Rate Control For People detection in video

[1]S. S. Joshi, [2]Prof. J. R. Panchal

[1](Electronics and Telecommunication, Siddhant College of engineering, / PuneUniversity , India)
[2]Asst. Professor (Electronics and Telecommunication, Siddhant College of engineering, / PuneUniversity , India)

Abstract: *People Detection in video is generally used in high level multimedia applications like intelligent surveillance systems, augmented reality ,etc. People detection is based on background subtraction. Generally videos are available in compressed form due to which noise will be added in it. Many algorithms are used to detect people and control their rate. There are three key steps in video analysis detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior. The ability of human visual system to detect visual saliency is extraordinarily fast and reliable. However, computational modeling of this basic intelligent behavior still remains a challenge. In this paper we put our attention into background subtraction and Gaussian grouping of pixels for detection of people in low quality video i.e. improve accuracy.*
Keywords: peopledetection,segmentation,background subtraction, gaussian grouping, kalman filter

I. Introduction

Video surveillance system is based on the ability to detect moving objects or people in video. Detection of moving objects is first step in surveillance systems. Each image is then segmented i.e. each image is split into number of frames. Video is converted into sequence of frames by image analysis techniques. Background subtraction is widely used technique for detecting moving objects in videos from static camera. The background subtraction method is to use the method of difference between current image and background image. This method is effective to improve the effect of moving object detection. Background subtraction is nothing but foreground detection. After detecting foreground image, grouping of pixels as per Gaussian distribution will be done. Kalman filter is used to track the people in video.

Background subtraction technique find the foreground object from video and then classify it into categories like human, animal, vehicle etc., based on shape , color , motion or other features. Most of the multimedia videos are available in compressed format. Usually higher the compression rate, lowers the correct hits and quality of video due to noise added in it. A modern object detection algorithm can be divided into five parts: pre-processing and normalization, local rectification and compensation of small shape variations, computation of descriptor set, machine learning classification, and post-processing to fuse multiple detections. In this paper, we focus on detection schemes based on background subtraction because of their widespread use and the possibilities they offer in implementing real-time object detection systems. background subtraction is nothing but foreground detection. Unfortunately, images and video are usually available in compressed format which makes object detection more difficult because of the additional distortion noise. In this paper we propose a saliency map algorithm and compare it with background method to improve accuracy.

Object detection has been studied for about four decades producing a wide range of object detection techniques. This research effort has been fostered by the many application scenarios where object detection can be employed (e.g., active video surveillance, assisted or autonomous drive, database search, data classification) and by the need of increasing the robustness of existing algorithms to different lighting conditions, poses, scales, locations, and geometries . Object detection has been studied for about four decades producing a wide range of object detection techniques. This research effort has been fostered by the many application scenarios where object detection can be employed (e.g., active video surveillance, assisted or autonomous drive, database search, data classification) and by the need of increasing the robustness of existing algorithms to different lighting conditions, poses, scales, locations, and geometries (deformable objects). As a matter of fact, a modern object detection algorithm can be divided into five parts: pre-processing and normalization, local rectification and compensation of small shape variations, computation of descriptor set, machine learning classification, and post-processing to fuse multiple detections.

II. Related Work

Many detection algorithm algorithms have been proposed in order to evaluate performance of object detection for video surveillance [3]. A survey of techniques for human detection [6]. Most of the presented algorithms depend on the the statistics of the orientation of edges (Histograms of Oriented Gradients or HOGs)

and color histograms[9]. Background subtraction algorithm is used to improve the effectiveness of human motion detection[8]. In this paper ,we improve the correct hits and PSNR as compare to saliency based algorithm[1].

T. Bouwmans, F. El Baf, B. Vachon state that a survey of Background Modeling using Mixture of Gaussians for Foreground Detection[13]. many background modeling methods have been developed and the most recent surveys can be found in . These background modeling methods can be classified in the following categories: Basic Background Modeling , Statistical Background Modeling , Fuzzy Background Modeling and Background Estimation . Other classifications can be found in term of prediction , recursion, adaptation , or modality.

III. Technique

1.1 .**Implimentation Scheme**

1.2 Background Modeling Using Mixture of Gaussians

In the context of a traffic surveillance system, each background pixel using a mixture of three Gaussians corresponding to road, vehicle and shadows. This model is initialized using an EM algorithm . Then, the Gaussians are manually labeled in a heuristic manner as follows: the darkest component is labeled as shadow; in the remaining two components, the one with the largest variance is labeled as vehicle and the other one as road. This remains fixed for all the process giving lack of adaptation to changes over time. For the foreground detection, each pixel is compared with each Gaussian and is classified according to it corresponding Gaussian. The maintenance is made using an incremental EM algorithm for real time consideration. Stauffer and Grimson generalized this idea by modeling the recent history of the color features of each pixel $X1,...,Xt$ { } by a mixture of K Gaussians.

There are so many methods to obtain the initial background subtraction. It computes the absolute difference between the current image and a static background image and compares each pixel to a threshold. Background subtraction is very adaptive to stable environments but is extremely sensitive to dynamic scene changes due to lighting and other conditions.
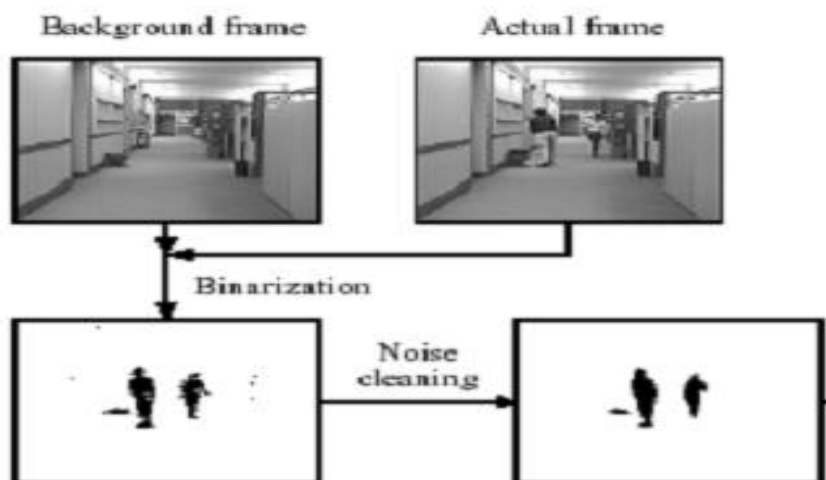
3.3 .**Gaussian Distribution**

Among the high complexity methods , mixture of Gaussian is more beneficial, as it can handle multi-modal distribution. In this, the background is not a frame of values. Rather ,the background model is parametric. Each pixel location is represented by a number of Gaussian functions that sum together to form a probability distribution function. A certain pixel value x, at time t by means of a mixture of Gaussians that sum together to form a probability distribution function as

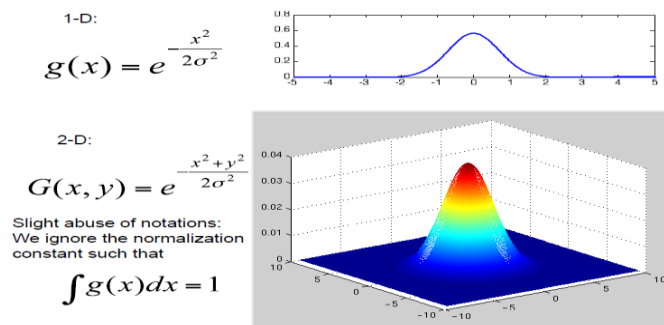$P(x_t)= \Sigma^k_{t=1} \omega_{i,t}.\eta (x_t-\mu_{i,t},\Sigma_{i,t})$………[11]

Where,

K= Gaussian distributions to describe one of the observable background or foreground objects.

μ= Mean of Gaussian function.

$\omega_{i,t}$= weight of the ith Gaussian in the mixture.

$\mu_{i,t =}$ Mean value of the ith Gaussian in the mixture.

The Gaussian distribution for 1dimentional and 2 dimentional image is as follows:

1-D:

$$g(x) = e^{-\frac{x^2}{2\sigma^2}}$$

2-D:

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Slight abuse of notations:
We ignore the normalization constant such that

$$\int g(x)dx = 1$$

This function is simple and of low computational complexity. However, the object is hard to be precisely detected when both of the background and the foreground are complicated.

3.4. **Tracking**

Kalman filter is used to track the people in video. It is also known as linear quadratic estimation.This filter operates recursively on streams of noisy input data and produces estimates of the current state variables. This filter is widely used for guidance,navigation and control of vehicles, aircraft and spacecraft.

The Kalman filter model assumes that the state of a system at a time t evolved from the prior state at time t-1 according to the equation

xt = Ftxt-1 + Btut + wt , ……….[12]

where

xt = state vector containing the terms of interest for the system (e.g., position, velocity, heading) at time t

ut = vector containing any control inputs (steering angle, throttle setting, braking force)

Ft = state transition matrix which applies the effect of each system state parameter at time t-1 on the system state at time t (e.g., the position and velocity at time t-1 both affect the position at time t)

Bt = control input matrix which applies the effect of each control input parameter in the vector **u**t on the state vector(e.g., applies the effect of the throttle setting on the system velocity and position)

wt = vector containing the process noise terms for each parameter in the state vector. The process noise is assumed to be drawn from a zero mean multivariate normal distribution with covariance given by the covariance matrix Qt.

3.5. **Compression Rate**[1]

Most of the object detection algorithms are based on analyzing the distribution of gradients and colors along object borders. Having this in mind, it is possible to design a saliency metric that combines different features into a normalized value. In particular, the input image is divided into blocks and the included pixels are processed computing three different metrics. In fact, it is possible to observe that object detection strategies are sensitive to edge strength, the stationarity of edge directions, and color contrast along the main orientations of an image.

3.5.1   Edge strength

In order to characterize the significance of edges in different regions of the image I(x; y), it is possible to compute horizontal and vertical Sobelian gradients (named Sx(x; y) and Sy(x; y), respectively) on the whole

image using a 3 * 3 Sobel operator. Then, these two gradient maps are merged into a common measure of the gradient strength named

$$G(x,y)= round \left[\frac{(|S_x(x,y)| + |S_y(x,y)|)}{16} \right]$$

3.5.2 Regularity of edge strength

Re-using the previous results, the orientation of borders A(x; y) can be computed as A(x; y) = $\tan^{-1}(Sy(x; y)=Sx(x; y))$. From these values, it is possible to compute the local regularity of edges

$$R_A(x,y) = \left[\frac{|A(x,y) - A(x+1, y)| + |A(x,y) - A(x, y+1)|}{2} \right]$$

For each image block (referenced with the index b), it is possible to compute to variance.

3.5.3 Contrast along edges

For every pixel position (x; y) such that G(x; y) is greater than Tq, the algorithm evaluates the color differences along the edge orientation, i.e., it computes

$$C(x,y) = \frac{|R(x+\delta_x , y+ \delta_y) - R(x- \delta_x, y- \delta_y)|}{3}$$
$$+ \frac{|G(x+\delta_x , y+ \delta_y) - G(x- \delta_x, y- \delta_y)|}{3}$$
$$+ \frac{|B(x+\delta_x , y+ \delta_y) - B(x- \delta_x, y- \delta_y)|}{3}$$

3.5.4 Final saliency metric

The input (uncompressed) image is divided into pixel blocks (indexedwith the variable b), and for each of these thealgorithm computes$\sigma R;b$ and Cb.Starting from this, the saliency value assigned to the current block is

$$S_b = 1- \begin{cases} \frac{C_b}{K} & if 5 < \sigma_{R,b} \\ \frac{C_b * 10}{K} & otherwise \end{cases}$$

In this work, we implemented this strategy considering the emerging video coding standard HEVC. This saliency algorithm is then compared with pixel based Gaussian distribution.

IV.   Result

The proposed approach has been tested using the object detector in [1] for people tracking on different video sequences the whole algorithm inherits the key idea of the HOG-based approach [8] but extends object models introducing improvement of people detection by using Gaussian distribution and kalman filter. Graphs shows the percentage of true hits vs target bit rate and rate distortion performance.
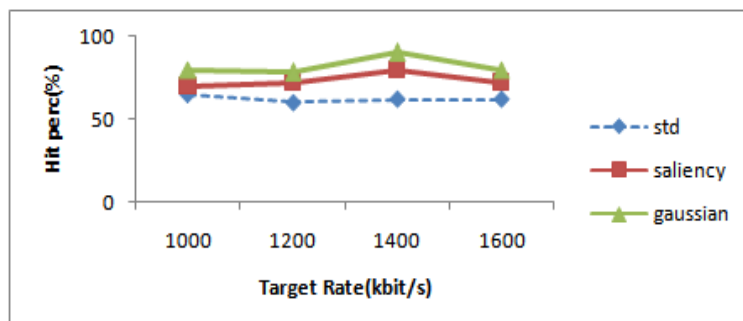

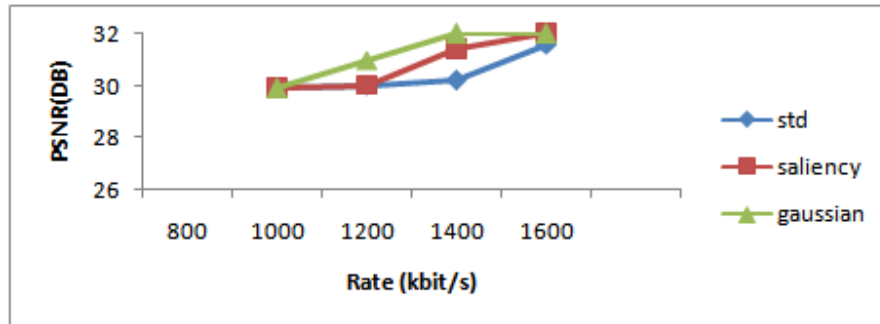Figure2: Graph of %true hits vs target rate

Figure3: Rate distortion performance

V.   Conclusion

The   saliency map is based on a set of features which give to the encoder an evaluation of the significance of the different regions in a frame. In this paper, we have presented people detection in video by using background subtraction method which effectively improve the accuracy and performance.

References

Journal Papers:
[1].   Simone Milani, Riccardo Bernardini, Roberto Rinaldo, "Saliency   based rate control for people detection in video", DIEGM - University of Udine Via Delle Scienze, 208 - 33100 Udine – Italy,IEEE2013
[2].   P.F. Felzenszwalb, R.B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained partbased models," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 32, no. 9, pp. 1627 –1645, sept. 2010.
[3].   Jacinto Nascimento, Jorge Marques, "Performance evaluation of object detection algorithms for video surveillance", IST/ISR, Torre Norte, Rovisco Pais, 1049-001, Lisboa Portugal.
[4].   David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[5].   M. Reisert and H. Burkhardt, "Equivariant holomorphic filters for contour denoising and rapid object detection," Image Processing, IEEE Transactions on, vol. 17, no. 2, pp. 190 –203, Feb. 2008.
[6].   Neeti A. Ogale, "A survey of techniques for human detection from video", Department of computer science ,university of Maryland, college park, MD 20742.
[7].   Yongxia Zuo, Guoqiang Wang, and Chuncheng Zuo, "Wavelet packet denoising for pavement surface cracks detection," in Computational Intelligence and Security, 2008. CIS '08.
[8].   Rupali S. Rakibe, Bharati D. Patil, "Background subtraction algorithm based human motion detection", International journal of Scientific and research publications, Volume3, May2013.
[9].   Navneet Dalal and William Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05), vol. 1, no. 3, pp. 886–893, 2004.
[10].   Simone Milani, Riccardo Bernardini, and Roberto Rinaldo,"Adaptive denoising filtering for object detection applications," in Proc of ICIP 2012, Oct. 2012.
[11].   Sanjivani Shantaiya, Kesari Verma, Kamal Mehta, "Study and analysis of methods of object detection in video", International journal of soft computing and engineering, volume2, Jan 2013.
[12].   Ramsey Faragher, "Understanding the basis of the kalman filter via a simple and intuitive derivation",IEEE signal processing magazine, Sep2012
[13].   T. Bouwmans, F. El Baf, B. Vachon, "Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey" Laboratoire MIA, Université de La Rochelle, Avenue M. Crépeau, 17000 La Rochelle, France