NONLINEAR OPERATOR FOR BLOB TEXTURE SEGMENTATION

P. Kruizinga and N. Petkov

Institute of Mathematics and Computing Science, University of Groningen
P.O. Box 800, 9700 AV Groningen, The Netherlands
peterkr@cs.rug.nl, petkov@cs.rug.nl

1. INTRODUCTION

Texture is an important part of the visual world of animals and men and their visual systems successfully detect, discriminate and segment texture. Relatively recently progress was made concerning structures in the brain which are presumably responsible for texture processing. Von der Heydt et al. (von der Heydt et al. 1992) reported on the discovery of a texture processing neuron in areas V1 and V2 of the visual cortex of monkeys which they called *grating cell*. Grating cells respond vigorously to gratings of bars of appropriate orientation, position and periodicity. In contrast to other orientation selective cells, grating cells respond very weakly or not at all to single bars which do not make part of a grating. This behaviour of grating cells cannot be explained by linear filtering followed by half-wave rectification as in the case of simple cells, neither can it be explained by three-stage models of the type used for complex cells. Elsewhere we proposed a model of this type of cell and demonstrated the advantages of grating cells with respect to the separation of texture and form information (Kruizinga & Petkov 1995, Petkov & Kruizinga 1997). Tanaka et al. (1991) found another type of texture processing neuron, that responds to dot-patterns. These texture cells, which we call *blob-texture cells* in the following, have similar characteristics as grating cells. They do not react to single dots but only to a pattern consisting of a number of dots. Neurophysiological experiments revealed a preference of the cells for a regular dot pattern in comparison to more random patterns of dots. Grating cells are not activated by these random dot patterns, though regular dot patterns cause a slight grating cell response.

In this paper we propose a computational model of blob-texture cells that is capable of explaining the results of neurophysiological experiments. Our model of blob-texture cells consists of three consecutive stages. The final stage (blob-texture cells) receives its inputs from multiple units in the second stage, the so-called blob-pattern subunits, which in turn receive their input of the first stage (blob detectors).

Furthermore, the model is used as an image processing operator and compared with existing texture operators like the Gabor-energy operator and the cooccurrence matrix operator. This evaluation is done by comparing the results of a texture segmentation task, in which an image containing a number of blob textures is segmented on the basis of features obtained with the three texture operators. The method is similar to the evaluation of the grating cell operator, with respect to the processing of oriented texture (Kruizinga & Petkov 1998).

2. COMPUTATIONAL MODEL OF BLOB DETECTORS

Though most of the cells in the primary visual cortex (V1) are orientation selective, about 10-20% of the cells do not show any orientation preference. Most of these 'non-oriented cells' have a centre-surround receptive field profile (impulse response), which can be modelled by means of a Difference-of-Gaussians (DoG) function as follows:

$$u_{\xi,\eta,\sigma,\gamma}(x,y) = \frac{1}{2\pi\gamma^2\sigma^2} e^{\frac{-((x-\xi)^2+(y-\eta)^2)}{2\gamma^2\sigma^2}} - e^{\frac{-((x-\xi)^2+(y-\eta)^2)}{2\sigma^2}} \quad (1)$$

where x and y specify the position of a light impulse in the visual field and ξ, η, σ and γ are parameters as follows:

The centre of the receptive field within the visual field is specified by the pair (ξ, η). The parameters σ and γ specify the standard deviations

$\sigma_c = \gamma\sigma$ ($\gamma < 1.0$) and $\sigma_s = \sigma$ of the centre and the surround Gaussians, respectively. In our experiments we used a value of $\gamma = 0.5$. The normalisation factor $\frac{1}{2\pi\gamma^2\sigma^2}$ is used to obtain a function with a zero DC component. In our experiments we used two types of blob detectors: one with an excitatory, and the other with an inhibitory central region. The latter detects dark blobs on a light background. These cells are modelled by a two-stage model consisting of a first, linear filtering stage and a second, non-linear stage which includes thresholding and contrast normalisation. The linear stage consists of computing an integral:

$$s_{\xi,\eta,\sigma,\gamma} = \iint f(x,y)u_{\xi,\eta,\sigma,\gamma}(x,y)\mathrm{d}x\mathrm{d}y \qquad (2)$$

where $f(x,y)$ is the intensity distribution of the input image. In the second stage, contrast normalisation is performed by dividing the response $s_{\xi,\eta,\sigma,\gamma}$ by the average grey level in the image within the receptive field of the modelled cell. The average gray level $a_{\xi,\eta,\sigma}$ is computed as follows:

$$a_{\xi,\eta,\sigma} = \iint f(x,y)e^{-\frac{(x-\xi)^2+(y-\eta)^2}{2\sigma^2}}\mathrm{d}x\mathrm{d}y \qquad (3)$$

In order to implement contrast normalisation, we use the hyperbolic ratio function to calculate the output of the unoriented cell from the ratio $l_{\xi,\eta,\sigma,\gamma} = \frac{s_{\xi,\eta,\sigma,\gamma}}{a_{\xi,\eta,\sigma}}$ which is proportional to the local contrast within the receptive field of the cell:

$$v_{\xi,\eta,\sigma,\gamma} = \begin{cases} 0 & \text{if } a_{\xi,\eta,\sigma} = 0 \\ \chi\left(\frac{l_{\xi,\eta,\sigma,\gamma}\cdot R}{l_{\xi,\eta,\sigma,\gamma}+C}\right) & \text{otherwise} \end{cases} \qquad (4)$$

where $\chi(z) = 0$ for $z < 0$, $\chi(z) = z$ for $z \geq 0$ (thresholding) and R and C are the maximum response level and the semi-saturation constant, respectively.

The modelled unoriented cell will react strongly to a blob which is located entirely in the centre, excitatory, region of the receptive field, though the cell will also react to other features in its receptive field such as lines or edges. The function of an ideal blob detector is, however, to signal only blobs. Our computational model of blob detectors is based on the hypothesis that this separation of blob features from other image features is induced by a lateral inhibition mechanism. A blob detecting subunit $v'_{\xi,\eta,\sigma,\gamma}$ gets its input from the modelled unoriented cell $v_{\xi,\eta,\sigma,\gamma}$ and a number of similar modelled cells with the same preferred blob size but with the centre of their receptive fields located in the vicinity of (ξ,η). The concerned subunit has the same output as the unoriented cell with the same location (ξ,η) if the other cells show no response. This will only be the case if the unoriented cell signals a blob in its receptive field. Other image features may also invoke a reaction of the cell, but they will cause a reaction of nearby cells as well. In that case, the output of the blob detecting subunit $v'_{\xi,\eta,\sigma,\gamma}$ is influenced by the outputs of nearby unoriented cells, in such a way that if at least one of these cells reacts, the subunit response is inhibited, *i.e.* the response is set to zero. In our model the lateral inhibition scheme involves a fixed number of nearby unoriented cells lying in a circle around the centre of the receptive field (ξ,η):

$$v'_{\xi,\eta,\sigma,\gamma} = \begin{cases} v_{\xi,\eta,\sigma,\gamma} & \text{if } \forall i,\ i \in \{1..N\}, \\ & v_{\xi,\eta,\sigma,\gamma}(\xi+\Delta\xi_i,\eta+\Delta\eta_i) < \rho v_{\xi,\eta,\sigma,\gamma} \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

$$\Delta\xi_i = R_{\mathrm{lat}}\cos\left(\frac{2\pi i}{N}\right) \qquad (6)$$

$$\Delta\eta_i = R_{\mathrm{lat}}\sin\left(\frac{2\pi i}{N}\right) \qquad (7)$$

where ρ is a fixed fraction (in our experiments we used $\rho = 0.8$) of the response $v_{\xi,\eta,\sigma,\gamma}$ and R_{lat} is the distance between the centre of the considered cell and the nearby cells ($R_{\mathrm{lat}} = 1.36\sigma$). The number of nearby cells that are involved in the inhibition process is set to $N = 15$. This value is high enough to guarantee that the blob-detectors do not react to features other than blobs.

In our experiments we use blob detectors with different values of σ to enable the detection of blobs on different scales. This introduces however a redundancy in the coding of blob localisation since blob-detectors on more than one scale, but with the receptive field centred at the same position, may react to a blob in their receptive field. Not only the blob detector with the appropriate size will react, but all blob detectors with larger values of σ (given the fact that no other image features appear in the inhibitory region of the larger receptive field) will show a response. The redundancy can be eliminated by suppressing all outputs of non-optimal blob-detectors at the same position. This is implemented by a winner-takes-all mechanism across all blob detectors with the same receptive field centre

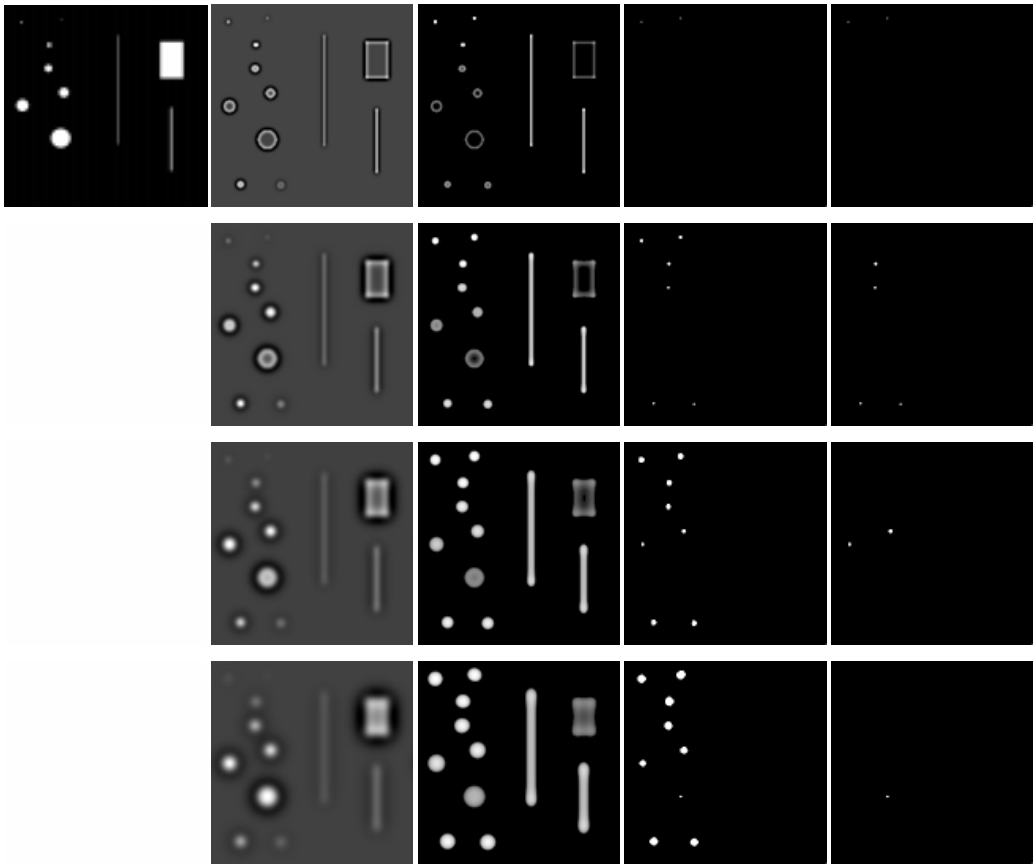

Figure 1: The single image in top-left position is a synthetic input image. The images in the first column are the results of convolutions with centre-surround kernels of four different sizes (*a*), the results are normalised for contrast (*b*), the modelled cell responses are inhibited by the responses of nearby modelled cells (*c*) and modelled cells with the same centre of the receptive field but with different sizes compete with each other in a winner-takes-all scheme (*d*).

but with different values of the size parameter σ. The output of a blob detector $\tilde{v}_{\xi,\eta,\sigma,\gamma}$ is computed as follows:

$$\tilde{v}_{\xi,\eta,\sigma,\gamma} = \begin{cases} v'_{\xi,\eta,\sigma,\gamma} & \text{if } v'_{\xi,\eta,\sigma,\gamma} = \max_{\sigma'}(v'_{\xi,\eta,\sigma',\gamma}) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

The winner-takes-all mechanism will cause the information concerning the location of blobs in the image to be separated into different channels, depending on the size of the blobs. The sensitivity of the blob-texture cells to blobs with different sizes will therefore depend on the sampling on the scale-range. In our experiments we used four different scales.

The model as it is presented above will detect blobs of a specific size in the visual field, independent of the contrast and discarding all other image features as lines and edges. The visual information processing by the modelled blob-detectors is illustrated in Fig. 1 together with the results at intermediate stages. The input image (upper-left image) is convolved with centre-surround DoG profiles with four different sizes. As can be seen from the resulting image (Fig. 1a), the modelled DoG cells react to blobs, but also to other image features. Furthermore, the strength of the response depends on the local contrast of the features. The results are than normalised in order to get contrast independency (Fig. 1b). To ensure that only blobs are detected, the modelled cell response is inhibited by

responses of neighbouring cells at the same scale (Fig. 1c). Finally, to make sure that there will be a response at only one scale at every position in the visual field, a winner-takes-all mechanism inhibits all sub-maximum responses (Fig. 1d).

3. COMPUTATIONAL MODEL OF BLOB-TEXTURE CELLS

In the second stage of our model, the outputs of the blob detectors are combined by so-called blob-pattern subunits using an AND-type nonlinearity. In the final stage, the actual blob-texture cells sum the responses of a large number of blob-pattern subunits in the vicinity of their receptive field centre.

This means that modelled blob-texture cells will only react if a number of blobs with a specific size are present in the receptive field of the cell. The response is dependent on the number of blobs up to a given maximum. This model of blob-texture cells is next explained in more detail.

The activity of a so-called blob-pattern subunit, $t_{\xi,\eta,\sigma,\gamma,\zeta}$, with position (ξ,η) and with preferred blob size specified by σ, is calculated as follows:

$$t_{\xi,\eta,\sigma,\gamma,\zeta} = \begin{cases} 1 & \text{if Card}\{\tilde{v}_{\xi+\Delta\xi_i,\eta+\Delta\eta_i,\sigma,\gamma} : \\ & i = 1\ldots n : \\ & \tilde{v}_{\xi+\Delta\xi_i,\eta+\Delta\eta_i,\sigma,\gamma} > 0\} > 3 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where the position of the of the respective blob detector cells is taken at random within the neighbourhood of (ξ,η):

$$\begin{aligned} \Delta\xi_i &= (\sigma\zeta + r_i)\cos\alpha_i \\ \Delta\eta_i &= (\sigma\zeta + r_i)\sin\alpha_i \ , \ i = 1\ldots n \end{aligned} \tag{10}$$

where $\sigma\zeta$ is a fixed radius (ζ specifies the spatial spreading of the blobs in the pattern) and r_i are random numbers taken from a normal distribution with zero mean and standard deviation 0.5 and α_i are random numbers taken from a uniform distribution between 0 and 2π. The number of locations n within the receptive field that is taken into account with the determination of blob-texture presence in the receptive field is larger than the number of blobs to be detected. In our experiments we set the number of inspected locations to $n = 30$. Only if three or more blobs were detected in these 30 locations, the blob-pattern subunit is activated.

Finally, the response a blob-texture cell $b_{\xi,\eta,\sigma,\gamma,\zeta}$, which is centred at position ξ,η in the visual field and has preferred blob size specified by σ, is calculated by weighted summation of the blob-pattern subunits.

$$b_{\xi,\eta,\sigma,\gamma,\zeta} = \iint e^{-\frac{(\xi-\xi')^2+(\eta-\eta')^2}{2\beta\sigma^2}} t_{\xi,\eta,\sigma,\gamma,\zeta} \mathrm{d}\xi' \mathrm{d}\eta' \tag{11}$$

The parameter β specifies the size of the region in which the weighted summation takes place. Larger values of β result in a uniform response in a blob-texture area even with larger discontinuities in the blob pattern.

4. TEXTURE OPERATOR EVALUATION

The quantities computed with the blob-texture cell operators can be used as texture features. We next compare the following set of features:

- **Blob-texture cell operator features:**
  A set of blob-texture cell operators with four different preferred blob sizes, three values of the spatial spreading and selective for both black and white blobs, is applied to an image, yielding a vector of 24 features in each point.

- **Gabor-energy features:**
  A popular set of texture features is based on the use of Gabor filters (Jain & Farrokhnia 1991). In this case, an image is filtered with a set of Gabor filters with different orientations, spatial frequencies and phases. Using eight orientations and three preferred spatial-frequencies and combining the results of symmetric and antisymmetric filters, this multi-channel filtering scheme yields a feature vector of 24 Gabor-energy quantities. The values of the preferred orientations and spatial-frequencies are taken to ensure a good coverage of the spatial-frequency domain.

- **Cooccurrence matrix features:**
  A classic method for texture segmentation is based on the gray-level cooccurrence matrices (Haralick et al. 1973). In each point of a texture image, a set of gray-level cooccurrence matrices is calculated for different orientations and inter-pixel distances. From these matrices, a number of features is extracted which characterise the neighbourhood of the concerned pixel. In our experiments eight

a) b) c) d) e)

Figure 2: Results of a segmentation experiment using the K-means clustering algorithm. The left-most image (*a*) shows the input image containing nine blob textures, of which the perfect segmentation (ground truth) is given in image (*b*). In the next three images, the segmentation results are shown based on the use of blob-texture cell operator features (*c*), Gabor-energy operator features (*d*) and cooccurrence matrix features (*e*).

gray-level cooccurrence matrices were calculated in each point using a neighbourhood of size 12×12. From each of the matrices three features (energy, inertia and entropy) were extracted resulting in a vector of 24 features in each image point.

In order to evaluate the quality of the features obtained with these three texture operators, with respect to texture segmentation, an image containing nine blob textures is segmented by means of the general purpose K-means clustering algorithm. First, a given texture operator is applied to the input image yielding a field of 24-dimensional feature vectors. The clustering algorithm then allocates the feature vectors to one of the k clusters. Figure 2 shows the result of this segmentation experiment using the three texture operators. As can be seen from this image, the segmentation based on the blob-texture cell operator features is better than the segmentations based on the features of the other two operators. Only at the texture borders pixels are misclassified.

5. REFERENCES

Haralick, R., Shanmugam, K. & Dinstein, I. (1973), 'Textural features for image classification', *IEEE Transactions on Systems, Man and Cybernetics* **3**(6), 610–621.

Jain, A. & Farrokhnia, F. (1991), 'Unsupervised texture segmentation using gabor filters', *Pattern Recognition* **24**(12), 1167–1186.

Kruizinga, P. & Petkov, N. (1995), A computational model of periodic-pattern-selective cells, *in* J. Mira & F. Sandoval, eds, 'Proc. IWANN '95', Lecture Notes in Computer Science, vol.930, Springer-Verlag, pp. 90–99.

Kruizinga, P. & Petkov, N. (1998), Grating cell operator fetures for oriented texture, *in* A. Jain, S. Venkatesh & B. Lovell, eds, 'Proc. of the Int. Conf. on Pattern Recognition', Brisbane Australia, pp. 1010–1014.

Petkov, N. & Kruizinga, P. (1997), 'Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells', *Biological Cybernetics* **76**(2), 83–96.

Tanaka, K., Saito, H., Fukada, Y. & Moriya, M. (1991), 'Coding visual images of objects in the inferotemporal cortex of the macaque monkey', *Journal of Neurophysiology* **66**(1), 170–189.

von der Heydt, R., Peterhans, E. & Dürsteler, M. (1992), 'Periodic-pattern-selective cells in monkey visual cortex', *Journal of Neuroscience* **12**, 1416–1434.