J Neurophysiol 119: 1795–1808, 2018.
First published January 31, 2018; doi:10.1152/jn.00502.2017.

REVIEW | *Control of Coordinated Movements*

Modeling auditory-visual evoked eye-head gaze shifts in dynamic multisteps

Bahadir Kasap and A. John van Opstal
Radboud University, Donders Institute for Brain, Cognition and Behavior, Department of Biophysics, Nijmegen, the Netherlands

Submitted 5 July 2017; accepted in final form 30 January 2018

Kasap B, van Opstal AJ. Modeling auditory-visual evoked eye-head gaze shifts in dynamic multisteps. *J Neurophysiol* 119: 1795–1808, 2018. First published January 31, 2018; doi:10.1152/jn.00502.2017.—In dynamic visual or auditory gaze double-steps, a brief target flash or sound burst is presented in midflight of an ongoing eye-head gaze shift. Behavioral experiments in humans and monkeys have indicated that the subsequent eye and head movements to the target are goal-directed, regardless of stimulus timing, first gaze shift characteristics, and initial conditions. This remarkable behavior requires that the gaze-control system *1*) has continuous access to accurate signals about eye-in-head position and ongoing eye-head movements, *2*) that it accounts for different internal signal delays, and *3*) that it is able to update the retinal (T_E) and head-centric (T_H) target coordinates into appropriate eye-centered and head-centered motor commands on millisecond time scales. As predictive, feedforward remapping of targets cannot account for this behavior, we propose that targets are transformed and stored into a stable reference frame as soon as their sensory information becomes available. We present a computational model, in which recruited cells in the midbrain superior colliculus drive eyes and head to the stored target location through a common dynamic oculocentric gaze-velocity command, which is continuously updated from the stable goal and transformed into appropriate oculocentric and craniocentric motor commands. We describe two equivalent, yet conceptually different, implementations that both account for the complex, but accurate, kinematic behaviors and trajectories of eye-head gaze shifts under a variety of challenging multisensory conditions, such as in dynamic visual-auditory multisteps.

dynamic feedback; gaze saccades; motor map; reference frames; spatial-temporal transformation; spatial updating; superior colliculus

INTRODUCTION

This paper deals with spatial updating of target locations for rapid goal-directed eye-head gaze shifts under static and dynamic localization conditions of visual and auditory targets. Gaze is defined as the eye orientation in world coordinates, and is given by the sum[1] of the head-in-world and eye-in-head vectors: $G_W = H_W + E_H$ (see GLOSSARY for a listing of variables and their definitions). We describe spatial coordinates in the world by their azimuth (horizontal plane; α) and elevation (vertical, median plane; ε) angles, defined in a double-pole coordinate system relative to the center of the upright head and body, when pointing straight ahead (Knudsen and Konishi 1979).

In a standard goal-directed gaze-orienting task, the subject aligns eyes and head on a central fixation spot, before generating a gaze shift to a peripheral target (T, either visual, or auditory; Fig. 1*A*). In this case, the world-, body-, head-, and eye-reference frames are all aligned (i.e., $G_0 = H_0 = 0$). The coordinates of the goal for the eye- (ΔG) and head (ΔH) motor responses will then be identical too, and correspond to the initial sensory coordinates on the retina, or at the ears. However, when the eyes are deviated from straight ahead, to G_0, as in Fig. 1*B*, the error signals for eyes and head will differ, and be (partly) dissociated from the sensory coordinates. For example, in case of a visual target on the retina at T_E, the coordinates for the goal-directed head movement will be $\Delta H = T_E + E_0$, with E_0 the initial eye-in-head orientation. In contrast, when the target is a sound, presented at head-centered location T_H, the signals that should drive the eye movement are given by $\Delta G = T_H - E_0$. Experiments have demonstrated that gaze- and head-movement trajectories are, indeed, expressed in their appropriate eye-centered and head-centered reference frames, as both remain goal-directed, irrespective of initial

Address for reprint requests and other correspondence: Radboud Univ., Donders Institute for Brain, Cognition and Behavior, Dept. of Biophysics, Heyendaalseweg 135, HG00.831, 6525 AJ, Nijmegen, the Netherlands (e-mail: j.vanopstal@donders.ru.nl).

[1] For simplicity, we confine our description to two-dimensional commutative, linear coordinate transformations, thereby neglecting the nonlinear noncommutativity of three-dimensional rotational kinematics.

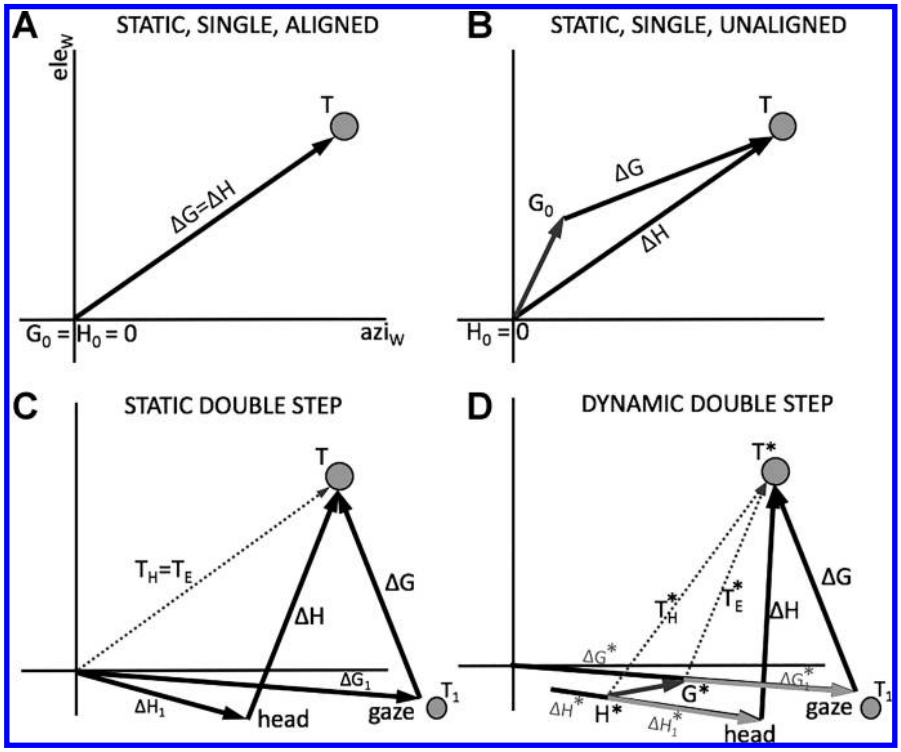

Fig. 1. *A–D*: four scenarios for programming a goal-directed eye-head gaze shift to a peripheral target, *T*, in increasing order of complexity. *A* and *B*: gaze shifts to a single target with different initial conditions. *C*: in the static double-step, the complete intervening gaze shift, ΔG_1, ΔH_1, should account for the target updates. *D*: in the dynamic double-step, however, only the unpredictable part of the ongoing movements after t^*, (ΔG_1^*, ΔH_1^*), should be compensated. See text and Table 1 for further explanation.

conditions, and sensory modality (Goossens and van Opstal 1997).

An additional complexity arises when an intervening gaze shift follows target presentation, but precedes the goal-directed eye-head movement to the target. Now, the initial sensory coordinates will be fully dissociated from the required motor commands of eyes and head. This is what happens, for example, in a static double-step trial (Fig. 1*C*). Here, two brief stimuli are presented in rapid succession, evoking two consecutive gaze shifts ("static" refers to the condition that nothing moves during target presentation). In this case, the coordinates of the second gaze shift have to incorporate the intervening eye- and head movements, elicited by the first target (at T_1). Because eye- and head-movement trajectories typically differ in a gaze shift, the orientations of eyes and head at the end of the gaze shift will be different too, and subsequent goal-directed movements will have to be generated in quite different directions. Depending on the sensory modality of the target (visual vs. auditory), the required coordinate transformations for the gaze and head movements will differ too. Behavioral experiments in humans (Goossens and van Opstal 1997) and monkeys (Van Grootel et al. 2012) have demonstrated that these coordinate transformations are, indeed, accurately performed.

Arguably, a more challenging localization problem arises when a brief visual or auditory target is presented in midflight of a fast, ongoing eye-head gaze shift (Fig. 1*D*; Vliegen et al. 2004, 2005). In this dynamic double-step paradigm the target is presented at time t^*, when the eye, on its way to T_1, has reached gaze position G^*, while the head may be at H^*. To subsequently orient toward the extinguished target location with a goal-directed gaze shift is not trivial, as timing (t^*), location, and modality of the target are all unpredictable to the system. As a result, neither the target coordinates ($T_E^*$and T_H^*), nor the appropriate eye-head motor commands (ΔH, ΔG), can be programmed beforehand, which is, in principle, possible for the other three scenarios (Fig. 1, *A–C*). Instead, the sensory signals, coordinate transformations, and updated motor commands have to be determined on the fly, during the rapid gaze shift, within a few tens of milliseconds.

Table 1 summarizes the required coordinate transformations for the eye and head, in case of a visual and an auditory target, for the static and dynamic double-step paradigms, respectively, where the symbols refer to Fig. 1, *C* and *D*.

Behavioral experiments in humans and monkeys have demonstrated that eye and head movements remain, indeed, goal directed, regardless stimulus timing, first gaze-shift characteristics (fast, slow, accurate, or inaccurate), modality (visual vs. auditory), and initial conditions (aligned, unaligned, eye-head onset delay; Vliegen et al. 2004, 2005; Van Grootel et al. 2012). This remarkable behavior requires that the gaze-control system: *1*) has continuous access to accurate signals about eye-in-head orientation, EH(t), and eye and head motor errors

Table 1. *Required coordinate transformations for the eye and head in space for static and dynamic double-steps, with a visual or auditory second target*

Second Target Modality	Gaze Coordinates, ΔG	Head Coordinates, ΔH
	Static Double-Step	
Visual	$\Delta G = T_E - \Delta G_1$	$\Delta H = T_E - \Delta G_1 + E_1$
Auditory	$\Delta G = T_H - \Delta H_1 - E_1$	$\Delta H = T_H - \Delta H_1$
	Dynamic Double-Step	
Visual	$\Delta G = T_E^* - \Delta G_1^*$	$\Delta H = T_E^* - \Delta G_1^* + E_1$
Auditory	$\Delta G = T_H^* - \Delta H_1^* - E_1$	$\Delta H = T_H^* - \Delta H_1^*$
Partial movement after t^*	$\Delta G_1^* = \Delta G_1 - \Delta G^*$	$\Delta H_1^* = \Delta H_1 - \Delta H^*$

E_1 is the eye-in-head position after the first gaze shift; ΔG^*, ΔH^* are the partial gaze- and head movements up to target presentation at t^*. Note that ΔG_1^*, ΔH_1^* correspond to the instantaneous gaze- and head motor errors of the intervening gaze shift at t^* (gray arrows in Fig. 1*D*).

(see Table 1); *2*) accounts for different internal delays for visual and auditory signals; and *3*) updates visual (T_E) and acoustic (T_H) target coordinates into the appropriate eye-centered and head-centered motor commands on millisecond time scales.

We have argued that a predictive, feedforward remapping strategy of target coordinates (like that proposed, e.g., by Goldberg and Bruce (1990) and Duhamel et al. (1992), for head-restrained visuomotor tasks) cannot account for this behavior, as the system has no prior access to the unpredictable stimulus properties (timing and modality), nor to the intrinsic, highly variable, partial eye and head movements that precede and follow target presentation (Vliegen et al. 2004, 2005).

One possible way to understand goal-directed orienting under these challenging conditions, expands on the early ideas of Robinson (1973, 1975), by assuming that targets are not kept in their original sensory coordinates, but are mapped into a world-centered reference frame, as soon as the sensory information becomes available (at $t = t^*$). In such a scheme, illustrated in Fig. 2, *left*, the coordinates of retinal targets are mapped into

$$T_W^{VIS} = T_E(t^*) + E_H(t^*) + H_W(t^*) \quad (1)$$

and the craniocentric acoustic coordinates of a sound source become

$$T_W^{AUD} = T_H(t^*) + H_W(t^*) \quad (2)$$

It is not immediately obvious how *Eqs. 1* and *2* are embedded in the system, given that the visual (~60–70 ms) and auditory (10–15 ms) sensory delays to the primary cortices and midbrain are quite different. Because the feedback signals should refer to the onset of the stimulus (t^*), they should also account for these modality-dependent delays.

Yet, because the world-centered target coordinates of *Eqs. 1* and *2* are invariant to any further intervening eye- and head movements, the system could store target information in a stable spatial memory, and plan a series of spatially accurate gaze shifts to different targets in the absence of any further sensory input. Another advantage of a single coordinate transformation of targets is its insensitivity to the accumulation of updating errors, which would occur when target updates would have to be made after every intervening movement (Van Grootel and van Opstal 2009).

Figure 2 incorporates these world-centered coordinate transformations into a conceptual scheme, in which the recruited neural population in the gaze-motor map of the midbrain superior colliculus (SC) specifies a common dynamic gaze-displacement signal for the brain stem (oculomotor system) and spinal cord (head-motor system). As the eyes and head move, the stable world-centered target coordinates for the next gaze shift are continuously updated into appropriate eye-centered target coordinates, through instantaneous feedback about any changes in eye and head position, e.g., during an ongoing gaze shift:

$$\Delta G(t) = T_W - E_1(t) - H_1(t) \quad (3)$$

At gaze-shift offset (around $t = t_1$), this updated signal becomes available to the motor map for the next gaze-shift command in a double-step: $\Delta G_2 = \Delta G(t_1)$.

We implemented a computational model, based on the scheme of Fig. 2, in which a population of neurons in the midbrain SC drives the eyes and head to multiple stored target locations. The SC cells encode the desired gaze trajectory, and its kinematics, through their cumulative spike counts, and instantaneous firing rates, respectively (Goossens and van Opstal 2006, 2012). Despite a number of simplifications in our model regarding the details of brain stem and cerebellar involvement in the downstream circuitry, and the eye and head motor plants, it accounts for the complex, yet accurate, kinematic behaviors and trajectories of measured eye-head gaze shifts under challenging multisensory conditions, such as obtained in static and dynamic visual-auditory double-steps (Fig. 1). We also discuss an alternative scheme, based on instantaneous gaze and head motor-error feedback.

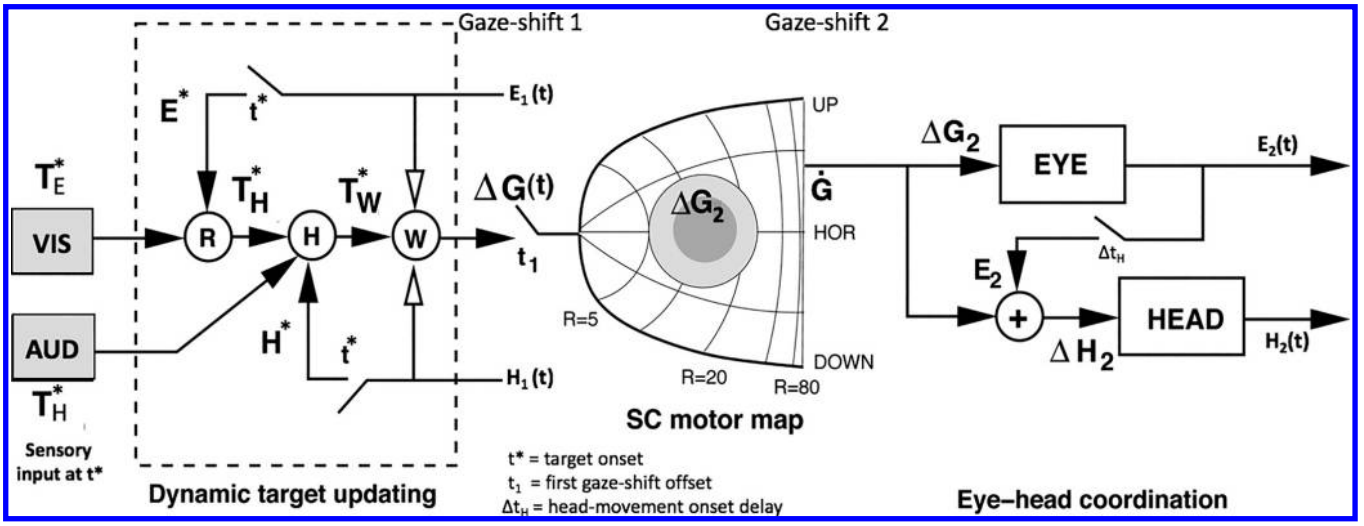

Fig. 2. Outline of dynamic target-updating and eye-head gaze control, as described by Vliegen et al. (2005). The scheme incorporates a dynamic update of auditory and visual targets into stable world-centered coordinates, T_W, (*Eqs. 1* and *2*) to program the next eye-head gaze shift (*Eq. 3*). The SC motor map generates a dynamic gaze command, $\dot{G}$, that rapidly drives eyes (ΔG_2) and head (ΔH_2) in their appropriate reference frames to the target (note that all signals represent two-dimensional vectors). Details of the eye and head motor systems have been omitted for clarity (see METHODS and Figs. 3 and 4).

METHODS AND RESULTS

Simulations were performed in MATLAB (MathWorks, Natick, MA, version 14a). The MATLAB script of the dynamic updating model in world coordinates is made available from the corresponding author upon request.

Internal Delays

Suppose a gaze shift is made in some arbitrary direction, with gaze and head movement trajectories, $G(t)$ and $H(t)$ (t is the current time), and that at time $t = t^*$, a visual target appears on the retina, with coordinates $T_{E,RET}(t^*)$. This visual signal reaches the perceptual system (cortex) at time t_1, after a delay of ~60 ms, so that $T_{E,PERC}(t_1) = T_{E,RET}(t^* + 60)$.

According to *Eq. 1*, specifying the world-centered location of a visual target at the time of the flash needs eye- and head-position signals that refer to t^*. However, for a rapid gaze shift, these positions may differ substantially from those measured at the perceived time $t = t_1$. Suppose that the percept of both position signals is delayed by 20 ms, then $G_{PERC}(t_1) = G(t_1 - 20)$ and $H_{PERC}(t_1) = H(t_1 - 20)$. As the signals should be referred to the target flash in the visual field, i.e., at $t^* = t_1 - 60$, we get $G(t^* + 40)$. Thus, the visuomotor system should possess a memory buffer for gaze and head positions of ~40 ms to correctly map the retinal target into a world-centered reference frame: $T_W^* = T_E^* + G^*$. For sounds, the sensory-cortical delay is ~20 ms, so that the memory span for head position to auditory targets would be ~20 ms: $T_W^* = T_H^* + H^*$. In what follows, we assumed that these internal delays are appropriately compensated in the system, and we discarded them in our calculations, for simplicity.

Generating the Eye-Head Gaze Shift

Once T_W is constructed (and tagged for the next gaze shift), its associated gaze error is calculated from the instantaneous eye- and head positions by *Eq. 3*. At "go" time, t_2, the population of cells in the SC motor map, thus, represents the desired gaze-displacement command ΔG_2. While this gaze shift is executed, the changing eye and head positions are used to continuously update the gaze coordinates (*Eq. 3*) for the next target (if stored in memory).

The gaze shift is driven by the linear spike-count model of the SC, as proposed by Goossens and van Opstal (2006, 2012). Although this model was originally formulated for head-restrained eye saccades, here, we extend this concept to eye-head gaze shifts.[2] Thus, each spike (s) from each cell (k) in the SC contributes a fixed (small) "spike vector," $\vec{m}_k$, to the desired gaze command, the size and direction of which depend solely on the cell's location in the motor map. The desired gaze shift is then encoded by the linear cumulative sum of all spike vectors from all spike trains in the saccade-related SC bursts:

$$\Delta G(t) = \eta \cdot \sum_{k=1}^{N} \sum_{s=1}^{S_k<t} \vec{m}_k \cdot \delta(t - \tau_{k,s}) \qquad (4)$$

with N the number of cells in the recruited population, η a fixed scaling constant, and $\delta(t - \tau)$ is a spike at time τ; thus, $\tau_{k,s}$ is the timing of spike s from cell k, and $S_k < t$ is the cumulative number of spikes from cell k, up to the current time, t.

Note, from *Eq. 4*, that the firing rates of the neurons directly influence the planned gaze kinematics. Indeed, our head-restrained recordings have shown that the SC motor map encodes the kinematic main-sequence relations of saccades by a spatial gradient in the peak firing rates of the neurons (high at the rostral zone for small-amplitude saccades, and lower at the caudal SC, for large saccades; Goossens and van Opstal 2006), in combination with a fixed number of spikes in the burst (see below, implementation, and Fig. 5). As a result, the instantaneous firing rates of all SC neurons together encode the (desired) gaze-velocity profile, while the cumulative number of spikes (the time integral of the bursts) encodes the amplitude of the gaze shift (illustrated in Fig. 3*A*, *inset*). The kinematics of gaze shifts are, thus, specified by the SC, leaving the brain stem and spinal cord burst generators essentially linear (performing vector decomposition; Goossens and van Opstal 2006).

Figures 3 and 4 show the different steps that specify how a common oculocentric collicular signal generates the goal-directed movements of gaze and head. The head movement is driven by a craniocentric motor error, which is transformed into a desired head movement vector in the same direction, ΔH_{DES} (Fig. 3*B*, *inset*). The amplitude of the head movement

[2] Recordings in head-unstrained monkeys indicate similar spike-count behavior of SC cells for gaze shifts than for head-restrained eye saccades, reported by van Opstal and Goossens (2006, 2012): the cumulative number of spikes in the burst encodes a straight gaze shift (unpublished data).

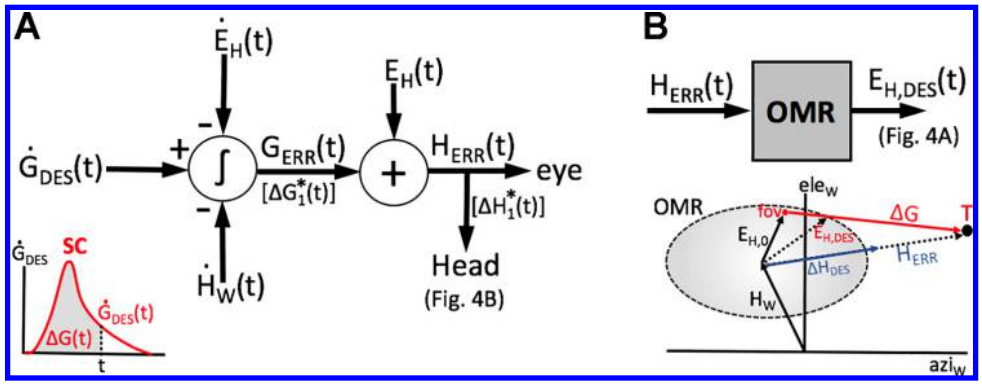

Fig. 3. Decomposition of the desired gaze-velocity output of the SC population into goal-directed oculocentric and craniocentric motor commands. *A*: difference between desired gaze-velocity (SC output, *inset*) and feedback from the actual gaze-velocity (sum of the ongoing eye and head velocities) is integrated [as in Scudder's (1988) model, albeit that the SC command represents a dynamic signal], yielding dynamic gaze-motor error, $G_{ERR}(t)$. To prevent the eyes from running into the oculomotor limits, this signal has to be transformed into a craniocentric error, $H_{ERR}(t)$, by combining it with instantaneous eye position. *B*: this craniocentric error subsequently determines the desired eye position within the oculomotor range (OMR) for the gaze shift. Note that $E_{H,DES}$ changes during the gaze shift, because of the gaze and eye position feedback. fov, fovea; $E_{H,0}$: initial eye position; ΔH_{DES}: desired head displacement.

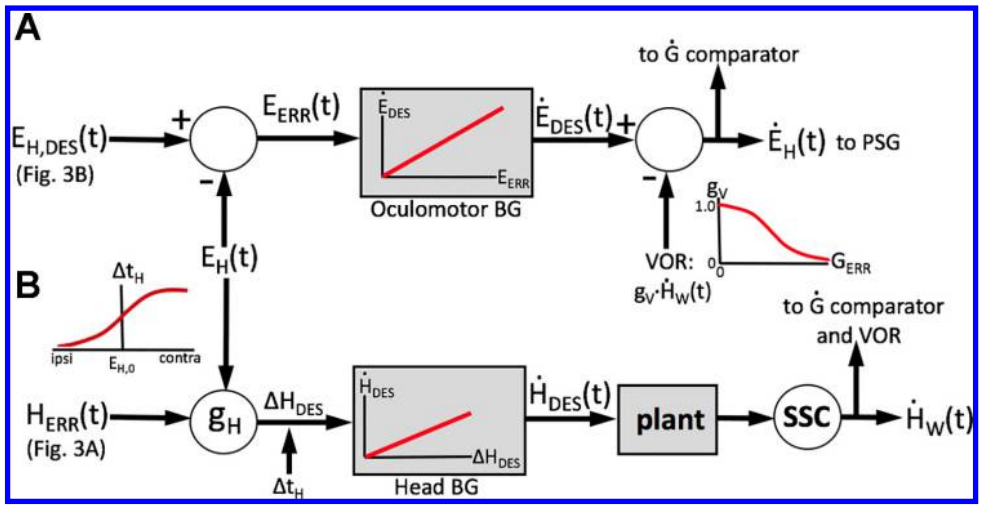

Fig. 4. *A*: linear oculomotor burst generator (Goossens and van Opstal 2006, 2012) is driven by instantaneous eye-motor error, $E_{ERR}(t)$, derived from the (dynamic) desired eye position (Fig. 3*B*). Its output is a desired eye-velocity signal, which is modified by the vestibular ocular reflex (VOR) into the actual eye-in-head velocity, $\dot{E}_H(t)$ (g_V is the VOR gain, which depends on current gaze error, see *inset*). This latter signal acts as an efference copy for the gaze-velocity comparator (Fig. 3*A*) and is integrated by the neural integrator in the pulse-step generator (PSG) to provide an efference copy of eye position, $E_H(t)$. *B*: head-movement controller is driven by a desired head displacement, ΔH_{DES}, which is proportional to the initial head motor error (Fig. 3*A*), with a head-onset delay, Δt_H, which depends on initial eye position (*inset*). The head-velocity feedback for the gaze comparator is derived from the vestibular system, which measures the actual head velocity with respect to the world (SSC, semicircular canals).

is determined by an eye-position-dependent gain: $\Delta H_{DES} = g_H \cdot H_{ERR,0}$ (Fig. 4*B*). Experiments have demonstrated that if the eye looks contralateral to the target, the head-movement contribution to the gaze shift will be relatively small, and late, as most of the gaze shift will be carried by the eye. Conversely, if the eye looks into the direction of the target (ipsi), the head-movement contribution will be larger and earlier (and, consequently, the gaze shift will be slower; Goossens and van Opstal 1997).

Model Implementation

The simulations ran at a time resolution of 1 ms (i.e., feedback signals arrived with a delay of one sample). The set of equations specifying the eye- and head-motor commands in the model is described as follows.

The population of recruited cells in the SC motor map effectively issues a gaze-velocity pulse (by the summed instantaneous firing rates of all neurons), which is given by $\dot{G}_{DES}(t)$ (see e.g., Fig. 3*A*, *inset*). In the most straightforward simulations, this profile was approximated by a simple rectangular pulse, issuing a constant output of 20 spikes/s, with an amplitude (after scaling by ΔG) that specifies the peak gaze velocity, and a duration that increases with gaze amplitude, and is modulated by eye position:

$$\dot{G}_{DES}(t) = \begin{cases} G_{PK} & \text{for } 0 \le t \le D \\ 0 & \text{elsewhere} \end{cases} \text{ with } \begin{cases} D = a + b \cdot \Delta G + c \cdot \Delta E_{G,0} \\ G_{PK} = \dfrac{20}{D} \end{cases} \tag{5}$$

where a (20 ms), b (1.5 ms/°), and c (0.3 ms/°) are constants, and $\Delta E_{G,0}$ is the initial eye-in-head position component along the gaze-shift vector (Fig. 5):

$$\Delta E_{G,0} = \frac{\left(\vec{E}_0 \cdot \Delta\vec{G}\right)}{\|\Delta G^2\|} \cdot \|E_0\| \tag{6}$$

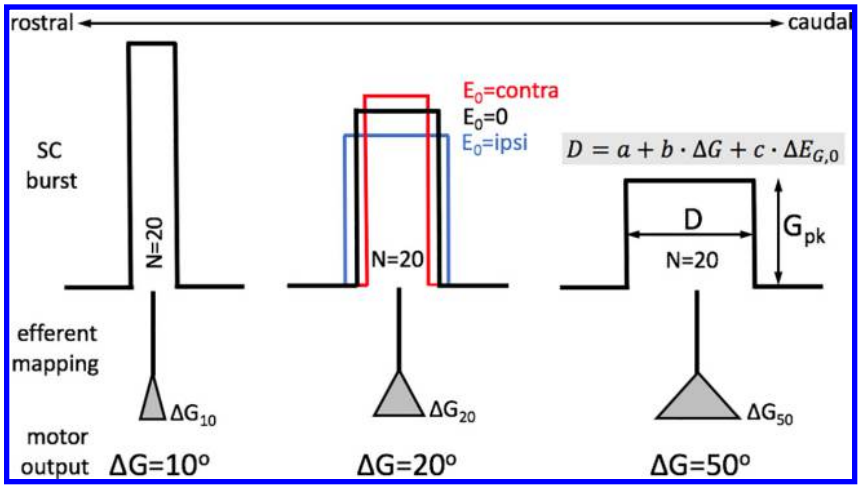

Fig. 5. Example SC burst profiles, as used in the simulations, for different gaze shifts (small, medium, and large), and for different initial eye positions (shown for the center bursts only). Each burst has the same number of spikes (n = 20), but its height systematically decreases from the rostral-to-caudal end of the motor map. The efferent projections to the brain stem burst generators weigh the bursts to specify the gaze-shift amplitude and its velocity. Initial eye position influences the burst duration and its peak firing rate: duration increases/decreases for eye positions in the same/opposite direction (ipsi/contra) as the upcoming gaze shift.

For contralateral eye positions, $\Delta E_{G,0} < 0$ (decrease in SC burst duration, increase of the peak firing rate), and for ipsilateral positions, $\Delta E_{G,0} > 0$. In this way, the cells in the SC motor map specify the kinematics of saccadic gaze shifts (a straight-line amplitude-duration function, and a saturating amplitude-peak velocity relation), by adjusting the peak firing rates according to a decreasing spatial gradient along the rostral-caudal (gaze-amplitude) axis of the map (Goossens and van Opstal 2012; van Opstal and Goossens 2008), as well as the dependence on initial eye position. The efferent mapping (synaptic projections to the motor systems) ensures that the integral of the burst encodes the correct gaze-shift amplitude.

The gaze-motor error (Fig. 3*A*) is then determined by the cumulative time integral of the difference between the desired and actual gaze velocity:

$$G_{ERR}(t) = \int_0^t \left[\dot{G}_{DES}(\tau) - \dot{E}_H(\tau) - \dot{H}_W(\tau)\right] \qquad (7)$$

Clearly, the gaze error may specify a movement that would exceed the mechanical limits of ocular motion (Fig. 3*B*, *inset*). To prevent this from happening, it is transformed into a craniocentric error, to determine the desired eye-in-head position, $E_{DES}(t)$, constrained by a "soft" oculomotor range (OMR; Fig. 3*B*). In our simulations, we described the OMR by a simple rectangle, with the horizontal eye position confined to $|E_{H,x}| < \mathrm{OMR}_H$ and $|E_{H,x}| < \mathrm{OMR}_H$. When the target remains within the OMR, the desired eye position is simply $E_{DES} = E_H + \Delta G$. When it's outside the OMR (like in Fig. 3*B*), E_{DES} is determined by the intersection of the line between fovea and target, and the OMR.

The desired eye-position signal drives the eye-velocity output of the brain stem burst generator within an eye-position feedback loop, just like in Robinson's influential eye-position feedback model (Van Gisbergen et al. 1981). However, in contrast to Robinson's original concept: *1*) the desired eye position (the "goal") is a dynamic signal that changes during the gaze shift, *2*) it was derived from a desired oculocentric gaze-displacement signal, *3*) the input-output characteristic of the burst generator is linear, and *4*) the actual eye-in-head velocity is modified by the head movement on the output of the oculomotor burst generator (the desired eye velocity signal) through the (gain-modulated) vestibular-ocular reflex (VOR):

$$\dot{E}_H(t) = \dot{E}_{DES}(t) - g_V\left[G_{ERR}(t)\right] \cdot \dot{H}(t) \qquad (8)$$

where the VOR gain, g_V, varies in a sigmoid fashion with the instantaneous gaze-motor error: it is close to zero when the gaze error is large, and close to one when the error approaches 0 (e.g., Laurutis and Robinson 1986; Tabak et al. 1996).

Together with the eye movement, also the desired head-motor command (issued with an onset delay, Δt_H, regarding gaze onset) is derived from the collicular desired gaze-shift command (Fig. 3*A*). It is calculated as follows:

$$\Delta H(t - \Delta t_H) = g_H \cdot \left[\int_0^t \dot{G}_{ERR}(\tau)d\tau + E_H(t)\right] + g_E \cdot \dot{E}_{DES}(t) \qquad (9)$$

where the head-onset delay, Δt_H, depends on the initial eye-position component along the gaze shift, and on the sensory modality:

$$\Delta t_H = p_{MOD} - \mathrm{d} \cdot \Delta E_{G,0} + N(0, 15) \qquad (10)$$

where we took $p_{MOD} = 30$ ms for a visual target, and 10 ms for an auditory target (e.g., Goossens and van Opstal 1997), and $d = 0.3$ ms/deg. $N(0,15)$ is a Gaussian random variable with zero mean and std = 15 ms. The attenuating head gain, g_H, was taken as 0.6 for the horizontal component, and 0.4 for the vertical component. The result of *Eq. 10* implicitly determines the actual contribution of the head movement to the gaze shift: it will be large for short delays (auditory targets and ipsilateral eye positions), and small for longer delays (visual targets and contralateral eye positions).

It should be noted that the desired head displacement of *Eq. 9* typically differs from the actual contribution of the head movement to the gaze shift. As the latter is determined at gaze-saccade offset, the former may be considerably larger. Especially at long head delays (e.g., for contralateral eye positions), the head movement contribution, $\Delta H_{\mathrm{Goff}} = \Delta G - \Delta E_{H,\mathrm{Goff}}$, will be relatively small, although the total head movement may still be considerable.

The head velocity (output from a simple first-order head plant) is measured by the VOR, and is subsequently used *1*) to modulate the desired eye velocity from the burst generator ($\dot{E}_{DES}$; *Eq. 8*), and *2*) to construct the gaze-velocity feedback. The right-hand term in *Eq. 9* accounts for the (small) influence (g_E) of the oculomotor burst on the head-motor response (reported by Goossens and van Opstal 1997). The integrated outputs of the actual eye and head velocities are instantaneous eye and head position, respectively, both of which are used in the world-centric target updates of Figs. 2 and 6.

Target Updating

Sampled eye and head positions. The switches in the target update mechanism of Fig. 2 symbolize the sampling of the eye- and head-position feedback signals at stimulus presentation time. Figure 6 illustrates four subsequent target events: two visual targets presented on the retina (at t_1 and t_3), and auditory targets presented at t_2 and t_4, where $t_1 < t_2 < t_3 < t_4$.

The system detects and stores each target event, whenever the temporal derivative in sensory space exceeds a threshold. In Fig. 6*A*, we illustrated this for step-like sensory changes, for which the derivatives are mere delta functions that sample the eye and head positions precisely at stimulus onset, according to (Fig. 6*B*):

$$\begin{aligned} T^V_{W,1} &= T^E_1 + E(t) \cdot \delta(t - t_1) + H(t) \cdot \delta(t - t_1) \\ T^A_{W,2} &= T^H_2 + H(t) \cdot \delta(t - t_2) \\ T^V_{W,3} &= T^E_3 + E(t) \cdot \delta(t - t_3) + H(t) \cdot \delta(t - t_3) \\ T^A_{W,4} &= T^H_4 + H(t) \cdot \delta(t - t_4) \end{aligned} \qquad (11)$$

where

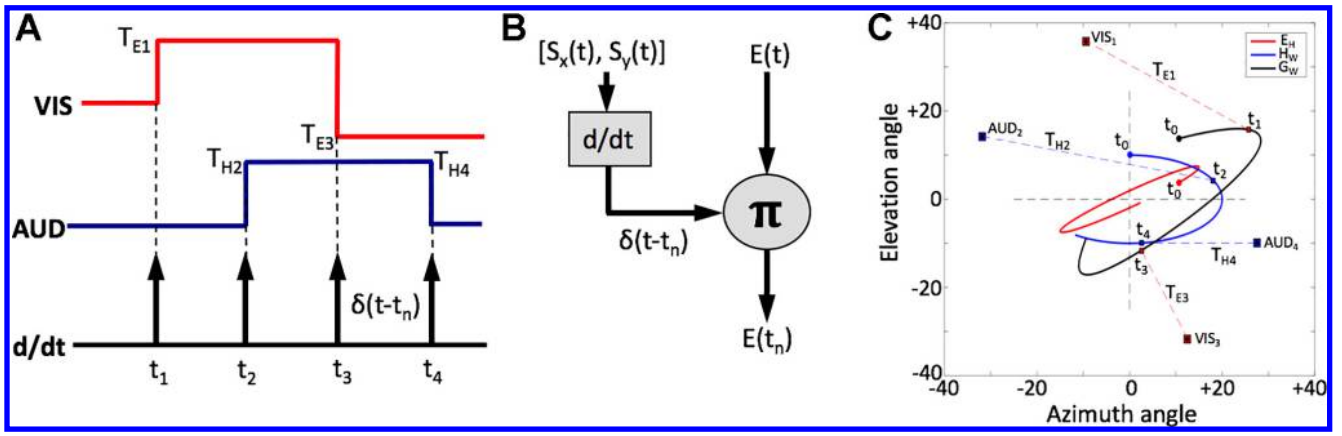

Fig. 6. *A*: example target appearances in the visual and auditory channels, with the associated triggers at t_1–t_4, representing the sampling switches in Fig. 2. *B*: multiplying ongoing eye position with a measure for the rapid sensory change, which approximates a delta function, yields the sampled signal (here: eye position) at target appearance. *C*: calculation of the world-centered target locations, presented during ongoing gaze (black trace) and head (blue trace) movements (generated by a sum of sines, for illustration purposes). Red trace: $E_H(t)$. Small squares within the gaze and head trajectories: sampled gaze (red) and head (blue) positions at the four target onsets.

$$\delta(t - t_N) \approx \left|\vec{S}_N(t)\right| = \left|\left(\frac{\partial S_{x,N}(t)}{\partial t}, \frac{\partial S_{y,N}(t)}{\partial t}\right)\right| \qquad (12)$$

for visual and auditory targets. The derivative reports a sudden sensory change at any (retinal or head-centered) location. Note that, in principle, this could be any change in target position, stimulus intensity, visual contrast, and interaural level difference.

Model Responses

The model responses faithfully mimic the natural behavior of measured eye-head gaze shifts. Figure 7*A* illustrates a number of gaze shifts in different directions and amplitudes, all starting from aligned initial conditions from straight ahead. We also added two example trajectories of gaze shifts having unaligned initial conditions. In these two examples, it can be seen that both gaze and head make goal-directed movements, and, therefore, follow different spatial trajectories. In Fig. 7*B*, we plotted the main-sequence peak-velocity behavior of gaze shifts for amplitudes between 5° and 65° for three different conditions: initial positions of the eye and head aligned (red), the eye looking in the ipsilateral direction (between +16 and +50%) of the upcoming gaze shift (blue), and in the contralateral direction (between −20 and −100%; green). Note that there is no unique main-sequence relation for eye-head gaze shifts, as the peak velocities depend strongly on the initial eye position: the contralateral condition yields the fastest gaze shifts, while ipsilateral initial conditions produce markedly slower gaze shifts.

The influence of initial eye position on the detailed kinematics of simulated gaze shifts and the associated head and eye movements is further illustrated in Fig. 8, for 50° oblique gaze

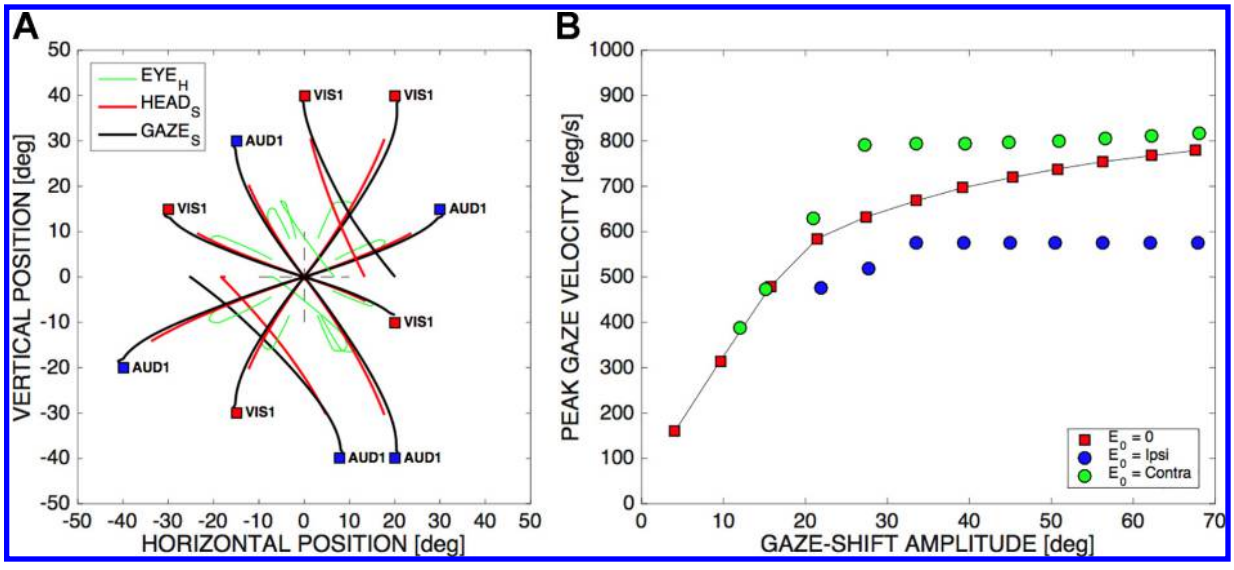

Fig. 7. Single-target gaze shifts. *A*: spatial trajectories for eight gaze shifts to visual and auditory targets in different directions and amplitudes with eye and head aligned at straight ahead (center of the cross). Two example trajectories are shown for incongruent initial conditions: rightward eye fixation for a visual target, and a leftward eye fixation for an auditory target. Note that the gaze and head trajectories are both goal-directed. *B*: gaze kinematics depend on the initial conditions with the eyes and head aligned; the gaze shifts follow a nonlinear main-sequence relation (red squares). Faster gaze shifts occur when the eyes fixate in a direction contralateral to the gaze shift (green symbols). Gaze shifts become markedly slower when the eyes fixate in the ipsilateral direction of the gaze shift (blue symbols). See also Fig. 8.

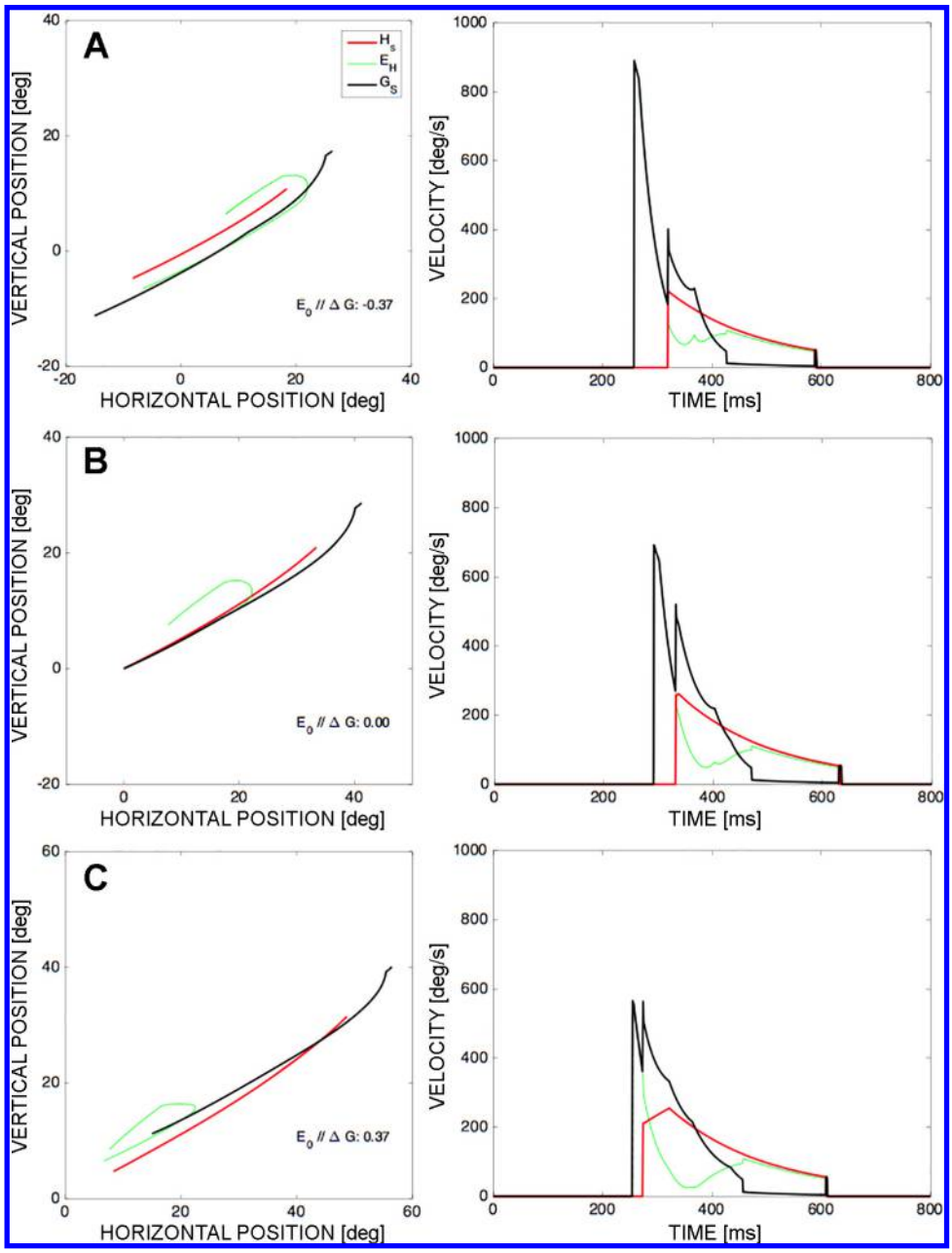

Fig. 8. Simulated gaze shifts with fixed polar coordinates, $R = 50°$, $\Phi = 37°$, but for three different initial eye-in-head fixations. *A*: initial eye position in the contralateral direction (−37%) of the gaze shift, at $E_0 = (-15, -11.2°)$. *B*: eyes are centered in the head, $E_0 = (0,0)$. *C*: eyes look in the ipsilateral direction (+37%) of the gaze shift, at $E_0 = (+15, +11.2°)$. Left-hand column: eye, head, and gaze trajectories. Right-hand column: eye, head, and gaze-vectorial velocities. Note the strong eye-position dependence of the contributions of the eye and head to the gaze shift, as well as to the gaze kinematics and head-onset delay. The fastest gaze shift, with the largest eye movement, and smallest and latest head movement is obtained for the eye in the contralateral initial position (*A*). The slowest gaze shift with the smallest eye movement and the largest head movement is obtained for the eye in the ipsilateral direction (*C*).

saccades, generated for three initial eye-in-head positions: contralateral (−37%; Fig. 8*A*), aligned (Fig. 8*B*), and ipsilateral (+37%; Fig. 8*C*). Note the different head movements and the associated changes in the gaze-velocity profiles for the different initial conditions.

An example for static multistep head and gaze trajectories, made by a monkey toward a series of three visual targets (taken after Bremen et al. 2010) is shown in Fig. 9. As the visual targets had a duration 900 ms in this experiment, the three consecutive gaze shifts were all executed under closed-loop static visual-feedback (as in Fig. 1*B*).

Figure 10*A* shows a simulation of the eye, head, and gaze trajectories, when the multiple target steps ($n = 5$) were made in an open-loop static localization mode, as in Fig. 1*C*, and to both visual and auditory targets that alternated in the sequence. In this simulation, the brief target flashes (and sound bursts) were presented before the onset of the intervening gaze shift to the previous target (Fig. 10*B*). The model produces spatial trajectories, shown in Fig. 10, which have a clear resemblance to the head and gaze trajectories of the monkey (Fig. 9). Table 2 gives the sensory coordinates at the moment of target presentation [for visual (V): retinal coordinates, for auditory (A): craniocentric coordinates], as well as the calculated world-centered target coordinates, and evoked gaze shifts.

Figure 11 illustrates the behavior of the model under dynamic localization conditions in a visual-visual (VV) sequence (like in Fig. 1*D*), in which target V_2 appeared in midflight of the first visual-evoked eye-head gaze shift. The trajectories of the gaze- and head-movements were all goal-directed, despite the considerable differences in initial conditions at the end of the intervening gaze shifts (Fig. 11*A*). Even though both gaze shifts could have amplitudes that would carry the eye beyond the oculomotor range, the complex eye-in-head movement trajectory stayed within the OMR of [± 25, ± 20]°.

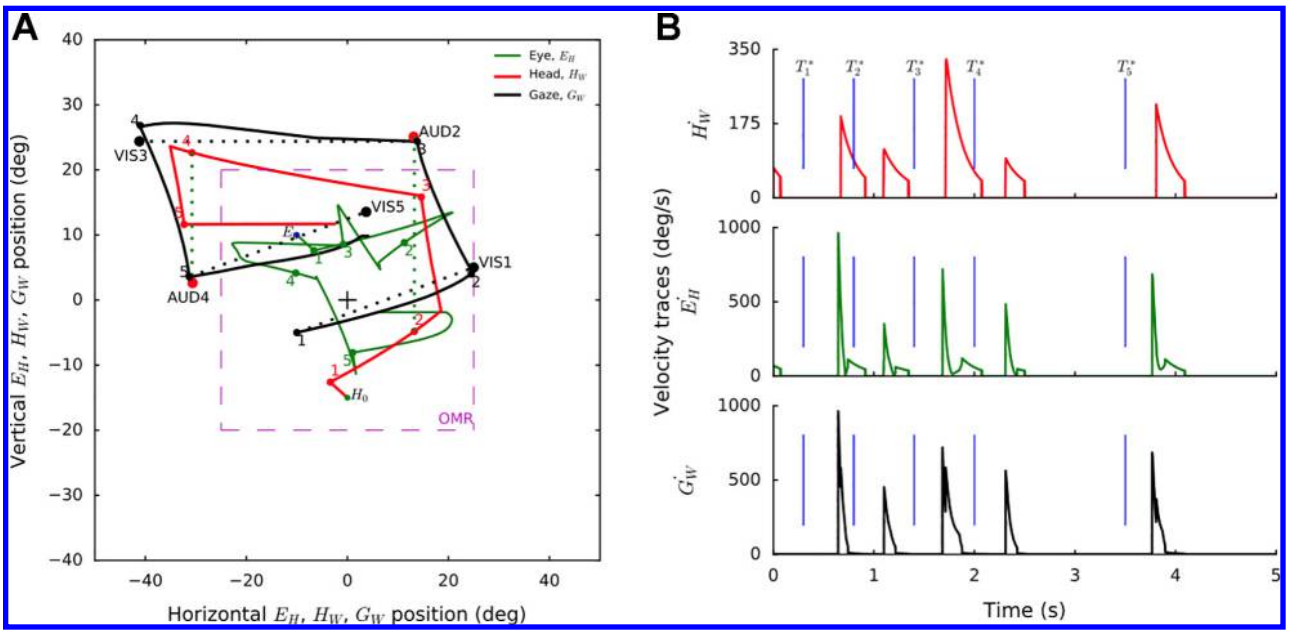

Fig. 10. *A*: simulated open-loop eye, head, and gaze trajectories to a series of five static multiple-step targets (VAVAV), presented as in Fig. 1*C*. *B*: track velocities of the gaze shifts (black) and of the head (blue). Note the correspondence with the measured trajectories of the monkey. Magenta lines: onsets of the brief visual and auditory targets.

The stimulus timings (magenta lines), as well as the horizontal/vertical time traces of the gaze, head, and eye movements, are shown in Fig. 11*B*. In this panel, we also included the simplified SC bursts (rectangular pulses), each scaled for the upcoming gaze-amplitude, according to *Eq. 5*. The kinematics of gaze, eye, and head for the two gaze shifts are presented in panel Fig. 11*C*.

Table 3 provides the target and movement coordinates at the moment of presentation during the gaze shifts.

DISCUSSION

We described a model for eye-head gaze control that copes with dynamic multisensory and multistep localization problems. As illustrated in Fig. 1*D*, and in Table 1, these sensory conditions impose several nontrivial updating challenges to the system. Nevertheless, behavioral experiments with humans (Vliegen et al. 2004, 2005) and monkeys (Van Grootel et al. 2012) have indicated that accurate and precise dynamic target updating can operate at millisecond time scales.

To our knowledge, our model is the first to deal with multitarget and multisensory dynamic feedback behavior; the large majority of models so far had been designed to explain the generation of combined eye-head gaze shifts to single (typically visual) targets (as in Fig. 1, *A* and *B*; Daye et al. 2014; Freedman 2001; Galiana and Guitton 1992; Guitton 1992; Guitton and Volle 1987; Lefèvre and Galiana 1992). Moreover, most models were restricted to horizontal gaze shifts, without the additional complexities of cross-coupling the horizontal and vertical movement components, which arise for oblique gaze shifts (with some notable exceptions, e.g., Daye et al. 2014; Goossens and van Opstal 1997; Tweed 1997).

Major Differences with Earlier Models

Our model extends earlier proposals in several ways:

1) The midbrain SC acts as a nonlinear vectorial pulse generator that encodes through its recruited population not only the amplitude and direction of the upcoming gaze shift vector, but also specifies, through the distribution of its spike trains, the desired instantaneous gaze-shift kinematics and trajectories. As head-restrained saccades, with aligned initial conditions, have characteristic main-sequence properties (a saturating amplitude-peak velocity relation), the SC has been proposed to implement this nonlinear behavior through a topographic organization (i.e., a spatial, rostral-caudal gradient) of the peak firing rates, burst durations, and burst skewness,

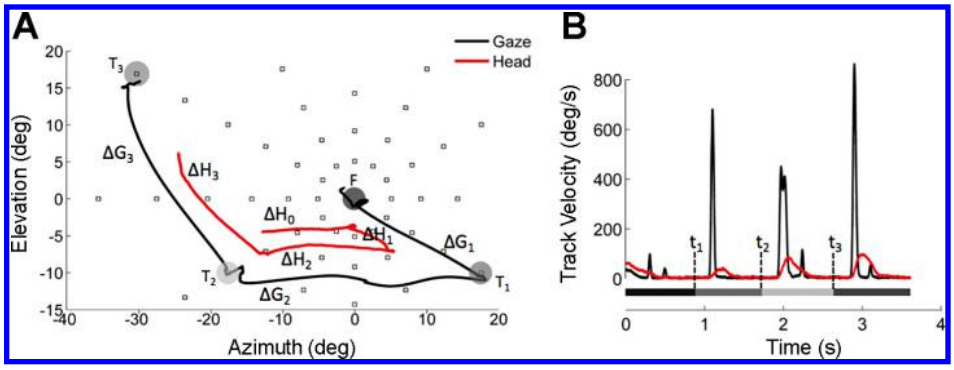

Fig. 9. *A*: closed-loop head and gaze trajectories to a sequence of three visual targets, measured in monkey (after Bremen et al. 2010). Targets were presented as in Fig. 1*B*; gaze shifts could be planned on the basis of visual feedback. *B*: corresponding gaze and head track velocities. The target presentation times are indicated by the gray horizontal bars, and their onsets are indicated by the dotted lines.

Table 2. *Sensory coordinates at the moment of target presentation*

	Sensory Coordinates		World Coordinates		Gaze Shifts	
Target	Hor., °	Vert., °	Hor., °	Vert., °	R, °	Φ, °
V1	35.0	10.0	25.0	5.0	36.4	16.0
A2	0.0	30.0	11.0	24.0	24.1	124.6
V3	−55.0	0.0	−43.4	23.6	55.0	180.0
A4	0.0	−20.0	−34.7	2.1	25.4	−69.9
V5	35.0	10.0	−0.2	12.7	36.4	16.0

Columns 2 and 3: sensory coordinates (in °) at the time of stimulus presentation (*t**), either on the retina (visual), or relative to the head (auditory). In Fig. 7*A* the sensory coordinates are shown as red (visual) and blue (auditory) dotted lines between the gaze (visual) and head (auditory) traces, taken at the moment of target presentation (*t**; small squares), and the associated target location (as in Fig. 5*C*). *Columns 4 and 5*: The world coordinates correspond to the red (visual) and blue (auditory) squares in Fig. 7*A*. *Columns 6 and 7*: The updated gaze-shift vectors are given in polar coordinates (R = amplitude, Φ = direction; −90° = down, 0° = right, +90° = up; 180° = left). Hor., horizontal; Vert., vertical.

Table 3. *Simulation results for dynamic multisteps, shown in Fig. 11*

	Sensory Coordinates		World Coordinates		Gaze Shifts	
Target	Hor., °	Vert., °	Hor., °	Vert., °	R, °	Φ, °
V1	35.0	−10.0	35.0	−10.0	36.4	−16.0
V2	−10.0	30.0	9.4	22.2	43.2	126.2

Simulation results for dynamic multisteps, shown in Fig. 11, in the same format as Table 2. In Fig. 11*A*, the sensory coordinates are shown as red dotted lines between the gaze (visual), taken at the moment of target presentation (*t**, small squares in the appropriate traces), and the associated target location (as in Fig. 6*C*). Note that *target 2* appeared in midflight of the first gaze shift. Hor., horizontal; Vert., vertical.

within the motor map: cells at the rostral zone, encoding small gaze shifts, are endowed with high peak firing rates and short burst durations, whereas caudal cells fire long-duration bursts at much lower firing rates. Yet, the number of spikes in the SC bursts does not systematically vary across the motor map. This hypothesis was forwarded on the basis of head-restrained collicular recordings, and was supported by quantitative analyses of single-unit responses, and computational modeling (Goossens and van Opstal 2006, 2012; Kasap and van Opstal 2017; Van Gisbergen and van Opstal 1989; van Opstal and Goossens 2008).

Here, we extended this idea to head-unrestrained gaze shifts, by proposing that the SC population specifies an abstract, desired, straight gaze trajectory (the sum of the eye-in-space and head-in-space trajectories), for which the specific motoric details (the actual oculomotor and head-motor signals) are subsequently extracted downstream from the motor SC within independent brain stem, spinal cord and cerebellar circuitries (Figs. 3 and 4). Recent recordings in head-unrestrained monkeys have revealed that, indeed, the cumulative number of spikes in the SC bursts of single units is much better described by the instantaneous straight-line gaze-displacement vector than by the eye-in-head saccade (van Opstal AJ, unpublished data).

Note that our model does not imply that the SC burst (e.g., Fig. 5) encodes the detailed changes in the velocity profiles for different contributions of the head, as illustrated in Fig. 8. Indeed, our simplified rectangular SC bursts produce clearly differently shaped gaze-velocity profiles (Figs. 7–11), which result from the (nonlinear) interactions between the eyes and head motor systems, the constraints imposed by the OMR, and the involvement of the VOR. Yet, as the head contribution to the gaze shift is strongly determined by the initial eye-in-head position (contra- vs. ipsi-gaze, e.g., *Eq. 10*, Figs. 4*B*, 7, and 8), by the sensory modality (visual vs. auditory), and also by top-down task-related factors (Goossens and van Opstal 1997; Laurutis and Robinson 1986), the SC burst already reflects the resulting influence of the head movement to the gaze shift in its spike trains on the basis of an eye-position signal that enters the SC motor map (introduced in *Eq. 5*). Evidence for the presence of an eye-position signal in the SC stems from single-unit recordings for saccades under head-restrained conditions ("gain field" tuning; Stuphorn et al. 2000; van Opstal et al.

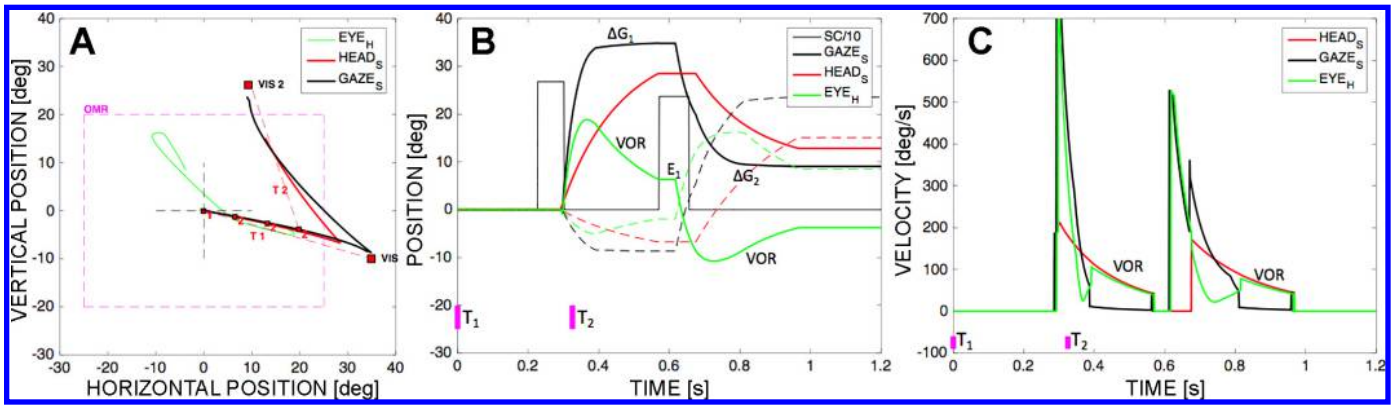

Fig. 11. Model performance under dynamic double-steps to two subsequent visual targets (VV), where the second target is presented in midflight of the first gaze shift. *A*: visual (red squares) targets are indicated in their order of presentation, in the world-centered coordinate system of the plot. Eye-in-head (green trace) and head-in-space (red trace) started with a gaze shift (black trace) from an aligned initial condition (at 1) to the first target (VIS 1). Note that the gaze and head traces are goal directed. OMR, oculomotor range in the head. *B*: time traces. Solid lines denote horizontal position of eye-in-head (green), gaze (black), and head-in-space (red) denote dashed lines vertical position traces. The thin rectangular pulses show the SC bursts for the two gaze shifts, for which the duration and height depend on the gaze-shift amplitude and initial eye position (which is E_1 for the second gaze shift; *Eq. 5*). Magenta lines denote stimulus presentation times. VOR, vestibular ocular reflex after gaze offset. Note that the second gaze shift (ΔG_2) starts during the second SC burst. *C*: vectorial velocity profiles of the different movement components during the two gaze shifts. The VOR is indicated. The head movement in the first gaze shift starts simultaneously with the eye at gaze onset; in the second gaze shift, it is delayed by ~50 ms, resulting in a double-peaked gaze-velocity profile.

1995). Recently, we obtained further support for a similar modulation in the head-unrestrained monkey SC, as schematically illustrated in Fig. 5 (van Opstal AJ, unpublished data).

2) Whereas most models assume nonlinear feedback burst generators to control the eyes and head, in our model, the eye and head feedback control circuits are driven by independent linear burst generators. Although this specific feature is not a critical aspect of our model, it emphasizes that there is no specific need for additional kinematic nonlinearities in the control of the eye and head motor systems to explain the major kinematic properties of gaze shifts (saturating peak velocities for large movements, with longer movement durations and increased skewness; e.g., Fig. 7*B*), as these are already accounted for by the firing patterns of the SC population. Yet, the limiting oculomotor range clearly imposes a nonlinearity in our model, as well as the (small) coupling of the oculomotor system on the head-movement trajectory (*Eq. 9*), the involvement of the VOR, and the eye-position dependent influence of the head. Further, the inclusion of more realistic (position-dependent) plant models will add additional nonlinearities to the control circuits. In this paper, we have not attempted to work out the tedious details to include all these additional factors, but instead to focus on the major functional units that are required to explain the observed behaviors.

3) The eyes and head are controlled by signals specified in their respective reference frames: an oculocentric gaze-error signal drives the eyes, and a craniocentric head-motor error signal drives the head (Figs. 3*A* and 7*A*). However, to constrain all eye positions to the oculomotor range, the desired gaze error is first transformed into a desired eye-in-head position goal. In this way, the model appears to revive Robinson's original proposal (Robinson 1973, 1975; Van Gisbergen et al. 1981) that the pontine oculomotor burst generator is driven by a desired eye-position signal, albeit through an entirely different transformation sequence, as the common drive for the eyes and head is now derived from the oculocentric SC motor map.

4) In classical saccadic eye-movement models (Jürgens et al. 1981; Scudder 1988; Van Gisbergen et al. 1981), the brain stem oculomotor burst generator produces the eye-velocity signal, which appears as a pulse on the oculomotor neurons (the direct pathway), while its integral provides the tonic step-signal to hold the eye in its final position. Note that in our eye-head gaze model, this is no longer the case: the oculomotor burst cells instead generate a desired eye-velocity signal. This signal is modified by the VOR during the gaze shift (*Eq. 9*), to produce the actual eye-in-head velocity for the oculomotor neurons. As a consequence, our model predicts that the output of the oculomotor burst generator, $B(t)$, should correlate with both the eye- and head-velocity signals:

$$B(t) = p \cdot \dot{E}_H(t - \tau_E) + q \cdot \dot{H}_W(t - \tau_H) \qquad (13)$$

with τ_E and τ_H short delays, and p and q coefficients that depend on the gain of the VOR. Such a systematic correlation has, indeed, been observed across the population of horizontal burst cells in the pons (Cullen and Guitton 1996). Yet, as p and q are not identical, vary from cell to cell, and from movement to movement, the pontine burst does not seem to encode horizontal gaze velocity. The latter depends on the involvement of the VOR: $B(t)$ equals gaze velocity when $g_V = 1$, i.e., when the VOR is fully engaged. However, during rapid gaze shifts, the VOR gain is typically low (Fig. 4*A*, *inset*; e.g., Laurutis and Robinson 1986; Tabak et al. 1996).

Alternative Spatial Updating Strategy

The world-centered spatial updates of *Eqs. 1–3*, and Fig. 2 (left-hand side) rely on instantaneous feedback of (absolute) signals about the eye-in-head and the head-on-neck orientations. A potential problem with this scheme, however, already emphasized by the work of Goldberg and colleagues (e.g., Goldberg and Bruce 1990) is the lack of neural signatures reflecting a world-centered (or even craniocentric) reference frame. Neural responses in the brain stem, midbrain superior colliculus, parietal cortex, and frontal eye fields, all seem to possess signals expressed in oculocentric (displacement) coordinates. This has prompted researchers to propose spatial updating models for saccades, which rely entirely on relative (eye-) displacement signals, rather than on feedback from absolute eye position (Goldberg and Bruce 1990). It is conceivable, however, that spatial reference frames are encoded by large populations of neurons, rather than by the response properties of single neurons. For example, subtle multiplicative eye-position gain-field modulations of the firing rates of neurons with retinocentric receptive fields could potentially embed the neural representation for a craniocentric target position (e.g., Andersen et al. 1985; Zipser and Andersen 1988). Similarly, if head orientation would also modulate such neurons, the population code could represent a world-centered target. Although gain-field eye-position modulations have been reported for several stages in the oculomotor pathways of head-restrained monkeys (e.g., parietal cortex, Andersen et al. (1985); superior colliculus, Stuphorn et al. 2000; van Opstal et al. 1995), similar gain-field modulations of head position have not yet been studied in head-unrestrained animals, and, therefore, remain speculative.

Still, keeping targets in absolute world coordinates is not the only way in which the system could ensure dynamic spatial accuracy during fast eye-head gaze shifts. As indicated in Table 1, the calculations could also be based (at least in part) on feedback from (relative) dynamic motor errors of the ongoing gaze and head movements. For head-restrained saccades, such models have been around since the early eighties (Duhamel et al. 1992; Goldberg and Bruce 1990; Jürgens et al. 1981; Scudder 1988). If also applicable to head-unrestrained gaze control, there would be no need for a suprasensory mapping stage of the goal into common world coordinates. Instead, the goal would be kept in its own sensory reference frame: retinocentric for visual targets and craniocentric for auditory targets. How could this be implemented?

The central defining feature of our model remains the SC motor map, which programs a feedforward desired dynamic gaze-velocity signal for brain stem and spinal-cord circuits. Like in Scudder's (1988) saccade model, this gaze-velocity signal is compared with the current gaze velocity, by integrating the difference. This comparison yields the dynamic gaze motor-error, $G_{\mathrm{ERR}} = \Delta G_1(t^*)$. The presence of such a signal in

the system, therefore, raises an alternative possibility to maintain spatial accuracy, without transforming the target into a suprasensory reference frame. For a visual target, it would, thus, suffice to subtract the dynamic gaze error from the retinal target location, perceived at time t^* (Fig. 12):

$$\Delta G_2(t^*) = T_E^* - G_{ERR}(t^*) \quad \text{and } \Delta H_2 = \Delta G_2 + E_1 \quad (14)$$

Likewise, auditory targets can be kept in their original craniocentric coordinates, as its head-motor error, $H_{ERR}(t)$, is constructed from the dynamic gaze-motor error, by adding the eye position, estimated at t^* (Fig. 3*A*). Also, this signal would be available within the system, as it was required to prevent the eyes from hitting the limits of the oculomotor range. Hence, for sounds, the dynamic updating algorithm reads:

$$\Delta H_2(t^*) = T_H^* - H_{ERR}(t^*) \quad \text{and } \Delta G_2 = \Delta H_2 - E_1 \quad (15)$$

The multisensory updating process described by *Eqs. 14* and *15* is reminiscent to the proposal of relative coordinate transformations, forwarded earlier by Jürgens et al. (1981) and Goldberg and Bruce (1990), for head-restrained visual-evoked saccades. Here, that concept is extended in several ways, to enable dynamic updating of target locations for auditory and visual eye-head gaze shifts.

In the earlier proposals, dynamic feedback of eye-motor error was used exclusively to drive the ocular kinematics, encoded by the brain stem oculomotor burst generator (through the so-called local feedback loop). Instead, in our extended model, the instantaneous gaze- and head-motor errors are used to update target locations in the global feedback loop as well (Van Gisbergen and van Opstal 1989). Note, however, that accurate knowledge about the dynamic eye-in-head position, $E_H(t)$, and about the eye position at the start of the gaze shift, E_1, is still required for either updating strategy. The need for these signals would explain why sound-evoked eye-head gaze shifts to pure tones vary in a systematic way with the initial eye position, despite the fact that changes in eye position have no influence on the sensory (acoustic) input (Van Grootel et al. 2012).

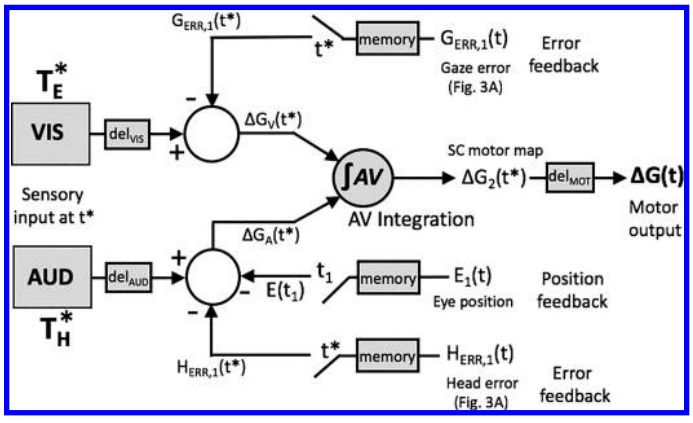

Fig. 12. Dynamic spatial updating schema relying on current gaze- and head-motor errors. Under multisensory conditions, audiovisual integration takes place in a unified, oculocentric reference frame, specifying the desired gaze-displacement vector for the selected target. Note that information about absolute eye position is still required to ensure goal-directed head movements. For reference, the scheme also includes the different delays that should be accounted for in the updating process: different sensory delays for visual and auditory inputs, and a motor delay for the gaze shift. The feedback signals, thus, have to be stored in short-term memory buffers, from which the correct timing relative to the stimulus event (t^*) may be retrieved to ensure accurate updating also in dynamic localization conditions, such as in Figs. 1 and 11.

Multisensory Integration

Although not explicitly modeled in this paper, auditory and visual inputs are combined at the SC motor map (stage "AV integration"). This notion is in line with the idea that the midbrain SC is a multisensory-motor gateway for gaze orienting (Bell et al. 2005; Frens and van Opstal 1998; Groh and Sparks 1996; Jay and Sparks 1987; Meredith and Stein 1986; Stein and Meredith 1993). In the scheme of Fig. 12, both sensory modalities express the goal in unified, oculocentric coordinates. The multisensory integration stage is thought to incorporate different factors, which have been shown to include the reliability of each sensory modality (target uncertainty, noise, and variability). Psychophysical evidence has shown that in simple two-stimulus AV localization paradigms, the system constructs a multimodal percept that, in the absence of any prior information (i.e., a uniform prior), resembles a weighted average of the visual and auditory target estimates. The unisensory response variances then act as weighting factors (e.g., Alais and Burr 2004; Hillis et al. 2002). In the AV stage in our scheme this would translate to

$$\Delta G_{AV}(t^*) = \frac{\sigma_A^2 \Delta G_V(t^*) + \sigma_V^2 \Delta G_A(t^*)}{\sigma_A^2 + \sigma_V^2} \quad (16)$$

in which σ_X^2 represents the uncertainty (variance) of sensory modality $X = (A,V)$. *Eq. 16* states that when uncertainty about the visual target is much smaller than for audition ($\sigma_V^2 \ll \sigma_A^2$), the weighted average estimate, ΔG_{AV}, will be close to the visual estimate, ΔG_V, and vice versa. Conversely, if both modalities are equally noisy, the bimodal estimate corresponds to the average of the unisensory estimates.

Because the construction of the goal is embedded within dynamic sensorimotor feedback loops, our model further proposes that the multisensory integration stage does not only rely on the (multi-)sensory coordinates, but should also incorporate dynamic information about gaze- and head-motor errors. Clearly, given the different coordinate-transformation problems for visual and auditory stimuli with respect to the eyes and head (Table 1), this poses interesting challenges for multisensory integration, which, so far, have received little attention in the literature (Corneil and Munoz 1996; Corneil et al. 2002; Stein and Meredith 1993; van Opstal 2016; Van Wanrooij et al. 2009).

A further interesting consequence of this concept is that the estimates of the unimodal variances in *Eq. 14* may be time-dependent too, i.e., $\sigma_X^2 = \sigma_X^2(t^*)$. For example, the longer the multisensory processing time (say, for long reaction times), the smaller the sensory uncertainties, and, hence, the lower the variance in the sensory estimates. However, it is likely that this process strongly depends on the sensory modality, as the sensory processing times (and internal delays) for vision and audition are markedly different.

So far, however, multisensory integration studies have typically dealt with the processing of audiovisual stimuli under static and aligned localization sensorimotor conditions only (as exemplified by Fig. 1*A*). New experimental

data are needed that reveal the dynamics of multisensory integration in paradigms that require rapid eye-head spatial updating like in Fig. 1*D*, and under conditions in which eye, head, and audio-visual stimuli may be either aligned or misaligned in space and time.

Glossary

$B(t)$	Oculomotor burst generator output (desired eye velocity)
D	Burst duration of SC output
E_{DES}	Desired eye-in-head position within the OMR
G^*, H^*	Ongoing gaze and head positions at target presentation
$\dot{G}_{DES}(t)$	Gaze-velocity pulse from SC motor map (rectangular pulse)
g_H	Attenuating gain for the head-motor error vector
G_0, H_0, E_0	Initial gaze, head and eye-in-head positions
G_{PK}	Peak activity of SC output burst (*Eq. 5*)
$G(t)$, $H(t)$, $E(t)$	Gaze-, head-, and eye position at time t
$\dot{G}(t)$,$\dot{H}(t)$,$\dot{E}(t)$	Gaze-, head-, and eye velocity at time t.
$g_v(G_{ERR})$	Gaze-error-dependent VOR gain (Fig. 4*A*)
G_W, H_W, E_H	Gaze (eye-in-world), head-in-world, and eye-in-head vectors
OMR_H, OMR_V	Horizontal, vertical oculomotor ranges around straight ahead
SC	Superior colliculus
t^*	Time of target presentation during ongoing gaze shift
T_E, T_H	Retinal (visual) and head-centric (auditory) target coordinates
T_W^V,T_W^A	Visual and auditory targets-in-world
T_W	Target in world coordinates
α, ε	Azimuth and elevation angles (deviation from straight ahead)
$\Delta E_{G,0}$	Initial eye-position component in the direction of ΔG
ΔG^*, ΔH^*	Ongoing gaze and head motor errors at target presentation
ΔG_{DES}, $G_{ERR}(t)$	Desired gaze displacement, and current gaze motor error
ΔH_{DES}, $H_{ERR}(t)$	Desired head displacement, current head motor error
$\Delta t_H(MOD, \Delta E_{G,0})$	Modality, and eye position-dependent head-onset delay

GRANTS

This work was supported by the European Commission through an FP7 Marie Curie ITN project "NETT" (Grant 289146 to B. Kasap), and through a Horizon 2020 ERC Advanced Grant for project "ORIENT" (Grant 693400 to A. J. van Opstal).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

B.K. analyzed data; B.K. and A.J.V.O. interpreted results of experiments; B.K. and A.J.V.O. drafted manuscript; B.K. and A.J.V.O. edited and revised manuscript; B.K. and A.J.V.O. approved final version of manuscript; A.J.V.O. conceived and designed research; A.J.V.O. prepared figures.

REFERENCES

Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14: 257–262, 2004. doi:10.1016/j.cub.2004.01.029.

Andersen RA, Essick GK, Siegel RM. Encoding of spatial location by posterior parietal neurons. *Science* 230: 456–458, 1985. doi:10.1126/science.4048942.

Bell AH, Meredith MA, van Opstal AJ, Munoz DP. Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *J Neurophysiol* 93: 3659–3673, 2005. doi:10.1152/jn.01214.2004.

Bremen P, Van der Willigen RF, Van Wanrooij MM, Schaling DF, Martens MB, Van Grootel TJ, van Opstal AJ. Applying double-magnetic induction to measure head-unrestrained gaze shifts: calibration and validation in monkey. *Biol Cybern* 103: 415–432, 2010. doi:10.1007/s00422-010-0408-4.

Corneil BD, Munoz DP. The influence of auditory and visual distractors on human orienting gaze shifts. *J Neurosci* 16: 8193–8207, 1996. doi:10.1523/JNEUROSCI.16-24-08193.1996.

Corneil BD, Van Wanrooij M, Munoz DP, van Opstal AJ. Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J Neurophysiol* 88: 438–454, 2002. doi:10.1152/jn.2002.88.1.438.

Cullen KE, Guitton D. Inhibitory burst neuron activity encodes gaze, not eye, metrics and dynamics during passive head on body rotation. Evidence that vestibular signals supplement visual information in the control of gaze shifts. *Ann N Y Acad Sci* 781: 601–606, 1996. doi:10.1111/j.1749-6632.1996.tb15735.x.

Daye PM, Optican LM, Blohm G, Lefèvre P. Hierarchical control of two-dimensional gaze saccades. *J Comput Neurosci* 36: 355–382, 2014. doi:10.1007/s10827-013-0477-1.

Duhamel JR, Colby CL, Goldberg ME. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255: 90–92, 1992. doi:10.1126/science.1553535.

Freedman EG. Interactions between eye and head control signals can account for movement kinematics. *Biol Cybern* 84: 453–462, 2001. doi:10.1007/PL00007989.

Frens MA, Van Opstal AJ. Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Res Bull* 46: 211–224, 1998. doi:10.1016/S0361-9230(98)00007-0.

Galiana HL, Guitton D. Central organization and modeling of eye-head coordination during orienting gaze shifts. *Ann N Y Acad Sci* 656: 452–471, 1992. doi:10.1111/j.1749-6632.1992.tb25228.x.

Goldberg ME, Bruce CJ. Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *J Neurophysiol* 64: 489–508, 1990. doi:10.1152/jn.1990.64.2.489.

Goossens HHLM, van Opstal AJ. Human eye-head coordination in two dimensions under different sensorimotor conditions. *Exp Brain Res* 114: 542–560, 1997. doi:10.1007/PL00005663.

Goossens HHLM, Van Opstal AJ. Dynamic ensemble coding of saccades in the monkey superior colliculus. *J Neurophysiol* 95: 2326–2341, 2006. doi:10.1152/jn.00889.2005.

Goossens HHLM, van Opstal AJ. Optimal control of saccades by spatial-temporal activity patterns in the monkey superior colliculus. *PLOS Comput Biol* 8: e1002508, 2012. doi:10.1371/journal.pcbi.1002508.

Groh JM, Sparks DL. Saccades to somatosensory targets. III. Eye-position-dependent somatosensory activity in primate superior colliculus. *J Neurophysiol* 75: 439–453, 1996. doi:10.1152/jn.1996.75.1.439.

Guitton D. Control of eye-head coordination during orienting gaze shifts. *Trends Neurosci* 15: 174–179, 1992. doi:10.1016/0166-2236(92)90169-9.

Guitton D, Volle M. Gaze control in humans: eye-head coordination during orienting movements to targets within and beyond the oculomotor range. *J Neurophysiol* 58: 427–459, 1987. doi:10.1152/jn.1987.58.3.427.

Hillis JM, Ernst MO, Banks MS, Landy MS. Combining sensory information: mandatory fusion within, but not between, senses. *Science* 298: 1627–1630, 2002. doi:10.1126/science.1075396.

Jay MF, Sparks DL. Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *J Neurophysiol* 57: 35–55, 1987. doi:10.1152/jn.1987.57.1.35.

Jürgens R, Becker W, Kornhuber HH. Natural and drug-induced variations of velocity and duration of human saccadic eye movements: evidence for a control of the neural pulse generator by local feedback. *Biol Cybern* 39: 87–96, 1981. doi:10.1007/BF00336734.

Kasap B, van Opstal AJ. A spiking neural network model of the midbrain superior colliculus that generates saccadic motor commands. *Biol Cybern* 111: 249–268, 2017. doi:10.1007/s00422-017-0719-9.

Knudsen EI, Konishi M. Mechanisms of sound localization in the barn owl (*Tyto alba*). *J Comp Physiol* 133: 13–21, 1979. doi:10.1007/BF00663106.

Laurutis VP, Robinson DA. The vestibulo-ocular reflex during human saccadic eye movements. *J Physiol* 373: 209–233, 1986. doi:10.1113/jphysiol.1986.sp016043.

Lefèvre P, Galiana HL. Dynamic feedback to the superior colliculus in a neural network model of the gaze control system. *Neural Netw* 5: 871–890, 1992. doi:10.1016/S0893-6080(05)80084-X.

Meredith MA, Stein BE. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol* 56: 640–662, 1986. doi:10.1152/jn.1986.56.3.640.

Robinson DA. Models of the saccadic eye movement control system. *Kybernetik* 14: 71–83, 1973. doi:10.1007/BF00288906.

Robinson DA. Oculomotor control signals. In: *Basic Mechanisms of Ocular Motility and Their Clinical Implications*, edited by Lennerstrand G, Bachy-Rita P. Oxford, UK: Pergamon, 1975, p. 337–374.

Scudder CA. A new local feedback model of the saccadic burst generator. *J Neurophysiol* 59: 1455–1475, 1988. doi:10.1152/jn.1988.59.5.1455.

Stein BE, Meredith MA. *The Merging of the Senses*. Cambridge, MA: MIT Press, 1993.

Stuphorn V, Bauswein E, Hoffmann KP. Neurons in the primate superior colliculus coding for arm movements in gaze-related coordinates. *J Neurophysiol* 83: 1283–1299, 2000. doi:10.1152/jn.2000.83.3.1283.

Tabak S, Smeets JB, Collewijn H. Modulation of the human vestibuloocular reflex during saccades: probing by high-frequency oscillation and torque pulses of the head. *J Neurophysiol* 76: 3249–3263, 1996. doi:10.1152/jn.1996.76.5.3249.

Tweed D. Three-dimensional model of the human eye-head saccadic system. *J Neurophysiol* 77: 654–666, 1997. doi:10.1152/jn.1997.77.2.654.

Van Gisbergen JAM, Robinson DA, Gielen S. A quantitative analysis of generation of saccadic eye movements by burst neurons. *J Neurophysiol* 45: 417–442, 1981. doi:10.1152/jn.1981.45.3.417.

Van Gisbergen JAM, van Opstal AJ. Models of saccadic control. In: *The Neurobiology of Saccadic Eye Movements*, edited by Wurtz RH and Goldberg ME. Amsterdam, the Netherlands: Elsevier, 1989, p. –69–101.

Van Grootel TJ, Van der Willigen RF, van Opstal AJ. Experimental test of spatial updating models for monkey eye-head gaze shifts. PLoS One 7: 47606, 2012. doi:10.1371/journal.pone.0047606.

Van Grootel TJ, Van Opstal AJ. Human sound-localization behaviour after multiple changes in eye position. *Eur J Neurosci* 29: 2233–2246, 2009. doi:10.1111/j.1460-9568.2009.06761.x.

van Opstal AJ. *The Auditory System and Human Sound Localization Behavior*. Amsterdam, the Netherlands: Elsevier, Academic, 2016.

van Opstal AJ, Goossens HHLM. Linear ensemble-coding in midbrain superior colliculus specifies the saccade kinematics. *Biol Cybern* 98: 561–577, 2008. doi:10.1007/s00422-008-0219-z.

Van Opstal AJ, Hepp K, Suzuki Y, Henn V. Influence of eye position on activity in monkey superior colliculus. *J Neurophysiol* 74: 1593–1610, 1995. doi:10.1152/jn.1995.74.4.1593.

Van Wanrooij MM, Bell AH, Munoz DP, van Opstal AJ. The effect of spatial-temporal audiovisual disparities on saccades in a complex scene. *Exp Brain Res* 198: 425–437, 2009. doi:10.1007/s00221-009-1815-4.

Vliegen J, Van Grootel TJ, Van Opstal AJ. Dynamic sound localization during rapid eye-head gaze shifts. *J Neurosci* 24: 9291–9302, 2004. doi:10.1523/JNEUROSCI.2671-04.2004.

Vliegen J, Van Grootel TJ, Van Opstal AJ. Gaze orienting in dynamic visual double steps. *J Neurophysiol* 94: 4300–4313, 2005. doi:10.1152/jn.00027.2005.

Zipser D, Andersen RA. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* 331: 679–684, 1988. doi:10.1038/331679a0.