Using mobile phone cameras
to interact with ontological data [*]

Pierluigi Grillo[1], Silvia Likavec[1], and Ilaria Lombardi[1]

Università di Torino, Dipartimento di Informatica, Torino, Italy
{grillo,likavec,lombardi}@di.unito.it

Abstract. In addition to using mobile phones to access data and services on the Internet, they can also be used to retrieve information about physical objects. In this paper we present a framework and a prototype implementation of a reality browsing system that exploits mobile camera phones as access points to ontological knowledge with the corresponding physical counterparts. The use of ontologies to represent the information about the objects of the domain permits for application of reasoning and querying techniques on domain elements. Users can identify a desired product by framing its logo, obtaining further information about the object in a direct, natural and intuitive way. In addition, users can perform Web 2.0 actions on objects, contributing to the creation and development of the network of objects and users. We also provide a brief description of the implemented interface and a real world user evaluation.

1 Introduction

With the integration of cameras into mobile phones, new horizons are emerging in the world of mobile interaction. Due to the possibility of constant connection mobile phones are being more and more used to access data and services on the Internet, as well as to retrieve the information about physical objects surrounding us. On the other hand, mobile devices have limited input and output capabilities (small keyboards and screens), thus making it tedious for users to type in huge amounts of data or visualize complex information. Innovative applications and interaction techniques are needed to support usage of mobile phones as interaction devices which can be used in new contexts and environments.

Users usually access the information about objects of interest either by using an Internet search engine (e.g. Google) or by typing an exact URL into a browser. In any case, the information obtained is often overwhelming and not directly related to the object in question. We try to overcome these limitations by proposing mobile camera phones as access points to ontological knowledge. The users can identify a desired object by framing its logo with the camera phone, thus obtaining further information about the object and more interaction possibilities. In our case, not only the access to the information is faster and easier, it is also more direct, providing the users with relevant information about the object in question. For other approaches using camera phones to interact with physical objects and access the related information see [11, 13, 9, 15].

We implemented *WantEat*, a prototype application [1, 3] that uses the Apple iPhone integrated camera to obtain information about products in a specific domain. The work presented here is a part of PIEMONTE Project [10], in which a Social Networking Application in the domain of gastronomy is being developed, integrating the ideas of the Social Web, Semantic Web and Augmented Reality (se also [5]). In this way a dynamic network of people and objects of the domain emerges, where the communication is possible in both directions, thus making objects of the domain active parts of the "society". The goal of the *WantEat* application is to enable users to access information about the products they might encounter while in mobility (e.g. being at the local farmer's market or visiting the local producer of a certain product), providing them with augmented reality experience. We want to offer the users intuitive and easy access to desired information, playful and innovative at the same time. On the other hand, users should be able to tag, vote and comment the objects they interact with, contributing to the creation and enhancement their social networks.

The main contributions of our work are:
- use of augmented reality on a mobile device for interaction with objects of interest which enables live recognition of the object in focus;
- use of Semantic Web aspects, i.e. ontologies to represent the information about the objects of the domain and their relations with other objects and users, hence having the possibility to use reasoning and querying techniques on such data;
- possibility to obtain information about the products not easily searchable with standard techniques;
- inclusion of Web 2.0 features, such as tagging, voting, and commenting on objects.

Furthermore, the prototype has been used in a real world context by a relatively large number of participants, providing realistic assessment of the approach.

The rest of the paper is organized as follows. In Section 2 we embed our work in the context of related work. Section 3 gives a brief overview of the system design, followed by the details of the recognition algorithm in Section 4. In Section 5 we present the results of the real world evaluation and we conclude in Section 6.

2 Related work

Due to their portable nature, smartphones have always been favored candidates for ubiquitous applications. In recent years their evolution has made possible a number of advanced applications, including augmented reality (AR).

Many AR applications for mobile devices are built for commercial and social purposes. For example, the *Layar Reality Browser*[1] is a free application for iPhone and Android OS smartphones that displays real time digital information on top of phone camera view. The *Layar* platform is the basis for a number of mobile location applications (such as tourist guides, live games) because the information is coded as points of interest organized in thematic layers.

---

[1] http://www.layar.com

Another iPhone application is proposed by the social review service *Yelp*[2]. It provides a feature called *Monocle* which uses the phone GPS and compass to display markers for restaurants, bars and other nearby businesses on top of the camera view. In a similar way, *Wikitude World Browser*[3] displays information about users' surroundings in camera view. In this case, the content comes from different sources like *Starbucks*, *Youtube*, *Qype* or *Wikipedia*. *Wikitude* proposes also *Wikitude Drive*, a navigation system that overlays video captured through the camera with driving instructions and other information.

The possibility to provide information about the surroundings in a natural way opens the door to discovery-based learning. The European research project *iTacitus*[4][8] studied how augmented reality can enhance user's experience in cultural heritage sites by overlaying 3D objects (like missing paintings or architecture models) or abstract context sensitive information on the real scene.

Our system does not simply provide information, it lets the user navigate the net of items (products, people, places) freely and easily, using the relations that link them. These connections are described in the domain ontology or can arise from users' actions and drive the user to serendipitous discovery and exploration of relationships among items.

Recently a number of mobile visual search applications have been developed, for example Google *Goggles*[5], *Kooaba*[6], Nokia *Point and Find*[7] and *SnapTell*[8]. In [4] the authors propose *Wikireality*, a system that, using GPS data and live camera input, recognizes a landmark and augments live video with a Wikipage icon pinned to the query object using orientation sensor data. The corresponding Wikipedia page is automatically opened in the background. The server-side image recognition has a preprocessing phase to prune the set of images used for the matching.

All these systems send a query photo to a remote server which compares it against a database. Even though various compression techniques lead to data of small dimensions, this server-centered approach suffers from network latency. On the contrary, in [14], similarly to what we propose, the recognition is performed on the mobile device. This allows the system to augment the phone camera live video with information about the objects it recognizes, in order to improve user's experience with faster live feedback.

3 System design

This section presents a short overview of the modules employed in our framework. The client-server architecture of our framework is based on the Resource Oriented Architecture approach [12] (see Figure 1). The main component of the

---

server side is the *Service Logic Layer* which coordinates various components. The *Exposure Layer* manages the communication between the *Service Logic Layer* and the client applications from the *Application Layer*, by exposing the features of *Service Logic* and monitoring the interaction with the user. The architecture has been realized using Java Environment where the Restlet tool [7] and Apache HTTP libraries collaborate in order to answer the HTTP requests from clients.
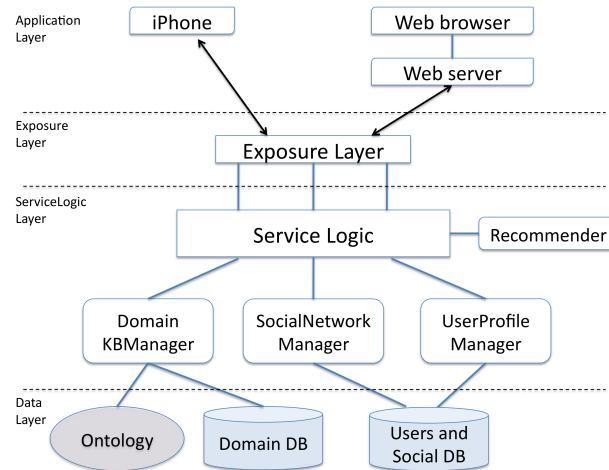

Fig. 1. System architecture

The *Data Layer* is the place where the knowledge about our system is stored. The domain ontology, represented in OWL[9], is used to model domain entities and relations among them, and is integrated with a relational MySQL database used to manage the information about the users and the social network features. The *DomainKB Manager* synthesizes a uniform representation of the two data sources in the *Data Layer* using the D2RQ tool [2] which allows to view a part of a relational database as a set of OWL individuals. Depending on the user interests, the content can be personalized employing the specialized modules *UserProfile Manager* and *Recommender*. Representing the knowledge base as ontology ensures that the domain entities are described as machine recognizable content, which facilitates querying and reasoning on them.

The domain of our application is gastronomy with many different kinds of domain items such as products (e.g. cheese or wine), actors related to products (restaurants, production companies, farms etc.), as well as recipes and geographical data. Our domain ontology consists of one upper ontology which imports a set of light-weight ontologies, each one describing a specific part of the domain. It contains very general classes such as `Salame`, as well as very specific ones, such as `Salame della Turgia di Bruno Tetti`. Further, a set of SWRL rules is used to define associations between classes. Hence, we were able to describe the domain entities with a fine level of granularity, including the products not easily searchable with standard techniques.

---

[9] `http://www.w3.org/TR/owl2-overview/`

4 Logo recognition

One of the main aims of the prototype we developed is to give users tools which would help them get in touch with products and all the elements of a territory, employing natural and easy interaction. In case of a product, the main and immediately characterizing property is its label or logo.

In the last years, many applications were developed based on recognition of barcodes, such as $DataMatrix$[10] or $QRCode$[11]. Although this approach is very fast and robust, barcodes are merely meaningless black-and-white images. Moreover, users need to take a photo in order to make the application retrieve the information or redirect them to a URL, making the interaction not so natural and detached from the real world.

With our approach we want to offer users an interaction modality which uses a mobile phone camera and is, on the one hand, natural, quick and immediate and, on the other hand, closely related and immersed into the surrounding environment. Interaction takes place in augmented reality mode where real world live images are enriched with information related to the identified object (Figure 2).

Fig. 2. Logo framing

The recognition algorithm is implemented on the device using the ARToolK-itPlus library,[12] and we designed the markers, recognizable by the algorithm, with the logos of the products (Figure 3). The simple marker images are constrained to have a black border, needed by the algorithm to locate the marker in the scene and to primarily estimate its position with respect to the camera in order to overlay the live scene with digital information placed in the right position, in accordance with augmented reality concepts. Although complex color images affect the effectiveness of recognition, we preferred to use the logos since they enable the users to easily recognize them and get in contact with products directly. This provides the right balance between the slow full remote image recognition and a real-time experience, even though this is a limitation with respect to the number of recognizable logos, since they must be stored on the device. Once the object is recognized, the user can tap on the overlaid logo to show more details

---

[10] http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44230

[11] http://www.qrcode.com/index-e.html

[12] http://studierstube.icg.tu-graz.ac.at/handheld_ar/artoolkitplus.php

of the product (Figure 4(b)) or, if she isn't interested in it, she can simply frame another marker to be recognized. Further, the user can use the action sheet to perform Web 2.0 actions on the object, such as voting, tagging or commenting (Figure 4(c)).

Fig. 3. Logos of some products

Apart from framing a product label with the phone camera, the user can start from a GPS position, a textual search or from a bookmarked or a randomly selected item, figure 4(a) shows the application starting page where the user can select the preferred input modality to begin the exploration of the gastronomic domain. In addition, it is possible to explore the social network of the object (Figure 4(d)) with an innovative interaction model based on the metaphor of a *wheel*. This novel interface makes it possible to visualize the recognized object with some related items around it and the relationships linking them to the object in question (the description of the application interface is not the subject of this paper, for more details see [1, 3]).
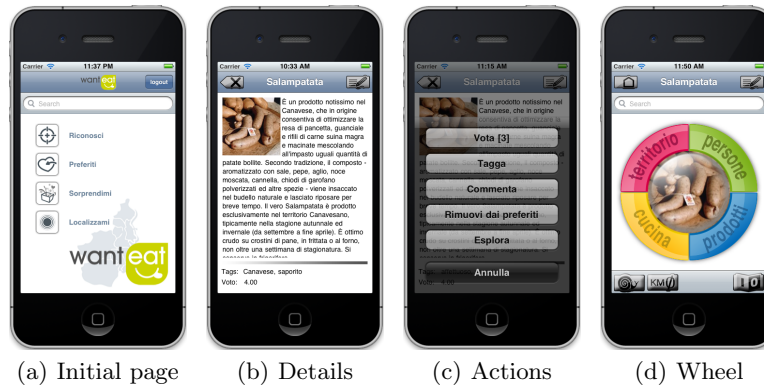

(a) Initial page (b) Details (c) Actions (d) Wheel

Fig. 4. User interface

5 Evaluation

We evaluated our application in the real world setting, during the gastronomical fair "Salone del Gusto 2010"[13] in Torino, Italy. 684 randomly selected subjects used mobile phone cameras to obtain information about the gastronomical products at the event. Each person tried the application for 3-5 minutes and rated its four dimensions: easiness of use, comprehensibility, usefulness and pleasantness, using a 1-4 scale. The results are promising, with all average ratings over 3.

---

[13] http://www.salonedelgusto.it

More specifically, most of the subjects (91.4%) judged the application easy to use, with the average of 3.42. In particular, regarding the interaction in augmented reality mode, 85.1% of the users found it easy and natural to use, with satisfying image recognition speed. 91% of the users found the application comprehensible, with the average of 3.46. Regarding usefulness, 78.6% of users were satisfied with the average of 3.12, This was especially due to their interest in the specific domain and the tools that permit to explore it and to receive the information in mobility. Pleasurable and enjoyable interface with innovative graphical elements had 91.8% of positive opinions with the average of 3.43.

The users, who expressed a negative judgment, complained about not feeling in control of the application, due to the experimental interface different from the standard ones for the chosen device, and suggested to improve intuitiveness of the navigation. On the other hand, the application is easy to use and the users learnt quickly how to use it effitiently.

The users showed great interest in the application. In fact, 76.5% of the users would personally use the application in their daily lives, while 88.2% of the users would recommend the application to friends and relatives. They found especially interesting the possibility to explore the network of related products and people and the relationships among them, discovering new things about the products, the territory they belong to, the people who produce, use and like them.

6   Conclusions

Mobile devices are gaining popularity due to advances in their development: bigger displays, natural interaction models through touch screens, better network capabilities, improved localization features etc. All these components and characteristics foster the implementation of new kinds of context-aware applications which can serve for information retrieval. Many of these applications allow the users to enjoy augmented reality experience through objects in the real world.

Broad gastronomic domain contains many types of domain items with a great number of relations between them, not easily searchable by standard techniques. In this paper we describe a prototype mobile application for reality browsing of gastronomic products and a system architecture that supports the application. We exploit the advantages of augmented reality interaction, employing product recognition using its logo. We present a simple, innovative interaction modality to explore the network of objects and people with the support of a mobile phone. By using our prototype, users are guided in the discovery of the knowledge of a territory in a natural and serendipitous way, staying in contact with surrounding environment and interacting with it directly. In our approach, we use ontologies to represent the domain knowledge in order to enable fluid description and integration of data, as well as reasoning and querying techniques on them. The real world evaluation of our application provided a valuable feedback from users and directions for future development.

In addition, one of possible uses of logo recognition would be to insert data into a storytelling application [6]. Once the product is recognized, users can tell the system about their experience with the product, for example if they tasted

or bought it. This would enable collecting structured contributions from users, which can be later aggregated into simple stories and presented to future users.

References

1. G. Biamino, P. Grillo, I. Lombardi, A. Marcengo, A. Rapp, R. Simeoni, and F. Vernero. "The Wheel": an innovative visual model for interacting with a social web of things. In *Proc. of Visual Interfaces to the Social and Semantic Web, VISSW '11*, 2011.
2. C. Bizer and A. Seaborne. D2RQ - Treating Non-RDF Databases as Virtual RDF Graphs. In *In Proc. of the 3rd International Semantic Web Conference, ISWC '04 (Poster)*, 2004. http://www4.wiwiss.fu-berlin.de/bizer/d2rq/.
3. E. Chiabrando, R. Furnari, P. Grillo, S. Likavec, and I. Lombardi. Dynamic interface reconfiguration based on different ontological relations. In *14th International Conference on Human-Computer Interaction, HCII '11*, volume 6771 of *LNCS*, pages 538–547. Springer-Verlag, 2011.
4. D. Gray, I. Kozintsev, Y. Wu, and H. Haussecker. Wikireality: augmenting reality with community driven websites. In *Proc. of the 2009 IEEE Intl. Conf. on Multimedia and Expo*, ICME'09, pages 1290–1293. IEEE Press, 2009.
5. D. Guinard, V. Trifa, F. Mattern, and E. Wilde. From the Internet of Things to the Web of Things: Resource Oriented Architecture and Best Practices. In *Architecting the Internet of Things*. Springer, 2011.
6. S. Likavec, I. Lombardi, A. Nantiat, C. Picardi, and D. Theseider Dupré. Threading facts into a collective narrative world. In *Proc. of the 3rd Joint Int. Conf. on Interactive Digital Storytelling 2010*, volume 6432 of *LNCS*, pages 86–97. Springer-Verlag, 2010.
7. J. Louvel and T. Boileau. *Restlet in Action*. Manning Publications Co., 2009.
8. M. Megliola, S. Toffolo, M. Zoellner, B. Hodgson, and L. Speller. Intelligent tourism through ubiquitous services: Lessons learnt from two field studies. In *Proc. of eChallenges e-2009*. IIMC, 2009.
9. M. Pielot, N. Henze, C. Nickel, C. Menke, S. Samadi, and S. Boll. Evaluation of camera phone based interaction to access information related to posters. In *Proc. of Mobile Interaction with the Real World - MIRW '08*, pages 61–72, 2008.
10. PIEMONTE Team. WantEat: interacting with social networks of intelligent things and people in the world of enogastronomy. In *Proc. of Interacting with Smart Objects Workshop 2011*, 2011.
11. J. Rekimoto and Y. Ayatsuka. Cybercode: designing augmented reality environments with visual tags. In *Proc. of DARE 2000 on Designing augmented reality environments, DARE '00*, pages 1–10, 2000.
12. L. Richardson and S. Ruby. RESTful web services. In O. Media, editor, *http://www4.wiwiss.fu-berlin.de/bizer/d2rq/*, 2007.
13. M. Rohs and B. Gfeller. Using camera-equipped mobile phones for interacting with real-world objects. In *Advances in Pervasive Computing*, pages 265–271, 2004.
14. G. Takacs, Y. Xiong, R. Grzeszczuk, V. Chandrasekhar, W. chao Chen, K. Pulli, N. Gelfand, T. Bismpigiannis, and B. Girod. Outdoors augmented reality on mobile phone using loxel-based visual feature organization. In *In Proc. of ACM Intl. Conf. on Multimedia Information Retrieval*, pages 427–434, 2008.
15. E. Toye, R. Sharp, A. Madhavapeddy, D. Scott, E. Upton, and A. Blackwell. Interacting with mobile services: an evaluation of camera-phones and visual tags. *Personal and Ubiquitous Computing*, 11(2):97–106, 2007.