Personalized In-store E-Commerce with the PromoPad: an Augmented Reality Shopping Assistant

Wei Zhu[1], Charles B. Owen[1], Hairong Li[2], Joo-Hyun Lee[2]

Michigan State University
East Lansing, Michigan, USA

[1]Department of Computer Science and Engineering
zhuwei@cse.msu.edu
cbowen@cse.msu.edu

[2]Department of Advertising
hairong@msu.edu
leejooh2@msu.edu

Abstract: This paper presents an in-store e-commerce system that provides shopping assistance and personalized advertising through the use of a new concept in context aware computing, *dynamic contextualization*. This system, PromoPad, utilizes augmented reality technologies on a hand-held Tablet PC to provide for dynamic modification of the contextual settings of products on store shelves through the use of see-through vision with augmentations. This real-time modification of the perception of context, *dynamic contextualization*, moves beyond the traditional concept of context-aware computing into context modification. The technical requirements for realizing dynamic contextualization using augmented reality technologies are described in detail. The target design of the PromoPad is a consumer friendly shopping assistant that requires minimum user effort and is practical in a public environment such as a shopping mall or a grocery store.

Keywords: Augmented reality, dynamic contextualization, context-aware, pervasive computing, e-commerce

1. Introduction

This paper describes the design of the PromoPad, an ubiquitous computing device and e-commerce system that performs in-store personalized advertising and shopping assistance. The PromoPad is a prototype hand-held device that provides context sensitive shopping assistance. This assistance is in the form of augmented imagery using augmented reality (AR) technologies and the content of the assistance is built around the concept of *dynamic contextualization.* Dynamic contextualization is an extension of the conventional advertising concept of contextualization to include real-time modification of perceived context on the basis of interactions among the user, the environment and focal objects using augmented reality technologies. Augmented reality differs from virtual reality in that there is no attempt to replace the real world so users can still interact with the real world, and at the same time perceive enhanced views with augmentations.

Augmented reality enhances the human perception of reality in this application by contextualizing individual objects that are encountered in the real world with virtual complements so as to make the real objects more meaningful and appealing. The PromoPad is a tablet PC with a camera mounted on the back. The display on the tablet provides a modified version of the camera image. This image is modified using augmented reality technologies that add new imagery relative to a focal product or remove elements of the image that may distract from the focal product. Thus augmented reality technologies offer the technical capabilities necessary for realizing dynamic contextualization. In traditional context-aware computing, context represents the situation of the user and is static, providing only input to the context-aware computing system. With dynamic contextualization, the context can be modified to be more meaningful for the focal objects and more interesting to the users. Context becomes both an input and output of the system.

Researchers in the field of context-aware computing and e-commerce systems have been seeking ways to provide handy and natural electronic assistance for shoppers. Kourouthanassis and Roussos developed MyGrocer [1], a pervasive retail system, that can manage shopping lists, monitor the total cost of cart contents, popup promotion information, and navigate consumers within the store. Project Voyager examines the use of context-aware computing as a shopping assistant [2]. The PSA is another experimental system that provides personalized shopping assistance [3]. These projects focus on discovering the store-wise context so as to provide an electronic and automated shopping aid that can ease or assist the shopping process. The discovered context in these projects, however, plays a passive role as the situation, such as the location, of the user that can be discovered and utilized. In addition to the context-sensitive content delivery, the PromoPad synthesizes a virtual experience utilizing dynamic contextualization in order to increase product knowledge, brand attitude, and purchase intent [4].

A good e-commerce system does not only provide passive information. The Point-of-Purchase Advertising Institute's research shows that 70% of buying decisions are made in the store [5]. Hence a good e-commerce system should also be able to trigger impulse purchase decisions. Augmented reality powered dynamic contextualization presents 3-D visualizations registered to actual products in the store and in the proper context for impulse decision making. The dynamic contextualization can be blended to reproduce consumption situations that can affect shoppers' perception of a brand and affect the purchase decision.

Dynamic contextualization is made possible by augmented reality technologies that modify the perception of the real world in real time [6]. Several empirical studies on the effectiveness of augmented reality technologies in human computer interaction provide evidence that augmented reality systems improve human performance. Tang et al showed statistical significance to support the hypothesis that augmented technologies improve the operational performance in instructing assembly tasks [7]. The Archeoguide system is an outdoor augmented reality guide that offers personalized tours of archaeological sites. It uses augmented reality technologies to improve information presentation, simulate ancient environment, and recover destroyed sites [8]. Considerable research in augmented reality covers a broad range of application domains. This work explores the technical feasibility and benefits of augmented reality for advertising and consumer experiences.

The remainder of this paper is organized as follows. Section 2 introduces the concept of dynamic contextualization and provides the theoretical basis for its application. Section 3 describes the PromoPad system and the technical details that make the augmented shelf view work. Section 4 discusses the implementation of dynamic contextualization on the PromoPad. Section 5 summarizes the results in this paper and discusses future issues and research.

2. Dynamic contextualization overview

From the perspective of consumer psychology, dynamic contextualization with the PromoPad using augmented reality technology can simulate an enhanced product experience. "Enhanced", in this setting, implies a combination of both direct experience and virtual experience. Traditionally, product experiences have been dichotomized as direct or indirect. *Direct product experience* is the unmediated interaction between consumers and products in a full sensory capacity, including visual, auditory, taste and smell, haptic and orienting [9]. Direct product experience is often obtained from personal inspection of a real product in a store setting. *Indirect product experience* is experience gained through secondary sources such as advertising. Compared with indirect experience, direct experience is much richer for several reasons. Direct experience product information is self-generated by the shopper and thus is perceived as more trustworthy. Direct examination of a product allows the shopper to see, feel and touch a product and get input from multiple sensory channels. The shopper can inspect a product in a sequence and pace of her choice and customize the information to her cognitive needs. However, direct experience from personal inspection is not perfect from a consumer learning perspective in that it is often limited to the product per se and it is not easy to incorporate external information such as the background, users and use scenarios of a product. The consumer can examine the product, but not necessarily remove it from packaging and rarely place it into its intended context. These disadvantages can be overcome with virtual experiences as simulated in 3-D visualization.

The beauty of augmented reality is that it enables a shopper to inspect a product personally and at the same time, view additional objects in 3-D visualization on the Tablet PC display of the PromoPad. Objects in 3-D visualization are found to be able to generate a new form of mediated experience – virtual experience [10]. A virtual experience is a form of indirect experience, because both are mediated experiences [11]. However, virtual experience tends to be richer than indirect experience rendered by printed ads, television commercials, or even two-dimensional (2-D) images on the Web. Li, Daugherty, and Biocca indicate that virtual experience, as simulated in 3-D visualization, consists of more active cognitive and affective activities than 2-D marketing messages [12]. They attribute these psychological and emotional effects to the interface properties of 3-D advertising, as well as to the psychological sensation of presence.

3. The PromoPad system

The PromoPad is a mediated device that provides in-store virtual experience with 3-D product visualization. The system consists of a front-end client component and a back-end server component. The front-end component is a light-weight display device that slips into a cradle in the shopping cart. Tablet PC technologies are used to implement this front-end device. With a camera attached to the back of the Tablet PC, the client device is aware of the position and orientation of the shopper relative to the shopping cart and store shelves. It is also capable of providing the shopper a see-through view of the shelves and additional information that is related to the items in the view. The back-end components consist of one or more servers that contain inventory databases, customer profiles and business logic, from which information in the databases is filtered and returned to the front-end component. The PromoPad employs augmented reality technologies and passes an augmented camera image from the rear of the Tablet PC to the display.

The goal of this design is an intelligent shopping aid that provides shoppers automatic and meaningful help when needed, as well as minimizing human interference and effort. With wireless communication, a Tablet PC can have different modes for shoppers in different shopping orientations. Planned shoppers, those entering the store with specific purchase in mind, may use a tablet PC to optimize their shopping routes in a store to quickly find items they plan to buy. Bargain shoppers, those seeking sales and clearance items, may use the tablet to find sale items that they are interested in with ease. Recreational shoppers, those entering the store simply to browse, can use a Tablet PC to obtain product information that is not on the packaging. For example, an augmented display of a bottle of wine might include images of the winery and/or wine ratings or reviews. Content customization and personalization of a tablet PC can greatly facilitate the convenience of all types of shoppers, enhancing the shopping experience. It is important to recognize that the vast majority of grocery and convenience store purchases are impulse purchases. Even slight improvements in marketing performance can result in massive increases in sales.

To achieve this goal of a Tablet PC as a see-though augmentation device, several technical issues have to be addressed. First, the means of tracking the location context needs to be robust, scalable and stable. Second, the real image in the Tablet PC display should be adjusted to offer a true see-through view as if the Tablet PC display was transparent so that the device is well integrated with the environment and in harmony with real products. Third, the virtual objects should be accurately registered to the real image. Thus, the Tablet PC display can act as a 'magic frame' that allows the user to 'look through' the frame with additional information displayed around the product that cannot be seen otherwise. Finally, the system needs to be able to deal with a variety of different virtual and real composition methods, including overlays, occlusion, and diminishment.

3.1 In-store tracking

The location of a shopper as a 3-D position and orientation relative to product and store shelves is acquired by an in-store tracking system. When the shopper is using the PromoPad, it is reasonable to assume that the position and the orientation of the Table PC are a good approximation of the position and orientation of the shopper. Considerable research has explored the use of ultrasonic, RFID, infrared, and vision-based technologies to achieve location-awareness (as reviewed in [13]). A variety of existing technologies can be scaled for this application to store-size volumes with large quantities of PromoPads. In our experiments, we use a vision-based fiducial system designed by Owen, Xiao, and Middlin [14].

Fiducials are markers that provide visual cues of the position and orientation of the user in a vision-based tracking system. As reported, this fiducial system is robust to partial occlusion and noise, computationally efficient, and scalable. Fiducial systems work well with inexpensive cameras and are easy to deploy. They are also a good match to the monitor-based augmented reality approach utilized in this application [15].

Location information required for the PromoPad is considerably more rigorous than that required by traditional context-aware computing systems. Dynamic contextualization and augmented reality require modification of the camera image. Owen, et al. [16] discuss many issues relative to augmentation of imagery for augmented reality applications. To achieve pixel-resolution registration of virtual elements with store shelf contents, the PromoPad system requires high-accuracy knowledge of the location and orientation of the client device. Visual fiducial systems provide sufficient accuracy for high-quality image modifications and natural image features can be used to *tweak* accuracy if necessary.

3.2 Video see-through systems

The view as seen on the Tablet PC display is derived from the image captured by a camera mounted on the rear of the tablet. The characteristics of the camera view are, in turn, determined by the camera intrinsic and extrinsic parameters. The intrinsic parameters of a camera describe how the camera will convert objects within the camera's field of view into an image. The extrinsic parameters describe the position and orientation of the camera in space. Figure 1 illustrates a perspective camera projection model. The optical axis, which is orthogonal to the retinal plane $\Re$, passes through the center of projection C and intersects with $\Re$ at the principal point c on the image plane. The distance between the center of projection C and the retinal plane $\Re$ is the camera focal length f. Let M denote the world coordinate of some point on the tip of the wine bottle. The corresponding point m on the retinal plane is the intersection of the line that passes through M and C and the retinal plane $\Re$. Thus, intuitively, what the camera can see is the volume inside the infinite pyramid whose apex is C and the four lines that form the edges of the pyramid pass through the four corners of the retinal plane, as illustrated in Figure 1. A detailed derivation of the projection matrix that maps the world coordinate to the retinal plane coordinate can be found in computer vision books such as [17].

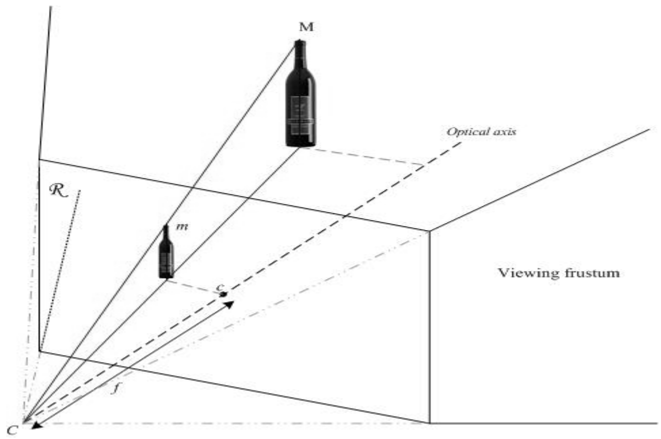


Figure 1 Perspective camera projection model

This pyramid is referred to as the *graphics frustum.* Graphics frustums for rendering are often truncated with near and far clipping planes, where the near clipping plane avoids rendering of objects too close to the camera or at the singularity point C and the far clipping plane avoids rendering objects so far away from the camera as to be considered no longer visible. The near and far clipping planes also serve a practical function of limiting the required precision of depth buffer values. The graphics frustum for the virtual elements must match the camera viewpoint and intrinsic parameters in order to mimic the view of the camera and have the virtual objects accurately registered with the camera image. The far and near clipping planes must clear any rendered virtual elements.

3.2.1 Registration

Assuming a calibrated camera, a reasonable assumption as camera calibration is a common process,, the camera focal length *f*, the center of projection *C*, the principal point *c* and the size and position of retinal plane $\Re$ are known to the system. Before setting up the viewing frustum, the viewport rectangle needs to be set as the same resolution (viewport rectangle size in pixels) as that of the camera retinal plane in order to have the same view of the camera if the frustum is set according to the camera's intrinsic parameters. For example, if the camera resolution is 640 by 480 pixels, then the size of the viewport rectangle needs to be set as 640 by 480 as well. The parameters defining a viewing frustum that match the camera point of view are (*l, b, -n*), which specifies the 3-D coordinates of the lower left corner of the near clipping plane, and (*r, t, -n*), which specifies the upper right corner of the near clipping plane [18]. The values of *l, r, t,* and *b* are as follows.

$$l = -n \cdot \frac{c_x}{f};$$

$$r = n \cdot \frac{w - c_x}{f};$$

$$t = -n \cdot \frac{c_y}{f};$$

$$b = n \cdot \frac{h - c_y}{f};$$

(c_x, c_y) is the 2-D coordinate of principal point c in the retinal plane $\Re$, w and h are the width and height of the retinal plane $\Re$, f is the camera focal length. These parameters much be measured in the same unit, usually pixels. n is the distance from the camera to the near clipping plane and is of the unit as l, r, t, and b. The virtual objects that are rendered in this frustum are well aligned with the camera image.

3.2.2 Zooming

Given this model, both the camera image of the real world and the virtual augmentations appear on the Tablet PC display with the camera point of view. It will provide users a more realistic view if the image can be properly zoomed and shifted as if it was seen from the user's point of view, yielding the effect of a 'magic frame' or magnifying glass. Moreover, the Tablet PC window is usually larger than the camera image and therefore able to accommodate more augmentations. Figure 2 illustrates the different viewing frustums caused by the camera and the user's point of view.

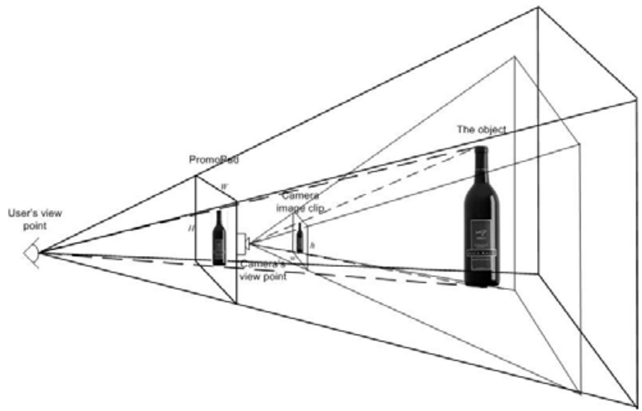


Figure 2 Perspective view frustums from different point of views

The construction of a frustum from the user's point of view has two steps. First, the viewport rectangle needs to be set as the same size of the display window. In other words, if the display is in full screen mode and the screen resolution is 1024 by 768, then the viewport rectangle has to be set as 1024 by 768 too.

The frustum that matches the user's point of view is defined as follows.

$$l = -\frac{W}{2 \cdot p_x}$$

$$r = \frac{W}{2 \cdot p_x}$$

$$t = -\frac{H}{2 \cdot p_y}$$

$$b = \frac{H}{2 \cdot p_y}$$

W and *H* are the window size in pixels and p_x, p_y are the number of pixels horizontal and vertical per measurement unit respectively. The distance to the near clipping plane is set to the distance from the user's point of view (d_u) to the display window, which can be normalized in the application or obtained by existing eye-tracking systems [19]. This frustum model matches to the user's point of view.

The camera image has to be adjusted to match the frustum model from the user's point of view. Since the camera is rigidly attached at the central axis of the Tablet PC, it is reasonable to assume that the user's point of view and the camera's point of view are along the optical axis. Figure 3 shows the vertical 2-D view of the perspective view with two situations. If the camera captures a bigger view as shown in Figure 3a, then the mapping is just a truncation of the invisible part and projection to the display window. If the camera captures a smaller view as shown in Figure 3b, than the camera image will be mapped to the display window according to the proportion without any truncation. However, a small area near the borders is out of sight of the camera and will not be displayed.

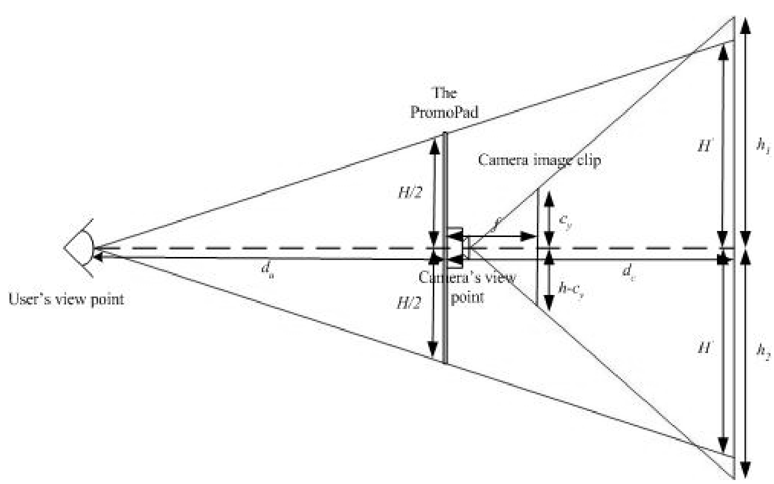
Let d_u be the distance between user's point of view and the tablet and d_c be the distance between the camera view point and the focal object. Then:

$$H' = \frac{H \cdot (d_u + d_c)}{2 \cdot d_u}$$

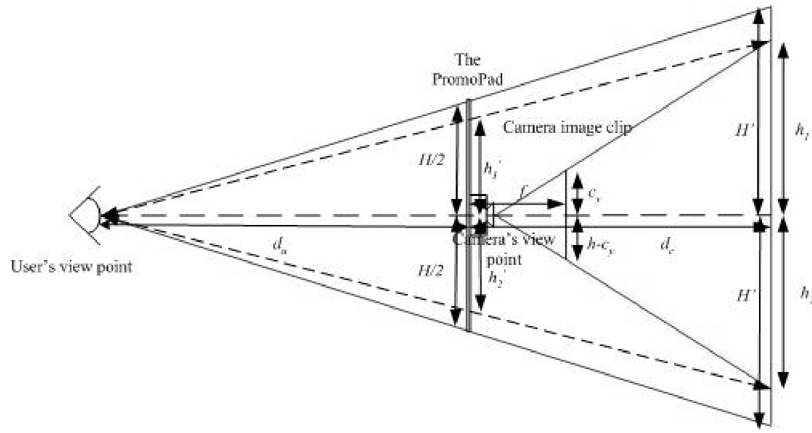
$$h_1 = \frac{c_y \cdot d_c}{f}$$

$$h_2 = \frac{(h - c_y) \cdot d_c}{f}$$

A comparison between H' and h_1, h_2 tells whether the image will be larger or smaller than the viewport on the display. If the image is larger than the display, then the upper $h_1 - H'$ and lower $h_2 - H'$ are truncated. If the image is smaller than the display, the camera image maps to $\left(-h_1', h_2'\right)$, where $h_1' = \frac{h_1 \cdot d_u}{d_u + d_c}$, $h_2' = \frac{h_2 \cdot d_u}{d_u + d_c}$. The horizontal projection can be calculated the same way.



a. camera captures bigger view



b. camera captures smaller view

Figure 3 Vertical 2-D view of the perspective frustums

Although there will be some blank area near the border in some cases, this is negligible in the limited depth of product shelves and can often be covered over with augmentations. Another limitation of this mapping is that it scales identically over the entire camera image, since the captured image is 2D. The zoom operation doesn't differentiate the depth of different objects in the camera image. However, the depth is correct for the focal object, which is the focus of the display in this application. Since the object and the tablet are simultaneously tracked, the depth, and therefore the zoom factor, is continuously varied to keep the focal object the correct size. Figure 4 shows the different effects of the Tablet PC window when the display is from the camera's point of view and adjusted to the user's point of view. A cereal box is behind the Tablet PC and in the scene of the camera. The PromoPad captures the cereal box and augments the view with a nutrition bar and a piece of advertising information. Figure 4a shows the augmented view from the camera's viewpoint. The adjusted view is shown in Figure 4b with the effect of a 'magic frame'.

a. display from the camera's point of view *b. display zoomed to user's point of view*

Figure 4 Tablet PC displays from different viewpoints

3.2.3 Composition

At this point both the real image and virtual elements are properly aligned and adjusted to the user's point of view. The next step is to compose the real image and virtual elements to produce a final image that convoy dynamic contextualization. Augmenting context in the foreground is relatively straightforward since it does not involve any 'mixing' of the different sources, the virtual elements can be simply overlaid onto the camera image to provide augmented context.

Placing augmentations in the background or immersing into the shelf display, however, is more technical challenging. The contour of the front objects needs to be determined and modeled using an *occlusion model* so that the front objects accurately occlude the virtual object in the background. The occlusion model is rendered in the graphics system as a transparent (invisible) object. Since the invisible object is in front (virtually) of the background, the occlusion model creates a hole in the overlay image such that the underlying graphics show through without occlusion. An occlusion model simulates the occlusion that would have occurred had the real object been a virtual object in the graphics system. In an immersive setting, the depth of the virtual object needs to be compared with all of the real objects or other virtual objects that may occlude it. Diminishing context is achieved by augmenting over the competition with background or contextual settings of the focal product to yield a virtual diminished view.

4. Dynamic contextualization in the PromoPad

With dynamic contextualization, the administrators of the system can control the interest flow of the users by virtually modifying the focal entity's context information. In a shopping environment, dynamic contextualization is a business strategy that the retailers can use to virtually change the product settings, placed objects in more complementary settings or removing competing or uncomplimentary products that may be in close proximity to the focal product.

4.1 Context-awareness in the PromoPad

Context-aware computing takes advantage of context information to provide services that are relative to the spatial or temporal context. Context is the situation of a person, place, or object at one time [20]. Given the requirements of a shopping assistant in a shopping environment, *location context, user context* and *product context* are utilized in this project. Location context of the user and product is acquired by the in-store tracking system.

User context is the consumer's profile in the store including such data as brand preference and buying history. User context also includes individual and aggregate behaviors based on shopping habits and demographics. Each time the consumer checks out, purchases are recorded in the store membership database. These systems are already common in many stores that include *loyalty cards* and there is evidence that many consumers utilize these systems [21]. From loyalty card systems or future, automated variations, stores can create personal profiles based on the previous purchases that the consumer has made. For non-member consumers, a generic profile or demographics sensitive profile can be used. The PromoPad system learns the consumer’s shopping pattern, and furthermore, the brand or product preferences. Then it is able to deliver personalized information that is tailored to the customer’s individual needs instead of flooding the shopper with a large amount of generic information.

Product context is the use of a set of *complementary products* that are associated with the focal product or the product under inspection. A complementary product is a product that enjoys an associative relationship with the focal product. By contextualizing the focal product with a matching product, image or symbol, the consumer’s attitude toward the focal product can be influenced [22].

Functional complementary products are products that can be consumed jointly in order to facilitate some operational relationship. For example, golf clubs can be functionally complemented by golf balls, bag, shoes, etc. A user purchasing hot dog buns is likely to purchase the hot dogs to place in them as well as chili, mustard, and ketchup. Hence, functionally complementary products can have very close relationships that influence simultaneous purchase.

Aesthetic complementary products are products that are consumed because they form inherently pleasant relationships with each other. Consumers’ motivation in using these products is the aesthetic pleasure derived from their juxtaposition. For example, a baroque painting in a baroque design house is aesthetically complementary to the house. Aesthetic complementary is often highly subjective; hence it is not currently included in our experiment design, though use of experts may allow for aesthetic suggestions [22].

Sociocultural complementary products are groups of products that involve consumption activities and/or products that hold little or no inherent relationship to each other, but are instead related through a sociocultural process of association and ascription of meaning. Groupings are valued for the ability to communicate social messages within a particular culture at a particular historic moment. For example, we may easily socioculturally associate BMWs with MBAs, Rolex watches, etc. And tie-dyed t-shirts are always socioculturally associated with patched blue jeans, army fatigue jackets etc. Table 1 lists some examples of product complementarity as used in the PromoPad evaluation products database.

Table 1 Product complementarity examples

Focal Products	**Functional Complementarity**	**Sociocultural Complementarity**
Digital camera	Photo papers, memory card, printer for digital camera, picture-editing software	Vacation package, plane ticket, ball park tickets
PDA	PDA keyboard, PDA software, Wireless Internet access, memory	Tie, pen, cell phone, laser pointer pen
Perfume	Body wash, deodorant, antiperspirant	Jewelry, candles
Pen	Notebook, highlighter, pencil jar	Hair tie
Candy bar	Soda, popcorns, ice cream	Ball park tickets, Big ‘n’ Tall clothes or shoes
Wine	Wine stand, cork screw, glasses	Crystal container, romantic dinner, travel package to winery
Detergent	Fabric softener, stain remover	Glass cleanser, floor cleaner

4.2 Dynamic contextualization with augmented reality

Product contextualization is the placement of a product in a particular setting that will resonate with consumers and make clear the product consumption practices or situations [23]. Product contextualization is often seen in store displays and advertising. In electronic commerce, product contextualization can be simulated with 3-D visualization, which can offer a variety of ways for the consumer to arrange a focal product with other complimentary products on the computer screen. Researchers use this *virtual contextualization* to place complimentary products with a focal product in 3-D visualization in order to affect the user's perception of the focal product [4]. For example, the user can arrange a set of furniture in different settings in 3-D on a website to select the preferable combination. Research has demonstrated that virtual contextualization can lead to better consumer experience, brand attitude, and hence influence purchase intention [24].

Augmenting context is the most common implementation of augmented reality systems, as suggested by the name augmented reality. By adding context to the focal product, the PromoPad is able to give the consumer information about the focal product that is not possible in traditional media. Theoretically, the added context can be coupons, advertisements or complementing products as discussed in previous sections. Based on the advertiser's needs, these pieces of information could be 2D pictures or 3D objects that appear alongside, in front of, or behind the focal product. Information can even appear to be immersed into the shelf display. Content in the display can even have depths deeper than the physical shelf, allowing a virtual extension of the store space. In this case the occlusion model is omitted for parts of the shelf that are in front of extended depth material, so that physical objects can be occluded by virtual objects that are actually at a deeper physical depth. Likewise, PromoPad can place information such as complementary settings of the product into the background of the focal product with proper occlusion. Although it may not draw the consumer's active attention, the new information affects the consumer's attitude towards the product.

Whereas augmenting context highlights the focal product by delivering augmented virtual objects to the consumer, ***diminishing context*** emphasizes the focal product by hiding the surrounding product items, most likely non-complementary products or competing brands. Removing the competition allows more room to display information for the product that the retailer plans to introduce to the consumer or increase the sales volume of at that period of time. Diminishments are accomplished through manipulation of the occlusion model so as to selectively occlude elements of the foreground that are not desirable in the final image. The PromoPad can replace the surrounding competition with complementary settings of the product or specific suggested complementary products.

Table 2 lists several possible examples of augmentations and diminishments of the focal products from Table 1.

Table 2 Examples of augmentations and diminishments

Focal Products	Augmentations	Diminishments
Digital camera	Picture slideshow, feature demonstration, accessories	Outmoded models, security locks and latches, film camera
PDA	PDA keyboard, PDA software, Wireless Internet access, memory	Security locks and latches, laptop computers
Perfume	Flowers, romantic pictures	Disliked brands or scent of the consumer*
Pen	Notebook, grade report, back to school picture	Crayon, scissors
Candy bar	Cartoon characters, ice cream	Mint drops, energy bar
Wine	Glasses, roses, picture of a grand banquet	All other than the bottle under inspection
Detergent	Picture of silk or wool, movie clip shows the effect after use	Unfavorable ingredient varieties*

* This is determined by user context, hence it is user dependent.

5. Summary

This paper presents the concept of a shopping assistant that utilizes augmented reality technologies to provide personalized advertising and in-store shopping assistance based on dynamic contextualization. This PromoPad system is a step towards pervasive and ubiquitous computing in the highly lucrative grocery shopping segment. The development goal is to offer a pleasant and inviting shopping experience that is mediated by an augmented reality-based Tablet PC. The paper describes the technical implementation of the video see-through augmentation system and how this technology makes possible the concept of dynamic contextualization, the modification of context to direct the interest flow of users. Dynamic contextualization, the real-time modification of context, can be enabled by augmented reality technologies with augmentations and diminishments of the perceived visual context. Dynamic contextualization is based on, but extends beyond, the spatial and temporal context of the user. Location context, user context, and product context are integrated in this design to address the requirements of an intelligent context-aware shopping assistant.

This work describes the underlined concept and technical details of the PromoPad. Subsequent user studies and in-store testing will be addressed in later paper.

The concept of dynamic contextualization and the design methodology of the PromoPad system can be extended to other circumstances such as tourism guides, training assistants, etc. Nevertheless, designers of other systems need to carefully consider the context factors based on the requirements of an application domain.

Although this article has addressed several important issues in designing the PromoPad, there are still some important issues that must be addressed in the future development of PromoPad as an e-commerce system.

User privacy has to be protected. The privacy issue arises when the retailers collect the consumption activities and attempt to predict the consumer's interest based on her previous shopping behavior. It is necessary to balance the tradeoff between automation and privacy to meet the needs of both retailers and consumers. Consumers may be willing to sacrifice certain degree of their privacy in return for certain perceived value, and retailers definitely should respect the privacy of their customers. A lot of on-going research works are concerned of privacy and security of online e-commerce systems [25, 26]. The goal of the initial development of PromoPad, however, is to maximize the automation of the shopping experience and explore the potential of dynamic contextualization and the possible applications of augmentations and diminishments in the perception of a short shelf image. Hence the privacy issue is beyond of the scope of this initial work and this article.

Another dilemma in this design is the tradeoff between user flexibility and automation. Maximizing the automation requires little user effort but limits the user's flexibility at the same time. Although the design has been deliberated in tailoring the information flow to fit individual needs, some advanced user would like to have more control over the augmentations and diminishments in the PromoPad. This dilemma can be arbitrated by analyzing a survey of usability questionnaire to samples of consumers with various education backgrounds. Also, options giving consumer-appropriate control over the PromoPad can be provided for advanced users to balance the tradeoff.

Acknowledgements

This work is supported by a Microsoft Research Grant.

References

1. Kourouthanassis, P. and G. Roussos, *Developing Consumer-Friendly Pervasive Retail Systems*, in *PERVASIVE computing*. 2003. p. 32-39.

2. Chan, W., *Project Voyager: Building an Internet Presence for People, Places, and Things*, in *Media Laboratory*. 2001, Massachusetts Instititute of Technology: Cambridge, MA. p. 57.
3. Asthana, A., M. Cravatts, and P. Krzyzanoowski. *An indoor wireless system for personalized shopping assistance*. in *IEEE Workshop on Mobile Computing Systems and Applications*. 1994. Santa Cruz, California: IEEE Computer Society Press.
4. Li, H., T. Daugherty, and F. Biocca, *Characteristics of Virtual Experience in Electronic Commerce: A Protocal Analysis.* Journal of Interactive Marketing, 2001. **15**(3): p. 13-30.
5. Armata, K., *Signs that Sell.* Progressive Grocer, 1996. **17**(21).
6. Azuma, R.T., *A Survey of Augmented Reality.* In Presence: Teleoperators and virtual Environments, 1997. **6**(4): p. 355-385.
7. Tang, A., et al. *Comparative Effectiveness of Augmented Reality in Object Assembly*. in *Proceedings of ACM CHI '2002*. 2002. Darmstadt, Germany.
8. Vlahakis, V., et al., *Archeoguide: An Augmented Reality Guide for Archaeological Sites*, in *IEEE Computer Graphics and Applications*. 2002. p. 38-50.
9. Gibson, J.J., *The Senses Considered as Perceptual Systems*. 1966: Boston: Houghton Mifflin.
10. Li, H., T. Daugherty, and F. Biocca, *Impact of 3-D Advertising on Product Knowledge, Brand Attitude, and Purchase Intention: The Mediating Role of Presence.* Journal of Advertising, 2002. **31**(3): p. 43-58.
11. Heeter, C., *Interactivity in the Context of Designed Experience.* Journal of Interactive Advertising, 2000. **1**(1).
12. Li, H., T. Daugherty, and F. Biocca, *The role of virtual experience on consumer learning.* Journal of Consumer Psychology, 2003. **13**(4).
13. Hightower, J. and G. Borriello, *A Survey and Taxonomy of Location Systems for Ubiquitous Computing*. 2001, University of Washington, Computer Science and Engineering.
14. Owen, C.B., F. Xiao, and P. Middlin. *What is the best fiducial?* in *The First IEEE International Augmented Reality Toolkit Workshop*. 2002. Darmstadt, Germany.
15. Tuceryan, M., et al., *Calibration Requirements and Procedures for a Monitor-Based Augmented Reality System.* IEEE Transactions on Visualization and Computer Graphics, 1995. **1**(3): p. 255-273.
16. Owen, C.B., et al., *Augmented imagery for digital video applications*, in *CRC Handbook of Video Databases*. 2003, CRC Press LLC.
17. Shapiro, L.G. and G.C. Stockman, *Computer Vision*. 1st edition (January 23, 2001) ed. 2001: Prentice Hall.
18. Neider, J., T. Davis, and M. Woo, *OpenGL Programming Guide*. 1994: Addison-Wesley Publishing Company.
19. Haro, A., M. Flickner, and I.A. Essa. *Detecting and Tracking Eyes By Using Their Physiological Properties, Dynamics, and Appearance*. in *Proceedings IEEE CVPR 2000*. 2000. Hilton Head Island, South Carolina.
20. Dey, A.K. and G.D. Abowd. *Towards a Better Understanding of Context and Context-Awareness*. in *the Workshop on The What, Who, Where, When, and How of Context-Awareness, as part of the 2000 Conference on Human Factors in Computing Systems (CHI 2000)*. 2000. Hague, Netherlands.
21. Mauri, C., *Card loyalty. A new emerging issue in grocery retailing.* Journal of Retailing and Consumer Services, 2003. **10**(1): p. 13025.
22. *Queer eye for the straight guy*. 2003, Bravo Television Network.

23. Englis, B.G. and M.R. Solomon, *Using Consumption constellations to Develop Integrated Marketing Communications.* Journal of Business Research, 1996. **37**(3): p. 183-191.
24. Host, B. *The Impact on Consumer Behavior by Virtual Reality: Survey in the German Furniture Market*. in *Proceedings of the 2001 Experiential E-commerce Conference*. 2001. East Lansing, MI.
25. Duh, R.-R., K. Jamal, and S. Sunder, *Control and Assurance in E-Commerce: Privacy, Integrity, and Security at eBay*, in *Taiwan Accounting Review*. 2002. p. 1-27.
26. Smith, S.W., *WebALPS: A Survey of E-Commerce Privacy and Security Applications*, Department of Computer Science/ Institute for Security Technology Studies, Dartmouth College: Hanover, New Hampshire.