Normalized Vector Codes for Object Recognition using Artificial Neural Networks in the Framework of Picture Description Languages

G.D. Jasmin[1] and E.G. Rajan[2]

[1] Mysore University

Received: 11 April 2013 Accepted: 3 May 2013 Published: 15 May 2013

Abstract

Your Understanding how biological visual systems recognize objects is one of the ultimate goals in computational neuroscience. People are able to recognize different types of objects despite the fact that the objects may vary in view, points, sizes, scale, texture or even when they are translated or rotated. In this paper we focus on syntactic approach for the description of objects as Normalized Vector Codes using which objects are recognized based on their shapes.

Index terms— pattern recognition, formal representation of images, object recognition.

Introduction attern recognition could be formally defined as categorization of input data into identifiable classes via extraction of significant features or attributes of the data from the background of irrelevant detail. A pattern class is a category, determined by some given common attributes. When a set of patterns falling into disjoint classes are available, it is desired to categorize these patterns into their respective classes through the use of some automatic device. It is important to note that learning or training takes place only during the design (or updating) phase of a pattern recognition system. Once acceptable results have been obtained, the system is engaged in the task of actually performing recognition on samples drawn from the environment in which it is expected to operate. The main objective of this research is to investigate and develop a general approach formally from a new theory based on Structural or Syntactic Pattern Recognition.

The problem of pattern recognition is divided into the following sub problems:

1. Image pre processing 2. Object Description 3. Classification

1 D

The pre-processing also known as low level image processing is performed on the input image to improve the quality of the image and to simplify the image for further processing. The pre-processing steps involve noise removal, conversion of gray scale and colour images into binary images and the extraction of contour from the binary image. The object description module takes an input contour image and gives a vector of direction and length called a knowledge vector as output. This plays an important role in the whole process as the knowledge vector gives more information about the objects present in the image which is used to characterize the pattern. We use a syntactic approach for the description of objects. The knowledge vector obtained here gives information about the direction and the length of lines in the contour of objects which can then be given as input to a classifier module. The normalization of this vector plays an important role in the classification process. The vector has to be analyzed and normalized in such a way that it better suits for more variance of objects. The classification module takes the normalized vector as input and identifies them as a member of one of the predefined classes depending on the set of attributes that they hold. The design of a classifier involves the selection of the classifier and the estimate of parameters for the classifier. The feedforward artificial neural networks with back-propagation learning algorithm could be used for classification. Details of such neural networks could be obtained from standard literature.

of noise in the image. The pixel point processing techniques such as binary contrast enhancement, stretching or shrinking and histogram equalization are used to improve the contrast of the image.

The input image to the system may be a colour image, a gray scale image or a binary image. These images have to be converted to binary images for further processing. Segmentation is an important process in our object recognition system as it highlights the objects present in the image from its background. The function im2bw() in matlab is used to convert the images to binary. The graythresh() function computes a global threshold (level) that can be used to convert an intensity image to a binary image with im2bw. Level is a normalized intensity value that lies in the range [0, 1].

The contours are extracted as they give the outlines of shapes of the objects present in the image. The function bwtraceboundary() in matlab is used for finding the contours. The initial point of the object is identified, and the outline tracing algorithm is used. bwtraceboundary() traces the outline pixels of an object in binary image bw. Nonzero pixels belong to an object and zero pixels constitute the background. Figure 2 shows a sample binary image and its contour. The notion of 'Picture Description Languages' refers to the language of digital images defined over a regular array of vertices. Having extracted the contour of an image this process involves running through the contour pixel by pixel and performing some form of calculations using every pixel and its surrounding pixels. This involves finding the length of the contour in every possible direction. The eight possible directions are two horizontal directions, two vertical directions and four diagonal directions. As the lines other than straight lines have more variations in their directions, the length of the descriptor becomes large. It is then reduced to a single occurrence of eight directions and normalized to fixed number of pixels for making it invariant. Following is the knowledge vector of a hexagon.

2 GENERATION OF DIRECTION LENGTH CODE (DLC)

The connectivity between the pixels is an important concept used in the field of object recognition. Two pixels are said to be connected, if they are adjacent in some sense and their intensity levels satisfy a specified criterion of similarity. If pixel p with coordinates (x,y) and pixel q with coordinates (s,t) are pixels of object (image subset) S, a path from p to q is a sequence of distinct pixels with coordinates (x0, y0), (x1, y1),???.., **??**xn, yn) where (x0, y0) = (x,y) and (xn, yn) = (s,t), (xi, yi) is adjacent to (xi-1, yi-1), 1 ? i ? n, and n is the length of the path. If there a path exists from p to q consisting of pixels in S, then p is said to be connected to q. The gap of one or two pixels marks the presence of the next component in the same object or the beginning of the next object which can be then found by analysing the relationship among the components.

With this background, we shall see now as to how to apply Look Ahead Tracer (LAT) algorithm to various image types and extract knowledge from them. Let (x,y) be the initial point. The algorithm continues to find the immediate pixel connected to (x,y) in one of the 8 directions. In this technique, a previously recognized direction during a scan is given the first priority status when compared to the other preferential ones. If R is the identified direction, then the other preferential directions are DR, D, DL, L, UL, U, UR.

The trace continues until there is no pixel connectivity to the current point (x,y), after removing the One can easily say that the above knowledge vector represents a square from the fact that the four directions R,D,L and U has the same length except the direction U that has 1 pixel less than other 3 sides. This is because of the immediate removal of the scanned pixels. The starting pixel at position <104,109> is removed after its scan.

3 III.

4 KNOWLEDGE VECTOR ANALYSIS

The direction codes in the knowledge vector obtained by tracing the contour of an object are R, DR, D, DL, L, UL, U and UR. For better processing, let us categorize these directions into basic directions R,D,L and U and diagonal directions DR,DL,UL and UR. For example, consider the knowledge vector of a square in figure **??**. R100*D100*L100*U99 Figure **??** : The knowledge vector of a square This shows the square consists of four basic directions with all lengths equal. Shapes with all the basic directions but with equal alternate length can be identified as a rectangle. Shapes with four sides consisting of only diagonal directions can be considered as rhombus.

The task of identifying regular shaped objects is simple. Objects in a real time environment may not be regular always. For the effective identification of those objects, the shapes of the objects can be approximated to some regular polygons to which it is closer to. This is possible by predefining the classes of regular shapes with some pre conditions set and approximating the other shapes to one of these shapes by identifying them as a member of the predefined class.

It is important to note that, the diagonal sides present in the shape of an object do not necessarily appear alone as the connection of only diagonal pixels. The occurrence of other pixels depends on the angle at which the diagonal lines appear in the image. Possibly D and R pixels appear more with DR, D and L pixels appear more with DL, U and L pixels appear with UL and U and R pixels appear with UR. For example consider a line with the direction code DR. As shown in fig: if the line bends more towards right the more R pixels appear, in the same way if it bends more towards down the more D pixels present with DR as shown in figure 6. In this respect, when the lineis rotated with a small change in the slope, they can be approximated to one of these shapes. Four basic sides or four diagonal sides with equal alternate lengths.

In the same manner, shapes with more sides can be predefined. When the shape of an object appears it can be then classified by artificial neural networks as a member of one of these classes.

5 THE SHORT COMINGS OF TRADITIONAL PATTERN RECOGNITION SYSTEMS AND THE NEED FOR ADAPTIVE PROCESSING

Objects in the real world from an image or image sequence of the world are detected with the advanced software tools the imaging applications like Automatic Object Recognition (AOR) that uses different object models which are known a-priori. Humans can perform this task of object recognition effortlessly and instantaneously. However, such a task of implementation based on machines is performed using algorithms and has been very difficult in networks for real time imaging problems which are non-linear in nature. In some applications, the patterns/objects to be recognized are so fuzzy that they cannot be modelled with conventional tools. In order to solve these neural network processors can be used as the best tool because of the fact that the processor can build a recognition engine from simple image annotations made by the programmer. It then extracts the characteristics or feature vectors from the annotated objects and sends them to the neural network. Neural networks having adaptive processing elements are capable of performing generalization and can consequently classify situations never seen before by associating them to similar learned situations.

6 a) Artificial Neural Networks

Learning from a set of examples is an important attribute needed for most pattern recognition systems. Artificial neural network is an adaptive system being widely used in pattern recognition systems that changes its structure based on external or internal information that flows through the network. Artificial neural networks are adjusted, or trained, so that a particular input leads to a specific target output. The neural network design part consists of two processes, training and application. The training of the neural network continues until the mean squared error reduces to a certain threshold or until the maximum number of iterations is reached. Once training is completed, the network can be applied for the actual classification of the data. The classification technique used may be one of the following:

1. Supervised classification -in which the input pattern fall as a member of a predefined class. 2. Unsupervised classification -in which the pattern falls into an unknown class as there are no predefined classes. The learning or training takes place only during the design phase of a pattern recognition system. Once the results obtained are satisfactory, the system is ready to perform the task of recognition on samples drawn from the environment in which it is expected to operate. b) Feed-forward neural networks Feed-forward networks are commonly used for pattern recognition. A three-layer feed forward neural network is typically composed of one input layer, one output layer and one hidden layer. In the input layer, each neuron corresponds to a given input pattern while in the output layer each neuron corresponds to a predefined pattern. Once a certain sample is input into the network, he best situation is that the output will be a vector with all elements as zero only except the one corresponding to the pattern that the sample belongs to. Of course, it is very complex to construct such types of neural networks. The commonly used networks for minimizing the cost are multi-layer-feed-forward neural networks, which uses the back-propagation learning algorithm for training neural networks. and transmits it to all units in the hidden layer z j (j=1,?.,p). 5. Each hidden unit z j sums its weighted input signals and applies its activation function and sends this signal to all units in the output layer y k (k=1,?.,m). [1] [2] [3]

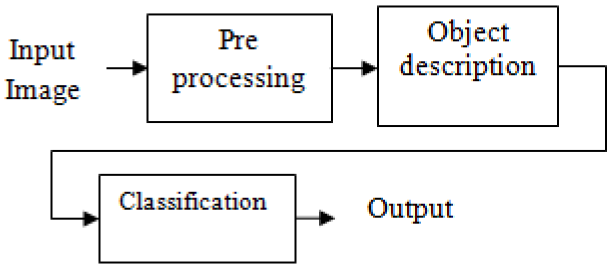
[1] DNormalized Vector Codes for Object Recognition using Artificial Neural Networks in the Framework of Picture Description Languages

[2] DNormalized Vector Codes for Object Recognition using Artificial Neural Networks in the Framework of Picture Description Languages

[3] DNormalized Vector Codes for Object Recognition using Artificial Neural Networks in the Framework of Picture Description Languages

1

Figure 1: Figure 1 :



2

Figure 2: Figure 2 :

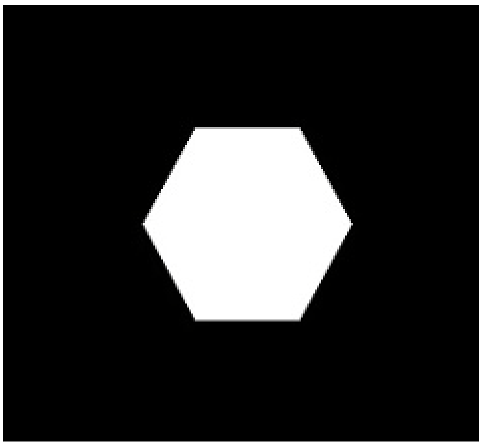


Figure 3:

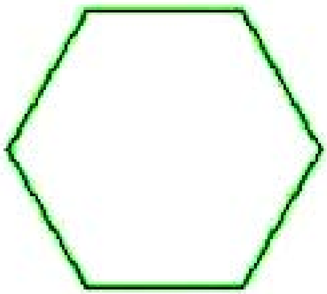


Figure 4:

34

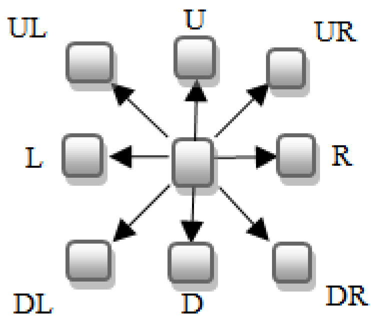


Figure 5: Figure 3 :Figure 4 :

6

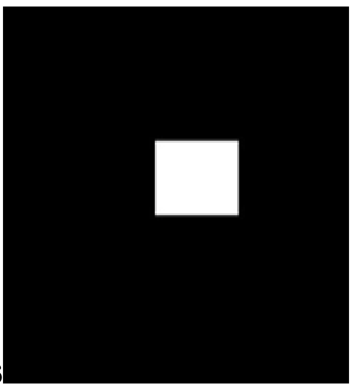


Figure 6: Figure 6 :D

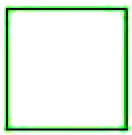


Figure 7:

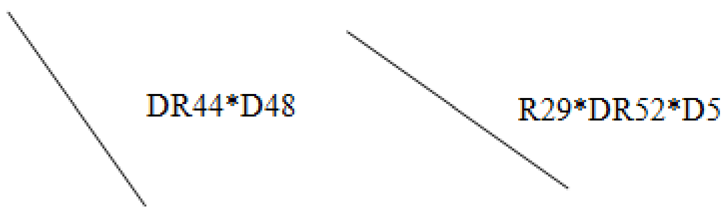


Figure 8:

6

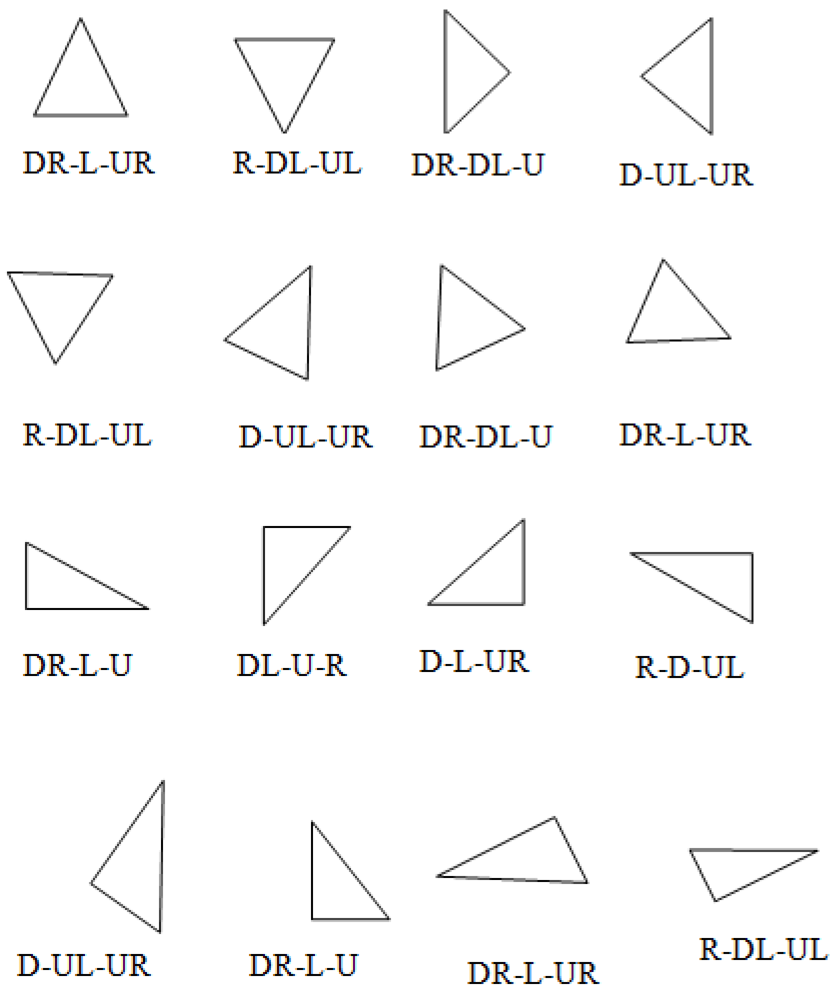


Figure 9: 6 .

7

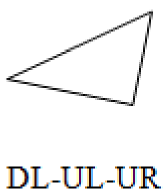


Figure 10: DFigure 7 :

1

Figure 11: Table 1

1

R	DR	D	DL	L	UL	U	UR	Class

Figure 12: Table 1 :

In the same way the DLC of real time images also can be approximated and normalized to fixed number of pixels. Such code gives us the input vector that can be given to the neural networks for the classification of their shapes as the shape in one of the 9 classes. Later more analysis can be done for further classification of these shapes inside the classes to produce more detailed classes for the actual identification of objects.

.1 CONCLUSION

In this Paper a syntactic approach is proposed for shape based object recognition. The knowledge vector is reduced to the input vector to a neural network by some vector approximation and normalization processes. The research efforts during the last decade have made significant progresses in both theoretical development and practical applications. The method presented here may offer a promising solution for object recognition problem.

[*R1*DR6*R1*DR2*R1*DR1*D1*DR1*R2*D1*DR1*D1] **R1*DR6*R1*DR2*R1*DR1*D1*DR1*R2*D1*DR1*D1*,

[*UL4*U1*UL1*L2*U1*UL1*L1*UL1*U3*UR5*R1*U R1*U1*UR1*R1*UR1*U1*UR2*R1*UR5*R1*UR1*R2* UR2*R1*UR3*R1 **UL4*U1*UL1*L2*U1*UL1*L1*UL1*U3*UR5*R1*U R1*U1*UR1*R1*UR1*U1*UR2*R1*UR5*R1*UR1*R2* UR2*R1*UR3*R1*UR2*R1*UR1*R2*UR1*R2*UR2*R2* UR1*R2*UR2*R1*UR1*U1*UL1*U2*UL2*U3*UL1*U3* UL2*U5*UL1*U20*UR1*U8*UR1*R11*UR1*R26*DR1* R7*DR1*R5*DR1*R5*DR1*R3*DR1*R2*UR3*D1*U2* UR2*U1*UR3*U1*UR4*R1*UR1*U2*UR1*R1*UR2*U1*,

[*UR1*U1*UR2*R1*UR1*R2*UR1*/<65,183> <64,184>/R26+160*DR74*D19+91*DL82*L29+153* UL74*U20+94*UR77*/<6 **UR1*U1*UR2*R1*UR1*R2*UR1*/<65,183># <64,184>/R26+160*DR74*D19+91*DL82*L29+153* UL74*U20+94*UR77*/<65,183># References Références Referencias*,

[Cohen and Wang] *3-D recognition and shape estimation from image contours*, F S Cohen , J.-Y Wang . (in Proc)

[<64,184>/R4*DR2*R2*D1*DR1*R2*DR4*R2*DR1*R1] *<64,184>/R4*DR2*R2*D1*DR1*R2*DR4*R2*DR1*R1*,

[Rajan (1993)] 'Cellular Logic Array Processing Techniques for High-Throughput Image Processing Systems'. E G Rajan . *The Indian Academy of Sciences* June 1993. 18 (2) p. . (Invited paper, SADHANA, Special Issue on Computer Vision)

[Mori et al. (2005)] 'Efficient shape matching using shape contexts'. Greg Mori , Member , Serge Ieee , Belongie , Ieee Member , Jitendra Malik , Member , Ieee . *IEEE Transactions on pattern analysis and machine intelligence*, November 2005. 27.

[IEEE Conf Computer Vision Pan. Recognition (1992)] *IEEE Conf Computer Vision Pan. Recognition*, June 1992.

[Hafiz et al.] 'Intelligent Object and Pattern Recognition using Ensembles in Back Propagation Neural Network'. T Hafiz , Muhammad U Hassan , Kashif Khalid , Imran . *International Journal of Electrical & Computer Sciences IJECS-IJENS* 10 p. 6.

[Rowley et al. ()] 'Neural network-based face detection'. H Rowley , S Baluja , T Kanade . *IEEE Transactions on Pattern Analysis and Machine intelligence* 1998. 20 p. .

[Belongie et al. (2002)] 'Shape matching and object recognition using shape contexts'. S Belongie , J Malik , J Puzicha . *IEEE Transactions on Pattern Analysis and Intelligence* April 2002. 4 (24) p. .

[Chen et al. ()] 'Two-Dimensional Packing For Irregular Shaped Objects'. Ping Chen , Zhaohui Fu , Andrew Lim , Brian Rodrigues . 0-7695-1874-5/03 $17.00. *Proceedings of the 36th Hawaii International Conference on System Sciences -2003*, (the 36th Hawaii International Conference on System Sciences -2003) 2003. IEEE.