Exploring the potential contribution of mobile eye-tracking technology in enhancing the museum visit experience

Moayad Mokatren
The University of Haifa
Mount Carmel, Haifa, 31905
+97248288511
moayad.mokatren@gmail.com

Tsvi Kuflik
The University of Haifa
Mount Carmel, Haifa, 31905
+97248288511
tsvikak@is.haifa.ac.il

ABSTRACT

An intelligent mobile museum visitors' guide is a canonical case of a context-aware mobile system. Museum visitors move in the museum, looking for interesting exhibits, and wish to acquire information to deepen their knowledge and satisfy their interests. A smart context-aware mobile guide may provide the visitor with personalized relevant information from the vast amount of content available at the museum, adapted for his or her personal needs. Earlier studies relied on using sensors for location-awareness and interest detection. This work explores the potential of mobile eye-tracking and vision technology in enhancing the museum visit experience. Our hypothesis is that the use of the eye tracking technology in museums' mobile guides can enhance the visit experience by enabling more intuitive interaction. We report here on satisfactory preliminary results from examining the performance of a mobile eye tracker in a realistic setting – the technology has reached a reliable degree of maturity that can be used for developing a system based on it.

Author Keywords

Mobile guide; Mobile eye tracking; Personalized information; Smart environment; Context aware service.

ACM Classification Keywords

H.5.2. Input devices and strategies (e.g., mouse, touchscreen)

1. INTRODUCTION

The museum visit experience has changed over the last two decades. With the progress of technology and the spread of handheld devices, many systems were developed to support the museum visitor and enhance the museum visit experience. The purpose of such systems was to encourage the visitors to use devices that provide multimedia content rather than use guide books, and as a consequence focus of the exhibits instead of flipping through pages in a guide book, as surveyed in [Ardissono et al. 2012].

Understanding the museum visitors' motivations plays a crucial role in the development and designing of systems that support their needs and could enhance their visit experience.

Falk and Dierking [2000] and Falk [2009] tried to answer the question of what do visitors remember from their visit and what factors seemed to most contribute to visitors' forming of long-terms memories: "when people are asked to recall their museum experiences, whether a day or two later or after twenty or thirty years, the most frequently recalled and persistent aspects relate to the physical context-memories of what they saw, what they did, and how they felt about these experiences.". Stock et al. [2009], and Dim and Kuflik [2014] explored the potential of novel, mobile technology in identifying visitors behavior types in order to consider what/how/when to provide them with relevant services.

A key challenge in using mobile technology for supporting museum visitors' is figuring out what they are interested in. This may be achieved by tracking where the visitors are and the time they spend there [Yalowitz and Bronnenkant, 2009]. A more challenging aspect is finding out what exactly they are looking at [Falk and Dierking, 2000]. Given todays' mobile devices, we should be able to gain access seamlessly to information of interest, without the need to take pictures or submit queries and look for results, which are the prevailing interaction methods with our mobile devices. As we move towards "Cognition-aware computing" [Bulling and Zander 2014], it becomes clearer that eye-gaze based interaction may play a major role in human-computer interaction before/until brain computer interaction methods will become a reality [Bulling et al. 2012]. The study of eye movements started long almost 100 years ago, Jacob and Karn [2003] presented a brief history of techniques that were used to detect eye movements, the major works dealt with usability researches, one of the important works started in 1947 by Fitts and his colleagues [Fitts et al. 1950] when they began using motion picture cameras to study the movements of pilots' eyes as they used cockpit control and instruments to land an airplane. "It is clear that the concept of using eye tracking to shed light on usability issues has been around since before computer interfaces, as we know them" [Jacob and Karn 2003]. Certain mobile eye tracking devices that enables to detect what someone is looking at and stores the data for later use and analysis, have been developed and could be found in the market nowadays [Hendrickson et al. 2014]. In recent years, eye tracking and image based object recognition technology have reached a reliable degree of maturity that can be used for developing a system based on it, precisely identifying what the user is looking at [Kassner et al. 2014]. We shall refer to this field by reviewing techniques for image matching and extend them for location-awareness use and we will follow the approach of "What you look at is what you get" [Jacob 1991].

With the advent of mobile and ubiquitous computing, it is time to explore the potential of this technology for natural, intelligent interaction of users with their smart environment, not only in specific tasks and uses, but for a more ambitious goal of integrating eye tracking into the process of inferring mobile users' interests and preferences for providing them with relevant services and enhancing their user models, an area that received little attention so far. This work aims at exploring the potential of mobile eye tracking technology in enhancing the museum visit experience by integrating and extending these technologies into a mobile museum visitors' guide system, so to enable using machine vision for identifying visitors' position and their object of interest in this place, as a trigger for personalized information delivery.

2. BACKGROUND

2.1 Museum visitors and their visit experience

Understanding who visits the museum, their behaviors and the goal of the visit can play an important role in the design of museums' mobile guide (and other technologies) that enhances the visit experience, "the visitors' social context has an impact on their museum visit experience. Knowing the social context may allow a system to provide socially aware services to the visitors." [Bitgood 2002; Falk 2009; Falk and Dierking 2000; Leinhardt and Knutson 2004; Packer and Ballantyne 2005]. Falk [2009] argued that many studies have been done on who visits museums, what visitors do in the museum and what visitors learn from the museum, and tried to understand the whole visitor and the whole visit experience as well as after the visit. Furthermore, he proposed the idea of visitors "identity" and identified five, distinct, identity-related categories:

- *Explorers: Visitors who are curiosity-driven with a generic interest in the content of the museum. They expect to find something that will grab their attention and fuel their learning.*
- *Facilitators: Visitors who are socially motivated. Their visit is focused on primarily enabling the experience and learning of others in their accompanying social group.*
- *Professional/Hobbyists: Visitors who feel a close tie between the museum content and their professional or hobbyist passions. Their visits are typically motivated by a desire to satisfy a specific content-related objective.*
- *Experience Seekers: Visitors who are motivated to visit because they perceive the museum as an important destination. Their satisfaction primarily derives from the mere fact of having 'been there and done that'.*
- *Rechargers: Visitors who are primarily seeking to have a contemplative, spiritual and/or restorative experience. They see the museum as a refuge from the work-a-day world or as a confirmation of their religious beliefs.*

In addition, he argued that the actual museum visit experience is strongly shaped by the needs of the visitor's identity-related visit motivations, and the individual's entering motivations creates a basic trajectory for the visit, though the specifics if what the visitor actually sees and does are strongly influenced by the factors described by the Contextual Model of Learning:

- Personal Context: The visitor's prior knowledge, experience, and interest.
- Physical Context: The specifics of the exhibitions, programs, objects, and labels they encounter.
- Socio-cultural Context: The within-and between-group interactions that occur while in the museum and the visitor's cultural experiences and values.

Nevertheless the visitor perceives his or her visit experience to be satisfying if this marriage of perceived identity-related needs and museum affordance proves to be well-matched. Hence, considering the use of technology for supporting visitors and enhancing the museum visit experience, it seems that these aspects need to be addressed by identifying visitors' identity and providing them relevant support.

2.2 Object recognition and image matching

Modern eye trackers usually record video by a front camera of the scenes for further analysis [Kassner et al. 2014]. Object recognition is a task within computer vision of finding and identifying objects in an image or video sequence. Humans recognize a multitude of objects in images with little effort, despite the fact that the image of the objects may vary somewhat in different viewpoints, in many different sizes and scales or even when they are translated or rotated. Objects can even be recognized when they are partially obstructed from view. This task is still a challenge for computer vision systems [Pinto et al. 2008]. Many approaches to the task have been implemented over multiple decades. For example, diffusing models to perform image-to-image matching [Thirion 1998], parametric correspondence technique [Barrow 1977] and The Adaptive Least Squares Correlation [Gruen 1985] were presented as a techniques for image matching. Techniques from [Naphade et al. 1999], [Hampapur et al. 2001] and [Kim et al. 2005] were presented for image sequence matching (video stream). A related field is visual saliency or saliency detection, "it is the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention." [Laurent 2007]. Goferman et al. [2012] proposed a new type of saliency which aims at detecting the image regions that represent the scene. In our case, we can exploit the use of eye tracking to detect salience in an efficient way since we have fixation points representing points of interests in a scene.

3. RELATED WORK

As mentioned above, many studies were conducted in detecting eye movements before considering their integration with computer interfaces, as we know them today. The studies have been around HCI and usability and techniques were presented that can be extended for further eye tracking studies and not just in the field of HCI. Jacob [1991] presented techniques for local calibrating of an eye tracker, which is a procedure of producing a mapping of the eye movements' measures and wandering in the scene measures. In addition, he presented techniques for fixation recognition with respect to extracting data from noisy, jittery, error-filled stream and for addressing the problem of "Midas touch" where people look at an item without having the look "mean" something. Jacob and Karn [2003] presented a list of promising eye tracking metrics for data analysis:

- Gaze duration - cumulative duration and average spatial location of a series of consecutive fixations within an area of interest.
- Gaze rate – number of gazes per minute on each area of interest.
- Number of fixation on each area of interest.
- Number of fixation, overall.
- Scan path – sequence of fixations.
- Number of involuntary and number of voluntary fixations (short fixations and long fixations should be defined well in term of millisecond units).

Using handheld devices as a multimedia guidebook in museums has led to improvement in the museum visit experience. Researches have confirmed the hypothesis that a portable computer with an interactive multimedia application has the potential to enhance interpretation and to become a new tool for interpreting museum collections [Evans et al. 2005, Evans et al. 1999, Hsi 2003]. Studies about integration of multimedia guidebooks with eye tracking have already been made in the context of museums and cultural heritage sites. Museum Guide 2.0 [Toyama et al. 2012] was presented as a framework for delivering multimedia content for museum's visitors which runs on handheld device and uses the SMI viewX eye tracker and object recognition techniques. The visitor can hear audio information when detecting an exhibit. A users' study was conducted in a laboratory setting, but not in a real museum. We plan to extend this work by integrating an eye tracker into real museum visitors' guide system and experiment it is realistic setting.

Brône et al. [2011] have implemented effective new methods for analyzing gaze data collected with eye-tracking device and how to integrate it with object recognition algorithms. They presented a series of arguments why an object-based approach may provide a significant surplus, in terms of analytical precision. Specifically they discussed solutions in order to reduce the substantial cost of manual video annotation of gaze behavior, and have developed a series of proof-of-concept case studies in different real world situations, each with its own challenges and requirements. We plan to use their lessons in our study. Pfeiffer et al. [2014] presented "EyeSee3D", where they combined geometric modelling with inexpensive 3D marker tracking to align virtual proxies with the real-world objects. This allowed classifying fixations on objects of interest automatically while supporting a free movement of the participant. During the analysis of the accuracy of the pose estimation they found that the marker detection may fail from several reasons: First, sometimes the participant looked sideways and there simply was no marker within view. More often, however, swift head movements or extreme position changes were causing these issues. Ohm et al. [2014] tried to find where people look at, when navigating in a large scale indoor environment, and what objects can assist them to find their ways. They conducted a user study and assessed the visual attractions of objects with an eye tracker. Their findings show that functional landmarks like doors and stairs are most likely to be looked at. In our case we can use these landmarks as reliable points of interest that can be used for finding the location of the visitor in the museum. Beugher et al. [2014] presented a novel method for the automatic analysis of mobile eye-tracking data in natural environment for object recognition. The obtained results were satisfactory for most of the objects. However, a large scale variance results in a lower detection rate (for objects which were looked at both from very far away and from close by.)

Schrammel et al. [2011, 2014] studied attentional behavior of users on the move. They discussed the unique potential and challenges of using eye tracking in mobile settings and demonstrated the ability to use it to study the attention on advertising media in two different situations: within a digital display in public transport and towards logos in a pedestrian shopping street as well as ideas about a general attention model based on eye gaze. Kiefer et al. [2014] also explored the possibility of identifying users' attention by eye tracking in the setting of tourism – when a tourist gets bored looking at a city panorama – this scenario may be of specific interest for us as locations or objects that attracted more or less interest may be used to model user's interest and trigger further services/information later on. Nakano and Ishii (2010) studied the use of eye gaze as an indicator for user engagement, trying also to adapt it to individual users. Engagement may be used as an indicator for interest and the ability to adapt engagement detection to individual users may enable us also to infer interest and build/adapt a user model using this information. Furthermore, Ma et al. [2015] demonstrated an initial ability to extract user models based on eye gaze of users viewing videos. Xu et al. [2008] also used eye gaze to infer user preferences in the content of documents and videos by the users attention as inferred from gaze analysis (number of fixations on word/image).

As we have seen, there is a large body of work about monitoring and analyzing users' eye gaze in general and also in cultural heritage setting. Moreover, the appearance of mobile eye trackers opens up new opportunities for research in mobile scenarios. It was also demonstrated in several occasions that eye gaze may be useful in enhancing a user model, as it may enable to identify users' attention (and interests). Considering mobile scenarios, when users also carry smartphones - equipped with various sensors - implicit user modeling can take place by integrating signals from various sensors, including the new sensor of eye-gaze for better modeling the user and offering better personalized services. So far sensors like GPS, compass, accelerometers and voice detectors were used in modeling users' context and interests, (see for instance [Dim & Kuflik. 2014]). When we mention mobile scenarios, we refer to a large variety of different scenarios – pedestrians' scenario differs from jogging or shopping or cultural heritage scenario. The tasks are different and users' attention is split differently. The cultural heritage domain is an example where users have long term interests that can be modeled and the model can be used and updated during a museum visit by information collected implicitly from various sensors, including eye-gaze. In this sense, the proposed research extends and aims at generalizing the work of Kardan and Conati [2013]. Still, even though a lot of research effort was invested in monitoring, analyzing and using eye gaze for inferring user interests, so far, little research attention was paid to users gazing behavior "on the go". This scenario poses major challenges as it involves splitting attention between several tasks at the same time – avoiding obstacles, gathering information and paying attention to whatever seems relevant, for many reasons.

4. RESEARCH GOAL AND QUESTIONS

Our goal is to examine the potential of integrating the eye tracking technology with a mobile guide for a museum visit and try to answer the question: **How can the use of mobile eye tracker enhance the museum visit experience?** Our focus will be on developing a technique for location awareness based on eye gaze detection and image matching, and integrate it with a mobile museum visitor's guide that provides multimedia content to the visitor. For that we will design and develop a system that runs on handheld device and uses Pupil Dev [Kassner et al. 2014] eye tracker for identifying objects of interest and delivering multimedia content to visitor in the museum. Then we will evaluate the system in a user study in a real museum to find out how the use of eye tracker integrated with a multimedia guide can enhance the museum visit experience. In our study, we have to consider different factors and constraints that may affect the performance of the system, such as the real environment lighting conditions which are different from laboratory conditions and can greatly affect the process of object recognition. Another aspect may be the position of the exhibits relative to the eye tracker holder, since the eye tracker device is mounted as this is constrained by the museum layout. While having many potential benefits, a mobile guide can also have some disadvantages [Lanir et al, 2013]. It may focus the visitor's attention on the mobile device rather than on the museum artifacts [Grinter et al, 2002]. We will also examine this behavior and try to review whether the use of eye tracker in mobile guide can increase the looking time at the exhibits. In addition, we will try to build a system that runs in various real environments with different factors and have the same constraints such as the light and the position constraints.

5. TOOLS AND METHODS

A commercial mobile eye tracker will be integrated into a mobile museum visitors' guide system as a tool for location awareness, interest detection and focus of attention by using computer vision techniques. Our hypothesis is that the use of the eye tracker in mobile guides can enhance the visit experience. The system will be evaluated in user studies, the participants will be students from University of Haifa. The study will be conducted in Hecht museum[1], which is a small museum, located at the University of Haifa that has both an archeological and art collections. The study will include an orientation about using the eye tracker and the mobile guide, then taking a tour with the eye tracker and handheld device, multimedia content will be delivered by showing information on the screen or by listening to audio by earphones. Data will be collected as follows: The students will be interviewed and asked about their visit experience, and will be asked to fill questionnaires regarding general questions such as if it is the first time that they have visited the museum, their gender and age, and more. Visit logs will be collected and analyzed for later use, we can come to conclusions about the exhibit importance and where the visitors tend to look, the positioning of the exhibits, and the time of the visits or explorations. The study will compare the visit experience when using two different system versions – a conventional one and one with an integrated eye tracker, we will choose the work of [Kuflik et al. 2012] that was conducted in Hecht museum and which uses "light weight" proximity based indoor positioning sensors for location-awareness as a comparison system for examining the user experience.

6. PRELIMINARY RESULTS

It was important to examine the accuracy of eye gaze detection when using the Pupil Dev mobile eye-tracker device. For that, we have conducted several small-scale user studies onsite.

6.1 The Pupil eye tracker

Pupil eye tracker [Kassner et al. 2014] is an open source platform for pervasive eye tracking and gaze-based interaction. It comprises a light-weight eye tracking headset that includes high-resolution scene and eye cameras, an open source software framework for mobile eye tracking, as well as a graphical user interface to playback and visualize video and gaze data. The software and GUI are platform-independent and include algorithms for real-time pupil detection and tracking, calibration, and accurate gaze estimation. Results of a performance evaluation show that Pupil can provide an average gaze estimation accuracy of 0.6 degree of visual angle (0.08 degree precision) with a processing pipeline latency of only 0.045 seconds.

Figure 1. Pupil eye-tracker (http://pupil-labs.com/pupil)

[1] http://mushecht.haifa.ac.il/

6.2 User study 1: Look at a grid cells

Five students from the University of Haifa, without any visual disabilities participated in this study. They were asked to look at wall-mounted grid from a distance of 2 meters and track a finger (see figure 2). On every cell that the finger pointed at, they were asked to look at for approximately 3 seconds. Data was collected for determining the practical measurement accuracy. The results were as follows: on average, fixation detection rate was ~80% (most missed fixations were in the edges/corners – see table 1 for details about misses). In addition, average fixation point error rate, in terms of distance from the center of grids, was approximately 5 cm (exact error rate can be calculated using simple image processing techniques for detecting the green circle and applying mapping transform to the real word).

Figure 2. User study 1. The finger points at the grid where the participant were asked to look at. The green circle is a fixation point given from the eye tracker. The size of each cell is 20x20 cm.

Cell #	6	18	19	23	24
Missed	5	5	3	5	5

Table 1. Experiment details.

During the study we ran into several practical problems. The Pupil Dev eye tracker that we are using is not fitted for every person. The device consists of two cameras, the first for delivering the scene and the second directed to the right eye for detecting fixations. In some cases when the device is not fitted correctly, the vision range got smaller and parts of the pupil got out from the capture frame (see figure 3 for example). As a consequence no fixations were detected. Another limitation was that when using the eye tracker with tall participants, they have to step back from the object which negatively affects the accuracy.

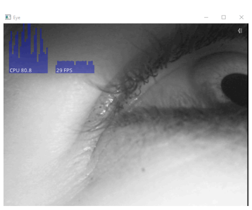

Figure 3. Screen capture from eye camera.

Figure 4. Gallery exhibition

6.3 User study 2: Look at an exhibit

In this study we examined the accuracy of the eye tracker in a realistic setting. One participant (1.79m tall) was asked to look at exhibits at the Hecht museum. Several exhibits where chosen with different factors and constraints (see figure 4, 5, and 6). The main constraint in this case is the distance from the exhibit since the visual range gets larger when the distance grows, and mainly we have to cover all the objects that we are interested in. Table 2 presents the objects height from the floor and the distance of the participant from the object. The next step was to examine fixations accuracy after making sure that the participant is standing in a proper distance. The participant was asked to look at different points in the exhibit/scene. In the gallery exhibits, the scan path has been set to be the four corners of the picture and finally the center of it. Regarding the vitrine exhibits, for each jug one point at the center has been defined

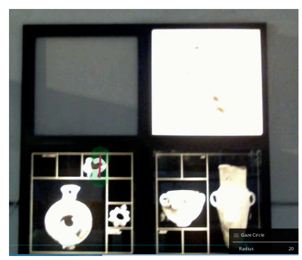

Figure 5. Mounted backlighted images exhibition

It's important to note that the heights/distances relation is for visual range (having the objects in the frame of the camera) and not for fixations detections. Since missed fixations could be as a result of a set of constraints and not the distance from the object, thing that we have not examined yet.

Figure 6. Vitrine backlighted exhibition.

Exhibit type	width (cm)	height (cm)	Height from floor (cm)	Stand distance (cm)
Vitrine	80	25	150	150

Exhibit type	width (cm)	height (cm)	Height from floor (cm)	Stand distance (cm)
shelf	80	15	120	230
	80	20	90	310
	80	15	40	390
Gallery	60	67	150	200

Table 2. Experiment details – we considered the three most left shelves in the vitrine exhibit shown in figure 6.

7. SYSTEM DESIGN

A smart context-aware mobile museum visitors' guide may provide the visitor with personalized relevant information from the vast amount of content available at the museum, adapted for his or her personal needs. Furthermore, the system may provide recommendations and location-relevant information. However, the potential benefit may also have a cost, the notifications may interrupt the user in the current task and be annoying in the wrong context. Beja et al. [2015] examined the effect of notifications in a special leisure scenario - a museum visit. Following Beja et al [2015], we will consider three different scenarios:

I. The Visitor is looking at an exhibit. The region of interest will be defined as the region from the scene around the gaze fixation point. Then object matching procedure will be applied (see section 8). It will enable us to determine both the visitor's position and the object of interest.

II. The visitor is looking at the tablet. This could be done in two ways: 1) the visitor is watching multimedia information, in this scenario there is nothing to do for him. 2) The visitor may need service from the system or a recommendation, so it is the right time to deliver him.

III. The visitor is wandering in the museum. According to Beja et al. [2015], it is the best time for sending notifications.

As a basic system we will use the PIL museum visitor's guide system [Kuflik et al 2012; Lanir et al. 2013]. The system is a context aware, mobile museum visitors' guide system. Its positioning mechanism is based on proximity based RF technology that enables to identify the visitor's position – when the visitor is near a point of interest. As vision is the main sense for gathering information, we plan to replace the system's positioning component with an eye-tracker based positioning and object of interest identification component. Hence we will enhance the positioning system by providing the system the ability to pin-point the object of interest. The rest of the system will remain unchanged. Having these two versions of systems will enable us to compare and evaluate the benefits of the eye-tracker as a positioning and pointing device in the museum.

8. OBJECT MATCHING PROCEDURE

8.1 Data-set preparation

A set of images of the exhibits will be taken, each image may contain one or more objects. Each image will be given a distinct label value and size of region around the object (in terms of width and height – rectangular shape).

8.2 Object matching

The matching procedure will be done in three steps:

1. Eye-tracker scene camera frame is taken (figure 7) and image-to-image matching applied. The result is an image with labeled regions in the current scene's frame (figure 8).

2. Mapping transformation – We need to transform the fixation point in the eye-tracker scene camera to a suitable/matched point in the image that we got in step one (image from the data-set with labeled regions), since the viewpoint of the objects can be different from this in the data set. For example one image is rotated relative to the other or one is zoomed in/out as a result of standing in different distance from the object when the data-set image was taken.

3. Finding the object - This is step is simple since we have a mapped fixation points and labeled regions. What remains is determining for which object the point does it relates (or it relates to nothing).

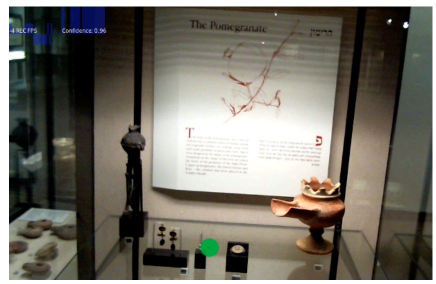

Figure 7. Example of eye-tracker scene camera. The green point is the fixation point.

9. DISCUSSION

We conducted these small-scale user studies in order to gain initial first-hand experience with the eye-tracker in a realistic setting. Furthermore, we tried to clarify which exhibits are appropriate to be included in our future study and, given the limitation of the device, what portion of the museum exhibits may be included in general. Not surprisingly, we got 100% accuracy rate when we examined the device in the art wing since all the pictures are placed in ideal height. Regarding the archeological wing, it is considerably more challenging environment, since objects are placed in different heights and have unequal sizes. As a result the visitor may have to stand far away from the objects in order to get them into the eye-tracker front camera frame, a fact that can negatively affect the visit experience. In the case of archeological wing we approximate that about 60% of the exhibits may be detectable with the current device. Regarding the low-height exhibits we don't know yet whether they can be considered or not. More challenging exhibits are these that are placed in harsh light conditions or placed in low height (see figure 9 for example) and/or these that are too large to fit in one frame (see figure 10 for example).

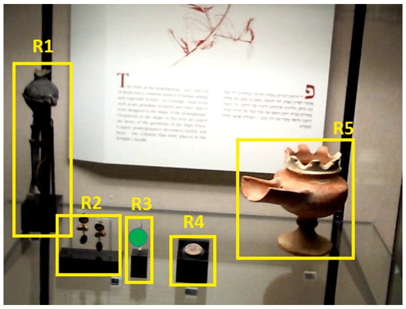

Figure 8. Image-to-image matching. The yellow rectangles are the regions around each object. The green point is the fixation point after performing proposed mapping transformation. The corresponding region would be R3.

Figure 9. A challenging exhibition: harsh light conditions and low-height.

Figure 10. A challenging exhibit: too big to fit in one frame.

10. CONCLUSIONS AND FUTURE WORK

This paper presents a work-in-progress that aims at exploring the potential contribution of the mobile eye tracking technology in enhancing the museum visit experience. For that we have done small-scale experiments in order to get an understanding of the performance of the system in realistic setting. We got satisfactory results from these studies and an understanding of the limitations of the equipment. The next step in the study is to design and build a museum mobile guide that extends the use of mobile eye tracking as a tool for identifying the visitor position and points of interests. We will use the eye-tracker scene camera captures and the collected gaze data to develop a technique for location-awareness. The system will run on tablet, and the multimedia content will be delivered to the participants by listening to audio guide via earphones or by watching slides. Furthermore, knowing exactly where the visitor look in the scene (specific object) will allow us to deliver personalized information. Our research will be a supplement to the nowadays mobile museum guide that uses location-awareness technology and techniques that enhances the visit experience. The system can also be extended and used in other venues such as outdoors cultural heritage sites as well as shopping centers/markets after further validation.

REFERENCES

[1] Ardissono, L., Kuflik, T., & Petrelli, D. (2012). Personalization in cultural heritage: the road travelled and the one ahead. User modeling and user-adapted interaction, 22(1-2), 73-99.

[2] Barrow, H. G., Tenenbaum, J. M., Bolles, R. C., & Wolf, H. C. (1977). Parametric correspondence and chamfer matching: Two new techniques for image matching (No. TN-153). SRI international, Menlo Park CA, Artificial Intelligence center.

[3] Beja, I., Lanir, J. and Kuflik, T. Examining Factors Influencing the Disruptiveness of Notifications in a Mobile Museum Context. *Human–Computer Interaction* just-accepted (2015).

[4] Bitgood, S. (2002). Environmental psychology in museums, zoos, and other exhibition centers. *Handbook of environmental psychology*, 461-480.

[5] Brône, G., Oben, B., Van Beeck, K. Goedemé T. (2011). Towards a more effective method for analyzing mobile eye-tracking data: integrating gaze data with object recognition algorithms. UbiComp '11, Sep 17–Sep 21, 2011, Beijing, China.

[6] Bulling, A., Dachselt, R., Duchowski, A., Jacob, R., Stellmach, S., & Sundstedt, V. (2012). Gaze interaction in the post-WIMP world. In CHI'12 Extended Abstracts on Human Factors in Computing Systems, 1221-1224.

[7] Bulling, A., & Zander, T. O. (2014). Cognition-aware computing. Pervasive Computing, IEEE, 13(3), 80-83.

[8] De Beugher, S., Brône, G., & Goedemé, T. (2014). Automatic analysis of in-the-wild mobile eye-tracking experiments using object, face and person detection. In Proceedings of VISIGRAPP 2014, 1, 625-633.

[9] Dim, E., & Kuflik, T. (2014). Automatic detection of social behavior of museum visitor pairs. *ACM Transactions on Interactive Intelligent Systems*, *4*(4), 17.

[10] Evans, J., Sterry, P. Portable computers & interactive media: A new paradigm for interpreting museum collections. In: Bearman, D., Trant, J. (eds.) Cultural Heritage Informatics 1999: Selected papers from ICHIM 1999, 93–101. Kluwer Academic Publishers, Dordrecht.

[11] Falk, John H., and Lynn D. Dierking. Learning from museums: Visitor experiences and the making of meaning. Altamira Press, 2000.

[12] Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. Aeronautical Eng. Review 9(2), 24–29.

[13] Goferman, S., Zelnik-Manor, L., & Tal, A. (2012). Context-aware saliency detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 34 (10), 1915-1926.

[14] Grinter, R. E., Aoki, P. M., Szymanski, M. H., Thornton, J. D., Woodruff, A., & Hurst, A. (2002). Revisiting the visit: understanding how technology can shape the museum visit. In Proceedings of CSCW 2002,146-155.

[15] Gruen, A. (1985). Adaptive least squares correlation: a powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography, 14(3), 175-187.

[16] Hampapur, A., Hyun, K., & Bolle, R. M. (2001, December). Comparison of sequence matching techniques for video copy detection. In Electronic Imaging 2002 . 194-201. International Society for Optics and Photonics.

[17] Hendrickson, K., & Ailawadi, K. L. (2014). Six lessons for in-store marketing from six years of mobile eye-tracking research. Shopper Marketing and the Role of In-Store Marketing. Review of Marketing Research, 11, 57-74.

[18] Kardan, S., & Conati, C. (2013). Comparing and Combining Eye Gaze and Interface Actions for Determining User Learning with an Interactive Simulation. In proceedings of UMAP 2013, 215-227.

[19] Kassner, M., Patera, W., & Bulling, A. (2014, September). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. 1151-1160. ACM.

[20] Katy Micha, Daphne Economou (2005). Using Personal Digital Assistants (PDAs) to Enhance the Museum Visit Experience. In proceedings of PCI 2005, Volas, Greece, November 11-13, 2005. Proceedings. 188-198.

[21] Kiefer, P., Giannopoulos, I., Kremer, D., Schlieder, C., & Raubal, M. (2014). Starting to get bored: An outdoor eye tracking study of tourists exploring a city panorama. In Proceedings of the Symposium on Eye Tracking Research and Applications (pp. 315-318). ACM.

[22] Kim, C., & Vasudev, B. (2005). Spatiotemporal sequence matching for efficient video copy detection. Circuits and Systems for Video Technology, IEEE Transactions on, 15(1), 127-132.

[23] Kuflik, T., Lanir, J., Dim, E., Wecker, A., Corra, M., Zancanaro, M., & Stock, O. (2012). Indoor positioning in cultural heritage: Challenges and a solution. In Electrical & Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of. 1-5. IEEE.

[24] Lanir, J., Kuflik, T., Dim, E., Wecker, A. J., & Stock, O. (2013). The influence of a location-aware mobile guide on museum visitors' behavior. Interacting with Computers, 25(6), 443-460.

[25] Laurent I.(2007) Visual salience. Scholarpedia, 2(9):3327.

[26] Leinhardt, G., and Knutson K. Listening in on museum conversations. Rowman Altamira, 2004.

[27] Ma, K. T., Xu, Q., Li, L., Sim, T., Kankanhalli, M., & Lim, R. (2015). Eye-2-I: Eye-tracking for just-in-time implicit user profiling. arXiv preprint arXiv:1507.04441.

[28] Nakano, Y. I., & Ishii, R. (2010). Estimating user's engagement from eye-gaze behaviors in human-agent conversations. In Proceedings of IUI 2010.139-148.

[29] Naphade, M. R., Yeung, M. M., & Yeo, B. L. (1999). Novel scheme for fast and efficient video sequence matching using compact signatures. Electronic Imaging. 564-572. International Society for Optics and Photonics.

[30] O. Stock, M. Zancanaro, F. Pianesi, D. Tomasini, and C. Rocchi. 2009. Formative evaluation of a tabletop Display meant to orient casual conversation. Journal of Knowledge, Technology and Policy 22(1), 17–23.

[31] Ohm, C., Müller, M., Ludwig, B., & Bienk, S. (2014). Where is the Landmark? Eye Tracking Studies in Large-Scale Indoor Environments.

[32] Packer, Jan, and Roy Ballantyne. (2005) "Solitary vs. shared: Exploring the social dimension of museum learning." Curator: The Museum Journal 48.2 177-192.

[33] Pinto, Nicolas, David D. Cox, and James J. DiCarlo. "Why is real-world visual object recognition hard?." (2008): e27.

[34] Pfeiffer, T., & Renner, P. (2014). Eyesee3d: A low-cost approach for analyzing mobile 3d eye tracking data using computer vision and augmented reality technology. In Proceedings of the Symposium on Eye Tracking Research and Applications. 369-376. ACM.

[35] ROBERT J. K. JACOB (1991). The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get, ACM Transactions on Information Systems, 9(3), April 1991, 152-169.

[36] Robert J. K. Jacob, Keith S. Karn (2003). Eye Tracking in Human–Computer Interaction and Usability Research: Ready to Deliver the Promises, Elsevier Science BV.

[37] Schrammel, J., Mattheiss, E., Döbelt, S., Paletta, L., Almer, A., & Tscheligi, M. (2011). Attentional behavior of users on the move towards pervasive advertising media. In Pervasive Advertising, 287-307..

[38] Schrammel, J., Regal, G., & Tscheligi, M. (2014). Attention approximation of mobile users towards their environment. In CHI'14 Extended Abstracts on Human Factors in Computing Systems, 1723-1728.

[39] S.Hsi (2003). A study of user experiences mediated by nomadic web content in a museum. The Exploratorium, 3601 Lyon Street, San Francisco, CA 94123

[40] Takumi Toyama, Thomas Kieninger, Faisal Shafait, Andreas Dengel. Gaze guided object recognition using a head-mounted eye tracker. ETRA '12 Proceedings of the Symposium on Eye Tracking Research and Applications, 91-98, ACM New York, NY, USA 2012

[41] Thirion, J. P. (1998). Image matching as a diffusion process: an analogy with Maxwell's demons. Medical image analysis, 2(3), 243-260.

[42] Xu, S., Jiang, H., & Lau, F. (2008). Personalized online document, image and video recommendation via commodity eye-tracking. In Proceedings of RecSys 2008

[43] Yalowitz, S.S. and Bronnenkant, K. (2009) Timing and tracking: unlocking visitor behavior. Visit. Stud., 12, 47–64.

[44] Zhang, Z., Deriche, R., Faugeras, O., & Luong, Q. T. (1995). A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial intelligence, 78(1), 87-119.