

Title: Beyond the keyhole

**Positional paper for the W3C Workshop: Augmented Reality on the Web
June 15 - 16, 2010 Barcelona**

Author: Klas Hermodsson
(firstname.lastname@sonyericsson.com)

Introduction

We think that today's mobile augmented reality (AR) experience falls short of fulfilling the potential of AR. Watching the world from a keyhole is limiting in so many ways:

- **Limited view.** Instead of augmenting the user's world, the user looks at an augmented reality through a keyhole.
- **Awkward interaction.** Similarly to how most people would feel awkward when standing in a public spot and holding up a camera in front of them for extended periods of time, an AR user should not need to hold a device in front of them. It is both socially awkward and physically tiring.
- **Relying on a camera sensor.** When the display is showing an augmented camera view then the view of the world is degraded to the quality and speed of the camera sensor. A camera sensor drains battery power and is inferior to the human eye for sensing the world around us.
- **Limited use.** The user must actively initiate the use of the AR application and point the device in the desired direction for there to be any augmented information to be available. This usage method results in use during short time periods and only when the user has decided that they would like to know more about something.

Full peripheral view AR (e.g. glasses, contact lenses) allows for better interaction and the possibility to truly assist an end user with guidance to the real or virtual world. This positional paper is written from the point of view that the full view AR device experience is the key use case to design for.

Moving beyond the keyhole usage of AR means that typical usage is more likely to be during longer periods of time and while the user is performing a number of different tasks and daily chores. The current opt-in stream model of most AR browsers and applications is not suitable for continuous use. User context is ever shifting and to be stuck in the "nearby restaurants" stream until you actively change to another stream is far from the true potential of AR.

Example: Think of a visit to a city for the first time. Bringing a guide book is great but to have a close friend that knows both you and the city is priceless. The keyhole opt-in stream AR model is like the guide book. It works if you are in a certain spot and you want to find out if there are any great jazz clubs nearby. In contrast your personal guide will walk with you around town and say "by the way, I know you love jazz. Over there on the right is the best jazz club by far, why don't we go have a look?". With the guide book it is a conscious decision that makes you look for a jazz club while the personal guide is pointing out things that you may be interested in even if you yourself did not ask for it.

Relevance, relevance, relevance

One of the primary challenges of good AR user experience is to provide relevant information or entry points to information. Relevance is completely dependent on the user and her context. We believe that in order to provide relevance we must have the full wealth of information available and then apply a user contextual filter on this data. Filtering by user context is one part of the AR experience. Another is answering questions like "What is that building that I am looking at right now?". Both should be covered to create good AR.

Example continued: The difference between applying user context or not is like having a personal close friend as guide versus having a guide that does not know you. Your friend will not tell you every single piece of information about your surroundings but will try hard to tell you things that he or she knows you have an interest in. A good AR experience should be like having an old trusted friend with you.

Web of data

There is already a wealth of data available. The web as we know it today is vast. However, in order to apply a user context, in a non manual way, the web must have data that is classified and understandable to machines. The web must be a web of data and not just a web of documents. This is what the semantic web is all about. With a web browser on a desktop computer, web of data is nice to have but for AR we would say that it is essential. Without some kind of classification and attributes to what is on the web we can never apply a proper user context.

With AR as another view on the existing web the distributed nature of the web remains. This is as it should be but it will present a substantial challenge. In order to break out from the grip of proprietary opt-in AR streams effective aggregation must be performed.

Filtering

For web data to be visible in AR space we need indexed attributes so that we can retrieve objects from one or a combination of attributes values. We believe attributes of special interest to AR includes:

- **Geo-location** (including altitude). A must for retrieving items close to the location which the user wants to view.
- **ID** (e.g. barcode, ISBN, visual fingerprint). Not all objects will be anchored to a physical location. This could cover lookup through object recognition. Example: in the bookshop you look at the book cover of an interesting title and the AR browser retrieves information on it by using the cover art as the lookup key.
- **Timestamp** (creation, content related, etc). Some information, e.g. news articles or friends' status updates, have limited value unless it is recently published. This will allow time based filtering.
- **Class**. Is this a news article, chat message, book or a building? Classification and semantic web is tightly coupled and classification is a research topic in its own right.
- **Author/Originator**. Enables filtering on your circle of friends or information from certain companies. It is all about who you trust.
- **Service**. Is this a service A chat message or a status update from service B? Filtering on service could be useful.

Unfiltered, most locations will result in unmanageable amounts of possible items. Strict filtering should be employed to provide relevance.

Possible user independent aspects to filter items include:

- **Geographical distance**. We do not want all the items in the world. Distance filtering will cut down the number of items greatly.
- **Line of sight/occlusion**. How far can the user see? Displaying items for discovery that are two blocks away and out of sight may not be useful. If we are indoors maybe we just want items related to what is in the room, floor or building.
- **Available screen real estate**. Display size and resolution will impact how much information can be visible without impairing user vision. Not all information will be presented visually of course so this may or not be valid for the way that the information is made available to the user.

Possible user dependent aspects to filter items include:

- **User attention focus**. Where is the user looking?
- **AR volume control**. A user control for how busy and intrusive a user allows the AR experience to be.
- **Own and circle of friends ranking**. "What café does my friends like? I don't really care if the majority of people like it or not."
- **User interests, mood, agenda etc**. The user context.

The user context could be applied in the mobile AR device but with multi device users and the limited scalability of such a solution we believe that user context should be stored and at least partially applied in the cloud.

User participation

We should not neglect the power of user participation in the world of AR. Users creating, reviewing and ranking must be seen as natural enhancement to published content. It could even be the way for most content to be created or classified. Considering how important the user context is for the AR case, the effect of user participation is naturally noticeable, creating a big incentive for user engagement.

Open questions

This positional paper points out some open questions regarding AR on the web:

- How can aggregation be done to enable the true AR mashup experience and alleviate the burden on the mobile AR device without compromising the distributed nature of the web?
- How to handle the user context for maximum scalability and performance while ensuring that user privacy is respected?
- Should the assumed importance of user content creation and ranking affect the AR infra structure or should we leave it up to the social services to wrap these social aspects around the content out there?