www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3 Issue 5 may, 2014 Page No. 5905-5912

OBJECT DETECTION AND TRACKING IN VIDEOS : A REVIEW

Nagalakshmi.C.K, Hemavathy.R , Shobha.G

Mtech, Dept of CSE, RVCE, Bangalore

Assistant Professor, Dept of CSE, RVCE, Bangalore

Prof & Head,Dept of CSE, RVCE, Bangalore

Abstract:***The detection of moving object is important in many tasks, such as video surveillance and moving object tracking. In this paper, a review has been made on a video surveillance scenario with real-time moving object detection and tracking. The design of a video surveillance system is directed on automatic identification of events of interest, especially on tracking and classification of moving objects. The object tracking and detection is used to establish a correspondence between objects or object parts in consecutive frames and to extract temporal information about objects such as trajectory, posture, speed and direction.Tracking is detecting the objects frame by frame in video. It can be used in many regions such as video surveillance, traffic monitoring and people tracking. In static environment segmentation of object is not complex. In dynamic environment due to dynamic environmental conditions such as illumination changes, shadows and waving tree branches in the wind object segmentation is a difficult and significant problem that needs to be handled well for a robust visual surveillance system.***

Keywords*: object detection, object tracking, occlusions, video surveillance*

1.INTRODUCTION

Video surveillance systems have long been use to monitor security sensitive areas. These videos provides usefulinformation that can be extracted for knowledge discovery and predictions. Detection and tracking forms a major usage in computer vision such as video surveillance, vision-based control, human-computer interfaces, medical imaging, augmented reality, and robotics. Moving object detection is the basic step for tracking the objects in video. In static environment segmentation of object is not complex. Due to dynamic environmental conditions such as illumination changes, shadows and waving tree branches in the wind object segmentation is a difficult and significant problem that needs to be handled well for a robust visual surveillance system.

Object tracking is to track an object (or multiple objects) over a sequence of images. Object tracking is defined as the process of segmenting an object of interest from a video scene and keeping track of its motion, orientation, occlusion etc. in order to extract the useful information. Difficulties in tracking objects can arise due to abrupt object motion, changing appearance patterns of both the object and the scene, non-rigid object structures, object-to-object and object-to-scene occlusions, and camera motion. Tracking is usually performed in the context of higher-level applications that require the location and/or shape of the object in every frame. Real time object tracking is recently becoming more and more important in the field of video analysis and processing. The technique of motion detection and object tracking can be applied to video surveillance system to prevent against threats.

Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. A common approach for object detection is to use information in a single frame. However, some object detection methods make use of the temporal information computed from a sequence of frames to reduce the number of false detections. This temporal information is usually in the form of frame differencing, which highlights changing regions in consecutive frames. Tracking involves registering the movements of the segmented object from initial frame to the last frame in a video.

2. LITERATURE REVIEW

Target tracking is an important topic in the computer vision and information fusion research field. The output of target tracking is the basis of many real-world applications including target classification, behavior understanding, etc. It contains large amount of information of the moving objects in each frame and is the building block for the visual surveillance system as in[1]. It is of significant value in both military and civilian applications, and has been a hot research topic for a long time. The traditional algorithms often cannot track moving objects accurately in real time, in order to overcome this problem, they proposed a new method based on wavelet features for target tracking. According to the previous information obtained by particle filter, the possible location of the target in the frame is predicted. Multi-scale two-dimensional discrete wavelet is used to characterize the possible target regions. Then the means and variances of the decomposed image are computed. Finally, Principal Component Analysis (PCA) is used to build an effective subspace. Tracking is achieved by measuring the similarity function between the target and the image regions. In addition, to combat complex background and occlusion, the characterization vector is updated based on the similarity between the object model and candidate object regions. The strength of this approach is the proposed algorithm is robust and can significantly improve the speed and accuracy of target tracking.

In this method some improvement in color based tracking has been incorporated and employed to track a moving object. The object state has been taken as the object position, speed, size, object size scale and the appearance condition of the object. The target model update condition and adaptive likelihood had been calculated to ensure the proper tracking of an object[2]. Advantage of this approach is it can handle non-rigid deformation of targets, partial occlusion and cluttered background. Disadvantage of this approach is that it does not take into account the color histogram in different regions of the target.

An automatic bandwidth adjustment method based on canny edge detection, which is focused on the defect of the fixed kernel-bandwidth of the traditional Mean-Shift object tracking algorithm and the important information of the object edge is proposed in[3]. The canny operator is calculated to extract the diagonal distances between diagonal edge points, the object changes are judged, then adjust the searching box size. Advantage of this approach is that it is efficient than traditional Mean Shift algorithm and it has a better adaptive effect. Disadvantage of this approach is that it is only applicable to the situation which the distinction between target and background is evident and less background interference.

Object detection has been carried out using twodifferent methods[4]. Background subtraction method is relatively less time consuming but it requires all the frames in advance, which is a major drawback to implement it in a real time. Considering, optical flow based method for object detection, it requires only two frames at a time to detect objects at the cost of time as it is a pyramidal approach. Another advantage of optical flow based approach is that it provides correct detection of objects in situations like temporal movement and partial background change. The Tracking algorithm is a robust approach for multiple object(s) tracking using color based probability matching in a video sequences. Advantage of this approach is that tracking is efficient when there are numerous issues such as multiple objects with same color, occlusion, merging and splitting of objects.

A method called as surf which is scale and rotation invariant and has a higher tolerance of illumination changes, it can track outdoor object[5]. A fixed template is used to get correct location of object so that it tracks occluded object and can reduce calculation of errors. Advantage of this method is that it is robust andcan track object accurately in complex environments. Disadvantage of this algorithm are: as the establish of template cache and two-stage tracking, it increases the complexity of the algorithm and reduce tracking efficiency.

A novel algorithm based on discrete cosine transform (DCT) to handle illumination variations is presented[6].Since illumination variations are mainly reflected in the low-frequency band. For instance, low illumination in a frame leads to low value DC coefficient and vice versa. The DC coefficient is modified to achieve illumination invariance. Average of DC coefficients of particular numbers of neighboring frames of current frame is taken. The correction in DC coefficient is performed using maximum eigen value of the image covariance matrix over N frames. The videos with corrected illumination are than used to track objects of interest using the Mean shift algorithm. Advantage of this approach is that it worked well for challenging videos such as for sudden and drastic changes in the illumination in video and changes in illumination condition as the targets are moving. Disadvantage of this approach is it is not adaptive for handling the scaling and orientation of the target and occlusion and illumination changes within the frame.

A unique technique for detection and tracking multiple objects from a dense cluttered area like crowd by deploying greed algorithm[7]. Understanding the complexity of deploying various image attributes e.g. edge, color etc, the proposed system will illustrate cost effective and robust procedure of using low-level attributes which takes very less computational time in order to produce autonomous rational mobility region as resultant. The technique also considers various difficult real-time scenarios in the dense crowd in order to design a highly cost effective algorithm. Performance analysis is carried out with different set of video sequences to find that proposed system has gradual robust detection rate as well as highly cost-effective computationally. The advantage of this approach is that it can be used in challenging scenarios of inter-object occlusion for object-counting as well as object localization. The disadvantage of this approach is the model recognizes two-conjoined objects as one.

Object detection in natural scene and image is playing an important role in computer vision. The traditional object detection method is more about local feature extraction and supervised learning. Using this method, the detection rate of the image with complex scenes is low. But the reality scene is complex and the real-time detection system need to handle a large number of images. In order to remedy the deficiency of the traditional test methods in the object detection, a new approach which uses patches of object and its relative position with the object's center as the feature and a new improved gentleBoost classifier, which enables it to work with better detection result is proposed[8]. In this method, linear regression stump as the weak classifier in learning algorithm, weights to the prediction model from the positive and negative classification, but not to weight it only from the positive aspect. The advantage of this approach is that it uses a new feature extraction method that the patch and it's position provides excellent performance relative to other existing feature sets. Second the method explore an addition of gentleBoost classifier to improve the accuracy in object detection. The algorithm is better than other recent methods on object detection in complex natural scene.

A novel pedestrian detection system that uses Haar-like feature ex-traction and a covariance matrix descriptor to identify the distinctions of pedestrians is proposed as in[9]. An approach that adopts an integral image is also applied to reduce the computational loads required in both the Haar-like feature extraction and evaluation of the covariance matrix descriptor. Based on the Fisher linear discriminant analysis (FLDA) classification algorithm, the proposed system can classify pedestrians efficiently. Additionally, the detection procedure of the proposed system is accelerated using a two-layer cascade of classifiers.The front end, constructed based on Haar-like features, can select candidate regions quickly wher- ever pedestrians may be present. Moreover, the back end, constructed based on the co-variance matrix descriptor, can determine accurately whether pedestrians are positioned in candidate regions. If a region tests positive through the two-layer cascade classifiers, pedestrian images are likely captured. The disadvantage of this approach is recognition accuracy degrades somewhat in dim light.

An object tracking algorithm based on surf is presented in[10]. Interest points are detected bu surf detector in reference region which is located in first frame manually. In the following the SURF features are extracted in a larger window which is selected as test region. In the matching stage we calculate the euclidean distance between descriptor vectors of interest points in the test image and ones in the reference image. The advantage of this approach is it implements real time tracking with robustness against appearance variations, scale change and cluttered regions.

An algorithm which uses the matching trajectory to realize the target tracking which based on the Gaussian mixture model is used as in [11]. Firstly use background subtraction to detect the target. Secondly, the feature of the color information, position information, shape information and modified Hough arithmetic are used to find the matching trajectory. At last, the result of target tracking is obtained. The advantage of this approach is that it has shorter matching time based on the good detection rate in target detection and tracking. Disadvantage of this approach is it can not track the targets well when the number of the targets is large.

An improved MeanShift and SIFT feature matching and calibration method for effective tracking is used in[12]. The former is used in initial localization for object tracking, which determines block method by the size of the latest enclosing rectangle and determined its weight coefficient by the Bhattacharyya coefficient of each block. The latter is used to refine the initial localization. The advantage of this method is that it efficiently addresses object tracking of an object whose motion, illumination and appearance change drastically and combinatorially. Finally the new algorithm used linear weighted method to fuse the improved MS's tracking result and SIFT's.

A method which uses Kalman filter to predict the location of object, and then introduce a CamShift-based adaptive local search method for robust tracking, using the parameter resulted from the Kalman filter is proposed in[13]. Advantage of this approach is CamShift can automatically adjust the object scale during tracking, this algorithm tracks objects effectively even with target deformation.

A method for object tracking with serious and long-time occlusion in image sequences based on occlude modeling is proposed in[14]. Occluder is modeled by detecting and evolving its rough partial contour represented by snake points, through minimizing the proposed energy function in which two novel terms are introduced: the push force and constraint force. Then the tracked object is tracked around the neighborhood of the occluder contour until the object reappears. Advantage of this approach is that the object is tracked in total and long-time occlusions. The disadvantage of this approach is if the tracking object appearance is drastically changed during occlusion the tracking precision is lost.

A new robust method for moving object tracking using an adaptive colour filter is presented in[15]. This adaptive colour tracking method is achieved by determining a unique colour of the target object comparing to the background and updating it when required. Advantage of this approach is that the method can perform robustly in tracking a moving object using a robotmounted camera in a crowded environment. Disadvantage of this approach is that it does not work well for outdoor environment.

An object tracking algorithm that learns a set of appearance models for adaptive discriminative object representation is presented[16]. The algorithm utilizes an adaptive discriminative representation to account for thenonlinear appearance change of an object over time. To reduce tracking drift, a two-stage particle filtering method is presented which makes use of both the static appearance information obtained at the outset and image observations acquired online. Advantage of this approach is that it works well under image blur, camera motion, change of pose, illumination, and scale conditions.

An approach that deals with changing backgrounds by using a quick training phase with user interaction at the beginning of an image sequence.is used. During this phase, some background clusters are learned along with object representations for those clusters. Next, for the rest of the sequence the best fitting background cluster is determined for each frame and the correspond- ing object representation is used for tracking[17]. One advantage of this approachis its ability to handle situations when an object appearsin different forms in different background. Disadvantage of this approach is that it does not take into account the bottom-up salient regions

A novel method for object tracking using global and local states of object in video surveillance application is proposed[18]. Most traditional object models using global appearance cannot handle partial occlusion effectively. The unoccluded part of partially visible object retains invariable appearance. Therefore, global and local dynamics model is introduced as object model to overcome partial occlusion using local feature, and apply it to Bayesian tracking problem using motion-based particle filtering. Advanatage of this approach is that it can reliably track pedestrian and handle partial occlusion in continuous image sequence taken from single camera in surveillance application. Disadvantage of this approach is that it does not use more features (such as *SIFT*, HOG) in this model to get better result.

A novel SIFT-Gabor Region Covariance Matrices based particle filter is proposed. First, a new SIFT-Gabor feature is presented for object representation in[19]. In this step, RCMs is used for fusing SIFT and Gabor to take advantage of the two features. Then the particle filter framework is applied for object tracking. In the tracking process, the target modelis updated automatically according to the matching resultbetween target model and candidate targets. Advantage of this approach is that it tracks the object of which illumination and scale are drastically changing, effectively, accurately and robustly. Disadvantage of this approach is it fails to track the target as the SIFT points are scarce. Although the window generated by Gabor based particle filter approximately catch the target, it is much large and does not give a precise result.

The traditional tracking methods in Unmanned Underwater Vehicles (UUVs) performpoor in controlling the distance, the value of which depends onthe random sensor observation noise. Based on the noise model ofthe sensor observation, this method proposes a tracking method tominimize the mean squared-error (MSE) of the distance, which reduces the effect of the sensor observation noise on the distance.As the optimal tracking method may be too computationallyintensive to implement in practice, we derived a suboptimaltracking method, which has a closed form[20]. Advanatage of this approach is that this tracking methods achieves much less MSE compared with the previous tracking method by keeping distance.

Marine biologists commonly use underwater videosfor their research, video analysis is typically based on visual inspection. This incurs prohibitively large user costs, and severely limits the scope of biological studies. There is a need for developing vision algorithms that can address specific needs of marine biologists, such as fine-grained categorization of fish motion patterns as in[21]. This is a difficult problem, because of very small inter-class and large intra-class differences between fish motion patterns. Their approach consists of three steps. First they apply new fish detector to identify and localize fish occurrences in each frame, under partial occlusion, and amidst dynamic texture patterns formed by whirls of sand on the sea bed. Then a tracking-by-detection is conducted. Given the similarity between fish detections, defined in terms of fish appearance and motion properties, they formulate fish tracking as transitively linking similar detections between every two consecutive frames, so as to maintain their unique track IDs. Finally, they extract histograms of fish displacements along the estimated tracks. The histograms are classified by the Random Forest technique to recognize distinct classes of fish motion patterns. The advantage of this work is that the new fish detector outperforms the shape- and part-based detectors. The fish tracking has been formulated as finding the maximum weight clique of a graph composed of pairs of fish detections from all consecutive video frames. The presented approach provides a viable solution for automated video analytics in highly specialized studies, like those in marine biology.

A new moving objects tracking algorithm is proposed as in[22] which combines improved local binary pattern texture and hue information to describe moving objects and adopts the idea of Camshift algorithm. In order to reduce matching complexity on the premise of satisfying the accuracy, many kinds of LBP and hue are cut down. The advantage of this approach is traditional Camshift algorithm is based on the colour histogram of moving objects. When an object and its corresponding background have similar colour, the illumination varies rapidly or an interferential object and the tracked object have similar colour, the tracking accuracy can be greatly affected. This approach combine LBP texture and hue to describe moving objects in order to solve these troubles.

An algorithm based on camshift and kalman filter to track moving objects in conditions such as occlusion and interference problems is proposed in[23]. In the occlusion processing, the program of linear prediction is combined with kalman filter. This approach improves robustness when moving object is under occlusion. The advantage of this method is that the system is stable for a longtime tracking moving object, even if under the condition of occlusion or camera movement. Disadvantage of this approach is that it cannot track if the object and background color is same.

The MeanShift embedded ParticleFilter tracker method produces satisfactory tracking performance and computationally inexpensive in[24]. The implementation is demonstrated with the use of color and edge information as the feature representations of the target. This tracker is capable to track abrupt direction and velocity changing object. The advantage of this approach is that it greatly reduces the number of particles required for robust tracking compared to conventional ParticleFilter approach. The disadvantage of this approach is that the tracking performance could be affected if the target accelerates at the moment it changes the moving direction as the motion behaviour would deviate from the average velocity model.

Multi-feature fusion strategy based object rough location and exact region-based object contour extraction are combined for accurate and robust object contour tracking. The object rough location is realized by color histogram and Harris corner features fusion method in particle filter framework[25]. Region based TDM is applied for exact contour detection, which is simpler than active contour models. The exact contour trackingresult can be used to guide the particle propagation of next frame,to enable more efficient particle redistributions and reducingparticle degeneration. Advantage of this approach is to re-obtain satisfactory tracking results even after a short whole object occlusion. Disadvantage of this approach is to improve on contour edge accuracy.

A method in which each object is tracked using Kalman filter in normal conditions. Also,appearance model is considered for each track, which is used at the time of occlusion. By presenting a suitable distance function and using the global nearest neighbor data association method the technique could identify segmentation errors and occlusions. In case of occlusion, with template matching by correlation coefficient method, using track appearance models and making measurement noise of Kalman filter adaptive compared to the result of template matching, could continue tracking objects[26]. Advantage of this approach is that it is robust for MOT and occlusion reasoning.

A new method that implemented the ideas ofboosting and propose a weight saturation technique to make boosting more robust to outliers is used in[27]. In order to be adaptive to the appearance variations, a new classifier is trained after the target is captured in each frame. The prior classifier and the classifier trained in the previous frame jointly predict the labels for training new classifiers. Advantage of this approach is that it outperforms in in some hostile conditions such as cluttered backgrounds, overexposure, and occlusion. Disadvantage of this approach is if features of the object and features of the background are locally indistinguishable in the feature space the discriminative tracking models will fail.

Mean-shift technique is used for initial registration, and the particle filter is used to improve the performance of tracking when the tracking result with meanshift is unconvincing[28]. RGB color histogram is exploited as image feature and Bhattacharyya coefficient is used for measuring the similarity between object model and candidate regions. Advantage of this approach is that this method works well under shift-variant, rotation and scaling, and achieves a satisfying tracking speed. Disadvantage of this approach is that color features and LBP features are not considered.

Region-based method and the contour-based method is combined called as the two-stage object tracking method for efficient object tracking. The kernel-based method, is used to locate the object in complex condition such as camera motion, partial occlusions, clutter, etc., but the tracking accuracy is not high when the object severely deforms. In order to improve the tracking accuracy, contour based method is used to track the object contour precisely after the target localization. In the first object localization stage, the initial target position is predicted and evaluated by the Kalman filter and the Bhattacharyya coefficient, respectively. In the contour evolution stage, the active contour is evolved on the basis of an object feature image generated with the color information in the initial object region. In the process, similarities of the target region are compared to ensure that the object contour evolves in the right way[29]. The advantage of this approach is it can achieve a higher tracking precision than that of the kernel-based method. The disadvantage of this approach is object feature image is based on the color information so it cannot effectively track the object when the color feature of the object is very similar to that of the background.

The absolute frame subtraction and histogram matching technique is used to detect the appearance of the object and to save computational time[30]. This method of extracting an object reduces the computation time by a factor of 10 with respect to using other methods of segmentation like N-cut. Also the processing time is saved since this algorithm starts working only when the object is detected and comes in the vicinity of the camera. The advantage of absolute frame subtraction gives better results even with low quality videos. The disadvantage of this approach is that if the object is deviated more from its path the tracking precision is decreased.

A novel visual tracking scheme, that combines the local feature point correspondences in foreground and background sets and an enhanced mean shift for global object appearance model, is used[31]. By using dynamic maintenance to both feature point sets, enhancing the mean shift by making use of local features and by dynamically updating the reference object model, as well as utilizing an optimal selection criterion for the final tracking, the results have shown very robust tracking performance. The advantage of this approach is this method tracks the objects in complex scenario such as long-term partial occlusion, intersection, large deformation, and/or complex background clutter. The disadvantage of this approach is that this method cannot track when there is full object occlusion and for large object for e.g., a face, large pose changes could potentially cause tracking degradation.

A method that combined shape and feature-based object tracking method is proposed in[32]. The adaptive background generation module embedded in this algorithm serves as a fundamental building block for robust tracking. This method is robust to the object's sudden movement or the change of features. This becomes possible by tracking both feature points and their neighbouring regions. Combination of background and shape boundary information significantly improves the tracking performance because the target object and the corresponding feature points on the boundary can be easily found. The shape control points (SCPs) are regularly distributed on the contour of the object, then comparing and updating the centroid during the tracking process is done, where straying SCPs are removed, and the tracking continues with only qualified SCPs. Advantage of this approach is that it works well under the conditions such as noisy and low-contrast environment. Disadvantage of this approach is that it does not select optimal regions for motion analysis

Non-rigid object tracking is an important issue in animation, surveillance, human–computer interaction and medical diagnosis. The difficulty of tracking is addressed on the high dimensional uncertainty and complexity embedded in the image sequences. Recently, approaches based on deformable templates have been established to solve this tracking problem. During the tracking procedure, the shape of the tracked object is represented by an ellipse. In spite of its success in some applications, the mean-shift still faces the challenge that the ellipse cannot be properly adapted if the shape or size of the tracked object varies. This can result in loss of object track of the object of interest due to dramatic intensity or colour changes in a pre-defined neighbourhood. To solve this problem a Lagrangian based method to integrate a regularising component into the covariance matrix to be computed as in[33] . Technically, we intend to reduce the residuals between the estimated probability distribution and the expected one. We argue that, by doing this, the shape of the ellipse can be properly adapted in the tracking stage. The advantage of this approach is favourable performance can be achieved in shape adaption and object localisation. The disadvantage of this approach is the proposed method has the ability of dealing with occluded images but its performance is affected by the neighbouring subject.

A study of simple algorithms for tracking objects in a video sequence, based on the selection of landmark points representative of the moving objects in the first frame of the sequence to be analyzed is proposed[34]. The movement of these points is estimated using a sparse optical-flow method. Methods of this kind are fast, but they are not very robust. Particularly, they are not able to handle the occlusion of the moving objects in the video. To improve the performance of optical flow-based methods, we propose the use of adaptive filters and neural networks to predict the expected instantaneous velocities of the objects, using the predicted velocities as indicators of the performance of the tracking algorithm. The efficiency of these strategies in handling occlusion problems are tested with a set of synthetic and real video sequences. The advantage of this approach is if the optical flow method fails to compute reasonable values for the velocities, the velocity values predicted by the filter can be used as reliable values for the velocities of the object.

Segmentation and tracking of multiple humans in crowded situations is made difficult by interobject occlusion. A model-based approach to interpret the image observations by multiple partially occluded human hypotheses in a Bayesian framework[35]. This method defines joint image likelihood for multiple humans based on the appearance of the humans, the visibility of the body obtained by occlusion reasoning, and foreground/background separation. The optimal solution is obtained by using an efficient sampling method, data-driven Markov chain Monte Carlo (DDMCMC), which uses image observations for proposal probabilities. Knowledge of various aspects, including human shape, camera model, and image cues, are integrated in one theoretically sound Framework. The advantage of this method is the integration of the top-down Bayesian formulation following the image formation process and the bottom-up features that are directly extracted from images. The integration has the benefit of both the computational efficiency of image features and the optimality of a Bayesian formulation.

The problem of tracking multiple objects poses a number of challenges due to the ambiguity of the observations and the presence of partial or complete occlusions. It introduces a novel extension to the Particle Filter algorithm for tracking multiple objects with a vision system as in[36]. This approach instantiates separate particle filters for each object and explicitly handles partial and complete occlusion for nontransparent objects, as well as the instantiation and removal of filters in case new objects enter the scene or previously tracked objects are removed. As opposed to single particle filters or mixture particle filter approaches which estimate a single multi-modal distribution, the proposed filter extension allows the continued tracking of objects through occlusion situations as well as the tracking of multiple objects of different types. To handle occlusion without an increase in computational complexity beyond the one of the Mixture Particle Filter, this approach addresses occlusions by projecting particles into the image space and back into the particle space, thus avoiding the use of a joint distribution. The advantage of using the proposed technique is that experiments performed with a color-based instantiation of the tracker demonstrated the filter's capability to track multiple targets, even when they are temporarily occluded, and to automatically generate new target filters and remove filters for targets that have been removed from the scene.

The main problem in video analysis is to track moving objects and to give description in video sequence. Occlusions is another problem to overcome while tracking. Previous approaches work on a pixel-by-pixel basis, yielding unusable in real-time situations as well as not much expressive at a semantic level of the video as in[37]. A novel method of tracking objects overcoming occlusions that exploits the information due to the spatial coherence between pixels, using a graph-based, multi-resolution representation of the moving regions is proposed.

A new method for detecting and tracking multiple moving objects based on discrete wavelet transform and identifying the moving objects by their color and spatial information is proposed in[38]. Many tracking algorithms have better performance under static background but get worse results under background with fake motions. Therefore, most of the tracking algorithms are used indoors instead of outdoor environment. Since discrete wavelet transform can divide a frame into four different frequency bands without loss of the spatial information, it is adopted to solve this problem due to the fact that most of the fake motions in the background can be decomposed into the high frequency wavelet sub-band. In tracking multiple moving objects, many applications have problems when objects pass across each other. Color and spatial information are used in this paper to solve this problem. The advantage of this approach when tracking multiple objects these objects pass across each other, color and spatial information are used to solve this problem. The disadvantage of this method is that it cannot identify the object when it enters the surveillance system. The extracted features in this case are not enough to be used to identify the moving objects.

Remotely operated vehicles (ROVs) and underwater observatories routinely record several hours of video material every day. Manual processing of such large amounts of video has become a major bottleneck for scientific research based on this data. This method develops an automated system that detects, tracks, and classifies objects that are of potential interest for human video annotators[39]. By pre-selecting salient targets for track initiation using a selective attention algorithm, the complexity of multi-target tracking is reduced. Then, if an object is tracked for several frames, a visual event is created and passed to a Bayesian classifier utilizing a Gaussian mixture model to determine the object class of the detected event.

Mean-shift tracker, which gained attention recently, is known for tracking objects in a cluttered environment. A new method to track objects by combining two well-known trackers, sum-of-squared differences (SSD) and color-based mean-shift (MS) tracker is proposed[40]. In the proposed combination, the two trackers complement each other by overcoming their respective disadvantages. The rapid model change in SSD tracker is overcome by the MS tracker module, while the inability of MS tracker to handle large displacements is circumvented by the SSD module. The performance of the combined tracker is illustrated to be better than those of the individual trackers, for tracking fast-moving objects. Since the MS tracker relies on global object parameters such as color, the performance of the tracker degrades when the object undergoes partial occlusion. To avoid adverse effects of the global model, we use MS tracker to track local object properties instead of the global ones. Further, likelihood ratio weighting is used for the SSD tracker to avoid drift during partial occlusion and to update the MS tracking modules. The advantage of this approach is able to track the object correctly without any shift even when the object got partly occluded. The disadvantage of this approach is without model update tracks the object till the end of sequence but suffers deviation from the object due to luminance/color change of the object.

A gradient image describes the differences of neighboring pixels in the image. Extracting edges only depending on a gradient image will result in innoised and broken edges. In this approach two-stage edge extraction approach with contextual-filter edge detector and multiscale edge tracker to solve the problems is proposed[41]. The edge detector detects most edges and the tracker refines the results as well as reduces the noise or blurred influence; moreover, the extracted results are nearly thinned edges which are suitable for most applications. Based on six wavelet basis functions, qualitative and quantitative comparisons with other methods show that the proposed approach extracts better edges than the other wavelet-based edge detectors and Canny detector extract. The advantage of this approach is that in edge tracking, if the starting points are lost or merged together in the coarser images, the tracking will fail. Thus this method first detect edges in the fine image and then track edges in the coarser images to fetch missed edges such that we can extract more reasonable edges, avoid the noised or blurred influence, and solve the problems of broken edges.

A robust visual object tracking algorithm based on the particle filter is presented in [42]. The method characterizes the tracked objects using color and edge orientation histogram features. To accelerate the algorithm while retaining robustness they adopt several enhancements in the algorithm. This method uses integral images for efficiently computing the color features and edge orientation histograms, which allows a large amount of particles and a better description of the targets. Next, the observation likelihood based on multiple features is computed in a coarse-to-fine manner, which allows the computation to quickly focus on the more promising regions. Quasi-random sampling of the particles allows the filter to achieve a higher convergence rate. The advantage of using the proposed algorithm is it maintains multiple hypotheses and offers robustness against clutter or short period occlusions.

A new method for object tracking in image sequences using template matching is proposed as in[43]. To update the template, appearance features are smoothed temporally by robust Kalman filters, one to each pixel. The resistance of the resulting template to partial occlusions enables the accurate detection and handling of more severe occlusions. Abrupt changes of lighting conditions can also be handled, especially when photometric invariant color features are used. The method has only a few parameters and is computationally fast enough to track objects in real time. The advantage of this approach is that it works well under changes of object orientation and lighting, zooms, partial occlusions, cluttered background and severe occlusions, abrupt changes of the lighting conditions. The disadvantage of this approach is that it does not work foe long-time occlusions.

Tracking objects underwater is a very hard task because of the hostile environment that the water presents. Many parameters should be considered in order to lessen the effect of the hostility. Such parameters underwater are not considered in the vision tracking and/or positioning of objects underwater as long as the image taken is not that distorted. This study proposes an image-based positioning system using neural network for colored objects submerged underwater[44]. The sample data for the Artificial Neural Network model is gathered by empirical methods using actual experimental set-up. The neural network is represented by the following variables: HSI components as input and the coordinates of each colored objects as outputs.

3. CONCLUSION

In this paper an analysis is presented for different image processing techniques which can employed for object detection and tracking in videos.

REFERENCES

[1] Ting Rui a,n, QiZhang a, YouZhou b, JianchunXing a, "*Object tracking using particle filter in the waveletsubspace*", Neurocomputing vol. 119,pp. 125–130,2013

[2] Ashwani Kumar, Sudhanshu kumar Mishra, Pranjna parimita Dash, "*Robust Detection &Tracking of Object by Particle Filter using Color Information*", 4th ICCCNT Tiruchengode, India, IEEE, July 4-6, 2013

[3] Le Zhang, Danhong Zhang, Yixin Su, Fei Long, "*Adaptive Kernel-bandwidth Object Tracking Based on Mean-shift Algorithm*", Fourth International Conference on Intelligent Control and Information Processing (ICICIP) Beijing, China, pp. 413-416, June 9 – 11, 2013

[4] Deval Jansari, Shankar Parmar and Geetali Saha, "*Real-Time Object Tracking Using Color-Based Probability Matching*", IEEE, 2013

[5] ZHOU Dan, HU Dong, "*A Robust Object Tracking Algorithm Based on SURF*", IEEE, 2013

[6] Gargi Phadke, Rajbabu Velmurgan, "*Illumination Invariant Mean-Shift Tracking*", IEEE, pp.407-412,2013

[7] Pushpa D.1, Dr. H.S. Sheshadri2, " *Precise multiple object identification and tracking using efficient visual attributes in dense crowded scene with regions of rational movement*", IJCSI Vol. 9, Issue 2, No 2, March 2012

[8]Li Guoa, Yu Liaoa, Daisheng Luob, Honghua Liaoa, "*Generic Object Detection Using Improved Gentleboost Classifier*", Physics Procedia, vol. 25,pp: 1528 – 1535,2012

[9]Ying-Che Kuo∗, Zih-Yi Yang, Chih-Hung Yen, "*Fastpedestrian detection system with a two layer cascade of classifiers*", Computers and Mathematics with Applications, vol. 64,pp. 1311–1323,2012

[10] Hu Shuo, Wu Na, Song huajun, "*Object tracking method based on SURF",*AASRI conference on modelling, Identification and control, 2012

[11] Min Huanga, Gang Chena, Guo-feng Yanga ,Rui Caoa, "*An Algorithm of the Target Detection and Tracking of the Video"*, International Workshop on Information and Electronics Engineering, Procedia Engineering, vol. 29,pp. 2567 – 2571, 2012

[12] Kai Du, Yongfeng Ju, Yinli Jin, Gang Li, Yanyan Li, Shenglong Qian, "*Object Tracking Based on Improved MeanShift and SIFT*", pp. 2716-2719, IEEE, 2012

[13] Xin Chen, Xiang Li, Hefeng Wu, and Taisheng Qiu, "*Real-time Object Tracking via CamShift-Based Robust Framework*", IEEE International Conference on Information Science and Technology ,Wuhan, Hubei, China, pp. 527-530, March 23-25, 2012

[14] Peng Wang, Wanyi Li, Wenjun Zhu and Hong Qiao, "*Object Tracking with SeriousOcclusion Based on Occluder Modeling*", Proceedings of IEEE International Conference on Mechatronics and Automation, Chengdu, China, pp. 1960-1965, August 5 – 8, 2012

[15]Feng Su and Gu Fang, "*Moving Object Tracking Using an Adaptive Colour Filter*", 12th International Conference on Control, Automation, Robotics & Vision Guangzhou, China, pp. 1048-1052, 5-7th December 2012

[16] Qing Wang, Wenli Xu, "*Object Tracking via Partial Least Squares Analysis*", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 10, pp. 4454-4456,OCTOBER 2012

[17] Ali Borji, Simone Frintrop, Dicky N, SihiteLaurent Itti, *"Adaptive Object Tracking by Learning Background Context*", IEEE, 2012

[18] NING WANG, JIN-HAI XIANG, WEI-PING SUN, JING-LI ZHOU, "*Object Tracking With Global And Local Dynamics Model*", Proceedings of the 2012 International Conference on Wavelet Analysis and Pattern Recognition, Xian, pp. 233-237,15-17 July, 2012

[19] Xinying Liu, "*Particle Filter Object Tracking Based on*

SIFT-Gabor Region Covariance Matrices", IEEE, pp. 201-204, 2012

[20] Xinyuan Wang, "*New Tracking Methods of Unmanned Underwater Vehicles (UUV*s)", IEEE, pp. 806-810, 2012

[21] Mohamed Amer1, Emil Bilgazyev2, "*Fine-grained Categorization of Fish Motion Patterns in Underwater Videos"*, IEEE International Conference on Computer Vision Workshops ,2011

[22] Guo-wu YUAN, Yun GAO, Dan XU, "*Improved local binary pattern texture and hue information to describe moving objects and adopts Camshift algorithm*", Procedia Engineering, vol. 15,pp. 3964 – 3968, 2011

[23] Shengluan Huang, Jingxin Hong, "*Moving Object Tracking System Based On CamshiftAnd Kalman Filter*", IEEE, pp: 1423-1426, 2011

[24] Y.S. Chia W.Y. Kow W.L. Khong A. Kiring K.T.K. Teo, "*Kernel-Based Object Tracking via Particle Filter and Mean Shift Algorithm*", IEEE, pp:522-527, 2011

[25] Xiaofeng Lu1,3, Li Song1,2, Songyu Yu1, Nam Ling2, "*Object Contour Tracking Using Multi-feature Fusion based Particle Filter*", pp. 237-242, IEEE,2011

[26] Mohammad Azari , Ahmad Seyf, Amir Hossein Rezaie, "*Real Time Multiple ObjectTracking and Occlusion Reasoning Using Adaptive Kalman Filters*", IEEE, 2011

[27] Ren C. Luo and Ching C. Kao, "*Object Tracking-by-Detection under Cluttered Environments Based on a Discriminative Approach*", IEEE, pp. 928-933, 2011

[28] Da Tang, Yu-Jin Zhang, "*Combining Mean-shift and Particle Filter for Object Tracking*", Sixth International Conference on Image and Graphics, pp.722-776, 2011

[29] Qiang Chen, Quan-Sen Sun, Pheng Ann Heng, De-Shen Xia, "*Two-Stage Object Tracking Method Based on Kernel and Active Contour*", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 20, NO. 4, pp. 605-609, APRIL 2010

[30] Madhur Mehta, Chandni Goyal, M.C. Srivastava, R.C. Jain, "*Real Time Object Detection and Tracking: Histogram Matching and Kalman Filter*", IEEE, pp.796-801, 2010

[31] Zulfiqar Hasan Khan, Irene Yu-Hua Gu*, "Joint Feature Correspondences and Appearance Similarity for Robust Visual ObjectTracking*",IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 5, NO. 3, pp. 591-606, SEPTEMBER 2010

[32] T. Kim, S. Lee, J. Paik, "*Combined shape and feature-based video analysis and its application to non-rigid object tracking*", Published in IET Image Processing, pp. 87-100, 2009

[33] Huiyu Zhoua, Yuan Yuanb, Yi Zhangc, Chunmei Shid*," Non-rigid object tracking in complex scenes*", Pattern Recognition Letters, vol. 30, pp.98–102, 2009

[34] Eduardo Parrilla, Damián Ginestar, José Luis Hueso, Jaime Riera, Juan Ramón Torregrosa, "*Handling occlusion in optical flow algorithms for object tracking*", Computers and Mathematics with Applications,vol. 56 , pp. 733–742,2008

[35]Tao Zhao, Member, IEEE, Ram Nevatia, Fellow, IEEE, and Bo Wu, Student Member, IEEE, "*Segmentation and Tracking of Multiple Humans in Crowded Environments*", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 30, NO. 7, JULY 2008

[36] HwangRyol Ryu and Manfred Huber, "Particle *Filter Approach for Multi-Target Tracking*", EEE/RSJ International Conference on Intelligent Robots and Systems, Oct 29 - Nov 2, 2007

[37] Donatello Contea, Pasquale Foggiab, Jean-Michel Jolionc, Mario Vento*,"graph-based, multi-resolution algorithm for tracking objects in presence of occlusions",* Pattern Recognition vol. 39,pp. 562–572,2006

[38] Fang-Hsuan Cheng,Yu-Liang Chen, "Real time multiple objects tracking and identification based On discrete wavelet transform", Pattern Recognition, vol. 39,pp. 1126–1139,2006

[39]Danelle E. Cline, Daniel Davis, Ishbel Kerkez, Jerome Mariettecx, "*Detecting, Tracking and Classifying Animals in Underwater Video*",IEEE,2006

[40] R.Venkatesh Babua, Patrick Pe´rezb, Patrick Bouthemy, "*Robust tracking with motion estimation and local Kernel-based color modeling*", Image and Vision Computing, 2006

[41] Ming-Yu Shih, Din-Chang Tseng, " *A wavelet-based multi resolution edge detection and tracking*", Image and Vision Computing, vol. 23, pp. 441–451,2005

[42] Ramani Duraiswami and Larry Davi,"*Fast Multiple Object Tracking via a Hierarchical Particle Filter Changjiang Yang",IEEE* International Conference on Computer Vision, 2005

[43] Hieu T. Nguyen andArnold W.M. Smeulders, "*Fast Occluded Object Tracking by a Robust Appearance Filter*", IEEE transactions on pattern analysis and machine intelligence, VOL. 26, NO. 8, AUGUST 2004

[44] Cyrus M. Delos Santos1, Elmer P. Dadios2, "Artificial Neural Network Color-based Positioning System for Multiple Objects Underwater", unpublished