Incremental 3D Model Generation Using Depth Cameras

Abstract
The ability to directly acquire 3D scene information from a single viewpoint could have a large impact on many computer vision applications. Depth cameras are a new sensing device facilitating such data acquisition, but are limited by occlusion, noise, and poor sampling of obliquely viewed surfaces. To extend the applicability of a single depth camera, we propose a method of retaining knowledge of surfaces from depth camera images acquired over time. This method draws on techniques from computer vision and computer graphics to reconstruct a 3D point cloud from a depth image, smooth and consolidate the point cloud into a good model, estimate surface normals of the model points, and register each new depth image to the model for tracking and model update.

Algorithm

Point Cloud Generation
Point clouds are generated from pre-processed depth images. Pixels which represent edges in the depth image are discarded due to blurring that occurs at edges. 3D points are recovered from Z-depth data by similar triangles shown here. Where f is the focal length and Δu is the difference in column pixels from the image center.

Surface Normals
Surface Normals are determined by fitting a plane to local neighborhoods around each point. The centroid is chosen as a point on the plane and the normal is computed as the direction of minimum variance between each point and the centroid. The normal is found through the following minimization:

\[ \hat{n} = \arg \min_n \sum_i ((p_i - c)^T n)^2 \quad \text{subject to} \quad ||n|| = 1 \]

Iterative Closest Point Algorithm
This algorithm was chosen in order to register a new depth view with the current model. By registering new views, we are able to build up a new comprehensive model over time.

Outline
1. Select 5% of points in new data set.
2. Match each point to the point in the model that is closest to it by Euclidean distance.
3. Reject pairs whose model point does not contain a surface normal or whose distance is greater than 0.005.
4. Construct the point-to-plane error metric.
5. Minimize error metric.
6. Apply optimal rotation and translation.

The rotation and translation are found by minimizing

\[ R, t = \arg \min \sum_i [ (Rp_i + t - q_j)^T n_j]^2 \]

Consolidation
The new model is created by the Weighted, Locally-Optimal Projection (WLOP) algorithm from the registered old model and the new input point cloud. WLOP has the advantages that its input and output are unorganized point clouds without surface normals, it projects noisy points onto an underlying surface, and it produces a uniformly sampled point cloud. It iteratively updates a set of points to be as close as possible to the original data, while also employing a repulsion term in order to enforce more uniform sampling.

The results are best seen in the final results where the noise introduced was sufficiently smoothed such that the models with an without noise appear very similar.

Results

Point Cloud Model from Noiseless Data
Point Cloud Model from Noisy Data