Journal of Theoretical and Applied Information Technology
15th May 2018. Vol.96. No 09

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A MULTI-CLASS SPATIAL CLASSIFIER FOR IMPROVING MOVING OBJECTS DETECTION AND TRACKING

[1]MATHESWARI RAJAMANICKAM, [2] DR . AISHWARYA PALANIAPPAN

[1] Assistant Professor, Department of Computer Science,

M.E.S College of Arts, Commerce and Science, Bangalore, Karnataka, India.

[2]Professor and Head, Department of Computer Science and Engineering,

Atria Institute of Technology, Bangalore, Karnataka, India.

E-mail: [1]matheswarir@gmail.com, [2]aishwarya.p@atria.edu

ABSTRACT

Object detection and tracking aims at detecting and tracking several objects over a sequence of images (i.e. video). The research works focuses on image segmentation and classification for moving object detection and tracking. Batch incremental classifier is used for classifying four classes of moving objects. Object fusion technique developed to solve the detection and tracking of moving objects failed to generate the visual descriptors of the image. In order to obtain efficient and robust moving object detection and tracking with minimum time, a Multi-Class Spatial Classifier (MCSC) method is introduced. Initially, the Markov Random Field (MRF) principle is applied in MCSC method for both fixed and moving objects in a sequence of images to provide label field fusion that exhibits the shape of the objects in the scene. Next, the fuzzy edge strength of each pixel location is identified with MRF modeling, which preserves the object boundary to classify the images of the segmented objects. Finally, the spatial classification is performed through the maximum a posteriori probability (MAP) estimation principle. This helps to improve the classification accuracy with minimum time. The performance of the proposed MCSC method achieves higher object detection and tracking accuracy in terms of classification accuracy, false positive rate and classification time.

Keywords: *Moving Object Detection and Tracking, Multi-Class Spatial Classifier, Object Classification, Markov Random field, Fuzzy Edge Strength, MAP Estimation.*

1. INTRODUCTION

Detecting the objects in the video (i.e. sequence of images) and tracking its motion to identify its features is a challenging research area in the domain of image processing and computer vision. The classification of objects is carried out under spatial characteristics relating to the position, area, and size of things. For measuring the behavior of objects in the frame, classification between the numbers of objects in specific frame is necessary. Classification is performed with respect to object shape, motion, color and texture. In our proposed methods, classification is performed to detect and track the moving objects.

Object tracking is an essential task in several computer vision applications, such as intelligence video surveillance, human machine interfaces, indexing and so on. Robust object tracking significantly increases the performance of intelligence video surveillance, human machine interfaces and action analysis. However, designing robust object tracking techniques is still an open issue; especially in the case of complicated deviations that occur in the dynamic scenes like illumination variations, pose changes, background clutters, occlusions, etc. Hence, a texture based approach determines the occurrences of gradient directions in the localized part of an image for improved classification accuracy. It provides better accuracy but it consumes more time. Therefore, the proposed Multi-Class Spatial Classifier method is introduced to improve the object detection accuracy with the help of improved classification techniques.

An intelligent system was designed in [1] using Batch incremental SVMs (BISVM) classifier for classifying four classes of moving objects using a new feature descriptor that considers both object form and moving information to improve the classification accuracy. However, the feature extraction and classification approach was not combined with other tracking algorithms for applications in crowded urban regions.

A multi-sensor fusion framework was introduced in [2] for Moving Object Detection and Tracking and also reduced the number of false

Journal of Theoretical and Applied Information Technology
15th May 2018. Vol.96. No 09

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

detections. However, the visual descriptors of the image are not generated for detecting the moving objects.

Moving object area detection method introduced in [3] can efficiently handle a variation of the object size, camera shake problem, slow and fast moving object. However it does not support for moving camera which resulted in lack of object detection and tracking. Adaptive block background making model and dynamically adaptive threshold technique was developed in [4]. However this algorithm needs improvement for more robust and efficient detection on night videos.

In [5], detection and Tracking Moving Objects (DATMO) algorithm was designed to reduce the number of false positives. Moving objects detector uses small layer laser scanner which provides very little information. Hence it takes more time to classify the images.

Detection and tracking method was designed in [6] to detect and track sea-surface obstructions by processing forward-looking sonar images. The problem of wake detection method increased false-alarm rate.

A temporal-spatial variable scale algorithm was designed in [7] for moving objects detection from complex backgrounds. However, it failed to obtain the 3-D factors and the multiple moving objects detection becomes highly difficult.

A novel video stabilization and moving object detection system was designed in [8] based on camera motion estimation using Scale Invariant Feature Transform (SIFT). But it reduced the processing speed as well as robustness and also increased the complexity.

A novel moving object detection algorithm was introduced in [9] using illumination change approach and chromaticity difference method, and a brightness ratio model. But, it failed to estimate exact thresholds for false foreground pixels.

In [10], an analysis of different methods for object detection and tracking was designed in Video-Based Surveillance System. However, developing a real-time video technique for handling large amounts of data becomes very difficult.

From the above said literature, the issues like false detection, high false alarm rate, high complexities etc., makes the existing methods a failure one to generate the visual descriptors of the image and also the inability to handle large number of objects. In order to overcome these issues in object detection and tracking, an efficient Multi-Class Spatial Classifier (MCSC) technique is introduced.

The contribution of the proposed MCSC is described as follows,

- Efficient and robust moving objects detection and tracking technique called Multi-Class Spatial Classifier (MCSC) is developed. The MCSC method contains three major contributions for improving the moving object detection and tracking accuracy in a given video frame. The sequence of video frame is taken as input with different size. By applying the input video frames, motion of the object is detected.
- Initially, Markov Random Field (MRF) principle is applied for both fixed and moving objects in a video frame to provide the label field fusion. The MRF principle is used to show the shape of the moving or fixed object in a given video frame. This helps to reduce the classification time.
- The second step is to apply fuzzy approach to preserve the objects' boundary with the help of 'if-then' rules. The fuzzy triangular membership function is used to find the edge-strength with three fuzzy sets namely "low", "medium," and "high". Object boundaries are preserved based on these high edge-strength values.
- The final contribution of MCSC method is to classify the moving objects from the video frames. The spatial classification is performed using Maximum a Posteriori (MAP) estimation. By applying MAP estimation, spatial classifier predicts the probabilities of different classes with higher likelihood values. The likelihood estimation is carried out with Bayes' theorem. This helps to improve the classification accuracy with minimum time and also reduces the false positive rate.

The structure of the research work is organized as follows: Section 2 presents the related work. Section 3 presents the proposed method multi-class spatial classifier (MCSC) technique. Section 4 presents the experimental settings. Section 5 presents the performance evaluation. Section 6 presents conclusion and future work.

2. RELATED WORKS

In [11], a novel framework for explaining two environmental awareness tasks using local volumetric hybrid map based SLAM with MOT was developed. However, the moving object detection and tracking was not improved at a required level. The MCSC method performs effective classification to improve the moving object detection and tracking.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A bag-of-regions (BoR) representation was carried out in [12] to perform video classifications using SVM. However, efficient classifiers were not used to offer significant improvement. The MCSC method uses an effective classifier to improve the classification accuracy.

A removed and abandoned blob classifier (RABC) was introduced in [13] to classify static foreground blobs in surveillance circumstances. But it failed to perform object detection and tracking. The MCSC method improves the moving object detection and tracking accuracy.

A tracking technique using multi-view learning approach using multiple support vector machines (SVM) was introduced in [14]. However, the classification time remains unaddressed. The multi class spatial classifier classifies the object with minimum time.

A combined approach for model adaptation was introduced in [15] for video tracking but the object detection remained unsolved. The proposed classifier effectively classifies the labels using MAP estimation to detect and track the moving object.

A spatiotemporal saliency in moving object detection was performed in [16] using hierarchical method for spatial distribution. However, false alarms rate was not reduced. The MCSC method reduces false positive rate.

A visual detection and tracking approach was designed in [17] for surveillance applications. But the classification of moving objects remained unaddressed. The proposed MCSC method performs effective classification of the moving objects with minimum time.

A new part-based model called as And-Or graph technique was introduced in [18] to distinguish object shapes in images. However, a moving object tracking was not performed. The MCSC method effectively tracks the moving object through the classification approach.

A systematic structure for Intelligence Video Surveillance System (IVSS) was presented in [19] with a multi-camera network. But an effective classification was not performed. The MCSC method performs classification of moving object detection effectively.

An extended MCMC method called as EMCMC was introduced in [20] for robust tracking but the boundary and edge strength of the object was not measured. The proposed MCSC performs edge-strength measurement to detect and track the object.

Background Subtraction Frame Difference Algorithm (BFSD) was designed in [21] to identify and extract the moving objects detection from input video frame. However, the location of object tracking was not performed. An integration of optical flow estimation and background modeling technique was developed in [22] for moving object detection. However, the performance of this method does not improve moving object detection and tracking.

3. PROPOSED METHOD

3.1 A Multi-Class Spatial Classifier for Object Detection and Tracking

Object detection and tracking is done to determine the presence of objects in video frame and to track the location of those objects. Object detection is labeled into set of classes such as humans, vehicles and other moving objects in video frame. Object tracking involves the process of classifying a segmented image from video frames and tracking its movement and position. In recent days, moving object tracking from the video sequences plays an essential role due to its huge practical applications such as visual surveillance, perceptual user interface, content-based image storage and retrieval, athletic performance analysis etc. Batch incremental SVMs (BISVM) classifier falsely detects the moving objects and failed to preserve object boundaries. The other existing multi-sensor fusion framework method used for Moving Object Detection and Tracking failed to create the visual descriptors of the images. The motivation of the research work is to reduce the false detection of moving objects by performing edge-strength estimation for preserving the object boundary. In order to address the above said issues, Multi-Class Spatial Classifier (MCSC) method is proposed for improving the moving object detection and tracking accuracy with reduced false positive rate. Spatial classification is done based on the investigation of spatial objects related to its spatial features like areas region, roads, and ponds or rivers. The MCSC method performs the classification of object based on their segmented texture features which results in improved object classification accuracy with minimum time. The diagram of the proposed MCSC method is shown in figure 1.

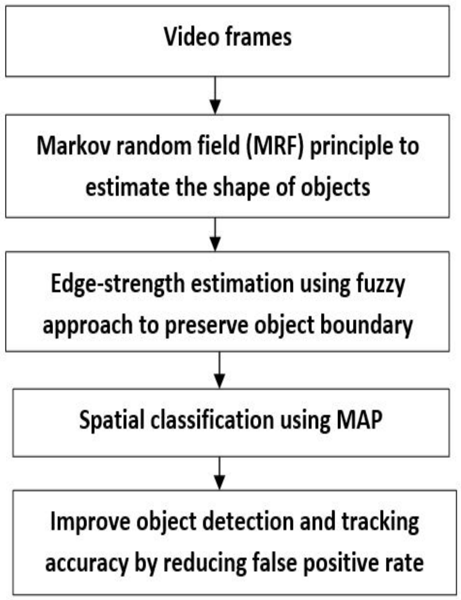

Figure 1. Flow Processing Diagram of Multi-Class Spatial Classifier Technique

Figure 1, illustrates the flow processing of the Multi-Class Spatial Classifier (MCSC) technique. In segmented video frames, the objects are being classified using Multi-Class Spatial Classifier technique. MCSC technique takes the video file as input as shown in figure 2.

Figure 2. Input video file

Initially, a Markov random field (MRF) principle is applied for fixed and moving object which provides a label field fusion. This exhibits shape of the main objects in the video frame. Next the spatial classifier technique is carried out by integrating the fuzzy edge strength of pixels in MRF modeling to preserve the object boundary. Finally, spatial classification is done using Maximum a' posteriori (MAP) to classify the objects based on the number of class labels. The main objective of the spatial classification is to generate an equivalent grid of class labels that recognize the classes present at each location in the frames, which has higher likelihood values. This helps to track the moving object effectively in video frame. The brief description about the Multi-Class Spatial Classifier technique is presented in forthcoming sections.

3.2 Markov Random Field (MRF) Principle

The first process in the design of MCSC method is the application of Markov Random Field (MRF) principle. The MRF principle is used to apply for both fixed and moving object in order to provide a label field fusion. A Markov random field fusion approach groups the label fields collectively. A label field consists of dissimilar information. In MCSC method, the MRF principle is used to fuse the label field in a flexible manner. More specifically, the MRF considers two label fields such as a region plot (R) which is obtained after segmenting the input frames and a spatial region plot ('S') of the application label field. The spatial regional plot 'S' is application specific and it contains the information about the occlusion labels (i.e. boundary labels), motion labels, or any other high-level information. These two label fields are combined together to reduce few local minima.

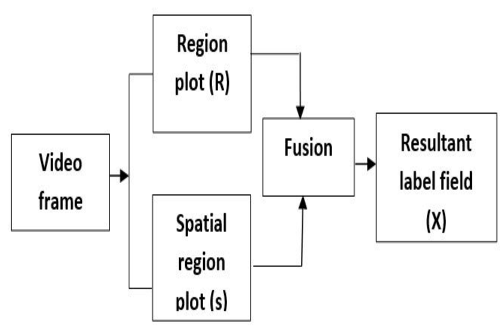

Figure 3. Schematic View of the Fusion Framework

Figure 3 shows the schematic view of the fusion approach. The input video frame is segmented in to region plot (R) and the spatial region plot (S). These two labels are combined to provide the resultant label field (X). Let us consider, R and S is realization of a pair of joint Markov random field (MRF) and it is estimated by the property of the Hammersley–Clifford theorem. The joint probability density of the Markov random field is described as follows:-,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$P(R_l, X_l, \varphi_l) = \frac{1}{Z}\exp(-E_l(R, X|\varphi_l)) \quad (1)$$

From (1), where Z represents the normalization factor, $E_l(R, X|\varphi_l)$ denotes the local energy function measuring *how well R and S fist together around the location* (l), φ_l is the local joint neighborhood function around the location in (l). More specifically, the value of E_l is small when the region X fits locally i.e. there is no edges in X cross a uniform region R. In addition, E_l is the energy function that counts the number of neighboring pixels 'P' around the location *l*, whose label 'R_t' and X_t is same as R_l and X_l. Therefore, the energy function is measured as follows:-,

$$E_l(R, X|\varphi_l) = -\sum_{t\in\varphi_l} \delta(R_t, R_l)\,\delta(X_t, X_l) \quad (2)$$

From (2), $\delta(X_t, X_l)$ and $\delta(R_t, R_l)$ is the Kronecker delta function. This delta function is expressed as,

$$\delta(X_t, X_l) = \begin{cases} 1 & if\ X_t = X_l \\ 0 & otherwise \end{cases} \quad (3)$$

Here, equation (3) is called as the fusion model that measures the spatial similarity of the joint pair of Markovian random fields. Such a way, the resultant label field is measured as follows:-,

$$X = \arg\min_{X} \sum_{l\in L} E_l(R, X|\varphi_l) \quad (4)$$

For each location, the resultant value (X) that maximizes the local conditional probability density function and it considerably reduces the classification time. Therefore, The MCSC method uses MRF principle to show the shape of the moving or fixed object. The output of the shape extraction using fusion approach is shown in figure 4.

Figure 4 (a) Markov random field object of frame 1 (b) Markov random field object of frame 2

Figure 4 shows the Markov random field object of frame 1 and Markov random field object of frame 2. From the figure, shape of the object is identified in the particular frame.

3.3 Fuzzy Rules to Estimate the Edge-Strength

Once the shape of the object is estimated using MRF principle, the proposed MCSC method determines the edge-strength for preserving the object boundary. Hence the proposed MCSC method uses fuzzy *"IF . . . THEN . . ."* rules and calculates the edge-strength. In general, a fuzzy approach considers a set of "antecedent clauses". Let us consider, the two gradient operators G_x and G_y at a point *(x, y)* in the image. An object gradient is a directional change in the intensity of the image. The gradient of the object is one of the basic structure blocks in object tracking. The gradient of the object is mathematically defined as,

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (5)$$

From (5), $\frac{\partial f}{\partial x}$ denotes the derivative with respect to gradient in *x* direction, $\frac{\partial f}{\partial y}$ represents the derivative with respect to gradient in *y* direction. These two gradient values of the object are linked and it is used for edge-strength measure using fuzzy analysis. The proposed MCSC method measures the edge-strength of pixels by using fuzzy logic based on a set of four rules. The proposed fuzzy system consists of two-input and one-output system. Let us consider two input variables (antecedent clauses) denoted as two gradients G_x and G_y, and the resulting output variable (consequent clause) as the normalized edge-strength $N(x, y)$, at pixel location *(x, y)* in the sequence of image. The four *IF-THEN* rules in fuzzy system are used to determine the edge strength. In Rule 1, if G_x *AND* G_b is low then the normalized edge-strength $N(x, y)\ is\ low$. Rule 2, if G_a is low *AND* G_y is high then the normalized edge-strength $N(x, y)\ is$ medium. In rule 3, if G_a is high *AND* G_y is low then the normalized edge strength $N(x, y)\ is$ medium. In rule 4, if G_x *AND* G_y is high then the normalized edge strength $N(x, y)\ is$ high. These four IF-THEN fuzzy rules include two conditions about the input variables and the output variable. The input variables are linked by fuzzy '*AND*' operator. Every clause is fully defined by the shape and location of a fuzzy set, which maps the equivalent variable to the real interval [0, 1]. Since, the two input variables G_x and G_y are measures of the similar quantity (intensity gradient) differing only in their directions of measure (horizontal (x) and vertical (y) directions). The membership function plots are shown in figure 3.

Journal of Theoretical and Applied Information Technology
15th May 2018. Vol.96. No 09

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

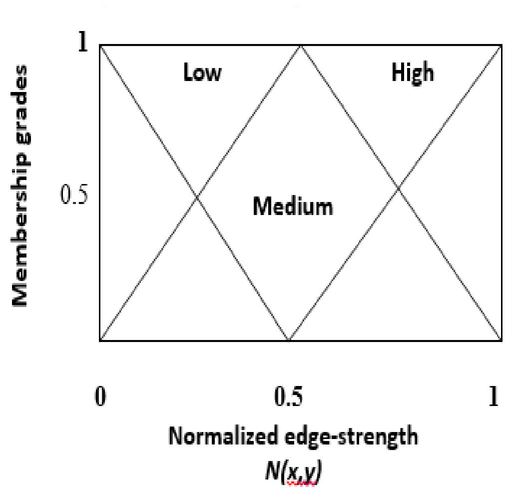

Figure 5 Fuzzy Membership Functions with Edge-Strength

Figure 5 shows the fuzzy triangular membership function to determine the edge-strength and it is used for preserving the object boundary. As shown in figure 3, fuzzy sets provides the resultant values as "low", "medium," and "high". Based on the edge-strength measurement, the location of each pixel of the object is identified. Therefore, the normalized edge-strength *N (x,y)* is measured based on the amount of horizontal and vertical intensity gradients at pixel location *(x, y)* in the object. The input image with fuzzy edge strength measure is shown in figure 6.

Figure 6 .Fuzzy Normalized Edge-Strength for Frame 1

Figure 6 shows the normalized edge-strength for frame 1 by using fuzzy triangular membership function. The algorithmic representation of fuzzy based edge strength is shown in algorithm 1.

Input : *Two gradients G_x and G_y,*
Output: *Preserve object boundary*
Step 1: Begin
Step 2: *Consider two gradient G_x and G_y,*
Step 3: ***If*** *G_x AND G_y low then*
Step 4: *Low edge-strength $N(x,y)$ is obtained.*
Step 5: ***If*** *G_x low AND G_y high then*
Step 6: *Medium edge-strength $N(x,y)$ is obtained.*
Step 7: ***If*** *G_x high AND G_y low then*
Step 8: *Medium edge-strength $N(x,y)$ is obtained.*
Step 9: ***If*** *G_x AND G_y high then*
Step 10: *Higher edge-strength $N(x,y)$ is obtained.*
Step 11: End if
Step 12: End if
Step 13: End if
Step 14: End if
Step 15: End

Algorithm: 1 Fuzzy based edge-strength algorithm

As shown above, fuzzy algorithm is designed for measuring the edge-strength in order to preserve the object boundary. Initially, two gradient values of objects are calculated with respect to pixel location *(x, y)* for determining the edge-strength at that point. The fuzzy IF-THEN rules in fuzzy system are applied to measure the edge-strength as low, medium and high. As a result, the higher edge-strength is obtained for preserving the boundary of the object.

3.4 Spatial Classification Using MAP Estimation Principle

The final step in the design of the MCSC method is the spatial classification using MAP estimation to improve the classification accuracy and to reduce the false positive rate. This MAP estimation function classifies the objects from the video frames effectively based on different class labels. Likelihood of each pixel's class label is derived from the classification model to produce the final classified results. This approach is used to perform the spatial classification based on the analysis of moving objects related to its spatial characteristics, such as areas region, roads, and ponds or rivers. Bayes' theorem is used for classifying the segmented objects in the estimation process for reducing the false positive rate.

$$P(X|D) = \frac{P(D|X)\,P(X)}{P(D)} = \frac{\text{likli hood} * \text{prior information}}{\text{denormalizing factor}}$$
(6)

From (6), P(X) is said to be a probability of prior information about the true classification and it's ratio of the number of location and the different

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

class labels. Here, $P(X|D)$ represents the posterior probability. From (6), the numerator denotes the product of the likelihood term and the prior information about the class labels. These classifications are used to determine a set of class likelihoods for every pixel, and the neighborhood relations.

In case of the location prediction problem, where a single class label is estimated for each location, a decision process is taken by the Bayes' rule based on the most likelihood value for a particular location. This solution is termed as Maximum a Posteriori estimate (MAP). Therefore, Bayes rule inference produces the maximum a posteriori (MAP) estimation which is expressed as,

$$\arg\max P(X|D) = \arg\max \; P(X|D)\,P(D) \quad (7)$$

The MAP finds maximum value of the posterior, where the posterior is a probability density function. Based on the above formulation, the Maximum posterior probabilities for different class labels are obtained.

Let us consider, a set of segmented objects $O_i = o_1, o_2, o_3, \ldots, o_n$, and the corresponding class labels $c_i \in C$. By using MAP estimation (c_{MAP}), the objects are classified by using Bayes classifier, and is expressed as follows:-,

$$c_{MAP} = \arg\max_{c_i \in C} P(c_i | O_i = o_1, o_2, o_3, \ldots, o_n) \quad (8)$$

From (8) the classifier predicts the probability of the class which has most likelihood value. Therefore, MAP returns class value where the probability is highest in the given frames. Spatial classification of input image is shown in figure 7.

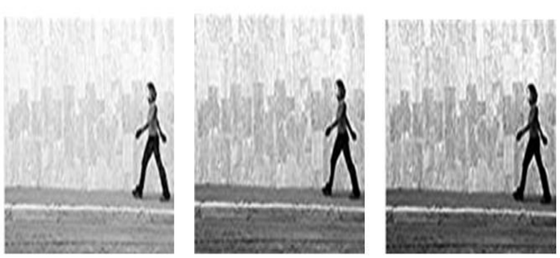

Figure 7 Spatial Classification Level 1, Spatial Classification Level 2 and Spatial Classification Level 3

As shown in figure 7, the classification model to produce final classified results of input moving object. This reduces the false positive rate in figure 7.

Input: *Video frames* $VF_i = VF_1, VF_2, VF_3, \ldots, VF_n$, *Segmented objects* $O_i = o_1, o_2, o_3, \ldots, o_n$
Output: *Improved Classification accuracy*
Step 1: Begin
Step 2: *For each segmented objects*
Step 3: *Perform MAP estimation using (6),*
Step 4: *Bayesian interference provides the MAP using (7)*
Step 5: *By using MAP estimation, the objects are classified using (8)*
Step 6: Classifier predicts the probability of the class with higher likelihood value
Step 7: Improved classification accuracy and reduced false positive rate
Step 8: End for
Step 9: End

Algorithm: 2 Maximum a' Posteriori Estimation Algorithm

As shown above, the Maximum a' posteriori estimation algorithm is designed to perform spatial classification. For each segmented objects in the video frames, the Maximum a' posteriori estimation is carried out to classify the objects with different class labels. By applying MAP estimation, classifier predicts the probability of the class with higher likelihood values. As a result, the MAP estimation technique improves the classification accuracy and also reduces the false positive rate. Finally, the scenario of object tracking is performed based on the classification results related to its spatial characteristics is shown in figure 8.

Figure 8 Object Tracking

As shown in figure 8, the object in the video frames is correctly tracked through the classification. This improves the object tracking accuracy effectively.

4. EXPERIMENTAL SETTINGS

Experimental evaluation of Multi-Class Spatial Classifier (MCSC) method is implemented using MATLAB coding to improve the moving object

Journal of Theoretical and Applied Information Technology
15th May 2018. Vol.96. No 09
© 2005 – ongoing JATIT & LLS
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

detection and tracking accuracy. The 168VJ Clips video dataset is used to perform the experimental evaluation. The video file is taken from the 168VJ Clips video dataset to perform classification of moving object detection and tracking. This dataset contains different video clips in different file size. The (MCSC) method is used to improve the video quality and simulation conducted on .avi file format. The evaluation of MCSC method is carried out using Batch incremental (BISVM) classifier [1] and multi-sensor fusion framework [2]. The experiment is conducted on the factors such as classification accuracy, false positive rate, classification time and moving object detection and tracking accuracy.

5. RESULTS AND DISCUSSIONS

The Multi-Class Spatial Classifier (MCSC) method is analyzed with existing BISVM classifier [1] and multi-sensor fusion framework [2]. The performance is carried out on the factors such as classification accuracy, false positive rate, classification time and moving object detection and tracking accuracy. The performance is evaluated by means of tables and graph values.

5.1 Impact of Classification Accuracy

Classification accuracy is defined as the ratio of the number of objects that are correctly classified to the total number of objects in video frames. The formula for classification accuracy is defined as follows: -,

$$CA = \frac{\text{Correctly classified objects}}{\text{No.of objects in video frames}} * 100 \quad (9)$$

From (9), CA denotes classification accuracy and it is measured in terms of percentage (%). Let us Consider 10 input video frames. The number of objects in those frames is considered as 10. From these 10 objects, the numbers of correctly classified objects are measured.

Table 1. Tabulation for Classification Accuracy

No. of Video frames/sec	Classification accuracy (%)		
	MCSC	BISVM	Multi-sensor fusion framework
10	70.25	60.10	51.36
20	75.64	65.85	55.73
30	82.10	68.67	60.11
40	84.36	72.12	65.47
50	85.22	75.98	68.32
60	86.14	80.11	73.65
70	88.79	83.65	75.82
80	90.12	85.79	78.65
90	93.65	86.55	80.12
100	95.48	88.74	84.65

Table 1 shows the classification accuracy with respect to the number of video frames ranging from 10 to 100 frames/sec. The classification accuracy is improved in the proposed MCSC method than the existing methods [1], [2]. This is due to the application of MAP in MCSC method. In MCSC method, spatial classification is effectively performed to classify objects with their corresponding class labels. By applying MAP estimation, the object which is related with spatial characteristics is classified in accordance with class label and it has higher likelihood value. For the experimental consideration, the number of frames taken, say 10 and the objects in those frames, say 10. Out of 10 objects, the 7 objects are correctly classified by our proposed MCSC method whereas 6 objects from BISVM and 5 objects from multi-sensor fusion framework. The research work shows significant improvement in the classification of moving objects based on their multiple features like size, shape, boundary, edges. The classification accuracy is considerably increased by 11% and 24% using MCSC method compared to existing BISVM classifier [1] and multi-sensor fusion framework [2] methods respectively.

5.2 Impact of false positive rate

False positive rate is defined as the ratio of the number of objects incorrectly classified to the total number of objects in video frames. The formula for false positive rate is defined as follows: -,

$$FPR = \frac{\text{Incorrectly classified objects}}{\text{No.of objects in video frames}} * 100 \quad (10)$$

From (10), FPR denotes a false positive rate and it is measured in terms of percentage (%). Consider 10 input video frames, the number of objects in that frames are considered as 10. For these 10 objects, the numbers of incorrectly classified objects are measured.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2. Tabulation for False positive rate

No. of Video frames/sec	False positive rate (%)		
	MCSC	BISVM	Multi-sensor fusion framework
10	31.24	42.65	51.23
20	35.45	48.12	53.48
30	37.46	50.32	55.24
40	40.12	52.48	56.21
50	42.57	55.64	60.75
60	45.69	58.75	62.44
70	48.10	60.12	65.10
80	50.24	63.45	68.46
90	52.67	65.71	72.12
100	54.79	68.45	75.68

Table 2 shows the false positive rate with respect to the number of video frames. The incorrectly classified objects are measured with three different techniques namely MCSC method BISVM classifier [1] and multi-sensor fusion framework [2].

In MCSC, a Markov random field principle is applied to estimate the shapes of both moving objects and fixed objects in the video frames. MRF principle provides label fields in a flexible manner thereby exhibiting the shape of objects in the video frame. Moreover, an object boundary is preserved by our Fuzzy rules. The process in MCSC method helps to classify the object in video frame. Moreover, the MAP estimation is performed to classify objects with higher likelihood values. This helps to perform effective classification and reduces the false positive rate. The false positive rate is considerably reduced in the proposed MCSC method by 23% and 30% compared to the existing BISVM classifier [1] and multi-sensor fusion framework [2] respectively.

5.3 Impact of Classification time

Classification time is defined as the amount of time taken to classify the object in video frames which is measured in terms of milliseconds (ms).

$$\text{classification time} = \text{No. of objects in video frame} * \text{time taken for classifying one object} \quad (11)$$

Table 3. Tabulation for classification time

No.of Video frames/sec	Classification time (ms)		
	MCSC	BISVM	Multi-sensor fusion framework
10	12.6	15.6	18.5
20	15.4	18.7	20.4
30	18.9	21.5	25.7
40	21.3	24.8	28.9
50	22.7	27.9	32.4
60	24.8	28.6	35.7
70	30.1	35.4	38.9
80	32.4	38.7	40.2
90	35.3	40.1	44.5
100	38.8	42.8	46.7

Table 3 illustrates the classification time of our proposed MCSC method, BISVM classifier [1] and multi-sensor fusion framework [2]. It's observed that while increasing the number of video frames, the object classification time also gets increased in all the methods. But the classification time is comparatively reduced in our proposed MCSC method. This is because; the spatial classification is performed through the MAP estimation to classify the objects in the video frames based on the texture feature. The MAP determines the maximum posterior value. The Maximum posterior probabilities provides class labels for object in the video frames. By performing MAP estimation, the objects in video frames are classified by using Bayes classifier with minimum time. Moreover, the classifier predicts the probability of the class with higher likelihood values. This helps to improve the classification with minimum time. The classification time is considerably reduced by 15% and 25% compared to the existing BISVM classifier [1] and multi-sensor fusion framework [2] respectively.

5.4 Impact of Object detection and tracking accuracy

Moving object detection and tracking accuracy is defined as the ratio of number of objects detected and tracked correctly to the total number of objects in the video frames. The object detection and tracking accuracy is formulated as,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$ODTA = \frac{\text{Object being Detected and tracked}}{\text{Number of objects in video frames}} * 100$$

(12)

From (12), '$ODTA$' signifies the object detection and tracking accuracy. It is measured in terms of percentage (%).

Table 4. Tabulation for object detection and tracking accuracy

No. of Video frames/ sec	Object detection and tracking accuracy (%)			
	MCSC	**BISVM**	**Multi-sensor fusion framework**	**BFSD**
10	80.25	70.31	61.85	51.32
20	82.43	73.56	64.58	54.52
30	83.52	75.14	66.20	57.65
40	85.10	77.32	69.14	60.12
50	86.67	78.10	72.46	63.65
60	88.12	80.65	75.35	66.78
70	90.65	82.12	78.26	69.52
80	92.58	84.63	80.64	72.45
90	94.45	85.41	81.65	75.41
100	96.13	87.46	83.20	78.65

Table 4 clearly shows the object detection and tracking accuracy for our proposed MCSC method along with the two existing methods BISVM classifier [1] and multi-sensor fusion framework [2]. It is observed that the proposed method gives higher object detection and tracking accuracy using Multi-Class Spatial Classifier method than the existing methods. This is because the Multi-Class Spatial Classifier uses MAP estimation that measures the likelihood value of the moving object. In addition, the MCSC method improves the object classification accuracy with minimum time and also reduces the false positive rate. Experimental results show that the performance improvement of the MCSC method is improved by 11% when compared to BISVM classifier [1] method and 21% improvement when compared to multi-sensor fusion framework [2], 37% when compared to existing BSFD [21] respectively.

Hence our proposed method achieves more robust and improved moving object detection and tracking accuracy.

6. CONCLUSION AND FUTURE WORKS

An efficient Multi-Class Spatial Classifier (MCSC) method is introduced to achieve efficient and robust moving object detection and tracking. In MCSC method, three processing steps are designed to obtain higher accuracy. At first, MCSC method uses Markov random field (MRF) principle to provide the label field fusion. This helps to estimates the shape of the moving and fixed object. After that, fuzzy based approach is used to identify the edge-strength of each pixel location. This helps to preserve the object boundary. Finally, the MCSC method uses MAP estimation principle to perform spatial classification. The moving objects in the video frames are correctly classified to improve the object detection and tracking accuracy. The performance of our proposed method is evaluated using various parameters like classification accuracy, false positive rate, classification time, and object detection and tracking accuracy. Experimental results reveal that the proposed MCSC method significantly improves the classification accuracy with minimum time and reduces false positive rate. The MCSC method also improves the accuracy of moving object detection and tracking compared to the existing state-of-the-art methods. When large number of objects and their shadow images in video frame are considered, the moving object detection and tracking accuracy of the proposed system may reduce considerably. Hence, the proposed work can be a combined with advanced pattern matching technique for handling large objects and their shadows to detect and track the moving object more efficiently.

Declarations

Availability of data and Materials

The dataset supporting the conclusions of this article are available in the 168VJ Clips video dataset repository and the hyperlink for this dataset is given below.

https://archive.org/details/Architects_of_Tomorrow

Competing interests

The authors declare that they have no competing interests.

Funding

This research work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] Chung-Wei Liang and Chia-Feng Juang, "Moving Object Classification Using a Combination of Static Appearance Features and Spatial and Temporal Entropy Values of Optical Flows", IEEE Transactions on Intelligent Transportation Systems, Volume 16, Issue 6, 2015, Pages 3453 – 3464

[2] Ricardo Omar Chavez-Garcia and Olivier Aycard, "Multiple Sensor Fusion and Classification for Moving Object Detection and Tracking", IEEE Transactions on Intelligent Transportation Systems, Volume 17, Issue 2, Year 2016, Pages 525 – 534

[3] Sandeep Singh Sengar and Susanta Mukhopadhyay, "Moving object area detection using normalized self adaptive optical flow", Optik - International Journal for Light and Electron Optics, Volume 127, Issue 16, August 2016, Pages 6258–6267.

[4] Mohammad Farukh Hashmi, Ritu Pal, Rajat Saxena and Avinash G. Keskar , "A new approach for real time object detection and tracking on high resolution and multi-camera surveillance videos using GPU", Journal of Central South University, Volume 23, Issue 1, January 2016, Pages 130-144

[5] Victor Vaquero, Ely Repiso, Alberto Sanfeliu, John Vissers and Maurice Kwakkernaat, "Low Cost, Robust and Real Time System for Detecting and Tracking Moving Objects to Automate Cargo Handling in Port Terminals", Advances in Intelligent Systems and Computing, Volume 418, November 2015, Pages 491-502

[6] Imen Karoui, Isabelle Quidu, and Michel Legris, "Automatic Sea-Surface Obstacle Detection and Tracking in Forward-Looking Sonar Image Sequences", IEEE Transactions on Geosciences and Remote Sensing, Volume 53, Issue 8, Year 2015, Pages 4661 – 4669.

[7] Tianxu Zhang , Gang Zhou, Chao Zhang, Gaofei Li , Jianchong Chen and Kuan Liu, "A novel temporal-spatial variable scale algorithm for detecting multiple moving objects", IEEE Transactions on Aerospace and Electronic Systems , Volume 51, Issue 1, January 2015, Pages 627 – 641

[8] Ahlem walha, Ali wali and Adel .Alimi, "Video stabilization with moving object detecting and tracking for aerial video surveillance", Multimedia tools and application, Volume 74, No. 17, 2015, Pages 6745-6767

[9] JinMin Choi, Hyung Jin Chang, Yung Jun Yoo, Jin Young Choi, "Robust moving object detection against fast illumination change", Computer Vision and Image Understanding, Volume 116, 2012, Pages 179–193

[10] Sabri M. A. A. Ahmed and Othman O Khalifa, "Moving Object Detection and Tracking in Video Surveillance System", Journal of Scientific and Engineering Research, Volume 3, Issue 6, 2016, Pages 426-432

[11] Jaebum Choi and Markus Maurer, "Local Volumetric Hybrid-Map-Based Simultaneous Localization and Mapping with Moving Object Tracking", IEEE Transactions on Intelligent Transportation Systems, Volume 17, Issue 9, 2016, Pages 2440 - 2455

[12] Min-Kook Choi, Ziyu Wang, Hyun-Gyu Lee, Sang-Chul Lee, "A bag-of-regions representation for video classification", Multimedia Tools and Applications, Springer, Volume 75, Issue 5, 2016, Pages 2453–2472

[13] Alex Pereira, Osamu Saotome and Daniel Sampaio, "Patch-based local histograms and contour estimation for static foreground classification", EURASIP Journal on Image and Video Processing, Volume 6, 2015, Pages 1-11

[14] Shunli Zhang, Xin Yu, Yao Sui, Sicong Zhao, and Li Zhang," Object Tracking with Multi-View Support Vector Machines", IEEE Transactions on Multimedia, Volume 17, Issue 3, 2015, Pages 265 – 278

[15] Samuele Salti, Andrea Cavallaro, Luigi Di Stefano, "Adaptive Appearance Modeling for Video Tracking: Survey and Evaluation", IEEE Transactions on Image Processing, Volume 21, Issue 10, 2012, Pages 4334-4348

[16] Hao Shen, Shuxiao Li, Chengfei Zhu, Hongxing Chang, Jinglan Zhang, "Moving object detection in aerial video based on spatiotemporal saliency", Chinese Journal of Aeronautics, Elsevier, Volume 26, Issue 5, October 2013, Pages 1211–1217

[17] Carlos R. del-Blanco, Fernando Jaureguizar , Narciso Garcia, "An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications", IEEE Transactions on Consumer Electronics , Volume 58, Issue 3, 2012 , Pages 857-862

[18] Liang Lin, Xiaolong Wang, Wei Yang, and Jian-Huang Lai, "Discriminatively Trained and - Or Graph Models for Object Shape Detection", IEEE Transactions on Pattern

Analysis and Machine Intelligence (TPAMI), Volume 37, Issue 5, 2015, Pages 959 – 972.

[19] Long Zhao, Li Fei Liu, Qing Yun Wang, Tie Jun Li, and Jian Hua Zhou, "Moving Target Detection and Active Tracking with a Multicamera Network", Discrete Dynamics in Nature and Society, Hindawi Publishing Corporation, Volume 2014, April 2014, Pages 1-11,

[20] Hidetomo Sakaino, "Video based Tracking, Learning, And Recognition Method For Multiple Moving Objects", IEEE Transactions on Circuits and Systems for Video Technology, Volume 23, Issue 10, 2013, Pages 1661 – 1674.

[21] D Stalin Alex and Amitabh Wahi, "BSFD: Background Subtraction Frame Difference Algorithm For Moving Object Detection And Extraction", Journal of Theoretical and Applied Information Technology, Volume 60, Issue 3, 2014, Pages 623-628.

[22] Hyojin Lim, Yeongyu Choi, Cuong Nguyen Khac, Ho-Youl Jung, "Moving Object Detection Using Cell Based Optical Flow For Static And Dynamic Scenes", Journal of Theoretical and Applied Information Technology, Volume, 95, Issues 18, 2017, Pages 4416-4426.