The utilization of grouped n-tuple logic nodes in artificial neural Single Layer Nets

Enhanced pattern recognition characteristics

Bruce A. Wilkie
Dept. of Electronic and Computer Engineering.
Brunel University,
Uxbridge, Middlesex UB8 3PH, UK
wilkie@kae.zcu.cz

Radek Holota
Dept. of Applied Electronics and Telecommunications
University of West Bohemia
306 14 Pilsen, The Czech Republic
holota5@kae.zcu.cz

Abstract – **Based upon the n-tuple methodology originated by Bledsoe and Browning, this paper outlines the benefits obtained by using grouped n-tuple nodes to enhance the pattern selectivity of single-layer n-tuple networks.**

The principles of the n-tuple and grouped node techniques are outlined and comparative results are presented.

From the documented results it is confirmed that, compared with a conventional n-tuple system and with no higher memory requirements, the grouped node, Pseudo Huge Tuple (PHT) technique significantly improves the recognition confidence levels. Additionally, this technique tends to provide faster training and allows the user more flexibility in choosing system parameters.

Keywords-pseudo huge tuple, n-tuple, pattern recognition, single-layer net

I. Introduction

The technique of grouping n-tuple nodes to approximate the operation of much larger sizes of n-tuple was first implemented in 1988. As these grouped nodes were not true large tuple nodes they were referred to as Pseudo Huge Tuple (PHT) nodes. This type of compound node probably has been employed in various pattern recognition applications but, because of company confidentiality regarding Intellectual Property Rights (IPR) at that time academic publications were not ethically viable.

The PHT technique is an extension of the n-tuple methodology originated by Bledsoe and Browning, which is outlined in their classical paper of 1959 [1]. The initial purpose of this methodology, implemented in software, was that of recognizing binary-encoded printed alpha-numeric characters. Later, in the late 1960's and early 1970's, it was realised by Aleksander [2] that the n-tuple methodology could be implemented in hardware and thereby could provide 'real-time' recognition speeds [2]. Consequently, although of limited resolution (16x16) several hardware systems were configured. These were used for many aspects of pattern recognition such as Chromatography, Postal Code character recognition, medical diagnose of abdominal pain and the recognition of piece parts on conveyor belts to assist robotic 'pick and placing'. Later hardware systems were capable of processing images of TV resolution (512x512) at rates of 4.1667 to 12.5 recognitions per second [2,3,4].

The very significant contribution by Aleksander was that the software algorithms of Bledsoe and Browning could be configured in hardware as Single Layer Nets (SLNs) consisting of 'deterministic' logic nodes. In its simplest form, a logic node of these SLNs can be realised by a Random Access Memory (RAM) consisting of 'n' address spaces. This infers that for 'full logical functionality' an n-tuple node requires 2^n bits of memory. The patterns applied to the address inputs comprise of 'n' sample points taken pseudo-randomly from the binary input pattern data (as shown in Fig. 1). Pseudo- random mapping is employed, as in the case of digital communication systems, to maximise the entropy of the input pattern vector.

During training, each RAM stores the relevant sampled n-tuple 'n-bit' data words. After training, the RAM, in 'read' mode, operates as a look-up table with a speed only limited by the access time of the RAM.

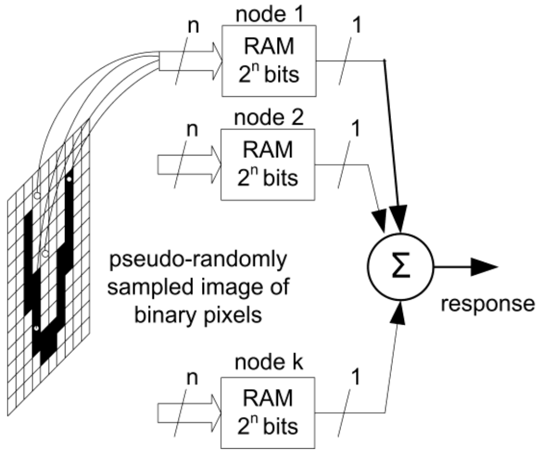

Figure 1. Single Layer Net pseudo-randomly sampling the input image

The summation of the logic nodes' outputs (zero or one) represents the response of that particular SLN (or 'discriminator').

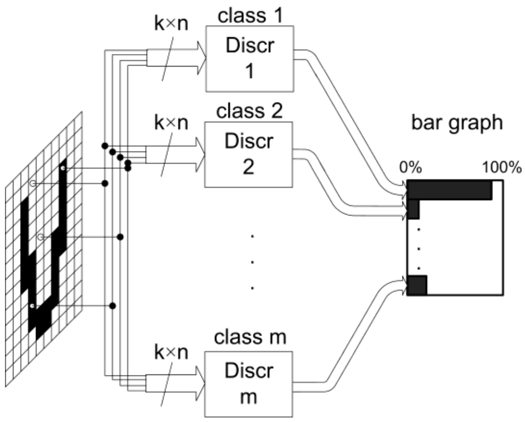

Figure 2. Multi-discriminator configuration

In general, 'n-tuple' classifier systems operate in a multi-discriminator configuration (as shown in Fig 2) where each discriminator can contain from several hundreds to several hundreds of thousands of logic nodes. During training, each discriminator is trained on each class, or classes, of patterns it is later to classify. After training and when in the classification mode, the discriminator which gives a higher response than any of the others is considered to represent the correct decision.

By 1988 it had been well established that a large size of n-tuple was required to obtain highly selective responses, However the costs for this, in hardware and non-optimised software systems, were large memory requirements and long training times (due to increased n-tuple size and the effects of 'edge' noise in a binary image). However, if the tuple size is too small, the nodes can become saturated and therefore perform no useful discriminatory properties. Consequently, some sort of compromise must be made to choose the optimum size of n-tuple. Earlier hardware systems were inflexible and, consequently, the concept of 'grouped' logic nodes was considered as a possibility of providing adequate recognition performance without the requirement of large memory overheads. The principle of the grouping method is shown in Fig. 3.

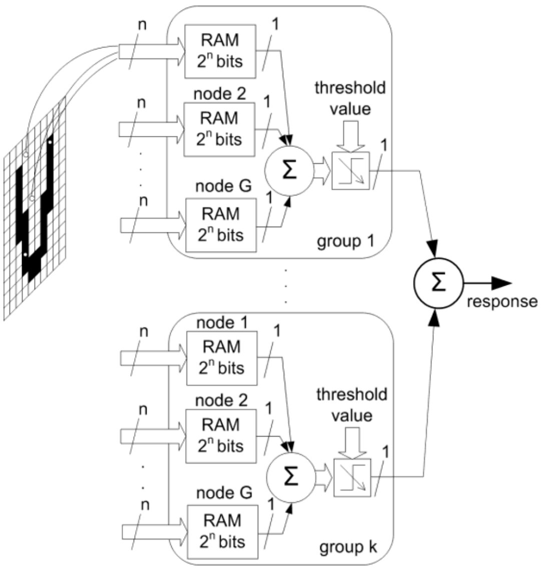

Figure 3. Single Layer Net pseudo-randomly sampling the input image and utilizing grouped nodes

The principles behind the recognition characteristics of the n-tuple and grouped node PHT techniques are outlined in section II. Some typical results are tabulated in section III.B and these verify the effectiveness of SLNs consisting of grouped n-tuple logic nodes.

II. The 'Pseudo Huge Tuple' (PHT) net

Initially, the typical performance characteristics of the conventional n-tuple SLN must be considered.

From Aleksander [2] assuming that only one pattern had been used in the training set then the most likely response of a trained discriminator (R) to an input pattern of similarity (S) in terms of area overlap or Hamming Distance is given by the following approximation.

$$\boldsymbol{R} \approx \left(\frac{S}{T}\right)^{\boldsymbol{N}} \tag{1}$$

Where S=Similarity, N=tuple size and T is the reference \training pattern (all values range from 0 to 1). By definition, T=1.

The confidence (C) is given by the following approximation.

$$\boldsymbol{C} \approx \boldsymbol{1} - \left(\frac{S_2}{S_1}\right)^{\boldsymbol{N}} \tag{2}$$

Where S1 is the maximum response and S2 is the response of the next maximum discriminator.

Generally, the responses are given as percentages i.e.

$$\boldsymbol{R} \approx \boldsymbol{100} \cdot \left(\frac{S}{T}\right)^{\boldsymbol{N}} \ [\%] \tag{3}$$

$$\boldsymbol{C} \approx \boldsymbol{100} \cdot \left[\boldsymbol{1} - \left(\frac{S_2}{S_1}\right)^{\boldsymbol{N}}\right] \ [\%] \tag{4}$$

Again, it must be emphasised that it is assumed that only one pattern has been used in the training set. However, from some several thousand hours of operational time with the original WISARD [2] and later software simulations and hardware, despite many patterns being included in the training set, the approximations derived by Aleksander have been found to be quite adequate for practical purposes.

However, the question which arose in 1988 was how would an n-tuple network perform if the output nodes were to be logically 'ANDed' together? For example, if the outputs of two 4-tuple nodes were logically 'ANDed', would the resultant discriminatory function of this compound node be similar to an 8-tuple node?

In fact, the performance of this compound 'pseudo 8-tuple' node would be identical to a true 8-tuple node providing the training set consisted of only one pattern. For more than 1 pattern in the training set, then the generalisation set of a single layer net of pseudo 8-tuple nodes would normally be considered to be the same as that of a net consisting of true 4-tuple nodes.

Unfortunately, for patterns of medium spatial frequency and high resolution, it is very difficult, and nearly impossible, to assess relative saturation and generalisation sets. In effect, the relative saturations and generalisation sets are highly dependent upon the spatial frequency statistics of the patterns.

Fortunately, in practice, it would appear that in the great majority of instances, patterns of different classes are sufficiently dissimilar and have relatively low saturation levels and generalisation sets.

Consequently, in 1988, the TV-based WISARD hardware system was modified to provide programmable Pseudo Huge Tuple (PHT) sizes from 2 to 256 (in multiples of 2).The initial results surpassed all the best expectations. It was even possible, with limited and careful training, to use 1-tuple nodes and then implement pseudo tuples up to a size of 256. Subsequently, a commercial company (PI Ltd) patented the methodology [5] and implemented it in a Transputer-based system.

Of importance, all the experimental results from various hardware and software systems utilizing the PHT technique were consistent and verified that the methodology was valid. Later, in 2000, a user-friendly PC-based system incorporating a commercial TV frame store was developed for evaluation purposes [6, 7]. This system was used to obtain the results presented in this paper.

Earlier notation referred to PHT nodes. For example, PHT8 simply meant a group of eight n-tuple nodes, and PHT256 would infer groups of 256 n-tuple nodes.

Presently, with subsequent experience, it is much easier to consider the nodes as being 'grouped' in groups of 'G'. Also, a threshold 'Gth' may be applied to the summed output of each group (normally, the same value of Gth is applied to each group and is equivalent to the 'AND' function). The value of Gth is dependent upon the application and, if required, may be adjusted for an optimum compromise between generalisation and selectivity.

The concept of 'pattern saturation' is difficult to define. Pattern saturation is only indirectly related to the actual memory saturation of each n-tuple node. In order to define 'Pattern Saturation', the alpha character 'u' has been used. The 'u' symbolises the pattern 'utilisation' of the memory as apposed to the actual saturation of the n-tuple node memory.

For example, if a single- layer net consisting of 1-tuple nodes gave a response of 50% to a white noise (equal black/white) input pattern of 50% then the actual 'Pattern Saturation' is 0% and not the memory saturation of 50%. If this net gave a response of 75% then the 'Pattern Saturation is 50% and not the memory saturation of 75%.

Similarly, if a 2-tuple net gives a response of 50% to white noise then the pattern saturation or utilisation (u) is 33.33%.

In general

$$\boldsymbol{u} \approx \mathbf{100} \cdot (\boldsymbol{r} - \mathbf{0.5}^{\boldsymbol{N}}) \cdot \left(\frac{\mathbf{2}^{\boldsymbol{N}}}{\mathbf{2}^{\boldsymbol{N}-\mathbf{1}}}\right) \ [\%] \qquad (5)$$

Where 'r' is the average response of the discriminators to an input of random 'white' noise.

In order to make an approximation, equal black/white input training patterns are assumed. In practice, this can be partially achieved if either auto thresh-holding is implemented on the input video stream or by using the most significant bit of each pixel following Histogram Equalisation.

Consequently, on classifying on a 'white / black' random noise input, the discriminators' responses are usually equivalent to those obtained from the averaged responses of the discriminators for ' all white' and 'all black' input patterns.

Similar to equation (1) the value 'u' can be used in the following empirically based approximation.

$$\boldsymbol{R} \approx \mathbf{100} \cdot (\boldsymbol{u} + (\mathbf{1} - \boldsymbol{u}) \cdot \boldsymbol{S}^{\boldsymbol{N}}) \ [\%] \qquad (6)$$

(T has been omitted in this and all following equations because, by definition, its value is always '1').

The response 'R' with grouping 'G' can be considered in the following manner.

Initially, a gross assumption is made in analogue terms that an average node has the probability 'pr' of [u + (1-u) ^{SN}] to output a '1'. Analogous to the principle of multi-mode redundancy, if a grouping of 'G' is used then the summation of a net consisting of several hundreds of nodes is most likely to give a response given by the following equation.

$$\boldsymbol{R} \approx \mathbf{100} \cdot \{[\boldsymbol{u} + (\mathbf{1} - \boldsymbol{u}) \cdot \boldsymbol{S}^{\boldsymbol{N}}]^{\boldsymbol{G}}\} \ [\%] \qquad (7)$$

However, if the n-tuple nodes are grouped by the logic 'AND' function and if 'u' is considered to represent the average number of nodes which will respond with the value of '1' then it may be assumed that only (1-u).G groups are effective. This is represented by the following approximation.

$$\boldsymbol{R} \approx \mathbf{100} \cdot \left\{[\boldsymbol{u} + (\mathbf{1} - \boldsymbol{u}) \cdot \boldsymbol{S}^{\boldsymbol{N}}]^{[(\mathbf{1}-\boldsymbol{u})\cdot \boldsymbol{G}]}\right\} [\%] \ (8)$$

Hence from equation (4) the confidence level 'C' for a given discriminator is give by the following equation/expression.

$$\boldsymbol{C} \approx \mathbf{100} \cdot \left[\mathbf{1} - \left(\frac{\boldsymbol{S}_2}{\boldsymbol{S}_1}\right)^{[(\mathbf{1}-\boldsymbol{u})\boldsymbol{N}\cdot\boldsymbol{G}]}\right] \ [\%] \qquad (9)$$

If 'u' has a low value and N and G are sufficiently high (for example, if N≥4 and u≤0.2) then the following first-order approximation can be made...

$$\boldsymbol{R} \approx \mathbf{100} \cdot \boldsymbol{S}^{\boldsymbol{G}\cdot\boldsymbol{N}} \ [\%] \qquad (10)$$

It should be noted that modifications of the PHT concept led to feasibility studies including the "Monitoring of Railway Level Crossings", "Faces Recognition/verification", "Security surveillance" and the "Monitoring of the contents of TV advertisements". These tasks were facilitated because large PHT sizes could be executed without the large memory requirements which would be required by an unmodified conventional n-tuple system.

III. Measurements and Results

A. Training and Test

The training and test sets were recorded in PAL TV format and therefore were consistent (apart from the inevitable additional analogue noise). They consisted of images of a manufacture's 'cup of soup' cartons. These patterns were chosen because, subjectively, they were similar both in monochrome and colour.

Figure 4. Images of a manufacturer's 'soup in a cup' cartons

The array size was 170 x 256, 8-bit encoded and the pseudo-random mapping coverage was chosen to be 100%.

'Snapshots' of typical Images used for the training and test sets are shown in Fig. 4.

B. Experimental Measurements

The presented results have been limited to the average responses obtained from 4-tuple and 8-tuple nets and are shown in tables 1, 2, 3 and 4.

TABLE I. Responses for N=8, G=1 and N=4, G=1

Class input	Discr. Responses NT=8 G=1				Discr. Responses NT=4 G=1			
	1	*2*	*3*	*4*	*1*	*2*	*3*	*4*
1	95	60	52	51	98	84	79	79
2	61	96	62	59	85	99	79	79
3	51	53	98	61	77	78	97	84
4	50	59	65	96	76	80	83	97
BLACK	6	9	6	3	33	41	33	27
WHITE	2	1	0	2	19	13	14	19

TABLE II. Responses for N=4, G=2, G_{TH} =2 and N=4, G=3, G_{TH} =3

Class input	Discr. Responses NT=4 G=2 G_{th} =2				Discr. Responses NT=4 G=3 G_{th} =3			
	1	*2*	*3*	*4*	*1*	*2*	*3*	*4*
1	98	71	63	62	96	59	49	49
2	73	98	72	70	61	96	61	58
3	61	64	99	72	48	51	99	61
4	62	66	73	95	49	54	63	91
BLACK	11	16	10	7	3	6	3	2
WHITE	4	5	4	3	1	0	0	1

TABLE III. Responses for N=8, G=2, G_{TH} =2 and N=4, G=6, G_{TH} =6

Class input	Discr. Responses NT=8 G=2 G_{th} =2				Discr. Responses NT=4 G=6 G_{th} =6			
	1	*2*	*3*	*4*	*1*	*2*	*3*	*4*
1	90	36	27	24	92	35	25	25
2	37	92	37	33	39	94	37	34
3	24	26	88	36	23	25	93	36
4	24	28	39	86	29	32	40	91
BLACK	6	9	6	3	6	0	0	0
WHITE	2	1	0	2	2	0	0	0

The background responses 'r' for a 'white noise' input are 25% for N=4 and 3% for N=8.

From equation (5) these values of 'r' infer that 'u'=20% for N=4 and 'u'= 2.6 % for N=8

Using a PC incorporating a commercial TV frame store, the recognition rate was 4.1667 per second.

IV. Conclusions

From the tables the following equivalences can be observed.

N=4, G=3, G_{th}=3 is comparable with N=8, G=1.

N=4, G=6, G_{th}=6 is comparable with N=8, G=2, G_{th}=2 .

It should be noted that when N=8, G=2, G_{th}=2 this is nearly comparable with an n-tuple size of 16.

If N=4, G=1 is taken as the reference and the measured values used in equations (10) and (11) this verifies that these approximations are valid.

In general, as long as the training does not result in over-generalisation then the PHT net provides a very powerful tool when using a conventional n-tuple system. The main advantages are that the pattern selectivity of a net can be increased without an increase of n-tuple memory. Also the training times tend to become shorter.

The method is also suitable for Min/Max nodes [8,9] and, for colour recognition, 'trixel' n-tuple nodes and 'trixel' Min/Max nodes [10, 11, 12, 13]. Indeed, grouping is necessary when using Min/Max nodes otherwise the recognition performance is no better that that of 'template matching'. Also, it must be noted that pseudo-random mapping must be used otherwise the technique will not provide an adequate recognition performance.

Little has been written concerning the application of the group threshold 'G_{th}' to each grouped node. The presented tables were obtained using the maximum 'G_{th}' applied to each grouped node. The selected value of G_{th} is a matter of 'trial and error' to obtain the best compromise between pattern selectivity and over-generalisation.

In practice, for any useful effect, the value of 'G' is high and the value of 'G_{th}' is extremely critical. For example, using the same training and test sets, confidence levels 'C' approaching 100% were obtained when N=4, G=128, G_{th} =119 and when N=8, G=128 and G_{th} =106. For these results, a confidence level approaching 100% infers that the maximum responses of the correct discriminators are in the order of 98% whilst the responses of any of the other discriminators are below 3%.

In summary, this paper documents the principle of the 'grouped' node (PHT) technique and presents some typical results.

The main advantage of the technique is that, compared with a conventional n-tuple system, it does not require large memory overheads, training times are shorter and it allows the user more flexibility to maximise the recognition performance. The latter point is an important issue because often it may be necessary to make a compromise between the 'pattern selectivity' and the 'generalisation' characteristics of the nets.

ACKNOWLEDGMENTS

This research has been supported by the Department of Applied Electronics and Telecommunications at the University of West Bohemia, and is based upon the original work funded and patented by PI Ltd [5].

REFERENCES

[1] W. W. Bledsoe, I. Browning, "Pattern recognition and reading by machine", in Proceedings of the Eastern Joint Computer Conference, pp. 232 – 255, 1959.

[2] I. Aleksander, "Artificial vision for robots", edited by Professor I. Aleksander, chapters 9, 10, 11, Kogan Page, ISBN 0-85038-757-4, 1984.

[3] I. Aleksander, T.J. Stonham, B.A. Wilkie. "Recognition Apparatus", GB Patent, GB 211219413, 17th July, 1985.

[4] I. Aleksander, W.V. Thomas, P.A. Bowden, "WISARD: a radical step forward in image recognition", Sensor Review, pp.120-124, July 1984.

[5] S. Turner, J. Lilley, "Recognition and Procedure and an Apparatus for carrying out the recognition procedure", United States Patent 5056147, 8th October 1991.

[6] R. Holota, "Position, size and rotation normalisation", Diploma project report, Dept. of Electrical Engineering, University of West Bohemia, May 2000.

[7] J. Trejbal, "Software realization of a colour image recognition system using n-tuple and MIN/MAX node neural networks ", Diploma project report, Dept. of Applied Science, University of West Bohemia, May 2000.

[8] B.A. Wilkie, "The use of arrays of grouped 'MAX/MIN' nodes as an aid to pattern recognition", in Proceedings of the International Conference "Applied Electronics '98'", University of West Bohemia, Sept.1998, pp.214-218, ISBN 80-7082-443-3.

[9] R. Holota, "Software and hardware realisation of neural network based on MIN/MAX nodes for image recognition", in Neural Network World, 2006, Vol. 16, No. 3, pp. 215-225. ISSN: 1210-0552.

[10] B.A. Wilkie, "A proposed colour recognition system utilising networks of grouped 'MIN/MAX' nodes", in Proceedings of the International Conference "Applied Electronics '99'", University of West Bohemia, Sept. 1999, pp.211-219, ISBN 80-7082-512-X.

[11] R. Holota, J. Trejbal, B.A. Wilkie, "Software realisation of a colour image recognition system with an image normalisation stage", in Proceedings of the University of West Bohemia , section 1, vol.4/2000, pp.75-86, ISBN 80-7082-718-1, ISSN 1211-9652.

[12] B.A. Wilkie, R. Holota, J. Firt, "A colour image recognition system utilising networks of trixel n-tuple and MIN/MAX nodes", in Proceedings of the International Conference "Applied Electronics '2001'", 5-6 Sept.2001, University of West Bohemia, Pilsen, Czech Republic, pp.262-269, ISBN 80-7082-758-0.

[13] B.A. Wilkie, "The Comparative Recognition of Monochromatic and Chromatic Images using Networks of both Trixel N-Tuple and Trixel 'MIN/MAX' nodes", in Proceedings of the International Conference "Applied Electronics '2002", University of West Bohemia, Sept. 2002, 196-219, ISBN 80-7082-881-X.