

Early vision and texture perception

James R. Bergen* & Edward H. Adelson**

* SRI David Sarnoff Research Center, Princeton,
New Jersey 08540, USA

** Media Lab and Department of Brain and Cognitive Science,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA

Texture perception has frequently been studied using textures constructed by repeated placement of micropatterns or texture elements. Theories have been developed to explain the discriminability of such textures in terms of specific features within the micropatterns themselves. For example, Beck^{1,2} observed that a region filled with vertical Ts is readily distinguished from one filled with tilted Ts but not from one filled with vertical Ls. He attributed this to the different distribution of oriented line segments preventing the former case but not in the latter. However, Bergen and Julesz³ found that a region of randomly oriented Xs segregated from one filled with randomly oriented Ls, in spite of the identical distribution of oriented line segments in the two cases. They suggested that this discrimination might be based on the density of such features as terminators, corners, and intersections within the patterns. We note here that simpler, lower-level mechanisms tuned for size may be sufficient to explain this discrimination. We tested this by varying the relative sizes of the Xs and the Ls; when they produce equal responses in size-tuned mechanisms they are hard to discriminate, and when they produce different size-tuned responses they are easy to discriminate.

Figure 1a shows an example of a texture composed of Xs within a texture composed of Ls, similar to the texture used by Bergen and Julesz. The X and L micropatterns are each made of two perpendicular bars, and the bars making the Xs have the same length and thickness as the bars making the Ls. The textures are easily distinguished, and Bergen and Julesz suggested that discrimination could be accomplished by mechanisms that measured the differing densities of micropattern features; for example the Xs have four terminators each, while the Ls only have two; the Ls have a corner while the Xs do not; and so on. This type of description was motivated by analysis of many textures constructed of small micropatterns made up of line

segments, dots or other discrete components.

One major difficulty with this approach is that it is based on a verbal description of image features rather than on the raw intensity values in the image itself. This makes it difficult to test under more general conditions. In order to apply this analysis to a more general class of images, it would first be necessary to construct operators that extract the feature descriptions being invoked—a task that has yet to be accomplished. Before embarking on such a difficult approach it is worth asking whether simpler extracted properties, such as those derivable from linear filters, will suffice (see refs. 4-7).

When inspecting this texture one may observe that the Xs look smaller than the Ls, and that they break up the background differently. This suggests that very simple size-tuned mechanisms, such as cells with center-surround receptive fields could play an important role in the discrimination. We changed the relative sizes of the Xs and the Ls to see whether we could increase and decrease the discriminability of the patterns.

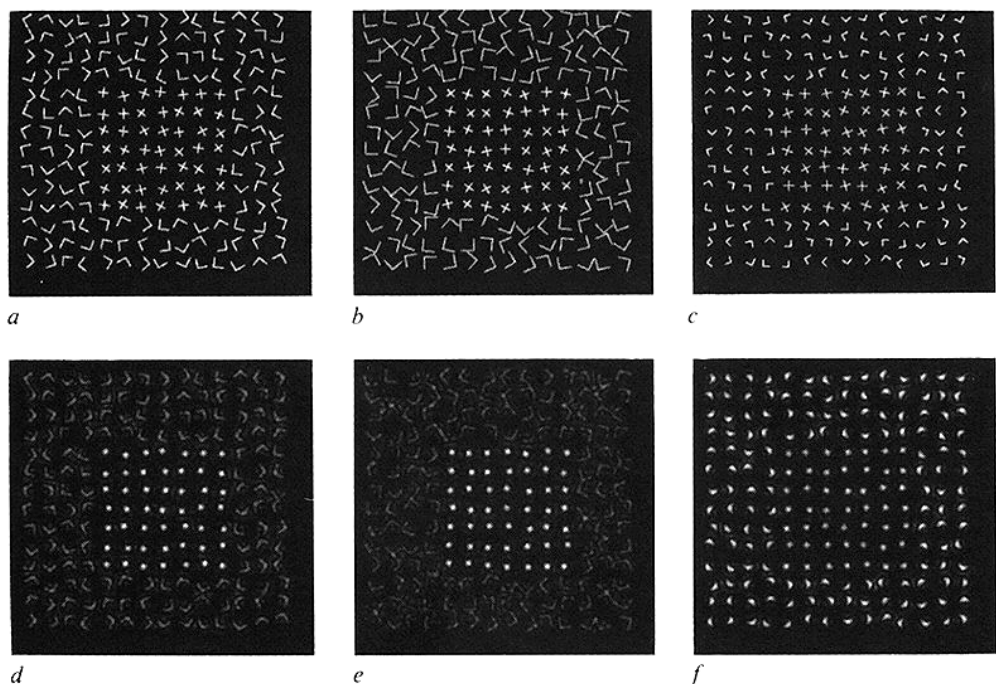
Figure 1b shows the result of lengthening the bars of the Ls by 25%. The bar intensities have been compensated so that the overall density of the micropatterns (that is, the equivalent amount of ink in each) is unchanged. The discrimination becomes easier. Thus, although the micropatterns still have the same number of terminators, corners, and so on, the manipulation of size has a significant impact on the discriminability of the texture.

Figure 1c shows the result of making the bars of the Ls 25% shorter than in the original textures, again with compensation in the intensity of the bars. Now the discrimination is more difficult.

Figure 1d-f shows the response of perhaps the simplest size-tuned mechanism we can construct: a linear centre-surround receptive field followed by full-wave rectification. Figure 1d shows the response to the stimulus of Fig. 1a; the mechanism responds more strongly to the patch in the centre. Figure 1e shows the response to Fig. 1b; now the differences are even more apparent. Figure 1f shows the response to Fig. 1c; in this case the size-tuned mechanism gives responses of similar strength to the two textures.

For this particular set of textures, then, the discriminability can be predicted fairly well from the activities of size-tuned units, without reference to more feature-like properties of the micropatterns. We suggest that the visual system uses a two-stage

Fig. 1 Top row, Textures consisting of Xs within a texture composed of Ls. The micropatterns are placed at random orientations on a randomly perturbed lattice. *a*, The bars of the Xs have the same length as the bars of the Ls. *b*, The bars of the Ls have been lengthened by 25%, and the intensity adjusted for the same mean luminance. Discriminability is enhanced. *c*, The bars of the Ls have been shortened by 25%, and the intensity adjusted for the same mean luminance. Discriminability is impaired. Bottom row: the responses of a size-tuned mechanism *d*, response to image *a*; *e*, response to image *b*; *f*, response to image *c*.



cascade of local energy measures (similar to the cascade of orientation measures discussed by Knutson and Granlund⁵). In the first stage, linear filters are followed by a rectifying nonlinearity (as in fig. 1 *d-f*); spatial averaging provides primary energy measures. These responses are then treated as image arrays for input to a further layer of linear filters, which compute secondary energy measures that indicate the locations of texture boundaries.

Models for texture perception that are based on concepts such as 'terminators' and 'corners' have been important in motivating research in early vision, but the models have proven

difficult to formalize in such a way that they can be applied to wide classes of textures. Although we do not present a full model of texture perception here, the above demonstration indicates that simple filtering processes operating directly on the image intensities can sometimes have surprisingly good explanatory power. The accompanying paper by Voorhees and Poggio⁸, based on a computational investigation into texture analysis, offers an example of a more fully elaborated theory and further demonstrates the potential power of simple processes in early vision.

Received 20 January; accepted 26 February 1988

1. Beck, J. *Percept. Psychophys.* **1**, 300-302 (1966).
2. Beck, J. *Percept. Psychophys.* **2**, 491-495 (1967).
3. Bergen, J. R. & Julesz, B. *Nature* **303**, 696-698 (1983).
4. Caelli, T. *Spatial Vision* **1**, 19-30 (1985).
5. Knutson, H. & Granlund, G. H. (1983) *IEEE Workshop for Computer Architecture for*

- Pattern Analysis and Image Data Base Management*, 12-14, Pasadena (IEEE Computer Society Press, Washington DC, 1983).
6. Voorhees, H. & Poggio, T. *Proceedings, First International Conference on Computer Vision*, 250-258, London (IEEE Computer Society Press, Washington DC, 1987).
7. Bergen, J. R. & Adelson, E. H. *J. Opt. Soc. Am.* **A3**, 99 (1986).
8. Voorhees, H. & Poggio, T. *Nature* **333**, 364-367 (1983).