Band-limited noise suppresses contour perception not only of letters but of any objects

Nicolai Petkov and Michel A. Westenberg

Institute for Mathematics and Computing Science,
University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
`petkov@cs.rug.nl`

Introduction

- Band-limited noise suppresses the recognition of letters.
- High and low frequencies mask the text less effectively than intermediate frequencies, Fig. 1.
- In [2], the effect was related to the size of the letters: noise at 3 cycles per letter was found to be most effective.

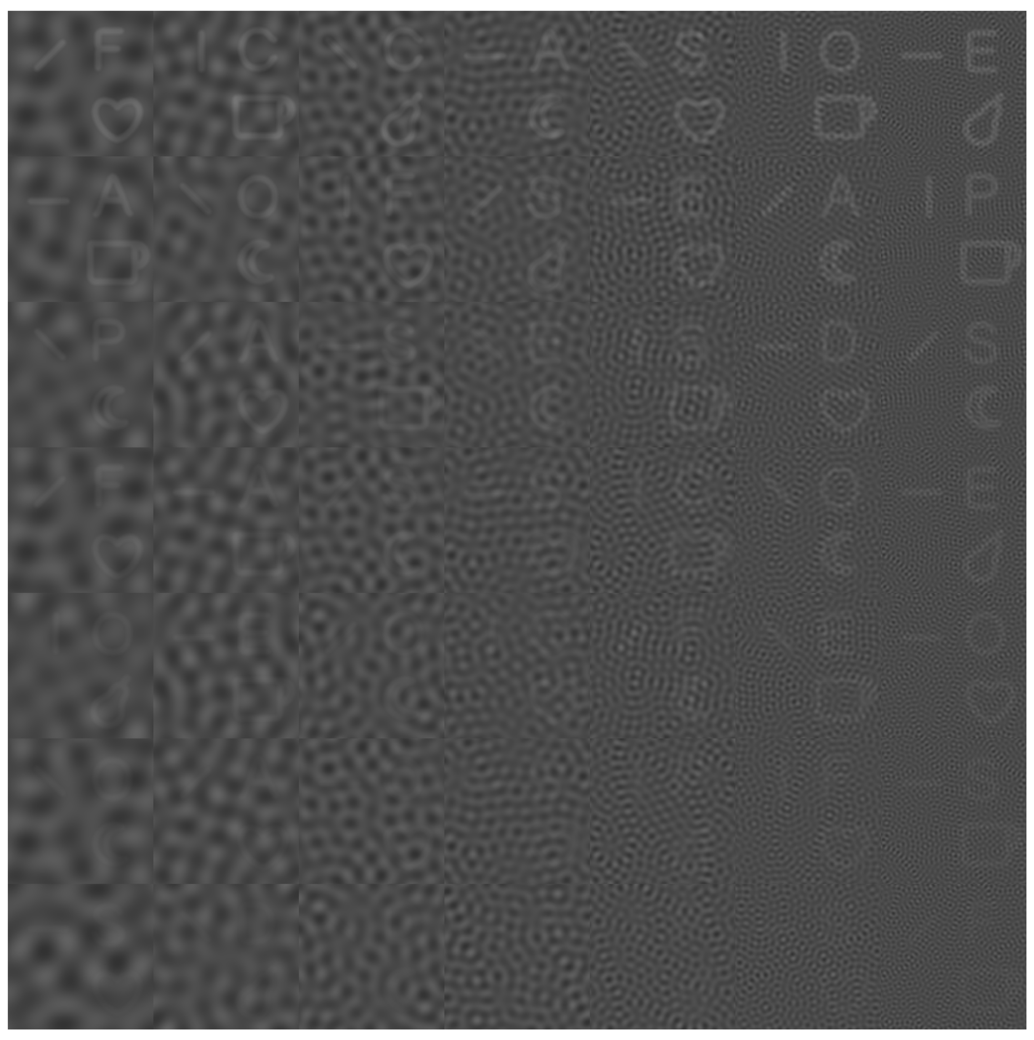


Figure 1: *The masking effect of band-spectrum noise on the perception of objects does not depend on the type of visual stimuli embedded in the noise: the same effect is observed simultaneously for letters, icons of familiar objects and bars. Vertically, the contrast ranges from 0.05 to 0.25. Horizontally, the noise frequency ranges from 0.38 to 3.05 cycles per degree (cpd) when a single subimage is viewed on a computer screen from a distance of 0.55 m.*

Open questions

- Is the effect specific to text?
  Answered by psychophysical experiments.
- What is the origin of the effect in terms of underlying neural processes?
  Answered by computational models.

Experimental setup

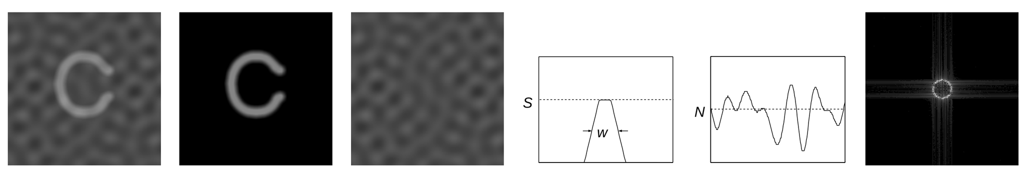


Figure 2: *(a) A test image is the sum of (b) an image of a light letter on a black background and (c) a noise image. The noise image is a superposition of a constant component of luminance N and 100 sine waves of randomly selected different orientations and spatial frequencies within a narrow band. (d) The letter contour width w is measured at half height of the maximum luminance S of the trapezoid luminance profile of the stroke. (e) Characteristic 1-D luminance profile of the noise in the space domain and (f) its 2-D spectrum.*

- Images had visual angle dimensions of $14^\circ \times 14^\circ$.
- Observer was presented a series of images of constant letter size and contour width.
- 5 letter contrasts and 11 noise spatial frequencies were used.
- 20 images were presented for each point in the spatial frequency vs. luminance contrast space.
- One of the letters A, C, D, E, F, O, P, R, S or no letter (blank) was presented; all represented with equal probability.
- The 1100 images were presented in random order to prevent accommodation to a specific spatial frequency, luminance contrast, or letter.
- The images were presented for one second after which the observer indicated the letter seen.
- Experiments were carried out for different letter sizes and contour widths.
- The experiments were also carried out with object icons and with bars of different orientations.
- The experiments were conducted with two observers.

Results of psychophysical experiments

The suppression effect is simultaneous for letters, bars, and object icons. This suggests a relation between a low-level interaction of the band-limited noise with the object contours rather than with the object as a whole. The stroke width of the contour was found to determine the frequency domain of most effective inhibition. The object size has an effect only on the threshold contrast at which an object can be recognized, Fig. 3.

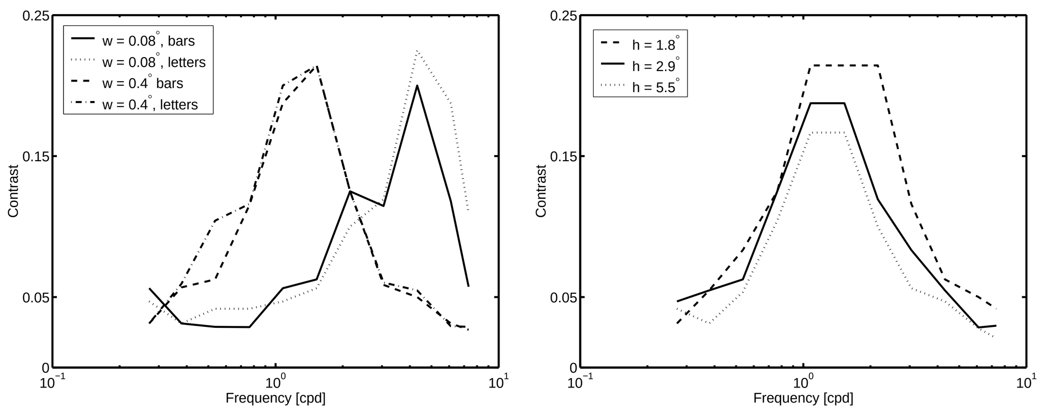


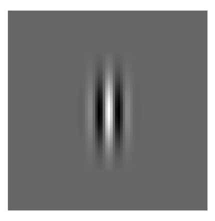
Figure 3: *(a - Observer 1) Bar and letter contrast needed for 90% recognition rate as a function of the spatial frequency of the noise. Stroke width is denoted by w. (b - Observer 2) Bar contrast needed for 90% recognition rate as a function of the spatial frequency of the noise; h denotes the bar length. Bar width was kept constant at $w = 0.4^\circ$. The same results were obtained for letters and object icons.*

Computational model

Simple cell

Gabor functions model the receptive fields of simple cells:

$$g_{\lambda,\sigma,\theta,\varphi}(x,y) = e^{-\frac{\tilde{x}^2+(\gamma\tilde{y})^2}{2\sigma^2}}\cos\left(2\pi\frac{\tilde{x}}{\lambda}+\varphi\right)$$
$$\tilde{x} = x\cos\theta + y\sin\theta$$
$$\tilde{y} = -x\sin\theta + y\cos\theta$$



where γ is the spatial aspect ratio, σ determines the size of the receptive field, λ is the preferred wavelength, θ is the preferred orientation, and φ determines the symmetry.
The response of a simple cell to a stimulus $f(x,y)$ is computed by convolution:

$$r_{\lambda,\sigma,\theta,\varphi}(x,y) = (f * g_{\lambda,\sigma,\theta,\varphi})(x,y)\ .$$

Complex cell

The response of a complex cell is a combination of the responses of two simple cells with a phase difference of $\pi/2$:

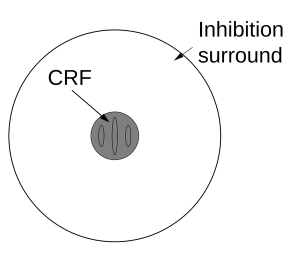
$$E_{\lambda,\sigma,\theta}(x,y) = \sqrt{r^2_{\lambda,\sigma,\theta,0}(x,y) + r^2_{\lambda,\sigma,\theta,-\frac{\pi}{2}}(x,y)}\ .$$

Figure 4: *(a) An input image. (b) Responses of complex cells with (b) vertical and (c) horizontal preferred orientation. (d) The maximal responses across all orientations.*

Non-CRF inhibition

The surround of the CRF is defined by the weighting function $\mathrm{w}_\sigma(x,y)$:

$$\mathrm{w}_\sigma(x,y) = \frac{1}{\|H(\mathrm{DoG}_\sigma)\|_1}H(\mathrm{DoG}_\sigma(x,y))$$
$$H(z) = \begin{cases} 0 & z < 0 \\ z & z \geq 0 \end{cases}$$
$$\mathrm{DoG}_\sigma(x,y) = \frac{1}{2\pi(4\sigma)^2}e^{-\frac{x^2+y^2}{2(4\sigma)^2}} - \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$$



We compute an inhibition term over the annular area by a convolution of $\mathrm{w}_\sigma(x,y)$ with the responses of complex cells with a preferred orientation, wavelength, and CRF:

$$t_{\lambda,\sigma,\theta}(x,y) = (E_{\lambda,\sigma,\theta} * \mathrm{w}_\sigma)(x,y)$$

The response of a complex cell with anisotropic surround inhibition is then computed as

$$\tilde{b}^A_{\lambda,\sigma,\theta}(x,y) = H(E_{\lambda,\sigma,\theta_i}(x,y) - \alpha t_{\lambda,\sigma,\theta}(x,y))\ ,$$

where α controls the strength of the inhibition [1].

Results of computer experiments

We applied the computational model to the test images we used in the psychophysical experiments, and found that the performance of the operator correlates very well with the perceptual sensitivity of humans.

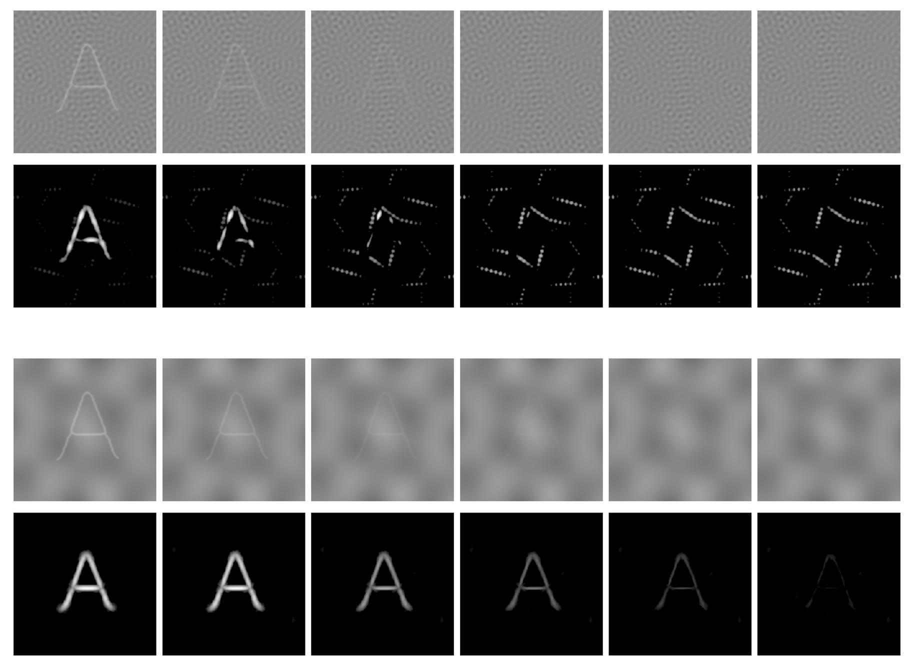


Figure 5: *Input stimuli and corresponding responses of the anisotropic inhibition operator. All input images contain the letter A with a stroke width $w = 0.08^\circ$. The input images of the first row contain noise with a spatial frequency of 4.3 cpd. The noise cannot effectively be removed by the operator without removing the letter contour as well, and already for $C = 0.25$ the letter is unrecognizable in the output. The third row shows input images with superimposed noise with a spatial frequency of 0.5 cpd, and the fourth row shows the corresponding responses. The operator effectively removes the background noise, while letting the letter contour pass, except for the bottom-right image in which the computed cell response reduces to almost zero.*

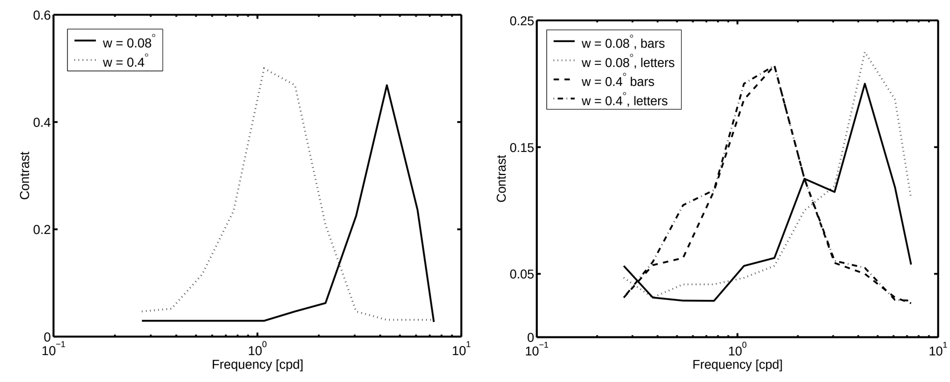


Figure 6: *(a) Letter contrast as a function of the spatial frequency of the noise. The plot shows letter contrast for which the anisotropic inhibition operator filters out noise and allows a sufficient part of the letter contour to pass, so that the letter can be recognized (automatically or by a human observer) in the output of the operator in 90% of the cases. The performance of the operator correlates almost perfectly with the psychometric curves shown in (b).*

Conclusions

- The width (and *not* the size) of a contour determines the frequency domain of most effective inhibition.
- The size of the object only scales the contrast sensitivity.
- The inhibition effect is also observed for single bars and object icons.
- The effect need not be attributed to some higher cognitive ability, such as reading.
- The effect can be attributed to properties of the visual system at a low level at which bars, lines and contours are detected.

References

[1] N. Petkov and M. A. Westenberg. Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition. *Biological Cybernetics*, 88(3):236–246, 2003.

[2] J. A. Solomon and D. G. Pelli. The visual filter mediating letter identification. *Nature*, 369:395–397, 1994.

ECVP 2004, 27th European Conference on Visual Perception, 22–26 August 2004, Budapest, Hungary

RuG