Concept and Various Types of Video Segmentation

[a] AslamYakubSuriya, [b] S. B. Kishor
[a]Researcher, Ballarpur Institute of Technology, Ballarpur, Maharashtra, India
[b] HOD, Dept. of Computer science, S. P. College, Chandrapur, Maharashtra, India,442402

Abstract

In the present world video information plays a vital role. Storage of video data has become significant for wide variety ofapplications such as in Forensic, Navy and Military, Medical, Multimedia applications.The video segmentation is an important technique for improvement of quality of video on the basis of segmentation. Significance of this segmentation is analysis of video as deformation Video quality improvement and blurring of video. Various authors improve the quality of video using sliding window technique and using different noise filter but all these method are not accurate and suffering for object prediction in noise video. The segmentation task is accomplished through event detection in a frame-by-frame processing setup.

KEYWORDS: Video segmentation, Temporal, clustering, Object Tracking

---

I. Introduction

The video segmentation is a technique of detecting changing frames in video and one of the important techniques required for efficient management of video data [1] several video segmentation algorithms have been proposed. They can be classified into types shape based video segmentation, edge information based video segmentation, image based video segmentation, texture based video segmentation and color based segmentation [2]

Video segmentation is a process to perform on multimedia data. Video refers to a set of technique which accepts video as input. The result of the processing can be new video or data extracted from the input video. In which all the analysis and the designing of simulink model for particular type of segmentation of video work has been done already but we have to design combine (color segmentation and shape base segmentation) simulink model for improvement of quality of video on the basis of segmentation.

Video is a continuous series of picture displayed sequentially at a fixed rate. Video is just a time sequence of images or all the pictures in a video files have equal size the pictures are called frames. Digital video information consists of a series of 25 frames per second. All video processing technique can be applied to frame.

Image segmentation aims to group perceptually similar pixels into regions and is a fundamental problem in computer vision. Video segmentation generalizes this concept to the grouping of pixels into spatio-temporal regions that exhibit coherence in both appearance and motion. Such segmentation is useful for several higher-level vision tasks such as activity recognition, object tracking, content-based retrieval,and visual enhancement. To illustrate the complexity of video segmentation, we identify three major challenges

Temporal coherence: Image segmentation approaches applied to each frame independently produce unstable segmentation results, owing to the fact that even small frameto-frame changes cannot be expressed as a continuous function in general. Consequently, posing video segmentation as spatial region matching problem cannot always enforce consistency of region boundaries over time in the same way as volumetric approaches can. For volumetric techniques, short-term coherence (~5 frames) can be obtained by generalizing image segmentation methods to a 3-D domain. However, we demonstrate that for long-term coherence, it is imperative to go beyond pure pixel-level approaches to a hierarchical approach.

Automatic Processing: Segmenting perceptually homogeneous regions in dynamic scenes is related to tracking regions over time. In contrast to tracking, however, it is not known a priori, which regions to track, what frames contain those regions, or the time-direction for tracking (forward or backward). We develop a fully automatic approach to segmentation, while leaving selection and tracking of specific regions as a post-process that may involve a user.

Scalability: Given the large amount of pixels or features in a video, video segmentation approaches tend to be slow and have a large memory footprint. Consequently, previous advances concentrate on short video sequences (usually less than a second) or reduce complexity, which can adversely affect long-term temporal coherence. We achieve scalability by employing a graph-based approach with linear time complexity and develop memory-efficientalgorithms that enable reliable segmentation of long videos [3].

Generic video segmentation systems are implemented in three stages. First, a set of time-indexed low-level featuresare extracted. For simplicity, we assume that the time index corresponds to the frame index, although this is notrequired. Next, features corresponding to pairs of frames are compared. In the simplest case, the first difference iscomputed by comparing adjacent frames using an appropriate distance measure. Finally, the differences betweenframes are analyzed to detect boundaries, usually with thresholds. Often, thresholding methods are customized forspecific low-level descriptors or inter-frame similarity measures, thus making it difficult to associate performancevariations with the choice of features or similarity measures

There are lot of video segmentation techniques

- Real time segmentation
- Automatic video segmentation
- Efficient hierarchical graph based video segmentation
- Temporal Video Segmentation

Other than the above mentioned methods there is other also

Real time segmentation of moving regions in image sequences is a fundamental step in many vision systems including automated visual surveillance human-machine interface and very low bandwidth telecommunications. Background identification is a common feature in many video processing systems. One of the most important back ground identification algorithm is the Gaussian Mixture Model algorithm (GMM).On implementation of the Gaussian mixture model on FPGA results in Reduction of the processing capability of the overall system. Trainable Segmentation is adapted to improve the processing capability.

Most segmentation methods are based only on colorinformation of pixels in the image. Humans use much moreknowledge than this when doing image segmentation, butimplementing this knowledge would cost considerablecomputation time and would require a hugedomain-knowledge database, which is currently notavailable. In addition to traditional segmentation methods,there are trainable segmentation methods which can modelsome of this knowledge. Neural Network segmentation relies.

On processing small areas of an image using an artificialneural network or a set of neural networks. After suchprocessing the decision-making mechanism marks the areasof an image accordingly to the category recognized by theneural network. A type of network designed especially forthis is the Kohonenmap[4].

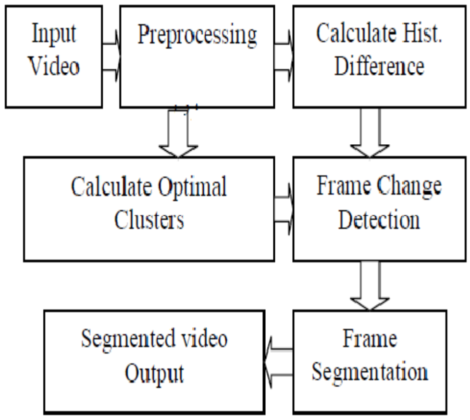

Fig. 1 :Block diagram of Video Segmentation

Graph- Based Algorithm

Spatio-temporal video segmentation builds upon Felzenszwalb and Huttenlocher's [5] graph-based algorithm for image segmentation. We start with a brief overview of their approach. Their objective is to group pixels that exhibit similar appearance, where similarity is based on color difference but also takes the color variation within a region into account. For example, homogeneous regions should not be merged with pixels of different color, but the merging process should be more tolerant to textured regions. Consequently, the notion of internal variation of a region is introduced, whereby regions are merged only if their color difference is less than each region's internal variation. Specifically, for image segmentation, a graph is defined with the pixels as nodes, connected by edges based on 8- neighborhood. Edge weights are derived from the per-pixel normalized color difference. The internal variation Int(R) of a region R is defined as the maximum edge weight emax of its Minimum Spanning Tree (MST):

$$IntR := \max_{e \in MST(R)} w(e)$$

with w(e) being the edge weight of e. The motivating argument is that since the MST spans a region through a set of edges of minimal cost, any other connected set of same cardinality will have at least one edge with weight ≥ emax. Therefore emax defines a lower bound on the maximal internal color variation of the region. We quickly review the original segmentation algorithm.

Initially, a graph is constructed over the entire image, with each pixel p being its own unique region {p}. Subsequently, regions are merged by traversing the edges in a sorted order by increasing weight and evaluating whether the edge weight is smaller than the internal variation of both regions incident to the edge. If true, the regions are merged and the internal variation of the compound region is updated. Sincethe internal variation of a single node is zero (its MST has no edges), only edges of zero weight can cause an initial merge [3].

Our novel video segmentation algorithm addresses all of the above challenges. We build a 3-D graph from the video volume and generalize Felzenszwalb and Huttenlocher's [5] graph-based image segmentation to obtain an initial oversegmentation of the video volume into relatively small space-time regions. Instead of employing a regular grid graph, we use dense optical flow to modify the graph structure along the temporal dimension, accounting for the distortion of the spatio-temporal volume caused by sweeping motions. We propose a hierarchical segmentation scheme that constructs a region graph from the previous level of segmentation and iteratively applies the same segmentation algorithm. By combining a volumetric over-segmentation with a hierarchical re-segmentation, we obtain regions thatexhibit long-term temporal coherence in their identities and boundaries. The use of optical flow as a region descriptor for graph nodes further improves coherence. We use a tree-structure to represent the segmentation hierarchy, effectively enabling subsequent systems to choose the desired granularity post-segmentation, as opposed to re-running the algorithm with different parameters. Granularity could alsobe specified as a desired minimum or average region size, which may be application dependent [3].

I. Temporal Video Segmentation in Uncompressed Domain

The majority of algorithms process uncompressed video. Usually, a similarity measure betweensuccessive images is defined. When two images are sufficiently dissimilar, there may be a cut.

Gradual transitions are found by using cumulative difference measures and more sophisticatedthresholding schemes.

Based on the metrics used to detect the difference between successive frames, the algorithms can bedivided broadly into three categories: pixel, block-based and histogram comparisons.

I.I. Pixel Comparison

Pair-wise pixel comparison (also called template matching) evaluates the differences in intensity orcolor values of corresponding pixels in two successive frames.

The simplest way is to calculate the absolute sum of pixel differences and compare it against athreshold [6]

I.II. Block-based comparison

In contrast to template matching that is based on global image characteristic (pixel by pixeldifferences), block-based approaches use local characteristic to increase the robustness to cameraand object movement. Each frame i is divided into b blocks that are compared with theircorresponding blocks in i+1.

1.3 Histogram comparison

A step further towards reducing sensitivity to camera and object movements can be done bycomparing the histograms of successive images. The idea behind histogram-based approaches isthat two frames with unchanging background and unchanging (although moving) objects will have little differencein their histograms. In addition, histograms are invariant to image rotation andchange slowly under the variations of viewing angle and scale [7]. As a disadvantage one can notethat two images with similar histograms may have completely different content. However, theprobability for such events is low enough, moreover techniques for dealing with this problem have already been proposed in [8].

References

[1] Won-Heekim,Tae-II Jeong and Jong-Nam Kim “Video Segmentation Algorithm Using Threshold and Weighting Based on Moving Sliding Window” ICACT,2009,pp. 1781-1784.

[2] YasiraBeevi CP and Dr.S.Natrajan “An efficient Video Segmentation Algorithm with Real time Adaptive Threshold Technique” International Journal of Signal Processing,Image Processing and Pattern Recognition.Vol.2,No.4,December 2009,pp. 13-26.

[3] M.Grundmann,V. Kwatra,M.Han and I. Essa “Efficient Hierarchical Graph-Based Video Segmentation,” http://www.cc.gatech.edu/cpl/projects/videosegmentation

[4] R. S. Sabeenian and S. Lavanya“High Definition Video Segmentation Techniques – A Review” International Journal of Computer and Electrical Engineering, Vol. 5, No. 6, December 2013,pp. 559-562.

[5] P. Felzenszwalb and D. Huttenlocher.“Efficient graph-based image segmentation”.IJCV, 59(2), 2004.

[6] T. Kikukawa, S. Kawafuchi, Development of an automatic summary editing system for theaudio-visual resources, Transactions on Electronics and Information J75-A (1992) 204-212.

[7] M. J. Swain, Interactive indexing into image databases, in: Proc. SPIE Conf. Storage andRetrieval in Image and Video Databases, 1993, pp.173-187.

[8] G. Pass, R. Zabih, Comparing images using joint histograms, Multimedia Systems (1999),in press.