A POMDP Model of Eye-Hand Coordination

Abstract

This paper present a generative model of eye-hand coordination. We use numerical optimization to solve for the *joint behavior* of an eye and two hands, deriving the optimal motion pattern from first principles, without imposing hand-crafted heuristics. We model the planar scene as a POMDP with 17 continuous state dimensions. Belief-space optimization is facilitated by using a nominal-belief heuristic, whereby we assume (during planning) that the maximum likelihood observation is always obtained. Since a globally-optimal solution for such a high-dimensional domain is computationally intractable, we employ local optimization in the belief domain by using Differential Dynamic Programming. By solving for a locally-optimal plan through belief space, we generate a motion pattern of mutual coordination between hands and eye: the eye's saccades disambiguate the scene in a task-relevant manner, and the hands' motions anticipate the eye's saccades. Finally, the model is validated through a behavioral experiment, in which human subjects perform the same eye-hand coordination task. We show how simulation is congruent with the experimental results.

1 Introduction

Eye-hand coordination is an integral part of many human activities, and has been the subject of scientific inquiry for more than a century. This domain poses an interesting challenge of motor intelligence, because the uncertain state of the world requires disambiguation, which can only be achieved by combining information-gathering activity (gaze shift) and goal-directed behavior (hand reaching). Since foveal vision is a limited resource, an intelligent, task-dependent allocation is necessary.

When trying to predict the eye's motion, a common null hypothesis is to assume the gaze is directed to visually-salient features of the scene (Koch and Ullman 1985). However, here we consider tasks whose goal is the motion of the hand, with the eye playing a supportive role. In this case, the assumption that only image heuristics (such as saliency) account for the eye's movement seems unlikely. The top-down effects of the motor task on visual behavior are an active area of study (Peters and Itti 2008; Rothkopf, Ballard, and Hayhoe 2007). Here we step away from neuroscientific investigation of visual processing, and abstract the eye's effect as a localized reduction in observation noise (section 3.2). This allows us to take a broader perspective on tasks which require coordination of information-seeking behaviors.

Todorov and Jordan (2002) propose the paradigm of optimal control as a framework for the study of motor coordination, and normative models (Körding 2007) harness simulation and numerical optimization for the study of neural mechanisms. Algorithms of optimal control often rely on the principle of *certainty equivalence* (Stengel 1994), which posits a separation between estimation and control: given an estimate of the current state of the world, the best action can be identified by considering a deterministic, fully-observable system in the same state. This separation allows for efficient computation, because optimal control can focus only on the deterministic system, and estimation can be safely ignored during motion planning. However, this separation does not hold in tasks which optimize information-seeking behavior.

In order to model the coupling between perception and action inherent to eye-hand coordination, we model this domain as a continuous-state partially-observable Markov decision processes (POMDP) (Sondik 1971; Kaelbling, Littman, and Cassandra 1998). The POMDP framework is designed to tackle domains with state uncertainty, and allows us to consider goal-directed actions and task-relevant information-pickup in a single optimization problem. The continuous POMDP we use here has 17 continuous state dimensions, and it is solved using the deterministic belief update heuristic (section 2.2).

The domain of eye-hand coordination exhibits a large degree of task-specific behavioral diversity (Carpenter 1977), e.g., saccades vs. smooth pursuit. This is a modeling challenge, because a candidate model must allow for the emergence of a variety of possible solutions, depending on the specific instantiation of task parameters. Our POMDP model meets this requirement, as different parameter settings generate different motion patterns. In section 4 we discuss the role of the various parameters in shaping the resulting behavior.

Our model finds an optimal motion pattern of the hands and the eye, allowing for the emergence of coordination

from first principles, without imposing hand-crafted heuristics. In order to test the model, we conducted a behavioral experiment, in which human subjects perform the same eye-hand coordination task. The experiment and results are described in section 5; we find that simulation and experimental results are mostly congruent, demonstrating the validity of the model.

2 Solving high-dimensional, continuous POMDPs

In POMDP terminology, the agent is said to occupy a *belief* state, which is a distribution over all possible states, representing the agent's ambiguous sense of the world. For the most part, the POMDP literature focuses on finding globally-optimal solutions for discrete domains. Previous studies of continuous POMDPs (Porta et al. 2006; Brunskill et al. 2008; Brooks 2009) focus on finding a globally-optimal solution for domains with only one or two dimensions. Several recent studies (Prentice and Roy 2009; Platt et al. 2010; Miller, Harris, and Chong 2009) propose an alternative approach to continuous POMDP optimization, finding the optimal behavior by deterministically planning optimal trajectories through belief-space. In this paper, we use the Nominal-Belief heuristic (Miller, Harris, and Chong 2009), replacing the stochastic observation with its maximum-likelihood counterpart (Platt et al. 2010) during planning. The resulting belief dynamics are deterministic, and therefore amenable to efficient optimization algorithms. However, since planning takes place in the belief domain, the optimization still accounts for the state's ambiguity (as this information is represented by the various belief states). Therefore, the resulting behavior strikes a balance between information-seeking and goal-directed action, despite the marginalization of the stochastic processes.

2.1 Definitions

Formally speaking, we consider a discrete-time POMDP defined by a tuple $\langle S, A, Z, T, \Omega, R, N\rangle$, where: S, A and Z are the state space, action space and observation space, respectively; $T(s', s, a) = \Pr(s'|s, a)$ is a transition function describing the probability of the next state given the current state and action; $\Omega(z, s, a) = \Pr(z|s, a)$ is the observation function, describing the probability of an observation given the current state and action; and $R(s, a)$ is a reward function, and a terminal reward $R^N(s)$. In this paper we consider an undiscounted optimality criterion, where the agent's goal is to maximize the expected cumulative reward within a fixed time horizon N.

The *belief state* $b \in B$ is a probability distribution over S, where $b^i(s)$ is the likelihood of the true state being s at time i. The reward associated with a belief is simply the expected value over this state distribution:

$$R^i(b, a) = \mathop{\mathrm{E}}_{s\sim b}\left[R^i(s, a)\right]. \quad (1)$$

Given the current belief b, an action a and observation z, the updated belief b' can be calculated by applying Bayes's rule. However, in the continuous case B is infinite-dimensional, and therefore the belief update must be approximated by some estimation filter.

2.2 The deterministic belief update heuristic

We study continuous stochastic dynamics of the form $ds = f(s, a)dt + q(s, a)d\zeta$, where ζ represents continuous-time Brownian motion. For a given state s and action a, integrating this continuous dynamics over a small time-step τ results in a normal distribution over the next state s': $T(s', s, a) = \mathcal{N}(s'|F(s, a), Q(s, a))$, where the mean is propagated with the Euler integration

$$F(s, a) = s + \tau f(s, a), \quad (2)$$

and the covariance $Q = \tau q^{\mathsf{T}} q$ is a time-scaling of the continuous stochastic process $qd\zeta$.

Similarly, we focus on observation distributions of the form $\Omega(z, s, a) = \mathcal{N}(z|w(s), W(s, a))$, where w is a deterministic observation function, and W describes how the current state and action affect the observation noise.

Given a Gaussian prior on the initial state, we approximate the infinite-dimensional b by a single Gaussian: $\hat{b}(s) = \mathcal{N}(s|\hat{s}, \Sigma)$, where the covariance Σ belongs to the space of symmetric, positive-semidefinite matrices $\mathcal{M} \subset \mathbb{R}^{n\times n}$. Therefore, the belief space $\hat{B}$ is parameterized in this case by the product space $\nu \in S \times \mathcal{M}$. In the limit of $\tau \to 0$, and given a Gaussian prior, this approximation is accurate.[1]

In order to approximate the belief update, we use the Extended Kalman Filter (EKF) (Stengel 1994). Given the current belief $\hat{b}$, action a and observation z, we calculate the partial derivatives of the dynamics and the observation functions around $\hat{s}$: $w_s = \partial w/\partial s$ and $F_s = \partial F/\partial s$. We find the uncorrected estimation uncertainty $H = F_s \Sigma F_s^{\mathsf{T}} + Q(\hat{s}, a)$ and calculate the new mean $\hat{s}'$ by the innovation process:

$$\hat{s}' = F(\hat{s}, a) - K(z - w(\hat{s})), \quad (3)$$

where $K = H w_s (w_s^{\mathsf{T}} H w_s + W(\hat{s}, a))^{-1}$ is the *Kalman gain*. Finally, the new covariance Σ' is given by:

$$\Psi(\hat{s}, \Sigma, a) = H - H w_s (w_s^{\mathsf{T}} H w_s + W(\hat{s}, a))^{-1} w_s^{\mathsf{T}} H^{\mathsf{T}}. \quad (4)$$

The deterministic belief update is obtained by taking the expectation of equations (3) and (4) with respect to the observation variable z. Since equation (3) is linear in z, we can replace z with its mean $w(\hat{s})$, causing the second term of equation (3) to vanish. Therefore, the maximum-likelihood estimate of the next belief's mean is (2) (the deterministic dynamics). By virtue of the EKF being a first-order filter, the calculation in (4) is independent of z, and so the next belief's covariance is the same, regardless of the value of z. In summary, the maximum-likelihood esimate for the next belief is formed by by combining (2) and (4): $\hat{b}'(s) = \mathcal{N}\big(s|F(\hat{s}, a), \Psi(\hat{s}, \Sigma, a)\big)$.

[1] Note that this single-Gaussian approximation fails when the domain includes discontinuities and unilateral constraints (e.g., joint limit constraints, or confining walls); in such cases, a different formulation of the belief is needed (Erez and Smart 2010).

For further discussion of the deterministic belief update heuristic, the interested reader is referred to the work of Platt et al. (Platt et al. 2010) and Miller et al. (Miller, Harris, and Chong 2009).

2.3 Planning in the belief domain

The belief update heuristic of the previous section (together with equation 1) define a problem of deterministic optimal control in a high-dimensional continuous space, with non-linear dynamics and non-quadratic reward. To find a locally-optimal solution, we may use trajectory optimization; here we use Differential Dynamic Programming (DDP), a second-order algorithm that has been successfully applied to real-world high-dimensional, non-linear control domains (Abbeel and Ng 2005).DDP represents the policy as an open-loop sequence of actions, together with locally-linear feedback controllers at every state. The interested reader may find an in-depth description of the algorithm in the book by Jacobson and Mayne (1970).

3 A POMDP model of eye-hand coordination

We model the position of the gaze target in the frontal plane, as well as the position of the hands' end-effectors in that plane. In addition, the model's state includes the (a priori unknown) planar positions of a target and four obstacles. The agent's task is to guide both hands to the target while passing between a pair of obstacles, using the eye's gaze to locally disambiguate portions of the scene (figure 1(a)). By solving for an optimal motion plan, we generate a coordinated movement of hands and eye through the scene. The model and the resulting motion are best illustrated by a short movie, submitted as supplementary material.

3.1 State, action, and transition function

The system's state is the concatenation of the planar positions of the agent's state and the environment's state. The agent has kinematic control of the hands (s_{h_1} and s_{h_2}) and gaze fixation point (s_e); the agent's state also includes a scalar time-lag variable (s_k), which measures the time from the last saccade (section 3.2). The environment's state specifies the positions of the target (s_T) and obstacles ($s_{b_i}, i = 1 \dots 4$). This leads to a 17-dimensional state space.

The state update is subject to process noise whose magnitude is constant but varies between the elements of the scene. For instance, the state update equation for the target position is:

$$s'_T = s_T + \sigma_c \xi,$$

with ξ being a two-element vector of zero-mean normally-distributed random variables with unit variance, and σ_c being the model parameter that scales the process noise for the target and obstacles (see discussion in section 4.2).

The agent controls six continuous action variables, specifying the displacement of the hands and eye at every time step. For example, the state update equation for the hand's position is:

$$s'_{h_1} = s_{h_1} + a_{h_1} + \sigma_h \xi,$$

where σ_h is a model parameter that sets the magnitude of the process noise affecting the hand; the same equation holds for s_{h_2} and s_e, and their corresponding controls. We eliminate any uncertainty in the eye's position by not subjecting it to process noise ($\sigma_e = 0$).

It is known that the brain's processing of visual information is impaired (even if not completely inhibited) during a saccade. Furthermore, it has been shown (Thorpe, Fize, and Marlot 1996) that even after the eye's gaze settles on the new target, it takes some time before visual information is available for some tasks. We model this effect with an auxiliary state variable s_k, which measures the time elapsed from the last saccade. This variable integrates linearly when the eye's velocity is zero, and becomes very low otherwise:

$$s'_k = \tau + \frac{s_k}{1 + \alpha \|a_e\|}, \tag{5}$$

where τ is the time-step length, and α is some large coefficient (we use $\alpha = 1000$). Like the eye's position, this auxiliary variable is also not subject to any process noise, and therefore needs no estimation.

3.2 Observation

In a POMDP, the agent receives stochastic observations, through which it infers the true state of the system. Here, the agent may observe the state of all the scene's elements, and the observation of every element's position is normally-distributed around the true underlying value; the agent has full observation of the eye's position, as well as the time since the last saccade.

The covariance of the observation of each scene element depends on its position relative to the eye's gaze. Let $d_* = \|s_* - s_e\|$ be the Euclidean (planar) distance between a scene element and the eye's gaze point ($*$ standing for either a hand, a target or an obstacle), and let $g(d)$ be a function that scales the width of this element's observation distribution; this function models foveated vision, so it is low around $d = 0$, and high farther away. We chose to model g as a sigmoid: given a width parameter η and a slope parameter l, we look at the scaled distance

$$\tilde{d}(s_*, s_e) = (\|s_* - s_e\| - \eta)/l_d,$$

and compute:

$$g(\tilde{d}) = \sigma_o \Big(0.5 + \frac{\tilde{d}}{2\sqrt{\tilde{d}^2 + 1}}\Big), \tag{6}$$

with σ_o being the maximal observation covariance due to peripheral vision (see discussion in section 4). The width of the human fovea is about 2 degrees; in our experimental setup, this translates to 7% of the scene's width, and so we set $\eta = 0.035$, and $l_d = 0.005$.

Proprioception. Proprioception is an independent channel of observation for the hands' positions. We model this by capping the covariance of the hands' observation (as given by equation (8)) by a fixed value σ_p which is the observation noise of proprioception.

Modeling post-saccadic perceptual delay. In order to model the time delay due to processing of visual information, we added a term to the observation's covariance that

is sigmoidal in the elapsed time since the last saccade. We scale the elapsed time s_k: $\tilde{t} = (\mu - s_k)/l_k$, and compute

$$k(s_k) = \alpha\Big(0.5 + \frac{\tilde{t}}{2\sqrt{\tilde{t}^2+1}}\Big). \quad (7)$$

We set $\mu = 150$ms and $l_k = 30$ms following Thorpe, Fize, and Marlot (1996), and α is some large coefficient (we use $\alpha = 1000$). Therefore, the observation's covariance remains high for the first 120 ms after a saccade; at 150 ms this term drops, allowing the elements within the fovea to be observed accurately.

In conclusion, the distribution of the observation for any scene element $*$ is:

$$\omega(o_*|s_*, s_h, s_k) = \mathcal{N}\Big(s_*, \mathbf{I}_2 \cdot \max(\sigma_p, k(s_k) + g\big(\tilde{d}(s_*, s_h)\big)\Big) \quad (8)$$

where the max operation models the integration of proprioceptive and visual observation, and is therefore dropped for the obstacles and target.

While this model makes use of a relatively-large set of parameters $(\sigma_h, \sigma_c, \sigma_p, \sigma_o)$, this allows us to generate a diverse set of motion patterns as optimal solutions in different parameter settings (see discussion in section 4).

3.3 The reward function

The reward function has two parts: running reward and terminal reward. The terminal reward is a penalty (cost) that is quadratic in the distance of the hands from the target. The running reward has two components: action cost and obstacle cost. The action cost is quadratic in the displacement of the hands and the eye at every time step. The obstacle cost is modeled as a delta-function: the agent is penalized if the hand's position is equal to an obstacle's position. Overall, the running reward function is:

$$R(s, a) = -a^{\mathsf{T}} M a - \sum_{i=1}^{2}\sum_{j=1}^{4} \delta(s_{h_i}, s_{b_j}) \quad (9)$$

where δ is the delta function, and M is a diagonal matrix assigning weight 1 to the hands' displacement, and the weight 0.01 to the eye's displacement.

The only non-quadratic term is the obstacle cost. In a deterministic system, the use of a delta-function would yield unrealistic motion, as the hand might pass at epsilon distance from the obstacle. However, since we plan in belief space, we use the belief cost function (1). The obstacle-related belief cost is the expectation of this delta-function under the normal distribution describing the joint uncertainty in the positions of the hand and the obstacle. Therefore, the belief-cost due to the obstacles is a Gaussian, which can be computed in closed form.

4 Exploring the model's behavioral diversity

As mentioned in section 1, the domain of eye-hand coordination exhibits a diverse array of behaviors, and a candidate model should allow for an equally-diverse array of solutions. Our model employs three categories of parameters: the coefficients of the cost function, the coefficients describing the process noise, and parameters that determine the state-dependent observation noise. In this section, we discuss the model's parameters, and the behavioral diversity they enable. This behavioral diversity is demonstrated (in part) in the movie submitted as supplementary material.

4.1 The cost function parameters

Our simulations suggest that the cost parameters are simple to choose, as the wrong value often leads to absurd behavior: for example, if the target cost is not big enough compared to the hands' action cost, the hands do not reach the target; increasing the target cost fixes the problem, and increasing it even more makes no difference (because the residual is already zero). The obstacle cost leads to interesting behavioral diversity — when it is very big (relative to the hand action cost), the hands increase their velocity as they pass between the obstacles, and if it is very small, the hands choose a shorter path that gets dangerously close to the obstacle. Such behaviors are conceivable in certain scenarios, but the human subjects in our experiment (section 5) exhibited neither.

4.2 Process noise

While the observation noise is state-dependent, the process noise in equation (2) is constant. Two values need to be determined — σ_h, the process noise associated with the hands, and σ_c, the process noise associated with the static elements of the scene (target and obstacles). When the hands are not subject to process noise, the eye only looks at the obstacles, and when the obstacles are not subject to noise, the eye's gaze shifts preemptively (before the hand reaches the obstacles), as the required disambiguation took place and there is no risk of accumulating new uncertainty.

In contrast, when the hands are subject to process noise, the eye performs smooth pursuit, following each hand along its path as it approaches the obstacles. This allows the agent to be certain of the hand's position at the mission-critical moment, when it is near the obstacles. Similarly, applying process noise to the static scene elements causes the eye's gaze to remain fixed on the obstacles until the hand passed through the obstacle pair, so as to ensure that no uncertainty accumulates before the hand reaches the obstacle.

Another two parameters of behavioral relevance are σ_p (the maximal observation covariance of the hands' positions) and σ_o (the maximal observation covariance of the static scene elements). When the first is small, no smooth pursuit will emerge, as the agent has a reliable source of information of the hands' positions. When the second is small, saccades will be inhibited, as peripheral vision provides observations that are good enough.

5 Comparing the model's solution to human subjects' behavior

In order to conflate simulation results with human motor behavior, we tested six subjects in an eye-hand coordination task. Subjects were equipped with two controllers to move a left and right virtual hand in a two-dimensional scene projected on a large screen (figure 1(a)). Subjects were asked

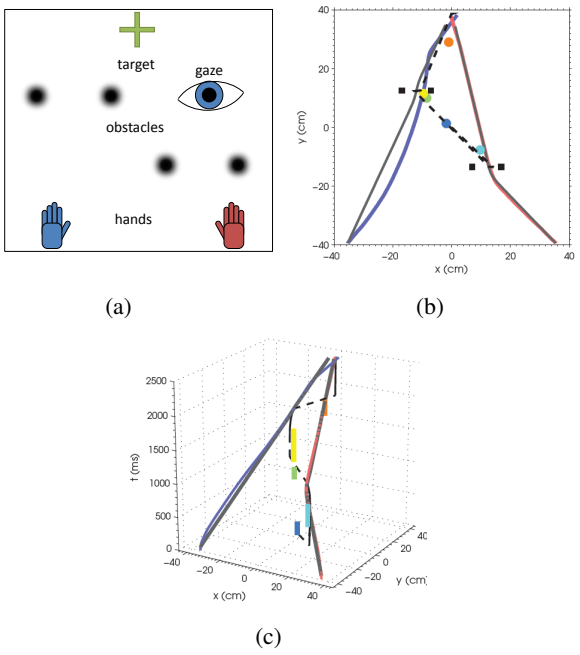

Figure 1: **(a)** The agent (model or subject) is required to bring the hands from their start positions (at the bottom) to the target (top center) while passing in between a pair of obstacles. **(b)** Gaze fixations (colored dots) and left/right hand trajectories (blue/red line) of subject S5 for a single trial, plotted in a two-dimensional plane. The obstacles are depicted by black squares. The dashed and solid lines represent the model's solution for gaze and hand trajectories, respectively. **(c)** Same data adding the dimension time, such that fixations at a particular position become bars.

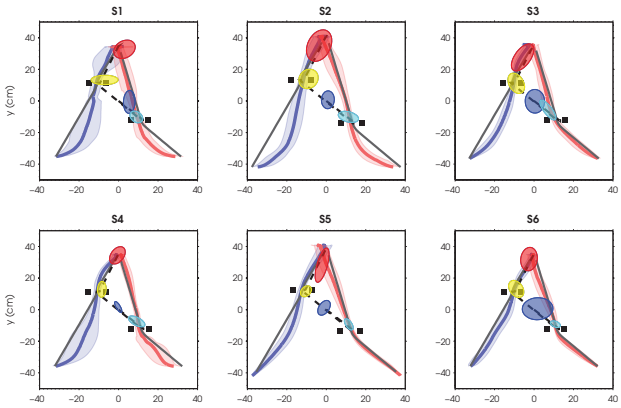

Figure 2: Trial averages of gaze fixations and hand trajectories for all subjects for the same scene as in figure 1. Clusters of gaze fixations are represented by colored ellipses (two SD). Hand trajectory is represented by the trial average (solid line) and the variability (shaded area, size of one SD) for left (blue) and right (red) hand.

to move the left and right virtual hand through two narrow gates (formed by a pair of obstacles) toward a final common goal within 3 seconds. The position of the two gates changed at every trial, randomly rotating between 12 different arrangements (each scene was repeated thirty times). The scene was presented to the subjects at a field-of-view of 45 degrees. Before each trial, subjects had to fixate at the center of the screen. No instructions were given regarding eye movements during a trial. Eye movements were recorded using the double magnetic induction (DMI) method (Bour et al. 1984).

Figure 1(b-c) shows the results of subject S5 for a single trial. Panel (b) shows the xy-view, with x the horizontal and y the vertical direction. The blue and red line represents the subject's trajectory of the left and right hand, respectively. Gaze fixations are represented by colored dots. In addition, we plotted the hand trajectories (gray lines) and gaze trajectory (black dashed line) as solved by the model. Panel (c) shows the same data, adding the dimension time. Here, gaze fixations are represented by colored bars.

In this trial, the subject's gaze saccades from the initial position (dark blue fixation point) toward the right obstacle pair (light blue) to assist the right hand in passing the obstacles. Before the right hand has reached the obstacle pair, gaze already jumps to the left obstacle pair. Note that this jump consist of one large saccade (from the light blue to the green fixation point) followed by a small, so-called correction saccade (from green to yellow) which is a well known phenomenon for large saccades (Carpenter 1977). Before the right hand approaches the obstacles, a saccade is made toward the target to guide both hands to the final goal.

In order to test the predictions of the POMDP model, we compared the model predictions on three criteria: (1) the order and location of fixations, (2) the trajectory of left and right hand, and (3) the relative timing of gaze and hands. Figure 2 shows the trial average of all subjects for the same scene. These responses are representative for the behavior in other scenes. The variability in hand position is represented by the shaded area (one standard deviation). Clusters of gaze fixations are represented by colored ellipses (two standard deviations). For all subjects, the order of fixations (i.e., initial fixation, right obstacle pair, left obstacle pair, target) is in agreement with the model predictions. This may not be surprising, since the order is determined by the spatial location of the obstacles: the right hand arrives earlier at the obstacles compared to the left hand. However, the correct order of fixations was well predicted by the model, irrespective of the location of the pair of obstacles.

The subjects fixate approximately between an obstacle pair, in agreement with the model predictions. Note that the precise fixation location differs slightly between subject and model. Some subjects make slightly smaller saccades than the model. This undershoot is a known phenomenon (Carpenter 1977) and is thought to reflect properties of the human saccadic system, which were not included in our model.

The right hand's trajectory predicted by the model matches the measured path for subjects S3, S5 and S6 almost perfectly, whereas the left hand shows a slightly curved trajectory. For the other subjects, both hands reach the target via a more curved trajectory than predicted by the model. A

control experiment revealed that even in absence of any obstacles, subjects tend to move the hands in an inward-curved trajectory whereas the model predicts a straight line (i.e., the shortest path). In that case subjects fixate at positions in the middle, between both hands. A plausible explanation for this behavior is that subjects use their peripheral vision to guide both hands to the target, relying on the heightened capacity to perceive motion through peripheral vision (McKee and Nakayama 1984) (a feature which we did not try to model).

The model not only correctly predicts the spatial location of eye fixations and hand positions, but also the timing of gaze relative to hand position. This is illustrated in figure 1 (right panel), which shows that gaze jumps from the initial gaze position at the start of each trial to the right obstacle pair, and from there to the left pair of obstacles, approximately 200 ms before the right hand has reached the obstacles. Similarly, the gaze jumps from the left pair of obstacles about 200 ms before the left hand reaches the left pair of obstacles. By tuning the value of the parameter describing the obstacles' process noise, we recover this temporal pattern in our model.

6 Conclusion

This paper presents a POMDP model of hand-eye coordination. While we demonstrate that the optimal solution is congruent with the behavior of human subjects performing the same task, it is important to note that we do not argue for the biological plausibility of the computational techniques; instead, this normative model may allow us to test our understanding of what (we believe) the brain "should" do.

Experiments of eye-hand coordination tasks yield a diverse set of behaviors, according to the particular experimental setup. Our model captures this feature, as it can produce qualitatively-different behavior when the values of these parameters change (as explained in section 4). Here, we demonstrate how the model can be congruent with one particular experimental setup; this naturally guided our choice of parameter values to a particular region.

While we could perform further tuning of the parameters (whether by hand or through learning), we forego that in favor of a broader message: even with minimal assumptions on the actual parametric forms and numerical values, the solution of this POMDP model is in qualitative agreement with experimental data. The quantitative accuracy of the model's predictions will be the goal of future studies.

References

Abbeel, P., and Ng, A. Y. 2005. Exploration and apprenticeship learning in reinforcement learning. In *International Conference on Machine Learning (ICML)*, 1–8.

Bour, L.; van Gisbergen, J.; Bruijns, J.; and Ottes, F. 1984. The double magnetic induction method for measuring eye movement–results in monkey and man. *IEEE Trans Biomed Eng* 31(35):419–427.

Brooks, A. 2009. *Parametric POMDPs*. VDM Verlag.

Brunskill, E.; Kaelbling, L.; Lozano-Perez, T.; and Roy, N. 2008. Continuous-state POMDPs with hybrid dynamics. In *Tenth International Symposium on Artificial Intelligence and Mathematics (ISAIM)*.

Carpenter, R. 1977. *Movements of the eyes*. Pion Limited, London.

Erez, T., and Smart, W. D. 2010. A scalable method for solving high-dimensional continuous pomdps using local approximation. In *Proceedings of the 26th Conference in Uncertainty in Artificial Intelligence (UAI)*.

Jacobson, D. H., and Mayne, D. Q. 1970. *Differential Dynamic Programming*. Elsevier.

Kaelbling, L. P.; Littman, M. L.; and Cassandra, A. R. 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101(1-2):99–134.

Koch, C., and Ullman, S. 1985. Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology* 4:219–227.

Körding, K. 2007. Decision theory: What "Should" the nervous system do? *Science* 318(5850):606 –610.

McKee, S., and Nakayama, K. 1984. The detection of motion in the peripheral visual field. *Vision Research* 24(1):25–32.

Miller, S. A.; Harris, Z. A.; and Chong, E. K. P. 2009. A POMDP framework for coordinated guidance of autonomous uavs for multitarget tracking. *EURASIP Journal of Advanced Signal Process* 2009:1–17.

Peters, R. J., and Itti, L. 2008. Applying computational tools to predict gaze direction in interactive visual environments. *ACM Transactions on Applied Perception* 5(2):1–19.

Platt, R.; Tedrake, R.; Kaelbling, L. P.; and Lozano-Perez, T. 2010. Belief space planning assuming maximum likelihood observations. In *Robotics: Science and Systems (R:SS)*.

Porta, J. M.; Vlassis, N.; Spaan, M. T.; and Poupart, P. 2006. Point-based value iteration for continuous POMDPs. *Journal of Machine Learning Research* 7:2329–2367.

Prentice, S., and Roy, N. 2009. The belief roadmap: Efficient planning in belief space by factoring the covariance. *The International Journal of Robotics Research* 28(11-12):1448–1465.

Rothkopf, C.; Ballard, D.; and Hayhoe, M. 2007. Task and context determine where you look. *Journal of Vision* 7(14).

Sondik, E. 1971. *The Optimal Control of Partially Observable Markov Processes*. Ph.D. Dissertation, Stanford.

Stengel, R. F. 1994. *Optimal Control and Estimation*. Dover Publications.

Thorpe, S.; Fize, D.; and Marlot, C. 1996. Speed of processing in the human visual system. *Nature* 381:520–522.

Todorov, E., and Jordan, M. I. 2002. Optimal feedback control as a theory of motor coordination. *Nature Neuroscience* 5(11):1226–1235.