Mobile Interaction: Automatically Adapting Audio Output to Users and Contexts on Communication and Media Control Scenarios

Tiago Reis, Luís Carriço, and Carlos Duarte

LaSIGE, Faculdade de Ciências, Universidade de Lisboa
treis@lasige.di.fc.ul.pt, {lmc,cad}@di.fc.ul.pt

Abstract. This paper presents two prototypes designed in order to enable the automatic adjustment of audio output on mobile devices. One is directed to communication scenarios and the other to media control scenarios. The user centered methodology employed on the design of these prototypes involved 26 users and is also presented here. Once the prototypes were implemented, a usability study was conducted. This study involved 6 users that included our prototypes on their day-to-day lives during a two-week period. The results of the studies are presented and discussed on this paper, providing guidelines for the development of audio output adjustment algorithms and future manufacturing of mobile devices.

Keywords: Media Control, Communication, Automatic Volume Adjustments, Context Awareness, Hand-held Devices, User Centered Design, Contextual Evaluation.

1 Introduction

Nowadays, mobile devices are strongly integrated in peoples' lives. The ubiquitous nature of these devices enables humans to use them in an enormous variety of contexts, which are defined by various sets of dynamic characteristics (contextual variables) that heavily affect user interaction. However, the differences amongst users (e.g. preferences, capabilities) and the frequent context mutations, which occur during device utilization (e.g. moving from a silent to a noisy environment), usually result on users' adaptation to both contexts and interfaces available, and not the other way around. Most mobile user interfaces are unable to adapt effectively and automatically to the mutations of their utilization contexts, introducing difficulties on user interaction and, many times, inhibiting it [1, 2, 3]. Accordingly, it is necessary to explore new approaches for user interface design and development, aiming usability and accessibility improvements on mobile applications. To achieve this, applications must be constantly aware of their utilization contexts and respective mutations, naturally providing users with adequate interaction modalities, combining and configuring them according to the contexts in which the devices are used.

The work presented in this paper addresses the contextual adaptation of audio output on mobile devices, focusing communication and media control scenarios.

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2009, LNCS 5615, pp. 384–393, 2009.

Although there are solutions available for similar purposes, especially regarding communication scenarios, they consider only few, insufficient, dimensions of context: noise levels are used in order to adapt different aspects of audio output (e.g., ringtone volume, earphone volume) [4, 5, 6]. The personal dimension of context, which represents a non-trivial issue on user interface adaptation, is not addressed. Peoples' preferences and capabilities are not considered. Nevertheless, nowadays, the available technology enables the transparent gathering of the information needed for this purpose [7, 8] and its correct employment may significantly increase usability and accessibility, consequently, improving user experience by assuring interaction's adequateness to both utilization contexts and user needs. This motivated the user centered design, development, and contextual evaluation of two context-aware prototypes. These prototypes gather noise levels from the surrounding environment, adapting audio output according to scenarios, user needs and environmental noise. They present a proof of concept that can be achieved through the utilization of the technology available on most mobile devices, and can be considered and improved on the manufacturing of new devices in order to overcome the limitations found by the studies conducted and presented on this paper.

This paper starts by presenting and discussing the related work developed in this area. Following, it presents a strongly user centered design process that enabled the definition of a set of requirements, guidelines and design decisions that were considered during the prototypes' implementation. Afterwards, it presents the prototypes created, and their underlying volume adjustment algorithms. It details and discusses the contextual evaluation of these prototypes, emphasizing the limitations and advantages of the algorithms created for audio output adaptations, as well as their users' acceptance. Finally, the paper is concluded providing future work directions within this domain.

2 Related Work

From simple headphones and headsets with physical noise canceling [9] to headphones and headsets that include very complex algorithms of noise elimination [10, 11], different solutions have been proposed and created in order to reduce the impact of noise on audio interaction.

Firstly, it is important to emphasize that there are significant differences between noise canceling and noise-based automatic volume adjustments. While the first is of utmost importance in many scenarios (e.g. communicating inside a plane or close to a helicopter), it demands an aggressive noise elimination that is extremely complex and, consequently, extremely expensive, sometimes making the headphones themselves more pricey than regular mobile devices [9, 10, 11]. Accordingly, the latter presents a reliable and more affordable solution for contexts in which the environmental noise varies significantly, but not tremendously. Moreover, noise canceling solutions can become dangerous in several contexts. For instance, while a user is walking on the street, if he/she is completely inhibited from hearing the cars passing by, he/she can be injured.

The work presented on this paper focus the contextual-based adaptation of volume in communication and media control scenarios. Regarding communication scenarios, Chris Mitchell developed an application that is capable of monitoring noise levels on the surrounding environment, consequently adjusting the ringtone volume of a mobile phone [4]. Regarding media control scenarios, there are several car stereos that adjust the volume according to the estimated noise generated by the car. Regarding both scenarios, Apple recently published a new patent with the US Patent Office [12]. They claim to show a new technology feature that may be included on their future products. Their concept is similar to the ones proposed on this paper: automatic volume adaptation on communication and media control scenarios. However, they envision the inclusion of sound sensor (an extra microphone) in order to capture noise, while we used only the hardware available on every mobile device including communication and media control capabilities. Moreover, this paper focuses the user centered design and contextual evaluation of such concepts.

Finally, all the existent solutions consider noise as the only contextual variable affecting audio output preferences, which is, as we show on the following section, an incomplete assumption, especially in media control scenarios.

3 Early Design

This section is dedicated to the early design stages of the two prototypes created. The design processes followed a strongly user centered methodology, involving 26 non-impaired users: 16 male, 10 female, ages between 14 and 45 years old, familiar with mobile devices, especially mobile phones and mobile MP3 players. The users involved answered a questionnaire regarding important aspects related to the contextual adaptation of volume in both communication and media control scenarios. It was important to understand which context variables have an impact on users' voluntary volume modifications, how the automatic volume modifications should be performed, and how useful the users' believed the concepts proposed to be. The resultant information was carefully analyzed, culminating on a set of requirements, design decisions and guidelines, which were considered during the implementation of the prototypes.

3.1 Questionnaires

Firstly, concerning users' decisions regarding volume's modification on the considered scenarios, it was important to understand which contextual variables have an impact on these decisions. Furthermore, it was necessary to quantify the impact of each variable on the different scenarios, defining a level of importance for each one of them. Accordingly, for the different scenarios, users were asked to rate the impact of a set of contextual variables on a scale from 0 (null impact) to 100 (great impact).

The results (Fig. 1) indicate noise, with an impact of 100, as the only contextual variable affecting voluntary volume modifications on communication scenarios. Conversely, on media control scenarios, the variables influencing voluntary volume modifications are significantly more, and their impact considerably different. Noise remains the primary variable; however, its impact is reduced to 82. The task the user is engaged reveals an impact of 70 and interruptions generated by third parties an impact of 58. The remainder factors only apply to media control scenarios and are all related to specific characteristics of the media file being played (the song itself, album, artist,

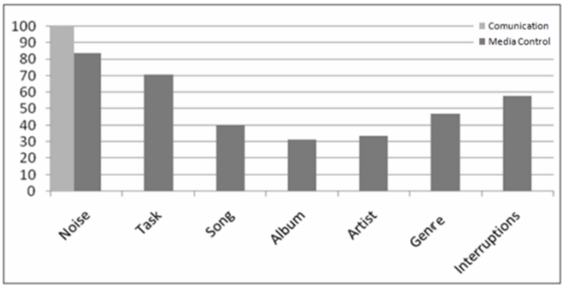


Fig. 1. User ratings: the impact of different contextual variables on voluntary volume modifications in both, communication and media control scenarios

and genre of music), presenting impacts between 30 and 45. Most users involved added that their emotional condition also influences voluntary volume modifications significantly, and can even influence the impact of the other context variables.

Secondly, it was fundamental to understand users' preferences regarding the automatic volume adaptation on the different scenarios considered. Two alternatives were implemented on a very simple application: gradual and direct volume adaptation. Fig. 2 a) depicts the differences between these alternatives. The gradual volume adaptation increases volume gradually (in approximately 0.8 seconds) from its current value (50 on the example provided) to the value suggested by the adaptation algorithm (80 on the example provided). On the other hand, the direct volume adaptation performs the same action in a quarter of that time (approximately 0.2 seconds).

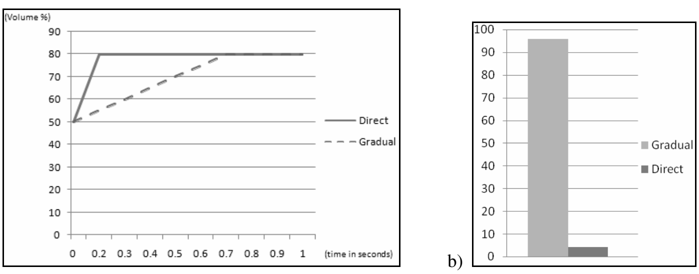
a)


Fig. 2. Example of gradual and direct volume adaptations (a) and user preferences (percentage) regarding the volume adaptation alternatives on both scenarios (b)

A small laboratory experiment was conducted with all the users involved, in order for them to experiment and be aware of the differences amongst these alternatives. The results of such experiment (Fig. 2 b)) show substantial differences regarding the appropriateness of the alternatives implemented. 96% of the users involved prefer the gradual volume adaptation on both communication and media control scenarios.

Finally, it was in the best interest of our team to understand how useful users believed the automatic volume adjustments to be. The results (Fig. 3) indicate a strong user acceptance for the concept proposed on both communication and media control scenarios. On communication scenarios 50% of the population involved rated the concept very useful, 30% rated it useful, 20% rated it slightly useful, and none of the persons involved considered the concept useless. On media control scenarios 58% of the population involved rated the concept very useful, 38% rated it useful, 4% rated it slightly useful, and none of the persons involved considered the concept useless.

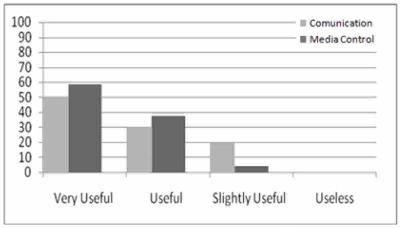


Fig. 3. Users' opinions about the usefulness of the concept proposed on communication and media control scenarios (percentage)

When asked about the discrepancies on the answers for each scenario, users emphasized the fact that they usually avoid noise on communication scenarios by moving to a silent place after answering a call and realizing that the noise is affecting the communication. Moreover, the users who rated the concepts useful and very useful were the ones which used mobile devices including communication and media control capabilities more often, and on contexts with significant environmental noise mutations (e.g. street, subway), while the remainder users used these devices mostly at work and at home.

4 The Prototypes

This section explains the noise monitoring process of the prototypes created, details the two automatic volume adjustment algorithms created, and the logging mechanisms implemented in order to ease the contextual evaluation of the prototypes. The mentioned prototypes are available for devices running Windows Mobile and were written in C#, using Microsoft's .Net Framework.

4.1 Noise Monitoring

Both prototypes created use the noise monitor available in [4]. This monitor gathers sample values from the device's microphone, consequently calculating loudness values in order to adjust the ring tone volume on a mobile phone (on a 0 to 5 scale). The measure used to calculate loudness is root-mean-square (RMS).

4.2 Algorithms for Automatic Volume Adjustments

The environmental noise was considered the primary context variable influencing users' decisions regarding volume modifications on both communication and media control scenarios. However, as described on section 3, there are other context variables that have a significant impact on these decisions. Moreover, users' hearing capabilities must also be considered. Accordingly, despite behaving differently, the two algorithms created take all these dimensions into account, sharing some principles.

For both algorithms, the noise spectrum varies from 0 to 127.5 [4] and the volume spectrum varies from 0 to 100. Moreover, both algorithms can be configured by defining 4 parameters that are accessible to the users (Fig. 4):

- Minimum: defines the minimum volume that can be set by the algorithm. This boundary is defined in order to avoid adaptations that are inconvenient for the users (e.g. setting the volume to low due to the absence of noise).
- Maximum: defines the maximum volume that can be set by the algorithm. This boundary is defined for the same reason as the above mentioned (e.g. setting the volume to high due to excessive noise on the environment).
- Sensibility: defines the coefficient dividing the noise spectrum into noise levels. For instance, if sensibility is defined to 3 and the noise spectrum's range is 127.5, the spectrum is divided in 43 noise levels.
- Volume Step: defines the increase or decrease of volume whenever the noise on the surrounding environment goes up or down one level, respectively.

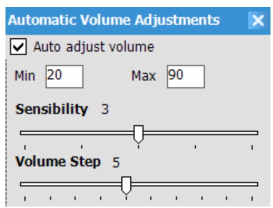


Fig. 4. Configuration screen presenting default values of the algorithms

The only contextual variable that is directly monitored by our prototypes is the environmental noise. Nevertheless, the remaining context variables are also considered (e.g. preferences, hearing capabilities, third party interruptions, etc.). These are indirectly expressed by the users whenever they perform a voluntary volume modification. For instance, if for any reason a user is not satisfied with the volume automatically set by the algorithm, his/her natural behavior would be to manually set the volume according to his/her preferences. When this happens, the algorithm registers a user preference, which is composed by a pair noise/volume, modifying the adaptation table accordingly and registering it on a XML file for posterior use. The developed algorithms behave differently on such situations and will be further explained:

Automatic Volume Adjustments on Communication Scenarios. This functionality is achieved through the utilization of a non-continuous preference based algorithm. The non-continuity derives from the constraints imposed by the scenario considered (the user is talking on the phone) and the device used to create the prototype, which includes only one microphone. Accordingly, the microphone used to communicate is the same one monitoring the environmental noise. The automatic volume adjustment is direct and effectuated based on the noise levels gathered immediately before the user answers the phone.

The preference base emerges from the last voluntary volume modification performed during a call. On the end of the call, a preference entry is registered and the adaptation table is modified according to that preference. This only happens on the end of the call because the noise monitoring is stopped in the meantime, while the user might move between contexts with different values of environmental noise.

Automatic Volume Adjustments on Media Control Scenarios. This functionality is achieved through the utilization of a continuous preference based algorithm. This algorithm was specifically designed for media control scenarios in which the users are wearing headphones. Accordingly, the sound produced by the media being played does not influence the noise monitoring process, enabling a continuous utilization of the device's microphone in order to monitor noise. Therefore, the automatic volume adjustments are applied continuously and gradually, while the user is controlling the media.

The preference base emerges from users' voluntary volume modifications. When these take place, the algorithm assumes that the volume set by the user represents his/her preferences for the noise captured at that moment, overriding his/her previous preference. Accordingly, the table noise/volume defined by the algorithm's sensibility and volume step is instantaneously modified in order for the noise registered to match the volume set, continuing the adaptation according to the modified table.

4.3 Logging Mechanisms

The logging mechanisms were implemented in order to ease the evaluation process of the prototypes. These mechanisms enabled middle term studies to be conducted, removing the need for direct monitoring. Both prototypes created include these mechanisms and are able to register the users' and the algorithms' behaviors on a XML file. Every user action is registered and associated with a contextual stamp, which includes time and noise information.

5 Contextual Evaluation

The prototypes created were evaluated through a strongly user centered procedure. The users selected to participate on this procedure were very familiar with mobile phones and media players. Moreover, there was a strong concern from our team on selecting users which used these devices on a broad variety of contexts (e.g. home, street, bus, subway, gym, etc.). These concerns emerged from the necessity of having a basis for comparison of our solutions with the existing technology, on several real contexts. Six users were involved: 3 male, 3 female, with ages between 18 and 35 years old. They used our prototypes on their day-to-day lives during a two week period. Accordingly, the logs gathered during the process represent the utilization of the prototypes on real contexts, under real, constantly mutating contextual constraints. In the end of the process, the users returned the utilization logs, which were carefully analyzed. Moreover, these users provided their feedback and opinions about the automatic volume adjustments, emphasizing situations where these were, and were not, satisfactory.

5.1 Contextual Evaluation on Communication Scenarios

Considering situations where the environmental noise decreases significantly during a call, all the users were slightly uncomfortable with the algorithm's behavior, reporting that the volume became too loud and they had to manually set it down. Such situations occurred mostly on the beginning of the study, when the users would still bond to their natural behavior, moving to a more silent place after answering a call. However, as the users continued to use the prototype, they have all changed their behavior regarding this issue, moving to a more silent place only in situations where they didn't feel comfortable discussing the topic of the call in front of other people.

Regarding situations where the environmental noise doesn't change significantly during a call and the topic of that call is not private, all the users were very satisfied with the algorithm.

On situations where the environmental noise increased significantly during a call, users were very uncomfortable with the algorithm's behavior. Such situations emerged mostly on public transportations (e.g. bus, subway). Users reported that the volume was too low, and that they had to manually modify it.

Finally, on situations where the environmental noise was extremely loud, users reported that the maximum volume was not enough to maintain the conversation, suggesting the creation of alerts that advise users not to answer calls in such situations.

5.2 Contextual Evaluation on Media Control Scenarios

Users reported several situations where they had to move between contexts with significantly different environmental noise values, without the need to manually modify the volume configuration. Such situations include going from home to the users' workplaces, having to walk on the street, ride public transportations, get in and out of different buildings, etc. The log analysis corroborates these reports, showing no manual volume adjustments during long periods of time (until 2 hours), characterized by very discrepant environmental noise values.

The only situations where users felt uncomfortable with the algorithm's behavior were situations where they were interrupted by third parties, engaging conversations. They explained that on such situations the volume would start to increase (due to the increasing environmental noise generated by the conversation), culminating on a manual volume modification or on users removing their headphones. These situations are also corroborated by the logs gathered, where in some situations of increasing environmental noise the users manually set the volume to mute, and then back to its previous value.

5.3 Discussion

The algorithm directed for communication scenarios was the one raising more questions. This occurred due to the non-continuity of the noise monitoring, implied by the hardware available on the device used, and the constraints of the scenario for which the prototype was developed (the only microphone available was being used to talk). However, the limitations of the algorithm, except the ones regarding the privacy of the topics being discussed during a call, could be overcome with the inclusion of another microphone on the mobile device. This microphone should be used in order to continuously capture noise, enabling continuous volume adjustments during the calls. Despite the limitations of the prototype created, users considered it better than their personal mobile phones, explaining that in the worst case scenarios their behavior was very similar to the one they had when using their phones: manually modifying the volume configuration.

The continuous noise monitoring of the algorithm directed to media control scenarios revealed an excellent user acceptance in most contexts. However, third party interruptions generated noise, consequently leading to an automatic increase of volume, which resulted on uncomfortable situations for the users. This problem can be solved by separating speech from environmental noise as in [10] (not only the speech of the user using the device but also of the third parties engaged in conversations with this user). Nonetheless, such solution would clearly increase the complexity of the algorithms and the amount of hardware used.

Overall, the studies conducted revealed that noise monitoring should be performed continuously, in order to enable accurate and continuous volume adjustments. There was a strong user acceptance of both prototypes created, and despite being able to modify all the algorithms' parameters, the users involved on the evaluation procedures only personalized the maximum and minimum volumes, being very satisfied regarding the default sensibility and volume step of the algorithms. The utilization logs gathered also corroborate these affirmations.

6 Conclusion and Future Work

In this paper we described the user centered design of two context-aware prototypes directed for communication and media control scenarios. These prototypes were built on top of a regular mobile device, including both communication and media control capabilities. They are capable of adjusting volume according to different aspects of the context in which they are being used, monitoring noise directly through the microphone, and considering user preferences and capabilities, which are expressed indirectly on the users' voluntary volume modifications. Issues regarding the amount of contextual information directly monitored by the algorithms responsible for the volume modifications, are left untied and will be studied on our future work. The contextual evaluation of the prototypes revealed a strong user acceptance of the concept proposed, especially on media control scenarios. However, the studies conducted also point some issues that could not be overcome using only the hardware available on most mobile devices nowadays. Accordingly, the study also provides important information to be considered on the manufacturing of future mobile devices.

Acknowledgments. This work was supported by LaSIGE and FCT through the Multiannual Funding Programme and individual scholarships SFRH/BD/44433/2008.

References

1. Barnard, L., Yi, J.S., Jacko, J., Sears, A.: Capturing the effects of context on human performance in mobile computing systems. Personal and Ubiquitous Computing 11(2), 81–96 (2007)

2. Reis, T., Sá, M., Carriço, L.: Multimodal Interaction: Real Context Studies on Mobile Digital Artefacts. In: HAID 2008. LNCS, vol. 5270, pp. 60–69. Springer, Heidelberg (2008)
3. Schmidt, A., Aidoo, K.A., Takaluoma, A., Tuomela, U., van Laerhoven, K., Van de Velde, W.: Advanced interaction in context. In: Gellersen, H.-W. (ed.) HUC 1999. LNCS, vol. 1707, pp. 89–101. Springer, Heidelberg (1999)
4. Mitchell, C.: Mobile Apps: Adjust Your Ring Volume For Ambient Noise. MSDN Magazine (2008), http://msdn.microsoft.com/en-us/magazine/cc163341.aspx
5. Kumar, Larsen, Infotelimcithed, T.: Smart Volume Tuner for Cellular Phones. IEEE Wireless Communications (2004)
6. US Patent 7023984 - Automatic volume adjustment of voice transmitted over a communication device, http://www.patentstorm.us/patents/7023984/description.html
7. Baldauf, M., Dustdar, S., Rosenberg, F.: A survey on context-aware systems. International Journal of Ad Hoc and Ubiquitous Computing 2(4), 263–277 (2007)
8. Chen, G., Kotz, D.: A survey of context-aware mobile computing. Technical Report TR2000–381, Dartmouth College, Department of Computer Science (2000)
9. Review on noise canceling headphones, http://www.seatguru.com/articles/noise-canceling_review.php
10. The Jawbone Headset, http://eu.jawbone.com/epages/Jawbone.sf
11. The Boom Headset, http://www.theboom.com/
12. Patent for automatic noise-based volume adjustments, http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearchbool.html&r=1&f=G&l=50&co1=AND&d=PG01&s1=20090022329.PGNR.&OS=DN/20090022329RS=DN/20090022329