Augmented Reality 3D Interactive Advertisements on Smartphones

Fadi Chehimi, Paul Coulton, Reuben Edwards
Department of Communication Systems, InfoLab21,
South Drive, Lancaster University, Lancaster, LA1 4WA
f.chehimi, p.coulton, r.edwards@lancaster.ac.uk

Abstract

Whilst Augmented Reality (AR) has been a prevalent research topic it has proved difficult to implement and apply in commercial situations as it generally requires complex and expensive hardware. With the proliferation of mobile phones amongst the world population with ever increasing sets of advanced features such as cameras and high resolution screens AR has the opportunity to emerge from laboratories and enter the high street. In this paper we discuss a unique system that will allow complex and highly interactive visual 3-D adverts to be viewed on mobile phones equipped with cameras. The 3-D adverts are contained within a novel visual code design which allows the system to be deployed without the requirement for additional network infrastructure making it both practical and affordable for advertisers.

1. Introduction

Augmented Reality (AR) describes technologies that enable users to see and interact with virtual computer generated content which is superimposed on the real (physical) world. It enhances users’ perceptions of the world by mixing a view of it with virtual elements relevant to their context [1]. In other words, AR introduces a bridge between the real world and the virtual one allowing content that exists only virtually to be contextually associated with real places or objects for simple and efficient access.

AR is a well established academic research topic but has yet to make an impact on the general public as it often requires highly specialized tools and equipment. Traditionally, the overlaid virtual objects are viewed through wearable head mounted displays (HMD), processed by portable computing units, and interfaced with via mice, tracker-balls, digital gloves [2], stylus pens, buttons, or specially designed visual code markers [3].

With the rapid development of mobile phone capabilities and their feature sets, AR applications have become achievable on a larger scale [4,5,6]. In fact many current mobile phones integrate all the required lab equipment for AR onto one device, albeit in a less complex form. For example, most have full-colour LCD displays and integrated cameras, representing the HMD; increasingly powerful processors, replacing the portable computing units; and an array of possible input mechanisms, though generally limited to ITU keypad and joy-pad which are more restrictive than the previously defined options.

As the number of worldwide mobile phone subscribers has surpassed the two billion milestone (and rapidly approaching three billion), advertisers have the mouth watering possibility of accessing nearly one third of the estimated world population [7]. A previous research [7] has highlighted possible strategies for advertisers to reach this audience and how they might deploy successful mobile advertising campaigns and the possible limitations they may face. Further, we have previously presented “3D M-Ads” in [8] which is the first system capable of delivering intuitive 3-D advertising onto mobile phones. In this paper we extend this approach by introducing another collaborative, context and privacy aware system that will open AR as a viable option for mobile advertising. It is an opt-in, client-server advertising service with virtual objects superimposed onto views of the real world through the screen of mobile phones.

In this paper we will concentrate on the mobile client side of the system which is a camera-based mobile phone application targeting the Symbian OS platform. It is designed to locate and read specially designed coloured tags which will allow us to display 3D advertising content on top of the captured camera frames. In section 2 we describe the operation of the application and how customers are able to interact with it. Sections 3, 4 and 5 present the solutions to the main questions related to the system design:

- What are the design requirements for the tag?
- How does the mobile phone decode such a tag?
- And, how is the superimposed content being rendered on screen?

In section 6 we identify some possible difficulties that could face the system. Then we propose two distribution mechanisms alongside the application’s potential future development and usability in section 7 before we draw our overall conclusions in section 8.

2. The AR approach to mobile ads

AR systems have been prototyped mainly in fields such as education [9], entertainment [10], communication [11] and remote control [12,13]. Purely
business-oriented applications have rarely been pursued due to the commercial difficulties and cost of supplying AR systems in sufficient numbers to the public. However, with the current generation of mobile phones existing in such large numbers and becoming the most carried items along with keys and money AR ads have become a viable business opportunity. This project presents an AR system for interactive, entertaining and informative advertising on mobile phones where customers involve with 3D-oriented experience in a mixture between the real and virtual worlds.

2.1. Application rationale

The general public wear their mobile phones everywhere. They get exposed to casual advertising also everywhere: billboards, TVs, busses, newspapers etc… Instead of limiting this experience to the quick-sight, flip-over and likely forget operation the proposed system would allow individuals to actively interact and communicate with a new media form that embeds advertisements within its content using the communication devices they are already familiar with and carry in their pockets all the time.

Wherever traditional ads may be found, we believe the specially designed tags developed for this project, shown in Figure 1, can replace or augment them for a greater visual experience, more subliminal effect and better satisfying results. The interaction process is simple and easy-to-use for all age categories. The user will only ‘target’ the tag with his/her camera phone running the associated AR software, termed ARMAD (Augmented Reality Mobile Advertisements), which will then automatically download and ultimately superimpose 3-D model(s)/information on the location of the tag on screen. The downloaded content could be, but not restricted to, one of the following:

1. A 3-D model representing the product or the brand being advertised enabling better viewing, awareness and affection
2. Interactive textual information directing users to location-based services and/or offers
3. Visual links to events’ and/or promotions’ mobile web sites
4. Location-based nodes for product-awareness games
5. Guiding avatars used for shopping centres/malls providing offers and promotions in addition to information about sales and directions

2.2. The interaction process

In order to view and manipulate 3-D models on screen the user simply has to rotate around the tag, if it is fixed, or rotate the tag, if it is mobile. The products/info in the content of the ad will then be viewable from all angles as in Figures 2. Users can also zoom in/out of the model by moving their phones closer to or further away from the tag, but keeping it within the boundary of the screen, Figure 3.

It is possible for the user to select one model if many are displayed on screen by simply moving the phone so that s/he points the crosshair in the centre of screen, which can be viewed and hidden, to the top of the desired model and then presses the selection button.
If available, ‘floating’ text will be displayed on screen to enable the user to retrieve further details of the product/promotion, or to download it for future use. This is achieved in the same manner as for selecting a model by using the crosshairs to select the augmented text.

The user is able to delete or save instantly the downloaded model for future retrieval. Currently the download process is limited to Bluetooth. Ultimately an over-the-air protocol via GPRS (General Packet Radio Service) will be implemented to provide greater flexibility. Other technologies such as Wi-Fi and Wi-MAX would be addressed once sufficient amount of mobile phones supporting these technologies is widely available in the market. Having a wider span of accessibility is the turning point for ARMAD to be a successful mass advertising channel.

3. Tag design

The target tag used to view AR ads with ARMAD is composed of four regions as shown in Figure 4 which are:

1. a coloured area in the centre
2. a black border surrounding the coloured area
3. a code area as a boundary of the centre
4. and 4 black corners of the code area

The coloured area is composed of four right-angle-triangles: two red, whose tips are facing each other, one white to the right and one blue to the left. These triangles are used to find the centre point of the tag, which is the point where their tips meet, and to determine its orientation. The centre point is used later to find the corners of the tag with the help of the black border. The black border region is used to find the corners of the coloured area first then will help in determining the corners of the tag as a whole. The next section will describe the details of the detection.

The code area contains the data associated with each tag depicted in bits. It is composed of black and white blocks with four black bits reserved for tag corners. The blocks represent binary data and are capable of representing $2^{58}$ word options.

Specific sections of the code sequence can be reserved for such things as tag ID or the URL/Bluetooth address where the model may be obtained from. The rest of the code area is left for customized usage. Coded data distribution is flexible according to requirements, with slight modification to application code.

The overall dimension of the tag is 11cm by 11cm for height and width divided as follows:

1. 5cm of both width and height for the coloured area
2. 1cm width and 1cm height for each bit block (black or white)
3. 1cm thickness for the black border of the coloured area (its total dimension is 7cm by 7cm with the coloured area inside)

The relatively large size of the tag is useful for outdoor placements of ads which will facilitate detection and viewing from far distances, and would allow more than one user to have access to in the placement proximity. However, it is not practical when applied to print media like magazines and newspapers, or even digital media like PC monitors. Work will be conducted to optimize the dimensions of the tags for such space-limited mediums depending on potential customer requirements.

4. Tag detection process

The detection process of the tag starts with the coloured triangles. Once ARMAD is launched, the camera on the phone starts capturing continuous frames and the application starts analyzing each frame individually to find a potential tag. The detection process is applied through the following six steps by using optimized image processing techniques for coloured-target detection.

4.1. Filtering frames

Before initiating any detection process there are two filters applied on every video frame captured by the camera in order to optimize tags detection process. First every pixel in a frame is compared to a colour threshold to classify it in one of the five colour buckets: blue, red, white black and yellow. The first four will be used literally: for instance whatever is ‘redish’ in the captured frame will be placed in the red bucket. The yellow colour on the other hand will replace whatever is not red, blue, white or black in a frame.

The second filter is applied to fix the yellowed pixels on edges in the initial image where anti-aliasing is applied. As the colours used in this effect most of the time do not blend to any of the four buckets they are yellowed. Usually these edges are on the lines that separate adjacent triangles in the colour area. Having yellow pixels in that critical region will affect the detection process by falsifying true targets.

![Figure 4. Tag design, regions and dimensions](image-url)
4.2. Finding tag centre point and orientation

The centre of a tag is identified, as long as it lies within the captured frame, by finding the pixel where red pixels are above and below it, a blue pixel is to its left and a white one is to its right with an appropriate threshold (4 pixels used here). The threshold is necessary to facilitate centre detection even when blurring occurs while moving the phone relatively fast.

The positions of the coloured pixels surrounding the centre point represent the triangles of the coloured area and they are relative to the orientation of the tag in regards to the phone. The previously mentioned red-red-white-blue sequence of colours is applicable when the tag is upright (rotated 0°). This will change according to its rotation degree. For example, if the tag is rotated 90° clockwise (CW), the detection of colours will take a different sequence: white above, blue below and red left and right. Table 1 below shows all possible orientations with their specific triangles colour sequences to use when detecting tags.

<table>
<thead>
<tr>
<th>Angle of rotation (CW)</th>
<th>Triangle colours (above, below, right, left)</th>
</tr>
</thead>
<tbody>
<tr>
<td>315° &lt; Θ ≤ 45°</td>
<td>Red, red, white, blue</td>
</tr>
<tr>
<td>45° &lt; Θ ≤ 135°</td>
<td>White, blue, red, red</td>
</tr>
<tr>
<td>135° &lt; Θ ≤ 225°</td>
<td>Red, red, blue, white</td>
</tr>
<tr>
<td>225° &lt; Θ ≤ 315°</td>
<td>Blue, white, red, red</td>
</tr>
</tbody>
</table>

4.3. Finding corners of the coloured area

The logic used to find corners traverses the tag vertically upwards and downwards from the centre point. Once it encounters an edge pixel where black colour is above it and red is below it, in case of moving upwards, or black below and red above, in case of moving downwards, the traversal is stopped and the location of that pixel is stored. The next step is to move right and left from that pixel until the end points of that edge, which represent the coloured area corners, are reached.

In case the tag is rotated finding the two end points of the edge will need to perform one additional step. A zigzag movement is taken from the centre point to get as close as possible to the end points required.

4.4. Estimating tag corners

Once the corners of the coloured area are identified finding the tag corners becomes a matter of mathematical replacement. In the design specifications above it was stated that the coloured area has a width and height of 5cm each. Given that the code area has a depth of 2cm and the black border has a depth of 1cm the position of the code corner can be calculated as a \( \frac{8}{5} \) factor of the distance between two diagonal colour area corners. For instance, the tag corner c1 can be approximated using the following equation from the coloured area diagonal corners a1 and a3:

\[
c1(x, y) = a3(x, y) + \frac{8}{5} (a1(x, y) - a3(x, y))
\]

4.5. Determining the starting point for reading

Having had the orientation classified and the tag corners located enable the application to start reading its embedded data. Reading always starts from the upper-left corner of the tag. The system will be able to find this point even if the tag is rotated or skewed since the software has already flagged each coloured area corner with its appropriate position in step 2: upper-left, upper-right, lower-left and lower-right.

4.6. Extracting the code

The last step in detecting a tag performs reading the code data in a similar fashion to approximating the tag corners in step 4. Each tag edge is traversed between its two end points, i.e. tag corners, with one eleventh of its length, as each bit block has 1cm width and 1cm height. The one-eleventh ratio is applied vertically and horizontally while scanning the code from top to bottom and from left to right. The tag is scanned line by line and each block’s colour is retrieved to construct the data embedded in it.

5. Rendering the tag content

The first initial step of detecting a tag in a camera frame is accomplished and next comes the process of rendering the content embedded in it onto the mobile phone screen. This is achieved by using the graphics library OpenGL ES which is a cut-down version of desktop OpenGL tailored to meet the limited capabilities and resources on embedded systems (ES) [14].

The main concept applied here is to find 3 points on screen, 2-D surface, and try to reverse-project them to their 3-D position in the 3-D world of OpenGL ES. Once those 3-D points are calculated, it will be possible to find a transformation method that would render the 3-D object of the ad in its appropriate perspective relative to the surface the tag is placed on. The three points we use for this purpose are the centre point of the tag, the lower-right corner and the lower-left corner.

6. Possible difficulties

There are two main streams of difficulties that may face ARMAD: technical and social difficulties. As the tag design contains colours and the detection process depends solely on them, it is essential to have tags well illuminated with white light. This is a technical issue which having failed to comply with may cause the
application not to detect coloured areas and hence not to display objects. This is a limitation of using colours in the design but it comes as a trade-off to excessive processing, heavy power consumption and jittering view frames that other well-known AR techniques may generate when applied on mobile phones [3].

The unavailability of OpenGL ES on many mobile phones is another technical issue that may deprecate the usability of ARMAD. It limits at the mean time the use of the application to a small but considerable portion of the mobile phones population, 100 million phones [15]. However, with the massive industry interest in OpenGL ES standard this is expected to be overcome or almost resolved in the future. A potential solution could be porting the application to J2ME (Java 2 Micro Edition) which is a cross-platform common to most phones which suffers from reduced accuracy and performance.

The other category of difficulties that may be encountered is related to the social response to the system. Many mobile phone users are not necessarily well-educated about the security and privacy issues in the mobile medium and often take a conservative view in regards to the technology. This could have a negative impact on any mobile service not only ARMAD. The system could be considered as a spam or a virus by users. They might feel worried of using the system fearing that their personal details could be stolen or compromised.

Add to this the concern of costs and the delivery mechanisms to use. With over-the-air delivery of ad content via GPRS users may be charged according to the amount of bytes downloaded/uploaded whilst in the retrieval process. This is expected to be overcome in the future when mobile operators provide flat-rate tariffs for wireless Internet access [16] like T-Mobile’s ‘Web & Walk’ first-of-kind service in the UK [17] and NTT DoCoMo’s wide range of such plan in Japan [18].

Another social issue relates to people getting bored of intrusive and redundant services or ads which repel their future interaction with them. Generally, people enjoy experiences that when contemplate in they find themselves learning or inferring something about the object being represented or advertised [19]. ARMAD addresses this concern as it mixes adverts with rich media, 3D graphics, and new forms of interaction. It provides, as a result, advertisers with an unprecedented opportunity to offer messages that include information in a supporting knowledge environment, and entertainment in an enjoyable format [19]. The 3D contents in the ads define a virtual, psychological experience that consumers undergo while interacting with 3D products in computer-mediated environment [19].

Nevertheless, most newly introduced technological services or products were considered complex and hard to co-op with when first put in the market (e.g. microwaves). Complexity often has negative effect on any product [19]. This issue applies in the advertising spectrum where complexity may arise from new, innovative advertising services, such as ARMAD, becoming more interactive and more user-interfaced. However, researchers have found that more involvement and more interactivity provide greater user learning experience [20] and open up potentials for new dialog forms between marketers and consumers [21].

As the public gets more exposed to mobile services in the future opt-in advertising systems such as ARMAD, or even opt-out, systems would become familiar to them.

7. Usability and potential benefits

Advertisers are expected to spend $11 billion on mobile advertising by 2011 according to a recent report by the telecom division of Informa group [22]. However, mobile advertising is unlikely to be successful unless it offers something new and interesting for the customer. The message or content should be relevant, contextual, permission-based, targeted, personalized, and user-friendly. Also, the content of the ad must be attractive and interactive which are the core of imaginative message design aimed at stimulating consumers’ responses such as attention and attitude towards the ad or the brand, as well as behavioural activity in the market place [19]. These responses are mental imagery processes by which sensory information is represented in memory very like picturing and very unlike describing. With the advent of new media possibilities on the mobile phones such as 3D graphics a new dimension of media-specific marketing messages can be created to stimuli this mental process by which customers translate messages into mental images that in turn stimulate purchase intentions [19].

Advertisers must realize that people seeking mobile services generally have only short periods of time to access such content, say while waiting for a doctor appointment or for a bus. That is an opportunity for advertisers to communicate with these potential customers. The real communication is made only by means of attractive media-richness that are the sole of conveying the spontaneous, intensive perception of reality depicting a vivifying freshness of impressions reconstructed as real pictures in the mind’s eye [19]. Thus, advertisers must make the most out of this little yet crucial time by providing something compelling, exciting, engaging and imaginative to attract users’ attention and provoke their acquisition or willingness to interact with the ad. The experimental poser of vivid adverts lies in their presentation of intellectual and emotional complex in an instant of time [19].

Several mobile advertising models and applications have been introduced and tried so far in the mobile market by different companies and/or organizations. Most of these models are just an inheritance from the Internet industry: banners have started to appear on mobile web pages, ad-links are being hosted on wirt...
Users operate ARMAD by personally dragging their phones to see what the tag in front of them is hiding in its context. It is similar to opening a present box expecting something pleasant. When in action, the user views 3-D models of products, animations presenting a product/brand, human-like avatars giving commercial information, or ad-linked texts which s/he has never experienced before. This engaging and exciting content may unconsciously tighten the user’s relationship with the brand or product being displayed on screen, or at least may influence his/her awareness. This is served for by with 3-D graphics and augmented content in ARMAD. Both provide unique virtual experience that brings into the customer-brand relation more effectiveness through the psychological states of presence and belonging to the proximity of the ad, involvement with its brand/product advertised, and entertainment offered [20]. Studies suggest that there is a constant connection within any medium between the human mind, the technology and the environment that serves to immerse users resulting in augmented learning, altered behaviour, and a perceived sense of control [27]. ARMAD addresses this by mixing users’ experience with the physical environment and the cutting-edge technology.

The application of ARMAD can be applied in wide arenas such as shopping centres, malls, public parks and even museums. In such places people are most likely free and under no time pressure. This gives an advantage to capture their attention and encourage them to interact with this entertaining ad medium, at least to add more fun to their time, via the formation of the virtual experience from 3D advertising [27].

7.2. ARMAD usability scenarios

The method here is based on reflecting advertisers’ message on their customers’ unconsciousness. The concept is to show or deliver messages of products more interactively and uniquely in such a way that is original and unknown to competitors via the medium perfectly understood and mostly adopted by the new generation. Average revenue per person (ARPU) is not the main issue for success here. Rather, ARMAD is more of a marketing-mix vector rather than a standalone marketing method.

Viewing 3D objects on the surrounding environment on people’s mobile phones will be a lucrative and more informative approach to subliminally deliver the qualities of products or services without forcing them on customers. They will be driven by curiosity to interact with the unusual virtual scenes shown on their phones. Building on some work in psychology, many researchers have suggested that in certain antecedent conditions consumers may feel a heightened degree of intrinsic motivation, intense concentration and enjoyment while engaging in technology interactions. This experience could lead to many positive outcomes positive attitude’
increased learning and participation [24]. With ARMAD users will be twinkled with the superimposed content on the environment around them and the new way of engagement with brands and products, which will add fun to their experiments. This indirect tapping on emotions and the sense of absurdity offered are what make customers recall ads and may drive immediate behavioural action with respect to the product [24].

In this sense ARMAD complies with the requirements stated above of entailing something overwhelming, informative and novel to the advertising experience customers pass through which stimulates curiosity [24]. For them it will be amusing, new, eye-catchy and fun, while for advertisers it will be distinguishing, distributable, manageable, and widely profitable.

7.3. Distribution models

Having discussed the technology behind ARMAD and shed some light on how it is meant to operate to make the most effect, we propose here two models of how to distribute the application among the public.

7.3.1. Packaged with operator-branded mobile phones: many operators ship mobile phones branded with their logos and special content and/or applications. Having ARMAD shipped with such packages can open the door for wider spectrum of cooperation between operators and the different poles in the mobile business competition. It would drive more revenue to partners and would make the application locally standardized if not globally. A good example of this is the QR barcode reader provided by the Japanese operator NTT DoCoMo. Almost all phones by the operator are shipped with the capability to reading such codes and any one can create customized codes with the publicly available utilities provided by NTT DoCoMo [28].

7.3.2. Location-based: another option of distribution could be using a back-end system that would send ARMAD manually or automatically to client phones. Users may request the download of the application over the air by typing special short-code number(s), advertised in the appropriate locations, where they would view and interact with visual ads. Or, they may request downloading it from a server in the proximity via Bluetooth, or other wireless technologies once implemented.

In order to have the application sent automatically to phones a dedicated back-end server and central database are needed to track the IDs of phones that have previously downloaded the application. If a phone in the proximity is discovered and found not to be registered in the database it will receive a message recommending the download. However, this method can be costly and may be disturbing to customers. It would also contradict what has been discussed earlier about user-privacy. It may in addition require a nationwide management which in turn costs time and money.

8. Conclusion

Mobile ads by being personal, creative and intimate could solve some of the frustration advertisers have with customers ignoring other traditional advertising formats. However, this interactive medium unlike others has a complex and sophisticated infrastructure for management, delivery and optimization of advertisements. Also, its personal nature is both a plus and a minus. It involves creating exciting chances to talk directly to customers via highly tailored messages. On the other hand it could be intrusive and irrelevant and would be discarver by customers causing the medium to be alienated and destroyed. We believe that interactive and entertaining LBS systems will have the greatest impact of all mobile marketing techniques on customers’ experiences and businesses’ logistic marketing mixes and although this project is in the early stage of evolution it highlights the potential of the field..

9. References


