International Journal of Research in Advent Technology (IJRAT)
(E-ISSN: 2321-9637)
Special Issue National Conference "CONVERGENCE 2017", 09th April 2017

Video Frame Segmentation and Object Tracking using ANN for Surveillance System

Ms. Priya Chaudhari
Department of Electronics & Telecommunication
J. T. Mahajan COE Faizpur

Prof. K. S. Bhagat
Department of Electronics & Telecommunication
J. T. Mahajan COE Faizpur

Abstract- This survey describes the current state-of-the-art in the development of automated visual surveillance systems so as to provide researchers in the field with a summary of progress achieved to date and to identify areas where further research is needed. The ability to recognise objects and humans, to describe their actions and interactions from information acquired by sensors is essential for automated visual surveillance. The increasing need for intelligent visual surveillance in commercial, law enforcement and military applications makes automated visual surveillance systems one of the main current application domains in computer vision. The emphasis of this review is on discussion of the creation of intelligent distributed automated surveillance systems the survey concludes with a discussion of possible future directions.

1. INTRODUCTION

Now a day‟s computer vision has been applied to every organisation. Such that the all in security systems, computers are widely used regarding to this for the security purpose every organisation are used different monitoring system i.e. surveillance system, suspicious monitoring system etc. [1]. In practical application, since the camera moves and rotates, it needs to track objects in a dynamical background. In this situation the visual object tracking is one of the most important parts of the video processing. In video processing system there are various problems occurred due to some insufficient configuration of hardware as well as software such as How to select the initial target objects in video footage automatically and establish objects motion model, and how to update object and background models at each frame are the key in real-time Video object tracking with an active camera, some cases due to limitations of hardware of a CCTV cameras the video there are too much noise in the footage, video footage the most difficult task is to track the object or segment because in the footage the object in moving condition that‟s why the captured image of the particular object is unable to visualize in cleared form due to its background.[1]

To resolve this problem the various algorithms are published for tracking and segmentthe video objects efficiently such as Threshold decision and diffusion distance, partial list square analysis, CAMSHIFT, etc. [2]. Such as Continuously Adaptive Mean Shift algorithm (CAMSHIFT) [2] is a popular algorithm for visual tracking, providing speed and robustness with minimal training and computational cost. An adaptive robust object tracking algorithm based on active camera is proposed. At the other algorithm Partial least squares analysis is a statistical method for modelling relations between sets of variables via some latent quantities. In PLS analysis, the observed data is assumed to be generated by a process driven by a small number of latent variables. In Threshold Decision and Diffusion Distance algorithm [7] video object segmentation and tracking framework for smart visual surveillance cameras is proposed with two major contributions. First, we propose a robust threshold decision algorithm for video object segmentation with a multi background model. Second, we propose a video object tracking framework based on a particle filter with diffusion distance (DD) [7] for measuring colour histogram similarity and motion clues from video object segmentation. For better tracking of no rigid objects, we include colour histogram in our object model as it is more stable for no rigid moving objects.

2. LITERATURE REVIEW

To implement this object tracking system are used various algorithms. In this system object tracking is most important field in surveillance systems. In surveillance system cameras capture the footage for tracking suspicious movement in organisation, in this condition the videos prepare with the help of surveillance cameras the most difficult task is to tracking the object from the video and make the another image so that image should be vague to identification. Generally the surveillance system work in client server architecture, at the client side, video is captured by surveillance cameras [6].Such cameras can be either analogue or digital. Digital camera has become more and more popular, mainly because the captured video by digital surveillance camerasis easier to track and analyse with object detection and content analysis tools. The captured video is sent to the server for further processing. At the server side, video data is used for object detection as well as tracking. The working flow of the surveillance system shown in fig2.1.[6]

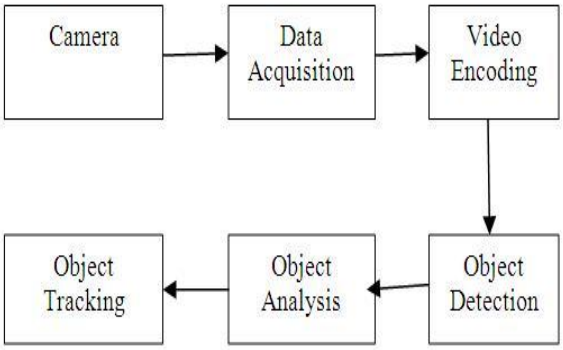


Fig.2.1 working flow of surveillance system

To tracking object, a method is implement which is object tracking station. OTS Stations implement the camera level and the scene level intelligent functions, like object tracking based on multiple camera views, shape classification, and detection of crossing perimeters of virtual zones. [1] Results of the OTS calculations are collected and processed in a central **Site Wide ObjectTracking Server**. There is another problem is toobject tracking that is nothing but the background just because of background the colour of object and their background are same in this situation the system is unable track the object to avoid this problem one solution is applied that is the segmentation. Due to segmentation the background of the image are segmented in number of images so that objects are free form background.

2.1 Working Process of Video Processing

Object Tracking stations Receive analogue or IP camera images directly and perform the most time consuming video content analysis and image processing tasks. OTS Stations implement the camera level and the scene level intelligent functions, like object tracking based on multiple camera views, shape classification, and detection of crossing perimeters of virtual zones. Thus with the help of all these methods input data are collected on this basis the video processing has been done. Results of the OTS [4] calculations are collected and processed in a central **Site Wide ObjectTracking Server**, which implements the site levelintelligent functions, like identity tracking, evaluation of complex rules based on the identity and security clearance of moving persons orvehicles, and moreover this module is capable of recognizing the suspicious activities. The optional **3D World Model Server** provides 3D calculation services for the other IDENTRACE modules. It can help OTS Stations in automatic calibration of the positions and viewing angles of cameras, while it can provide 3D location information to the SWOT Server, [5] and the generation of synthesized virtual images for the central monitoring console. The **RemoteIdentification Server** helps the SWOT Server tohandle the inevitable uncertainties of object tracking (e.g. when the system loses track of persons in blind areas, like rest rooms).

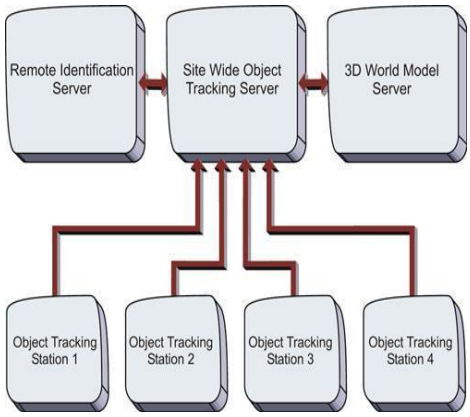


Figure 2.2Online video objects tacking for live video and movies

This not only creates a focus of attention for higher level processing but also decreases computation time considerably. Commonly used techniques for object detection are background subtraction, statistical models, temporal differencing and optical flow. Object classification step categorizes detected objects intopredefined classes such as human, vehicle, animal, clutter, etc. It is necessary to distinguish objects from each other in order to track and analyse their actions reliably. Currently, there are two major approaches towards moving object classification, which are shape-based and motion-based methods. The objects 2D spatial information whereas motion-based methods use temporal tracked features of objects for the classification solution. Detecting natural phenomenon such as fire and smoke may be incorporated into object classification components of the visual surveillance systems.

The advances in the development of these algorithms would lead to breakthroughs in applications that use visual surveillance. Monitoring of banks, department stores, Airports, museums, stations, private properties, parking lots for crime prevention, detection patrolling of highways, Railways for accident detection.[8]

2.2 Object Detection And Tracking:

The current system is able to distinguish transitory and stopped foreground objects from static background objects in dynamic scenes, detect and distinguish left and removed objects, classify detected objects into different groups such as human, human group and vehicle, track objects and generate trajectory information even in multi-occlusion cases and detect fire in video imagery. [8] The computational complexity and even the constant factors of the algorithms we use are

important for real time performance. Hence decisions on selecting the computer vision algorithms for various problems are affected by their computational run time performance as well as quality. Furthermore, the current system˝s use is limited only to stationary cameras and video inputs from Pan/Tilt/Zoom cameras where the view frustum may change arbitrarily are not supported. The system is initialized by feeding video imagery from a static camera monitoring a site. Most of the methods are able to work on both colour and monochrome video imagery. [7] The first step of our approach is distinguishing foreground objects.

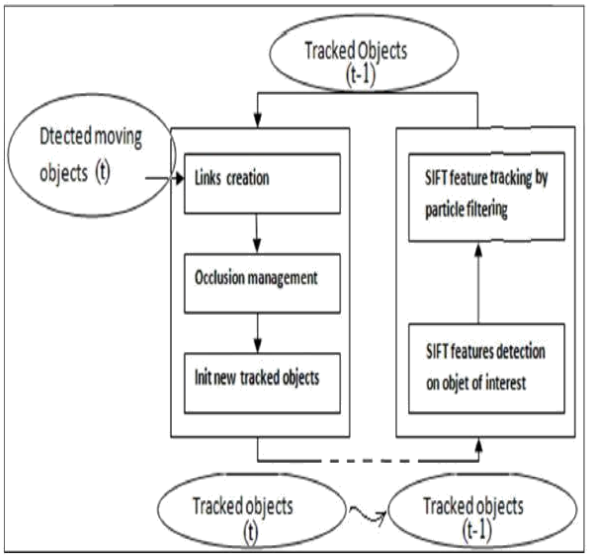


Figure 2.2.1 Process of object detection

Distinguishing foreground objects from the stationary background is both a significant and difficult problem. Almost the entire visual surveillance systems first step is detecting foreground objects. This both creates a focus of attention for higher processing levels such as tracking, [5] classification and behaviour understanding and reduces computation time considerably since only pixels belonging to foreground objects need to be dealt with. Short and long term dynamic scene changes such as repetitive motions, light repentance, shadows, camera noise and sudden illumination variations make reliable and fast object detection Difficult. Hence, it is important to pay necessary attention to object detection step to have reliable, robust and fast visual surveillance system. The aim of object tracking is 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.

3. PROPOSED ALGORITHM

The increasing rate of multimedia data and transmission facility induces some problem of data loss and delay of delivery. Now in the process of video object detection background updating is important factor for analysis. For the background updating used segmentation process and segmentation used clustering technique. Now in our dissertation used RBF neural network model for segmentation process and reduces the loss of frame and video data during object tracking process. The basic processing elements of neural networks are called artificial neurons, or simply neurons or nodes. In a simplified mathematical model of the neuron, the effects of the synapses are represented by connection weights that modulate the effect of the associated input signals, and the nonlinear characteristic exhibited by neurons isrepresented by a transfer function. The neuron impulse is then computed as the weighted sum of the input signals, transformed by the transfer function. The learning capability of an artificial neuron is achieved by adjusting the weights in accordance to the chosen learning algorithm. The basic architecture consists of three types of neuron layers: input, hidden, and output layers. In feed-forward networks, the signal flow is from input to output units, strictly in a feed-forward direction. The data processing can extend over multiple (layers of) units, but no feedback connections are present. [5] Recurrent networks contain feedback connections.

3.1 NEURAL NETWORK

Neural networks are an approach to computing that involves developing mathematical structures with the ability to learn. The methods are the result of academic investigations to model nervous system learning. Neural networks have the remarkable ability to derive meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse. This expert can then be used to provide projections given new situations of interest and answer "what if" questions. [5]

Neural networks have broad applicability to real world business problems and have already been successfully applied in many industries. Since neural networks are best at identifying patterns or trends in data, they are well suited for prediction or forecasting needs including:

- sales forecasting
- industrial process control
- customer research
- data validation
- risk management
- Target marketing etc

The structure of a neural network looks something like the following:

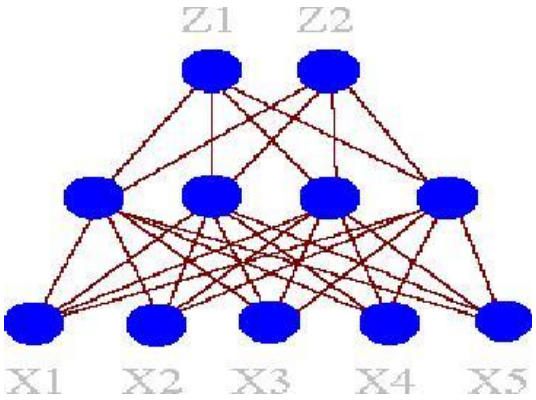


Figure 3.1 Interconnection of neural network

The bottom layer represents the input layer, in this case with 5 inputs labels X1 through X5. In the middle is something called the hidden layer, with a variable number of nodes. It is the hidden layer that performs much of the work of the network. The output layer in this case has two nodes, Z1 and Z2 representing output values we are trying to determine from the inputs. For example, predict sales (output) based on past sales, price and season (input).[5]

Each node in the hidden layer is fully connected to the inputs which mean that what is learned in a hidden node is based on all the inputs taken together. Statisticians maintain that the network can pick up the interdependencies in the model. The following diagram provides some detail into what goes on inside a hidden node.

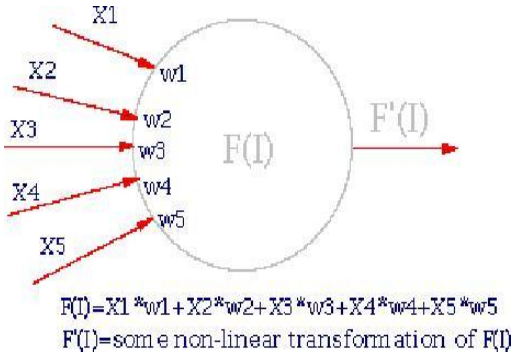


Figure 3.2 Interconnection

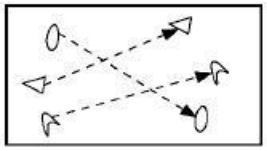
weight of neural network

Simply speaking a weighted sum is performed: X1 times W1 plus X2 times W2 on through X5 and W5. This weighted sum is performed for each hidden node and each output node and is how interactions are represented in thenetwork. The issue of where the network gets the weights from is important but suffices to say that the network learns to reduce error in its prediction of events already known (i.e., past history).

3.2 Object Tracking Technique

3.2.1 Point Tracking

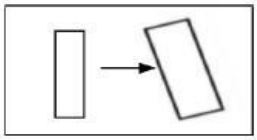
Objects detected in consecutive frames are represented by points, and a point matching is done. This approach requires an external mechanism to detect the objects in every frame.



3.2.2 Kernel Tracking

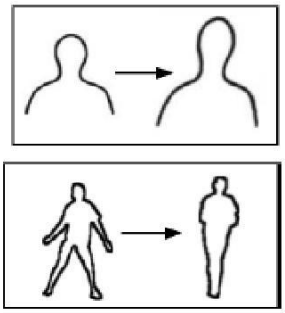
Kernel = object shape and appearance.

E.g. kernel = a rectangular template or an elliptical shape with an associated histogram. Objects are tracked by computing the motion (parametric transformation such as translation, rotation, and affine) of the kernel in consecutive frames.



3.2.3 Silhouette Tracking

Such methods use the information encoded inside the object region (appearance density and shape models). Given the object models, silhouettes are tracked by either shape matching (c) or contour evolution (d). The latter one can be considered as object segmentation applied in the temporal domain using the priors generated from the previous frames.



4. PROPOSED SOFTWARE DESCRIPTION

In the field of technical computing the various languages are used to implement the various algorithms. As we developing this system so we have to implement high performance environment so that the MATLAB [4] is a high-performance language for technical computing. MATLAB is a modern programming language environment: it has sophisticated data structures, contains built-in editing and debugging tools, and supports object-oriented programming. These factors make MATLAB an excellent tool for teaching and research. MATLAB [4] provides many functions for image processing and other tasks. Most of these functions are written in the MATLAB language and are publicly readable as plain text files. Generally the image processing algorithms are developing in MATLAB just because in image processing algorithm most complicated mathematical expressions. To implement these algorithms are MATLAB is the best frame work because in MATLAB framework the mathematical expressions are easily develop.
That‟s why to implement inaccessible algorithm, MATLAB is the only easiest way to use this framework for development proposed algorithm. MATLAB has many advantages compared to conventional computer languages (e.g., C, FORTRAN) [4] for solving technical problems. MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. The software package has been commercially available since 1984 and is now considered as standard tool at most universities and industries worldwide.

5. APPLICATION

The name of this topic itself is one of the applications of the system. The term surveillance is a most applicable to security system because in the security system, to watch on every suspicious movement in the specific organisation. The advances in the development of these algorithms would lead to breakthroughs in applications that use visual surveillance.[6] Monitoring of Banks, Departmental Stores, Airports, Museums, stations, private properties and parking lots for crime prevention and detection patrolling of highways and railways for accident detection, Measuring traffic flow. In such organisation this system is most not only applicable but also more reliable. The demand for remote monitoring for safety and security purposes has received particular attention, especially in the following areas:

- Transport applications such as airports, auditorium environments, railways, and motorways to survey traffic [6].
- Public places such as banks, supermarkets, homes, department stores and parking lots.
- Remote surveillance of human activities such as attendance at football matches or other activities [8].

6. FUTURE SCOPE

This organisation are only develop for providing security to every organisation, in upcoming days this system may be given some expert work for security such as tracking the object automatically and warn to their administrator so that the suspicious object may be detect at the instinct location. Another scope to this system may be used to control traffic system, consider one who breaks the traffic rules on road at that time this system will track that object as well as location of that object so that this system will alert traffic control organisation [6]. There is also one scope of this system that is in cricket matches to recognise the decision which is given by the third umpires.

7. CONCLUSION

This paper has presented the state of development of surveillance systems, including a review of current image processing techniques that are used in different modules that constitute part of surveillance systems. As far as this image processing tasks is concern it has identified research areas that need to be investigated further such as adaptation, data fusion and tracking methods in surveillance system. Thus the artificial neural network is best concept to develop surveillance system.

The growing demand for safety and security has led to more research in building more efficient and intelligent automated surveillance systems. Therefore, a future challenge is to develop a wide-area distributed multi-sensor surveillance system which has robust, real-time computer algorithms able to perform with minimal manual reconfiguration on variable applications. Such systems should be adaptable enough to adjust automatically and cope with changes in the environment like lighting, scene geometry or scene activity. The system should be extensible enough, be based on standard hardware and exploit plug-and-play technology.

REFERENCES

[1] Lee L, Romano R, Stein G. "Monitoring activities from multiple video streams: Establishing a common coordinate frame", IEEE Transactions on PAMI, Cornell University, 2000, pp. 758–767.

[2] R.Stolkin, I. Florescu, G. Kamberov, "An adaptive background model for CAMSHIFT tracking with a moving camera."Center for Maritime Systems (Vol.1) (Hobken: Stevens Institute of Technology), 2000, pp. 50.

[3] H. Z. Ning, L. Wang, W. M. Hu, and T. N.Tan, "Articulated model based people tracking using motion models", *Proc. Int.Conf. Multi-Model Interfaces* , pp.115 -1202002.

[4] TheMathWorks Inc. *MATLAB 7.0 (R14SP2)*. The MathWorks Inc., 2005

[5] J. Zupan, J. Gasteiger, Neural Networks for Chemists: An Introduction, VCWeinheim,

[6] AnanyaPthak, kandarpasharma, "Human Detection

And Surveillance System Using Real Time Video Feed AndAnn"IRNet Transactions on Electrical and Electronics Engineering (ITEEE) ISSN 2319 – 2577, Vol-1, Iss-2, 2012

[7] Shao-Yi Chien, Wei-Kai Chan, "Video Object Segmentation and Tracking Framework With Improved Threshold Decision and Diffusion Distance" IEEE Transactions On Circuits And Systems For Video Technology, Vol. 23, No. 6,June 2013