OBJECT EXTRACTION USING EDGE, MOTION AND SALIENCY INFORMATION FROM VIDEOS

Rashmi J¹, Mangala C N², Sushma J³

¹Department of Computer Science, East West Institute of Technology, Karnataka, India ²Department of Computer Science, East West Institute of Technology, Karnataka, India ³Department of Computer Science, Vivekananda Institute of Technology, Karnataka, India

Abstract

Object detection is a process of finding the instances of object of a certain class which is useful in analysis of video or image. There are number of algorithms have been developed so far for object detection. Object detection has got significant role in variety of areas of computer vision like video surveillance, image retrieval`. In this paper presented an efficient algorithm for moving object extraction using edge, motion and saliency information from videos. Out methodology includes 4 stages: Frame generation, Pre-processing, Foreground generation and integration of cues. Foreground generation includes edge detection using sobel edge detection algorithm, motion detection using pixel-based absolute difference algorithm and motion saliency detection. Conditional Random Field (CRF) is applied for integration of cues and thus we get better spatial information of segmented object.

Keywords: Object detection, Saliency information, Sobel edge detection, CRF.

1. INTRODUCTION

Object extraction is the key element in analysis of videos in computer vision and multimedia application. In recent years, object detection has got considerable role in the field of video processing. Detecting moving object from the series of images is not a simple task. In most cases, object detection is become difficult when both background and foreground objects are in movement. For example if trees and sky are the objects of background region, if there any movement found in branches of trees or clouds of sky, and this leads to overlapping in object extraction. The region of interest focused on the part of the image that gets changed from one frame to another. Sometimes change in the background objects are not at all required. Hence, efficient segmentation of object is still a challenge.

According to procedure of the algorithm, object detection techniques can be categorized into three groups:

- Spatiotemporal.
- Motion segmentation.
- Morphological and model-matching techniques.

Spatiotemporal segmentation technique does not bother about motion information. For segmentation, it concentrates on space and time information of the frame. Motion segmentation is based on estimating the differences from one frame to the next frame. It uses motion parameters such as color information, luminance information, etc. Morphological techniques make use of morphological filters or watershed segmentation. Compared to three groups, morphological technique is more efficient from the computational point. Edge change detection [1, 11] and pixel based difference [15] are most widely used approaches for moving object extraction. Edge change detection provides inaccurate information about boundary, as the number of objects increases boundary inaccuracy also increases. Most conventional algorithms were only able to detect most dominant structures in images and failed to detect other structure of object. By considering this problem the [5] an algorithm is implemented to find regions which are important for indexing in a video database system. The algorithm works on shot basis. First segmentation is done for each frame to get similar regions in terms of color and texture. Then, extraction of features like set of regional and inter-regional color, shape, texture and motion of entire regions is done. These features are classified has being important and not important using Support Vector Machine (SVM), at last each important region is tracked within each shot for trajectory generation and testing effectiveness. But this technique only aims to matching particular features with targets and appropriate to detect objects of specific classes. These techniques are based on training the specific set of database and restrict to detect those set objects only. Saliency map [3, 9] is the efficient algorithm for detecting moving object without knowledge of its type. Human vision has noticeable ability to automatically detect only salient regions. Researcher's goal is to model the human visual system in machine vision. However, without knowledge of type of object or training data, it is difficult to extract the foreground object in video. In [3] proposed an algorithm based on the representation of log spectra of images by finding the spectral residual and its general ability. Author decomposed the image using spectral residual into two parts. Can be expressed as:

H(image)=H(innovation)+H(prior knowledge)(1)

H(innovation) is the required part and H(prior knowledge) is the redundancy that will be rejected by the coding system. Saliency map is very effective in extracting the foreground region in almost all the respective images submitted.

Normally object extraction in video uses either supervised or unsupervised approaches. Supervised methods need prior information on the interested subject and have to gather training data before the implementation of algorithm. Supervised methods require user interaction for classifying foreground and background regions. For example image segmentation algorithms [2, 9, 14]. Segmentation improves the quality of object detection algorithm. Random field models are widely used approaches in segmentation. In [14] proposed unsupervised image segmentation using simple MRF. Traditionally MRF was using training data to estimate necessary model parameters; hence it was unsuitable for unsupervised segmentation. The implementation scheme uses function-based weighting parameter between two components. By this technique the model is able to estimate the parameters automatically and produce précised result. Conditional Random Field (CRF) [2, 9] is the variance of MRF and it is more flexible than MRF. In our work, we are using CRF for labeling of extracted object. CRFs have wide range of applications in variety of domains, including text processing, computer vision, and bioinformatics.

In this paper we introduced multiple feature detection based approach for foreground segmentation from sequence of frames. CRF is applied for labeling of extracted region, to get clear structure of the foreground object. The proposed work is directly motivated by the works [4, 9]. In [4], the system makes use of multiple features: color and motion information based on the characteristic and combined those features for moving object extraction. And in [9], system used CRF for combining the saliency induced features.

2. PROPOSED SYSTEM

Proposed object extraction framework utilizes multiple features of foreground region and combines those features to get resultant frame. Below Figure 1 shows the block diagram of system design for moving object extraction.



Fig1: Proposed Architecture

In first step a video is taken as input and decomposed into sequence of frames, and then frame sequence is preprocessed.

2.1 Foreground Extraction

From preprocessed frames, foreground extraction is done by using edge detection, motion change detection and saliency detection. At last CRF is applied to get resultant sequence of frames. After the frame generation, each frame is preprocessed. Preprocessing involve color conversion. When color conversion is done, following detection procedures encountered to detect foreground region:

2.1.1 Sobel Edge Detection:

In this paper, we used sobel operator for edge detection. Most edge detection methods work based on assumption that the edge to be found where there is a sudden change or discontinuity in intensity in the image. Sobel utilizes this assumption and uses the derivative approximation to estimate edges. Hence, it yields edges at the points where the gradient of the image is maximum. Here vector as a gradient, whose components measure how pixel value changes rapidly with distance in the direction x and y. Therefore, the gradient components are to be found using the approximation:

$$\frac{\partial f(x,y)}{\partial x} = \Delta x = \frac{f(x+dx,y) - f(x,y)}{dx}$$
(2)

$$\frac{\partial f(x,y)}{\partial x} = \Delta y = \frac{f(x, y+dy) - f(x,y)}{dy}$$
(3)

where dx and dy calculate the distance along x and y directions respectively. The distance is measured in terms of number of pixel between the points. dx = dy = 1 is the point where coordinates of pixel are (i, j) therefore,

$$\Delta x = f(i+1,j) - f(i,j) \tag{4}$$

$$\Delta y = f(i, J+1) - f(i, j) \tag{5}$$

In order to find the discontinuity of gradient, we need to calculate the change in the gradient at (i, j). This is completed by estimating the magnitude measure M and gradient direction θ , are given by:

$$M = \sqrt{\Delta x^2 + \Delta^2 y} \tag{6}$$

$$\theta = \tan^{-1} \left[\frac{\Delta x}{\Delta y} \right] \tag{7}$$

Sobel detector is magnificently sensitive to noise in pictures, hence edges are effectively highlighted.

2.1.2 Motion Detection

Motion detection algorithm detects the moving object by calculating difference between the current frame and previous frame. It is most widely used method in motion detection. The method works based on pixel based difference.

 I_n is assumed to be the value of n^{th} frame and I_{n+1} is the value of $n + 1^{th}$ frame in the sequence of image respectively. The absolute difference of two consecutive frames is given by

$$I_{d(n,n+1)} = |I_{n+1} - I_n| \tag{8}$$

By using equation (8) we obtain frame that contain region of moving object.

2.1.3 Saliency Map Detection

Saliency detection measures visual saliency of given image. In our system we used optical flow based motion saliency mapping. This approach directly utilizes the conventional optical-flow model. The model predicts, among others,

$$I(x + u, y + v, t + 1) - I(x, y, t) = 0$$
(9)

Or

$$l_x u + l_y v + lt = 0 \tag{10}$$

where (u, v), to be calculated at every pixel (x, y) and is the displacement from the t frame to t + 1 frame, (l_x, l_y) is the gradient of spatial which is calculated by two adjacent frames. Equation (9) is the gray-value constancy, and

equation (10) is the linearization which is referred to as constraint of the optical flow. The field of flow over the image domain in terms of (u,v) is then described by variation model, where condition is associated with regularization terms by energy function on smoothness. For example an energy function f, the equation is given by,

$$min_{(u,v)}f(u,v) = \phi(dI(u,v)) + \lambda\phi(||\nabla(u,v)||^2)$$
(11)

where, dl(u, v) = l(x + u, y + v, t) - l(x, y), ϕ , is 2norm, convex and differentiable function of flow field and ∇ represents $(\partial_x, \partial_y, \partial_t)$. $\partial_t(u, v)$ is the temporal gradient component, which become absent when (u, v) is not available at t - 1. When the flow field is obtained at the time t, m extract the saliency map by the magnitude |u(x, y)| + |v(x, y)|. The solution for equation is intensive in the operations of array and in general iterative in nature.

2.2 Conditional Random Field (CRF)

Conditional random field is robust technique to measure the structural information of a group of variables with the related observations. For moving object segmentation in a video, CRF has been applied estimate the label of observed pixel in a given image I.



Fig 2: Segmentation in CRF

As shown in Figure 2, pixel I in a frame of video is associated with observation z_i , F_i is the hidden node which represents label of foreground or background. z_i Observations calculates the label F_i . Simultaneously the spatial coherence between this resultant and associated observations z_j and labels F_j are taken to consideration. Thus, estimating the label of an observation node is similar to maximizing the posterior probability function.

$$p(F|I,\psi) \alpha exp\left\{-\left(\sum_{i \in I} (\psi_i) + \sum_{i \in I, j \in Neighbor} (\psi_{i,j})\right)\right\}$$
(12)

Where ψ_i denotes the unary term which decides the possibility of F_i with observation z_i . $\psi_{i,j}$ is the pairwise term

that describes the relation between the z_i and z_j which are associated each other between the output labels F_i and F_j . One important thing is observation z can be represented by specific feature or set of features.

CRF is utilized to combine all the extracted features and hence the problems in object extraction are solved automatically.

3. RESULT



Fig 3: Original Frame.



Fig 4: Frame after segmentation.

Figure 3 shows the original frame of a video and Figure 4 shows the moving object in frame of video after applying the foreground extraction and segmentation algorithm.

4. CONCLUSION

In this paper, we proposed an approach for automatic extraction of moving object from video which utilizes multiple information such as edges, motion change and saliency information of the frame. First we achieved foreground extraction by using edge detection, motion detection and saliency measure. After the results obtained by these three algorithm, the results are combined and formed single output. Finally we applied powerful technique CRF for segmentation. The major benefit of our proposed method is that we no need to collect training data (i.e. without any prior knowledge of object of interest), also no need of user interaction during the segmentation. Combining the multiple outputs can improve the quality of result.

REFERENCES

- [1] Kavitha Ganesan and Shanmugam Jalla, "Video Object Extraction Based on a Comparative Study of Efficient Edge Detection Techniques", International Arab Journal of Information Technology, Volume 6, Issue 2, 2009.
- [2] Xuming He, Richard S. Zemel and Miguel A. Carreira-Perpinan, "Multiscale Conditional Random Fields for Image Labeling", IEEE, Volume 2, 2004.
- [3] Xiaodi Hou and Liqing Zhang, "Saliency Detection: A Spectral Residual Approach", IEEE, PP. 1-8, 2007.
- [4] Jinhui Pan, Shipeng Li and Ya-Qin Zhang, "Automatic Extraction of Moving Objects Using Multiple Features and Multiple Frames", IEEE, Volume 1, PP. 36-39, 2000.
- [5] Muhammet Bastan, Ugur Gudukbay and Ozgur Uluso, "Segmentation-Based Extraction of Important Objects from Video for Object-Based Indexing", IEEE, PP. 1357-1360, 2008.
- [6] William Brendel and Sinisa Todorovic, "Video Object Segmentation by Tracking Regions", IEEE, PP. 833-840, 2009.
- [7] Abhishek Kumar Chauhan and Deep Kumar, "Study of Moving Object Detection and Tracking for Video Surveillance", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, 2013.
- [8] Wei Zeng, Wen Gao and Debin Zhao, "Automatic Moving Object Extraction in MPEG Video", IEEE, Volume 2, PP. 524-527, 2003.
- [9] Wei-Te Li, Haw-Shiuan Chang, Kuo-Chin Lien, Hui-Tang Chang and Yu-Chiang Frank Wang, "Exploring Visual and Motion Saliency forAutomatic Video Object Extraction", IEEE, Volume 22, Issue 7, 2013.
- [10] Hari Hara Santosh, Venkatesh, Poornesh, Narayana Rao and Arun Kumar, "Tracking Multiple Moving Objects Using Gaussian Mixture Model", International Journal of Soft Computing and Engineering, Volume 3, Issue 2, 2013.
- [11] Y.Ramadevi, T.Sridevi, B.Poornima, B.Kalyani, "Segmentation and Object Recognition Using Edge Detection Techniques", International Journal of Computer Science & Information Technology, Volume 2, Issue 6, 2010.
- [12] Arnab Roy, Sanket Shinde and Kyoung-Don Kang, "An Approach for Efficient Real Time Moving Object Detection", ESA, PP. 157-162, 2010.
- [13] Ashish Ghosh, Ajoy Mondal and Susmita Ghosh, "Moving object detection using Markov Random Field and DistributedDifferential Evolution", Elsevier, PP. 121-136, 2014.
- [14] Huawu Deng and David A. Clausi, "Unsupervised image segmentation using a simpleMRFmodel with a newimplementation scheme", Elsevier, PP. 2323-2335, 2004.
- [15] Nishu Singla, "Motion Detection Based on Frame Difference Method", International Journal of Information & Computation Technology, Volume 4, Issue 5, 2014.

BIOGRAPHIES



Rashmi J, graduated in computer science and engineering in East West Institute of Technology from Visveswaraya University, Belgaum, India.



Mangala C N, Assistant professor in East West Institute of Technology from Visveswaraya University, Belgaum, India.



Sushma J, student in computer science and engineering in Vivekananda Institute of Technology from Visveswaraya University, Belgaum, India.