Outdoor Dynamic Scene Video Object Segmentation Using SRM Algorithm

A.Niranjil Kumar
Department of ECE
P.S.R.Rengasamy College of Engineering for Women
Sivakasi Tamilnadu, India
niranjil.publications@gmail.com

Dr.C.Sureshkumar
Department of CSE
J.K.K.Nattraja College of Engineering and Technology
Komarapalayam, Namakkal, Tamilnadu, India

Abstract—In video surveillance, image segmentation in outdoor scenes is a most important and complex task. A novel approach for video object segmentation in outdoor environment is described by using SRM (Statistical Region Merging) algorithm in this paper. Here we are going to identify both structured (e.g. persons, buildings, car, etc.) and unstructured background objects (e.g. sky, road, grass, etc.) which are containing the some characteristic based on color, intensity and texture in sequence. Our main objective of this work is to solve the over segmentation of objects while segmenting images in outdoor environment in a video surveillance system. This work is divided into four modules; preprocessing, bottom up segmentation, boundary detection and region merging. In pre-processing the input image is converted into CIE (Commission Internationale d'eclairage) color space technique. Bottom-up segmentation process is used to capture the structured and unstructured image characteristics. Another process is the Ada boost classifier which is used to classify the background objects in outdoor environment scenes. Then the contour maps are used to detect the boundary energy. Finally the statistical region merging provides the groupings of images to identify the computer vision.

Keywords-component; Video Surveillance; Image Segmentation; Boundary Detection; Region Merging; Background Subtraction

I. INTRODUCTION

Many structures have a boundary that shows certain level of bilateral regularity, a property that has been applied to resolve many computer-vision tasks. A main and important challenge in imaging is vision grouping or image segmentation. The purpose of grouping and segmentation is to recognize and describe automatically, the salient objects in an image. Now-a-days image segmentation is measured to be a basic problem of the computer vision. It refers to the operation of partitioning an image into module or into separate objects. It is equal to partitioning the set of pixels forming the image, or clustering the image pixels. High quality clustering is commonly described by multiple attributes. The objective is to take out desirable structural boundaries from real images and to solve the over segmentation problem.

Generally the outdoor scenes can be divided into two groups, structured objects and unstructured objects.The background objects usually have homogenous surfaces and are dissimilar from the structured object in images. Currently many approaches are based on different methods that have achieved by high accuracy in recognizing these background object classes [7] [10]. In order not to contain certain knowledge about an object in outdoor scene image segmentation the objects are difficult to group these parts together. These difficult object specific models are studied [2] [3] [8]. The spatial support of objects decision is done by using segmentation algorithms and probability of unnecessary computational weight. Generally image processing is carried out to detect the low level features [7]. The goal of the Label Me project was to collect a large set of images and to research the computer graphics. These projects are comparing the dataset objects and perform the detection, recognition [9] [7]. The gestalt law emphasize to perceive objects as well as organized patterns rather than to separate the component parts (e.g. proximity, similarity, continuity, symmetry, closure etc.). The main role of perceptual organization is convexity because the real world objects are tends to have convex shapes [2] [3]. The gestalt law is followed for non accidentalness images, in which the structures are most likely to produce by an object [7]. To detect boundaries of the objects in the natural environment is the salient feature [8]. The multi-class segmentation is to discover image collections without prior knowledge [3] [1] [8]. The SVM (Support Vector Machines) classifiers are used to give the probability results of multi class objects. [6] [1]. A large number of object classes are classified by using Ada boost classifier that is based on textons and layout filters. It can be invariant to view the object classes in both structured and unstructured images [10]. The segmentation method is based on region merging technique. It provides good results, with high time and space complexity [11]. The testing of image databases consists of a wide variety of outdoor scenes and object classes.

The proposed system of Chang Cheng and Andreas Koschan [12] based on background recognition and perceptual organization (POM) to indentify background object evaluate the huge number of structured object classes, there are only a small number of common background objects in outdoor scenes. These background objects have short visual variety and hence they can be reliably recognized. As a result of the POM could not detect any spatial relationship between the small parts and hence could not piece them together. Ali et al proposed a system based on H-minima Transform and Region Merging technique for medical image segmentation [13].Watershed transform is a normally used image segmentation technique. The major problem with this segmentation method is that over-segmentation and its sensitivity to noise. In this paper the over-segmentation problem is defeat by combing pre-processing and post-processing procedures along with watershed transform. Sheng Liu Rui Cao et al proposed a novel quality evaluating method for over-segmentation approaches using real-time boundary information [14]. They explore the segmentation act of four state-of-the-art unsupervised over-segmentation techniques by estimating the quality of over-segments. To achieve this goal, they intend a novel real-time evaluating framework; this method is integrating the classical goodness methods. Our proposed new difference method based on real-time boundary removal, which is able to suppress the minor edges found within a homogeneous region, while being clever to locate the important edges in real-time.

Here we are going to use SRM for outdoor scene video object segmentation. Assume that a bottom-up method is employed to segment an image into uniform patches, and then the majority structured objects should be over segmented to multiple patches. After the background patches are spotted in the image, the greater part of the staying image patches correspond to the constituent parts of structured objects. POM (Perceptual Organization Method) refers to ability of the vision system to arrange features noticed in images based on perceptual phenomenon, Gestaltic law and viewpoint consistency. It is a computational structure for pre attentive perceptual organization. Outdoor segmentation obtained from the structured objects that are commonly having multiple parts, with each part having different surface characteristics. Without sure knowledge about an object, it is complicated to group these parts together. One aim is to get quantitative and objective measures of these grouping laws. The Gestalt laws are in descriptive forms. Therefore, one needs to quantify them for scientific use. Another challenge consists of finding a way to combine the various grouping factors since object parts can be attached in many different ways. Under different situations, different laws may be applied. Therefore, a perceptual organization system needs combining numerous Gestalt laws. We use Normalized cut method to provide the segmented and boundary detected objects. Finally the SRM used recognizes the each and every object from the outdoor scenes. The experimental datasets are achieved accurate segmentation quality on various outdoor natural scenes. The outdoor scene background objects such as both structured and unstructured objects are identified. These are based on the gray scale information that is color and intensity.

II. Overview of Proposed Method

In this proposed work the over segmentation of images while doing image segmentation in outdoor scene is rectified in an efficient way by using SRM algorithm. The low computational difficulty and properties are very feasible for real-time image segmentation process. The proposed system pre-processes image by using the SRM algorithm to form segmented regions that save the wanted discontinuity characteristics of the image. Then the segmented regions are signified by using the graph structures, and then the Normalized cut method is used to execute globally optimized clustering. As the number of the segmented regions is smaller than that of the actual image pixels, the proposed method permits a low-dimensional image clustering with significant reduction of the difficulty matched up to conventional graph partitioning methods that are directly applied to the image pixels. In addition, the image clustering using the segmented regions, instead of the Image pixels, also reduces the concern to noise and results in enhanced Image segmentation performance. The bottom up segmentation methods utilizes low level features such as colors, textures, and edge to decay an image into uniform regions. The proposed work is divided into 4 modules; pre-processing, bottom-up segmentation, boundary detection and region merging.

Bottom up segmentation can be divided into two categories namely region based and contour based approaches. Edge detection is a very significant and an important feature-extraction technique that has been extensively used in computer vision applications. The fundamental scheme of most existing edge detectors is to place some local object-boundary information in an image by threshold the pixel-intensity variation map. Furthermore, to pass up some inappropriate partitioning when allowing for every region as only one graph node, we build up an advanced segmentation strategy using multiple child nodes for each region. The lead of the proposed method is examined and established through a large number of experiments using natural scenes. To achieve region based and boundary based measurement, we collect a set of real images each of which consists of an unambiguous foreground salient object and a noisy background. In these images, the ground truth object boundary can be unambiguously extracted by manual processing, which enables an objective and quantitative measurement of the boundary-detection performance. In our framework a major component is to obtain a reliable algorithm for detecting a salient closed boundary from the edge-detection method results.

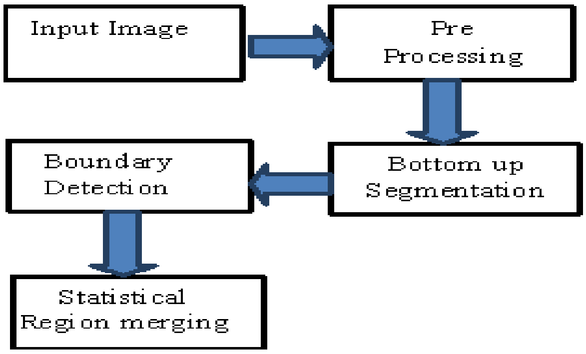

Fig. 1 Flow Chat of Proposed Method

III. PRE-PROCESSING

In pre-processing input video frame is converted as CIE (Commission Internationale d'eclairage) at lab colour space (Fig.3). It contains the same characteristics based on L-lightness, a-red/green value and b-yellow/blue value. L*a*b* can clearly separate gray-scale information from color information. It was designed so that the Euclidean distance in this space corresponds reasonably well with perceived difference between colors. Because of this property the l, a, b colour space are perceptually uniform.

Fig. 2 Input Image

Fig. 3 RGB Color Conversion Image

The CIE device-independent colour spaces are processed by using Matlab image processing toolbox. This process is depends on three characteristics brightness, intensity and contrast as shown in Fig4, Fig5 and Fig6 respectively. As a corollary, L*values relate linearly to human perception of brightness and a*value represent an intensity, b*represent a contrast. CIE specification that efforts to create the luminance scale more perceptually uniform L* is a nonlinear scaling of L, normalized to a reference white point.

Fig. 4 Brightness Image

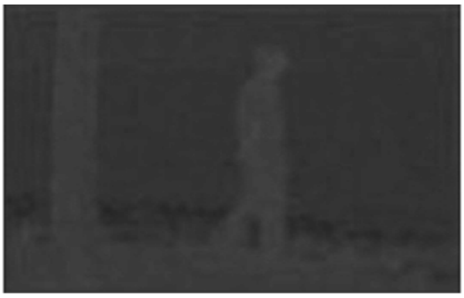

Fig. 5 Intensity Image

Fig. 6 Contrast Image

IV. BOTTOM-UP SEGMENTATION

The images are segmented as first level in the bottom up approach to identify the background object as shown in Fig. 7. We need to first segment an image into regions so that each region approximately corresponds to an object part. In our implementation, we make use of Felzenszwalb and Huttenlocher's approach to generate initial super pixels for an outdoor scene image. We choose this method because it is very efficient and the result of the method is comparable to the mean-shift algorithm. The image is modelled as a weighted, undirected graph. Usually a pixel or a group of pixels are related with nodes and edge weights describe the similarity between the neighbourhood pixels. Then the image is partitioned according to a criterion designed to representation clusters. Each partition of the pixels output from these algorithm is considered an object segment in the image.

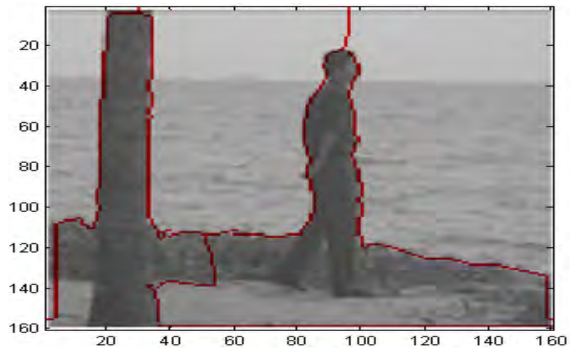

Fig. 7 Segmented Image

V. BOUNDARY DETECTION

Boundary detection is a very important feature-extraction method that has been extensively used in image processing applications. The fundamental scheme of most existing boundary detectors is to place some local object-boundary information in an image by thresholding and skeletonizing the pixel-intensity variation map.

A two dimensional isotropic Gaussian role is described as,

$$G\sigma\ (x,y) = (\left(\frac{1}{\sqrt{2\pi\sigma}}\right) \exp\ [-\frac{x^2+y^2}{2\sigma^2}]) \qquad (1)$$

First derivative of Gaussian along the x axis is given by,

$$GD\sigma\ (x,y) = \frac{\partial G\sigma(x,y)}{\partial x} = -\frac{x}{\sigma^2 G\sigma}(x,y) \qquad (2)$$

Mainly boundary detection process identifies and locates sharp discontinuities in an image. There are extremely large numbers of edge detection operators available in the algorithm and each designed to be sensitive types of edge both horizontal and vertical.

Here it finds the boundary of the given image. It is based on the breath first search strategy. Instead of finding the region with the global minimum boundary energy, the algorithm tries to find a region with the local minimum boundary energy. The images are separated by image patch and are given to the boundary detection. It will measure boundary energy of the combinations of and its immediate neighbors. It stops when no combination of and a single neighbors have smaller boundary energy. Then it tests the combination of a pair of connected neighboring regions as shown in Fig. 8.

Fig. 8 Boundary Detection in Image

VI. Region Merging

Region growing and region merging techniques work with a statistical test to choose the merging regions. A novel model of image generation and the segmentation approach is proposed. We approached segmentation problem by finding boundaries between regions based on cutoffs in intensity levels. Segmentation was accomplished via thresholds based on distribution of pixel properties. This technique those is used to find the regions directly.

The statistical region merging is based on a module of image generation that captures the idea of formulating image approaches usually work with a statistical regions and merging techniques. It contains the merging theorem that can be used to solve the over segmentation and strong reflection. A region merging algorithm has to conclude a good balance between preserving the perceptual units and the possibility of over merging for the remaining region [4]. This algorithm provides the homogeneity property in optimal regions. The every color channel makes available the same expectation of statistical pixels inside a region that means all distributions linked to each can be different as long as the homogeny property is satisfied. It is a quick and strong algorithm to segment an image into regions of similar intensity or color. The algorithm is used to evaluate the values with in a region and group together based on the merging criteria. It can correctly separate the regions and provides the original images which have clear edges. Based on the different colors we can identify both structured (e.g. persons, buildings, car, etc.) and unstructured background objects (e.g. sky, road, grass, etc.). Here the person (object) is segmented properly without over segmentation.

A. Technical analysis:

Statistical Region merging algorithm based on two essential components; merging predicate and order in merging, these are used to improve the performance of that algorithm.

B. Order in merging:

It refers to an order to be followed to check the merging predicate. When any test between two true regions is performed, then all tests should have been performed previously inside each of the true regions. The observed image I, there is coupled with adjacent pixels in connectivity. The simplest sort function f is defined as

$$f(p,p') = \max\ |Na(pa) - Na(pa')| \qquad (3)$$

Where $Na(pa) is\ t$he observed mean of the region defined by the set of points in channel a. That are within manhattans distance $\leq \boldsymbol{\nabla}$ to p and that are closer to p than p'. Merging predicate used to improve the speed and quality of segmentation.

C. Merging predicate:

It refers how to merge undetermined region that has used in theorem. We propose a new method for solving the perceptual grouping problem in computer vision, rather than spotlighting on local features and their consistencies in the image data. The proposed approach aims at extracting the global impression of an image; we use video image segmentation as a graph partitioning problem and propose global criterion and normalized cut criterion to measure the total variation between the different groups and the total affinity within the groups. We show that an efficient computation technique based on the generalized eigen value problem can be used to optimize this criterion. We have used this approach to segment static images, as well as motion sequences and the experimental results are very encouraging. Similarity measure between elements in each region is quite high, and the similarity between elements across regions is low.

Similarity: element × element = R^+ (4)

Intensity

$$aff(x, y) = e^{-\frac{||I(x)-I(y)||^2}{2\sigma 1^2}} \quad (5)$$

Distance

$$aff(x, y) = e^{-\frac{||x-y||2}{2\sigma d^2}} \quad (6)$$

Let A, B partition G. Therefore $A \cup B = V$ and $A \cap B = \emptyset$

Similarity between $A \ and \ B$ is

$$cut(A, B) = \sum_{i \in A, j \in B} Wij \quad (7)$$

The optimal bi-partition of G is the one that minimizes cut and this cut biased towards small regions.

VII. EXPERIMENTAL RESULTS

We have experimented on a variety of natural images with SRM. We tested the problems over segmentation of the objects. We provided a comprehensive contribution of each step and introduced a well-organized low level feature representation of the image segments. We propose the system to represent the segments using color, intensity and texture. The proposed system performs the quantitative estimation of our approach and compares it with the present state of the art. After merging the estimated probabilities are given for improved quality of result. The normalized cut method performs the segmentations for each region. A difficult pattern image that is light can transmit and reflect different spectrum relative to color image such as person body and clothes etc. The object is segmented properly by using Ncut method and SRM algorithm by reducing the over segmentations during the segmentation process as shown in Fig 9. In this video frame sky, sea, plants, land, tree and person are almost segmented properly by different colors to identify the computer vision.

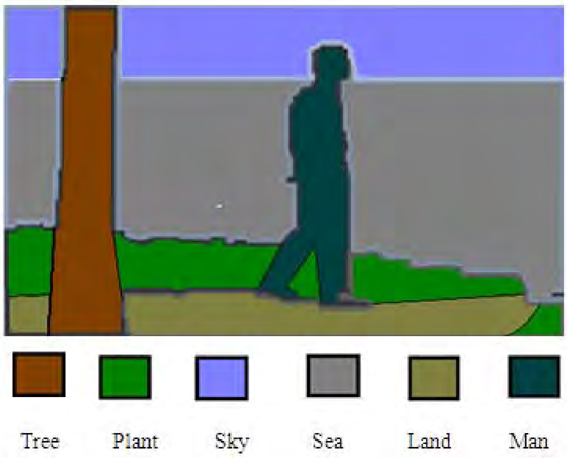

Fig. 9 Region Merging

VIII. CONCLUSION

Here we proposed SRM based outdoor scene video image segmentation. This project mainly used for pattern recognition and computer vision applications. Over segmentation images are not possible to set segmentation process parameters such as threshold value. So that all the objects are extracted from the background each other without over segmentation. The proposed system is cable and more efficient to reduce over segmentation of object. Finally we have improved the segmentation accuracy. However some object classes such as bicycles, motorcycles have very complex structures. Some parts of the objects are not strongly attached to other parts. For these object classes the SRM may not be able to piece the whole object together. Instead it may only piece some semantically meaningful parts of the objects together. For these objects higher level object specific knowledge is still required to segment entire objects.

REFERENCES

[1] Yang, P.Meer and D.J.Foran, “Multiple Class Segmentation Using a nified framework over Man-shift Patches”, in Proc IEEE CVPR, 2007 PP1-8.

[2] S. Gould, R. Fulton and D. Koller, “Decomposing a Scene into Geometric and Semantically Consistent Regions”, in Proc, IEEE ICCV, 2009 PP1-8.

[3] B.C.Russell, “Using Multiple Segmentation to Discover object and their Extent in Image Collections”, in Proc. IEEE CVPR 2006, Vol. 2 PP1605-1614.

[4] Jenny Benois-Pineau, “Segmentation Based Multiclass Semantic objects Detection Multimedia Tools and Applications”2011.

[5] Felipe Calderero, “Region Merging Techniques Using Information Theory Statistical Measures”, IEEE Trans.Vol.19, No.6, June 2010.

[6] Tiber Neoravcik,“Image Segmentation in Programming Environment,” xi International Conference Workshop October 2009.

[7] T. Malisiewicz and A.A. Efros, “Improving Spatial Support For objects Via Multiple Segmentations”, in Proc. BMVC, 2007..

[8] S.Gould, J.Rodgers, D.Cohen, G.Elidan and D.Koller, “Multiclass Segmentation with Relative Location Prior”, Int.J. Computer vision, Vol 80, No.3, PP300-316, Dec-2008.

[9] B.C. Russell, A.B. Torralbe, K.P. Murphy and W.T. Freeman,La-belme, ss“A Database and Web-based Tool for Image Annotation”, Int.J. Computer Vision, Vol 77, No.1-3, PP157-173.

[10] J.Shotton, J.Winn, C.Rother, and A.Crimnisini, “Textonboost for Image Understanding: Multiclass object recognisation and segmentation by jointly modeling texture layout and context”, J .Comput. Vis, Vol.80, N0.1.PP2-23, Jan 2009.

[11] J.B.shi and J.Malik,”Normalized cuts and image segmentation,”IEEE Trans.pattern Anal.Mach.Intel.,Vol.22, No.8,pp 888-905,Aug. 2000.

[12] Chang Cheng, Andreas Koschan, “Outdoor Scene Image Segmentation Based on Background Recognition and Perceptual Organization” in IEEE, Transaction on Image Processing,Vol,21,No.3,March 2012.

[13] Ali, K. Jalil, A.; Gull, M.U.; Fiaz, M. ” Medical Image Segmentation Using H-minima Transform and Region Merging Technique” Frontiers of Information Technology (FIT), 19-21 Dec. 2011.

[14] Sheng Liu Rui Cao; Shengyong Chen; Zichen Chen “A Novel Quality Evaluating Method for Over-Segmentation Approaches Using Real-Time Boundary Information” Digital Media and Digital Content Management (DMDCM), 2011 Workshop on 15-16 May 2011.

[15] Haji, M. Sahoo, K.A.; Bui, T.D.; Suen, C.Y.; Ponson, D. “Statistical Hypothesis Testing for Handwritten Word Segmentation Algorithms” Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on 18-20 Sept. 2012.

[16] Richard Nock and Frank Nielsen,” Statistical Region Merging,” IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 26, No. 11, November 2004.

[17] Jianbo Shi and Jitendra Malik, Member, IEEE.” Normalized Cuts and Image Segmentation,”IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, August 2000.

[18] Ali, K. Jalil, A.; Gull, M.U.; Fizz, M. ” Medical Image Segmentation Using H-minima Transform and Region Merging Technique” Frontiers of Information Technology (FIT), 19-21 Dec. 2011

[19] F. Lang, J. Zhao, “Hierarchical classification of polar metric sar image based on Statistical Region Merging”, ISPRS Vol 1-7,Sep 2012.

A. Niranjil Kumar received M.E. degree specialized in Applied Electronics from Anna University, Chennai, Tamil Nadu, India, in 2006. He is working as Assistant Professor in the department of ECE, P.S.R.Rengasamy College of Engineering for Women, Sivakasi, Tamil Nadu, India. He has published many papers in video surveillance, background subtraction, image quality measure and image segmentation. His research is mainly focused on video surveillance.

Dr. C. Sureshkumar received Ph.D. in Computer Science and Engineering, from Anna University, Tamil Nadu, India, in 2011. He is working as Principal in J.K.K.Nattraja College of Engineering and Technology, Namakkal, Tamil Nadu He has published many papers in computer vision applied to automation, motion analysis, image matching, image classification and view based object recognition and management oriented empirical and conceptual papers in leading journals and magazines. His present research is focused on statistical learning and its application to computer vision and image understanding, problem recognition and video surveillance.