

Robust Real Time Multiple Human Detection and Tracking for Automatic Visual Surveillance System

LAKHYADEEP KONWAR, ANJAN KUMAR TALUKDAR, KANDARPA KUMAR SARMA
Dept. of ECE GUIST, Gauhati University, Guwahati, Assam, INDIA

Abstract—Detection of human for visual surveillance system provides most important rule for advancement in the design of future automation systems. Human detection and tracking are important for future automatic visual surveillance system (AVSS). In this paper we have proposed a flexible technique for proper human detection and tracking for the design of AVSS. We used graph cut for segment human as a foreground image by eliminating background, extract some feature points by using HOG, SVM classifier for proper classification and finally we used particle filter for tracking those of detected human. Our system can easily detect and track humans in poor lightening conditions, color, size, shape, and clothing due to the use of HOG feature descriptor and particle filter. We use graph cut based segmentation technique, therefore our system can handle occlusion at about 88%. Due to the use of HOG to extract features our system can properly work in indoor as well as outdoor environments with 97.61% automatic human detection and 92% automatic human detection and tracking accuracy of multiple human.

Keywords: Segmentations, Graph cut, HOG, SVM, Human detection, Human tracking, Particle filter, Occlusion handling.

Received: March 27, 2021. Revised: August 15, 2021. Accepted: August 31, 2021.

Published: September 15, 2021.

1. Introduction

Object detection is a challenging task in case of machine vision application. It is difficult to detect an object from an image if there is a crowded scene in the image [4]. For that purpose we have to use properly trained classifier, so that it can detect only that kind of objects [6]. Human detection from image is a newly identified challenging task owing to their different poses, color, size, shape [8] etc. It is also difficult to detect human from images due to the variable appearance [11] and disappearance of human in different images. Some other technique that has already been implemented gives some amount of accuracy but system provides slow response. Some other techniques are there where the system is unable to work in various lightening conditions as well as outdoor environment [14]. So we have to implement an algorithm where the accuracy will be better, system provides better response, provides proper result of detection in various lightening conditions, and works properly in indoor as well as outdoor environments.

All over the world including India several researchers are working in the area of designing a proper automatic visual surveillance system that includes human detection [8] [10], usual and unusual event detection [16], behaviour recognition [17] [23], action recognition or activity recognition [18] [19], human tracking [2] [15] [20] [21] [22] etc.

Bokov et. al. [3] used graph cut segmentation techniques to segment boundary and regions of an image from an N-dimensional images. In [4] Zhang used graph cut segmentation technique to segment multiple objects that can be used for occlusion handling in the area of tracking of multiple moving object from videos. Combination of graph cut based object

segmentation and Histogram of Oriented Gradient (HOG) for proper human detection was introduced by Lakshmi et. al. [5]. After that Ramya et. al. [6] and Kharabe et. al. [7] found that graph cut segmentation technique provides proper result for human as well as detection of other objects in the case of moving object detection purpose. The first concept of human detection was introduced by Dalal et. al. [8] that is the combination of HOG feature descriptor and Support Vector Machine (SVM) [9] classifier for proper human detection. They found that HOG feature descriptor provides better result of feature extraction and therefore human detection accuracy is increases. After that a method was introduced by Zhu et. al. [11] which is based on variable block size HOG feature descriptor that captures salient features of human object automatically. They identified the appropriate set of blocks from a large set of possible blocks and the use of Adaboost based feature selection by using integral image representation and a rejection of cascade that significantly speed up the computation. Kachouane et. al. [12] presented an algorithm for human detection and recognition in real time images that is based on the combination of HOG feature descriptor and SVM classifier are provides good results of detection and generally used for robotic tasks.

An another system that can properly work in terms of variation of size, partial occlusion, amount of context, noise and clutter situations was developed by Bell et. al. [13], where the method can automatically detect dismantled human at long range from a single, highly compressed images. Beaugendre et. al. [20] and Saboune et. al. [21] used particle filter based tracking of human for their system implementations. The particle filter based object tracking system provides good estimation of the 3D positions with the help of optimization of particles which is depend upon those of the likelihood function applied and therefore this technique can track people

that are newly enter to a scene and those are recovering from occlusions.

In this paper we have implemented an automatic multiple human detection and tracking algorithm to design an AVSS that can work properly in different lightening conditions, does not depend on shape variations, handle the occlusion, and improve the human detection and tracking accuracy. Our system design involves mainly four parts that comprises of maximum flow based graph cut technique for segmentation of human image, HOG as a feature descriptor, SVM as a linear classification purpose and finally we use particle filter based technique for human tracking. Due to the combinations of these, our system provides better accuracy as well as system provides better performance. This system is also robustly familiar with variation of lightening condition based detection, handle occlusions as well as it works in any environments.

The remaining part of this paper is systemized in the following ways: Section II provides the step by step theoretical considerations of proposed work on robust real time multiple human detection and tracking for automatic visual surveillance system or video surveillance system. Section III provides the experimental outputs and the experimental comparison with various testing videos as well as other existing methods. Finally section IV concludes the paper.

2. Theoretical Consideration

2.1 Human Detection and Tracking Model

Figure 1 shows the overall block diagram of proposed system; where first one is the input video dataset section, than next one is the graph cut segmentation section, third one is the feature extraction section, fourth one is the classification section and fifth one is the tracking section.

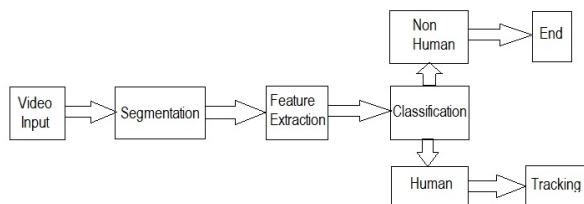


Fig. 1. Basic block diagram of the proposed system.

Our system is mainly design for video surveillance purpose. Therefore we need some of the basic hardware as well as software systems such as a mobile phone camera or CCTV camera, a PC, openCV 3.4.1 library, CodeBlock 10.12 library on C/C++ programming language by using Microsoft Visual Studio 2008 and MATLAB 2017a for calculating some graphical representations.

The automatic multiple human detection and tracking for AVSS section generally comprises of six blocks. They are input training dataset, input testing dataset, Graph cut segmentation and morphological operation, feature extraction, linear SVM based classification and Particle filter based tracking.

1) *Input Training Dataset:* We trained our detector by using the well-established MIT pedestrian dataset as positive dataset, where 888 positive images are there and we created 1000 negative images; that contains car, bike, house, tree, chair, staircase, animal etc. The size of each image is 128×64 which are in *height* \times *width* format and all are in color (RGB). Figure 2 and 3 shows the some samples of positive and negative dataset that we used to train our detector.



Fig. 2. Sample of some positive input dataset from MIT database.



Fig. 3. Sample of some negative input dataset from our database.

2) *Input Testing Dataset:* First we collect some video captured by using a mobile phone camera. The length of each video stream is about 5 minutes and all of those videos are in .mp4 format. The captured video quality is about 1024×768 and with 25 frames per second. We test our system for 10 videos. The system gives approximately equal accuracy for all of those video streams.

3) *Graph cut segmentation and morphological operation:* Graph cut segmentation is one of the image segmentation technique that can properly extract foreground from an image by eliminating the background [5]. Graph cut technique is basically based on the graph theory that can minimize the energy function by the use of max flow min cut theorem [6]. Generally a graph is a set of vertices V and edge E that connect various pairs of vertices. A graph can be written as

$$G = (V, E) \quad (1)$$

Where each edge can be represented by a pair of vertices. i.e.

$$E \subset V \times V \quad (2)$$

Here the graphs are often drawn as a set of point with curves connecting the points and the degree of a vertex is the number of edges incident on that vertex. We use mathematical morphology to refine the segmentation mask by using

differently shaped structuring elements [1]. The structuring element are some small sets or sub-images which is used to probe an image under the study for properties of interest. The structuring elements are a systematic about the direction of its origin. We use mathematical morphology operations opening, closing, erosion and dilations to reconstruct the loss portions of an image.

4) *Feature Extraction:* The basic blocks for HOG feature descriptor is shown in figure 4; where there are five major blocks are present. They are: gamma correction, gradient computation section, binning the orientations or weighted vote into spatial and orientation cells, contrast normalization over the overlapping blocks and feature vector collection blocks. Figure 5 shows the graphical representation of each steps of HOG feature descriptor for human detection.

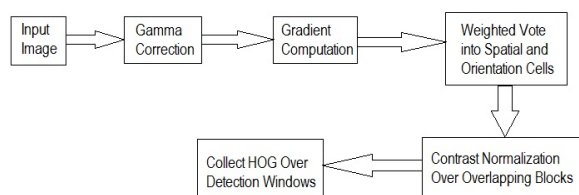


Fig. 4. Overall block diagram of HOG feature descriptor.

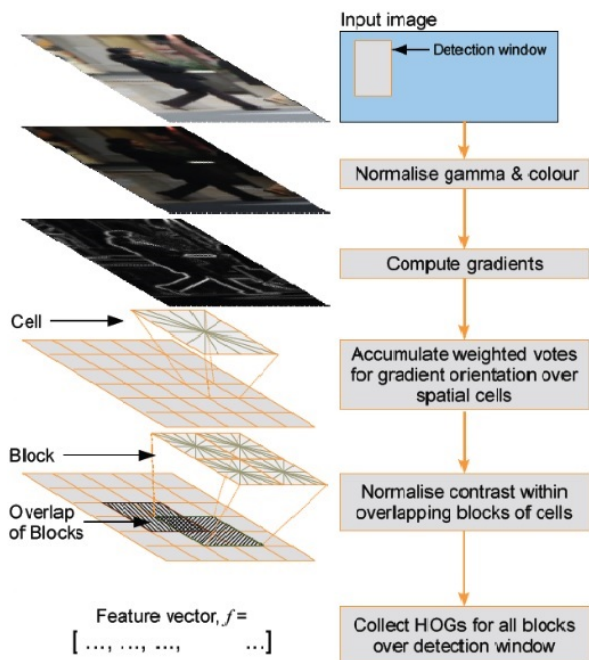


Fig. 5. Graphical representation of each steps of HOG feature descriptor for human detection (Source: Dalal and Triggs, 2005).

Table 1 shows the specifications of HOG feature descriptor that we have used for our human detection purpose. For block normalization section we use the L2-norm of block normalization technique. Where we have to consider a non normalized vector v that contain all of the histograms in a given block $\|v\|_k$ be the k^{th} number of norms, where $k=1,2$

TABLE I

SOME SPECIFICATIONS OF OUR HOG FEATURE DESCRIPTOR

Bin size used	16×16
Bin range	0 to 255
Color space	RGB
Pixel per cell	16
Cell per block	4
No of histogram bin	9
Angle of notation	0 to 180 degree
Histogram range	0 to 255

and ϵ be a small constant. Then the normalization factor be the following forms:

$$f = \frac{v}{\sqrt{\|v\|_2^2 + \epsilon^2}} \quad (3)$$

5) *Linear SVM based Classification:* Our main goal is to detect and track human in videos or sequence of images, therefore we have to classify human and non-human data from our input video dataset for human detection purpose. First we classify the human and non human data by using SVM classifier and then pass it to the next step for tracking purpose. Here we use linear SVM with polynomial kernel, and human and non-human classes are considered as +1 and -1.

6) *Particle Filter Based Tracking:* The basic algorithm which is used for person tracking are discussed below:

- **Initialization:-** First we represent $P(X_0)$ by a set of N number of samples $(s_0^{k,-}, w_0^{k,-})$, where $s_0^{k,-} \sim P_s(S)$ and $w_0^{k,-} = P(w_0^{k,-})/P_s(S = s_0^{k,-})$. Ideally $P(X_0)$ has a simple form where $s_0^{k,-} \sim P(X_0)$ and $w_0^{k,-} = 1$.
- **Prediction:-** Represent $P(X_i \setminus y_0, y_{i-1})$ by using $(s_i^{k,-}, w_i^{k,-})$. Where $s_i^{k,-} = f(s_{i-1}^{k,+}) + \xi_i^k$ and $\xi_i^k \sim N(0, \sum d_i)$.
- **Correction:-** Represent $P(X_i \setminus y_0, y_i)$ by using $(s_i^{k,+}, w_i^{k,+})$. Where $s_i^{k,+} = s_i^{k,-}$ and $w_i^{k,+} = P(Y_i = y_i \setminus X_i = s_i^{k,-})w_i^{k,-}$.
- **Re-sampling:-** Normalize the weight to set $\sum_i w_i^{k,+} = 1$ then compute the variance of the normalized weights. If the variance exceeds some threshold value then construct a new set of samples by drawing and replacement N numbers of samples from the old set by using the weights as the probability that of a sample will be drawn. Then the weight of each samples will be $\frac{1}{N}$.

3. Experimental Outputs and Comparisons

Some of the detection and tracking results of automatic multiple human detection and tracking for visual surveillance videos that are captured under different environments are shown in figure 6 to 10. We have tested our system for 10 video streams. Here human detection and tracking accuracy is found to be satisfactory for all of the videos.

Figure 6 shows detection and tracking of human in different frames such as frame no. 1, 375, 975 and 1525. Here video was captured on trunk road on a holiday, where a person is walking far against the camera positioned. In this videos other peoples are also there, not all but some of them are detected. This is mainly due to the limitation of detection range of our system. Figure 7 shows detection and tracking of human in



Fig. 6. Human detection and tracking result for video 1.

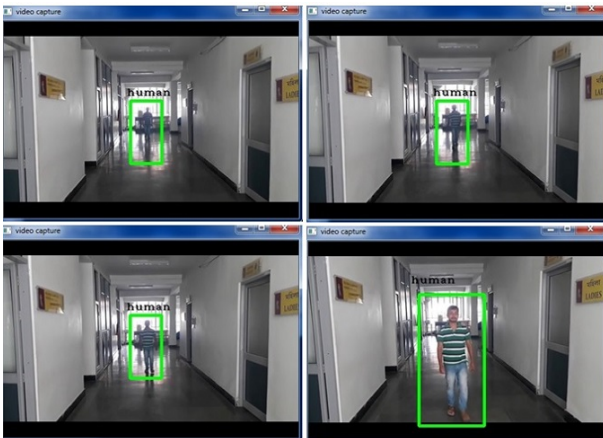


Fig. 7. Human detection and tracking result for video 2.

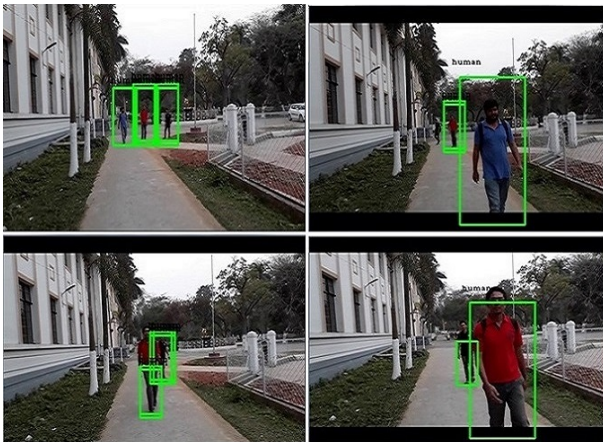


Fig. 8. Human detection and tracking result for video 4.

frame no. 1, 250, 725 and 900. Here video was captured in a corridor of a medical cum research center, where one human was walking toward the camera positioned. Figure 8 shows detection and tracking of human in frame no. 1, 1275, 4825 and 6015. Here three humans were present in this video. In frame no. 1 all of three human are detected, frame no. 1275 two walking human are detect and tracked, next in frame no. 4825 and 6015 two occluded human are detected and tracked. Here detection and tracking of humans at about 88% under occlusions takes place. Figure 9 shows detection and tracking



Fig. 9. Human detection and tracking result for video 5.

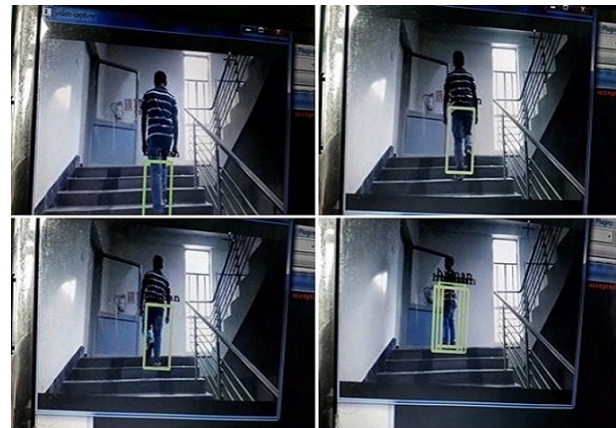


Fig. 10. Human detection and tracking result for video 6.

of human in different frames such as frame no. 1, 425, 975 and 1600. Here two humans were walking on a footpath against the camera positioned. Similarly, in figure 10 shows detection and tracking of human in different frames such as frame no. 1, 151, 297 and 514, where a human was walking on a staircase of a shopping mall against the camera positioned.

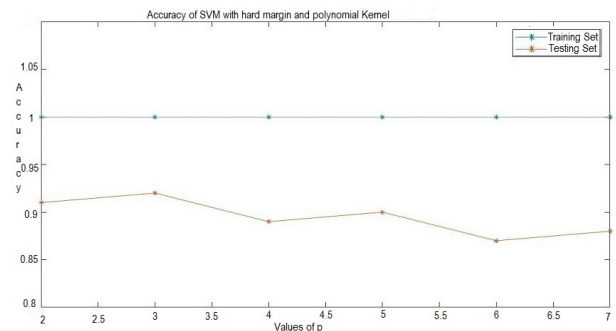


Fig. 11. Accuracy of training and testing for human detection and tracking by using hard margin with different values of polynomial kernel.

Figure 11 shows that the accuracy of training and testing for human detection and tracking with polynomial kernel [9]. Here we used hard margin with different values of polynomial kernel. Here we consider training accuracy is about 1 or 100%. Then we have found that detection and tracking accuracy is maximum at polynomial value $p = 3$, which is about 92%.

Figure 12 shows that the accuracy of training and testing for human detection and tracking by using soft margin [9], where we use different kernel value of polynomial kernel against C and p value. Here p is the different values of polynomial kernel and C be the regularization parameter. Regularization parameter C is used to control the trade off between the achieving a low training and testing error that is the ability to generate a classifier to unseen data. Here accuracy of human detection and tracking is found to be maximum at polynomial values from $p = 4.5$ with value of regularization parameter $C = 0$ to 2.5 which is about 92%.

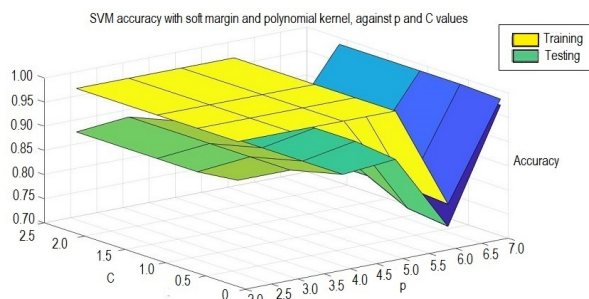


Fig. 12. Accuracy of training and testing for human detection and tracking by using soft margin with different kernel value of polynomial kernel against C and p value.

3.1 Performance Evaluation of Human Detection

Performance evaluation/measurement is necessary for detection of object as well as classify different objects. We use confusion matrix for the calculation of those parameters that we needs. Confusion matrix has a link between the actual values and predicted values of object detection.

1) *Precision*: Here precision is used to describe the accuracy in picking out a particular type of target (e.g. human or non-human) from data that contains both the human and non-human data. If TP be the true positive rate and FP be the false positive rate of detection, than we have the precision value from the following equation as:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

2) *Recall*: Recall is used to describing the success in finding an item in a database (e.g. human or non-human database). Therefore we have the recall value by using following equation:

$$Recall = t_{pr} = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (5)$$

3) *F-measure*: To indicate overall performance of a classification system we can use the F-measure parameter. We have the equation for F-measure parameter equation as:

$$F - measure = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} \quad (6)$$

We have captured videos for our system in different scenarios. Table 2 shows the total no. of frames of each videos, positive samples, negative samples, true positive, true negative,

TABLE II
TABLE FOR SERIAL NO.(SN), NO. OF FRAMES PER VIDEOS(NFPV), POSITIVE SAMPLES(PS), NEGATIVE SAMPLES(NS), TRUE POSITIVE(TP), TRUE NEGATIVE(TN), FALSE POSITIVE(FP), FALSE NEGATIVE(FN) VALUES OF ALL TESTED VIDEOS

SN	NFPV	PS	NS	TP	TN	FP	FN
1	1567	2616	3265	2519	3150	115	97
2	1049	1049	4430	1024	4296	134	25
3	496	992	6944	952	6890	54	40
4	6082	14160	92271	13657	90806	1465	503
5	1656	3529	19217	3399	18547	670	130
6	621	621	2215	599	2181	34	22
7	5260	15018	20968	14758	20378	590	260
8	943	2829	9208	2749	9007	201	80
9	1543	14301	21960	13823	81499	461	478
10	3727	20124	79715	19438	78802	913	686

TABLE III
TABLE FOR SERIAL NO.(SN), TRUE POSITIVE RATE(TPR), TRUE NEGATIVE RATE(TNR), FALSE POSITIVE RATE(FPR), FALSE NEGATIVE RATE(FNR), PRECISION(P), RECALL(R), F-MEASURE(F), ACCURACY(A) IN % OF ALL TESTED VIDEOS

SN	TPR	TNR	FPR	FNR	P	R	F	A
1	96.30	96.50	3.50	3.70	95.63	96.30	95.96	96.40
2	97.60	97.00	3.00	2.40	88.40	97.62	92.78	97.10
3	96.00	99.23	0.77	4.00	96.00	94.60	95.29	98.81
4	96.40	98.40	1.60	3.60	90.30	96.40	93.30	98.20
5	96.30	96.50	3.50	3.70	80.00	96.30	87.40	96.48
6	96.46	98.50	1.50	3.54	97.72	96.46	97.10	98.02
7	98.30	97.20	2.80	1.70	96.20	98.30	97.24	97.60
8	97.17	97.72	2.18	2.89	93.20	97.17	95.15	97.67
9	96.66	97.90	2.10	3.34	96.80	96.66	96.73	97.40
10	96.60	98.85	1.15	3.40	95.51	96.60	96.05	98.40

TABLE IV
COMPARISON TABLE FOR PERFORMANCE EVALUATION OF OUR METHOD WITH SOME EXISTING METHODS. (HERE P=PRECISION, R=RECALL, F=F-MEASURE, DR=DETECTION RATE, DTR=DETECTION AND TRACKING RATE, POD=PURPOSE OF DESIGN)

Authors	P in %	R in %	F in %	DR in %	DTR in %	POD
Lakshmi et. al. [5]	92	96	NA	NA	NA	Human detection from video
Ramya et. al. [6]	89.11	92.83	89.25	NA	NA	Moving object detection
Dalal et. al. [8]	NA	NA	NA	89	NA	Human detection from image
Davis et. al. [10]	100	92.73	NA	95.29	NA	Human detection for robotic tasks
Kachoune et. al. [12]	NA	NA	NA	86	NA	Fast human detection from video
Kushwaha et. al. [15]	NA	NA	NA	NA	87.44	Human detection & tracking
Beaugendre et. al. [20]	NA	NA	NA	NA	80	Human tracking only
Mishra et. al. [24]	NA	NA	NA	77.78	NA	Human motion detection
Wo et. al. [25]	NA	NA	NA	80	80	Head shoulder detection to track multiple human in meeting
Proposed	92.98	96.64	94.70	97.61	92	Robust real time multiple human detection & tracking for Automatic visual surveillance system

false positive, false negative values for all of those tested video streams that are captured in different locations. Table 3 shows the true positive rate, true negative rate, false positive rate, false negative rate, precision, recall, F-measure and accuracy of human detection for all of those tested video streams that are captured in different locations. Table 4 shows the comparison table of performance evaluation for our system with different existing methods, which shows that our system provides better response as well accuracy for robust real time multiple human detection and tracking for AVSS. Here NA means not available, not applicable or not announced. In our system, we have tested 10 videos, all of them were captured in outdoor as well as indoor environments to test our system may work or nor in different environments.

4. Conclusion

Object detection from a video stream became a challenging task in case of the design of AVSS. We have mainly focused our work to solve two problems, they are: our system must work in different lightening conditions and occlusion handling. Our system can give approximately 92% accurate result of automatic multiple human detection and tracking with an average automatic multiple human detection accuracy of 97.61%. We use Graph cut segmentation so that the system can automatically detect and track multiple humans in occluded scenarios. This means that our system can handle occlusion properly. The system can work properly in different environments such as poor lightening conditions, variation of lightening conditions, and can handle occlusions.

References

[1] X. Benavent, E. Dura, F. Vegara, and J. Domingo. (2012, Nov.). Mathematical morphology for color images: an image-dependent approach. *Mathematical Problems in Engineering* [Online]. vol. 2012, Article ID 678326, 18 pages, 2012. Available: <https://www.hindawi.com/journals/mpe/2012/678326>.

[2] D. V. Thombre, J. H. Nirmal and L. Das, "Human detection and tracking using image segmentation and kalman filter," *2009 Int. Conf. on Intelligent Agent & Multi-Agent Systems*, Chennai, 2009, pp. 1-5.

[3] Y. Y. Boykov and M. P. Jolly. "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," in *Proc. 8th IEEE Int. Conf. on Computer Vision. ICCV 2001*, Vancouver, BC, Canada, 2001, pp. 105-112 vol. 1.

[4] Q. Zhang and K. N. Ngan (2011, Nov.). Segmentation and tracking multiple objects under occlusion from multiview video. *IEEE Trans. on Image Processing* [Online]. 20(11), pp 3308-3313. Available: <https://ieeexplore.ieee.org/document/5872044>.

[5] N. D. Lakshmi, Y. M. Latha and A. Damodaram, "Silhouette extraction of a human body based on fusion of HOG and graph-cut segmentation in dynamic background," *3rd Int. Conf. on Computational Intelligence and Information Technology (CIIT 2013)*, Mumbai, 2013, pp. 527-531.

[6] A. Ramya and P. Raviraj, "Performance evaluation of detecting moving objects using graph cut segmentation," *2014 Int. Conf. on Green Computing Communication and Electrical Engineering (ICGCCEE)*, Coimbatore, 2014, pp. 1-6.

[7] S. R. Kharabe, P. S. Hanwate, D. S. Gaikwad and K. P. kaliyamurthie, "Human image segmentation," *2017 Int. Conf. on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET)*, Chennai, 2017, pp. 1-3.

[8] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for human detection," *2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005, pp. 886-893 vol. 1.

[9] C. P. Diehl and G. Cauwenberghs, "SVM incremental learning, adaptation and optimization," in *Proc. of the Int. Joint Conf. on Neural Networks, 2003.*, Portland, OR, 2003, pp. 2685-2690 vol. 4.

[10] M. Davis and F. Sahin, "HOG feature human detection system," *2016 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*, Budapest, 2016, pp. 002878-002883.

[11] Q. Zhu, S. Avidan, M. C. Yeh, and K. T. Cheng, "Fast human detection using a cascade of Histograms of Oriented Gradients," *2006 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'06)*, New York, NY, USA, 2006, pp. 1491-1498.

[12] M. Kachouane, S. Sahki, M. Lakrouf and N. Ouadah, "HOG based fast human detection," *2012 24th Int. Conf. on Microelectronics (ICM)*, Algiers, 2012, pp. 1-4.

[13] A. E. Bell, "Robust feature vector for efficient human detection," *2013 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Washington, DC, 2013, pp. 1-5.

[14] T. Zhao and R. Nevatia (2004, Sept.). Tracking multiple human in complex situations. *IEEE Trans. on pattern analysis and machine intelligence* [Online]. 26(9), pp. 1208-1221. Available: <https://ieeexplore.ieee.org/document/1316854>.

[15] A. K. S. Kushwaha, C. M. Sharma, M. Khare, R. K. Srivastava and A. Khare, "Automatic multiple human detection and tracking for visual surveillance system," *2012 Int. Conf. on Informatics, Electronics & Vision (ICIEV)*, Dhaka, 2012, pp. 326-331.

[16] A. Adam, E. Rivlin, I. Shimshoni and D. Reinitz (2008, March). Robust real time unusual event detection using multiple fixed location monitor. *IEEE Trans. on pattern analysis and machine intelligence* [Online]. 30(3), pp. 555-560. Available: <https://ieeexplore.ieee.org/document/4407716>.

[17] A. Galata, N. Johnson, and D. Hogg, "Learning behaviour models of human activity," in *Proc. British Machine Vision Conf.*, 1999, pp. 12-22.

[18] A. K. R. Chowdhury and R. Chellappa, "A factorization approach for activity recognition," *2003 Conf. on Computer Vision and Pattern Recognition Workshop*, Madison, Wisconsin, USA, 2003, pp. 41-41.

[19] O. Chomat and J. L. Crowley, "Probabilistic recognition of activity using local appearance," in *Proc. 1999 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (Cat. No PR00149)*, Fort Collins, CO, USA, 1999, pp. 104-109 Vol. 2.

[20] A. Beaugendre, H. Miyano, E. Ishidera and S. Goto, "Human tracking system for automatic video surveillance with particle filter," *2010 IEEE Asia Pacific Conf. on Circuits and Systems*, Kuala Lumpur, 2010, pp. 152-155.

[21] J. Saboune and R. Laganieri, "People detection and tracking using the explorative particle filtering," *2009 IEEE 12th Int. Conf. on Computer Vision Workshops, ICCV Workshops*, Kyoto, 2009, pp. 1298-1305.

[22] S. Rahimi, A. Aghagolzadeh and H. Seyedarabi, "Human detection and tracking using new features combination in particle filter framework," *2013 8th Iranian Conf. on Machine Vision and Image Processing (MVIP)*, Zanjan, 2013, pp. 349-354.

[23] Y. Chen, G. Liang, K. K. Lee and Y. Xu, "Abnormal behaviour detection by multi-SVM based bayesian network," *2007 Int. Conf. on Information Acquisition*, Seogwipo-si, 2007, pp. 298-303.

[24] S. K. Mishra and K. S. Bhagat (2015, July). Human Motion Detection and Video Surveillance Using MATLAB. *International Journal of Scientific Engineering and Research (IJSER)* [Online]. 3(7), pp. 154-157. Available: <https://www.ijser.in>.

[25] Bo Wu and Ram Nevatia. "Tracking of Multiple Humans in Meetings," in 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW06), New York, NY, USA, 2006, pp. 143-148. doi:10.1109/CVPRW.2006.205.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US