Modeling Embodied Visual Behaviors

NATHAN SPRAGUE
Kalamazoo College
and
DANA BALLARD and AL ROBINSON
University of Rochester

To make progess in understanding the operations of the human brain, we will need to understand its basic functions at an abstract level. One way to achieve such an understanding is to create a model of a human that has a sufficient amount of complexity so as to be capable of interpreting abstract behavioral models. Recent technological advances have been made that allow progress to be made in this direction. Virtual reality(VR) graphics models that simulate extensive human capabilities can be used as platforms from which to develop synthetic models of visuo-motor behavior. Currently such models can capture only a small portion of a full behavioral repertoire, but for the behaviors that they do model, they can describe complete visuo-motor subsystems at a useful level of detail. The value in doing so is that the body's elaborate visuo-motor structures greatly constrain and simplify the specification of the abstract behaviors that guide them. The result is that, essentially, one is left with proposing an embodied "operating system" model for picking the right set of abstract behaviors at each instant. This paper outlines one such model. A centerpiece of the model uses vision to aid the behavior that has the most to gain from taking environmental measurements. Preliminary tests of the model against human performance in realistic VR environments show that the main features of the model show up in human behavior.

Categories and Subject Descriptors: I.2.10 [**Vision and Scene Understanding**]: Perceptual reasoning

1. INTRODUCTION

All brain operations are situated in the body [Clark 1997]. Even when the operations are purely mental, they reflect a developmental path whereby symbols are grounded in concrete interactions in the world. The genesis of this view is attributed to the philosopher Merleau-Ponty [Merleau-Ponty 1962], but more recently it undergone extensive development[O'Regan and Noe 2001; Noe 2005; Clark 1999; Roy and Pentland 2002; Yu and Ballard 2004; Ballard et al. 1997]. These authors argue that not only is the body a source of mechanisms for grounding the experiences that we describe symbolically in language, but in fact is a *sine qua non*. We may have music in sheet form, but we can only experience it with a body to play it or listen to it. The same goes for the brain's abstract symbols.

There is a great boon when embodiment is taken as a tenet of research programs: once the body is modeled, tremendous computational economies result. However there is no free lunch. The reason that embodied models can forgo computation is that it is done implicitly by the body itself. Thus modeling embodiment places an enormous demand on the researcher to simulate the body's prodigious computational abilities. To this end, research programs that focus on embodiment have been facilitated by the development of virtual reality (VR) graphics environments.

These VR environments can now run in real time on standard computing platforms. The value of VR environments is that they allow the creation of virtual agents that implement complete visuo-motor control loops. Visual input can be captured from the rendered virtual scene, and motor commands can be used to direct the graphical representation of the virtual agent's body[Terzopoulos and Rabie 1997; Faloutsos et al. 2001; Sprague and Ballard 2003b]. Embodied control has been studied for many years in the robotics domain, but virtual agents have major advantages over physical robots in the areas of experimental reproducibility, hardware requirements, flexibility, and ease of programming.

Embodied models can now be tested using new instrumentation. Linking mental processing to visually-guided body movements at a millisecond timescale would have been impractical just a decade ago, but recently a wealth of high resolution monitoring equipment has been developed for tracking body movements in the course of everyday behavior, particularly head, hand and eye movements (e.g. [Pelz et al. 2001; Babcock and Pelz 2000]). This allows for research into everyday tasks that typically have relatively elementary cognitive demands but require elaborate and comprehensive physical monitoring[Hayhoe et al. 2003; Triesch et al. 2003; Ballard et al. 1995]. In these tasks, overt body signals provide a direct indication of mental processing.

During the course of carrying out tasks, humans engage in a wide variety of subtasks, each of which requires certain perceptual and motor resources. Thus there must be mechanisms that allocate resources to subtasks. Understanding this resource allocation requires an understanding of the ongoing demands of behavior, as well as the nature of the resources available to the human sensori-motor system. The interaction of these factors is complex, and that is where the virtual human platform can be of value. It allows us to imbue our artificial human with a particular set of resource constraints. We may then design a control architecture that allocates those resources in response to task demands. The result is a model of human behavior in temporally extended tasks that may be tested against human performance.

We refer to our own virtual human model as 'Walter.' Walter has physical extent and programmable kinematic degrees of freedom that closely mimic those of real humans. His graphical representation and kinematics are provided by the DI-guy package developed by Boston Dynamics. The crux of the model is a control architecture for managing the extraction of information from visual input that is in turn mapped onto a library of motor commands. The model is illustrated on a simple sidewalk navigation task that requires the virtual human to walk down a sidewalk and cross a street while avoiding obstacles and collecting litter. The movie frame in Figure 1 shows Walter in the act of negotiating the sidewalk which is strewn with obstacles (blue objects) and litter (purple objects) on the way to crossing a street.

The goal of this paper is to describe both the structure and usefulness of the Walter model and thus it is divided into two main parts. The first part, sections 2-8, describes Walter's control architecture and resource allocation mechanisms in detail. The main result is to show how learned behaviors, when referenced to the body, can be easily composed and that such a composition leads directly to a

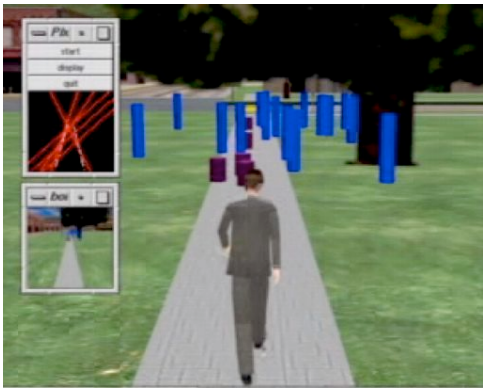

Fig. 1. The Walter simulation. The main panel shows a single video frame from the real-time simulation that has Walter negotiating a sidewalk strewn with litter and obstacles. The insets show the use of vision to guide the humanoid through a complex environment. The upper inset shows the particular visual routine that is running at any instant. The lower insert shows the visual field in a head-centered frame.

novel interpretation of gaze allocation. The second part of the paper, section 9, presents eye tracking data collected from a human subject engaged in the same sidewalk navigation task, and compares this to the output of the virtual human model. The main result is to show that the main features of the model are present in human data. Overall, the paper describes a way of modeling natural behavior over extended time periods and suggests new roles for gaze in the management of a human's task agenda.

2. THE ROLE OF EMBODIMENT

Walter's vision system uses fixations that can be changed very three hundred milliseconds. Thus he approximates human vision which uses fixations that have an average duration of 200 to 300 milliseconds e.g. [Aivar et al. 2005]. Walter can only work on one task at each gaze point. This is obviously a simplification, but may not be too much so. Several researchers have shown e.g. [Roelfsema et al. 2003; Fabre-Thorpe et al. 2001] the time to carry out a basic task is about 200 milliseconds. Humans can process items at the rate of 25 ms per item when searching for a target, but only when they can process items in parallel. As [Palmer 1995; Eckstein et al. 2000; Zelinsky 1996] have separately shown, the exact number that can be so processed depends on signal to noise conditions.

The fixational system brings home the key role of embodiment in behavior. Although the phenomenological experience of vision may be of a seamless three-dimensional surround, the system that creates that experience is discrete. Furthermore as humans are binocular, they spend most of their time fixating objects in the near distance. That is, the centers of gaze of each of the eyes meet at a point in three dimensional space and, to a first approximation, rest on that point for three hundred milliseconds. This ability to dynamically fixate has at least three important consequences:

(1) *Visual Routines* Visual computations have to be efficient since their result has to be computable in 300 milliseconds. Thus for the most part, the computations are task-dependent tests since those can use prior information to simplify the computations[Ullman 1985; Roelfsema et al. 2000].

(2) *Dynamic reference frames* Huge debates have ranged over the coordinate system used in vision- is it head based, eye based or otherwise? The fixational system shows that it must be dynamic. Depending on the task at hand, it can be any of these. Imagine using a screwdriver to drive a screw into hardwood. The natural reference frame for the task is the screw head where a pure torque is required. Thus the gaze is needed there *and* all the muscles in the body are constrained by this purpose: to provide a torque at a site remote to the body. To keep this idea in mind, consider how the visual system codes for disparity. Neurons that have disparity-sensitive receptive fields code for zero, negative and positive disparities. However zero disparity is at the fixation point, a point not in the body at all[Ballard et al. 1997].

(3) *Simplified Computation* The visual system has six separate systems to stabilize gaze. This of course is an indication of just how important it is to achieve gaze stabilization but the third consequence is that the visual computations can be simplified as they do not have the burden of dealing with gaze instability. Despite the fact that the human may be moving and the object of interest may be independently moving as well, the algorithms used to analyze that object can assume that it remains in a fixed position near the fovea.

The summary description of these advantages has vision as a discrete system that samples three dimensional space every 300 milliseconds performing purposeful computations. This description can also serve for the motor system.

The motor system is comprised of an extensive musculo-skeletal system that consists of over two hundred bones and six hundred muscles. One of the most important properties of this system is the passive energy that can be stored in muscles. This spring-like system has at least two important properties: 1) it can lead to very economical operation in locomotion, and 2) it can be used in passive compliant collision strategies that save the system from damage. Moreover it can be driven by a discrete strategy whereby set points for the spring-muscle system are communicated at a low bandwidth. The consequence of this organization is that the musculo-skeletal system confers on the motor system advantages similar to that of the fixational system in vision:

(1) *Motor Routines* Motor routines can be efficient for the same reason that visual routines are. Since they are also goal directed, they can make extensive use of expectations that can be quickly tested [Newell et al. 2001].

(2) *Dynamic reference frames* Like visual routines, motor routines use dynamic frames of reference. The screw driving example used for vision also applies for motor control. The multitude of muscles work together to apply a pure torque at the driver end. Consider also that for upright balance the motor system must make sure the center of gravity(CG) is over the base of the stance. But of course the CG is a dynamic point that moves with postural changes. Nonetheless the control system must refer its computations to this point.

(3) *Simplified Computation* There are many examples that could be mentioned to illustrate how the computation done by the motor system makes the job of motor routines easier, but one of the most obvious is the extensive use of passive compliance in grasping. If the motor system was forced to rely on feedback in the way that standard robot systems do then the grasping strategies would have to be far more delicate. Instead, grasp planning can be far easier as the passive conformation of the multi-fingered hand provides great tolerances in successful grasps.

3. BEHAVIOR BASED CONTROL

The body allows for simplified representations of behaviors but how are these behaviors organized? Any system that must operate in a complex and changing environment must be compositional[Newell 1990], that is it has to have elemental pieces that can be composed to create its more complex structures. Figure 2 illustrates two broad compositional approaches that have been pursued in theories of cognition, as well as in robotics. The first decomposition works on the assumption that the agent has a central repository of symbolic knowledge. The purpose of perception is to translate sensory information into symbolic form. Actions are selected that result in symbolic transformations that bring the agent closer to goal states. This sense-plan-act approach is typified in the robotics community by early work on Shakey the robot [Nilsson 1984], and in the cognitive science community by the theories of David Marr [Marr 1982]. In principle, the symbolic planning approach is very attractive, since it suggests that sensation, cognition and action can be studied independently, but in practice each step of the process turns out to be difficult to characterize in isolation. It is hard to convert sensory information into general purpose symbolic knowledge, it is hard to use symbolic knowledge to plan sequences of actions, and it is hard to maintain a consistent and up to date knowledge base.

The difficulties with the symbolic planning approach led Brooks [Brooks 1986] to suggest a radically different decomposition, illustrated in Figure 2B. Brooks' approach is to attempt to describe whole visuo-motor behaviors that have very specific goals. This behavior-based control involves a different approach to composition than planning-based architectures: simple behaviors are sequenced and combined to solve arbitrarily complex problems. The best approach to attaining this sort of behavioral composition is an active area of research. Brooks' own *subsumption* architecture organizes behaviors into fixed hierarchies, where higher level behaviors influenced lower level behaviors by over-writing their inputs.

Subsumption works spectacularly well for trophic, low-level tasks, but generally fails to scale to handle more complex problems [Hartley and Pipitone 1991]. For that reason we have chosen a more flexible control architecture. that follows more recent work on behavior based control (e.g. [Firby et al. 1995; Bryson and Stein 2001]). Our architecture allows the agent to address changing goals and environmental conditions by dynamically activating a small set of appropriate behaviors that we term *microbehaviors*.

A microbehavior is a complete sensory/motor routine that incorporates mechanisms for measuring the environment and acting on it to achieve specific goals. For

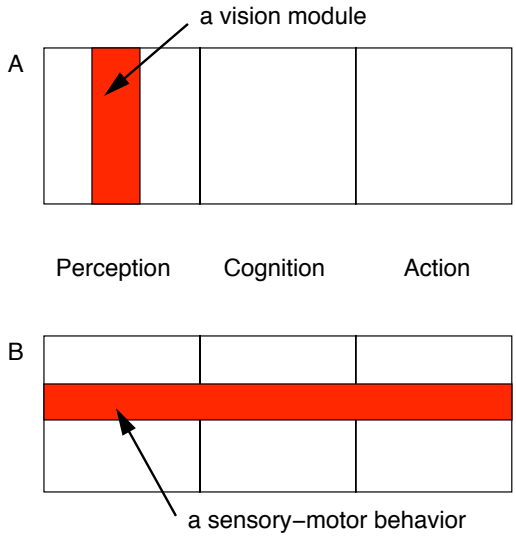

Fig. 2. Two approaches to behavioral research contrasted. A) In the Marr paradigm individual components of vision are understood as units. B) In the Brooks paradigm the primitive unit is an entire behavior.

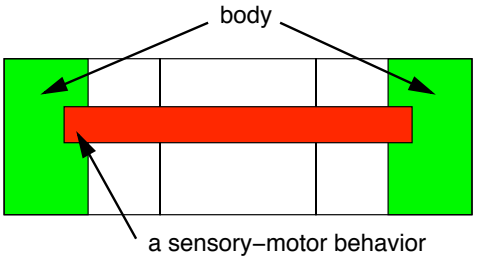

Fig. 3. Walter's behavioral primitives can be abstract as their primitives can be interpreted by his body's sensory motor system.

example a collision avoidance microbehavior would have the goal of steering the agent to avoid collisions with objects in the environment. A microbehavior has the property that it cannot be usefully split into smaller subunits. As suggested by Figure 3, the specification of microbehaviors is greatly simplified by embodiment. Microbehaviors can assume access to the bodies fixation system for input, and can generate low bandwidth output that is interpreted by the motor system.

4. THE HUMAN OPERATING SYSTEM MODEL

We think of the control structure needed to implement microbehaviors in terms that are similar to that of an operating system as three of Walter's abstractions have analogous roles. Firstly the behaviors themselves, when they are running, each have distinct jobs to do. Each one interrogates the sensorium with the objective

of computing the current state of the process. Secondly, once the state of each process is computed then the action recommended by that process is available. Such actions typically involve the use of the body. Thus an intermediate task is the mapping of those action recommendations onto the body's resources. Thirdly, the behavioral composition of the microbehavior set itself must be chosen. We contend that, similar to multiprocessing limitations on silicon computers, that the brain has a multiprocessing constraint that allows only a few microbehaviors to be simultaneously active. This constraint, we believe, is related to that for working memory. Addressing the issues associated with this vantage point leads directly to the three-level abstract computational hierarchy shown in Table 1. The *behavior level* of the hierarchy addresses the issues in running a microbehavior. These are each engaged in maintaining relevant state information and generating appropriate control signals. Microbehaviors' policies are represented as state/action tables, so the main issue is that of computing state information needed to index the table. The *arbitration level* addresses the issue of managing competing behaviors. Since the set of active microbehaviors must share perceptual and motor resources, there must be some mechanism to arbitrate their needs when they make conflicting demands. The *context level* of the hierarchy maintains an appropriate set of active behaviors from a much larger library of possible behaviors, given the agents current goals and environmental conditions. The composition of this set is evaluated at every simulation interval, which we take to be 300 milliseconds.

Abstraction Level	**Problem Being Addressed**	**Role of Vision**
Behavior	Need to get state information	Provide State Estimation
	The current state needs to be updated to reflect the actions of the body	None
Arbitration	Active behaviors may have competing demands for body, legs, eyes. Conflicts have to be resolved	Move gaze to the location that will minimize risk
Context	Current set of behaviors B is inadequate for the task. Have to find a new set	Test for off-agenda exigencies

Table I. The organization of human visual computation from the perspective of the microbehavior model.

The issues that arise for vision are very different at the different levels of the hierarchy. Moving through the levels, starting with the least abstract:

(1) At the level of individual behaviors, vision provides its essential role of computing state information. The issue at this level is understanding how vision can be used to compute state information necessary for meeting behavioral goals. Almost invariably, the visual computation needed in a task context is vastly

simpler than that required general purpose vision and, as a consequence, can be done very quickly.

(2) At the arbitration level, the principal issue for vision is that the center of gaze is not easily shared and instead generally must be allocated sequentially to different locations. Eye tracking research increasingly is showing that all gaze allocations are purposeful and directed toward computing a specific result [Land et al. 1999; Hayhoe et al. 1998; Johansson et al. 1999]. Our own model [Sprague and Ballard 2003a] shows how gaze allocations may be selected to minimize the risk of losing reward in the set of running microbehaviors.

(3) At the context level, the focus is to maintain an appropriate set of microbehaviors to deal with internally generated goals. One of these goals is that the set of running behaviors be response to rapid environmental changes. Thus the issue for vision at this level is understanding the interplay between agenda-driven and environmentally-driven visual processing demands.

This hierarchy immediately presents us with a related set of implementation 'questions: How do the microbehaviors get perceptual information? How is contention managed? How are sets of microbehaviors selected? Subsequent sections use the hierarchical structure as a framework to address each of these in turn, emphasizing implications for vision.

5. STATE ESTIMATION USING VISUAL ROUTINES

The first question that must be addressed is how individual microbehaviors map from sensory information to internal state descriptions. The position we adopt is that this information is gathered by deploying visual routines. These are a small library of special-purposed functions that can be composed. The arguments for visual routines have be made by [Ullman 1985; Roelfsema et al. 2000; Kosslyn and Shwartz 1977; Ballard et al. 1997]. The main one is that the representations of vision such as color and form, are problem-neutral in that they do not contain explicitly the data upon which control decisions are made.[1] and thus an additional processing step must be employed to make decisions. The number of potential decisions that must be made is too large to pre-code them all. Visual routines address this problem in two ways: 1) routines are composable and 2) routines process visual data in an as-needed fashion.

To illustrate the use of visual routines, we describe the ones that create the state information for three of Walter's microbehaviors: *collision avoidance*,*sidewalk navigation* and *litter collection*. Each of these requires specialized processing. This processing is distinct from that used to obtain the feature images of early vision even though it may use such images as data. The specific processing steps are visualized in Figure 4.

—**Litter collection** is based on color matching. Litter is signaled in our simulation by purple objects, so that potential litter must be isolated as being of the right color and also nearby. This requires combining and processing the hue image with

[1]Marr recognized this difficulty of processing visual data prior to knowing what it will be needed for implicitly in his 'principle of least commitment' [Marr 1982].

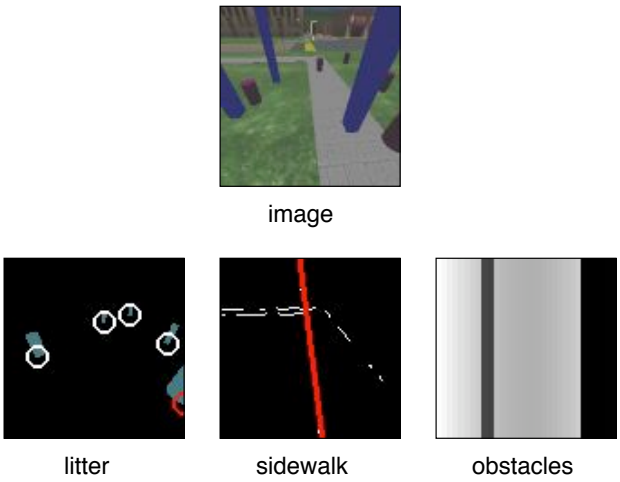

Fig. 4. The Visual Routines that compute state information. a) Input image from Walter's viewpoint. b) Regions that fit the litter color profile. Probable litter locations are marked with circles. c) Processed image for sidewalk following. Pixels are labeled in white if they border both sidewalk and grass color regions. The red line is the most prominent resulting line. b) One dimensional depth map used from obstacle avoidance (not computed directly from the rendered image).

depth information. Depth information may be obtained from any of a number of cues, (stereo, kinetic depth, parallax depth, etc.) but here, since the image processing is not our central focus, we use the prosthesis of the graphic simulator which allows us to sample depth from the scene graph directly. The result of this processing is illustrated in Figure 4b.

—**Sidewalk navigation** uses color information to label pixels that border both sidewalk and grass regions. A line is fit to the resulting set of pixels which indicates the estimated edge of the sidewalk. The result of this processing is illustrated in Figure 4c.

—**Collision detection** uses a one-dimensional depth image. For collisions, it must be processed to isolate potential colliders. As in litter collection, depth information is obtained from the scene graph. The result of this processing is illustrated in Figure 4d. A study with human subjects shows that they are very good at this, integrating motion cues with depth to ignore close objects that are not on a collision course [Ballard and Sprague 2002].

Regardless of the specific methods of individual routines, each one outputs information in the same abstract form: the state needed for its encompassing microbehavior that associates states with actions. The next section describes how Walter can learn these associations.

Outcome	Immediate Reward
Picked up a litter can	2
On sidewalk	1
Collision free	4

Table II. Walter's reward schedule

6. LEARNING MICROBEHAVIORS

The basic structure of a microbehavior is a table that associates states with actions together with the value of taking the action. Such tables can be learned by reward maximization algorithms: Walter tries out different actions in the course of behaving and remembers the ones that worked best in the appropriate table. The reward-based approach is motivated by studies of human behavior that show that the extent to which humans make such trade-offs is very refined [Trommershuser et al. 2003] as well as studies using monkeys that reveal the use of reinforcement signals in a way that is consistent with reinforcement learning algorithms [Suri and Schultz 2001].

Formally, the task of each microbehavior is to map from an estimate of the relevant environmental state s, to one of a discrete set of actions, $a \in A$, so as to maximize the amount of reward received. For example the the obstacle avoidance behavior maps the distance and heading to the nearest obstacle $s = (d, \theta)$ to one of three possible turn angles, that is, $A = \{-15^o, 0^o, 15^o\}$. The *policy* is the action so prescribed for each state. The coarse action space simplifies the learning problem.

Our approach to computing the optimal policy for a particular behavior is based on a standard reinforcement learning algorithm, termed Q-learning[Watkins and Dayan 1992]. This algorithm learns a value function $Q(s, a)$ for all the state-action combinations in each microbehavior. The Q function denotes the expected discounted return if action a is taken in state s and the optimal policy is followed thereafter. If $Q(s, a)$ is known then the learning agent can behave optimally by always choosing $\arg\max_a Q(s, a)$(See Appendix for details). Figure 5 shows the table used by the litter collection microbehavior, as indexed by its state information.

Each of the three microbehaviors has a two-dimensional state space. The litter collection behavior uses the same parameterization as obstacle avoidance: $s = (d, \theta)$ where d is the distance to the nearest litter item, and θ is the angle. For the sidewalk following behavior the state space is $s = (\rho, \theta)$. Here θ is the angle of the center-line of the sidewalk relative to the agent, and ρ is the signed distance to the center of the sidewalk, where positive values indicate that the agent is to the left of the center, and negative values indicate that the agent is to the right. All microbehaviors use the logarithm of distance as a state table index in order to devote more of the state representation to areas near the agent and they all use the same three-heading action space described above. Table II shows Walter's reward contingencies. These are used to generate the Q-tables that serve as a basis for encoding a policy. Figure 6 shows a representation of the Q-functions and policies for the three microbehaviors.

When running the Walter simulation, each Q-table associated with a behavior is indexed every 300 milliseconds and the action that is its policy is selected and

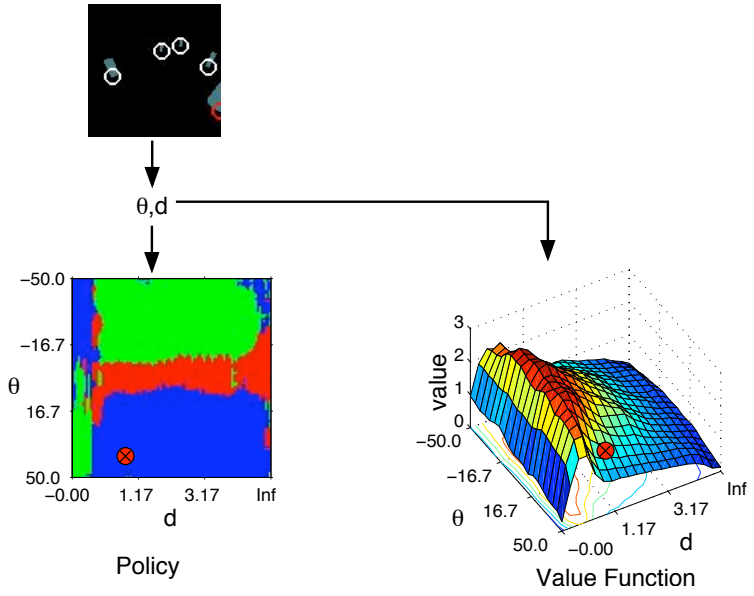

Fig. 5. The central portion of the litter cleanup microbehavior after it has been learned. The color image is used to identify the heading to the nearest litter object as a heading angle θ and distance d. This state information is used to retrieve the appropriate action as indicated in the policy table on the lower left. Green regions correspond to $turn = -15^o$, red regions to $turn = 0^o$, and blue regions to $turn = -15^o$. In this case the selected action is $turn = -15^o$. The fact that the model is embodied means that we can assume there is neural circuitry to translate the abstract heading into complex walking movements. This is true for the graphics figure that has a 'walk' command that takes a heading parameter. The state information can also be used to retrieve the expected return associated with the optimal action as illustrated on the lower right.

submitted for arbitration. The action chosen by the arbitration process is executed by Walter. This in turn results in a new Q-table index for each microbehavior and the process is repeated. The path through a Q-table thus evolves in time and can the visualized as a thread of control analogous to the use of the term thread in computer science. The thread concept will be very useful when we address the issue of how many microbehaviors can be active simultaneously.

7. MICROBEHAVIOR ARBITRATION

A central complication with the microbehavior approach is that concurrently active microbehaviors may prefer incompatible actions. Therefore an arbitration mechanism is required to map from the recommendations of the individual microbehaviors to final action choices. The arbitration problem arises in directing the physical control of the agent, as well as in handling gaze control. Each of these requires a different solution because in Walter's environment, his heading can be a compromise between the actions recommended by different microbehaviors but his gaze

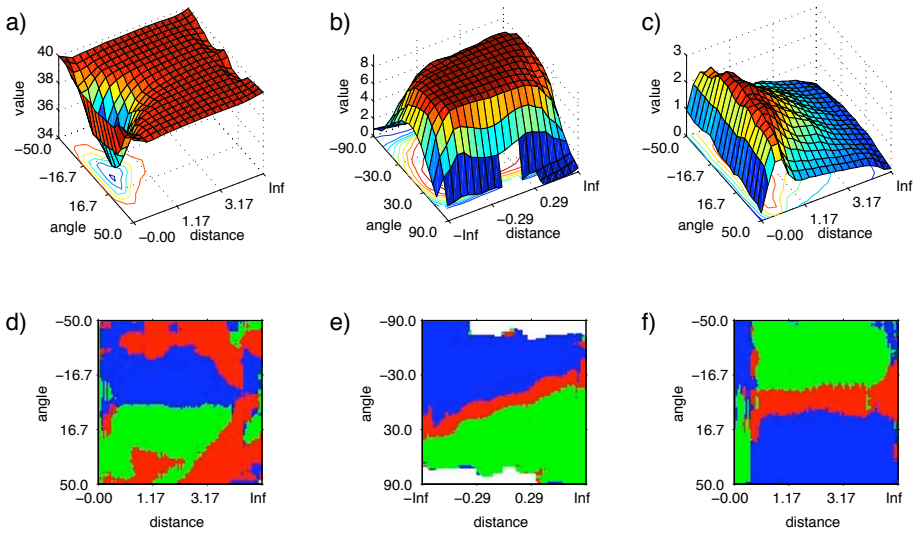

Fig. 6. Q-values and policies for the three microbehaviors. Figures a)-c) show $\max_a Q(s,a)$ for the three microbehaviors: a) obstacle avoidance, b) sidewalk following and c) litter collection. Figures d)-f) show the corresponding policies for the three microbehaviors. The obstacle avoidance value function shows a penalty for nearby obstacles and a policy of avoiding them. The sidewalk policy shows a benefit for staying in the center of the sidewalk $\theta = 0, \rho = 0$. The litter policy shows a benefit for picking up cans that decreases as the cans become more distant. The policy is to head toward them.

location cannot. A benefit of knowing the value functions for the individual behaviors is that the they can be used to handle the arbitration problem in each of these cases.

Heading Arbitration Since in the walking environment each behavior shares the same action space Walter's heading arbitration is handled by making the assumption that the Q-function for the composite task is approximately equal to the sum of the Q-functions for the component microbehaviors [Sprague and Ballard 2003c]:

$$Q(s,a) \approx \sum_{i=1}^{n} Q_i(s_i,a), \tag{1}$$

where $Q_i(s_i,a)$ represents the Q-function for the ith active behavior. The idea of using Q-values for multiple goal arbitration was independently introduced in [Humphrys 1996] and [Karlsson 1997].

The state that indexes the table can deviate from the true state for the following reason. In order to simulate the fact that only one area of the visual field may be foveated at a time, only one microbehavior is allowed access to perceptual information during each 300ms simulation time step. That behavior is allowed to update its state information with a measurement, while the others propagate their estimates

and suffer an increase in uncertainty. Note that we are not modeling the spatial targeting of eye movements. We are addressing only the issue of scheduling: which microbehavior should be given access to perception during each 300ms interval.

The mechanics of maintaining state estimates and tracking uncertainty are handled using Kalman filters - one for each microbehavior. In order to simulate noise in the estimators, the state estimates are corrupted with zero-mean normally distributed random noise at each time step. The noise has a standard deviation of .2m in both the x and y dimensions. When a behavior's state has just been updated by its visual routine's measurement, the variance of the state distribution will be small, but as we will demonstrate in simulation, in the absence of such a measurement the variance can grow significantly.

Since Walter may not have perfectly up to date state information, he must select the best action given his current estimates of the state. A reasonable way of selecting an action under uncertainty is to select the action with the highest expected return. Building on Equation (1) we have the following: $a_E = \arg\max_a E[\sum_{i=1}^{n} Q_i(s_i, a)]$, where the expectation is computed over the state variables for the microbehaviors. By distributing the expectation, and making a slight change to the notation we can write this as:

$$a_E = \arg\max_a \sum_{i=1}^{n} Q_i^E(s_i, a), \tag{2}$$

where Q_i^E refers to the expected Q-value of the ith behavior. In practice we estimate these expectations by sampling from the distributions provided by the Kalman filter.

Gaze Arbitration The reinforcement learning algorithms that are used for learning microbehavior controllers cannot be applied directly to the problem of allocating gaze. Eye movements are difficult to put in the same framework because they have only indirect consequences: they do not change the physical state of the agent or the environment; they serve only to obtain information. The alternative is to use the Q-tables learned for physical control to estimate the value of possible gaze allocations. Simply put, as time passes the uncertainty of the state of a microbehavior grows, introducing the possibility of low rewards. Deploying gaze to measure that state reduces this risk.

Estimating the cost of uncertainty is equivalent to estimating the expected cost of incorrect action choices that result from uncertainty. Given that the Q functions are known, and that the Kalman filters provide the necessary distributions over the state variables, it is straightforward to estimate, this factor, $loss_b$, for each behavior b by sampling (See Appendix). The maximum of these values is then used to select which behavior should be given control of gaze.

Figure 7 gives an example of seven consecutive steps of the sidewalk navigation task, the associated perceptual decisions, and the corresponding state estimates. Perception is allocated to reduce the uncertainty where it has the greatest potential negative consequences for reward. For example, the agent attends to the obstacle as he draws close to it, and shifts perception to the other two microbehaviors when the obstacle has been safely passed. Note that the regions corresponding to state estimates are not ellipsoidal because they are being projected from world-space into

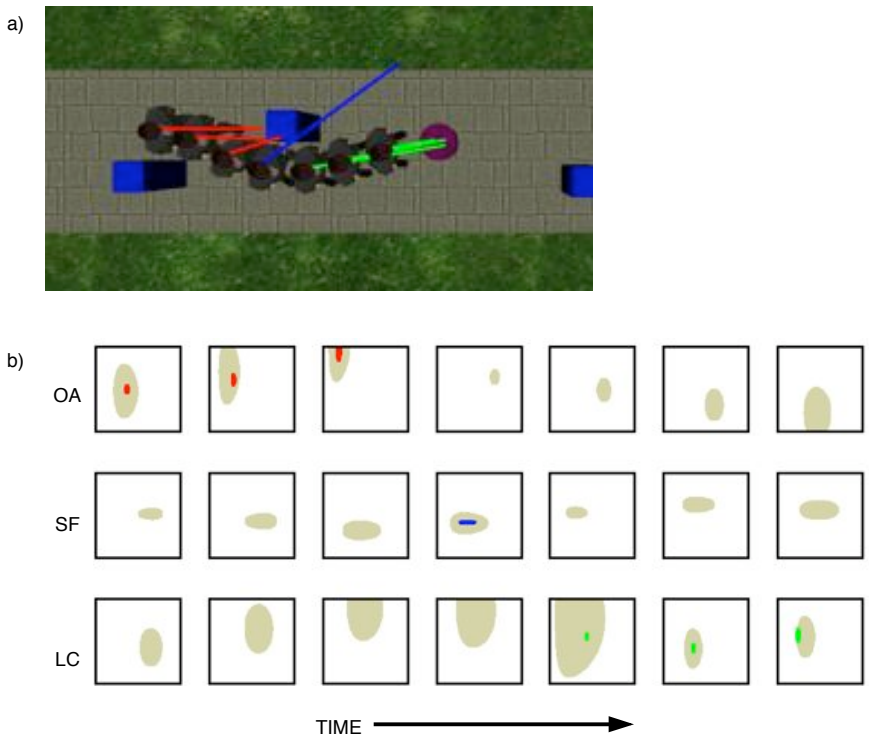

Fig. 7. a) An overhead view of the virtual agent during seven time steps of the sidewalk navigation task. The blue cubes are obstacles, and the purple cylinder is litter. The rays projecting from the agent represent gaze allocation; red correspond to obstacle avoidance, blue correspond to sidewalk following, and green correspond to litter collection. b) Corresponding state estimates. The top row shows the agent's estimates of the obstacle location. The boxes are abstracted versions of the state tables shown in Figure 6. For example, the top row shows seven successive instances of the obstacle avoidance (OA) state space used in Figure 6 *a* and *b* The beige regions correspond to the 90% confidence bounds before any perception has taken place. The red regions show the 90% confidence bounds after an eye movement has been made. The second and third rows show the corresponding information for sidewalk following(SF) and litter collection(LC).

the agents non-linear state space.

One possible objection to this model of eye movements is that it does not address the contribution of extra-foveal vision. One might assume that the pertinent question is not which microbehavior should direct the eye, but which location in the visual field should be targeted to best meet the perceptual needs of the whole ensemble of active microbehaviors. There are a number of reasons that we address the former question and not the latter. First, eye tracking studies in natural tasks show little evidence of "compromise" fixations. That is, nearly all fixations are clearly directed to a particular item that is task relevant. Second, results in [Roelf-

sema et al. 2003] suggest that simple visual operations such as local search and line tracing require a minimum of 100-150ms to complete. This time scale roughly corresponds to the time required to make a fixation. This suggests that there is little to be gained by sharing fixations among multiple visual operations. Such sharing can be demonstrated under controlled experimental conditions, but for the reasons outlined above it is unlikely to play a major role in natural behavior.

8. MICROBEHAVIOR SELECTION

The successful progress of Walter is based on having a running set of N microbehaviors that are appropriate for the current environmental and task contexts. For example, on the sidewalk the set $\{OA, SF, LC\}$ suffices, but later on when crossing the street, a different set is needed. The view that visual processing is mediated by a small set of microbehaviors immediately raises two questions: 1) What is the exact nature of the context switching mechanism? and 2) What should the limit on N be to realistically model the limitations of human visual processing?

Answering the first question requires considering to what extent visual processing is driven in a top down fashion by internal goals, versus being driven by bottom up signals originating in the environment. The Walter model reflects our view that vision is predominantly a top-down process. Thus the model of the switching mechanism is that it works as a state machine as shown in Figure 8. For the planned tasks, certain microbehaviors keep track of the progress through the task and trigger new sets of behaviors at predefined junctures. For example, the microbehavior "Look for Crosswalk" triggers the state NEAR-CROSSWALK which contains three microbehaviors: "Follow Sidewalk", "Avoid Obstacles", and "Approach Crosswalk." Figure 8B shows when the different states were triggered on three separate trials.

The Walter model is sufficient for handling simple planned tasks, but it does not provide a straightforward way of responding to off-plan contingencies. Interrupts from dynamic scene cues can automatically attract the brain's "attentional system" in order to make the correct context switch e.g [Itti and Koch 2000; Navalpakkam and Itti 2005]. The difficulty with predominantly bottom-up interrupts is that what constitutes a relevant cue is highly task dependent. Nonetheless, to be more realistic, the model would require at least two additions. First, microbehaviors should be designed to error-check their sensory input. In other words, if a microbehavior's inputs do not match expectations, it should be capable of passing control to a higher level procedure for resolution. Second, there should be a low latency mechanism for responding to certain unambiguously important signals such as rapid looming.

We now take up the question of the possible number of active microbehaviors. There are at least two reasons to suspect that the maximum number that are simultaneously running might be modest. The first reason is the ubiquitous observation of the limitations of spatial working memory (SWM). The original capacity estimate by Miller was seven items plus or minus two [Miller 1956], but current estimates favor the lower bound [Luck and Vogel 1997]. We hypothesize that this limitation is tied to the number of independently running microbehaviors which we have termed threads. The identification of the referents of SWM has always been problematic, since the size of the referent can be arbitrary. This has lead to the denotation of

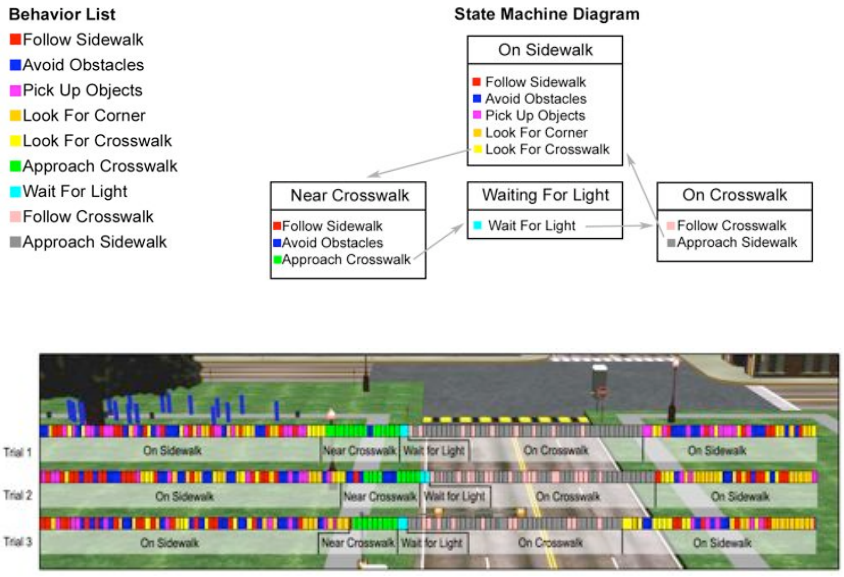

Fig. 8. (Top left) A list of microbehaviors used in Walter's overall navigation task. (Top right) The diagram for the programmable context switcher showing different states. These states are indicated in the bands underneath the colored bars below. Bottom) Context switching behavior in the sidewalk navigation simulation for three separate instances of Walter's stroll. The different colored bars denote the different microbehaviors that are in control of the gaze at any instant.

the referent as a 'chunk,' a jargon word that postpones dealing with the issue of not being able to quantify the referents. The thread concept is clearer and more specific as it denotes exactly the state necessary to maintain a microbehavior.

The second factor limiting the number of running microbehaviors is that large numbers of active microbehaviors may not be possible given that they have to be implemented in a neural substrate. Cortical memory is organized into distinct areas that have a two-dimensional topography. Furthermore spatial information is usually segregated from feature based information so that the neurons representing the colors of two objects are typically segregated from the neurons representing their location. As a consequence there is no simple way of simultaneously associating one object's color with its location together with another object's association of similar properties (This difficulty is the so-called "binding problem" [von der Malsburg 1999]). Some proposals for resolving the binding problem hypothesize that the number of active microbehaviors is limited to one, but this seems very unlikely. However the demands of a binding mechanism may limit the number of simultaneous bindings that can be active. Thus it is possible that such a neural constraint may be the basis for the behavioral observation.

Although the number of active microbehaviors is limited there is reason to believe that it is greater than one. Consider the task of walking on a crowded sidewalk.

Two fast walkers approaching each other close at the rate of 6 meters/second. Given that the main source of advanced warning for collisions is visual and that eye fixations typically need 0.3 seconds and that cortical processing typically needs 0.2-0.4 seconds, during the time needed to recognize an impending collision, the colliders have traveled about 3 meters, or about one and a half body lengths. In a crowded situation, this is insufficient advance warning for successful avoidance. What this means is that for successful evasions, the collision detection calculation has to be ongoing. But that in turn means that it has to share processing with the other tasks that an agent has to do. Remember that by sharing we mean that the microbehavior has to be simultaneously active over a considerable period, perhaps minutes. Several elegant experiments have shown that there can be severe interference when multiple tasks have to be done simultaneously, but these either restrict the input presentation time [VanRullen et al. 2004] or the output response time [Pashler 1998]. The crucial issue is what happens to the internal state when it has to be maintained for an extended period.

9. TESTING THE MODEL

The goal of developing the model is to gain insight into human sensori-motor processing by designing an artificial agent that is capable of handling realistic temporally extended tasks under a set of constraints similar to those faced by a real human. Validating a model of this sort is vastly more difficult than validating a traditional psychophysical model that makes predictions about one isolable aspect of sensori-motor control. As a result, our goal here is not to demonstrate that our virtual human model is correct in all of its particulars. Instead, we hope to demonstrate that the model's performance is sufficiently consistent with observed human performance to suggest that the model can provide a valuable starting point for understanding human behavior in natural tasks.

To do this we introduced humans into the virtual environment and had them walk Walter's walk. The humans wear a head-mounted binocular display (HMD) that contains monocular eye tracking capability. in addition the rotational and translational degrees of freedom of their heads are monitored with a Hi-ball tracker. The head tracker has a latency of a few milliseconds so that the experience in the HMD has no detectable lags. The HMD has a diagonal field of view of 55^o. This is much smaller than the regular human field of view of more than 180^o but is ameliorated by two factors: 1) The free head very low latency HMD means that the subjects can have a sense of access to a larger field of view by making head movements and 2) the sidewalk and crosswalk that the subjects walk on subtends a visual angle considerably less than the display limit (See insert Figure 1).

The biggest problem faced by the overall setup is that the linear track of Walter's path is many times longer than the 7 meter width of the laboratory. Our solution to this discrepancy was to map a curved path in motor space onto a linear path in visual space. That is, in order to experience a linear path in visual space, the subjects have to walk a circular path in the laboratory. A typical transit of Walter's path takes about four laps of this path. Subjects are very unaware of this manipulation and, when asked, drastically underestimate their transit as a lap or two. Eye movement data for each of six subjects is collected and scored on a

Fig. 9. The head mounted display worn by human subjects has eye tracking capability so that gaze can be tracked in virtual environments. The display has a 55^o diagonal field of view.)

frame-by-frame basis. As shown in Figure 8, Walter's path consists of a sidewalk portion where he has to handle staying on the sidewalk, obstacle avoidance and litter, and then a clear sidewalk segment followed by a crossing of a street. the crossing of the street is regulated with a large traffic light. Three subjects walked the sidewalk portion only and three additional subjects subsequently walked the entire segment.

The first claim of the model is that the use of fixation is to gather information for a small set of active microbehaviors. This claim is borne out in a number of ways in the subjects' data. In the first place over 95% of the fixations can be interpreted as gathering information for one of the ongoing tasks. For example in the initial segment, the fixations are invariably on the edge of the sidewalk or on a blue pillar (the obstacle) or a purple box (the litter). Figure 10 shows examples of the scored fixations.

The lack of off-task fixations cannot be explained by bottom-up mechanisms. The virtual environment is visually rich and includes many possible fixation targets that are not task related. This point is highlighted by comparing the record of fixations to the predictions of a bottom-up saliency model. This is accomplished by sampling individual frames and running the Itti saliency computation on each frame[Itti and Koch 2000]. This software is a reasonable model for human eye fixations in images. Its central claim is that constellations of image features define locations of "saliency." Observed points of fixation in an image can be explained as being chosen from the most salient of these locations. Our situation is different than an image-based test since our subjects are immersed in a 3D environment and have a very specific task agenda.

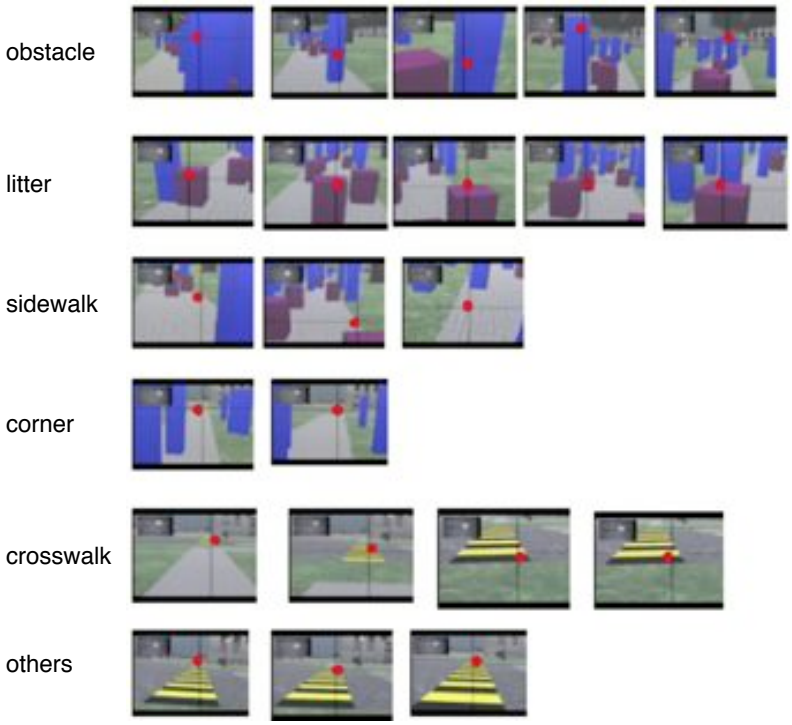

Fig. 10. Sample scored frames from the human video augmented with a red dot to highlight the fixation point.)

In a limited sample, we compared the actual points chosen by human subjects to the points recommended by the Itti algorithm. Our comparison was generous: if one of the top five points recommended by the algorithm was on the same object as the human fixation it was scored as a match, otherwise it was denoted a non-match. Of eighteen points tested only eight matched under these criteria. The non-match points were in the majority. As all of the human data could be readily interpreted as being directly relevant to one of the three tasks, we take this as evidence for task-directed visual routines.

The second claim of the model is that in the course of natural behavior a small number of microbehaviors are active and these are competing for the gaze vector. Again this is supported by the scored human data. On the initial sidewalk segment the fixation data was directed predominantly between one of the kinds of locations relevant to the three tasks. While our data cannot rule out all alternate interpretations of gaze control besides the 'most-to-gain' strategy, such as looking at the nearest of the three objects, the model data does show similar behavior patterns to that of the human subjects. Figure 12 shows the histogram of fixations for three subjects in the initial sidewalk task compared to three runs of Walter over the same data. The figure shows that the subjects used more fixations than the model reflecting that Walter's walking speed was higher. More importantly it shows that

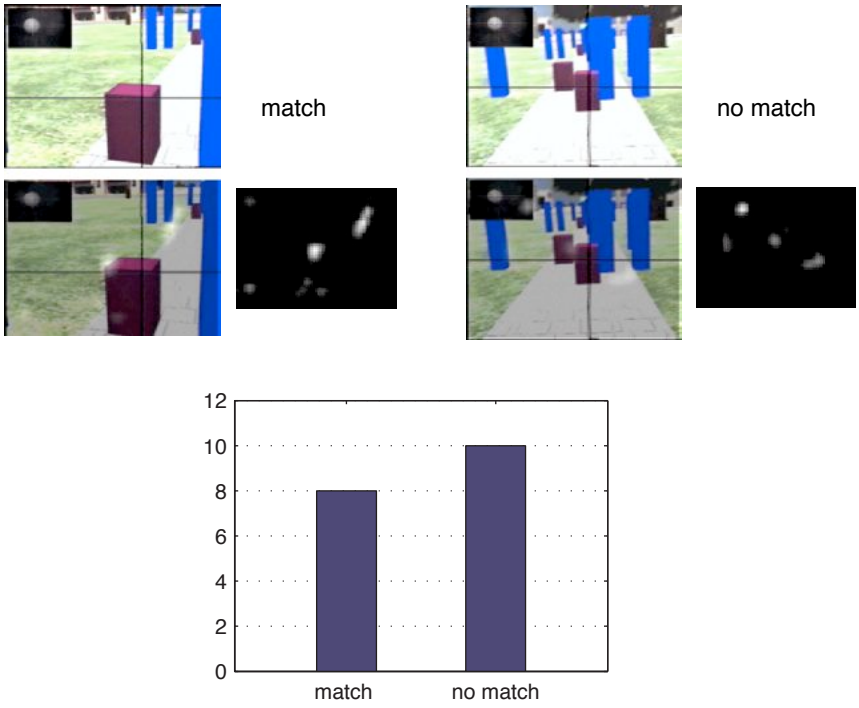

Fig. 11. Comparing human gaze locations to those found by the Itti saliency detector. The small inserts show the saliency maps that are overlaid as transparencies on the lower versions of the images. In a sample of 18 frames, more than half show fixation locations that are not detected by the maps. The saliency program was provided by Dr. Laurent Itti at USC.

the relative proportions of fixations on locations relevant to each of the three tasks was similar. Of course we chose the relative rewards in Table 6 to model the human data but the coarseness of values in that table shows that no extensive tuning was done.

The one discrepancy in the table is that the humans use fewer sidewalk fixations than suggested by the model. Our explanation is that the human subjects make some litter and obstacle fixations do double duty. For example, if you are on the sidewalk and fixating an unobstructed litter can that is also on the sidewalk, you can confidently walk toward it knowing you will remain on the sidewalk. This highlights one of the challenges in interpreting human fixations in complex scenes: it is not always possible to uniquely identify a given fixation with a given task.

10. CONCLUSIONS

The focus of this paper was to introduce the issues associated with using a graphical agent as a proto-theory of human visuo-motor behavior. One possible criticism of such a project is that, even though the system is vastly reduced from that needed to capture a substantial fraction of human behavior, the model as it stands is

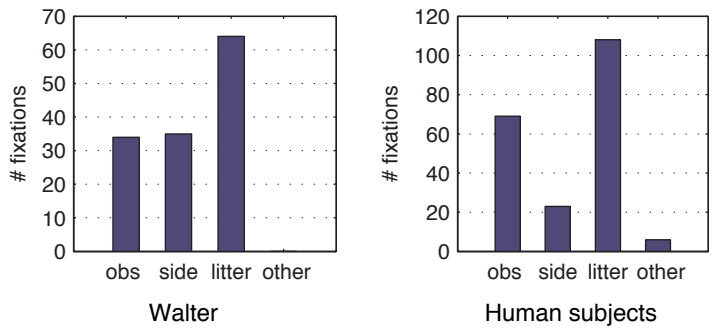

Fig. 12. Comparing the model and human subjects' fractional gaze allocation to different tasks)

complicated and has enough free parameters so that any data from real human performance would be easy to fit. Although the system is complex, most of the constraints follow from the top-level assumption of composable microbehaviors. Once one decides to have a set of running microbehaviors, the questions of how many and when are they running are immediate. Furthermore they have ready answers in observations of human behavior in the classic observations of working memory and eye movements: Working memory suggests the number of simultaneous microbehaviors is small; eye movements suggest which behavior is running as each fixation is an indication of the instantaneous problem being addressed. The restricted number of active microbehaviors means that there must be a mechanism for making sure that a good behavioral subset has been chosen. Such a mechanism must interrogate the environment and 1) add needed microbehaviors as well as 2) drop microbehaviors if needed to meet the capacity constraint.

The reinforcement learning venue provides a different perspective on gaze allocation. One of the original ideas was a bottom-up view that gaze should be drawn to the most salient locations in the scene as represented in the image, where salience was defined in terms of the spatial conjunction of many feature points. However recent measurements have shown that eye movements are much more agenda driven than that predicted by bottom-up saliency models. For example Henderson has shown that subjects examining empty urban scenes for people examine places where people *might be* even though these can have very low feature saliency [Henderson 2003].

We think that the paper makes several important contributions to the understanding of natural behavior:

(1) The current notion of attention per se as a general resource is too simplistic to capture the complexity of even simple behaviors in complex environments. As the model shows, resource allocation is a complex issue that raises questions at different levels of abstraction. To our knowledge, we are the first to propose a specic computational hierarchy in the brain to address the different resource management problems that occur at different time scales.

(2) Our simulation reveals that it is useful to make a distinction between a men-

tal program that is active and its momentary demands to access the bodys resources, such as head, hand and eye gaze. The simulations suggest that the interplay between multiple running programs can be much more subtle than that observed in traditional single-trial psychophysics.

(3) Walter's use of Q-tables suggest that to interpret gaze allocation, an additional level of indirection may be required. For example, the controller for sidewalk navigation uses gaze to update the estimate of the location of the sidewalk. In order to predict when gaze might be allocated to do this, in our model, requires knowing the uncertainty in the current estimate of the sidewalk location.

(4) Our model of gaze is the only model that addresses why eyes go to a place in terms of competing tasks demands. All the other models regard the problem of addressing gaze allocation as a function of current image features. To the extent that task effects are modeled they are only done so as a modulation of image features. Furthermore these experiments do not directly address the use of eye movements to maximize task reward even though there has been a great deal of research that does tie expected reward to saccade timing and to neural signals correlated with saccade timing [Hikosaka et al. 2000; McCoy et al. 2003; Watanabe et al. 2003; Itoh et al. 2003].

(5) We show that individual reinforcement learning tables could be composed to produce more complex behaviors. when their actions are expressed in a common language of body heading. This is an important illustration of the value of the motor system as a final common arbiter. It is also an illustration of the way potentially very large libraries of tables could be combined in small subsets.

The human subject experiments do not definitively validate our model, but are consistent with the model, and they illustrate the methodology that we propose for studying human sensori-motor behavior. Future work will focus on experimentally manipulating task parameters and observing the effect on both human and model performance. An example of a possible experiment would be to leave the environment unchanged, but alter the task relevance of different aspects of the environment. For example, repeat the sidewalk task, but remove the instruction that litter should be collected. Our model would predict a redistribution of fixations frequencies that could be tested against human results.

Perhaps the most important theme in recent vision research, is that no component of the visual system can be properly understood in isolation from the behavioral goals of the organism. Therefore, properly understanding vision will ultimately require modeling complete sensori-motor systems in behaving agents. The model presented in this paper abstracts away many details unspecified but it does provide an abstract framework for thinking about action-oriented human vision.

Appendix: Reinforcement Learning Details

Learning behaviors There are a number of algorithms for learning $Q(s, a)$ [Kaelbling et al. 1996; Sutton and Barto 1998] the simplest is to take random actions in the environment and use the Q-learning update rule [Watkins 1989]:

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a'))$$

Here $\alpha \in (0,1)$ is a learning rate parameter, $\gamma \in (0,1)$ is a term that determines how much to discount future reward, and s' is the state that is reached after action a. As long as each state-action pair is visited infinitely often in the limit, this update rule is guaranteed to converge to the optimal value function. The Q-learning algorithm is guaranteed to converge only for discrete case tasks with Markovian transitions between states. Walter's tasks are more naturally described using continuous state variables. The theoretical foundations of continuous state reinforcement learning are not as well established as for the discrete state case. However empirical results suggest that good results can be obtained by using a function approximator such as a CMAC along with the Sarsa(0) learning rule: [Sutton 1996]

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha(r + \gamma Q(s',a'))$$

This rule is nearly identical to the Q-learning rule, except that the max action is replaced by the action that is actually observed on the next step. The Q-functions used throughout this paper are learned using this approach. A more detailed account of the learning procedure can be found in [Sprague 2004].

Choosing behaviors for a state update Whenever Walter chooses an action that is sub-optimal for the true state of the environment, he can expect to lose some return. We can estimate the expected loss as follows:

$$loss = E[\max_a \sum Q_i(s_i,a)] - E[\sum Q_i(s_i,a_E)]. \quad (3)$$

The term on the left-hand side of the minus sign expresses the expected return that Walter would receive if he were able to act with knowledge of the true state of the environment. The term on the right expresses the expected return if he is forced to choose an action based on his state estimate. The difference between the two can be thought of as the cost of the agent's current uncertainty. This value is guaranteed to be positive, and may be zero if all possible states would result in the same action choice.

The total expected loss does not help to select *which* of the microbehaviors should be given access to perception. To make this selection, the loss value can be broken down into the losses associated with the uncertainty for each particular behavior b:

$$loss_b = E\left[\max_a \left(Q_b(s_b,a) + \sum_{i \in B, i \neq b} Q_i^E(s_i,a)\right)\right] - \sum_i Q_i^E(s_i,a_E). \quad (4)$$

Here the expectation on the left is computed only over s_b. The value on the left is the expected return if s_b were known, but the other state variables were not. The value on the right is the expected return if none of the state variables are known. The difference is interpreted as the cost of the uncertainty associated with s_b.

REFERENCES

Aivar, M. P., Hayhoe, M. M., Chizk, M. M., and Mruczek, R. E. B. 2005. Spatial memory and saccadic targeting in a natural task. *Journal of Vision. 4*, 1–3.

Babcock, J. and Pelz, J. B. 2000. The rit wearable eyetracker. *ACM SIGCHI Eye Tracking Research and Application Symposium*.

Ballard, D., Hayhoe, M., and Pelz, J. 1995. Memory representations in natural tasks. *Cognitive Neuroscience. 7*, 66–80.

Ballard, D., Hayhoe, M., and Pook, P. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences 20*, 723–767.

Ballard, D. and Sprague, N. 2002. Attentional resource allocation in extended natural tasks [abstract]. *Journal of Vision 2,* 7, 568a.

Brooks, R. A. 1986. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation RA-2,* 1 (Apr.), 14–23.

Bryson, J. J. and Stein, L. A. 2001. Modularity and design in reactive intelligence. In *International Joint Conference on Artificial Intelligence*. Seattle, Washington.

Clark, A. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Clark, A. 1999. An embodied model of cognitive science?l. *Trends in Cognitive Sciences 3*, 345–351.

Eckstein, M. P., J. P., T., Palmer, J., and Shimozaki, S. S. 2000. A signal detection model predicts effects of set size on visual search accuracy for feature, conjunction, triple conjunction and disjunction displays. *Perception and Psychophysics 62*, 425–451.

Fabre-Thorpe, M., Delorme, A., Marlot, C., and Thorpe, S. 2001. A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience 13*, 1–10.

Faloutsos, P., van de Panne, M., and Terzopoulos, D. 2001. The virtual stuntman: Dynamic characters with a repertoire of motor skills. *Computers and Graphics 25*, 933–953.

Firby, R. J., Kahn, R. E., Prokopowicz, P. N., and Swain, M. J. 1995. An architecture for vision and action. 72–79.

Hartley, R. and Pipitone, F. 1991. Experiments with the subsumption architecture. In *Proceedings of the International Converence on Robotics and Automation*.

Hayhoe, M. M., Bensinger, D., and Ballard, D. H. 1998. Task constraints in visual working memory. *Vision Research 38*, 125–137.

Hayhoe, M. M., Shrivastava, A., Mruczek, R., and Pelz, J. 2003. Visual memory and motor planning in a natural task. *Journal of Vision. 3*, 49–63.

Henderson, J. M. 2003. Human gaze control in real-world scene perception. *Trends in Cognitive Sciences 7*, 498–504.

Hikosaka, O., Takikawa, Y., and Kawagoe, R. 2000. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Reviews 80,* 3 (July).

Humphrys, M. 1996. Action selection methods using reinforcement learning. In *Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*.

Itoh, H., Nakahara, H., Hikosaka, O., Kawagoe, R., Takikawa, Y., and Aihara, K. 2003. Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior. *Journal of Neurophysiology 89*, 1774–1783.

Itti, L. and Koch, C. 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research 40,* 10-12 (May), 1489–1506.

Johansson, R., Westling, G., Backstrom, A., and Flanagan, J. R. 1999. Eye-hand coordination in object manipulation. *Perception 28*, 1311–1328.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. 1996. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research 4*, 237–285.

Karlsson, J. 1997. Learning to solve multiple goals. Ph.D. thesis, University of Rochester.

Kosslyn, S. M. and Shwartz, S. 1977. A simulation of visual imagery. *Cognitive Science 1*, 265–269.

Land, M., Mennie, N., and Rusted, J. 1999. The roles of vision and eye movements in the activities of daily living. *Perception 28*, 1311–1328.

Luck, S. J. and Vogel, E. K. 1997. The capacity of visual working memory for features and conjunctions. *Nature 390*, 279–281.

Marr, D. 1982. *Vision*. W.H. Freeman and Co., Oxford.

McCoy, A. N., Crowley, J. C., Haghighian, G., Dean, H. L., and Platt, M. L. 2003. Saccade reward signals in posterior cingulate cortex. *Neuron 40*, 1031–1040.

Merleau-Ponty, M. 1962. *Phenomenology of Perception*. Routledge & Kegan Paul.

Miller, G. 1956. The magic number seven plus or minus two: Some limits on your capacity for processing information. *Psychological Review 63*, 81–96.

Navalpakkam, V. and Itti, L. 2005. Modeling the influence of task on attention. *Vision Research 45,* 2, 205–231.

Newell, A. 1990. *Unified Theories of Cognition*. Harvard University Press.

Newell, F. N., Ernst, M. O., Yjan, B. S., and Blthoff, H. H. 2001. Viewpoint dependence in visual and haptic object recognition. *Psychological Science 12*, 37–42.

Nilsson, N. 1984. Shakey the robot. Tech. Rep. 223, SRI International,.

Noe, A. 2005. *Action in Perception*. MIT Press.

O'Regan, J. K. and Noe, A. 2001. A sensorimotor approach to vision and visual consciousness. *Behavioral and Brain Sciences 24*, 939–973.

Palmer, J. 1995. Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science 4*, 118–123.

Pashler, H. 1998. *The Psychology of Attention*. Cambridge, MA: MIT Press.

Pelz, J. B., Hayhoe, M. M., and Loeber, R. 2001. The coordination of eye, head, and hand movements in a natural task. *Exp. Brain Research 139*, 266–277.

Roelfsema, P., Lamme, V., and Spekreijse, H. 2000. The implementation of visual routines. *Vision Research 40*, 1385–1411.

Roelfsema, P. R., P.S., K., and Spekreijse, H. 2003. Subtask sequencing in the primary visual cortex. *Proceedings of the National Academy of Sciences USA 100*, 5467–5472.

Roy, D. and Pentland, A. 2002. Learning words from sights and sounds: A computational model. *Behavioral and Brain Sciences 26*, 113–146.

Sprague, N. 2004. Learning to coordinate visual behaviors. Ph.D. thesis, University of Rochester Computer Science Dept.

Sprague, N. and Ballard, D. 2003a. Eye movements for reward maximization. In *Advances in Neural Information Processing Systems 15*.

Sprague, N. and Ballard, D. 2003b. Multiple-goal reinforcement learning with modular sarsa(0). Tech. Rep. 798, University of Rochester Computer Science Department.

Sprague, N. and Ballard, D. 2003c. Multiple-goal reinforcement learning with modular sarsa(0). In *International Joint Conference on Artificial Intelligence*.

Suri, R. E. and Schultz, W. 2001. Temporal difference model reproduces anticipatory neural activity. *Neural Computation 13*, 841–862.

Sutton, R. 1996. Generalization in reinforcement learning: Successful examples using sparse coarse coding. In *Advances in Neural Information Processing Systems*. Vol. 8.

Sutton, R. and Barto, A. 1998. *Reinforcement Learning: An Introduction*. MIT Press.

Terzopoulos, D. and Rabie, T. F. 1997. Animat vision: Active vision in artificial animals. *Videre: Journal of Computer Vision Research 1,* 1, 2–19.

Triesch, J., Ballard, D., Hayhoe, M., and Sullivan, B. 2003. What you see is what you need. *Journal of Vison. 3*, 86–94.

Trommershuser, J., Maloney, L. T., and Landy, M. S. 2003. Statistical decision theory and tradeoffs in motor response. *Spatial Vision 16,* 255–275.

Ullman, S. 1985. Visual routines. *Cognition 18*, 97–159.

VanRullen, R., Reddy, L., and Koch, C. 2004. Visual search and dual tasks reveal two distinct attentional resources. *Journal of Cognitive Neuroscience 16*.

von der Malsburg, C. 1999. The what and why of binding: the modeler's perspective. *Neuron 24*, 95–104.

Watanabe, K., Lauwereyns, J., and Hikosaka, O. 2003. Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *The Journal of Neuroscience 23,* 31 (November), 10052–10057.

Watkins, C. J. C. H. 1989. Learning from delayed rewards. Ph.D. thesis, King's College, Oxford.

Watkins, C. J. C. H. and Dayan, P. 1992. Q-learning. *Machine Learning Journal 8,* 3/4 (May).

Yu, C. and Ballard, D. 2004. A multimodal learning interface for grounding spoken language in sensorimotor experience. *ACM Transactions on Applied Perception 1*, 57–80.

Zelinsky, G. 1996. Using eye saccades to asses the selectivity of search movements. *Vision Research 36*, 2177–2187.