IORD Journal of Science & Technology
E-ISSN: 2348-0831
Volume 3, Issue 3 (MAR-APRIL 2017) PP13-19
IMPACT FACTOR 1.719
www.iord.in

"RESEARCH PAPER ON IMPLEMENTING OBJECT TRACKING FOR VISIUAL SURVEILLANCE USING MEAN SHIFT ALGORITHIM"

Miss. Anand Gupta [1],Prof.Pravin Kulurkar [2]

[1](PG Student, VIT, Nagpur)
[2](Prof. VIT ,Nagpur)

Abstract- . This paper presents efficient object tracking in video sequences using multiple features by embedding mean shift into particle filters. When clutter background and occlusions are present. Particle filtering is used because it is very robust and performs well for non-linear and non-Gaussian dynamic state estimation problems. The image features, such as shape, texture, color, contours, and random motion appearance can be used to track the moving object(s) in videos. We proposed a smart video surveillance system with real-time as well as offline stored video database for moving object detection, classification and tracking capabilities. The graphical user interface of the proposed system using MATLAB 2012b is implemented that operates on both color and gray scale video images from a stationary camera. In this method, we proposed real time moving object detection and tracking system using static webcam that can processes 800*600, 640*480, and 320*240 resolution video sequences for capturing live scene as well as stored database of segmented videos. Issues related with object detection and tracking is jointly employed using mean shift and particle filters. Tracker will estimate the dynamic shape and random appearance of objects. Tracking requires location and shape of object in every segmented frame. Rectangular bounding box is used to improve the estimate of its shape and appearance. Target object may experience partial occlusions, intersection with other objects with similar color distributions, abrupt motion speed changes, and cluttered background. Test result shows improvement in terms of robustness, tracking drifts, accuracy of tracked bounding box to occlusions.

Keywords—*(Object Tracking, Mean Shift Algorithim,Local Binary Pattern,)*

INTRODUCTION

In a Visual object tracking is a fundamental task in the field of computer vision. Its objective is to identify the location of the object of interest from frame to frame. There exist many applications, e.g., video surveillance, human–computer interaction, traffic pattern analysis, and robotics. The goal of object tracking is to estimate the states of a target object in an image sequence. It plays a critical role in numerous vision applications such as motion analysis, activity recognition, visual surveillance and intelligent user interfaces. While much progress has been made in recent years, it is still a challenging problem to develop a robust algorithm for complex and dynamic scenes due to large appearance changes caused by varying illumination, camera motion, occlusions, pose variation and shape deformation. The goal of object tracking is to estimate the states of a target object in an image sequence. It plays a critical role in numerous vision applications such as motion analysis, activity recognition, visual surveillance and intelligent user interfaces. While much progress has been made in recent years, it is still a challenging problem to develop a robust algorithm for complex and dynamic scenes due to large appearance changes caused by varying illumination, camera motion, occlusions, pose variation and shape deformation.Several factors need to be considered for an effective appearance model. First, an object can be represented by different features such as intensity, color, texture, superpixels, and Haar-like features. Meanwhile, the representation schemes can be based on holistic templates or local histograms. In this work, we use intensity values to represent objects due to its simplicity and efficiency. Furthermore, our approach exploits the strength of holistic templates to distinguish the target from the background, and the effectiveness of local patches in handling partial occlusion.

Object Tracking

Video tracking is the process of locating a moving object (or multiple objects) over time using a camera. It has a variety of uses, some of which are: human-computer interaction, security and surveillance, video communication and compression, augmented reality, traffic control, medical imaging and video editing Nowadays video surveillance systems are installed worldwide in many different sites such as airports, hospitals, banks, railway stations and even at home (see figure 1). The surveillance cameras help a supervisor to oversee many different areas from the same room and to quickly focus on abnormal events taking place in the controlled space. However one question arises: how can a security officer analyse and simultaneously dozens of monitors with a minimum rate of missing abnormal events (see figure 2) in real time? Moreover, the observation of many screens for a long period of time becomes tedious and draws the supervisor's attention away from the events of interest. The solution to this issue lies in three words: intelligent video monitoring. The term "intelligent video monitoring" expresses a fairly large research direction that is applied in different fields: for example in robotics and home-care. In particular, a lot of researches and works are already achieved in video surveillance applications. Figure 3 presents a processing chain of a video interpretation system for action recognition. Such a chain includes generally different tasks: image acquisition, object detection, object classifi- cation, object tracking and activity recognition. This paper studies the mobile object tracking task. The aim of an object tracking algorithm is to generate the trajectories of objects over time by locating their positions in every frame of video. An object tracker may also provide the complete region in the image that is occupied by the object at every time instant. Mobile object tracking has an important role in the computer vision applications such as home care, sport scene analysis and video surveillance-based security systems (e.g. in bank, parking, airport). In term of vision tasks, the object

Mean Shift Algorithim:

The mean-shift algorithm is an efficient approach to tracking objects whose appearance is defined by histograms. This section provides an intuitive idea of Mean shift and the later sections will expand the idea. Mean shift considers feature space as a empirical probability density function. If the input is a set of points then Mean shift considers them as sampled from the underlying probability density function. If dense regions (or clusters) are present in the feature space , then they correspond to the mode (or local maxima) of the probability density function. We can also identify clusters associated with the given mode using Mean Shift.For each data point, Mean shift associates it with the nearby peak of the dataset's probability density function. For each data point, Mean shift defines a window around it and computes the mean of the data point . Then it shifts the center of the window to the mean and repeats the algorithm till it converges. After each iteration, we can consider that the window shifts to a more denser region of the dataset.

At the high level, we can specify Mean Shift as follows :

1. Fix a window around each data point.

2. Compute the mean of data within the window.

3. Shift the window to the mean and repeat till convergence.

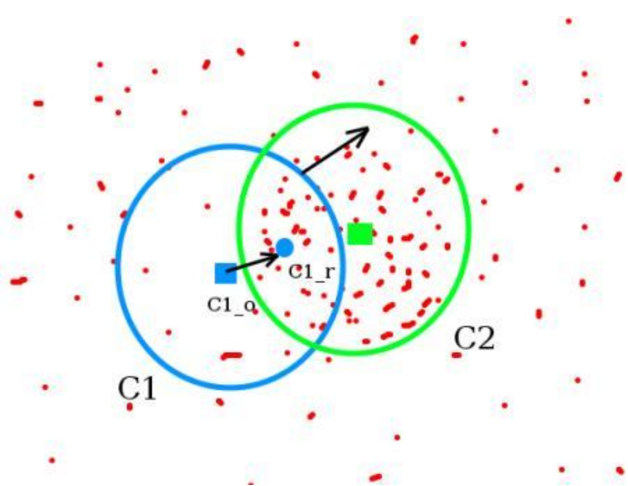
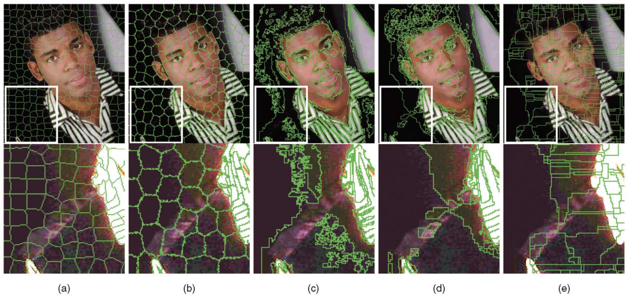


Fig.1. Mean Shift Algorithim Set a point.

IORD Journal of Science & Technology
E-ISSN: 2348-0831
Volume 3, Issue 3 (MAR-APRIL 2017) PP13-19
IMPACT FACTOR 1.719
www.iord.in



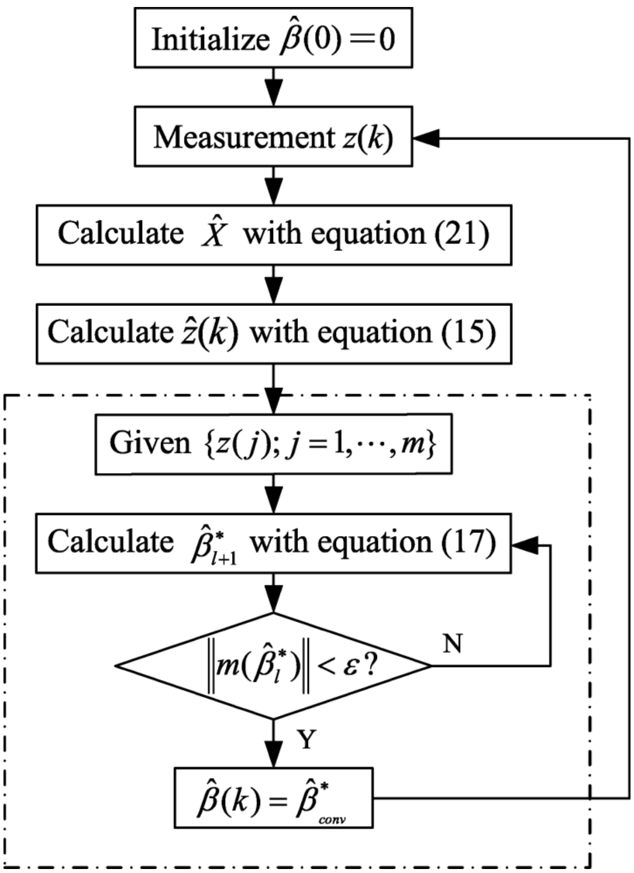


Fig2. Mean Shift Algorithim Implementation Flowchart

1. Problem Statement

Video tracking can be a time consuming process due to the amount of datathat is contained in video. Adding further to the complexity is the possible need to use object recognition techniques for tracking, a challenging problem in its own right. Also the difficulties arising from noise, occlusion, clutter and changes in foreground object or in the background environment

2. Literature Survey

2.1 In the year SEPTEMBER 2016 Bo Ma, Member, IEEE, Hongwei Hu, Jianbing Shen, Senior Member, IEEE, Yangbiao Liu, and Ling Shao, Senior Member, IEEE [1] "Generalized Pooling for Robust Object Tracking" majority of sparse coding based tracking algorithms computes final feature vectors only by low-order statistics or extreme responses of sparse codes. The high-order statistics and the correlations between responses to different dictionary items are neglected. We present a more generalized feature pooling method for visual tracking by utilizing the probabilistic function to model the statistical distribution of sparse codes. Since immediate matching between two distributions usually requires high computational costs, we introduce the Fisher vector to derive a more compact and discriminative representation for sparse codes of the visual target. We encode target patches by local coordinate coding, utilize Gaussian mixture model to compute Fisher vectors, and finally train semi-supervised linear kernel classifiers for visual tracking. In order to handle the drifting problem during the tracking process, these classifiers are updated online with current tracking results.a robust object tracking algorithm based on a sparse collaborative model that exploits both holistic templates and local representations to account for drastic appearance changes. Within the proposed collaborative appearance model, we develop a sparse discriminative classifier (SDC) and sparse generative model (SGM) for object tracking.

2.2 In the year 2010 Mr.Di.Zang [2] proposed an illumination invariant object tracking based on multi scale phase representation is used for capturing the video feature information. The presented approach can tackle the illumination changes, experimental results. Video tracking is widely used for various applications such as vehicle navigation traffic control and face detection.

2.3 In the year 2011 Kun wang, Xiaopingp.liu [3]proposed method visual object tracking based on filtering methods. Basically process determined in two approach 1} deterministic method 2} stochastic method.The deterministic method is based ontarget representation and localization techniques. The stochastic method is based on filtering and data association .

The advantages of particle filter include accuracy, flexibility in non-linear and non Gaussian problems.

2.4 In 2006Juan Zho,Dany Yao [4] method of occlusion adaptive object tracking based on video image .occlusion tracking is difficult in image processing ,and this paper gives a system using radar data to segment the object data .this paper focuses on developing a traffic tracking system . In video surveillance, occlusion object tracking is difficult because two objects overlap to each other in an image. This paper improves background acquisition and shadow clutter rejection method.

2.5 In 2011 Jiajun Wen, Yong Xu [5] proposed ghost, occlusion removal and distracters handling in object tracking. For real time tracking system, interference such as ghost, occlusion and distracters have a great influence on tracking system. Occlusion handling is done using merge/splitting and the hue histogram can well in many cases in our videos.Distractors problem is developed by utilizing moving direction.

2.6 In 2009 Saad M .Khan and Mubarak Shah [6] proposed method based on tracking multiple occluding people y by localizing on multiple scene. This algorithm can track multiple people in a complex environment. This is achieved by resolving occlusions on multiple scène. Our method of object detection and occlusion resolution is based on geometrical constructs, which is obtained by using foreground and background modeling techniques.

RECENT WORKS AND LIMITATIONS

Drifting problem

Drifting problem occurred when an object abruptly changes its direction to reverse. In that case, it's become very difficult to track the object; because the motion model doesn't work. Figure shows the object drifting problem. Here, the dotted rectangle represents predicted object location in next frame that is predicted by object motion model.

1. **Function: Examine if the tracking window is in the range of image**

parameters:

- rmin,ramx,cmin,cmax: window of tracking result
- height,widht: the size of image

2. Function: Calculate the target model with joint color-texture histogram

parameters:

- Frame: the 1th image
- center: the center of target window
- w_halfsize: the bandwidth of target window
- lbpThreshold,
- redBins,greenBins,blueBins quantification scheme of rgb space
  Function returns:
- q_u: target model

3. **Function: mean shift tracking with joint color-texture joint histogram**

parameters:

- Image: current frame to track
- center: initial location of target
- w_halfsizet: bandwidth of target window
- q_u: target model
- redBins,greenBins,blueBins: quantification scheme of rgb space
- minDist: convergence condition
- maxIterNum: maxmimal iterative number
- incre: increase the candidate winodw to robust tracking
- lbpThreshold

IMPLEMENTATION ON PREVIOUS WORK

we have done a comparison between LBP(Local Binary Pattern) algo& Mean Shift Algo& Find that time consume required for tracking is less by mean shift algorithm & also show the RGB correlation bet these two algorithims.

Local Binary Pattern

Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings. In the LBP approach for texture classification, the occurrences of the LBP codes in an image are collected into a histogram. The classification is then performed by computing simple histogram similarities. However, considering a similar approach for facial image representation results in a loss of spatial information and therefore one should codify the texture information while retaining also their locations. One way to achieve this goal is to use the LBP texture descriptors to build several local descriptions of the face and combine them into a global description. Such local descriptions have been gaining interest lately which is understandable given the limitations of the holistic representations. These local feature based methods are more robust against variations in pose or illumination than holistic methodsthe occurrences of the LBP codes in an image are collected into a histogram. The classification is then performed by computing simple histogram similarities. However, considering a similar approach for facial image representation results in a loss of spatial information and therefore one should codify the texture information with their locations. One way to achieve this goal is to use the LBP texture descriptors to build several local descriptions of the face and combine them into a global description. Such local descriptions have gained interest lately which is understandable given the limitations of the holistic representations. These local feature based methods seem to be more robust against variations in pose or illumination than holistic methods..The basic methodology for LBP based face description proposed by Ahonen et al. (2006) is as follows: The facial image is divided into local regions and LBP texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face, as shown in Fig.

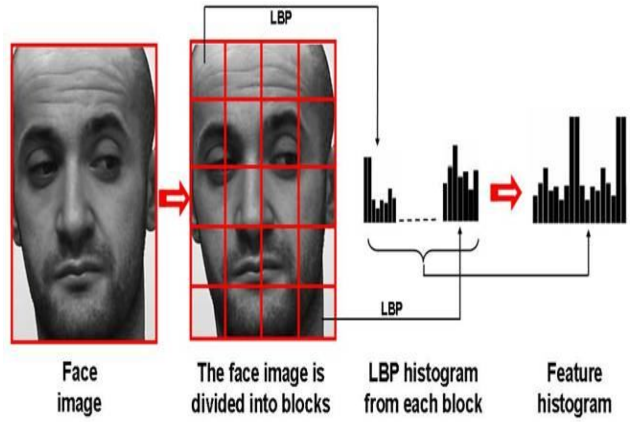


Figure 3: Face description with local binary patterns.

This histogram effectively has a description of the face on three different levels of locality: the LBP labels for the histogram contain information about the patterns on a pixel-level, the labels are summed over a small region to produce information on a regional level and the regional histograms are concatenated to build a global description of the face.

It should be noted that when using the histogram based methods the regions do not need to be rectangular. Neither do they need to be of the same size or shape, and they do not necessarily have to cover the whole image. It is also possible to have partially overlapping regions.The two-dimensional face description method has been extended into spatiotemporal domain (Zhao and Pietikäinen 2007). Fig. depicts facial expression description using LBP-TOP. Excellent facial expression recognition performance has been obtained with this approach.

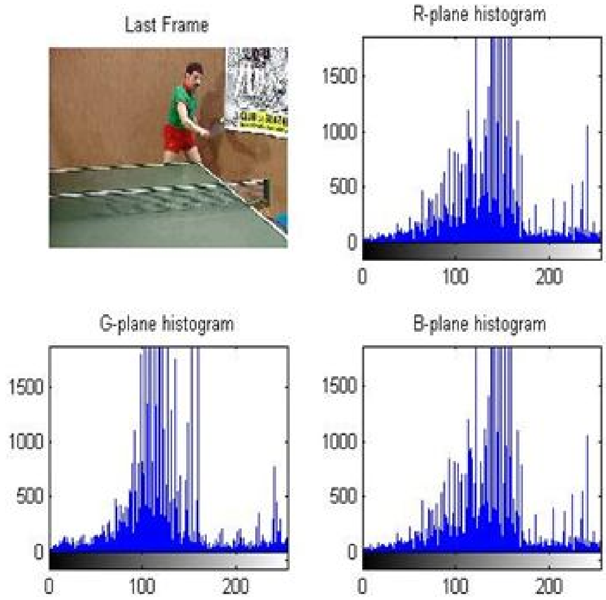


Fig. Illustration of how image is viewed by people with RGB deficiencies

From the analysis, we can see some gray scale values in the image, area of the pixels, minimum and maximum values of RGB colors and the mean of the colors that make up the image. This affirms that an image contains shades of gray and colour

Fig.(a).Selected Region

Different features extracted from the image: (a) the original image; (b) the ordinary LBP feature image in the intensity channel

Fig.LBP based facial representation

The basic methodology for LBP based face description is as follows: The facial image is divided into local regions and LBP texture descriptors are extracted from the each region independently. The descriptors are then concatenated to a global face description, as shown in Fig.

Conclusion

In this paper we implemented the mean shift algorithm with certain improvements. Firstly, the frames of the avi video file are generated. Secondly, the window size is calculated to track a target accurately when the target's shape and orientation are changing. Various results of the experiment conducted on the "Samplevideo.avi" file are shown. For example, if the motion of the target from frame to frame is known to be larger than the operational basin of attraction, one should initialize the tracker in multiple locations in the neighborhood of basin of attraction, according to the motion model.

References

1. Robust Object Tracking via Sparse Collaborative Appearance Model Wei Zhong, Huchuan Lu, Senior Member, IEEE, and Ming-Hsuan Yang, Senior Member, IEEE 2014
2. Optimal color-based mean shift algorithm for tracking objects Xiaowei An, Jaedo Kim, vPublished in IET Computer Vision 2014

3.]Oct. 2008. Y. Li, H. Ai, T. Yamashita, S. Lao, and M. Kawade, “Tracking in lowframe rate video: A cascade particle filter with discriminative observersof different life spans,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30,no. 10, pp. 1728–1740
4. Kernel-Based Structural Binary Pattern Tracking Dae-Hwan Kim, Hyo-Kak Kim, Seung-Jun Lee, Won-Jae Park, and Sung-Jea Ko, Fellow, IEEE
5. 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.J. Kwon and K. M. Lee, “Visual tracking decomposition,” in Proc. IEEEConf. CVPR, Jun. 2010, pp. 1–10.
6. B. Babenko, M.-H. Yang, and S. Belongie, “Visual tracking with online multiple instance learning,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jan. 2009, pp. 983–990.
7. A. Adam, E. Rivlin, and I. Shimshoni, “Robust fragments-based tracking using the integral histogram,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2006, pp. 798–805.
8. B. Liu, J. Huang, L. Yang, and C. Kulikowsk, “Robust tracking using local sparse appearance model and k-selection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 1313–1320
9. Q. Yu, T. B. Dinh, and G. G. Medioni, “Online tracking and reacquisition using co-trained generative and discriminative trackers,” in Proc. 10th ECCV, 2008, pp. 678–691.
10. W. Zhong, H. Lu, and M.-H. Yang, “Robust object tracking via sparsitybased collaborative model,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 1838–1845.