Australian Journal of Basic and Applied Sciences, 9(21) Special 2015, Pages: 115-119

ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

A Survey on Compressed Video Segmentation

[1]First K. Suganya Devi and [2]Second P. Srinivasan

[1]Department of Computer Science and Engineering, University college of Engg Panruti, Panruti 607106, Tamilnadu, India
[2]Department of Physics, University college of Engg Panruti, Panruti 607106, Tamilnadu, India

ARTICLE INFO

Article history:
Article Received: 12 January 2015
Revised: 1 May 2015
Accepted: 8 May 2015

Keywords:
Compressed Domain, Compressed video, Markov Random Field, Motion Compensation, Motion Estimation, Motion segmentation, Motion vectors, Pixel Domain.

ABSTRACT

Segmentation is important concern in the field of image processing. It is important problem to segment compressed video and many approaches has been proposed to solve the problem of segmentation techniques for compressed video. State of the art object segmentation methods can be broadly classified into pixel domain approaches and compressed domain approaches and combinations of both have also been proposed to balance complexity and accuracy. This paper made survey on these methods and various algorithms are compared and discussed with their merits and demerits.

To Cite This Article: First K. Suganya Devi and Second P. Srinivasan, A Survey on Compressed Video Segmentation. ***Aust. J. Basic & Appl. Sci.,*** *9(21): 115-119, 2015*

INTRODUCTION

Segmentation of moving object was important problem in the variety of applications such as video surveillance; video database browsing and object based video transcoding. The broad classified approaches are pixel domain and compressed domain approaches of segmentation. The pixel domain extracts object by exploiting visual features such as shape, color,and texture. In the pixel domain approach (Brouard, 2008; Kato, 2001), the compressed video has to be fully decoded prior to segmentation. This leads to high computational overload and over segmentation of these approaches.

On the other hand, compressed domain approach (Zeng, 2005; Chen, 2009) exploit compressed domain data, such as motion vectors (MVs) and DCT Coefficients , to facilitate segmentation. These methods have low complexity however they suffer from poor localization of object boundaries and inconsistency in the number of segmented regions from frame to frame. Combinations of both operations also have been proposed to balance complexity and accuracy (Zhong, 1999; Chen, 2009). These methods first create coarse segmentation map from sparse MV field and refine it in the pixel domain. Although these methods segment the compressed video accurately than purely compressed domain approaches, maintaining a consistent number of segmented regions across frames can still be a challenge. conference.

Compressed video segmentation methods:

A. Pixel domain methods:

In this kind of methods of segmentation of compressed video, the decoding of compressed video takes place and there is high computational complexity.

Spatio-temporal segmentation Based on MRF Model:

This method proposed in (Brouard, 2008) is based on pixel domain approach of segmentation. It extracts features by exploiting visual features such as shape, texture and color. Here compressed video has to be fully decoded prior to segmentation. This incorporates MRF model to segment and track video objects. First motion based segmentation is realized for Group of nine frames. Next MRF model is applied to improve MBs(Macro blocks) using spatial features and to keep consistency between successive GOF segmentation maps. A video object tracking is then achieved. The advantages of this approach are accurate segmentation and video objects tracking. However the demerits of this method are high computational load and over segmentation.

Unsupervised segmentation of Color-Texture Regions:

This method given in (Deng, 2001) involves two independent steps such as color quantization and spatial segmentation. In first step, colors in image are quantized to several representative classes that can be used to differentiate regions in image. The class map

Corresponding Author: First K. Suganya Devi, Department of Computer Science and Engineering, University college of Engg Panruti, Panruti 607106, Tamilnadu, India
E-mail: ssuganya.ucep@gmail.com

was formed and spatial segmentation of class map is done using JSEG algorithm on real images and videos. A Criterion for good segmentation is proposed. Here also video is fully decoded and high complexity and time wasted on decompression. For JSEG algorithm, three parameters need to be specified by user

1. Threshold for color quantization
2. Number of scales desired for image
3. Threshold for region merging.

These are necessary because of varying image characteristics in different applications. The disadvantages of the above approach are varying shades due to illumination. The problem is difficult to handle because in many cases not only illuminant component but also chromatic components of pixel change their values due to spatially varying illumination. Over segmentation also occur. Error generated in one frame can affect subsequent frames.

Color image Segmentation and parameter estimation in a Markovian framework:

This method proposed an unsupervised image segmentation algorithm (Kato, 2001) and it is defined in markovian framework and uses a first order potential derived from a tri-variate Gaussian distribution in order to tie final segmentation to observed image. In this approach, color MRF Process consists of color difference of neighboring pixels. In an MRF segmentation model with color and line processes is proposed and use of three different lattice schemes(squares, hexagon and triangles) are discussed. The segmentation is obtained through simulated annealing. To estimate model parameters, we use an iterative algorithm, which subsequently generates a labeling and then recomputed the parameter values. This process requires good initial mean values. The methods have been tested on variety of real and synthetic images and results are very close to supervised segmentation algorithms.

The advantages of this method are that it does not require any human intervention. The demerits of the above approach are that it is time consuming. Half of the computing power is used for estimation of β in MRF segmentation and estimation of Gaussian parameters. Larger images or images with more classes require more CPU Time.

*B. **Compressed Domain Methods:***

Here there is no need of decoding of compressed video. The motion vectors are extracted from video and motion information are obtained from motion vectors and the computational complexity was reduced.

Robust Moving object segmentation on H.264/AVC compressed video using Block based MRF model:

In this method proposed in (Zeng, 2005), decoding of compressed video is avoided and it employs a block based MRF model to segment moving objects from sparse motion vector field obtained directly from bit stream. Object Tracking is integrated in uniform MRF model and exploits object temporal consistency simultaneously

The merits are remarkable performance, and can extract moving object efficiently and robustly. However, segmentation precision must be improved for accuracy in segmentation. This was the demerit in this method.

Motion based Video object tracking in compressed domain:

Iterative Rejection is concept used in this mehod (Ritch, 2007) which is computationally efficient technique used to identify local motion. This algorithm is particularly useful for compressed domain processing as motion vector information is readily available. Tracking Macroblocks back through number of frames can be introduced to rectify errors introduced in encoder’s motion estimation algorithm. Here the method identify and track an object of interest within compressed video using only motion information. Here also we need not fully decode each frame and the system performed better than using iterative rejection alone as segmentation method. The merits are low complexity and proper segmentation. The future work of this paper includes extending system to identify and track multiple objects and to handle problems such as occlusion.

Representing Moving Images with Layers:

This representation [6] describes some methods for decomposing image sequences into layers using motion analysis and it can be used for applications. Analysis of an image sequence consists of 2 stages i) Robust motion segmentation & ii) Synthesis of the layered representation. Segmentation involves motion estimation technique based on robust estimation and K-means clustering in affine parameter space. At each iteration of motion estimation, our multiple model frameworks identify multiple coherent motion regions simultaneously, Iteratively, motion model parameters are calculated within these coherent regions and segmentation is refined. The advantages of these approaches are stability and robustness in segmentation by rejecting outliers at different stages in algorithm. The demerits are that the method was not successful for complex image sequences.

Empirical Bayesian Motion Segmentation:

This proposed method in (Vasconcelos) introduces simultaneous segmentation of an observed motion field and estimation of hyper parameters of a markov random field prior. The EB Method develop framework for estimating prior parameters that best explain observed data. This eliminates need for trial and error strategies for parameter setting and leads to better segmentations in less iterations. The clustering parameter should be carefully chosen to obtain good segmentation.

The demerits are as follow: Smaller values leads to very noisy segmentation and larger values originate segmentation with reduced accuracy near region boundaries. So we should be careful in choosing clustering parameter.

Content based video object segmentation and Retrieval:

This method impacts on effective semantic object segmentation and tracking in general video sequences (Zhong, 1999). It provides integrated approach for semantic object segmentation and content based search based on region based video object model. Here a model AMOS is discussed, a generic video object segmentation system which combine low level automatic region segmentation with user input for defining, and tracking semantic video objects. The system effectively fuses color, edge, and motion information in its region segmentation and utilizes an iterative region aggregation and boundary alignment process to generate and track accurate object boundaries. K-Means clustering algorithm is used in Motion estimation. Edge map is generated by canny edge detector and the motion field is generated by hierarchical block matching algorithm. The advantages are good and accurate segmentation and it is automatic without any human intervention.

Real time outlier Removal from MV fields for 3D Reconstruction:

This method in (Dante, 2003) obtain motion vectors from real time hardware and fast method is implemented to filter out more than 99.7% of all outliers and two filter criteria are intended to remove outliers from motion vector fields of real image sequences.

1. Smooth change it filters correctly even near null vector, however it reject some inlier vectors at depth discontinuities
2. Neighborhood is rather insensitive to large proportion of outliers and to depth discontinuities.

The merits of this approach are that it removes more percentage of outliers from the moving region and give accuracy in segmentation.

Statistical analysis of Dirty pictures:

This method reconstruct true scene (Besag, 1986), with additional knowledge that pixels close together tends to have same or similar colors. A simple, iterative method for reconstruction is proposed which does not depend on large scale characteristics. A continuous two dimensional region is partitioned into a fine rectangular array of sites or “pixels”, each pixel having particular “color” belonging to prescribed finite set. The demerit of this method is large computational complexity.

C. ***Combined Pixel domain and Compressed Domain approaches:***

Markov Random Field Classification:

This Motion Segmentation in compressed video using MRF Classification (Chen, 2010) focuses on unsupervised segmentation algorithm for extracting moving regions from compressed video using markov random field classification. The Segmentation method Combines both pixel domain and compressed domain approach. The distinctive feature of this paper is use of MV quantization based on local motion similarity to find most likely number of moving objects and use statistics of resulting clusters to initialize prior probabilities for subsequent MRF Classification.It overcomes over segmentation, under segmentation and inconsistency in the number of segmented regions. It provides good balance between accuracy and complexity. The demerits are that there is influence of camera motion on MV field and there is some over segmentation due to motion bias introduced by camera movement.

Low complexity global motion estimation from Block motion vectors:

This method proposed two global motion estimation methods (Smolic, 2009) that are based on block matching motion estimation. These methods estimate different global motion models. Model parameters are estimated sequentially allowing performing additional filtration to increase robustness and maintaining high processing speed and reasonable memory requirements. First algorithm estimates 4 parameter motion model and other estimates six parameter motion model. Algorithms use high speed approach of motion estimation and minimize the errors of global motion compensation introduced by low reliable motion vectors. The merit of algorithm is that it gives better result than least square method in robustness and accuracy. But both methods require reasonable amount of memory for processing. This is demerits in these methods.

Global motion Estimation for camera motion characterization:

In this method (Haller, 2009), two pixels based and 5 motion vector based GME algorithms are proposed by means of PNSR values for global motion compensation on five sequences with and without moving foreground objects. Subsequently, we evaluate motion vector based GME with highest PSNR value and 2 parameters with system for camera motion characterization. Gradient descent approach and least square solution approach are analyzed using Gaussian Newton algorithm and RANSAC algorithm are also used to obtain global motion parameters. Pixel based GME methods such as Gaussian Newton gradient descent algorithm is an energy minimization method and is used because of its good performance and simplified GME on down sampled image pairs are also obtained. It lowers computational cost even more and accelerates estimation process. Then camera motion parameterization was used for evaluation of different GME algorithms. The merits of this approach are that it provides good balance between computational complexity and motion estimation accuracy as well as highly satisfying camera motion characterization results.

Model fitting with applications image analysis and automated cartography:

A new model was introduced (Fischler, 1981) for fitting model to experimental data. It is capable of interpreting data containing significant percentage of gross errors and ideally suited for applications in automated image analysis (cartography). RANSAC (RANdom SAmple Consensus) was applied to solve Location determination problem. Given an image depicting a set of landmarks with known locations, determine point in space from which image was obtained.

The merit is that it solves LDP.

Reversible jump MCMC sampling:

A novel method called reversible jump MCMC (Markova chain Monte Carlo) has been proposed[20] for color segmentation. These methods make it possible to construct reversible markov chain samplers than jump between parameter subspaces of different dimensionality. This technique is used to solve unsupervised color image segmentation problem in Markovian framework. We have established Bayesian segmentation model using MRF modeling of underlying label field. The final estimates, satisfying MAP (Maximum A Priori) criterion are obtained through simulated annealing. The advantage of this approach is accuracy in segmentation.

Table I:

Different techniques	Methods	Strength	Weaknesses	Processing Time	Computational Complexity
Pixel Domain Methods	Spatia-temporal segmentation	Accurate segmentation.	Over segmentation	-	High
	Unsupervised segmentation	Segmentation was clear.	Varying shades due to illumination, over segmentation also occur.	0.33-0.50 sec per frame	High
	Markovian framework	It does not require any human intervention.	Time consuming. It takes more CPU time.	58.94sec per frame	High
Compressed Domain methods	Block based MRF Model	Stability, Robustness and Remarkable performance.	No accuracy.	Above 700ms per frame	Low
	Iterative Rejection	Proper segmentation	Occlusion occurs.	-	High
	Empirical Bayesian motion segmentation	Fast and better segmentation in less iteration.	Reduced accuracy near region boundaries.	-	Low
	Automatic segmentation	Good and accurate segmentation and automatic without any human intervention.	More computation.	20 sec per frame	High
Combination of pixel and compressed Domain Methods	Markov Random Field Classification	It provides good balance between accuracy and complexity and removes inconsistency in number of segmented regions from frame to frame.	Influence of camera motion on motion vector leads to over segmentation.	80ms per frame	Low
	Global motion Estimation	Robustness and accuracy.	Need reasonable amount of memory for processing.	43.1ms per frame	Low

Conclusion:

This paper surveys on various segmentation methods for compressed video based on two domain approaches such as pixel domain and compressed domain approach. Both methods have their merits and demerits. The above Fig.1 clearly explains about processing time needed for segmentation of video in various methods of segmentation. Although we need automatic segmentation without decoding of compressed video, accuracy and robustness and inconsistency in number of segmented regions was major concern. Thus we conclude combination of both pixel and compressed domain offer low computational complexity as well as accuracy in segmentation.

Comparison of Various Segmentation methods:

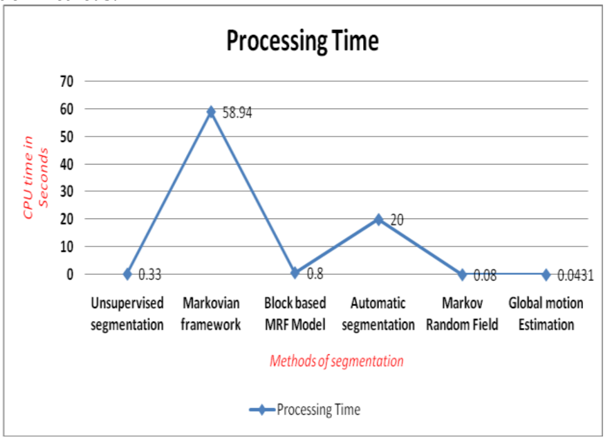

Fig. 1: Comparison chart for processing time for various models of segmentation.

References

Babu, R.V., K.R. Ramakrishnan and S.H. Srinivasan, 2004. "Video object segmentation: A compressed domain approach," *IEEE Trans. Circuits Syst. Video Technol*, 14(4): 462–474.

Besag, J., 1986. "On the statistic analysis of dirty pictures," *J. Roy. Statist.Soc. B.*, 48(3): 259–302.

Brouard, O., F. Delannay, V. Ricordel and D. Barba, 2008. "Spatio-temporal segmentation and regions tracking of high definition video sequences based on a Markov Random Field model," in *Proc. IEEE ICIP'08*, Dec, pp: 1552–1555.

Chen, Y.M., I.V. Bajić and P. Saeedi, 2009. "Coarse-to-fine moving region segmentation in compressed video," in *Proc. IEEE WIAMIS*, London, U.K., pp: 45–48.

Chen, Y.M., I.V. Bajić and P. Saeedi, 2010. "Motion segmentation in compressed video using Markov random fields," in *Proc. IEEE ICME*, Singapore, Jul., pp: 760–765.

Dante, A. and M. Brookes, 2003. "Precise real-time outlier removal from motion vector fields for 3-D reconstruction," in *Proc. IEEE ICIP*, pp: 393–396.

Deng, Y. and B.S. Manjunath, 2001. "Unsupervised segmentation of colortexture regions in images and video," *IEEE Trans. Pattern Anal.Mach. Intell*, 23(8): 800–810.

Ewerth, R., M. Schwalb, P. Tessmann and B. Freisleben, 2007. "Segmenting moving objects in MPEG videos in the presents of camera motion," in *Proc. ICIAP*, Sep., pp: 92–96.

Fischler, M. and R. Bolles, 1981. "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, 24(6): 381–395.

Haller, M., A. Krutz and T. Sikora, 2009. "Evaluation of pixel- and motion vector-based global motion estimation for camera motion characterization," in *Proc. WIAMIS*, May, pp: 49–52.

Kato, Z., 2008. "Segmentation of color images via reversible jump MCMC sampling," *Image Vis. Computing*, 26(3): 361–371.

Kato, Z., T.C. Pong and J.C.M. Lee, 2001. "Color image segmentation and parameter estimation in a Markovian framework," *Pattern Recognit. Lett.*, 22(3–4): 309–321.

Ritch, M. and N. Canagarajah, 2007. "Motion-based video object tracking in the compressed domain," in *Proc. IEEE ICIP*, Oct., 6: 301–304.

Shi, X., Z. Zhang and L. Shen, 2007. "Multiresolution segmentation of video objects in the compression domain," *Opt. Eng.*, 46(9): 097401.

Smolic, A., M. Hoeynck and J.R. Ohm, 2009. "Low-complexity global motion estimation from P-frame motion vectors forMPEG-7 application," in *Proc. IEEE ICIP*, Sep., pp: 271–274.

Su, Y., M.T. Sun and V. Hsu, 2005. "Global motion estimation from coarsely sampled motion vector field and the applications," *IEEE Trans. Circuits Syst. Video Technol.*, 15(2): 232–242.

Vasconcelos, N. and A. Lippman, "Empirical Bayesian motion segmentation," *IEEE Trans. Pattern Anal. Mach. Intell*, 2(2): 217–221.

Wang, J. and E. Adelson, 1994. "Representing moving images with layers," *IEEE Trans. Image Process*, 3(9): 625–638.

Zeng, W., J. Du, W. Gao and Q. Huang, 2005. "Robust moving object segmentation on H.264/AVC compressed video using the block-based MRF model," *Real-Time Imaging*, 11: 290–299.

Zhong, D. and S.F. Chang, 1999. "An integrated approach for content-based video object segmentation and retrieval," *IEEE Trans. Circuits Syst. Video Technol*, 9(8): 1259–1268.