A Hybrid Toolkit for Urban Augmented Reality

Sam Kronick
University of California San Diego
skronick@ucsd.edu

Abstract

The goal of this project is to develop a platform for mobile augmented reality (AR) applications that enable a user to explore and manipulate the image of an urban environment in real time. Most vision-based AR tools that exist today (including ARToolkit and Qualcomm's AR SDK) assume that the objects to be tracked are not stationary and therefore no external frame of reference is used to guide object recognition and tracking. If, however, the objects to recognize are facades of buildings and signs in a city, we can use sensor data (GPS/compass/etc) to roughly estimate our position and pose (as is done with Layar) and use vision techniques to more precisely align a 3D data overlay.

1. Qualifications

My academic background is in art and architecture, but technology has been a consistent theme in my work. I have extensive experience programming in Java and, more recently, Objective-C development for iOS along with some C/C++. I have done some work with with sensor fusion (the Kalman filter in particular) in the context of navigation systems for robots. Moreover, I have been working on this project specifically for the past five months, becoming familiar with OpenCV and marker and markerless AR systems. My approach to this project will likely be much more pragmatic than theoretical, with a focus on designing a toolkit or system usable by other artists/designers who are a part of the "creative code" community like myself.

2. Milestones

2.1. Ferns-based Feature Tracking and Pose Estimation on iOS Platform - Week 2

- Build OpenCV for iOS platform
- Customize the find_obj_ferns.cpp example for this platform
- Convert homography into OpenGL transformation matrix for simple 3D rendering
- Basic memory/processing optimizations to achieve an acceptable frame rate (15fps?)
- This stage is already well underway.

2.2. Client-Server model for retrieving local Ferns based on geolocation - Week 4

- Load potential features (building facades and/or street signs) to track based on a combination of GPS and compass data.
- Offload the process of finding the best features to the server.

2.3. Refined pose estimation with sensor fusion - Week 6

- Combine GPS/compass/accelerometer/gyro readings with vision-based pose estimates using an Extended Kalman Filter

2.4. User interface for data input - Week 8

- Develop a basic UI for users to geotag and insert new building/sign imagery into the database
- Develop a basic UI for users to add new augmented images/annotations onto the real world from the mobile device

2.5. Toolkit beta version packaging - Week 10

- Clean up and document the core classes
- Package the most stable code as an open source toolkit for release to the public

3. Questions

- What should the architecture of a hybrid vision- and sensor-based augmented reality system look like if designed for a modern

mobile platform? What parameters need to be tuned and optimizations need to be used given the limited CPU and RAM on these devices?
- At what stage of the AR pipeline should fusion of vision- and sensor-based pose estimation occur?
- What can these tools, packaged as a unified toolkit, contribute to the fields of art and design-- the "creative code" community?

4. Software and Hardware Resources

This project will use the iOS frameworks and the Objective-C language to target the Apple iPad 2 and iPhone 4 platforms. These frameworks provide high-level access to mobile device sensor data and the device's camera.

Vision processing will use the OpenCV libraries, including the Ferns-based classifier and matcher. OpenCV also may be used for its Kalman filter implementations.

Rendering of 3D overlays will use the Ogre3D rendering engine.

Sample urban imagery may be taken from Google Streetview, although some local, real-world photos will be necessary for testing.

5. Relevant Papers

[1] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, Fast Keypoint Recognition using Random Ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.

[1] M. Calonder, V. Lepetit, and P. Fua, Compact signatures for high-speed interest point description and matching. *IEEE 12th International Conference on Computer Vision*, 2009.

[2] D. Wagner, G. Reitmayr, et al, Pose tracking from natural features on mobile phones. *ISMAR '08 Proceedings of the 7th IEEE/ ACM International Symposium on Mixed and Augmented Reality,* 2008.

[3] G. Takacs, V. Chandrashekar, et al, Outdoors augmented reality on mobile phone using loxel-based visual feature organization. *MIR '08 Proceeding of the 1st ACM international conference on multimedia information retrieval,* 2008.