# Dealing with robustness in mobile robot guidance while operating with visual strategies

Giovanni Bianco

Servizio Informatico
University of Verona
Via San Francesco 8, I-37138 Verona
Italy
Bianco@chiostro.univr.it

Alexander Zelinsky

Dept. of Systems Engineering, RSISE
The Australian National University
Canberra, ACT 0200
Australia
Alex.Zelinsky@anu.edu.au

## Abstract

*This paper introduces a theory to formally and practically analyze the robustness issues of visual guidance methods for robot navigation. The first aspect is related to the convergence of the navigation system to the goal. It will be shown how the dynamic system which drives the strategies can be analyzed by using classical concepts such as the Liapunov functions. The second aspect concerns the conservativeness of the resulting navigation vector fields. It will be shown how this deals with the repeatability of the trials. Furthermore, the selection of the best landmarks to perform the navigation processes strongly affects the conservativeness thus providing a formal way to do landmark learning. The theory has been tested with two different visual methods that have been derived from the biological world: the snapshot model and the landmark model. The former considers a portion of the full panorama taken by a color camera to accomplish navigating actions. The latter is a more sophisticated approach which uses the most suitable visual landmarks to calculate navigation movements.*

## 1 Introduction

A key problem in robotics is the development of effective visual navigation strategies and a crucial aspect is concerned with the search for robust yet practical solutions [6]. Interestingly, animals, including insects, are robustly able to achieve visual navigation. Entomological studies have discovered several mechanisms of visual navigation that offer valuable ideas for robot navigation (see review [21]).

Therefore, the selection of relevant visual features, their use in visual navigation models and the ability to work with practical solutions have been widely addressed in literature often starting from a biological background (e.g. [15, 22, 14, 20, 3, 16, 9, 8, 5]).

All these aspects are of utmost importance but a thorough study on the reciprocal effects for the final visual behavior still has to be performed. In [2] we show how the feature selection phase strongly affects the final navigation process. In [4] and [5] we have shown how different visual guided navigation strategies operate with the same driving principle.

In this paper we present a theory that studies two of the main aspects related to the concept of *robustness*: i.e. *repeatability* and *convergence* [1]. Two main results are presented:

- the *visual potential function* that represents the driving principle to perform visual guidance is proven to be a *Liapunov* function and therefore can state about the convergence of the navigation system to the goal.
- The *conservativeness* of the navigation vector field deals with the concept of repeatability of the trials and provides key information to perform landmark learning.

These results are not limited to only visual guidance, but apply to broader classes of algorithms whose driving principle is similar to the system dynamic model reported in Equation 1.

The main contribution of this paper is the theory itself which can provide a useful base to develop new visual guided strategies or tune existing visual mechanisms measuring their effects involving robustness.

The organization of this paper is as follows. In Section 2 the mathematical aspects related to visual guidance strategies are detailed. In particular, in Section

2.1 issues concerning the visual potential function and how this can be interpreted as a Liapunov function are presented. In Section 2.2 formal studies related to the concept of conservativeness of the navigation field are introduced. In Section 3 the results of two different tests involving both convergence and repeatability are detailed. Concerning the experimental aspects, both methods use a color camera with approximately the same field of view and resolution. The different guidance approaches have been tested in different environments of comparable sizes while the agents are different (details can be found in Section 3) as described later.

# 2 Theory

Starting with local sensor information, a vector needs to be computed by the agent which will use it to perform the next movement. If vector $\vec{V}$ represents the next movement with a module and a direction relative to the actual robot position, considering an agent with two D.O.F., for sake of simplicity, the system dynamical model to perform guidance is therefore given by:

$$\begin{cases} x(k+1) = x(k) + V_x(x(k), y(k)) \\ y(k+1) = y(k) + V_y(x(k), y(k)) \end{cases} \quad (1)$$

where $x(k)$ and $y(k)$ represent the coordinates of robot at step $k$; $x(k+1)$ and $y(k+1)$ represent the new positions the robot will move to. Clearly, an important equilibrium point $(x^*, y^*)$ for the system is given by the coordinates of the goal position.

Lastly, $V_x(x(k), y(k))$ and $V_y(x(k), y(k))$ represent the displacements computed at step $k$ following a generic guidance model. Those displacements are related to the position at step $k$ given by $(x(k), y(k))$.

Some examples of computing vector $\vec{V}$ starting from visual inputs can be found in e.g. [3, 16] for the so called *snapshot model* and in e.g. [5] for the *landmark model*. The working hypotheses for the methods is that visual information about the goal has been previously grabbed. Details of both models are presented in Section 3.

Both approaches, however, share a similar navigating principle. The computation over the whole environment of vector $\vec{V}$ defines a vector field $\mathsf{V}$. Several interesting considerations can be suitably extracted by analyzing its properties. Those studies can be further generalized to address theoretical aspects related to the concept of robustness.

Let us consider a partial set of equivalent statements about a generic vector field $\mathsf{V}$ [13].

- any oriented simple closed curve $c$: $\oint_c \mathsf{V} \cdot \mathsf{ds} = 0$
- $\mathsf{V}$ is the gradient of some function $U$: $\mathsf{V} = \nabla \mathsf{U}$

Both these conditions are of utmost importance for guidance strategies. The former is related to the concept of *conservativeness* of the field. The latter is concerned with the existence of a *potential function* that uniquely generates the field.

From another point of view, the former is concerned with the repeatability of the experiments, the latter is concerned with their convergence to the goal. The following Sections address the two aspects in turn.

## 2.1 Convergence

Supposing that all the necessary hypotheses hold, the dynamic system presented in Equation 1 can be considered continuous-time with the following (leaving out the vector notations):

$$\dot{x}(t) = V(x(t)) \quad (2)$$

where $x$ represents the generic coordinates and an equilibrium point $x^*$ is located at the goal position.

The basic idea for verifying the stability and the convergence of a dynamic system is to seek an aggregate summarizing function on the states of the system itself [12].

In particular, when a dynamic system can be represented by $\dot{x} = f(x)$ with a fixed point $x^*$, and it is possible to find a *Liapunov function*, i.e. a continously differentiable, real-valued function $U(x)$ with the following properties [19]:

1. $U(x) > 0$ for all $x \neq x^*$ and $U(x^*) = 0$
2. $\dot{U}(x) < 0$ for all $x \neq x^*$ and $\dot{U}(x^*) = 0$ (all trajectories flow *downhill* toward $x^*$)

then $x^*$ is globally stable: for all initial conditions $x(t) \to x^*$ as $t \to \infty$.

The system depicted in Equation 2 is of type $\dot{x} = f(x)$ but, unfortunately, there is no systematic ways to construct Liapunov functions.

In our case, by considering the vector field the guidance system produces an important Liapunov function that can be constructed by integrating the right-hand side of the system Equation 2 as reported in [12].

Our approach consists of calculating the shape of $U$ *a posterior*, starting from the vector field $\mathsf{V}$. Classically, $U$ is mathematically specified once given the

sensors and the environment (see e.g. [10] and review in [11]).

The scalar function $U$ can thus be given by integrating the conservative field $\mathsf{V}$ in question [17] taking into account that the field must be inverted in sign [11]:

$$U(x,y) = -\int_{(0,0)}^{(x,y)} \mathsf{V} \cdot \mathsf{ds} \tag{3}$$

where the path of integration, following the infinitesimal piece of motion $d\mathsf{s}$, is arbitrary. The scalar function $U$ is referred to as the *potential* of the conservative field $\mathsf{V}$ in question.

The scalar product reported by Equation 3 can be further simplified by following a *particular* curve $c$. Therefore, more effectively, it can be written as:

$$U(x,y) = -\int_{p_x}^{x} V_x(X,p_y)\cdot dX - \int_{p_y}^{y} V_y(x,Y)\cdot dY \tag{4}$$

where $U(x,y)$ is the potential function and the path of integration is along the semi-perimeter of the rectangle connecting $(p_x,p_y)$ to $(x,y)$, i.e. the horizontal line segment from the initial point $(p_x,p_y)$ to the vertical line through $(x,y)$ and then along the vertical line segment to $(x,y)$.

An advantage of this method is the use of the goal position as the referring point $(p_x,p_y)$. Every other point is thus referred to in terms of potential in reference to the goal position.

If the visual potential function has a basin of attraction where the minimum is at the goal position then for the considerations expressed above homing is intrinsically stable, when starting navigating from part of the environment.

## 2.2 Repeatability

As described above, a vector field $\mathsf{V}$ is said to be *conservative* when the integral computed on any closed path (*circuitation*) is zero.

If the field is not conservative (when the Equation stating that the circuitation is null does not hold anymore for at least one curve) then the integration detailed in Equation 3 can lead to an infinite number of results depending on the integrating curve $c$.

This means that, the potential function is no longer entirely determined by the extreme points of the integration process. This potential function can be considered as *multi-valued*: for every reference position, a general position will have more than one potential value according to the path chosen for integration. This sort of multi-valued potential function can be translated in *non-uniqueness* of the vector field. In other words, repeatedly placing the robot in the same point within a non-conservative area, different paths can be followed.

Theoretically, as the field is not unique, the repeatability of navigation paths does not hold and this can lead to unpredictable results.

A practical measure of the conservativeness of the field appears to be essential to assess the quality of the navigation process. To this extent, if the vector field is defined on a connected set then the Equation stating that the circuitation must be null is equivalent to [18]:

$$\frac{\partial V_x(x,y)}{\partial y} = \frac{\partial V_y(x,y)}{\partial x} \tag{5}$$

In other terms:

$$\frac{\partial V_x(x,y)}{\partial y} - \frac{\partial V_y(x,y)}{\partial x} = 0 \tag{6}$$

Hereafter, the result of Equation 5 will be referred to with the term of *conservativeness*.

# 3 Experiments with visual guidance

In this Section we present two models of visual guidance to test the theory above presented. In choosing the experiments we have the following objective: showing how to practically assess the robustness of a particular guidance model.

The approaches we chose to deal with are the snapshot model and the landmark model. In addition, two different agents are used to accomplish the experiments: a tripod with a camera for the snapshot model and a *Nomad200* for the landmark model. Details on the experiments can be found in [3, 16, 5, 2].

For both models, in order to measure the vector field, the agent was manually placed at various points in the environment. The position of the agent progressively covered a grid in the environment where each cell is approximately as big as the base of the agent itself.

From those points, applying one of the navigation methods, a displacement vector is computed. The iteration of the method over the whole environment and the collection of every displacement vector produces a vector field, as shown in Figure 1.

After collecting the whole set of vectors, it is necessary to compute the cross derivatives and evaluate the conservativeness condition and integrate the values to calculate the potential function, given by Equations 6 and 4 respectively.

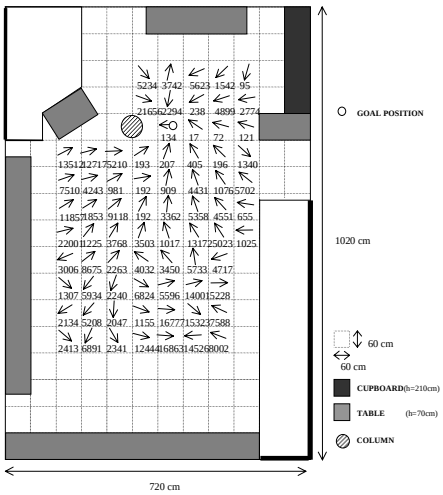


Figure 1: An example of a navigation field: directions and modules (numbers)

## 3.1 Tests with the snapshot model

For the snapshot approach the main idea is that an estimate of the vector pointing from the current position of the agent to the pre-learnt goal can be computed comparing position and amplitude of matching areas in the considered images. The matching between the goal image and the actual view is performed using an affine model. All possible affine transformations and shifts of the actual image in the allowed range are computed and the one that best fits the goal image is chosen. From the parameters of the affine transformation the algorithm computes an estimate of the robot displacement from the goal position, i.e. its current position.

Examples of images used by the algorithm are shown in Figures 2 and 3. The images were acquired at the goal position and in a generic starting point of the environment. A decimation process was applied in order to speed up the affine matching.

Figure 2: Vertically sub-sampled goal position image

Figure 3: An example of a vertically sub-sampled starting point image

Some constraints, namely fixed heading and constant height of the camera, allow for the use of a simplified affine model to perform matching:

$$\begin{cases} S_{\mathsf{X}}(X,Y) = a_{0\mathsf{X}} + a_{1\mathsf{X}} \cdot X + a_{2\mathsf{X}} \cdot X \\ S_{\mathsf{Y}}(X,Y) = a_{0\mathsf{Y}} + a_{1\mathsf{Y}} \cdot Y + a_{2\mathsf{Y}} \cdot Y \end{cases} \tag{7}$$

where $S_{\mathsf{X}}$ and $S_{\mathsf{Y}}$ are the displacement components for the matching along $x$ and $y$ axes respectively; $(X, Y)$ are the pixels co-ordinates, $a_{0\mathsf{X}}$, $a_{0\mathsf{Y}}$ represent translations in pixels and $a_{1\mathsf{X}}$, $a_{2\mathsf{X}}$, $a_{1\mathsf{Y}}$ $a_{2\mathsf{Y}}$ represent expansions (a-dimensional). Furthermore, other working hypotheses allow for additional simplifications of the affine model [16]. The computation of the estimated displacement components from the actual position of the agent to the goal is given by:

$$\vec{V} = [V_{\mathsf{x}}\ V_{\mathsf{y}}] = [K \cdot a_{0\mathsf{X}}\ \ H \cdot a_{1\mathsf{X}}] \tag{8}$$

where $K$ and $H$ are constants which derive from the central projection theorem as reported in [16].

All the tests have been performed in the same environment where the goal position is located in co-ordinate $(20, 30)$.

In Figure 4 the conservativeness of the navigation field is shown. The goal has been acquired while facing towards the maximum value for $y$. Figure 3 has been acquired approximately at co-ordinate $(20, 5)$ facing towards the goal. A large part of the field can be considered conservative. However, there are some regions of the environment where this condition does not hold. This affects the repeatability of the navigation paths starting from those areas: from the same starting point different paths can be selected, possibly missing the goal position and getting trapped somewhere else.

The visual potential function is shown in Figure 5. It is calculated by integrating the field, whose conservativeness is shown in Figure 4, as specified by Equation 3.

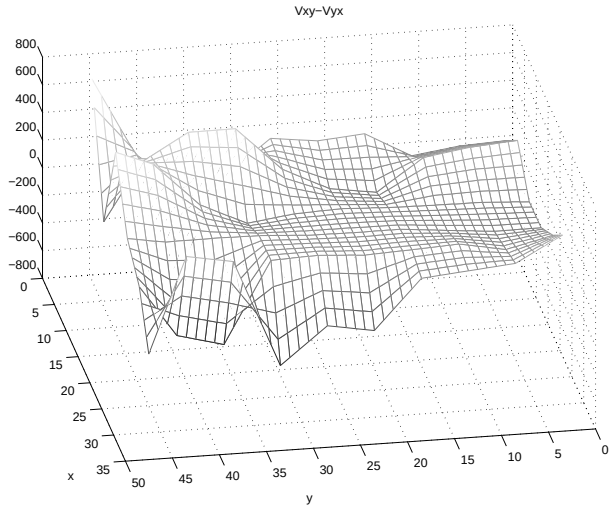


Figure 4: Conservativeness of the navigation field for the snapshot model

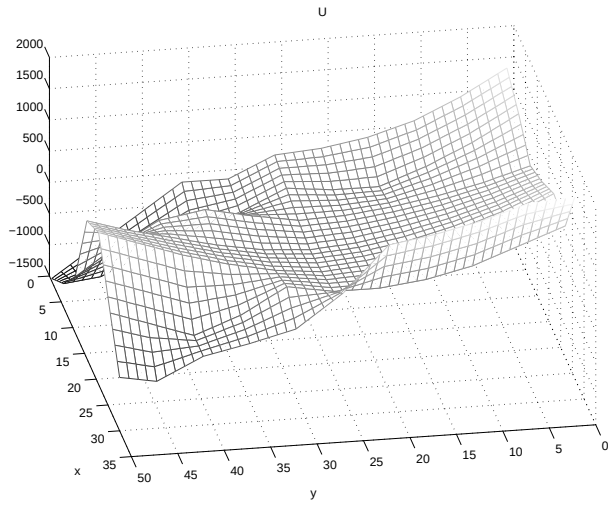


Figure 5: The visual potential function with the snapshot model

There are large areas of the environment from which the navigation phase can be successfully performed. Nevertheless, starting from some areas the robot can get trapped by *false* goals.

Stability is not guaranteed for the whole environment, as can be seen in Figure 5. The method cannot deal with the intrinsic limitations of the guidance strategy when the agent is placed in some regions of the environment (in Figure 5 when $x \leq 10$ and $y \geq 35$). The use of a camera with a limited field of view can lead to bad navigation results: the agent is attracted to *false* goals because the actual view might not contain enough information for a good comparison with the stored snapshot taken at the goal position. This can be overcome by acting either on the matching method or, more likely, by widening the camera field of view.

## 3.2 Tests with the landmark model

For the landmark model, after *reliable* landmarks have been chosen [2] then navigation information can be extracted from them. The underlying biological principle is that a real movement is represented by an *attraction* force. It is produced by taking into account the fact that the agent tries to restore the original position and size of every landmarks [7]. The data can be fused together by weighed addition.

Figure 6 summarizes the situation where the picture represents a typical frame taken during a navigation test. In particular, the circle at the bottom-center represents the overall attraction exerted by the goal. Above the circle the variance of that attraction is reported and under the circle the attraction vector is broken down into a magnitude and an angle. In the circle on the right the single attraction exerted by each landmark (box-shaped) is drawn. Each landmark has a number associated with it given by the value of the sigmoid function applied on its reliability measure. The arrows at the top-center of the Figure represent the motion commands given to the robot. In the rectangle on the left the visual potential field profile which has been followed so far is drawn.

The overall navigation vector can be thus calculated as (see Equation 1):

$$\vec{V} = [V_x \; V_y] = \frac{\sum_{l=1} \vec{v}_l \cdot s(r_l)}{\sum_{l=1} s(r_l)} \qquad (9)$$

where $L$ is the number of landmarks chosen after the

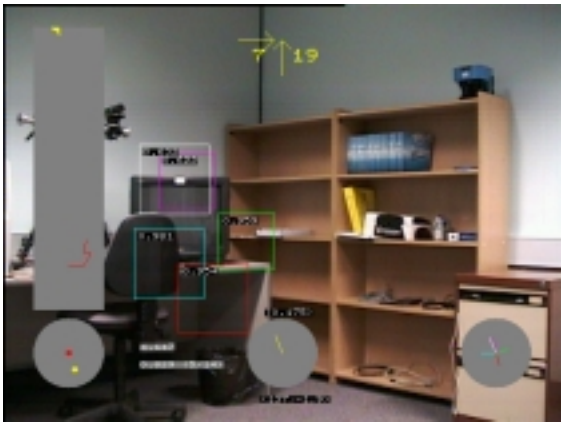

Figure 6: A frame taken by a real navigation

selection phase, $s(r_l)$ is a confidence value continuously associated to landmark $l$ and $\vec{v}_l$ is the attraction force *felt* by landmark $l$. Details of the computation can be found in [5, 2].

All the tests are performed in the same environment where the goal position is located at $(25, 30)$.

With this model, the measure of conservativeness is representative of the quality of the selected landmarks other than the repeatability of the trials as reported by Figures 7 and 8.

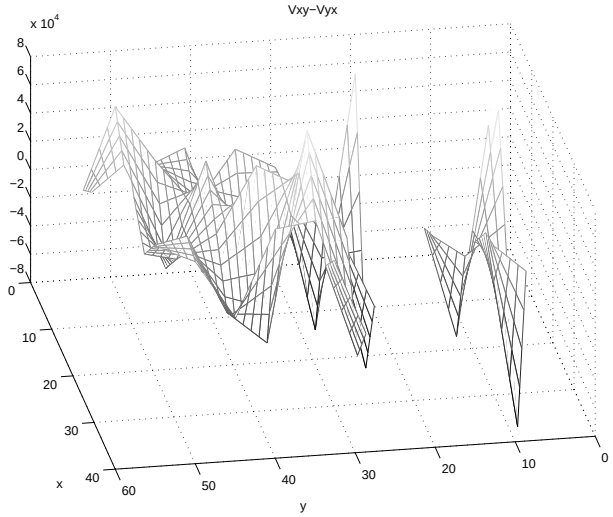


Figure 7: Conservativeness of the field before the dynamic learning selection phase

Figure 7 shows the conservativeness of the field when landmarks have been chosen from the goal image which still had to be further refined by a dynamic learning phase. Good landmarks, in fact, can be adequately selected following a learning phase where the agent follows a set of stereotyped movements [5, 2]. Figure 8 shows the conservativeness of the field after the learning phase.

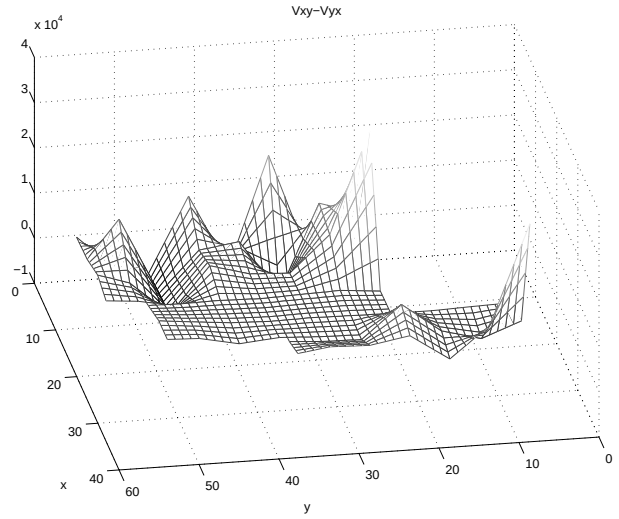


Figure 8: Conservativeness of a vector field after the learning phase

The computation of the visual potential functions is performed only on those areas of the environment where conservativeness holds. The tests previously performed reveal that the snapshot model allows us to immediately consider the vector field it produces. On the other hand, the landmark model does not behave that way. A learning phase is necessary for the field to be conservative at least for a large part of the environment. The visual potential function for the landmark model is shown in Figure 9.

The shape of the potential function tends to produce a minimum around the goal. In addition, the basin of attraction of the goal is the whole environment, i.e. apart from some isolated cases, all the starting points lead to the goal.

### 3.3 Issues on the visual potential as a Liapunov function

From the potential function previously plotted it can be easily understood why the system can sometimes be trapped by false goals or what are the regions of convergence for the main goal position.

This implicitly states that the system does not have overall stability in the whole environment. Therefore,

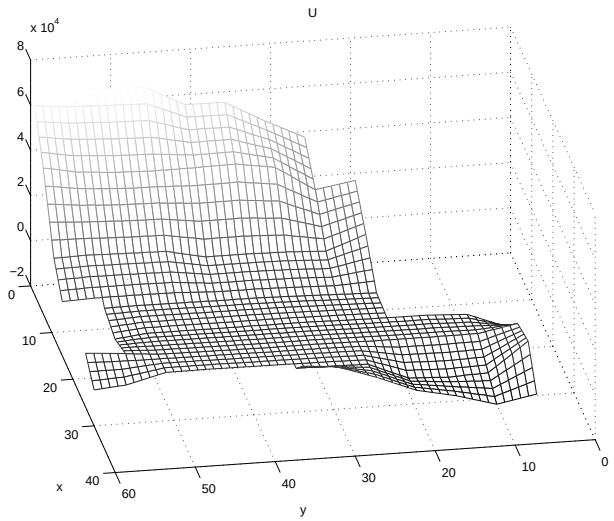


Figure 9: Potential function computed for the landmark model after a dynamic selection phase

the visual potential function itself cannot be considered Liapunov compliant unless reducing its domain of application to a region around the goal position, starting from which the system converges.

This restricted visual potential function can be formally adopted as a Liapunov function, though in a restricted area of the environment.

## 4 Conclusions and perspectives

The development of visual navigation strategies have been widely addressed in the robotics literature despite the lack of formal yet practical guidelines to cover various and different aspects.

In this paper it has been shown that a common driving principle, namely the visual potential function, can be regarded to as the *engine* of the visual guidance methods. Furthermore, the visual potential function itself can be regarded to as a summarizing function (specifically, a Liapunov function) to assess the stability of the different strategies.

Important considerations based on the conservativeness of the vector fields can also be stated. It has been shown how different aspects, e.g. landmark learning, can affect its characteristics. To this extent, the snapshot model is intrinsically more conservative (i.e. experiments are more repeatable) than the landmark model. But the performance of the latter can be strongly improved with a dynamic landmark learning phase.

## Aknowledgements

This research would not be possible without the suggestions and the efforts of Dr. Yoshio Matsumoto, Dr. Jochen Zeil, Dr. Gordon Cheng, Mr. Jochen Heinzmann, Dr. Miriam Lehrer and Prof. Mandyam Srinivasan. We are particularly indebted to Dr. Paolo Fiorini for his help in reviewing the content of this paper.

## References

[1] J. Andersen. Specifications. In R. Dorf, editor, *International Enclicopedy of Robotics*. Wiley, 1988.

[2] G. Bianco and R. Cassinis. Biologically-inspired visual landmark learning for mobile robots. In *Proceedings of the European Workshop on Learning Robots (EWLR 8)*, Lausanne (Switzerland), September 18 1999.

[3] G. Bianco, R. Cassinis, A. Rizzi, N. Adami, and P. Mosna. A bee-inspired robot visual homing method. In *Proceedings of the Second Euromicro Workshop on Advanced Mobile Robots (EUROBOT '97)*, pages 141–146, Brescia (Italy), October 22-24 1997.

[4] G. Bianco, A. Rizzi, R. Cassinis, and N. Adami. Guidance principle and robustness issues for a biologically-inspired visual homing. In *Proceedings of the Third Workshop on Advanced Mobile Robots (EUROBOT '99)*, Zurich (Switzerland), September 1999.

[5] G. Bianco and A. Zelinsky. Biologically-inspired visual landmark navigation for mobile robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Kyongju (Korea), October 17-21 1999.

[6] J. Borenstein, H. Everett, and L. Feng. *Where am I? Sensors and Methods for Mobile Robot Positioning*. The University of Michigan, April 1996.

[7] B. Cartwright and T. Collett. Landmark learning in bees. *Journal of Comparative Physiology*, A(151):521–543, 1983.

[8] P. Gaussier, C. Joulain, and J. Banquet. Motivated animat navigation: a visually guided approach. In *Simulation of Adaptive Behavior*, Zurich (Switzerland), August 1998.

[9] D. Jung, J. Heinzmann, and A. Zelinsky. Range and pose estimation for visual servoing of a mobile robot. In *Proceeding of the International Conference on Robotics and Automation*, Bruxels (Belgium), May 1998.

[10] O. Khatib. Real-time obstacle avoidance for manipulators and mobile robots. *Int. Journal of Robotics Research*, 4(1):90–98, 1986.

[11] J. Latombe. *Robot motion planning*. Kluwer Academic Publisher, Boston/Dodrecht/London, 1991.

[12] D. Luenberger. *Introduction to dynamic systems - theory, models, and applications*. John Wiley and Sons, New York Chichester Brisbane Toronto, 1979.

[13] J. Mardsen and A. Tromba. *Vector calculus*. W.H. Freeman and Company, 1996.

[14] Y. Matsumoto, M. Inaba, and H. Inoue. Visual navigation using view-sequenced route representation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 83–88, 1996.

[15] T. Mori, Y. Matsumoto, T. Shibata, M. Inaba, and H. Inoue. Trackable attention point generation based on classification of correlation value distribution. In *JSME Annual Conference on Robotics and Mechatronics (ROBOMEC 95)*, pages 1076–1079, Kavasaki (Japan), 1995.

[16] A. Rizzi, G. Bianco, and R. Cassinis. A bee-inspired robot navigation using color images. *Robotics and Autonomous Systems*, (25):159–164, 1998.

[17] C. Ross. *Differential Equations An Introduction with Mathematica*. Springer-Verlag, New York Berlin Heidelberg London Paris Tokyo Hong Kong Barcelona Budapest, 1995.

[18] G. Strang. *Introduction to Applied Mathematics*. Wensley-Cambridge Press, 1986.

[19] S. Strogatz. *Nonlinear dynamics and chaos*. Addison-Wesley Publishing Company, New York, 1994.

[20] S. Thrun. A bayesian approach to landmark discovery and active perception in mobile robot navigation. Technical report, School of Computer Science Carnegie Mellon University, 1996.

[21] O. Trullier, S. Wiener, A. Berthoz, and J. Meyer. Biologically based artificial navigation systems: Review and prospects. *Progress in Neurobiology*, 51:483–544, 1997.

[22] A. Zelinsky. Incorporating real-time vision into robotic task. In *Proceedings of the 1995 Joint Australia-Korea Workshop on Manufacturing and Technology*, pages 61–66, Wollongong (New South Wales, Australia), November 20-21 1995.