In: Connectionist Models in Cognitive Neuroscience, Proc. of the 5th Neural Computation and Psychology Workshop (NCPW'98). Hrsg. von D. Heinke, G. W. Humphreys, A. Olson. University of Birmingham, England, September 1998. London: Springer Verlag 1999, S. 252-261.

The role of feedback connections in task-driven visual search

F. H. Hamker[*]
Department of Neuroinformatics, Technische Universität Ilmenau
D-98684 Ilmenau, Germany
e-mail: fred@informatik.tu-ilmenau.de

Abstract

Studies of attention suggest a model in which attention emerges from a parallel, distributed competition. Within that framework, this contribution tries to explain the findings of the experiment by CHELAZZI ET AL. [1] by taking into account feedback from successive stages. It is shown that the developed model can qualitatively obtain the same results as measured in the experiment. Furthermore, the model shows promising similarities to human reaction times in visual search tasks.

1 Introduction

Recent neurophysiological findings show that at very early stages, attention modulates the visual processing [2,3]. Among others, the reason is to create a representation that serves to filter out irrelevant information and to emphasise information needed for the current task [4]. For example, V4 neurons show an enhanced activity if the presented colour or luminance items match the features of the cue [5]. This dynamic effect occurs in parallel across the visual field and it seems to segment the scene into possible candidates and background, before a final selection takes place. Relating these observations to psychophysical studies such as visual search, this effect seems to be a physiological counterpart of restricting search to subgroups, which share the same feature [6].

If the task requires to perform a saccadic eye movement or a grasping operation, competition in the dorsal stream has to be coordinated with the processing in the ventral stream. Or in general, the processing in different brain areas must be integrated into a joint behavioural result [7,8]. This contribution aims to illustrate, how a distributed competition for a task-relevant item can be integrated by using backward connections.

A possible basis for goal-directed competition among neurons in inferior temporal cortex (IT) is shown in an experiment by CHELAZZI ET AL. [1]. Monkeys were presented a cue to hold in memory. During a delay period the neuron sensitive for the target showed a higher rate of activity, presumably a top-down projection from the working memory. After the delay period, the monkeys were given a picture containing the cue

---

[*] Since September 1998 he joined the medical data analysis project MEDAN at the J.W.Goethe- Universität Frankfurt (email: hamker@cs.uni-frankfurt.de).

and other objects and instructed to initiate an eye movement towards the target. The recordings of average firing rates show that the neuron sensitive for the target and the neuron sensitive for a distractor became active but only the neuron sensitive for the target remains high or was even enhanced while the neuron sensitive for the distractor was suppressed.

It is supposed by USHER and NIEBUR that this behaviour occurs due to lateral interactions in IT cortex [9]. They present a model which consists of self-excitatory cells coupled with an inhibitory pool. Apart from a bottom-up input, the cells receive a goal directed signal. Their simulations show similar responses as observed in the above experiments. They came to the conclusion that the experiment of CHELAZZI ET AL. shows an example of a parallel competition but in case of a search for conjunctions of features a serial scan is needed: "Assuming that each feature dimension (colors, shapes, etc.) is represented by a different memory module, it is easy to see why the selection mechanism reaches its limitations. In the example [...] red vertical target among red horizontal and blue vertical distractors, the top-down attentional input is provided to both the red and the vertical assemblies. Therefore both assemblies will win the interaction whether there is or there is not a conjunctive target in the display. The parallel stage is useless in this case and the decision has to be based on a serial scan of all objects." [9]

However, two important questions remained unsolved:

- How can a competition among IT neurons define the location of the saccade? Although the receptive field of IT neurons does not cover the complete visual field, it is large enough that even a distributed coding can not clearly indicate the target location.
- How may a target defined by a conjunction of features emerge as a winner by a parallel competition?

This contribution addresses the above questions and presents a solution based on feedback connections from successive layers. In this hypothesis, competition among IT neurons can not be regarded as an isolated process without influence of the preceding layer.

2 The Model

2.1 Overview

This model consists of three simulated areas: an object-sensitive, a location-sensitive and a feature-sensitive area. The neurons in the object-sensitive area may be regarded as IT neurons of two different dimensions (e.g. colour and form) located within the same receptive field (Fig. 1). The feature-sensitive area can be interpreted as V4/V2 neurons at different locations within this receptive field. Again two different dimensions were simulated. No direct cross-dimensional interactions were taken into consideration. The location-sensitive area is responsible for the selection of a location, e.g. for an eye movement.

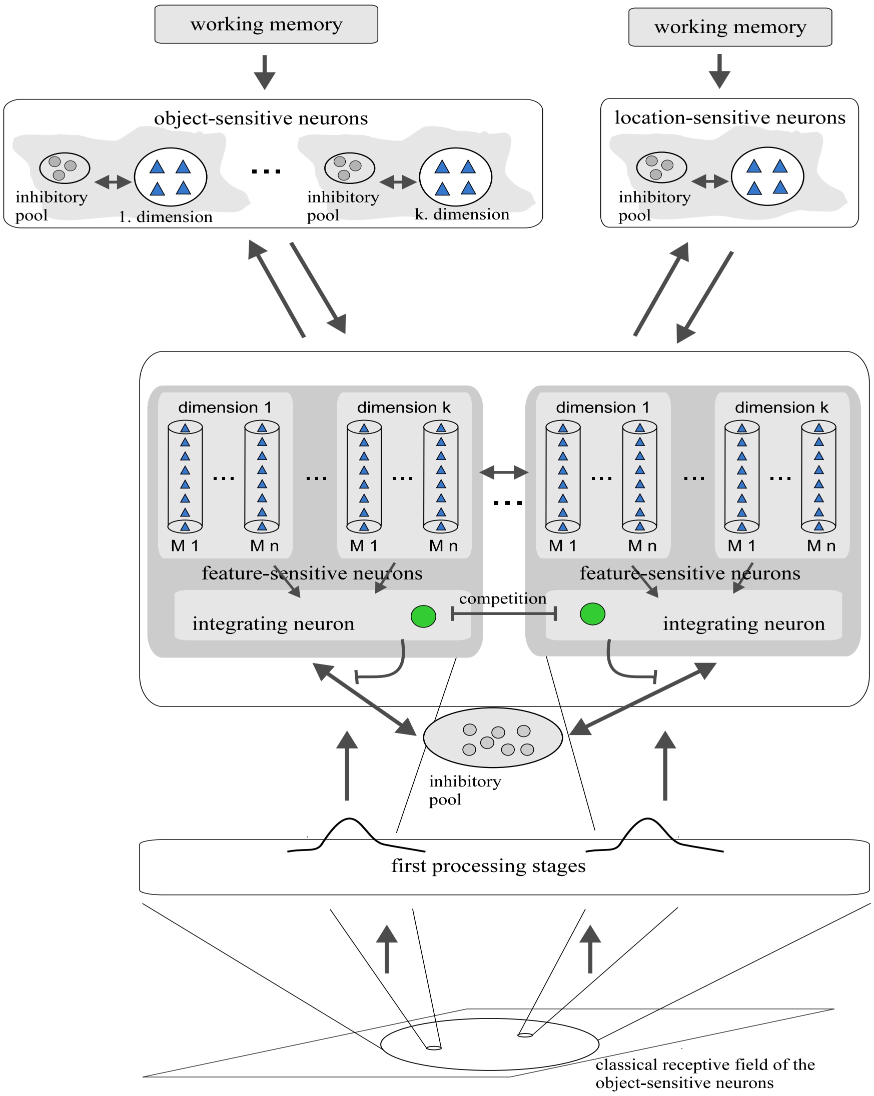

Figure 1. Overview of the model. It consists of three functional blocks which are reciprocially connected. Feature-sensitive neurons of the topographically ordered feature maps code in parallel features in different dimensions and project their activity to neurons of higher complexity and much larger receptive fields called object-sensitive neurons. They simulate processing in the ventral pathway. Similar to the model in [9], they receive a top-down activity from short-term memory coding the features of the cue and compete for visual processing via an inhibitory pool. Feature-sensitive neurons also project to an integrating neuron within their receptive field which sends its activity to location-sensitive neurons simulating one function of the dorsal stream: Their competition determines the location to which attention is directed. A high activity of a location-sensitive neuron allows the feature-sensitive neurons in the same receptive field to enhance their activity while feature-sensitive neurons at other locations are suppressed. Thus, each competition in a stage is not independent from the competition in other stages, they build a closed loop. Similar to the theory in [10], the features in different dimensions are integrated via a common location.

2.2 Feature-sensitive area

Neurons are known to be broadly tuned to variations of stimulus parameters and sharp tuning can be achieved by recurrent cortical excitation [11]. Thus, in this model each feature is represented by a population code within a neural field of local excitation and global inhibition [12] (Fig. 3), and determined by the area of activation. Because the input of the model should simulate the output of the previous stage, again a population code is used where the location determines the feature and the height the strength of each feature (Fig. 2). Thus, the more difficult the task, the broader the input curve and the more salient an item the higher the curve. This kind of coding also relates to the findings of target-distractor and distractor-distractor similarity [13].

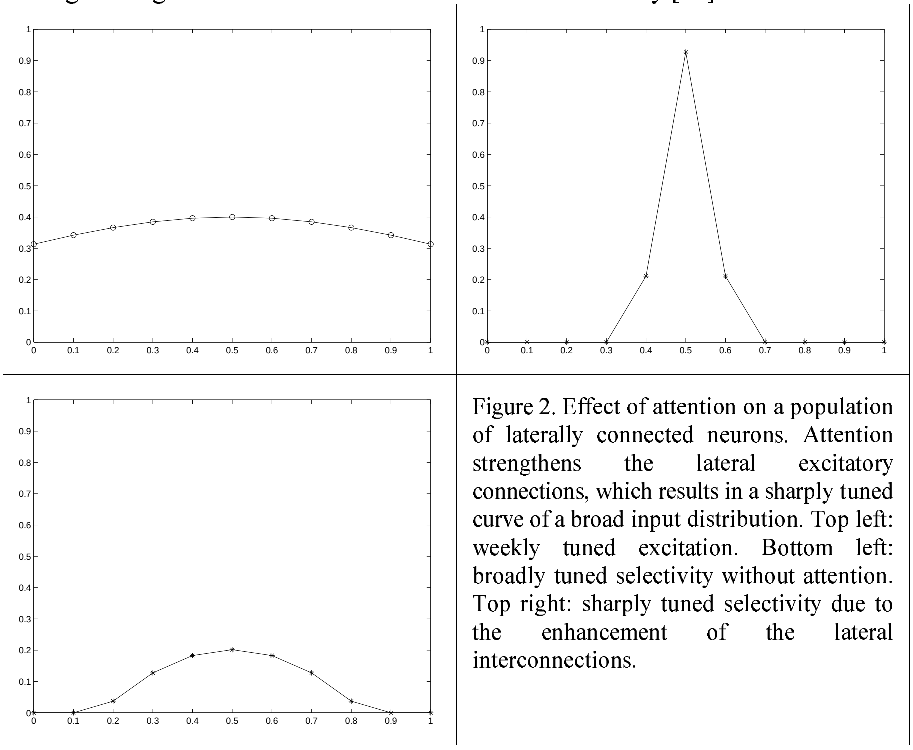

Figure 2. Effect of attention on a population of laterally connected neurons. Attention strengthens the lateral excitatory connections, which results in a sharply tuned curve of a broad input distribution. Top left: weekly tuned excitation. Bottom left: broadly tuned selectivity without attention. Top right: sharply tuned selectivity due to the enhancement of the lateral interconnections.

A high activity of an integrating neuron strengthens the lateral weights in the neural fields of all dimensions located in the same receptive field (Fig. 3). This results in a sharply tuned curve and thus, in an unambiguously coded feature (Fig. 2). Because the integrating neuron also prevents its feature-sensitive neurons from global inhibition within the receptive field of the object-sensitive neurons, the competition for processing leads to an ambiguity resolution [14].

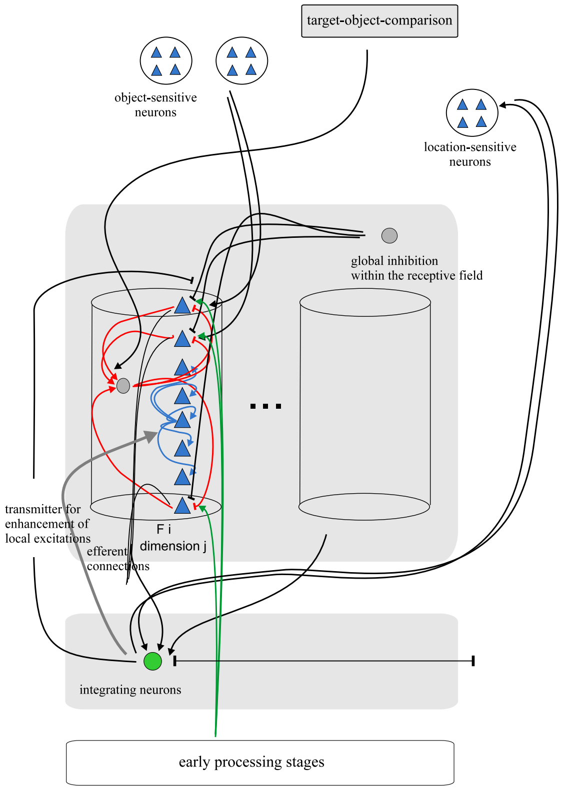

Figure 3. The structure of a neural field. Each column codes a feature by a neural field. It is interconnected with its previous stage (the input of this model) and with its higher stage (the object-sensitive neurons). Every neuron within this field receives a local excitation from its neigbours and an inhibition from all neurons via an interneuron. The strength of the local excitatory weights are modified by the activity of the integrating neuron. Furthermore a high firing rate of the integrating neuron prevents all feature-sensitive neurons at the same location from the global inhibition within the receptive field of the object-sensitive neurons.

2.2 Object-sensitive and location-sensitive area

Different dimensions (e.g. colour or form) were located on different maps. Each map consists of several neurons coding a different object related feature. They compete against each other via a low global inhibition (Fig. 4). External inputs come from the feature-sensitive neurons and from working memory holding the features of the cue. Thus, task-driven visual search is obtained by a brief activation of object-sensitive neurons. This refers to the results in [1], where it was observed that IT neurons coding the relevant features of the task are primed by an external source, which is assumed to bias the overall competition in favour of the goal [7].

All integrating neurons from the feature-sensitive area project to location-sensitive neurons (Fig. 4), which compete for the location of a planned action (e.g. an eye movement). The firing rates of the integrating neurons hold all information needed for determining the location of an action, because they sum up the activity of the feature-sensitive neurons from different dimensions, which code the bottom-up saliency and the task-driven top-down importance by their activity. For the purpose of an anticipation of a planned action the decision enhances processing in the ventral stream by a feedback from the location-sensitive neurons to the integrating neurons, which appears to shrink the receptive field around the target [15].

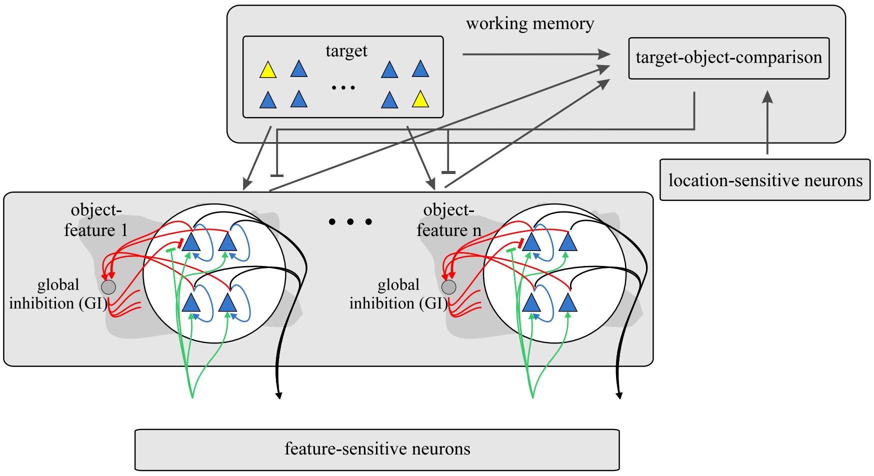

Figure 4. Connections in the object-sensitive area. The input from the feature-sensitive neurons project in two ways onto the object-sensitive neurons. The first component is a regular additive input and the second prevents the neuron from the global inhibition. The latter works very nonlinearly and has only an effect if the firing rate of a feature-sensitive neuron is high. After the onset of visual search and after every wrong selection, the object-sensitive neurons are activated for a short time from working memory.

3 Simulation results

The results presented here were restricted to a target defined as a conjunction of features. Thus, both dimensions in the object-sensitive area (say red and vertical) receive a top-down input, assumed to be directed from working memory.

In extension to the USHER and NIEBUR model [9], the neurons of the feature-sensitive area compete for projecting to the next higher stage, i.e., within the receptive field of an object-sensitive neuron the locations of the feature-sensitive neurons compete for an enhanced processing. By recurrent connections resonance occurs only at those locations where both dimensions code the properties of the target.

The top row in figures 5 and 6 illustrates the strength of overlap between different features in the input of the model. The broader the input, the more the maximum changes due to noise. In each dimension the target and a distractor have one feature in common. Although the target object-sensitive neurons receive a top-down input in both dimensions and gain a competitive advantage (3rd row in fig. 5), a distractor which is equal to the target in the first dimension wins the competition first (fig. 5): A selection in the location modifies the competition in the feature-sensitive area as explained. The emerging sharply tuned neural curve causes the object-sensitive neurons to represent the selected object features. By inhibition of return the winner breaks down and the search is continued unless the target is detected.

Comparing this experiment with the second experiment (fig. 6) in which the overlap of the features in the input is reduced in the second dimension, the top-down activation is now strong enough to guide the competition directly to the target.

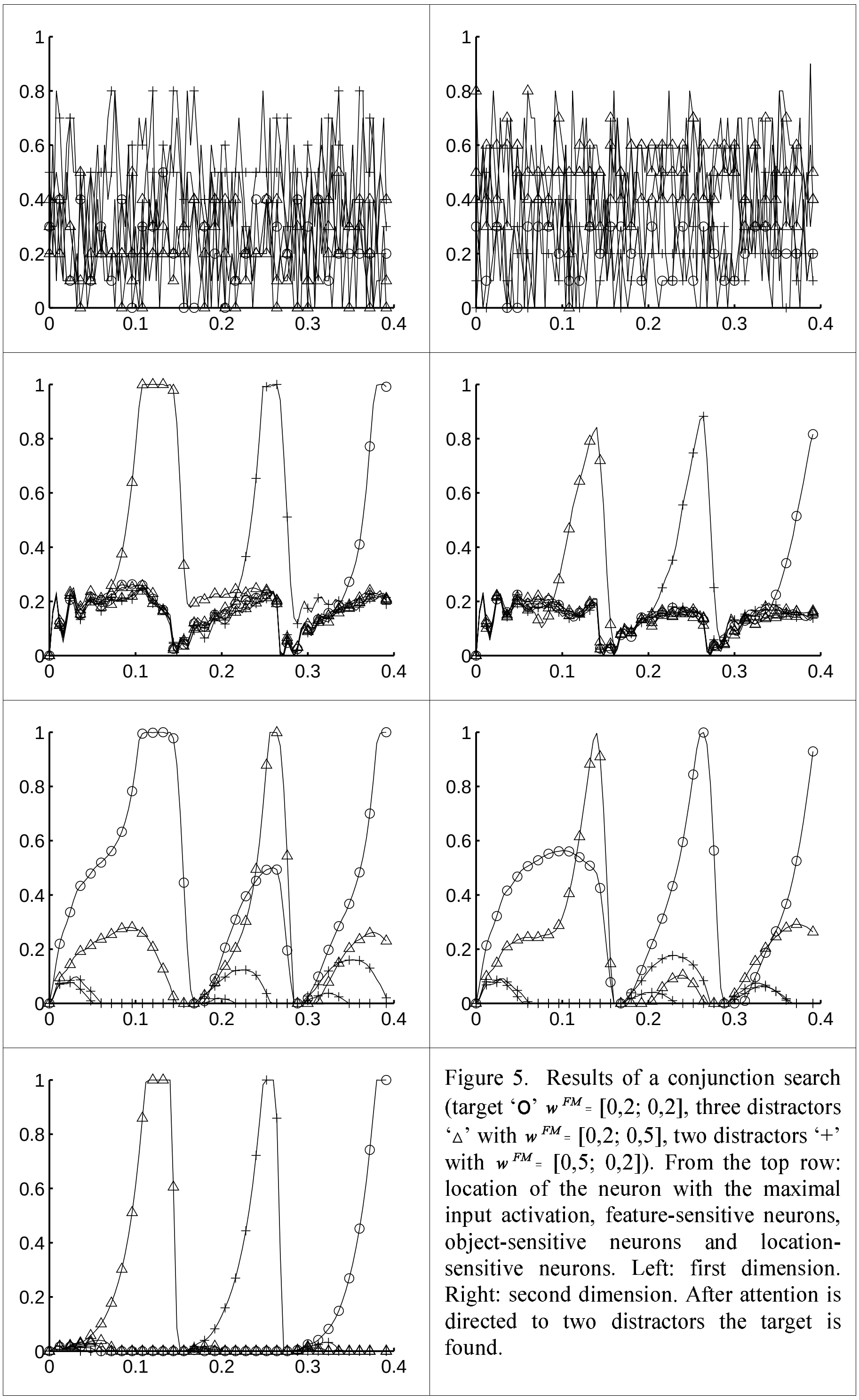

Figure 5. Results of a conjunction search (target ‘O’ $w^{FM}=[0,2;\ 0,2]$, three distractors ‘△’ with $w^{FM}=[0,2;\ 0,5]$, two distractors ‘+’ with $w^{FM}=[0,5;\ 0,2]$). From the top row: location of the neuron with the maximal input activation, feature-sensitive neurons, object-sensitive neurons and location-sensitive neurons. Left: first dimension. Right: second dimension. After attention is directed to two distractors the target is found.

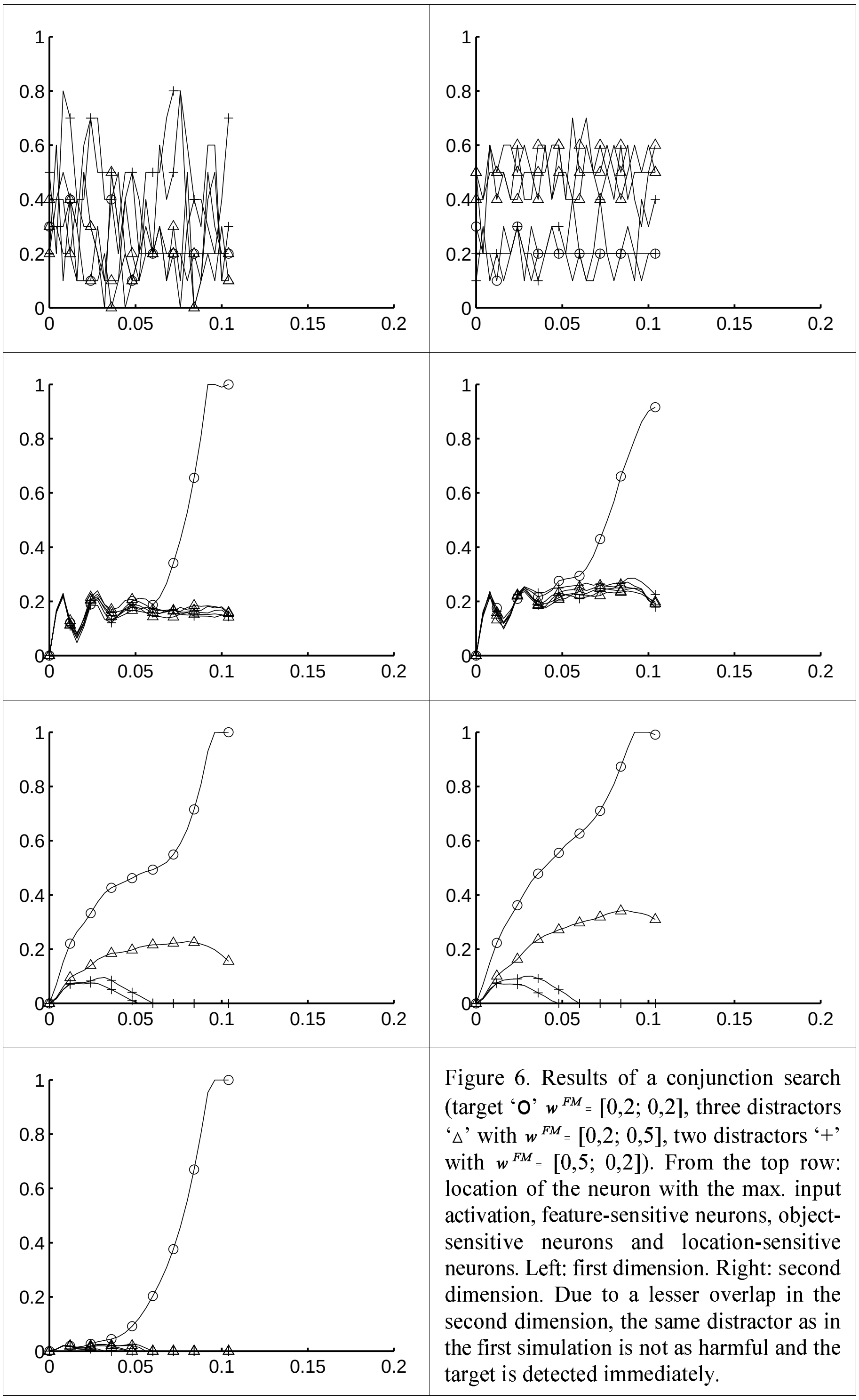

Figure 6. Results of a conjunction search (target 'O' w^{FM} = [0,2; 0,2], three distractors '△' with w^{FM} = [0,2; 0,5], two distractors '+' with w^{FM} = [0,5; 0,2]). From the top row: location of the neuron with the max. input activation, feature-sensitive neurons, object-sensitive neurons and location-sensitive neurons. Left: first dimension. Right: second dimension. Due to a lesser overlap in the second dimension, the same distractor as in the first simulation is not as harmful and the target is detected immediately.

4 Discussion

4.1 Relation to experiments

Coming back to the introductory questions, the model explains how a competition among IT neurons can define the location of the saccade: The computational advantage is transferred via feedback to layers with smaler receptive fields. These neurons code the selective weight of an object feature by their activity but in a finer locational resolution. Because different dimensions are linked together due to their common location, the competition in separate areas is integrated into a uniform perception.

In extension to explaining a possible role of feedback to lower stages in the Chelazzi et al. [1] experiment, several visual search tasks have been simulated showing promising similarities to human reaction times. Fig. 6 shows the average times. It is assumed that the arrangement and choice of the items influences the neural representation of a population code to some extend as in this model, which improves or deterioates the capability of a parallel competition resulting in different detection times.

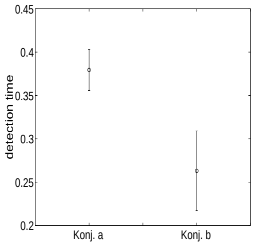

Figure 6. Mean and stdv. of detection times with five distractors (10 simulations with different noise). In the first experiment the target is found on average after the detection of two distractors – often termed as "serial" search. But it is known that conjunction search can be more efficient than "serial" seach times, e.g. if saturated colour stimuli are used (exp. 7 in [16]). These easier conditions result in a faster detection of the target, as illustrated in the second experiment.

4.2 Comparison to other models

Mozer and Sitton distinguish between two classes of models: the *spatial-selection perspective* and the *ubiquitous-competition perspective* [17]. In the *spatial-selection perspective* attention selectively routes the neural activation; attention selection and the representation of object features are separated modules. Most models of attention belong in this category. In the *ubiquitous-competition perspective* attention emerges from a competition among representations of stimuli in different areas. An example of this category is [18].

The model presented here belongs the the latter category. Although it shows some similarities to [19], it suggests an underlying neural circuit how object representation at different hierarchies and location may interact. A feedback of a location area to the represenation of stimuli is also presented in [20], but the competition in the identification pathway is disregarded.

Although there are no findings that clearly settle between those models, the *ubiquitous-competition perspective* seems to be more in line with current results and theories, compare [7,8,3].

5 Conclusion

The results of this model suggest that visual search is always a parallel competition among objects and no serial mechanism is needed but serial processing occurs by the enhancement of desired patterns and suppression undesired patterns. This model supports the theories of a parallel and integrated competition in which different areas bias the competition, e.g. [7, 8]. Future research has to extend the model to simulate more than one receptive field and to consider experimental results more precisely.

References

1. Chelazzi L, Miller EK, Duncan J, Desimone,R. A neural basis for visual search in inferior temporal cortex. Nature 1993; 363: 345-347.
2. Motter BC. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. Journal of Neurophysiology 1993; 70:909-919.
3. Luck SJ, Chelazzi L, Hillyard SA, Desimone, R.: Mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex. Journal of Neurophysiology 1997; 77:24-42.
4. Maunsell JHR. The brains's visual world: Representation of visual targets in cerebral cortex. Science 1995; 270:764-769.
5. Motter BC. Neural correlates of attentive selection for color or luminance in extrastriate area V4. Journal of Neuroscience 1994; 14:2178-2189.
6. Egeth HE, Virzi RA, Garbart H. Searching for conjunctively defined targets. Journal of Experimental Psychology: Human Perception and Performance 1984; 10:32-39.
7. Desimone R, Duncan J. Neural mechanisms of selective attention. Anu. Rev. of Neurosc. 1995; 18:193-222.
8. Duncan J, Humphreys GW, Ward R. Competitive brain activity in visual attention. Current Opinion in Neurobiology 1997; 7:255-261.
9. Usher M, Niebur E. Modeling the temporal dynamics of IT neurons in visual search: A mechanism for top-down selective attention. Journal of Cognitive Neuroscience 1996; 8(4):311-327.
10. Schneider WX. VAM: A neuro-cognitive model for visual attention, control of segmentation, object recognition and space-based motor action. Visual Cognition 1995; 2:331-375.
11. Somers DC, Nelson SB, Sur M. An emergent model of orientation selectivity in cat visual cortical simple cells. The Journal of Neurocience 1995; 15:5448-5465.
12. Amari S, Arbib MA. Competition and cooperation in neural nets. In: Systems Neuroscience. Metzler J (ed). Academic Press, San Diego, 1977, pp. 119-165.
13. Duncan J, Humphreys GW. Beyond the search surface: Visual search and attentional engagement. Journal of Experimental Psychology: Human Perception and Performance 1992; 18: 578-588.
14. Luck SJ, Girelli M, McDermott MT, Ford MA. Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. Cognitive Psychology 1997; 33:64-87.
15. Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. Science 1985; 229:782-784.
16. Wolfe J, Cave K, Franzel S. Guided Search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology 1989; 15:419-433.
17. Mozer MC, Sitton M. Computational modeling of spatial attention. In: Pashler H (ed) Attention. Psychology Press, East Sussex, UK, 1998, pp. 341-393.
18. Phaf RH; van der Heijden AHC, Hudson PTW. SLAM: A connectionist model for attention in visual selection tasks. Cognitive Psychology 1990; 22:273-341.
19. Wolfe J. Guided search 2.0 A revised model of visual search. Psy. Bulletin & Review 1994; 1:202-238.
20. van der Heijden AHC. Selective attention in vision. Routledge, London, New York, 1992.