ISSUES AND CHALLENGES IN HANDHELD AUGMENTED REALITY APPLICATIONS

Stan Kurkovsky
Central Connecticut State University, Department of Computer Science, New Britain, CT, U.S.A.

Keywords: Mobile Augmented Reality, Requirements, Navigation, Location Tracking, Interaction Design, Usability, Content Creation, User Evaluation.

Abstract: Equipped with powerful processors, cameras for capturing still images and video, and a range of sensors capable of tracking user location, orientation and motion smartphones offer a sophisticated platform for implementing handheld augmented reality applications. Despite the advances in research and development, implementing handheld augmented reality applications remains a challenge due to many unsolved problems related to navigation, context-awareness, visualisation, usability and interaction design, as well as content creation and sharing, which are surveyed in this paper.

1 INTRODUCTION

Augmented reality (AR) refers to a real-time representation of the real world that is digitally augmented by adding graphics, sound or video (van Krevenlen and Poelman, 2010). Handheld augmented reality systems often utilize smartphones equipped with powerful processors, high-resolution cameras, and a range of sensors including Global Positioning System (GPS), accelerometers and magnetometers. Unlike other AR systems, handheld AR applications do not require the users to carry or wear any special equipment and do not constrain the applications to any specialized physical area. Handheld AR systems that utilize location and position information are often used to augment the view of the real world with relevant information about the currently visible points of interest (POI).

Interaction design challenges exemplified by the “magic lens” metaphor present just a small sample of issues and open research problems related to navigation, context-awareness, visualisation and content creation in handheld AR applications. These challenges need to be addressed before AR systems can make a transition from research and academic labs to the domain of everyday users. This paper outlines many of the open design questions that developers and researchers might encounter building handheld AR applications using currently available technologies and tools.

2 HANDHELD AR APPLICATIONS

The architecture of a typical handheld AR application consists of three components: the *mobile AR browser* for end-user interaction, the *AR server* responsible for identifying and querying one or more *POI repositories/servers*. AR browsers provide the user with a choice of information channels; upon selecting a channel, the browser sends a query requesting relevant POIs, which are bounded by the channel selection, current location and a certain spatial range. The AR server acts as a broker and selects an appropriate POI provider/repository to which the query is forwarded. Similarly, POI content is returned to the mobile AR browser via the server. Finally, the mobile AR browser overlays the POI-related content over a real-time view of the physical world.

The task of scene identification to determine the correct location and orientation of the user is fundamental to any AR application and may be implemented on the mobile device, on the AR server, or distributed between the two. Marker-based scene identification techniques rely on previously placed artificial visual tags (e.g. Kan et al, 2009). Non marker-based scene identification relies on computer vision (e.g. Gammeter at al, 2010), geopositioning (e.g. You et al, 2008) or a combination of these two techniques (e.g. Seo et al, 2010).

Kurkovsky S..
ISSUES AND CHALLENGES IN HANDHELD AUGMENTED REALITY APPLICATIONS.
DOI: 10.5220/0003937007990802
In *Proceedings of the 8th International Conference on Web Information Systems and Technologies* (WEBIST-2012), pages 799-802
ISBN: 978-989-8565-08-2

3 ISSUES AND CHALLENGES

Careful examination of the existing AR platforms and applications reveals a number of open research problems, each with a number of alternative solutions offering a specific set of trade-offs. The remainder of this section outlines these open research problems and challenges, along with different promising ways to address them.

3.1 Indoor and Urban Navigation

Continuous localization of the user is a key component of any AR system. The vast majority of outdoor handheld AR systems utilise GPS for navigation because of its wide availability and relatively high accuracy. Indoor navigation for AR systems does not have a similar commonly accepted solution. Satellite signals used by the GPS are too weak or unavailable indoors unless special High Sensitivity GPS (HSGPS) or Ultra-wide Band (UWB) location sensors are used. Furthermore, it has long been recognized that no single sensor technology is currently capable of providing robust tracking with high enough precision both indoors and outdoors (Welch and Foxlin, 2002). Deploying a specialized hardware infrastructure could be costly and unfeasible, in which case developers of handheld AR systems may resort to using the sensors already available on the mobile device. Images or video captured with the built-in camera can be processed to recognize the features of indoor environment or previously placed QR (or similar) codes (e.g., Kan et al, 2009). Multiple WiFi signal triangulation could be used for approximate localization (Arth, et al, 2009). Finally, localization can be achieved by combining sparsely placed 'info points' whose precise location is known, accelerometer and compass data, with activity-based instructions, such as "walk five steps and turn right" (Mulloni et al, 2011). Gee et al, 2011, describe an approach where GPS and UWB-based location sensing is combined with vision-based tracking that offers a reliable platform for both indoor and outdoor handheld AR applications.

3.2 Computer Vision-based Tracking

Although a tracking solution based on computer vision could offer the best precision, real-time object recognition from a live video feed may be too taxing for a smartphone CPU. Wither et al, 2011, propose a compromise solution, Indirect AR, which replaces a true AR based on the live camera feed with a previously captured panoramic view of the environment. A solution suggested by Gammeter at al, 2010, suggests using a remote server to split the tasks of object tracking and recognition: tracking is performed on the mobile device that periodically sends still images to the server, which is responsible for object recognition. Such an approach could have several advantages: instead of keeping a database on the device, objects can be retrieved from large server-side databases in close to real-time; the bandwidth usage is reasonable since only still images are transmitted to the server instead of a constant video feed. An approach suggested by Takacs et al, 2008, performs on-device object recognition using a local database of previously captured location-tagged images, which helps to limit the search only to the objects in the close proximity to the user. In case if no match is found, the system offers an option to send the image to the server along with a label describing the relevant POI. It is possible to extend this approach by equipping the server with a larger image database and/or a more robust content-based image retrieval algorithm that would be impractical to implement on a mobile device. Unlike GPS-based tracking, computer vision could offer accurate information about the user location, as well as the pose of the user, with a refresh rate exceeding that of a GPS-based solution. Langlotz et al, 2010, propose a computer vision-based solution that enables high-precision tracking and object registration without the need to construct a 3D object database. Instead, this approach takes advantage of natural-feature mapping performed on the device that enables tracking with three degrees of freedom. Natural features of the surrounding environment are mapped to the panoramic view captured by the device in real time.

3.3 Content Creation

In many existing handheld AR systems, only the application developers can add new content because this requires access to the application backend along with programming skills for linking existing systems to he data sources. A truly mobile AR system would allow regular users, such as tourists and small business owners, to add their own content on the go with a minimal technical effort. Such a system could also provide an easy way for the users to mash up user-created content from multiple sources into a uniform handheld AR view. Belimpasakis et al, 2010 describe a handheld AR system that addresses these concerns by creating a generic Mixed Reality Web Service Platform enabling users to geo-register new content without a substantial expertise in AR systems.

Platforms like Wikitude and Layar help solving the problem of location tracking and visualization. However, AR applications will not be able to gain much traction with the end users without a broad availability of diverse sources of content. Active user participation in content authoring is leading the evolution of the Worldwide Web. A similar trend could be applied to the AR applications. Schmalstieg et al, 2010, introduce the concept of AR 2.0 or Social AR, in which regular user can actively participate and create their own content instead of only consuming the content authored by a select group of professional AR modellers and developers. Langlotz et al, 2011, describe a handheld AR system for on-the-go, on-device content authoring and sharing. Using this system, end users can create 2D and 3D content on a mobile device and publish it to their private library on a remote server that supports ARML (described below). Users are then free to share this content with others or reuse the objects they created for marking other real world locations.

3.4 Integration and Reuse of Content

In a typical AR application utilizing multiple POI repositories, the AR server acts as the only point of interaction between the POIs from different repositories. For example, the only connection between a bus stop and a nearby restaurant will be their close proximity that will only become apparent when the AR server processes both sets of POIs. There is no logical or symbolic relationship between such two POIs, although it could be of a great benefit. A possible solution to this problem could be to utilize the Linked Open Data (LOD) principles (Berners-Lee, 2007), which suggest using URIs as names for all data elements, including POIs, as well as cross-referencing among them. Augmented Reality Markup Language (ARML) used by Wikitude provides a native LOD support and it is gaining traction among AR system developers as the Open Geospatial Consortium uniting over 440 international industry, government and academic organizations has established the ARML 2.0 Standards Working Group in September 2011.

3.5 Using Context Information

One of the key features of AR applications is the ability to present a subset of available information in the current geospatial context. Research in context awareness focuses on creating intelligent systems that can adapt to the surrounding environment and the user behaviour, thereby reducing information overload and providing the user with the services and information are relevant in the current context. Although all AR systems take advantage of the user location context, it should be possible to provide the users with a more personalized experience by utilizing other contextual dimensions, such as user intention based on the past behavioural profile (e.g., Lee and Woo, 2008). In addition to improving the level of personalization, context-awareness in handheld AR applications could facilitate sharing of personalized content and social collaboration among the users (Suh et al, 2007).

3.6 Usability Issues

Current applications address only the most obvious and simplest challenges that could be solved by handheld AR systems. Nack, 2010, notes that many of them take advantage mainly of the contextualized user position and orientation, provided that the correct information channel is available. Smartphone GPS sensors have the accuracy of about 20 meters, while the magnetometers enable compass orientation within about 20 degrees. This could lead to problems with calculating the correct camera field of view making real and digital objects not perfectly aligned. Consequently, current mobile AR systems may not offer the precision necessary to identify the specific location of the entrance door or even distinguish between different entrances to a building.

Although modern smartphones are equipped with high-resolution cameras, they provide a limited field of view, which is significantly smaller than that of the human eye. Consequently, current handheld AR applications can only augment a small portion of the mobile user's field of view The "magic lens" design of the current handheld AR applications requires the user to stretch out their hand while holding the device and pointing it in various directions. This problem could possibly be resolved by "freezing" the augmented view to allow the user to see it in a more comfortable position.

In order to see the augmented view of real-world objects that are currently to the sides or behind the user, the user needs to either change their orientation or use a mini-map showing all nearby POIs that is typically displayed on the screen by many current handheld AR applications. Having to rotate around while holding the phone in an outstretched hand may be rather awkward, while interpreting the POIs on the mini-map and matching them to the augmented view and the rest of the unfamiliar real-world surroundings might require a substantial mental effort. Schinke et al, 2010, suggests using arrows embedded in the AR view to point at the surrounding off-screen POIs, which can make the task of interpreting such information much less demanding for the user.

4 SUMMARY

Although the concept of AR was first developed over four decades ago, wide availability of mobile devices with adequate processing power and a multitude of sensors is attracting an increased interest to handheld AR applications. However, many ongoing research projects and off-the-shelf handheld AR solutions limit themselves to leveraging only the user geographic location and orientation information. Today, smartphones are already quite capable of providing tracking services using computer vision algorithms, fusing different methods of location tracking for robust indoor and outdoor navigation, providing tools for easy on-device content creation and sharing, leveraging user and location context, using heterogeneous sources of POI and other data, and providing more unobtrusive user interaction than what is currently offered by the existing handheld AR applications.

REFERENCES

Arth, C., Wagner, D., Klopschitz, M., Irschara, A., Schmalstieg, D., 2009. Wide Area Localization on Mobile Phones, *Int'l Symp. on Mixed and Augmented Reality*, IEEE Comp. Soc.

Belimpasakis, P., You, Y., Selonen, P., 2010. Enabling Rapid Creation of Content for Consumption in Mobile Augmented Reality. In *4th Int'l Conf. on Next Generation Mobile Applications, Services and Technologies (NGMAST)*, IEEE Comp. Soc.

Berners-Lee, T., 2007. Linked Data. Retrieved from http://www.w3.org/DesignIssues/LinkedData.

Gammeter, S., Gassmann, A., Bossard, L., Quack, T., Van Gool, L., 2010. Server-Side Object Recognition and Client-Side Object Tracking for Mobile Augmented Reality, *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, IEEE Comp. Soc.

Gee, A., Webb, M., Escamilla-Ambrosio, J., Mayol-Cuevas, W., Calway, A., 2011. A Topometric System for Wide Area Augmented Reality. *Computers & Graphics*, 35(4):854–868, Elsevier.

Kan, T., Teng, C., Chou, W., 2009. Applying QR Code in Augmented Reality Applications. *8th Int'l Conf. on Virtual Reality Continuum and its Applications in Industry*, ACM Press.

van Krevenlen, D., Poelman, R., 2010, A Survey of Augmented Reality Technology, Applications and Limitations, *The Int'l Journal of Virtual Reality*, 9(2):1-20, IPI Press.

Langlotz, T., Mooslechner, S., Zollmann, S., Degendorfer, C., Reitmayr, G., Schmalstieg, D., 2011. Sketching up the World: In Situ Authoring for Mobile Augmented Reality, *Personal and Ubiquitous Computing*, DOI:10.1007/s00779-011-0430-0, Springer.

Langlotz, T., Wagner, D., Mulloni, A., Schmalstieg, D., 2010. Online Creation of Panoramic Augmented Reality Annotations on Mobile Phones, *Pervasive Computing*, to appear, IEEE Comp. Soc.

Lee, W., Woo, W., 2008. Exploiting Context-Awareness in Augmented Reality Applications, *Int'l Symp. on Ubiquitous Virtual Reality*, pp. 51-54, IEEE Comp. Soc.

Mulloni, A., Seichter, H., Schmalstieg, D., 2011. Handheld Augmented Reality Indoor Navigation with Activity-Based Instructions. *13th Int'l Conf. on Human-Computer Interaction with Mobile Devices and Services*, ACM.

Nack, F., 2010. Add to the Real. *IEEE Multimedia*, 17(1):4-7, IEEE Comp. Soc.

Schinke, T., Henze, N., Boll, S., 2010. Visualization of Off-screen Objects in Mobile Augmented Reality. *12th Int'l Conf. on Human-Computer Interaction with Mobile Devices and Services*, ACM.

Schmalstieg, D., Langlotz, T., Billinghurst, M., 2010. Augmented reality 2.0. In: Coquillart. S., Brunnett. G., Welch. G. (eds), *Virtual Realities*. Springer, Vienna, pp. 13–37.

Seo, B., Kim, K., Park, J., Park, J., 2010. A Tracking Framework for Augmented Reality Tours on Cultural Heritage Sites. *9th ACM SIGGRAPH Conf. on Virtual-Reality Continuum and its Applications in Industry*, pp. 169-174. ACM.

Suh, Y., Park, Y., Yoon, H., Chang, Y., Woo, W., 2007. Context-Aware Mobile AR System for Personalization, Selective Sharing, and Interaction of Contents in Ubiquitous Computing Environments, *12th Int'l Conf. on Human-Computer Interaction*, Springer.

Takacs, G., Chandrasekhar, V., Gelfand, N., Xiong, Y., Chen, W., Bismpigiannis, T., Grzeszczuk, R., Pulli, K., Girod, B., 2008. Outdoors Augmented Reality on Mobile Phone Using Loxel-based Visual Feature Organization, *1st ACM Int'l Conf. on Multimedia Information Retrieval*, pp. 427-434. ACM Press.

Welch, G., Foxlin, E., 2002. Motion Tracking: No Silver Bullet, But a Respectable Arsenal. *IEEE Computer Graphics and Applications*, 22(6):24–38, IEEE Comp. Soc.

Wither, J., Tsai, Y., Azuma, R., 2011. Mobile Augmented Reality: Indirect Augmented Reality, *Computers & Graphics*, 35(4):810-822, Elsevier.

You, Y., Chin, T., Lim, J., Chevallet, J., Coutrix, C., Nigay. L., 2008. Deploying and Evaluating a Mixed Reality Mobile Treasure Hunt: Snap2Play, *10th Int'l Conf. on Human Computer Interaction with Mobile Devices and Services*, pp. 335-338, ACM.