Dynamic Object Tracking and Occlusion Detection Based on Extended and Advanced Approach

CH. Bindusri1, K. Srinivas2

Computer Science Department, VR Siddhartha College, Student, Vijayawada, India

ABSTRACT

In general Occlusion Detection is a challenging problem. Although different Detection algorithms were proposed, they all have problems in detecting occluded Objects. Some of those are Random Subspace Method (RSM), Support Vector Machine (SVM), Active Contour Model. RSM Approach is used to detect the Pedestrians. To implement these mainly three types of datasets are to be taken PobleSec, INRIA, and Daimler Multicue dataset, Additionally used linear SVM for Classification. Active Contour Model is used for finding object outline from an image. This Approach takes the advantage of the point distribution model to limit the shape. This algorithm is highly sensitive to the initialization of tracking, making it difficult to start tracking automatically. So to overcome this problem, in this project a novel technique is introduced namely Histogram Oriented Gradients Descriptor and Adabooster (HOGA) For Occlusion Handling and integrate this in a sliding window detection framework using HOG features and linear classification. The Proposed Tracking algorithm performs favorably against various methods that can be demonstrated by both qualitative and quantitative estimations challenging in a video sequences. The input for this project is live video. And the output is occlusion detection for the give video sequence.

Keywords : object tracking, object detection, object recognition, background subtraction, feature extraction, occlusion handling.

I. INTRODUCTION

In this project, a novel technique is used to track and detect the occluded objects, called HOGA (Histogram of Oriented Gradients descriptor (HOG) for object detection And ADABOOSTER (The goal of boosting is to improve the accuracy of any given learning algorithm)). Visual surveillance system is used to detect, recognize, and track certain objects in a scene. This type of system was mainly used in applications such as security for human, important building, military target detection and traffic surveillance in cities. It is essentially a video recording system that is used for post-event analysis. In the earlier stages, human beings are watching the videos in such type of systems to check for any unusual activities. These systems cannot provide sufficient security by various issues. The aim of efficient systems is to replace passive video surveillance. An efficient video surveillance system must be fast, reliable and use of robust algorithms for moving object detection, classification, tracking and activity analysis. So such system has to raise warning on the occurrence of any suspicious events. Moving object detection is the important phase for further analysis of the video. Occlusion is one of the main performance reduction problems in video surveillance systems. All automated occlusion detection system should
accurately monitor occlusion. When the detected objects in a scene come behind another object, some parts in the objects become undetected due to occlusion. Under occlusion human bodies will be overlapped and walking together in a scene. Occlusion can be of three types: self occlusion, inter-object occlusion, back ground occlusion [1]. When some parts of object is occluded, called self occlusion. This will occur frequently. When two or more objects occluded each other, inter object occlusion occurs. Background occlusion occurs if objects are hided due to back ground objects in a scene. There are verities of occlusion detection algorithms to monitor the objects from visual video. Face recognition algorithm has so much importance in video surveillance [2].

The faces might be masked either purposely using sun glasses or mask unintentionally like scarves or crowded places. Depending up on the places such as banks, the occlusion2 maybe suspicious. Because of face occlusion, the performance degradation of system will occur. So researches in the last decade have concentrated on improving the performance of the human detection system under conditions like occlusion. Occlusion is the major challenges for such systems. The systems that have addressed occlusion can be classified according to how they handle occlusion are human detection without occlusion detection, with occlusion detection and localization of occlusion. The first category does not distinguish occlusion and non occlusion. Second category can detect occlusion partially. The third category can not only detect the occlusion but also locate the occluded region. The main problem in occlusion detection is that occlusion cannot be detected directly. To detect occluded part, the pixels under goes or going to be occluded is detected. Occlusion was mainly detected for the purpose of restoring the occluded parts in an image. Occlusion detection in complex environment can be improved by using multiples view from different sensors or cameras. Accurate occlusion detection and interpretation will help the user to get the information accurately.

Object tracking is a very essential task in many applications, of computer vision. such as surveillance, vehicle navigation, autonomous robot navigation, etc., it contains detection of moving objects, and tracking of such objects from frame to frame.[1].

**Occlusion:**
Occlusions occur when, the view of a moving object is blocked completely or partially by other objects. [1].

**Partial occlusion:** An occultation is an event that occurs when one object is partially hidden by another object that passes between it and the observer.

**Short term occlusion:** with a short period of time the object may be occluded.

**Long term occlusion:** with a long period of time the object may be occluded.

**Complete occlusion:** An occultation is an event that occurs when one object is completely hidden by another object that passes between it and the observer.

**Short term complete:** with a short period of time the object may be occluded.

**Long term complete:** with a long period of time the object may be occluded.

![Figure 1. Procedure for object classification](image-url)
A continuous wide range of devices leads to a more predicted benefit such projects. Late perform has designed to discover different devices by executing the test of all peculiarities inside the details set or for every picture class and choosing the most continuous ones. This technique needs an extra identifying venture with an inalienable demand on memory and modifying time subject to the quantity of peculiarities. This option might along these lines give particular desire to strong devices lowering the total wide range of monetary dedication focuses used for picture recovery. We suggest color economical dedication focuses to purchase a scanty picture representation. Consequently, we reduce the affectability to picture circumstances, light-invariant economical dedication focuses are recommended. For color helped focuses, the point is to experience color details deduced from the event chance of colors. Color strengthened focuses are gotten through saliency-based technique choice. The career of color data allows extricating repeatable and scale-invariant economical dedication concentrates.

The aim of effective techniques is to replace passive movie monitoring. A proficient movie monitoring program must be fast, reliable and use of robust methods for shifting item identification, category, monitoring and activity research. So such program has to raise warning on the existence of any dubious events. Moving item identification is the essential phase for further research of it clip.

Occlusion is one of the primary efficiency reduction problems in movie monitoring techniques. All automated closure identification program should perfectly observe closure. When the recognized things in a field come behind another item, some areas in the things become unnoticed due to closure. Under closure people will be overlapped and walking together in a field. Occlusion can be of three types: self closure, inter-object closure, history closure [1]. When some areas of item is occluded, called self closure. This will happen frequently. When two or more things occluded each other, inter item closure happens. Background closure happens if things are hided due to history things in a field. There are verities of closure identification methods to observe the things from visual movie. Experience identification criteria have so much importance in movie monitoring [2]. The encounters might be hidden either intentionally using sun glasses or mask accidentally like neckties or populated locations. Depending up on the locations such as banks, the closure may be dubious. Because of face closure, the efficiency deterioration of program will happen. So studies in the last decade have concentrated on improving the efficiency of a persons identification program under conditions like closure. Occlusion is the major challenges for such techniques. The techniques that have addressed closure can be classified according to how they handle closure are individual identification without closure identification, with closure identification and localization of closure. The first category does not distinguish closure and non closure. Second category can identify closure partially. The third category can
not only identify the closure but also locate the occluded region. The problem in closure identification is that closure cannot be recognized directly.

To identify occluded part, the part under goes or going to be occluded is recognized. Occlusion was mainly recognized for the purpose of repairing the occluded areas in an image. Occlusion identification in complex environment can be improved by using many view from different receptors or cameras. Accurate closure identification and presentation will help the user to get the information perfectly. Object monitoring is a very essential task in many programs, of computer vision. such as monitoring, vehicle routing, independent robot routing, etc., it contains identification of shifting things, and monitoring of such things from structure to frame.

II. METHODS AND MATERIAL

Factor tracking program and kernel tracking program methods. In accordance with the purpose of an item at a certain place that is point is started at a particular part of the item. To monitor the item. These are the Strategies used for the Factor Tracker.GOA tracking program, Kalman Narrow, MHT (Multiple Theories tracking).are the methodologies. Benefits of point tracking program should be as falls, Manage closure and misdetection, Monitor points in loud pictures, and Deal with records of new item and quit existing item. Drawbacks of point tracking program should be as falls, Assume no item records and prevails, State is distributed by Gaussian, Computationally rapid both in time and memory.

point tracking program and kernel tracking program methods. An repetitive localization procedure depending on the maximization of a likeness measure.Methodology is,Color Histogram.Advantages of Kernel Tracking should be as falls, Effective to closure, Mess, diversion. Drawbacks of Kernel Tracking should be as falls, Spatial details of the objective is lost, Cannot give good performance when an item & its qualifications have similar color.

the form of the item is acquired in the first structure. Once, a difficult form of the desired structure is available on the first picture of the succession, the program instantly describes the shapes on these pictures at movie rate.[6][7] Contour base Object tracking is as movement centered identification, automated monitoring, useful in many areas such individual computer interaction, traffic monitoring, vehicle routing etc. Complexity in the issue may occur due to noise in pictures, complicated item movement, complicated item shapes, closure etc.[6][7][8].Advantages of this technique is as follows, Effective form design is used for finding item summarize from a picture, and This strategy takes the benefit of the purpose submission design to limit the form. Drawback is This criteria is highly delicate to the initialization of tracking, making it difficult to start tracking instantly.

it symbolizes the item depending on the shade. The use of areas as primitives for tracking enables to directly handle consistent object-level organizations. [7][8][9][10] A motion-based segmentation process depending on normal moves and first order movement models provide immediate dimensions. Shape, position and movement of each area present in such segmented pictures are approximated with a recursive criteria along the succession. [8-9] Occlusion situations can be managed. [11] Benefits is: It is computationally efficient. Drawbacks are Its performance is deteriorated when several things move together in the picture series, It is not possible to achieve precise tracking when several things move due to closure.[12][13]

a new technique for monitored item segmentation in movie series. In the suggested technique the user feedback item summarize is considered as movie item. In shifting item tracking, the design integrated the
object’s area segmentation and the movement evaluation. Effective form design is also employed for form fine-tuning.

Gu and Lee [15] presented movie item tracking program using in reverse area removal, movement evaluation depending on area, area category and post-processing of the area. Semantic movie item border is found using a combination of morphological segmentation tool and individual assistance. Motion evaluation, semantic movie item compensation and Iframes border details are taken to find out other movie things in the remaining frames.

An item removal plan mainly comprises of two trackers. Using Adaboost-based global shade feature selection the pixel-wise tracking program ingredients an item. To regionalize each structure K-means clustering is performed by the region-wise tracking program at the beginning. Using a bidirectional marking plan area tracking is achieved.

Discovering and tracking shifting things using automatic initialization depending on qualifications modelling. Their suggested area competition level-set technique was used for movement identification and tracking depending on the mathematical details of picture intensity within each part instead of searching geometric limitations. Before going to item segmentation and tracking qualifications modelling is done.

A region-based compound filter for general item tracking and segmentation. Their criteria brings together shade centered compound filter and area centered compound filter. The criteria monitored things in a reliable manner and also provides a precise segmentation of the objective during the succession. The compound filters uses several hypotheses for tracking things.

3D tracking design which is capable of extract item independent movement velocity under out of control environment. They have designed two novel methods, including a motion-based segmentation and a region-based Mean-shift tracking strategy. A Kalman filter is applied to merge their tracking results of the two methods.

B. Sai Chandana, Dr. K. Srinivas, [25] provided four clustering algorithms namely K-means, Moving K-means, Fuzzy K-means and Fuzzy Moving K-means combined with ECEMD for the classification of remote sensing image. The qualitative and quantitative analysis done proved that Fuzzy Moving K-means has higher classification quality than other clustering algorithms. Clustering algorithm combined with ECEMD overcomes the problem of random selection of number of clusters and initialization of centroids. The proposed method reduces the number of iterations for classifying an remote sensing image and costs less execution time.

K. Srinivas, R. KiranKumar [26] provided hyperspectral image enhancement method based on evolutionary algorithm has been proposed. The hyperspectral image data set involves many consecutive narrow bands, resulting in a continuous reflectance spectrum for each pixel. EMD is used for decomposition of an image in a specific band into IMFs. Each IMF is multiplied by a specific weight and the summation of these IMFs produces an enhanced image. The weights are determined by genetic algorithm optimizing the information entropy in the image. The experiment result shows significant enhancement in the image and thus producing more classification accuracy.
III. PROPOSED APPROACH

The procedure for color quantization is for the most part involved of two stages: plan review (the commitment of a little set of colors that speaks to the first picture shades) furthermore pixel implementing (the execute of each one information pixel to one of the plan colors). The important concentrate on is to decrease the amount of novel colors, N', in a picture to C, C ≪ N', with little mutilation. In many applications, 24-bit p in the first picture are decreased to 8 items or less. Various picture keep color quantization techniques have been designed previous times three years. These could be classified into two families: preclustering techniques and post-clustering techniques. Re-clustering techniques are usually focused around the particular dissection of large conveyance of pictures. Divisive pre-clustering exercises start with a personal group that keeps all N' picture colors. This beginning selection is recursively separated until C groups are obtained. In this papers, we look at the efficiency of difficult what's more at ease c-implies calculations within the connection of color quantization. We understand a few effective variations of both calculations, every particular with an different instatement technique, and after that look at the following quantizes on an various set of pictures.

The main objective of this project is to identify the objects which are partially occluded by the other objects in a video sequence. The scope of this project is moving object tracking under partial occlusion.

IV. OBJECT TRACKING USING HOGA

In this paper, we are trying to do the monitoring of the occluded things in videos clip series. In past research only occluded pictures can be taken place to give the results centered on their incorrect advantages etc.

For movie item monitoring study, there are mainly three methods:

A. Technique centered on design related.
B. Technique centered on category.
C. Technique centered on item state estimation.

A. Technique depending on design matching:
Is the process which converts visible monitoring into item related of subsequent movie supports [34]. Mean Move [15, 26] is the most common item design related criteria. This approach has relatively little determining quantity and can achieve fast people recognition and monitoring in fixed qualifications. However, it is difficult for people recognition and monitoring in moving qualifications, which boundaries the application range of this method.

B. Technique depending on classification:
Transforms item monitoring into forefront and qualifications category and usually assumes machine learning means for handling,[17-29]. But there are three problems of this method: first, development of classifier needs a lot of good and bad examples to learn and how to choose examples is a key issue; second, the computation has high complexness and huge computation amount; thus it is hard to fulfill real-time needs; third, it needs to do item search within the opportunity of item area. It still needs to study how to improve the opportunity to the size which is neither too little to impact item monitoring precision[31-32] nor too huge to reduce searching performance.

C. Technique depending on item state estimation:
Relies on Bayesian concept. [30]. So to get rid of this problem I am going to do this venture on it series. By using a novel technique. HOGA Algorithm. Below is the prevent plan of our technique used to identify and monitoring of the things in videos series. A detailed process to be listed in that plan for our
method which is to be taken as our suggested approach. Procedure of the HOGA shown in figure 3.

**Figure 3.** Architecture of the proposed approach

- Sample a set of picture areas.
- \( t-1 \) is the monitoring place at the \( (t-1) \) -th structure, and draw out the functions with low dimensionality.
- Use classifier to each function vector to get the monitoring place \( lt \) with the maximum classifier reaction.
- Sample two categories of picture areas.
- For feature removal objective we are using HOG.
- Extract the functions with these two categories of examples increase the classifier factors.
- Now from these 3 things will be taken to consider. These are given as feedback to the classifier parameter(ADABOOSTER).
- Finally the classifiers will categorizes the feedback which is given from the HOG and it gives the outcome i.e., partly occluded object

V. RESULTS AND DISCUSSION

Background subtraction, also known as foreground detection, is a technique in the fields of image processing and computer vision wherein an image’s foreground is extracted for further processing (object recognition etc.). Generally an image's regions of interest are objects (humans, cars, text etc.) in its foreground. After the stage of image preprocessing (which may include post processing like morphology etc.) object localization is required which may make use of this technique.

Figure 4 shows that the input of our project, we are observing that a pole is occluded by the objects which are moving on that road. This is the input for our project.

**Figure 4.** Static video input for different sequences.

And when the objects are moving through or crossing that pole, Pole has to be occluded. Now based on our project scenario the occluded object will be displayed in the second window. Here the expected output is pole. That will be displayed in figure 5. As an output the pole is to be the main object of this video.

**Figure 5.** Static video output results at different frames.

The fundamental believed behind the histogram of focused gradients descriptor is that regional item overall look and form within an picture can be described by the submission of strength gradients or advantage guidelines. The picture is split into little linked areas known as tissues, and for the \( p \) within each mobile, a histogram of slope guidelines is collected. The descriptor is the concatenation of these histograms. For enhanced precision, the regional histograms can be contrast-normalized by determining even of the strength across a bigger area of the picture, known as a prevent, and then using.
this value to stabilize all tissues within the prevent. This normalization outcomes in better invariance to changes in lighting and following their every move. The HOG descriptor has a few key benefits over other descriptors. Since it runs using regional tissues, it is invariant to geometrical and photometric changes, except for item orientation.

![Figure 6](image1.png)

**Figure 6.** Output for live video sequences in real time scenario.

Figure 6 Reveals that the outcome of the taken feedback movie. Here fingertips are the things which are concealing the boll, and the boll is the occluded item that is shown on the resulting screen. The sensitivity of the detection of objects in relevant video or audio streaming with different streams shown in Table 1.

<table>
<thead>
<tr>
<th>Video Frames</th>
<th>Traditional</th>
<th>Adabooster Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>2</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>4</td>
<td>0.7</td>
<td>0.9</td>
</tr>
<tr>
<td>5</td>
<td>0.8</td>
<td>1.24</td>
</tr>
<tr>
<td>6</td>
<td>0.9</td>
<td>1.35</td>
</tr>
</tbody>
</table>

**Table 1.** Sensitivity values for different video streams.

Graphical performance evaluation with respect to sensitivity for different video frames shown in figure 7.

**Figure 7.** Performance evaluation in terms of sensitivity for different video frames

AdaBoost is a well-known enhancing strategy which allows you merge several “weak classifiers” into only one “strong classifier”.

**VI. CONCLUSION**

The main sections in occlusion detection system are segmentation and occlusion detection procedures. Background subtraction techniques show very good results on motion segmentation with static backgrounds. Patch creation on objects can detect occlusion in a good way. Occlusion detection system gives better results while employing blob or bounding box creation around the objects in addition to patch based frame work.

To develop a robust active surveillance system, we proposed to use single camera for image capturing and back ground subtraction can be used for object detection in static backgrounds. But an algorithm has to be used for shadow removal, illumination changes and ghosts to obtain good results. Patches and bounding box can be created and it is useful to detect occlusion. Multiple cameras can also be used to take videos in a frame. Occlusion can be solved in a better way using the different viewpoints but camera co-ordination is difficult.
III. REFERENCES


17. Rashid M, Ramasamy S, PS Raghava G. A simple approach for predicting protein-protein


