Video Segmentation Using Global Motion Estimation and Compensation

Pallavi R. Wagh¹, Shubhangi Vaikole²

¹M.E. Student, Department of C.E., Datta Meghe College of Engineering, Airoli, Navi Mumbai, India
²Assistant Professor, Department of C.E., Datta Meghe College of Engineering, Airoli, Navi Mumbai, India

Abstract: Emerging multimedia applications demands content-based video processing. Video has to be segmented into objects for content-based processing. A number of video object segmentation algorithms have been proposed such as semiautomatic and automatic. Semiautomatic methods adds burden to users and also not suitable for some applications. Automatic segmentation systems are still a challenge, although they are required by many applications. The proposed work aims at contributing to develop a system to segment video objects automatically from the background given a sequence of video frames. The proposed work will identify and solve different problems in change detection like uncovered background, global motion of background, temporary poses. To improve the resulting change detection masks, a post-processing method will be used. This technique is used to fill open areas inside object regions with uniform intensity. To improve the boundary of segmentation masks, morphological operations will be used.

Keywords: Automatic segmentation, global motion estimation and compensation, global motion of background, semantic video object

1. Introduction

Two methods are mostly used in video object segmentation, one is semi-automatic, in which some kind of user intervention is required to define the semantic object and other one is automatic, where segmentation is performed without user intervention, but usually with some a priori information. Automatic segmentation of video objects is required by most applications, especially those with real time requirements.

A number of Video object Segmentation algorithms have been proposed, most aiming to specific applications, and trying to fulfill specific requirements. Promising results have been obtained so far in semiautomatic methods, since there is also human assistance in the segmentation process. However, the human assistance involved in these methods is not required because it adds burden to users and also it is not suitable for some applications. On the other hand, fully Automatic Segmentation (AS) systems are still a challenge, although they are required by many applications.

Many automatic segmentation systems are designed for specific problems and with simplified assumptions like videos with fixed background. So it is necessary to have flexible automatic segmentation system for different types of videos. Most of the existing AS systems involves complex techniques. Also each stage of the segmentation process involves computationally intense operations to obtain good segmentation results. Thus, reducing the complexity of the techniques involved is required while keeping accuracy of segmentation results. This can be done by selecting efficient algorithms with reduced computational intensity in each step of the segmentation process. Accuracy of segmentation can be improved by applying post-processing.

2. Related Work

A number of video segmentation algorithms have been proposed. Automatic video object segmentation algorithms developed are roughly classified into three types: the edge-feature based segmentation, spatial-temporal based segmentation, and change detection based segmentation. This section provides a critical review of the various approaches available for video segmentation.

[2] Kim and Hwang(2002) utilises Canny edge detector to find edge information of each frame and it can obtain correct segmentation object for stable moving-object, but this approach will require an absolute background from video sequence and involve a computation-intensive processing.


[4] Guo J, Kim J and Kuo CCJ. (1999) combined colour and motion features for foreground and background separation. Global motion of background is done according to colour similarity using a non-parametric gradient-based iterative colour clustering algorithm called the mean shift algorithm. Moving regions are then identified by a motion detection method, which is developed based on frame intensity difference.


[6]Neri et al. (1998) proposed an algorithm based on change detection which separates potential foreground regions employing a higher order statistics (HOS) significance test to inter-frame differences. Comparison could be performed on a global level, so methods based on histograms were also proposed.
3. Proposed System

Proposed system is used for effective automatic SVO segmentation system. The design considers videos with fixed and moving backgrounds to make the system more flexible. Depending on whether there exists camera motion or not in video sequences, and whether there is an initial background, three cases are examined in the design of the segmentation system. Block diagram of the entire proposed system is shown in Figure and the description of overall working for each of the cases is given below:

A. Statistical Change Detection

Change detection is used in a number of applications including video surveillance, remote sensing, medical diagnosis, driver assistance and video segmentation. However, the application is limited to video segmentation in this discussion. The goal here is to identify the set of pixels that are significantly different between an image and another reference image in video sequences. A comparison is applied between the image under concern and the reference image to detect those pixels. The changes may result from (1) camera movement, (2) appearance/disappearance of objects, (3) motion of object relative to background, (4) shape changes of objects, or (5) changes in illumination.

B. Post-Processing

As mentioned in the above sections, the result of change detection algorithms is not accurate. Due to noise and uncovered background, there exist some noise regions in the initial object mask in both background region and object region. Also object boundary is not very smooth. Therefore a post-processing step to eliminate these noise regions and filter out the ragged boundary is necessary. The post-processing can be applied either to the binary mask only or to both the binary mask and original frame. Post-processing using binary mask only is the simplest approach to remove irregularities. It preserves contours and reduces spurious regions of the mask. It does not use any information from the original image. It consists of simple morphological opening or closing or a more complex composition of morphological filters.

2) Case 2: NO GMOB, without initial background:

In this case, the architecture of the system aims at removing uncovered background by integrating two change detection masks (obtained by forward and backward mask generation). These change detection masks are combined by a logical operator. The operator removes all areas except the foreground object detected which is the region that overlaps in the two change detection masks. This can be explained using Figure 2 as follows. Consider the two change detection masks (forward and backward) shown in Figure 2 a) and b). Region A is uncovered background revealed during forward mask generation, and region C is uncovered background revealed by backward mask generation. Region B is the object which is detected in both change detection masks. When these two masks are combined using a logical AND operator.

Figure 2: Removal of uncovered background by combining two change detection masks a) forward change detection mask; b) backward change detection mask; c) combination of a and b.

A. Frame Normalisation

Perhaps the most critical issue in comparison is selection of the GMOB. Two choices are possible: the previous frame relative to the current frame in the sequence, or an image representing background of the scene. The background frame can be fixed, or can be updated periodically. Fixed background frames can be obtained from the sequence prior to entrance of MOs to the scene. Background can be updated...
periodically by temporal integration of previous and next frames, to produce large background image.

3) Case 3: With GMOB:
For video sequences with moving backgrounds, the system first compensates the global motion. The rest of the segmentation process is just similar to Case-2. The forward and backward Global Motion Estimation and Compensation (GMEC) are applied to align previous and next frames, In-1 and In+1 respectively, towards the current frame In. These two blocks are the Forward GMEC and the Backward GMEC blocks. It is assumed that in video sequences with moving backgrounds, obtaining an initial reference background frame is very difficult and so the system combines three consecutive frames just similar to that of Case-2. For video sequences that contain camera motion, which causes GMOB, this global motion is compensated by the Motion estimation and compensation blocks. These are the Forward GMEC and Backward GMEC blocks in Figure 2. There are three phases of the motion estimation and compensation process proposed in this design: 1) Dense motion vector estimation 2) Parameter estimation 3) Motion Compensation or Frame warping.

a) Dense Motion Vector Estimation:
For the dense motion vector estimation, a three-level Hierarchical block based algorithm is used for its compromised performance in computational complexity, speed and accuracy. Mean pyramids are constructed using the simple averaging equation. Then the pyramidal images are searched from top to bottom using MAD for block matching criterion. A full search within two pixels distance around the concerned block is done for better accuracy of matching block. The most commonly used 16x16 pixels block size is used at level-0 of the hierarchy. The process is shown in Figure 3.

Figure 3: Three-level hierarchical mean pyramid based ME

b) Parameter Estimation
The six parameter motion model, the affine model is used for the global motion. For estimation of the affine parameters, the least squares minimization criterion over the background frame is used to minimize the error between the dense and parametric model motion vectors. The previous background mask Bn-1 is used to estimate the background part of the current frame where the global motion is valid.

c) Frame Warping
The parameters estimated from the above step are used to align the previous/next frame to the current frame. In this process, a new frame is calculated from previous frame by transforming the coordinates of the pixels of the previous frame into a new frame.

4) Region Based Segmentation
The Algorithm Partition the image into regions with common features suitable for further analysis. The Extracted Region Are Uniform with respect to some characteristic, such as intensity, color, texture.

4. Conclusion
This paper presents the video segmentation using global motion estimation and compensation. As in the previous methods there were many drawbacks. To overcome the drawbacks of previous methods for static background, new method is used which is based on moving camera i.e. affine motion model and also find the uniform features by using Region Based Segmentation.

References


