Personalized Interaction on Large Displays: The StreetSmart Project Approach

Paolo Cremonesi[1], Antonella Di Rienzo[1], Cristina Frà[2], Franca Garzotto[1], Luigi Oliveto[1], Massimo Valla[2]

1. Department of Electronics, Information and Bioengineering – Politecnico di Milano
Via Ponzio 34/5 20133, Milano, Italy
{firstname.lastname}@polimi.it

2. Open Innovation Research – Telecom Italia S.p.A.
Via Golgi 42, Milano, Italy
{firstname.lastname}@telecomitalia.it

ABSTRACT
The StreetSmart Project develops information services that integrate multiple (touch and touchless) interaction paradigms on personal devices and large public displays. It exploits personalization techniques in order to offer new engaging user experiences involving large amounts of multimedia contents.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: Multimedia Systems, User Interfaces

General Terms
Design, Experimentation, Human Factors

Keywords
Motion-based touchless interaction, large screen, mobile devices, personalization, recommender system, mobile interaction

1. INTRODUCTION
Digital multimedia information is more and more woven into the fabric of our living contexts through large displays in public indoor and outdoor spaces, e.g. shops, galleries, cultural sites, airports [4]. The StreetSmart Project offers new User Experiences (UXs) that make the consumption of multimedia contents on shared large displays more engaging and useful.

The StreetSmart approach includes a set of core features. First, our UXs integrate large public displays with personal devices (tablets or smartphones) and combine *multiple interaction paradigms* to explore the information through the screens, individually or in group. We exploit the interpretation of body movements (*touchless gestural interaction* [1][5]) as well as touch gestures on personal devices as control mechanisms (*multi-touch remote interaction*). In addition, we *personalize* contents on the large displays to increase users' engagement, to facilitate contents exploration and to reduce information overload, "recommending" the items that are likely to fit users' interests and characteristics [2] [3]. Personalization techniques are applied to create a user profile of the current user or group in front of the screen; based on the profile a recommendation engine suggest the most relevant items for the user/group.

The user profile considers various types of information with increasing levels of details: demographic data (gender and approximate age detected using image processing techniques), behavioral data inferred from the history of interactions with other displays connected to the system, or preferences explicitly declared by users or retrieved from their social networks profiles.

AVI '14, May 27 – 30, 2014, Como, Italy
ACM 978-1-4503-2775-6/14/05.
http://dx.doi.org/10.1145/2598153.2600055

2. Examples
Some of the UXs supported by StreetSmart applications are exemplified in the following scenarios.

Scenario 1

A family (a couple and two young children) is visiting Milano for the first time. In the Duomo gallery, large interactive walls display multimedia information about major attractions: museums, shopping areas, fashion districts, gourmet food spots, music halls or events. When the family is in front of the large interactive screen, image processing techniques are used to recognize gender and age of the group members. The proposed contents are personalized to the characteristics of the group, for example suggesting a set of restaurants appropriate for families. A single person or the entire group can then interact with screen content using movements and gestures.

Scenario 2

At the entrance of the main exhibition area of EXPO 2015[1], large displays welcome the visitors and present information about scheduled events. A visitor, who has previously downloaded the StreetSmart app on her smartphone and logged in using her Facebook account, moves close to a large screen. Her personal device is sensed by the StreetSmart system and the large display starts showing events or stands that are more appropriate to her profile, which is created using the preferences she has explicitly declared on the app or the interests mined from her Facebook page. The visitor can use her mobile phone to control the large display and to explore the various contents. Once she has decided which stands to visit and which events to attend, she can download on her personal device the related information (e.g. event and stand descriptions, maps and routes to reach them, etc.).

[1] EXPO will be held in Milan from May 1st to October 31st , 2015

3. SYSTEM ARCHITECTURE

The general software architecture of StreetSmart system is depicted in Figure 1. Applications modules can be hosted both on users' personal devices and on public large screens. The application on public screens exploits Microsoft's Kinect motion sensing technology to implement touchless gestural interactions. The available Kinect APIs have been adapted to meet the requirements of the gesture language of StreetSmart UXs and have been integrated with additional components. The detection of user group participants and their age and gender from the images captured by the Kinect is currently supported invoking Betaface, an image processing web service [6]. NFC technology enables the identification of personal devices. The mobile app on the smartphone allows the user to login and to be recognized by the system.

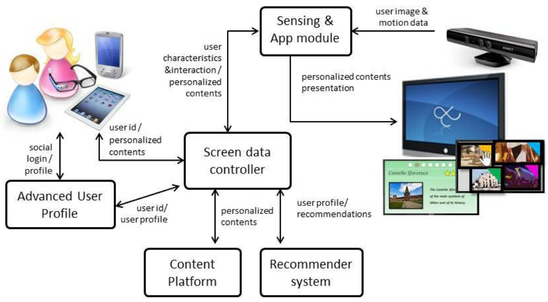

Figure 1: StreetSmart System Architecture.

The Screen Data Controller component is responsible of creating the user profile, invoking the Recommender System that identifies the relevant items according to this profile, and retrieving the corresponding multimedia contents to be proposed on the shared and personal devices. The user profile is created by elaborating interaction history data and complementing them with additional information: either age/gender generated by the image processing service or more advanced features. The latter are obtained by communicating with the user personal device and retrieving the preferences configured on the app or mined from user social accounts. Multimedia contents come from the Content Platform fed by services editors with official items, points of interest, events, advertisings, etc.

4. CURRENT STATUS AND FUTURE WORK

An example of StreetSmart application enables the exploration of restaurants and points of interest in Milano (**Figure 2**). Some limited sets of items and metadata have been retrieved from public sites to validate the approach. The application is delivered in two versions, each one supporting a different interaction paradigm: touch-less gestural interaction or touch interaction using personal devices.

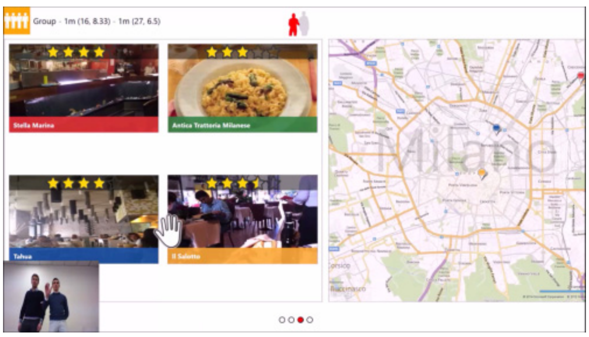

Figure 2: StreetSmart Large Screen Application: Interface.

After a set of incremental prototypes and iterative evaluations done in our labs, the first beta version of the application has been tested in a public space. In the cafeteria of our university (Figure 3), 12 users have been involved in assisted test sessions. Data gathered from observations and semi-structured interviews show a good degree of usability of both versions, and comparable levels of user satisfaction and engagement. Still, a number of aspects need to be improved: the aesthetics of the visual interface on the large display (e.g. layout and visual quality of multimedia contents) and the performance and precision of both the body movements processing and the age and gender interpretation component.

Figure 3: Evaluation of StreetSmart applications: touchless interaction (left) and multi-touch remote interaction (right).

From the experience gained so far, the research challenge of StreetSmart application relies upon the intrinsic complexity of

- enabling *touchless* interaction with *large* amount of multimedia contents;
- supporting the interaction of single and group of users;
- combining all these aspects with personalization features.

A key issue of touchless interaction is how to support movement and mid-air gestures that are intuitive and natural. This problem is exacerbated when interacting with large amounts of multimedia contents, which need to be organized in non-trivial information architectures. In these contexts, the amount of interaction tasks is higher compared to those involved in the interaction with simple, linear information structures, and tasks are semantically more complex. In addition, supporting the interaction of both a single person and a group raises the issue of discriminating between individual's and group's movements and of interpreting them. Finally, an open problem is how to identify the characteristics of both individuals and groups that are appropriate for profiling purposes, to meet the algorithmic requirements of recommendation engines and to build effective recommendations.

5. ACKNOWLEDGMENTS

This work is supported by the EC commission through EIT ICT Labs Project 14183/2014 "StreetSmart".

6. REFERENCES

[1] Bellucci, A., Malizia, A., Diaz, P., Aedo, I. 2011. Human-display interaction technology: Emerging remote interfaces for pervasive display environments. *Pervasive Computing, IEEE* 9, 2, 2010, 72 –76.

[2] Cremonesi, P., Garzotto, F., Turrin, R., 2012. Investigating the Persuasion Potential of Recommender Systems from a Quality Perspective: an Empirical Study, *ACM TIIS* 2(2) June 2012, 41-73.

[3] Cremonesi, P. and Garzotto, F., Smoothly Extending e-Tourism Services with Personalized Recommendations: A Case Study. *Proc. EC-Web 2013*. Springer, 2013, 174-181.

[4] Kuikkaniemi K., Jacucci G., Turpeinen M., Hoggan E. 2011 From space to stage: How interactive screens will change urban life. *Computer* 44, 6, 2011, 40 –47.

[5] Marquardt N., Diaz-Marino R., Boring S., Greenberg S. 2011. The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. *Proc. UIST 2011*, ACM, 315-326.

[6] Betaface: face detection and recognition - http://betafaceapi.com/