Feasibility Study on a Hyperacuity Device With Motion Uncertainty: Two-Point Stimuli

Lei Wei, *Senior Member, IEEE*, Dennis M. Levi, Roger W. Li, and Stanley A. Klein

***Abstract*—Can a device that can perform hyperacuity vision tasks be built? In this paper, a feasibility study based on separation discrimination is conducted. Two types of ideal detectors are considered. The first is the stimulus defined statistically (SDS) detector by Geisler and Davilla. The second is one that estimates the uncertainty and then takes out its effect. In the first method, an array of many ideal stimulus defined exactly (SDE) detectors covers uncertainties and forms the ideal SDS detector. When the separation distance between the SDE detectors is around 1 arcmin, the SDS detector can achieve nearly optimal performance. To cover the motion uncertainty with nearly optimal performance, the SDE detector at each position needs to cover 16 directions, and at each direction, it needs to cover speeds with an increment of 0.5°/s. Typically, the SDS detector needs 7776 SDE detectors to deal with a speed up to 2°/s stimulus movement with a randomly selected direction and a 9 min × 9 min position uncertainty region. This ideal observer can achieve a hyperacuity threshold of 2–4 arcsec. Its threshold is almost constant over the range of speeds covered by the SDS detector. Using the second method, position estimation and motion tracking capability is examined. With perfect position estimation and motion tracking, the SDS detector can be reduced to a single SDE detector that is tuned to correct position and motion direction and speed. Two lower bounds on the estimation variance are examined, namely: 1) the Cramer–Rao bound and 2) the Ziv–Zakai bound. The results showed that if an estimation algorithm that can achieve the performance of bounds can be found, then the second method could achieve a hyperacuity capability of 1 s or less. The human visual system may more likely adopt the first method, but the second is simpler to use in building up a device using microprocessors. In this paper, dot-pair templates are used, which precisely match to dot-pair stimuli, to compute the likelihood and make a decision. The main difference between this paper and that of Geisler and Davilla lies in using a discrete sum over an array of SDE detectors to closely approximate continuous integration over the uncertain region, which makes it much easier in hardware implementation.**

***Index Terms*—Distributed detection, image resolution, signal detection, visual system.**

I. Introduction

THE HUMAN visual system is capable of judging relative position with remarkable accuracy. Thresholds for these tasks could come down to 1 arcsec [1]–[3], which is about 1/30 of the intercone spacing. For this reason, Westheimer [4] has coined the term hyperacuity to describe a variety of tasks that involve sensing the direction of spatial offset of a line or point relative to a reference. Provided target energy or target visibility is kept constant, the hyperacuity capability is very robust to the exposure duration, luminance [5], [6], and motion [7].

High visual acuity is very important in many professions. For example, a diamond trader often needs to detect crack lines or clouds using a 10× loupe. An operator of a precision metal-polishing machine needs to examine the metal surface constantly to ensure a perfect flat finish. Many scientists in a space center are constantly monitoring the space shuttle during takeoff. Misalignment between insulation tiles could be an indicator of problems. Over the decades, many advanced tools have been developed for these workers, mainly magnifiers, microscope tools, or digital cameras with image processing units. Nevertheless, in many cases, examination using human vision is still essential at the final stage of the process, which may be limited by the observer's visual capability. Can we build a device that a worker can wear at his or her forehead to assist him or her in decision making? For example, if the surface is not sufficiently smooth, a beep sound would be generated to alert the worker to pay closer attention to surface. Another example, if a scientist is monitoring the contour of the flying spacecraft or space shuttle on a screen (the displaced image is often averaged over many frames), a beep sound may indicate the abnormal misalignment in a contour line. Such a device will not only take out the influence of observer's emotion but can also improve the performance. Research has shown that an ideal observer can often detect a misalignment ten times smaller than ordinary human observers.

Three mechanisms in vision research have been proposed to explain the hyperacuity, namely: 1) the optimal discriminator methods [8]–[12]; 2) the local sign method [2]; and 3) the contrast-sensitive spatial filter method [1], [13].

The purpose of this paper is to study the feasibility of such a device based on the ideal observer concept. We will focus on the detector part of system design. We assume the optical part of the device and other features as close to the human visual system as possible. That is, we consider the influences of the eye's point spread function, Poisson noise, fixational eye movements without head constraint, and movements of the stimulus. Furthermore, in order to compare the results in [12], we select a simulation environment identical to that in [12], including two-dot stimuli and separation discrimination.

Manuscript received May 14, 2005; revised February 14, 2006 and June 26, 2006. This paper was recommended by Associate Editor S. Sarkar.

L. Wei is with the School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816 USA (e-mail: lei@ee.ucf.edu).

D. M. Levi, R. W. Li, and S. A. Klein are with the School of Optometry, University of California, Berkeley, CA 94720 USA (e-mail: jqaut@osa.org).

Digital Object Identifier 10.1109/TSMCB.2006.883871

In [11], Geisler showed that the hyperacuity can be achieved by a likelihood detector if the detector knows the exact position and configuration of the stimuli [called the stimulus defined exactly (SDE) observer]. In [12], Geisler and Davilla used an array of many SDE detectors that covers position uncertainties to the ideal stimulus defined statistically (SDS) detector.

The device aims to extend the human vision capability one step forward. Specifically, we will examine the following issues in this paper.

1) Can we improve performance under the condition the human visual system fails, for example, stimuli with low intensity or high speed?
2) Can we implement a detector with low complexity? Two methods are considered to deal with uncertainty. Which one is simpler and better?
3) How does the detector behave in the regions of high accuracy? Typically, vision research focuses on performance level of 60%–90% of accuracy, since it is difficult to perform experiments with human subjects for high accuracy. For a device, we would like to understand its behavior in the regions that it can achieve 99.9% accuracy. If the device is built for 60%–80% accuracy, then human observers may tend to ignore the device.

This paper is organized as follows. In Section II, we describe the setting for our study and select the parameters for the stimuli, point spread, and fixational movements. In Section III, we derive two ideal detectors for moving stimuli. One is a direct extension of the SDS observer in [12]. Another consists of one SDE detector and one estimator. The detector will discriminate the separation of two dots (i.e., δ and $\delta + \Delta\theta$), respectively. In Section IV, we use the ideal observer array to cover the motion of stimuli (i.e., method 1) and examine its performance and complexity. In Section V, we examine the performance of method 2 by evaluating two lower bounds on variance of estimations. Section VII concludes this paper.

II. Study Setting and Parameter Selection

Fig. 1 shows the study setting that consists of an artificial pupil and a lens, a photoreceptor layer, and a detector that processes the signal's input from the photoreceptor layer to make a decision.

A. Stimuli Selection

We select the two-dot stimuli identical to that for separation discrimination [12]. That is, in stimulus A, two dots are separated by δ arcmin with orientation along the x axis; whereas in stimulus B, two dots are separated by $\delta + \Delta\theta$ arcmin. The diameter of each dot is 0.2 arcmin. The study method is the two-alternative forced choice (2AFC) method. We randomly select one of two stimuli, present it to the device, and record the probability of a correct decision. The $\Delta\theta$ value corresponding to 75% correct decision is recorded as the threshold.

Identical to the setting in [12], we will consider the stimuli in two situations, namely: 1) no background illuminance

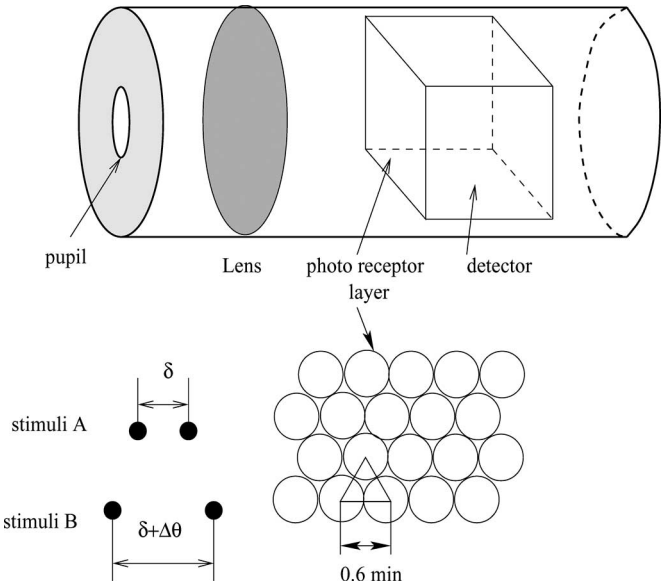


Fig. 1. Study setting.

(i.e., a dark background, $T_B = 0$) and 2) a background level of 20.8 Td (i.e., 7.41 log quanta/s/$^{\circ -2}$ at the cornea or $T_B = 2.6 \times 10^7$ quanta/s/$^{\circ -2}$). For a sampling rate of 1000 samples/s and a cone diameter of 0.6 min, the background level is then 2.6 quanta per sample time per cone.

To ensure a fair comparison between the human data in [12] and the detectors, we use all parameters closely related to the human visual system given in [12].

B. Receptor Lattice

It is assumed that the receptors are positioned at the nodes of a hexagonal array such that the internodal distance equals the diameter of the receptors. This provides the tightest possible packing of photoreceptors (see Fig. 1). The diameter of the receptors is assumed to be 0.6 arcmin [11]. The total size of the receptor lattice, depending on the stimuli, was always picked, so any increases in lattice size produced no further improvements in performance.

C. Point Spread Function

The line spread function is approximated by the sum of two 2-D Gaussian functions [11], [12], i.e.,

$$h(x) = \frac{a_1}{2\pi a_3} \exp\left[\frac{-0.5x^2}{a_3^2}\right] + \frac{a_2}{2\pi a_4} \exp\left[\frac{-0.5x^2}{a_4^2}\right]. \quad (1)$$

For a 3-mm pupil, the coefficients are $a_1 = 0.417$, $a_2 = 0.583$, $a_3 = 26.58$ s, and $a_4 = 122.4$ s. The curve fits well to the measurement data from Campbell and Gubisch [21]. That is, if a tiny dot stimulus is presented, then the image is spread at the retina, as shown in Fig. 2.

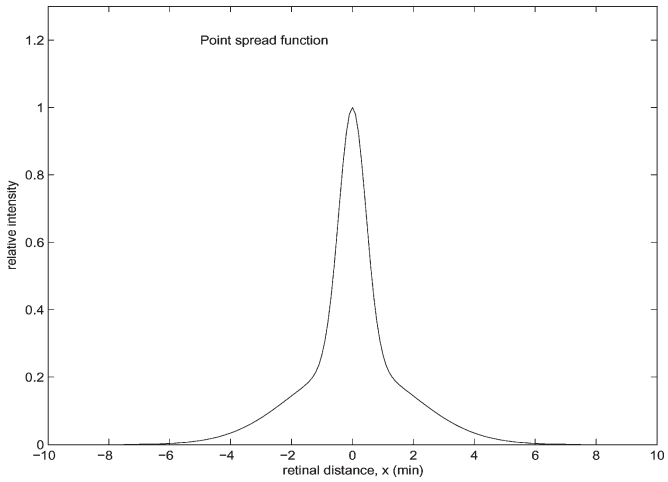


Fig. 2. Point spread function.

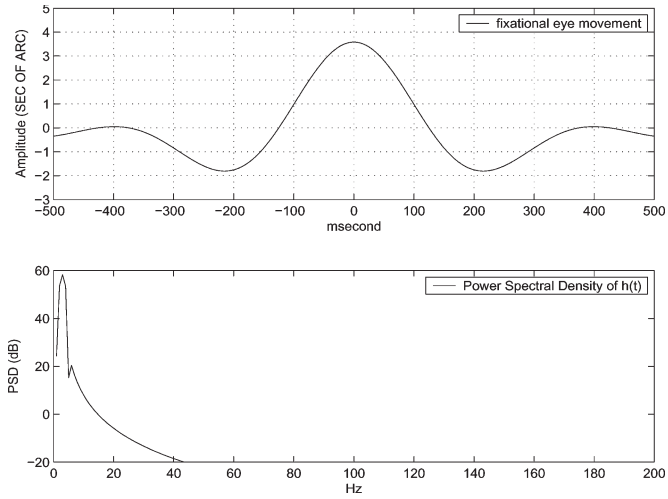


Fig. 3. Impulse responses of the filter to generate fixational eye movements.

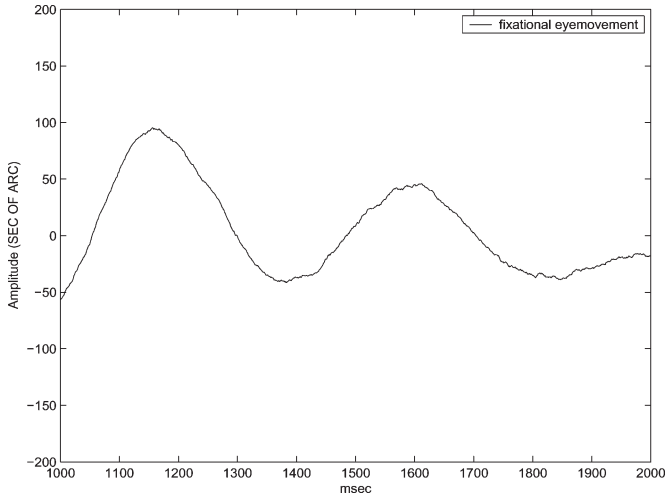


Fig. 4. Eye movement random-walk waveforms (1-s duration).

D. Poisson Distribution

In a fixed time period, the number of photons (R^i) absorbed in the ith photoreceptor of the lattice is described by the Poisson density: $p(r) = F^r \exp(-F)/r!$, where $r = 0, 1, 2, \ldots$, and F is the mean number of quanta absorbed. In order to compute the mean number of quanta absorbed, we need to study the effect of eye movement.

E. Model of Fixational Eye Movements

Fixational eye movements mainly drift. Under different situations such as standing, sitting, or while on a bite board, drift rates are different. Under the same situation, different persons may have different drift patterns and rates [14]. Typically, the drift rate is 2–5 Hz, and peak-to-peak amplitude is less than 5 arcmin [15]. The drifts along the horizontal and vertical directions are independent. The drifts between the two eyes are uncorrelated [17]. Under the condition of bite boards or head or chin rests, the percentile of samples exceeding eye movement over certain seconds or minutes of arc has been measured as a function of exposure time [16]. With an unrestrained head, gaze is less stable than with a head rest; typically, the amplitude increases by 1.5–3 times. A good review of eye movement can be found in [19].

We model the drift-like eye movement as a Gaussian random-walk function as follows:

$$x(t) = h(t) \otimes \phi_x(t) \tag{2}$$

$$y(t) = h(t) \otimes \phi_y(t) \tag{3}$$

where $\otimes$ denotes the convolution operation, $h(t)$ is the impulse response (3), and ϕ_x and ϕ_y are independent Gaussian random processes with zero mean and unitary variance. Engineers often model a random-walking stochastic process by passing a whitening Gaussian process through a filter with impulse response $h(t)$. In this paper, we tune $h(t)$ until the process fits well with the observation results in [16]. A good $h(t)$ is a Gabor-like function, i.e.,

$$h(t) = \left(\frac{2400}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{t^2}{2\sigma_i^2} \right) - 0.7 \right) \cos(4\pi t/1000) \tag{4}$$

where $h(t)$ is in seconds of arc, $\sigma_i^2 = 50\,000$ ms^2. In Figs. 3–5, we present the impulse response and drift-like movement over 1 and 10 s. We examine the percentage time of samples exceeding the criterion amplitude of movement, and the results are given as follows: for a 10-ms exposure time, 8.3% of sample exceeding $10''$ and 39.5% exceeding $5''$; for 20 ms, 9.3% exceeding $20''$ and 71% exceeding $5''$; and for 50 ms, 5% exceeding $60''$, 18% exceeding $40''$, 51% exceeding $20''$, and 78% exceeding $10''$. These results fit well with experimental data [16]. For fixational eye movement with free head, we will increase the amplitude by two times.

The power spectrum in Fig. 3 shows the main components at the range of 2–4 Hz. The curve of the spectrum closely approximates to the main curves in [18]. For simplicity, we did not consider other kinds of fixation instability studied in [19] and [20].

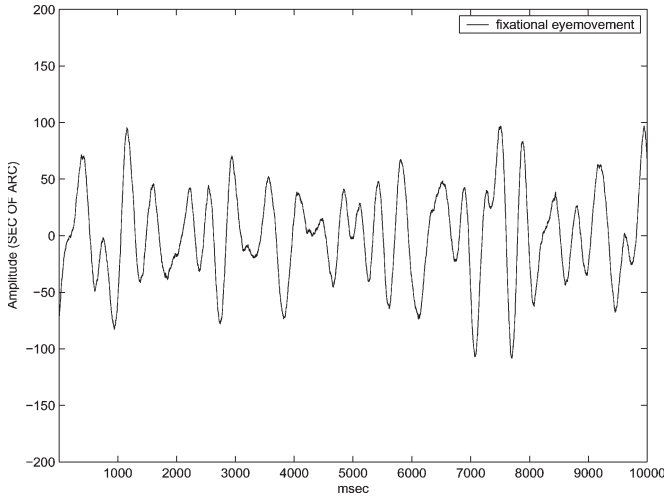


Fig. 5. Eye movement random-walk waveforms (10-s duration).

It is worth mentioning here that the images of the stimuli under these conditions are very close to those images that have already passed through the human eye, reaching the neuronal layers. They are severely blurred by point spread functions and eye movement. If we build a system using a charge-coupled device (CCD) camera, then the images fed into the detector could be much clearer. The principles for achieving hyperacuity using the CCD camera could be very different from those in human eyes. On the other hand, if we can deliver a device that can work with a much poorer image quality, then this device would be much more valuable. For this reason, in this paper, we assume that the images fed into the detector are as poor as those in the human eye.

III. Ideal Detectors for Motion Uncertainty

In this section, we first derive an ideal detector for position and motion uncertainty, which is a direct extension of Geisler and Davilla's SDS detector [12]. Two synchrony mechanisms are considered in this paper. The first is by using an array of SDE detectors. Each SDE detector is optimal for the stimuli at a given position, moving direction, and speed. Hundreds or thousands of SDE detectors cover the uncertainty region of position, moving speed, and direction. The other is by estimating the uncertainty from snapshots, then combining all snapshots with compensated speed, direction, and position uncertainties to form a picture. When we take away the uncertainty, it is reduced to Geisler's SDE observer [11]. If we take away the motion uncertainty, it is reduced to Geisler and Davilla's SDS detector [12].

To model the effect of combining both eye and stimulus movements, we assume that during the observation time, the stimulus stops at N_t times, spends $T_s = T/N_t$ ms at each stop, and at the end of each stop time, jumps to the next stop, where T is the duration of the stimulus in milliseconds. We focus on 555 nm. Effectively, the mean number of absorbed quanta for the receptor located at the position (x, y) is given by

$$F^{(x,y)} = TS\Upsilon_{555}E_{555}0.3478l(x,y) * h(x,y)\xi_r/N_t \qquad (5)$$

where S is the pupil area in square millimeters, Υ_{555} is the transmittance of the ocular media at 555 nm, E_{555} is the quantum efficiency of the photopigments at 555 nm, $l(x, y)$ is the luminance distribution of a stimulus in candelas per square meter; $h(x, y)$ is the point spread function, and ξ_r is the receptor area (in this case, it is a cylinder with a diameter of 0.6 min). We take $\Upsilon_{555} = 0.67$, $E_{555} = 0.5$, and $S = 7.07$ min^2 [12]. The symbol $*$ represents the operation of convolution. Equation (5) computes the effective absorbed photons per photocell at location (x, y).

In a period of a single stimulus stop T/N_t, the number of effective quantum absorptions $r^{(x_i,y_i)}$ in the ith photoreceptor of the lattice is described by the Poisson density defined as follows:

$$p\left(r^{(i)}\right) = \left(F^{(i)}\right)^{r^{(i)}} \exp\left(-F^{(i)}\right)/r^{(i)}! \qquad (6)$$

where the superscript (i) denotes an abbreviation for the superscript (x_i, y_i).

If we do not know the position of each stop, then we shall average it over all possible positions. Assuming the position of stop k is (c_{x_k}, c_{y_k}), the parameter v in (7) [12] is then replaced by $\mathbf{v} = [v_1, \ldots, v_{N_t}] = [(c_{x_1}, c_{y_1}), \ldots, (c_{x_{N_t}}, c_{y_{N_t}})]$. Letting $p(\mathbf{v})$ denote the probability density function of $\mathbf{v}$, then we have the likelihood ratio of the optimum discriminator as

$$L = \frac{\int_{\mathbf{v}} \exp\left[\sum_{k=1}^{N_t} l(B,k)\right] p(\mathbf{v})d\mathbf{v}}{\int_{\mathbf{v}} \exp\left[\sum_{k=1}^{N_t} l(A,k)\right] p(\mathbf{v})d\mathbf{v}} \qquad (7)$$

where

$$l(X,k) = \sum_i r_k^{(x_i,y_i)} \ln\left(F_{X^{(0,0)}}^{(x_i-c_{x_k},y_i-c_{y_k})}\right) - F_{X^{(0,0)}}^{(x_i-c_{x_k},y_i-c_{y_k})}. \qquad (8)$$

Here, $X = A$ or B denotes stimuli A or B, $F_{B^{(c_{x_1},c_{y_1})}}^{(x,y)}$ denotes the number of quanta absorbed in by the receptor located at (x, y) when stimulus B is at position (c_{x_1}, c_{y_1}), and $l(X, k)$ is the likelihood function.

We can obtain the SDE detector from (7) by setting $p(\mathbf{v}) = \delta(\mathbf{v})$, where $\delta(x)$ is the impulse function, i.e., $\delta(x) = 1$ if $x = 0$, otherwise, $\delta(x) = 0$, i.e.,

$$L_{\mathrm{SDE}} = \frac{\exp\left(l_{\mathrm{SDE}}(B)\right)}{\exp\left(l_{\mathrm{SDE}}(A)\right)} \qquad (9)$$

where

$$l_{\mathrm{SDE}}(X) = \sum_i \left(\left(\sum_{k=1}^{N_t} r_k^{(x_i,y_i)} \right) \times \ln\left(N_t F_{X^{(0,0)}}^{(x_i,y_i)}\right) - N_t F_{X^{(0,0)}}^{(x_i,y_i)} \right). \qquad (10)$$

When the ideal discriminator does not know the initial position of the stimuli, it needs a matrix of SDE observers covering the entire uncertainty region. Since it does not know which SDE observer will receive the signal, it makes its decision

based on the average of all the SDE observers. Geisler and Davila's SDS observer is the ideal detector for this case. We can obtain the SDS detector from (7) by setting $p(\mathbf{v}) = p(v_1)\delta(v_2 - v_1)\ldots\delta(v_{N_t} - v_1)$. That is, if during the observation the stimulus position does not move, then we have $(c_{x_k}, c_{y_k}) = (c_{x_1}, c_{y_1})$ for $k = 2, \ldots, N_t$. Thus, we have Geisler and Davila's observer as

$$L_{\mathrm{SDS}} = \frac{\int_{v_1 \in R^u} \exp\left[\sum_i l_{\mathrm{SDS}}(B,i)\right] p(v_1)dv_1}{\int_{v_1 \in R^u} \exp\left[\sum_i l_{\mathrm{SDS}}(A,i)\right] p(v_1)dv_1} \tag{11}$$

where R^u denotes the region of uncertainty and

$$l_{\mathrm{SDS}}(X,i) = \left(\sum_{k=1}^{N_t} r_k^{(x_i,y_i)}\right) \ln\left(\Psi(X,i)\right) - \Psi(X,i). \tag{12}$$

Here, $\Psi(X,i) = N_t F_{X^{(0,0)}}^{(x_i - c_{x_1}, y_i - c_{y_1})}$.

In (11), the stimulus is fixed at the uncertain position during the entire presentation. If the stimulus moves during the presentation, then two methods can be used to deal with the motion uncertainty, namely: 1) use SDE detectors to cover the entire position and motion uncertainty regions and 2) estimate the motion uncertainty and then cancel its effect.

For the first method, we only need to construct an array of SDE detectors in which each SDE detector is ideal for a given initial position, moving direction, and speed. If we know the initial position, moving speed, and direction, then the SDE observer can be obtained by shifting each position accordingly. For example, if the stimulus will stay in two positions, i.e., 1 and 2 (50% of duration each) during 0.1 s of presentation, then we have $F_{B^{(c_{x_1}, c_{y_1})}}^{(x,y)}$ and $F_{B^{(c_{x_2}, c_{y_2})}}^{(x,y)}$. If we do not know the exact position of (c_{x_1}, c_{y_1}) but know the difference precisely, i.e., $(c_{x_2} - c_{x_1}, c_{y_2} - c_{y_1})$, then we can obtain $F^{(x_i,y_i)} = F_{B^{(c_{x_1}, c_{y_1})}}^{(x_i,y_i)} + F_{B^{(c_{x_2}, c_{y_2})}}^{(x_j,y_j)}$, where $x_j = x_i + c_{x_2} - c_{x_1}$, $y_j = y_i + c_{y_2} - c_{y_1}$. We can view that each duration is a snapshot or a frame of cone inputs. The motion is reflected in position differences of dots in frames. By adjusting the position of each frame, we can stack all frames in such a way that the positions of dots are identical over all frames. One can vertically thread a needle through one dot of all frames. After this combination, it is reduced to the SDE observer in (9). In summary, we have

$$L_{\mathrm{SDE}_m} = \frac{\exp\left(\sum_i l_{\mathrm{SDE}_m}(B,i)\right)}{\exp\left(\sum_i l_{\mathrm{SDE}_m}(A,i)\right)} \tag{13}$$

where

$$l_{\mathrm{SDE}_m}(X,i) = \left(\sum_{k=1}^{N_t} r_k^{(x_{i_k}, y_{i_k})}\right) \left(\ln\left(\Psi(X,i)\right) - \Psi(X,i)\right). \tag{14}$$

Here, $\Psi(X,i) = N_t F_{X^{(0,0)}}^{(x_i - c_{x_1}, y_i - c_{y_1})}$, $x_{i_k} = x_i - c_{x_1} + c_{x_k}$, and $y_{i_k} = y_i - c_{y_1} + c_{y_k}$.

The first ideal detector can be obtained by forming an array of SDE detectors given in (13), as shown in Fig. 6. To implement this type of detector, we need to examine how many SDE detectors are sufficient to achieve nearly optimal performance.

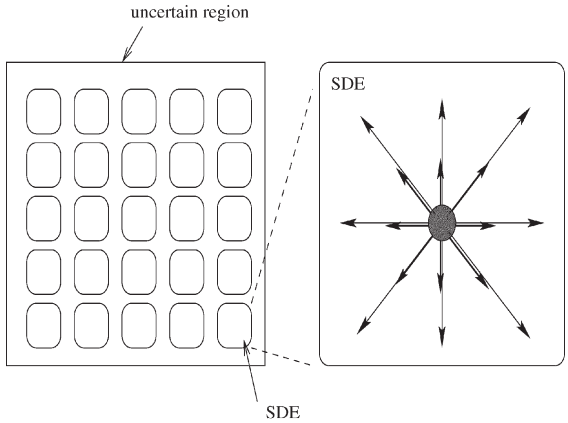


Fig. 6. Illustration of the first method SDS detector for motion uncertainty.

The second detector needs to estimate the position at each step k, i.e., $c_k = (c_{x_k}, c_{y_k})$. If the estimation is perfect, then the second detector is reduced to (13). If the estimation is not perfect, which is often classified as a random variable with a PDF of $p(c)$, then we have

$$L_{D2} = \frac{\int_c \exp\left(\sum_i l_{D2}(B,i,c)\right) p(c)dc}{\int_c \exp\left(\sum_i l_{D2}(A,i,c)p(c)dc\right)} \tag{15}$$

where

$$l_{D2}(X,i) = \left(\sum_{k=1}^{N_t} r_k^{(x_{i_k}, y_{i_k})}\right) \left(\ln\left(\Psi(X,i)\right) - \Psi(X,i)\right). \tag{16}$$

Here, $\Psi(X,i) = N_t F_{X^{(0,0)}}^{(x_i - c_{x_1}, y_i - c_{y_1})}$, $x_{i_k} = x_i - c_{x_1} + c_{x_k}$, and $y_{i_k} = y_i - c_{y_1} + c_{y_k}$, and c_k is generated from the distribution $p(c)$. Clearly, this detector just needs one estimator and one SDE detector. Its performance is directly related to the accuracy of estimation, which is often evaluated by the estimation variance.

In summary, the optimality of Geisler and Davilla's detector [12] can only be realized when we have infinite SDE detectors to cover the uncertainty region. If we have a finite number of SDE detectors, the first detector is suboptimal. The second detector is also suboptimal if we cannot perfectly estimate its position, speed, and direction.

IV. Performance and Complexity of the First Detector

In this section, let us focus on the first detector, i.e., using an array of SDE detectors to cover position and motion uncertainties. We will consider fixed stimulus first, and then moving stimulus afterward.

A. Stationary Stimulus

Now, let us study the effect on the hyperacuity threshold by decreasing the density of the SDE detector array in the SDS observer. Consider an uncertain region of 4.5 min × 4.5 min (roughly 20 min^2 of arc). The position of the stimulus is uniformly distributed in this region, and the stimulus is fixed

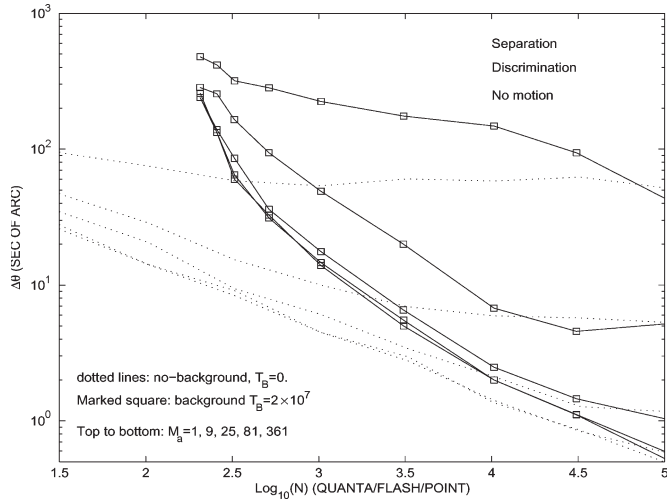


Fig. 7. Threshold for different numbers of SDE observers in the SDS detector, without motion.

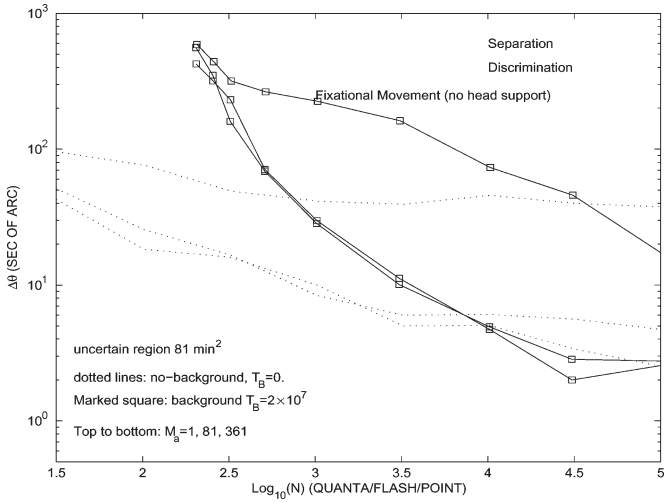


Fig. 9. Threshold for different numbers of SDE observers in the SDS detector, with fixational movement without head support.

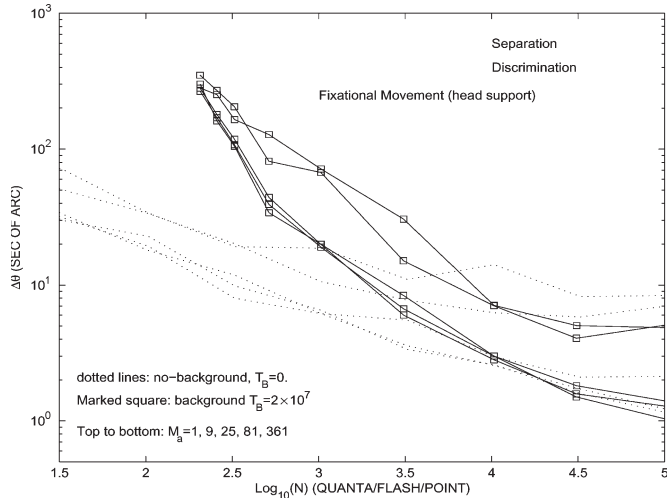


Fig. 8. Threshold for different numbers of SDE observers in the SDS detector, with fixational movement with head support.

for a duration of 0.1 s. We select 0.1 s since the eye has very poor capability of pursuing a moving target during 0.1 s without prior information of motion direction and speed [7]. Thus, we can model the fixational eye movement during this period as a smooth random-walking process.

The SDS observer consists of an array of $M_a = M_x \times M_y$ SDE observers that are evenly distributed over the uncertain region. In Fig. 7, we show the threshold as a function of $\log_{10}(N)$, with $M_a = 1$, 9, 25, 81, 361, respectively, and the corresponding separation distance between the SDE detectors is $2'$, $1'$, $0.5'$, and $0.25'$ for $M_a = 9, 25, 81, 361$, respectively. N is defined as the half sum of $F^{(x,y)}$ over all cells and over a period of T, i.e., quanta per stimulus dot per flash. The results show that with nine SDE detectors, the SDS observer can perform within 10 arcsec for large Ns.

Even if the stimulus is fixed, eye movements will cause the stimulus drift at the photoreceptor layer. Now, we repeat the same simulations under the fixational movement [i.e., (2) and (3)] with head support (see Fig. 8) and without head support (see Fig. 9). Surprisingly, with a single SDE detector, it is possible to achieve a threshold of a few seconds of arc with large N. Clearly, it is much better than without fixational movement. It is easily understandable. Fixational movements increase the chances that the stimulus lands on the SDE observer. For the case without head support, the amplitude of motion is twice that with head support so that the uncertain region becomes 9×9 min^2. The data is for $M_a = 1$, 81, 361, respectively. The corresponding separation distance between the SDE detectors is $1'$ and $0.5'$ for $M_a = 81$ and 361, respectively. Clearly, the increase of motion degrades the threshold gradually across all cases.

The results in these figures show that nearly optimal performance can be achieved if the separation distance in position between SDE detectors is around 1 min. One can marginally improve the threshold by increasing the SDE detector density.

B. Moving Stimulus

What are the speed and direction incremental values? We construct four SDE detectors to detect the threshold when the stimulus starts from a position uniformly and randomly selected from the uncertain region (9 min $\times$ 9 min, i.e., fixational movement without head support) and moves along the x axis with a speed of $2^\circ/$s. The four SDE detectors are for the cases that move at a speed of $2 \pm \Delta_s/2^\circ/$sec along the directions of $\pm\theta/2$.

Another way to look at the first detector is given as follows. We can view that the first detector uses a number of cameras. Each camera sits at a given initial position and moves in a given direction and at a given speed. Each camera takes one snapshot per millisecond, and we then combine 100 snapshots together to form a picture. After that, we get some good picture out of those pictures in order to examine which stimulus it is. When there is no motion, with 81 cameras (to cover 4.5 min $\times$ 4.5 min), it can produce the results in [12] (less than 1 s separation threshold). However, when there is stimulus motion plus eye movement, each camera would not be optimal. Performance would be reduced. More cameras would be better to cover the

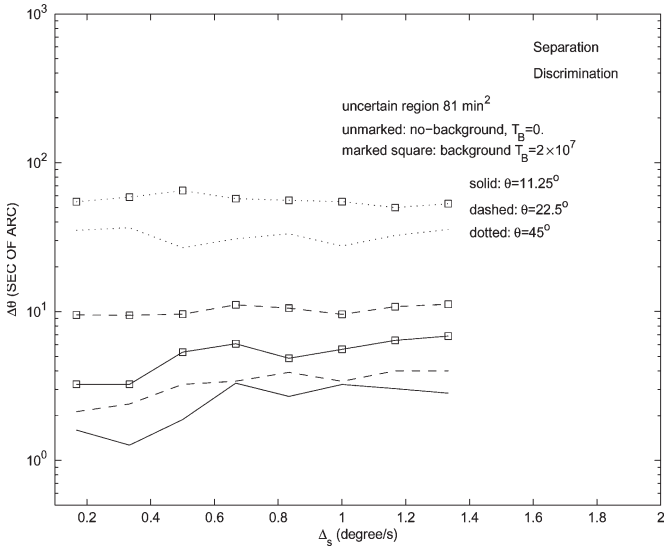


Fig. 10. Effects on threshold by increments in speeds and directions.

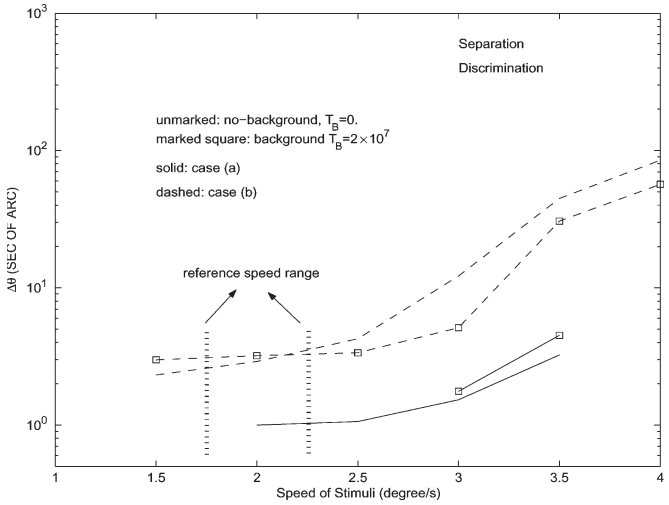


Fig. 11. Effects on threshold by increments in speeds and directions.

uncertainty, but there would also be more pictures to sort out, which could be very difficult.

Fig. 10 shows the effects of Δ_s and θ on the threshold. This shows that we can select 16 equally spaced directions and five equally spaced speeds ($0^\circ/$s, $0.5^\circ/$s, $1^\circ/$s, $1.5^\circ/$s, and $2^\circ/$s) as reference directions and speeds. What will happen if the speed of the stimulus is above the maximum reference speed value? In Fig. 11, we examine the threshold as a function of stimulus speed from $1.5^\circ/$s to $4^\circ/$s, where in case (a), the reference speed is 1.75 and $2.25^\circ/$s, the reference directions are $\pm\pi/4$, and the stimulus moves in the direction of 0° and in case (b), the reference speeds are $1.75^\circ/$s, $1.85^\circ/$s, $1.95^\circ/$s, $2.05^\circ/$s, $2.15^\circ/$s, and $2.25^\circ/$s, the reference direction is set to $\theta = 45^\circ$, and the stimulus moves with a direction randomly selected from 0 to 2π. In case (b), we consider the fixational movement without head support.

Another view of case (b) is given as follows. There are 6 speeds ($1.75^\circ/$s to $2.25^\circ/$s) $\times$ 16 directions (45° apart) $= 96$ cameras. Each of them is moving at a given speed (between $1.75^\circ/$s and $2.25^\circ/$s) and a given direction. The stimulus moves

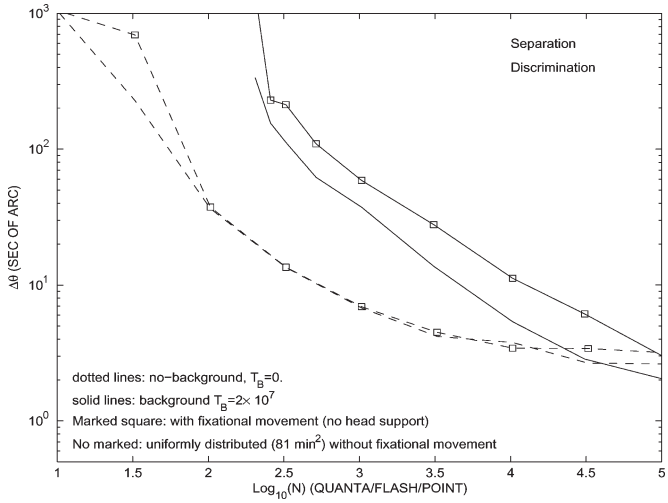


Fig. 12. Effects of uncertainties in speeds and directions on the threshold.

at a speed from $1.5^\circ/$s to $4^\circ/$s at a random direction. When the speed is over the maximum camera speed, no camera can follow it up. Thus, the threshold is degraded. If it is below the maximum camera speed, then the threshold shall not be affected. The camera speed in the human visual system may not cut off completely at $2.25^\circ/$s. Thus, the kink point could be higher.

What kind of impact on SDE and SDS detectors is derived from motion uncertainty? The results in [12] showed that position uncertainty does not affect the threshold when N is large. This is true even with motion uncertainty since, for large N, the uncertainty of speed, direction, and position can be significantly reduced by examining the likelihood value of each component SDE observer. However, the results in [12] showed that when N is small, position uncertainty affects the case with background illumination but not the case with dark background. In Fig. 12, we present thresholds for the SDS detector with an uncertain region (9 min $\times$ 9 min). The SDS detector is constructed as follows. The reference speeds are from $0^\circ/$s to $2.5^\circ/$s with an increment of $\Delta_s = 0.5^\circ/$s; the phase separation between reference direction is $\theta = 22.5^\circ$; the SDE detector array is formed with horizontal space of 1.2 min and vertical space of $1.2\sqrt{3}/2$. The stimulus moves at the speed of $2^\circ/$s and in a direction randomly selected from [0, 2π]. We also check the case in which the stimulus does not move, or move at different speeds, the results are almost identical. The results clearly show that the uncertainties in speed and direction will affect the threshold at small N with dark background. This clearly demonstrates that motion uncertainty is one of the causes for the difference between human observers (Figs. 3–5) [12] and the ideal SDS observer at small N with dark background. We can achieve a three to four times improvement over human observers, but we would not be able to achieve the superior performance predicted in Figs. 3–5 [12] for regions of small N when the background is dark.

The phenomenon stated above can be explained as follows: When there is motion uncertainty, a photon received at a particular location could be weighted almost equally in many SDE detectors even without background illumination. For example, if a photon hits the center, then it can either be the central SDE

detector without motion or the surrounding SDE detectors with motion. Thus, it presents to the SDS observer many possible choices, and the SDS observer will make more errors. If there is no motion uncertainty, then it will be very rare to be weighted similarly (unless the photons land on a place that has an equal space between two SDE detectors). Thus, the SDS detector shall have a clear picture about the position of the stimulus.

In summary, we need $9 \times 9 \times 16 \times 6 = 7776$ SDE detectors to cover position uncertainty of 9 min $\times$ 9 min (assuming that the separation distance between SDE detectors is 1 min), direction uncertainty, and speed uncertainty (assuming six speeds of 0°/s, 0.5°/s, 1°/s, 1.5°/s, 2.0°/s, and 2.5°/s). Performance is about three to four times better than human observers.

V. Performance Bounds for the Second Detector

The second detector uses one estimator and one camera. The estimator tells the camera where to take a snapshot. The camera takes one snapshot per millisecond and then combines 100 snapshots together to form a picture. In the end, we have one picture to make a decision on which stimulus it is. When the absorbed photons are very high, there is no uncertainty on position, speed, and direction over a few snapshots, so the detector can perform close to a stationary stimulus and produce accuracy of less than 1 s.

To understand the performance of the second detector, we need to examine the accuracy of estimation. We study two lower bounds, namely: 1) the Cramer–Rao bound (CRB) [22] and 2) the Ziv–Zakai bound (ZZB) [23], [24], on the variance of eye capability of tracking a moving dot. CRB is a lower bound for all unbiased estimators, which may not be tight for low signal-to-noise ratio (SNR) regions. ZZB is a tighter lower bound for the maximum likelihood estimator. Optimal estimators often perform very close to the bounds. Thus, based on the variance bounds, we can evaluate the performance of the second detector.

In order to examine the limit, we need to observe $\mathbf{r}(t)$, which can be represented by sampled values $\mathbf{r}(k)$, $k = 1, \ldots, N_t$, if the sampling rate is high enough [e.g., 1000 Hz, i.e., the sampling interval is 1 ms (i.e., $T_s = 1$ ms)]. Furthermore, we assume that the stimulus does not move between the sample interval. That is, the stimulus stays in one place in this sampling time and moves to the next one in the next sampling time. Position uncertainty is 2-D (x_k, y_k), so does the velocity uncertainty (v_x and v_y). Assuming that the velocity and orientation of the stimulus are fixed over D_s samples, we have $x_k - x_{k-1} = v_x T_s$, $y_k - y_{k-1} = y_x T_s$. Thus, estimating the position difference between two adjacent samples is equal to estimating the velocity.

A. CRB

The variance of velocity estimation using an unbiased estimator is bounded by the CRB as

$$\sigma_{v_x}^2 \geq \sigma_{\mathrm{CRB},v_x}^2 = \frac{1}{E\left\{\left[\frac{\partial \ln p(\mathbf{r}|B,v_x,v_y)}{\partial v_x}\right]^2\right\}} \tag{17}$$

where $p(\mathbf{r}|B, v_x, v_y)$ denotes the probability of $\mathbf{r}$ given the condition that stimulus B is moving at speed (v_x, v_y), i.e.,

$$\partial \ln p(\mathbf{r}|B, v_x, v_y) = \sum_{k=2}^{D_s} \ln p\left(\mathbf{r}_k|B^{(x_1+v_x kT_s, y_1+v_y kT_s)}\right) - \ln p\left(\mathbf{r}_k|B^{(x_1+(v_x+dv_x)kT_s, y_1+v_y kT_s)}\right) \tag{18}$$

and $p(\mathbf{r}|B^{(x,y)})$ denotes the probability of $\mathbf{r}$ given the condition that stimuli B is located at the position (x, y), i.e.,

$$\begin{aligned}\ln p\left(\mathbf{r}_k|B^{(x,y)}\right) &= \sum_i \ln p\left(r_k^{(i)}|B^{(x,y)}\right) \\ &= \sum_i \left[r_k^{(i)} \ln F_{B^{(x,y)}}^{(x,y)} - F_{B^{(x,y)}}^{(x,y)} - \ln\left(r_k^{(i)}!\right)\right].\end{aligned} \tag{19}$$

Equation (17) shows a lower bound called CRB on the variance of estimation for any unbiased estimator.

The following properties for cases with dark background are based on the CRB. Proofs can be found in Appendices A–C.

Property 1: With dark background, we have

$$\frac{1}{\sigma_{\mathrm{CRB},v_x}^2(N)} = N^2\left[\frac{1}{\sigma_{\mathrm{CRB},v_x}^2(1)} - \chi\right] + N\chi \tag{20}$$

where

$$\chi = \sum_{k=2}^{D_s}\sum_i F_{B^{(x,y)}}^{(x,y)}(1) \ln\left(\frac{dF_{B^{(x,y)}}^{(x,y)}(1)}{dx}\right)^2. \tag{21}$$

By examining $\sigma_{\mathrm{CRB},v_x}^2(1)$, and χ, we conclude that for a small N, $\sigma_{\mathrm{CRB},v_x}^2(N)$ decreases at the rate of N^{-1}, and for a larger N, it decreases as fast as N^{-2}.

Property 2: With dark background, if two dots are widely separated (i.e., $\theta \to \infty$), then $\sigma_{\mathrm{CRB},v_x}^2$ for two dots is equal to $\sigma_{\mathrm{CRB},v_x}^2$ for a single dot, provided that N is a constant.

Clearly, for widely separated dots, we can treat them as one dot with equivalent N. The next property shows that when two dots are close to each other, then position uncertainty $\sigma_{\mathrm{CRB},v_x}^2$ is larger when the moving direction is along the orientation of the dots than when the moving direction is orthogonal to the orientation of the dots.

Property 3: With dark background, if a two-dot stimulus has its orientation along the x axis, then $\sigma_{\mathrm{CRB},v_x\neq 0,v_y=0}^2 \geq \sigma_{\mathrm{CRB},v_y\neq 0,v_x=0}^2$.

This property shows that if the dots are positioned along the direction of stimuli movement, then for a certain speed range, the dots will cancel each other (see Fig. 17). Thus, the estimation accuracy would suffer. If the second detector uses this estimator as a reference, then threshold will increase for those combinations of moving speed, dot separation distance, and moving direction, in which dots are cancelled out. In most of our simulation, we examine the threshold averaging over motion with random direction. Thus, this effect is negligible.

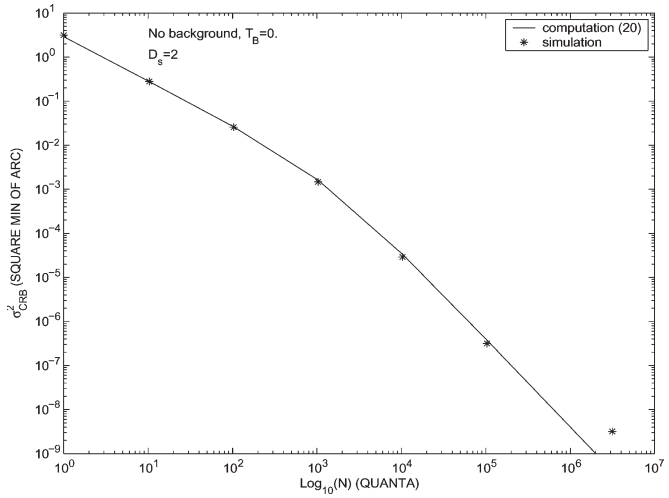


Fig. 13. CRB for movement estimation accuracy. Computation versus simulation.

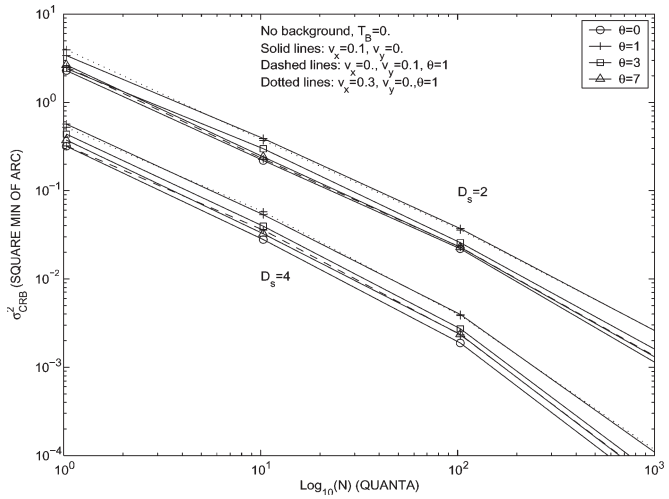


Fig. 14. CRB for movement estimation accuracy for $D_s = 2, 4$, respectively.

In the following study, initially, we assume that we know the position, and then the stimulus moves D_s steps with a constant speed (v_x, v_y) (in minutes of arc per millisecond). Since each step lasts T_s ms, the position difference between speed v_x and $v_x + dv_x$ is $dv_x k T_s$, for $k = 1, 2, \ldots, D_s$. The presented variance is for $T_s = 1$ ms and for $D_s = 2$ and 4, respectively. We fix the total absorbed photons for all D_s cases. The figures show that the variance reduces marginally as D_s increases. This is due to the fact that the larger the value of D_s, the larger the position difference $dv_x D T_s$.

In Fig. 13, we show the computed CRB based on (20), and we also simulated them using (17). An excellent agreement has been achieved. When $N < 10^3$, the CRB decreases at the rate of N^{-1}. When $N > 10^4$, it changes at the rate of N^{-2}. Since we will focus on the cases with small N, we will not run into the $N > 10^4$ range. We further show in Fig. 14 the simulated results for many different situations, which confirm all the properties listed above. One observation worth mentioning is that when $\theta = 1$ arcmin, the CRB of the x axis is twice that of the CRB of the y axis.

Other results are given as follows: 1) CRB has marginal effects by selection of a small value of dv_x (say 0.01–0.1 arcmin/ms), but CRB increases when d_x increases from 0.1 to 0.4 arcmin. In all our simulations, we select $dv_x = 0.1$ arcmin/ms. 2) CRB is largely independent of v_x and v_y. Property 2 discusses only cases in which CRB increases significantly.

B. ZZB

In many applications, it has been shown that the simple CRB is often very loose at the low-SNR regions (i.e., small N regions). In Section V-A, the CRB could be too optimistic for small Ns. At the low-SNR regions, the ZZB often provides a much tighter bound for the maximum-likelihood estimator [25]. Thus, in this subsection, we will examine the ZZB.

A genie-aided method has been widely applied in analyzing the performance of telecommunication systems. In general, there is a large number of factors that can affect the performance of a system. The key concept of the genie-aided method is to construct a simplified system called a genie-aided system in which a genie tells us all factors except one and two. Thus, the performance of the genie-aided system will serve as a lower bound for the real system.

The genie system that we will introduce in this section is the two-dot stimulus moving in one direction with uniformly distributed speeds and a uniformly distributed position uncertainty. This reduces the estimation of multivariables (speeds, positions, and directions) to a two-variable estimation problem. CRB does not depend on prior probability, but ZZB depends on prior probability.

If the speed and position are uniformly distributed over the uncertain region $[0, U_{v_x}]$ and $[0, U_x]$, respectively, then the mean square error of the maximum-likelihood estimator is bounded by

$$\mathbf{a}\Sigma\mathbf{a}^{\mathbf{T}} \geq \sigma_{zzb}^2 = \frac{1}{U_T}\int_0^{U_T} \nu\left\{\max_{a_0 h_{v_x} + a_1 h_x = h} P_{e_{\text{ave}}}\right\} h dh \quad (22)$$

where $\nu\{\}$ denotes the valley-filling function [25], $P_{e_{\text{ave}}} = (1/U_x)\int_{-h_x}^{U_x - h_x} P_e((v_x, x_0), \quad (v_x + h_{v_x}, x_0 + h_x))dx$, $\mathbf{a} = [a_0, a_1]$ is a vector describing the direction of the error estimation probe, and we set $a_0 = a_1 = 1$, Σ is the error covariance in all directions, and $P_e((v_x, x_0), (v_x + h_{v_x}, x_0 + h_x))$ denotes the minimum probability of error for deciding between the actual moving speed v_x and position x_0 and of $v_x + h_{v_x}$ and $x_0 + h_x$, assuming both of them are of equal probability (i.e., unbiased). The minimum probability can be achieved by the maximum-likelihood detection, i.e., the likelihood ratio is

$$L = \frac{p(\mathbf{r}|B, x_0, v_x, v_y)}{p(\mathbf{r}|B, x_0 + h_x, v_x + h_{v_x}, v_y)}. \quad (23)$$

Equivalently, we have log-likelihood

$$\ln(L) = \sum_{k=2}^{D_s}\sum_i r^{(i)} \ln\left(\frac{F^{(i)}_{B^{(x_a, y_a)}}}{F^{(i)}_{B^{(x_c, y_a)}}}\right) \quad (24)$$

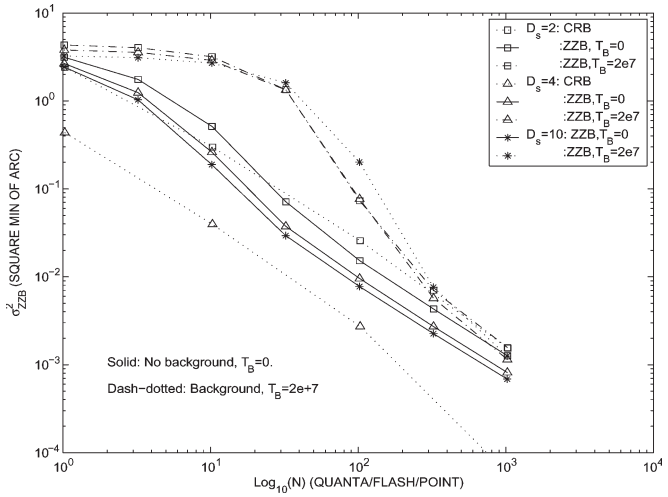


Fig. 15. ZZB for movement estimation accuracy for $D_s = 2$, 4, 10, respectively.

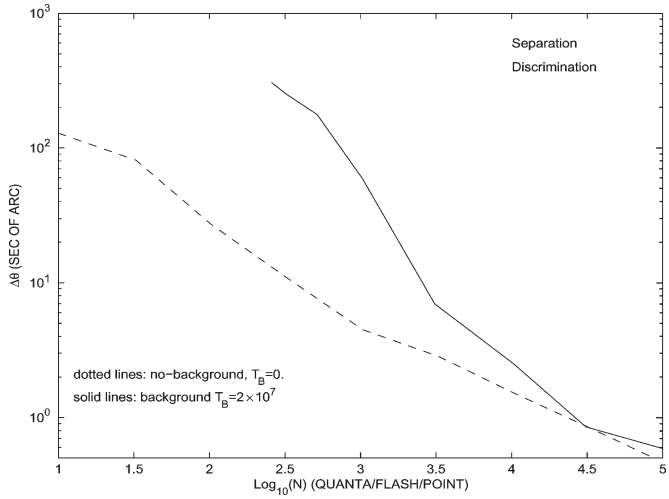


Fig. 16. Hyperacuity threshold under the motion estimation uncertainty.

where $x_a = x_1 + v_x kT_s$, $y_a = y_1 + v_y kT_s$, and $x_c = x_1 + h_x + (v_x + h_{v_x})kT_s$. Since L depends only on the position differences, we have

$$\ln(L) = \sum_{k=2}^{D_s} \sum_i r^{(i)} \ln\left(\frac{F^{(i)}_{B^{(0,0)}}}{F^{(i)}_{B^{(kh_{v_x}+h_x,0)}}}\right). \qquad (25)$$

Exact evaluation of $P_e(v_x, v_x + \Delta)$ is complicated. We can apply Gaussian approximation, i.e., approximate $\ln(L)$ as a Gaussian variable with mean $\mu_j = \sum_{k=2}^{D_s} \sum_i F^{(i)}_{B^{(0,0)}} \times (\ln(F^{(i)}_{B^{(0,0)}}) - \ln(F^{(i)}_{B^{(kh_{v_x}+h_x,0)}}))$ and variance $\sigma_j^2 = \sum_{k=2}^{D_s} \sum_i F^{(i)}_{B^{(0,0)}} (\ln(F^{(i)}_{B^{(0,0)}}) - \ln(F^{(i)}_{B^{(kh_{v_x}+h_x,0)}}))^2$.

In Fig. 15, we show the computed ZZB values for $D_s = 2$, $4, 10$. Compared with CRB, there are two distinguishing features for ZZB, namely: 1) at low N, the variance degrades not as fast as CRB and 2) the variance does not change much as D_s increases.

C. Performance of the Second Detector With Motion Estimation Uncertainty

In this section, we will examine the Vernier accuracy thresholds for the second detector with motion estimation uncertainty. The procedure is given as follows. The total observation T (in seconds) is divided into many subintervals (typically $T = 0.1$ s is divided into 100 subintervals and $T_s = 1$ ms). Assume that the stimulus stays in each place T_s ms and then moves from one to another according to the defined eye movements in Section II. Thus, we can estimate the CRB every D_s frames. Assuming that the estimation error is a Gaussian variable with zero mean and variance of CRB, then we generate position biases in both the x axis and y axis. That is, as in (15), we shift each frame according to the randomly distributed biases. Thus, if we pin a needle through the center of our estimation, then the dot in each frame will be located around the needle, and the distance between the dot and the needle is a Gaussian distribution. Now, our task is to examine how the Vernier accuracy is affected by such an imperfect combination. Since no estimation method can produce a smaller estimation variance than CRB and ZZB, our threshold would be the best possible threshold for any ideal detector with motion estimation uncertainty.

During the simulation, we assume that $D_s = 10$ and $T_s = 1$ ms; thus, the effective quanta per estimation is one-tenth of total quanta per flash. Fig. 16 clearly confirms the result in Fig. 12. That is, at low N, the threshold tends to be unbounded even through the dark background. Such a phenomenon is partly due to the eye or device not being able to track the movement of the stimulus accurately.

In summary, the second detector is much simpler and offers much better performance than the first detector if we can find an estimation algorithm that can achieve the performance of bounds. Therefore, it may be a good candidate for our hyperacuity device.

We believe that the second detector can produce a threshold of less than 1 s if we can estimate the motion achieving the performance bound. Human visual systems may use both (e.g., an array of SDE detectors for position uncertainty [12] and an estimator for motion uncertainty). However, for an engineering device, the second detector could be a better choice.

VI. Simple Analysis Used in Vision Sciences

In vision science, some simple but effective computations are used to evaluate the result. Here, we present a simple computation that may be useful to provide a mathematical explanation for the ideal observer result.

Property 4: For a stationary stimulus, the threshold for 75% correction is $\Delta\theta = 27/(\sqrt{N'/2})$ arcsec, where N' is the number of photons in the narrow part of the receptive field, i.e., $0.42N$. Note that N is this section is the number of photons from each dot absorbed in the cones, whereas, previously, it was the number of photons entering the cornea (matched in size to the pupil). For a moving stimulus, the threshold is unchanged when the moving speed is less than the maximum speed that can be resolved by the oriented motion detector; whereas the threshold would be linearly increased as $(v - 2.25)t$, where t is the temporal resolution for this task.

Proof: Let N be the number of photons absorbed by the cones for each stimulus dot during the stimulus presentation. The point spread function of (1) has 42% of the absorbed photons having a standard deviation of 27 arcsec.

When the target is stationary, the standard error (SE) of the location of the dot is $27/\sqrt{0.21N}$. The exact calculation would take into account the size of the cone, slightly increasing the SE, and the role of the second term in the PSF for the remaining 58% of the photons, slightly decreasing the SE. These corrections are small compared to the dominant term.

The Vernier judgment requires the comparison of the location of two dots so the SE of the Vernier judgment is increased by $\sqrt{2}$. Thus, the predicted Vernier acuity for an ideal observer is the threshold $\Delta\theta = 27/\sqrt{0.21N}$ arcsec. Note that this is the threshold for getting 75% correct in the 2AFC method.

When the target is moving, we assume that the visual system has oriented motion detectors that can effectively stabilize the moving target up to velocities of $2.25^\circ/\text{s}$. Therefore, Vernier thresholds should be the same as the stationary thresholds for velocities below $2.25^\circ/\text{s}$. For larger velocities, the stimulus would be motion-blurred and thresholds will increase linearly. Thus, the motion blur would be $(v - 2.25)t$, where t is the temporal resolution for this task (possibly about 50 ms). In this paper, the worst case assumption was made whereby the motion blur would last the full 100-ms duration of the stimulus. This motion blur has a rectangular profile, so to convert it to a Gaussian SE, one would need to divide the total blur extent by $1/\sqrt{2\pi}$. ■

For $N = 10^5$, we have the threshold $\Delta\theta = 27/\sqrt{0.21N} = 0.18$ s, which is for stationary stimuli without position uncertainty. Compared with numerical results in Figs. 7 and 16 (i.e., those cases in the absence of both eye and stimuli movements and with the largest number of SDE detectors), we can see that the initial position uncertainty nearly doubles the threshold. The rest of the results demonstrate how the performance degrades due to the reduction of the number of SDE detectors, the stimuli motion uncertainty, eye movement, etc.

VII. Conclusion

We examined two detectors for a hyperacuity device. The first is the SDS detector by Geisler and Davilla [12]. The second estimates the uncertainty and then takes out its effect.

In the first method, when the separation distance between the SDE detectors is around 1 arcmin, the SDS detector can achieve nearly optimal performance. To cover motion uncertainty with nearly optimal performance, the SDE detector at each position needs to cover 16 directions, and in each direction, it needs to cover speeds with an increment of $0.5^\circ/\text{s}$. Typically, the SDS detector needs 7776 SDE detectors to deal with a speed up to $2^\circ/\text{s}$ stimuli movement with randomly selected direction and 9×9 min position uncertainty region. The ideal observer described above can achieve a hyperacuity threshold of 2–4 arcsec, which is about three times better than human observers, when the intensity of the stimulus is high and the stimulus is under the influence of fixational movement without head support. Its threshold is almost constant over the range of speeds covered by the SDS detector. When the speed of the stimulus is faster than the maximum reference speed, the threshold gradually increases, initially followed by a sharp upturn. Unless we can find new and inexpensive units, the method would be too complicated to be implemented.

In the second detector, position estimation and motion tracking capability are examined. With perfect position estimation and motion tracking, the SDS detector can be reduced to a single SDE detector, which is tuned to correct position and motion speed and direction. We examine two lower bounds on the estimation variance, i.e., CRB and ZZB. The results showed that with the second method, it is possible to achieve a hyperacuity capability of 1 s or less, which is ten times better than human observers. Therefore, it is a good candidate for our device if we build it using a microprocessor and if we can find an estimator to achieve the performance described by the bounds. Furthermore, in this paper, we assume that the detectors have no knowledge of impulse response of eye movement [i.e., $h(t)$ in (2) and (3)]. Certainly, this knowledge would help detectors to improve performance, which is worthwhile for further exploration in the future.

Many challenges still lie ahead to implement this device. However, as we know in the vision study, if the device uses all possible cues in detection rather than parts of cues as the human visual system does, then better acuity results can be achieved. Furthermore, hyperacuity is one of visual acuity capabilities, and our ultimate goal is to build a device to achieve a wide range of visual acuity that not only can improve the acuity but also may improve the life quality of many visually impaired patients.

Appendix A
Proof of Property 1

Let $X(N)$ denote the X function with the total number of absorbed photons equal to N. Then, we have $E[r^{(i)}(N)] = NE[r^{(i)}(1)]$, $E[(r^{(i)}(N))^2] = N\sigma^2_{r^{(i)}(1)} + N^2E^2[(r^{(i)}(1))]$, and $F^{(x,y)}_{B^{(x,y)}}(N) = NF^{(x,y)}_{B^{(x,y)}}(1)$. Let $x_a = x_1 + v_x kT_s$, $y_a = y_1 + v_y kT_s$, and $x_b = x_1 + (v_x + dv_x)kT_s$. We then have

$$\begin{aligned}
&\partial \ln p(\mathbf{r}|B, v_x, v_y, N) \\
&= \sum_{k=2}^{D_s}\sum_i \left\{ r_k^{(i)}(N) \ln \frac{F^{(x_a,y_a)}_{B^{(x_a,y_a)}}(N)}{F^{(x_b,y_a)}_{B^{(x_b,y_a)}}(N)} \right. \\
&\qquad \left. - F^{(x_a,y_a)}_{B^{(x_a,y_a)}}(N) + F^{(x_b,y_a)}_{B^{(x_b,y_a)}}(N) \right\} \\
&= \sum_{k=2}^{D_s}\sum_i \left\{ r_k^{(i)}(N) \ln \frac{F^{(x_a,y_a)}_{B^{(x_a,y_a)}}(1)}{F^{(x_b,y_a)}_{B^{(x_b,y_a)}}(1)} \right. \\
&\qquad \left. - NF^{(x_a,y_a)}_{B^{(x_a,y_a)}}(1) + NF^{(x_b,y_a)}_{B^{(x_b,y_a)}}(1) \right\} \\
&= \frac{\sum_{k=2}^{D_s}\sum_i \left\{ r_k^{(i)}(N)G_k^{(i)}(1) - NH_k^{(i)}(1) \right\}}{\partial v_x}. \qquad \text{(A1)}
\end{aligned}$$

Substituting this to (17), we obtain

$$E\left\{\left(\frac{\partial \ln p(\mathbf{r}|B,v_x,v_y,N)}{\partial v_x}\right)^2\right\}$$

$$= E\left\{\left(\frac{\sum_{k=2}^{D_s}\sum_i\left\{r_k^{(i)}(N)G_k^{(i)}(1)-NH_k^{(i)}(1)\right\}}{\partial v_x}\right)^2\right\}$$

$$= N^2E\left\{\left(\frac{\sum_{k=2}^{D_s}\sum_i\left\{r_k^{(i)}(N)/NG_k^{(i)}(1)-H_k^{(i)}(1)\right\}}{\partial v_x}\right)^2\right\}$$

$$= N^2\left(E\left\{\left(\frac{\partial \ln p(\mathbf{r}|B,v_x,v_y,1)}{\partial v_x}\right)^2\right\}-\chi\right)+N\chi. \quad \text{(A2)}$$

Thus, the property holds.

Appendix B
Proof of Property 2

Let A denote the single-dot stimulus and B denote the two-dot stimulus. Then, $F_B = F_{A_1} + F_{A_2}$ and $r_B = r_{A_1} + r_{A_2}$, where A_i denotes the value due to the ith dot. When two dots are widely separated, we have $F_{A_1} \times F_{A_2} = 0$ and $r_{A_1} \times r_{A_2} = 0$. That is, for any given cone, the mean absorbed photon must only come from one stimulus dot, not from both dots.

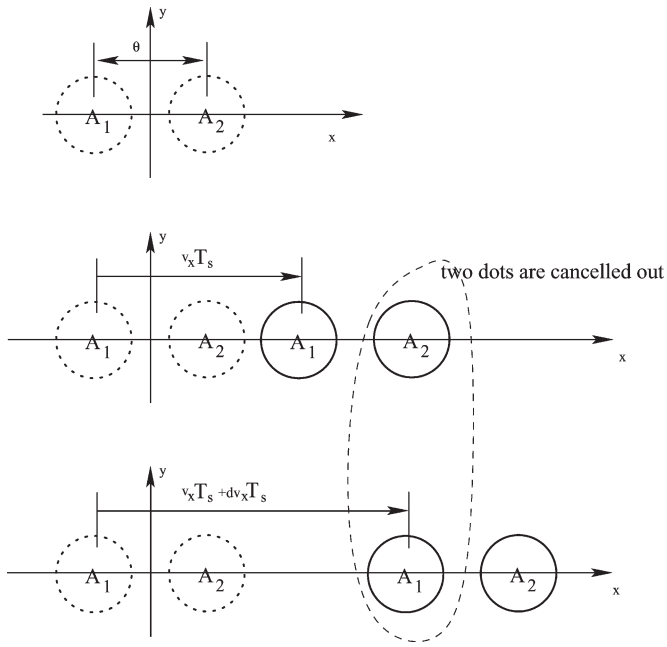


Fig. 17. Cancellation of dots in CRB computation.

Therefore, we have (B1), shown at the bottom of the page. Thus, the property holds.

Appendix C
Proof of Property 3

The proof is simple. If we examine the computation in (17) as illustrated in Fig. 17, then we can see the possible cancellation

$$\begin{aligned}
\frac{1}{\sigma^2_{\mathrm{CRB},v_x}} &= E\left\{\left[\frac{\partial \ln p(\mathbf{r}|B,v_x,v_y)}{\partial v_x}\right]^2\right\} \\
&= E\left\{\left(\frac{\sum_{k=2}^{D_s}\sum_i\left\{r_{k,B}^{(i)}(N)G_{k,B}^{(i)}(1)-NH_{k,B}^{(i)}(1)\right\}}{\partial v_x}\right)^2\right\} \\
&= E\left\{\left(\frac{\sum_{k=2}^{D_s}\sum_i\left\{r_{k,A_1}^{(i)}\left(\frac{N}{2}\right)G_{k,A_1}^{(i)}(1)-\frac{N}{2}H_{k,A_1}^{(i)}(1)\right\}}{\partial v_x}+\frac{\sum_{k=2}^{D_s}\sum_i\left\{r_{k,A_2}^{(i)}\left(\frac{N}{2}\right)G_{k,A_2}^{(i)}(1)-\frac{N}{2}H_{k,A_2}^{(i)}(1)\right\}}{\partial v_x}\right)^2\right\} \\
&= E\left\{\left(\frac{\sum_{k=2}^{D_s}\sum_i\left(r_{k,A_1}^{(i)}\left(\frac{N}{2}\right)+r_{k,A_2}^{(j)}\left(\frac{N}{2}\right)\right)G_{k,A_1}^{(i)}(1)-NH_{k,A_1}^{(i)}(1)}{\partial v_x}\right)^2\right\} \\
&= E\left\{\left(\frac{\sum_{k=2}^{D_s}\sum_i\left(r_{k,A_1}^{(i)}(N)\right)G_{k,A_1}^{(i)}(1)-NH_{k,A_1}^{(i)}(1)}{\partial v_x}\right)^2\right\} \\
&= E\left\{\left[\frac{\partial \ln p(\mathbf{r}|A,v_x,v_y)}{\partial v_x}\right]^2\right\} \quad \text{(B1)}
\end{aligned}$$

between dots when the stimulus is moving along the x axis only (i.e., A_2 when the speed is v_x, and A_1 when the speed is $v_x + dv_x$). Thus, for the worst case in which two dots are completely overlapped, the CRB computation is only based on the single dot with half of N. That is, the variance could be four times worse than the single-dot case. When the stimulus is moving along the y axis only, there will be no cancellation between dots. Thus, the property holds.

REFERENCES

[1] S. A. Klein and D. M. Levi, "Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 7, no. 2, pp. 1170–1190, Jul. 1985.

[2] H. Wang and D. M. Levi, "Spatial integration in position acuity," *Vis. Res.*, vol. 34, no. 21, pp. 2859–2877, Nov. 1994.

[3] D. R. Williams and N. J. Coletta, "Cone spacing and the visual resolution limit," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, no. 8, pp. 1514–1523, Aug. 1987.

[4] G. Westheimer, "Visual acuity and hyperacuity," *Invest. Ophthalmol.*, vol. 14, no. 8, pp. 570–572, Aug. 1975.

[5] I. Hadani, A. Z. Meiri, and M. Guri, "The effects of exposure duration and luminance on the 3-dot hyperacuity task," *Vis. Res.*, vol. 24, no. 8, pp. 871–874, 1984.

[6] S. Waugh and D. M. Levi, "Visibility, timing and Vernier acuity," *Vis. Res.*, vol. 33, no. 4, pp. 505–526, Mar. 1993.

[7] G. Westheimer and S. P. McKee, "Visual acuity in the presence of retinal-image motion," *J. Opt. Soc. Amer.*, vol. 65, no. 7, pp. 847–850, Jul. 1975.

[8] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. New York: Wiley, 1960.

[9] J. Nachmias, "Signal detection theory and its application to problems in vision," in *Handbook of Sensory Physiology*, vol. VII/4, H. Autrum and Alpern *et al.*, Eds. New York: Springer-Verlag, 1972, pp. 56–77.

[10] R. Laming, "On the limits of visual detection; signal detection theory; Theoretical basis of the processing of simple visual stimuli; contrast sensitivity," in *Limits of Vision*, J. J. Kulikowski *et al.*, Eds. Boca Raton, FL: CRC, 1991, pp. 6–43.

[11] W. S. Geisler, "Physical limits of acuity and hyperacuity," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 1, no. 7, pp. 775–782, Jul. 1984.

[12] W. S. Geisler and K. D. Davilla, "Ideal discriminations in spatial vision: Two point stimuli," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 2, no. 9, pp. 1483–1497, Sep. 1985.

[13] H. Wilson, "Response of spatial mechanisms can explain hyperacuity," *Vis. Res.*, vol. 26, no. 3, pp. 453–469, 1986.

[14] E. Kowler, "The stability of gaze and it implications of vision," in *Eye Movements*, R. H. S. Carpenter, Ed. Boca Raton, FL: CRC, 1991, pp. 71–92.

[15] F. Ratliff and L. A. Riggis, "Involuntary motions of the eye during monocular fixation," *J. Exp. Psychol.*, vol. 40, no. 6, pp. 687–701, Dec. 1950.

[16] L. A. Riggis and J. C. Armington, "Motions of the retinal image during fixation," *J. Opt. Soc. Amer.*, vol. 44, no. 4, pp. 315–321, Apr. 1954.

[17] J. Krauskopf, T. N. Cornsweet, and L. A. Riggis, "Analysis of eye movements during monocular and binocular fixation," *J. Opt. Soc. Amer.*, vol. 50, no. 6, pp. 572–578, Jun. 1960.

[18] M. Eizenman, P. E. Hallett, and R. C. Frecker, "Power spectra for ocular drift and tremor," *Vis. Res.*, vol. 25, no. 11, pp. 1635–1640, 1985.

[19] S. Martinez-Conde, S. L. Macknik, and D. H. Hubel, "The role of fixational eye movements in visual perception," *Nat. Rev. Neurosci.*, vol. 5, no. 3, pp. 229–240, Mar. 2004.

[20] R. Engbert and R. Kliegl, "Microsaccades keep the eyes' balance during fixation," *Psychol. Sci.*, vol. 15, no. 6, pp. 431–436, Jun. 2004.

[21] F. W. Campbell and R. W. Gubisch, "Optical quality of the human eye," *J. Physiol. London*, vol. 186, no. 3, pp. 558–578, Oct. 1966.

[22] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.

[23] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 3, pp. 386–391, May 1969.

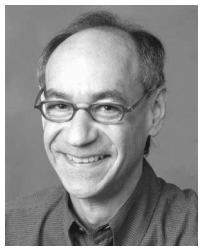
[24] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bounds on signal parameter estimation," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 1, pp. 90–93, Jan. 1975.

[25] K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Tree, "Extended Ziv–Zakai lower bound for vector parameter estimation," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 624–637, Mar. 1997.

Lei Wei (S'93–M'95–SM'99) received the M.E. degree from the University of New South Wales, Sydney, Australia, in 1993, and the Ph.D. degree from the University of South Australia, Adelaide, in 1995, both in electrical engineering.

He is an Associate Professor at the School of Electrical Engineering and Computer Science, University of Central Florida, Orlando. His research interests include error control coding, ad hoc network, wireless communications, and biologically inspired signal processing.

Dr. Wei was a member of the Organization Committee for many international conferences and an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1999 to 2003.



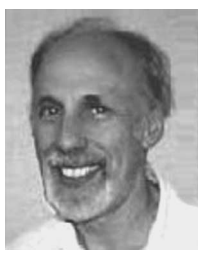
Dennis M. Levi received the O.D. and Ph.D. degrees from the University of Houston, Houston, TX.

He is a Professor of optometry and vision science and the Dean of the School of Optometry, University of California, Berkeley. His research on amblyopia and spatial vision has been continuously funded by the National Eye Institute for almost 30 years. He has published over 200 scientific papers. He is the Editor-In-Chief of *Vision Research* and has served on a number of editorial boards (e.g., *Journal of Vision*, *Ophthalmic and Physiological Optics*, and *Investigative Ophthalmology and Visual Science*).

Dr. Levi was elected Fellow of the Optical Society of America for his research contributions in the areas of amblyopia and spatial vision. He has received numerous awards including the Glenn Fry Award and the Garland Clay Award from the American Academy of Optometry.

Roger W. Li received the B.Sc. degree in optometry and the Ph.D. degree from the Hong Kong Polytechnic University, Kowloon, in 1995 and 1999, respectively.

He spent the following two years with the Centre of Myopia Research, Hong Kong Polytechnic University, as a Research Associate. He was a Clinical Instructor in the Optometry Clinic, Hong Kong Polytechnic University, from 1995 to 2000. In 2001, he moved to the College of Optometry, University of Houston, TX, to continue his postdoctoral research training in the area of spatial vision. Since 2002, he has joined the School of Optometry, University of California, Berkeley, as a Postdoctoral Research Associate, and later Associate Research Specialist. He has numerous publications in leading scientific and professional journals. He has a wide range of research interests including amblyopia treatment, perceptual learning, spatial and pattern vision, clinical visual psychophysics and electrophysiology, visual aging, myopic control, and ocular accommodation.



Stanley A. Klein received the Ph.D. degree from Brandeis University, Waltham, MA.

He is a Professor of optometry and vision science and bioengineering. He is the Chair of the Vision Science Graduate Program and the Chair of the Program Advisory Committee of the Center for Adaptive Optics, University of California at Berkeley. His earlier research was in the field of theoretical particle physics (early string theory), and his present research is on spatial vision and electroencephalogram source localization.

Prof. Klein was elected Fellow of the Optical Society of America for his research contributions in spatial vision.