New Perspectives on Mobile Robot Navigation with Visual and Inertial Information

Jorge Dias, Inácio Fonseca, Jorge Lobo
Institute of Systems and Robotics
University of Coimbra, 3030 COIMBRA, Portugal
jorge@isr.uc.pt

Abstract

Advanced sensor systems, exploring high integrity and multiple sensorial modalities, have been significantly increasing the capabilities of autonomous vehicles and enlarging their application potential. The article describes two relevant sensors for mobile robot navigation - active vision systems and inertial sensors. Vision and inertial sensing are two sensory modalities that can be explored for navigation. This article presents our results on the use and integration of those two modalities. In a first example we present a computational solution for the problem of visual based guidance of a moving observer, by detecting the orientation of the cameras set that maximises the value of visual information. The algorithm explores the geometric properties of log-polar mapping. The second example, relies on the integration of inertial and visual information to detect the regions in the scene that we can drive a mobile platform: in our case the ground plane. The solution is based on information about the scene that could be obtained during a process of the visual fixation, complemented by the information provided by inertial sensors. The tests were performed with a mobile platform equipped with one active vision system and inertial sensors. The paper presents our recent results on simulation of visual behaviours for navigation.

1 Introduction

Navigation may be defined as the process of directing the movements of a vehicle from one point to another and a remarkable variety of physical principles and sensors have been utilised within navigational equipment. Some depend on receipt of information from somewhere outside the vehicle itself, either on or the earth or in the sky. They are therefore subject to error or in-operativeness, when such information is erroneous or is lacking, whether from natural or artificial induced sources. Others are based on internal sensing (dead-reckoning [1]), and do not explicitly depend on external references, overcoming some of their associated problems, but having others, such as drift. These systems provide relative position measurements and not absolute positioning. The present location of a vehicle is determined by advancing a previous position through known course and velocity information over a given length of time.

Vision systems are a class of sensors that can be used to obtain information from outside a vehicle. The informations obtained by these sensors are very rich and useful for vehicle navigation. In this article we propose an algorithm to detect the direction of the translation in the images when a vehicle moves in a stationary scene. In this situation the observer senses the displacement projection on the images as illustrated in Fig. 1. If the observer maintains constant its translation velocity, the pole in the front, equivalent to a focus of outflow, is stationary. For each object in the scene at equal distance from the observer, the images of objects outside the pole, move faster the more peripherally they are located in the visual field.

However if the eyes, head or body rotates (respectively the cameras, of a vision system within a mobile vehicle), a rotational motion is induced and the pole will not be in the direction of heading.

The pole **P** is a point in the sphere and can be geometrically defined by the interception of the translation vector with the image sphere surface. If, by controlling the cameras, this point **P** is maintained in the image sphere, we increase the knowledge of observer's surroundings and the information required for a safe navigation task.

Since the pole is crucial for the guidance of locomotion, it is usually kept in the center of the visual

[1] The origin of the term is "deduced reckoning" from sailing days.

0-7803-4484-7/98/$10.00

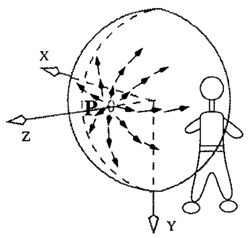


Figure 1: The optical flow field of an observer in locomotion towards the pole **P**.

field. This article reflects our efforts to implement a computational solution for this problem using, for testing, a mobile platform equipped with an active vision system. The article describes the technique used to detect the preferred direction of movement of an observer, and establishes a solution to maintain it constant with observer's motion. The algorithm detects the movement's direction and explores the geometrical properties of log-polar mapping to detect the pole **P** - see Fig. 1.

Inertial sensors are a class of sensors useful for internal sensing and they are not dependent on external references. In human and other animals the ear vestibular system gives inertial information essential for navigation, orientation or equilibrium of the body. In humans this sensorial system is crucial for several visual tasks and head stabilisation. This is a sensorial modality, which co-operates with other sensorial systems and gives essential information for everyday tasks. One example of co-operation is between the vestibular sensorial system and the visual system. It is well known that, in humans, the information provided by the vestibular system is used during the execution of navigational and visual movements, as described by Carpenter [7]. However the inertial information is also important for head-stabilisation behaviours, including the control of posture and equilibrium of the body.

This kind of sensorial information is also crucial for the development of tasks with artificial autonomous systems where the notion of horizontal or vertical is important, see [8] for one example. The inertial system prototype described in this article is used within a mobile robot equipped with an active vision system as illustrated in figure 2. The article describes a technique for inertial and vision data fusion, to identify the ground plane.

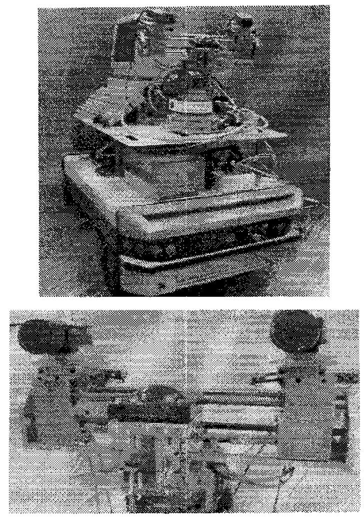


Figure 2: The mobile system with the active vision system. The inertial system prototype was designed to use with this system.

2 Direction of Translation

In visually-guided navigation, the knowledge of the Focus Of Expansion (FOE) parameter, can be used to determine and control the translation motion in space. Using this point, it is possible to estimate the system's velocity parameters or align the visual sensor with the system's translation direction velocity and maximize the visual information.

The method used to detect the FOE is based on log-polar images. Vision systems with spatially homogeneous resolution are not able to provide a real time response in a dynamically changing environment. A reactive behaviour needs selective sensing in space. Such a selection can be achieved by the combination of a space-variant resolution scheme. The field of view is split into a homogeneous high resolution area - the fovea - and the periphery with decreasing resolution. The most convincing model for the resolution decrease can be obtained by the logarithmic polar mapping. Basically, this mapping is accomplished by a transformation which maps the points measured from the center of the image proportional to the distance from the center, and also proportional to the angular aperture or resolution. This mapping is made between the image plane and the log-polar plane and the transformation between the two planes in the most simple form, is a mapping from image plane coordinates to polar coordinates, followed by the application of the loga-

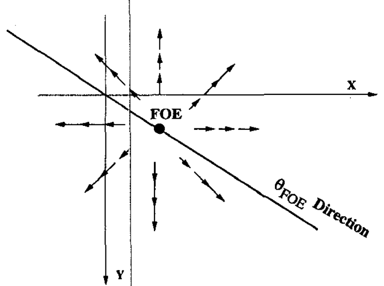


Figure 3: The geometric property of the flow vectors when the observer is moving with a translation velocity

rithmic function to the radial ρ coordinate ($\rho \in \Re^+$, $\theta \in [0..2\pi[$, $a, b > 1$). This kind of mapping reduces significantly the amount of data to handle, and we can generate some feedback signals that can be useful for visual based navigation.

In some special conditions (lower rotational velocity) the self-induced flow vectors in the image will present special geometric properties that can be useful to extract the FOE position in the image. Those geometric properties are based on the fact that the flow vectors will radiate from the FOE position, and that the line defined by the point in the center of the image coordinate system and the FOE position will present only radial flow vectors, as seen in Fig 3.

Using these facts, we can establish a computational mechanism to search the FOE position. The approach is to detect the orientation of the line (θ_{FOE} direction) by scanning every possible orientation until we find the line with the lowest circular flow. Then to detect the FOE position, we will translate the coordinate system for each point of this line, until we find the point where there is only radial flow radiating from the chosen position (this method is what we call Dynamic Retina).

In visually-guided navigation, the knowledge of the FOE parameter, can be used to determine and control our own translation motion in space.

2.1 Detecting the FOE position

To detect the FOE position, we have tested 3 processes:

- *The line algorithm* - this process just uses the flow vectors in the θ_{FOE} line, to detect where the radial flow vectors change their direction. If any point is found, that will be considered as the FOE.
- *The FFT algorithm* - explores the "Dynamic Retina" concept, by scanning every point in the θ_{FOE} line, and for each point, performing an histogram with the radial vectors. If the histogram is constant, then we are in the FOE position. The FFT is used to give a measurement of the histogram's shape.
- *The Standard Deviation algorithm* - this algorithm is very similar with the FFT algorithm, but using standard deviation calculations, to give a measurement of the histogram's shape.

2.2 Visual Alignment Control

All the experiments used a experimental set-up with a mobile robot and one active vision system. These two basic systems are interconnected by a computer that controls the movements of the active vision system and communicates with the robot's on-board computer. The computer is responsible for processing the images provided by the active vision system.

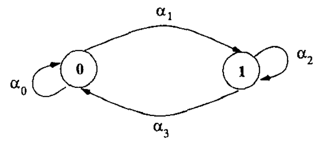
The tested the FOE detection algorithms to align the cameras with the direction of translation motion when the mobile robot moves. In the majority of cases the alignment is performed quickly in an highly textured scene.

The control system was designed as system of discrete events. The formalism used to model this kind is also used in our system. The figure Fig. 4 illustrates the different system states and transitions conditions. The control is based on the measurements obtained by the algorithms described to detect the FOE position. These measurements are signals on X and Y axis and used for camera alignment. The system uses PID controllers to align the cameras to adjust the FOE position (X,Y) in the image centre. These two variables are controlled separately.

2.3 Experiments with the mobile robot

We have tested the FOE position to be used as feedback to align a camera with the translation motion of a mobile robot in a simulated environment. In the most general cases the align as been done in a successful way, but using an highly textured scene and a small motion, to detect the optical flow. The Fig. 5 presents a sequence of images captured by the system with the system performing the camera alignment.

Alignment Control Model



$\Sigma_u = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$

State Description

0: Mobile platform moves and Vision System static

1: Vision System Control (camera alignment)

Figure 4: Discret event model for camera alignment control.

3 Visual and Inertial Sensing Integration

To study the integration of the inertial information in artificial autonomous systems that include active vision systems it was decided to develop an inertial system prototype composed of low-cost inertial sensors. Their mechanical mounting and the necessary electronics for processing were also designed. The sensors used in the prototype system include a three-axial accelerometer, three gyroscopes and a dual-axis inclinometer.

The three-axial accelerometer chosen for the system, while minimising eventual alignment problems, did not add much to the equivalent cost of three separate single-axis sensors. The device used was Summit Instruments' 34103A three-axial capacitive accelerometer. In order to keep track of rotation on the x-, y- and z-axis three gyroscopes were used. The piezoelectric vibrating prism gyroscope Gyrostar ENV-011D built by Murata was chosen. Since orientation is obtained by integration of the angular velocity over time, drift errors will occur. In the prototype, a magnetic flux-gate compass was used to overcome this problem, providing an external reference - see [9] for details. To measure tilt about the x and y-axis a dual axis AccuStar electronic inclinometer, built by Lucas Sensing Systems, was used.

To handle the inertial data acquisition, and also enable some processing, a micro-controller based card was built. This card has analogue filters, an A/D converter as is based on Intel's 80C196KC micro-controller. The robot's master processing unit has an EISA bus interface, where the card is connected

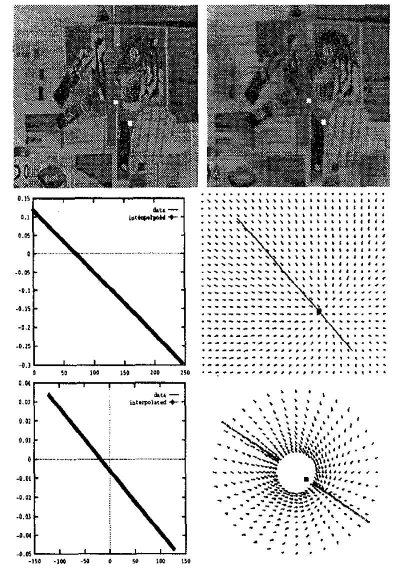


Figure 5: Real data experiments. The movement is $\vec{v} = [10.0, 10.0, 4.0]^T$ and $\vec{\omega} = [0,0,0]^T$,$f_x = f_y =$ 15.0. The real FOE in this case is in $(37.5, 37.5)$. At the top is the sequence with the white points marking the FOE position. The middle images represent the detection in the Cartesian plane. At the bottom is the detection in the Log-Polar plane.

along with another for image acquisition and processing. This card is a frame grabber module with two Texas Instruments TMS320C40 DSPs that handles the video processing.

Figure 6 shows the architecture of the system and the computer that supervises the active vision, moving platform and inertial system. The inertial sensors were mounted inside an acrylic cube, enabling the correct alignment of the gyros, inclinometer (mounted on the outside) and accelerometer, as can be seen in the close up of figure 2. This cube is connected to, and continuously monitored by, the micro-controller card in the host computer. The inertial unit is placed at the middle of the stereo camera baseline. The head coordinate frame referential, or Cyclop $\{C_y\}$is defined as having the origin at the centre of the baseline of the stereo cameras.

The inclinometer data can be used to determine the orientation of the ground plane. In order to locate this plane in space at least one point belonging to the ground plane must be known. When the vehicle is stationary or subject to approximately constant speed the inclinometer gives the direction of $\vec{g}$ relative to

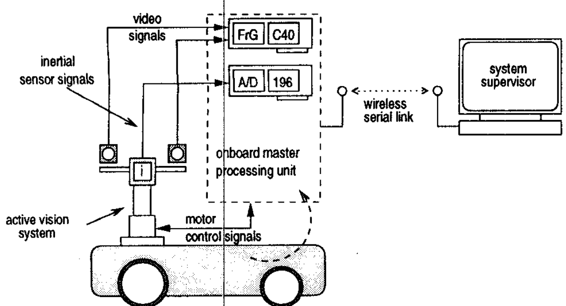


Figure 6: System Architecture. The inertial system processing board uses the Master processing unit as host computer.

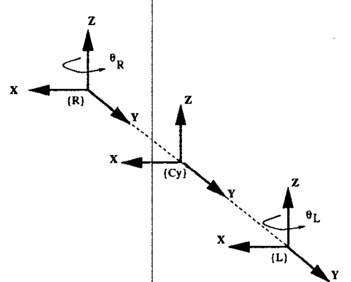


Figure 7: System Geometry

the Cyclop referential $\{C_y\}$. Assuming the ground is levelled, and with α_x and α_y being the sensed angles along the x and y-axis, the normal to the ground plane will be

$$\hat{n} = -\frac{\vec{g}}{||\vec{g}||} = \frac{1}{\sqrt{1-\sin^2\alpha_x\sin^2\alpha_y}}\begin{bmatrix} -\cos\alpha_x\sin\alpha_y \\ -\cos\alpha_y\sin\alpha_x \\ \cos\alpha_y\cos\alpha_x \end{bmatrix} \quad (1)$$

given in the Cyclop frame of reference. Using this inertial information the equation for the ground plane will be given by

$$\hat{n}.\vec{P}_f + h = 0 \quad (2)$$

where $\vec{P}_f$ is a point in the plane and h is the distance from the origin of $\{C_y\}$ down to the ground plane.

To obtain a point belonging to the ground plane it will be necessary to find the correspondence between points in the image. Establishing this correspondence will give us enough equations to determine the 3D co-ordinates, if a few vision system parameters are known. However if visual fixation is used, the geometry is simple and the reconstruction of the 3D fixated point is simplified, as can be seen in figure 8. Notice that the visual fixation can be achieved by controlling the active vision system and the process allows a fast and robust 3D reconstruction of the fixation point. This mechanism was developed in the ISR laboratory and is described in [10] and [11]. If the active vision system fixates in a point that belongs to the ground plane, the ground plane could be determined in the Cyclop referential $\{C_y\}$ using the reconstructed point $\vec{P}_f$ and the inclinometer data. Hence, any other correspondent point in the image can be identified as belonging or not to the ground plane.

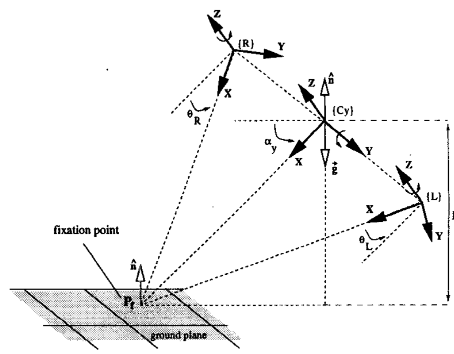


Figure 8: Ground plane point fixated. The point $\vec{P}_f$ in the ground plane is visualised by the active vision system. The geometry of this visualisation corresponds to a state, named visual fixation.

Figure 9 shows a pair of stereo images where fixation was obtained for a ground plane point. In this example α_x is null and $\alpha_y = 16.05^o$, $\theta = 2.88^o$ and $b = 29.6cm$. Making $h = b.\cot(\theta)sin(\alpha_y)/2 \simeq 81,3cm$. The points shown in figure 9 were interest points obtained with a corner detector. These points are tested by an algorithm that classifies them as belong or not, to the ground plane. Any point in the ground plane verifies a geometric constraint that could between established between two stereo images, easily obtained from the stereo vision geometry. The ground plane is thus determined. Figure 9 shows the matched ground plane points of interest.

4 Summary and Conclusions

Navigation plays an important role in all mobile robot's activities and tasks. For navigation the mobile robots can use external sensors (ex. vision systems) or internal sensors (ex. inertial sensors). The former allows position correction with respect to structures

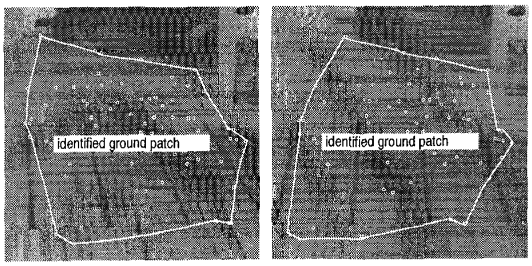


Figure 9: Detected ground points.

or landmarks. Vision systems are sensors that can be used to obtain information from outside a vehicle. The information obtained by these sensors is very rich and useful for vehicle navigation. The process is similar to an artificial observer. Any artificial observer, with forward facing cameras and when moves with translation velocity among stationary scene, senses the projection of the displacement in the images. If the observer maintains constant its translation velocity, a pole in the front, equivalent to a focus of outflow, will be stationary. This point is defined by the interception of the translation motion vector with the image surface. If, by controlling the cameras, this point is maintained in the image, we increase the knowledge of observer's surroundings and the information required for a safe navigation task. Since the pole is crucial for the guidance of locomotion, it is usually kept in the center of the visual field. This article reflects our efforts to implement a computational solution for this problem. The algorithms to detect the FOE position are based on log-polar images.

Inertial sensors are based on internal sensing and do not depend explicitly on external references, overcoming some of their associated problems. The integration of inertial systems in mobile robots opens a new field for the development of applications based or related with inertial information. Here we demonstrated one of these applications in an autonomous system for ground floor detection.

References

[1] Jorge Dias, Helder Araújo, Carlos Paredes, Jorge Batista *Optical Normal Flow Estimation on Log-Polar Images. A Solution for Real-Time Binocular Vision*, Real-Time Imaging, Vol 3, pp.213-228, 1997.

[2] Inácio Sousa, Jorge Dias, Helder Araújo, *Normal Optical Flow Based on Log-Polar Images*, 4th International Symposium on Intelligent Robotics Systems, Lisbon,22-26 July,1996.

[3] Rami Guissin, Shimon Ullman, *Direct Computation of the Focus of Expansion From Velocity Field Measurements*, IEEE Transactions,1991,pp. 146-155.

[4] Michal Irani, Benny Rousso and Shmuel Peleg, *Recovery of Ego-Motion Using Image Stabilization*, IEEE Transactions,1994,pp. 454-460.

[5] J. Santos-Victor, G.Sandini, F.Curotto, S.Garibaldi, *Divergent Stereo in Autonomous Navigation: From Bees to Robots* VisLab-TR 01/95, International Journal of Computer Vision, 14, pp 159-177, March 1995.

[6] V. Sundareswaran, P. Bouthemy and F. Chaumette, *Visual servoing using dynamic image parameters*, INRIA, RR n^o 2336 -Août 1994.

[7] H. Carpenter, *Movements of the Eyes*, London Pion Limited, 2nd edition, 1988, ISBN 0-85086-109-8.

[8] T. Viéville and O.D. Faugeras, "Computation of Inertial Information on a Robot", *Fifth International Symposium on Robotics Research*, ed. H. Miura, S. Arimoto, pg. 57-65, MIT-Press,1989

[9] Jorge Lobo, Paulo Lucas, Jorge Dias, A. T. de Almeida, "Inertial Navigation System for Mobile Land Vehicles", *Proceedings of the 1995 IEEE International Symposium on Industrial Electronics*, 1995, July, Athens, Greece, pg. 843-848.

[10] Jorge Dias, Carlos Paredes, Inácio Fonseca, H. Araujo, J.Batista, A. T. de Almeida, "Simulating Pursuit with Machines - Experiments with Robots and Artificial Vision", *IEEE Trans. on Robotics and Automation* , Feb. 1998, Vol 14, No1, pg. 1-18.

[11] Carlos Paredes, Jorge Dias, A. T. de Almeida, "Detecting Movements Using Fixation", *Proceedings of the 2nd Portuguese Conference on Automation and Control*, 1996, September, Oporto, Portugal, pg. 741-746.

[12] Arlindo Silva, Paulo Menezes and Jorge Dias, "Avoiding Obstacles Using a Connectionist Network", *Proceedings of the 1997 IEEE International Conference on Intelligent Robots and Systems (IROS'97)*, 1997, pg. 1236-1242.