Occlusion Detection with Background Elimination and Moving Object Tracking

K.Padmapriya, P.V.Gopirajan, K.Suresh Kumar

Abstract: ***An occlusion is said to be a situation when few parts of an object cannot be viewed by human vision. Detecting an object is quite difficult to identify and not so easy to track the moving objects during occlusion. It can be achieved in two steps. The first step is to identify the foreground or background image frame by frame, by naming each and every pixel in the frame. The next step is to compare the observations at each point in the sequence of the moving object with occlusion. It can be done by subtracting the background which yields pixel as combination of Gaussians. Then the distribution of Gaussian have to be calculated to conclude the result from the background process. This method is very much useful to identify the motions in surveillance camera, repeated suspicious movement and elongated changes in the scene.***

Keywords: Occlusion, Gaussian, Human vision, Image processing

I. INTRODUCTION

Occlusion means that there is something we like to see, but can't due to some disturbance of our sensor setup, or some event. There are two types of occlusion: 1.Self occlusion - a part of an object is being hidden by another part of the same object shown in Figure 1(a). 2. Inter person occlusion – a part of an object is being hidden by another different object shown in Figure 1(b).

Object tracking is the process of segmenting an object in a video, identifying its motion, orientation and occlusion to get our desired information.

Our proposed technique is used for tracking the movement of different human objects depending upon motion detection, navigation, analyzing and estimating by subtracting the background, eliminating the shade and detecting the occlusion. Initially Reference model is framed as background process.

Manuscript published on November 30, 2019.
* Correspondence Author
Dr.K.Padmapriya*, Assistant Professor(SG), Department of Computer Science and Engineering, Saveetha School of Engineering, Chennai, India.
Email: dr.k.padmapriya@gmail.com
Mr.Gopirajan.P.V, Assistant Professor(SG), Department of Computer Science and Engineering, Saveetha Engineering College, Chennai, India.
Email: gopirajan@saveetha.ac.in
Dr.K.Suresh Kumar*, Associate Professor, Department of Information Technology, Saveetha Engineering College, Chennai, India.
Email: sureshkumar@saveetha.ac.in

Fig. 1(a). Example of Self occlusion Fig. 1(b) Inter person occlusion

When a new object is found in the video, the foreground information is retrieved by deleting the reference model which is the constant background. It is the process of distinction between foreground and background images.This type of object detection is applied widely in retrieval of images, computer vision applications, traffic monitoring, automatic camera surveillance and face recognition

The ability of finding moving objects in a sequence of video is a basic and critical problem of many vision systems which include surveillance scenarios as defined by I. Haritaoglu, D. Harwood, L.S.Davis and P.L. Rosin. [1,2], N. Friedman and S. Russell determines traffic monitoring [3], C.R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland discussed about human detection and tracking for video teleconferencing or human-machine interface [4], J. Ohya and J.Davis and A.. Bobick [5, 6] proposed video editing, among other applications.

In general, applications that uses fixed cameras using constant background, a general methodology is to use background elimination to attain a preliminary estimation of moving things. Mostly, background elimination containing comparison of each frame with a demonstration of the scene background, significant distinctions with respect to foreground images. Preferably, background elimination should find out the genuine moving things with higher accurateness, restraining fake negatives; at the same time, it must retrieve pixels of moving things with the maximum possibilities, avoiding uncovering of fake things like unwanted noise and shades.

II. PROPOSED METHOD

Object Tracking

Basic steps involved in object tracking are:

A. Partitioning
B. Environment removal
C. Camera modelling
D. Tracking the extracted feature

A. Partitioning

Partitioning is the process of recognizing components of an image. Partitioning includes several processes like margin discovery, associated element tagging, applying the threshold etc. Margin discovery discovers edges of an image. To detect the margin any kind of differential operator could be used [7, 8]. Thresholding is the method of decreasing the grey levels in an image. Already so many algorithms are being used for thresholding [7, 8]. Refer [8] for associated element tagging algorithms. We can make use of that algorithm in our methodology.

B. Environment removal

This is the process of splitting up the foreground and background of the image. Here it has been assumed that foreground includes the objects of interest. Once foreground is removed, the remaining part is the background of that object [1]. An alternate method can also be used is Background learning. But it is useful only static cameras are applied for video capturing. In this approach, a preliminary training method is applied before organizing the scheme. In this training method, the background is continuously recorded to find it. Once the background has been learnt, the system has completed (or almost complete) information about the background. Even though it is a lengthy step, it has more advantages when compared with other algorithms.

C. Camera modelling

Camera model is an important feature of any object-tracking algorithm. Already existing object tracking systems make use of pre-set camera model which is derived from the field knowledge. It has been achieved in [10]. For a dynamic camera, knowledge about camera motion is required. If we have accurate data regarding camera motion, then it will be incorporated in the form of conversion. Since we have many dynamic cameras, it leads to very complicated condition. So we require a method to deploy the motion given by all the cameras and also to integrate the final result.

D. Tracking the extracted feature

Here we apply some algorithms to find and separate different portions of a digitized image. A feature is said to be an important part of data pulled out from an image. It is useful to understand the image clearly. This methodology includes the description of resources needed to identify the data exactly. But surely it yields large set of complex data only, since we are using large number of variables. Then form the grouping of different variables in order to find the desired data accurately.

III. WORKING PRINCIPLE

The below given flowchart in fig. 1 explains the working principle of this paper. It has been assumed that the input image is attained from some background elimination. Each process involved in our proposed methodology has been given in the following sections.

Suppose each resultant pixel from a surface with fixed lighting, a simple Gaussian is enough for modelling it when we try to get the acquisition noise. If the lighting has been changed, adaptive Gaussian for each pixel is required. In real time, lighting will be changed on different surfaces while viewing a particular pixel. Hence we need many adaptive Gaussians. We use the combination of adaptive Gaussian to achieve our goal. Each and every time, the variables and parameters are getting updated and the new Gaussians are retrieved. Sometimes few pixels will not be matched with Gaussian values which are combined using related components. As a result, the related components are tracked over the frames of video.

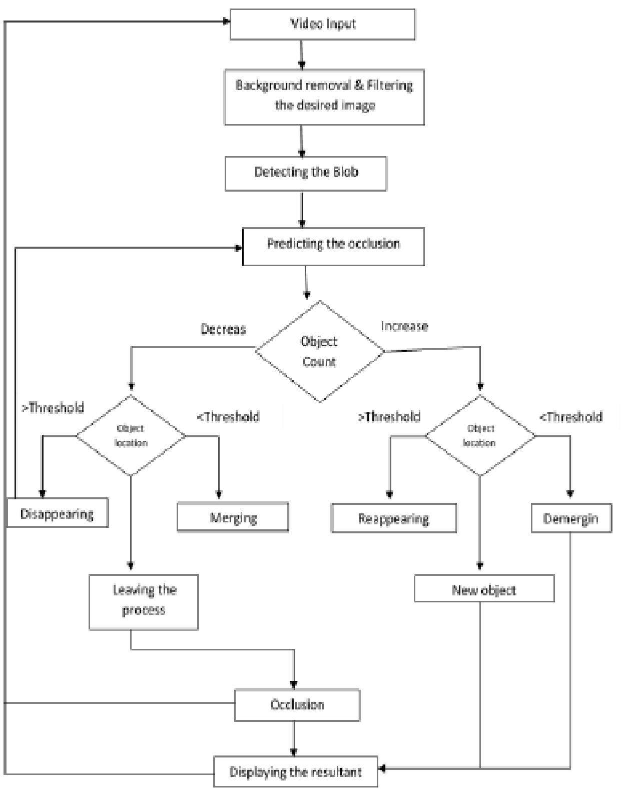

Fig. 2. Flow Chart

Assume a “Pixel Process” is said to be the values of particular pixel overtime. Let t be the time and I be the image, then pixel process can be defined as

$$\{X1, \ldots, Xt\} = \{I(X0, Y0, i; 1 \leq i \geq t\}$$

Hence the probability of particular pixel will be

$$P(Xt) = \sum \sum_{i-1}^{n} wi,t * 1/(2\pi \left| \sum \right|)e^{1/2(Xt-\mu t)\sum(Xt-\mu t)}$$

Where n is the total number of distributions, Wi,t is the weight of the Gaussian of an image at time t, X_t is the pixel at the particular time, μ_t is the mean value and $\sum(X_t-\mu_t)$is the covariance value.

In this way, distribution have been studied for each pixel and categorized by the combination of Gaussians. Assume the pixel process is a static process, the general method of the observed data is likely to be maximized. Suppose any changes in the lighting, inserting or removing any static objects will lead to decrease dependence on what we have learnt so far. To overcome these two problems, we can make use of online k-means approximation algorithm [16] to update the result.

A pixel value is defined when we found the exact match of each pixel with the existing K Gaussian distributions. If a match is found for a pixel with the values of 2.5 as standard deviation for the distribution of 3, then the threshold value may give some trouble on its performance.

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

If there is no match found, the probable distribution value is replaced by the current value as its mean value, initial variance and low weight. The major advantage of this method is that while something is entered as a background of the desired object, it will not affect the existing background model.

Environmental study

Suppose the parameters are getting updated for each pixel, then it is little bit difficult to choose the appropriate Gaussian model. In this situation we used to select the model which has lowest variance with highest supporting evidence. This idea can be applied when the object is static. When background is occluded by a new object, it will not match with any existing distributions. Hence the new distribution value is created. And also the variance value becomes high when the object is getting moved. To model this, we require a method to decide which portion should be represented as background process.

Initially the Gaussian values are sorted by the value of weight. This value will be increased when the variance value decreases and distribution gets extra evidences. After revaluating the parameters, it is enough to order the distribution which gets matched with the distribution of possible background. This model successfully shows an open-ended ordered list where distributions of background have been kept on top and the possible transient background distributions are at the bottom.

Hence the initial distribution of background will be

$$D = \text{Argmin}_d(\textstyle\sum_{k=1}^{d} wk>P)$$

Here P denotes the data portion which should be considered for by the background. It accounts the optimal distributions of particular portion P with recent data. If P is very small, the background model will be unimodal. In this case, using most possible distribution is the only way to save the processing. If P is high, then it is said to be multi-modal where the distribution is caused by the repetitive motion of the background which produce many colors. In this case it should be added as background example, movement of leaves in a tree. Hence we got more than one color as background.

Related components

With the help of above said method, we can easily find out the foreground image in each frame when the distribution of each pixel gets updated. The flow of data is shown in figure 5. These foreground pixels are clustered into regions by the algorithm given for connected components in [17]. This algorithm is very productive while determining the entire dynamic objects, moving regions and shape details. This process is useful during tracking.

Fig 3.Example of changing occlusion regions

Occlusion detection

Using multiple hypotheses tracking algorithm, we can establish the connectivity of related components among the frames with the help of size and its position. We can make use of Kalman Filters to implement our model shown in figure 4.

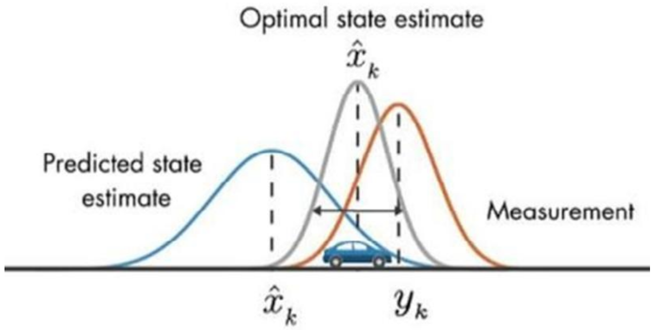

Fig 4. Kalman Model

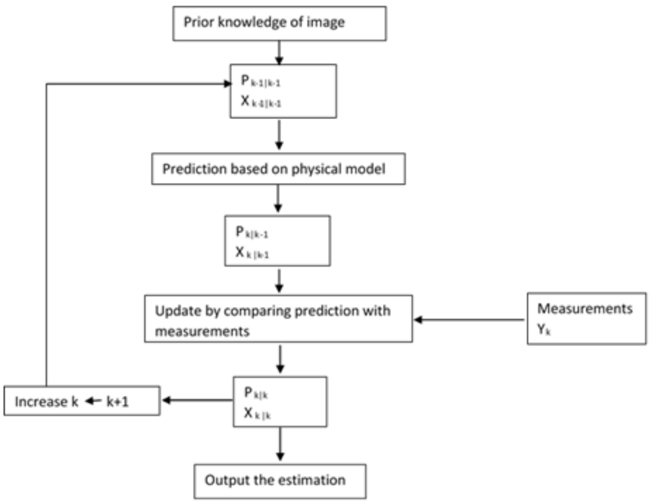

Fig. 5. Flow of Data

In each frame, we have a set of Kalman models and a new set of related components. Initially the models are being matched with the related regions. Then check for the not connected regions with the Kalman models. At last, remove the model whose matching is below the threshold value.

By checking each model with the set of related regions consists of more than one pixel yield the matching model. Matches with small error can also be considered to update the respective model. Suppose the updated model has adequate matches, then it can be considered for the prediction. If no model is having null matches, then it will be hypothesized. If an image reappears in predictable area, the model will get back the image again. The unmatched image from the present frame and the previous frames are used to hypothesize the new models. Suppose the present frame has adequate matches with previous frames, updation is done in the existing model.

IV. EXPERIMENTAL RESULTS

The image shows the occluded behind a baby. Successively Background object (Sitting baby), images in the frame, marked image, occluded object is shown and marked below in figure 6. In figure 7 Scatter plot for debugging pixel are shown and in figure 8 The pixel statistics for debugging pixel is shown.

Fig 6. Experimental results shown occluded object

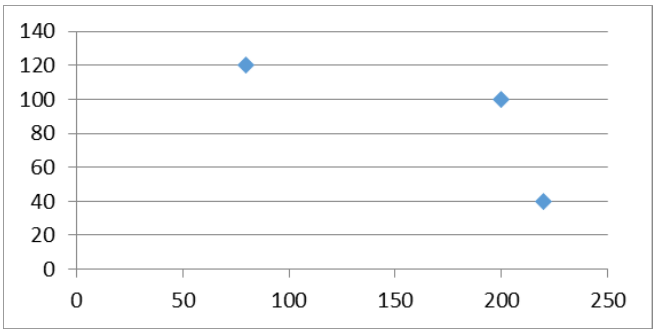

Fig 7. Scatter plot for debugging pixel shown in the result

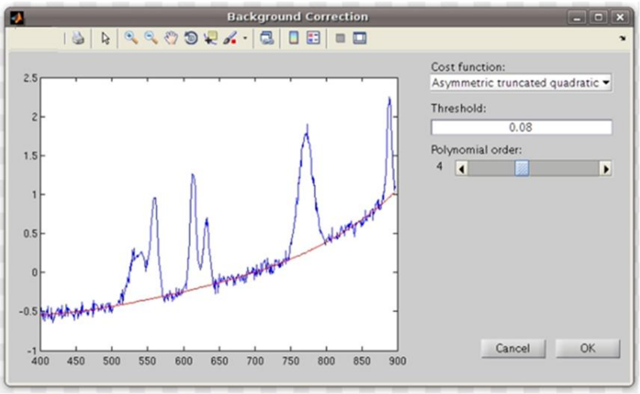

FIG 8. THE PIXEL STATISTICS FOR DEBUGGING PIXEL IS SHOWN IN THE RESULT

V. CONCLUSION

In this paper, we have discussed the elimination of background. It considers modelling pixel as a mixture model. It is implemented with the help of a static image. This method also can be dealt with moving of light slowly by adapting the Gaussians value. It also discussed the distribution of multimodal using Kalman features to overcome the problems given by shadows, tree leaves and other problematic features of the real world. It updates easily and quickly when the background gets reappear. Hence our method proved that it is necessary to detect occlusion detection most easily. As we have experimented with our methodology using different images, different vision, different cameras, and different lightings, we didn't get any problems in our method. Moreover, it shows good performance with a short period, that too without any human interaction.

REFERENCES

1. I. Haritaoglu, D. Harwood, and L.S.Davis. W4: Who? when? where? what? a real-time system for detecting and tracking people. In Proc. the thrid IEEE Int'l Conf. Automatic Face and Gesture Recognition (Nara, Japan), pages 222–227. IEEE Computer Society Press, Los Alamitos, Calif., 1998.
2. P.L. Rosin. Thresholding for change detection. In Proc. IEEE Int'l Conf. on Computer Vision, 1998.
3. N. Friedman and S. Russell. Image segmentation in video sequences: A probabilistic approach. In Proc. 13th Conf. Uncetainty in Artificial Intelligence. Morgan Kaufmann, 1997.
4. C.R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):780–785, July 1998.
5. J. Ohya and et al. Virtual metamorphosis. IEEE Multimedia, 6(2):29–39, 1999.
6. J.Davis and A.. Bobick. The representation and recognition of action using temporal templates. In Proc. The Computer Vision and Pattern Recognition, 1997.
7. R. Gonzalez and R. Woods – Digital image processing – second edition – Prentice Hall.
8. L. G. Shapiro and R. M. Haralick –Computer and robot vision – volume 1 and 2 – 1 st edition – Prentice Hall.
9. C. Ridder, O. Munkelt, and H. Kirchner Adaptive Background Estimation and Foreground Detection using Kalman-Filtering - Proceedings of International Conference on recent Advances in Mechatronics (ICRAM), pp. 193-199, 1995.
10. A. Turolla, L. Marchesotti and C.S.Regazzoni – Multicamera object tracking in video surveillance applications.
11. Mohamed Ibrahim M, Anupama R,"Scene Adaptive Shadow Detection Algorithm." Transactions on Engineering, Computing and Technology, V2, December, pp 88-91, 2004.
12. Alessandro Bevilacqua, "Effective Shadow Detection i in Traffic Monitoring Applications." Journal of WSCG, VI1, pp 57-64, 2003.
13. A. Bevilacqua, M. Roffilli, "Robust denoising and moving shadows detection in traffic scenes." Proc. Of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai Marriott, Hawaii, pp 1-4, 2001.
14. Falah E.Alsaqre,Yuan Baozong,Moving Shadows Detection In Video Sequences." Proc. of 7th ICSP 04, V2, pp 1306-1309, 2004.
15. R. Cucchiara, C. Grana, M. Piccardi, A. Prati: " Detecting moving objects,ghosts, and shadows in video streams" IEEE Transactions on Pattern Analysis and Machine Intelligence(PAMI), Vol. 25(10), 1337 - 1342, 2003.
16. Cran H.D. and steele, C.M.(1968), "Translationtolerant Mask Matching using Noncoherent Optics" Pattern Recognition, Vol. 1, No. 2, pp. 129-136.
17. Comaniciu, D., Ramesh, V. and Meer, P. (2000) Realtime Tracking of Non-rigid Objects using Mean shift", IEEE Conference on Computer Vision and Pattern Recognition (CVPR'00), Vol.2, pp.142-149.
18. Heisele, B. (2000) "Motion based Object Detection and Tracking in Color Image Sequence" 4th Asian Conference on Computer Vision.

AUTHORS PROFILE

Dr. K. Padmapriya has more than 12 years of experience in teaching. She is currently working as Assistant Professor in Saveetha School of Engineering, Tamil Nadu, India. She completed her Master of Computer Science and Engineering in Saveetha

Engineering College and completed her Ph.D. in Sathyabama University. She has published more than 20 Articles. Research Areas includes Data mining, Database technology, Watermarking, Wireless sensor network.

Mr. P.V. Gopirajan has more than 12 years of experience in teaching. He is working as Assistant Professor in Saveetha Engineering College, Tamil Nadu, India, He completed his Master of Computer Science and Engineering in Saveetha Engineering College and Currently pursuing Ph.D. in SSN Engineering College, Anna University. He has presented various papers in National and International Conferences. His research area includes Image Processing, Cloud Computing.

Dr K. Suresh Kumar has more than 16 years of teaching and 2 years of industry experience. He is working as Associate Professor in the Department of Information Technology at Saveetha Engineering College, TamilNadu, India. He did his Postgraduate M. Tech in Information Technology at Satyabama University, Chennai. He did his Ph.D. from Anna University, Chennai. He has published more than 20 research papers in various International and National conferences and journals. He has authored 3 books. His areas of interest include Computer Security, Mobile Ad hoc Networks, Mobile Computing and Cloud Computing