

Automatic Volume Adjustments for Mobile Media Players

Tiago Reis

Luís Carriço

Carlos Duarte

LASIGE e Departamento de Informática, Universidade de Lisboa
Edifício C6, Campo Grande, 1749-016, Lisboa, Portugal
treis@lasige.di.fc.ul.pt lmc@di.fc.ul.pt cad@di.fc.ul.pt

Abstract

The automatic adjustment of mobile user interfaces to the contexts in which they are used presents a possible solution for the interaction difficulties introduced by the constantly mutating characteristics of such contexts. Here we present a prototype of a portable media player. This player automatically adjusts its volume according to the users' preferences and the noise of the surrounding environment. We focus on the details of the volume adjustment algorithm and its – laboratorial and contextual - user centered evaluations. We detail the problems found on each phase of the evaluation procedure, and present a set of Wizard of Oz experiments exploring possible solutions for those problems.

Keywords

Mobile HCI, Automatic Volume Adjustments, Context Awareness, 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, defined by various sets of dynamic characteristics (contextual variables) that heavily affect user interaction. However, the differences amongst users and the frequent context mutations, which occur during device utilization, often 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 interaction difficulties [Barnard07, Reis08, Schmidt99]. Accordingly, it is necessary to explore new approaches for user interface design, aiming at usability improvements. To achieve this, applications must be constantly aware of the relevant aspects of their utilization contexts and respective mutations, naturally providing their users with adequate interaction modalities, combining and configuring them accordingly.

Audio-based mobile interaction processes are significantly affected by noise, a contextual aspect that varies considerably on mobile scenarios. Amongst these processes, the use of portable media players is a reality that remounts to the late 70's when the Sony Walkman became available to the public. Nowadays, these players evolved and became an integrated component of several mobile devices, which serve other purposes as well. The volume controller on portable media players also suffered several changes since the late 70's volume wheel. This component evolved to physical buttons and, later, to virtual controls presented on touch screens. The substantial utilization of touch screens eradicated the tactile awareness and feedback provided by mechanical controls, making it often impossible to control such devices without having

to look at them. Nevertheless, some headphones include small controllers (e.g.: volume, forward, etc.) placed on their wire in order to reduce the users' need to look at the device in order to interact. Despite the evolution of portable media players and respective controls, noise still affects interaction considerably, and the mutations of this context variable often result on manual volume modifications. The reduction or eradication of these modifications may introduce significant improvements to the portable media players' user experience.

From simple headphones and headsets with physical noise canceling [Review10] to headphones and headsets that include very complex algorithms of noise elimination [JawBone10, TheBoom10], different solutions have been proposed and created in order to reduce the impact of noise on audio interaction. Noise canceling and elimination are paramount in critical scenarios, such as: communicating on a landing strip. However, this solution demands an aggressive noise elimination that is extremely complex and, consequently, extremely expensive, sometimes making the headphones themselves more pricey than regular mobile devices [Review10, Jawbone10, TheBoom10]. Moreover, noise canceling solutions may become dangerous in every-day mobile contexts. For instance, while a user is walking on the street, if (s)he is completely inhibited from hearing the cars passing by, (s)he can be injured. Accordingly, noise-based automatic volume adjustments present a more affordable and safe solution for contexts in which the environmental noise varies significantly enough to force users to manually modify the volume of their portable media players. Moreover, several mobile phones function also as portable media players, and all of them include a microphone that can be continuously used in order to monitor the environmental noise.

The work presented in this paper addresses the contextual adaptation of volume on portable media players. It is materialized on a functional prototype and a set of user-centered studies. The prototype is a noise-aware media player that attempts to reduce the manual volume modifications generated by noise, gathering noise levels from the surrounding environment, and automatically adjusting the volume accordingly. These adjustments also consider the users' preferences, which are transparently identified whenever a manual volume adjustment is performed. The studies conducted focused on the details of the automatic volume adjustment algorithm and its employment on real life scenarios. These studies revealed a strong user acceptance of both the concept proposed and its implementation. However, for specific situations where the users engage in conversations, our implementation of the concept generated slightly uncomfortable reactions on the users. Building on top of such results, we also present a set of Wizard of Oz studies, conducted in order to evaluate a set of possible solutions for this problem.

2. THE PROTOTYPE

The prototype created includes a noise monitor and a continuous preference-based volume adjustment algorithm. Logging mechanisms were also implemented in order to ease the contextual evaluation. The prototype is an application for devices running Windows Mobile 5. The noise monitor employed is the one available in [Mitchell07], which gathers sample values from the device's microphone and calculates loudness values in a spectrum that varies from 0 to 127.5 root-mean-squares (RMS). The algorithm responsible for the automatic volume adjustments considers noise as the primary context variable. It 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 the continuous monitoring of the environmental noise. The automatic volume adjustments are applied continuously within a volume spectrum that may vary from 0 to 100. The algorithm is configurable, 4 parameters are accessible to the users (Figure 1): minimum, maximum, sensibility, and volume step. The first two define the minimum and maximum volumes that can be set by the algorithm. These boundaries are 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). The sensibility parameter ranges from 1 to 5, defining **when** the algorithm reacts. This value is employed on the definition of the coefficient dividing the noise spectrum into a set of discrete noise levels. Each of these levels defines a line on the adaptation table used by the algorithm and has a correspondent volume level that is defined by the volume step. The volume step parameter ranges from 1 to 10. This value defines **how** the algorithm reacts, corresponding to the amount that the volume is automatically increased/decreased, whenever the noise of the surrounding environment goes up/down one level. Accordingly, the automatic volume adjustment algorithm behaves ac-

ording to a dynamic table of pairs noise/volume. The number of lines on this table is defined by sensibility, and the correspondent volume levels are defined by the volume step, varying between the boundaries defined by the minimum and maximum values.

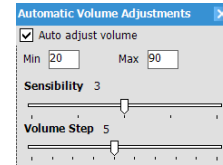


Figure 1. Configuration screen of the algorithm.

This algorithm is also preference-based. The user's preferences are indirectly expressed whenever (s)he performs a manual volume modification. When this happens, the algorithm assumes that the volume set represents the user's preference for the noise captured at that moment, overriding his/her previous preference, and instantaneously modifying the table noise/volume. From that moment on, the adaptation continues according to the modified table, which is saved on a XML file for posterior use.

The logging mechanisms register the users' and algorithm behaviors on XML files. Every action is registered and associated with a contextual stamp, which includes time and noise-related information.

3. EVALUATION

The evaluation process comprised two strongly user centered procedures (laboratorial and contextual). The main goals of these procedures were to understand: (1) if the automatic volume adjustments reduce the number of manual volume adjustments performed in order to mitigate noise disturbances; (2) if the configuration parameters were adequate; and, (3) if the users believed the concept proposed to be useful.

3.1 Laboratorial Evaluation

The laboratorial evaluation involved four users: 3 male, 1 female, with ages between 20 and 29 years old. The experiment conducted in order to fulfill the first goal comprised two phases. On the beginning of both phases the users were asked to set the volume to a comfortable level. On the first phase, the users listened to a song while an audio stream was played in order to simulate the background noise. This stream contained sounds that are typically found on real mobile interaction scenarios (e.g.: birds, cars driving by, trains, noise generated by people, etc.). During this phase the users performed in average 5 manual volume adjustments in order to mitigate the noise disturbances. On the second phase, the automatic volume adjustments were activated with the default values, significantly reducing the number of manual volume adjustments. The users listened to the same song while the same audio stream was played, averaging less than 1 manual volume adjustment. The manual volume adjustments performed on this phase emerged strictly because 2 of the

users were not comfortable with the default maximum value of the algorithm.

Following, in order to fulfill the second goal, the users were presented to the configuration screen of the automatic volume adjustment. All users considered the minimum and maximum volume parameters of utmost importance, emphasizing that they allow an important level of personalization. Regarding the sensibility and volume step, users were confused with the existence of these two parameters. All of them were expecting just one parameter. After experimenting different configurations for these two parameters, all the users reported that high sensibility values should correspond to low volume steps, and low sensibility values should correspond to higher volume steps. The users reported to be very comfortable with the default values for the sensibility and volume step, considering them the most adequate combination of all the ones they have experimented. Moreover, they stated that this feature should either be available through one control only, or not be available at all.

Finally, when asked about the usefulness of the automatic volume adjustments all the users considered them very useful, and were very enthusiastic about the possibility of integrating our prototype in their every-day lives.

3.2 Contextual Evaluation

The evaluation procedure conducted in real life scenarios involved six users: 3 male, 3 female, with ages between 18 and 35 years old. The users involved included our prototype on their day-to-day lives during a two week period summing a total of 40 utilizations on real contexts. In the end of the process, the users participated on an informal interview where they provided their feedback and opinions about the automatic volume adjustments, emphasizing situations where these were, and were not, satisfactory.

All the users reported that the number of manual volume adjustments performed in order to mitigate noise disturbances was significantly reduced when compared to their own portable media players. 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 users reported other contextual aspects that generate manual volume adjustments. Amongst these, they emphasized their emotional states, the song being played, and the primary task they were involved in (e.g.: reading a book) to affect their preferences regarding the volume set. Nevertheless, all the users stated that these situations were well covered by the prototype behavior, namely by the preference base of the algorithm.

There was one particular type of situation where users felt uncomfortable with the algorithm's behavior. Whenever the users were interrupted by people, engaging in conversation, the volume would increase due to the increasing environmental noise generated by the conversation. These situations often culminated on a manual volume modification or on users removing their headphones. The users did not consider this to be a major problem. They all reported that the improvements achieved on the majority of the scenarios justified such inconvenience. These uncomfortable situations are also patent on the logs gathered, where in some situations of increasing environmental noise the users manually set the volume to mute.

Despite being able to modify all the algorithms' parameters, the users only personalized the maximum and minimum volumes. Again, they reported some confusion regarding the existence of the two remaining parameters. Nevertheless, all the users reported smooth and accurate automatic volume adjustments, showing a high satisfaction regarding the default values of such parameters. The logs gathered corroborate these statements, showing modifications only for the maximum and minimum parameters.

Regarding the usefulness of the automatic volume adjustments, all the users acknowledged this functionality to be very useful. All of them reported improvements of the interaction process, namely a significant reduction on the number of interactions with the device, and, consequently, a better user experience. Moreover, when the users started using their own portable media players again they stated that once familiarized with the automatic volume adjustments having to go back to performing them manually became significantly more "annoying".

4. WIZARD OF OZ EXPERIMENTS

Despite the strong user acceptance of our prototype, its contextual evaluation procedure revealed uncomfortable situations for the users. These situations were generated whenever the users engaged in conversations with other people. In order to explore automatic solutions for this problem, a series of Wizard of Oz experiments was conducted. During the process, both the volume and the automatic volume adjustment algorithm were remotely controlled in order to simulate the particularities of each experiment.

The first experiment considered the perfect identification of the beginning and end of conversations involving the user. The experiment focused the behavior the prototype should assume during the conversations the users engage in. Three alternatives were experimented: 1) pause; 2) muting the volume; and, 3) lowering the volume to the minimum value, and disabling the volume adjustment algorithm. Four users preferred the first solution, one preferred the second, and one voted for the third, suggesting that such behavior should be configurable since its selection is strongly affected by the users' preferences. Regardless of the behavior assumed by the prototype, the perfect identification of the beginning and end

of conversations involving the user revealed a strong user acceptance; all the users reported this to be the perfect solution for the problem.

However, the fully reliable identification of such situations is unfeasible if we only consider the technology available in most mobile devices nowadays (one microphone and limited processing capabilities). Accordingly, the following two sets of experiments considered some problems that could emerge from the employment of solutions based on voice activity detection, namely the existence of: 1) false positives (e.g.: identifying a conversation that involves the user when the user is not involved in the conversation); and, 2) false negatives (e.g.: not identifying a conversation involving the user when the user is actually involved in a conversation).

The experiments considering only the existence of false positives on the identification of conversations where the users participate showed that unless the mechanism responsible for this identification is fully (100%) reliable the uncomfortable situations increase when compared to the actual implementation of our prototype. Accordingly, the existence of false positives was strongly rejected by all the users. The experiments considering only the existence of false negatives, presented improvements when compared to the actual implementation of the prototype. The users reported that the identification of some conversations and the consequent automatic behavior improved the interaction process, reducing the manual volume adjustments. Such reduction is directly proportional to the reliability of the mechanism responsible for the identification of conversations where the users participate. However, the existence of false negatives leaves us with the same problem found during the contextual evaluation, implying the creation of a mechanism to overcome the interaction difficulties they introduce.

Accordingly, we conducted a new set of experiments focusing the manual control of the situations where the users engage in conversations. Different solutions were experimented for the activation/deactivation of the behavior assumed by the prototype during a conversation, namely: 1) the use of a button available on the wire of the headphones; 2) the use of a button on the device; 3) the use of gestures on the device's screen; and, 4) the use of voice commands. After trying all the solutions proposed, users rated them according to their usability and compared them to the solutions adopted during the contextual evaluation. The use of voice commands was considered very poor regarding its usability. The users reported to feel very uncomfortable talking to the device, especially in public places where there are more people. The use of gestures on the device screen and the use of a button on the device were considered poor regarding their usability. The users stated that solutions considering the direct manipulation of the device, which is often on their pockets or bags, are not adequate. The control through a button available on the headphones' wire was considered the most usable solution. When asked to compare the solutions proposed with the ones adopted during the contex-

tual evaluation, all the users still considered the use of a button available on the headphones' wire as the most adequate solution. However, they all stated that the second most adequate solution would be the manual control of the volume through a volume controller placed on the headphones' wire. Finally, the users stated that removing the headphones is a better solution than voice commands.

Overall, the above presented experiments enabled us to conclude that a fully reliable solution for the identification of situations where the users engage conversations would solve the problem found during the contextual evaluation procedure. However, this is currently unfeasible from a technological point of view. Accordingly, the manual control of such situations presents the most adequate solution for the problem.

5. CONCLUSION AND FUTURE WORK

In this paper we described a media player that automatically adjusts its volume according to the environmental noise and to the users' preferences. The prototype presented was evaluated through two strongly user centered procedures. These procedures revealed a strong user acceptance of the prototype, pointing some issues that should be improved. Following, a set of Wizard of Oz experiments was conducted in order to explore a set of possible solutions for these issues. Our future work comprises the development of plug-ins for mobile media players. These shall enable automatic noise- and preference-based volume adjustments, employing the knowledge assembled on the experiments here presented.

6. ACKNOWLEDGEMENTS

LaSIGE and FCT. Multiannual Funding Programme and individual scholarships SFRH/BD/44433/2008.

7. REFERENCES

- [Barnard07] Barnard, L., Yi, J S., Jacko, J., Sears, A.: Capturing the effects of context on human performance in mobile computing systems. In *Personal and Ubiquitous Computing*, Vol. 11, 2, pp. 81 - 96.
- [Reis08] Reis, T., Sá, M., Carriço, L.: Multimodal Interaction: Real Context Studies on Mobile Digital Artifacts. In *Procs. of the 3rd HAID*.
- [Schmidt99] Schmidt, A., Aidoo, K.A., Takaluoma, A., Tuomela, U., van-Laerhoven, K., and Van-de-Velde, W.: Advanced interaction in context. In *Procs. of the 1st HUC*, pp. 89-101.
- [Review10] Review on noise canceling headphones. Available at: http://www.seatguru.com/articles/noise-canceling_review.php
- [JawBone10] The Jawbone Headset. URL: <http://eu.jawbone.com/pages/Jawbone.sf>
- [TheBoom10] Boom <http://www.theboom.com/>
- [Mitchell07] Mitchell, C.: Mobile Apps: Adjust Your Ring Volume For Ambient Noise. In *MSDN Magazine*.