



**HAL**  
open science

## Tracking a Mobile Robot Position Using Vision and Inertial Sensor

Francisco Coito, António Eleutério, Stanimir Valtchev, Fernando Coito

► **To cite this version:**

Francisco Coito, António Eleutério, Stanimir Valtchev, Fernando Coito. Tracking a Mobile Robot Position Using Vision and Inertial Sensor. Luis M. Camarinha-Matos; Nuno S. Barrento; Ricardo Mendonça. Technological Innovation for Collective Awareness Systems: 5th IFIP WG 5.5/SOCOLNET Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2014, Costa de Caparica, Portugal, April 7-9, 2014. Proceedings, AICT-423, Springer, pp.201-208, 2014, IFIP Advances in Information and Communication Technology, 978-3-642-54733-1. 10.1007/978-3-642-54734-8\_23 . hal-01274775

**HAL Id: hal-01274775**

**<https://inria.hal.science/hal-01274775>**

Submitted on 16 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Tracking a Mobile Robot Position Using Vision and Inertial Sensor

Francisco Coito, António Eleutério, Stanimir Valtchev, Fernando Coito,

Faculdade de Ciências e Tecnologia – Universidade Nova de Lisboa,  
2829-516 Caparica, Portugal  
francisco.coito@sapo.pt, a.eleuterio@campus.fct.unl.pt, ssv@fct.unl.pt,  
fjvc@fct.unl.pt

**Abstract.** Wheeled mobile robots are still the first choice when it comes to industrial or domotic applications. The robot's navigation system aims to reliably determine the robot's position, velocity and orientation and provide it to control and trajectory guidance modules. The most frequently used sensors are inertial measurement units (IMU) combined with an absolute position sensing mechanism. The dead reckoning approach using IMU suffers from integration drift due to noise and bias. To overcome this limitation we propose the use of the inertial system in combination with mechanical odometers and a vision based system. These two sensor complement each other as the vision sensor is accurate at low-velocities but requires long computation time, while the inertial sensor is able to track fast movements but suffers from drift. The information from the sensors is integrated through a multi-rate fusion scheme. Each of the sensor systems is assumed to have its own independent sampling rate, which may be time-varying. Data fusion is performed by a multi-rate Kalman filter. The paper describes the inertial and vision navigation systems, and the data fusion algorithm. Simulation and experimental results are presented.

**Keywords:** Mobile robotics, multi-rate sampling, sensor fusion, vision, inertial sensor.

## 1 Introduction

The number of industrial mobile robots applications has increased significantly in the last decades and its use in other indoor environments also shows a growing interest. Due to its simple mechanical structure and agility wheeled robots are still the first choice when it comes to industrial or domotic applications. Knowing exactly where a mobile entity is and monitoring its trajectory in real-time is an essential feature. Most applications, however, still rely on a more or less complex infrastructure to fulfill this task and that reduces the robot's autonomy [1-3].

The guidance system for a mobile robot is usually composed by three modules: trajectory generation and supervision system, control system and navigation system. The navigation system aims to determine reliable information on the robot's position, velocity and orientation and provide it to the other modules. Knowing the location of

an autonomous mobile equipment and monitoring its trajectory in real-time is still a challenging problem that requires the combined use of a number of different sensors. The most frequently used sensors are inertial measurement units (IMU) [1-8]. With inertial sensors the robot's position and velocity can be computed by integrating the acceleration measurements given by the sensors, but the position and the orientation errors grow over time as a result from the accumulation of the measurements' noise and bias. As the computation position from acceleration requires two integration steps, the position error grows much faster than the error of orientation, which requires only one integration step. Thus, when an IMU is used for position measurement it is usually combined with another position measurement system with lesser drifting error.

To tackle with this task a number of solutions are proposed in the literature. Some authors use multiple IMUs or combine them with electromechanical odometers [2,7]. As all the sensors are prone to drift errors the problem is not eliminated but the precision is improved. The use of GPS as an absolute positioning system is also proposed [5,8], however this is not an option for indoor operation. A number of solutions for indoor applications use active beacons as a means to determine the absolute position of the robot [1-3]. This type of approach requires the existence of a fixed infrastructure that involves additional cost and reduces the robot's autonomy.

### **1.1 The Proposed System**

In this work we propose the use a vision based system in combination with the inertial system to determine the robots position and velocity. The use of vision together with inertial sensors has been used by other authors [4,6]. Most systems consider the use of one or more cameras with known positions on the robot's environment. An IMU attached to a moving camera is proposed in [6] to determine the grip position of an industrial robotic manipulator. Experimental results also make use of measurements from mechanical odometers included on the robot traction system.

### **1.2 Multi-rate Sensor Fusion**

When fusing the information from multiple sensor systems each sensor presents a different time delay. While inertial systems are fast, allowing sample rates in the order of several kHz at least, when vision algorithms are involved the computation time associated with feature evaluation and coordinate computation causes a significantly larger time delay. As a result the vision system position estimates are delayed, both to the robot real position, as well as to the IMU's information.

In summary, while inertial measurements are frequent and accurate on the short term, but present drift, vision information has long (variable) time delays, which renders them inappropriate for driving control purposes. Thus, the two systems are complementary to each other and the fusion of their information can improve the system overall performance.

The integration of the IMU and vision systems is done by means of a Kalman filter (KF). As the KF is not computationally intensive, it is easily applied in real-time applications such as inertial navigation systems.

## 2 Contribution to Collective Awareness Systems

The impact information technology on every day's life is enormous and has jumped from plants to offices and the household. Most equipment embeds one or more computers and the ability to access and use web based information is increasing daily. However, the equipment ability to integrate the human environment is still limited by their lack of mobility. Households are unstructured environments and, in spite of the progress in the mobile robotics area, this is still a challenge for the autonomy of mobile equipment and the ability to move around freely without supervision.

## 3 Position Estimation

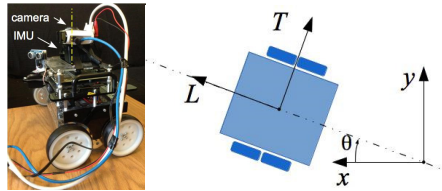
In this work a wheeled robot is considered, moving on an almost plane horizontal surface, located indoors.

### 3.1 The Inertial Measurement System

The mobile robot dynamics is represented in two coordinate frames: the robot frame and the world frame. The robot frame is fixed onto the moving robot, usually with the origin at the center of mass, which is assumed to lie between the wheels. The world frame is fixed on the field. Fig. 1 illustrates how the two frames are related.

The IMU sensor is placed at the center of the robot. The gyroscope is aligned with the vertical axis. The position of the robot is given by:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \cos(\theta(t)) & \sin(\theta(t)) \\ -\sin(\theta(t)) & \cos(\theta(t)) \end{bmatrix} \begin{bmatrix} v_L \\ v_T \end{bmatrix}, \quad (1)$$



**Fig. 1.** Left: Mobile robot. Right: Coordinate frames:  $\langle x, y \rangle$  world frame;  $\langle L, T \rangle$  robot frame;  $\theta$  orientation.

Where  $\langle x, y \rangle$  is the robot position on the world frame;  $\langle v_L, v_T \rangle$  are the longitudinal and transversal robot speed on it's own frame and  $\theta$  is the orientation.

The values for  $\theta$ ,  $v_L$  and  $v_T$  result from the integration of the measured rotation rate and translational accelerations. In order to minimize the error these integral values are computed through a KF.

### 3.2 The Vision System

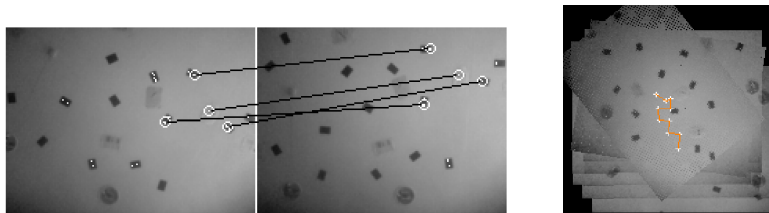
The vision based odometry system uses the FAST corners method [9] to extract features from each image. Two consecutive images are compared and their features are matched. From this comparison the camera displacement is computed through a kinematic transformation of a frame onto the next.

The FAST corner features are referred to as corners. Each corner corresponds to a central pixel location that, on its surrounding circle of pixels (with a certain diameter), contains an arc of  $N$  or more contiguous pixels that are all much brighter than the central pixel. A corner is also found if the all the pixels on the arc are much darker than the center. Each corner on a frame is characterized by a set of features such as coordinates, pixel arc size, the arc average color and corner color.

In order to find a match between two corners from different images, all the corners in both images are compared. The comparison takes into account all the above features, as well as each corner signature. The signature of a corner in an image is the set of distances between that corner and every other corner in the same image. This feature is computationally burdensome, but the probability of two corners having a similar signature is small.

In the comparison each feature is weighted by the reverse of the frequency of occurrence. Hence, the features that are less probable to occur in different corners are given a higher relevance for corner comparison. When the same corner is found in two different images it is said to be a matching. Two matching (two points) allows the definition of corresponding coordinate frames in each of the images.

Experimental results are presented in Fig. 2. The displacement of the coordinate frame between the two images allows computing the camera movement. To increase the precision the frame displacement is computed as an average from several matching. By calculating the motion of the robot between frames, a track of the motion can be drawn.

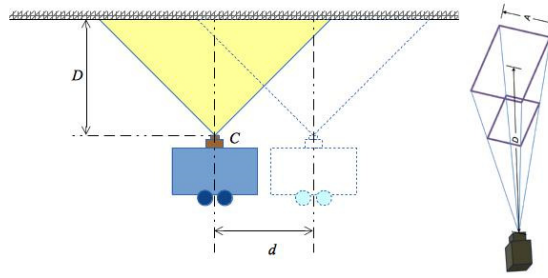


**Fig. 2.** An example of corner detection and matching – Left: Sequence of two images taken from a ceiling. All FAST corners feature points are marked as white points. Those in the middle of white circumferences are chosen for matching between frames, and black lines point out corner matching pairs. Right: Mapping of the robot motion.

Fig. 3 represents schematically the vision system geometry. As proposed by [10] the camera is located on top of the robot directed to the ceiling. It is aligned with the middle of the wheel axis line (where the IMU is also located). A known limitation stems from the fact that the pixel displacement depends not only on the robot movement, but also on the distance to the ceiling ( $D$ ). As seen in the figure, there is a linear relation between the distance  $D$  (from the robot to the ceiling) and the camera generic field of view  $A$ . Hence, it is possible to calculate  $A$  by knowing  $D$ .

At present it is assumed that this value is known in advance. In the future the camera will be integrated with a distance sensor. Precision is also affected if the camera angle is not precisely perpendicular to the ceiling. This is not a limitation as the robot is aimed to work on a horizontal surface. The use of a 3-axis accelerometer-gyroscope IMU allows the detection of strong disturbances and the correction of constant displacement angles.

In order to minimize error accumulation, the first image of the set used to measure the robot's movement is kept as long as the number of matchings is large enough to ensure precision. Only when that fails to happen a new initial image is selected.



**Fig. 3.** Vision system image acquisition geometry. C: camera; D: distance from the target (ceiling); d: robot displacement.

### 3.3 Multi-Rate Sensor Fusion

The two sensor systems present very different sampling rates. The vision system's sampling period is long and changes randomly. Due to its computation time the position measurements are time delayed. The proposed fusion scheme uses a discrete multi-rate KF [4] designed to tackle with these characteristics.

The state estimation is based on a discrete-time state-space description of the system dynamics:

$$\begin{aligned} \mathbf{Z}_{k+1} &= \mathbf{A}_k \mathbf{Z}_k + \mathbf{G}_k n_k, \\ \mathbf{Y}_k &= \mathbf{C}_k \mathbf{Z}_k + w_k \end{aligned} \quad (2)$$

where  $\mathbf{Z} = [v_x \ x \ v_y \ y \ \theta]^T$  is the system state;  $\mathbf{Y}_k$  is the vector of measured values at time  $k$ ;  $n_k$  and  $w_k$  are considered to be noise. The time varying matrices are:

$$\mathbf{A}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \Delta t_k & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \Delta t_k & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{G}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

The sampling period  $\Delta t_k$  is the interval between consecutive data fusion moments, and may not be constant. As in [4], at each fusion moment the set of available measured signals depends on their own sampling rates, therefore the output equations may vary because only available measurements are considered. To use the position data provided by the vision system it is necessary to take into account the computation delay. As proposed by [11], model (2) is used to embed the time delay of the vision measurements in the multi-rate KF.

## 4 Results

### 4.1 Simulation Results

The simulation results presented in this section consider the inertial system's sampling rate to be 100 Hz (average), with a sampling time uniformly distributed in the range from 9-11 milisecond. Inertial data samples are synchronous with the fusion moments. The vision system is simulated through a block that provides 2 position measurements per second, with a variable time delay on the range 0.38-0.42 second. Vision measurements are uniformly distributed within 5 cm of the true position. The noise and bias of the IMU's measurements are larger than the upper values taken from the manufacturer datasheet. For simulation purposes image position information is assumed to lie within  $\pm 3$ cm from the real position. The robots initial position and orientation are assumed as the origin of the real world coordinate frame. The trajectory is performed at a constant linear speed of 0.04 m/s. Different rotation rates are used along the test. Fig. 4 presents the simulated results. The robot's trajectory is tracked without significant bias. If there is any bias it does not increase with time. The position error is under 3 cm over most of the test (see Fig. 5). The deviation of the position measurement increases only when the robot performs a fast turn, with a small radius. The error is strongly affected by the vision system sample rate and delay. A faster vision system leads to smaller errors.

### 4.2 Experimental Results

The mobile robot from Fig. 1 is used to develop the position measurement system.

The measurement unit couples an inertial sensor (IMU) and upward directed camera positioned on top of the IMU. Both sensors are vertically aligned. The robot has two mechanically connected wheels on each side. The measurement of the wheel speed for both sides of the robot is available. The robot is coupled with a device to graphically mark the performed path.

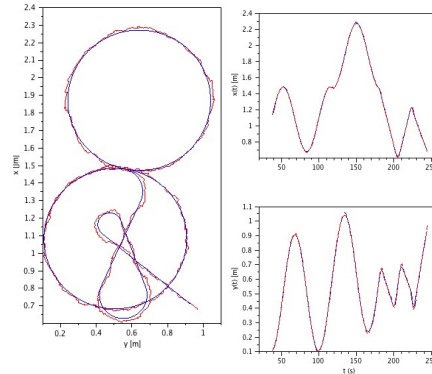


Fig. 4. Simulation results: Robot position vs. measured position. Left: Robot movement and measured position. Right:  $x$  and  $y$  measurements and true values.

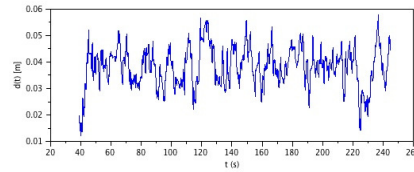


Fig. 5. Simulation results: Position error (in meters).

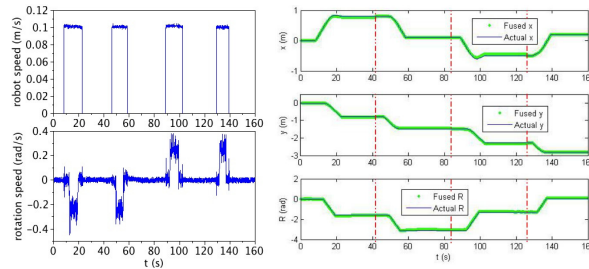
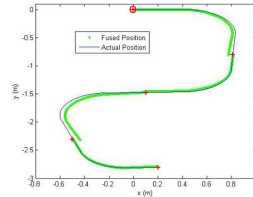


Fig. 6. Experimental results - Left: Raw measurement data. Right:  $x$ ,  $y$  and  $\theta$  estimation and actual values; vertical lines indicate when absolute position measurement are obtained.

At the present stage the robot's absolute position is determined only when it stops at a still position. In Fig. 6 are presented experimental results obtained with the position measurement system. The robot performs a trajectory with four sections. Within each section the robot moves with a linear speed of 10 cm/s and turns with an angular speed of 0.2 rad/s (approx.). At the end of a section the robot stops and absolute position is determined. A Kalman filter is used to estimate the robot's position ( $x$ ,  $y$ ) and heading ( $\theta$ ) from the IMU and wheel speed measurements. The estimates are computed every 25 ms. Absolute position measurements are also included in the estimation process when they are available. The time gap between these measures is of 40 s (approx.).





**Fig. 7.** Experimental results - Line: Robot position. Dot: position estimate. Cross: Absolute position measurement. Circle: Initial position.

## 5 Conclusions

The paper addresses the problem of measuring the position and speed of an autonomous mobile robot for indoor applications. In order to enhance the robot autonomy it is assumed that the position measurements do not require an external infrastructure. Simulation results show that proposed solution is presents a viable way to determine the robot position. The precision and sample rate adequate for most control applications.

## References

1. Kim, S. J., Kim, B. K.: Dynamic Ultrasonic Hybrid Localization System for Indoor Mobile Robots. *IEEE Trans. Ind. Elect.* 60-10, pp. 4562 - 4573 (2013)
2. Lee, T., Shirr, J., Cho, D.: Position Estimation for Mobile Robot Using In-plane 3-Axis IMU and Active Beacon. *IEEE Int Symp Ind Electronics*, pp. 1956-1961. IEEE Press, (2009)
3. Tennina, S., Valletta, M., Santucci, F., Di Renzo, M., Graziosi, F., Minutolo.: Entity Localization and Tracking: A Sensor Fusion-based Mechanism in WSNs. In: *13<sup>th</sup> IEEE 13th Int Conf on Digital Object Identifier*, pp. 983 – 988. IEEE Press, New York (2011)
4. Armesto, L., Chroust, S., Vincze, M., Tornero, J.: Multi-rate fusion with vision and inertial sensors. In: *IEEE Int Conf on Robotics and Automation*, pp. 193 - 199. IEEE Press (2004)
5. Bancroft, J. B., Lachapelle, G.: Data Fusion Algorithms for Multiple Inertial Measurement Units. *Sensors*. 11-7, pp. 6771-6798 (2011)
6. Hol, J., Schön, T., Luinge, H., Slycke, P., Gustafsson, F.: Robust real-time tracking by fusing measurements from inertial and vision sensors. *J. Real-Time Image Proc.* 2-3, pp. 149-160 (2007)
7. Cho, Bong-Su, Moon, Woo-sung, Seo, oo-Jin, Baek, Kwang-Ryul: A dead reckoning localization system for mobile robots using inertial sensors and wheel revolution encoding. *Sensors*. 11-7, 6771-6798 (2011) . *J. Mech. Sc. Tech.* 25-11, pp. 2907-2917 (2011)
8. Marín, L., Vallés, M., Soriano, Á., Valera, Á., Albertos, P.: Multi Sensor Fusion Framework for Indoor-Outdoor Localization of Limited Resource Mobile Robots. *Sensors*, 13-10, pp. 14133-14160 (2013)
9. Rosten, E., Porter, R., Drummond, T.: Faster and Better: A Machine Learning Approach to Corner Detection. *IEEE Tans. Pattern Anal. Mach. Int.*, 32-1, pp. 105-1019 (2010)
10. Lucas, A., Christo, C., Silva, M.P., Cardeira, C.: Mosaic based flexible navigation for AGVs. In: *IEEE Int Symp. Industrial Electronics*. pp. 3545 - 3550. IEEE Press, (2010)
11. Yanyan Hu, Zhansheng Duan, Donghua Zhou: Estimation Fusion with General Asynchronous Multi-Rate Sensors. *IEEE Trans. Aerosp. El. Sys.* 46-4, pp.2090-2102 (2010)