# Mobile Visual Recognition, the future of Mobile AR

David Marimon, Tomasz Adamek, Kerstin Göllner and Carlos Domingo, Telefonica R&D

## 0. Introduction

Location-based services for mobile devices started to grow as GPS antennas became commonly embedded in smartphones. A similar effect has occurred with Mobile Augmented Reality with the digital compass.

The year 2009 will be remembered by many as the year when all blogs talked about Augmented Reality. AR consists in the overlay of information such as text or images on top of the reality, usually captured by the camera. Having known the usual *wow* effect of AR for a long time, the current wide-world spread of AR comes as no surprise,... but what's next?

A large number of companies in the most heterogeneous markets have raised interest in AR as a path to innovation. The question for many is whether AR is full of smoke or there is real ground for commercial products and services.

## 1. The evolution of Augmented Reality

From a technical point of view, Augmented Reality depends on sensors to capture the relation between the user and the environment. Those sensors span from large magnetic trackers, to mobile-embedded GPS antennas.

AR started using complex tracking systems probably due to its heritage from Virtual Reality. However, the tendency went quickly to leverage the use of special equipment and exploit the benefits of the core sensor: a video camera. Computer Vision techniques started to play a central role in the evolution of AR. Firstly, marker-based approaches introduced a great leap for AR applications and interaction prototypes. The first tracking systems that have appeared on mobile devices use marker-based tracking due to the low computational complexity of those approaches. The AR research community has however moved its interest to another area of computer vision which is marker-less tracking. During the last few years, webcam-based tracking systems have been developed with great success. In the past two years, we have also seen lab prototypes with mobile devices tracking the pose of natural objects or scenes.

But, the real explosion of Mobile AR, the one that has gone into all those blogs is completely disconnected from such ground-breaking academic research. It is solely based on GPS antennas and digital compass embedded into mobile phones. Those two sensors are enough to determine the rough location and point of view of the user; enough to overlay information about the environment onto the image captured by the camera. The question is whether it is really enough or we should be looking into that cutting edge research community.

## 2. The user experience or why GPS and compass is not enough

The unarguable success of certain mobile devices (and failure of others) comes from the user experience. On top of that, people are getting used to quick access to relevant information. The sum of both obviously determines those applications that have an impact on the mobile market.

User experience is usually fulfilled in the first time somebody uses a Mobile AR browser. Once this *wow* effect is passed, one discovers the other side of the coin. One can be sitting at home,

seeing the walls of the room and, overlaid, a picture of the major touristic attraction of the city. One could ask "How is that linked to my room and relevant to the wall my mobile is pointing at?". Some might argue that it is valuable to know in what direction and distance that attraction is. In general it is, but what is the real benefit as compared to a map?

We covered a rather controversial point, the utility of browsing content unrelated to the view. Let us now cover another aspect, in this case, a fact. GPS antennas have an error of several meters (depending on many factors, such as weather or sky visibility). Such error introduced into the point of view of the user is non-relevant for distant points of interest (POI). What happens with closely located places is another story. An error of 2 or 3 meters is enough for the display to show information about a POI behind us, just in the middle of the screen (i.e. as if it was in front of us). Imagine we are looking for a review of a restaurant that is in front of us, the question is: should we be seeing the review of another restaurant one block away? Here is where the user experience breaks.

## 3. Mobile Visual Recognition: the door to relevant information

On a separate track but naturally linked to AR, the computer vision community has been working on visual recognition for many years now. This technology enables to identify an object by taking a picture of it and querying a database of reference images. The step from academic research to commercial products has in this case been much shorter. There are several mobile applications that perform visual recognition of objects such as posters, cd/dvd covers or books and which can be perfectly used for commercial purposes, e.g. for advertising.

Mobile Visual Search applications provide a natural interaction with the environment. Firstly, because the information or services provided are specific to THAT object that we are interested in; secondly, because it skips the sometimes cumbersome step of typing text on a mobile device.

## 4. The future of MobileAR

The market for Mobile AR services is estimated to reach $732 million by 2014. Revenues will derive from a combination of paid app downloads, subscription-based services and especially advertising, where Mobile AR has started to become a hot topic recently.

Once the door of identifying what a user is seeing is opened, the opportunities to enhance the mobile experience and services are almost limitless. From the perspective of the user: relevant services in the display, just by aiming with the camera. From the perspective of a provider: information, offers, and reviews, all this can be provided right on the spot.

Another advantage of adding visual recognition into Mobile AR applications is that both location-based and non-location based AR is possible. On one hand, GPS antennas can be used to filter-out information that is surely unavailable to the user according to her location. On the other, a user can access services linked to an object just next to her regardless of her location.

## 5. Conclusion

We strongly believe that the future of Mobile AR goes naturally through Mobile Visual Recognition. New services will be able to provide an enhanced mobile view of the reality, increasing the users´ engagement and opening up new revenue sources for app developers, advertisers, publishers and the likes. Identifying what the user is seeing is the key factor for an advanced Mobile AR user experience and for the successful commercialization of the technology.