Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Sensor Fusion for Augmented Reality ⋆

Fredrik Gustafsson, Thomas B. Schön, Jeroen D. Hol *

* *Division of Automatic Control, Linköping University, SE-581 83 Linköping, Sweden (e-mail: {fredrik, schon, hol}@isy.liu.se)*

Abstract: The problem of estimating the position and orientation (pose) of a camera is approached by fusing measurements from inertial sensors (accelerometers and rate gyroscopes) and a camera. The sensor fusion approach described in this contribution is based on nonlinear filtering using the measurements from these complementary sensors. This way, accurate and robust pose estimates are available for the primary purpose of augmented reality applications, but with the secondary effect of reducing computation time and improving the performance in vision processing. A real-time implementation of a nonlinear filter is described, using a dynamic model for the 22 states, where 100 Hz inertial measurements and 12.5 Hz vision measurements are processed. An example where an industrial robot is used to move the sensor unit, possessing almost perfect precision and repeatability, is presented. The results show that position and orientation accuracy is sufficient for a number of augmented reality applications.

Keywords: Sensor fusion, nonlinear filtering, tracking, Kalman filter, augmented reality

1. EXTENDED ABSTRACT

This contribution deals with estimating the position and orientation (pose) of a camera in real-time, using measurements from inertial sensors (accelerometers and rate gyroscopes) and a camera. A system has been developed to solve this problem in unprepared environments, assuming that a map or scene model is available. For a more detailed description of the overall system and the model building we refer to Hol et al. [2007].

Existing pose tracking algorithms in literature stem from the computer vision society, where the primary information is taken from the image stream. IMU's are in some studies [Davison et al., 2007] used as a supporting sensor. We have taken a reverse approach inspired by navigation systems in aircraft, where the IMU is the primary sensor, and vision information is the support information used in a Kalman filter to stabilize inherent drift when integrating IMU measurements of accelerations and angular velocities. This system design has proved to give a very robust (to occlusion and fast movements of the camera) and accurate pose estimate, but it also as a side effect decreases the computation burden in the computer vision algorithms, in that an accurate prior estimate of feature locations in the image can be computed.

A further fundamental contribution is to use SLAM (simultaneous localization and mapping) for on-line learning of a scene model, which is required for the pose estimator. SLAM [Thrun et al, 2005] was developed in the mobile robotics community, where only a few dynamical states (typically three) are used, while we have developed marginalization techniques enabling complex dynamics models with a high-dimensional (22 in this application) state vector.

⋆ This work has been performed within the MATRIS consortium, which is a sixth framework research program within the European Union (EU), contract number: IST-002013.

The supporting video illustrates:

- Broadcasting applications of augmented reality in sports, news and entertainment.
- The difference of augmented and virtual reality.
- Prior state of the art in form of expensive marker based system.
- The individual contributions from the MATRIS consortium, explaining and providing an overview of the MATRIS system.
- In particular, the IMU sensor which is the core of the hardware is shown in two different versions. The first one is to be attached to the film camera. The second one can be used standalone, as it integrates synchronized and calibrated three-dimensional accelerometers, gyroscopes, magnetometers and one miniature camera in one single housing.
- Our development platform, where the sensor kit is attached to an industrial robot. In this way, preprogrammed motions can be repeated indefinitely in a Monte Carlo fashion, where for instance robustness to changes in lightning and scene can be evaluated.
- Finally, the performance of the Matris demonstrator is exemplified.

The performance of the developed pose tracking algorithm satisfies the television performance requirements of less than two centimeters and one degree error at all times, independent of the camera motion.

REFERENCES

Davison, A. J., Reid, I., Molton, N., and Strasse, O. (2007). MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Patterns Analysis and Machine Intelligence*, 29(6):1052–1067.

S. Thrun, W. Burgard, D. Fox. Probabilistic Robotics. Intelligent Robotics and Autonomous Agents. The MIT Press, Cambridge, MA, USA. 2005.

Hol, J.D., Schön, T. B., Luinge, H., Slycke, P., and Gustafsson, F. (2007). Enabling real-time tracking by fusing measurements from inertial and vision sensors. *Journal of Real-Time Image Processng.*