An Efficient Object Detection and Tracking System Based on Immune Genetic Algorithm

Madhavi Gajula and A. Jhansi Rani

1 Research Scholar, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India
2 Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

Abstract: Particle filter algorithm is widely used for target tracking using video sequences which is of great importance for intelligent surveillance applications. However, there is still much room for improvement, e.g., the so-called "sample impoverishment". A novel algorithm, the Immune Genetic Algorithm (IGA) is proposed based on the theory of immunity in biology which mainly constructs an immune operator accomplished by two steps: a vaccination and an immune selection. In our proposed methodology a sample video is given as input for the purpose of object tracking. The steps to be carried out in our proposed method are given: motion segmentation, noise removal and object tracking. Motion segmentation involves the generation of background model for the purpose of segmentation. Noise removal involves the particle filters. The main process that takes place in the particle filter is re-sampling. Then the object was tracked by using immune Genetic algorithm. Immune Genetic algorithm involves vaccination and the immune selection process including the process of ordinary Genetic algorithm. This efficiently involves the multiple object detection.

Key words: Particle filter algorithm, immune Genetic algorithm, motion segmentation, noise removal, re-sampling, efficiently

INTRODUCTION

Video surveillance is an important research area in the commercial sector. Technology has reached a stage where mounting cameras to capture video imagery is cheap but finding available human resources to sit and watch that imagery is expensive. Surveillance cameras are already prevalent in commercial establishments with camera output being recorded to tapes that are either rewritten periodically or stored in video archives (Youssef et al., 2010). One of the goals in video surveillance is to detect and track all the relevant moving objects in the scene and to generate exactly one track per object. This may involve detecting the moving objects, tracking them while they are visible and re-acquiring the objects once they emerge from an occlusion to maintain identity. This is a very difficult problem, even more, so, when the sensor is moving in aerial surveillance scenarios (Lin et al., 2011).

The motion analysis is a basis of all intelligent video surveillance technologies. In particular, it provides the fundamentals for automatic detection and tracking of moving objects and automatic detection of new or disappeared objects of observed scene. It is the well-studied area of computer vision including many different techniques (Manohar et al., 2006). Object detection and tracking focuses on detecting the position of a moving object in a video sequence. It is the first step accomplished by a event recognition system that extracts semantic content from video (Vishnyakov et al., 2012). The purpose of detecting moving objects in video stream is to be able to track the objects over time and derive a set of properties from their trajectory such as their behaviours (Cohen and Medioni, 1999). A good detection measure should capture the performance in terms of both overall detection (number of objects detected, missed determines and false alarms) and goodness of detection for the detected objects, i.e., spatial accuracy (how much of the ground truth is detected) and spatial fragmentation (object splits and object merges) (Manohar et al., 2006). In general, there are two types of approaches in motion detection: the region-based approach and the boundary-based approach (Yokoyama and Poggio, 2005). Other approaches to object detection for surveillance, including feature-based object detection, template-based object detection and background subtraction or inter-frame difference-based detection (Huang et al., 2008).

The goal of the tracking stage is to use the observations to generate the tracks that correspond to the moving objects. Some of the issues that make this a challenging problem include missed observations due to poor video quality, noisy observations due to changes in
illumination, occlusions, objects that are moving in close proximity to each other, etc. (Kamath et al., 2005). Commonly used approaches for tracking are token-based when a geometric description of the object is available or intensity-based (optical flow, correlation). These techniques are not appropriate for blob tracking, since, a reliable geometric description of the blobs cannot be inferred. On the other hand, intensity-based techniques ignore the geometric description of the blob (Cohen and Medioni, 1999). The tracking-based methods encounter the problems of merging, splitting, entering, leaving, occlusion and correspondence. These problems are not easy to solve in many cases. And it is difficult to track all the objects precisely in crowded situations (Zin et al., 2012). To track from a moving camera, we need to describe the motion of moving objects in common coordinates. The mosaic space which is derived from the image coordinates of one frame is usually selected as the tracking coordinates. Without further refinement, accumulated errors are inevitable if fixed coordinates are selected. More importantly, an object motion in image coordinates is not physically meaningful (Lin et al., 2011).

A typical surveillance application consists of three building blocks, responsible of: moving object detection, object tracking and higher level motion analysis. Applications include car and pedestrian traffic monitoring, human activity surveillance for unusual activity detection, people counting, etc. (Jasani et al., 2005). Detection of moving objects in video streams is the first relevant step of information extraction in many computer vision applications including video surveillance, people tracking, traffic monitoring and semantic annotation of videos. In these applications, robust tracking of objects in the scene calls for a reliable and effective moving object detection that should be characterized by some important features: high precision with the two meanings of accuracy in shape detection and reactivity to changes in time; flexibility in different scenarios (indoor, outdoor) or different light conditions and efficiency, in order for detection to be provided in real-time (Cucchiara et al., 2003). Detecting and tracking moving objects are widely used as low-level tasks of computer vision applications such as video surveillance, robotics, authentication systems, user interfaces by gestures and a pre-stage of MPEG4 image compression. Software development of low-level tasks is especially important because it influences the performance of all higher levels of various applications (Yukooyama and Poggio, 2005).

**Literature review:** Video surveillance is defined as the task of analysing video sequences to detect abnormal and unusual activities. Moving object detection is the basic step for further analysis of video. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. Object tracking is the process to track the object over the time by locating its position in every frame of the video in surveillance system. A brief review of some recent researches is presented here.

Khassman (2008) has presented an automatic Moving Object Detection, Extraction and Recognition system (aMODERs). The system received still images captured every 2 sec from a surveillance camera which monitors a restricted zone such as international border crossings, buffer zone in conflict areas or any monitored area where movement across that area is to be detected. The potential objects which have been considered in his work were classified into three groups: human (female, male), animals (goat) and vehicles (motorbike, car, loader, jeep). The use of (aMODERs) in a real life application was not limited to those objects because the neural network classifier could be quickly and easily trained to recognize as many different objects as required. The system provided solutions to the problem of monitoring secured areas; Namely, the detection of movement across the area, the extraction of the moving object and the recognition of the object.

Patel (2012) has formed a series of algorithm, to track the feature of motion detection under surveillance system. A pixel variant played a vital role in detection of moving frame of a particular clip. If there was a little bit motion in a file then it was detected very easily by tracking pixel variance. The algorithm detected the zero variation only when there was no motion in a frame clip. It was simple and easier for motion detection in the frames of moving object having Avi file. There were different levels at which tracking can be performed. At the highest level, the whole body was tracked without paying attention to the details of the posture and limbs. At a lower level, the posture and limbs were tracked. At an even lower level, one or two parts of the body (such as hands) were tracked. The finest level would be tracking the fingers of a hand or facial features. The scope of his study was dealt with the highest level of tracking (tracking the whole body as one unit). His study answered the crucial question regarding the tracking of an object and suggested the requirement of surveillances in the terms of security and in high tech world.

Video surveillance is gaining its importance in almost every field of day to day life. Surveillance is being done not only in military areas or airports but also in offices, schools, shopping areas old age home and many more areas. Sharma (2012) has presented a general review on the overall process of a surveillance system used in
the present time. The processing framework of the video surveillance system included the following stages: moving object detection, object segmentation, representation, classification, tracking of objects, activity recognition and prediction. The video surveillance system could be classified in a number of areas considering the environment in which they operate: indoor, outdoor or airborne.

Aliabadian et al. (2012) have proposed and tested a modified object-tracking algorithm that uses the flexible metric distance transform kernel and multiple features for the mean shift procedure. The faithful target separation based on RGB joint pdf of the target region and that of a neighborhood surrounding the object was obtained. The non-linear log-likelihood function mapped the multimodal object/background distribution as positive values for colors associated with foreground while, negative values were marked for background. This replaced the more usual Epanechnikov kernel (E-kernel), improving target representation and localization without increasing the processing time, minimizing the similarity measure using the Bhattacharya coefficient.

Pande et al. (2012) have proposed a new method to detect moving object based on background subtraction. First of all, they have established a reliable background updating model based on statistical and used a dynamic optimization threshold method to obtain a more complete moving object. And then, morphological filtering was introduced to eliminate the noise and solve the background disturbance problem. At last, contour projection analysis was combined with the shape analysis to remove the effect of shadow the moving human bodies were accurately and reliably detected. In a single static camera condition, they combined dynamic background modelling with dynamic threshold. Selection method based on the background subtraction and up-date background on the basis of accurate detection of object, the method was effective to enhance the effect of moving object detection.

Joshi and Thakore (2012) have presented a survey of various techniques related to video surveillance system improving the security. The goal of their study was to review of various moving object detection and object tracking methods. They have focused on detection of moving objects in video surveillance system then tracking the detected objects in the scene. Moving object detection was first low level important task for any video surveillance application. Detection of moving object was a challenging task. Tracking was required in higher level applications that require the location and shape of object in every frame. They have described background subtraction with alpha, temporal differencing, statistical methods. They gave valuable insight into the important research topic and encouraged the new research in the area of moving object detection as well as in the field of computer vision. Their research on object tracking could be classified as point tracking, kernel tracking and contour tracking according to the representation method of a target object.

Chih-Hsien and Jen-Shiun (2012) have presented a new approach, Direct LL-mask Band Scheme (DLLBS) for the detection and tracking of moving objects using a low resolution image. Moving object detection was an important basic task for intelligent video surveillance systems because it provided a focus of attention for post-processing. However, the successful detection of moving objects in a real environment was a difficult task, due to noise caused by false motion such as the motion of leaves in trees. Many methods have been developed in constrained environments for the detection and tracking of moving objects. The DLLBS method could effectively reduce the noise with low computing cost, in both indoor and outdoor environments. For circumstances where occlusions occur, they have proposed a new approach, Characteristic Point Recognition (CPR). Together with DLLBS and CPR, the problems associated with occlusions were alleviated.

Ritika and Selvon (2012) have discussed an efficient and effective approach for identifying and tracking of moving object from a video. A video was captured by stationary camera. Moving object tracking and detection from video sequences has applications in several areas such as automatic video surveillance, motion-based recognition, video indexing, human-computer interaction, traffic monitoring and vehicle navigation. They have presented a computer vision-based approach for object tracking and detection. A method was proposed to detect and track moving object through video even if background was changed at any instant and capable of plotting a 3D graph mesh based on the moving object in between any number of frames per second. They have used consecutive frame analysis technique to detect background changing criteria and use morphological filtering for image enhancement. Finally, they got the co-ordinates for the moving object and these co-ordinates are imported to any other 3D Software's like MAYA, etc., to analyse or edit the results calculated by the algorithm.

Zhang et al. (2012) have proposed a two-step accurate moving target detection method. Their method obtained a rough target area using a colour background subtraction method with a feedback to background estimation, followed by a modified SUSAN method to estimate the accurate target edge. In the first step, they used the colour background subtraction method and morphology method to detect the rough area of target. In the second step, a modified SUSAN method was proposed to detect the accurate edge of target in the
rough area. In the case of noisy video frames, some pixels are recognized as background, holes appear in the estimation results. Holes need to be filled to get the whole target area. After filling the holes, we get the accurate area of moving target. The background area is returned to background estimation as a feedback to update the background. The modified SUSAN method used the grads magnitudes to achieve an adaptive threshold and the grey barycentre criterion for denoising.

The surveillance detection related events from video input is a sophisticated technology in real time security related applications. Many current and existing solutions to this problem are simply slight variations on frame differencing concept. This proves to be difficult to configure and operate effectively. Sahiti and Mutiyalo (2012) have presented a new approach based on extracting and classifying the background contents of each video frame using a CAM equipped with a standard frame grabber. The action of objects classified as people was further categorized into a series of events such as person leaving, entering, deposits object and so forth. A backround model was used to obtain candidate surveillance objects from input video. The focus of their research was to detect and predict anomalous behaviour of individuals in an environment using a camera as a way to capture the information about the environment.

MATERIALS AND METHODS

Problem definition: One of the critical tasks in computer vision is detection and tracking of objects. But still now, the issues related to this are developing. For the automatic detection of moving objects, some of the monitoring systems cannot able to find the difference when the difference of brightness between the background and the moving objects is small. The costs are very high in these systems. Most of the previous methods, only concentrated on detecting rough area of targets. The accurate moving target detection cannot be achieved. It makes the result to be shown with noise and heal and the computation time is also increased. Many systems are unable to solve critical solutions such as partial occlusions and cross targets. Even if the partial occlusions happened because of other objects or the crossing of other moving objects, the systems cannot have the ability to keep tracking its target. The object tracking can be complex, due to camera motion, loss of information caused by the projection of the 3D world on a 2D image, variations of target scale, partial occlusions, real-time processing requirements, clutter and so on. Therefore, it is difficult to make the efficient tracker and also, we cannot track the object accurately.

RESULTS AND DISCUSSION

Video tracking of objects using immune Genetic algorithm: Video tracking is the process of locating a moving object (or multiple objects) over time using a camera. It has a variety of uses, some of which are: human-computer interaction, security and surveillance, video communication and compression, augmented reality, traffic control, medical imaging and video editing. Video tracking can be a time consuming process due to the amount of data that is contained in video. Adding further to the complexity is the possible need to use object recognition techniques for tracking. Particle filter algorithm is widely used in single object tracking of video sequences in the video surveillance systems. The greatest disadvantage of the particle filter is the particle degradation. In our research multiple object tracking is done. In order to overcome this disadvantage we have proposed immune Genetic algorithm. The particle diversity is increased by adding an evolutionary particle filter with the immune Genetic algorithm. The particle diversity refers to the number the particle increased in front of the re-sampling process. Particles were referred to as the antibodies of the immune system. New antibodies were generated using the crossover and mutation process of the immune Genetic algorithm. Regulatory mechanisms of antibodies, such as promotion and suppression, ensure the diversity of the particle set. In our proposed algorithm the particles optimized by the immune Genetic algorithm provides high trusted information for about the target. The effectiveness and robustness of the proposed particle filter are verified by target tracking experiments.

In our proposed methodology a sample video is given as input for the purpose of object tracking. The steps to be carried out in our proposed method are given as:

- Motion segmentation
- Noise removal
- Object tracking

Motion segmentation: For motion segmentation process, average background model is used (Fig. 1). Firstly, the background model was generated for the frames. Then the current frame is subtracted from the background model, so that, the threshold difference can be achieved:

\[
\text{bg}^x(y, x) = \frac{1}{N} \sum_{j=1}^{N} I^x(j, y, x)
\]  

(1)
Fig. 1: a, b) Original frame

Where:
\[ b(x, y) = \text{The background values of pixels in the frame} \]
\[ N = \text{Total Number of frames} \]
\[ I(x, y) = \text{Pixel values of frames} \]

Noise removal: Noise removal of the frames was done by using particle filters. A particle filter, also known as a Sequential Monte Carlo method (SMC) is a sophisticated model estimation technique based on simulation. Monte Carlo methods (or Monte Carlo experiments) are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results, i.e., by running simulations many times over in order to calculate those same probabilities heuristically just like actually playing and recording your results in a real casino situation, hence, the name. Particle filters are usually used to estimate Bayesian Models in which the latent variables are connected in a Markov chain which is similar to a Hidden Markov Model (HMM) but typically where the state space of the latent variables is continuous rather than discrete and not sufficiently restricted to make exact inference tractable (as, for example, in a linear dynamical system where the state space of the latent variables is restricted to Gaussian distributions and hence, exact inference can be done efficiently using a Kalman filter) (Fig. 2 and 3). The steps in the particle filter algorithm.

Initialization: Select an object initially in a reference frame and for that frame.

Fig. 2: Architecture of our proposed methodology

Particle state transition: Develop a new particle set using random drift model.

Particle weight calculation: The weight of each particle was calculated.

MMSE estimation: A Minimum Mean Square Error (MMSE) estimator is an estimation method which minimizes the Mean Square Error (MSE). Here, MMSE was applied to the targets.

Re-sampling process: Changing the pixel dimensions of an image is called re-sampling. When you resample up or increase the number of pixels in your image, new pixels are added based on colour values of existing pixels. The particles were re-sampled according to the particle weight and new particle set was obtained. The noise removal process using the particle filter was explained in Fig. 4. The steps are repeated until the noise in the segmented image was removed.

Object tracking: Object tracking was done by using immune Genetic algorithm. Genetic algorithm is a family of computational models based on principles of evolution.
Fig. 3: Architecture for the particle filter algorithm process

Fig. 4: Noise removal process

and natural selection. These algorithms convert the problem in a specific domain into a model by using a chromosome-like data structure and evolve the chromosomes using selection, crossover and mutation operators. In immune Genetic algorithm an immune operator is used. The immune operator consists of vaccination and the immune selection process. The new particle set was optimized using immune Genetic algorithm (Fig. 5).

**Selection process:** Selection is the first stage in Genetic algorithm. In the selection process the MMSE estimated output is given as input. In the selection process the fitness function is calculated (Fig. 6). The fitness function was calculated using Eq. 2:

$$ F(i) = \frac{1}{\epsilon_i} $$  \hspace{1cm} (2) 

Here, the fitness function is the inverse of the root mean square error between the state estimate and the true state:

$$ \epsilon_k = \left| X_k - \hat{X}_k \right| $$  \hspace{1cm} (3)
Fig. 7: Mutation the chromosome

**Crossover:** It is a Genetic operator used to vary the programming of a chromosome or chromosomes from one generation to the next. It is analogous to reproduction and biological crossover, upon which Genetic algorithms are based. Crossover is a process of taking more than one parent solutions and producing a child solution from them. The best solutions were selected using the fitness function calculation. Crossover is calculated by using Eq. 4:

\[
\begin{align*}
  x_1' &= r x_1 + (1-r) x_2 \\
  x_2' &= (1-r) x_1 + r x_2
\end{align*}
\]

where, \( r \), random number generated from interval \([0, 1]\).

**Mutation:** Mutation is the final stage in Genetic algorithm. Mutation is a Genetic operator which is used to maintain Genetic diversity from one generation of a population of algorithm chromosomes to the next. A Genetic operator is an operator used in Genetic algorithms to maintain Genetic diversity. In mutation the chromosome value remains the same. Mutation is calculated by using Eq. 5 (Fig. 7-9):

\[
\begin{cases}
  x = x + \Delta \left[ g_x \cdot r(k) - x \right], \text{ sign } = 0 \\
  x = x + \Delta \left[ g_x \cdot x - l(k) \right], \text{ sign } = 1
\end{cases}
\]

In this study Gaussian mutation is taken as the mutation. Mutation in binary string.

**Vaccination:** Vaccination means modifying the genes on some bits in accordance with priori knowledge, so as to gain higher fitness with greater probability. This operation must satisfy the following two conditions.

Firstly, if the information on each gene bit of an individual \( y \) is wrong, i.e., each bit of it is different from that of the optimal one and then the probability of transforming from \( x \) to \( y \) is 0.

Secondly, if the information on each gene bit of \( x \) is right, i.e., \( x \) is the optimal one, then the probability of transforming from \( x \)-\( x \) is 1.
Immune selection: Immune selection was done by using following two steps:

Immune test: If the fitness is smaller than that of the parent which means serious degeneration must have happened in the process of crossover or mutation and then instead of the individual the parent will participate in the next competition.

Annealing: Annealing is selecting an individual in the parents. After that reproduction is done and the process is repeated all the individuals were processed.

CONCLUSION

In this study, we propose a new evolutionary particle filter with IGA for video tracking. The study of new algorithm focuses on the sample impoverishment brought by re-sampling. The input video is motion segmented to track the motion of the object. Noise removal was done by means of the particle filter. After the noise was removed the object was tracked by means of immune Genetic algorithm. Immune Genetic algorithm improves the fitness value, so that objects can be tracked with high efficiency.

RECOMENDATIONS

In future from multiple videos multiple objects can be tracked as a development. Immune Genetic algorithm can be efficiently used for this purpose. This will become important for the purpose of particle diversity.

REFERENCES


