A NOVEL IMPLEMENTATION OF PEDESTRAIN DETECTION USING HOGD AND SVM ALGORITHMS

Rajana Kanaka Raju*, V.Lakshmi,**
Department of Computer Science Engineering, GVPCDPGC(A). Visakhapatnam.
Department of Computer Applications, GVPCOE(A). Visakhapatnam.

Received on: 01/08/2023 Accepted on:05/09/2023

Keyword:
Template matching
Support vector machine
Gaussian mixture model
Pedestrian detection

ABSTRACT (9 PT)
Human detection in real time is an extensive, hard and essential area of research. It offers a wide variety of human recognition and video surveillance applications, etc. Computer vision is a scientific subject which seeks to replace human vision. The great problem for humans is recognising or counting people in a crowded environment. There are several CCTV monitoring systems placed in different areas, such as religious gatherings, airports, and public sites, that allow for improved crowd monitoring systems. The system suggested identifying and counting the number of people in each frame. In every frame, the system can detect people. The system built using HOGD and SVM algorithms can estimate and anticipate the high-density crowd area in real time. It gives individual estimations and various statistical measurement combinations. The techniques of classification and regression analysis are used for the support vector machine (SVM). The trained SVM is the key approach for the human detection algorithm that searches for human-like motion patterns in optical flows. The system creates a human model based on the regional change characteristic of many pictures of pedestrians and detects two or more individuals who intersect with the image of difference and the human model. There are numerous advantages, such as safety monitoring, suspect detection of behaviour, people in densely crowded regions, etc.

Corresponding Author: Email: rkanakaraju@gvpcdpgc.edu.in

INTRODUCTION

Nearly every day, people encounter crowds or mass meetings at different locations, such as entertainment events, airports, universities, sports stadiums, and theme parks. The events are very varied and span social, cultural, and religious activities. In contrast to social and sports-related events, it can not be prevented from the multitude of people experiencing important religious events. It is therefore important to have an intelligent Crowd Monitoring System that guarantees public safety, maintains high flow rates for pedestrians to prevent stampeded traffic, provides better emergency services in the event of a multitude-related emergency, and optimises resources to ensure good accessibility by avoiding congestion.A crowd may expand in no time and the organisers' management of the crowd might become extremely difficult. Problems like aberrant behaviour or even severe conditions may arise in these instances. The crowds may now be observed in real time 1 through the vector holder. The task has been downloaded to a distant node rather than to local video data processing. Similarly, an SVM crowd monitoring system has been created. Their technique was able to correctly detect the crowd. In many real-world applications, human video detection (i.e., picture series) plays an important part (e.g., visual surveillance and automated driver assistance). For many reasons, the task of detecting humans in a series of aerial pictures is difficult.

LITERATURE REVIEW

In paper [1] Robust detection and tracking with an active camera human faces they propose an effective frame for the detection and tracking in an active camera of human faces. The coefficient Bhattacharyya is used as a measure of similarity between the colour distribution of the face pattern and the face candidate. The correct derivation of these distributions permits the application of the Bhattacharyya spatial gradient to lead the quick search for the best candidate in face. Optimization, based on a mean shift analysis, needs just a few converging rounds. The scale variations in the tracked face are managed by using the invariable scale of the measure of similarity and the calculated brightness gradient at the boundary of the hypothesised face area. The detection and tracking modules are almost similar, since the detection includes mean shift optimization with repeated initialations. The camera control system is implemented in dual mode and is designed to transition between pan, tilt and zoom camera smooth pursuits and saccadic movements, depending on the objective presence in the fovea area. The resultant system operates on a normal PC in real time and is resistant to partial occlusion, conflict, fluctuations in face size, rotating depths and rapid changes in the subject / camera location..in paper [2] The objective of this paper is to examine the state-of-the-art developments in visual tracking techniques, divide them into several categories and identify future trends. In many computer vision applications visual tracking is a key task and has been extensively researched in recent decades. Although several methods have been suggested, effective visual tracking remains an enormous problem. Visual tracking difficulties may occur because of rapid object movements, changing appearance pattern, non-rigid object structures, occlusion and camera movement. We first examine the state-of-the-art feature descriptors that depict the look of tracked objects in this paper. Then we classify the progress in tracking into three categories, give comprehensive descriptions of typical approaches in each category, and discuss their good and bad features. Finally, we describe potential directions for research on visual tracking

In paper [3] author has explained about The detection and tracking of faces and facial characteristics in video sequences is a major and difficult computer vision issue. This study field has numerous applications for face recognition systems, model-based code, gaze detection, interaction between human and computer, teleconferencing, etc. This study offers a real-time system for facial detection and tracking in video sequences. A skin colour model is utilised in the face for the segment facing candidate areas. The existence or absence of a face in each area is checked by an eye sensor based on an effective template. The pupils, nostrils and lip corners are identified after a face is recognised and these facial characteristics are monitored in the picture sequence. The suggested system is effective in the real- time tracking of a mobile individual under various circumstances including lighting, movement and rotation.In paper [4] author has explained about This paper presents a human face detection and recognition system for colour picture series. The system is made up of two subsystems: a subsystem for human face detection and a subsystem for human face identification. The face detection subsystem consists of two modules: face and face check. Using the skin colour analysis and motion analysis, the module for identifying the human face identifies the face areas of many individuals in colour picture series. The To validate the identified faces of human beings by evaluating the eclipse and supporting vector machine (SVM) and to accurately locate human faces by finding the eyes and mouths on the basis of generalised symmetry transformation. The characteristics of the relationship between face patterns may be retrieved and chosen using the main component analysis. With these chosen characteristics, we can ultimately categorise human faces by training multiple SVMs. Moreover, many basic and sophisticated techniques are employed in these modules to decrease the search space. The system may thus operate at a fast speed, high detection and detection rate. Under controlled lighting circumstances, human face detection accuracy of the system is 97.2 percent. The accuracy of the human face recognition for 70 individuals is 96,5% (with 20 vectors) and 98,3%. (with 30 eigenvectors). In paper [5] author present a sub window scanning technique and a face/non-face classification using a facial colour filter in this paper. The scanning process skips sub- windows that do not include potential faces based on the facial colour feature. The face-to-face classificator in facial colour has minimal computer costs and in the early approach lowers the total time required to calculate the face detection and eliminates false alarms.

METHODOLOGY

The system being proposed is a novel rapid human detection system and an environmentally-related counting system where many persons have a system view. In the area, we build a human model for numerous pictures of pedestrians who move in different directions and identify two or more persons who overlap by matching the difference image and the human model. And we propose filtering the axial direction to simplify the matching procedure. In the counting system, we also propose the algorithm which counts the pedestrian line passage utilising passage information. The experiment in the environment where numerous persons exist in the view of the camera demonstrates that the proposed systems can function with adequate precision in real time. In Figure 1 shows the flowchart of the proposed method, which can be illustrated in detail as follows:

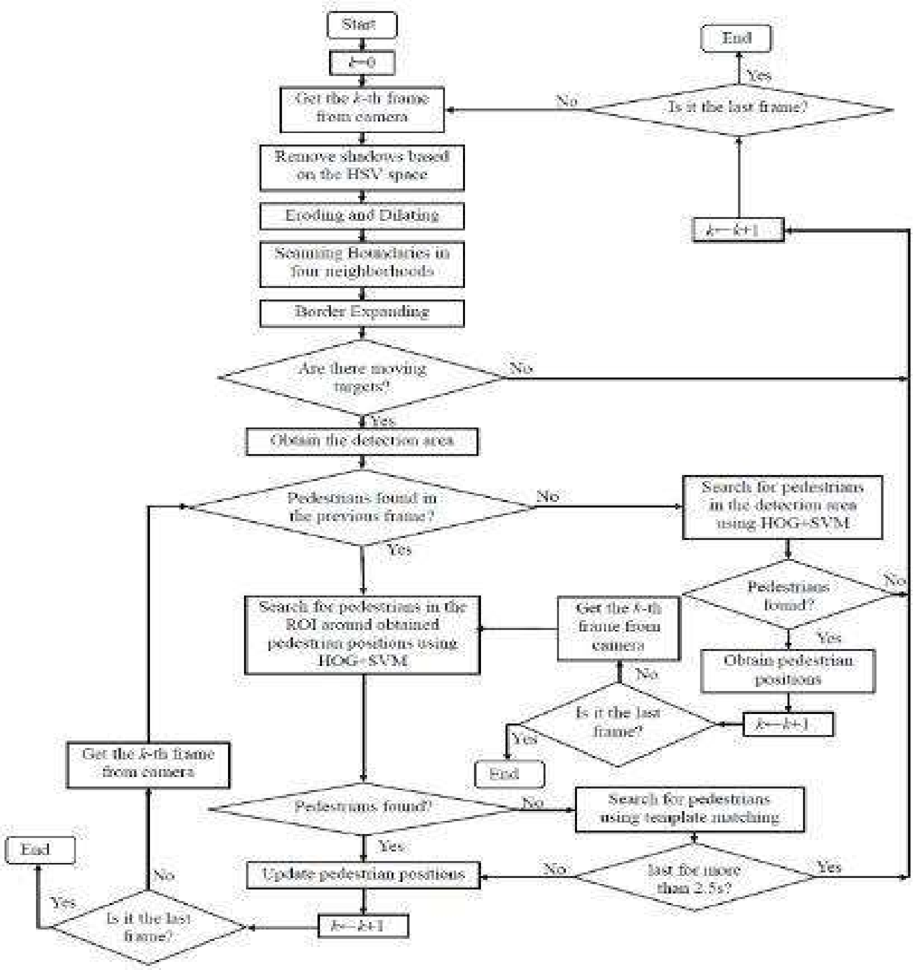

FIG 1: The flowchart of the proposed method

Proposed algorithm
Step 0: *set k to 0.* ***Step 1:*** *Get the k-th camera frame.* ***Step 2:*** *Remove shadows using HSV space, erode and dilate, scan boundaries in four neighborhoods, and increase borders.Check for moving targets in the preprocessed frame. If affirmative, receive the detection area and proceed to Step 4. Otherwise, continue to Step 3.* ***Step 3:*** *Update k by k+1, then stop the program if k reaches the final frame. Otherwise, step 1.* ***Step 4:*** *If the preceding frame had pedestrians, move to Step 7. Otherwise, continue to Step 5.* ***Step 5:*** *Use HOG+SVM to find pedestrians in the detection area. Step 6: If pedestrians are detected, place them. Otherwise, continue to Step 3.* ***Step 6:*** *Update k by k+1, then stop the procedure if k reaches the final frame.Otherwise, receive the camera's k-th frame. Proceed to Step 7.* ***Step 7:*** *Use HOG+SVM to find pedestrians in the ROI surrounding pedestrian spots. If pedestrians are spotted, update positions and continue to Step 8.Otherwise, step 9.* ***Step 8:*** *Update k by k+1 If k reaches the final frame, halt the program.Otherwise, receive the camera's k-th frame. Proceed to Step 4.* ***Step 9:*** *Use template matching to find pedestrians. Step 3 if template matching lasts longer than 2.5s frames and no pedestrians are discovered.Otherwise, update pedestrian placements, Step 8.*

The Oriented Gradients Histogram technique is primarily used for facial recognition and image detection to categorise images as above. This area has a number of uses, from independent cars to surveillance methods to more intelligent advertising. This essay gives a deep insight into the implementation of the HOG detection technique and explains why it is still used so often.

Feature descriptor simply means the representation of an image that simply for HOG functional descriptors, we may additionally convert the image to a feature vector of the length n specified by the user (width x height x channels). Although the images may be difficult to view, they are ideal for image classification algorithms such as SVMs to provide good results.

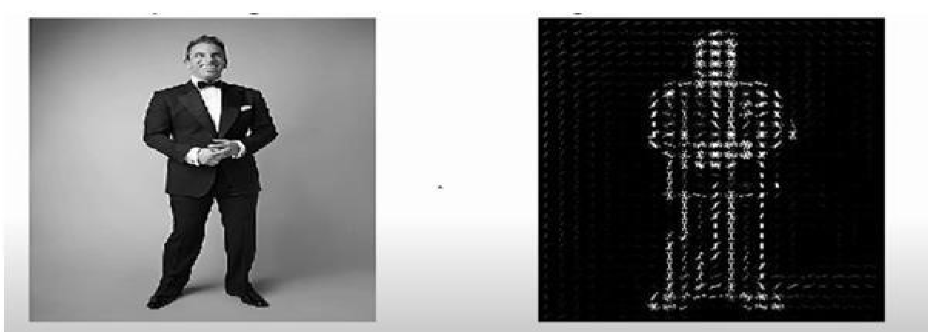

FIG 2: CONVERT THE IMAGE INTO A FEATURE VECTOR OF LENGTH

Preprocessing

A major error that people frequently make when they detect HOG objects is to neglect to pre-process the image to a defined aspect ratio. A typical aspect ratio is 1:2, so that your images may range from 100x200, 500x1000, etc.

Make sure you identify the section you want for a specific image so it matches the aspect ratio properly and enables better accessibility in the long run.

The section is identified by FIG 3 and the aspect ratio fits.As shown above, we must compute the corresponding horizontal and vertical gradients to provide the histogram that may be utilised later on in this method to make the HOG feature descriptor possible. You may accomplish this by simply filtering the image through the kernels:

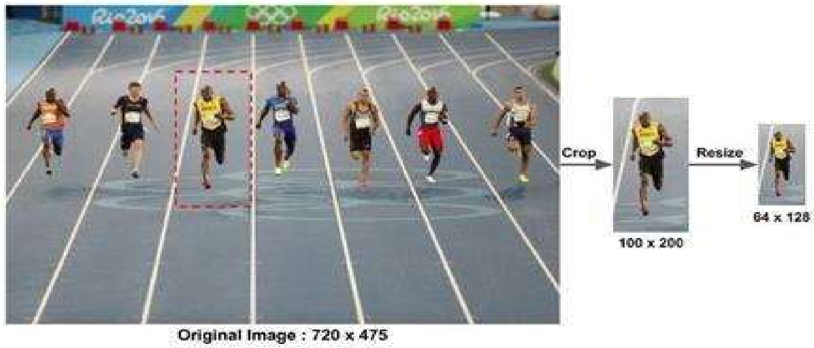

FIG 3 identify the section and correctly fits the aspect ratio.

Kernels such as these are usually employed in the categorization of images, primarily in convolutionary neural networks, to identify the borders and key points of a specific image. The magnitude and direction of the gradients may then be simply determined using the following formulas.While the following formulas may seem quite intimidating, just remember that most of this implementation is done by the computer, thus it is not really necessary to calculate these gradients ourselves.The picture below shows an optimal example of gradients as they fire throughout these gradients along the edges of the images. The superfluous material is eliminated as the backdrop and only the most important portions remain.

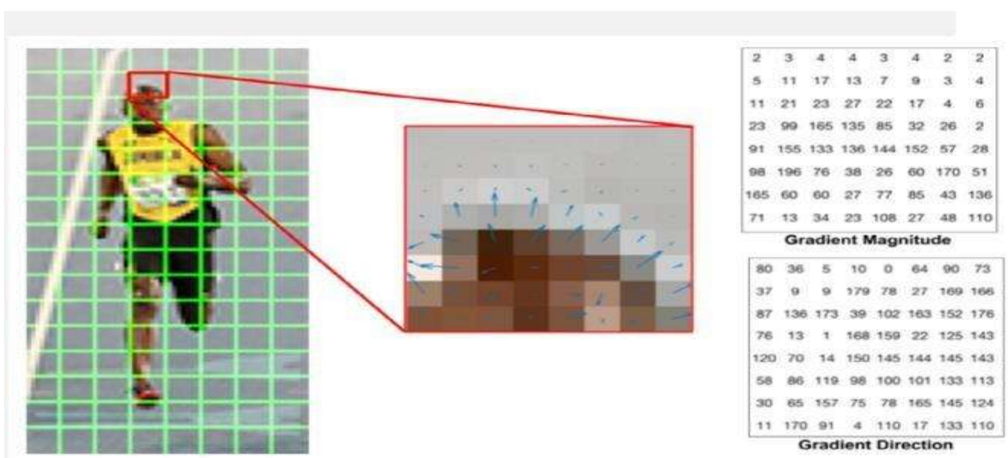

FIG 4:IMAGE TO GRADIENT MAGNITUDE AND GRADIENT DIRECTION

Making A Histogram From These Gradients,Make a histogram of the gradients.To proceed to the next stage of the HOG method, ensure that you split the image into cells so you can compute the histogram of gradients for each cell. If, for example, you have a 64x128 image, split your image into 8x8 cells (this means a little tinkering and guessing!).Feature descriptors enable a simple and concise description of specific patches of the

picture; an 8x8 cell may simply be described by 128 integers, using our example from above (8x8x2 where the last 2 are from the gradient magnitude and directional values). By transforming these values to histograms, we enable an image patch to be considerably more robust and compact.
Make sure the histogram is divided into nine distinct bins, each with an angle of 0-160 increments of twenty. This is an example of what an image with the appropriate size and gradient might appear like (notice the arrows get larger depending on the magnitude).Now, how can we determine where the histogram contains each pixel? It's simple. It's simple. A bin is picked according to the specified direction and the value within the bin depends on the magnitude. Note that if the pixel is between two bins, the magnitudes are divided according to their distance from each bin. After this process, a histogram may be created and the weight bins can be readily observed.

Lighting changes are another significant issue that may disturb the calculation of these gradients. For example, if the picture was darker by 1/2 of the present gradient, the degrees of gradients and the histogram magnitudes would all be halved. We thus want our descriptor to be free of changes in illumination so that it is impartial and effective.The usual normalization process happens simply by calculating the length of a vector by its magnitude and then simply by dividing all components of the vector by length. For example, if you had a vector of [1,2, 3], then the vector length would be the square root of 14, following the fundamental mathematical concepts. By splitting the vector at this length, you reach your normalized new [0.27, 0.53, 0.80] vector.This normalization process may be done according to your choice (whether you want to perform on an 8x8 block or even a larger 16x16 block). Just remember to convert these blocks into element vectors so that the above-mentioned normalization may be carried out.The descriptors of HOG are frequently seen with the image on the right to ensure a precise depiction of the person's form. This visualization may be quite helpful in understanding where the gradients are moving and where the items are within the image.

CONCLUSION

A crowd may grow in no time and managing the crowd can be extremely difficult for the organisers. In these instances, there may be issues such as aberrant behaviour or even stamping. The crowds may now be observed in real time through vector support. Instead of processing local video, the job has been downloaded to a distant node. Similarly, a crowd surveillance system has been created using SVM. Their technique was able to correctly identify the crowd. The system will enable organisations and branches of crime and the police to know how many individuals are moving. Since numerous calculations are necessary, it would be quite difficult to calculate, yet for a good result or accuracy, the system would be built. The system limitation is that the images of input are rotated and can not be detected. The second limitation of the system is that it may not be good if we provide poor quality, and it would be tough to extract the functionality from hogs. It works for black and white pictures and movies, but it can not detect when the darkness is strong in a black and white image or video. In future, we shall test this thesis against robust constraints and detect several human beings.

REFERENCES

[1] D. Comaniciu and V. Ramesh, “Robust detection and tracking of human faces with an active camera,” in Proceedings of Third IEEE International Workshop on Visual Surveillance, pp. 11–18, 2000.
[2] H. Yang, L. Shao, F. Zheng, L. Wang, and Z. Song, “Recent advances and trends in visual tracking: A review,” Neuro Computing, pp. 1–9, 2011.
[3] I. Hemdan, S. Karungaru, and K. Terada, “Facial features-based method for human tracking,” Workshop on Frontiers of Computer Vision, 2011.

[4] J. Yang, X. Ling, Y. Zhu, and Z. Zheng, “A face detection and recognition system in color image series,” Mathematics and Computers in Simulation, pp. 531–539, 2008.
[5] Y. N. Chae, J. Chung, and H. S. Yang, “Color filtering-based efficient facedetection,” Pattern Recognition, 2008.
[6] C. N. RaviKumar and B.A, “An efficient skin illumination compensation model for efficient face detection,” IEEE Industrial Electronics IECON, pp. 3444–3449, 2006.
[7] M. Yap, H. Ugail, R. Zwiggelaar, and B. Rajoub, “Facial image processing for facial analysis,” Crown, 2010.
[8] S. Srisuk and W. Kurutach, “A new robust face detection in color images,” in Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002.
[9] E. Saber and A. M. Tekalp, “Face detection and facial feature extraction using color,shape and symmetry-based cost functions,” Proceedings of ICPR, pp. 654–658, 1996.
[10] S. Sirohey and A. Rosenfeld, “Eye detection in a face image using linear and nonlinear filters,” Pattern Recognition, pp. 1367–1391, 2001.
[11] Y. Wang and B. Yuan, “A novel approach for human face detection from color images under complex background,” Pattern Recognition, pp. 1983–1992, 2001.
[12] R. Vranceanu, R. Condorovici, C. Patrascu, F. Coleca, and L. Florea, “Robust detection and tracking of salient face features in color video frames,” Int. Symposium on Signals, Circuits & Systems, 2011.
[13] J. Tang and J. Zhang, “Eye tracking based on grey prediction,” First International Workshop on Education Technology and Computer Science, pp. 861–864, 2009.
[14] L. Ke and Y. Huang, “Eyes location based on dual-orientation gabor filters andtemplates,” 2 nd International Congress on Image & Signal Processing, 2009.
[15] C. Ming and C. Yisong, “A hardware/software co-design of a face detection algorithm based on fpga,” in Int. Conf. on Measuring Technology and Mechatronics Automation, pp. 109–112, 2010.
[16] M. Fini, M. Kashani, and M. Rahmati, “Eye detection and tracking in image with complex background,” in 3rd International Conference on Electronics Computer Technology ICECT, pp. 57–61, 2011.
[17] Amiripalli, S. S., Bobba, V., & Potharaju, S. P. (2019). A novel trimet graph optimization (TGO) topology for wireless networks. In Cognitive Informatics and Soft Computing: Proceeding of CISC 2017 (pp. 75-82). Springer Singapore.
[18] X. Xie and R. S. H. Zhuang, “Real time eye feature tracking from a video image sequence using kalman filter,” IEEE transactions on systems, man, and cybernetics, vol. 25, no. 12, pp. 1568–1577, 1995.
[19] Z. Zhu and Q. Ji, “Robust real-time eye detection and tracking under variable lighting conditions and various face orientations,” Journal of Computer Vision and Image Understanding, pp. 124–154, 2005.
[20] L. Xia, D. Yongqing, and L. Su, “Rapid human-eye detection based on an integrated method,” in International Conference on Communications and Mobile Computing, pp. 3–7, 2010.