



Observing the Context of Use of a Media Player on Mobile Phones using Embedded and Virtual Sensors

Jakob Eg Larsen, Michael Kai Petersen, Nima Zandi, and Rasmus Handler

Wi: Journal of Mobile Media, Spring 2011

The online version of this article can be found at:

<https://doi.org/10.65968/OGCW5531>

Larsen, Jakob Eg, Michael Kai Petersen, Nima Zandi, and Rasmus Handler . "Observing the Context of Use of a Media Player on Mobile Phones using Embedded and Virtual Sensors". *Wi: Journal of Mobile Media*. Spring (2011). Web. <https://doi.org/10.65968/OGCW5531>

Observing the Context of Use of a Media Player on Mobile Phones using Embedded and Virtual Sensors

Jakob Eg Larsen, Michael Kai Petersen, Nima Zandi, and Rasmus Handler

Introduction

Mobile phones have become ubiquitous and an integrated part of our everyday life. In the last couple of years smart phones have received increased attention as application stores (e.g. Apple App Store, Android Market, and Nokia Ovi Store) have enabled easy distribution of mobile applications. New smart phone features and sensors have enabled a wide range of novel mobile applications, especially within games and media consumption domains. In the area of music applications a number of different mobile applications have been created. One example is mobile applications for the Internet radio Last.fm, which enable recommendation of music similar to the user's favorites based on social networking features. MoodAgent is an example of an application that enables the user to navigate a music collection in terms of mood, rhythm, and style of the music.

However, getting an understanding of actual behavior and use of such mobile applications is difficult as the situations of use are inherently mobile. This makes it relevant to observe music listening habits "in-the-wild", as laboratory studies may

prove insufficient for reproducing realistic scenarios to capture and getting insights into actual contextual user experience. Thus methods and techniques to acquire information about actual mobile use in context are needed to obtain a better understanding of the mobile user experience.

In the study described in this paper we use a software framework we have developed, which is capable of acquiring contextual data from embedded mobile phone sensors during daily life use by a mobile phone user. Our software runs transparently in the background on the mobile phone and is logging data from multiple embedded sensors. This allows us to describe and understand information about surrounding people (devices) and places (locations), as well as application and media usage. In the present study our focus is on studying the use of the media player application on mobile phones, and in particular to understand the context in which it is being used. We hypothesize that contextual information obtained from embedded mobile phone sensors can offer useful information in terms of understanding the situations of mobile use involving the media player application. Furthermore, we discuss how such contextual information can provide inspiration for new types of context-aware mobile user interfaces and interaction, such as in the present case user interfaces for music recommendation applications.

Related Work

Field versus laboratory evaluation of mobile applications has been debated in the HCI literature. Kjeldskov et al. [7] has shown that few additional usability problems were found in field experiments compared to laboratory experiments and suggested that establishing the right laboratory settings can provoke similar findings [6]. On

con-trary Duh et al. [2] found that significantly more (and more severe) usa-bility problems were iden-tified in field ex-pe-riments. This is supported by Bernard et al. [1] that found that contextual changes had a strong impact on behavior and performance. Roto and Olasvirta [12] studied mo-bile users on the move using web-browsers on mobile pho-nes by employing mul-tiple cameras worn by the test par-ti-ci-pants and a moderator stay-ing in proximity of the test par-ti-ci-pant to monitor the ex-pe-riment. Shorter attention span was observed when using applications on the move com-pa-red to use in la-bo-ra-to-ry settings [11]. Froehlich et al. [3] took an approach re-sem--bling ours in combining quan-ti-ta-tive and qualitative me--thods for “in-the-wild” col-lec-tion of data about usage, in--clu-ding device logging of app-lication use and context using embedded sensors. In [], [] and [] this was con-si-dered in music scenarios. Such studies under-line the im-por-tance of undertaking studies of observed “ac-tu-al use” in con-text, instead of “learning to use” an un-fa-mi-li-ar app-li-ca-tion in a controlled environment as described above. Such stu-dies ty--pi-cal-ly only cap-ture use over a short period of time and re-veal little about actual use or use pat-terns over an ex--ten-ded time pe-riod where learning and ha-bi-tu-a-tion has ta-ken place.

Mobile Context Logging

To observe test participants while using mobile phones and app-lications in real-world settings, we have created a mobile con-text logging software framework for mobile phones. The Mobile Context framework for mobile phones [8] allow us to carry out “in-the-wild” experiments where acquiring data from mul-ti-ple em-bed-ded mobile phone sensors is required in order to esta-blish information about the context of use. The software fra--me--work installed on a mobile phone can obtain in-for-ma-ti-on from

embedded sensors including accelerometer, GPS, Bluetooth, WLAN, microphone, call logs, calendar, and additional sensors can be added for specific experiments. An overview of the software framework as used in the present study is shown in Fig. 1, and illustrates the emphasis on aggregating multiple sensor data into higher-level context descriptions.

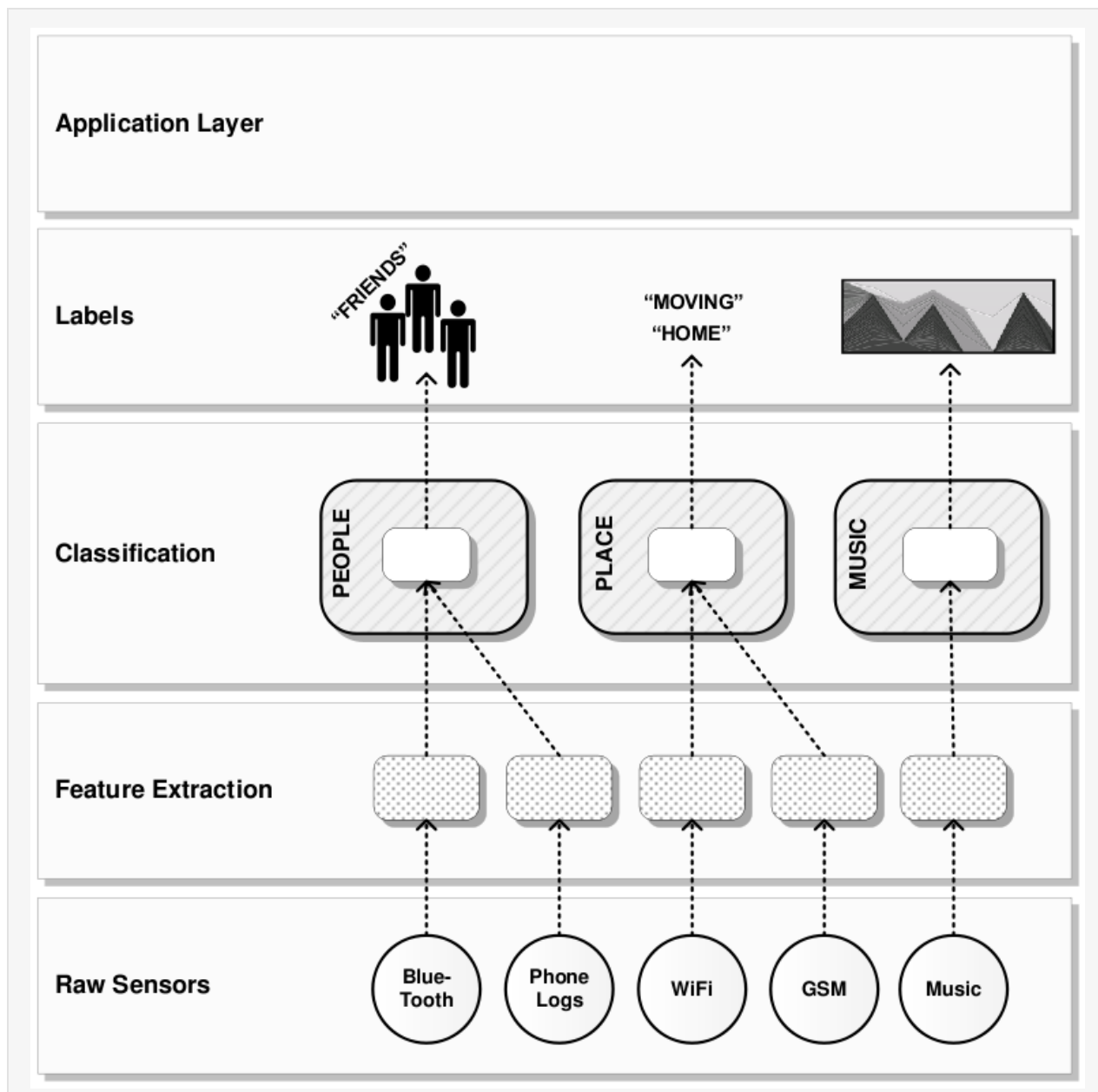


Fig. 1. Mobile Context Framework taking raw sensor data, extracting features, and translating into contextual labels describing people, places, and music

Method and Experiment

In our experiment 7 participants (4 male and 3 female, ages 18-29) were provided with a Nokia N95 mobile phone with our software framework installed. They were instructed to carry and use the mobile phone as they would normally use their own mobile phone for a two-week duration. In addition they were told to use the mobile phone as their MP3 player device and encouraged to upload their own music collection to the mobile phone. Thereby encouraging them to listen to the music that they liked and they would typically listen to on a daily basis. The software was running silently and continuously acquiring data from the embedded mobile phone sensors for the duration of the experiment. The framework included a virtual sensor component capable of acquiring data from the embedded media player application on the mobile device. The component obtained data about music tracks being played, the duration of the song, and current playback position. In addition ID3 meta-data from each song including artist and title information was retrieved. As shown in Fig. 1, information was obtained from the Blue-tooth sensor and phone logs in order to extract features describing the user's social relations. Features describing places were extracted from information acquired from Wi-Fi and GSM cellular network information. Each of these features was translated into labels describing people and places.

Results

The data acquired during the two-week duration for the 7 participants in the experiment is shown in Table 1. The participants were fairly active using the media player listening to 94-292 songs in total, meaning 7-21 songs on a daily basis on average.

The played tracks indicate the number of tracks that were started on the media player, but we only denote a track as listened to when more than 50% of the track has been played. The number of unique music tracks indicates that some participants listened to a smaller set of tracks repeatedly. Each track played was logged with a time-stamp meaning that the time of day where the media player was being used could be analyzed.

| Participant | Gender | Age | Tracks played | Listened tracks | Unique tracks | Known places |
|-------------|--------|-----|---------------|-----------------|---------------|--------------|
| 1 | M | 25 | 337 | 160 | 85 | 7 |
| 2 | M | 26 | 474 | 153 | 100 | 11 |
| 3 | M | 26 | 375 | 190 | 48 | 5 |
| 4 | F | 18 | 524 | 292 | 68 | 6 |
| 5 | M | 29 | 173 | 110 | 58 | 7 |
| 6 | F | 23 | 742 | 167 | 124 | 6 |
| 7 | F | 24 | 198 | 94 | 65 | 7 |

Table 1. Overview of music listening for the 7 test participants

The information on music listening was coupled with the analysis of the contextual labels acquired from the logs of embedded sensor data. The GSM cellular information and Wi-Fi access points were analyzed in order to determine locations. Based on the

analysis it was possible to determine the places in which the participants spend the most time. Thus it was possible to determine if a participant was at “ho-me”, in a “known place” (a place where time was spent re-peatedly), an “unknown place”, or “in a transition” bet-ween places (continuous changes in GSM and Wi-Fi data in mi-nu-te size time windows). As an example Fig. 2 illustrates 7 known pla-ces discovered for a participant by analyzing the co-occur-rences of Wi-Fi access points. Each node re-pre-sents a Wi-Fi access point and edges denote access points dis-covered concurrently. In this example access points seen below a threshold has been excluded for clarity. Using this method an average of 7 known places were identified for the participants, as shown in Table 1. The time where most time was spent was assumed to be the “home” place.

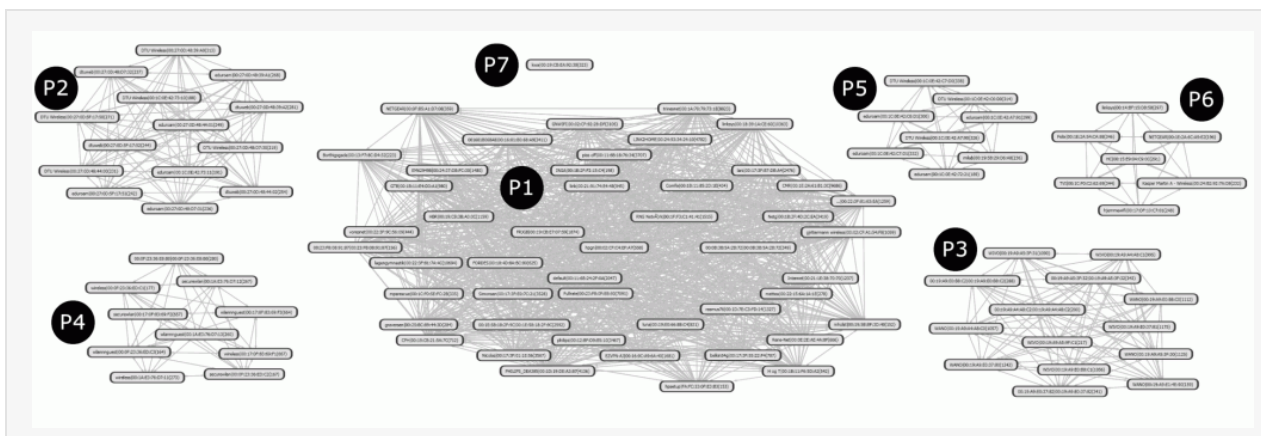


Fig. 2. Seven known places P1-P7 discovered for a participant by analyzing the co-occurrences of Wi-Fi access points over two weeks. Each node represents a unique Wi-Fi access point.

The social relations were mapped based on the data ac-qui-red from correspondence logs (phone calls and SMS mes-sa-ges) in terms of who was calling and sending mes-sages to whom. Based on Bluetooth device discoveries (of par-ti-ci-pant mo-bile phones) it was possible to map out when the par-ti-ci-pants were in physical proximity of each other or in proximity of other people. The social relations of the par-ti-ci-pants

based on the mapping of the Bluetooth data are shown in Fig. 3. The numbers on the edges denote the number of Bluetooth discoveries indicating the time spent in proximity and the number of correspondences (calls and SMS messages).

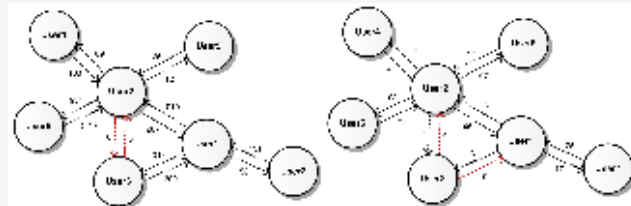


Fig. 3. Social relations of the 7 participants mapped based on Bluetooth-based proximity (left) and call/SMS logs (right)

Based on this data it was possible to establish the context of use of the mobile media player on the mobile phone described as time, places, and people. It was possible to determine the time and places in which the music was being played and it was possible to determine the people present when the media player was being used for music playback.

The analysis of the music used the track metadata that was acquired from the media player during playback. This was enhanced by social tags, which offer a large-scale source of descriptive labels about music exceeding the information available in typical track metadata. Based on the artist and song title (ID3 tag) metadata the collaborative tagging of music tracks available from Last.fm was used to enhance and capture the underlying semantic aspects of the music. Based on a latent semantic analysis of co-occurring tags Levy and Sandler [] have described a method to enrich the original metadata with additional descriptions of related genres that facilitate a comparison of the songs. The model involves 90 specific semantic aspects describing the music based on social tags. For each aspect the ratio of tags present was calculated

and combined to obtain a track signature. Subsequently we collected tracks into sessions if the participant has been listening to at least 3 tracks in a row. Thereby combining the above track signature into a session signature constituting an average of the ratios of tag co-occurrences defined in the track signature. For clarity we generalized the tag co-occurrences into 8 broader categories of musical style, which allows us to observe the changing genre characteristics over time [], as illustrated in Fig. 4.

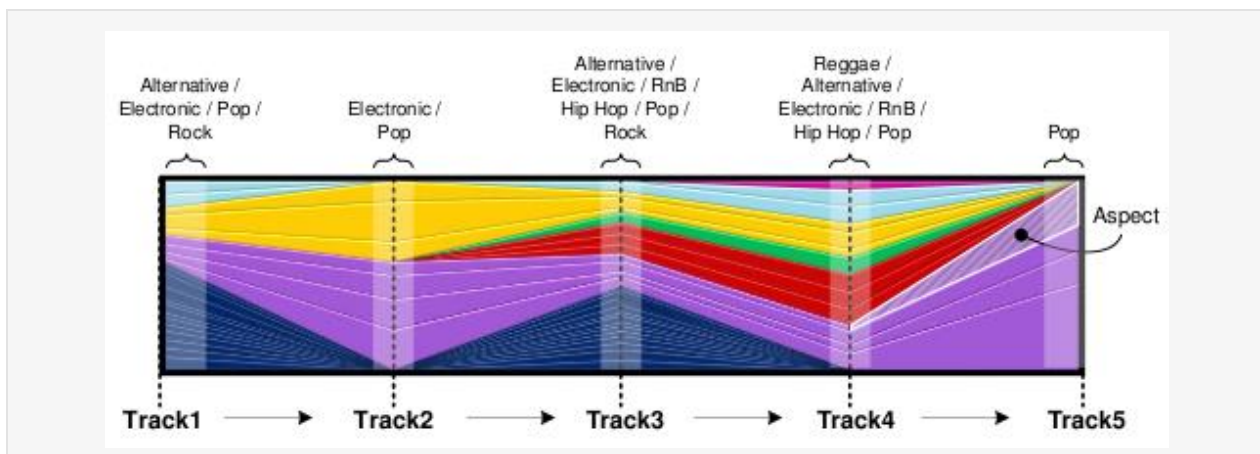


Fig. 4. Example five track sequence profile derived by enhancing the track metadata with Last.fm tags. Each color represents a music genre, as annotated at the top.

This approach introduces a way to describe the music being listened to in terms of a few generalized genres. This allows us to consider and compare which genres of music are being played over time by the participants and the particular contextual settings in which it is being played.

Our analysis of the data on music listening behavior indicated interesting patterns. First of all, the genres of music that were listened to over time highly depended on the context of the user. In some places one set of genres were typically played, whereas in transition between places a different set of genres were being played. The level of

inter-action (e.g. skipping songs) also depended on the context of use. Transitions between places were characterized by frequent interaction (skipping and choosing tracks), whereas in known places the interaction was less frequent. The time of day also had an influence on the music listening behavior.

Discussion

Our initial findings have indicated an inter-play of time, context (places), social relations, and the media (music) being played. Based on these initial findings we propose that media player user interfaces could benefit from involving parameters including time, places, people, music genre as multiple dimensions along which the user could navigate the available music in a music collection. This means that the mobile context could potentially play a much more prominent role in mobile applications, such as the media player. In this case, an obvious example is recommendation systems that not only recommend music based on music similarity, but also contextual similarity. Current music recommendation interfaces typically rely mostly on music similarity, but for instance Last.fm partly involve the social dimension enabling the user to discover music that similar users have listened to. We find the potential of involving more contextual parameters discussed in this study intriguing and suggest that these findings could inform designers and inspire new kinds of context-aware music navigation interfaces for navigating and discovering content in the media player.

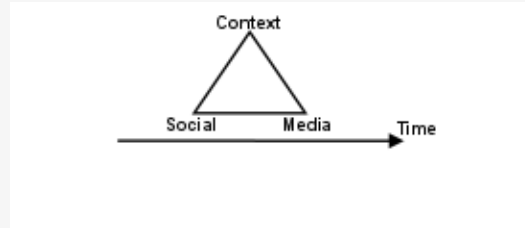


Fig. 5. Illustration of the 4-dimensional feature space that could shape novel context-aware music navigation interfaces

Fig. 5 outlines a feature space that could be considered to capture the four contextual parameters identified as shaping the music listening behavior in the study. Next generation interfaces designed along these lines, could allow users to browse playlists purely based on the similarity of tracks, incorporate listening habits found within similar contexts or alternatively take into account what other relevant users are listening to based on contextually derived social graphs. Thereby allowing the user to discover music based on si-mi-la-ri-ty along multiple dimensions.

The initial findings in this study indicated the importance of studying the “actual use” of mobile app-li-cations “in-the-wild” over labo-ratory studies should be em-pha-sized. Studying the mobile user experience in con-text could potentially be a valuable source of inspiration for de-signers. This study has provided valuable insights on how the context of use had im-pli-ca-tions for the in-ter-action with and the content chosen in the media player application. Since our study is based on two weeks of data acquired from 7 participants, it has pro-vi-ded some initial indications on the potential for alternative con-text-aware music inter-fa-ces. Future work could involve lar-ger-scale studies with mo-re par-ti-ci-pants, as well as ex-plore novel context-aware mu-sic navi-ga-tion interfaces. It would also be relevant to ap-ply these methods in the study of other application domains.

Conclusions

We have observed 7 mobile phone users using a mobile phone application in real-world settings. Based on logged information from multiple embedded mobile phone sensors we have been able to establish information about the time and context of use of a media player application on mobile phones and we have been able to identify how the context has implications for music listening behavior and the use of the application. We conclude that contextual information can offer valuable insights to the where and when of mobile use and provide valuable insights forming input for new approaches to mobile context-aware applications and user interfaces.

Acknowledgments

We would like to thank the participants that took part in our experiments. Also thanks to Nokia Denmark and to Forum Nokia for the equipment used in the experiments.

References

- Barnard, L., Yi, J.S., Jacko, J.A. and Sears, A. Capturing the effects of context on human performance in mobile computing systems. *Pers. & Ubiqu. Comp.* 11(2) (2007).
- Baumann, S., Jung, B., Bassoli, A. and Wisniowski, M. Bluetuna: let your neighbour know what music you like, in *CHI'07 extended abstracts on Human factors in computing systems* (2007) 1941–1946.
- Duh, H.B.L., Tan, G.C.B. and Chen, V.H. Usability evaluation for mobile device: a comparison of laboratory and field tests. *Proc. of the 8th conf. on HCI with mobile devices and services* (2006) 186.
- Froehlich, J., Chen, M. Y., Consolvo, S., Harrison, B. and Landay, J. A. MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. *Proc. 5th Int. conf. on Mobile systems, applications and services* (2007) 57-70.
- Kjeldskov, J. and Skov, M. B. Studying usability in vitro: Simulating real world phenomena in controlled environments. *Int. Journal of HCI* 22(1) (2007) 7-36.
- Kjeldskov, J., Skov, M. B., Als, B. S. and Høegh, R. T. Is it worth the hassle? Exploring the added value of evaluating the usability of context-aware mobile systems in the field. In *Proc. MobileHCI* (2004) 529-535.

Larsen, J. E. and Jensen, K. Mobile Context Tool-box: an extensible context framework for S60 mo-bi-le phones. In *Proc. EuroSSC, Springer LNCS 5741* (2009) 193-206.

Levy, M. and Sandler, M. Learning latent semantic models for music from social tags. *Journal of New Music Research* 37(2) (2008) 137–150.

Oulasvirta, A., Tamminen, S. and Roto, V. and Kuorelahti, J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. *Proc. of the SIGCHI conf. on Human factors in computing systems* (2005) 928.

Reddy, S. and Mascia, J. Lifetrak: Music in tune with your life, in *Proc. of the 1st ACM Int. workshop on Human-centered multimedia* (2006) 25–34.

Roto, V. and Oulasvirta, A. Need for non-visual feed-back with long response times in mobile HCI. *Special interest tracks and posters of the 14th Int. conf. on World Wide Web* (2005) 775-781.

Seppänen, J. and Huopaniemi, J. Interactive and context-aware mobile music experiences. *Proc. of the 11th Int. Conf. on Digital Audio Effects* (2008).

Zandi, N., Handler, R., Larsen, J. E. and Petersen, M. K. People, Places and Playlists: modeling soundscapes in a mobile context. *2nd Int. Workshop on Mobile Multi-media Processing* (2010).

Biographies

Michael Kai Petersen's research focus is cognitive modeling of digital media, combining "bottom-up" machine learning elements of latent semantics with "top-down" behavioral aspects of emotions, in order to emulate how we perceive content. An approach that might potentially be applied in applications ranging from web 2.0 social network interaction and sentiment analysis to cognitive neuroscience.

Petersen has 30 years of experience within digital media and has since 2004 been associated with DTU where he received his PhD doctorate in 2010 and was appointed Assistant Professor in the Cognitive Systems Section at DTU Informatics. He has a combined technical and creative background, holding M.Mic master degree in mobile internet communication from DTU 2004, while previously being trained in digital sound engineering as a producer in DR Danish Broadcasting Corporation, after completing his studies in music graduating from the soloist class of the Royal Danish Academy of Music in 1982. As an entrepreneur he has founded three start-up companies over the period 1993-2003 and has furthermore produced 100 plus CD albums. At DTU he is working with aspects of personalized context awareness at the MILAB mobile informatics lab, as well as teaching courses within the Digital Media Engineering master program related to building collective intelligence based on metadata, sensors and web 2.0 mobile interaction.