

# Measuring the amplification of attention

ERIK BLASER\*, GEORGE SPERLING\*†‡§, AND ZHONG-LIN LU¶

Departments of \*Cognitive Sciences, and †Neurobiology and Behavior, and ‡Institute of Mathematical Behavioral Sciences, University of California, Irvine, CA 92697; and §Laboratory of Brain Processes, Department of Psychology, University of Southern California, Los Angeles, CA 90089

Contributed by George Sperling, August 6, 1999

**ABSTRACT** An ambiguous motion paradigm, in which the direction of apparent motion is determined by salience (i.e., the extent to which an area is perceived as figure versus ground), is used to assay the amplification of color by attention to color. In the red–green colored gratings used in these experiments, without attention instructions, salience depends on the chromaticity difference between colored stripes embedded in the motion sequence and the yellow background. Selective attention to red (or to green) alters the perceived direction of motion and is found to be equivalent to increasing the physical redness (or greenness) by 25–117%, depending on the observer and color. Whereas attention to a color drastically alters the salience of that color, it leaves color appearance unchanged. A computational model, which embodies separate, parallel pathways for object perception and for salience, accounts for 99% of the variance of the experimental data.

“Visual attention” refers to a class of mechanisms that select incoming information for subsequent processing from a particular set of locations in space and an interval in time. When the attended locations are determined by previous instructions to attend to a location, we say the mechanism is “top-down” attention, because the attentional selection depends on the interpretation, memory, and execution of the instructions; all high-level cognitive processes. Locations in space and time can also be selected “preattentively” by “bottom-up” mechanisms. For example, attending to red items in a many-colored scene would be an example of a top-down process. On the other hand, a single red item in a field of gray items would be selected on the basis of its uniqueness by preattentive, bottom-up processes.

An abstract representation of visual space, which we term a “salience map,” records, at each location of the visual field, the total attentional strength resulting from bottom-up plus top-down selection. The salience map is assumed to record only the attentional strength at a location; it does not record information about particular attributes (shape, color, texture, etc.) that produced this strength. A representation of total attentional strength is an essential component of theories for a variety of attentional tasks and phenomena, such as modeling the sequence of locations searched in visual search tasks (1), accounting for the ability to perceive motion in alternating-feature stimuli (2) and in isoluminant color gratings (3), gaining access to short-term visual memory (4, 5), figure-ground segmentation (2), and shape processing/object recognition (6).

The term “salience map” was popularized by Koch and Ullman (6), who used the concept to describe a winner-take-all network that determined a spatial location from which information from various topographic feature maps is combined and directed to a central processor. Related concepts have

emerged independently as an “attention map” (7), a “priority map” (8), a “selective tuning mechanism” (9), a “hierarchical pruning mechanism” (10), and under other names, with different authors giving somewhat different interpretations to these concepts. In *Discussion*, we consider in what respects the salience map and computations proposed here differ from these others.

The present study poses two questions: First, how can the impact of top-down, attentional selection be measured? In particular, when an observer is instructed to attend to a feature of the visual scene, e.g., red, can the resulting increase in visual salience be expressed as an amplification of the feature, as if the experimenter had actually increased the redness of the stimuli? Second, what is the spatial resolution of attention, i.e., of the salience map? If regions of high salience occur adjacent to regions of low salience, at what point is the resolution of the salience map exceeded? These questions are addressed within an ambiguous motion paradigm that lays the groundwork for a general computational model of visual attention similar to the kinds of theories found in low-level vision [e.g., visual optics (11), contrast sensitivity (12–15), and motion perception (16)].

We chose the red–green color axis as the dimension along which attention exerts its influence. An ambiguous motion stimulus (2) is used to assess the effectiveness of attention to color. Without attention instructions, the probability of perceiving motion in a particular direction depends on the chromaticity difference (redness or greenness) between a colored stimulus embedded in the motion sequence and the background (3). In the stimuli in these experiments, selective attention to red (or to green) is found to be equivalent to increasing the physical redness (or greenness) by 25–117%, depending on the observer and color. The data are used to construct a computational salience-map model that very accurately accounts for the experimental data.

## METHODS

**Overall Plan.** The paradigm for measuring attention utilizes a sensitive assay method, the “alternating-feature” motion paradigm (2). The paradigm uses motion stimuli that are completely neutral to the first-order (luminance-based) and second-order (texture-based) motion systems, but nonetheless give rise to clear and consistent apparent motion (see ref. 17 for a review). The remarkable fact is that the direction and strength of apparent motion in these “third-order” motion stimuli can be strongly influenced by attention (2, 17). Therefore, we can calibrate attention (a top-down factor) against bottom-up factors, such as color saturation and spatial frequency, to determine attentional color amplification prior to the third-order motion computation.

**Motion Systems.** In the ideal case, a motion system receives a time-varying image as input and produces a motion flow field as output. The flow field is a map of two-dimensional visual space in which each local neighborhood is represented by a

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at [www.pnas.org](http://www.pnas.org).

§To whom reprint requests should be addressed. E-mail: [sperling@uci.edu](mailto:sperling@uci.edu).

vector that indicates the direction and velocity of movement in that location. The output of a motion system does not directly indicate anything about the identity of the objects or events that were responsible for the flow field. Such information must be extracted by other subsystems that deal with texture, color, shape, and other image properties. Ultimately, information from many different subsystems is combined to generate perception.

The first-order motion system generates its flow field from a relatively raw luminance input in which an eye's multicolored visual image is represented simply as the luminance (amount of light) at each point  $x, y, t$ , relative to the overall mean luminance. The second-order system generates a flow field, analogously, not from the amount of light, but from the amount of texture. The third-order system generates a flow field from figure-ground information. That is, the perceptual system segregates most visual images into figure (the important part designated for further processing) and ground (not so designated). According to Lu and Sperling (2, 17–19), the results of this computation are stored in a salience map where figure is represented, for example, by 1 and ground by 0.

"Figure" and "ground" are binary concepts, and not every point on every image is so clearly divided. Therefore, we use a real-valued variable, salience, to indicate the relative importance of each image point (in space and time). At any moment in time, the set of instantaneous values of salience constitutes a salience map of the visual field, corresponding, in the obvious cases, to a figure-ground map. Third-order motion system is assumed to use this dynamic map as its input and to compute a flow field that gives the direction and the magnitude of salience movement at each point as a function of time. In effect, the third-order motion system computes the motion of those parts of the visual field that are designated as "figure."

**Controlling Salience.** There are two different ways to increase the salience of a region: an adjustment of a physical property or an adjustment of attention. For example, increasing the redness of a slightly reddish patch embedded in a large background of homogeneous yellow will make the reddish area more salient. The effectiveness of this manipulation depends primarily on bottom-up, preattentive processing. Deliberately attending to red (versus attending to green) in a stimulus that contains both red and green patches will also make the red regions more salient. This depends on a two-stage top-down process; that is, the high-level cognitive process of attending alters the low-level processing of visual inputs. [We know that this is not a high-level process operating directly on salience because salience is altered even in stimuli that are too brief for the observer to consciously discriminate the to-be-attended features (2).]

The hallmark of the study of spatial resolution in low-level vision is the application of techniques from linear systems analysis. For a linear system, the response to a sum of input components is the same as the sum of the system's response to each component individually. Because Fourier analysis can decompose an arbitrary input into its sinusoidal components, knowledge of a linear system's response to a basis set of sinusoids is sufficient to predict the system's response to any input whatsoever. Our approach, therefore, is to measure the extent to which the salience map can resolve sinusoidal stimuli spanning a wide range of spatial frequencies. For any requested distribution of visual salience, these data will, in principle, enable the prediction of the observer's achievable distribution of salience.

**Stimuli.** In the present study, ambiguous motion stimuli are used to assess the relative effectiveness of bottom-up and top-down factors. The motion stimuli consist of a temporal sequence of five spatially coincident frames, each of which contains a vertical sinusoidal grating. The gratings appear inside a rectangular aperture 10.7 cm wide and 6.6 cm tall. Temporal frequency was fixed at 2.5 Hz, stimulus width was 4

cycles, and spatial frequency was 0.50 cycles per degree at a viewing distance of 0.75 m. The stimuli used two types of gratings, an isoluminant red–green grating and a contrast-modulated noise grating. A motion sequence was constructed by alternating between red–green and contrast-modulated texture frames, with each successive frame displaced 90° consistently to the right or left, relative to its predecessor. In such a stimulus, the high-contrast texture patches are perceived as figure. Alternating with the texture stimuli are isoluminant color stimuli (20) containing side-by-side red and green regions (Fig. 1). In a red–green stimulus, the color of the area perceived as figure (i.e., having greater salience) by the motion system depends on which color differs more from the yellow background, which is itself simply a 50/50 mixture of red and green. In such a color stimulus, the amount by which a reddish or greenish area differs from the background is here called "chromaticity difference" and is designated as  $|R|$  or  $|G|$ .  $|R|$  and  $|G|$  vary from 0.0 (yellow, the background color) to +1.0 (purest available green or purest red) (Fig. 1).

**Procedure.** Three kinds of experimental sessions were conducted in sequence. First, in neutral "baseline" sessions, no attentional instructions were given, and observers simply made left/right direction judgments. An individual trial consisted of a 500-msec blank frame containing a fixation point, followed by a 5-frame motion stimulus (5 frames @ 100 msec per frame), followed by another fixation frame. Following the motion sequence, observers were required to enter a direction of motion judgment. On each trial, a chromaticity difference of red or green was chosen randomly from one of three chromaticity differences: 0.32 (pale red or green, barely discriminable from yellow), 0.60 (intermediate red or green), to

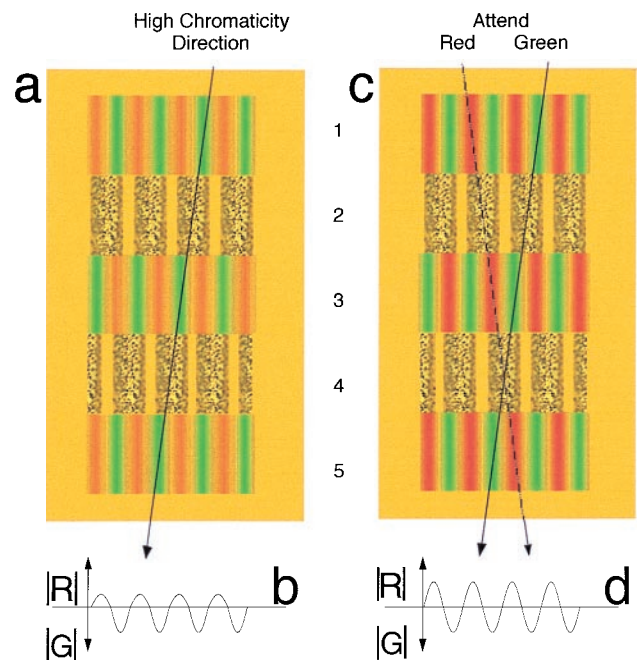


FIG. 1. Stimuli. Motion stimuli were composed of two types of gratings, an isoluminant red–green grating and a contrast-modulated noise grating (with same expected luminance throughout). (a) A five-frame motion sequence with a green advantage; each successive frame displaced 90° consistently to the right or left relative to its predecessor. Frames are presented one on top of the other. The color frames consist of saturated green,  $|G| = 1$ , and pale red,  $|R| = 0.32$ . The arrow indicates the typical direction of apparent motion. (b) Graphic representation of color in a green-advantage frame:  $|R| - |G|$  vs.  $x$ . (c) A neutral five-frame motion sequence:  $|R| = |G| = 1$ . The direction of apparent motion in this display is determined by attention, as indicated by arrows. (d) Graphic representation of color in a frame in which the red advantage is zero:  $|R| - |G|$  vs.  $x$ .

1.0 (highly saturated red or green). The chromaticity difference of the other color was chosen as 1.0, so that stimuli varied from having a large green advantage ( $|R| = 0.32$ ,  $|G| = 1.0$ ) to having a large red advantage ( $|R| = 1.0$ ,  $|G| = 0.32$ ). The starting frame type (texture or color) was chosen randomly. In subsequent sessions, observers viewed similar sequences of stimuli at different distances to yield four spatial-frequency conditions (number of spatial cycles per degree of visual angle).

After the baseline sessions, observers were instructed to attend to a particular color; again, sessions were conducted at four distances. Finally, the entire procedure was repeated with instructions to attend to the opposite color. Two observers completed the entire series in about 64 sessions; 17,600 observations (including practice); a third observer completed only the neutral and the “attend to red” conditions.

Stimulus sequences were designed so that there were always two potential interpretations on every trial. When red areas are rendered with a greater chromaticity difference from the background than green areas,  $|R| > |G|$ , we say that the stimulus has a “red advantage.” Red-advantage stimuli appeared to move in one direction, henceforth called the red direction. When green regions had a greater chromaticity difference,  $|R| < |G|$ , there was a green stimulus advantage, and apparent movement was in the opposite direction. Attending to red produced a similar effect to increasing the chromaticity difference of red. In other words, increasing the salience of an area either by increasing its redness or greenness, or by selectively attending to red or to green, increased the probability of perceiving motion in the direction consistent with that color.

## RESULTS

Observers’ responses are plotted as psychometric functions: the percent of red-consistent apparent motion judgments versus the red stimulus advantage  $|R| - |G|$ , (green advantage is indicated as a negative value of red advantage). Obviously, the greater the red stimulus advantage, the more likely observers are to perceive motion in the red-consistent direction. In Fig. 2, the only difference between the psychometric functions within a panel is the attentional state—the stimuli are identical for all three curves. The size of the lateral shift of a curve indicates the size of the attentional effect. The slope of the psychometric functions indicates sensitivity, the steeper the slope, the greater the sensitivity. Analysis of the data in Fig. 2 indicates that sensitivity (slope) decreases as spatial frequency increases (panels from top to bottom). The size of the attentional effect (lateral shift) is independent of spatial frequency. As spatial frequency increases from top to bottom in Fig. 2, the horizontal shift of the curves is the same, although it appears to be smaller when the psychometric functions are shallower at high spatial frequencies.

## THE MODEL

**Components.** For each subject, the continuous curves in Fig. 2 account for 99% of the variance of the data. They are derived from a model that includes both a bottom-up and a top-down attentional process (Fig. 3). In the model, there are two types of inputs: visual stimuli and attentional instructions. Visual stimuli are analyzed along various dimensions. Shown are channels that carry depth, orientation, texture, and color signals. In the present experiments, only the color and texture channels are critical for the motion computation.

The texture input into the salience computation consists of a texture grabber (21–23), which is composed of a linear spatial filter and a rectifier. The filter selectively responds maximally to textures of a specific coarseness-fineness; the rectifier computes the absolute value (removes the sign) of the filter

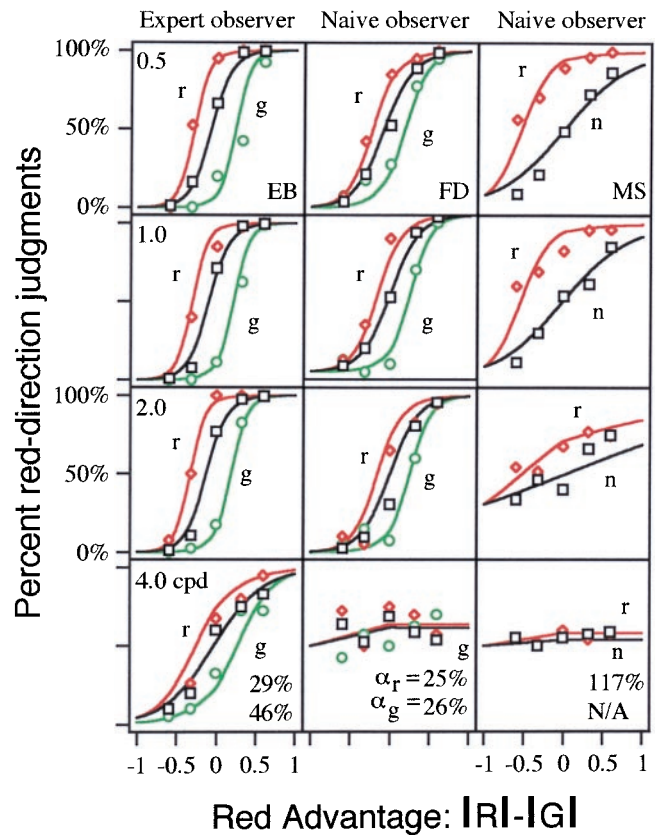


FIG. 2. Results. Psychometric functions are plotted as the percent of red-consistent motion judgments vs. the red stimulus advantage ( $|R| - |G|$ , the difference  $|R|$  of red from background yellow – the difference  $|G|$  of green from background yellow). As red advantage increases, observers are more likely to perceive motion in the red-consistent direction. Data points are shown for four spatial frequencies (rows; in cycles per degree, cpd) and three observers (columns). Solid curves are model fits (see Fig. 3). Black (middle) curves indicate the baseline condition (no attention instructions or n); red (r) and green (g) curves are model fits for the attend-red and attend-green conditions, respectively. The estimated model parameters for the increase in effective  $|R|$  and  $|G|$  because of attentional amplification,  $\alpha_r$  and  $\alpha_g$ , are indicated in the *Bottom* for each observer.

output (which could be positive or negative). For specificity, we assume that texture grabbers have outputs of 1 in areas of stimulus frames with the highest contrast texture, 0 in areas that do not contain texture, and intermediate outputs in areas of intermediate-contrast texture.

There are two color channels, red and green. Equivalently, the two color channels can also be conceptualized as a single green-minus-red channel, in which positive and negative values are carried along separate lines, and rectified, just as in the texture channel. Fully saturated red and fully saturated green are both represented by +1; the background is represented by 0; intermediate colors are represented by intermediate values proportional to their chromatic differences from the background.

At each location and instant in time  $x, y, t$ , texture, color, and other outputs sum up their contributions to the total salience (represented in the salience map) at that location. When there are no attention instructions, only these bottom-up processes are active. A field of standard motion-energy detectors (24, 25) (the third-order motion system) takes as its input the spatio-temporal distribution of salience and computes the third-order motion flow field. For a particular viewing distance, the texture frames in all stimulus sequences are equivalent, therefore, the directional motion energy in the flow field is determined by the amplitude and phase of the color signal. Random

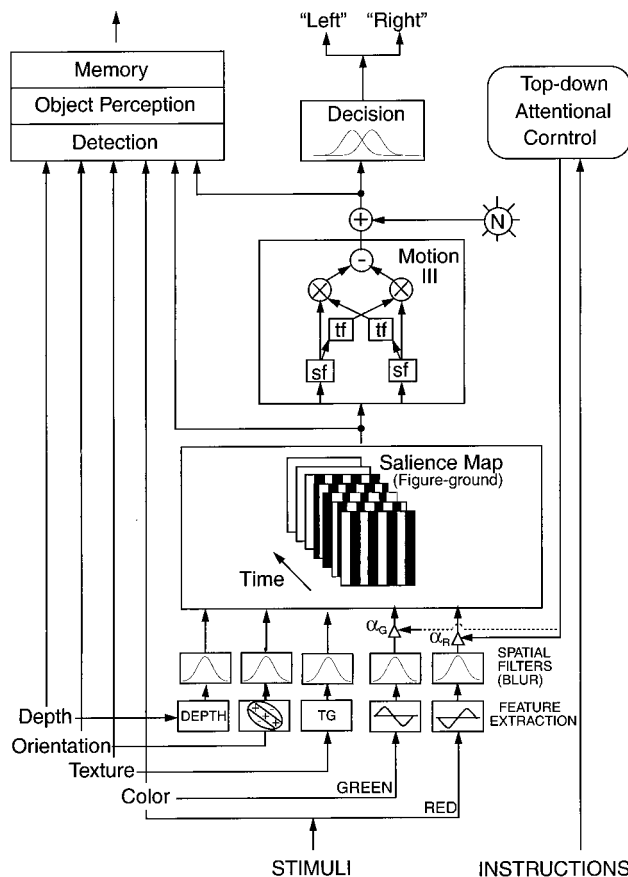


FIG. 3. A computational model of attentional processes in third-order motion embedded in a more comprehensive model of visual processing. The inputs to the model are stimuli and attentional instructions; the computational output is a direction-of-motion judgment. There are two types of inputs: visual stimuli and attentional instructions. Stimulus inputs are analyzed along various dimensions: depth, orientation, texture (TG, texture grabber), and color channels are indicated. For the present experiments, it is only necessary to consider color (red and green) and texture processing. Instructions to attend to a color are assumed, in only the saliency pathway, to increase the gain of the attended color signal to a value greater than 1.0, so that the attended color input is amplified by  $1 + \alpha_G$  or  $1 + \alpha_R$ . The "Saliency Map" is the sum of all the stimulus inputs in the saliency pathway; its output goes to the "Motion III" (third-order) computation, and also joins the stimulus inputs in the object-processing pathway. Motion III is represented as a Reichardt model (24); it produces a real-valued output that indicates a direction of motion and is perturbed by additive noise (N). "sf" denotes a spatial frequency filter; "tf" denotes a temporal frequency filter. A decision process outputs a response "Right" if its input is greater than a criterion, and "Left" otherwise. The third-order motion signal is also available to subsequent perceptual processes, as indicated.

noise (error) in the flow field is represented by the addition of noise to the motion flow field; a decision process evaluates the net flow field output and arrives at a decision of "left" versus "right" motion.

**Top-Down and Bottom-Up Control of Saliency.** In the model, instructions to attend to a color, say red, are assumed to increase the gain of the attended color signal to a value greater than 1.0, so that the attended color input is amplified. For example, if top-down attention to red were to amplify the red-filter output by 1.3, this would indicate an attentional amplification of 30%. Top-down attentional amplification is independent of spatial frequency and of the color (or texture) composition of the input stimuli.

**Figure/Ground.** The bottom-up process described above is assumed to be automatic figure/ground segmentation: Small

areas, areas that contain high-contrast texture, and areas that differ markedly from their surroundings are assigned higher saliency values relative to large, homogeneous areas, and tend to become segmented as figure rather than as ground. The present theory does not offer specific algorithms to determine figure/ground marking in general, but such rules would be similar to those that have been proposed in the first half of the century by the Gestaltists and that have been proposed, for example, by attention theorists (e.g., refs. 1 and 6–10). For the restricted stimuli used in the present experiment, the only computation needed is the amount by which the local texture (in the texture stimuli) or the local color (in the color stimuli) differs from the homogeneous background—a much simpler computation.

**Computational Efficiency.** Although the model is cast within a physiologically plausible framework, it is a computational model that describes the microprocesses of attention. As a computational model, it accounts for 60 data points for each observer (5 stimuli, 3 attention states, 4 viewing distances) with minimally only four parameters estimated from an observer: the amplification factors  $\alpha_G$ ,  $\alpha_R$ , for selective attention to green or to red; the amount of noise N, which determines the slope of the psychometric functions; and the spatial "corner" frequency  $f_c$ , which describes the spatial frequency at which sensitivity has been reduced by 1/2. (Spatial filtering in texture and color is assumed to be identical.) These parameters would have been sufficient to generate the predictions similar to Fig. 2. However, our primary interest is not in obtaining the absolutely most parsimonious computational account, but in demonstrating that a simple model that is consonant with reasonable physiological processes can provide a quite accurate account.

**Attentional Amplification Is Constant, Only the Manifestation Varies with Spatial Frequency and with the Red/Green Composition of the Stimuli.** We wish to test the hypothesis that a single parameter suffices to describe attentional amplification of an attended color ( $\alpha_G$  and  $\alpha_R$ ) under all conditions. To test this hypothesis, it is necessary to make all other parts of the model as accurate as possible so that an attention parameter is not used by the fitting algorithm to correct deficiencies elsewhere in the model. In order that small errors in describing the shape of the spatial filter function with a single corner parameter would not affect the above conclusion, we used three parameters to completely describe the spatial filter function for each observer (Fig. 4b). These spatial "tuning functions" represent the sensitivity of the channels as a function of spatial frequency. The corner frequency is between 2 and 4.5 cycles per degree for the three observers. All the changes in slope of the psychometric functions with spatial frequency are a consequence of the limited spatial resolution of this single filter. Additionally, a bias parameter was used for observer EB to account for the fact that the stimuli were not perfectly equal for him in the neutral attention condition—red was somewhat more salient. Given accurate specification of the nonattentional aspects of performance, a single attention parameter for "attention to red" and another parameter for "attention to green" was sufficient to predict for each of the observers, 99% of the variation in performance with attention for the 5 stimuli in each of 4 viewing distances.

Obviously, in Fig. 2, the apparent effect of selective attention is smaller when stimuli are smaller (high spatial frequencies) and when the red and green stripes in the stimuli differ greatly in saturation. For example, when the red stimulus advantage is very large, differences in selective attention make very little difference in performance. That one parameter per attentional state suffices to account for all the data means that these apparent differences in the effects of attention in different conditions are merely incidental manifestations of attentional states. All the data are accounted for by just three

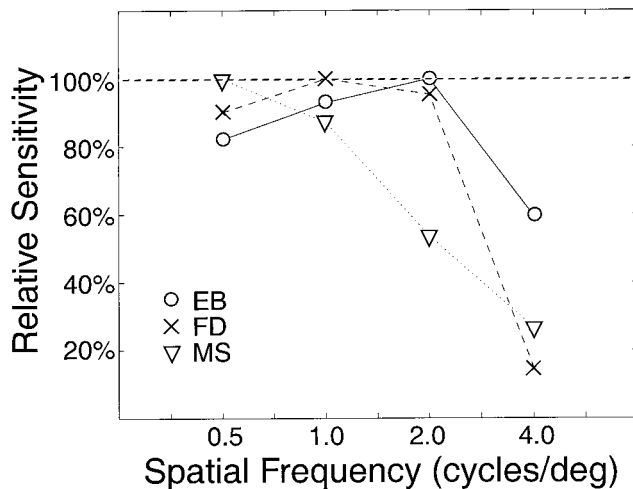


FIG. 4. Visual salience modulation transfer functions. The model parameter corresponding to bottom-up amplification (i.e., filter throughput without attentional amplification) is plotted against spatial frequency on log-linear coordinates. Data, shown for three observers, are typical of spatial tuning functions for third-order motion in other tasks (17).

states of attention: neutral, attention-to-red, and attention-to-green.

## DISCUSSION

**Attention Modifies Only Salience, Not Appearance.** It is critical to note that, in the model, the color input to only the salience map is amplified: color inputs to object perception processes are unchanged. This corresponds to the empirical fact that attention to color does not change the appearance of color (26), although it may make judgments of appearance slightly more reliable from trial to trial. It is perfectly obvious in observing our stimuli that they do not change their appearance when attention is selectively directed toward red or green. A change of a few percent in redness or greenness would be easily visible and does not occur when deliberately attending to red or green in stationary stimuli. Certainly, the equivalent changes in salience, 25% and more, are orders of magnitude larger than any possible change in appearance because of attention. This makes perfect sense from functional and physiological points of view: Attention to color should direct processing to colored locations so that whatever is present there (including color) is processed more accurately, but attention should not alter the appearance of the attended color. The distinction between the large effect of selective attention in altering the importance of an object or area of the visual field (its salience), and the small effect of attention in altering the representation of the object itself is absolutely

critical, especially in gaining understanding of the neurophysiological correlates of attention.

The salience map, which provides the input to the third-order motion computation, also provides the input to other processes. Because most scenes contain enormously more information than can be remembered (27), there is a selection of to-be-remembered locations; we assume that this selection is determined by the salience map. The salience map is also assumed to control input to object-recognition processes. It directs object-discrimination processes to compute the shape of those parts of the field that have been categorized as figure, and not to compute the shape of those parts that have been categorized as ground.

**Parallel Attentional Processes.** The selection of some parts of the visual input for object processing and, ultimately, for storage in memory, obviously involves attenuation of the to-be-forgotten items (or equivalently, amplification of the to-be-remembered items). At the point where the to-be-forgotten items are excluded from memory, they will have been attenuated almost to 0, relative to the to-be-remembered items. Thus, according to the computational model set forth here, there are two conceptually different representations of the stimulus: First, there is the complex representation of the stimulus itself, in which the local relations between features are unaltered by selective attention. Second, there is the relatively simple salience map representation, in which internal relations between features are dramatically altered by attention (bottom-up or top-down). Corresponding to these two representations are two different kinds of attentional amplification: first, attentional amplification that determines salience; second, attentional attenuation that determines selection for subsequent processing.

Previous computational theories (e.g., refs. 1 and 6–10) have dealt primarily with selection, and have not been concerned with the possibility of two qualitatively different representations, one for salience, another for selection. Typically, these models have a “winner-take-all” architecture, in which one item or object at a time is selected for further processing according to “preattentive” (bottom-up) and attentive (top-down) mechanisms. Can physiological data discriminate between the inner workings of all these proposed attentional mechanisms? There has been explosive growth of research that deals with attention effects measured in brain-imaging studies of human subjects, and in single cell records of behaving primates, including quantitative modeling of single cell responses in an attentional system (30); nevertheless, the physiological basis of the salience map and of the attentional processes proposed here remains to be discovered.

**How the Salience Model Applies to Common Paradigms** In third-order motion, output of the salience map is the input to a field of motion detectors that compute a motion flow field (Fig. 3). In other tasks, other recipients of the salience map output become critical. For example, in a partial report (iconic memory) experiment, an observer views a stimulus with, say,

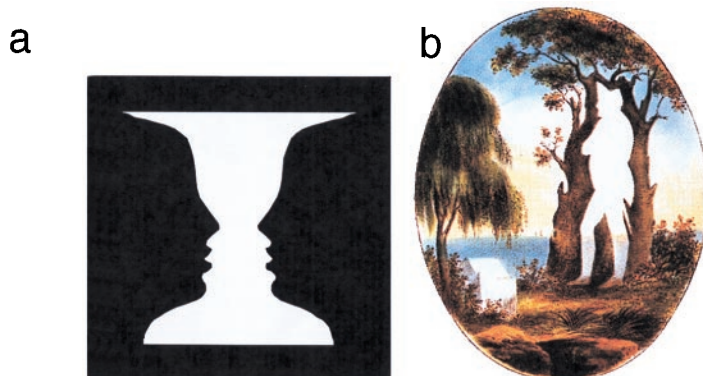


FIG. 5. Figure-ground ambiguities. (a) Ambiguous face-vase (31). (b) Normally, trees are seen as figure, and the space between as ground. However, when the space is attended, or when it has a meaningful shape, it also can be seen as figure. (Courier & Ives, ca. 1835).

three rows of letters, and a cue directs him/her to report a particular row. Here, we assume that the salience map controls the input to short-term memory. That is, the attentional cue produces top-down activation of the appropriate region of the salience map; this is the mechanism by which access to short-term memory is controlled. In experiments where only one row is displayed, top-down activation is unnecessary; bottom-up activation of the salience map determines that it is the row with letters, not a blank row, that is recorded in memory. Insofar as third-order motion reveals dynamic properties of the salience map rather than of the third-order motion-detection process, it is enormously more efficient to derive the properties of the salience map from motion experiments than from partial-report experiments.

In pattern recognition, the salience map is assumed to determine what part of the field is sent to object recognition mechanisms. Pattern/shape recognition processes are assumed to operate only on the parts of the field that are designated as figure, i.e., those parts that have high values of salience. Normally, figure/ground segregation is a bottom-up process, determined by the texture, shape, and other properties of the visual input. Sometimes, figure/ground segregation is ambiguous, as in Fig. 5*a*; more often, it is not. For example, in a forest, we normally see the trees as "figure" and the spaces between the trees as "ground." However, when we wish to know whether the space between two trees will enable us—and not whatever is chasing us—to fit through, the space has to become figure. In terms of the salience map computation, attending to the space between trees produces a sufficient increase in the salience value of that area, so that the open space, not the trees, is sent forward to pattern-recognition processes. That observers can compute the shape of the space between trees is nicely illustrated in Fig. 5*b*.

**Guided Search.** Perhaps visual search is the paradigm that has been used the most to study visual pattern recognition. Typically, an observer views an array of items that includes a target and nontargets (distracters). The search process consists of examining the elements, either in serial or in parallel, or in some combination, to discover the target. Theories of visual search involve the strategic allocation of processing resources (1, 28, 29), according to priorities that are determined by *a priori* probabilities of finding particular kinds of targets in particular locations and by the properties of the stimulus being viewed. Because it combines both bottom-up and top-down influences, the salience map provides an ideal mechanism to implement precisely this kind of "guided search."

**Summary and Conclusions.** The full panoply of predictions that can be derived from the computational model still remains to be tested. However, the experimental paradigm that has been put forward here has enabled us to measure the functional amplification produced by visual attention to color in a variety of conditions. The model draws an important distinction between the very large attentional amplification of the salience of an attended color, while the appearance of the color itself is not significantly changed. The attentional amplification of salience is considerable: The maximum amplification values for each of the three observers were 46%, 26%, and 117%. The computational model put forward here is consistent with our

overall understanding of the attentional components of the brain processes involved in attentionally determined apparent motion, in object recognition, and in short-term memory, and it provides a highly accurate account of a considerable data set.

This work was supported by the U.S. Air Force Office of Scientific Research, Life Sciences, Visual Information Processing Program.

1. Cave, K. R. & Wolfe, J. M. (1990) *Cognit. Psychol.* **22**, 225–271.
2. Lu, Z.-L. & Sperling, G. (1995) *Nature (London)* **377**, 237–239.
3. Lu, Z.-L., Lesmes, L. A. & Sperling, G. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 8289–8294.
4. Rumelhart, D. E. (1970) *J. Math. Psychol.* **7**, 191–218.
5. Sperling, G. & Weichselgartner, E. (1995) *Psychol. Rev.* **102**, 503–532.
6. Koch, C. & Ullman, S. (1985) *Hum. Neurobiol.* **4**, 219–227.
7. Mozer, M. (1991) *The Perception of Multiple Objects: A Connectionist Approach* (MIT Press, Cambridge, MA).
8. Ahmad, S. & Omohundro, S. (1991) International Computer Science Institute Technical Report tr-91-040 (Univ. California, Berkeley).
9. Tsotsos, J. K., Culhane, S. M., Wai, W. Y. K., Lai, Y., Davis, N. & Nuflo, F. (1995) *Artificial Intelligence* **78**, 507–545.
10. Burt, P. (1988) in *9th International Conference on Pattern Recognition*, Rome, Italy (IEEE Computer Society Press, Washington, DC), pp. 977–987.
11. Campbell F. W. & Gubisch, R. W. (1966) *J. Physiol.* **186**, 558–578.
12. Lowry, E. M. & DePalma, J. J. (1961) *J. Opt. Soc. Am.* **51**, 740–746.
13. Lowry, E. M. & DePalma, J. J. (1962) *J. Opt. Soc. Am.* **52**, 328–335.
14. Davidson, T. M. (1966) *Science* **152**, 797–799.
15. Campbell, F. W., Nachmias, J. & Jukes, J. (1970) *J. Opt. Soc. Am.* **60**, 555–559.
16. Anderson, S. J. & Burr, D. C. (1985) *Vision Res.* **25**, 1147–1154.
17. Lu, Z.-L. & Sperling, G. (1995) *Vision Res.* **35**, 2697–2722.
18. Lu, Z.-L. & Sperling, G. (1996) *Curr. Dir. Psychol. Sci.* **5**, 44–53.
19. Sperling, G. & Lu, Z.-L. (1998) in *High-Level Motion Processing*, ed. Watanabe, T. (MIT Press, Cambridge, MA), pp. 153–183.
20. Anstis, S. & Cavanagh, P. (1983) in *Color Vision*, eds. Mollon, J. D. & Sharpe, E. T. (Academic, New York), pp. 155–166.
21. Chubb, C. & Sperling, G. (1988) *J. Opt. Soc. Am. A* **5**, 1986–2006.
22. Chubb, C. & Sperling, G. (1989) *Proceedings: Workshop on Visual Motion* (IEEE Computer Society Press, Washington, DC), pp. 126–138.
23. Werkhoven, P., Sperling, G. & Chubb, C. (1993) *Vision Res.* **33**, 463–485.
24. van Santen, J. P. H. & Sperling, G. (1984) *J. Opt. Soc. Am. A* **1**, 451–473.
25. Adelson, E. H. & Bergen, J. R. (1985) *J. Opt. Soc. Am. A* **2**, 284–299.
26. Prinzmetal, W., Amiri, H., Allen, K. & Edwards, T. (1998) *J. Exp. Psych. Hum. Percept. Perform.* **24**, 261–282.
27. Sperling, G. (1960) *Psychol. Monogr.* **74**, 1–29.
28. Koopman, B. O. (1957) *Oper. Res.* **5**, 613–626.
29. Sperling, G. & Doshier, B. (1986) in *Handbook of Perception and Performance*, eds. Boff, K., Kaufman, L. & Thomas, J. (Wiley, New York), Vol. 1, pp. 1–65.
30. Reynolds, J. H., Chelazzi, L. & Desimone, R. (1999) *J. Neurosci.* **19**, 1736–1753.
31. Rubin, E. (1915) *Synsoplevede Figurer: Studien i psykologisk analyse* (Gyldendalske, Copenhagen). German trans (1921): *Visuell wahrgenommene Figuren: Studien in psychologischer Analyse* (Gyldendalske, Copenhagen).