In: Proceedings in Artificial Intelligence, Vol. 9. Dynamische Perzeption, Workshop der GI-Fachgruppe 1.0.4 Bildverstehen. Hrsg. von G. Baratoff, H. Neumann. Berlin: AKA, Akademische Verlagsgesellschaft, 2000, pp. 39-44.

Distributed competition in directed attention

Fred H. Hamker
CALIFORNIA INSTITUTE OF TECHNOLOGY
Division of Biology 139-74
Pasadena, CA 91125

Abstract

Studies of attention suggest a model in which attention emerges from a parallel, distributed competition. The central purpose of this contribution is to explain the findings of single cell recordings gained in two different experimental paradigms by taking into account feedback from successive stages. It is shown that the developed model can quantitatively obtain similar results as measured in the experiment.

1 Introduction

The neural processes of feature analysis, object selection by attention and object recognition have traditionally been decomposed into distinct, often sequential stages. Recent neurophysiological findings, however, show that attention modulates the visual processing at early stages [1],[2].

Taken the results of neurophysiological recordings together, the effect of attention turns out not to be a simple enhancement of processing within a spotlight of attention. The effect of attention depends on whether a presentation of stimuli within the receptive field is sequential or simultaneous [5], [2]. Furthermore, attention reflects the current interest of the viewer, i.e. a top-down component adds and modifies the pattern of the visual world as reflected in experiments recording from V4 [6].

Models proposed so far have only partially constrained by neurophysiological recordings. Competition might inevitably entangled with the processing of stimuli, as suggested by Usher and Niebur [7]. They suppose that a parallel competition based on lateral interactions within one stage is sufficient for simple feature search, but they kept us dark about what mechanism might guide a search for conjunctions. Instead of switching off the input of non-attended locations, a different approach is to assume a temporal tagging in which all neurons within a focus of attention receive a higher synchronous firing probability [8]. This model was used to reproduce experimental data obtained from Moran and Desimone [3]. A more fine graded model presented by Grossberg and Raizada [9] illustrates the dynamics of an attended location task, e.g. as in the experiment of Reynolds et al. [5], and grouping in V1 and V2. Similar as in the previous model [8], attention is implemented as a locational bias, but the attentional capturing mechanism is not explained.

White spots on the map of models cover the issue of top-down activation by targets and the role of feedback connections, expected to play a major role in the cortical information flow. Recently, in extension to the model of Usher and Niebur [7], it was shown that even a search for conjunctions can be based on a parallel competition if the features are bound by their common location due to feedback [10]. The sequential selection, that is often presupposed, is a result of a parallel competition constrained by noise and a high overlap in the feature space.

2 Brief description of methods

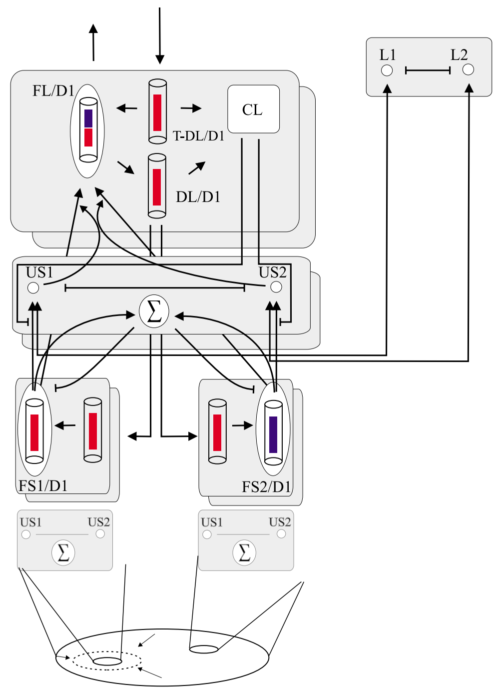

Figure 1: The model consists of four strongly interconnected functional blocks with interwinded bottom-up and top-down pathways. Feature-sensitive neurons (FS) of the topographically ordered units dynamically represent features by a population code. Each dimension has its own unit. They project their activity to neurons of higher complexity and larger receptive fields (FL). This simulates the general idea of bottom-up processing in the ventral pathway. Each stage reads out its features to filter and project back the strongest feature (DL). Thus, the feature-sensitive units (F) receive a top-down activity from the following stages or finally from short-term memory coding the features of a cue. Between two stages and within each dimension unspecific cells (US) modulate the projection by increasing the projection weight in dependence of their location and activation. They also send their activity to location-sensitive neurons (L) simulating one function of the dorsal stream or subcortical areas. Their firing rates contain all information needed for determining the location of an action. They unite the activity of the feature-sensitive neurons via the unspecific neurons from different dimensions, which code the bottom-up saliency and the task-driven, top-down importance. In fact, the whole processing is completely parallel, there is no sequential order in the blocks.

This contribution suggests a general principle how the bottom-up and top-down flow cooperate and how attention emerges within these pathways. Since recent results have demonstrated cooperative and nonlinear effects, 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.

In essence, the flow of activation and the role of attention within the receptive field is addressed. Figure 1 gives an overview of the proposed model containing two stages of growing receptive field size. Given a stimulus, each feature-sensitive unit F consists of several neurons n describing the population response u of the bottom up signal $u^{\uparrow}$, top-down signal $u^{\downarrow}$, lateral inter unit communication $u^{\circlearrowleft}$, and receptive field interactions $u^{\leftrightarrow}$:

$$\dot{u}(t) = p \cdot u^{\uparrow}(t) \times u^{\downarrow}(t) + r \cdot u^{\uparrow}(t) \times u^{\circlearrowleft}(t) + s \cdot u^{\uparrow}(t) - q \left(u^{\leftrightarrow}(t) + u(t) \cdot \sum u^{\circlearrowleft}(t) \right)$$

Thus, the bottom-up input pattern is continuously compared to the top-down target $u^{\downarrow}$ and lateral feature memory $u^{\circlearrowleft}$. The top-down pathway and the lateral inter unit connections act in a multiplicative fashion which enhances matching bottom-up patterns. They implement a Bayesian inference operation, i.e. an input pattern is compared with the prior or expected information encoded in the population [11]. An input pattern that does not match a top-down or lateral pattern results in a lower activation of the feature detector. The current implementation considers only competitive (inhibiting) connections within the receptive field $u^{\leftrightarrow}$. This simplification only applies for very sparse presentation scenes. Finally, depending on the intrinsic activation of each neuron, the activation is counterbalanced by the unit activity. The rationale behind this rule is a feature detector with the ability to superimpose different input patterns and simultaneously to perform a competition without erasing the minor pattern, as done by other models that read out a population code. One of those additive activation rule models [12] gains more importance for the top-down projection of a feature representation (see unit DL in Fig. 1). In comparison to the bottom-up feature pathway, the top-down expectation pathway must not convey superimposed features, but it has to urge a decision in receptive field competition by prescribing a target to the previous layer. Thus, unit DL reads out the strongest feature and projects this population code back to all locations within the receptive field of the current unit.

The unspecific neurons z calculate the computational weight of the features in each dimension Dl and each location i separately. Their activation determines the bottom-up projection from layer Lk in location i:

$$u_i^{\uparrow Lk,Dl}(t) = w_i^{\uparrow Lk,Dl}(z(t)) \cdot u_i^{Lk,Dl}(t) \qquad with \;\; w_i^{\uparrow Lk,Dl}(z(t)) = w^{\uparrow} \cdot (1 + z_i^{Lk,Dl}(t))$$

The unspecific neurons are reciprocally linked to location coding neurons, whose response is no further specific to a particular dimension.

This theoretical work is proven to be relevant by simulating the dynamics while performing an attended location task, the same as in the experiment of Luck et al. [2], and a guided feature selection task, the same as in the experiment of Motter [4], [6].

3 Results

3.1 Attended location task

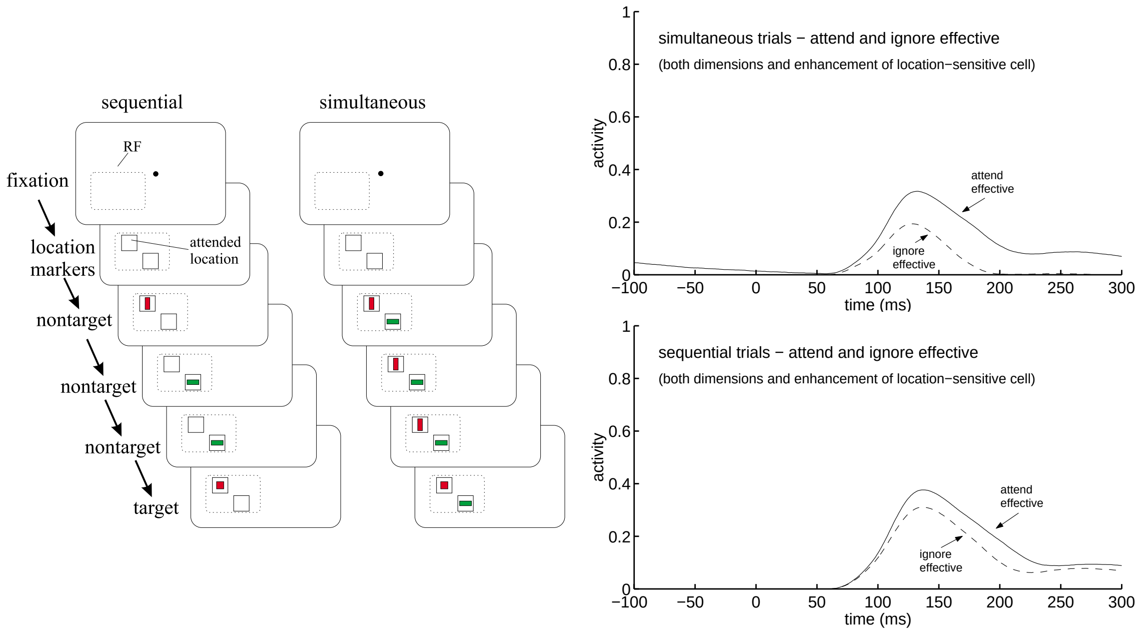

Figure 2: Simultaneous and sequential conditions in the experiment of Luck et al. [2]. *Left*: Monkeys were trained to attend a specific location within a receptive field of a recorded neuron in V4 while ignoring the other. A typical sequential trial consists of an attended-location nontarget, two ignored-location nontargets, and an attended location target. A simultaneous trial differs by presenting items at both locations simultaneously. The input pattern of the model is chosen to reflect the same condition as in the experiment. *Right*: The feature-sensitive cells with a large receptive field show a quantitatively similar behavior as the recordings of the neurons in V4 from the experiment of Luck et al. The peak activity in the simultaneous trials is lower than in the sequential trials. However, since the effect of attention is to reduce the suppression within the receptive field, the influence of attention is higher in the simultaneous trials, as indicated by the larger difference between ignore and attend conditions.

The attended location task is currently the most widespread experiment for neurophysiological recordings [3], [2], [5]. This task requires to attend a location while the activation of a cell, whose receptive field covers this location and beside it, being measured. Thus, the attentional capturing process has already been done. It is reasonable to assume that attending to one location is reflected by the activation of corresponding location-sensitive cells of the model, although attention in the presented model is more than a simple competition within one layer. As can be seen in the guided feature selection task attentional effects are possible without any activation of the location-sensitive cells. But they integrate activity from all stages and different dimensions to determine a location for action. Thus, The simulation was chosen to be alike the experimental setup of Luck et al. [2] (fig. 2). The average activity of the neuron with the best match among the feature-sensitive cells with a large receptive field (FL) closely resembles the recordings within the experiment.

3.2 Guided feature selection task

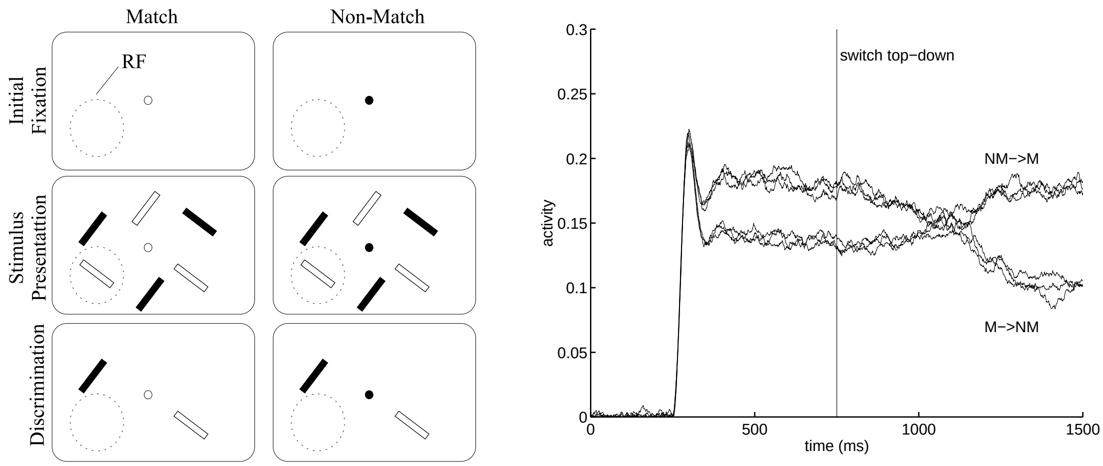

Figure 3: Match and non-match conditions in the experiments of Motter [4], [6]. *Left*: In either task a monkey is initially faced with a fixation spot, whose color also serves as a cue. The conditional discrimination task can be performed after the presented array of oriented colored bars is reduced to two bars. During the stimulus presentation period, several items could be the final target. In the match condition a color or luminance match occurs between the fixation spot and the item in the receptive field. In another experiment a cue switch paradigm is used to investigate the behavior if the target changes during the stimulus presentation phase. Stimuli previously represented potential targets change into distractors and vice versa. *Right*: The simulation is performed with three potential targets and three potential distractors. After the scene is presented the activity of all neurons coding the target and the distractor raises. Due to top-down activation of one feature (e.g. color), the neurons representing the target remain stronger activated. After changing the top-down activation, the attentional system switches in order to bias the current potential targets.

The guided feature selection task was mostly discussed in terms of a yet unknown attentional effect, I suppose, due to the unimportant locational information – the dominant selection criterion in most models. Therefore, reliable models have to simulate both, attended location and feature selection tasks.

The simulation is in line with the experiment of Motter (see fig. 3) [4], [6]. The results closely resemble the average temporal development of the activation in the match (M) and non-match (NM) condition during the stimulus presentation phase of the experiment. After the cue switches, a delayed response occurs due to an activity change, similar as the activity of the feature-sensitive cells (FS) in the model. Thus, the simulation demonstrates a possible explanation of the measured parallel bias and the switch of the bias in case of a changing top-down activation.

4 Conclusion

Several models have been proposed to account for distinct attentional effects. These assumptions are now constrained by neurophysiological recordings which give a larger insight into the underlying mechanisms. The model presented here, bears on this. The performed simulations quantitatively resemble the results, even the recorded dynamics, of currently fundamental experiments, for instance [2], [4], [6]. Most models assume an exclusive stage in which selection occurs – that is a separation between parallel and serial processing – whereas this model suggests that attention might be better interpreted as a graded process. Selection occurs when a constraint demands this. There is no unitary part in the model that accounts for attention, but competition in different parts converge to let the system operate on the same event, which resembles in several aspects the biased competition [13] and the integrated competition [14] theory.

On a long-term basis models incorporating feedback should outperform the dominant, but for real visual scenes weak feedforward paradigm in object recognition. Directed attention mediates this recognition process and abolishes the overestimated separation between segmentation and recognition.

References

[1] Motter B. C. (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. Journal of Neurophysiology 70, pp. 909-919.

[2] Luck S. J., Chelazzi L., Hillyard S. A., Desimone R. (1997) Mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex. Journal of Neurophysiology 77, pp. 24-42.

[3] Moran J., Desimone R. (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229, pp. 782-784.

[4] Motter B.C. (1994) Neural correlates of attentive selection for color or luminance in extrastriate area V4. Journal of Neuroscience 14, pp. 2178-2189.

[5] Reynolds J. H., Chelazzi L., Desimone R. (1999) Competetive mechanism subserve attention in macaque areas V2 and V4. Journal of Neuroscience 19, pp. 1736-1753.

[6] Motter B. C. (1994) Neural correlates of feature selective memory and pop-out in extrastriate area V4. Journal of Neuroscience 14, pp. 2190-2199.

[7] Usher M., Niebur E. (1996) Modeling the temporal dynamics of IT neurons in visual search: A mechanism for top-down selective attention. Journal of Cognitive Neuroscience 8, pp. 311-327.

[8] Niebur E., Koch C. (1994) A model for the neuronal implementation of selective visual attention based on temporal correlation among neurons. Journal of Computational Neuroscience 1, pp. 141-158.

[9] Grossberg S., Raizada R. (2000) Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. Vision Research 40, pp. 1413 - 1432.

[10] Hamker F. H. (1999) The role of feedback connections in task-driven visual search. In: Connectionist Models in Cognitive Neuroscience. London: Springer Verlag, pp. 252-261.

[11] 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, pp. 25-44.

[12] Amari S., Arbib M. A. (1977) Competition and cooperation in neural nets. In: Systems Neuroscience. Academic Press, San Diego, pp. 119-165.

[13] Desimone R., Duncan J. (1995) Neural mechanisms of selective attention. Anu. Rev. of Neurosc. 18, pp. 193-222.

[14] Duncan J., Humphreys G. W., Ward R. (1997) Competitive brain activity in visual attention. Current Opinion in Neurobiology 7, pp. 255-261.