Considerations of Generic Framework for AR on the Web

A Position Paper for W3C's AR on the Web Workshop

Jonghong Jeon, Sunghan Kim, Seungyun Lee

Standardization Research Center, ETRI,

161, Kajong-Dong, YuSong-Gu, Daejeon, Republic of Korea

{ hollobit, sh-kim, syl@etri.re.kr }

Background & Motivation

Augmented reality (AR) is a term for a live direct or indirect view of a physical real-world environment whose elements are merged with (or augmented by) computer-generated virtual imagery – creating a mixed reality. In the case of Augmented Reality, the augmentation is conventionally in real-time and in meaningful context with environmental elements. With the help of advanced Computer Vision technology the information about the surrounding real world of the user becomes interactive and digitally usable.

Augmented reality is getting a lot of attention these days - particularly the use of AR with smartphones. The main hardware components for augmented reality are: display, tracking, input devices, and computer. Combination of powerful CPU, camera, accelerometers, GPS and solid state compass are often present in modern smartphones, which make them prospective platforms for augmented reality [1].

GPS technology combined with mobile camera can do wonders in bringing augmented reality experience. There are many applications which can interact with the video being streamed through mobile camera to the database present in the Internet about the location of the mobile phone.

Type of Augmented Reality Application

With growth of smartphone-based mobile AR, there are internal tensions within the traditional AR community over whether current mobile AR qualifies as true augmented reality.

Traditional AR community, using Military grade HMD (Head Mounted Display), and industrial AR applications often understand what they are "seeing." They can analyze a stream of visual data to identify that a chassis on an assembly line is an actual chassis, or that a missile streaking toward a fighter plane is indeed a missile. Despite the ergonomic shortcomings with this, there has been success in certain applications areas, such as industrial assembly, surgery training or games. However, the truly radical use of AR based on mobile technology that allows "Anywhere Augmentation" away from the desktop has not yet been realized [2, 5].

Conversely, the current of mobile AR browsers for smartphones use location technologies in order to display information on a video stream. The smartphone doesn't "see" what is in front of it; it just knows its geographic position. It uses geolocation data to place digital content, 2-D and 3-D objects along with links to other information and services into the user's field-of-vision.

Currently, there are four main categories of mobile AR applications: navigation, location overlays, geo-information services, and gaming [2].

1) **Navigation** - Augmented Reality can be seamlessly applied to existing navigation and mapping tools like Google Maps and MapQuest. The idea is to make these services increasingly user-friendly by displaying arrows and basic navigation information on the screen of the mobile device while the user is en route to the desired destination.

2) **Location Overlay** – The basic principle of this type of AR application is to overlay a specific location or site with information or images. The user points their phone at a target location such as a skyline, building, mountain or other geographical feature and supplemental content will be displayed on the screen

3) **Geo-Information Services** – The goal of geo-information-focused AR is to take the idea of location overlays to a new level. By utilizing user-generated content and web sites like Wikipedia, these apps will display vast amounts of information to the users, like digital notes, the best deals in the surrounding area, nearby landmarks or local business reviews.

4) **Gaming** – Extensive R&D is being invested into games utilizing AR. In their basic form, AR-based games display a game over a video stream of the real world. But instead of being confined to the screen of the mobile device, AR gaming places the user into the center of the action by expanding the game field to include the whole world.

Interoperability issue of Augmented Reality

There are still a number of technical and practical issues to be overcome to create a seamless AR experience for the user. The mobile industry should address the core issues of processing speed, memory and also the battery life issues created by the camera being in preview mode. Developers should ensure that their applications have the ability to compensate for the physical issues; loss of GPS or network and latent inaccuracies.

If AR applications are going to be deployed on a massive scale, there are several key areas of technology that are needed [5]:

(1) A low cost platform that combines AR display, tracking and processing
(2) Mobility as key requirement for the platform to realize AR in a global space
(3) Backend infrastructure for distributing of AR content and applications
(4) Easy to use authoring tools for creating AR content
(5) Large scale AR tracking solutions which work in real time

Among issues, Interoperability issue is one of major challenges faced by the small but fast-growing industry of Augmented Reality. For example, people cannot share the augmented information between different AR browsers and AR service environment.

The Lack of interoperability in AR systems and services, such as 3D modeling markup, AirTag and POI description method, has been identified as a major obstacle to the widespread take-up of AR applications [4, 6].

There are twelve interoperability issues between AR Provider and AR User Agent, such as figure1.

1. Composite AR Contents Markup & Format
2. AR contents Transport/Interaction method
3. Representing 3D Interactive AR/MR Contents
4. Event Scripting
5. Local Caching method
6. Additional functionability of AR User Agent
7. Device Capability Access APIs
8. AR Data Mashup method
9. AR Data Format – POI(images, text, 3D models, URLs), Person..
10. AR Data Service API
11. Open Marker Database
12. Security & Privacy
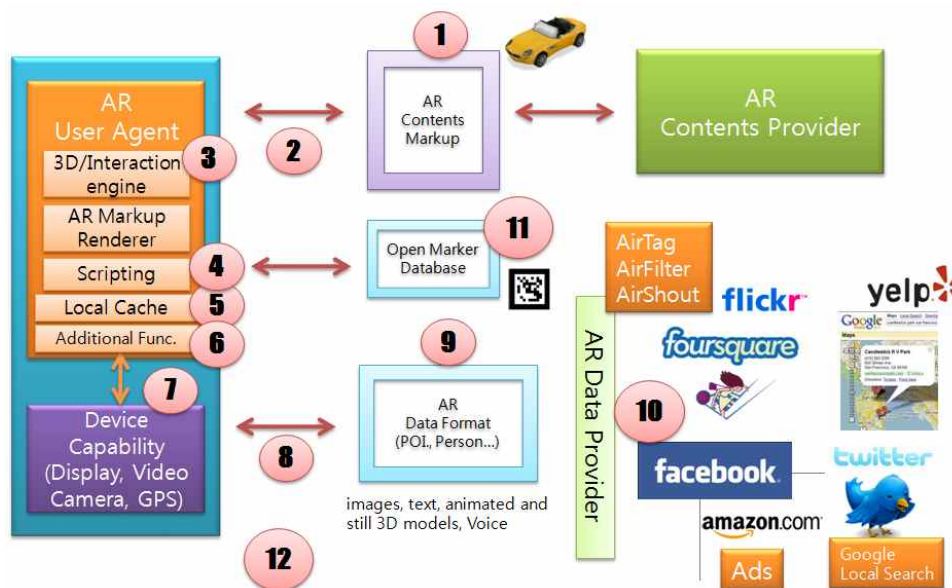
(Figure 1: interoperability issues between AR Provider and AR User Agent)

Interoperability Decreases Costs and Improves Information Access

There are a number of benefits of AR interoperability, including:

- **Improved user access**: Getting users the right information they need quickly. With interoperable AR systems, users do not need to launch multiple AR browsers to access AR content throughout the day. They can use a single AR User Agent on own device and access to interoperable AR contents, as like World Wide Web.

- **Cost Reduction**: Reductions in development costs for AR content. AR Standards can be reducing the expensive burden of duplicated content authoring that should support to multiple AR services and environment.

What Are the Main Benefits to Using the Web Technology for Augmented Reality?

Network-based Augmented Reality refers to technology (mobile devices, in-windshield displays, etc.) that can overlay information from the Web on top of objects in the real world. It is very similar with the World Wide Web.

But In the past year, a number of mobile AR applications have been developed native application. The main reason for this native app style development has been the integration of device capabilities and the ability to access this functionality such as the camera. One of the main weaknesses of native AR applications is the inability to transfer them across over multiple devices and multiple platforms [3].

One possible solution to solving the fragmentation problem of native application is to use the components of the general Web Users Agents (e.g. web browsers) to provide cross-platform capabilities, such as figure 2. In this always-connected age, most of the useful additional information for informative-AR application you need to display is most like going to reside on the web.
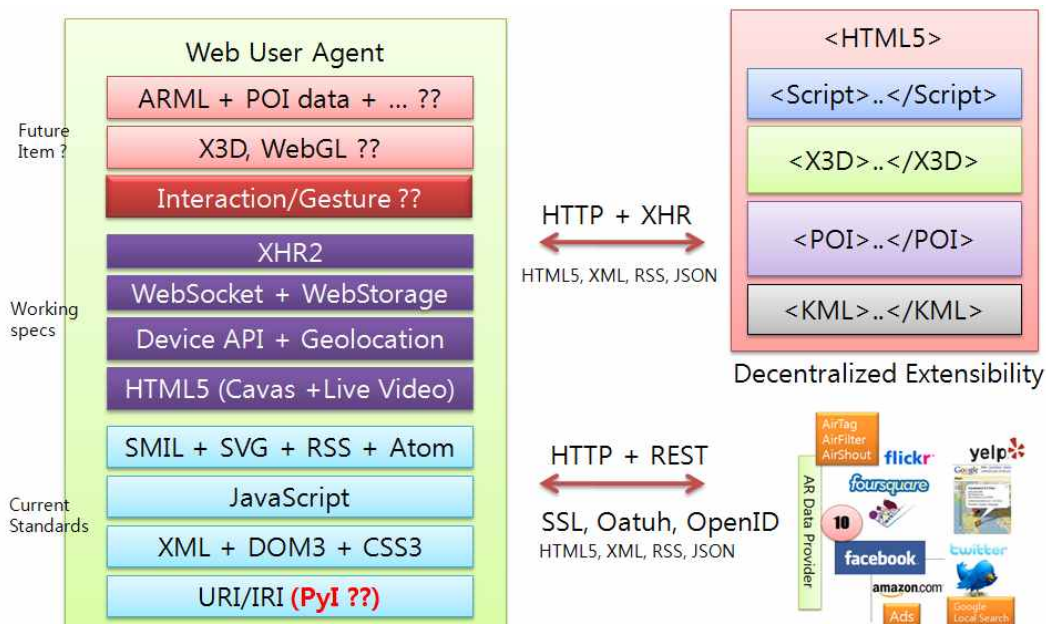
(Figure 2: Generic Framework of the AR on the Web)

General requirements for AR on the Web

Before progressing W3C standardization activity, argument reality services need to consider in several perspectives, e.g. AR contents visualization, architecture model, privacy, context awareness, as well device resource capability. Recently, many kinds of AR mobile services are emerging for mobile users. Until now, many research and projects are basically based on sensor technologies, GPS information, camera and display devices. Besides, user's context awareness information is mixing with basic AR mechanism and it makes possible to support more dynamic and user centric services.

In market, there are already popular AR services, and some services are globally providing with mobile phone.

For examples, bionic eye service for searching near coffee shops, car finder service for easy finding parking location, red laser service for comparing prices with barcode, etc. So, it is not difficult expecting more diverse services to appear in near days. Therefore, after reviewing the service types, features and functionalities for AR standardization, and then some general requirements can be suggested for common AR services. Though now can not define the general requirements in detail, W3C seems necessary to consider further for the under standardization issues [7].

- AR content visualization issue
  - ✓ Consideration for 2D/3D object visualization scripts(methods)
  - ✓ Live Video Streaming in HTML5
  - ✓ Registration between Live(video) object and Virtual object.
  - ✓ Interaction Mechanism with Virtual object
  - ✓ Description format of Augmented Information Mark (e.g. AirTag, POI)
  - ✓ Harmonization with traditional standards with OpenGL, X3D, COLLADA, JavaScript, PNG, JPEG, etc.
- General architecture model
  - ✓ Common framework to support transport/interaction methods
  - ✓ Protocols, data formats and parameters to query image, marker and location information and to get the results (Query and result API)
  - ✓ Local caching method
- Privacy and Context awareness
  - ✓ Protections mechanism for user's personal private information
- Device capability access
  - ✓ Access to Camera, Geolocation, Gesture API
  - ✓ Harmonization with DAP WG, HTML 5 WG, Web application WG in W3C
- Data base issue: data structure for Object data and POI
  - ✓ Key value and format for matching each object images/markers to another resources (eg. ISBN, UPC/EAN)
  - ✓ Data structure for basic and/or detailed information related to each object
  - ✓ Key value and format for matching each POI to another resources
    (The lat/long pair is NOT an unique key!)
  - ✓ Data structure for basic and/or detailed information related to each POI
  - ✓ Data structure for the POI list for a specific local query with lat/long/viewpoint and/or Street-view image query (eg. KML/ARML)

References

[1] Augmented Reality, Wikipedia, http://en.wikipedia.org/wiki/Augmented_reality
[2] GigaOM Report, "The World Is the Desktop: Mobile Augmented Reality", December 2009
[3] David Rogers, "BONDI Augmented Reality", Mobile AR Summit, MWC 2010
[4] Martin Lechner, Markus Tripp, "ARML - An Augmented Reality Standard", Mobile AR Summit, MWC 2010
[5] Dieter Schmalstieg, Tobias Langlotz, "Augmented Reality 2.0", Innovation Workshop Proceedings , ISMAR 2009
[6] 1st AR/MR Standardization Workshop, http://www.w3c.or.kr/~hollobit/ARKR/201004-workshop/
[7] Junghee Ryu, "AR standard requirements", 1st AR/MR Standardization Workshop, http://www.w3c.or.kr/~hollobit/ARKR/201004-workshop/pdf/arsw1-5.pdf