IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 12, December 2016

Denoising and Analysis of Surveillance Video for Security Applications

Ashwini R. Deoghare[1], Pritesh R. Gumble[2]

PG Student, Electronics & Telecommunication Dept, Sipna College of Engineering & Technology, Amravati, India[1]

Associate Professor, Electronics & Telecommunication Department, Sipna College of Engineering & Technology, Amravati, India[2]

Abstract: Nowadays because of increase in crime activities, higher demand and greater awareness on security problems lead to the study of more secure, high performance, reliable and flexible systems. To reduce such crime activities processing of videos involved in such activities will be necessary. It aims at the surveillance video of payee's abnormal behaviour based on the video denoising algorithm. In this work, we will be able to distinguish the normal and abnormal activities like money snatching, harm to the person by virtue of fight, or attack on the person. The proposed system will detect the unusual event which occurs in the security centre. By developing a system that will give an alarm when something abnormal is occurring it could eliminate the need for extra people to monitor surveillance feeds. It has growing demand for applications to support monitoring indoor and outdoor environments such as ATM center, shopping malls etc. The significant efforts in the field of tracking and moving object detection will have been done to make the security applications robust, reliable and efficient.

Keywords: Surveillance video, Denoising algorithm, unusual event detection, Moving object detection.

I. INTRODUCTION

A fundamental challenge in intelligent video surveillance is to automatically detect abnormal events in long video streams. Video anomaly detection is also important as it is related to other interesting topics in computer vision, such as dominant behaviour detection, visual saliency and interestingness prediction. A typical approach to tackle the anomaly detection task is to learn a model which describes normal activities in the video scene and then discovers unusual events by examining patterns which distinctly diverge from the model. However, the complexity of scenes and the deceptive nature of abnormal behaviours make anomaly detection still a very challenging task. Several anomaly detection approaches are based on analyzing individual moving objects in the scene. Surveillance is carried out, e. g., to ensure public safety or for safety-oriented supervision of private environments where people may live alone. In fact, the increasing desire in public security over the past decades has motivated the installation of sensors such as cameras or microphones in public places (stores, subway, airports, etc.). Thus, the need of unsupervised situation assessment stimulated the signal processing community to experiment with several according automated frameworks. Data-driven classification approaches, relying on a-priori classification of the data, were applied for a successful operation and recognition of the events. Usually, the research

In the area of automatic surveillance systems is focused on detecting abnormal events based on the acquired video. Using accurate tracking algorithms, trajectory extraction can be carried out to further perform trajectory clustering analysis or design representative features model, typical activities and subsequently discover anomalies. In trajectories which are spatially close and have similar motion patterns are identified and used for detecting unusual events. It proposes a "shape activity" model to describe moving objects and detect anomalies. The basic idea in this approach is detecting the moving objects from the difference between the current frame and the reference frame. High quality video denoising method in the context of motion based exposure controlled by combining spatial denoising and temporal denoising in a novel way. The processing of the recorded video sequence will do in MATLAB software.

Absence of automated security mechanism leads to post incident forensic analysis by the law agencies. Many a time law enforcement authorities become aware of the crime after several hours after the incident. This is a major problem in the urban areas as well as in the rural areas. Video based human activity analysis has gained lots of attention amongst the researchers. The goal of human activity recognition is to analyze different activities automatically from an unknown video. Analysis of various activities involves recognition of motion pattern and generation of high level description of actions. There are various approaches caught by the police red-handed because they are informed about the crime instantly. In addition, the system can be used to generate automated alarm that can alert security guard deputed at the security location as well as other people around the premise to obtain immediate security. Increase in nefarious activities like robbery, murder, and other crimes have raised an

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 12, December 2016

urgency to install an effective system that can protect people. Generally security centres are equipped with CCTV cameras that keep a watch on the activities. Unfortunately, CCTV is not sufficient to provide security due to their inability to recognize unusual behaviours themselves and hence monitoring authority needs to monitor these feeds 24 × 7 which is a challenging task. Today, we need an advanced system that can effectively monitor and automatically recognize unusual crime activities in public places and can also report to the nearest monitoring firm before an offender could elope. Another approach to handle this situation could be an alarm system or electrical buzzer. Each premise can be equipped with an electric buzzer. In addition, most alarms require a noticeable effort to operate, presenting an uncertainty that the perpetrator can simply physically stop the victim from triggering the alarm or may take a belligerent action against the victim if the victim is seen to initiate an alarm signal like manifold approaches, spatiotemporal interest of feature points, motion history images, accumulated motion image, and bag of words model which are recently used by many researchers for effective human action recognition and representation . In this project, we will present a system that can amend the current trends of the surveillance system. The system can automatically recognize different actions or number of persons through a CCTV camera like single normal, multiple normal, and multiple abnormal and generate signal.

Obtaining correct tracking information of moving foreground object is a difficult task in events like activity recognition and modelling. The basic idea in this approach is detecting the moving objects from the difference between the current frame and the reference frame, which is called background modelling. Event recognition will be the ultimate purpose of a fully automated video surveillance system. In event recognition the objects are detected by using background subtraction. Motion activities of segmented blobs can be utilized in event recognition and detection such as fight or theft, walking or running, overcrowding etc.

II. LITERATURE REVIEW

Various authors made a research on the implementation and modifications of video denoising for efficient results. They have analysed the utility of reliable motion estimation to establish temporal correspondence across frames in order to achieve high-quality video denoising.

David BARTOVČAK, Miroslav VRANKIĆ [1] had proposed video denoising algorithm based on adaptive, pixel-wise, temporal averaging. The algorithm decomposes videos into a set of 1-D time dependent signals and then removes the noise by establishing temporal averaging intervals throughout each signal from set. Temporal averaging Intervals had established by simple, yet effective comparison processes which include two-way thresholding. The proposed algorithm is simple, easy to use and doesn't require more complex video denoising procedures that are numerically more demanding. Low-end camera market is growing rapidly (digital cameras, web-cams, cell phones etc.) and there is a need now more than ever for fast, effective and reliable image and video enhancement technologies to improve their output. Even high-end and professional equipment (surveillance cameras, medical devices) have to cope with image degradation and noise corruption (especially in extreme conditions). Nowadays, practically every image-capturing device incorporates some sort of noise removal technology**.**

Erik Marchi1, Fabio Vesperini2, Florian Eyben1, Stefano Squartini2, Bj¨orn [2] had showed spectral features are processed by a denoising auto encoder with bidirectional Long Short-Term Memory recurrent neural networks. They used the reconstruction error between the input and the output of the auto encoder as activation signal to detect novel events. The auto encoder is trained on a public database which contains recordings of typical in-home situations such as talking, watching television, playing and eating. the information given by the acoustic signal offers several advantages, such as low computational needs or the fact that the illumination conditions of the space to be monitored do not have an effect on the sound; the same applies for possible occlusion or fast events like shots or explosions. The statistical approach is the most widely used for this problem. It consists of modelling data based on its statistical properties and using this information to estimate whether a test sample comes from the same distribution.

Travis portz, Li Zhang, Hongrui Jiang [3] had proposed new digital cameras, such as Canon SD1100 and Nikon COOLPIX S8100, have an Auto Exposure (AE) function that based on motion estimation. The motion estimation helps to set short exposure and high ISO for frames with fast motion, thereby minimizing most motion blur in recorded videos. This AE function largely turns video enhancement into a denoising problem. It will be useful to investigate a real-time implementation of this approach so that denoising can be executed before compression. It shows how to achieve high-quality video denoising in the context of motion-based exposure control. This weighted combination scheme makes our method robust to optical flow failure over regions with repetitive texture or uniform colour and combines the advantages of both spatial and temporal denoising. Our method also exploits high quality frames in a sequence to effectively enhance noisier frames. In experiments using both synthetic and real videos, our method outperforms the scale of the art.

Ce Liu1, William T. Freeman1 [4] had proposed recent advances in the sparse representations of images have achieved outstanding denoising results, removing real, Structured noise in digital videos remains a challenging problem. We showed the utility of reliable motion estimation to establish temporal correspondence across frames in order to achieve high-quality video denoising. Experimental results show that our system has comparable

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 12, December 2016

with the state of the art in removing AWGN, and significantly outperforms the state of the art in removing real, structured noise. Our system is easy to implement, with broad applications in digital video enhancement.

Dan Xu1, Ellisa Ricci [5] had presented a novel unsupervised deep learning framework for anomalous event detection in complex video scenes. While most existing works merely use hand-crafted appearance and motion features, we proposed Appearance and Motion Deep Net (AMDN) Which utilizes deep neural networks to automatically learn feature representations? The proposed method has based on multiple stacked auto encoder networks for learning both appearance and motion representations of scene activities. A double fusion scheme is designed to combine the learned feature representations.

Xiaogang Chen, Sing Bing Kang, Jie Yang, and Jingyi Yu [6] had presented a novel fast patch-based denoising technique based on Patch Geodesic Paths (Patch GP).Patch GPs treat image patches as nodes and patch differences as edge weights for computing the shortest (geodesic) paths. The path lengths can then be used as weights of the smoothing/denoising kernel. Patch GPs can be effectively approximated by minimum hop paths (MHPs) that generally correspond to Euclidean line paths connecting two patch nodes. To construct the denoising kernel, we further discretize the MHP search directions and use only patches along the search directions.

III. SYSTEM ARCHITECTURE

Our goal is to achieve an efficient, adaptive and high-quality video denoising algorithm that can effectively remove real, structured noise introduced by low-end camcorders and digital cameras. We can show the utility of reliable motion estimation to establish temporal correspondence across frames in order to achieve high-quality video denoising.

A. Block diagram

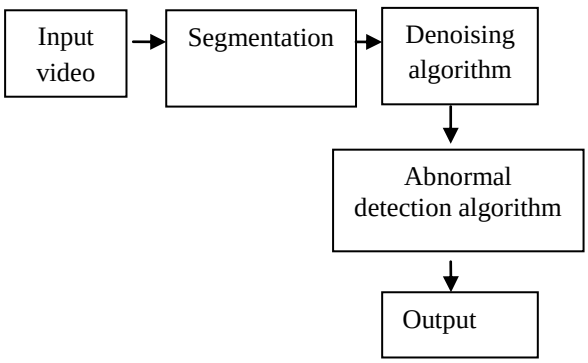

Fig1.block diagram

B. System description

The raw input video which contains some noise and it is given to the segmentation. Video object segmentation have emerged as one of the most imperative and challenging area of research. The principal objective of video object segmentation is to facilitate content-based representation by extracting objects of interest from a series of consecutive video frames.

Recently, a number of video object segmentation algorithms have been discussed and unfortunately most existing segmentation algorithms are not adequate and robust enough to process noisy video sequences. Competence of most segmentation techniques is affected by the presence of noise in frames which is a critical issue of edge preservation. This project will presents a novel video object segmentation approach for noisy colour video sequences towards effective video retrieval. Initially, the noisy video frames are denoised using a strategy based on an enhanced sparse representation in transform domain.

Afterwards, the background is estimated from the denoised frames using the Expectation Maximization (EM) algorithm. Then, the foreground objects i.e.) moving video objects are segmented with the aid of the novel approach presented. The biorthogonal wavelet transform and the L2 norm distance measure are employed in the foreground object segmentation. The experimental results demonstrate the effectiveness of the presented approach in segmenting the video objects from noisy colour video sequence .Our denoising algorithm is based on the following intuition. If an image region approaches are based on analyzing individual moving objects in the scene. Tracking is usually an initial step for this class of methods. By accurate tracking algorithms, trajectory extraction can be carried out to further perform trajectory clustering analysis or design representative features to model typical activities and subsequently discover anomaly. In analysis with tracking, objects are being found in sequential frames and the direction and speed can be computed using e.g. blob tracking, kernel-based tracking or contour tracking. This can then be classified which will indicate if the behaviour is an abnormal. When using tracking methods it is necessary to have a frame-rate that is high enough to ensure that the object will still be in the scene. Furthermore in crowded environments and scenes with much occlusion tracking can be difficult to achieve with good results.

Video anomaly detection is also important as it is related to other interesting topics in computer vision, such as dominant behaviour detection, visual saliency and interestingness prediction. A typical approach to tackle the anomaly detection task is to learn a model which describes normal activities in the video scene and then discovers unusual events by examining patterns which distinctly diverge from the model. Nature of abnormal behaviours make anomaly detection stills a very challenging task. Event recognition is the ultimate purpose of a fully automated video surveillance system. In event recognition the objects are detected by using background subtraction. Motion activities of segmented blobs can be utilized in event recognition and detection such as fight or theft, walking or running, overcrowding etc.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 12, December 2016

Architectures have been successfully used to tackle various computer vision tasks, such as image classification object detection and activity recognition. However, these works are mainly based on Convolutional Neural Networks and consider a supervised learning scenario.

Unsupervised deep learning approaches based on auto encoder networks have also been investigated .For the better understanding of algorithm this section is segregated into the various sub-sections .They are

- Rolling Background Subtraction Technique
- Close Morphological Operation (o)
- Thresholding and Standard Deviation (σ)
- GSM module

A. Rolling Background Subtraction Technique

The rolling background subtraction technique is used here. It dynamically updates the background model and then does temporal differencing. It is very adaptive to dynamic environments but there may be holes left inside moving objects. The background modelling and subtraction technique has various steps. They are Video Acquisition, Frame conversion, Pre-processing, Background modelling, Background subtraction, and Post processing, Foreground extraction.

B. Morphological Operation

Morphological Operations are generally used to fill small gaps inside the moving object, connecting disjoint objects and also helpful in reducing noise. Here the closing morphological operation is used which is composed of two sub-operations dilation followed by erosion. Dilation allows the objects to expand, thus potentially connecting disjoint objects and filling in small holes. Erosion shrinks the objects by (eroding) etching away their boundaries. By the proper selection of the structuring element these Operations can be customized for a specific application.

The closing morphological operation with disk shape structuring element on segmented areas are used here to remove pepper noise and other inaccuracies.

The close morphological operation can be expressed as

$Ft = Ft \; o \; D\rho$

Here o = morphological closing operation.

$D\rho$ = disk like structuring element.

C. Thresholding

Thresholding is the simplest method of image segmentation .Thresholding can be generally used to create binary images from a gray scale image. Segmentation involves separating an image into contours (or their regions) corresponding to objects. In each frame the every pixel is classified as either foreground (1) or background (0) using a simple thresholding functions.

$$Ft = \begin{cases} 1 & \text{if } |ut - Bt| > T \\ 0 & \text{otherwise} \end{cases}$$

Here 'T' (threshold) difference between a pixel in current frame and background model.

'ut'-pixel value in the current frame.

'Bt'-pixel value in the background model

In this proposed approach the clustering based thresholding is used. In Clustering-based methods, where the gray level samples are clustered into two parts as foreground (object) and background object tracking and face alignment . We obtain anomaly detected output which is a normal condition.

D. Standard deviation

Standard deviation (σ) is a measure shows the amount of variation or dispersion exists from the average values. A low standard deviation (standard deviation close to 0) indicates that the data points tend to be very close to the mean and high standard deviation indicates that the data points are spread out over a large range of values. The standard deviation is calculated by the following formula

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Where N -number of samples in population

μ -mean value

xi - sample value

σ - Standard deviation

In this technique the standard deviation (σ) of the population of centroids of the bounding box of the blobs in 'n' consecutive frames are calculated. If the standard deviation is above than a threshold value continuously 'x' times then it signifies the presence of unusual event.

E. GSM Module

The unusual output generated from the software is sent to the GSM modem to distribute the information to the concerned person or the concerned authority.

IV. CONCLUSION

This paper presents a review of the proposed new framework that will be able to detect unusual events such as overcrowding situation and fight. In case of occurrence of unusual activity the beep sound will occur from the buzzer and the GSM modem will send the message to the concerned authority. The need of developing such security system is the increasing number of suspicious actions at the security centres. An algorithm efficiently applicable on (LR)low resolution video and there is no need of using any training datasets, classifiers and high computational schemes that enhance low resolution videos by super resolution techniques.

DOI 10.17148/IARJSET.2016.31211

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 12, December 2016

REFERENCES

[1] David BARTOVČAK–Miroslav VRANKIĆ: "Video denoising based on adaptive temporal averaging" Vol. 32, Issue 2, 64-69, 2012.
[2] Erik Marchi1, Fabio Vesperini2, Florian Eyben1, Stefano Squartini2, Bj¨orn: "A novel approach for automatic acoustic novelty detection using Denoising auto encoder with bidirectional stm neural networks".
[3] Travis Portz, Li Zhang, Hongrui Jiang: High-quality Video Denoising for Motion-based Exposure Control."
[4] Ce Liu1, William T. Freeman1, 2, "A high quality video denoising algorith m based on reliable motion estimation".
[5] Dan Xu, Ellisa Ricci: Learning deep representations of appearance and motion for anomalous Event Detection".
[6] Xiaogang Chen1,3,4, Sing Bing Kang2, Jie Yang1,3, and Jingyi Yu4: "Fast patch-based denoising using approximated patch geodesic paths".CPVR2013
[7] A. Buades, B. Coll, and J. M. Morel: "A review of image denoising algorithms, with a new one". Simulation, 2005.
[8] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, 2014.
[9] M. Ahmad and S.-W. Lee, "Human action recognition using shape and CLG motion flow from multi-view image sequences," Pattern Recognition, vol.41, 2008.
[10] P. Bottoni, "A Dynamic Environment for Surveillance," 24-28 August, 2009, pp. 892-895.
[11] V. Saligrama and Z. Chen. "Video anomaly detection based on local statistical aggregates". In CVPR, 2012.
[12] Robert T. Collins, Alan J. Lipton, and Takeo Kanade, "Introduction to the Special Section on Video Surveillance". VOL.22, NO. 8, AUGUST 2000