

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Topological Mapping and Navigation using a Developmental Learning Approach based on Imitation through Sensory-motor Maps

Raquel Frizera Vassallo¹, Hans Jörg Andreas Schneebeli¹
and José Santos-Victor²

¹*Universidade Federal do Espírito Santo and* ²*Instituto Superior Técnico*
Brazil and Portugal

1. Introduction

Over the past decades, we have witnessed a massive utilization of robots in industrial settings. The use of robotic manipulators with high precision and repeatability were fundamental for the increase of efficiency of many manufacturing processes and mass production. Due to the complete predictability of the factory environment, each robot is programmed a priori, and repeats the same set of movements for weeks or even months. In spite of the limited sensory information available, these robots can complete their tasks in an extremely efficient manner. Recent progress in computer vision, robot navigation, estimation theory, reasoning and computing have afforded engineers with the tools to build more flexible robots able to “live” in more unpredictable environments and interact with (non-technical) humans in a rich and sophisticated manner. In the long-run, this research effort will contribute to the ubiquitous presence of robots in our homes and offices to assist, entertain and work together with humans. For this to become true, there are two very challenging questions that need to be addressed. Firstly, there is the issue of complexity, as such systems are likely to be much more complex than industrial robots with regard to the number of motor and perceptual degrees of freedom and will have to operate in a highly unpredictable environment. Secondly, there is the question of programming or commanding these robots, as they will interact over extended periods of time with non-technical users, who are not expected to write sophisticated computer programs each time the robot is required to perform a new task. As some of these challenges are common to many biological social species, we will borrow inspiration from biology, developmental psychology and neuroscience in our approach.

Many animals can be seen as complex and flexible systems, able to learn and adapt through a simultaneous motor, sensorial and cognitive development. Studies in developmental psychology clearly show how human infants progressively learn how to acquire more sophisticated skills over time, through interaction with their own body, the environment and caretakers. At birth, many motor and perceptual capabilities are not yet in place and the acquisition of some skills is a pre-requisite to acquire more complex ones. It would otherwise be very difficult to handle the overwhelming flow of sensory-motor information.

In this work, we adopt the approach of developmental robotics Fitzpatrick et al. (2003); Lungarella & Pfeifer (2001); Metta, Sandini, Natale, Manzotti & Panerai (2001) to build systems

able to acquire complex sensory, motor and cognitive skills, in a progressive manner, guided through different levels of interaction with the environment.

With regard to the second challenge of “programming” such complex robotic systems without the need to explicitly writing computer programs, we can again see how knowledge and skills are transferred amongst individuals of some animal species. Imitation seems to be a powerful social learning and adaptation modality in social species, Dautenhahn (1995); Dautenhahn & Nehaniv (2002). Imitation avoids the need to undergo through extensive trial and error, since the imitator can learn directly from the teacher’s experience.

Similarly, endowing a robot with the ability to imitate tasks performed by a demonstrator could become an intuitive and yet extremely powerful way of “programming” such robots by non-technical users. Learning by imitation has been addressed by several researchers with interesting results for task learning, skill acquisition and communication Billard (2002); Billard & Hayes (1997); Matáric (2002). Other works in robotics inspired by imitation in animals can be seen in Dautenhahn & Nehaniv (2002).

It is clear how imitation can be a powerful learning mechanism and how it can avoid having to pre-program the robot for every new task. Then, the developmental approach would define the set of skills the robot needs to acquire, in an incremental manner, modulated by the interaction with the environment, objects and people.

Recent findings in neuroscience have shed new insight as to the brain mechanisms possibly implied in imitative behavior. We are referring to the discovery of the *mirror neurons* in area F5 of monkey’s brain Fadiga et al. (2000); Gallese et al. (1996), that fire during the execution of a goal directed (grasp) action by the animal as well as with the observation of similar actions performed by others Rizzolatti & Arbib (1998); Rizzolatti et al. (2002).

These findings suggest that the recognition of actions may be facilitated by the ability to execute that same action, the same brain circuitry being used for both purposes. In addition, these results show the importance of motor information for perception and action understanding tasks. It indicates that recognition is done with motor data as opposed to visual information. In our work, we explore the inspiration from mirror neurons and rely on motor representations for our robot to execute or recognize “actions”. We will show how this becomes a very natural way of representing, learning and executing robotic tasks. The fact that mirror neurons are located in the motor area of the brain, suggests that observed actions are first “converted” to motor information, before recognition may take place.

The existence of a visuo-motor mapping that converts visual information into motor measurements is also supported by biology. When newborns look to their hands and own movements, they are probably learning the relationship between motor action and visual stimuli (sensory-motor coordination). Through this visuo-motor mapping, children become able to recognize and repeat movements. Some works have explored these visuo-motor mappings for learning how to grasp objects with manipulators, Fitzpatrick et al. (2003); Lopes & Santos-Victor (2003a;b); Metta, Sandini & Natale (2001).

In this work we propose an approach for building robots able to learn and adapt in an open-ended manner, while interacting with humans. The approach is based on three main ideas: (i) use of artificial ontogenetic development as a means for the robot to acquire complex skills under controlled complexity, (ii) use of imitation learning as the main form of “programming” such robots and (iii) exploiting the use of motor representations for action recognition, learning and executions. The proposed developmental architecture is illustrated in Figure 1.

The first stage in this developmental pathway is that of sensory-motor coordination or learning sensory-motor maps. Sensory-motor maps relate motor actions to sensor (visual, propri-

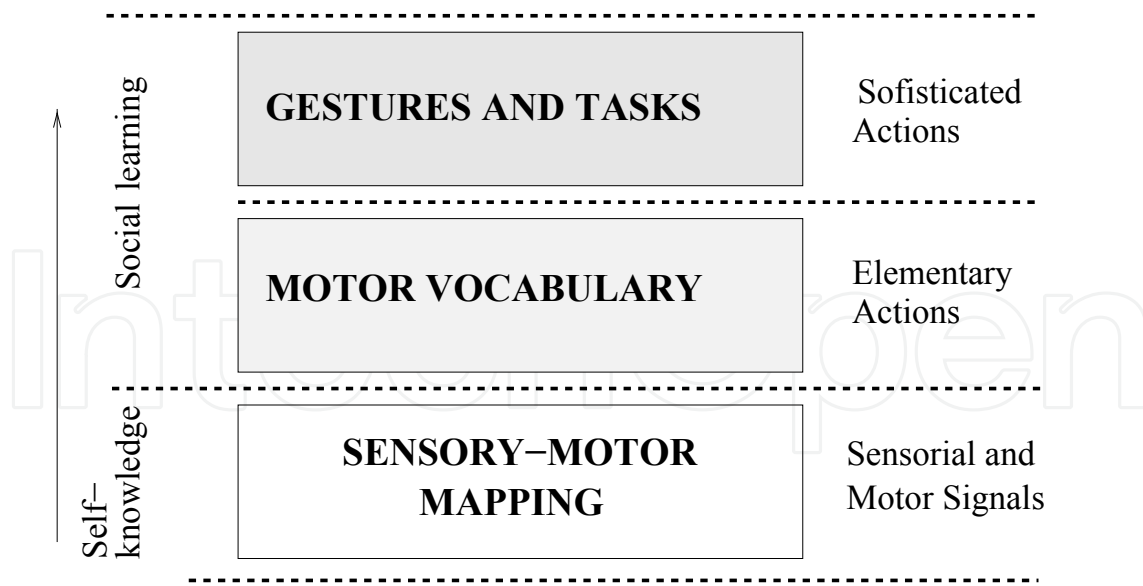


Fig. 1. Robot learning and adaptation approach.

ceptive or auditive) percepts and vice-versa. The forward map allows to predict the new sensory observations resulting from a certain motor action (e.g. the arm configuration in the retina arising from the activation of a set of muscles). Conversely, the inverse map is useful to determine which motor actions need to be taken to produce given sensory percept (e.g. determine the muscle activations to reach a certain point in space) or to recognize movements from sensory information. Figure 2 illustrates the use of the forward and inverse versions of sensory-motor maps.

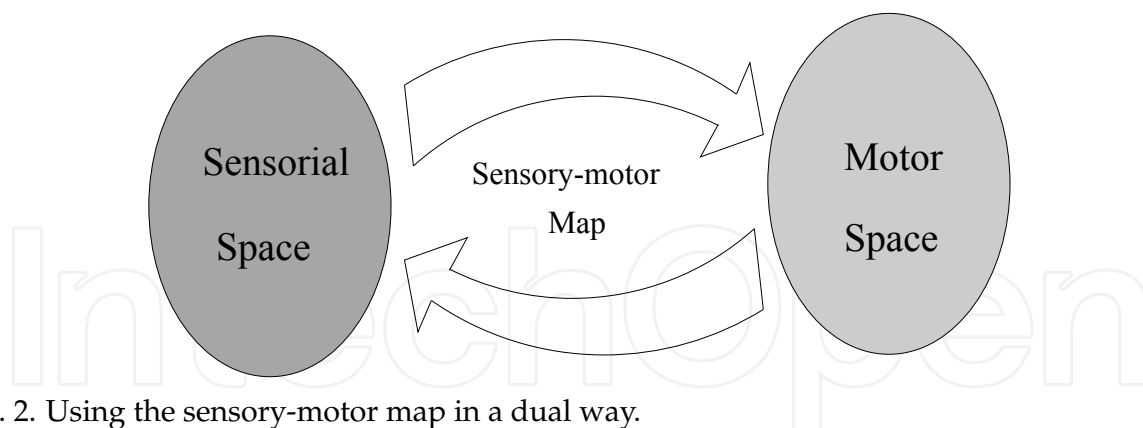


Fig. 2. Using the sensory-motor map in a dual way.

Sensory-motor maps can either be computed analytically (whenever calibration information is available) or learned directly from data through the observation of the robot's own movements.

Once the sensory-motor maps are in place, the system can move to the next developmental stage and learn a repertoire (or a vocabulary) of elementary motor actions that require the coordination of various degrees of freedom. It is important to notice that this repertoire results from two types of constraints. It is tightly related to the robot's own kinematic constraints but, more importantly, it is modulated by the interaction with the environment (co-development) and other agents (either humans or robots).

Hence, interaction allows the robot to organize its own (low-level) movements and learn task-specific actions that are not defined a priori. These actions are coded according to the robot's motor capabilities (using the learned sensory-motor maps) and this approach can thus be applied to different robots, working in different environments and addressing different tasks. Whenever the application changes, a new set of actions can be taught. There is no need to program every action needed to accomplish a task.

At this level, the robot represents actions in a symbolic manner, while the fine motor control is handled by the sensory-motor maps previously learned. This hierarchy of movements/actions follows a close parallelism with the central and peripheral nervous system of many animals, where the brain triggers high-level actions and the spinal cord is responsible for the individual muscle activation (e.g. in locomotion).

In the next stage of development, the system will be able to learn how to use this action vocabulary to solve specific tasks. The idea is that the system can observe actions performed by others and map the observed actions to those it knows how to perform (the action vocabulary). Usually, this requires combining these elementary actions in a task-specific manner, to solve different day to day tasks.

This strategy for learning and adaptation is a very flexible way of mapping the desired task (and the way it should be performed), to the robot's motor capabilities. Different robots can learn the same action vocabulary and different vocabularies can be taught if very different environments or task domains are required.

1.1 Application to mapping and navigation

In order to illustrate the approach applicability, we chose the task of topological mapping and navigation for a mobile robot, where vision is the primary sensor modality.

The level of sensory-motor coordination corresponds to the system's ability to relate visual stimuli (image motion) to motor/proprioceptive data (egomotion). The process to estimate the motor information from image flow is based on the particular geometry of an omnidirectional camera and benefits from the use of enlarged fields of view.

In the second level of development, the system relies on vision to follow a person. Through this stage, an egomotion estimation process is utilized to convert visual motion to motor data. In the motor space, the system will learn how to categorize the most frequent actions. Those actions result both from the system's kinematics as well as the shape of the environment and the guidance of the human assistant. For example, if the system would be taught in a world composed of circular hallways, then rotational motion would be very dominant in the action vocabulary.

In the final level of development, the system engages in social learning through imitation. The idea is that one person can now act as a guide and show the workspace to the robot. By following the person, the robot will then be able to create a topological map of the environment, expressed in terms of the motor vocabulary. Then, for a navigation task like "going to a place B", the system will just need to localize itself using the visual information and invoke the motor programs (actions) to move to the next sub-goal (e.g. turn right, go ahead, etc). The implementation steps are illustrated in Figure 3.

By using the proposed learning and adaptation approach, the robot performs mapping and navigation without the need for programming every action required for the task. The basic actions are learned while interacting with an user. Later on, these actions are used for mapping and navigating. Besides illustrating the proposed methodology, the task of topological mapping has not been addressed in many previous works done on imitation.

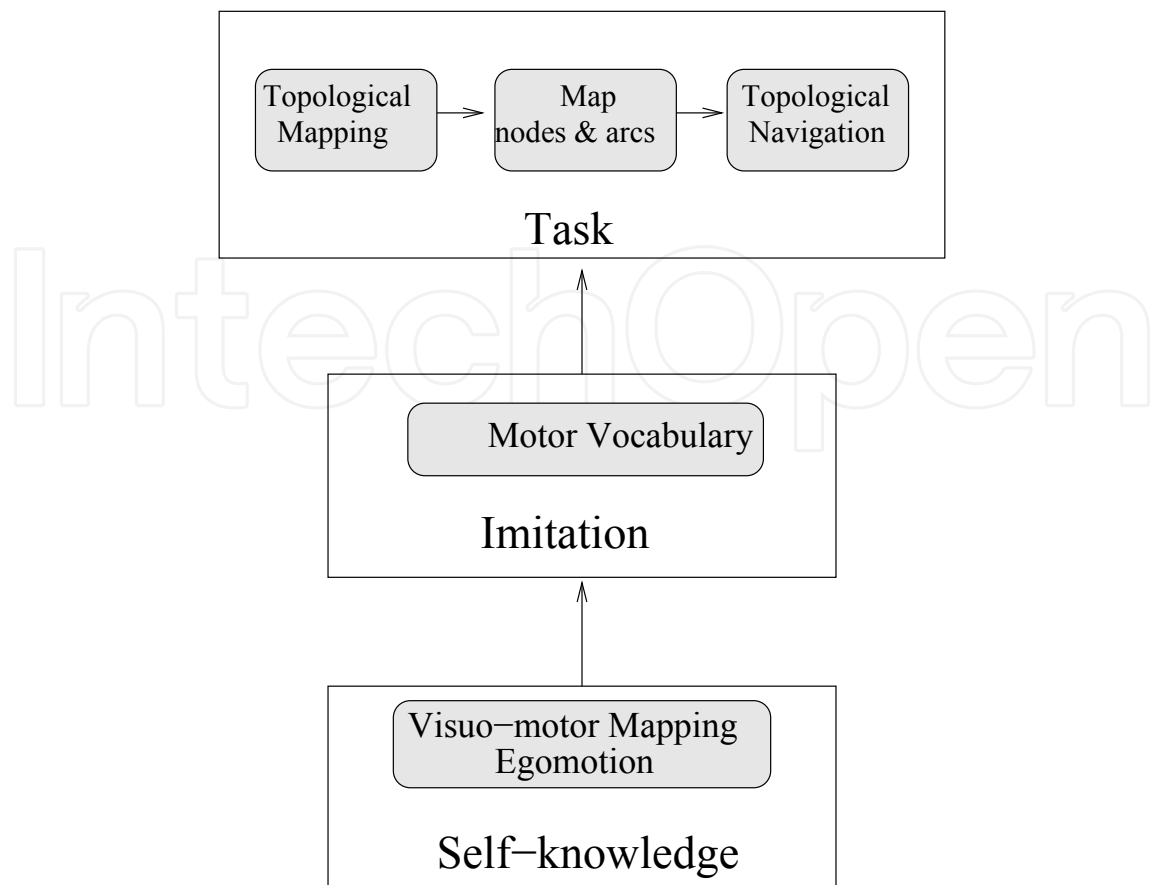


Fig. 3. Topological mapping and navigation application.

1.2 Structure of this chapter

In Section 2 we present the motor and sensory spaces defined for the used robot and the proposed egomotion estimation process responsible for the visuo-motor mapping. Sections 3 and 4 describe how to learn the motor vocabulary and the mobile robot application. Experiments and results are listed in Section 5 while our conclusions and future work are discussed in Section 6.

2. Visuo-motor coordination: egomotion estimation

As we discussed previously, the goal of the visuo-motor coordination is to build maps that relate visual stimuli to motor data. In the task of mapping and navigation for a mobile robot, this corresponds to estimating the camera egomotion from optical flow measurements obtained from a sequence of images.

The robot used in this work is a differential platform, capable of moving in the ground plane and rotating about the Z-axis by receiving motor commands for the linear velocity T_y and angular velocity ω_z . Thus its motor space can be defined by its linear velocity on the XY plane and the angular velocity about the Z-axis.

The robot is equipped with an omnidirectional camera aligned with the platform's rotation axis. It is a catadioptric system formed by a B&W camera and a spherical mirror Baker & Nayar (1998). The robot also has a color (perspective) camera pointing to the forward direction,

that is used for the person-following behavior. The robot and both vision systems can be seen in Figure 4.

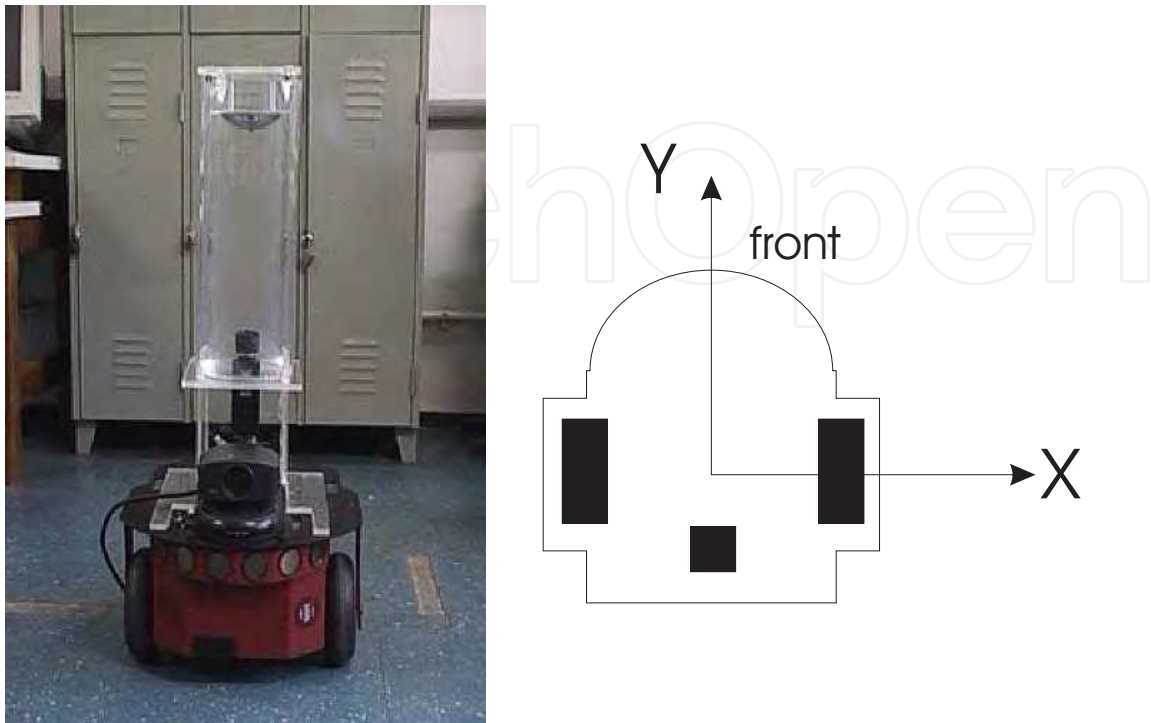


Fig. 4. The robot, the vision systems and the adopted robot reference frame.

The sensory space is defined by spherical optical flow measurements, calculated from omnidirectional images and used for egomotion estimation. In what follows, we will show how to obtain a spherical optical flow from a sequence of omnidirectional images, and how the visuo-motor mapping converts visual information into motor data.

The reason for using omnidirectional images for egomotion estimation is that the problem becomes easier if a spherical motion field is used instead of a planar field obtained with perspective cameras Nelson & Aloimonos (1988). We start by calculating the optical flow field from a sequence of omnidirectional images. Then, the image flow vectors are remapped to the surface of the unit sphere, through an image Jacobian matrix. On such hemispherical motion field, we know that either the focus of expansion (FOE) or the focus of contraction (FOC) must be visible. Finally, motor information is estimated from the motion field adapting an egomotion algorithm designed for planar projection to spherical projection Vassallo et al. (2002a).

In previous works, the Jacobian matrix needed to remap image flow vectors was defined according to the system projection model Gluckman & Nayar (1998). Instead, we use a general projection model defined by Geyer & Daniilidis (2000a;b) to define a general Jacobian. This projection model can represent different omnidirectional systems (with a single projection center) by combining a mapping of a 3D point P to a sphere, followed by a projection to the image plane. The center of the sphere $C(0,0,0)$ lies on the optical axis of the projection to the plane and represents the origin of the reference frame. The general projection model is illustrated in Figure 5.

The parameters l and m can be adjusted to model different camera types (mirror shapes) and correspond to the distances from the sphere center C to the projection center O and to the

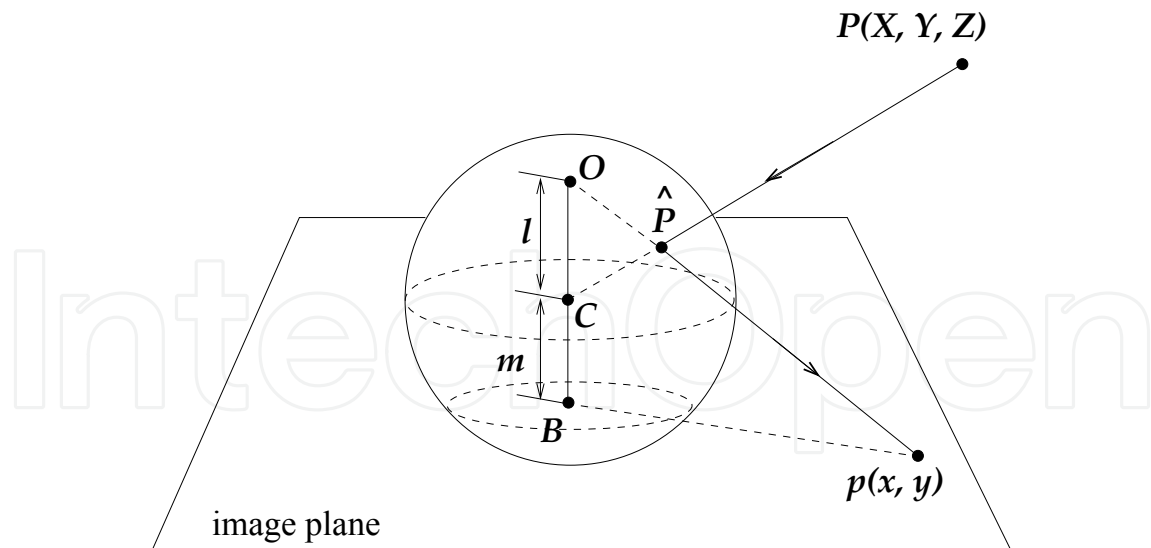


Fig. 5. The general projection model for single projection center cameras.

projection plane. According to the model, the image projection $p(x, y)$ of a 3D point $P(X, Y, Z)$, can be defined by:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{l+m}{lR-Z} \begin{bmatrix} X \\ Y \end{bmatrix} \quad \text{with} \quad R = \sqrt{X^2 + Y^2 + Z^2} \quad (1)$$

Back-projecting an image point (x, y) , we can obtain a point on the unit sphere $\hat{P}(\hat{X}, \hat{Y}, \hat{Z})$ corresponding to the direction of the incoming ray from the original 3D point $P(X, Y, Z)$ Gaspar et al. (2001). The back-projection is described by Equation 2.

$$\begin{bmatrix} \hat{X} \\ \hat{Y} \end{bmatrix} = \frac{lA + \text{sign}(A)\sqrt{(x^2 + y^2)(1 - l^2) + (A)^2}}{x^2 + y^2 + (A)^2} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$A = l + m \quad \hat{Z} = \pm \sqrt{1 - \hat{X}^2 - \hat{Y}^2} \quad (2)$$

where \hat{Z} becomes negative if $|l+m|/l > \sqrt{x^2 + y^2}$ and positive otherwise. Note that the camera intrinsic parameters, image center and focal length are not considered in the above expression.

To reproject flow vectors from the image plane to the sphere surface, a general *Jacobian* matrix J is defined by differentiating the spherical coordinates $(\hat{X}, \hat{Y}, \hat{Z})$ on the back-projection equation with respect to the image coordinates (x, y) Vassallo et al. (2002a). It maps image velocity vectors to the unit sphere surface, transforming a planar flow field to a hemispherical motion field that will help estimate egomotion (see Equation 3 and Figure 6).

$$\begin{bmatrix} \dot{\hat{X}} & \dot{\hat{Y}} & \dot{\hat{Z}} \end{bmatrix}^T = J \begin{bmatrix} \dot{x} & \dot{y} \end{bmatrix}^T \quad \text{with} \quad J = \begin{bmatrix} \frac{\partial \hat{X}}{\partial x} & \frac{\partial \hat{Y}}{\partial x} & \frac{\partial \hat{Z}}{\partial x} \\ \frac{\partial \hat{X}}{\partial y} & \frac{\partial \hat{Y}}{\partial y} & \frac{\partial \hat{Z}}{\partial y} \end{bmatrix}^T \quad (3)$$

Egomotion is estimated by adapting the Bruss & Horn (1983) algorithm designed for planar perspective projection to spherical projection. The motion field U at a point \hat{P} on the unit

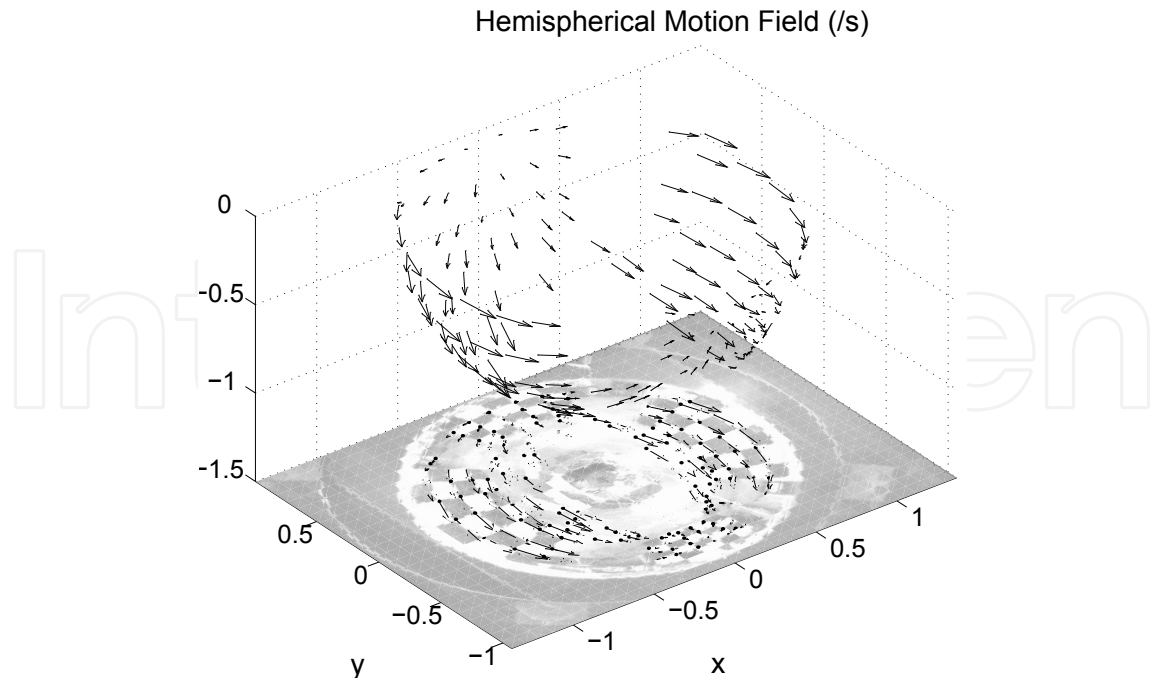


Fig. 6. Image velocities remapped to the unit sphere surface.

sphere is a function of the camera rotation Ω , translation T and the corresponding 3D point depth $R = \sqrt{X^2 + Y^2 + Z^2}$ (Eq. 4).

$$U(\hat{P}) = \frac{1}{R}((T \cdot \hat{P})\hat{P} - T) - \Omega \times \hat{P} \quad (4)$$

Depth dependency is removed by taking the cross product with \hat{P} and the dot product with T , resulting in Equation 5. Estimation was done through an iterative process using non-linear minimization considering $|T| = 1$, since the linear velocity can only be determined up to a scale factor.

$$T \cdot (\hat{P} \times (U + (\Omega \times \hat{P}))) = 0 \quad (5)$$

The process for egomotion estimation that we have just described can be interpreted as a visuo-motor map that converts visual information to motor measurements, $T(T_x, T_y, T_z)$ and $\Omega(\omega_x, \omega_y, \omega_z)$. It is important to stress that this process is intimately related to the robot's motor and visual capabilities.

Some egomotion results obtained with the mobile robot used for the experiments are shown in Figure 7. Image flow vectors were calculated using Lucas & Kanade (1981) method and then remapped to the unit sphere by the general Jacobian (Equation 3). The first example corresponds to a translation, the second is a rotation and the third a combined motion, translation plus rotation. The robot can translate in the XY plane (the ground plane) and rotate around the Z -axis. The vectors (T_x, T_y) and ω_z obtained by egomotion estimation can also be seen in Figure 7.

Table 1 presents estimated values and errors calculated by comparing the results with the nominal values. For the translation vectors, just the error on direction was computed, once it is not possible to recover the absolute velocity values from egomotion estimation. For the angular velocity, the errors were measured along the Z -axis in $^\circ/s$.

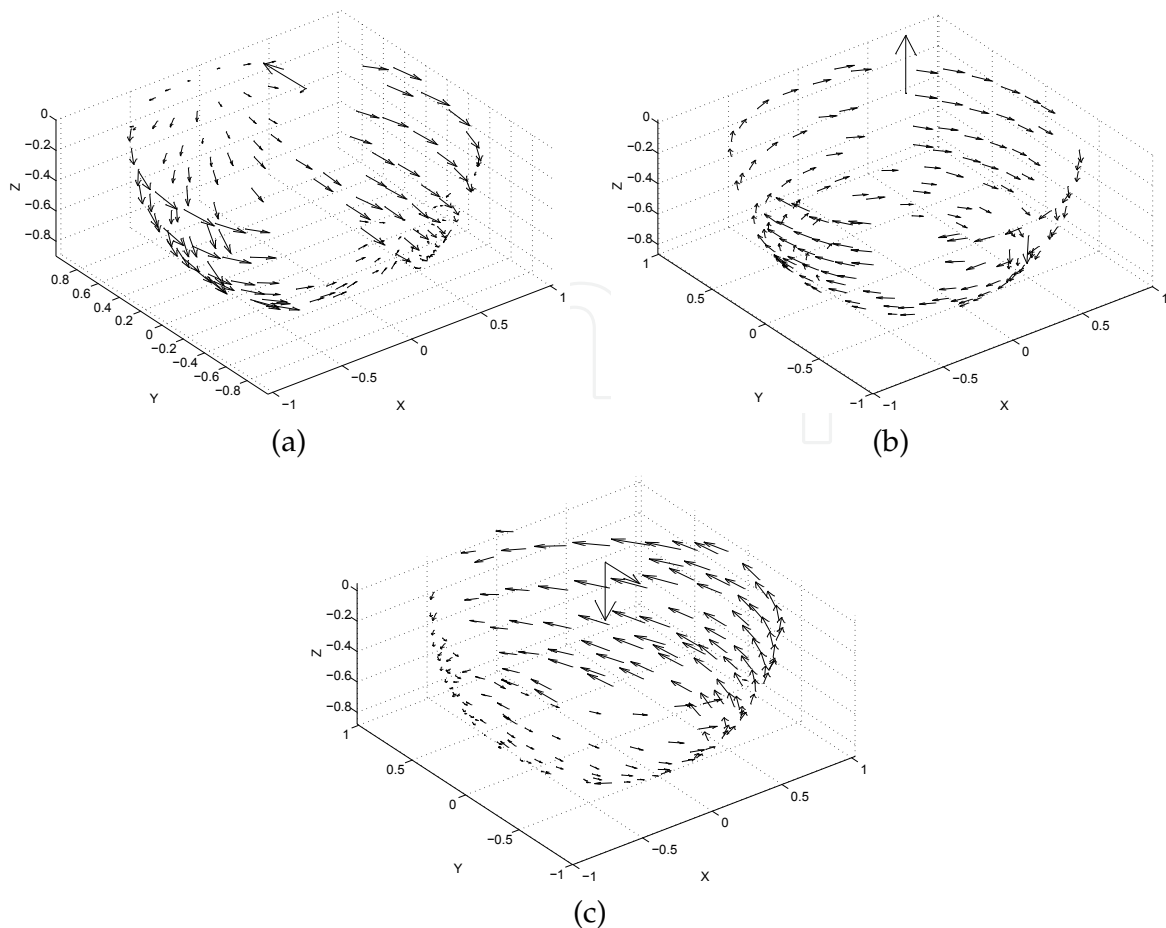


Fig. 7. Examples of (a) Translation, (b) Rotation and (c) Combined move. Egomotion vectors (T_x, T_y) and ω_z indicated in the middle of the hemispherical flow.

	Translation	Rotation
(a)	$T (T_x \ T_y)$	$\Omega (\omega_z) (^\circ/s)$
nominal	$T = [0 \ 1]$	$\Omega = [0]$
estimation	$\hat{T} = [0.0128 \ 0.9999]$	$\hat{\Omega} = [-0.0132]$
error	$e_t = 0.734^\circ$	$e_\Omega = [-0.0132]$
(b)	$T (T_x \ T_y)$	$\Omega (\omega_z) (^\circ/s)$
nominal	$T = [0 \ 0]$	$\Omega = [-3.1255]$
estimation	$\hat{T} = [0 \ 0]$	$\hat{\Omega} = [-3.4965]$
error	$e_t = 0^\circ$	$e_\Omega = [-0.371]$
(c)	$T (T_x \ T_y)$	$\Omega (\omega_z) (^\circ/s)$
nominal	$T = [0 \ -1]$	$\Omega = [3.8776]$
estimation	$\hat{T} = [0.0252 \ -0.9997]$	$\hat{\Omega} = [4.3284]$
error	$e_t = 1.45^\circ$	$e_\Omega = [0.4508]$

Table 1. Egomotion estimations/errors.

A set of 608 different movements were executed for testing the egomotion method used as visuo-motor mapping. All the results were compared to the nominal values and the errors

analysed. The mean error found for the direction of translation was 1.44° and for the angular velocity was $0.56^\circ/s$. The errors for translation and rotation are shown in Figures 8 and 9.

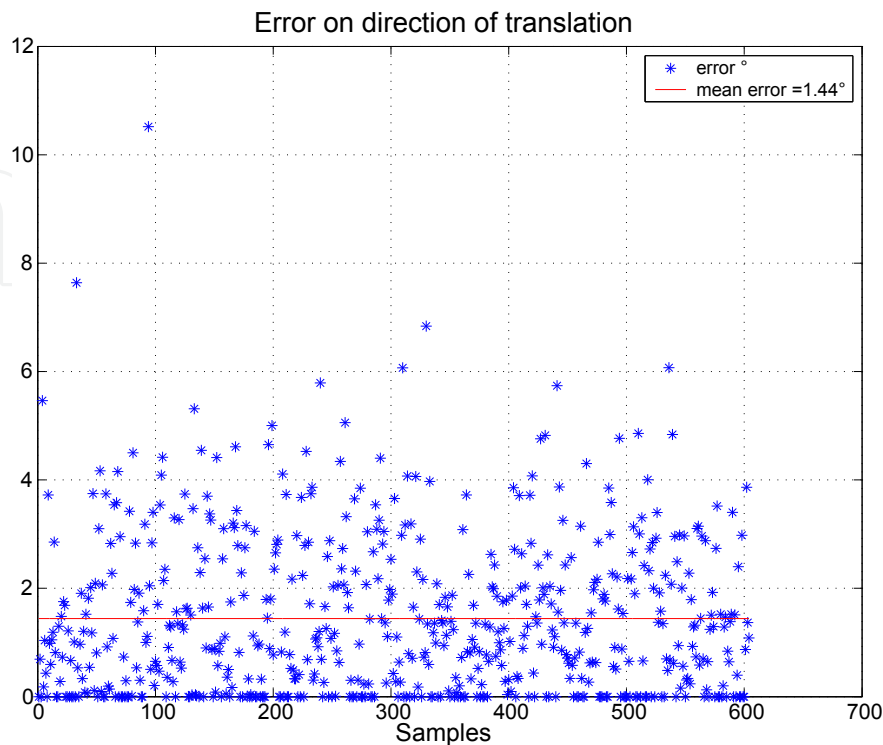


Fig. 8. Errors on direction of translation.

Although the translation and rotation estimation errors were in the order of 5° and $2^\circ/s$ for some experiments, the results were satisfactory for the envisaged application. The mean errors do not have a large impact on the direction of translation and on the angular velocity estimation, when the robot travels small distances (instantaneous measurements).

At this point, we have defined the nature of the visuo-motor map that will be used throughout the remainder of the chapter. It allows the system to convert purely visual information to its motor variables, the egomotion.

3. Learning a purposive motor vocabulary through imitation

The first stage of development endows the system with the ability to map the motion of the visual world onto the robot's motor variables. In this second level of development, this motor information will be organized onto more complex actions. From that point onwards, such actions can be used to solve more sophisticated tasks, without the need to consider the details of motion execution.

The way motor signals are organized to build a vocabulary of actions depends on the environment where the robot usually operates and will be modulated by the typical tasks it will execute. Hence, we rely on imitation, in the form of a person-following mechanism, to drive this process.

The idea is that the robot is guided through the workspace by a human operator. The correspondence problem for imitation was solved by a person-following behavior. While following the guide, the robot continuously maps the observed visual motion to its own motor variables

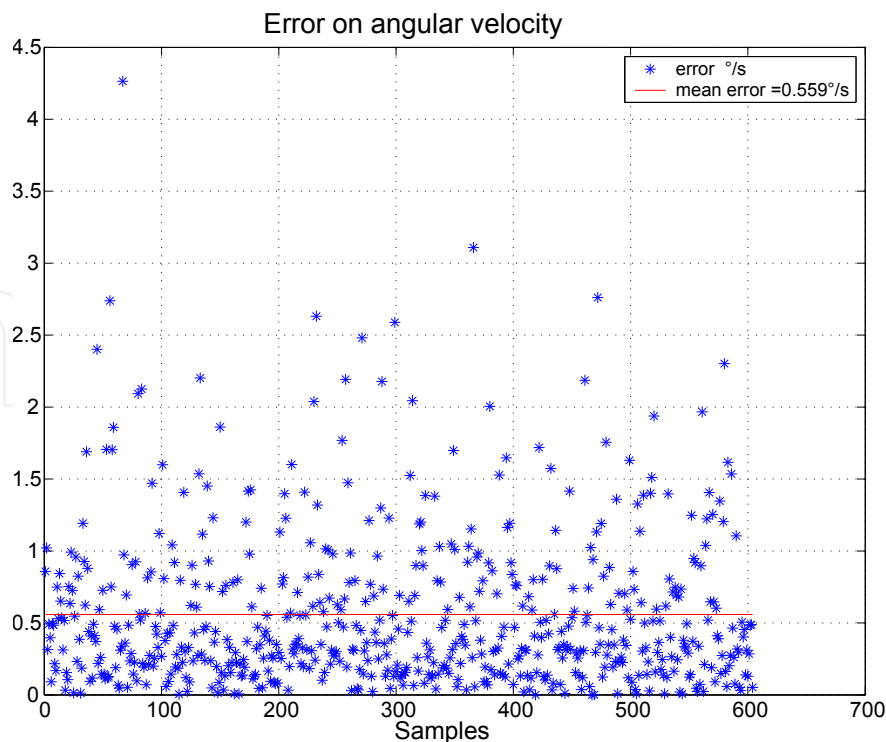


Fig. 9. Errors on angular velocity.

using the visuo-motor map (egomotion estimation). In this motor space, one can identify what are the most typical executed movements to create a basis of elementary actions.

For simplicity, we assume that the person-to-follow carries a distinctive green-colored rectangle. The target is first detected using the hue channel of frontal images captured by a color camera. Noise in the resulting binary image is filtered through morphological operators and the largest remaining blob is selected. The contour is detected and the rectangle lines are estimated by a robust fitting procedure Fischler & Bolles (1981). Finally, the corners coordinates are determined from the lines intersection, as shown in Figure 10. A visual servoing strategy was used so that the robot could follow the green rectangle at a predefined distance (1m) and oriented toward its center Vassallo et al. (2002b).

While the robot follows a person, it uses its visuo-motor map (egomotion) to perceive its own movements. This motor information is classified into clusters by an unsupervised learning method based on K-means. In our implementation, the number of desired clusters is defined by the user but it could be learned in an automatic manner. The different clusters (and associated centroids) represent the learned characteristic movements and will constitute the motor vocabulary. If needed, labels can also be associated to each movement as they were motor words.

During the tests, a set of person-following experiments were done to learn this motor vocabulary. Egomotion values were normalized and classified into clusters defining the motor words. As we discussed before, the motor variables for a robot with differential kinematics, are the angular velocity ω_z and the robot forward velocity, T_y . We estimate both the values for T_x and T_y . The values for T_x are usually non-zero (albeit small) due to sliding of the robot wheels or noise in the estimation process. Once these values are small, they will not be considered for defining the motor vocabulary, but just T_y and ω_z .

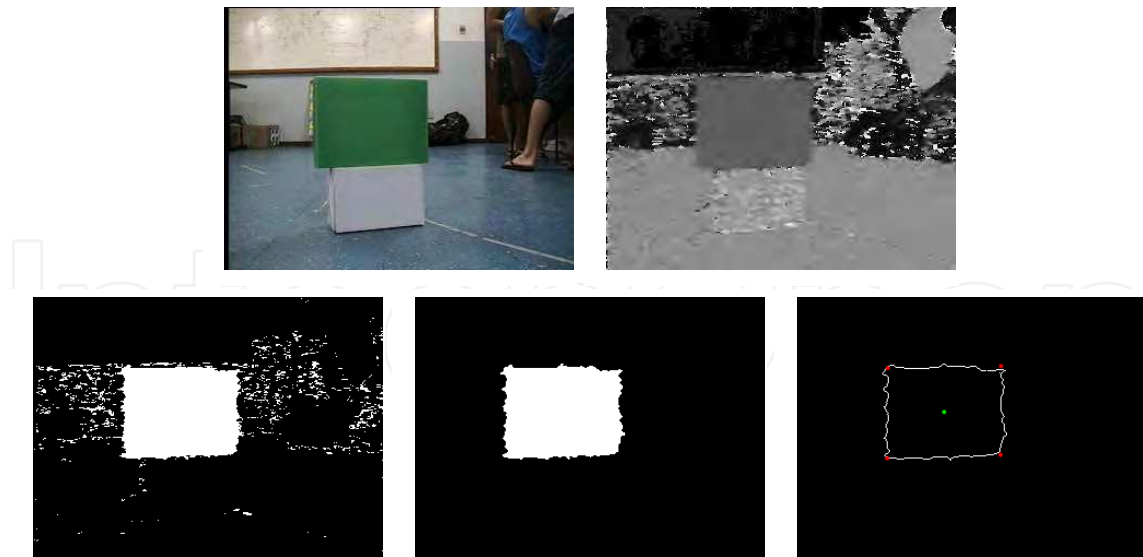


Fig. 10. The green rectangle's corners detection: the rgb-image, the hue channel, the image segmentation and processing until the corners estimation.

The clusters found in the (T_y, ω_z) space are shown in Figure 11. The mean values are represented by asterisks and Voronoi lines separating the various clusters are drawn. The centroids represent the purposive motor vocabulary. Names were given to each cluster defining motor words which are detailed in Table 2.

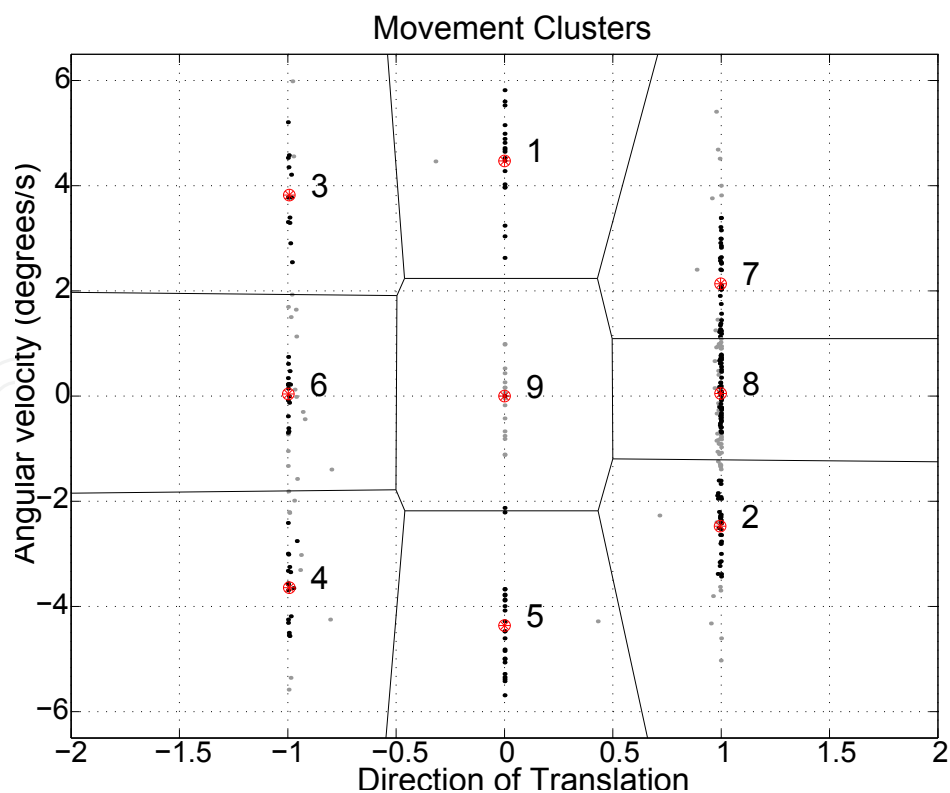


Fig. 11. Clusters of the created vocabulary. Black/gray points indicate inliers/outliers.

N	Motor Word	T_y	ω_z (degrees/s)
W 1	left turn	0	4.4722
W 2	front right	0.9948	-2.4749
W 3	back left	-0.9927	3.8223
W 4	back right	-0.9927	-3.6440
W 5	right turn	0	-4.3646
W 6	back	-0.9972	0.0386
W 7	front left	0.9973	2.1350
W 8	front	0.9974	0.0488
W 9	stopped	0	0

Table 2. Motor words for each cluster.

It is important to stress that this level of development allows the system to organize movements onto more complex motor actions. Motor actions can be seen as the result of several motor programs and treated in a more symbolic level when compared to individual movements.

The motor vocabulary is determined by two main factors: the environment and the user/demonstrator. If the environment were exclusively constituted of curved sections, rectilinear movements would not be part of the vocabulary. Similarly, if the demonstrator always uses right-turns, then left-turns would not be represented in the vocabulary. This is advantageous because the vocabulary adapts to the structure of the environment as well as to the “behavior” of the demonstrator.

Now that the motor action vocabulary is created, the robot can define and recognize movements of interest and use those to perform a desired task. Movement recognition is done in motor space, using the Euclidean distance as the discriminant function, instead in visual space, which depends on frame orientation and position, and environment illumination.

4. Solving a task: Topological Mapping and Navigation

This section shows how we used the visuo-motor map and the learned motor vocabulary for a mobile robot application: topological mapping and navigation. This task appears at the end of a sequence of (developmental) steps where the robot acquires progressively more sophisticated skills:

1. Definition of a visuo-motor map for the mobile robot, in the form of an egomotion estimation process.
2. A person-following (imitation) behavior was used to allow the robot to learn a purposive motor action vocabulary. These actions depend both on the environment structure and the demonstrator pattern of actions. These actions are “discrete” and are built upon the elementary motor signals acquired through the visuo-motor mapping process.
3. Finally, the learned vocabulary is used for building topological maps and for navigation.

During the map building phase, we assume that the robot is guided through the environment by a user or a demonstrator. While moving, the robot creates a topological map, in the form of a graph. Nodes of this graph correspond to omnidirectional images captured during the

robot motion, while the links are associated to motor actions, resulting from the egomotion estimates projected onto the motor vocabulary.

The decision of whether or not inserting a new node in the map is taken based on two criteria: (i) a comparison with the previously stored reference image or (ii) abrupt changes in motion. A new node is added to the map whenever the *sum of squared differences* (SSD) between successive images exceeds a threshold or the robot motion changes suddenly. Once a new node is stored, the most frequently recognized motor word (to introduce robotsness) is attributed to the link between the current and previous nodes.

Navigation with the created map comprises several steps. First, the robot determines its initial position. Then, the robot searches a path in the graph leading from the initial position to the goal destination. During navigation, progress along the route is monitored by comparing the captured images against the current image node and the subsequent one in the path. Image correlation is based on the same SSD metric that is used for map building. To determine whenever the robot position should be updated, we adopt a hill-climbing strategy in such a way that a new node is accepted whenever a persistent increase in the SSD values is detected after a clear-cut minimum ("valley").

Every time a new map position (node) is reached, the motor word stored in the subsequent link determines the next motor command. This behavior goes on until the final location is reached. Navigating through the map can be seen as executing a concatenation of the motor actions sampled from the motor vocabulary. Such commands are directly retrieved from the map, as that information was stored during the mapping phase. In other words, it can be considered a "motor program" formed by "motor words". Some experiments and results are shown in the next section, a natural way of expressing and representing navigation tasks.

5. Experiments and Results

The robot used in the experiments described in this paper is a Pioneer DX2 equipped with an onboard computer (Pentium II MMX - 266 MHz - 128 MRAM). An omnidirectional camera is mounted on the top of the robot with its axis aligned with the platform's rotation axis. It is a catadioptric system formed by a B&W camera and a spherical mirror Baker & Nayar (1998). In addition, the robot is equipped with a color (perspective) camera pointing to the forward direction. The robot and both vision systems were already shown in Figure 4.

The omnidirectional system is used for defining the visuo-motor mapping, based on the egomotion estimation process, and for performing the topological mapping task. The color camera is used for implementing the imitation/following behavior.

As we have seen before the mobile robot imitation (person following) behavior was used in two distinct ways. First, it is used to build a visuo-motor mapping and for the robot to learn elementary motor actions, coded according to the robot's motion repertoire. Secondly, the following behavior is used to combine such basic actions for topological mapping and navigation. These maps become motor representations of the environment, which main advantage is to be adapted to the used robot and environmental structure.

Several mapping and navigation experiments were conducted for testing the proposed method. For that purpose, we have used an experimental arena adapted to the robot size and camera height. One of the mapping tests is shown in Figure 12. The executed path for map building is plotted and the places where images were selected as topological map nodes are indicated by asterisks

In this experiment, a total of 16 omnidirectional images were selected for map nodes. These images were subsampled twice down to 166 x 166 pixels. As described, motor words extracted

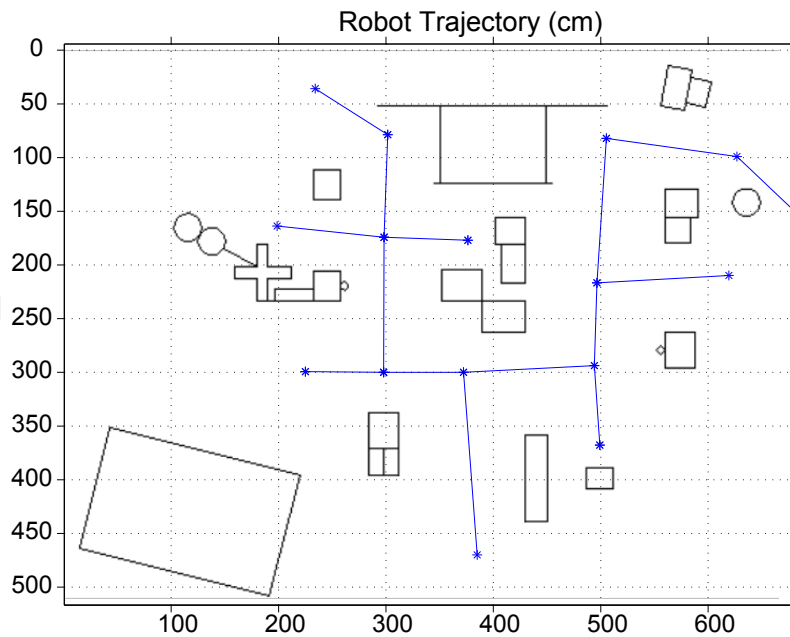


Fig. 12. Robot trajectory (odometry) during map building. Asterisks indicate map nodes.

from the learned vocabulary were associated to the links between images, according to robot motion. Whenever possible, the dual (symmetric) motion of each motor word was also associated to the map link. This strategy allows the robot to navigate in each link in opposite directions. Figure 13 shows the omnidirectional images overlaid the map nodes.

Some navigation experiments for the created map are shown in Figure 14. The robot was asked to navigate between different points across the map. It involved traversing some of the map links in the direction opposite to that learned during mapping.

The places where the robot updated its (qualitative) position are indicated by asterisks. Although some points correspond to the same node in the map, the asterisks did not happen exactly in the same coordinates but in the same region, associated to a common “qualitative” location. The robot solved the navigation tasks successfully in about 90% of the trials. It only fails in cases where it encountered obstacles during the mission, since we did not include any obstacle avoidance scheme. Instead, our efforts were directed to testing the task learning method and the use of motor information for topological map building and navigation. We therefore found the results very encouraging.

6. Conclusions and Future Work

In this work, we have proposed an approach that allows a robot to use motor representations for learning a complex task through imitation. Our approach is inspired after developmental psychology and some findings in neuroscience. The framework relies on development as the process allowing the robot to acquire sophisticated capabilities over time, as a sequence of simpler (learning) steps. At the first level, the robot learns about sensory-motor coordination. Then, motor actions are identified based on lower level, raw signals. Finally, these motor actions are stored in a topological map and retrieved in an efficient way during navigation. The entire methodology is grounded on hierarchical motor representations as suggested by experiments in Mirror Neurons.

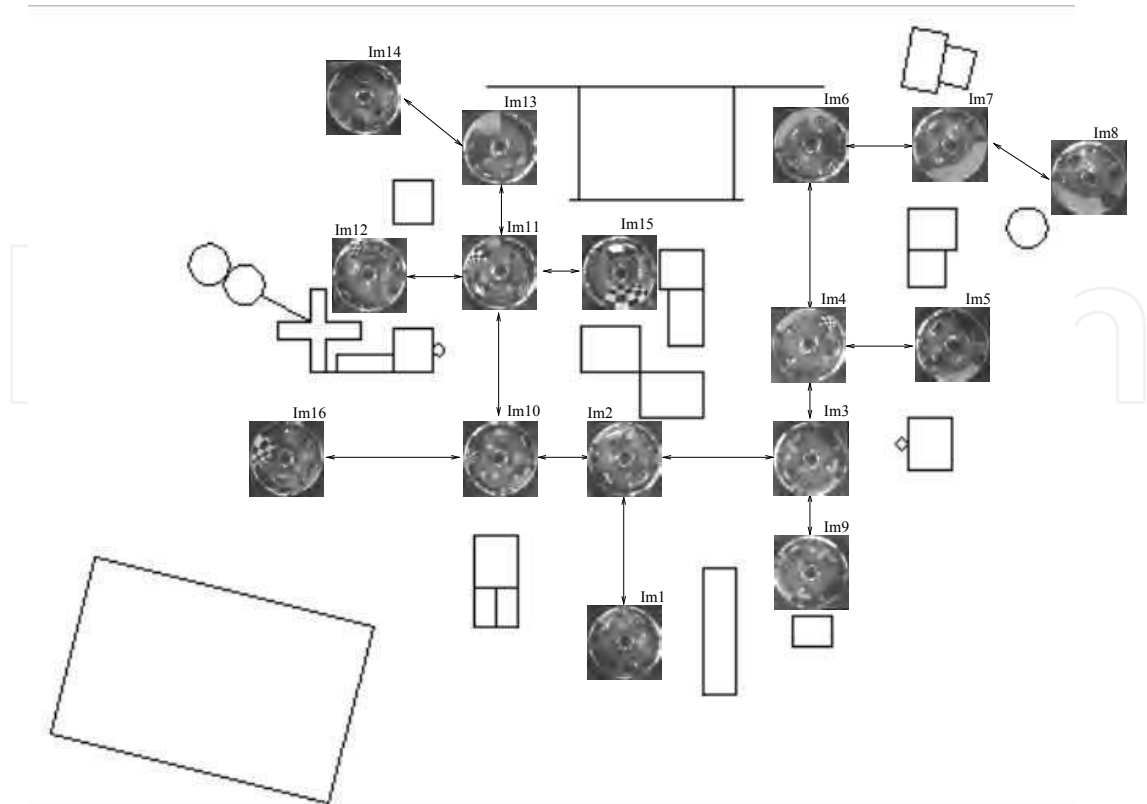


Fig. 13. The 16 omnidirectional images of each map node.

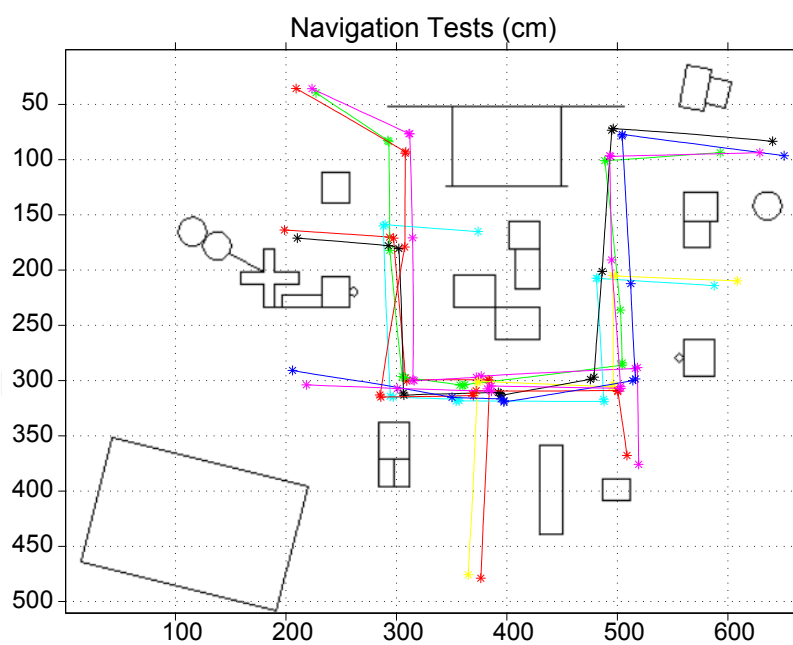


Fig. 14. Robot trajectories (odometry) when map navigating. Asterisks are position updates.

We illustrated the approach with an application with a mobile robot. An egomotion estimation process was used to perform visuo-motor mapping. It transforms optical flow measurements

on omnidirectional images to motor information, matching the motor capabilities (kinematics) of the differential robot chosen for the task. Extensive tests were performed to test and analyze the egomotion implementation validating the visuo-motor map.

In a subsequent stage of development, a person following behavior was implemented. The robot was able to learn a set of elementary actions, based on the visuo-motor mapping, that form a motor vocabulary.

Finally, the robot builds a topological map of the workspace, while it is guided through the environment by a human operator. During this process, the robot selects omnidirectional images to store as map nodes and uses the motor vocabulary to associate basic actions to links. After that, navigating through the map was considered a higher level action, composed by basic motor movements.

It is important to stress that the learning process occurs throughout the entire action hierarchy: during the self-knowledge phase when the visuo-motor map was defined; learning through imitation when basic actions are encoded in motor space and, finally, executing a desired task by composing more complex actions based on the elementary ones. This approach thus avoids the need to program every action or situation that robot must execute or encounter during its lifetime.

We consider that the main contributions of this work are multi-fold: (i) to make a robot able to rely on interaction with a user to learn sets of movements that were not programmed or defined a priori; (ii) to provide a natural manner of adapting the desired application and the way it is executed to the robot's motion repertoire; (iii) a methodology for open-ended robot learning and adaptation that can be applied to different robots, applications and environments. The sensory-motor maps and the created motor vocabulary are defined according to robot's body and sensory system. Whenever the application or environment changes, one can either teach the robot a new set of movements or modify the current motor vocabulary.

More specifically, we show that the use of omnidirectional images is beneficial both for egomotion estimation as well as for the mapping and navigation processes. Large fields of view help finding and tracking objects and allows more of the environment to be caught in just one image. Also, using motor representations to define the links in a topological map allows representing the environment in a motor way, adapted to the robot used for the task.

The results obtained are encouraging. Future work will focus on further extending these ideas and conducting tests with other robots and applications. We are also working on a way of learning the visuo-motor mapping instead of computing it explicitly through egomotion equations. We have already got some results using neural networks based on perceptrons Vassallo et al. (2004).

Acknowledgements

This work is partly supported by CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior via a cooperation between Instituto Superior Técnico, Lisbon - Portugal, and Universidade Federal do Espírito Santo, Vitória - Brazil; and also by CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil, through the project 472727/2003-6.

7. References

Baker, S. & Nayar, S. K. (1998). A theory of catadioptric image formation, *Proc. of 6th International Conference on Computer Vision - ICCV98*.

- Billard, A. (2002). *Imitation in Animals and Artifacts*, The MIT Press, chapter 11 - Imitation: A Means to Enhance Learning of a Synthetic Protolanguage in Autonomous Robots, pp. 281–310.
- Billard, A. & Hayes, G. (1997). Learning to communicate through imitation in autonomous robots, In Proc. of ICANN97 - 7th International Conference on Artificial Neural Networks.
- Bruss, A. & Horn, B. K. P. (1983). Passive navigation, *Computer Vision, Graphics and Image Processing* **21**: 276–283.
- Dautenhahn, K. (1995). Getting to know each other - artificial social intelligence for autonomous robots., *Robotics and Autonomous Systems* **16**: 333–356.
- Dautenhahn, K. & Nehaniv, C. L. (eds) (2002). *Imitation in Animals and Artifacts*, The MIT Press.
- Fadiga, L., Fogassi, L., Gallase, V. & Rizzolatti, G. (2000). Visuo-motor neurons: ambiguity of the discharge or 'motor' perception?, *Int. J. Psychophysiology* **35**: 165–177.
- Fischler, M. A. & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography., *Communications of ACM* **24**(6): 381–395.
- Fitzpatrick, P., Metta, G., Natale, L., Rao, S. & Sandini, G. (2003). Learning about objects through action - initial steps towards artificial cognition, *IEEE International Conference on Robotics and Automation - ICRA 2003*.
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex, *Brain* **119**: 593–609.
- Gaspar, J., Grossmann, E. & Santos-Victor, J. (2001). Interactive reconstruction from an omnidirectional image, In Proc. of the 9th International Symposium on Intelligent Robotic Systems - SIRS2001.
- Geyer, C. & Daniilidis, K. (2000a). Equivalence of catadioptric projections and mappings of the sphere, Proc. of IEEE Workshops on Omnidirectional Vision.
- Geyer, C. & Daniilidis, K. (2000b). A unifying theory for central panoramic systems and practical implications, *ECCV 2000*, Vol. 2, pp. 445–461.
- Gluckman, J. & Nayar, S. K. (1998). Ego-motion and omnidirectional cameras, *Proc. of 6th International Conference on Computer Vision - ICCV98*, pp. 999–1005.
- Lopes, M. C. & Santos-Victor, J. (2003a). Motor representations for hand gesture recognition and imitation, *IROS Workshop on Robot Programming by Demonstration*.
- Lopes, M. C. & Santos-Victor, J. (2003b). Visual transformations in gesture imitation: what you see is what you do, *IEEE Intl. Conference on Robotics and Automation - ICRA2003*.
- Lucas, B. & Kanade, T. (1981). An iterative image resgistration technique with an application to stereo vision., *IJCAI'81*.
- Lungarella, M. & Pfeifer, R. (2001). Robot as cognitive tools: Information theoretic analysis of sensory(-motor) data, *EDEC 2001 - Symposium at the International Conference on Cognitive Science*.
- Matáric, M. J. (2002). *Imitation in Animals and Artifacts*, The MIT Press, chapter Sensory-Motor Primitives as a Basis for Imitation: Linking Perception to Action and Biology to Robotics, pp. 391–422.
- Metta, G., Sandini, G. & Natale, L. (2001). Sensorimotor interaction in a developing robot, *1st Intl. Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, pp. 18–19.
- Metta, G., Sandini, G., Natale, L., Manzotti, R. & Panerai, F. (2001). Development in artificial systems, *EDEC 2001 - Symposium at the International Conference on Cognitive Science*.

- Nelson, R. & Aloimonos, J. (1988). Finding motion parameters from spherical motion fields (or the advantages of having eyes in the back of your head), *Biological Cybernetics* **58**: 261–273.
- Rizzolatti, G. & Arbib, M. A. (1998). Language within our grasp, *Trends in Neurosciences* **21**(5): 188–194.
- Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. (2002). *The Imitative Mind: development, evolution, and brain bases*, Cambridge University Press, chapter 14 - From mirror neurons to imitation: facts and speculations, pp. 247–266.
- Vassallo, R. F., Santos-Victor, J. & Schneebeli, H. (2002a). A general approach for egomotion estimation with omnidirectional images, *IEEE Workshop on Omnidirectional Vision 2002 (with ECCV)*, pp. 97–103.
- Vassallo, R. F., Santos-Victor, J. & Schneebeli, H. J. (2004). Task learning and adaptation based on visuo-motor mapping, *CBA 2004 - XV Congresso da Sociedade Brasileira de Automática*.
- Vassallo, R., Santos-Victor, J. & Schneebeli, H. (2002b). Using motor representations for topological mapping and navigation, *Intl. Conference on Intelligent Robots and Systems - IROS2002*, pp. 478–483.

IntechOpen

IntechOpen

IntechOpen



Mobile Robots Navigation

Edited by Alejandra Barrera

ISBN 978-953-307-076-6

Hard cover, 666 pages

Publisher InTech

Published online 01, March, 2010

Published in print edition March, 2010

Mobile robots navigation includes different interrelated activities: (i) perception, as obtaining and interpreting sensory information; (ii) exploration, as the strategy that guides the robot to select the next direction to go; (iii) mapping, involving the construction of a spatial representation by using the sensory information perceived; (iv) localization, as the strategy to estimate the robot position within the spatial map; (v) path planning, as the strategy to find a path towards a goal location being optimal or not; and (vi) path execution, where motor actions are determined and adapted to environmental changes. The book addresses those activities by integrating results from the research work of several authors all over the world. Research cases are documented in 32 chapters organized within 7 categories next described.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Raquel Frizera Vassallo, Hans Jorg Andreas Schneebeli and Jose Santos-Victor (2010). Topological Mapping and Navigation Using a Developmental Learning Approach Based on Imitation through Sensory-Motor Maps, Mobile Robots Navigation, Alejandra Barrera (Ed.), ISBN: 978-953-307-076-6, InTech, Available from: <http://www.intechopen.com/books/mobile-robots-navigation/topological-mapping-and-navigation-using-a-developmental-learning-approach-based-on-imitation-throug>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen