COMPARISON OF SOME METHODS FOR PROCESSING 'GREY LEVEL' DATA IN WEIGHTLESS NETWORKS

R.J.MITCHELL, J.M.BISHOP, S.K.BOX AND J.F.HAWKER
Department of Cybernetics, The University of Reading, PO Box 219, Whiteknights, Reading, RG6 6AY, UK

Weightless neural networks have been used in pattern recognition vision systems for many years. The operation of these networks requires that binary values be produced from the input data, and the simplest method of achieving this is to generate a logic '1' if a given sample from the input data exceeds some threshold value, and a logic '0' otherwise. If, however, the lighting of the scene being observed changes, then the input data 'appears' very different. Various methods have been proposed to overcome this problem, but so far there have been no detailed comparisons of these methods indicating their relative performance and practicalities. In this chapter the results are given of some initial tests of the different methods using real world data.

1 Introduction

The use of weightless networks for pattern recognition is well known and the first commercial hardware neural network vision system, WISARD[1], used weightless networks. These networks require binary values for forming the 'tuples' used to address their RAM neurons. In a simple system, the input to the network comprises a set of boolean values, and the tuples are formed by sampling these values. If however the input comprises 'grey level' data, the values used to form the tuples must be obtained by processing the input data in some manner.

The simplest method, 'thresholding', is to generate a logic '1' if the input value exceeds a specified threshold value, and a logic '0' if otherwise. One drawback of thresholding is that the tuples so formed can be very different when the d.c. level of the input data changes, which might occur in a vision system due to changes in the ambient light of a scene as viewed by a video camera.

This problem can be partly reduced by automatic thresholding, that is, by determining the threshold value of each input data. This can be achieved by automatic controls on a video camera, but if that option is not available suitable processing of the input data is required. Ideally the threshold value should be the median of the input data so that half of the data are logic '1'. However, calculating the median is relatively difficult, often involving sorting the data[2]. Therefore, as a compromise, the mean of the input data is often used, but this reduces the performance of the system.

One robust method of processing grey level data is to use 'thermometer coding'[3]. This involves converting each sampled point of the input into an array of Boolean values, where the greater the amplitude of the source, the more Boolean values are true. In a 4-bit thermometer code, the input data are quantised into five values, represented by 0000, 0001, 0011, 0111 and 1111; similarly, a 16-bit thermometer code quantises the data into seventeen values. This coding is equivalent to maintaining many threshold systems, each with a different threshold value. A major drawback of the method is that, for a t-bit thermometer code, t times as much memory is required for the RAM neurons than for thresholding.

Therefore, a different technique has been developed at the University of Reading, called Minchinton cells[4]. These cells are general purpose simple pre-processing elements which are placed between the input data and the RAM neurons and which can make the system more tolerant of changes in the d.c. level of the data without the need for any extra memory and for very little extra processing. The method has been used successfully, for example, in a hybrid weightless system which attempted to find the position of eyes within images of faces[5], and as part of an auto-associative weightless neural network[6].

Three types of Minchinton cells have been defined[4], all of which process values from the non-Boolean input data and which produce a Boolean result. In the following, let **I** be the input data and **I**[*x*] be the value at position *x* within these data: typically *x* is chosen randomly. The cell types are:

	Definition	Function
i)	$\mathbf{I}[x] > \text{constant}$	Threshold
ii)	$\mathbf{I}[x1] > \mathbf{I}[x2]$	Type 0 cell
iii)	$(\mathbf{I}[x1] - \mathbf{I}[x1 + 1]) > (\mathbf{I}[x2] - \mathbf{I}[x2 + 1])$	Type 1 cell

The first of the above is thresholding, which is thus one form of the general Minchinton cell. As thermometer coding can be achieved with multiple thresholds, such coding can also be implemented using Minchinton cells.

The Type 0 cell returns a Boolean true if the value at position *x1* in the input data exceeds that at position *x2*: this type gives better tolerance to changes in overall lighting for the following reason. Suppose the lighting of the input data increases. It is likely therefore, if saturation effects are ignored, that the values at the two positions will increase by about the same amount, hence the difference between the values will hardly change and so the output of the cell is likely to be unchanged. Any changes in the cell outputs that do occur can, of course, be compensated for by the generalisation properties of a standard weightless recognition system.

The Type 1 cell simply compares two points in the input data by calculating the difference between adjacent points, each pair being chosen randomly. The effect of the Type 1 cell is to render the network insensitive to zonal d.c. changes in the input, such as shadows, except at zonal boundaries. However, as the Type 1 cell involves a second differential process, the network becomes more sensitive to high frequency changes in the image. This form of cell has not yet been used in practice.

The three types of Minchinton cell can, like weightless networks, be easily implemented in hardware, for example, by using the modular hardware system developed at Reading[7].

Another potential advantage to the Minchinton cell is that it uses more information than the other methods. If simple thresholding is used and, for example, the input data are 8-bit values and the threshold is 100, the system cannot distinguish between an input value of 120 and one of 160. Thermometer coding effectively provides more threshold values, so this problem is reduced, but only slightly. With the Type 0 Minchinton cell, each input data value is compared with another data value, not a constant, so overall the system may be able to detect higher frequency information than the other methods. This can become more significant when the input data are oversampled, that is, each value from the input space is sampled more than once; then each sample is likely to be compared with many different values.

Although these Minchinton cells have been used successfully, little research has as yet been done to compare their performance with that of thresholding or thermometer coding. The purpose of this chapter, therefore, is to report on such comparisons. These comparisons are achieved by testing images from a video camera, looking at the success of the network is terms of pattern recognition and discrimination. The tests were done on weightless networks which employed thresholding, thermometer coding and the Type 0 and Type 1 Minchinton cells.

2 The Tests

2.1 Test Data

The aim of the research is to investigate the response of weightless networks in recognising and discriminating different objects which are subject to changes in how they are illuminated over time. Although software simulations could be used, it was felt that using real world data would add verisimilitude to the work. As vision systems are often used on production lines, it was decided to choose some suitable products and record their appearance under different lighting conditions using a standard video camera. The products chosen were four disk boxes as they were convenient and of appropriate shape and size.

These boxes were placed under a video camera such that all four were visible at one time and then the image from the camera was captured every 10 minutes over a 24 hour period and stored for later processing. Each complete image comprised 512×512 8-bit values, so the image of each disk box was of size 256×256 bytes. Then, when each image was tested, the data for the four disk boxes were separated and processed by VISIWIN[8] a Windows™ based product in which various configurations of weightless networks can be simulated.

Figure 1 shows four of the images of data used in the tests. Each image comprises the four diskette boxes, and each image was taken at a different time of the day. As can be seen from these images, the system is being asked to recognise images subject to different amounts of illumination and to shadows. In the tests four discriminators were used: the Fuji MF2HD box was taught in the first discriminator, the 3M box to the second, the Fuji MF2DD to the third, and the Dysan box to the fourth.

Figure 1: Four images, each of four diskette boxes, taken during the 24 hour period

2.2 Weightless networks tested

Nine network configurations were tested, these were:

Thresholding, where the threshold was the mean of each data input

4-bit Thermometer coding
16-bit Thermometer coding
Type 0 Minchinton cell with no oversampling
Type 0 Minchinton cell with 4 times oversampling
Type 0 Minchinton cell with 16 times oversampling
Type 1 Minchinton cell with no oversampling
Type 1 Minchinton cell with 4 times oversampling
Type 1 Minchinton cell with 16 times oversampling.

For each network type, the software was configured with four discriminators and images from each of the four disk boxes were trained in the associated discriminator.

The full range of results will be reported elsewhere. Here however a summary of the important results will be given. For these tests a suitable tuple size was chosen, namely 8, each network was trained on a few images, and then the response of each discriminator was recorded. This process was performed twice: in test A the training data were a few images at the start of the sequence; in test B the training data were images taken once each hour during the day.

3 Results

A multi discriminator system seeks to recognise and discriminate data. The best measure for this purpose is relative confidence[9], defined as follows, where the response of a given discriminator is the number of RAM neurons which report a '1' when the data are presented:

$$\frac{\text{Response of best discriminator - Response of next best discriminator}}{\text{Response of best discriminator}}$$

Note, where a network misrecognised the input image, the relative confidence measure was set to 0.

In the tests done, the three Type 0 Minchinton Cell networks responded almost identically, as did the three Type 1 Minchinton Cell networks, and the 4-bit thermometer code network was significantly better than the 16-bit thermometer code network. The graphs below thus show the relative confidence of the four discriminators for the network using thresholding, the 4-bit thermometer code network, and the Type 0 and Type 1 Minchinton cells with no oversampling.

Figure 2 shows graphs of Relative Confidence for the four discriminators when the networks were trained on data at the start of the sequence, that is test A; and figure 3 shows the responses of the discriminators when the network was trained on data once every hour, that is test B.

4 Discussion

The graphs clearly show that the Type 0 and Type 1 Minchinton cells are better than thermometer coding which itself is better than thresholding. Usually the Type 0 cell is better than the Type 1 cell, however, in some of the graphs in test B, the Type 1 cell performs best. In test B, the network is trained on more representative data: thus it seems that the Type 0 cell is better at generalisation.

As mentioned earlier, the 4-bit thermometer code has a better relative confidence than the 16-bit code. This is because the response of all discriminators using the 16-bit thermometer code were consistently higher than those when the 4-bit code was used and hence the relative difference between the 'best' and the 'next best' discriminator is smaller.

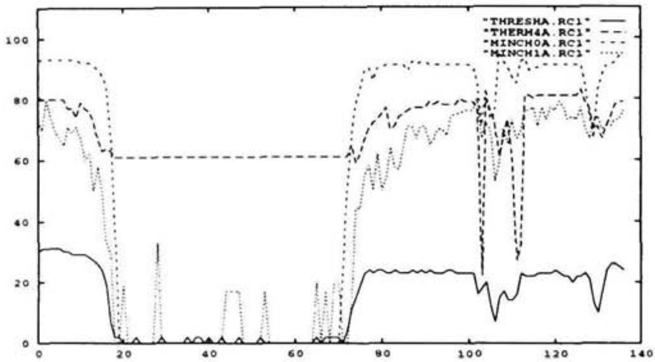

Figure 2a: test data were the 'Fuji MF2HD'

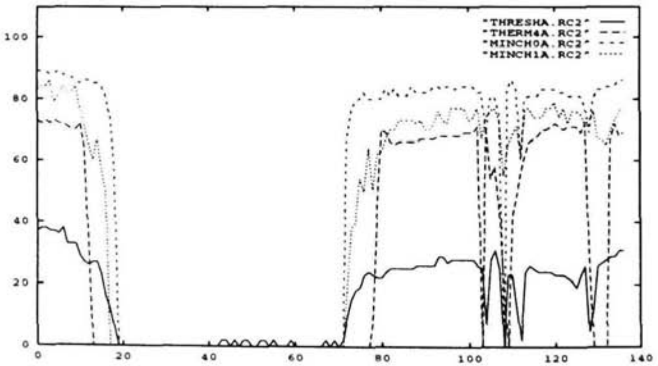

Figure 2b: test data were the '3M'

Figure 2: Relative Confidence for four discriminators for Test A

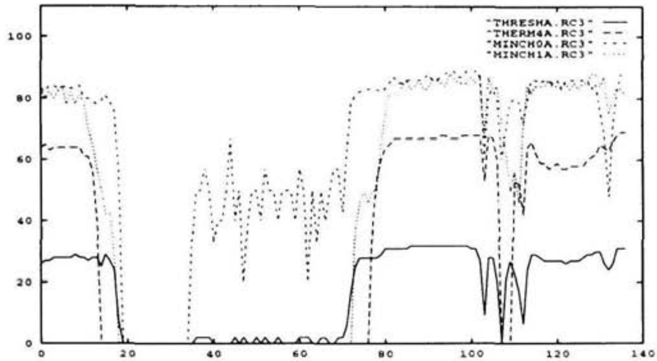

Figure 2c: test data were the 'Fuji MF2DD'

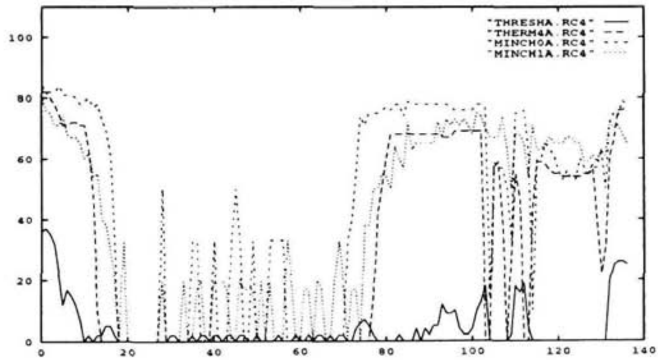

Figure 2d: test data were the 'Dysan'

Figure 2: Relative Confidence for four discriminators for Test A

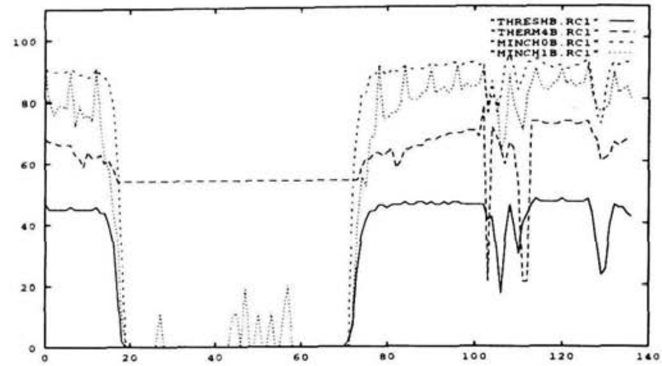

Figure 3a: test data were the 'Fuji MF2HD'

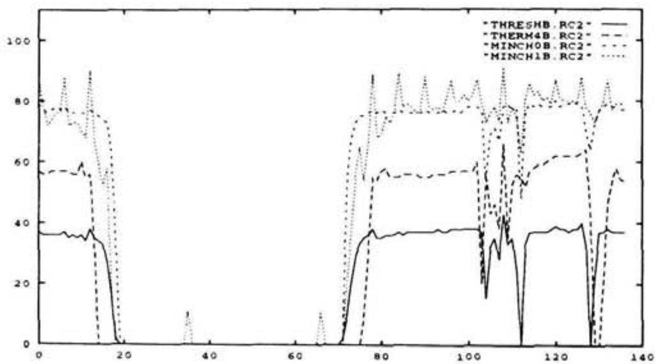

Figure 3b: test data were the '3M'

Figure 3: Relative Confidence for four discriminators for Test B

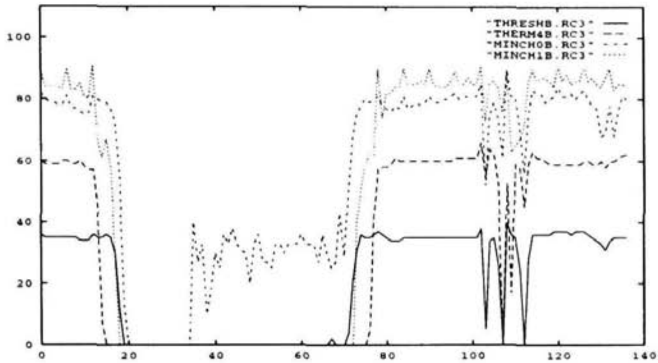

Figure 3c: test data were the 'Fuji MF2DD'

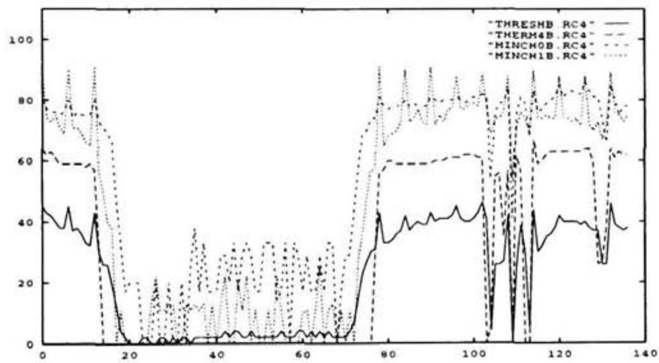

Figure 3d: test data were the 'Dysan'

Figure 3: Relative Confidence for four discriminators for Test B

The period on each graph labelled between 20 and 70 is at night, when the disk boxes were illuminated very poorly. As one would expect, the performance of all networks was reduced. However, Thermometer coding performed well on image 1 because most of it was dark.

In the period between 100 and 110, significant shadows occurred in the images, which is the cause of the sharp dips in the performance of all networks. On the A set of data, the Type 0 cells performed best over this interval; on the B set, the Type 1 cells were slightly better.

5 Conclusion

These results clearly demonstrate that the best method for pre-processing continuous data for use by a weightless network is the Type 0 Minchinton cell. In general this gives higher relative confidence, generalises better and is both more computationally and memory efficient than the other methods investigated.

References

1. I. Aleksander, W.V. Thomas. and P.A. Bowden (1984) 'WISARD: a radical step forward in image recognition', Sensor Review, pp120-124.
2. N.Wirth (1986) 'Algorithms and Data Structures', Prentice-Hall.
3. I. Aleksander and R.C. Albrow (1968) 'Pattern Recognition with AdaptiveLogic Elements', IEE-NPL Conf. Pattern Recognition.
4. J.M. Bishop, P.R. Minchinton & R.J. Mitchell (1991) 'Real Time Invariant Grey Level Image Processing Using Digital Neural Networks', Proc IMechE Conf, EuroTech Direct '91 -Computers in the Engineering Industry, pp187-199, Birmingham.
5. J.M. Bishop and P. Torr (1992) 'The Stochastic Search Network', in 'Neural Networks for Vision, Speech and Natural Language', Ed. D.J.Lingard and C.Nightingale, pp 370-387.
6. J.M. Bishop and R.J. Mitchell (1995) 'Auto-associative memory using n-tuple techniques', Intelligent Systems Engineering, Vol 3, No 4, pp 222-229.
7. R.J. Mitchell, P.R. Minchinton, J.M. Bishop, J.W.Day, J.F.Hawker and S.K.Box (1993) 'Modular Hardware Weightless Neural Networks', Proc. WNNW'93, pp123-128, University of York.
8. Cybertronix Ltd (1995) 'VISIWIN'.
9. J.M. Bishop (1989), 'Anarchic Techniques for Pattern Classification', PhD Thesis, University of Reading, UK.