POINT-AND-SHOOT MODEL ACQUISITION

Voicu Popescu and Elisha Sacks
Computer Science Department, Purdue University
{popescu | eps@cs.purdue.edu}
http://www.cs.purdue.edu/cgvlab/

We are developing a point-and-shoot device for producing high-fidelity 3D geometric models of real-world scenes in real time. The *depth camera* consists of a digital video camera augmented with an array of laser pointers and linked to a portable computer. A texture-mapped model will be constructed automatically as the operator scans the scene. Our goal is to make 3D modeling as easy and inexpensive as home video.

Real-time geometry acquisition is an enabling technology for computer graphics, which removes the modeling bottleneck in generating realistic images of complex scenes. The applications include modeling and simulation, photorealistic rendering, and virtual environments for training and telecollaboration. Current 3D acquisition devices are slow, bulky, expensive, and produce redundant yet incomplete data sets. Moreover some only work in controlled environments and some do not acquire color. Our depth camera should solve these problems.

The depth camera will operate as follows.

1. *Grab a video frame* (Figure 1)
2. *Extract depth* (Figure 2):
   - Locate laser blobs in the frame.
   - Compute corresponding 3D points.
3. *Register frame* (Figure 3):
   - Track camera motion.
   - Place frames in common coordinate system.
4. *Update model*:
   - Integrate depth and color information from new frame.
5. *Display model:*
   - Guide operator with real time visual feedback.

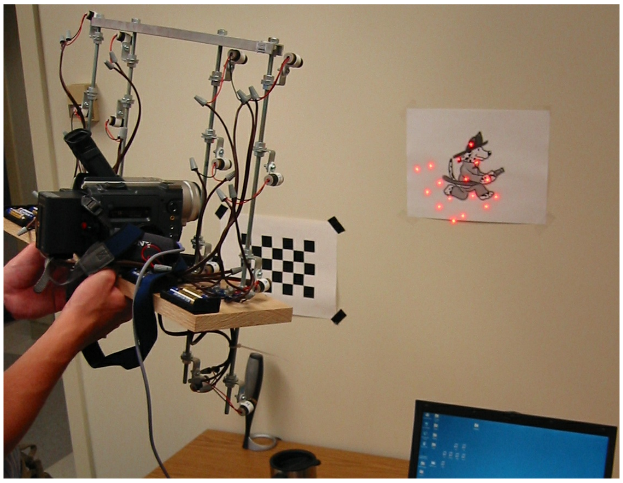

Figure 1. Point-and-shoot scene acquisition

Making the device work poses several research challenges.

The first challenge is to compute the 3D positions where the lasers hit the scene. Our approach is to detect the laser blobs in the video image via epipolar geometry and image processing, then to triangulate the blobs to obtain 3D positions.

The second challenge is to register the frames in a common coordinate system by computing the camera motion in real time. Our approach is to minimize the discrepancy in position and color between consecutive frames, while exploiting frame coherence for rapid convergence.

The third challenge is to build a complete scene model from a large set of frames, each of which contains detailed color and sparse depth. Our approach is to build the model interactively from the incoming frames to provide real-time visual feedback to the operator. The tight feedback loop should make it easy for the operator to acquire a complete model at a specified level of detail.

We have built a prototype depth camera from commodity components and are using it to investigate these issues. We have encouraging preliminary results for depth extraction and registration.

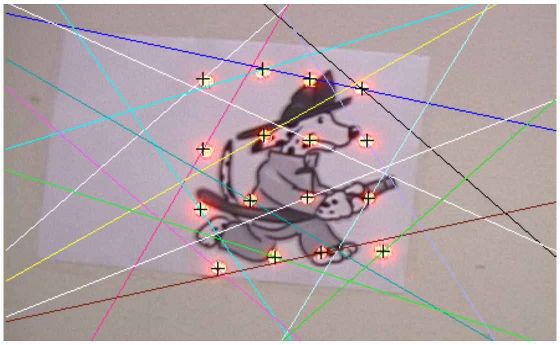

Figure 2. Blob detection on epipolar lines

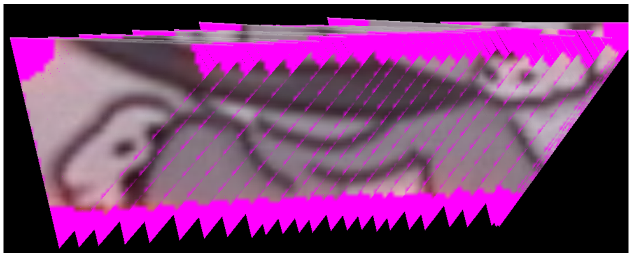

Figure 3. Frame registration