MODELING HEAD TRACKING OF VISUAL TARGETS

K. J. Chen¹, E. A. Keshner², B. W. Peterson³ and, T. C. Hain⁴

¹Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208, USA

²Sensory Motor Performance Program, Rehabilitation Institute of Chicago, Chicago, IL 60611, USA

³Department of Physiology, Northwestern University Medical School, Chicago, IL 60610, USA

⁴Department of Otolaryngology, Northwestern Memorial Hospital, Chicago, IL 60611, USA

Running Head: Head Tracking of Visual Targets

Address correspondence to:
Kelvin J. Chen
Sensory Motor Performance Program
Rehabilitation Institute of Chicago
Room 1406
345 East Superior Street
Chicago IL 60611
e-mail: kelvin@ieee.org
Fax: 312-238-2208
Abstract

Control of the head involves somatosensory, vestibular, and visual feedback. The dynamics of these three feedback systems must be identified in order to gain a greater understanding of the head control system. We have completed one step in the development of a head control model by identifying the dynamics of the visual feedback system. A mathematical model of human head tracking of visual targets in the horizontal plane was fit to experimental data from seven subjects performing a visual head tracking task. The model incorporates components based on the underlying physiology of the head control system. Using optimization methods, we were able to identify neural processing delay, visual control gain, and neck viscosity parameters in each experimental data subject.

Keywords: mathematical model, head control, vision, gaze.
Introduction

Visually guided tracking of targets is a skill used in nearly all activities of daily living in order to maintain gaze on objects moving with respect to the eyes. Tracking may be performed with the eyes, head, trunk, entire body, or even by control of a vehicle. Depending on the behavioral context, one or more of these body segments may be engaged. For example, when looking at a computer screen, the eyes alone may be engaged. When watching a tennis match, the eyes and head might be used to maintain gaze on the ball. When preparing to catch a ball, the entire body might be oriented towards the ball. For optimal performance, the activity of all body segments must be coordinated to provide gaze stabilization.

Most activities require only coordinated movement of the head and eyes for gaze stabilization. Eye tracking has been studied extensively [1,2]. The behavior of the eye is dominated by muscular control as torques produced by inertia, and the passive viscoelastic properties of soft tissue are relatively small compared to the torques that can be generated by the ocular muscles. Accordingly, models of visual pursuit need to mainly consider central nervous system control. Generally, they have focused on methods of dealing with biological delays in the context of the powerful actuators provided by the ocular muscles [3-6].

Relatively little experimental attention has been focused on head tracking despite its large contribution to gaze stabilization. Because of its greater inertial properties, greater viscoelastic properties, and more complex kinematics, control of the head involves many more factors than the eye. Work by Guitton and associates [7], as well as two later studies of our group [8,9] characterized the overall dynamics of head control when tracking a randomly moving target with the head. Head with respect to trunk transfer functions resembled a second order system. Subjects were able to effectively track the target at frequencies below 1 Hz, but at higher
frequencies, subjects exhibited decreasing head gains and phases. These studies did not attempt to describe in detail the underlying neural and mechanical contributions to the tracking response.

Compared to the eye, the head is underpowered and behavior is more related to inertia and the viscoelastic properties of the neck, especially at high frequencies [7-13]. Biological delays related to sensory input and motor output are greater for the head than the eye due to longer distances over which neurological impulses must travel [14]. Furthermore, spatial orientation processing involving coordinate transformations between eye, head and trunk-based coordinate systems may require additional delays [15-18]. Nevertheless, head tracking is an activity that must be effective, coordinated with eye tracking, and somewhat predictable so that processes that compensate for delays can extrapolate head position and produce appropriate commands to optimize tracking behavior [14,19].

Existing models of head control do not explicitly define contributions from the visual system [12,20,21]. The lack of an explicit solution may relate to the complexity of head control that requires consideration of multiple sensory inputs (vestibular, proprioceptive and visual), voluntary input, and biomechanics. While experimental data exists describing the frequency characteristics of the visual system in head control [7-9], the more complete understanding that comes from model based simulations and reproduction of results has so far not been accomplished. In order to gain more insight into visually guided head control, we developed a physiologically based mathematical model of visual tracking that extends a previously published model of head/neck control incorporating proprioceptive and vestibular control loops [20]. This augmented model was used to fit experimental visual head tracking data from previously published data from our laboratory [8].
Methods

Experimental Data

The experimental data that we modeled has been reported previously, but as the method by which it was obtained may have implications for our modeling results, we will briefly describe the experimental paradigm. Seven healthy subjects located at the center of a circular screen, sat with their head free to move. Subjects wore a head mounted laser that was projected onto the screen, and provided visual feedback of head position and motion. They were asked to track another laser target on the screen at the same level as the head target. In this situation, the task is to minimize error between a head mounted target and a target in space. The advantage of this paradigm is that eye movement need not be considered when modeling the responses. The laser target moved horizontally according to a sum of sines waveform (SSN) containing 10 frequencies ranging between 0.185 and 4.117 Hz. Amplitudes of component sine waves decreased from 20°/s at 0.185 Hz to 13°/s at 4.117 Hz. Peak velocity of the SSN was ±30°/s, and peak position was ±7°. Further methodological details can be obtained from [8]. The range of experimental responses is shown by the grayed area in Figure 1, summarizing data previously reported in [8]. Subjects showed a wide range of responses but all had a similar pattern relative to frequency. Gains close to 1 and phases close to 0° signify accurate tracking. Above 0.8 Hz, gains decreased at approximately 40 dB/decade, suggesting a 2nd order system. Also, the progressive phase lag that exceeds 180° suggests an additional delay operation.
**Model Development**

The general framework of the model was loosely based on previously published head control models [20,22]. It was implemented in a commercial mathematical modeling software package, Matlab/Simulink (Mathworks, Natick, MA). Figure 2a represents a simplified control systems representation of head control in the horizontal plane. A forward path represents the inertia and viscoelastic properties of the head/neck system and feedback pathways represent neurological control systems that process sensory input. Feedback pathways utilizing input from the vestibular, somatosensory, and visual systems provide reflex and voluntary control activation of neck muscles to move the head. In figure 2a, the dynamics of these pathways are contained within blocks named “vestibular feedback dynamics”, “somatosensory feedback dynamics”, and “visual feedback dynamics” respectively.

Our goal was to characterize the visual feedback dynamics. For simplicity, we abstracted from figure 2a the parts relevant to our experimental paradigm: visual feedback dynamics, torque converter, and head mechanics. The resulting model is shown in figure 2b. Implicit to this abstraction is the assumption that vestibular and somatosensory feedback does not contribute significantly to head responses for this experimental paradigm. There are several reasons that this might be the case. First, vestibular and somatosensory reflexes are not large contributors to head stability, which is dominated by mechanics. Second, these loops should ideally be turned off or cancelled out by feedforward circuits as these systems stabilize the head in space and would thus be antagonistic to visual tracking. It is well known that normal persons can reduce the gain of the vestibular ocular reflex during combined head-eye tracking, when it is inappropriate for the context [23,24]. Nevertheless, it is also possible that these antagonistic
reflexes are overcome by the visual tracking system. Our experimental paradigm did not provide enough data to differentiate between the possibilities that they are turned off or overcome by the visual system. Accordingly, we fit the data initially to a model configured with these reflexes turned off, and then we compared the results to the full model of 2a configured with the reflexes configured with normal gains in an attempt to infer the significance of these loops.

We hypothesized that input to the neurological visual head control system is *head velocity error*, which is the difference between the relevant visual input target velocity and head velocity. There must also be a delay in the signal due to neural conduction times and neural processing. This time delay element is shown in Figure 2b as the *Neural Delay*. The time-delayed signal is processed by what we call the *Visual Head Controller*, the output of which represents the neural command signal to the muscles of the neck. This is shown in Figure 2b as the Visual Head Controller box. The command signal must then be transformed into torque on the head. The operation is represented in Figure 2b as the *Torque Converter*. The torque produced by neck muscles acts on the head/neck structure, represented as the *Neck Plant*, to produce head motion output.

To be consistent with the previous models of visual tracking [5,14,25], we chose the time dimension of the visual input to be velocity. Although not explicitly represented, it seems most likely that signals representing retinal slip velocity of both laser targets are subtracted centrally, providing a target velocity with respect to head velocity signal. We constructed this signal from ideal representations of head and target velocity and used it to drive the visual subsystem.

The delay component, $N_{\text{delay}}$, represents the lumped processing and neural conduction delays that must exist within the visual head control system. Delays of 100-250 ms have been measured in ocular pursuit studies with unpredictable stimuli [4,14,18,26-28]. Because the head
receives the input signal from the eyes, we would expect similar delays in the head control system.

The visual controller represents the lumped neural processing dynamics that transforms the head velocity error signal to a neural activation command signal to neck muscles. For clarity, we separated the gain term, $K_{vis}$, from the dynamic terms of the controller. There is no experimental data concerning this central processor, so we designed the controller as the simplest possible circuit having the necessary dynamics to produce the second order output response of the head observed in the experimental data.

The Torque Converter and Neck Plant are identical to those in the model of head stabilization published by Peng and associates [20,29]. The Torque Converter represents muscle contraction dynamics. The Neck Plant represents the steady state dynamics of the head/neck system. This not only includes the passive properties of soft tissue, but non-time varying active muscle properties as well. The $I$ parameter represents inertia of the head. The $B$ and $K$ parameters represent the lumped damping and elasticity properties of soft tissue surrounding the neck.

The model was implemented using the Matlab mathematical software (Mathworks, Natick, MA) with the Simulink toolbox. To fit model simulations to experimental data, we used optimization (Matlab Optimization Toolbox) to iteratively adjust parameters to minimize an error function based on the magnitude of the difference between experimental data and simulation results. There are four parameters whose values may differ across subjects, and therefore should be optimized. $N_{delay}$ (neural delay) and $K_{vis}$ (visual gain) were optimized because experimentally they are unknown and clearly critical to reproduce behavior. $K$ (joint stiffness) and $B$ (joint viscosity) are parameters that subjects may be able to control. In the
frequency range tested, head inertia exerts only a small contribution to head dynamics.

Consequently, the neural delay and remaining first order dynamics in the neck plant dominate the response.

A unique solution cannot be found with four free parameters. The maximum number of free parameters that will yield a unique solution in this situation is three. The mathematical redundancy is clear when you consider only the transfer function of the first order dynamics of the neck plant with the visual gain:

\[ \frac{K_{vis}}{s^2 + Bs + K} \approx \frac{K_{vis}}{Bs + K} \text{ for } B, K >> 1 \]  

(1)

Both the numerator and denominator can be scaled by a constant to yield different values of \(K_{vis}, B, \) and \(K,\) resulting in an equivalent equation. We eliminated one free parameter by assigning a value of 1 N/rad to \(K.\) Peng and associates [20] estimated \(K\) to be approximately 2 N/rad in a mental distraction paradigm. For the voluntary paradigm of this experiment, it is reasonable to assume a lower stiffness since stiffness imposes an opposing force to voluntary movement.

**Results**

Results of model optimization are shown in Table 1. The model was able to fit the experimental data well with percent variance accounted for values greater than 80% in all but subject S1. This subject had a low gain of 0.26 and phase lead of 27° at 0.185 Hz. Since it is unlikely that a subject can lead a pseudo-randomly moving target, the phase lead is more likely due to noise. When we removed this low frequency point from the model optimization, a fit with 81% VAF was obtained.
As expected from the variability of the experimental data, parameter values varied greatly. The neural delay ($N_{\text{delay}}$) parameter ranged from 153 ms to 252 ms, which is reasonable given delays of 100 ms to 250 ms measured in visual pursuit studies [4,14,18,26-28]. Neck plant viscosity ($B$) ranged from 0.22 to 1.88 Nsec/rad. These values are higher than reported in previous studies [20,30-32], and may indicate that the $B$ parameter may be representing other properties of the visual head controller that are otherwise not modeled. The visual gain ($K_{\text{vis}}$) ranged from 0.14 to 4.25. At frequencies below 0.8 Hz where voluntary control is effective, greater ratios between $K_{\text{vis}}$ and $K$ were related to higher gains.

To help us understand how the visual controller gain and neural delay effect control of the head, model simulations were performed using optimization results from subject S4. Figure 1 illustrates the simulated frequency response (Bode plot) of the head. The model (S4 fit) shows a good fit to the experimental data. When we increased the visual control gain by 2 times, shown by lines labeled "$K_{\text{vis}}*2$", response gains increased. However, at approximately 1.5 Hz, there is a resonant peak in gain that is greater than 1, indicating greater head motion than target motion. The line labeled '0 delay' shows the effect of removing the neural delay. Removing the neural delay results in less phase lag, particularly above 0.8 Hz, with a slight damping of the resonant peak in the gain.

The inset in figure 1 shows a simulated time domain response of the head to a step change in target velocity. It illustrates some points better than the Bode plot. The overshoot is typical of an underdamped second order system. The steady-state response has a gain of 0.4. Experimentally, the step-response task is not comparable to our SSN waveform. The step response has a wider bandwidth, and the steady-state portion of the response is predictable.
Experimentally, a step waveform would be likely to elicit a head-saccade rather than a tracking response.

The simulations above were performed with the vestibulocollic and cervicocollic reflexes turned off. As mentioned in the methods, turning off these reflexes might be behaviorally desirable in this paradigm, as they would oppose head tracking. Unfortunately, the present experimental data is not adequate to determine whether they are still operating. In order to address this question, we took the optimized model and configured it with normal reflexes \((K_{vcr}=20, K_{ccr}=0.1)\) [20]. When this was done, phase was unaffected, and gain was only slightly affected, being reduced by a maximum of 3 DB. This suggests that this model is insensitive to the influence of the vestibulocollic and cervicocollic reflexes, and that they can be neglected for this experimental paradigm.

**Discussion**

We have presented a simple model of visually guided head control. The model was able to fit the experimental data well. Thus, it is a feasible representation of visual head control, and allows us to gain insight into characteristics of head control that experimental data alone is unable to provide. The simple dynamic properties of the experimental data preclude the identification of a more complex model.

The wide range of experimental data and parameter values suggests that these parameters may be adjustable, perhaps under conscious control. Subjects may set these parameters based on their behavioral goal, physical ability, and perturbation characteristics. For example, if a subject is anxious about joint pain, he may increase his neck stiffness and viscosity to limit motion. Our group found in a previous study of reflex control of the head that subjects modulate neck plant
elasticity and the vestibulocollic reflex gain to compensate for additional inertia placed on the head [32]. Future experiments may test the ability to modulate visual head control parameters.

Because \( K \) was fixed, the optimized \( B \) and \( K_{vis} \) values do not necessarily represent the underlying physiological properties. As \( K \) is varied, there is always a \( B \) and \( K_{vis} \) that provides the same optimal fit due to the redundancy in equation 1. This defines a line of optimal fits in the 3-dimensional parameter space, rather than a single best fit point. In order to obtain a unique fit for \( K \) in addition to \( B \) and \( K_{vis} \), experimental data is needed from higher frequencies, where the effect of inertia is more prominent. Although unique values of \( K, B, \) and \( K_{vis} \) could not be identified in our experiment, their relative ratios are meaningful. For any given \( K \), the ratio of \( K/B \) and \( K_{vis}/K \) together with the optimized \( N_{delay} \) uniquely characterizes the response.

The experimental data show a wide range of gains with low phase lags at frequencies below 0.8 Hz. This is the frequency range where voluntary control is most effective [8]. The model predicts that response gain in this frequency range is determined by the ratio between \( K_{vis} \) and \( K \). Since the optimized value of \( B \) is related to \( K \), it is also true that performance increases with a greater \( K_{vis} \) to \( B \) ratio. Larger ratios correspond to greater gains and greater performance. The variability of responses may be indicative of individual effort level or physical limitations.

The ability to track the target depends on accurate detection of error between the head and target, and an individual’s ability to correct for the error. Our model assumes perfect detection of error. This may not be a realistic assumption as ocular pursuit studies have found amplitude and frequency limitations to eye tracking [33]. The anomalous neck plant viscosity values may be indicative of dynamics of visual error detection that were not modeled, or other central processes that was represented by our simple model as additional viscosity. A more
detailed model of visual error detection is needed to investigate the effect of visual error
detection on head tracking.

The ability to correct for error depends on neck viscoelasticity, visual controller gain, and
response delay. The viscoelastic forces produced by the neck plant that the head controller must
counteract may be reduced in several ways. Reducing the level of cocontraction of neck
muscles, changing active muscle groups, and changing active muscle fiber types may change the
viscoelastic properties of the neck joint. Increasing the head controller gain relative to the
viscoelastic properties of the neck, as shown in Figure 1, improves the gain (i.e. amplitude)
performance. This may be accomplished by increasing the level of effort during the task.
Decreasing response delay increases performance by decreasing the phase lags. Response delay
may be changed by practice or increasing alertness.

Neither the model nor experimental data illustrate “perfect” tracking at low frequencies,
i.e. with a gain of one and phase shift of 0, although intuitively one might expect that low-
frequency gain should be one. There are several possible explanations. First, the visual tracking
system that we have studied here might not normally stand alone. In particular, at low
frequencies, tracking might be supplemented by head saccades and prediction. Second, there
may be interactions between the experimental paradigm and frequency response. Barnes and
Lawson noted that for head-free pursuit of SSN targets that gain of lower frequency responses
was dependent on the relative proportion of low and high frequency components[33]. If head
tracking has similar dependence on the frequency content of the SSN input, then the
experimental findings of gain less than one at low frequencies could be an artifact of the
experimental paradigm. Nevertheless, at this writing there is no experimental data available for
frequencies below 0.1 Hz. An SSN input limited to frequencies below 0.1 Hz would test this prediction and might be a useful direction for future investigation.

Our model is a feasible representation of visual head control that generally reflects what is known about the underlying physiology. The identification of this system is a necessary step in developing a complete model of the head control system. Visual input is just part of several partially redundant sensory inputs to the head. For example, static visual field information produces small differences in head stability when absent or present [7-9]. Nevertheless, in disease processes and the elder population where other senses may be impaired, head stability and whole-body postural stability becomes more dependent on vision [7,13,34-44]. A better understanding of head control can provide insight into the more complex task of whole body posture and movement.
Acknowledgments

This study was supported in part by research grants DC01125 and DC02072 from the National Institute on Deafness and Other Communication Disorders, and NS22490 from the National Institute on Neurological Disorders and Stroke, National Institutes of Health.
References


Table 1. Optimized parameter values for model fit to experimental data shown in Figure 1.

<table>
<thead>
<tr>
<th>Subject</th>
<th>B</th>
<th>Kvis</th>
<th>Ndelay (sec)</th>
<th>vaf</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.25</td>
<td>0.14</td>
<td>0.248</td>
<td>64.4%</td>
</tr>
<tr>
<td>S2</td>
<td>1.71</td>
<td>1.43</td>
<td>0.153</td>
<td>86.4%</td>
</tr>
<tr>
<td>S3</td>
<td>0.36</td>
<td>0.45</td>
<td>0.163</td>
<td>84.7%</td>
</tr>
<tr>
<td>S4</td>
<td>0.22</td>
<td>0.63</td>
<td>0.154</td>
<td>95.4%</td>
</tr>
<tr>
<td>S5</td>
<td>0.60</td>
<td>0.59</td>
<td>0.239</td>
<td>80.3%</td>
</tr>
<tr>
<td>S6</td>
<td>0.73</td>
<td>1.66</td>
<td>0.252</td>
<td>90.8%</td>
</tr>
<tr>
<td>S7</td>
<td>1.88</td>
<td>4.25</td>
<td>0.239</td>
<td>90.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Kvis</th>
<th>Ndelay (sec)</th>
<th>vaf</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.82</td>
<td>1.31</td>
<td>0.207</td>
<td>84.6%</td>
</tr>
<tr>
<td>st. dev.</td>
<td>0.69</td>
<td>1.41</td>
<td>0.047</td>
<td>10.1%</td>
</tr>
</tbody>
</table>
**Figure Legends**

**Figure 1.** Bode plot of experimental data and model simulations of head with respect to target motion. Gains of 1 indicate head motion amplitude equal to target motion amplitude. Responses with phases of 0° indicate head motion in the same direction as target motion. Responses with phases at less than 0° indicate head motion direction that lags target motion. The gray area represents range of experimental responses from [8]. The model with optimized parameters (S4 fit) shows a good fit to the experimental data from S4. Increasing the visual control gain by a factor of 2 (Kvis*2) results in higher head response gains. Removing the neural processing delay results in less phase lag, especially above 0.8 Hz.

**Figure 2.** Simplified control systems model of horizontal plane head control (a). The forward path represents the inertia and viscoelastic properties of the head/neck system. Feedback pathways represent neurological control systems utilizing visual, vestibular, and somatosensory information. Model of visually guided head control (b). The model details the visual feedback control loop in (a). Input is visual target acceleration, output is head acceleration with respect to space. Parameters that were optimized were neural delay, visual control gain (Kvis), and lumped neck viscosity (B). Parameter values are given in Table 1.
Figure 1 is on next page
The image contains two plots with the same x-axis labeled "Frequency (Hz)", ranging from 0.1 to 10 Hz. The y-axis of the top plot is labeled "Gain" and ranges from 0 to 10 on a logarithmic scale, while the y-axis of the bottom plot is labeled "Phase (deg)" and ranges from -540 to 90 degrees.

The plots show data points and curves labeled "S4", "S4 fit", "S4 Kvis^2", and "S4 0 delay". The curves are shaded to indicate confidence intervals.

In the bottom inset graph, the x-axis is labeled "time (s)" and ranges from 0 to 3 seconds, while the y-axis is labeled "amplitude" and ranges from 0 to 0.6.
Figure 2 is on next page
Model of Horizontal Plane Head Control

Head Control Model Simplified for Visual Head Tracking