View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by CiteSeerX

 ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Segmentation of Structured Objects in Image

[1]Rajeshwari Rasal, [2]Prof. N.M.Shahane
[1]ME Student, Dept. of Computer Engineering, KKWIEER, Savitribai Phule Pune University, India
[2]Associate Professor, Dept. of Computer Engg., KKWIEER, Savitribai Phule Pune University, India

Abstract
Detection of foreground structured objects in the images is an essential task in many image processing applications. This paper presents a region merging and region growing approach for automatic detection of the foreground objects in the image. The proposed approach identifies objects in the given image based on general properties of the objects without depending on the prior knowledge about specific objects. The region contrast information is used to separate the regions of the structured objects from the background regions. The perceptual organization laws are used in the region merging process to group the various regions i.e. parts of the object. The system is adaptive to the image content. The results of the experiments show that the proposed scheme can efficiently extract object boundary from the background.

Keywords
Segmentation, Perceptual Organization, Region grouping, Symmetry, Proximity, Contrast

I. Introduction
Image segmentation is the task of dividing an image into coherent regions so that each region corresponds to an object or area of interest. The detection of structured foreground objects in an image is useful in many applications like image classification, image retrieval, content-aware image resizing etc.
The structured objects are more difficult to identify as they are composed of multiple parts. Different parts of object may have distinct surface characteristics (e.g. colors, texture etc.). The proposed approach tries to detect the structured foreground objects in the image using region contrast information and perceptual organization of object parts. The approach is based on the general properties of real-world objects (e.g. similarity, proximity etc.) and hence does not depend on specific properties of objects.
The structured objects usually have high contrast to their background. Hence region contrast information is used to separate the regions that belong to unstructured and structured objects. Perceptual organization is the basic capability of human visual system to identify relevant groupings and structures from an image without any prior knowledge of its content. The Gestalt Laws are based on human visual perception of objects. In the proposed approach Law of similarity, Law of Symmetry, Law of Alignment & Law of Proximity are used to group the regions together to form a region that constitutes a structured object in an image.
The proposed approach applies Gestalt laws to image regions and merge the regions together to identify one single region that represents a structured object. The accuracy of region merging is measured by using boundary energy function.

II. Related Works
In the top down approach of image segmentation prior knowledge about an object such as its shape, color or texture is used to guide the segmentation. In bottom-up approach the pixels are grouped according to the similarity among low-level features such as color, textures etc.
There are two types of bottom-up segmentation methods, region-based and edge based segmentation. In the graph based image segmentation approach each pixel of the image is equivalent to a node in the graph and edges represent adjacent pixels. Weights on each edge are the dissimilarity between pixels. The boundaries between regions are defined by measuring the dissimilarity between the neighboring pixels. In the Ncut method the nodes are arranged into groups so that within the group the similarity is high and between the groups the similarity is low.
In the region-based approach, each pixel is assigned to a particular region. In region growing method of image segmentation a seed region is selected at start and a new region is identified by merging as many neighboring regions with the seed region.
The boundary detection of objects was also implemented as a supervised learning problem [12]. A large data set of human-labeled boundaries in natural images is used to train a boundary model. The model can then identify boundary pixels based on a set of low-level cues such as brightness, color and texture extracted from local image patches.
In the multi-class image segmentation technique [9] a number of classes (e.g., road, sky, water, etc.) are defined for labeling every pixel in an image. In this method the image is first segmented into multiple coherent initial regions and then each region is assigned to one of predefined classes. Gestalt grouping laws can also be used for segmenting an image. They used boundary energy functions to identify the regions that can be grouped together. The energy function includes information about the region. The energy function measures the accuracy of grouping the regions.

III. Problem Statement
Accurate detection of foreground objects in an image is a difficult task as images may contain several different types of objects. Also an object is composed of multiple parts with each part having different surface characteristics.
In this paper we propose a novel approach to identify boundary of foreground objects in an image by using general properties of real world objects.
Therefore the problem is to design, develop and implement a system, that takes an image as input, separates the background and foreground object regions and detects boundary of structured objects solely based on perceptual organization laws, without depending on a priori knowledge of the specific objects.

IV. Objective
Given an image, the objective is to detect boundary of structured objects in the image solely based on some general properties of the real world objects, without depending on the a priori knowledge of the specific objects.

V. Methodology
The proposed approach uses region contrast information and Gestalt perceptual organization laws to identify the structured objects in the image. Contrast is defined as the difference between maximum and minimum intensities in the image. Perceptual organization is the basic capability of human vision system to identify relevant groupings in an image without prior knowledge of its content. The Gestalt laws reflect the general properties of man-made and natural objects.

Figure 1 shows the major steps in the system:

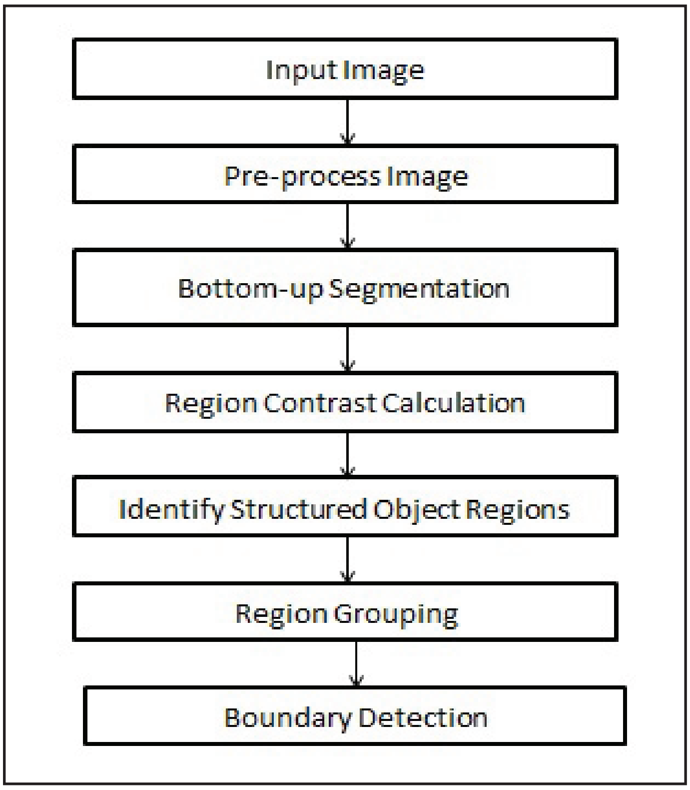


Fig. 1: Steps for Boundary Detection of Structured Objects

The input to the system is an image in the RGB color format. The preprocessing involves applying a smoothing filter to remove the noise from the input image.

The image is first segmented into multiple uniform regions (i.e. superpixels) by using the Felzenszwalb and Huttenlocher's [3] segmentation algorithm. These initial super-pixels carry more information as compared to pixels as well as use of superpixels reduces execution time as number of superpixels is less than number of image pixels. Also the superpixels having similar color are merged together. These initial regions approximately correspond to parts of the object.
As the object detection is done based on color contrast information, a color histogram is created for each region. The structured objects have high contrast to background. Hence based on the region contrast information we can identify regions that belong to structured object [6]. For each region its color contrast is calculated based on region's histogram, by measuring its color contrast to all the other regions in the image.
The proposed approach used thresholding technique to identify the regions that form the structured object. A threshold of region contrast is calculated as,

$$\text{Threshold} = \text{mean}_{\text{region\_contrast}}$$

The regions having region contrast more than threshold are considered as regions that belong to structured objects. The initial regions or superpixels roughly correspond to object parts. The regions having region contrast less than threshold are considered as regions that belong to background (i.e. unstructured objects).

Next these multiple regions (i.e. object parts) are grouped together to detect a structured object[7]. To start the grouping process, we select a seed region from the set of structured object regions.

The seed region is a region with highest region contrast. The regions are grouped together by using perceptual organization laws (i.e. similarity, symmetry, alignment and proximity). The regions are grouped together based on their structural relationships.

A boundary energy function is defined to measure the accuracy of region grouping. The energy function incorporates region information. This region grouping procedure is repeated for multiple rounds until no region can be grouped with other region. The function tries to find the best region that corresponds to structured object. The structured object is extracted from the background when the merging process ends. The output generated is an image with detected structured objects.

VI. Results & Discussion

The system is tested on Berkley dataset. The Berkley dataset is widely used as a benchmark dataset for many boundary detection and segmentation algorithms.

The dataset contains images with a wide variety of man-made and biological objects with various types of backgrounds. The size of the images are 481 X 321.This dataset also provides ground truth object segmentation.

The input to the system is an image file and output is an image with foreground structured object separated from the background. The accuracy of boundary detection is measured by using F-Measure. F measure considers both the precision and the recall of the test to compute the score.
Precision p is defined as

$$\frac{\textit{Total number of boundary pixels correctly detected}}{\textit{Total number of boundary pixels detected}}$$

Recall is defined as

$$\frac{\textit{Total number of boundary pixels correctly detected}}{\textit{Total number of boundary pixels in groundtruth}}$$

The F score can be interpreted as a weighted average of the precision and recall, where an F score reaches its best value at 1 and worst score at 0. The F score is given as,

$$\text{F score} = 2 * \frac{precision * recall}{precision + recall}$$

The images in the dataset are falling into 3 categories :
- Images containing one object (70 images)
- Images containing two objects (34 images)
- Images containing three or more objects (196 images)

ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Figures 2 shows the results of the system on images under different categories.

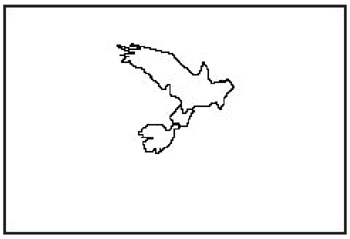
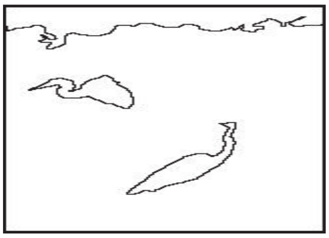
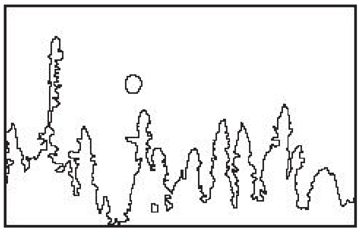
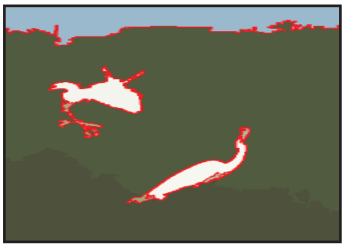
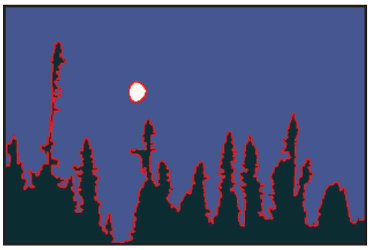
	Image with 1 object	Image with 2 objects	Image with 3 or more objects
Input Image			
Ground truth			
Output Image			
Precision	0.82	0.83	0.81
Recall	0.92	0.98	0.95
F Measure	0.87	0.90	0.88

Fig. 2: Sample of Structured Object Detection Using Berkley Dataset

Table 1 shows average precision, recall and F-measure values for each image category.

Table 1: Test Results of Proposed System

Image Category	Total No. of Images	Avg.Precision	Avg. Recall	Avg. F-measure
1 object	70	0.84	0.91	0.87
2 objects	34	0.86	0.92	0.88
3 or Multiple objects	196	0.80	0.88	0.83
Total	300	0.83	0.90	0.86

Table 2 shows comparison between existing method and the proposed method based on precision, recall and F-measure values. The results of both the systems are tested on Berkley dataset with total 300 images.

Table 2: Comparison of Test Results

Parameter	Existing Method [1]	Proposed Method
Precision	0.75	0.83
Recall	0.69	0.90
F-Measure	0.72	0.86

V. Conclusion

We have presented a system for boundary detection of foreground structured objects. The system follows a region merging & growing approach. The system uses region contrast information and perceptual organization laws to group the regions.

The experimental results show that, the system can identify various types of structured objects in the image. The objects are identified based on general properties of real world objects without requiring object-specific knowledge.

The system can detect different types of objects with variety of complex backgrounds. The system is adaptive to the image content. The system can segment an object without the need of recognizing it.

In future, the system accuracy will be improved by combining region and edge information.

References

[1] C. Cheng, A. Koschan, D. L. Page, M. A. Abidi,"Outdoor Scene image segmentation based on background recognition and perceptual organization", IEEE Transaction on Image Processing Vol. 21, No. 3. March 2012.

[2] Ming-Ming Cheng, Guo-Xin Zhang, Niloy J. Mitra, Xiaolei Huang Shi-Min Hu,"Global contrast based salient region detection", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2011.

[3] E. Borenstein and E. Sharon, "Combining top-down and bottom-up segmentation", In Proc. IEEE Workshop Perceptual Org. Comput. Vis.,CVPR, 2004.

[4] X. Ren,"Learning a classification model for segmentation", in Proc. IEEE ICCV, 2003.

[5] P. Felzenszwalb, D. Huttenlocher,"Efficient graph-based image segmentation", Int. Journal of Computer Vision, Vol. 59, No. 2, Sep.2004.

[6] Rajeshwari Rasal, Prof. N.M.Shahane,"Structured Object Detection in Images Using Region Contrast & Perceptual Organization", In proceedings of cPGCON March 2014.

[7] Rajeshwari Rasal, Prof. N.M. Shahane,"Boundary Detection of Structured Objects", IJARCCE, Vol. 3, Issue 6, June 2014.

[8] J. B. Shi, J. Malik,"Normalized cuts and image segmentation", IEEE Transaction Pattern Anal. Mach. Intell., Vol. 22, No. 8, Aug. 2000.

[9] S. Gould, J. Rodgers, D. Cohen, G. Elidan, D. Koller, "Multi-class segmentation with relative location prior", Int. J. Computer Vision, Vol. 80, No. 3, Dec. 2008.

[10] J. D. McCafferty, Human and Machine Vision: Computing Perceptual Organization", Chichester, U.K.: Ellis Horwood, 1990.

[11] R. Mohan, R. Nevatia,"Perceptual organization for scene segmentation and description", IEEE Trans. Pattern Analysis & Machine Intelligence, Vol. 14, No. 6, Jun. 1992.

[12] D. R. Martin, C. C. Fowlkes, J. Malik,"Learning to detect natural image boundaries using local brightness, color, and texture cues", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 26, No. 5, pp. 530–549, May 2004.

[13] I. H. Jermyn, H. Ishikawa,"Globally optimal regions and boundaries as minimum ratio weight cycles", IEEE Trans. Pattern Analysis & Machine Intelligence, Vol. 23, Oct. 2001.

[14] B. C. Russell,"Using multiple segmentations to discover objects and their extent in image collections", In Proc. IEEE CVPR, 2006, Vol. 2, pp. 1605–1614.

[15] D. R. Martin, C. C. Fowlkes, J. Malik,"Learning to detect natural image boundaries using local brightness, color, and texture cues", IEEE Trans. Pattern Analysis & Machine Intelligence., Vol. 26, No. 5, pp. 530–549, May 2004.

[16] Z. L. Liu, D. W. Jacobs, R. Basri,"The role of convexity in perceptual completion: Beyond good continuation", Vis. Res., Vol. 39, No. 25, pp. 4244–4257, Dec. 1999

[17] C. Cheng, A. Koschan, D. L. Page, M. A. Abidi,"Scene image segmentation based on perceptual organization", In Proc. IEEE ICIP, 2009, pp. 1801–1804.

[18] L. Yang, P. Meer, D. J. Foran,"Multiple class segmentation using a unified framework over man-shift patches", In Proc. IEEE CVPR, 2007, pp. 1–8.

[19] S Jayaraman,"Image Segmentation", In Digital Image Processing: McGraw-Hill, 2010