Assessing Feedback for Indirect Shared Interaction with Public Displays

Jorge C. S. Cardoso[1], Rui José[2]

[1] Escola das Artes / CITAR – Univ. Católica Portuguesa, 4169-005 Porto, Portugal
jorgecardoso@ieee.org
[2] DSI – Univ. do Minho, Campus de Azurém, 4800-058 Guimarães, Portugal
rui@dsi.uminho.pt

Abstract. Interaction is a key element in turning public displays into a platform for social interaction, making them more engaging and valuable. However, interactive features are still rare in public displays, due to the lack of generalised abstractions for incorporating interactivity. In our work, we explore to what extent the concept of interaction widget, which was so successful on desktop computers, could also be used as an abstraction for remote, shared interaction with public displays. A particular challenge is presenting input feedback in this shared, multi-user, and indirect interaction setting. In this paper, we present a study on the feedback mechanisms of these widgets, to determine if users are able to understand the results of interactions in single and multi-user settings. We have evaluated three feedback mechanisms and the results indicate that the general mechanism provides an appropriate sense of what is happening and could in fact provide general awareness of the interaction alternatives and current status, even in multi-user scenarios.

Keywords: interactive public displays, widgets, remote interaction

1 Introduction

Public digital displays are becoming increasingly ubiquitous artefacts in public and semi-public spaces. Interactive public displays are particularly well suited for social interactions, because they are part of a share physical space and naturally provide a shared digital medium that anyone can see and use. Most of them, however, do not support any type of interactive feature, despite interaction being clearly regarded as a key element in making them more engaging and valuable. A key reason behind this apparent paradox is the lack of generalised abstractions for supporting interaction, which means that each new application developer will need to create his or her own approach for dealing with a particular interaction objective using a particular interaction modality. The consequence is too much development work outside the core application functionality to support even basic forms of interaction and an obvious waste of development effort, potentially leading to poor designs. This is also a problem for users, as they need to deal with inconsistent interaction models across different displays, representing a major obstacle to the emergence of any expectations and practices regarding interaction with public displays. Given that interaction is considerably more limited than it is in desktop environments, and thus there are not many interface exploration possibilities, it is even more important to be able to convey the interaction affordances of the display in a way that people can easily understand.

Early desktop computer programmers had to make a similar effort to support their interaction with users. This was quickly recognised as a problem and addressed with the emergence of reusable high-level interaction abstractions, including widgets, which provided consistent interaction experiences to users and shielded application developers from low-level interaction details [1].

In our work, we are exploring to what extent the concept of interaction widget, which was so successful on the desktop, could also be used as an abstraction for some forms of interaction in public displays. While this may eventually lead to totally new widgets, we started by investigating to what extent some well-known desktop widgets could migrate to this new usage context. Well-known widgets are particularly attractive as an entry strategy because people are already familiar with them and may easily recognise them as interaction opportunities.

We are considering specifically the case in which multiple people may be concurrently interacting with a shared display using remote devices, such as mobile phones or portable computers. Multiple interaction techniques may be used, e.g. Bluetooth or SMS, and a key design goal is that widgets should be as much as possible independent from specific input channels. Creating widgets for this type of indirect interaction model raises two fundamental issues that are not present in traditional widgets. The first is that the direct manipulation paradigm does not apply because people are not acting directly on the widget. Feedback, or at least the type of feedback that is normally associated with widgets, is either non-existent, or would never be as immediate as the expectations that people have when interacting with traditional widgets under direct manipulation paradigms. A second issue is that multiple people may be acting concurrently on the same widget and therefore the type of feedback given may need to indicate to whom, from the potentially multiple people interacting at a particular moment, a specific feedback is directed to. The main implication of these issues is the need to create a specific feedback model that enables people to understand the effect of their actions and possibly the actions of others.

In this paper, we focused on the feedback for these widgets and we evaluated three different mechanisms for presenting visual feedback about input events directly on the public display. Results indicate that users are generally able to understand the feedback to their input and that of other users, giving us confidence that the proposed feedback model is appropriate and can be further explored.

2 Related Work

Widgets have been created for many specific uses [2, 3], but none for general-purpose public displays. For example, Rohs [4] has implemented a set of widgets for visual marker-based interaction that allows users to activate actions or select options encoded in a visual marker and send it via SMS (using a custom mobile application). In this case the visual marker encodes the type (menu, radio or check button list, sliders, etc.) and layout (vertical or horizontal menu, number of options, etc.) of the widget, so that the mobile phone application can immediately superimpose graphical information about the currently selected item or value. However, these widgets are limited to camera-equipped mobile phones and were not developed with multi-user interaction in mind. This approach requires users to be relatively close to the display in order to be able to use the mobile phone's camera with enough quality (or the use of larger visual markers); also, system feedback is shown only on the personal mobile phone and not on the public display limiting other user's understanding of what is happening in the display application.

Dearman & Truong [5] developed Bluetone: a widget that is activated through dualtone multi-frequency (DTMF) over Bluetooth. Users interact with an application by changing the Bluetooth name of their device to a system command, wait for the display to pair with the user's phone as an audio gateway, and then pressing the keys on the keypad of their phone. Bluetone supports several users, being limited only by the Bluetooth protocol (7 connections). This widget is limited to the DTMF interaction mechanism, and has been developed for an environment where a single application executes at a time; graphically, it consists of a single widget that encapsulates all the interactive features of the application. Also, Bluetone does not directly address how feedback is presented in a multi-user setting.

3 PuReWidgets

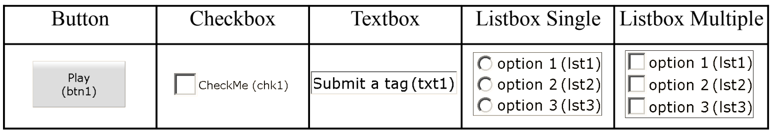
PuReWidgets is an initial implementation of a widget toolkit for public display applications. PuReWidgets keeps the visual appearance and basic affordances of their desktop counterparts, but adapts them to address the specificities of public display interaction. Like traditional WIMP widgets, these widgets are easily identifiable as possible actions to be performed on the public display and provide applications and users with a single abstraction for multiple input mechanisms. Unlike traditional widgets, they are not based on direct manipulation and multiple people can use them simultaneously.

The association between interaction events and specific widgets is achieved through a generic referencing scheme, which consists of automatically generating unique textual references whenever an application adds an instance of a widget to the interface. This reference is graphically shown on the widget instance (e.g., "btn1" in the button of Fig. 1) and it serves as an addressing mechanism for users and for the display system. Whenever a user enters the reference code in an input command, the display system routes the command to the appropriate widget instance. This referencing scheme is common for many SMS services, so most users will be familiar with it. Although, in this study, we focused on SMS as the input mechanism, the referencing scheme could be used in many other text input mechanisms such as Bluetooth naming, email, instant messaging but also by mechanisms such as visual markers or even custom mobile applications.

In our initial set, we included five of the most common input controls in traditional

GUI interfaces: button, checkbox, textbox and two variants of a listbox. Figure 1 shows the initial widget set. The reference code is shown between brackets in all widgets. Widgets with several options, such as listboxes, have several reference codes assigned to them so that users may address each of the available options. A widget may be enabled or disabled. Widgets should generally be enabled, but applications may choose to disabled a widget temporarily, for example, to signal that a particular operation is not available at the moment. Disabled widgets use the traditional "greyed out" style.

Fig. 1. Initial widget set showing the enabled state.



The PuReWidgets feedback model allows multiple users to interact concurrently, while being able to correctly display feedback to each user's input in an unambiguous manner. Feedback is shown directly on the public display, next to the widget that received the input. The user identification (taken from the input channel, e.g. phone number) is included in the feedback information, masked in a way that enables the sender to recognise his input in the feedback, but other users will only be able to distinguish different users, not to identify them. This shared feedback on the public display, next to the widget (or inside it), instead of, for example, directing it to the user's mobile device, is independent of the input mechanism: it does not assume that a return channel is available. It also provides a simple awareness mechanism that conveys information about other people's interactions. This should serve to entice more people to interact [6] and also to enable viewers to make sense of the display behaviour.

Feedback is seen here as system level feedback, i.e., it translates an acknowledgement that a particular input has been received and how it has been interpreted, although not necessarily accepted. Input can either be accepted by the widget, or rejected, but it always generates feedback to the user. Accepted input is passed on to the application, translated into an application event, while rejected input does not result in any application event. Input can be rejected for two reasons: the widget could be disabled; or the user made an error in a parameter needed by the widget (for example, forgetting to enter a string in the texbox widget).

3.1 Usage scenario

The initial widget set provides a large range of possible interactive features for public displays. Although their exact graphical appearance may change considerable, their basic affordances can be used for a variety of situations. Buttons for "liking" items on social or even regular webpages are ubiquitous now and can also have a big impact on public display content. Textboxes can be used to signal the possibility of entering comments, perhaps filtered before they show up next to a content item on a public display. Listboxes are the building blocks for polls and questionnaires that can be used by public displays to collect users' preferences and generating discussion about current national or local hot topics. Checkboxes are an easy way to let users rapidly configure a public profile by checking or un-checking profile items.

The following scenario gives an example of how some of these widgets might be useful for public display interaction:

Sophia has just entered her university's main hall and is looking for a place to sit down and wait for her colleagues -- the first class is only due in 15 minutes. As she sits down, she looks at the large display across the hall noticing it is showing a list of local news related to her school. One of the entries catches her eye -- it's about Adam, a colleague on the robotics class, which has won the national robot dancing contest. Sofia notices a button next to the news entry's header and recognises it from her favourite social website: is a ``like'' button with three letters underneath. The instructions on the top of the display tell her how to interact so she fetches her mobile phone and sends a text message to the number on the instructions. A few seconds later, text pops up on the button: something resembling a phone number with some digits erased appears, and she recognises it as her own. She knows her ``like'' will increase the news visibility on the school's website and on the display. Adam deserves it! At the same time, a group of students passes by, and one of the girls points to the display and makes some remarks about it. The button's animation triggered by Sophia's interaction apparently caught the girl's attention.

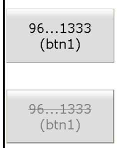
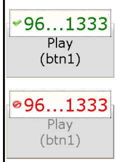
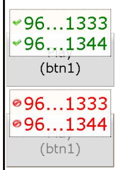
Sophia also notices a vertical bar on the left of the display, showing a list with the last people that interacted with the display. She recognises her phone number as the last entry, and because the display does not have much information about her, there is also a textbox that allows her to enter a nickname. Sophia hears someone calling her name. Looking away from the display she sees one of her classmates signalling her. It's time for classes; filling in the nickname will have to wait.

4 User Study

This study's main purpose was to allow us to have an initial assessment of the feedback model for PuReWidgets and possible implementations. To assess the effectiveness of the feedback, we evaluated three alternative implementations: Internal, External, and External-Cumulative (described in Table 1). The study was focused on the following research questions:

- Will users be able to identify and understand the feedback to their input?
- Will users be able to identify and understand feedback directed at other users?
- Will there be a difference in error rate between the several feedback mechanisms?
- What will be the preferences of users regarding the feedback mechanisms?

Table 1. The three implementations of the feedback model. The right column shows a button with feedback to an accepted input (top), and with feedback to a rejected input (bottom).

Internal feedback (Int) uses the widget's internal text components and temporarily changes them to show input feedback. Accepted input is shown in the same style as the label. In the case of the listbox, feedback is shown directly in the label of the selected option. Feedback to rejected input is formatted with a strike-through text style and grey colour. This mechanism is the one that requires fewer changes to the visual appearance of the widget because it simply reuses, temporarily, a text component of the widget.	
External feedback (Ext) uses a pop-up panel that is displayed next to the widget (by default, over the widget, slightly raised). In this case, since it is not bound to the existing components of the widget, feedback can use a richer visual style. We use a green "check mark" icon and green text to indicate an accepted input. To indicate rejected input, the style of the feedback is changed to show a red "prohibited" icon and red text.	
External-Cumulative (E-C) is an extension to the external feedback mechanism. In order to facilitate the comparison of multiple input feedbacks, the pop-up panel maintains the previous feedback texts (within a pre-configured period of time) and is incremented with the current feedback text. The style for the accepted and rejected input is the same as the External mechanism. When there is only a single input, the External-Cumulative mechanism behaves in the same way as the External mechanism.	

4.1 Procedure and Tasks

The study was structured in three parts: a first part in which participants were asked to interact with widgets in a single user setting, a second part in which participants were asked to interact with widgets in a multiple user setting, and finally a questionnaire in which users were asked to indicate the best and worst feedback mechanism for each widget. Participants were told about the objectives of the study and were given minimal information about how to use the widgets: they were told that to activate a widget they had to send an SMS message with the reference code that was shown inside the widget. Participants were divided into three separate groups according to the feedback they would be subjected to: Internal, External, and External-Cumulative. For part 1, the group External-Cumulative was incorporated into group External because both mechanisms behave the same way when only feedback to one input is shown. This meant that part 1 had two groups, and parts 2 and 3 had three groups.

Participants and Apparatus We recruited 24 participants (all daily computer users) from the university campus (8 female, 16 male; ages from 22 to 41). The experiment was done in a room equipped with a 23 inches LCD display. Participants sat at about two meters away from the display and were given a mobile phone to use during the experiment. Interaction was done via SMS.

Part 1 – Single Input For part 1 of the study, participants were presented with a sequence of five screens with one widget on each screen. Each screen depicted a simple mock-up application to give some context to the interaction, with a single widget placed in it. The application itself did not respond to any interaction, only the widget reacted to the input. Each screen presented a task to be completed by the participant. The five screens were presented in a random order. In two randomly selected screens, the widget was disabled (i.e., any input would fail). Participants were told to complete the presented task (they could perform several trials) for each screen and, after receiving the widget's feedback, were asked: "Did you successfully activate the widget?" We recorded the responses (yes, no, don't know) and number of trials that each participant took before answering. We recorded an error for each wrong response (a "don't know" counted as wrong) and for each extra trial needed.

Part 2 – Multiple Input In this part of the study we simulated the case were several users would interact with the same widget at the same time. We chose to simulate the interactions of other users instead of performing a multi-user experiment, in order to have full control over the interaction sequence (guaranteeing the occurrence of collisions between the input of multiple users). Participants were presented with the same five screens of part 1 (in a randomised order) and were told that there would be other users interacting at the same time as they were, but not necessarily making the same choices. All the screens were configured so that the widget would become automatically disabled after receiving the first input. Again, participants were told to complete the presented task (they could perform several trials) for each screen. After each screen (after the user has issued the input and seen all feedback), participants were asked: "Did you successfully activate the widget?", "How many other people were interacting with the display?" and "How many of those other people successfully activate the widget?"

Part 3 - Subjective preference In the final part of the study we showed participants the three alternative ways to display feedback for all the widget types. We asked participants to observe the different feedback mechanisms for as long as they needed and then to indicate the best and the worst mechanism for each widget type. We recorded the participants' answers, allowing them the choice of not selecting any as the best or worst.

5 Results

For part 1 and 2 of the study, we calculated the total number of errors (wrong answers and extra trials) made by each participant over the five screens in each part (shown in the boxplots of Fig. 2). For part 1, a right-tailed t-test was performed to test the hypothesis that the average error is higher in the Int condition than in the Ext condition. The test was significant ($t(22) = 5.87$, $p < 0.0001$), suggesting that users have more difficulty with the Internal mechanism.

For part 2, we did a pairwise comparison of the three conditions using Welch's Two Sample t-test and then applied Bonferroni's correction to the p-values. The result was statistically significant for the Int vs Ext ($p = 0.0134$) and Int vs E-C ($p = 0.0076$) tests and nonsignificant for the Ext vs E-C ($p = 0.6982$) test. As in part 1, this suggests that the participants have significantly more difficulty with the Internal mechanism than with the External or External-Cumulative ones.

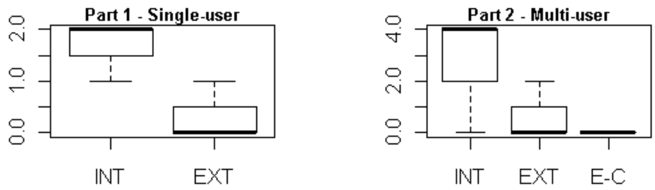


Fig. 2. Boxplots of the total number of errors committed by participants in part 1 and part 2

In part 3 we asked participants to indicate the best and worst feedback mechanisms for each of the widget types. The relative frequency of the occurrence of each mechanism in the "best" and "worst" classifications is shown in Table 2. From these results it is apparent that the External-Cumulative mechanism had the highest positive influence in the "best" classification, and the Internal mechanism resulted in the highest influence for the "worst" classification.

Table 2. Relative frequency of the occurrence of each feedback mechanism in the "best" and "worst" classifications (percentages were rounded to the nearest integer.)

	Internal	External	External-Cumulative
Best	28%	6%	66%
Worst	60%	31%	8%

6 Discussion

The results from part 1 and 2 clearly indicate that users perform worse under the Internal feedback mechanism. Participants subjected to the Internal feedback mechanism often expressed confusion about the meaning of the feedback. Another source of frequent errors was telling the number of users who were interacting, in part 2. Participants often missed a feedback event when faced with the Internal mechanism. These results were expected since the graphical cues used in the Internal mechanism are subtler than in the External and more propitious to being unnoticed.

Results do not show a statistically significant performance difference between the External and External-Cumulative mechanisms. However, results from part 3 suggest that users generally prefer the External-Cumulative mechanism (the results also suggest that users generally consider the Internal feedback mechanism to be the worst).

In some widgets, participants were divided about which mechanism was more suited: the Internal or the External-Cumulative. Participants liked the Internal mechanism when applied to textboxes and listboxes. In the textbox case, some participants felt it was more natural to present the feedback (which included the submitted text) inside the textbox, as it would happen if a user were entering text using a keyboard on a desktop computer, resulting in a preference for the Internal mechanism, for textboxes. In the listbox case, some participants liked the fact that the feedback appeared exactly on the options that were selected giving a more direct cue about what options were affected. Still, usage of textboxes and listboxes did not result in a statistically significant difference in performance, which suggests that, although users would prefer a slightly different visual cue, it did not affect how well the information was perceived.

The low error rate of the External and External-Cumulative mechanisms also suggests that the basic idea behind the feedback mechanisms (on screen with users identified by the masked phone number) is understandable and does not present any significant difficulty to users.

7 Conclusion

We have evaluated three different on screen feedback mechanisms for public display widgets in a controlled user study. The performance of the External and External-Cumulative mechanisms indicates that users are able to identify and understand the feedback for their input and for other people's input. The study also showed that users generally prefer the External-Cumulative feedback mechanism, although, for some widget types, there is still room for improvement regarding the feedback.

Future work on this subject will need to evaluate other aspects of the interaction with these widgets. Namely, it is important to study if, in a real setting, these widgets are effectively perceived as interaction opportunities and to what extent a real social context may affect that perception. The number of widgets used by an application, the number of users submitting input simultaneously to the same or different widgets in an application, as well the level of distraction, and the feedback latency are all variables that could affect the performance of this feedback approach in a real setting.

Acknowledgements

The research leading to these results has received funding from FCT under the Carnegie Mellon - Portugal agreement. Project Wesp (Web Security and Privacy (Grant CMU-PT/SE/028/2008). Jorge Cardoso has been supported by “Fundação para a Ciência e Tecnologia” and “Programa Operacional Ciência e Inovação 2010” (POCI 2010), co-funded by the Portuguese Government and European Union by FEDER Program and by “Fundação para a Ciência e Tecnologia” training grant SFRH/BD/47354/2008.

References

1. Swick, R.R., Ackerman, M.S.: The X Toolkit: More Bricks for Building User Interfaces, or Widgets for Hire. Proceedings of the Usenix Winter 1988 Conference. pp. 221-228 (1988)
2. Conner, B.D., Snibbe, S.S., Herndon, K.P., Robbins, D.C., Zeleznik, R.C., Dam, A. van.: Three-dimensional widgets. I3D ’92: Proceedings of the 1992 symposium on Interactive 3D graphics. pp. 183-188 ACM, New York, NY, USA (1992)
3. Ryall, K., Esenther, A., Forlines, C., Shen, C., Shipman, S., Morris, M.R., Everitt, K., Vernier, F.: Identity-Differentiating Widgets for Multiuser Interactive Surfaces. IEEE Computer Graphics and Applications. 26, 5, 56-64 (2006)

4. Rohs, M.: Visual Code Widgets for Marker-Based Interaction. ICDCSW '05: Proceedings of the Fifth International Workshop on Smart Appliances and Wearable Computing. pp. 506-513 IEEE Computer Society, Washington, DC, USA (2005)
5. Dearman, D., Truong, K.N.: BlueTone: a framework for interacting with public displays using dual-tone multi-frequency through bluetooth. Ubicomp '09: Proceedings of the 11th international conference on Ubiquitous computing. pp. 97-100 ACM, New York, NY, USA (2009)
6. Brignull, H., Rogers, Y.: Enticing People to Interact with Large Public Displays in Public Spaces. In: Rauterberg, M., Menozzi, M., and Wesson, J. (eds.) INTERACT '03. pp. 17-24 IOS Press (2003)