IEEE Region 10 Conference, Tencon 92
11th - 13th November, 1992
Melbourne, Australia

Applying Neural Networks to Colour Image Data Compression

Keith R. L. Godfrey and Yiannis Attikiouzel

Centre for Intelligent Information Processing Systems
Department of Electrical and Electronic Engineering
The University of Western Australia
Nedlands 6009 Australia

This work has been supported in part by OTC Limited and the Australian Telecommunications and Electronics Research Board.

Abstract

With the advent of complex medical imaging and high speed video teleconferencing, there is an emerging need for fast colour image compression algorithms. This paper employs a self-organising neural network to achieve colour image segmentation and image data compression, with an adaptive codebook for faster training. Neural network architectures are well suited to high speed processing because they are massively parallel.

1. Introduction

Neural networks [1] are becoming increasingly popular in Engineering because they can adaptively solve real world problems without the need for algorithmic programming [2-6]. Neural networks have been applied to handwritten character recognition [3], oil field reserve estimation [4], robotic control [5], pronunciation of written text [6] and many other tasks. Neural network architectures and training algorithms have evolved to suit each type of application.

Neural network paradigms fall into three broad categories: supervised, reinforced and self-organising. The categories identify the amount of corrective data required by a network during its training phase. A supervised network relies on the error between the actual and desired outputs and therefore requires *a priori* knowledge of its desired output [7]. A reinforced network requires a measure of the overall error but does not need its exact output [8]. A self-organising network determines its own internal representations of the input data and does not rely on external guidance to a specific output format [9].

Colour image segmentation involves an analysis of the pixels in an image to determine a codebook for pixel quantisation. By definition, the pixel colours in the codebook are not previously known. Thus if a neural network is to achieve this goal, it must incorporate a self-organising paradigm.

Self-organising networks have been applied to vector quantisation of monochrome images [10] by exploiting spatial correlation. This paper shows how self-organising networks can also be applied to quantisation of individual pixels, which is especially important in colour images.

Colour image data compression arises from colour segmentation if a unique code is assigned to each codebook colour, using a minimal number of data bits. There is typically a trade-off between the extent of compression and the speed of the algorithm. Although a high degree of compression is desirable, the delay imposed by the compression algorithm must be small for real time applications. Neural network architectures are well suited to the task because they are inherently parallel.

2. Segmentation

Colour images are segmented by classifying each pixel into one of a set of fixed colours. A self-organising neural network achieves this by matching each pixel's colour to the closest colour in a self-organised array. Prior to matching each pixel, the network must first be trained so that it builds a codebook of the preferred colours in the image.

A self-organising architecture for colour image segmentation is developed as follows. The network uses the type of processing element introduced by Teuvo Kohonen [9,11] and comprises two slabs: an input slab and a self-organising slab. If the input is the colour of a single pixel then, after training with random presentations of colours from an image, each element in the self-organising slab will correspond to a single time-averaged colour. The set of colours thus formed in the self-organising slab becomes the codebook for pixel classification. The size of the codebook is determined by the size of the self-organising slab, so an image can be segmented to any desired extent by *a priori* specification of the dimensions of the neural network. The segmentation network is shown in Figure 1.

Colour images commonly use 24 bits per pixel, giving a total of 16,777,216 colours mixed from 256 shades of reds, greens and blues. Each pixel therefore contains three 8-bit numbers. Prior to input to the segmentation network, the pixel colour is transformed to be independent of luminance so that the network forms segments of uniform chromaticity. The transformation is:

incorporates a red-green difference and a yellow-blue difference, so the chromaticity is represented in two independent dimensions.

$$X = (R-G) / (R+G+B) \qquad \text{\{the red-green differential\}}$$

$$Y = (R+G-B) / (R+G+B) \qquad \text{\{the yellow-blue differential\}}$$

where R, G and B are the red, green and blue components of the pixel. The input to the network is therefore two-dimensional.

During classification, the input pixel chromaticity is compared to each chromaticity in the codebook by means of the Euclidean distance. The distance between the input chromaticity and the i-th codebook vector is:

$$distance_i^2 = (X-X_i)^2 + (Y-Y_i)^2$$

where X_i and Y_i are the chromaticity coordinates of the i-th codebook vector. The winning codebook vector is the one with the smallest distance measure.

3. Data Compression

Without compression, digital images consume vast quantities of data space. A medium resolution image of 512 by 512 pixels for example consumes three quarters of a megabyte.

The information content of an image is typically much smaller than its physical size, for two reasons. First, most images contain a high degree of spatial correlation and, the larger the image, the larger the correlation will be. The second reason is that the number of colours present in an image can normally be represented in much fewer than 24 bits. The remainder of this analysis concentrates on the latter point.

If an image has size 512 by 512 then it is physically impossible for that image to contain 16,777,216 distinct colours because there are only 262,144 pixels in the image. Furthermore, it is unlikely that there will be even that many colours. The individual pixels can therefore be represented in fewer than 24 bits with no loss of information. Information Theory is used to predict the minimum number of bits required. If C represents the set of colours present in an image, then the information content of each colour C_i is given by $log_2(1 / P(C_i))$, and the average entropy of each pixel is:

$$entropy = \sum_i P(C_i) * log_2(1 / P(C_i))$$

where $P(C_i)$ is the probability of meeting each colour in the image.

The entropy represents the minimum number of bits for lossless representation of each pixel or, alternatively, the maximum number of bits if loss is permitted.

Data compression is achieved by coding, in binary, the colours from a segmentation codebook. The segmentation neural network is extended to perform compression by adding an extra slab for binary output. Each element in the self-organising slab is connected to the binary output slab through a unique binary combination. Strong lateral inhibition between elements in the self-organising layer ensures that only one self-organising element responds to each input colour, so that there will be a unique binary output for each codebook colour.

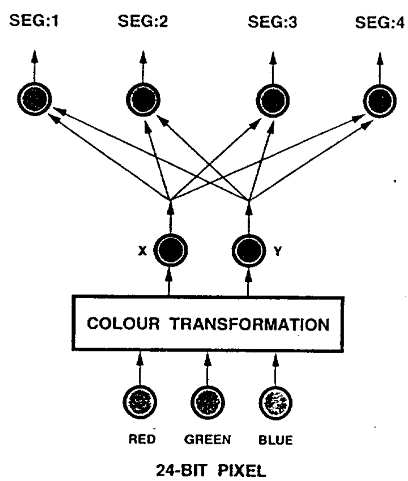

Figure 1:
The segmentation neural network.

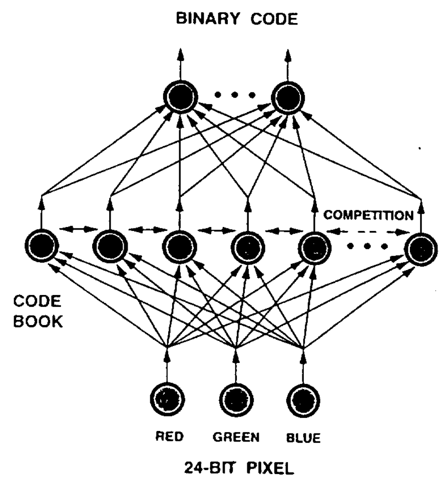

Figure 2:
The compression neural network.

The compression network is shown in Figure 2. An important difference between the compression network and the segmentation network is that the pixel intensity is required for compression. Thus the inputs to the compression network are the red, green and blue colour components, instead of the X and Y chromaticities.

4. An Adaptive Codebook

The standard self-organising architecture is simple and effective but its codebook must be determined prior to classification. This typically requires a training run through random selections of pixels from the image, until the codebook converges to an average representation of the whole image. The training phase represents a considerable overhead in the segmentation or compression process. For example, a video frame is normally transmitted from top to bottom and left to right, so that the pixels form a continuous sequence. If a codebook has to be determined from this sequence prior to its transmission, then the whole sequence must be received and stored before random training can commence, and the training must finish before the coded sequence can be generated.

The segmentation and coding process is much faster if the training is performed adaptively. This means that the neural network should learn from the pixels while they are being coded, so that the codebook adapts to the emerging pixel sequence. An adaptive codebook may not be as optimal as a codebook derived from a preview of the entire image but it greatly reduces the delay to the transmitted pixel sequence.

The self-organising neural network will require a reinforced input to decide whether a new pixel is sufficiently close to a codebook colour or whether the new pixel should be added to the codebook. This is implemented by a threshold distance measure. The encoder and decoder must follow a strict communication protocol to ensure that they stay aligned with the threshold decisions and hence the evolving codebook.

5. Example

Figure 3 shows an image of yachts moored on the Swan River in Perth, Western Australia. The image uses 24-bit colour but has an entropy of only 10.325 bits per pixel.

The yachts image was passed through a compression neural network with 1024 codebook vectors. After selecting a suitable distance threshold for adaptive coding, only 974 of the codebook vectors were filled. The resulting image, shown in Figure 4, was encoded with almost no visual degradation at 6.39 bits per pixel. The reason for the non-integral number of bits per pixel is that the codebook has to be transmitted with the image.

Both images are normally displayed on an IRIS-4D workstation in full 24-bit colour. As colour cannot be reproduced in this publication, only the green components of the images are shown. The images have been half-toned through a 8x8 dithering grid.

6. Conclusion

Self-organising neural networks have been shown to segment and compress 24-bit colour images. By adding an external threshold decision, a compression network can build its codebook adaptively and therefore speed the compression process. A 24-bit colour image, reproduced here in monochrome, was compressed to 6.39 bits with virtually no visual degradation. As the analysis in this paper has concentrated on the individual pixels, the spatial correlation between pixels could still be exploited.

6. References

[1] R. P. Lippmann, "An Introduction to Computing with Neural Nets", IEEE ASSP Magazine, pp 4-22, April 1987.

[2] Teuvo Kohonen, "An Introduction to Neural Computing", Neural Networks, Vol. 1, No. 1, pp 3-16, 1988.

[3] Kunihiko Fukushima, Sei Miyake and Takayuki Ito, "Neocognitron: a neural network model for a mechanism of visual pattern recognition", IEEE Trans. System, Man and Cybernetics, vol. SMC-13, pp 826-834, 1983.

[4] "Introductory Neurocomputer Applications Workshop", Hecht-Nielsen Neurocomputer Inc., 5501 Oberlin Drive, San Diego, California, 1989

[5] Teruo Fujii and Tamaki Ura, "SONCS: Self-Organizing Neural-Net-Controller System for Autonomous Underwater Robots", Proc. Int. Joint Conf. on Neural Networks (IJCNN'91), pp 1973-1982, Singapore 1991.

[6] Terrence J. Sejnowski and Charles R. Rosenberg, "NETtalk: a parallel network that learns to read aloud", Technical Report JHU/EECS-86/01, Electrical Engineering and Computer Science, The Johns Hopkins University, 32 pp, 1986.

[7] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning Internal Representations by Error Propagation", Parallel Distributed Processing: Explorations in the Microstructures of Cognition, Vol. I, pp 318-362, MIT Press, Cambridge MA, 1986.

[8] A. G. Barto, "Learning by Statistical Cooperation of Self-Interested Neuron-Like Computing Elements", Human Neurobiology, Vol. 4, pp 229-256, 1985.

[9] Teuvo Kohonen, "Self-organized formation of topologically correct feature maps", Biological Cybernetics, vol. 43, pp 59-69, 1982.

[10] Nasser M. Nasrabadi and Yushu Feng, "Vector Quantization of Images Based Upon the Kohonen Self-Organising Feature Maps", Proc. Int. Conf. on Neural Networks (ICNN'88), pp I-101 - I-108, San Diego, 1988.

[11] Teuvo Kohonen, "Self-Organization and Associative Memory", Springer-Verlag, 1984.

original:

24 bits per pixel

entropy:

10.325 bits p. p.

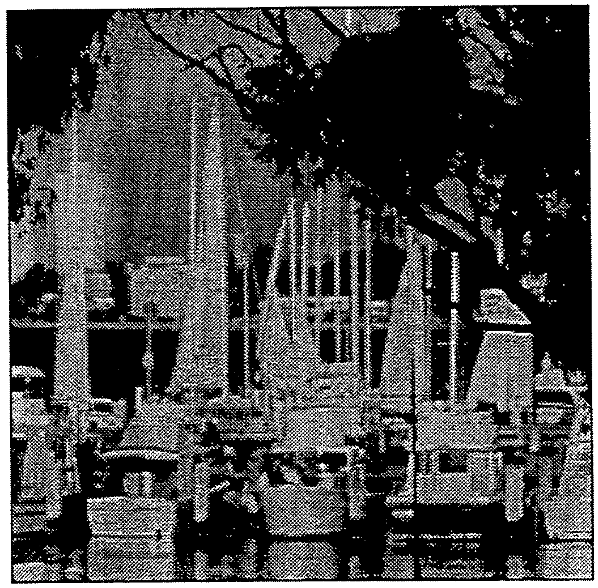

Figure 3: Original yachts image (green component).

encoded as:

6.39 bits per pixel

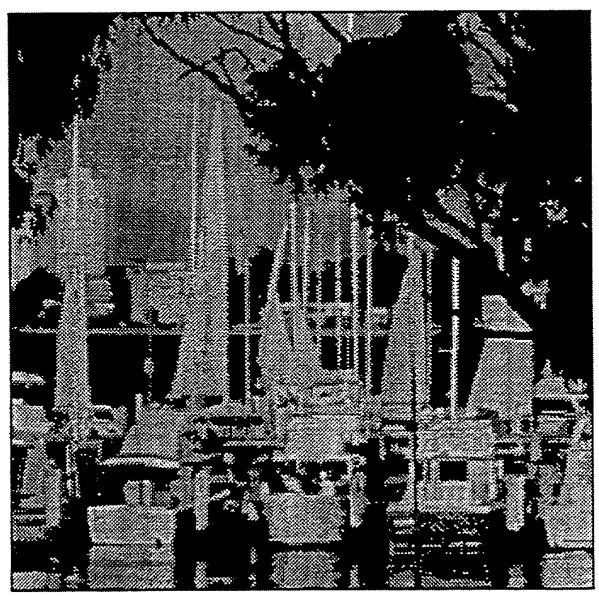

Figure 4: Compressed yachts image (green component).