Proceedings European Symposium on Artificial Neural Networks, Bruges 1999, pages 295-300

Feature Binding and Relaxation Labeling with the Competitive Layer Model

Heiko Wersing and Helge Ritter

University of Bielefeld, Technical Faculty
P.O.-Box 10 01 31, D-33501 Bielefeld, Germany
{hwersing,helge}@techfak.uni-bielefeld.de

Abstract. We discuss the relation of the Competitive Layer Model (CLM) to Relaxation Labeling (RL) with regard to feature binding and labeling problems. The CLM uses cooperative and competitive interactions to partition a set of input features into groups by energy minimization. As we show, the stable attractors of the CLM provide consistent and unambiguous labelings in the sense of RL and we give an efficient stochastic simulation procedure for their identification. In addition to binding the CLM exhibits contextual activity modulation to represent stimulus salience. We incorporate deterministic annealing for avoidance of local minima and show how figure-ground segmentation and grouping can be combined for the CLM application of contour grouping on a real image.

1 Introduction

A major challenge of computational neuroscience is the question which neural mechanisms could facilitate the process of *dynamic feature binding*. From the viewpoint of brain theory [12], feature binding may provide one of the basic sensory information processing principles. This has raised much interest in using similar mechanisms for pattern recognition applications like image segmentation and object recognition. While a lot of neural network research focuses on binding models based on temporally correlated neural activity [12], also supported by experimental data [2], successful applications to real-world data are still rather exceptional. This is mainly caused by the high dynamical complexity of these models, which makes their simulation costly and their analytic study a difficult task. In the field of image segmentation, many successful approaches rely on the minimization of a suitable cost function by iterative algorithms. A cost function yields a very direct way of controlling the desired groupings by merging contextual constraints into an energy landscape with minima as possible output states. Relaxation labeling [11] (RL) is a family of such iterative procedures which has become a standard technique in pattern recognition and machine vision domains [8].

The competitive layer model (CLM) [10] provides an energy-based recurrent network approach to feature binding which has been applied to Gestalt-motivated models of perceptual grouping [13,6]. In this contribution we discuss the relation between the well-established theory of RL [3] and the competitive recurrent neural circuit of the CLM. In the CLM, binding is achieved by a collection of competitive layers, which produces feasible solutions to the labeling or binding problem as stable attractors. The central advantages of the CLM approach are i) analytical results concerning dynamics and attractors, ii) a straightforward neural circuit interpretation, and iii) a very simple computer implementation with a rapidly converging asynchronous iteration routine.

2 Feature Binding as a Labeling Problem

Relaxation labeling (RL) [11], is an approach to solve the following problem: Given a set of N features $r = 1, \dots, N$ and a set of L labels $\alpha = 1, \dots, L$, find a labeling of the features which embodies contextual information in an optimal way. The contextual constraints are given by a set of compatibility coefficients $\mathsf{f}_{rr'}^{\alpha\beta}$ which denote the mutual compatibility of assigning label α to feature r and label β to feature r'. The coefficients may be derived by heuristic arguments, statistical considerations or learning [8]. In this framework, we interpret the attachment of the same labels as a binding of features and restrict ourselves to interactions between features with the same labels $\mathsf{f}_{rr'}^{\alpha\beta} = \delta_{\alpha\beta}\mathsf{f}_{rr'}^{\alpha}$. This is to be distinguished against matching problems [4], where one-to-one constraints between features and labels require inter-label interactions.

If we define $x_{r\alpha} \geq 0$ as the certainty of the assignment of label α to feature r, the space of weighted labelings is defined by the condition $\sum_\alpha x_{r\alpha} = 1$. The task of an RL algorithm is to find an unambiguous and consistent labeling, that is $x_{r\alpha(r)} = 1$, $x_{r\beta\neq\alpha(r)} = 0$, $r = 1, \dots, N$ where $\alpha(r)$ denotes the unambiguously assigned label to feature r, and consistency of the unambiguous labeling is defined as $F_{r\alpha(r)} > F_{r\beta\neq\alpha(r)}$, where the $F_{r\alpha} = \sum_{r'} \mathsf{f}_{rr'}^{\alpha} x_{r'\alpha}$ are called the linear support functions which accumulate the pairwise contributions of all other certainties weighted with their compatibility. For symmetric compatibilities with $\mathsf{f}_{rr'}^{\alpha} = \mathsf{f}_{r'r}^{\alpha}$ these conditions are necessary and sufficient for a local maximum of the average local consistency $A = \sum_{r\alpha} x_{r\alpha} F_{r\alpha}$. In that case we can formulate the problem of feature binding as an optimization problem. Conventional RL algorithms as that of Hummel and Zucker (HZ) [3] and Rosenfeld [11] converge to local maxima of A and achieve good results, if it is either not necessary to find a global optimum, or the corresponding energy function has not many local maxima. Alternative approaches are mean-field-annealing algorithms [9,14] which improve the chance of finding optimum or near optimum solutions.

A central feature of these algorithms is the explicit reprojection of the current state onto the space of weighted assignments upon each iteration. The CLM achieves this by enforcing the constraint only approximately, but still ensuring convergence to consistent and unambiguous labelings. The result is a context-dependent activity modulation, where activity represents the degree of salience of a feature. This property which results in more flexible responses than the conventional RL approach, is also in accordance with experimentally observed context-dependent activity modulations [5]. Deterministic annealing can be incorporated in the CLM by a simple self-inhibitory loop, the strength of which can be interpreted as an inherent temperature in analogy to mean-field-annealing.

3 The CLM Architecture

The CLM consists of a set of L identical layers of feature-selective neurons which are replicas of an input layer (see Fig. 1). The neurons in the input layer are labelled by r. Driven by an external input, each input neuron responds with a value h_r which indicates activity ($h_r > 0$) in the presence of the corresponding feature r or silence ($h_r \leq 0$). In a simple setting we may think of h_r as encoding the light intensity at position $\mathbf{r}$ in

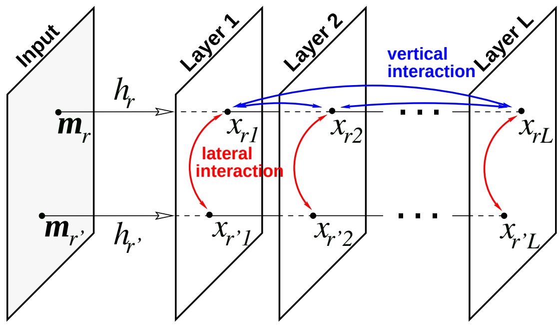


Fig. 1. The CLM architecture

some "imaginary" retina. For each of the input neurons at position r there is one neuron per replica layer α whose activity is $x_{r\alpha} \geq 0$ and represents the certainty of assigning feature r to layer α. We denote the L neurons $x_{r\alpha}$, responding to feature r as the *column* r. The activities are subject to the following constrained gradient dynamics:

$$\dot{x}_{r\alpha} = -\frac{\partial E}{\partial x_{r\alpha}} = J(h_r - \sum_{\beta} x_{r\beta}) + \sum_{r'} \mathsf{f}^{\alpha}_{rr'} x_{r'\alpha} - T\, x_{r\alpha} \quad , \tag{1}$$

subject to $x_{r\alpha} \geq 0$, where the energy is given by

$$E = \frac{J}{2} \sum_{r} \Big(\sum_{\alpha} x_{r\alpha} - h_r \Big)^2 - \frac{1}{2} \sum_{\alpha} \sum_{rr'} \mathsf{f}^{\alpha}_{rr'} x_{r\alpha} x_{r'\alpha} + \frac{1}{2} T \sum_{r\alpha} x^2_{r\alpha} \quad . \tag{2}$$

J and T are positive constants and $\mathsf{f}^{\alpha}_{rr'} = \mathsf{f}^{\alpha}_{r'r}$ are the components of a symmetric weight matrix. The gradient in (1) can be split into $-\frac{\partial E}{\partial x_{r\alpha}} = JV_r + F_{r\alpha} + Tx_{r\alpha}$, where $V_r = h_r - \sum_{\beta} x_{r\beta}$, $F_{r\alpha} = \sum_{r'} \mathsf{f}^{\alpha}_{rr'} x_{r'\alpha}$, and $-Tx_{r\alpha}$ are three basic interactions in the model, which can be interpreted with regard to the energy E:

i) The *"vertical"* interaction JV_r implements a dynamical winner-take-all circuit within each column. Unlike in a standard penalty function, J should not be chosen large, but slightly above a critical value $J_c = \max_{r\alpha} \sum_{r'} \max(0, \mathsf{f}^{\alpha}_{rr'})$ which ensures convergence of the dynamics [13] and allows for a modulation of the input h_r by the lateral interactions.

ii) The lateral interaction $F_{r\alpha}$ couples activities within layers by the symmetric weight matrix $\mathsf{f}^{\alpha}_{rr'}$. The compatibility coefficients $\mathsf{f}^{\alpha}_{rr'}$ determine which pattern configurations, if elicited as activity pattern within layers, will be mutually supporting among their constituent parts ($\mathsf{f}^{\alpha}_{rr'} > 0$) or instead suffer mutual inhibition ($\mathsf{f}^{\alpha}_{rr'} < 0$).

iii) The self-inhibitory interaction $-Tx_{r\alpha}$ biases the minima of E towards ambiguous assignments. $T > 0$ can be regarded as a temperature, for very large T the global minimum of (2) is given by $x_{r\alpha} = h_r/L$, the maximally unassigned state. For $T \to 0$ the dynamics converges towards a proper consistent labeling in the sense of RL , as we will prove in the next section; by gradually lowering T we can perform deterministic annealing to avoid falling into local minima.

4 Efficient Relaxation Labeling with the CLM

The CLM dynamics can be simulated in principle by standard differential equation integrators like the Euler or Runge-Kutta method and can be computed in parallel. The piecewise linear dynamics, however allows also for a sequential asynchronous update [6] which shows rapid convergence and can be very easily implemented:

1. Set $T(0) = T_c$, where $T_c = \lambda_{max}(\mathsf{f}^{\alpha}_{rr'})$.
   Initialize all $x_{r\alpha}$ with random values $x_{r\alpha}(t=0) \in [h_r/L - \epsilon, h_r/L + \epsilon]$.
2. Do $N \cdot L$ times: Choose (r, α) randomly and update $x_{r\alpha}(t+1) = \max(0,\, \xi)$, where
   $\xi = \frac{1}{J - \mathsf{f}^{\alpha}_{rr} + T}\Big(J\big(h_r - \sum_{\beta \neq \alpha} x_{r\beta}(t)\big) + \sum_{r'} \mathsf{f}^{\alpha}_{rr'} x_{r'\alpha}(t) - \mathsf{f}^{\alpha}_{rr} x_{r\alpha}(t)\Big)$
3. Set $T(t+1) = \eta T(t)$, with $0 < \eta < 1$. Go to step 2 until convergence.

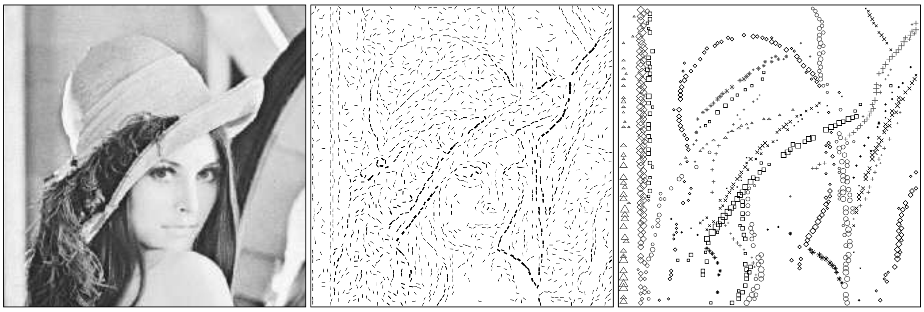
Step 2 corresponds to solving the linear equation (1) $\dot{x}_{r\alpha} = 0$ independently for a randomly chosen activity $x_{r\alpha}$. If $\mathsf{f}^{\alpha}_{rr} > T$, we can be sure that this asynchronous dynamics converges towards an attractor of the continuous model (1) according to a recent convergence result [1,6]. This holds also for $T = 0$. The exact computation of the largest eigenvalue of the compatibilities can also be replaced by a simple conservative approximation. Since this update procedure converges to a feasible CLM attractor we can now reconsider an earlier result from [13] in the RL framework.

Theorem 1. *If the compatibilities satisfy* $\mathsf{f}^{\alpha}_{rr} + \mathsf{f}^{\beta}_{rr} > 2\,T$ *for all* r, α, β*, then the asynchronous CLM update converges to a consistent and unambiguous modulated RL labeling with i) at most one positive activity* $x_{r\alpha(r)} = h_r + F_{r\alpha(r)}/J$ *in a column where* $\alpha(r)$ *is the index of the maximally supporting layer with* $F_{r\alpha(r)} > F_{r\beta \neq \alpha(r)}$ *or ii) for all activities in a column* $x_{r\alpha} = 0,\ F_{r\alpha} \leq 0$.

The dynamical coupling to the input results in a modulated activity $x_{r\alpha(r)}$ of the final assignment. This is useful since it introduces an auto-associative component into the binding process. Features that receive low input h_r, may develop a higher output activity due to strong lateral feedback $F_{r\alpha} > 0$, but the network still remains sensitive to variations in the input intensities. We emphasize that annealing in T is not necessary for convergence to consistent labelings which is also guaranteed for $T = 0$ and constant. It only reduces the chance of finding suboptimal, but feasible labelings.

5 Application to Contour Grouping

Contour grouping is an important objective for models of feature linking, where compatibilities between edge features generally express the degree of continuity of a curve passing over them. Synchronization-based models, however, still face major difficulties in delivering a controllable grouping for the complex excitatory and inhibitory interactions as encountered in real world images. There is now a long tradition in the pattern recognition community of using RL for the process of contour integration [7], which aims only at the detection of contours in noisy images. We will show now, how a similar RL-motivated approach can be used in the recurrent CLM network to combine a binding of salient contour groups with a mechanism of figure-ground segmentation.



a) input b) edge features c) CLM binding

Fig. 2. CLM contour grouping on a real image. From the input image a) edge features b) are generated by a sobel x-y operator and sub-sampling. Edge thickness corresponds to the gradient intensity. The result of the binding with a CLM with 14 layers + ground layer is shown in c). Black/grey symbols code for different layers, with size corresponding to output activity. The ground layer is omitted. Note the enhancement of low-input edges at the brim and the upper part of the hat. Lateral interaction parameters are $R = 0.15$, $S = 300$, $I = 0.6$, $k = 1$, $m = 3$.

The lateral interactions $\mathsf{f}^{\alpha}_{rr'}$ are given as $\mathsf{f}^{1}_{rr'} = m\delta_{rr'}$ for the *ground* layer and as a co-circular interaction [7] with lateral inhibition $\mathsf{f}^{\alpha>1}_{rr'} = \mathsf{f}_{rr'}$ in the other layers. The parameter $m > 0$ defines a self-coupling against which lateral interactions in the figure layers must compete to "pop out" a feature from the ground layer. The co-circular interaction of two edges at positions $\mathbf{r}_1 = (r_1^x, r_1^y)$ and $\mathbf{r}_2 = (r_2^x, r_2^y)$ with a difference vector $\mathbf{d} = \mathbf{r}_1 - \mathbf{r}_2$, $d = |\mathbf{d}|$ and $\hat{\mathbf{d}} = \mathbf{d}/d$, and unit orientation vectors $\hat{\mathbf{n}}_1 = (n_1^x, n_1^y)$, $\hat{\mathbf{n}}_2 = (n_2^x, n_2^y)$ is given by $f((\mathbf{r}_1, \hat{\mathbf{n}}_1), (\mathbf{r}_2, \hat{\mathbf{n}}_2)) = \theta(a_1 a_2 q)\big(e^{-d^2/R^2 - C^2 S}\big) - Ie^{-2d^2/R^2} - k/N$, where $a_1 = n_1^x \hat{d}_y - n_1^y \hat{d}_x$, $a_2 = n_2^x \hat{d}_y - n_2^y \hat{d}_x$, $q = \hat{\mathbf{n}}^1 \cdot \hat{\mathbf{n}}^2$ and $\theta(x) = 1$ for $x \geq 0$ and $\theta(x) = 0$ otherwise is necessary to exclude skewed symmetric edges. The parameter R controls the spatial range, which is smaller for the inhibitory component. The degree of co-circularity is given by $C = |\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{d}}| - |\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{d}}|$ which is equal to zero if both edges lie tangentially to a common circle. The parameter $S > 0$ controls the sharpness of the co-circularity constraint. Parameters $I > 0$ and $k > 0$ control the strength of the local and global inhibition, respectively.

The application to a real image, scaled into a unit square, is shown in Fig. 2. The edge input intensities h_r are chosen as the absolute value of the local gradient intensity and $\hat{\mathbf{n}}_r$ and $\mathbf{r}_r$ as orientation and position respectively. The constant J is chosen as $J = 1.1J_c$ (see Sect.3), which results in contextual modulation of the output intensities and enhances edges with low input, but strong lateral support. Annealing was started with $T_c = \lambda_{max}\{\mathsf{f}_{rr'}\}$ with a schedule of $\eta = 0.99$. Faster lowering leads to more fragmented groupings. Raising the ground layer coupling m from zero to higher values suppresses less salient groups, until at $m \approx J_c$ only the ground layer remains active.

6 Conclusion

We showed how feature binding with the CLM can be reconsidered in the RL framework. The stochastic asynchronous update converges to consistent labelings and pro-

vides a highly efficient simulation procedure that might also prove to be very useful for other RL applications in labeling tasks. As compared to our earlier work [13], the incorporation of deterministic annealing leads to a better and less fragmented grouping quality. An interesting result is that simple additional quadratic terms in the energy which lead to linear modifications of the dynamics give comparable performance for labeling problems as the more complex Potts-Mean-Field annealing. In the application section we show how contour grouping and figure-ground segmentation can be performed with the CLM on a complex real image. Our results show that the layered topology leads to a stronger uncoupling of formed groups, which we consider essential for a robust representation of multiple bindings. A combination of the presented spatial mechanisms with temporal mechanisms provides an interesting future perspective. The link to RL also offers the application of a recently proposed learning scheme [8] for the compatibility coefficients which opens the door for supervised learning of lateral interactions for feature binding.
Acknowledgements: H. Wersing is supported by DFG grant GK-231.

References

1. J. Feng. Lyapunov functions for neural nets with nondifferentiable input-output characteristics. *Neural Computation*, 9(1):43–49, 1997.
2. C. M. Gray, P. König, A. K. Engel, and W. Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338:334–337, 1989.
3. R. A. Hummel and S. W. Zucker. On the foundations of relaxation labeling processes. *IEEE Trans. Pattern Analysis and Mach. Intell.*, 5(3):267–286, 1983.
4. B. Kamgar-Parsi and B. Kamgar-Parsi. On problem solving with Hopfield networks. *Biological Cybernetics*, 62:415–423, 1990.
5. M. K. Kapadia, M. Ito, and G. Westheimer. Improvement of visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron*, 15:843–856, 1995.
6. J. Ontrup and H. Ritter. Perceptual grouping in a neural model: Reproducing human texture perception. Technical Report SFB 360 98-6, University of Bielefeld, 1998.
7. P. Parent and S. W. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE Trans. Pattern Analysis and Mach. Intell.*, 46(17):763–770, 1989.
8. M. Pelillo and A. M. Fanelli. Autoassociative learning in relaxation labeling networks. *Pattern Recognition Letters*, 18:3–12, 1997.
9. C. Peterson and B. Soderberg. A new method for mapping optimization problems onto neural networks. *Int. J. Neural Systems*, 1(1):3–22, 1989.
10. H. Ritter. A spatial approach to feature linking. In *Int. Neur. Netw. Conf. Paris*, 1990.
11. A. Rosenfeld, R. A. Hummel, and S. W. Zucker. Scene labeling by relaxation operations. *IEEE Trans. Systems, Man, Cybernet.*, 6(6):420–433, 1976.
12. C. v.d. Malsburg. The correlation theory of brain function. Technical Report 81-2, MPI Göttingen, 1981.
13. H. Wersing, J. J. Steil, and H. Ritter. A layered recurrent neural network for feature grouping. In W. Gerstner et al., editor, *Proc. ICANN Lausanne*, pages 439–444, 1997.
14. Alan L. Yuille and J. J. Kosowsky. Statistical physics algorithms that converge. *Neural Computation*, 6(3):341–356, 1994.