Radek Holota, Jan Trejbal, Bruce A. Wilkie

Software realisation of a colour image recognition system with an image normalisation stage

Key words: image processing, image recognition, retino-cortical transform, N-tuple node, MIN/MAX node

Abstract

The task of pattern recognition can be greatly facilitated if position, rotation and size normalisation are incorporated within the pre-processing stage. Having obtained the required normalisation, then either feature extraction, template matching, artificial neural network systems or n-tuple logic node techniques may be used for object recognition.

This paper outlines the general concepts of position, rotation and size normalisation followed by n-tuple and MIN/MAX node techniques for colour image recognition. The developed software system, which includes these normalisation and recognition methods, is briefly described.

Finally, some results are presented which illustrate the performance obtainable from the developed software colour recognition system.

1 Introduction

In order to assist the operation of pattern recognition, it is beneficial to ensure that the objects to be recognised are initially placed in a predefined position and orientation. Additionally, in certain applications, it may be necessary to incorporate some form of size normalisation.

If conventional image processing techniques are employed then, initially, after edge tracing, the centroids of the object are calculated to obtain the position, followed by finding the maximum, or minimum, radius vector to assess object rotation. The image may then be rotated by an amount depending upon the position of the maximum, or minimum, radius vector. Also, the use of retino-cortical transforms (polar co-ordinates mapped to rectangular co-ordinates) has also yielded satisfactory results [1,2,3,4] and, in order to increase speed, optical methods have been investigated [5,6,7,8] and special-purpose radial scan CCDs have been developed [9,10,11].

Adopting similar approaches, software simulations and prototype hardware systems have been developed at Brunel University [12,13,14,15,16,17,18,19]. However, these have become dated as they used the limited available computer power and technology of the early 1990's and hence they required considerable 'up-dating'.

Consequently, designs for new software systems were initiated at Brunel University [20,21] but, unfortunately, due to the lack of resources at that time, the systems were only partially operational. Subsequently, the presently documented system, which operates successfully, was implemented and includes an image pre-processing stage [22] and an image recognition stage [23]. The image pre-processing stage uses the method of 'running sums', for position normalisation, and the retino-cortical transform for rotation and size normalisation. Alternatively, also included, is the minimum enclosing rectangle method which implements both position and size normalisation. All these methods can be utilised by the image recognition stage; this part uses n-tuple and MIN/MAX node methods for image recognition. Sub-grouping of the nodes is also available in both methods. This integrated system works with images in a RGB (Red, Green, and Blue) colour format of 24bits.

An overall block diagram of the system is shown in Figure 1 and a flow diagram is shown in Figure 2.

2 Image Processing

2.1 Position normalisation.

This program enables one to choose from two algorithms for position normalisation.

The first of them is used when the 'position normalisation check', in the 'normalisation settings' dialogue, is enabled. In this algorithm, the 256x256 image is taken and stored in a buffer. Each pixel is then compared with the threshold and the running sums for each row and column are calculated. The total running sum is calculated as the sum of all the running sums of rows. If the total running sum is not equal to zero then the x- and y-centroids are calculated. The x-centroid is the number of the column where the running sum is greater or equal to half of the total running sum. The same principle is used for the y-centroid: in this case, the y-centroid is obtained when the running sum of the rows reaches a value equal to, or greater than, half the total running sum. When the x- and y-centroids are known then the new image, which has the object in the centre, is extracted and stored. This method solves the problem of the wrap around effect, which occurred in previous systems. A disadvantage is that of lower operational speed because the new image must be extracted and stored.

The second algorithm is used when only the size normalisation check is enabled. This algorithm covers position and size normalisation. Part of this programme, which performs this algorithm, is described in section 2.3 relating to Size Normalisation.

2.2 Rotation normalisation using the retino-cortical transform

In the field of image processing, a great deal of work has been directed towards object recognition that is independent of size, position and orientation. One approach, which provides the means for size and rotation normalisation, is that of retino-cortical mapping. In essence, this method transforms changes of size and angular rotation of the retinal image into orthogonal shifts of the cortical image which simplifies subsequent processing.

The example in figure 3 illustrates the radial scan (retino-cortical) transform of a skeletonised square. It is assumed that the object has been first centralised within the field of view (i.e. position-normalised). Figure 3 indicates how the retino-cortical transform (the cortical image) can be normalised and how the inverse transformed image would appear. If the upper right-hand diagram of figure 3 is considered, then the subtraction of the address of the first maximum radial scan (corresponding to the first maximum radius vector) implements a vertical 'wrap-around' shift to the start address. This results in the rectangular display as is shown in the bottom right-hand diagram of figure 3. When the inverse retino-cortical transform is performed on this rectangular co-ordinate image, then the output display is as shown in the bottom left-hand diagram of figure 3. By inspection, if the 'square' in the upper left-hand diagram of figure 3 is rotated, then after the requisite address offset manipulation, the normalised image always will be the same as that shown in the bottom left-hand diagram of figure 3. A manual pre-set offset may be also added to rotate the normalised image to any required angle of rotation.

2.3 *Size normalisation*

Two methods for size normalisation are available in this system.

The first is mainly intended for use with rotation normalisation. This method processes the co-ordinates of the image in the cortical plane. This image is expanded as shown in figure 4 such that the size of the MRV (Maximum Radius Vector) is equal to the maximum radius. The expansion is only in one direction. First of all, the scale factor is calculated from the size of the MRV and maximum radius. Each pixel from the non-normalised buffer is read and its new position is calculated. All pixels, between the previous and present new position, are filled with the previous pixel. So the expansion is achieved. When the pixels are stored to the normalised buffer, the rotation normalisation algorithm is included.

The second method can be used when rotation normalisation is not required. Known as the 'Minimum Enclosing Rectangle' technique, this method combines both position and size normalisation. Firstly, all pixels are read and compared with the threshold. The minimum and the maximum values of the x- and y- co-ordinates are stored. These values determine the minimum rectangle, which includes all black pixels. The scale factor is calculated from the size of the greater side of the rectangle and the size of the window (256x256 in this system). The rectangle is shifted to the left upper corner and then the algorithm from the first method is used. The algorithm is only modified for 2-D expansion. There is a possibility that reading from outside of the image (256x256) is required and, if so, then the white pixel is read.

As shown in figure 5, this algorithm combines size and position normalisation in a single architecture. The object is not centred but shifted to the top-left-hand corner of the image, after which, the size normalisation is applied.

The expansion ratios are calculated for size normalisation. The same value for vertical and horizontal stretching must be used if the shape is not to be distorted. The lesser of the two ratios is used as the size ratio for normalisation.

$$\text{Vertical expansion ratio} = \frac{\text{height of window}}{\text{height of object}}$$

$$\text{Horizontal expansion ratio} = \frac{\text{width of window}}{\text{width of object}}$$

3 Image recognition

3.1 *The N-Tuple Methodology*

The n-tuple methodology was originated by Bledsoe and Browning in 1959 [24] for the purpose of recognising printed characters. In principle, their technique is equivalent to that of using Single Layer Networks (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. The pattern applied to the address inputs is comprised of 'n' sample points taken pseudo-randomly from the input data and is termed an 'n-tuple'. 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 representing a full functionality logic node. The summation of the logic node outputs represents the response of that particular SLN (or 'discriminator').

In general, ‘n-tuple’ classifier systems operate in a multi-discriminator configuration where each discriminator can contain from several hundred to several tens 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. The discriminators’ responses may be represented either numerically or as a bar graph (histogram) display. It should be noted that each bar in the bar graph display represents the sum of the individual responses obtained from the logic nodes that constitute each discriminator.

In order to perform colour recognition, the networks are configured from ‘Trixel’ N-tuple nodes (TNT nodes) as shown in figures 6 and 7.

3.2 The 'MIN/MAX' node

Networks of grouped MIN/MAX nodes incorporate many of the techniques employed by the n-tuple classifier methodology. The MIN/MAX node technique may be considered similar to template matching, which measures the nearest distance of the test pattern to stored reference patterns. However, unlike template matching, for MIN/MAX nodes, no distance measure calculations are required as each MIN/MAX node operates as a simple ‘look up table’ and thereby enables fast operational speeds. Also, several training images can be stored within each net and thereby provide several patterns per class (as in an n-tuple system).

Early software and hardware implementations of networks of grouped MIN/MAX nodes at Brunel University [25,26,27,28] indicated that they provided powerful pattern recognition properties. Of importance, if sub-grouping of these nodes was implemented and variable thresholding applied to these groups, then both the generalisation and discriminatory characteristics could be controlled and optimised for specific applications.

Subsequently, work at both the University of West Bohemia and Brunel University [20,21,22,23,29,30] was undertaken to provide ‘user-friendly’ software operational systems.

In principle, during training, for the relevant class, each node stores the absolute Minimum and Maximum (MIN/MAX) values, which occur. On classification, each node will then respond with a '1' to any level between, and including, the 'MIN/MAX' values, which were stored during training. In order to improve generalisation, offsets, greater than the maximum value and less than the minimum value, may be added, or subtracted, to allow for noise, illumination, translational and rotational variations not included in the training set.

The classical n-tuple techniques adopted are in that the pixels must be pseudo-randomly mapped from the input pattern space, and a group of 'G' 'MIN/MAX' nodes corresponds to an n-tuple node; for example, if G = 8, then this is analogous to an 8-tuple node.

In order to perform colour recognition, the networks are configured from ‘Trixel’ MIN/MAX nodes (TMM nodes) as shown in figures 8 and 9.

4 Results

4.1 Examples of the results obtained

Figures 10, 11, and 12 show some of the system results obtained. In this example, three BP ‘points awarded’ (swipe-type) cards were recognised. These cards were chosen because they were of a similar colour and only had minor differentiating features. The minimum enclosed rectangle method for the position and size normalisation was used. The processing window, the normalised image and the train and recognition dialogues are displayed in the figures.

For example, figure 10 shows the recognition of the first BP card. The response of the 8-tuple discriminator number ‘1’ is 99%, and the responses of the next discriminators are 29% and 35%. The response of the MIN/MAX node discriminator number ‘1’ is 99%, and the responses of the next discriminators are 52% and 41%.

4.2 System performance

The operational speed of the software system is very dependent on the performance of the computer used. All rates, which are tabulated, were achieved with a Pentium III operating at 500MHz with a memory of 128MB.

System performance including normalisation						
Image size	Number of discriminators	Coverage %	Tuple size	N-tuple speed	MIN/MAX speed	Combined speed
256x256	4	10	6	3.95/3.98	4.05/4.65	3.74/3.84
256x256	4	5	8	4.76/4.90	4.26/4.38	3.91/4.61

5 Conclusions

The ‘Visual C++’ software realisation of the image recognition system with the image normalisation part was developed and successfully tested. The results indicate that the software realisation is suitable for the presentation, testing and research of image recognition and normalisation techniques, but, as yet, it is not suitable for real-time applications. To reach a ‘ real-time’ TV frame rate of 25 frames per second, then either the use of specialised hardware is necessary or a higher processor speed is required. Nevertheless, the developed system is sufficient for evaluation purposes before proceeding to any hardware implementation of the software methods employed.

Acknowledgement

This research was supported by the grant GA ÈR 102/99/1074.

Resumé

S potøebou rozpoznávání obrazu se setkáváme èím dál èastìji v mnoha oborech lidské èinnosti. Zpracování a rozpoznání obrazu je napøíklad èasto využíváno v lékaøské elektronice. Hledání nových metod rozpoznávání obrazu a zdokonalení stávajících algoritmù je proto pøedmìtem dalšího vývoje v této oblasti. K rozpoznávání obrazu byly zde použ ity metody n-tuple node a MIN/MAX node. Ke zvýšení úèinnosti byla do systému zaèlenìna èást pro pøedzpracování obrazu. Vtomto pøípadì se jedná o normalizaci pozice, natoèení a velikosti pozorovaného objektu.

References

1. WEIMAN, C.F.R., and CHAIKIN, G., 1979, ''Logarithmic Spiral Grids for Image Processing and Display'' Computer Graphics and Image Processing 11, Academic Press, pp.197-226.
2. BRACCINI, G., GAMBARDEYIA, G., SANDINI, G., and TAGLIASCO, V., 1982, "A model of the early stages of the human visual system: functional and topological transformations performed in the peripheral visual field'', Biological Cybernetics, Springer-Verlag, pp.47-58.
3. MASSONE, L., SANDINI, G., and TAGLIASCO, V., 1985, "Form-invariant topological mapping strategy for 2-D shape recognition '', Computer Vision, Graphics and Image Processing 30, Academic Press, pp.169-188.
4. HECHT-NIELSON, R., 1988, "Neuro-computing: picking the human brain'', IEEE Spectrum, pp.36-41.
5. CASASENT, D., and PSALTIS, D., 1976, " Position, rotation and scale invariant optical correlation'', Applied Optics, vol. 15, no. 7.
6. WATSON, A. B., 1987, "The cortex transform: rapid computation of simulated neural images'', Computer Vision, Graphics and Image Processing 39, Academic Press, pp.311-327.
7. UDOMKESMALEE, S., 1989, "Intelligent optical correlator'', Optical Engineering, vol.28, no.5.
8. SHENG, Y., 1989, "Fourier-Mellin spatial filters for invariant pattern recognition", Optical Engineering, vol. 28, no. 5.
9. MASSIMO, T., and SANDINI, G., 1990, "On the estimation of depth from motion using an anthropomorphic visual sensor''. Lecture notes in Computer science, Computer Vision – ECCV90, 1st Conference on Computer Vis ion, Antibes, France.
10. FERRARI, F., NIELSEN, J., QUESTA, P., and SANDINI, G., 1995, "Space invariant imaging'', Sensor Review, vol.15, no. 2, pp.18-20, ISSN 0260-2288.
11. BRAGGINS, D., 1999, "Intelligent Cameras Report", IIP News Letter, Issue 17 (SIRA TECHNOLOGY CENTRE), June 1999.
12. MAY, D.S., 1990, " A prototype 'radial scanning' image processing system'', B.Eng. final year report, Dept. of E.E.& E., Brunel University, April 1990.
13. WILKIE, B. A., 1990, "Rotation normalisation using radial circular scanning techniques'', Internal 'feasibility design' report, Dept of E.E.& E., Brunel University, June, 1990.
14. REICHERT, G., 1990, " Implementation of a rotation normalisation image preprocessor'', M.Sc. dissertation, Dept of E.E.& E., Brunel University, Sept. 1990.
15. LUCKIN, M., 1992, "Implementation of a position-normalisation module and investigation of radial scanning", B.Eng. final year report, Dept. of E.E & E., Brunel University, April 1992.
16. WALKER, C.W.M., 1992, Real-time object characterisation system", M.Sc. dis sertation, Dept. of E.E.& E., Brunel University, Sept. 1992.
17. WILKIE, B. A., STONHAM, T.J., 1994, "Position, rotation and size normalisation of TV images at half frame rate'', 'Specialised Processors for Real-Time Image Analysis', Springer Verlag, chapt. 6, pp.57-69, ISBN 3-540-57016-0 and ISBN 0-387-57016-0.

18. WILKIE, B.A., 1997, “The position, size and rotation normalisation of a single object within an image at TV frame rate”, Proceedings of the International Conference “Applied Electronics ‘97’”, University of West Bohemia, pp.208-234, ISBN 80-7082-337-2.
19. WILKIE, B.A., 1977, “Object recognition after normalisation for position, size and rotation”, Proceedings of the International Conference “Applied Electronics ‘97’”, University of West Bohemia, pp.200-207, ISBN 80-7082-337-2.
20. FIRT, J., 1999, “System for image recognition using the TMS 320 C30 DSP”, Diploma project report, Department of Electrical Engineering, University of West Bohemia, May 1999.
21. KOTALA, Z., 1999, “A software-based image processing and pattern recognition system utilising MIN/MAX node networks”, Diploma project report, Department of Applied Sciences, University of West Bohemia, May 1999.
22. HOLOTA, R., 2000, “Position, size and rotation normalisation”, Diploma project report, Dept. of Elecrical Engineering, University of West Bohemia, May 2000
23. TREJBAL, J., 2000, “Software realisation of a colour image recognition system using n-tuple and MIN/MAX node neural network”, Diploma project report, Dept. of Applied Sciences, University of West Bohemia, May 2000.
24. BLEDSOE, W. W., BROWNING, I., 1959, “Pattern recognition and reading by machine”, Proceedings of the Eastern Joint Computer Conference, pp.232-255.
25. WANG, Y.S., 1994, “Multiple Adaptive Grouped Nodes Image Matching System MAGNIMS”, Internal report no. 1, Dept. of Manufacturing and Engineering Systems, Brunel University, Jan. 1994.
26. WANG, Y.S., 1994, “Multiple Adaptive Grouped Nodes Image Matching System MAGNIMS”, Internal report no. 2, Dept. of Manufacturing and Engineering Systems, Brunel University, March 1994.
27. OLLIER, R.J., 1994, “Grey Scale Template Matching System (GSTMS)”, B.Eng. final year report, Dept. of E.E.& E., Brunel University, April 1984.
28. WILKIE, B.A., 1998, “The use of arrays of grouped ‘MAX/MIN’ nodes as an aid to pattern recognition”, Proceedings of the International Conference “Applied Electronics ‘98’”, University of West Bohemia, pp.181-218, ISBN 80-7082-443-3.
29. SAJTOS, P., 1998, “Software Implementation of a Hybrid Neural Network”, Proceedings of the International Conference “Applied Electronics ‘98’”, University of West Bohemia, pp.181-183, ISBN 80-7082-443-3.
30. Wilkie. B.A., 1999, “A proposed colour recognition system utilising networks of grouped MIN/MAX nodes”, Proceedings of the International Conference “Applied Electronics ‘99’”, University of West Bohemia, pp.211-219, ISBN 80-7082-512-X.

Doc. EurIng. Bruce A. Wilkie, PhD. CEng. MIEE
Dept. of Electrical Engineering & Electronics
Brunel University
Uxbridge, Middlesex, UB8 3PH
United Kingdom
Tel. ++44 1895 274000
Fax. +44 1895 258728
E-mail: bruce.wilkie@brunel.ac.uk

Ing. Radek Holota Ph.D. student
Department of Applied Electronics
Sedláèkova 15, PLZEÒ 306 14
Tel.: 019/7236881 linka 242
Fax.: 019/7223315
E-mail: radek.holota@post.cz
Ing. Jan Trejbal
E-mail: trejbal4@centrum.cz

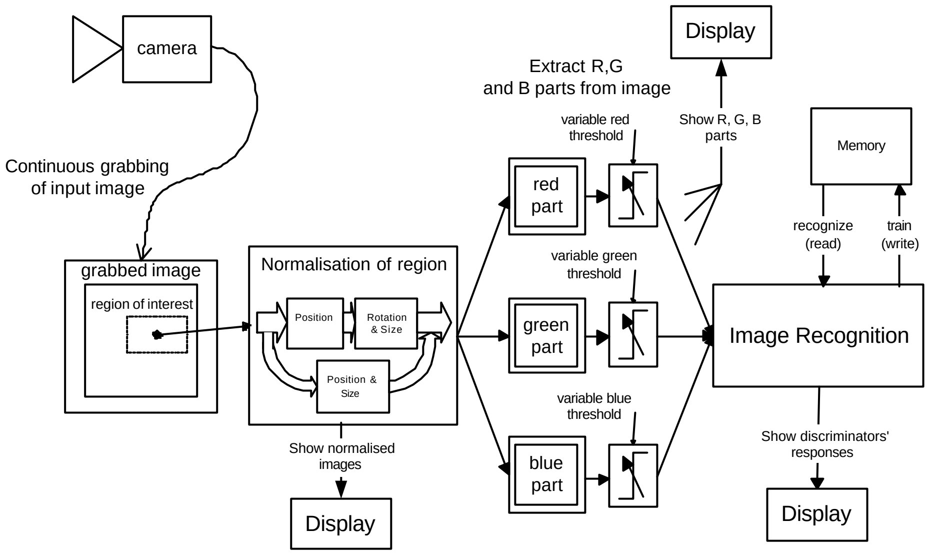

Figure 1. Overall block diagram of system

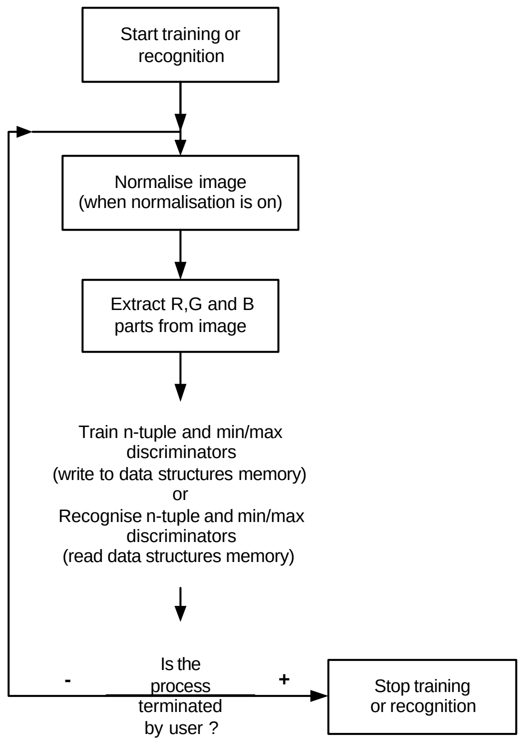

Figure 2. Flow diagram

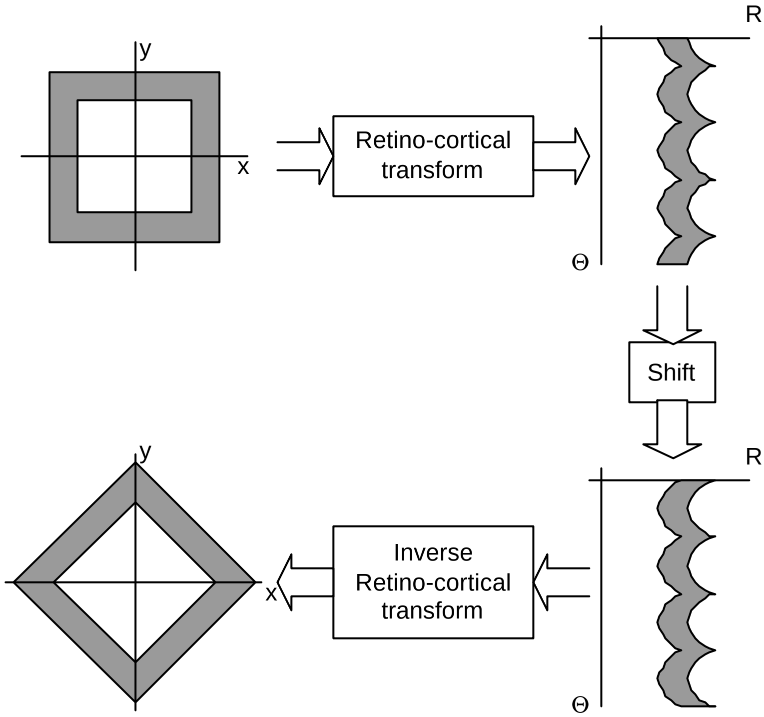

Figure 3. Usage of the retino-cortical transform for rotation normalisation.

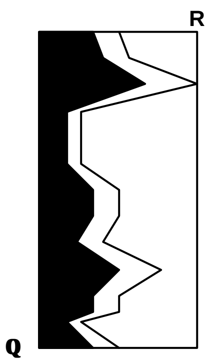

Figure 4. Size normalisation of the cortical image

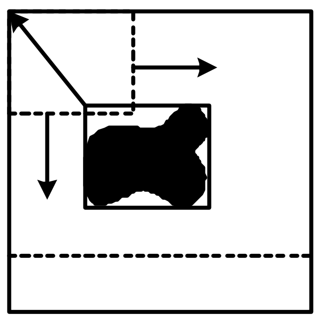

Figure 5. Minimum Enclosing Rectangle method

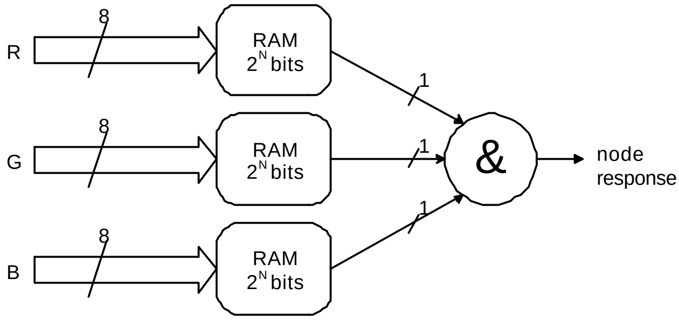

Figure 6. RGB 8-tuple node (the 'Trixel' N-tuple node)

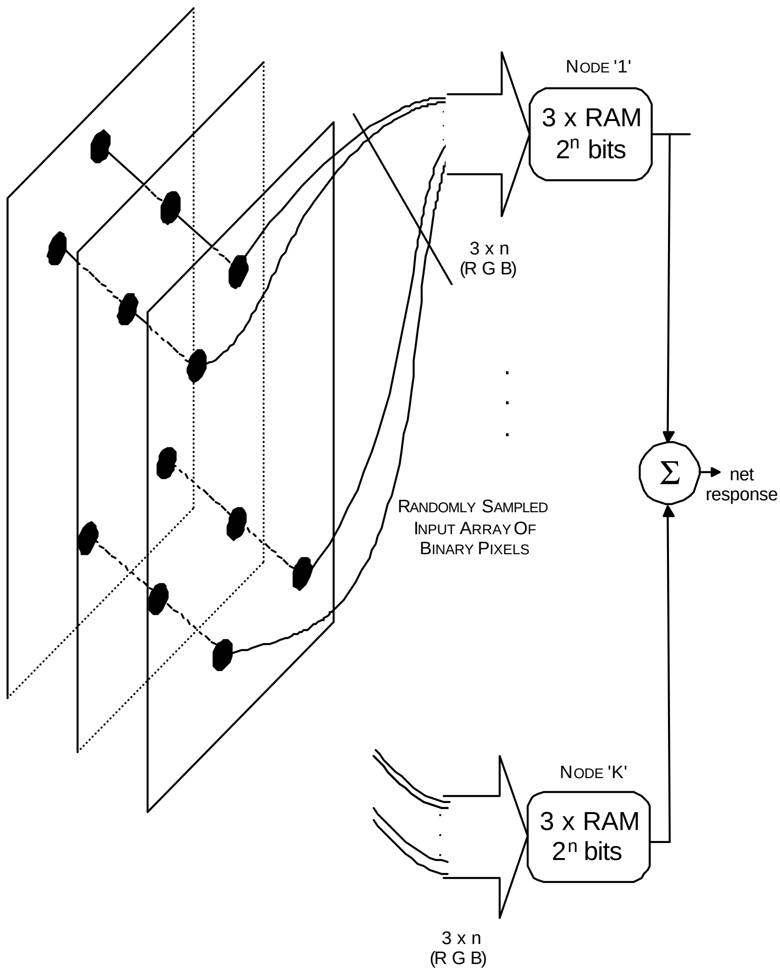

Figure 7. Single discriminator 'Trixel' N-tuple node SLN

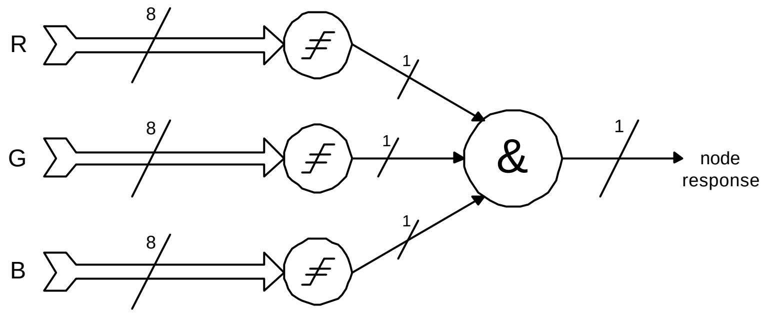

Figure 8. RGB Min/Max node (the ‘Trixel’ Min/Max node)

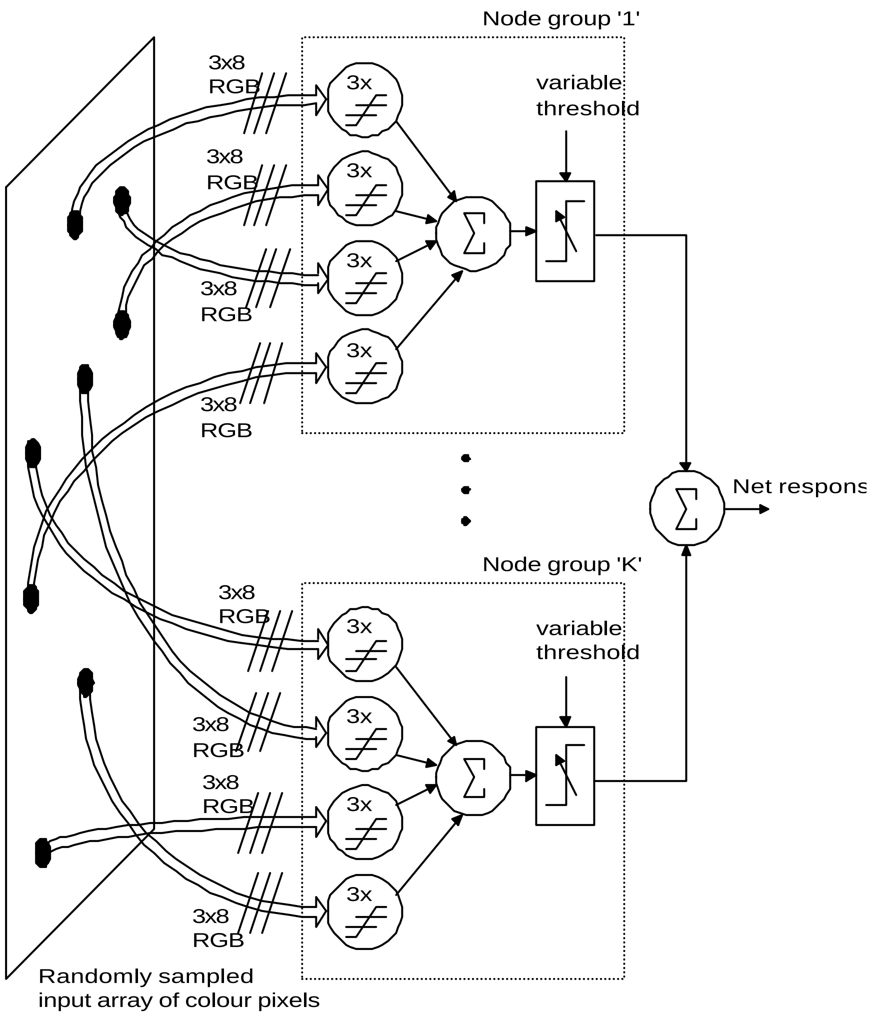

Figure 9. Single discriminator TMM node SLN

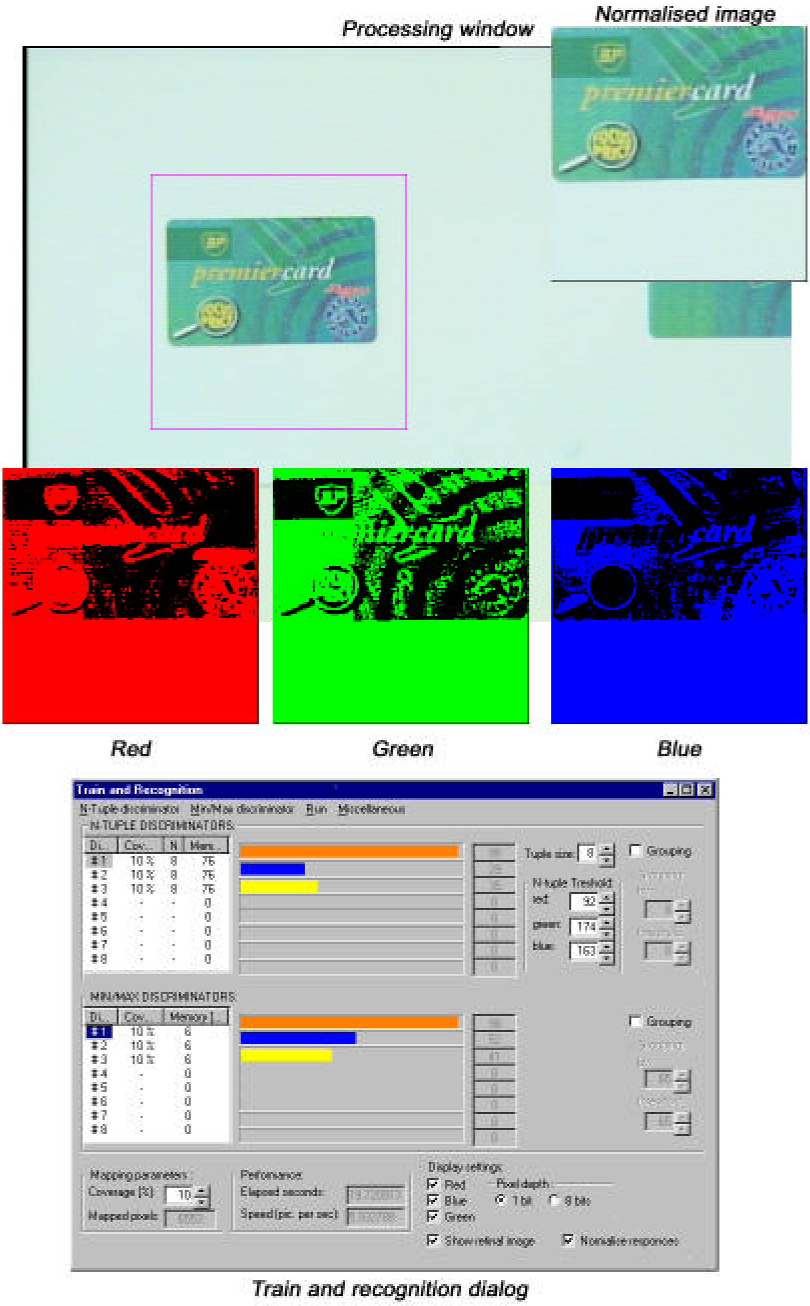

Figure 10. Image recognition with position and size normalisation

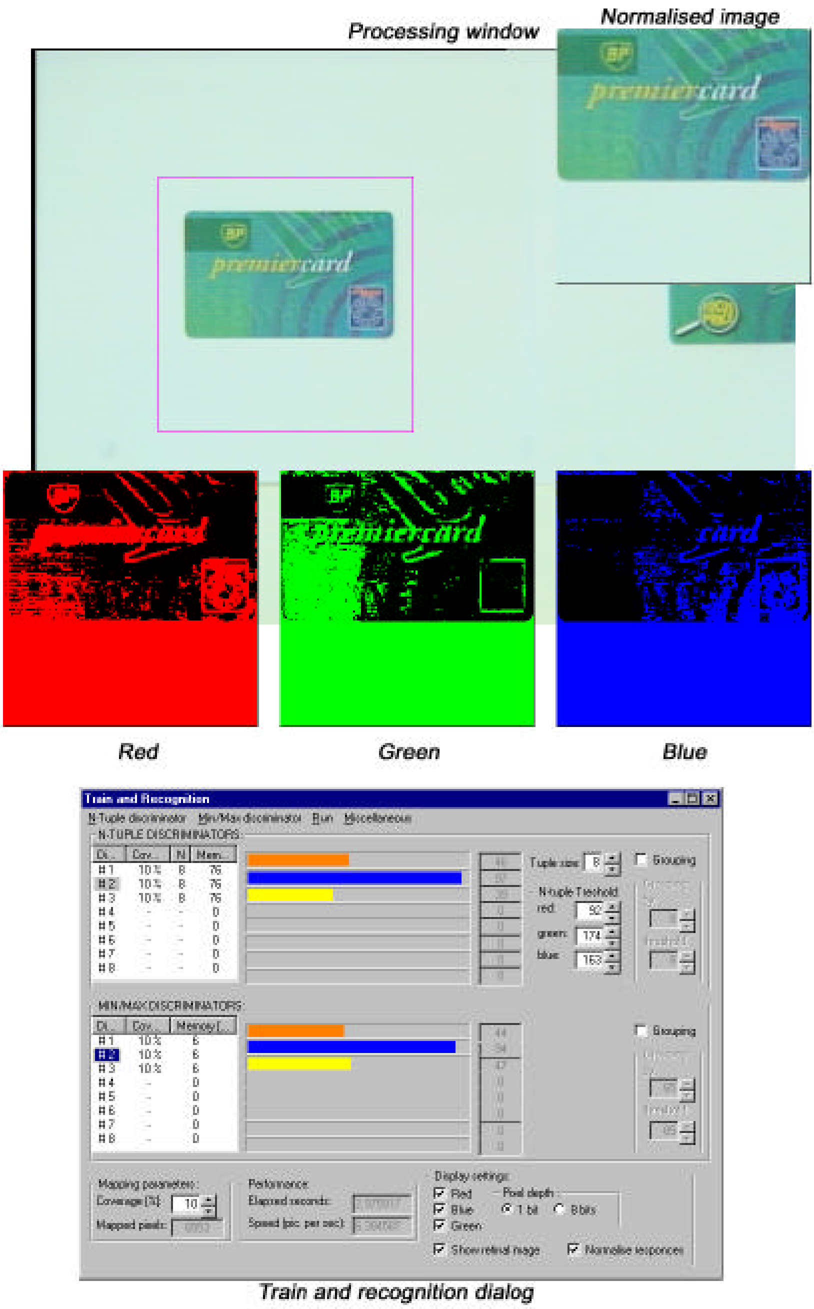

Figure 11. Image recognition with position and size normalisation

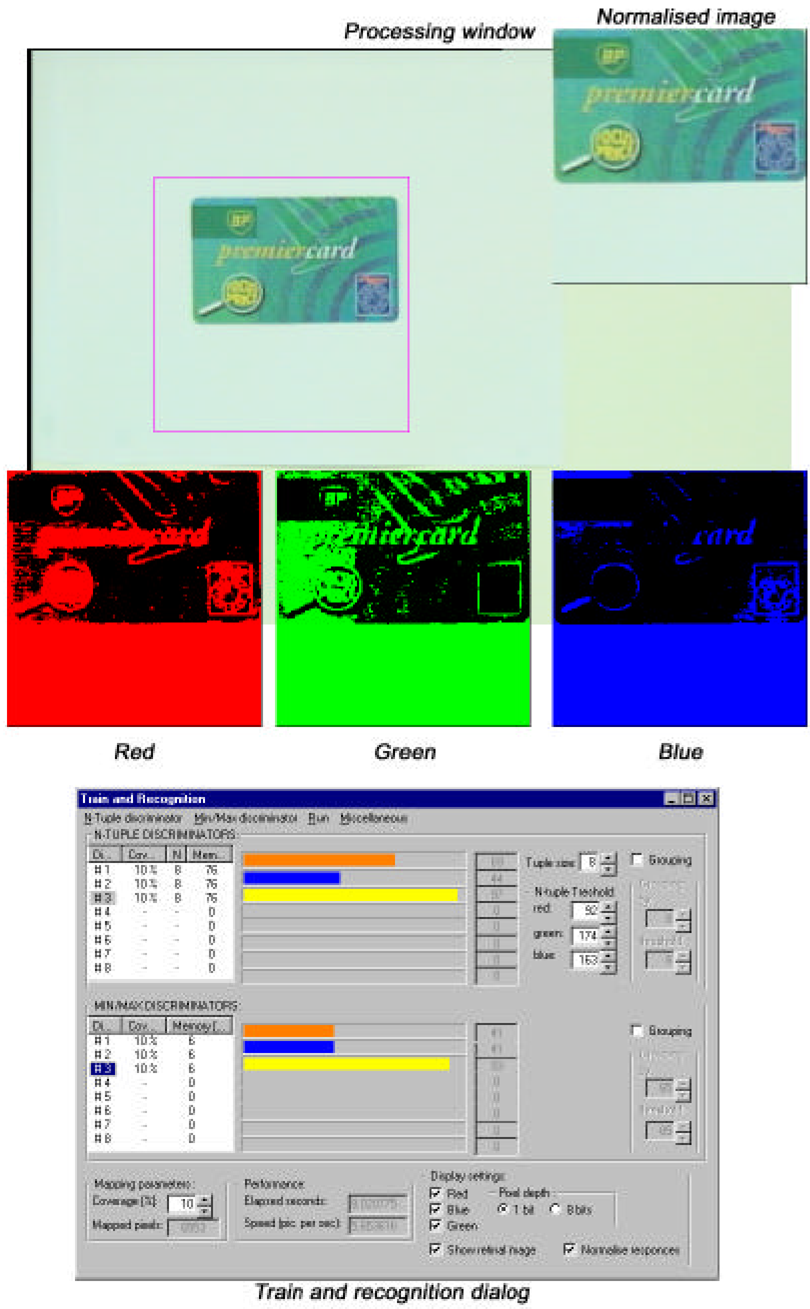

Figure 12. Image recognition with position and size normalisation