

# What do we compare when comparing separate objects?

M. V. Danilova

Laboratory for Vision Physiology, I. P. Pavlov Institute of Physiology, St. Petersburg, Russia

J. D. Mollon

Department of Experimental Psychology, University of Cambridge, Cambridge, UK

(Submitted March 10, 1999)

Opticheskiĭ Zhurnal **66**, 15–21 (October 1999)

Little is known about how well we can compare visual objects in the visual field and what are the mechanisms underlying such comparisons. In psychophysical experiments we studied the precision of such comparisons when discriminating spatial frequency or contrast and found that the precision is as good at 10° separation as it is when the stimuli are juxtaposed. This suggests that subjects rely on the same mechanism over the full range of separations. This mechanism is unlikely to depend on the lateral interactions in the primary cortex that are sometimes thought to underlie segregation of texture, since the density of such connections declines with distance. We suggest instead that the comparison is made at a level where stimulus properties are represented by an abstract code rather than by the activities of particular labeled lines. Our results rule out models in which the subject makes an absolute judgment of only one of the two stimuli, assessing it against a stored template or matched filter that represents the average stimulus over many trials; but we cannot rule out a model in which each of the two stimuli is individually assessed against a stored template and the subject's decision is based on the outcome of these two comparisons. © 1999 The Optical Society of America. [S1070-9762(99)00310-3]

## 1. INTRODUCTION

How well can we discriminate stimuli that are well separated in the visual field? This aspect of visual perception has not been studied systematically. Almost always the objects to be compared are either juxtaposed, as in classical photometry, or replace each other in time, as in, for example, the separation discrimination task.<sup>1</sup> But suppose that one patch appears several degrees to the left of fixation, and the observer's task is to compare its spatial frequency to that of a second patch presented several degrees to the right of fixation. Would performance deteriorate compared to the situation where the patches are juxtaposed or temporally interleaved? Recent psychophysical and physiological studies suggest that comparisons of stimuli separated by small and large distances may be carried out by different mechanisms.

### A. Local comparisons

A large body of psychophysical work on texture segmentation shows that rapid, pre-attentive segmentation occurs if the subarrays differ in the orientation, the spatial frequency, the size, or the color of their elements (see, for example, Refs. 2 and 3). Several authors have proposed a central representation of the visual field that records local discontinuities, whether they are discontinuities in orientation, color, brightness, or some other primary attribute—the “common representation” of Cavanagh, Arguin and Treisman<sup>4</sup> or the “salience map” of Lu and Sperling.<sup>5</sup> These psychophysical results for texture segmentation may be related to electrophysiological studies of the non-classical zones of receptive fields early in the visual system.<sup>6,7</sup> At the level of the lateral geniculate nucleus, Sillito *et al.*<sup>8</sup> found cells that give a stronger response if a central grating and a

surrounding annular grating were of different orientation; the effect was mediated by efferent cortical signals. Such psychophysical and physiological findings suggest that local comparisons occur at an early level where stimulus attributes may be represented by a small number of single neurons.

### B. Large separations

However, are the mechanisms that detect local discontinuities likely also to subserve comparison at a distance? The lateral connections in the primary visual cortex may certainly be rather extensive, sometimes extending for several millimeters: in the case of the cat, Gilbert and Wiesel<sup>9,10</sup> cite some examples of axons extending up to 8 mm, while slightly shorter lengths (3–4 mm) have been reported for the monkey.<sup>11,12</sup> The axonal terminations may cluster in regions of cortex with the same preferred orientation and spatial frequency.<sup>9,13</sup> Nevertheless, such lateral connections become systematically sparser with distance. This is suggested by consideration of the number of labeled boutons with increasing distance from the injection site: from the example of Bosking *et al.*,<sup>14</sup> it is clear that a doubling of distance led to a threefold decrease in the number of boutons.

If a subject is asked to compare two stimuli at different separations, and if the comparison does depend on lateral interactions in the primary visual cortex, then we might expect the precision of comparisons to deteriorate as separation increases. In the present experiments, we measured the efficiency with which separated objects are compared over distances of up to 10° of visual angle. In the first series of measurements, the task was one of spatial frequency discrimination. In the second, the task was contrast discrimination.

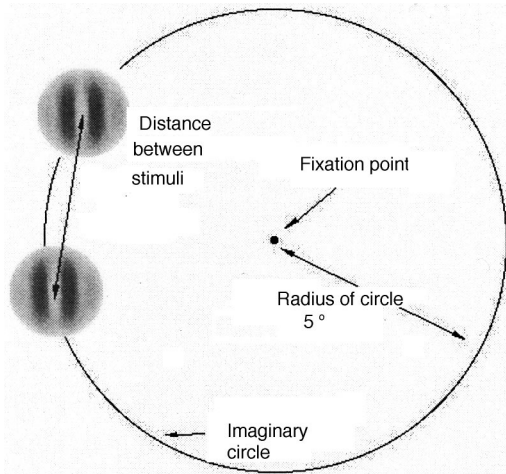


FIG. 1. A schematic example of the stimuli used in the experiments.

## 2. METHODS

### A. Stimuli

The stimuli were vertical Gabor patches presented on a monitor screen (SONY Trinitron, 21 in., with a refresh rate of 60 Hz). All parameters of the stimuli were controlled by a VSG 2/4 graphics board (Cambridge Research Systems). The display had an average luminance of 7.4 cd/m<sup>2</sup>. The screen resolution was 1280×960. At the 114-cm viewing distance, the size of one Gabor patch was 2°, and its sigma was equal to the period of the referent grating and was held the same in all experiments. The Gabors had the same space-averaged luminance as the background.

In order not to confound eccentricity and separation, we presented the two Gabor patches on an imaginary circle of 5° radius. The centers of the patches were always on the circle. The separations between the two Gabors were measured as the separations between the centers of the two patches. Thus, when the separation was 2°, the two patches were juxtaposed, and when their separation was 10°, they lay on a diameter of the imaginary circle. An example of the stimuli is given in Fig. 1. The circle represents an imaginary circle, on which the centers of the two compared Gabor elements lie. The arrowed line between the two Gabors shows the separation measured between their centers.

Of the two simultaneously presented stimuli, one was always of a fixed reference value. In the frequency discrimination task the referent stimulus was of 2 c/deg; its contrast was fixed at 0.3. The contrast of the test stimulus was jittered in the range ±0.03 in order to minimize contrast cues. In the contrast discrimination task, the referent stimulus was the same as in the first experiment, but the phase of the test stimulus was chosen randomly from the range between +90° and -90°. We introduced this jitter to remove any special advantage that might arise when the two vertical Gabor elements were close together and aligned (i.e., when they fell near 3 o'clock or 9 o'clock).

In the contrast discrimination task, we also measured the time between the stimulus offset and the moment when the observer pressed a button on the response box.

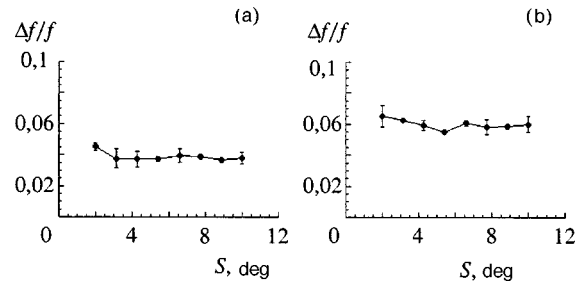


FIG. 2. Thresholds measured in the spatial frequency discrimination task. (a) observer MVD, (b) observer JDM.

### B. Procedure

The discrimination thresholds were measured with a staircase procedure. Two independent staircases were randomly interleaved to estimate the ascending and descending thresholds from the reference value. The measurements were made in the 3-1 scheme; i.e., three correct responses resulted in a change of the parameter toward the reference value, whereas one incorrect response caused the parameter to change in the opposite direction. A fixation point was always present at the center of the screen, between presentations and during the 100-ms stimulus interval. The observer's task was to report whether the more clockwise Gabor patch was of higher or lower spatial frequency or contrast than the other one.

On a given trial, the subject did not know which of the two stimuli was the referent, nor whether the trial was measuring a descending or an ascending threshold. These maneuvers obliged the subject to make an actual comparison of the two stimuli, because, when the reference stimulus is the more clockwise, the correct answer can be found only by consideration of the other stimulus (see Discussion).

In different blocks of trials, the separation varied between 2° and 10°. Several repetitions were accumulated for each condition on different experimental days. Each data point in the plots represent the mean values.

### C. Observers

Two highly trained observers (the two authors of the paper) participated in the experiments.

## 3. RESULTS

### A. Spatial frequency discrimination

For each separation  $S$  between stimuli, we measured two individual thresholds in interleaved ascending and descending staircases and repeated these measurements six times for each observer. The Weber fractions  $\Delta f/f$  for the ascending and descending staircases do not differ significantly. Figure 2 shows the Weber fractions for the two observers averaged over six experimental days and for ascending and descending staircases. Apart from the initial increase of discrimination

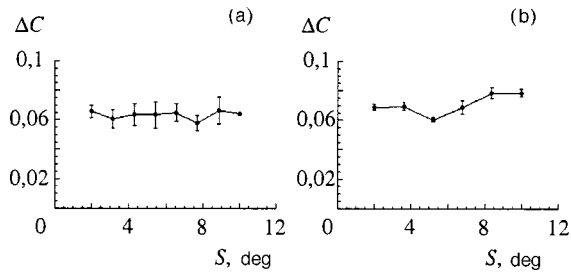


FIG. 3. Thresholds measured in the contrast discrimination task. (a) observer MVD, (b) observer JDM.

thresholds at 2° center-to-center separation (where the Gabor patches are adjacent), the curves are flat. An analysis of variance was performed to test the effect of separation. The analysis showed that separation was not a significant factor ( $F[7] = 1.24, P = 0.286$  for MVD;  $F[7] = 0.64, P = 0.721$  for JDM). The conclusion from this set of measurements is that, for separations up to 10°, there is no change in the precision with which we discriminate the spatial frequency.

### B. Contrast discrimination

Figure 3 shows mean threshold differences  $\Delta C$  from the referent contrast value (0.3) averaged over ascending and descending staircases. JDM's data were based on six repetitions, MVD's on eight. For observer MVD, the contrast discrimination threshold did not vary significantly with the spatial separation of the discriminanda ( $F[7] = 0.5438, P = 0.799$ ). For JDM, the effect of separation was significant ( $F[5] = 3.40, P = 0.0085$ ), reflecting an increased threshold at the two largest separations, but quantitatively this variation was small.

Figure 4 shows the reaction times measured in the contrast discrimination task. We did not ask the subject to respond as quickly as possible, so there was no specific time pressure. The mean reaction times for all conditions lie between 750 and 800 ms and do not vary with separation ( $F[5] = 1.764, P = 0.13$  for JDM;  $F[7] = 0.9399, P = 0.48$  for MVD), which would not be the case if the precision of judgment were gained in exchange for an increase in time needed to make the comparison.

## 4. DISCUSSION

The results of the two experiments show that the spatial frequency and the contrast of simultaneously presented Ga-

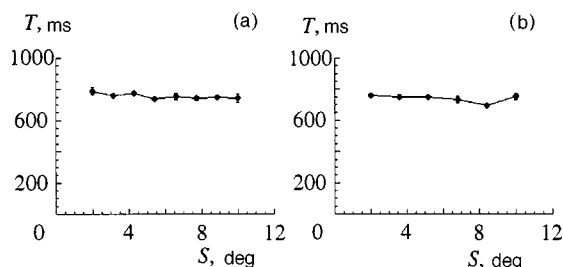


FIG. 4. Reaction times measured in the contrast discrimination task. (a) observer MVD, (b) observer JDM.

bor patches can be compared with similar precision independently of whether the two patches are juxtaposed or their centers are separated by 10° of visual angle. Before considering how this comparison is achieved, we must first establish that both targets are taken into account on each trial.

### A. Evidence that both stimuli are used

An analogy might be made between the present experiments and those that show that subjects can discriminate spatial frequencies with the same precision whether or not the reference and variable stimuli are presented simultaneously or are separated in time by many seconds or minutes.<sup>15,16</sup> In the case of such experiments, it has been suggested by Lages and Treisman<sup>17</sup> that the subject does not perform a true comparison between the two stimuli presented but instead judges the target stimulus against an internal criterion. According to Lages and Treisman, subjects are able to set and retain in memory "a specification for modifying the criterion used in later decisions." In a variant form of this model, one could suppose that the subject, over many trials, constructs an internal template, a neural representation of the average stimulus, and that it is against this template, rather than an external reference stimulus, that the target is judged. Such strategies might paradoxically be the more efficient, since the judgment would then depend on only one sample of external noise, whereas it will depend on two samples of external noise if the external reference is used.

However, the design of our experiments makes it unlikely that the observer considers only one of the two stimuli present on a trial. If he or she simply attends to the more clockwise target and makes a single absolute judgment, then he or she should score at chance when the more clockwise target is the reference frequency or contrast. In this case, the spatial frequency of the referent stimulus lies at the center of the frequency interval of all the presented stimuli—both the test and the referent stimuli—i.e., it corresponds to the chosen sample. For the correct response on such trials can be determined only by knowledge of the spatial frequency or contrast of the less clockwise stimulus, the variable stimulus. In Fig. 5, we have combined, for each observer, the data for different separations in the spatial-frequency task, but have distinguished those trials on which the variable is the more clockwise of the two stimuli and those trials on which the referent is the more clockwise. Figure 5 shows the psychometric functions thus obtained. The horizontal axis shows the spatial frequency  $f$  of the Gabor test element, and the vertical axis shows the percent  $L$  of the responses that are greater for stimuli that are more clockwise. Each psychometric function thus obtained was fitted with a single sigmoid curve using a standard fitting program (SigmaPlot). For both observers, the data for the two subsets of trials are approximately symmetrical, MVD's data being more closely symmetrical than those of JDM. The standard deviations of the two fitted functions are 0.0482 and 0.0437 for MVD and 0.0623 and 0.0923 for JDM. In other words, the observers are approximately as efficient when the reference stimulus is the more clockwise as they are when the target is the more clockwise. This result rules out the possibility that the observer merely attends to the more clockwise target and makes an absolute judgment,

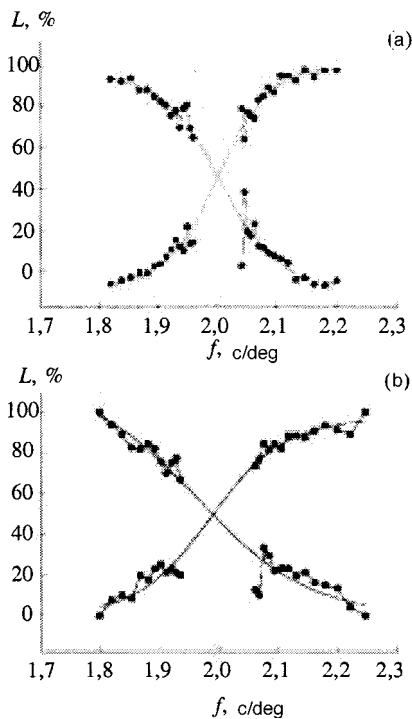


FIG. 5. Psychometric functions reconstructed from the staircase procedure that we used in the experiments. ●—referent stimulus has to be judged; ■—test stimulus was judged. (a) observer MVD, (b) observer JDM.

i.e., compares a single stimulus to a remembered template or criterion.

The data of Fig. 5 also rule out the possibility that the observer attends to only one of the two stimuli but attends randomly to the more clockwise or to the less clockwise: for the observers achieve nearly 100% correct performance when the referent is the more clockwise, provided the difference in spatial frequency is large enough.

### B. The nature of the comparison

From the argument of the preceding section, we conclude that the observer is taking both stimuli into account and thus, in some sense, is making a comparison. But what is the nature of this comparison, at a physiological and at a psychological level? Our primary experimental finding is that discrimination thresholds are nearly independent of separation. This result suggests that observers rely on the same mechanism over the full range of separations tested, and it places constraints on what the mechanism might be.

It is unlikely that the comparison depends on lateral interactions between spatial-frequency analyzers in primary cortex. Reciprocal inhibition between nearby cells might provide the basis for a local comparison, and might, say, underlie texture segregation, enhancing the difference between stimulus regions of different spatial-frequency content. But laterally extending fibers become sparser with distance in the cortex, and it would seem odd if such connections could sustain a constant threshold of discrimination over a range of  $10^\circ$  of visual angle. In Fig. 6, we caricature one specific model of this class, in which the comparison depends on specific connections between individual

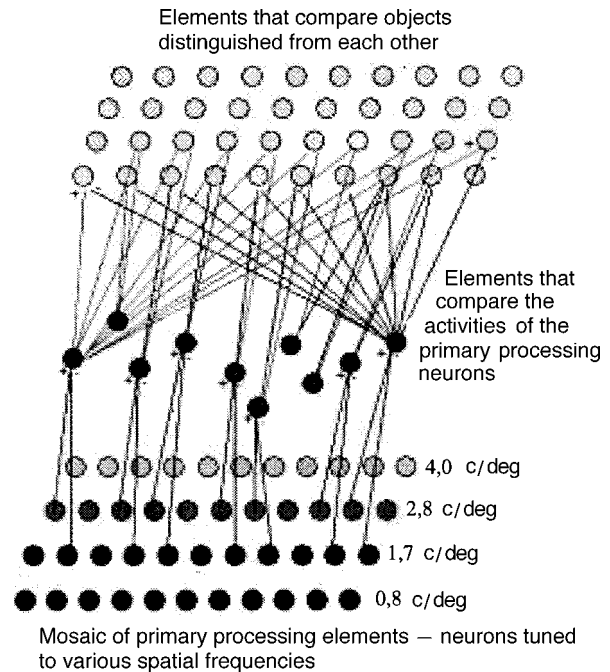


FIG. 6. A neural model for comparisons of separated objects.

cortical cells. At the lowest level is represented an array of primary units tuned to particular bands of spatial frequency, the frequency of peak sensitivity being indicated to the right. Each local patch of visual space may be represented by 4–6 such units. To explain the human ability to discriminate differences in frequency of a few percent, the model postulates a second level of comparators shown in the middle of the figure. The latter are opponent cells drawing inputs of opposite sign from primary units tuned to adjacent bands of frequency.

Each local region of the visual field is sampled by several spatial-frequency analyzers in parallel. We draw here only four out of six neural filters postulated by Wilson and Gelb for spatial frequency and width discriminations.<sup>18</sup> They are tuned to the low and medium frequencies: 0.8–4.0 c/deg. It is unlikely that there are as many analyzers at each position as there are JND's in a psychophysical task. So at the next stage (in the middle of the figure) we postulate, as in the model of Morgan and Regan,<sup>19</sup> local comparators that extract the ratio of activity in the primary analyzers and so support psychophysical thresholds of a few percent. In order to explain the results of the present experiment, we must then postulate a third level (at the top in Fig. 6) in which, for every possible pairing over the central  $10^\circ$  of the visual field, there exists a second-order comparator connected to two first-order comparators. Such a mechanism seems to us implausible. The problem is not only that the number of second-order comparators must be very large and proportional to the square of the number of primary units, but also that a large part of the cortical mass would be given to long axons.

We are led to an alternative hypothesis, though one not often considered in current neuroscience. Beyond the cortical stage where local stimulus features are represented by the activity of single units, there may be a radical transforma-

tion, so that stimulus parameters are represented in an abstract code that can run on a cerebral bus in the way that a computer bus carries different messages for different destinations. We do not, of course, know the physiological nature of this code; but the experimental psychologist can independently ask what information the code might represent. Thus, in the case of the present experiments, one could either postulate that the observer independently compares the code for each stimulus to an interval standard, or template, basing a decision on the categorized values for the two stimuli; or one could postulate that a direct comparison is made between the codes that represent the sensory properties of the stimuli, their spatial frequencies, and contrasts. These last two models cannot be distinguished by the present experiments, but could in principle be distinguished by experiments in which there were a large number of possible reference stimuli.

- <sup>1</sup>J. Hirsh and R. Hylton, "Limits of spatial-frequency discrimination as evidence of neural interpolation," *J. Opt. Soc. Am.* **72**, 1367 (1982).
- <sup>2</sup>J. Beck, "Textural segmentation, second-order statistics, and textural elements," *Biol. Cyber.* **48**, 125 (1983).
- <sup>3</sup>D. Sagi and B. Julesz, "'Where' and 'what' in vision," *Science* **228** 1217 (1985).
- <sup>4</sup>P. Cavanagh, M. Arguin, and A. Treisman, "Effect of surface medium on visual search for orientation and size features," *J. Exp. Psychol.: Human Percep. & Perform.* **16**, 479 (1990).
- <sup>5</sup>Z. L. Lu and G. Sperling, "Attention-generated apparent motion," *Nature* **377**, 237 (1995).
- <sup>6</sup>C. Blakemore and E. A. Tobin, "Lateral inhibition between orientation detectors in the cat's visual cortex," *Exp. Brain. Res.* **15**, 439 (1972).
- <sup>7</sup>V. D. Glezer, *Vision and Mind: Modelling Visual Functions* (Erlbaum Associates, inc. Hillsdale, N.J., 1995).

- <sup>8</sup>A. M. Sillito, K. L. Grieve, H. E. Jones, J. Cudeiro, and J. Davis, "Visual cortical mechanisms detecting focal orientation discontinuities," *Nature* **378**, 492 (1995).
- <sup>9</sup>C. D. Gilbert and T. N. Wiesel, "Clustered intrinsic connections in cat visual cortex," *Neuroscience* **3**, 1116 (1983).
- <sup>10</sup>C. D. Gilbert and T. N. Wiesel, "Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex," *Neuroscience* **9**, 2432 (1989).
- <sup>11</sup>J. S. Lund, T. Yoshioka, and J. B. Levitt, "Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex," *Cerebral Cortex* **3**, 148 (1993).
- <sup>12</sup>A. Grinvald, E. E. Lieke, R. D. Frostig, and R. Hildesheim, "Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex," *Neuroscience* **14**, 2545 (1994).
- <sup>13</sup>D. Shoham, M. Hubner, S. Schulze, A. Grinvald, and T. Bonhoeffer, "Spatio-temporal frequency domains and their relation to cytochrome oxidase staining in cat visual cortex," *Nature* **385**, 529 (1997).
- <sup>14</sup>W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick, "Orientation selectivity and arrangement of horizontal connections in tree shrew striate cortex," *Neuroscience* **17**, 2112 (1997).
- <sup>15</sup>S. Magnussen, M. W. Greenlee, R. Asplund, and S. Dyrnes, "Perfect visual short-term memory for periodic patterns," *Eur. J. Exp. Psych.* **2**, 345 (1990).
- <sup>16</sup>S. Magnussen and S. Dyrnes, "High-fidelity perceptual long-term memory," *Psych. Sci.* **5**, 99 (1994).
- <sup>17</sup>M. Lages and M. Treisman, "Spatial-frequency discrimination: visual long-term memory or criterion setting," *Vision Res.* **38**, 557 (1998).
- <sup>18</sup>H. R. Wilson and D. J. Gelb, "Modified line-element model for spatial-frequency and width discrimination," *J. Opt. Soc. Am. A* **1**, 124 (1984).
- <sup>19</sup>M. J. Morgan and D. Regan, "Opponent model for line interval and vernier performance compared," *Vision Res.* **37**, 107 (1987).

Translation submitted by Russian Editorial Office. Published here with stylistic changes by Translation Editor.