An improved model for contour completion in V1 using learned feature correlation statistics

Vadas Gintautas[1,3], Shawn Barr[5], Michael Ham[2], Benjamin Kunsberg[6], Steve Zucker[6], Steven Brumby[4], Craig Rasmussen[5], Luis M. A. Bettencourt[3] Garrett T. Kenyon[3,5]

[1] CNLS; [2] P-21; [3] T-5; [4] ISR-2 Los Alamos National Laboratory, Los Alamos, NM

[5] New Mexico Consortium, 4200 W. Jemez Rd., Los Alamos, NM

[6] Center for Computational Vision and Control, Yale University, New Haven, CT

vadasg@lanl.gov

1. Background and new image set

Limitations of existing photographic image sets:

- Limitations in image resolution and number of images during set creation
- Reference to exogenous knowledge, such the definition of "animal"
- Non-parametric complexity or difficulty

Features of our data set:

- Not limited in size
- Relies on no explicit outside knowledge
- Has tunable parameters so that the difficulty can be varied

1.1 Generated Image Set ("Amoeba / No Amoeba")

- An "amoeba" is a deformed, segmented circle (with gaps).
- The radius varies with polar angle.
- A distractor "no-amoeba" image is created by rotating groups of amoeba segments through random angles.
- Randomly superimposed no-amoeba images serve as background clutter.

The segments are composed of a series of points from a periodic random function, with period 2π:

$$r(\phi) = A_0 + \sum_{n=1}^{N} A_n \cos(n\phi) + B_n \sin(n\phi).$$

The average radius length is drawn from a Gaussian distribution with mean m_0 and standard deviation c_0: $A_0 \sim N(m_0, c_0)$. Changing the speed α and the amount c that the radii fluctuates determines how much the segments wiggle: $A_n \sim N(0, c/n^{\alpha})$; $B_n \sim N(0, c/n^{\alpha})$. The difficulty of the task can be parametrically varied by adjusting the statistical parameters of the line segments.

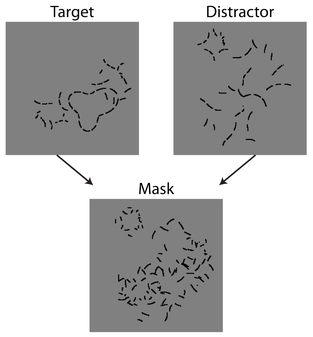

Figure 1: *Left: A generated test image that contains a target, or "amoeba." Right: A generated test image that does not contain an amoeba. Center and below: For human psychophysics experiments, a mask is created by superimposing the target and distractor images, then rotating each segment by a random angle.*

2. Methods

2.1 Human Psychophysics Experiments

- Images are presented to subjects in a light-controlled environment.
- One target image and one distractor image are shown side by side, followed by a mask.
- The task is 2 Alternative Forced Choice.
- The subject indicates which side had the target and confidence.
- Images are presented on a 19 Hitachi CRT monitor.
- The monitor resolution was 1024×768, with a refresh rate was 100 Hz.
- All participants were given a viewing angle of $7^{\circ} \times 7^{\circ}$.
- Each subject was shown 1200 images divided into 10 blocks of 120 images.

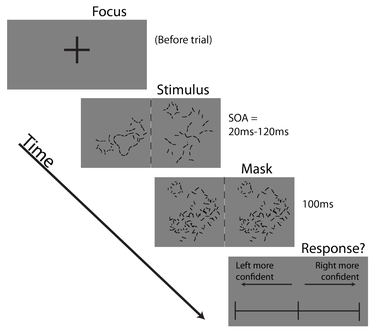

Figure 2: *First the subject is instructed to focus on the center of the screen. The stimulus consists of 1 distractor image and 1 target image (left vs right is randomized). After 20ms -120ms SOA stimulus presentation, a mask is shown for 100ms. Then the subject reports the position of the target by using a mouse to click along a horizontal line. The position of the click corresponds to the subject's confidence in his or her response.*

2.2 Hierarchical models of primate visual cortex

Building on Serre, et al. [?] we developed PANN (Peta-scale Artificial Neural Network, a feed-forward model of the primate visual system.

- Images are first fed through a model of V1, then V2, then classified using a Support Vector Machine (SVM).
- V1 and V2 feature simple cells (S1 and S2), which respond to oriented bars and edges.
- V1 and V2 also feature complex cells (C1 and C2) cells, which correspond to striate complex cells.
- The simple cells project patches of the image into a useful basis.
- The complex cells group simple cells with similar orientations.
- The output of V2 is classified using a SVM.
- The bases used in S1 and S2, along with the optimal SVM parameters, are learned during training.

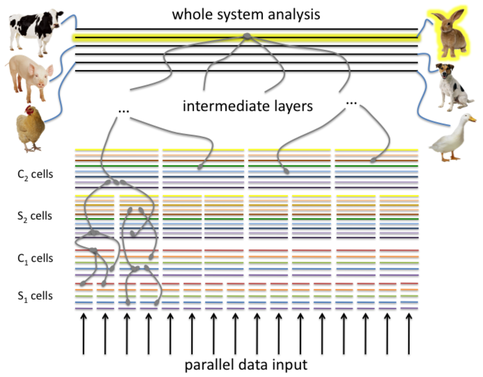

Figure 3: *Schematic of hierarchical feed-forward model of the primate visual cortex.*

2.3 Learned lateral connections

Following Geisler and Perry [?], we create a simple model for learning lateral connections in S1.

- During training, PANN S1 measures the response to Gabor filters.
- For each patch that responds (above a threshold), we measure co-occurance statistics for other patches as a function of distance and orientation.
- We use these statistics to create an excitation/inhibition kernel for lateral interactions.
- During testing, any S1 patch that responds above a threshold is given support based on the trained kernel.
- We this as a new, filtered S1 layer and interface with PANN.

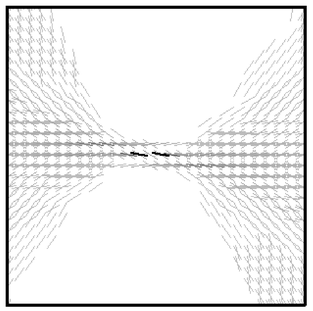

Figure 4: *A typical S1 kernel learned during training. Darker color indicates stronger excitation. Contrast has been enhanced for printing.*

3. Results

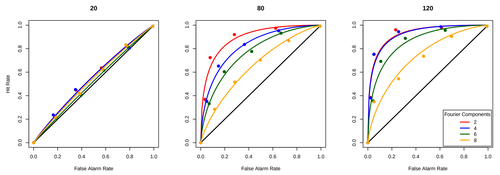

Figure 5: *Human performance for the amoeba/no amoeba task. Five different human test subjects were used. Receiver operating characteristic (ROC) curves are shown (points); the effectivee d' is calculated and idealized ROC curves based on this are plotted (solid lines).*

Humans perform well on this task for long SOA, but poorly for short SOA. This indicates that the mask we use effectively interrupts processing as expected. The lateral interactions addition to PANN significantly improves performance, which even exceeds that of humans for difficult amoebas.

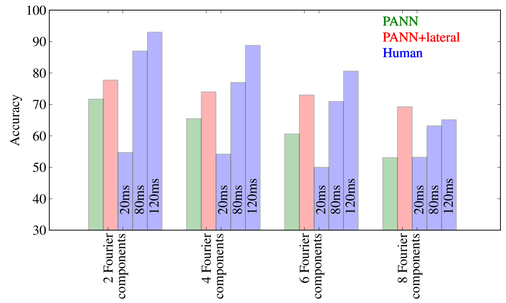

Figure 6: *A comparison of performance for the amoeba/no amoeba task.*

This work was supported by the the Los Alamos National Laboratory LDRD program under project 20090006DR.

References

[1] W. S. Geisler and J. S. Perry. Contour statistics in natural images grouping across occlusions. *Vis. Neurosc.*, 26(1):109–121, 2009.

[2] T. Serre, A. Oliva, and T. Poggio. A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, 104(15):6424–6429, 2007.

LA-UR 09-00042