Touchless interfaces for public displays: can we deliver interface designers from introducing artificial push button gestures?

Vito Gentile
Dipartimento di Ingegneria Chimica, Gestionale, Informatica e Meccanica (DICGIM)
Università degli Studi di Palermo
Viale delle Scienze, ed. 6
90128 – Palermo (Italy)
vito.gentile@unipa.it

Salvatore Sorce
Dipartimento di Ingegneria Chimica, Gestionale, Informatica e Meccanica (DICGIM)
Università degli Studi di Palermo
Viale delle Scienze, ed. 6
90128 – Palermo (Italy)
salvatore.sorce@unipa.it

Alessio Malizia
Human Centred Design Institute (HCDI)
Brunel University London
UB8 3PH, Uxbridge, Middlesex, UK
Alessio.Malizia@brunel.ac.uk

Dario Pirrello
Dipartimento di Ingegneria Chimica, Gestionale, Informatica e Meccanica (DICGIM)
Università degli Studi di Palermo
Viale delle Scienze, ed. 6
90128 – Palermo (Italy)
d.pirrello146@gmail.com

Antonio Gentile
Dipartimento di Ingegneria Chimica, Gestionale, Informatica e Meccanica (DICGIM)
Università degli Studi di Palermo
Viale delle Scienze, ed. 6
90128 – Palermo (Italy)
antonio.gentile@unipa.it

ABSTRACT
Public displays have lately become ubiquitous thanks to the decreasing cost of such technology and public policies supporting the future development of smart cities. Depending on form factor, those displays might use touchless gestural interfaces that therefore are becoming more often the object of public and private research. In this paper, we focus on touchless interactions with situated public displays, and introduce a pilot study on comparing two interfaces: an interface based on the Microsoft Human Interface Guidelines (HIG), a de facto standard in the field, and a novel interface, designed by us. Differently from the HIG, our interface displays an avatar, which does not require an activation gesture to trigger actions. Our aim is to study how the two interfaces address the so-called interaction blindness — the inability of the users to recognize the interactive capabilities of those displays. According to our pilot study, although providing a different approach, both interfaces proven effective in the proposed scenario: a public display in a campus building's hall.

CCS Concepts
• Human-centered computing → Interaction design → Interaction design process and methods → User interface design.

Keywords
Touchless interfaces; public displays; gestural input; user interface design

1. INTRODUCTION
In the last decade, many authors have investigated and studied touchless and gestural interactions as a novel tool for interacting with computers. According to the definition by de la Barré et al. [1], "interaction is said to be touchless if it can take place without mechanical contact between the human and any part of the artificial system". This means that, for instance, interacting with a system using some controller, such as the Nintendo Wiimote or any other similar device, cannot be considered touchless interaction. On the contrary, eye trackers (such as Tobii EyeX and similar) or Kinect-like devices [2] have been widely accepted as valid examples of devices that enable for touchless interactions.

Recently, several authors proposed touchless interaction as a new way for interacting with public displays [8] [9]. One of the main advantages of this idea is the possibility of offering interactive solutions to users also if the display is placed in a non-touchable or non-reachable area. This approach can be useful, for instance, in order to provide wheelchair users with accessibility and/or to prevent vandalisms. Moreover, very large displays (e.g. media façades [8] [11]) can still be interactive via touchless-enabled technologies.

The aforementioned scenarios, however, are often complicated by several typical issues of public displays and touchless interactions. Most of them may be solved by an appropriate design of the visual interface, and by using appropriate mechanisms.

In this paper, we present a novel interface for enabling touchless interactions with public displays. We have compared our interface with another one based on the Microsoft Human Interface Guidelines (HIG) [6], which in our opinion can be considered a *de facto* standard for applications developed using Microsoft Kinect devices[1]. Our comparative study have been based on users' opinions, collected in a pilot study via interviews and questionnaires. After a brief section dedicated to related works, we describe the compared interfaces, the collected qualitative results and our plans for future works.

2. RELATED WORKS
Nowadays public displays are everywhere. We can found them in squares, malls and many other public places. Moreover, many of them are augmented with interactivity. Despite the wide adoption of touchscreens as main input (and output) devices, new

[1] This statement is supported by the high number of existing interfaces based on the HIG, specifically for gesture-controlled games.

interaction modalities have emerged to fulfill specific needs of public display systems. For instance, the increasing number of interactive *media façades*, defined as installations in which displays are integrated into architectural structures [8] [11], have implied the need of interacting from distance, and without any physical input device. Many authors proposed interaction methods based on detection of users' position and their body movements, as well as by using gestures or mobile devices (see for instances Aarhus by Light [12], DTW [13] and/or Climate Wall [14]).

However, such kind of interaction modalities are still rarely used in *situated public displays*. The latter term refers to smaller displays (size ranges from TV- to billboard-sized screens [8]), still placed in public spaces (both indoor and outdoor) that more often include touch-sensing features. Touchless gestural technologies have been less often studied for this kind of devices, probably due to the size and position of the screens, that allow for interactions by touch. However, despite the widespread of touch-based technologies for enabling interactivity on situated public displays, some authors have investigated touchless interactions. In some circumstances applications are very specific (see [16] and [8]), and it is difficult to design interfaces by following any kind of "standard" guidelines. This is the case of games and other forms of entertainment systems. For information provision systems, the definition and application of general guidelines may be more straightforward. In addition, they may result in different systems with the same interaction paradigm. Several general-purpose applications for public displays have been proposed by many authors, and quite a lot of them are based on the Microsoft HIG. For instance, most of features and controls used in the gestural system proposed by Cremonesi et al. [9], are implementations of HIG. Similar ideas have been adopted in [15].

Despite the applications, and guidelines used for designing them, touchless interactions must be studied by keeping in mind other peculiar issues of public displays. First of all, experimenters must take into account all the influencing factors of a public place. The best results may be achieved conducting the study *in-the-wild* [10], although this increase costs and efforts in terms of time and set up challenges. Furthermore, researchers must take into account issues like the need of overcoming the *display blindness* [4] and – probably more complicated – the *interaction blindness* [3]. Solutions to the first one have been studied in the field of persuasive computing. The interaction blindness – to which we have focused our interests – have been investigated by many authors. They agree it is one of the most relevant issue with public displays, and proposed different solutions to overcome it. Among them, Müller et al. [16] studied the so-called *remote honeypot effect*, which can be observed when multiple public displays are interconnected. Authors noticed that if the silhouette of a user who interacts with a display A is also shown in another display B on the same network, users in front of display B are encouraged to interact, guessing the interactivity of the display. Other mechanisms are based on giving users explicit indications about the display interactivity, e.g. using introductive video tutorials or posters. Moreover, in [17] authors show that displaying users' silhouettes may help in communicating display interactivity to passers-by.

3. INTERFACES COMPARISON

In the following, we describe the main features of our proposed interface. Next, we will compare it with another touchless interactive interface, based on the Microsoft HIG. Both of them are aimed to the same goal.

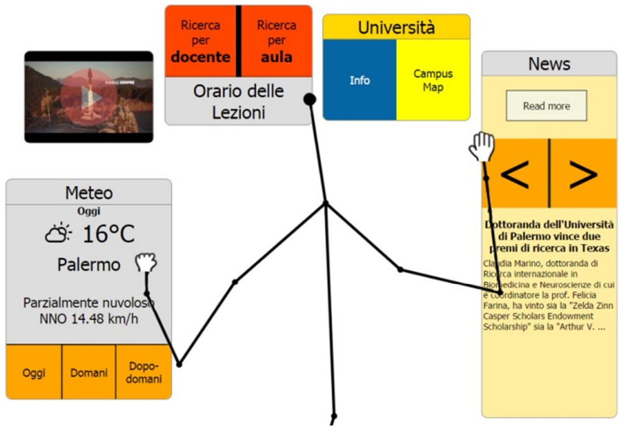

Figure 1. Layout of our avatar-based interface.

In more detail, we have developed two interfaces for the same system to be used in a public space inside a University campus. The main goal of our system is to provide an easy access to useful information for students, such as news, events, weather data and lecture timetables.

3.1 Avatar-based Interface

In order to design a proper touchless interface for public displays, we focused on two of the main issues related to public interactive displays. The first problem is the interaction blindness, for which users do not understand the display interactivity and its touchless nature. While in touch-based displays, claims like «touch me» or «touch screen» may be helpful, there is the need for a different approach for touchless interactions. The second problem arises from the need for novel visual interfaces expressly designed for touchless gesture-based and natural interactions, that should outdo the WIMP paradigm commonly used in desktop-based systems.

To address both issues, we have designed a novel interface that uses only *in-air direct manipulation*, as defined in the following. According to [5], one promising solution to implement interactions that are more natural is the use of *direct manipulations*, instead of symbolic gestures. Such paradigm, however, is appropriate in touch-based or tangible systems, where "touching" actions allow for the direct manipulation of objects in the interface. We thought that this paradigm could be extended to touchless interfaces, thus becoming what we call *in-air direct manipulations*. By means of body movements and in-air gestures, it is possible to imitate the direct manipulation of an object, as we would do in the real life, without actually grabbing or touching them.

To support our choice of using in-air direct manipulations, besides our considerations, we have been also inspired by dontclick.it [7], a website that allows its users to browse contents without the need of a single click. In dontclick.it it is possible to open sections, select items and animate objects, just by moving the mouse over them and without pushing any buttons. Interestingly, the statistics of the website show that the majority of users do not miss the click. By observing dontclick.it, we supposed that if it is possible to interact naturally with a web page without a single click (that is an activation gesture for that interface), it should be possible to interact even more naturally with a touchless interface without any activation gesture. Indeed, we think that it should be more difficult to avoid the use of the 'click' having a mouse in a hand, rather than avoid any activation gesture having nothing in the hand. In other words, we based the design of our proposed interface on the hypothesis that in-air direct manipulation will improve the naturalness of touchless interactions.

Our interface is based on the presence of an avatar placed in the middle of the screen, which continuously replays user's movements (see fig. 1). The avatar appears whenever a user approaches the display. This idea is inspired by the User Viewer control proposed by Microsoft HIG. The difference is that our avatar, once appeared, is permanently present in the middle of the screen, and other interface components are placed all around it. The hands of the avatar are depicted as two distinct hand-shaped cursors, by means of which the user can interact with the available tiles just by placing them on top of these tiles – with no activation gestures. There is also the possibility to close the hand (the so-called *grip* gesture) to trigger specific (but not primary) actions. This is useful to implement the zoom feature, by closing both hands and bringing them nearer or farther.

3.2 HIG-based Interface

In HIG, Microsoft recommends the use of one or two Kinect cursors, i.e. hand-shaped cursors by means of which users can interact with several tiles in the interface. Microsoft also suggest the use of an activation gesture, the so-called *push-to-press*, consisting in emulating a pushing action that, if executed when a Kinect cursor overlays an interactive tile, triggers the corresponding event.

Another interesting feature is the presence of the User Viewer control, a small frame in the middle upper border of the window that shows the user silhouette (taken from Kinect depth camera). The presence or the absence of this silhouette suggests users if they are detected or not.

Using the Microsoft Kinect SDK (which includes several controls ready to be used for implementing HIG), we have developed another interface (see fig. 2) that allows for the same operations of our avatar-based system.

3.3 Tasks

We implemented both interfaces to provide users with exactly the same functionalities and to accomplish the same tasks, and in particular:

- reading news;
- reading university information;
- displaying and navigating campus map;
- displaying lecture timetable;
- displaying weather data;
- displaying a video.

Of course, the layouts of the two interfaces were different, due to the presence of the avatar in the middle of the screen, and to reduce the Midas touch issue [5] due to involuntary interactions with the tiles.

3.4 Experiment Setup

In order to evaluate which of the aforementioned interfaces better fit users' needs, we set up a pilot experiment to collect users' opinions.

A public display was installed in a public transit area inside a building within the University campus in Palermo. This is an area where students of several disciplines and different age typically hang out. We asked 12 students (7 male, 5 female) to interact with our system, testing both interfaces for each participant ("within subject" set up).

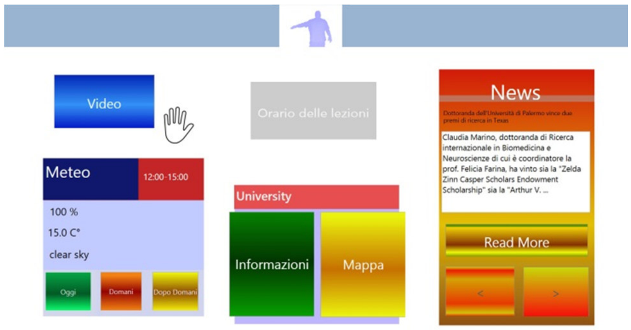

Figure 2. Layout of the HIG-based interface used in our comparison study.

Each user performed two 5-minutes-long interaction sessions (one per interface), both followed by semi-structured individual interviews. In order to randomize the users sample and to variegate their technology-related skills, we enrolled students attending various courses, in particular Computer Engineering, Arts and Theatre.

Concerning the hardware, we used a 32-inch monitor placed at eye-sight, with a Microsoft Kinect sensor (clearly visible to all users) above the monitor.

In the interaction sessions, we asked each participant to execute the following tasks:

1. find and read a specific news;
2. find and read university information;
3. find the timetable for a specific class;
4. play a video;
5. find and read the weather forecast for the next day.

We asked users to perform these tasks without any suggestions or hints on how to achieve such goals, especially in terms of interaction modality.

In the following, we present some interesting findings that will guide our future works.

3.5 Lessons Learnt

Among all the differences between the interfaces, our study were focused on understanding users' preferences related to two crucial aspects: 1) the use/non-use of an activation gesture to execute a command such as the selection or the click on a button, and 2) if and how the presence of an avatar in the middle of the screen helps in overcoming interaction blindness.

During the interviews after the interaction session with the avatar-based interface, 6 users assessed they miss a gesture that allows to "click" on the tiles shown in the interface. Among the remaining 6, only 2 assessed they were comfortable in interacting without activation gestures. Users explained the need by explicitly referring to the habit of using mice and touch-based systems. Interestingly, others used some activation gestures also if they were not necessary, with the consequence of complaining about the fact that "some buttons activates by themselves".

On the other hand, guessing the activation gesture could result frustrating. None of the participants used the "push-to-press" gesture on which the HIG-based interface was based, starting by using other gestures (e.g. closing the hand, using a single finger, etc.). The difficulties in guessing the gesture to use may convince users to stop any further interactions. For this reason, we believe that the idea of avoiding activation gesture should be still pursued.

Furthermore, we noticed another interesting users' behavior: all users except one preferred the use of both hands while interacting with the avatar-based interface, whilst all users except one interacted by a single hand with the HIG-based interface. Some of them explained this behavior as a consequence of their habit in using a mouse (which is always dragged by the same hand). Because of the presence of a cursor in the HIG-based interface, they used their gestures as if they were moving a mouse. On the other hand, being able to see the avatar in the screen seemed to elicit the use of both arms.

Another interesting characteristic of the avatar we observed, was its ability to communicate the touchless interactivity supported by the interface. Most of the users knew the Kinect sensor, and it was the main clue to understand the possibility of interacting with gestures, but several users explained that they were immediately able to guess the touchless nature of the system after having seen the avatar. However, its presence was perceived as annoying, muddler and useless when users interacted with the video or wanted to read a long text: in such cases, the avatar continued to be in the middle of the interface, while users assessed that using hand-shaped cursors would be a better choice.

As a result, we can imagine a sort of "fusion" of the two approaches we tested. The use of only hand-shaped cursors (as in HIG, except for dropping any activation gesture) could be a good choice for interacting with videos, images, texts and other contents. The avatar seems to be the best choice to interact with menus, allowing also for the reduction of the interaction blindness-related issues.

4. CONCLUSION

We presented a pilot study comparing two alternative interface models for touchless interaction with situated public displays. Users appreciated both models, but their opinions suggest that the use of an avatar that mirrors user's movements is a good choice for interacting with menus and to communicate the touchless interactivity of such systems. On the other hand, using an HIG-based interface is less confusing, but the use of an activation gesture may discourage users for further interactions.

In conclusion, a hybrid solution should be the best choice for addressing both users' preferences and solving interaction blindness-related issues.

In the future we are going to improve our avatar-based interface by removing the avatar when it overlays on texts, images or videos, in order to avoid clumsy interactions.

We will also reinforce our analysis by including the time-to-task data, the evaluation of error rates during the interactions, to obtain a better and more reliable quantitative evidence of users' preferences.

5. REFERENCES

[1] de la Barré, R., Chojecki, P., Leiner, U., Mühlbach, L., & Ruschin, D. (2009). Touchless Interaction-Novel Chances and Challenges. In J. A. Jacko (Ed.), *Human-Computer Interaction. Novel Interaction Methods and Techniques* (pp. 161–169). Springer Berlin Heidelberg

[2] Gentile, V., Sorce, S., & Gentile, A. (2014). Continuous Hand Openness Detection Using a Kinect-Like Device. In *2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)* (pp. 553–557). Birmingham, UK: IEEE. doi:10.1109/CISIS.2014.80

[3] Ojala, T., Kostakos, V., Kukka, H., Heikkinen, T., Linden, T., Jurmu, M., … Zanni, D. (2012). Multipurpose Interactive Public Displays in the Wild: Three Years Later. *Computer*, *45*(5), 42–49. doi:10.1109/MC.2012.115

[4] Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., & Krüger, A. (2009). Display Blindness: The Effect of Expectations on Attention towards Digital Signage. In H. Tokuda, M. Beigl, A. Friday, A. J. B. Brush, & Y. Tobe (Eds.), *Proceedings of the 7th International Conference on Pervasive Computing Pervasive '08* (Vol. 5538, pp. 1–8). Newcastle, UK: Springer Berlin Heidelberg.

[5] Wigdor, D., & Wixon, D. (2011). *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann Publishers Inc.

[6] Microsoft. *Human Interface Guidelines*. Retrieved from http://download.microsoft.com/download/6/7/6/676611B4-1982-47A4-A42E-4CF84E1095A8/KinectHIG.2.0.pdf

[7] Dontclick.it, http://dontclick.it

[8] Gehring, S., & Wiethoff, A. (2014). Interaction with Media Façades. *Informatik-Spektrum*, *37*(5), 474–482.

[9] Cremonesi, P., Di Rienzo, A., Frà, C., Garzotto, F., Oliveto, L., & Valla, M. (2014). Personalized interaction on large displays. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces - AVI '14* (pp. 353–354). New York, New York, USA: ACM Press.

[10] Gentile, V., Malizia, A., Sorce, S., & Gentile, A. (2015). Designing Touchless Gestural Interactions for Public Displays In-the-Wild. In M. Kurosu (Ed.), *Human-Computer Interaction: Interaction Technologies* (pp. 24–34). Springer International Publishing

[11] Haeusler, M. H. Media facades: history, technology, content. Avedition, Ludwigsburg, 2009.

[12] Brynskov, M., Dalsgaard, P., Ebsen, T., Fritsch, J., Halskov, K. and Nielsen, R. Staging Urban Interactions with Media Facades. In Human-Computer Interaction - INTERACT 2009. Springer Verlag, Heidelberg, 2009, 154-167.

[13] Dalsgaard, P. and Halskov, K. Dynamically transparent window. In CHI EA '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems. ACM, New York, NY, USA, 2009, 3019-3034.

[14] Fritsch, J. and Dalsgaard, P. Media Facades Beyond Interaction. In Proceedings of OzCHI 2008.

[15] Hardy, J., Rukzio, E., & Davies, N. (2011). Real world responses to interactive gesture based public displays. In *Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia - MUM '11* (pp. 33–39). Beijing, China: ACM Press

[16] Müller, J., Eberle, D., & Tollmar, K. (2014). Communiplay: a field study of a public display mediaspace. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14* (pp. 1415–1424). New York, New York, USA: ACM Press.

[17] Müller, J., Walter, R., Bailly, G., Nischt, M., & Alt, F. (2012). Looking glass: a field study on noticing interactivity of a shop window. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12* (pp. 297–306). New York, New York, USA: ACM Press