Landmark recognition in VISITO: VIsual Support to Interactive TOurism in Tuscany

Giuseppe Amato ISTI-CNR
via G.Moruzzi, 1
Pisa, Italy
g.amato@isti.cnr.it

Paolo Bolettieri ISTI-CNR
via G.Moruzzi, 1
Pisa, Italy
p.bolettieri@isti.cnr.it

Fabrizio Falchi ISTI-CNR
via G.Moruzzi, 1
Pisa, Italy
f.falchi@isti.cnr.it

ABSTRACT

We present the VIsual Support to Interactive TOurism in Tuscany (VISITO Tuscany) project which offers an interactive guide for tourists visiting cities of art accessible via smartphones. The peculiarity of the system is that user interaction is mainly obtained by the use of images – In order to receive information on a particular monument users just have to take a picture of it. VISITO Tuscany, using techniques of image analysis and content recognition, automatically recognize the photographed monuments and pertinent information is displayed to the user. In this paper we illustrate how the use of landmarks recognition from mobile devices can provide the tourist with relevant and customized information about various type of objects in cities of art.

Categories and Subject Descriptors

H3.1 [**Information Storage and Retrievals**]: Content Analysis and Indexing; H3.5 [**Information Systems**]: Online Information Services—*Commercial services*

General Terms

Experimentation, Algorithms

Keywords

landmarks recognition, image classification, interactive tourism

1. INTRODUCTION

In the last few years, the problem of recognizing landmarks has received growing attention by the research community. As an example, Google presented its approach to building a web-scale landmark recognition engine [6] that was also used to implement the Google Goggles service [5].

VISITO Tuscany (VIsual Support to Interactive TOurism in Tuscany[1]) also aims at addressing this interesting issue and investigates and realizes technologies able to offer an interactive and customized advanced tour guide service to visit the cities of art in Tuscany. More specifically, it focuses on offering services to be used (see Figure 2):

[1] http://www.visito-tuscany.it/

Figure 1: Tourist information on a smartphone

During the tour – through the use of mobile devices of new generation, in order to improve the quality of the experience. As shown in Figure 1, the mobile device is used by the user to get detailed information about what he's watching, or about the context he's placed in. While taking pictures of monuments, places and other close-up objects, the user points out what, according to him, seems to be more interesting. When a picture is taken it is processed by the system to infer which are the user's interests and to provide him relevant and customized information. For example, if a user takes a picture of the bell tower of Giotto, he can get detailed information describing the bell tower, its structural techniques, etc.

Before the tour – to plan the visit in a better way. Both the information sent by other users and their experiences, can be employed by the user to better plan his own visit, together with the information already included in the database system and, more generally, on the web. The interaction will take place through advanced methods based on 3D graphics.

After the tour – to keep the memory alive and share it with other people. The user can access the pictures and the itinerary he followed through advanced mode of interaction based on 3D graphics. Moreover, he might share his information and experiences with other users by creating social networks.

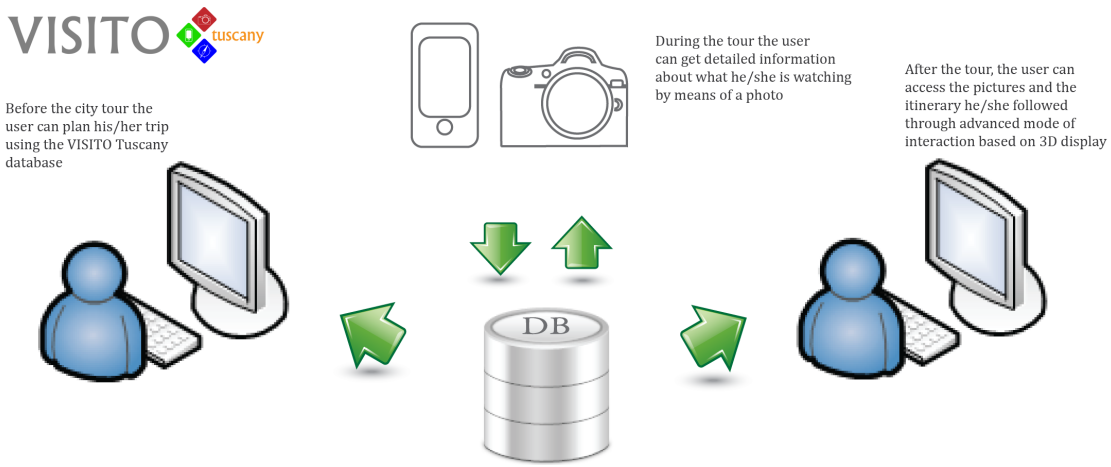

Figure 2: The VISITO Tuscany project services.

Even if the general objective of the VISITO Tuscany project is broader, in this demonstration we will mainly focus on the use of the Smartphone to obtain information during a visit in an tourist place. The user can obtain information on monuments by simply pointing the landmark of interest with the embedded camera and taking a picture. The acquired image is analyzed and the landmark recognized so that the user can be provided with related information.

The demonstrated system is composed of three main components: a client application that runs on a mobile phone, an image classifier that recognizes landmarks contained in pictures, and a digital library containing descriptions of various monuments. At the moment of writing, we have created recognizers for hundreds of monuments in three cities in Tuscany: Florence, Pisa, and San Gimignano. For these monuments we also populated the digital library with descriptions consisting of text and images that can be easily read from a mobile device. When the user takes a picture of a monument, the picture is first sent to the classifier that checks if one of the available monuments is recognized. In case a monument is recognized, the description is retrieved from the digital library and sent back to the mobile device.

Landmark recognition is performed using local features and kNN based classification algorithms. We defined a new approach that relies on a revision of the single label kNN classification algorithmn. More specifically, as better discussed in [1, 2], we propose an algorithm that first assigns a label to each local feature of an image query. The label of the image is then assigned on the basis of the labels and confidences assigned to its local features. In other words, our kNN approach is based on the similarity among the local features of the query image and the ones in the training set rather than similarity among whole images. Even if we do not rely on an Image-to-Class distance, our approach is similar to the one described in [3]. Moreover, for bag of words approaches, the importance of considering relations between local features belonging to different images of the same class, has been recently studied in [4] where visual synonyms are considered for landmark image retrieval.

Acknowledgments

This work was partially supported by the VISITO Tuscany project, funded by Regione Toscana, in the POR FESR 2007-2013 program, action line 1.1.d. VISITO Tuscany is coordinated by ISTI-CNR. Its consortium includes three ISTI-CNR laboratories (Networked Multimedia Information Systems, Visual Computing, High Performance Computing), the security laboratory of IIT-CNR, and three private companies: Alinari24Ore, Hyperborea, and 3Logic MK.

We also thank the municipalities of Florence, Pisa, and San Gimignano that provided us with all needed permissions to build the demonstrator.

2. REFERENCES

[1] G. Amato and F. Falchi. kNN based image classification relying on local feature similarity. In *SISAP '10: Proceedings of the Third International Conference on SImilarity Search and APplications*, pages 101–108, New York, NY, USA, 2010. ACM.

[2] G. Amato and F. Falchi. Local feature based image similarity functions for kNN classfication. In *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence (ICAART 2011)*, 2011. to appear.

[3] O. Boiman, E. Shechtman, and M. Irani. In defense of nearest-neighbor based image classification. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR08)*, pages 1–8, 2008.

[4] E. Gavves and C. G. M. Snoek. Landmark image retrieval using visual synonyms. In *ACM International Conference on Multimedia*, 2010.

[5] Google. Goggles. `http://www.google.com/mobile/goggles/`, 2011. last accessed on 3-March-2011.

[6] Y. Zheng, M. Z. 0003, Y. Song, H. Adam, U. Buddemeier, A. Bissacco, F. Brucher, T.-S. Chua, and H. Neven. Tour the world: Building a web-scale landmark recognition engine. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, pages 1085–1092, 2009.