Wearable Ambient Sound Display: Embedding Information in Personal Music

George Chernyshov[1], Benjamin Tag[1], Jiajun Chen[1], Vontin Noriyasu[2], Paul Lukowicz[3], Kai Kunze[1]
[1]Keio Media Design, Japan; [2]Fujitsu Design, Japan; [3]DFKI, Germany
(chernyshov, tagbenja, cyrilchan)@kmd.keio.ac.jp, vontin.noriyasu@jp.fujitsu.com, paul.lukowicz@dfki.de, kai.kunze@gmail.com

ABSTRACT

In this paper we explore how to embed information into users' music playlists while limiting the obtrusiveness to the user. We focus on continuous over time information rather than discrete information (e.g. "monitoring traffic flow" versus "new email received"). We are presenting results for a traffic monitoring task (4 levels) for 10 users that are engaged in either a cognitive (Sudoku solving) or one of two physical tasks (running on a treadmill or playing table tennis), with up to a 98 % detection rate.

Categories and Subject Descriptors

H.5.2. [**Information Interfaces and Presentation**]: User Interfaces-Auditory (non-speech) feedback, evaluation/ methodology

Keywords

Sonification; wearable audio display; process monitoring; distraction free; frequency filtering

1. INTRODUCTION

Delivering information from the digital domain to users who are engaged in primary tasks within the physical world, in a minimally disruptive way, is a core component of the wearable computing vision. For simple information such as individual notifications (e.g. "new email") many possibilities exist and have been realized, e.g. in smart watches. A much more difficult question is how to allow the user to `continuously` monitor a process over time, similar to a progress bar throughout the entire download process rather than a discrete notification when a task has been completed (or failed). Not only is it a fundamental human desire to be able to predict and anticipate rather than wait for things to happen "out of the blue", but also are temporal patterns in certain situations of more interest than the arrival of a discrete event.

In this paper we investigate a way to transmit information in a minimally disruptive way by subtly altering the music people listen to. This work is motivated by two observations. First, it is well known that people are capable of perceiving subtle, complex information through an audio channel while being cognitively engaged in complex tasks. Second, people like to listen to music on their personal devices while being physically active: from sports through work. Thus, encoding information in music on a wearable device would allow people to access it in a minimally obtrusive way.

The basic idea is to encode the information by altering the equalizer settings. Extreme changes are easily noticeable, however people are used to listening to their music on different devices with different equalizer settings. Therefore, as long as the changes are not too significant they are unlikely to have too much impact on the listening experience. We are focusing on the specific use case of traffic monitoring; showing that basic patterns can be perceived. Whilst being engaged in cognitive and physical activities, namely solving Sudoku puzzles, running on a treadmill, or playing table tennis, participants listen to popular music presenting different equalizer patterns.

2. RELATED WORK

Most research in the ubiquitous computing domain is concerned with conveying notifications to the user [3, 4]. These works mostly focus on discrete events (e.g. receiving a phone call). They convey status but in a discrete manner. Mostly, researchers do not aim to use music the user already listens to, but synthesize specific soundscapes based on the information content they want to convey. Mynatt et al. use voice, music, sound effects or a combination of these as notification sounds in their audio aura system [6]. In terms of music, they utilize different short melodies carrying different meanings. There is also a great amount of research on using sound and audio for localization and communication between devices [7, 1], of which the most interesting is by Madhavapeddy et al. [5]. However the focus of the work is still device to device communication. Russel et al. show how to implement a location based augmented reality soundsphere [8]. We see this as complementary to our approach since the sounds and music adoption also convey information, yet their main goal is entertainment. Butz et al. present work that is close to ours, namely encoding information in music [2]. However, the music piece in itself is completely generative. Therefore, listening comfort highly depends on the information content.

3. EXPERIMENTS

For each of the following experiments we invited 10 participants. The average age was 26.58, median: 25, range:23-38. 80% male 20% female. The first two experiments are performed in order to find the limits of perceivable changes using a frequency band filtering technique that would still not affect the pleasure of listening

ISWC '16, September 12-16, 2016, Heidelberg, Germany
© 2016 ACM. ISBN 978-1-4503-4460-9/16/09...$15.00
DOI: http://dx.doi.org/10.1145/2971763.2971789

perceived levels (columns); actual levels (rows)

Physical	0	1	2	3
0	75	25	0	0
1	12.5	46.9	34.4	6.24
2	0	21.1	51.6	27.2
3	3.12	0	41.8	55.1

Mental	0	1	2	3
0	54.5	45.5	0	0
1	14.3	57.1	28.6	0
2	4.16	33.3	41.7	20.8
3	0	4.76	42.9	52.4

Figure 1: Confusion matrices for traffic monitoring with simulation of physical and mental load.

to music. The main findings are that 7-10db muffling or boosting of a chosen frequency band is noticeable while not being too obtrusive. Another finding is that participants can easily notice the trend of muffling or boosting the db level change over time. However, most of the subjects are not able to recognize absolute db values of the introduced frequency muffling or boosting effect nor can they easily compare two effects and tell which one is stronger.

In the main experiment we investigate if people can interpret the perceived changes as information. Considering the findings from the first two experiments, we decided to convey information on traffic conditions using the amplitude modulation of a sine-shaped wave. This is basically the same technique that is used in AM radio broadcasting. We picture a setting where an employee wants to leave the office when the traffic conditions on his/her way home are the best. Participants are asked to monitor a hypothetical traffic condition while performing one of two activities: a mental or physical activity. The traffic conditions are split into four levels: no traffic, little traffic, heavy traffic and traffic jam with corresponding EQ-amplitudes: 0, 3, 6, and 9 db. Muffling/boosting is applied to the 0.5-1.4 kHz band only, which corresponds to mid-range frequencies. Mid-range frequencies are reproducible even with low-end headphones and speakers and are present in a vast majority of music pieces. For the mental activity task the participants are asked to solve Sudoku puzzles before, after and during the experiment. The solving time is used to measure how distracting the proposed approach is. However, the changes of the solving time were not statistically significant. For the physical activity simulation, participants are asked to play table tennis or jog on a treadmill. Due to significantly noisier environments, we decided to make the changes more audible by increasing the amplitudes to 0, 4, 8, and 12 db and widening the frequency band to 0.35 - 2 kHz. In total, every participant is asked twice about each of the 4 levels introduced. In every experiment order and period between questions are random and not known to the participants. The results of this experiment are presented in Figure 1. 54% out of 88 answers are correct. However 97.6% of the answers are only one level different from the correct one. Most participants (except 2) agree that the proposed approach is not bothersome and that they can use it to encode information they care about while still being able to enjoy the music.

4. CONCLUSIONS AND FUTURE WORK

This paper introduces an approach to encoding continuous information in personal music using frequency modulation. We explore what constitutes noticeable changes, how to encode curves and evaluate them in a traffic monitoring example with simulation of physical and mental load. For future work, we are working on a wearable system, using a smart phone and bone conductive head phones for easier use throughout the day. We are also looking into analyzing the music before encoding information, to make sure there are significant levels of a specific frequency band before dampening/boosting it.

5. ACKNOWLEDGEMENTS

We would like to thank Fujitsu Design Japan and all participants for their support and involvement.

6. REFERENCES

[1] Gaetano Borriello, Alan Liu, Tony Offer, Christopher Palistrant, and Richard Sharp. 2005. WALRUS: wireless acoustic location with room-level resolution using ultrasound. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services*. ACM, 191–203.

[2] Andreas Butz and Ralf Jung. 2005. Seamless user notification in ambient soundscapes. In *Proceedings of the 10th international conference on Intelligent user interfaces*. ACM, 320–322.

[3] Ádám Csapó and György Wersényi. 2013. Overview of auditory representations in human-machine interfaces. *ACM Computing Surveys (CSUR)* 46, 2 (2013), 19.

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

[5] Anil Madhavapeddy, David Scott, and Richard Sharp. 2003. Context-aware computing with sound. In *UbiComp 2003: Ubiquitous Computing*. Springer, 315–332.

[6] Elizabeth D Mynatt, Maribeth Back, Roy Want, Michael Baer, and Jason B Ellis. 1998. Designing audio aura. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 566–573.

[7] Aditya Shekhar Nittala, Xing-Dong Yang, Scott Bateman, Ehud Sharlin, and Saul Greenberg. 2015. PhoneEar: interactions for mobile devices that hear high-frequency sound-encoded data. In *Proceedings of the 7th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. ACM, 174–179.

[8] Spencer Russell, Gershon Dublon, and Joseph A Paradiso. 2016. HearThere: Networked Sensory Prosthetics Through Auditory Augmented Reality. In *Proceedings of the 7th Augmented Human International Conference 2016*. ACM, 20.