Ecological Robotics

Andrew P. Duchon, Leslie Pack Kaelbling and William H. Warren

Adaptive Behavior 1998; 6; 473
DOI: 10.1177/105971239800600306

Ecological Robotics

Andrew P. Duchon*
Brown University

William H. Warren*
Brown University

Leslie Pack Kaelbling†
Brown University

There are striking parallels between ecological psychology and the new trends in robotics and computer vision, particularly regarding how agents interact with the environment. We present some ideas from ecological psychology, including control laws using optic flow, affordances, and action modes, and describe our implementation of these concepts in two mobile robots that can avoid obstacles and chase or flee moving targets solely by using optic flow. The properties of these methods were explored further in simulation. This work ties in with that of others who argue for a methodological approach in robotics that forgoes a central model or planner. Not only might ecological psychology contribute to robotics, but robotic implementations might, in turn, provide a test bed for ecological principles and a source of ideas that could be tested in animals and humans.

Key Words: ecological psychology; behavior-based robotics; optic flow; obstacle avoidance; tag

Introduction

During the last decade, a new approach, called *behavior-based robotics*, has developed in robotics (Brooks, 1991b). Although some successes were achieved in earlier work on mobile robots [e.g., SRI's Shakey (Nilsson, 1984) and Moravec's CART (Moravec, 1981)], these robots were generally of the sense-model-plan-act variety, requiring intense computation for inferring the location and identity of objects, updating a central world model, and planning a course of action to achieve some defined goal state. In contrast, the new approach (Brooks, 1991a) attempts to build up robots through networks of simple, fully functional behaviors that map sensors to actuators, with no central model. Complex behavior emerges from the dynamic interaction between

* Department of Cognitive and Linguistic Sciences, Box 1978, Brown University, Providence, RI 02912; E-mail: duchon@cog.brown.edu
† Department of Computer Science, Box 1910, Brown University, Providence, RI 02912

the agent, with its simple mappings, and the environment, producing what appears to be goal-directed action. Admittedly, most of this research is attempting to solve a class of tasks different from those posed for earlier robots, but such tasks are, arguably, more central to and necessarily must precede successful planning in the real world.

Most work in robotics uses non-visual sensors other than visual. Sonar, infrared detectors, laser-light stripers, and dead-reckoning provide metric distance information, which, traditionally, the robot uses to place itself at a particular point in its world model and to plan a metric path through the environment. The typical role of vision in these robots is to create or augment the model. Using visual information as simply another way to obtain metric values allows one to treat computer vision as a separate task, one of scene analysis: creating a description of the three-dimensional world from two-dimensional images. The numerous means of constructing such models (e.g., shape from shading, structure from motion), as they are commonly formulated, often are ill-posed problems requiring assumptions and noise models that do not generalize to real-world vision (Aloimonos & Rosenfeld, 1991). However, with active control of the visual system (*active vision*), these problems become well-posed, usually with unique solutions and a few reasonable assumptions (Aloimonos & Rosenfeld, 1991; Ballard, 1991; Bajcsy, 1988). Nevertheless, this line of research essentially investigates more efficient means of obtaining the same end: a reconstruction of the environment. More recently, it has evolved into *purposive* or *animate vision* (Ballard & Brown, 1992; Aloimonos, 1993), which attempts to go one step further by posing the question, "What is vision for?" (Ballard, 1991). The first purpose of vision is to create those relationships between the animal and its environment that are necessary for survival. It may be that the animal does not need to model its environment to achieve these relationships. Furthermore, when internal representations are required, they will be derived from the ways in which vision is used in this primary function.

2 Ecological Psychology and Robotics

Much of the research in animate vision and behavior-based robotics was anticipated by the work of J. J. Gibson, and we would like to probe further into the relevance of his ideas [see Pickering (1992) and Effken & Shaw (1992) for other points]. Ecological psychology, as developed by Gibson (1950, 1966, 1979), views animals and their environments as "inseparable pairs" that should be described at a scale relevant to the animal's behavior. So, for example, animals perceive the layout of surfaces (not the coordinates of points in space) and what the layout *affords* for action (not merely its three-dimensional structure). A main tenet of the ecological approach is that the optic array, the pattern of light reflected from these surfaces, provides adequate information for controlling behavior without further inferential processing or model construction. This view is called *direct perception*: The animal has direct knowledge of, and a relationship to, its environment as a result of natural laws. The degree to which perception plays a role in cognition is an open question but, minimally, the information involved in both perception *and* action could ground other, nonperceptual tasks. The strategy is to push natural law as far as possible into cognition, thus placing more constraints on the cognitive system.

The Gibsonian approach can be summarized in the idea that it is more desirable to put the animal in its environment than to put the environment into the animal. Rather than internally representing detailed knowledge of the world, animals detect and use information about the world as required. This is the "fundamental hypothesis" of the ecological approach to vision:

> Optic structure specifies its environmental source and . . . therefore, *mobile* organisms with *active* visual systems that can pick up this information will see their environments and suitably adjust their activity, if and when they detect that information, and only then (Turvey, Shaw, Reed, & Mace, 1981, p. 253; emphases ours).

Now, if we replace "mobile organisms" with mobile robots or, more generally, agents, this hypothesis is just as applicable to behavior-based robots as it is to animals: That is, sufficient information is available in the robot-environment interaction to control the robot's behavior without further inference or reconstruction. In addition, appropriate perception-action dynamics in the robot provide a noninferential source of information on which other processes (planning, mapping, reasoning, etc.) could be based and by which they could be limited.

Similar hypotheses might be made in regard to the other senses or sensors, but it is primarily vision that appears to be the most promising for relating the fields of robotics and ecological psychology. Both would gain from such crosstalk. The latter can provide insights into the kinds of information that can control the actions of agents—that is, what ecological laws might be exploited in a given task; and the former can provide a test bed for ecological principles, such as exploring the viability of hypothesized control strategies and facilitating the discovery of new ones. The new robotics and the ecological approach share common concerns and complement each other well.

Optic Flow and Control Laws

A relevant case is the study of optic flow. As an observation point moves through the environment, the pattern of light reflected to that point changes continuously,

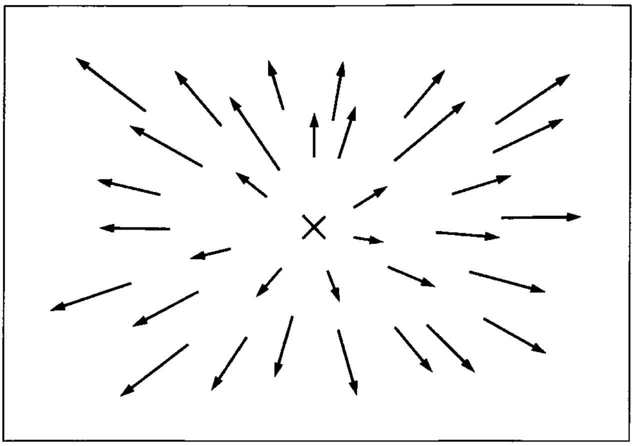


Figure 1
The projection of optic flow onto a two-dimensional surface, creating a velocity field. Shown is the instantaneous velocity field produced by pure observer translation toward a frontal plane. The $\times$ marks the focus of expansion, the point from which all the vectors radiate. In the case of pure translation, it marks the point on the surface toward which the observer is heading.

creating *optic flow* (Gibson, 1958; Lee, 1980). Optic flow (Fig. 1) contains information about both the layout of surfaces and the motion of the point of observation. For example, if an observer is translating, the *focus of expansion* (FOE), or center of the radial flow pattern, specifies the observer's heading (Warren, 1988). If the observer is moving at a constant velocity, then the time to contact with a surface is given by the relative rate of expansion $\tau = \beta/\dot{\beta}$, where τ is the "optic variable" *tau*-global (Lee, 1976; Tresilian, 1991), β is the visual angle between a point on the surface and the FOE, and $\dot{\beta}$ is the rate of change in this angle. The term $\dot{\beta}$ refers to the radial component of the optic flow (during pure translation); more generally, we will refer to optic flow vectors as $\vec{\mathbf{w}}$.

The observer's heading and time to contact are just two examples of information available in optic flow. One way an agent can use this information is by acting to achieve a certain type of flow. For example, to maintain ambient orientation, the type of optic flow required is no flow at all. If some flow is detected, then the agent should change the forces produced by its effectors (whether wings, wheels, or legs) so as to minimize this flow, according to a *law of control* (Warren, 1988):

$$\Delta F_{\text{internal}} = f(\Delta \mathbf{w}) \tag{1}$$

That is, the change in the agent's internal forces (as opposed to external forces such as wind) is a function of the change in the optic flow (here, from no flow to some flow). The Δ in this formulation implies control based on changes from initial conditions, rather than a fixed reference level, although either may be used in specific instances.

Gibson (1958, p. 187) described various control laws that an animal might use for locomotion:

> . . . to begin locomotion, therefore, is to contract the muscles so as to make the forward optic array flow outward. To stop locomotion is to make the flow cease . . . To aim locomotion at an object is to keep the center of flow of the optic array as close as possible to the form which the object projects.

These types of rules have been noted by scientists studying the control of balance, steering, and braking in humans (Lee, 1976; Lee & Lishman, 1977; Yilmaz & Warren, 1995; Warren, Kay, & Yilmaz, 1996) and the control of flight in flies (Collett & Land, 1975; Reichardt & Poggio, 1976; Wagner, 1986a,b). Ambient orientation, or hovering, is controlled by minimizing the global optic flow: Purely vertical flow (say, upward) will induce increased lift by the fly to minimize that flow (Srinivasan, 1977; Götz, Hengstenberg, & Biesinger, 1979). Similarly, a fly in a rotating drum will produce a differential thrust with the two wings, tracking the drum by rotating about its own vertical axis (Collett, 1980a,b).

Warren (1988) proposed a set of control laws that a fly might use for each of its major activities. For example, the laws of control for hovering in the face of vertical and horizontal flow, respectively, could be

$$\Delta U = (k/c)\Delta\vec{\mathbf{w}}_{\mathrm{v}} \tag{2}$$

$$\Delta(F_{\mathrm{L}} - F_{\mathrm{R}}) = (k/c)\Delta\vec{\mathbf{w}}_{\mathrm{h}} \tag{3}$$

where U is the amount of upthrust generated by the two wings, (k/c) is the ratio of the drag constant to an optic scaling coefficient, $\vec{\mathbf{w}}_{\mathrm{v}}$ is the vertical component of the optic flow, F is the forward thrust given by a wing, and $\vec{\mathbf{w}}_{\mathrm{h}}$ is the horizontal component of the optic flow. These laws assume that changes in force are what the fly can control and, given a critically damped system (e.g., the drag in insect flight), a constant force quickly equilibrates to a constant velocity. These kinds of control laws can be summarized by the idea that motion (visual) begets motion (of the agent).

Which control laws govern the fly's behavior at any one time depend on the goal, or "action mode," of the fly (Warren, 1988): cruising, landing, foraging, pursuing conspecifics, and so forth. For each of these action modes, objects in the environment will "afford" certain behaviors. "The affordance of anything is a specific combination of the properties of its substance and surfaces taken with reference to an animal" (Gibson, 1977, p. 67). Strictly speaking, the affordances of surfaces in the environment are constant for a particular animal (Gibson, 1979) and are discovered in the course of learning. Every action mode constrains which affordances of an environment an animal might use on a particular occasion. For example, while foraging, a fly might use a flower's affordance of nourishment and support, while avoiding all other surfaces. However, when tired, the fly might avoid flowers but use the affordance of a resting place that large stationary objects have. Once an action mode is adopted, laws of control determine the actual behavior of the fly.

We should consider now the objection that ecological psychology appears to be similar to stimulus-response psychology. The emphasis on direct perception of the environment, especially for motor behavior within the environment, sounds as if the environment is controlling the behavior of the animal, that the animal is simply responding to stimuli. This is a misconception. Rather, the animal is *using* information in the environment as a resource to control its own goal-directed actions. The animal perceives the environment in terms of affordances. A decision about the consequent behavior must be made based on this perception. Then, to carry out this behavior, the animal can adopt a particular control law. For example, through various forms of information, one can perceive that the hallway affords locomotion, the chair sitting, the steps climbing, and so forth; one will choose among these possibilities based on one's current action mode. To carry out the act of walking down the hallway, one uses control laws, directly relating what one sees to what one does, regulating posture, heading, and stride.

4 Ecological Robotics

In the following sections, we discuss some work demonstrating that control laws such as those outlined in the previous section can be used successfully to control a mobile robot. We call this *ecological robotics*, the practice of applying ecological principles to the design of a mobile robot. We can briefly summarize these principles as follows:

1. Because of their inseparability, the agent and the environment together are treated as a system.
2. The agent's behavior emerges out of the dynamics of this system.
3. Based on the direct relationship between perception and action, the task of the agent is to map available information to the control parameters at its disposal in order to achieve a desired state of the system.
4. The environment provides enough information to make adaptive behavior possible.
5. Because the agent is in the environment, the environment need not be in the agent. That is, no central model is needed, but this does leave room for task-specific memory and learning.

Our work in ecological robotics has followed a series of steps from the concrete to the abstract. First, we dealt with the problem of a real robot wandering and avoiding obstacles, using optic flow (Duchon & Warren, 1994) in a manner similar to that of others implementing ideas from the insect literature. Then we made some progress on a robot playing tag with a hand-held target without needing to segment the scene (Duchon, Warren, & Kaelbling, 1995). These two operations have been transferred easily to a new, larger robot with little change, attesting to their generality. Recently, we have taken the strategies that were implemented in the robots and explored in simulation their utility and limits.

Obstacle Avoidance

5.1 Control laws

Our first work (Duchon & Warren, 1994) investigated the use of control laws for the most crucial ability of a mobile agent: avoiding obstacles. Srinivasan, Lehrer, Kirchner, and Zhang (1991) have shown that bees will steer down a corridor by using a simple strategy that, in normal circumstances, enables them to fly down the exact middle of the corridor. However, when one of the walls is in motion, bees will fly closer to a wall moving in the same direction of their flight and farther from a wall moving in the opposite direction. Further tests showed that the bees were moving so as to equate the optic flow in the lateral portions of the optic array of each eye. This behavior is termed the *centering response*, but it would not be useful for obstacle avoidance in a cluttered environment, because objects need to be avoided before they are reached. Hence, we have investigated a control law called the *balance strategy*, which takes into account the entire field of view:

$$\Delta(F_L - F_R) = k\left(\frac{\sum \|\vec{\mathbf{w}}_L\| - \sum \|\vec{\mathbf{w}}_R\|}{\sum \|\vec{\mathbf{w}}_L\| + \sum \|\vec{\mathbf{w}}_R\|}\right) \tag{4}$$

where $\Delta(F_L - F_R)$ is the difference in forces on the two sides of the agent's body, k is a constant, and $\sum \|\vec{\mathbf{w}}\|$ is the sum of the magnitudes of optic flow in the visual hemifield on one side of the FOE.

The essential idea behind this strategy is that of motion parallax: When the agent is translating, closer objects give rise to faster motion across the retina than do farther objects. It also takes advantage of perspective in that closer objects also take up more of the field of view, biasing the average toward their associated flow. The agent turns away from the side of greater flow but only so that it does not turn into something on the other side.

This strategy is like many other gradient-descent mechanisms used by animals, what Gibson (1966) called "the principle of symmetrical stimulation in orientation," which include the taxes and tropisms, such as chemotaxis and phototropism. Here, the agent moves down the gradient of a difference in flow into the valley where the difference is zero (e.g., into the middle of a corridor).

5.2 Robot implementation

5.2.1 The robots We have successfully implemented the balance strategy on two mobile robots. The small robot, named Louie, is an RWI B-12 (30.5-cm) base on which is mounted a single camera having a 60-degree horizontal field of view (FOV), placed approximately 75 cm off the ground. A fast patch-matching optic flow algorithm (Camus, 1994) provides a dense 128 $\times$ 32 flow field at four frames per second, allowing us to control Louie moving at a speed of 4 cm/sec.

The large robot, Ramona, is an RWI B-24 (61-cm) base on which is mounted a single camera with a 110-degree FOV placed at a height of 120 cm. Optic flow input is obtained using the Teleos AVP-100 vision system, which gives motion information (128 $\times$ 92 pixels) at approximately 10 Hz. This allows Ramona to move safely at speeds up to 30 cm/sec.

5.2.2 The environments Louie was tested only in the Brown University Artificial Intelligence Laboratory (AI lab), which, typically, is poorly lit and tightly constrained with numerous chairs (having textureless backs), people, and various wires on the floor. The area beneath the tables is relatively dark, and the camera was placed on the robot at a height just under the table tops, which had black metal rims from which no motion could be registered. Because of these conditions, we equipped the robot with two "emergency" reflexes. The *intensity reflex* stopped and turned the robot 90 degrees when it became too dark or too bright (i.e., when no flow could be detected). We also estimated τ from the optic flow field but, for simplicity, we used $\eta = 1/\tau$, which can be thought of as immediacy: The greater η is, the more immediate is an impact. Thus, the *eta reflex* produced the same response (stopping and turning 90 degrees) when a crash was imminent ($\eta > m_\eta$, where m_η is a "margin value").

Ramona has been run in both the AI laboratory and in a more open atrium that contains tables, chairs, and a few trees. Most surfaces in the atrium are textureless, but it is generally well-lit. The same two reflexes were used, except that Ramona turned 180 degrees. Because the camera was now approximately 50 cm above most of the objects, the camera was tilted down at a 45-degree angle. Unrecorded runs in the AI lab have lasted up to 25 minutes without collision. The recorded runs discussed in the following section were all in the atrium.

5.2.3 Robotic implementation of the control laws The control laws previously described are idealizations that might be used by biological agents but that require approximations in artificial agents. The use of force in the control laws implies acceleration but, as noted earlier, in a damped system this quickly results in simple changes in speed (Δs), which is what was controlled in the robot's translation. For rotations, the $\Delta(F_L - F_R)$ in Equation 4 was reduced to the term Δr, which gave an angular velocity value in degrees per frame. With both robots, the input to the control laws was the matrix of vectors given by their respective optic flow algorithms. In practice, if $\sum \|\vec{\mathbf{w}}_L\| + \sum \|\vec{\mathbf{w}}_R\|$ did not exceed a threshold, then Δr was set to zero.

These control laws assume that only optic flow due to translation is used as input. With Louie, the rotations were made fast enough relative to the translational speed that flow from rotation overwhelmed that due to translation. In this manner, flow from rotation was essentially ignored, as the amount of optic flow was equal on the two sides, resulting in no rotation in the next frame. With Ramona, this method could not be used, but the frame rate was fast enough that it could simply make "saccadic" movements every three to five frames based on the flow detected in the current frame. This allowed time for a rotation to be made and motion to be detected from at least two frames of pure translational movement. Because we can thus assume an initial angular velocity of zero, the left-hand side of the control laws for rotation in the robots (and the simulations) can be further reduced from Δr to r.

Finally, the two hemifields in the control laws are given with respect to the FOE. We assume the FOE will be in the middle of the horizontal FOV of the camera so we can also split the vector field of flow values in the middle. We have occasionally used an algorithm similar to that of Camus (1994) to determine the location of the FOE in the vector field. If the FOE is not in the middle, which may happen if the camera is not perfectly aligned to the wheels, then the left and right hemifields are taken with respect to the calculated FOE and are the size of the smaller field. However, we did not perform this calculation for the work presented in the following subsections.

Louie wandering Duchon and Warren (1994) give numerous examples of Louie wandering through the AI lab using the balance strategy (among others). Dark regions under the tables and textureless chair backs caused problems. However, Louie could avoid hands placed suddenly in its path and cross the length of the cluttered AI lab (12 m) without collision, demonstrating the utility of the control law.

Ramona wandering Because a greater range of speeds could be used with Ramona than with Louie, we added a control law in which speed is a function of the total magnitude of optic flow detected:

$$s = s_{\max}\left(\frac{\sum \|\vec{\mathbf{w}}\|_{\max(\text{possible})} - \sum \|\vec{\mathbf{w}}\|}{\sum \|\vec{\mathbf{w}}\|_{\max(\text{possible})}}\right) \tag{5}$$

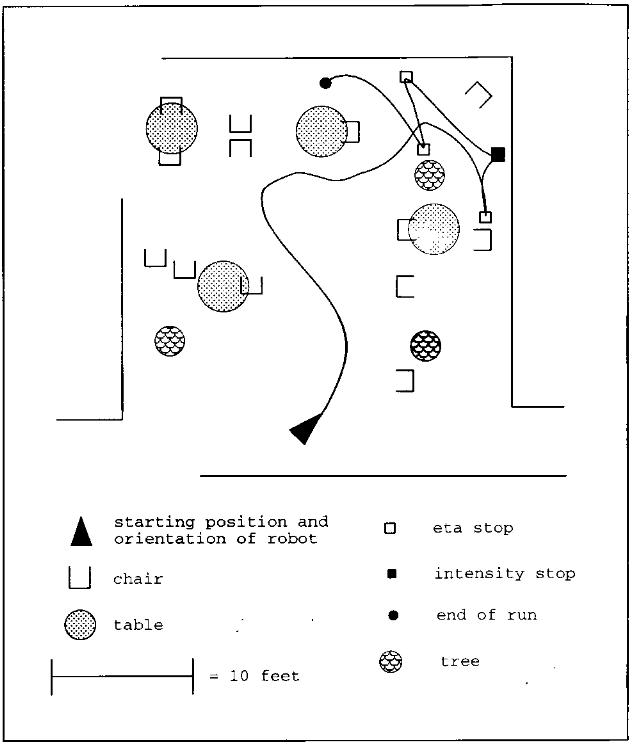


Figure 2
Ramona wandering in the atrium. Given the field of view and the particular angle at which Ramona approaches the chairs in the middle of the figure, Ramona turns right instead of into the open area on the left. Nevertheless, Ramona does manage well in the tight spaces of the corner, where sonars would have difficulty detecting chairs and tables. The trial is ended after approximately 2.5 minutes, owing to multiple emergency stops. The path is hand-drawn from video.

In cluttered environments (where more flow is detected), the robot will slow down; in open environments, it will move at the maximum speed allowed.

Two other modifications were made because the camera was tilted. First, the values input to the balance strategy were weighted more toward the bottom of the image so that Ramona would avoid obstacles at its "feet" more. Second, the *eta* calculations took into account the camera tilt. Figures 2 and 3 show the path of Ramona wandering in the atrium. The most surprising aspect of implementing the balance strategy on Ramona was the relative ease with which it was done, this despite the change in weight, size, height, and FOV.

5.2.4 Related work At least two other groups have independently implemented similar ideas. Sandini, Santos-Victor, Curotto, and Garibaldi (1993) built a robot that would balance the flow seen from two cameras facing laterally. Coombs, Herman, Hong, and Nashman (1995) have designed a robot with two cameras facing forward, one with a wide-angle lens (115 degrees) and one more foveal (40 degrees), both of which are controlled with active gaze stabilization. Whereas they balance the

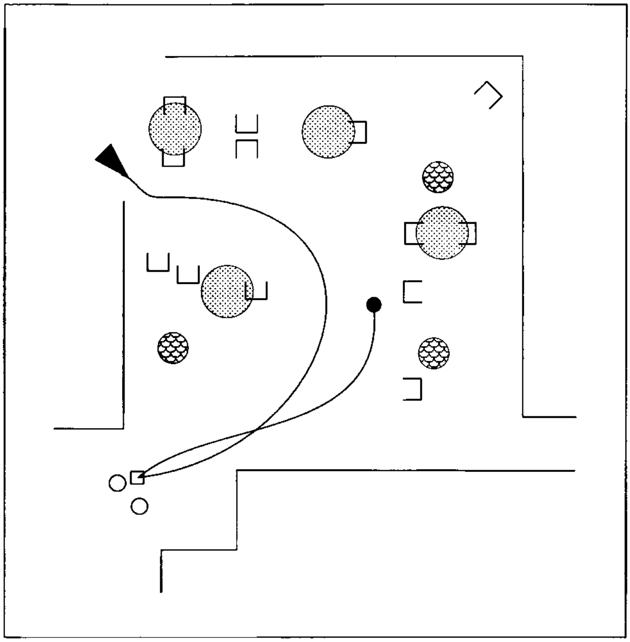


Figure 3
Ramona wandering in the atrium. Ramona wanders gracefully through the atrium, stops and turns around when she nears the human who suddenly appears in her path, then wanders back into the atrium, after which the trial is ended.

maximum values of the optic flow in the left and right peripheral fields, we balance the normalized sums of the entire left and right fields relative to the FOE. We have found that using the normalized sum is less susceptible to noise and incorporates information from more distant objects (not just the closest) in an appropriate way by adding a slight bias toward the more open space. Also, no gaze stabilization is required, as the robot itself makes saccadic movements, which, when the frame and rotation rates are fast enough, make the robot's movement appear to be smooth. Recently, Srinivasan's collaborators have themselves developed a mobile robot and simulator (Weber et al., 1997). Using a number of mirrors, the robot has a wide field of view for a similar implementation of the Balance Strategy. In addition, they use pseudo-stereo images to gauge distance in front of the robot, and their particular optic flow algorithm returns the slope of surfaces allowing them to determine the axis of a corridor. They have also added mechanisms for turning at corners, wall-following, and distance-travelled estimation using optic flow.

Other robotic implementations based on the insect literature (Franceschini, Pichon, & Blanes, 1992; Sobey, 1994) have used metric distance calculations and a mapping function for planning a (local) path or for tracking (cf. later). This information is totally absent here (at least explicitly). Aloimonos (1992) and Nelson and Aloimonos (1989) took a similar approach to ours but required an additional intermediate representation of a "hazard map" based on normal flow to find the heading of the safest path. Cliff (1991) has done simulations in a similar vein, exploring development of the optomotor response, but this is a very different mechanism from the balance strategy (Srinivasan et al., 1991).

5.3 Simulations

Robot experiments are slow to run and difficult to control precisely. Simulations are a good way to explore the space of possible behaviors, though the successful ones must eventually be tested on a real robot. Our simulations were run in a simple environment in which simulated agents could move around. The environment was designed to demonstrate the behaviors of avoiding a small object, going through a small aperture, negotiating a corner, going down a hall, and interacting with other agents.

5.3.1 Calculation of the optic flow For the simulations, the instantaneous optic flow alone was used, and it was calculated analytically. No rotational components were added, so no saccadic movements were required as with Ramona. The human visual system appears to solve this problem using both optic flow itself and extraretinal (camera position) information (Warren & Hannon, 1990; Royden, Banks, & Crowell, 1992), such that translational heading can be determined with some precision.

Because only the horizontal dimension and translational motion are simulated, $\vec{\mathbf{w}}$ and $\dot{\beta}$ are equivalent, and the latter will be used in the equations referring to the simulations. The usual equation for optic flow (Gibson, Olum, & Rosenblatt, 1955; Nakayama & Loomis, 1974) is given by

$$\dot{\beta} = \frac{\|\vec{V}_{\mathrm{O}}\| \sin \beta}{d} \tag{6}$$

where β is the angle from the heading, $\vec{V}_{\mathrm{O}}$ is the velocity of observer O, and d is the distance to the point P. The variables are represented by the solid lines and $\dot{\beta}_1$ in Figure 4. This equation gives the relative velocity of P projected onto the normal to the line of sight ($\gamma = \|\vec{V}_{\mathrm{O}}\| \sin \beta$), divided by the distance between P and O. If the point P is moving independently, then it is the relative motion ($\vec{V}_{\mathrm{R}} = \vec{V}_{\mathrm{P}} - \vec{V}_{\mathrm{O}}$) that is projected onto the normal.

Equation 6 is true only in the limit as Δt (the time between frames) goes to zero. This could not be assumed in the simulations, so $\dot{\beta}_2$, shown by the dotted lines, was used. It can be obtained from the law of cosines:

$$\dot{\beta} = \arccos\left(\frac{d^2 + x^2 - \|\vec{V}_{\mathrm{O}}\|^2}{2dx}\right) \tag{7}$$

Because the value of $\dot{\beta}$ will always be positive, we still need to determine the sign of the flow—that is, positive for outward (away from the heading that is assumed to be

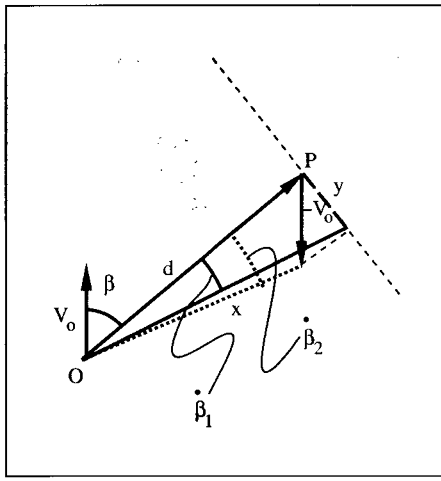


Figure 4
Variables used to calculate optic flow in simulation.

tied to the center of the FOV as it is in the robot) and negative for inward. If values of β are taken to be positive to the right of the heading, then

$$\dot{\beta} = \dot{\beta} \times \left[\text{sign}\left(\text{sign}(\beta) \times \left(V_{d_x} V_{r_y} - V_{d_y} V_{r_x}\right)\right)\right] \tag{8}$$

The final value that needs to be calculated is $\eta = 1/\tau = \dot{\beta}/\beta$ for small values of β. In general, the average of a few samples around the heading ($|\beta| < 5$ degrees) was used to estimate η.

5.3.2 The agent and the simulated environment The agent is an observation point with an FOV that can vary from 5 to 360 degrees and, although it is a point, a triangle (or other shape) is drawn around it for display purposes (and it is this that other agents "see"). The environment is modeled by straight-line segments parameterized as $(x, y) = (1-t)P_1 + tP_2$, so $0 \leq t \leq 1$. The FOV of the robot is sampled every few degrees left and right of center. We can then check for the intersection between a unit-parameterized line of sight and all line segments, which will return the distance to each intersected line segment, and then simply choose the segment that is closest. Included in this calculation are the line segments of each of the other agents, if any. The agent receives a number of samples from the optic flow (48 were used here), and these signals are the inputs for the appropriate control laws.

5.3.3 Observations

State attractors When we simulated a single agent moving at a fixed speed through a fixed environment, we were first struck by the fact that starting from a random position and orientation, the trajectory would settle into one of a few basins of attraction

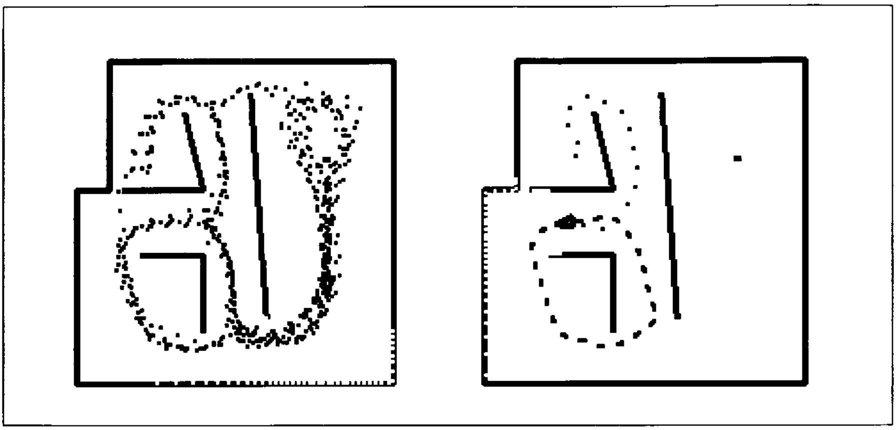


Figure 5 Path taken by a noiseless agent. The agent is an open triangle. A dot represents the agent's location every 20 frames. The notches in the walls indicate the points currently being sampled by the agent. (Left) An agent with 120-degree field of view. After starting in the lower right, the agent quickly produces a repeated path through the environment. The behavior is strikingly reminiscent of the pacing of caged animals. (Right) An agent with a 150-degree field of view. After starting in the upper left, the agent falls into one of a few states available (others exist in the upper left and on the right side).

(Fig. 5). This is a graphical representation of emergent behavior determined by the interaction between the agent's capabilities and the structure of the environment.

Adding noise to this system resulted in several attractive paths (Fig. 6), with occasional transitions between them. Gaussian noise (with the variability σ specified) in the inputs produces behavior dominated by the periphery where the magnitudes far exceed the noise. The balance strategy also is resistant to noise in the outputs. The resulting behavior of the agent could not be predicted from either the agent's capabilities or the environment taken separately (Smithers, 1995; Steels, 1995). In addition, having more than one agent present produced an even more complex pattern of motion through the environment (see Fig. 6, bottom), though here, too, one often can see repetitive patterns after longer periods.

Varying the FOV The previous few figures have shown paths of an agent having either a 120- or a 150-degree FOV. Narrowing the FOV to 60 degrees (while preserving the number of samples), the same as our robot implementation, resulted in much worse behavior. The most distressing aspect was that if the agent started in the right configuration, it would head directly into a corner, at which point an emergency *eta* reflex would be used to turn the agent around. In attempting to balance the flow with a narrow FOV, the corner has the same properties as a hallway, so the agent goes down the middle (into the corner). This was also true of the real

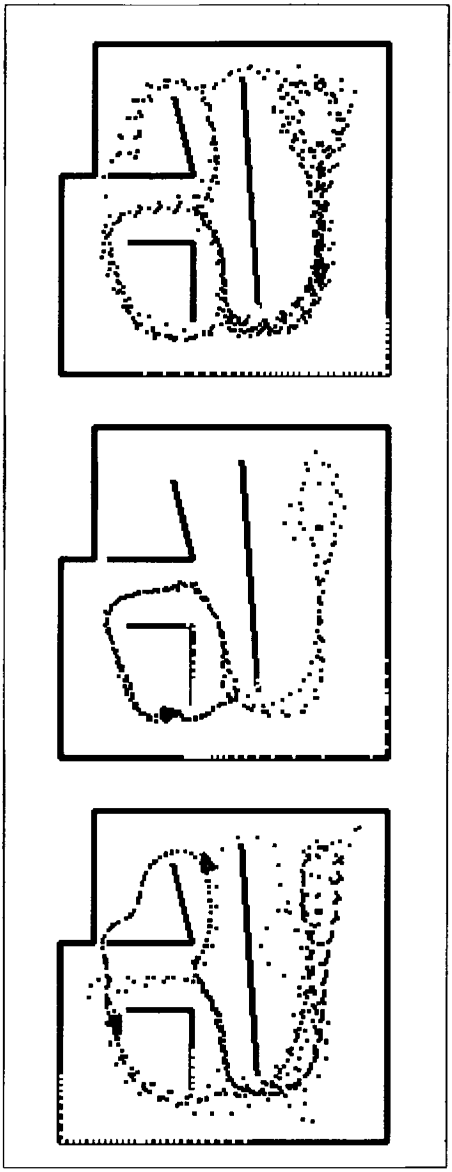


Figure 6 Path taken by an agent with input noise ($\sigma = 0.01$). (Top) Agent with 120-degree field of view (FOV). The agent is buried in the dots in the lower right. With noise, it can only occasionally go through the small opening in the upper left. Notice in Figure 5 (left) that the agent almost enters the passage on the middle left. Here, due to noise, it takes that passage most of the time. (Middle) Agent with 150-degree FOV. The agent makes occasional forays into the right half of the environment, after which it returns to the same circular path in the lower left as in Figure 5b. (Bottom) Two agents in the same environment. The dots indicate the path of the 120-degree FOV agent in the lower left. The path and sample points of the 150-degree FOV agent are not shown.

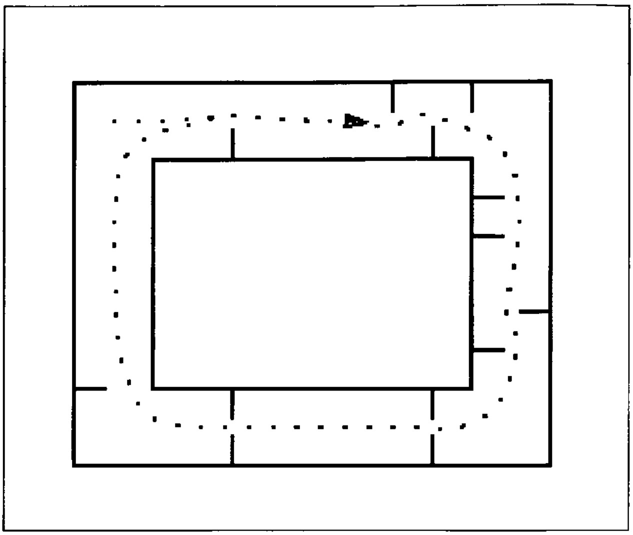


Figure 7
The circuit used to test path length and resistance to noise. Here, an agent with a 120-degree FOV and no noise makes a path with a length between the shortest and the safest routes.

robots. As the FOV increases, the range of angles of approach for which this is true decreases. Increasing the FOV to more than 200 degrees again resulted in reduced performance as the agent avoided things it had already passed.

Varying the speed of the agent Varying the speed of the simulated robot showed that the passability of various apertures (e.g., the small opening in the upper left of the environment shown in Figures 5 and 6) changed with speed: Slower agents would go through, whereas faster ones would turn around and leave the area the way they had come. One could speculate that a combination of an animal's most efficient speed, body width, FOV, motion perception thresholds, and "noticed" spacings of its environment are all related and perhaps invariant across species.

Optimal paths and resistance to noise We looked into two final properties of the balance strategy—namely, how "good" a path an agent would take through an environment, and how resistant to noise it would be. For this purpose, we created another environment, the circuit (Fig. 7), for which the agent always started in the upper left corner, pointing to the right, and had to make one loop through the circuit. The "best" path could be either the safest—a path that kept the maximum distance between the agent and the walls at the moment of passage—or it could be the absolute shortest path that led to no collisions.

We can measure these distances for the circuit. The outer walls (in arbitrary units) are 60 × 50, the inner walls are 40 × 30, the obstacles stick out 4 units from the wall, and the agent is 2 units wide. If the agent starts in the upper left aligned with the inner left wall and moves clockwise around the circuit to cross that line again, then the safest path is 184.58 units. The shortest path, with the agent just nicking the walls (i.e., the center of the agent is 1 unit from a wall) and making the "tightest" circle, is 147.93 units. Under the balance strategy with no noise and a 120-degree FOV, the agent makes the circuit in 164.85 units—approximately halfway between the shortest and the safest path lengths. Adding a small Gaussian noise to its inputs ($\sigma = 0.01$) leads to paths of 157.73 ± 0.53 units (from ten runs). With $\sigma = 0.02$, the agent completes the circuit approximately 50 percent of the time. The average path length of ten of the completed trials was 175.42 ± 5.90 units. So, up to the point where the agent can still complete the circuit, its path length is between the safest and the shortest path lengths, representing a compromise between safety and effort. However, these results depend on the constraints of the environment. In a less symmetrical environment, where the outer wall is far from the inner wall and the obstacles on the sides do not protrude much, the agent will approach the safest-path length.

The Game of Tag

Wandering around and not hitting things may help an agent avoid getting hurt, but survival requires more interactive behavior such as that exhibited in predator-prey relations. A lion attacking a herd of zebras will chase the closest or slowest animal in the herd and then use its claws to bring down the prey. The prey, for its part, must recognize that it is being attacked and ensure that the predator comes no closer. In humans, the children's game of tag is a (usually) gentler form of this interplay.

Our implementation of tag is slightly different from the typical game. We consider tag to be an action mode, similar to foraging, composed of three subordinate action modes: watching, chasing, and escaping. Instead of the concept of *it*, there is only a watcher and a wanderer. While in the watching mode, the watcher does not move until the wanderer appears in its FOV. It then fixates the wanderer by centering it in the FOV and tracking it throughout the trial. If the wanderer approaches the watcher, then the watcher flees until it is caught or successfully escapes. If the wanderer withdraws from the watcher, then the watcher will attempt to chase it. The wanderer is "caught" when η is above a certain threshold.

6.1 Control laws

Described here are theoretical control laws and the transition rules between them. Two very similar sets of control laws have been used, one with the small robot, Louie, described previously (Duchon et al., 1995), and one set used more recently with Ramona and in simulation.

6.1.1 Fixating the target The first goal of the watcher is to obtain a clear view of the wanderer in order to determine better whether it is attacking or fleeing the watcher: That is, the watcher should foveate the wanderer, or center it within the FOV. This has been termed *fixating* in the insect literature (Reichardt & Poggio, 1976; Collett, 1980a), but note that the control laws here are *not* based on specific works in the insect literature [e.g., Collett & Land (1975), which is the basis of the simulations done by Cliff (1992)].

With segmentation, fixation can be achieved by the simple position control law:

$$(F_L - F_R) = k(\beta_t) \tag{9}$$

where β_t is the optic position (on the horizontal axis) of the center of the target relative to the center of the FOV if the agent is stationary, relative to the FOE if chasing, or relative to the focus of contraction if escaping. The Δ is dropped here because we are not concerned with starting from any initial state (as we are in Equations 2 and 3) but with achieving a privileged position of the target on the retina.

A more advanced agent could take into account the optic velocity of the target itself, allowing for a certain amount of prediction of where the target will be after the rotation is made:

$$(F_L - F_R) = k(\beta_t + \dot{\beta}_t) \tag{10}$$

Without segmentation, the balance strategy that was used for obstacle avoidance (Equation 4) can be used, but with the signs reversed:

$$\Delta(F_L - F_R) = -k\left(\frac{\sum \|\vec{w}_L\| - \sum \|\vec{w}_R\|}{\sum \|\vec{w}_L\| + \sum \|\vec{w}_R\|}\right). \tag{11}$$

That is, more motion on the left or right side of the FOE or center of the FOV would induce the agent to turn toward that side. Once the amount of motion is equal on the two sides, rotation will stop. If the target moves again, the agent will follow it by moving to the side on which the target has moved (if the target is the only moving object in the scene). While chasing, docking, and even escaping, Equations 9 and 10 will act well enough, but Equation 11 will be functional only as long as the target is the sole or fastest moving object (e.g., a rabbit in a field), the target is the closest object, all the moving objects are potential targets (e.g., a school of fish), or the motion signals are subsequently filtered [see Prokopowicz, Swain, & Kahn (1994) for further discussion of how and when to use various kinds of information for tracking—though we would argue that motion signals can still be used when the agent is moving].

6.1.2 Escape or chase? Gibson (1958, p. 189) writes:

> From the point of view of the prey, the expansion of a textured contour in the optic array means the approach of something. This in itself may touch off the reaction of flight if it comes within the field of view ... There is one other geometrical possibility of stimulation: contraction of a contour in an otherwise static array. This means something going away.

The latter situation may, in fact, provoke a predator to pursue. Lorenz (1980/1952, p. 148) writes of the jackdaws he raised: "The urge to fly after an object moving away from it is very strong in a young jackdaw and almost takes the form of a reflex action." He also notes that to repel a flock of geese, "the loudest shouts, the wildest waving of arms have no effect whatever" (p. 3), but opening an umbrella "with a sudden jerk" (which would produce an expansion pattern) causes the geese to take to the air.

Looking at either the segmented target or a patch around the center of the FOE (given that fixation takes place in parallel, an assumption for all the variables used in the following sections), escape is triggered by a value of η above the margin value, m_e: That is, if the wanderer approaches the watcher at a rate causing a sufficient expansion in the center of the agent's FOV, then the watcher will attempt escape. Chase is similarly triggered by a sufficient contraction, indicated by $\eta < -m_e$.

6.1.3 Escaping In the following sections concerning escaping and chasing, only the change in the sum of the forces, $\Delta(F_L + F_R)$—that is, only the translational acceleration and not the rotational—need be considered, because the rotational acceleration, the difference in forces, is taken care of by the control laws for fixation.

If the watcher determines that it is being chased, then it must retreat. For ease of exposition, we will consider only the case in which the watcher is facing the target and backing away. The control law for escape is:

$$\Delta(F_L + F_R) = -k(\eta + \mu) \tag{12}$$

That is, as long as the flow of the patch is not inward, $\Delta(F_L + F_R)$ should be negative (i.e., accelerating backward if watching). If $\eta = 0$, the agent should still retreat; otherwise, it would simply be keeping a constant distance from the target. The minimum rate at which it should accelerate away is given by $\mu > 0$. If the direction of flow is inward and greater than μ, then $\Delta(F_L + F_R)$ is positive until

$(F_L + F_R) = 0$, at which point the escape is successful. The escape can also be successful if $\eta < -m_c$. If $\eta > m_c$, then the agent can assume that it has been caught.

6.1.4 Chasing We explored two methods of controlling the chase. The method used with Ramona and the simulations has only one phase:

$$\Delta(F_L + F_R) = \begin{cases} -k\eta & \text{if } \eta < 0 \\ -\mu & \text{otherwise} \end{cases} \tag{13}$$

This states that if there is a contraction, approach; otherwise, slow down. The chase is ended when the watcher has "captured" the target ($\eta > m_c$), or the watcher has stopped or is moving backward and η is below threshold, in which case the watcher has lost the target.

6.1.5 Shadowing With Louie, chasing is composed of two stages. The first stage is to shadow the wanderer, or move so as to match its speed. Thus, an agent might also shadow a rival off its territory or follow its mother using this control law. The rule is simple: If the flow is inward, $\Delta(F_L + F_R)$ is positive and, if outward, negative; otherwise, zero. The control law for shadowing is:

$$\Delta(F_L + F_R) = -k\eta \tag{14}$$

6.1.6 Docking Once the agent has matched the speed of the target ($|\eta|$ is below threshold for a period of time), or the agent and target have come to a stop, then the agent can begin to dock. Though not stated explicitly, the latter type of sequence seems inherent in the male hoverfly, which "rapes" another fly (male or female) only after shadowing the target and the target lands on a flower or hovers for a period of time (Collett & Land, 1975). In other words, it attempts to attack only a motionless fly; thus, the shadowing will have brought it to a standstill.

The intent of docking though, is to make a "soft" contact with the target, as opposed to attack, in which the agent makes a "hard" contact with the target. If τ specifies the time to contact with the target, then its derivative, $\dot{\tau}$ (which is unitless), can be used to control deceleration prior to contact (Lee, 1976). Generally, the observer's deceleration is adequate if and only if $\dot{\tau} \geq -0.5$. Approaches with $\dot{\tau} < -0.5$ will yield hard collisions, whereas $\dot{\tau} > -0.5$ will yield soft contacts. Kim, Turvey, and Carello (1993) showed that human observers make this qualitative distinction between types of contact based solely on $\dot{\tau}$, and Yilmaz and Warren (1995) found that subjects used the "$\dot{\tau}$ strategy" to control braking during an approach to a "stop sign" in a closed-loop display.

In the docking mode though, the target might change its speed. Therefore, docking is not purely a matter of deceleration but also of acceleration. The control law is:

$$\Delta(F_L + F_R) = \begin{cases} -k(\eta - \mu) & \eta \leq 0 \\ k(\dot{\tau} + 0.5) & \eta > 0 \end{cases} \tag{15}$$

If the flow is inward ($\eta \leq 0$), then $\Delta(F_L + F_R)$ should be positive and Equation 14 should be used with a minimum amount of acceleration. When the flow is outward ($\eta > 0$), then the agent should decelerate. If the agent is attacking, then 0.5 should be changed to a higher value reflecting the most force that the agent can apply to the target without hurting itself.

If the target is stationary, then there is little chance that $\eta \leq 0$, and the agent can simply use the latter part of Equation 15. Again, for simplicity, we assume that at the same margin value that the agent uses to detect that it has been captured ($\tau < m_c$), it also will detect that it has captured the target. Once the escape or dock is complete, the agent stops and the process begins again—the agent waits for a target to come into its FOV.

6.2 Small robot implementation

With Louie, we were able to perform a thorough robo-ethological study, examining the utility of these control laws in real-world interactions with a human. Thus, the following sections deal only with this work, which has been briefly reported (Duchon et al., 1995).

6.2.1 Control law conversion As with obstacle avoidance, rotations were made every few frames. Interestingly, this same kind of behavior, termed *body saccades*, has been observed in some flies when pursuing a target (Collett & Land, 1975; Wagner, 1986b) but typically only when the target is outside a "foveal" region. Within the foveal region (approximately 8 degrees in radius), the fly is able to maintain smooth tracking and keep the target well fixated (Collett & Land, 1975). As will be described later, Equation 11 appears to have this same effect when a higher sampling density is used for the center of the FOV.

Translational speed of the robot could be controlled only in 1-cm/sec steps, so the control laws had to be adjusted accordingly. Only the sign of the immediacy values was used and approximated by averaging the direction of flow from a patch around the center of the FOV or for the entire area of the segmented target. So, for example, during escape, speed was decreased one unit (i.e., to move backward faster) if the average direction of flow was outward or near zero, and was increased one unit if the average direction was inward. Santos-Victor and Sandini (1994) have implemented a similar visually guided docking strategy.

The results reported here use only Equations 9 and 11 for fixation. Equation 10 was not implemented because the ability to segment and find the center of the target was so crude that the optic velocity of the target, $\dot{\beta}_t$, was not reliable.

6.2.2 Segmentation Segmentation was done using the optic flow alone. A complicated algorithm could not be used owing to real-time limitations and because the flow information was noisy. Working in the horizontal direction alone, the flow vector magnitudes in a column of the array were added, and the difference between this sum and that of the vectors in the adjacent column was found. This is similar to Nakayama and Loomis's (1974) "convexity" function in one dimension. The difference should be higher at an object boundary than within an object. Moreover, the difference should be especially high at the boundary of a moving object as the motion on the two sides of the boundary could be in completely opposite directions. The two columns with the highest difference values were taken as the boundaries of the target (in one dimension).

6.2.3 Evaluation We videotaped Louie playing tag with a number of different targets. Results are reported from those trials using a 4 $\times$ 12-inch, hand-held cardboard target covered with floral print fabric. This same fabric was stapled to posterboards to provide a textured background, instead of the usual AI lab walls and paraphernalia. Without pause, the experimenter would bring the target into the FOV while either approaching or retreating from the robot.

At all times, the experimenter tried to move the target at an appropriate rate given the limitations of the robot. The videotapes were analyzed later, and each interaction was classified as an escape or chase, and as either a true or a false escape or chase. A false chase occurred when the target was approaching the robot but the robot gave chase anyway. A false escape occurred when the target was moving away or laterally in front of the robot and the robot retreated. Real chases and escapes were further divided in terms of their difficulty, which was based on the overall length of the chase and the number of changes in direction made by the target. The more difficult chases were further categorized into hits and misses, and the hits were classified by the way the chase ended: *contact*, running into the target; *η-dock*, stopping when $\eta > m_c$; or *far*, stopping far from the target owing to a false $\eta > m_c$. The more difficult escapes were classified as: *short*, in which the robot stopped before the target stopped; *long*, in which the robot continued long past the time the target stopped moving; and *reasonable*, in which the robot stopped at a reasonable time after the target stopped.

Table 1 Results from chases with and without segmentation

Capture Type	With Segmentation	Without Segmentation
Contact	5	2
η-dock	**1**	**4**
Far	1	1

6.2.4 Results

Escapes Without segmentation, 63% of the approaches led to a false chase, probably due to an excessively high m_e. Of the 11 trials in which an escape was triggered, the robot stopped at a reasonable distance in 10 trials and stopped short in 1.

With segmentation, false chases were less common (55%), owing to the fact that it is strictly flow-registered in the segmented region that is used. Of nine escapes, five ended reasonably, three short, and one long.

Chases Relatively few false escapes were made when a target retreated from the robot. The results of the difficult real chases are shown in Table 1. With and without segmentation, the robot achieved an equal number of hits (seven of ten chases) but, with segmentation, most of these hits resulted from actual contact with the target, whereas without segmentation a majority of hits were successful docks using η.

6.2.5 Discussion

Performance with and without segmentation It is difficult to compare the performance of the control laws with and without segmentation. Many of the variables (e.g., the number of chases, or their difficulty) are based on the experimenter's behavior, not the robot's. However, to get some idea of whether a difference exists, we can look at the chases, whether they were successful, and the kind of contact that was made with the target.

The higher percentage of η-docks without segmentation probably is due to the fact that the target was directly in front of the robot at the time and, because it was looking only at the center of the FOV, this would contain correct information. When up close, the target covers most of the FOV; therefore, a segment can be found anywhere, and the flow around the center of that segment will be skewed, giving incorrect η information.

Other reasons can be given for the difference in performance between the two methods. First and foremost, the quality of the segmentation algorithm was poor. The robot would often follow "ghosts": That is, it would segment some part of the visual field other than the target or, if the target were only partially in the FOV, one side of the segment would come from the target and the other side from something else in the field. This type of behavior would momentarily lead the robot astray. An animal in a natural environment has many other properties (e.g., color, shape, size, type of internal motion) to help segment the object that would increase the reliability of these control laws using segmentation. However, the possibility also exists for "implicit" segmentation by means of filtering. That is, the control law without segmentation (Equation 11) could be used if only those flow signals from regions of the FOV containing certain other properties (e.g., color) were included. Preliminary work on Ramona using a color camera has shown that this can be successful.

One might ask the question of performance in the opposite way: Why did the robot do so well *without* segmentation? Much of its ability can no doubt be attributed to the simplicity of the environment. The target was the only object nearby, and the visible background usually was two or more times farther away than the target. This strategy could be effective only in environments in which the prey make up a majority of all objects present, or in which the prey are very few and the medium is relatively empty, as was the case here. Here, too, other forms of information, such as auditory localization and odor differentials, could bias the amount or direction of rotation. If the agent is capable of segmenting (by any means), then only the probability of the prey running behind an object is important. From an ecological perspective, in this case there is information about the accretion and deletion of texture to inform the agent of where the target is behind the object. The ability to find and use this information would be very helpful in future ecological robotics research.

Control law interaction The main problem associated with using control laws is knowing which one(s) to use at any given time. The use of margin values can demarcate the transitions between action modes such as docking and stopping or between escaping and giving up. Chasing or escaping is triggered naturally by the direction of flow of the target. However, a biological agent obviously would not want to chase everything going away, nor would it want to escape from every approaching surface; the other affordances of the target (e.g., mate-ability) would help determine the changes in the action mode.

There is a problem during escaping, for example, in which the law of escape must be balanced with avoiding objects. This problem was avoided here by having the robot face the target and assuming an open medium behind the robot. While chasing, however, the agent could follow the target alone without consideration of the environment, because where the target goes, so can the agent (usually). This is a problem of "reflex interaction" and has been noted in some types of flies, for which the background flow does influence the ability to fixate the target (Collett & Land, 1975). It is possible that these control laws alternate [(Srinivasan & Bernard,

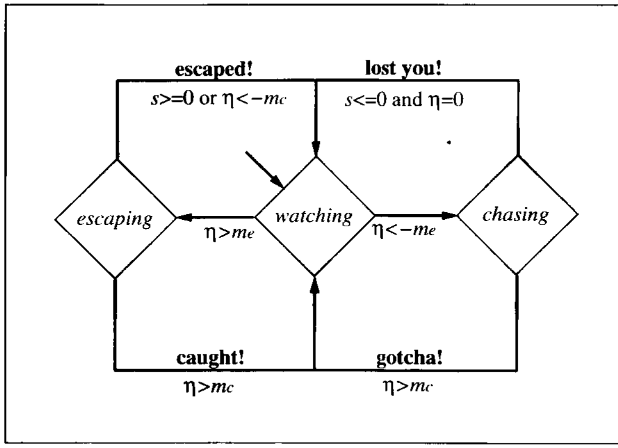


Figure 8
The action modes and transitions for tag on Ramona and in the simulations. Fixating takes place continuously throughout the game. The transition conditions and control laws for each mode are described in the text.

1977); a subsumption architecture would support this kind of behavior] or all are acting simultaneously with varying thresholds, constants (Virsik & Reichardt, 1976; Reichardt, 1986), time scales (Wagner, 1986b), dynamics (Schöner, Dose, & Engels, 1995), or schemas (Arkin, 1990). We have begun looking into these issues in the realm of maze navigation (Duchon, 1996).

6.3 Large robot implementation

After working with the simulations in section 6.4, we used a slightly different set of control laws with Ramona, which were implemented in a more continuous manner such that changes in speed (Δs) could be controlled directly by η (Fig. 8). Another change was in the use of a single control law for chasing (Equation 13), causing the watcher to increase its speed until it catches up with the wanderer, but to slow down only gradually. If the watcher tracks the wanderer well enough, it will make an η-dock; however, if it loses the agent, it eventually will slow down to a stop.

With a wide-angle lens, the camera near chest height, and speeds up to 30 cm/sec, it is possible to play tag directly with Ramona in a "natural" manner. In the atrium, chases up to 20 feet have been recorded, but a systematic study has not been made. When shadowing is implemented as a separate action mode, it is possible to interact with Ramona by pulling an extended hand toward one's chest to indicate "come to me" and by putting one's hand out toward the robot to indicate "stop."

6.4 Simulations

Simulations of tag were carried out in two arenas, one open (Fig. 9) and one with some walls (Fig. 10). In the open arena, tag was played in a manner very similar to that in the robot studies: The wanderer used the balance strategy to wander around

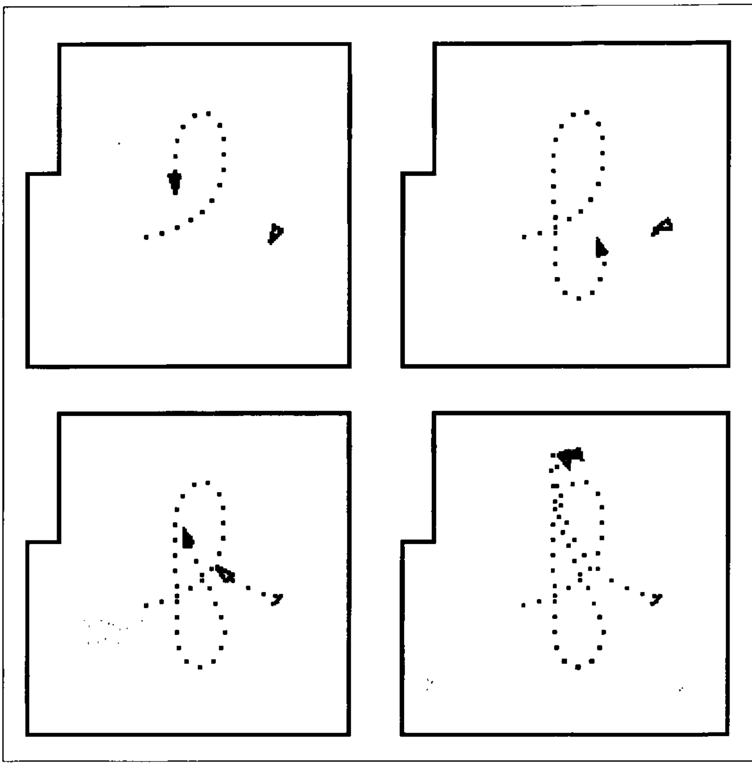


Figure 9 A simulated game of tag. The open triangle is the watcher; the filled triangle, the wanderer. (Top left) The wanderer blissfully moves around out of sight of the watcher. (Top right) The watcher fixates the wanderer, detects an expansion, and so escapes momentarily. (Bottom left) The wanderer moves away and the watcher gives chase. (Bottom right) The watcher captures the wanderer as the latter avoids a wall.

the arena, treating the watcher as simply another obstacle. If the wanderer came into the FOV of the watcher, the watcher fixated it. If the wanderer approached the watcher (by chance), the watcher backed up until the wanderer receded. If the wanderer moved away fast enough, the watcher would chase it down until $\eta > m_c$. At this point, the watcher and wanderer agents would switch roles.

6.4.1 Fovea In these simulations, the agents were given a higher-resolution "fovea." Sixteen samples were distributed uniformly in a 1-degree area around the center of the FOV. Thirty-two more samples were distributed uniformly throughout the entire FOV.

Figure 9 shows a few frames from a chase sequence without segmentation. Both agents were placed randomly in the arena, and one agent began as the watcher. As

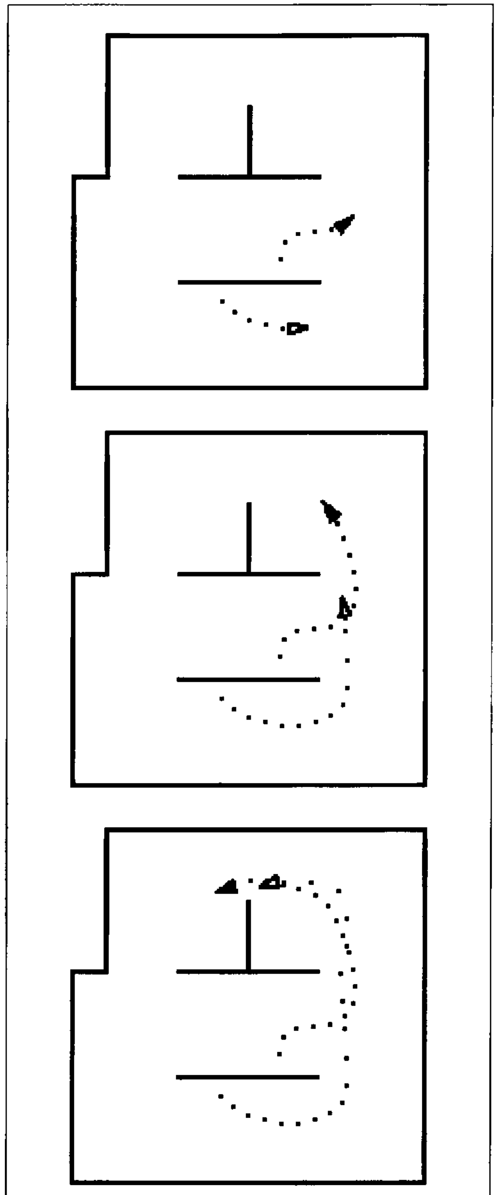


Figure 10 A simulated chase with segmentation. (Top) The two agents randomly start on opposites of the lower wall. (Middle) When the open agent rounds the corner and sees the other moving away, it begins chasing it. (Bottom) The chaser loses the target momentarily in the upper right, but finds a segment again and chases the other around the middle wall to make a hit a few frames later.

the wanderer came into the FOV, the fixation law caused the watcher to turn slowly toward the wanderer. Given the oversampling in the fovea, once the target was seen there, the watcher never lost track of it. Once the watcher started moving, however, an additional precaution was taken, which was to narrow the entire FOV from 180 to 90 degrees (analogous to attention). This ensured that Equation 11 did not cause the chaser to be drawn to a wall; because the target was avoiding walls as well, this worked perfectly in the open arena.

6.4.2 Problems with η As with obstacle avoidance, η was found by averaging the flow values from the fovea. This works fine if the wanderer is moving directly away. However, if the wanderer moves at an angle to the watcher's heading—say, to the right—then the right side of the wanderer will be closer than the left, causing more outward flow to the right of the heading than inward flow to the left of the heading. Thus, the watcher will see the target expand, overestimate η, and slow down. If the wanderer continues to move in an outward spiral like this, the watcher never will reach it during the docking phase. Equation 13 is not susceptible to this "illusion." However, if the near side of the wanderer is moving toward the watcher (as happens when the wanderer turns away from a wall), the opposite happens: The watcher underestimates η and "pounces" on the target.

A number of interesting conclusions can be drawn from these results. One such conclusion is that animals may also have this problem and that prey animals have developed zig-zagging as an effective escape mechanism to counteract this problem (Tinbergen, 1951): Zig-zagging not only allows animals to see their attacker (as they do not have a 360-degree FOV), but it confuses the predator by making the predator think that it is doing better than it is. Another conclusion is that, with animals that move on a ground plane, information from vertical expansion of the target may be more reliable for determining time-to-contact than information from horizontal expansion. However, this dimension was not simulated here.

6.4.3 Segmentation and obstacle avoidance Our final tag simulations were done to address the issue of control law interaction. The balance strategy gave a continuous signal of how the agent should turn to avoid obstacles. On top of this signal was added a bias from the segmentation algorithm (similar to the one used for the robot, but with the difference divided by the sum of the two columns and thresholded). The bias (γ), a slow-moving average of the angle of the center of the segment, was added to the angle of the turn determined by the balance strategy (r_{bal}); thus

$$r_{\text{out}} = r_{\text{bal}} + \gamma. \tag{16}$$

Moreover, if a target was segmented, then that portion of the flow was eliminated from the input to the balance strategy. The η value of the segmented section was used, as it was with the robot.

Both agents wandered around using the balance strategy. When an agent segmented a part of its view, it determined whether the target was approaching or retreating and escaped or chased appropriately. By zeroing out the flow from the segment, the chaser could avoid walls but approach the target. Figure 10 shows a successful pursuit. This method of simply adding the two output signals together worked successfully, because any signal turns the agent made toward an obstacle would be compensated in the next moment by a higher obstacle avoidance signal.

Conclusion

We have discussed behaviors such as obstacle avoidance and the game of tag, which can be produced in a robot with no reconstruction of the visual scene (Aloimonos, 1993). In addition, we have recently begun exploring the use of these behaviors as a basis for navigating a maze (Duchon, 1996). At a minimum, this work points to an approach eschewing a central model in favor of a tighter binding between action and perception. This methodology has been explored by a number of robotics researchers (Brooks, 1991a; Aloimonos, 1992; Horswill, 1993; Pfeifer & Verschure, 1993; Sandini et al., 1993; Coombs et al., 1995) and has even produced higher-order behaviors such as planning (Matarić, 1992; Meeden, 1994). The similarities between these approaches and ours, based independently on Gibsonian ideas, suggests that the application of the theories and results from 50 years of ecological psychology will surely enhance this endeavor.

Our work also ties in with recent physiological studies and lesion cases (Milner & Goodale, 1995) that suggest separate "what" and "how" pathways in the brain. The lesion cases have indicated a difference between knowing what an object is and knowing how to maneuver it. This change of emphasis is reflected also in some philosophical approaches to knowledge, whereby to "know that" first requires one to "know how" (Ryle, 1949; Bechtel, 1990). Our robot does not need to know what an object is in order to avoid it, nor does it need to identify a target before knowing how to control its escape. In essence, we have implemented a simple "how" pathway. Nonetheless, because an approaching conspecific may afford mating as well as escape, it is important that all the affordances of an object be recognized and one of them chosen to be acted on. Neural networks would be an ideal means of satisfying the many soft constraints (affordances) of an object and of choosing a single output (action mode). In any case, knowledge of the affordances of an environment provides a basis for a choice of action, and that action, once chosen, can be controlled without a central model of the world. Such procedural, functional knowledge seems necessarily prior to more abstract, declarative knowledge.

Finally, the fact that these control laws are essentially universal for mobile agents with perceptual systems capable of detecting optic flow means that these laws can be investigated in insects, animals, humans, and robots. We are beginning an interactive approach with the last two kinds of agents. For example, we have recently reported a study (Duchon and Warren, 1997) that demonstrates that under some circumstances, even humans will use the balance strategy. Robotic modeling helps us to determine the plausibility of control laws that have been hypothesized for biological agents and, from psychophysical studies, we hope to find new control laws that are useful in a robot. The study of control laws based on optic flow thus provides a unique opportunity for cognitive scientists, computer scientists, and engineers to work together on solving the same problems.

Acknowledgments

Tom Dean kindly provided the small robot and the cameras. Ted Camus' original work on real-time optical flow made this entire project possible. Bill Smart and Chris Cantor were very helpful with the programming and implementation. Thanks also go to Jak Kirman, Moises Lejter, Jon Monsarrat, Dan Gillmor, and Jason Lango for keeping the robots and computers working. This work was supported by a National Science Foundation Graduate Research Fellowship (A.P.D.) and, in part, by NSF grants IRI-9453383 and IRI-9312395 (L.P.K.).

References

Aloimonos, Y. (1992). Is visual reconstruction necessary? Obstacle avoidance without passive ranging. *Journal of Robotic Systems, 9*(6), 843–858.

Aloimonos, Y. (1993). Introduction: Active vision revisited. In Y. Aloimonos (Ed.), *Active perception.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Aloimonos, Y., & Rosenfeld, A. (1991). Computer vision. *Science, 253*, 1249–1254.

Arkin, R. C. (1990). The impact of cybernetics on the design of a mobile robot: a case study. *IEEE Transactions of Systems, Man, and Cybernetics, 20*(6), 1245–1257.

Bajcsy, R. (1988). Active perception. *Proceedings of the IEEE, 76*(8), 996–1005.

Ballard, D. (1991). Animate vision. *Artificial Intelligence, 48*, 57–86.

Ballard, D., & Brown, C. (1992). Principles of animate vision. *CVGIP: Image Understanding, 56*(1), 3–21.

Bechtel, W. (1990). Multiple levels of inquiry in cognitive science. *Psychological Review, 52*, 271–281.

Brooks, R. A. (1991a). Intelligence without representation. *Artificial Intelligence, 47*, 139–160.

Brooks, R. A. (1991b). New approaches to robotics. *Science, 253*, 1227–1232.

Camus, T. (1994). *Real-time optical flow.* Unpublished doctoral dissertation, Brown University, Providence, RI.

Cliff, D. (1991). The computational hoverfly: A study in computational neuroethology. In J.-A. Meyer & S. W. Wilson (Eds.), *From animals to animats: First international conference on simulation of adaptive behavior.* Cambridge, MA: MIT Press.

Cliff, D. (1992). Neural networks for visual tracking in an artificial fly. In F. J. Varela & P. Bourgine (Eds.), *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Intelligence.* Cambridge, MA: MIT Press.

Collett, T. S. (1980a). Angular traacking and the optomotor response: An analysis of visual reflex interaction in a hoverfly. *Journal of Comparative Physiology, 140,* 145–158.

Collett, T. S. (1980b). Some operating rules for the optomotor system of a hoverfly during voluntary flight. *Journal of Comparative Physiology, 138,* 271–282.

Collett, T. S., & Land, M. F. (1975). Visual control on flight behavior in the hoverfly, *Syritta pipiens* L. *Journal of Comparative Physiology, 99,* 1–66.

Coombs, D., Herman, M., Hong, T., & Nashman, M. (1995). Real-time obstacle avoidance using central flow divergence and peripheral flow. In *Proceedings of the Fifth International Conference on Computer Vision.*

Duchon, A. P. (1996). Maze navigation using optical flow. In P. Maes, M. J. Matarić, J.-A. Meyer, J. Pollack, & S. W. Wilson (Eds.), *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior.* Cambridge, MA: MIT Press.

Duchon, A. P., & Warren, W. H. (1994). Robot navigation from a Gibsonian viewpoint. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.* Piscataway, NJ: IEEE.

Duchon, A. P., & Warren, W. H. (1997). Poster presented at the International Conference on Vision and Action, York University, Toronto, Canada, June 24–28, 1997.

Duchon, A. P., Warren, W. H., & Kaelbling, L. P. (1995). Ecological robotics: Controlling behavior with optical flow. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates.

Effken, J. A., & Shaw, R. E. (1992). Ecological perspectives on the new artificial intelligence. *Ecological Psychology, 4*(4), 247–270.

Franceschini, N., Pichon, J. M., & Blanes, C. (1992). From insect vision to robot vision. *Philosophical Transactions of the Royal Society of London, Series B, 337,* 283–294.

Gibson, J. J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology, 49*(3), 182–194.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Gibson, J. J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology, 68*, 372–385.

Götz, K. G., Hengstenberg, B., & Biesinger, R. (1979). Optomotor control of wing beat and body posture in *Drosophila*. *Biological Cybernetics, 35*, 101–112.

Horswill, I. D. (1993). *Specialization of perceptual processes.* Unpublished doctoral dissertation, MIT, Cambridge, MA.

Kim, N., Turvey, M. T., & Carello, C. (1993). Optical information about the severity of upcoming contacts. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 179–193.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception, 5*, 437–459.

Lee, D. N. (1980). The optic flow field: The foundation of vision. *Philosophical Transactions of the Royal Society of London, Series B, 290*, 169–179.

Lee, D. N., & Lishman, R. (1977). Visual control of locomotion. *Scandinavian Journal of Psychology, 18*, 224–230.

Lorenz, K. Z. (1980/1952). *King Solomon's ring.* Chicago: Time-Life Books.

Matarić, M. (1992). Integration of representation into goal-driven behavior-based robots. *IEEE Transactions on Robotics and Automation, 8*(3), 304–312.

Meeden, L. A. (1994). *Towards planning: Incremental investigations into adaptive robot control.* Unpublished doctoral dissertation, Indiana University, Bloomington, IN.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* Oxford: Oxford University Press.

Moravec, H. P. (1981). Rover visual obstacle avoidance. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence.* Los Altos, CA: Morgan Kaufmann.

Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity-sensitive neurons and space perception: A hypothesis. *Perception, 3*, 63–80.

Nelson, R. C., & Aloimonos, Y. (1989). Obstacle avoidance using flow field divergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(10), 1102–1106.

Nilsson, N. J. (1984). *Shakey the robot* (Tech. Rep. No. 325). Menlo Park, CA: SRI International.

Pfeifer, R., & Verschure, P. (1993). Designing efficiently navigating non-goal-directed robots. In *Proceedings of the Second International Conference on the Simulation of Adaptive Behavior.* Cambridge, MA: MIT Press.

Pickering, J. (1992). The new artificial intelligence and biological plausibility. In S. S. Valenti & J. B. Pittenger (Eds.), *Studies in perception and action: II. Posters presented at the Eighth International Conference on Event Perception and Action.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Prokopowicz, P. N., Swain, M. J., & Kahn, R. E. (1994). Task and environment-sensitive tracking. In *Proceedings of the IARP/IEEE Workshop on Visual Behaviors.* Piscataway, NJ: IEEE.

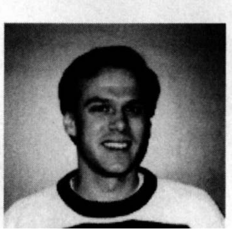
Reichardt, W. (1986). Processing of optical information by the visual system of the fly. *Vision Research, 26*(1), 113–126.

Reichardt, W., & Poggio, T. (1976). Visual control of orientation behavior in the

fly. *Quarterly Reviews of Biophysics, 9*(3), 311–375.
Royden, C. S., Banks, M. S., & Crowell, J. A. (1992). The perception of heading during eye movements. *Nature, 360*, 583–585.
Ryle, G. (1949). *The concept of mind.* New York: Barnes & Noble.
Sandini, G., Santos-Victor, J., Curotto, F., & Garibaldi, S. (1993). Robotic bees. In *Proceedings of IROS '93.* New York: IEEE.
Santos-Victor, J., & Sandini, F. (1994). *Visual behaviors for docking* (Tech. Rep. No. TR 2/94). Genoa, Italy: Laboratory for Integrated Advanced Robotics (LIRA)-Lab, Department of Communication, Computer, and Systems Science (DIST) University of Genoa.
Schöner, G., Dose, M., & Engels, C. (1995). Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems, 16*, 213–245.
Smithers, T. (1995). Are autonomous agents information processing systems? In L. Steels & R. Brooks (Eds.), *The artificial life route to artificial intelligence.* Hillsdale, NJ: Lawrence Erlbaum Associates.
Sobey, P. (1994). Active navigation with a monocular robot. *Biological Cybernetics, 71*, 433–440.
Srinivasan, M. V. (1977). A visually evoked roll-response in the housefly. *Journal of Comparative Physiology, 119*, 1–14.
Srinivasan, M. V., & Bernard, G. D. (1977). The pursuit response of the hoverfly and its interaction with the optomotor response. *Journal of Comparative Physiology, 115*, 101–117.
Srinivasan, M. V., Lehrer, M., Kirchner, W. H., & Zhang, S. W. (1991). Range perception through apparent speed in freely flying honeybees. *Visual Neuroscience, 6*, 519–535.
Steels, L. (1995). Building agents out of autonomous behavior systems. In L. Steels & R. Brooks (Eds.), *The artificial life route to artificial intelligence.* Hillsdale, NJ: Lawrence Erlbaum Associates.
Tinbergen, N. (1951). *The study of instinct.* Oxford: Clarendon Press.
Tresilian, J. R. (1991). Empirical and theoretical issues in the perception of time to contact. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 865–876.
Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981). *Cognition, 9*, 237–304.
Virsik, R. P., & Reichardt, W. (1976). Detection and tracking of moving objects by the fly *Musca domestica. Biological Cybernetics, 23*, 83–98.
Wagner, H. (1986a). Flight performance and visual control of flight of the free-flying housefly (*Musca domestica* L.): II. Pursuit of targets. *Philosophical Transactions of the Royal Society of London, Series B, 312*, 553–579.
Wagner, H. (1986b). Flight performance and visual control of flight of the free-flying housefly (*Musca domestica* L.): III. Interactions between angular movement induced by wide- and small-field stimuli. *Philosophical Transactions of the Royal Society of London, Series B, 312*, 581–595.

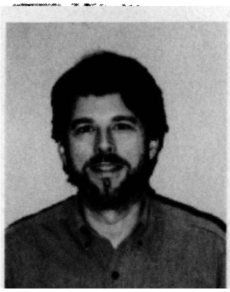
Warren, W. H. (1988). Action modes and laws of control for the visual guidance of action. In O. G. Meijer & H. Roth (Eds.), *Complex movement behaviour: The motor-action controversy.* Amsterdam: Elsevier Science/North-Holland.
Warren, W. H., & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America, Series A,* 7(1), 160–169.
Warren, W. H., Kay, B. A., & Yilmaz, E. H. (1996). Visual control of posture during walking: Functional specificity. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 818–838.
Weber, K., Venkatesh, S., & Srinivasan, M. V. (1997). Insect inspired behaviours for the autonomous control of mobile robots. In M. V. Srinivasan & S. Venkatesh (Eds.), *From Living Eyes to Seeing Machines.* Oxford: Oxford University Press.
Yilmaz, E., & Warren, W. H. (1995). Visual control of braking: A test of the $\dot{\tau}$ hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 21*(5), 996–1014.

About the Authors



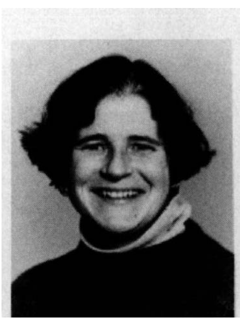
Andrew P. Duchon

Andrew P. Duchon has a B.A. in psychology and the integrated science program from Northwestern University. He is pursuing a Ph.D. in the Department of Cognitive and Linguistic Sciences at Brown University. He has conducted research in the psychophysics of motion perception at Northwestern University and neural network models of speech recognition at the Human Information Processing Labs at Advanced Telecommunications Research in Kyoto, Japan. Mr. Duchon is interested in navigation from on-line visual control to cognitive maps, combining research from animals, humans, and robots. He is a member of the International Society of Adaptive Behavior and has received a National Science Foundation Graduate Research Fellowship.



William H. Warren

William H. Warren is a professor of cognitive and linguistic sciences at Brown University. He received a B.A. in psychology, biology, and philosophy from Hampshire College in 1976 and a Ph.D. in experimental psychology from the University of Connecticut in 1982. Dr. Warren's research interests include the visual perception of optical flow, the visual control of locomotion, and the nonlinear dynamics of gait and posture. He is a member of the Psychonomic Society, the American Psychological Society, and the International Society for Ecological Psychology. He recently received a National Science Foundation Research Scientist Development Award.



Leslie Pack Kaelbling

Leslie Pack Kaelbling is an associate professor of computer science at Brown University. She previously held positions at the Artificial Intelligence Center of SRI International and at Teleos Research. She received a B.A. in philosophy in 1983 and a Ph.D. in computer science in 1990, both from Stanford University. Dr. Kaelbling has done substantial research on programming paradigms and languages for embedded systems, on mobile robot design and implementation, and on reinforcement learning algorithms. Her current research directions include integrating learning modules into systems programmed by humans, algorithms for learning and navigating using hierarchical domain representations, and methods for learning perceptual strategies. She is a National Science Foundation Presidential Faculty Fellow, a member of the AAAI Executive Council, a member of the International Joint Conference on Artificial Intelligence (IJCAI) Advisory Committee, and the 1997 recipient of the IJCAI Computers and Thought Award.