

Parallel processes within the ‘spot-light’ of attention

Yael Adini and Dov Sagi*

Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot 76100, Israel

Received 3 April 1991; revised 28 June 1991; accepted 1 July 1991

Abstract—Human ability to identify simultaneously two targets in the visual field is severely limited. Previous studies have shown that orientation identification of two targets takes twice the time needed for one target. Here we asked whether this seriality is imposed by the decision requirement of the task or by such stimulus properties as target spatial separation and similarity. Observers had to identify the orientations (vertical vs horizontal) of two Gabor patches presented at random positions. Performance on this double-task experiment was compared with performance on each of the tasks when carried out alone. We varied the spatial separation between the two targets for targets having identical or different spatial-frequencies and found that the orientation of two targets having different frequencies could be identified in parallel when occupying the same spatial position but not when separated in space by 4 deg of visual angle or more. Targets having the same frequency could be identified in parallel even when separated by 8 deg, demonstrating that decision factors do not impose seriality. This result can be taken as evidence for the existence of a grouping process operating prior to orientation identification. This grouping process operates according to classical Gestalt rules (proximity, similarity) and enables parallel attentive processing of large input chunks.

1. INTRODUCTION

According to a widely accepted model of pattern vision our visual system encodes the retinal image by means of local, multiple, parallel mechanisms (filters, channels) each of them sensitive to different spatial frequency and orientation ranges. Since Campbell and Robson's (1968) pioneering work, the properties of the hypothetical, psychophysically defined channels were investigated using several types of psychophysical paradigms. Most of them explore detection of near threshold simple and compound gratings using summation phenomena (Graham and Nachmias, 1971; Sachs *et al.* 1971; Bergen, *et al.* 1979; Watson, 1982), adaptation phenomena (Blakemore and Campbell, 1969; Blakemore and Nachmias, 1971), and masking phenomena (Stromeyer and Julesz, 1972; Henning *et al.*, 1981; Phillips and Wilson, 1984). A rough estimate of the bandwidth (full-width at half-maximum sensitivity) of a channel is taken to be 15–20 deg in the orientation dimension, 1–2 octaves in the spatial-frequency dimension, and no more than two cycles in spatial extent (see Olzak and Thomas 1986 for a review). The channels are assumed to be labelled according to their response to orientation and spatial-frequency (Watson and Robson 1981). There is some experimental evidence that observers can use labelling information in detection and identification tasks. Graham *et al.* (1985) and Yager *et al.* (1984), manipulating the uncertainty level in grating detection and identification tasks, suggested that the labelling information is accessible and that human observers can selectively attend to a subset of channels in order to reduce uncertainty and noise.

*To whom all correspondence should be addressed.

While there is a remarkable consensus about the parallel filtering (encoding) of spatial-frequencies and orientations of stimuli, it is not clear whether the information represented in this assumed parallel system can be used simultaneously by a higher-level decision stage. A spatial frequency or an orientation identification requires combining the channels' responses and making an identification judgment on the basis of a certain decision rule (Thomas, 1985). An example of a simple decision rule is a maximum response rule, according to which a stimulus is identified as having the value of the most activated channel. Alternatively, it is possible to use the activity ratio of two channel populations as a continuous measure for stimulus parameter (Foster and Ward, 1991). However, in order to better account for the observation that observers can detect tiny spatial-frequency (orientation) differences in comparison to the assumed bandwidth of channels, a more complicated decision rule has to be used. Such a decision rule can be based on a weighted average of the signals from all active pathways (Georgeson, 1980).

Human ability to simultaneously identify two targets in the visual field is severely limited. Previous experiments have shown that observers' ability to report about several targets is limited, even in the case when just simple features (orientations) are involved (Duncan, 1985; Sagi and Julesz, 1985; Braun and Sagi, 1990). Braun and Sagi (1990) showed, using a backward-masking paradigm, that observers' performance on a task involving two orientation identifications (a double-task condition), is much lower than on each of the single identifications when carried out alone (a single-task condition). This performance reduction was not observed when the double task involved two detection tasks, or when a detection task was combined with an identification task. Given that two orientation identification (discrimination) tasks are carried out serially (Braun and Sagi, 1990), we asked whether this seriality is imposed by: the decision requirements of the task, such as limitation on the number of identifications that can be carried out simultaneously (where identification can be implemented as a comparison between weighted sums of filter responses), or by the 'attentive window' size, i.e. limitation on the size of the spatial region that can be processed in parallel. The second account for seriality implies the possibility of carrying out two identification tasks concurrently when the targets are in a close spatial neighborhood, but not when they are separated by a large distance. Extension of the concept of attentive window into a multidimensional feature space (Bergen and Julesz, 1983) would require a more complicated definition of neighborhood. Two patterns may be processed simultaneously under this extension, if they are close to each other on some dimension other than spatial position, like spatial frequency or orientation. If the limitation on identification is a result of our inability to simultaneously apply weighting functions, or decision rules, to the output of two filters' populations, we would predict serial behavior independently of the distance between the two targets and their similarity.

In order to answer this question, we carried out experiments where observers had to identify the orientation (vertical vs horizontal) of two targets on each trial (2×2 AFC). We used Gabor signals as targets and manipulated the distance between them and their similarity (same or different spatial frequency). The stimuli were designed so that the two targets stimulated two different filter populations. Performance on these double-task experiments was compared with performance on experiments where the same observers had to identify only one of the targets, thus allowing

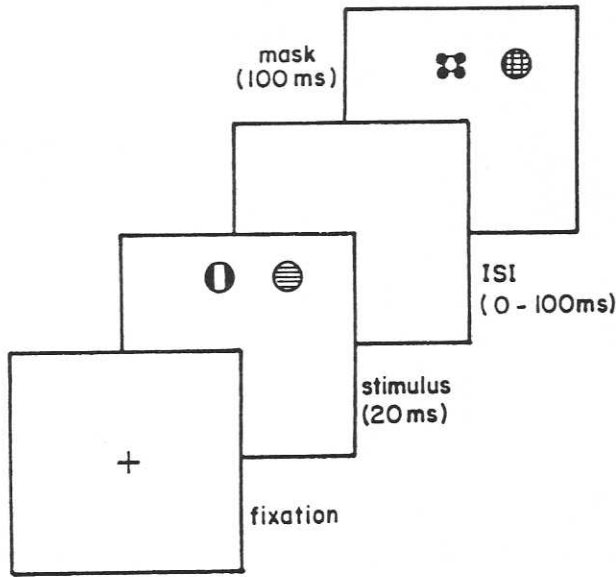


Figure 1. General Experimental Design. See text for details.

us to decide on the seriality of the process. We assumed that if two tasks cannot be performed simultaneously without a reduction in performance in one of the two tasks, they compete for some limited attentive resources.

All experiments were carried out using a masking paradigm, limiting the temporal availability of the stimulus (Bergen and Julesz, 1983; Braun and Sagi 1990).

2. METHODS

2.1. Observers

Five observers participated in these experiments, all of them having normal vision. Three of them were unaware of the purpose of the experiments. All observers were well practiced in the experiments reported here; observers usually had lower performance in the initial phase of the experiments, but later they reached a constant level of performance. Only the latter phase was used in the data analysis. Observers were tested on 100–1000 trials for each data point.

2.2. Display

The stimuli (Fig. 2) were displayed on the face of a Conrac video monitor, with an average luminance of 50 cd/m^2 , using an Imaging Technology frame buffer with a 256×256 pixels resolution at a frame rate of 50 Hz (noninterlaced). Stimulus generation and display were controlled by a Sun-2 workstation, using a special frame-buffer driver for stimulus timing control. Observers were seated at viewing distance of 150 cm.

2.3. Stimulus

The stimuli consisted of a pair of Gabor patches, presented in an otherwise empty visual field, each patch occupying 32×32 pixels (each pixel subtending 3.3 arc min). The Gaussian envelope had a spread of 16 pixels between $1/e$ points, and was

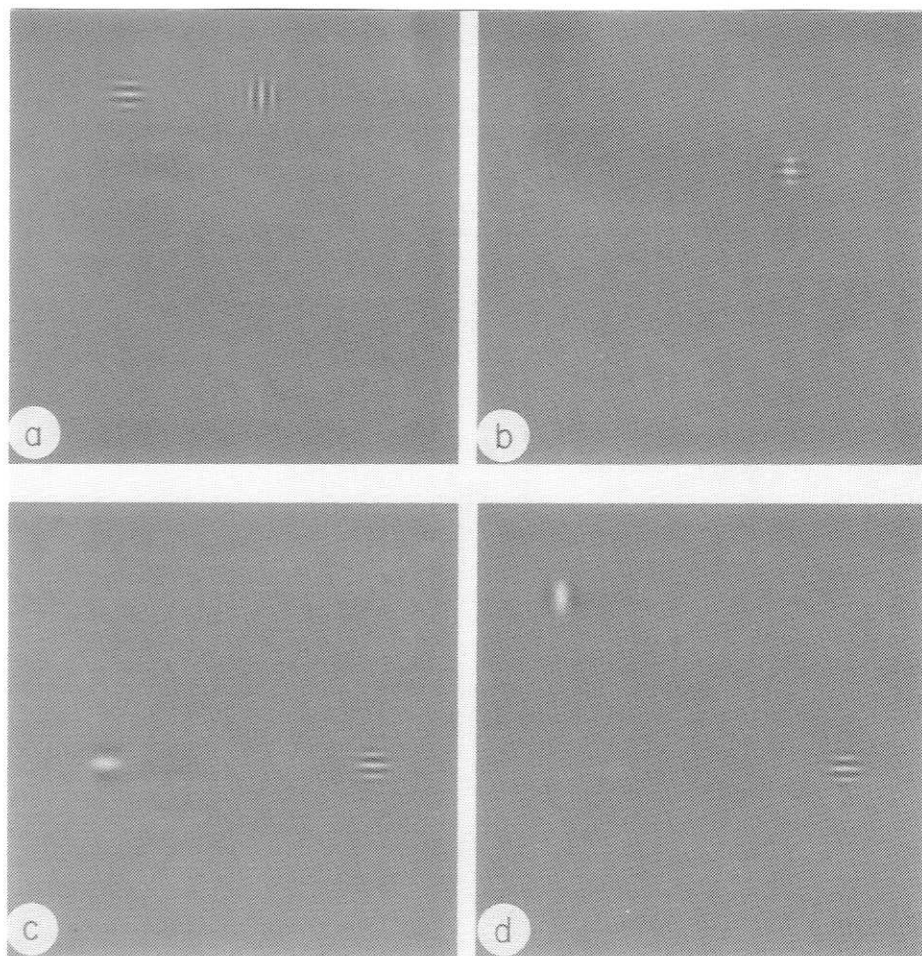


Figure 2. Examples of stimuli used in the reported experiments; (a) HH4, (b) LH0, (c) LH8, (d) LH9nf. Targets differed in their orientation and spatial frequency but not in their size (Gaussian envelope). Actual contrast was lower than shown.

modulated by a cosine function with a wavelength of 6 or 18 pixels (3 and 1 cpd). On each trial, patch-orientation was randomly selected and could be either vertical or horizontal with an equal probability. Two pair-configurations were used. In the *fixed-configuration* experiments, stimulus elements were horizontally aligned (i.e. the vertical separation between them was 0 deg) with a fixed (throughout the experiment) horizontal separation. In the *non-fixed-configuration* (9 deg nf) experiment, the horizontal separation was fixed (8 deg) while the vertical separation was random between 0 and 8 deg, yielding an average separation of 9 deg (Figure 2d). Patch position was randomized around the screen at eccentricities between 1 and 4 deg (constrained by pair configuration). The stimulus contrast was 30% (defined as the ratio of the cosine amplitude to the average luminance, thus reflecting the actual contrast only at the center of the Gabor patch).

2.4. Task

In double-task experiments (DT), observers were asked to identify the orientation (horizontal or vertical) of the left stimulus element (DT1) and to identify the orientation (horizontal or vertical) of the right stimulus element (DT2), i.e. two separate responses. Observers were instructed (but not forced) to give the two tasks equal priority, although responses ordered such that they first had to respond to the left element and then to the right element. In single-task 1 experiments (ST1) observers were asked to identify the orientation (horizontal or vertical) of the left stimulus element. In single-task 2 experiments (ST2) observers were asked to identify the orientation (horizontal or vertical) of the right stimulus element.

2.5. Stimulus presentation

Experiments were conducted in blocks of 50 trials. Before each stimulus presentation, a fixation point appeared at the center of the screen until the observer signalled readiness by pressing the space bar on a standard terminal keyboard. Then, after a random delay of 200–500 ms, the stimulus was briefly presented and masked. The time between the onset of stimuli and the onset of mask (Stimulus Onset Asynchrony or SOA) was varied. The brief presentation of the stimuli (20 ms) precluded a second fixation. The backward mask limited the visual availability of the stimuli and thus the 'processing time' available to the observer. The mask was composed of a pair of Gabor patches appearing at stimulus position. Each of the mask elements was composed of two overlapping Gabor patches with orthogonal orientations (0 and 90 deg) and of the same spatial frequency and size as the target element it covered. Each mask element was of a total contrast of 100% and 100 ms presentation time. As a result of high mask energy (contrast and duration) the task of discriminating the orientation of the stimulus proved to be impossible at some small SOA. At these SOAs, stimulus and mask were probably superimposed perceptually by visual persistence.

2.6. Experimental variables

Two experimental variables were considered: The spatial separation between the centers of the two stimulus elements and the degree of similarity between them. Experiments were carried out in three sets: fixed-configuration experiments, non-fixed-configuration experiments and a set of mixed-spatial-frequencies (mixed-sf) experiments. On the fixed-configuration condition, we used three different horizontal separations: 0, 4, 8 deg (vertical separation was 0 deg). On the non-fixed configuration condition, horizontal separation was 8 deg and vertical separation was random between 0 and 8 deg (average separation of 9 deg). We used patches of two different spatial frequencies: 1 cpd (low frequency) and 3 cpd (high frequency). Three conditions were examined. LH experiments: the left stimulus element was of the low frequency, the right stimulus element was of the high frequency; LL experiments: stimulus elements were of the same low frequency; and HH experiments: stimulus elements were of the same high frequency (Fig. 2). For each spatial separation (4, 8, 9 deg nf) the three possible patches' spatial frequency combinations were tested.

For spatial separation 4 deg with fixed configuration, two additional similarity conditions were used in a set of mixed-sf experiments. These were: the Same-Mixed condition, where within each block both patches were of the same spatial frequency

(LL or HH with an equal probability) and the Different-Mixed condition, where patches were of different spatial frequency (LH or HL with an equal probability).

In summary, the experimental condition we used were¹:

Fixed-configuration set: LH4, LL4, HH4, LH8, LL8, HH8, LH0;

Non-fixed configuration set: LH9nf, LL9nf, HH9nf;

Mixed-SF (fixed configuration) set: LH4|HL4, LL4|HH4.

2.7. Psychophysical procedure

Observers had to identify the orientations (vertical vs horizontal) of a pair of Gabor patches (Gaussian-modulated cosine patches) presented simultaneously at a random position. On each DT trial, two separate responses were given corresponding to each pair member (first for left target and then for right target). Response was given by typing 0 or 1 on a standard computer keyboard. Performance (per cent correct) on this double-task experiment was compared with performance on each of the tasks (identifying one of the targets; a single-response task) when carried out alone. Experiments were conducted in blocks of 50 trials. Each session lasted about 45 min. Experimental variables (configuration, inter-stimuli separation, patches' spatial-frequency combination) and task (DT, ST1 or ST2) were fixed for any one session. (Observers DS and YA were exceptions in the sense that data from different tasks were collected during each session.) From trial to trial the orientation of each stimulus element was randomly selected to be either vertical or horizontal (with an equal probability). A session consisted of 15–24 blocks. An observer started a session with a block at a high SOA (180–120 ms) where his or her performance was around 100%. Thereafter, SOA in each block decreased by 20-ms steps until chance performance was reached. An observer usually repeated this 'procedure' 3–6 times each session.

2.8. Data analysis

Observers' responses were collected from each of the trials and the percentage of correct responses was computed for each experimental block. The average (over blocks) percentage of correct responses and the standard error were computed for each experimental condition (i.e. type of display, task and SOA). Using this data, the psychometric function for each experiment was plotted.

In order to bring all different psychometric functions into a common form, and in order to obtain a more compact way to present the data collected, we replaced each psychometric function with an equivalent step function (a threshold function). An arbitrary psychometric function can be defined to be equivalent to a step function, if the area above this psychometric function (i.e. the area between the line which describes perfect performance at all SOAs and the observer's actual performance) is equal to the area above the step function. We define the stepping point of the step function to be the threshold time (T_{th}). (The data analysis was repeated under the assumption that the underlying psychometric curve is a Quick function, with practically the same results.)

¹ To shortly describe an experimental condition, we used the notation $L_1 L_2 D$, where L_1 indicates the spatial frequency of the left Gabor patch, L_2 indicates the spatial frequency of the right Gabor patch, and D indicates the inter-stimuli separation. Whenever a non-fixed configuration was used, the letters 'nf' were added.

$T_{th}(task_i)$ were calculated for each experimental condition. $T_{th}(task_i)$ is the threshold-time for $task_i$ ($task_i \in \{DT1, DT2, ST1, ST2\}$), computed from the experimental psychometric curve of each task.

We defined the degree of capacity limitation of a task as:

$$CLI = (T_{th}(DT1) + T_{th}(DT2)) - (T_{th}(ST1) + T_{th}(ST2))$$

For a parallel process $T_{th}(DT_i) = T_{th}(ST_i) \Rightarrow CLI = 0$.

2.8.1. Statistical analysis. The CLI parameter obtained under various experimental conditions was analyzed using a three-way analysis of variance (ANOVA) mixed model. We used the SAS procedure GLM (General Linear Model) and checked the hypothesis that the means of CLIs (over each of the model's effects) are the same. We then used Scheffé and Tukey multiple-comparison procedures to obtain homogeneous sub-groups. The model we used to analyze the non-overlapping configurations had two main fixed effects: Stimulus spatial frequency structure (at three levels: LH, LL, HH), and the inter-stimuli spatial separation (at two levels: 4 and 8 deg). The five observers were treated as a random effect.

The model we used to analyze all experimental conditions included in the fixed and non-fixed experimental set had two main fixed effects, Similarity and Configuration. Similarity effect had two levels: Same (same location or same spatial-frequency, in this category were included conditions: LH0, LL4, LL8, LL9nf, HH4, HH8, HH9nf), and Different (conditions LH4, LH8, LH9nf). The Configuration effect had two levels: Fixed (includes all experiments with a fixed configuration) and Non-Fixed (includes 9 deg nf experiments). The five observers were again treated as a random effect.

3. RESULTS

Experiments were carried out in three sets: fixed-configuration experiments, non-fixed-configuration experiments and a set of mixed-spatial-frequencies (mixed-sf) experiments.

3.1. Fixed-configuration experiments

In the fixed-configuration set of experiments, seven experimental conditions were carried out: LH0, LH4, LH8, HH4, HH8, LL4, and LL8. Each of these conditions was carried out under one double-task and two single-task conditions.

Psychometric curves for one observer (out of 5) are presented in Fig. 3. This particular naive observer had been trained for the longest period (more than eight months; 2–3 sessions in a week). Most of her training was on the LH4 condition, where she failed to achieve parallel (or capacity-free) performances, and on the LH0 condition, where she achieved capacity-free performance. The psychometric curves describe the observer's performance (per cent correct) as a function of SOA for all the different experimental conditions tested.

These psychometric functions show that the ability to simultaneously identify two orientations is not the same under different similarity and proximity conditions. On same-frequency experiments (LL4, LL8, HH4, HH8) and on same-location experiments (LH0), performance does not significantly differ between double- and single-task conditions. However, when stimulus components were of different spatial frequencies, and did not occupy the same location (LH4, LH8), performance under DT2 condition were significantly different from those obtained under ST2 condition.

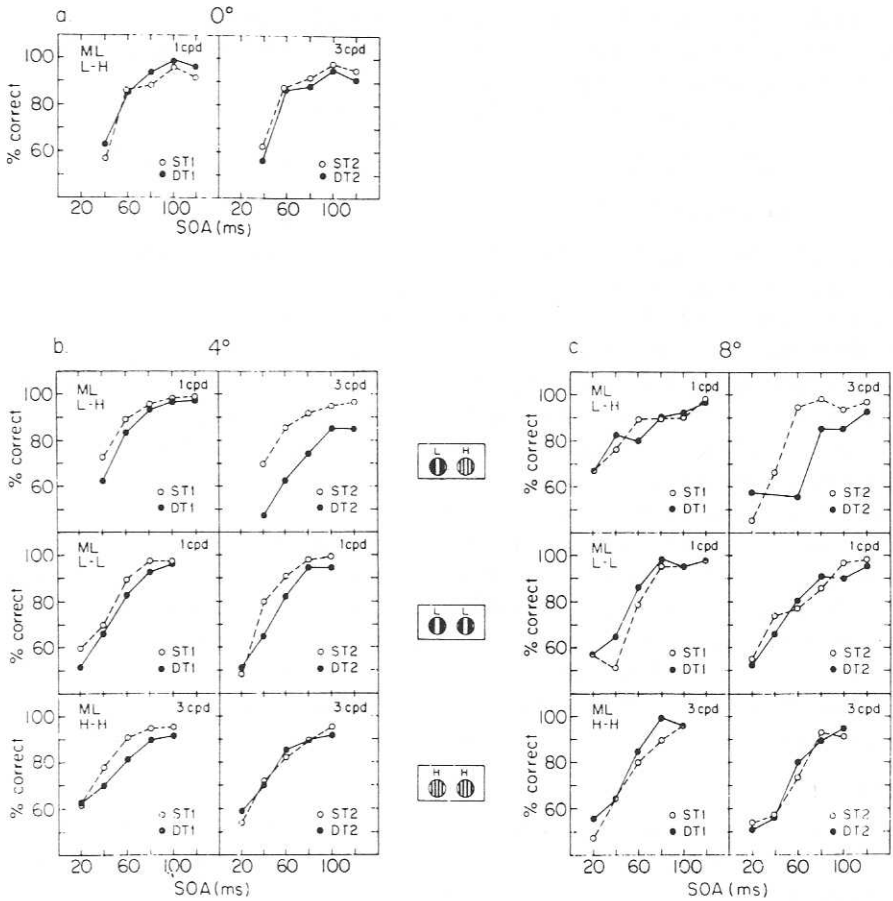


Figure 3. The psychometric functions obtained by observer ML for all fixed-configuration experiments. Curves are arranged in three groups, (a) LH0 curves, (b) LH4, LH4, and HH4 curves, and (c) LH8, LL8, and HH8 curves. For each experimental condition two pairs of curves are shown. The left pair describes performance on DT (double-task condition) and ST (single-task conditions) for the left Gabor patch, whereas the right pair of curves describes performance on DT and ST for the right Gabor patch. Filled symbols (●) represent DT performances, empty symbols (○) represent ST performance.

In general, observers showed reduced performance on both DT1 and DT2 conditions; thus we consider the sum of the temporal shifts of DT_i relative to ST_i , across both tasks as indicative of performance reduction (defined as Capacity Limitation Index in Section 2.8). The full CLI data for all observers are presented in Fig. 4; the average (across observers) of time thresholds, for all conditions are presented in Table 1. These results suggest little or no capacity limitations on parallel orientation identifications under the same-frequency or same-location conditions, and capacity limitations on parallel identifications, under the 'different' condition.

Since the CLI describes the difference between performance under DT and ST conditions, a low CLI may result from high performance under the DT condition (low T_{th}) which may reflect observers' ability to divide their attentions among the two tasks, or from poor performance under the ST conditions (high T_{th}) which may reflect

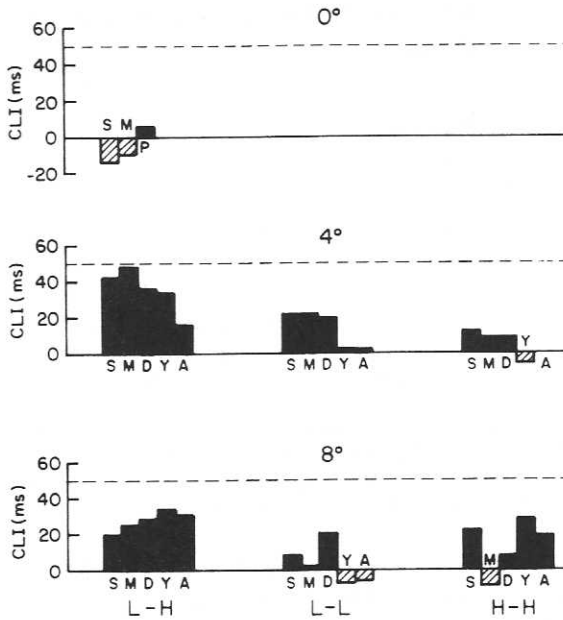


Figure 4. The effect of targets' similarity on observers' CLI values. CLI is defined as $(T_{th}(DT1) + T_{th}(DT2)) - (T_{th}(ST1) + T_{th}(ST2))$. Each bar represents data of a single observer identified by her or his first name initial. The dashed lines are reference lines, which describe the average threshold-time of a single orientation identification ($T_{th}(ST_i) = 50 \pm 4$ ms). This average $T_{th}(ST)$ (across all single-task experimental condition) was computed from Table 1. Note the relatively low Capacity Limitation Index (CLI) values obtained under LH0, LL and HH conditions relative to the high CLI values obtained under LH4 and LH8 conditions.

observers' inability to focus their attention in order to optimize the desired single task. A low CLI which is a result of observers' inability to focus their attention on the desired ST may reflect a strong interference among elements competing for the attentive system rather than capacity-free performances. An examination of the data presented in Table 1 shows that the average T_{th} s obtained under all ST conditions are about the same. This observation suggests that the low CLI values obtained under LL and HH conditions (at 4 and 8 deg inter-stimuli separations) are due to observers' ability to successfully divide their attention under these conditions.

We used a three-way ANOVA model to test the hypothesis that the mean capacity

Table 1.

Average threshold times and SEs (across observers) for the different experimental conditions.

Exp	$T_{th}(ST1)$	$T_{th}(ST2)$	$T_{th}(DT1)$	$T_{th}(DT2)$
LH0	67 ± 5	69 ± 10	62 ± 9	67 ± 12
LH4	50 ± 11	50 ± 8	65 ± 14	71 ± 9
LL4	45 ± 8	42 ± 6	48 ± 10	53 ± 8
HH4	53 ± 8	49 ± 6	57 ± 8	52 ± 7
LH8	49 ± 10	48 ± 6	56 ± 9	68 ± 6
LL8	53 ± 8	54 ± 11	51 ± 13	58 ± 12
HH8	53 ± 11	49 ± 6	54 ± 6	61 ± 5

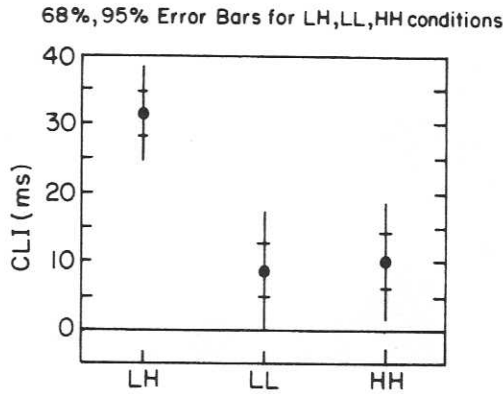


Figure 5. The effect of stimuli similarity on capacity limitations as indicated by CLI data. Data for each of the similarity conditions is averaged over separations of 4 and 8 deg across all observers. Note that mean CLI obtained under LH conditions is significantly higher than mean CLIs obtained under LL and HH conditions. Error bars indicate 68 and 95% confidence intervals.

limitations on performing two identifications (as indicated by observers' CLI values) are the same under all experimental conditions tested. The model we used had two main fixed effects: Stimulus spatial frequency structure (at three levels: LH, LL, HH), and the inter-stimuli separation (at two levels: 4 and 8 deg). The five observers were treated as a random effect. The stimulus spatial-frequency structure was found to be significant: $F(2, 8) = 14.8$; $P < 0.002$. The mean CLI (the difference in ms between DTs and STs) obtained for the LH level (31.4 ms), was found to be significantly different from the mean CLI for the HH level (9.6 ms) and for the LL level (8.4 ms); (by Scheffé and Tukey procedures $\alpha < 0.01$). Means for LL and HH levels were not found to be significantly different (see Fig. 5). The inter-stimuli-separation effect was found to be non-significant ($F(1, 14) = 0.61$, $P > 0.44$).

This result suggests that observers' ability to simultaneously perform two orientation identifications depends on similarity between stimulus components, but not on the spatial separation (in the range 4–8 deg) between them. The results show that parallel identification is not limited by the decision requirements of the task, nor by the response requirements, all these requirements being constant across experimental conditions. The results show also that concurrent identification is constrained by such stimulus properties as target similarity and proximity.

3.2. Non-fixed configuration experiments

In the previous section it was found that the orientations of two Gabor patches having the same spatial frequency or location, can be identified with little or no capacity limitations under the constraint of fixed configuration. In this section we ask to what extent this result depends on this constraint.

As in the case of the fixed-configuration experiments, observers had to identify the orientations of two Gabor patches having the same or different spatial frequencies. The horizontal spatial separation between the centers of the two stimulus elements was kept fixed (8 deg). However, vertical separation was random between 0 and 8 deg, yielding to a non-fixed configuration. In the non-fixed-configuration set of experiments, there were three experimental conditions: LH9nf, LL9nf and HH9nf. Each

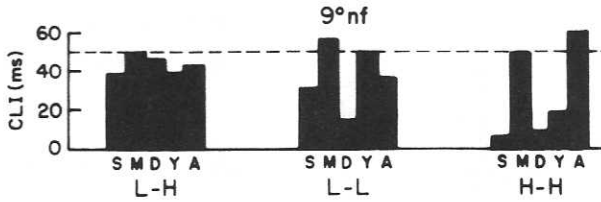


Figure 6. CLI data for 9 deg nf condition. From left to right bars describe observers' CLI under LH, LL, and HH experimental conditions. Letters indicates observers' first name initial.

of these conditions was implemented under one double-task and two single-task conditions.

To check the hypothesis that the configuration has no effect on the main result, we used a three-way ANOVA mixed model. The main effects were *Configuration* (in two levels: fixed or non-fixed), and *Pair-similarity* (in two levels: same, different). Observers were taken as a random effect. The analysis revealed a significant interaction between the Observers effect and the Configuration effect ($P < 0.03$). That is, we cannot decide generally on the effect of Configuration on observers' performances since this effect depends strongly on the particular observer. Also, in the non-fixed condition it is impossible to decide on the Similarity effect due to the large inter-observer differences.

The CLI data for the non-fixed configuration are presented in Fig. 6. A consistent high capacity-limited performance is observed when the two Gabor targets were of different spatial frequencies (LH9nf: CLI = 43 ms; sd = 5; $n = 5$). Note that under this condition (LH9nf), the average CLI (43 ms) is very close to the average threshold time needed for performing a single task (50 and 51 ms for fixed and non-fixed configurations respectively), suggesting that at least in this case all observers adopted the same *serial* strategy. The most inconsistent CLIs are obtained under HH9nf (average CLI = 28 ms; sd = 22), where some observers achieved capacity-free performances, while others showed a relatively high capacity-limited performance (higher than in the LH9nf condition). In the LL9nf condition the inter-observer variance is somewhat smaller (average CLI = 37 ms; sd = 15), but still larger than in the fixed-configuration conditions. On the average, the CLI variances in the non-fixed same-frequency conditions are 3.5 times larger than the equivalent fixed-configuration variances. The source of this inconsistency is not clear and should be a subject of further investigation. For the time being we can consider this inconsistency as evidence for top-down control over the size of the processed area (attentive window). It is possible that the variability in targets' spatial configuration and separation imposes modification of the attentive window on each trial, thus introducing an additional difficulty into the task.

3.3. Mixed-SF experiments

In the previous sections (in particular 3.1) it was shown that there is a difference between capacity limitation observed under different similarity conditions. However, this result was obtained under the constraint that each experimental condition was carried out in a separate set of blocks. Spatial-frequency channels are assumed to be labelled according to their response to orientation and spatial-frequency (Watson and

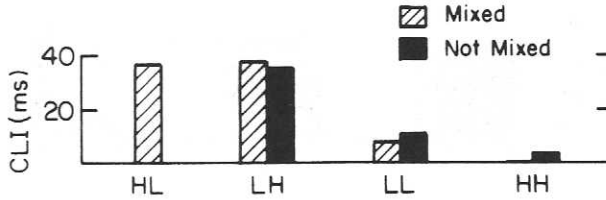


Figure 7. The effect of mixing stimuli having different spatial-frequency pair structure, within the same block, on the main result (the dependency of CLI values on patch similarity). Bars indicate average (of two observers) CLI values under mixed-sf (empty) and not-mixed (filled) experimental conditions.

Robson 1981) and there is some experimental evidence that observers can use labeling information in detection and identification tasks (Yager *et al.*, 1984; Graham *et al.*, 1985). That is, it could be argued that under LL and HH conditions, observers could selectively attend to just one set of spatial-frequency filters (low-frequency filters or high-frequency filters) and that the performance reduction in the LH case (between ST and DT) is due to increased uncertainty in the spatial-frequency domain. In the latter case observers are forced to identify two patches of different frequencies and although the two patches occupy different locations, this situation may force the system to integrate responses across both frequencies in both locations. This would result in non-optimal identification due to the additional noise originating from the nonrelevant filters. By this explanation, the similarity effect found in the fixed-configuration experiments might suggest that the low capacity limitation performances obtained under LL and HH are a result of pre-setting the system to the relevant frequencies (which should be distinguished from a dynamical setting by attentional mechanisms). To test this hypothesis, we ran a set of mixed-sf experiments, where LL and HH pairs had an equal probability of appearing within each block. Thus, if the low CLIs obtained under LL and HH conditions resulted from pre-setting the attentive system to the relevant frequencies, we would now expect an increase in CLI for these cases.

3.3.1. Methods. We used the same methods as for the 4 deg separation experimental condition, with only one exception. Two conditions were used: same-frequency condition, where targets were chosen from the set LL and HH (with equal probability), and a different-frequency, control condition, where targets were chosen from the set LH and HL (with equal probability). We separately analyzed the percentage of correct responses as a function of SOA, for each of the HH, LL, LH and HL targets. The two authors (DS and YA) served as observers in these experiments. Both of them were not trained on the mixed-sf conditions. Their results reflect the first 2–3 blocks (100–150 trials for each relevant SOA).

To analyze the effect of mixing two spatial frequencies within the same block (i.e. LL or HH pairs at the same block as one condition, and LH or HL pairs at the same block as a second condition), we used a three-way ANOVA ($3 \times 2 \times 2$) mixed model, with pair-spatial-frequency structure (in three levels: LH, LL, HH) and uncertainty (in two levels: high (mixed) and low (not-mixed)). Observers were taken as a random effect.

3.3.2. Results. As can be seen from Fig. 7, observers showed the same pattern of

performance under mixed and non-mixed conditions. As was found in previous sections, where stimulus' components were of a different spatial frequencies, observers obtained relatively high capacity limitations, while under LL and HH conditions the CLI values were small.

The ANOVA model we ran revealed *no* significant main effect for the uncertainty effect ($F(1, 7) = 0.11, P > 0.75$). As was found in previous sections, a significant main effect was found for the pair-spatial-frequency-structure ($F(2, 7) = 12.97, P < 0.005$). Mean CLI for LH cases, mixed and not-mixed, (37 ms) was found to be significantly different ($P < 0.012$) from means obtained for LL and HH conditions (9 ms and 2 ms respectively).

3.4. Main results

- Consistent (across observers), capacity-free (parallel) identification is observed when the two Gabor targets spatially overlap (CLI = -5 ms, sd = 6).
- Consistent (across observers), high capacity-limited (serial) identification is observed only when the two Gabor targets are not in a constant spatial relationship and are of different spatial frequencies (CLI = 43 ms, sd = 5 for LH9nf).
- When spatial relations between targets are constant through a set of experiments (and targets do not overlap, i.e. 4 and 8 deg), targets can be processed in parallel to different degrees. The ability to carry out two identification tasks in parallel (without a loss in performance) may depend on target similarity and proximity, and on the particular strategy adopted by the observer. On the average, targets having the same spatial frequency may be processed in parallel whereas targets having different spatial frequency are processed serially.

4. DISCUSSION

We have investigated the nature of capacity limitations involved in the performance of two identification tasks under various similarity and proximity conditions. A masking paradigm was used to encourage resource-limited rather than data-limited performance (Norman and Bobrow, 1975). We found that the orientations of two targets having different frequencies could be identified without a loss in performance, when occupying the same spatial position (overlap) but not when separated in space by 4 deg or more. Targets having the same frequencies could be identified in parallel even when separated by 8 deg of visual angle, at eccentricities up to 4 deg. In all these cases our targets stimulated two different filter populations.

The results show that parallel identification is neither limited by the decision requirements of the task (such as limitation on the number of identifications that can be carried out simultaneously, where identification can be implemented as a comparison between weighted sums of filter responses), nor by the response requirements, all these requirements being constant across experimental conditions. This finding is consistent with the claim that attention affects the quality of the representation of selected items and not merely the decision process involved in the tasks (Prinzmetal *et al.*, 1986). It is still possible that the ability to carry out two decisions simultaneously is constrained by some higher level of visual processing not affected by backward masking. However, our results show that concurrent identification is already constrained by stimulus properties such as target similarity, configuration and proximity.

One possible interpretation of this result is that parallel (but attentive) processing is limited to a region in a combined frequency-position space having a constant volume but flexible dimensions, so that frequency differences can be traded for spatial separation. That is, as the spatial-location uncertainty increases (a wider 'spot-light' in space), uncertainty in the spatial-frequency domain should decrease. More than that, the center of the processing region within this combined space can be shifted within time steps as fast as 50 ms. This model differs from earlier suggestions linking spatial-frequency channels to attention (Julesz, 1980; Braddick and Atkinson, 1982) by allowing for efficient parallel processing of broad-band localized stimuli. Only for global processing is it advantageous to select a specific frequency band for processing.

It is also possible to view our results in terms of a theory assuming some simple grouping processes operating before identification is taking place. In this theory, some processes link together stimulus parts over a large spatial range according to some simple rules. In agreement with the Gestaltists claims (Köhler, 1929), we may think that groups or 'gestalts' (perceived units) can be regarded as attentional units. That is, elements that are grouped together by these processes are selected as an attentional unit and identified simultaneously by the attentive system. Along this line of explanation, our results suggest that there are some low-level linking operations which are determined by the neural connectivity of the visual system. Cells which respond to the same sensory dimension (spatial-frequency in the present case) are connected directly or indirectly, through shared neighbors; cells responding to stimuli occupying the same location are connected, but cells of different sensory dimensions (with non overlapping receptive fields) are not connected, and thus are not linked by this hypothetical low-level grouping system. Such a connectivity may explain why a pair of stimuli of the same spatial frequency are processed by the visual system as one structural unit, whereas a pair of non-overlapping Gabor patches, of different spatial frequencies, cannot be processed as one unit (resulting in a serial processing of the pair's components). Note that we do not require a complete identity between grouped items. Targets may be grouped together when having only one or two features in common. In our experiments, targets having the same spatial frequency differed in their orientation on half of the trials, and targets having the same location differed always in their spatial frequencies. On the other hand, an identity of one feature (spatial frequency) may not guarantee an efficient linkage as revealed by the non-fixed configuration experiments (this argument holds also for experiments with targets having different frequencies where on half of the trials they assumed the same orientation). Thus, although grouping cannot be established without at least one sensory-linking feature, it requires some spatial and temporal consistency between the grouped elements.

Recently it was suggested that hierarchical segmentation and grouping is a basic, capacity-free stage in visual processing (Duncan and Humphreys, 1989). The visual selection theory, suggested by these authors, has three components: a parallel stage of perceptual description, a selection process and the entry of selected information into visual-short-term-memory (VSTM) which allows control of access to awareness. According to this theory, the perceptual description is made by a process of hierarchical segmentation of the image into linked groups and subgroups (structural units). Each structural unit is described by its elementary sensory properties (relative position, color, size, motion etc) and categorical properties (based on meaning). The

above-described process of segmentation and description is said to be parallel and resource free; selection starts when all parts of the display compete for access to the VSTM. Linked structural units tend to gain or lose resources (activity strength) together. Increased assignment of resources to any structural unit increases its speed and probability for access to VSTM. Once a structural unit emerges the 'winner' in the visual selection process it accesses VSTM. Duncan and Humphreys (1989) propose that "Structural units act as wholes, competing for and gaining access to VSTM with all their associated descriptions". That is, according to this theory, in our 2-Gabors experiments, capacity-limited identifications are predicted whenever each of the stimulus' components competes for attention separately (as two separate structural units), whereas a capacity-free performance is predicted under conditions where the stimulus components are grouped together and compete for attention as one structural unit (see also Kahneman and Henik, 1977; Treisman, 1982). Our results are in a good agreement with the basic claim of this theory that attention may work on structural units, rather than by optimizing the processing of each isolated target. Our results do not support the notion of effortless grouping according to meaning. The highly trained observers participated in our experiments could not learn to perceive a 'word' composed of two 'letters' (two Gabor patches), each with a different spatial frequency, as one structural unit. That is, in order to account for the difference between LH and LL/HH conditions we should assume that there is an absolute limit to the ability of early vision to produce structural units and descriptors.

Finally, an important distinction should be made between grouping processes involved in pattern identification and segmentation processes involved in detection of texture differences. Human performance on texture tasks (Julesz 1962; Beck, 1972; Gurnsey and Browse 1987) can be modeled by assuming linear spatial filters with short-range interactions between filters having *identical* spatial frequency and orientation selectivities (Malik and Perona 1990; Rubenstein and Sagi 1990). These processes seem to be efficient in detecting feature gradients in an otherwise uniform textural background without making use of any attentive resources (Braun and Sagi, 1990). For the grouping processes implied by the present study, it is sufficient to have only one linking feature being spatial frequency or orientation. This linkage takes place over much larger distances, up to 32 times the target wavelength (in the HH8 condition), compared with a maximum of 9 wavelengths found for orientation-gradient detection (Sagi, 1990). It remains to be seen how these two processes of segmentation and grouping cooperate in defining object candidates for recognition.

Acknowledgment

We are grateful to Dr. Edna Schechtman for her guidance with statistical analysis, Leah Mory Rauch for helping with the software used in the experiments, and Yehuda Barbut for the graphics. This work was supported by the National Institute for Psychobiology in Israel.

REFERENCES

- Beck, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. *Am. J. Psychology*, **85**, 1-19.
- Bergen, J. R. and Julesz, B. (1983). Rapid Discrimination of Visual Patterns. *IEEE SMC-13*, 857-863.
- Bergen, J. R., Wilson, H. R. and Cowan, J. D. (1979). Further evidence for four mechanisms mediating

- vision at threshold: sensitivities to complex gratings and aperiodic stimuli. *J. Opt. Soc. Am.* **69**, 1580–1587.
- Blakemore, C. and Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol. Lond.* **203**, 237–260.
- Blakemore, C. and Nachmias, J. (1971). The orientation specificity of two visual after-effects. *J. Physiol. Lond.* **213**, 157–174.
- Braddick, O. J. and Atkinson, J. (1982). Higher functions in vision. In: *The Senses*, H. B. Barlow and J. D. Mollon (Eds.). (Cambridge Texts in the Physiological Sciences, 3) Cambridge University Press, Cambridge.
- Braun, J. and Sagi, D. (1990). Vision outside the focus of attention. *Percept. Psychophys.* **48**, 45–58.
- Campbell, F. W. and Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *J. Physiol. Lond.* **197**, 551–566.
- Duncan, J. (1985). Visual search and visual attention. In: *Attention and Performance XI*, M. I. Posner and O. S. M. Marin (Eds.). Erlbaum, Hillsdale, New Jersey.
- Duncan, J. and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychol. Rev.* **96**, 433–458.
- Foster, D. H. and Ward, P. A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proc. R. Soc. Lond.* **B243**, 75–81.
- Georgeson, M. A. (1980). Spatial frequency analysis in early visual processing. *Phil. Trans. R. Soc. Lond.* **B290** 11–22.
- Graham, N. and Nachmias, J. (1971). Detection of grating pattern containing two spatial frequencies: A comparison of single-channel and multi-channel models. *Vision Res.* **11**, 251–259.
- Graham, N., Kramer, P. and Haber, N. (1985). Attending to spatial frequency and spatial position of near-threshold visual patterns. In: *Mechanisms of Attention: Attention and Performance XI*, M. I. Posner and O. S. M. Marin (Eds.). Erlbaum, Hillsdale, New Jersey.
- Gurnsey, R. and Browse, R. A. (1987). Micropattern properties and presentation condition influencing visual texture discrimination. *Percept. Psychophys.* **41**, 239–252.
- Henning, G. B., Hertz, B. G. and Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. Noise masks. *J. Opt. Soc. Am.* **71**, 574–581.
- Julesz, B. (1962). Visual pattern discrimination. *IRE Trans. Info. Theory* **IT-8**, 84–92.
- Julesz, B. (1980). Spatial frequency channels in one-, two- and three-dimensional vision: Variations on an auditory theme by Békésy. In: *Visual coding and Adaptability*, C. S. Harris (Ed.). Erlbaum, Hillsdale, New Jersey.
- Kahneman, D. and Henik, A. (1977). Effects of visual grouping on immediate recall and selective attention. In: *Attention and Performance VI*, S. Dornic (Ed.). Erlbaum, Hillsdale, New Jersey.
- Köhler, W. (1929). *Gestalt Psychology*. Horace Liveright, New York.
- Malik, J. and Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am.* **A7**, 923–932.
- Norman, D. L. and Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychol.*, **7**, 44–64.
- Olzak, L. and Thomas, J. P. (1986). Seeing spatial patterns. In: *Handbook of Perception and Human Performance, Vol. 1: Sensory Processes and Perception*, K. Boff, L. Kaufman and J. Thomas (Eds.). Ch. 7. Wiley, New York.
- Phillips, G. C. and Wilson, H. R. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *J. Opt. Soc. Am.* **A1**, 226–232.
- Prinzmetal, W., Presti, D. and Posner, M. (1986). Does attention affect visual feature integration? *J. Exp. Psychol. HPP*, **12**, 361–369.
- Rubenstein, B. S. and Sagi, D. (1990). Spatial variability as a limiting factor in texture discrimination tasks: Implications for performance symmetries. *J. Opt. Soc. Am.* **A7**, 1632–1643.
- Sachs, M. B., Nachmias, J. and Robson, J. G. (1971). Spatial-frequency channels in human vision. *J. Opt. Soc. Am.* **61**, 1176–1186.
- Sagi, D. (1990). Detection of an orientation singularity in Gabor textures: Effect of signal density and spatial-frequency. *Vision Res.* **30**, 1377–1388.
- Sagi, D. and Julesz, B. (1985). 'Where' and 'What' in vision. *Science* **228**, 1217–1219.
- Stromeyer, C. F. and Julesz, B. (1972). Spatial-frequency masking in vision: Critical bands and spread of masking. *J. Opt. Soc. Am.* **62**, 1221–1232.

- Thomas, J. P. (1985). Detection and identification: how are they related. *J. Opt. Soc. Am.* **A2**, 1457-1467.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and objects. *J. Exp. Psychol. HPP* **8**, 192-214.
- Watson, A. B. (1982). Summation of grating patches indicates many types of detector at one retinal location. *Vision Res.* **22**, 17-25.
- Watson, A. B. and Robson, J. G. (1981). Discrimination at threshold: Labelled detectors in human vision. *Vision Res.* **21**, 1115-1122.
- Yager, D., Kramer, P., Shaw, M. and Graham, N. (1984). Detection and Identification of Spatial Frequency: Models and Data. *Vision Res.* **24**, 1021-1035.