ELSEVIER

A dynamic model of how feature cues guide spatial attention

Fred H. Hamker

Division of Biology 139-74, California Institute of Technology, Pasadena, CA 91125, USA
Allgemeine Psychologie, Psychologisches Institut II, Westf. Wilhelms-Universität, Fliednerstrasse 21, 48149 Münster, Germany

Abstract

We will describe a computational model of attention which explains the guidance of spatial attention by feedback within a distributed network. We hypothesize that feedback within the ventral pathway transfers the target template from prefrontal areas into intermediate areas like V4. The oculomotor circuit consisting of FEF, LIP and superior colliculus picks up this distributed activity and provides a continuous spatial reentry signal from premotor cells. In order to test this hypothesis, we simulate two experiments that require a response given a color cue. The first experiment indicates a parallel feature-based enhancement prior to any spatial selection. If two targets are behaviorally relevant, as in the second experiment, experimental findings indicate that subjects split their attention between two locations containing the searched feature. Our simulation results suggest that the split in attention between two foci is a transient effect occurring during competition. We predict that the time after cue presentation determines the state of this competition and ultimately the distribution of attention at different locations. In addition we provide simulation results to explain how reentrant processing through the oculomotor circuit might lead to variations of the time for target detection in visual search.
© 2003 Elsevier Ltd. All rights reserved.

1. Introduction

Experiments have shown that we can easily use color information to guide our attention or gaze in visual search (Findlay, 1997; Kaptein, Theeuwes, & Van der Heijden, 1995; Motter & Belky, 1998). Such goal-directed attention can also be applied if the target shape is known (Caputo & Guerra, 1998; Findlay, 1997; Ghirardelli & Egeth, 1998), although the effect of using color is usually stronger. As a result search has only to be performed within a pre-determined subgroup (Cepeda, Cave, Bichot, & Kim, 1998; Duncan & Humphreys, 1992; Egeth, Virzi, & Garbart, 1984; Wolfe, Cave, & Franzel, 1989). Moore and Egeth (1998) showed that feature-based attention increased performance (decreased reaction time) when displays were present until the response was made. However, in data-limited conditions, in which stimuli are presented only briefly before being replaced by a mask, feature-based attention

did not increase performance (measured by the accuracy of report).

Previous models of attention have only touched possible underlying computational mechanisms of how feature cues (e.g. color) guide spatial attention. It was suggested that feature inhibition can suppress non target features within feature maps and thus, diminish their activity within a map of locations which determines the locus of attention (Treisman & Sato, 1990) or that bottom–up and top–down feature maps add up within an activation map which then selects the location of maximal activation (Wolfe et al., 1989). In most models attention is defined by determining a locus of a unique spatial focus (Ahmad, 1991; Koch & Ullman, 1985; Olshausen, Anderson, & van Essen, 1993; Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe et al., 1989). In addition, prior selection by spatial attention is typically considered to be necessary for items to be reported. For example: "If the task requires selecting the red item for report, then the system should be configured such that only activity from the 'red' feature map drives the attentional network, causing selection of red items. If the display contains only a single red item, it

E-mail addresses: fred@klab.caltech.edu, fhamker@uni-muenster.de (F.H. Hamker).
URL: http://www.klab.caltech.edu/~fred.html

will be selected, activity from all feature types in its spatial location will then be allowed to pass through the recognition network, and the output of the recognition network will be the identity of the red item." (Mozer & Sitton, 1998).

In the present manuscript we address in more detail the question of such a goal-directed selection. We suggest a distributed processing in which selection occurs within a continuous dynamic process. Specifically, we show how feedback within the ventral pathway and a slow reentry from premotor cells can account for the observations in two experiments.

We first describe two experiments that require a response given a feature cue. We then introduce our model and simulate these experiments with the model. In addition, we simulate a feature and conjunction visual search task in order to further illustrate parallel and serial search in our model.

1.1. Experiment 1

The effect of cuing by color has been investigated in an earlier electrophysiological experiment (Fig. 1). The task of this experiment required to report the orientation of an item that matches the color of the fixation point. Since the display during the stimulus presentation period contains several possible targets, the monkey had to wait until the display contained only one target. However, even during the stimulus presentation period,

V4 neurons showed an enhanced activity if the presented color or luminance items matched the target (Motter, 1994a, 1994b). This dynamic effect is thought to occur in parallel across the visual field, segmenting the scene into possible candidates and background.

The interpretation in terms of a parallel process is not proven. The process of mapping a receptive field may bias the location, such that the monkey tends to spatially attend to the observed location (Newsome, 1996). Since the recordings were not done concurrently with multiple electrodes, an alternative interpretation in terms of a fast moving but serial attentional focus cannot be completely excluded (Chelazzi, 1999; Motter, 1994b; Newsome, 1996).

Nevertheless, major variations of activity indicative of a serial activation of cells have not yet been observed (Chelazzi, Duncan, Miller, & Desimone, 1998; Motter, 1994b), weakening the fast moving spotlight theory. Furthermore, there is growing evidence for a spatially independent, feature-based, top–down component of attention, as shown in the motion system using single cell recordings in area MT (Treue & Martínez Trujillo, 1999) as well as in the motion and color dimension using fMRI (Saenz, Buracas, & Boynton, 2002) and in a divided attention task (Saenz, Buracas, & Boynton, 2003). Due to the unimportant locational information—the selection criterion in most models of attention—the experiment of Motter investigates a so far not fully explained attentional effect and Chelazzi (1999) demands
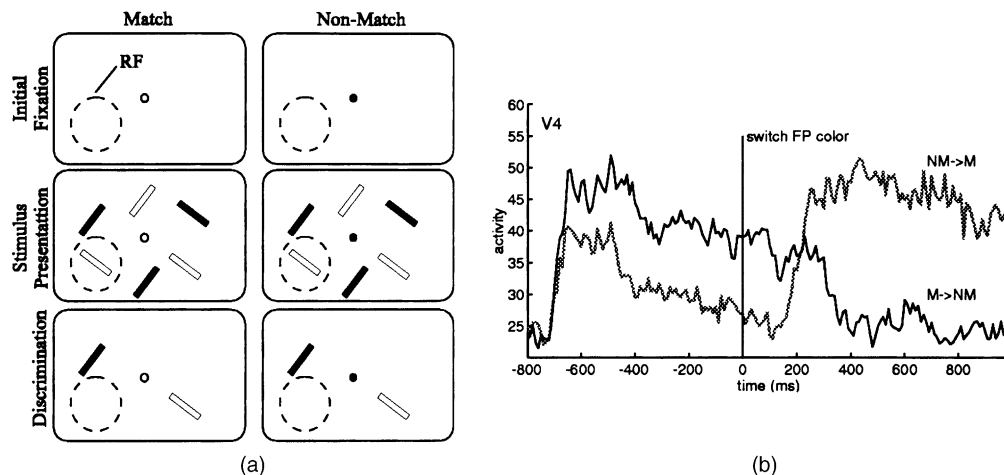

Fig. 1. Match and non-match conditions in the experiment of Motter (1994a, 1994b). (a) In either task a monkey is initially faced with a fixation point (FP), whose color also serves as a cue. Then, six oriented bars are presented. Three of them have the same color (e.g. red) as the fixation spot. Thus, during this stimulus presentation period, all three items could be the final target. Finally, the scene is reduced to a target and a distractor by removing all other stimuli and the monkey has to specify the orientation (right, left) of the target. In the match condition a color or luminance match occurs between the fixation spot and the item in the receptive field (dotted line). (b) The data shows the temporal course of activity in the match and non-match condition. The scene containing the six stimuli is presented at −800 ms. Cells representing the potential target object show an enhanced activity regardless of their location. If the fixation point color switches to another color at 0 ms, the activity follows the definition of the target. Neurons previously representing potential targets change into distractors and vice versa. The figure shows the original data (the published data (Motter, 1994b) does not show the activity after stimulus onset). The data reflects the average over 30 neurons, that showed the largest differences in the interval after the cue switch, out of the entire 109 cell population observed. 81 of the 109 cells showed a M/NM differentiation. The histograms are based on 10 ms bins.

that "it remains to be established what role it might play in visual search".

1.2. Experiment 2

Although attention can be cued by features, the selection itself might be location-based (but see Duncan (1984) and Vecera and Farah (1994) for alternative explanations). Spatial theories of attention consider selection by location as the ultimate source, "but selection by other features such as color is possible by first determining the location containing the target color, and then selecting that location" (Kim & Cave, 2001). In an experiment of Bichot, Cave, and Pashler (1999) subjects had to compare two target shapes among six distractors. Targets were indicated by their color. The display of shapes (cue) was followed by a "probe" display of letters appearing in these shapes in which the subjects had to report as many letters as possible (Fig. 2). The time between the cue and the probe is called the SOA. The latter task is intended to measure the amount of attention resulting from the primary task at each of those locations. Although subjects often reported only the letter within one of two target shapes correctly (about 60–70% error rate on the second letter when the first one

is correctly reported), the accuracy within the second target shape was still significantly higher than within the distractor shapes (85–90% error). This result does neither favor an object-based selection that occurs in higher, location-insensitive areas nor a selection of a single spatial location. Bichot et al. (1999) also exclude an interpretation of a fast moving visual spotlight: The probe display is only visible for 60 ms. Although measurements based on search slopes can indicate a processing of an item that lies much below 60 ms, actual estimates of the time needed to identify a stimulus at one location and shift to another are typically 200 ms or even longer (Kröse & Julesz, 1989; Eriksen & Yeh, 1985). Thus, Bichot et al. (1999) argue that their data cannot be explained by a sequential shift of attention and conclude that spatial attention can be split across two noncontiguous locations. Such a facilitated report of items at two noncontiguous locations has also been found in other experiments, but the items were cued directly by their locations (Awh & Pashler, 2000).

2. A model of feature feedback and spatially organized reentry

2.1. Outline

Our long term strategy is to develop a single model in order to explain multiple attentional phenomena. At present we investigate how feature feedback and spatially organized reentry can account for the observations in the aforementioned experiments. Both require the detection of a color feature before information at its location can be reported. Thus, similar brain mechanisms should be activated in either task, the search for objects that match the color of the fixation cue and the recognition of the two target shapes by their color. So this raises the question how a feature cue like color can capture spatial attention.

Motter (1994a) suggests that "a sequential combination of [...] two processes, initially a full-field prefocal attentive selection based on features followed by a spatially restricted focal attentive process [...], offers an interesting physiological model of selective attention expressed within single neurons in V4." This view is similar to the two-stage model of visual search proposed by Hoffman (1979). According to Hoffmans hypothesis a similarity measure originates from a parallel comparison of each item to all memory set items. A decision at this point is possible, but a low signal to noise ratio, as is often the case in visual search, produces a high error rate. Thus, items are serially transferred to a detailed processing (attended) in order of decreasing similarity. However, this raises the question of how the processing of features and the processing of the location are bound together in a distributed network.
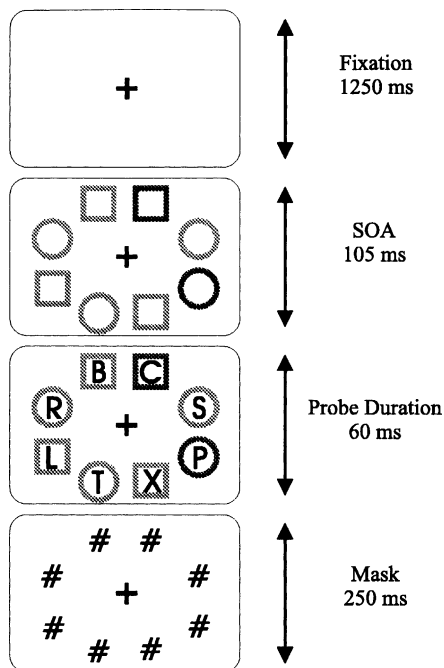

Fig. 2. Experiment of Bichot et al. (1999) to investigate how much attentional selection is mediated by location. After a fixation period the eight stimuli of the primary task appear. Two match the target color while the remaining have a uniform distractor color. The primary task was to determine whether the two target shapes were the same or not. After a fixed stimulus onset asynchrony (SOA) of 105 ms letters were presented within each shape. The secondary task was to report as many letters as possible.

Concepts pointing in this direction have been worked out within theories of selection for action (Allport, 1987), selection for parameter specification (Neumann, 1987, 1990; van der Heijden, 1992), selection for perception and motor-action (Goodale & Humphrey, 1998; Schneider, 1995), the anterior and posterior attention systems (Posner & Dehaene, 1994) as well as the biased competition (Desimone & Duncan, 1995; Desimone, 1999) and integrated competition (Duncan, Humphreys, & Ward, 1997) frameworks. However, no clear cut model explains how several biases in different brain areas converge to let the system operate on the same event and how the temporal dynamics of attention evolves.

2.2. Top–down feature feedback

We suggest a model that explicitly incorporates top–down feature feedback (Fig. 3). Anatomically massive feedback projections from TE to TEO, V4, V2 and even as far as V1 have been identified (Rockland & van Hoesen, 1994; Rockland, Saleem, & Tanaka, 1994). It has been suggested that top–down feedback plays an important role in feature binding and attention (Lammé & Roelfsema, 2000), but its influence has only recently begun to be investigated by models of visual processing and attention (Hamker, 1999, 2000, 2002; Hamker & Worcester, 2002; Kirkland & Gerstein, 1999; Koechlin & Burnod, 1996; Roelfsema, Lammé, Spekreijse, & Bosch, 2002; Ullman, 1995; van der Velde & de Kamps, 2001). Ullman (1995) suggested that the brain imple-

ments largely independent bottom–up and top–down pathways. Besides the bottom–up transformation of the input, a top–down pathway transforms the prototype in parallel and both pathways try to match possible alternatives on all stages until one match becomes dominant. We also assume that the brain uses prior knowledge in form of such a prototype or target template in order to connect planning with the outside world (Fig. 4). Thus, the brain generates a fast bottom–up wave transformed by several fixed filters and meets the alternatives of a goal-related prototype in V4 and IT, whereas the target template is generated and held in prefrontal cortex.

The receptive fields of cells in IT comprise large areas of the total visual field and thus can hardly indicate the target location. However, feedback connections transfer any target in IT downwards to V4. This top–down expectation pathway emphasizes the current strongest features. It uses a divergent backprojection to all locations within the receptive field. Thus, V4 represents a spatially organized map that indicates potential target features by an increased sensitivity. As a result the late part of the IT responses get tuned to space.

2.3. The source of spatial reentry

Typically the dorsal pathway is associated with spatial attention. A task-relevant increase has been reported in several fronto-parietal areas processing locations, like LIP (Gottlieb, Kusunoki, & Goldberg, 1998; Snyder, Batista, & Andersen, 2000) and FEF (Bichot & Schall,
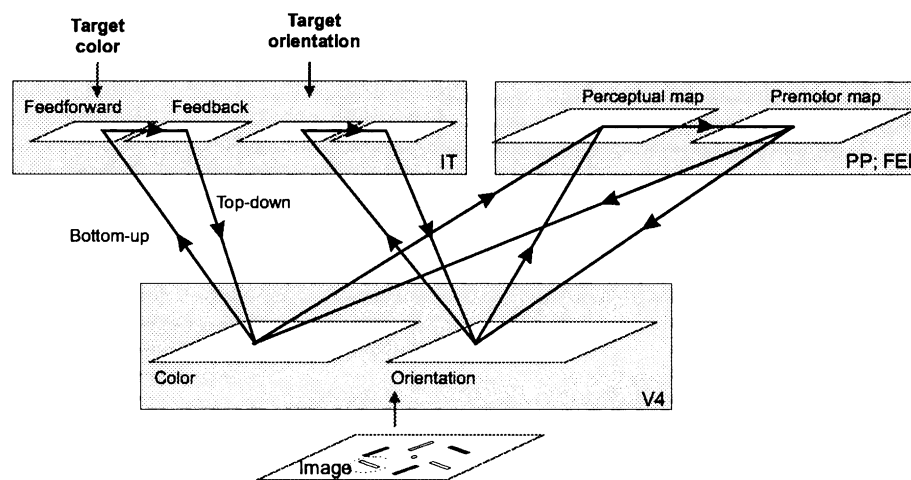

Fig. 3. The model consists of three interconnected functional blocks (V4, IT and PP/FEF). V4 cells encode features in different dimensions (here color and orientation) in parallel and project their activity to IT cells. IT cells have larger receptive fields and are modulated by the target template activated from prefrontal areas. The top–down pathway in IT cells shows a stronger competition among active populations to force one specific interpretation of the scene. Such feedback enhances the gain of populations encoding target features independent of their RF location in the display. With a short delay after stimulus onset, V4 cells encode task-relevant features. Perceptual maps (here one) in the fronto-parietal network (e.g. posterior parietal cortex and frontal eye field) combine the activity across dimensions to determine behaviorally relevant locations. Such locations feed a premotor map responsible for action preparation, which only encodes the most likely locations under strong competition. The premotor map feeds back to V4 and to IT (not shown). It enhances the gain of all cell populations within the movement field of the cells in the premotor map. Thus, after a delay the information about relevant stimuli reenters into V4. The whole processing is completely parallel; there is no sequential order in the blocks.
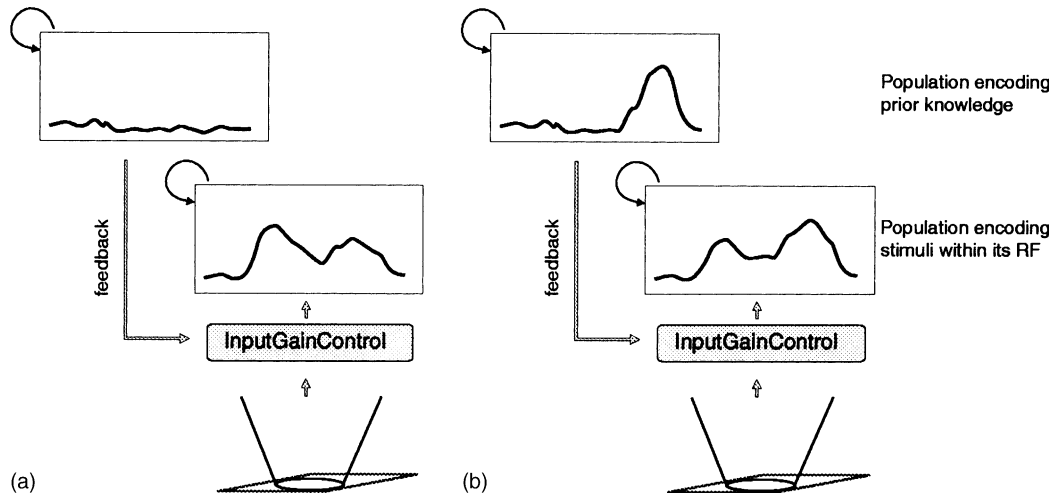
Fig. 4. Illustration how prior, top–down knowledge modulates processing (e.g. from PF to IT or from premotor maps to V4). A population of cells encodes a scene within its receptive field in parallel. The x-axis can represent feature space or spatial coordinates. The activity of a cell indicates that its preferred stimulus is present. Its firing rate is correlated with the stimulus salience and probability that the stimulus is in the scene. (a) Without a significant top–down influence the content is simply processed in a bottom–up manner. (b) Top–down, prior knowledge increases the gain of the cells and therefore emphasizes a specific pattern (or location). Due to receptive field interactions the population response encoding the non-supported stimuli (or location) decreases resulting in a dynamic attention effect.

1999). However, the major question is not which areas are modulated by attention but which areas are likely candidates for a spatially organized feedback signal—the source of spatial attention in the ventral pathway. According to our view, spatial attention is embedded in the visuo-motor system by competition and not by specific maps fully devoted to computing attention. We assume that spatial attention is generated in areas that serve for action selection and that the attentional feedback into the ventral stream is a premotor signal (Rizzolatti, Riggio, & Sheliga, 1994).

Visuo, visuo-movement and even movement cells in the FEF reflect stimulus-driven salience and goal-directed task relevance by their activation (Bichot & Schall, 1999; Bichot, Rao, & Schall, 2001; Burman & Segraves, 1994). Increased local activity in the ventral pathway enhances the visually responsive neurons in the frontal eye field and cells in the posterior–parietal cortex. These reflect the task-relevance of a location.

Areas responsible for action preparation use this activity as a guide for an appropriate action (Fig. 3). In our model we separate these areas into perceptual and premotor maps. The perceptual map signals the outcome of perceptual processing by a summation across all dimensions (e.g. "color" and "orientation"). Its cells show a strong phasic response. The premotor map is fed by the percpeptual map in an excitatory feedforward and surround inhibitory manner. It is decoupled from perception and signifies the locations used for planning an action. Its cells show a late, typically motor related response. Altogether, the perceptual and the premotor map constitute the equivalent of a "saliency" map in this model. However, we apply this term with caution since the perceptual and premotor map show fundamental differences to the saliency map used in other models.

According to our previous simulations, FEF cells with a strong phasic component, like visual and visuo-movement cells, are likely not the source of spatial attention (Hamker, 2002). We predict that spatial attention is tightly connected to premotor movement neurons in the FEF, since such a feedback signal allows a better target/distractor separation. Although we supported our prediction by a temporal analysis of the average cell activity, simulations cannot prove the spatial reentry from the frontal eye field movement cells. However, the strong overlap between eye movements and attention (Corbetta & Shulman, 1998; Deubel & Schneider, 1996; Hoffman & Subramian, 1995; Kowler, Anderson, Dosher, & Blaser, 1995; Kustov & Robinson, 1996; Perry & Zeki, 2000; Rizzolatti & Craighero, 1998; Rizzolatti, Riggio, Dascola, & Umiltá, 1987; Rizzolatti et al., 1994; Sheliga, Craighero, Riggio, & Rizzolatti, 1997) suggests that, if not the frontal eye field, other areas for planning a motor response are likely in the position to generate such a feedback signal. For the current study we therefore assume that the visual signal is transferred from a perceptual level to a premotor level without specifying the exact brain area.

2.4. The target of spatial reentry

The spatial focus of attention occurs from spatially organized reentrant processing. We deny an exclusive stage that separates between parallel and serial processing as suggested by other frameworks which emphasize the distributed character of visual processing

(Schneider, 1995). However, the spatially organized re-entry signal targets primarily areas of intermediate complexity.

The increasing activity of a premotor cell is sent into extrastriate visual areas and facilitates the processing within its movement field. It is known that neurons in V4 at the location of the intended eye movement increase their activity (Moore, 1999), whereas long-range competition (Desimone & Schein, 1987) could suppress others. Such a spatial reentry signal would allow a binding of features within different dimensions (Hamker, 1999; Tononi, Sporns, & Edelman, 1992) by linking bundles of features (compare the explanation by Wolfe and Cave (1999)) in different dimensions together. Thus, spatial reentry leads to a facilitated processing of certain stimuli, but does not fully suppress the activity of non-attended items. In our model this feedback modulates the activity in V4 and IT.

3. Results

Before we discuss the actual simulation results of the experiments we illustrate the temporal dynamics of the model and provide implementation details as far as they are helpful for the understanding of the results. A full mathematical description is found in the Appendix A.

3.1. Temporal dynamics and implementation

The analysis and description of attentional phenomena in this model is at the level of collective, dynamic activation variables. Generally speaking, such a population code is equivalent to the projection of the responses of cells into a functional parameter space. This means each item in a scene is represented by populations of active cells. For example a colored oriented bar is represented by a cell population in V4 encoding its orientation and by a cell population encoding its color (Fig. 3). It is known that some V4 cells are sensitive for color and orientation as well, but for simplicity we separate the representations into two different dimensions. Consistent with the experimental conditions (Bichot et al., 1999; Motter, 1994a), we assume that only one object is presented in each V4 receptive field.

In our model all visual features are computed in parallel, but follow competitive rules. At each location $x \in \{1, \ldots, 6\}$ and each dimension $d \in \{color, orientation\}$ in the simulation of the experiment of Motter (1994a, 1994b) and $d \in \{color, shape\}$ in the simulation of the experiment of Bichot et al. (1999) we simulate a neural population of cells as indicated by the feature-space axis in Fig. 5. The feature-space is simplified to a single dimension. Input stimuli are generated according to the experimental scenes and are encoded as V4 input populations determined by a

Gaussian distribution (Fig. 5). In the present version of the model we ignore stimulus-driven salience. Thus, the input activity of all stimuli is equal. For the sake of simplicity our present model also excludes lateral binding, which restricts comparisons with real data to experiments with visual displays containing simple, not densely covered arrangements of items. Long-range inhibition in V4 leads to a slight competition among active populations.

Let us assume the model is supposed to look for red items. This is implemented by generating a population of active PF cells representing a red target template (Fig. 5). Now we present the scene to the model by generating 6 input populations (3 representing red and 3 representing green). Green distractors and red targets are encoded by different "hills" on the feature axis (Fig. 5). We do not aim at modeling a realistic color space. The input activity travels up from V4 to IT. We model just one IT population, i.e. the receptive fields of our IT cells cover all V4 receptive fields. Once the activity from V4 enters IT, it meets the feedback signal from PF (the expected information).

Such expected information from the feedback pathways is incorporated into the population activity by continuously matching it with the input signal into the population (Fig. 4). This is implemented as a neural Bayesian inference operation, i.e. an input pattern is compared with the prior or expected information encoded in the network (Koechlin, Anton, & Burnod, 1999). An input pattern that matches a top–down pattern increases its activity.

In our example, the population in IT encoding the red items in the scene enhances its activity due to the feature feedback from PF and distractor populations in IT are suppressed (Fig. 5a). This is similar as proposed by the Biased Competition framework (Chelazzi et al., 1998; Desimone & Duncan, 1995). In extension to the Biased Competition framework we now assume that a feedback signal from IT travels down to V4 and again provides expected information for which the input into V4 is compared with. This feedback is distributed to all V4 cells within the receptive field of the IT population. Thus, it is feature specific but not location specific (at least within the RF of an IT cell). Again, input populations into V4 that match the expected information get enhanced. Thus, populations encoding red get enhanced and due to long range inhibition they suppress other populations (Fig. 5a). Without any other external influence the state of the network drifts into this fixed point solution where populations encoding the target feature are represented by a higher activity. We now assume that at $t = 0$ ms an exogenous cue indicates that the green stimuli are now relevant. We implement this by switching the PF population to represent green at $t = 150$ ms in order to account for the time the exogenous information is processed.
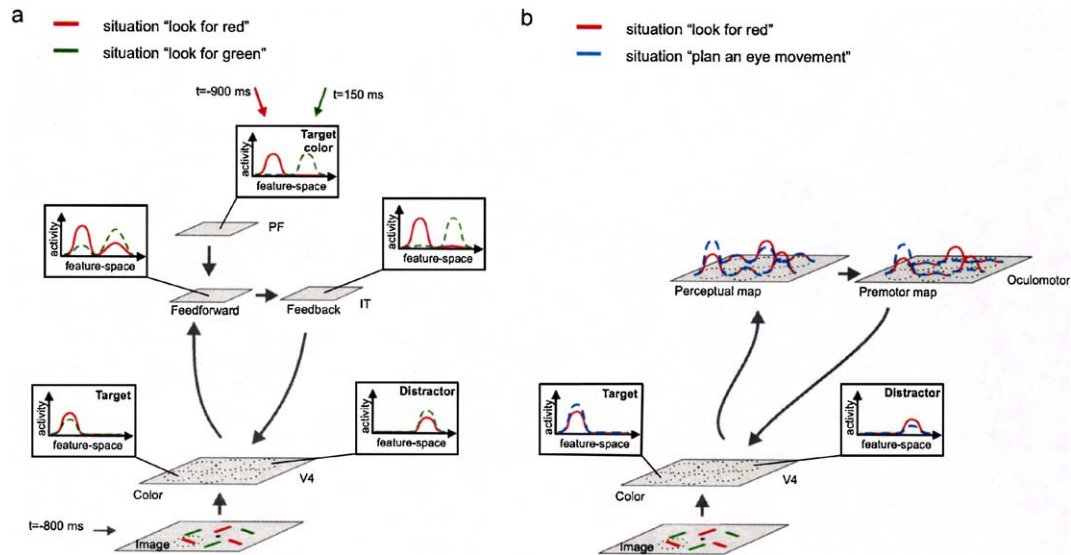
Fig. 5. Illustration of the coding and temporal dynamics of the proposed model. (a) Illustration of the pathway for "object recognition" in the model. We present artificial input stimuli to the V4 layer. The stimuli are generated to reflect the experimental condition, here illustrated as an image. At each V4 receptive field we model a population of cells which encode the feature space. Here the axis reflects an arbitrary color space. We only show one target and one distractor. PF and IT are each modeled with a single population. The first situation we describe is "looking for red items", indicated by the activity curves in red. At $t = -900$ ms we activate the target template in PF and present the inputs at $t = -800$ ms. The activity of the stimuli gets processed through V4 into IT. Due to the gain control by feedback the population encoding the target color gets enhanced and the one for the distractor is suppressed. Feedback from IT to V4 is based on the same mechanism. The networks settles in a state where all items of the target color in V4 are represented by a higher firing rate than those with a non-matching color. The second situation is "looking for green items", indicated by the activity curves in green. The inputs remain active. By changing the target template at $t = 150$ ms the model now switches into a state where again all items of the target color in V4 are represented by a higher firing rate than those with a non-matching color, so that the former distractor is now the target. Thus, feedback in the "object pathway" operates feature specific and location unspecific. (b) Oculomotor pathway in the model. The initial situation is "looking for red items". The perceptual map follows the activity distribution in V4 and encodes all locations of items that match the target by a higher activity. If an eye movement is planned, initially the target locations build up activity in the premotor map, which in turn reenters V4. This reentry signal enhances the gain of all encoded features, regardless of their dimension. However, this is not a stable state. Due to strong competition in the premotor map the network switches into a state where just one population in the premotor map is active, which then enforces just a single stimulus. Thus, the activity of all features at the target location get enhanced and due to long range inhibition in V4 other stimuli are encoded by less activity. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

The computationally challenging task is to enable the model to switch its internal representation. Our gain control mechanism supports rapid switches, because feedback acts on the excitatory input of a cell and not on its output activity. Due to the switch in the PF activity, the population in IT encoding red looses its feedback signal whereas the one for green receives support. As a result the prioritized encoding in IT changes and the whole system switches to a state where populations encoding the new target feature are represented by a higher activity. This is essentially the explanation our model gives for the experiment done by Motter (1994a, 1994b).

In extension to this experiment we now assume that the model should report the orientation of any target item. Let us continue from the state prior to the switch in PF and assume the red stimuli are still the targets. As a result of the feature specific feedback, perceptual cells in the oculomotor system show an enhanced activity at locations containing a potential target (Fig. 5b). A response towards a specific feature of a stimulus that is not the target feature is possible at this point by reading out the IT activity, but if the stimulus does not pop-out the target discrimination is typically too low for a reliable response, similar to Hoffmans hypothesis (Hoffman, 1979). So let us assume the target orientation is not reported at this time. The scene contains three potential targets and the model has to decide which orientation it reports. As a result of the required response it starts to plan an eye movement towards one of the targets, without necessarily executing it. The required type of response is encoded in the model prior to the experiment: A strong activity of the fixation cell prevents the premotor cells from firing at all. An intermediate activity only prevents the cells from reaching threshold activity that elicits eye movements. A full suppression of the inhibitory fixation cell disinhibits the premotor cells and allows them to reach threshold activity.

All cells with a potential target in their movement field raise their activity due to the target/distractor discrimination in the perceptual map. A spatial focus emerges from immediate reentry of the premotor activity into V4 and IT. Similar to the feedback from IT, but now specific in space and unspecific in feature and dimension, reentry from the premotor cells enhances the gain of V4

cells. Initially three weak foci emerge at locations of strong activity of perceptual cells due to the slight increased firing of the premotor cells. Strong competition among premotor cells eventually suppresses all but one cell (or in general one population). Thus, the activity of V4 populations encoding other features of the target item than just the redness get enhanced as well and the report of the orientation is facilitated. Again, if no other event occurs (e.g. eye movement, inhibition of return) the network drifts into a state where just one of the potential targets receives reentry activity (Fig. 5b). This hypothetical experiment is similar to the experiment of Bichot et al. (1999), but here we had three possible targets and no primary task.

In sum, our model suggests that shifts of attention are equal to the modification of the internal state of a dynamical system and feedback determines the state into which the system switches. Attention itself is not explicitly implemented but emerges from the mechanisms of feedback and competition within the visuo-motor network. We now show the simulation results on the two experiments and discuss the variability of the reentry component by means of an artificial visual search experiment.

3.2. Experiment 1

We first used the explained model to simulate the experiment done by Motter (1994a, 1994b) in order to demonstrate how the feedback from a color target influences the activity of V4 cells. As explained a match of the pattern in working memory with the one that enters IT leads to an increase of activity in IT cells encoding the target. Such an advantage in encoding the target is transferred further back to V4. The simulation results (Fig. 6), specifically the sum over the cells in the feedforward and feedback pathway (Fig. 6c), closely resemble the average temporal development of the activation in the match (M) and non-match (NM) conditions during the stimulus presentation phase (Fig. 1b).
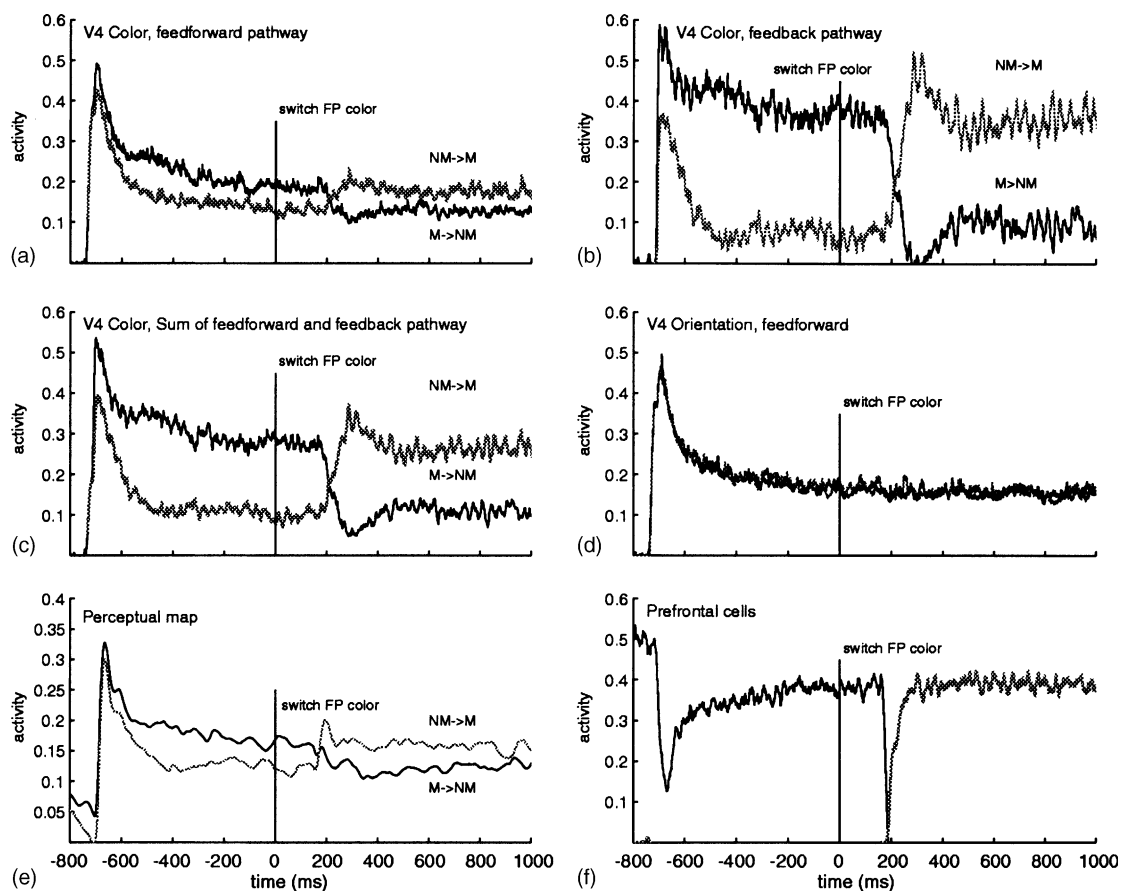

Fig. 6. Simulation results of the experiment of Motter (1994a, 1994b) with the FP color switch. Model V4 populations encoding the target color receive a top–down feature bias. As a result these populations increase their activity (see (a)–(c)). (a) Feedforward pathway in the color dimension. (b) Feedback pathway in the color dimension. (c) Average of the cells in the feedforward and feedback pathway. This data resembles best the electrophysiological data. (d) The orientation selective cells do not reflect any feedback bias, since during the stimulus presentation activity in the premotor cells is suppressed. (e) The cells of the perceptual map reflect the relevance of locations, whereas the activity of cells in the premotor map is at baseline and therefore not shown. (f) Activity of prefrontal cells encoding the target knowledge. Instead of changing the color of the cue, we switch this target template after 150 ms to account for the time the cue information must be processed.

Since the single cell data quantitatively strongly depends on the cells observed, we do not aim to achieve a quantitative match with the data. Three observations are relevant at this point: (i) As indicated in the data the model V4 cells should show a higher activity if the feature matches the target. (ii) A change of the target template should be followed by a switch of the activity in our model V4. (iii) The time needed for a switch should be appropriate. By assuming (physiologically plausible) 150 ms for the time the cue is processed before a change in prefrontal areas can be initiated the time of switch at around 200 ms in the simulation is consistent with the data.

Note that in this phase the monkey in Motters experiment has to observe the fixation point and no eye-movement is required. The knowledge that items which match the color of the fixation spot are potentially important is reflected in our model as a feedback signal from prefrontal areas (Fig. 6f) to IT and further back to V4. Such an expectation signal meets the stimulus-driven signal and increases the cells sensitivity. When the stimulus signal and the expected signal coincide we observe an enhancement of the cells response. Given the target is red, V4 cells sensitive for this color enhance their activity regardless of their location in the visual field (match condition). Since these cells feed the oculomotor circuit, the cells in the perceptual map follow the perception and represent the location of these potential target items through enhanced activity (Fig. 6e). Since fixation is the only task requirement during the presentation phase we expect that premotor cells in the oculomotor circuit, cells from which we expect the bulk of spatial reentry, are not activated. Attention within this phase is purely feature-driven.

After the cue switches, which we simulate by a change in the expected signal in prefrontal areas (Fig. 6f), the system responds to the new situation by shifting its internal dynamic state. The former distractors are now represented as potential targets and vice versa.

What does the model predict that goes beyond the findings of Motter (1994a, 1994b)? In his experiment he did not report any effect on orientation selective cells in V4. According to our hypothesis, knowing the effects on the orientation selective cells is crucial for understanding the difference of feature-based and spatial attention. If the features in different dimensions of an item group together as an entity we would expect an effect of feature-based attention on orientation selective cells as well. As soon as feedback enhances color sensitive V4 cells, orientation selective V4 cells should raise their activity as well. However, an immediate attention effect on the cells sensitive for orientation by strong grouping would indicate an enhancement of the sensory quality of that item and not only a prioritized processing. It seems that this is not consistent with the findings of Moore and Egeth (1998).

We assume no strong immediate cross-dimensional grouping in our model. V4 cells that are sensitive for orientation only show a target–distractor discrimination through spatial reentry. Due to the fact that the monkey in the experiment has to observe the fixation spot we assume that no eye-movement is planned. Thus, orientation sensitive cells should not reflect a selective enhancement (Fig. 6d).

3.3. Experiment 2

Regarding the aspect of cueing by feature, the experimental condition of Bichot et al. (1999) is very similar to the previous experiment. Here the red shapes are behaviorally relevant and have to be processed to decide if their shapes are same or different. We generate artificial inputs for the feature cue ("color") of the shapes and the probes ("letters"). We only model the "color" feature of the shape, not the shape itself. The model does not compute an explicit response (e.g. same or different for the primary task and a list of letters for the secondary task). This would require a much more complex model that reads out the activity in IT. In addition we would have to make additional assumptions about the generation of a response, which is beyond the scope of explaining how feature cues guide spatial attention. Thus, in the simulation of the experiment of Bichot et al. (1999) we focus on the dynamics that will influence a response, but we leave out the response generation. Studies in neuroscience have revealed a close relationship between neural activity and behavioral response (Parker & Newsome, 1998; Shadlen, Britten, Newsome, & Movshon, 1996). The activity of a small number of cells or in some cases a single neuron is sufficient to reliably discriminate a stimulus and to predict the behavioral response. In order to compare our simulations qualitatively with the experimental data, we therefore use the activity of the cells as a measure of the likelihood for a response. Since the target shapes typically have a different color than their surrounding items a stimulus-driven salience effect is possible. It was observed that voluntary attention overrides stimulus-driven effects by time (Müller & Findlay, 1988). Thus, we focus our analysis on SOA's longer than 100 ms.

The top–down target template enhances the activity of the populations in IT that match the target template ("red") (Fig. 7). Although the subjects are instructed not to make eye movements we assume that the primary recognition task automatically recruits oculomotor circuits. As a result of processing in the premotor map, spatial reentry occurs and enhances populations encoding the "letters" within the "shapes" (Fig. 7). In the experiment the letter-report task is used to probe the amount of attention at each location. Obviously, the feature cues that do not match the target (the behaviorally
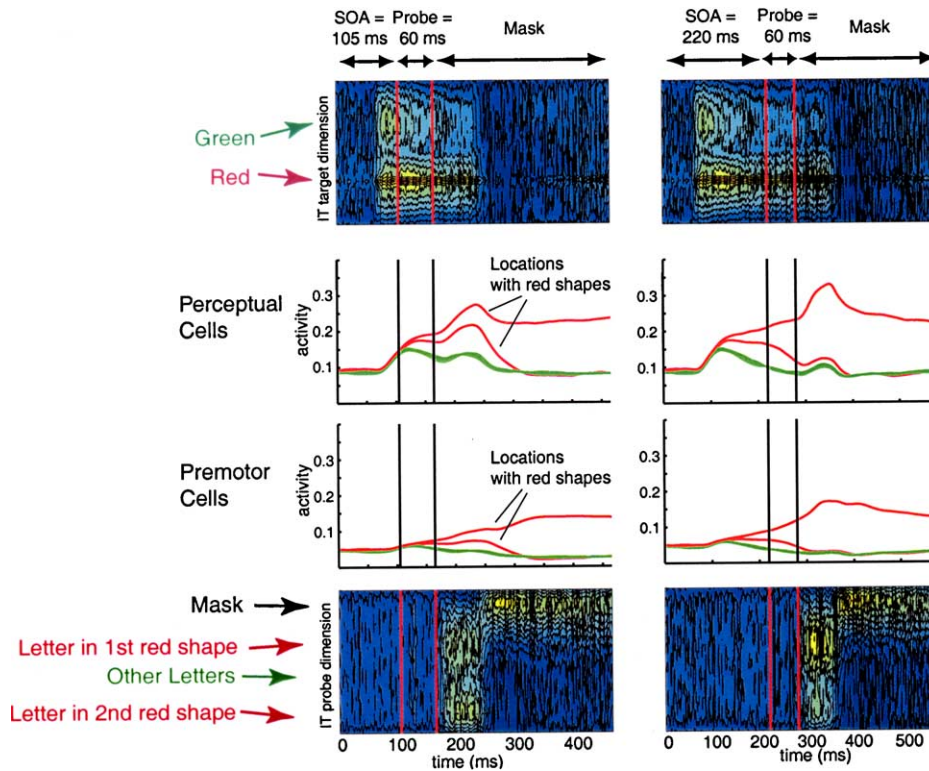
Fig. 7. The course of neural activity in the simulation of the experiment of Bichot et al. (1999) from stimulus onset until the end of mask presentation with two different SOA's. As explained in Fig. 5 stimuli are encoded in the cue and probe dimension (top and bottom of figure) by a population of cells in one feature axis. Please note that the activity encoding the stimuli needs time to reach these higher areas. At the top of the figure we show an IT cell population sensitive for color that has a receptive field which comprises the whole display. Initially, populations encoding the color of the target shapes (here red) are enhanced. As shown in Fig. 3 the advantage of an increased firing rate of cells encoding "red" is then transferred down to cells in V4 which have receptive fields that comprise only one stimulus. Therefore, perceptual cells (second graph) in areas responsible for oculomotor control indicate such behaviorally relevant locations by an enhanced activity. Later premotor cells in these areas (third graph), however, start to compete for the dominant location, but do not raise their activity to levels that initiate an eye movement. Reentry from those premotor cells enhances all populations within the movement field in V4 and IT. Thus, populations encoding letters (bottom of figure) within the behaviorally relevant shapes are enhanced according to the distribution of activity in the premotor cells. The mask is used to suppress activity in the probe dimension. Using a SOA of 220 ms the competition among the premotor cells is more settled towards one location at the time the letters appear, as compared to an SOA of 105 ms. Thus, a larger SOA increases the amount of spatial attention at one location. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

unimportant shapes) capture no spatial attention, since they are not processed in the premotor map.

In the following we restrict our analysis to the amount of attention within the two behaviorally relevant shapes. Our probe stimuli ("letters") at locations of the features that match the target (the ones within behaviorally relevant shapes) benefit from reentry and are thus more actively represented, which is in line with the experimental finding of an enhanced probability of report. Bichot et al. (1999) report about 40% correct for the probe task using an unique SOA of 105 ms. In a conditional analysis they find that when one letter within a behaviorally relevant location was correctly reported, the other one was reported correctly in about 35% of trials (slightly depending on the distance between both relevant shapes). They did not try to compute the probability of report of each of the probes independently. Thus, we do not know the balance of the individual probabilities. Our simulation results suggest that

they are similar with an SOA of 105 ms, which is consistent with the interpretation "that both target locations can be attended simultaneously, without distractor locations between them being selected" (Bichot et al., 1999).

What does the model predict in this experiment? We now compare the simulation with an SOA of 105 ms with one where the SOA is 220 ms. Our simulation results (Fig. 7) indicate that in this case the competition in the premotor map has almost settled towards one location at the time the probes appear in V4. Thus, the probe ("letter") at one location gets sensitized by a high level of reentry activity whereas the other one does not.

If we now compare the activity of those cells encoding the probe stimuli ("letters") across all simulated SOA's (Fig. 8), we observe an increase of V4 activity at one location and a slight decrease at the other location. With an SOA of 105 ms the peak activity of V4 cells is almost the same, which explains the result of Bichot et al.
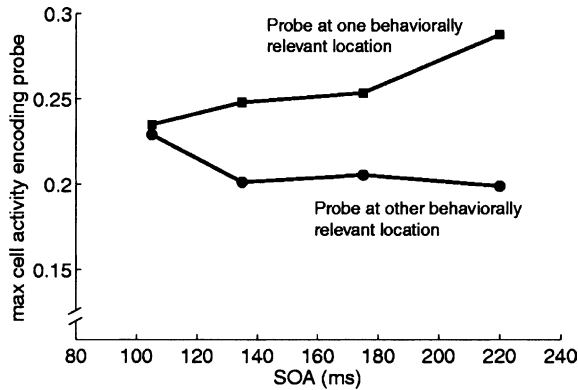
Fig. 8. The neural peak activity of cells encoding the letters within the behaviorally relevant shapes over the SOA's. An SOA of 105 ms results in an equally strong activation of the populations. Longer SOA's benefit the activity of one probe population as against the other.

(1999). With longer SOA's the model predicts an enhanced spatial reentry at one location and a less strong at the other location. Due to long range competition among the populations encoding the probe, however, the activity for the probe at one "less selected" location decreases.

We therefore suggest that the observed split of spatial attention (Bichot et al., 1999) could be a transient effect, at least in such a cuing task. At a SOA of 105 ms the amount of spatial attention is approximately the same at both locations cued by the redness. With an increase in SOA spatial reentry enhances the V4 activity at one location while it is slightly decreasing at the other one. We predict that the probability of reporting both letters within the behaviorally relevant shapes will decrease with increasing SOA, even though the subjects have more time to finish the primary task. Note that this prediction was recently confirmed experimentally (Hamker & VanRullen, 2002). In a conjunction search paradigm (Kim & Cave, 1995) reported no significant differences in the percentage of correct responses to letters appearing in shapes across SOAs of 75, 105 and 135 ms. Two reasons might be responsible for the difference to our simulations. First of all, our simulations indicate a general trend. We do not predict that at exactly 135 ms we should see that the split of attention starts to diminish—it could well be after 135 ms. Second, they presented the search array for only 60 ms regardless of the SOA of the probes. Thus, the different SOAs do not reflect relative changes in the selection process, but in the decay process after the buildup within the first 60 ms.

3.4. Visual search experiment

Our reentry hypothesis does also provide implications for the distinction of parallel and serial search. In a visual search experiment, many items are presented simultaneously and the subject must find a prespecified target. An increasing number of items can increase the reaction time or the detection rate. The search slope refers to the slope of a linear fit of the reaction time over set size. Search is typically assumed to be serial for steep slopes. Parallel search is typically assumed to show no set size effect, because the time for selection is regarded to be constant. However, search slopes show a high variability such that knowing the slope does hardly allow to predict the underlying search mode. Duncan and Humphreys (1989) rejected the dichotomy of parallel and serial search and introduced a continuum of search efficiency. They assume that visual processing is always resource free and parallel, but access to visual short term memory is strictly limited. Major factors that influence competition for visual short term memory are the target–distractor and distractor–distractor similarity.

We here show the response of our model when we increase the target–distractor similarity in a visual search experiment containing six stimuli. We use just one set size and do not compute search slopes. At this point we are interested in demonstrating the implicit mechanisms of the model that determine search. Note that, since we have not implemented any grouping effects in our model, varying the distractor–distractor similarity would be meaningless. We simulate a conjunction and a feature visual search task. In feature search the target feature is not shared by any other item in the display. In conjunction search a target is solely defined in conjunction
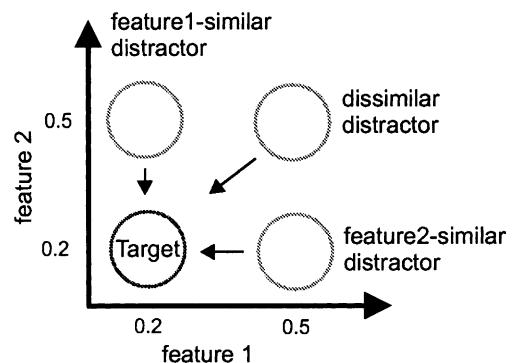

Fig. 9. Definition of the target and distractors in the simulation of a visual search task with increasing target–distractor similarity. In each dimension the feature space consists of a single axis. We construct a target item with input populations, defined by a Gaussian, around the artificially chosen value 0.2. The circle illustrates the width of the Gaussian (σ-value). The feature1-similar distractor activates the same neural population as the target in the first dimension and the feature2-similar distractor activates the same population as the target the second dimension. The target dissimilar features are initially chosen to be of the value 0.5. We increase the similarity of target and distractor with increasing stimuli sets in steps of 0.05 on the feature axis. In conjunction search the used six-item display contains a target, a dissimilar, and two feature1-similar and two feature2-similar distractors. In feature search the input contains a target and five dissimilar distractors. The target is stored as a top–down template.

with features. Conjunction search was often associated with serial search and feature serach with parallel search (Treisman & Gelade, 1980). We use eight stimulus sets in which the similarity of target and distractor increases (Fig. 9). In order to have an exact measurement of when an item is selected we now allow the model to reach threshold activity in the premotor cells (overt search), which indicates the time of internal target selection. After the first selection we abort the search.

In our analysis we distinguish between correct trials and incorrect trails. Correct trials contain only those trials in which the first selected item was the target. Incorrect trails are all other trials. Thus, correct trials are the result of our model's parallel selection process. In incorrect search, because the first selected item is not the target, a serial component is needed to terminate the search. With increasing similarity of target and distractor our model shows two effects for both, feature and conjunction search: an increase of the time to select the target in the correct trials and an increase in the number of incorrect trials (Fig. 10). The increase of the time to

select the target in the correct trials reflects the variability of the parallel component in our model. The number of incorrect trials reflects the influence of an additional serial component. The reason of this observation is that a high similarity of target and distractor reduces the efficiency of feedback in the models "what pathway", since with increasing similarity the distractors also get more sensitized. Since V4 projects to the perceptual cells the discrimination in the perceptual cells (e.g. FEF visual cells) is also poor. As a result the premotor cells (e.g. FEF movement cells) need more time to resolve the competition. Thus, the reentry component is delayed and any decision will likely to be delayed, too.

The increase of the time to select the target in the correct trials suggests that our model can produce a variety of search slopes with parallel search. At this point we do not aim to specify the possible range of slopes which can be produced by the model in parallel search mode. It is also beyond the scope of the present paper to suggest physiologically plausible mechanisms
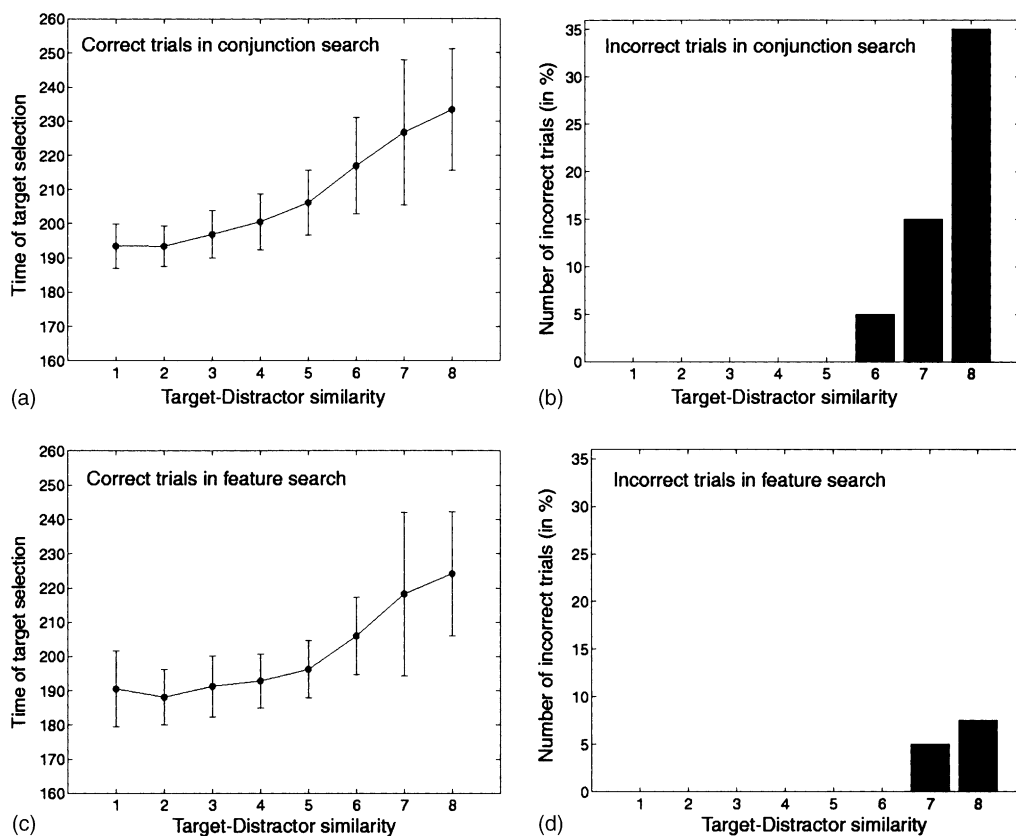

Fig. 10. Results of the visual search task (six stimuli) with increasing target–distractor similarity from input set 1–8. For each stimulus set we perform 40 simulations with different noise in the system. Correct trials contain only those trials in which the first selected item was the target. Incorrect trails are all other trials. (a): Conjunction search: Average time for target selection in correct trials. The bars show the standard deviation of the distribution across all trials. (b): Conjunction search: Percentage of error trials. (c): Feature search: Average time for target selection in correct trials. (d): Feature search: Percentage of error trials. Feature search shows slightly less time for target selection in correct trials but a similar increase for higher target–distractor similarity than conjunction search. The primary difference between both search modes is in the number of error trails. Although feature search can show error trials, conjunction search leads to much more incorrect trials with increasing the target–distractor similarity.

how attention shifts to another item, but an inhibition of return mechanism that inhibits the perceptual cells can lead to such shifts (Hamker & Worcester, 2002).

In addition to a delay in parallel search the error rate of the reentry component increases (Fig. 10). A poor discrimination does not only need more time to resolve competition, it also favors the selection of a distractor location. Feature search shows a similar increase of target selection in the parallel component, but much less error rates than in conjunction search (Fig. 10). This is presumably the reason for the steeper search slopes in conjunction search as compared to feature search. In extreme cases, feature search is not necessarily parallel. Similar to the proposal of Duncan and Humphreys (1989) our model shows a continuum of search efficiency depending on target–distractor similarity. However, unlike Duncan and Humphreys (1989), our model predicts that search can indeed be serial. Serial search in turn is neither defined by the task (e.g. feature or conjunction search) nor explicitly implemented in the model, it is defined by implicit factors that determine the discrimination of the target from the distractors.

4. Discussion

Many models of visual attention assume that attention is determined by a specific spatially organized map. If this map is properly defined this approach can be used to explain a number of experimental results. However, except for pure bottom–up approaches (e.g. Itti & Koch, 2000), explanations of how the information of where to attend gets into this map are kept simple. Knowledge about a task could lead to an inhibition of certain feature maps (Treisman & Sato, 1990) or an enhancement within "top–down feature maps" which are later combined to activation maps together with the bottom–up feature maps (Wolfe et al., 1989). The latter Guided Search model (Wolfe, 1994) was specifically designed to explain various search slopes through a serial selection of locations on an activation map. Its activation map is computed by applying a set of rules on the stimulus configuration and on the target–distractor relation. As an extension to Guided Search, others have suggested an activation map within each feature map (Cave, 1999; Deco & Zihl, 2000). However, these approaches leave the construction of the top–down feature maps undefined. An explanation of the processes leading to a cognitive penetrability of attention maps is lacking.

We propose a refinement that gives insight into how attention emerges from the knowledge about the task at hand. Our simulations show that a reasonable explanation is a feature-based top–down modulation within the same pathway. A behavioral anticipation sets a top–down activation which changes the response of feature-coding cells. Thus, a goal represented in higher brain areas modifies perception at early stages and increases the sensitivity of the pattern within a resonance loop. A top–down component was also proposed by Tsotsos et al. (1995). However, they implemented a WTA on each level, e.g. IT, in order to compute a spatially focused feedback (spotlight) within the ventral stream. Our feature-based top–down modulation operates in parallel, independent from any focus of attention. For example, in the experiment done by Motter (1994a, 1994b) we predict that the cells encoding the orientations of the red items do not gain an advantage from this feature feedback.

Referring to the introductory question of Chelazzi (1999) about the role of the findings of Motter (1994a, 1994b) in visual search, we propose that prior knowledge from high level areas is transferred into visual areas of intermediate hierarchy in order to prepare a dynamic switch of the state of the visual filters. There is no need to assume that the ventral pathway is only the target of spatial selection (Kim & Cave, 2001). According to our view, the ventral pathway also provides the source for spatial selection. Feedback within the ventral pathway can only facilitate the perception of the searched feature. If the feature to be reported is different from that to be searched, a reentry is typically necessary to facilitate the perception of the feature to be reported. As illustrated in the simulation of experiment 2 this reentry is continuous and thus explains a split of attention as a transient phenomenon, at least in such a cuing experiment.

Our model shows some similarities to the hypothesis of Caputo and Guerra (1998). They found that knowledge of the target shape reduces threshold detection with an additional distance-effect of a salient distractor on the target and suggested two components in selective attention. The first compares the candidate object with objects in its neighborhood and the second performs a parallel suppression on the objects having non-selected features (Caputo & Guerra, 1998). Since the shape information has to be projected downwards by divergent top–down pathways our model is consistent with their first stage. Reentry and long range competition in V4 are then responsible for distractor suppression.

The Biased Competition framework (Desimone & Duncan, 1995) became popular as an neurophysiologically motivated approach of explaining attention. Our model is directly influenced by the Biased Competition framework in that the competitive interactions in the model V4 and IT areas can be biased by spatial or feature feedback. Our approach even allows to simulate the time course of IT cell activity (Hamker, 2002) of experiments within the Biased Competition framework (e.g. Chelazzi et al., 1998). However, Biased Competition addresses the question of how an external signal, like spatial attention, modifies the processing through competitive interactions. It has not aimed to explain how the biasing signal (e.g. spatial attention) emerges. We go one step further and try to explain how the

spatial attention signal itself emerges on the systems level by means of competitive interactions.

We now discuss some debated issues of attention in more detail by comparing our approach with others. This should not be understood as a critic of the valuable contribution of earlier models: we simply pinpoint our suggested refinements.

4.1. Feature-based vs. spatial attention

Feature-based attention was suggested to result in a prioritized processing of items but does not enhance their sensory quality (Moore & Egeth, 1998; Shih & Sperling, 1996). Moore and Egeth (1998) based this claim on the finding that feature-based attention does not improve the performance under data limited conditions. Since an enhancement of the quality of a stimulus (e.g. increasing the brightness) improves the performance under such data-limited conditions, they conclude that a feature-based attention mechanism operates by prioritizing the item's processing, but does not enhance the sensory quality of that item possessing the feature of interest. How might this finding relate to the parallel enhancement of V4 cells observed by Motter (1994a, 1994b)? Furthermore, why does feature-based attention only allow a prioritized processing, i.e., in what respect is it different from spatial attention?

A feedforward model of spatial attention has difficulties in explaining the enhanced firing rates at all locations in V4 encoding red items (Motter, 1994a, 1994b). A selection of all potentially relevant locations must be assumed. However, this would contradict with the finding of Moore and Egeth (1998), since a selection in these models results in a sensory enhancement. Thus, it is more likely that the enhanced firing of cells has nothing to do with spatial selection, but with its guidance. Such guidance mechanisms have not been fully explained so far. The measured V4 activity cannot simply represent the top–down feature map in the Guided Search model (Wolfe, 1994) since the V4 activity is related to the stimulus.

Our model was not designed to account for the experiment of Moore and Egeth (1998), but we can give a reasonable interpretation why feature-based attention does not improve the performance under data limited conditions. Prioritized processing of feature-based attention (Moore & Egeth, 1998) is explained by our model on the basis of feedback connections within the ventral pathway which only enhance the features of the target template but not the whole item. This is important to recognize. Such an enhancement allows only an improved response towards the feature that is searched. In the experiment of Moore and Egeth (1998), however, the report of another feature than the searched one is required. For example, information about target color (e.g. blue or green) was given as a probability ratio and

subjects had to report a digit within blue and green letters. Since feature-feedback does not enhance the responses of the cells encoding the non-searched but to be reported feature (e.g. digit or letter) a subject has no immediate advantage of such feedback. Only spatial reentry enhances all populations encoding an item. As we have explained this reentry is guided by the activity within the ventral pathway. Since this spatial reentry is slow our model predicts no performance increase under data-limited conditions. Thus, the pysiological basis of the effect of prioritized processing by feature-based attention could be the feature-selective feedback followed by a spatially selective reentry.

4.2. Preattentive vs. attentive processing

Many models of attention assume a stage, which separates parallel from serial processing. Preattentive processing is assumed to be parallel and early while attentive processing is serial and occurs at higher levels after selection (Koch & Ullman, 1985; Theeuwes, 1993; Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe et al., 1989). For example, referring to the 'spotlight of attention' (Posner, Snyder, & Davidson, 1980), Koch and Ullman (1985) "postulate a 'switch' that routes the properties of a single location [...] into the central representation". Thus, those models separate preattentive from attentive vision in space. Other models assume that attention operates in a hierarchy of stages so that a distinction between preattentive and attentive processing is less distinct (Cave, 1999; Tsotsos et al., 1995). Lammé and Roelfsema (2000) suggest that preattentive processing is more likely reflecting a feedforward processing mode, and attentive processing occurs due to feedback.

In our model a feedforward signal can penetrate higher areas without being gated. Unlike the a model of Mozer and Sitton (1998) who assume an initial gating of several stimuli by spatial attention, which then is reduced to just one area, we assume an enhancement of targets by a reentrant signal. It is not clear which areas provide such a signal but there is growing evidence for the FEF. FEF primarily targets areas of intermediate complexity such as V4 and TEO (Schall, Morel, King, & Bullier, 1995). If there is no strong spatial selection by the Pulvinar, attention effects in V1 (Lammé & Roelfsema, 2000) might result from feature-feedback similar as we have described from IT to V4. This would explain the late occurrence of V1 effects (Martinez et al., 2001) and the poor spatial resolution of attention (Intriligator & Cavanagh, 2001). Our model is consistent with the idea that later stages strongly modify the patterns in a top–down fashion to suit the current interest of the viewer. Since attention emerges through the distributed interactions there is no explicit separation between preattentive and attentive processing in our model. In

this regard our model is best described by a space and time gradient from preattentive processing to attentive processing. The earlier the area the less attentive and the earlier in time the less attentive.

4.3. Parallel vs. serial search

Limited capacity models like Feature Integration Theory (Treisman & Gelade, 1980; Treisman & Sato, 1990) and Guided Search (Wolfe, 1994; Wolfe et al., 1989) assume a temporally serial mechanism that highlights one area after the other. Guided Search characterizes attention as a sequential, noisy selection process on the activation map. The time for selection is considered roughly constant. As an inherent result, non-flat search slopes are necessarily based on serial search. If such serial models assume in addition that only one item at a time can be processed they typically predict fast shifts of spatial attention in the range of 10–20 ms per item. If more than one item can be processed in parallel such models would predict approximately 30–40 ms per item. There has been an ongoing debate if search is based on a parallel mechanism or a serial scan. Advocates of the parallel hypothesis explain set-size effects in terms of a parallel, noisy decision process (Eckstein, Thomas, Palmer, & Shimozaki, 2000; Palmer, Ames, & Lindsey, 1993), or as a slow competitive mechanism for memory (Duncan, 1984), or as a slow distributed selection (Deco, Pollatos, & Zihl, 2002). Models assuming a noisy decision process are typically based on Signal Detection Theory and have been applied only in detection experiments using brief presentations that do not allow eye movements. It remains to be shown how accuracy measures map onto reaction times. Hybrid models suggest a parallel capacity limited 'one-view search' and an additional slow spatial shift of attention (Bundesen, 1990, 1999).

The search slope in experiments has turned out not to reveal a lot about the underlying mechanisms. Search slopes have been shown to depend on various low level factors. An analysis of roughly one Million trials of search shows an unimodal distribution of search slopes (Wolfe, 1999) which suggests that the slope does hardly indicate whether a search is parallel or serial. Thus, replicating search slopes is a big challenge that requires more knowledge about how items compete and group together. Many models have been shown to replicate some search slopes, but we still lack a fundamental understanding of the involved mechanisms. However, feature search shows on average less steep slopes than conjunction search (Wolfe, 1999). A model of attention should at least provide an explanation for such a difference.

It was suggested that the amount of guidance a preattentive process can provide, explains the slope (Wolfe et al., 1989; Wolfe, 1994). Thus, the amount of guidance

reduces the number of items to a subspace of items, which is similar to the effect of feature-specific feedback within the "ventral pathway" in our model, but search itself is explained on a fast serial scan. Our model predicts that feature search is not necessarily parallel and conjunction search not necessarily serial. In addition, our model shows a parallel selection that varies in time, which indicates that non-flat search slopes might not necessarily be the result of a serial mechanism. We illustrated this effect by increasing the target–distractor similarity in a visual search task. A serial search occurs only when feedback in the ventral pathway does not sufficiently enhance the relevant populations in order to allow the reentry process to focus on the correct item. These errors are typically higher in conjunction search, since distractors are more similar to the target. Our model does therefore not deny shifts of spatial attention in visual search, but it predicts slow shifts which is consistent with the observed change after 200 ms in the N2pc component of the event-related potential waveform during visual search (Woodman & Luck, 1999) and with late detection rates about 300 ms in a rapid visual presentation task (Weichselgartner & Sperling, 1987).

The effect of the degree of similarity between items in the display on attention (Duncan & Humphreys, 1992) would therefore be explained by both a parallel and serial mechanism. A high target–distractor similarity results in less efficient feedback operating in the ventral pathway and thus in a slow and error prone reentry process. We have not modeled effects of high distractor–distractor similarity which might result in a high stimulus-driven salience of the target, which reduces interference as explained.

Our model suggests a distinction between two conditions. Brief stimulus presentations (i.e., less than 100 ms) followed by a mask, do not allow the emergence of an effective spatial reentry component. In this condition our model is similar to the parallel decision process suggested by Signal Detection Theory in that set-size effects occur due to the increase in the probability of making an incorrect choice with the increase in the number of elements to choose from (Palmer et al., 1993). With longer stimulus presentation times the reentry signal has an increasing influence on the decision. The target–distractor discrimination determines the time the reentry signal gets significant and if the reentry signal gets locked onto a target. A correct reentry signal does explain variation in search times by a parallel mechanism. A series of wrong reentry signals would be responsible for steep search slopes. Realistic visual search experiments require the model to be equipped with more complex feature detectors as well as their interactions in order to find out the exact conditions that determine the discrimination of the target, the number of items that can be processed at once, and ultimately the distribution

of search slopes. This model is promising to be extended to visual search tasks since it predicts a continuous range of search slopes depending on implicit factors such as the specific arrangement of items as well as the target knowledge.

4.4. Spatial attention and saccades

There has been a long debate whether the system for attention overlaps with the one for programming saccades. Recent results indicate that directing a saccade to one location is influenced by attempting to attend to another (Corbetta & Shulman, 1998; Deubel & Schneider, 1996; Hoffman & Subramian, 1995; Kowler et al., 1995; Kustov & Robinson, 1996; Perry & Zeki, 2000; Rizzolatti et al., 1987, 1994; Rizzolatti & Craighero, 1998; Sheliga et al., 1997). This data can be interpreted by two overlapping systems, an attention and a saccade system or by an attention system that originates from a preparation of an eye movement (Rizzolatti et al., 1994). In terms of the assumption of two overlapping systems, spatial attention is often suggested to precede a saccade and perhaps select the endpoint. This could be useful to decide whether a particular object is a suitable target for a saccade. A tight coordination of both systems would be very beneficial in such a case. Kowler et al. (1995) suggested two possible models to explain the mutual interference of spatial attention and eye movements, one in which attention is devoted simultaneously to perceptual and saccadic targets (spatial model) and one in which attention shifts during the saccadic latency period (temporal model). In the spatial model attention can be allocated to saccadic and perceptual targets, but both sites compete with each other, so that too much attention at the perceptual site impairs the planning of the saccade such that accuracy decreases and latency increases. The temporal model suggests that both systems use the same spatial parameters such that when a saccade is elicited attention has to be settled on the saccadic endpoint. Issuing the go signal too early leads to errors.

Our reentry hypothesis fits with aspects of both models but has more similarity with the temporal model. Activity hills in our premotor map strongly compete with each other in order to ensure only one hill to determine the endpoint. In fast saccades the competition is not resolved so that the superior colliculus averages the planned directions. This would explain why saccades with latencies less than 270 ms can show spatial averaging if two targets are presented simultaneously (Findlay, 1997). Since we predict that the phenomena of spatial attention originate in the reentry signal from premotor cells, a task that requires to attend another location than the endpoint of a saccade, requires two different activity hills in the premotor map. The fact that both of them compete with each other resembles the spatial model. Similar to the temporal model is the fact

that an early competition between two activity hills is not harmful for determining the endpoint of the saccade. However, at the time one activity hill reaches the critical threshold activity for eliciting a saccade, remaining activity at other locations shifts the saccadic endpoint due to averaging. Thus, under natural viewing conditions spatial attention and eye movement selection are automatically coordinated such that prior to the eye movement the amount of reentry is maximized at the endpoint and minimized elsewhere. Spatial attention could be interpreted as a shortcut. In fact, we predict that there is no separate spatial attention system. Observed interference effects could be explained by our reentry hypothesis.

4.5. Spatial selection of one vs. multiple locations

Guided Search (Wolfe, 1994) and several other models (Itti & Koch, 2000; Koch & Ullman, 1985; Treisman & Gelade, 1980) define attention as the result of a winner-takes-all process on an activation or saliency map. Thus, by definition spatial attention is directed only to a single location. Recent experiments indicate that attention can be split (Awh & Pashler, 2000; Bichot et al., 1999). In order to explain the split of attention additional activation maps within each feature map have been suggested (Bichot et al., 1999). Each of these activation maps can be deliberately accessed by memory to induce a top–down activation. Cave (1999) justifies such additional activation maps as a way to independently control the gate to the next higher layer. Such additional activation maps indicate the distractor items to be suppressed (Bichot et al., 1999). Cave (1999) also suggested a FeatureGate model which determines the number of selected locations depending on the task by relevance and salience instead of irrelevance. When searching for a single target the model should employ a winner-take-all competition so that only one gate is open and the stimulus features can pass to the next higher stage. In the discussed shape comparison task (experiment 2), however, the winner-take-all should be relaxed so that every location with the target color is allowed to transmit its features. Both simulated tasks could be in principle explained by such a mechanism. Although we cannot rule out the existence of such a mechanism, we have to consider that it also requires to explain in detail the conditions that determine the selection of one or more locations (from a computational perspective this seems to be not trivial) and the possible costs. Given that one letter at the behaviorally relevant location was correctly reported, the probability of reporting the letter at the other behaviorally relevant location is about 30% (Bichot et al., 1999). Obviously there seems to be a cost. It is not clear if this cost is fully determined by processes after selection.

We offer an anatomically plausible explanation of the enhanced report of letters at multiple behaviorally relevant locations. Since two locations are enhanced due to the described feature-specific feedback, each of these locations "captures spatial attention" due to reentry as part of an automatic process for planning an eye movement (e.g. from the FEF movement cells) or an action in general. First of all, this reentry signal is continuous, not discrete. Second, since the premotor map has evolved to determine the location of a saccade, two activity regions compete with each other such that the amount of reentry for a single target is higher than for two competing target locations. This already predicts a cost prior to the recognition process of the letter probes. Third, competition decreases the firing rate of cells encoding the stimuli at all locations but one over time.

This leads us to differences of a feedforward model compared with the proposed reentry model. The feedforward model predicts that a variation of the SOA's in experiment 2 has no effect on the individual probabilities of report at the two behaviorally relevant locations. The proposed reentry model, however, predicts that a facilitated report of letters at two or more locations in space is only temporary (even when the eye movement is suppressed). With longer SOA's the probability of a correct report increases at one location and decreases at others. Indeed, our data obtained from a repetition of the experiment of Bichot et al. (1999) with different SOA's confirms this prediction (Hamker & VanRullen, 2002).

It was also shown in dual task experiments that under some conditions a high probability of report can be achieved in a central task (e.g. a T/L comparison) and a peripheral task as well. We now discuss why these findings should not be explained by a reentry from two locations.

4.6. Saliency map vs. reentry from premotor areas

Correlated with the separation of processing into preattentive and attentive, many models assume that for a stimulus to be processed in higher areas it has to be selected. Although the FeatureGate model (Cave, 1999) does not show a strong separation between both processing modes it also suggests a bottom–up activation system that enables a singleton to keep its gate open.

We do not question the the concept of saliency map itself, but rather the use of the map in order to explain attention. Spatial attention is not a constant (Müller & Findlay, 1988); it is a changing feedback process based on the encoded stimulus-driven salience and goal-directed knowledge in visual areas. This temporal characteristic of attention allows us to give a reasonable answer why the report of briefly presented items (followed by a mask) does not benefit from the known

target feature (Moore & Egeth, 1998; Shih & Sperling, 1996). Spatial selection needs time to evolve. Reentry arrives too late to facilitate processing of the whole item. Similarly, Weichselgartner and Sperling (1987) found high late detection rates 300 ms after target presentation in a rapid visual presentation task. Our simulation (Fig. 7) also indicates a peak reentry activity of the premotor cells at around 300 ms. The timing of the reentry in a model of attention is therefore essential to explain these effects.

Since our model defines spatial attention as a slow reentry signal we have to explain fast pop-out effects. As an alternative to a fast selection by a saliency map, stimulus-driven saliency might emerge from local and long-range interactions within the pathway from V1–V4 (Knierim & Van Essen, 1992; Li, 2002; Nothdurft, Gallant, & Van Essen, 1999; Super, Spekreijse, & Lammé, 2001; Wörgötter & Eysel, 2000), perhaps even in later stages. Such a temporally early, parallel enhancement of populations encoding salient stimuli would explain a high report ratio of simple salient objects (Braun & Julesz, 1998) and animals that sufficiently discriminate from the background (Li, VanRullen, Koch, & Perona, 2002), even when spatial attention is primarily directed to a central task. Selection by location facilitates but might not be necessary for report, especially if the target does sufficiently discriminate from distractors and if it is salient.

Yet, there had been no experiment that could rule out an explicit saliency map which selects and gates items for further processing. However, it seems that an early parallel stimulus-driven salience module which does not select just one item (and therefore differs from a selection of attention on a saliency map) but enhances the item's competitive weight and a later spatial reentry stage is a serious alternative to a selection on a saliency map. In addition, our model is well defined by anatomy and physiology, whereas previous models using a saliency map remained vague and did not fully link this map to known processes in the brain.

4.7. Conclusion

We conclude that the consideration of time as an additional dimension in models offers new possible explanations regarding the cause and effect of attentional selection. At present we have focused on the spatial reentry stage and how this process can be guided by feature cues. Our model, based on continuous temporal dynamics, predicts that attention builds up gradually by convergence and feedback. As a result, a split of spatial attention reflects an intermediate stage of target selection. Future experiments and simulations, of course, have to provide converging support for this model.

Appendix A

At each location $x \in \{1, \ldots, 6\}$ and each dimension $d \in \{\text{color}, \text{orientation}\}$ in the simulation of the experiment of Motter (1994a, 1994b) and $d \in \{\text{color}, \text{shape}\}$ in the simulation of the experiment of Bichot et al. (1999) we simulate a neural population of i cells by a mean field approach (Eq. (A.1)).

$$\tau \frac{d}{dt} y_{d,i,x}^{V4} = I_{d,i,x}^{V4} - (y_{d,i,x}^{V4} + 0.1) I_{d,x}^{\text{inh}} \tag{A.1}$$

The overall input $I_{d,i,x}^{V4}$ is a result of the bottom–up signal $I_{d,i,x}$ modulated by top–down gain control. The model that was used for the simulations did not contain mechanisms to compute stimulus-driven saliency (Fig. 3).

$$I_{d,i,x}^{V4} = S_{d,i,x}^{V4} \cdot I_{d,i,x} + I_{d,i,x} \cdot \sum_j w_{ij} y_{d,j,x}^{V4} + I_{d,i,x} \cdot y_{d,i}^{\text{ITt}}$$
$$+ I_{d,i,x} \cdot y_x^{\text{pmot}} + N \tag{A.2}$$

Expected information from the feedback pathways ($y_{d,i}^{\text{ITt}}$ for feature-feedback and y_x^{pmot} for location feedback) and from adjacent cells $\sum_j w_{ij} y_{d,j,x}$ is incorporated into the population activity by continuously comparing it with the bottom–up input pattern $I_{d,i,x}$ (Fig. 4). They implement a neural Bayesian inference operation, i.e. an input pattern is compared with the prior or expected information encoded in the network (Koechlin et al., 1999). An input pattern that matches a top–down or lateral pattern increases its weight. The populations with location feedback reduce the influence of other populations by inhibition as proposed by the Biased Competition framework (Desimone & Duncan, 1995). N is a noise term that leads to variations in the transmission from one cell to another. Each connection has an independent noise term.

Input stimuli $I_{d,i,x}$ are generated according to the experimental scenes and are encoded as populations determined by a Gaussian distribution. For realistic experimental conditions, we delayed the input for 30 ms to account for the time a stimulus needs to reach V2. Since cells typically fire very strongly in the beginning and then decrease in sensitivity, we use a short term synaptic depression $S_{d,i,x}^{V4}$.

$$\tau_S \frac{d}{dt} s_{d,i,x}^{V4} = I_{d,i,x} - s_{d,i,x}^{V4}; \quad S_{d,i,x}^{V4} = (1 - \alpha \cdot s_{d,i,x}^{V4}) \tag{A.3}$$

Inhibition of V4 cells (Eq. (A.1)) arises from a baseline inhibition B, from the local activity within the population $\sum_i y_{d,i,x}^{V4}$ and from long range activity $w_{\text{inh}}^{\text{RF}} z_{d,x}^{V4}$ (Eq. (A.4)).

$$I_{d,x}^{\text{inh}} = B + w_{\text{inh}} \sum_i y_{d,i,x}^{V4} + w_{\text{inh}}^{\text{RF}} z_{d,x}^{V4} \tag{A.4}$$

We assume that long range inhibition is controlled by a pool of inhibitory neurons $z_{d,x}^{V4}(t)$ which collect the activity of each population.

$$\tau_{\text{inh}}^{\text{RF}} \frac{d}{dt} z_{d,x}^{V4} = \sum_x \max_j [y_{d,j,x}^{V4}] - z_{d,x}^{V4} \tag{A.5}$$

The receptive field of IT cells comprises the receptive fields of all V4 cells $\text{RF}(x^{\text{IT}}) = x \in \{1, \ldots, 6\}$. We omit x^{IT} in the following equations, since we only simulate one population in IT. We do not increase the complexity of features from V4 to IT. The population of IT cells reads:

$$\tau \frac{d}{dt} y_{d,i}^{\text{ITs}} = I_{d,i}^{\text{ITs}} - w_{\text{inh}}(y_{d,i}^{\text{ITs}} + 0.1) \left(B + \sum_i y_{d,i}^{\text{ITs}} \right) \tag{A.6}$$

Similar to V4, IT cells are inhibited by a baseline term B and the activity within the population $\sum_i y_{d,i}^{\text{ITs}}$.

The V4 activity $y_{d,i,x}^{V4}$ drives the IT cells but feedback implements a gain control ($y_{d,i}^{\text{PF}}$ for feature-feedback and y_x^{pmot} for location feedback)

$$I_{d,i}^{\text{ITs}} = S_{d,i}^{\text{ITs}} \cdot \max_x(y_{d,i,x}^{V4}) + \max_x(y_{d,i,x}^{V4} \cdot y_{d,i}^{\text{PF}})$$
$$+ \max_x(y_{d,i,x}^{V4} \cdot y_x^{\text{pmot}}) + N \tag{A.7}$$

with

$$\tau_S \frac{d}{dt} s_{d,i}^{\text{ITs}} = \max_x(y_{d,i,x}^{V4}) - s_{d,i}^{\text{ITs}}; \quad S_{d,i}^{\text{ITs}} = (1 - \alpha \cdot s_{d,i}^{\text{ITs}}) \tag{A.8}$$

According to a previous study (Hamker, in press) we simulate a convergent projection from areas with smaller receptive filed sizes (V4) to areas with larger receptive field sizes (IT) with a max-pooling function. This prevents multiple distractors placed within the receptive field from adding up their feature weight and dominating the competition.

ITt is part of the feedback pathway and gets only input from ITs cells. Strong inhibition allows only a few active cells to feed back into V4.

$$\tau \frac{d}{dt} y_{d,i}^{\text{ITt}} = \sigma(w^{\text{ITs}} \cdot y_{d,i}^{\text{ITs}} - 0.1) + y_{d,i}^{\text{ITs}} \cdot \sum_j w_{ij} y_{d,j}^{\text{ITt}}$$
$$- w_{\text{inh}}\left(y_{d,i}^{\text{ITt}} + 2\right) \sum_j y_{d,j}^{\text{ITt}} \tag{A.9}$$

$$\sigma(a) = \max(a, 0)$$

Perceptual neurons receive convergent afferents from model V4 at the same retinotopic location. Different dimensions d add up.

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} y_x^{\mathrm{vis}} = w^{\mathrm{V4s}} \sum_d \max_i(y_{d,i,x}^{\mathrm{V4s}}) + w^{\mathrm{pmot}} y_x^{\mathrm{pmot}}$$
$$- w_{\mathrm{inh}} \max_x(y_x^{\mathrm{vis}}) + w_{\mathrm{inh}}^{\mathrm{map}} z^{\mathrm{vis}} \quad (A.10)$$
$$\tau_{\mathrm{inh}}^{\mathrm{map}} \frac{\mathrm{d}}{\mathrm{d}t} z^{\mathrm{vis}} = \sum_x y_x^{\mathrm{vis}} - z^{\mathrm{vis}}$$

Increased activity in premotor cells occurs when the input from the perceptual cells is larger than the slight surround inhibition.

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} y_x^{\mathrm{pmot}} = w^{\mathrm{vis}} y_{\mathbf{x}}^{\mathrm{vis}} - w_{\mathrm{inh}}^{\mathrm{vis}} \sum_x y_x^{\mathrm{vis}} - I^{\mathrm{inh}}$$
$$I^{\mathrm{inh}} = w_{\mathrm{inh}} \max_x(y_x^{\mathrm{pmot}}) + w_{\mathrm{inh}}^{\mathrm{map}} \sum_x y_x^{\mathrm{pmot}} \quad (A.11)$$
$$- w^{\mathrm{fix}} y^{\mathrm{fix}}$$

Since fixation is required, activity of a fixation cell y^{fix} inhibits the cells in the premotor map in order to avoid activity levels that initiate an eye movement.

The underlying circuits, which are responsible for feedback in the prefrontal circuits are still unclear. We simply assume that the feedback signal depends on the difference between the target signal $I_{d,i}^{\mathrm{Target}}$ and the ITs cell activity.

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} y_{d,i}^{\mathrm{PF}} = I_{d,i}^{\mathrm{Target}} - y_{d,i}^{\mathrm{ITs}} \quad (A.12)$$

References

Ahmad, S. (1991). *VISIT: An efficient computational model of human visual attention*. PhD Thesis, University of California, Berkley.

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). Lawrence Erlbaum Associates.

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 834–846.

Bichot, N. P., & Schall, J. D. (1999). Effects of similarity and history on neural mechanisms of visual selection. *Nature Neuroscience, 2*, 549–554.

Bichot, N. P., Cave, K. R., & Pashler, H. (1999). Visual selection mediated by location: Feature-based selection of noncontiguous locations. *Perception and Psychophysics, 61*, 403–423.

Bichot, N. P., Rao, S. C., & Schall, J. D. (2001). Continuous processing in macaque frontal eye cortex during visual search. *Neuropsychologia, 39*, 972–982.

Braun, J., & Julesz, B. (1998). Withdrawing attention at little or no cost: Detection and discrimination tasks. *Perception and Psychophysics, 60*, 1–23.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547.

Bundesen, C. (1999). A computational theory of visual attention. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space and action: Studies in cognitive neuroscience* (pp. 54–71). Oxford University Press.

Burman, D. D., & Segraves, M. A. (1994). Primate frontal eye field activity during natural scanning eye movements. *Journal of Neurophysiology, 71*, 1266–1271.

Caputo, G., & Guerra, S. (1998). Attentional selection by distractor suppression. *Vision Research, 38*, 669–689.

Cave, K. R. (1999). The FeatureGate model of visual selection. *Psychological Research, 62*, 182–194.

Cepeda, N. J., Cave, K. R., Bichot, N. P., & Kim, M. S. (1998). Spatial selection via feature-driven inhibition of distractor locations. *Perception and Psychophysics, 60*, 727–746.

Chelazzi, L. (1999). Serial attention mechanisms in visual search: A critical look at the evidence. *Psychological Research, 62*, 195–219.

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology, 80*, 2918–2940.

Corbetta, M., & Shulman, G. L. (1998). Human cortical mechanisms of visual attention during orienting and search. *Philosophical Transactions of the Royal Society of London Series B: Biological Science, 353*, 1353–1362.

Deco, G., & Zihl, J. (2000). Neurodynamical mechanism of binding and selective attention for visual search. *Neurocomputing, 32–33*, 693–699.

Deco, G., Pollatos, O., & Zihl, J. (2002). The time course of selective visual attention: Theory and experiments. *Vision Research, 42*, 2925–2945.

Desimone, R. (1999). Visual attention mediated by biased competition in extrastriate visual cortex. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space and action* (pp. 13–30). Oxford University Press.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective attention. *Annual Review of Neuroscience, 18*, 193–222.

Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *Journal of Neurophysiology, 57*, 835–868.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*, 1827–1837.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501–517.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Duncan, J., & Humphreys, G. W. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 578–588.

Duncan, J., Humphreys, G. W., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology, 7*, 255–261.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception and Psychophysics, 62*, 425–451.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 32–39.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 583–597.

Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research, 37*, 617–631.

Ghirardelli, T. G., & Egeth, H. E. (1998). Goal-directed and stimulus-driven attention in cross-dimensional texture segregation. *Perception and Psychophysics, 60*, 826–838.

Goodale, M. A., & Humphrey, G. K. (1998). The objects of action and perception. *Cognition, 67*, 181–207.

Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature, 391*, 481–484.

Hamker, F. H. (1999). The role of feedback connections in task-driven visual search. In D. Heinke, G. W. Humphreys, & A. Olson (Eds.), *Connectionist models in cognitive neuroscience* (pp. 252–261). London: Springer Verlag.

Hamker, F. H. (2000). Distributed competition in directed attention. In G. Baratoff & H. Neumann (Eds.), *Proceedings in artificial intelligence, dynamische perzeption* (vol. 9, pp. 39–44). Berlin: AKA, Akademische Verlagsgesellschaft.

Hamker, F. H. (2002). How does the ventral pathway contribute to spatial attention and the planning of eye movements. In R. P. Würtz & M. Lappe (Eds.), *Dynamic Perception* (pp. 83–88). St. Augustin: Infix Verlag.

Hamker, F. H. (in press). *Predictions of a model of spatial attention using sum- and max-pooling functions*. Neurocomputing.

Hamker, F. H., & VanRullen, R. (2002). The time course of attentional selection among competing locations (Abstract). *Journal of Vision, 2*(7), 7a.

Hamker, F. H., & Worcester, J. (2002). Object detection in natural scenes by feedback. In H. H. Bülthoff et al. (Eds.), *Lecture notes in computer science, Biologically motivated computer vision* (pp. 398–407). Berlin: Springer Verlag.

Hoffman, J. E. (1979). A two-stage model of visual search. *Perception and Psychophysics, 25*, 319–327.

Hoffman, J. E., & Subramian, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57*, 787–795.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*, 171–216.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.

Kaptein, N. A., Theeuwes, J., & Van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1053–1069.

Kim, M.-S., & Cave, K. R. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychological Science, 6*, 376–380.

Kim, M.-S., & Cave, K. R. (2001). Perceptual grouping via spatial selection in a focused-attention task. *Vision Research, 41*, 611–624.

Kirkland, K. L., & Gerstein, G. L. (1999). A feedback model of attention and context dependence in visual cortical networks. *Journal of Computational Neuroscience, 7*, 255–267.

Knierim, J. J., & Van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1. *Journal of Neurophysiology, 67*, 961–980.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Psychology, 4*, 219–227.

Koechlin, E., & Burnod, Y. (1996). Dual population coding in the neocortex: A model of interaction between representation and attention in the visual cortex. *Journal of Cognitive Neuroscience, 8*, 353–370.

Koechlin, E., Anton, J. L., & Burnod, Y. (1999). Bayesian inference in populations of cortical neurons: A model of motion integration and segmentation in area MT. *Biological Cybernetics, 80*, 25–44.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*, 1897–1916.

Kröse, B., & Julesz, B. (1989). The control and speed of shifts in attention. *Vision Research, 29*, 1607–1619.

Kustov, A. A., & Robinson, D. L. (1996). Shared neural control of attentional shifts and eye movements. *Nature, 384*, 74–77.

Lammé, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience, 23*, 571–579.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6*, 9–16.

Li, F.-F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Science of the United States of America, 99*, 9596–9601.

Martinez, A., DiRusso, F., Anllo-Vento, L., Sereno, M. I., Buxton, R. B., & Hillyard, S. A. (2001). Putting spatial attention on the map: Timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Research, 41*, 1437–1457.

Moore, T. (1999). Shape representations and visual guidance of saccadic eye movements. *Science, 285*, 1914–1917.

Moore, C. M., & Egeth, H. (1998). How does feature-based attention affect visual processing? *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1296–1310.

Motter, B. C. (1994a). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience, 14*, 2178–2189.

Motter, B. C. (1994b). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience, 14*, 2190–2199.

Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research, 38*, 1805–1815.

Mozer, M. C., & Sitton, M. (1998). Computational modeling of spatial attention. In H. Pashler (Ed.), *Attention* (pp. 341–393). East Sussex, UK: Psychology Press.

Müller, H. J., & Findlay, J. M. (1988). The effect of visual attention on peripheral discrimination thresholds in single and multiple element displays. *Acta Psychologica (Amst), 69*, 129–155.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 361–393). Lawrence Erlbaum Associates.

Neumann, O. (1990). Direct parameter specification and the concept of perception. *Psychological Research, 52*, 207–215.

Newsome, W. T. (1996). Visual attention: Spotlights, highlights and visual awareness. *Current Biology, 6*, 357–360.

Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia. *Visual Neuroscience, 16*, 15–34.

Olshausen, B., Anderson, C., & van Essen, D. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience, 13*, 4700–4719.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 108–130.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience, 21*, 227–277.

Perry, R. J., & Zeki, S. (2000). The neurology of saccades and covert shifts in spatial attention: An event-related fMRI study. *Brain, 123*, 2273–2288.

Posner, M. I., & Dehaene, S. (1994). Attentional networks. *Trends in Neuroscience, 17*, 75–79.

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.

Rizzolatti, G., & Craighero, L. (1998). Spatial attention: Mechanisms and theories. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in psychological science: Biological and Cognitive Aspects* (Vol. 2, pp. 171–198). East Sussex, England: Psychology Press.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umiltá & M. Moscovitch (Eds.), *Attention and*

performance XV. Concious and nonconcious information processing (pp. 231–265). Cambridge: MIT Press.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologica, 25*, 31–40.

Rockland, K. S., & van Hoesen, G. W. (1994). Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cerebral Cortex, 4*, 300–313.

Rockland, K. S., Saleem, K. S., & Tanaka, K. (1994). Divergent feedback connections from areas V4 and TEO in the macaque. *Visual Neuroscience, 11*, 579–600.

Roelfsema, P. R., Lammé, V. A., Spekreijse, H., & Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience, 14*, 525–537.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5*, 631–632.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision Research, 43*, 629–637.

Schall, J. D., Morel, A., King, D. J., & Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: Convergence and segregation of processing streams. *Journal of Neuroscience, 15*, 4464–4487.

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention, control of segmentation, object recognition and space-based motor action. *Visual Cognition, 2*, 331–375.

Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience, 16*, 1486–1510.

Sheliga, B. M., Craighero, L., Riggio, L., & Rizzolatti, G. (1997). Effects of spatial attention on directional manual and ocular responses. *Experimental Brain Research, 114*, 339–351.

Shih, S. I., & Sperling, G. (1996). Is there feature-based attentional selection in visual search? *Journal of Experimental Psychology: Human Perception and Performance, 22*, 758–779.

Snyder, L. H., Batista, A. P., & Andersen, R. A. (2000). Intention-related activity in the posterior parietal cortex: A review. *Vision Research, 40*, 1433–1441.

Supèr, H., Spekreijse, H., & Lammé, V. A. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience, 4*, 304–310.

Theeuwes, J. (1993). Visual selective attention. *Acta Psychologica, 83*, 93–154.

Tononi, G., Sporns, O., & Edelman, G. (1992). Reentry and the problem of integrating multiple cortical areas: Simulation of dynamic integration in the visual system. *Cerebral Cortex, 2*, 310–335.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 459–478.

Treue, S., & Martínez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399*, 575–579.

Tsotsos, J. K., Culhane, S. M., Wai, W., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence, 78*, 507–545.

Ullman, S. (1995). Sequence seeking and counter streams: A computational model for bidirectional flow in the visual cortex. *Cerebral Cortex, 5*, 1–11.

van der Heijden, A. H. C. (1992). *Selective attention in vision*. London, New York: Routledge.

van der Velde, F., & de Kamps, M. (2001). From knowing what to knowing where: Modeling object-based attention with feedback disinhibition of activation. *Journal of Cognitive Neuroscience, 13*, 479–491.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123*, 146–160.

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science, 238*, 778–780.

Wörgötter, F., & Eysel, U. T. (2000). Context, state and the receptive fields of striatal cortex cells. *Trends in Neuroscience, 23*, 497–503.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*, 202–238.

Wolfe, J. M. (1999). What can 1 Million trials tell us about visual search? *Psychological Science, 9*, 33–39.

Wolfe, J. M., & Cave, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron, 24*, 11–17.

Wolfe, J. M., Cave, K., & Franzel, S. (1989). Guided Search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 419–433.

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature, 400*, 867–869.