Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICONECT-2015 Conference Proceedings

Motion Detection

Arati Bodake[1,] Deepali Ghadigaonkar[2], Dipika Kini[3], Samta Patel[4], Prof - J R Mahajan[5]
[1-4]Students, [5]Assistant Professor, Department of Electronics &Telecommunication
MCT's, Rajiv Gandhi Institute of Technology
Off Juhu-Versova link road,Andheri(West),Mumbai-400053

Abstract— **This paper describes a real-time system for human detection, tracking and motion Analysis. The system is an automated video surveillance system for detecting and monitoring people in both indoor and outdoor environments. Detection and tracking are achieved through several steps: First, we design a robust, adaptive background model that can deal with lightning changes, long term changes in the scene and objects occlusions. This model is used to get foreground pixels using the background subtraction method. Afterwards, noise cleaning and object detection are applied, followed by human modeling to recognize and monitor human activity in the scene such as human walking or running.**

Keywords— Motion Detection; Tracking; Human Model; Motion Analysis; Image Processing

I. INTRODUCTION

Human detection and tracking in a complex environment is a hard task, since people interact with each other, form groups and may move in unexpected ways. This requires a robust method, which copies with the different motions, without being affected by occlusions and changes of environment features. To overcome changes in the environment monitored by the system, we have to design a robust background model that can deal with slow illumination changes like light changes between day and night, fast illumination changes like clouds blocking the sun.

There are several successful vision systems for people detection and tracking like Pfinder [1], W4 [2]. These systems use human features like head or body shape, leg symmetry analysis and statistical models which restrict them to human figures. They also need large number of pixels on target due to the shape based nature of the model which leads to misidentification of small objects. These drawbacks are alleviated by using the star skeleton model based on Fujiyosh & Lipton work [3]. The main idea is that a simple form of skeletonization, which only extracts the broad internal motion features of a target, can be employed to analyze its motion. This method does not require a priori human model, or a large Number of "pixels on target" Furthermore, it is computationally inexpensive, and thus ideal for real-world video application such as outdoor video surveillance. The system described in this paper consists of the following parts:

- Background modeling.
- Object tracking
- Human modelling
- Human motion analysis.

First a background model is maintained, and a foreground image is acquired, then the contour and the features are extracted. The moving objects are then tracked, and finally a human modeling "Star Skeletonization" is applied to detect human objects and motion analysis.

I. BACKGROUND MODELING

The background is the image which contains the non-moving objects in a video. Obtaining a background model is done in two steps: first, the background initialization, where we obtain the background image from a specific time from the video sequence, then, the background maintenance, where the background is updated due to the changes that may occur in the real scene.

A. Background Model Initialization

The median algorithm is widely used in background model initialization [4]. It is based on the assumption that the background at every pixel must be visible more than fifty percent of the time during the training sequence. The median algorithm results in wrong background intensity value, especially when a moving object stops for more than fifty percent of the training sequence time. More efficient algorithm is the Highest Redundancy Ratio (HRR) algorithm [5]. HRR considers that pixel intensity belongs to the background image only if it has the highest redundancy ratio among intensity values of the pixel taken from a training sequence. This assumption is very near to the actual meaning of the background. Therefore, HRR algorithm is more flexible and applicable for real events than median algorithm. Figure 1 and figure 2 give an example of using the median and the HRR methods..
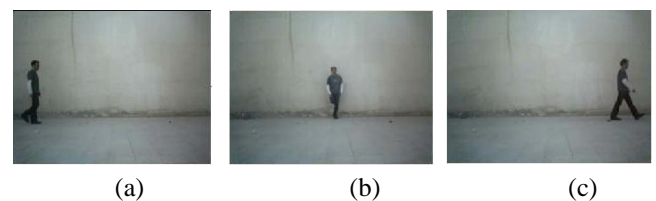

(a)   (b)   (c)

Fig 1. Example frames from a scenario. (a) a person enters the scene, (b) the person waits for a while, (c) the person goes away

(a)   (b)

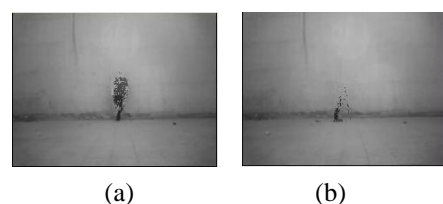Fig 2. Testing median algorithm and HRR algorithm on the scenario in figure 1. (a) Median algorithm, (b) HRR algorithm.

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICONECT-2015 Conference Proceedings

B. Background Model Maintenance

In any indoor or outdoor scene, there are many changes that may occur over time and may be classified as changes to the background scene. We can classify these changes according to their sources as follows: Illumination changes: like the change of the sun location, the change between cloudy and sunny weather, and turning the light on/off. Motion changes: like small camera displacement or tree branches moving. Changes introduced to the background: like objects entering the scene and stays without moving for a long period of time. The background model must tolerate these kinds of changes. The background maintenance helps the background model to adapt many to changes that may occur. The background maintenance model update uses two adaptations: the sequentially adaptation and the periodically adaptation. The first one is done by using a statistical background model which provides a mechanism to adapt to slow changes in the scene. This adaptation is performed using a low pass filter and is applied for each pixel. The periodically adaptation is used to adapt to high illumination and physical changes that may happen in the scene, like deposited or removed objects. In this adaptation the background model is re-initialized using the HRR algorithm and this is done in a periodical sequence (50-100 frames).

II. OBJECT TRACKING SUBSYSTEM

Object tracking is used to describe the process of

- Recording movement and translating that movement onto a digital model.
- Simulink with Video and Image processing block set enable to run fast simulations for real-time embedded video, vision, and imaging systems.
- It can create executable specifications for communicating the system to downstream design teams and to provide a golden reference for verification throughout the design process.
- The amount of work does not vary with the complexity or length of the performance to the same degree as when using traditional techniques.
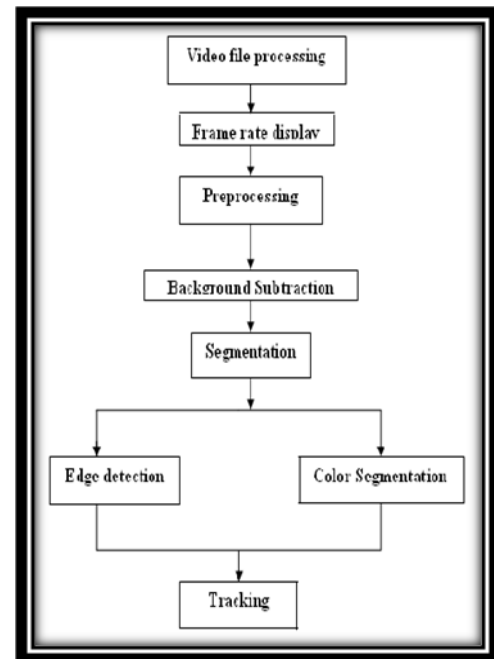

Fig 3.Object tracking

Our paper is in the same vein as given by above. Here the framework is carried out using the image processing steps such as; Video processing, frame display, background subtraction, edge Detection, segmentation and tracking such as: First, the videos are separated as frames and preprocessing method is used for the color conversion to subtract the foreground objects from the background, and background subtraction is used to find the total or sudden change in intensity in the video. Edge detection is performed as the middle wok and to extract the boundary of the object from the background segmentation module is executed. Finally tracking will be carried out. The main objective of the object tracking is to detect and track the moving object through video sequence. Most of the image processing techniques covered here are tracking related experimental results. This proposed framework combines the existing and recent techniques of video processing techniques for tracking system. The output of the system delivers a Matlab-based application development platform intended for tracking system. It allows the user to investigate, design, and evaluate algorithms and applications using object based videos. It offers standardized, which do not require detailed knowledge of the target hardware and is based on the following:

A. Pre-Processing

Pre-processing is mainly used to enhance the contrast of the image, removal of noise and isolating objects of interesting the image. Pre-processing is any form of signal processing for which the output is an image or video, the output can be either an image or a set of characteristics or parameters related to image or videos to improve or change some quality of the input. Pre-processing helps to improve the video or image such that it increases the chance for success of other processes.

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICONECT-2015 Conference Proceedings

B. Frame Rate Display

Frame rate or frame frequency is the frequency at which an imaging device produces unique consecutive images, frames that applies equally well to compute graphics, video cameras, film cameras and motion capture system. Frame rate is most often expressed in frames per second and is also uttered in progressive scan monitors as Hertz (Hz).The frame rate display block calculates and displays the average update rate of the input signals.

C. Background Subtraction

It is a computational vision process of extracting foreground objects in a particular scene. A foreground object can be described as an object of attention which help sin reducing the amount of data to be processed as well as provide important information to the task under consideration. Often, the foreground object can be thought of as a coherently moving object in a scene. We must emphasize the word coherent here because if a person is walking in front of moving leaves, the person forms the foreground object while leaves though having motion associated with them are considered background due to its repetitive behavior. In some cases, distance of the moving object also forms a basis for it to be considered a background, e.g. if in a scene one person is close to the camera while there is a person far away in background, in this case the nearby person is considered as foreground while the person far away is ignored due to its small size and the lack of information that it provides. Identifying moving objects from a video sequence is a fundamental and critical task in many computer-vision applications. A common approach is to perform background subtraction, which identifies moving objects from the portion of video frame that differs from the background model.

D. Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s)

Edge detection

Edge detection is the name for a set of mathematical methods which aim at identifying points in a digital image at which the image_brightness changes sharply or, more formally, has discontinuities. The points at which image brightness changes sharply are typically organized into a set of curved line segments termed *edges*. The purpose of detecting sharp changes in image brightness is to capture important events and changes in properties of the world. It can be shown that under rather general assumptions for an image formation model, discontinuities in image brightness are likely to correspond to:

- discontinuities in depth,

- discontinuities in surface orientation,

- changes in material properties and

- Variations in scene illumination.

Colour detection

Object detection and segmentation is the most important and challenging fundamental task of computer vision. It is a critical part in many applications such as image search, scene understanding, etc. However it is still an open problem due to the variety and complexity of object classes and backgrounds. The easiest way to detect and segment an object from an image is the color based methods. The object and the background should have a significant color difference in order to successfully segment objects using color based methods.

Fig 4.backgrond subtraction

III. HUMAN MODELING

We need a model that characterizes the human moving parts in a frames sequence so we can make a decision about

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICONECT-2015 Conference Proceedings

the object class and motion pattern. Our approach to the problem is called "Star Skeletonization" in which a star-like representation for human object is extracted, that models the five main parts of the human body (head, arms, and legs). This model was first proposed by [6] and since then it has been used in many modern tracking systems. Skeletonization can be done in different ways like thinning and distance transform methods but these methods are computation expensive which make them inadequate for real time applications. Here we define a star skeleton construction method which has low computation cost and proved to be robust when dealing with noise [7]. This is done by performing different steps of a subtraction on the objects like boundary extraction, external point's detection, and skeleton configuration.

Fig 5.An occlusion example.

a) The previous frame containing two regions, (b) the current frame containing one region which is the two regions in the previous frame merged in it, (c) the intersection between the two bounding boxes, (d). the occluded region splitting into two.

The location of the human body (external points) is defined by unwrapping the bounding contour into an Euclidean distance signal that represents the distance between the centroid of the object and the point in the object's contour. Next a median filter is applied to smooth the distance signal. Then the local maxima are extracted by monitoring the changes in the amplitude values of the distance signal, by detecting the zero crossing. Building the star skeleton is now easy by connecting each of the extracted local maximum to the centroid of the object, as it shown in figure 6. There are many advantages offered by this type of skeletonization such as scalability, computationally cheap, size insensitivity and it is applicable for different kind of objects especially to those which exhibits periodic movement and not only humans.
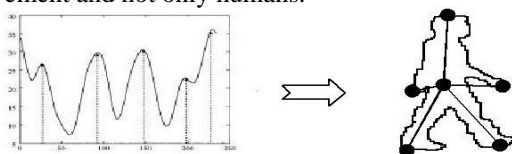

Fig 6. Detecting the local maxima and building the star skeleton model

CONCLUSION

Our approach is suitable for any real time video surveillance system especially for fast detection of moving objects in the video frame sequence. this module can be applied to any computer vision application for foreground extraction or moving object detection. This module's output can be utilized for the purpose of object detection and tracking. There are many approaches for motion detection in a continuous video stream. All of them are based on comparing of the current video frame with one from the previous frames or with something that we'll call background. If we think about the application for motion detectors, there is a lot to do with them and it depends on the imagination. One of the most straight forward applications is video surveillance, but it is not the only one.

FUTURE WORK

This work can be kept continuous for classifying detected objects and tracking trajectories of objects of the interest. Classification from all detected moving object in video can be done for human, animals, vehicles or any class for which system is aimed. In addition to classification, tracking of detected class can also be implemented. Some post processing task such as shadow removal can also be considered. After completion of detection, classification and tracking, this work could be utilized in various video surveillance applications like Visual security and surveillance,, Scene analysis and activity recognition, Event detection, Interpretation of video and logical inference.

REFRENCES

[1] C. Wren, A. Azerbaijanis, T. Darrell, and A. Paul Pent land "Pfinder: Real-Time Tracking of the Human Body" IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997.

[2] I. Haritaoglu, D. Harwood and L. S. Davis "W4: Who? When? Where? What? A Real Time System for Detecting and Tracking People" Computer Vision Laboratory, University of Maryland College Park, 1998.

[3] H. Fujiyoshi, A. J.Lipton and T. Kanade "Real-time Human Motion Analysis by Image Skeletonization" IEICE TRANS, 2004.

[4] D. Gutchessy, M. Trajkovi´cz, E. Cohen-Solalz, D. Lyonsz, A. K. Jainy "A Background Model Initialization Algorithm for Video Surveillance " Computer Science and Engineering, Michigan State University East Lansing, Adaptive Systems Philips Research Briarcliff Manor,2001.

[5] M. EKINCI, E. GEDIKL "Silhouette Based Human Motion Detection and Analysis for Real-Time Automated Video Surveillance" Dept. of Computer Engineering, Karadeniz Technical University, Trabzon, TURKEY, 2005.

[6] R. Fablet and M. J. Black "Automatic detection and tracking of human motion with a view-based representation" Department of Computer Science, Brown University, 2002.

[7] J. W. Davis, S. R. Taylor "Analysis and Recognition of Walking Movements" Dept. of Computer and Information Science, Ohio State University Columbus, 2002.

[8] Gonzalvis, Digital Image Processing

[9] Paresh M tank, Darshak G Thakore,A Fast Moving Object Detection Technique In Video Surveillance System " presented at ieee computer engineering department,