Evolving speed control in mobile robots: from blindness to kinetic vision

J.G. Donnett* and B.O. McGonigle**

*Dept. of Artificial Intelligence
**Laboratory for Cognitive Neuroscience
University of Edinburgh
Scotland

Abstract

Our approach to vision has derived from our attempts to engineer a robot capable of navigating autonomously in a local area environment. Begun essential as a 'blind' system, we first document its competences in obstacle avoidance and sound based navigation. As a derivation of the engineering, we have also succeeded in developing a form of sequential map 'representation' of obstacle distributions in the robot's environment, such that it can contour its operating speed to the local layout features of its terrain. As a further development of this competence, we also document our approach to a visually guided control system. Based on kinetically determined depth information and taking advantage of some of the established competences of the 'blind' robot system, the current system may serve to illustrate some of the productive interplay between visual and non-visual systems in the control of action.

1 Introduction

We are developing a vehicle intended for operation in 'local-area' environments such as offices/corridors, able to function as a robotic substrate in any of a variety of service roles (see [McGonigle 1990a, 1990b, 1990c] for the motivation and operating hypothesis of this enterprise). It is required to locate objects within its operating niche, although facilities for manipulation are not currently provided.

On the supposition that the robot would begin its 'evolution' as a blind agent, we began to explore the competences we could achieve without vision. The design issue of specifying which competences, and in which order they should be engineered to make system or adaptive sense has been addressed in [McGonigle 1989, 1990a, 1990b; Donnett and McGonigle 1991]. In this report, we confine ourselves principally to questions of fitting a robot's motion to an environment changing unpredictably with respect to the obstacles within it. We shall argue that an 'evolutionary' approach through implementation has led us naturally to an efficient system exploiting kinetic visual effects induced by the robot's own motion.

2 Agent Specification

From a biological perspective it is clear that many free-moving systems have no vision. In fact, detailed vision of any sort is an extrememly costly process to 'engineer' in biological terms, possessed only by advanced mammals. Engineering the process in robots appears to be similarly expensive in hardware/software complexity; this is demonstrated by the increasing number of roboticists who are attempting to sidestep camera-based mobile robot guidance and obstacle avoidance in favour of sonar (e.g., [Walter 1987, Crowley 1989]), infrared [Durieu *et al.* 1989], or laser-ranging techniques [Ozono *et al.* 1986, Hebert 1989].

Our vehicle (Fig. 1) has a moving head, currently fitted with two 'ears' (condenser microphones) selectively tuneable to a set of tones produced by acoustic beacons in the robot's environment. The presence of obstacles in the vehicle's path is registered by a collection of infra-red light emitter/detector pairs mounted on head and body. Direction and location of the robot relative to the beacons are determined by measuring beacon signal phase and intensity discrepancies between the 'ears'. As rotation of the head is low cost, relative to movement of the whole vehicle, hunting for signals can proceed whilst the vehicle is stationary. Once a signal is located, and directional symmetry achieved, the body can be rotated to dead-lock into a pre-determined head-body position. The robot then moves towards the beacon. However, this activity is interruptible by any 'object-approach' signal from the obstacle avoidance system, which has priority.

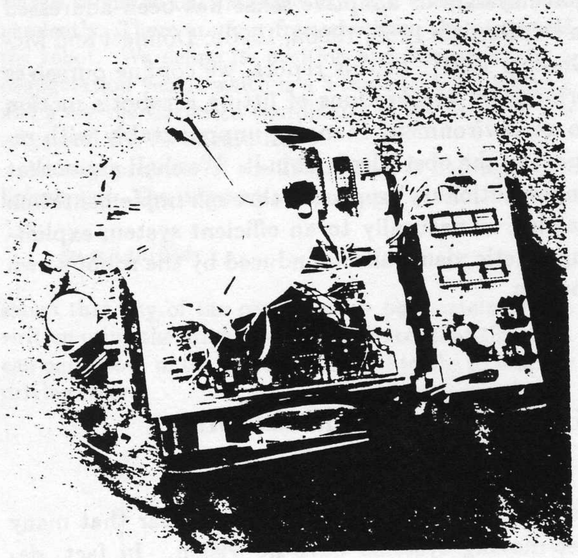

Fig. 1: The experimental vehicle.

The collision avoidance interrupt is therefore an 'object analogue' for the vehicle. Such a binary signal might at first glance appear to stand as a poor proxy for objects, inferior in all respects to the obvious power of a 'general-purpose' vision system, dealing both with object avoidance and map development. Yet the problems which emerge in designing such a system are formidable: for example, what sort of visual information is the robot going to use for obstacle avoidance, how is this different from map information, and what visual attributes will it use for specific locative identification within any map representation it may have?

3 Smooth Motion in the 'Blind' Robot

Given our control hierarchy and its prioritization, navigation is preempted temporarily when the presence of local objects demands this. It is these transitions between guidance and avoidance, qua object analogues, that constrain the operating speed of the vehicle. In an environment of varying obstacle density, the designer has two choices: set the vehicle to travel at some constant speed for all cases, or contour its speed to fit the obstacle distribution. This latter dynamic gearing of action to the environment is an important competence supported by vision. However, whilst vision potentially affords a 'simultaneous' view of environment lay-out, our blind system is limited to operating on sequential information.

3.1 Simulated Smooth Motion

Suppose for the purpose of illustration that the vehicle is traversing an environment of evenly distributed objects, roughly equal in size. In the limit, as the time actually spent negotiating an obstacle approaches zero, one might draw an analogy between the robot's trajectory (in two dimensions) and the mean free path of a particle in a gas (in three dimensions)[1].

If the objects are assumed to be n metres wide from the robot's point of view (i.e., they appear from above to have a circular zone of influence, of diameter n), if the robot can sense obstacles at distances of about n, and if the mean object density is s objects per m^2, then the robot can expect to travel

$$d = \frac{1}{ns}$$

metres between collisions. If it is travelling with

---

[1] This analogy was suggested to us by Dr. G. Hayes.

speed v m/s, it can expect a collision every

$$t = \frac{1}{vns}$$

seconds. As the vehicle travels faster within the same environment, therefore, the mean *distance* between collisions is unchanged, but the mean *time* decreases.

In reality, of course, the collision avoidance system must be active for more that zero-time, and even assuming only that collision processing takes a fixed time at any speed, it follows that for a given obstacle density, there exists a speed above which the vehicle is constantly avoiding, and never navigating. The behaviour of this vehicle is by definition nothing more than a random walk – entertaining perhaps, but of no practical use. Figure 2 shows how time spent navigating suffers as obstacle density increases. These are data from a simulation informed by the size and speed range characteristics of our robot, and show that even in moderately dense environments, the robot might spend half of its time just avoiding obstacles.

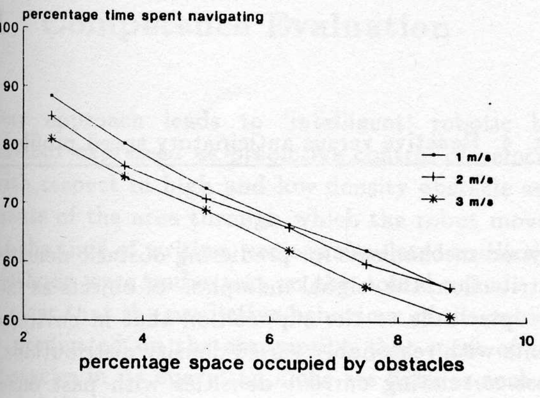

Fig. 2: Percentage time spent navigating (simulated).

3.2 Implementation

The density of environmental obstacles is available to the robot's control processors as the frequency of interruptions, by the collision avoidance system, of beacon tracking. Suppose there are m non-overlapping interrupts, happening at times t_i, $1 \leq i \leq m$. If we model interrupts as pulses of width t_c, the fixed time to process and execute an avoidance manoeuvre, then the interrupt function is

$$f(t) = \begin{cases} 1 & \text{if } t_i \leq t \leq t_i + t_c \quad (1 \leq i \leq m) \\ 0 & \text{otherwise.} \end{cases}$$

Clearly a first-order measure of obstacle density is the integral over time of this function,

$$s(t) = \int_0^t f(t)dt.$$

In terms of robot control, this observation leads to a first strike strategy for speed control:

1. From simulation and experimentation, generate an on-board table of 'best operating speeds' for a series of interrupt densities.
2. On line, low-pass filter the obstacle avoidance interrupts (this is loosely equivalent to integrating them over time).
3. Set the current speed to a function of the instantaneous value of this integral, by using the integrated interrupt frequency as an index into the 'best speed' table, as this integral corresponds to object density.

3.3 Performance

The actual robot fares worse than the simulation, as the latter does not model the inertial effects of starting and stopping; this is apparent from Fig. 3, which plots typical times-spent-navigating at two different speeds.

In any non-empty environment, higher speeds mean more time colliding and less time navigating, as both simulation and implementation clearly demonstrate. Controlling the speed of a mobile robot is thus an optimization across two conflicting variables, time to destination and energy consumption, both of which must be minimized. The minimal travelling time constraint attempts to ensure that the vehicle will eventually reach its destination, and reduces the viability of trivial speed control strategies like driving extremely slowly in order to avoid abrupt stops and starts. Energy consumption is minimized in order to maximize the life of the robot's power supply between recharges.

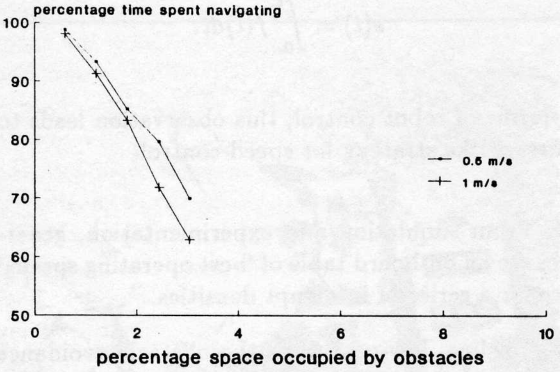

Fig. 3: Percentage time spent navigating (actual vehicle).

3.4 Predictive Control

The reactive speed-to-travel determination outlined above can be extended to a strategy for contouring speed to ***expected*** obstacle density, relative to some locative references, such as are afforded in our case by acoustic beacons. The densities along the simulated vehicle's trajectory are recorded serially, and averaged across several runs. Assuming a world in which obstacles are relatively fixed, this serial memory is used for predicting changes in density, and accelerating or decelerating in anticipation of density decreases or increases, respectively. Figure 4 contrasts speed profiles for a simulated vehicle able only to react to density changes, and one which averages densities over several runs and sets a speed appropriate to the anticipated density. The reactive profile includes sharp decelerations in response to zones of increased density, while the predictive controller, with a few seconds of look-ahead, can reduce speed less suddenly.

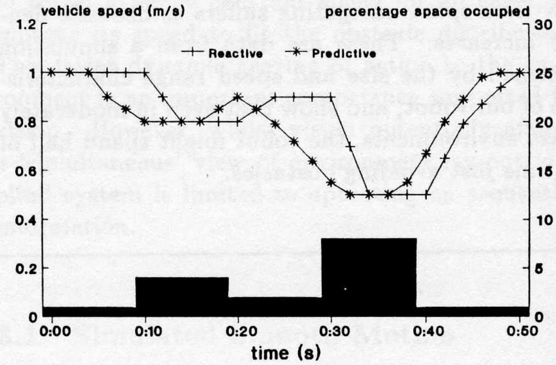

Fig. 4: Reactive versus anticipatory speed profiles.

Beyond mechanisms for predicting obstacle density distribution, the 'signal metaphor' of objects as interrupts leads to the supposition that in environments with reasonably varied density distributions, cross-correlating current densities with past ones can act as an associative memory for current position. Figure 5 depicts two typical interrupt versus time profiles, and their cross-correlation. For both plots, the vehicle followed the same trajectory between two locative references (beacons), but started at different points along it. The strong peak in the cross-correlation curve accurately reflects the magnitude of the positional shift between the two trajectories.

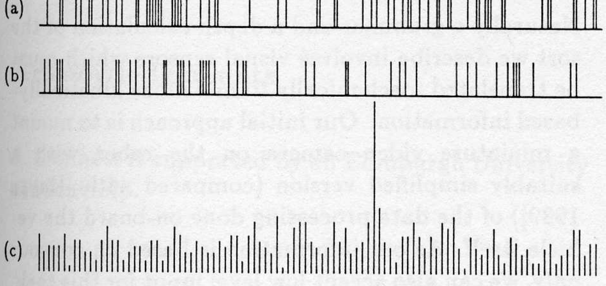

Fig. 5: A previous interrupt profile 'model' (a) and the current interrupt pattern along the same trajectory (b) can be cross-correlated (c) to determine vehicle position.

4 Competence Evaluation

Our approach leads to 'intelligent' robotic behaviour expressed as predictive control of velocity with respect to high and low density obstacle segments of the area through which the robot moves. At the time of writing, various simulated predictive methods were under test on the actual vehicle. It is clear that the predictive behaviour of the vehicle is predicated on the assumption that most of the obstacles in its operating niche are fixtures such as desks, doors, walls, etc. In the case of obstacles which are frequently displaced, the model is less adequate. However, there are several sorts of obstacle mobility to be discriminated before we can fully evaluate the system's limits. Local variation, where objects migrate but never very far from fixed objects, presents few problems, e.g., as in the usual relationship between garbage cans and office desks. Should such a constraint be removed, leaving most objects to migrate within (say) an office environment, then we would need to invest in some other form of distance object sensing to provide predictive information for motion control. In this context we have turned to some form of vision based control.

5 Vision Interface

Having arrived at a need for vision on the basis of the 'gearing speed to the environment' requirement brings us, of course, closer to issues central to machine vision research. What kinds of information may we profitably extract from the visual input to help guide our robot? Our task is constrained, because it is agent-centred within an environment which can be reasonably specified. What the agent needs to know is well limited by the nature of the locomotor task. First, it is knowledge of the relative position of obstacles as it moves towards them on a putative collision course. The faster it moves, the more it needs to 'look ahead'. As the environments we envisage the robot operating within are office regions and corridors approximately 100 metres long at maximum, we can usefully derive the nature of the predictive information which an agent travelling at a maximum speed of 1 metre per second would require to decelerate safely.

Deceleration is not enough, however. Overall, three response categories are now proposed. The first is *anticipatory* speed control where the *initial* response to the obstacle information is to decelerate (conversely it must accelerate into a clutter free region). The second component of the response is *circumnavigation*: to negotiate an obstacle, following deceleration. The third involves a longer term strategy of *optimizing* the agent's trajectory through its environment by avoiding regions of clutter in favour of comparatively open spaces (scene analysis).

The theoretical and practical advantages conferred on an active agent in terms of computational vision have been investigated with increasing interest in several laboratories, such as Rochester [Brown *et al.* 1988]. Aloimonos has analyzed the mathematical implications of active observation in detail [Aloimonos *el al.* 1987; Aloimonos and Shulman 1989]. At Edinburgh, Hayes has implemented a kinetic-based depth system [Hayes 1989; Hayes and Fisher 1990] comprising a video camera, a mechanism for translating camera position at a specifiable velocity in a fixating motion (so that a chosen point in space is always imaged at the centre of the camera's sensitive surface), and a dedicated image-processing system (a Datacube) linked to two Sun workstations. The distance to objects appearing in the image can be calculated from the space and time derivatives of the image intensity.

This system delivers depth information at a speed and resolution which a mobile robot would require, but the synthesis of a kinetic-based depth system such as this and a robot is not merely one of installation. While affluent in depth information, the implementation has, nonetheless, no interpreter so vital to the transduction of the depth values into ***prescriptions*** for obstacle avoidance: first to decelerate, then to guide the system round obstacles – alternatively to plan optimal routes through clutter whiles maintaining high speeds.

Our current method of dealing with the interpreter problem is first to compute the minimal depth values necessary for the successful deceleration of the agent when travelling at 1 metre per second. Given a stopping time of about 1 second, this distance is 1 metre. However, assuming a smooth deceleration takes 2 seconds and that the robot needs manoeuvre space round the obstacle, we have arrived at 3 metres as the first depth value which will critically control speed (deceleration). Two things follow from this. The first is that we can substantially reduce the data we need to control the agent in this way, from the depth resolution available in the described system. This brings economies in computation and weight of equipment necessary to perform the task. The second is that the semantics of depth analysis or interpretation are built into the system as a depth threshold or functional field – rationally determined by the niche, the weight and speed of the agent, and the kind of information available to it.

It will be apparent, however, that the interpreter function is not solely solved by one functional field alone. To achieve a sequence or syntax of responses to the same putative depth values clearly means that there is not a simple one to one mapping between input and response. Thus we have developed a task grammar [McGonigle 1990, 1991] designed to reinterpret the same obstacle as a function of the time course of agent approaching an obstacle. One duty of this grammar is to switch the criterion of obstacle sensing from 3 metres (triggering deceleration) to 1 metre (to enable circumnavigation).

Naturally a grammar and a depth calculation of the sort we describe involves visual sensors which must be translated mechanically to maximize kinetically-based information. Our initial approach is to mount a miniature video camera on the robot with a suitably simplified version (compared with [Hayes 1989]) of the data processing done on-board the vehicle itself. As our information is based on contour only, we can also accept low level input for this task. Therefore, a second approach is a system based on an analogue of a compound insect eye, consisting of rows of detectors or facets arranged in radial strips. The radial compound construction brings advantages over the use of a video camera because the spherical symmetry ensures that the direction of any obstacle detected is simply along the line joining the centre of the 'eye' to the facet registering the obstacle: no further geometrical calculation is required. The details of this design are documented in [Hayes, Donnett, and McGonigle 1991].

6 Summary

Arriving at the need for vision in the way we describe may certainly help establish a principled case for the precise role of vision in ecologically valid situations such as gearing speed to clutter. In addition, the resources already developed in the agent may illustrate important competences which visual systems may rely on to *ground* them to the niche over which the agent must be 'fit' or adaptive. The ability to specify movement independently of vision, for example, permits an agent to correlate head movement with visual (whole-field) transformation. These action contingent transformations (see [McGonigle 1978]) are a second source of invariance in addition to the parallax-based invariance available in the (kinetically-determined) visual array.

A further source of grounding is the sound-based competence of the robot we describe which allows the system to identify a critical region first, before analyzing it visually (say). The advantages of a sound-based orientation system like this would seem to circumvent many of the segmentational and attentional problems inherent in visual processing per se. However, these remain to be explored.

Acknowledgements

J. Donnett is supported by an Edinburgh University studentship.

References

Aloimonos, J., I. Weiss, and A. Bandyopadhyay, 'Active Vision', *Proc. 1st Int. Conf. on Computer Vision*, June 1987, pp. 35–54.

Aloimonos, J., and D. Shulman, *Integration of Visual Modules*, London: Academic Press, 1989.

Brown, C.M., D. Ballard, T. Becker, R. Gans, N. Martin, T. Olson, R. Potter, R. Rimey, D. Tilley, and S. Whitehead, 'The Rochester Robot', *Univ. of Rochester Comp. Sci. Dept. Tech Report 257*, August 1988.

Crowley, J.L., 'World modeling and position estimation for a mobile robot using ultrasonic ranging', *Proc. 1989 IEEE Conf. on Rob. and Autom.*, pp. 674-680.

Donnett, J.G. and B.O. McGonigle, 'A robot for local-area environments', in preparation.

Durieu, C., H. Clergeot, and F. Monteil, 'Localization of a mobile robot with beacons taking erroneous data into account', *Proc. 1989 IEEE Conf. on Rob. and Autom.*, pp. 1062- 1068.

Hayes, G.M., 'A real-time kinetic depth system', M.Sc. Dissertation, Dept. of Artificial Intelligence, University of Edinburgh, 1989.

Hayes, G.M. and R.B. Fisher, 'Evaluation of a real-time kinetic depth system', *Proc. British Machine Vision Conf.*, Oxford, Sept. 1990, pp. 315-318.

Hayes, G.M., J.G. Donnett, and B.O. McGonigle, 'Engineering a kinetic (motion-parallax) vision system for a mobile robot', *submitted.*

Hebert, M., 'Building and navigating maps of road scenes using an active sensor', *Proc. 1989 IEEE Conf. on Rob. and Autom.*, pp. 1136-1142.

McGonigle, B.O., 'A visual analogue of the diplophonic effect', Vision and the Senses, Section J, Brit. Assoc. for the Advancement of Sci., Lancaster, 1976.

McGonigle, B.O., 'A sound-based navigation system for robots', Lab. for Cognitive Neuroscience Memo, 1989.

McGonigle, B.O., 'Incrementing intelligent systems', Workshop on Biology, Cognition, and Robotics, Bonn, Germany, May 1990 [a].

McGonigle, B.O., 'Incrementing intelligent systems by design', *Proc. of Simulation of Adaptive Behaviour Conf.*, Paris, September 1990 [b].

McGonigle, B.O., 'A grammar for serial control in robots', Lab. for Cog. Neurosci. Memo, 1990 [c].

McGonigle, B.O. and J.G. Donnett, 'Giving a robot a grammar', in preparation.

Ozono, S. and K. Takamasu, 'Application of high-speed distance sensor to guidance for mobile robot', IMEKO TC14 Symposium, Budapest, Hungary, November 1986.

Walter, S.A., 'The sonar ring: obstacle detection for a mobile robot', *Proc. 1987 IEEE Conf. on Rob. and Autom.*, pp. 1574-1578.