A Survey On Occlusion Detection

Smitha Suresh [1], Dr. K. Chitra[2] and P. Deepak [3]

[1]Associate professor (Non- Cadre), Dept.of CSE,
SNG College of Engineering, Kadayiruppu,kerala,India

[2]Asst. Prof.   , Dept. of Computer Science,Govt.Arts College,Melur,India
[3]Associate professor (Non-Cadre), Dept.of ECE,
SNG College of Engineering, Kadayiruppu,Kerala,India

Abstract

When objects are occluded, some parts may not be visible to human vision. Object detection is a difficult job in real time tracking of multiple objects due to occlusion. This paper presents a survey on detection of occlusion. Similarly occlusion at faces may be leads to performance degradation of face recognition algorithms. Many methods were developed to detect occlusion in present and previous works. The aim of developing such intelligent system is to detect the object and identify them actively under various performance reducing parameters. The performances of the best techniques are reviewed in this paper.

Keywords

Occlusion, Human vision, Occlusion detection, Face recognition, intelligent system.

I. Introduction

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 occlusion may be 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.

Several approaches to detect occlusion were discussed in this work. The remaining of this paper is as follows. Section II gives a small introduction of the earlier and present works and gives a general frame work of video surveillance to detect occlusion using single and multiple camera. Section III is a review of the existing methods. Section IV is a comparison table for different methods. Section V concludes this paper and Section VI is the possible directions for the future works.

Ii.Present and earlier strategies

In most of the previous works, image was captured using single camera and foreground objects will be segmented. Fig 1 shows a frame work for occlusion detection using single camera.
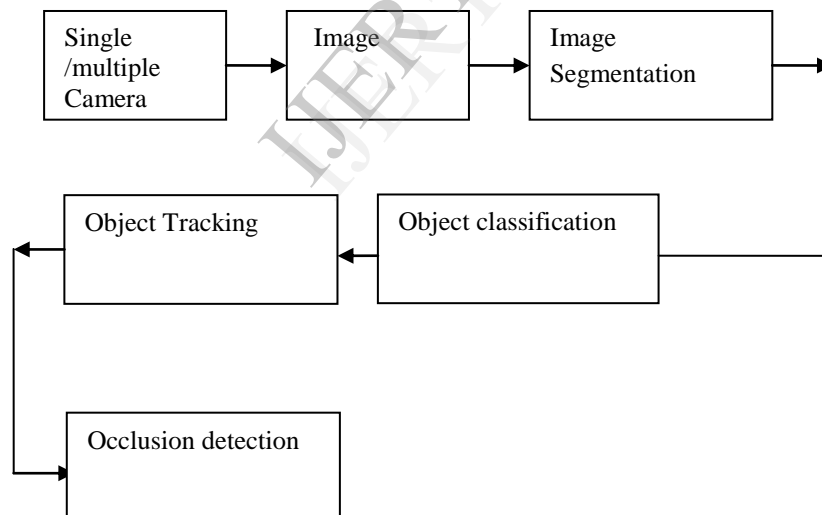

Figure 1. General Frame work for occlusion detection using single/multiple camera

In recent years, multiple cameras are used to segment and tracking the objects in a crowded scene and deals with occlusion in a useful way. Information from multiple cameras can reduce the uncertainty of using single camera [3]. Multiple cameras can solve the occlusion but some algorithms are needed to determine individual objects properly in complex scene. The one disadvantage of using multiple cameras is the co-alignment of camera Co- ordination is very difficult.

Segmentation is nothing but detecting image regions and it provides a focus of attention for later processes such as tracking and behavior analysis. Background subtraction is the popular method for segmenting the image from a motion frame. Background subtraction is

suitable for the situations with static background. Temporal differencing method of segmentation is suitable for dynamic backgrounds. Background subtraction techniques show very good results on motion segmentation with static backgrounds. Adaptive background subtraction gives very good and fast moving object segmentation. [3]

The image sequences captured by surveillance cameras was normally mounted in road traffic scenes include humans, vehicles and other moving objects such as flying birds and moving clouds, etc. To further track objects and analyze their behaviors, it is essential to correctly classify moving objects. Object classification can be considered as pattern recognition. There are two object classification techniques. They are Shape-based classification and motion based classification. In shape based classification, different descriptions of shape information of motion regions such as points, boxes, silhouettes and blobs are available for classifying moving objects. In general, non rigid articulated human motion shows a periodic property, so this has been used as a strong cue for classification of moving objects [3].

After motion detection, surveillance systems generally track moving objects from one frame to another in an image sequence. Tracking methods are divided into four major categories; region-based tracking, active-contour-based tracking, feature based tracking, and model-based tracking.

Previous methods [4], W. Hu, M. Hu, X. Zhou, T. Tan, J. Lou, and S. Maybank, of segmentation were affected by two problems. First, occlusion are common in places such as railway station, airports etc., where people are frequently walked together in most of the time. Second, people dressed in similar colors makes some problem when color models are used for segmentation. Segmentation under occlusion was neglected in the previous works or reduced by placing the camera at very high place to capture images from motion plane.

Tracking is very important in surveillance. Tracking of object is very difficult when objects are subjected to occlusion

Iii. Existing methods for occlusion detection

In the work [2], Zhaohua Chen, Tingrong Xu, and Zhiyuan Han have solved the problem of face recognition under occlusion due to sun glasses or scarves. Here the presence of sunglasses or scarves was detected and the non- occluded region only was processed. Occlusion can be dealt by selecting non occluded patches from the faces. Occlusion was detected by PCA and support vector machines (SVM). To detect the occluded region in the face, divide the images in to finite number of patches and examine each patch separately. As configuration and size of the patches are important in the performance of occlusion detection they have divided faces in to 6 symmetrical patches. Then dimension of these patches were reduced by using PCA.

In the work [5], Yi Deng and Qiong Yang proposed a patch based frame work to detect occlusion. Segmentation was applied to both images and warps the segment of one image to the other by disparity. The warped segment is then divided in to small patches on the boundaries of other images. The occlusion at boundary can be treated as the occlusion at patches. A frame work using graph cuts was used to find the disparity and occlusion. A symmetric frame work using graph cut is constructed to find the disparity and occlusion at patches.

This work [6] Xinting Pan*, Xiaobo Chen and Aidong Men proposed a particle filter for tracking the object accurately. The object pixels were first classified as foreground and background in all frames using background subtraction. The object in the scene is extracted from the background by background subtraction, and then the object is considered as the region of interest (ROI). An elliptical model was used to represent each object with parameters like center of the ellipse, length of the major axis of ellipse eccentricity of the ellipse. Occlusion can be detected by the merging and splitting of the ellipse.

In the work [7], Sherin M. Youssef, Meer A. Hamza and Arige F. Fayed, was proposed an algorithm to detect and track multiple objects under occlusion. They have detected and track the objects using discrete wavelet transform (DWT) and identifying the objects by their colour. A bounding box was created around each object and it was labeled. The object was scanned from right to left, left to right, top to bottom and bottom to top to create bounding box. The top left coordinates, height and width of the each bounding box was created after scanning. Occlusion was detected by analyzing the height and width of the bounding box.

In this work [8], the Tao Yang, Stan Z.Li, Quan Pan and Jing Li, proposed a real time system for multiple human tracking in dynamic scenes. This work can deal with complete and long duration occlusion. Here they have considered object sates in to three categories; before, during and after occlusion. The system consists of object segmentation part, merging and splitting detection module for occlusion detection. Occlusion was identified only after the merging and splitting of blobs in each object occurs. During occlusion the trajectories of each object is similar to the entire group. The authors proposed an algorithm for background maintenance to handle scene changes like ghosts and illumination changes.

In paper [9], Yao-Te Tsai, Huang-Chia Shih, and Chung-Lin Huang introduced a system to detect and track multiple objects with occlusion. Each object is represented by colour models of the two regions of the body. Guass models were used to model the colour of top and bottom region. They have also analyzed the optical flow of object to predict the position in the next frame. This system consists of segmentation, noise removal, optical flow based position estimation, occlusion detection and object separation from occlusion. They have created an object window for each object. The neighborhood pixels of an object are identified using Bayesian classification. The central vertical axis was calculated for object window. The position of the central vertical axis in the next frame was predicted using optical flow. Occlusion was detected using the distance between the central vertical axis of every two object.

Valtteri Takala and Matti Pietikainen [10] introduced a novel real-time tracker algorithm based on color, texture and motion information. RGB color histogram and correlogram are used to describe the object's color properties. The merging and splitting of objects are handled using the same set of features. The main sources of descriptors are color, texture, shape, and temporal (motion) properties. But the color has gained the most of attention as it is well distinguishable to human eye and seems to contain a good amount of useful information. Tracker used in this paper consists of two main elements background subtraction (detection) and tracking. The subtraction on the video data, which is first processed with a Gaussian filter to remove noise, is done by an adaptive algorithm. The subtracted foreground is enhanced by filtering the artifacts caused by noise and moving background using standard morphological operations. The tracking is done by matching features extracted from the subtracted foreground shapes. These are color, texture and motion. Instead of using the bounding boxes themselves for occlusion detection, this system surrounds the boxes with circles that have the radius of the half diagonals of the boxes and use them for event detection. If the object circles are occluding each other in the previous frame n-1 a merging event in frame n is possible. If the occlusion is true in frame n + 1 and the closest occluding object is a group object 1(1&2) then a split might have happened.

T.H. Tsai, Y.C. Liu, T.M. Chen and C.Y. Lin [11] used the Gaussian Mixture Model was used to get video of interesting (VOI) region in every new frame. To simplify the computational complexity of the mixture models, grayscale images were used instead of the three independent color channels. After converting the original color images to gray scale, the algorithm determines the background and foreground blobs and outputs a binary representation of the foreground. Then filter all of the noises (too small) below the threshold and labeling foreground with connected pixels. After labeling the foreground objects, every individual foreground blob without connected pixels has a unique number in the image. In this system, the occlusion was detected by the distance between each object and size of the blob, when objects are going to be occluded.

Pierre F. Gabriel, Jacques G.Verly, Justus.piater and Andre Genon [12] proposed the concept of blob. A blob acts as a container that can have one or more objects. Objects and blobs are described by a series of attributes such as position, velocity and appearance. Blobs are also characterized by operations, such as create, delete, merge, and split. The system uses a predicate Po that detects the fact that two or more blobs are occluding each other. Once an occlusion is detected, there appear to be two approaches for dealing with it. The first approach is the merge-split (MS) approach. As for the MS approach, one can use a Kalman filter to estimate the positions and velocities of the moving objects. As soon as blobs are declared to be occluding by Po, the system merges them into a single new blob. From that point on, the original objects are encapsulated into the new blob. The new blob is characterized by new attributes and is tracked as any of the active blobs in the system. In this system, assume that there is a predicate Ps that is able to decide whether a blob containing at least two objects must be split. In the MS approach, attributes of atomic blobs i.e. (containing a single object) are continuously updated until they come into an occlusion situation. When a split condition occurs, the problem is to identify the object that is splitting from the group. The systems uses only appearance features such as color, shape and texture to re-establish identity. The second approach is the straight through (ST) approach. Here, simply continue to track the individual blobs (containing only one object) through the occlusion without attempting to merge them. The system uses same occlusion predicate Po used in the MS approach. The main difficulty in MS approaches is to reestablish object identities following a split. In regions-based ST approaches, the main difficulty is the assignment to a specific object of pixels that could belong to several objects ("disputed" pixels). In contour-based ST approaches, the main difficulty is the assignment to a specific object of some partial contours. Occlusion was detected by the checking merge and splits of the blobs.

Tao Zhao and Ram Nevatia [13], proposed a method   to track multiple people in complex situations using a single stationary video camera after segmentation. Moving pixels are grouped into blobs according to their connectivity. Here they proposed to use a coarse 3D human shape model (an ellipsoid). First, the foreground blobs are extracted by a change detection method. Human hypotheses are computed by boundary analysis and shape analysis using the knowledge provided by the human shape model and the camera model. Each hypothesis is tracked in 3D in the subsequent frames with a Kalman filter using the object's appearance constrained by its shape. The binary map is filtered with a median filter and the morphology close operator to remove isolated noise, resulting in the foreground mask F. Connected components are then computed, resulting in the moving blobs. To track the human, a mapping is needed to align different ellipses for matching and updating. At each frame, the position is predicted by the Kalman filter. The system computes r, the visible fraction of the object. r is defined by N_v/N_c, where N_v is the number of visible (i.e., un occluded) foreground pixels in the elliptic mask and N_e is area, in pixel, of the elliptic mask of each object. Using two thresholds T_{o1} and T_{o2}, if $T_{o1} > r > T_{o2}$, the object is said to be partially occluded. If $r < T_{o2}$, the object is said to be completely occluded. One deficiency of this method is that the human orientation when standing is inferred poorly from its speed, which is close to zero.

Iv. Comparison of different methods

Author (first author)	Method	Type	Full/partial/self occlusion
Anandhakumar.P	Filter bank	Facial occlusion	full
Zhaohua Chen	SVM and Block weighted	Facial occlusion	partial
Weiming Hu	Principal axis based method	Human occlusion	Partial/ full
Yi Deng	Patch based stereo algorithm	Object occlusion	Partial/full
Xinting Pan	Partial filter	Human occlusion	full
Sherin M. Youssef	Object labeling & bounding box	Human occlusion	Partial/full
Tao Yang	Merging and splitting approach	Human occlusion	Partial/full
Yao-Te Tsai	Object window	Human occlusion	Partial/full
Valtteri Takala	Color,texture and motion	Human occlusion	Partial/full
T.H. Tsai	Object distance & power	Human occlusion	Partial/full
Pierre F. Gabriel	Merge split and straight through approaches	Human occlusion	Partial/full
Tao Zhao	Human shape model & Kalman filter	Human occlusion	Partial/full

"Table 1: Comparison Methods For Occlusion Detection"

V.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.

Vi. Directions for future works

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 view points but camera co-ordination is difficult.

References

[1]   Anandhakumar.P and Priyadharshini J, "Occlusion detection and object tracking using filter banks", International conference on recent trends in Information technology, IEEE 2011.

[2]   Zhaohua Chen, Tingrong Xu, and Zhiyuan Han, "Occluded Face Recognition Based on the Improved SVM and Block Weighted LBP", IEEE 2011.

[3]   Weiming Hu, Tieniu Tan, *Fellow, IEEE*, Liang Wang, and Steve MaybankA, "Survey on Visual   Surveillance of Object Motion and Behaviors", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. 34, NO. 3, AUGUST 2004.

[4]   W. Hu, M. Hu, X. Zhou, T. Tan, J. Lou, and S. Maybank, "Principal axis-based correspondence between multiple cameras for people tracking", Pattern Analysis and Machine Intelligence,IEEE Transactions 2006.

[5]   Yi Deng , Qiong Yang, Xueyin Lin and Xiaoou Tang, "A Symmetric Patch-Based Correspondence Model for Occlusion Handling", Proceedings of the Tenth IEEE International Conference on Computer Vision, 2005.

[6]   Xinting Pan*, Xiaobo Chen, Aidong Men,"Occlusion Handling Based on Particle Filter in Surveillance System", International Conference on Computer Modeling and Simulation, IEEE 2010.

[7]   Sherin M. Youssef, Meer A. Hamza and Arige F. Fayed, "Detection and Tracking of Multiple Moving Objects with Occlusion in Smart Video Surveillance Systems", IEEE 2010.

[8]   Tao Yang, Stan Z.Li, Quan Pan and Jing Li, "Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes", Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2005.

[9]   Yao-Te Tsai, Huang-Chia Shih, and Chung-Lin Huang, "Multiple Human Objects Tracking in Crowded Scenes", International Conference on Pattern Recognition, IEEE 2006.

[10]   Valtteri Takala and Matti Pietikainen,"Multi-Object Tracking Using Color, Texture and Motion" IEEE, 2007

[11]   T.H. Tsai, Y.C. Liu, T.M. Chen and C.Y. Lin,"Fast Occluded Object Tracking Technique with Distance Evaluation".

[12]   Pierre F. Gabriel, Jacques G.Verly, Justus.piater, Andre Genon , "The State of the Art in Multiple Objects Tracking Under Occlusion in Video   Sequences".

[13]   Tao Zhao and Ram Nevatia, "Tracking Multiple Humans in Complex Situations", Pattern Analysis and Machine Intelligence, IEEE Transactions 2004.