A MOBILE MAPPING DATA WAREHOUSE FOR EMERGING MOBILE VISION SERVICES

Linde Vande Velde[1*], Lucas Paletta[2], Katrin Amlacher[2], Patrick Luley[2], Alexander Almer[2], Dušan Omerčević[3], Roland Perko[3], and Aleš Leonardis[3]

[1]Tele Atlas, Gent, Belgium
[2]Institute of Digital Image Processing, Joanneum Research, Graz, Austria
[3]Visual Cognitive Systems Laboratory, University of Ljubljana, Slovenia

*Moutstraat 132, 9000 Gent, Belgium, phone +32/9 244 89 63
E-mail: linde.vandevelde@teleatlas.com

ABSTRACT

Mobile vision services are a type of mobile ITS applications that emerge with increased miniaturization of sensor and computing devices, such as in camera equipped mobile phones, and appropriate adoption of advanced computer vision methodologies. The interpretation of mobile imagery with reference to geo-coded information from augmented city maps offers new applications. In this paper we describe a central component in mobile geo-service application, the Mobile Mapping Data Warehouse (MOMA) as developed by Tele Atlas, and associated prototypical mobile vision services that make use of MOMA features and cognitive vision technology, such as, (i) vision based geo-indexed object recognition, (ii) image based localization, and (iii) hyperlinking reality via camera phones. We demonstrate the applicability of the state-of-the-art mobile vision technology and outline a roadmap towards the realization of future mobile ITS applications.

KEYWORDS

Mobile interfaces, digital city maps, geo-services, cognitive vision, mobile mapping.

A general framework about the requirements of geo-services that would use geo-spatial information for the support of mobile services, in particular, vision based functionalities that enable natural interfaces for situated annotation in a local urban environment, has been proposed in [1]. Geo-coded imagery, event and context information are stored as part of city maps in order to make the maps intelligent and to exploit further context information of objects. In the EU funded project MOBVIS [2], we develop advanced methodologies to aid mobile vision and context awareness tasks using the enhanced Tele Atlas 2D and 3D City Maps. Augmenting the maps with visual features, semantic and context information enables to support within a given context specifically targeted geo-services in order to deliver appropriate geo-referenced map information on demand.

Mobile vision services are an emerging technology within on-going market trends: Currently, camera phones are entering the rapid growth stage and they will soon be the most common image capture device in the world. Market research estimated that over 175 million camera phones were shipped in 2004, and that by the end of the decade there will be a global population of over one billion mobile imaging handsets - more than double the number of digital still cameras. Furthermore, in 2010 revenues from location-based services will reach € 622 million and account for 1.8 percent of the non-voice services. These figures demonstrate that a systematic way for the combination of mobile imagery and geo-coded information will provide a high impact on the future markets.

This work outlines the development of a central system component for mobile vision services, i.e., the Mobile Mapping Data Warehouse (MOMA) that provides geo-coded reference information and interfaces to geo-services in interaction with advanced computer vision (cognitive vision) technology. The MOMA consists of databases for raw mobile mapping data, image and position metadata, such as object identity, category, pose information and services to select arbitrary data via user interfaces and specifically tuned geo-service components.

Object Awareness provides a concept for semantic indexing into huge information spaces where standard approaches suffer from the high complexity in the search processing otherwise, thus providing a means to relate the mobile agent to a semantic aspect of the environment. Visual object recognition using innovative and robust pattern recognition methodologies is an emerging technology to be applied in mobile computing services [3]. Due to the many degrees of freedom in visual object recognition, it is highly mandatory to use a geo-context to focus object recognition on a specific set of object hypotheses. In this work, we describe the improved performance in geo-indexed object recognition [4] and specifically how this methodology interfaces to geo-referenced data using MOMA services.

MOBILE MAPPING IMAGERY AND ITS APPLICATIONS

A constituent part of the mobile mapping data warehouse is mobile mapping imagery (see Figure 1.a). The mobile mapping imagery is a set of images with known geographic position and camera orientation collected by a mobile mapping van (see Figure 1.b) during the process of map making.

(a) (b)

Figure 1. An example of mobile mapping imagery (a) and the Tele Atlas mobile mapping van (b).

While traditionally mobile mapping imagery has been used only for map-making, it has recently been put to some new uses. The most prominent new use for mobile mapping imagery is Google's *Streetview* [5], where users are given the option of browsing mobile mapping imagery themselves. The street level gives users a completely novel map experience by immersing them in a specific micro-location where they can explore the scene on a level of detail never before provided to users by a map.

In this paper we present three applications that employ computer vision techniques to extend the use of mobile mapping imagery beyond simple visualization. In the first application, a mobile image is used to return annotation about the most prominent point of interest found in the visual information. In the second application, imagery is used to enable localization of the user based on a single snapshot, while in the third application, the user's photo is augmented with hyperlinks to additional information about buildings, monuments, logos and other objects depicted on the photo.

IMAGE BASED GEO-INDEXED OBJECT RECOGNITION

Image based recognition of points of interest provides an intuitive interface to directly access annotation in urban scenes. While location based services have already been presented using GPS based indexing into, e.g., tourist databases alone, these kind of mobile services still require a considerable overhead in terms of interaction with the mobile device and result in lack of intuition in the access of relevant information. Purely image based recognition [1,3] only requires to direct the mobile camera to the object of interest, capture an image and receive annotation in return, therefore it represents to the mobile user an intuitive choice using visual information. The methodology behind these services is to match the visual features that were extracted from the captured image with features that were stored from reference images. Reference images can be retrieved from the mobile mapping data warehouse described in the previous Section.

A major issue for the performance of these services is uncertainty in visual information; covering large urban areas with naive approaches would require to refer to a huge amount of reference images and consequently to highly ambiguous visual features. Previous work on mobile vision services primarily advanced the state-of-the-art in computer vision methodology for the application in urban scenarios. However, so far it has not been considered to investigate the contribution of geo-information to the performance of the vision service. In [4] we proposed to exploit contextual information from geo-services with the purpose to cut down the visual search space into a subset of all available object hypotheses in the large urban area. Geo-information in association with visual features enables to restrict the search within a local context. The results from experimental tracks and image captures in an urban scenario prove a significant increase in recognition accuracy about more than 10% [4] and use of computational resources when using in contrast to omitting geo-contextual information.

Figure 2 presents the intuitive interaction with the mobile device (a,b) and a trajectory of user positions (c) during an evaluation session in the Inner City of Graz, Austria. Note that due to the urban canyon effect the GPS based position signal may deviate from the ground truth, however, this effect is irrelevant for the indexing into image databases and demonstrates that image based recognition is an important asset for mobile location based services. Other mobile vision services can be defined in a similar manner, such as mobile advertising (d) when capturing images from shops including brand information and retrieving annotation about, e.g., related offers and information about new products that are available at the shop.

The mobile mapping data warehouse is an important prerequisite for mobile vision applications, providing the mandatory information about geo-referenced images about the urban environment and offering opportunity to associate the visual information with related map features, such as points of interest, marketing and tourist information.

(a) (b) (c)

(d)

Figure 2. Image based recognition of urban points of interest is initiated by an image capture about the object of interest (a) and returns annotation about tourist sights (b). User positions are tracked via GPS (c) and index into relevant map information. With similar methodology (d), mobile interfaces inform about related map information.

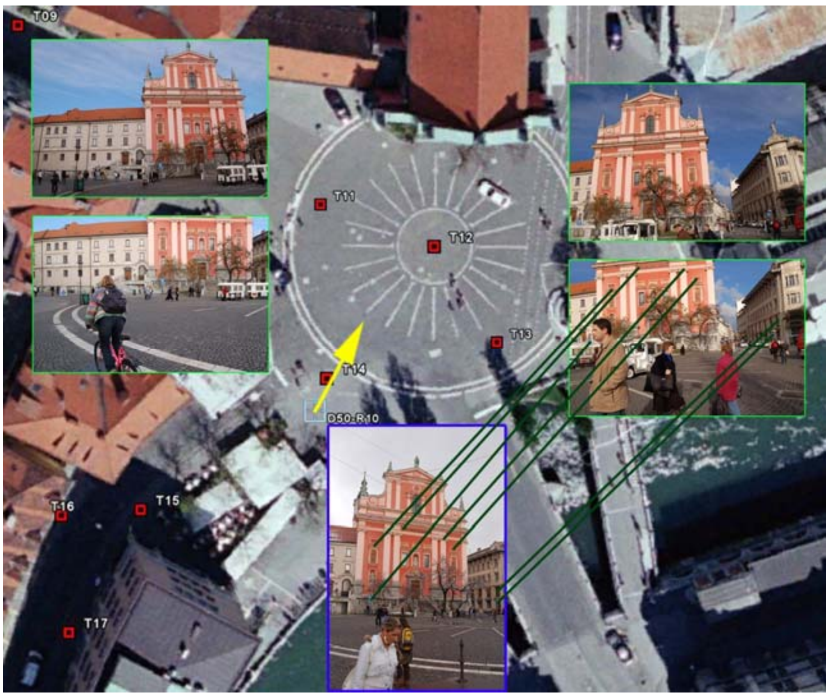

Figure 3. The user position and camera orientation (yellow arrow) is determined by relating the user image (blue frame) and reference images (green frames) to each other.

IMAGE BASED LOCALIZATION

Localization for navigation is usually associated with satellite based services, e.g. GPS, but there exist several other localization techniques such as positioning based on WiFi [6] and GSM-based positioning [7]. Recent advances in computer vision [8, 9, 10, 11] have enabled another alternative to satellite based positioning that uses positioning relative to a set of images with known geographic position and camera orientation. In our case we have used the mobile mapping imagery as reference images.

Image-based localization (Figure 3) starts by identifying corresponding image regions (dark green lines in Figure 3) in the user and reference images. Then, geometry relating the user and the reference images is estimated. Finally, the user's position and orientation is triangulated relative to the (known) position and orientation of the reference images. Please refer to [12] for further details about our approach to image-based localization and a quantitative comparison of several localization techniques that has shown that computer vision enables localization accuracies comparable to GPS. The image-based localization can therefore augment the GPS, or in some circumstances where GPS performs poorly, even replace it.

(a) (b)

Figure 4. To access information about objects in the surrounding, user just snaps a photo (a) and objects on the image become hyperlinks to information (b) that the user can access by simply tapping an icon.

HYPERLINKING REALITY VIA CAMERA PHONES

A further application that employs computer vision to extend the use of mobile mapping imagery is “Hyperlinking Reality via Camera Phones”. The objective of this application is to enable a novel interface for mobile devices. Instead of typing keywords on the small and clumsy keypad of a mobile device, the user just snaps a photo of his surroundings and objects on the image become hyperlinks to information (see Figure 4 for an example). This second application is a natural extension of the image-based localization application presented in the previous paragraph. The establishment of relations between the query image and the mobile mapping imagery enables a transfer of information from the mobile mapping imagery to the query image. The information that is annotated on the mobile mapping imagery are the location of interesting buildings, logos, banners and other objects of interest, thus enabling a natural interface for exploring an urban environment for the pedestrian user.

CONCLUSION

Mobile mapping data warehouses will provide geo-coded infrastructure information together with interfaces to access and select appropriate data for emerging mobile vision services. This infrastructure is mandatory to enable the application of advanced computer vision based functionalities, such as geo-indexed object recognition, vision based positioning and hyperlinking of reality as presented in this paper. These functionalities are prerequisite for numerous mobile services that require the matching of captured with stored geo-referenced imagery and the association of visual information with geo-information databases as outlined in this paper.

ACKNOWLEDGMENTS

This work is supported in part by the European Commission funded project MOBVIS under grant number FP6-511051 and by the FWF Austrian National Research Network on Cognitive Vision under sub-project S9104-N04.

REFERENCES

[1] Vande Velde, L., Luley, P., Almer, A., Seifert, C., and Paletta, L., Intelligent Maps for Vision Enhanced Mobile Interfaces in Urban Scenarios, *Proc. 12th World Congress on Intelligent Transportation Systems and Services*, ITS 2006, London, UK, October, 2006.

[2] URL http://www.mobvis.org.

[3] Fritz, G., Seifert, C., and Paletta, L., A Mobile Vision System for Urban Object Detection with Informative Local Descriptors, *Proc. 4th IEEE International Conference on Computer Vision Systems*, ICVS 2006, New York, NY, USA, January, 2006.

[4] Amlacher, K., Fritz, G., Luley, P., Almer, A., and Paletta, L., Mobile Object Recognition Using Multi-Sensor Information Fusion in Urban Environments, *Proc. IEEE International Conference on Image Processing*, ICIP 2008, San Diego, CA, October, 2008, in press.

[5] Vincent, L., Taking Online Maps Down to Street Level, *Computer*, 40(12):118-120, 2007.

[6] Cheng, Y.-C., Chawathe, Y., LaMarca, A., and Krumm, J.: Accuracy Characterization for Metropolitan-Scale Wi-fi Localization, *Proc. 3rd International Conference on Mobile Systems, Applications, and Services*, MobiSys 2005, Pages: 233 – 245, Seattle, WA, USA, June, 2005.

[7] Varshavsky, A., Chen, M.Y., de Lara, E., Froehlich, J., Haehnel, D., Hightower, J., LaMarca, A., Potter, F., Sohn, T., Tang, K., Smith, I., Are GSM Phones the Solution for Localization?, *Proc. IEEE Workshop on Mobile Computing Systems and Applications*, WMCSA 2006, Semiahmoo Resort, Washington, USA, April, 2006.

[8] Schmid, C., and Mohr, R. Local Grayvalue Invariants for Image Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5):530–535, May 1997.

[9] Lowe, D.G., Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision,* 60(2) (2004) 91–110.

[10] Omerčević, D., Drbohlav, O., and Leonardis, A., High-Dimensional Feature Matching: Employing the Concept of Meaningful Nearest Neighbors, *Proc. IEEE 11th International Conference on Computer Vision*, ICCV 2007, pages 1-8, Rio de Janeiro, Brazil, October, 2007.

[11] Zhang, W. and Košecka, J., Image based localization in urban environments, *Proc. International Symposium on 3D Data Processing Visualization and Transmission*, 3DPVT 2006, University of North Carolina, Chapel Hill, USA, June, 2006.

[12] Steinhoff, U., Omerčević, D., Perko, R., Schiele, B., and Leonardis, A., How Computer Vision Can Help in Outdoor Positioning, *Proc. European Conference on Ambient Intelligence* 2007, AmI 2007, pages 124-141, Darmstadt, Germany, November, 2007.