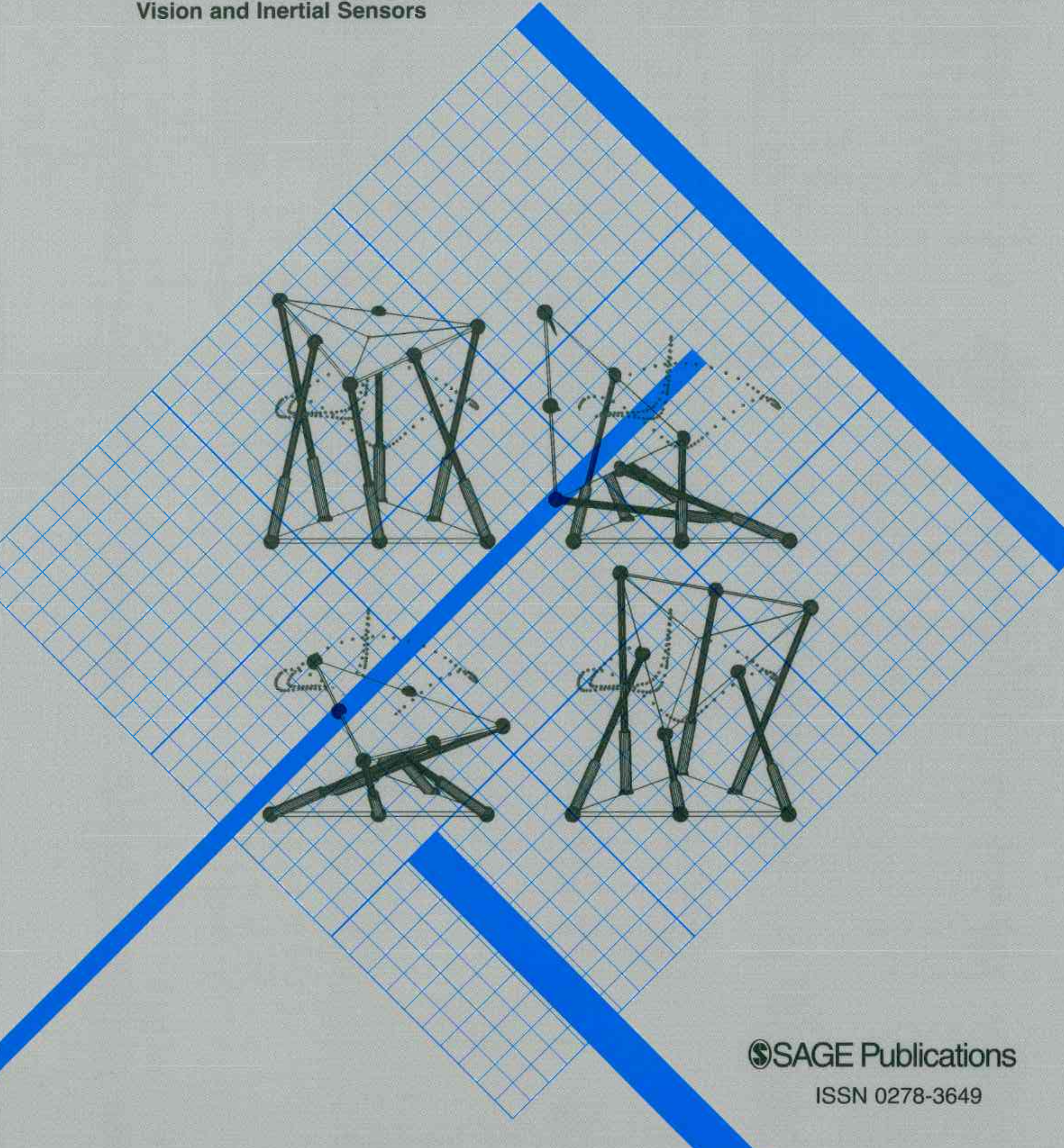
The International Journal of
Volume 26 Number 6 June 2007
www.ijrr.org

Robotics Research

Special Issue
2nd Workshop on Integration of Vision and Inertial Sensors



SAGE Publications
ISSN 0278-3649

The International Journal of

Robotics Research

Volume 26, Number 6, June 2007

Special Issue: 2nd Workshop on Integration of Vision and Inertial Sensors

•CONTENTS•

MM denotes that a paper includes multimedia material and can be viewed at *www.ijrr.org.*

Editorial

Special Issue: 2nd Workshop on Integration of Vision and Inertial Sensors.

This special issue of *The International Journal of Robotics Research* is devoted to the topic of Integration of Visual and Inertial Sensors.

For robots, as for animals, a sense of self-motion is critical to performing useful tasks in a complex and non-priory mapped world. The problems with sensing, particularly reliability and robustness, are well known to robotocists and sensor-fusion is a standard technique. Animals also make use of multiple sensory modalities, not just vision, smell, hearing, touch and taste but also joint position, muscle exertion, balance and motion. Some have even developed task- or environment-specific sensors such as acoustic ranging, magnetic dead-reckoning and electric fields.

In this special issue we discuss two sensors that are common to animals and robots – vision and inertial sensing – which when combined provide rich information about self-motion and also the structure of the world. The advantage of these sensors is that they require no external infrastructure such as GPS satellites or visual or acoustic navigation beacons. Neither vision nor inertial sensors alone are perfect; however they have nice complementarities where each compensates for a shortcoming in the other. While inertial sensing is able to follow motions of any speed, its uncertainty – in particular at slow motions – is high compared to visual input. On the other hand visual sensors have problems with fast motions and find it difficult to distinguish between rotational and translational motion.

Visual and inertial sensors today have high performance and are low cost and compact. They require no external reference and emit no radiation. Digital camera chips and micromachined accelerometers and gyroscopes are now commodities, and when combined with today's available computing can provide robust estimates of self-motion as well 3D scene structure.

This special issue is about visual and inertial sensing and it is motivated by the work presented at the second workshop on Integration of Visual and Inertial Sensors – InerVis2005. The second InerVis was held in Barcelona, Spain, alongside the International Conference on Robotics and Automation (ICRA05) on 18 April 2005. The first InerVis was held in June 2003 at the International Conference on Advanced Robotics (ICAR03) in Coimbra, Portugal. The numbers of papers accepted were small and the meeting is held in a single track format, in order to increase interaction and discussion among the participants, followed by a closing debate discussion on the topic. Papers submitted to the workshop had been rigorously refereed by the Program Committee, consisting of the co-chairs of InerVis2005 and the following researchers:

- Ernst Dickmanns (Universität der Bundeswehr München, Germany)
- William Hoff (The Colorado School of Mines, CO, USA)
- Axel Pinz (EMT, Graz University of Technology, Austria)
- Thierry Viéville (INRIA, Sophia Antipolis, France)
- François Berry (LASMEA, Université Blaise Pascal, France)
- Marnix Nuttin (Katholieke Universiteit Leuven, Belgium)
- Giorgio Metta (LIRA-Lab, DIST, University of Genova, Italy)

in addition to ourselves.

The authors of selected papers were invited to submit extended versions of their conference papers, which then under-

The International Journal of Robotics Research
Vol. 26, No. 6, June 2007, pp. 515–517
DOI: 10.1177/0278364907079903

went the regular review process of the IJRR, resulting in this Special Issue where we present eight articles.

The first paper in this special issue, by Corke, Lobo and Dias provides a tutorial style introduction to the topic. It describes the fundamentals of visual and inertial sensing from the perspectives of physical principles and the engineering and biological implementations. It shows that these sensors have useful complementarities, each able to cover the limitations and deficiencies of the other. Using a standard notation the paper describes the standard inertial only navigation system (INS), vision-only system (structure from motion) and two approaches to combine inertial and visual fusion. They are defined as *loosely* and *tightly* coupled. The loosely coupled approach uses separate INS and SFM blocks, running at different rates and exchanging information. The tightly-coupled systems combine the disparate raw data of vision and inertial sensors in a single, optimum filter, rather than cascading two filters, one for each sensor.

Chalimbaud et al. propose sensing hardware based on a CMOS camera and an artificial vestibular system, inspired by the neurological system of primates, which detects changes in posture by the central nervous system through a vestibular process. The whole sensor can be considered as an embedded sensor where one of the most original aspect of this approach is the use of a System On Chip implemented in a FPGA to implement the whole system. The sensing device is designed around a 4 million pixel CMOS imager and the artificial vestibular set comprises three linear accelerometers and three gyroscopes. With its structure, the system exhibits a high degree of versatility and allows the implementation of parallel image and inertial processing algorithms. The proposed approach is illustrated with feature depth estimation using a Kalman filter.

Harada et al. built a tiny absolute orientation estimating device equipped with a network function using accelerometers and magnetometers to estimate gravity and the geomagnetic field, respectively. They also propose an estimation method that excludes the effect of magnetic disturbances. An advantage of this estimation method is that models can be switched according to the environment. A Square-Root Central Difference Kalman Filter (SR-CDKF) was found to be the best method to estimate the orientation in the presence of motion and magnetic disturbances considering stability, accuracy, and calculation cost.

Lobo and Dias present a simple calibration method for relative pose between low-cost inertial sensors and cameras. The cue is to observe the vertical direction, using a well oriented visual pattern and gravity, so that the rotation between the two sensors can be estimated. To estimate translation a passive turntable is used, onto which the inertial sensor needs to be centered. Validation of the method includes simulation and real data, and a toolbox is provided.

Armesto et al. present a tracking system for egomotion estimation which compares the fusion of vision and inertial measurements using EKF and UKF (Extended and Unscented Kalman Filters) adapted to the multi-rate nature of the sensors. Inertial sensing is sampled at a higher frequency than visual sensing. It is shown that fusing both measurements achieves a significant improvement over the cases of using either sensor modality alone. It is also shown that both filters produce similar results, though EKF at less computational costs.

Gemeiner et al. extend this egomotion estimation with a subsequent structure estimation procedure, where new landmarks can be inserted on-line into the map of a scene. The benefit of this simultaneous structure and motion estimation are that the changes in aspect or illumination can be handled and fast camera motion can be performed. The sparse structure estimation approach is based on the detection of corner features in the images. A bank of Extended Kalman filters, one per corner feature, is used to estimate the position and the quality of structure points and to include them into the structure estimation process. The system is demonstrated on a mobile robot executing known motions, such that the estimation of the egomotion in an unknown environment can be compared to ground truth.

Tao et al. introduce a real-time system for tracking a human arm in 3D where the two modalities are used in a complimentary fashion: inertial sensing gives hints to vision on where to search for features. The advantage is that such a system would be cheap and could be used for rehabilitation or other purposes at home.

Finally, Meguro et al. present a 3D reconstruction method using omni-directional images taken from a moving vehicle. A multi-baseline stereo reconstruction is performed, with the camera position given by an integrated positioning system based on GPS (phase based) and dead reckoning using a fiber optic gyroscope and odometry. They show that this enables the use of variable baseline lengths, such that the accuracy of reconstruction can be improved. Results are given with the reconstruction of the shapes of vehicles parked alongside a road, where measurements were found to have a standard deviation of 140 mm within a range of 10 m.

The articles in this issue also contain a wealth of resources for the implementer. In addition to a large collection of references to relevant literature in the field the articles contain code and detailed formulations. Gemeiner et al. include expressions for the Jacobians required for extended Kalman filter implementation of the ego-motion estimator based on a quaternion formulation. Armesto et al. include a full Matlab implementation of an inertial-visual structure from a motion system, as well as data sets. Lobo and Dias provide a calibration toolbox. The introductory article by Corke et al. contains pointers to other tools for visual tracking and structure from motion.

These eight articles present some of the latest research on integration of vision and inertial sensors. We hope that you find these articles interesting and motivating, and that you will find the opportunity to participate in future InerVis workshops.

Finally we would like to thank the program committee, all the authors and the reviewers for their fine efforts and contributions, and SAGE Publications, with special thanks to Jennet Batten, the Managing Editor, for keeping us on track and managing the entire submission and review process.

Guest editors and Co-Chairs of InerVis2005:

Jorge Dias,
Institute of Systems and Robotics,
University of Coimbra,
Portugal

Markus Vinzce,
Automation and Control Institute,
Technical University of Vienna,
Austria

Peter Corke,
CSIRO ICT Centre,
Queensland Centre for Advanced Technologies,
Australia

Jorge Lobo,
Institute of Systems and Robotics,
University of Coimbra,
Portugal