Chapter 7
Design Guidelines for Mobile Augmented Reality: User Experience

Subhashini Ganapathy

1 Introduction: Mobile Augmented Reality

Mobile augmented reality (MAR) enabled devices have the capability to present a large amount of information in real time, based on sensors that determine proximity, visual reference, maps, and detailed information on the environment. Location and proximity technologies combined with detailed mapping allow effective navigation. Visual analysis software and growing image databases enable object recognition. Advanced graphics capabilities bring sophisticated presentation of the user interface. These capabilities together allow for real-time melding of the physical and the virtual worlds and can be used for information overlay of the user's environment for various purposes such as entertainment, tourist assistance, navigation assistance, and education [1].

In designing for MAR applications it is very important to understand the context in which the information has to be presented. Past research on information presentation on small form factor computing has highlighted the importance of presenting the right information in the right way to effectively engage the user [2–4]. The screen space that is available on a small form factor is limited, and having augmented information presented as an overlay poses very interesting challenges.

MAR usages involve devices that are able to perceive the context of the user based on the location and other sensor based information. In their paper on "Context-Aware Pervasive Systems: Architectures for a New Breed of Applications", Loke [5],

S. Ganapathy (✉)
Department of Biomedical, Industrial, and Human Factors Engineering, Wright State University, 207 Russ Engineering Center, 3640 Col. Glen Hwy, Dayton, OH 45431, USA
e-mail: subhashini.ganapathy@wright.edu

W. Huang et al. (eds.), *Human Factors in Augmented Reality Environments*,
DOI 10.1007/978-1-4614-4205-9_7,

talk about two approaches to context-aware systems based on where the information resides—self supported context awareness and infrastructure supported context awareness. In the case of self-supported context awareness the device has the ability to identify the context and alter its behavior based on the context. This is enabled through the hardware, software and other sensors present on the device. In the case of infrastructure supported context awareness, the devices acquire context-awareness by using the infrastructure external to the device.

Context aware computing has been applied in different contexts. Some examples include—the TEA project—mobile phone that detects situations of the phone such as "in hand," "on table," "in pocket," and "outdoors" [6]. Hakansson et al. [7] talk about cameras that can sense sound, pollution in the air, temperature and smell, and create visual effects in photographs given its context. Augmented reality has also been discussed in the context of clothing SensVest [8], worn during different sporting activities has sensors to measure physiological data such as movement, energy expenditure, heart rate, and temperature.

When designing for a MAR application it is important to understand the following segments:

- Usage scenario
  - Shopping assistant, tourism/vacation related usage scenarios, social networking related usage scenarios, and gaming
- Interaction modalities
  - Voice/audio, touch-based, tactile feedback
- Device form factor
  - Ultra mobile device, MID, camera, smartphones

The next couple of sections will go into the details of each of these segments. This chapter outlines the design guidelines related to mobile augmented reality based on the context in which the information has to be presented. The first section gives an overview of the different usage areas where mobile augmented reality is most relevant. The second section discusses the different input methods that are used for accessing the information. The third section talks about the different form factors that are in the space of mobile computing that needs to be understood to design effectively. The last section summarizes the different design aspects in developing persuasive usage model interactions for a great user experience on mobile devices.

2 Usage Scenarios

In the context of mobile computing, some of the key application areas include—shopping usage scenarios, tourism/vacation related usage scenarios, social networking related usage scenarios, and gaming. Understanding the context or the usage

Fig. 7.1 MAR shopping assistant usage scenario. Bottle selected to capture in an image. Wine information appears

scenario is very important when designing MAR applications as that provides the right user experience. By detailing the usage scenario the user requirements can be gathered which helps set the design of the product. The usage scenario and user requirements can be developed by understanding the users' likes, dislikes, perceptions, desirability, and expectations and their interaction with the product(s). The following sub sections discuss the different usage scenarios.

2.1 Shopping Usage Scenario

Shopping usage scenario typically involves a certain object of interest. The type of augmented information can be based on a particular object. It can be product information, reviews, or pricing comparisons based on the photo of an item or barcode. For example, as shown in Fig. 7.1, by taking a picture of a wine bottle, the MAR application on the phone would match the image of the bottle with a database listing and present related information, such as competitive pricing and reviews. The application can also act like a shopping assistant by developing a profile of your buying patterns, likes/dislikes, style, and guide you to selections that might appeal to you or match items you have or remind you if already have something similar. It can act like a virtual dressing room where you can superimpose items over an image of yourself to virtually try on items scanned.

2.2 Navigation Assistant: Tourism Related

Navigation assistant scenario provides a rich context for MAR. There are a variety of information that can be augmented such as restaurants, buildings, points of interest, attraction spots, public transport stops, and so on. It is also a very practical implementation that is now more commonly available on phones such as G1 Android and iPhone® [9]. Wikitude® is an AR application that is available on mobile phones that allows the user to view the environment with overlay information on points of interest and a brief description to it. Augmented GeoTravel application for the iPhone uses sensors to show the Augmented Reality view of a point of interest. The Layar Reality Browser is another application that shows the user augmented information of the environment.

Device can retrieve publicly available data and develop personal tour based on user interests/preferences. The user can actively request access to location relevant shared content from others such as text, photos, audio and videos regarding their experiences at that location at other points in time. The device can augment live video feed as seen through the device with overlays identifying landmarks, waypoints, etc. Some of the other compelling usages include:

- *Personal translator*: take snapshot of written information and receive real-time transcription. Speak into device (or type/select phrases) and receive real-time translation.
- *Virtual guide*: tour of an area pointing out things of interest and providing additional reference information. Can be visual/auditory.
- *Voyeur vacation*: view photos/videos, listen to voice commentary, or read blogs from others sharing their experiences at your location at other points in time. Video overlay of historical re-enactments (battles or see what people/places looked at a point in time).

2.3 Gaming

Gaming is another rich context in which augmented reality has been explored a lot. Specifically in the case of mobile augmented reality there are opportunities to create games that let you do a visual recognition of the landscape and place avatars in the space. It can be used for position tracking to hunt virtual treasure. The augmented information can be a display of other players' avatars or simulated scenes. As show in Fig. 7.2, it can be a simulated image of a treasure box when the person holds the mobile device in front of the scene as they are searching for treasure. Figure 7.3 shows an avatar of a character when a person holds a camera in front.

2.4 Social Networking Usage Scenarios

In the case of social media and networking scenario the user can provide information on friends with similar preference based on dynamic extraction of social

Fig. 7.2 Gaming augmented information

Fig. 7.3 MAR gaming: avatar

network. Based on your location, the device can automatically alert other users information (movies, restaurant). It can also provide real-time media recommendations and sharing based on user preference. The device can act like "People Minder" to provide reminders or details about an individual of interest (basic info, or tap into user database like last meeting, interests, etc.). The user can manage their location broadcast settings (on/off, who can see). User can enter additional information in individual's profile (type, scan biz cards, hit database, etc.). User can also ping others with information about their location/activities/plans for opportunistic socialization.

The next section gives an overview of the different input modalities that are important to consider in designing a mobile augmented reality system or application.

3 Input Modalities

With the human–computer interaction model moving from traditional input methods to more natural, ubiquitous input techniques there is a need for us to understand and increase the richness of the user experience with higher integration and functionality of these types of technologies with augmented information. There have been many recent developments in the field of multi-modal user interaction. Reactable™ is a multi-touch input based electronic music instrument that allows performers to simultaneously share complete control of the instruments by moving and rotating physical objects on a table surface. Microsoft's multi-touch, Surface™, allows users to manipulate digital content by the use of natural motions such as hand gestures. There has also been a significant improvement in speech recognition algorithms and software sophistication combined with computation power that allows natural language inputs to be useable.

Each of the input modalities will enable different and important experiences. Voice input allows easy navigation and searching without keypad. In the case of dirty hands or if your hands are occupied and you want to search of a specific information it is useful to use voice input. In that case the system should be able to identify the context in which the user is requesting information and present the right information. Voice can also be used to add precision to the manipulation of information.

Figure 7.4 shows an example of multi-modal input interaction on a mobile device. In the case of gesture input you can move your hand in a specific way to send command to your device. Gesture interaction can be very useful for browsing, scrolling and zooming. It is easy to use it for flipping through a lot of information and for 3D manipulation. For example, the current smartphone implement applications that can be easy moved through hand gestures instead of physically touching the device. You can also use a hand-held device, equipped with accelerometers and positional sensing, to wave in the air and create natural gestures that can interact with your laptop or desktop PC. Both voice and gesture are subject to involuntary/accidental input.

Fig. 7.4 Multi modal input output interaction model

When designing for augmented reality it is important to understand the strengths and challenges of these different modalities and the human interaction through these modalities.

4 Form Factor

In developing applications for new device classes such as tablets, it is important to pay early attention to the physical characteristics of the device. These include: ergonomics (feel, grip, balance and weight, hold-position for different hand-sizes); surface texture and screen vs. surround proportions. These design features quickly constrain other aspects of the solution, including: electronics, technology components feasibility, and so on. Figure 7.5 shows an example of UI on different screen size. As seen in Fig. 7.5 the UI designed shows better on a larger screen vs. a small screen size. The UI doesn't allow the user to see details of the different building in a small screen size. Hence it is important to make sure that the information that is presented is scalable across different form factors.

Also based on the orientation of the device the augmented information needs to be adjusted whether it is in the landscape mode or portrait mode. When a person is holding the device in the portrait mode for a mobile like that is of ~5″ screen size there is not a lot of information you can show on the image. If it is in the horizontal orientation then more amount of augmented information can be presented. Figure 7.6 shows an example of different screen size for mobile form factor.

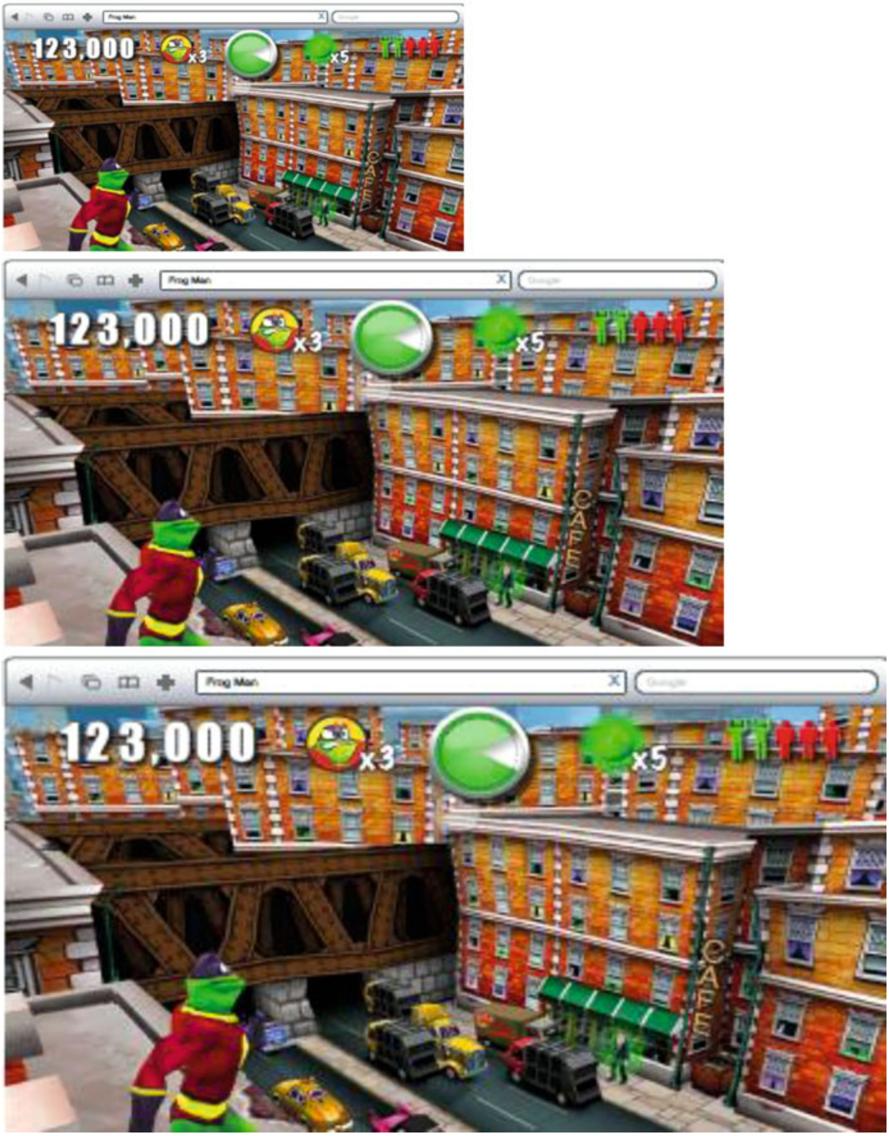

Fig. 7.5 User interface on different screen sizes

Baus et al. [10] discusses the restriction that the user has in terms of the time available to look at the augmented display information and the amount the time it takes for the user to understand and place that information in the spatial context of their environment. There can also be other parallel tasks that the user performs that can affect the viewing time. Figure 7.7, shows an example of different density of

Fig. 7.6 Example of different form factors

information that is presented to the user. Based on the context and the device the user may prefer to see a low level of density of information to a high density level of information. If it is a larger screen like 9.5″ then it will be easy to read the information with high density.

This section illustrates how this new classification can be used to classify existing applications. It also shows how it could guide a developer to place his application in the proper level in order to find out the resources needed as well as the density of competitors existing in that level.

5 Stages of MAR Evolution

Current methods of augmented information presentation on mobile devices are very crowded. This can increase the cognitive overload and frustration for the user. As shown in Fig. 7.8, there is no clear textual information, representation of image overlay. The dials can be confusing.

Hence there is a need to systematically design the MAR application to be most effective. The context or usage scenario in which the application is used must be understood that leads to the usage requirements which can then translate to the design requirements of the application. MAR usages can be mapped against the

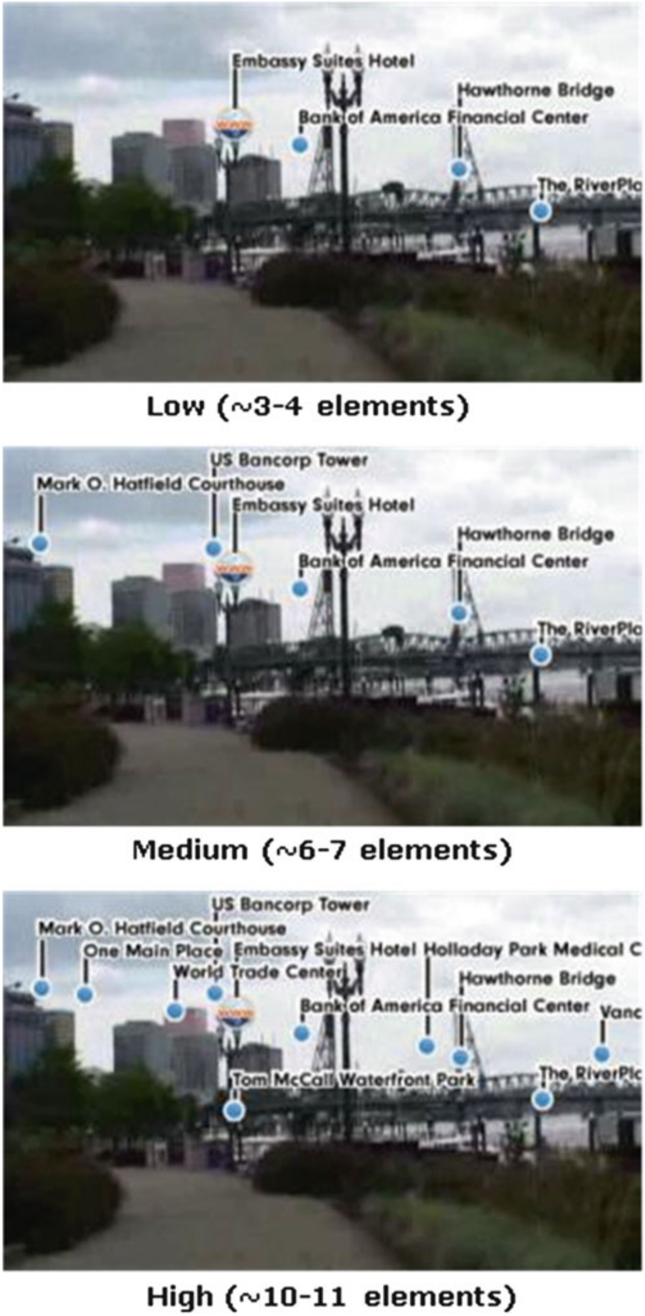

Fig. 7.7 Three levels of density

Fig. 7.8 Example of MAR system

different stages of MAR evolution from sensory data to visual search to data overlay and then to video overlay. As shown in Table 7.1, the different usages discussed in the previous section can be broken down into the different scenarios in order to understand the requirements.

Based on these usage scenarios it is important to create usage requirements. Below is an example UI for a MAR usage (shopping scenario), shown in Fig. 7.9, in order to demonstrate best practices and ideate on the potential interactions and requirements.

Example usage scenario: the device detects that the user is in a grocery store and begins to prefetch grocery related information for a potential search. The user pulls out the handheld device and select the application. The screen comes up with a viewfinder window for the device phone enabled and ready and providing video feed of whatever device is pointed at. The user points the device down toward the Eggplant on the refrigerated rack and the device snaps a photo. The user can then tap the highlighted region of the Eggplant and the device prompts for “limit search to grocery items?” The device prompts the user if he or she would like to add the ingredients to a shopping list and they can tap the case to select the Yes (default option), As the user picks up items she takes a photo of the barcode and the device checks it off her list and maintains a running total for the cost of the items that she is purchasing.

Table 7.1 Usage vs. technology matrix

Usage areas	Sensory data	Visual search	Data overlay	Video overlay
Tourism/travel	ID places in scene, (low data density) Use GPS directions, and orientation	Search based on recognition Geo tagging	ID places in scene (high data density) Self-guided tours identification of written language	Past/future representation of scene—historical recreations
Shopping assistant	Provide directions within store	Search based on recognition Use of profile	Identification of information on items in scene Comparative pricing	Live virtual dressing room Incorporating multiple items/ accessories/cosmetics
Social network/ people finder	Location reporting of self/ others ID places in scene Data push vs. pull	Recognition of faces (poor)	People recognition Social tagging	Social tagging Virtual meetings/ collaboration
Gaming/ entertainment	Integration with sensors for I/O Peripheral detection Translation of angular positioning/tilt/acceleration/motion	Role playing games, social interaction games	Treasure hunt, social interaction games	Integration of virtual game items into real scene Natural tracking algorithms

Fig. 7.9 MAR usage scenario—shopping assistant

6 MAR User Interface Guidelines

This section details the design guidelines when developing the user interface for mobile augmented reality. It is critical that MAR enabled device provides feedback to the user when performing the appropriate mode of activity to ensure that the device is in the right state to perform the task. It is also essential that when multiple objects are being displayed on the MAR enabled device, the user has a way of selecting a subset of the objects, or a subset of the screen for the MAR enabled device to process. Figure 7.10, shows example of using the design guidelines. The design guidelines to be considered are

6.1 Clear Textual Information

You need to make sure that the textual information presented is clear. The type of font used should be such that it is easy to read.

6.2 Contrast

Text/background visibility—visual overlay should be such that there is sufficient contrast between the overlay text and background. The text that is presented should

Fig. 7.10 Information presentation using the design guidelines

be visible across different background. It is important to understand the type of background before determining the text color. In certain cases it is useful to use a textbox to define the background for the text. The disadvantage of that is the background will be hidden.

6.3 Grouping

Organization of information is very important when presenting overlay of information.

6.4 Placement

It is important that the information that is presented should not obscure the item of interest.

6.5 Alert/Attention Sensitivity

Specifically calling out information that may need action. Especially in the case of training or medical overlay the information presented should be able to identify the areas which are important and need immediate attention.

6.6 Interaction Methods

Based on the context, the user should be able to switch to different interaction methods. The user may want to pause the information on the screen so that they can hold the smartphone more comfortably or may want to share the information with their friends/family. They could then toggle different labelling options to appear on the screen. For example, they could show all restaurants, and then switch to a type of restaurant, then switch to restaurants within 5 miles.

6.7 Distinct Icons

Augmented information, using icons, must make sure that the items are labeled with varied icons for easy (arsing of information). For example, all social media updates can be of a particular icon, listing of restaurants can be of a different icon, and so on. The user should be able to identify the items' category at a glance without reading the text label.

6.8 Visibility and Distance

In augmenting information, especially in an outdoor scenario, it is important to provide filter based on distance and visibility [11]. It would be helpful if the icon can indicate whether the labelled item is visible.

7 Conclusions

In this book chapter we have focused on addressing the factors that influence the design of mobile augmented reality applications. Future research can focus on integrating the findings from this research as a model-based decision support system that can be integrated into the computing algorithm to achieve high joint performance in the context modeled. The design variables can be dynamic and can support the decision support system for different usage scenarios. For example, a rule-based model can be developed to effectively present information on objects visually occluded without overloading the computing power based on user selection of how far to look for objects, placement of the information (on top of the screen), and so on.

Acknowledgements Special thanks to Delbert Marsh, Glen J. Anderson, Ashley McCorckle, and Igor V. Kozintsev who worked on the project.

References

1. S. Ganapathy, G.J. Anderson, I.V. Kozintsev. MAR Shopping Assistant Usage: Delay, Error, and Utility. In *Virtual Reality Conference (VR),* pp. 207–208, 2011.
2. B. Karstens, R. Rosenbaum, and H. Schumann. Information presentation on mobile handhelds. 1997.
3. D.A. Norman. Cognitive engineering. In *User centered system design: New perspectives on human computer interaction*, pp. 31–61, 1986.
4. A.W. Stedmon, R.S. Kalawsky, K. Hill and C.A. Cook. Old Theories, New Technologies: Cumulative Clutter Effects Using Augmented Reality. In *IEEE International Conference on Information Visualization,* 1999.
5. S. Loke. *Context-Aware Pervasive Systems: Architectures for a New Breed of Applications*. 2007.
6. H.W. Gellersen, A. Schmidt, and M. Beigl. Multi-Sensor Context-Awareness in Mobile Devices and Smart Artifacts. In *Mobile Networks and Applications (MONET),* 2002.
7. M. Hakansson, S. Ljungblad, and L.E. Holmquist. Capturing the invisible: Designing Context Aware Photography. In *Proceedings of DUX - Designing for User Experience*, 2003.
8. J.F. Knight, A. Schwirtz, F. Psomadelis, C. Baber, H.W. Bristow, and T.N. Arvanitis. The Design of the SensVest, In *Personal and Ubiquitous Computing*, volume 9, no 6, 2005.
9. Augmented Reality [Online]. Available: http://en.wikipedia.org/wiki/Augmented_reality. [Accessed on 05/07/2011].
10. J. Baus, A. Kruger, and C. Stahl. Resource-Adaptive Personal Navigation In *Multimodal Intelligent Information Presentation*, pp. 71–93, 2005.
11. T. H¨ollerer, S. Feiner, T. Terauchi, G. Rashid and D. Hallaway. Exploring MARS: Developing indoor and outdoor user interfaces to a mobile augmented reality system. *Computers and Graphics*, volume 23, no. 6, pp. 779–785, 1999.