Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 44 (2004) 541–556
Vision Research
www.elsevier.com/locate/visres

Visual working memory for image statistics

Jonathan D. Victor *, Mary M. Conte

Department of Neurology and Neuroscience, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA

Received 20 May 2003; received in revised form 3 November 2003

Abstract

To define the role of statistical features of images in visual working memory, we compared the ability of subjects ($N = 6$) to identify changes in arrays of black and white checks when these changes altered some aspect of their statistical structure, versus when these changes did not. Alteration of luminance statistics or local higher-order statistics improved performance, but alteration of the degree of bilateral symmetry did not. The dependence of performance on the degree of statistical change indicated that statistical information was represented in a graded, rather than categorical, fashion.

Keywords: Symmetry; Isodipole

1. Introduction

Early vision segments the retinal image into objects and represents these objects in a manner in which they can be recognized. Individual features such as lines and edges play a role in these processes, but often the statistics of visual images are at least as important. For example, in complex images, including naturalistic ones, only a small number of contrast contours represent object boundaries, and consequently objects are more reliably defined by discontinuities in image statistics, rather than by isolated features (Julesz, 1981a; Marr, 1982). Despite the impressive ability of the visual system to make use of scene statistics, much previous work (Caelli & Julesz, 1978; Caelli, Julesz, & Gilbert, 1978; Julesz, Gilbert, Shepp, & Frisch, 1973; Julesz, Gilbert, & Victor, 1978; Victor & Conte, 1991, 1996), which has focused on texture discrimination and segmentation, indicates that this statistical processing is limited and specific. In natural viewing, segmentation of an image proceeds along with an analysis of surface composition. This analysis (e.g., sand versus wood versus stone) is doubtless based on image statistics (Cho, Yang, & Hallett, 2000), rather than a pixel-by-pixel match with an exemplar, and thus, represents another situation in which image statistics play a crucial role.

To determine the extent to which visual image statistics are available for such processes as distinct from spatial segmentation or explicit discrimination tasks, we made use of a visual working memory task introduced by Cornelissen and Greenlee (2000). These authors showed that human subjects are able to determine whether a random array has changed over a brief time interval, but performance on this task is poor. Most likely, the poor performance on this task reflects the limited capacity of visual working memory for encoding and/or representing such images on a pixel-by-pixel basis. Consequently, if alteration in an array could be detected by a change in image statistics, then performance might improve dramatically. That is, we can assay the extent to which visual working memory makes use of image statistics by determining the extent to which introduction of specific kinds of statistical structure affect performance on this task.

Since there are far too many image statistics to attempt a rigorously comprehensive analysis (Cho et al., 2000; Harvey & Gervais, 1981), we adopt a "survey" strategy motivated by physiological principles and previous work with texture discrimination and segmentation. We will therefore consider exemplars of three classes of visual statistics (see Fig. 2 for examples). The first class consists of image statistics that influence the mean activity of a population of retinal ganglion cells. This includes luminance and second-order correlation structure; we will use luminance (the fraction of white checks) as an exemplar of this class (Fig. 2(A)). A

* Corresponding author. Tel.: +1-212-746-2343; fax: +1-212-746-8984.
E-mail address: jdvicto@med.cornell.edu (J.D. Victor).

0042-6989/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2003.11.001

second class of visual statistics requires cortical analysis for extraction, but it suffices that this analysis occur within a local region. We will use fourth-order correlations, as manifest by the "even and odd" isodipole textures (Julesz et al., 1978) as an exemplar of this class (Fig. 2(B)). Sensitivity of cortical (Purpura, Victor, & Katz, 1994) but not lateral geniculate neurons (Victor, 1986) to these statistics has been demonstrated experimentally. A third class of statistics can only be extracted via cortical analysis that is long-range. We will use bilateral symmetry, widely considered a salient and important visual feature (Attneave, 1954; Tyler, 1995; Wenderoth, 1994), as an exemplar of this class (Fig. 2(C)–(E)). Textures that isolate the first two kinds of statistical structure can readily be constructed by homogeneous Markov random fields (Zhu, Wu, & Mumford, 1998), while the third cannot. Nevertheless, as described below, each of these kinds of structure can be introduced into otherwise random arrays in a quantitative and graded fashion, allowing us to measure their effects independently and on a common footing. Moreover, because each kind of structure can be precisely quantified, an information-theoretic analysis can be used to compare the intrinsic difficulty of the psychophysical tasks. As shown below, only the first two classes of image statistics appear to be used in visual working memory, even though bilateral symmetry is widely considered to be visually salient.

2. A categorical representation?

If indeed image statistics are used to identify or classify surface materials (Cho et al., 2000), one might speculate that they play a role in object identification analogous to that of color. A distinctive feature of color perception, both in discrimination (Bornstein & Korda, 1984; Wandell, 1985) and visual memory tasks (Amano, Uchikawa, & Kuriki, 2002) is that the physical continuum of color space is represented in a *categorical* manner. In a categorical representation, performance depends primarily on where these stimuli are located with respect to one or more boundaries in a perceptual space (Berlin & Kay, 1969; Bornstein & Korda, 1984; Wandell, 1985). Stimuli are distinctive on opposite sides of such a boundary. Accuracy on a discrimination task is enhanced and reaction time decreases on trials in which stimuli are drawn from opposite sides of a category boundary, compared to trials in which both stimuli are drawn from the same category. The alternative is a graded representation, in which the parametric difference between two stimuli, not their position relative to a perceptual boundary, is the main determinant of performance. For those classes of image statistics that are used in visual working memory, our approach allows us to ask whether this representation is categorical or graded. Although the analogy to color and the role of textures in material or surface classification might suggest a categorical representation, the picture that emerges is largely a graded one.

3. Methods

3.1. Subjects

Studies were conducted in six normal subjects (two male, four female), ages 30–54. Other than author MC, the remaining subjects were naïve to the purpose of the experiments. Subjects were practiced psychophysical observers in a related task involving targets in the same positions relative to fixation (Victor & Conte, 2001), and had visual acuities (corrected if necessary) of 20/20 or better.

3.2. Stimuli

The stimulus frame S1 (Fig. 1) consists of four arrays of checks on a mean gray background. The arrays were positioned along the cardinal axes, with centers 200 min from fixation. In most experiments, each array subtended 160 min and contained 64 (8×8) contiguous checks, each of which was either black or white and subtended 20 min. The stimulus frame S2 also consisted of four arrays, three of which were identical to those in the S1 frame of the trial. The target array, determined at random, differed from the corresponding array in S1 by a contrast inversion of 16 of the 64 checks. In other experiments, the arrays were of equal size but contained only 16 (4×4) checks, each subtending 40 min. In these experiments, the target array differed from the corresponding array in S1 by inversion of eight of the 16 checks. The number of checks that differed between S1 and S2 was chosen so that performance was neither at chance nor ceiling, and at approximately the same level for the two array sizes.

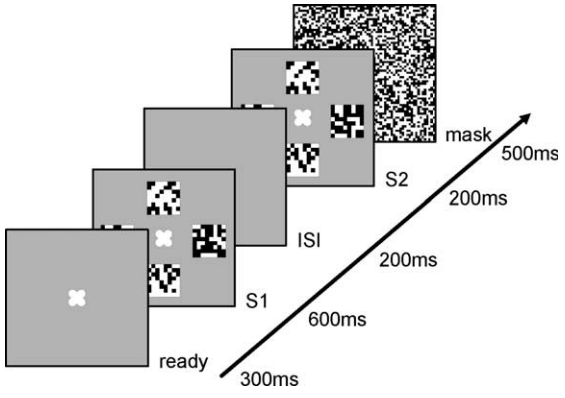

Fig. 1. A diagram of a typical trial. The subject's task is to determine which of the four arrays in S1 has changed in S2.

The main experimental manipulation consisted of the assignment of luminance values to the checks. For each experiment, a particular kind of statistical structure was introduced into the arrays: luminance bias, higher-order statistical structure (the "isodipole" textures), or symmetry. In each case, the strength of the statistical structure was parameterized by a quantity c, where $c = 0$ denotes a maximally random assignment, and $c = 1$ (or $c = -1$) denotes a maximally structured assignment. In the experiments that examined luminance statistics, an array corresponding to a value c had $\frac{1+c}{2}$ of its checks white, and $\frac{1-c}{2}$ of its checks black. In the experiments that examined higher-order statistics, $c = 1$ corresponded to a maximally "even" texture, while $c = -1$ corresponded to a maximally odd texture. In the symmetry experiments, $c = 1$ corresponded to a texture in which all pairs of checks that were related by the symmetry axis were matched in luminance, while $c = -1$ corresponded to a texture in which all such pairs were opposite in luminance. Appendix A provides details the construction of these stimuli, including the precise definition of the textures for intermediate values of c.

Within an experiment, trials were constructed with one or more *pairs* of c-values, as follows. To construct a trial based on the pair $(c_{\text{low}}, c_{\text{high}})$, each of the four arrays in S1 was constructed either with $c = c_{\text{low}}$ or with $c = c_{\text{high}}$. The c-value of the target array in S2 was also randomly assigned to one of these two values of c. Thus, for each pair of c-values, there were $128 = 2^5 \times 4$ varieties of stimuli, since for each of the five arrays (four in S1, and the target in S2) there were two possible values for c, and the target could occur in any of four locations. Each of these varieties was presented the same number of times. In exactly half of the trials, designated "different statistics" trials below, the c-value of the target changed from S1 to S2 by an amount $\Delta c = |c_{\text{high}} - c_{\text{low}}|$; in half of these trials, c changed by $+\Delta c$ (increasing from c_{low} to c_{high}), and in half it changed by $-\Delta c$ (decreasing from c_{high} to c_{low}). In the other half of the trials, designated "same statistics" trials, the c-value of the target remained at either c_{low} or c_{high}. Thus, neither the c-values chosen for the arrays in S1, nor the c-values chosen in S2, provided a cue as to which array was the target. As described in the Appendix A, it is possible to construct sets of arrays that not only have the requisite values of c, but also differ in a specific number of checks, as required by the experimental design.

3.3. Apparatus

The above visual stimuli were produced on a Sony Multiscan 17seII (17″ diagonal) monitor, with signals driven by a PC-controlled Cambridge Research VSG2/3 graphics processor programmed in Delphi II to display precomputed maps (generated in Matlab) for specified periods of time. The resulting 768×1024 pixel display had a mean luminance of 47 cd/m^2, a refresh rate of 100 Hz and subtended 11×15 deg (approximately 1 min/pixel) at the viewing distance of 114 cm. The intensity versus voltage behavior of the monitor was linearized by photometry and lookup table adjustments provided by VSG software. Stimulus contrast was 1.0.

3.4. Procedure

Essentially, our experimental design is a modification of the Cornelissen and Greenlee (2000) visual working memory task, in which (a) four stimuli are presented simultaneously, and (b) statistical cues are intentionally introduced. All experiments were organized as a sequence of 4-alternative forced choice trials, whose common features are as follows (Fig. 1). After binocular fixation on a uniform gray background, the subject initiated a trial via a button-press on a Cambridge Research CT3 response box 300 ms later, a stimulus (S1, described in detail above) appeared, consisting of four arrays of checks, surrounding a central "X" subtending approximately 30 min. After presentation of S1 for 600 ms, the display returned to mean luminance for 200 ms, following which a second stimulus S2 (described above) appeared, containing a "target" that differed from the corresponding array in S1. After presentation of S2 for 200 ms, a mask was presented for 500 ms, consisting of a full-field random checkerboard whose checks were half as large (linear dimension) as those in S1 and S2. The subject's task was to identify the target array via a button-press on a response box with four buttons, positioned corresponding to the stimulus arrays. Subjects were instructed to maintain central fixation and to respond as quickly as possible, but not to compromise accuracy for speed. Responses and reaction times (measured with respect to the onset of S2) were collected via the Delphi II display software. Trials in which the subject responded before the onset of S2, or after 8000 ms, were discarded and repeated.

An experimental session consisted of a single block of either 512 or 640 trials (4 or $5 \times$ the 128 varieties of trials that were required to examine each kind of statistical structure and each pair of c-values, presented in random order). In Experiment I, one pair of c-values was examined for each kind of statistical structure. In Experiment II, four or five pairs of c-values were examined for each kind of statistical structure. To accumulate a sufficient number of trials for each c-value pair in Experiment II, four sessions (on separate days) were required. In Experiment I, subjects were shown paper exemplars of trials with the maximal level of statistical structure at the beginning of each session. In Experiment II, subjects were shown exemplars that spanned the c-values to be used. In both cases, subjects were informed that the trials would be similar to these exemplars. For data analysis in Experiment II, results

from each subject were pooled across sessions. Thus, the critical comparisons between c-values in Experiment II were based on trials run in parallel.

4. Results

4.1. Experiments I: What kinds of image statistics are available in visual working memory?

In these five experiments, we examined the effects of large variations of statistics on visual working memory. We examined two kinds of local statistics (luminance and isodipole) and three examples of the third class of statistics (global symmetry). In each experiment, the target array either was constructed with statistics that differed by a large amount $\Delta c = |c_{\text{high}} - c_{\text{low}}|$ along one of these axes, or not at all.

Fig. 2 illustrates the stimuli and summarizes the results from the five sub-experiments. For the luminance experiment (Fig. 2(A)), we used $c_{\text{high}} = 0.25$ and $c_{\text{low}} = -0.25$, so $\Delta c = 0.5$ in the "different statistics" trials. This led to a substantially higher fraction correct, compared to the fraction correct in the "same statistics" trials, in each of the six subjects. On average, fraction correct improved from 0.69 to 0.89 for the 16-check arrays in which eight checks changed, and from 0.56 to 0.84 for the 64-check arrays in which 16 checks changed. The improvement in performance was highly statistically significant ($p < 10^{-4}$ for each subject and each array size individually, two-tailed χ^2). Here and below, we only consider differences significant if there are consistent differences in most subjects, when data are analyzed individually (without Bonferroni correction).

For the isodipole experiment (Fig. 2(B)), we used $c_{\text{high}} = 1$ (maximally "even") and $c_{\text{low}} = -1$ ("odd"), so $\Delta c = 2$ in the "different statistics" trials. As with luminance statistics, there was a substantial increase in fraction correct in the "different statistics" trials. On average, fraction correct improved from 0.79 to 0.86 for the 16-check arrays in which eight checks changed, and from 0.58 to 0.83 for the 64-check arrays in which 16 checks changed. The improvement in performance was highly statistically significant ($p < 0.02$ for four of six subjects for the 16-check array, and $p < 10^{-4}$ for each subject for the 64-check array). The finding that there is a smaller improvement for the 16-check array than for the 64-check array is entirely due to the fact that performance in the "same statistics" trials for 16-check arrays was higher than for 64-check arrays.

Increases in fraction correct were accompanied by decreases in reaction time (calculated by averaging across trials for which the response was correct). In luminance experiments, the reaction time decreases were large and statistically significant for both array sizes (16-check array: mean RT decrease across subjects, 144 ms; $p < 0.05$ for five of six subjects by one-tailed t-test; 64-check array: mean RT decrease across subjects, 191 ms; $p < 0.01$ for all six subjects). For the isodipole experiments, the change in reaction time was minimal for the 16-check array, but large for the 64-check array (16-check array: mean RT decrease across subjects, 20 ms; $p < 0.05$ for one subject; 64-check array: mean RT decrease across subjects, 80 ms; $p < 0.01$ for four of six subjects). This closely paralleled the pattern of findings for fraction correct. In sum, these experiments show that change in an array is easier to detect when its luminance statistics (Fig. 2(A)) or isodipole statistics (Fig. 2(B)) change.

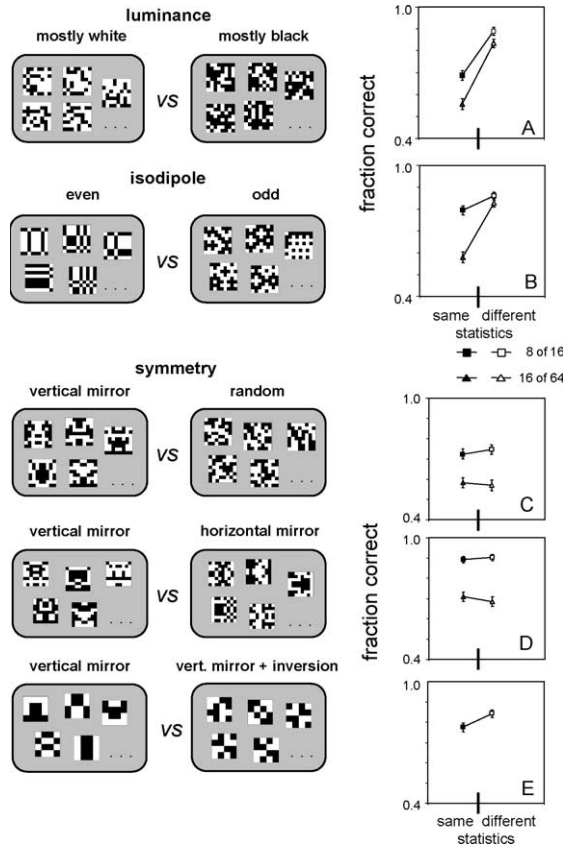

Fig. 2. The five kinds of statistical manipulations used in Experiment I (left), and a summary of performance (right). Fraction correct is pooled across subjects, and the error bars represent 95% confidence limits of the pooled values, based on binomial statistics. The size of the array and the number of checks that change is indicated by symbol shape: squares, eight checks change within an array of 16 checks; triangles, 16 checks change within an array of 64 checks. Open symbols: "different statistics" trials; filled symbols: "same statistics" trials. (A) Luminance statistics, (B) isodipole statistics, (C) vertical symmetry versus absence of symmetry, (D) vertical symmetry versus horizontal symmetry, (E) vertical symmetry versus contrast-inverted vertical symmetry.

For the symmetry experiments, the results were quite different. Fig. 2(C) shows an experiment in which some arrays had perfect bilateral symmetry along the vertical axis ($c_{\text{high}} = 1$), and others were random ($c_{\text{low}} = 0$). There was no suggestion of a significant difference in

performance between the "different statistics" trials ($\Delta c = 1$) and the "same statistics" trials ($p > 0.05$ in each subject for each array size; $p = 0.14$ pooled across subjects for the 16-check arrays; $p = 0.49$ pooled across subjects for the 64-check arrays). Reaction times were minimally decreased in the "different statistics" trials (16-check array: mean RT decrease across subjects, 40 ms; $p < 0.05$ for two of six subjects; 64-check array: mean RT decrease across subjects, 18 ms; $p < 0.01$ for one subject).

We sought to increase the magnitude of this effect in two ways. In the experiment of Fig. 2(D), all arrays were fully symmetric, but some had a vertical symmetry axis, while others had a horizontal symmetry axis. Overall performance was somewhat better, but again there was no difference in performance between the "different statistics" trials and the "same statistics" trials ($p > 0.05$ in each subject for each array size; $p = 0.34$ pooled across subjects for the 16-check arrays; $p = 0.13$ pooled across subjects for the 64-check arrays). In the experiment of Fig. 2(E), we implemented the maximum possible modulation of vertical symmetry, by using arrays that either had perfect perfect bilateral symmetry along the vertical axis ($c_{\text{high}} = 1$), or those in which checks related by the mirror pairing *mismatched* in luminance ($c_{\text{low}} = -1$). As described in Appendix A, such stimuli could only be realized with the 16-check array size. While there was a statistically significant difference in performance between the "different statistics" trials and the "same statistics" trials ($p < 10^{-4}$ across all subjects), this effect was not robust (significant at $p < 0.05$ in only two subjects, MC and EM). Moreover, this small difference might be due to the fact that all of the stimuli with mismatch along the mirror axis ($c_{\text{low}} = -1$) necessarily had a vertical contrast contour at the midline, while the stimuli with perfect vertical symmetry ($c_{\text{high}} = 1$) necessarily had no contour at this location.

Correspondingly, there was no consistent effect of a symmetry change on reaction time (vertical versus horizontal symmetry: RT decrease of 22 ms for the 16-check array and RT *increase* of 15 ms for the 64-check array; symmetry versus mismatch: RT decrease of 11 ms; only three of 18 comparisons significant within subjects at $p < 0.05$). Thus, in contrast to our findings with local statistics (Fig. 2(A) and (B)), gain, loss, or change of bilateral symmetry did not appear to provide a useful cue in this visual working memory task.

4.1.1. Role of spatial differences in image statistics

Differences between the statistical structure of the arrays might influence performance not only via aiding representation in working memory, but also via spatial differences within S1 or S2—for example, by guiding attention. The statistical structure of each of the four arrays ($c = c_{\text{high}}$ versus $c = c_{\text{low}}$) was assigned independently, and with equal probability. Thus, on one quarter of the trials, a triplet of the S1 arrays had $c = c_{\text{high}}$ and the remaining array (a "singleton") had $c = c_{\text{low}}$, and in a separate quarter of the trials, a triplet of S1 arrays had $c = c_{\text{low}}$ and the remaining singleton array had $c = c_{\text{high}}$. If the singleton array drew spatial attention by virtue of their distinctive statistics (i.e., elicited "pop-out"), one might expect that performance would be enhanced on trials in which the target was also a singleton array in S1 or S2 (Joseph & Optican, 1996; Treisman, 1982).

To identify the role played by spatial differences in image statistics, we separated the trials into three groups: those in which the target was a singleton in S1 (here designated "pop-out in S1"), those in which the target was a singleton in S2 (here designated "pop-out in S2"), and the remaining trails, in which the target was not a singleton in either S1 or S2 (here designated "no pop-out"). We use these terms as a shorthand, and not to imply a mechanism. They indicate not only the presence or absence of a singleton array, but specifically whether the singleton is also the target.

We found that for some of the statistical classes, fraction correct was higher in the pop-out trials than in the "no pop-out" trials. For the luminance experiments based on 64-check arrays, average fraction correct was 0.73 in the "pop-out in S1" trials, 0.76 in the "pop-out in S2" trials, and 0.60 in the "no pop-out" trials. For isodipole experiments, the corresponding fractions correct were 0.75 in the "pop-out in S1" trials, 0.78 in the "pop-out in S2" trials, and 0.61 in the "no pop-out" trials. Differences between the pop-out and "no pop-out" trials were highly significant ($p < 10^{-4}$ pooled across subjects for luminance and for isodipole statistics); differences between "pop-out in S1" and "pop-out in S2" were not significant ($p > 0.2$ pooled across subjects for luminance and isodipole classes).

Reaction time differences among these three kinds of trials were generally small, and not statistically significantly within subjects. Note that the number of correct pop-out trials was typically 50–60 for each subject (out of 64–80 pop-out trials), so small differences in reaction times might not reach statistical significance. Pooling data across subjects only revealed a statistically significant difference in RT in the luminance trials, and only when pop-out occurred in S2. Compared with the "no pop-out" trials, RT in "pop-out in S2" trials was reduced by 43 ms (16-check arrays, $p < 0.02$) and 49 ms (64-check arrays, $p < 0.02$).

In sum, an effect of spatial differences in image statistics on fraction correct was seen for both statistical classes that showed an effect of "different statistics" versus "same statistics", and on RT for the luminance class. However, the effect of statistical change (Fig. 2) was larger than, and not accounted for by, this phenomenon. This is seen in Fig. 3. As seen in Fig. 3(A) for the luminance experiments, the difference in fraction correct between the "same statistics" trials and in the

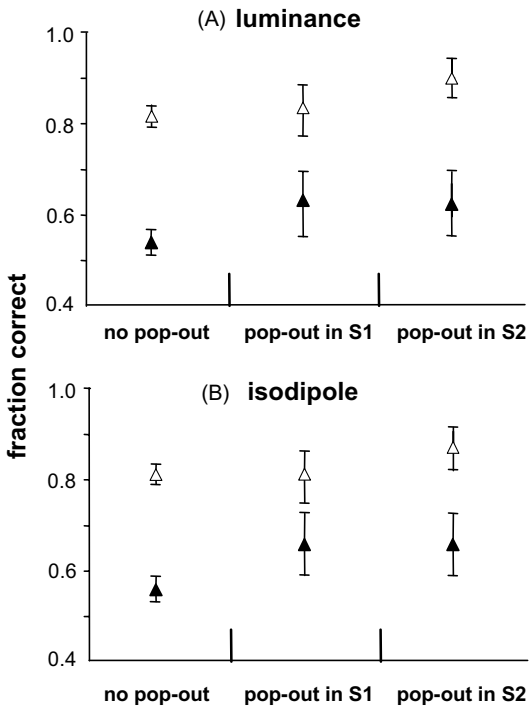

Fig. 3. Analysis of performance for trials in which the target's statistics can guide spatial attention by virtue of its statistics in S1, or in S2, or in neither trial. Open symbols: "different statistics" trials ($\Delta c \neq 0$); filled symbols: "same statistics" trials ($\Delta c = 0$). (A) Luminance statistics ($\Delta c = 0.5$ or 0), (B) isodipole statistics ($\Delta c = 2$ or 0). See text for details on how trials were classified as "pop-out" and "no pop-out" trials. Data for 64-element arrays. Error bars as in Fig. 2.

"different statistics" trials persisted when "pop-out in S1", "pop-out in S2", and "no pop-out" trials were separately analyzed. A similar pattern was seen for the experiment based on isodipole statistics, as seen in Fig. 3(B), and in the trials based on the 16-check arrays (not shown).

4.2. Experiments II: Structure of the perceptual space

The previous experiments established that certain image statistics, and not just the pixel-by-pixel details of an image, were represented in visual working memory. Since all instances of "different statistics" consisted of large changes ($\Delta c = 0.5$ for luminance statistics, $\Delta c = 2$ for isodipole statistics), this result leaves open the question of whether this representation is categorical or graded. The next set of experiments address this issue by examining the influence of small changes in statistics. We restricted consideration to the 64-element arrays, since this size provided a larger effect in Experiment I than the 16-element arrays (Fig. 2(A) and (B)). Four of the six subjects who participated in Experiment I provided data for all of the experiments described here. A fifth subject from Experiment I (KS) also provided data for the first isodipole statistics experiment described below.

4.2.1. Luminance statistics

Luminance statistics were investigated with $\Delta c = 0.25$, and stimulus pairings ranging from $(c_{low}, c_{high}) = (-0.625, -0.375)$ to $(c_{low}, c_{high}) = (+0.375, +0.625)$. Examples of the arrays used are shown in Fig. 4. As seen in Fig. 5(A), fraction correct in the "different

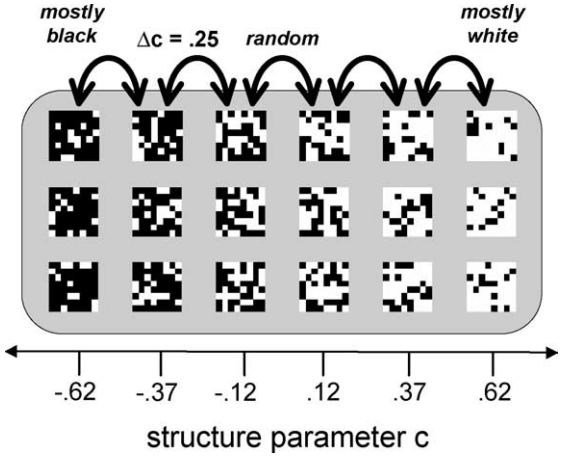

Fig. 4. Examples of stimuli for the luminance version of Experiment II. Each column contains three representative examples of arrays constructed with the indicated value of the structure parameter c. Adjacent columns are separated by $\Delta c = 0.25$, the amount of statistical change used in the experiment.

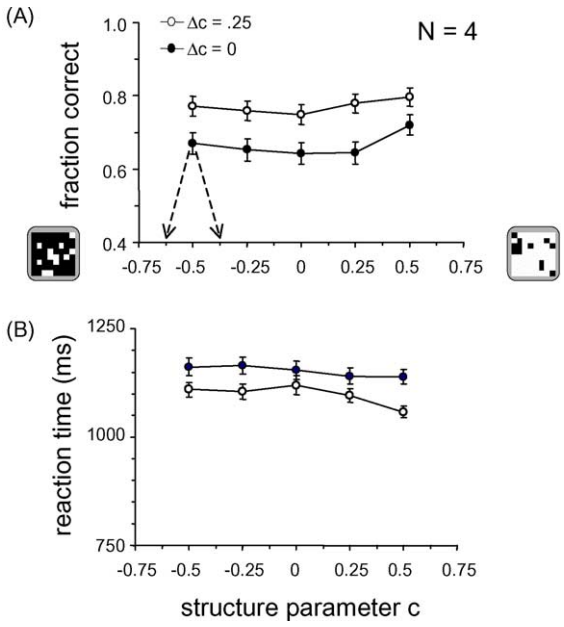

Fig. 5. Results of the luminance experiment with $\Delta c = 0.25$. Open circles: "different statistics" trials; filled circles: "same statistics" trials. (A) Fraction correct, pooled across four subjects. The dashed arrows indicate the (c_{low}, c_{high}) pair associated with an example data point; the other data points correspond to an equally separated pair of values. (B) Reaction time, pooled across four subjects. Error bars for fraction correct as in Fig. 2; error bars for reaction times are the means of the 95% confidence limits within each subject.

statistics'' trials was higher than that in the ''same statistics'' trials. This difference, about 0.1, was independent of the position of the stimuli along the range of luminance statistics. It was highly statistically significant ($p < 10^{-4}$) for each pairing of c-values in data pooled across subjects. The same pattern was seen in all individual subjects, though not all comparisons (3, 4 or 5 out of 5) reached statistical significance ($p < 0.05$) in data from individual subjects. Correspondingly, reaction time (Fig. 5(B)) was shorter in the ''different statistics'' trials than in the ''same statistics'' trials. The reaction time reduction, on average 54 ms, was also independent of the position of the stimuli along the range of luminance statistics. Reaction time changes were of only modest statistical significance across subjects ($p = 0.04$–0.07 via one-tailed paired t-test at four of the five pairings, $p > 0.2$ at the middle pairing) and within subjects ($p < 0.05$ at 2–5 of the pairings in individual subjects), most likely because of the intersubject variability of reaction times.

4.2.2. Isodipole statistics

A corresponding analysis for isodipole statistics, with $\Delta c = 0.4$ (Fig. 6) is shown Fig. 7. The only clear difference in fraction correct between ''different statistics'' and ''same statistics'' trials was for $(c_{\text{low}}, c_{\text{high}}) = (+0.6, +1)$ ($p < 10^{-3}$ pooled across subjects, $p < 0.05$ in four of five individual subjects). For seven of the other eight pairs $(c_{\text{low}}, c_{\text{high}})$ tested, there was a tendency, though not statistically significant, in the same direction. Reaction time data (Fig. 7(B)) showed no clear difference ($p > 0.05$) between conditions at all nine pairings, both within and across subjects.

These data suggest that there is something unique about the fully ''even'' ($c = 1$) stimulus, but leave open the possibility that a small difference (that failed to reach statistical significance) was also present at lower values of c. For this reason, we also measured performance for

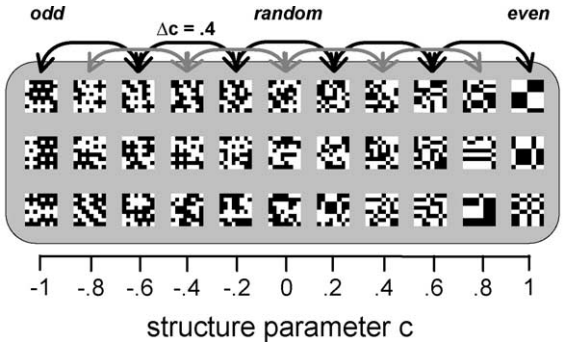

Fig. 6. Examples of stimuli for the isodipole version of Experiment II. Each column contains three representative examples of arrays constructed with the indicated value of the structure parameter c. Adjacent columns are separated by $\Delta c = 0.2$. Stimulus pairings (indicated by arrows) were constructed from examples drawn from next-nearest neighbors ($\Delta c = 0.4$).

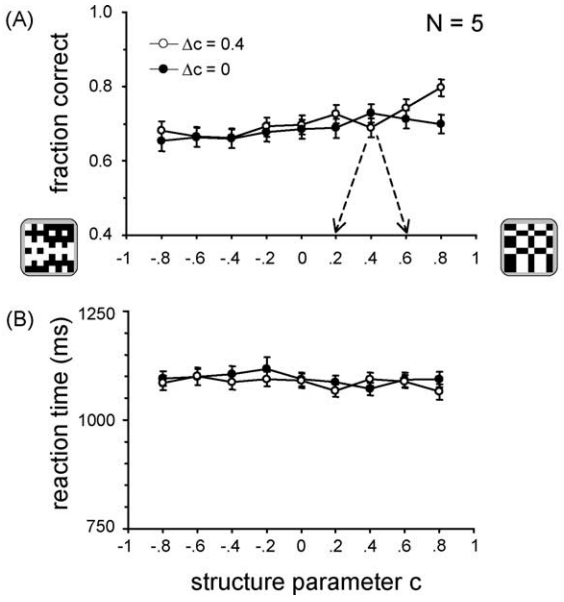

Fig. 7. Results of the isodipole experiment with $\Delta c = 0.4$. (A) Fraction correct, pooled across subjects. (B) Reaction time, pooled across subjects. Plotting conventions as in Fig. 5.

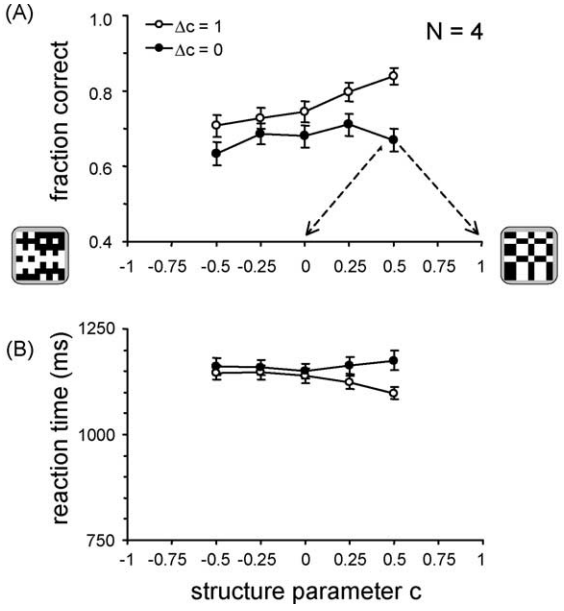

Fig. 8. Results of the isodipole experiment with $\Delta c = 1$. (A) Fraction correct, pooled across four subjects. (B) Reaction time, pooled across four subjects. Plotting conventions as in Fig. 5.

stimuli constructed with $\Delta c = 1$. As seen in Fig. 8(A), a change in image statistics does improve performance across the entire range of c ($p < 0.05$ at each pairing, pooled across subjects). However, this improvement is larger near the extreme ''even'' ($c = 1$) end, both in terms of magnitude and statistical significance. In particular, there was an improvement in performance for one or both of the pairings at the ''even'' end ($c = 1$) of the range at $p < 0.05$ in all four subjects, but only one

subject's individual data showed a comparable difference at the "odd" end ($c = -1$) of the range. There was a 31 ms reduction of reaction time in the "different statistics" trials across the entire range (Fig. 8(B)). In keeping with the previously observed close parallel of reaction time and fraction correct, the reduction in reaction time was greater at the "even" end of the range (77 ms, $p < 0.05$) than at the "odd" end (14 ms, $p > 0.15$).

4.2.3. Dependence of performance on the statistics of the target

We have shown that a change in the statistics of the target between S1 and S2 leads to improved performance, across the entire range of luminance and isodipole statistics. Here we ask whether there is an influence of the position of the target along the statistical range in S1 or in S2, in addition to the effect of a *change* in position between S1 and S2. To make this distinction, we separately consider the "same statistics" and "different statistics" trials in the above experiments.

We consider the luminance experiments (Fig. 9(A)) first. For the "same statistics" trials (upper panel), there is no overall trend of the fraction correct as a function of the c-value of the target (regression slope 0.036, $p > 0.1$, F-test). However, for trials in which the stimuli were at the dark ($c < 0$) end of the range, fraction correct was greater ($p < 0.03$) when the target was the brighter of the two alternatives ($c = c_{\mathrm{high}}$). Conversely, for trials in which the stimuli were at the bright ($c > 0$) end of the range, fraction correct was greater ($p < 10^{-4}$) when the target was the darker of the two alternatives ($c = c_{\mathrm{low}}$).

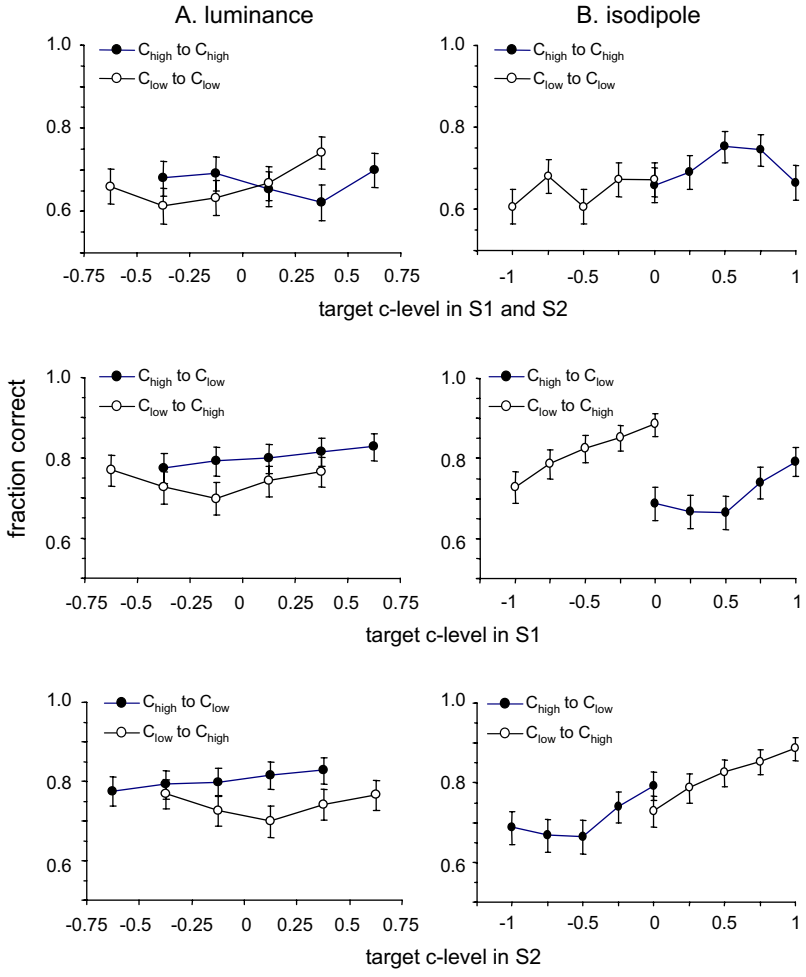

Fig. 9. Dependence of fraction correct on the c-value of the target. (A) Luminance experiment of Fig. 5 with $\Delta c = 0.25$. Negative and positive values of c represent, respectively, the dark and light ends of the range of luminance statistics. (B) Isodipole experiment of Fig. 8 with $\Delta c = 1$. Negative and positive values of c represent, respectively, the odd and even ends of the range of luminance statistics. Error bars as in Fig. 2. Upper plot: performance in the "same statistics" trials, as a function of the c-value of the target, which is identical in S1 and S2. Open symbols: target $c = c_{\mathrm{low}}$ in S1 and S2. Filled symbols: target $c = c_{\mathrm{high}}$ in S1 and S2. Middle plot: performance in the "different statistics" trials, as a function of the c-value of the target in S1. Open symbols: target $c = c_{\mathrm{low}}$ in S1 and $c = c_{\mathrm{high}}$ in S2. Filled symbols: target $c = c_{\mathrm{high}}$ in S1 and $c = c_{\mathrm{low}}$ in S2. Lower plot: performance in the "different statistics" trials, as a function of the c-value of the target in S2; same convention for symbols. The two lower plots differ only in the positions of the points along the abscissa.

This suggests a modest contribution of a contextual pop-out (i.e., statistics that deviate from the end of the range typified by the pair $(c_{\text{low}}, c_{\text{high}})$ that define the block of trials).

When the statistics of the target do change, then the target c must necessarily equal one of the two paired values c_{low} or c_{high} in S1, and the other value in S2. Thus, the statistical change necessarily nulls any effect of contextual pop-out. Analysis of fraction correct as a function of the c-level of the target in S1 (Fig. 9(A), middle panel) or in S2 (Fig. 9(A), lower panel) show that there is no overall dependence of performance on the position of the target along the range (regression slopes of 0.052 and −0.003, respectively, both $p > 0.1$, F-test). However, for target c-values of −0.125, 0.125, and 0.375, fraction correct is higher for trials in which the target is darker in S2 than in S1, than for trials in which the target is brighter in S2 than S1 ($p < 0.05$ if comparison is based on target statistics in S1, $p < 0.01$ if comparison is based on target statistics in S2).

The isodipole experiments with $\Delta c = 1$ (Fig. 9(B)) show a very different pattern. For the "same statistics" trials (upper panel), there is a small but significant increase in fraction correct as a function of the c-value of the target (regression slope 0.049, $p < 0.05$, F-test). There is only one c-value, namely $c = 0$, that is tested both as $c = c_{\text{high}}$ in one pairing $[(c_{\text{low}}, c_{\text{high}}) = (-1, 0)]$ and also as $c = c_{\text{low}}$ in another pairing $[(c_{\text{low}}, c_{\text{high}}) = (0, 1)]$. A comparison of performance in these two kinds of trials yields no suggestion of an effect of context. For the "different statistics" trials (Fig. 9(B), middle panel and Fig. 9(B), lower panel), there is a clear dependence of fraction correct on the c-value of the target. Regression analysis shows that this dependence is predominantly or exclusively related to the c-value of the target in S2 (regression slope 0.115, $p < 10^{-4}$, F-test), not S1 (regression slope −0.026, $p > 0.1$, F-test).

We also analyzed the isodipole experiment with $\Delta c = 0.4$ (Fig. 7) along these lines. Consistent with the above observations, there was a trend of similar magnitude towards greater fraction correct as the c-value of the target increased (regression slope 0.040 for the "same statistics" condition, 0.065 for the "different statistics" condition, each $p < 10^{-2}$, F-test). Since the c-values of the target in S1 and S2 were similar ($\Delta c = 0.4$), we could not separate the influence of the c-value of the target in S1 versus S2.

In sum, superimposed on the effect of whether target statistics change, there are additional influences of the position of the target's statistics along the range investigated. This dependence is remarkably different for the luminance and the isodipole experiments. In the luminance experiments, in "same statistics" trials, there a context effect: an increase in fraction correct when the target is closer to random than the typical array. In the "different statistics" trials, fraction correct is higher for targets that darken than for targets that brighten. In the isodipole experiments, there is an increase in fraction correct when the target is closer to the even end of the range, whether or not there is a change in statistics, and there is no effect of context.

5. Discussion

5.1. What kinds of image statistics are used in visual working memory?

In the experiments reported here, we used a modification of the Cornelissen and Greenlee (2000) task to demonstrate that visual working memory can make use not only of individual pixel values, but also of the statistical structure of the arrays. We considered three kinds of statistical structure: changes in luminance, changes in higher-order local correlations, and the presence or absence of bilateral symmetry. Only the first two kinds of statistical structure appeared to be useful as cues in this visual memory task. Surprisingly, removal or introduction of bilateral symmetry, despite is apparent salience and importance in visual tasks (Attneave, 1954; Tyler, 1995; Wenderoth, 1994), did not influence performance.

5.1.1. Improved performance is not due to stimulus set size

This finding cannot be accounted for by differences in the stimulus set size induced by the statistical structures we have introduced. The size of a stimulus set, weighted by the relative frequencies of the stimuli within the set, is naturally quantified (on a logarithmic scale) by its entropy. (For a general review of entropy and related concepts, see Cover and Thomas (1991).) The entropy of an ensemble of random $n \times n$ arrays, in which each check is assigned independently and with equal probability to one of two states, is $H_{\text{random}} = n^2$ bits, since there is one bit associated with each check's assignment. For an $n \times n$ array in which luminance statistics are controlled by the parameter c, the ensemble entropy is

$$H_{\text{lum}}(c) = -\frac{n^2}{\log 2}\left[\frac{1+c}{2}\log\left(\frac{1+c}{2}\right) + \frac{1-c}{2}\log\left(\frac{1-c}{2}\right)\right], \quad (1)$$

since each of the n^2 checks are independently assigned to two states, with probability $\frac{1+c}{2}$ and $\frac{1-c}{2}$. (Note that $H_{\text{lum}}(0) = H_{\text{random}}$.) We have seen (Fig. 5) that a change of c from 0 to 0.25 yields a significant cue in the working memory task. As seen from Eq. (1), this is only a modest reduction in entropy: $H_{\text{lum}}(0.25) \approx 0.95 H_{\text{random}}$. On the other hand, a change from a random texture to a completely symmetric $n \times n$ square array of checks fails to yield a cue to the working memory task (Fig. 2).

However, the ensemble entropy is reduced from H_{random} by a factor of 2 to $H_{\text{symm}} = \frac{n^2}{2}$ bits, since each check on one half of the array can be assigned with equal likelihood to one of two states. That is, even though the symmetry manipulation associated with $c = 1$ induces a much greater reduction in randomness than the luminance manipulation associated with $c = 0.25$, only the latter cue is available to improve performance in the working memory task.

5.1.2. An information-theoretic measure of statistical difference

Next we determine the extent to which the *differences* between the ensembles induced by the change in statistics, rather than the size of the ensembles themselves, might account for the observed performance. A natural notion of the differences between two statistical ensembles is the Kullback–Leibler divergence (Cover & Thomas, 1991). The Kullback–Leibler divergence $D_{\text{KL}}(P\|Q)$ is an information-theoretic measure of the extent to which stimuli drawn from an ensemble P have probabilities that are unanticipated if it is expected that they were drawn from an ensemble Q. The Kullback–Leibler divergence has an interpretation in terms of the performance of an ideal observer on a task that is conceptually related to the one we used here: it measures how readily an ideal observer can determine, by observing sample arrays from an ensemble P, that these samples did *not* come from an ensemble Q. This measure is based only on the extent to which arrays have unequal probabilities in the two ensembles, and not on the spatial or geometrical aspects of their structure. (However, we note that this is not an ideal-observer analysis of the task we used. For any memory task, ideal-observer performance would be perfect. That is, the Kullback–Leibler divergence does not address how well an ideal observer should perform on our task, but rather, the magnitude of the statistical cue.)

The Kullback–Leibler divergence $D_{\text{KL}}(P\|Q)$ is defined by

$$D_{\text{KL}}(P\|Q) = \sum_i p_i \log\left(\frac{p_i}{q_i}\right), \tag{2}$$

where p_i is the probability of the ith stimulus in the ensemble P, q_i is the probability of the ith stimulus in the ensemble Q, and the sum is over all stimuli. The Kullback–Leibler divergence is evidently not symmetric in P and Q, and is infinite if any stimuli occur in P but not in Q. For both reasons, it is customary to use a symmetrized form of $D_{\text{KL}}(P\|Q)$, the "resistor average" (Johnson, Gruner, Baggerly, & Seshagiri, 2001), to measure the difference between the distributions P and Q. The resistor-average divergence $D_{\text{RA}}(P\|Q)$ is twice the harmonic mean of $D_{\text{KL}}(P\|Q)$ and $D_{\text{KL}}(Q\|P)$, and is defined by

$$\frac{1}{D_{\text{RA}}(P\|Q)} = \frac{1}{D_{\text{KL}}(P\|Q)} + \frac{1}{D_{\text{KL}}(Q\|P)}, \tag{3}$$

with the understanding that if either Kullback–Leibler divergence ($D_{\text{KL}}(P\|Q)$ or $D_{\text{KL}}(Q\|P)$) is infinite, then the corresponding term in the above equation is set to zero.

The Kullback–Leibler divergences, and hence the resistor-average divergences, are readily calculated for the three series of stimuli used, for any pair of levels of structure. For the luminance series, the calculation is most straightforward. The assignments of states of the n^2 checks are completely independent. For an ensemble of arrays P_{lum} characterized by a level of structure c_P, the probabilities of the states are $\frac{1+c_P}{2}$ and $\frac{1-c_P}{2}$. Correspondingly, in an ensemble of arrays Q_{lum} characterized by a level of structure c_Q, the probabilities of the states are $\frac{1+c_Q}{2}$ and $\frac{1-c_Q}{2}$. Since these assignments are made independently at each of the n^2 checks, we find from Eq. (2) that

$$D_{\text{KL}}(P_{\text{lum}}\|Q_{\text{lum}}) = \frac{n^2}{\log 2}\left(\frac{1+c_P}{2}\log\frac{1+c_P}{1+c_Q} + \frac{1-c_P}{2}\log\frac{1-c_P}{1-c_Q}\right). \tag{4}$$

For the isodipole series, the states of the $2n-1$ checks in the first row and first column are assigned randomly and with equal probability to the two states, while the assignment of the interior checks requires $(n-1)^2$ independent choices biased by the value of c_P or c_Q. This leads to

$$D_{\text{KL}}(P_{\text{isodipole}}\|Q_{\text{isodipole}}) = \frac{(n-1)^2}{\log 2}\left(\frac{1+c_P}{2}\log\frac{1+c_P}{1+c_Q} + \frac{1-c_P}{2}\log\frac{1-c_P}{1-c_Q}\right). \tag{5}$$

Finally, for the symmetry series, the states of the $\frac{n^2}{2}$ checks in one half of the array are assigned at random, while the assignment of the checks on the opposite half of the array requires $\frac{n^2}{2}$ independent choices biased by the value of c_P or c_Q. Thus,

$$D_{\text{KL}}(P_{\text{symm}}\|Q_{\text{symm}}) = \frac{n^2}{2\log 2}\left(\frac{1+c_P}{2}\log\frac{1+c_P}{1+c_Q} + \frac{1-c_P}{2}\log\frac{1-c_P}{1-c_Q}\right). \tag{6}$$

5.1.3. Does performance depend on degree of statistical difference?

The above Eqs. (4)–(6), combined with the symmetrization of Eq. (3), yields a natural (but purely statistical) measure of the extent to which each of the pairs of ensembles differ. In Fig. 10, we compare this measure with the experimental findings. The divergence between the fully symmetric and fully random arrays was larger than that between all of the pairings of the luminance

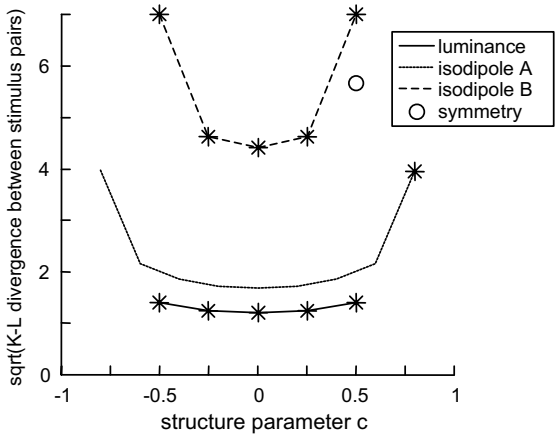

Fig. 10. A summary of the Kullback–Leibler divergences between the ensemble pairings studied in Experiment II (luminance and isodipole) and Experiment I (symmetry). For each pair of c-values studied, the ordinate shows the square root of the resistor-average divergences, Eq. (3), calculated from Eqs. (4)–(6), with $c_P = c_{\text{low}}$ and $c_Q = c_{\text{high}}$. The abscissa is the mean c-value, $\frac{c_{\text{low}}+c_{\text{high}}}{2}$. The asterisks mark the pairings for which a significant difference in performance between performance in "same statistics" and "different statistics" trials was observed. For the luminance series (Fig. 5), $\Delta c = |c_{\text{high}} - c_{\text{low}}| = 0.25$. For isodipole series A of Experiment II (Fig. 7), $\Delta c = 0.4$. For isodipole series B of Experiment II (Fig. 8) and the symmetry experiment (Fig. 2(C)), $\Delta c = 1$.

series with $\Delta c = 0.25$, but only the latter pairings provided a cue in the working memory task. Additionally, all of the Kullback–Leibler divergences in the isodipole pairings are substantially larger than those in the luminance series, but only the pairings in Experiment IIB ($\Delta c = 1$) and the most "even" pairing of Experiment IIA lead to improved performance in the "different statistics" condition. Finally, although many of the Kullback–Leibler divergences in the isodipole pairings of Experiment IIB are *smaller* than those of the symmetry experiment, only the former reveal an effect of image statistics.

Much as with ideal observer analyses in other contexts (e.g., (Geisler, 1984)), the discrepancies between our observations and those anticipated from the Kullback–Leibler analysis reveal the limitations (and strategies) of the visual system. The above analysis shows that the spatial arrangement of the correlations, and not merely their statistical strength, is crucial for the representation in visual working memory. This is entirely in parallel with the role of image statistics in texture segmentation and discrimination (Julesz, 1981a, 1981b; Julesz et al., 1973, 1978; Victor & Conte, 1991). Differences in power spectra (second-order spatial correlation structure) are potent cues to texture discrimination and segmentation. However, only very specific higher-order spatial correlations (as manifest by isodipole textures) can support these processes efficiently.

Even within the isodipole texture series, the perceptual consequences of a statistical change do not correspond to the Kullback–Leibler divergence. In the isodipole experiments of Figs. 7 and 8, the "even" end of the range (c near 1) showed a larger effect of a change in image statistics than the "odd" end of the range (c near −1). This asymmetry is not manifest in the Kullback–Leibler divergences of Fig. 10, which are independent of the sign of c. The greater salience of visual structure for positive values of c compared to negative values observed here corresponds to the larger size of the visual evoked potential elicited by interchange between even and random isodipole textures, compared with that elicited by interchange between odd and random isodipole textures (Victor & Conte, 1989a). We suspect that this asymmetry reflects the fact that the arrays at the even end of the range contain large homogenous patches (blobs), as well as extended contours (Victor & Conte, 1989a, 1989b, 1991)—features that are important in early visual processing (Field, Hayes, & Hess, 1993; Julesz, 1991; Kovacs & Julesz, 1993).

On the other hand, the Kullback–Leibler analysis may account for the interaction of n^2, the number of checks in the array and the size of the cue due to array statistics. As seen from Fig. 2, overall performance is better for the 16-element arrays than for the 64-element arrays for all statistical modalities studied. This main effect of array size could have multiple explanations, including a reduced load on a pixel-based mechanism. More significantly, in the two modalities for which there is a difference in performance between the "same statistics" and "different statistics" (Fig. 2(A) and (B)), there is *less* of a difference for the smaller arrays. This is anticipated from statistical considerations, since (as seen from Eqs. (4)–(6)) for smaller arrays (smaller n), it is more difficult to determine the ensemble from which an array is drawn. A ceiling effect may also contribute to the smaller effect of the statistical cue, but this is unlikely to be the full explanation, since performance in the isodipole experiments, even at the smaller array size (Fig. 2(B)), was worse than that in the luminance experiments (Fig. 2(A)).

5.1.4. Possible alternative explanations

The main message of the above analysis is that it allows for a comparison across statistical classes on an equal footing. Not surprisingly, the difference in luminance statistics needed to provide a cue to visual working memory is much smaller than the difference in isodipole statistics. But, given the salience of symmetry in other contexts (Attneave, 1954; Tyler, 1995; Wenderoth, 1994), it is particularly striking that a maximal difference in symmetry (and one which, on a statistical basis, is substantially larger than the threshold difference for isodipole statistics) provides no cue whatsoever.

One possibility that might account for this result is that symmetry is not fully processed within the constraints of our stimulus presentation (four simultaneous targets presented for 600 ms in S1) and interstimulus

interval (200 ms). Recent experiments support a contributing role for this factor: identification of a symmetric array is improved when the arrays are presented in RSVP fashion (Conte, Purpura, & Victor, 2002) compared to simultaneous presentation, or, when the interstimulus interval is increased (Conte & Victor, 2003)—though substantial differences between detection of symmetry and the other kinds of statistics persist even with processing times up to 1 sec. Along with the constraints implied by symmetry detection on multiple axes (Wenderoth, 1994) and the distinctive scaling behavior of symmetry perception (Rainville & Kingdom, 1999, 2002; Tyler, 1999), the slow dynamics of symmetry analysis compared with the more rapid analysis that suffices to extract luminance and isodipole statistics suggests that processing of symmetry is not based on a cascading of signals through a hardwired circuit, but more likely, reflects an iterative hypothesis-testing visual routine (Hayhoe, 2000).

The distinctive behavior of symmetry cannot be fully accounted for by detection differences. We recently performed (Victor, Hardy, & Conte, 2002) detection experiments based on a similar stimulus set, but a shorter presentation time (100 ms). Performance levels attained at $|\Delta c| = 1$ for symmetry ($\approx$40% correct) were typically achieved for isodipole statistical changes of $|\Delta c| = 0.4$, and luminance changes of $|\Delta c| < 0.1$. Thus, detection differences may account for the much smaller value of $|\Delta c|$ that is required for an improvement on the memory task for luminance changes, as compared with isodipole statistical changes. However, a change in symmetry ($|\Delta c| = 1$) that does not lead to an improvement on the memory task (Experiment I) is, by this measure, as detectable as a change in isodipole statistics ($|\Delta c| = 0.4$), that does lead to an improvement in performance (Experiment II). The possibility that detection differences coupled with differences in processing dynamics combine to account for what appears to be a difference in visual working memory is unlikely. This is because with the longer presentation times used here, detection differences are less prominent.

5.2. Relation to models of texture segregation

Many computational models for texture segregation have been proposed, with generally similar structure: an initial local stage, usually consisting of Gabor-like elements and perhaps local non-linear processing followed by a second stage, in which local signals are pooled, perhaps also in a non-linear fashion (Bergen & Adelson, 1988; Chubb & Landy, 1991; Graham, 1989; Graham, Beck, & Sutter, 1992; Grossberg & Mingolla, 1985; Malik & Perona, 1990; Victor & Conte, 1991; Wilson, 1993; Zhu et al., 1998). Such models have been successful in accounting for a wide range of texture discrimination phenomena.

In texture discrimination studies, performance is assayed by asking the subject to segment an image. This is, fundamentally, a statistical task: the outputs of the local filters *must* be pooled in order to identify gradients or discontinuities. Here, the subject is asked to determine whether an image has changed. Presumably, the same early visual mechanisms ("local filters") are used in both tasks. In principle, the outputs of these local filters could be retained individually, rather than collectively. Were this the case, there would be no difference in performance between conditions in which the statistical structure changed, and in which it did not—since there is the same degree of local change in each case. Thus, our finding that certain kinds of overall statistical structure provide a cue in a visual memory task indicates that the pooled signal is the basis not only for spatial comparisons, but also for comparisons across time.

5.3. Texture is represented continuously, not categorically

In many sensory and perceptual domains, a stimulus space that spans a continuum is represented in terms of discrete categories. Among non-visual domains, experimental evidence for categorical perception has been found in the processing of phonemes (Aaltonen, Niemi, Nyrke, & Tuhkanen, 1987; Pisoni & Lazarus, 1974) and somatosensation (Romo, Merchant, Zainos, & Hernandez, 1997). Within vision, evidence for categorical perception has been seen for color (Amano et al., 2002; Berlin & Kay, 1969; Bornstein & Korda, 1984; Wandell, 1985), orientation (Rosielle & Cooper, 2001), facial expression (Roberson & Davidoff, 2000), and animal form (Freedman, Riesenhuber, Poggio, & Miller, 2001, 2002).

Consider a discrimination task within a domain parameterized by a parameter c, in which the subject is to discriminate stimuli A and B (characterized, respectively, by c_A or c_B). In this setting, the hallmark of a categorical representation (Aaltonen et al., 1987; Bornstein & Korda, 1984; Friedman-Hill, Robertson, & Treisman, 1995; Pisoni & Lazarus, 1974; Wandell, 1985) is that there is a jump in fraction correct and a decrease in reaction time when c_A and c_B straddle a category boundary, compared to performance with the same $\Delta c = |c_A - c_B|$ when c_A and c_B are within the same category. Conversely, with a graded representation, performance depends primarily on $|c_A - c_B|$, but not on the particular values of c_A or c_B.

One might speculate that as indices of texture-like surface properties, image statistics may play a role similar to that of color (Cho et al., 2000; Harvey & Gervais, 1981). However, our results provide no support for a categorical representation of texture, either by the fraction correct or reaction time criteria. While the data of Fig. 7 ($\Delta c = 0.4$) raise the possibility that the pure even ($c = 1$) texture is in a category by itself, and that all

statistics corresponding to $c < 1$ are equivalent, the data of Fig. 8 ($\Delta c = 1$) show that this is a manifestation of a threshold. One possible basis of this difference is that categorical perception derives not from a functional role in surface (Cho et al., 2000; Harvey & Gervais, 1981) or object (Freedman et al., 2001, 2002; Rosielle & Cooper, 2001) identification, but rather, arises from a graded stimulus domain via verbal coding or storage of stimuli (Roberson & Davidoff, 2000). The latter authors showed, in a visual memory task, that verbal interference removed the hallmarks of categorical perception, both for color and facial expression. The rather abstract nature of our stimuli may have precluded such verbal processes, and thus, revealed an underlying graded representation.

Acknowledgements

Portions of this work were presented at the 2001 meeting of the Society for Neuroscience and the 2002 meeting of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, FL, and was supported by NIH NEI EY7977. We thank Caitlin Hardy for assistance with data collection and Jeff Tsai for programming assistance.

Appendix A. Details of stimulus construction

Here we describe the construction of stimulus arrays across the range of the structure parameter c, for the three kinds of statistical structure studied: luminance, isodipole textures, and symmetry. We have two goals: first, generation of an array that has a prescribed value of c (for use in S1), and second, generation of a second array (for use in S2) that has a second prescribed value of c, and in which only a prescribed number of checks have changed in luminance.

A.1. Luminance statistics

For the luminance experiments, the value of c determines the number of checks that are white ($\frac{1+c}{2}$) and the number of checks that are black ($\frac{1-c}{2}$). Since the total number of checks must be an integer (N), these ratios can only achieve certain discrete values. For the experiments described here, we only used values of c for which these ratios could be achieved exactly—that is, values of c from -1 to 1, in steps of $\frac{2}{N}$. Once c is specified, the number of white and black checks is then specified, as $N_w = N(\frac{1+c}{2})$ and $N_b = N(\frac{1-c}{2})$, respectively. Arrays were then constructed by random placement of the requisite number of white and black checks.

If k of the N checks change in luminance between S1 and S2, the maximum increase in c that can occur is $\frac{2k}{N}$. This occurs if all k of the altered checks change from black to white. Similarly, the maximal decrease in c that can occur is a change of $-\frac{2k}{N}$, which happens if all k altered checks change from white to black. More generally, changing the state of $\frac{k}{2} + \frac{N\Delta c}{4}$ checks from black to white and $\frac{k}{2} - \frac{N\Delta c}{4}$ checks from white to black (a total of k checks) leads to a net change in c of Δc.

For all of the experiments described here, the above quantities are non-negative integers. To create the S2 array from the S1 array, the number of black checks and white checks to be flipped in contrast is determined by the above rules, and their location is determined at random from the location of the N_b black checks and the N_w white checks in S1.

A.2. Isodipole statistics

For the other kinds of statistics, the goal was the same, but the details differ. For the arrays based on the even and odd isodipole textures (Fig. 2(B)), we proceeded as follows. In an even ($c = 1$) or odd ($c = -1$) isodipole texture array (Julesz et al., 1978), the state $a_{i,j}$ (+1 or −1) of the check in the ith row and jth column is forced to obey the recursion rule

$$a_{i,j}a_{i-1,j}a_{i,j-1}a_{i-1,j-1} = c. \tag{A.1}$$

This rule, along with a random assignment of the states of checks in the initial row ($a_{1,j}$) and initial column ($a_{i,1}$), generates a texture in which half of the checks (on average) are in either state, and in which there are no pairwise or third-order correlations. Isodipole textures with intermediate values of c were constructed according to the "propagated decorrelation" textures of Victor (1985). For these textures, the deterministic rule (A.1) is replaced by the probabilistic rule

$$\text{prob}\{a_{i,j}a_{i-1,j}a_{i,j-1}a_{i-1,j-1} = 1\} = \frac{c+1}{2}. \tag{A.2}$$

Note that since the value of the quadruple product in the above expression must be either +1 or −1, Eq. (A.2) reduces to Eq. (A.1) for $c = 1$ or $c = -1$. For intermediate values of c, the average value of the product in Eq. (A.2) (over an infinite sample of the texture) is c.

For an $n \times n$ square array of N elements, there are $(n-1)^2$ instances at which the probabilistic rule (A.2) can be applied: namely, all (i,j) pairs for which $2 \leqslant i \leqslant n$ and $2 \leqslant j \leqslant n$. Consequently, we replace the probabilistic rule of Eq. (A.2) by the deterministic rule

$$\frac{1}{(n-1)^2}\sum_{i=2}^{n}\sum_{j=2}^{n} a_{i,j}a_{i-1,j}a_{i,j-1}a_{i-1,j-1} = c. \tag{A.3}$$

Again, since the value of the quadruple product in the above expression must be either +1 or −1, this average can only have specific values, ranging from −1 to +1, in steps of $\frac{2}{(n-1)^2}$. To construct an array corresponding to a

particular value of c that is intermediate between these achievable values, we used either the next-highest exactly achievable value, c_{above}, or the next-lowest exactly achievable value, c_{below}, to determine exactly how many of the recursion products in Eq. (A.3) are equal to +1, and exactly how many are equal to −1. We select either c_{above} or c_{below} with probabilities p_{above} and p_{below} to ensure that the average value of the recursion products in Eq. (A.3) is exactly c. That is,

$$p_{\text{above}} = \frac{c - c_{\text{below}}}{c_{\text{above}} - c_{\text{below}}} \quad \text{and}$$
$$p_{\text{below}} = \frac{c_{\text{above}} - c}{c_{\text{above}} - c_{\text{below}}}. \tag{A.4}$$

For the isodipole textures, changing the state of k checks can alter the value of c in Eq. (A.3) by as much as $\frac{8k}{(n-1)^2}$, since each of the k checks can participate in up to four products, each of which can have the value $\frac{1}{(n-1)^2}$ or $-\frac{1}{(n-1)^2}$. However, since the checks involved in these quadruple products overlap, it is not straightforward to construct a pair of arrays that have prescribed values of c, and which differ at exactly k checks. The first step in doing this was to construct two arrays with the desired values of c without any constraints on the number of checks at which they differed. Since these arrays were constructed independently, they typically mismatched at approximately $\frac{N}{2}$ of the checks. Thus, the typical number of mismatches in the independently constructed arrays is larger than the number of desired mismatches, k. Then, we sought to apply a sequence of flips of entire rows and/or columns of checks to one of the arrays, so that the number of mismatched checks would equal k. Note that flipping entire rows or columns does not change the value of c, since these transformations invert the state of two checks in every affected quadruple product. This search combined a Monte Carlo strategy and strategies used in the solution of Berlekamp's switching game (Fishburn & Sloane, 1989): essentially, a game whose goal is to minimize the number of mismatches. We then sought (by iterating this strategy) to combine these pairs into quadruples of arrays (A_1, A_2, B_1, B_2) such that (i) A_1 and A_2 corresponded to c_A, (ii) B_1 and B_2 corresponded to c_B, (iii) the pairs (A_1, A_2), (A_2, B_1), and (B_1, B_2) each differed by k flips. A library of such quadruples was created via repeating this search procedure. The library was further enlarged by (a) flipping all the arrays within a quadruple along their horizontal and/or vertical axes, and (b) randomly multiplying all the arrays within a quadruple by a randomly chosen even $(c = 1)$ array. These transformations preserve the conditions (i), (ii), and (iii) above, and also ensured that the positions of the checks that were flipped between S1 and S2 were symmetrically distributed, and that the stimuli themselves were approximately luminance-balanced.

To generate the stimuli for individual trials, each array in S1 was drawn from the middle elements (A_2 and B_1) of a quadruple randomly selected from this library. (Each array was drawn from an independently constructed quadruple.) To create the S2 target in a "different statistics" trial, an A_2 array in S1 is replaced in S2 by the array B_1 drawn from *the same* quadruple, or a B_1 array in S1 is replaced in S2 by the A_2 array from *the same* quadruple. To create the S2 target in a "same statistics" trial, an A_2 array in S1 is replaced by an A_1 array in S2, or a B_1 array in S1 is replaced by a B_2 array in S2.

The reason for the elaborate construction described above is that it ensures that the exemplar chosen for an array in S1 gives no indication as to whether it is a target, nor whether the trial will be a "different statistics" trial or not. It leaves open the possibility that some overall difference between the A_1 arrays and the A_2 arrays, or between the B_1 arrays and the B_2 arrays, might provide an extra cue in S2 for the "same statistics" trial that does not require comparison with S1. However, this kind of artifact would produce the opposite of the result we observed. Additionally, such cues could not be identified by the investigators, who had full knowledge of the construction, even after extended scrutiny of the textures. We also verified that the changed and unchanged checks were approximately uniformly distributed throughout the arrays, and that nearly all arrays were within 1 or 2 checks of being perfectly luminance-balanced, even though we did not explicitly balance luminance.

A.3. Symmetry

We constructed the arrays based on symmetry (Fig. 2(C)–(E)) as follows. An array with perfect $(c = 1)$ twofold symmetry (with either a horizontal or a vertical mirror axis) can be constructed by randomly assigning the luminance values to the $\frac{N}{2}$ checks in one half of the array. Each of these checks is then paired (via symmetry) with another check, and the luminance assigned to this paired check must match to preserve the symmetry. Graded amounts of symmetry correspond to an array in which a fraction $(\frac{c+1}{2})$ of the paired checks match, and the rest of the pairs mismatch. Thus, for $c = 0$, exactly half of the pairs match, for $0 < c < 1$, more pairs match than would be expected by chance, and for $c = -1$, all of the paired checks mismatch. Changing the state of k checks can change the match versus mismatch state of up to k of the $\frac{N}{2}$ pairs, and thus, change the value of c by amounts ranging from $-\frac{4k}{N}$ to $\frac{4k}{N}$ (in steps of $\frac{8}{N}$) in either direction. Thus, to change array statistics from perfect mirror symmetry $(c = 1)$ to mirror symmetry with luminance inversion $(c = -1)$, as in Fig. 2(E) (corresponding to $\Delta c = 2$), half of the checks in the target array must be flipped between S1 and S2. This cannot be

accomplished with $N = 64$ checks and $k = 16$ flips, but only with the smaller array size ($N = 16$ checks, $k = 8$ flips).

To create arrays in which changing k checks resulted in a switch from vertical symmetry to horizontal symmetry (Fig. 2(D)), we randomly assigned luminance values to the upper left quadrant of the array ($\frac{N}{4}$ checks). Each of these checks was used to determine the value of the three other checks that were related by horizontal and vertical mirror operations. Quadruples of checks with configurations $\pm\begin{pmatrix} +1 & +1 \\ +1 & +1 \end{pmatrix}$ contributed $\frac{4}{N}$ to both vertical and horizontal symmetry; quadruples with configurations $\pm\begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix}$ contributed $-\frac{4}{N}$ to both vertical and horizontal symmetry; quadruples with configurations $\pm\begin{pmatrix} +1 & +1 \\ -1 & -1 \end{pmatrix}$ contributed $\frac{4}{N}$ to vertical symmetry but $-\frac{4}{N}$ to horizontal symmetry; quadruples with configurations $\pm\begin{pmatrix} +1 & -1 \\ +1 & -1 \end{pmatrix}$ contributed $-\frac{4}{N}$ to vertical symmetry but $\frac{4}{N}$ to horizontal symmetry; and the eight other quadruples with configurations like $\pm\begin{pmatrix} +1 & -1 \\ +1 & +1 \end{pmatrix}$ and its rotations contributed 0 to vertical and horizontal symmetry. By choosing appropriate fractions of these configurations in S1 and S2, the desired amounts of symmetry and number of checks that were flipped were obtained.

All of the trials made use of values of c (0, 1 and −1) that could be achieved exactly. Moreover, since N was a multiple of 16 and k was a multiple of 8, it was possible to arrange the luminance assignments so that S1 and S2 were both precisely luminance-balanced, for the stimuli involving only vertical symmetry (Fig. 2(C) and (E)) as well as those with the additional constraint of horizontal symmetry (Fig. 2(D)).

References

Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkanen, M. (1987). Event-related brain potentials and the perception of a phonetic continuum. *Biological Psychology, 24*(3), 197–207.

Amano, K., Uchikawa, K., & Kuriki, I. (2002). Characteristics of color memory for natural scenes. *Journal of the Optical Society of America A—Optics Image Science and Vision, 19*(8), 1501–1514.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review, 61*, 183–193.

Bergen, J. R., & Adelson, E. H. (1988). Early vision and texture perception. *Nature* (333), 363–364.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.

Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research, 46*(3), 207–222.

Caelli, T., & Julesz, B. (1978). On perceptual analyzers underlying visual texture discrimination: Part I. *Biological Cybernetics, 28*(3), 167–175.

Caelli, T., Julesz, B., & Gilbert, E. (1978). On perceptual analyzers underlying visual texture discrimination: Part II. *Biological Cybernetics, 29*(4), 201–214.

Cho, R. Y., Yang, V., & Hallett, P. E. (2000). Reliability and dimensionality of judgments of visually textured materials. *Perception & Psychophysics, 62*(4), 735–752.

Chubb, C., & Landy, M. (1991). Orthogonal distribution analysis: A new approach to the study of texture perception. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 291–301). Cambridge, MA: MIT Press.

Conte, M. M., Purpura, K. P., & Victor, J. D. (2002). Processing of image symmetry in an RSVP task. *Society for Neuroscience*, 28, Orlando, FL.

Conte, M. M., & Victor, J. D. (2004). Temporal stability of image statistics in visual working memory. *Vision Sciences Society 2003*, Sarasota, FL.

Cornelissen, F. W., & Greenlee, M. W. (2000). Visual memory for random block patterns defined by luminance and color contrast. *Vision Research, 40*(3), 287–299.

Cover, T. M., & Thomas, J. A. (1991). Elements of information theory. In *Wiley series in telecommunications* (p. 542). New York: Wiley.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research, 33*(2), 173–193.

Fishburn, P. C., & Sloane, N. J. A. (1989). The solution to Berlekamp's switching game. *Discrete Mathematics, 74*, 263–290.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 291*(5502), 312–316.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2002). Visual categorization and the primate prefrontal cortex: Neurophysiology and behavior. *Journal of Neurophysiology, 88*(2), 929–941.

Friedman-Hill, S. R., Robertson, L. C., & Treisman, A. (1995). Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. *Science, 269*(5225), 853–855.

Geisler, W. S. (1984). Physical limits of acuity and hyperacuity. *Journal of the Optical Society of America A, 1*(7), 775–782.

Graham, N. (1989). *Visual pattern analyzers*. Oxford: Clarendon Press.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research, 32*(4), 719–743.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics, 38*(2), 141–171.

Harvey, L. O., Jr., & Gervais, M. J. (1981). Internal representation of visual texture as the basis for the judgment of similarity. *Journal of Experimental Psychology—Human Perception and Performance, 7*(4), 741–753.

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition, 7*(1/2/3), 43–64.

Johnson, D. H., Gruner, C. M., Baggerly, K., & Seshagiri, C. (2001). Information-theoretic analysis of neural coding. *Journal of Computational Neuroscience, 10*(1), 47–69.

Joseph, J. S., & Optican, L. M. (1996). Involuntary attentional shifts due to orientation differences. *Perception & Psychophysics, 58*(5), 651–665.

Julesz, B. (1981a). Textons, the elements of texture perception, and their interactions. *Nature, 290*(5802), 91–97.

Julesz, B. (1981b). A theory of preattentive texture discrimination based on first-order statistics of texture. *Biological Cybernetics* (41), 131–138.

Julesz, B. (1991). Early vision and focal attention. *Reviews of Modern Physics, 63*(3), 735–772.

Julesz, B., Gilbert, E. N., Shepp, L. A., & Frisch, H. L. (1973). Inability of humans to discriminate between visual textures that agree in second-order statistics—revisited. *Perception, 2*(4), 391–405.

Julesz, B., Gilbert, E. N., & Victor, J. D. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics, 31*(3), 137–140.

Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences USA, 90*(16), 7495–7497.

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A, 7*(5), 923–932.

Marr, D. (1982). *Vision*. New York: W.H. Freeman & Co.

Pisoni, D. B., & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America, 55*(2), 328–333.

Purpura, K. P., Victor, J. D., & Katz, E. (1994). Striate cortex extracts higher-order spatial correlations from visual textures. *Proceedings of the National Academy of Sciences USA, 91*(18), 8482–8486.

Rainville, S. J., & Kingdom, F. A. (1999). Spatial-scale contribution to the detection of mirror symmetry in fractal noise. *Journal of the Optical Society of America A—Optics Image Science and Vision, 16*(9), 2112–2123.

Rainville, S. J., & Kingdom, F. A. (2002). Scale invariance is driven by stimulus density. *Vision Research, 42*(3), 351–367.

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition, 28*(6), 977–986.

Romo, R., Merchant, H., Zainos, A., & Hernandez, A. (1997). Categorical perception of somesthetic stimuli: Psychophysical measurements correlated with neuronal events in primate medial premotor cortex. *Cerebral Cortex, 7*(4), 317–326.

Rosielle, L. J., & Cooper, E. E. (2001). Categorical perception of relative orientation in visual object recognition. *Memory & Cognition, 29*(1), 68–82.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology—Human Perception and Performance, 8*(2), 194–214.

Tyler, C. W. (1995). Empirical aspects of symmetry perception. *Spatial Vision, 9*(1), 1–7.

Tyler, C. W. (1999). Human symmetry detection exhibits reverse eccentricity scaling. *Visual Neuroscience, 16*(5), 919–922.

Victor, J. D. (1985). Complex visual textures as a tool for studying the VEP. *Vision Research, 25*(12), 1811–1827.

Victor, J. D. (1986). Isolation of components due to intracortical processing in the visual evoked potential. *Proceedings of the National Academy of Sciences USA, 83*(20), 7984–7988.

Victor, J. D., & Conte, M. M. (1989a). Cortical interactions in texture processing: Scale and dynamics. *Visual Neuroscience, 2*(3), 297–313.

Victor, J. D., & Conte, M. M. (1989b). What kinds of high-order correlation structure are readily visible? *Investigative Ophthalmology & Visual Science, 30*(Suppl.), 254.

Victor, J. D., & Conte, M. M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research, 31*(9), 1457–1488.

Victor, J. D., & Conte, M. M. (1996). The role of high-order phase correlations in texture processing. *Vision Research, 36*(11), 1615–1631.

Victor, J. D., & Conte, M. M. (2001). Dynamics of selective spatial attention in a working memory task. *Association for Research in Vision and Ophthalmology*, 42 (p. 863). Ft. Lauderdale, FL.

Victor, J. D., Hardy, C., & Conte, M. M. (2002). Visual processing of image statistics: Qualitative differences between local and global statistics; quantitative differences between low- and high-order statistics. *Vision Sciences Society 2002*, Sarasota, FL.

Wandell, B. A. (1985). Color measurement and discrimination. *Journal of the Optical Society of America A, 2*(1), 62–71.

Wenderoth, P. (1994). The salience of vertical symmetry. *Perception, 23*(2), 221–236.

Wilson, H. R. (1993). Nonlinear processes in visual pattern discrimination. *Proceedings of the National Academy of Sciences USA, 90*(21), 9785–9790.

Zhu, S. C., Wu, Y., & Mumford, D. (1998). Filters, random fields and maximum entropy (FRAME): Towards a unified theory for texture modeling. *International Journal of Computer Vision, 27*(2), 107–126.