Current Biology 24, R622–R628, July 7, 2014 http://dx.doi.org/10.1016/j.cub.2014.05.020

Modeling Task Control of Eye Movements

Minireview

Mary Hayhoe[1] and Dana Ballard[2]

In natural behavior, visual information is actively sampled from the environment by a sequence of gaze changes. The timing and choice of gaze targets, and the accompanying attentional shifts, are intimately linked with ongoing behavior. Nonetheless, modeling of the deployment of these fixations has been very difficult because they depend on characterizing the underlying task structure. Recently, advances in eye tracking during natural vision, together with the development of probabilistic modeling techniques, have provided insight into how the cognitive agenda might be included in the specification of fixations. These techniques take advantage of the decomposition of complex behaviors into modular components. A particular subset of these models casts the role of fixation as that of providing task-relevant information that is rewarding to the agent, with fixation being selected on the basis of expected reward and uncertainty about environmental state. We review this work here and describe how specific examples can reveal general principles in gaze control.

Human vision gathers information in complex, noisy, dynamic environments to accomplish tasks in the world. In the context of everyday visually guided behavior, such as walking, humans must accomplish a variety of goals, such as controlling direction, avoiding obstacles, and taking note of their surroundings. They must manage competing demands for vision by selecting the necessary information from the environment at the appropriate time, through control of gaze. How is this done, apparently so effortlessly, yet so reliably? What kind of a control structure is robust in the face of the varying nature of the visual world, allowing us to achieve our goals? While the underlying oculomotor neural circuitry has been intensively studied and is quite well understood, we do not have much understanding of how something becomes a target in the first place [1].

It has long been recognized that the current behavioral goals of the observer play a central role in target selection [2–4]; however, obtaining a detailed understanding of exactly how gaze targets are chosen on the basis of cognitive state has proved very difficult. One reason was that, until recently, it has been difficult to measure eye movements in active behavior, so the experimental situations that were examined typically involved fixing the subject's head and measuring gaze on a computer monitor. Within this tradition it was natural first to consider how stimulus features such as high contrast or color might attract gaze. Formalizing this approach, Koch and Ullman [5] introduced the concept of a saliency map that defined possible gaze points as regions with visual features that differed from the local surround. For example, a red spot on a green background is highly salient and attracts gaze. It was quickly recognized that visual features alone are insufficient, and that stimulus-defined saliency or conspicuity is modulated by behavioral goals, or top-down factors, to determine the priority of potential gaze points.

Consequently, in later models of salience (or priority) the stimulus saliency computations were weighted by factors that reflected likely gaze locations, such as sidewalks or horizontal surfaces, or introduced a specific task such as searching for a particular object [6,7]. These models reflected the consensus that saccadic target selection is determined by activity in a neural priority map of some kind in areas such as the lateral intra-parietal cortex and frontal eye fields [8–10]. This kind of modeling has a critical limitation, however, in that it applies to situations where the subject inspects a static image on a computer monitor, and this situation does not make the same demands on vision that are made in the context of active behavior, where visual information is used to inform ongoing actions [11]. While there have been successful attempts to model specific behaviors such as reading or visual search, we need to develop a general understanding of how the priority map actively transitions from one target to the next as behavior evolves in time.

The use of a computer monitor was typically imposed by the limitation of eye tracking methodology, which required that the subject's head be in a fixed position. This limitation was removed, however, when Land [12] developed a simple head-mounted eye tracker that allowed observation of human gaze behavior in the context of everyday tasks. This development provided a more fertile empirical base for understanding how gaze is used to gather information to guide behavior. In the subsequent decades, improvements in eye tracking methodology have allowed a wide variety of natural visually guided behavior to be explored [11,13,14]. While these observations have provided very clear evidence for the control of gaze by the current cognitive agenda, a second critical roadblock has been the difficulty in developing a deeper theoretical understanding of how this agenda determines changes of gaze from one target to the next, given the inherent complexity of interactions with the world in the course of natural behavior. It is these gaze transitions that are hard to capture in standard experimental paradigms, and the problem that we address here is how to capture the underlying principles that control them.

The challenge of modeling tasks is at first blush intractable, given the diversity and complexity of visually guided behavior. Observations of gaze control in natural behavior, however, suggest a potential simplifying assumption, namely, that complex behavior can be broken down into simpler sub-tasks, or modules, that operate independently of each other, and thus must be attended to separately. For example, in walking, heading towards a goal and avoiding obstacles might be two such sub-tasks. The gaze control problem then reduces to one of choosing which sub-task should be attended at any moment; for example, whether to look towards the goal or to look for obstacles. In both these cases, gaze is taken as an indicator of the current attentional focus for subtask computation. Gaze and attention are very tightly linked [15,16] and there is now a significant body of work on natural gaze control suggesting that gaze is a good, although imperfect, indicator of the

[1]Department of Psychology, 108 E. Dean Keeton, Stop A8000, The University of Texas at Austin, Austin, TX 78712, USA. [2]Department of Computer Science, 2317 Speedway, Stop D9500, The University of Texas at Austin, Austin, TX 78712, USA.
E-mail: mary@mail.cps.utexas.edu (M.H.), dana@cs.utexas.edu (D.B.)

current attentional computation. Subjects sequentially interrogate the visual image for highly specific, task-relevant information, and use that information selectively to accomplish a particular sub-task [17–19]. In this conceptualization, vision is seen as fundamentally sequential. Thus, gazing at a location on the path ahead might allow calculation of either the current walking direction or the location of an obstacle relative to the body, but it is assumed that these computations are sequential rather than simultaneous, given that both are attentionally demanding.

While the assumption of independent sub-tasks is almost certainly an oversimplification, it is a useful first step, as it is consistent with a body of classic work on a central attentional bottleneck that limits simultaneous performance of multiple tasks (for example [20,21]). For the most part, a new visual computation will involve a shift in gaze. This is not always true, for example, when spatially global visual information is needed, or when peripheral acuity is good enough to provide the necessary information without a gaze shift; however, it is valid for many instances, and it is those cases we focus on here.

In this minireview, we shall describe models of gaze control that use the simplification of modules in two different but closely related contexts. In the first situation, we consider the problem of how to allocate attention, and hence gaze, to different but simultaneously active behavioral goals. In driving a car, simultaneously active goals might be to follow a lead car while obeying a speed limit and staying in the lane. Here, the challenge for gaze is to be in the right place at the right time, when the environment is somewhat unpredictable. This setting deals with the problem of competition between potential actions and has been labeled "the scheduling problem" [22]. The second situation tackles the problem of structuring elaborate behavioral sequences from elemental components. We consider a sequential task, such as making a sandwich, where gaze is used to provide information for an extended sequence of actions. The question asked is: given gaze location and hand movement information of a subject in the process of making a sandwich, can we determine the stage in the construction they are currently working on? It turns out that a probabilistic model, termed a dynamic Bayes network [23], provides sufficient information to identify each task stage. Thus, the observed data are used to infer the internal state that generates the behavior.

Task Modules, Secondary Reward and Reinforcement Learning

Multiple tasks that are ongoing simultaneously are a ubiquitous characteristic of general human behavior and consequently the brain has to be able to allocate resources between them. This scheduling problem can be addressed if there is a way of assigning value to the different tasks. It has been demonstrated that external reward, in the form of money or points in humans, and juice in monkeys, influences eye movements in a variety of experiments [1,24,25]. It remains to be established how to make the definitive link between the primary rewards used in experimental paradigms and the secondary rewards that operate in natural behavior, where eye movements are for the purpose of acquiring information [1,11]. In principle, the neural reward machinery provides an evaluation mechanism by which gaze shifts can ultimately lead to primary reward, and thus potentially allow us to understand the role that gaze patterns play in achieving behavioral goals. A general consensus is that this accounting is done by a secondary reward estimate, and a huge amount of research implicates the neurotransmitter dopamine in this role. It is now well established that cells in many of the regions involved in saccade target selection and generation are sensitive to expectation of reward, in addition to coding the movement itself [26–31]. The challenge is to distill this experimental data into a more formal explanation.

All natural tasks embody delayed rewards whereby decisions made in the moment must anticipate future consequences. The value of searching for a type of food must include estimates of its nutritional value, as well as costs in obtaining it. Furthermore, the value of a task at its initiation can only reflect the *expected* ultimate reward, because reward in the natural world is uncertain. Moreover, a consequence of this uncertainty is that the initial evaluation needs to be continually updated to reflect actual outcomes [32]. An important advance in this direction has been the development of reinforcement learning models. Recent research has shown that a large portion of the brain is involved in representing different computational elements of reinforcement learning models, and this provides a neural basis for the application of such models to understanding sensory-motor decisions [32–34]. Additionally, reinforcement learning has become increasingly important as a theory of how simple behaviors may be learned [33], particularly as it features a discounting mechanism that allows it to handle the problem of delayed rewards.

A central attraction of such reinforcement learning models for the study of eye movements is that they allow one to predict gaze choices by taking into account the learnt reward value of those choices for the organism, providing a formal basis for choosing fixations in terms of their expected value to the particular task that they serve. However, reinforcement learning has a central difficulty in that it does not readily scale up to realistic natural behaviors. Fortunately this problem can be addressed by making the simplifying assumption that complex behaviors can be factored into subsets of tasks served by modules that can operate more or less independently [22]. Each independent module, which can be defined as a Markov decision process, computes a reward-weighted action recommendation for all the points in its own state space, which is the set of values the process can take. As the modules are all embedded within a single agent, the action space is shared among all modules and the best action is chosen depending on the relative reward weights of the modules. The modules provide separate representations for the information needed by individual tasks, and their actions influence state transitions and rewards individually and independently. The modular approach thus allows one to divide an impractically large state space into smaller state spaces that can be searched with conventional reinforcement learning algorithms [35]. The factorization can potentially introduce state combinations for which there is no consistent policy, but experience shows that these combinations, for all practical purposes, are very rare.

Expected Reward as a Module's Fixation Protocol

The module formulation directly addresses the scheduling problem in that it allows fixation choices to be understood in terms of competing modules' demands for reward. In the driving scenario, where separate modules might