Defence Science Journal, Vol. 52, No. 3, July 2002, pp. 261-275

Unstructured Object Recognition using Morphological Learning

S. Kar
Bhabha Atomic Research Centre, Mumbai - 400 085

and

S. Chandran
Indian Institute of Technology Bombay, Mumbai - 400 076

ABSTRACT

A technique of object recognition which can detect absence or presence of objects of interest without making explicit use of their underlying geometric structure is deemed suitable for many practical applications. In this work, a method of recognising unstructured objects has been presented, wherein several gray patterns are input as examples to a morphological rule-based learning algorithm. The output of the algorithm are the corresponding gray structuring elements capable of recognising patterns in query images. The learning is carried out offline before recognition of the queries. The technique has been tested to identify fuel pellet surface imperfections. Robustness wrt intensity, orientation, and shape variations of the query patterns is built into the method. Moreover, simplicity of the recognition process leading to reduced computational time makes the method attractive to solve many practical problems.

Keywords: Object recognition, learning algorithm, morphological learning, unstructured object, gray structuring element, geometric structure, surface imperfections

1. INTRODUCTION

It is a challenging problem to develop a recognition method which is applicable to objects irrespective of their underlying structures. Many techniques suggested in the literature seem to work only on classes of images which have some form of geometric structure. Inspection of machine parts by robots is a common scenario where the geometric structure is exploited. Moreover, real-time implementation of these techniques becomes difficult, owing to their computational burden.

A method has been proposed for object recognition which is efficient in execution time and does not depend on the geometric structure of the objects to be recognised. In this method, a set of example images is presented for training during which a set of gray structuring elements (GSEs) is extracted. These structuring elements contain the signatures of the objects learnt. This set of GSEs can identify a query pattern to be one that it has seen before, and hence, recognises[1] the object in the query pattern. By including diverse examples in the example set, flexibility in the recognition process has been incorporated that follows the learning.

The present study is attractive on account of the following viewpoints:

- Firstly, this method is general in the sense that it works irrespective of the geometric structure

Received 28 February 2002

of the objects concerned. By including sufficient number of examples in the training set, the object recognition process can be made reasonably insensitive to object shape changes.

- The recognition method produces clean output by indexing the reference GSE. The method does not produce an intermediate image that needs to be further processed, but reports the reference class as the answer.
- The technique is robust against small rotation as well as local discontinuity, noise or deformation of the query pattern.
- The method of recognition requires few computations, since the training is viewed as pre-processing. The learnt GSEs contain few entries (the gray values at the corresponding points of the learnt input pattern), and for recognition, only simple differencing followed by accumulation of the differences are required. Therefore, the method can perform unstructured object recognition in real-time.

In short, the power of the proposed recognition method is in its capability to identify unstructured objects, its computational simplicity, and handling of the learning operation offline.

2. MOTIVATION

Many objects one encounters in day-to-day life are irregular in structure according to the notion of conventional Euclidean geometry. Considerable variations are noticed in the shape of these objects over the instances of their appearance. Such objects are termed as unstructured. The image of such an object is not divisible, for the purpose of modelling and description, into primitive geometrical entries like lines, points, arcs, and so on at the gross-level of imaging. It is difficult to have a definite quantitative measure of the size and shape parameters of unstructured objects.

The authors' interest was spurred by the problem of identifying surface imperfections in fuel pellets, as seen in Fig 1. The dark irregular longitudinal patch in the figure is a manifestation of a crack in the pellet.

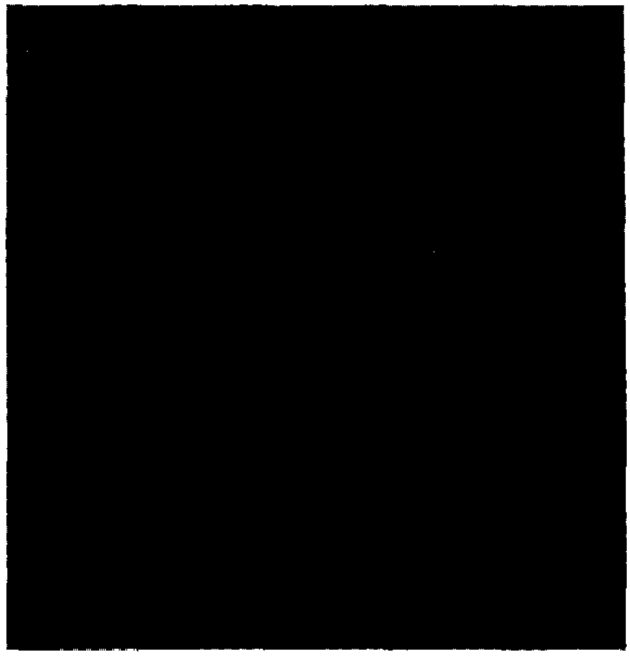

Figure 1. Surface of a fuel pellet showing an irregular longitudinal crack.

Unstructured objects cannot be modelled by CAD-like modeller, and so model-based recognition techniques cannot be applied on these. Most of the other conventional methods of object recognition depend either on finding the contour of the objects by sensing the gray-level transition and then identifying their shape, or on matching the ideal object templates with the real image. The process of finding the object contour in the first category of methods (by operators like the Sobel edge detector) produces additional spurious responses and becomes too much context-dependent for thresholding, to extract the relevant part of the contour. Recent operator like Canny[2] suffer from severe computational burden and do not necessarily provide results. The method of template matching is computationally too expensive for application in real-time situations, and by definition, is difficult to generalise. Trials to detect unstructured shapes using local and global gray-level statistics of the image data showed spurious responses[3-5] (along with the desired response in the output).

Here, a morphological learning-based object recognition technique is presented. This technique works irrespective of the geometric structure of the underlying objects and identifies these fast to be useful in many practical applications.

3. MORPHOLOGICAL LEARNING-BASED OBJECT RECOGNITION METHOD

In the morphological learning-based object recognition method, a set of images is first input as examples during the training phase to a morphological rule-based learning algorithm. The output of the algorithm is a set of signatures of the input patterns in the form of GSE. The GSE contains the gray values of the image at several points. This set of signatures can identify a query pattern, similar to one of the training examples, during the recognition phase. For the purpose of recognition, the set of GSE is applied on the query image, one by one, and a representation of the query is extracted corresponding to each GSE. These representations are used to find a match for the query against the GSEs by computing the gray distance between these. The query pattern, in effect, indexes the reference image in the database. By including diverse examples during training, one allows for flexibility in the shape of the query pattern during the recognition process.

3.1 Framework of Learning

Learning of an object is the process of identifying some of its characteristic features. It is possible to recognise similar objects occurring amidst others using these learnt features. In the training phase, one applies the learning algorithm with an ensemble of example images representing the respective classes that the objects in these images belong to. A learning algorithm has been developed that automatically extracts the characteristic features of the objects being learnt from their images in terms of the GSE defined as follows:

Definition 3.1: GSE is a subset of an image.

Definition 3.2: A point of a GSE is an element of the GSE.

3.1.1 *Morphological & Neural Learning*

Conventional morphological methods are useful for determining the existence or location of patterns with reasonably defined shapes. But if the shapes of patterns are ill-defined, the identification problem becomes more difficult. The automatic method of training GSE described here will allow the training of object patterns (unstructured in general) on the basis of some morphological criteria. This learning algorithm will generate the GSE by selecting pixels based on the object's intensity, which are adequate to distinguish one object from the other within the example set chosen for training. During the training of the example patterns, the learning algorithm derives an efficient representation (GSE) from the given set of examples by way of morphological transformations. The transformation process gets rid of redundant data in the example objects and extracts the salient features to make the representation of the input pattern set compact, and the identification process in the recognition phase more efficient. The most efficient GSE is considered to be the one which consists of a sparse sampling of points along the object under training, as a little localised distortion or miss in the shape of a query object will not affect the recognition performance too adversely.

In a conventional neural network, the neurons in the network learn from examples to derive a pattern representation most conducive to recognition later on. Drawing similarity from learning by an artificial neural network (ANN), where a connection strength is enhanced[6] if it is useful in establishing the identity of a given pattern, in morphological learning method, a point is added to a GSE if it is useful in identifying an example pattern amongst the full set of examples employed for training. The rules in the case of learning of a conventional ANN and the morphological learning are alike. Similar to the recognition capability acquired by the ANN after its learning is over, the morphologically trained GSE has the ability to find a match between an example pattern (for which the GSE is the signature) and similar query patterns, recognising the latter.

3.2 Overview of Training & Recognition

The training of the input objects in the chosen set of example images (each image containing one object representing a class) is performed sequentially, one after another. During the training phase, the GSE corresponding to the object under training is enhanced by a point if it is useful in identifying the pattern under training and has minimal effect in generating false recognition in the background and to other patterns. The point may be intuitively thought of as being one of the two types. (The algorithm itself simply records the gray value at the chosen point of the example into the GSE). A foreground point is added to the GSE when it corresponds to a certain feature point in the pattern being trained; it is helpful in identifying the particular object being learnt amidst the full set of objects employed during the training phase. A background point is added to the GSE when it corresponds to a point lying outside the object, which is useful for the GSE to reject some of the other example input.

Once the reference database consisting of a set of GSE (corresponding to the set of objects that are trained) is generated, the recognition algorithm receives query objects for identification. The set of GSE is applied on a query image sequentially and a representation of the query is extracted using each of these, which is used to find a match between the GSE and the query. A match indicates that the GSE and the query belong to the same class, and the query pattern is said to index the reference image (for which the said GSE is a signature). If the training is properly done, at the most one reference will be indexed from which the query object will be uniquely recognised (Fig. 2).

3.3 Erosion *vs* Gray Distance for Identification

To perform indexing of the reference database with the test object, a method is needed to apply a reference GSE on the test object for finding a match. The same method is to be utilised, as it so turns out, during the generation of the GSE as well. Erosion[7] is the basic pattern detector application mechanism in the case of binary objects. A particular shape in a binary image can be detected by eroding the image with a structuring element slightly smaller than the desired shape.

The authors[8] suggest that the distance from a problem p to a class C can be calculated as the norm of the difference between the two characteristic vectors, $d(p,C)$. Then p belongs to class C if $d(p,C) < \delta$, where δ is some threshold value adjusted depending on the reliability of the characteristic vectors. In the case of recognition of gray-level patterns, one computes the gray distance D (cumulative difference of gray values) of all the GSE points from the corresponding points of the test pattern to find a match between a GSE and a test pattern. If D is within a certain threshold

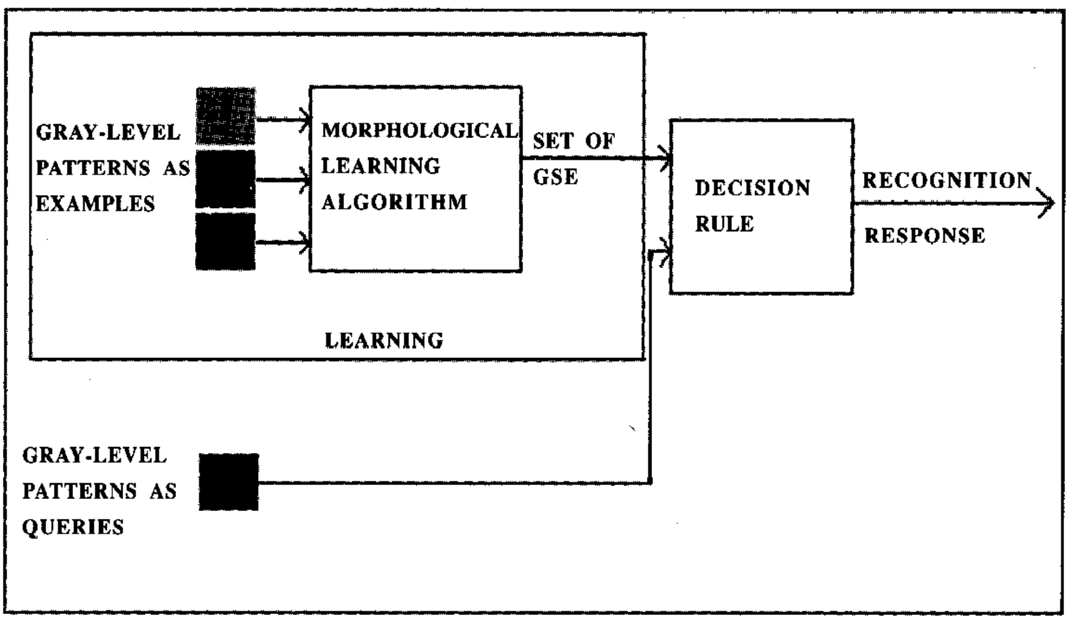

Figure 2. Block diagram of morphological learning-based object recognition method

(τ), an affirmative recognition response is decided. The gray distance between the query and the set of reference GSE are computed for all the references, one by one, to find a probable match. The matching result is output simply in the form of the specification of the object class that the matched GSE represents. The authors say that the query pattern indexes the reference database of objects to produce the recognition output and call such an output to be clean.

3.4 Method of Learning

Suppose in an ANN the connection strength (weight) between the i^{th} input and j^{th} output neuron is w_{ij}. Assume that for a given input, the network does not produce the correct output, signifying that the connection strength w_{ij} has to be changed, i.e., the network has to be trained. The delta rule[9,10] for training weights in neural networks is written as

$$\Delta w_{ij} = \eta (t_{sj} - o_{sj}) i_{si} \quad (1)$$

where Δw_{ij} is the connection strength increment from input neutron i to output neutron j; η is a learning rate constant; t_{sj} is the desired training output value in training cycle number s; i_{si} is the signal value for a particular training input, and o_{sj} is the output computed according to the current set of weights.

Δw_{ij}, called the error, is added to the existing w_{ij}, to get the modified value of connection weight and the procedure of finding out o_{sj} and Δw_{ij} thereafter, is repeated until the error term $(t_{sj} - o_{sj})$ in Eqn (1) reaches some pre-decided small value. Under this choice of the set of connection weights w_{ij}, the neutral network produces correct output for the training input, signifying that its learning is complete. The ANN now becomes capable of recognising correctly a query similar to the input with which the training is carried out.

This rule[11] was adapted for the morphological learning of objects. Let one considers the process of obtaining a GSE to learn an example pattern P_1, in the context of other patterns P_j, $2 \leq j \leq N$. The aim is to come out with a GSE which will recognise P_1, and will not produce any recognition response for all other P_j. (Similar GSEs have to be generated for all other P_j to complete the whole training process). The GSE now corresponds to the neutral network and the elements of the GSE correspond to weights. Similar to the process of updating the connection strengths during the training cycles of a neutral network, the GSE has to be grown during the learning of a pattern by this morphological method. Since the authors' are working with square images, they consider both the pattern examples used for training of the GSE and the GSE themselves (all of the same size) as 2-D square matrices. The desired training output T_j for all the patterns P_j other than P_1 corresponds to a 2-D zero square matrix (of the same size as that of the P_j): The trained GSE for P_1, when applied on all the patterns P_j, $j \neq 1$, should not produce any recognition response. In practice, until the learning of the GSE for P_1 is over, it will produce a non-zero output O_j (considered a 2-D unit square matrix of the same size as that of the P_j) to some of the P_j. These P_j are responding to the GSE. Starting with the initial entry in the GSE, as the GSE grows in iterative cycles, these output O_j corresponding to the input patterns other than P_1 will be reduced to zero matrix, one by one. The image-based delta rule becomes

$$\delta w_j = \eta (T_j - O_j).P_j \quad (2)$$

In Eqn (2), the matrix δw_j corresponds to the error term. The analogy of the learning process of GSE to that of an ANN follows immediately from Eqns (2) and (1). The desired training output T_1 for P_1 is defined to be $n \times I$, where I is the identity matrix of the size same as that of the P_j, and n is the number of patterns (including that under training), responding to the GSE grown till the current cycle of training. Also, $\eta = 1$ was choosen. The incremental errors (δw_j) were accumulated to get $\Delta W = \Sigma \delta w_j$, corresponding to all the example patterns responding to the current GSE. ΔW is the resultant error matrix for the current iteration of training of P_1. Finally, the weight is changed at a single position r of the GSE, that corresponding to the element with maximum magnitude in the error matrix ΔW, by making a fresh entry into the GSE. The value of the entry is the gray value of

P_1 at that position. If there are multiple maxima, a suitable value of r is chosen by a tie-breaking rule.

Once the GSE is updated with the fresh entry in a training cycle, computation of the response O_j, $2 \le j \le N$, followed by evaluation of the error matrix ΔW is repeated in the next cycle of iteration to choose the next entry into the GSE. An example pattern P_j is said to produce a non-zero response O_j to a GSE if the gray distance between these (i.e., P_j and the GSE) is less than or equal to a threshold value τ is equal to a constant integer C multiplied by the number of entries into the GSE. The value of C is chosen based on the range of gay value of the objects in the training set. Once an input ceases to respond to the growing GSE, it is never considered in the following training cycles. This way, the iterations are repeated till all the P_j other than that under training (P_1 under discussion) stops responding to the GSE.

3.4.1 *Sparsity Modification*

The problem with the above algorithm is that two successive elements in the learnt GSE could be arbitrarily close. As already mentioned, two neighbouring elements in the GSE should be placed sufficiently apart to ensure that a small local distortion in the shape of a query object (as compared to the example generating the GSE) does not degrade the recognition performance of the GSE much. Such a GSE is called as sparse. To place any two elements in the GSE sufficiently apart, before selecting a fresh element, the GSE is dilated (using eight connectivity[9]) by a square structuring element of size R (denoted by) and subtracted from the error matrix. This process is indicated in the modified training rule in Eqns (3) and (4) as

$$W_f = GSE \oplus \quad (3)$$

and

$$\Delta W' = \Delta W \pm W_f \quad (4)$$

In the binary operator in Eqn (4), the negative value is chosen for an element in the matrix ΔW if the element is positive, and the positive value is chosen if the element of ΔW is negative. To choose a position r in order to make an entry into the GSE, $\Delta W'$ is considered. In Eqn (3), the amount of dilation[13] R acts as some sort of repulsive force between the two adjacent elements in the learnt GSE. In the implementation, R = 10 was used to ensure that successive entries in the GSE are separated by at least 10 pixel locations.

3.5 Algorithmic Steps for Morphological Learning of GSE

Input: A set of N training images P_j, $1 \le j \le N$, each of size $M \times M$ pixels, and a constant integer C.

Output: A set of N gray structuring elements GSE_j, $1 \le j \le N$, each of size $M \times M$.

Notation

$i,\ k$	Pixel coordinates in the horizontal and vertical directions, respectively
$GSE_j^{i,k}$	$(i,k)^{\text{th}}$ element of GSE_j
$P_j^{i,k}$	$(i,k)^{\text{th}}$ pixel of P_j
$g_j(i,k)$	Gray value of P_j at pixel position (i,k)
$\oplus$	Operation of dilation (using eight connectivity) by a square structuring element of size R

Step 1: $GSE_j^{i,k} = 0$, for $1 \le j \le N$, $1 \le i \le M$, $1 \le k \le M$; $j = 1$ and $R = 10$.

Step 2: Select a pixel $P_j^{i,k}$, $1 \le i \le M$, $1 \le k \le M$. $GSE_j^{i,k} = g_j(i,k)$.

Step 3: $l = 1$.

Step 4: Check whether $\Sigma_{i,k} \mid P_l^{i,k} - GSE_j^{i,k} \mid \le \tau$, $\forall\, i,k \mid GSE_j^{i,k} \neq 0$. $\tau = C$ x current number of entries in GSE_j.

If condition is true, O_l = I, the identity matrix otherwise, $O_l = 0$, the null matrix.

Step 5: $l = l+1$.

If $l \le N$, GOTO *step* 4.

Step 6: If $O_l = 0,\ 1 \le l \ne j \le N$, GOTO *step* 13.

Step 7: Evaluate $\delta w_l = (T_l - O_l).P_p \quad 1 \le l \le N$ where $T_l = 0,\ l \ne j$, and $= n \times I,\ l=j$, n being the number of $O_l \ne 0, \quad 1 \le l \le N$.

Step 8: Compute $\Delta W = \sum_{l=1}^{N} \delta w_l$

Step 9: Evaluate $W_f = GSE_j \oplus$

*Step**10: Evaluate $\Delta W' = \Delta W \pm W_f$

Step 11: Select the element of $\Delta W'$ with the maximum magnitude located at position (m,n)

$$GSE_j^{m,n} = g_j(m,n) \quad .$$

Step 12: GOTO *step* 3

Step 13: $j = j+1$

Step 14: If $j \le N$, GOTO *step* 2.

3.5.1 *Termination of Learning Operation*

Learning of the GSE for each training example continues for several cycles (maximum number of cycles required for the training of one GSE is equal to the number of examples present in the training set). The learning process for one GSE gets over when only one example pattern–that under training–responds to the current GSE. In practice, some extra entries are allowed into the GSE to increase the quality of the response during recognition phase. This implies that the learning process for the GSE is to be continued for few more cycles after growing the GSE to the extent that only the example pattern under training responds to it. If there are N example patterns in the training set, the training of each GSE is continued till it attains $2N$–1 number of elements. This choice of the number of GSE elements is empirical, and is found to produce optimum results in terms of recognition quality as well as accommodability of small amounts of variations of test objects. These extra entries of the GSE will increase the computed gray distance of the query objects belonging to other classes, during recognition, without increasing the gray distance much for an object belonging to the same class as that of the GSE.

3.6 Convergence Complexity of Training Algorithm

It has been assumed that training set comprises N example patterns, each of size MxM pixels. The examples are subjected to training one by one. After each cycle of training for a particular example, using Eqn 4, at least one other example pattern ceases to respond to the growing GSE. Hence, in the worst case, after N cycles of the training operation (including the first cycle which selects the first entry into the GSE, but this GSE need not necessarily reject any of the examples), one will attain a unique GSE responding only to the particular example being learnt. Hence, the training of all the N example patterns will require a maximum of NxN cycles for convergence. In practice, a few extra elements are added to include a total of ($2N$–1) elements in each GSE. Thus, the upper bound for the number of required training cycles for all the examples is Nx($2N$–1). The time complexity of each cycle is of the order of M^2. Hence, the overall complexity of the morphological learning algorithm is $O[N(2N-1)M^2]$.

3.7 Recognition Decision

To recognise a query object, one tries to find a match between the query and one of the example objects. Given a query pattern P_q, one sees its response by computing its gray distance (gd_l) from GSE_l, the l^{th} GSE (there are a set of N such GSE kept in the reference database) using the equation

$$gd_l = \Sigma_{i,j} \mid P_q^{i,j} - GSE_l^{i,j} \mid \tag{5}$$

for $1 \le l \le N$ and $\forall\, i,j$ such that $GSE_l^{i,j}$ has an entry. The subscript (i,j) denotes the position of an entry (pixel) in $GSE_l(P_q)$. If $gd_l \le \tau$, a threshold, P_q is said to index the l^{th} reference, and it is decided that P_q belongs to the l^{th} class.

The value of τ is decided as follows. τ is equal to a constant integer C multiplied by the number of elements in the GSE. The choice of the value of the integer C is dependent on the intensity of the object in the images encountered in a particular context and is usually the same as that used during the training of the GSE.

*In the binary operation of $\pm$ in *step* 10, the negative value is chosen if the element of ΔW is positive, and the positive value is chosen if it is negative.

While computing gd_l by Eqn (5), the size of the query pattern P_q is assumed to be the same as that of the GSE (i.e., same as the size of the training patterns from which the set of GSEs is extracted).

3.7.1 *Accommodation of Variability*

The authors wish to recognise P_q in spite of some variations that might be there in it as compared to the example pattern closest to it. To account for this variation, the basic criterion of computation of gd_l is relaxed as indicated in Eqn (5) above. While computing gd_l one does not consider all the entries of the GSE, but only 75 per cent (this figure is chosen based on the performance of the recognition method on the images used in the experiments) of the entries corresponding to smaller differences of gray values between the GSE and the query image. Relaxing the distance criterion like this effectively helps the recognition process in achieving robustness against deformation and missing parts of real patterns, without degrading the overall recognition performance. In order to evaluate gd_l, then the differences of gray values between the GSE and the query image are computed for all the elements of the GSE, the difference values are sorted and the lower 75 per cent population of the difference values are added up. If $gd_l \leq \tau$, a threshold, P_q indexes the J^{th} reference pattern. Now the value of τ is chosen to be equal to C multiplied by the number of elements of the GSE considered to compute gd_l.

3.7.2 *Algorithmic Steps for Recognition*

Input: A query image Q of size $M \times M$ pixels and a set of N gray structuring elements GSE_j, $1 \leq j \leq N$, each of size $M \times M$, and a constant integer C.

Output: Integer j, specifying the reference object which is indexed by the query Q.

Step 1: $j = 1$

Step 2: Compute

$$D = \Sigma_{i,k} \left| Q^{i,k} - GSE_j^{i,k} \right|, \forall\, i,k \,\middle|\, GSE_j^{i,k} \neq 0$$

Step 3: Check whether $D \leq \tau$, $\tau = C \times$ the number of entries in GSE_j, if yes, output j and GOTO *step* 5.

Step 4: $j = j+1$. GOTO *step* 2, *if* $j \leq N$

Step 5: Stop.

3.8 Recognition amidst Translational, Rotational, Intensity & Contrast Variations

The morphological learning-based object recognition method discussed above has the capability to recognise objects in spite of small variations in their position, orientation, intensity, and contrast in the query image. These capabilities are attributed to the threshold τ permitted in the gray distance between the query image and the reference GSE, computed during the former's recognition. Identification of the query, even under large translational variations, can be achieved by positioning the object at different locations within the image frame, performed by applying successive linear shifts on the image in both the horizontal and the vertical directions (the portion of the image overflowing through one border is folded back to re-appear through the opposite border), and matching the object with the GSE (in the same way explained in Section 3.7) at each of these locations.

Object recognition in the presence of uniform brightness and contrast variations of considerable magnitude is achieved by performing the following computation in addition to that indicated in Section 3.7. During the recognition process, the differences (absolute values) of the gray levels of the GSEs and the corresponding pixels in the query image are tested to see whether the differences corresponding to all the GSEs, even if large, lie within a pre-decided range (i.e., whether the separation between the maximum and the minimum of these differences is less than a pre-decided value). If that is the case, the query image is recognised by the GSE.

4. RESULTS & DISCUSSION

The approach of object recognition based on morphological learning indicated above has been applied to both the binary and the gray-level objects,

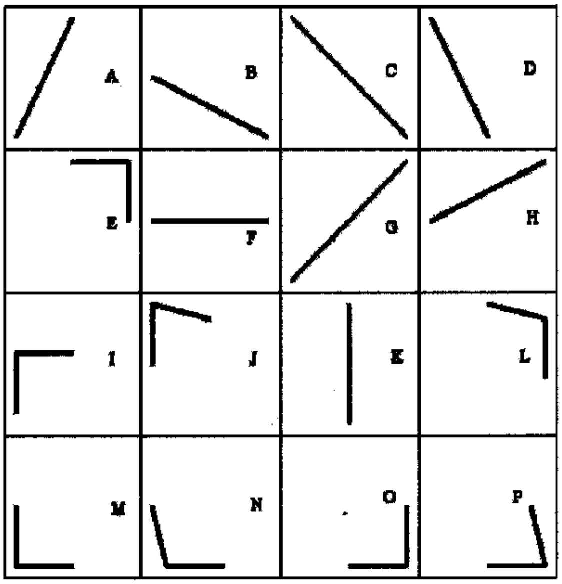

Figure 3. Binary training objects

all of size 121x121 pixels. It is tested on over 200 unstructured objects, few of which are included here for illustration. The effectiveness of this method on binary patterns is discussed.

Figure 3 shows a set of synthetic binary example objects and Fig. 4, the corresponding set of trained structuring elements. When a query similar to any of the objects in Fig. 3 is input, it is recognised correctly. Four such binary queries are depicted in

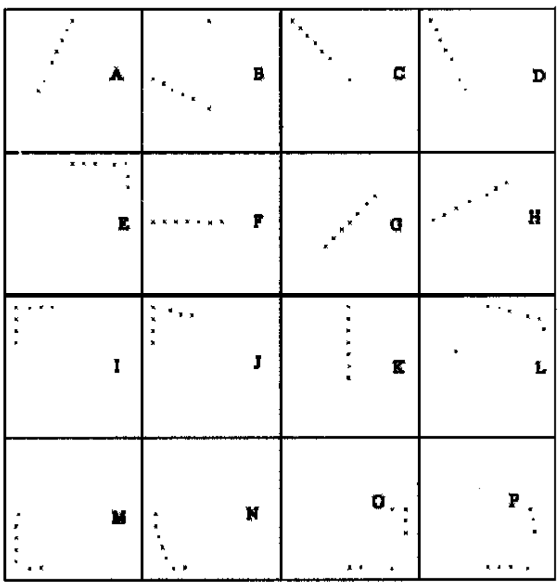

Figure 4. Structuring elements generated by the training algorithm.

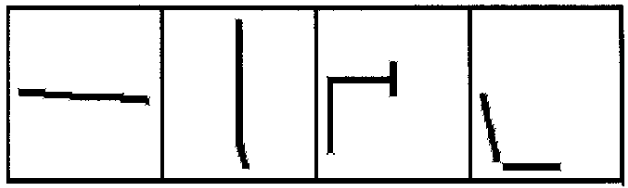

Figure 5. Sample binary queries

Fig. 5. Note that none of these is identical to any pattern in Fig. 3.

Next, illustrations of the learning method on unstructured gray patterns are considered, which show the robustness of the method against shape, orientation, and intensity variations.

Figure 6(a) depicts four unstructured gray-level objects imaged on the circumferential surface of fuel pellets. The objects in this and in the following figures are identified by (i), (ii), (iii), and (iv), wherever mentioned. The GSE for the object (i) of Fig. 6(a), trained in the context of the set of four objects of this figure, is shown in Fig. 6(b) in the form of a 3-D line plot. In Fig. 6(b), the abscissa and the ordinate represent the pixel coordinates in the horizontal and the vertical directions, respectively (the origin of these axes represent the top-left position in the image). The lines in the 3-D plot represent the pixel gray values, implying that the heights of these lines are proportional to the gray values of the

Figure 6 (a). Four unstructured gray-level objects

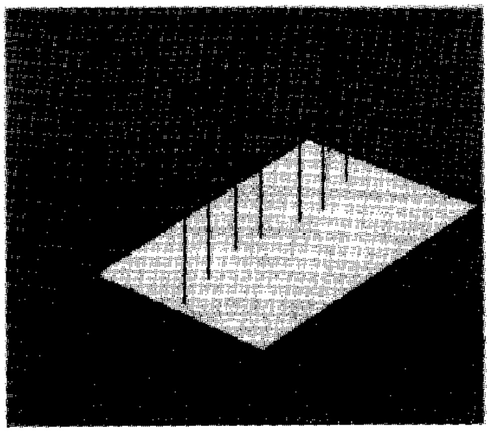

Figure 6 (b). Trained gray-structuring elements for the object (i) in Fig. 6(a).

individual pixels in the image. Figure 6 shows that the morphological learning method extracts pixel intensity values of the trained image mainly from within the object contour (in other cases, the trained GSE may have entries corresponding to the background). The GSE for the other three example objects of Fig. 6(a), as generated by the learning algorithm, are not shown here.

Six sample query object patterns are shown in Fig. 7. These resemble the objects of Fig. 6 with some variations in the form of local misses atthe end as well as at the middle of the patterns. Rotation and deformations wrt the objects of Fig. 6 are also seen. The vertical white line at the middle of the object (v) in Fig. 7 has been added to accentuate the small rotation of the object wrt object (i) in Fig. 6(a).

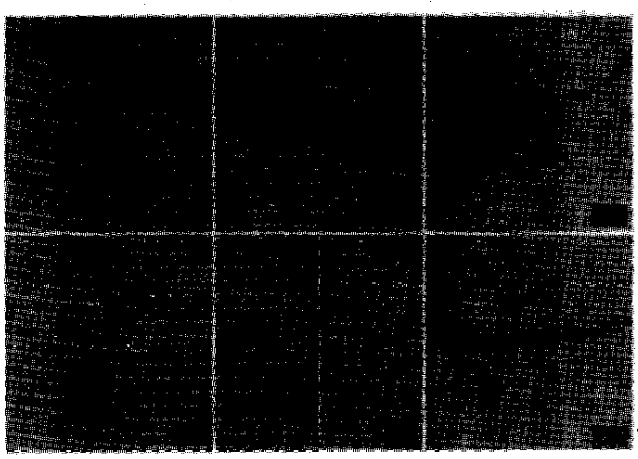

Figure 7. Sample queries

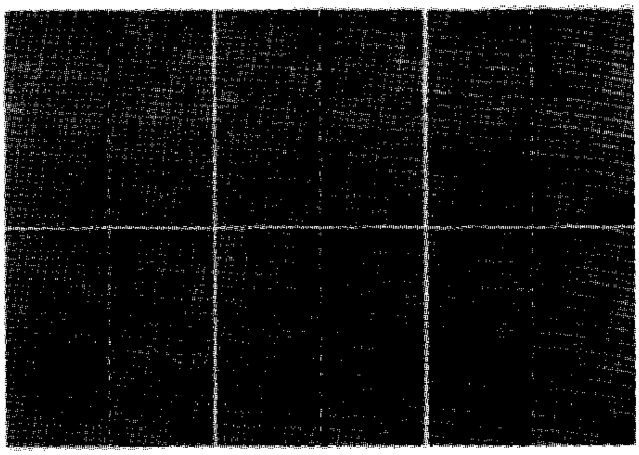

Figure 8. Six query patterns with rotational variations

These query objects are recognised using the learnt GSEs. Specifically, the objects (i), (iv) and (v) in Fig. 7 index and hence, are identified as, object (i) in Fig. 6(a). Similarly, the objects (ii) and (vi) in Fig. 7 are recognised by object (ii) in Fig. 6(a), and object (iii) in Fig. 7 by object (iv) in Fig. 6(a).

4.1 Tolerance to Rotation & Contrast Variations

4.1.1 *Illustration* 1

Figure 8 shows six queries generated from pattern (i) of Fig. 6 by rotations (about the centre of the image) of 5°, 10°, 15°, 17°, 20° and 25° respectively. (The white vertical lines through the middle of these patterns highlight the rotation). All but (vi) in this set of query patterns index the GSE in Fig. 6(b). When the morphological

Figure 9. Additional object used for learning

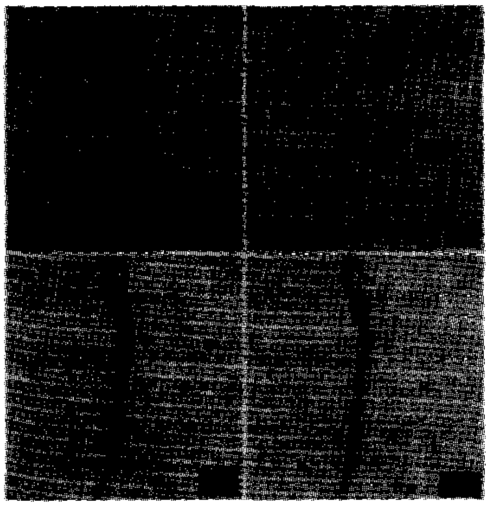

Figure 10. Query patterns with brightness and contrast variations.

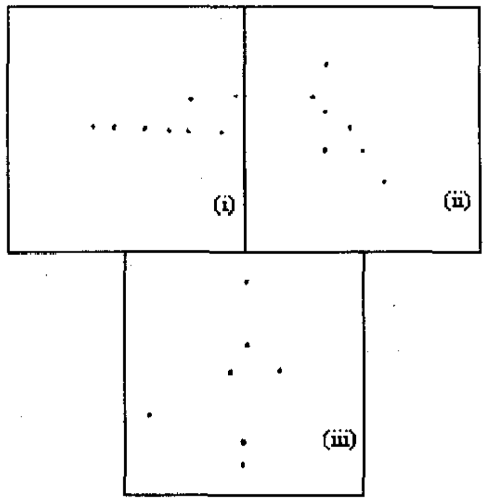

Figure 12. Set of GSEs obtained by learning the objects in Fig. 11.

learning is carried out with only the four examples of Fig. 6, query (vi) in Fig. 8 is not recognised by any of the generated GSEs. When the pattern in Fig. 9 is introduced in the training set, (vi) in Fig. 8 is recognised by the GSE learnt from this (additional) training pattern. Figure 10 contains four query objects which resemble (i) and (iii) in Fig. 6 with contrast and brightness variations. Thus, objects (i) and (iii) of Fig. 10 look similar to object (i) of Fig. 6 with change in contrast, whereas object (iv) of Fig. 10 is similar to object (i) of Fig. 6 with shift in brightness level. The object (ii) of Fig. 10 looks similar to object (iii) of Fig. 6 with change in contrast. Recognition of objects (i), (iii) and (iv) in Fig. 10 is performed by object (i) in Fig. 6(a), whereas, object (ii) in Fig. 10 is identified as object (iii) of Fig. 6(a).

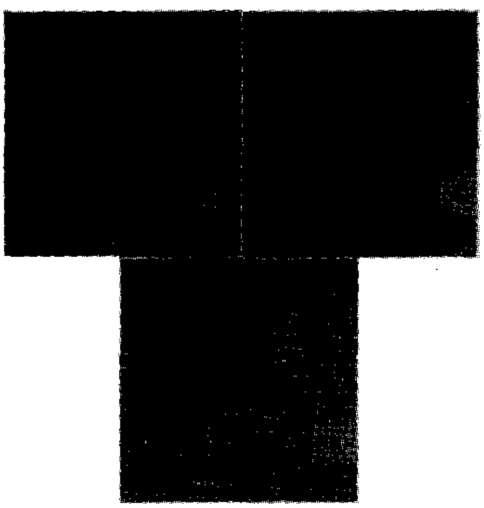

Figure 11. Three gray-level objects constituting an example set.

4.1.2 *Illustration* 2

Figure 11 shows three unstructured objects (again representing cracks in fuel surface images), which constitute a training set. Figure 12 shows the corresponding learnt GSEs with the GSE plotted as dots. Figure 13 depictsfive different crack patterns presented to the recogniser as query objects. The objects (i) and (v) of Fig. 13 index the GSE (i) of Fig. 12, the object (ii) in Fig. 13 indexes the GSE (ii) in Fig. 12, and the objects (iii) and (iv) of Fig. 13 index the GSE (iii) of Fig. 12, indicating correct recognition of all these queries.

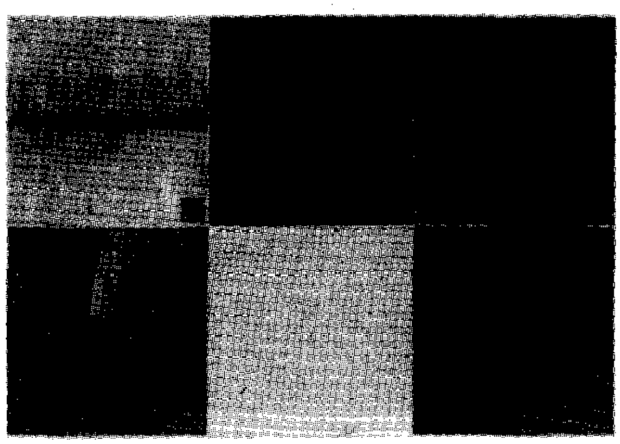

Figure 13. Query objects

4.3 Comparison with Other Methods

One of the main advantages of our morphological learning-based object recognition method is its speed of execution. In this method, the training is carried out offline. The online recognition of objects requires few computations (basically taking the termwise difference of gray values of the query and the GSE — the latter having only few entries, and summing the differences). Comparatively, the commonly used gray-level correlation-based method is computationally much more expensive. As a result, the method of object indexing by morphologically learnt GSE require about 0.2s. with a reference database size of 200 GSEs (which is 0.16 the time required for object recognition by gray-level correlation), tested on pentium-III computer running at 733 MHz.

For comparing the morphological learning-based object recognition method with two statistical methods, Figs 14 and 15 are included which depict the response of these statistical methods when applied on pattern (i) of Fig. 6. Figure 14 shows the noisiness of the result obtained by an adaptive thresholding[4,5] method.

Figure 15 shows the output of the Canny edge detector at a particular choice of the detector's parameters. Both these images require some post-processing to remove the noise, before the object of interest in these can be recognised. Figures 14 and 15 clearly bring out the advantage of the proposed learning-based pattern recogniser over the other two, as the former produces inherently clean output and the technique does not depend on the choice of any parameter.

Figure 14. Output produced by adaptive thresholding technique.

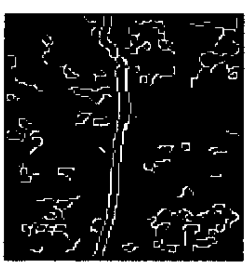

Figure 15. Result produced by Canny edge detector

4.4 Statistics on Successful Recognition

Tables 1 and 2 show the statistics on experimental results of the morphological learning-based recognition technique on binary and gray-level objects, respectively. The statistics justify the claim of 100 per cent success rate of the proposed recognition method, as the only case of false recognition occurring due to rotational variation of the query gray pattern, gets eliminated when the threshold value used on the gray distance, to consider a response as recognition, is increased from 20x to 25x the number of elements in the GSE.

Table 1. Statistics on binary image

Image size (pixels)	Trained examples	Queries	correct response (%)
121x121	16	100	100

Figure 16 is a plot of recognition response magnitude (minimum normalised gray distance) of the set of GSEs learnt by the indicated morphological method against a few sample queries (chosen randomly

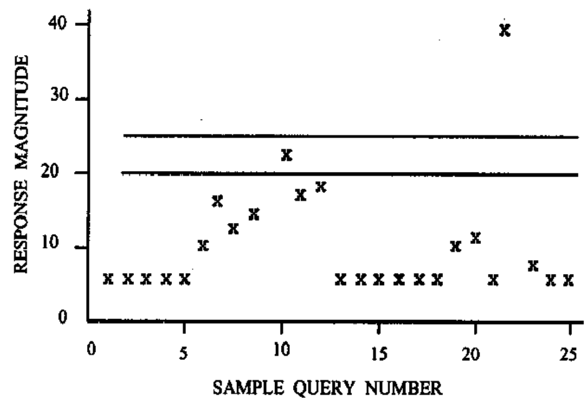

Figure 16. Plot of recognition response magnitude against sample query number.

Table 2. Statistics on gray-level image

Type of variation on the query	Amount of variation (degrees)	Number of queries	Number of false responses	Recognition success rate (%)
Rotation	up to 25	60	3	95
			0, $\tau = 25 \times n$	100
Brightness, contrast	± 40	40	0	100
Local discontinuity, deformation		80	0	100
Unknown pattern or background image		20	0	100

Image size used = 121 x 121 pixels. Image tonal resolution = 256.
Example set size (for training) = 4, 5. No. of queries tested upon = 200.
Threshold, $\tau = 20 \times n$, where n is the number of elements in the GSE.

from the pool of queries used in the experiments). For each of the queries chosen, the gray distance of the query pattern was computed from each of the GSEs, the minimum of such distances selected and divided (normalise) by the number of elements in the GSE. This normalised distance is plotted against an integer designating the query. The majority of response values lie below the threshold level of 20 to indicate correct recognition. The example corresponding to response value lying between 20x and 25x (i.e., false recognition when the threshold is set at 20x) can be considered to give positive recognition when the threshold level is raised to 25x. The query in this case corresponds to an object rotated as compared to an example object employed during training. The only response value situated above the threshold level of 25x (indicating no recognition) corresponds to a query without any object in it (i.e., pure background).

5. CONCLUSION

Similar learning-based object recognition scheme applied to structured noisy binary objects have been reported[14]. The technique presented here extends that approach to unstructured gray-level objects. This method can take care of shape fluctuations, local misses, and discontinuities in the objects, rotations and intensity/contrast variations while recognising the unstructured objects present in real images.

Computer vision is one of the potential application areas of neural networks. It has an important role in industrial automation, such as visual quality control and robotics, as well as in document processing[15], etc. The indicated pattern recognition technique based on learning shows a lot of promise in terms of computer vision applications in industrial inspection tasks involving identification of patterns without having quantifiable Euclidean structure. Once the training of GSEs are over, the recognition phase incorporates few computations. So the method is attractive for practical applications where throughput rate is of prime consideration. The GSE-based recognition technique is robust against noise and deformations that are bound to occur in real patterns.

Lastly, one shortcoming of the proposed method is mentioned. If there are multiple unstructured objects in the query pattern presented for recognition, more than one GSE will produce a recognition response. So the objects in the query will be detected but will not be categorised. As this study does not attempt to solve the general segmentation problem, the objects cannot be recognised, individually. In comparison, gray-level correlation-based recognition technique will fail to identify any of the objects in the query image as the correlation value will be reduced for any of the individual reference objects owing to the simultaneous presence of multiple objects in the query.

REFERENCES

1. Anzai,Y. Pattern recognition and machine learning. Academic Press, 1992.
2. Canny, J.F. A computational approach to edge detection, *IEEE Trans. Pattern Analy. Machine Intelli.*, 1986, **PAMI-8**. pp. 679–98.
3. Bovic, A.C.; Huang, T.S. & Munson, D.C. Jr. Nonparametric tests for edge detection in noise. *Pattern Recognition*, 1986, **19**(3), 209–19.
4. Kar, S.; Rao, Y.V.; Valli, G.; Joshi, D.G. & Chadda, V.K. Vision system for automatic inspection of fuel pellets. Seminar on Supercomputing for Scientific Visualisation, Mumbai, 15-17 February 1994.
5. Kar, S.; Rao, Y.V.; Valli, G.; Joshi, D.G.; Chadda, V.K; & Nigam, R.K. Automated visual inspection of nuclear material. International Conference on Automation, Indore, 12-14 December 1995.
6. Winston, P.H. Learning by training neural nets. *In* Artificial intelligence. Addison-Wesley Pub. Co., 1992.
7. Serra, J. Image analysis and mathematical morphology. Academic Press, 1982.
8. Joshi, A.; Ramakrishnan, N.; Houstin, E.N.; & Rice, J.R. On neurobiological, neuro-fuzzy, machine learning, and statistical pattern recognition techniques. *IEEE Trans. Neural Networks*, 1997, **8**(1), 18-31.
9. Rumelhart, D.E. & McClelland, J.L. Learning internal representations by error propagation. *In* Parallel distributed processing. MIT Press, 1986.
10. Wasserman, P.D. Perceptions. *In* Neural computing theory and practice. Van Nostrand Reinhold, New York, 1989.
11. Kar, S. & Chandran, S. Recognition of unstructured objects by morphological learning. 2nd National Conference on Communications, NCC'96. IIT, Bombay. Allied Publishers Ltd., 16-18 February 1996.
12. Horn, B.K.P. Binary images: Topological properties. *In* Robot Vision. The MIT Press, 1987.
13. Giardina, C.R. & Dougherty, E.R. Topological processing. *In* Morphological methods in image and signal processing. Prentice Hall, New Jersey, 1988.
14. Wilson, S.S. Training structuring elements in morphological networks. Mathematical morphology. *In* Image processing, edited by E.R. Dougherty. Marcel Dekker Inc., New York, USA, 1993.
15. Kohonen, T.; Oja, E.; Simula, O.; Visa, A. & Kangas, J. Engineering applications of the self-organizing map. *Proceedings IEEE*, 1996, **84**(10), 1358–384.

Contributors

Dr S Kar obtained his PhD from the Indian Institute of Technology (IIT) Bombay, Mumbai, in 2001. Presently, he is working as Scientific Officer (Grade F) at the Computer Division of Bhabha Atomic Research Centre, Mumbai. He is the past recipient of the *Homi Bhabha Award* (BARC Training School, Mumbai). His research interests include: Digital image processing and computer vision.