Augmenting the World using Semantic Web Technologies

Jens Grubert[+] Lyndon Nixon[†] Gerhard Reitmayr[+]

ABSTRACT

Creating and maintaining augmented scenes for mobile Augmented Reality browsers can be a challenging and time consuming task. The timeliness of digital information artifacts connected to changing urban environments require authors to constantly update the structural representations of augmented scenes or to accept that the information will soon be outdated.

We present an approach for retrieving multimedia content and relevant web services for mobile Augmented Reality applications at runtime. Using semantic web technologies we are able to postpone the retrieval of actual media items to the moment a user actually perceives an augmented scene. This allows content creators to augment a scene only once and avoid continous manual updates.

We also discuss the tradeoff between runtime content retrieval using Linked Data concepts and decreased control over the scene appearance at the time of authoring that comes along with this approach.

Keywords: augmented reality; authoring; thing of interest; semantic web.

1 INTRODUCTION

The creation of Augmented Reality (AR) scenes of high visual quality and with rich interaction possibilities typically involves an authoring process using proprietary authoring software or custom software development. The result of the process is most often a structural description (e.g., in xml) of the augmented scene connecting multimedia assets (like 3D models, videos) with tracking targets and predefined events (e.g., play animation, open webpage) that can be triggered by user's actions.

Creating augmented scenes for mobile AR browsers requires developers to manually edit these structural descriptions. Content authoring solutions such as Unity (for games) or Layar creator (for augmenting print products) help authors to create interactive AR experiences without the need to understand the underlying structural descriptions.

However, at the end of most authoring processes concrete multimedia assets are connected to tracking targets. While this might be desirable for some use cases such as interactive AR games we foresee a number of scenarios that would benefit from the retrieval of content at runtime.

Especially, the specification of prefixed assets for geo-located Points of Interest (POIs) might not meet the requirements of changing interaction contexts in urban environments for which many channels, layers or worlds are created. As an example, imagine a POI for a café that describes a part of the menu. This potentially helps users to make a decision whether to go into that café or not - if the café is open. However, it will be useless if the café is closed and the user wants to get a coffee at that point in time. In that case it might be helpful for the user to see information about a nearby associated café. Similarly, the graphical representation of the content could be an image of the menu or a smaller textual label – depending on the handheld device (tablet or smartphone) the user has.

Instead of linking concrete assets to tracking targets (locations or objects) we propose to instead link into the growing amount of structured, machine readable data that is available on the web and let the mobile AR application retrieve assets at runtime; allowing the context of interaction (in its simplest form time, place, device) to influence the selection of assets.

2 CONTENT RETRIEVAL USING SEMANTIC WEB TECHNOLOGIES

In this section we will discuss the approach used in the SmartReality project [8] of integrating content retrieved by semantic web technologies into a mobile AR application. This should serve especially the AR browser community as example of how existing mobile AR infrastructure could be adapted to include semantic data retrieval as one means to deploy content.

Our approach draws on the structured data being published as what is known as "Linked Data" on the Web (http://linkeddata.org). By using Linked Data, concepts can be referred to in structured descriptions of AR scenes via identifiers (URIs) just as concrete assets can be. While we use a format called Things of Interest [9] (similar to POIs) for describing concepts connected to the augmented scene and employ a Web-based authoring tool [2] those URIs could easily be integrated into other authoring solutions and structural descriptions. Furthermore, we chose to use a client-server model that lets a server crawl Linked Data repositories and return links to the final assets to the client. However, the functionality described next could be integrated into any AR client application.

The approach of turning concepts into assets consists of two stages. First, given a single concept URI, relevant metadata about that concept is crawled. Afterwards, the metadata is used to retrieve the final content that is displayed in the AR client.

2.1 Crawling Metadata

Metadata about the initial concept URIs can be crawled from the Web, or retrieved from Linked Data caching servers which already provide interfaces to the main Linked Data sources (like DBPedia [1]). Information collection can be done generically for any Linked Data concept by iterating over all linked URIs to a certain depth or triple limit, such as is implemented in the Linked Data crawler ldspider [1]. After relevant metadata for an initial concept (specified by a single URI) is crawled, associated assets have to be retrieved.

[+]: Institute for Computer Graphics and Vision, Graz University of Technology. {lastname}@icg.tugraz.at
[†] STI International. lyndon.nixon@sti2.org

2.2 Retrieving the assets

Since there is usually a lack of (consistent or usable) linkage to media resources about any concept in its own Linked Data metadata, we take a semantic services approach to deriving usable content for the annotated concepts. This means: we annotate Web APIs according to the sort of concepts they can provide content for, the sort of content they can provide in response and the means to generate an input request to the API. Finally, the output response in terms of content items is returned. We use the Linked Services approach for this supported by the tools OmniVoke, iServe and SmartLink [2]. Since most Web APIs do not directly support Linked Data URIs, we define a means to "lower" an URI to literals which form part of the API request, as well as "lift" the API response to a set of annotated media resources which can be used in the enriched Augmented Reality view.

Currently, the Linked Services infrastructure used by the SmartReality platform supports a number of general Web APIs (YouTube, Flickr, Ookaboo, Foursquare) but also two services specifically provisioned for a music event poster scenario to demonstrate how any Web-based API could be supported in this workflow. A music streaming service from the website play.fm [7] is called to link music artists annotated in the poster to music streams available from play.fm by that artist. Also a hotel booking service from Seekda [8] is called to link a music event to available hotel rooms, making use of the available machine processable metadata to directly select the most appropriate offers without additional input by the user; the event date is used for the date of the hotel stay, and the event location is used to ensure only close by hotels are offered. As a result, from a music event poster, our approach can directly provide a link to an available hotel room for the stay after that event. Again, the services approach is based on service annotations, so any service could be integrated into such a workflow as soon as an annotation of that service.

2.3 Optimizations

To make the use of online data and services more efficient, both steps described above make use of middleware technology which can also usefully cache previous Linked Data or Linked Services responses and make the respective metadata or content immediately available again. For Linked Data, we employ the Linked Media Framework [4] which is also a RDF repository capable of automatically caching metadata of Linked Data resources it has requested as well as syncing that metadata subsequently. For the services, the Linked Services infrastructure makes use of a Sesame interface to the OWLIM RDF repository to cache Web API responses for each concept, and hence if the same concept is queried for again the media resource information (service responses) can be directly loaded from the local repository instead of making again the API call. This local cache can also be regularly updated to keep the media resource links fresh.

3 Limitations

During an authoring process multimedia assets are typically placed inside a specific location (position, orientation, scale) within the coordinate system of the augmented scene and connected with events triggered by users' actions. Many events that can be specified with current AR browser scripting facilities can work with arbitrary assets (e.g., on proximity, on select) and therefore are also applicable to assets retrieved via semantic web technologies at runtime. Furthermore, if the augmented scene mainly consists of labels, images and videos as it is common with many AR browser channels semantically retrieved content can be integrated. Videos, images and textual description about many concepts are widely available in the Linked Open Data cloud.

However, if a tight visual integration of digital content within the physical scene is desired, e.g. in games, the assets should be known beforehand and content retrieval at runtime is not guaranteed to deliver satisfying results (imaging a video with alpha channel that blends into an existing magazine cover). Furthermore, until today not many 3D models are described in publicly available in Linked Data repositories and therefore cannot be retrieved via semantic web technologies.

An additional hurdle to integrate content retrieved via semantic web technologies into augmented scenes is the knowledge about specialized linked data sources needed by authors. Assets about general concepts available in repositories like DBPedia are suitable for a wide range of applications. However, if authors want to constrain the selection of assets to a specific field and type (e.g., music, videos and audio streams) they have to employ specialized ontologies (like BBC Music and Musicbrainz) and data sources (e.g., data.play.fm for audio streams). Furthermore, if the content retrieval should be guided by more than simple crawling patterns (follow a certain depth and breadth in the data graph) authors still have to provide specific crawling rules. For example, in the SmartReality project we use crawling rules for fetching DJ mixes of music artists. These rules can be expressed using path expressions or Datalog rules.

4 Conclusion

Leveraging the existing Web of data can lower costs for both individual authors as well as whole organizations seeking to offer Augmented Reality based services to users. Instead of relying on one's own generated and curated data, sources existing and maintained online can be dynamically and automatically integrated into the workflow authoring and finally viewing an Augmented Reality scene. We hope the presented approach will serve as example showing potentials and limitations of dynamically retrieving content via semantic web technologies.

References

[1] DBPedia. http://dbpedia.org. Last accessed on July 16th 2012.

[2] J. Grubert, R. Gründler, L. Nixon and G. Reitmayr. Annotate That: Preparing Event Posters for Augmentation. In *ISMAR 2011 Workshop on Authoring Solutions for Augmented Reality*. 2011.

[3] R. Isele, J. Umbrich, C. Bizer and A. Harth. LDSpider: An open-source crawling framework for the Web of Linked Data. In *ISWC 2010*.

[4] Linked Media Framework. http://code.google.com/p/lmf/. Last accessed on July 16th 2012.

[5] L. Nixon, J. Grubert and G. Reitmayr. SmartReality. Augmented Reality + Services + Semantics. In *2nd International Augmented Reality Standards*. 2011.

[6] C. Pedrinaci, D. Liu, M. Maleshkova, D. Lambert, J. Kopecky and J. Domingue. iServe: a linked services publishing platform. In *Ontology Repositories and Editors for the Semantic Web Workshop*. 2010.

[7] Play.fm - Online Radio for Electronic Dance Music. http://www.play.fm. Last accessed on July 16th 2012.

[8] seekda Technologie - e-Tourismus, Multi Channel Management, Konferenzmanagement, Registrierung. https://www.seekda.com. Last accessed on July 16th 2012.
[9] SmartReality. Things of Interest. http://www.smartreality.at/rdf/toi. Last accessed on July 16th 2012.