Mobile Augmented Reality browsers should allow labeling objects A Position Paper for Mobile AR Summit, MWC 2010

Pedro de-las-Heras-Quirós   Raúl Román-López
Roberto Calvo-Palomino
Universidad Rey Juan Carlos, Madrid, Spain
{pheras,rroman,rocapal}@libresoft.es

February 9, 2010

Abstract

LibreGeoSocial is a new FLOSS (Free/Libre Open Source) mobile social network with a Mobile Augmented Reality (MAR) interface that offers functionality beyond the one available on commercial MAR browsers. The nodes of the social network (people, text notes, audio snippets, pictures,...) are both geolocated and registered at different altitudes. LibreGeoSocial nodes can be shown through a traditional list based view or through a magic lens MAR UI. This interface allows the user not only to browse tags associated with objects of the reality perceived through the camera, but also enables to use the mobile to link media to objects, enabling the labeling of things at different altitudes and situated not necessarily in the vicinity of the user. The system works both outdoors and indoors, in this latest case with location provided by BIDI codes.

We argue that similar functionality to labeling mechanisms available in LibreGeosocial is a requirement for many kinds of MAR applications yet to come, and that it is in fact its absence in current MAR browsers one of the factors that nowadays could hamper the development of this nascent industry.

1 Introduction

Augmented Reality (AR) intends to provide richer experiences by overlying labels or virtual objects over the scene observed through a camera attached to a computer. Many AR systems follow the approach of analyzing in real time the video stream provided by a camera to recognize objects. This way they figure out what virtual objects must be drawn, and where, overlaid to the video stream. These systems need a model of the world to match against the video stream. Advances in the HW have made it possible to perform this computationally expensive image processing in mobile phones [1], although no commercial system has yet appeared, and only controlled demonstrations have been shown.

By using sensors such as accelerometers, GPS and digital compasses available in mobile phones such as the Android devices from HTC and Motorola, or the Apple iPhone 3GS, an alternative approach can be followed to augment the reality perceived through the camera in mobile phones. The information provided by these sensors can be used to know where the mobile is located and towards where it is oriented. Nearby geotagged labels found in the vicinity of the scene can be found in a database of POIs and then shown on the screen on top of the object to which they are associated, without requiring intensive image processing.

It is well known that the A-GPS of these devices can report big offsets from the actual position, but compared to visual AR approaches, it is available everywhere outdoors and does not require a model of the world to match against, being also immune to moving objects that can affect visual approaches. Digital compass can also report significant errors, but it provides good enough orientation information for many applications that users of mobile phones desire to use.

Nowadays social networks are increasingly being used from mobile phones. These have big touch screens and powerful microprocessors, but it is specially the geo-location information provided by GPS, and the orientation information provided by compass and accelerometers, what enables new interesting functionality of mobile social networks that is not available in the desktop. We are developing a FLOSS [1] implementation of a mobile social network called LibreGeoSocial whose client side runs on Android and communicates through a REST API with the backend.

A key feature of LibreGeoSocial is that all nodes of the social graph are stored alongside its position and its altitude. Preexisting layers of geolocated content from other sources such as Panoramio or 11870.com are also integrated through channels, a mechanism similar to Wikitude's overlays or to Layar's layers, but which does not require to push content to the platform; it can pull in real-time the content from the original sources of geolocated data.

[1] Free/Libre Open Source Software, available at http://libregeosocial.morfeo-project.org/

Apart of the normal GUI that users employ to browse and create new geolocated content, LibreGeoSocial can show the contents of the social network through a magic lens [2] interface where labels representing the content of the social network (users, text notes, audio snippets, pictures,...) are drawn over the video stream of the camera, in the 3D position at which they are located.

A key functionality of LibreGeoSocial that is not present in other commercial mobile browsers is the ability to create labels that are linked to specific objects seen through the camera of the mobile.

Next section explains why we consider labeling is a must in mobile augmented reality browsers and how the Open Source nature of LibreGeoSocial could help to advance in this direction. Then section 3 introduces the labeling mechanisms implemented in LibreGeoSocial and section 4 briefly describes additional interesting functionality of LibreGeoSocial, with a final review of related work and conclusions.

2 What MAR labeling functionality similar to Open Source LibreGeoSocial will enable

At first sight it seems unnecessary to defend in 2010 why enabling users to individually upload content to the network is a good thing. The old Web, the new Web2.0 crowdsourcing applications, social networks, Wikipedia, etc. support this point of view. But it seems that in the area of commercial Mobile Augmented Reality things are not so clear. Most experiences provided by Mobile Augmented Reality browsers consider the user as a sink of media and not as a potential producer of it. This is sad from our point of view.

Labeling the real world from mobile phones through a magic lens interface as the one provided by LibreGeoSocial paves the way towards new applications that currently don't exist due to the limited functionality of existing Mobile Augmented Reality browsers in the market.

Tourists uploading media to places, monuments or pictures they are sightseeing, teachers and students uploading media to things of the reality in the context of augmented learning experiences, or players uploading and capturing digital objects in a yet to come wave of MAR games are just a few examples of what can be done with applications that incorporate the labeling functionality. The success of throwdown applications such as MyTown, FourSquare or Gowalla that allow the user to label the places she is visiting and to throw or retrieve virtual things, are just a hint of what's to come. The labeling functionality of LibreGeoSocial enables to throw or retrieve virtual goods at a much finer grain, throwing objects at a window of a building, leaving objects hooked on a monument, etc.

According to iSuppli [3], global shipments of magnetic sensors for electronic compasses is experimenting a huge increase, rising from 8.7 million in 2008 to 80.1 in 2010, with an estimated shipment of 540.7 million units in 2013. Using mobile augmented reality applications not only on smartphones but also on feature phones is thus a certainty more than a possibility. Let's allow them not only to browse content, but also to individually help us nurture the digital commons of Mobile Augmented Reality applications.

It is perfectly known that Internet technology was developed through open standards and open source software. In the current state of development of the Mobile Augmented Reality industry we argue that open standards should be defined as soon as possible in order to establish urgent interoperability measures in the benefit of users. Lessons should be learnt from the recent social networking community, where isolated silos rapidly emerged. Working alongside W3C to create standards should be pursued in order to promote competition among contenders and to avoid trouble to both users and developers.

We think that having available a rich Open Source MAR code base would help everybody, from users to developers to content providers. It is so why LibreGeoSocial code is distributed as Open Source code, with the GNU/GPL 3.0 license, although other licenses could be studied in case it is requested. Either LibreGeoSocial or other similar Open Source frameworks could help to speed up this emergent market as it has already happened so many times in the past decades: Internet protocols, Mosaic, Firefox and Chrome browsers, Apache Web server, Android mobile OS, Linux kernel,...

3 Labeling the world in LibreGeoSocial

The labeling process will not only have to be able to discover the position of the object we want to label, but also its elevation, by using the azimuth and roll data provided by the sensors of the device, and a position provided either by GPS in outdoors or by markers in indoors/outdoors. A critical feature of these devices is that we can only find a single camera in them, what makes it very computationally expensive to estimate distance from it. Thus we chose to estimate based on the information provided by those other sensors.

There are two significant facts that we have to bear in mind before starting the development of these techniques. The first one is that we must calculate the geographic coordinates of the target from the distance between the user an it, so we can come close to this by considering techniques based on different models of the Earth's shape, as Hayford-ellipsoid, WGS84 or ED50. The same models are used to calculate the inverse situation: find the distance from a geographic location. The second one is about the elevation of the targets: we will calculate it from ours, which is looked up from our geographic coordinates in web services based on different models (as SRTM3 or GTOPO30) stored in their databases, and adding the height of the user holding the phone, a parameter that the user can configure in the application, as well as he can configure the floor of the building where it is located. Once the distance and the elevation are known, we can calculate the elevation of the object using the roll information provided by the device and applying trigonometry.

Below we describe the different techniques we have implemented for creating labels, all of them accesible in the LibreGeoSocial mobile social network:

- **Immediate tagging mode**

  The Immediate tagging mode is the fastest and simplest way to attach a label to an object. It assigns the current user's position and elevation to the labeled resource, so it is useful only when the object and the user share the same position.

- **Fast tagging**

  In the Fast tagging mode the user must estimate the distance from the object. He just has to focus in the camera view and move a slide bar to indicate the estimated distance, which is used together with the azimuth, roll, elevation and geo-location of the mobile to estimate where is the object to be labeled.

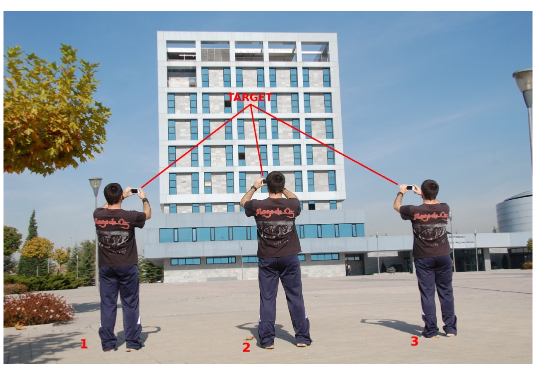
- **Map tagging**

  In this mode, after the user focuses the target, a map is presented to her showing the geographic area where she is located. The user then must identify the object on the map and tap over it. This position is stored as the object's location, and is then used to calculate de distance between the user and the object. Then elevation of the resource can be calculated knowing the roll of the device and its elevation. The map shown to the user has a coloured region which shows the user's sight range as viewed from the camera, rotated according to the user's azimuth to make it easier and more intuitive for him to discover the object in the map.

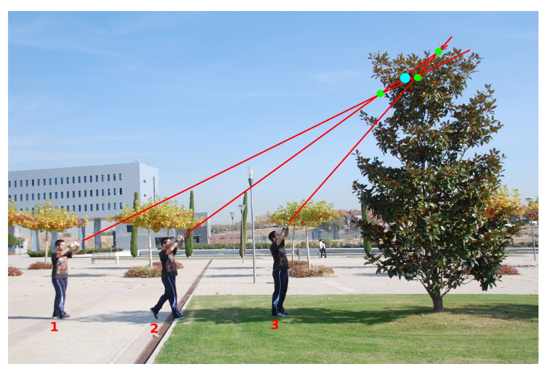
- **Accurate tagging**

  In this case users must focus the target resource from several positions (from 2 upto 10, although 2 is usually enough), saving the position, azimuth and roll of each sample. Then, lines are traced from these positions, that should intersect at the object's position. Of course, the inaccuracy of the sensors makes it almost imposible that they intersect in the same point. We have developed two different accurate modes: accurate side and accurate line.

  Figure 1(a) shows an user taking three samples from different positions while he tries to label a window in a building. The intersection of the lines should be in the window to be labeled. Usually we have errors in these samples due to the sensors of the device, so the intersections will produce more than one point. We calculate the mean of these points to get the real target location.



(a)



(b)

Figure 1: Accurate side (a) and accurate line (b) techniques.

  Figure 1(b) shows another example where the user takes three samples along the line that links him to the object, what gives us three lines. Their intersections produces many points (the green ones) due to sensors errors around the point where the label must be located. Then we calculate the mean of these points (the light blue one) to approximate it.

  After the label position has been obtained it is shown in the screen. In case the user sees that the calculated position is not the one expected, he can repeat the process, perhaps using a differen mode, before proceeding to upload the associated content (audio note, text, picture, etc.)

4 Indoors labeling and other functionality of LibreGeoSocial

The GPS device is a very important element when we want to know our location, but there are some scenarios where we can not use it, specially indoors. The mobile mixed reality code used in LibreGeoSocial supports location in these cases using fiduciary markers with QR code stamps that encode location coordinates and elevation values.

They are easily captured from the application interface using the camera, and then decoded. Once the location of the phone is known, the rest of labeling and visualization techniques of the application work as usual. In places such as museums or malls, markers encoding positions can be sticked to walls. Once a user has found its position through one of them, he can label objects or see preexisting labels attached to objects in the room just by pointing the phone towards the different objects. Right now, users are forced to stay close to a marker in order to visualize or create labels with accuracy because we can only update the position through markers.

LibreGeoSocial provides a fine grain mechanism for privacy enforcing, enabling the user to choose not only which objects, but which attributes of objects (think about the profile of a user of the social network) are available to whom, in a group by group basis, or in a person by person basis.

Thanks to the alarms system implemented in LibreGeoSocial users don't need to use continuously the magic lens interface in order to be aware of nearby interesting information. An alarm is fired when you are near interesting places, objects or multimedia content that you previously specified to the alarm system.

The magic lens interface of LibreGeoSocial has been recently improved to incorporate a mechanism to deal with lots of nearby labels on the screen, allowing to browse through crowded screens with ease. You are encouraged to experiment this functionality either by downloading the application or by attending the MAR showcase at MWC 2010.

As for interconnectivity with other systems, LibreGeoSocial does not require the user to create an new account, being able to login through the Facebook account. The user can also export media information stored in LibreGeoSocial to external systems such as Twitter, FaceBook, etc.

5 Related work

Wither [4] has developed a taxonomy of applications providing annotation in outdoor augmented reality. They detected that there are very few applications that provide either editable or creatable annotations in the context of augmented reality. Only some researchers have used several samples and triangulation [5, 6] or aerial maps[7] to estimate distances in mobile computers as we do. In the context of mobile phones, in 2006 the Nokia's MARA project [8] added external GPS and sensors to a Nokia 6680 mobile device to show labels through a magic lens interface. During 2009 several commercial applications such as Wikitude [9], NearestTube [10] or Layar [11] have demonstrated the use of GPS, accelerometers and digital compass to show through a magic lens interface labels, although most of them don't allow the creation of labels from the mobile as we do. Only the Sekai Camera [12] application running on iPhone allows a limited form of labeling, although it only allows to associate media to a general location, and not to specific objects.

6 Conclusions

The labeling mechanisms described in this position paper have been incorporated to the mobile augmented reality interface of a FLOSS mobile social network called LibreGeoSocial, developed for the Android platform, that not only manages latitude and longitude but also elevation for geolocating contents. Instead of using image processing algorithms we leverage the data measured by the digital compass, GPS and accelerometer sensors included in the Android HTC Magic and Dream devices. By using those sensors we obtained a reasonable solution for both visualizing and creating labels both indoors and outdoors. These techniques can be used today by millions of people on their mobile phones because of the minor HW requirements and because pre-existing models of reality are not needed, ensuring an almost universal applicability in time and space.

We seem to be living a tipping point in the area of mobile mixed reality, with certain similarities with the appearance of the Web: more mature and sophisticated technologies existed in the area of hypermedia, but a good enough solution changed the world even though it was technically not the best solution in multiple aspects.

We hope that the Open Source code of LibreGeoSocial and specially its labeling functionality will help paving the future of new MAR applications either through the integration of this code in other projects, or through the usage of LibreGeoSocial as a reference implementation where to explore new ideas.

References

[1] Gabriel Takacs, Vijay Chandrasekhar, Natasha Gelfand, Yingen Xiong, Wei-Chao Chen, Thanos Bismpigiannis, Radek Grzeszczuk, Kari Pulli, and Bernd Girod. Outdoors augmented reality on mobile phone using loxel-based visual feature organization. In *MIR '08: Proceeding of the 1st ACM international conference on Multimedia information retrieval*, pages 427–434, New York, NY, USA, 2008. ACM.

[2] Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 73–80, New York, NY, USA, 1993. ACM.

[3] isuppli web page, 2010. http://www.isuppli.com.

[4] Jason Wither, Stephen DiVerdi, and Tobias Höllerer. Annotation in outdoor augmented reality. *Computers and Graphics*, 2009.

[5] Yohan Baillot, Dennis Brown, and Simon Julier. Authoring of physical models using mobile computers. In *ISWC '01: Proceedings of the 5th IEEE International Symposium on Wearable Computers*, page 39, Washington, DC, USA, 2001. IEEE Computer Society.

[6] Wayne Piekarski and Bruce H. Thomas. Interactive augmented reality techniques for construction at a distance of 3d geometry. In *EGVE '03: Proceedings of the workshop on Virtual environments 2003*, pages 19–28, New York, NY, USA, 2003. ACM.

[7] Jason Wither, Stephen Diverdi, and Tobias Hollerer. Using aerial photographs for improved mobile ar annotation. In *ISMAR '06: Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 159–162, Washington, DC, USA, 2006. IEEE Computer Society.

[8] Kate Greene. Hyperlinking reality via phones. *Technology Review*, November/December 2006.

[9] Wikitude web page, 2009. http://www.wikitude.org.

[10] Acrossair web page, 2009. http://www.acrossair.com/apps_nearesttube.htm.

[11] Layar web page, 2009. http://layar.com.

[12] Sekai web page, 2009. http://www.tonchidot.com/Sekai_Camera.html.