Cameras with inertial sensors

Inertial sensors attached to a camera can provide valuable data about camera pose and movement. In biological vision systems, inertial cues provided by the vestibular system play an important role, and are fused with vision at an early processing stage. Micromachining enables the development of low-cost single-chip inertial sensors that can be easily incorporated alongside the camera's imaging sensor, thus providing an artificial vestibular system. As in human vision, low-level image processing should take into account the ego motion of the observer. In this article we present some of the benefits of combining these two sensing modalities.

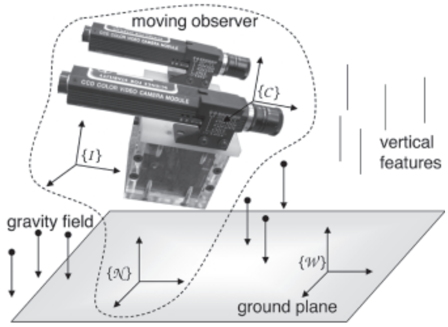

Figure 1. Stereo cameras with an inertial measurement unit used in experimental work.

Figure 1 shows a stereo-camera pair with an inertial measurement unit (IMU), assembled with three capacitive accelerometers and three vibrating structure gyros. The 3D-structured world is observed by the visual sensor, and its pose and motion are directly measured by the inertial sensors. These motion parameters can also be inferred from the image flow and known scene features. Combining the two sensing modalities simplifies the 3D reconstruction of the observed world. The inertial sensors also provide important cues about the observed scene structure, such as vertical and horizontal references. In the system, inertial sensors measure resistance to a change in momentum, gyroscopes sense angular motion, and accelerometers change in linear motion. Inertial navigation systems obtain velocity and position by integration, and do not depend on any external references, except gravity.

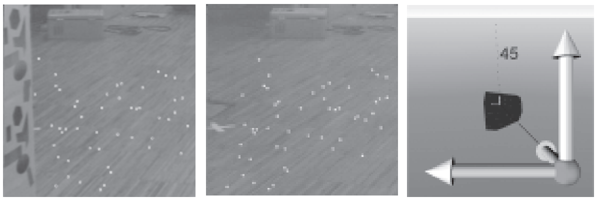

Figure 2. Ground-plane 3D-reconstructed patch.

The development of Micro-Electro-Mechanical Systems (MEMS) technology has enabled many new applications for inertial sensors beyond navigation, including aerospace and naval applications. Capacitive linear acceleration sensors rely on proof mass displacement and capacitive mismatch sensing. MEMS gyroscopes use a vibrating structure to measure the Coriolis effect induced by rotation, and can be surface micromachined providing lower-cost sensors with full signal-conditioning electronics. Although their performance is not suitable for full inertial navigation, under some working conditions or known system dynamics they can be quite useful.

In humans, the sense of motion is derived both from the vestibular system and retinal visual flow, which are integrated at very basic neural levels. The inertial information enhances the performance of the vision system in tasks such as gaze stabilisation, and visual cues aid spatial orientation and body equilibrium. There is also evidence that low-level human visual processing takes inertial cues into account, and that vertical and horizontal directions are important in scene interpretation. Currently-available MEMs inertial sensors have performances similar to the human vestibular system, suggesting their suitability for vision tasks.[1]

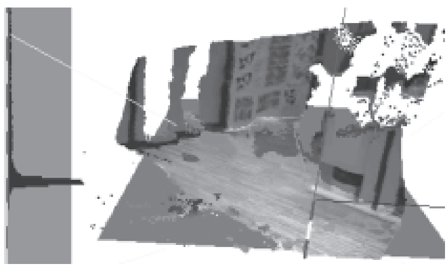

Figure 3. Aligned depth map showing histogram for ground-plane detection.

The inertial-sensed gravity vector provides a unique reference for image-sensed spatial directions. If the rotation between the inertial and camera frames of reference is known, the orthogonality between the vertical and the direction of a level plane image vanishing point can be used to estimate camera focal distance.[1] When the rotation between the IMU and camera is unknown from construction, calibration can be performed by having both sensors measuring the vertical direction.[2] Knowing the vertical-reference and stereo-camera parameters, the ground plane is fully determined. The collineation between image ground-plane points can be used to speed up ground-plane segmentation and 3D reconstruction (see Figure 2).[1] Using the inertial reference, vertical features starting from the ground plane can also be segmented and matched across the stereo pair, so that their 3D position is determined.[1]

The inertial vertical reference can also be used after applying standard stereo-vision techniques. Correlation-based depth maps obtained from stereo can be aligned and registered using the vertical-reference and dynamic-motion cues. In order to detect the ground plane, a histogram in height is performed on the vertically-aligned map, selecting the lowest local peak (see Figure 3). Taking the ground plane as a reference, the fusion of multiple maps reduces to a 2D translation and rotation problem. The dynamic inertial cues can be used as a first approximation for this transformation, providing a fast depth-map registration method.[3] In addition, inertial data can be integrated into optical flow techniques. It does this by compensating camera ego motion, improving interest-point selection, matching the interest points, and performing subsequent image-motion detection and tracking for depth-flow computation. The image focus of expansion (FOE) and centre of rotation (COR) are determined by camera motion and can both be easily found using inertial data alone, provided that the system has been calibrated. This information can be useful during vision-based navigation tasks.

Studies show that inertial cues play an important role in human vision, and that the notion of the vertical is important at the first stages of image processing. Computer-vision systems for robotic applications can benefit from low-cost MEMS inertial sensors, using both static and dynamic cues. Further studies in the field, as well as bio-inspired robotic applications, will enable a better understanding of the underlying principles. Possible applications go beyond robotics, and include of artificial vision and vestibular bio implants.

Jorge Lobo and Jorge Dias
Institute of Systems and Robotics
Electrical and Computer Engineering Department
University of Coimbra, Portugal
E-mail: {jlobo, jorge}@isr.uc.pt

References

1. J. Lobo and J. Dias, *Vision and Inertial Sensor Cooperation, Using Gravity as a Vertical Reference*, **IEEE Trans. on Pattern Analysis and Machine Intelligence 25** (12), 2003.
2. J. Alves, J. Lobo, and J. Dias, *Camera-Inertial Sensor modelling and alignment for Visual Navigation,* **Proc. 11th Int'l Conf. on Advanced Robotics,** pp. 1693-1698, 2003.
3. J. Lobo and J. Dias, *Inertial Sensed Ego-motion for 3D vision,* **Proc. 11th Int'l Conf. on Advanced Robotics,** pp. 1907-1914, 2003.