Indian Journal of Innovations and Developments Vol 5 (4), April, 2016 ISSN (online) : 2277-5390
ISSN (Print): 2277-5382

Video object co-segmentation

M.Suryaprabha*[1], Dr.V.Shunmughavel[2]

[1]*PG Scholar, Department of CSE, V.S.B Engineering College, Karur, Tamilnadu*
[2]*Professor, Department of CSE, V.S.B Engineering College, Karur, Tamilnadu*
suryaprabhacse12@gmail.com[1], shunsvel@gmail.com[2]

Abstract

Objective: The main motivation of this research is to discover and segment out common object regions in different videos.
Methods: The proposed system introduced a spatio-temporal scale-invariant feature transforms (SIFT) flow descriptor which is used to incorporate across-video correspondence. In order to improve the system performance particle swarm optimization (PSO) 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.
Findings: The experimental results show that the proposed system achieves better performance compared with existing system in terms of accuracy, precision, recall and f-measure.
Improvement: The proposed algorithm increases the overall system performances by spatio-temporal scale-invariant feature transform flow descriptor and particle swarm optimization algorithm prominently.

Key word: salient object, Object Refinement, spatio-temporal SIFT flow and particle swarms

1. Introduction

Object segmentation plays an essential and significant role in solving lot of high-level vision troubles, such as object detection, object recognition and scene understanding [1]. Video object segmentation is the difficulty of routinely segmenting the objects in an unannotated vides [2]. With increasing of video data, well-organized and automatic extraction of the interest object from multiple videos is significant and very difficult process [3]. The problem of simultaneously segmenting a common category of objects from two or more videos is known as video object co-segmentation [4]. Various methods are implemented to harness such information for video object co-segmentation.

In [5] introduced a new model for Temporal Video Objects Segmentation. This system present a temporal hierarchy model for object motion description. The object based segmentation method having one or more foreground objects and the corresponding background object(s). The temporal hierarchy model consists of low-level elementary motion units (EMU) and high-level action units (AU). To segment video object into EMU a new algorithm is proposed. By using single representative parametric model the dominant motion can be determined.

In [6] introduced a new Video Object Segmentation method which is based on Key-Segments. Unannotated video is taken as an input. Based on the static and dynamic cues the video object segmentation method finds out object-like regions in any frame**.** To find out hypothesis groups with persistent appearance and motion, perform a serious of binary separation between those Candidates. Each ranked partition robotically defines a foreground and background model. Finally we have to compute pixel-level object.

In [7] introduced a new Multiple View Object Co segmentation method using Appearance and Stereo Cues. A new piecewise planar layer-based stereo algorithm determine the dense depth map. It consists of set of 3D planar surfaces. By using energy minimization framework the algorithm is computed. This algorithm combines stereo and appearance cues. The planar surface is considered as structural elements of the scene. The segmentation is processed by fusing information across multiple views.

In [8] introduced a Video Object Detection and Tracking mechanism which is based on the A Change Information Based Fast Algorithm. The designed algorithm consists of two phases. In first phase spatio-temporal spatial

ISSN (online) : 2277-5390
ISSN (Print): 2277-5382

segmentation is performed and temporal segmentation is performed in another phase. These two schemes are combined for find out and track the moving objects. The Markov random field (MRF) model [9] is used for take care of spatial allocation of color, temporal color coherence. This is for spatio-temporal spatial segmentation [10].

2. Proposed methodology

The proposed methodology consists of the following phases. They are object discovery, object refinement and object segmentation. Here spatio-temporal SIFT flow incorporates optical Flow and conventional SIFT flow.

Object Discovery

To discover the common object region saliency and spatio-temporal SIFT flows are used. The following properties also used for object discovery.

1. Intra-frame saliency– the foreground pixel in frames is different to others.
2. Inter-frame consistency– the foreground pixels are more reliable within a video;
3. Across-video similarity– the foreground pixel is related to other pixels between different videos

To find out relatedness between various videos spatio-temporal SIFT flow algorithm is introduced, which incorporates saliency, SIFT flow and optical flows.

Here, **V** = {V_1, V_2, ..., V_N } which is set of N input videos. **F***n* = {F_n^1, F_n^2 , ..., F_n^i, ...} is a set of frames belong that video V V_n. we have to calculate normalized saliency map $m_n^i$for frame F_n^i . To determine foreground or background pixel cost of labeling the saliency term is constructed.

$$A_n^i(x) = \exp -\{ M_n^i (x)\}.\ l_n^i (x) + \exp -\{1-\{M_n^i(x)\} .(1-l_n^i (x))$$

Where,

M_n^i $-$ Saliency map

Optical flow is defined as a 2D vector,. Based on the colour consistency optical flow denotes the pixel motion information. It reflect the inter frame motion of the pixel. The flow field between frame F_n^i and F_n^{i+1} described by v_n^i. To construct dense correspondence map athwart various scenes and object appearances SIFT flow can be used. The consistent correspondences $w_{nn'}^{ii'}$ = ($u_{nn'}^{ii'}$, $v_{nn'}^{ii'}$) between the pixels of frame $F_n^i$and $F_{n'}^{i'}$ from various videos are recognized through spatio-temporal SIFT flow.

Through spatio-temporal SIFT flow, consistent correspondences $w_{nn'}^{ii'}$ = ($u_{nn'}^{ii'}$, $v_{nn'}^{ii'}$) between the pixels of frame $F_n^i$and $F_{n'}^{i'}$ from different videos are established. In other words, pixel **x** of frame $F_{n'}^{i'}$ is associated with the pixel **x**+$\mathbf{w_{nn'}^{ii'}}$*(***x***)* of frame$F_{n'}^{i'}$.

The energy functions for spatio-temporal SIFT flow is

E= $E_s + \alpha_1 E_{os} + \alpha_2 E_{Disp} + \alpha_3 E_{smooth} + \alpha_4 E_{sal}$

The optical flow compensated SIFT is

$$E_{os}\left(w_{nn'}^{ii'}\right) = \sum_{x \in R_n^i} ||\ S_n^{i+1}(X + V_n^i(x)) - S_{n''}^{i''+1}\left(xw_{nn'}^{ii'} + V_{n'}^{'i}(xw_{nn'}^{ii'})\right)$$

Saliency term $E_{sal}(w_{nn'}^{ii'})$= $\sum_{x \in R_n^i} M_{n'}^{i'}(xw_{nn'}^{ii'}(x))) + (1 - M_{n'}^{i'+1}\ (xw_{nn'}^{ii'}(x))) + (V_{n'}^{i'}(xw_{nn'}^{ii'}))) + \quad w_{nn'}^{ii'}(x)))$

To characterize the common object appearance Gaussian mixture models (GMM) is used.
The GMM of the frame $f_n^k$is set as:

$$GMM_{f_n^k}^{f} = GMM_{f_n^{k-1}}^{f}$$

$$GMM_{f_n^k}^{b} = GMM_{f_n^{k-1}}^{b}$$

ISSN (online) : 2277-5390
ISSN (Print): 2277-5382

Object Refinement

To refine the estimated object regions the object refinement process. Based on their variations the objects are divided into sub regions for filter background pixels. The pair of videos (V_n , $V_{n'}$) is arbitrary selected from dataset. And spatio-temporal SIFT flow between frames $f_n^k and f_{n'}^k$)is computed. To represent image local spatial structure the texture region with background and object area using LBP features are compared. Two normalized histograms are computed to design texture of foreground and background in frames. Thus the probability *of* pixels $\mathbf{x}_t \in \mathbf{t}$ for foreground is computed through LBP histograms which is follows as:

$$l_n^k\ (x_t) = \frac{H_t^f[x_t]}{H_t^f[x_t] + H_t^b[x_t]}$$

Where

$H_t^f\ [x_t]$ - Value of histogram H_t^f at pixel $\mathbf{x}_t$.

We have to update $\{ GMM_{f_n^k}^f, GMM_{f_n^k}^b \}$ after the frame refinement. It gives direction for the object segmentation process. To achieve accurate estimation for foreground object this refinement process is used.

Object Segmentation by Optimization

The correct estimation for object in each video, the appearance of foreground object and background $\{GMM_{f_n^k}^f, GMM_{f_n^k}^b\}$ for frame f_n^k. It is very useful for segmentation in next five or ten frames of f_n^2. After the foreground estimation the graph-cut based method is used for segmentation. To achieve final segmentation result we have to update the labeling $\{ l_n^i \}$ for all pixels in a video. Based on spatio temporal graph, the object segmentation is defined as

$$F_n(x) = \sum_i \{\sum_x U_n^i(x) + \gamma_1 \sum_{x,y\ \epsilon N_s} v_n^i(x,y) + \gamma_2 \quad \sum_{x,y\ \in N_t} W_n^i\ (x,y) W_n^i(x,y)\ \}$$

Where

$\mathbf{N}_s$ have all the 8-neighbors within one frame and the set $\mathbf{N}_t$ has backward nine neighbors in pairs of adjacent frames.
γ -Positive coefficient

Compute optimal flow by using PSO

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. Here inter-frame motion is estimated by using Particle swarm optimization. Particle swarm optimization (PSO) is a computation technique that optimizes a problem by iteratively trying to enlarge 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 are moved towards the optimal point with a velocity. 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.

ISSN (online) : 2277-5390
ISSN (Print): 2277-5382

Algorithm

1. Initialize particles
2. For each particles
3. do
4. if f(x_i) < f (pb_i) then
5. $pb_i = x_i$
6. end if
7. update global best position
8. if f (pb_i) < f (gb) then
9. gb= pb_i
10. end if
11. end for
12. update particle velocity and position
13. For each particle i
14. Do
15. $V_{i,d} = V_{i,d} + c_1 * Rand(0,1) * [pb_{i,d} – x_{i,d}] + c_2 * Rand(0,1) * [gb_d – x_{i,d}]$
16. $x_{i,d} = x_{i,d} + V_{i,d}$
17. end for
18. end for
19. it= it+1
20. until it >MAX_Iterations

High Quality Original Image

By utilizing object discovery, object refinement and object segmentation 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

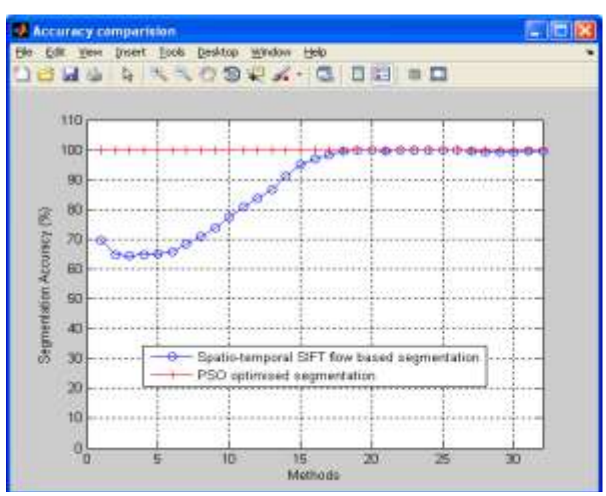
3. Experimental results and Discussion

In existing system, spatio-temporal SIFT flow based segmentation is used. In proposed system, the PSO based segmentation is used. The experimental results show that the proposed system achieves high performance compared with existing system in terms of segmentation accuracy.

Segmentation Accuracy comparison

The degree to which the delineation of the object corresponds to the truth

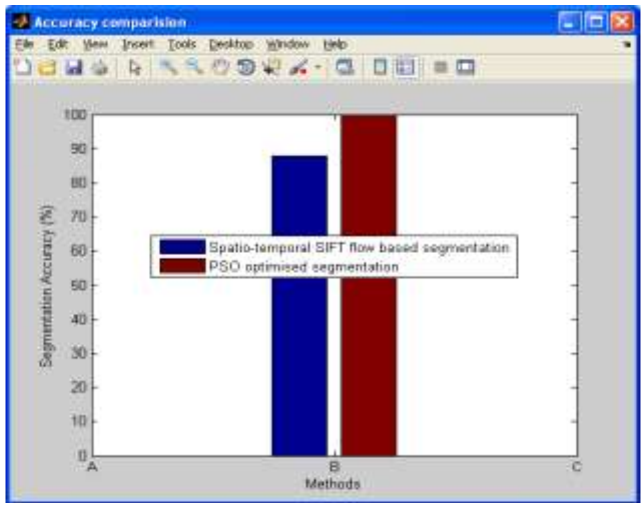
Figure 1. Segmentation Accuracy comparison



ISSN (online) : 2277-5390
ISSN (Print): 2277-5382

(Figure 1) In this graph, x axis will be the number of input samples and y axis will be segmentation accuracy in %. In existing system, spatio-temporal SIFT flow based segmentation is used. In proposed system, PSO based segmentation is used. From the graph see that, segmentation accuracy of the proposed PSO based segmentation is better than existing one.

Figure 2. Segmentation Accuracy comparison



(Figure 2) In this graph, x axis will be the two approaches of segmentation and y axis will be segmentation accuracy in %. In existing system, spatio-temporal SIFT flow based segmentation is used. In proposed system, PSO based segmentation is used. From the graph see that, segmentation accuracy of the proposed PSO based segmentation is better than existing one.

4. Conclusion

The proposed systems find out the common object over an entire video dataset. And also segment the objects from backgrounds. The segmentation process consists of object discovery; object refinement and object segmentation phases. Then a particle swarm optimization captures 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, the segmented objects are combined to form a high quality video. The experimental results show that the proposed system achieves higher performance compared with existing system in terms of segmentation accuracy.

5. References

1. Hong Li, Bee June Tye, Ee Ping Ong, Weisi Lin, Chi Chung Ko. Multiple motion object segmentation based on homogenous region merging. In IEEE International Symposium on Circuits and Systems.2011; no. 5, 175-178.
2. Alex Levinshtein. Low and mid-level shape priors for image segmentation. University of Toronto, 2010.
3. AnestisPapazoglou, Vittorio Ferrari. Fast object segmentation in unconstrained video. In Proceedings of the IEEE International Conference on Computer Vision, 2013; 1777-1784.
4. JiamingGuo, Zhuwen Li, Loong-Fah Cheong, Steven Zhiying Zhou. Video Co-segmentation for Meaningful Action Extraction. IEEE International Conference on Computer Vision, 2013; 2232-2239.
5. Yue Fu, AhmetEkin, A. Murat Tekalp, Rajiv Mehrotra. Temporal Segmentation of Video Objects for Hierarchical Object-Based Motion Description. IEEE Transactions on image processing, 2001; 11(2), 135-145.
6. Yong Jae Lee, Jaechul Kim, Kristen Grauman. Key-Segments for Video Object Segmentation. Proceedings of the International Conference on Computer Vision (ICCV), 2011; 1995-2002.

ISSN (online) : 2277-5390
ISSN (Print): 2277-5382

7. AdarshKowdle, Sudipta N. Sinha, Richard Szeliski. Multiple View Object Cosegmentation usingAppearance and Stereo Cues. *Computer Vision–ECCV 2012*; 789-803.
8. Badri Narayan Subudhi, Pradipta Kumar Nanda Ashish Ghosh. A Change Information Based Fast Algorithm for Video Object Detection and Tracking. IEEE Transactions on circuits and systems for video technology. 2011; 21(7), 993-1004.
9. Shaikh RA, Li JP, Khan A, Kumar K. Content Analysis using Shape and Spatial Layout with Markov Random Field. Indian Journal of Science and Technology, 2016 Mar 3;9(7),1-6.
10. M. Sindhana Devi, M. Soranamagesweri. An Analysis of Various Noise Removal Techniques to Enhance Image Quality, Indian Journal of Innovations & Developments, 2015; 4(7), 1-7.

The Publication fee is defrayed by Indian Society for Education and Environment (iSee). www.iseeadyar.org

Citation:

M.Suryaprabha, Dr.V.Shunmughavel. Video object cosegmantation. *Indian Journal of Innovations and Developments.* 2016; 5 (4), April.