

HTML5 – AUGMENTED REALITY, A NEW ALLIANCE AGAINST THE OLD WEB EMPIRE?

Alexandru Tăbușcă¹

ABSTRACT

The current paper presents the 2013 realities of the Augmented Reality paradigm in the context of the use of modern web technologies. The article starts with an educated summary, presentation style, of the present worldwide recognized standards for Augmented Reality (AR) applications. I will not only present but also compare and comment on the current state problems, advantages, development directions and guidelines for AR web applications. The focus of this article is set towards the development and use of AR web applications especially for mobile devices like smartphones, tablets, special factor laptops or ultra-books.

Keywords: html5, augmented reality, canvas, video stream, web extras, qr code

1. INTRODUCTION

One of the hottest subject in the IT trends today is represented by the “augmented reality” applications. The most talked about project of the IT giant company, Google, the iconic Google Glass is of course the spear head of this news but, by far not, not the only gadget that is based on augmented reality applications for introduction of a new paradigm in the IT consumer products and services. This trend comes as a normal step forward - today is almost unthinkable that some previous generations were able to cope without the use of computers in their lives. Today, almost every aspect of our life has something to do with a computer. We do not only rely on computers but even the internet became something that we think about as a normal thing, like running water or electricity. Some years ago the internet was even established as a citizen right in Finland, assured by the law [1]. After these facts we really need something new and exciting to drive forward the crowds of developers and users that live on the internet environment. As a consequence, the AR seems the best option to light a new spark on the sky of web interactivity.

Arguably, the best and most successful hardware product already on the market in high volumes today is Microsoft’s Kinect Console. The applications and games developed for this revolutionary device are already taking advantage of face tracking, geo-location and other software features that are available in connection with the augmented reality paradigm. Our children, and not only them, are now able to interact with the software in ways that seemed science fiction only several years ago. We can now virtually travel inside the old Egyptian pyramids, admire the medieval history recreations, walk on the moon or visit the Louvre museum directly from our home PC or even with the help of our super powerful smartphones. Even more, in front of the Kinect device, we can swipe, touch or

¹ Lecturer, PhD, School of Computer Science for Business Management – Romanian-American University; e-mail: alextabusca@rau.ro

push a button, in the air, and we get a lot more detailed information that is available on demand. We can now go from Bucharest to Florence, frame the old Dome inside the sensor of our smartphone and we get all the historical information available in any respectable tourist guide; and that is not all – in certain areas we can even get, through the phone's display, an entire image with the room's setup three hundreds of year ago.

The “augmented reality” concept is now known as AR, abbreviated like almost any other term related to the IT environment ☺. The term actually refers not exactly to “augmentation” but more precisely to “duplication”. The augmented reality, AR, is described as focusing on the replication of real world environment within a computer virtual reality. The AR usually composes a new complex image by mixing the real details (such as location, part of the image, weather condition etc.) with a setup completely generated by the machine, producing a result with “added” information that suggest an “augmented” image – therefor the A in the AR. The virtual composite setup created by the computer has the aim to improve the user's image of a real thing by making him see more details or even interact at some level with real or virtual items. The final aim of the AR concept would be to recreate an entire ecosystem inside of which we cannot differentiate between the virtual and the real elements – a sort of a scenario widely presented in the iconic StarTrek television show, tens of years ago, by the use of a so called “holodeck”. The actual uses of these technologies today are quite often related to entertainment and gaming but we have to remember that a very important role is also played by their insertion within the fields robotics, micro-manufacturing, military simulation or engineering design.

AR was, at the beginning, based on the implementation of code in high level programming languages, like C++ or Java. Today, a new trend is raising – the use of the HTML family of languages for the development of AR applications. The AR can be actually considered the heir of the QR codes. The “Quick Response Codes” can now be regarded as AR beta 1 – they represent 2D codes that triggered an event when scanned and read by a specific QR reader application. Usually the QR codes send the user to a website or another web based resource directly linked with the developer of the QR code. The most distinctive part separating the AR and the QR is actually the destination that the user is pointed to; the QR code let you see the real thing (an object, a catalogue page, a poster etc.) and if you interpret it through a dedicated application you are sent to another location for more information related to the topic. Switching to the AR, when you interpret the AR code through its software application you do not get sent away to another location – you actually receive the new detailed information an see it as a layer over the reality layer that you started from. We can say that QR sends you away while AR brings over other things to you.

The QR code insertion on the market became quite relevant during 2011, mainly because of the very easy way to implement it; just encode the address of a website within a QR code and print or publish it somewhere for the client to see it. AR, on the other hand, is not so easily integrated into our mainstream software or printed environments; the AR in truth delivers a way more complex and exciting way for you to actually interact with you potential clients than in the case of QR but it might not be the right choice for all scenarios. Most printed elements are still relying on QR (magazine publishers, consumer packaged goods, printed catalogues, e-commerce websites etc.) to deliver content, information and support for a virtual shopping experience. The number of cases that a company relies on

augmented reality support is very small – this thing, in itself, bringing also an unexpected bonus for those companies: some complex and unique experiences are quite rare and represent something special for a client. The professionals of the modern advertising industry are agreeing that AR helps create some very interactive experiences, more complex and immersive than the mere QR codes and that are excellent facts for a campaign that aims to dazzle the potential clients. However, on the same agreeing tone, the professionals argue that QR is still the more pragmatic solution. QR is able to provide quite a comparable experience with the benefit of being more versatile and much easier to implement.

AR and QR are mainly regarded as two different evolutionary steps on the same ladder – you first go on with the QR and when you feel ready step over to the AR. In my opinion the matter cannot be viewed as settled just yet. I actually think both QR and AR codes will coexist for quite some time, both being related and not exactly competitors. While the QR is much more suited for the pragmatic and time-aware clients, the AR is going to have more appeal for the enthusiast and gadget-oriented clients. From my personal experience with a world recognized brand such as IKEA², the use of the most modern AR coding can become a very frustrating issue. The 2013 IKEA catalog has some items with attached codes that should trigger an AR experience – just scan a certain image with their dedicated application and you can see the coffee table, for example, imposed as a virtual layer on the live image of your living room seen through the smartphone's display... it sound interesting but it can actually take quite a lot of time to get it right and in some cases it does not work at all. On the other hand, lots of products have QR codes besides them and then scan of these images works 99% of the times and send you to a dedicated webpage with lots of information. From the point of view of usability and user experience the QR is much more mature than the AR, at least at this moment. QR code interpretation requires little, almost no effort from the part of the user in order to access it and offers a great way to see the product “acting” before making a decision to purchase something or not. Also, a big advantage of the QR, as I already mentioned before, comes from the seller's perspective – these codes are quite inexpensive and very easy to implement. The AR can be quite a challenging experience, with the clients required to hold up a phone in front of their face to find the right position for digital overlaps on the physical objects. The production costs of AR can also be much higher, a negative aspect for the seller, but the investment might be worth it if the target client is prone to the use of a fun and immersive experience.

The retail business is the area where both QR codes and AR have found a good breeding ground. Companies always look to upgrade the clients' experiences beyond the physical aspect and into the electronic environment, in the meantime also bringing an amount of modern and added excitement to their stores.

For the following year, during 2014, I think that the largest retailers will continue to spread the use of QR codes as a simple and practical way to engage users and deliver useful content while AR is likely to continue to be used as a means to drive on the user's excitement. I think that a real game-changer for the AR can be the Google Glass. Already, excitement is building around this project for a while now, with almost everybody agreeing the Google

² IKEA – one the largest furniture and accessories retailer in the world

Glass project will finally offer the AR experience in a very simple and easy to use way - via the special glasses that users wear.

While the excitement about QR will decrease, partly because they are not “news” anymore and partly because they will be regarded simply as just a quick short-cut to a digital site, the interest in QR might come to new heights with the public release of the Google Glasses – and for sure with the entire plethora of copies and similar products from other companies.

2. FACTS

The use of AR is now targeted at mobile devices. Most of the test scenarios preview the use of AR with the help of the smartphones or tablets. While indeed the personal computers have much more computing power they also have a larger range of options for bringing the virtual elements to the screen and, most important, these classic computers are quite difficult to imagine in situation when AR is best suited. It would be quite difficult and completely strange to take out the laptop from your backpack, open it, switch its camera towards a mall and start an AR experience while you struggle to keep you laptop in front of you. Even with a smartphone, a phablet or a tablet, the experience can be quite strange – as you might be required to make an entire series of adjustments to the position of the camera in order to obtain the best results.

The best suited technology for implementing AR applications can become, arguably, the HTML5. While in truth HTML5 did not, and will not for quite a long time, replace the aging HTML4 it is best suited for the development of this new type of applications. The HTML5 might actually be described as a mix of web technologies that matured over time: HTML, CSS and JAVASCRIPT. A very good analogy sees the HTML5 “anatomically”: the HTML represents the bones, the CSS represents the skin and the JAVASCRIPT represents the muscles. They were here long before this 5 version but now they seem to have been treated with growth steroids ☺. In fact, I can mention three main reasons why HTML5 might be considered as revolutionary rather than just evolutionary, why the concept has the right to be called HTML5 rather than 4.5.

First of all, in HTML5 the good old webpage does not have to look and act like a “classic” webpages. The rise of the Adobe Flash technology over the last years has mainly been just an attempt to overcome the natural limitations of the things that HTML5 was able to provide. Flash was initially targeted at animating websites and producing special visual effects. Soon after, complete websites emerged as being produced entirely in Adobe Flash, allowing the developers to obtain a different type of navigation and organization of the webpage, more complex access to the certain sections of a webpage, and the ability to avoid the use of JavaScript (actually one can say that they replaced JavaScript with ActionScript).

Secondly, webpages no longer represent, in the majority of cases, one person or organization as regarding the content. The common webpage today is usually an accessory of somebody’s digital world, comprising text content, images, audio or video materials etc.

Lastly, webpages today have to be able to function easily across a lot of display devices. HTML5 grabbed that boat and navigates under the banner of the “mobile” concept.

Actually, at this moment, smartphones, phablets or tablets should be able to provide the same experience as a computer browser.

The HTML5 introduced a new concept of programming interconnection that overpassed the previous limitations. HTML5 and CSS3 provide a very powerful and comprehensive set of utilities, tools and effects that can compete on par with everything that Adobe Flash had to offer. Actually, almost every website that was completely designed in Adobe Flash can now be developed only by implementing HTML5, CSS3 and JavaScript. The most important advantages of this fact lies in exactly the parts where the Adobe Flash was somehow vulnerable: the user can select the text like in classical HTML, you can bookmark any page directly, source code is very accessible again and not wrapped inside a nice .swf package.

All these facts make the HTML5 developed websites completely and easily indexable, the audio and video content is now very easily linked intra or extra website, the overall experience of the website is almost the same on either your office Dell PC, your iPad tablet or your Samsung smartphone.

As I mentioned before, mobile devices can view HTML5 webpages or provide audio and video content without special issues. In fact, HTML5 owes a lot of its interest to mobile devices, and in particular to Steve Jobs³ and his active campaign against Adobe Flash. From the beginning of the war against Adobe Flash, Steve Jobs' Apple provided the iPhone with the Safari browser, which had HTML5 support but lacked the Adobe Flash plugins. Today all mobile phones have HTML5 browsers or the capability to install one. On the other hand, I have to mention some facts that seem to be overlooked quite all the time – actually, HTML5 does not offer a special set of tools for working with mobile devices. Indeed, we have the capability to detect the orientation and screen size of a mobile device but those commands also allow for the detection of the size of a browser window or a tablet display. The canvas element and the new audio and video features make it possible for the developer to deliver content to a mobile device. In truth, HTML5 is not targeted for mobile use – it just works normally on both classical and mobile devices; the developer does not have to build two versions of the website, one for PC and one for smartphones anymore.

As a test case, I will present a very simple application, developed by implementing HTML5 for an AR application. The application develops a virtual slideshow based on an AR code that should be read with any camera, either on a smartphone or a webcam.

For creating this test application I chose a quite simple AR scenario: show the current camera an AR code and the application will impose a layer with another image on the camera live preview; take the code out of the camera's sight and then bring it again – a new image will be imposed on the live preview. Actually, the development is based on the

³ Steve Jobs – founder of Apple Corporation, an innovator and visionary that changed the digital paradigm of the 21st century; among the products that became market anchors and were backed by him we can mention the iPhone or the iPad.

capabilities of the WebRTC⁴ API and the JSARToolKit⁵ library and on the implementation described by Ilmari Heikkinen [2].

To shortly present the schematic algorithm of the application I can outline the main steps involved:

- We have to send an HTML5 canvas element to JSARToolKit
- JSARToolKit returns a list of AR markers found in the image and the corresponding transformation matrices; on top of these markers we will place the AR layers with the images
- we send the transformation matrix to the 3D library responsible for the image creation
- in order to analyze the video that is taken from the camera live preview we have to place the respective video on a canvas element and send it to JSARToolKit for frame-by-frame analysis.

In order to make the application work we need to use a WebRTC capable browser – at the moment of writing this article the WebRTC is fully supported by Google Chrome, Firefox and Opera browsers.

The first step is to download the JSARToolKit and save the corresponding .js file inside our working folder, together with the magi.js file that is also required by the JSARToolKit. Then, we have to build out main webpage, called “RAU-AR-test.html”. Next, we have to select the pictures that we want to impose as the AR layer over the live preview of the camera. We save them in the same folder as the html file and the .js files. Next, we build our AR marker and print it on any piece of paper:



Figure 1. AR code to be printed

⁴ WebRTC – a free open project that enables web browsers with Real-Time Communications (RTC) capabilities via simple JavaScript APIs; information available publicly at <http://www.webrtc.org>

⁵ JSARToolKit – a JavaScript port of FLARToolKit, operating on canvas images and video element contents; available publicly at <https://github.com/kig/JSARToolKit>

Then, we launch our web page in a WebRTC enabled browser - my favorite browser is Firefox - and allow it to access our webcam for creating the background of the AR application:

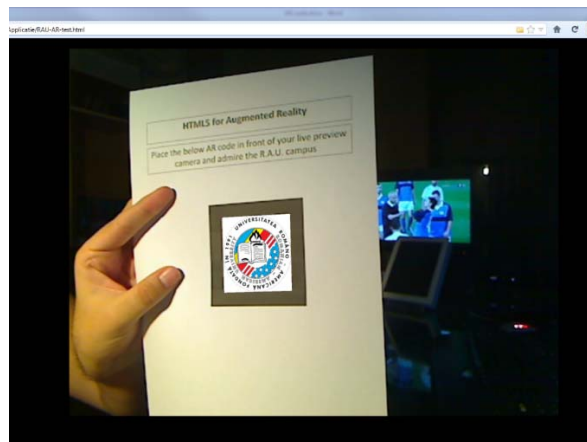


Figure 2. The first imposition of the AR image layer over the marker

Each time that we hide the marker from the camera and show it again, the AR application shall place another image from the batch that we have selected on the markers position. Actually, each detection of the marker in front of the camera would trigger another step in a “for” element that brings about the images that are preloaded by the use of a JavaScript array:

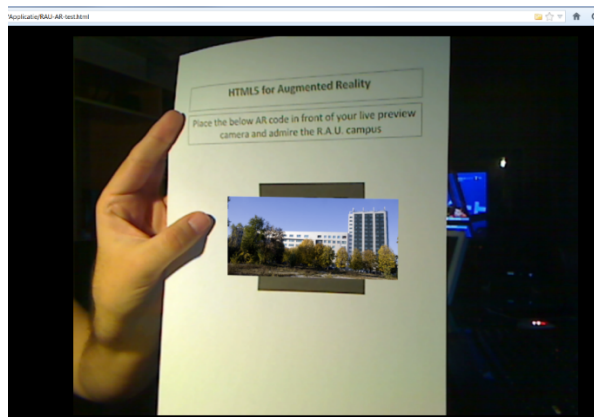


Figure 3. Another image imposed over the marker, at a later show of the AR code to the application inside the html webpage

3. CONCLUSION

The new alliance between the Augmented Reality environment and the HTML5 is the best bet at the moment for most web developers. This statement is fully arguable – only if you check out the sources for the above mentioned application we can see that we can build something with a great potential for a good user experience with quite a little experience in

the field and in a relatively short time frame. Right now, the HTML5 is the best mix of web technologies, (mostly) standardized by a set of common specifications, is relatively easy to learn but complex enough in order to be able to provide solutions for the most scenarios met during a web application development cycle.

I have started using HTML5 a couple of years ago and started to poke around the AR applications only about a year ago. From my own experience I can say that HTML5 can give you a very enthusiastic start and let you progress very quickly, even stepping over some previous mandatory requirements that a web developer needed to know just a few year back. But, I also have to mention that not everything is peachy in this AR-HTML5 alliance – as for any big potential alliance, downturns appear from time to time. The HTML5 is not yet fully supported, or even supported in the same manner on every important browser. I would personally recommend developing by the means of HTML5 approach in the case of a trendy project, with targeted and predictable customers that are tech-enthusiasts. On the other hand, I would warn about the use of this approach for enterprise-grade applications. The implementation of the HTML5 elements is not yet graded as enterprise-level and the cost involved in developing an enterprise-grade web application based on such an immature framework could be quite high. As a direct result, I foresee that at least for the next couple of years the AR based on HTML5 will continue to gain popularity but will not replace any other technology already available and mature. I do not think that a lot of retailers, for example, will scrape off the price labels from their shelves and rely only on a AR code for convincing a client to buy a product. The online retailing, usually the most modern and tech savvy environment, is not very likely to drop their so hard-worked security implementations [3] for the fast integration of AR only shopping.

More important, some other support technologies used for AR development are still not fully supported by all our computer-like devices. Today, the statistics show that around 1.4 billion mobile devices are HTML5 enabled and support the WebRTC standard while only this year alone the mobile devices sold on the global market numbered 1.86 billion⁶.

All in all, after carefully considering all the facts and factors involved, I am sure that any dedicated web applications developer should envisage the field of AR implementations and, in order to make the AR work they can use the HTML5 as one of the best tools available. A next step that I see coming in just a couple of years is related to hardware products. I am quite sure that the Google Glass project will be a success from the start – the first iterations of the device will bring the wearable devices to the mainstream consumer, will make him comfortable with the idea and the interfaces involved. The next step will be to develop real sci-fi applications that would need a considerable amount of computing power that, for now, cannot be built into a smaller and smaller wearable device. But there already is a solution for this – we just have to link to Google Glasses, or the devices of this kind, to a powerful processor that stays idle for a long time while being able to produce amazing numbers as regarding the computing power. We can link the computing needs to a specific chipset, inside a PC or a dedicated one, based on the Nvidia CUDA architecture [4]. This

⁶ Source: <http://www.ccsinsight.com/press/company-news/1655>

CUDA enabled devices can produce an enormous amount of power that could probably benefit a lot to the advanced AR applications of tomorrow.

4. REFERENCES & BIBLIOGRAPHY

1. [Tabusca, Silvia Maria - "The Internet Access as a Fundamental Right"; published in "Journal of Information Systems and Operations Management", Vol.4. No.2 / 2010, pp 206-212, ISSN 1843-4711.](#)
2. Heikkinen, Ilmari – “Writing Augmented Reality Applications using JSARToolKit”; available online at http://www.html5rocks.com/en/tutorials/webgl/jsartoolkit_webrtc/; last accessed on November 15th, 2013
3. [Pirjan, Alexandru – “Electronic commerce security in the context of the means of payment dematerialization”, published in “Journal of Information Systems & Operations Management”, Vol.4 No.1 / 2010, pp. 184-194, ISSN 1843-4711](#)
4. [Pirjan, Alexandru; Petroșanu, Dana-Mihaela_ – “Solutiouns for Optimizing the Radix Sort Algorithm Function using the Compute Unified Device Architecture”, published in “Journal of Information Systems & Operations Management”, Vol.6 No.2 / 2012, ISSN 1843-4711](#)