
Attention and spatial resolution: A theoretical and experimental study of visual search in hierarchical patterns

Gustavo Deco

Computational Neuroscience, Department of Technology, Institutió Catalana de Recerca i Estudis Avançats (ICREA), Universitat Pompeu Fabra, Passeig de Circumval.lació 8, E 08003 Barcelona, Spain

Dietmar Heinke ¶

Behavioural and Brain Sciences Centre, University of Birmingham, Birmingham B15 2TT, UK;
e-mail: d.g.heinke@bham.ac.uk

Received 9 November 2005, in revised form 17 January 2006; published online 28 February 2007

Abstract. Our aim has been to study visual search in hierarchical patterns through a computational neuroscience perspective. In this way we were able to analyse the relationship between the global precedence effect and attention. The theoretical part of our investigations extended a model built upon the biased-competition hypothesis. An important feature of this extension is that different spatial frequencies are processed with different time constants. Simulations with this model resulted in specific predictions that were confirmed with a novel visual-search experiment utilising hierarchical patterns. Both experiment and simulations suggest that the global-precedence effect prevails under the condition of visual search.

1 Introduction

In everyday life humans show an astonishing ability to select objects among numerous other objects. In laboratories this ability is examined in visual-search experiments where participants have to search for a predefined target among nontargets (distractors). Numerous studies have shown that the efficiency of the search depends on several factors, such as the ability to form groups of distractors (eg Duncan and Humphreys 1989) or the difference in orientation between target bar and distractor bar (see Wolfe et al 1998 for a recent review). However, typical items used in these experiments do not exhibit an important property of most natural objects. Natural objects tend to be structured in a hierarchical fashion: local elements (eg edges) form perceptual groups and, in turn, perceptual groups are joined together to form objects. How humans process this hierarchical nature of objects is typically examined in experiments with Navon stimuli (Navon 1977). Navon stimuli are large, global patterns composed of small, local patterns, eg a large letter constructed of local letter elements (see figure 1 for an example). Results of experiments on hierarchical patterns demonstrate that the processing of the global level is faster than processing of the local level—the so-called global-precedence effect (see Kimichi 1992 for a review). In contrast to visual-search studies, these experiments tend to use displays with single objects. Hence so far the question has not been raised if global precedence prevails in scenes with multiple objects. Also, the influence of hierarchical items on the efficiency of search has not been examined.

In this paper we approach this question in the theoretical framework of a computational neuroscience model (Deco and Zihl 2001; Corchs and Deco 2002, 2004; Deco and Lee 2002; Deco et al 2002, 2004; Rolls and Deco 2002). Amongst other applications, this model has been used to successfully simulate data on visual search. Here, we extend it by adding to it the ability to search through hierarchical objects and demonstrate that the model can successfully predict the results of such experiments.

¶ Author to whom all correspondence should be addressed.

The advantage of this approach is that we can not only derive concrete predictions about behavioural experiments but also make predictions about physiological evidence. The ability of the model to combine the physiological level and the behavioural level stems from the fact that it is based on the hypothesis of biased competition for selective attention (Reynolds et al 1999). In short, this hypothesis postulates that objects compete for attention. As an expression of this competition, the neural activation related to the winning object is enhanced, whereas activation that relates to ignored objects is decreased. This competition between objects is biased by numerous factors, such as bottom–up salience, expectation, memory, etc. Based on this theory, our model covers a broad range of experimental evidence on attention and memory at behavioural, fMRI, and single-cell level (Corchs and Deco 2002; Rolls and Deco 2002). For instance, Deco and Lee (2002) demonstrated that object-based and space-based attention, often seen as a dichotomy, can be unified in a computational model (see Heinke and Humphreys 2003 for a similar argument). Deco and Zihl (2001) and Deco et al (2002) applied the model to visual-search experiments. They demonstrated that experimental results normally interpreted as evidence for serial or parallel search (eg search for single-feature and conjunction targets) originate from different latencies of the competitive dynamics in the parallel-operating model.

Here we extend this model of visual search so that it is able to search through hierarchical patterns. In order to do this, we extend the competitive dynamics to operate on different spatial-frequency channels in parallel. This extension is based on the ‘resolution hypothesis’ which will be introduced in the following section. However, as the review of physiological evidence for the resolution hypothesis will show, it is not clear whether the speed of the competitive dynamics is the same at each level of spatial frequency or whether it can be assumed that the dynamics at coarse spatial frequencies is faster than dynamics at finer spatial frequencies. The latter hypothesis can be seen as a neurodynamic implementation of the global-precedence effect. Simulations with the new model will later show that these two hypotheses lead to different predictions for visual search in hierarchical patterns. Two experimental tests support the implementation of the global precedence. In the conclusion we will discuss the implications of these results for the physiological level.

2 The resolution hypothesis

Since Navon’s (1977) seminal work, a number of psychophysical studies have been devoted to eliciting the role of low-frequency channels in the processing of global image structure (eg Wong and Weisstein 1983; Shulman et al 1986; Shulman and Wilson 1987; Lovegrove et al 1991). Navon (1977) developed a task that permits easy manipulation of global and local features of an image by using hierarchical patterns. A hierarchical pattern refers to a large, global pattern composed of small, local patterns; a classic example is a large letter constructed of local letter elements (Kinchla 1974; Navon 1977; Shulman and Wilson 1987; Badcock et al 1990). Figure 1 shows typical examples of hierarchical patterns formed with letters H and T. Participants are required to report either the global letter or the local letter in the display. Experimental results demonstrate that the processing of the global level is faster than the processing of the local level. Subsequent studies have shown that manipulation of low spatial frequencies selectively influences the global task, thus supporting the hypothesis that the global

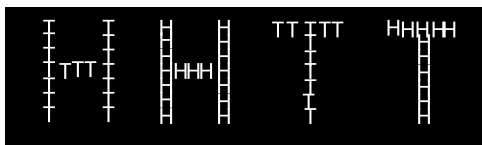


Figure 1. Hierarchical patterns composed of letters (Navon 1977).

aspect of a stimulus is processed by low-spatial-frequency mechanisms (eg Shulman et al 1986; Shulman and Wilson 1987; Badcock et al 1990; Lovegrove et al 1991; Michimata et al 1999). Shulman and Wilson (1987) studied the relationship between spatial frequency and selective attention to local and global structures by an experiment in which observers had to respond to the global level of hierarchical stimuli, followed by a detection response to sinusoidal gratings that differed in spatial frequency. They reported that high-frequency gratings were better detected when observers had just responded to a local feature of the hierarchical stimulus, whereas low-frequency gratings were better detected when observers had just responded to a global feature of the stimulus. Badcock et al (1990) also confirmed the role of low-spatial-frequency information in the processing of global aspects of hierarchical patterns. They showed that the global precedence effect was lost after a selective removal of low spatial frequencies from the hierarchical stimuli. A very elegant demonstration of the relationship between global processing and low-spatial-frequency mechanisms was presented by Michimata et al (1999). Their experiment was designed on the basis of findings about the function and characteristics of the magnocellular pathway in the visual system. The lateral geniculate body in primates and humans has a six-layer structure, which can be divided into two major portions. Four layers, called parvocellular layers, contain cells with relatively small receptive fields and have sustained response characteristics. The other two layers, called magnocellular layers, contain cells with large receptive fields and have transient response characteristics. Both kinds of layers provide input to the primary visual cortex (V1) where the division is preserved. The magnocellular pathway contains cells of type IV (Livingstone and Hubel 1984) which have receptive fields with a tonic red surround mechanism (ie imposing diffuse red light causes tonic suppression of the activity of these cells). Breitmeyer and Breier (1994) provided experimental evidence that the magnocellular pathway is the physiological basis of the low-spatial-frequency channel of the visual system. Michimata et al (1999) repeated the reaction-time experiment of Navon but manipulated the background colour of the stimulus. The control conditions with a green background reproduced basically Navon's results, but a red background, which blocks the processing of the low-spatial-frequency magnocellular channel, produced the cancellation of the global precedence effects.

In summary, numerous psychological experiments with a single hierarchical pattern as stimulus indicate that low spatial frequencies are processed faster than high spatial frequencies. However, it is not known how this global precedence effect is implemented at the physiological level. In the following section we suggest a possible implementation of this effect.

3 Computational model

3.1 Neurodynamics of attention

We now briefly describe a cortical model of visual attention for object recognition and visual search based on the biased-competition hypothesis (Desimone and Duncan 1995; Duncan et al 1997; Reynolds et al 1999), and outline the corresponding neurodynamic mechanisms. The system operates in an autonomous fashion, with each of its functionalities explicitly described in a complete mathematical framework. The overall architecture representation of the model is shown in figure 2.

The model is essentially composed of three modules structured to resemble the two known main pathways of the mammalian visual cortex. Information from the retinogeniculo-striate pathway enters the visual cortex through area V1 in the occipital lobe and proceeds into two processing streams. The occipito-temporal stream (the so-called 'what' pathway) leads ventrally through V2, V4, and IT (inferotemporal cortex) and is mainly concerned with object recognition, independently of position and scaling.

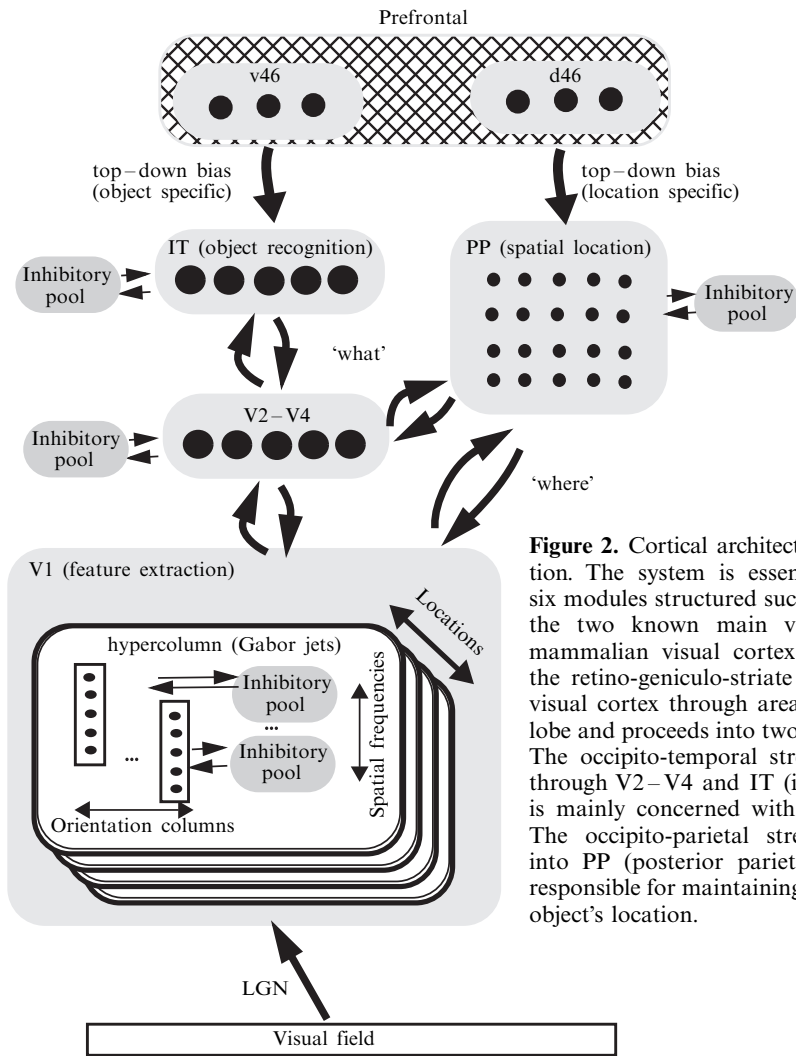


Figure 2. Cortical architecture for visual attention. The system is essentially composed of six modules structured such that they resemble the two known main visual paths of the mammalian visual cortex. Information from the retino-geniculo-striate pathway enters the visual cortex through area V1 in the occipital lobe and proceeds into two processing streams. The occipito-temporal stream leads ventrally through V2–V4 and IT (inferotemporal), and is mainly concerned with object recognition. The occipito-parietal stream leads dorsally into PP (posterior parietal complex) and is responsible for maintaining a spatial map of an object's location.

The occipito-parietal stream (the so-called 'where' pathway) leads dorsally into PP (posterior parietal complex) and codes the location and spatial relationships between objects (eg Ungerleider and Mishkin 1982; Desimone and Ungerleider 1989). The first module (V4) of our system is concerned with the extraction of features and consists of pools of neurons with Gabor receptive fields tuned to different positions in the visual field, orientations, and spatial-frequency resolutions. The 'where' pathway operates through the mutual connection between the early visual areas and the second (PP) module that consists of pools coding the positions of stimuli. The connections with the first module generate a top-down biasing attentional modulation based on the location of stimuli. The third module (IT) of our system is engaged with the recognition of objects and consists of pools of neurons which are sensitive to the presence of a specific object in the visual field. The pools of neurons in IT are synaptically connected with translational invariant receptive fields with pools of the first module (V4) such that specific objects are invariantly recognised, irrespective of their lateral positions in the field (eg Gross et al 1972). The mutual connections between the IT and V4 modules represent a top-down biasing attentional modulation associated with specific objects.

The model can operate in two different modes: the learning mode and the recognition mode. During the learning mode, the synaptic connections between V4 and IT are trained by Hebbian learning during several presentations of specific objects that appear at different positions in the visual field. When there are multiple objects in the field, recognition of any given object can operate in two ways. First, an object can be recognised by biasing the system with an external top–down component in the IT module, which drives the competition in favour of the pool associated with the specific object being searched for. Owing to this object-based bias, the intermodular attentional modulation between V4 and IT will enhance the activity of the pools in V4 associated with the features of the specific target object, and the intermodular attentional modulation between V4 and PP drives the competition in favour of the pool localising the target. Second, the system can be biased by a spatial cue within the PP module which drives the competition in favour of the pool associated with the specific location of one object. As a consequence, the intermodular attentional modulation between V4 and PP will favour the pools in V4 associated with the features of the object at that location, and intermodular attentional modulation between V4 and IT will favour the pool that recognises the particular object. These two top–down biases represent known effects on visual selection from (i) object-based cues to attend to a particular target (see Chelazzi et al 1993 for physiological evidence; see Downing 2000 for psychological evidence), and (ii) spatial cues to attend to a particular location (see Posner et al 1980; Moran and Desimone 1985 for psychological and physiological evidence, respectively). We assume that both forms of the top–down bias originate from regions of frontal cortex that are not explicitly modelled. The mathematical formulation of the system is given in the Appendix.

3.2 Neurodynamics of the resolution hypothesis

The experimental evidence reviewed earlier suggests that attentional mechanisms actively enhance the spatial resolution at the attended location. On the other hand, we have assumed in our model that top–down attentional mechanisms can be understood as the parallel mutual modulation between neuronal pools resulting from intramodular and intermodular dynamic interactions. Hence, a plausible neuronal realisation of the resolution hypothesis involves a modification of the underlying dynamics in such a way that the neuronal pools associated with the processing of different spatial resolutions show also different latencies. We implement the resolution hypothesis and the global precedence effect in our theoretical model by assuming that the neuronal pools in module V1 which correspond to different octaves (ie different spatial resolution) have also different velocities of evolution. This can be achieved by defining different time-membrane constants for different octaves. The new update equations for neuronal pools in V1 are similar to equation (A6) of the Appendix but now with explicit different τ_k :

$$\begin{aligned} \tau_k \frac{\partial}{\partial t} A_{k p q l}^{V1}(t) = & -A_{k p q l}^{V1}(t) + aF[A_{k p q l}^{V1}(t)] - bF[A_k^{I, V1}(t)] \\ & + I_{k p q l}^{V1}(t) + I_{p q}^{V1-PP}(t) + I_{k p q l}^{V1-IT}(t) + I_0 + v. \end{aligned} \quad (1)$$

We pose that the neuronal pools corresponding to low spatial frequency (higher octaves) evolve faster than the neuronal pools corresponding to high spatial frequency (lower octaves). Consequently, the time-membrane constants corresponding to higher octaves are smaller than the ones corresponding to the lower octaves:

$$\tau_k < \tau_{k'} \quad \text{for } k > k'. \quad (2)$$

In other words, we assume that the processing of coarse spatial information is faster than the processing of finer spatial details. Figure 3 shows the implementation of the resolution hypothesis in the cortical architecture introduced before.

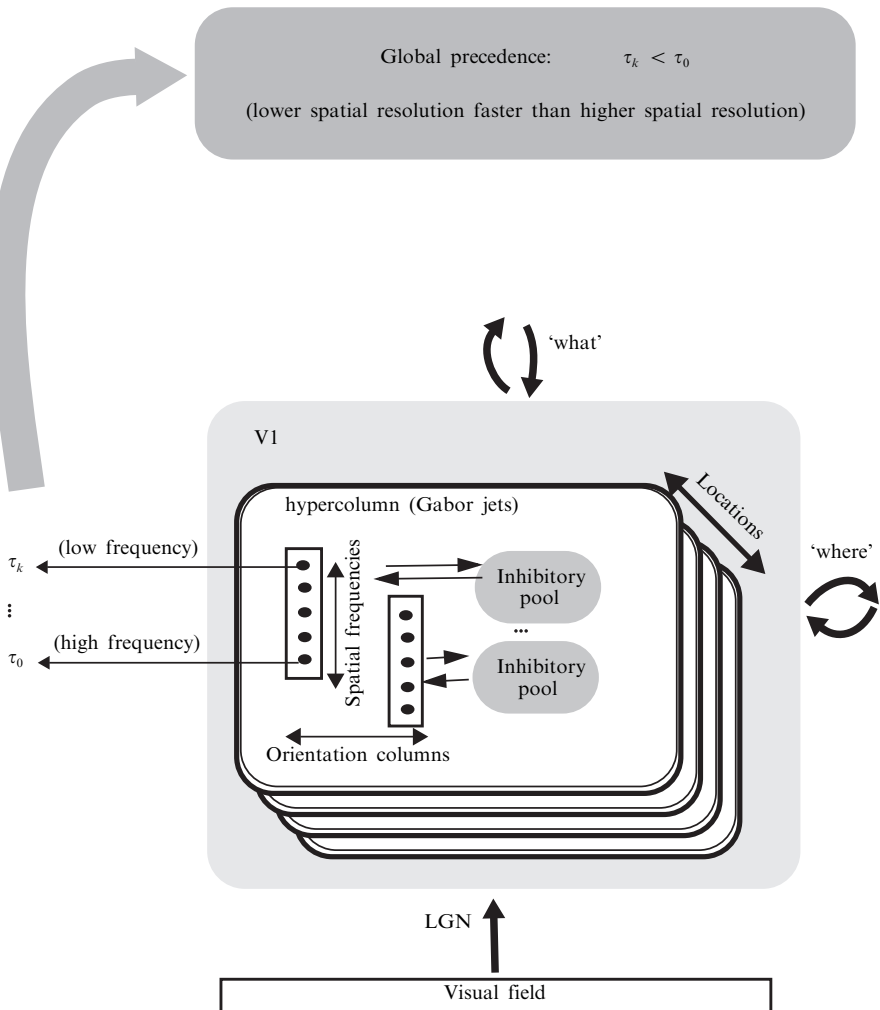


Figure 3. Implementation of the resolution hypothesis in our cortical model for visual attention. The neuronal pools in module V1 which correspond to different octaves (ie different spatial resolution) are assumed to have also different velocities of evolution. This can be achieved by defining different time-membrane constants for different octaves. See text for details.

The dynamic evolution of the whole visual system is similar as before but with the important difference that now the top-down attentional intramodular and inter-modular dynamic modulation is considered at the beginning with major-weight coarse spatial features, because the lower spatial frequencies update much faster. After a while, the slower high spatial frequencies reach also a level of activity capable of influencing the whole dynamic system, which means that finer features are now considered in the visual processing. Owing to the dynamic competition existing from the very beginning, only the spatial regions already primed by coarser spatial resolution are able to be further analysed with finer resolution. The lower spatial resolution therefore guides the attentional modulation to concentrate on certain spatial regions for a further fine analysis later. Intuitively, this seems to be an implementation of the global precedence effect. In the following section this intuition will be tested in visual-search experiments.

4 Visual search in the framework of the resolution hypothesis

In the previous section we introduced an implementation of the resolution hypothesis. In this section we propose a new psychological experiment which allows us to generate specific experimentally testable predictions.

4.1 A new experimental paradigm

In visual-search experiments, the number and the characteristics of the distractors can be manipulated in order to test different dynamic behaviours of attention. Hence, the attentional dynamics that underlies the resolution hypothesis can be demonstrated by combining the processing of multiple spatial resolutions with visual search. In concrete, we propose a visual-search task in the context of the global–local paradigm, ie by using hierarchical patterns as targets and distractors. As hierarchical patterns we use global letter shapes constructed of local letter elements. Let us denote as Xy an item corresponding to a global letter X and composed of local letters y , X and y being arbitrary letters. The aim of our psychophysical paradigm is to find out whether our implementation of the spatial resolution has an effect on the reaction time during a search experiment, such that, if active attentional experiments are able to discriminate, we can predict whether the hypothesis of different update velocities for different spatial frequencies does or does not hold. We define two different conditions in a visual-search experiment. The first condition (VS1) defines the distractors and the target such that the target can be found by analysing only the global (coarse) level. We say in this case that the display shows global disparity. Figure 4a shows a typical screen for this kind of visual-search experiments containing three items. In this case the target is defined as the pattern Hh and the distractors are patterns Fh. The global similarity of the patterns (ie global H or global F) can be determined reliably at a coarse level of feature extraction by using only low frequencies. In this situation, if the processing precedence of low frequencies holds, we expect the search efficiency to be determined by the global level. For instance, if the target ‘popped-out’ on the global level, search would be highly efficient. The neuronal pools in module V1 extracting the global coarse features cause an early full polarisation of the spatial attention in PP at the position where now the only item showing the global characteristics of the target is the target itself. The processing of high frequencies (finer details) will only play a role at the spatial position of the early localised target and serves for object recognition.

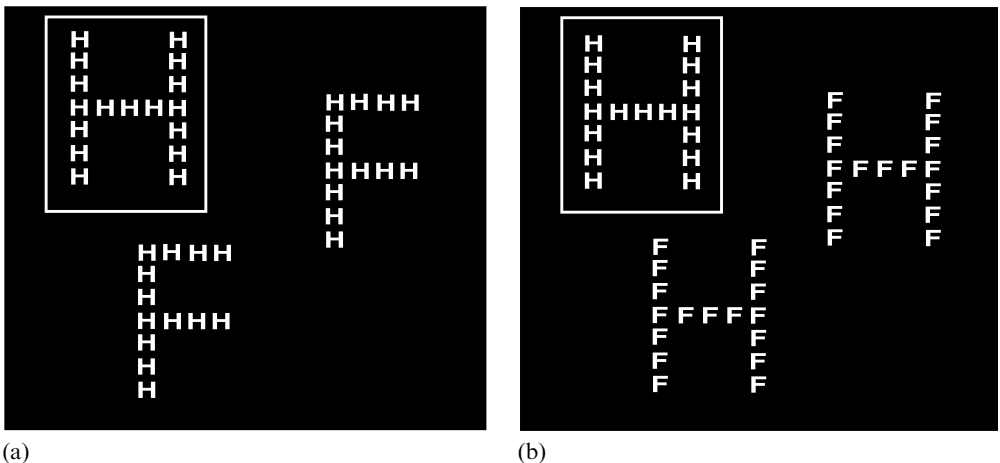


Figure 4. Examples of visual-search screens analysed in the context of the global–local paradigm: (a) global disparity (VS1), (b) global similarity (VS2).

On the contrary, if the processing precedence of low frequencies does not hold, then finer details of distractors and the target increase the attentional competition from the very beginning and a search should be more inefficient than a simple search on the global level.

The second condition (VS2) defines the distractors and the target such that their global characteristics are indistinguishable, and they only differ at the local level. We say in this case that the display shows global similarity. Figure 4b shows a typical display for this case. The target is defined as the pattern Hh and the distractors are patterns Hf. To find the target one has to process high-frequency information along the entire screen because, in this case, the coarse feature information extracted by the low frequencies does not polarise the attention and assigns equal levels of activity in PP to each item location (all global forms are H). The competition between distractors and target is only solved when the high-frequency processing pools are activated. The level of competition at the high-frequency level increases with the number of distractors and therefore an inefficient (serial) search is expected, even when the processing precedence of low frequencies holds.

4.2 Computational results and experimental predictions

In this section we present the simulation results obtained in the visual-search experiments on hierarchical patterns introduced in the last section. The input display is given by a pixelised image. The V1 hypercolumns cover the entire image uniformly. They are distributed in 33×33 locations and each hypercolumn is sensitive to three spatial frequencies and to two different orientations. Consequently, the V1 module has 6534 pools and three inhibitory pools. The IT module utilised has four pools for recognition of the patterns Hh, Hf, Ff, and Fh, and one common inhibitory pool. Finally, the PP module contains pools corresponding to each possible spatial location, ie to each of the pixels, and a common inhibitory pool. Each of the four hierarchical objects (shown in figure 4) is isolated in order to define four categories to be associated with four different pools in the IT module. During the learning phase, these four objects are presented randomly and at random positions in order to achieve translation-invariance responses. The system required 3 000 000 different presentations for training the IT pools. During the search mode, two types of implementation were tested: (i) without resolution hypothesis, and (ii) with resolution hypothesis. In the case without resolution hypothesis, the time-membrane constants were the same for all spatial frequencies. In the implementation with resolution hypothesis, the time-membrane constants for the V1 pools corresponding to the two highest spatial frequencies were higher than for the lowest spatial frequency. Hence, in this last case, the lowest spatial frequencies represent a processing advantage.

Figure 5 shows the dynamic evolution of the polarisation of the pools in the PP module, in terms of differences of the previously defined maximal activity of target and distractors. The reaction time is obtained by determining the point at which the level of polarisation of the pools in the PP module reaches a certain threshold, meaning that the searched target is localised. Figures 5a and 5b show the dynamic evolution for experiments VS1 (global disparity) and VS2 (global similarity) and different numbers of distractors. In figure 5c the simulated reaction time is plotted as a function of the display size. When the resolution hypothesis is implemented, VS1 shows an efficient search instead of the inefficient search observed when the different octave latencies in V1 are equalised. On the other hand, in the case of VS2 always an inefficient search is observed. Only in the case of global disparity the faster processing of the global structure causes a significant facilitation of the search task owing to an early polarisation of the competition for the location of the target. The implementation of the resolution hypothesis can be described phenomenologically as a kind of impressionist search, in the sense that first the coarse global structure is searched in a scene and

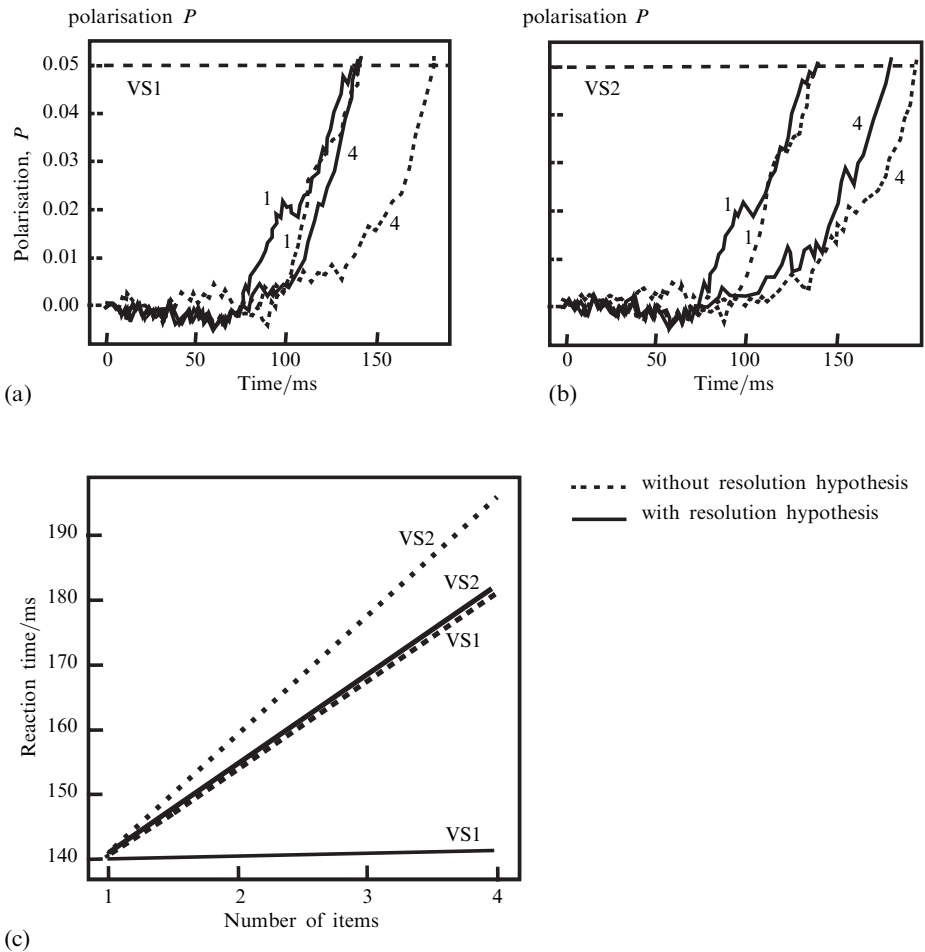


Figure 5. Neurodynamics of visual search of hierarchical patterns. (a) Dynamic evolution of the pool activity in the PP module during global disparity visual search VS1; (b) dynamic evolution of the pool activity in the PP module during global similarity visual search VS2. (Number at the curves indicates the number of items on the screen.) (c) Simulated reaction time for different types and numbers of distractors.

the finer details of the globally preselected spatial regions are analysed afterwards. It should be noted that the impressionist search is only efficient when the global structure of the target is very different from the global appearance of the distractors in the scene. If the scene contains objects which have similar global aspects, impressionist coarse search does not produce any advantage. The simulations explain the neurodynamics underlying this kind of impressionist visual processing, which involves the effect of the resolution hypothesis in the active context of a visual-search experiment.

In summary, our cortical model and neurodynamic implementation of the resolution hypothesis predict that the efficiency of search through hierarchical patterns is largely determined by the global level when the distractors show global disparity. In contrast, when the search display exhibits global similarity, the search is expected to be highly inefficient.

We conducted two experiments to test the prediction of the model. In the first experiment we used the same hierarchical patterns as in the simulations. However, the outcomes of the experiment were not as clear-cut as the simulations suggested. This resulted mainly from the fact that the human behaviour did not exhibit the same search

efficiency on the global level as the simulations assumed. Therefore, we conducted a second experiment with items that show a pop-out effect at the global level. This experiment clearly confirmed the prediction of the model.

5 Experiment 1

This experiment consisted of two blocks. Block 1 implemented the experimental paradigm explained earlier (see figure 6). Block 2 used displays consisting of either small or large non-hierarchical letters (see figure 7 for examples). As the large letters in block 2 match the size of the global letters in block 1, block 2 is assumed to give a good estimate of the search efficiency on the global level of block 1. The displays with small letters were included in block 2 to equate the two blocks with respect to their difficulty. Now, following the predictions from the simulations, the search in block 2 should be as efficient as the search through the hierarchical patterns in VS1 condition.

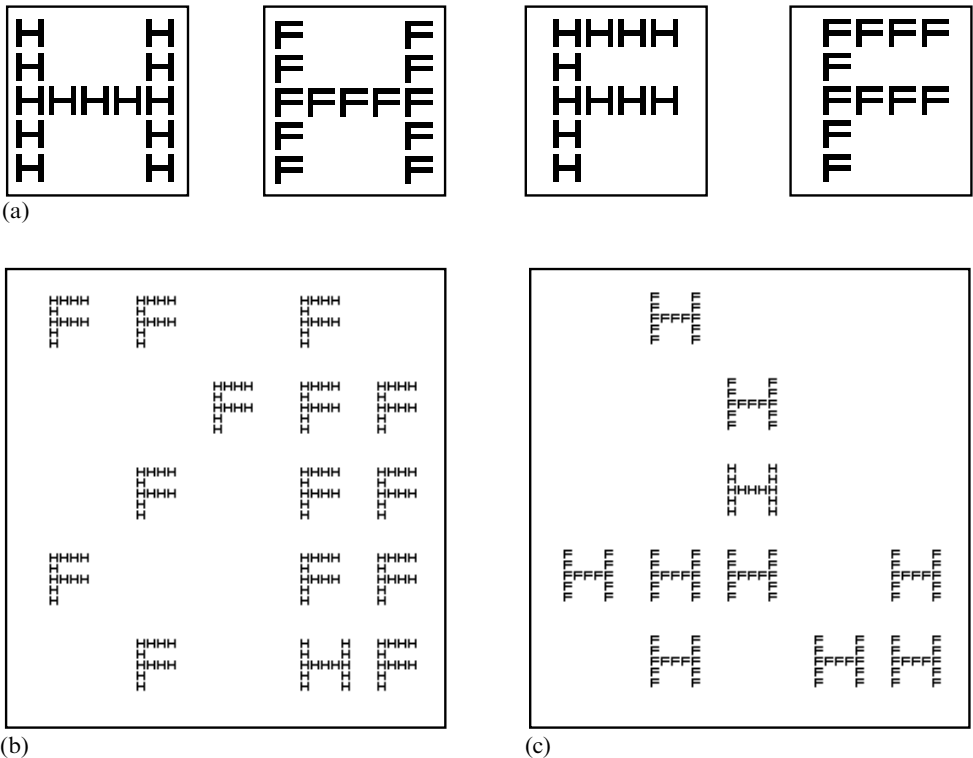


Figure 6. Experiment 1. (a) Hierarchical stimuli used as target (Hh) and distractors (Hf, Fh, Ff) for the experimental validation of the resolution hypothesis (block 1); (b) typical display for the condition of global disparity (VS1); (c) typical display for the condition of global similarity (VS2) with Hf as distractor.

5.1 Procedure

5.1.1 Participants. The participants were eighteen undergraduate students from the University of Birmingham. All participants were reported having normal or corrected-to-normal vision.

5.1.2 Stimuli. In block 1 the stimuli consisted of large letters (H or F) constructed of small letters (H or F). Figure 6a shows the four items used as target and distractors for the visual-search task. The side length of the global structure consisted of five letters. The small letters covered an area of $0.28 \text{ deg} \times 0.3 \text{ deg}$. The large H was $1.4 \text{ deg} \times 1.4 \text{ deg}$.

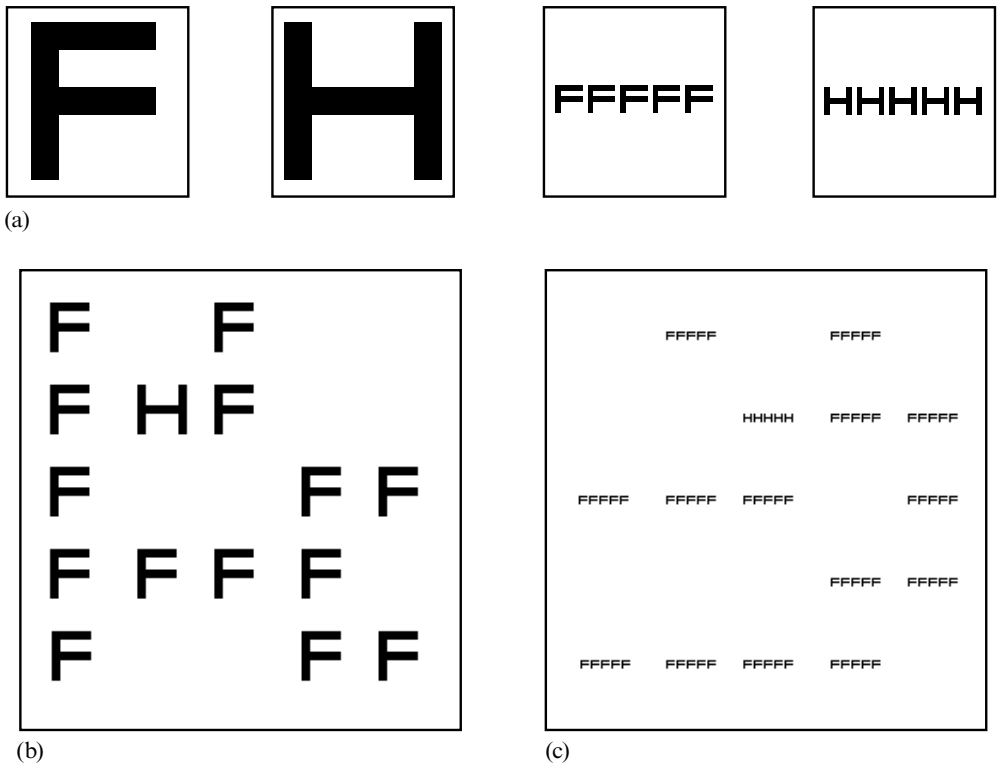


Figure 7. Experiment 1. (a) Stimuli used in block 2 (baseline). (b) Typical display with large letters with the target being defined as H. (c) Typical display with small letters.

The large F was $1.1 \text{ deg} \times 1.4 \text{ deg}$. The stimuli are defined as follows: (i) ‘Hf’—a global H constructed of small F letters, (ii) ‘Hh’—a global H constructed of small H letters, (iii) ‘Ff’—a global F constructed of small F letters, and (iv) ‘Fh’—a global F constructed of small H letters. The search displays were made of 5, 10, or 15 items. Figure 6 shows two examples for 15 and 10 items. The display size was $11 \text{ deg} \times 13.5 \text{ deg}$.

In block 2 size-matched non-hierarchical versions of the letters in block 1 were used (see figure 7). The number of items and the display size were the same as in block 1.

5.1.3 Design. The presentation of the stimuli and the timing and data collection were under the control of a standard PC with a 15 inch monitor. ‘Z’ and ‘M’ were used as response buttons for target present and target absent, respectively. The monitor was viewed from a distance of 50 cm. For block 1, figures 6b and 6c show typical displays for the global disparity and similarity conditions, respectively. The target was a large H made out of small letters H (Hh) and could be present or absent. Each display remained until the computer picked up a key-press. The subjects were instructed to respond fast and accurately. Each experiment was run as a within-subjects design. There were three conditions: one corresponding to global similarity (VS2) and two corresponding to global disparity (VS1). In the case of global similarity, the distractor items are defined by a large H letter constructed of small F letters (Hf). In the case of global disparity, the distractors are given by large F constructed of small letters, which were either H or F (Fh or Ff). After 18 practice trials, the subjects had to do 288 trials interrupted after 144 trials by a break. No feedback was given.

In block 2 the same design and procedure were used with the exception that the displays were made out of non-hierarchical letters H and F (see figure 7). The two blocks were counterbalanced.

5.2 Results

5.2.1 Block 1: Hierarchical patterns. A three-way ANOVA revealed three significant main effects for distractors ($F_{2,34} = 66.21, p < 0.001$), number of items ($F_{2,34} = 55.03, p < 0.001$), and target (present/absent) ($F_{1,17} = 29.93, p < 0.001$). There was a significant interaction between distractor and number of items ($F_{4,68} = 12.65, p < 0.001$), between distractor and target (present/absent) ($F_{2,34} = 64.87, p < 0.001$), and between number of items and target (present/absent) ($F_{2,34} = 24.16, p < 0.001$). The three-way interaction between distractors, number of items, and target (present/absent) was significant ($F_{4,68} = 10.39, p < 0.001$). Figure 8 shows the search function together with the slope for each condition.

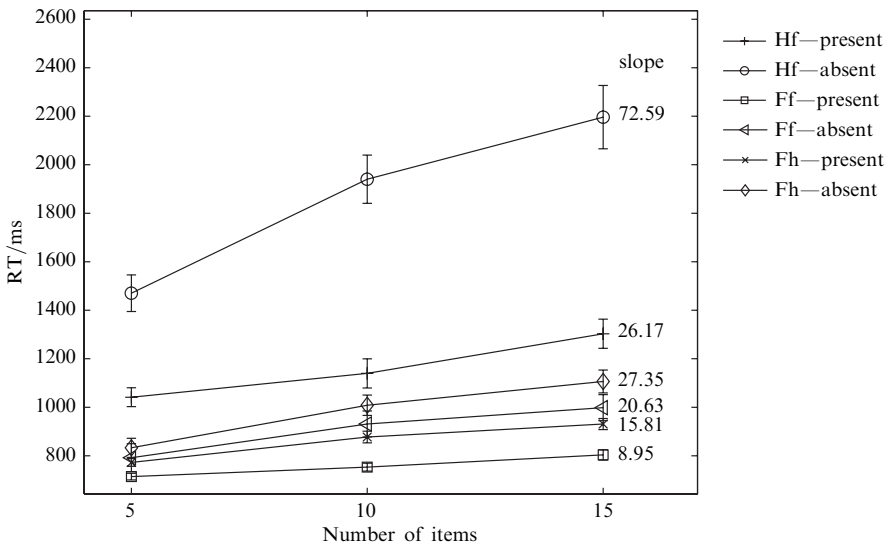


Figure 8. Experimental results for block 1 in experiment 1. The search functions depict the search efficiency for search displays with different distractor items (see legend).

5.2.2 Block 2: Large and small letters. A three-way ANOVA showed three significant main effects for item size ($F_{1,17} = 149.48, p < 0.001$), number of items ($F_{2,34} = 61.13, p < 0.001$), and target (present/absent) ($F_{1,17} = 47.86, p < 0.001$). There was a significant interaction between item size and number of items ($F_{2,34} = 25.67, p < 0.001$), between item size and target (present/absent) ($F_{1,17} = 76.26, p < 0.001$), and between number of items and target (present/absent) ($F_{2,34} = 21.57, p < 0.001$). The three-way interaction between distractors, number of items, and target (present/absent) was not significant ($F_{2,34} = 1.82, p = 0.178$). Figure 9 shows the search function together with the slope for each condition. A comparison between the global disparity condition (Fh as distractor) and global similarity (Hf as distractor) shows no significance ($F_{2,34} = 1.784, p = 0.183$).

5.2.3 Comparison between block 1 and block 2. The object of this analysis is to test whether the predictions of the computer simulations are confirmed by the experiment. A two-way ANOVA compared the search for large letters (block 2) and the condition VS1 in block 1 (global disparity). The results indicated significant main effect of display size ($F_{1,17} = 31.56, p < 0.001$) and block ($F_{1,17} = 75.31, p < 0.001$). More importantly, there is no significant interaction between display size and block ($F_{2,34} = 1.44, p = 0.251$). This confirms the prediction of the simulations.

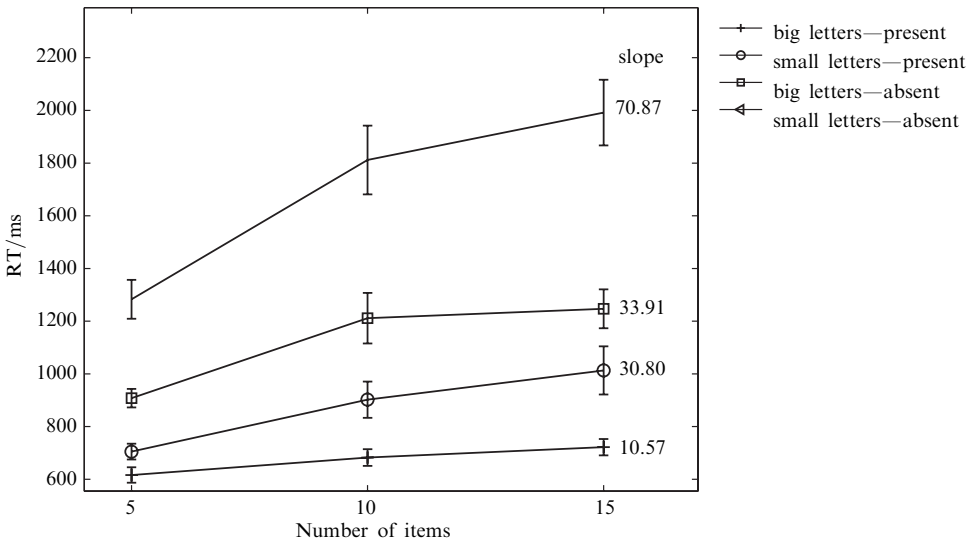


Figure 9. Experimental results: visual search through large and small letters (block 2).

5.3 Discussion

The results showed that search in the global-disparity condition was as efficient as search through items matching the size of the hierarchical patterns at the global level. Therefore, the experiment confirms the prediction of the model with resolution hypothesis that, in the global disparity condition, the search is dominated by properties of search display at the global level. However, there are several problems with this experiment. First, using displays with large items as baseline is a questionable approach, as it makes a wide range of assumptions about visual search as a whole. Second, not rejecting the null-hypothesis obviously is a very weak support for a theoretical prediction. Third, contrary to the simulation results, search in the global-disparity condition is not different from search in the global-similarity condition. This is due to the fact that search at the global level is not very efficient, in particular not as efficient as in the simulations. The parameters of the simulations were chosen in a way that the target pops out at the global level, whereas search in the condition of global similarity was highly inefficient. This setting allowed us to demonstrate clearly the difference between ‘without resolution’ hypothesis and ‘with resolution’ hypothesis. Human behaviour did not show such a clear-cut effect with the same stimuli. In order to provide a clearer proof for the prediction of the model, we conducted an additional experiment in which we expected a pop-out search at the global level and, thus, an efficient search in the global-disparity condition, if the prediction of the model with the resolution hypothesis is confirmed.

6 Experiment 2

In this experiment we aimed at reproducing the conditions of the simulations by making search at the global level highly efficient. Therefore, we chose two geometric shapes, diamond and square, as it is known that search for diagonal lines amongst horizontal and vertical lines is highly efficient (eg Wolfe et al 1992).

6.1 Procedure

6.1.1 Participants. The participants were nine students from the University of Birmingham. All participants were reported having normal or corrected-to-normal vision.

6.1.2 *Stimuli*. The stimuli consisted of global geometric shapes (diamond or square) constructed of small letters (F or E). Figure 10a shows the four items used as target and distractors for the visual-search task. The global shapes, squares and diamonds, cover an area of $1.7 \text{ deg} \times 1.7 \text{ deg}$ and $2.4 \text{ deg} \times 2.4 \text{ deg}$, respectively. The small letters covered an area of $0.28 \text{ deg} \times 0.28 \text{ deg}$. The distance between the letters was 0.17 deg . The stimuli are defined as follows: (i) ‘diamond-f’ (Df)—a global diamond constructed of small F letters, (ii) ‘diamond-e’ (De)—a global diamond constructed of small E letters, (iii) ‘square-f’ (Sf)—a global square constructed of small F letters and (iv) ‘square-e’ (Se)—a global square constructed of small E letters. The search displays were made of 1, 5, 12, 18, or 25 items that appeared randomly anywhere on a 19 inch monitor. Figures 10b and 10c show examples of a 12-items display.

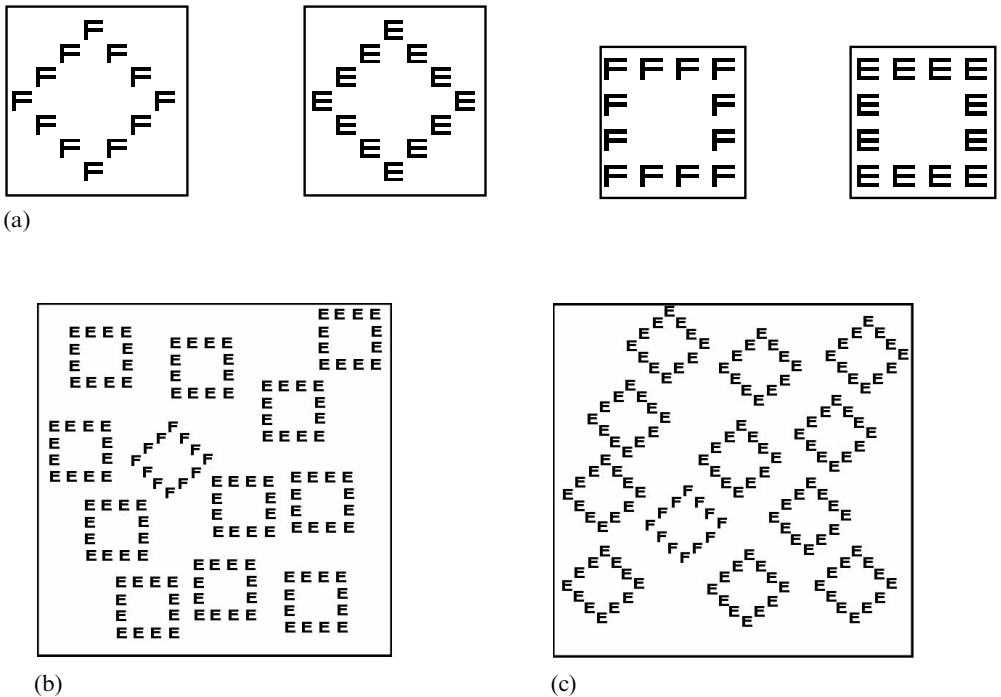


Figure 10. Stimuli in experiment 2. (a) Hierarchical stimuli used as target (Df) and distractors (De, Sf, Se) in the second experiment; (b) typical display for the condition of global disparity (VS1) with Se as distractor; (c) typical display for the condition of global similarity (VS2) with De as distractor.

6.1.3 *Design*. The presentation of the stimuli and the timing and data collection were under the control of a standard PC with a 19 inch monitor. ‘Z’ and ‘M’ were used as response buttons for target present and target absent, respectively. The monitor was viewed from a distance of 50 cm. The target was a large diamond made out of small letters F and could be present or absent. Each display remained up until the computer picked up a key-press. The subjects were instructed to respond as fast and accurately as possible. There were three conditions: one corresponding to global similarity (VS2) and two corresponding to global disparity (VS1). In the case of global similarity the distractor items are defined by a large ‘square’ constructed of small F letters (square-f). In the case of global disparity the distractors are given by a large diamond constructed of small letters, which were either E or F (diamond-e or diamond-f). After 15 practice trials the subjects had to do 480 trials interrupted every 120 trials by a break. No feedback was given.

6.2 Results

The results are shown in figure 11. A three-way ANOVA revealed that three main effects were significant: that of distractor ($F_{2,16} = 66.69$, $p < 0.001$), display size ($F_{4,32} = 43.62$, $p < 0.001$), and target (present/absent) ($F_{1,8} = 37.74$, $p < 0.001$). There was a significant interaction between distractor and display size ($F_{8,64} = 36.21$, $p < 0.001$), between distractor and target (present/absent) ($F_{4,32} = 40.50$, $p < 0.001$), and between display size and target (present/absent) ($F_{4,32} = 24.45$, $p < 0.001$). The three-way interaction between distractors, display size, and target (present/absent) was significant ($F_{8,64} = 11.18$, $p < 0.001$) as well.

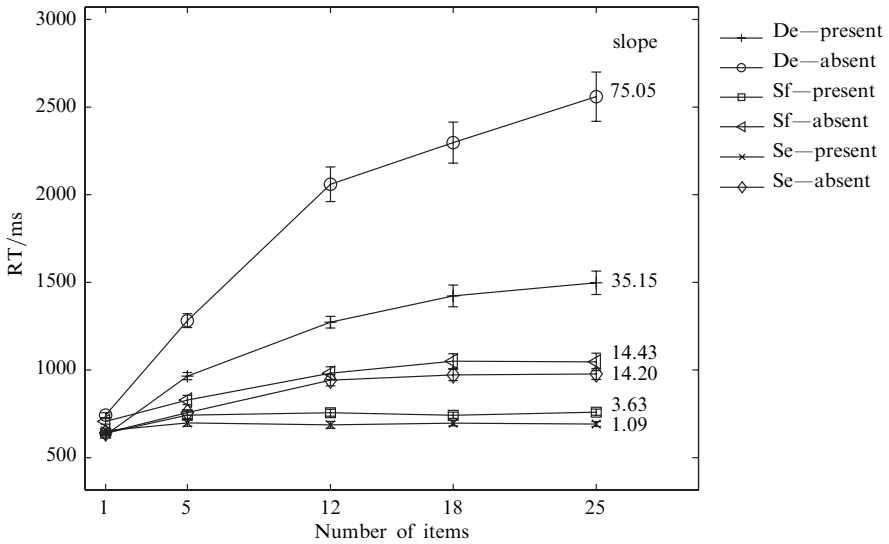


Figure 11. Search functions of experiment 2. The legend indicates the different distractor items.

An ANOVA between global disparity (Se as distractor) and global disparity (Sf as distractor) showed a significant effect ($F_{4,32} = 20.22$, $p < 0.001$), thus confirming the prediction of the model.

6.3 Discussion

The experiment shows that search in the global-disparity condition is highly efficient whereas search in the global similarity condition is significantly less efficient. This second experiment confirms the prediction of the model with resolution hypothesis whereby search efficiency in the global disparity condition is dominated by the search efficiency at the global level.

7 Conclusion

In this paper we have presented a neurodynamic model based on the biased-competition hypothesis that allowed us to study the interaction between attention and the spatial-resolution hypothesis in the context of visual search with hierarchical patterns. The spatial-resolution hypothesis was implemented in a simplified way by assuming different time constants for processing the different spatial frequencies. Subsequent simulations predicted efficient search for the case of global disparity, ie when the global information was sufficient for finding the target amongst distractors. On the contrary, assuming the same time constants for processing the different spatial frequencies predicted inefficient search for the same search condition (global disparity). On the other hand for the global similarity, ie when target and distractor were indistinguishable at the global level, visual search was inefficient with or without implementation

of the resolution hypothesis. These predictions were tested in two novel visual-search experiments. The experimental results confirmed the hypothesis that different spatial frequencies are processed with different time constants.

Our results are also consistent with evidence from visual-search experiments that participants can restrict their search to items with set features.⁽¹⁾ For instance, Egeth et al (1984) showed that when participants search for a red O amongst black Os and red Ns, the length of reaction time is only determined by the number of red Ns and not by the number of the black Os. However, this assertion assumes that the processing levels are equivalent to simple features. This certainly remains an open research question. Moreover, our simulations show that there are possible alternative implementations for processing levels conceivable.

Finally, we would like to point out that the different time constant assumed by the spatial-resolution hypothesis can also emerge from a network as suggested by the ‘high-resolution buffer theory’ (Lee et al 1998; Lee and Mumford 2003). In the high-resolution buffer theory recurrent feedforward/feedback loops between V1 and extrastriate areas serve to integrate global and local information in a Bayesian inference framework. The results of these network dynamics can be described in a simplified way by assuming different time constants for different spatial frequencies in V1 neurons as in our implementation of the spatial-resolution hypothesis. As reviewed in Lee and Mumford (2003), a large body of experimental evidence at neurophysiological and behavioural levels, together with the fact that it is consistent with mathematical and computational concepts from computer vision makes this implementation, in our view, highly plausible.

Acknowledgments. We thank Glyn Humphreys for valuable discussions. This work was supported by a grant from the Royal Society No. 15619.

References

- Badcock J C, Whitworth F A, Badcock D R, Lovegrove W J, 1990 “Low-frequency filtering and the processing of local–global stimuli” *Perception* **19** 617–629
- Breitmeyer B, Breier J, 1994 “Effects of background color on reaction time to stimuli varying in size and contrast: Inferences about human M channels” *Vision Research* **34** 1039–1045
- Chelazzi L, Miller E, Duncan J, Desimone R, 1993 “A neural basis for visual search in inferior temporal cortex” *Nature* **363** 345–347
- Corchs S, Deco G, 2002 “Large-scale neural model for visual attention: Integration of experimental single-cell and fMRI data” *Cerebral Cortex* **12** 339–348
- Corchs S, Deco G, 2004 “Feature-based attention in human visual cortex: simulation of fMRI data” *NeuroImage* **21** 36–45
- Deco G, Lee T S, 2002 “A unified model of spatial and object attention based on inter-cortical biased competition” *Neurocomputing C* **44–46** 775–781
- Deco G, Pollatos O, Zihl J, 2002 “The time course of selective visual attention: theory and experiments” *Vision Research* **32** 2925–2945
- Deco G, Rolls E T, Horwitz B, 2004 “Integrating fMRI and single-cell data of visual working memory” *Neurocomputing* **58–60** 729–737
- Deco G, Zihl J, 2001 “Top–down selective visual attention: A neurodynamical approach” *Visual Cognition* **8** 119–140
- Desimone R, Duncan J, 1995 “Neural mechanisms of selective attention” *Annual Review of Neuroscience* **18** 193–222
- Desimone R, Ungerleider L, 1989 “Neural mechanisms of visual processing in monkeys”, in *Handbook of Neurophysiology* volume 2 (Amsterdam: Elsevier) pp 267–299
- Downing P E, 2000 “Interactions between visual working memory and selective attention” *Psychological Science* **11** 467–473
- Duncan J, Humphreys G, Ward R, 1997 “Competitive brain activity in visual attention” *Current Opinion in Neurobiology* **7** 255–261
- Duncan J, Humphreys G W, 1989 “Visual search and stimulus similarity” *Psychological Bulletin* **96** 433–458

⁽¹⁾ We would like to thank the anonymous reviewer for drawing our attention to this similarity.

- Egeth H E, Virzi R A, Garbat H, 1984 "Searching for conjunctively defined targets" *Journal of Experimental Psychology: Human Perception and Performance* **10** 32–39
- Gross C G, Rocha-Miranda C E, Bender D B, 1972 "Visual properties of neurons in inferotemporal cortex" *Journal of Neurophysiology* **35** 96–111
- Heinke D, Humphreys G W, 2003 "Attention, spatial representation and visual neglect: Simulating emergent attention and spatial memory in the Selective Attention for Identification Model (SAIM)" *Psychological Review* **110** 29–87
- Kaplan E, Shapley R, 1982 "X and Y cells in the lateral geniculate nucleus of macaque monkeys" *Journal of Physiology* **330** 125–143
- Kimichi R, 1992 "Primacy of wholistic processing and global/local paradigm" *Psychological Bulletin* **112** 24–38
- Kinchla R, 1974 "Detecting target elements in multi-element arrays: A confusability model" *Perception & Psychophysics* **15** 149–158
- Lee T S, 1996 "Image representation using 2d Gabor wavelets" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(10) 959–971
- Lee T S, Mumford D, 2003 "Hierarchical Bayesian inference in the visual cortex" *Journal of the Optical Society of America A* **20** 1434–1448
- Lee T S, Mumford D, Romero R, Lamme V A F, 1998 "The role of primary visual cortex in higher level vision" *Vision Research* **38** 2429–2454
- Livingstone M, Hubel D, 1984 "Anatomy and physiology of a color system in the primate visual cortex" *Journal of Neuroscience* **4** 309–356
- Lovegrove W, Lehmkuhle J, Baro S, Garzia R, 1991 "The effect of uniform field flicker and blurring on the global precedence effect" *Bulletin of the Psychonomic Society* **29** 289–291
- Michimata C, Okubo M, Mugishima Y, 1999 "Effects of background color on the global and local processing of hierarchically organized stimuli" *Journal of Cognitive Neuroscience* **11** 1–8
- Moran J, Desimone R, 1985 "Selective attention gates visual processing in the extrastriate cortex" *Science* **229** 782–784
- Navon D, 1977 "Forest before trees: the precedence of global features in visual perception" *Cognitive Psychology* **9** 353–383
- Posner M I, Snyder C R R, Davidson B J, 1980 "Attention and the detection of signals" *Journal of Experimental Psychology: General* **109** 160–174
- Reynolds J, Chelazzi L, Desimone R, 1999 "Competitive mechanisms subserve attention in macaque areas V2 and V4" *Journal of Neuroscience* **19** 1736–1753
- Rolls E, Deco G, 2002 *Computational Neuroscience of Vision* (Oxford: Oxford University Press)
- Shulman G L, Sullivan M A, Gish K, Sakoda W J, 1986 "The role of spatial-frequency channels in the perception of local and global structure" *Perception* **15** 259–273
- Shulman G L, Wilson J, 1987 "Spatial frequency and selective attention to local and global information" *Perception* **16** 89–101
- Ungerleider L G, Mishkin M, 1982 "Two cortical visual systems", in *Analysis of Visual Behavior* Eds D I Ingle, M A Goodale, R J W Manfield (Cambridge, MA: MIT Press) pp 549–586
- Wolfe J M, Friedman-Hill S R, Stewart M I, O'Connell K M, 1992 "The role of categorization in visual search for orientation" *Journal of Experimental Psychology: Human Perception and Performance* **28** 34–49
- Wolfe J M, O'Neill P, Bennett S C, 1998 "Why are there eccentricity effects in visual search? Visual and attentional hypotheses" *Perception & Psychophysics* **60** 140–156
- Wong E, Weisstein N, 1983 "Sharp targets are detected better against a figure, and blurred targets are detected better against a background" *Journal of Experimental Psychology: Human Perception and Performance* **9** 194–202

APPENDIX: Mathematical description of the model

The mathematical formulation of the system was as follows. We consider a pixelised grey-scale image given by a matrix. The subindices denote the spatial position of the pixel. Each pixel value is given a grey value coded on a scale between 0 (black) and 255 (white). The first step in the pre-processing consists in removing the DC component of the image (ie the mean value of the grey-scale intensity of the pixels), a process probably done in the lateral geniculate nucleus (LGN) of the thalamus. The visual representation in LGN is essentially a contrast-invariant pixel representation of the image, ie each neuron encodes the relative brightness value at one location in visual space referred to the mean value of the image brightness (Kaplan and Shapley 1982). Feedforward connections to a layer of V1 neurons serve to extract simple features. Previous investigations have suggested that simple cells in the primary visual cortex can be modelled by 2-D Gabor functions (Lee 1996). The 2-D Gabor functions are local spatial bandpass filters that achieve the theoretical limit for conjoint resolution of spatial and frequency information, ie in the 2-D spatial and 2-D Fourier domains. The Gabor receptive fields have five degrees of freedom given essentially by the product of an elliptical Gaussian and a complex plane wave. The first two degrees of freedom are the 2-D location of the centre of the receptive field, the third is the size of the receptive field, the fourth is the orientation of the boundaries separating excitatory and inhibitory regions, and the fifth is the symmetry of the field. This fifth degree of freedom is given in the standard Gabor transformation by the real and imaginary part, ie by the phase of the complex function representing it, whereas biologically this can be done by combining pairs of neurons with even or odd receptive fields. Let us consider the experimental neurophysiological constraints. There are three constraints fixing the relation between width, height, orientation, and spatial frequency. The first constraint postulates that the aspect ratio of the elliptical Gaussian envelope is 2 : 1. The second constraint postulates that the orientation is aligned with the long axis of the elliptical Gaussian. The third constraint assumes that the half-amplitude bandwidth of the frequency response is about 1 to 1.5 octaves along the optimal orientation. Further, we assume that the mean is zero in order to have an admissible wavelet basis. A family of discretised 2-D Gabor wavelets that satisfy the wavelet theory and the neurophysiological constraints for simple cells as given by Lee (1996) is:

$$G_{kpql}(x, y) = a^{-k} \Psi_{\theta_l} [a^{-k}(x - 2p) - a^{-k}(y - 2q)],$$

$$\Psi_{\theta_l} = \Psi [x \cos(l\theta_0) + y \sin(l\theta_0) - x \sin(l\theta_0) + y \cos(l\theta_0)] \quad (\text{A1})$$

and the mother wavelet is given by

$$\Psi(x, y) = \{\exp[-\frac{1}{8}(4x^2 + y^2)] [\exp(ikx) - \exp(-\frac{1}{2}\kappa^2)]\} / \sqrt{(2\pi)}. \quad (\text{A2})$$

In the above equations θ_0 denotes the step size of each angular rotation, l the index of rotation corresponding to the preferred orientation θ_l , k denotes the octave, and pq the position of the receptive field centre. In this form, the receptive fields at all levels cover the spatial domain in the same way. In this work we chose $\kappa = \pi$, a spatial-frequency bandwidth corresponding to one octave. The neurons in the pools in V1 have receptive fields performing a Gabor wavelet transform. Let us denote by I_{kpql}^{V1} the sensorial input activity to a pool in V1 which is sensitive to a determined spatial frequency given at octave k , to a preferred orientation defined by the rotation index l , and to stimuli at the centre location specified by the indices pq . The sensorial input activity to a pool in V1 is therefore defined by the module of the convolution between the corresponding receptive fields and the image. Since, in our numerical simulations, the system needs only to learn a small number of objects (usually 2–4), in our current implementation, for simplicity, we temporarily eliminate the V4 module and fully connect

V1 and IT cell assemblies directly together. We implement translation invariance by the attentional intermodular biasing interaction between pools in the modules V1 and PP. For each V1 neuron, the gain modulation observed decreases as the actual point where attention is being focused moves away from the centre of the receptive field in a Gaussian-like form. Consequently, the connections with the pools in the PP module are specified such that the modulation is Gaussian. Let us define in the PP module a pool A_{ij}^{PP} for each location ij in the visual field. The mutual connections (ie bilateral) between a pool A_{kpql}^{V1} in V1 (or V4) and a pool A_{ij}^{PP} in PP are therefore defined by

$$W_{pqij} = A \exp\{[(i-p)^2 + (j-q)^2]/2S^2\} - B. \quad (\text{A3})$$

Let us now define the neurodynamic equations that regulate the evolution of the whole system. The activity level of the input current in the V1 module is given by

$$\tau \frac{\partial}{\partial t} A_{kpql}^{\text{V1}}(t) = -A_{kpql}^{\text{V1}}(t) + aF[A_{kpql}^{\text{V1}}(t)] - bF[A_k^{I, \text{V1}}(t)] + I_{kpql}^{\text{V1}}(t) + I_{pq}^{\text{V1-PP}}(t) + I_{kpql}^{\text{V1-IT}}(t) + I_0 + v, \quad (\text{A4})$$

where the attentional biasing due to the intermodular ‘where’ connections with the pools in the parietal module PP, $I_{pq}^{\text{V1-PP}}$, is given by

$$I_{pq}^{\text{V1-PP}} = \sum_{i,j} W_{pqij} F[A_{ij}^{\text{PP}}(t)], \quad (\text{A5})$$

and the attentional biasing due to the intermodular ‘what’ connections with the pools in the temporal module IT, $I_{kpql}^{\text{V1-IT}}$, is defined by

$$I_{kpql}^{\text{V1-IT}} = \sum_{c=1} w_{ckpql} F[A_c^{\text{IT}}(t)], \quad (\text{A6})$$

where w_{ckpql} is the connection strength between the V1 pool A_{kpql}^{V1} and the IT pool, A_c^{IT} corresponding to the coding of a specific object category c . We assume that the IT module has c pools corresponding to different object categories. We implement the resolution hypothesis by assuming different latencies τ in equation (A4) for pools corresponding to different octaves. Higher-resolution octaves have a smaller τ than the ones corresponding to lower resolutions. For each spatial-frequency level, a common inhibitory pool is defined. The current activity of the inhibitory pools obeys the following equations:

$$\tau_p \frac{\partial}{\partial t} A_k^{I, \text{V1}}(t) = -A_k^{I, \text{V1}}(t) + c \sum_{p,q,l} F[A_{kpql}^{\text{V1}}(t)] - dF[A_k^{I, \text{V1}}(t)]. \quad (\text{A7})$$

The current activity of the excitatory pools in the posterior parietal module PP are given by

$$\tau \frac{\partial}{\partial t} A_{ij}^{\text{PP}}(t) = -A_{ij}^{\text{PP}}(t) + aF[A_{ij}^{\text{PP}}(t)] - bF[A^{I, \text{PP}}(t)] + I_{ij}^{\text{PP-V1}}(t) + I_{ij}^{\text{PP,A}} + I_0 + v, \quad (\text{A8})$$

where $I_{ij}^{\text{PP,A}}$ denotes an external attentional-object-specific top-down bias, the intermodular attentional biasing $I_{ij}^{\text{PP-V1}}$ through the connections with the pools in the module V4 is

$$I_{ij}^{\text{PP-V1}}(t) = \sum_{k,p,q,l} W_{pqij} F[A_{kpql}^{\text{V1}}(t)], \quad (\text{A9})$$

and the activity current of the common PP inhibitory pool evolves according to

$$\tau_p \frac{\partial}{\partial t} A^{I, \text{PP}}(t) = -A^{I, \text{PP}}(t) - dF[A^{I, \text{PP}}(t)] + c \sum_{i,j} F[A_{ij}^{\text{PP}}(t)]. \quad (\text{A10})$$

The dynamics of the inferotemporal module IT is given by

$$\tau \frac{\partial}{\partial t} A_c^{\text{IT}}(t) = -A_c^{\text{IT}}(t) + aF[A_c^{\text{IT}}(t)] - bI^{I, \text{IT}}(t) + I_c^{\text{IT-V1}} + I_c^{\text{IT,A}} + I_0 + v, \quad (\text{A11})$$

where $I_c^{\text{IT},A}$ denotes an external spatial attentional top–down bias and the intermodular attentional biasing $I_c^{\text{IT-V1}}$ between IT and V1 pools is

$$I_c^{\text{IT-V1}}(t) = \sum_{k,p,q,l} w_{ckpql} F[A_{k p q l}^{\text{V1}}(t)], \quad (\text{A12})$$

and the activity current of the common PP inhibitory pool evolves according to

$$\tau_p \frac{\partial}{\partial t} A^{I,\text{IT}}(t) = -A^{I,\text{IT}}(t) + c \sum_c F[A_c^{\text{IT}}(t)] - dF[A^{I,\text{IT}}(t)]. \quad (\text{A13})$$

Training

During a learning phase a representation for each object is stored. The external attentional-location-specific bias in PP, $I_{ij}^{\text{PP},A}$, is set so that only the pool ij corresponding to the spatial location where the training object is, receives a positive bias. In this way, spatial attention defines the location of the object to be learnt. The external attentional-object-specific bias in IT, $I_c^{\text{IT},A}$, is similarly set so that only the pool c that will identify the object receives a positive bias. Therefore, we define in a supervised way the identity of the object. After presentation of a given stimulus, ie a specific object at a specific location, and the corresponding external bias, the system evolves until convergence. After convergence, the V1–IT connections w_{ckpql} are trained through the following Hebbian rule:

$$w_{ckpql}(\text{new}) = w_{ckpql}(\text{old}) + \eta F[A_c^{\text{IT}}(t)] F[A_{k p q l}^{\text{V1}}(t)] \quad (\text{A14})$$

with t large enough (ie after convergence) and η the learning coefficient. This procedure is repeated for all objects at all possible locations until the weights converge. During the recognition phase there are two possibilities: (a) to search for a specific object (visual search), or (b) to identify an object at a specified spatial location (object recognition). In the case of visual search (a) the stimuli are presented and the external attentional-object-specific bias in IT, $I_c^{\text{IT},A}$, is set so that only the pool c corresponding to the category of the object to be searched receives a positive bias; the external attentional-location-specific bias in PP, $I_c^{\text{PP},A}$, is set to zero everywhere. The external attentional bias, $I_c^{\text{IT},A}$, drives the competition in the IT module so that the pool corresponding to the searched object wins. The intermodular attentional modulation between IT and V1 biases the competition in V1 so that all features detected from the retinal inputs at different positions that are compatible with the specific object to be searched now win. Finally, the intermodular attentional bias between V1 and PP drives the competition in V1 and in PP so that only the spatial location in PP and the associated V1 pools compatible with the presented stimulus, and with the category of the object to be searched, will remain active after convergence. In this form, the final activation state is neurodynamically driven by a combination of stimulus, external top–down bias and intermodular bias, in a fully parallel way. ‘Attention’ is not a mechanism explicitly involved in the competition but rather an emergent effect that supports the dynamic evolution to a state where all constraints are satisfied (see also Heinke and Humphreys 2003). For object recognition following cueing to a spatial location [(b) above], the external attentional bias in PP, $I_{ij}^{\text{PP},A}$, is set so that only the pool associated with the cued spatial location receives a positive bias. The other external bias, $I_c^{\text{IT},A}$, is zero everywhere. In this case, the dynamics evolve such that, in PP, only the pool associated with the top–down biased spatial location will win. This in turn drives the competition in V1 such that only the pools corresponding to features of the stimulus at that location will win, and this biases the dynamics in IT such that only the pool identifying the class of the features at that position will remain active. This indicates the category of the object at that predefined spatial location.

ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 36 2007

www.perceptionweb.com

Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.