INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD
ISSN(O): 2455-0620 [Impact Factor: 7.581]
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value : 86.87
Volume - 9, Issue - 4, April - 2023 Publication Date: 20/04/2023

DOIs:10.2015/IJIRMF/202304012 --:-- Research Paper / Article / Review

Study of Silhouette Based Method for Human Object Classification

[1]Dr.Smitha Suresh, [2]Deepak P.

[1]Professor & Head, CSE Department, Sree Narayana Gurukulam College of Engineering, Ernakulum, India
[2]Associate Professor, ECE Department, Sree Narayana Gurukulam College of Engineering, Ernakulum, India
Email - [1]smithadeepak2006.sngce@gmail.com, [2]deepakpeeth@gmail.com

Abstract: *Object classification is a crucial step that allows for better tracking and more accurate description of évents. Accurately identifying humans in a video surveillance system is essential for a variety of application areas, such as public safety and abnormal event detection, human gait characterisation, traffic analysis, person identification, gender classification, and fall detection for senior individuals. Moving object recognition is a critical and vital phase in a visual surveillance system that separates the moving foreground object from the background. Using any object segmentation technique, find the moving foreground objects as the initial stage in the detection procedure. Humans, human groups, animals, vehicles, flying birds, moving clouds, tree leaves, and other noises could all be present in the retrieved moving foreground item. To categorise the moving foreground object into specified categories, such as a single person, a group of people, or vehicles, etc., is the goal of the classification step in the visual surveillance system. The necessity of object classification in video surveillance system is to differentiate objects from each other to track their actions reliably. To categorise human objects following moving object segmentation, we adopted the more effective shape-based object classification approach called Silhouette Template based object classification algorithm.*

Key Words: *Video surveillance, moving object segmentation, object classification, silhouette template*.

1. INTRODUCTION:

Object classification is the process by which an object in motion needs to be classified in the image frame as what object of interest is needed. The classification of moving objects is crucial in many computer vision applications such as event understanding (Chapman et al., 2019; Xu et al., 2019), human action recognition (Zhang and Tao, 2019; Avola et al., 2019; Liu et al., 2019), pedestrians and vehicles tracking (Zhao et al., 2019; Cui et al., 2019), a smart video surveillance systems for adaptive foreground extraction (Muchtar et al., 2019) and video surveillance systems against anti-terrorism (Karthikeswaran et al., 2019). The classification rules should be used to accurately classify the relevant objects in the environment with computationally inexpensive and sophisticated manner. The two important object classification methods are: shape based classification and motion based classification. The objects' 2D spatial information is used in shape based classification to categorize the objects whereas the temporary tracked features of objects are used in motion based classification. The shape based classification is normally used for surveillance application such as action recognition. The self-similarity based techniques are used in motion based classification to analyze the periodic tracking of interesting objects. The classification rules should be used to accurately classify the relevant objects in the environment with computationally inexpensive and sophisticated manner. The moving region extracted after background subtraction may have different kind of targets like pedestrians, vehicles, clutter and other moving objects such as birds etc. If the required targets are a person or a group of persons it can be recognized by the most suitable classification algorithm. A classifier is an algorithm that takes a set of factors or features that characterize objects and can be used to determine the type (or class) of each object. Each object have number of its own properties such as brightness, size etc. The classifier then uses these properties to correctly classify the objects. Outdoor environments are more challenging due to the uncontrollable condition, incomplete appearance details of moving objects in object tracking system. The most commonly used approaches to classify the objects are shape based classification, Motion based classification, colour based classification and texture based classification. During object classification the occurrence of shadows will change the size and shape of the moving objects because the shadow will move along with the moving objects and it will result in incorrect output due to false classification and it will degrade the performance of video surveillance system.

This paper is organized as follows: first in Section 2, we review the related works. The methodologies for Silhouette shape based object classification process is described in detail in section 3. Then in Section 4, we present the obtained results finally, we give the conclusion

2. LITERATURE REVIEW :

According to our knowledge, the three primary categories of the state of the art for improving the categorization accuracy of moving objects in surveillance films are motion-based approaches, colour-based techniques, and shape-based approaches. The following is how we display these recent studies.

Motion based classification schemes are used to distinguish non rigid objects such as human from rigid objects like vehicles. Rigidity and periodicity of moving objects can be analyzed by residual flow (Weiming, 2004).The average residual flow generated by human motion will be high than rigid objects such as vehicles, pedestrians or heads. According to this residual flow the authors (Cutler et al., 1998) describes a similarity based procedure to identify the periodic motion. In motion base classification scheme, it does not require predefined pattern templates but it needs greater effort to identify a non-moving human. The motion based classification scheme requires some temporal amount of motion data before classification and this classification method do not performs well when classifying a group of people. Motion based scheme can be used in many tracking applications because the color is relatively constant under viewpoint changes and its computational cost is also low. To detect and track vehicles and pedestrians, a color histogram based technique is used. Color histograms have extremely popular to describe a large image region (Weiming et al., 2004). Since it require less target translation rotation or the target scale variation. In this method, Mixture Model is created to describe the color distribution of images and segment the image from background. In shape based object classification method different classification schemes of shape information of moving regions such as points, bounding box, silhouettes and blobs are existing. The classification should be performed on each blob at each frame by applying the efficient shape based classification scheme. Among the number of classification schemes the appropriate classification features of the detected regions are used and comparing the similarity against the entities on a database. The commonly used features for object classification in video frames are silhouette and this type of shape classification is mainly used in video surveillance systems to classify human object based on shape based classification methods. People have very different shape, appearance and motion patterns compared to other objects like animals car etc. One can use a static shape analysis such as size, area, and ratio or the dynamic motion analysis such as speed or periodicity of the movement to differentiate people from other objects. Different metrics found in literature for object classification and people have very different shape, appearance and motion patterns compared to other objects like animals, car etc. and the size can be used as a classification metric when the objects in the environment have different sizes. As per the survey, (Grabner et al., 2006, Krahnstoever et al., 2006) and (Smith et al., 2006) are uses the size metric, to classify human and luggage in crowded environment such as airports railway stations. The paper proposed by (Grabner et al., 2006), large individuals in the environment are colored as blue whereas abandoned small baggage is colored as red. The main problem of this metric is that when the objects with similar size or when occlusion happens resulting in the change of an object appearance. The next available object classification metric is speed. When the objects have distinct speed, the speed metric of an object can be used to classify different objects. The authors (Bodor et al., 2006) the speed metric is used to classify people on foot and people riding on bicycles. This metric misclassify the objects when speed of the objects become same. Another important metric is Dispersedness. It is a ratio that can be used as a classification metric to different sized objects. The approach presented by (Lipton et al., 1998) uses objects silhouette contour length and information of area to classify objects into human groups, vehicles and other. According to the assumption that the width and length of human is smaller than vehicles and the shape is very complex. The static shape features are directly calculated from the silhouette and its bounding box. By using Dispersedness metric, human objects classification are performed easily.(Dedeoğlu1 et al.) proposed an approach to classify objects as human, human group and vehicles based on comparing silhouette database. In this shape based classification method, classification metric used to measures the similarity of object based on comparing silhouette of the extracted object regions with object silhouettes stored in a database. Then the distance function is measured between query silhouette to be classified and the database.

3. METHODOLOGY :

In visual surveillance system, moving object detection is crucial and very important step that classifies the moving foreground object from the background. The extracted moving foreground object may have human, human groups, animals, vehicles, flying birds, moving clouds, leaves of a tree, or any other noises etc. The purpose of classification stage in the visual surveillance system is to classify the moving foreground object into predetermined categories such as single person, group of people, or vehicles etc. In any human object tracking video surveillance system, the human

object classification is required for further processing. Therefore after foreground object detection the latter task the personal identification can be done efficiently and accurately. In this paper the existing Silhouette Template based Classification method (Yiğithan Dedeo˘glu et al., 2006) is adopted to classify human objects after foreground object segmentation. In the Silhouette Template based Classification the distance signal of the foreground object is computed and compares the similarity with the stored templates of the foreground object is calculated using the minimum distance and is classified into human and human groups. This approach can classify single person or a group of people in a well manner.

3.1. Object Classification Based on Silhouette Template

The Classification metric used in the system to measure the object similarity based on Silhouettes of the detected object regions extracted from the foreground pixel with manually classified template stored in a database. This object classification method is a two-step process of which the first one is an offline step and the other is an online step. In the offline step, the sample object Silhouettes of various objects such as human, group of people and vehicles are stored in a data base by manually labeling the object type whereas in online step, the object Silhouettes are extracted in each frame then comparing its Silhouette based features with template database for recognizing its type. After the comparison a template with minimum distance is found. To achieve the temporal consistency of classification results, the object tracking results are used.

3.2. Extraction of Object Silhouette

In offline and online classification algorithm the Silhouettes of the detected object regions are detached from the foreground pixel by contour tracing algorithm (F. Heijden., 1996). Figure. 1 shows the sample foreground object Silhouette for human object extraction.

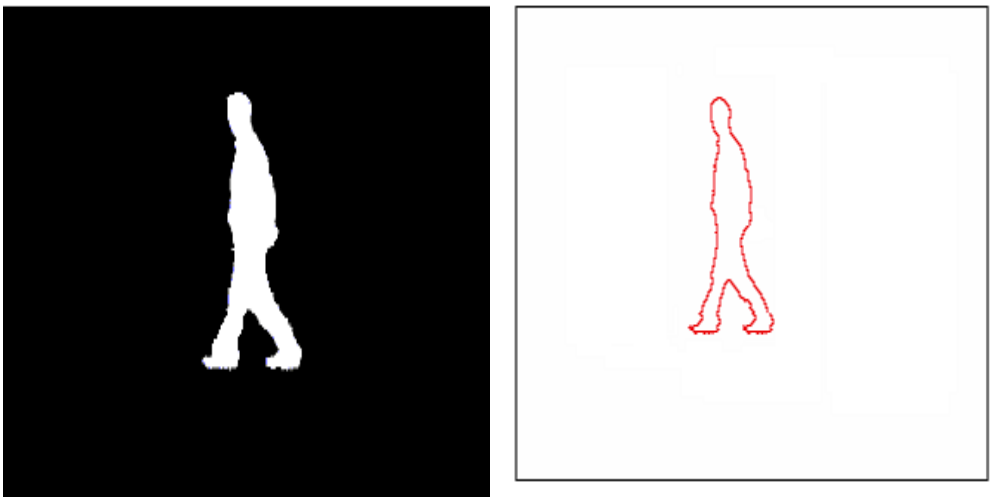


Figure. 1. Sample Foreground Objects and its Corresponding Silhouette

3.3. Silhouette Database Template

In off line step, the template silhouette database is created by extracting different objects contours from different scene. Considering human type, the different shapes of human with different poses are added to the template database to increase the chance of a query object of type human to be categorized correctly. For example, if all the samples are human shapes with vertical positions, it may be misclassified when a human is sitting on a chair. So different pose for human, human group and vehicles are needed to create a database which will help to classify the objects correctly. In classification step, the distance signal and the corresponding type information of each Silhouettes are stored in the template database. Figure 2 shows some sample object silhouette template database with Centre of mass.

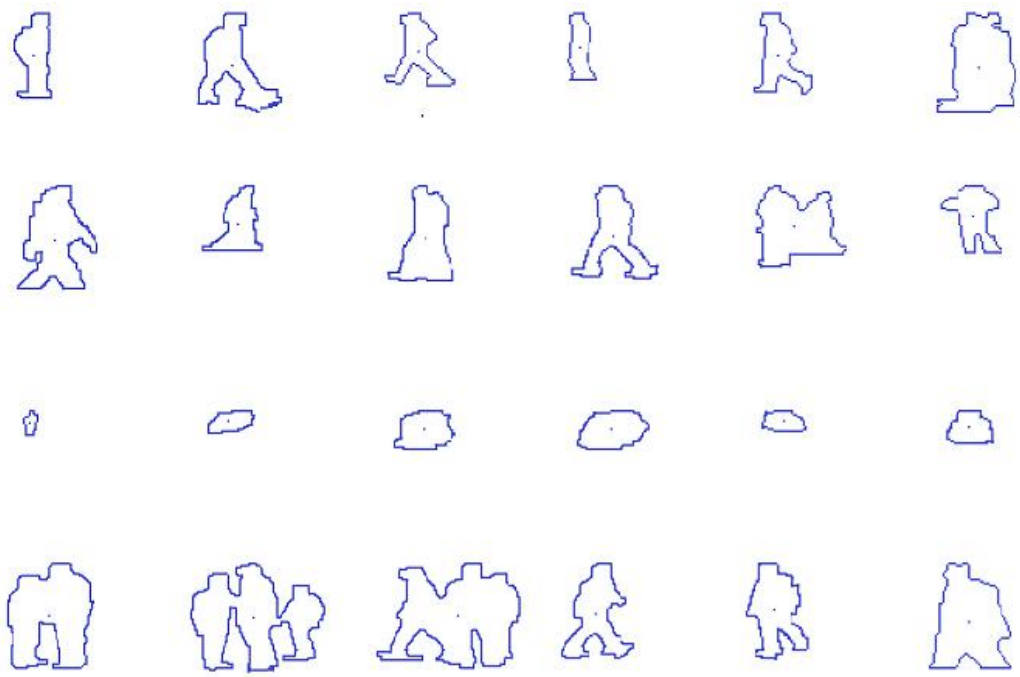


Figure 2. Template Database of Human, Human Groups and Vehicles

Let P= {p1, p2…,pn} be the Silhouette of an object Obj consisting of n points that are ordered from top Centre point of the detected region in clockwise direction and Centre of mass Cm =(Xc,Yc) of object Obj is detected. Silhouettes of each foreground objects after background subtraction is converted into distance signal. The distance signal DS= {d1, d2… dn} is generated by determining the Euclidean distance between CM and the boundary points of the silhouette as follows.

$$Di=Dist(Cm,P_i) \quad) \quad \forall\, i \in [1\ldots.. n] \tag{1}$$

The Silhouettes of each objects have varying sizes because different object have different shapes. Similarly the contour size of same object may vary from frame to frame. In order to compare the signals corresponding to different sized objects and to make a comparison matric scale-invariant clearly the size of the distance signal should be fixed and the fixed sized distance signal DSF is calculated.The fix-sized distance signal DSF is calculated by sub sampling or super sampling the original distance signal DS as follows.

$$DSF[i] = DS\ (i \times N/C) \quad \forall\, i \in [1\ldots.. C] \tag{2}$$

Where N- Original size of a distance signal DS and C is a fixed length signal.The corresponding normalized length signal NSD is represented as,

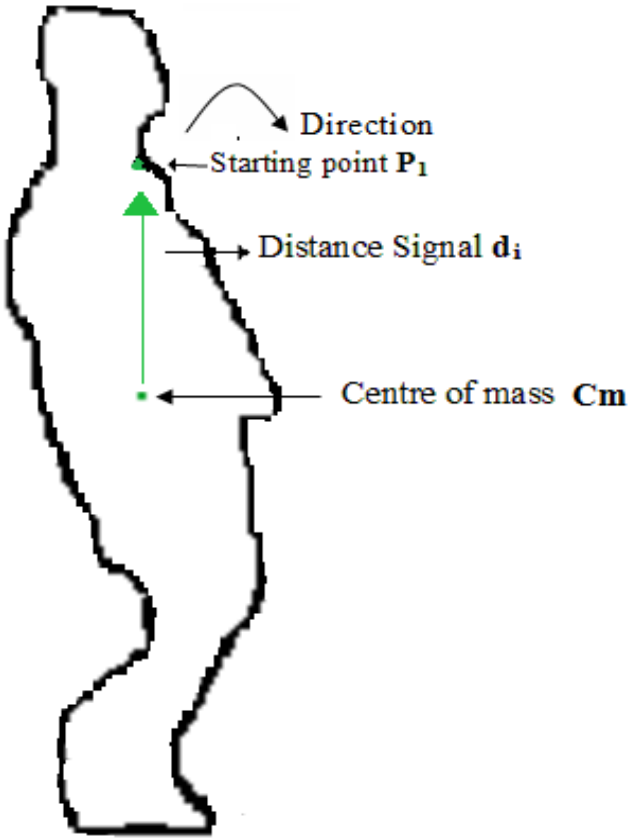
$$NDS\ [i] = DSF\ [i] \,/\, \sum_{i=1}^{n} (DSF\ [i]) \tag{3}$$

Figure 3 shows a sample Silhouette and its corresponding original and scaled distance signal. An object classification is performed by finding out the similarity between the shapes of two objects X and Y, by finding the distance between their corresponding distance signals, DS_X and DS_Y.

The distance between the scaled and normalized distance signals DS_X and DS_Y is calculated as follows.

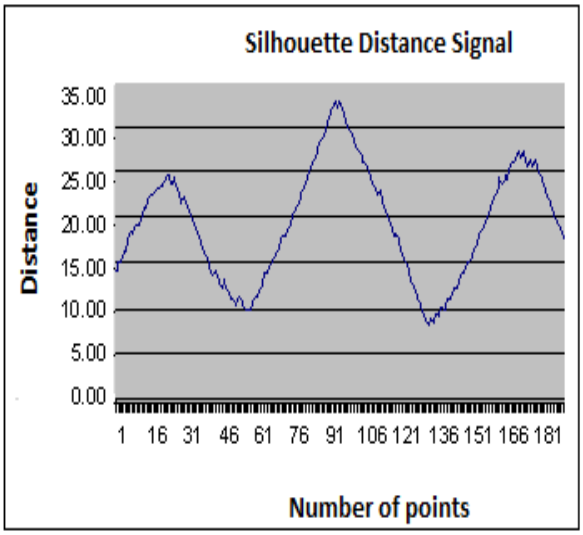
$$Dist_{XY} = \sum_{1}^{n} \text{mod}\ (DSF_X\ [i] - DSF_Y\ [i]) \tag{4}$$

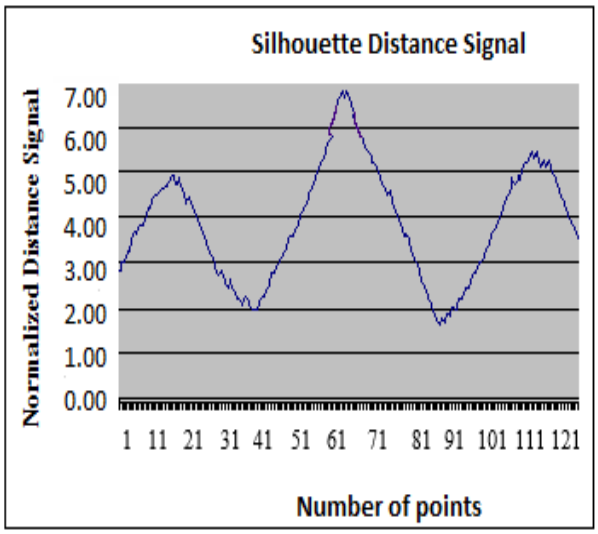
INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD
ISSN(O): 2455-0620 [Impact Factor: 7.581]
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value : 86.87
Volume - 9, Issue - 4, April - 2023 Publication Date: 20/04/2023



(a) Object Silhouette

(b)





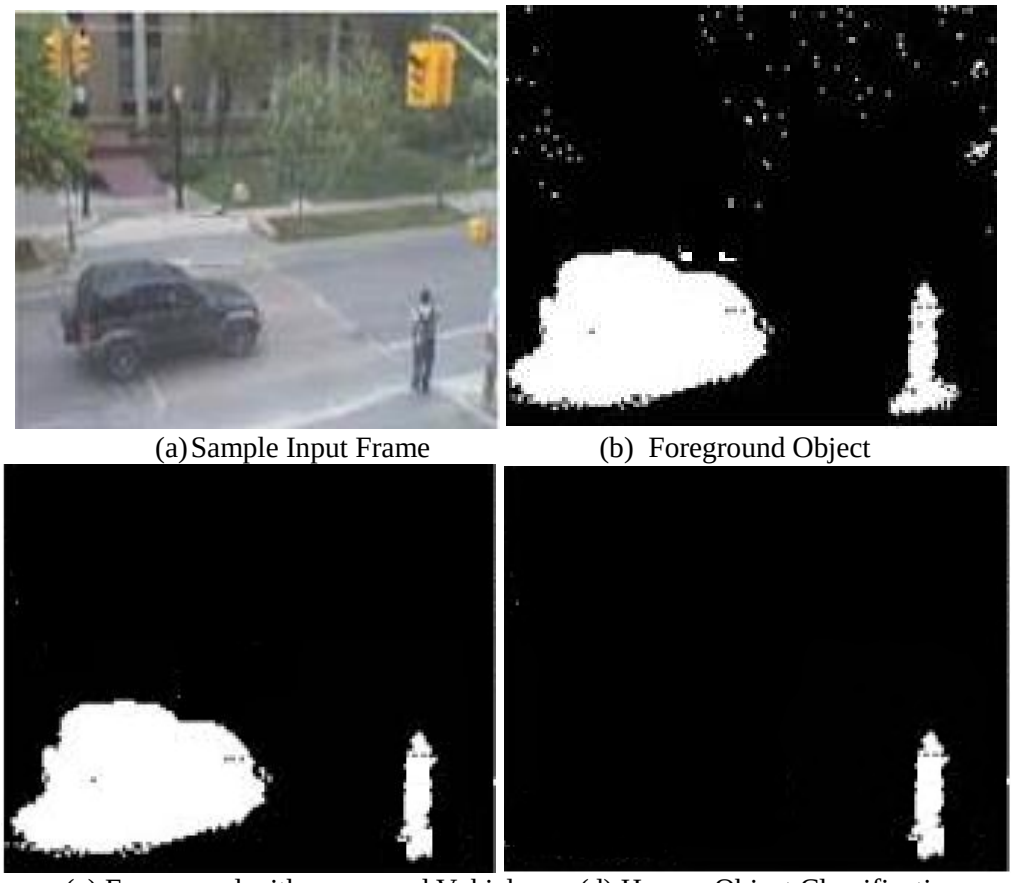
(b) Distance Signal **(c) Normalized Distance Signal**

Figure. 3 Sample Object Silhouette and its Original and Normalized Distance Signals

In order to evaluate the type To of an object Obj , compare its distance signal DSFo with all the distance signals in the template database. The template object P of its type Tp where P satisfies the following;

$$Distop < Distoi \ , \forall \ I \text{ is the template database.} \quad (5)$$

Figure. 4 shows that the sample Human object classification results.



(a) Sample Input Frame (b) Foreground Object

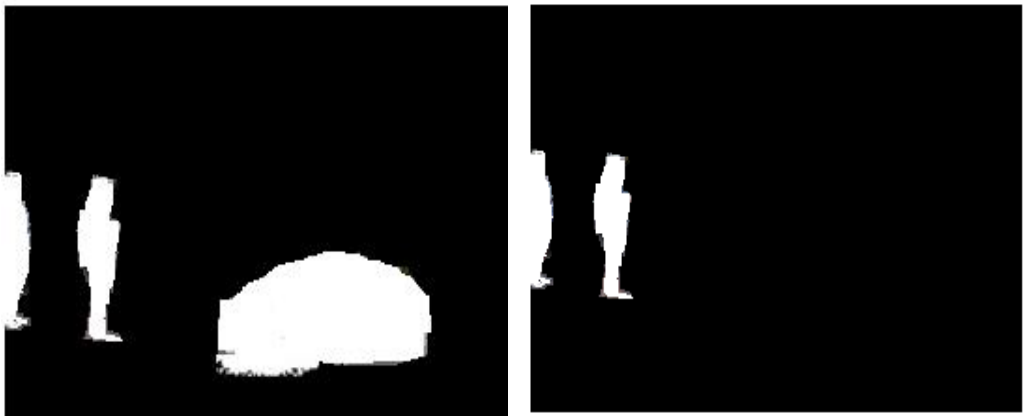
(c) Foreground with a man and Vehicle (d) Human Object Classification

Figure 4 Sample Human Object Classification Result

To show the Object Classification results the input sample (a) is collected from the web. The result of moving Object Segmentation using Gaussian Mixture Model shown in (b) The Gradient based Shadow Removal result is represented in (c) and finally the Silhouette based Human Object Classification result in (d).

4. EXPERIMENTAL RESULT :

When the system is working in a complex public environment the object classification is very essential to classify the human objects. The existing Silhouette based classification provides a better effect and to test the object classification algorithm the template database is created to extract and label object silhouettes. The sample object database classify the objects in different scenes containing human, human groups and vehicles. The foreground object and the corresponding human object classification result of a data sample is shown in Figure 5.



(a) Foreground Object (b) Human Object Classification

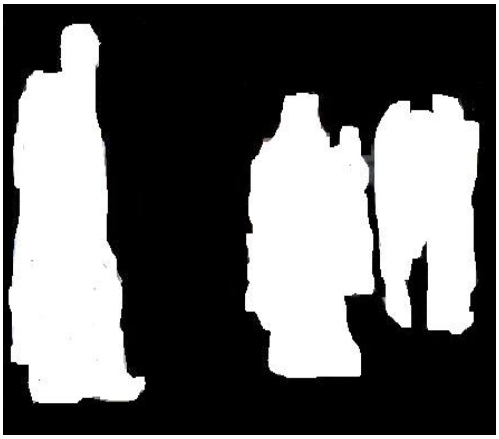
Figure 5 Human Object Classification Result

4.1 Object Extraction from Segmented Human Object

To perform object extraction after object classification check whether each pixel in the foreground objects is one then change the corresponding pixel in the current input frame to one, otherwise it is changed to zero. To show the results two different video samples from different outdoor environments are chosen and the corresponding results are shown in Figure 6 (a) and Figure 6(b).

Sample 1 Foreground Image

(b) Extracted Colour Image



Sample 2 (c) Foreground Image

(d) Extracted colour image

Figure 6 Object Extraction result after Human object classification

5. CONCLUSION :

The classification of moving objects in a video surveillance system is crucial. However, it is sensitive to the used features and number of classes. To solve this problem, researchers proposed the use of several features with a feature selection method. Others added fresh features, and the majority of them made use of binary classification. They do, however, have some drawbacks. The moving objects are extracted from the background using background subtraction method. The existing Shallotte template shaped based classification algorithm provide accurate results in human objects extraction.

REFERENCES :

1. Chapman, A.B., Peterson, K.S., Alba, P.R., DuVall, S.L., Patterson, O.V, (2019). Detecting adverse drug events with rapidly trained classification models. Drug Safety 42,147–156.
2. Xu, H., Li, B., Ramanishka, V., Sigal, L., Saenko, K. (2019). Joint event detection and description in continuous video streams. 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp. 396–405.
3. Zhang, Z., Tao, D. (2019). Slow feature analysis for human action recognition. arXiv preprint arXiv:1907.06670.
4. Avola, D., Bernardi, M., Foresti, G.L. (2019). Fusing depth and colour information for human action recognition. Multimedia Tools Appl. 78, 5919–5939.

5. Liu, L., Lei, X., Chen, B., Shu, L. (2019). Human action recognition based on inertial sensors and complexity classification. J. Inform. Technol. Res. (JITR) 12, 18–35.
6. Zhao, J., Xu, H., Liu, H., Wu, J., Zheng, Y., Wu, D. (2019). Detection and tracking of pedestrians and vehicles using roadside lidar sensors. Transp. Res. Part C: Emerging Technol. 100, 68–87.
7. Cui, Y., Xu, H., Wu, J., Sun, Y., Zhao, J. (2019). Automatic vehicle tracking with roadside lidar data for the connected-vehicles system. IEEE Intell. Syst.Dataset 2014 (2014). Dataset 2014. http://www.changedetection.net.
8. Muchtar, K., Rahman, F., Munggaran, M.R., Dwiyantoro, A.P.J., Dharmadi, R., Nugraha, I.(2019). A unified smart surveillance system incorporating adaptive foreground extraction and deep learning-based classification. 2019International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE, pp. 302–305.
9. Karthikeswaran, D., Sengottaiyan, N., Anbukaruppusamy, S. (2019). Video surveillance system against anti-terrorism by using adaptive linear activity classification (alac) technique. J. Med. Syst. 43, 256.
10. Weiming Hu, Tieniu Tan, Liang Wang, and Steve Mayban.(2004) : A Survey on Visual Surveillance of Object Motion and Behaviors, In: IEEE Transactions on Systems, Man, and Cybernetics-Applications and Reviews, vol. 34, no. 3, pp. 334-352.
11. Cutler .R and Davis L.(1998) :View-based detection and analysis of periodic motion,In: Fourteenth International Conf. on Pattern Recognition, vol. 1, pp. 495-500.
12. Grabner H., Roth P. M., Grabner M. And Bischof H. (2006): Autonomous Learning of a Robust Background Model for Change Detection. In: Proc. of the 9th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, vol. 1, pp.39-46.
13. Krahnstoever N., Tu P., Sebastian T., Perera A. And Collins R.(2006):Multi-View Detection and Tracking of Travelers and Luggage in Mass Transit Environments, In: IEEE International Workshop on Performance Evaluation of Tracking and Surveillance & CVPR, pp.67-74
14. Smith, K. C., Quelhas, P., and Gatica-Perez, D(2006):Detecting abandoned luggage items in a public space, In: Performance Evaluation of Tracking and Surveillance (PETS) Workshop at CVPR.
15. Bodor Robert, Bennett Jackson Nikolaos Papanikolopoulos(2003):Vision-Based Human Tracking and Activity Recognition," In: Proc. of the 11th Mediterranean Conference on Control and Automation, vol. 1, pp. 18–20.
16. Lipton A. J., Fujiyoshi H. And Patil R. S.(1998): Moving Target Classification and Tracking from Real-Time Video, In: Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision, pp. 8–14.
17. Yiğithan Dedeoğlu B., Uğur Töreyin, Uğur Güdükbay and Enis Etin A.(2006): Silhouette-Based Method for Object Classification and Human Action Recognition in Video, In: International Workshop on Human-Computer Interaction, Graz, Austria, (in conjunction with ECCV 2006). Lecture Notes in Computer Science, vol. 3979, pp. 64-77, Springer-Verlag GmbH.
18. F. Heijden.(1996): Image based Measurement Systems: Object Recognition and Parameter Estimation, Wiley, vol.17, no.16, pp.348.