc. Ultrasonic Rangefinders Feature Custom Beam Width. Technical report, July 2007. Press Release.

[MES09] MESA Imaging AG. *SR4000 Data Sheet*, February 2009.

[Min] Mindrum Precision, Inc. Technical Specifications. Internet: http://www.mindrum.com/TechSpecsInfo.htm. Last Visited: April 26, 2009.

[Posa] Postar Limited. Research. Internet: http://www.postar.co.uk/research. Last Visited: April 25, 2009.

[Posb] Postar Limited. Research: Panel Audit. Internet: http://www.postar.co.uk/research/classification. Last Visited: April 25, 2009.

[Posc] Postar Limited. Research: Traffic Counts. Internet: http://www.postar.co.uk/research/traffic-counts. Last Visited: April 25, 2009.

[Posd] Postar Limited. Research: Travel Survey. Internet: http://www.postar.co.uk/research/survey. Last Visited: April 25, 2009.

[Pose] Postar Limited. Research: Visibility Study. Internet: http://www.postar.co.uk/research/visibility-study. Last Visited: April 25, 2009.

[RF03] Deva Ramanan and D. A. Forsyth. Finding and Tracking People from the Bottom Up. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 467–474, June 2003.

[Rit07] Gabriele Ritter. Field management systems for Face-to-Face samples. In *Proceedings of the Annual European Media Research Organisations (EMRO) Congress 2007*, May 2007.

[SBGPO08] Kevin Smith, S. Ba, D. Gatica-Perez, and J.M. Odobez. Tracking the Visual Focus of Attention for a Varying Number of Wandering People. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)*, 30(7), July 2008.

[Sch05] Christian Scheier. *Focus-Jahrbuch 2005*, chapter Wie wirken Plakate?, pages 265–289. Focus Magazin Verlag, 2005.

[SGMR01] Claudio Sacchi, Gianluca Gera, Lucio Marcenaro, and Carlo S. Regazzoni. Advanced image-processing tools for counting people in tourist site-monitoring applications. *Signal Processing*, 81(5):1017–1040, 2001.

[TEM] TEM-Technologie. Phas-100 Laser Phaseshift Measurement for large Distances. Internet: http://www.tem-technologie.de/product.php?p=p3&lan=en. Last Visited: April 26, 2009.

[YnG03] Danny B. Yang, Héctor H. González-Ba nos, and Leonidas J. Guibas. Counting People in Crowds with a Real-Time Network of Simple Image Sensors. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 1, Los Alamitos, CA, USA, 2003. IEEE Computer Society.

Modelling Missing Values for Audience Measurement in Outdoor Advertising Using GPS Data

Michael May, Christine Körner, Dirk Hecker,
Martial Pasquier, Urs Hofmann and Felix Mende

{michael.may, christine.koerner, dirk.hecker}@iais.fraunhofer.de,
martial.pasquier@idheap.unil.ch, u.hofmann@spr-plus.ch,
felix.mende@affichage.com

Abstract: GPS technology has made it possible to evaluate the performance of outdoor advertising campaigns in an objective manner. Given the GPS trajectories of a sample of test persons over several days, their passages with arbitrary poster campaigns can be calculated. However, inference is complicated by the early dropout of persons. Other than in most demonstrations of spatial data mining algorithms where the structure of the data sample is usually disregarded, poster performance measures such as reach and gross impressions evolve continuously over time and require non-intermittent observations. In this paper, we investigate the applicability of survival analysis to compensate for missing measurement days. We formalize the task of modeling the visit potential of geographic locations based on trajectory data as our variable of interest results from dispersed events in space-time. We perform experiments on the cities of Zurich and Bern simulating different dropout mechanisms and dropout rates and show the adequacy of the applied method. Our modeling technique is at present part of a business solution for the Swiss outdoor advertising branch and serves as pricing basis for the majority of Swiss poster locations.

1 Introduction

Research in the area of outdoor advertising has undergone a number of developments in the past ten years, many of which have been triggered by technological progress. First, the rapid development and commercialization of GPS technology permits a new form of audience calculation. Second, the advancement in digital processing software for geographical data is necessary to evaluate the obtained data. The innovative usage of new technologies is important to sustain the competitiveness of outdoor advertising in the advertising landscape. For example, in 2007 the Swiss outdoor advertising branch made up about 11% of the total advertisement net sales in Switzerland, generating about 663 million Swiss Franc (about 440 million Euro) [WEM08]. The pricing of poster sites is a critical business task and must be justified by objective performance measures. The two predominant indicators for poster performance are gross impressions and reach. They specify the total number of contacts of a population with a given poster campaign and the percentage of population that passes at least one poster of the campaign within a given period of time, respectively.

The evaluation of poster contacts using GPS data is complicated by the fact that test persons tend to drop out of the study before the end of the surveying period. As a result, the data set decreases with advancing time. Missing measurement days pose a serious problem to our application as reach and gross impressions are defined with respect to a given time span. As opposed to most case studies of spatial data mining algorithms, the temporal structure of measurements cannot simply be disregarded in our application. If the missing data are ignored, i.e. missing measurement days are treated as immobility, gross impressions and reach will clearly be underestimated. Also, the removal of test persons with less than seven measurement days is not an option as this leads to a strong reduction of available test persons. A third option, typically applied in data mining, is to estimate missing values from the distribution of available measurements. However, this approach is not easily applied to mobility data as it implies the reconstruction of individual trajectories for the missing measurement days. We therefore treat missing data explicitly in the modeling step. In this paper we consider the estimation of reach from incomplete trajectory data using methods from survival analysis. Estimating reach is more challenging than estimating gross impressions because reach depends more strongly on continuous measurements. For evaluation we apply a technique from the area of survival analysis, namely Kaplan-Meier. We show that this approach is adequate to model the reach of poster campaigns for a given audience.

This paper is organized as follows. In the next section we outline the setup of the Swiss audience measurement study. Section 3 then gives an overview of related work. We begin with general concepts of missing data and continue with a discussion of related analyses of mobility data. Section 4 provides a general problem statement and formulates the task with respect to outdoor advertisement. Section 5 introduces the applied modeling technique and Section 6 provides experimental results for the conurbations Zurich and Bern. We conclude the paper with a summary and outlook on future work.

2 Swiss Audience Measurement Study

In 2003 the two leading Swiss outdoor advertising companies commissioned a pilot study in the conurbation Winterthur to measure the performance of poster campaigns using GPS technology. A representative sample of persons was selected and equipped with a GPS logger for a period between 7-10 days. Since then, further GPS studies have been conducted, which include the largest metropolitan areas in Switzerland as well as a number of smaller conurbations. In total, the survey includes more than 10.000 participants which form a representative sample for about two thirds of the Swiss population. Figure 1 displays the 13 Swiss conurbations with GPS measurements and the resulting GPS traces.

A second part of the empirical data contains information about poster sites. In total, the study includes about 60,000 sites which cover over 99% of the Swiss out-of-home market. Besides geographic coordinates, a visibility area for each panel is defined which is based on empirical values of other countries (see Figure 2 left). The visibility areas are intersected with a building layer in order to obtain realistic areas from within which each panel can be seen (see Figure 2 middle and right).

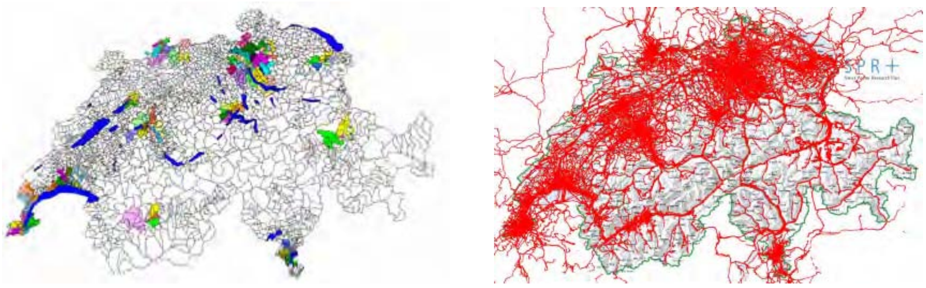


Figure 1: Swiss Conurbations included in the GPS survey (left) and the resulting total mobility measurements (right)

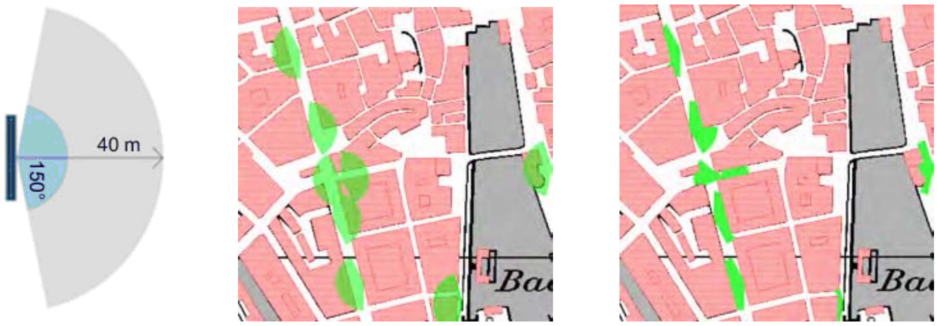


Figure 2: Standardized visibility area of a panel (left), overlay of visibility areas and building layer (middle) and visibility areas after intersection with building layer (right)

Given the trajectories of an individual and a the visibility area of a poster panel, all resulting passages can now be calculated by geographic intersection. However, passing the visibility area of a panel does not imply that a person actually looks at the poster. Depending on passage angle, speed, time of day (only some posters are illuminated at night) and the number of panels at the location (many panels increase the distraction), each passage is weighted. A thus qualified passage constitutes a poster contact, which serves as basis to evaluate reach and gross impressions of poster campaigns. Further details and an example calculation of poster contacts from poster passages can be found in [PHM$^+$08].

3 Related Work

3.1 Concepts of Missing Data

The first major works on missing data appeared in the 1970s. Rubin [Rub76] introduced a typology for missing data and discussed their influence on the inference process. In general, three variants of missing data are distinguished: *missing completely at random*

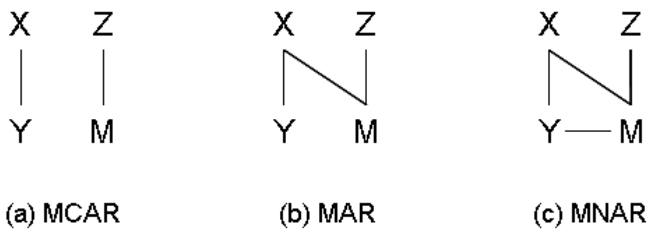


Figure 3: Graphical representation of MCAR, MAR and MNAR (adopted from Schafer and Graham [SG02])

(MCAR), *missing at random* (MAR) and *missing not at random* (MNAR) [LR87, SG02]. We will start with an intuitive explanation of the concepts given a simple data set and then proceed to a more general formulation. Assume a data set with one explanatory variable X and one dependent variable Y for n objects. While X is completely observed, Y may contain missing values. We can encode the missingness of Y within a separate variable M which assumes a value of 1 if Y is observed and 0 if Y is not observed. Further, we define a variable Z of random noise which is unrelated to X and Y. MCAR occurs if the missingness is completely independent of the data, i.e. $P(M|X,Y) = P(M)$ (see Figure 3a). If a relationship between M and X exists but M is still independent of Y, the data are defined to be MAR. MAR denotes a conditional independence of missingness given a fixed value of X (see Figure 3b). Note however, that under MAR a relationship between M and Y may exist due to their mutual dependency on X. This relationship disappears once the value of X is taken into account. Finally, if the distribution of missing values depends on Y, the data are said to be MNAR (see Figure 3c). MCAR and MAR are also referred to as ignorable missingness (or noninformative dropout in longitudinal studies) while MNAR is termed nonignorable missingness (respectively informative dropout).

Usually, data sets contain several variables of which more than one variable may be subject to missingness (e.g. as in longitudinal data). The above illustrated concepts of missing data can then be generalized as follows. Let $Y = (Y_1, Y_2, \ldots, Y_p)$ denote a set of variables with observations for n objects and M denote a $(n \times p)$ matrix encoding (arbitrary) missingness as defined above. Given complete knowledge, we can partition the complete data set $Y_{com} = (Y_{obs}, Y_{mis})$ into subsets containing the values for observed and unobserved parts of the data. The missing values are said to be MCAR if $P(M|Y_{com}) = P(M)$. If the missingness does not depend on the values of unobserved data, i.e. $P(M|Y_{com}) = P(M|Y_{obs})$, the missing values are MAR. Otherwise, if the missingness depends upon the missing value itself, the missing data are MNAR [SG02].

Depending on the type of missingness and the method of inference, estimated parameters of the data may be biased. In general, MCAR results in a correct sampling distribution for Y_{obs} and poses no problem for parameter estimation (except of resulting in reduced sample sizes). Missing values that are MAR produce a correct likelihood distribution, and unbiased parameter estimation conditioned on the observed values is possible. In case of MNAR, parameter estimation is a serious problem and requires an explicit specification of the missingness distribution. However, in many cases the mechanism that leads to missing values in a data set is unknown. The assumption of MAR is then often reasonable but its robustness should be assured.

3.2 Mobility Analysis

To our best knowledge, the application of survival analysis in order to compensate for missing measurement days in GPS surveys has not been described in literature yet. However, survival analysis has been used by Schönfelder and Axhausen [SA01] to analyze rhythmic patterns of travel behavior based on travel-diaries. Within the diaries the test persons noted down a categorical purpose of each trip (activity), means of transportation, destination address, time and duration etc. over a period of six weeks [PFSA00]. The authors then studied the periodicity of leisure and shopping activities by estimating the belonging survival and hazard functions. Hereby, the activities correspond directly to the events of interest and their geographic location is not considered within the analysis. In contrast, we consider the problem to estimate the visit potential for an arbitrary but fixed set of locations. The number and geographic distribution of the locations play an essential role in our application. In this paper we formalize the general task to estimate visit potential of geographic locations and analyze the appropriateness of survival analysis for our application based on a simulation study with the application data.

Fraunhofer IAIS conducts a similar project for the German outdoor advertising media. On behalf of ag.ma, a joint industry committee of German advertising vendors and customers, trajectory data of a nationwide survey are evaluated with methods from survival analysis [Arb09].

4 Problem Formalization

In most general terms we consider the following problem. Given a set of trajectories Tr of a set of persons P and a set of locations L that may be visited by the persons along their tracks. We are interested in the events of 1^{st}, 2^{nd}, ..., k^{th} passage that the persons produce with the location set over time. We seek aggregated values such as the total number of visits or the percentage of persons that visit the locations $1, 2, \ldots, k$ times within a given time span. Both measures can be derived from the distribution of visits with respect to the time axis.

Note, that this definition is independent of the recording technology and data format of the provided trajectory and location data. Trajectories may be given as raw or pre-processed GPS data, as sequence of street segments, sequence of radio cells when using GSM technology or directly as events recorded at various locations using RFID. Also, the definition of a passage or visit is application dependent and can be specified according to needs.

In the context of our outdoor advertising application, the trajectory and location sets are instantiated as follows. The trajectory set consists of pre-processed GPS data, which has been aligned to the street network. Further, small gaps in the trajectories (e.g. caused by tunnels or when passing below a bridge) have been closed by routing. Partitioning the data into daily routes, we can denote the trajectory set as $Tr = \{tr_{ij} \mid i = 1..n,\ j = 1..sd\}$ where tr_{ij} represents the movement of test person p_i $(i = 1..n)$ that has been recorded on day j within the survey duration sd. Note, that a trajectory tr_{ij} may be the empty set if a person has stayed at home during a day.

A location set $L \subseteq \mathcal{L}$ represents a specific poster campaign and is a subset of all existing poster locations $\mathcal{L}$. A single poster location $l \in \mathcal{L}$ is thereby modeled by the reduced visibility area as described in Section 2. A passage occurs if the intersection of a trajectory tr_{ij} and a poster location l is not empty over some time span. The rating of poster locations in outdoor advertising relies on qualified passages. I.e. a poster contact is not automatically generated by *passing* a poster location but by actually *looking* at a poster [SB02]. This condition is implemented by weighting each passage with visibility criteria such as distance and angle of passage. However, for simplicity we shall not consider visibility criteria in this paper as they do not affect the general problem setting. Instead, we concentrate on the *coverage* of a campaign, which is a preliminary state of the poster rating index reach.

Definition 4.1 *The coverage of a poster campaign L in a given target group of persons P over a duration of d days is the percentage of persons that produce at least one passage with the poster campaign within the specified time span d.*

The evaluation of coverage is based on *first* passages, i.e. a single passage of a person through a location of the poster campaign suffices to increase the coverage. Note, that in the above definition the coverage is defined with respect to the test persons P. Given a representative sample of test persons, the extrapolation to the total population is straightforward as the point estimator of coverage in the population is equal to the coverage in the sample.

Given a complete data set, the coverage for a duration d smaller or equal to the surveying period sd can easily be calculated. However, our data shows that only few people produce measurements for the whole surveying period. For example, in Zurich, about two thirds of the test persons drop out of the study early. For the remaining surveying period the movement behavior of these persons is unknown. The application challenge lies in the appropriate treatment of missing location visits that are caused by incomplete mobility data.

Note, that in general several types of missing GPS data may occur: a) short interruptions due to bridges or tunnels, b) single missing trajectories when GPS persons forget to carry their device for parts of a day and c) participant dropout. In this paper we consider only the last case. Case a) is treated within data preprocessing where short gaps are closed by routing. Missing data of type b) cannot be identified from the data itself but requires user interviews at the end of each day. As this information is not available, we assume that the provided measurement days are complete. In consequence, the calculated poster passages possibly underestimate the true poster passages.

5 Modeling Technique

5.1 Data Layout

For a given target audience and campaign our data can be arranged in a table, the rows corresponding to test persons and the columns to observed variables. Some variables $X = (X_1, \dots, X_p)$ are completely observed such as gender or age. Other variables $Y = (Y_1, \dots, Y_{sd})$, corresponding to the aggregated number of location visits per measurement day, are available only in part (see Figure 4). The data possess a monotone pattern of nonresponse as for any test person and any $j \in 2..sd$ the following property with respect to Y holds: if Y_j is missing, then $Y_{j+1}, \dots, Y_{sd}$ are missing as well. A monotone data layout is a special case of a missing data pattern and allows the application of survival analysis with right censoring as explained in detail in the next section.

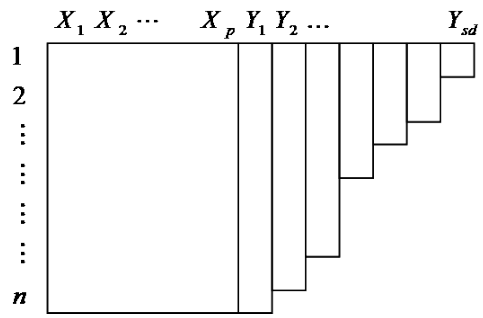


Figure 4: Monotone dropout pattern for GPS test persons

5.2 Survival Analysis

Survival analysis (also: event history analysis) is a branch of statistics that investigates the occurrence of events as they take place over time. More precisely, survival analysis considers the individual time from an initiating event to an event of interest for a group of objects [ABG08, KK05]. Such events denote, for example, the occurrence of some disease in a clinical study or the failure of a device in quality control. One typical method of survival analysis is Kaplan-Meier [KM58], which estimates the probability that some event does not occur (i.e. the object of interest "survives") within a given period of time allowing for dropout behavior. For example, Kaplan-Meier can be used to calculate the life expectancy after a cancer treatment. Naturally, people enter a medical study at different points in time and therefore possess differing lengths of participation. In addition, people can drop out of the study when moving into another city or dying from a different cause. In our application the event of interest denotes the first passage of a person with a given poster campaign. Dropout occurs if the provided mobility data covers a period of less than seven days. In survival analysis, missing measurements are also termed censored data.

More formally, let T denote a random variable that states the survival time of an object, i.e. the time until the occurrence of the event of interest. The function

$$S(t) = P(T > t) \tag{1}$$

is called the survival function and denotes the probability that the specified event occurs later than some time t. For a given data set, Kaplan-Meier analyzes at which times t_i events occur (with $t_0 = 0$) and determines the following variables

- r_i - number of objects at risk at time t_i,
- e_i - number of events at time t_i,
- c_i - number of dropouts between t_{i-1} and t_i.

In our application the objects at risk at t_0 are all persons in the survey, further $e_0 = c_0 = 0$. At each point in time when events occur, the number of objects at risk is reduced by the objects with events as well as by the objects that drop out in the preceding time interval, i.e. $r_{i+1} = r_i - e_i - c_i$. Kaplan-Meier adapts to differing sample sizes by calculating conditional probabilities between two consecutive events. Objects that drop out of the study between two events are assumed to survive until the next event occurs and are then removed. The conditional probability p_i to survive time point t_i given that t_{i-1} has been survived is then calculated as

$$p_i = P(T > t_i | T > t_{i-1}) = \frac{r_{i-1} - e_i}{r_{i-1}}. \tag{2}$$

In the above formula the denominator corresponds to the number of objects in the study before the event at t_i occurs and the numerator subtracts from this number the amount of objects with an event at t_i, i.e. the numerator still contains all objects without an event that drop out between t_{i-1} and t_i. Given the conditional probabilities p_i, the total probability to survive some time point t_k is

$$S(t_k) = P(T > t_k) = \prod_{i=1}^{k} p_i. \tag{3}$$

The transformation from survival probability to poster coverage is straightforward. So far, $S(t)$ states the probability that people in the data sample do not pass any poster location of the campaign until t. Consequently, the coverage of a campaign is given by the probability of the complimentary event

$$F(t) = P(t \leq T) = 1 - S(t). \tag{4}$$

Figure 5 shows the development of coverage for a campaign of 50 randomly selected posters in Zurich over a period of 7 days. The coverage increases on each day on which an event occurs. The dotted lines show 95% confidence intervals for the estimated coverage.

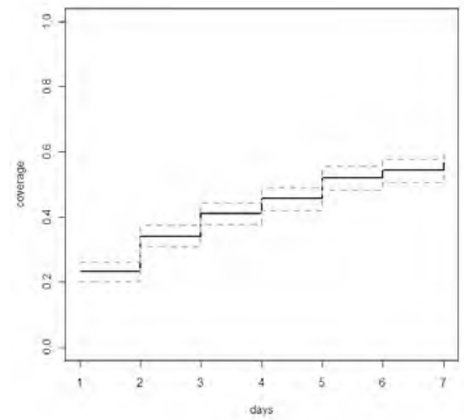


Figure 5: Development of coverage over 7 days

5.3 Properties

Kaplan-Meier is a non-parametric estimator and can adapt to arbitrary event rates over time. However, Kaplan-Meier presupposes that the target variable and the dropout behavior are independent of each other, i.e. the method requires MCAR or MAR missingness. (Note, that in case of MAR Kaplan-Meier should be applied to individual strata in order to perform proper conditioning.)

Is MAR a reasonable assumption for our mobility data? Can we presume that dropout behavior and mobility are independent of each other, that mobile and immobile persons are similarly willing to carry GPS devices? At first sight, this supposition seems reasonable. However, differences may arise considering, for example, different age groups. It is known that the mobile behavior of young and old people differs [Bun04]. In addition, young people are usually more technology enthusiastic than old people, which may keep them longer in a survey. Yet, people in their middle years are usually more reliable than young persons.

Within the pilot study in Winterthur we analyzed the dropout behavior using subgroup analysis. In addition to gender and age group, further sociodemographic attributes were provided. The results showed that the most reliable test persons (i.e. persons with a high number of valid measurement days) belonged primarily to the occupational group of officials and employees. As this occupation is correlated with the age (and thus mobility) of a person, we may expect that Kaplan-Meier shows some bias in the results. However, this effect should be lessoned when conditioning on the attribute age group.

It is therefore important for our application to analyze the effects of different types of missingness on Kaplan-Meier. How robust will the estimated parameters be in case of violated requirements? What degree of missingness may still produce acceptable results? In the next section we try to answer these questions by simulating different dropout mechanisms and dropout rates in the mobility data.

6 Experiments

6.1 Setup

We conducted experiments for the Swiss conurbations Zurich and Bern. Figure 6 shows the distribution of test persons with 1, 2, ... valid measurement days for both conurbations. In order to verify our modeling approach, we used only persons with seven valid measurement days and introduced artificial missingness. This resulted in a total of 807 and 635 test persons in Zurich and Bern, respectively. We simulated different dropout behavior and dropout rates and compared the estimated coverage to the coverage in the entire data set. In order to realize different passage probabilities, we experimented with varying campaign sizes (50 and 100 posters in Zurich and 20 and 50 posters in Bern).

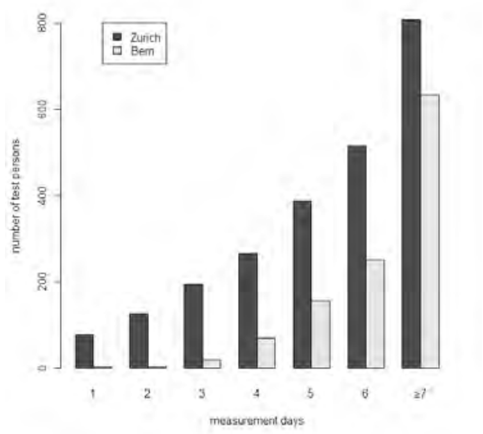


Figure 6: Number of test persons with 1, 2, ... valid measurement days in Zurich and Bern

We implemented three different dropout strategies, simulating MCAR, MAR and MNAR dropout behavior. In general, we first selected a group of dropout persons according to a given dropout rate and then chose per person a random day (day 2 till 7) from which on all trajectories were censored (i.e. removed). For MCAR, the selection of dropout persons took place completely at random. For MAR, we set up different dropout rates for sociodemographic groups. We used the attributes sex (male, female) and age group ($< 30, 30-49$, ≥ 50 years). Within each group, the dropout persons were chosen randomly. Finally, to simulate MNAR we correlated the dropout rate with the total number of passages the persons produced during the surveying period with a given campaign. We simulated positive correlation, censoring preferably persons with many poster passages (mobile persons), and

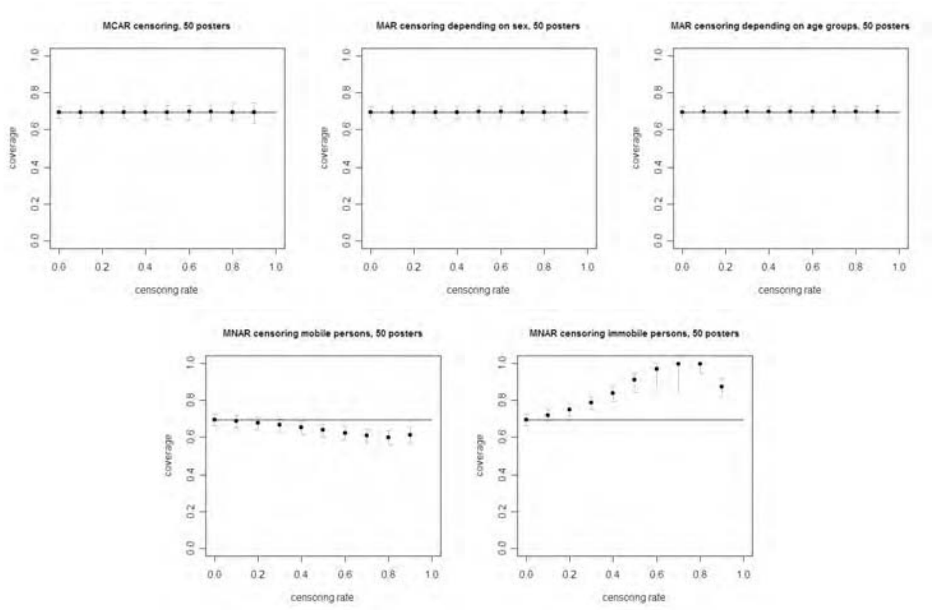


Figure 7: Coverage in Zurich for campaigns of size 50 using the Kaplan-Meier method for MCAR, MAR and MNAR dropout strategies

negative correlation, censoring persons with few passages (immobile persons).

We varied the dropout rate between 0.1 and 0.9 (a rate of 0.0 corresponding to the full data set for comparison). For MAR, an increasing dropout rate was implemented for females and the youngest age group, while the dropout rate for the other groups was held constant at 0.2. Note, that the dropout rate in the tables and figures always refers to the dropout rate of the person group for which we changed the rate in the experiment series, i.e. in case of MAR this rate does not correspond to the dropout rate across all test persons. The results are averages of 100 randomly drawn campaigns of fixed size. For each campaign we performed 10 simulations per dropout behavior to reduce the variability in the results. The calculation of survival rates and confidence intervals was conduced with the statistic software R.

Tables 1-4 contain the estimated coverage for Zurich and Bern. As the results are similar for all four setups, only diagrams for Zurich campaigns of size 50 are shown in Figure 7. The horizontal line denotes the coverage in the full data set. The points denote the mean estimated coverage under different degrees of dropout. The gray lines show 95% confidence intervals calculated based on the cumulative hazard.

6.2 Interpretation

The estimation of coverage for the random dropout strategies MCAR and MAR is unbiased under all tested dropout rates. For MCAR this behavior was expected. For MAR some bias may have been possible due to different mobility within the groups as we conducted all experiments without stratification. In fact, the average coverage between males and females as well as between age groups are slightly different in the data set. For example, in Zurich male test persons produced on average a coverage of 0.712 while female test persons were less mobile and produced a coverage of 0.676 for poster campaigns of size 50. This result implies some degree of robustness of Kaplan-Meier with respect to informative dropout, however this needs to be assessed in further experiments. The diagram for MCAR shows that the confidence intervals increase with advancing dropout rate. This effect is due to the smaller number of test persons that are still at risk at the end of the surveying period. The effect is not visible for MAR because the depicted dropout rate refers only to one gender or age group. For the other groups, dropout was held constant at 0.2, leading to a lower total dropout rate.

For informative censoring the coverage decreases with increasing dropout rate if preferably mobile persons are censored and vice versa for censoring of immobile persons. However, the bias grows more slowly when predominantly mobile persons are censored. This effect is due to the different influence of censoring on mobile and immobile persons. Mobile persons possess a higher chance to produce their first poster passage in the beginning of the surveying period than immobile persons. As the dropout day is chosen randomly once a person has been selected for censoring, the probability that a passage occurs before censoring takes place is higher for mobile persons than for immobile persons. In consequence, mobile persons are less affected by random censoring of days. This behavior is also reflected in the confidence intervals. While the intervals increase only slightly when censoring preferably mobile persons, censoring immobile persons immediately reduces the persons at risk and results in larger confidence intervals. Both experiment series for MNAR show that the bias starts to decrease at a dropout rate around 0.8. This behavior is natural as the increasing dropout rate lessens the structural effect of deliberate censoring. In the extreme case with a censoring rate of 1.0 censoring would take place for each test person.

In summary our experiments confirm the applicability of Kaplan-Meier to mobility data if an uninformative censoring mechanism can be assumed. Further, different censoring rates in distinct sociodemographic groups show no influence on the results even though stratification was not applied. Finally, effects of MNAR censoring depend on the censoring mechanism and seem still acceptable for low censoring rates.

7 Conclusion

In this paper we consider the problem to rate arbitrary poster campaigns in outdoor advertising using GPS mobility data that is affected by dropout behavior. We give a formal problem definition and analyze the applicability of survival analysis in an extensive simulation study on a part of the application data. The simulation shows that the proposed method gives unbiased results under systematic censoring in sociodemographic groups even for high censoring rates. Informative censoring leads as expected to biased estimations which, however, may be acceptable for low censoring rates.

In future work we will explore the more general setting of arbitrary patterns of missingness. So far, we have examined a monotone dropout pattern, where test persons quit the study completely. However, persons may also forget to carry the GPS device for a single day within the survey, resulting in intermittent missing values. A second challenge is the estimation of poster ratings for time spans that are longer than the surveying period. In this case, the estimation procedure has to be combined with an appropriate extrapolation model.

Table 1: Estimated coverage for campaigns with 50 posters in Zurich

dropout rate	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MCAR	0.696	0.696	0.696	0.696	0.696	0.696	0.696	0.696	0.695	0.695
MAR (sex)	0.696	0.695	0.696	0.696	0.695	0.696	0.696	0.694	0.694	0.696
MAR (age)	0.695	0.696	0.696	0.696	0.696	0.696	0.697	0.697	0.696	0.697
MNAR (mob.)	0.696	0.687	0.677	0.666	0.653	0.638	0.622	0.607	0.599	0.613
MNAR (imm.)	0.696	0.720	0.750	0.786	0.839	0.910	0.970	0.998	0.995	0.871

Table 2: Estimated coverage for campaigns with 100 posters in Zurich

dropout rate	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MCAR	0.832	0.832	0.832	0.832	0.832	0.832	0.832	0.832	0.832	0.833
MAR (sex)	0.832	0.833	0.833	0.832	0.832	0.832	0.831	0.831	0.831	0.831
MAR (age)	0.831	0.832	0.833	0.833	0.833	0.833	0.833	0.833	0.832	0.835
MNAR (mob.)	0.832	0.828	0.822	0.816	0.809	0.800	0.791	0.779	0.764	0.756
MNAR (imm.)	0.832	0.859	0.892	0.933	0.973	0.996	0.999	0.984	0.943	0.846

Table 3: Estimated coverage for campaigns with 20 posters in Bern

dropout rate	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MCAR	0.710	0.710	0.710	0.709	0.710	0.710	0.710	0.710	0.710	0.710
MAR (sex)	0.710	0.709	0.709	0.709	0.709	0.711	0.713	0.712	0.713	0.717
MAR (age)	0.710	0.710	0.711	0.709	0.709	0.710	0.709	0.709	0.708	0.710
MNAR (mob.)	0.710	0.702	0.693	0.682	0.671	0.657	0.643	0.629	0.620	0.631
MNAR (imm.)	0.710	0.734	0.764	0.797	0.849	0.905	0.959	0.985	0.977	0.879

Table 4: Estimated coverage for campaigns with 50 posters in Bern

dropout rate	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MCAR	0.874	0.874	0.874	0.874	0.874	0.875	0.875	0.874	0.874	0.873
MAR (sex)	0.874	0.875	0.875	0.874	0.875	0.874	0.873	0.874	0.875	0.875
MAR (age)	0.874	0.875	0.874	0.874	0.874	0.874	0.876	0.875	0.875	0.876
MNAR (mob.)	0.874	0.871	0.868	0.863	0.858	0.853	0.845	0.838	0.825	0.813
MNAR (imm.)	0.874	0.901	0.937	0.969	0.991	0.995	0.989	0.974	0.912	0.867

References

[ABG08] O. O. Aalen, Ø. Borgan, and H. K. Gjessing. *Survival and Event History Analysis*. Statistics for Biology and Health. Springer, 2008.

[Arb09] Arbeitsgemeinschaft Media-Analyse e.V. (ag.ma), 2009. http://www.agma-mmc.de.

[Bun04] Bundesministerium für Verkehr, Bau und Stadtentwicklung. Mobilität in Deutschland 2002, Ergebnisbericht (Mobility in Germany 2002, report on results), 2004. http://www.mobilitaet-in-deutschland.de.

[KK05] D. G. Kleinbaum and M. Klein. *Survival Analysis*. Statistics for Biology and Health. Springer, 2005.

[KM58] E. L. Kaplan and P. Meier. Non-parametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–481, 1958.

[LR87] R. J. A. Little and D.B Rubin. *Statistical Analysis with Missing Data*. Wiley Series in Probability & Mathematical Statistics. John Wiley & Sons, 1987.

[PFSA00] PTV AG, B. Fell, S. Schönfelder, and K.W. Axhausen. Mobidrive questionnaires. Technical report, Institut für Verkehrsplanung, Transporttechnik, Strassen und Eisenbahnbau, ETH Zürich, 2000.

[PHM+08] M. Pasquier, U. Hofmann, F. H. Mende, M. May, D. Hecker, and C. Körner. Modelling and prospects of the audience measurement for outdoor advertising based on data collection using GPS devices (electronic passive measurement system). In *Proceedings of the 8th International Conference on Survey Methods in Transport*, 2008.

[Rub76] D. B. Rubin. Inference and Missing Data. *Biometrika*, 63(3):581–592, 1976.

[SA01] S. Schönfelder and K.W. Axhausen. Analysing the rhythms of travel using survival analysis. In C. Kaspar, C. Laesser, and T. Bieger, editors, *Jahrbuch 2000/2001 Schweizerische Verkehrswirtschaft*, pages 137–162. Universität St. Gallen, 2001.

[SB02] J. Z. Sissors and R. B. Baron. *Advertising Media Planning*, chapter 4-5. McGraw-Hill, 2002.

[SG02] J. L. Schafer and J. W. Graham. Missing Data: Our View on the State of the Art. *Psychological Methods*, 7(2):147–177, 2002.

[WEM08] WEMF AG für Werbemedienforschung. Werbeaufwand Schweiz (Advertising expenditure Switzerland). Press release by Stiftung Werbestatistik Schweiz, 2008. http://www.wemf.ch/de/pdf/Pressetext-d.pdf.

The Perception of Information and Advertisement Screens Mounted in Public Transportation Vehicles - Results from a Mobile Eye-tracking Study

Norman Höller*, Johann Schrammel*, Manfred Tscheligi*, Lucas Paletta±

*CURE
Hauffgasse 3-5
1110 Wien
{hoeller, schrammel, tscheligi}@cure.at

±Joanneum Research
Wastiangasse 6
8010 Graz
lucas.paletta@joanneum.at

Abstract: This research deals with perception of information and advertisement screens mounted in public transport vehicles. We have conducted an exploratory field study with 106 participants. Our main research question was aimed at finding out to what extent people do look at such screens and for how long they fixate them. Further we investigated correlations between content type and the amount of focus time as well as the amount of time a person fixates on such a screen and the number of things one can reproduce freely or recognize shortly after the exposition. We researched whether certain content types can be considered as attention catchers and if certain combinations of content-types have the power to bind persons' focus-time longer than others. Results suggest high awareness of the info-screens among participants but no correlations between fixation time and content respectively fixation time and recall/recognition of content.

1 Introduction

Display Screens have become a very common element of public space recently. A walk through almost any modern urban area is sufficient to confirm this. A glance at the "Digital Signage in Europe: Opportunities for digital out-of-home advertising" [SG07] report confirms this notion by stating as one of it's key findings that digital out of home advertising revenues in Western Europe will quadruple over the next five years. Technological progress, enabling us to build screens that are flatter and lighter than previous ones, paired with increasing affordability of digital displays can be seen as the key to this development.

One application area for digital displays is public transport. In public transport such screens often are used to communicate with passengers, providing passenger information, news, entertainment, and also advertising. As large numbers of people everyday use public transport in urban areas to travel from A to B display screens mounted within these vehicles have the potential to reach a big audience. However, it is unclear whether passengers actually do pay attention to these screens and how much information presented they ultimately remember. Also, questions dealing with optimal content representation for this rather new form of advertising and the identification of specific attention catchers remain unanswered so far. In order to find answers for these questions we conducted a field study using state-of-the-art mobile eye tracking equipment. We hereby had the opportunity to work with a large sample of test subjects: 106 subjects were recruited for the project and conducted a test ride using a tram with mounted displays. Of all test subjects we captured eye tracking data and recorded the user's field of view (scene video). Scene data then was analysed using computer vision algorithms for the detection of the information screens. This way we had the possibility to study the users gaze behaviour in a naturalistic context but still have the ability to process the data of a big sample.

What sets our study apart from many other eye-tracking studies is the context in which it was conducted. Many studies using eye-tracking technology are performed in the lab. That allows researchers a good control of experimental settings and therefore enables the study of very specific aspects of human behaviour under controlled conditions. This approach, as good as it is, is somewhat limiting because we can not expect results being applicable to the real world or being representative for it. Too many factors are being eliminated and the context is highly artificial. In contrast to that approach we decided to go out into the field to study gaze deployment and attention behavior, thereby allowing for all the unforeseeable events and distractions that form and shape our everyday life. Clearly, this approach has its drawbacks as it does not support good control of experimental conditions. On the other side, our experimental design allowed for a 3D representation of the context in which the display screens were presented. In the lab often a 2D representation of the context and the stimulus is used. In order to make the findings of our study as representative and relevant as possible we recruited more than 100 subjects covering different age spans.

2 Related Work

Related work in the domain of out of home media stem from a broad range of disciplines and research techniques. Marketing research, Neuroscience and Psychology are all contributing to the field. Prominent research methods include the use of eye-tracking technology and questionnaires. Shavitt et al. conducted a survey to assess attitude towards different kinds of media channels for advertisement. The authors categorized available media channels regarding self selected (e.g., catalogue) vs. highly intrusive (e.g., tv). Results of the survey suggested that media channels allowing for self selected experiences were favoured over highly intrusive ones [SVL04]. The Digital Out of Home Awareness Study, conducted in 2007 in the USA [OS07], investigated the effectiveness of digital out of home media by focusing on awareness, attention, impact and attitude towards advertising on digital signage and other media. Key findings of the study suggest that digital signage has the "power to stop people" and catch their attention (63% of adults). That is the highest level reported of all media surveyed in the study, including TV, the Internet, billboards, magazines, newspapers, radio, and mobile phone advertising. Digital signage was rated more positively than any other type of media. The awareness with digital signage advertising, according to the study is very high. Sixty-two percent of adults say they have seen ads on digital signage over the past 12 months, and the figure is even higher for young adults between 18 and 24, at 75%.

Eye tracking as a means to measure the movement of the eye is being used for more than 100 years now (it began in ca. 1879) [Da02]. The reason why deployment of human gaze for a long time has been studied almost exclusively limited to 2-D picture viewing is because only in the last few years systems have become compact enough for use outside of labs.

The Outdoor advertising association of America (OOOA) conducted several research studies on the topic. One two year study, carried out over the period of 1999 to 2000 in Los Angeles, New York and Minneapolis revealed that 70% of outdoor posters in the visual field of subject were seen. Of these 63% were likely to be read. The authors of the study made a visibility vs. recall study and the results indicated that visibility was about 3 times higher than aided recall levels (26%) would indicate. Additionally subjects claimed to recall boards and brands that were never posted [OO00].

Maughan et al. have investigated – also using eye-tracking - if people in a street scene (the scene was presented as a virtual image on a computer screen) would notice bus shelter advertisements. Their findings suggested that although posters did get attention they were not the most salient features in the street scenes: only in 11.1 % of cases participants did make their first fixation on the poster. The authors conclude that the emotional appeal of a poster had effect on both exposure and memory [MGS07].

Wedel M. and Pieters R. developed a model of the processing that takes place to store information in long term memory: It is assumed that the number of fixations not their duration is related to the amount of information a consumer extracts from an ad. The authors applied the model to a sample of 88 consumers who were exposed to 65 print ads in its natural context, two magazines. Results support the assumptions [WP00].

Grundland G. and Eizenman M. conducted a study on consumer exposure and awareness of Outdoor advertising: eye-tracking was used to measure peoples attention towards out of home media such as trio boards, standard posters, super boards and bus shelter advertisements while driving a certain route in a car under various driving conditions and different times of the day. Data came from both drivers of the car and passengers. As in our work subjects were not informed about the aim of the study. Over 55% of the ads were seen. Passengers were more likely to see ads, than drivers. Among those, who looked at the advertising, on average they looked 2.04 times on a single drive by [GE06].

Visual perception research currently is not so much focused on where we look, but more on why we look (where we look). In the field this new focus in research emerged some years ago: away from the experimental understanding from where in a scene the eyes fixate in an image, to why the eyes choose a location [HB05]. According to Ballard D. and Hayhoe M. three principal complementary advances in research can be seen as the cause of the shift in focus:

The description of the role of eye movements in executing everyday visually guided behavior driven by the development of portable eye-trackers, the recognition of the role of internal reward in guiding eye and body movements and thirdly theoretical developments in the field of reinforcement learning. Together, these three developments have allowed the simulation of reward based systems incorporating realistic models of eye-movements over extended time scales [HB05].

Findings from this new approach have resulted in two fundamentally different approaches that exist in current research: bottom up and top-down [RBH07]. “Attention can be focused volitionally by "top-down" signals derived from task demands and automatically by "bottom-up" signals from salient stimuli.” [BM07]

Also other theories and hypothesis are being investigated. Karacan and Hayhoe [KH08] investigate whether attention is drawn to changes in familiar scenes. They examined mechanisms that control attention in natural scenes. The results of the study support the hypothesis that humans learn the structure of natural scenes over time and that attention is attracted by deviations from the stored scene representation. Though promising concepts and models of human gaze have been developed, none of them so far can explain fully what mechanisms direct eye movement.

3 Goals

The driving idea behind our work was to investigate human attentional behavior towards digital display systems in urban public transport under natural conditions. Are people aware of them and what is the attitude towards them? Can those displays act as attention catchers? Do different content-types result in different fixation times? Is some content more appealing and how long do subjects fixate on the screen on average? Further we investigated correlations between content type and the amount of focus time as well as the amount of time a person fixates on such a screen and the number of things one can reproduce freely or recognize shortly after the exposition.

4 Method

In order to answer these questions we conducted a field study using state-of-the-art mobile eye tracking equipment. In the rest of this section we will describe the field study in detail.

Field Study Context. The info-screen is a screen mounted in trams in the city of Graz in Austria and it is used to display out of home media. A typical configuration is shown in Figure 1. There are 4 info-screens installed in each tram. The size of the screen is 17 inch. Two of them are mounted towards the middle of the train and two in the front part. They are either pointing against the driving direction or sideways, as seen in the Figure 1. Visibility towards the screens is different depending on the location in the tram. Subjects were standing in different places within the tram and so the viewing distance was different as well. The info-screen is broadcasting its content in cycles that get repeated. One cycle lasts for 14 minutes (840 seconds), which is about the length subjects were in the tram in one direction. The categories forming a cycle at the time of the experiment were the following: Advertisement, News, Culture, Entertainment, Sports/Weather, Event-tips and passenger information.

Figure 1: Two info-screens in a tram

Participants. Invitation of test participants was based on a random sample of the population of the Austrian city of Graz, where the study took place. We started with the intention to have a representative sample, but unfortunately due to practical reasons (mainly willingness to participate in the study) we had to allow for an unbalance towards younger people. Alltogether 106 persons (47,9% female, 52,1% male) participated in our study. Mean age was 30 years, and participants were drafted from different educational (bachelor, master) and professional (apprenticeship, teacher) backgrounds. Due to technical problems (malfunction of the information displays in the tramway respectively problems with the mobile eye-tracking system) the date of 6 subjects could not be used for analysis.

Procedure. Participants were contacted by phone based on a random sample. Study participants were left uninformed about the detailed aims of the study, but they were provided with a general explanation of what the study was about: "to measure the attention behaviour in public spaces."

Participants were invited for a certain time and date to come to one of the researcher's office. After the participant arrived at the institute we would show the mobile eye-tracker and provide a description of the study, not telling that it was about the measurement of attention towards the info-screen in trams.

After this outline of the experiment subjects were asked if they still would like to participate in the experiment or not. If they agreed to do so the next steps consisted of calibrating the eye-tracker once subjects had the helmet on their head. The calibration was done using a 5-point calibration target. We checked the calibration again at the time subjects left the institute and when they came back from the public library to see if the helmet had gotten out of place.

The eye-tracker that was used in the experiment was an iView X HED mobile eye-tracker from SMI that consists of a size adjustable, lightweight helmet and a tablet pc that is carried in a daypack. The sampling rate of this eye-tracker is 50 Hz. Figure 2 shows the helmet with the mounted eye-tracker that was used in the experiment.

Figure 2: The mobile eye-tracker

Each subject was asked to carry out a simple task. We had borrowed books and magazines from a public library that is located about 25 minutes away from the institute were we met subjects. Subjects were asked to bring back a book to the library and then return to the institute using tramways. Participants were also instructed to use only the newest generation of tramway types, as in the older generation trams info-screens are not mounted in all vehicles.

We had prepared a detailed description of the route that subjects were supposed to follow including pictures describing the way from the institute to the tram station, the model of the tram subjects were supposed to take and the way from the station where subjects got of the tram to the public library. Outside of the public library a box is located enabling people to return books 24 hours a day. Subjects were told to just use this box to return books. Also this was documented using pictures in the description of the task. Subjects were also asked to make a note of the number of the tram they were riding in. This description was given to subjects together with a mobile phone in case they would have any additional questions on their way. Two more things were given to subjects: a ticket for the tram ride and an umbrella for the case of sudden rain to protect the eye-tracking system.

In total subjects were on their way for about 50 minutes. Weather conditions and time of day were recorded in a protocol.

When subjects returned they were asked to fill out a questionnaire consisting of 22 questions covering the following topics:

- had subjects realized and seen the info-screens in the tram
- subjective measurement of how long they in their opinion had paid attention to the info-screen
- recall and recognition of content shown on the info-screen; for recognition we provided a list with the ads actually shown on the info-screen including fake ones
- did subjects feel bothered by the helmet and the daypack
- general feeling about the info-screens
- questions about their usual behavior on the tram
- general questions about the content – what content did they like and what not
- questions covering advertisement – whether or not subjects like advertisement, what kind of advertisement they prefer (emotional vs. informative), which medium they prefer for advertisement
- demographic data

The questionnaire – presented in an electronic form - was filled out by us with subjects sitting nearby providing the answers. This was done that way because we also had subjects that were not familiar with computer technology.

Data resulting from the experiment. As described above subjects wore a mobile eye-tracker while on their way. The data from the eye-tracker, consisting of the scene video and the data of the eye movements, was recorded on the tablet pc that had been carried in a daypack. The company operating the info-screens provided us with the corresponding time-stamped log-files. The log-files, put together with the data from the eye-tracking provided us with a solid set of quantitative data about each participant's behavior. This data was supplemented by data resulting from the above described questionnaire.

5 Results

2.1 Eye Tracking Results

Of the 100 subjects 61% fixated the info-screen. Subjects spent a total of 48.7 hours in the tram. Of this time in 23.06 %, that is 11.23 hours, the info-screen was in the visual field of subjects. In 3.64% of the 48.7 hours attention was directed towards the info-screen. That is: in 16% (1.7 hours) of the time in which the info-screen had been in the visual field of subjects they were focusing on it. Attentive participants focused on the info-screen between 0.08 seconds and 9 minutes. Since the distribution is right-skewed we use median and interquartile range (IQR) to describe location and dispersion: the median of the sum of fixation time on the info-screen over all participants is 74 seconds (1853 frames); the IQR is 135 seconds (3399 frames). Figure 3 on the horizontal axis shows fixation time in minutes and on the vertical axis the according proportion of participants in percent, e.g. 42.6% of attentive participants fixated on the info-screen up to one minute.

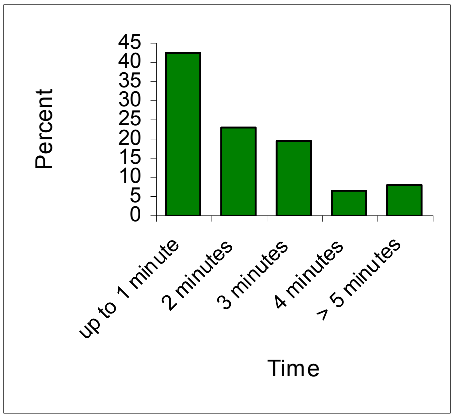


Figure 3: Fixation proportion in percent

As described above the content on the info-screen was composed of 7 different categories, not of equal length. Figure 4 provides an overview of the distribution of the different content-types over all participants in descending order, e.g. “News” was broadcasted in 27.6% of the total time.

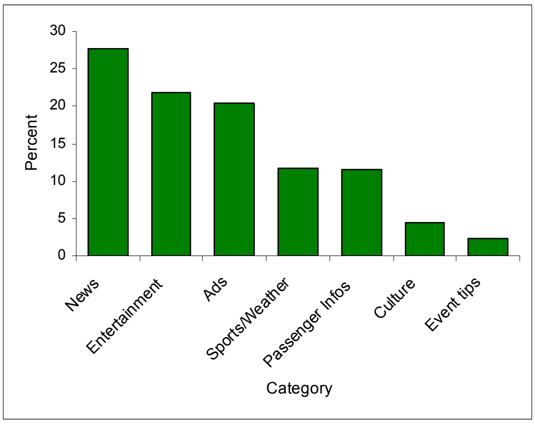


Figure 4: Distribution of different Content-types in percent

One of the main questions in this research was to find out about correlations between focus time and content-type. Does the type of content have an influence on the time participants fixate on the info-screen? To answer this question we have compared the proportions of fixation time for each of the 7 content types. To explain further: We have summarized fixation time (measured in frames) over all participants for each category. Then we computed for each category the ratio between the sum of fixated frames over all participants and the sum of frames this category had been displayed and expressed this in percent. So results indicate the percentage of fixation for each category. Figure 5 shows the resulting distribution of fixation-time against content types in percent. The content type "News" was fixated in 4.54% of the time it has been displayed, followed by "Sports/Weather" (3.98%) and "Entertainment" (3.81%). Remaining categories got around 3 percent except for "Culture" which was fixated in 2.65 % of the available time.

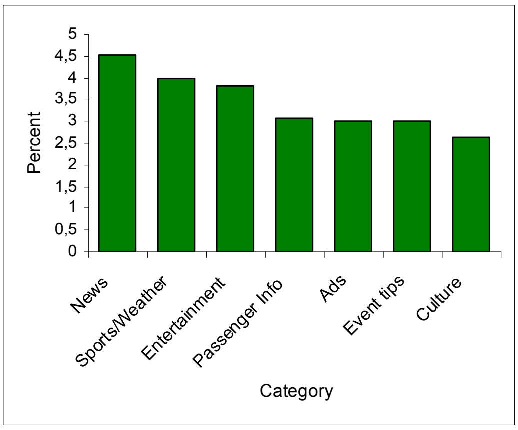


Figure 5: Distribution of proportion of attention against content types in percent

Next we investigate the same question but address potential influence of category on fixation time for each individual participant: We have computed the ratios between the sum of frames a participant fixated a certain category and the sum of frames this category was displayed and expressed this in percent. We only considered participants having fixated more than zero frames. This way we obtain for each participant and each category the proportion s/he was fixating on the info-screen, based on the total amount the info-screen has been available. The following image resulted: Over all categories the distributions are right-skewed and the medians reflect relative similarity (min=3.4; max=7.4). The skewness of a distribution indicates outliers in the data; in this case they can be interpreted as participants focusing many more frames compared to the average participant. The similarity of the medians indicates no big differences in focus time between categories. To test whether differences between content types are statistically significant we performed a Kruskal Wallis Test. The result (p = .392) is not significant at an α-level of 0.05. Figure 6 shows for each category over all participants individual the distribution of proportions of fixated frames over participants.

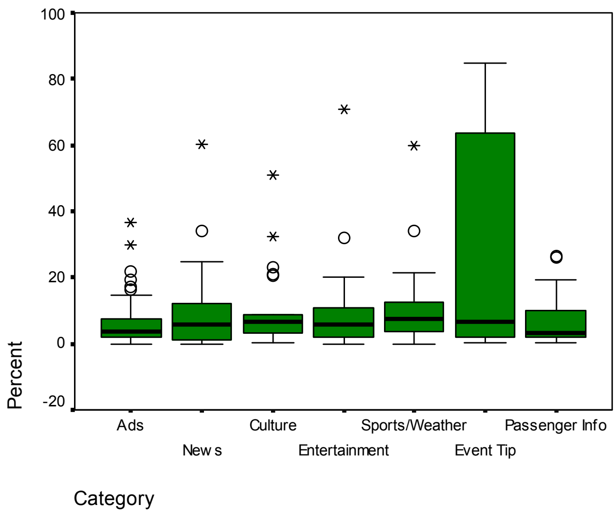


Figure 6: box-plots showing distribution of proportion of fixated frames (in percent) for each category and all persons

From the obtained results we conclude that the content type does not necessarily have an influence on the fixation time on an info-screen under the given circumstances. We addressed the question whether one content type is of greater interest than another also in the questionnaire: we asked participants to indicate whether a given content type interests them or not. Due to an organizational error in the beginning of the study we were left with data from only 96 participants instead of 106. “News” and “Event-tips” were rated by most participants as interesting content-types, followed by “Weather” and “Sports”. The questionnaire data and the data from the eye-tracker do not correspond fully here. There are some discrepancies between the actual fixation-behavior in the tram and the way participants rated contents at the end of the study. For example, few participants found “Advertisement” interesting (see Figure 7). None the less the number of participants’ fixating on “Ads” did not differ from any other content type. One possible explanation for this could be the fact that fixation time on the info-screen does not automatically mean that the participants’ eyes were actively following the content presented; it just means that the participants’ eyes were focusing on the info-screen. So, even while participants were having their eyes on the info-screen they might have thought about something completely else. We address the issue of length of fixation time and its correlation with the memorized content in greater detail below.

Figure 7: Interest in Content shows in percent how many of the participants rated a given category as interesting or not.

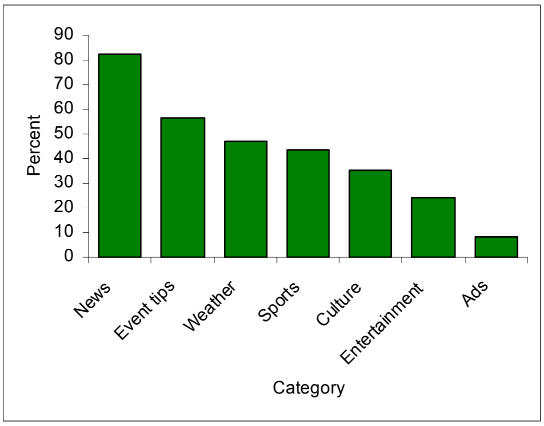


Figure 7: Interest in Content-category in percent

We were also interested in finding out whether time of day has an influence on the fixation behavior towards the info-screen. Figure 8 (x*100 = percent) shows peaks of mean fixation time on the info-screen between 8 and 10 a.m. and again between 14 and 16 p.m. A possible explanation for this might be the fact that in the morning-hours participants’ interest for news is higher than during the rest of the early day hours. However, the even higher peak in the afternoon remains unexplained.

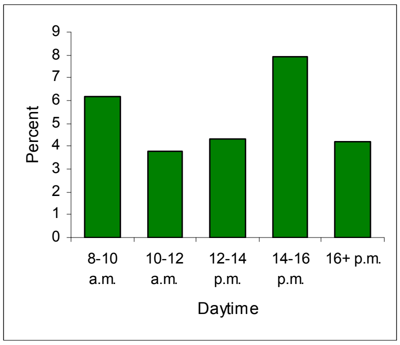


Figure 8: Influence of day-time on fixation on info-screen

2.2 Questionnaire

As described above additionally to the eye-tracking data we have used a questionnaire in this research. We have done so to complement and compare the two sources of data and to assess participants' thoughts on the subject.
Participants were not bothered much by the helmet and the daypack: 23% of the participants were not disturbed at all, 70% reported to be disturbed a little bit and only 5% felt severely disturbed by the equipment. Also, most participants reported to not feeling nervous (64%) during the ride at all, 34% felt a little nervous and only 1% reported having felt very nervous.

88 participants reported to have seen the info-screen; that number is not in line with the data we obtained from the eye-tracking (61 participants were fixating the info-screen for at least one frame (1/25 of a second) on the info-screen). Subjects were asked to measure their subjective feeling about the attention-span towards the info-screen. Four possibilities were given: no attention, up to 2 minutes, between 2 and 5 minutes and more than 5 minutes. Most subjects thought about their attention-span being up to 2 minutes, which corresponds to the measured average value of attention-span of 104 seconds. Figure 3 also shows that most subjects fixate the info-screen up to 2 minutes.

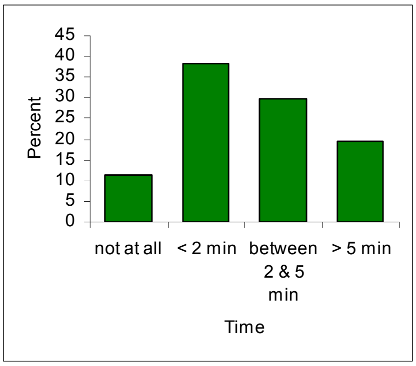


Figure 9: how long (do you think) have you looked at the info-screen

As already mentioned above we researched an eventual correlation between the length of fixation on the info-screen and the amount of content memorized. To measure what participants had memorized from the content presented on the info-screen we asked them after they returned from the library to tell us in no specific order which content they had seen during their ride. The remembered contents were categorized by us. Additionally, subjects were asked to recall what specific ads, logos and brand names they could remember. From all participants 80 were able to recall categories. On average, subjects recalled three different categories. 27 subjects were able to recall specific spots and logos. 18 persons recalled brand names.

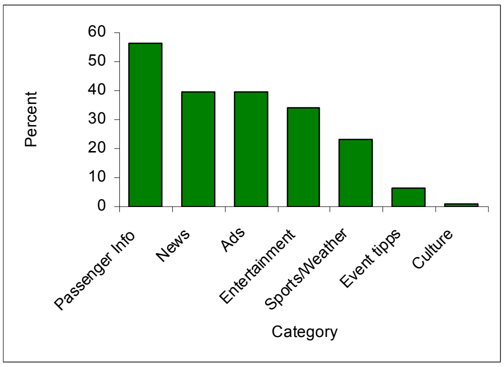


Figure 10: recall of categories in percent

Additionally to letting participants reproduce what they had remembered we measured recognition. For this purpose we composed a list of ads actually being broadcasted with ads that were not shown on the info-screen. 43 of the participants were able to recognize spots correctly; however, they only recognized few spots, between 1 and 4 ads (average: 1.37).

The hypothesis of an eventual correlation between fixation time on the info-screen and recall/recognition was tested next. Although intuition might suggest a correlation between the two our data does not support that hypothesis; correlations are r = 0,102 between number of frames fixated and number of content recalled respectively r = 0,341 between the number of frames fixated and number of ads recognised correctly. One explanation for this result might lie in the experimental setup: participants had no idea about the purpose of their ride and no order to look at the info screen or to even remember any of the content being broadcasted. Another explanation could be the fact that, as already mentioned above, fixating the eye-tracker does not automatically mean perceiving its content.

Ultimately we would like to report on results describing the phenomenon info-screen and what participants think about it in general as well as preferences regarding the medium of advertisement. Two thirds of the participants rated the presence of the info-screen as being good, 20% where indifferent and 10% didn't like it at all. Subjects in the questionnaire were asked about their preferred media for advertising: TV-ads and posters were the preferred medium.

6 Discussion

The results we obtained suggest that more than half of the participants did focus on the info-screen and nearly nine out of ten of our participants reported to have noticed the screen. Content type does not influence length of fixation time, nor does any noteworthy correlation between length of fixation on the screen and memory or recall exist. One has to take into account that an eye-tracker only can track where a person does look, not what the person is doing when looking. Regarding content types as attention catchers we identified "News" as the content type on which most attention chains started; a change from "News" to "Sports/Weather" provided the content combination that was followed through most often. We think that - considering the fact that participants did not know what the research was aimed at (attention behavior towards info-screen) and the natural environment the study was carried out in - those results are very interesting because of their explorative character. We learn from this study that fixation time on a digital screen in public transport seems not to depend on content-type. Also, fixation time per se does not seem to be a suitable indicator for content-memorization. This result directs towards the findings from [WP00]. Future research could focus on an eventual correlation between the frequency of fixations on content and its memorization. Despite the eye-tracking results we have found that on average participants recalled three different categories out of seven. That seems to be a good proof that the content presented on the info-screens comes through to people. The findings from this work can serve as a basis for future research; e.g. can we actively and reliable catch people's attention in a natural environment like this one (or any other urban environment) and how can we optimize content for such a new and challenging form of advertising to improve content memorization? We shall not forget that digital screens positioned in the urban area despite their similarities with TVs do miss a channel for communicating with its audiences - the audio part.

7 Acknowledgements

This work is part of the project LOOKAT and has been partly funded by the Austrian Research Promotion Agency (FFG) under contract number 815462. We would like to thank all partners of the LOOKAT consortium and the Austrian Research Promotion Agency for support.

References

[BM07] Bushman T.; Miller E.: Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. Science, 03/2007.

[OS07] Digital Out-of-Home Media Awareness & Attitude Study United States, OTX and SeaSaw Networks, 2007.

[SG07] Digital Signage in Europe: opportunities for digital out-of-home advertising, Report. Produced by screendigest in partnership with German research company Goldmedia, 2007.

[Da02] Duchowski A.: A Breadth-First Survey of Eye Tracking Applications. Behavior Research Methods, Instruments, and Computers, 2002.

[GE06] Grundland G.; Eizenman M.: Outdoor Advertising Consumer Exposure Study (OACES), Marketing Magazine, 10/2006.

[HB05] Hayhoe M.; Ballard D.: Eye movements in natural behaviour. TRENDS in Cognitive Sciences, Vol. 9, April 2005.

[KH08] Karacan H.; Hayhoe M.: Is attention drawn to changes in familiar scenes? Visual Cognition, 2008.

[MGS07] Maughan L.; Gutnikov S.; Stevens R.: Like more, look more. Look more, like more: The evidence from eye-tracking. Journal of Brand Management, Vol. 14, 2007.

[RBH07] Rothkopf A.; Ballard D.; Hayhoe M.: Task and context determine where you look. Journal of Vision 2007.

[SVL04] Shavitt, S.; Vargas, P.; Lowrey, P.: Exploring the Role of Memory for Self-Selected Ad Experiences: Are some advertising Media Better Liked than Others? Psychology & Marketing, 2004.

[OO00] The PRS Eye-Tracking Studies. Outdoor advertising association of America, 1999 to 2000.

[WP00] Wedel M.; Pieters R.: Eye Fixations on Advertisements and Memory for Brands: A Model and Findings. Marketing Science, Vol. 19 2000.