

# Object Recognition Technique based on Level Set Method and Neural Network

V.N. Pawar

A. C. Patil College of Engineering  
Navi Mumbai ,India 410210

S.N. Talbar

S.G.G.S. Institute of Engineering & Tech.  
Nanded, India 431606

## ABSTRACT

The Object recognition is the task of finding and labeling parts of a two-dimensional (2D) image of a scene that correspond to objects in the scene. In this paper, we have proposed an efficient approach using level set method for extracting object shape contour and convex hull as a shape invariant features to the Feed forward Neural Network classifier for object recognition. We extracted the shape contour by level set method. Then, we have obtained invariant shape feature, convex hull of the objects. This convex hull set serves as a pattern for the Neural Network. Initially Feed forward neural network trained on the odd data set and tested on even data set. Our approach is evaluated on the Amsterdam Library of Object Images collection, a large collection of object images containing 1000 objects recorded under various imaging circumstances. The experimental results clearly demonstrate that our approach significantly outperforms. The proposed method is shown to be effective under a wide variety of imaging conditions.

## Keywords

Object Recognition, Shape feature, Convex Hull, Feed Forward Neural Network.

## 1. INTRODUCTION

Object recognition plays a crucial role in Computer Vision, especially in the semantic description of visual content [1], [2]. Building appropriate object models is central to object recognition, which is a fundamental problem in computer vision. Desirable characteristics of a model include good representation of objects, fast and efficient learning algorithms that require as little supervised information as possible. Although object recognition has been intensely studied, it still remains a hard and computational expensive problem. The main difficulty in the description of image content is the lack of information about the kind and the number of objects possibly present. Moreover, objects can appear at different locations in the image and they can be deformed, rotated, rescaled, differently illuminated or also occluded with respect to a reference view. The problem is to build a system able to recognize, as much as possible in an automatic way, the objects under various conditions. In order to simplify object detection and to reduce computational cost, many systems (e.g. [3]) limit the recognition to specific classes of objects. In these cases, a priori knowledge permits to select the most descriptive features for the objects at hand and to circumscribe the search space. However, even under this restriction, high classification performance is seldom reached. Moreover, many object recognition systems rely on user interaction to label as wrong or correct the returned items or to improve system response [4]. In [5], Nelson presented a technique to automatically extract a geometric description of an object by detecting semi- invariants at localized points. A

new concept was presented by Schmid and Mohr [6], where, instead of using geometric features, the authors use a set of differential invariants extracted from interest points Rao and Ballard [7] explore the use of local features for recognizing human faces.

In this paper, we have proposed object recognition technique with level set method and feed forward Neural Network. Proposed technique has four different phases namely pre-processing, extraction of shape of object, obtaining convex hull set of shape and used as a pattern to train and test the feed forward neural network for object recognition.

The organization of the paper is as follows. Section 2 describes feature extraction using Level set method, Section 3 describes feed forward neural network. Section 4 the experimental results and the conclusion are given in Section 5.

## 2. FEATURE EXTRACTION

Feature is anything that is localized, meaningful and detectable. In an image noticeable features include corners, lines, objects, colour, shape, spatial location, motion and texture. Features extracted from images define and describe the image content. Feature extraction is an important part in object recognition, which locates pixels in an image/object that have some distinctive characteristics. Shape of an object is the characteristic surface configuration as represented by the outline or contour. Shape recognition is one of the modes through which human perception of the environment is executed. It is important in object recognition because it corresponds to region of interests in images.

Shape feature representations are categorized according to the techniques used. They are boundary-based and region-based. Boundary-based technique describes the shape region by using its external characteristics, for example pixel along the object boundary, while the region-based technique describes the shape region by using its internal characteristics, for example the pixel contained in the region.

### 2.1 Level Set Method

The level set method (LSM) is a numerical technique for tracking interfaces and shapes. The main benefit of the level set method is that performance of numerical computations involving curves and surfaces on a fixed Cartesian grid without having to parameterize these objects, the level set method makes it very easy to follow shapes that change topology, i.e. when a shape splits in two, develops holes, or the reverse of these operations. All these make the level set method a great tool for modelling time-varying objects, like inflation of an airbag, or a drop of oil floating in water.

The level set method was first introduced by Osher and Sethian in 1987 [8], for capturing moving fronts of nonlinear hyperbolic PDEs. The method has been used for many applications, some of which are related to image processing.

The idea of the level set method is to implicitly represent fronts by the zero level of higher dimensional functions sum  $\Phi$ . Consider representation of a curve  $\mathbf{r}$  in  $\mathbb{R}^2$ . This curve is one dimensional. Using a level set approach, this curve is represented by the zero level of a two dimensional surface  $(x,y,\Phi(x,y))$  defined over  $\mathbb{R}^2$  i.e. the curve is the intersection of the plane  $(x, y, 0)$  and  $\Phi(x,y)$ . This is shown in Fig. 1 and 2. The implicit representation of  $\mathbf{r}$  makes it possible to treat the set interior to  $\mathbf{r}$ ,  $\text{int}(\mathbf{r})$ , as the set  $(x, y) : \Phi(x,y) < 0$  and exterior of  $\mathbf{r}$ ,  $\text{ext}(\mathbf{r})$  as the set  $(x, y) : \Phi(x,y) > 0$ , Note that using the level set representation of  $\mathbf{r}$ , these sets need not be connected. By introducing a time parameter  $t$ ,  $\mathbf{r}(t)$  can be implicitly moved by evolving  $\Phi(x,y)$ . How  $\Phi(x,y)$  is moved, is application dependent. In order to implicitly move the curve, the convection equation given below is solved.

$$\Phi_t + v \cdot \nabla \Phi = 0 \quad \text{--- (1)}$$

where  $v$  is a velocity field which can be dependent of time, geometry, internal or external physics, etc. When considering the evolvement of the curve in  $\mathbb{R}^2$ .

$$v \cdot \nabla \Phi = v_1 \Phi_x + v_2 \Phi_y \quad \text{--- (2)}$$

Equation (2) only says how to move the zero level of  $\Phi$  [9].

### 2.2 Convex Hull

Region based compact shape representation is given by convex hull. The convex hull of a geometric object (such as a point set or a polygon) is the smallest convex set containing that object. The convex hull of a set  $Q$  of points is the smallest convex polygon  $P$  for which each point  $Q$  is either on the boundary of  $P$  or in its interior [10].

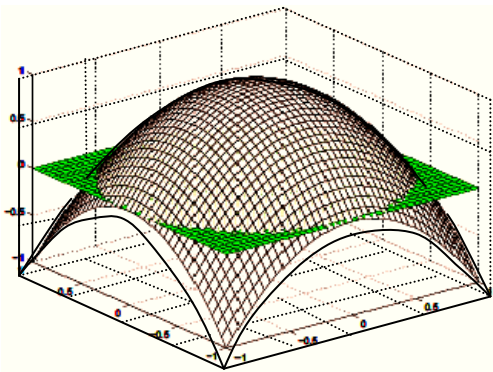


Fig 1 The level set function  $(x, y, \Phi(x, y))$  and  $(x,y,0)$

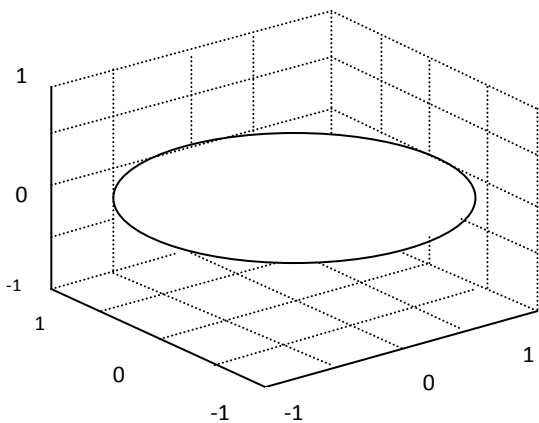


Fig 2  $\mathbf{r} : \Phi = 0$

### 3. NEURAL NETWORK

A Neural Network (NN), is a mathematical model that is inspired by the structure and functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases an NN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are nonlinear statistical data modelling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data.

Minsky Papert showed in 1967 that a two layer feed forward network can overcome many restrictions but did not present a solution to the problem of how to adjust the weights from input to hidden units. The central idea behind this solution is that the errors for the units of the hidden layer are determined by back propagating the errors of the units of the output layer. For this reason the method is often called the back propagation learning rule. Back propagation can also be considered as a generalisation of the delta rule for nonlinear activation functions and multilayer networks. In supervised learning assumes the availability of a labelled set of training data made up of  $M$  input output examples :

$$T = \{(x_i, d_i)\}_{i=1}^M \quad \text{----- (3)}$$

Where  $x_i$ = input vector of  $i^{\text{th}}$  example  
 $d_i$  = desired (target) response of  $i^{\text{th}}$  example, assumed to be scalar for convenience of presentation  
 $M$  = sample size

Given the training sample  $T$ , the requirement is to compute the free parameters of the neural network so that the actual output  $y_i$  of the neural network due to  $x_i$  is close enough to  $d_i$  for all  $i$  in a statistical sense. In the experimental result section we are using the mean-square error  $E(n)$  as the index of performance to be minimized.

$$E(n) = \frac{1}{M} \sum_{i=1}^M (d_i - y_i)^2 \quad \text{----- (4)}$$

### 3.1 The back-propagation algorithm

A multilayer perceptron (MLP) has an input layer of source nodes and an output layer of neurons (i.e., computation nodes) shown in Figure 3; these two layers connect the network to the outside world. In addition to these two layers, the multilayer perceptron usually has one or more layers of hidden neurons, because these neurons are not directly accessible. The hidden neurons extract important features contained in the input data. The training of an MLP is usually accomplished by using a backpropagation (BP) algorithm that involves two phases [11,12]:

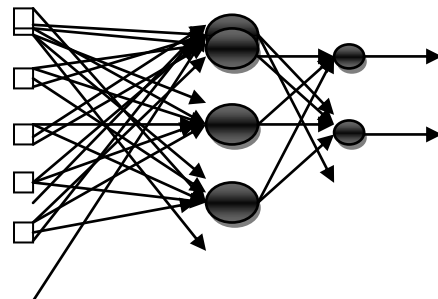


Fig 3 . Feed forward neural network with one hidden layer and one output layer.

Forward Phase: In this phase the free parameters of the network are fixed, and the input signal is propagated through the network of Figure 3 layer by layer. The forward phase finishes with the computation of an error signal

$$e_i = d_i - y_i \quad \text{--- (5)}$$

where  $d_i$  is the desired response and  $y_i$  is the actual output produced by the network in response to the input  $x_i$ .

Backward Phase : In this second phase, the error signal  $e_i$  is propagated through the network of Figure 3 in the backward direction, hence the name of the algorithm. It is during this phase that adjustments are applied to the free parameters of the network so as to minimize the error  $e_i$  in a statistical sense.

The back-propagation learning algorithm is simple to implement and computationally efficient in that its complexity is linear in the synaptic weights of the network. However, a major limitation of the algorithm is that it does not always converge and can be excruciatingly slow, particularly when we have to deal with a difficult learning task that requires the use of a large network.

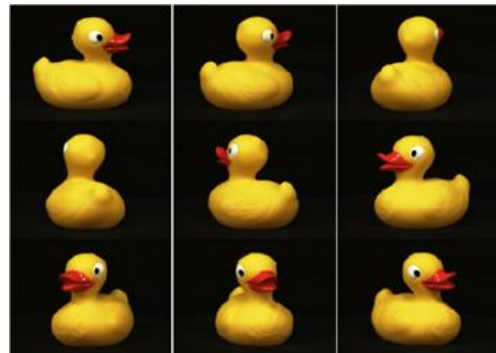
#### 4. EXPERIMENTAL RESULTS

In this section, several experiments are carried out to demonstrate the effectiveness of our proposed object recognition technique based on level set method and neural network approaches. The experiments are conducted in different levels such as Shape and convex hull set extraction, feedforward neural network training and testing for performance evaluation of proposed object recognition technique under varying conditions. Performance is evaluated with object recognition task on the Amsterdam Library of Object Images (ALOI) dataset [13, 14]. This public database has been widely used in object recognition literature. The ALOI collection consists of 1000 objects recorded under various imaging circumstances. For each object the viewing angle, illumination angle, and illumination colour are varied. The ALOI database consisting of 72000 images of 1000 different objects with 72 diverse views each. The sizes of the images are 256x256 with 8 bits per colour.

The images show a considerable amount of shadows, shading, and highlights. The combination of a large image dataset with a considerable variety of appearance, offers a formidable challenge for object recognition. Object recognition is the problem of matching one appearance of an object against a standardized version. One object may give rise to millions of different images, as camera conditions may be assorted endlessly. In our recognition experiment, one prototypical version of each object in the ALOI dataset is indexed and for querying, the diversity of the recorded object variations in the collection is being used. An object is accurately recognized when for all different variations the correct indexed object is returned. In this case, one may assume that the object can be recognized under a wide variety of real-life imaging circumstances. The objects of the data set are presented in Figure. 4. As an example, the sample views of the object are displayed in Figure. 5.

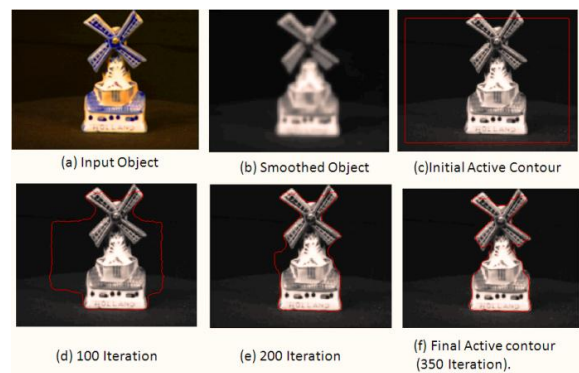


**Fig 4. Sample Objects from the ALOI Database**



**Fig 5. Example of One object in the ALOI database with different rotations**

The proposed variational level set method has been applied to a variety of images from ALOI database in different modalities. A new variational formulation for geometric active contours that forces the level set function to be close to a signed distance function, and therefore completely eliminates the need of the costly re-initialization procedure. We used a region based initialization of level set function, which is not only computationally more efficient than computing signed distance function, but also allows for more flexible applications. Our variational formulation consists of an internal energy term that penalizes the deviation of the level set function from a signed distance function, and an external energy term that drives the motion of the zero level set toward the desired image features, such as object boundaries. The resulting evolution of the level set function is the gradient flow that minimizes the overall energy functional. The intermediate results for geometric active contours are depicted in Figure 6.



**Fig 6. Intermediate Results of Shape Invariant Feature Extraction.**

The proposed System performance computed: the dataset is divided in two mutually disjoint sets, one used for the training phase and another for testing. The fact that the system is tested on a set, different from the one used for training, is a crucial factor for assessing the capability of the system in generalizing to different views. In particular, odd views are used to build the classifier and even views are used for testing. The experiments results of our proposed approach listed in the Table.1. The used ALOI database comprises of 1000 class image objects; for each objects 72 views are gathered, with a separation of 50. In our work, for analysis we have chosen 15 sample class objects randomly from the ALOI database. Four different datasets are formed from the ALOI dataset. Each consists of 6,8,12,18,24 and 36 images with a separation of  $60^0, 45^0, 30^0, 20^0, 15^0$  and  $10^0$  respectively are used for training the system. Afterwards, all 72 images in the test dataset are given as input for examining the performance of the proposed Object recognition techniques. The results of the experiments are illustrated in Table 2.

**Table 1 . Proposed technique performance for 15 image class**

Image Class	Mean Square Error	Percentage Accuracy
Apple teabags	0.00001595	99.71
Christmas bear	0.00001857	99.69
Shell	0.00001835	99.78
Green box	0.00001735	99.71
Shoe	0.00001215	99.74
Matchbox	0.00001484	99.78
Baby cream tube	0.00002811	99.63
Yellow bear	0.00001278	99.72
Blue bear	0.00002564	99.55
Optic lamp box	0.00002448	99.54
Colorful toy square	0.00002256	99.62
Chocolate box	0.00001616	99.7
Brown bear	0.00001436	99.74
Chess horse	0.0000228	99.63
Vitamin pills	0.000004059	99.91

**Table 2 . Performance of proposed technique with training of 6, 8, 12, 18, 24 and 36 data size**

Training Data size	Mean Square Error	Percentage Accuracy
6	0.0000225	99.56
8	0.00002912	99.57
12	0.00001836	99.61
18	0.00002159	99.69
24	0.00001856	99.72
36	0.0000132	99.74

## 5. CONCLUSION

We have proposed an efficient object recognition technique using level set method and neural network. In this paper, object shape contour is extracted by level set method and convex hull set descriptors have been proposed which are invariant to the viewpoint, geometry of the object and illumination conditions, this set have been served as the pattern for the forthcoming processes involved in the proposed approach. Subsequently we have trained feed forward neural network using different class of data set and tested using all set of data. We have attained distinct results concerned with different class of the objects. We have evaluated our proposed approach on the ALOI collection, a large collection of object images consists of 1000 objects recorded under various imaging circumstances. The experiments clearly demonstrated that our proposed approach significantly outperforms the state-of-the- art methods with level set method for shape features. The proposed method is shown to be effective under a wide variety of imaging conditions.

## 6. REFERENCES

- [1] P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth, Object Recognition as Machine Translation: Learning a lexicon for a fixed image vocabulary, In European Conference on Computer Vision (ECCV) Copenhagen, 2002
- [2] D. A. Forsyth and J. Ponce, Computer Vision: a modern approach, Prentice Hall,2002
- [3] A. Hoogs, R. Collins, R. Kaucic, and J. Mundy, A common set of perceptual observables for grouping, figure - ground discrimination, and texture classification, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 4, pp. 458 ,474, 2003.
- [4] B. Ko and H. Byun, Extracting Salient Regions And Learning Importance Scores In Region-Based Image Retrieval, International Journal of Patter Recognition and Artificial Intelligence, Vol. 17, No. 8, pp. 1349 ,1367, 2003.
- [5] R. C. Nelson, Memory-based recognition for 3-d objects, In ARPA Image Understanding Workshop, pp: 1305 ,1310, February 1996.
- [6] C. Schmid and R. Mohr, Local gray value invariants for image retrieval, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.19, Issue.5, 1997, pp: 530,535.
- [7] R.P.N. Rao and D.H. Ballard, Natural basis functions and topographic memory for face recognition, In International Joint Conference on Artificial Intelligence, pp: 10,17, 1995.
- [8] S. Osher and J. Sethian, Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulation. J. Comput. Phys., vol. 79, pp. 12, 49, 1988.
- [9] Johan Lie Variational Image Segmentation using Discontinuous Level Set Formulations, Cand. Scient. Thesis in Applied Mathematics, University of Bergen, December 12th 2003
- [10] Thomas H. Cormen Introduction to Algorithms, Second edition, MIT Press, 2001

- [11] Werbos, P. J., Beyond regression: New tools for prediction and analysis in the behavioral sciences, Ph.D. Thesis, Harvard University, Cambridge, MA. 1974
- [12] Rumelhart, D. E., G. E. Hinton, and R. J. Williams, Learning internal representations by error propagation, in D. E. Rumelhart and J. L. McClelland, eds. (Cambridge, MA: MIT Press), vol. 1, Chapter 8. 1986
- [13] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, The Amsterdam library of object images, *Int. J. Comput. Vision*, Vol. 61, No. 1, pp. 103, 112, January, 2005.
- [14] Amsterdam Library of Object Images (ALOI) Datasets from <http://staff.science.uva.nl/aloi/>

## **7. AUTHORS PROFILE**

**V. N. Pawar** is a PhD student at SGGS Institute of Engineering & Technology, Vishnupuri, Nanded, India doing research in the field of object recognition systems. He

received B.E. Electronics and M.E. (Control Systems) from Shivaji University, Kolhapur, India in 1990 and 1997 respectively. He is life fellow member of IETE and ISTE. His research interests are in Computer Vision, Embedded Systems, Image Processing and Robotics.

**S. N. Talbar** has received his B.E and M.E degrees from SGGS Institute of Technology, Nanded, India in 1985 and 1990 respectively. He obtained his PhD (Highest degree) from SRTM University, India in 2000. He is recipient of the Young Scientist Award by URSI, Italy in 2003. He had Collaborative research programme at Cardiff University Wales, UK. Currently he is working as Professor in Electronics & Telecommunication Department of SGGS Institute of Technology Nanded, India. He is a member of many prestigious committees in academic field of India. His research interests are Signal & Image processing, Pattern Recognition and Embedded System Design. He has published about 40 papers in Journals and about 105 papers in Conferences.