A Semantically Enriched Augmented Reality Browser

Tamás Matuszka[1,2], Sándor Kámán[1], and Attila Kiss[1]

[1] Eötvös Loránd University, Budapest, Hungary
{tomintt,kasraai,kiss}@inf.elte.hu
[2] Inter-University Centre for Telecommunications and Informatics, Debrecen, Hungary

Abstract. Owing to the remarkable advancement of smartphones, Augmented Reality applications have become part of everyday life. Augmented Reality browsers are the most commonly used among these applications. The users can search and display interesting places from the physical environment surrounding them by means of these browsers. Some of the most popular AR browsers use only one data source and the openly available datasets are not used. In contrast, the main objective of Linked Open Data community project is to link knowledge from different data sources. This pursuit makes it easier to retrieval information, among others. In this paper, an Augmented Reality browser was presented. Information derived from Linked Open Data was used by the browser as data source. Due to this, the system is able to handle more data sources.

Keywords: Augmented Reality, Semantic Web, Location-based Services, Linked Data.

1 Introduction

Augmented Reality applications, which combine the real and virtual worlds in real-time, are more and more widespread nowadays. The history of Augmented Reality dating back to the last few decades, when the hardware possessed by researchers was typically weaker than nowadays. As a result, the use of different, expensive as well as uncomfortable wearable devices (for instance, computer in backpack, head-mounted display) was necessary. Due to the advancement of hardware, computer graphics and mobile information technology, Augmented Reality has been widespread in the everyday use as well. Instead of expensive, inconvenient devices; the small, affordable and convenient smartphones can be used as a platform for Augmented Reality.

In the past few years, several Augmented Reality applications have been published in the field of medical applications, education, tourism, etc. [15], [7], [3]. Among these, perhaps the most widely used application is the Augmented Reality browser that combines the traditional Augmented Reality application with the Internet browsing. This kind of Augmented Reality takes advantage of the

R. Shumaker and S. Lackey (Eds.): VAMR 2014, Part I, LNCS 8525, pp. 375–384, 2014.

user's current geographical location and location-based information can be superimposed into the real life view. A typical example is when the user looks around with the mobile phone and could see the icons which represent restaurants located near in the real-life view. The most important criterion of a browser is the amount and variety of accessible data [12]. The current augmented reality browsers (e.g. Junaio[1], Layar[2], Mixare[3]) use only one data source and the openly available datasets are not used. Wikitude built on Augmented Reality Markup Language (ARML)[4], Mixare and Layar use hidden and proprietary data structures [24]. In contrast, the main objective of Linked Open Data (LOD) [5] community project is to link knowledge from different data sources. Several publicly available data source can be found on the Internet in semantically represented format. A wider knowledge base can be obtained by means of interlinked data sources. The Linked Open Data can be used for data source of Augmented Reality browser as well, as opposed to one data source, which is used by the recently browsers.

In this paper, a Linked Data-driven Mobile Augmented Reality browser will be discussed. The Semantic Web was used for data integration and data retrieval purposes. The Semantic Web is able to manage the data available on the Internet [2]. A lot of publicly available datasets have been published in semantically form currently. The Semantic Web stores the information in RDF (Resource Description Framework) statements about resources in the form of subject-predicate-object expressions. These expressions are known as triples in RDF terminology [13]. Information can be obtained from these data sources with SPARQL [20] query language that is able to retrieve and manipulate data stored in RDF format. With the help of this application, a user can navigate and collect local-aware information. A sensor-based tracking approach was combined with RDF processing of related geographical data. The used data come from semantically represented data source from the Linked Open Data. A map-based navigation is also provided by the system. We have implemented our browser on Android operation system which allows for a widespread of usability. One of the main challenges was correctly positioned the displayable virtual elements on the screen of mobile phone. For this purpose, a mathematical model [24] was used.

The organization of the rest of this paper is as follows. After the introductory Section 1, we outline the preliminary definitions in Section 2. Section 3 deals with the related work. Then, the details of our system is described in Section 4. Finally, the conclusion and our future plans are described in Section 5.

2 Preliminaries

In this section, the concepts that are necessary for understanding are defined. We provide insight into the basic concepts of Semantic Web and Augmented Reality.

[1] http://www.junaio.com/
[2] https://www.layar.com
[3] http://www.mixare.org
[4] http://openarml.org

A possible way to manage the data available on the Internet is to use the Semantic Web [2]. The Semantic Web aims for creating a "web of data": a large distributed knowledge base, which contains the information of the World Wide Web in a format which is directly interpretable by computers. *Ontology* is recognized as one of the key technologies of the Semantic Web. An *ontology* is a structure $\mathcal{O} := (C, \leq_C, P, \sigma)$, where C and P are two disjoint sets. The elements of C and P are called *classes* and *properties*, respectively. A partial order $\leq_C$ on C is called class hierarchy and a function $\sigma: P \rightarrow C \times C$ is a signature of a property [23]. The Semantic Web stores the knowledge base as *RDF triples.* Let I, B, and L (IRIs, Blank Nodes, Literals) be pairwise disjoint sets. An *RDF triple* is a $(v_1, v_2, v_3 \in (I \cup B) \times I \times (I \cup B \cup L))$, where v_1 is the subject, v_2 is the predicate and v_3 is the object [19].

Augmented Reality applications are more and more widespread nowadays. With its help the real physical environment can be extended by computer generated virtual elements creating the illusion that the two worlds coexist. Augmented Reality has two different types. The first one is the marker-based Augmented Reality and the second one is the position based Augmented Reality. The marker-based one uses a so-called marker. This marker allows the registration of the virtual object in the physical space. The position based Augmented Reality depends on the user's physical position which is determined by GPS coordinates [16].

According to Azuma's definition, an Augmented Reality system combines real and virtual reality; is interactive in real-time and is registered in 3D [1]. Another, formal definition for the same concept was given by us [18]. A quintet $\langle M, VE, T, \varphi, \xi \rangle$ is called as Augmented Reality system, where M is the set of markers, VE is the set of virtual elements, T is the set of transformations, φ is the mapping function and ξ is the transformation function. Let IB, PB (image-based markers and position-based markers) be two disjoint sets. Then, M can be written as follows: $M = IB \cup PB$. Let I, V, S and K (images, videos, sounds, knowledge base) be pairwise disjoint sets. In this case, $VE = I \cup V \cup S \cup K$. The T set contains geometric transformations, namely translation (τ), rotation (ρ) and scale (σ). In addition, let L be the set of 3D vectors. Every $v \in VE$ virtual element has an $l \in L$ vector. This l vector stores the position of v virtual element. The function $\varphi: M \rightarrow VE \times L$ maps a virtual element and its relative initial position to a marker. The last part of the quintet is the transformation function ξ. The function $\xi: M \times VE \times L \times T \rightarrow VE \times L$ transforms a virtual element corresponding to the given marker with a given transformation in real-time. The current Augmented Reality systems can be modeled by the above definition.

In this paper, we present an $\langle M|_{PB}, VE|_{KB}, T, \varphi, \xi \rangle$ Augmented Reality system, where the set of markers is restricted to position-based markers (i.e. the markers are limited to position-based type) and the set of virtual elements is restricted to knowledge base (i.e. the virtual elements are solely location-aware information). This knowledge base is stored in RDF triples and it derived from DBpedia [4] that is included in Linked Open Data. The function φ assigns relevant information based on the latitude and longitude coordinates of the places represented by markers.

3 Related Work

Augmented Reality has been thoroughly researched over the last few years. Several early projects have been started which aims to create augmented reality browser. The [9] presents a location-, and spatial-aware system which enables the integration of existing information systems with virtual objects in easily way. Feiner et al. describe a wearable Augmented Reality prototype and a campus information system case study in [6]. A backpack, head-worn display and hand-held display were required by the user of prototype to the navigation. Rekimoto, Ayatsuka and Hayashi in [22] show a system that is able to assign virtual content (e.g. audio, video) to physical location. The user of system can see these virtual contents as an Augmented Reality post-it. A head-mounted display (HMD) has to be worn by the user also in this case. Kooper and MacIntre present in [11] a general Augmented Reality system, which connects physical locations with World Wide Web. Head-mounted display is also required in this case.

The usage of inconvenient backpack and HMD has become unnecessary due to the improvement of hardware of mobile phones and the spread of smartphones. The smartphones can be used for browsing purpose instead of uncomfortable wearable devices. Kähäry and Murphy in [10] show a hand-held, video-see through Augmented Reality system, which can interact with the surrounding environment through a simple interface. Luna et al. describe a navigation system in [14] that has an Augmented Reality view as well. They claim that their system is friendlier than the similar applications which enable the users to navigate along streets. They argue that the navigation along known places (e.g. shops, monuments) is friendlier than the original solution. Traditional navigation system cannot deal with the user's preferences, the environment and omitted important information: the user context. A system based on the POIs which provides different paths depending on the user preferences was proposed by the authors, so that path leads to a specific target through the well-known places. An open architecture and corresponding prototype is presented by Hill et al. in [8]. Hill's system allows quickly mobile Augmented Reality application development. HTML authoring tool, client-side JavaScript and database connection with AJAX were used for this purpose. The main advantage of the system is the easy content creation according to the authors.

Today, there have been some attempts to combine Semantic Web and Augmented Reality. In a previous paper [17] we have tried to combine the advantages of Augmented Reality and Semantic Web to provide indoor navigation. In this work, we describe an indoor navigation system which uses Augmented Reality for visualization of navigation. The storage of data which are necessary to the navigation was based on ontology. In addition, the possible paths are generated by rule-based inferences. In another previous work [16] we present the conception of a general, Semantic Web-based Augmented Reality system. Due to this framework, the user would navigate through arbitrary areas and get context-aware information. Martín-Serrano, Hervás and Bravo show a touristic Android application in [15] that uses Web 3.0 technology tools to data extraction from various data sources with help of publicly available services on the Internet.

Ontology was used for determining the user's context. In addition, the application provides a recommendation system that based on rule-based inferences. Schubotz and Harth describe a case study in which Linked Open Data [5] is combined with spatial and encyclopedic data sources. The resulting data was transformed to Web3D renderable format. The most LOD browser supports only navigation and exploratory discovery and the query visualization happened in table and list view. Instead, they present a system that uses Web3D for visualization the result of a LOD query [21].

4 Details of the Prototype

In this section, we overview the motivation and the basic functions of our system. Afterwards, we describe the architecture of the prototype. Finally, we present the mathematical model which was used to position the displayable virtual elements on the screen of mobile phone.

4.1 Motivation

Imagine, that you are in a foreign city or you just have a few free hours and you do not know what interesting places are in the near. Maybe you are hungry too. It would be great, when you could somehow find out, what kind of places are in the near, which you are interested in. Of course you can ask somebody, but he or she may not know every place in the city. It would be great, if you could reach a database, from which you could get a lot of information about your environment.

In these or similar situations our application can help for you. You have to nothing else to do, just grab your smartphone with Android operating system and start the application. The browser connects to the DBpedia database and downloads the information about your near environment. You can see where are the sights and attractions. With the help of Google Maps, the program can easily plan a direction for you, so you will not be lost. You can also get information from the selected point, which is also downloaded from DBpedia.

4.2 The Basic Functions of the Browser

The main objective of our Linked Data-drive Mobile Augmented Reality browser is to help navigate the users through the environment and to gather local-aware information. These information come from the most well-known semantic dataset, namely DBpedia, which contains the knowledge of Wikipedia in semantic form. DBpedia contains the latitude and longitude coordinates of numerous places, therefore it can be used as geographical data source. A sensor-based tracking approach was combined with RDF processing of DBpedia. The GPS (Global Positioning System) is used for get the objects position on the whole surface of Earth. It is supported by the satellites around the planet. In order to set our position, the GPS receiver has to communicate with 4 satellites in the case of 2D positioning, and 5 satellites in case of 3D positioning. In our application, we only use 2D positioning. There are weaker and cheaper GPS receivers in the cellphones, which are perform better with network access. The phone downloads the orbital position of the satellites and that way the weaker receiver can find them faster.

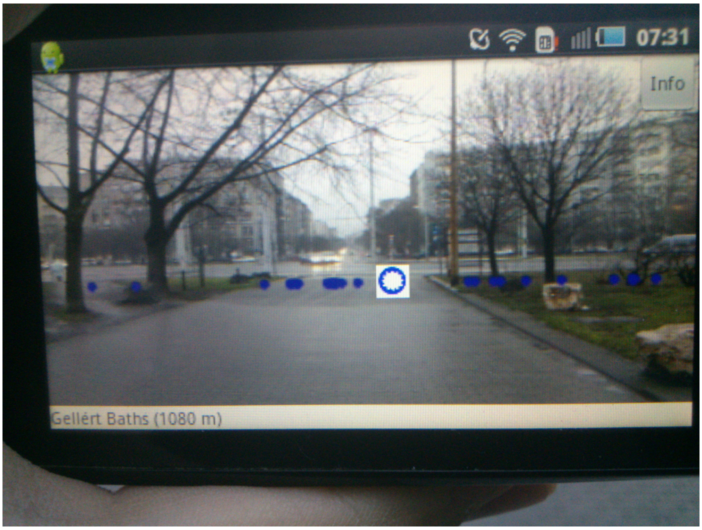

Fig. 1. The Augmented Reality view of the browser

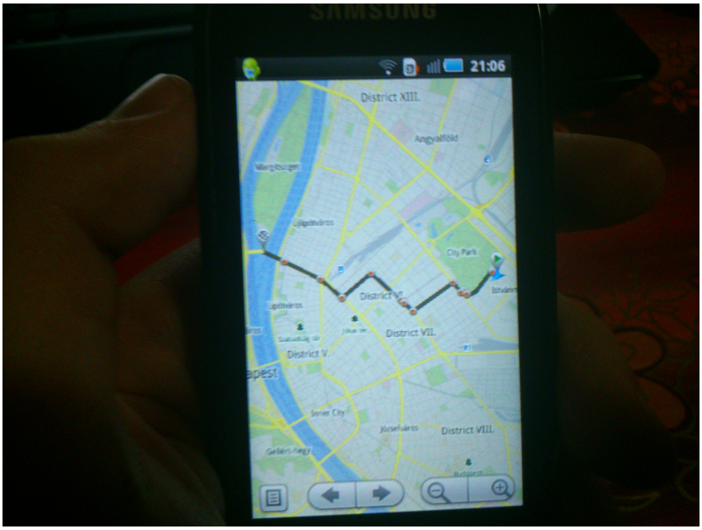

Fig. 2. Map view, the direction is provided by Google Maps

When the user starts the application, he or she can see the POI-s on the real-world view via camera's display. The data of POI-s are originally stored in DBpedia and were converted to POI by a transformation step. The main activity includes a special POI, which has a differential icon, depicted in Figure 1. This POI is the active one, the user can read informations about them in the bottom of the screen which contains the name of active POI as well as the distance from the current position of the user. Further information can be obtained from the active POI based on the DBpedia on the Info activity. This view contains the corresponding photos and a short description about the POI. A map-based navigation is also provided by this view, the user can get the direction with the help of Google Maps from the Info activity which can be viewed on Figure 2.

4.3 The Architecture of the System

The system is separated into two parts: the database and the client. The database is the DBpedia itself, while the client was implemented on Android operation system which allows for a widespread of usability. The schematic diagram of the architecture can be seen on the Figure 3.

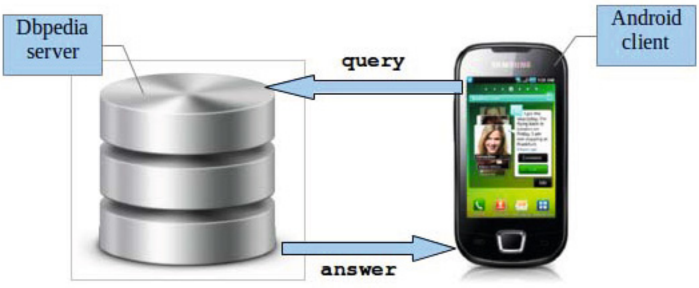

Fig. 3. The system architecture

The database contains RDF triples. The DBpedia is built on the Virtuoso database handler system which is the first cross-platform server, which implements web, file and database server functionalities. The DBpedia-based data is queried from the client with using AndroJena API which is a porting of Jena semantic web framework to the Android platform. Jena is a free and open source Java framework for building Semantic Web and applications. The client connects to DBpedia and sends a SPARQL query to it, which will run on the DBpedia server. Figure 4 depicts an example SPARQL query. The given result was processed by the client. The client makes POI objects from these data. To take account of the low computing capacity of the cell phones, we limited the size of the queries in 100 rows.

```
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX dbpedia-owl: <http://dbpedia.org/ontology/>
SELECT ?label ?lat ?long ?abstract ?pic WHERE {
        _:a geo:lat ?lat .
        _:a geo:long ?long .
        _:a rdfs:label ?label .
        _:a dbpedia-owl:abstract ?abstract .
        _:a dbpedia-owl:thumbnail ?pic .
        FILTER(?lat - 35.341846 <= 0.05 &&
                35.341846 - ?lat <= 0.05 &&
                ?long - 25.148254 <= 0.05 &&
                25.148254 - ?long <= 0.05 &&
                lang(?label) = "en" &&
                lang(?abstract) = "en"
                )
} LIMIT 100
```

Fig. 4. SPARQL query which collects data of POI-s

4.4 The Mathematical Model

Several difficulties occur during on the correctly positioning of the virtual elements. In order to know, where to put the POI on the screen, we used a lot of mathematical calculations. A lot of parameters have to be used for determine the position of the virtual element on the screen of mobile device, for instance the longitude, latitude, altitude of POI and camera; range, azimuth and pitch of device. The details of the parameters, the formal definitions as well as the mathematical model can be read in [24].

5 Conclusion

In this paper a mobile Augmented Reality browser was presented, which extracts the information from Linked Open Data, in contrast to the similar current browser. In that case, only one data source is used by the recently applications. Semantic Web was used for information retrieval as well as a sensor-based tracking approach was combined with RDF processing. Due to the Augmented Reality browser, the user is able to display the surrounding related geographical data from DBpedia, extract information about them, and navigate to the selected place as well. The browser was implemented on Android operation system which allows for a widespread of usability.

In the future, we are planning to integrate another data sources by using Semantic Web technologies. In order to validate the work we have carried out, a more in-depth investigation into data integration is needed. We are intending to use not only semantic data sources, but also arbitrary geographical data sources (e.g. Foursquare, Google Places API, Facebook API) as well. We also want to implement a touristic, sightseeing application with recommendation system. The user can set certain conditions (for instance the available time, maximal distance, interested sights) and the system proposes some sightseeing trips which contain the maximal number of given attractions. We would like to formalize this task and give a mathematical solution for it as well.

Acknowledgments. This work was partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013) and TAMOP-4.2.2.C-11/1/KONV-2012-0001 supported by the European Union, co-financed by the European Social Fund.

References

1. Azuma, R.T.: A survey of augmented reality. Presence 6(4), 355–385 (1997)
2. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American 284(5), 28–37 (2001)
3. Billinghurst, M., Dünser, A.: Augmented reality in the classroom. Computer 45(7), 56–63 (2012)
4. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S.: DBpedia-A crystallization point for the Web of Data. Web Semantics: Science, Services and Agents on the World Wide Web 7(3), 154–165 (2009)
5. Bizer, C., Jentzsch, A., Cyganiak, R.: State of the LOD Cloud, `http://wifo5-03.informatik.uni-mannheim.de/lodcloud/state/`
6. Feiner, S., MacIntyre, B., Höllerer, T., Webster, A.: A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. Personal Technologies 1(4), 208–217 (1997)
7. Hansen, C., Wieferich, J., Ritter, F., Rieder, C., Peitgen, H.O.: Illustrative visualization of 3D planning models for augmented reality in liver surgery. International Journal of Computer Assisted Radiology and Surgery 5(2), 133–141 (2010)
8. Hill, A., MacIntyre, B., Gandy, M., Davidson, B., Rouzati, H.: Kharma: An open kml/html architecture for mobile augmented reality applications. In: 2010 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 233–234. IEEE (2010)
9. Hohl, F., Kubach, U., Leonhardi, A., Rothermel, K., Schwehm, M.: Next century challenges: Nexus an open global infrastructure for spatial-aware applications. In: Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking, pp. 249–255. ACM (1999)
10. Kähäri, M., Murphy, D.J.: Mara: Sensor based augmented reality system for mobile imaging device. In: 5th IEEE and ACM International Symposium on Mixed and Augmented Reality (2006)
11. Kooper, R., MacIntyre, B.: Browsing the real-world wide web: Maintaining awareness of virtual information in an AR information space. International Journal of Human-Computer Interaction 16(3), 425–446 (2003)
12. Langlotz, T., Grubert, J., Grasset, R.: Augmented reality browsers: essential products or only gadgets? Communications of the ACM 56(11), 34–36 (2013)

13. Lassila, O., Swick, R.R.: Resource Description Framework (RDF) Schema Specification,
http://www.w3.org/TR/rdf-schemahttp://www.w3.org/TR/rdf-schema
14. Ma Luna, J., Hervás, R., Fontecha, J., Bravo, J.: A friendly navigation-system based on points of interest, augmented reality and context-awareness. In: Bravo, J., López-de-Ipiña, D., Moya, F. (eds.) UCAmI 2012. LNCS, vol. 7656, pp. 137–144. Springer, Heidelberg (2012)
15. Martín-Serrano, D., Hervás, R., Bravo, J.: Telemaco: Context-aware System for Tourism Guiding based on Web 3.0 Technology. In: 1st Workshop on Contextual Computing and Ambient Intelligence in Tourism, Riviera Maya, Mexico (2011)
16. Matuszka, T.: Augmented Reality Supported by Semantic Web Technologies. In: Cimiano, P., Corcho, O., Presutti, V., Hollink, L., Rudolph, S. (eds.) ESWC 2013. LNCS, vol. 7882, pp. 682–686. Springer, Heidelberg (2013)
17. Matuszka, T., Gombos, G., Kiss, A.: A New Approach for Indoor Navigation Using Semantic Webtechnologies and Augmented Reality. In: Shumaker, R. (ed.) VAMR 2013, Part I. LNCS, vol. 8021, pp. 202–210. Springer, Heidelberg (2013)
18. Matuszka, T., Kiss, A.: Alive Cemeteries with Augmented Reality and Semantic Web Technologies. In: Proceedings on ICHCI 2014 (accepted, 2014)
19. Pérez, J., Arenas, M., Gutierrez, C.: Semantics and Complexity of SPARQL. In: Cruz, I., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L.M. (eds.) ISWC 2006. LNCS, vol. 4273, pp. 30–43. Springer, Heidelberg (2006)
20. Prud'hommeaux, E., Seaborne, A.: SPARQL Query Language for RDF, http://www.w3.org/TR/rdf-sparql-query/
21. Schubotz, R., Harth, A.: Towards Networked Linked Data-Driven Web3D Applications. In: Dec3D (2012)
22. Rekimoto, J., Ayatsuka, Y., Hayashi, K.: Augment-able reality: Situated communication through physical and digital spaces. In: Second International Symposium on Wearable Computers, Digest of Papers, pp. 68-75. IEEE (1998)
23. Volz, R., Kleb, J., Mueller, W.: Towards Ontology-based Disambiguation of Geographical Identifiers. In: I3 (2007)
24. Zander, S., Chiu, C., Sageder, G.: A computational model for the integration of linked data in mobile augmented reality applications. In: Proceedings of the 8th International Conference on Semantic Systems, pp. 133–140. ACM (2012)