International Journal of Future Innovative Science and Engineering Research (IJFISER)
Volume - 2, Issue - II ISSN (Online): 2454- 1966

Research Manuscript Title

A PARTICLE SWARM OPTIMIZATION BASED VIDEO OBJECT CO-SEGMENTATION

M.Surya Prabha[1], Dr.V.Shunmughavel[2]

PG Scholar, Professor,

Department of Computer Science and Engineering,

V.S.B Engineering College,Karur.

E-Mail: *suryaprabhacse12@gmail.com*, *shunsvel@gmail.com*

JUNE – 2016

www.istpublications.com

A PARTICLE SWARM OPTIMIZATION BASED VIDEO OBJECT CO-SEGMENTATION

M.Surya Prabha[1], Dr.V.Shunmughavel[2]
PG Scholar, Professor,
Department of Computer Science and Engineering,
V.S.B Engineering College,Karur.
E-Mail: *suryaprabhacse12@gmail.com*, *shunsvel@gmail.com*

ABSTRACT

Segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as superpixels). The goal of segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to analyze. The existing system introduces a video object co segmentation framework to discover and segment out common object regions across multiple frames and multiple videos in a joint fashion. In this system the spatio-temporal scale-invariant feature transforms (SIFT) flow descriptor is used to integrate across-video correspondence from the conventional SIFT-flow into inter frame motion flow from optical flow. However it does not captures optimal motion of inter-frame for accurate result. So, in the proposed system a particle swarm optimization (PSO) algorithm is used to capture the optimal inter-frame motion based on the position and velocity updation of the particle. The particle swarm optimization extracts the common object from multiple frames and removes the noisy data from the extracted image. Further it segmented the image accurately. Then consider the tracking image as the input from the multiple frames and group the segmented image to form a single video. The proposed system gives the accurate result which compared to the existing system.

Keywords— Video object segmentation, Energy optimization, object refinement, Particle Swarm Optimization (PSO) flow.

1. INTRODUCTION

Segmentation of a single image is a highly unconstrained problem. Image co-segmentation trades the need for such knowledge for something much easier to obtain, namely additional images showing the object from other view points. The faster growth video data and automatic extraction of object from the multiple videos is very challenging. An image is an array, or a matrix, of square pixels(picture elements) arranged in columns and rows. In a grey scale image each picture element has an assigned intensity that ranges from 0 to 255. A grey scale image is normally called a black and white image, but emphasizes that such an image will also include many shades of grey. Image Processing is a technique to enhance raw images received from cameras/sensors placed on satellites, space probes and aircrafts or pictures taken in normal day-to-day life for various applications such as remote sensing, medical imaging, non-destructive evaluation, textiles, material science, military, film industry, document processing, graphic arts, printing industry, etc.

Analog image processing refers to the alternation of image through electrical means. The most common example is the television image. The television signal is a voltage level which varies in amplitude to represent brightness through the image. Digital image processing are use to process the image. The image will be converted to digital from using a scanner –digitizer and then process it. The term digital image processing generally refers to processing of a two-dimensional picture by a digital computer.

2.RELATED WORK

X. Bai., et al [1] suggested that interactive object segmentation and matting in still images and video is a critical and challenging task that has received significant attention in recent years. Accurately extracting dynamic objects in video remains a very challenging problem. Previous video cut out systems present two major limitations: (1) Reliance on global statistics, thus lacking the ability to deal with complex and diverse scenes; and (2) Treating segmentation as a global optimization, thus lacking a practical workflow that can guarantee the convergence of the systems to the desired results. We present Video Snap Cut, a robust video object cut out system that significantly advances the state-of-the-art. In our system, segmentation is achieved by the collaboration of a set of local classifiers, each adaptively integrating multiple local image features. It consists of a group of overlapping windows around the foreground object boundary, each associated with a local classifier which only segments a local piece of the foreground boundary. The spatial locations of these classifiers are pushed along with the motion vectors estimated across frames, and the segmentation of a new frame is achieved by aggregating local classification results together. Furthermore, each local classifier carries local image features such as color, shape, and motion, and adaptively integrates them together to generate the optimal classification. For applications such as compositing, the binary segmentation results need to be further processed to create soft alpha mattes, a process known as alpha matting. The merit of this paper is to give high quality result. The demerit of this paper is, it does not consider texture feature for further improvement.

Vasileios Mezaris., et al [2] suggested that digital video is an integral part of many newly emerging multimedia applications. New image and video standards, such as MPEG-4 and MPEG-7, do not concentrate only on efficient compression methods, but also providing better way to represent, integrate, and exchange visual information. These efforts aim to provide the user with greater flexibility for "content-based" access and manipulation of multimedia data. Many multimedia applications benefit from this content-based approach, including efficient coding of regions of interest in digital video, personalized user-interactive services, and sophisticated query and retrieval from image and video databases. A novel unsupervised video object segmentation algorithm is presented, aiming to segment a video sequence to objects: spatiotemporal regions representing a meaningful part of the sequence. The proposed algorithm consists of three stages: initial segmentation of the first frame using colour, motion, and position information, based on a variant of the K-Means-with-connectivity- constraint algorithm; a temporal tracking algorithm, using a Bays classifier and rule-based processing to reassign changed pixels to existing regions and to efficiently handle the introduction of new regions; and a trajectory-based region merging procedure that employs the long-term trajectory of regions, rather than the motion at the frame level, so as to group them to objects with different motion. The unsupervised video segmentation algorithm resulting from combining the above algorithms handles fast-moving, newly appearing, and disappearing regions efficiently. The merit of this paper is to give more accurate result. The demerit of this paper is that it does not suitable for complex backgrounds.

Yue Fu., et al [3] said that the system describes a hierarchical approach for object- based motion description of video in terms of object motions and object-to-object interactions. We present a temporal hierarchy for object motion description, which consists of low-level elementary motion units (EMU) and high-level action units (AU). Likewise, object-to-object interactions are decomposed into a hierarchy of low-level elementary reaction units (ERU) and high-level interaction units (IU). A novel way to use dominant affine motion parameters to segment the lifespan of a video object into EMUs. An EMU is a set of consecutive frames within which the dominant motion of the object can be represented by a single parametric model. An ERU is a set of consecutive frames within which two video objects have a predefined interaction. An AU is defined as a time-ordered sequence of EMUs, while an IU is that of ERUs. We then propose an algorithm for temporal segmentation of video objects into EMUs, whose dominant motion can be described by a single representative parametric model. In the proposed framework, the static content of an object-based segment consists of one or more foreground objects and the corresponding background object(s). The motion of each object and a set of object- to-object interactions describe the dynamic content of the segment. The merit of this paper is high-level visual summaries. The demerit of this paper is low accuracy.

Katerina Fragkiadaki., et al [4] said that the goal of this work is to segment a video sequence into moving objects and the world scene. Motion, as the gestaltic principle of "common fate" suggests, is a strong perceptual cue for video segmentation. We propose a novel embedding discretization process that recovers from over-fragmentations by merging clusters according to discontinuity evidence along inter-cluster boundaries. For segmenting articulated objects, we combine motion grouping cues with a centred surround saliency operation, resulting in "context-aware", spatially coherent, saliency maps. Figure-ground segmentation obtained from saliency thresholding, provides object connectedness constraints that alter motion based trajectory affinities, by keeping articulated parts together and separating disconnected in time objects. Centred-surround filtering on per frame flow magnitude has been used by numerous works for spatio-temporal figure-ground segmentation. Our contribution lies in coupling the centred-surround saliency computation with the trajectory embedding. The merit of this paper is the saliency thresholding provides object connectedness constraints. The demerit of this paper is, the optimal motion flow does not described.

Badri Narayan Subudhi., et al [5] says that detection and tracking of moving objects from a video scene is a challenging task in video processing and computer vision. It has wide applications such as video surveillance, event detection, activity recognition; activity based human recognition, fault diagnosis, anomaly detection, robotics, autonomous navigation, dynamic scene analysis, path detection, and others. Moving object detection in a video is the process of identifying different object regions which are moving with respect to the background. We present a novel algorithm for moving object detection and tracking. The proposed algorithm includes two schemes: one for spatio-temporal spatial segmentation and the other for temporal segmentation. Spatial segmentation helps in determining the boundary of the regions in the scene accurately, and temporal segmentation helps in determining the foreground and the background parts of it. For temporal segmentation, a CDM is obtained by taking the difference between two consecutive frames, where information from the previous frame is fed back and the label of the current spatial segmentation result is used to modify the CDM. The modified CDM itself represents a binary mask of foreground and background region where VOP is extracted by superimposing the original pixels of the current frame on the foreground part of the temporal segmentation. A combination of these schemes is used to identify moving objects and to track them. A compound Markov random field (MRF) model is used as the prior image attribute model, which takes care of the spatial distribution of colour, temporal colour coherence and edge map in the temporal frames to obtain a spatio-temporal spatial segmentation. The merit of this paper is it provides better spatial segmentation. The demerit of this paper is, that it does not estimate MRF model parameters and VOP generation of the initial frame.

Yong Jae Lee., et al [6] suggested that video object segmentation is the problem of automatically segmenting the objects in an unannotated video. While the unsupervised form of the problem has received relatively little attention, it is important for many potential applications including video summarization, activity recognition, and video retrieval. We present an approach to discover and segment foreground object(s) in video. Given an unannotated video sequence, the method first identifies object-like regions in any frame according to both static and dynamic cues. Then compute a series of binary partitions among those candidate "key-segments" to discover hypothesis groups with persistent appearance and motion. Finally, using each ranked hypothesis in turn, to estimate a pixel-level object labelling across all frames, where the foreground likelihood depends on both the hypothesis's appearance as well as a novel localization prior based on partial shape matching. The background likelihood depends on cues pulled from the key-segments' (possibly diverse) surroundings observed across the sequence. The merit of this paper is, that it discovers hypothesis groups with persistent appearance and motion. The demerit of this paper is, that the salient object regions does not consider.

Frederic Precioso., et al [7] said that image and video segmentation region-based active contours is used. The goal is to extract image regions corresponding to semantic objects. Image and video segmentation can be cast in a minimization framework by choosing a criterion which includes region and boundary functional. A parametric active contour method based on B-Spine interpolation has been proposed to highly reduce the computational cost, but this method is sensitive to noise. The curves preserve the implementation advantages as the B-spines while softening the interpolation constraint. The relaxation of the interpolation condition is traded for an optimal increase of the smoothness of the spine snake. A smoothness parameter controls the amount of relaxation. The merit of this paper is low computational cost. The demerit of this paper is that it does not give robustness.

D. Baswaraj., et al [8] suggested that image segmentation is a basic task, responsible for the separating process. The function of segmentation is to dividing an image into its basic and disjoint sub-regions, which are identical according to their property, e.g. intensity, colour, and quality. Segmentation algorithms are usually based on either discontinuity with sub regions, i.e. edges, or equality within a sub-region, though there are a few segmentation algorithms depends on both discontinuity and equality. Active contours have been widely used as attractive image segmentation methods because they always produce sub-regions with continuous boundaries, while the kernel-based edge detection methods, e.g. Sobel edge detectors, often produce discontinuous boundaries. The merit of this paper is high performance. The demerit of this paper is local minimum problems still exist.

3. SYSTEM DESIGN

EXISTING SYSTEM

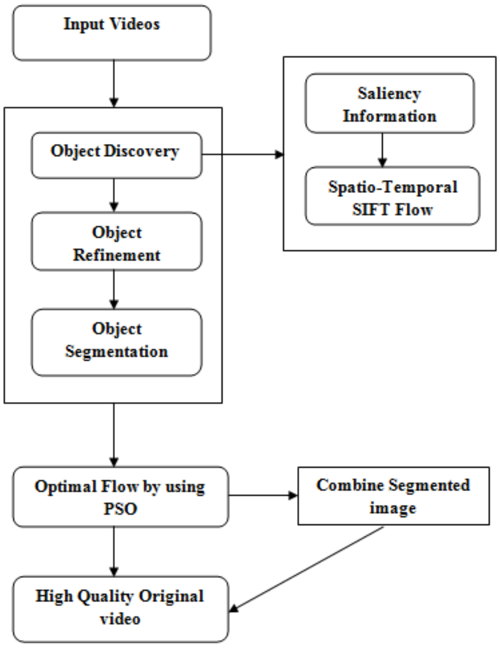
With the faster growth of video data, efficient and automatic extraction of the interest object from multiple videos is quite important and very challenging. Maybe these objects of interest exhibit drastically different in their appearance or motions. The proposed system is used to jointly segment multiple videos containing a common object in an unsupervised manner. In this process, we use a spatio-temporal SIFT flow that integrates optical flow, which captures inter-frame motion, and conventional SIFT flow, which captures across-videos correspondence information. The algorithm has three main stages: object discovery among multiple videos, object refinement between video pairs, and object segmentation on each video sequence. Here, to present a co segmentation framework to discover and segment out common object regions across multiple frames and multiple videos in a joint fashion. The limitations of the existing system is optimal flow not efficient and produce low accuracy.

PROPOSED SYSTEM

Optical flow methods are accurate algorithms for estimating motion of objects, being their performance dependent on the configuration of a set of parameters. Optimal motion estimation is important for effective Co segmentation of the Video Object. Here, inter-frame motion is estimated by using Particle swarm optimization. This can be used as a search algorithm based on stochastic processes , where the learning of social behaviour allows each possible solution (particle) 'fly' onto that space (swarm) looking for other particles that have the best features and thus minimizing or maximizing the objective function. The particle of the swarm fly through hyperspace and have two essential reasoning capabilities: their memory of their own best position - local best (lb) and knowledge of the global or their neighbourhood's best - global best (gb). Position of the particle is influenced by velocity. The position of the particle is changed by adding a velocity, to the current position

$$x_i\,(t+1) = x_i\,(t) + v_i\,(t+1)$$

Let denote the position of particle I in the search space at time step t ; unless otherwise stated, t denotes discrete time steps. All particles move towards the optimal point with a velocity. Initially all of the particle velocity is assumed to be zero. This mechanism can be summarized in three principles: (i) evaluation, (ii) comparison, and (iii) imitation. Each particle can evaluate others within your neighbourhood through some objective function; it can compare with own value and finally decide whether it is a good choice to imitate it or not. The merits of the proposed system is optimal flow is efficient and produce high accuracy.



1.1 **System Achitecture**

4. SYSTEM IMPLEMENTATION

OBJECT DISCOVERY

In this module, the method explores the video dataset structure and associates the global information with the intra-frame information like saliency to discover the common object from multiple videos, even in the presence of some frames without the common object. Three main properties of targeted object are helpful for object discovery: a) intra-frame saliency–the pixels of foreground should be relatively dissimilar to other pixels within a frame; b) inter-frame consistency–the pixels of foreground should be more consistent within a video; c) across-video similarity–the pixels of foreground should be more similar to other pixels between different videos (with possible changes in color, size and position).To propose a new spatio-temporal SIFT flow algorithm that integrates saliency, SIFT flow and optical flow to explore the correspondences between different videos.

OBJECT REFINEMENT

First, a pair of videos is randomly selected from dataset. The spatio-temporal SIFT flow between the frames is constructed. As discontinuities of spatio-temporal SIFT flow field reflect the variation of object structure (but not color variation) yet robust to object details. This property of spatio-temporal SIFT flow field is very important. Through the computation of the discontinuities of spatio-temporal SIFT flow field, divide the object-like area into a few regions depending on the structure variation. This enables us to estimate every part of the object-like area whether belongs to foreground using GMMs. Based on the visualization of spatio-temporal SIFT flow field, numerous over-segmentation methods can be introduced and the object-like area can be efficiently partitioned into regions.

OBJECT SEGMENTATION BY OPTIMIZATION

Once the correct estimations for foreground of each video are obtained, a graph-cut based method is employed to get per-pixel segmentation results. we select frame every other five or ten frames from video. After the object refinement process, we get more correct estimation for common object and update the appearance model of the object and background for frame, which can be used to conduct the segmentation in next five or ten frames. For frame, we obtain the likelihood of pixel for foreground as *pin* using our appearance models estimated by its temporally nearest frame. For video., we update the labelling for all pixels to obtain the final segmentation results through an object segmentation function. This object segmentation function based on spatio-temporal graph by connecting frames temporally.

OPTIMAL FLOW BY USING PSO

The proposed method a particle swarm is used which captures the optimal inter-frame motion based on the position and velocity updation of the particle. In this optimization process, we use a spatio-temporal SIFT flow that integrates optical flow, which captures inter-frame motion, and conventional SIFT flow, which captures across-videos correspondence information.

Here inter-frame motion is estimated by using Particle swarm optimization. Particle swarm optimization (PSO) is a computation method that optimizes a problem by iteratively trying to increase a candidate solution with regard to a given measure of quality. This uses a number of particles that set up a swarm moving everywhere in an N dimensional search space looking for the best solution. Every particle takes track of its coordinates in the solution space, which are related with the best solution that is achieved to this point by that particle is called as personal best position (pbest) and the other best value achieved until now by any particle in the neighbourhood of that particle is called as global best position (gbest).

All particles move towards the optimal point with a velocity. Initially all of the particle velocity is assumed to be zero. To improve the searching effectiveness, initially the object model is projected into a high-dimensional feature space, and then PSO-based algorithm is used to search over the high dimensional space and congregate to some global features of the object. This PSO based algorithm considers even the inter frame motion estimation to speed up the searching procedure. The proposed algorithm can be used to estimate the inter-frame motion at each pixel in a video sequence. This novel spatio-temporal SIFT flow generates reliable estimations of common foregrounds over the entire video data set.

HIGH QUALITY ORIGINAL IMAGE

By utilizing object discovery, object refinement, object segmentation, optimal flow by using PSO process the common objects in the video are segmented. In this module, the segmented objects are combined to form an original video with high quality.

5. CONCLUSION

The proposed video co-segmentation method discovers the common object over an entire video dataset and segments out the objects from the complex backgrounds. The optimization process consists of object discovery; object refinement and object segmentation which are executed on the whole set of videos. In the proposed method, a PSO method is used to capture the optimal inter-frame motion based on the position and velocity updation of the particle. To achieve optimization process, we use a spatio-temporal PSO flow that integrates inter-frame motion process, and across-videos correspondence information. Finally a novel object discovery energy function is proposed to discover the common object with this situation by utilizing the proposed spatio-temporal PSO flow and those properties of foreground object. Both the quantitative and qualitative experimental results have shown that the proposed algorithm creates more reliable and accurate video co-segmentation performance than the existing system.

REFERENCES

[1] X. Bai, J. Wang, D. Simons, and G. Sapiro, "Video SnapCut: Robust video object cutout using localized classifiers," *ACM Trans. Graph.,*vol. 28, no. 3, Aug. 2009, Art. ID 70.

[2] T. Brox and J. Malik, "Object segmentation by long term analysis of point trajectories," in *Proc. ECCV, 2010, pp. 282–295.*

[3] M. Grundmann, V. Kwatra, M. Han, and I. Essa, "Efficient hierarchical graph-based video segmentation," in *Proc. IEEE Conf. CVPR, Jun. 2010,*pp. 2141–2148.

[4] K. Fragkiadaki, G. Zhang, and J. Shi, "Video segmentation by tracing discontinuities in a trajectory embedding," in *Proc. IEEE CVPR,* Jun. 2012, pp. 1846–1853.

[5] D. Tsai, M. Flagg, and J. Rehg, "Motion coherent tracking with multilabel MRF optimization," in *Proc. BMVC, 2010, pp. 56.1–56.11.*

[6] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski,"A database and evaluation methodology for optical flow," *Int. J.Comput. Vis.*, vol. 92, no. 1, pp. 1–31, Mar2011.

[7] L. S. Silva and J. Scharcanski, "Video segmentation based on motion coherence of particles in a video sequence," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 1036–1049, Apr. 2010.

[8] D. Tsai, M. Flagg, and J. Rehg, "Motion coherent tracking with multilabel MRF optimization," in *Proc. BMVC*, 2010, pp. 56.1–56.11.

[9] T. Wang and J. Collomosse, "Probabilistic motion diffusion of labeling priors for coherent video segmentation," *IEEE Trans. Multimedia*, vol. 14, no. 2, pp. 389–400, Apr. 2012.

[10] Y. J. Lee, J. Kim, and K. Grauman, "Key-segments for video objectsegmentation," in *Proc. IEEE ICCV*, Nov. 2011, pp. 1995–2002.

[11] T. Ma and L. J. Latecki, "Maximum weight cliques with mutex constraints for video object segmentation," in *Proc. IEEE CVPR*, Jun. 2012,pp. 670–677.

[12] C. Liu, J. Yuen, and A. Torralba, "SIFT flow: Dense correspondenceacross scenes and its applications," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 33, no. 5, pp. 978–994, May 2011.

[13] W. Wang, J. Shen, and F. Porikli, "Saliency-aware geodesic video object segmentation," in *Proc. IEEE CVPR*, Jun. 2015, pp. 3395–3402.

[14] Y. Huang, Q. Liu, and D. Metaxas, "Video object segmentation byhypergraph cut," in *Proc. IEEE CVPR*, Jun. 2009, pp. 1738–1745.