# Temporal mechanisms underlying flicker detection and identification for red–green and achromatic stimuli

**Andrew B. Metha and Kathy T. Mullen**

*McGill Vision Research, Department of Ophthalmology (H4-14), McGill University, 687 Pine Avenue West,*
*Montreal, Quebec H3A 1A1, Canada*

We have simultaneously measured detection and temporal frequency identification for both red–green isoluminant and achromatic stimuli over a range of temporal frequencies for two observers. Results show that temporal frequency identification can be made along the temporal frequency dimension for both red–green and achromatic stimuli at contrasts close to detection threshold. In general, temporal frequency identification was better for the achromatic than for the red–green stimuli; however, the level of chromatic identification performance was still sufficient to permit us to reject the notion that the red–green mechanism embodies a single temporal filter. We have developed a model based on signal detection theory that assumes that detection and identification both depend on the properties of the temporal filters underlying each mechanism. From this we have derived putative underlying shapes and sensitivities for the temporal filters of the red–green and achromatic mechanisms that comprise a low-pass and a bandpass filter for red–green color vision and two bandpass filters for luminance vision. Finally, we suggest that the relative perceived slowing of isoluminant stimuli may be accounted for by a common motion analysis subserved by different front-end temporal filters for red–green and achromatic motion signals. © 1996 Optical Society of America.
*Key words:* color, luminance, isoluminance, temporal frequency, speed, detection, identification, model.

## 1. INTRODUCTION

There have been a number of attempts to deduce the temporal impulse response for the red–green (RG) color-opponent mechanism of human vision by single-pulse[1] or double-pulse[2–5] methods. From these studies both monophasic and biphasic impulse response functions (IRF's) have been inferred for color vision, the differences between them being attributed principally to different levels of light adaptation and to other details of the experimental methods employed.[5] However, despite differences in reported impulse response shapes, these investigations all assume that for red–green color detection a single chromatic IRF is based on the existence of a unitary chromatic filter that is tuned for temporal frequency (TF). Whereas such a filter may explain the shape of the temporal IRF for one- and two-pulse detection data, the assumption of a single chromatic TF-tuned filter is dubious for two reasons:

First, many of these studies have also assumed a single TF filter for luminance vision, but experiments on TF detection and identification, temporal masking, and subthreshold summation have demonstrated the existence of a small number of temporally tuned mechanisms that subserve the temporal contrast sensitivity function,[6–14] albeit with some debate concerning the exact shape and number of filters.[15] The outputs of these filters have been used to model luminance TF discrimination, and it has further been suggested that they serve as the basis for metric of velocity coding.[16–19] The number and the bandwidth of the temporal filters that subserve color vision has yet to be addressed.

Second, in a single filter system, changes in filter output that are due to the shifts in stimulus TF would be confounded with changes that are due to contrast. Thus a single univariant mechanism would not permit TF discrimination or the identification of stimuli moving at different speeds, in the absence of other cues. To code TF independently requires at least two filters tuned to overlapping TF ranges.[19,20] There are no reports of RG chromatic TF discrimination in the literature. However, although it was originally thought that color information provided no useful cues for motion processing, many subsequent experiments have shown that color vision does seem to support at least some degree of motion perception (for a review, see Ref. 21) and can also discriminate among different velocities.[22] It is thus enticing to investigate the proposal of multiple underlying temporal filters for color vision.

In this study we address the question of whether the RG chromatic system can support the discrimination of TF at and close to detection threshold. We simultaneously measure detection and TF identification thresholds for RG and achromatic stimuli and develop a model to derive from these data the likely bandwidths and sensitivities of the temporal filters that subserve color vision. Our data show that TF discrimination is possible close to detection threshold for chromatic stimuli, and we have therefore used two temporal filters in the model, representing the minimum number required for TF and contrast disambiguation. Based on computational considerations of efficient signal processing and not on any *a priori* physiological result, we further suppose that these filters share the property of temporal orthogonality. We

assume that detection performance is based on probability summation over time and that the upper envelope of underlying filter sensitivity must correspond to the measured temporal contrast sensitivity functions. The simultaneously measured identification thresholds further constrain the underlying filter shapes. From this model we derive a set of temporal IRF's that subserve the RG and achromatic mechanisms. The implications of the model for how these temporal filters can be used in flicker and speed perception are discussed.

## 2.    METHODS

### A.    Stimuli

The stimuli were cardinal RG or achromatic defined horizontal Gabor patches. They were generated by a digital waveform generator (CRS VSG 2/2) and were presented on a red–green–blue (RGB) monitor (Barco Calibrator CCID 7751) running at a frame rate of 120 Hz. The mean chromaticity (1931 CIE: $x = 0.3377$, $y = 0.3184$) and luminance (62.2 cd/m$^2$) of the display were not altered by the presentation of the stimuli, which had a spatial frequency of 0.25 cycles/degrees, spatially windowed by a Gaussian with a standard deviation (SD) of 4°. The spatial window was truncated at $\pm2$ SD (16° diameter). The vertical profile of the stimuli had 12-bit resolution, and the horizontal profile was generated by use of frame-by-frame dynamic dithering of 12 statistically independent 1-bit Gaussian masks (pixel size $0.54 \times 0.54$ mm).[23] The screen display size was 35.4 cm $\times$ 26.2 cm. Calibration and gamma correction were achieved by methods described by Metha et al.[24]

In the two-interval forced-choice (2-IFC) tasks, the stimuli were presented sequentially in 1.0-s intervals with contrast ramped on and off according to a raised cosine envelope and separated by 500 ms. The stimuli were counterphase flickered by sinusoidal contrast modulation. The flicker rates used for the luminance-defined stimuli were 1.0–32.0 Hz in octave steps. For the RG-defined stimuli the same TF's were used, except that the 32.0-Hz stimulus was replaced by a 22.6-Hz stimulus.

### B.    Subjects

The two authors served as subjects. Testing was performed monocularly at a distance of 90 cm from the CRT face under dim ambient room illumination. A small fixation marker (2-mm-diameter black spot) centered on the screen aided fixation. Subjects wore their prescribed corrective spectacles if necessary. Both subjects had normal color vision.

### C.    Psychophysical Procedures and Analysis

#### 1.    Determination of Red–Green Isoluminance

The cone contrast weights to the luminance postreceptoral mechanism are not fixed but depend on the observer, the state of chromatic adaptation, and the TF content of the stimulus,[23,25,26] and thus we determined the luminance cone contrast ratio for each TF tested for each observer. This measurement was made by a minimum motion paradigm, by a method of adjustment.[27] A stimulus that equally modulated L- and M-cone contrast in spatial antiphase at a fixed surpathreshold contrast, drifting at

the test TF, was continuously presented. To this stimulus we superimposed an L-and-M-cone in-phase stimulus, the contrast of which was controllable by the observer using a computer mouse. Observers adjusted the contrast of this luminance stimulus to obtain the minimum motion, and the mean of 10 settings was obtained. The resulting stimulus defines the RG isoluminant direction.

Figure 1 shows the L:M ratio cone contrast contribution ratio of the luminance mechanism as measured by the above technique for both observers as a function of TF. The error bars represent the SD's of 10 minimum motion settings. The L:M ratio is found to vary significantly for both observers as a function of TF and, interestingly, in opposite directions. These findings are in line with previous reports that isoluminance settings can vary markedly among individuals[28] and also depend on temporal frequency.[23,25,26,29]

The cardinal direction for the luminance mechanism was assumed to be the achromatic stimulus direction, which modulates L-, M-, and S-cone contrasts equally and in phase. Because the RG isoluminant direction changes with TF, to relate threshold measurements across TF's we base the reported cone contrast sensitivities on the projections of the cardinal stimuli onto the inferred achromatic and RG mechanism directions (the direction in cone contrast space that optimally excites each postreceptoral mechanism[30]).

#### 2.    Detection and Identification Thresholds

We measured detection and identification performance simultaneously for pairs of different TF stimuli as a function of contrast, using a $2 \times 2$-IFC procedure. In any session one temporal frequency from the set (TF1) was paired with another (TF2), and simultaneous detection and TF identification judgments were made following each two-interval trial, at five different contrast levels. This was done for 15 paired combinations of TF1 and TF2, separately for both RG and achromatic cardinal directions. To acquaint the observer with the appearance of stimuli close to threshold and to establish approximate detection thresholds, we performed simple interleaved staircases for each TF pair immediately beforehand.
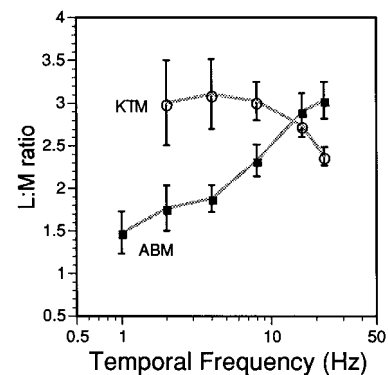


Fig. 1.    RG isoluminance ratios as a function of TF. Each point represents the mean of 10 minimum motion settings in the L:M plane of cone contrast space. The error bars represent 1 SD. The filled squares are the results for observer ABM, and the open circles are for observer KTM.
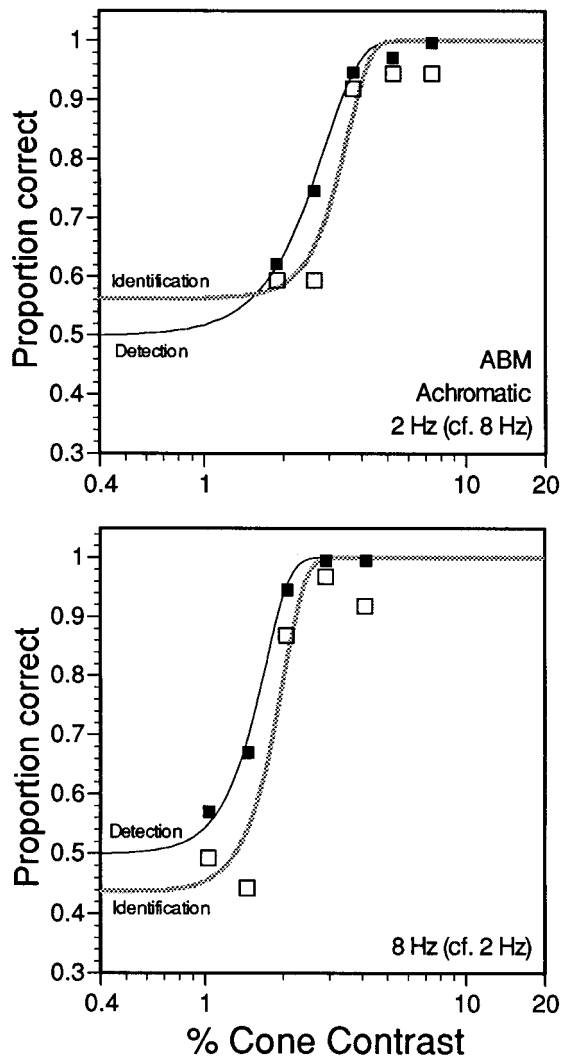
Fig. 2. Example of the four psychometric functions generated by each 2 × 2-IFC experimental block. The filled squares represent detection performance (Question 1) for the 2-Hz (top) and the 8-Hz (bottom) stimuli, which were presented 40 times at five contrast levels in random order spanning predetermined threshold ranges. Detection psychometric functions were fitted by use of Weibull functions constrained to give chance performance at low contrast levels. Weighting was applied by binomial SD estimates. Open squares represent identification performance (Question 2). Because identification performance can be subjectively biased, Weibull functions fitted to these data were constrained by complementary guess rates, as explained in the text.

These staircase-derived thresholds were used to place five contrast levels for each TF, spanning detection threshold in 0.15-log-unit steps. Because the contrasts used in the experiments straddle detection threshold for both TF's, contrast itself cannot be used to aid in identifying which stimulus was presented in any one 2-IFC trial.

For each stimulus at each contrast level, 40 2-IFC trials were randomly presented. The observer was asked to indicate by a button press, first in which interval the stimulus appeared (detection task) and second whether the stimulus was the faster or the slower of the pair under consideration (identification task). Feedback was given after each response, so the observer could maintain an updated impression of each TF during the task. The inter trial interval was 500 ms, and in each session the 400 trials (2 TF's × 5 contrast levels × 40 trials, requiring 800 responses) typically took ~30 min.

### 3. Fitting the Psychometric Functions

The detection and identification experiments yield four psychometric functions per pair of temporal frequencies. An example is shown in Fig. 2, which shows performance for detecting and identifying which of a 2- or an 8-Hz achromatic stimulus was presented, as a function of contrast. The detection psychometric functions are each fitted with two-parameter Weibull functions (base-2) constrained to asymptote at 50% levels for low contrasts. The parameters reflect the threshold contrast at 75% correct performance and the slope of the psychometric function ($\beta$). However, the identification question is susceptible to bias. For example, even at low contrast levels that yield chance detection performance, Fig. 2 shows that the 2-Hz stimulus was reported more often than the 8-Hz stimulus, even though both were presented randomly at equal rates. The difference in guess rate (or bias) can be compensated for because the guess rates must sum to unity. The identification psychometric functions were therefore each fitted with Weibull functions that were constrained to asymptote at complementary guess rates for low contrasts. The threshold parameter thus specifies the contrast at which performance reaches a midlevel between guessing and perfect performance, and, when they are adjusted for 50% guess rates, the detection and identification thresholds both refer to the same level of performance in each task.

Psychometric functions were fitted simultaneously by a least-chi-squared procedure employing the estimated binomial SD's at each point as weight factors. This resulted in determination of nine parameters with associated SD's for each TF pair (four threshold estimates, four psychometric function slopes, and the identification bias rate). We then used the fitted threshold parameters and their estimated SD's to determine the multiplicative increase in contrast above detection threshold, permitting threshold identification performance (which we call factor alpha) and an estimate of the SD for this measure.

## 3. RESULTS

Detection threshold data for all 15 paired combinations of TF's results in 5 estimates at each frequency. Figures 3(a) and 3(b) show the average and the SD of the detection thresholds for achromatic and RG mechanisms, respectively, plotted as cone contrast sensitivity versus TF on log–log axes. Both of the achromatic temporal contrast sensitivity functions display the typical bandpass shape for this low spatial frequency, with sensitivity peaking at ~8 Hz.[30] Likewise, Fig. 3(b) shows the detection results for stimuli isolating the RG mechanism that display a characteristic low-pass shape.[32,33] The use of cone contrast allows the sensitivity of the luminance and RG mechanisms to be compared directly. For stimuli below ~8 Hz the RG mechanism is more sensitive than the luminance mechanism under these conditions. Although the extrapolated TF cutoff is only slightly lower for the

RG than for the luminance mechanism, because of the high degree of overlap in L- and M-cone spectral sensitivity RG performance much above 22.6 Hz cannot be realized. The thicker curves in Fig. 3 represent the resulting model fits (described in detail in Section 4) for detection performance, based on the inferred underlying filters whose modulation transfer functions (MTF's) are shown by the thinner curved lines and whose impulse response profiles are shown in the insets.

As the temporal frequencies of two stimuli (TF1 and TF2) get closer together, they begin to appear more similar, and so their correct identification becomes increasingly difficult at detection threshold levels. However, it is generally found that at suprathreshold contrast levels correct classification of the two TF's can again be made. The ratio of threshold contrast for identification/detection performance (factor alpha) is a measure of how distinguishable TF1 and TF2 are at low contrast levels. If cor-



(a)                                                                    (b)
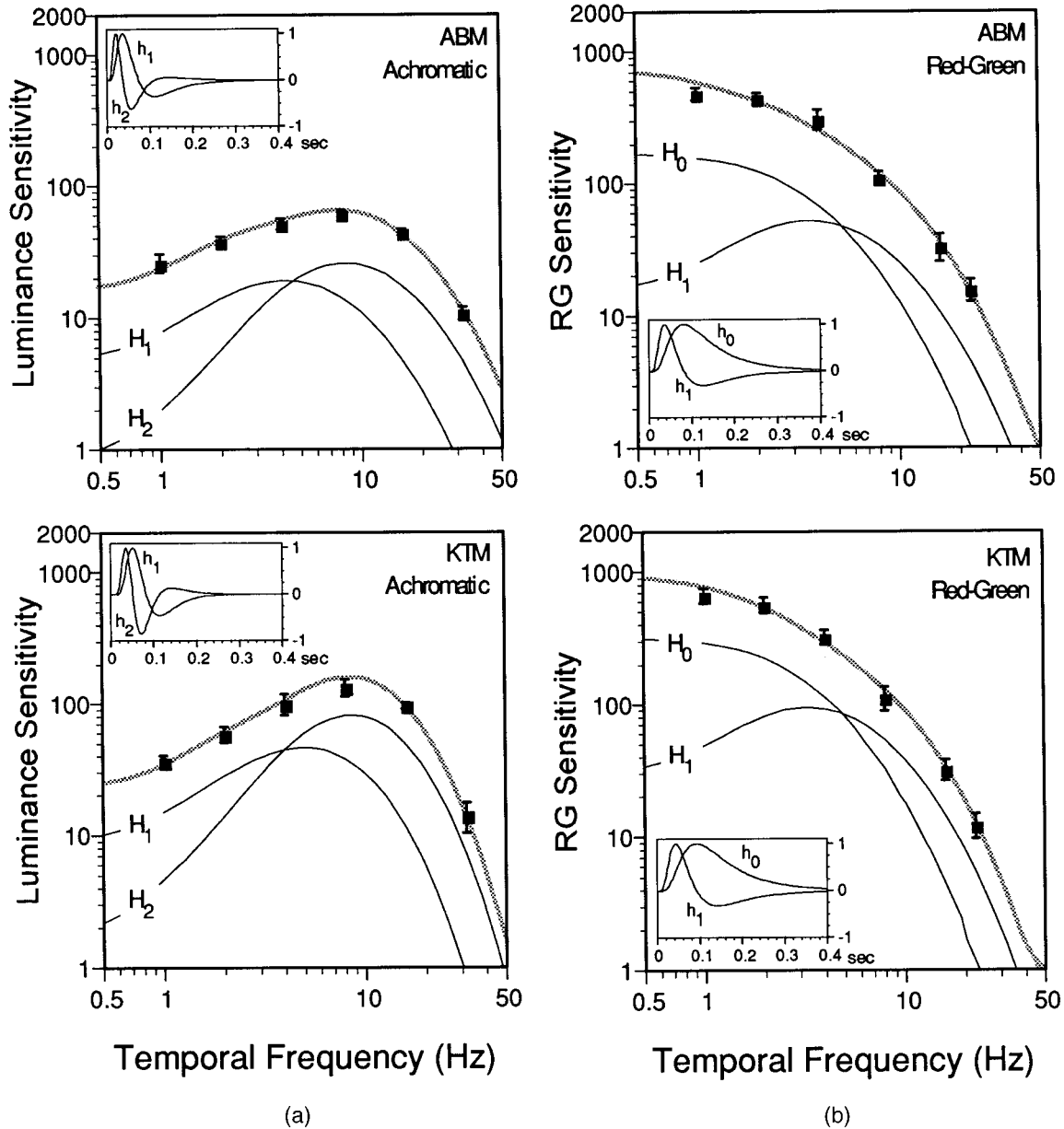
Fig. 3.   Temporal cone contrast sensitivity functions for both observers for (a) achromatic and (b) RG isoluminant stimuli.   The filled squares represent the mean and the SD of the five detection thresholds determined for each temporal frequency during the $2 \times 2$-IFC comparison sessions.   The thicker, lighter curves represent the model predictions for detection performance after parameters were adjusted to give the best fit for both detection and identification data.   The curves labeled $H_1$ and $H_2$ in (a) and $H_0$ and $H_1$ in (b) are the MTF's of the inferred filters underlying the luminance and the RG mechanisms, respectively.   The normalized IRF's of the best-fitting filters are shown as an inset in each case; the filter gain factors have not been applied (refer to the text and to Table 1 for details). Figure 1 shows that the luminance mechanism receives varying L- and M-cone contrast input as a function of TF; therefore luminance sensitivity is given here as the reciprocal (in cone contrast units) of the projection magnitude of the threshold achromatic cardinal stimuli onto the measured luminance mechanism for each TF.   RG sensitivity is given as the reciprocal of the projection magnitude of the threshold isoluminant stimuli onto the RG mechanism, which we assume receives fixed (equal and opposite) L- and M-cone contrast input at all TF's.
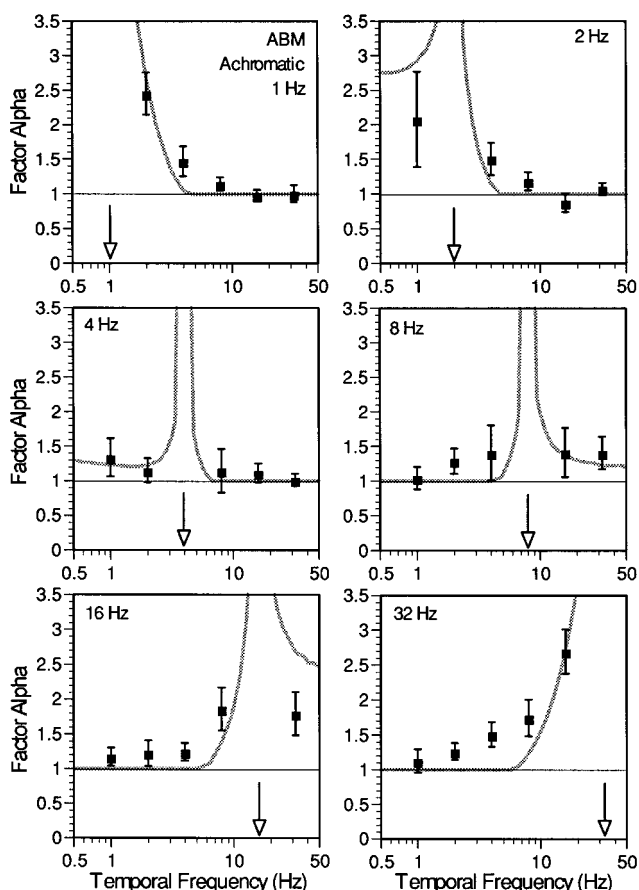
Fig. 4.   Identification performance among achromatic TF pairs for observer ABM.   For conditions comparing the TF labeled and marked by an arrow in each panel (TF2), the symbols represent alpha (the multiplicative factor by which contrast must be raised above detection threshold to permit 75% correct identification) for each compared frequency (TF1).   The error bars are SD estimates derived from the detection and identification psychometric function fits.   The thicker, lighter curves represent the model prediction after parameters were adjusted to give the best fit simultaneously for both the detection and the identification data, resulting in the filter shapes shown in the insets of Fig. 3.

As expected for achromatic stimuli, alpha approaches unity for different frequencies and increases when TF's are closer together.   The data for both observers show that correct identification can occur at detection threshold for each TF tested, as long as the comparison TF was sufficiently different.   An interesting case occurs for the 4-Hz (and to a lesser extent for the 8-Hz) comparison TF whereby identification performance for both higher and lower TF's approaches threshold levels at detection threshold contrasts.   This is particularly true for observer KTM; for example, in sessions containing the 4-Hz stimulus all other tested TF's could be correctly identified as often as they could be detected.   This finding has implications for later modeling, especially concerning the number of underlying filters involved.

Alpha for RG chromatic stimulus pairs is shown in Figs. 6 and 7 for each observer.   These show the same general trends as for the achromatic stimuli in that alpha is high when TF1 and TF2 are close together and that this factor decreases when the difference between TF1 and TF2 increases.   The data show that TF identification is possible among RG stimuli at contrasts approaching detection threshold.   Overall, performance at detection threshold levels was slightly better for observer KTM, a trend also evident in the achromatic data.   The thicker, lighter curves in Figs. 4–7 represent the results of a two-filter model capable of predicting both detection and identification performance, which is discussed next.

rect identification is possible at detection threshold, alpha is unity.   An alpha value of 2.0 for TF1 means that TF1 contrast must be raised to twice its 75% detection threshold level before it can be correctly identified (75% of the time) as TF1 and not TF2.

Figures 4 and 5 show identification performance for observers ABM and KTM, respectively, for achromatic stimulus pairs.   The filled squares indicate alpha for each frequency (TF1) given on the abscissa when paired in sessions containing TF2 (marked by arrows and indicated at the top of each panel).   The error bars are SD estimates derived from the parameters of the detection and identification psychometric function fits.   Alpha values were not averaged across conditions: i.e., alpha for a 1-Hz stimulus with a comparison TF of 2 Hz is plotted independently of the alpha that arises when TF1 is 2 Hz and TF2 is 1 Hz, resulting in the 30 data points shown.   These factors are not always the same, and this point is taken up later in the discussion.
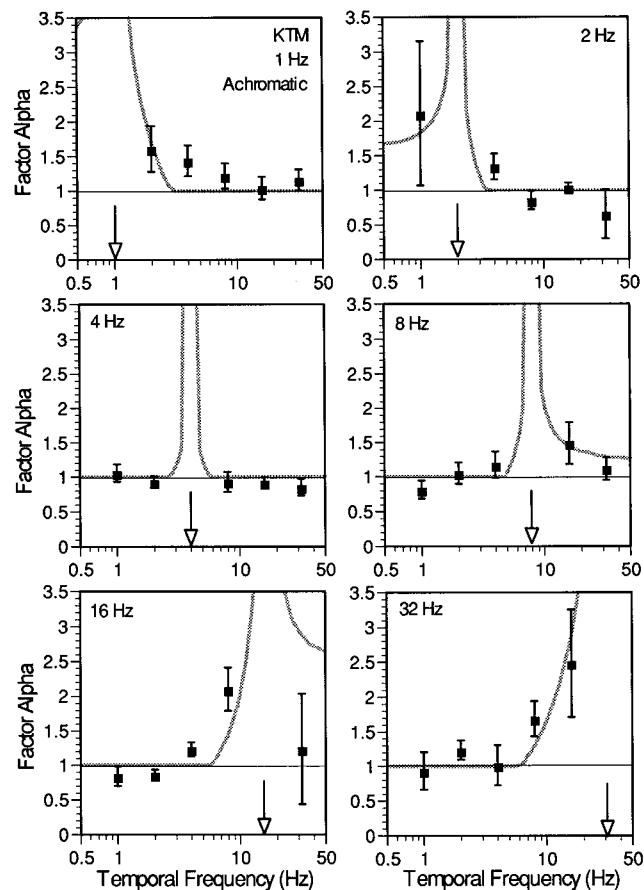


Fig. 5.   Identification performance among achromatic TF pairs for observer KTM.   Other details are the same as for Fig. 4.
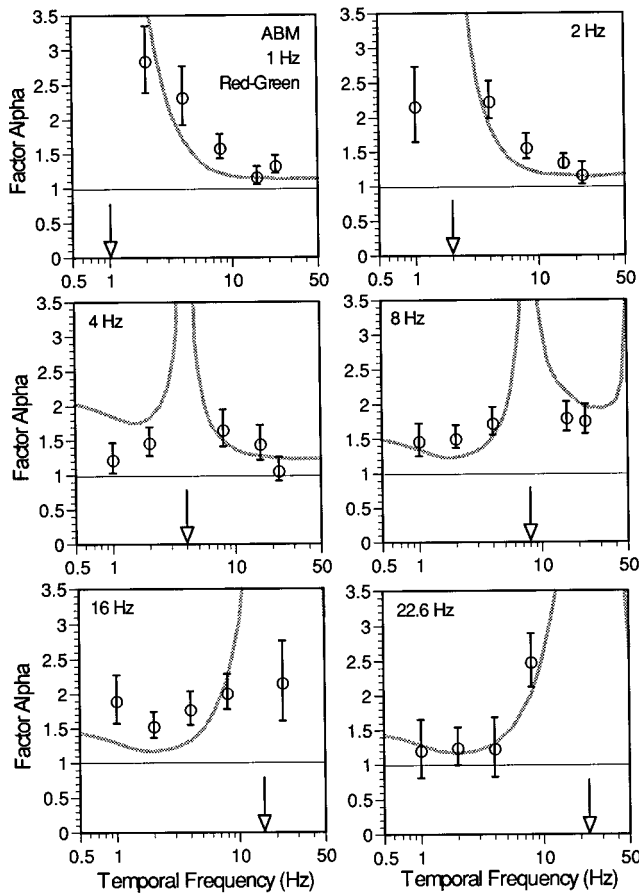
Fig. 6.   Identification performance among RG TF pairs for observer ABM.   Other details are the same as for Fig. 4.

## 4.  MODEL

The model that we propose considers the tasks of detection and TF identification as two separate processes, both drawing on information derived from the outputs of two independent TF-tuned linear filters.

Our results show that TF's can be distinguished close to detection threshold in the absence of contrast cues for both RG and achromatic stimuli.   This argues against the operation of a unitary temporal filter as the basis of all temporal processing, as such a system would confound TF and contrast.   The most parsimonious way to model these results is to propose that both detection and TF identification performance rely on the output of at least two independent linear temporal filters.   The property of independence is not an absolute requirement but is based on considerations of information processing theory to ensure that the principle of minimum redundancy is upheld as well as allowing analytical solutions for calculations involving probability summation over time.

Each temporal filter is fully described by its temporal impulse response function $h(t)$.   The proposed filters come from the set of log Gaussians and their temporal derivatives, as suggested by Koenderink[34] and used by others because of their inherent causality and for other theoretical reasons.[34–36]   Any temporal IRF and its first temporal derivative are orthogonal and will therefore give

statistically independent responses to random white noise.   By choosing two such filters as our basis set we maintain the independence requirement, although it must be noted that this does not mean that the filters' MTF's do not overlap because these profiles represent only magnitude spectra; the temporal phase information is not displayed.

### A.  Filter Form and Response
A log Gaussian temporal impulse response has the form

$$h_0(t) = A_0 \, \exp\left[ -\left( \frac{\ln(t/\tau)}{\sigma} \right)^2 \right], \qquad (1)$$

where $\tau$ and $\sigma$ are parameters that determine the peak position and width of the IRF and $A_0$ is a scaling factor that controls the sensitivity of the filter.   The subscript zero denotes that this is the generator function, and the first and the second temporal derivatives, $h_1$ and $h_2$, are described by

$$h_1(t) = 2A_1 \left[ \frac{\ln(t/\tau)}{t\sigma^2} \right] \exp\left[ -\left( \frac{\ln(t/\tau)}{\sigma} \right)^2 \right], \qquad (2)$$
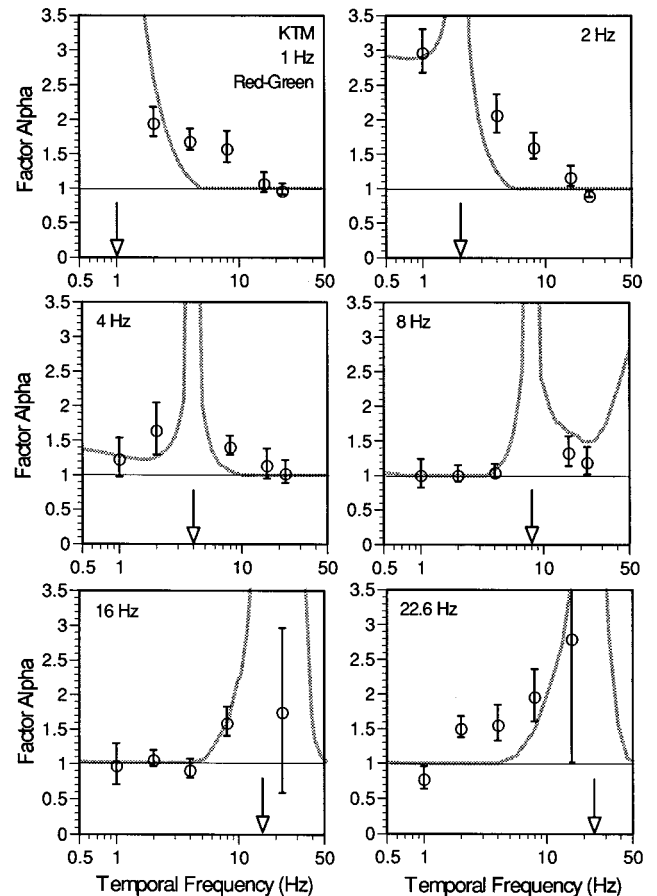


Fig. 7.   Identification performance among RG TF pairs for observer KTM.   Other details are the same as for Fig. 4.

$$h_2(t) = A_2 \left\{ \frac{-2}{(t^2\sigma^2)} + \frac{2\ln(t/\tau)}{(t^2\sigma^2)} + \frac{4[\ln(t/\tau)]^2}{(t^2\sigma^4)} \right\}$$

$$\times \exp\left\{ -\left[ \frac{\ln(t/\tau)}{\sigma} \right]^2 \right\}, \tag{3}$$

where $A_1$ and $A_2$ control the sensitivities of the filters. Note that, apart from the sensitivity factors, the whole family of IRF profiles is governed by just two parameters, $\tau$ and $\sigma$.

The Gabor stimuli have a normalized temporal waveform given by

$$g(t, f) = [1/2 - 1/2\cos(2\pi t)]\cos[2\pi f(t - 1/2)], \tag{4}$$

where $f$ is the stimulus temporal frequency in hertz and $t$ represents time in seconds.

Following Watson,[37] it is convenient to describe the stimulus contrast waveform $C(t)$ as the product of $g(t, f)$ and a positive constant $C$ equal to the peak contrast of the waveform. According to linear systems theory, the response of each filter $i$, as a function of time $t$ and stimulus frequency $f$, will be given by

$$R_i(t, f) = Cg(t, f) * h_i(t), \tag{5}$$

where the asterisk denotes the convolution operation. Thus for any given stimulus of the form $g(t, f)$ and contrast $C$ we can calculate the output response $R$ of each filter over time.

## B. Detection Performance
We accomplish the detection task by separately monitoring these filters over time according to the probability summation model formalized by Watson.[37,38] A detection response occurs when the instantaneous response of either filter is above some independently noise-perturbed criterion level for each filter. The noise can be considered to be introduced to the response signals themselves or, equivalently, to the criterion level. This fluctuation is also considered to be sufficiently rapid that the noise is statistically independent from instant to instant. Further, assuming that the probability associated with a detection outcome as a function of absolute filter response follows a Weibull distribution, one can make calculations to determine the contrast at which stimulus detection occurs, incorporating the effects of probability summation over time and over the two independent filters.

Formally, it is assumed that the probability of detection that is due to filter $i$ at any one instant of time $t$ is given by a Weibull function:

$$p_i(t) = 1 - 2^{(-|R_i(t)|^\beta)}. \tag{6}$$

A common parameter $\beta$ is assigned for all filters in modeling a particular data set. This homogeneity condition reduces the number of free parameters in the model and furthermore is justified because there is no evidence in our data for a systematic change in $\beta$ (detection psychometric function slope) as a function of temporal frequency.

Over the period of time from which the impulse response is initiated by the stimulus ($T_{\text{start}}$) until the re-

sponse decays completely ($T_{\text{end}}$) the total probability of each filter response $R_i$ being detected is given by

$$P_i = 1 - 2^{[-\int_{T_{\text{start}}}^{T_{\text{end}}} |R_i(t)|^\beta dt]}. \tag{7}$$

Given the independence of filter outputs $R_A$ and $R_B$, the total probability for detection of either response is

$$P_{\text{Det}} = 1 - (1 - P_A)(1 - P_B)$$

$$= 1 - 2^{[-\int_{T_{\text{start}}}^{T_{\text{end}}} (|R_A|^\beta + |R_B|^\beta)dt]}. \tag{8}$$

Thus $P_{\text{Det}}$ represents the total probability of detection. However, when the probability of correct performance ($P_C$) is measured in a 2-IFC task, these probabilities are related by the relationship $P_C = 0.5 + 0.5P_{\text{Det}}$. Taking the 75% level of performance to define the psychophysical threshold, we require that

$$1 = \int_{T_{\text{start}}}^{T_{\text{end}}} (|R_A|^\beta + |R_B|^\beta)dt. \tag{9}$$

This expression can be solved analytically for the contrast level at which performance is likely to reach the 75% performance level:

$$C_{\text{thr}} = \left\{ \int_{T_{\text{start}}}^{T_{\text{end}}} [|g(t, f) * h_A(t)|^\beta \right.$$

$$\left. + |g(t, f) * h_B(t)|^\beta]dt \right\}^{-1/\beta}. \tag{10}$$

In practice, however, a numerical solution to Eq. (10) is sought to model contrast threshold detection performance.

## C. Identification Performance
In this subsection we model how one uses the responses among the filters to determine identification performance. The first requirement is to select a metric translating the filter output over time into a single output measure. A minimal requirement is that this metric be a monotonically increasing function of stimulus contrast. We used a squaring of the peak filter response to represent an accelerating nonlinearity. This choice provided a better fit to the data than a simple linear transform in all cases and is a feature of models that predict increment contrast discrimination data.[39,40] Because the filter responses are independent, they can be represented in orthogonal directions in a perceptual space, as shown in Fig. 8. The filled squares indicate the squared peak responses of the two filters at detection threshold. The straight lines radiating from the origin correspond to the ratio of transduced filter outputs for our stimuli at the temporal frequencies indicated. When the transductions of the two filter responses are identical (in this case both are squared peak responses) this ratio is constant, resulting in a contrast-invariant measure of each temporal frequency. Following general line element and signal detection theory, each point in Fig. 8 corresponds to a perceptual state, and points close together are more likely to be confused than points farther apart. In our model we propose a zone of confusion in this space (depicted by the shaded circle) that describes the boundary between perceptual events likely to be confused and those that are reliably distin-
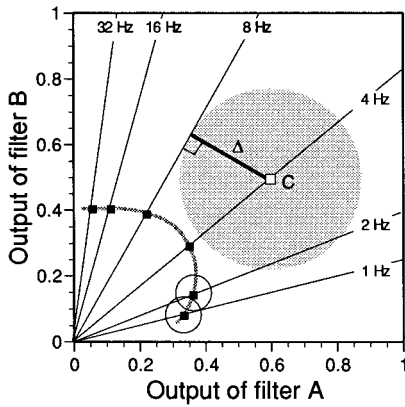
Fig. 8. Representation of internal space governed by the transduced output of the two basis filters. The thick curve indicates the squared peak responses of model filters A and B at detection threshold; the filled squares show this condition at the temporal frequencies indicated. Refer to the text for further details.

guished. For simplicity we take this zone to be circular with a constant radius $\Delta$ (i.e., the zone is invariant on translation and rotation within this space).

The $2 \times 2$-IFC task used here does not allow for simultaneous comparison of points in this space. We must therefore postulate that this space provides a map in memory onto which we lay down an internal representation of each temporal frequency, given by a straight line through the origin, so sequentially presented stimuli can be compared. In Fig. 8 the straight line labeled 8 Hz represents the pattern of filter responses that is due to an 8-Hz stimulus, regardless of contrast. Given the confusion zone depicted by the shaded circle, the open square labeled C denotes the smallest contrast of a 4-Hz stimulus that is reliably distinguished from the 8-Hz internal representation. The factor by which contrast must be raised above detection threshold to yield 75% correct identification corresponds to the factor alpha measured in the psychophysical task. At this contrast level (C) the 4-Hz stimulus will remain confused with the 2-Hz representation but will be distinguished from the 1-Hz representation. This model also predicts that, as contrast increases, TF identification becomes more reliable, as is generally observed.

This output space also provides for a degree of asymmetrical performance in identifying pairs of stimuli. For example, consider the task of identifying whether a 1- or a 2-Hz stimulus is presented in a trial; we assume that the

observer has practiced and is familiar with both stimulus types. Further suppose (as shown by the two small open circles in Fig. 8) that the zone of confusion is just large enough that the 2-Hz stimulus at detection threshold is just reliably distinguished from the 1-Hz internal representation; i.e., the value of $\Delta$ equals the radius of a circle centered on the 2-Hz detection point, which just grazes the 1-Hz line. Thus we predict accurate identification of the 2-Hz stimulus at detection threshold. The same zone of confusion centered on the 1-Hz stimulus, however, cuts across the 2-Hz representation line, and so we predict that reciprocal accurate identification will not ensue for the 1-Hz case. Asymmetries such as these are observed in the data, even after correction for bias as described in Subsection 2.C.3.

### D. Fitting the Model

The adjustable parameters of the model are the overall sensitivity factors of the two filters ($A_A$ and $A_B$), the values of $\tau$ and $\sigma$ that together control the shape of the impulse responses, the psychometric function slope parameter $\beta$, and the size of the confusion zone $\Delta$. In practice the psychometric function slope parameter was set to the mean of that measured for each observer's detection performance, leaving only five free parameters to be constrained by the detection thresholds (measured at six temporal frequencies) and the alpha factors (for 30 combinations of those temporal frequencies). Another fundamental choice to be made concerning the model is which pair of adjacent derivative log Gaussian filters to use. By calculating the best resulting fits, using all combinations of filter functions, we found that the optimal fit was always afforded to the RG data when the generator and the first derivative of the log Gaussian temporal impulse were used, and for the achromatic condition when the first and the second log Gaussian derivatives were used. That is, for modeling the RG data $h_0$ and $h_1$ were used, whereas for the achromatic data $h_1$ and $h_2$ were used.

The free model parameters were iteratively adjusted by Powell's method[39] to minimize the total chi-squared statistic for both the detection and the identification data simultaneously. As the model uses 5 free parameters to fit 36 data points, there are 31 degrees of freedom associated with the model fit for each observer and cardinal stimulus condition. The best-fitting model parameters and the least chi-squared values that resulted from the detection and the identification data for each fit are given in Table 1, and the outcomes are shown graphically in Fig. 3 and

**Table 1. The Best-Fitting Parameters and the Least Chi-Squared Values Resulting from the Model Fit to the Detection and TF Identification Data[a]**

| Stimulus | Observer | Mean $\beta$ (measured) | Best-Fitting Model Parameters | | | | | Least Chi Square | | | $Q$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\tau$ | $\sigma$ | $A_A$ | $A_B$ | $\Delta$ | Det. | Ident. | Total | |
| Achromatic | ABM | 3.97 | 0.0736 | 0.7446 | 4.213 | 8.120 | 0.0798 | 1.79 | 54.33 | 56.11 | $2.65 \times 10^{-3}$ |
| Achromatic | KTM | 5.08 | 0.0831 | 0.5503 | 9.602 | 20.38 | 0.0901 | 1.77 | 43.83 | 45.60 | $3.39 \times 10^{-2}$ |
| RG | ABM | 3.14 | 0.0815 | 0.8262 | 12.69 | 10.06 | 0.1151 | 5.22 | 89.52 | 94.74 | $1.24 \times 10^{-8}$ |
| RG | KTM | 4.13 | 0.0931 | 0.8065 | 21.88 | 16.29 | 0.1213 | 2.46 | 82.69 | 85.15 | $3.51 \times 10^{-7}$ |

[a] The $\beta$ values used in the model were derived from the detection psychometric functions and were not varied in the least-chi-squared fitting procedure. For modeling the RG data filters $h_0$ and $h_1$ were used, whereas for the achromatic data $h_1$ and $h_2$ were used.

Figs. 4–7, respectively. $Q$ in Table 1 indicates the probability that, if the model is true, the observed data vary from the model predictions as a matter of chance.[41] The low values of $Q$ indicate that the model does not account for all the variance in the data and that there is room for improvement in the model performance. Nevertheless, the fitted model does afford an excellent account for the detection data in all instances and also provides a qualitative account of the TF identification data, incorporating the essential observation that different TF's can be identified correctly near detection thresholds and that suprathreshold contrasts are required for accurate classification of closer TF's. Because the present model is at best statistically qualitative, estimates of the errors associated with the best-fitting parameters are not given in Table 1.

As mentioned above, the model performance was improved in all cases when a squaring transducer was applied to the peak responses than in the case of a simple linear transduction. Allowing this power transducer function to vary freely afforded slightly better fits to the data but, in addition to its adding extra parameters, the transducer power parameter was highly codependent on the value of $\Delta$, suggesting these data do not lend themselves to constraining the form of filter response transduction into the internal space defined above. Other types of experiment such as contrast masking are required for a fuller understanding of this transduction aspect of the filter responses, which, as discussed below, affects the modeled contrast invariance of subsequent speed perception.

## 5.  DISCUSSION

The results clearly show that stimulus identification can be made solely on the basis of TF content for both RG isoluminant and achromatic defined stimuli. This was shown to be true at contrast levels close to threshold, and we also observed that, as contrast increases, TF identification becomes more reliable. From these results we dismiss the notion that the RG mechanism depends on a single univariant temporal filter and instead infer that (like the achromatic channel) the RG channel must embody more than one TF filter to allow the discrimination performance to be observed. The model framework that we have adopted allows us also to propose the likely shapes of the temporal IRF's that subserve RG and achromatic function at low contrast levels.

This interpretation of the data necessarily rests on the assumption that the isoluminant performance that we measure is based on the operation of chromatic mechanisms and not on the vestiges of some luminance response, either of the type suggested to account for residual luminance flicker after color fusion in heterochromatic flicker photometry[42] or that which is due to the frequency-doubled responses of magnocellular-projecting retinal ganglion cells.[43] To investigate this possibility, and to determine whether thresholds are indeed mediated by chromatic mechanisms, we performed further experiments examining the chromatic nature of the threshold percept for our stimuli.[44]

We suppose that, if chromatic mechanisms are involved, stimuli should appear distinctly chromatic (alternately red and green) at detection threshold. In a method completely analogous to the TF identification procedure described above we determined color identification performance by using threshold level nominally isoluminant stimuli randomly intermixed with similar threshold level achromatic stimuli. If an intruding luminance mechanism mediates detection for both achromatic and nominally isoluminant stimuli, then correct identification performance would not closely follow detection performance but would become possible only when a chromatic mechanism's response finally became significant, if at all.

We performed this test for both observers at a sample of frequencies in our TF range (2, 8, and 22.6 Hz) and found that in all cases color-identification performance closely followed detection performance, even at high temporal frequencies. Average separations ($\pm$ estimated SD) of detection and identification psychometric functions were for observer KTM: $0.024 \pm 0.048$, $0.040 \pm 0.042$, and $0.078 \pm 0.028$ $\log_{10}$ unit, and for observer ABM: $0.033 \pm 0.054$, $0.018 \pm 0.059$, and $0.042 \pm 0.045$ $\log_{10}$ unit, for 2, 8, and 22.6 Hz, respectively. These very small differences in detection and identification thresholds give us confidence that chromatic mechanisms mediated performance for our isoluminant stimuli.

### A.  Achromatic Impulse Response Functions
The similarity of the achromatic detection and identification data measured for both observers is reflected in the similarity of the best-fitting temporal IRF's resulting from the model, as shown in Fig. 3 and in Table 1. Filter $h_1$ has a biphasic IRF, resulting in a MTF that peaks between 4 and 5 Hz. The low-frequency limb of this function has a much shallower sensitivity decline than the high TF limb, but it still remains clearly bandpass. Filter $h_2$ has a triphasic IRF, but it should be noted that the secondary positive part of this function is small compared with the initial positive portion, although it does extend in time to just beyond 250 ms. The resulting MTF of this filter is distinctly bandpass in shape and peaks at ~8 Hz.

The crossover point in sensitivity for the underlying achromatic filters occurs at ~4 Hz in both subjects, in accordance with the results of Thompson[8] inferred from suprathreshold velocity-discrimination experiments at a range of spatial frequencies. At the crossover TF we would expect best sensitivity to small changes in TF, and therefore the most accurate TF identification should be centered about this frequency, as is observed in the data. Indeed, both subjects could accurately identify 8-Hz or faster stimuli as well as 2-Hz or slower stimuli at or very close to detection threshold when these stimuli were paired with 4-Hz achromatic stimuli (see panels for 4 Hz in Figs. 4 and 5). According to the labeled line high-threshold model[7] used previously to interpret this class of data, this level of categorization along the TF dimension would constitute evidence for three separate TF-tuned filters operating at detection threshold levels. Using the same procedure as ours, Mandler and Makous[10] found that at least one observer viewing a uniform 1° field was able to identify three distinct TF's (1, 4, and 45 Hz) at detection threshold and that, in general, identification im-

proved with increasing TF separation.[10]   Similarly, using 0.2-cycle/degree achromatic gratings, Hess and Plant[12] also found three steps in TF classification at detection threshold (0, 4, and 32 Hz).[12]   Both of these investigations used the labeled line high-threshold model of Watson and Robson[7] to conclude that at least three filters were in operation and modeled their data accordingly. On the other hand, the model used here embodies a different approach based on signal detection, whereby the output of a filter that itself does not determine the detection response still provides useful information for a separately determined decision concerning signal identification.   Thus identification, even at detection threshold, is based on a combination of filter outputs.   Furthermore, the high-threshold model does not readily account for the observation found in all data sets that identification performance improves with increasing contrast, whereas this feature of the results arises naturally from the framework of the present model.   Because of these differences in underlying theory, we can model our data, which comprise three discrete categorical steps of TF at detection threshold, by using the output of only two independent filters. These filters are compatible in shape with those derived psychophysically with masking procedures.[11,13]

## B.   Red–Green Chromatic Impulse Response Functions

Despite the observation that chromatic TF identification performance was mostly better for observer KTM than for observer ABM, Fig. 4 shows that the inferred underlying IRF's derived by the model are consistent between the observers.   For both subjects the inferred filter $h_0$ has a monophasic IRF peaking at 81–93 ms after stimulus onset, gradually decaying over ~450 ms.   This results in a low-pass MTF that would have a maximum acuity of ~22 Hz if high chromatic contrasts could be generated to drive it this fast.   Filter $h_1$ has a biphasic IRF, which peaks at 38–44 ms, crosses zero at 81–93 ms, and slowly decays after ~350 ms, resulting in a bandpass MTF with maximum sensitivity at 3–4 Hz.   The shallow negative lobe of this filter's IRF reaches a minimum at 125–145 ms. Although filter $h_1$ has a bandpass MTF, because of the greater sensitivity of $h_0$ the overall sensitivity profile of the RG mechanism remains low pass.   The crossover point in sensitivity for the underlying RG filters is at ~5 Hz for both observers, as expected given the identification data in Figs. 7 and 8 from which the model is derived.

It is interesting to note that the properties of the inferred biphasic chromatic IRF are similar to those reported by Eskew et al.,[5] who used a double-pulse procedure.   Also, the $h_0$ IRF is similar in general shape but peaks at a time intermediate between that of the monophasic IRF's reported by Uchikawa and Ikeda[2] and that of Burr and Morrone[4] which had maxima at ~55 and ~120 ms, respectively.   The peak times of the monophasic chromatic IRF's computed by Swanson et al.[1] varied according to background adaptation level, a factor that could be responsible for the differences in the IRF shapes determined by these other studies.   The biphasic IRF reported by Swanson et al.[1] at 900 Td does not resemble the biphasic IRF inferred from the data and the model presented here.   As noted above, these previous studies all reported the action of a single temporal filter under any

experimental condition, a situation that is untenable according to the thesis of this paper.   It would be interesting to ascertain what filter shapes would be modeled from these pulse detection data under the assumption that more than one filter were operable.

## C.   Red–Green and Achromatic Speed Computations

Using our model framework, we can suggest how differences in the inferred RG and achromatic front-end temporal filters can provide a link to understanding differences in chromatic and achromatic speed perception.   There is considerable psychophysical evidence that both chromatic and achromatic information is combined in a common motion-processing stage at some level in the visual system.   This is supported by observations that adding color to an achromatic stimulus slows it down, whereas adding luminance contrast to an isoluminant chromatic grating speeds it up,[27,45] and that motion aftereffects induced by isoluminant chromatic motion can be transferred to and nulled by achromatic motion and vice versa.[46–48]

If we consider the outputs for the RG and achromatic filters to be measured and compared in the same filter output space, then this model can provide an explanation for the phenomenological differences reported between chromatic and achromatic motion perception.   In the representation of internal space shown in Fig. 8 the ratio of filter outputs, or equivalently the angle in this space, uniquely encodes each stimulus along the TF dimension independently of contrast.   Using the angle in a common internal space as a measure of stimulus speed, we illustrate in Fig. 9 how higher chromatic TF's are consistently required to derive the same speed signal as a slower achromatic stimulus for observer KTM.   The speed predictions of this model can be directly compared with the results of an earlier published report of measured velocity matching performed for isoluminant RG gratings by the same observer.[32]   There it was shown that, although perception of smooth motion was restricted to high chromatic contrasts, a 0.3-cycle/degree RG grating drifting at 3.2 Hz was velocity matched with an achromatic grating drifting
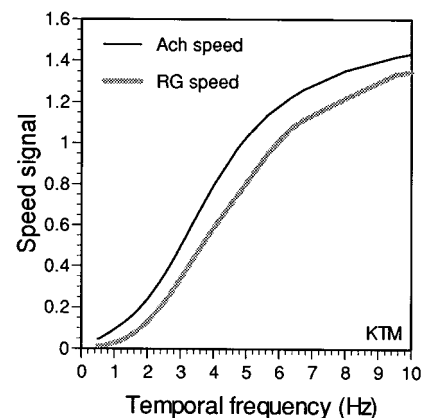


Fig. 9.   Speed signal as determined by the ratio of underlying filter outputs for RG and achromatic (Ach) stimuli derived for observer KTM.   Higher chromatic TF's are required for the same speed signal as for a slower achromatic stimulus, assuming a common internal response space.

at ~2 Hz, whereas a 1.6-Hz RG grating appeared to drift at an equivalent achromatic speed of 0.7–1.1 Hz. These speed comparisons are in good agreement with the predictions shown in Fig. 9, although it should also be noted that the Mullen–Boulton study,[33] along with others,[45–49] also reported that chromatic speed perception, especially at low speeds, is not contrast independent near threshold, signifying an important failure of our model.

Likewise, it is well established that achromatic velocity perception is also not strictly contrast invariant near detection threshold.[17,45,49,50] However, simple adjustments to the model can be made to incorporate these phenomena: Within both the RG and the achromatic channel the property of contrast-invariant TF coding is a direct result of identical transduction for both temporal filters. One can thus address the failure of the model by considering different nonlinear transducers for each of the two underlying filters. For stimuli to be more readily perceived as slower than they really are it is required that the power exponent related to the low TF filter be smaller than that for the higher TF filter. However, as we have seen, the types of experimental data reported here are not able to constrain adequately this important aspect of the model. Future experiments using contrast masking paradigms are planned to address this issue and to further develop the model.

## ACKNOWLEDGMENTS

## REFERENCES

1. W. H. Swanson, T. Ueno, V. C. Smith, and J. Pokorny, "Temporal modulation sensitivity and pulse detection thresholds for chromatic and luminance perturbations," J. Opt. Soc. Am. A **4**, 1992–2005 (1987).
2. K. Uchikawa and M. Ikeda, "Temporal integration of chromatic double pulses for detection of equal-luminance wavelength changes," J. Opt. Soc. Am. A **3**, 2109–2115 (1986).
3. K. Uchikawa and T. Yoshizawa, "Temporal responses to chromatic and achromatic change inferred from temporal double-pulse integration," J. Opt. Soc. Am. A **10**, 1697–1705 (1993).
4. D. C. Burr and M. C. Morrone, "Impulse-response functions for chromatic and achromatic stimuli," J. Opt. Soc. Am. A **10**, 1706–1713 (1993).
5. R. T. Eskew, Jr., C. F. Stromeyer III, and R. E. Kronauer, "Temporal properties of the red-green chromatic mechanism," Vision Res. **34**, 3139–3144 (1994).
6. P. E. King-Smith and J. J. Kulikowski, "Pattern and flicker detection analyzed by subthreshold summation," J. Physiol. (London) **249**, 519–548 (1975).
7. A. B. Watson and J. G. Robson, "Discrimination at threshold: Labelled detectors in human vision," Vision Res. **21**, 1115–1122 (1981).
8. P. Thompson, "Discrimination of moving gratings at and above detection threshold," Vision Res. **23**, 1533–1538 (1983).
9. M. B. Mandler, "Temporal frequency discrimination above threshold," Vision Res. **24**, 1873–1880 (1984).
10. M. B. Mandler and W. Makous, "A three channel model of temporal frequency perception," Vision Res. **24**, 1881–1887 (1984).
11. S. J. Anderson and D. C. Burr, "Spatial and temporal selectivity of the human motion detection system," Vision Res. **25**, 1147–1154 (1985).
12. R. F. Hess and G. T. Plant, "Temporal frequency discrimination in human vision: evidence for an additional mechanism in the low spatial and high temporal frequency region," Vision Res. **25**, 1493–1500 (1985).
13. R. F. Hess and R. J. Snowden, "Temporal properties of human visual filters: number, shapes and spatial covariation," Vision Res. **32**, 47–59 (1992).
14. S. J. Waugh and R. F. Hess, "Suprathreshold temporal-frequency discrimination in the fovea and the periphery," J. Opt. Soc. Am. A **11**, 1199–1212 (1994).
15. S. T. Hammett and A. T. Smith, "Two temporal channels or three? A re-evaluation," Vision Res. **32**, 285–291 (1992).
16. M. G. Harris, "Velocity specificity of the flicker to pattern sensitivity ratio in human vision," Vision Res. **20**, 687–691 (1980).
17. L. Stone and P. Thompson, "Human speed perception is contrast dependent," Vision Res. **32**, 1535–1549 (1992).
18. A. T. Smith, "Velocity perception and discrimination: relation to temporal mechanisms," Vision Res. **27**, 1491–1500 (1987).
19. A. T. Smith and G. K. Edgar, "Antagonistic comparison of temporal frequency filter outputs as a basis for speed perception," Vision Res. **34**, 253–265 (1994).
20. W. Richards, "Quantifying sensory channels: generalizing colorimetry to orientation and texture, touch and tones," Sensory Process. **3**, 207–229 (1979).
21. P. Cavanagh, "Vision at equiluminance," in *Vision and Visual Dysfunction: Limits of Vision*, J. J. Kulikowski, V. Walsh, and I. J. Murray, eds. (Macmillan, London, 1991), Vol. 5, pp. 234–250.
22. S. J. Cropper, "Velocity discrimination in chromatic gratings and beats," Vision Res. **34**, 41–48 (1994).
23. A. B. Metha, "Detection and direction discrimination in terms of post-receptoral mechanisms," Ph.D. dissertation (University of Melbourne, Melbourne, Australia, 1994).
24. A. B. Metha, A. J. Vingrys, and D. R. Badcock, "Calibration of a color monitor for visual psychophysics," Behav. Res. Methods Instrum. Computers **25**, 371–383 (1993).
25. P. Cavanagh, D. I. A. MacLeod, and S. M. Anstis, "Equiluminance: spatial and temporal factors and the contribution of blue-sensitive cones," J. Opt. Soc. Am. A **4**, 1428–1438 (1987).
26. C. F. Stromeyer III, R. E. Kronauer, A. Ryu, A. Chapparo, and R. T. Eskew, Jr., "Contributions of human long-wave and middle-wave cones to motion detection," J. Physiol. (London) **485**, 221–243 (1995).
27. P. Cavanagh, C. W. Tyler, and O. E. Favreau, "Perceived velocity of moving chromatic gratings," J. Opt. Soc. Am. A **1**, 893–899 (1984).
28. H. L. DeVries, "The heredity of the relative numbers of red and green receptors in the human eye," Genetica **24**, 199–212 (1948).
29. D. Regan and C. W. Tyler, "Some dynamic features of color vision," Vision Res. **11**, 1307–1324 (1971).
30. G. R. Cole, T. Hine, and W. McIlhagga, "Detection mechanisms in L-, M-, and S-cone contrast space," J. Opt. Soc. Am. A **10**, 38–50 (1993).
31. J. G. Robson, "Spatial and temporal contrast sensitivity functions of the visual system," J. Opt. Soc. Am. **56**, 1141–1142 (1966).
32. D. H. Kelly, "Spatiotemporal variation of chromatic and achromatic contrast thresholds," J. Opt. Soc. Am. **73**, 742–750 (1983).
33. K. T. Mullen and J. C. Boulton, "Absence of smooth motion perception in color vision," Vision Res. **32**, 483–488 (1992).
34. J. J. Koenderink, "Scale time," Biol. Cybern. **58**, 159–162 (1988).
35. A. Johnston and C. W. G. Clifford, "A unified account of

three apparent motion illusions," Vision Res. **35**, 1109–1123 (1995).
36. R. E. Fredericksen and R. F. Hess, "Two, three or four temporal channels?  A re-re-evaluation," Invest. Ophthalmol. Vis. Sci. **36**, S16 (1995).
37. A. B. Watson, "Temporal sensitivity," in *Handbook of Perception and Human Performance*, K. R. Boff, L. Kaufman, and J. P. Thomas, eds. (Wiley, New York, 1986), Vol. 1, pp. 1–43.
38. A. B. Watson, "Probability summation over time," Vision Res. **19**, 512–522 (1979).
39. G. E. Legge and J. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. **70**, 1458–1471 (1980).
40. G. E. Legge, "A power law for contrast discrimination," Vision Res. **21**, 457–467 (1981).
41. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge U. Press, Cambridge, 1988).
42. P. Kaiser, M. Ayama, and R. L. P. Vimal, "Flicker photometry: residual minimum flicker," J. Opt. Soc. Am. A **3**, 1989–1993 (1986).
43. B. B. Lee, J. Pokorny, V. C. Smith, P. R. Martin, and A. Valberg, "Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers," J. Opt. Soc. Am. A **7**, 2223–2236 (1990).
44. A. B. Metha, A. J. Vingrys, and D. R. Badcock, "Detection and discrimination of moving stimuli: the effects of color, luminance and eccentricity," J. Opt. Soc. Am. A **11**, 1697–1709 (1994).
45. K. T. Mullen and J. C. Boulton, "Interactions between colour and luminance contrast in the perception of motion," Ophthal. Physiol. Opt. **12**, 201–205 (1992).
46. K. T. Mullen and C. L. Baker, "A motion aftereffect from an isoluminant stimulus," Vision Res. **25**, 685–688 (1985).
47. P. Cavanagh and O. E. Favreau, "Colour and luminance share a common motion pathway," Vision Res. **25**, 1595–1601 (1985).
48. A. M. Derrington and D. R. Badcock, "The low level motion system has both chromatic and luminance inputs," Vision Res. **25**, 1879–1884 (1985).
49. M. J. Hawken, K. Gegenfurtner, and C. Tang, "Contrast dependence of colour and luminance motion mechanisms in human vision," Nature (London) **367**, 268–270 (1994).
50. P. G. Thompson, "Perceived rate of movement depends on contrast," Vision Res. **22**, 377–380 (1982).