ISSN:0975-9646

Neha Dabhi et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (6), 2012,5419-5423

A Review on Outdoor Scene Image Segmentation

Neha Dabhi[#1], Prof.HirenMewada[*2]

VTP Electronics & communication Dept., CHARUSA University, Changa,,Anand, India

Abstract—**Image segmentation is the process of partitioning or divisioning of an image into homogeneous and self consistent region which doesn't overlap with each other.The segmentation is based on color, texture, motion, depth, gray level etc. The level of detail in which the partitioning is carried out is depend on the application. Several general-purpose algorithms and techniques have been developed for image segmentation. This paper describes various segmentation techniques for outdoor scene images.**

Keywords—**Image segmentation, superpixels,signed distance function,Gestalt laws, Perceptual Orgnisation**

I. Introduction

Digital image processing means image is processed by digital computer[1]. digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the building-up of noise and signal distortion during processing.

Image segmentation is related with analysis of an image. In that the input is an image and the output is features extracted from the image[1].Natural scenes can be roughly divided into two categories:

1. Unstructured objects
2. Structured objects

Unstructured objects usually have nearly homogenous surfaces while structured objects usually consist of multiple parts with each part having distinct appearances (e.g. colors, textures).The common backgrounds in outdoor natural scenes are those unstructured objects like skies, roads, trees, grasses, etc. Generally the structured objects are the foreground objects and the unstructured objects are the background objects.

Fig 1: Classification of natural scene image a)Unstructured Objects b)Structured Objects Adapted from[2]

It is difficult to segment an outdoor scene image as it consist both the structured and unstructured objects. The structured object is the ultimate challenge as it consists multiple part with each part having different surface characteristics. Another challenge in the segmentation of the unstructured objects because they are not seen before.Image segmentation can be performed using one of the following approaches:

1. Top down approach
2. Bottom up approach

Top down approach uses prior knowledge about an object required.Then the segmentation is carried out based on some high level features such as its symmetry, proximity ,alignment of an object etc.. While in the Bottom up approach, the image is first segmented into regions and then the image regions that correspond to a single object are identified. Here prior knowledge of an object is not required.In the bottom up approach the image is first segmented based on some low level features such as color, texture, edges etc. The complexity in this approach is that an object may be segmented into numerous regions some of which may occlude with the background.

Borestein et al.[2] integrates the top down and bottom up segmentation techniques and constructs a classification map K(x,y). In the top-down approach, the requirement is to make K as close as possible to the initial top down classification map T. The bottom up constraint requires K to match the image structure, so that pixels within the homogeneous image regions, as defined by the bottom up process are likely be segmented together into either the figure or background part of the image. The advantage of this approach is that it provides a reliable confidence map indicating the regions of residual ambiguity with no additional computation cost. The problem is that the top down and bottom up approaches may conflict with each other.

II. Image Segmentation Techniques

The different types of image segmentation techniques for outdoor scenes can be classified as follows:

- Region based image segmentation
- Contour based approaches
- Boundary detection based
- Multiclass image segmentation
- Image segmentation based on perceptual organization

A. Region based approaches

The graph based image segmentation approach defines the boundaries between regions by measuring the dissimilarity between the neighboring pixels. Each pixel is equivalent to a node in the graph. The early Graph based method segments the images based on local features and fixed threshold. Thus the results were not accurate. Wu and Leahy[4] first introduced a minimum cuts criterion and global cost function in graph . But the weakness of it is :1.Method favors small components And 2.It fails to find global optima for these high dimension.To overcome this

Shi and Malik[5]proposed a new graph-theoretic criterionfor measuring thegoodness of an image partition known as the "normalized cut".The goal is partition the graph into a sets vertices (regions), such that the similarity within the region is high – and similarity across the regions is low.

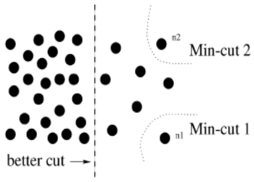
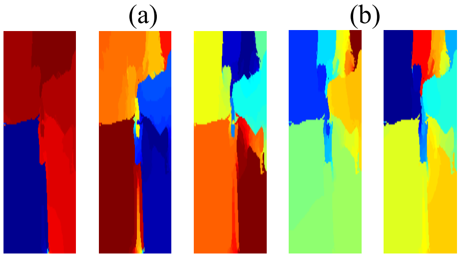


Fig 2 : A case where minimumcut gives a bad partition.

Ncut method organizes nodes into groups so that within the group the similarity is high and/or between the groups the similarity is low. This method is relatively robust and can be recursively applied to get more than two clusters. Each time the subgraph is partitioned that has the maximum number of nodes (random selection for tie breaking). When the bound on the number of clusters is reached or Ncut value exceeds some threshold, the process terminates.
The normalized cut algorithm can be explained as follows: First the image is represented as a weighted graph G=(V,E) and then compute the weight of each edge. Also summarize the values D and W where D is a N*N diagonal matrix and W is a N*N symmetrical matrix.Next solve the $(D - W)y = \lambda Dy$ for the eigenvector with the second smallest eigenvalue. Finally the graph is bipartitioned using the entries of the eigenvector. The basic idea behind this is big clusters have higher associativity and less ncuts. The computations are hard to compute and take much time to complete and also this approach works well only for relatively small images. Other problems are the high storage requirement and this approach is bias towards partitioning into equal segments. Fig 3 illustrates the image segmentation using Normalized cut approach. The weaknesses are : 1.it has problems with textured backgrounds and 2.The algorithm is able to extract the major components of scene while ignoring small intra component variations.

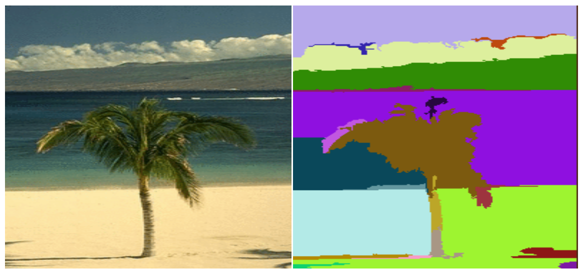
(a) (b)



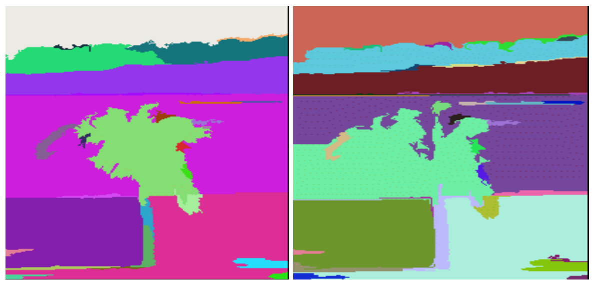
(c)

Fig 3: Image Segmentation using Normalized Cut approach: (a) Input Image (b) Segmentation of the image.(c) Eigen-vectors

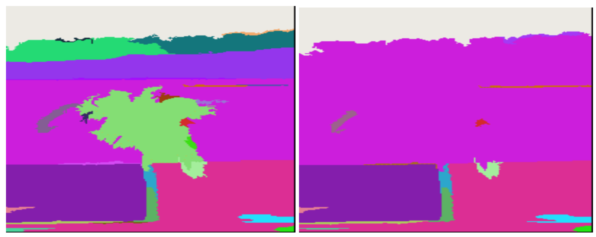
Felzenszwalb and Huttenlocher [6] proposed an efficient graph-based generic image segmentation algorithm. The success of this method is due to the adaptive thresholding. Kruskal's algorithm is used to generate the minimum spanning tree of data points. In that if any edges with length greater than a given hard threshold are detached. The joined components become the clusters in the segmentation. It has many advantages like it is Good for thin regions, Easy to control coarseness of segmentations by changing the parameter of it, Can include both large and small regions. And the disadvantage is it produces very large errors when the parameter are not well define.



(a) Segmentation parameters: sigma = 0.5, K = 500, min = 50.



(b) Improper value of σ



(c)Improper value of threshold

Fig 5: Image Segmentation using graph based method.(a) Input Image and segmented image (b)Resultant Images achieved by improper value of σ (c)Resultant Images achieved by improper value of threshold

B. Contour-based segmentation approach

In this method the segmentation is carried out based on Computer generated curve or model that will find the object boundaries. Zhu and Yuille[7] first used both boundary and region information within an energy optimization model. It is Bottom up segmentation method means segmentation is based on some low level features. Here the set of initial seeds are necessary to be placed correctly inside each homogeneous region to achieve good result. It uses Gradient Descent algorithm to find energy function thus it only finds local minima. To overcome this Jermyn and Ishikawa [8] proposed new technique. It uses two polynomial-time digraph algorithms for finding the global minima of energy. In this neither algorithm requires initialization. Here the Energy model depends on the boundary (intensity gradients), interior of the region (texture, homogeneity)The weakness of this algorithm is it can not capture structure context information. By incorporating the Gestalt laws it can be removed.

The figure shows a 256 × 256 pixel image and its resultant image. Note that the extracted boundaries are contrast reversed, i.e., for the left-hand-side boundary of the crater, the inside is lighter, while for the right hand side the inside is darker. In another figure the same energy detected a non-contrast-reversed boundary (176 × 256 pixel image). Adapted from [8]

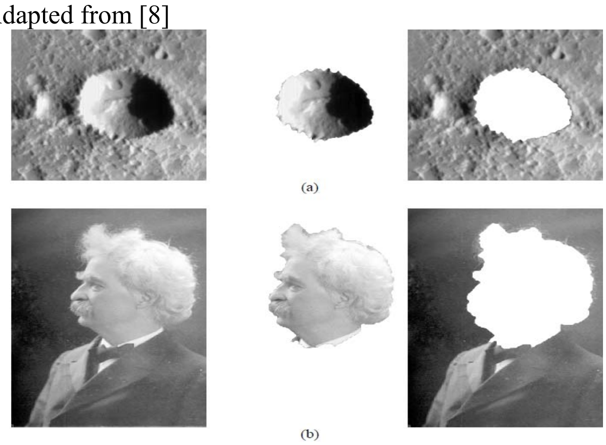


Fig 4 : Image segmentation through Contour based segmentation. Adapted from[8]

The other technique which uses the level set function for active contour segmentation method its capability of automatically handling changes in topology. In the traditional level set methods, the level set function is initialized to be a signed distance function (SDF) to its interface in order to prevent it from being too steep or flat near its interface, and because the existing re-initialization schemes were not able to re-initialize the level set function, if the initial level set function is significantly different from a signed distance function. Furthermore, it is difficult to decide when and how to apply the re-initialization. In addition, re-initialization is a very computationally expensive operation (consumes approximately 90% of total execution time). Thus, re-initialization is not only, time consuming operation, but it also affects the accuracy of final segmentation. For that the method is proposed [9] .

(a) (b) (c)

Image size=240 X 320

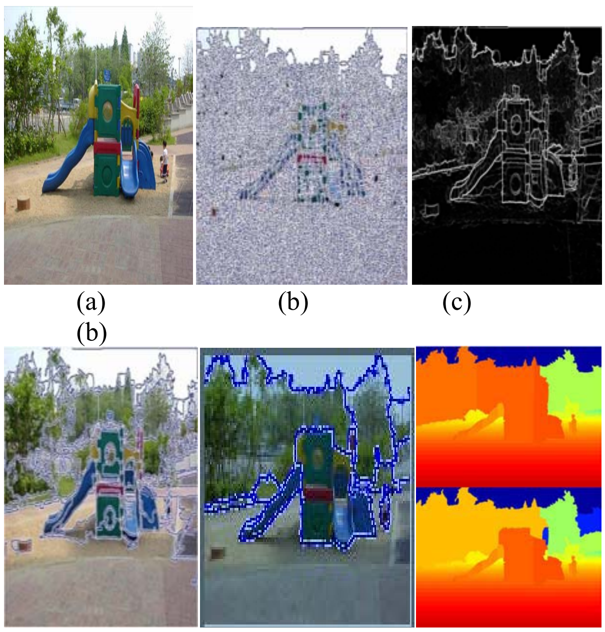
Time taken by the model= 1.85

Fig 6:(a) Original image with initial contour (b) Unsuccessful result obtained with standard C-V model [10] (c) Successful segmentation result obtained with proposed model [9] Adapted from[9]

The advantages are it generates continuous edge in compared to kernel based edge function. It can handle complex topology and noisy image can also be segmented which is not accurate in kernel based and edge based segmentation technique. The disadvantages are it fail to segment image having intensity inhomogeinity. It does not work well for medical images. The required initialization of contour which is near to the desired object. Computationally expensive due to re-initialization of the level cut function.

C.Boundary Detection approach

It is the statistical learning method for image segmentation.Martin et al.[11] proposed the technique which uses a large data set of human-labeled boundaries in natural images to train a boundary model. The model can then predict the possibility of boundary at each pixel based on a set of low-level cues like brightness, color and texture extracted from local image patches.The weakness of this algorithm is that 1.Object parts have wide variation of surface characteristics (e.g., brightness, color, texture etc.) 2.Fails to detect some true boundaries. To overcome this Hoiem et al. [12] the technique which is based on occlusion and the algorithm utilizes both 2D and 3D cues to estimate the surface characteristic and depth.The algorithm is for the occlusion recovery. For that the steps are as this : In this first the scene is over segmented into thousand of regions. In this gradually progress toward the final result by iteratively computing cues over boundaries, regions and produces the soft boundary map by performing interference over CRF model.At the end occlusion boundaries can be estimated.The weakness are: 1 .Suffers from the wide variation of the surface properties of object parts.2.problem of accurately detecting object boundaries in many cases.



(a) (b) (c)
(b)

(d) (e) (f)

Fig 7: Boundary detection approach (a)Original Image (b)Initial overSegmentation(c)Soft Boundary Map(d)Next Segmentation (e)Final result (f)depth estimates.Adapted from [12].

D. Multiclass image segmentation Or Semantic Image Segmentation

Unlike object recognition methods that aim to find a particular object multi-class image segmentation methods are aimed at concurrent multi-class object recognition and attempt to classify all pixels in an image[17].Multi-class image segmentation uses one of a number of classes (e.g., road, sky, water, etc) for pixel-labeling in an image.It first over-segment the image into superpixels (or small coherent regions) and then classify each region since classifying every pixel can be computationally expensive. Shotton et al. [13] described the need to label each pixel in the image with one of a set of predefined object class labels. In this approach, Shotton assigned a class label to a pixel based on a joint appearance, shape and context model. The aim of this approach is, given an image, the system should be capable of automatically partitioning it into semantically meaningful regions each labeled with a specific object class. For this a discriminative model for object class is learned incorporating texture, layout and context information efficiently. The learned model is then used for automatic visual understanding and semantic segmentation of images. This technique can model very long range contextual relationship extending over half the size of the image.

Gould et al. [14] proposed a superpixel-based conditional random field to learn the relative location offsets of categories. Unlike object recognition methods that intend to find a particular object, multi-class image segmentation methods are intended at concurrent multi-class object recognition and attempt to classify all pixels in an image. For each superpixel region, this method first extracts appearance (color and texture), geometry and location features. Then boosted classifiers are learnt over these features for each region class. Finally, a CRF or logistic model is learned using the output of the boosted classifiers as features. It does not distinguish between objects at different scales. Fig 8 shows the example results of simultaneous object class recognition and segmentation algorithm.

Fig 8: Example results of simultaneous Object Class Recognition and Segmentation algorithm. Adapted from [13]

E. Image Segmentation based on perceptual organization

Perceptual organization refers to a basic capability of the human visual system to obtain relevant groupings and structures from an image without having prior knowledge of the contents of the image. This technique utilizes Gestalt laws for image segmentation. Gestalt law describes A physical, biological, psychological, or symbolic configuration or pattern of elements so unified as a whole that its properties cannot be derived from a simple summation of its parts. The Weaknesses are : 1. Descriptive in nature so it needs to produce the Scientific model produced from it.2.Combine the various grouping factors because there is no systematic attempt has been made to categorize the relative importance of each Gestalt Laws of Organization. The advantages are : 1.Classification of the object without prior knowledge. 2. Detect various salient objects for different scenes.

The Gestalt psychologists summarized some underlying principles (e.g., proximity, similarity, continuity, symmetry, etc.) that lead to human perceptual grouping. Maire et al. [15] suggests a state-of-the-art solution for the problems related to finding contours (segmentation curves), and finding junction (points joined by multiple contours). The contours are found by combining the local and global features. The local cues are combined in a multi-scale oriented signal including brightness, color and texture gradients. The global information is considered to be in the first nine generalized eigenvectors, from which a signal is extracted with Gaussian directional derivatives at multiple orientations. The local and global information are then linearly combined, resulting in a globalized probability of boundary, which claims the top spot in the standard Berkeley segmentation benchmark. The procedure for contour detection can be briefly stated as:

- Collect Data Set of Human segmented images.
- Learn the local boundary model for combining brightness, color and texture.
- Global framework to capture closure, continuity.
- Detect and localize junctions.
- Integrate the low, mid and high-level information for grouping and figure-ground segmentation.

Chang et al. [1] explores detecting object boundaries in outdoor scene images based on some general properties of the real world number of structured object classes, there are only a few common background objects in outdoor scenes. These background objects have low visual variety and hence can be reliably recognized. After background objects are identified, it is roughly know where the structured objects are and delimit perceptual organization in certain areas of an image. For many objects with polygonal shapes, such as the major object classes appearing in the streets (e.g., buildings, vehicles, signs, people, etc.) and many other objects, this method can piece the whole object or the main portions of the objects together without requiring recognition of the individual object parts. In other words, for these object classes, this method provides a way to separate segmentation and recognition.

Fig 9: Example result of outdoor scene image segmentation using background recognition and perceptual organization. Adapted from [16]

Objects such as perceptual organization without depending on a priori knowledge of the object. It is well accepted that segmentation and recognition should not be separated and should be treated as an interleaving procedure. This method basically follows this scheme and requires identifying some background objects as a starting point. Compared to the large number of structured object classes, there are only a few common background objects in outdoor scenes. These background objects have low visual variety and hence can be reliably recognized. After background objects are identified, it is roughly know where the structured objects are and delimit perceptual organization in certain areas of an image. For many objects with polygonal shapes, such as the major object classes appearing in the streets (e.g., buildings, vehicles, signs, people, etc.) and many other objects, this method can piece the whole object or the main portions of the objects together without requiring recognition of the individual object parts. In other words, for these object classes, this method provides a way to separate segmentation and recognition.

III. CONCLUSION

This review highlights different segmentation techniques from all aspects with its pros and cons which are used in the segmentation of outdoor scene images . Boundary based methods are sensitive to noise and texture, and usually depend on good pre-processing .It gives good results with urban zones with high contrast. Region based methods have difficulty with transition zones. Region growing was less sensitive to texture (good for high resolution (1m) images). Multi-level techniques are the only way to get all objects without over-segmentation .In Contour analysis (e.g. edge detection) may be adequate for untextured images, but in a textured region it results in a meaningless tangled web of contour. Active Contour based segmentation is widely used in Segmentation of Soft Tissues in Medical Images. The other segmentation techniques are sensitive to noise and initial manual initialization are required so they does not give good results in addition to that it uses level set algorithm, which can automatically handle topology change. While in Multi-class image segmentation has significant advances in recent years through the combination of local and global features. In this object recognition and classification of all pixels in an image is carried out. It can capture structured context information efficiently. The another is human visual system based image classification called the perceptual organization. In that the prior knowledge of an object is not required and it can detect various salient objects for different scenes. In that gestalt law based image segmentation is carried out. The segmentation quality depends on the choice of the law and the number of the law which are incorporated for the segmentation. In future we can incorporate more gestalt laws into that and can improve the accuracy of the segmentation and also by solving the problem caused by strong reflection and over-segmentation.

REFERENCES

[1] *Rafel C. Gonazalez , Richard E.Woods, DigitalImage Processing,Pearson* Publication.

[2] G. Griffin, A. Holub and P. Perona, "Caltech-256 object category dataset," Technical report Caltech, 2007.

[3] E.Borenstein and E. Sharon, "Combining top-down and bottom-up segmentation," in *Proc. IEEEWorkshopPerceptual Org. Comput. Vis.,CVPR*, 2004, pp. 46–53.

[4] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: theory and its Applications to image segmentation," IEEE Trans. PAMI, 15(11), pp. 1,101-1,113, 1993.

[5] J. B. Shi and J. Malik, "Normalized cuts and image segmentation", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[6] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation,", *IJCV,*vol. 59, no. 2, 2004.

[7] S.C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing and Bayes/MDL for multi-band image segmentation," IEEE Trans. PAMI, 18(9), pp. 884-900, 1996.

[8] I.H. Jermyn, and H. Ishikawa, "Globally optimal regions and boundaries as minimum ratio weight cycles," IEEE Trans. PAMI, *23(10):1075-1088, 2001.*

[9] Rahul Patel ,Hiren Mewada and Suprava Patnaik, "Fast and Regularization less Active Contour" *International Journal of Computer Applications (0975 – 888)Volume 47– No.6, June 2012*

[10] Lankton, S., and Tannenbaum, A. 2008. Localizing region based active contour. IEEE Transactions on Image Processing, Vol. 17, No. 11, pp. 2029-2039.

[11] D.R. Martin, C.C. Fowlkes, and J. Malik, " Learning to detect natural image boundaries using local brightness, color, and texture cues," IEEE Trans. PAMI*, 26(5) pp. 530-549, 2004*

[12] D. Hoiem, A. N. Stein, A. A. Efros, and M. Hebert, "Recovering occlusion boundaries from a single image," in *Proc. IEEE ICCV, 2007,* pp. 1–8.

[13] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context," *Int. J. Comput. Vis.*,vol. 81, no. 1, pp. 2–23, Jan. 2009.

[14] S. Gould, R. Fulton, and D. Koller, "Decomposing a scene into geometric and semantically consistent regions," in *Proc. IEEE ICCV*, 2009, pp. 1–8.

[15] M.Maire, P. Arbelaez, C. C. Fowlkes, and J.Malik, "Using contours to detect and localize junctions in natural images," in *Proc. IEEE CVPR*, 2008, pp. 1–8.

[16] C. Cheng, A. Koschan, D. L. Page, and M. A. Abidi, "Outdoor scene image segmentation based on background recognition and perceptual organization," in *Proc. IEEE Trans*, vol.21, no.3,pp. 1007–1019, March2012

[17] Stephen Gould, Jim Rodgers, David Cohen, Gal Elidan, Daphne Koller ," Multi-Class Segmentation with Relative Location Prior," International Journal of Computer Vision DOI 10.1007/s11263-008-0140-x