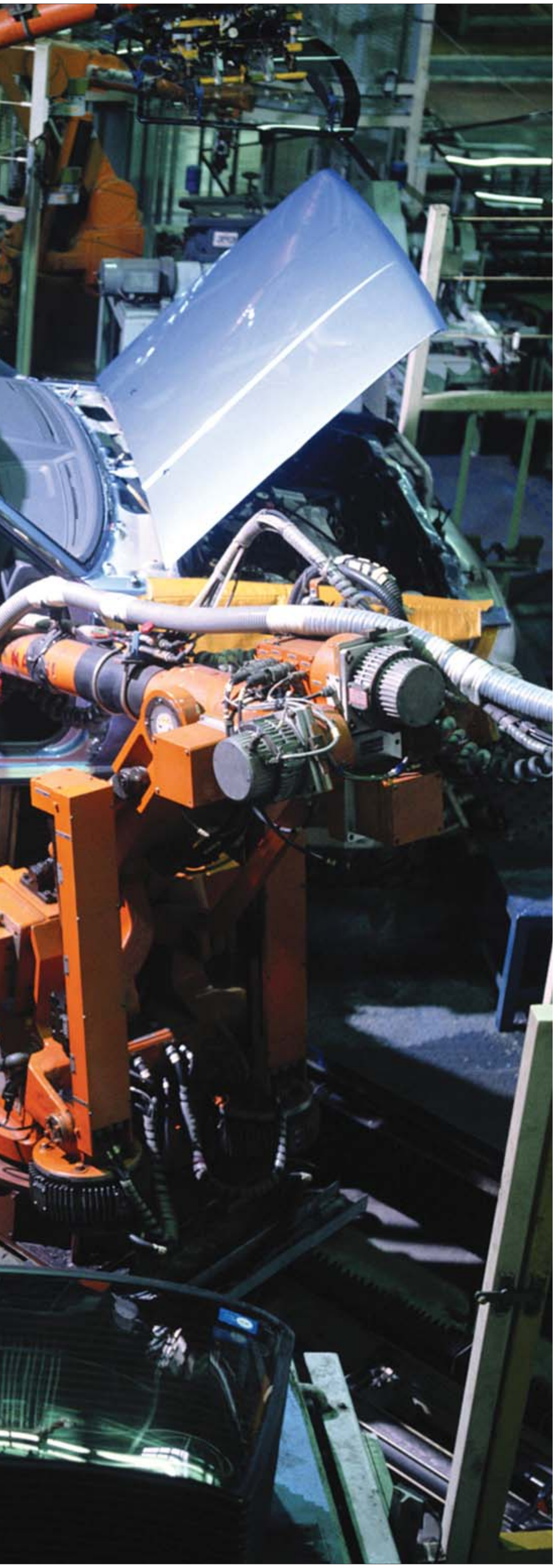
Bad experiences from the past – where changes in ambient light adversely affected system performance – have made engineers sceptical about adopting machine vision into their production processes. The latest generation of cameras and system software removes many of these issues.

By Earl Yardley

Next generation machine vision

Coping with changes in light and surface quality



A large car manufacturer produced seven different engine components down one of its modular production lines, and to avoid tool damage it was necessary to confirm the identification and position of the piston rods prior to entering the next process. The solution was an image processing system that would reliably identify the position and type of differing piston rods. However, strong surface fluctuations from rust and debris, plus varying degrees of reflection, made automated visual identification very difficult (see examples in figs 1-3).

The piston rods are traversed to the station via a fixture system. To capture the image of the part, the rod is briefly held by two pneumatic cylinders. To determine the correct type and position a number of essential inspections must take place.

A piston rod always contains a large and small 'bore', in direct relation to each other. At the end with the small bore, a 'nose' feature is present. This feature is only located on one side of the rod and is used by operators and automated machinery to determine the true orientation of the part. Using this feature it is possible to ensure that the part is always transferred downstream to the next process with the small bore end leading and the nose face down. Any deviation from this is quickly communicated to the control system and the rod re-positioned accordingly.

None of this would place too much of a strain on any modern machine vision system, however the situation is different when surface texture, reflections and rust are present on the part—all of which are very common problems. The variations range from a matt, rusty finish through to a shiny, silver finish making the machine vision task difficult.

THE NEURAL NETWORK SOLUTION

Optimised diffused LED lighting was used to give overall clarity to the image and offered the best overall solution given the changing surface conditions. The latest Firewire camera technology combined with NeuroCheck software offers technological advantages in changing surface conditions. Firewire allows two-way communications between the camera and the image processor, so the camera set-up can be dynamically changed depending on the brightness and reflection of the piston rod.

This is a step forward compared to some older vision system solutions available; and crucial for this application. The decision to dynamically change the camera set-up is automatically controlled and triggered by the software as part of the inspection process. Frame capture takes about 40 milliseconds and, 90% of the time, the image is of sufficient quality and the processing can continue. However in 10% of the images the quality is poor and the software sends a signal back to the camera with appropriate adjustments so that the frame can be captured again. The software is capable of scrutinising a frame and deciding whether or not it should be retaken in about 100 milliseconds. Although it is rare, the system may recapture as many as five or six frames before it is satisfied with the results.

Upon acquisition of an image, the software assigns appropriate pre-processing algorithms based on the quality of the captured image. In this initial step it uses pre-defined →

Variations in the Piston Rod surface

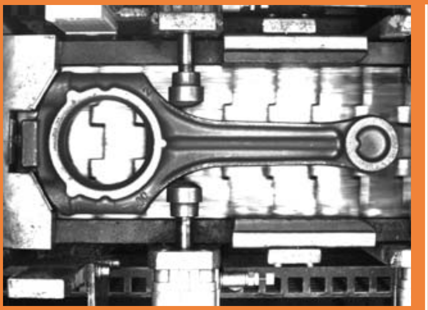
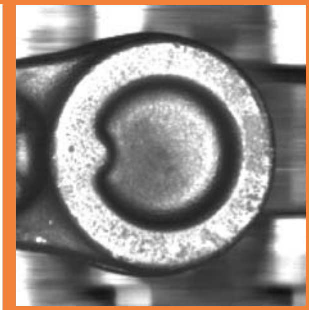


Fig 1: This is a good image of the piston rod, with optimal contrast conditions between bright and dark zones



This close-up shows optimal contrast conditions. About 90% of the time a good image is acquired on the first frame.

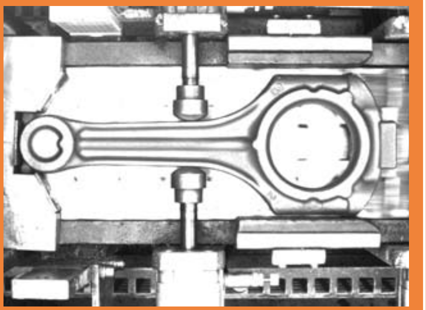
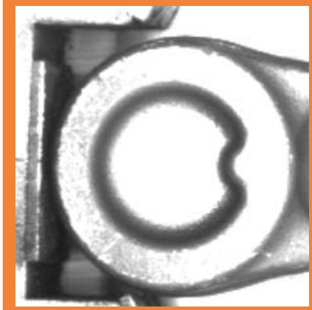


Fig 2: With a highly reflective, shiny piston rod it is difficult to separate the background from the part



In this close-up view of the nose, the grey levels are near the same level, which make identification difficult. In this case the software would send a signal to the camera to adjust exposure and re-capture the frame.

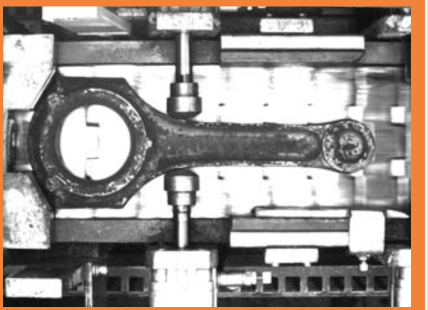
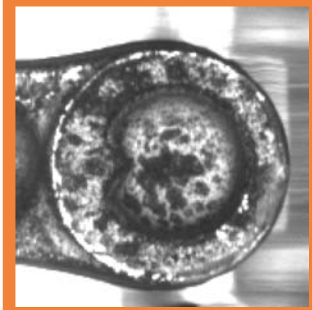


Fig 3: The rusty piston rod and nose are hardly recognisable. In this case, a well-trained neural network will successfully evaluate the image and process the available information.



In this close-up of the nose, surface defects diminish any information for separating the nose from the background. Adjusting the camera for a different exposure will not improve the image.

templates to increase detection certainty. The 'template matching' stage is simply a way to find the area to be processed.

The next stage utilises neural network classification for the reliable detection of the nose and other features. This 'evaluation' stage is also called 'classification'. It is where the neural network classifier is run to check the part, the most important step in the process.

Neural networks enhance the solution with the capability of cognitive intelligence. Contrary to template matching, the neural network has the ability to automatically train itself against the presented images, but does not utilise specific differences for individual samples.

Development of the Mach

During the development of the system the versatility of machine vision software is important because changes invariably require reconfiguration. The solution developed in the pilot tests had to be changed twice, completely altering the mounting and production conditions but the core of the check routine was easily adapted to the changing environments. This meant the time between the development stage and final production was greatly reduced.

A check routine with a simple structure was written but had the full flexibility to cope with any changes to either the communication with the programmable controller (PLC) or in the components themselves. The solution consists of two separate sections. The first, the so-called "start actions" are necessary only for communication with the PLC and for capture of images based on changing lighting and surface conditions. This is where the various digital handshaking signals are set. The individual functions are labelled with user-defined names to give the operators of the system a better understanding of the check routine.

The second section is where the actual image processing takes place. The check routine executes in the following way:

Delay execution – This is where a delay is implemented. The system waits for slightly more than a video frame so that the parts can come completely to rest. This prevents any movement within the image. This could also be achieved by shuttering if required.

Capture image/transfer image – The image is then captured and transferred into NeuroCheck . The size of the region to be transferred can be adjusted to suit the application.

Thus it offers an advantage where surface quality is poor and changing. This information is enough to develop a system with the necessary intelligence to guarantee reliable recognition.

According to statements from the automotive company the system runs 100% reliably. In relation to the installed system the down-times from subsequent processes could be decreased to zero due to the correct identification that the vision system provides; and as a result of this productivity has increased accordingly. ■

Earl Yardley, Industrial Vision Systems Ltd; www.industrialvision.co.uk

ine Vision Solution

Define regions of interest – For each of the possible encoding positions, an individual work area is defined for further processing. In the course of the development cycle, it was found that the positioning of the parts was sufficiently precise. Therefore, an additional position correction on the basis of the global position of the part was not needed. The areas are coloured differently in order to indicate that each area is treated individually in the further course of the check routine.

Template matching (model comparison) – For each of the defined regions a template is produced (as part of the set up) and saved. Multiple templates can be saved to increase the detection certainty, although for most applications no more than 10 are usually needed. These templates are then used to compare each new part in turn. Brightness compensation and sub pixel correlation can also be used to enhance the matching process as well as assigning different classes to the individual regions. In this particular application the customer was achieving about 90% accuracy.

Evaluate classes – In this function the classes created in template matching are again used to check that the right match has been found. The specific classes needed can also be verified in the correct sequence to improve the overall accuracy. Due to the group function being checked in the 'Define ROI' each region has its own class evaluated. Each area is then verified for positional movement in x and y, as well as additional pivotal and angular movement.

Count regions of interest, work areas count – Here each region is then counted and verified against the correct quantity just to clarify that the check routine has been completed correctly.

The structure and training of neural networks for machine vision applications

Neural networks try to exploit some of the principles of information processing in the brain. The information processing power of the brain stems from the large number of neurons and the even larger number of interconnections between these basically rather simple processors. In analogy to this, an artificial neural network is constructed from a large number of simple units connected by many weighted links. This architecture allows the realisation of practically arbitrary transfer functions, i.e. relationships between input signals and output signals. The transfer function actually created depends on the weights of the internal connections. One of the main advantages of neural networks is that it is not necessary to construct this function explicitly. Training algorithms enable neural networks to derive the weights needed to create the desired relationship between input and output from a set of training patterns.

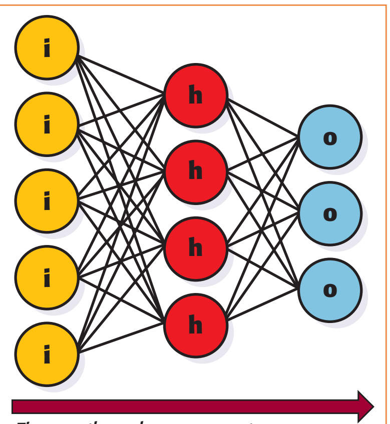


Fig a: a three-layer perceptron

Machine vision software applications such as NeuroCheck use networks of the multilayer perceptron type, a very widely and successfully applied type of network. They consist of three layers of processing units. The sole function of the first layer is to receive the input signals and transfer them to the second layer. The second layer is also called 'hidden layer', because it is not directly visible to the user. It does the actual processing of the signals. The internal representation of the pattern created by this processing is translated by the final layer into an output signal, directly encoding the class membership of the input pattern. Neural networks of this type can realise arbitrarily complex relationships →

between feature values and class designations. The relationship actually incorporated in the network depends on the weights of the connections between the layers and can be derived by special training algorithms from a set of training patterns. Fig a shows a simple network with five inputs, four hidden units and three outputs. Typical networks for digit classification have between 100 and 300 inputs, 10 to 50 hidden units and 10 output units (one for each digit).

NEURAL NETWORK TRAINING

A training pattern consists of an input signal, i.e. a collection of feature values, and the correct class information for the object described by the feature values. The training patterns are repeatedly processed by the network. In case of an error, i.e. a deviation between the actual network output and the correct class information, the internal weights of the network are changed in such a way that the deviation is reduced, until after a number of passes through the training set all training patterns are recognised correctly.

CONSTRUCTING A NEURAL NETWORK APPLICATION

Fig b illustrates the process of building a neural network application.

The first step is a precise specification of the classification task, usually defining the required classes, such as 10 classes for digit recognition, 26 for a complete alphabet, good-bad for a basic quality assessment.

The next step is to choose the features describing the objects. The features have to represent properties of the objects essential for distinguishing the classes. Apart from this requirement the feature selection largely depends on the specific problem and may have to be revised, if it turns out that the objects cannot be recognised reliably using the selected features.

In the next step training data have to be generated, i.e. a collection of objects described by the selected feature values has to be stored. These objects then have to be assigned to the correct classes by hand.

In the NeuroCheck software, network configuration restricts itself to setting the number of hidden units, because input and output configuration are already given by the problem specification. Afterwards the network can be trained. If a network of the defined size proves unable to learn the classification task, the network size may have to be changed.

Finally the classifier has to be tested, preferably with a set of pattern data not used for training. A possible reason for an unsatisfactory recognition rate in the test is an inappropriate selection of features, which should be revised in such a case.

FireWire Cameras

Representing the latest developments in digital camera technology, the NeuroCheck FireWire series of machine vision cameras allows the controlling software and camera to become one unit. Digital image capturing and transmission permits noise-free, sharp images with high precision in colour and grey level imaging applications.

The IEEE 1394 bus (FireWire) standard simplifies the connection to the point where the camera can be installed in a system with only one mouse click. Data cables and mains power supply are combined in one connector so that additional connections are not necessary. The camera setup parameters can be saved in XML format and therefore can be visualised and managed in a standard browser such as Internet Explorer.

A significant advantage of the FireWire bus is that it has two-way communication capability. This allows for online changes of camera parameters based on intermediate results in the inspection cycle. Adjustments can automatically be made for image size, image capturing time, amplification and so on. The two-way communications coupled with dynamic settings allow for new concepts in image evaluation.

The product line ranges from simple CMOS cameras for standard applications to high-resolution colour cameras and FireWire line-scan cameras. Within an application, all cameras can be operated in combination.

USING CLASSIFIERS

Classification tries to model aspects of human reasoning, therefore it is a complex subject. NeuroCheck makes applying a classifier to a problem as easy as possible, performing many of the tasks necessary for creating and using classifiers automatically, but nevertheless a specific procedure has to be observed and some thought has to be invested in how to make the best use of this technology.

WHEN TO USE A CLASSIFIER

Not every problem needs a classifier. If the distinction between objects of different classes can be made by simply comparing some measurements to certain thresholds, a screening process will be sufficient as it is realised by function Screen ROIs (region of interest). As soon as there are more complex, possibly non-linear relationships between measurements, one might think of using a classifier. The same holds, when not only simple features, but the overall appearance of the object has to be taken into consideration as is the case in character recognition.

OTHER WAYS OF GENERATING CLASS INFORMATION

Standalone classifiers (neural networks in NeuroCheck) are not the only way to generate class information. For example, after an object has been found using template matching NeuroCheck has the information about which template was most similar to the object and thus is able to attach a class to the object, namely the class of the template. The correlation algorithm used by template matching is by nature a linear classifier and therefore it is not as powerful and robust as a neural network.

TYPES OF CLASSIFIERS

The task of a classifier in image processing is to assign an object described by a large set of feature values to one of comparatively few possible classes. Fig c shows the assignment of a digit to an image consisting of 216 pixels. The object description comprises 216 bytes (using one byte per pixel, as is typical for a standard grey level image). The object can belong to one of 10 possible classes. Two bytes are sufficient for this information.

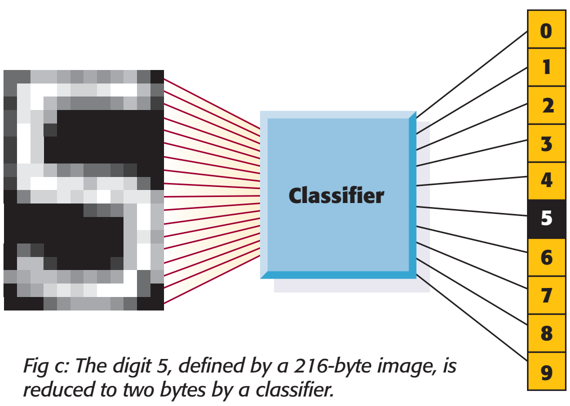


Fig c: The digit 5, defined by a 216-byte image, is reduced to two bytes by a classifier.

THRESHOLD CLASSIFIERS

There are very simple classification problems, which can be solved by checking whether all feature values lie within certain prescribed ranges. An example would be the inspection of the homogeneity of a dark surface. Possible criteria for this application are the maximum size of bright objects, the maximum overall brightness of the surface or a maximum standard deviation of the brightness values, in conjunction with a maximum average brightness. Another such example would be the dimensional accuracy of a geometrically complex object. This type of application can be solved in NeuroCheck using functions like Screen ROIs (for simple geometrical and brightness measurements computed by function Measure ROIs) and Check allowances (for complex geometrical measurements computed by the functions Gauge ROIs and Derive measures) The weakness of such simple thresholding methods is that they are unable to consider complex relationships between the different feature values.

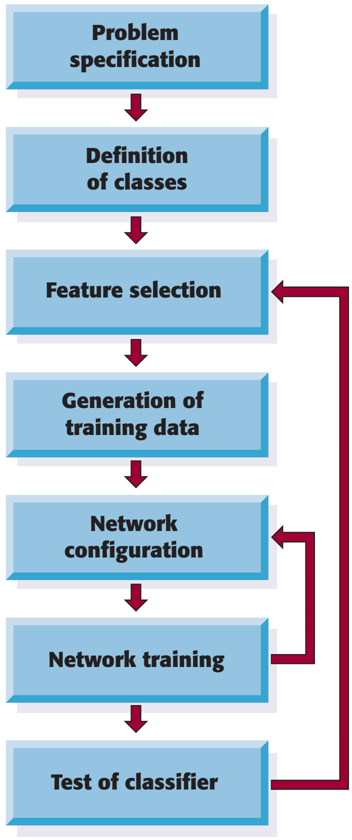


Fig b: building a neural network application

DISTANCE CLASSIFIERS

Distance classifiers store a set of typical patterns, so-called prototypes, for each class. For every object to be classified, they compute the distance of the object to each of the prototypes, usually as the sum of the squared deviations of all feature values. Classifiers of this kind are very easy to construct and can be augmented by additional prototypes at any time, but for difficult classification problems they need a large number of prototypes, requiring a lot of memory and computing time. ■