12TH INTERNATIONAL JOINT CONFERENCE on ARTIFICIAL INTELLIGENCE
Darting Harbour, Sydney, Australia 24 - 30 August 1991

IJC I-91

Australian National Committee
Secretariat
Ms. Beverley Parrott
Parrish Conference Organizers
120 BourkeStreet
Woolloomooloo NSW 2011
AUSTRALIA
(+61-2) 357-2600
(+61-2) 357-2950

All mail to:
PO Box 787
Potts Point NSW 2011
AUSTRALIA

Proceedings

Workshop on Computer Vision - From Cognitive Science to Industrial Automation

Edited by R.A. Jarvis

Multi-Scale Versus Scale Invariance

Harvey A. Cohen
Computer Science and Computer Engineering
La Trobe University
Melbourne Victoria 3083

Workshop: ***Computer Vision, from Industrial Automation to Cognitive Science***

Multi-Scale Versus Scale Invariance

Harvey A. Cohen

Computer Science and Computer Engineering

La Trobe University

Melbourne Victoria Australia 3083

Multi-Scale Versus Scale Invariance

Harvey A. Cohen
Computer Science and Computer Engineering
La Trobe University
Melbourne Victoria Australia 3083
H.Cohen@latrobe.edu.au

Abstract:

It is argued that the multi-scale approach of Marr and co-workers is more neurologically plausible than the mathematically rich but neurologically sterile concept of scale invariance. Key to this view-point is the use of multiple spatial channels - with features detected in one channel confirmed in another. For Marr's edge detector the neurological work of Wilson supplied strong support for the actual existence of such channels.

The writer (with I You) has applied the multi-scale approach to the problem of the definition of texture, to successfully segment images containing textures of unknown scale. This multi-scale texture recognition and segmentation works is outlined.

The work with texture, determines features common to multiple scales, which is consistent with the application of fractals and contemporary dynamics in a specifically limited way to image analysis. The scope of image coding via fractals is discussed both as a practical tool for image compression, and a conceivable mechanism for neurological coding is sketched.

1 To Start with Marr's Edge Detector

The pioneering work of Marr and co-workers was most striking in the development of a multi-scale approach to detection. What Marr sought was a biologically plausible edge detector for early vision. One part of Marr's model is the use of the mechanism of the DOGS filter - which could naturally model either a single retina or a group, with the line being determined by zero crossings. But the output from such a process includes many poor matches to lines. Marr - in his brilliance - hit on the idea of using confirmation of the output of multiple channels as the means of determining the most plausible edge points.

This multi-scale approach attracted much attention, with much work _ being devoted to scale invariance in many forms. Yet out of all this work, mathematically fascinating as it has been, there has not been a single major new algorithm for computer vision.

The naive neurophysiologist might well have predicted this evident failure - the jump from the discrete channels of Marr (and of the neurophysiologist Wilson) to a continuum of values lead to wondrous scope for analysis - but far removed from neural plausibility.

2. Texture

Early work on texture involved the concept of texture as being quasi-repetitive. This lead to attempts to characterise texture via the fourier transform. Although applicable to artificial highly ordered textures, this approach was of little merit in the analysis of natural textures. However, the basic scheme was extended through the adjoining of probability, with moderately successful work with co-occurrence matrices.

A simple and more biologically plausible approach was introduced by Laws, whose work can with hindsight be characterised as being concerned with the use of a local averages and differences at multiple scales. What Laws had actually done was to define a quantity called, the texture energy, to be associated with individual pixels (in image analysis). Formally the texture energy of the pixel at (i,j) with gray-scale E(i,j) is best (re)-defined as

$$TE(i,j) \quad \sum_{rs} (A * E[i+r,j-Fs])^2 \{r,s \in W\}$$

where W is a square window centred on i,j while A is a zero sum convolution mask.

Laws used 3*3 and 5*5 convolution masks, and windows of size 15*15 and 17*17. The actual zero sum masks used by Laws were of the form of the direct product of 1D difference operators with 1D binomial filters.

Laws found different mask could better distinguish certain texture pairs, and proposed the use of classifier theory to more optimally determine pixel classification. The problem was, as Cohen and You determined (1987) Laws Texture energy based on Laws 'best' mask has a ratio of standard deviation (SDV) to average value (VAL). often of the order of 20% leading to difficulties in segmentation.

In a series of papers Benke and Skinner introduced the notion of a 'tuned' mask - a single mask for a group of textures whose parameters have been selected to optimise some figure of merit. This work was extended and improved by You and Cohen, by a scheme whereby group of textures as large as 15 might be recognised.

3- Multi-scale texture recognition

How might a multi-scale approach be applied to texture discrimination and image segmentation (by texture). be developed? The basic algorithm for generating a tuned mask involves training over large samples of pure texture. Over the training period stochastic means are used to change the mask parameters so that a figure of merit is optimised. You and Cohen have developed performance indexes that maximise the dispersion between different textures, while minimising the SDV of texture energy for a given texture. A particular data structure of linked lists is used to dynamically alter the 'ranking' of textures to achieve maximum dispersion.

To extend this scheme to multi-scale one must somehow 'confirm' the parameters at different scales. This we have achieved by augmenting the training system so that for each texture considered there samples available at different scales. In the initial work we have been concerned with scales of $(1:1)^2$ $(2:2)^2$ and $(3:3)^2$. We have been able to tune mask so that the dispersion between 'different' textures is maximised, while the same texture energy to a few percent is given by the 'same' texture at different scales. Textures from the set of 15 Brodatz textures in Fig(i) have been effectively ranked for similarity in the determination of single mask for all 5 at different scales. As part of an evolving hierarchical approach to image segmentation, further mask have been developed that discriminate in a multi-scale manner between rank adjacent textures - the textures that otherwise are not well distinguished. Applying such mask to collages of rank adjacent Brodatz textures yields the effective segmentation indicated in Fig(2). Note that in these segmentation examples, simple thresholding of the texture energy ONLY has been applied.

4 . Fractals

Barnsley and Sloan have proposed the use of Iterated Function Systems as a means of encoding images. An encoder based on this scheme has recently become commercially available.
There is a class of models of conceivable neural processes - expressed by modern dynamical systems:
Formally suppose there are a number of nodes with mappings W[i] between each with probability p[i],

Then a system

[i] --> [j] is a dynamical system

Such a random walk is a plausible process for neural networks - but overall very different from what is currently construed to be connectionist.

Barnsley and co-workers have shown that arbitrary images can be in principle be encoded via sets of such recurrent processes, and have shown that remarkable image compression can be achieved.
In this context it is salient to point out that the predominant feature of fractals is their essential invariance under various scales.

Conclusions

In the multi-scale approach the key feature is the use of multiple spatial channels - with features detected in one channel confirmed in another. Biologically plausible, it has been somewhat ignored since Marr. However the writer has been able to apply this powerful notion to effect a segmentation of collages of Brodatz textures at unknown scale. Likewise, using the multi-scale philosophy, Barnsley and co-workers have been able to demonstrate a new, potent, object oriented approach to image compression_

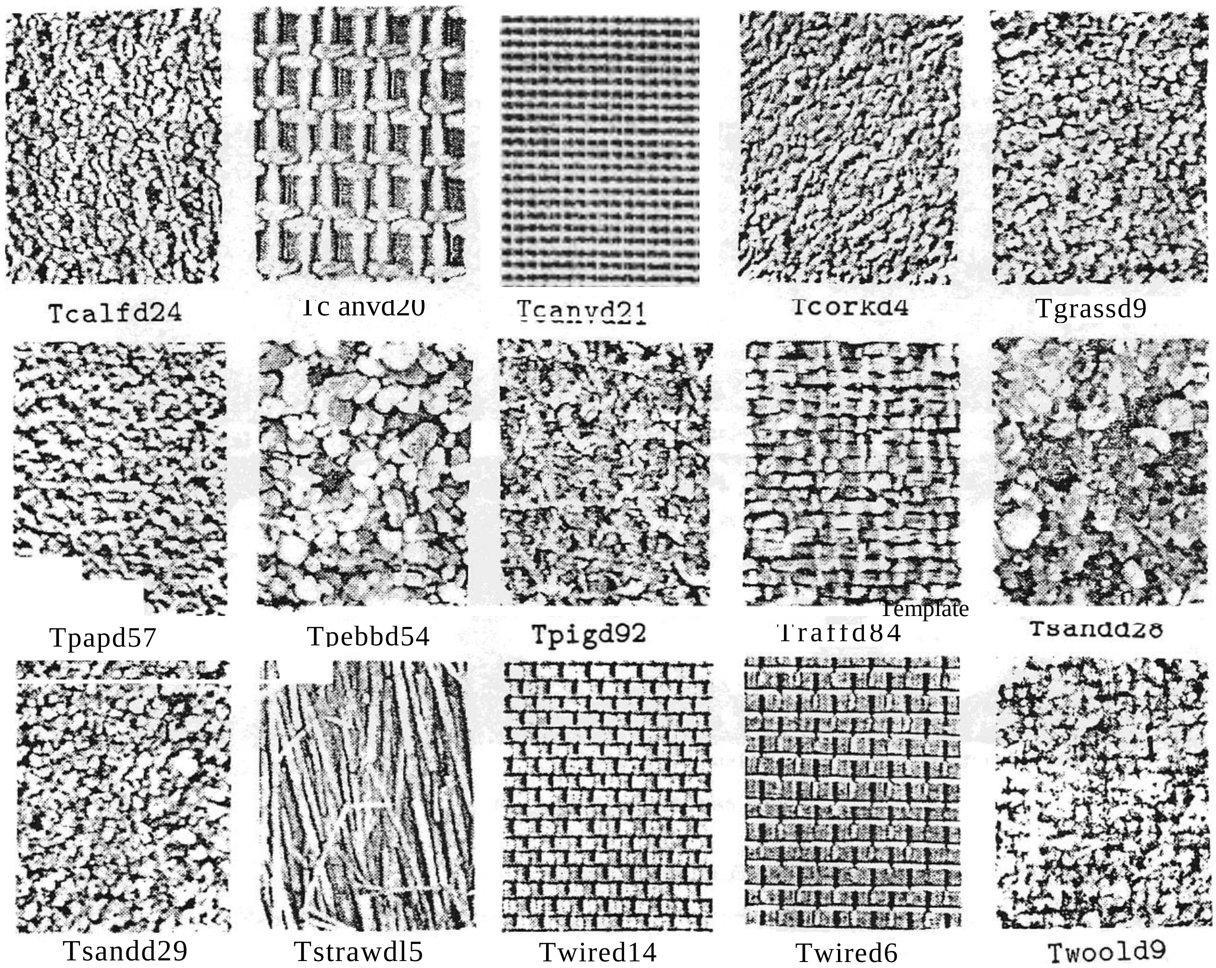

Figure 8. 15 Texture Samples

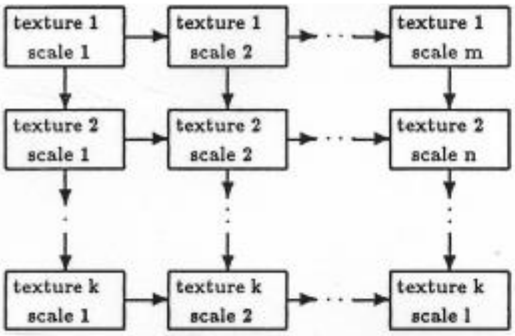

Figure 1 Two dimension link list

Template

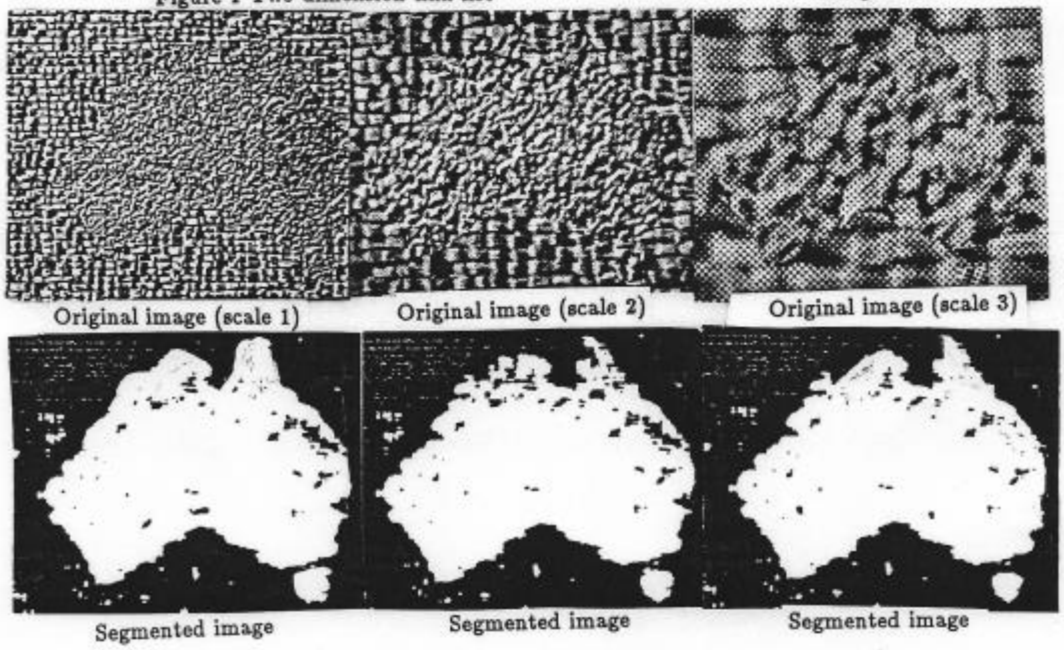

Figure 2 Segmentation results

Table 1 Texture Energy Dispersion

M_{seg}

$$\begin{matrix} -11 & 60 & -6 & -21 & -22 \\ -41 & 58 & -57 & -26 & 66 \\ -27 & -27 & 62 & 35 & -43 \\ 87 & -68 & -81 & 44 & 18 \\ -77 & 5 & 66 & -45 & 51 \end{matrix}$$

	Mask M_{seg}		Mask $R5R5$	
Texture	TE	SDV/TE	TE	SDV/TE
s1raffiad84	1331	18%	131	29%
s1corkd4	2677	18%	486	51%
s2raffiad84	447	17%	5	30%
s2corkd4	1154	16%	320	37%
s3raffiad84	202	21%	6	65%
s3corkd4	569	20%	28	38%

$R5R5$

$$\begin{matrix} 1 & -4 & 6 & -4 & 1 \\ -4 & 16 & -24 & 16 & -4 \\ 6 & -24 & 36 & -24 & 6 \\ -4 & 16 & -24 & 16 & -4 \\ 1 & -4 & 6 & -4 & 1 \end{matrix}$$

References

[1] Benke, K.K., Skinner, D.R., and Woodruff, C.J., "Convolution operators as a basis for objective correlates for texture perception", *IEEE Trans. Systems, Man, and Cybernetics* Vol. SMC-18, 1988, **pp**. 158-163.

[2] Cohen, H.A., and You, J., "Texture statistic selective masks", *Proc. 6th Scandinavian Conference on Image Analysis,* Oulu, Finland, 1989, pp. 930-935.

[3] Laws, K. **I.,** "Textured Image Segmentation", *Ph.D Thesis,* University of Southern California, January 1980.

[4] You, J. and Cohen, H.A., "Image Segmentation via Convolution Masks", *Proc. of the Third National Conference on Robotics,* Melbourne, Australia, 1990, **pp.** 305-316.