Non-linear processes in perceived region segregation: orientation selectivity of complex channels

Norma Graham, Anne Sutter, Charu Venkatesan and Maria Humaran

Department of Psychology, Columbia University, New York City, NY 10027, USA

(Received 27 September 1991)

Models incorporating linear spatial-frequency- and orientation-selective channels explain many aspects of visual texture segregation. The inability of such models to fully explain texture segregation results, indicates that non-linear processes are also involved. One non-linearity that has been suggested is complex channels consisting of two stages of linear filtering separated by a rectification-type non-linearity (much like cortical complex cells). Here we further demonstrate the usefulness of complex channels in explaining texture segregation results and investigate the orientation-selectivity of the first stage of such complex channels. Our results suggest that the first stage is much more selective for orientation than are lateral geniculate nucleus cells, but that the first-stage orientation bandwidth is rather wide with some interaction occuring between perpendicular orientations.

An important, relatively low-level process in visual perception may be region segregation; that is, there may be an early stage in which the visual system breaks a visual scene into spatial parts (meaningful regions) before going on to do higher-level computations. The laboratory task of texture segregation may depend largely on this low-level region-segregation process. (For a thorough review of previous literature on texture segregation, see reference 1) Spatial-frequency- and orientation-content are important determinants of perceptual texture segregation, and models incorporating spatial-frequency- and orientation-selective channels explain many texture-segregation results. (The response of a linear channel to a compound stimulus is the sum of its responses to the stimulus components.) It has become clear, however, that linear channels, by themselves, are insufficient to explain segregation. Thus satisfactory models must include non-linear processes. However, little is yet known about the properties of these non-linear processes. We have previously proposed the existence of at least two kinds of non-linear processes; a spatial non-linearity and an intensity-dependent non-linearity[2,3]. Others have proposed alternative processes that may perform quite similar functions[4–6]. Other work suggests what may be still a third kind of non-linear process that operates over larger distances[7,8]; here we review briefly simple linear models and then go on to explore the properties of the proposed spatial non-linearity, namely complex channels.

Simple-channel model

In our simple-channel model of region (texture) segregation[9], the first part of the model is a set of multiple, linear, spatial-frequency and orientation-selective channels. The second part of the model is a higher level process that compares the channels' aggregate activity in response to different spatial regions of the pattern, and asks whether the outputs of channels differ from one region to another. These simple-channel models are approximate embodiments of Julesz's original statistical conjecture about texture segregation[10], put into the context of spatial-frequency sensitivity channels first suggested by Campbell and Robson[11,12]; the class we consider contains many models suggested by other investigators. Chubb and Landy[13], have referred to models like these as backpocket models of texture segregation, that is, they are the models numerous people pull out of their back pockets to account for texture segregation. As many people have shown, this class of simple-channel models can account for many phenomena of texture segregation.

Figure 1 shows examples of patterns containing two textures, each texture composed of the same two element types but differently arranged (a checkerboard in the middle and stripes in the left and right). In the top pattern, only one element type is visible, as the other element type has a contrast of zero. We will call this a one-element-type-only pattern. The elements composing the patterns in *Figure 1* are balanced, that is, their space-average luminance is equal to the background luminance. According to the simple-channel models, all element-arrangement textures made from balanced elements should segregate

---

Correspondence to: Dr Norma Graham

0275-5408/92/020142-05

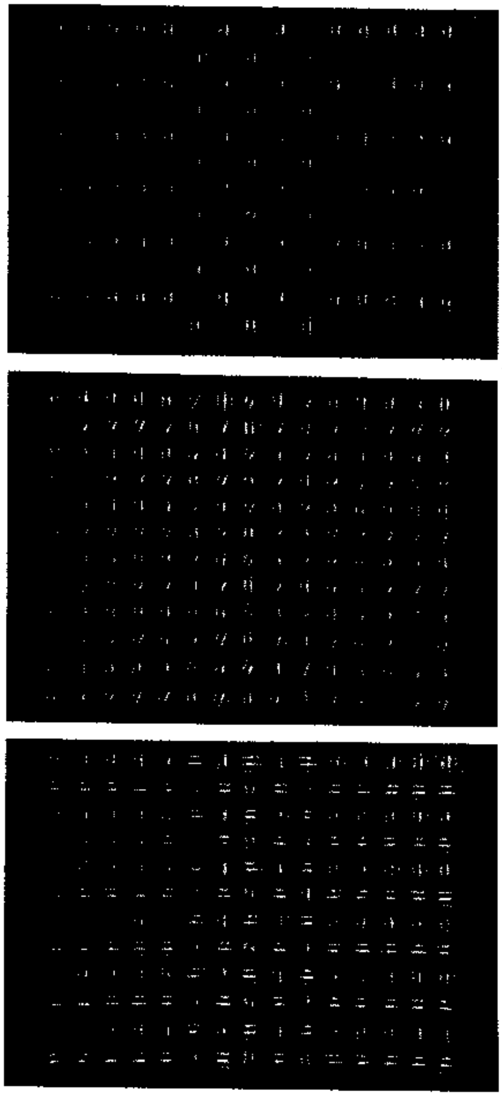

Figure 1 Examples of element-arrangement pattern where the elements are gratings. In the top panel only one type of element (a grating of vertical orientation) is visible since other element has zero contrast. In the middle and bottom panels the two types of elements are of the same contrast and differ in orientation by 22.5 and 90 degrees of rotation respectively. In the stimuli used in the experiments, the luminance averaged across each grating was the same as the background luminance. Reproduction will have distorted the stimuli somewhat in these figures, however

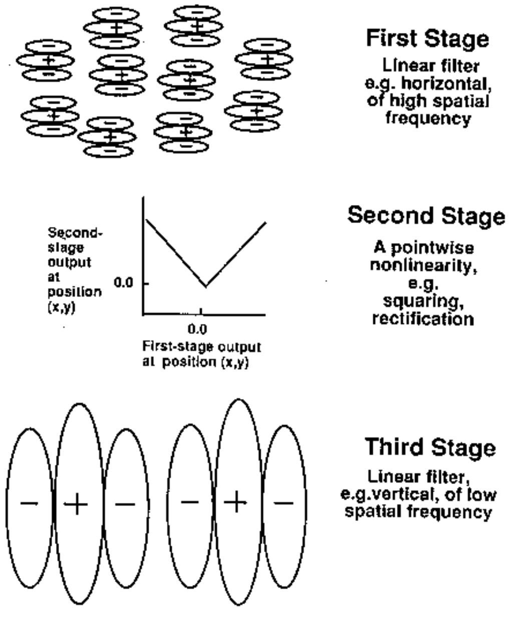

Figure 2 Diagram of complex channel

poorly, if at all. The intuition behind this prediction is explained at some length elsewhere[3], although it is applied there to somewhat different patterns. Briefly, the intuition is as follows. Only the filters tuned to high spatial frequencies respond to patterns composed of balanced elements. These high-spatial-frequency filters respond with a mixture of positive and negative responses to much the same degree in both the checkerboard and the striped regions (although the precise spatial arrangement of the response differs in the two regions). The second part of our simple-channels models (or the analogous stage in all backpocket models of texture segregation) then pools over local regions (i.e. as in computing the standard deviation of the responses at different nearby positions), thus erasing the differences in precise arrangement between the responses in the checkerboard and the striped regions. Thus the simple model predicts that these balanced-element patterns should segregate very poorly.

Yet, many such textures segregate very well indeed (e.g., one-element-type-only patterns like the top pattern in *Figure 1*). This result, along with other failures of simple-channel models, suggests the existence of complex channels or very similar calculations[2,3,9,14–19].

Complex-channels models

In the complex-channel model, the simple channels are augmented by (or perhaps even replaced by) complex channels like that in *Figure 2*. Each complex channel consists of two stages of linear filtering separated by a rectification-type non-linearity that is dramatic near zero. This structure is like that suggested for complex cells in the cortex[20]. The final part of the complex-channel model (the pooling rules for comparing channel outputs in different regions) is identical to the analogous part of the simple-channel model. Why such complex channels predict segregation with balanced-element patterns is explained more fully elsewhere[3]. Briefly, however, such channels can respond to low-frequency arrangements of high spatial frequency elements (as in amplitude modulated radio transmission).

To date, little is known in detail about these complex channels, although one study[19] does suggest that, in any particular complex channel, the first stage is maximally sensitive to spatial frequencies three or four octaves higher than those to which the second filtering stage is maximally sensitive (i.e., the receptive fields at the second stage are 8–16 times larger than those at the first stage).

If the physiological substrate for the first stage of complex channels is a set of lateral geniculate nucleus (LGN) cells, the first stage would show little orientation preference (since LGN receptive fields are approximately concentric). If, however, the physiological substrate for the first stage is a set of simple cortical cells (as implied by the diagram in *Figure 2* where the receptive fields are elongated), the first stage should show considerable selectivity for orientation. Here we report experiments designed to measure the first-stage's orientation selectivity.

Experiment

Observers rated the degree of segregation between the checkerboard and striped regions in patterns like those of *Figure 1* where the two element types were both patches of sinusoidal grating differing in orientation.

To understand the reasoning behind this experiment, consider a pattern where the two orientations of grating patches are very different (i.e. the bottom panel of *Figure 1*). Suppose that the first-stage filters of complex channels are not oriented. Then the pattern should not segregate since the same complex channels are responding to both elements. (The second-stage of complex channels spatially pools the rectified outputs from the first stage and thus erases differences between the elements if the first stage responds to both element types.)

If, however, the first-stage filters are orientation-selective, these two element types (of very different orientation) will stimulate entirely different subsets of complex channels. Hence, from the point of view of any one complex channel, this pattern will be a one-element-only type pattern. And the second-stage filter will respond quite differently in the striped and checkerboard regions. Therefore the pattern is predicted to segregate very well indeed.

More generally, the degree of orientation selectivity of the first-stage filter can be measured by varying the difference between the orientations of the two elements.

Methods and procedures

For the experiments reported here, element 1 was always a patch of vertical grating. Element 2 generally differed from element 1 in orientation. Its orientation could be one of seven different values varying from vertical to horizontal. *Figure 1* shows cases where element 2 has contrast zero (top panel), has the same contrast as element 1 but is oriented at 22.5 degrees clockwise from vertical (middle panel), has the same contrast as element 1 and horizontal orientation (bottom panel).

Details of experiment. The elements were Gabor patches with a concentric Gaussian window having a half-width half-height of 8 pixels (0.25 degrees at a viewing distance of 0.91 m) truncated at ±16 pixels so as not to overlap with the neighbouring elements. The centre-to-centre spacing between neighbouring elements was 32 pixels (1 degree); thus the fundamental frequency of the striped region (one period of which consists of two rows of elements and two rows of inter-element spaces) was 0.5 per degree. The numbers, spacing, and arrangements of the elements can be seen in *Figure 1*.

The spatial frequency of the grating patches used in the experiments was 8 c/deg (a period of 4 pixels). (To ensure their visibility, the period used in the grating patches in *Figure 1* was 8 pixels; thus each patch in *Figure 1* contains only half as many cycles as in the experiments reported here).

Values of the base contrast for each orientation were roughly equated for visibility on the basis of pilot experiments with each observer. When the element was vertical, the base contrast was 5%.

Two types of session were used. In one type, the contrast of element 1 could be one of three equally-spaced contrasts (zero, three, or six times the base contrast, see definition of base contrast below); the contrast of element 2 could be one of ten equally-spaced levels including zero where the lowest non-zero level equalled the base contrast. In the other type of session, the contrast of both elements could be one of four equally-spaced levels including zero where the lowest non-zero level was the base contrast. In this second type of session, the second element could be present in one of two phases 180 degrees apart (with either a dark or a light bar just to the left of centre.) Either type of session contained two replications of every stimulus, and two sessions were used in an experiment (four replications of each stimulus). Two observers (M.H. and C.V.) ran in both types of session. The orientation-selectivity estimated from both types of session was very similar (much closer than the differences between subjects illustrated in *Figure 3*) and so the results from both types have been averaged together. The two other subjects only ran in the second type of session.

Each trial started when the subject pressed the appropriate part of a response device (an 'unmouse'). The stimulus was presented for 1s with abrupt onset and offset. A 1 s delay was then enforced after which a beep signalled to the observer that a response would be accepted. The observer indicated the degree of perceived segregation by pressing the appropriate position within a four-inch rectangle on the response device (where the left edge indicated minimal and the right edge maximal perceived segregation). The patterns were generated by a Macintosh IIci on a standard Apple Monitor.

Estimating channel sensitivity

Precise estimates of the sensitivity function of a complex channel's first stage necessarily depend upon many assumptions of the model. Here we rely on an approximation shown in the inset of *Figure 3*. For the curve shown in the inset of *Figure 3*, the contrast of element 1 was held constant at a value greater than zero while the contrast of element 2 was varied as indicated on the horizontal axis. The observer's rating is shown on the vertical axis. Let r (for range) be the amount of segregation when element 1 (the vertical grating element) is the only type of element present. Let d (for dip) be the maximal amount by which the segregation is decreased due to the presence of the second element.

Note that d/r is intuitively a measure of how much the second element can interfere with the first element. In the complex-channels model, such interference occurs when both elements excite the same complex channels to some extent (because the channels' first stage is somewhat sensitive to both).

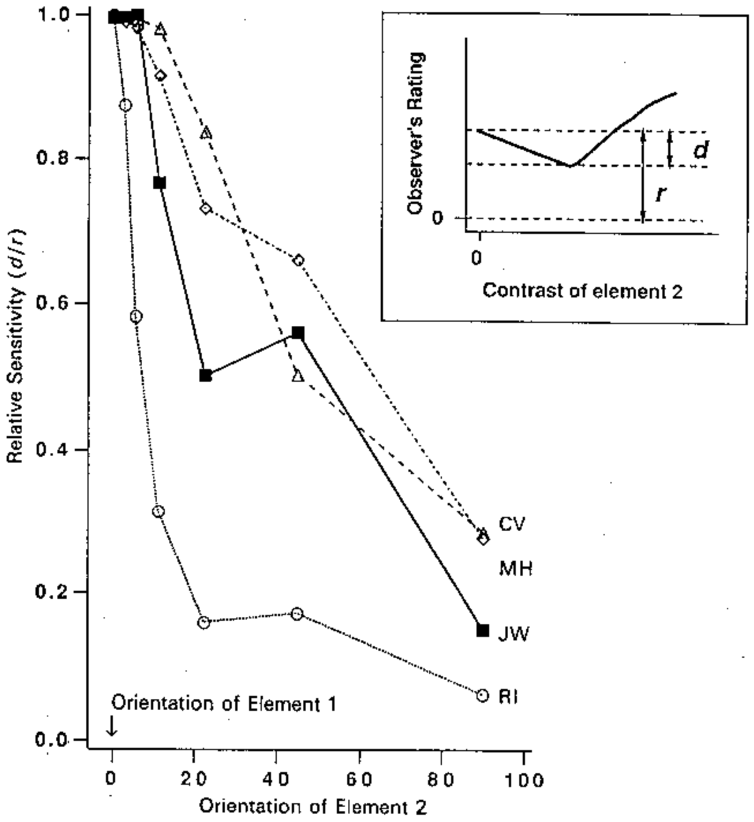

Figure 3 Estimate of orientation sensitivity of first stage of complex channel. Vertical axis of main figure is the d/r ratio as shown in the inset. See text for further explanation

More formally, for at least one class of reasonable model (Graham and Sutter, in preparation), it can be shown that:

$$\frac{d}{r} \approx \frac{S_A(2)}{S_A(1)} \times \frac{S_{obs}(1)}{S_{obs}(2)}$$

where Channel A is a channel in which the first stage is more sensitive to element 1 than to element 2; $S_A(i)$ is the sensitivity of Channel A to a pattern in which only element i is present (that is, it equals the reciprocal of the contrast which produces a response of magnitude 1.0 from Channel A); and $S_{obs}(i)$ is the sensitivity of the observer to a pattern in which only element i is present (that is, it equals the reciprocal of the contrast which produces a criterion-sized response of magnitude 1.0 from the observer).

The relative sensitivities reported here should be treated with caution for several reasons. The model leading to the above equation is incomplete; among other things it ignores the intensity-dependent non-linearities clearly displayed in previous work with patterns of this general type[2,3]. Also, for the results reported here, the value of d was estimated from the minimal rating in a curve rather than by fitting the model predictions to the whole curve.

Results and conclusions

As shown above, d/r is an approximate measure of the relative sensitivity (weighted by the contrast sensitivity) of the complex channels' first stage. These values are plotted for four different observers in the main part of *Figure 3*. As mentioned above, these values should be treated with some caution. Nevertheless, it seems clear that the first stage of complex channels is much more narrowly tuned for orientation than are LGN cells since, for all four observers, the channels were not nearly as sensitive to horizontal as to vertical.

For at least three of the four observers here, however, definite interaction occurs between orientations 45 degrees apart, and some interaction occurs even for perpendicular orientations. Indeed, the orientation bandwidth at half-peak-sensitivity is about 45 degrees for three of the four observers. For the complex-channel model of *Figure 2*, where all the receptive fields at a given stage are identical except in position, the first-stage bandwidth estimated here is that characterizing the single receptive fields composing the first filtering stage. Roughly speaking, the wider the bandwidth, the less elongated the excitatory and inhibitory regions in that receptive field. Perhaps, however, the first-stage filter's receptive fields differ somewhat from each other in orientation; perhaps receptive fields of several different orientations (centred at approximately the same position) feed into each second-stage neurone. If so, the first-stage bandwidth estimated here would characterize the conglomerate composed of several different orientations of receptive fields; thus it would be wider than the bandwidth characterizing any one of the receptive fields.

The bandwidth estimated here for three of the four observers is considerably broader than that estimated in many other kinds of psychophysical experiments (particularly those done with near-threshold contrast, as reviewed in Graham[21]. These other experiments may well be measuring the bandwidth of simple channels and/or of the receptive fields at the second stage of certain complex channels (perhaps somewhere the first and second stages of filtering are quite similar). On the other hand, the bandwidth from the fourth observer is a good deal narrower and within the range of estimates from these other experiments. We plan to explore these individual differences further by running this experiment on other subjects as well as doing further experiments on the subjects reported here.

The psychophysical results reported here suggest that the physiological substrate for the first stage of complex channels involved in region segregation is not LGN cells, but it might be simple cells in cortical area V1. The substrate for the second stage might be the complex cells in V1. It is still too early to be sure, however, that the physiological substrate is not higher in the cortex.

Acknowledgments

We are grateful to Hugh Wilson for giving us the copies of the Pascal programs which we modified to do the experiments reported here. The research was supported by USA National Eye Institute Grant I RO1 EYO8459.

References

1. Bergen, J. R. Theories of visual texture perception. In *Vision and Visual Dysfunction, Vol. 10B: Spatial Vision.* (ed. D. Regan) Macmillan, New York (1991).
2. Graham, N. Complex channels, early local nonlinearities, and normalization in perceived texture segregation. In *Computational Models of Visual Processing* (eds. M. S. Landy and J. A. Movshon) MIT Press; Cambridge, MA, USA, pp. 273–290 (1991).
3. Graham, N., Beck, J. and Sutter, A. Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Sign and amount of contrast. *Vision Res.* (1992) (in press).
4. Bergen, J. R. and Landy, M. S. Computational modeling of visual texture segregation. In *Computational Models of Visual Processing* (eds. M. S. Landy and J. A. Movshon) MIT Press, Cambridge, MA, USA, pp. 253–271 (1991).
5. Landy, M. S. and Bergen, J. R. Texture segregation for filtered noise patterns. *Invest. Ophthalmol. Visual Sci.*, **30 (Suppl. 3)**, 160 (1989).
6. Malik, J. and Perona, P. Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am. A*, **7**, 923–932 (1990).
7. Victor, J. D. and Conte, M. Cortical interactions in texture processing: scale and dynamics. *Visual Neurosci.* **2**, 297–313 (1989).
8. Victor, J. D. and Conte, M. M. Spatial organization of nonlinear interactions in form perception. *Vision Res.* **31**, 1457–1488 (1991).
9. Sutter, A., Beck, J. and Graham, N. Contrast and spatial variables in texture segregation: testing a simple spatial-frequency channels model. *Percept. Psychophys.* **46**, 312–332 (1989).
10. Julesz, B. Visual pattern discrimination. *IRE Transactions in Information Theory*, **IT-8**, 84–92 (1962).
11. Campbell, F. W. and Robson, J. G. Application of Fourier analysis to the modulation response of the eye (abstract). *J. Opt. Soc. Am.* **54**, 581 (1964).
12. Campbell, F. W. and Robson, J. G. Application of Fourier analysis to the visibility of gratings. *J. Physiol.* **197**, 551–556 (1968).
13. Chubb, C. and Landy, M. S. Orthogonal distribution analysis: A new approach to the study of texture perception. In *Computational Models of Visual Processing* (eds. M. S. Landy and J. A. Movshon) MIT Press, Cambridge, MA, USA, (1991).
14. Fogel, I. and Sagi, D. Gabor filters as texture discriminators. *Biol. Cybern.*, **61**, 103–113 (1989).
15. Grossberg, S. and Mingolla, E. Neural dynamics of perceptual grouping: textures, boundaries, and emergent features. *Percept. Psychophys.* **38**, 141–171 (1985).
16. Robson, J. G. Neural images: the physiological basis of spatial vision. In *Visual Coding and adaptability* (ed. C. D. Harris) Erlbaum, Hillsdale NJ 177–214 (1980).
17. Sperling, G. Three stages and two systems of visual processing. *Spatial Vision* **4**, 183–207 (1989).
18. Sperling, G. and Chubb, C. Apparent motion derived from spatial texture. *Invest. Ophthalmol. Visual Sci.* **30 (Suppl. 3)** 161 (1989).
19. Sutter, A., Sperling, G. and Chubb, C. Further measurements of the spatial frequency selectivity of second-order texture mechanisms. *Invest. Ophthalmol. Visual Sci.* **32 (Suppl. 4)** 1039 (1991).
20. Hochstein, S. and Spitzer, H. One, few, infinity: linear and nonlinear processing in the visual cortex. In *Models of the Visual Cortex* (eds. D. Rose and V. G. Dobson) Wiley, New York, pp. 341–350 (1985).
21. Graham, N. *Visual Pattern Analyzers.* Oxford University Press, New York (1989).