Real-Time Position and Orientation Estimation with a Monocular Camera and Motion Sensors over a Structured, Planar Surface

Marco Grimm, Rolf-Rainer Grigat

Contact Information

Marco Grimm · Vision Systems · Technical University of Hamburg-Harburg · Germany
grimm@tu-harburg.de · Phone: +49 40 42878-35 36 · Fax: +49 40 42878-29 11
http://www.tuvision.de

A brief description of the demonstration

This project is a co-operation between Philips Semiconductors Hamburg and the Technical University of Hamburg-Harburg, Vision Systems department.

Motivation

The output signals of motion sensors and a camera are used to realize a pen-like human-computer interface with a maximum of spatial freedom. Utilizing three-dimensional detection of the pen's position and orientation (pose) several applications are possible, e.g. ergonomic human-computer interfaces in 3D, devices for handwriting input, image mosaicing applications or remote-controls of PC-applications in 2D or 3D [1].

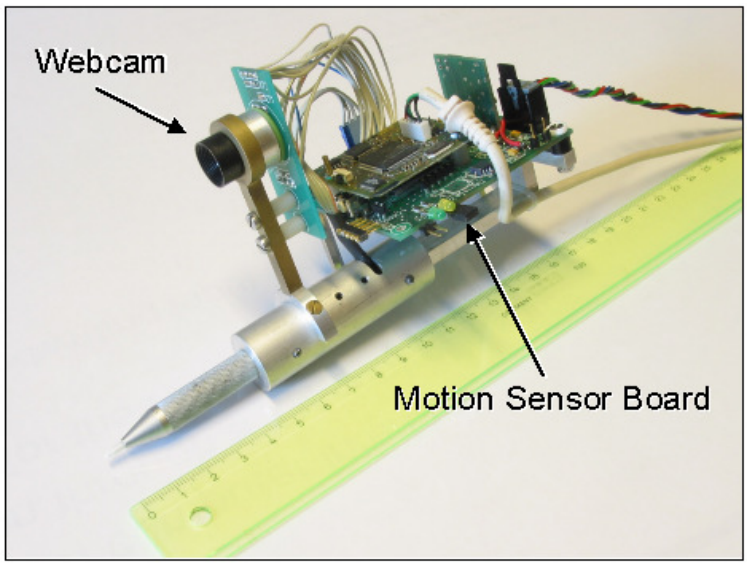

Figure 1: Pen-like sensor device prototype with webcam and motion sensors.

Approach and Methods

A hybrid method to determine the three dimensional pose is used. The sensor device shown in figure 1 contains a camera and motion sensors (accelerometers). The pen-like interface can be used over a structured, planar surface, e.g. a wooden desk or an image. As features Harris corners are used, so the structured surface must contain enough corners. The camera pose is calculated using a modified plane-to-image homography from point correspondences. The critical part of this calculation is the estimation of the orientation. Therefore the accelerometer signals are used to estimate the tilt of the complete system. The remaining four parameters are then estimated from the correspondences. Real-time Simultaneous Localization and Map-building (SLAM) is used to organize the world coordinates of the recognized feature points similar to [2].

Results

The results of the hybrid pose estimation can be visualized with a Visual C++ software project in real-time on a desktop PC. The three-dimensional device position can be visualized with OpenGL over a synthetic planar surface. The pose can be used for image mosaicing of the captured surface as shown in figure 2.

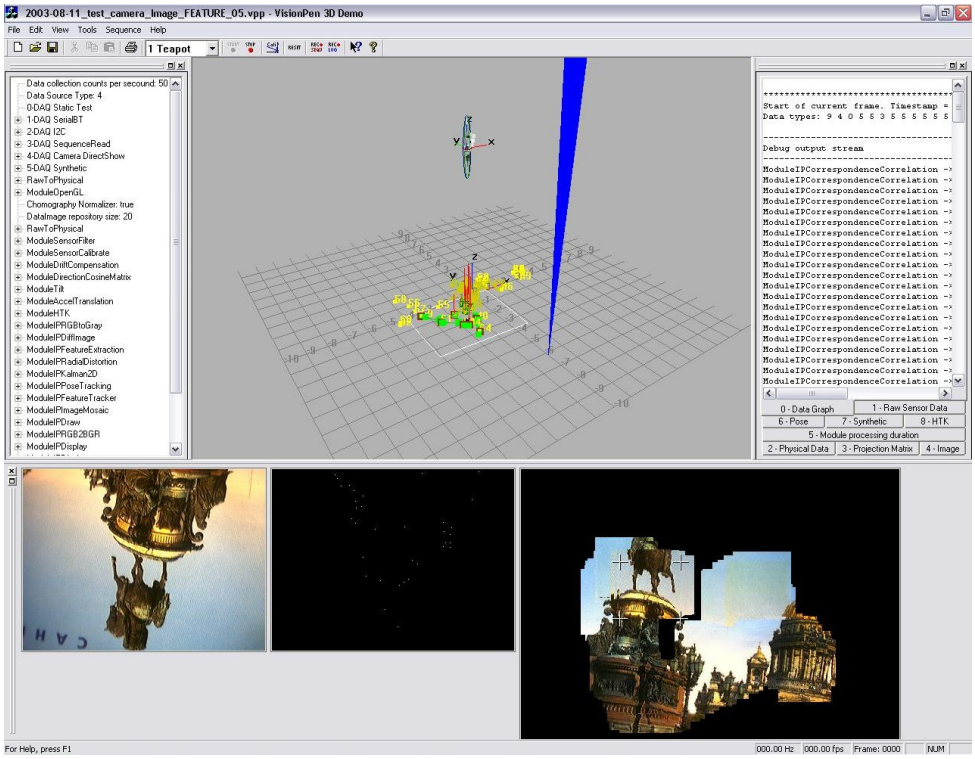

Figure 2: Visualization of processing results in real-time

References

[1] I. Shdaifat, M. Grimm and R.-R. Grigat. *Man-Machine Interface Using Motion Sensors and Monocular Camera for 6D Motion Estimation to Control Multimedia Applications.* Demonstration on 8th IEEE International Conference on Computer Vision (ICCV), Vancouver, Canada, 2001.

[2] A. J. Davison. *SLAM with a Single Camera*. Workshop on Concurrent Mapping and Localization for Autonomous Mobile Robots, in conjunction with ICRA, 2002.