Image Processing Supports HCI in Museum Application

Niki Martinel, Marco Vernier, Gian Luca Foresti and Elisabetta Lamedica

Department of Mathematics and Computer Science, University of Udine, Via Delle Scienze 206, Udine, Italy

Keywords: Augmented Reality, Information Visualization, User Interface Design, Mobile HCI.

Abstract: This work introduces a novel information visualization technique for mobile devices through Augmented Reality (AR). A painting boundary detector and a features extraction modules have been implemented to compute paintings signatures. The computed signatures are matched using a linear weighted combination of the extracted features. The detected boundaries and the features are exploited to compute the homography transformations. The homography transformations are used to introduce a novel user interaction technique for AR. Three different user interfaces have been evaluated using standard usability methods.

1 INTRODUCTION

The continuous increase in power and graphics capabilities of mobile devices has introduced novel visualization techniques. Visualizing contents for mobile devices is a challenging tasks because of the limitation of screen size, etc. According to (Chittaro, 2006) it's impossible to follow a trivial porting approach from desktop computers to mobile devices. Due to the limitation of the screen sizes (Marois and Ivanoff, 2005) it is very hard to display a lot of information such that it does not overload the human cognitive capabilities. According to (Chittaro, 2006) to correctly display information on mobile devices, the visualization design rules have to be followed.

AR techniques introduce many challenging tasks especially for mobile devices due to the computational costs and small displays. The main goal of the proposed work is to introduce a novel information visualization technique that allows end-users to access paintings' characters-based information through AR.

The proposed technique is highly innovative and there are no similar works where the end-user is supported by an AR technique to visualize information related to painting characters. In (Dahne and Karigiannis, 2002) the mobile "archeoguide" application has been introduced to reconstruct historical sites. In (Takacs et al., 2008) an outdoor augmented reality system for mobile phones using Loxel-Based visual feature organization is has been proposed.

To reach the proposed goal a dataset of paintings has been built by taking pictures of paintings available from the web. The system exploits the signature computation module and the signature matching module to compute and match paintings' signatures. Given a query painting and the matched painting, the homography estimation module is used to estimate the homography such that information can be displayed through AR. The Human Device Interface (HDI) module has been designed to provide a novel information visualization technique. Three different designs have been proposed and evaluated using standard evaluation methods.

2 SYSTEM DESCRIPTION

As shown in Figure 1, the proposed system exploits four modules to achieve the proposed goal: i) the boundary detector module, ii) the signature computation and signature matching module, iii) the homography computation module and iv) the Human Device Interface module. Before extracting the features and computing the signature the relevant painting region of a given image is extracted using the standard Hough Transform. The signature computation module aligns the image boundaries such that they are orthogonal to each other and extracts visual based features from such relevant regions to compute a discriminative signature. The computed signatures are stored in a dataset that is used for comparison with real-time acquired images. The signature matching module computes the distance between signatures to detect the best match. Given a match between feature vectors the homography computation module estimates the homogrpahy transformation between the
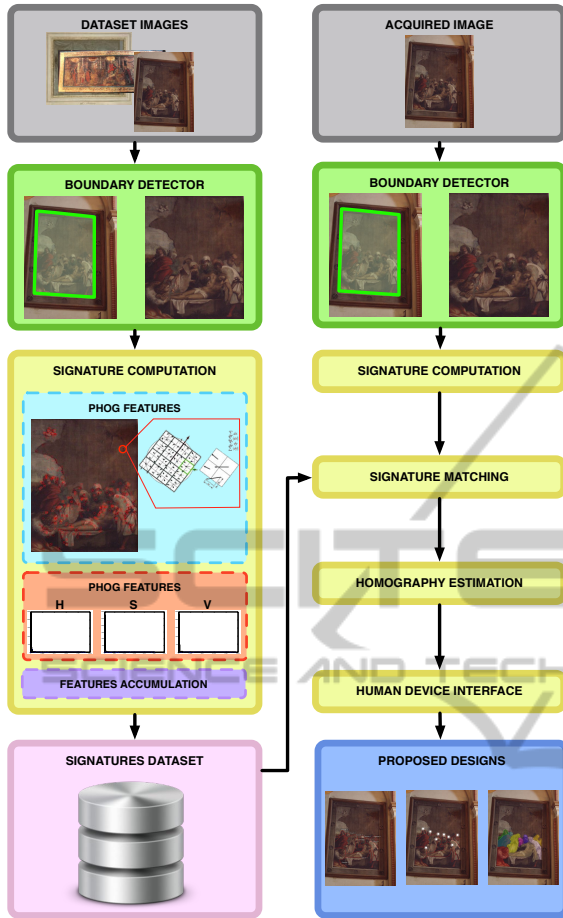
Figure 1: System overview. Each image is processed by the boundary detector module and the detected relevant painting region is used to compute the painting signature. The real-time acquired image is processed to compute the signature and to match it with the dataset signatures. Matching signatures and features are used to compute the homography transformation used to display content through AR.

original image and the acquired image. The estimated homography transformation is finally exploited by the HDI module. This module is used to overlap content-based information to the currently acquired images through AR.

3 SIGNATURE COMPUTATION AND MATCHING

The signature computation module extracts image features to compute the discriminative signature. Given the relevant painting region R of image I two features are computed: i) Speeded-Up Robust Features (SURF) and ii) Pyramid of Histogram of Oriented Gradients (PHOG). Before extracting the fea-

tures the transformation $R' = TR$ is applied such that the regions boundaries are orthogonal to each other. Then R' is projected to the HSV color space to achieve illumination and color invariance. SURF features $F^{(I)}$ are extracted from the given painting region R' as the sum of the Haar wavelet response around each detected point of interest. The PHOG feature matrix $P^{(I)} \in \mathbb{R}^{c \times 3}$ is extracted by concatenating the PHOG histograms extracted from the three image channels. c is the total number of histogram bins. The signature $S^{(I)}$ of image I is defined as $S^{(I)} = \langle F^{(I)}, P^{(I)} \rangle$.

The *signature matching module* matches a query signature $S^{(Q)}$ with the dataset signatures. Let $S^{(Q)}$ and $S^{(I)}$ be the paintings signatures of the query painting Q and the dataset painting I, respectively. Let q, i be a match between two SURF feature descriptors such that the L^2 norm distance between $F_q^{(Q)}$ and $F_i^{(I)}$ is lower than a fixed threshold Th. $F_q^{(Q)}$ and $F_i^{(I)}$ are the q-th and the i-th SURF features of the two given signatures. The overall SURF features distance is computed as

$$d_{SURF}(F^{(Q)}, F^{(I)}) = \frac{1}{m}(d_{L^2}(F_q^{(Q)}, F_i^{(I)}) < Th) \quad (1)$$

where m is the total number of matching SURF features.

PHOG features are matched through a weighted χ^2 distance as proposed in (Martinel and Micheloni, 2012). Given the PHOG feature matrices of two signatures $P^{(Q)}$ and $P^{(I)}$, the PHOG distance is computed as

$$d_{PHOG}(P^{(Q)}, P^{(I)}) = \sum_i \lambda_i \chi^2(P_i^{(Q)}, P_i^{(I)}) \quad (2)$$

where $P_i^{(Q)}$ and $P_i^{(I)}$ are the PHOG feature vectors computed for the signatures $P^{(Q)}$ and $P^{(I)}$ on channel i. λ_i is the normalization weight.

Finally a match between a query signature $S^{(Q)}$ and a dataset signature $S^{(I)}$ is computed as

$$\arg\min_I d(S^{(Q)}, S^{(I)}) = \alpha d_{SURF}(F^{(Q)}, F^{(I)}) \quad (3)$$
$$+ \beta d_{PHOG}(P^{(Q)}, P^{(I)})$$

α and β are the normalization weights.

4 HOMOGRAPHY ESTIMATION

The homography estimation module is used to compute the homography transformation matrix that allows overlap painting characters information to the device display. The homography estimation is achieved by exploiting a feature matching technique.

Let Q and I be the query image and the dataset matching painting. Given all the matches q, i between SURF feature descriptors the goal is to estimate the homography transformation $H^{Q,I}$ such that $K_i^{(I)} = H^{Q,I} K_q^{(Q)}$. $K_i^{(I)}$ and $K_q^{(Q)}$ are the detected SURF interest points of the matching features $F_q^{(Q)}$ and $F_i^{(I)}$. The approach proposed in (Brown and Lowe, 2006) is exploited to achieve such goal.

Given a point in the original coordinate frame of the dataset image Q, the inverse transformation matrices $H^{Q,I^{-1}}$ and T^{-1} can be used to display it onto the image region R.

5 HUMAN DEVICE INTERFACE

Given the computed inverse homography transformations, the HDI module is used to display the information related to a painting character through AR. Standard Human-Computer Interaction methods have been used to find the correct way display the information such that users can easily interact with the user interface without any cognitive effort. Three different user interfaces have been designed and evaluated respecting the usability rules.

The three proposed user interfaces have been designed as follows: i) painting characters edges are highlighted with the same color and their names are shown close to themselves. The end-user can access character information by selecting the displayed label. ii) painting characters edges are highlighted with the same color as before, but name labels are replaced by blinking white circles. The end-user has to select the circle to access the character information. iii) characters edges are displayed with different colors, and characters silhouettes are overlapped with semi-transparent coloured and blinking silhouettes. The end-users have to select the semi-transparent coloured silhouette to access the character information.

6 EXPERIMENTAL RESULTS

To evaluate the proposed designs two type of tests have been performed: i) inspection tests and ii) end-user tests. Inspection tests have been performed by usability experts without the direct involvement of the end-users.

Two types of inspection tests have been exploited: i) *heuristics* tests, i.e., analytical evaluation techniques that provide opinions, and ii) *cognitive walkthrough* tests where the HCI experts examine the ele-
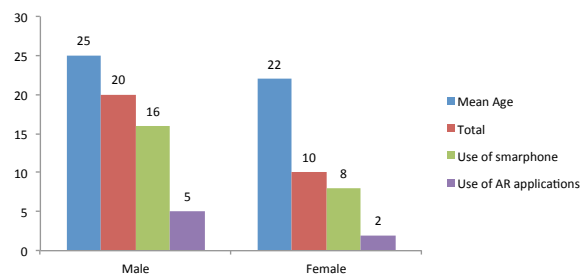

Figure 2: Testers' profiles. 30 testers have been selected to perform the required task.

mentary actions that end-user needs to take to achieve a given goal.

The proposed system has been evaluated with a total of 30 users (Figure 2) without loss of generalization (Nielsen and Landauer, 1993). During the briefing participants were informed about the purposes of the test, the task and its duration. Users were also asked to fill a screening questionnaire to get information about them. The "think-aloud" technique has been used for test sessions, each of which lasts about fifteen minutes. After each test a debriefing is exploited to investigate unusual or interesting events that occurred.

The first user interface has been designed as shown in Figure 3(a). Six participant out of ten completed the given task with an average execution time of 8'33". As shown in Figure 4, 25% of the users that failed to complete the task selected different areas other than the character labels; 25% selected the menu button; and the remaining 50% didn't complete the task at all. After debriefing, 90% of the participants was satisfied about the application but 40% of them stated that the user interface was not clear.

The second designed user interface is shown in Figure 3(b). Only one tester out of ten failed the test selecting the menu button. 90% of the testers stated that the proposed user interface was clear and it was easy to reach the information related to a character. One single tester suggested to display the white circles with different colors. As shown in Figure 5 the second designed user interface achieves the best performance both in terms of success rate and average execution times. The average execution time required to complete the task was about 4.1".

The third designed user interface (Figure 3(c)) achieved the worst results. Only one tester successfully completed the task. According to debriefing questionnaire inspection, 70% of the participants stated that the interface was not clear and 80% of them had difficulties in recognizing the silhouette as a selectable element. Most of the testers agreed that the character recognition task was difficult due the over-

Figure 3: In (a) painting characters edges are highlighted and characters names are used as selectable elements to access character information. In (b) painting characters edges are highlighted as (a). Selectable characters names are replaced by blinking white circles. In (c) painting characters edges and silhouettes are highlighted using different colours. Characters silhouettes are used as selectable elements.
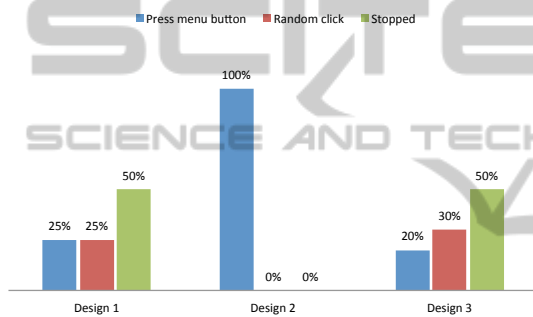

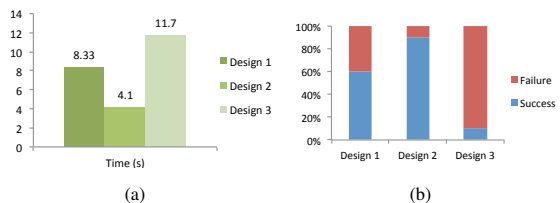Figure 4: Most relevant user interface issues.

Figure 5: Evaluation results: (b) average execution time; (a) success and failures.

lapping of the colored silhouette.

Figure 4 and Figure 5 show that, for the third user interfaces, testers that failed to complete the task stopped before completing it or they randomly touched the screen or selected the menu button.

7 CONCLUSIONS

The proposed work introduced a novel information visualization for mobile devices that allows end-users to access painting character-based information through AR. A set of paintings signatures is computed extracting two different types of features from the de-

tected relevant painting regions. The real-time computed signatures are matched to the dataset signatures using a weighted distance. Matching signatures are used to estimate the homography transformation that allows to display content through AR. Three different designs have been evaluated to propose a novel information visualization technique. Each design has been evaluated using standard Human Computer-Interaction evaluation techniques.

REFERENCES

Brown, M. and Lowe, D. G. (2006). Automatic Panoramic Image Stitching using Invariant Features. *International Journal of Computer Vision*, 74(1):59–73.

Chittaro, L. (2006). Visualizing Information on Mobile Devices. *Computer*, 39(3):40–45.

Dahne, P. and Karigiannis, J. (2002). Archeoguide: system architecture of a mobile outdoor augmented reality system. In *International Symposium on Mixed and Augmented Reality*, pages 263–264.

Marois, R. and Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in cognitive sciences*, 9(6):296–305.

Martinel, N. and Micheloni, C. (2012). Re-identify people in wide area camera network. In *Conference on Computer Vision and Pattern Recognition Workshops*, pages 31–36, Providence, RI.

Nielsen, J. and Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *SIGCHI conference on Human factors in computing systems*, pages 206–213.

Takacs, G., Chandrasekhar, V., Gelfand, N., Xiong, Y., Chen, W.-C., Bismpigiannis, T., Grzeszczuk, R., Pulli, K., and Girod, B. (2008). Outdoors augmented reality on mobile phone using loxel-based visual feature organization. *ACM International Conference on Multimedia information retrieval*, page 427.