Robust Energy-Based Image Segmentation Using Nonparametric Joint Shape and Feature Priors

DR. B. HEMANTH KUMAR[1], MOPURI SRINIVAS[2], SHAIK ASADHULLA[3], PAMMI PRANAY[4], PANGULURI RAJESH[5]

[1, 2, 3, 4, 5] *Information Technology, R.V.R & J.C College of Engineering*

Abstract- Image segmentation is a key task in computer vision that identifies important structures in images. Traditional methods like thresholding and region-based segmentation often struggle with noise and changes in shape. This paper introduces a segmentation approach that uses nonparametric joint shape and feature priors to enhance accuracy. Shape priors are learned from the MNIST dataset, and feature representations are extracted through Principal Component Analysis (PCA). We formulate the segmentation as an energy minimization problem that combines data fidelity, shape, and feature terms. Experimental results, measured using Dice coefficient and Hausdorff distance, show the effectiveness of this method.

Index Terms- Image Segmentation, Shape Priors, Feature Priors, Principal Component Analysis (PCA), Energy Minimization, MNIST Dataset, Dice Coefficient.

I. INTRODUCTION

Image segmentation is a key task in computer vision. It aims to split an image into meaningful regions for easier analysis and interpretation. This technique finds use in various areas, including medical imaging, pattern recognition, object detection, and document analysis [1]. Traditional segmentation methods include thresholding, edge-based techniques, and region-based approaches. Among these, thresholding methods, such as the Otsu method, are popular for their simplicity and effectiveness in distinguishing the foreground from the background [2]. However, these methods often struggle when dealing with noise, irregular shapes, or complex backgrounds. Region-based segmentation methods group pixels with similar features but can yield incorrect results when object boundaries are unclear [3].To overcome these limitations, shape prior models have been developed. They incorporate prior knowledge about object structures during segmentation. These methods enhance segmentation accuracy by limiting possible shapes based on training data [4]. Additionally, feature extraction techniques like Principal Component Analysis (PCA) capture key variations in image features and lower dimensionality [5].

Inspired by these approaches, this paper presents a segmentation framework that uses nonparametric joint shape and feature priors. The proposed method learns shape templates from training images and employs PCA-based feature priors to guide segmentation through an energy minimization framework.

1.1 Applications of Image Segmentation

Image segmentation is commonly used in areas like medical image analysis, object detection, satellite image interpretation, and handwritten character recognition. Accurate identification of objects is essential for further processing and decision making [1]. It is also used in autonomous driving, surveillance systems, and document analysis to pull out useful information from complex visual data [2].

1.2 Literature Survey

The scientific community has introduced different methods for image segmentation which enable precise detection of objects located in the foreground while removing all background elements. The Otsu algorithm together with other thresholding methods serves as a popular choice for gray-level image segmentation because of its straightforward design and effective performance [1]. Region-based segmentation techniques group neighboring pixels with similar properties to form homogeneous regions, but they may struggle with noisy images and complex boundaries [2]. The edge-based methods establish object boundaries through their ability to detect abrupt changes in image intensity, which enables them to identify sudden shifts that occur throughout visual material [3]. The development of segmentation techniques that use shape priors has become increasingly popular during the last few years because these methods combine existing information about object shapes to achieve better results in segmenting materials [4]. In image analysis tasks, researchers have applied statistical methods like Principal Component Analysis (PCA) to extract essential features while decreasing the number of dimensions to be analyzed [5]. The integration of shape and feature priors through a nonparametric framework has improved segmentation accuracy because it enables segmentation to use both structural and statistical data obtained from training databases [6].

1.3 Motivation

Traditional image segmentation techniques such as thresholding and edge-based methods face challenges when they attempt to segment images which contain noise and complex shapes and different object structures in the image [1] [2]. The need for segmentation methods which use shape and feature priors to achieve better object extraction results from images drives this research. The need for segmentation methods which use shape and feature priors to achieve better object extraction results from images drives this research [5],[7].

1.4 Objectives of the Project

The objective of this work is to develop an image segmentation framework that integrates nonparametric shape priors together with PCA-based feature priors to enable more precise and dependable segmentation of handwritten digit images [5],[7],[8].

1.5 Scope of the Work

The proposed segmentation framework can be applied to handwritten character recognition, document analysis, and other image analysis tasks which need shape and feature priors to achieve better segmentation results which show increased accuracy and improved efficiency of operations [1], [5], [8].

1.6 Segmentation Databases

The researchers of this study performed their experiments with the MNIST dataset, which includes grayscale images of handwritten digits from 0 to 9 and serves as a standard testing set for image processing and pattern recognition studies [8]

II. VARIANTS OF IMAGE SEGMENTATION

2.1 Thresholding

2.1.1 Introduction

Thresholding stands as the most basic method which researchers commonly use for image segmentation because it separates pixels into two groups according to their brightness levels. The method establishes a threshold value which allows the system to identify pixels that belong to either the object or background area. Global thresholding methods apply a single threshold to the entire image, while adaptive thresholding methods determine thresholds based on local image characteristics. The Otsu method serves as a widely used thresholding technique which automatically determines the best threshold by maximizing pixel class variance according to [2]. Thresholding methods achieve high computational efficiency and straightforward implementation, yet they struggle to handle images contaminated by noise and featuring inconsistent lighting and shared brightness patterns according to [1]. The thresholding method persists as an essential basic technique used within various image segmentation tasks despite its limitations.

2.1.2 Methodology

Thresholding serves as a basic method for image segmentation which uses pixel intensity values to distinguish between objects and their background. The method converts a grayscale image into a binary image by selecting a threshold value that classifies pixels into foreground and background regions. The process starts with converting the input image into a grayscale format which helps to make segmentation easier. Gaussian filtering serves as a noise reduction technique which helps to smooth the image while eliminating undesired image changes that interfere with threshold selection.

The segmentation process can be mathematically expressed as:

$$g(x,y) = \begin{cases} 1, & \text{if } f(x,y) \geq T \\ 0, & \text{if } f(x,y) < T \end{cases}$$

The function f(x,y) shows the intensity value for a pixel located at the coordinates (x,y) while g(x,y) displays the binary image that has been processed through segmentation and T serves as the threshold value.

The probabilities of the two classes are expressed as:

$$\omega_0 = \sum_{i=0}^{T} p(i)$$

$$\omega_1 = \sum_{i=T+1}^{L-1} p(i)$$

The class means are given by:

$$\mu_0 = \frac{\sum_{i=0}^{T} i\,p(i)}{\omega_0}$$

$$\mu_1 = \frac{\sum_{i=T+1}^{L-1} i\,p(i)}{\omega_1}$$

The between-class variance is calculated as:

$$\sigma_b^2 = \omega_0\omega_1(\mu_0 - \mu_1)^2$$

The optimal threshold is selected such that the between-class variance is maximized:

$$T^* = \arg\max \sigma_b^2$$

After finding the best threshold point the researchers changed the grayscale image into a binary image which showed the object region through one class and the background through another class. Small artifacts which disrupt the segmented outcome can be eliminated through the application of morphological operations which include erosion and dilation.

2.2 Region Based Segmentation

2.2.1 Introduction

The image processing technique known as region-based segmentation divides an image into different regions based on the intensity and color and texture similarities of its pixels. This method partitions an image into multiple regions which display similar characteristics between their adjacent pixels. Region-based segmentation identifies all connected areas in an image whereas edge-based methods identify only the boundary lines between different regions. The method uses three main techniques which include region growing and region splitting and region merging to combine pixels and regions that meet specific similarity requirements. Region growing starts with chosen seed points which enable the technique to grow regions through the addition of nearby pixels that exhibit matching traits. The accuracy of region-based methods enables them to create precise image segmentation results for homogeneous images but their performance decreases when dealing with noisy images or images that contain undefined boundaries between regions [1],[4] ,[10].

2.2.2 Methodology

Region-based segmentation is an image segmentation technique that partitions an image into multiple regions based on similarity criteria such as intensity, color, or texture. The primary goal of this system works to combine adjacent pixels which share identical characteristics into complete areas that show important elements from the photograph. The input image undergoes conversion into a grayscale format because this process helps to make segmentation easier while decreasing the amount of processing power needed for the task. The system uses Gaussian filtering as a noise reduction method to create a clear image which removes all unwanted changes that disrupt the creation of different regions. The step before main processing helps to improve segmentation results because it makes pixel values more uniform. Region-based segmentation begins with the selection of one or more seed points that serve as the initial regions. The selection of seed points allows users to choose between manual selection and automatic selection which uses pixel intensity and gradient magnitude and other feature descriptors as selection criteria. Each seed point represents the starting location for the region-growing process. The region expands through time as the system adds pixels from adjacent areas which match the seed pixel or current region average intensity. The system uses a similarity condition to decide which neighboring pixels will join the existing region.

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715105

The similarity condition can be expressed as:

$$|I(x, y) - I_s| \leq T$$

Where the function I(x,y) shows the intensity value of the adjacent pixel which exists at the coordinates (x,y). The intensity value of the seed pixel or the average region intensity is represented by I_s. The function requires T which serves as its fixed threshold value.

0	1	2	0
2	5	6	1
1	4	7	3
0	2	5	1

→

0	0	0	0
0	1	1	0
0	0	1	0
0	0	1	0

Figure 1- Region Growing

The process of segmentation enables the combination of two neighboring areas which possess identical characteristics to create one bigger region that maintains uniform characteristics. The method decreases excessive segmentation while it enhances the accuracy of displaying object shapes.

The merging condition can be defined as:

$$|\mu_i - \mu_j| < T$$

Where μ_i and μ_j represent the mean intensity values for the two neighboring regions. If the difference in their mean intensities were less than the predefined threshold, then the regions would be merged to form a single segment.

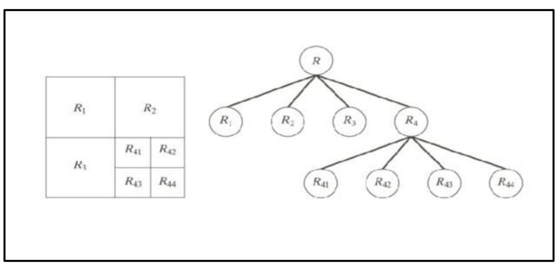


Figure 2- Region Merging

III. PROPOSED METHODOLOGY

3. Nonparametric Joint Shape and Feature Priors for Image Segmentation

3.1 Introduction

Researchers developed nonparametric joint shape and feature prior models to enhance image segmentation accuracy through the integration of existing knowledge about object structures and statistical attributes. The traditional methods of image segmentation depend solely on pixel intensity which leads to incorrect outcomes when working with noisy images or images that contain intricate shapes. Shape priors establish boundaries for object segmentation by providing restrictions that limit potential object shapes to previously documented patterns [5]. Feature priors employ Principal Component Analysis (PCA) [7] to extract statistical patterns that exist within image data. The combination of structural and feature data enables the segmentation process to achieve greater reliability through enhanced stability. The nonparametric framework enables the model to construct shape templates from training data because it does not require any specific parametric distribution to be set beforehand. This method has demonstrated better results for both handwritten digit segmentation tasks and medical image processing tasks [5], [8].

3.2 Methodology

The proposed segmentation framework integrates shape priors and feature priors within an energy minimization framework to improve segmentation accuracy. The approach learns representative object structures from training images and uses these learned patterns to guide the segmentation of new input images.The segmentation model uses the MNIST handwritten digit dataset for training which includes grayscale images of handwritten digits that range from 0 to 9.Each image is normalized to a fixed resolution to maintain consistency in shape representation.

The training dataset be represented as:

$$D = \{I_1, I_2, I_3, \ldots, I_n\}$$

where I_i represents the i^{th} training image and n is the total number of images in the dataset. These images are used to construct the shape prior database.

The mean shape for each class is calculated as:

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715105

$$S_{mean} = \frac{1}{N}\sum_{i=1}^{N} S_i$$

where N denotes the shape of the i^{th} training sample and S_i denotes the total number of samples within that class.

Each image is represented as feature vector as below:

$$X = [x_1, x_2, x_3, \ldots, x_m]$$

The covariance matrix is calculated as:

$$C = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)(X_i - \mu)^T$$

Where μ symbolizes the mean vector feature of the dataset.

Segmentation needs to be solved through energy minimization which requires three specific components to be used as the solution method. They are Data fidelity term, Shape prior term, Feature prior term.

The Energy function is calculated as:

$$E = E_{data} + \lambda_s E_{shape} + \lambda_f E_{feature}$$

where

- E_{data} measures similarity between the observed image and the template,
- E_{shape} enforces shape consistency using the learned shape prior,
- $E_{feature}$ incorporates PCA-based feature constraints,
- λ_s and λ_f are weighting parameters controlling the influence of shape and feature priors.

The input image requires multiple preprocessing steps before it can be segmented into separate parts. The first step involves converting the image into a grayscale format. The second step uses Gaussian filtering to eliminate undesirable noise from the image. The third step of the process establishes a standardized level of contrast throughout the image. The fourth step uses Otsu's Method to create a threshold for the image. The image quality improvements from these preprocessing steps enable precise segmentation work. The algorithm assesses each input image by matching it against every shape template that exists in its database.

The optimal match is chosen as the template that minimizes the energy function:

$$T^* = \arg\min(E)$$

Where T^* represents the optimal template.

IV. COMPARATIVE STUDY OF IMAGE SEGMENTATION

4.1 Introduction

Researchers have developed various image segmentation methods which enable precise object extraction from digital image backgrounds. The traditional method of thresholding uses pixel intensity values to detect objects while region-based segmentation divides pixels into groups that share visual and spatial characteristics [1], [2]. The methods demonstrate computational efficiency and simplicity but face challenges when processing images that contain noise and complex shapes and overlapping intensity patterns. The solution to this problem requires advanced methods which use nonparametric joint shape and feature prior models to combine existing object structure knowledge with statistical feature data. The models use training data to assist the segmentation process which leads to better accuracy results [5]. The process of method comparison enables researchers to identify the strengths and weaknesses of different approaches used for practical image segmentation tasks.

4.2 Performance Evaluation

4.2.1 Performance Metrics

The proposed segmentation method performance assessment uses quantitative metrics which include Dice coefficient and Hausdorff distance to measure segmentation accuracy and boundary similarity. The Dice coefficient evaluates the overlap between the predicted segmentation and the ground truth, while the Hausdorff distance measures the maximum boundary deviation between two segmented regions. The metrics enable assessment of image segmentation algorithm reliability and precision through their evaluation process [1],[10].

The Dice Coefficient is calculated as:

$$Dice = \frac{2|A \cap B|}{|A| + |B|}$$

A depicted segmented region serves as the predicted segmenting output while B shown to its actual boundary. The ground truth region serves as the actual boundary of the B area.

The Hausdorff distance is calculated as:

$$H(A,B) = \max\left\{\sup_{a\in A}\inf_{b\in B} d(a,b), \sup_{b\in B}\inf_{a\in A} d(a,b)\right\}$$

The formula compares two data sets which contain boundary points from predicted segmentation and ground truth data. The function d defines the Euclidean distance between two points a and b.

4.2.2 Comparison

Thresholding serves as an effective method for segmenting images because it requires minimal computational resources yet succeeds in separating objects through its analysis of pixel intensity values, yet the method fails to deliver reliable results when applied to images containing noise or inconsistent lighting conditions [2]. Region-based segmentation improves this by grouping neighboring pixels with similar characteristics, producing more coherent regions but sometimes failing when object boundaries are unclear [1]. The nonparametric joint shape and feature prior method uses training data to establish segmentation guidelines through learned information. The combination of shape templates with PCA-based feature representations results in this method, which delivers enhanced segmentation performance. The two techniques achieve different accuracy levels, with prior-based methods delivering superior results in structured image segmentation tasks compared to standard approaches [5], [7].

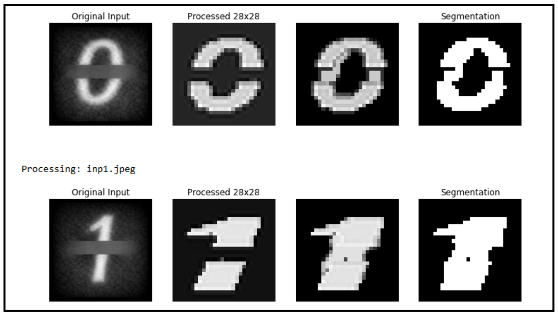
4.2.3 Segmentation Analysis

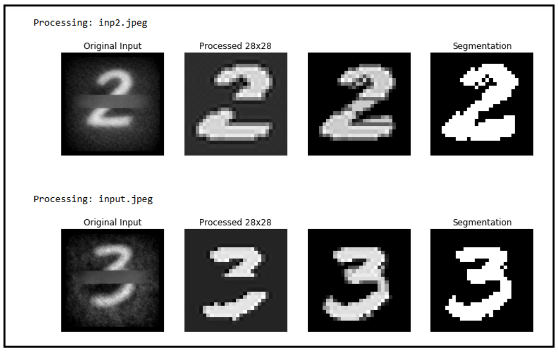
The process of segmentation analysis tests how well various algorithms can isolate objects from their backgrounds in images. The basic segmentation results of traditional methods which include thresholding and region-based segmentation face challenges when they deal with complex shapes and noisy images [1]. The nonparametric joint shape and feature prior method enhances segmentation results through its ability to use structural and statistical information derived from training data [5]. The performance of these methods can be quantitatively analyzed using metrics such as Dice coefficient and Hausdorff distance to measure segmentation accuracy and boundary similarity.

4.3 Results and Discussion

The study results prove that different segmentation methods for image datasets produce different accuracy results. Thresholding methods provide fast segmentation but they frequently encounter difficulties with noise and changing light conditions according to reference [2]. Region-based segmentation improves region continuity through pixel grouping of similar pixels yet it encounters challenges when object boundaries lack distinct edges according to reference [1]. The nonparametric joint shape and feature prior method which the researchers developed demonstrates better segmentation results through its capacity to extract structural and statistical data from training samples. The method achieves better segmentation results through the application of shape templates and PCA-based feature priors to structured image content which includes handwritten digits according to references [5] and [7]. The proposed method proved effective through its quantitative evaluation which used Dice coefficient and Hausdorff distance metrics. The results demonstrate that using shape and feature priors together leads to better segmentation results than using traditional methods.

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715105





================ SUMMARY TABLE ================

Dice	Hausdorff	Image
0.8462	2.00	inp0.jpeg
0.9031	5.39	inp1.jpeg
0.8517	2.24	inp2.jpeg
0.9341	1.41	input.jpeg

V. CONCLUSION AND FUTURE WORK

5.1 Conclusion

The research evaluated thresholding, region-based segmentation, and nonparametric joint shape and feature prior methods as different image segmentation techniques. Basic segmentation results arise from traditional methods which include thresholding and region-based segmentation but these methods encounter problems when they must segment noisy images with complex object shapes. The proposed nonparametric approach improves segmentation accuracy by using shape priors together with feature representations which were derived from training data. The experimental results demonstrate the method's effectiveness through evaluation with Dice coefficient and Hausdorff distance metrics. The combination of shape and feature priors enables better image segmentation accuracy through its implementation as a robust solution.

5.2 Future Scope

Future work can focus on improving segmentation accuracy by integrating deep learning techniques with shape and feature prior models. The proposed framework can handle additional complex datasets which include medical and satellite image data. The algorithm will gain more practical value when it achieves better performance for real-time image processing tasks.

REFERENCES

[1] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed., Pearson Education, 2018.

[2] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62–66, 1979.

[3] J. Canny, “A Computational Approach to Edge Detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679–698, 1986.

[4] L. G. Shapiro and G. C. Stockman, Computer Vision, Prentice Hall, 2001.

[5] D. Cremers, S. Osher, and S. Soatto, “Kernel Density Estimation and Intrinsic Alignment for Shape Priors in Level Set Segmentation,” International Journal of Computer Vision, vol. 69, no. 3, pp. 335–351, 2006.

[6] T. F. Chan and L. A. Vese, “Active Contours Without Edges,” IEEE Transactions on Image Processing, vol. 10, no. 2, pp. 266–277, 2001.

[7] I. T. Jolliffe, Principal Component Analysis, Springer, 2002.

[8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-Based Learning Applied to Document Recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[9] A. K. Jain, “Data Clustering: 50 Years Beyond K-Means,” Pattern Recognition Letters, vol. 31, no. 8, pp. 651–666, 2010.

[10] D. L. Pham, C. Xu, and J. L. Prince, “Current Methods in Medical Image Segmentation,” Annual Review of Biomedical Engineering, vol. 2, pp. 315–337, 2000.