Augmenting Images with Context on Mobile Devices

Harlan Hile[†‡] Radek Grzeszczuk[‡] Ramakrishna Vedantham[‡]
Jana Košecka[§‡] Alan Liu[†‡] Gaetano Borriello[†]

[†] University of Washington
Dept. of Computer Science and Engineering
{harlan,aliu,gaetano}@cs.washington.edu

[‡] Nokia Research Center Palo Alto
radek.grzeszczuk@nokia.com
ramakrishna.vedantham@nokia.com

[§] George Mason University
Dept. of Computer Science
kosecka@cs.gmu.edu

Mobile phones are an attractive platform that include image capture, location sensing, increasing processing power, and network connectivity. This set of features is ideal for navigation and other location-based services. We present a system to provide additional contextual information to images displayed on these mobile devices. By utilizing reconstruction and alignment techniques, we can produce highly accurate georegistered camera poses for existing sets of images. This pose information can then be used to accurately render new details onto the images. In the area of navigation, we use this set of images to automatically generate landmark-based navigation instructions, providing additional context to the images in the form of directional arrows, surrounding map information, and building annotations. Our system also enables live registration of views acquired by a camera phone, allowing users to conceptually point at and click on a building in their environment through the window of their phone's camera. We demonstrate a light-weight mobile client that can utilize a touch-based user interface for easy access to all of the functionality supported by our context-rich navigation system.

This work makes the following contributions to the area of augmenting images with contextual information:

• We present an approach to automatically compute camera poses from unstructured datasets and to correct poor GPS readings by leveraging on algorithms for robust image-based matching and 3D structure and motion computation. Our system automatically reconstructs camera pose by using the Photo Tourism system [2]. We extend this reconstruction by automatically aligning the reconstructed geometry to the world using (possibly noisy) image tags containing GPS data. We can further refine alignment by fitting the reconstructed points to building outlines available from Geographic Information Systems (GIS).

• This accurate camera pose allows us to automatically make intelligent selections of images to use in a landmark-based navigation plan. The poses are also used to render new items into the images, such as navigational arrows to highlight the path to follow [1]. Images may also be synthesized for new views by warping an existing image to a plane as seen from a different location. This generates understandable directions in more locations with a smaller database of images.

• By using available building extent information and the calculated camera poses, we can highlight portions of the image corresponding to the landmark. This allows us to clearly label the landmark referenced in navigation instructions when there may be multiple landmarks visible. It also allows us to make the image interactive, giving the user the ability to select a building and open a URL or other information, allowing landmarks to be hyperlinked.

Figure 1: Light-weight mobile client that can display navigational instructions as well as additional context by labelling landmarks and allowing them to be hyperlinked.

• We support automatic live augmentation of images taken by the user. Given a new image, we can quickly calculate its camera pose and then render augmentations into the image. Instead of using a system based on matching to 3D structure similar to Photo Tourism and other previous work [2], we propose a new method based on image-to-image matching that can work with a system designed to run in real-time on a camera phone [3].

• We built a prototype mobile client for both a mobile phone and a mobile tablet with a touch-based user interface, as shown in Figure 1. The client gives easy access to all of the functionality supported by our context-rich navigation system, including selectable image annotations for easy access of information, and live matching using the built-in camera.

1. REFERENCES

[1] H. Hile, R. Grzeszczuk, A. Liu, R. Vedantham, J. Košecka, and G. Borriello. Landmark-Based Pedestrian Navigation with Enhanced Spatial Reasoning. In *7th Int. Conf. on Pervasive Computing (to appear)*. Springer, 2009.

[2] N. Snavely, S. M. Seitz, and R. Szeliski. Photo Tourism: Exploring Photo Collections in 3D. In *SIGGRAPH Conference Proceedings*, pages 835–846. ACM Press, 2006.

[3] G. Takacs, V. Chandrasekhar, N. Gelfand, Y. Xiong, W-C. Chen, T. Bismpigiannis, R. Grzeszczuk, K. Pulli, and B. Girod. Outdoors Augmented Reality on Mobile Phone using Loxel-Based Visual Feature Organization. *ACM International Conference on Multimedia Information Retrieval (MIR'08)*, 2008.