Automatic VOE Techniques using CRF and Visual Saliency Method

Dr. Swathi. K[1], Sivabalan .T[2], Vijayanathan .R[3]

[1]Associate Professor, Department of Computer Science & Engineering,
Cauvery College of Engineering and Technology, Trichy (Tamilnadu), India

[2]M.E (Scholar), Department of Computer Science & Engineering,
Cauvery College of Engineering and Technology, Trichy (Tamilnadu), India

[3]Senior Librarian and Head, Department of Library and Information Science,
Cauvery College of Engineering and Technology, Trichy (Tamilnadu), India

Abstract: ***Object extraction from active videos is by and large an unexplored area in computer vision and pattern recognition. We have designed several analysis algorithms for video object detection and segmentation in the general framework of multimedia content analysis. The algorithms are developed in two directions, i.e., fully automatic algorithms and semi-automatic algorithms. In this project we can propose a new method of video object detection to automatically extract the objects from acquired videos. The existing video object detection techniques operate under homogenous object motions. In contrast, our proposed video object extraction approach effectively detects the object under motion sequences and robust a Hidden Markov Model (HMM), Support Vector Machine (SVM) and Gaussian Mixture Model is applied to automatically determine the label of each pixel based on the observed models. Then we deal with multiple object instances and this system does not require the previous knowledge of the object.***

Keywords: Object Detection, Object segmentation, object classification

1. Introduction

Detection and segmentation of objects of one or more given classes is of fundamental interest in computer visualization. Conventionally, one would section an image into areas, and then try to categorize these areas as belonging to one of the desired classes at a glance; human can easily determine the subject of notice in a video, even though that topic is obtainable in an unknown or cluttered background or even has never been seen before. The representation of video information, in terms of its content, is at the foundations of many multimedia applications such as broadcasting content-based information retrieval, video editing , activity recognition and its entertainment In existing system use the unsupervised approaches do not train any specific object detectors or classifiers in progress. For videos imprison by a fixed camera, removal of foreground substance can be treat as a background calculation difficulty. In other terms, foreground substance can be notice purely by subtract the current frame from a video progression. However, if the setting is constantly altering or is occluded by foreground substance, setting model becomes an extremely demanding task. For such cases, researchers usually aspire at learning the setting model from the contribution video, and the foreground substance is considered as outliers to be notice. For instance, an auto regression moving average model (ARMA) that approximates the native look of dynamic textures and area was anticipated and it mainly covenant with situation in which the background consists of natural scene similar to sea waves or trees. However, discovery visual saliency in pictures or video frames would offer promising results and deduce the area of the foreground substance. However, because real-world videos capacity encounters low contrast or inadequate lighting, etc. Troubles, one may not be capable to get popular visual saliency maps for recognize applicant foreground substance. As a consequence, one cannot basically apply visual saliency techniques for segmenting foreground substance in real world videos.

Figure 1: Example of object extraction

2. Related Work

In [1] B. Wu and R. Nevatia et al. For the cases with partial, inter-object occlusions, part based representations can be used. For each part, a detector is learned and the part detection responses are combined to form object hypotheses. The part detectors are typically applied to overlapping windows and the windows are classified independently. Consequently, one local feature may contribute to multiple overlapped responses for one object. Some false detection may also occur, as local features may not be discriminative enough. Because of poor image cues or partial occlusions, some object parts may not be detected. To get a one-to-one mapping from part detection responses to object hypotheses, we need to group the responses and explain inconsistency between the observation and the hypotheses. When objects are close to one another, both the one object-multiple-response problem and the part-object assignment problem require joint consideration of multiple objects, instead of treating them independently. One important component of the joint consideration is the analysis of the occlusion relation between multiple objects in the 2-D image space. To obtain an accurate occlusion model, pixel level segmentation of objects is necessary.

In [3] Y. Y. Boykov, O. Veksler, and R. Zabih et al. The major difficulty with energy minimization for early vision lies in the enormous computational costs. Typically these energy functions have many local minima. Worse still, the space of possible labeling has dimension, which is many thousands. There have been numerous attempts to design fast algorithms for energy minimization. Simulated annealing was popularized in computer vision, and is widely used since it can optimize an arbitrary energy function. Unfortunately, minimizing an arbitrary energy function requires exponential time, and as a consequence simulated annealing is very slow. In practice, annealing is inefficient partly because at each step it changes the value of a single pixel. The energy curves as a function of time are very similar to the diamond example shown above, but are omitted to save space and also include the ratio between annealing energy and ours. The third row for each image gives the best energy that annealing eventually achieves, when run until it is making very minimal progress.

3. Gaussian Mixture Models (GMMs)

Gaussian Mixture Models (GMMs) are among the most statistically mature methods for. In this tutorial, we introduce the concept of clustering, and observe how one appearance of clustering in which we assume that individual data points are generated by first choosing one of a set of multivariate Gaussians and then sampling from them can be a definite computational operation. This optimization method is called Expectation Maximization (EM). Spend some time giving a few high level explanations and demonstration of EM, which turn out to be precious for numerous other algorithms beyond Gaussian Mixture Models.

Advantages

1. Component distributions have high "peaks".
2. The mixture model "covers" the data well

4. Hidden Markov Models (HMM)

HMM is a popular technique widely used in signal processing. HMMs are a formal foundation for making probabilistic models of linear sequence "labeling" problems and they are especially known for their applications in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics. They are mostly used for classifying sequential data to capture the temporal relationships of the extracted features. In our research, we extended it to video analysis and classification. In an HMM, there are a finite number of states, each of which is associated with a transition probability to the others. Every time, the HMM stays in one definite state. The state at time t is directly influenced by the state at time *t*-1. After each translation from one state to another, an output observation is generated based on an observation probability distribution associated with the current states. Formally, a HMM is defined to be:

HMM=$\{N, B, II\}$

where N is the set of states, B is the number of observation symbols and II is set of state transition probabilities. First, an observation sequence based on the speed of color changes computed from video frames is generated from this input video. Then, the observation sequences are fed into each HMM. Finally, computing the log-likelihood of the test sequences, incoming videos are classified. In every step, the system determines the probability distribution of the next phase from the given state. Classification process of videos consists of two step. In the first step algorithm is used to estimate the most likely parameters for the HMM that generates the training set. In the second step the probability of the observation of that sequence given the HMM is compute by backward algorithm. Property groups identified as using similar color distributions of the groups were calculated with Gaussian mixture.

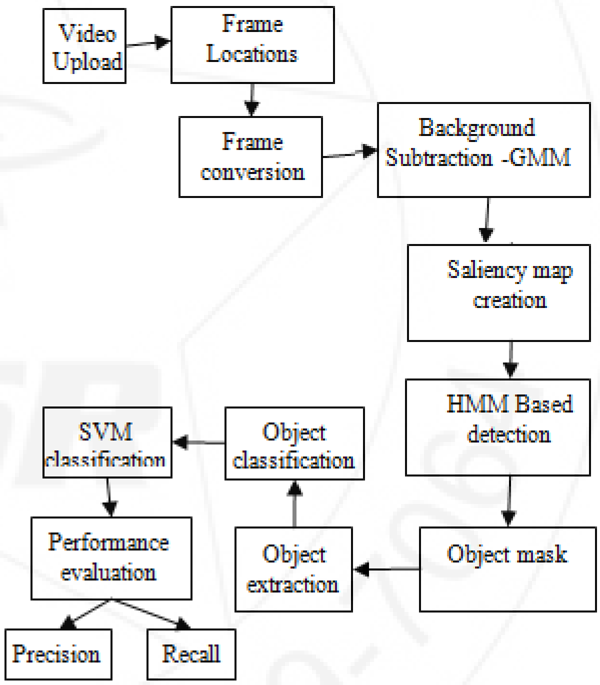

Figure 2: System Architecture

5. Support Vector Machine (SVM)

Support Vector Machines (SVM) is a state-of-the-art learning machine based on the structural risk minimization induction principle. In recent years, SVM has been extensively used as a classification tool and has found a great deal of success in a wide range of applications including pattern recognition, communications, and image/video analysis. Support Vector Machines are based on the concept of decision planes that define conclusion limits. A resolution plane is one that separate among a set of objects having dissimilar set memberships. In this example, the objects go any to class GREEN or RED. The disconnection line explain an edge on the right side of which all substance are GREEN and to the left of which all substance are RED. Any innovative item falling to the right is labeled, i.e., classified, as GREEN. The above is a classic example of a linear classifier, i.e., a classifier that divides a set of objects into their respective gathering with a line. Most classification tasks, although, are not too easy, and frequently additional composite structures are needed in order to make a most favorable separation, i.e., appropriately categorize new substance on the basis of the examples that are available. This situation is depicted in the illustration below. Compared to the before representation, it is clear that a full severance of the GREEN and RED objects would require a curve. Classification tasks based on drawing separating lines to distinguish between objects of different class memberships are known as hyper plane classifiers. Support Vector Machines are mainly suited to handle such tasks.

Advantages:

1) By introducing the kernel, SVMs gain flexibility in the choice of the form of the threshold separating solvent from ruin companies, which needs not, are linear and still desires not having the same functional form for all data, because its function is non-parametric and operates locally.
2) As a consequence they can work with economic ratios, which demonstrate a non-monotone relation to the score and to the possibility of prevaricate, or which are non-linearly dependent and this without needing any specific work on each non-monotone variable.

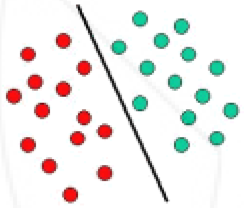

Figure 3: Example of SVM

6. Experimental Results

In the experimental results show the object extraction of the existing system and the proposed system. Then show the results in the figure.

Table 1: Comparison Table

Method	*Object extraction*
Existing System	50%
Proposed System	95%

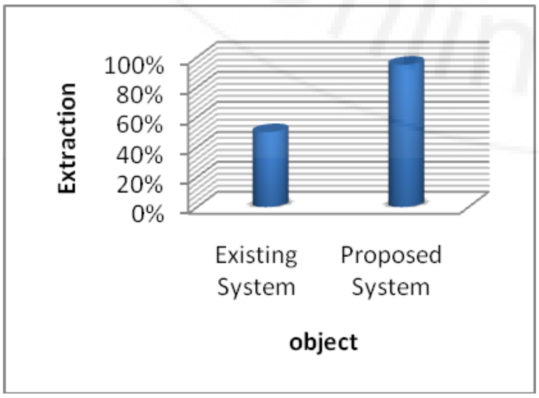

Figure 4: Comparison results

7. Conclusion

Fully automatic extraction of semantically meaningful objects from visual data is one of the ultimate aspirations in computer vision and pattern detection. The noticeable academic anxiety in this difficulty, there is an extensive array of practical applications that can benefit tremendously from successful object extraction algorithms. We conclude that our proposed approach based on automatic VOE approaches and analyze the multiple motions and represent the visual saliency map for extracted foreground objects. We use the Conditional random field to analyze the features and restriction. From the new results our anticipated method performs efficient object detection with saliency features. Our most important come close to is VOE and do not need previous information of the object of interest.

References

[1] B. Wu and R. Nevatia, "Detection and segmentation of multiple, partially occluded objects by grouping, merging, assigning part detection responses," Int. J. Comput. Vis., vol. 82, no. 2, pp. 185–204, 2009.

[2] Z. Lin and L. S. Davis, "Shape-based human detection and segmentation via hierarchical part-template matching," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 4, pp. 604–618, Apr. 2010.

[3] Y. Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[4] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut': Interactive foreground extraction using iterated graph cuts," ACM Trans. Graph., vol. 23, no. 3, pp. 309–314, 2004.

[5] A. Criminisi, G. Cross, A. Blake, and V. Kolmogorov, "Bilayer segmentation of live video," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2006, pp. 53–60.

[6] P. Yin, A. Criminisi, J. M. Winn, and I. A. Essa, "Bilayer segmentation of webcam videos using tree-based classifiers," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 1, pp. 30–42, Jan. 2011.

[7] X. Bai and G. Sapiro, "A geodesic framework for fast interactive image and video segmentation and matting," in Proc. IEEE Int. Conf. Comput. Vis., Oct. 2007, pp. 1–8.

[8] M. Gong and L. Cheng, "Foreground segmentation of live videos using locally competing 1SVMs," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 2105–2112.

[9] T. Bouwmans, F. E. Baf, and B. Vachon, "Background modeling using mixture of Gaussians for foreground detection—A survey," Recent Patents Comput. Sci., vol. 3, no. 3, pp. 219–237, 2008.

[10] F.-C. Cheng, S.-C. Huang, and S.-J. Ruan, "Advanced background subtraction approach using Laplacian distribution model," in Proc. IEEE Int. Conf. Multimedia Expo, Jul. 2010, pp. 754–759.

Author Profile

Dr. K. Swathi obtained her under-graduation in B.E., (Computer Science & Engineering) from Bharathidasan University, Trichy in 1999. She obtained her M.E. degree in Computer and Communication Engineering from Anna University, Chennai in 2004.She obtained Ph.D degree in Faculty of Information & Communication Engineering from Anna University, Chennai in 2014. Presently, she is working as Associate Professor in Computer Science & Engineering, Cauvery College of Engineering and Technology, Trichy in 2008. She has 14 years teaching experience and also she had attended many workshops, seminars and conferences on Research issues in Image processing. She has published papers in international journals and presented papers in various Conferences. She is the life member of Indian Society for Technical Education (ISTE). Her areas of interest include Image processing, Data mining, network security and Software engineering.

T. Siva Balan, received hem B.E in computer science engineering from periyarmaniammai university , vallam in 2011 and doing him M.E in Computer Science Engineering in Cauvery College of Engineering and Technology, Trichy. Her area of interest includes Image Processing and Networks.

Vijayanathan. R. received his Master of philosophy Library and Information Science from Annamalai University in 1999 Also he obtained his post-graduate degree in Master of Economics in Bharathidasan University in 1996.He is working as a Sr. Librarian, Department of library and Information Science in Cauvery College of Engineering and technology, Trichy from 2009. He has guided many M.Phil scholars and Committee member in various Universities in Tamil Nadu. He served more than 14 years as Senior Librarian in reputed Engineering College and also he had attended many workshops, seminars and conferences on Research issues in Bibliometric analysis. He has published 21 research papers in National and International journals. His areas of interested are networking; Cloud computing, IC Technologies, Environmental study, Library Automation, Webometric study, Scientometric study, Bibliometric analysis, and citation study.