

An Augmented Reality Exhibition Guide for the iPhone

Mohamed Ashraf Nassar, Fatma Meawad

German University in Cairo

mohamed.nassar@alumni2009.guc.edu.eg, fatma.meawad@guc.edu.eg

Abstract

Handheld Augmented Reality (AR) can enable intuitive browsing and annotation of objects in an exhibition through the visitors' in-hand mobile devices. Several researchers explored Handheld AR technologies in museums and exhibition-like environments. However, despite the proliferation of smart phones that can act as magic lenses for augmented objects, AR technologies are not widely adopted in exhibitions. This paper investigates the possible techniques to build a reliable, scalable and cost effective solution for an indoor marker based exhibition guide on the iPhone. After reviewing possible tracking technologies, available open source marker based tracking toolkits on the iPhone are explored. The paper concludes with a proposed design for dynamic content creation to augment and annotate exhibition objects.

1. Introduction

Augmented reality has become a rich field of research in recent years, so have the technologies enabling it to become widely used. Smart phones, for example, with their growing capabilities, built-in sensors and wide spread represent one of these enabling technologies. Much research has been carried out on how handheld augmented reality can be used to enhance user experience in museums, exhibitions or in other indoor environments. Some research attempted to contextually overlay 3D characters on the mobile screen to interact with users and guide them in the environment [18]. Others used games to get many users to interact and enjoy a tour inside a building [14]. However, existing attempts in indoor environments are merely prototypes that cannot be easily adopted in real life contexts.

The aim of this work is to provide an augmented reality application on the iPhone to allow exhibitions' visitors to enjoy access to dynamic information about the displayed items and share their own comments if they please. Enabling these forms of interactions in a reliable, yet seamless manner is envisaged to promote wide adoption of the technology in any indoor exhibition.

Tracking techniques are the biggest challenge to the usability of an AR client on mobile devices. Thus, we start this paper with a discussion of existing tracking techniques to give a basic

background that supports our subsequent design decisions. With a clear description of the system requirements and its essential success criteria, we discuss a simple scenario on a prototype of our proposed design. Based on some experimental results of testing the prototype in different scenarios, we highlight some future directions for this on-going work.

2. Tracking Techniques

Choosing a tracking technique for an Augmented Reality application is an important design decision. Such choice can either cause or resolve many challenges for the application. For example, reliable and accurate solutions might come with a compromise of usability and performance. Additionally, using a free tracking toolkit, purchasing one or building one from scratch would impact reliability, performance and the expected cost of the whole solution. We aim to present a solution that combines the benefits of usability, reliability and cost effectiveness, in order to achieve wide adoption. In this section, we give an overview of existing tracking techniques and their suitability for our main goal.

2.1. Indoors Location Based Tracking (LBT)

Despite the wide success of Global positioning system (GPS) with location based mobile applications in outdoor settings, GPS does not provide the required accuracy indoors. Alternately, various approaches were used for indoor location based tracking, either through the use of Infra-Red (IR) networks [1], Wireless-LAN (Wi-Fi) networks for triangulating the location [2] or active Radio Frequency Identification (RFID) tags [8]. The existing solutions for indoors location based tracking require expensive infrastructure without a guarantee of an acceptable level of accuracy for an AR application [8].

2.2. Marker-less Tracking

Marker-less tracking, or feature tracking does not require special markers to be introduced in the application setting. It is based purely on computer vision algorithms extracting the specific required features from the incoming camera feed, for

example, faces, buildings, objects, etc. Mobile based feature tracking is gaining more attention as the processing power of current smart phones increases to allow it [12, 4].

Even though, the concept behind marker-less tracking could provide the ultimate augmented reality experience for users, both indoors and outdoors, there are some major challenges for this approach that hinder its adoption, especially with mobile based applications. The machine learning techniques used for feature detection demand high processing powers and advanced memory requirements. Scalability is another challenge for this approach since complex algorithms need to be developed for learning different features. Finally, marker-less tracking can be severely crippled in mono-vision applications that use a single camera, due to the difficulty in depth perception [10].

2.3. Marker-based Tracking

Markers, also known as Fiducials, are symbols with specific known patterns that are designed to be easily recognized by machines [11, 6]. The marker-based technique is the most commonly used indoor tracking technology. This is because it relies on relatively simple software algorithms that do not require high processing powers nor expensive special built-in sensors. The following subsections provide an overview of some of the most common types of markers: template, 2D Barcodes and topological markers.

2.3.1. Template Markers

Template markers are one of the earliest fiducial markers developed. Template markers are basically a square with a black or white border with a contrasting background (white/black respectively) that can have anything inside of it. The whole square with the pattern inside it are recognized as a marker, (see figure 1(a)). Being a first however, template markers suffer from some drawbacks. Firstly, when a marker is detected, its pattern is extracted and cross-correlated with all known patterns. Consequently, as the patterns used increase and as the markers in the input image increase, the application becomes slower [6, 17]. Secondly, template markers must first be designed and then trained for the toolkit to use them. Finally, the complexity of the marker pattern affects the efficiency of the tracking, where patterns with large black and white regions (i.e. low frequency patterns) are the most effective [3]. Template markers were originally developed for the ARToolKit [9] and currently supported by the ARToolKitPlus [17] and Studierstube ES [15].

2.3.2. 2D Barcodes

2D Barcodes are gaining more popularity as fiducial markers. The ISO standard Data Matrix barcode (see figure 1(b)) can store up to 2KB of information [16]. However, in environments where only little information needs to be encoded, simpler markers are preferred. For example frame, split, BCH (Bose, Ray-Chaudhuri & Hocquenghem) (see figure 1(d)) and Standard markers (see figure 1(c)). These markers can be used for an ID-aware environments where the ID number is encoded as bits in the marker itself. There are several benefits to such approach. Firstly, detection of ID-markers is always faster than for template-markers since no image matching is required [6]. Secondly, the user does not have to provide marker images, but can freely choose any marker from a fixed set of patterns. Finally, in contrast to template markers, the user is not required to train the toolkits with new patterns since any valid marker is implicitly known to the system [17]. Frame and Split markers offer an extra benefit of enabling human understandable information in addition to the encoded ID. However, they are only supported by a few toolkits of which none is available for licensing. Section 2.4 provides more details about BCH and Standard ID markers.

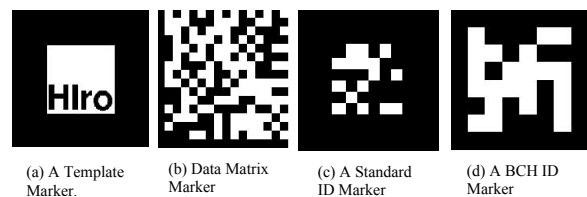


Figure 1. Several Types of Markers used in marker based tracking.

2.3.3. Topological markers

Topology based markers, such as D-Touch [6], are a class of markers that are recognized based on topological features of the markers rather than their geometry. Marker recognition is not based on shape, but on the relationship of dark and light regions. As a consequence, the visual design of the markers is less constrained and can be governed by aesthetic principles. That's form and function, rather than just function. This is their distinctive feature compared to similar systems. The markers can be visually designed to convey meaning to people [5]. Topological markers seem as good candidates for our application context since they allows the usage of markers with better aesthetic qualities. Their main drawback however is that each marker must be independently and uniquely designed and validated. This requirement would incur additional effort, cost and time on initial deployment and on scalability, as in the case of template markers.

2.4. Toolkits

An augmented reality toolkit is a software library that can be used to calculate camera position and orientation relative to physical markers in real time. We are interested in open source toolkits that can demonstrate an acceptable performance on a mobile device and that is either compatible with or can be ported to the iPhone platform. Existing literature introduced several non-commercial marker based toolkits for augmented reality, for example, ARTag[7], Studierstube ES [13] and ARToolKitPlus [17]. Studierstube ES has been written from scratch with high performance for PCs as well as mobile phones in mind. It is about twice as fast as ARToolkitPlus on mobile devices and supports many different marker types [13]. However, Studierstube ES is a closed source library and is not currently being licensed. ARTag library was announced to be unavailable for use by its author, Mark Fiala, despite its benefits over other similar toolkits, for example: its edge-based detection algorithm that make it both lightning and occlusion immune [7].

ARToolkitPlus is the predecessor of Studierstube; it is an extended version of the original ARToolKit's vision code that was built to suit the needs of mobile devices. ARToolkitPlus has several additional features over the original ARToolkit. Most important of which are: Firstly, the support for mobile devices pixel formats such as RGB565. This allows image processing on the native pixel formats of the devices instead of incurring a penalty for converting to other formats. Secondly, it is based on fixed point arithmetic operations. This speeds it up considerably since mobile devices usually do not have a dedicated hardware floating point unit. Finally, it supports automatic thresholding and vignetting that compensate for the sometimes low resolution and the distortion of mobile device cameras.

ARToolkitPlus offers two types of ID markers. Standard ID markers (STD ID), also known as simple ID, use simple four-fold redundancy and encode 9 bits in a 6x6 pattern. The second type of ID markers is the BCH, this type uses cyclic redundancy checks (CRC) which require less redundancy for similar robustness. BCH markers can therefore store 12 bits in the 6x6 pattern image. ARToolkitPlus supports up to 512 STD ID markers and 4096 BCH markers. As mentioned earlier, ID markers have the benefit that they encode the ID within the marker so there is no speed penalty incurred when using a large number of markers. Consequently, ID markers stand out as the right choice of makers for this work.

3. System Design Considerations

Exhibitions are information rich environments that can vastly vary in their settings. We target indoor environments where exhibits can change frequently and are not tied to a specific location. Thus, the deployment of our system with minimal effort and cost is an essential requirement for our solution. Such requirement rules out the use of indoor location based tracking and marker less tracking techniques for our solution, leaving us with marker based tracking as the best suit for our solution.

Our proposed system is based on a client server architecture, where the marker identification is done on the mobile device (client), and the ID is sent to a database server to query for information about the displayed item. Figure 2 presents a typical usage scenario; the user points the iPhone at the augmented object (the Porsche Poster in this case). Since the information is stored in a database, dynamically updating the information, adding new ID's, or even changing the information associated with a certain ID can be done quickly and independently from the rest of the system. This, in addition to the fact that deploying printed black and white markers is quick and inexpensive, simplifies dynamic content generation and association with the exhibited items.

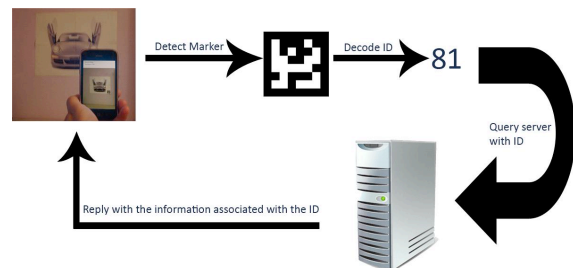


Figure 2: A typical scenario for an exhibition visitor who points the iPhone to the displayed item to browse related information.

Data retrieval and display on mobile devices is a major usability consideration. The client application on the mobile should handle the intuitive browsing of information about a recognized item. Moreover, the client application should allow visitors to share their comments on the items they are seeing. Finally, it is important to target the visitors' in-hand mobile device instead of offering special devices at the beginning of their visit. This solution is not only cheaper, but also ensures the users' familiarity with the devices.

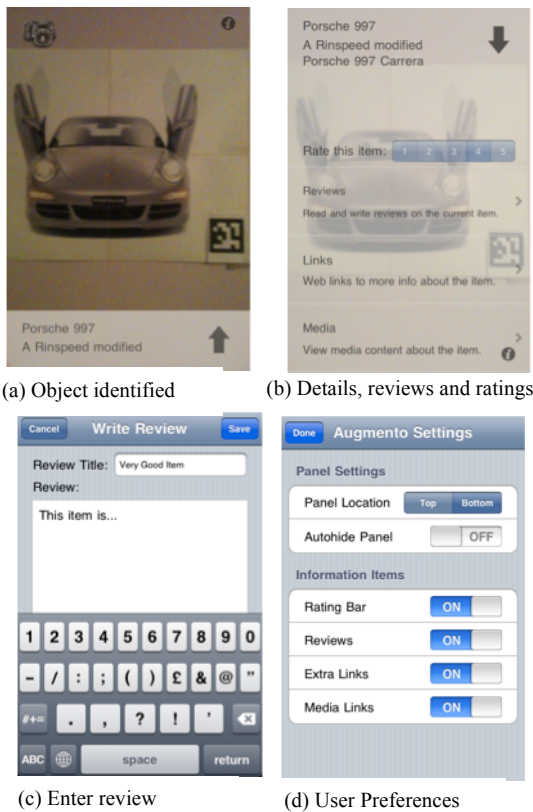


Figure 3: Snapshots from the iPhone Mobile client

A prototype was built using ARToolkitPlus on the iPhone to test our envisaged solution. Our choice of the iPhone stems from its wide adoption worldwide [19]. Figure 3 shows several snapshots from our iPhone mobile client. The client application running ARToolKitPlus, detects and decodes the marker, then queries the server with the marker ID and retrieves the information over WiFi. The iPhone screen, as shown figure 3(a), displays the information on a transparent white panel that fades in when a marker is detected and fades out when no marker is detected. Additionally, the panel can be fixed at the top or button according to user preference. The panel is easily scrollable and expandable to show more details about an object, or to show existing reviews and ratings on this specific item (see figure 3(b)). The user can choose to enter ratings and reviews. Additionally, the user can personalize the interface through the display settings (see figure 3(d)). Entering data in such panel is relatively simple for an iPhone user. The use of a fixed panel was initially a fast prototyping decision. Then, it was found to combine good aesthetics with good usability. The iPhone here acts as a “magic-lens”, where the user holds the device and perceives the physical exhibition and the augmented information through the iPhone screen.

Type	Marker Size	MRR At Angle: 90 degrees (Straight)	MRR At angle: 45 degrees
STD ID	Small	55 cm	40 cm
	Medium	126 cm	85 cm
	Large	300 cm	198 cm
BCH	Small	72 cm	56 cm
	Medium	187.5 cm	100 cm
	Large	300+ cm	247 cm

Table 1: A Table showing the results of measuring the maximum robust range (MRR) in different scenarios.

4. Experiments

The proposed system is aimed at exhibition visitors who are assumed to have no technical background.

Hence we put the following criteria that our system needs to fulfill. The first and most important is the recognition accuracy. The system needs to be able to recognize markers correctly at a comfortable distance to the user. Secondly, the system must be fast in detection. Thirdly, it needs to be robust in the sense that it does not exhibit jittering or false recognition.

4.1. Setup

For the purpose of our experiment, the two types of ID markers used by ARToolKitPlus are tested: STD ID and BCH. We are interested in the maximum robust range (MRR) at which the marker is correctly recognized and is tolerant to smooth medium speed movements. We will measure the MRR for both the BCH and STD ID markers with three different marker sizes: small (2.5x2.5 cm), medium (6x6 cm) and large (13x13 cm). Lightning conditions are reminiscent of typical indoor lightning and are constant throughout different trials. This is an acceptable setting since lightning can, in most cases, be easily controlled indoors.

Each marker is fixed perfectly flat to the bottom right corner of a poster. Such arrangement is chosen to mimic a painting displayed in a museum or gallery as an example exhibition setting. Then, the experimenter moves further away from the marker till a MRR is reached and then it is measured. This step is repeated for each size of each type of the markers and from two different angles: 90 degrees and 45 degrees to the marker. The process is conducted using two different generations of the iPhone the 3G and the 3GS models.

4.2. Observations

Table 1 shows the measured MRR for each marker type, size and angle using the iPhone 3G. The table clearly shows that BCH markers tend to offer more overall reliability and robustness than STD ID markers. They are recognized more quickly and their MRR is larger than the STD ID in all cases. This is probably due to the CRC error correction embedded in the BCH markers as discussed earlier in section 2.4. Thus, we interpret the results taking the BCH as a point of reference since it represents the best measurement.

The small markers only offer an MRR of up to 72 cm when facing the markers and up to 56 cm when viewing them from an angle. This restricts the movements of the users in this small distance. The user has to be too close to the item. With such measurements, small markers can be suitable for smaller objects/pieces of art that cannot be viewed from a distance. Naturally, the medium sized markers showed better MRR measurements than the small ones due to their size advantage. Offering up to approximately two meters in a direct view and up to one meter from an angle of 45 degrees provides a comfortable range for users to move around an exhibit item. Their slightly bigger size makes them more obtrusive than their smaller counterparts; however they offer a good compromise between obtrusiveness and MRR in many cases. Finally, the large markers are recognized in distances greater than three meters. This makes them ideal when trying to augment a large item or when requiring augmentation from large distances. However, they are even more obtrusive than the medium sized markers. The overall results of both the iPhone 3G and 3Gs are similar. The only notable difference is in the speed of launching the application where it is faster on the 3Gs model.

5. Discussion

We have presented an on-going work towards an interactive handheld augmented reality system for enhancing the experience for exhibitions' visitors and to promote wide adoption of the technology. We emphasized several criteria as the success factors for our proposed system: usability, reliability, scalability and cost effectiveness. We presented a prototype of our system using the iPhone platform as an example of current smart phones. We favored the marker based approach over other tracking techniques due to its relative simplicity, ease of deployment and cost effectiveness.

Our proposed system is based on a client server architecture, where the marker identification is done on the mobile device (client), while data storage and management takes place on a server. Our experimental results were promising for the three

different sizes of the markers used. Plus, there are no restrictions on the marker size chosen for printing which allows for a great deal of flexibility for setting up the exhibition. For example a small art gallery may choose to use medium sized markers because its visitors tend to move around the paintings in a 2 meter radius for example. While a big museum may use very large markers to augment a huge statue and allow viewing it from a larger distance.

5.1. Challenges

Using marker based tracking has its limitations. Markers can be intrusive in an exhibition's environment. Our approach offers flexibility in the choice of the size. Such control allows the organizer of an exhibition to place the markers in the least intrusive manner according to the application context. Another potential limitation is that the number of possible markers is limited (4096 in case of ARToolkitPlus BCH markers). This limits the number of objects that could be augmented if the target exhibition is really large. Although ARToolkitPlus is open source, increasing the number of markers is not an option since this requires changing the encoding/decoding algorithms of the BCH markers themselves. This is roughly equivalent to writing our own library from scratch.

Finally, when augmenting a large piece, one marker may not be enough. Picture the scenario where a user is standing in front of a painting spanning an entire wall, either he/she will back up until the whole painting is in his field of vision, or he/she will stand up close and observe each part separately. In the first case we might need to use a larger marker in order for it to be recognized from a larger distance. In the second case we might need to place several markers so that recognition is preserved wherever the user looks. A similar problem arises when augmenting a 3D piece like a vase for example; a marker is only tracked when it is in view, so if the user moves about the piece, the augmentation is lost. A solution for this would be to use four markers on the four sides of the object. However, this may be obtrusive and unwanted in some applications. In this case we might provide an option for the user to lock on a specific marker ID so that the information associated with that ID is visible even if the user moves around.

One major technical challenge that we faced while building the prototype is getting a camera feed on the iPhone. The iPhone SDK does not allow direct camera access. Thus, we decided to program the phone to snap periodic frames in order to search them for markers. However, this proved inefficient because the iPhone camera takes about 1.5-2.5 seconds (iPhone 3G) or 1-2 seconds (iPhone 3GS) to capture a photo, which makes the program jittery. Plus, the camera capture is done in the main thread,

suspending the application until it completes. As a workaround, we programmed the phone to take a self-snap (a snapshot of its own display) instead. The self-snap is done in a separate thread so it does not impact performance, and is sufficiently fast to allow us to take up to 5 pictures per second on the iPhone 3G. This enables real-time perception for the user. The result image from each snap is a 480 * 320 pixel (which is the iPhone screen resolution) bitmap image. The raw image bytes are then extracted and fed into the vision code of ARToolkitPlus. The color space of the toolkit is set to BGRA (Blue, Green, Red, Alpha) in order to allow fast processing at the native capture format and color space.

5.2. What is Next?

The work described in this paper is a work in-progress. Hence, we conclude by highlighting our next steps:

- Finalizing the prototype to add intelligent recommendations to the visitor based on viewed pieces and ratings.
- Adding support for multi-marker tracking is another point that should be taken in consideration. Currently, if two markers are recognized by the application, the one with the highest confidence will have precedence. In case both markers achieve the same confidence, the user should have the choice which identified object to be in focus.
- Conducting a variety of usability tests with many users in real life conditions in an information rich environment. Such tests will help ensure a great level of usability and high acceptance of the application.

6. REFERENCES

- [1] Roy Want, Bill N. Schilit, Norman I. Adams, Rich Gold, Karin Pedersen, David Goldberg, John R. Ellis, and Mark Weiser. The ParcTab ubiquitous computing experiment. *IEEE Personal Communications*, 2:28-43, 0.
- [2] P. Bahl and V. N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, volume 2, pages 775–784. IEEE, 2000.
- [3] Artoolkit documentation, <http://www.hitl.washington.edu/artoolkit/documentation/evmulti.htm>.
- [4] E. Bruns, B. Brombach, and O. Bimber. Mobile phone-enabled museum guidance with adaptive classification. *IEEE Computer Graphics and Applications*, 28(4):98–102, 2008.
- [5] E. Costanza. Telling a story on a tag: the importance of markers' visual design for real world applications. In *Mobile Interaction with the Real World (MIRW)*, 2006.
- [6] E. Costanza and J. Huang. Designable visual markers. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1879–1888, New York, NY, USA, 2009. ACM.
- [7] M. Fiala. Artagrev2 fiducialmarker system: Vision-based tracking for ar. In *ISMAR (October 2005)*
- [8] J. Hightower and G. Borriello. A survey and taxonomy of location systems for ubiquitous computing. *Computer*, Vol. 34, No. 8. (07 Aug 2001), pp. 57-66.
- [9] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. *Augmented Reality, International Workshop on*, 0:85–94, 1999.
- [10] J. M. V. Misker and J. Ster. Authoring immersive mixed reality experiences. In E. Dubois, P. Gray, and L. Nigay, editors, *The Engineering of Mixed Reality Systems*, chapter 14, pages 275–291. Springer London, London, 2010.
- [11] C. Owen, F. Xiao, and P. Middlin. What is the best fiducial? In *Augmented Reality Toolkit, the First IEEE International Workshop*, 2002.
- [12] B. Ruf, E. Kokiopoulou, and M. Detyniecki. Mobile museum guide based on fast sift recognition. In *6th International Workshop on Adaptive Multimedia Retrieval*, 2008.
- [13] D. Schmalstieg, A. Fuhrmann, G. Hesina, Z. Szalav'ari, L. M. Encarnac, M. Gervautz, and W. Purgathofer. The studierstube augmented reality project. *Presence: Teleoperators and Virtual Environments*, 11(1):33–54, February 2002.
- [14] D. Schmalstieg and D. Wagner. A handheld augmented reality museum guide. In *International Conference Mobile Learning. IADIS*, 2005.
- [15] D. Schmalstieg and D. Wagner. Experiences with handheld augmented reality. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–13, Washington, DC, USA, November 2007. IEEE.
- [16] D. Wagner, T. Langlotz, and D. Schmalstieg. Robust and unobtrusive marker tracking on mobile phones. In *ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, volume 0, pages 121–124, Washington, DC, USA, 2008. IEEE Computer Society.
- [17] D. Wagner and D. Schmalstieg. Artoolkitplus for pose tracking on mobile devices. Technical report, Institute for Computer Graphics and Vision, Graz University of Technology, February 2007.
- [18] R. Wojciechowski, K. Walczak, M. White, and W. Cellary. Building virtual and augmented reality museum exhibitions. In *Web3D '04: Proceedings of the ninth international conference on 3D Web technology*, pages 135–144, New York, NY, USA, 2004. ACM
- [19] Gartner. Worldwide mobile phone sales 2009. <http://www.gartner.com/it/page.jsp?id=1306513>