PaperView: Augmenting Physical Surfaces with Location-Aware Digital Information

Dimitris Grammenos[a], Damien Michel[a], Xenophon Zabulis[a], Antonis A. Argyros[a,b]
[a] Institute of Computer Science –FORTH, Heraklion, Crete, Greece
[b] Department of Computer Science, University of Crete, Heraklion, Crete, Greece
{gramenos, michel, zabulis, argyros}@ics.forth.gr

ABSTRACT
A frequent need of museums is to provide visitors with context-sensitive information about exhibits in the form of maps, or scale models. This paper suggests an augmented-reality approach for supplementing physical surfaces with digital information, through the use of pieces of plain paper that act as personal, location-aware, interactive screens. The technologies employed are presented, along with the interactive behavior of the system, which was instantiated and tested in the form of two prototype setups: a wooden table covered with a printed map and a glass case containing a scale model. The paper also discusses key issues stemming from experience and observations in the course of qualitative evaluation sessions.

Author Keywords
Interactive museum exhibits; interaction design; magic lens

ACM Classification Keywords
H5.2. User Interfaces; I.4 IMAGE PROCESSING AND COMPUTER VISION; I.4.9 Applications

General Terms
Design, Experimentation

INTRODUCTION
In the past few years, a number of museums worldwide started exploring new ways for integrating novel interactive exhibits in their spaces, moving beyond the typical "multimedia information kiosk" paradigm of the past (e.g., [5, 6, 8, 10, 13]). The basic goal of this approach is to substitute sheer didactic information provision with captivating experiences that utilize a combination of active user participation through natural interactions and dynamic media in order to support constructive and engaging edutainment (i.e., entertainment designed to educate).

Maps and scale models constitute two types of exhibits that are frequently met in museums. A related need is to provide their visitors with location-related information about such exhibits. Typically, this is achieved by: (a) overlaying information on the exhibits and (b) putting up support material in nearby walls. In the first case, one problem is the limited available space and another one the fact that the actual exhibit content becomes visually cluttered. In the second case, a key drawback is that visitors have to look back and forth between the exhibit and the information and mentally correlate the two. An unsolvable problem shared by both approaches is multilingualism, since each language requires additional space. In this context, this paper suggests PaperView, an augmented-reality approach for supplementing physical surfaces with digital information through the use of pieces of plain paper that act as personal, location-aware, interactive screens.

TEI'11, January 22–26, 2011, Funchal, Portugal.

RELATED WORK
A number of interactive exhibits have been deployed in museums worldwide. The "Re-Tracing the Past" exhibition [5] of the Hunt Museum in Limerick, Ireland comprised two room-sized spaces and the "Fire and the Mountain" [6] exhibition at the Civic Museum of Como, Italy, comprised four hybrid exhibits. The Austrian Technical Museum in Vienna opened a digitally augmented exhibition [8] and ARoS, an art museum in Denmark, employed four interactive exhibits [10]. The Ragghianti Foundation in Milan, Italy also held an exhibition entitled "Puccini Set Designer" [13] that used new technologies.

PaperView implements a "magic-lens"-type augmented-reality approach, building upon the paradigm originally suggested by Wellner [15] for DigitalDesk. In this direction, Reitmayr et al. [11] augmented paper maps using a PDA and a rectangular piece of cardboard with a black border to browse images. A limitation of their implementation was that only a single PDA and one cardboard could be tracked. Following a different line of thinking, Bimber et al. [1] presented an approach for digitally augmenting pictorial artwork, also supporting user interaction through a secondary screen and a mouse. Holman et al. [7], project a windowing environment on physical paper to simulate the use of digital paper displays. The pieces of paper and the user fingers are augmented with IR markers in order to be tracked. The Magic Lenses framework [4] employs a handheld mirror-like prop with an attached magnetic tracking sensor and a "selection icon" with a secondary display to augment a 3D image projected on a workbench. In order to tack this icon, a surface coated with the retro-reflective material is used, along with an IR camera and IR illumination. Izadi et al. [9] present a rear-projection surface that can project a second image through it onto a sheet of projection film placed on it, acting as a magic lens. In order to be tracked, the magic lens is augmented with either passive or active tags.

THE PAPERVIEW SYSTEM

PaperView is a tabletop augmented reality system that builds upon Wellner's [15] DigitalDesk concept in combination with Reitmayr et al. [11] augmentation of maps using real world objects, but taken further, since multi-user interaction, finger-based input and concurrent tracking of diverse border colors are supported. In contrast to PaperWindows [7] and SecondLight [9] no IR markers are used to augment the paper or the users' fingers (something that would not be practical in a museum setting). Also, unlike Magic Lenses [4], no magnetic tracker, or special coating of the table surface are used. PaperView employs a projector and rectangular pieces of plain white paper upon which visual information and interactive controls are presented, digitally augmenting physical surfaces with context-sensitive data. Two alternative prototype setups were developed; a large wooden table covered with a printed map and a glass case containing a scale model. The key difference between the two is that the latter can easily be installed with practically no construction requirements or physical interventions.

Technical Overview

The system comprises the surface to be augmented, a projector, an RGB camera, a PC and speakers. Additionally, rectangular pieces of plain white cardboard of various sizes are used (20x20 to 30x30cm^2). Cardboard pieces are surrounded by a colored frame (1-2cm). A computer vision system is used to track the position and pose of the paper surfaces, as well as the activation of any related interactive areas by the users' fingers. This information is passed through a custom middleware layer to an application developed in Actionscript 3.0 using Adobe Flash, which implements the system's user interface.

Each cardboard appears in the camera field of view as a quadrilateral. Upon setup, the projector-camera system is spatially calibrated. The camera is set to adjust gain automatically, so as to adapt to scene illumination. The first goal of the vision system is to detect the quadrilaterals corresponding to the cardboards (see Fig. 1). This detection is based on the knowledge of the color of the border of each cardboard. A color similarity metric [12] is applied to each pixel, estimating its similarity to a reference color and resulting in a binary image B. A small colored palette in a predefined location of the scene provides the reference colors of the cardboards, invariantly to changes of global scene illumination (i.e. due to light coming in from the windows). The silhouettes of the contours appearing in image B are traced and fitted with straight lines. A check is performed regarding whether the arrangement of the resulting segments conforms to the hypothesis of two nested quadrilaterals. This prevents from ambiguities that may arise due to the presence of similarly colored objects. The detection of the quadrilaterals is tolerant to occlusions that might arise due to, e.g., users' hands.

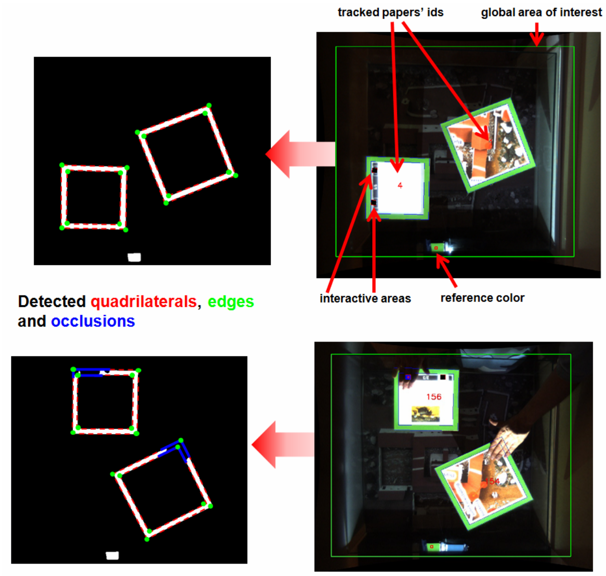

Figure 1. Computer vision system: (top-right) camera image; (top-left) color similarity image B; (bottom, right/left) images of quadrilaterals partially occluded by hands.

The system estimates the transformation that maps the detected quadrilateral(s) to rectangular regions, using a homography. The corners of the quadrilateral are estimated as the points of intersection of consecutive straight lines. The four corners are used to estimate this homography, and through the induced transformation the projected images appear undistorted on the slanted cardboard. The identity of each cardboard is maintained over time, based on polygon intersection tests across subsequent frames. The trajectories of the polygon corners are tracked using Kalman filtering modeling location and velocity, as in [14], to increase robustness and suppress jitter. The 3D location and pose of the cardboard is estimated given the size of its edges in the image and its size, during setup, by placing the cardboard in a frontoparallel posture on the surface of the table. By comparing the apparent sizes to its initial size, the slant and distance of the cardboard are estimated. Additionally, interactive areas are implemented as follows. Using the homography computed for each cardboard, various rectangular areas (dynamically defined by the application) are continuously checked for differences using a background subtraction method. The image projections of these areas are tested for the occurrence of skin-colored blobs, using the method in [1]. At the same time the neighboring area of the nested quadrilateral is checked for occlusions (such as the one shown at the bottom row of Fig.

1). Combining these two cues provides a detection event that is triggered when a user's finger continuously occurs above the interactive area.

Setup A: PaperView Table

The system (Fig. 2 and 3a) comprises a wooden table (1.8x1m^2), the surface of which is covered by a printed map of Macedonia, Greece, which does not contain any text or other kind of data. Underneath the table there are two wireless speakers and high above it, inside a false ceiling, there are an HD video projector, an IR camera and an RGB camera. Next to the table lies a 56'' HD TV screen, bearing a shelf with a stack of rectangular framed pieces of white cardboard. The frame color is used as a means for implicit language selection. Multiple users can concurrently use the table. Initially, the projector overlays on the map the location (as red spots) and names of ancient Greek cities with archeological interest (in English and Greek). If a visitor places a cardboard on the table surface, an image is projected on it, showing a satellite view of the respective printed map area ("map mode"). Furthermore, a circled crosshair is projected on the paper's centre along with a virtual red string connecting the paper with the closest site of interest. If a site's name intersects with a paper, then it is rotated so that it is aligned with the paper's orientation.

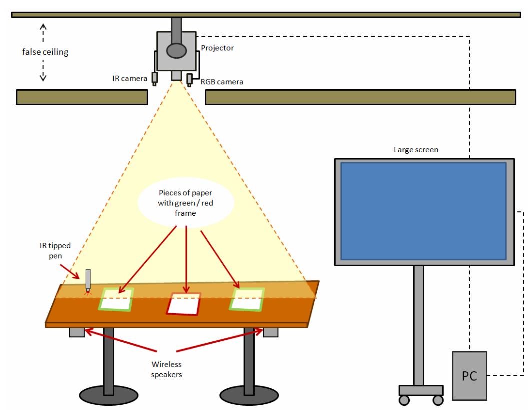

Figure 2. PaperView Table schematic installation layout

If the visitor moves the paper so that the site of interest lies within the boundaries of the crosshair, a multimedia slideshow starts ("info mode"). The slideshow comprises a series of pages, each of which may contain any combination of texts, images, and videos. When a cardboard piece is lying on the table, a toolbar is projected at its lower bottom area, containing an indication of the current page and the total number of pages (e.g., 4/23), and buttons for moving to the next/previous page. The user can interact with these "soft" buttons using his/her fingers (see Fig. 3b). If the paper is taken off the table's surface, the buttons disappear and the user can move to the next/previous page, by tilting the paper right or left, respectively. In this case, the projection is appropriately distorted (Fig. 4 - left), so that the visual content registers correctly on the paper surface. In order to avoid accidental browsing actions, page change does not happen instantly. Instead, an arrow-shaped progress bar is presented on the paper (Fig. 4 - right) and takes about 1 second to fill. In order to visually link information presented on a cardboard piece to the site it refers to, a red connecting string is used. If the cardboard is moved beyond a minimum distance, the string "breaks" and the paper's surface returns back to "map mode". The nearby TV screen presents a Google Maps view of the geographical area covered by the printed map, in order to help visitors correlate the areas presented on it to their location in the real world. Visitors can use the pen on the table surface to navigate in Google Maps. If the user selects a point of interest, related multimedia information is presented. Additionally, if the user keeps the pen at the same position for more than 1 second, a virtual remote control appears, through which she can zoom in/out in Google Maps and select alternative map views.

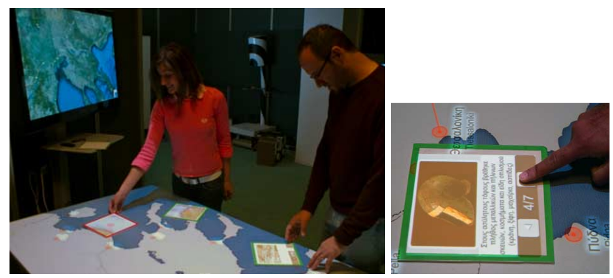

Figure 3. (a) Two users concurrently interacting with the map in different languages; (b) pressing the soft buttons

Figure 4. (left) lifting the paper above the table surface; (right) tilting the paper to browse content.

Setup B: PaperView Glass Case

The main goal of this setup was to create a self-contained "add-on" system, requiring minimal interventions to the installation space (see Fig. 5). In order to accommodate the space limitations, an ultra short throw projector was used, and the camera was equipped with a wide angle lens. For our prototype an exact replica of a museum glass case (0.83x0.83m^2) was constructed, inside of which a 1-to-1 scale printout of the original scale model (an ancient farmhouse of the 4th century BC) was placed. The system behaves similarly to the table setup, with just the following differences: (a) due to the glass case, no information is projected outside the tracked papers; (b) papers cannot be used when higher than a few centimeters above the surface, since the projector's beam is located very close and to the side of it – thus, the paper titling function is not supported; and (c) no secondary display, nor an IR pen are used.

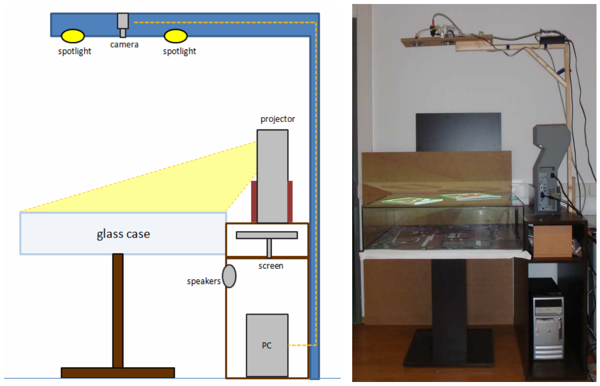

Figure 5. PaperView Glass Case schematic installation: (left) layout and (right)prototype system

DISCUSSION

Prior to installation in a real setting, both prototypes were tested and evaluated in an in-house space simulating a museum space, using a combination of the "observer participant" and "participant observer" approach [3]. During a 6-month period, more than 100 persons of various ages and backgrounds have interacted with the systems and their opinion ranged from positive to enthusiastic. The fact that interaction was supported through plain pieces of paper made a big impression, since the papers had a nice and soft feeling, and provided predictable, straightforward, interactions. Language selection is a challenging task for interactive exhibits, and is rarely addressed by previous efforts. Initially, PaperView users had to select their language from a dialogue on the cardboard, the first time it was detected. Thus, when the cardboard was taken outside the camera's viewport, the selection had to be repeated. During the evaluation sessions, this was spotted as a key usability problem. As a result, the "language selection through frame color" approach was devised, resulting in much simpler and smoother interaction.

Updated versions of both presented systems are now installed at the Archaeological Museum of Thessaloniki, Greece, as part of a permanent exhibition of interactive systems (for more information visit www.makedonopixels.org, exhibits "Multimodal Diverse travel" and "One day in a farmstead"). In order to accommodate the museum's requirements some modifications were made; e.g., the PaperView Table uses a short-throw projector, a much larger table and does not include the secondary screen. Additionally, based on *in situ* observations, additional software changes were made in order to improve the usability of the systems; e.g., to disambiguate the orientation of the projection on the cardboard, one of the colored edges is painted with a different color. Regarding usability improvements, the crosshair was replaced by a magnifying glass, the size of which changes depending on its distance from a site of interest (SOI), three different colours (green, orange and red) and sizes were used for visualizing the connecting string and when its connection to a SOI would "break", and animations were introduced to provide feedback when entering and leaving a SOI.

ACKNOWLEDGMENTS

This work has been supported by the FORTH-ICS internal RTD Programme "Ambient Intelligence and Smart Environments". Authors express their gratitude to M. Sifakis for the archaeological content of the prototype systems and to A. Katzourakis for the graphical designs.

REFERENCES

1. Argyros, A.A., Lourakis, M.I.A. 2006. Vision-based Interpretation of Hand Gestures for Remote Control of a Computer Mouse, in Proc. of the HCI'06 workshop, 40-51.
2. Bimber, O., Coriand, F., Kleppe, A., Bruns, E., Zollmann, S., and Langlotz, T. 2006. Superimposing pictorial artwork with projected imagery. In ACM SIGGRAPH 2006 Courses, 16-26.
3. Blomberg, J., Giacomi, J., Mosher, A. and Swenton-Wall, P. 1993. Ethnographic Field Methods and Their Relation to Design. In: Participatory design, LEA, 123-155
4. Brown, D., L. and Hua, H. 2006. Magic Lenses for Augmented Virtual Environments. IEEE Comput. Graph. Appl. 26, 4 (July 2006), 64-73.
5. Ferris, K., Bannon, L., Ciolfi, L., Gallagher, P., Hall, T. & Lennon, M. 2004. Shaping experiences in the hunt museum: a design case study. In Proc. of DIS '04, 205-214.
6. Garzotto, F. and Rizzo, F. 2007. Interaction paradigms in technology-enhanced social spaces: a case study in museums. In Proceedings of DPPI '07, 343-356.
7. Holman, D., Vertegaal, R., Altosaar, M., Troje, N., and Johns, D. 2005. Paper windows: interaction techniques for digital paper. In Proc. of CHI '05, 591-599.
8. Hornecker, E. and Stifter, M. 2006. Learning from interactive museum installations about interaction design for public settings. In Proceedings of OZCHI '06, 135-142.
9. Izadi, S., Hodges, S., Taylor, S., Rosenfeld, D., Villar, N., Butler, A., and Westhues, J. 2008. Going beyond the display: a surface technology with an electronically switchable diffuser. In Proc. of UIST '08, 269-278.
10. Kortbek, K. J. and Grønbæk, K. 2008. Interactive spatial multimedia for communication of art in the physical museum space. In Proceeding of MM '08, 609-618.
11. Reitmayr, G., Eade, E., and Drummond, T. 2005. Localisation and Interaction for Augmented Maps. In Proc. of 4th IEEE/ACM int/nal Symposium on Mixed and Augmented Reality, 120-129.
12. Smith, J. R. and Chang, S. 1996. VisualSEEk: a fully automated content-based image query system. In Proc. of ACM Multimedia '96, 87-98
13. Sparacino, F. 2004. Scenographies of the past and museums of the future: from the wunderkammer to body-driven interactive narrative spaces. Proc. of MM '04, 72-79.
14. Trucco , E., Verri, A. 1998. Introductory Techniques for 3-D Computer Vision, Prentice Hall, 201.
15. Wellner, P. 1993. Interacting with paper on the DigitalDesk. Commun. ACM 36, 7 (Jul. 1993), 87-96.