Hallucinating Image Features to Supplement Perceptual Groups

Martin Meier, Robert Haschke, Helge Ritter
{mmeier,rhaschke,helge}@techfak.uni-bielefeld.de

August 29, 2011

Abstract

In this paper we present an approach towards cognitive reasonable figure amendments utilizing the Gestalt-based dynamics of the Competitive Layer Model.

1 Introduction

When a human perceives incomplete shapes, for example the ones from Fig. 1, no effort is needed to recognize the meant geometric primitives, although they are far from being complete. In this paper, we propose an human-like approach to fill these "gaps". Based on Gestalt Theory (e.g. see [1] for an overview), especially the law of continuity, we strive to amend these sparse informations through modelling missing parts utilizing the neural dynamics of the Competitive Layer Model (CLM).

The CLM [3] has been proven feasible in a wide spectrum of recognition tasks. Previous works successfully applied the CLM to simulate various grouping tasks based on Gestalt Laws like contour grouping in noisy settings [5] or action segmentation [2].

Based on the approaches for contour grouping, we make use of the internal binding dynamics of the CLM to evaluate the quality of hallucinated features with respect to previously grouped contours.

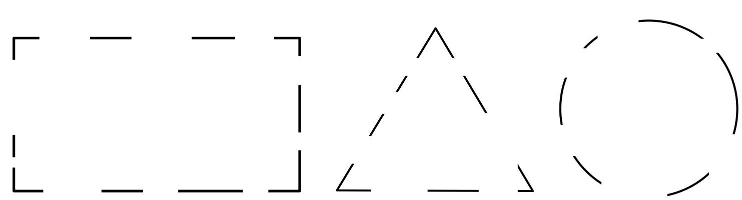


Figure 1: Gestalt Law of continuity: Although the shapes are not complete, they are easily recognized as a rectangle, triangle and circle.

2 The Competitive Layer Model

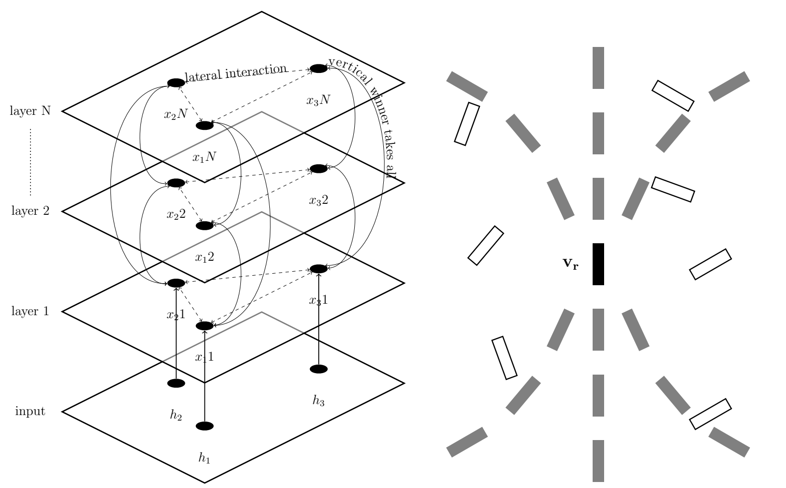
The CLM uses an internal recurrent dynamics to group similar features. To this end, a set of $L \times N$ linear threshold units are arranged in L neuron layers. We denote the activity of a neuron with $x_{r\alpha}$, where $r = 1..N$ denotes the feature index and $\alpha = 1..L$ the layer index. Hence, for each feature r exists a column of neurons across all L layers. The significance of a feature r is determined by the external input h_r (cf. Fig. 2(a)).

Within each layer a lateral interaction $f_{rr'}$ is defined according to the compatibility or similarity of features v_r and $v_{r'}$. If both features are considered similar, a positive connection weight between $x_{r\alpha}$ and $x_{r'\alpha}$ is used, realizing a positive feedback loop. This compatibility measurement is domain specific for the type of used features v and must therefore be explicitly specified in a symmetric interaction function:

$$f_{rr'} = \mathbf{f}(v_r, v_{r'}) = \mathbf{f}(v_{r'}, v_r) \tag{1}$$

This mutually reinforces activity of neurons representing similar features. All layers employ the same lateral interaction weights.

Grouping of features is realized by collecting positive neuronal activity within layers. To enforce activation of a neuron related to a particular feature v_r within a single layer only, the lateral layer-wise interaction is augmented by a column-wise *winner-takes-all (WTA)* interaction. The combination of the vertical WTA



(a) Competitive Layer Model (b) Compatibility for oriented edges

Figure 2: (a) The Competitive Layer Model with three inputs $h_{1...3}$ and the corresponding neurons $x_{r\alpha}$ in each layer. (b) Compatibility for oriented edges. Emanating from the centered feature v_r, dark filled edges indicate a high compatibility whereas unfilled edges indicate low compatibility.

dynamics and the lateral interactions leads to a linear threshold dynamics of

$$\dot{x}_{r\alpha} = -x_{r\alpha} + \sigma(J(h_r - \sum_{\beta} x_{r\beta}) + \sum_{r'} f_{rr'} x_{r'\alpha}) \tag{2}$$

with $\sigma(x) = max(0, x)$, where $h_r - \sum_{\beta} x_{r\beta}$ represents the vertical *WTA* interaction, weighted by a (usually small) constant J and $\sum_{r'} f_{rr'} x_{r'\alpha}$ represents the lateral interaction.

Since the lateral interactions $f_{rr'}$ are identical in each layer, they can be calculated once and stored in a symmetric interaction matrix

$$M_{rr'} = \mathbf{f}(v_r, v_{r'}) \tag{3}$$

An exemplary interaction function is shown in Fig. 2(b), displaying the interaction of oriented edges. Starting from the centered feature v_r, features with a similar orientation w.r.t. to their distance have a higher compatibility than nearly perpendicular features in close proximity.

3 Hallucinating Features

We strive to use the CLM binding dynamics to "imagine" well matching amendments for sparse geometric shapes. In order to achieve this goal, we apply the CLM to a set of geometric shapes, let it converge and then induce hallucinated features to evaluate their compatibility using the binding dynamics.

The induction of hallucinated features is currently done without a priori knowledge about the distribution of known features from the CLM grouping. Therefore the search space is narrowed to a finite set and the search for well matching hallucinated features is currently done with a "brute force" approach. For each possible element the compatibility to the existing groups is evaluated.

To evaluate the compatibility of a new feature vector v_{new}, an interaction vector

$$m = (\mathbf{f}(v_{new}, v_0), \mathbf{f}(v_{new}, v_1), \ldots, \mathbf{f}(v_{new}, v_r))^T \tag{4}$$

is created to extend the interaction matrix $M_{rr'}$:

$$M_{new} = \begin{pmatrix} M_{rr'} & m \\ m^T & 1 \end{pmatrix} \tag{5}$$

The support for the hallucinated feature from the existing neurons is then calculated as:

$$x_{v_{new}\alpha} = m^T \cdot \vec{x}_{\alpha} \tag{6}$$

4 Preliminary Results

To evaluate the proposed approach, we applied a CLM with ten layers to a set of sparse circles composed of oriented edges, as depicted in Fig. 3(a), with an

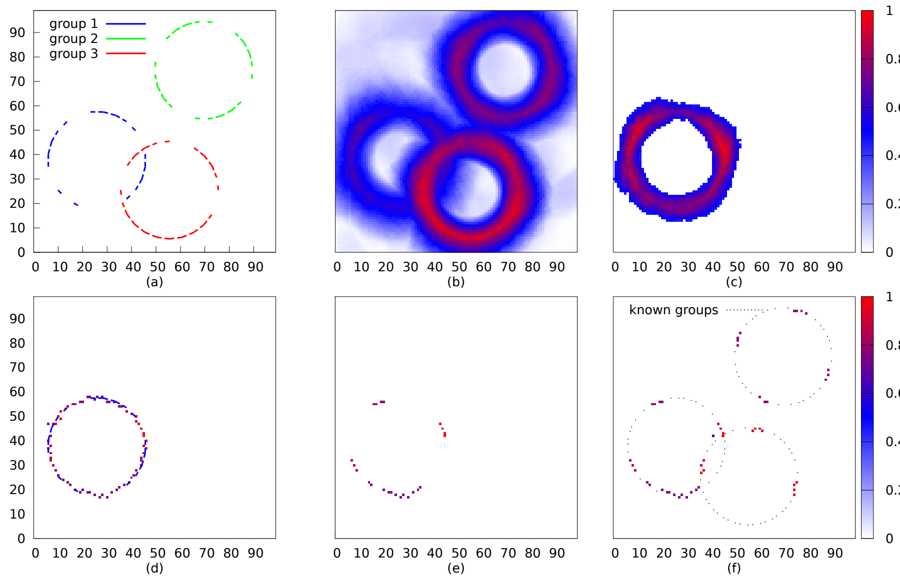


Figure 3: (a) CLM grouping of three sparse circles. (b) Activity of hallucinated features merged over all layers. (c) Activation from hallucinated features for a single layer after applying a threshold of 0.5. (d) Local maximum in a 5×5 neighbourhood with known features from group 1. (e) Known features subtracted from previous maxima. (f) Best matching features in unoccupied areas for all layers with known features from (a).

oriented edge defined by a 2D position (x, y) and orientation θ. Different layers are represented with different colors. For each position in the 100×100 input space 36 features with different orientations in a range from $0°$ to $175°$ were imagined and evaluated for their compatibility with existing groups.

Fig. 3(b) shows the maximal activity at each position (x, y) over all possible orientations θ. Please note that Fig. 3(b) is furthermore a combination of all layers.

To reduce the noise from not well matching hallucinated features, a single layer is selected in Fig. 3(c) and a threshold is applied, which sets every activity smaller than 0.5 to zero.

To narrow down the result of the thresholding, a filter which selects the maximum in a 5×5 neighbourhood is utilized. This new local maximum is then used as point of origin for a new filtering step in which already visited positions are omitted. This enables the filter to "follow" local maxima. Of course hallucinated features in close proximity to already known features are selected by this filter, too. This is shown in Fig. 3(d), where the result of the filtering process is overlaid with group 1 from Fig. 3(a).

In an additional step depicted in Fig. 3(e), hallucinated features in close proximity to existing groups are removed, leaving only good amendments. Fig. 3(f) shows the above mentioned steps for all groups, including the original CLM grouping results of from Fig. 3(a). In the interests of clarity, all groups are displayed with the same symbols.

These results show the feasibility of using the CLM dynamics in conjunction with hallucinated features to amend sparse informations.

5 Conclusion

Inducing hallucinated features into the CLM opens an interesting foundation to amend sparse informations, which is not only limited to the completion of geometric shapes but can also be generalized to much more complex scenarios. For example given the action segmentation from [2], it is imaginable to use the CLM for action generation, given a set of incomplete action segments.

It also introduces a lot of new questions for research, e.g. how to overcome the current "brute force" approach to initially generate hallucinated features, as well as a more general technique to finally find good amendments in contrast to the feature specific method presented here.

Also of interest will be a combination of learning the lateral interactions as presented in [4] with amendment through hallucinated features to gain a better generalization.

Acknowledgements

This work has been conducted within and funded by the German collaborative research center "SFB 673: Alignment in Communication" granted by Deutsche Forschungsgemeinschaft.

References

[1] A. Desolneux, L. Moisan, and J.M. Morel. *From gestalt theory to image analysis: a probabilistic approach.* Springer Verlag, 2008.

[2] M. Pardowitz, R. Haschke, J. Steil, and H. Ritter. Gestalt-based action segmentation for robot task learning. In *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS Int. Conf. on*, pages 347–352. IEEE, 2008.

[3] H. Ritter. A spatial approach to feature linking. In *Int. Neural Network Conference, Paris*, 1990.

[4] S. Weng, H. Wersing, J.J. Steil, and H. Ritter. Learning lateral interactions for feature binding and sensory segmentation from prototypic basis interactions. *Neural Networks, IEEE Transactions on*, 17(4):843–862, 2006.

[5] H. Wersing, J.J. Steil, and H. Ritter. A competitive-layer model for feature binding and sensory segmentation. *Neural Computation*, 13(2):357–387, 2001.