International Journal of Nuclear Science and Technology

https://trpub.org/index.php/IJNST

Vol 1, Issue 1, 2026

A METHOD FOR SEPARATING THE IMAGES OF OUTSIDE SCENES USING PERCEPTUAL ORGANIZATION AND BACKGROUND RECOGNITION

[1]G Bruhaspathi, [2]E. Soumya, [3]P. Devasudha, [4]Dr. M. Vadivukarassi

[1,2,3]Assistant Professor, [4]Associate Professor
Department of Computer Science and Engineering,
St. Martin's Engineering College, Secunderabad, Telangana, India

ABSTRACT

Using color and texture information, a new outdoor scene image segmentation method based on background recognition and perceptual organization is utilized to identify background items including the ground, sky, and plants. A perceptual organization model was created for structurally challenging objects, which typically have several constituent parts. This model is able to identify the non-accidental structural relationships between the constituent parts of the structured objects and, as a result, group them together appropriately without requiring prior knowledge of the particular objects. According to the experimental results, the suggested method obtained correct segmentation quality on a variety of outdoor natural scene contexts and outperformed two cutting-edge picture segmentation techniques on two difficult outdoor databases (the Berkeley segmentation data set and the Gould data set).

Keywords: Boundary energy, image segmentation, perceptual organization.

I. INTRODUCTION

The term digital image refers to processing of a two dimensional picture by a digital computer. In a broader context, it implies digital processing of any two dimensional data. A digital image is an array of real or complex numbers represented by a finite number of bits. An image given in the form of a transparency, slide, photograph or an X- ray is first digitized and stored as a matrix of binary digits in computer memory. This digitized image can then be processed and/or displayed on a high-resolution television monitor. For display, the image is stored in a rapid-access buffer memory, which refreshes the monitor at a rate of 25 frames per second to produce a visually continuous display.

The Image Processing modeled in the form of Multidimensional Systems. An image may be defined as a two-dimensional function, f(x, y) where x and y are spatial coordinates, and the amplitude off at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the amplitude values of f are all finite, discrete quantities, we call the image a digital image. The field of digital image processing refers to processing digital images by means of a digital computer. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements, pels, and pixels. Pixel is the term used most widely to denote the elements of a digital image.

II. LITERATURE SURVEY

T. Malisiewicz and A. A. Efros introduce a smethod, “Improving spatial support for objects via multiple segmentations”,Sliding window scanning is the dominant paradigm in object recognition research today. But while much success has been reported in detecting several rectangular- shaped object classes (i.e. faces, cars, pedestrians), results have been much less impressive for more general types of objects.

Several researchers have advocated the use of image segmentation as a way to get a better spatial support for objects. In this paper, our aim is to address this issue by studying the following two questions: 1) how important is good spatial support for recognition? 2) can segmentation provide better spatial support for objects? To answer the first, we compare recognition performance using ground-truth segmentation vs. bounding boxes. To answer the second, we use the multiple segmentation approach to evaluate how close real segments can approach the ground-truth for real objects, and at what cost. Our results demonstrate the importance of finding the right spatial support for objects, and the feasibility of doing so without excessive computational burden.

E. Borenstein and E. Sharon proposes “Combining top-down and bottom- up segmentation,” In this method combine bottom-up and top-down approaches into a single figure-ground segmentation process. This process provides accurate delineation of object boundaries that cannot be achieved by either the top-down or bottom-up approach alone. The top-down approach uses object representation learned from examples to detect an object in a given input image and provide an approximation to its figure-ground segmentation. The bottom-up approach uses image-based criteria to define coherent groups of pixels that are likely to belong together to either the figure

or the background part. The combination provides a final segmentation that draws on the relative merits of both approaches: The result is as close as possible to the top-down approximation, but is also constrained by the bottom-up process to be consistent with significant image discontinuities. They construct a global cost function that represents these top-down and bottom-up requirements. We then show how the global minimum of this function can be efficiently found by applying the sum-product algorithm. This algorithm also provides a confidence map that can be used to identify image regions where additional top-down or bottom-up information may further improve the segmentation. Our experiments show that the results derived from the algorithm are superior to results given by a pure top- down or pure bottom-up approach. The scheme has broad applicability, enabling the combined use of a range of existing bottom- up and top-down segmentations.

III. OUTDOOR SCENE IMAGE SEGMENTATION

In computer vision, segmentation is the process of partitioning a digital image into multiple segments (sets of pixel also known as super pixels). The goal of image segmentation is to cluster pixels into salient image regions, i.e., regions corresponding to individual surfaces, objects, or natural parts of objects. In general, objects in outdoor scenes can be divided into two categories, namely, unstructured objects (e.g., sky, roads, trees, grass, etc.) and structured objects (e.g., cars, buildings, people, etc.). Unstructured objects usually comprise the backgrounds of images. The background objects usually have nearly homogenous surfaces and are distinct from the structured objects in images``. The challenge for outdoor segmentation comes from the structured objects that are often composed of multiple parts, with each part having distinct surface characteristics (e.g., colors, textures, etc.).Without certain knowledge about an object, it is difficult to group these parts together.
In outdoor scene image segmentation based on background recognition and perceptual organization objects appearing in natural scenes can be roughly divided into two categories: homogenous surfaces, whereas structured objects typically consist of multiple constituent parts, with each part having distinct appearances (e.g., color, texture, etc.). The common backgrounds in outdoor natural scenes are those unstructured objects such as skies, roads, trees, and grasses. These background objects have low visual variability and are distinguishable from other structured objects in an image. For instance, a sky usually has a uniform appearance with blue or white colors; a tree or a grass usually has a textured appearance with green colors. Therefore, these background objects can be accurately recognized solely based on appearance information.

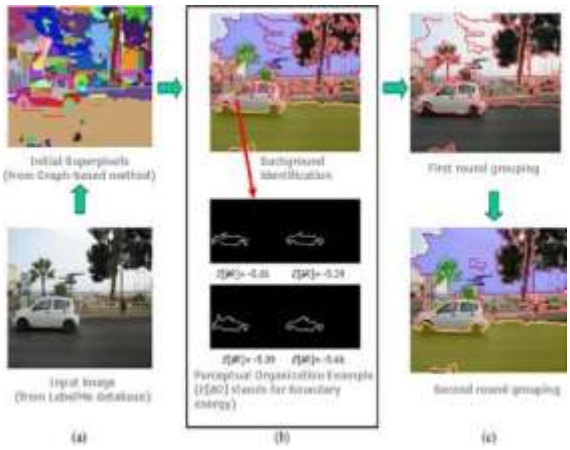

Figure: 1 Segmentation Process Based on Background Identification and POM

IV. BLOCK DIAGRAM

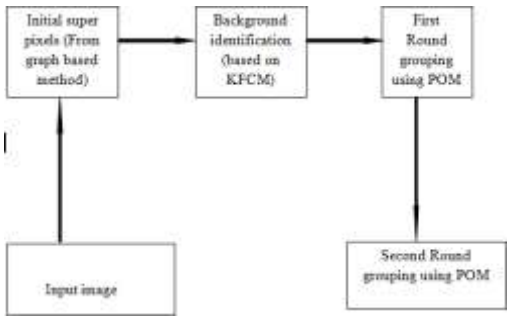

Figure:2 Block diagram of Segmentation based on KFCM and POM Superpixel Generation Using Segment Merging Method

Superpixel image segmentation techniques have been recognized as useful for segmenting digitized images into discrete pixel clusters to aid and enhance further image analysis. The discrete pixel clusters, called superpixels, represent contiguous groups of digital image pixels sharing similar characteristics, such as color, texture, intensity, etc. Because the superpixels of a segmented image represent clusters of pixels sharing similar characteristics, further image processing tasks may be carried out on the superpixels themselves, rather than the individual pixels. Thus, superpixel segmentation serves to lower the memory and processing requirements for various image processing tasks, which in turn may permit either greater throughput or more in-depth image analysis. Tasks that may benefit from superpixel segmentation include feature extraction, object recognition, pose estimation, and scene classification tasks. Conventional superpixel segmentation techniques are not well suited to process images that have a high texture content. Image texture is a function of spatial variation in image pixel intensity. Highly textured images are notable in that the colors are not highly varied, but variations in pixel intensity produce well-defined images. Cloth, tree bark, and grass are examples of highly textured images. It is also desirable to provide a superpixel segmentation method that is computationally efficient, and able to produce compact superpixel segmentations. A compact superpixel segmentation is one in which the superpixel size is kept as uniform as possible, and the total number of superpixels remains small variability and, in most cases, are distinguishable from other structured objects in an image. For in-stance, a sky usually has a uniform appearance with blue or colors; a tree or a grass usually has a textured appearance with green colors. Therefore, these background objects can be accurately recognized solely based on appearance information. The key for this method is to use textons to represent object appearance information. The term texton is first presented in for describing human textural perception. The whole textonization process proceeds as follows: First, the training images are converted to the perceptually uniform CIE color space. Then, the training images are convolved with a 17-D filter bank which consists of Gaussians at scales 1, 2 and 4.The derivatives of Gaussians at scales 2 and 4; and Laplacians of Gaussians at scales 1, 2, 4, and 8. The Gaussians are applied to all three color channels, whereas the other filters are applied only to the luminance channel. By doing so, we obtain a 17-D response for each training pixel. The 17-D response is then augmented with the CIE, channels to form a 20-D vector. This is different from because we found that, after augmenting the three color channels, we can achieve slightly higher classification accuracy. Then, the Euclidean-distance Kernelized Fuzzy C Means Algorithm (KFCM) is performed on the 20-D vectors collected from the training images to generate cluster centers. These cluster centers are called textons. Finally, each pixel in each image is assigned to the nearest cluster center, producing the texton map. After this textonization process, each image region of the training images is represented by a histogram of textons. We then use these training data to train a set of binary Ad- boost classifiers to classify the unstructured objects (e.g., skies, roads, trees, grasses, etc.).

CIELAB

CIE L*a*b* (CIELAB) is the most complete Color space specified by the International Commission on Illumination (ICI). It describes all the color visible to the human eye and was created to serve as a device-independent model to be used as a reference. The three coordinates of CIELAB represent the lightness of the color (L* = 0 yields black and L* = 100 indicates diffuse white; specular white may be higher), its position between red/magenta and green (a*, negative values indicate green while positive values indicate magenta) and its position between yellow and blue (b*, negative values indicate blue and positive values indicate yellow). The asterisk (*) after L, a and b are part of the full name, since they represent L*, a* and b*, to distinguish them from Hunter's L,

a, and b, described below.
Since the L*a*b* model is a three- dimensional model, it can only be represented properly in a three-dimensional space. Two- dimensional depictions include chromaticity diagrams: sections of the color solid with a fixed lightness. It is crucial to realize that the visual representations of the full gamut of colors in this model are never accurate; they are there just to help in understanding the concept. Because the red-green and yellow- blue opponent channels are computed as differences of lightness transformations of cone responses, CIELAB is a chromatic value color space.
Fuzzy C-Means Clustering
In fuzzy clustering, each point has a degree of belonging to clusters, as in fuzzy logic rather than belonging completely to just one cluster. Thus, points on the edge of a cluster, may be in the cluster to a lesser degree than points in the center of cluster. An overview and comparison of different fuzzy clustering algorithms is available. Any point x has a set of coefficients giving the degree of being in the kth cluster wk(x). With fuzzy c- means, the centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster:

The degree of belonging, wk(x), is related inversely to the distance from x to the cluster center as calculated on the previous pass. It also depends on a parameter m that controls how much weight is given to the closest center. The fuzzy c-means algorithm is very similar to the k-means algorithm:
Choose a number of clusters.
Assign randomly to each point coefficients for being in the clusters.
Repeat until the algorithm has converged (that is, the coefficients' change between two iterations is no more than , the given sensitivity threshold) :
Compute the centroid for each cluster, using the formula above.
For each point, compute its coefficients of being in the clusters, using the formula above.
The algorithm minimizes intra-cluster variance as well, but has the same problems as k-means; the minimum is a local minimum, and the results depend on the initial choice of weights. Fuzzy c-means has been a very important tool for image processing in clustering objects in an image. In the 70's, mathematicians introduced the spatial term into the FCM algorithm to improve the accuracy of clustering under noise.

Kernelized Fuzzy C Means Algorithm

The kernel methods are one of the most researched subjects within machine learning community in the recent few years and have widely been applied to pattern recognition and function approximation. The main motives of using the kernel methods consist of: (1) inducing a class of robust non-Euclidean distance measures for the original data space to derive new objective functions and thus clustering the non- Euclidean structures in data; (2) enhancing robustness of the original clustering algorithms to noise and outliers, and (3) still retaining computational simplicity. The algorithm is realized by modifying the objective function in the conventional fuzzy c-means (FCM) algorithm using a kernel- induced distance instead of Euclidean distance in the FCM, and thus the corresponding algorithm is derived and called as the kernelized fuzzy c-means (KFCM) algorithm, which to be more robust than FCM. In FCM, the membership matrix U is allowed to have not only 0 and 1 but also the elements with any values between 0 and 1, this matrix satisfies the constraints:

$$\sum_{i=1}^{C} u_{ij} = 1, \forall j = 1,\ldots, N$$

Since the K-means method aims to minimize the sum of squared distances from all points to their cluster centers, this should result in compact clusters. We use the intra-cluster distance measure, which is simply the median distance between a point and its cluster centre. The equation is given as:

$$\text{intra} = midean(\sum_{i=1}^{C}\sum_{x\in c_i}\|x - v_i\|^2)$$

AdaBoost generates and calls a new weak classifier in each of a series of rounds. For each call, a distribution of weights is updated that indicates the importance of examples in the data set for the classification. On each round, the weights of each incorrectly classified example are increased, and the weights of each correctly classified example are decreased, so the new classifier focuses on the examples which have so far eluded correct classification.

PERCEPTUAL ORGANIZATION METHOD

Perceptual Organization Model (POM) is used to detect a boundary. The POM quantitatively incorporates a list of Gestalt laws and therefore is able to capture the no accidental structural relationships among the constituent parts of a structured object. With this model, we are able to detect the boundaries of various salient structured objects under different outdoor environments. To developed a POM first pick one part and then keep growing the region by trying to group its neighbors with the region. The process stops when none of the region's neighbors can be grouped with the region. To achieve this, we develop a measurement to measure how accurately a region is grouped. The region goodness directly depends on how well the structural relationships of parts contained in the region obey Gestalt laws.

V. RESULTS

Our 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, 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, people, etc.) and many other objects.our method can piece the whole object or the main portions of the objects together without requiring recognition of the individual object parts.Background object such as skies, roads, trees, grasses, etc detected by kernelized fuzzy c means clustering algorithm.

VI. CONCLUSION

A novel method for image segmentation algorithm for outdoor natural scenes. Our main contribution is that develop a POM. It is well accepted that segmentation and recognition should not be separated and should be treated as an interleaving procedure. Our 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, 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, our 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, our method provides a way to separate segmentation and recognition. This is the major difference between our method and other class segmentation methods that require recognizing an object in order to segment it. This paper shows that, for many fairly articulated objects, recognition may not be a requirement for segmentation. The geometric relationships of the constituent parts of the objects provide useful cues indicating

VII. REFERENCES

[1]E. Borenstein and E. Sharon, "Combining top-down and bottom-up segmentation," in Proc. IEEEWorkshop Perceptual Org. Comput. Vis.,CVPR, 2010,

[2]T. Cour, F. Benezit, and J. B. Shi, “Spectral segmentation with multi scale graph decomposition,” in Proc. IEEE CVPR, 2005, vol. 2, pp.1124–1131.

[3]P. Felzenszwalb and D. Huttenlocher, “Efficient graph-based image segmentation,” Int. J. Comput. Vis., vol. 59, no. 2, pp. 167– 181, Sep.2004.

[4]S. Gould, R. Fulton, and D. Koller, “Decomposing a scene into geometric and semantically consistent regions,” in Proc. IEEE ICCV,2002, pp. 1–8.

[5]D. Lowe, Perceptual Organization and Visual Recognition. Dordrecht,The Netherlands .

[6]T. Malisiewicz and A. A. Efros, “Improving spatial support for objects via multiple segmentations,” in Proc. BMVC, 2011.

[7]M.Maire, P. Arbelaez, C. C. Fowlkes, and J.Malik, “Using contours to detect and localize junctions in natural images,” in Proc. IEEE CVPR,

[8]D. Martin, C. Fowlkes, D. Tal, and J.Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in Proc. IEEE ICCV,2000S, vol. 2, pp. 416–423.

[9] J. D. McCafferty, Human and Machine Vision: Computing Perceptual Organization. Chichester, U.K.: Ellis Horwood, 2004.

[10] R. Mohan and R. Nevatia, “Perceptual organization for scene segmentation anddescription,” IEEE Trans. Pattern Anal. Mach. Intell., vol.14, no. 6, pp. 616–635, Jun.

[11] C. Pantofaru, C. Schmid, and M. Hebert, “Object recognition by integratingmultiple image segmentations,” in Proc. ECCV, 2008, pp.481–494.

[12]X. Ren, “Learning a classification model for segmentation,” in Proc.IEEE ICCV, 2009, vol.5.