ProFORMA: Probabilistic Feature-based On-line Rapid Model Acquisition

3D Photography FS2012
Project proposal by Gábor Sörös and Tobias Widmer

Introduction
There is lots of room for applying augmented reality techniques to consumer products like food, books, or other shopping items, to enhance them with additional information such as nutrition facts, price comparison, advertisements, descriptions indifferent languages, or services like an interactive user guide, etc. For realistic augmentation, we need to accurately register the real objects with their virtual counterparts. This registration process requires a 3D model of the real object, but creating such models is still a time-consuming task with contemporary modeling tools. Modeling is usually done in two stages, first taking photos of the object, and later loading them into a professional modeling software. If the photos lack some information about the object, there may be no chance to collect that missing data any more. We wish to accelerate this modeling process by melting the two stages together and implement a fast online 3D modeling approach called ProFORMA [1].

Approach
The ProFORMA method assumes a static camera, in front of which the user moves a rigid object around. This setup makes the segmentation of the object from the background easier because of the geometric constraints on the motion of the rigid body (epipolar constraints). Another advantage of this approach is that the background does not change a lot during the capturing process. No further assumptions are made and no prior information about the object is known, although it must be sufficiently textured.

The system consists of three main parts: tracking, reconstruction and visualization. The tracker uses FAST features[5] and tracks model points between keyframes and actual frames. At the same time, the epipolar geometry is estimated and a new keyframe is taken when sufficiently away from the previous keyframe (approx. 10 degrees). The reconstruction algorithm works with keyframes and triangulates 3D model points. The 3D points and the key camera poses are refined using bundle adjustment. The point cloud is then converted into a 3D mesh using Delaunay tetrahedralization and probabilistic triangle carving. When the 3D model is ready, it gets fed back to the tracker that estimates the actual camera pose with respect to the object. The usage of the temporal 3D model during tracking leads to a more robust object tracking system. The virtual model together with additional information is visualized from the estimated viewpoint (see Figure below from [1]).

Figure 1: Stages of interactive model reconstruction for a partial model. Left to right: (a) Object augmented with user instructions and mesh model (b) Point cloud from feature tracks and bundle adjustment (c) Delaunay Tetrahedralisation (d) Carved Mesh (e) Textured partial model.
The details of the algorithm can be found in [1]. We will implement the project in C++ and if time permits, we will port it to a smartphone model.

**Tasks and Milestones**

The list below should be seen as a guideline. Due to the fact that the project will be realized by the two of us, it’s likely that we partition the Milestones.

- Search for available code resources (PTAM [2], CGAL [3], FastCV[4], OpenGL ES etc..)
- Camera input
- FAST feature detection and matching  
  **Milestone 1 (end of March)**
- Epipolar geometry estimation
- Keyframe extraction and triangulation
- Bundle adjustment  
  **Milestone 2 (end of April)**
- Delaunay tetrahedralization
- Probabilistic triangle carving  
  **Milestone 3 (middle of May)**
- Pose estimation
- Visualization  
  **Milestone 4 (end of May)**
- Interaction and blending with additional information
- (optional) Porting to a smartphone model (e.g., Android)
- Testing
- Documentation
- Presentation  
  **Milestone 5 (end of June)**

**References**

[5] Ed Rosten, Tom Drummond – Fusing points and line for high performance tracking, in International Conference on Computer Vision (ICCV), Beijing, China, 2005