Two separate processing streams in a cortical-type architecture

Tobias Rodemann, Edgar Körner

2001

Preprint:

This is an accepted article published in Computational Neuroscience: Trends in Research 2001. The final authenticated version is available online at: https://doi.org/[DOI not available]

Published in: James. M. Bower, editor, Computational Neuroscience: Trends in Research, 2001, volume 38-40, pages 1541-1547, Amsterdam, Elsevier Science

Two separate processing streams in a cortical-type architecture

Tobias Rodemann [1] and Edgar Körner

Future Technology Research
Honda R&D Europe (Deutschland)
Carl-Legien-Strasse 30
63073 Offenbach (Main) / Germany

Abstract

We present a model of a cortical-type architecture that uses a latency coding scheme to process inputs as fast as possible given certain neuronal constraints while at the same time keeping the potential to perform a detailed analysis. The basic processing elements of our network are columnar structures, called minicolumns. We propose that these minicolumns are the prime constituents of cortical architecture.

Key words: Hierarchical Network, Latency Coding, Cortical Architecture

1 Introduction

Information processing in the mammalian cortex must satisfy two opposing objectives: to extract the behaviorally most important features in minimum time and to provide a detailed input description for fine analysis and learning. Furthermore, the system has to function within the framework of cortical architecture: a processing hierarchy comprising spiking neurons as basic elements. We propose a model of a hierarchical processing system of spiking neurons that fulfills these requirements.

[1] Corresponding author. EMail: Tobias.Rodemann@hre-ftr.f.rd.honda.co.jp

2 Latency encoding for fast processing

It was noted in [9] that visual information can be analyzed extremely fast (within about 100–150 ms). Considering cortical hierarchy, a coarse recognition of sensory input must be possible based on the first elicited afferent spikes with about 10 ms for the information exchange between two network elements. Due to the low firing rates of cortical neurons (far less than 100 Hz [1]) only one spike per neuron has to transmit the information, see also [8]. We propose that the most efficient coding scheme within these constraints is a latency encoding [2,4], where the reliability of a local decision is transmitted in the time of the spike. A prime requisite of such a coding strategy is a temporal reference frame that is accessible to both sending and receiving neuron. It is often implicitly assumed that the first, strong input transient provides such a clocking signal. However, this signal is available only once for each input stimulus. If the need arises to analyze the same, quasi-stationary input again a new reference signal is needed. Furthermore, to decode the latency signal, the reference frame must be explicitly accessible for every coding element. We propose that the input triggers a rhythmic oscillation that provides a global, centralized clocking signal to all cortical areas. We identify the ILN (intra laminar nuclei) as the source of these oscillations [3,6]. These structures generate gamma (40 Hz) frequency oscillations and have afferents to all cortical areas. We model the influence of the ILN by generating a rhythmic, sine-like modulation of the membrane potential (see [4]). Initially, the ILN are triggered by a new sensory stimulus. The phase of the modulation varies between different processing levels and the minimum of ILN modulation roughly matches the arrival time of the first feedforward input spikes.

3 Minicolumns

The requirement to finish the first, coarse analysis of the input in about 150 ms not only necessitates a one-spike-only coding strategy. It is furthermore obvious that there is not much time to go through several processing iterations using feedback. It must be possible to extract the behaviorally relevant features based on a single feedforward-only processing cycle. There should be a specialized processing system that is optimized to provide a coarse but fast initial categorization of the input based on those spikes with minimum latency (i.e. highest reliability) only. This type of analysis can't provide a detailed analysis, which would involve iterated processing using feedback [5]. We therefore conclude that there must be a second processing system that is able to include the details (i.e. the later arriving spikes) and context information (feedback) in its analysis. We propose that both processing streams are combined in one columnar unit that acts as the basic element of corti-

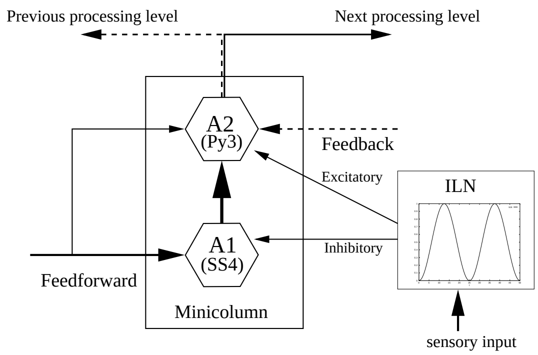

Fig. 1. The minicolumn consists of two subsystems, $A1$ and $A2$, both of which receive feedforward input. Subsystem $A2$ also receives feedback. Spikes from $A1$ are relayed to other processing elements via the $A2$ system. Both subsystems are modeled as spiking neurons and receive inputs from the ILN.

cal architecture – the minicolumn. Minicolumns (see Fig. 1) receive and send both feedforward and feedback signals. Within each minicolumn there are two subsystems, the feedforward categorizer $A1$ and the refinement system $A2$. Subsystem $A1$ analyses only the strongest (fastest) feedforward signals and only triggers a spike if this input is sufficiently strong. A maximum of one spike with minimum latency is transmitted. Subsystem $A2$ also receives feedforward input but takes all spikes into account. Furthermore, feedback is also integrated (the effect is additive). $A2$ uses the full processing period (half of the gamma period, see [5]) and might need several processing cycles (gamma phases) with varying feedback to reach a final decision. Input analysis is performed in several iterations, corresponding to consecutive upbeat phases of gamma modulation. In every iteration, the input passes all processing levels, triggering local hypotheses at each level. These are fed back as prediction to preceding levels to improve analysis in subsequent iterations. Spiny stellate cells of layer IV ($SS4$) match the requirements (strong feedforward input, no feedback, lateral inhibition) for the first, rapid categorization system. We consider pyramidal cells from lower layer III (Py3) to be the first stage of the second, bidirectional refinement system.

4 Model architecture

We implemented a simplified version of the model described in [5] with the two processing streams for the special purpose of a vision task. The general network structure is a hierarchy of six processing levels (see Fig. 2), each consisting of an array of minicolumns. Minicolumns (see Fig. 1) are modeled as two spiking neurons corresponding to $A1$ and $A2$. We employ standard I&F model neurons [10] with exponentially decaying EPSPs (time constant 20 ms) and an absolute refractory time (3 ms). Both neurons receive the same feedforward input (same receptive field, same connection structure) but $A2$'s connections are generally weaker (80% of $A1$). $A2$ also receives long-lasting EPSPs (time constant 80 ms) from the next higher processing level as feedback. There is also a very strong connection from $A1$ to $A2$. It is used to transmit the information from $A1$ via

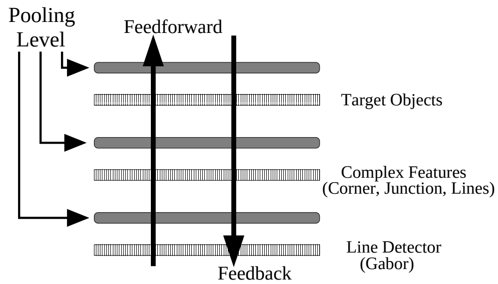

Fig. 2. Processing hierarchy for the vision task. The architecture consists of six processing levels which are connected bidirectionally (feedforward and feedback). Levels 1, 3 and 5 are standard combination detector levels responding to lines (Gabor), complex features, and finally different views of the objects to be recognized. The other three levels perform pooling operations [7] to provide a limited invariance to shifting, rotation and deformation.

$A2$ to other processing levels. That means that if $A1$ is activated, $A2$ will be activated immediately afterwards. ILN modulation has an inhibitory effect on $A1$. The minimum strength of ILN is reached when the first spikes from the preceding level arrive. The quick rise of ILN effectively disables processing in $A1$ within a few ms until ILN again reaches its minimum. The effect of ILN on $A2$ is excitatory and also starts with the arrival of the strongest signals. Then ILN effectively pushes $A2$ neurons with weaker inputs above threshold, with the weakest inputs triggering the last spikes. Therefore, ILN modulation implements the latency encoding. To increase performance and to enforce sparse activity we use a fast and strong winner-take-all inhibition for the $A1$ system. The first active $A1$ inhibits all other $A1$ neurons in other minicolumns with the same receptive field. The feedforward connections are effectively the features of the minicolumns. For every receptive field there is a fixed number of minicolumns with a set of different features. The feedback connections are fully symmetric to the feedforward connections.

5 Vision system

The model architecture (see Fig. 2) was tested for a moderately complex vision task: invariant recognition of a small number of objects. Because we did not employ learning strategies the objects and all filters (connections) were 'hand-designed'. Some of the objects to be detected are shown in the right part of Fig. 3. There are eight different objects stored in eight different orientations. The first processing level uses a standard Gabor-wavelet operation to extract lines. The third level reacts to combinations of lines (longer lines, corners, junctions). The fifth level then detects the target objects at different positions and 8 different orientations (Fig. 3). Levels 2, 4, and 6 perform pooling operations to deal with variations in the input. The latency coding scheme is used to detect the maximum input (minimum latency) out of a group of minicolumns with similar features (neighboring position, similar orientation). This is in effect

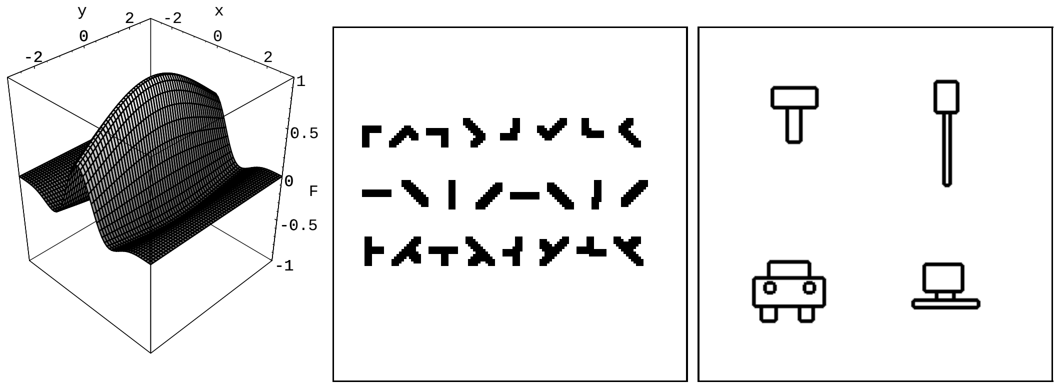

Fig. 3. Features used for the vision system. **Left:** A gabor-wavelet used as a filter for the first processing level. In total 8 different orientations have been used. **Center:** The 24 objects that are used as features for the third level of the processing hierarchy. These are simple combinations of the level one line features. **Right:** Four of the eight high-level objects that are to be recognized by the systems.

a spiking-neuron version of the MAX operator introduced in [7]. The input consists of grey-scale images size 128×128. The total number of minicolumns is about 250,000 with 500,000 neurons.

6 Results

The system was tested with the eight stored high-level objects undergoing certain transformations. We now shortly summarize the results:

- The system is fully translation invariant.
- The system can easily tolerate a rotation of up to 10 degrees from one of the stored views.
- A scaling of up to 10% in both directions does not reduce performance.
- Noise tolerance: White noise poses no real problem; objects are still recognized by the system even when using normally distributed noise with a standard deviation of 30% of the image contrast.
- Distractor objects. Additional low level objects do not degrade performance unless they strongly overlap with the main object.
- Several high-level objects can be detected at once, although in case of a strong overlap the detection rate decreases.

In most cases the first processing cycle was enough to recognize the object and in later iterations feedback could be used to 'correct' local decisions (e.g. suppressing noise or low-level distractor objects). The threshold for the final level was set very high so that the false acceptance rate was very low.

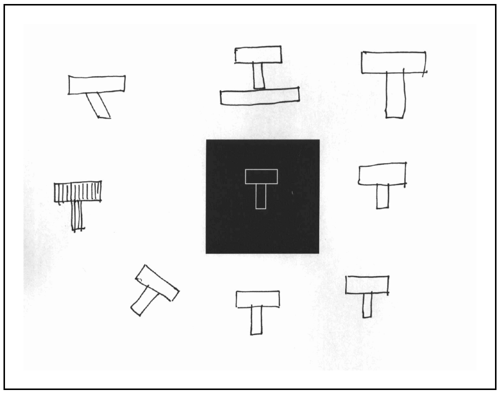

Fig. 4. Variations of the 'hammer' object (the stored form is depicted in the black square in the center).

Finally, we present the results from a set of experiments made with hand-drawn versions of the stored object 'hammer'. The hammer was drawn in eight different variations (on paper) and then scanned (Fig. 4 is the actual scanned image) and preprocessed. All eight samples were correctly detected by the system. It is important to note that not only the higher-level detectors produced the correct result but also intermediate level (i.e. combination feature) detectors produced the correct output (often with the help of feedback).

7 Summary

We presented a biologically plausible model architecture which combines fast recognition with the potential for a feedback controlled, iterative refinement of analysis. The main aspects of our model are latency encoding and columnar processing elements. It was demonstrated that our model processing system can even be applied to large-scale networks (order of 10^5 neurons). Although we employed a very primitive vision system with manually designed features, recognition performance was very good.

References

[1] R. Baddeley, L. F. Abbott, M. C. A. Booth, F. Sengpiel, T. Freeman, E. A. Wakeman, and E. T. Rolls. Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proceedings of the Royal Society of London - Series B: Biological Sciences*, 264:1775–1783, 1997.

[2] Jacques Gautrais and Simon Thorpe. Rate coding versus temporal order coding: a theoretical approach. *BioSystems*, 48:57–65, 1998.

[3] L. L. Glenn and M. Steriade. Discharge rate and excitability of cortically projecting intralaminar thalamic neurons during waking and sleep states. *Journal of Neuroscience*, 2:1387–1404, 1982.

[4] J. J. Hopfield. Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 376:33–36, July 1995.

[5] Edgar Körner, Marc-Oliver Gewaltig, Ursula Körner, Andreas Richter, and Tobias Rodemann. A model of computation in neocortical architecture. *Neural Networks*, 12(7–8):989–1005, 1999.

[6] R. Llinas and D. Pare. *The Mind-Brain Continuum*, chapter The brain as a closed system modulated by the senses, pages 1–18. MIT Press, Cambridge, London, 1996.

[7] Maximilian Riesenhuber and Tomaso Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025, 1999.

[8] Rufin Van Rullen, Jacques Gautrais, Arnaud Delorme, and Simon Thorpe. Face processing using one spike per neurone. *BioSystems*, 48:229–239, 1998.

[9] S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381:520–522, 1996.

[10] Henry C. Tuckwell. *Introduction to Theoretical Neurobiology*, volume 1. Cambridge University Press, Cambridge, 1988.