D. Suter and H. A. Cohen *Modelling of Texture Perception*, Proceedings Int. Conference **Modelling and Simulation**, pp. 430-435, Melbourne, Australia, 14,15,16 Oct. 1987.

Modelling of Texture Perception

D. Suter* and H. A. Cohen
La Trobe University
Bundoora 3083

Introduction

Many advances have been made in machine speech and vision processing through modelling. In speech processing technology we find that many of the advances have been made possible by accurate modelling of the human speech production, the speech sound wave characteristics, and of the receptive mechanisms of the human ear (see for e.g. [1]). Much of the progress made in machine vision has stemmed from modelling of the properties of surfaces, their projections, and of various lighting conditions; as well as modelling of the human visual system itself [2]. Indeed, similar models are often used by those working in both computer graphics and computer vision [3] [4]: these being in some sense analogous to speech production and speech recognition.

The majority of techniques used in vision for texture recognition or discrimination model the texture itself rather than the human visual system. This paper first briefly surveys these approaches before considering how one can relate these models more closely to the emerging model of low level human vision. In particular we describe our approach to unifying texture models, scale space analysis, and computational models of low level human vision.

Textures

It is well appreciated by the computer graphics community that uniformly shaded scenes are unnatural. Real life scenes are highly textured and non-uniformly illuminated [1], and textured scenes give important visual information including depth cues [6]. The graphics artist tries to improve their pictures by including texture; the computer vision researcher tries to improve the performance of a system by developing methods of texture discrimination. The purpose of this section is not to provide an exhaustive summary of these models, but rather to provide a background to our current research.

Texture has proved to be a very difficult feature to model effectively. The initial attempts in this direction were the so called structural models of texture. In these, a texture is modelled as being a collection of primitives (rectangles, blobs or grains) and a rule or statistic describing the spatial distribution of these primitives. An indication of attempts at texture discrimination based on this type of approach can be found in [7]. In this reference the size of blob primitives is estimated by taking a histogram of the difference of average pixel values in inner and outer boxes of different sizes. If the histogram contains modes then texture elements of those corresponding sizes are detected. No attempt was made to extract information on the spatial distribution of these elements.

*Cadet Research Scientist Dept. Defence. DRCS.

[1] An illustrative collection of textures (used by many vision researchers for testing their discrimination procedures) can be found in Brodatz [5]

The other major group of texture models are the statistical models; a global characterisation of texture is attempted through statistical properties of the distribution of gray levels or elementary features such as line terminators or corners. Much research has been devoted to extracting the identity of the statistical models (and their parameters) that explain discrimination of texture by humans [8]. Furthermore, the models proposed by this research have been used to design computer vision algorithms and to devise methods of realistic artificial texture generation [9] [10]. The random models include those based on Markov and Gibbs Random Fields [11] [12] [13] [14] (15].

One perceptual feature of textured surfaces are that they are in some sense rough. Attempts to model surface roughness have included models based on surface curvature statistics [16]. A related method is that of Law's [17] using an energy measure based on the output of detectors designed to be sensitive to the presence of blobs and bars and ridges. This is an attempt to capture empirically the intuitive notion of surface roughness and will be discussed further in section 3.

A theoretically attractive model of surface roughness is the fractal model. A true fractal surface or line is rough even when viewed at any scale; the level of detail and roughness is repeated at each scale. A real life object can only display fractal nature over a range of scales. This range of scales is limited even further by the inherent sampling and smoothing of the human optical system and capturing or display devices. Thus a fractal model is only approximate, and valid only over a range of scales. However, over this range, by the very nature of the model, the fractal nature is preserved and can be represented by a single statistic: the fractal dimension. Rather striking pictures of terrain have been rendered using this model - see Mandelbrot [18]. The traditional method of drawing fractals involves a recursive subdivision of a polygonal boundary (this is the linear analogue of the block faulting method or recursive triangular subdivision used for generation of terrain or texture - see [18] (19] [20]). Fractals can also be generated by fourier synthesis and by simulation of fractional Brownian motion [18] [19]. Fournier, Fussell, and Carpenter [19] use stochastic interpolation to render coastlines, and example given in this reference renders a coastline of Australia from 8 points (digitised from a map). A more novel method is the method of generation of attractor sets of iterated mappings [21] [22] [23].

Many methods of extracting the fractal dimension have been proposed [24], [25]. Pentland [25] shows that under the Lambertian model of surface reflectance the image is also fractal with the same fractal dimension as the surface from which the image arose. Whereas Mandelbrot merely remarks that the images derived from this model appear quite natural, Pentland reports that actual measurements from real images support the model.

2 Scale-Space

The scale space model of visual perception is currently receiving much attention in vision research. In this approach, the image is transformed into a hierarchical representation based on the location of zeros or extrema of a quantity (e.g. image intensity, contour curvature, image intensity gradient) . The idea is to try to capture the most salient features at each scale; allowing matching to be performed from coarser to fine level detail. Analysis of contour curvature will be given as an example.

Analysis of contours in scale space enables segmentation of an image at salient points. Usually, these salient points are taken as corresponding to regions of local maxima/minima in curvature or zeros of curvature. The necessary smoothing to reduce detail in the coarser levels can be gaussian convolution or other low pass filtering operations. An example of a scale space representation of the coastline of Australia is given in figure 1, following much the same procedure as outlined in [26]. The resulting "fingerprint representation" is scale

D. Suter and H. A. Cohen *Modelling of Texture Perception*, Proceedings Int. Conference **Modelling and Simulation**, pp. 430-435, Melbourne, Australia, 14,15,16 Oct. 1987.

and orientation independent (up to a shift of axes). Other scale space representations can be found in [27] [28]. Segmentation of a curve in this manner (by choosing zeros of second derivative of curvature) enables the hierarchical segmentation, where segmentation points at high levels corresponds to coarse scale perceptual significance; proceeding downwards to ever finer detail and less perceptual significance.

We suggest that the scale space representation can be unified with the work on fractal models in vision, by suggesting that the perceptual system extracts and stores a hierarchy of salient segmentation points, together with a measure of the roughness of the region in between these points. An illustration of this can be made using the fractal interpolation methods of [29]; this method allows one to specify a set of fixed points (that could perhaps mark the extreme convexities and concavities of a mountain range skyline at a given scale) and then interpolate with a set fractal roughness between these points. We argue that this is very similar to the way humans would draw a map of Australia, for example, plotting first such features as major capes and bights, and then interpolating with a curve of appropriate roughness between these perceptually salient points.

3 DOOG Filters and Laws Energy Measures

Young [30] recently proposed a method that is biologically plausible for the efficient implementation of oriented filters in the human visual system. These Difference Of Offset Gaussian (DOOG) operators are defined recursively in terms of the gaussian filter *G(x):*

$$D000_0 = G(x)$$

$$DOOG_i = G(x + [\mathrm{Ix}) — G(x — Az)$$

$$DOOG_2 = G(x\ 2\mathrm{dx})\ 2G(x) + G(x\ 2\mathrm{Ax})$$

The offsets *dx* are said to be of one standard deviation *(a);* and the coefficients are binomial. Such an offset scheme is shown to approximate the gaussian and its derivatives. Indeed, the various DOOGS can be related to the combination of gaussian filtering followed by scaled versions of the central difference approximations to differentiation.

Having noted this similarity, we suggest that a natural extension is to construct DOOG filter versions of Laws texture measure filters [17]. This provides a natural way to incorporate scale into these measures. More than this, such a scheme is easily implemented in hardware, both artificial and biological. Many workers have suggested a Laplacian pyramid type architecture for vision systems [31]; here, the image is repeatedly low pass filtered and subsampled in going from one level to the next higher (larger scale and coarser detail). If we low pass filter, then sub-sample at a distance of *a*, followed by Laws filters; then we effectively implement scaled versions corresponding to the DOOG paradigm. Such a scheme reflects the pyramid model of neural visual processing. Here, successive layers of neurons receive weighted signals from a neighbourhood below (a type of combined low pass filtering and sub-sampling), and lateral inhibition and facilitation by horizontal connections within a layer provides the appropriate mask weightings.

Conclusion

There is no truly comprehensive model of vision. There are a number of successful models of limited aspects. Texture recognition has been attempted using a variety of models that are not well integrated with other models of perception. In our approach the scale-space model of perception is extended to incorporate aspects of fractal texture models, similarly, the scale-space mechanism provides a natural framework for providing a sense of appropriate scale to Laws texture measures. This extension naturally encompasses the DOOG model of neural vision processing. Thus, we are working towards an integrated set of models that incorporates models of objects, imaging, and perceptual processes.

References

[1] R. Linggaard, "Electronic Synthesis of Speech," Cambridge University Press, Cambridge, 1985.

[2] D. Man, "Vision," Freeman, San Francisco, 1981.

[3] D. Suter and H. Cohen, "Fractals: representations for rendition and for recognition," *Proceedings Ausgraph '87, May* 1987.

[4] D. Suter, "CAD and Vision: the interplay between rendering, representation, and recognition," - presented at Ausgraph '87 and La Trobe University Dept. Computer Science Technical Report.

[5] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover, New York, 1966.

[6] A. Rosenfeld, "A Note on Automatic Detection of Texture Gradients," *IEEE Trans. on Computers*, Oct. 1975, pp.988-991.

[7] S. W. Zucker, A. Rosenfeld, and L. S. Davis, "Picture Segmentation by Texture Discrimination," *IEEE Tran. on Computers*, Dec. 1975, pp. 1228-1233.

[8] B.Julesz, "Spatial nonlinearities in the instantaneous perception of textures with identical power spectra," *Phil. Trans. R. Soc. Lond. B*, 290, 1980, pp. 83-94.

[9] A. Gagalowicz and S. De Ma, "Sequential Synthesis of Natural Texture," *Computer Vision, Graphics, and Image Processing*, Vol. 30, 1985, pp. 289-315.

[10] N. Ahuja and A. Rosenfeld, "Mosaic Models for Textures," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 3 No. 1, Jan. 1981, pp. 1-11.

[11] R. Chelappa and S. Chatterjee, "Classification of Textures Using Gaussian Random Fields," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 33 No. 4, Aug. 1985, pp. 959-963.

[12] T. S. Yu and K. S. Fu, "Recursive Contextual Classification Using a Spatial Stochastic Model," *Pattern Recognition*, Vol. 16 No. 1, 1983, pp. 89-108.

[13] H. Derin and H. Elliot, "Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 9 No. 1, Jan. 1987, pp. 39-55.

[14] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6 No. 6, Nov. 1984, pp. 721-741.

[15] R. L. Kashyap and P. M. Lapsa, "Synthesis and Estimation of Random Fields Using Long-Correlation Models," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6 No. 6, Nov. 1984, pp. 800-809.

D. Suter and H. A. Cohen *Modelling of Texture Perception*, Proceedings Int. Conference **Modelling and Simulation**, pp. 430-435, Melbourne, Australia, 14,15,16 Oct. 1987

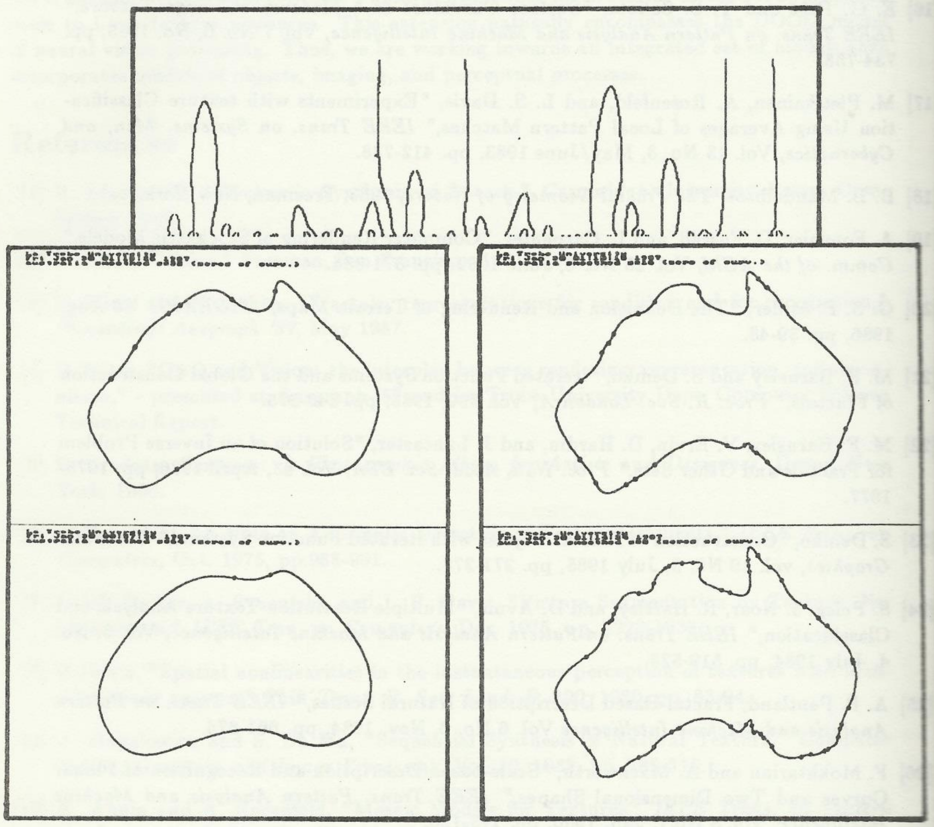

Fig. 1. Scale-Space Representation of Australia
(Top) Zeros of curvature, scale vetical and arc length horizontal.
(Bottom) Four stages of smoothing showing zeros of curvature.