Technical Paper

Neural classifiers for automated visual inspection

D T Pham, BE, PhD and **E J Bayro-Corrochano**, MS, PhD
School of Electrical, Electronic and Systems Engineering, University of Wales College of Cardiff

This paper discusses the application of a back-propagation multi-layer perceptron and a learning vector quantization network to the classification of defects in valve stem seals for car engines.
Both networks were trained with vectors containing descriptive attributes of known flaws. These attribute vectors ('signatures') were extracted from images of the seals captured by an industrial vision system. The paper describes the hardware and techniques used and the results obtained.

1 INTRODUCTION

Some critical automotive components require 100 per cent inspection. The use of human inspectors is economically not feasible when the volume of production is high.

The purpose of the study reported in this paper was to explore fast neural processors to assist the automated visual inspection (AVI) of such high-volume critical components. Only supervised neural nets were considered, namely the back-propagation multi-layer perceptron (**1, 2**) and the learning vector quantization network (**3, 4**).

The study focused on two tasks: the recognition of the shape of the inner perimeter of a valve stem seal in a car engine (see Fig. 1) and the classification of surface flaws on a seal. These tasks are currently carried out manually on a sampling basis. They would be difficult to perform using conventional algorithmic AVI techniques as the artefacts to be recognized or classified are not defined in precise mathematical or quantitative terms. Instead, classes of artefacts are subjectively determined from representative samples provided by quality control personnel.

The paper is organized as follows. A brief review of neural network applications in image processing and automated visual inspection is provided in Section 2. The back-propagation multi-layer perceptron is described in Section 3 and the learning vector quantization network in Section 4. The test equipment is explained in Section 5. Section 6 presents the results obtained. Finally, the conclusions of the study are given in Section 7.

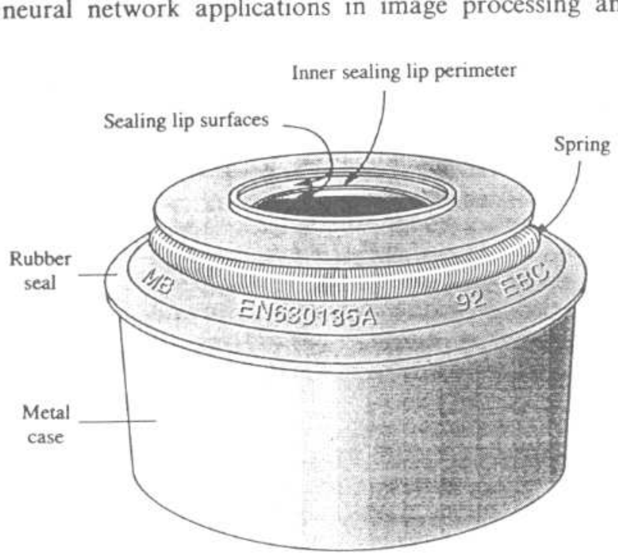

Fig. 1 Valve stem seal

The MS was received on 23 November 1992 and was accepted for publication on 8 September 1993.

2 NEURAL NETWORKS AND MACHINE VISION

A neural network is a computing system made up of several interconnected elementary processing units operating in parallel. Due to this parallelism neural networks are eminently suitable for machine vision applications, including automated visual inspection, which can benefit from a high degree of concurrent operation.

A neural network is taught a particular task by being shown examples rather than through programming. This considerably simplifies its use and is one of the most attractive features of neural computing. Another often-mentioned advantage of neural networks is their ability to a certain extent to make generalizations from the examples with which they have been trained. This enables them to deal with noisy input data and to provide solutions to problems that they have not exactly previously encountered.

In automated visual inspection, the main problem is that of pattern recognition. Prior to a pattern of an appropriate format being isolated for recognition purposes, images must be processed (filtered, enhanced, restored and segmented). Neural networks have been described for a wide range of image processing and pattern recognition applications, some examples of which are presented in the following paragraphs. None of these examples is directly related to the problem of valve stem seal inspection which, to the knowledge of the authors, has not previously been solved using neural computing. The examples are cited to illustrate the scope of the neural network approach and identify general areas where it can most profitably be employed in practice.

Tenorio and Hughes (**5**) developed a method for real-time noisy image segmentation. They employed a non-causal, Markov-field, neural-network-based segmentation method which is insensitive to rotation, scaling, translation and multiplicity of objects. Eichmann *et al.* (**6**) used Kohonen's associative memory method (**3**) for classification and restoration of degraded images. There are diverse applications using multi-layer perceptrons and the back-propagation method described by Werbos (**1**) and Rumelhart *et al.* (**2**). One relevant property of this approach is that it can solve recognition problems in complex decision regions. For example, Yang and Guest (**7**) have used a back-propagation neural network for rotation-invariant pattern recognition. Pham and Bayro-Corrochano (**8**) have developed a back-propagation neural network for noise filtering and edge operations.

Other basic pattern operations performed by a neural net include pattern completion, filtering and recognition (**9, 10**). Aleksander *et al.* (**11**) developed WISARD, a multi-discriminator system which has each of its discriminators (simple RAM networks) trained to recognize a different class of object. WISARD has been reported to have applications in security monitoring and industrial visual inspection. Recently, Glover (**12**) reported the development of a hybrid optical Fourier/electronic neurocomputer machine vision system. In this work, Glover attempted to demonstrate the capability of the system to perform effectively multi-class pattern discrimination for manufacturing and packaging inspection tasks based on global analysis of image texture and shape information at speeds of up to 15 images/second. Back-propagation and counter-propagation networks (**13**) were chosen for the experiment. The back-propagation algorithm was found consistently to perform as well as or better than the Fisher linear-discriminant statistical pattern recognition technique (**14**).

Using a different approach, Gouin (**15**) presented a system consisting of a two-stage arrangement of 'restricted Coulomb energy' (RCE) networks implementing a type of hierarchical filtering. One application was the inspection of rivet-head formations in aircraft floorboard assemblies, which is very subjective and hard to describe algorithmically.

Another interesting application of neural networks has been presented by Beck *et al.* (**16**). This is a self-training visual inspection system using a standard back-propagation neural-network-based classifier. The system consists of a control unit and a signal processing unit, together with a connectionist classifier. The control unit can both generate the training set and perform the function of teacher to the classifier. The signal processing unit compresses a two-dimensional image into a one-dimensional signal, extracts potentially significant flaws and sends them to the classifier. The system has been applied to two inspection tasks involving two-dimensional surfaces characterized by a known intensity distribution.

From the above review, it can be noted that implementations of neural-network-based *image processing* only exist in laboratory environments (**5–8**). This is because the large amount of raw image data to be processed by a neural network would necessitate large-scale hardware to achieve the high speeds required for industrial use. This hardware is not yet available at economic prices. On the other hand, *feature extraction* and *pattern recognition* applications [such as those described by Gouin (**15**) and Beck *et al.* (**16**)] are practically feasible since the amount of data involved is in general much smaller and either modest hardware networks or simulated software networks can be employed. In an industrial application, such as the inspection of valve stem seals described in this paper, the most expedient approach tends to be to use dedicated conventional image processing hardware for routine tasks such as filtering and segmentation and to reserve neural networks for the high-level tasks of feature extraction and pattern recognition.

3 BACK-PROPAGATION MULTI-LAYER PERCEPTRON

The back-propagation multi-layer perceptron (BMLP) is based on the perceptron, the oldest type of artificial neural network (**17**). A BMLP normally consists of an input layer, an output layer and one or more hidden layers of neurons (see Fig. 2). Signals propagate in one direction from the input through the hidden layers to the output layer. Consequently, the network is known as a feed-forward network.

The neurons in a BMLP usually have non-linear output activation (that is a non-linear transfer function). This enables a BMLP to perform arbitrary mappings which could not be achieved by the original single-layer perceptron.

As its name implies, a BMLP is trained to carry out a particular mapping by applying the back-propagation supervised learning algorithm (**1, 2**). Errors, or differences between the actual output of the network and the desired output corresponding to some training input, are back-propagated from the output layer towards the input layer to determine the necessary adjustments to the strengths (or weights) of the connections between neurons in the network. The adjustments are made by following the error gradient. The aim of the training is to find the set of weights yielding the smallest error.

Training is controlled by a learning rate (η) and momentum constant (α), both in the range 0 to 1. The learning rate affects the amount of weight modification in response to a training input. Large values of η cause network instability and conversely too small a value of η slows the learning process unacceptably. In some cases it might be useful to start with a large η and then reduce it to achieve a gradual convergence to the global minimum. The momentum constant α acts to smooth the weight modifications. In general, a high value of α will speed up the training.

The BMLP is one of the most popular neural networks due to the ease with which it can be imple-

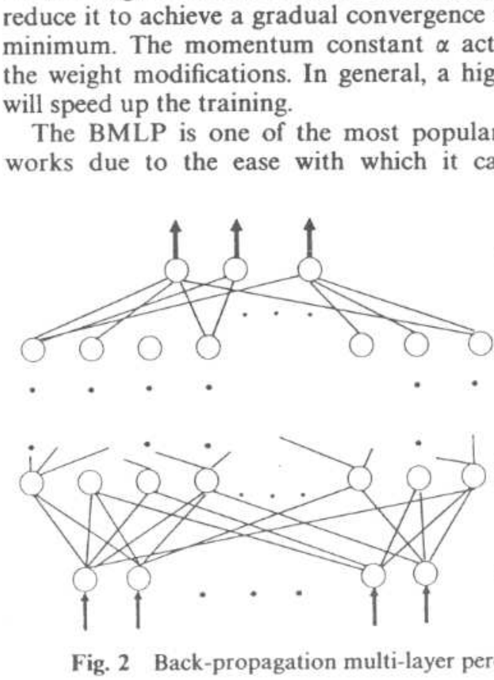

Fig. 2 Back-propagation multi-layer perceptron

mented. However, the application of BMLPs can be problematical and necessitates that attention be paid to the proper scaling of the input patterns, appropriate use of activation functions, prevention of network paralysis (input saturation caused by large weight values), correct choice of the learning rate and momentum constant, and avoidance of local error minima. There is no theoretical basis to guide a user who generally has to resort to trial and error in determining the best way of applying this type of neural network.

For the classification of inner sealing lip perimeters, a three-layer BMLP was used. This consisted of 20 input neurons, 10 hidden neurons and 3 output neurons. The 20 input neurons were to accommodate the 20-dimensional feature vectors representing the geometric features of the perimeters. (Each feature vector is an array of 20 components. As explained further in Section 5.2, these components are tallies for the numbers of occurrences of particular geometric features in the perimeter being inspected.) The 3 output neurons correspond to the 3 classes of perimeters: good, oval and irregular. 'Good' perimeters are approximately circular, 'oval' perimeters have a smooth shape but are not circular; 'irregular' perimeters are neither 'good' nor 'oval'. Note the qualitative definitions of the perimeter classes.

For the classification of seal surface flaws a similar net was used. This had 25 input neurons, 10 hidden neurons and 3 output neurons. The 25 input neurons were to accommodate the 25-dimensional feature vectors characterizing the geometric features of the flaws. (The first 20 components of each feature vector are of the same kind as for the feature vector of a sealing lip perimeter. As detailed in Section 5.2, the additional five components give further explicit geometric information such as the perimeter and area of the flaw and the dimensions of the smallest rectangular box enclosing the flaw.) The 3 output neurons correspond to the 3 types of flaws: veins, circular marks and rough patches. 'Veins' have a thick elongated shape; 'circular marks' are thin and approximately circumferential lines; 'rough patches' are amorphous blobs which could be made up of smaller spots. Again, note the fuzzy definitions of the flaw types.

In both BMLPs, the momentum value was 0.8 and the learning rate was 0.7. All hidden and output neurons had sigmoidal activation functions. (When a neuron has sigmoidal activation, its output/input curve has a shallow slope in the small-input range, followed by a steep rise in the medium-input range and flattening out in the high-input range. The sigmoidal function simulates the thresholding action observed in biological neurons.)

4 LEARNING VECTOR QUANTIZATION NETWORK

The learning vector quantization (LVQ) network is a supervised-learning network developed by Kohonen (3, 4). An LVQ network comprises three layers (see Fig. 3). The first layer is the input layer, the number of neurons in which is equal to the dimension of the input space. The second layer, called the Kohonen layer, is a hidden layer. The neurons in this layer are known as the Kohonen neurons. Their number determines the quantization level of the network, that is the finite number of vectors that an input vector can be mapped into. The third layer is the output layer. The number of neurons here is equal to the number of classes used to differentiate between the various input vectors. Each output neuron is connected to a cluster of neurons in the Kohonen layer. All neuron clusters are disjointed from one another and all have the same number of Kohonen neurons. Thus the number of output classes determines the number of neurons in the Kohonen layer, which, as already mentioned, is the number of vectors that the input space can be quantized into.

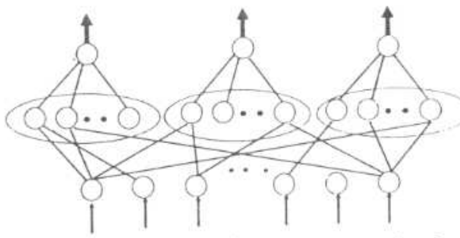

Fig. 3 Learning vector quantization network

In an LVQ network, linear activation (that is a proportional input–output transfer function) is used for the output neurons and a type of variable-threshold activation function is employed for the Kohonen neurons, as explained below. The input neurons do not perform any processing function and simply transmit the input signal directly to the Kohonen neurons. The weights of the connections between the Kohonen neurons and the output neurons are fixed at unity, while those of the input-to-Kohonen neuron connections are modified during training.

The training of an LVQ network starts with the randomizing of all the weights of the connections between the input and Kohonen neurons. The Euclidean distances between the training input vector and the weight vectors of the Kohonen neurons are computed. The neuron with its weight vector closest to the input vector is called the winner. If the winning neuron is within the cluster assigned to the output neuron representing the training vector's class, its weight vector is moved towards the training vector. Conversely, if it is the wrong cluster, its weight vector is moved away from the training vector. The effect of this training method is to make the neurons associated with a class migrate to the correct region dedicated to that class. This method implements a form of 'competitive learning' with the neurons competing against one another to have their weights modified. It can be contrasted with the back-propagation technique described previously where weight modification is carried out by following the direction that most reduces the error between the desired and actual outputs of the neural network.

In the recall mode, the Euclidean distance between the input vector and the weight vector of each Kohonen neuron is computed. The neuron associated with the shortest distance is again taken as the winner and made to cross its 'firing' threshold. This activates the output neuron connected to the cluster containing the winning neuron, thus revealing the class of the input vector.

For the classification of inner sealing lip perimeters, an LVQ network was used, consisting of 20 input neurons, 30 Kohonen neurons (10 for each class of perimeter) and 3 output neurons (1 for each class). For the classification of surface flaws a similar net was employed. It had 25 input neurons, 30 Kohonen

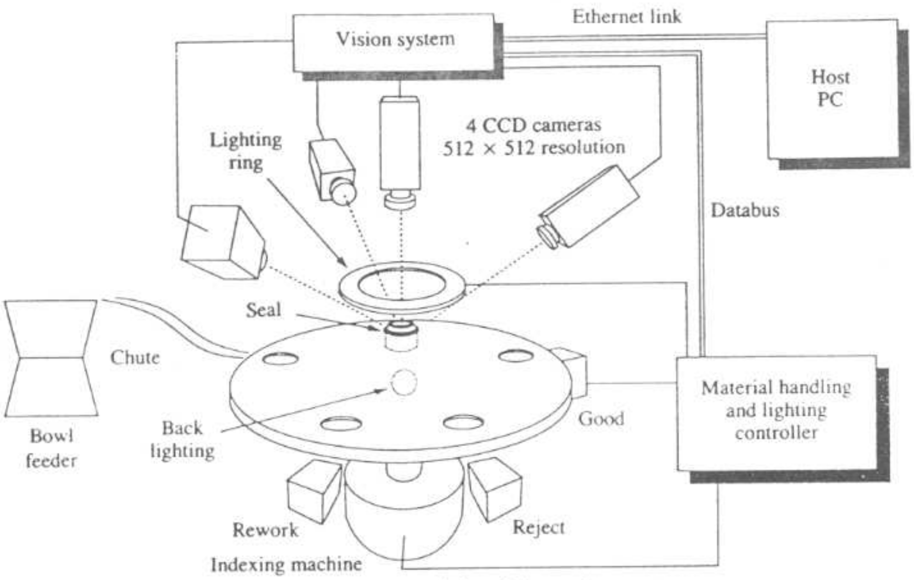

Fig. 4 Automated visual inspection system

neurons (10 for each class of flaws) and 3 output neurons (1 for each class).

5 INSPECTION SYSTEM AND IMAGE PROCESSING

5.1 Inspection system

The experimental hardware of the inspection system is depicted in Fig. 4. There were three modules in the system: the vision module, the lighting module and the materials handling module.

The vision module comprised an image processor connected to a host microcomputer and four CCD cameras each with 512×512 pixels. The image processor was a standard commercially available unit (**18**) for storing images and carrying out low-level processing tasks (filtering, thresholding, etc.). It was also responsible for the real-time implementation of the executable code for the BMLP neural networks described in Section 4. The host microcomputer was committed for neural networks training and peripheral communication functions. One of the CCD cameras was used to capture the top view of the seals which revealed the inner sealing lip perimeter. (The top view of the seal completely filled the 512×512 array so that the sealing lip diameter only corresponded to 256 pixels approximately. A better resolution would have been achieved had the camera been focused just on the sealing lip and its immediate surrounding. However, this was not done as the camera also had to inspect other parts of the top of the seal in the same operation.) The other CCD cameras were for viewing the sealing lip surface. They were positioned around a pitch circle so as to capture equal segments of the sealing lip image.

The lighting module was in two parts, a diffuse halogen lighting unit mounted beneath the seal being inspected for back-lighting the inner lip and a ring of red light emitting diodes located above the seal and directing the illumination at the sealing lip surface to highlight any blemishes.

The materials handling module comprised a bowl feeder and an indexing machine. The bowl feeder delivered 60 seals per minute via a gravity chute to the indexing machine. The latter had five stations, the inlet station fed by the bowl feeder, the inspection station located below the top-view CCD camera and three outlet stations, one for good seals and two for rejects and for parts to be reworked.

5.2 Image segmentation and feature signature extraction

The feature vectors required as inputs to the neural networks were obtained as follows. The grey-level camera images were binarized using a thresholding technique based upon modal analysis of the grey-level histogram (**19**). Isolated noise pixels were removed from the binarized images. A blob labelling operation for connectivity analysis was undertaken to isolate the different objects in the images. A Laplacian operator was applied to these objects to obtain their 1 pixel wide outline. The outline image of each object was separately fed to a feature-extraction n-tuple network (**8**) which scans it to construct the feature vectors for the original objects. This network was implemented using the hardware look-up truth tables of the vision system. The complete preprocessing operation took approximately 350 ms. As mentioned earlier, the feature vector for an inner sealing lip perimeter had 20 components. The value of each component was equal to the number of times a particular geometric feature was present in the outline image being analysed. The 20 different geometric features looked for are illustrated in Fig. 5. Only exact matches between these standard features and segments of the outline were counted.

The feature vector for a surface flaw had five additional components. These corresponded to other geometric attributes, namely the area of the flaw, its perimeter, the length and width of the smallest enclosing rectangular box and the area of the latter. The aim of providing both the linear dimensions and the area of the rectangular enclosing box was to facilitate the training of the neural networks.

Examples of different feature vectors for different artefacts can be found in Figs 6 and 7. (Note that only the first 20 components are shown for the feature vectors of Fig. 7.)

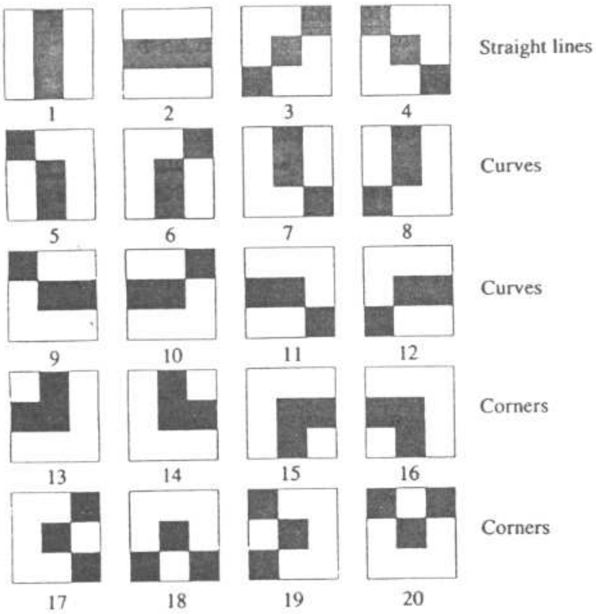

Fig. 5 Set of geometric features

6 RESULTS

Two training sets, each containing 180 feature vectors, were used to teach the neural networks the sealing lip perimeter classification and the surface flaw recognition tasks respectively. For the sealing lip perimeter classification task, the BMLP took 150 000 training iterations to reduce its global output error to below 0.01. The training time was approximately 13 minutes. The LVQ network required 3600 iterations and 8 minutes to reach its minimum global output error of 0.02. For the surface flaw recognition task, the BMLP needed 130 000 training iterations to achieve a global output error less than 0.001. The training time was almost 8 minutes. The LVQ network took 2800 iterations and around 5 minutes to reach a minimum global output error of 0.001.

Two test sets, each with 100 feature vectors, were employed to evaluate the generalization capability of the neural networks in the sealing lip classification and surface flaw recognition tasks respectively. None of the test feature vectors were in the training sets. The BMLP was consistently better than the LVQ network at generalizing, that is dealing with new feature vectors. It was

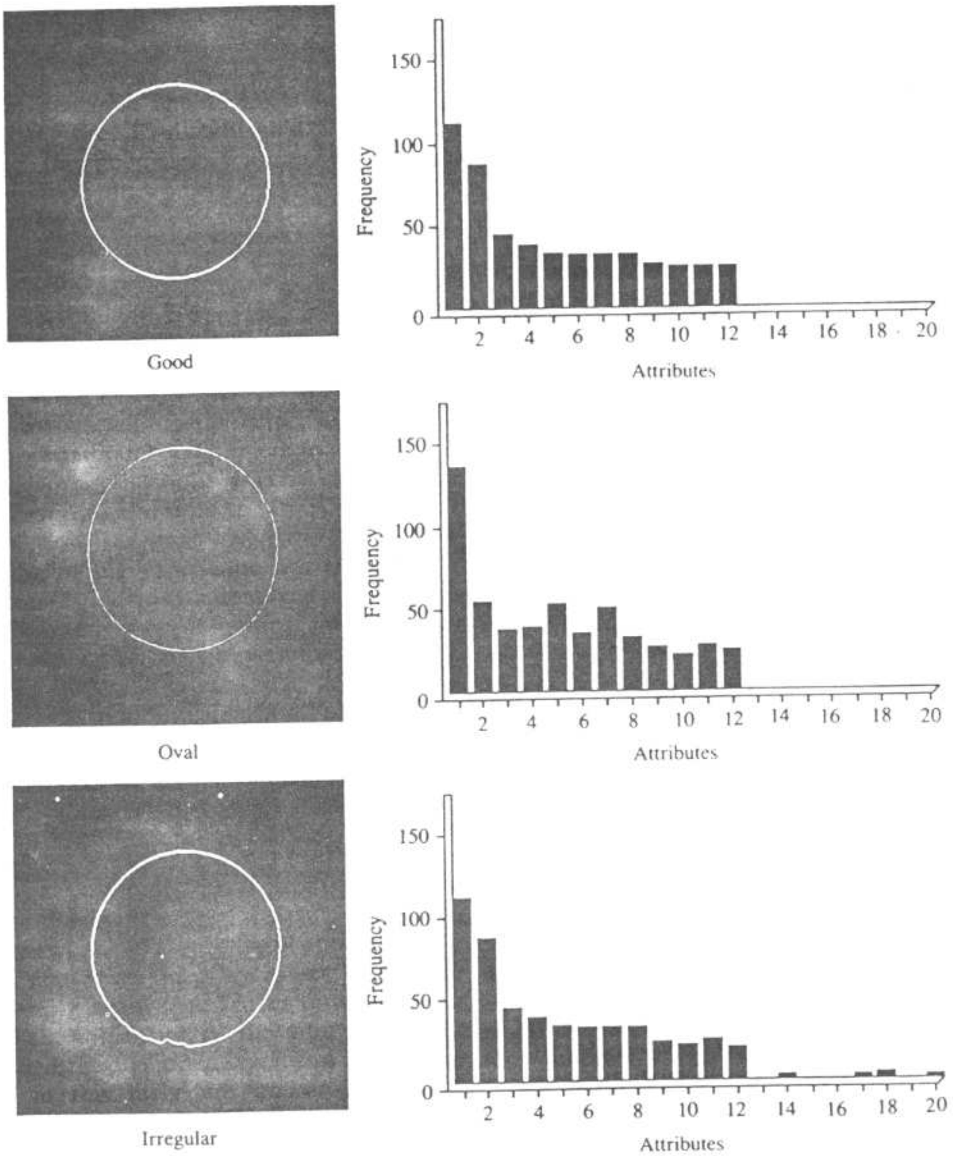

Fig. 6 Perimeters and feature vectors

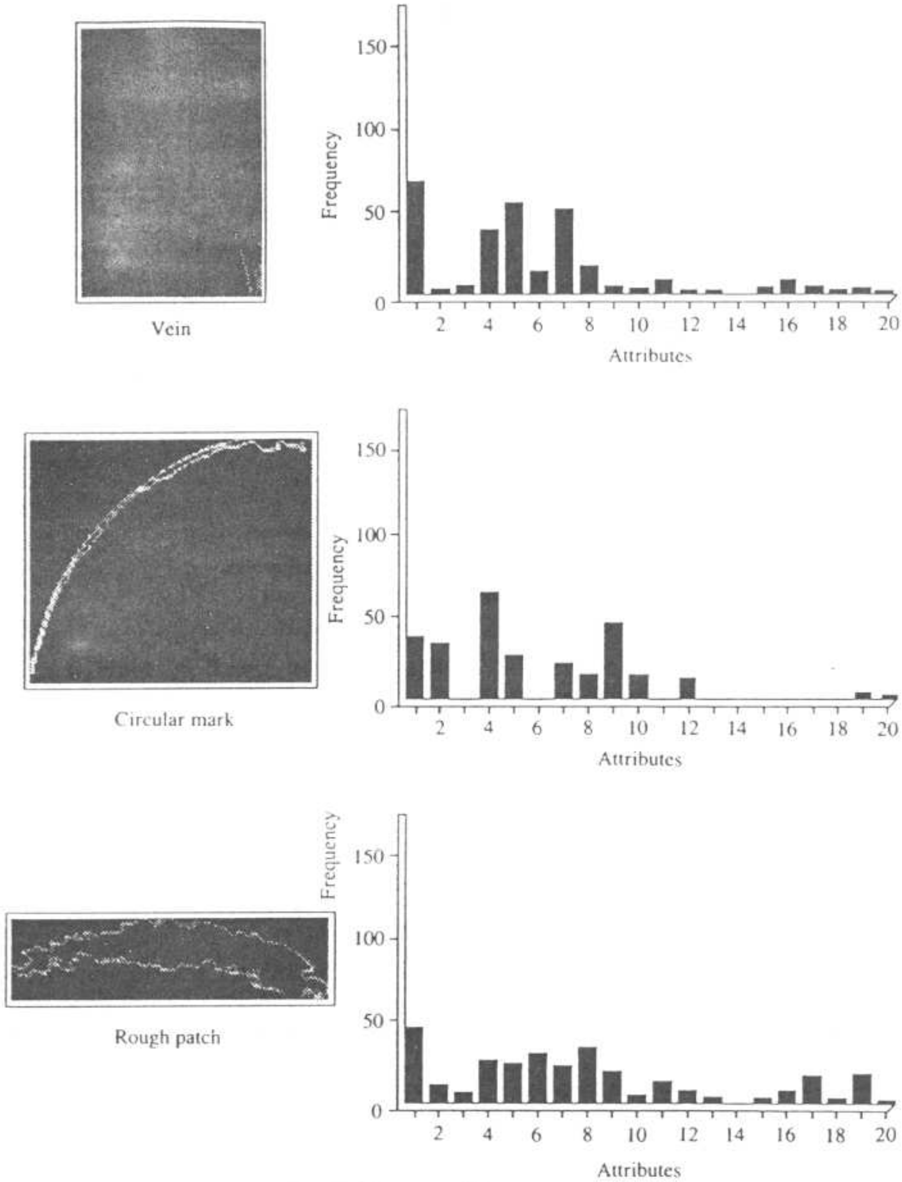

Fig. 7 Flaws and feature vectors

able correctly to classify 83 feature vectors for different categories of inner sealing lip perimeters. The LVQ network only succeeded with 74 feature vectors. Note that the majority of the mistakes were in confusing a perimeter containing small flaws (1 or 2 pixels in size) with a good perimeter. In the surface flaw recognition task, the BMLP accurately recognized 93 test feature vectors and the LVQ network 85. The poorer performance of the LVQ network in both tasks could be attributed to its use of the Euclidean distance metric which is not appropriate for classification problems involving non-linear decision boundaries. However, even the use of the LVQ network would significantly improve on the existing situation, which only involved sampling inspection, as already mentioned.

The time required for both networks to reach a decision, or recall time, was approximately 4 ms in the inner sealing lip perimeter classification and surface flaw recognition tasks. Although the LVQ network is slightly larger than the BMLP in its hidden layer, it took a similar amount of time to process a feature vector because its processing is simpler, involving only the evaluation of linear activation functions rather than the sigmoidal functions of the BMLP modules.

7 CONCLUSION

Automated visual inspection is a computer-intensive task involving both image processing and pattern recognition. This paper has focused on the pattern recognition part and has described two types of neural network pattern classifiers. The use of neural networks in this kind of application has achieved three main benefits: high speed of operation, ease of implementation and high reliability. In this latter respect, the back-propagation multi-layer perceptron has proved to be superior to the learning vector quantization network.

ACKNOWLEDGEMENTS

The authors wold like to thank the ACME Directorate of the SERC, Seal Technology Systems Limited, Performance Vision Limited, the British Council and the Bolivian Science Academy for supporting this work.

They also thank the Editor and referees for their helpful comments on the paper.

REFERENCES

1 **Werbos, P. J.** Beyond regression: new tools for prediction and analysis in the behavioural sciences. PhD thesis, Harvard University, Cambridge, Mass., 1974.

2 **Rumelhart, D. E., Hinton, G. E.** and **Williams, R. J.** Learning internal representation by error propagation. In *Parallel distributed processing: explorations in the microstructures of cognition* (Eds D. E. Rumelhart and J. L. McClelland), Vol. I, 1986, pp. 318–362 (MIT Press, Cambridge, Mass.).

3 **Kohonen, T.** *Self-organization and associative memory*, 1984 (Springer Verlag, Heildelberg).

4 **Kohonen, T.** An introduction to neural computing. *Neural Networks*, 1988, **1**, 3–16.

5 **Tenorio, M. F.** and **Hughes, C. S.** Real time noisy image segmentation using an artificial neural network model. Proceedings of the IEEE First International Conference on *Neural networks*, San Diego, Calif., 1987, Vol. 4, pp. 357–363.

6 **Eichmann, G., Jankowski, M.** and **Stojancic, M.** Shape description with an associative memory. *SPIE Hybrid Image Processing*, 1986, **638**, 76–82.

7 **Yang, H.** and **Guest, C. C.** Performance of backpropagation for rotation invariant pattern recognition. Proceedings of the IEEE First International Conference on *Neural networks*, San Diego, Calif., 1987, Vol. 4, pp. 365–370.

8 **Pham, D. T.** and **Bayro-Corrochano, E. J.** Neural computing for noise filtering, edge detection, and signature extraction. In *Journal systems engineering*, Vol. 2, 1992, pp. 111–122 (Springer Verlag, London).

9 **Lansner, A.** and **Ekeberg, O.** Reliability and speed of recall in an associative network. *IEEE Trans. Pattern Anal. Machine Intell.*, July 1985, **PAMI-7**(4), 490–498.

10 **Carpenter, G. A.** Neural network models for pattern recognition and associative memory. *Neural networks*, 1989, **2**(4), 243–257.

11 **Aleksander, I., Thomas, W. V.** and **Bowden, P. A.** WISARD—a radical step forward in image recognition. *Sensor Review*, July 1984, pp. 120–124.

12 **Glover, D. E.** A hybrid optical Fourier/electronic neurocomputer machine vision inspection system. ROBOTS 12 in VISION'88 Conference Proceedings, Detroit, Mich., 5–9 June 1988, SME, Vol. 2, pp. 8-77–8-103.

13 **Hecht-Nielsen, R.** Counterpropagation networks. *Applied Optics*, 1987, **26**, 4979–4984.

14 **Duda, R. O.** and **Hart, P. E.** *Pattern classification and scene analysis*, 1973 (John Wiley, New York).

15 **Gouin, P. R.** Performing a part inspection task with a two-staged neural network learning system. Electronic Imaging 88, International Exposition and Conference on *Advance printing of paper-summaries*, Boston, Mass., October 1988, Vol. 1, pp. 60–65.

16 **Beck, H., McDonald, D.** and **Brzakovic, D.** A self-training visual inspection system with a neural network classifier. IEEE International Conference on *Neural networks*, Washington, 1989, Vol. 1, pp. I-307–I-311.

17 **McCulloch, W. S.** and **Pitts, W.** A logical calculus of the ideas imminent in nervous activity. *Bull. Math. Biophys.*, 1943, **9**, 127–147.

18 Performance Imaging Inc. *Image analysis library manual*, Revision 2.1, January 1992 (Performance Vision Limited, Solihull, West Midlands).

19 **Rosenfeld, A.** and **Kak, A. C.** *Digital picture processing*, 2nd edition, 1982, Vol. 2 (Academic Press, New York).