Geometry and Texture from Thousands of Images

Recovering three-dimensional information from two-dimensional images is the fundamental goal of stereo techniques. The problem of recovering depth (three-dimensional information) from a set of images is essentially the correspondence problem: *Given a point in one image, find the corresponding point in each of the other images*. Finding potential correspondences usually involves matching some image property. If the images are from nearby positions, they will vary only slightly, simplifying the matching process.

Once a correspondence is known, solving for the depth is simply a matter of geometry. Real images are composed of noisy, discrete samples, therefore the calculated depth will contain error. This error is a function of the baseline or distance between the images. Longer baselines result in more precise depths. This leads to a conflict: short baselines simplify the matching process but produce imprecise results; long baselines produce precise results but complicate the matching process.

In this talk, I will present a method for recovering both geometry and texture from large sets (1000's) of pose images (images annotated with absolute position and orientation) acquired from arbitrary locations within the scene. By using sets of images we get the benefits of both long and short baselines. I will also discuss techniques which allow us to use sets of images which 1) contain significant calibration errors, 2) are captured under widely varying illumination conditions, and 3) may include occlusion. I will also present the results of applying these techniques to a large dataset acquired in an urban environment.