Current Status of Standards for Augmented Reality

Christine Perey, Timo Engelke, and Carl Reed

Abstract This chapter discusses the current state, issues, and direction of the development and use of international standards for use in Augmented Reality (AR) applications and services. More specifically, the paper focuses on AR and mobile devices. Enterprise AR applications are not discussed in this chapter. There are many existing international standards that can be used in AR applications but there may not be defined best practices or profiles of those standards that effectively meet AR development requirements. This chapter provides information on a number of standards that can be used for AR applications but may need further international agreements on best practice use.

1 Introduction

Standards frequently provide a platform for development; they ease smooth operation of an ecosystem in which different segments contribute to and benefit from the success of the whole, and hopefully provide for a robust, economically-viable, value chain. One of the consequences of widespread adoption of standards is a baseline of interoperability between manufacturers and content publishers. Another is the ease of development of client applications.

In most markets, standards emerge during or following the establishment of an ecosystem, once a sufficient number of organizations see market and business value in interoperating with the solutions or services of others.

Publishers of content that support AR applications are motivated to make their content available when there is an assortment of devices that support the content for different use cases and this translates into the maximum audience size. Standards

C. Perey (✉)
Spime Wrangler, PEREY Research & Consulting
e-mail: cperey@perey.com

L. Alem and W. Huang (eds.), *Recent Trends of Mobile Collaborative Augmented Reality Systems*, DOI 10.1007/978-1-4419-9845-3_2,

enable such device and use case independence, thereby reducing implementation costs and mitigating investment risks.

As of early 2011, the mobile AR solutions available to users and developers are based on a mixture of proprietary and open standards protocols and content encodings, without interoperable – or standards based - content, platforms or viewing applications. It is a field of technology silos and, consequently fragmented markets.

For mobile collaborative AR, the needs for standards are compounded by the fact that the content of shared interest must travel over a communications "bridge" which is, itself, established between end points between and through servers, client devices and across networks. The more interoperable the components of the end-to-end system are, the less the need for the participants in a collaborative session to use technologies provided by the same manufacturer. More interoperability translates directly into more enabled people, hence more potential collaborators, and more service and application providers.

2 Guiding principles of an open AR industry

Open AR, or interoperable systems for viewing content in real time in context, is a design goal for the evolution of the AR market. Currently, there are numerous standards that can be used in the development and deployment of open AR applications and services. However, there are still interoperability gaps in the AR value chain. Further, work needs to be done to determine best practices for using existing international standards. In some cases in which there are interoperability gaps, new standards will need to be defined, documented, and tested. Developing new standards and pushing them through the development process required in a standards development organization may not be appropriate for the needs of the AR community. In this case, perhaps profiles of existing standards would be more appropriate. Further, the development of an over-arching framework of standards required for AR may be beyond the resources of any single body. And, as AR requires the convergence of so many technologies, there are numerous interoperability challenges. As such, there will not be one "global" AR standard. Instead, there will be a suite of standards for use in AR applications.

Many technology participants in the AR ecosystem desire to leverage existing standards that solve different interoperability issues. For example, standards which permit an application to learn the locations of users, how to display objects on the users' screen, how to time stamp every frame of a video, how to use the users' inputs for managing behaviors, and which are proven and optimized… to be extended to address new or related issues which AR raises.

One of the strongest motivators for a cross-standard, multi-consortium and open discussion about standards and AR is time-to-market. Re-purposing existing content and applications is critical. The use of existing standards or profiles of these standards is driven by the need to avoid making mistakes and also use of currently deployed and proven (and emerging) technologies to solve/address urgent issues for AR publishers, developers and users.

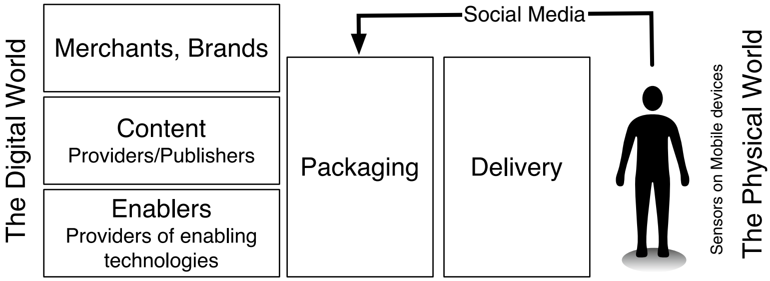

Fig. 1 Ecosystem of mobile AR Segments

However, the time-to-market argument is only valid if one assumes that there is a motivation/agreement on the part of most or all members of the ecosystem that having open AR—the opposite of technology "silos"—is a good thing. Based on the participation of academic and institutional researchers, companies of all sizes and industry consortia representing different technology groups, there is agreement across many parts of the AR ecosystem regarding the need for standards[1].

2.1 *Suggested model of a General AR ecosystem*

The AR ecosystem is composed of at least six interlocking and interdependent groups of technologies. Figure 1 shows how these interlocking and interdependent groups bridge the space between the digital and physical worlds in a block diagram.

Beginning on the far right side of the figure, there is the "client" in the networked end-to-end system. The user holds or wears the client, a device that provides (1) an interface for the user to interact with one or more services or applications and the information in the digital world, and (2) integrates an array of real-time sensors. The sensors in the users' devices detect conditions in the users' environments as well as in some cases the users' inputs. There may also be sensors (e.g., cameras, pressure sensors, microphones) in the environment to which the applications could provide access. The client device is also the output for the user, permitting visualization or other forms of augmentation such as sounds or haptic feedback.

Manufacturers of components and finished AR-capable devices (e.g., Smartphones) occupy both the client segment of the ecosystem as well as, in some cases, the "technology enablers" segment to the far left of the figure. At the right side, the client devices are frequently tightly connected to the networks.

[1]This conclusion is based on the results of discussions at two recent multi-participant AR workshops.

Network providers, providers of application stores and other sources of content (e.g., government and commercial portals, social networks, spatial data infrastructures, and geo-location service providers) provide the "discovery and delivery" channel by which the user receives the AR experience. This segment, like the device segment described above, overlaps with other segments and companies may occupy this as well as the role of device manufacturer.

Packaging companies are those that provide tools and services permitting viewing of any published and accessible content. In this segment we can imagine sub-segments such as the AR SDK and toolkit providers, the Web-hosted content platforms and the developers of content that provide professional services to agencies, brands and merchants.

Packaging companies provide their technologies and services to organizations with content that is suitable for context-driven visualization. For the case of collaborative AR, this is probably not an important segment since, in effect, users themselves are the creators of content[2].

For some purposes, packaging companies rely on the providers of enabling technologies, the segment represented in the lower left corner of the figure. Like the packaging segment, there are sub-segments of enabling technologies (e.g., semiconductors, sensor providers, algorithms, etc). This segment is rich with existing standards that can and are already being assimilated by the companies in the packaging segment of the ecosystem.

Content providers include a range of public, proprietary (commercial), and user-provided data. Traditionally, proprietary content was the primary source of content for use in AR applications. More recently, more and more content is being provided by government agencies (e.g. traffic data or base map information) and volunteered sources (e.g. Open Street Map). An excellent example of this evolution from proprietary to a mixed content platform is the map data used in AR applications.

The traditional AR content providers, brands and merchants who seek to provide their digital information to users of AR-enabled devices, are reluctant to enter the AR ecosystem until they feel that the technologies are stable and robust. The adoption of standards for content encoding and access of content by platforms and "packaging segment" providers is a clear indicator of a certain market maturity for which content providers are waiting. The use of standard interfaces and encodings also allows application providers to access content from many more sources, including proprietary, user-provided, and government sources.

Furthermore, content providers and end users of AR applications will benefit when AR content standards are available to express the provenance and quality of the source content in a consistent fashion.

[2] Also known as "user provided" content.

3 AR Requirements and Use Cases

For the proper development of standards for AR, there needs to be a very clear understanding of AR requirements and use cases. Different domains have different AR requirements. For example, mass market mobile AR tourist applications requirements and related use cases may be different from those required by first responders in an emergency scenario. Such analysis will also permit identification of the most common requirements. By specifying use cases and requirements, standards organizations will have the information necessary to determine which standards can best be used in given situations or workflows. Given that different AR ecosystem segments have different requirements for standards, different standards bodies and industry consortia have been working on various aspects of the AR standards stack. Therefore, stronger collaboration between the various standards bodies is required. By conducting face-to-face open meetings of interested parties, such as AR DevCamps and the International AR Standards Meetings, people from vastly different backgrounds are convening to exchange (share) information about what they have seen succeed in their fields and how these may be applied to the challenges facing interoperable and open AR. Having a common set of use cases and related requirements provides the "lingua-franca" for collaboration and discussion.

This collaboration between Standards Development Organizations (SDOs) and their related expert communities is crucial at this juncture in the growth of AR. If we can benefit from the experience of those who wrote, who have implemented and who have optimized a variety of today's most popular standards already in use for AR or AR-like applications (OpenGL ES, JSON, HTML5, KML, GML, CityGML, X3D, etc.), the goal of interoperable AR will be more quickly achieved and may avoid costly errors.

Discussions on the topic of standards to date indicate that the development of standards specifically for AR applications is necessary in only a small number of cases. Instead, re-purposing (profiles) and better understanding of existing standards from such organizations as the Khronos Group (Khronos, 2011), the Web3D Consortium, the World Wide Web Consortium (W3C), the Open Geospatial Consortium (OGC), the 3GPP, the Open Mobile Alliance (OMA) and others is the way to proceed with the greatest impact and assurance that we will have large markets and stable systems in the future.

3.1 AR Content publisher requirements

The content sources from which the future AR content will be produced and deployed are extremely varied. They range from multi-national information, news and media conglomerates, device manufacturers, to national, state, and local government organizations, user-provided content, to individual content developers who wish to share their personal trivia or experiences. Clearly, publisher sub-segments will have needs for their specialized markets or use cases.

As a broad category, the content publisher's needs are to:

- Reach the maximum potential audience with the same content,
- Provide content in formats which are suited to special use cases (also known as re-purposing),
- Provide accurate, up-to-date content,
- Control access to and limit uncontrolled proliferation (pirating) of content.

For content publishers, a simple, lightweight markup language that is easily integrated with existing content management systems and offers a large community of developers for customization, is highly desirable.

In collaborative AR, the case can be made that the users themselves are the content that is being enhanced. In this view, the users will rely on real time algorithms that convert gestures, facial expressions, and spoken and written language into objects or content, which is viewed by others at a distance.

Real time representation of 2D and 3D spaces and objects at a distance will rely on projection systems of many types and for remote commands to appear in the view of local users. These remote commands could leverage the existing work of the multimedia telecommunications manufacturers and videoconferencing systems adhering to the ITU H.3XX standard protocols.

3.2 Packaging segment requirements

This is the segment of the AR ecosystem in which the proprietary technology silos are most evident at the time of this study and where control of the content development platforms is highly competitive. There are the needs for differentiation of the providers of tools and platforms that are substantially different than those of the professional service providers who use the tools to gain their livelihoods.

Tools and platform providers seek to be able to:

- Access and process content from multiple distributed repositories and sensor networks. This may include repackaging for efficiency. However, for certain content types, such as maps or location content, the ability to access the content closest to source allows the end user to use the latest, best quality content;
- Offer their tools and platforms to a large (preferably existing) community of developers who develop commercial solutions for customers;
- Integrate and fuse real time sensor observations into the AR application;
- Quickly develop and bring to market new, innovative features that make their system more desirable than a competitor's or a free solution.

Professional developers of content (the service providers who utilize the SDKs and platforms for publishing) seek to be able to:

- Repurpose existing tools and content (managing costs as well as learning curves) to just make an "AR version" of their work

- Provide end users rich experiences that leverage the capabilities of an AR platform but at the same time have features tying them to the existing platforms for social networking, communications, navigation, content administration and billing.

3.3 AR system and content users

This is the most diverse segment in the AR stack in the sense that users include all people, related services, and organizations in all future scenarios. It is natural that the users of AR systems and content want to have experiences leveraging the latest technologies and the best, most up-to-date content without losing any of the benefits to which they have grown accustomed.

In the case of collaborative mobile AR users, they seek to:

- Connect with peers or subject matter experts anywhere in the world over broadband IP networks,
- Show and manipulate physical and local as well as virtual objects as they would if the collaborator were in the same room, and
- Perform tasks and achieve objectives that are not possible when collaborators are in the same room.

4 Approaches to the AR Standards challenge

To meet the needs of developers, content publishers, platform and tool providers and users of the AR ecosystem, the experts in hardware accelerated graphics, imaging and compute, cloud computing and Web services, digital data formats, navigation, sensors and geospatial content and services management and hardware and software devices must collaborate.

4.1 Basic tools of the standards trade

Standards that are or will be useful to the AR ecosystem segments will leverage know-how that is gained through both experimentation, and creation of concrete open source, commercial and pre-commercial implementations. In most standards activities, the process of developing a recommendation for standardization begins with development of core requirements and use cases. This work is then followed by development of a vocabulary (terms and definitions), information models, abstract architectures and agreement on the principle objectives.

An AR standards gap analysis must be performed. The results of the gap analysis combined with known requirements will reveal where the community should concentrate future standardization efforts. Finally, any new standards work designed to fill the gaps can begin in existing standards organizations to support AR experiences.

4.2 Standards gap analysis

A gap analysis begins with detailed examinations of available standards and to determine which standards are close to and which are distant from meeting the AR ecosystem requirements. The gap analysis process began during the International AR Standards Meeting in Seoul, October 11-12, 2010.

The gap analysis exercise divided the scope of the problem into two large spaces: those related to content and software, and those that are most relevant to hardware and networks.

Existing standards search results were grouped according to whether the standard addresses a content/software service related issue or a network and hardware issue. In some cases, there is overlap.

5 Content-related standards

5.1 Declarative and imperative approaches to Content Encoding

First, it is important to the success of the gap analysis to clarify the differences between the declarative and the imperative approaches of standards.

The imperative approach of description usually defines how something is to be computed, like code. It features storage, variables, states, instructions and flow control. Usually it benefits from a high potential of possibilities and user driven variations. In use, imperative code can be designed in any manner, as long as they conform to the common rules of the interpreting background system. A typical example is JavaScript (JS), a highly popular implementation of the ECMAScript (ECMA-262) language standard, which is part of every Web browser on mobile and desktop systems today.

Declarative approaches are more restrictive and their design usually follows a strict behavior scheme and structure. They consist of implicit operational semantics that are transparent in their references. They describe what is to be computed. Declarative approaches usually do not deliver states and, thus, dynamic systems are more difficult to achieve using declarative approaches. On the other hand, they tend to be more transparent and easy to use and generate. A common declarative language in use today is the W3C XML standard which defines a hierarchical presentation of elements and attributes. Another coding form for declarative data is the JavaScript Object Notation (JSON), which benefits from being a lightweight

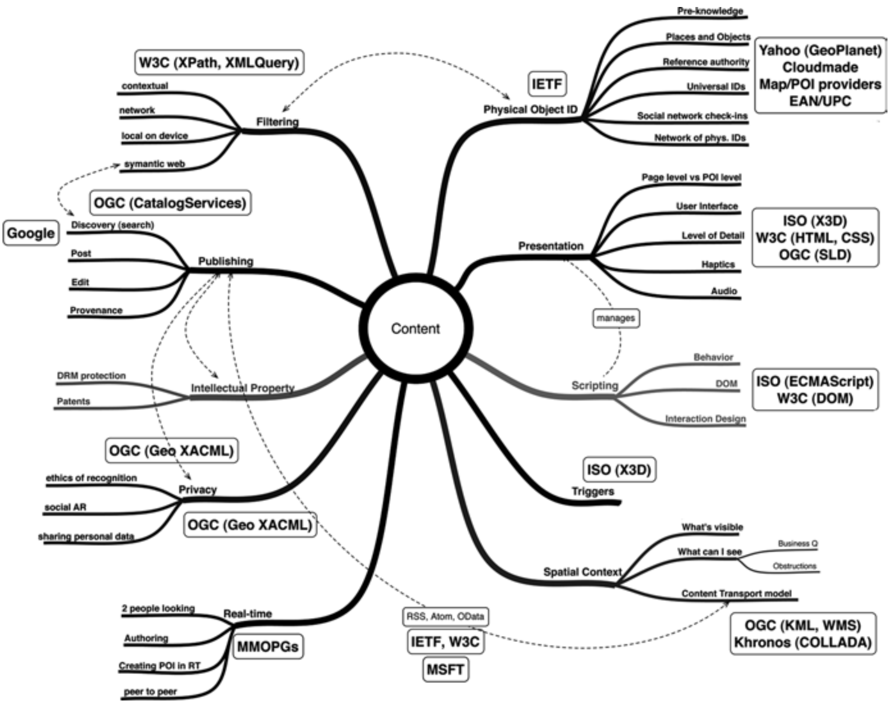

Fig. 2 Standards Landscape of Impact to Mobile AR

and easy-to-port data interchange format. It builds upon the ECMAScript specification.

Both approaches,or a combination of these, are likely to be used for creation of AR content encoding and payload standards.

5.2 Existing Standards

There are many content or payload encoding related standards that could be used for AR applications. The diagram below suggests an initial inventory of such standards and their possible relationships.

The table below shows Geo Information System (GIS)-based standards and other standards used within the system along with the Web addresses where further definitions can be found. The organizations mentioned are potential providers of experience and knowledge in specialized fields. Today's mapping software is usually based on these standards; consequently, AR services and applications that rely on the user's location also leverage these standards. There are also many standards that define the position and interactivity of virtual objects in a user's visual space.

Table 1 Existing Standards for Use in Geo-location-based Mobile AR

Standards	Organization	url
Geography Markup Language	OGC and ISO	http://www.opengeospatial.org/standards/gml
CityGML	OGC	http://www.opengeospatial.org/standards/citygml
KML	OGC	http://www.opengeospatial.org/standards/kml
SensorML	OGC	http://www.opengeospatial.org/standards/sensorml
Sensor Observation Service	OGC	http://www.opengeospatial.org/standards/SOS
Web Map Service	OGC and ISO	http://www.opengeospatial.org/standards/wms
OpenGL	Khronos	http://www.opengl.org/
SVG	W3C	http://www.w3.org/Graphics/SVG/
Style Layer Descriptor	OGC	http://www.opengeospatial.org/standards/SLD
ECMAScript	ISO	http://www.ecmascript.org/
HTML	W3C	http://www.w3.org/html/
Atom	IETF	http://tools.ietf.org/html/rfc4287
X3D	Web3D/ISO	www.web3d.org
GeoRSS	Georss	www.georss.org
COLLADA	Khronos	www.collada.org

6 Mobile AR Standards Considerations

A lot of promising technologies have been developed and integrated in mobile device hardware. New sensor technologies allow delivering sensor data for mobile AR applications in a format that can be processed. The processing power in mobile devices, network and memory bandwidth supporting the latest mobile devices and applications have expanded exponentially in recent years. Software frameworks and platforms for mobile application development have also made huge advances, permitting developers to create new user experiences very quickly. This provides a huge potential for context- and location-aware AR applications or applications that are extended to take advantage of new capabilities.

Although these developments have accelerated the growth of the number and diversity of mobile AR applications, this growth has come at a cost. There is clearly a lack of standards for implementing mobile AR applications for users of multiple, different platforms and in different use scenarios.

In the next subsections we describe the use of standards in mobile AR sensing, processing and data presentation.

6.1 Mobile AR and Sensors

A sensor is an entity that provides information about an observed property as its output. A sensor uses a combination of physical, chemical or biological means in order to estimate the underlying observed property. An observed property is an identifier or description of the phenomenon for which the sensor observation result provides an estimate of its value. Satellites, cameras, seismic monitors, water temperature and flow monitors, accelerometers are all examples of sensors. Sensors may be in-situ, such as an air pollution monitoring station, or they may be dynamic, such as an unmanned aerial vehicle carrying a camera. The sensor observes a property at a specific point in time at a specific location, i.e. within a temporal and spatial context. Further, the location of the sensor might be different from the location of the observed property. This is the case for all remote-observing sensors, e.g. cameras, radar, etc.

From a mobile AR perspective, sensors may be onboard (in the device) or external to the device and accessed by the AR application as required, Regardless, all sensors have descriptions of the processes by which observations and measurements are generated, and other related metadata such as quality, time of last calibration, and time of measurement. The metadata, or characteristics, of the sensor are critical for developers and applications that require the use of sensor observations. The ability to have a standard description language for describing a sensor, its metadata, and processes will allow for greater flexibility and ease of implementation in terms of accessing and using sensor observations in AR applications.

Sensors behave differently on different and distinct device types and platforms. Due to differences in manufacturing tolerances or measurement processes, dynamic, or mobile, sensor observations may also be inconsistent even when observing the same phenomenon. Calculation of user location indoors is one example where wide variability may occur. Different location measurement technologies provide different levels of accuracy and quality. The problem is exacerbated by a variety of factors, such as interference from other devices, materials in the building, and so forth.

Approaches combining inaccurate geo-positioning data along with computer vision algorithms are promising for increasing accuracy of mobile AR, but require the definition of new models for recognition, sensor-fusion and reconstruction of the pose to be defined. Ideally, an abstraction layer which defines these different "sensor services" with a well-defined format and its sensor characteristics, would address existing performance limitations.

For some AR applications, real-time processing of vision-based data is crucial. In these cases, direct camera data is not appropriate for processing in a high level programming environment, and should be processed on a lower level. Since the processing is performed at the lower level, algorithms that create an abstracted sensor data layer for pose will be beneficial. In summary, sensor fusion and interpretation can happen on different levels of implementation[3].

Sensors with different processing needs can contribute to the final application outcome. In parallel, the higher level application logic may benefit from taking data from multiple sensors into account. Standards may provide direct access to sensor data or higher-level semantic abstractions of it.

[3] As an example of fusion requirements, consider the OGC "Fusion Standards Study Engineering Report". http://portal.opengeospatial.org/files/?artifact_id=36177

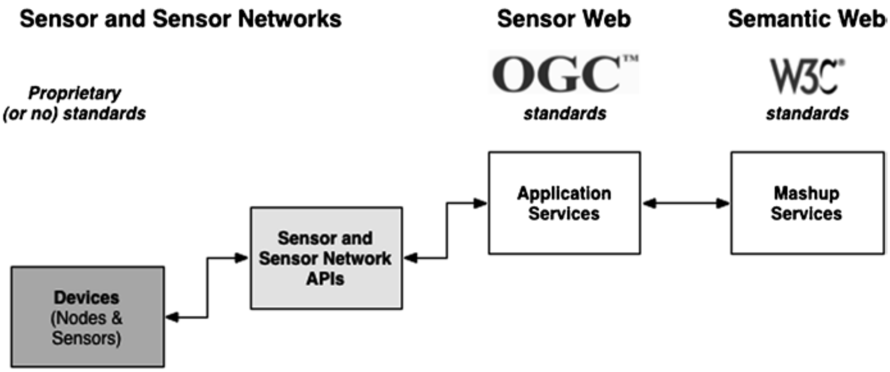

Fig. 3 The OGC Sensor Web Enablement Standards Landscape

A critical content source for many AR applications, independent of domain, will be near real-time observations obtained from in-situ and dynamic sensors. Examples of in-situ sensors are traffic, weather, fixed video, and stream gauges. Dynamic sensors include unmanned aerial vehicles, the mobile human, and satellites. Already, the vast majority of content used in AR applications is obtained via some sensor technology, such as LIDAR[4], that is subsequently processed and stored in a content management system. There are many other sources of sensor data that are (or will be) available on demand or by subscription. These sensor observations need to be fused into the AR environment in real time as well. As such, there is a need for standards that enable the description, discovery, access, and tasking of sensors within the collaborative AR environment.

The following is a simple diagram depicting one widely implemented sensor standards landscape.

A sensor network is a computer-accessible network of many, spatially-distributed devices using sensors to monitor conditions at different locations, such as temperature, sound, vibration, pressure, motion or pollutants. A Sensor Web refers to web-accessible sensor networks and archived sensor data that can be discovered and accessed using standard protocols and APIs.

There is a suite of standards that support Sensor Web Enablement (SWE) maintained by the OGC. SWE standards include:

1. Observations & Measurements Schema (O&M) – Standard models and XML Schema for encoding observations and measurements from a sensor, both archived and real-time.
2. Sensor Model Language (SensorML) – Standard models and XML Schema for describing sensors systems and processes; provides information needed for discovery of sensors, location of sensor observations, processing of low-level sensor observations, and listing of taskable properties.

[4]LIDAR is an acronym for LIght Detection And Ranging.

3. Transducer Markup Language (TransducerML or TML) – The conceptual model and XML Schema for describing transducers and supporting real-time streaming of data to and from sensor systems.
4. Sensor Observations Service (SOS) - Standard web service interface for requesting, filtering, and retrieving observations and sensor system information. This is the intermediary between a client and an observation repository or near real-time sensor channel.
5. Sensor Planning Service (SPS) – Standard web service interface for requesting user-driven acquisitions and observations. This is the intermediary between a client and a sensor collection management environment.

These standards could be used for low level descriptions of sensors and their fusion, and combined with visual processing for pose estimation and tracking, supply automatically-generated data for augmenting the users' immediate environment.

6.2 *Mobile AR processing standards*

Considerable standards work has previously been done in the domains of situational awareness, sensor fusion, and service chaining (workflows). This work and some of these standards can be applied to processes in an AR workflow. For example, the OGC Web Processing Service (WPS) provides rules for standardizing how inputs and outputs (requests and responses) for geospatial processing services, such as polygon overlay. The standard also defines how a client can request the execution of a process, and how the output from the process is handled. It defines an interface that facilitates the publishing of geospatial processes and clients' discovery of and binding to those processes. The data required by the WPS can be delivered across a network or they can be available at the server. The WPS can be used to "wrap" processing and modeling applications with a standard interface. WPS can also be used to enable the implementation of processing workflows.

In addition, the OpenGIS Tracking Service Interface Standard supports a very simple functionality allowing a collection of movable objects to be tracked as they move and change orientation. The standard addresses the absolute minimum in functionality in order to address the need for a simple, robust, and easy-to-implement open standard for geospatial tracking.

Other approaches for descriptions of vision-based tracking environments with its visual, camera constraints have been made first through Pustka et al., by introducing spatial relationship patterns for augmented reality environment descriptions.

There are several standards that could be applied to the presentation and visualization workflow stack for AR applications. There are service interfaces that an AR application can use to access content, such as a map for a specific area. Then there are lower level standards that enable standard mechanisms for rendering the content on the device.

Not all AR requires use of 3D. In some use cases and, especially on low processor devices unable to render 3D objects, 2D annotations are preferable. A very convenient declarative standard for the description of 2D annotation is W3C's

Scalable Vector Graphic (SVG) standard or even HTML. This is a large field in which many existing standards are suitable for AR use.

6.3 Mobile AR Acceleration and Presentation Standards

Augmented Reality is highly demanding in terms of computation and graphics performance. Enabling truly compelling AR on mobile devices requires efficient and innovative use of the advanced compute and graphics capabilities becoming available in today's smartphones.

Many mobile AR applications make direct and/or indirect use of hardware for acceleration of computationally complex tasks and, since the hardware available to the applications varies from device to device, standard Application Programming Interfaces (APIs) reduce the need for customization of software to specific hardware platforms.

The Khronos Group is an industry standards body that is dedicated to defining open APIs to enable software to access high-performance silicon for graphics, imaging and computation. A typical AR system with 3D graphics uses several Khronos standards and some that are under development. For example:

- OpenGL ES is a streamlined version of the widely respected desktop OpenGL open standard for 3D graphics. OpenGL ES is now being used to provide advanced graphics on almost every 3D-capable embedded and mobile device;
- OpenMAX provides advanced camera control, image and video processing and flexible video playback capabilities;
- OpenCL provides a framework for programming heterogeneous parallel CPU, GPU and DSP computing resources. Already available in desktop machines, OpenCL is expected to start shipping on mobile devices in 2012, becoming mainstream in mobile in 2013;
- OpenCV is a widely used imaging library that will potentially join Khronos to define an API to enable acceleration of advanced imaging and tracking software;
- StreamInput is a recently initiated Khronos working group that is defining a high-level, yet flexible framework for dealing with multiple, diverse sensors, enabling system-wide time-stamping of all sensor samples and display outputs for accurate sensor synchronization, and presenting high-level semantic sensor input to applications;
- COLLADA is an XML-based 3D asset format that can contain all aspects of 3D objects and scenes including geometry, textures, surface effects, physics and complex animations. COLLADA can be used to transmit 3D data over a network – or can be encoded to suit a particular application or use case;
- EGL is a window and surface management API that acts as an interoperability hub between the other Khronos APIs – enabling images, video and 3D graphics to be flexibly and efficiently transferred for processing and composition;
- OpenSL ES, not shown on the visual flow diagram, is an advanced native audio API that provides capabilities from simple alert sounds, through high-quality audio mixing through to full 3D positional audio that interoperates well with OpenGL ES 3D visuals.

Using these Khronos APIs, it is now becoming possible to create a mobile AR application that uses advanced camera and sensor processing from any device using the APIs to feed an accelerated image processing pipeline, that in-turn inputs to an accelerated visual tracker, that drives an advanced 3D engine that flexibly composites complex 3D augmentations into the video stream – all accompanied with a fully synchronized 3D audio stage.

6.4 Making Browser's AR Capable

Many developers and middleware vendors have strived to create application frameworks to enable content that is portable across diverse hardware platforms. The collection of standards and initiatives known as HTML5 is turning the browser into an application platform capable of accessing platform resources such as memory, threads and sensors – creating the opportunity for web content to be highly capable as well as widely portable.

Enabling browsers to support AR requires the Khronos native system resources be made available to web developers, typically leveraging the main components of a modern browser: JavaScript, the DOM and CSS.

The first 'connect point' between the native world of C-based acceleration APIs and the web is WebGL from the Khronos Group that defines a JavaScript binding to the OpenGL ES graphics rendering API – providing web developers the flexibility to generate any 3D content within the HTML stack without the need for a plug-in.

6.5 Declarative Programming

Some content creators, particularly those working on the web, prefer to use declarative abstractions of 3D visualizations. For instance, X3D is an ISO standard that allows the direct description of scenes and flow graphs. X3D can be used in conjunction with other standards, such as MPEG-4 and OGC CityGML. Besides a scene-graph description of objects, X3D also allows logic to be described in a declarative way.

Another declarative approach for describing soundscapes is the Audio Markup Language (A2ML) proposed by Lemordant.

7 Mobile AR architecture options

In the creation of AR applications, many disciplines may converge: computer vision recognition and tracking, geospatial, 3D etc. Algorithms for computer vision and sensor fusion are provided by researchers and software engineers or as third party services, while content for presentation and application logic are likely to be created by experience designers.

The most common standard for low-level application development on mobile devices today is C++. The problem with developing in C++ is that presentation and sensor access is different on each device and even the language used to access different sensors is not standardized. This results in a huge effort on the part of application developers for maintaining code bases. Thus, it would be desirable to have declarative descriptions of how AR logic is processed in order to have reusable blocks for AR application development. It would also be desirable to have a common language for definition of AR content.

All smart phones and other mobile devices today have a Web browser. Many already support elements of the HTML5 standard and this trend towards full HTML 5 support will continue. This represents a very clear and simple option for use by AR developers. Within a Web browser, JavaScript directly allows accessing the Document Object Model (DOM) and thus observing, creating, manipulating and removing of declarative elements. An option in standardization for AR could be the integration of declarative standards and data directly into the DOM.

A complete X3D renderer that builds on the WebGL standard and uses JavaScript has been implemented; it is called X3Dom (WebGL is a JavaScript interface for encapsulating OpenGL ES 2). X3Dom is completely independent of the mobile platform on which the application will run. This does not imply that an output solution, such as X3Dom, is all that's needed to make a universal viewer for AR. For example, a convenient interface extension for distributed access of real-time processed, concrete or distributed sensor data would be required.

A promising way for data synchronization of collaborative AR data within the network will be using existing standards, like the Extensible and Presence Protocol (XMPP), which has been established by the IETF. It has been implemented in chat applications and already delivers protocols for decentralized discovery, registration and resource binding. The work of the ARWAVE project could produce interesting results for mobile collaborative AR in the future.

8 Future mobile collaborative AR standards architectures

While in most of AR applications content is simply rendered over a user's camera view, AR can also be much more complex, particularly for mobile collaborative scenarios.

Technologies can, in the future, produce "reality filtering" to allow the user to see a different "reality," a view which provides a different position relative to the scene without the user moving, but maintaining perspective and context, in order to expose or diminish other elements in the scene. Work in this area will help maintain privacy, especially in collaborative applications, or help to focus users' attention more narrowly on the main task in context.

Another active area of research of potential value is the occlusion of objects when rendering over an image. This technology could improve the feeling of immersion of the augmentation. The user's hand might appear over or on top of the augmentation, rather than being covered by the projected virtual image. Additional sensors for depth perception would be required in order to present a correctly occluded composition. Imminent mobile devices with stereo cameras may bring occlusion correction closer to reality.

The application of existing standards or their extensions for AR, on the sensor side in mobile devices requires more research on delivery of improved context for fixed objects. Additionally, processing of moving objects in the scene and variable light conditions needs to be improved. Improvements such as these, relying on both hardware and software, are currently the subject of research in many laboratories.

In a collaborative virtual or physical environment, there may be communicating hardware devices that interchange feature data (at a pixel level), annotations, or other abstractions of task specific data, in order to enhance creativity, capacity of shared spaces, resources, and stability. When networked, such collaborative devices could be provided to increase the perception of immersion and to provide a fluid and productive environment for collaboration.

These potential future AR applications will require deep interoperability and integration of sensors for data acquisition and presentation and significantly enhanced use of advanced silicon for imaging, video, graphics and compute acceleration while operating at battery-friendly power levels. Further, AR will benefit greatly from the emergence of interoperable standards within many divergent domains. Other standards topics that will need to be addressed by the community's collaboration with other domains in the future include rights management, security and privacy.

9 Conclusions

In this chapter we have discussed the current status of standards that can be used for interoperable and open AR, the issues and directions of development and use of international standards in AR applications and related services. We have identified the different standard bodies and players in different fields of interest involved in the development of AR. We have also analyzed the current state of standards within the mobile AR segment, specifically.

From widely available standards and the numerous potential applications, it is clear that for the industry to grow there must be further research to agree on standards, profiles suited to AR and for there to be discussion among AR experts on many different levels of development. While extending existing standards will be highly beneficial to achieve the ultimate objectives of the community, there must also be room for the inclusion of new ideas and evolving technologies. Therefore, standardization meetings for finding the best interconnection and synergies have emerged (i.e. International AR Standards Meetings).

These standards coordination meetings enable all players to not only discuss requirements, use cases, and issues but also the establishment of focused working groups that address specific AR standards issues and the generation and coordination of AR-related work items in the cooperating standards bodies. This approach fosters and enhances the process of standardization of AR in specialized fields in order that the community develops seamless and stable working products in the market.

References

1. ARWave Project Web page http://arwave.org
2. CityGML. 2008. OGC CityGML Encoding Standard. Open Geospatial Consortium, http://www.opengeospatial.org/standards/CityGML.
3. ECMA-262. 1999. Standard ECMA-262 3rd Edition. European Association for Standardizing Information and Communication Systems, http://www.ecma.ch/ecma1/STAND/ECMA-262.HTM KML. 2008. http://www.opengeospatial.org/standards/kml/
4. GML. 2006. OGC Geography Markup Language Encoding Standard 3.2.1. Open Geospatial Consortium, http://www.opengeospatial.org/standards/gml.
5. HTML5. 2011. Vocabulary and associated APIs for HTML and XHTML, http://dev.w3.org/html5/spec/.
6. KRONOS. 2008. *COLLADA 1.5.0 Specification.*
7. http://www.khronos.org/collada/
8. Pusta et al.. Spatial Relationship Patterns: Elements of Reusable Tracking and Calibration Systems, ISMAR '06 Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality.
9. SOS. 2008. OGC Sensor Observation Service Interface Standard 1.0. Open Geospatial Consortium, http://www.opengeospatial.org/standards/sos.
10. WebGL. 2011. Open Khronos Group. WebGL Specification, http://www.khronos.org/webgl/specs/
11. WMS. 2007. OGC Web Map Service Interface Standard 1.3. Open Geospatial Consortium, http://www.opengeospatial.org/standards/wms.
12. WPS. 2009. OGC Web Processing Service Interface Standard. Open Geospatial Consortium, http://www.opengeospatial.org/standards/wps.

http://www.springer.com/978-1-4419-9844-6

Recent Trends of Mobile Collaborative Augmented Reality Systems
Alem, L.; Huang, W. (Eds.)
2011, XIII, 176 p., Hardcover
ISBN: 978-1-4419-9844-6