A Novel Approach to Automated Video Analysis using DECOLOR Algorithm

Mr. Dhere Satish Shivaji¹ Ms. Chavan Asawari Satish² Mr. Khandagale Hridyanath.P³

¹²³Assistant Professor
¹²³Department of Technology
¹²³Shivaji University, Vidyanaagar, Kolhapur 416 004

Abstract—Video analysis is very important for many applications such as video surveillance, traffic analysis, vehicle navigation etc. Object detection is fundamental step in video analysis. Object detection is the process of finding instances of real world objects such as faces, bicycle, and building in images or videos. It is challenging to support moving object detection in timely manner. Moreover to automate the video analysis, object detection requires training on some labeled examples. Without this training it is very difficult to analyze moving object. Some methods have been proposed to tackle this problem and these are based on motion information; but these methods are not useful when there is non-rigid motion or dynamic background. These challenges can be addressed in framework called as Detecting Contagious Outlier in the Low-rank representation. It is assumed in DECOLOR that the underlying background images are linearly correlated and the low rank matrix is formed from vectorized video. The moving objects can be detected as an outlier in low-rank representation. This method also integrates the object detection and background learning into single process.

Key words: Moving object detection, Background subtraction, Object segmentation

I. INTRODUCTION

Image processing plays an important role in many fields like Government organizations, Security System, Traffic Monitoring System etc. In image processing object detection is one of the key steps for automated video analysis. There are three key steps for machine controlled video analysis: Object Detection [1], Object Trailing & Behavior Recognition [2].

As the first opening steps: object detection aims to locate & phase fascinating object in video. Then such object is often traced from frame to border and the track scan can be analyzed to acknowledge object behavior. Object detection is sometimes achieved by object detector or background subtraction [7], [12]. An object detector is commonly a classifier that scans the image by a window and tags every sub image outlined by the window as either object or background. Mainly the classifier is made by offline learning on separate datasets [8] or by online learning initialized with manually tagged frame at the beginning of the video [9], [13].

The background subtraction is performed on an image sequence taken with a static camera to detect changes as object background subtraction compares image with background model. It always assumes that no object seems in pictures once building the background model [10].

Another type of object detection methods which will avoid the coaching phase is motioned based strategies [7] which mainly use motion data to separate object from background. The most natural means for motion based object detection is to classify pixels with respect to motion pattern which is usually named motion segmentation. An alternative motion based strategy is background estimation [11] different from background subtraction. It eliminates a background model directly from testing sequence. This approach is also depending on the idea of static background. Thus object detection has been approached using methods based on background subtraction, frame differences, histograms of oriented gradients (HOG). Partial least squares analysis, Haar Wavelets with SVM, SIFT descriptors. In this paper we propose an automatic moving object detecting system which uses novel rule for moving object detection which falls into class of motion based methods.

We will develop computer vision system which identifies the only moving objects, not all objects from video frame. It will mainly intend for moving object detection and tracing in fixed, moving camera. Therefore object detection being very useful for video surveillance application. We propose a new formulation of outlier detection in the low rank representation in which the outlier support and low rank matrix are estimated simultaneously.

II. LITERATURE REVIEW

Pramod Sharma and Ram Nevatia in their paper [14] discussed efficient techniques to detect objects from video. They have worked on human detection as object from videos. Obtained detection responses, are tracked by applying a simple tracking by- detection method, which only considers the association of detection responses in consecutive frames based on the size, position and appearance of the object. Computational efficiency is compared with other methods.

KhinThandarLwin, Dr. Than Htike and Dr. Zaw Min Naing, in their paper [16] demonstrated the multiple object detection from video. Multiple objects like vehicles are detected for speed detection of vehicles on a highway. Multiple object problems and noisy in data are also considered for performance calculation.

Osama Masoud and Nikolaos P. Papanikolopoulos [9] used multiple object detection for pedestrian tracking and counting. The system outputs the spatio-temporal coordinates of each pedestrian during the period the pedestrian is in the scene. Processing is done at three levels: raw images, blobs, and pedestrians. Pedestrians are modeled as rectangular patches with a certain dynamic behavior. Kalman filtering is used to estimate pedestrian parameters. System robustness is tested for outdoor and indoor videos.

Young-Kee Jung, Kyu-Won Lee, and Yo-Sung Ho in their paper [8] propose object segmentation and tracking algorithm for visual surveillance applications. In order to detect moving objects from a dynamic background scene which may have temporal clutters such as swaying plants, we devised an adaptive background update method and a motion classification rule. A two-dimensional token-based
tracking system using a Kalman filter is designed to track individual objects under occlusion conditions.

John Wright et.al [5] they developed a few representative examples of how the interaction between sparse signal representation and computer vision. They describe sparse representation to obtain a compact high fidelity representation observed similar and to extract semantic information. They also highlight algorithm based on sparse representation can often achieve state of art performance.

Wanceng Zhang, et.al [2] was developed rotation invariant parts based model to detect objects with complex shape in high resolution remote sensing image. The pose variance of each part relative to the object is defined in this model. They also proposed extended histogram oriented gradients and clustering method can be used to fuse the detection result. The rotation invariance of parts in rotated objects by defining the displacement deformation cost and rotation cost for parts.

“Layered Dynamic Textures” was developed by Antoni B. Chan and Antoni B. Chan [6]. layered dynamic texture (LDT) was proposed for video representation. The LDT is a generative model, which represents a video as a collection of stochastic layers of different appearance and dynamics. Each layer is modeled as a temporal texture sample from a different linear dynamical system. The ability of the LDT to segment videos into layers of coherent appearance and dynamics is also evaluated, on both synthetic and natural videos. They proposed two alternatives for inference with this model: a Gibbs sampler and an efficient variational approximation. The combination of LDT and variational inference has been shown to outperform a number of state-of-the-art methods for video segmentation.

Jichan Lee, et.al was developed “Moving Object Detection Using Background Subtraction and Motion Depth Detection in Depth Image Sequences” [12]. Proposed method detects moving objects using background subtraction and motion depth detection using the MD distribution in depth image sequences. Since a distribution of MDs is updated as the Gaussian distribution, the moving regions that are in close proximity to the background surfaces can be successfully detected.

Dibyendu Mukherjee et.al developed Multimodal 3D Histogram for detecting Moving Object [13]. They proposed a real-time method based on 3D histogram and temporal multiple mode selection, suited towards a vast majority of dynamic and noisy backgrounds, congested backgrounds and slow foregrounds. Multiple mode selection process involves identifying the dominant modes of the temporal distribution and assigning them to a multimodal background. They have provided an improved temporal mode selection method for moving object detection incorporating multiple modes from a 3D histogram. Straight-forward implementation and provides a high resistance to noise and dynamic nature of background.

Bo-Hao Chen and Shih-Chia Huang was proposed “An Advanced Moving Object Detection Algorithm for Automatic Traffic Monitoring in Real-World Limited Bandwidth Networks” [3]. They proposed an approach for motion detection which utilizes an analysis-based radial basis function network as its principal component. This approach is applicable not only in high bit-rate video streams, but in low bit-rate video streams, as well. The proposed approach consists of a various background generation stage and a moving object detection stage. This proposed approach in both low bit-rate and high bit-rate video streams. The proposed method can be easily achieved for real-time application.

Xiang Zhang, et.al developed statistical background subtraction based on imbalanced learning [11]. They studied the class imbalance problem in statistical background subtraction. They developed a spatio-temporal over-sampling method to resolve the class imbalance in background subtraction. A new index to measure the change of imbalance level during over-sampling.

“Detection and Imaging of ground moving target with real SAR data” was developed Jian Yang, et.al [4]. They developed two step ground moving target indication (GMTI) algorithm and a practical ground moving target imaging (GMTIm) algorithm with motion error compensation. The two step GMTI algorithm has the ability of indicating multiple moving targets particular those submerged by the clutter. A new classification of the moving target in airborne SAR has been presented which take the impact of the velocity ambiguity into consideration. Based on the classification, a novel ground moving target processing strategy. They developed to detect and focus the moving target with real airborne SAR data, including a two-step GMTI algorithm and a GMTI algorithm.

Osama Masoud and Nikolaos P. Papanikolopoulos was developed “a novel method for tracking and counting pedestrians in real-time using a single” [10] A real-time system for pedestrian tracking in sequences of grayscale images acquired by a stationary CCD camera. The objective is to integrate this system with a traffic control application such as a pedestrian control scheme at intersections. The proposed approach can also be used to detect and track humans in front of vehicles system was implemented on a Datacube MaxVideo 20 equipped with a Datacube Max860 and was able to achieve a peak performance of over 30 frames per second. A real-time model-based pedestrian tracking system capable of working robustly under many difficult circumstances such as occlusions and ambiguities. Performing simple classification based on the location and the direction of motion of the object being tracked.

Peter Ochs and Thomas Brox were proposed a hierarchical variational approach for turning point trajectories into dense regions for object segmentation in video. [15] Point trajectories have emerged as a powerful means to obtain high quality and fully unsupervised segmentation of objects in video shots. They can exploit the long term motion difference between objects, but they tend to be sparse due to computational reasons and the difficulty in estimating motion in homogeneous areas. In this paper they introduced a variational method to obtain dense segmentations from such sparse trajectory clusters. Information is propagated with a hierarchical, nonlinear diffusion process that runs in the continuous domain but takes super pixels into account. They show that this process raises the density from 3% to 100% and even increases the average precision of labels.
III. IDEAS OF MOVING OBJECT DETECTION

A. System Architecture:
The input to the system is given by web camera. The captured video stream may be saved, viewed or may be send on to other network via internet, an email as an attachment. In the process of segmenting first step is background detection which will be in data storage. The 4 individual models are used to train the model, update model, foreground/background classification and post processing. The first K video frames are used to coach the background model to get a model that represents variation in background during this period. The video frames from K+1 onwards are processed by background subtraction module to produce mask. Mask describes the foreground regions identified by comparing incoming frame with background model. The information from frames K+1 and onwards are used to update background model. The output of the background subtraction is masked that is processed further in post processing model which minimize the effect of noise in mask.

The next step in system is apply the segmentation algorithm by DECOLOR. After applying the DECOLOR algorithm the moving object detected. The output of DECOLOR algorithm are low rank representation of the image. The k-means clustering algorithm is used to cluster the low rank images. The k-means clustering generates specific numbers of disjoint, flat cluster.

The final step in a system compares the video frame and alters the central control unit or user through SMS using GSM model after detecting changes in the video frame. The following fig. 1 shows overall system architecture.

![Fig. 1: The proposed system architecture.](image)

B. Object Detection Using Frame Differencing:
The task to identify moving objects in a video sequence is serious and vital for a general object detection system. Frame differencing technique is applied to the successive frames, which identifies all the moving objects in consecutive frames. This basic technique employs the image subtraction operator which takes two images or frames as input and produces the output. This output is simply a third image produced after subtracting the second image pixel values from the first image pixel values. The general operation performed for this rationale is given by:

\[\text{DIFF}[X, Y] = I_i[X, Y] - I_f[X, Y]\]

\[\text{DIFF}[X, Y]\] represents the difference image of two frames.

\[\text{DIFF}[X, Y]\] image is first converted into a binary image by using binary threshold and the resultant binary image is processed by morphological operations.

I) Algorithm For Object Detection:
Input: All previous frames are stored in a memory buffer and the current frame in video is \(F_i\)
Step 1: Take \(i^{th}\) frame \((F_i)\) as input.
Step 2: Take \((i-3)^{th}\) frame \((F_{i-3})\) from the image buffer.
This image buffer is generally a temporary buffer used to store some of previous frames for future use.
Step 3: Now, perform Frame Differencing Operation on the \(i^{th}\) and \((i-3)^{th}\) frame. The resultant image generated is represented as:

\[\text{DIFF}_i = F_{i-3} - F_i\]

This method removes the limitation to detect slow moving object, which makes it independent of speed of moving object and more reliable. After the frame differencing the binary threshold operation is performed to convert difference image into a binary image with some threshold value and thus the moving object is identified with some irrelevant non-moving pixels due to flickering of camera. The binary image \((F_{\text{bin}})\), in which the pixel corresponding to moving object is set to 1 and rest is treated as background which sets to 0.

This threshold technique work as, a brightness Threshold \((T)\) is chosen with the \(\text{DIFF}[X, Y]\) to which threshold is to be applied:

- If \(\text{DIFF}[X, Y] \geq T\) then
  \[F_{\text{bin}}[X, Y] = 1\] //for object
- Else
  \[F_{\text{bin}}[X, Y] = 0\] //for background

This assumes that the interested parts are only light objects with a dark background. But for dark object Having light background we use:

- If \(\text{DIFF}[X, Y] \leq T\) then
  \[F_{\text{bin}} = 1\] //for object
- Else
  \[F_{\text{bin}} = 0\] //for background

The threshold taken here is not fixed it can vary according to our perception. The use of threshold \(T\) is just to separate the objects’ pixels from the background.

C. Decolor:
DECOLOR on several real sequence selected from public datasets of background subtraction is tested by [1]. In this paper, DECOLOR integrates object detection and background learning into a single process of optimization, and it can naturally model complex background and avoid the complicated computation of foreground motion. It turns out that the optimization can be solved by an alternating algorithm efficiently. Also, we explain the relations between DECOLOR and other sparsity-based methods. The DECOLOR aim to segment moving object from image sequence. It avoids complicated motion computation by formulating the problem as outlier detection and makes use
of the low rank modeling to deal with complex background. The low rank representation of background makes it flexible to accommodate the global variations in the background.

D. Proposed Work:

1. In proposed system we are presenting a Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation which is used for efficient object detection.
2. In proposed system we are using DECOLOR algorithm.
3. In proposed system we are taking video as input.
4. Sends a message (MMS) to the registered mobile number along with the image captured of the object.
5. Web camera storage can be reduced.

IV. SYSTEM ANALYSIS

The following snapshots show the video analysis using frame differencing algorithm.

![Input video](image1.png)

![Object detection](image2.png)

Fig. 2: Input video

Fig. 3: Object detection

V. CONCLUSION

The proposed system detects the moving as long as it can be capture by the stationary camera. DECOLOR minimize a non-convex energy via alternation optimization. Thus DECOLOR performs object detection and background estimation simultaneously without training sequences.

REFERENCES


[11] Xiang Zhang, Zhi Liu†, Hongsheng Li, Xu Zhao‡, Ping Zhang, “STATISTICAL BACKGROUND SUBTRACTION BASED ON IMBALANCED LEARNING”, IEEE


[23] Mind’s dataset, http://www.visint.org/datasets#Year_1_Videos