Live Ego-Motion Estimation

David Nistér
Sarnoff Corporation
CN5300, Princeton, NJ 08530
dnister@sarnoff.com

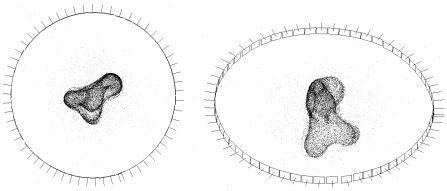

Figure 1: Reconstruction made from a sequence of turntable images.

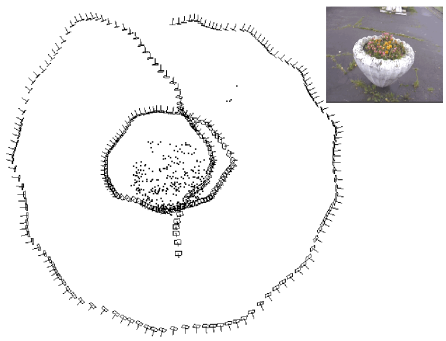

Figure 2: Reconstruction made at real-time rate from a freehand motion. The camera makes an outer circle and then an inner circle with some forward motion in between.

1. Description

This demo will show live (real-time, low delay) estimation of the motion of a perspective camera. Some scene points are also reconstructed. The reconstruction of structure and motion is based on the video input alone and performed on a regular PC-laptop without special hardware. The estimation will be shown in three ways:

- Real-time directly from a video file stored on the hard-drive.
- Real-time through the capture card with camera playback as the external video source.
- Live from camera using both turntable and freehand motion.

The system uses the techniques described in [1] and [2], among other things. Examples of reconstructions made from disk are shown in Figures 1 and 2. A screenshot of the demo running is shown in Figure 3. An example of real-time reconstruction from camera playback is shown in Figure 4. An example of live reconstruction is shown in Figure 5.

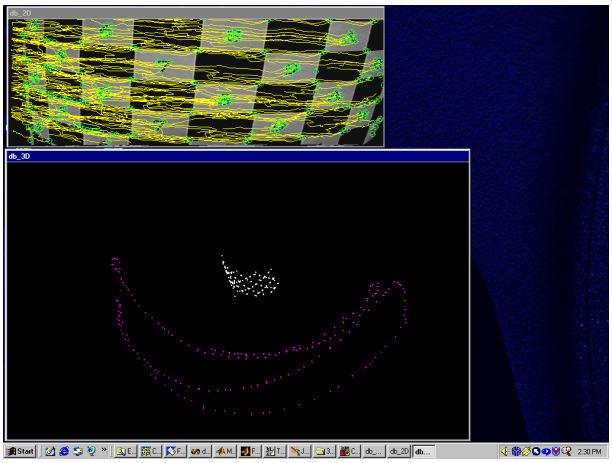

Figure 3: A screenshot of the demo running. The window at the top displays feature tracks on the original video. The window at the bottom displays the camera positions and some structure points. The graphic display is done concurrently with the estimation of camera motion, including feature detection, feature tracking, robust estimation and local bundle adjustment. The estimation is done directly on video from the capture card, with camera playback or live videofeed as the external video source.

References

[1] D. Nistér. An Efficient Solution to the Five-Point Relative Pose Problem, *IEEE Conference on Computer Vision and Pattern Recognition*, Volume 2, pp. 195-202, 2003.

[2] D. Nistér. Preemptive RANSAC for Live Structure and Motion Estimation, *IEEE International Conference on Computer Vision*, to appear, 2003.

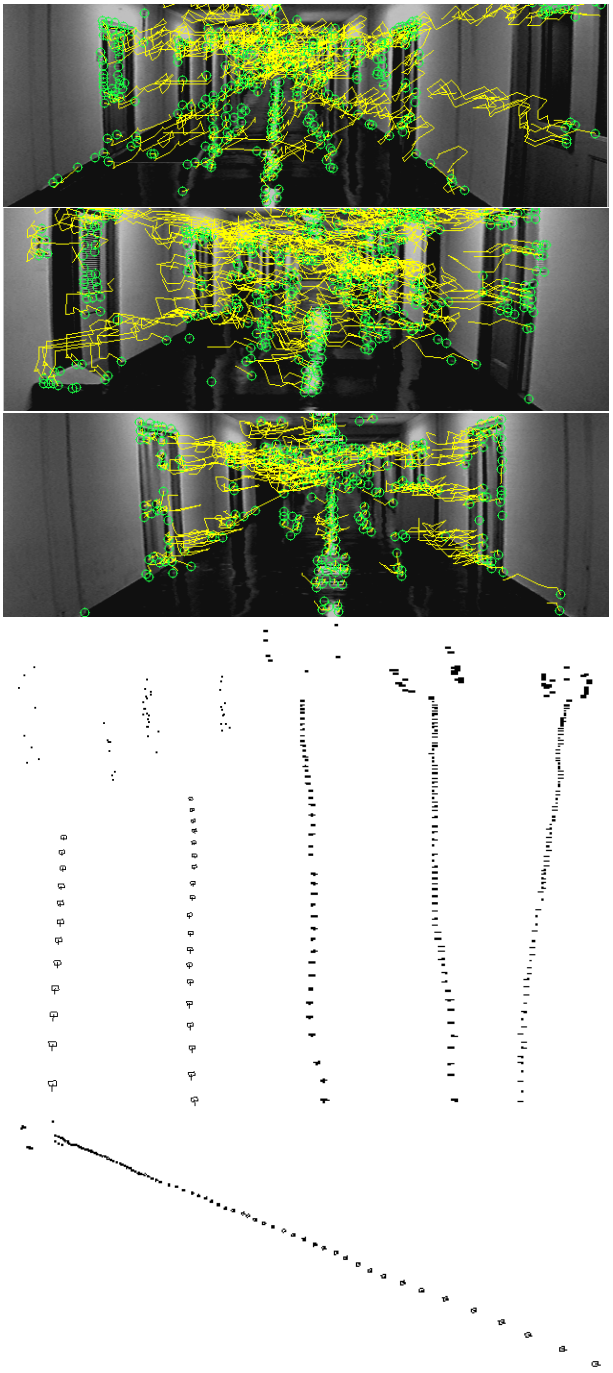

Figure 4: Estimation of the motion of a handheld camera, induced by walking down a 160 meters long corridor. The whole motion is successfully integrated into the same coordinate frame. Note how the straight trajectory builds up from left to right. The estimation was done in real-time from the tape playback. Thus, the system has to deal with capture, dropping frames if not able to keep up etc. The estimation was made with both the feature-track window and the 3D scene window displaying. In this mode, the delay from the tape to the reconstruction displaying on the screen is ~ 1 second. No prior knowledge of the motion was used and the system is identical to the one used in Figure 5.

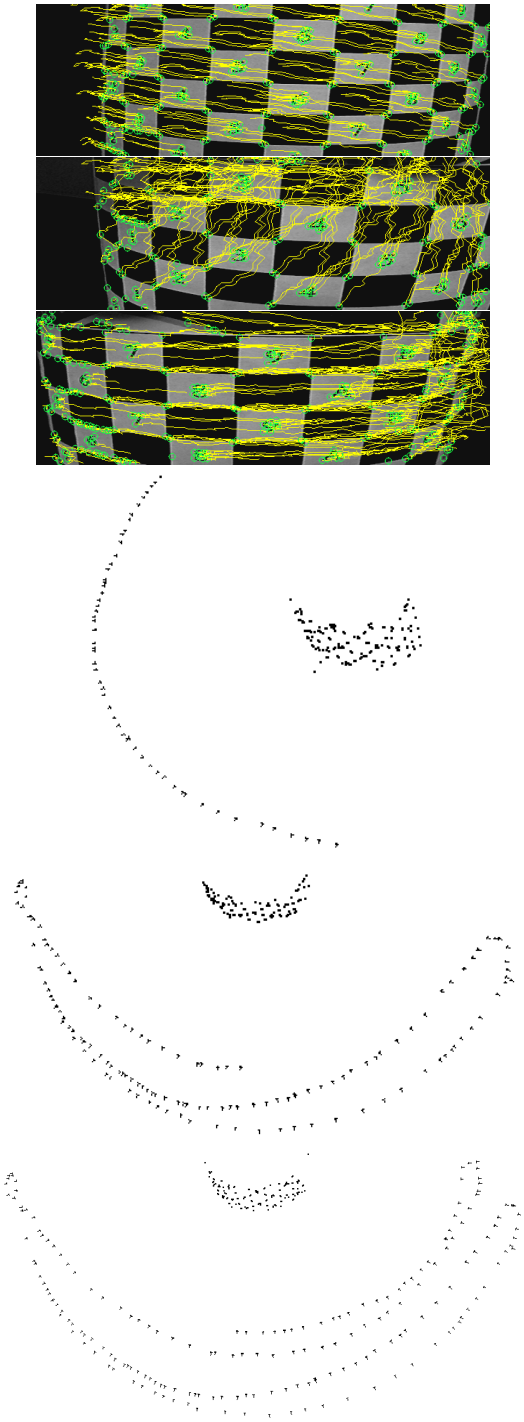

Figure 5: Estimation of a freehand motion created with a calibration cylinder in one hand and a camcorder in the other. The estimation was done live. Thus, the system has to deal with capture, dropping frames if not able to keep up etc. The estimation was made with both the feature-track window and the 3D scene window displaying. Note how the trajectory builds up from top to bottom. No knowledge of the structure or motion was used and the system is identical to the one used in Figure 4