ISSN 2348 – 7968

Segmentation of Video Object by Using Scale Invariant Feature Transform

M.Surya Prabha[1] and Dr.A.Padma [3]

[1] PG Scholar, CSE, V.S.B Engineering College,
Karur,Tamilnadu,India

[2] CSE, V.S.B Engineering College,
Karur, Tamilnadu, India

Abstract

Video object segmentation is a challenging problem. Without human annotation or other prior information, it is hard to select a meaningful primary object from a single video, so extracting the primary object across videos is a more promising approach. The existing system introduces a 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 to integrate across-video correspondence from the conventional SIFT-flow into inters frame motion flow from optical flow. However it does not captures optimal motion of inter-frame for accurate result. So, in 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. This novel spatio-temporal SIFT flow generates reliable estimations of common foregrounds over the entire video data set. The experimental results show that the proposed system achieves high accuracy compared with the existing system.

***Keywords: Video object segmentation, Energy optimization, object refinement, spatio-temporal Scale-invariant feature transform (SIFT) 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

Many techniques have been proposed to detect the particular object from the multiple videos. For example, foreground appearance or motions from various videos are much different and low contrast with the background. This cause difficulties on the existing video segmentation [2], [3], [4], [5], [6], [8], they are benefit from visual cues as motion or appearance. The lack of taking into account the joint information between videos leads to unsatisfactory.

In previous object segmentation methods for a single video, video co-segmentation has been proposed to extract the common object from a set of related videos. The few methods are till in problem [12], [13], [15]. Rubio et al. [12] make that the foreground objects from different videos have similar motion patterns and similar appearance model which is distinct from the background. Chen et al. [13] emphasize that the coherent motion of regions and similar appearance to conduct the segmentation.These approaches [12], [13] is the set of videos is assumed to similar background and foreground. The underlying properties of video objects in three levels: intra-frame saliency, inter-frame consistency and across-video correspondence.

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.

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

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 .

3. System Design

Existing System

In existing scenario, the co-segmentation framework is used for detecting and segmenting out common object from multiple related videos. The spatio-temporal SIFT flow and inter frame consistency to discover the common objects. To optimize the inter frame motion process, the particle swarm optimization is used.

With the faster growth of video data, efficient and automatic extraction of the interest object from multiple videos is quite important and very challenging. May be 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 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 demerits of this existing system is optical flow does not efficient and produce a 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 behavior 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

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

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 neighborhood through some objective function; it can compare with own value and finally decide whether it is a good choice to imitate it or not.

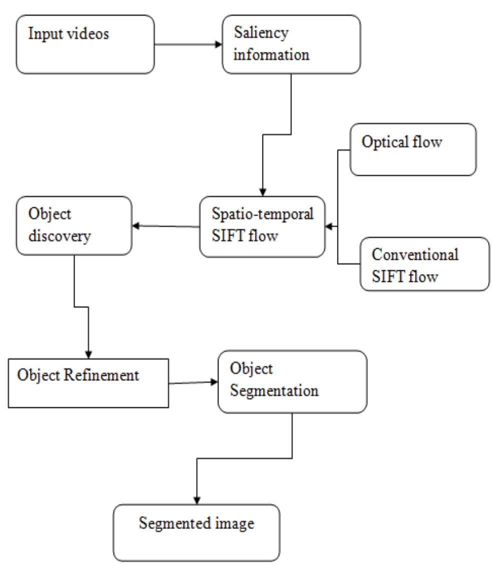

Figure 3.1: System Architecture

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.

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 SIFT 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 SIFT 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 SIFT 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

Acknowledgments

Ms. SuryaPrabha - Received the Bachelor Degree in Computer Science and Engineering from M.P.nachimuthu M.Jaganathan Engineering College, Erode in 2014. Currently I am doing Master of Engineering in Computer Science at V.S.B Engineering College, Karur under Anna University of chennai. My research interests include Image Processing and Network Security.I participated in several workshops and presented papers.

Padma received her B.E in Computer science and Engineering from Madurai Kamaraj University ,Tamilnadu, India in the year 1990. She received her M.E in Computer science and Engineering from Madurai Kamaraj University, Tamilnadu, India in 1the year 997. She received her P.hD in the Department of Information and Communication Engineering from Anna University, Chennai in the year 2014. She is working as a Professor in the Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, India Tamilnadu. Her area of interest includes Image processing, Neural Networks. She has published more than 10 Journals in International level.

References

1.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.

2. 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.

3. D. Tsai, M. Flagg, and J. Rehg, “Motion coherent tracking with multilabel MRF optimization,” in *Proc. BMVC*, 2010, pp. 56.1–56.11.

4. 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.

5. Y. J. Lee, J. Kim, and K. Grauman, “Key-segments for video objectsegmentation,” in *Proc. IEEE ICCV*, Nov. 2011, pp. 1995–2002.

6. 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.

7. 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.

8. W. Wang, J. Shen, and F. Porikli, “Saliency-aware geodesic video object segmentation,” in *Proc. IEEE CVPR*, Jun. 2015, pp. 3395–3402.

9. Y. Huang, Q. Liu, and D. Metaxas, “Video object segmentation byhypergraph cut,” in *Proc. IEEE CVPR*, Jun. 2009, pp. 1738–1745.

10. T. Brox and J. Malik, “Large displacement optical flow: Descriptor matching in variational motion estimation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 500–513, Mar. 2011.

11. I. Endres and D. Hoiem, “Category independent object proposals,” in *Proc. ECCV*, 2010, pp. 575–588.

12. J. C. Rubio, J. Serrat, and A. López, “Video Co-segmentation,” in *Proc. ACCV*, 2012, pp. 13–24.