Title

Dynamic Estimation of Oncoming Vehicle Range and Range Rate: An Assessment of the Human Visual System's Capabilities and Performance

Author

Barton, Joseph Edward

Publication Date

2004

Dynamic Estimation of Oncoming Vehicle Range and Range Rate: An Assessment of the Human Visual System's Capabilities and Performance

by

Joseph Edward Barton

B.S. (University of Texas, Austin) 1974
M.S. (Rensselaer Polytechnic Institute) 1977
S.M. (Massachusetts Institute of Technology) 1984

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy
in

Engineering-Mechanical Engineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Theodore E. Cohn, Co-Chair
Professor Masayoshi Tomizuka, Co-Chair
Professor J. Karl Hedrick
Professor Brian Barsky

Fall 2004

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Dynamic Estimation of Oncoming Vehicle Range and Range Rate:
An Assessment of the Human Visual System's Capabilities and Performance

by

Joseph Edward Barton

Abstract

Dynamic Estimation of Oncoming Vehicle Range and Range Rate:
An Assessment of the Human Visual System's Capabilities and Performance

by

Joseph Edward Barton

Doctor of Philosophy in Engineering-Mechanical Engineering

University of California at Berkeley

Professor Theodore E. Cohn, Co-Chair
Professor Masayoshi Tomizuka, Co-Chair

The detection of impending collisions and the subsequent choice and regulation of maneuvers to deal with them are general problems of locomotor control that arise in many situations, both human and non-human. When an object moves towards an observer, the size of the image that it projects onto the retina of the observer's eyes increases, providing a powerful sensation of motion. Physiological and psychophysical research into this "looming" effect provides strong evidence for the existence of neural "looming detectors" that are used by humans and non-humans alike to detect and respond to oncoming objects. Automotive applications constitute an important context for the study of the visual perception of looming. To date, however, this aspect of the driver's performance has largely been neglected, and human driver models typically incorporate representations of the visual system that are based upon idealized behavior and in some cases questionable assumptions.

In this three part study we begin to address this deficiency by quantifying the visual system's ability to detect and track an object's approach, as represented by the rate of change of the angle θ that its image subtends on the retina of the eye. In the first part we tested a long-standing assumption of an absolute threshold in the human's ability to detect $d\theta/dt$, below which humans are unable to discern that θ is changing (and thus that a collision is imminent). The results provide evidence contradicting the threshold assumption, and indicate instead that the detection task is more accurately described as one of *signal detection* (detection of the signal $d\theta/dt$ in the presence of noise) with no threshold limitation. Collision *avoidance* requires that an observer accurately and

continuously track an approaching object's distance and closing speed. In the second part of this study we investigated the dynamic response of the visual system to changes in θ, employing both psychophysical and classical frequency response techniques. We found that the visual system exhibits a band-pass characteristic in this task that is well described by a linear, minimum phase, second order transfer function. Further analysis revealed that this aspect of the visual system exhibits a biphasic impulse response, which is the focus for the third part of our study. According to the model, certain pairs of "impulsive" stimuli presented in the proper sequence will reinforce one another, and thus be more easily detected, while others will cancel each other and be less so. This final series of experiments provided evidence consistent with this hypothesis.

The shortcomings of human driver models based upon current assumptions are discussed, and the development of improved models based the dynamic response characteristics of the visual system and the principles of signal detection are described. To focus our efforts we have assumed a fairly constrained driving scenario (the 'Lead Vehicle Braking" scenario), but these results are applicable to any scenario (automotive or not) in which the observer has an unobstructed view of the approach of an object or stationary obstacle.

1 INTRODUCTION

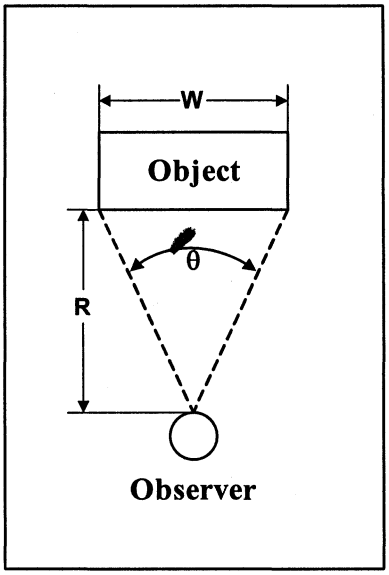
The detection of impending collisions and subsequent regulation of maneuvers to deal with them is a general problem of locomotor control that arises in many situations, both human and non-human. A bird landing on a perch, a snake striking at its prey, and a human driver braking to avoid a collision are all examples of this. When an object moves towards an observer the size of the image that it projects onto the retina of the observer's eyes increases. Even though this effect could be the result of an increase in the physical size of a stationary object, it provides a powerful sensation of motion. Much research has concentrated on this "looming" effect (1). Several psychophysical[1] studies (2, 3, 4) strongly suggest the presence of "looming detectors" in humans that interpret expanding retinal images as an approaching object. Furthermore, these studies find that the visual system is sensitive to looming in both an absolute ($\dot{\theta}=d\theta/dt$) and relative ($\dot{\theta}/\theta$) sense, where from Figure 1,

$$\frac{\theta}{2}=\tan^{-1}\left(\frac{W/2}{R}\right), \qquad \theta\approx\frac{W}{R} \qquad (\theta \text{ small}),$$

$$\dot{\theta}=-\frac{W\dot{R}}{R^2}=-\frac{\theta\dot{R}}{R}, \tag{1}$$

$$\frac{\dot{\theta}}{\theta}=-\frac{\dot{R}}{R}.$$

[1] *Psychophysics* is that branch of psychology that deals with the relationships between physical stimuli and sensory response.



Definition of Approach Velocity
Figure 1

The information provided by $\dot{\theta}$ and $\dot{\theta}/\theta$ alerts the observer that the object is moving towards him or her but it is not possible with this alone to sense the velocity of approach (or Range Rate) $\dot{R}$, which is necessary to deal effectively with the object's arrival (for example, by catching a ball or stopping just in time to avoid striking the object). This can be seen more clearly by rearranging the second of Equs. (1),

$$\dot{R} = -\frac{\dot{\theta}}{\theta} R . \qquad (2)$$

Information about the object's distance, or Range, (R) is also required, which cannot be provided by monocular cues alone. Binocular cues could help, but they are ineffective beyond about 10 m. They therefore do not explain how observers can respond to far-away, fast approaching objects—if the observer had to wait for the binocular cues to become effective there would in many cases be too little time left to react once they were received. The visual system must therefore be relying on some other cue or cues in the collision detection/avoidance task.

It is possible to extract information about the timing of an object's approach, however, from the simple relation [2]

$$\tau = \frac{\theta}{\dot{\theta}}. \tag{3}$$

τ, the "time to collision", is the time it will take the object to reach the observer. The ability of humans to perceive τ directly has been the subject of several psychophysical studies, the most compelling of which is that by Regan and Hamstra (5). Though these and others have provided persuasive evidence that humans' behavior in collision detection/response situations is *consistent* with a knowledge of τ, they fail to demonstrate conclusively that τ was being sensed directly and not inferred by other means (by estimating the necessary parameters, for example, or relying on experiential knowledge of them). Because of this [and also because, unlike looming detectors, no neurophysiological correlates to τ have to date been identified in primates (6)] the direct perception of τ remains the subject of ongoing research and debate. It is plausible to assume, though, that if this is possible it is accomplished with the same sensory mechanisms [3] that detect $\dot{\theta}$ and $\dot{\theta}/\theta$. It is apparent, then, that looming and the mechanisms that detect it play an important, if not exclusive, role in the collision detection/response task.

Automotive applications constitute an increasingly important context for the study of the visual system's capabilities with regard to collision detection/avoidance, especially

---

[2] This was first presented in the science fiction novel "The Black Cloud", by Sir Frederick Hoyle (Harper, 1957).

[3] *Mechanism* refers to the physiological entity responsible for detecting a stimulus, or discriminating a difference between stimuli. It commonly refers to a single neural unit or perhaps to a small pool of neural units with nearly identical tuning for particular stimulus dimensions.

given the trend towards automated driving. Though the current automation technologies are admirably suited to overcome some of the more obvious human driving deficiencies (slow response, distractibility, fatigue, impaired performance), humans bring a number of truly remarkable skills to the driving task that no technology can yet duplicate (superior information acquisition and processing capabilities and the cognitive-based abilities to adapt and anticipate). Attempting to automate even part of the overall driving task with current technology will therefore solve some problems but create others. Ideally, of course, we would like to integrate the human driver with automation technologies in a way that keeps the best of both worlds. The most efficient route to this objective is to develop a comprehensive, quantitative model of human driving behavior. Though human driver models of varying scope and complexity have been developed, none has to date incorporated a scientifically grounded representation of the visual system, which, as with the human driver they seek to represent, constitutes their chief source of input. This is somewhat surprising, given how dependent the performance of these models will necessarily be on that of their visual system component.

In this study we begin to address this deficiency by quantifying the visual system's ability to detect and respond to changes in visual angle θ (and thus defer the aforementioned issues surrounding τ and $\dot{R}$ to a subsequent investigation). To focus this effort we consider a very specific driving scenario, the "Lead Vehicle" (LV) Braking Scenario, in which two automobiles are traveling relatively close together at the same speed (zero range rate) and in the same direction along a straight section of road. At some point the LV unexpectedly engages in maximum braking to come to a complete stop, and in order to avoid a rear end collision (REC), the following vehicle's (FV's)

driver must brake to a stop just before making contact with the by-then fully stopped LV.

The motivation for this approach is several-fold:

1. The driving scenario under consideration has a definite, well defined beginning (onset of LV braking) and end (FV comes to a complete stop);
2. It allows us to isolate $\dot{\theta}$ (as much as possible) as the only cue available to the FV driver;
3. It also allows us to restrict the collision avoidance maneuvers available to the FV driver to only one, braking, whose regulation requires the continued monitoring of $\dot{\theta}$.

Even though this is a rather constrained scenario, the General Estimates System (7) reported that in 2001 about 4% of all REC's involved a decelerating LV and an attentive FV driver. This represents approximately 76,000 collisions and 2300 fatalities. Furthermore, the results of this study are quite general and can be applied to any scenario—driving or not—in which in which the observer views the approach of another object or stationary obstacle.

1.1 Contribution

In this study we conducted an experimental program to develop a model of the visual system as it relates to the collision detection/avoidance task (i.e., detection of and response to changes in visual angle θ.) The model, consisting of a linear filter, a nonlinear criterion detector, and an adjustment mechanism (see Figure 31), exhibits the following characteristics:

1. The linear filter exhibits a band-pass frequency response characteristic (Figure 34) which is well described by a minimum phase, second order transfer function. This in turn gives rise to a damped, second order response characteristic (Figure 39).
2. The criterion detector monitors the output of the filter and signals the onset of looming ($\dot{\theta} > 0$) when some aspect of the filter's output reaches some *criterion value*. Our experiments indicate that it senses either the absolute magnitude of the filter output or its peak-to-peak excursion.
3. Detection of the onset of looming is based on a signal/noise paradigm. This is in contrast to the widely held assumption that a threshold is associated with the detection task. This carries two significant implications for the detection/monitoring task:

- The criterion value is not fixed but instead is a function of the expected costs and benefits associated with the detection task. It can be set by the adjustment mechanism to any value necessary to maximize the task's expected net benefit.
- The filter's output is not altered in any way by the criterion detector. Information loss thus does not occur and the filter's output is available in its entirety for other visual tasks, whether the criterion detector senses the onset of looming or not.

PART I: BACKGROUND

The concept and design of the experiments performed in this study were guided by prior research in both the engineering and vision sciences. The relevant findings from each field are reviewed in the next three chapters so that the results of the experiments performed in this study can be more fully understood and appreciated. We begin in Chapter 2 with a review of three important experimentally based engineering models of manual control. The earliest, McRuer's crossover model (1965), is notable in that it separates the function of the human operator into separate sensory (visual), control, and actuation components, with separate experimentally derived representations for each. This basic structure can be discerned in many subsequent manual control models, even up to the present day, including Kleinman's LQG model (1970) and Fancher and Bareket's headway control model (1994), which are also described. McRuer's and Kleinman's models were originally developed to assist in the design of pilot interfaces and controls for high performance aircraft, but in later years were adapted to address the automotive lateral control (lane keeping) problem. Fancher and Bareket's model, which is much more recent, differs in that it addresses the automotive longitudinal control (rear end collision avoidance) problem.

Chapter 3 provides a brief overview of the human visual system's "motion pathway" and describes the neural mechanisms along it that underlie our ability to perceive motion in the environment. Our performance in this regard is predicated upon the capabilities and limitations of these mechanisms, and to better understand them we review prior studies that have quantified their response characteristics. (Linear systems theory and signal detection theory have proven to be especially effective and powerful tools with which to conduct this research and understand its results.) This chapter concludes with a

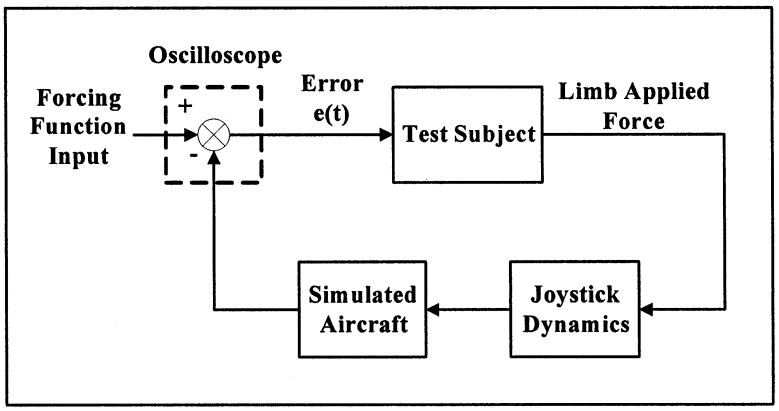
description of a neural looming detection mechanism, which forms the basis for the perceptual model developed for this study.

Chapter 4 begins by summarizing and comparing the results of two related sets of studies. First, the response characteristics of an retinal ganglion cell and human observers for the detection of temporally varying light stimuli are compared. Then the response characteristics of human observers for the detection of temporally varying light stimuli and motion stimuli are compared. Given that the motion pathway begins with the retinal ganglion cell receptive field, the performance of all downstream mechanisms will be dictated by that of the retinal ganglion cell. These three sets of response characteristics, then, should be similar. We next review the flicker detection models of DeLange (36) and Kelly (37), and Stork and Falk's (38) follow-on analysis of Kelly's data for insight into the structure of the looming detection mechanism that will be investigated in this study.

2 ENGINEERING MODELS OF MANUAL CONTROL

2.1 McRuer's "Crossover" Model

One of the first comprehensive human operator (HO) models was McRuer, et al's 1965 "crossover" model (8). This was the product of an experimental program in which test subjects (all experienced pilots) controlled a simulated aircraft using a simple joystick. Their task was to track an input forcing function using the position error between this function and the aircraft's response to their control action as their only input signal. This error signal appeared as a spot on an oscilloscope display constrained to move only along its horizontal axis. This is described schematically by the block diagram in Figure 2. The spot's distance from the center of the screen represented the

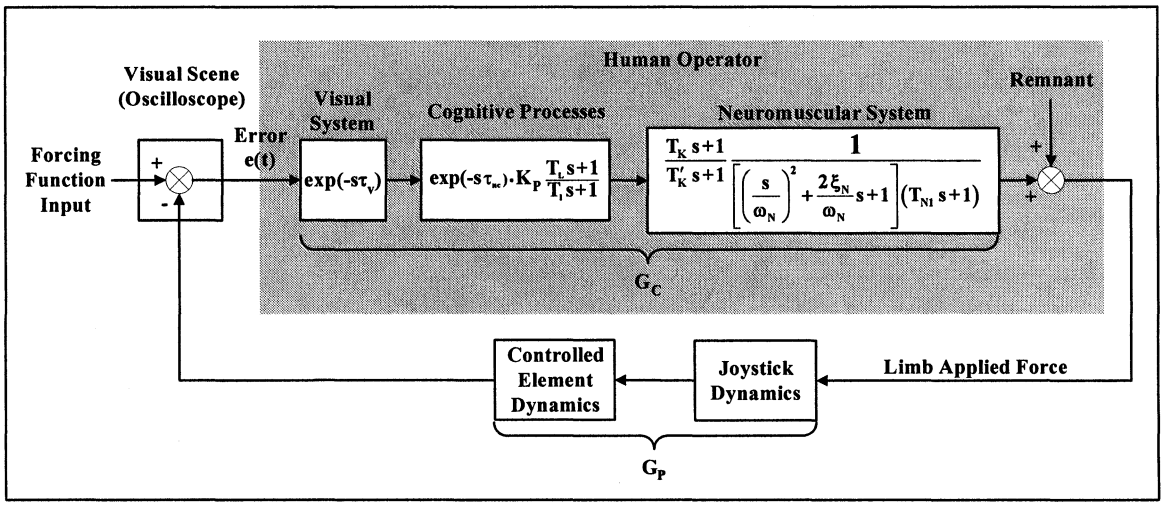


Test Procedure Schematic
Figure 2

direction and magnitude of the error. The joystick was a spring-resistant device, constrained to move only left or right, with very little mass or damping. Hence its contribution to the overall dynamics of the system could be neglected. The Forcing Function input consisted of 10 independent sine waves with a Gaussian amplitude distribution, and with frequencies chosen to make the resultant signal appear random, so that the test subject could discern no repeatable pattern that would allow him to anticipate the future trajectory of the spot. Two first order and two second order aircraft models

were employed. One of each was marginally stable, and the other was unstable. For each combination of Forcing Function Input and aircraft model the HO's frequency response was obtained.

The final crossover model is shown in Figure 3. (Here s is the Laplace variable.) This is a quasi-linear model consisting of a linear describing function (the Visual System,



McRuer's Crossover Model
Figure 3

Cognitive Processes, and Neuromuscular System blocks in the figure) and a Remnant, which takes into account time-varying and non-linear behavior (including noise). The parameters of the describing function are adjusted according to a set of verbal rules to match the task at hand, which is defined by specifying a particular Forcing Function Input and Controlled Element Dynamics. McRuer and others adapted the crossover model for use in automobile steering applications (9). No further elaboration of the human operator components (and in particular the visual component) shown in Figure 3 was undertaken, however.

Several of this study's results and assumptions are pertinent to the current study:

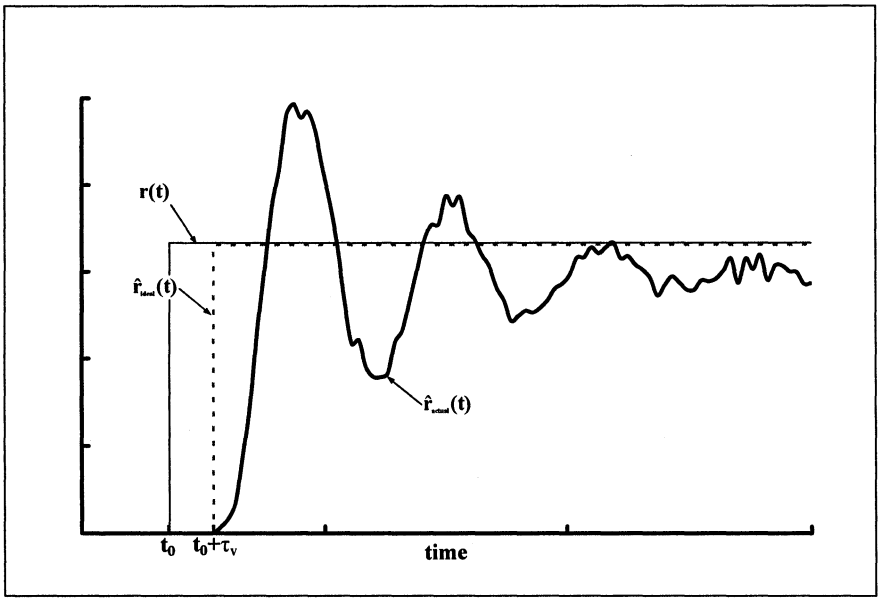
1. The HO exhibits an adaptive capability which allows him to compensate the system for the Controlled Element Dynamics and his own reaction time delay. This is accomplished in the model by the adjustment of the Cognitive Processes's parameters K_P, T_L, and T_I. The effect of this is to make the open loop transfer function $G_C G_P$ in the vicinity of the crossover frequency ($\omega_c \approx 7$ rad/sec) approximately that of a simple integrator plus a time delay, i.e.,

$$G_C G_P(s) = \frac{\omega_c e^{-s\tau_e}}{s},$$

where ω_c and τ_e are both variable. The effect of this adaptation is to achieve absolute system stability and minimization of steady state error.

2. Interestingly, the study found no evidence for significant nonlinear behavior on the part of the HO. The primary sources of remnant were nonstationary behavior (i.e., time-varying components in the effective time delay and gain), and noise. For extremely unstable controlled elements HO's also exhibited a pulsing control action, which was reflected in the remnant as well.
3. The visual system is represented here as a pure time delay—specifically, the shortest time that the signal from a flash stimulus can travel from the retina to the visual cortex ($.04 \leq \tau_v \leq .07$ sec). In other words, the visual system is represented as a perfect sensor, responding instantly (except for this delay) and perfectly to any visual input.

We know from the vision sciences that the representation described in this point is not the case. The visual system typically exhibits temporal dynamics in its response, information loss, self-generated noise, and steady state error. One way in which these imperfections might manifest themselves is shown in Figure 4. Since the crossover model was developed from actual experimental data taken from human subjects, it contains these additional characteristics. Since they are not represented in the Visual System element, they must be embedded in one or more of the Cognitive Processes block, the Remnant, and the Parameter Adjustment Rules. This brings us to the crux of the matter, and our ultimate goal, of which this study is a first step: to separate the visual system functions from the other elements and contain them in a single, comprehensive sub-model for inclusion in HO models such as this one. This will result in a more



Actual vs. Assumed Visual System Dynamic Response
Figure 4

accurate and robust model which relies less on heuristic rules of thumb for its operation, and which gives greater insight into the processes underlying human performance in the manual control task.

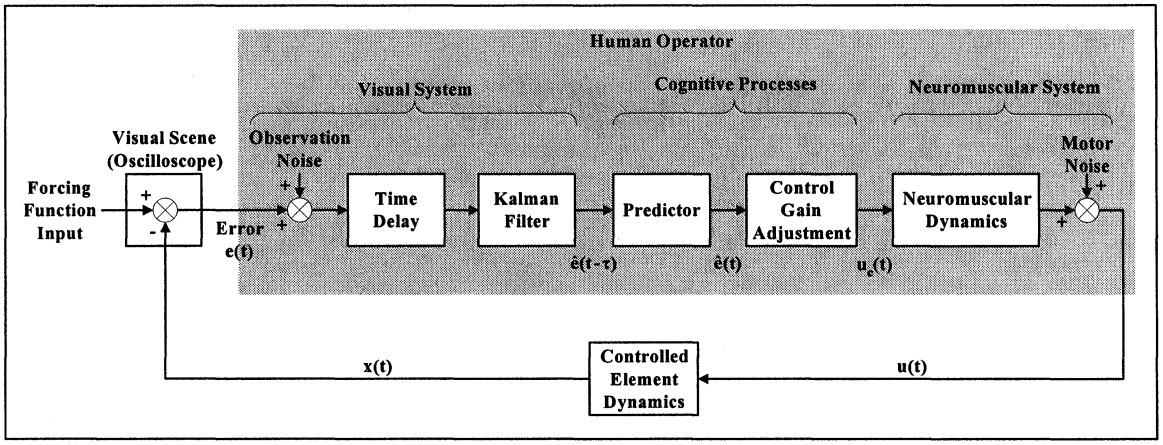
2.2 Kleinman's LQG Model

Kleinman, et al (10) extended McRuer's model; using optimal control theory and state space techniques. The resulting two-input, single-output, Linear Quadratic Gaussian (LQG) model is shown in Figure 5. The most significant differences between the two models are:

1. Kleinman's is a linear model and does not provide for discontinuous or nonlinear behavior.
2. Kleinman represents McRuer's remnant as zero-mean, Gaussian observation and motor noise, whose variances must be determined empirically[4].
3. The visual system is now represented by a time delay, a Kalman filter, and observation noise.

[4] In a related study the authors found that for a wide range of foveal viewing conditions employing this type of visual input, observation noise was well represented by zero-mean Gaussian noise whose covariance was equal to .01 π times the variance of the visual signal being observed.

4. The time delayed observation emanating from the Kalman filter is explicitly extrapolated to the present by a Predictor, which presumably resides in the higher level processing centers of the brain.
5. The HO behaves as an optimal controller, seeking to minimize some endogenous cost function subject to the inherent limitations imposed by the observation noise and time delay.



Kleinman's Linear Quadratic Gaussian Model
Figure 5

As with McRuer's model, the HO here acts as a "null operator" in that he observes the lateral position and velocity of a point on a display (an error signal) and applies control action to place and keep it at the display's origin. Both the Kalman filter and the Predictor operate on the basis of minimizing this deviation. The cost function that the HO seeks to minimize is assumed to take the form

$$J(u)=\lim_{T\to\infty} E\left\{\frac{1}{T}\int_0^T\left[\sum_{i=1}^{n} q_i x_i^2 + r u^2(t) + g\dot{u}^2(t)\right]dt \;\middle|\; y_p(\sigma),\ \sigma\le t\right\}, \qquad (4)$$

where, as indicated in Figure 5, x_i is the controlled element state [which here would be the deviation between the desired ($x_i=0$) and actual states], u the control action, $\dot{u}$ the rate of change of control action, and y_p the Kalman filter input. The weightings q_i, r, and

g are unknown and must be determined empirically. Note that this is tantamount to specifying the HO's control objectives.

Kleinman applied this model to the task of maintaining the longitudinal position of a hovering VTOL aircraft, presenting the HO with two visual displays:

1. Deviation in longitudinal position from desired;
2. Deviation in aircraft pitch from horizontal.

Both signals were presented, as in McRuer's study, as spots on separate oscilloscope screens that could move only laterally about the origin. The HO controlled the aircraft by varying its pitch, based upon the information provided by the two signals. He was thus required to scan back and forth between the two oscilloscopes to obtain this information. Agreement between the model's predictions and the experimental data was very good.

Both the Kalman filter and the Predictor require a model of the system being controlled, implying that the HO is capable of developing an internal model of his environment. The requirement to project the delayed observation forward to the present is also significant. McRuer's model also incorporates a delay in perceiving the visual stimulus, so this projection capability must also be present in it as well, residing most likely in the Equalization System. Kleinman's model is an improvement over McRuer's in that it expands upon and separates out the visual system function. As previously noted, there is scientific evidence that both transmission delays and noise are present in the visual system. (There is also evidence, as we shall discuss, that the visual system is required to filter its input). Given this, the existence of a mechanism to project the delayed visual signal forward appears a logical necessity, though to date no one has actually identified such an mechanism in the visual system. Still missing is a

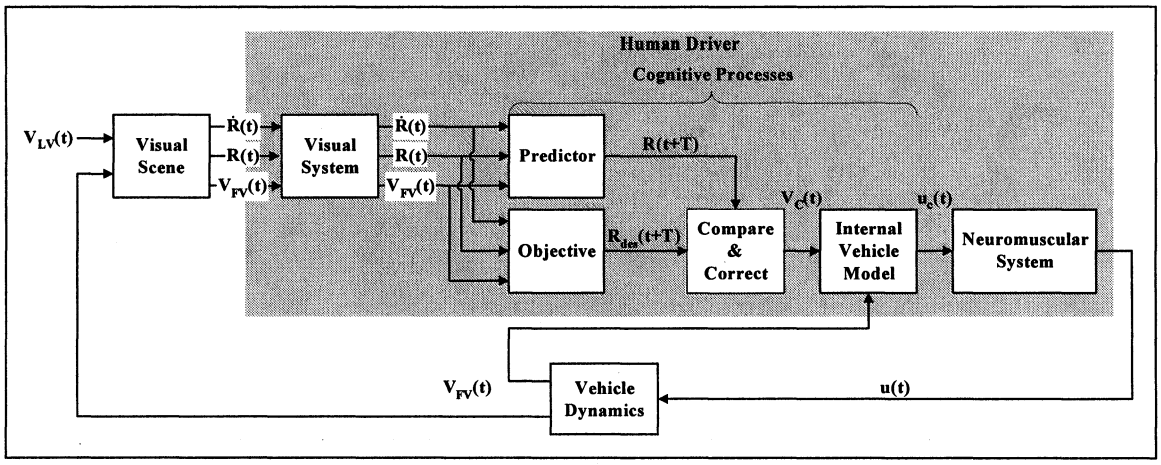
representation of the visual system's temporal dynamics. The form of the filter, predictor, and control law that Kleinman has assumed are of course not neurophysiologically grounded. As with McRuer's model, Kleinman's model represents the result that a control systems engineer would arrive at were he asked to design a controller for these tasks. Even so, these models constitute an important step in the study of human control and provide valuable insights into the HO's function.

Levison, one of Kleinman's co-authors, continued to work on the multiple input problem, and developed the Integrated Driver Model (11) to assist in the design of automotive information systems. The model is the same as Kleinman's, except that it has been adapted to simulate a constant speed lane-keeping task and presents more than two visual inputs, not all related to the driving task. Longitudinal control is not addressed.

2.3 Fancher and Bareket's Headway Control Model

Fancher and Bareket's Driver Headway Control Model (12) is based upon concepts they had previously derived from kinematic and dynamic relationships to aid in the design and evaluation of Automatic Cruise Control (ACC) (13). Their model (which has been recast to facilitate comparison with McRuer's and Kleinman's models) is shown in Figure 6. Here the FV driver acts not upon an error signal, but instead upon LV Range (R), Range Rate ($\dot{R}$), and FV velocity (V_{FV}), which he receives simultaneously from the scene before him. With these, the LV driver predicts what his range will be some period T into the future, and also establishes a desired range R_{des}, according to

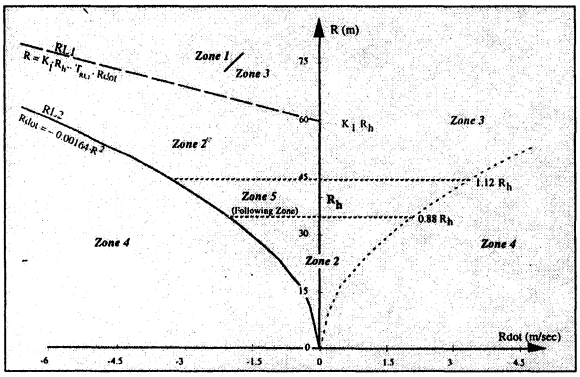
$$\begin{aligned} R(t+T) &= \dot{R}(t)\cdot T + R(t), \\ R_{des}(t+T) &= V_{FV}\cdot T_{des}. \end{aligned} \tag{5}$$



Fancher and Bareket's Driver Headway Control Model

Figure 6

Here T, the "preview time" (14), is not specified and must be obtained empirically. T_{des} is the FV driver's desired headway time, also obtained empirically, but typically in the range of one to two sec. The velocity command V_C is then determined to make $R(t+T)=R_{des}(t+T)$. This is accomplished with reference to the $\dot{R}(t)-R(t+T)$ phase plane, developed by the authors and shown in Figure 7 and Table 1. The driver then sets



$\dot{R}(t)$ - $R(t+T)$ Phase Plane

Figure 7

Zone	Command Equation
1	$V_C = V_{free}$
2	$V_C = V_{LV} - \frac{R_{des}(t+T)}{R(t+T)} \cdot T_2$
3	$V_C = V_{LV} + \frac{R_{des}(t+T)}{R(t+T)} \cdot T_3$
4	$V_C = V_{FV} + [R(t+T) - R_{des}(t+T)] \cdot T_4 + \dot{R}(t)$
5	$dV_C = \text{sgn}[R(t+T) - R_{des}(t+T)] \cdot a_{des}$

Computation of V_C

Table 1

the accelerator or brake pedal to achieve the desired velocity. In the table V_{free} is the speed the FV driver would choose if there were no LV in front of him. The constants T_1 to T_4 are empirically obtained "time constants", a_g is chosen by the FV driver to affect a

comfortable acceleration/deceleration, and "sgn" is the sign function. It is assumed that V_C is achieved via a feedback loop that incorporates some internal model of the vehicle's dynamics. The time constants for this inner loop are assumed to be much shorter than those for the outer loop that establishes V_C, hence V_C is achieved well before any subsequent decisions to alter it are made.

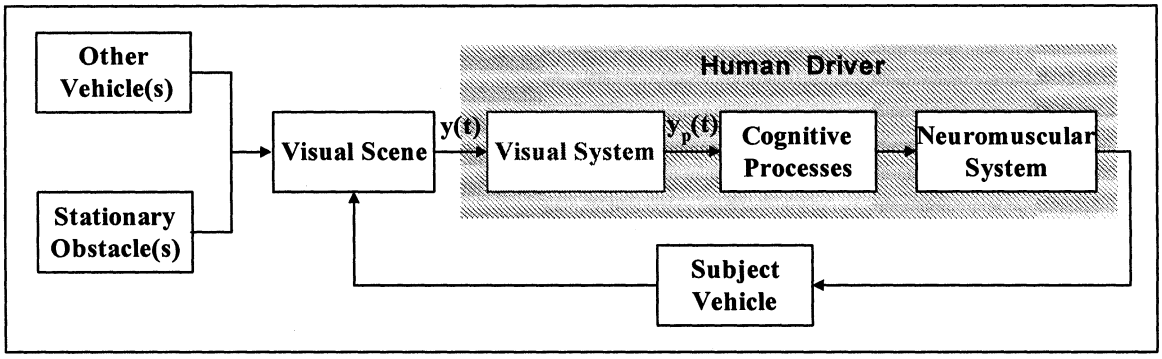
In its present configuration the model's visual component operates without noise, time delay, or temporal dynamics. (We noted that the Predictor function projects the non-delayed quantities R and R_{des} into the future, as opposed to Kleinman's model, which projects delayed quantities to the present. Though the authors note the possibility that the FV driver's function may involve a pure time delay, they do not incorporate one, nor do they specify which aspect of the driver's function it might be associated with.) The Visual System block in Figure 6 thus acts as a perfect sensor. The authors, however, have incorporated two assumptions of non-ideal behavior from Hoffman (15) and Hoffman and Mortimer (16), which form the basis for the zone boundaries in Figure 7:

1. There exists a threshold in the perception of $\dot{\theta}$ such that for $|\dot{\theta}| \leq .003$ rad/sec the FV driver is unable to distinguish it from zero;
2. The LV's range R(t) can only be perceived to within 12% of its true value.

We are skeptical of these, however, and investigate them in Chapter 5. The performance of the model, of course, is critically dependent upon these assumptions and should they prove unwarranted, a much different model of human driving behavior would result. Hoffman and Mortimer's assumptions have been widely disseminated and can be found in a number of other human driver models. [For example, (17), (18), and (19)].

2.4 Summary

In this chapter we reviewed three engineering models of manual control whose development and basic structure guided the design of this study's design. All can be reduced to the basic structure of Figure 8. The loop begins with the subject vehicle,



A Basic Human Driver Model
Figure 8

which interacts with other vehicles and objects in the environment to produce the actual visual scene viewed by the subject vehicle's driver. The driver's cognitive processes decide on a course of action based not upon the actual scene [y(t)], but the perceived scene [$y_p(t)$], and issue commands to the driver's body (the neuromuscular system) to implement it. The subject vehicle responds to the driver's actions and the loop repeats.

We noted that to date the vast body of relevant research from the vision sciences have not been incorporated into these models, whose visual elements are instead based upon idealized behavior and in some cases questionable assumptions. It should also be pointed out that in developing comprehensive models such as these in one step, the experimenter has at hand many free parameters with which to fit the model to the data. This almost ensures reasonably good agreement between the two. Predictions based on such models, then, must be made with great care. The approach taken in this study is to isolate the visual function and separately develop a model of its ability to detect and track looming

based upon the visual cues associated with such motion. Once complete, this representation can be incorporated back into the larger human driver model. For guidance in this effort we now turn to a review of the relevant research from the vision sciences.

3 ANATOMY OF THE VISUAL SYSTEM'S MOTION PATHWAY[5]

The *magnocellular* or motion pathway is one of perhaps 20 pathways or *streams* in the visual system (22). They are distinguished by their functionality: each is specialized to analyze a different aspect of vision. They are not isolated from one another, however, and a good deal of interaction can occur between them throughout their length. The motion pathway is one of the most extensively studied visual pathways.

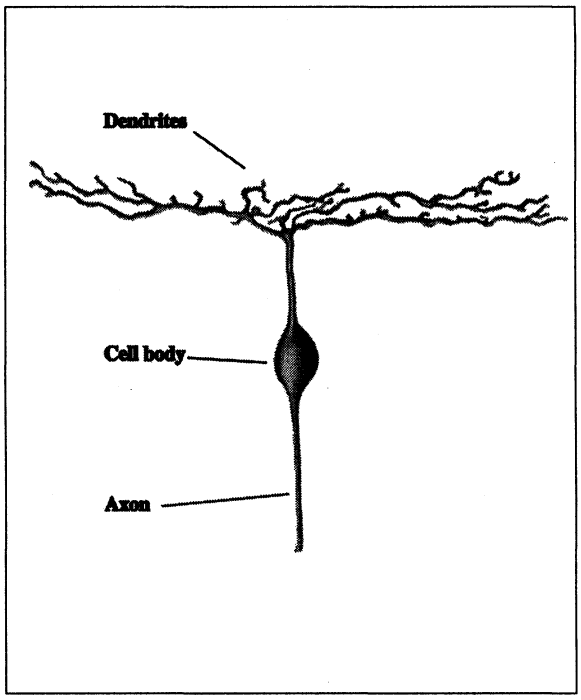
Vision Research has revealed a number of important "principles", which give a broader perspective of the brain's activity regarding vision generally and motion processing in particular. Key among these are:

1. The existence (referred to above) of a "processing stream" made up of interconnected regions within the brain that are specialized for the detection and interpretation of motion signals.
2. A cascading structure of increasing functionality, in which upstream neural mechanisms capable of detecting and analyzing simpler aspects of motion converge to form more complex downstream mechanisms that can perform more sophisticated detection and interpretation tasks.
3. An adaptive capability that optimizes the system's limited sensory capabilities for the viewing conditions at hand.

3.1 The Neuron

The *neuron* is the fundamental element of the visual system, as it is of the rest of the nervous system (which includes the brain). It is a type of cell found only in the nervous system, and is unique from all other cells in that it can communicate with other neurons. A prototypical neuron is shown in Figure 9. The *axon*, *dendrites*, and *synapses* (not shown) are the unique features that enable it to communicate with other neurons. A neuron receives its input from the axon of a sending neuron, which makes contact either directly with the receiving neuron's body or, more typically, with one of its dendritic

[5] Except where otherwise noted, Chapters 3 and 4 and the figures contained therein are based on Smith & Snowden (1), Purves, et. al. (20), Wandell (21), and De Valois & De Valois (24).



A Prototypical Neuron
Figure 9

branches. A neuron issues its output via its axon, an extension of the cell body specialized for signal conduction. Neurons can have any number of dendritic branches, but no more than one axon, which can branch at its terminus to form as many as 1,000 separate contact points (not shown in Figure 9). (A few types of neurons have no axons at all.) Axons are typically only a few millimeters long, though a few are much longer. Retinal ganglion cell axons, which extend from the retina to the mid brain-region, are about five centimeters long. Signals travel along the axon by means of self-generating electrical waves called *action potentials*.

Neurons transmit information by means of generated electrical potentials, of which there are two basic types: *graded potentials* and *action potentials*. Graded potentials are generated in sensory receptors (such as the photoreceptors of the eye) and in dendrites. They are usually sustained in nature, last as long as the stimulus, and have an amplitude

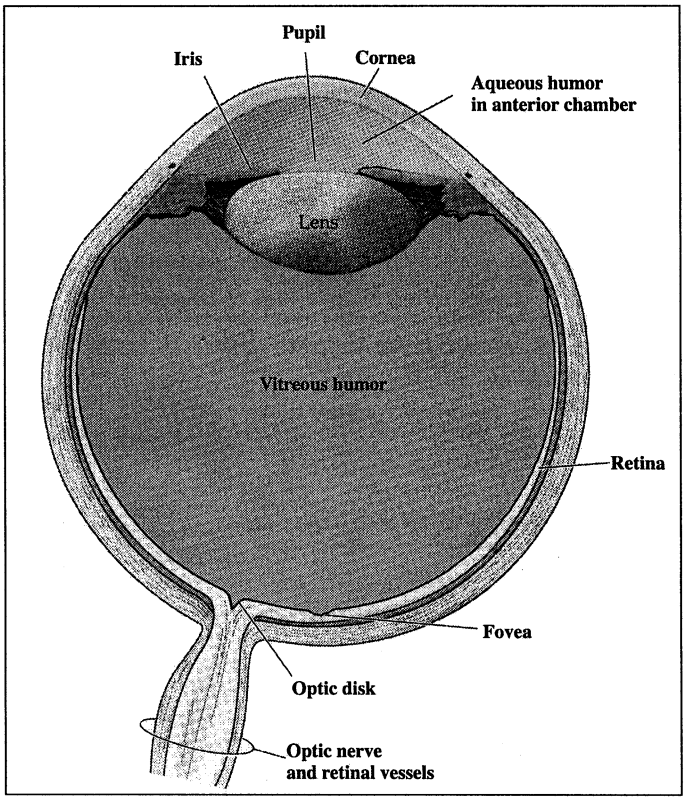
that is proportional to the strength of the stimulus. Graded potentials are not *regenerated*, and so their amplitudes diminish with the distance from their site of generation. They add together, and may be either positive of negative. Graded potentials have no *threshold*; for example, a single photon of light absorbed by a photoreceptor generates a small, graded potential. Action potentials, often called *impulses* or *spikes*, are relatively large (.1 volts) transient (1 to 2 milliseconds in duration) potentials that are transmitted along axons. Despite its intrinsically poor electrical characteristics, the axon constitutes a kind of "booster system" that continually regenerates these signals along its length, providing for their transmission over long distances without any reduction in amplitude. Axon potentials signal the strength of a stimulus by their frequency rather than their amplitude. They also have a threshold, typically requiring a change in cell membrane potential of about 15 mV before they can be generated. Some neurons with short axons or no axons transmit signals only by way of the graded potentials developed at their dendritic synapses. Most neurons, however, have both kinds of potentials. The graded potentials in the dendrites sum to change the membrane potential of the neuron by a sufficient amount to trigger the firing of action potentials in the axons. Graded potentials thus generate action potentials in these neurons.

Underlying both graded and action potentials is a *resting potential*, which is the membrane potential of the unexcited neuron. In its unstimulated state, the electrical potential across a neuron's plasma membrane is typically –40 to –90 mV: the cell's inside charge is negative relative to the outside charge. Action potentials are elicited when an ionic current passes across the membrane of the neuron. Under normal circumstances, such a current is generated by another neuron (at the synapse between the

two nerve cells), or by the transduction of an external stimulus in sensory neurons. If the current is such as to make the membrane potential more positive than the resting potential (*depolarization*), then at a certain point, called the *threshold potential*, an action potential occurs.

3.2 The Eye

Figure 10 gives a schematic representation of the imaging components of the eye. The formation of focused images on the photoreceptors of the retina is accomplished by the refraction (bending) of light by the *cornea* and the *lens*. The cornea is responsible for most of this refraction. The lens has considerably less refractive power than the cornea;

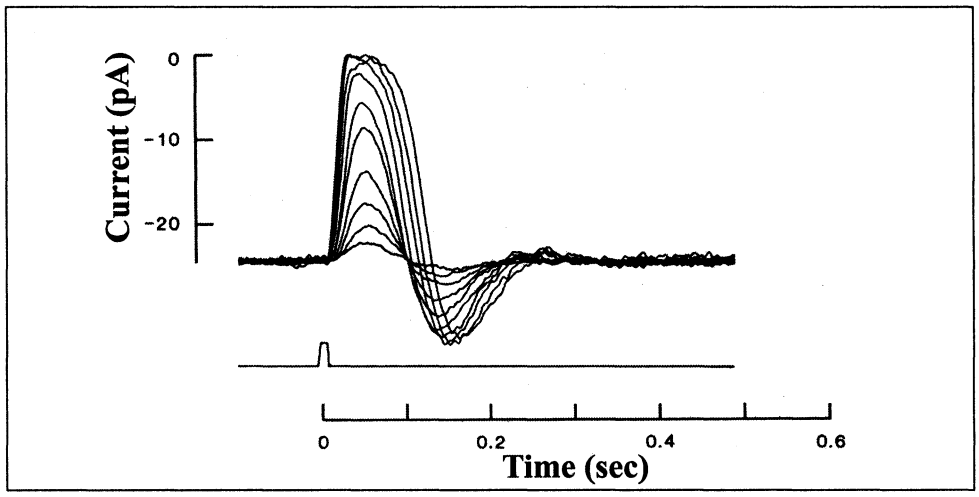


Imaging Components of the Eye
Figure 10

but it can be adjusted through the action of the *ciliary muscle* (not shown) that connects around its circumference. Objects that lie at various distances from the observer can thus

be brought into sharp focus on the retinal surface. Adjustments in the size of the *pupil* (the circular opening in the iris) also contribute to the clarity of images formed on the retina.

The *retina* lines approximately 200° along the back of the eye and despite its peripheral location, is actually part of the brain. It contains several different types of neurons that convert light first into graded electrical potentials, and then into action potentials. Two of these, the *rod* and *cone photoreceptors*, are the only elements of the retina that are sensitive to light. They contain different types of *photopigment*, a light-absorbing substance. Absorption of light by these photopigments initiates a cascade of events, known as *phototransduction*, that results in the generation of a graded potential, as shown in Figure 11. Information about the retinal stimulation is ultimately communicated (through several classes of intermediary neurons) to the *retinal ganglion cells* (RGC's), whose axons form the *optic nerve*. The RGC's communicate this



Response of a Cone Photoreceptor to Brief Flashes of Light (23)
Figure 11

information (via action potentials) to the rest of the central nervous system, and are the only cells to do so.

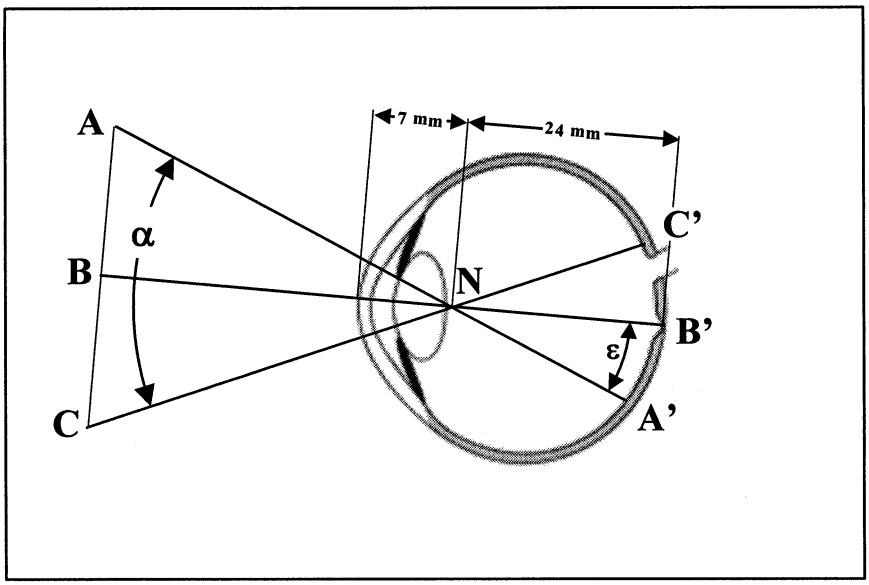
The rod and cone *systems* (by which we mean the photoreceptors and their connections within the retina) are specialized for different aspects of vision. Rod photoreceptors are able to capture more light, and the transduction mechanism in rods provides greater amplification than that in cones. (A rod can respond to a single photon.) The rod system is therefore extremely sensitive to light. Rod systems exhibit very low spatial resolution, however. In contrast, the cone system exhibits high spatial resolution but low sensitivity to light. (More than 100 photons are required to activate a cone.)

Both rods and cones transmit information about the wavelength of light as a function of the types of the photopigments they contain. All rods contain the same photopigment—rhodopsin—whereas individual cones contain one of three different photopigments, collectively called *cone opsins*, that have different but overlapping absorption spectra. The relative activity of these three sets of cones [referred to as S, M, and L cones for the short (blue), middle (green), and long (red) wavelengths] generates retinal signals that ultimately give rise to the sensation of color. Besides it aesthetic appeal, color vision makes it possible to distinguish objects that might be difficult to contrast with their surroundings.

When light arrives at the photoreceptor its intensity varies sinusoidally but superimposed on this are high frequency variations due to quantum fluctuations. In addition, there are also high frequency fluctuations due to the probabilistic nature of photon capture (less than 10% of the available light is actually captured by the

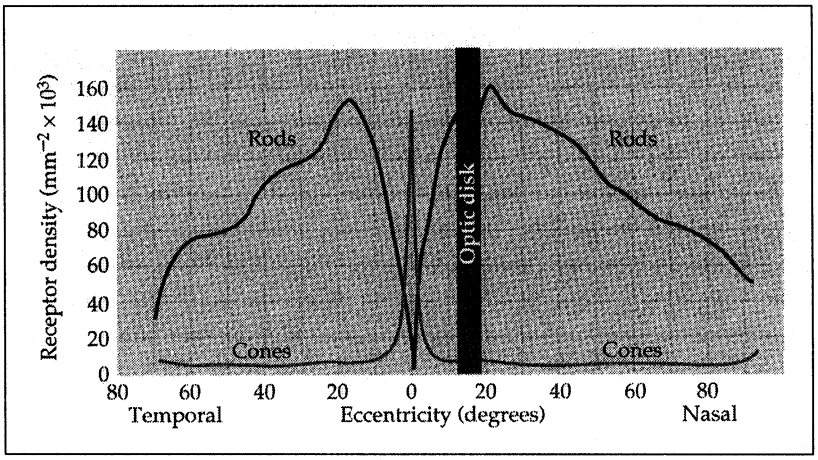
photoreceptors) and the whole phototransduction process. The electro-chemical processes involved in phototransduction have the effect of averaging information over time (over a period of some milliseconds), and filtering out this quantum and physiological noise. These filtering properties change as a function of background light level. At low light levels, there is not enough information available (few photons per unit time) to resolve high temporal frequencies and the high frequency cutoff is therefore relatively low. The high frequency cutoff shifts to higher temporal frequencies as the background light level increases, to take advantage of the greater information content (more photons per unit time) and the concomitant increase in signal to (quantum) noise ratio (24).

The region of highest visual acuity in the retina is the fovea (See Figure 10). When we turn our head and eyes to focus on a point in the environment, we are placing its image on the fovea. Figure 12 shows how the image of object ABC is formed on the retina. (To a good degree of approximation we can ignore the lens and consider only the refractive effects of the cornea.) The central rays from points A, B, and C strike the cornea normal to its surface and pass through it without being refracted, falling on points A', B', and C' of the retina. Here the observer is assumed to be fixing his gaze on point B, so B' falls on the fovea. These central rays all intersect at the nodal point N of the eye. (The dimensions shown in the figure are typical for a normal, adult eye.) Ray B-N-B' constitutes the *visual axis* of the eye. The angle subtended by the object, α, is referred to as the *visual angle*. Specific locations on the retina are typically described in terms of their angular distance from the fovea. This is referred to as *eccentricity* (ε), as shown in Figure 12.



Visual Angle and Eccentricity
Figure 12

There are approximately 120 million rods and 6 million cones in the retina of each eye. Their distribution across the surface of the retina varies markedly, as Figure 13 shows. The central human fovea contains approximately 50,000 cones and no rods,



Distribution of Rods and Cones in the Human Retina
Figure 13

which explains why it is the region of highest visual acuity. Although cones are not restricted to the fovea, their lower density outside the fovea, as well as the lower density of the ganglion cells that they can connect to, explains why visual acuity declines so

markedly as we move away from it. Acuity is reduced by 75% just 6° eccentric to the line of sight.

3.3 Retinal Ganglion Cell Receptive Fields

Signals from the approximately 126 million photoreceptors in each eye converge upon approximately one million retinal ganglion cells (RGC's), which are communicated to the rest of the brain. This compression is accomplished by way of the *receptive field* structure. The receptive field of an RGC consists of an approximately circular array of photoreceptors divided into two groups: a circular *center* and an annular *surround*. Not all of the photoreceptors contained within this circular region connect to the RGC in question, and the photoreceptors that do can also be connected to other RGC's. Thus an individual photoreceptor can be a member of a number of RGC receptive fields. Receptive fields can vary in both their size and in the number of photoreceptors making them up. In the fovea, an RGC contacts only a single photoreceptor (though again, that photoreceptor may contact a number of RGC's). Receptive fields of varying sizes are encountered outside the fovea. Thus the entire retina of each eye is covered by many overlapping receptive fields of different sizes. In general, however, receptive field size increases and the number of receptive fields decrease with increasing eccentricity. (Foveal RGC's don't exhibit this center-surround receptive field structure, since they are connected to only a single photoreceptor in this region.)

RGC's respond to illumination of their receptive fields with a series of action potentials, which were previously defined as a series of voltage spikes with a constant magnitude of approximately .1 volts. The strength of an RGC's response is reflected by the frequency of this impulse series, in units of impulses, or *spikes*, per second. Action potential frequency, or *spike rate*, changes over the time course of the RGC's response.

RGC action potentials have been recorded in cats and monkeys using a microelectrode which is placed either near the RGC's body or near its axon in the optic nerve. As with all neurons, noise and variability are inherent in an RGC's response . Therefore a number of such responses to the same stimulus are usually recorded and then averaged together. Spike rate as a function of time is then estimated by first dividing the time course of this averaged response into increments and then dividing the number of spikes occurring in each increment by its duration. A plot of this averaged response versus time is called a *peri-stimulus time histogram (PSTH).*

RGC's respond poorly to uniform illumination of their receptive field, though response (as measured by spike rate) does increase with increasing luminance level. The response at any given level of uniform illumination is called the *spontaneous firing rate.* The situation is dramatically different if, in addition to this uniform field, a small spot of light is directed at various points within the cell's receptive field. Kuffler (25), who was the first to investigate RGC receptive fields in this way, found that when a spot of light was placed in the center of its receptive field, some RGC's responded with a distinct increase in spike rate over the spontaneous firing rate. Moving the spot to the surround resulted in a response that was below the spontaneous rate. Such receptive fields were labeled *on-center, off-surround,* and about half the RGC's have receptive fields are of this type. The other half have *off-center, on-surround* receptive fields, and behave in an opposite manner. These results show that an RGC is sensitive to the *difference* between the level of illumination falling on its receptive field center and the level falling on its surround. Thus, the signal supplied by the retina to central visual structures does not give equal weight to all regions of the visual scene; rather it emphasizes the regions that

contain the most information—namely, the regions where there are differences in luminance.

Stimulus contrast is defined as the ratio of the luminance at a particular point to that of the region surrounding it. To place this on a firmer mathematical footing, let $\ell(x,y)$ be the luminance of a pencil of light falling on a point (x,y) anywhere within the area of an arbitrarily shaped stimulus. The contrast of that point with respect to the entire stimulus region is defined as

$$c(x,y) = \frac{\ell(x,y) - \ell_{avg}}{\ell_{avg}} = \frac{\Delta\ell}{\ell_{avg}}, \tag{6}$$

where

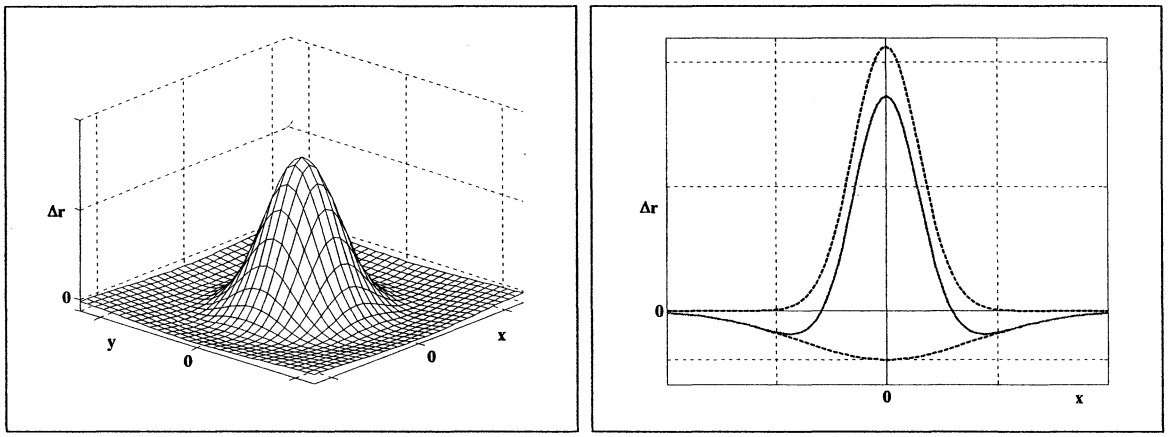
$$\ell_{avg} = \frac{1}{A}\iint_{A} \ell(x,y)\,dx\,dy. \tag{7}$$

Here ℓ_{avg} is the integrated average luminance falling on the stimulus and A the area of the stimulus. Note that $\Delta\ell$ can be either positive or negative.

Suppose that an RGC receptive field is uniformly illuminated with light having luminance ℓ_{avg}, evoking a spontaneous firing rate r_{sfr}. At some point (x,y) we then superimpose a pencil of light having a contrast ratio c(x,y). This will evoke some response r(x,y) from the RGC, so that $\Delta r(x,y) = r(x,y) - r_{sfr}$. For a given *background*, ℓ_{avg}, and for stimulus contrast ratios in the range of approximately $\pm 30\%$, the derivative $\frac{\partial(\Delta r)}{\partial c}$ is approximately constant with respect to c (though it will vary across x and y). A linear relationship thus exists between $\Delta r(x,y)$ and $c(x,y)$, which can be written as

$$\Delta r(x,y) = \frac{\partial(\Delta r)}{\partial c}(x,y) \cdot c(x,y). \tag{8}$$

Kuffler's work has been extended by directing individual pencils of light of equal intensity to various locations within an RGC's receptive field that is otherwise subjected to uniform illumination, and then measuring $\Delta r(x,y)$. A surface plot of these individual responses vs. the location of the stimulus that evoked them (for an on-center RGC) has an appearance very similar to that shown in Figure 14a. It is referred to as the RGC's *two-*



a) Two-Dimensional Steady State Receptive Field

Difference of Gaussians Approximation

Figure 14

dimensional, steady state receptive field. Superposition also holds, and once $\frac{\partial(\Delta r)}{\partial c}(x,y)$ is determined for every position (x,y) within an RGC's receptive field (which can be done experimentally), the response Δr to a general stimulus whose contrast varies across the receptive field according to $c_1(x,y)$ can be estimated from

$$\Delta r = \sum_{x,y} \frac{\partial(\Delta r)}{\partial c} \cdot c_1(x,y). \tag{9}$$

In the mid-1960's Rodieck (26), Enroth-Cugell and Robson (27), and Enroth-Cugell, et al (28) developed a receptive field model that continues to be the basis for most analytical work in this area. Their "Difference of Gaussians" model relates the receptive field's steady state response to point stimulation anywhere within it by the expression (26)

$$\Delta r_{xy} = k_c\, e^{-\frac{x^2+y^2}{\sigma_c^2}} - k_s\, e^{-\frac{x^2+y^2}{\sigma_s^2}}, \qquad (10)$$

where k_c and k_s are the sensitivities of the center and surround region, respectively, and σ_c and σ_s measures of their respective sizes. (Note that in this model the surround region includes the center region.) This is indicated in Figure 14b. Given its linear behavior, the response to more complicated stimuli can be obtained through integration of Equ. (10).

These receptive fields constitute, to a good degree of approximation, a quasilinear system. At any given level of average luminance, its response varies linearly with stimulus contrast. Moving to a different average luminance results in a different linear relationship between response and stimulus contrast. This is true of the visual system in general, and this responsive adjustment is referred to as *adaptation*. This "scheme" fits the visual system to its environment very well. Over the course of the day the background intensity can vary by as much as 14 log units (from starlight to midday sun). Under any particular set of viewing conditions, however, the variation in light intensity is much less—of the order of two log units. RGC neurons give their full range of output over a two log unit range of light intensity. Furthermore, at any particular level of background intensity, its output to small perturbations within this two log unit range is approximately linear. The previously mentioned constraint on using the linearity

assumption should not be viewed as overly severe, for two reasons. First, within any two log unit range of illumination, contrast variations of ±30% about the mean level of illumination constitute approximately 30% of this range. Secondly, the detection tasks considered here involve far less than a 30% variation.

The particular level of average luminance that the visual system is operating at is often referred to as the *adaptation* or *background level*. At the level of the RGC receptive field, adaptation manifests itself in three important ways:

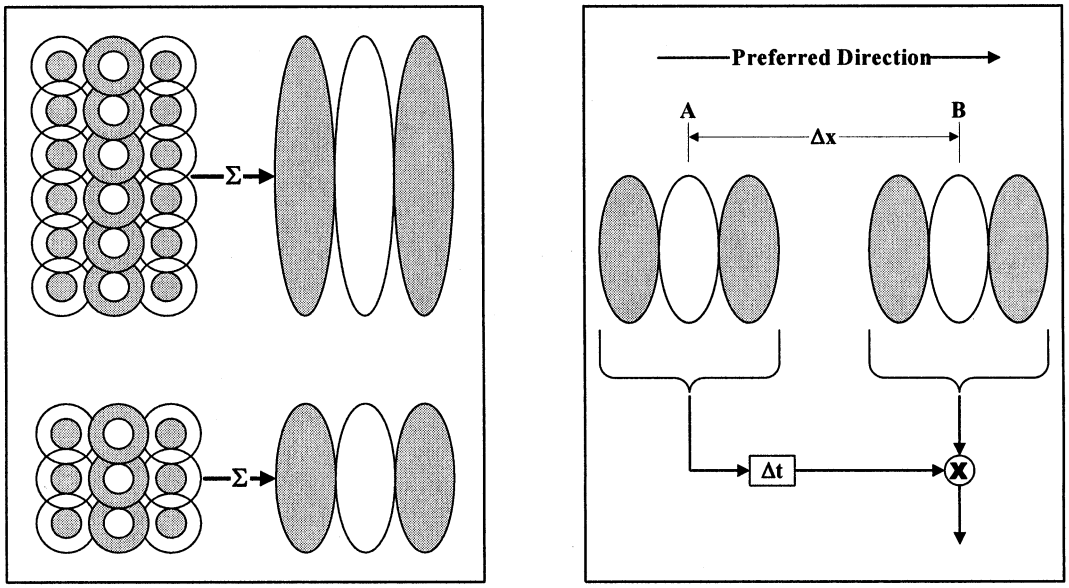
1. Both rods and cones exhibit a decrease in sensitivity with increasing light intensity.
2. Cones (but not rods) exhibit a shift in their operating range such that their whole response range is utilized for the range of intensities about each adaptation level.
3. The temporal frequency response characteristics of the individual photoreceptors, and hence the receptive fields that they make up, change with background intensity, as discussed in Section 3.2.

This is accomplished by a variable gain control mechanism, a large part of which exists at the level of the photoreceptors themselves (24). A complete specification of RGC performance, then, must stipulate these relationships as well.

3.4 Downstream Receptive Field Structures

We have seen that RGC receptive fields detect regions of light-dark contrast. "Downstream" neurons in higher level visual processing centers of the brain combine individual RGC receptive fields in various ways to perform more sophisticated detection tasks. Two such neurons pertinent to this study are *orientation selective* (OS) and *direction selective* (DS) neurons. Orientation selective neurons have adjacent excitatory and inhibitory regions (as opposed to the concentric excitatory/inhibitory regions of RGC neurons), which are longer in one dimension than the other. Lines or edges in the visual field whose orientation, or angle of rotation, coincide with its *main axis* invoke the strongest response from such a neuron. Two possible schemes are shown in Figure 15a.

OS neurons retain the basic linear response characteristics of RGC neurons whose outputs they inherit. They are an example of *simple cells*.



a) Orientation Selective Receptive Field

b) Basic Reichardt Detector with OS Receptive Fields

Figure 15

DS neurons respond strongly to motion in one direction but not the opposite direction. DS neurons are thus *motion detectors*. Their operation, first postulated by Werner Reichardt and colleagues (29), is shown schematically in Figure 15b. An object's image moving across the retina from left to right evokes a response from the receptive field located at position A on the retina. This response is transmitted through a time delay Δt to a multiplication junction. (Except for the time delay, all transmissions are for simplicity assumed to be instantaneous.) Later the image passes across the receptive field at location B, eliciting a response from it. If the speed of the retinal image is

$$v = \frac{\Delta x}{\Delta t},$$

where Δx is the distance between the two receptive field locations, both responses will arrive at the multiplication junction at the same time, yielding a large positive value. If the speed is significantly different from v, or the direction is significantly different from the preferred direction the multiplication will yield a value close to zero. Image speed, and thus the object's speed, is inferred by the distance Δx between the two receptive field locations. Direction is inferred by the orientation of the two receptive fields with respect to one another in retinal x-y space.

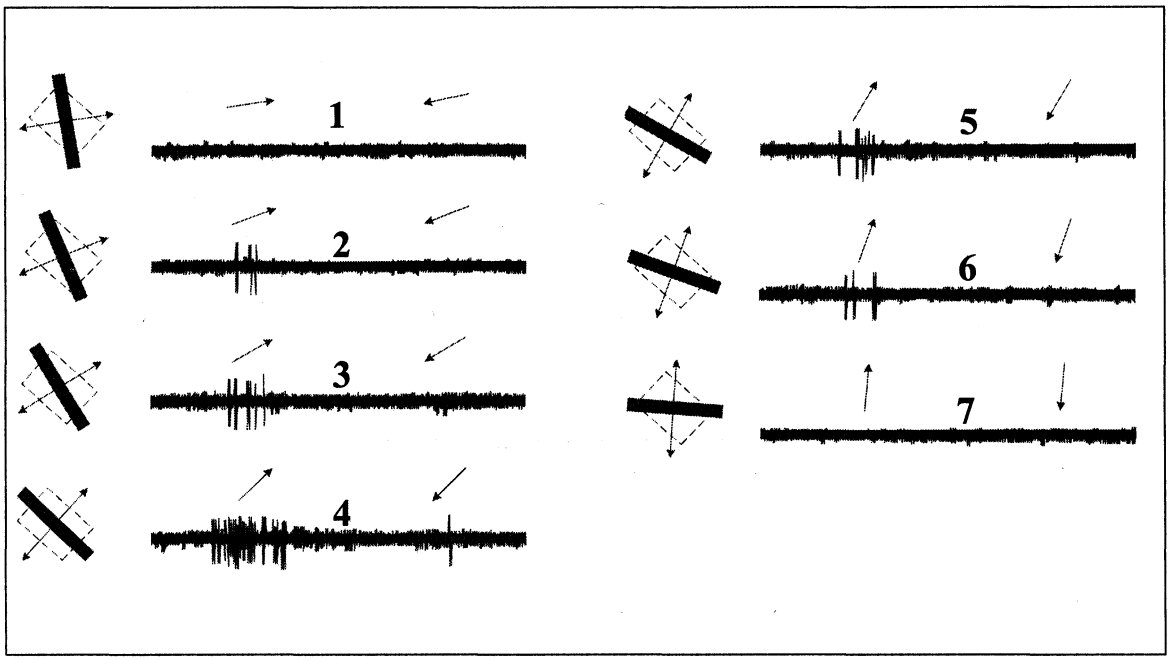
Reichardt, et. al. determined that for such a mechanism to work, it must incorporate a spatially asymmetric component and a nonlinear component [Chapter 2 of (6)]. Spatial asymmetry simply says that if, as in this example, a neuron is to be sensitive to left-right motion, then something must be different between left and right. The simplest scheme to accomplish this is shown in the figure: inputs from only two locations are polled, and the one coming from the left is time-delayed (the spatial asymmetry). Time delays are the most common form of spatial asymmetry employed in motion detection models, but other forms have also been suggested. For example, the left input could act to enable (or *gate*) a response from the neuron to the right input, but be unable to evoke such a response on its own. Alternatively, one input could be inhibitory and the other excitatory, and their interaction nonlinear.

Poggio and Reichardt (30) later proved mathematically that a purely linear motion detector would be incapable of determining the direction of motion, even with a spatially asymmetric element. Large portions of the detector could be linear, but at some stage nonlinearity is required. The simplest type of nonlinearity that can be employed is a quadratic one. Multipliers or sum-and-square operations are in turn the simplest forms of

quadratic nonlinearities, and the ones most frequently employed in motion detection models. Other types of nonlinearities are also possible, of course.

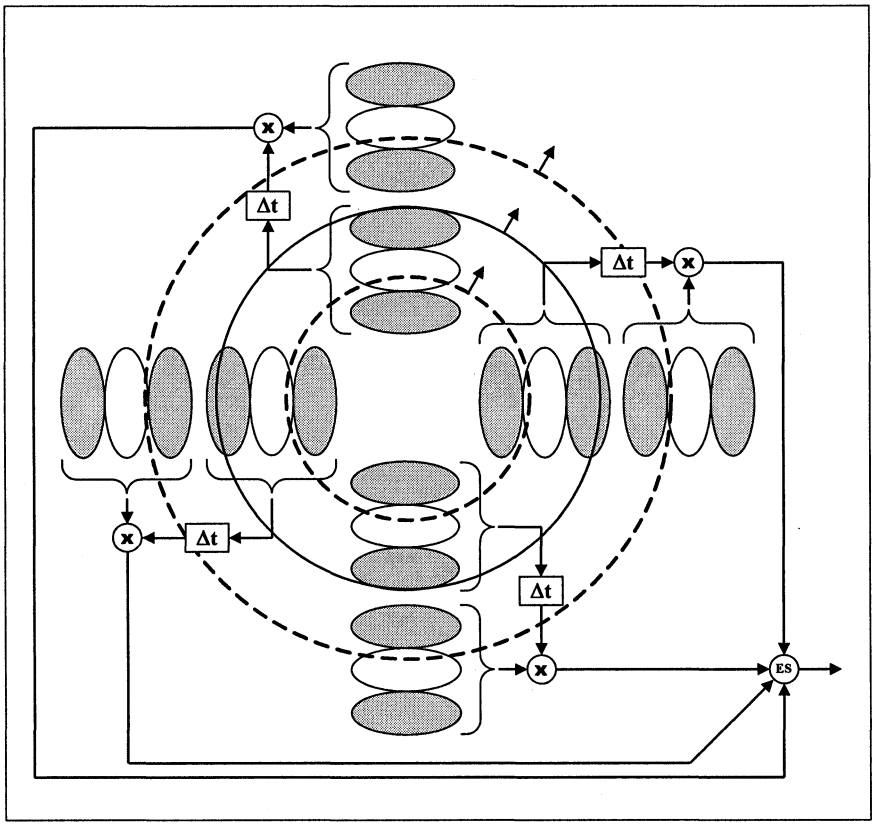
DS neurons appear to be constructed from other simple neurons having either circularly symmetric receptive fields (like those of RGC's) or OS receptive fields. If the input receptive fields are circularly symmetric, the DS neuron would be sensitive to the movement of blobs of light of the appropriate dimension, and lines or edges at any orientation. If they consist of OS receptive fields, the DS neuron would be sensitive to the movement of blobs of light or lines and edges at the preferred orientation. The latter scheme is shown in Figure 15b. The response characteristics of a monkey DS neuron, recorded by Hubel and Wiesel (31), are shown in Figure 16. Here the strongest response is evoked when the line is oriented with the neuron's receptive field (indicated by the dashed rectangle) and moving up and to the right (Case 4). The greater the line/edge's departure from the preferred orientation, direction, or speed, the lower the DS neuron's response. Because of their nonlinear response characteristics, DS neurons have been classified as *complex cells* to distinguish them from the linear simple cells. Note, though, that their basic inputs (simple cells) are linear. Because of its temporal attribute, the overall receptive field of a DS neuron is known as a *space-time receptive field.*

Beyond the DS neurons are neurons sensitive to image contraction/expansion, from which longitudinal motion can be inferred. Though still the subject of ongoing research, it is plausible to assume that these contraction/expansion receptive fields are constructed from DS inputs. Perhaps the simplest way in which such inputs could be combined to construct the receptive field of an expansion sensitive neuron, or "looming detector", is shown in Figure 17. It arranges DS receptive fields such that one with a sensitivity to



Response Characteristics of a Monkey DS Neuron with OS Receptive Fields
Figure 16

upward motion forms the top of the looming detector's receptive field, one sensitive to rightward motion forms the right side, and so on. When an image of the appropriate size expands within this composite receptive field, and its edges expand over the component receptive fields at the appropriate (lateral) speed, each component will send a maximal signal to the expansion selective neuron (ES in the figure), which combines these in a nonlinear fashion and in turn issues a maximal signal. Though this structure is plausible in view of known neurophysiology, this is still an area of active research. Sekuler (32),for example, performed psychophysical studies of speed discrimination in looming displays, and concluded that a simple linear summation of the individual lateral motion inputs of as few as four well distributed DS neurons is sufficient to explain the results predicted by the properties of looming detectors. The preponderance of evidence at this point, however, argues in favor of an arrangement similar to that of Figure 17 (See, for example, 33).



A Basic Looming Detector
Figure 17

3.5 Summary

In this chapter we introduced the brain's fundamental processing element, the neuron, noting the differences between it and other cells and describing its signaling capabilities and characteristics. We next described the receptive field structure and identified its most distinctive feature: its sensitivity to stimulus contrast as opposed to absolute stimulus intensity. We described how light signals entering the eye are transformed and processed by progressively more complex and sophisticated receptive field structures to extract information from complex motion stimuli. Our review culminated in the presentation of physiologically grounded neural models for motion detection (Figure 15b) and looming detection (Figure 17). These two models form the physiological and theoretical basis for this study.

The construction and response characteristics of OS neurons, DS neurons, and looming detectors reveal a general theme that applies throughout the visual system: downstream neurons combine the outputs of simpler upstream neurons to form more sophisticated and complex receptive field structures, with which to better discriminate and disambiguate the complex visual signals that the eyes receive. Here we have seen that OS neurons integrate the outputs of RGC neurons to form OS receptive fields, DS neurons integrate the outputs of OS neurons to form DS receptive fields, and looming detectors integrate the outputs of DS neurons to form receptive fields sensitive to image expansion.

4 NEURAL AND PSYCHOPHYSICAL RESPONSE TO TIME-VARYING VISUAL STIMULI

There are two important conclusions that can be drawn from our review of the motion pathway that form the logical basis for this study's approach:

1. *All visual detection tasks (including the detection of motion and looming) are accomplished by specialized receptive fields that are built up from RGC receptive fields. Their response characteristics should therefore be similar to the RGC's basic response characteristics.*
2. *Luminance patterns that may vary over space and/or time are the only stimuli that RGC receptive fields are sensitive to. The detection of any feature or event in the environment must therefore reduce to the detection of a particular luminance pattern or patterns. Thus, the response characteristics of any detection task (such as motion or looming) should therefore be similar to the RGC's response characteristics to luminance variation.*

In this chapter we present two sets of experiments to demonstrate these similarities. We first demonstrate the similarity between the frequency response characteristics of an RGC and a human test subjects to temporally varying light stimuli. We next demonstrate the similarity between the frequency response characteristics of human test subjects to temporally varying light stimuli and motion stimuli.

4.1 Neural and Human Psychophysical Response to Flicker

Flicker refers to the periodic temporal variation of a spatially uniform light stimulus. If the variation is sinusoidal then the luminance of the stimulus can be described by

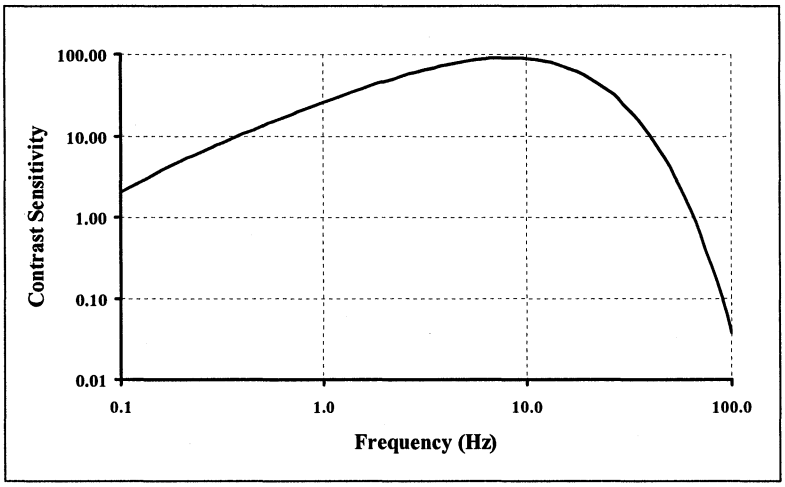
$$\ell(t) = \ell_{avg}\left[1 + \frac{a}{\ell_{avg}} \cdot \sin(2\pi f_t\, t)\right] = \ell_{avg}\,[1 + c(t)], \qquad (11)$$

where ℓ_{avg} is the average level of luminance, a the amplitude of the sinusoidal variation, and f_t the temporal frequency in cycles per second. An RGC's frequency response characteristic to such a stimulus can be obtained by determining the contrast amplitude $a(f_t)/\ \ell_{avg}$ at each frequency required to elicit some steady state *criterion level of response* from it. The criterion level of response is an arbitrarily assigned spike rate. It is set to a

value sufficiently far away from the neuron's spontaneous firing rate to allow the two to be reliably (in a statistical sense) distinguished from one another, given the presence of noise. (We identified two sources of noise in Sections 3.2 and 3.3, and will discuss this further in Section 5.4.1). The contrast amplitude necessary to evoke this response is called the *contrast threshold*, and its reciprocal the *contrast sensitivity*.

4.1.1 Neural Response to Flicker

Derrington and Lennie (34) performed an experiment whose stimulus was very similar to that described by Equ. (11)[6], using the RGC of a macaque monkey. (the corresponding experiment can not be performed on human subjects, but the behavior of human RGC's is believed to closely resemble that of the macaque.) This produced the result shown in Figure 18. As was the case with receptive field response to spatially



Temporal Contrast Sensitivity Function for a Macaque Monkey RGC
Figure 18

[6] In their experiment the stimulus varied in both space and time according to

$$\ell(x,t) = \ell_{avg}\left[1 + \left(.02 \cdot \sin(5\pi x) \bullet \frac{a}{\ell_{avg}} \cdot \sin(2\pi f_t t)\right)\right],$$

where x is measured in degrees of visual angle.

varying contrast patterns (Section 3.3), response to temporal contrast patterns is also linear for a given adaptation level and small changes in contrast (35).

The low-frequency cut-off exhibited by the RGC is noteworthy. Based upon the considerations outlined in Section 3.2 (neural coding and conduction time delays) we expect a high frequency cutoff. In fact, at very high temporal frequencies the stimulus takes on the appearance of a uniform field whose equivalent luminance ℓ_{eq} is equal to the time averaged luminance, that is,

$$\ell_{eq} = \frac{1}{T}\int_0^T \ell(t)\,dt\,. \tag{12}$$

This is a statement of the *Talbot-Plateau Law*. Different sized receptive fields will have different numbers of neurons and different transmission distances associated with them, and will thus have different temporal characteristics in this regard. There are no obvious physiological grounds, however, for a low pass cutoff.

We do know, however, that the visual system is tuned to identify those features of the visual scene that convey the most information, i.e., features that are undergoing significant change over space/time. Other elements of the scene are typically passed over and ignored. The low-frequency cut-off here may be a manifestation of this, indicating that objects whose luminance changes very slowly over time are not of much interest and that information regarding them is thus being discarded in favor of those whose luminance changes more quickly.

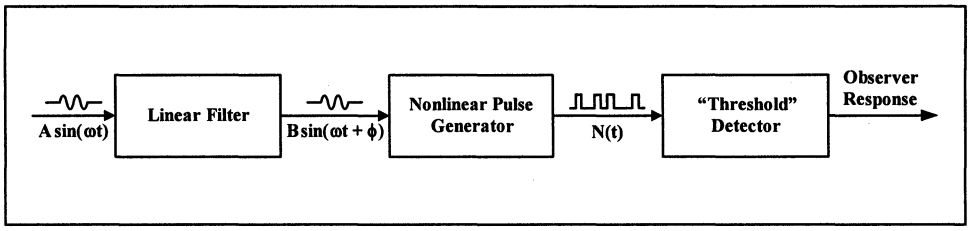
4.1.2 Human Psychophysical Response to Flicker

As with an RGC's response to temporally varying stimuli, above some *critical frequency* a human observer can no longer perceive the temporal variation and instead senses a constant and uniform field (*fusion*) whose luminance is given by Talbot-Plateau's Law [Equ. (12)]. The critical frequency at which this transition takes place is given by *Porter's Law*:

$$f_c = k_1 \cdot \ln \ell_{eq} + k_2 , \tag{13}$$

where ℓ_{eq} is the time averaged luminance of the stimulus [from Eq. (12)], or adaptation level, and the constants k_1 and k_2 depend on the form of the temporal variation, the adaptation level, and the location on the retina that the stimulus' image falls on. In their experiments, DeLange (36) and Kelly (37) established for various flicker frequencies f_t the amplitude a for which human test subjects could "just detect" flicker. Though these experiments eschew spatial variations in stimulus luminance, they are nonetheless comparable to the result described in previously (Section 4.1.1). These experiments and the follow-on analysis of Stork and Falk (38) served as prototypes for the motion detection experiments and analyses performed in this study, given their close similarity in both concept and design.

We have seen in Section 3.2 that the RGC receptive field acts as a temporal filter. Furthermore, if Talbot-Plateau's Law [Equ. (12)] is to be satisfied, this filter must be linear and precede any nonlinear element. Based on these observations DeLange and Kelly both assumed the same underlying neural mechanism for the psychophysical flicker detection task. It consists of a linear filter followed by a nonlinear pulse generator and a "threshold" detector, as shown in Figure 19. Flicker input [$A\sin(\omega t)$] to the visual



Generic Flicker Detection Model Employed by DeLange and Kelly
Figure 19

system is temporally filtered by the RGC receptive field and converted to graded potentials (Section 3.2) by the photoreceptors making it up. The filter's continuous output [$B\sin(\omega t+\varphi)$] is the combination of these graded potentials. The nonlinear pulse generator transforms the filter's output to a series of action potentials whose frequency [N(t)] codes the strength of the filter output. The pulse generator's output is assumed to vary instantaneously and monotonically with its input. There is therefore no information loss through the pulse generator, and the transient performance of the filter/pulse generator is just that of the filter. The pulse generator represents all of the neural elements between the photoreceptors and the threshold detector, which is assumed to be located somewhere before the visual cortex at the back of the brain. DeLange noted that as a result of the requirement that the linear element (the filter) precede the nonlinear element (the pulse generator) the two will behave as a linear system to stimuli at or above the critical frequency, whose average luminance is constant. Under these conditions the filter attenuates any varying input to such a small fluctuation about its mean that the pulse generator output N(t) is approximately linearly related to its input $B\sin(\omega t+\varphi)$.

Neither researcher addressed the threshold detector in detail. Kelly described it as a neural element beyond which information required for an appropriate detection response

could not pass, under sub-threshold conditions[7]. Their models applied only up to the input of the threshold detector. Even so, the models were not viewed as inherently limited to threshold stimuli, and Kelly applied his model to quite different stimuli (e.g., impulse and step inputs).

Neither DeLange nor Kelly provided for the existence of pure transmission delays, which are known to exist for all visual processes (and which the engineering models described in Section 2 correctly incorporated.) Temporal response to a stimulus, then, begins with the onset of the response, not the presentation of the stimulus. Methods can be devised to measure this time delay (see Section 9.1).

The linearity of the filter permits the use of classical frequency response techniques to measure its frequency response to periodic stimuli. A straightforward application of these techniques, however, is not possible because the output of the filter cannot be measured directly, only the observer's response. Using the assumptions underlying their model, however, DeLange and Kelly were able to circumvent this difficulty. Recall first that if we excite a linear system with a sinusoidal input of the form $A\sin(\omega t)$, its response will be a sinusoid of the form $B\sin(\omega t+\varphi)$. Repeating the process for a number of different frequencies provides the system's gain and phase characteristic, $B/A(\omega)$ and $\varphi(\omega)$. Since the system is linear, the output amplitude remains proportionate to the input amplitude for each frequency. The ratio B/A is thus unaffected by the particular value that we choose for A. We could therefore keep A constant for each frequency that we test for, or we could vary A so as to keep the output amplitude B constant. Since under

[7] This represents the historical view of an *absolute* threshold, which is described in Section 5.2.

these experimental conditions the criterion value C is assumed to remain constant, the input amplitude A can be adjusted until the maximum pulse generator frequency N_{max} is just equal to C. N_{max} will thus remain constant throughout the experiment, as will the filter's output amplitude B that produced it. The fact that the magnitude (and for that matter, the units) of B remains unknown is not critical to the analysis, since it remains constant throughout the experiment. It can thus be taken to be unity. (This means that the derived results will be known to within a multiplicative constant.)

4.1.2.1 DeLange's Electrical Analog Flicker-Fusion Model

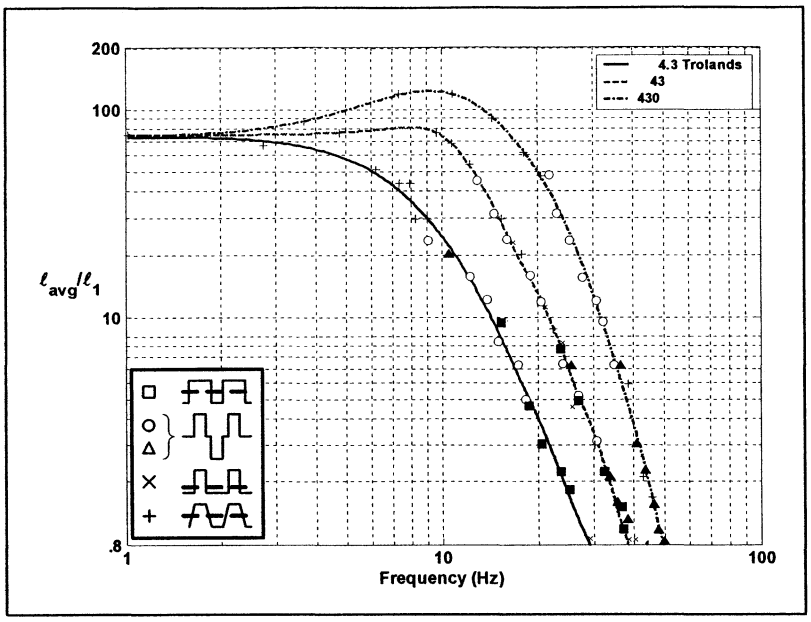
Delange is credited with being the first to apply linear systems theory to visual psychophysics, and with it unified a substantial body of data gathered by previous investigators. Using his own experimental results he showed that the visual system behaves approximately linearly near the flicker detection threshold.

DeLange measured the just-detectable amplitude of a flickering stimulus consisting of a bright, circular, sharp-edged spot of light superimposed upon a dark background[8]. The central spot fell on the central 2° of the retina[9]. This stimulus was modulated according to four different square wave patterns (shown in the inset to Figure 20). The luminance of the background was made equal to the time-averaged luminance of the central spot, such that background luminance also set the adaptation level. The frequency response characteristic that DeLange measured is shown in Figure 20. He found that for all but the lowest flicker frequencies the filter's attenuation characteristic was so great that only the first harmonic of the stimulus, given by

---

[8] In these experiments stimulus luminance is measured in trolands, the retinal illuminance that results when a surface luminance of 1 cd/m^2 is viewed through a pupil area of 1 mm^2.

[9] The fovea occupies the central 2° of the retina, so this stimulus entirely fills it.



DeLange's Experimental Results
Figure 20

$$\ell(t) = \ell_{avg}\left[1 + \frac{\ell_1}{\ell_{avg}} \cdot \sin(2\pi f_t\, t)\right],$$

was perceptible to the observer. The ordinate in the figure is thus ℓ_{avg}/ℓ_1. Detection performance, for the 430 troland case in particular, approximately mirrors that previously presented for an individual RGC (Figure 18). Peak sensitivity occurs at approximately the same frequency. Low frequency attenuation is not as pronounced as it is for the RGC, and high frequency attenuation declines more rapidly.

Based in these results DeLange postulated an electrical analog for the linear filter, consisting of n, loosely coupled RC elements. (Only resistors and capacitors were chosen for this model because the filtering action is in reality a chemical process that includes diffusion, and these processes are governed by the same laws as electrical networks containing only resistances and capacitors.) The individual resistance and capacitance values can in principle be different from one another, but for convenience DeLange

assumed each type of element had the same value, which along with n were determined by the adaptation level. The transfer function for the filter thus takes the form

$$G(s) = \frac{1}{(1 + RCs)^{n/2}}, \tag{14}$$

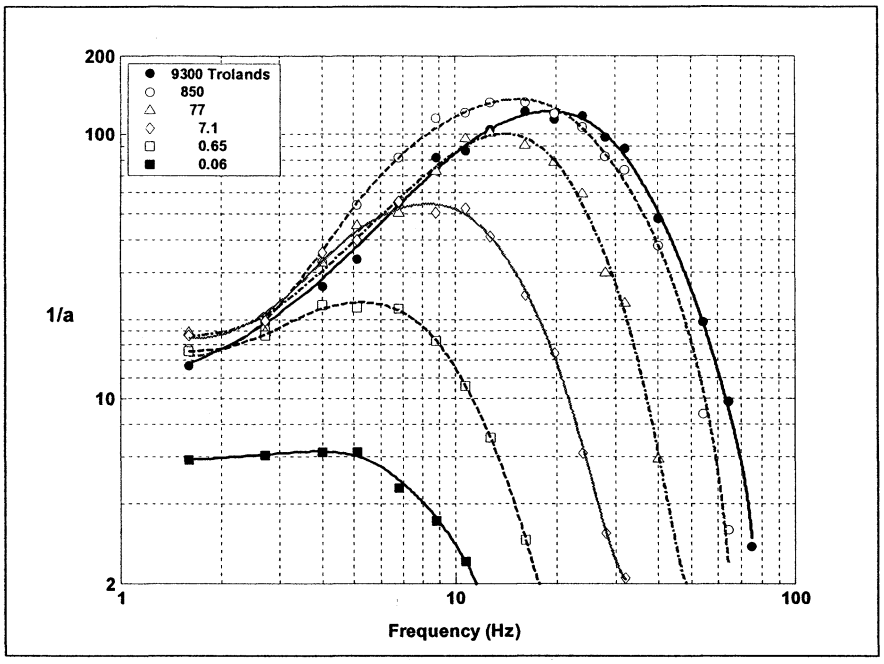
where s is the Laplace variable. For the 430 troland data n=12 and RC=.0035 sec.

4.1.2.2 Kelly's Single Channel Flicker-Fusion Model

In contrast to DeLange's stimulus, Kelly employed a wide field, edgeless stimulus. Here the instantaneous luminance of the flickering field was uniform for the first 50° of visual angle. Between 50° and 68°, the luminance was gradually reduced to zero, giving the boundary of the flickering field a blurred appearance. Kelly's experiments also differed in that he used a purely sinusoidal stimulus, whose luminance varied according to

$$\ell(t) = \ell_{avg}\left[1 + \frac{a}{\ell_{avg}} \cdot \sin(2\pi f_t t)\right]. \tag{15}$$

Kelly's results are shown in Figure 21. Both the low and high frequency attenuation exhibited here are more consistent with that of the RGC (Figure 18) than DeLange's data (Figure 20), but peak sensitivity occurs at a lower frequency. He attributed the more pronounced bandpass characteristic of his response curves primarily to the elimination of a sharp stimulus edge. This precludes the effect of "spatial derivatives" on detection performance and gives results that are more consistent with those obtained from individual *retinal channels*, which Kelly defined as some retinal spatial unit (i.e., one or a group of photoreceptors) that transmits signals independently of its neighbors, up to a



Kelly's Experimental Results
Figure 21

certain point. The "single channel" model that Kelly derived from this data, then, should be representative of an RGC receptive field.

Based on his results and physiological considerations, Kelly assumed a filter described by a minimum phase transfer function and constrained by the following observations and conditions:

1. It behaves as a second order differentiator for low frequencies and a second order integrator for high frequencies.
2. $g(t=0)=0$ and $\left.{}^{dg}/_{dt}\right|_{t=0} < \infty$, where g(t) is the filter's temporal response.
3. Since neither the magnitude nor the units of the filter's output were known, Kelly for convenience assumed it to be numerically equal to the input at zero frequency, so that $G(0)=1$.

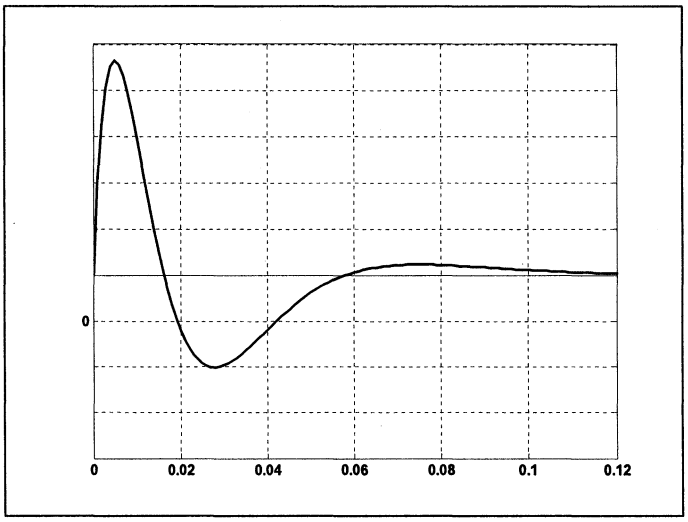
With these the filter transfer function describing the 850 troland response characteristic becomes

$$G(s) = k^4 \cdot \frac{as^2 + bs + 1}{(k+s)^4} = (89)^4 \cdot \frac{.00372s^2 + .0725s + 1}{(89+s)^4}. \tag{16}$$

Taking the inverse Laplace Transform of Equ. (16) gives the impulse response,

$$\begin{aligned} h(t) &= k^4 \left[at + \frac{1}{2}(b - 2ak)t^2 + \frac{1}{6}(1 - bk + ak^2)t^3 \right] e^{-kt}, \\ &= 2.3 \cdot 10^5 \left[t - 79.2t^2 + 1071t^3 \right] e^{-89t}. \end{aligned} \tag{17}$$

h(t) is plotted in Figure 22. Since, as we discussed in Section 4.1.2, we know neither the



Kelly's Single Channel Impulse Response
Figure 22

magnitude nor the units of the filter's output, the ordinate of this plot reflects relative scaling. Further, since transmission delays have not been addressed in this model the initial time in the plot coincides with the onset of the response, not the presentation of the stimulus.

4.1.2.3 Stork and Falk's Model-Independent Impulse Response Derivation

In their studies DeLange and Kelly developed physiologically based models whose free parameters were then specified with reference to experimental data. The danger of such an approach is that the experimenter's task can become one of fitting the data to the model, causing experimental results that are inconsistent with the model to be obscured. Stork and Falk employed the Kramers-Krönig (39) relations to develop a psychophysically determined temporal impulse response so as to place strict model-independent constraints on these models.

Let f(z) be a complex function given by

$$f(z) = f(x + iy) = u(x,y) + i\,v(x,y). \tag{18}$$

The Kramers-Krönig relations show that under a given set of assumptions (analyticity and boundedness), the real and imaginary parts of $f(x,0) = u(x) + i\,v(x)$ are Hilbert Transforms of one another. In their simplest form they can be written as

$$\begin{aligned} u(x_0) &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{v(x)}{x - x_0}\,dx, \\ v(x_0) &= -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{u(x)}{x - x_0}\,dx, \end{aligned} \tag{19}$$

where the Cauchy principle value of the integral is assumed. This is actually a statement of Titchmarsh's (40) Theorem. Let $G(\omega)$ be the Fourier Transform of $g(t)$:

$$G(\omega) = \int_{-\infty}^{\infty} g(t)\, e^{i\omega t}\, dt.$$

Titchmarsh's Theorem states that if $G(\omega)$ is square integrable over the real ω–axis, then any one of the following three conditions implies the other two:

1. The inverse Fourier transform g(t) of G(ω) is zero for t<0 (causality);
2. Replacing ω by z=x+iy, G(z) is analytic in the complex z–plane for y>0 and approaches G(x) almost everywhere as y→0. Further

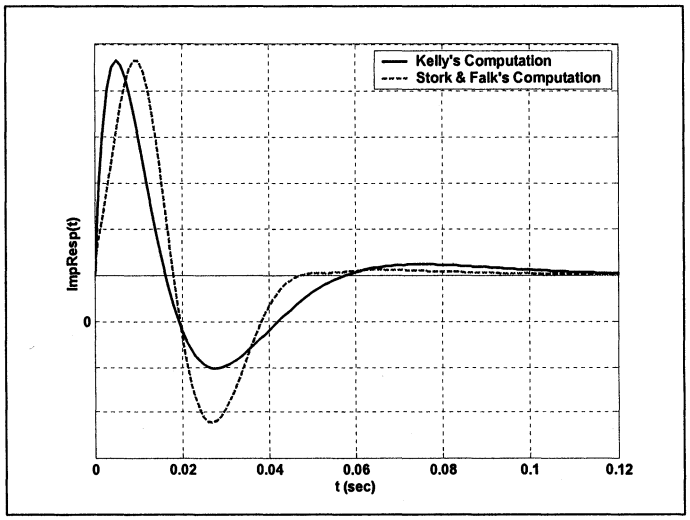
$$\int_{-\infty}^{\infty} |G(x + iy)|^2 \, dx < C \qquad (y > 0) .$$

That is, the integral is bounded.

3. The real and imaginary parts of G(z) are Hilbert transforms of one another [Equs. (19)].

Thus if G(ω) is causal (condition 1), and absolutely stable (condition 2), then by Titchmarsh's Theorem, the third condition applies. This result is generally applicable to any analytic, bounded, complex function, independent of the details of its particular application.

Using these relations Stork and Falk derived a psychophysically-based impulse response corresponding to Kelly's data for 9300 trolands (Figure 21). Their result is plotted in Figure 23. For comparison Kelly's result (Figure 22) for 850 troland is also plotted. The two approaches give quite similar results.



Stork and Falk's Model-Independent Impulse Response
Figure 23

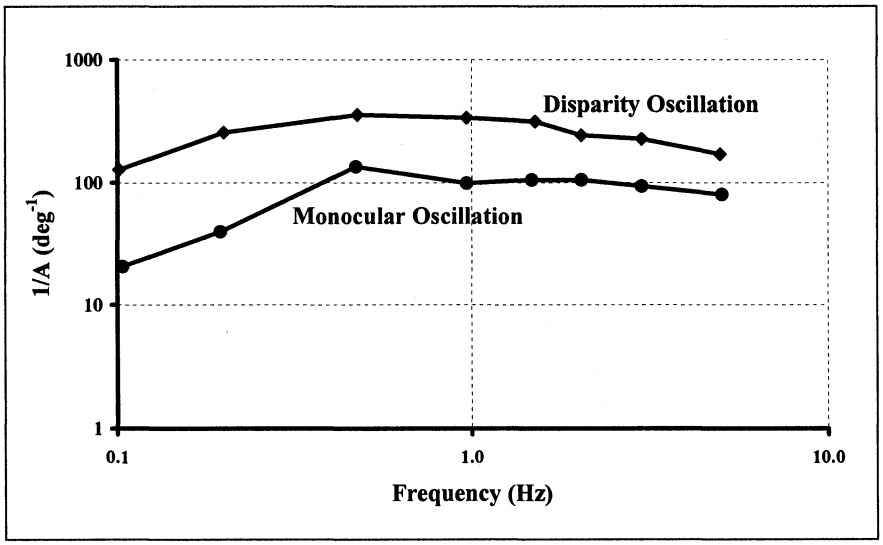
The conditions of Titchmarsh's Theorem are also satisfied by a minimum phase transfer function. Such systems are optimal in the sense that they exhibit the fastest response of any system having the same gain characteristic. It is not possible to verify that this aspect of the human visual system is minimum phase without employing invasive experimental techniques to actually measure its phase characteristic. It is reasonable to assume that this is the case, however, since the detection of luminance variation is a critical task for the survival of any organism. The ability to respond to such a stimulus in the shortest possible time would thus confer an advantage to those organisms possessing this capability and one could conclude from this that the visual system has evolved to exhibit minimum phase behavior. Based on these considerations we employ the assumption here that the system under study here is minimum phase. The Matlab μ-tools function *fitmag* (41) constructs a minimum phase transfer function of a specified order using as its input the system gain characteristic, based upon the algorithm described by Oppenheim (42). Using *fitmag*, a second order transfer function was fitted to Kelly's 9300 troland data. Its impulse response was identical to Stork and Falk's result.

4.2 Human Psychophysical Response to Motion

4.2.1 Sensitivity to Monocular and Disparity Oscillation

Tyler (43) measured human visual sensitivity to two types of apparent motion. For each the stimulus consisted of two bright, thin lines drawn parallel to one another on an oscilloscope display. Each was then displaced laterally in a sinusoidal fashion at the same frequency and amplitude. The test subject viewed the display through a set of conventional orthogonal polarizers. When the lines were modulated in phase with one another, one line appeared to be stationary while the other moved sinusoidally from side

to side (*monocular oscillation*). When they were modulated out of phase, one line continued to appear stationary, but now the other appeared to move sinusoidally forward and backward in space (*disparity oscillation*). For a range of frequencies between .1 and 5 Hz, the test subject adjusted the amplitude of the modulation to the level at which it was just detectable. Detection performance for a single observer is shown in Figure 24 (where A is the just detectable modulation amplitude in degrees of visual angle).



Sensitivity to Monocular and Disparity Oscillation
Figure 24

Neither of these motion stimuli would be detected by the looming detector (Figure 17) that we assume will be active in the experiments we will conduct. The monocular motion stimulus, however, would be detected by the basic Reichardt detector (Figure 15b), that we assume makes up our looming detector. This experiment are therefore quite relevant to our own, and we expect our results to be somewhat similar to these. The stereoscopic motion stimulus would be detected by neurons sensitive to binocular disparity, whose study we have deferred (See Section 9.1). A comparison of the two results does confirm our prioritization: sensitivity to the monocular stimulus is

approximately three times a great as that to the stereoscopic stimulus. We expect that this would carry over to the case of looming detection as well, and that even if these two cues were pooled by the visual system in the detection of looming, the improvement due to the stereoscopic cue would not be that great.

4.2.2 Sensitivity to Changes in Lateral Velocity

Harrison, et al (44) measured sensitivity to periodic velocity modulation of the base lateral velocity of an object for three periodic waveforms (sine, square, and triangle). For modulation frequencies ranging from one to nine Hz and base velocities between three and nine deg/sec. They found that human test subjects were able to detect each type of modulation equally well. Frequency response characteristics were bandpass with peak sensitivity at approximately three Hz. These results are again similar to those of the previous experiments.

4.3 Summary

In this chapter we showed that human performance in the flicker detection task is similar to that of the RGC (as measured by frequency response). We then showed that human performance in the detection of certain motion stimuli are in turn similar to flicker detection performance. (These motion stimuli are similar but not identical to the looming stimuli that we will be using in the present study.) These results are expected given the observations (Section 3.5) that all detection mechanisms are built up from RGC receptive fields, and that all visual stimuli reduce to changing luminance patterns. We then reviewed the flicker detection models proposed by DeLange (36) and Kelly (37). Based upon these results and observations we conclude that the basic structure of these models is applicable to the looming detection task. Finally, Stork and Falk's (38) follow-on analysis of Kelly's data showed that the assumption of minimum phase behavior is

justified, and presents an equivalent, simpler approach to the development of the filter model.

PART II: EXPERIMENTAL PROCEDURES AND RESULTS

Introduction

As noted in the Introduction (Section 1), expansion of an object's image in an observer's visual field, as represented by $\dot{\theta}$ increasing from zero, provides a powerful sensation of its approach. This isn't the only visual cue available for the detection of looming objects, but it is generally regarded as one of the most important. In these experiments we consider only the cue of two-dimensional image expansion (which needs only one eye to appreciate), and we study it in isolation from any other. Because it is two-dimensional, presentation of an expanding shape on a two dimensional visual display produces qualitatively the same effect as an object approaching an observer in three-dimensional space. This more tractable approach will be taken in these experiments.

A complete model would of course combine the results of separate studies quantifying the visual system's detection and response capabilities for other cues to looming. Two that we will discuss later (Section 0) are:

1. "Second order" motion;
2. Binocular disparity.

5 A TEST OF THE THRESHOLD ASSUMPTION

5.1 Introduction

It is generally accepted that humans detect and avoid impending collisions through the perception and regulation of *Time-to-Collision* (τ). Judgments of τ are in turn assumed to involve judgments of the speed at which the FV and LV are approaching one another. Earlier (see Section 1) we showed that when an object approaches an observer the monocular image that it projects onto the retina of the observer's eyes expands at a rate $\dot{\theta}$ given by

$$\dot{\theta} = -\frac{W\dot{R}}{R^2} = -\frac{\theta\dot{R}}{R}, \tag{20}$$

where θ is the visual angle subtended by an object of width W, at a distance or range R, and approaching the observer at a speed, or range rate $\dot{R}$. τ is in turn given by

$$\tau = \frac{\theta}{\dot{\theta}}. \tag{21}$$

[See Equs. (1) and (3)]. In their study, Hoffman and Mortimer (16) assert the existence of a just-noticeable "threshold" in the perception of $\dot{\theta}$, of approximately 0.003 rad/sec. Fancher and Bareket (12) apply this to Equ. (20) to assert the existence of a threshold in the perception of $\dot{R}$ of approximately $\pm 0.00164R^2$, which they employ in their Headway Control Model (Section 2.3).

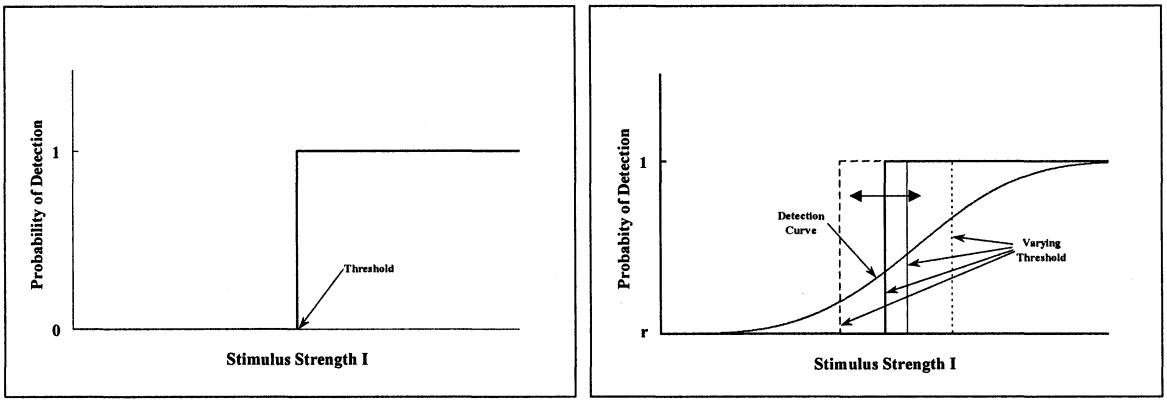
Implicit in these treatments is the assumption that the *calculations* of $\dot{\theta}$ and τ describe how a human observer actually *perceives* these quantities. In reality, these are two different concepts. The *equations* given above for $\dot{\theta}$ and τ are concepts of geometry, while the *percepts* $\dot{\theta}$ and τ obtained by a human observer in performing a certain

judgment are *behaviorally determined* quantities. The assumption linking the two, while attractive, is unfounded. Certainly, no experimental evidence has been presented in support of it, and in fact a number of studies have presented evidence contradicting it. Regan and Hamstra (5), for example, have studied the intrinsic visual act of judging τ and have found evidence for independent mechanisms in the human brain sensitive separately to τ and to $\dot{\theta}$ (though, as we noted in Section 1, the existence of a neural correlate to τ has yet to be conclusively proven). A second assumption, again presented without evidence, is that there is a threshold associated with the detection of $\dot{\theta}$, and below which the human can perceive nothing. Again, a number of studies have presented contradictory evidence. Evans and Rothery (45) showed evidence that drivers were sensitive to relative speed and spacing even when angular change was "too small" to be seen. Schiff and Detwiler (46) have also pursued studies using the threshold concept. These issues have significance for human driver models because, if the assumptions are incorrect, the models will not effectively convey what drivers actually do, even if calculations based upon them are approximately accurate.

In this study we conducted two experiments to test the existence of a threshold for detecting $\dot{\theta}$. These experiments are based on the methods of Tanner and Swets (47), who investigated the existence of a "classical" threshold associated with the detection of light signals. (As we explained in Section 4, all visual detection tasks ultimately reduce to this.) The implications that these results have for human driver models are far reaching: they suggest the need for a fundamentally different kind of model whose output lends insight not only into the manner in which humans perform this driving task, but also the ways in which their performance can go awry.

5.2 Background

The concept of threshold was first formalized by Gustav Fechner (48), who asked the most fundamental question of all concerning the visual system: what is the least amount of light that can be seen? Fechner's hypothesis, that there must exist some limit below which a light signal was too weak to be perceived, is shown schematically in Figure 25a.



a. Absolute Threshold b. Variable Threshold

Figure 25

This plots an observer's ability to detect a stimulus as a function of stimulus strength, I. Below some threshold value, the visual system is unable to detect the stimulus and it goes unseen. Above this threshold, "phenomenal seeing" takes place, and the observer is always able to detect the signal. Fechner and others disproved this hypothesis, however, when they attempted to measure this threshold. Instead of measuring the same sharply defined threshold each time, different ones were measured at different times with no apparent pattern to the changes. Presenting an observer with a large number of stimuli at each of a number of different signal strengths did produce a fairly repeatable result—the "smeared" detection curve shown in Figure 25b. Notice first the lack of a distinct threshold, which when discovered was explained by the fact that the threshold is not static, but randomly fluctuates over time. Secondly, when stimulus strength is very low, the observer isn't sure if the stimulus is there or not, and so resorts to guessing. The

probability that he will correctly detect the stimulus approaches not zero but some guessing rate r. It is important to keep in mind, however, that even in this modified view a *specific* threshold is assumed to be in effect every time a stimulus is presented. The uncertainty reflected in the ordinate of Figure 25b reflects the uncertainty associated with the particular location of the threshold at any particular instant. The observer is still assumed able to detect the stimulus with complete certainty whenever the stimulus strength exceeds the threshold in effect at the particular instant the stimulus is presented in.

The repeatability of the experimentally derived detection curve of Figure 25b allows us to establish stimulus levels that will produce any desired detection rate. Formal procedures called staircase algorithms have been developed to accomplish this. This fact, along with the ideas of an instantaneous threshold and of guessing when the signal strength is below threshold, form the basis for the experiments performed here.

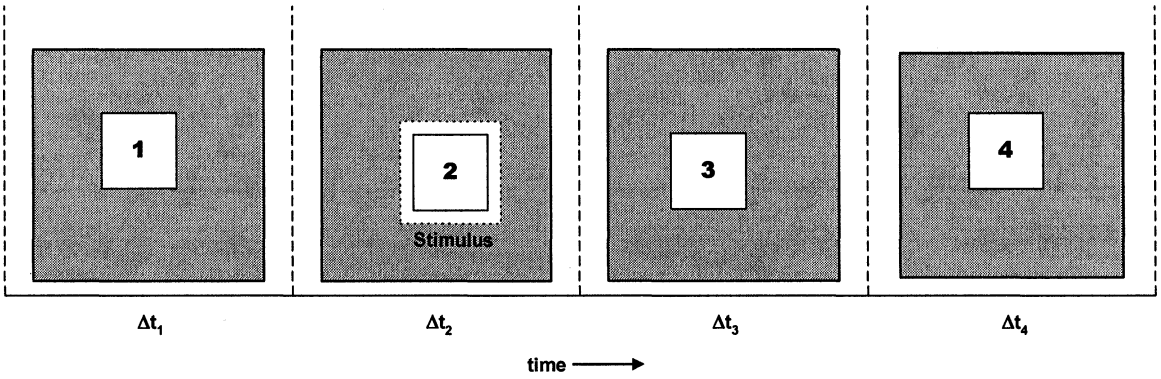
5.3 Experiment 1

Consider an experimental trial in which an observer is presented with a visual display divided into four time epochs. During a particular epoch a stimulus of strength I is presented and later the observer is asked to indicate which epoch it was presented in. If the stimulus strength was set to achieve a 75% detection rate, then over many such trials the observer will correctly identify the epoch 75% of the time. On those trials for which the observer chose incorrectly, the threshold concept asserts that he "saw" nothing. If the observer were asked on these trials to try again, then the probability that he could achieve a correct choice would be no better than chance—in this case ⅓. If, on the other hand, the observer managed to glean some information when the stimulus was shown (as for example, if there was no threshold), his success rate on the second try would be greater

than ⅓, contradicting the threshold assumption. In this first series of experiments we will test the null hypothesis, that the second choice success rate is ⅓, against the alternative that it is greater than ⅓.

5.3.1 Methodology and Procedure

Three young adults, all in their early 20's and having (corrected) normal vision, participated (they are identified as TS1, TS2, and TS3 in Table 2 and Table 3.) The stimuli were generated by a personal computer and consisted of four white squares presented sequentially against a gray background, an example of which is shown in Figure 26. Each of the squares was randomly displaced vertically and horizontally from



Stimulus Presentation
Figure 26

the center of the screen by a small amount (exaggerated in the figure) to prevent the observer from establishing any fixed points of reference. The time increment during which each square was presented was the same: 500 msec. One of the squares, chosen at random, expanded by a small amount on all sides. This expansion took place 200 msec after the presentation of the square, and occurred over a single screen refresh cycle (1/72 or 1/75 sec). (The expanded square thus remained visible for the last 300 msec of the 500 msec time interval.) Test subjects were instructed to fix their gaze on the number at the center of each square as it appeared. The initial size of the squares, the amount by which

it increased, and the distance between the test subject and the display were determined in pilot tests beforehand to achieve approximately a 75% detection rate on the first choice. All test parameters and the manner in which they were adjusted based on the pilot tests are given in Table 2a.

The small number of test subjects used in this study is justified by the nature of the tests performed. These tests are intended to assess the capabilities of the human visual system, and we know based on longstanding experience that these capabilities do not vary appreciably (in statistical terms) between individuals with normal vision, as selected here. (*Perception*, that is, the *interpretation* of the incoming sensory data, may indeed vary, but the neural events triggered by incoming photons of light do not.) Forced choice experiments ("Which square expanded?") involving stimuli of limited dimensionality (an expanding square) of the kind performed here minimize the test subjects' opportunity to voice subjective judgments and push them instead to accept one of four alternative hypotheses without elaboration. There are, of course, factors that do affect sensory function—age and neural disease being two prominent examples, and if we were investigating these effects we would want to diversify the type of subjects studied. This, though, is not the goal of the present study, which is intended to quantify the best performance that can be achieved by the normally functioning human visual system.

Each observer viewed the display in a darkened room, and took approximately 75 practice trials to become familiar with the stimulus display and test regimen. Five to eight 300-600 trial tests were then conducted over a three day period. The tests were somewhat tedious and thus were generally limited to 300 trials to ensure that fatigue did

Table 2

a) Experiment 1 Setup

#	Test	Resolution (mm/pix)	Refresh Rate (Hz)	Viewing Distance (mm)	Initial Square Size (Side)		Square Incremental Expansion (Side)		θ (rad)	Δθ (rad) x 10^{-3}	Δθ/θ x 10^{-3}	Δθ/Δt (rad/sec)
					Pixels	mm	Pixels	mm				
1	TS1-001	0.272	72	761	500	136	2	0.544	0.178	1.418	7.957	0.102
2	TS1-003	0.272	72	761	500	136	1	0.272	0.178	0.709	3.979	0.051
3	TS1-005	0.240	75	761	500	120	2	0.480	0.157	1.254	7.967	0.094
4	TS1-006	0.240	75	761	500	120	2	0.480	0.157	1.254	7.967	0.094
5	TS1-007	0.240	75	761	500	120	2	0.480	0.157	1.254	7.967	0.094
6	TS2-001	0.272	72	761	750	204	1	0.272	0.266	0.702	2.635	0.051
7	TS2-002	0.240	75	924	1,000	240	1	0.240	0.258	0.511	1.978	0.038
8	TS2-004	0.240	75	1,464	1,000	240	2	0.480	0.164	0.651	3.982	0.049
9	TS2-005	0.240	75	1,464	1,000	240	2	0.480	0.164	0.651	3.982	0.049
10	TS2-006	0.240	75	1,464	1,000	240	2	0.480	0.164	0.651	3.982	0.049
11	TS2-007	0.240	75	1,464	1,000	240	2	0.480	0.164	0.651	3.982	0.049
12	TS3-002	0.272	72	761	500	136	2	0.544	0.178	1.418	7.957	0.102
13	TS3-003	0.272	72	761	500	136	2	0.544	0.178	1.418	7.957	0.102
14	TS3-007	0.272	72	761	500	136	2	0.544	0.178	1.418	7.957	0.102

b: Experiment 1 Results

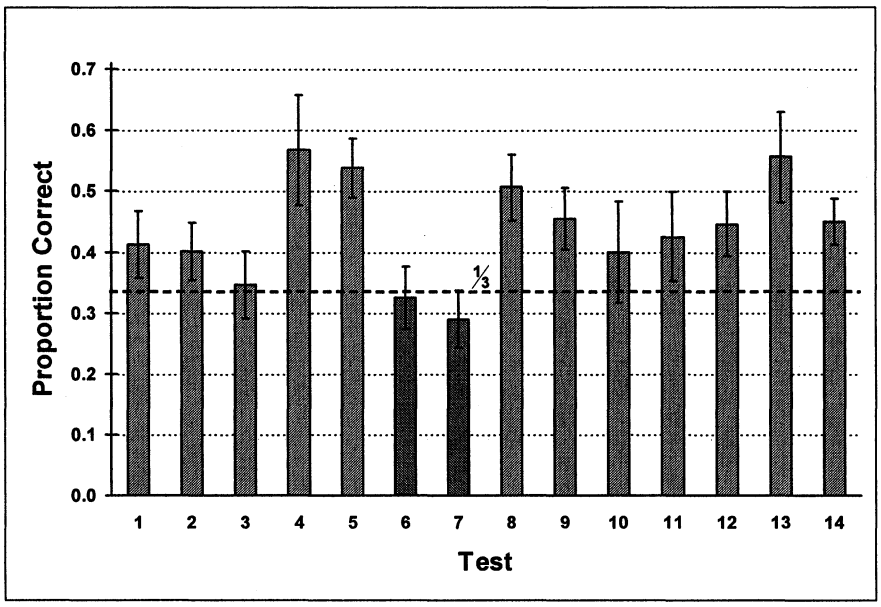
#	Test	1st Choice				2nd Choice (When 1st Choice Incorrect)				r
		# Trials n_1	# Correct n_{c1}	P_1	Std Error S_1	$n_2=n_1-n_{c1}$	n_{c2}	P_2	S_2	
1	TS1-001	300	220	0.733	0.026	80	33	0.413	0.055	-0.0879
2	TS1-003	300	188	0.627	0.028	112	45	0.402	0.046	+0.2804
3	TS1-005	300	225	0.750	0.025	75	26	0.347	0.055	-0.3321
4	TS1-006	300	270	0.900	0.017	30	17	0.567	0.091	-0.1040
5	TS1-007	600	492	0.820	0.016	108	58	0.537	0.048	+0.0609
6	TS2-001	300	217	0.723	0.026	83	27	0.325	0.051	+0.0815
7	TS2-002	300	207	0.690	0.027	93	27	0.290	0.047	-0.2859
8	TS2-004	300	213	0.710	0.026	87	44	0.506	0.054	+0.1769
9	TS2-005	300	201	0.670	0.027	99	45	0.455	0.050	+0.3926
10	TS2-006	300	265	0.883	0.019	35	14	0.400	0.083	-0.5551
11	TS2-007	300	253	0.843	0.021	47	20	0.426	0.072	+0.0759
12	TS3-002	300	208	0.693	0.027	92	41	0.446	0.052	+0.5095
13	TS3-003	300	255	0.850	0.021	45	25	0.556	0.074	-0.1996
14	TS3-007	600	420	0.700	0.019	180	81	0.450	0.037	+0.0843

not compromise the results. (Test subjects were also allowed to rest during the course of a test as needed.) On several occasions test subjects reported that they were not fatigued and could continue, so the test was extended to 400 or 600 trials. The tests took approximately 20 minutes per 300 trials to complete. The number of trials performed for each test is given in Table 2b. At the end of each trial the test subject was given two chances via prompts to identify the square that had expanded via the keyboard. Different pitched tones sounded to indicate correct vs. incorrect choices. (There was no time limit placed on the subjects' response, but they typically responded within a couple of seconds of each prompt.) Test subjects were paid 4¢ for each correct first choice, and 2¢ for each correct second choice (when the first choice was incorrect). This ensured that they would choose correctly on the first choice if they could, but also encouraged them to try their best on their second choice in the event that the first was incorrect. At these detection rates and payouts the test subjects could earn approximately $10 per 300 trial test. An experimenter was present at all times to monitor the test. The correct choice and the responses for each trial were logged into a data file for later processing.

5.3.2 Experiment 1: Results

We measured the ability to identify which of the four squares expanded. Performance was about 75% correct on the first choice. A second choice was also recorded. The results of these experiments are summarized in Table 2b and Figure 27. The individual trials making up a given test constitute Bernoulli Trials, for which the Proportion Correct and the Standard Error are given by

$$P_i = \frac{n_{ci}}{n_i}, \qquad S_i = \sqrt{\frac{P_i(1-P_i)}{n_i}},$$



Second Choice Performance
Figure 27

where i = 1,2 (see Table 2b). [For a review of all the statistical methods employed in this study, see (49).] If it is true that our test subjects perform no better than chance (⅓) on their second choice, then in a series of tests like the ones conducted here their second choice performance should be above chance about half the time and at or below chance the other half (assuming that the probability distribution the second choice results actually correspond to has a median of ⅓). The 14 tests performed here then constitute another set of Bernoulli Trials in which

$$\text{Probability}\{\text{Success}\} = \text{Probability}\left\{P_2 \le \frac{1}{3}\right\} = \frac{1}{2},$$

$$\text{Probability}\{\text{Failure}\} = \text{Probability}\left\{P_2 > \frac{1}{3}\right\} = \frac{1}{2}.$$

By this reasoning, then, we would expect only 7 of the 14 tests to yield correct response rates in excess of ⅓. We find, however, that 12 of the tests did. Since the underlying probability distribution for Bernoulli Trials is the Binomial Distribution, the probability

that we could encounter 12 (or more) failures if indeed the true probability of failure is ½ is given by

$$\text{Probability}\{x \geq 12\} = \sum_{x=12}^{n} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = .0065.$$

where n=14 and p=½. Since this probability is less than 1%, we can reject with 99% confidence that these subjects are able to do no better than chance on their second choice in favor of the alternative, that their performance is better than chance.

One possible explanation for the results of Experiment 1 is that test subjects, being aware of the experimenter's interest in the 2nd choice detection rate, "saved" responses for some trials that they were sure of for the 2nd choice. This would inflate the 2nd choice detection rate but at the same time reduce the 1st choice detection rate, causing the two to be negatively correlated. To investigate this, the data for each test was subdivided into 10 non-overlapping 30-trial "windows", and p_1 and p_2 calculated for each. The first window thus contains trials 1-30, the second window trials 31-60, and so on to the tenth window, which contains trials 271-300. The correlation coefficient r measures the correlation between p_1 and p_2, and is given by

$$r = \frac{s_{p_1 p_2}}{s_{p_1} s_{p_2}},$$

where the covariance and variance terms are given by

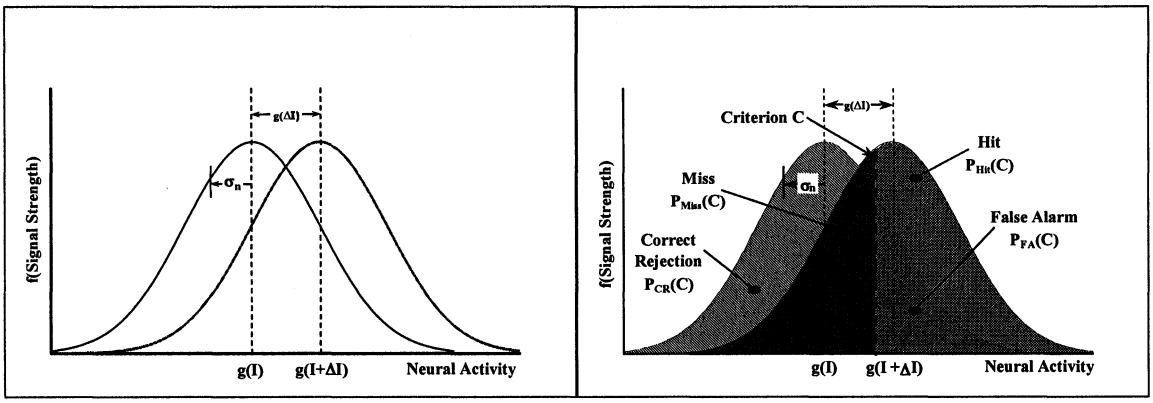
$$s_{p_1 p_2} = \frac{\sum_{i=1}^{n_w} (p_{1_i} - \bar{p}_1)(p_{2_i} - \bar{p}_2)}{n_w}, \qquad s_{p_1}^2 = \frac{\sum_{i=1}^{n_w} (p_{1_i} - \bar{p}_1)^2}{n_w}, \qquad s_{p_2}^2 = \frac{\sum_{i=1}^{n_w} (p_{2_i} - \bar{p}_2)^2}{n_w},$$

and where n_w is the number of trial windows and $\bar{p}_1$ and $\bar{p}_2$ the overall detection rates of the two series. The far right column of Table 2b gives the correlation coefficient for all of the tests. We can again construct a test of the hypothesis that there is no correlation between p_1 and p_2. Here, then, we would expect half the computed values of r to be non-negative and the other half to be negative. According to Table 2b this is almost exactly the result that we observe, leading us to conclude that test subjects were not saving correct responses for their second choice. Apparently, contrary to threshold theory, observers can do better than chance if given a second choice.

5.4 Experiment 2

5.4.1 Introduction

In their investigation, Tanner and Swets (47) hypothesized that the detection problem was really one of distinguishing between two signals in the presence of noise. In their experiments the first signal consisted of a uniform light background of luminance I, and the second consisted of a circular target of luminance $I+\Delta I$ superimposed on this background. The noise that they spoke of arises from many sources: it is intrinsic to all visual stimuli (Section 3.2), and physiological noise is also present in every phase of the visual detection process (Section 3.3). A useful way in which to treat noise, both for experimental and analytical purposes, is to aggregate it together from all its various sources and associate it with the domain of the visual stimuli, as indicated schematically in Figure 28a. A common assumption in this regard is that this "equivalent" noise is normally distributed and additive, so that both distributions have the same variance. Introduction of the target of strength $I+\Delta I$ thus has the effect of shifting the background distribution to the right by $g(\Delta I)$. The observer distinguishes between the two signals by selecting a criterion level C and deciding that neural activity corresponding to signal

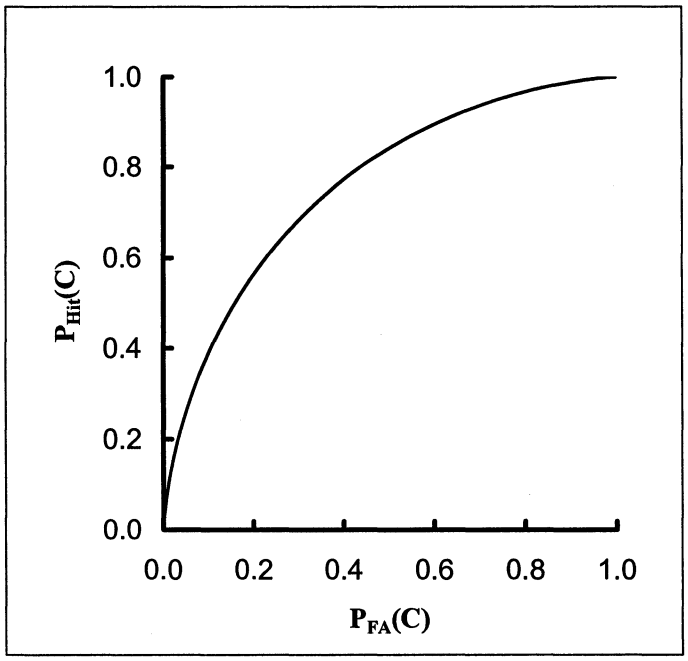


a. b.

Presentation of Visual Stimuli in the Presence of Noise
Figure 28

strengths less than C indicates background alone, while neural activity corresponding to signal strengths greater than C indicates background plus target[10]. In so doing, the observer accepts that only a fraction of all the targets presented will be correctly identified ("hits"). Those that aren't identified are labeled "misses". Furthermore, some background-only stimuli will be mistakenly included as well (a "false alarm"). Background-only stimuli that are correctly identified are labeled "correct rejections". These possible outcomes are indicated in Figure 28b. The actual level at which the observer sets the criterion will depend both on the expected benefit derived from each type of correct assessment of the signal and the expected cost incurred from incorrect ones. The observer can thus vary C (and be influenced to vary C), and in so doing obtain different proportions $P_{Hit}(C)$ and $P_{FA}(C)$. I and ΔI are under the control of the experimenter and thus known. If in addition the variance σ_s can be determined, then it is an easy matter to compute $P_{FA}(C)$ and $P_{Hit}(C)$ for different C (while holding I and ΔI fixed). Plotting them against one another gives a single *Receiver Operating Characteristic* (ROC) curve, as shown in Figure 29, thus allowing us to predict in

[10] Implicit in this is the assumption that neural activity is a monotonically increasing function of signal strength.



Receiver Operating Characteristic (ROC) Curve
Figure 29

advance the observer's performance for any C. (Repeating this calculation for different ΔI, leaving I constant, would yield a family of such curves.) False alarms are thus not guesses made in the absence of any information, but rather the result of an informed cost/benefit tradeoff. An easy way to influence this tradeoff, and hence the observer's selection of the criterion level, is to change the costs and benefits associated with the various outcomes of the detection task.

Threshold theory specifies a different relationship between P_{FA} and P_{Hit} [Green, Swets (50)], which we can use as the basis for a second test of its validity. Under the (varying) threshold theory, signals whose strength exceeds the threshold in effect at the instant they are presented will always be seen, and those that do not will never be seen. When we conduct experiments such as the kinds performed in this study and those performed by Tanner and Swets (47), we observe a hit rate less than the actual percentage of times a signal was presented, as well as a non-zero false alarm rate. The former is easily

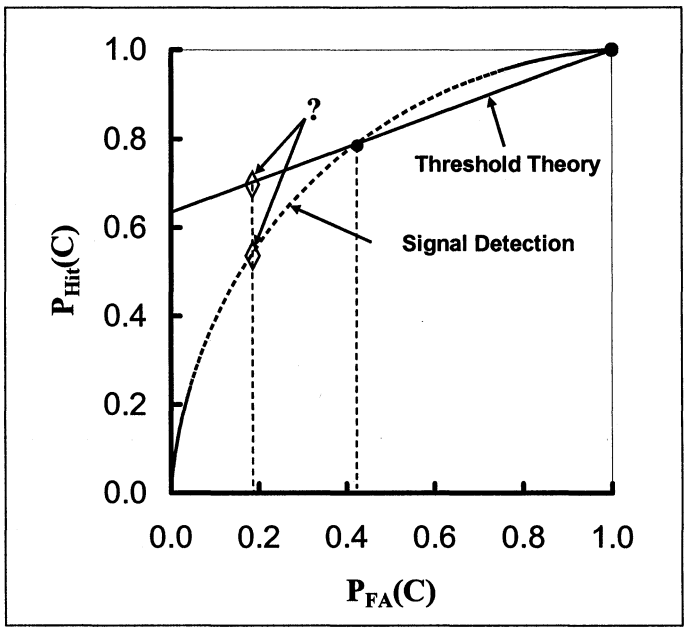
explained according to the theory already presented: the strength of the signal presented during a particular trial did not exceed the threshold in effect at the instant it was presented. The non-zero false alarm rate is more problematic, because according to this theory observers should never "see" a signal that isn't there. To explain this, threshold theory was augmented by the assumption that in addition to a varying threshold, a certain amount of guessing must be taking place as well. But if guessing inflates the false alarm rate (from zero), it must inflate the hit rate too. The observed hit rate, then, results from "true detections" and "guesses". The mathematical expression of this relationship is straightforward:

$$P_{Hit} = P^t_{Hit} + r\left(1 - P^t_{Hit}\right).$$

P_{Hit} here is the observed hit rate. P^t_{Hit} is the true hit rate, which depends only on I and ΔI (which are under the experimenter's control). r is a guessing factor that operates in the absence of a true hit. By definition, though, this is the observed false alarm rate P_{FA}, so we have

$$P_{Hit} = P^t_{Hit} + P_{FA}\left(1 - P^t_{Hit}\right).$$

Since P^t_{Hit} remains fixed for a given I and ΔI, the relationship between P_{FA} and P_{Hit} is linear. Experimentally obtaining one pair of points (P_{FA}, P_{Hit}) and observing that $P_{Hit}\left(P_{FA} = 1\right) = 1$ then allows us to plot this relationship. This is shown in Figure 30, along with the previously derived ROC curve for comparison. [We note too that by extrapolating this plot to the abscissa, we can obtain P^t_{Hit}, since $P_{Hit}\left(P_{FA} = 0\right) = P^t_{Hit}$.] These considerations lead us to another way to test the threshold theory: Perform the



Comparison of Threshold Theory and Signal Detection Theory
Figure 30

"rating scale" version [Green and Swets (50), described below] of the first experiment to obtain two pairs of points (P_{FA}, P_{Hit}). Construct a straight line passing through the rightmost pair and (1, 1), and plot it as in the Figure 30 (filled circles). If the leftmost pair (open diamond) falls on this line, it would lend support to the threshold theory. If it falls significantly below this line, however, it would contradict this theory and lend support to the signal detection theory. This was pursued in our second set of experiments.

5.4.2 Methodology and Procedure

The same subjects participated in this experiment as the first. The stimulus was generated by a personal computer and consisted of a single white square presented in the center of the display monitor, against a gray background. The square, which appeared for a total of 150 msec, expanded by a small amount on all sides or not, in a random fashion. If it expanded, it did so 75 msec after it first appeared, and the expansion took place over a single screen refresh cycle (1/72 sec), giving an angular expansion of approximately 0.05 rad/sec. The expanded square then remained visible for the last 75 msec of its presentation. Test subjects were instructed to fix their gaze at the center of the square as it appeared. At the end of each trial the test subject was asked to indicate whether or not the square had expanded in one of three ways:

- Certain
- Possibly
- Certain It Didn't

A tone sounded if the subject incorrectly chose *Certain* or *Certain It Didn't*. No tone sounded for *Possibly*. The subjects were paid +4¢ for a correct choice, +2¢ when they indicated *Possibly*, and penalized -3¢ for an incorrect choice.

The inclusion of the *Possibly* category is an efficient way to obtain two pairs of points (P_{FA}, P_{Hit}) in a single test. The *Possibly* responses represent those trials that the test subject was most ambivalent about. In the absence of this choice, he would have recorded these trials as either *Certain* or *Certain It Didn't*, depending upon the relative magnitudes of the reward and penalty offered. To see how two sets of points can be derived from this data, we first note all the possible stimulus-response outcomes for a trial:

- **Signal** (Square Expanded) ----------Certain---------- **Hit**
- **Signal** ----------Possibly ---------- **MaybeHit**
- **Signal** ----------Certain Didn't---------- **Miss**
- **NoSignal** (Square Didn't Expand)------Certain---------- **FalseAlarm**
- **NoSignal** ----------Possibly ---------- **MaybeFalseAlarm**
- **NoSignal** ----------Certain Didn't---------- **CorrectReject**

The two pairs of data points are then given by

$$P_{Hit,1} = \frac{\#\,Hit}{\#\,Signal}, \qquad P_{Hit,2} = \frac{\#\,Hit + \#\,MaybeHit}{\#\,Signal},$$

$$P_{FA,1} = \frac{\#\,FalseAlarm}{\#\,NoSignal}, \qquad P_{FA,2} = \frac{\#\,FalseAlarm + \#\,MaybeFalseAlarm}{\#\,NoSignal}.$$

Furthermore, because all of the data is derived from a single sequence of test trials, this scheme has the added benefit of ensuring that the same psychophysical parameters (e.g., motivation, fatigue, and external conditions) apply to each of the computed pairs of points.

In order for this scheme to provide the desired information, the various test parameters and payments/rewards must be set properly to influence the appropriate utilization of the *Possibly Expanded* category. Too few responses in this category (brought about by either a too easily detected square expansion and/or by an insufficient penalty for wrong choices) will yield data points that plot too close together in Figure 30, making it difficult to discern the leftmost point's location vis-à-vis the straight line threshold prediction constructed through the rightmost data point. Too many responses in this category, on the other hand, places the rightmost pair of points too close to (1, 1), calling into question the exact location of the straight line. For this reason the test parameters cited above (the presentation times, initial size of squares, amount by which the square expanded, and the distance between the test subject and the display) and

payments/penalties were adjusted from the stated nominal values to provide an adequate mix of responses for each test. This adjustment had to be performed at the beginning of each test because the performance of the test subjects could vary from test to test. (No adjustments were made during the course of a test.) The results of these adjustments for each test are indicated in Table 3a. The procedure was very similar to that of Experiment 1. After approximately 75 practice trials to become familiar with the stimulus display and test regimen, each test subject completed three to five 300 - 600 trial tests.

5.4.3 Experiment 2: Results

We measured the ability to detect a randomly expanding square according to a three-category rating scale. The results of these experiments are summarized in Table 3b. The far right column indicates whether the pair $\left(P_{FA,1}, P_{Hit,1}\right)$ (the open triangle in Figure 30) fell on/above (+) or below (-) the straight line constructed through $\left(P_{FA,2}, P_{Hit,2}\right)$ and $(1,1)$. Assuming that the threshold theory holds, the pair $\left(P_{FA,1}, P_{Hit,1}\right)$ should be collinear with $\left(P_{FA,2}, P_{Hit,2}\right)$ and $(1,1)$. For these tests, then, we would expect $\left(P_{FA,1}, P_{Hit,1}\right)$ to plot on or above this straight line half the time and below it the other half. Thus, we can again treat these tests as a set of Bernoulli Trials for which

$$\text{Probability}\{\text{Success}\} = \text{Probability}\left\{\left(P_{FA,1}, P_{Hit,1}\right) \text{ lies on or above the line}\right\} = \frac{1}{2},$$

$$\text{Probability}\{\text{Failure}\} = \text{Probability}\left\{\left(P_{FA,1}, P_{Hit,1}\right) \text{ lies below the line}\right\} = \frac{1}{2}.$$

According to Table 3b, we encountered only one "success" among the 12 tests (TS3-014). Since the probability of getting 11 (or more) failures when the true probability of success is ½ is

Table 3

a) Experiment 2 Setup

#	Test	Resolution (mm/pix)	Refresh Rate (Hz)	Viewing Distance (mm)	Initial Square Size (Side)		Square Incremental Expansion (Side)		θ (rad)	Δθ (rad) x 10^{-3}	Δθ/θ x 10^{-3}	Δθ/Δt (rad/sec)
					Pixels	mm	Pixels	mm				
1	TS1-009	0.272	72	1,040	500	136	1	0.272	0.131	0.521	3.989	0.038
2	TS1-010	0.272	72	1,040	500	136	1	0.272	0.131	0.521	3.989	0.038
3	TS1-011	0.272	72	1,040	500	136	1	0.272	0.131	0.521	3.989	0.038
4	TS1-012	0.272	72	800	750	204	1	0.272	0.254	0.669	2.638	0.048
5	TS2-011	0.272	72	800	750	204	1	0.272	0.254	0.669	2.638	0.048
6	TS2-012	0.272	72	800	750	204	1	0.272	0.254	0.669	2.638	0.048
7	TS2-013	0.272	72	800	750	204	1	0.272	0.254	0.669	2.638	0.048
8	TS2-014	0.272	72	640	750	204	1	0.272	0.316	0.829	2.622	0.060
9	TS2-015	0.272	72	640	850	231	1	0.272	0.357	0.823	2.303	0.060
10	TS3-013	0.272	72	640	1,000	272	1	0.272	0.419	0.813	1.942	0.059
11	TS3-014	0.272	72	640	1,100	299	1	0.272	0.459	0.806	1.755	0.058
12	TS3-015	0.272	72	640	1,100	299	1	0.272	0.459	0.806	1.755	0.058

b) Experiment 2 Results

#	Test	# Trials	$P^{o}_{FA,1}$	$P^{o}_{Hit,1}$	$P^{o}_{FA,2}$	$P^{o}_{Hit,2}$	On/Above (+) Below (-)
1	TS1-009	300	7.43%	69.08%	16.89%	80.92%	-
2	TS1-010	300	12.08%	78.15%	24.16%	86.09%	-
3	TS1-011	300	20.27%	58.55%	29.73%	70.39%	-
4	TS1-012	300	17.93%	45.16%	53.79%	74.84%	-
5	TS2-011	300	3.52%	79.11%	11.97%	86.71%	-
6	TS2-012	300	2.07%	79.35%	9.66%	87.10%	-
7	TS2-013	600	10.25%	72.66%	18.94%	77.34%	-
8	TS2-014	600	18.71%	59.31%	42.58%	75.17%	-
9	TS2-015	400	11.52%	73.68%	29.32%	83.25%	-
10	TS3-013	600	15.11%	46.84%	37.16%	63.57%	-
11	TS3-014	600	17.28%	60.87%	39.20%	71.24%	+
12	TS3-015	300	20.78%	77.40%	36.36%	86.99%	-

$$\text{Probability}\{x \geq 11\} = \sum_{x=11}^{12} \frac{12}{x!(12-x)!}(.5)^x\,(1-.5)^{12-x} = .0032,$$

We can with 99% confidence reject that $\left(P_{FA,1}, P_{Hit,1}\right)$ is collinear with $\left(P_{FA,2}, P_{Hit,2}\right)$ and $(1,1)$, and accept instead that it lies below the line formed by these latter pairs of points. This result suggests that the data is better described by an ROC curve, consistent with a signal detection theory that includes underlying noise but no sensory threshold, than by the linear relationship derived from threshold theory.

5.5 Discussion

5.5.1 Implications for Threshold Theory

These experiments tested the assertion of the existence of a threshold associated with the detection of expansion rate, $\dot{\theta}$, and provide evidence contradicting it. Hoffman and Mortimer (16) asserted that the just-noticeable expansion rate is 0.003 rad/sec. This is a statement of static threshold theory, according to which rates below this value can't be detected at all while those above it can be detected all of the time. "Threshold" thus takes the form of the step function presented in Figure 25a. The expansion rates used in these experiments ($0.038 \leq \Delta\theta/\Delta t \leq 0.102$ rad/sec) are over an order of magnitude greater than Hoffman and Mortimer's threshold and yet correspond to an approximately 75% detection rate. This highlights the problem for human driver models that incorporate the absolute threshold assumption. Though at expansion rates less than 0.003 rad/sec detection levels may very well be near zero, human drivers will continue to err significantly in detecting expansion rates quite a bit larger than 0.003 rad/sec. The models, however, will fail to pick this up, and will instead predict perfect detection at these higher levels. The threshold concept is problematic on at least three levels because it is at variance with the large body of widely accepted evidence amassed by the visual

and perceptual sciences, particularly since Tanner and Swets explored the issue. The first difficulty, which was noted in the introduction to this study, is that implicit in this treatment is the assumption that Hoffman and Mortimer's *calculation* of $\dot{\theta}_{thresh}$ describes how a human observer actually *perceives* this quantity. These, as we pointed out, are two different concepts.

5.5.1.1 Static Threshold

The second difficulty has to do with the fundamental concept of a static threshold. As we also noted, the hypothesis' author, Gustav Fechner, and his contemporaries disproved it when they attempted to measure it. We now know that noise (a concept unknown to Fechner) is responsible for this result. In the absence of noise, it has been found that the photoreceptors of the eye are capable of sensing light right down to the level of a single photon. Hence, if we wish to speak of a static, absolute limit to detection, it would be the limit on the divisibility of light itself. As discussed in Section 4, motion, like all visual stimuli, is ultimately perceived by the photoreceptors of the eye as light signals of varying intensity. Thus we should be skeptical of an assertion of a threshold associated with any aspect of the visual system. In the present study, we disproved the existence of a static threshold in our experimental setup phase. If such a threshold existed it wouldn't have been possible to establish conditions under which an expanding square was detected 75% of the time—it would have either been detected all of the time or never. [In addition to the postulation of an expansion rate threshold, Hoffman and Mortimer (16) also asserted the existence of a threshold in perceiving changes in subtended visual angle of $\Delta\theta/\theta = .12$. Our study employed quantities in the range $0.002 \leq \Delta\theta/\theta \leq 0.008$ — 15-60 times *less* than this amount—which were routinely detected by all of the test subjects.]

We did not comment on this at the time and tested instead the more recent (but still unproved) concept of a varying threshold.

5.5.1.2 Prior Tests Purporting to Demonstrate a Threshold

The third difficulty has to do with the type of experiment Hoffman and Mortimer conducted to verify the existence of an expansion rate threshold: pairs of four-second film clips of oncoming vehicles were shown to test subjects, who were asked to identify the clip exhibiting the faster speed of approach (i.e., the relative speed $\dot{R}$ in the previously developed expressions for $\dot{\theta}$ and τ) and then to estimate the ratio of the faster to the slower speed of approach. The authors noted that "at low relative speeds, the subjects were unable to scale the relative speed between the vehicles". They thus defined a "threshold relative speed" $\dot{R}_{thresh}$ as "the point at which the subject starts to scale the relative speed, i.e. where the subjective speed increased above a ratio of unity". The threshold angular speed $\dot{\theta}_{thresh}$ was then calculated as a function of $\dot{R}_{thresh}$:

$$\dot{\theta}_{thresh} = \frac{W \dot{R}_{thresh}}{R^2}.$$

Thus the existence of an expansion rate threshold was never tested directly but instead inferred from a calculation based on the unfounded assumption of a relative speed threshold. Furthermore, other visual cues, most notably changes in subtended visual angle and changes in headway (which Hoffman and Mortimer themselves noted), were also present in these film clips, further confounding the issue. Though headway changes can and should be eliminated from an experiment investigating the rate of expansion of visual angle, subtended visual angle will always be present. Its effect can be minimized, however, by reducing the time interval over which the expanding object is shown. [It's

interesting to note in this regard that Hoffman and Mortimer expressed concern that the film clip durations (4 seconds) were too short. The concern could also be the other way around—4 seconds is a rather long stimulus duration for studies aimed at quantifying the visual system's detection capabilities.]

5.5.2 Implications for Criterion Detection

Experiment 2 showed that the criterion level can be shifted at will by the observer. (The experiment produced two different points on the ROC curve, corresponding to two different criterion levels). This is significant because it refutes a fundamental concept upon which threshold theory is based. Threshold theory holds that because of inherent limitations in the visual system, sub-threshold information is lost and inaccessible to the observer. The threshold level is thus fixed and unalterable. Experiment 2, however, showed that the observer can place the criterion level wherever he desires[11]. There is no lower limit, then, to the information the observer has access to. The criterion level is a necessary aid that the observer uses to judge whether neural activity reflects the presence of noise alone or a signal in addition to the noise. The process of setting a criterion level and arriving at a detection decision based on it thus does not alter the neural activity that this decision is based on. Neural activity below the criterion level is not lost, and the observer remains aware of it even though he may have judged that it does not reflect the presence of a signal. Finally, it should be noted that the observer does not make detection decision with complete certainty. He is aware that his choice could be incorrect and will alter the criterion level should he determine that a greater net benefit will ensue (See

[11] To better see this, consider the strategy an observer would employ if offered $1,000 for every hit, and no penalty was extracted for false alarms. Compare this to the strategy employed if a penalty of $1,000 was extracted for each false alarm but no reward given for hits.

Section 9.1). This has important implications for the looming detection model that we propose in the next chapter (Section 6.2.1).

6 SENSITIVITY TO SIZE MODULATION OF A TWO-DIMENSIONAL SHAPE

6.1 Introduction

To better understand and quantify the visual system's dynamic response to looming, we conducted experiments measuring observers' ability to detect sinusoidal modulation of an image's size on the retina of the eye, as would be induced by an object moving towards and away from an observer in three dimensional space. As noted in the Introduction to Part II, this can be satisfactorily represented by the expansion/contraction of a circle presented on a two-dimensional analog CRT.

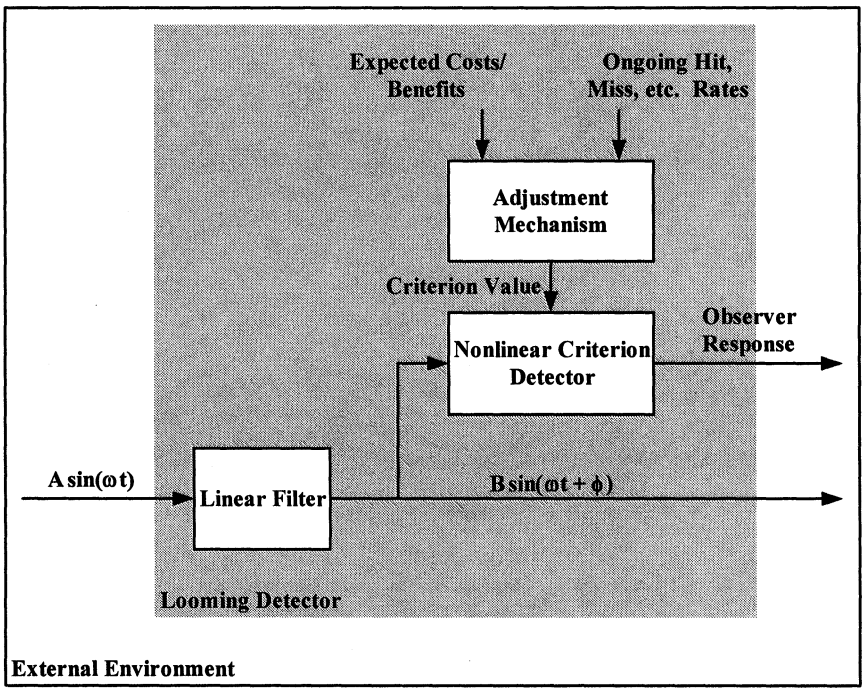
6.2 Background

The experimental approach employed here assumes that the dynamic response we seek to measure is governed by a linear system, which is consistent with the previously cited neurophysiological and psychophysical research (Sections 4.1.1 and 4.1.2). Beyond this, however, we keep our assumptions to a minimum, so as not to impose a predetermined structure on our analytical results. The assumption of a linear element allows us to employ traditional frequency response techniques to ascertain its dynamic response characteristics, but as we cannot measure the output of this element directly, we resort to the psychometric function and a well established adaptive staircase technique to surmount this difficulty. These points are described further in the remainder of this section.

6.2.1 The Looming Detector

We assume that the visual system detects longitudinal motion (looming objects) by means of the "looming detector" shown schematically in Figure 31. It consists of a linear filter, a non-linear "criterion detector" [12], and an adjustment mechanism that sets the

[12] A pulse generator is not included in this model, as it was in DeLange's (Section 4.1.2.1) and Kelly's (Section 4.1.2.2). Assuming, as they did, that the pulse generator output varies instantaneously and



Looming Detection Model
Figure 31

detector's criterion value. The filter, located in the retina of the eye, senses the size modulation of an image on the retina of the eye, converts it to a continuous graded potential, and then sends this signal to the criterion detector and other higher level processing centers of the brain via the optic nerve. The criterion detector reacts instantaneously to this input and signals motion whenever the magnitude of some characteristic of the incoming signal meets or exceeds the criterion value. The adjustment mechanism compares observer responses to correct responses subsequently obtained from the environment and establishes ongoing "hit", "miss", "false alarm", and "correct rejection" rates. It then forms expectations of the overall net benefit associated with these rates, based on internally and externally supplied cost/benefit information. This net benefit is then maximized via adjustment of the criterion level, which in turn

monotonically with its input, and that its performance is approximately linear, it does not alter the filter output and can be excluding without affecting the performance of the model.

shifts the proportion of hits and misses. We typically think of the costs and benefits as coming from the environment, in the form of tangible, performance-based rewards and penalties conferred upon the observer. They can also have observer-based components, however (e.g., the "benefit" of "getting the right answer" versus the "cost" of the effort involved in being especially attentive and focused). It is important to note here that the criterion value is only affected by the benefits and costs associated with hit and miss rates, and not by any parameters associated with the stimulus, most notably the modulation amplitude A or frequency ω. Conversely, the criterion detector does not affect the filter's performance in any way. It only aids the observer in making a judgment regarding the presence of looming The looming detector's dynamic response in this task is therefore that of the linear filter, whose response characteristic can be obtained using classical frequency response techniques.

The particular characteristic that the criterion detector senses is the subject of ongoing research. Three that have been suggested, and that we will consider in this study, are:

1. The peak value of the signal (maximum of the absolute value of the signal, as in the flicker detection model);
2. The peak-to-peak excursion of the signal (|maximum value - minimum value|);
3. The integral of the absolute value of the signal with respect to time.

We assume that the value of the characteristic sensed by the criterion detector increases monotonically with the amplitude of its input, B, which of course increases monotonically with stimulus amplitude A. Thus, increasing A will always make a stimulus more detectable to an observer, while decreasing A will make it less so. If during the course of a detection experiment the benefits and costs of correct and incorrect choices are held constant, the observer's criterion level will quickly reach and then

remain at some constant value. With the criterion level constant the input amplitude A can be adjusted until the filter's output amplitude B is made equal to it. A staircase procedure will be employed to maintain a constant criterion level.

This approach is patterned after DeLange's (36) and Kelly's (37) flicker-fusion experiments (Sections 4.1.2.1 and 4.1.2.2). Here, however, we performed a "Yes-No" experiment using Watson and Pelli's (52) adaptive staircase algorithm to establish, for each of a number of radius modulation frequencies, the modulation amplitude that could be detected with a probability of 79.6% (Sections 5.2 and 6.2.2 below).

We don't know the modulation amplitude of the filter's output. Since it remains constant throughout our experiment we can take it to be unity. (This means that the results we derive will each be known to within a multiplicative constant.) We are also unable to measure $\varphi(\omega)$ directly but we can calculate it based on the assumption that the filter exhibits a minimum phase characteristic (Section 4.1.2.3).

6.2.2 The Psychometric Function

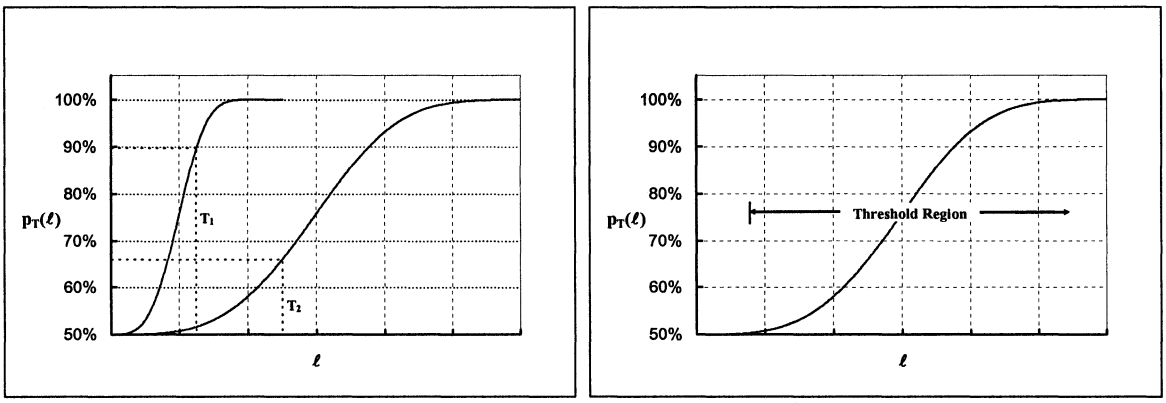
Though threshold theorists incorrectly interpreted the "variable threshold" curve of Figure 25b, it is an experimentally derived result (unlike the absolute threshold curve of Figure 25a). It is more formally known as a *psychometric function*. In general, psychometric functions describe the relation between some physical measure of a stimulus and the probability of a particular psychophysical response. Here the physical measure is the luminance of the light stimulus, ℓ, and the psychophysical response is "detect" or "didn't detect". Two of the more common types of experiments used to obtain psychometric functions are the *Yes-No* experiment and the *Two Alternative, Forced Choice* experiment:

1. A Yes-No experiment consists of a number of individual trials. For every trial a uniform background of constant luminance ℓ_{avg} is displayed to the observer. For approximately half, chosen at random, a light stimulus of luminance ℓ is superimposed on the background. At the end each trial the observer indicates "Yes, I saw the stimulus", or "No, I didn't".
2. A Two Alternative, Forced Choice experiment also consists of a number of trials. Here, though, each trial is divided into two time epochs. In each a uniform background of constant luminance ℓ_{avg} is displayed. For one of the two, again chosen at random, the light stimulus is superimposed on the background. At the end of the trial the observer indicates in which time epoch the stimulus appeared.

In each type of experiment, the stimulus strength for the next trial is generally decreased for correct choices and increased for incorrect ones. *Staircase methods* are commonly employed to control this adjustment process. The experiment proceeds until the stimulus strength becomes sufficiently close to that corresponding to a particular detection rate. The detection rate that is approached depends on the particular form of the staircase method employed, and any factors (such as monetary rewards or penalties) that would influence the observer in setting his criterion level (Section 5.4.1). A functional form based on the Weibull (51) distribution has proven to be an excellent empirical description of psychometric data obtained from such experiments:

$$p_T(\ell) = 1 - \frac{1}{2} e^{-\left(\frac{\ell}{T}\varepsilon\right)^{\beta}} . \qquad (22)$$

Here $p_T(\ell)$ is the probability of a correct detection when a stimulus of strength ℓ is presented. The parameters ε and β depend on the particular type of staircase method employed. The pair $[\ell, p_T(\ell)]$ produced by the experiment then allows determination of T, the *threshold measure*. T sets the "position" of the function along the abscissa, as well as its "spread", as indicated by the plots in Figure 32a. T is therefore a convenient "landmark" for the psychometric function it defines.



a. The Psychometric Function b. Threshold Region

Figure 32

The *adaptive staircase* algorithm developed by Watson and Pelli (52) can quickly and efficiently establish T directly, within any desired confidence limit, by adjusting ℓ in a series of experimental trials until $\ell \approx T$, thus fitting Equ. (22) to the data for that experiment. When this algorithm is employed in a Yes-No experiment, the parameters ε and β become

$$\varepsilon = 1.142\,, \qquad \beta = 3.5\,,$$

and when $\ell = T$,

$$p_T(\ell = T) = .796\,.$$

For a Two Alternative, Forced Choice experiment,

$$\varepsilon = 1.189\,, \qquad \beta = 3.5\,,$$

$$p_T(\ell = T) = .920\,.$$

In view of these considerations we can see that an absolute or even a varying threshold, as proposed by Fechner and others before Tanner and Swets' study, does not exist. The term continues to be used, however, and generally refers to the narrow region (perhaps covering an octave in contrast or energy) in stimulus parameter space wherein

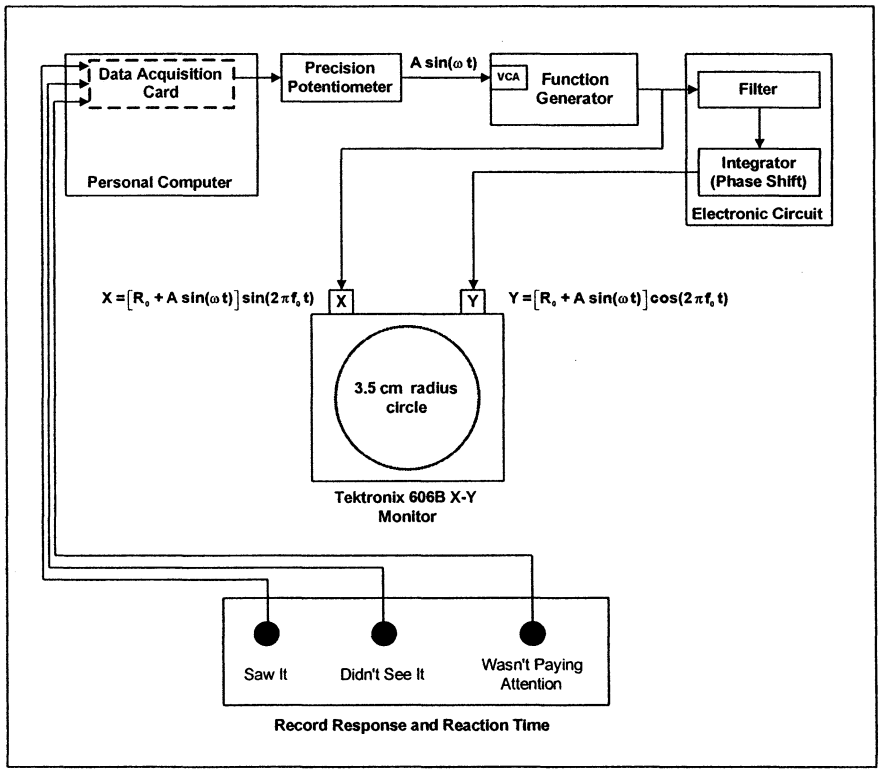
performance varies from close to chance to nearly perfect. This region is indicated in Figure 32b, bounded by [ℓ_{min}, ℓ_{max}]. Each coordinate pair [ℓ, $p_T(\ell)$] within it is associated with a unique criterion value, as set by the observer. The term is also be used to identify a particular point within this region, as in "threshold measure" or "criterion level of response". The latter term is often used to define the predetermined detection rate whose associated stimulus luminance we seek in a staircase experiment. In Watson and Pelli's algorithm, the criterion level of detection equals the threshold measure.

6.3 Experiment 3

6.3.1 Methodology and Procedure

The experimental apparatus is shown schematically in Figure 33. It consists of the following components:

- Standard personal computer;
- 50 kΩ precision potentiometer;
- National Instruments 6024E Data Acquisition Card (NI-DAC);
- Global Specialties PW-2120 Function Generator;
- In-house produced filter/integrator circuit consisting of standard electronic components;
- Tektronix 606B X-Y Monitor;
- In-house produced button box consisting of standard electronic components.



Experimental Apparatus
Figure 33

To understand the operation of the apparatus we first consider the display of an unmodulated circle on the Tektronix monitor. This is accomplished by providing a sine wave and cosine wave of the form

$$V_x = R_0 \sin(2\pi f_0 t),$$

$$V_Y = R_0 \cos(2\pi f_0 t)$$

to its X- and Y-inputs, respectively, where R_0 is the voltage corresponding to a 3.5 cm deflection on the monitor and $f_0 = 36$ kHz. With this input the monitor will display a 3.5 cm radius circle, drawing it at a frequency of 36 kHz. The function generator is manually set to provide V_X directly. V_Y is obtained by splitting V_X and passing one of the resulting signals through the filter/integrator circuit to obtain a 90° phase shifted signal.

To modulate the circle's radius, we generate a second sine wave using the NI-DAC and input this to the "VCA" input of the function generator. This sinusoid takes the form

$$A \sin(\omega t) = A \sin(2\pi f t),$$

where A is the modulation amplitude (volts) and ω and f the modulation frequency (rad/sec and Hz, respectively). The corresponding modulation of the circle's radius is then given by

$$\ell \sin(2\pi f t) \text{ cm.}$$

The NI-DAC can generate a discrete sinusoidal waveform with a maximum amplitude of 10 V, divisible into 2,048 increments. (Thus the minimum incremental voltage change that the NI-DAC can produce is .00488 V.) The value of A corresponding to the threshold measure T will be much less than this, of the order of .01 V. Thus, in order to maintain the full resolution of the NI-DAC for the experiment, its output is passed through a precision potentiometer to attenuate the signal to the desired range. The amount of attenuation, and thus the setting of the potentiometer, is varied for each frequency ω to keep it within an optimal range. The generated waveform consists of

1000 discrete points per cycle. This upstream circuitry thus inputs the amplitude-modulated waveforms

$$V_x = [R_0 + A \sin(2\pi f t)] \sin(2\pi f_0 t),$$
$$V_Y = [R_0 + A \sin(2\pi f t)] \cos(2\pi f_0 t)$$

to the Tektronix monitor. Finally, three of the NI-DAC's digital input ports and its internal clock/timer are utilized to record the test subject's response to a particular trial, as well as his reaction time (the time between the initiation of the stimulus and his response via the push buttons).

If the Tektronix monitor's phosphors do not "turn on" and off quickly enough, the modulating circle will have a "smeared" appearance, thus compromising the findings. To ensure that this is not the case, the dynamic response of the phosphors to a square wave input was measured with a Photo Research Spectra Pritchard photometer (Model 1980A). The phosphors were found to decay to less than 10% of their peak luminance value in .05 msec. For the highest frequency modulating sine wave (20 Hz), the phosphors will have decayed to less than 10% of their peak value by the time each subsequent point on the waveform is presented. This is fast enough to make negligible any effect due to phosphor latency.

Three subjects, all males having (corrected) normal vision, participated. (They are hereafter identified as TS4, TS5, and TS6.) They were seated 110 cm from the display (at this distance the circle subtended a visual angle of approximately 3.6°), and viewed it in a darkened room. Each was instructed to fix his gaze on the center of the circle as it appeared. The circle was shown for about two seconds, and throughout this time its

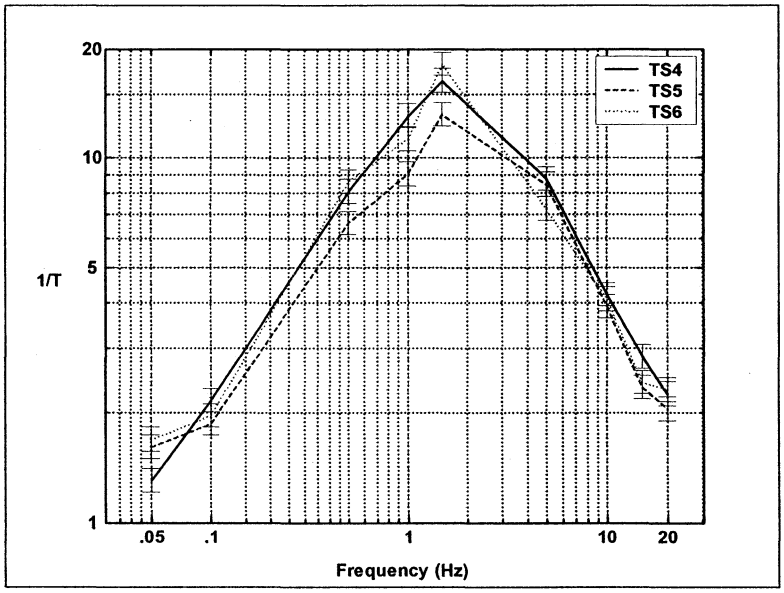
radius either remained constant or varied sinusoidally (each with 50% probability). The test subject then indicated whether the circle's radius oscillated or not. Different pitched tones sounded to indicate correct and incorrect choices. Depending upon his response (correct or incorrect), the modulation amplitude was either decreased, increased, or left unchanged, and another trial was run. The amount by which the amplitude A was varied was determined using Watson and Pelli's (52) algorithm. The experiment continued until the threshold measure T could be estimated to within 7.5%, with 95% confidence. (This typically required approximately 150 trials.) For each multi-trial experiment, the frequency of modulation was kept constant. For each test subject, nine experiments were run for each of nine different frequencies between 0.05 Hz to 20 Hz, as shown in Table 4.

We chose not to offer subjects an explicit monetary reward/penalty for each individual response (e.g., 5¢ for a correct response and -3¢ for an incorrect one), relying instead on each to balance his own internal benefits and costs, as described earlier. It was only important for our purposes that the overall "values" of these costs and benefits be consistent across all three test subjects, and as the data will show, we accomplished this. We did, however, pay test subjects $10 per frequency experiment. An experimenter was present at all times to monitor the test. Data pertaining to each trial was logged into a file for later processing.

6.3.2 Experiment 3: Results

We measured the ability of three observers to detect modulation in the radius of the circle, at each of nine different modulation frequencies from .05 Hz to 20 Hz. For each frequency, we established the threshold measure $\ell = T$ (the amplitude at which the observer could detect modulation at a rate of 79.6%). The overall results of this experiment are shown in Figure 34 and Table 4. As the figure shows, the test subjects all

performed quite similarly to one another, supporting our contention that each had arrived at similar estimates for the costs and benefits involved.



Experiment 3 Results
Figure 34

	Threshold Measure T (cm)					
	TS4		TS5		TS6	
f (Hz)	T	Std Error	T	Std Error	T	Std Error
0.05	0.767	0.029	0.622	0.023	0.593	0.022
0.10	0.464	0.018	0.537	0.020	0.511	0.019
0.50	0.124	0.005	0.151	0.006	0.116	0.004
1.00	0.077	0.003	0.111	0.004	0.089	0.003
1.50	0.061	0.002	0.076	0.003	0.055	0.002
3.25	0.091	0.000	0.101	0.000	0.084	0.003
5.00	0.114	0.004	0.118	0.005	0.138	0.005
10.00	0.237	0.009	0.256	0.010	0.245	0.009
15.00	0.351	0.013	0.431	0.016	0.414	0.015
20.00	0.445	0.017	0.489	0.019	0.433	0.016

Experiment 3 Results
Table 4

6.3.3 Experiment 3: Discussion

Detection performance in this experiment exhibited a bandpass characteristic, just as it did for the flicker-fusion experiments (Sections 4.1.2.1 and 4.1.2.2). In Section 4.1.1 we conjectured that the visual system discards information regarding objects whose luminance changes very slowly over time in favor of information regarding objects whose luminance changes more quickly. It appears that the same strategy is in effect here, and that very slowly moving objects are of less interest than faster moving objects.

6.3.3.1 Integrated Luminance Confound

As the circle radius is modulated, its space-integrated luminance will change in direct proportion to its perimeter. It is therefore possible that the test subjects were detecting integrated luminance change instead of (or in addition to) modulation of the circle's radius, thus confounding our results. Let L(t) be the circle's integrated luminance at time t and K a constant of proportionality. Then

$$L(t) = K\,[2\,\pi\,r(t)]\,,$$

where r is the circle's radius at the instant t. For any modulation frequency ω, the integrated luminance will vary between

$$K \cdot 2 \cdot \pi \cdot [3.5 - T(\omega)] \le L(t) \le K \cdot 2 \cdot \pi \cdot [3.5 + T(\omega)]\,,$$

since the circle's base radius is 3.5 cm and at the threshold measure for that frequency this radius varies by $\pm T(\omega)$. Table 5 shows this variation as a percentage of the unmodulated circle's integrated luminance. Previous research has demonstrated that the visual system is sensitive to variations in this range. To investigate whether or not integrated luminance changes significantly influenced these results we conducted a second set of experiments in which integrated luminance did not vary.

f (Hz)	TS4	TS5	TS6
0.05	±22%	±18%	±17%
0.10	±13%	±15%	±15%
0.50	±4%	±4%	±3%
1.00	±2%	±3%	±3%
1.50	±2%	±2%	±2%
5.00	±3%	±3%	±4%
10.00	±7%	±7%	±7%
15.00	±10%	±12%	±12%
20.00	±13%	±14%	±12%

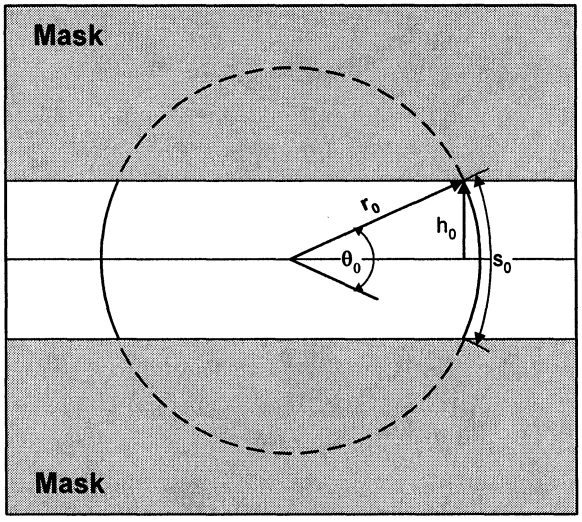
Integrated Luminance Variation (% of Unmodulated Integrated Luminance) at the Threshold Measure

Table 5

6.4 Experiment 4

6.4.1 Methodology and Procedure

The same experimental setup that was used in Experiment 3 was employed here, except that now a mask was placed over the CRT display allowing only the right- and left-most segments of the circle to be displayed. This is shown in Figure 35. For the



Masked CRT

Figure 35

unmodulated circle ($r=r_0=3.5$ cm) the length of each segment s_0 was taken to be approximately 2 cm, for which θ_0 and h_0 (as defined in the figure) were

$$\theta_0 = \frac{s_0}{r_0} = \frac{2}{3.5} \approx 33°, \qquad h_0 = r_0 \sin\left(\frac{\theta_0}{2}\right) \approx 1 \text{ cm}.$$

For any other r, the values of θ and h required to keep $s = s_0$ were thus

$$\theta = \frac{r_0\,\theta_0}{r}, \qquad h = r \sin\left(\frac{r_0\,\theta_0}{2\,r}\right).$$

The profile of the mask boundary to the left of the circle's center was the mirror image of that to the right, and the lower mask profile was the mirror image of the upper mask profile[13]. The length of the segments showing through the mask were thus the same regardless of the size of the circle, and since the luminance of the CRT's beam remained constant across the display and over time (assumed), a stimulus with a constant integrated luminance resulted. The same three subjects that participated in Experiment 3 took part here, and except for the presence of the mask, the identical apparatus and test procedure were employed. Threshold measures for only the four middle frequencies of Experiment 3 (where test subjects showed the greatest sensitivity) were obtained.

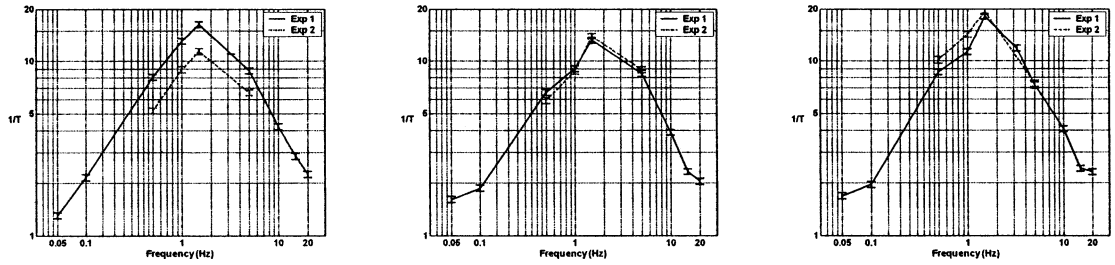
6.4.2 Experiment 4: Results

As in Experiment 3, we measured the ability of three observers to detect modulation in the radius of a stimulus, this time under conditions of constant integrated luminance. At each of four different modulation frequencies (listed in Table 6), we established the threshold measure T (the amplitude at which the observer could detect modulation at a rate of 79.6%.) The overall results of this experiment are shown in Table 6 and plotted in Figure 36 (along with the Experiment 3 results for comparison). From the figures, TS5 and TS6 exhibit nearly identical performance in Experiments 1 and 2, and even though

[13] In actually constructing the mask, it was found that the variations in the height h as a function of circle radius were so small (for the range of radii of interest) that h could be taken to be h_0 without introducing appreciable luminance change.

	Threshold Measure T (cm)					
	TS4		TS5		TS6	
f (Hz)	T	Std Error	T	Std Error	T	Std Error
0.50	0.1921	0.0073	0.1687	0.0064	0.0987	0.0041
1.00	0.1121	0.0043	0.1148	0.0043	0.0700	0.0026
1.50	0.0879	0.0033	0.0718	0.0027	0.0517	0.0021
5.00	0.1518	0.0057	0.1126	0.0043	0.1350	0.0051

Experiment 4 Results
Table 6



TS4 TS5 TS6

Experiment 4 Results
Figure 36

TS4's quantitative performance for Experiment 4 is significantly below that for Experiment 3, his qualitative performance is nearly identical. From this we can conclude that luminance changes provided no information that wasn't already available from the motion of the circle's perimeter, and that the Experiment 3 data reflects the subjects' ability to detect modulation in the circle's radius.

6.4.3 Experiment 4: Discussion

Subject TS4 showed the same bandpass character as in Experiment 3 but lesser modulation sensitivity. The most likely explanation for the difference in TS4's performance is that the time between Experiments 3 and 4 was substantially longer for him than it was for TS5 and TS6. The latter two test subjects performed the experiments concurrently and began Experiment 4 immediately after completing Experiment 3. TS4, on the other hand, performed Experiment 3 before TS5 and TS6 began, and then

performed Experiment 4 after they finished. Thus a period of approximately 4½ months elapsed between TS4's performance of Experiment 3 and Experiment 4. The numerous physiological and psychophysical conditions associated with these tests could easily have shifted in this time, producing the observed variation in quantitative results.

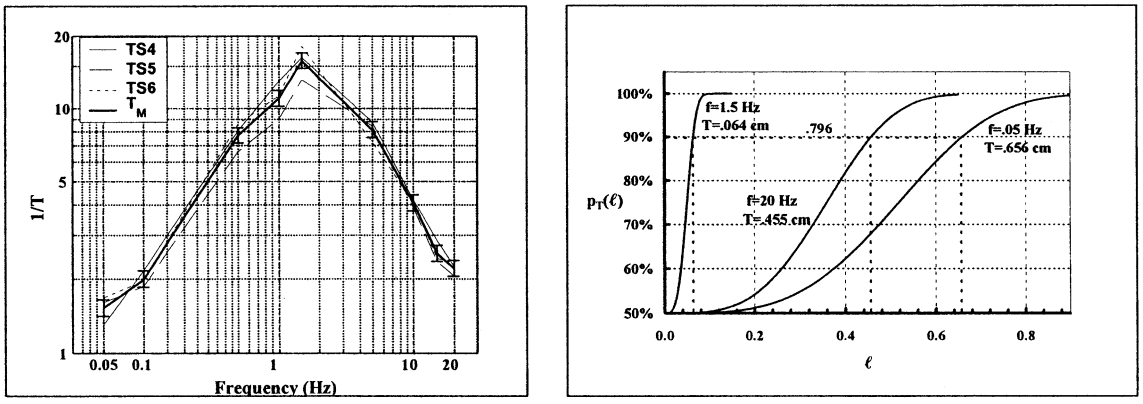
As noted, previous research has established that the visual system is quite sensitive to integrated luminance changes—of magnitudes substantially smaller than those listed in Table 5. Both integrated luminance change and looming, then, were available to the test subjects in Experiment 3, and their performance must have reflected the use of both. Why, then, did performance not degrade when integrated luminance was kept constant in Experiment 4? It may be that the number of neural detectors available for this task was so great that the luminance cue provided no additional information. A second possibility is that the response characteristics of the two detection systems are so similar (as would be the case, for example, if the same visual processes and mechanisms were involved) that again, the information provided by the luminance detection system was redundant. The fact that the two sets of responses both exhibit a band pass characteristic with peak sensitivity occurring at or near the same modulation frequency is consistent with this conjecture. A more detailed investigation would be required to establish whether or not the two detection systems indeed exhibit the same performance characteristics.

6.5 Experiment 3: Analysis

Returning now to Experiment 3, since the performance of all three subjects is so similar, we will take the arithmetic mean T_M of the three as representative. T_M thus represents that value of modulation amplitude whose probability of detection is, according to Equ. (22),

$$p_{T_M}(\ell = T_M) = .796.$$

T_M and the experimentally obtained threshold measures TS4-TS6 from which it is derived are shown in Figure 37a. In going from .05 Hz to 1.5 Hz to 20 Hz, the representative psychometric function moves from the rightmost curve in Figure 37b to the leftmost curve, and then back to the middle curve.



a b

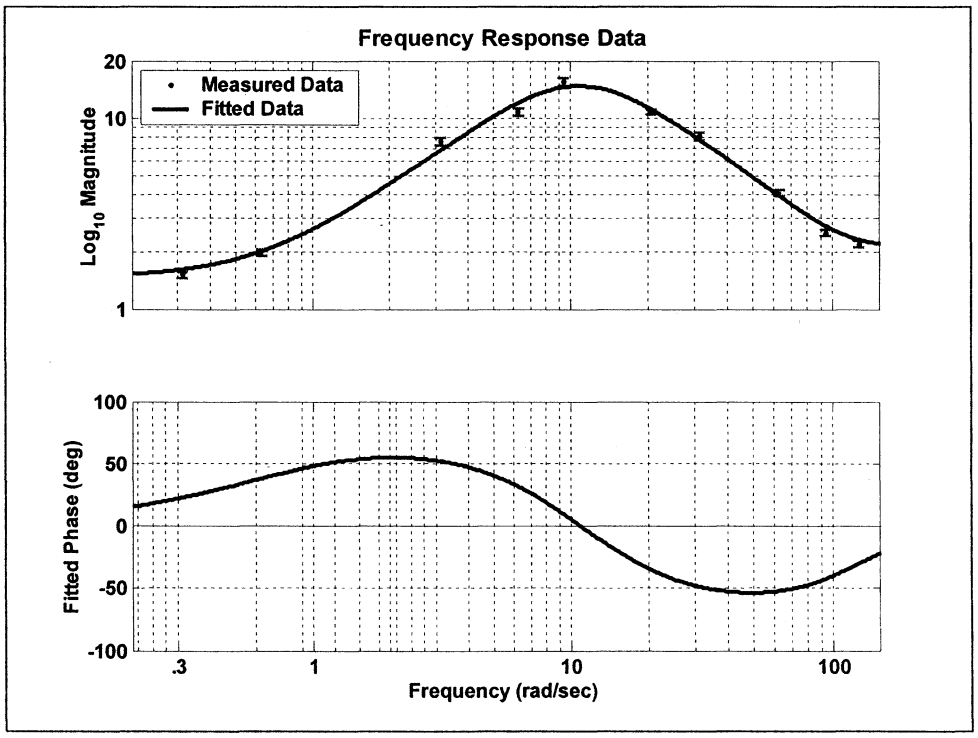
Representative Threshold Measure and Psychometric Functions
Figure 37

The ultimate goal of our study is to quantify the dynamic response characteristics of an observer's visual system to changes in an oncoming object's visual angle, $\theta(t)$. This is accomplished by deriving the visual system's transfer function based on the experimentally derived gain characteristic $1/T_M(\omega)$. The derivation of the system transfer function in general requires both the gain characteristic, which we have measured, and the phase characteristic $\varphi(\omega)$, which we have not. Given that the filter is assumed to exhibit a minimum phase characteristic (Section 4.1.2.3), however, we can use the Matlab function *fitmag* to fit such a transfer function to the gain characteristic. After modifying *fitmag's* result to remove poles and zeros far removed from the origin, we obtain

$$G(s) = 258 \cdot \frac{s + 0.6645}{s^2 + 17.46\,s + 117.5}, \qquad (23)$$

$$= 258 \cdot \frac{s + 0.6645}{[s + (8.728 + 6.427\,i)][s + (8.728 - 6.427\,i)]}.$$

The magnitude and phase characteristic of this transfer function are plotted in Figure 38



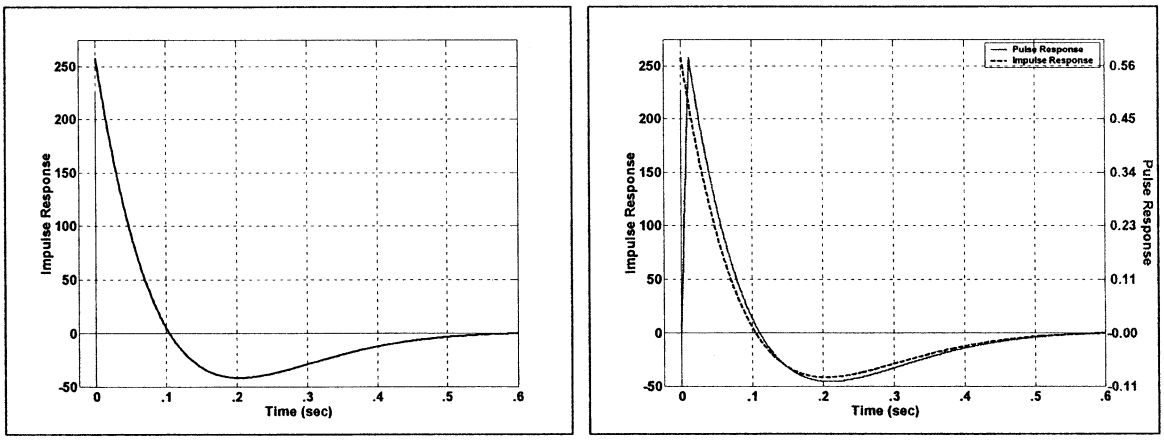
Bode Diagram of Visual System Frequency Response
Figure 38

as a standard bode plot, along with the experimentally derived gain characteristic $1/T_M(\omega)$ for comparison. The agreement between the measured and fitted gain characteristics is seen to be very good.

The impulse response h(t) of the filter is given by

$$h(t) = 258\, e^{-8.728\,t}\left[\cos(6.427\,t) - 1.255\sin(6.427\,t)\right]. \qquad (24)$$

This shows that the filter's frequency of response is 6.427 rad/sec and that its settling time (the time for the response to reach and stay within 2% of its maximum value) is approximately .5 sec. Referring again to Equ. (23), transient response is dictated by the pair of second order poles, which gives the system a damping ratio of .8. In engineering design, there is a general rule of thumb that damping ratios between .4 and .8 provide optimal transient performance in terms of fast response and acceptable overshoot. It's not unreasonable to expect this in a natural setting. From an evolutionary perspective, the ability to make accurate assessments of movement in the environment in the shortest possible time is a critical task for the survival of any organism, and organisms possessing this capability would have an advantage over others that were less capable in this regard. These same considerations were previously offered in support of the assumption of a minimum phase filter (Section 4.1.2.3). h(t) is plotted in Figure 39a. As with the flicker



a b

Looming Detection System Impulse and Pulse Responses
Figure 39

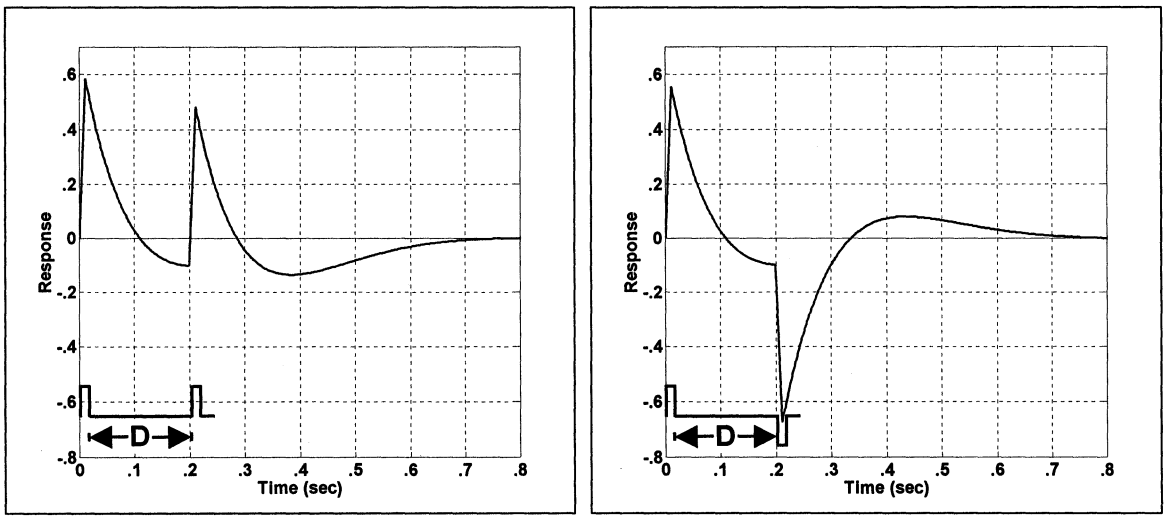
detection models (Sections 4.1.2.1 and 4.1.2.2), the ordinate of this plot reflects relative scaling, and the initial time in the plot coincides with the onset of the response, not the presentation of the stimulus. These results are consistent with those measured in the

flicker detection studies, adding credence to our hypothesis that both detection tasks employ the same or similar basic sensory mechanisms. An impulse is a mathematical artifice that cannot be realized in a real setting. It can, however, be satisfactorily approximated by a *pulse*. This is demonstrated in Figure 39b, which plots the filter's response to a pulsed increase of the circle's radius of .225 cm, lasting .01 sec. The impulse response of Figure 39a is shown again for comparison. Though the ordinal values of the two plots differ, their temporal characteristics are nearly identical.

Figure 39 displays another important characteristic of this filter—its "biphasic" nature (i.e., its excursion to negative values before damping out). This presents an opportunity not only to investigate the limits of our model, but also to better understand how the visual system uses such a signal to detect looming. If the filter is indeed linear, then superposition applies and the filter's response to two successive pulsed expansions (hereafter referred to as a "Plus-Plus", or "PP" pulse) separated in time by D will be

$$r(t) = r_1(t) + r_2(t - D),$$

where r_1 and r_2 are the responses of the two expansions presented individually. If D were chosen equal to the time between r_1's maximum and minimum values, (approximately .2 sec from Figure 39), the result would be as in Figure 40a. Conversely, presentation of a pulsed expansion followed by a *contraction* (a "Plus-Minus", or "PM" pulse) would result in the response shown in Figure 40b. Though we don't know which aspect of the filter's output signal the criterion detector senses, we can easily see that "more signal" is present in the second case. (Recalling from Section 6.2.1 the previously proposed candidates, the peak magnitude, peak to peak magnitude , and absolute integral of the



a) PP Pulse Response b) PM Pulse Response

Filter Response to PP and PM Pulse Stimuli

Figure 40

second signal are all greater than those of the first.) We therefore hypothesize that the second signal will be more detectable than the first. Generalizing from this then, we expect that when the delay is such that the maximum of the second pulse's response coincides with the minimum of the first pulse's response, a PM pulse will be more detectable than a PP pulse. This forms a prediction for our final set of experiments.

7 AN INVESTIGATION OF THE BIPHASIC IMPULSE RESPONSE CHARACTERISTIC

7.1 Experiment 5

If the looming detectors under study actually exhibit a biphasic impulse response characteristic, then for certain time delays (Figure 40a) a PP pulse should be harder to detect than a PM pulse. For this experiment pulse inputs with durations of .01 sec were used and the threshold magnitude of the pulse pairs as a function of the time delay between the first and second pulse was determined using the adaptive staircase algorithm described previously. If our model is valid, then for the delays in the vicinity of that shown in Figure 40, a PP pulse should be more difficult to detect, and hence have a greater threshold magnitude, than a PM pulse.

7.2 Methodology and Procedure

The same experimental setup and test subjects that were employed in the previous experiments were employed here as well. Here, though, a two-alternative, forced choice format was followed, in which a single trial consisted of the presentation of two .5 sec time epochs in sequence. During both epochs a 3.5 cm radius circle was shown, as in the previous experiments. In one of the epochs (selected at random with a 50% probability), the radius was subjected to a pair of pulsed perturbations and during the other it remained constant. The test subject's task was to identify the epoch during which the radius was perturbed. The adaptive staircase algorithm was used here with parameters corresponding to a two-alternative, forced choice experiment (Section 6.2.2).

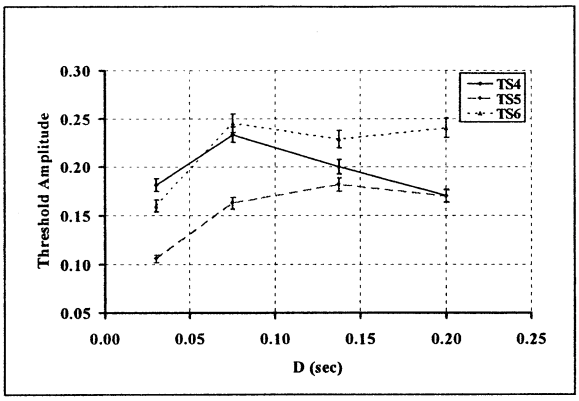
As with Experiments 3 and 4, we relied on each test subject to establish his/her own estimate of the costs and benefits, assuming that these estimates would be consistent across the three. Because the stimulus presented for this experiment differs so from that used in Experiment 3, however, it is quite possible that the expected costs and benefits

associated with its detection will also differ. Further, for the two-alternative, forced choice format the adaptive staircase algorithm will converge to a different hit rate (92.0% as opposed to 79.6%). Referring to our model in Figure 31, we see that the corresponding inputs to the adjustment mechanism will change, leading us to conclude that the criterion level established for this experiment will differ from that of Experiment 3. This is not of concern, however. As with Experiment 3 we only want the criterion levels to remain consistent between test subjects and constant throughout the course of Experiment 5.

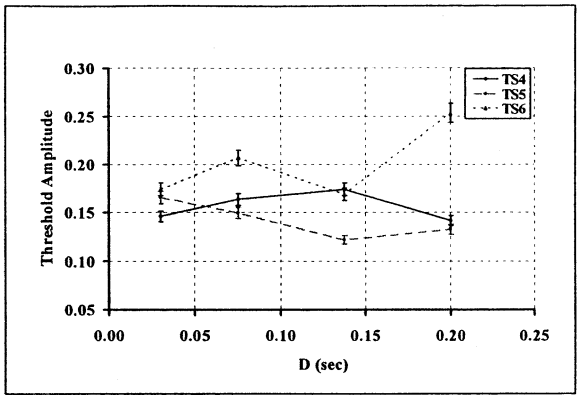
A single experiment consisted of a number of trials, each initiated by the observer, in which PP and PM pulse pairs were randomly mixed with one another and continued until threshold magnitudes for each pair were arrived at (satisfying the same criteria as in Experiment 3). The interleaving of PP and PM pulses in this manner ensured that the same criterion level was in effect for each, at each time delay. The adaptive staircase algorithm employed previously was modified to deal with each case separately. For a given experiment, the time delay between pulses was kept constant for all trials. Four different experiments were run, corresponding to time delays of .0300 sec, .0750 sec, .1375 sec, and .2000 sec. Test subjects were paid $10 per experiment.

7.3 Results

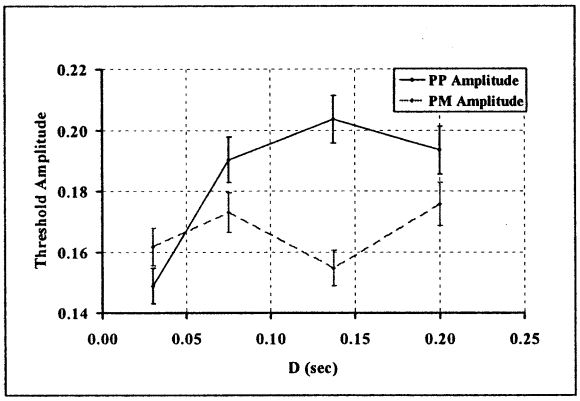
We measured the ability of three observers to detect pairs of pulsed perturbations in the radius of the circle, at each of four different time delays from .0300 sec to .2000 sec. For each time delay we established the threshold pulse magnitude for each type of perturbation (PP or PM). The overall results of this experiment are shown in Figure 41 and Table 7. Figure 41a and Figure 41b show the performance of the individual test



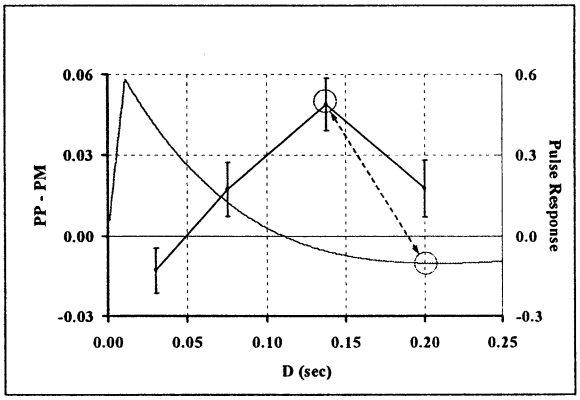
a) Individual Results:
PP Threshold Amplitudes



b) Individual Results: PM
Threshold Amplitudes



c) Average PP and PM
Threshold Amplitudes



d) PP – PM

Experiment 5 Results
Figure 41

subjects in the detection task, Figure 41c averages their performance, and the difference PP_{avg} - PM_{avg} between these averages is shown in Figure 41d. Also shown in Figure 41d is the pulse response of Figure 39. Positive differences in PP-PM mean that at that value of D the PM change was easier to detect than the PP change. According to the model (as represented by the pulse response), the difference should start out negative, increase to some maximum positive value at the minimum of the pulse response, and then decrease again. Qualitatively, this is what we get, so this result supports the hypothesis of a linear filter having a biphasic impulse response. The data suggest, however that the model's

Test Subject	D (sec)	PP Response		PM Response		PP - PM	
		Amplitude	Std Error	Amplitude	Std Error	Amplitude	Std Error
TS4	0.0300	0.1813	0.0069	0.1461	0.0055	0.0352	0.0088
	0.0750	0.2336	0.0080	0.1632	0.0059	0.0704	0.0100
	0.1375	0.2002	0.0076	0.1740	0.0066	0.0262	0.0101
	0.2000	0.1704	0.0064	0.1416	0.0053	0.0288	0.0083
TS5	0.0300	0.1058	0.0039	0.1653	0.0063	-0.0595	0.0074
	0.0750	0.1623	0.0062	0.1493	0.0056	0.0131	0.0084
	0.1375	0.1818	0.0069	0.1217	0.0046	0.0601	0.0083
	0.2000	0.1695	0.0064	0.1326	0.0051	0.0369	0.0082
TS6	0.0300	0.1596	0.0060	0.1740	0.0066	-0.0144	0.0089
	0.0750	0.0750	0.0097	0.2070	0.0081	0.0388	0.0126
	0.1375	0.2291	0.0087	0.1686	0.0064	0.0605	0.0108
	0.2000	0.2408	0.0102	0.2535	0.0099	-0.0127	0.0142
Average	0.0300	0.1489	0.0057	0.1618	0.0062	-0.0129	0.0084
	0.0750	0.2139	0.0081	0.1731	0.0066	0.0408	0.0105
	0.1375	0.2037	0.0078	0.1548	0.0060	0.0489	0.0098
	0.2000	0.1936	0.0079	0.1759	0.0071	0.0177	0.0106

Experiment 5 Results
Table 7

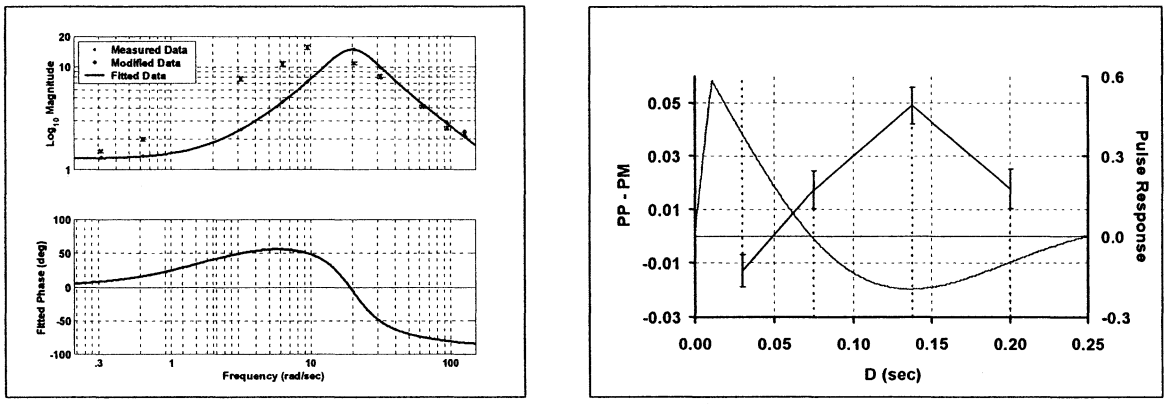
temporal response is too slow. To be completely consistent with Experiment 5's results, we would expect the pulse response to be such that it achieves its minimum not at .2 sec but where the difference PP - PM reaches its maximum. According to Figure 41d, this maximum point lies somewhere in between the second (.0750 sec) and fourth (.200 sec.) data points, say between .1 and .18 sec.[14] We will investigate this further in the next section.

7.4 Analysis

If the filter transfer function were modified to exhibit a faster temporal response, its performance should more closely match the measured data. To investigate this we return to the frequency response characteristic obtained in Experiment 3 (Figure 38). A faster

[14] Since only one data point (.1375 sec.) lies within this range it is not clear where the actual maximum lies. Additional data would have to be taken to find out.

temporal response implies that the magnitude response for the low frequencies should be lower relative to that for the higher frequencies. Figure 42a shows a frequency response characteristic in which this modification has been affected, to yield a filter model with the temporal response of Figure 42b. Note that its minimum approximately coincides with



a) Modified Frequency Response Characteristic

b) Pulse Response

Figure 42

the maximum of the measured amplitude difference PP - PM of Experiment 5 (Figure 41d). The transfer function for this modified filter is given by

$$G_{mod}(s) = 258 \cdot \frac{s + 1.942}{s^2 + 17.46\, s + 393.0}, \qquad = 258 \cdot \frac{s + 1.942}{[s + (8.728 + 17.80\, i)][s + (8.728 - 17.80\, i)]}. \tag{25}$$

and its impulse response by

$$h_{mod}(t) = 258\, e^{-8.728\, t}\, [\cos(17.8\, t) - .3812 \sin(17.8\, t)]. \tag{26}$$

It's interesting to note here that this modified system exhibits a damping ratio of .4. This is at the lower end of the "optimum" range cited in Section 6.5. The original system exhibited a damping ratio of .8, which was at the upper end, so both models are within

this optimal range. Based upon this analysis, the implication is that either something about Experiment 3 led to a consistent underestimation of low frequency sensitivity, or something about Experiment 4 led to a decrease in low frequency sensitivity. Assuming the model is correct, one or more of the following may account for the discrepancy:

1. First, the maximum point PP-PM is not necessarily where it's indicated here. It could be anywhere between the second and fourth data points.
2. In line with this, there is of course measurement uncertainty associated with both sets of data that may be combining to contribute to this discrepancy.
3. Thirdly, human performance is inherently variable to some degree, and could have shifted in approximately four months that passed between the two experiments. It is therefore possible that the modified frequency response characteristic of Figure 42 was in effect for Experiment 5, instead of the one originally obtained in Experiment 3. (This is the same reasoning that we applied in explaining TS4's performance difference between Experiments 3 and 4.)
4. Finally, the two experiments aren't identical in nature, and there could be subtle differences that could have caused observer performance to change.

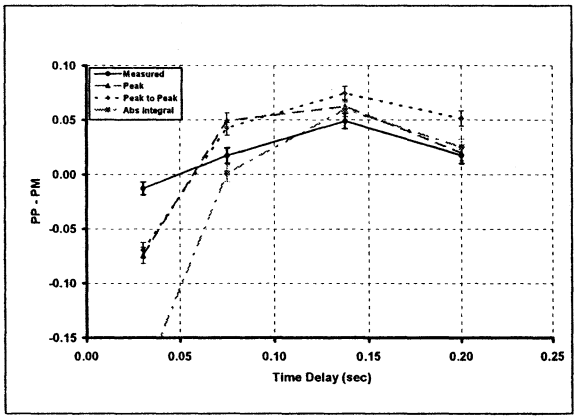
But it could also be that the existing model is deficient, and that some aspect of the detection mechanism has not been captured adequately. It may be, for example, that the visual system is able to tune the filter's parameters to better match it to the task at hand. In Section 3.3 we noted that for temporally varying luminance stimuli, a variable gain control mechanism shifts the filter's temporal characteristics with background light level. Further investigation will be required to see if a similar mechanism is at work in the case of looming.

To assess the predictive power of our model, we will use it to compute the PM pulse amplitude for each time delay, based upon the measured PP pulse amplitude and each of the three proposed criterion schemes. The differences $PP_{meas} - PM_i$ (where the subscript, i, refers to one of the three criterion schemes) will then be compared to the measured

difference shown in Figure 41d. The algorithm that we employ for this purpose proceeds as follows:

1. For each time delay compute the filter's response to a PP pulse pair, where each pulse has a duration of .01 sec and an amplitude equal to the measured threshold amplitude for that time delay.
2. Compute for this response the proposed criterion values:
   - The maximum of its absolute value;
   - Its peak-to-peak excursion (absolute value of the maximum plus the absolute value of the minimum);
   - The integral of its absolute value with respect to time.
3. Determine the amplitude of a PM pulse that would produce criterion values equal to those calculated for the PP pulse in step 2.
4. Form the differences $PP_{meas} - PM_i$ for each criterion and compare to the measured differences.

The result is shown in Figure 43 and Table 8. All three criterion schemes capture the trend of the measured data, the Peak and Peak-to-Peak particularly so: their correlation



Time Delay	Measured PP-PM	Computed PP - PM		
		Peak	Peak to Peak	Absolute Integral
0.030	-0.0129	-0.0743	-0.0697	-0.1883
0.075	0.0173	0.0494	0.0427	0.0013
0.138	0.0489	0.0623	0.0744	0.0597
0.200	0.0177	0.0199	0.0514	0.0252
RMS Error		0.0355	0.0502	0.0884
Correlation Coeff.		0.8980	0.9100	0.9068

Comparison of Modified Model to Experiment 5 Data

Figure 43 Table 8

coefficients are approximately 90%. These results support the validity of our proposed model (Figure 31). They also point to a very simple criterion scheme (Peak or Peak-to-Peak) for looming detection. It should be noted, though, that all of the criterion schemes considered here are correlated with one another. Further fine-tuning of the model can be

expected to yield even closer agreement with the measured data, but will not be undertaken as it would offer little additional insight into the underlying process.

7.5 Discussion

In this experiment we investigated the biphasic impulse response characteristic. The results are consistent with the model after modifying the frequency response characteristic, and showed that the model fairly accurately predicts the Experiment 3 results via either a peak or peak-to-peak criterion detection scheme. Recalling our argument for a minimum phase filter (Section 4.1.2.3), since organisms that can detect looming more quickly would be advantaged compared to those who detect looming less quickly, it is reasonable to expect this result. The peak detection schemes should be faster than the integration scheme for two reasons: First, peak or peak-to-peak amplitudes occur fairly early in the filter's output signal (certainly for sudden size changes of the type studied here), so there would be little delay in waiting for the requisite information to present itself. Alternatively, a scheme involving an integration would involve a relatively more substantial delay in that the integration could not be completed until some significant portion of the filter's output signal was presented. Secondly the "computation" of a peak amplitude is faster than that of an integral. Such an argument does not constitute proof that a peak detection scheme of some sort is actually employed, of course, but it does offer guidance in determining the most efficacious lines of inquiry to pursue next.

Further investigation will have to be performed to see if such a modification is warranted, or if further elaboration of the model is instead necessary. Variability due to the time that passed between the experiments 3 and 5 could be eliminated by conducting a two alternative forced choice experiment in which both the Experiment 3 and

Experiment 5 stimuli were mixed together, with costs and benefits made explicit by providing monetary rewards and penalties for the outcome of each experimental trial. Because of the large number of trials that would be required, care would have to be taken in designing the experiment to ensure test subjects did not become overly fatigued. Data from trials presenting the Experiment 3 stimuli would be used, as before, to construct a model of the filter, which would then be used to predict PM pulse amplitudes based on the corresponding PP pulse amplitudes, as was done in Experiment 5. If the predictions fit the experimental data significantly better than in the analysis performed here, it would tend to confirm the hypothesis that time-induced performance variations were at least partly responsible in causing the two sets of data to differ.

PART III: CONCLUSIONS AND AVENUES FOR FUTURE RESEARCH

8 CONCLUSIONS

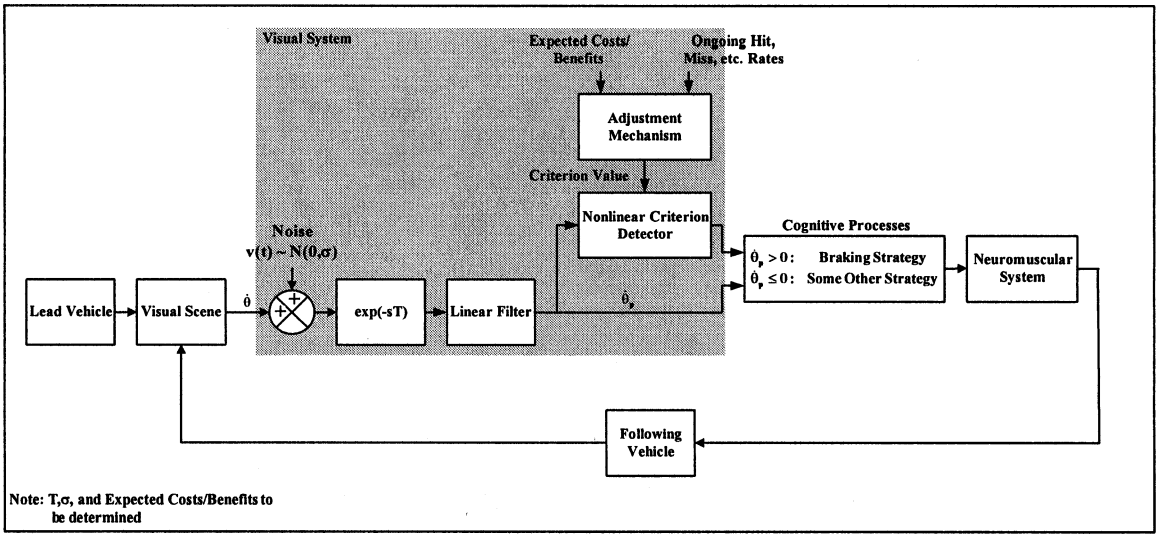
8.1 Contribution

In this study we conducted an experimental program to develop a model of the visual system as it relates to the collision detection/avoidance task (i.e., detection of and response to changes in visual angle θ.) The model, consisting of a linear filter, a nonlinear criterion detector, and an adjustment mechanism (see Figure 31), exhibits the following characteristics:

1. The linear filter exhibits a band-pass frequency response characteristic (Figure 34) which is well described by a minimum phase, second order transfer function. This in turn gives rise to a damped, second order response characteristic (Figure 39).
2. The criterion detector monitors the output of the filter and signals the onset of looming ($\dot{\theta} > 0$) when some aspect of the filter's output reaches some *criterion value*. Our experiments indicate that it senses either the absolute magnitude of the filter output or its peak-to-peak excursion.
3. Detection of the onset of looming is based on a signal/noise paradigm. This is in contrast to the widely held assumption that a threshold is associated with the detection task. This carries two significant implications for the detection/monitoring task:
   - The criterion value is not fixed but instead is a function of the expected costs and benefits associated with the detection task. It can be set by the adjustment mechanism to any value necessary to maximize the task's expected net benefit.
   - The filter's output is not altered in any way by the criterion detector. Information loss thus does not occur and the filter's output is available in its entirety for other visual tasks, whether the criterion detector senses the onset of looming or not.

8.2 Towards a Complete Model of Car-Following

Figure 44 shows how the results of these experiments can be used to improve upon the basic human driver model that was originally presented in Figure 8. Here the original visual system block has been expanded to incorporate noise, a transmission delay, a linear filter, and a nonlinear criterion detector. A quantification of the noise, transmission delay, and the expected costs and benefits associated with this detection task are the subject of a future study (See Section 9.1). The basic operation of the model is similar to

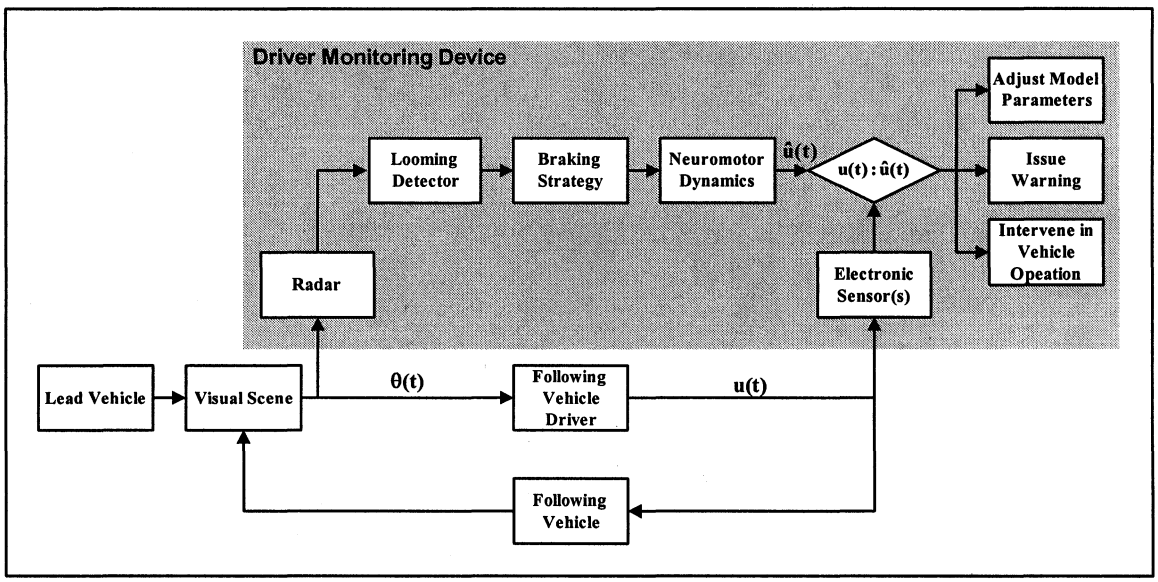


A More Complete Model of Car Following
Figure 44

the original. The Following and Lead Vehicle dynamics interact to produce the actual visual scene before the FV driver. Here the relevant variable is $\theta(t)$, the angle that the LV subtends in the FV driver's field of view. If $\dot{\theta} > 0$ the FV is approaching the LV and the driver implements a breaking strategy to avoid a rear end collision. A number of strategies have been proposed, perhaps the most widely known is the "tau-dot" strategy developed by Lee (53). Based upon this the driver issues commands to his body to implement the strategy and the loop repeats.

8.3 Other Uses

This model could be used not only for design and evaluation, but also as an onboard driver monitoring device, as shown in Figure 45. The model, shown in the dashed box, takes in $\theta(t)$ via a radar and predicts the driver's response, $\hat{u}(t)$. This is then compared with the driver's actual response, u(t), which is determined by monitoring the movement of the accelerator and brake pedals. The system could then perform a number of functions, depending on the "mode" it is in:



Monitoring Driver Behavior
Figure 45

1. Adjust the model's parameters in an adaptive fashion to better match the model to a particular driver's behavior.
2. If the driver's predicted response differs markedly from his actual response, it could issue a warning.
3. In extreme cases it could intervene in the vehicle's operation.

In the introduction to this study we noted that these results are not limited to the Lead Vehicle Braking Scenario. For example, this model could be used to better understand how drivers perform in continuously following another vehicle at an approximately constant range. Given the low frequency cutoff we observed in frequency response characteristic, for instance, the model should predict that drivers would be relatively insensitive to slowly gaining on or falling behind the vehicle in front, at least as far as the looming detector is concerned. It could also be applied to other perceptual problems in driving, such as judging the approach of oncoming trains at crossings, and detecting gaps in a line of vehicles that a driver seeks to enter.

9 AVENUES FOR FUTURE RESEARCH

This study constitutes an initial step towards a complete quantification of the visual system's ability to detect and respond to looming objects, as would be required for collision avoidance applications. We proposed a model for looming detection consisting of a minimum phase linear filter, a nonlinear criterion detector, and a nonlinear adjustment mechanism. Further research into the function of each is warranted.

9.1 Linear Filter

Noise

Previously (Sections 3.2 and 3.3) we noted that quantum and physiological noise are present in any detection task. It is possible to estimate the magnitude of the (assumed Gaussian) noise associated with the looming detection task with Pelli and Farrell's (54) technique. Here we would introduce enough external noise (by superimposing Gaussian fluctuations on the circle's radius) until the threshold measure (Section 6.2.2) at each modulation frequency is halved. The magnitude of the external noise at this point is assumed equal to that of the internal noise, in the sense that all noise is referred to the input of the linear filter.

Transmission Delay

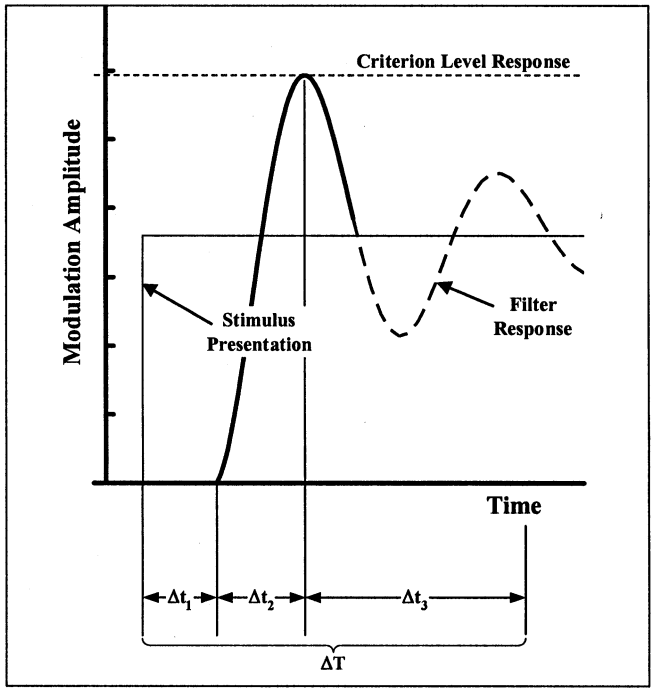
We also noted the presence of a transmission delay between the stimulus presentation (the instant at which the circle begins to modulate) and the onset of the filter's response (Section 6.5). This too can be estimated. The delay that can most conveniently be measured is that between the stimulus presentation and depression of the "detect" button by the test subject. This could be obtained with a modified version of the size modulation experiment, in which:

1. A monetary reward or penalty is conferred based on the outcome of each trial, each of which becomes more negative with time. This will encourage test subjects to reach a decision and then respond as quickly as they are able to, ensuring consistency of response from trial to trial.

2. The time between the presentation of the stimulus and the depression of the "detect" button is measured.

A representative, overall time delay ΔT can then be computed from these results. This overall delay consists of three components, as shown in Figure 46, so that

$$\Delta T = \Delta t_1 + \Delta t_2 + \Delta t_3 \,. \tag{27}$$



Time Delays Associated with the Looming Detection Task (not drawn to scale)
Figure 46

The first, Δt_1, is the delay we seek. Assuming a simple peak criterion scheme (Section 6.2.1), Δt_2 is the time from the onset of the filter response to the first peak in that response. This is known from the existing model and experimental results. Finally, assuming that the perception of motion coincides with the neural command to depress the "detect" button in the experiment, Δt_3 corresponds to the neuromuscular delay in actually carrying out the command. This last delay could be estimated with reference to physiologically-based studies that have already been conducted on the subject. [See, for

example, Vickers (55) and McRuer, et. al. (56).] Alternatively, these studies could be used as guides in developing an experimental procedure to obtain data specific to the conditions of the present study's detection experiments. Δt_1 is thus the only unknown in Equ. (27), from which it can then be obtained.

Expected Costs/Benefits for the Detection Task

The criterion C, as noted previously, is set by the FV driver as a consequence of the tradeoff between the benefits of correct assessments of $\dot{\theta}$ and the costs of incorrect ones. If B_{Hit}, B_{Miss}, B_{FA}, and B_{CR}, are the benefits/costs $\left(B_i > 0 \Rightarrow \text{benefit}, B_i < 0 \Rightarrow \text{cost}\right)$ and P_{Hit}, P_{Miss}, P_{FA}, and P_{CR} the probabilities of the respective outcomes (See Figure 28b), then the expected net benefit realized from this detection task is

$$B_{net} = \left[B_{Hit} \cdot P_{Hit}(C)\right] + \left[B_{Miss} \cdot P_{Miss}(C)\right] + \left[B_{FA} \cdot P_{FA}(C)\right] + \left[B_{CR} \cdot P_{CR}(C)\right].$$

In the car following scenario there are no material benefits conferred upon the FV driver for hits and correct rejections ($B_{Hit}=B_{CR}=0$), but he does realize penalties for misses (potentially an REC) and false alarms (unnecessary braking). These penalties will be functions of LV and FV speed (V_{FV} and V_{LV}), R, and $\dot{R}$, some of which may themselves be dependent upon C. B_{net} thus takes the form

$$B_{net} = \left[B_{Miss}(V_{LV}, V_{FV}, R, \dot{R}, C) \cdot P_{Miss}(C)\right] + \left[B_{FA}(V_{LV}, V_{FV}, R, \dot{R}, C) \cdot P_{FA}(C)\right]. \quad (28)$$

The inputs to the model, then, are the quantification of these benefits/costs. The criterion value C is the value that maximizes Equ. (28).

Eventually, all occurrences of LV braking will be detected by the FV driver, the question is when. If the conditions of a particular car following scenario are such as to make $P_{Hit}(C)$ relatively low, we would expect that the detection of LV braking would be

delayed. This would of course affect the resulting FV braking profile (less time and space to stop in), and in extreme cases leave insufficient time to avoid an REC. It is interesting too to see how the fundamental premise for our experiments (observers receive some information on angular expansion, even though it may not be sufficient for them to conclude such an expansion is actually taking place) factors into this. During the time between the actual onset of LV braking and the FV driver's detection of it, he is nonetheless receiving some information to this effect from the scene before him. Thus he is acquiring information that B_{Miss} is becoming more negative, which should in turn cause him to shift C to the left in Figure 28b, reducing P_{Miss} (but at the expense of increasing P_{FA}). This would make it more likely that he would detect the LV's braking in the next instant. Finally, the model also provides insight into the ways in which human error can enter into the problem. The FV driver can err in estimating the probabilities, costs, and benefits associated with the various outcomes. He can also err in identifying the criterion level that maximizes the net benefit of the detection task.

Eccentricity

In Section 3.3 we observed that as eccentricity increases, photoreceptor density decreases[15], RGC receptive field size increases, and the number of RGC receptive fields decreases. Sensitivity to any type of stimulus, including looming stimuli, should therefore decrease with eccentricity. A second line of inquiry would be to quantify the effects of eccentricity on sensitivity to size modulation, and this could be accomplished by repeating the size modulation experiment but with different circle radii.

[15] From Figure 13, cone density fall dramatically as soon as we leave the foveal region (approximately 2° eccentricity) while rod density falls after reaching its maximum at approximately 17° eccentricity.

Stimulus Contrast and Background Light Level

The size modulation experiment was carried out using a single stimulus contrast at a single adaptation level. The flicker-fusion studies (Sections 4.1.2.1 and 4.1.2.2) showed that detection performance varied with adaptation level. (Stimulus contrast was the dependent variable that was measured in those experiments.) Similarly, detection performance in the size modulation experiment should vary, both with background light level and with stimulus contrast at a given background level. A complete quantification of the visual system's ability to detect and respond to size modulation characterized by first order motion stimuli would show how all the performance measures identified here—threshold measure, time delay, and quantum/physiological noise—vary with modulation frequency, eccentricity, background light level, and stimulus contrast.

Looming based on first order motion stimuli represents one of a number of independent *cues* signaling the approach of an object in three dimensional space. Other cues are also available, and the visual system utilizes these as well in the looming detection/response task. The set of experiments described in this section constitutes a template that can be used to quantify the visual system's detection/response capabilities to looming based these other cues. Two such cues, looming stimuli based on "second order" motion stimuli and looming based on binocular disparity, and the manner in which different independent cues can be pooled together are described in the following sections.

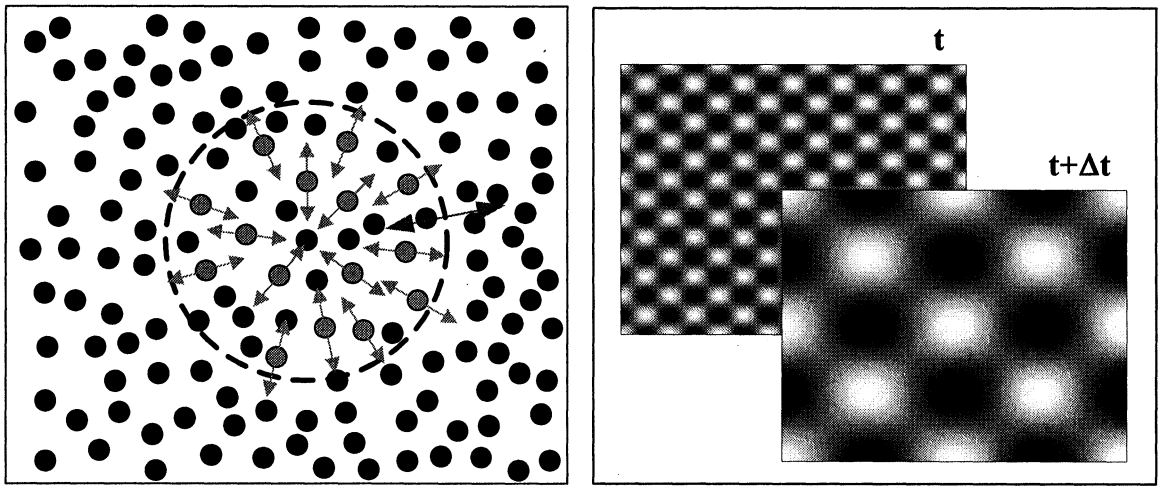
Second Order Motion Cues

Size modulation can be detected with second order motion cues when no first order motion cues are present. This would involve images with no discernible luminance boundaries, but instead composed of texture patterns. In the actual environment object surfaces do not exhibit uniform luminance but instead contain variations in luminance

intensity which form certain repeated patterns called *texture*. These patterns can be the result of physical surface properties such as roughness or oriented strands which often have a tactile quality, or they can be the result of reflectance differences such as the color on the surface. Exactly how the visual system perceives second order motion is the subject of ongoing research, but there is strong evidence that the detection of second order motion stimuli is carried out by separate neural detection mechanisms, independently from that of first order motion stimuli. (57, 58). Since different mechanisms are involved, we can expect somewhat different performance in the detection task. Size modulation of such an image could be effected in either of the following ways (59):

1. A circular image consisting of a pattern of random dots is presented on a CRT against a background of identical, static, random dots. The dots comprising the image then expand outward/contracts inward against the static background, simulating longitudinal motion. Note that luminance does not change as a result of the "motion", nor are there any sharp boundaries present to distinguish the circular image from its background. This constitutes a *divergence of the optic flow field.*
2. The image consists of a random pattern of elements, whose size expands and contracts to simulate longitudinal motion. The edges of the elements, as well as the boundaries of the image, are blurred to eliminate the presence of localized lines or edges. Care must be taken in this case to ensure that the overall luminance remains constant. (Every expanding/contracting light region in the image is matched with a nearby expanding/contracting dark region.) This is referred to as *dilation of scale.*

Examples of each are shown in Figure 47. (The different colors employed in Figure 47a are not part of the actual stimulus.) A quantification of the visual system's detection/response capabilities to looming based on second order motion stimuli can be carried out using the previously described experiments, using these two stimuli (separately and in combination) in place of the first order motion stimuli.



a) Divergence of Optic Flow Field b) Dilation of Scale

Second Order Motion Stimuli Applied to the Detection of Size Modulation

Figure 47

Binocular Disparity Cue

When an object in three dimensional space is viewed by an observer it casts an image onto different locations within the retina of each eye. This *disparity* serves as a cue with which to infer the object's distance from the observer. Far objects project to more nasal positions on the retina than near objects. (At about 10 meters, however, the light rays emanating from the object become nearly parallel, and the disparity cue ineffective beyond this point.)

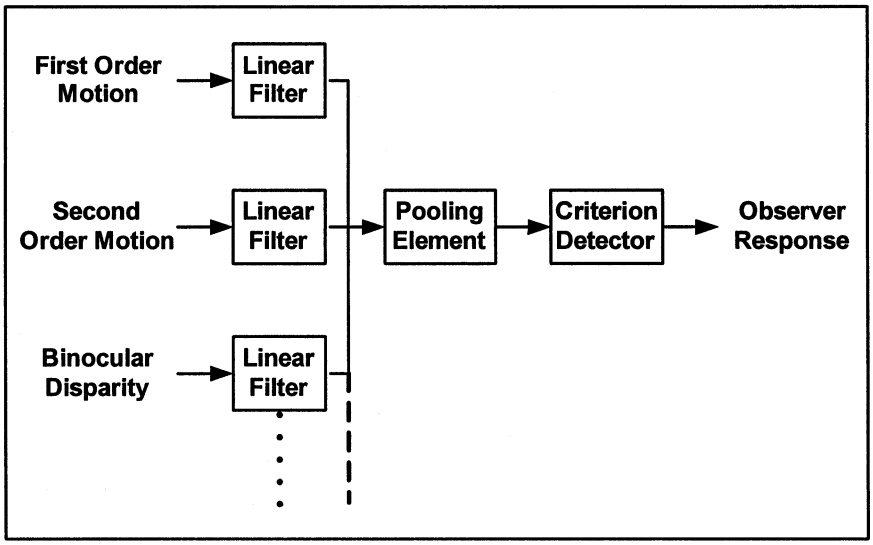
The effects of binocular disparity can be ascertained by repeating the first and second order motion experiments in a simulated three dimensional environment. This can be accomplished either with a haploscope [16] or with special goggles that deliver separate images to each eye. A three dimensional effect is achieved by first delivering to each eye only that portion of the stimulus' image that would be visible to it. These images are then placed on a slightly different region of each eyes' retina, consistent with the stimulus'

[16] A haploscope is an instrument for presenting separate fields of view to the two eyes so that they may be seen as one, continuous, superimposed, integrated, or fused field, and hence useful for measuring or stimulating various binocular functions. Many specially designed experimental and clinical models provide for elaborate controls of the accommodation, convergence, and fusion stimuli, the color, brightness, and size of target and field, and stereo-producing disparity.

"distance" from the observer. In this way a very realistic simulation of a disk moving towards and away from the observer in three dimensional space can be achieved. Using such equipment all of the previously described first and second order monocular motion experiments could be repeated, for each eye separately and then for both eyes together. These results could be compared and combined with the monocular test results. They could also be used separately to formulate a model of, and gain insight into, the manner in which information from each eye is combined for this visual task.

Combination of Cues

In this section we pointed out that a number of different cues, each detected by a different mechanism in the visual system, are associated with looming. We identified three: monocular image expansion (the subject of this study), second order motion, and binocular disparity. Others cues may also be present. The general looming detection task, then, will involve detections of these individual cues (some of which may signal looming and others not) followed by a pooling of the results. An expanded looming detection model that incorporates the pooling of different cues is shown in Figure 48. (The adjustment mechanisms are omitted for clarity.) The overall pooling element and criterion detector are here shown as separate mechanisms, but it is also possible that both functions are performed by the same neural mechanism. Green and Swets (50) show how such cues can be combined to determine the likelihood that looming has occurred. Other cues can be incorporated into the model in the same fashion.



Expanded Looming Detection Model (Adjustment Mechanism not shown)
Figure 48

9.2 Nonlinear Criterion Detector

Further Test of Linearity

We have assumed that the criterion detector acts on the output of the linear filter without altering it, and that no other significant nonlinearities exist between the filter and the criterion detector. It is not possible to test this assumption directly (in humans, at least) since we cannot directly measure the neural signal between the filter and the criterion detector. An indirect test can be performed, however. A square wave of the form

$$S(f) = \begin{cases} 1 & 0 < 2\pi f\, t < \pi \\ -1 & -\pi < 2\pi f\, t < 0 \end{cases}$$

has the Fourier series representation

$$S(f) = \frac{4}{\pi}\sin[2\pi f\, t] + \frac{4}{3\pi}\sin[2\pi (3f)\, t] + \frac{4}{5\pi}\sin[2\pi (5f)\, t] + \cdots$$

Referring to Figure 37a, the higher harmonics of square wave inputs having frequencies above approximately three Hz will be attenuated to insignificant levels, leaving only the

first harmonic. If the system is linear, then performing the modulation sensitivity experiment with square wave inputs at modulation frequencies greater than three Hz should yield threshold measures $4/\pi$ times greater than those for purely sinusoidal inputs.

Criterion Detection Scheme

In introducing the psychophysical looming detection model (Section 6.2.1) we noted that researchers have not yet identified the particular signal characteristic that the criterion detector bases its detection decision on. Based on our analysis of the Biphasic Impulse Response Experiment (Section 7.4), we showed that a peak detection scheme was better correlated with the experimental results than an integration scheme, noting as well (Section 7.5) that detection based on the former schemes should proceed more quickly than that based on the latter. If this is so, then the *duration* of the signal presentation should not affect detection performance. One way to test this would be with a modified version of the size modulation experiment (Section 6.3) where now threshold measures of stimulus duration are obtained as a function of modulation amplitude and frequency. Reliance on a peak detection scheme, an integration scheme, or both together would be indicated by the following experimental outcomes:

1. No correlation between modulation amplitude and stimulus duration for any modulation amplitude would lend support for a peak detection scheme alone.
2. A negative correlation (threshold duration time increases with decreasing modulation amplitude) would lend support of an integration scheme.
3. No correlation for high modulation amplitudes and a negative correlation for low modulation amplitudes would indicate that both are being used.

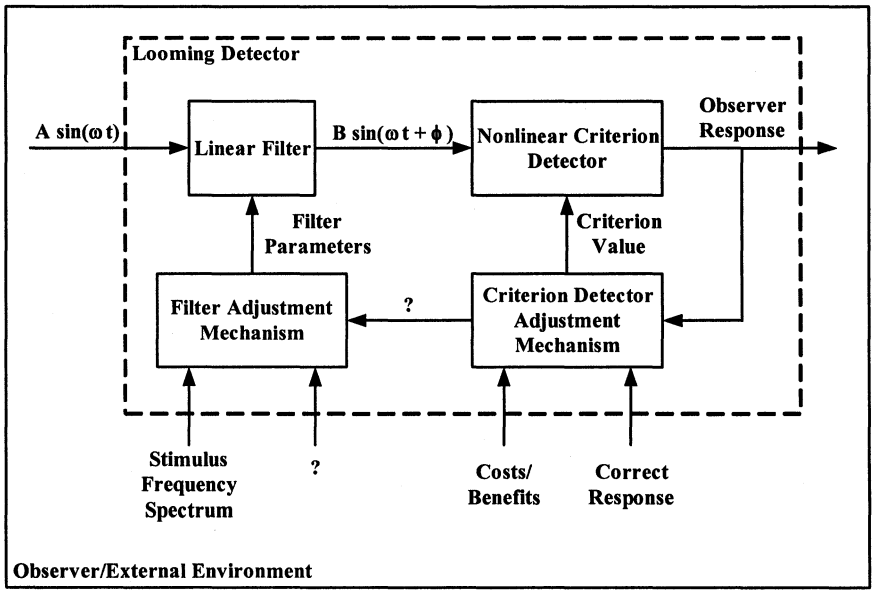
If the third result were encountered, our next task would be to identify the way in which the information generated by the separate schemes is combined to arrive at a final detection decision. A practical implication for the looming detection model is that the

time to detect a particular looming stimulus [Δt_2 in Equ. (27)] will vary depending upon which scheme "predominates" in the detection process.

9.3 Nonlinear Adjustment Mechanism

In our analysis of the Biphasic Impulse Response experimental results (Experiment 5, Section 7.4) we modified the filter transfer function obtained in the Size Modulation experiment (Experiment 3, Section 6.5) by reducing the filter's low frequency response characteristic (Figure 42). The resulting ability of the model based on this modification to predict the Experiment 5 results was observed to be very good. Several experimental sources for the discrepancy between the two experiment's results were suggested, as well as the possibility that further elaboration of the model itself was necessary. To this latter point it was suggested that a mechanism capable of adjusting the temporal characteristics of the filter to better match it to the task at hand may exist. Figure 49 shows one way in which this could be accomplished. The frequency spectrum of the input would be one input to this mechanism. The possibility that it receives other inputs and/or that its actions are coordinated with those of the criterion detector adjustment mechanism is also indicated. If such an adjustment mechanism exists, then it would have been active during the Size Modulation experiment, affecting those results. More investigations would have to proceed before we could come to a definitive conclusion either way, but the possibility leads us to pose the following questions:

1. Whether such a shift in the filter's frequency response characteristic is random or not, what is the range of its variability?
2. What are the Filter Adjustment Mechanism's inputs and outputs?
3. Are the Filter Adjustment Mechanism's actions coordinated with those of the Criterion Adjustment Mechanism to match the detector's performance to the task at hand? If so, how?



Modified Model of the Looming Detector
Figure 49

The discussion in these final two sections has focused on the initial detection of looming. Once looming is detected it will be necessary to continuously monitor an object's approach in order to respond effectively to it. Is the looming detector involved in this task as well? It appears to possess the necessary neural machinery for the purpose, and the existence of a separate mechanism possessing many of the same capabilities would not appear to be an efficient use of the visual system's limited resources. If the looming detector is active in the monitoring task, what role do the criterion detector and adjustment mechanisms play? Further research is required to answer these questions as well.

10 REFERENCES

1. Smith, A.T. and Snowden, R.J. "*Visual Detection of Motion.*", Academic Press, London, 1994, Chapter 9.
2. Regan, D. and Beverly, K.I. "Illusory Motion in Depth: Aftereffect of Adaptation to Changing Image Size." *Vision Research*, 1978, Vol. 18, pp. 209-212.
3. Regan, D. and Beverly, K.I. "Looming Detectors in the Human Visual Pathway." *Vision Research*, 1978, Vol. 18, pp. 415-421.
4. Regan, D. and Beverly, K.I. "Visual Responses to Changing Size and to Sideways Motion for Different Direction of Motion in Depth: Linearization of Visual Responses." *Journal of the Optical Society of America*, 1980, Vol. 70, pp. 1289-1296.
5. Regan, D. and Hamstra, S. "Dissociation of Discrimination Thresholds for Time to Contact and for Rate of Angular Expansion." *Vision Research*, 1993, Vol. 33, pp. 447-462.
6. Smith, A.T. and Snowden, R.J. "*Visual Detection of Motion.*", Academic Press, London, 1994, Chapter 12.
7. National Center for Statistics and Analysis General Estimates System. http://www-nrd.nhtsa.dot.gov/departments/nrd-30/ncsa/GES.html. Accessed November 8, 2003.
8. McRuer, D.T., Graham, D., Krendel, E.S. and W. Reisener. "Human Pilot Dynamics in Compensatory Systems". AFFDL-TR-65-15.
9. McRuer, D.T., Allen, R.W., Weir, D.H. and R.H. Klein. "New Results in Driver Steering Control Models". *Human Factors*. 1977, Vol. 19, No. 4, pp. 381-397.
10. Kleinman, D.L., Baron, S. and W.H. Levison. "An Optimal Control Model of Human Response. Part I: Theory and Validation, & Part II: Prediction of Human Performance in a Complex Task". Automatica; Vol.6, No. 3; pp. 357 – 384.
11. Levison, W.H. and Cramer, N.L. "A Description of the Integrated Driver Model". Report FHWA-RD-94-092. FHWA, U.S Department of Transportation. July, 1995.
12. Fancher, P.S. and Z. Bareket. "Evolving Model for Studying Driver-Vehicle System Performance in Longitudinal Control of Headway". *Transportation Research Record*. No. 1631, 1998, pp. 13-19.
13. Fancher, P.S. and Z. Bareket. "Evaluating Headway Control Using Range Versus Range-Rate Relationships". *Vehicle System Dynamics*. Vol. 23, No. 8, November, 1994, pp. 575-596.
14. Guo, K. and H. Guan. "Modeling of Driver/Vehicle Directional Control System". *Vehicle System Dynamics*. Vol. 22, No. 3-4, May/July, 1993, pp. 141-184.
15. Hoffman, E.R. "Note on the Detection of Vehicle Velocity Changes". *Human Factors*. Vol. 8, No. 2, April 1966, pp. 139-141.
16. Hoffman, E.R. and R.G. Mortimer. "Scaling of Relative Velocity Between Vehicles". *Accident Analysis and Prevention*. Vol. 28, No, 4, July 1996, pp. 415-421.
17. Cody, D. "Human Driver Model Development". *California PATH Project TO 4222*.

18. Kourjanski, M.A. and J.A. Misener. "Modeling the Driver: A Microsimulation Approach". *ITS America (1998: Detroit, Michigan)*. Transportation Technology for Tomorrow: Conference Proceedings.
19. Fancher, P., Z. Bareket and R. Ervin. "Human-Centered Design of an ACC-With-Braking and Forward-Crash-Warning System". *Vehicle System Dynamics*. Vol. 36, No. 2-3, 2001, pp. 203-223.
20. Purves, D., et. al. "Neuroscience", Sinauer Associates, 1997.
21. Wandell, B.A., "Foundations of Vision". Sinauer Associates, 1995.
22. Rodieck, R. and M. Watanabe. "Survey of the Morphology of Macaque Retinal Ganglion Cells that Project to the Pretectum, Superior Colliculus, and Parvicellular Laminae of the Lateral Geniculate Nucleus". *Journal of Comparative Neurology*. 1993, Vol. 338, pp. 289-303.
23. Baylor, D.A. "Photoreceptor Signals and Vision", *Investigative Ophthalmology and Visual Science*. Vol. 28, No. 1, 1987, pp. 34-49.
24. DeValois, R.L. and K.K. DeValois. "Spatial Vision", Oxford University Press, 1990, Chapter 3.
25. Kuffler, S.W. "Discharge Patterns and Functional Organization of Mammalian Retina". Journal of Neurophysiology. Vol. 16, 1953, pp. 37-68.
26. Rodieck, R.W. "Quantitative Analysis of Cat Retinal Ganglion cell response to Visual Stimuli". *Vision Research*. Vol. 5, 1965, pp. 583-601.
27. Enroth-Cugell, C. and J.G. Robson. "The Contrast Sensitivity of Retinal Ganglion Cells of the Cat". *Journal of Physiology*. Vol. 187, 1966, pp. 517-885.
28. Enroth-Cugell, C., Robson, J.G. and A.B. Watson. "Spatio-Temporal Interactions in Cat Retinal Ganglion Cells Showing Linear Spatial Summation". *Journal of Physiology*. Vol. 341, 1983, pp. 279-307.
29. Hassenstein, B. and W.E. Reichardt. "Systemtheoretische analyse der zeitreihenfolgen- und vorzeichenauswertung bei der bewegungspwezeption des rüsselkafers". *Chlorophanus. Z. Naturf.* Vol. 11b, 1956, pp. 513-524.

    Reichardt, W.E. "Autocorrelation, a Principle for the Evaluation of Sensory Information by the Central Nervous System". *Sensory Communication*. W.A. Rosenblith (Ed.), 1961, Wiley, New York.
30. Poggio, T. and W.E. Reichardt. "Visual Control of Orientation Behavior in the Fly" Part II: Towards the Underlying Neural Interactions". *Quarterly Review of Biophysics*. Vol. 9, 1976, pp. 377-438.
31. Hubel, D.H. and T.N. Wiesel. "Receptive Fields and Functional Architecture of Monkey Striate Cortex". *Journal of Physiology*. Vol. 195, 1968, pp. 215-243.
32. Sekuler, A.B. "Does the Visual System Possess Looming Detectors?" *Annual Meeting Abstract Issue: The Association of Research in Vision and Ophthalmology*. 1991, Vol. 32, No. 4, p. 831.

    Sekuler, A.B. "Simple-Pooling of Unidirectional Motion Predicts Speed Discrimination for Looming Stimuli". *Vision Research*. 1992, Vol. 32, No. 12, pp. 2277-2288.

33. Beverly, K.I., and D. Regan. "Separable Aftereffects of Changing-Size and Motion-In-Depth: Different Neural Mechanisms". *Vision Research*. Vol. 19, 1979, pp. 727-732.

Beverly, K.I., and D. Regan. "Visual Sensitivity to the Shape and Size of a Moving Object: Implications for Models of Object Perception". *Perception*. Vol. 9, 1980, pp. 151-160.

34. Derrington, A.M. and P. Lennie. "Spatial and Temporal Contrast Sensitivities of Neurons in Lateral Geniculate Nucleus of Macaque". *Journal of Physiology*. Vol. 357, 1984, pp. 219-240.

35. Enroth-Cugell, C. and L.H. Pinto. "Algebraic Summation of Centre and Surround Inputs to Retinal Ganglion Cells of the Cat". *Nature*. Vol. 226, 1970, pp. 458-459.

36. De Lange Dzn, H. "Experiments on Flicker and Some Calculations on an Electrical Analogue of the Foveal Systems". *Physica*, Vol. 18, No. 11, 1952, pp. 935-950.

De Lange Dzn, H. "Relationship Between Critical Flicker-Frequency and a Set of Low-Frequency Characteristics of the Eye". *Journal of the Optical Society of America*, Vol. 44, No. 5, 1954, pp. 380-389.

De Lange Dzn, H. "Research into the Dynamic Nature of the Human Fovea—Cortex Systems with Intermittent and Modulated Light. I. Attenuation Characteristics with White and Colored Light". *Journal of the Optical Society of America*, Vol. 48, No. 11, 1958, pp. 777-789.

37. Kelly, D.H. "Visual Responses to Time-Dependent Stimuli I: Amplitude Sensitivity Measurements". *Journal of the Optical Society of America*. Vol. 51, No. 4, 1961, pp. 422-429.

Kelly, D.H. "Visual Responses to Time-Dependent Stimuli II: Single-channel Model of the Photopic Visual System". *Journal of the Optical Society of America*. Vol. 51, No. 7, 1961, pp. 747-754.

Kelly, D.H. "Visual Responses to Time-Dependent Stimuli III: Individual Variations". *Journal of the Optical Society of America*. Vol. 52, No. 1, 1962, pp. 89-95.

38. Stork, D.G. and Falk, D.S. "Temporal Impulse Response From Flicker Sensitivities." *Journal of the Optical Society of America A*. Vol. 4, No. 6, June, 1987, pp. 1130-1135.

39. Kramers, M.H.A. "La Diffusion de la Lumière par les Atomes". *ATTI Congress Internazionale Dei Fisici*. 1927, pp. 545-557.

Krönig, R. de L. "On the Theory of Dispersion of X-Rays". *Journal of the Optical Society of America*. Vol. 12, No. 6, 1926, pp. 547-557.

40. Arfken, G. "Mathematical Methods for Physicists". Academic Press, New York, 1985.

Nussenzveig, H.M. "Causality and Dispersion Relations". Academic Press, New York, 1972.

41. Balas, G.J., Doyle, J.C., Glover, K., Packard, A. and Smith, R. "*μ-Analysis and Synthesis Toolbox*". MUSYN Inc. and The MathWorks, Inc., Natick, MA, 1998, pp. 8-52 – 8-55.

42. Oppenheim, A.V., and Schafer, R.W. "Digital Signal Processing", Prentice-Hall, Englewood Cliffs, NJ, 1975, Chapter 7.

43. Tyler, C.W. "Stereoscopic Depth Movement: Two eyes Less sensitive than One". *Science*, 1971, Vol. 174, No. 4012, pp. 958-961.

44. Harrison, J., Rensink, R.A., and M. van de Panne. "Detecting Changes in Velocity of Smoothly Moving Objects". *Vision Sciences Society Third Annual Meeting*, May 9-14, 2003, Sarasota, FL, p. 116.

.45 Evans, L. and R. Rothery. "Detection of the Sign of Relative Motion When Following a Vehicle". *Human Factors*, 1974, Vol. 16, pp. 161-173.

46. Schiff, W. and M.L. Detwiler. "Information Used in Judging Impending Collision". *Perception*, 1979, Vol. 8, pp. 647-658.

47. Tanner, W.P. and J.A. Swets. "A Decision-Making Theory of Visual Detection", *Psychological Review*. Vol. 61, No. 6, 1954, pp. 401-409.

48. Fechner, G. "Elemente de Pshchophysik", Vol. 2, Leipzig: Breitkopf and Härtel, 1860, p. 559 (Reprinted Bristol: Thoemmes Press, 1999).

49. Spiegel, M.R., J. Schiller, and R.A. Srinivasan. Theory and Problems of Probability and Statistics, Second Edition. Schaum's Outline Series. McGraw-Hill, New York, 2000.

50. Green, D.M. and J.A. Swets. Signal Detection Theory and Psychophysics, Krieger Publishing. Huntington NY, 1966.

51. Weibull, W.A. "A Statistical Distribution Function of Wide Applicability." *Journal of Applied Mechanics*. Vol. 18, 1951, pp. 292-297.

52. Watson, A.B. and D.G. Pelli. "QUEST: a Baysesian Adaptive Psychometric Method", *Perception & Psychophysics* Vol. 33, 1983, pp. 113-120.

53. Lee, D.N. "A Theory of Visual Control of Braking Based on Information About Time-to-Collision". *Perceptio*n. 1976, Vol. 5, pp. 437-459.

54. Pelli, D.G. and B. Farell. "Why Use Noise?". *Journal of the Optical Society of America – A*. 1999, Vol. 16, No. 3, pp. 647-653.

55. Vickers, W.H. "A Physiologically Based Model of Neuromuscular System Dynamics". *IEEE Transactions on Man-Machine Systems*. 1968, Vol. MMS-9, No. 1, pp. 21-23.

56. McRuer, D.T., Magdaleno, R.E. and G.P. Moore. "A Neuromuscular Actuation System Model". *IEEE Transactions on Man-Machine Systems*. 1968, Vol. MMS-9, No. 3, pp. 61-71.

57. Ledgeway, T. and A.T. Smith. "Separate Mechanism for the Detection of first and Second Order Motion in Human Vision". *ARVO Abstract issue, Investigative Ophthalmology and Vision Science*. Vol. 34, 1993, p. 1363.

58. Von der Heydt, R., Perterhans, E. and G. Baumgartner. "Illusory contours and Cortical Neuron Responses". *Science*. Vol. 224, 1984, pp. 1260-1262.

Von der Heydt, R. and E. Peterhans. "Mechanisms of Contour Perception in Monkey Visual Cortex. I. Lines of Pattern Discontinuity". *Journal of Neuroscience*, Vol. 9, 1989, pp.1731-1748.

59. Schrater, P.R., Knill, D.C., and E.P. Simoncelli. "Perceiving visual Expansion Without Optic Flow". *Nature*. 2001, Vol. 410, pp. 816-819.