International Journal of Advanced Computer Research (ISSN (print):2249-7277 ISSN (online):2277-7970)
Volume-3 Number-2 Issue-10 June-2013

Video Object Tracking based on Automatic Background Segmentation and updating using RBF neural network

Pushpender Prasad Chaturvedi[1], Amit Singh Rajput[2], Aabha Jain[3]

Abstract

In this paper, the problems associated with the automatic object segmentation of the video sequences are considered. Towards this objective, a unique method that combines of image and video processing techniques ranged from noise filtering to data clustering is developed. The method also addresses a number of challenging issues along with computational complexity, accuracy, generality, and robustness. One of the primary aims of this paper is to find segmentation of color, texture, motion, shape, frame difference, and other methods of video segmentation for automatic detection considering the real-time processing requirements. In contrast to frame-wise tracking techniques, the employment of a spatiotemporal data that is constructed from multiple video frames introduces new degrees of freedom that can be exploited in terms of object extraction and content analysis. The current notions of region segmentation are extended to the spatiotemporal domain, and new models to estimate the object motion are derived.

Keywords

Video Processing, Wavelet, RBF

I. Introduction

Video object tracking play an important role in security surveillance in current scenario. The explosion of successful digital device, the ease of use of high quality and economical video cameras, and the increasing need for computerized video analysis has generated a great deal of interest in video tracking methods. There are three techniques for video analysis: exposure of interesting moving target, tracking of such target from frame to frame, and analysis of target tracks to identify their activities [1].

Pushpender Prasad Chaturvedi, Department of Electronics & Communication, RITS, Bhopal, India.
Amit Singh Rajput, Department of Electronics & Communication, RITS, Bhopal, India.
Aabha Jain, Department of Electronics & Communication, RITS, Bhopal, India.

The successful video object tracking system faced a problem of false detection of moving video object. The false video object detection arises due to drastic change of background of moving video. For the maintenance of background updating various authors proposed a method for automatic background updating. The process of segmentation and clustering technique improved the automatic background updating of targeted video frame. The variation of frame cycle induced the problem of background updating for tracking of object[3]. Backgroundsegmentation plays an important role in video object tracking. The automatic changing of background creates a difficulty for capturing and tracking of object, due to automatic background loss of frame and generates error. The generation of error creates a false detection of object tracking. Video tracking is one of the most important applications in computer vision, and has been widely applied to traffic surveillance system, suspicious person monitoring system etc. In practical application, since the camera moves and rotates, it needs to track objects in a dynamical background. How to select the initial target objects automatically and establish objects' motion model, and how to update object and background models at each frame are the key in real-time visual tracking with an active camera [5]. Recent years have seen significant progress in background segmentation using portion and clustering technique for video object tracking. Background segmentation is to exploit features in a low-dimensional space for object detection. However, the computational complexity is likely to increase significantly as a result of low dimensionality of features. Since object tracking can be posed as a binary classification problem with the goal to separate the target object from the background, a discriminative object representation scheme greatly facilitates this task. Therefore, feature selection is of crucial importance for generating an effective low-dimensional discriminative subspace. The creation of noise due to frame loss of video and free environments of camera distance of object. Noise induced a problem during updating a background process of video tracking. Video segmentation are crucial factor in video tracking, the part of segmentation generates temporal video segmentation. The majority of algorithms process uncompressed video [7]. Usually, a similarity measure between successive 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 sophisticated threshold schemes. Based on the metrics used to detect the difference between successive frames, the algorithms can be divided broadly into three categories: pixel, block-based and histogram comparisons. Pair-wise pixel comparison evaluates the differences in intensity or color values of corresponding pixels in two successive frames.

A step further towards reducing sensitivity to camera and object movements can be done by comparing the histograms of successive images[8]. The idea behind histogram-based approaches is that two frames with unchanging background and unchanging (although moving) objects will have little difference in their histograms. In addition, histograms are invariant to image rotation and change slowly under the variations of viewing angle and scale.The segmentation process attempt form two or more process such as edge detection, feature extraction etc. The improved background updating factor varies during frequent change of frame cycle and loss of frame and detect fake object, so improvement of video background for automatic updating we apply RBF neural network classification technique for segmentation process for background detection. The majority voting process of RBF neural network classification maintains the frequent change of background of video tracking object. RBF neural network classification is extended form of neural network classification. The rest of this paper is organized as follows. In section II related technique for video tracking. Section III gives a proposed method. Section IV experimental result analysis V concludes this paper.

II. Related Work

In this section we discuss some related work with video object tracking method based on filter and segmentation method. The process of video segmentation and updating process proposed by different approach. Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn and Andrew Zisserman entitled "The PASCALVisual Object Classes (VOC) Challenge"Authors has evaluated methods for both detection and classification which analyses whether the methods are statistically different, what they are learning from the images, and what the methods find easy or confuse[1]. Karim Ali, David Hasler and Francois Fleuret entitled "FlowBoost – Appearance Learning from Sparsely Annotated Video" propose a learning method to learn a complex appearance model from a sparsely labeled training video with temporal consistency [2]. Authors use repeatedly a Boosting procedure to improve appearance-based model. This method is demonstrated to reduce the labeling requirement by one to two orders of magnitude. ZdenekKalal, ZdenekKalal and KrystianMikolajczyk entitled "P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints" a novel paradigm for training a binary classifier from labelled and unlabeled examples that we call P-N learning is proposed [3]. The learning process is guided by positive (P) and negative (N) constraints which restrict the labeling of the unlabeled set. P-N learning evaluates the classifier on the unlabeled data. Joao Carreira, Fuxin Li and CristianSminchisescu entitled "Object Recognition by Sequential Figure-Ground Ranking" a problem for segmentation and recognition in different categories of objects in images is proposed [4]. An approach to visual object-class segmentation and recognition based on a pipeline that com-bines multiple figure-ground hypotheses with large object spatial support generated by bottom-upcomputational processes that do not exploit knowledge of specific categories and sequential categorization based on continuous estimates of the spatial overlap between the image segment hypotheses and each putative class., is presented. ZdenekKalal, KrystianMikolajczyk and Jiri Matas entitled "Tracking-Learning-Detection" problem of tracking of an unknown object in a video stream, where the object changes appearance frequently moves in and out of the camera view is proposed [5]. A novel tracking framework (TLD) that explicitly decomposes the long-term tracking task into tracking, learning and detection is proposed by the authors. The tracker follows the object from frame to frame. SamueleSalti, Andrea Cavallaro and Luigi Di Stefano entitled "Adaptive Appearance Modeling for Video Tracking: Survey and Evaluation" a unified conceptual framework for appearance model adaptation that enables a principled comparison of different approaches is proposed [6]. A key component for achieving long–term tracking is the tracker's capability of updating its internal representation of targets (the appearance model) to changing conditions. Qing Wang, Feng Chen, WenliXu and Ming-Hsuan Yang entitled "Object Tracking via Partial Least Squares Analysis" a problem of object tracking as a binary classification problem [7]. Object tracking is an important problem in image analysis with numerous applications. It is concerned with low-level visual processing and high-level image analysis, and is widely used in image understanding, human-computer interaction, surveillance, and robotics, to name a few. Yi We I and Zhao Long entitled "Robust objects tracking algorithm based on adaptive background updating" an algorithm called Continuously Adaptive Mean Shift (CAMSHIFT) is proposed to solve the false tracking problem [8]. It uses the disparity of global and local motion to detect the motion area.

III. Proposed Method

We proposed a novel methodology for video object tracking based on wavelet thresholding and radial biases neural network. Initially the discrete wavelet transform function is applied into input video. Now input video decomposed in to layer structure form. After that we calculate horizontal, vertical and diagonal coefficient of input video, after that we apply soft thresholding technique and generate trained pattern using ACP algorithm. In RBF network we used Gaussian based kernel function. The ACP algorithm generates a trained pattern for the removal of noise. In that process the variance factor of frame is increase and the target tracking area is achieved. As known, the high-order statistical relationship does play an important part in video filtration technique area. So in order to take advantage of the high-order statistical relationship among variables, so we used ACP algorithm for training the network. Proposed segmentation filter is a three-layer neural network with inputs derived from an NxN neighborhood of the transformed video and appropriately selected neuron activation functions. As shown inFigure 1, the network takes Ypand ΔYkas the inputs, where Ypis the wavelet transform coefficient under consideration, which is the center of a N x N processing window, and ΔYk=Yk -Ypis the difference value between Ypand the coefficient Yk (k=0,1,…,N2-1, kΔp) of the other points in the N x N window. Figure 2 shows an example of a processing window with a size of contents frame cycle. In this example, Y12 is the center of the window, and ΔYk Y12(k=0,1,…,24, kΔ12).ablest, so we used ACP algorithm for training the network.

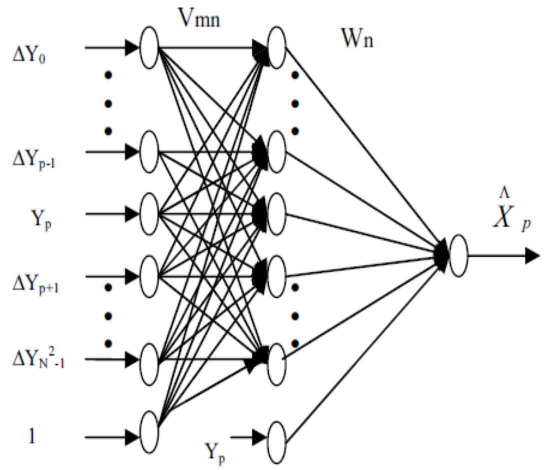

Figure 1: Neural network structure.

Y0	Y1	Y2
Y3	Y4	Y5
Y6	Y7	Y9

Figure 2: shows that unit input vector frame cycle.

the output of network is linear activation function.that activation function perform the targeted output of PSNR value. Step for proposed methodology.

Input video for object detection

1. Perform wavelet transform and video decomposed in layers.
2. Find horizontal, vertical and diagonal coefficient of wavelet.
3. Apply soft thresholding of wavelet
4.Check value of coefficient of wavelet
5.Decide the size of vector input for frame cycle
6.Trained the network.
7.Apply target value of activation function
8.Find frame loss
9.Segmented video traced

IV. Experimental Result

This section demonstrates the performance of the PLS model and our proposed algorithms on a video segmentation. The sequence shown here are 30 frames per cycle. We used an adaptive mixture of five Gaussian components. The *L* was set at 500 frames (α=0.002 in Grimson et al's) and the threshold *T* was set at 0.6. In the shadow detection module, the brightness threshold, τ of 0.7 was used. To show the performance of the background models, higher level processes such as noise cleaning or connected component analysis algorithms were not introduced to the results of background subtractions. Figure 1 shows a sequence of busy outdoor scene containing people playing a football in a central room. Because of no clean video at the beginning, an artifact of the initial image left in PLS tracker lasted for over a hundred frames. Better segmentation can be seen from our method. The performance enhances dramatically with the shadow detection module.

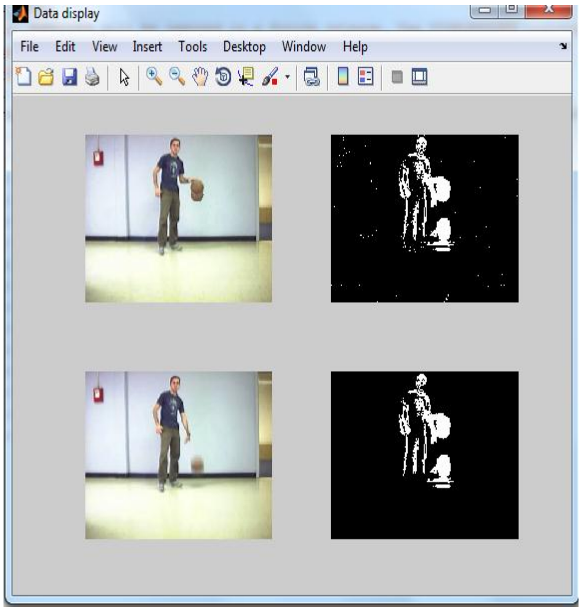

Figure 3: shows that video segmentation of both methods PLS and proposed method.

Table 1: shows that performance parameter of man football video in form of segmented area and loss of frame.

Method	Segmented area (%)	Frame loss
PLS	89	33.37
GMM	86	30.75
Proposed	92	29.62

Table 2: shows that performance parameter of man walking in street video in form of segmented area and loss of frame

method	Segmented area (%)	Frame loss
PLS	91	10.23
GMM	88	14.34
Proposed	94	8.64

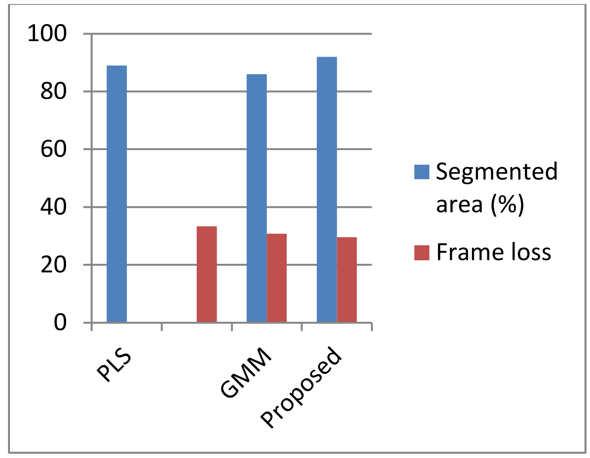

Figure 4: shows that comparative result of video object detection of PLS, GMM and proposed method for playing ball man video.

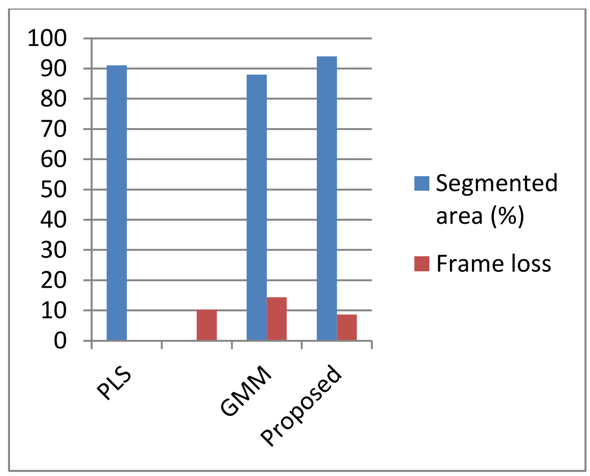

Figure 5: shows that comparative result of video object detection of PLS, GMM and proposed method for walking man video.

V. Conclusion and Future Work

In this paper we proposed a novel method for video segmentation and background removal for video object tracking. The proposed method is very efficient in compression of frame loss and segmentation area for video object tracking. The proposed method comes along with wavelet filter and RBF neural network. So the complexity of method is increase in terms of segmented area and frame loss minimisation. The performance can be further improved by fusing multi-modal information such as by applying the vehicle classification result to constraining the size of the object and vice versa.

Futurework will include applying the algorithm to a larger numberof data and performing comparative studies on variousapplications with various vision- and other sensor-based approaches.

References

[1] G. D. Hager and P. N. Belhumeur, "Efficient region tracking with parametric models of geometry and illumination," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 10, pp. 1025–1039, Oct. 1998.

[2] J. Ho, K.-C. Lee, M.-H. Yang, and D. Kriegman, "Visual tracking using learned linear subspaces," in Proc. IEEE Comput. Vis. Pattern Recognit. Conf., vol. 1. Jul. 2004, pp. 782–789.

[3] A. D. Jepson, D. J. Fleet, and T. F. El-Maraghi, "Robust online appearance models for visual tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 10, pp. 1296–1311, Oct. 2003.

[4] Z. Kalal, J. Matas, and K. Mikolajczyk, "P-N learning: Bootstrapping binary classifiers by structural constraints," in Proc. IEEE Comput. Vis. Pattern Recognit. Conf., San Francisco, CA, Jun. 2010, pp. 49–56.

[5] J. Kwon and K. Lee, "Visual tracking decomposition," in Proc. IEEE Comput. Vis. Pattern Recognit. Conf., San Francisco, CA, Jun. 2010, pp. 1269–1276.

[6] X. Li, W. Hu, Z. Zhang, X. Zhang, and G. Luo, "Robust visual tracking based on incremental tensor subspace learning," in Proc. IEEE Comput. Vis. Pattern Recognit. Conf., Rio de Janeiro, Brazil, Oct. 2007, pp. 1–8.

[7] R.-S. Lin, D. A. Ross, J. Lim, and M.-H.Yang, "Adaptive discriminative generative model and its applications," in Proc. Adv. Neural Inf. Process. Syst. Conf., 2004, pp. 801–808.

[8] L. Matthews, T. Ishikawa, and S. Baker, "The template update problem," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 6, pp. 810–815, Jun. 2004.

[9] X. Mei and H. Ling, "Robust visual tracking using 1 minimization," in Proc. IEEE Int. Comput. Vis. Conf., Sep. 2009, pp. 1436–1443. D. Nguyen and D. Rocke, "Tumor classification by partial least squares using microarray gene expression data," Bioinformatics, vol. 18, no. 1, pp. 39–50, 2002.

[10] R. Rosipal and N. Kramer, "Overview and recent advances in partial least squares," in Latent Structures Feature Selection. New York: Springer-Verlag, 2006, pp. 34–51.

[11] R. Rosipal and L. Trejo, "Kernel partial least squares regression in reproducing kernel Hilbert space," J. Mach. Learn. Res., vol. 2, pp. 97–123, Mar. 2002.

[12] D. Ross, J. Lim, R.-S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking," Int. J. Comput.Vis., vol. 77, nos. 1–3, pp. 125–141, 2008.

[13] H. Saigo, N. Kramer, and K. Tsuda, "Partial least squares regression for graph mining," in Proc. ACM SIGKDD Int. Knowl. Discovery Data Mining Conf., 2008, pp. 578–586.

[14] W. Schwartz, A. Kembhavi, D. Harwood, and L. Davis, "Human detection using partial least squares analysis," in Proc. IEEE Int. Comput. Vis. Conf., Oct. 2009, pp. 24–31.

[15] O. Williams, A. Blake, and R. Cipolla, "Sparse bayesian learning for efficient visual tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 8, pp. 1292–1304, Aug. 2005.

[16] H. Wold, "Partial least squares," in Encyclopedia of Statistical Science, vol. 6, S. Kotz and N. L. Johnson, Eds. New York: Wiley, 1985, pp. 581–591.