Text S1

In text S1, we provide details for the stochastic optimal control framework used to model reach and eye movements.

1 Hand modeling.

To preserve interpretability, it is important to use a model no more complex than absolutely necessary. In the current study, we modeled the dynamics of the hand as a “point of mass” (m = 1 Kg), in Cartesian coordinates, with 2-dimensional position $p^h_t = [x^h(t), y^h(t)]^T$ and velocity $\dot{p}^h_t = [\dot{x}^h(t), \dot{y}^h(t)]^T$. The combined actions of all muscles is represented by the force vector $f_t = [f_x(t), f_y(t)]^T$. The motor command $u_t$ is transformed into these forces $f_t$ through second-order muscle-like low-pass filters with constants $\tau_1$ and $\tau_2$, and by adding control-dependent multiplicative noise $\epsilon$.

$$\tau_1 \tau_2 \ddot{f}_t + (\tau_1 + \tau_2) \dot{f}_t + f_t = u_t$$  \hspace{1cm} (1)

The second-order low pass filters can be written as a pair of coupled first-order filters with outputs $g$ and $f$, as:

$$\tau_1 \dot{g}_t + g_t = u_t, \hspace{0.5cm} \tau_2 \dot{f}_t + f_t = g_t$$  \hspace{1cm} (2)

Given that $p^G = [x^G, y^G]$ is the position that the hand is planned to arrive, at the end of the movement (i.e., the current goal of the task), the discrete-time state is described by the following $10^{th}$-dimensional vector $x_t$. 

1
The discrete-dynamics of the hand are given by Eq. (4):

\[
\begin{align*}
    p_{t+\delta t}^h &= p_t^h + \dot{p}_t^h \delta t \\
    \dot{p}_{t+\delta t}^h &= \dot{p}_t^h + f_t \frac{\delta t}{m} \\
    f_{t+\delta t} &= f_t \left( 1 - \frac{\delta t}{\tau_2} \right) + g_t \frac{\delta t}{\tau_2} \\
    g_{t+\delta t} &= g_t \left( 1 - \frac{\delta t}{\tau_1} \right) + u_t \left( 1 + \sigma_c \epsilon_t \right) \frac{\delta t}{\tau_1}
\end{align*}
\]  

(4)

where \( \delta t = 0.01 s \) is the sampling period of discretization. In the current framework, we assume that the motor control signal \( u_t \) is corrupted by multiplicative noise described by the product term \( \sigma_c \epsilon_t u_t \), with \( \sigma_c = 1 \), which is a unitless variable defined as the noise magnitude related to the control signal magnitude.  

2. Reaching to a single goal.

In the current study, each controller is linked with a neuron \( i \) from the “motor plan formation field”. The purpose of the controller is to generate an optimal policy that drives the hand to a distance \( r \) from the current location, towards the preferred direction \( \phi \) of the neuron \( i \). Within the stochastic
optimal control framework, control policies originate as the solution of an optimization problem.

The basic idea is to find a sequence of motor commands that acquire as much reward as possible, while spending as little effort as possible. For reaching tasks, given the kinematics of the hand and the sensory and motor noise in estimating and controlling the state of the hand, the stochastic optimal control framework finds an optimal policy $u^*$, for time instances $t = [t_1, t_2, \cdots t_T]$ that optimizes the cost function $J$ described in Eq. (5), for each state of the hand and the environment $x_t$.

$$J(x_t, \pi) = \|p^h_T - p^G\|^2 + \|\dot{p}^h_T\|^2 + \|f_T\|^2 + \sum_{t=1}^{T} \pi(x_t)^T R\pi(x_t)$$

The first 3 terms define that current task of the controller - i.e., drive the hand to the position $p^G = [rcos(\phi) r sin(\phi)]$ (first term) and stop there (second and third term). The last term is the motor command cost (i.e., action cost) that penalizes the effort required to arrive at the position $p^G$ after time $T$ starting from the current state. The matrix $R = \frac{1}{T} \begin{bmatrix} r_x & 0 \\ 0 & r_y \end{bmatrix}$ is the control-dependent cost of the hand motion in the x and y dimension.

We can write Eq. (5) in the general form of optimal control cost function as follows:

$$J = (x_T - Sp^G)^T Q_T (x_T - Sp^G) + \sum_{t=1}^{T-1} \pi(x_t)^T R\pi(x_t)$$
where $S$ matrix picks out the actual position of the hand and the current goal-position $p^G$ at the end of the movement, from the state-vector. The time-varying matrix $Q_T$ describes the state-dependent cost and is the zero matrix for any time $t < T$ and is equal to the Hessian matrix of the cost function evaluated at the end of the movement $T$.

To minimize the cost function in Eq. (6), a model of the system dynamics and sensory feedback must be incorporated. A plethora of experimental studies provide evidence that the sensory system uses an internal forward model to predict the next state of the system at time $t + 1$, $\hat{x}_{t+1|t}$, based on the sensory feedback $y_t$, the current state estimate $\hat{x}_t$ and the control commands $u_t^2$. This prediction is necessary to overcome control instabilities due to noisy sensors and temporal delays.

In the current study, we modeled the hand and the state space using linear dynamics and measurement as a discrete linear system, Eq. (7), considering that the motor commands are corrupted by multiplicative noise, normally distributed with zero mean and standard deviation proportional to the magnitude of the control commands and the state variables $^1$.

$$
x_{t+1} = Ax_t + Bu_t + \xi_t + C(u_t)\epsilon_t
$$

$$
y_t = Hx_t + \omega_t
$$

(7)

where $A$, $B$ and $H$ are the actual system dynamics and observation matrices, respectively.
The noise variables $\xi_t$, $\omega_t$ and $\epsilon_t$ are normally distributed variables with zero mean and covariance $\Omega^\xi \geq 0$, $\Omega^\omega \geq 0$ and $\Omega^\epsilon = I$, respectively. $C(u_t)$ is a scaling matrix for control-dependent
system noise, such as \( C(u)\epsilon = \sum_i C_i u\epsilon_i \), where \( \epsilon_i \) is the \( i^{th} \) component of the random variable \( \epsilon \).

Given the belief about the state at time \( t \) and the current task of the controller, the controller will suggest an optimal policy \( \pi^*(x_t) = u^* \) that minimizes the expected cost function in Eq. (5). This form of optimal control is a modified version of the Linear Quadratic Gaussian (LQG) regulator, since the dynamics of the system are linear, the expected cost function is quadratic and the noise is Gaussian, but with signal-dependent noise \(^1\).

The optimal policy is incorporated into the system model to generate a feedback controller that uses its forward model to make predictions \( \hat{y}_t \) from knowledge of controls, dynamics and sensory measurements. These predictions are combined with actual sensory feedback \( y_t \) using Eq. (9) to update the belief about the state in time \( t + 1 \).

\[
\hat{x}_{t+1|t+1} = (A - BL_t)\hat{x}_{t+1|t} + K_t (y_t - H\hat{x}_{t+1|t})
\]  

(9)

where \( K_t \) is the Kalman gain at time \( t \).

### 3 Eye modeling

In the current study, we modeled the globe and the surrounding tissues of the eye by an elastic element with stiffness \( k \) that pulls the eye away from the equilibrium point \( (x = 0) \), and a viscous element \( \mu \) that resists in this motion (for more details about this eye model see \(^3\)). The one-dimensional dynamics of the eye is given by Eq. (10).

\[
m\ddot{x} = -kx - \mu\dot{x} + u
\]

(10)
where $m$ is the inertia of the eye and $u$ are the motor commands.

We used a simple model to translate motor commands $u$ into forces: $\alpha_1 \dot{f} + \alpha_2 f = u$, where $f$ describes the instantaneous force generated by the extra-ocular muscles to move the eye.

Let’s assume that the current goal is to move the eye at the position $p^G = [x_G, y_G]^T$. The discrete-time state is given by the 8th-dimensional vector $x_t$, Eq. (11).

$$x_t = [x_e(t), \dot{x}_e(t), y_h(t), \dot{y}_e(t), f_x, f_y, x_G, y_G]^T$$ (11)

We extended this one-dimensional model of the eye to a more realistic two-dimensional model, in which the discrete-time dynamics of the eye are given by Eq. (12):

$$p_{t+\delta t}^e = p_t^e + \dot{p}_t^e \delta t$$
$$\dot{p}_{t+\delta t}^e = -\delta t \frac{k}{m} p_t^e + \left(1 - \delta t \frac{\mu}{m}\right) \dot{p}_t^e + \frac{\delta t}{m} f_t$$
$$f_{t+\delta t} = \left(1 - \delta t \frac{\alpha_2}{\alpha_1}\right) f_t + u_t (1 + \sigma_c \epsilon_t) \frac{\delta t}{\alpha_1}$$ (12)

where $p_t^e = [x_e(t), y_e(t)]^T$ and $\dot{p}_t^e = [\dot{x}_e(t), \dot{y}_e(t)]^T$ is the 2-dimensional position and velocity of the eye, respectively. Similarly, with the hand model, the product term $\sigma_c \epsilon_t u_t$ describes the multiplicative noise added to the control signal $u_t$, with $\sigma_c = 1$ and $\epsilon_t$ a vector of zero-mean random variables with covariance $\Omega^c = I$. Additionally, $\delta t = 0.001s$ is the sampling period of discretization.
4 Saccade to a single goal

Even though saccades are brief in time and sensory information seems to have no role in the control of eye movements, recent findings suggest that saccades are not open-loop movements. Instead, the motor commands that generate saccades benefit from a forward internal model that monitors the motor commands and predicts their consequences.4

We modeled eye movements within the stochastic optimal control framework using a similar approach with the hand movements presented in a previous section. However, the difference with eye movements is that the sensory feedback does not affect the control of the eye movements. Each controller is linked to a neuron $i$ from the motor plan formation field, and generates a policy that drives the eye towards the preferred direction $\phi$ of the neuron $i$. The cost function is the same as the one used in reaching movements (see Eq. (5)).

The dynamics of the eye can be transformed into the form of Eq. (7), with the following matrices:
\[ A = \begin{bmatrix}
1 & \delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & \delta t & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & \frac{\delta t}{m} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{\delta t}{m} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

\[ B = \begin{bmatrix}
0_{4 \times 2} \\
\frac{\delta t}{\alpha_1} & 0 \\
0 & \frac{\delta t}{\alpha_1} \\
0_{2 \times 2} \\
\end{bmatrix}
\]

For a monkey eye, we used time constants \( \tau_1 = 0.260 \) s and \( \tau_2 = 0.012 \) s as proposed by Keller \(^5\) and Robinson \(^6\). These time constants are related with the constants of the system dynamic matrices \( A \) and \( B \) as follows: \( \mu = \tau_1 + \tau_2, \ m = \tau_1 \tau_2 \). Based on these studies \(^5,6\), we set \( k = 1, \ \alpha_1 = 0.004 \) and \( \alpha_2 = 1 \).
References

1. Todorov, E. *Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system*, Neural Comput., 17, 1084-1108 (2005)


