

Image Segmentation by Complex-Valued Units

Cornelius Weber and Stefan Wermter

Hybrid Intelligent Systems, SCAT, University of Sunderland, UK

Abstract. Spike synchronisation and de-synchronisation are important for feature binding and separation at various levels in the visual system. We present a model of complex valued neuron activations which are synchronised using lateral couplings. The firing rates of the model neurons correspond to a complex number's absolute value and obey conventional attractor network relaxation dynamics, while the firing phases correspond to a complex number's angle and follow the dynamics of a logistic map. During relaxation, we show that features with strong couplings are grouped by firing in the same phase and are separated in phase from features that are coupled weakly or by negative weights. In an example, we apply the model to the level of a hidden representation of an image, thereby segmenting it on an abstract level. We imply that this process can facilitate unsupervised learning of objects in cluttered background.

1 Introduction

Object recognition is a key task in everyday situations and for robotics applications. Unsupervised learning of object classes from natural data is performed by young living beings and has a chance of becoming a convenient and flexible method of learning to categorise sensory data by an artificial agent. A hierarchy of increasingly complex feature detectors is one aspect of the visual recognition process. In many models, such a feature extracting step performs an almost linear vector transformation. So in order to achieve noticeable achievements in their serial application, a strong non-linearity must be introduced at every level.

The non-linear response properties observed in cortical cells are explained in model studies by intra-area horizontal connections. A mathematical advantage to implementing these as an attractor network is that its activations recover noisy input with maximum likelihood [1]. Contrast-invariant orientation tuning curves and shift invariant responses can be obtained [2], as in V1 neurons.

Unsupervised learning of objects is possible if objects are shown on a plain background, but still fails with a noisy background [3]. While the competitive effect of the attractor network reduces background noise, in realistic conditions further percepts are the rule in addition to the object to be learnt. We therefore aim to separate these simultaneous percepts in the dimension of *phase* in order to separate an object from its background. In a hierarchical model this would allow higher levels to learn only the object at certain phases or only background elements at other phases and will facilitate unsupervised learning of objects.

Detailed spiking neuron models are attractive for segmentation purposes. In addition to the neurons' firing rate their code provides information that can

be used for mutual binding and segregation. Computationally it is efficient to incorporate these additional capabilities of spikes in a single variable per neuron which we call a neuron's *phase*. A process to adjust this variable efficiently is deterministic chaos with a dual role of (i) supplying a process of pattern creation by synchronising the phases of coupled units and at the same time (ii) revolting against convergence into stereotyped synchronised states [4].

Our approach has the following characteristics: (i) Coupling strengths are represented by connection weights that can be trained to represent correlated activations and which can be negative. (ii) The weighting of the neural inputs is performed by complex number algebra. Complex-valued neural networks have advantages for chaotic and brain-like systems, image processing and quantum devices [6]. We identify the absolute value of a complex neuron activation with its firing rate and the phase of its activation with the phase at which its spikes are emitted. (iii) The *rates* follow a conventional update dynamics of a recurrent neural net. (iv) A logistic map provides the chaotic dynamics to synchronise and separate the *phases*, as in the "Divide and Conquer" model [5]. (v) Finally, we show an application where an image is segmented on its "hidden", first cortical representation on area V1, similar to the competitive layer model [7].

2 Real Valued Relaxation Procedure

We use one layer of fully connected units to define a recurrent update dynamics for the neurons' rates and phases. After some, possibly random, initialisation of the firing rates, the rate r_k of any unit k is governed over time steps t by

$$r_k(t+1) = f(\sum_j w_{kj} r_j(t) - \theta_k) \quad (1)$$

where w_{kj} is the connection weight from unit j to unit k and θ_k is its threshold. The transfer function is $f(x) = 1/(1 + e^{-x})$. A learning rule for the weights and thresholds is given in [8] and weights which sustain a bell-shaped hill of activation rates will have a Mexican hat shaped profile. Thus, we have a continuous valued attractor network and since weights are approximately symmetric, $w_{kj} \approx w_{jk}$, our intuition is that activations r_k relaxate to a stable state corresponding to a minimum of some energy function. In the following, we will introduce a second variable, the phase φ_k , with different dynamics.

3 Complex Valued Interactions

A complex number as displayed in Fig. 1 a) can be written as $z = x + iy = r e^{i\varphi}$ with $i^2 = -1$ and the relations $r^2 = x^2 + y^2$ and $\tan \varphi = \frac{y}{x}$. We express our neuronal activation as $z_k = r_k e^{i\varphi_k}$ where the complex number's length r_k is the neuron's firing rate and its phase φ_k is the phase at which the neuron spikes. Similar phases of two neurons would correspond to similar firing times if their rates were the same, however, we regard these phases as abstract.

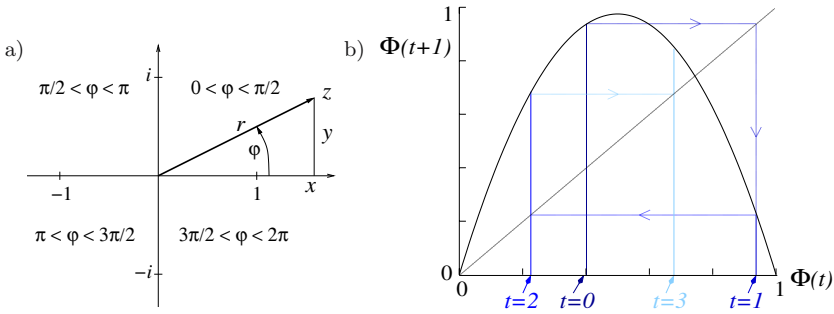


Fig. 1. a) A complex number z can be expressed by Cartesian coordinates x, y or polar coordinates r, φ . We identify its length r with a neuronal firing rate and angle φ with a firing phase. The ranges of φ are displayed as used in the text. **b)** The logistic map. Iterations according to $\varphi(t + 1) = 3.9 \varphi(t) (1 - \varphi(t))$, indicated for $t = 0 \dots 3$ starting at $\varphi(t=0) = 0.4$, will lead to a desired chaotic behaviour.

The dynamics of the rates follows Eq. 1 and is independent of the phases. In the following we will first define how phases between neurons interact (Eq. 2) and then impose a local update dynamic on every neuron (Eq. 3).

At the beginning of the relaxation procedure, the neurons' phases φ_k are initialised with random values between 0 and 2π . Then at each relaxation step a neuron receives an influence from all other neurons, which we express as:

$$z_k^{wf} = \sum_j w_{kj} r_j e^{i\varphi_j} \tag{2}$$

This weighted field is a complex number which is a sum of the complex number contributions by the other neurons weighted by the connection weights w_{kj} which are real values. Using $e^{i\varphi_j} = \cos \varphi_j + i \sin \varphi_j$ it can be expressed as:

$$z_k^{wf} = \sum_j w_{kj} r_j \cos \varphi_j + i \sum_j w_{kj} r_j \sin \varphi_j \equiv x_k^{wf} + i y_k^{wf}$$

We obtain the phase of the neuron's weighted field as: $\varphi_k^{wf} = \text{atan} \frac{y_k^{wf}}{x_k^{wf}}$ which we shift to range between 0 and 2π according to Fig. 1 a).

4 Logistic Coupled Map

The logistic map maps a value between 0 and 1 to another, different value within this interval, as shown in Fig. 1 b). Iterative application leads to desired chaotic development of these values for most settings of the map parameter A between 3.57 and 4.0. Nevertheless, if several values undergoing this mapping are coupled, they can maintain proximity and thus display structured mutual behaviour while displaying chaotic individual behaviour [4].

Since the logistic map takes values ranging between 0 and 1 while the neurons' phases range from 0 to 2π , we make the technical definitions: $\Phi_k \equiv \frac{\varphi_k}{2\pi}$, $\Phi_k^{wf} \equiv \frac{\varphi_k^{wf}}{2\pi}$.

We will now use the scaled phase Φ_k^{wf} of the weighted field at each neuron k in order to determine its scaled phase Φ_k at the next iteration time step. This is done via the logistic map:

$$\Phi_k(t+1) = A \Phi_k^{wf}(t) (1 - \Phi_k^{wf}(t)) \quad (3)$$

where its actual phase value φ_k is scaled back to the range from 0 to 2π . We have set $A = 3.9$. Having obtained the phases at the next time step, another iteration for the phases is performed starting with computing the weighted field (Eq. 2). While the rates develop concurrently according to Eq. 1 to a stable state, the phase values never converge.

5 Network Activation with Synchronising Phases

Figure 2 a) shows the activations of a network with 25 units as their rates have converged to a stable state according to Eq. 1. This is in no way influenced by the phases but only by the weight profile, which is also displayed for one neuron.

The weight profile with strong positive weights between neighbouring units should synchronise the phases between such connected units. A single unit's phase behaves random-like from one time step to the next. Fig. 2 b) shows a plot of time averaged phase differences between pairs of units, while the network is maintaining the rates shown in Fig. 2 a). It shows that adjacent units within the hill of activation have similar phases, while adjacent units at its boundary have differing phases. The phases are thus clustering regions of strong activity that are linked by strong positive weights (phase influence is weighted by weights *times* rates, cf. Eq. 2). Regions with negative connections to such a cluster have differing phases, as we see units with zero activation sharing an own phase. Note that since adjacency is defined functionally by mutual strong positive connections, long-range connections could mediate synchronisation over large distances.

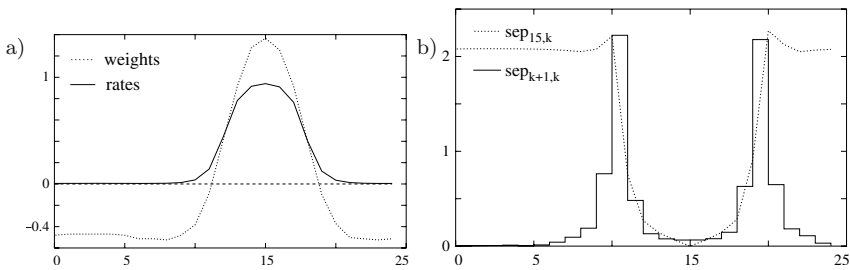


Fig. 2. a) shows the weight profile (dotted line) of neuron number 15 and a bell-shaped hill of activation *rates* as sustained by the network (solid line). b) shows the average *phase separation* $sep_{k+1,k} = \langle |\varphi_{k+1} - \varphi_k| \rangle$ between neighbouring units (solid line) and $sep_{15,k} = \langle |\varphi_{15} - \varphi_k| \rangle$ between any unit and unit 15 (dotted line), where the $\langle \cdot \rangle$ -brackets denote a time average over 500 iteration steps while sustaining the rates in a).

6 Segmentation in a Feature Space

In the following experiment we apply the attractor network to the activations of neurons which have been trained to extract features from natural images and which thereby resemble V1 “simple cells” [2]. Given an image as input, these units have sparse activations with values between 0 and 1 while responding to edges and colour features. V1 lateral weights have been trained on the same set of natural images to memorise these codes in the attractor activation patterns. The resulting weights are short-range excitatory and long-range inhibitory along cortical distance, as well as in feature space of orientation and spatial frequency [2]. This leads, during relaxation, to focused patterns of the activation rates on the simulated V1, after initialising with a somewhat irregular activity pattern obtained from presenting an image. From this V1 representation, a virtual reconstruction of the image can be obtained by projecting these rates back to the image. Fig. 3 a) shows an example.

Fig. 3 c) shows the hidden units’ activations where each frame shows only those units’ activations r_k which have a phase φ_k within a range shown in Fig. 3 b). The active units in each frame are thus a subset of all active units shown in Fig. 3 a), middle. It can be seen that within any selected phase range, preferably units from a certain region are active. The functions of the neurons within these active clusters can be seen by projecting their rates to the image. As a result, Fig. 3 d) shows partial reconstructions of the image which is hereby segmented into elements which belong together by having similar phases on the model V1. At different phases we find elements of the background (frames 1 and 2) or of the ball at the right (frames 4 and 5). This implies a segmentation which accounts for learnt proximities (via the V1 lateral weights) in an abstract representation of an image with the potential of separating objects from their background.

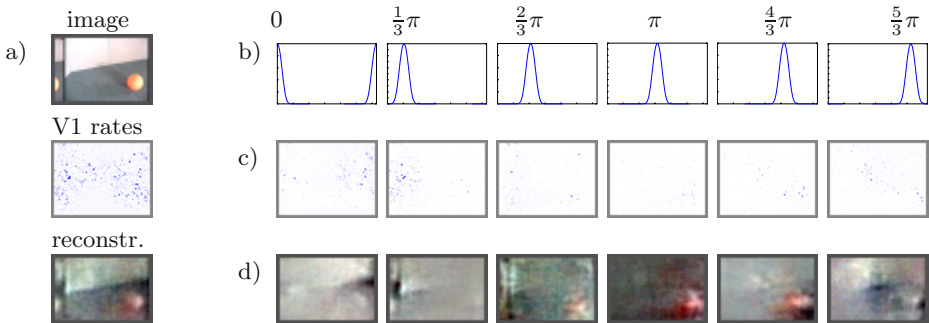


Fig. 3. a) shows full representations. Top, the original image, middle, the rates of the full hidden code and bottom, the reconstruction of the image from the full hidden code. b), c), d) show partial phase-dependent representations. b) shows Gaussian-like functions on an axis of 0 to 2π used to determine which phases contribute to the hidden code presented in c) and thus to the image reconstruction in d). c) shows the partial hidden code corresponding to selected phases and d) shows their partial reconstruction from those units in c) which are active at the selected phases.

7 Discussion

We have demonstrated a network of simplified *rate* and *phase* coding neurons which segments a neural code efficiently using the connection strengths between the units. The computational load is that of a network in which two activation values develop concurrently. The most demanding operations are the scalar products in Eqs. 1 and 2, while all other computations are local.

Since image segmentation involves top-down directed information flow, how does our model for intra-area lateral connections deal with this? Previously we have extended the lateral connections to link a “what”- to a “where” area for object localisation [9]. In the cortex, such lateral connections correspond to those originating from pyramidal cells in layers 2/3 of the cortex and arriving in the same layer, possibly in a different area. Hierarchically arranged areas have characteristically asymmetric connections, however, they also have characteristic horizontal connections and only with increasing hierarchical level difference the intensity of these horizontal connections decreases [10]. Thus, connections of a horizontal character may relay top-down information. If we would apply the lateral connections of our model to a larger hierarchical model, therefore, we might observe top-down influences such as stabilisation of consistent attractors.

Acknowledgements. This work is part of the MirrorBot project supported by the EU in the FET-IST programme under grant IST-2001-35282.

References

- [1] Deneve, S., Latham, P., Pouget, A.: Reading population codes: a neural implementation of ideal observers. *Nature Neurosci.* **2** (1999) 740–5
- [2] Weber, C.: Self-organization of orientation maps, lateral connections, and dynamic receptive fields in the primary visual cortex. In Dorffner, G., Bischof, H., Hornik, K., eds.: *Proc. ICANN, Springer-Verlag Berlin Heidelberg* (2001) 1147–52
- [3] Stringer, S., Rolls, E.: Position invariant recognition in the visual system with cluttered environments. *Neural Networks* **13** (2000) 305–15
- [4] van Leeuwen, C., Steyvers, M., Nooter, M.: Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *J. Mathematical Psychology* **41** (1997) 319–44
- [5] Raffone, A., van Leeuwen, C.: The “divide and conquer” model of image segmentation: object-bounded synchrony propagation in coupled map lattices. In: *Proc. KES Knowledge-Based Intelligent Information & Engineering Systems*. (2002)
- [6] Hirose, A., ed.: *Complex-Valued Neural Networks: Theories and Applications*. World Scientific Publishing Co. (2003)
- [7] Ontrup, J., Wersing, H., Ritter, H.: A computational feature binding model of human texture perception. *Cognitive Processing* **5** (2004) 32–44
- [8] Zhang, K.: Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *J. Neurosci.* **16** (1996) 2112–26
- [9] Weber, C., Wermter, S.: Object localization using laterally connected “what” and “where” associator networks. In: *Proc. ICANN/ICONIP, Springer* (2003) 813–20
- [10] Felleman, D., Van Essen, D.: Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* **1** (1991) 1–47