

Scale Independence in the Visual System

Raul C. Mureşan

Nivis Research, Gh. Bilaşcu Nr. 85, 3400 Cluj-Napoca, Romania

Abstract. We briefly present some aspects of information processing in the mammalian visual system. The chapter focuses on the problem of scale-independent object recognition. We provide a simple model, based on spiking neurons that make use of shunting inhibition in order to optimally select their driving afferent inputs. The model is able to resist to some degree to scale changes of the stimulus. We discuss possible mechanisms that the brain could use to achieve invariant object recognition and correlate our model with biophysical evidence.

Keywords. Object recognition, scale independence, spiking neurons, shunting inhibition, visual system

1. Introduction

The visual system performs continuously amazing tasks such as object recognition and it seems to achieve these tasks effortlessly. The last two decades of research revealed some of the mechanisms involved, both at cellular and microcircuit structure level. Numerous studies have been performed on the retina, optic nerve and the first levels of visual processing, such as the primary visual cortex (V1) [8]. However, to this date, we have no complete or correct understanding of the complex processes that take place at higher areas in the brain such as area V4 or the infero-temporal cortex. It is not yet clear how the brain is capable of achieving attentional and recognition tasks.

Many hypotheses have been formulated trying to explain the way our brain can recognize objects. One such theory is based on the RBC (Recognition By Components) model developed in 1987 by Biederman [1]. The RBC model constructs intermediate primitives from primary features. These primitives have a finite number, complex objects being later constructed and described in terms of these primitive decompositions. Attractive as it may be, this theory is still far from solving the problems. It has been thought that the primitive based recognition would eliminate the combinatorial problem of hierarchic recognition. However, it is quite obvious that the RBC model only transfers the problem to higher areas that have to combine primitives. Primitive combination still suffers from the combinatorial explosion for real-world complex visual scenes. Also, it is unclear how

the visual system might achieve primitive decomposition. One of the hardest problems the visual system has to solve is scale-independent recognition. Should it be primitive extraction or direct recognition, scale-independence must still be achieved.

The temporal correlation hypothesis does not constrain the recognition process. It provides some basic mechanisms in order to solve the "binding problem" formulated by Malsburg and Milner [13,14,15,16,17]. The correlation hypothesis states that synchronization and correlated activity might signal aspects of the visual world that "belong together". Assemblies of such synchronized neurons emerge dynamically leading eventually to perceptual grouping and neural routing to achieve recognition. However, the hypothesis has been widely criticized and the main argument for opposing it has been the difficulty of replicating the experiments that favored it [7].

There are several problems that have no straightforward solution, each approach to solving them proving that the method quickly becomes intractable. Probably the most striking feature of our visual system is the way it combines information. Ultimately we are reduced to a context problem. Simple cells in the lower areas of the visual cortex, such as V1, have small receptive fields. They integrate information from a limited, small area of the visual scene. The problem is how successive layers in the processing hierarchy integrate these separate, small spatial contexts. Combining different feature detectors could rapidly lead to a combinatorial explosion (the combinatorial problem). Riesenhuber and Poggio [25] created a model (HMAX) based on Fukushima's "neocognitron" [6]. These models combined simple feature detector responses into more and more complex representations. The HMAX model combined the features according to a maximum response function, selecting optimal responses for later stages. HMAX is in fact a winner-take-all model that segments the visual scene into approximated representations that are later combined. To some degree, such a hierarchic approach resembles the RBC model. Olshausen et al. provided attention based dynamical routing models in order to achieve spatial context integration [23]. However, all these models are only partial solutions. The HMAX can hardly capture scale variance while Olshausen's model of dynamic routing does not clearly provide an answer for preattentive visual search.

The visual system is a very powerful processing unit that is able to solve all of the visual tasks with limited resources. One of the most striking aspects of our visual processing is its speed. We are able to segment the visual scene in a few tens of milliseconds, to modulate attention and achieve recognition under 150 milliseconds [28]. All these findings impose even greater constraints on models. The speed of visual processing seriously questions rate-based coding of information in neural systems. Temporal coding might be, in this context, a much better candidate [29].

2. Visual scale independence

One very little studied aspect of visual processing is scale independence. Resizing an object leads to big variations in the local organization of the stimulus. Locally speaking, like translation and rotation, scaling effectively changes "where" features will be located. Unlike translation and rotation, scaling also changes "what" local detectors perceive. The same object, at different scales, generates a whole different pattern of stimulation at the lower stages of visual processing (Fig. 1). We might say that only the global "characteristic" of the stimulus is preserved.

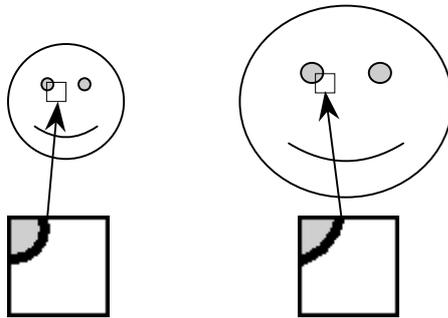


Fig. 1. Local feature variation due to stimulus scaling

Recognizing objects in a scale independent way is a very hard problem. The visual processor has to solve both "what" and "where" problems. There are at least two possible approaches to invariant recognition:

1. the creation of invariant representations;
2. the inference of stimulus characteristics.

Although the second approach is more likely to be used by our visual system, we do not understand yet how it is able to "infer" stimulus characteristics. Accepting that inference processes take place in the visual system would radically change our understanding of visual processing. Most of the present models ignore feedback projections and lateral interactions in visual areas. However, these processing pathways are at least equally important as the feed-forward ones. Quantitatively we have evidence of strong feedback and lateral interactions. It just might be that large, distributed, recurrent and dynamical neural systems can achieve inference. Object representations, in this context, would not be clearly separable. It is much likely that objects and relationships between them generate "mental states". The "liquid state machine" and "far from equilibrium" systems can offer nontrivial processing capabilities, while largely respecting the wiring rules of our cortical microcircuits [12]. Also, the "coherence theory", under development, could offer a new perspective on visual processing and object recognition based on the belief that visual recognition is an inference process [22].

In this chapter, we will focus mainly on the first approach, the creation of invariant representations. To some degree, invariant representations are biologically plausible and are easier to understand. We will provide neural mechanisms that could account for scale invariant recognition. Also, we will describe a model, based on spiking neurons that is fairly biologically plausible and reproduces some known experimental evidence. The neural coding paradigm applied is rank-order coding, while shunting inhibition as an underlying mechanism is extensively used.

3. Neural mechanisms

Studies on the mammalian visual cortex had shown that cortical simple cells might project to complex and hipercomplex cells. Since Hubel and Wiesel [8] it has been thought that such a hierarchical composition might lead to transformation invariant recognition. On the other hand, the speed of processing in the visual system [28] leads us to the conclusion that some feed-forward mechanisms should exist that could achieve, at least to some degree of precision, object categorization and recognition. The speed problem imposes a strong constraint on the neural code that visual neurons might use. Rank-order coding and shunting inhibition could be the answer. Next, we will present the neural mechanism that is able to encode a huge amount of information in a very short time.

Evidence have been accumulated that neurons are able to select their synaptic inputs on the basis of afferent timing. Inhibition could be an important mechanism that contributes to timing sensitivity. Some researchers treat inhibition as having a stabilizing function [26,27]. Although this is probably an important function of inhibition, it can also generate shunting and hyperpolarizing effects that shape the temporal structure of neural activity on a millisecond time scale [2,3,5,11]. Shunting inhibition in general and fast shunting inhibition in particular could offer a strong mechanism for temporal coding rather than firing-rate coding.

The primary question behind all neural systems is how they encode information. At this time, there is widely accepted that neurons encode information by the means of their firing rate. In this framework, we speak about stability and fixed points in the oscillatory processes. However, numerous recent studies seriously question such a neural code because it is unable to account for ultra-rapid brain processes. It is true however that the firing rate of a neuron is directly linked to the stimulus and excitatory as well as inhibitory afferent activity. But we can understand firing rate as a side effect of millisecond time scale processes.

Thorpe and colleagues [29] have proposed a new scheme for information coding in the cortex: rank-order coding. Rank-order coding is based on shunting inhibition to shape the response of a target neuron as a function of afferent spike timing. In principle, considering a target neuron and a pool of inhibitory interneurons,

each time an afferent spike arrives, the inhibitory pool shunts down the efficacy of spike integration of the target neuron by the means of shunting inhibition. Simple implementations of such a neural coding proved their power in dealing with large amounts of information [4,18,19,21]. The enormous encoding power of this scheme has been outlined by S.J. Thorpe, showing that rank order codes can be used, in principle, to transmit up to $\log_2(N!)$ bits of information (where N is the number of neurons).

The rank order code scheme presented above has two main advantages. First, due to progressive shunting effects, a normalization process is automatically generated. As afferent activity increases, the shunting effect is also increased and the target neuron desensitized so that its response will be rectified. Secondly, because each time an afferent neuron spikes, the target neuron is desensitized, the response of the target neuron can be made dependent on the timing of afferent spikes (using different combinations of synaptic strengths) [29].

4. A scale invariant recognition model

Next, we will present a simple model, based on a biologically plausible framework that can construct scale invariant (to some degree) representations of objects. The model uses spiking neurons as processing units and has a retinotopic hierarchic organization. A pathway is constructed from the retina, to V1 simple cells and "end-stopped" bar detectors, up to V4 and infero-temporal cortex. Depending on the available computing resources (number of units) used, the model is able to construct scale invariant representations of complex objects with different degrees of precision.

4.1. Building blocks of the model

Spiking integrate-and-fire neurons were used in the recognition system. We have applied a similar model as Thorpe and colleagues for the individual neurons used in simulations. Each neuron is characterized by a small set of parameters:

- membrane potential (U);
- instantaneous sensitivity (S);
- synaptic modulation factor (M).

The update rule adds the current post-synaptic potential (PSP) to the current membrane potential. The PSP is computed by multiplying the synaptic weight (W) by the sensitivity factor (S). After each afferent spike the sensitivity factor is decreased by multiplying it with the synaptic modulation ($M \in (0..1)$). For simplicity, no refractory period or leakage are included in the model.

$$U^{(t+1)} = U^{(t)} + W \cdot S^{(t)} \quad (1)$$

$$S^{(t+1)} = S^{(t)} \cdot M \quad (2)$$

Neurons are organized in maps. Each map has a given, well-determined functional role. Maps are interconnected with each other using synaptic kernels. Because we used retinotopic mapping, each pair of neural maps that are interconnected has exactly one connection kernel associated. This connection kernel is used as a "rule" of interconnection between maps. Each time spikes are being processed, the appropriate connection kernel is selected and used in computing the PSP's. Simulation is iterative and event-driven at the same time. In each iteration a set of spikes that are emitted is computed for every map. In the next iteration, these spikes are integrated according to the position of the spike events. Only those neurons that are affected by the previous spikes are updated. The update is spike-event driven. The simulator we constructed is called RetinotopicNET and is able to simulate large populations of spiking simple neurons [20].

4.2. The architecture of the model

The architecture of the model contains 7 levels of processing, following the retinal, V1, V4 pathway up to the infero-temporal cortex. The seven layers of processing correspond to an ascending feed forward processing with lateral interactions at some levels (Fig. 3). The key feature of the model is the use of extensive competition between different elements of the objects to be recognized. The only information used at this time is contour information but blob-type cells could also be included to account for color or intensity patches as well.

4.2.1. Retinal processing

At the first level of processing the retinal ganglion cells process the incoming image intensities (only intensity 8 bit grayscale images were used). The ON-OFF effect has been achieved by using a classical difference-of-gaussians (DOG), center-ON-surround-OFF and vice versa filter with a ratio of standard deviations: 1 to 3 (eq. 3). Then, the image intensity for the two maps has been converted into spike latency and spikes were fed into the "RetinotopicNET" simulator.

$$DOG(x, y) = \frac{e^{-\frac{x^2+y^2}{2\cdot\sigma_1^2}}}{\sqrt{2\cdot\pi}\cdot\sigma_1} - norm \cdot \frac{e^{-\frac{x^2+y^2}{2\cdot\sigma_2^2}}}{\sqrt{2\cdot\pi}\cdot\sigma_2} \quad (3)$$

where: $-x, y$ are the coordinates in the kernel space

- σ_1 and σ_2 are the standard deviations ($\sigma_2 = 3 \cdot \sigma_1$)
- $norm$ is a normalization constant

Spike latencies are subsequently computed using a linear mapping. The pixel with the highest contrast (highest intensity after applying DOG filtering) generates the first spike (zero latency). The lowest contrast (0) generates the last spike. Intermediate latencies are computed using a linear approximation from the contrast level.

4.2.2. V1 Area

The second layer of processing corresponds to the V1 primary cortex area where different orientation channels are selected by oriented Gabor-like receptive fields (eq. 4). These are the corresponding simple cells, which detect different orientation contrasts.

$$G(x, y) = e^{-\frac{(x^2+y^2)}{2\sigma^2}} \cdot \sin(2\pi f \cdot (x \cdot \cos(\phi) + y \cdot \sin(\phi))) \quad (4)$$

- where:
- x and y are the coordinates in the kernel space
 - $G(x, y)$ represents the Gabor filter kernel value
 - f is the spatial frequency of the filter
 - ϕ the angle that influences the orientation selectivity of the filter

One key feature is the lateral connection within each orientation map. We have used a butterfly-like lateral connection, which has the property of improving contours. This is a form of primitive contour-integration, but due to the lack of iterative loops only a feed forward contour completion is used. Important work on this topic had been conducted by Zhaoping Li [9]. Further improvement on the system may be achieved by implementing a stronger contour integration mechanism. The Gabor patches were all at the same scale and had a spatial frequency of 0.5 pixels. They covered the range of 0 to 180 degrees with a total of 8 orientations.

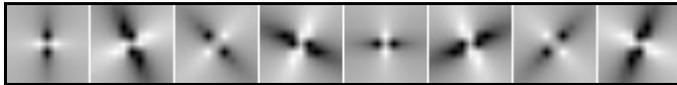


Fig. 2. Lateral "butterfly-like" receptive field profiles. White denotes excitatory areas and black inhibitory ones.

Approximately 2 to 5% of V1 neurons responded to illusory contours. While the contour integration effect is well known to occur in the V2 area, V1 neurons are also known to have some degree of contribution to the integration of contours [10]. Since we haven't modeled the V2 area, our simple model of V1 cells includes also a restricted contour-integration function. Our observations showed that

the performance of the model had improved by including lateral interactions between the V1 neurons.

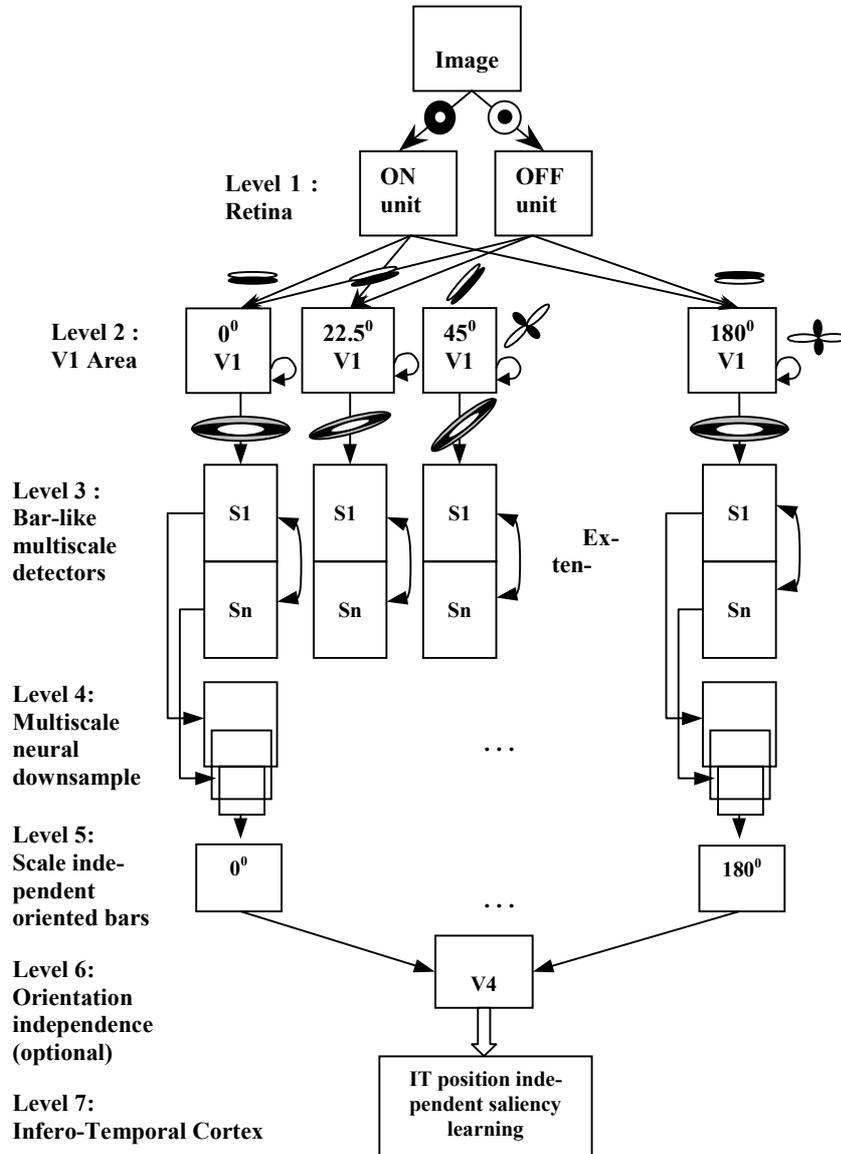


Fig. 3. Model architecture.

4.2.3. Bar-like detectors

For each orientation, a corresponding set of different scaled-receptive-field maps was used to extract the bar-like feature at the corresponding position. Bar like detection, across different length channels, provides the solution to the "what" problem. It is in fact a classification and measurement at the same time.

Each receptive field had an oriented bar-like, end-stopping type. The central, elongated bar, corresponds to excitation. The surrounding area corresponds to weak inhibition proportional to the level of blackness (Fig. 4). This type of receptive field tunes the neuron to the bar that best matches its excitatory size. For the same orientation, multiple scaled bar-detectors were used.

Considering one set of oriented scaled maps (multiscale maps for the same orientation), lateral inhibition has been introduced from large to small sizes, generating a size competition at that orientation. The priority in terms of timing varied from large to small bars. In other words, the maps with a large receptive field had the chance of firing first and, by the means of inter-map lateral inhibition, the smaller bar detectors were inhibited. Such a mechanism ensures that the largest size possible is always detected (instead of composing it from multiple smaller size bars).



Fig. 4. Bar-like detector for 0° orientation. The white bar in the center corresponds to excitatory synapses; the gray and black areas correspond to inhibitory synapses. The black denotes the strongest inhibition, the gray denotes intermediate strength of inhibition.

4.2.4. Multiscale downsample

The fourth layer of processing is responsible for bringing every detail to the same level of spatial importance. We deal here with the "where" problem. Once the constructing elements were detected (bars), we have to create invariance to distance changes between these features that occur due to scaling. In other words, if two long lines are detected, the distance between them has to be brought down to the same distance as the one between the scaled down versions of the same lines (in an object scaling operation). The key mechanism of scale independence is exactly the equivalence of feature distance with feature size. This level is the most important one for scale independence and we will describe the mechanism in detail.

Let us consider, as an example, a simple object, formed of just two lines, oriented at 0^0 , as shown in figure 5.

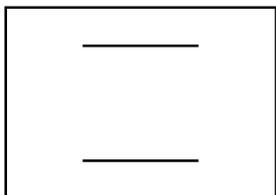


Fig. 5. An object formed of 2 lines, of size 20 pixels each.

For explanation purpose, let us consider that our scaled bar detectors range over 10 to 30 neurons (pixels) and that the lines in the original image (Fig. 5) have a size of 20 pixels. We have 3 maps of 0^0 orientation with bar detectors at 10, 20 and 30 neurons (pixels). The response of the 3 maps to the original image is presented in Fig. 6.

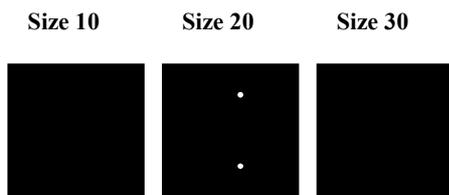


Fig. 6. The response of the 3 maps to the original object with lines of size 20.

Because of the strong lateral inhibition and competition, neurons in the map of size 10 can't fire. Neurons in the map of size 30 have not enough stimulation in the excitatory area to be driven by the 2 lines. Thus, only in the map of size 20 the activity will exist. Now let us scale down by a factor of two the original image (Fig. 7).

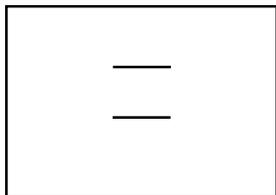


Fig. 7. The original object scaled down by a factor of 2.

The result of the down scaling is that activity will move down to lower sized detectors by a distance proportional to the ratio between the size of different detectors of the maps (Fig. 7).

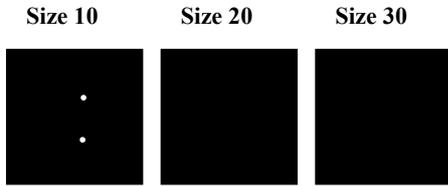


Fig. 8. The response of the 3 maps to the scaled object with lines of size 10.

We take a look at the distance between the two lines: by scaling the object down, the distance between its parts is also scaled down yielding a cortical response with scaled distances between bar detector neurons. All the system has to do, in this simple case, is to scale down the second map by a factor of 2 and the third by a factor of 3 and feed all of these resulting maps into a scale invariant map (Fig. 9).

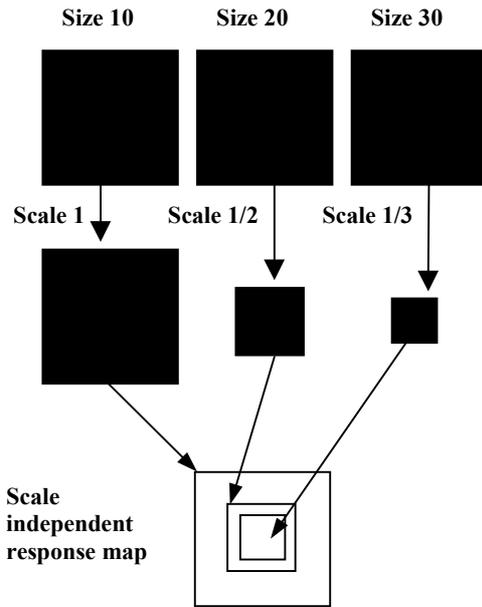


Fig. 9. Neural downsampling and combination to achieve scale independence.

Some may argue that for an object that has lines which will be detected by different sized detectors at the equivalent locations (as dictated by the scale of detection), the neural responses will overlap each other in the final scale invariant map. But we have to take into account that for that specific location, the final neuron will have a multiplied feeding and thus an increased firing rate (the multiplication

depends on the number of responses that overlap). The firing rate captures the equivalence information and the location captures the relative position of features.

The neural down sampling is achieved by using window-like receptive fields which could be associated with center-ON-surround-OFF receptive fields in area V4, taking into account the fact that the surround-OFF is a very small silent inhibition which could be used for stability and normalization purposes.

4.2.5. Scale independence maps

At the next layer, at each orientation, the downsampled oriented maps are combined into a scale independent map corresponding to that specific orientation (Fig. 9). The mechanism of combining them should take into account the scaling center of the object. Further improvement, like position independence of features, could be implemented at this level.

4.2.6. Orientation independence

This layer is optional, and is used only for reducing the number of synapses with the infero-temporal map. It corresponds to the final stage in the V4 area. There is no reason for which one might consider different orientations as being equivalent but this type of combination can be used to simulate the hypercomplex cells. Orientation equivalence can be used to improve the generalization capability of the recognition in the IT cortex.

4.2.7. Infero-Temporal Cortex

The infero-temporal cortex is responsible for object recognition. In the architecture presented, learning is performed by increasing the synapse strength with the current sensitivity value as resulted from successive modulator effects generated by the firings in the level 6 map. Every neuron in the final IT map has a retinotopic type of receptive field, covering most of the level 6 map. The synaptic strengths are shared among all neurons, yielding good position independence.

4.2.8. Simulation results

Using the "RetinotopicNET" simulator we calibrated the system for face detection (and recognition). The test database had been generated using a QuickCam web

camera and consisted of the faces of three different persons with different face expressions. Image sizes were fixed at 92 x 112 grayscale bitmaps and the faces were scaled in a range 1 to 0.58 the original scale (Fig. 10).

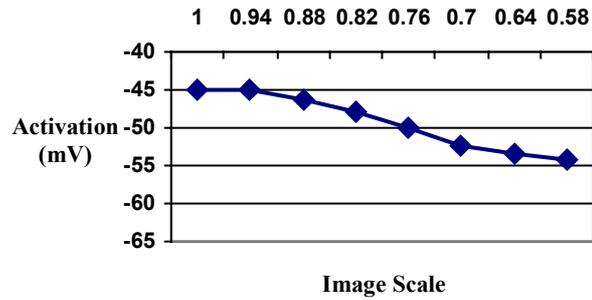


Fig. 10. Activation of infero-temporal cortex for different image scales. The IT layer has been trained to reach the exact threshold activation of -45 mV for the face at the original size.

The number of scales used at level 4 was 7, bar-like detectors ranging from a length of 7 to 13 neurons (7,8,...13). At the infero-temporal level, a strong shunting inhibition had been used to provide enhanced selectivity on learning (for the face recognition case).

The selectivity map of the trained infero-temporal neuron is shown in figure 11 and different sized details can be observed at different positions.



Fig. 11. Training image and the selectivity of the infero-temporal cortex

The given results are surprisingly good, taking into account the limited scale levels used and the size of the details in the image (the details are by far larger than the detectors used). Increasing the size of the bar-detectors and their number (to cover a wider field of scales) can increase the accuracy of recognition. At the same time, we expect that the usage of more orientations can increase accuracy because of the better localized bar detection at the third level of the model.

5. Discussion

We have seen that neural mechanisms like shunting inhibition can contribute to spatial competition processes between different feature detector neurons. However, the model presented in this chapter is clearly not very plausible biologically. The first problem this model encounters is the great redundancy that is generated by the requirement for specialized detector cells. There are many types of features that the visual cortex extracts from the incoming visual signals. Thus, having a good approximation of this signal (in this model) requires an enormous amount of specialized feature-detection cells. Such enormous redundancy and waste of resources is unlikely to be the case in our visual system. It is supposed that many cells change their selectivity from broad to sharp tuning on millisecond time scale. It might be the case that the visual system adapts the feature tuning of neurons dynamically, under coherent attentional feedback supervision. Such adaptation is much likely to occur in the V4 area where complex feature detectors and encoders might exist [24].

The second problem of our model is that at the fifth layer, the neural downsampling should be performed around the center of the object. Such a mechanism would be possible only with attentional involvement. It is not clear however, how attention could modulate the responses and reroute the information flow in order to achieve object based attention as a step in object recognition processes. Nonetheless we have evidence that attention is very important, maybe essential, also in the object recognition process.

Probably the biggest pitfall of the described mechanisms is that they are too algorithmic. Our brain continuously processes information, dynamically reallocating resources and having parallel, multisensorial, integrative capabilities. Algorithmic approaches are good for modeling but unlikely to be used by our visual system.

Although many aspects stand against such simplified models of vision, optimality selection, competition and limited-context scale independence might actually be based on similar mechanisms. Ultimately, such simple analyzers might dynamically arise in a huge hierarchical and recurrent neural system. Pairs of content-context pools might be generated during adaptation to stimulus. Attentional modulation could play an important role, sometimes mixed with recurrent feedback projections that might contribute not only to stability but also to inference processes [22]. The presence of a particular object and the relationship to the overall scene would be captured by the transitions of the dynamical system (as mental states), yielding a characteristics-based recognition capability. Mechanisms similar to our model could be, at least in part, involved in such a complex dynamical picture of visual processing.

As a final statement, we have to mention our belief that vision is not about filtering or matching, *vision is about thinking*.

6. References

1. Biederman I (1987) Recognition by components: A theory of human image understanding. *Psychol. Rev.* 94:115–147.
2. Bush P, Sejnowski TJ (1996) Inhibition synchronizes sparsely connected cortical neurons within and between columns in realistic network models. *J Comput Neurosci* 3: 91–110.
3. Chrobak JJ, Buzsaki G (1998) Gamma oscillations in the entorhinal cortex of the freely behaving rat. *J Neurosci* 18:388–398.
4. Delorme A, Thorpe SJ (2001) Face identification using one spike per neuron: resistance to image degradations. *Neural Networks* 14:795–803.
5. Freeman WJ (1975) *Mass Action in the Nervous System*. Academic Press, New York.
6. Fukushima K (1980) Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern* 36:193–202.
7. Gray CM (1999) The Temporal Correlation Hypothesis of Visual Feature Integration: Still Alive and Well. *Neuron* 24:31–47.
8. Hubel D, Wiesel T (1965) Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J Neurophysiol* 28:229–289.
9. Li Z (1998) A neural model of contour integration in the primary visual cortex. *Neural Comput* 10(4):903–40.
10. Li Z (1999) Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network: Comput Neural Syst* 10:187–212.
11. Lytton WW, Sejnowski TJ (1991) Simulations of cortical pyramidal neurons synchronized by inhibitory interneurons. *J Neurophysiol* 66:1059–1079.
12. Maass W, Legenstein RA, Markram H (2002) A new approach towards vision suggested by biologically realistic neural microcircuit models. In Buelthoff HH, Lee SW, Poggio TA, Wallraven C (eds) *Biologically Motivated Computer Vision*. Proc of the Second International Workshop BMCV 2002 Tübingen, Germany, vol. 2525 of *Lecture Notes in Computer Science*, Springer, Berlin, pp 282–293.
13. von der Malsburg C (1981) The correlation theory of brain function. *MPI Biophysical Chemistry Internal Report* 81–2.
14. von der Malsburg C (1985) Nervous structures with dynamical links. *Ber Bunsenges Phys Chem* 89:703–710.
15. von der Malsburg C (1986) Am I thinking assemblies? In: Palm G, Aertsen A (eds) *Proceedings of the Trieste Meeting on Brain Theory*. Springer, Berlin.
16. von der Malsburg C (1999) The What and Why of Binding: The Modeler's Perspective. *Neuron* 24:95–104.
17. Milner P (1974) A model for visual shape recognition. *Psychol Rev* 81:521–535.
18. Mureşan RC (2002) Complex Object Recognition Using a Biologically Plausible Neural Model. In: Mastorakis NE (eds) *Advances in Simulation, Systems Theory and Systems Engineering*. WSEAS Press, Athens, pp 163–168.
19. Mureşan RC (2002) Visual Scale Independence in a Network of Spiking Neurons. *ICONIP '02 Proceedings*, Singapore, 4:1739–1743.
20. Mureşan RC (2003) RetinotopicNET: An Efficient Simulator for Retinotopic Visual Architectures. *ESANN'03 Proceedings*, Bruges, pp 247–254.

21. Mureşan RC (2003) Pattern Recognition Using Pulse-Coupled Neural Networks and Discrete Fourier Transforms. *Neurocomputing* 51C:487-493.
22. Mureşan RC (2003) The Coherence Theory: Simple Attentional Modulation Effects. *CNS'03 Proceeding*, in press.
23. Olshausen BA, Anderson CH, van Essen DC (1993) A Neurobiological Model of Visual Attention and Invariant Pattern Recognition Based on Dynamic Routing of Information. *J Neurosci* 13(11):4700-4719.
24. Pasupathy Anitha, Connor CE (1999) Responses to Contour Features in Macaque Area V4. *J Neurophysiol* 82: 2490–2502.
25. Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2(11):1019-1025.
26. Shadlen MN, Newsome WT (1994) Noise, neural codes and cortical organization. *Curr Opin Neurobiol* 4:569–579.
27. Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18:3870–3896.
28. Thorpe SJ, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381(6582):520-522.
29. Thorpe SJ, Gautrais J (1998) Rank order coding. In: Bower J (eds) *Computational neuroscience: Trends in research*. Plenum Press, New York, pp 113-118.