THREE PROCESSING CHARACTERISTICS OF TEXTURE DISCRIMINATION

Terry M. Caelli

Department of Psychology, The University of Alberta

Edmonton, Alberta T6G 2E9

Abstract

In this paper we examine the idea that texture segmentation comes about by the differential outputs of detectors (non-linear associative filters) computed at each resolvable position on the textured surface. Further, we consider some of the conditions under which "primary" detector outputs are dynamically compared and associated to develop into a smaller set of "texton" profiles which capture the predominant differentiating features of the texture regions. Comparisons to human psychophysical results are made.

KEY WORDS: texture segmentation, associative networks, orientation detectors, adaptability

1. Introduction.

For a biological visual system endowed with a multitude of cells which apparently act as feature extractors or filters, it seems reasonable to presume that visual texture segmentation may come about by the differential responses of such detectors over the textured region. This proposal has received experimental and mathematical attention over the past decade with one-dimensional grey-scaled textures (Richards, 1979; Harvey & Gervais, 1978) and two-dimensional textures (Caelli & Julesz, 1978; Caelli 1982, 1985). However, only until recently has a full computational model been proposed which produces segmentation as a function of such "texton" (Julesz, 1981) outputs, and this paper extends the above analyses in a number of ways (Caelli, 1985).

Here texture segmentation is viewed as having three component processes: (1) spatial decomposition, (2) dynamical associative processing, and (3), classification of textured regions. The specific aims of this model are to enable segmentation when the textures consist of sparse micropatterns; to create networks which will extract, or adapt to, the predominant features of the texture; and to use a classification procedure which is adaptive to the outputs of such detectors.

2. The Model.

2.1 Level I processing: Spatial decomposition and activity profiles.

The initial process of texture segmentation is envisaged to involve the registration of the input (foveal) texture through the parallel outputs of many detectors whose responses are determined by some non-linear transformation of their cross correlation with the input. Assuming a relatively fixed "retinal pre-processor", having opponent center-surround receptive fields, the primary information to be processed must have differential, or band-pass, components emphasized. Further to this, we assume the existence of a relatively fixed primary projection area where such image derivative information is further classified (encoded) by cortical edge and bar detectors whose outputs are a non-linear function of the cross-correlation of the detector's profile with the input image. That is, we assume that the response $R_i(x,y)$ of a detector d_i at retinotopic position (x,y) is determined by:

$$\sum_{\alpha,\beta} d_i(\alpha,\beta) = \text{const.} \qquad (1)$$

and $$R_i(x,y) = \text{const} + \Upsilon\psi\{d_i o I\},\ \Upsilon \equiv \text{constant}. \qquad (2)$$

o denoting cross-correlation between the detector and image (I)

$$d_i oI(x,y) = \sum_{\alpha,\beta} d_i(\alpha,\beta) I(x+\alpha,y+\beta), \qquad (3)$$

and

$$-1 \leq \psi_\delta(z) = \frac{1-e^{-\delta z}}{1+e^{-\delta z}} \leq +1,\ \delta \equiv \text{constant}. \qquad (4)$$

In our simulations we have used

$$R_i(x,y) = 128 + 127\ \psi\delta\{d_i oI\},\ \delta = 0.03, \qquad (5)$$