



REVIEW OF HUMAN DETECTION TECHNIQUES

R. S. Lomte¹, Prof. Kalpana Malpe²

¹M.Tech.-IInd Year, Department of Computer Science and Engineering,
NUVA College of Engineering and Technology,
Nagpur University, Nagpur, INDIA

²Department of Computer Science and Engineering, GNIT,
Nagpur University, Nagpur, INDIA

Abstract: In this paper we focuses on overall methods and general characteristics involved in the human detection from surveillance real time video, we provide detailed introduction to human detection techniques. We also provide more detailed discussions on advantages and drawbacks of different techniques in human detection.

Keywords: Background subtraction, statistical methods, temporal differencing, optical flow, shape based Analysis and HOG &SVM detector.

I. Introduction

Now days automatic security growing at high rate, automated security is possible due to real time video processing system. It is important that surveillance should be in house, bank, railway station or even a small shop. Video surveillance system record the day to day activities in corporate sectors, bank, government farms, shops where security needed and if the event such as theft, threats of violence and vandalism to happen it should catch the video. Surveillance video can capture suspects behavior of human we can analysis distinctive features of our guess. Video system also monitoring Visitor in banks, video cameras are connected to a video processing unit such as mobile or laptop, to extract high-level information identified with alert situation; this processing unit could be connected throughout a network to a control and visualization center that manages and send alert.

Need of video surveillance due to some reason

1. Continuous real-time monitoring – Authorized employee can monitor critical areas continuously from their mobile and laptop
2. Security Cameras- We can positioned security cameras throughout bank to prevent crimes
3. Building/venue security – Surveillance cameras monitoring restricted access areas can help to ensure that only registered guests and staff members gain entry to specific areas of a facility.
4. Remote video monitoring – Remote monitoring is an extremely helpful tool. IP surveillance allows your employees to view security camera footage remotely from any PC with network access. Multiple sites can even communicate over the same network with all of the camera views accessible online via the Internet.
5. Crowd control – Large crowds may create by violence and disruptive behavior. A video surveillance network can play a significant role by providing security workers with clear views of the action, and the ability to focus in on specific scenes and individuals.
6. Safe Entry and Exit-Safe entry and exit in Video surveillance areas helps in interruptions and disturbances for those coming and going.
7. Surveillance cameras can provide invaluable visual evidence for investigations of criminal activity and other specific events that have taken place within or around facilities.

So there is a need for a system which can analyze the real time videos. To analyze the real time video for surveillance, we need to detect the humans. Here we are discussing various human detection techniques and human motion detection techniques.

II. Human detection techniques

There are different ways we can detect human, by detecting shape and size of person or by detection of the moving object in the video surveillance and capture the events. There are different methods are used to detect moving object

1. Background subtraction
2. Optical flow
3. Temporal differencing
4. Statistical Analysis
5. Shape Based Analysis
6. HOG and SVM Detection Analysis

III. Training Classifiers

A. Background subtraction Techniques

Background subtraction is a simple method for human detection when still cameras are used for static background. In this method Human are detected by frame differencing, subtracting current frame from background frame. The background image can be calculated by subtracting the current frame from the previous frame or from the average image of a number of frames. Applying threshold value we can store some frame.

Chih-Chang Chen et al. [4] to model light variations and to determine pedestrian objects from a static scene a dynamic background subtraction module is used to built adaptively using pixel gray level values.

Yanling Wang et al. [9] used the background subtraction method to separate background and foreground objects. A new updating algorithm is used for background approximating a temporal low-pass filter is used to upgrade background model.

Osama Masoud et al. [21] developed a new method for tracking and counting pedestrians in real time system using a single camera. The background subtraction method and thresholding is used to produce the differential image. The threshold value is considered by subtracting background frame and current frame by measuring the maximum fluctuation of pixel values during this training period.

J. L. Raheja et al. [13] used Multimodal background for a bidirectional people counting algorithm. The Gaussian mixture model is used to describe a pixel of background.

Background subtraction can be updated by

$$S_i(x,y) = F_i(x,y) - C_i(x,y)$$

Where i is the index of current frame,

$F_i(x,y)$ denotes the pixel gray level value on (x,y) on first frame

$C_i(x,y)$ denotes the pixel gray level value on (x,y) on current frame

$S_i(x,y)$ denotes the background pixel gray level value calculated from current frame subtracted from first frame[4].

Drawback of Background subtraction

1. It only works well in particular conditions of object speed and frame rate and is very sensitive to the threshold[12][9].
2. It is sensitive to changes of dynamic scenes due to lighting and extraneous events, gradual and sudden changes in lighting conditions.
3. If a moving object is present during the initialization of the background scene.
4. Any shadows that is present at any time.
5. Slow-moving objects.
6. If a foreground object becomes motionless, it would appear to be the same as a background object that moves and then becomes motionless.
7. Background objects that are inserted or removed from the scene that become/was part of the background.
8. If an object that was part of the background is moved, both it and the part of the background that was behind the object appear to change.
9. If a foreground object's pixel characteristics are almost the same as the background [7].
10. This method could handle some of the inconsistencies due to lighting changes in dynamic background.
11. Background subtraction was then performed to cope with shadows and unreliable color cues effectively.

B. Optical flow Techniques

Optical flow can be used when a human movement is distinguishable, from its background provided Optical flow is the amount of image movement within a given time period [11].

Meyer et al. [26] performed monotonic operation which computed the displacement vector field to initialize a contour-based tracking algorithm, called active rays, for the extraction of articulated objects which would be used for gait analysis.

Rowley and Reh [10] also focused on the segmentation of optical Flow fields of articulated objects. Its major contributions were to add kinematic motion constraints to each pixel, and to combine motion segmentation with estimation in expectation maximization (EM) computation.

Bregler's work [3], each pixel was represented by its optical Flow. These Flow vectors were grouped into blobs having coherent motion and characterized by a mixture of multivariate Gaussians.

Barron's work [15]. In addition to the basic methods described above, there are some other approaches to motion segmentation. Using the extended EM algorithm [8],

Friedman and Russell [20] implemented a mixture of Gaussian classification model for each pixel. This model attempted to explicitly classify the pixel values into three separate predetermined distributions corresponding to background, foreground and shadow. Meanwhile it could also update the mixture component automatically for each class according to the likelihood of membership. Hence, slow-moving objects were handled perfectly, meanwhile shadows were eliminated much more effectively.

Stringa [7] also proposed a novel morphological algorithm for scene change detection. This proposed method allowed obtaining a stationary system even under varying environmental conditions. From the practical point of view, the statistical methods are better choice due to their adaptability in more unconstrained applications.

1. It is computationally expensive and therefore not practicable for real-time systems based on a lowcost hardware [17].
2. Optical Flow computation methods are computationally complex and very sensitive to noise, and cannot be applied to video streams in real-time without specialized hardware.
3. Optical Flow methods can be used to detect independently moving objects even in the presence of camera motion.
4. Flow computation methods are computationally complex and very sensitive to noise, and cannot be applied to video streams in real-time without specialized hardware.

C. Temporal differencing

In temporal differencing the reference image is the previous images. Hence the previous frame is subtracted from the current image and the subtraction value must be greater than a threshold value in order to give a difference image Temporal differencing [9] is very adaptive to dynamic environments.

The approach of temporal differencing [18] makes use of pixel-wise difference between two or three consecutive frames in an image sequence to extract moving regions. Temporal differencing is very adaptive to dynamic environments.

Lipton et al. [1] detected moving targets in real video streams using temporal differencing. After the absolute difference between the current and the previous frame was obtained, a threshold function was used to determine change. By using a connected component analysis, the extracted moving sections were clustered into motion regions. These regions were classified into predefined categories according to image-based properties for tracking to use three-frame differencing instead of two-frame differencing [7,46].

VSAM [18] has successfully developed a hybrid algorithm for motion segmentation by combining an adaptive background subtraction algorithm with a three-frame differencing technique. This hybrid algorithm is very fast and surprisingly objective for detecting moving objects in image sequences.

Drawback

1. Temporal differencing does a poor job of extracting all relevant feature pixels e.g. Holes in moving object [11].

D. Statistical methods

In statistical method we can consider characteristic of individual pixel or group of pixel to construct background frame, statistics of background can automatically update during processing use here for dynamically.

1. This approach is becoming increasingly popular due to its robustness to noise, shadow, change of lighting conditions, etc.
2. Stauoer and Grimson [5] presented an adaptive background mixture model for real-time tracking. In their work, they modelled each pixel as a mixture of Gaussians and used an online approximation to update it.
3. L. Wang et al. [2] The Gaussian distributions of the adaptive Mixture models were then evaluated to determine the pixels most likely from a background process, which resulted in a reliable, real-time outdoor tracker which can deal with lighting changes and clutter.
4. Haritaoglu et al. [9] built a statistical model by representing each pixel with three values: its minimum and maximum intensity values, and the maximum intensity difference between consecutive frames observed during the training period. The model parameters were updated periodically.

5. McKenna et al.[19] used an adaptive background model combining color and gradient information, in which each pixel's chromaticity was modelled using means and variances, and its gradient in the x and y directions was modelled using gradient means and magnitude variances. Background subtraction was then performed to cope with shadows and unreliable color cues effectively.

E. Shape-based classification

For classifying moving object, for point silhouette analysis is used and for box blob analysis is available. From Both shape, box are more suitable for human detection. In image processing, a blob is defined as a region of connected pixels. Blob analysis is the identification and study of these regions in an image. The algorithms discern pixels by their value and place them in one of two categories: the foreground (typically pixels with a non-zero value) or the background (pixels with a zero value) [16].

Collins et al.[18] classified moving object blobs into four classes such as single human, vehicles, human groups and clutter, using a viewpoint-specific three-layer neural network classifier. Input features to the network were a mixture of image-based and scene-based object parameters such as image blob area, apparent aspect ratio of the blob bounding box, and camera zoom. Classification was performed on each blob at every frame, and the results of classification were kept in histogram.

Lipton et al. [1] area of image blob as classification metrics to classify all moving object blobs into humans, vehicles and clutter. Temporal consistency constraints were considered so as to make classification results more precise.

Kuno and Watanabe [25], silhouette-based shape representation for object classification, described a reliable method of human detection for visual surveillance systems. The merit of this method was to use simple shape parameters of human silhouette patterns to classify humans from other moving objects such as butterflies and autonomous vehicles, and these shape parameters were the mean and the standard deviation of silhouette projection histograms and the aspect ratio of the circumscribing rectangle of moving regions.

Advantages

1. It includes high flexibility and excellent performance.
- 2 It is possible that the image will contain some noise in it even though the human is detected in foreground.

Limitations

Clear background-foreground relation requirement and pixel-precision.

Drawback of Blob

1. The main difficulty in tracking is the discontinuity in the detected objects.
2. Due to movement of human in the video, it is possible that some parts of the human will appear separate. Also during background subtraction process, some errors can occurs. Thus legs or hands could appear as the separate objects from the body. It can be also caused by any noise or shadows present in the video. Due to this, the parts of body of a single object could be get tracked separately in a video.
3. Error could be occurred in counting also.

F. HOG and SVM Detection Techniques for HUMAN detection

In these techniques we can search human and vehicles.

Papageorgiou et al.[100] the human body is divided into the face, left arm, right arm and leg four parts, and then were trained four parts detectors, finally, according to the geometric constraints between the components to detect the entire human body.

Ronfard et.al.[100] "graphic structure" to describe the relationship between the human body the various components, using a first and second order gradient characterization of the appearance characteristics of the various components, and then use SVM classifiers to construct a human body detector [11].

Mikolajczyk [25] people use the direction. Location of the establishment of a joint histogram features based on human face, head and shoulders and other parts of the upper detector.

Leibel et al [25] extracted using key points, Hough voting and Chamfer distance from the template matching method, the establishment of a bottom-up and top-down combination of the human body detection method.

Papageorgiou et al [25] proposed the use of a modified Haar features and polynomial SVM as the classifier of the human body detector. Viola and others to use Haar-like features and movement characteristics of a cascade of Ad boosting combined classifier.

HOG by Dalal and others made a kind of human-performance single-window detection methods, the Chinese translation of the gradient histogram [6]. The method is to use small gradient on the direction of the image

histogram to describe the experimental results show that the combination of SVM classifiers described method can effectively separate the human and non-human regions.

Drawback

1. Due to light, goals and objectives of the changes in their attitude colors the impacts of the diversity of the target in different changes in different scenes is very large, so human or vehicle detection is a very difficult problem.

2 HOG & SVM detector requires a large representative sample of plus or minus to train classifier, quality are considered so need is to use many samples.

IV. Conclusion

We address here different techniques for motion detection and object classification: background subtraction, statistical methods, temporal differencing, optical flow, shape based Analysis and HOG & SVM detector. We compare these techniques and find out more suitable and convenient techniques for our project. The background subtraction is more convenient for static background. Statistical methods may be a better. In Shape based techniques Blob analysis are more suitable for tracking human. Optical flow and temporal isn't more suitable for real time image capturing. HOG & SVM detector need different quality samples to check accuracy

At the end of this survey, we have finding out some drawback of different techniques, in detailed discussions we will find more suitable techniques for human detection.

References

- [1] A.J. Lipton, H. Fujiyoshi, R.S. Patil, "Moving target classification and tracking from real-time video", Proceedings of the IEEE Workshop on Applications of Computer Vision, 1998, pp. 8–14.
- [2] A.M. Baumberg, D. Hogg, "Learning spatio-temporal models from training examples", Technical Report of University of Leeds, September 1995.
- [3] C. Bregler, "Learning and recognizing human dynamics in video sequences", Proceedings of the IEEE CS Conference on Computer Vision and Pattern Recognition, 1997, pp. 568–574.
- [4] Chih-Chang Chen, Hsing-Hao Lin and Oscar T.-C. Chen, "TRACKING AND COUNTING PEOPLE IN VISUAL SURVEILLANCE SYSTEMS", 978-1-4577-0539-7/11/\$26.00 ©2011 IEEE 1425 ICASSP 2011.
- [5] C. Staufer, W. Grimson, "Adaptive background mixture models for real-time tracking", Proceedings of the IEEE CS.
- [6] Cohen, I., Medioni, G.: "Detecting and tracking moving objects for video surveillance", in Proc. IEEE Computer Vision and Pattern Recognition, Fort Collins (CO), USA, June (1999).
- [7] E. Stringa, "Morphological change detection algorithms for surveillance applications", British Machine Vision Conference, 2000, pp. 402–411.
- [8] G.J. McLachlan, T. Krishnan, "The EM Algorithm and Extensions", Wiley Interscience, New York, 1997.
- [9] Guanglun Li, Yanling Wang and Weiqun Shu, "Real-Time Moving Object Detection for Video Monitoring Systems", IEEE 2008, 978-0-7695-3497-8/08.
- [10] H.A. Rowley, J.M. Rehg, "Analyzing articulated motion using expectation-maximization", Proceedings of the International Conference on Pattern Recognition, 1997, pp. 935–941.
- [11] I. Haritaoglu, D. Harwood, L.S. Davis, "real-time surveillance of people and their activities", IEEE Trans. Pattern Anal. Mach. Intell. 22 (8) (2000) 809–830.
- [13] J. L. Raheja, Sishir Kalita, Pallab Dutta and Solanki Lovendra, "A Robust Real Time People Tracking and Counting Incorporating Shadow Detection and Removal", International Journal of Computer Applications, May-2012, 0975-8887, Vol. 46, No. 4.
- [14] Julio Cezar Silveira Jacques Jr., Claudio Rosito Jung and Soraia Raupp Musse, "A BACKGROUND SUBTRACTION MODEL ADAPTED TO ILLUMINATION CHANGES", 1-4244-0481-9/06/\$20.00 C2006 IEEE 1817 ICIP 2006
- [15] J. Barron, D. Fleet, S. Beauchemin, Performance of optical flow techniques, Int. J. Comput. Vision 12 (1) (1994) 42–77.
- [16] Kavita P. Mahajan, S. V. Patil, "Tracking and Counting Human in Visual Surveillance System," International Journal of Electronics and Communication Engineering & Technology (IJECE), Volume 3, Issue 3, October- December (2012).
- [17] M. Kilger, 'A Shadow Handler in a Video-based Real-time Traffic Monitoring System', Siemens AG, Corporate Research and Development, ZFE ST SN 3 Otto-Hahn-Ring 6, D-8000 Munich 83, 0-8186-2840-5/92, 1992 IEEE
- [18] R.T. Collins, "A system for video surveillance and monitoring: VSAM Anal report", CMU-RI-TR-00-12, Technical Report, Carnegie Mellon University, 2000.
- [19] S.J. McKenna, "Tracking groups of people", Computer Vision Image Understanding 80 (1) (2000) 42–56.
- [20] N. Friedman, S. Russell, "Image segmentation in video sequences: a probabilistic approach", Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence, August 1–3, 1997.
- [21] Osama Masoud, Nikolaos Papanikolopoulos, "A Novel Method for tracking and Counting Pedestrians in Real-Time Using a Single Camera", IEEE 2001, 0018-9545/01, Vol. 50, No. 5.
- [22] Yuming Chen, "Study of moving object detection in Intelligent Video Surveillance System", 978-1-4244-6349-7/10/\$26.00 _c 2010 IEEE V7-62
- [23] Ya-Li Hou, *Student Member, IEEE*, and Grantham K. H. Pang, *Senior Member, IEEE*, "People Counting and Human Detection in a Challenging Situation", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, VOL. 41, NO. 1, JANUARY 2011.
- [24] Y.H. Yang, M.D. Levine, "The background primal sketch: an approach for tracking moving objects", Mach. Vision Appl. 5 (1992) 17–34.
- [25] Y. Kuno, T. Watanabe, Y. Shimosakoda, S. Nakagawa, "Automated detection of human for visual surveillance system", Proceedings of the International Conference on Pattern Recognition, 1996, pp. 865–869.