THE INTERNATIONAL JOURNAL OF
SCIENCE & TECHNOLEDGE

Systematic Evaluation and Comparative Analysis of Moving Object

Abitha P.
Final Year M.Tech CSE, JCET, Lakkidi, Palakkad, Kerala, India
Anu T. P.
Assistant Professor, Department of CSE, JCET, Lakkidi, Palakkad, Kerala, India

Abstract:
Object detection is the process of finding instances of real-world objects such as faces, bicycles, and buildings in images or videos. In an automated video analysis it is the fundamental step.An object detector requires manually labelled examples to train a binary classifier, while background subtraction needs a training sequence that contains no objects to build a background model. In order to achieve better performance object detector without training sequence i.e motion based methods came into existence. But when comparing situations such as non rigid motion and dynamic background, motion based object detection becomes a failure. So here an efficient algorithm is discussed which will integrates object detection and background learning into a single process of optimization. It can work effectively on a wide range of complex scenarios. An efficient alert system is implemented so that the moving object can be detected in real time.

Key words*: Moving object detection, low-rank modelling, Markov Random Fields, motion segmentation, server monitoring*

1. Introduction

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Image processing is among rapidly growing technologies today, with its applications in various aspects of a business. Digital Processing techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital technique are Pre-processing, enhancement and display, information extraction. Application of digital image processing is vast including Intelligent transport system, Moving object tracking, Remote sensing, Defense surveillance, Biomedical imaging techniques, Automatic visual inspection system.

Automated video analysis is mainly used for vision based applications. The three key Steps in automated video analysis are: object detection, object tracking, and behaviour recognition. Object detection aims to locate and segment interesting objects in a video. These objects can be tracked from frame to frame, and the tracks can be analyzed to recognize object behaviour. So, object detection plays a fundamental role in practical applications. Object detection is performed by object detectors or background subtraction techniques. Object detector is a classifier that scans the image by a sliding window and labels each sub image defined by the window as either object or background. The classifier is built by offline learning on separate datasets, or by online learning initialized with a manually labelled frame at the start of a video. And background subtraction technique compares current images with a background model and detects the changes as objects. Thus the requirements of training examples for object or background modelling limit the applicability of methods in automated video analysis.

Motion-based methods are object detection methods that can avoid training phases, in that it only use motion information to separate objects from the background. But a problem can be occurred such as: A sequence of images are given in which foreground objects are present and moving differently from the background, It is not possible to separate the objects from the background automatically. Motion-based object detection generally means to classify pixels according to motion patterns, which is known as motion segmentation. Motion Segmentation is a means of separating a moving object in an image from a static background. The background image is accumulated over multiple images , and the invariant values of each pixel are taken to be the background either after a certain number of frames or as a median value for the pixel over say 20 frames. An object which moves dynamically from frame to frame will change the value of pixels in the image and obscure the background which it moves across and an outline of the changed values can be used to segment out moving objects from the non moving background.But they assume rigid motion or smooth motion in respective regions, which is not true in practice. There are situations in which the foreground motion can be very complicated with non rigid shape changes. And also the background may be complex, including illumination changes and varying textures such as waving trees and sea waves. In such situations the motion based object detection becomes a failure.

Another method for motion-based approach is background estimation. It estimates a background model directly from the testing sequence. Many applications simply require that there be introductory frames in the sequence which contain only background elements. If pure background frames are available, pixel-wise statistics in colour and depth can be computed directly. The more difficult case is computing the background model in sequences which always contain foreground elements.Here it tries to seek the temporal intervals inside which the pixel intensity is unchanged and uses image data from such intervals for background estimation. It generally assumes a static background. So, it is difficult to handle the scenarios with complex background or moving cameras.

Here an efficient algorithm for moving object detection which falls into the category of motion-based methods is discussed and assume that the underlying background images are linearly correlated. The matrix composed of video frames can be approximated by a low-rank matrix, and the moving objects can be detected as outliers in this low-rank representation. The low-rank representation of background makes it flexible to accommodate the global variations in the background. DECOLOR performs object detection and background estimation simultaneously without training sequences.

2. Related Works

The motion segmentation aims at decomposing a video in moving objects and background. In many computer vision algorithms this decomposition is the first fundamental step. It is an essential building block for robotics, inspection, metrology, video surveillance, video indexing, traffic monitoring and many other applications. A great number of researchers has focused on the segmentation problem and this testifies the relevance of the topic. However, despite the vast literature, performances of most of the algorithms still fall far behind human perception.

In paper [3] a natural way of grouping in images is done. The pure bottom-up segmentation from static cues is well known to be ambiguous at the object level, the story changes as soon as objects move ,Here in [3] it present a method that uses long term point trajectories based on dense optical flow. Defining pair-wise distances between these trajectories allows to cluster them, which results in temporally consistent segmentations of moving objects in a video shot. In contrast to multi-body factorization, points and even whole objects may appear or disappear during the shot. In order to compute the trajectories, run a tracker, which is based on large displacement optical flow. It provides sub pixel accurate tracks on one hand, and can deal with the large motion of limbs or the background on the other. Moreover, in contrast to traditional feature point trackers, it provides arbitrarily dense trajectories, so it allows to assign region labels far more densely. With these long term point trajectories at hand, differences in how the points move can be measured. But in these methods require point trajectories as input and only output a segmentation of sparse points. The performance relies on the quality of point tracking and post processing is needed to obtain the dense segmentation. Also, they are limited when dealing with noisy data and non-rigid motion.

Object tracking is a well studied problem in computer vision and has many practical applications. The problem and its difficulty depend on several factors, such as the amount of prior knowledge about the target object and the number and type of parameters being tracked (e.g. location, scale, detailed contour).In paper [2], the problem of tracking an object in a video given its location in the first frame and no other information is discussed it is done using online MIL algorithm. In the paper[2] it focus on the problem of tracking an arbitrary object with no prior knowledge other than its location in the first video frame (sometimes referred to as "model-free" tracking). The goal was to develop a more robust way of updating an adaptive appearance model It will make the system to be able to handle partial occlusions without significant drift, and for it to work well with minimal parameter tuning. To do this, turn to a discriminative learning paradigm called Multiple Instance Learning (MIL) that can handle ambiguities in the training data.

Background subtraction, also known as Foreground Detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further processing (object recognition etc.). Generally an image's regions of interest are objects (humans, cars, text etc.) in its foreground. After the stage of image preprocessing (which may include image denoising etc.) object localisation is required which may make use of this technique.Several methods based on the sparse representation for background modelling .One pioneering work is the eigen backgrounds model [7], where the principal component analysis (PCA) is performed on a training sequence. When a new frame arrives, it is projected onto the subspace spanned by the principal components, and the residues indicate the presence of new objects. As a result the validity of this approach relies on the assumption of static background. So, it is limited when processing dynamic background or videos captured by a moving camera.

3. Motion Segmentation Using Decolor Algorithm

A sequence of images D is given ,objective is to estimate the foreground support S as well as the underlying background images B .Here the following models are required to describe the foreground, the background, and the formation of observed signal. Segmentation of the images are done using DECOLOR algorithm [1].Steps in the DECOLOR algorithm are:

3.1. Background Model

The intensity of the image should be unchanged over the sequence except for variations arising from illumination change or periodical motion of dynamic textures. So, background images are linearly correlated with each other, forming a low-rank matrix B.

$\mathrm{rank}(B) \le K$,

Where K is a constant to be defined, K is the complexity of the background model.

3.2. Foreground Model

Foreground means as any object that moves differently from the background. Foreground motion gives intensity changes which cannot be fitted into the low-rank model of background. Thus, they can be detected as outliers in the low-rank representation. The binary states of entries in S can be modelled by a Markov Random Field.

3.3. Signal Model

The signal model means the formation of D, given B and S. The three models are combined, and the energy should be minimal while performing these operations. Thus the operation is that the background images should form a low-rank matrix and fit the observed sequence in the least squares sense except for foreground regions that are sparse and contiguous.

For energy minimization, nuclear norm is used in the rank operator B. It will avoid overfitting in the rank operator by performing the nuclear norm. The function defined is non convex and it includes both continuous and discrete variables. Joint optimization over B and S is very difficult. Such that an alternating algorithm that separates the energy minimization over B and S into two steps is done. B-step is a convex optimization problem and S-step is a combinatorial optimization problem. It can be said that the optimal solutions of B-step and S-step can be computed efficiently. Estimation of the Low-Rank Matrix B can be efficiently done using SOFT-IMPUTE [4] algorithm. The optimal solution is obtained after that. At every iteration SOFT-IMPUTE decreases the value of the objective function towards its minimum, and at the same time gets closer to the set of optimal solutions of the problem. With warm-starts SOFT-IMPUTE computes the entire regularization path very efficiently along a dense series of values.

Estimation of outlier support S, can be found out using graph cuts since the energy minimization of outlier will be of the form first-order MRFs with binary labels. Graph cut means spatial and temporal smoothness be done by connecting all pairs of nodes in G which correspond to all pairs of spatially or temporally neighbouring pixels in the sequence. But, this will make extremely large and difficult to solve. In implementation, connect spatial neighbours only. So, G can be separated into sub graphs of single images, and the graph cuts can be operated for each image separately. This will reduce the computational cost.

There are 2 parameters to control the complexity of the system they are: α and β. Parameter α controls the complexity of the background model. A larger gives a B with smaller nuclear norm. Parameter β controls the controls the sparsity of the outlier support. The choosing of β depend on the noise level in images. With the model estimation getting better and better, decrease the threshold and declare more and more outliers. Here 2 parameters must be choosen such as K and ϒ. DECOLOR works stably if K and ϒ are in proper ranges. Here the fixed parameters should minimize a single lower bounded energy in each step. The convergence property of SOFT-IMPUTE is used here. So, the system must converge to a local minimum. For efficient parameter tuning, our strategy guarantees that the coefficients keep decreasing for each change. Thus, the energy decreases monotonically with the algorithm running.

4. System Architecture

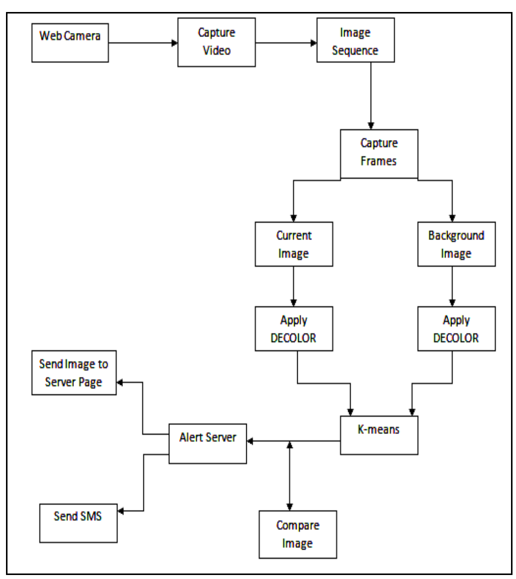

Fig ure 1: Block diagram of Moving object Detection system

Through the web camera the video is given to the system.A webcam is a video camera that feeds or streams its image in real time to or through a computer or computer network. When "captured" by the computer, the video stream may be saved, viewed or sent on to other networks via systems such as the internet, and email as an attachment. When sent to a remote location, the video stream may be save, viewed or on sent there. Unlike an IP camera (which uses a direct connection using ethernet or Wi-Fi), a webcam is generally connected by a USB cable, FireWire cable, or similar cable, or built into computer hardware, such as laptops. Their most popular use is the establishment of video links, permitting computers to act as videophones or videoconference stations. Other popular uses include security surveillance, computer vision, video broadcasting, and for recording social videos. Webcams are known for their low manufacturing cost and flexibility, making them the lowest cost form of video telephony. They have also become a source of security and privacy issues, as some built-in webcams can be remotely activated via spyware.Here through the web camera the video of the place can be captured.

Background subtraction is the first step in the process of segmenting and tracking people. Distinguishing between foreground and background in a very dynamic and unconstrained outdoor environment over several hours is a challenging task. The background model is kept in the data storage and four individual modules do training of the model, updating of the model, foreground/background classification and post processing. The first k video frames are used to train the background model to achieve a model that represents the variation in the background during this period. The following frames (from k + 1 and onwards) are each processed by the background subtraction module to produce a mask that describes the foreground regions identified by comparing the incoming frame with the background model. Information from frames k + 1 and onwards are used to update the background model either by the continuous update mechanism, the layered Updating, or both. The mask obtained from the background subtraction is processed further in the post processing module, which minimizes the effect of noise in the mask.

After applying the segmentation algorithm by DECOLOR the moving object can be detected. The DECOLOR makes low rank representation of the images. Then that images are clustered using k-means clustering. *K*-means clustering aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. The problem is computationally difficult (NP-hard), however, there are efficient heuristic algorithms that are commonly employed and converge quickly to a local optimum. These are usually similar to the expectation-maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both algorithms. Additionally, they both use cluster centers to model the data, however, *k*-means clustering tends to find clusters of comparable spatial extent, while the expectation-maximization mechanism allows clusters to have different shapes. K-Means clustering generates a specific number of disjoint, flat (non-hierarchical) clusters. It is well suited to generating globular clusters. The K-Means method is numerical, unsupervised, non-deterministic and iterative.

4.1. GSM SMS Alert System is Known as EnviroGSM System

This system is useful tool to remotely monitor critical systems like Industrial Automation, Remote security & Environmental monitoring. EnviroGSM has up to 8 input and 8 potential free Alarm outputs which can be connected to Buzzer, CCTV, Flood Lights, and Willkie-Talkie. Almost all security systems have safety cut off unit. The Potential free "NO" alarm contact from safety cut-off unit can be connected as Input to GSM Alarm SMS System. As soon as safety cut-off goes on, a SMS will be sent to user mobile whose mobile number is programmed in GSM System. Once alarm condition on chamber gets restored second SMS saying "Alert OFF" will be sent to user by GSM Alarm SMS System allowing user the access to real time information. Here after detecting the changes in video frames, we are alerting the central control unit or the user through SMS using the GSM Modem. A GSM modem is a wireless modem that works with a GSM wireless network. A wireless modem behaves like a dial-up modem

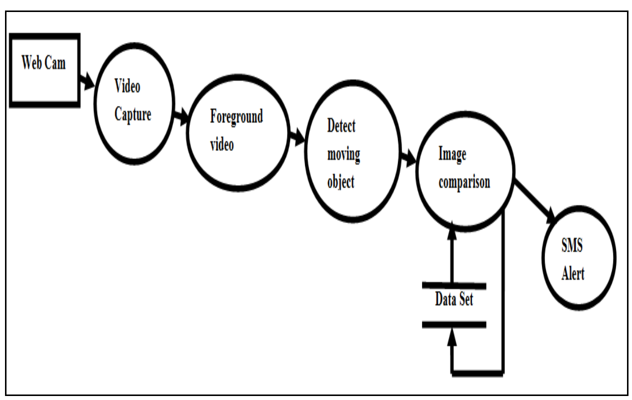

Figure 2: SMS Alert System

Here in figure 2 depicts the various steps in the SMS alert system. The main difference between them is that a dial-up modem sends and receives data through a fixed telephone line while a wireless modem sends and receives data through radio waves. Typically, an external GSM modem is connected to a computer through a serial cable or a USB cable. Like a GSM mobile phone, a GSM modem requires a SIM card from a wireless carrier in order to operate. For better security, as soon as the SMS is send to the mobile phone of the user, email is also send to the user with the compressed images of the moving object detected. Thus intruder can be detected easily.

5. Conclusion

Thus DECOLOR aims to segment moving objects from image sequences. It avoids complicated motion computation by formulating the problem as outlier detection and makes use of the low-rank modelling to deal with complex background. The outlier detection allows to get rid of many assumptions on the behaviour of foreground. The low-rank representation of background makes it flexible to accommodate the global variations in the background. Thus, DECOLOR performs object detection and background estimation simultaneously without training sequences. DECOLOR achieves better accuracy in terms of both object detection and background estimation compared against the other algorithms of RPCA. Here an efficient SMS alert system is discussed which will perform moving object detection and the intruder can be found out efficiently since on every movement occurs in the system, an SMS will be generated and send to the user. So the user can work accordingly. Since object detection is done using DECOLOR algorithm and the results are send to SMS alert system, the result produced will have advantages such as :fewer false positives, higher accuracy, larger performance when SNR ratio is relatively high, and time required to perform these operations is less. An e-mail will also be send to the user with the compressed images of the moving object to ensure better security.

6. References

1. Xiaowei Zhou, Can Yang, and Weichuan Yu,"Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation", IEEE Trans.Pattern Analysis and Machine Intelligence, March 2013.
2. B. Babenko, M.H. Yang, and S. Belongie, "Robust Object Tracking with Online Multiple Instance Learning," IEEE Trans.Pattern Analysis and Machine Intelligence, Aug. 2011.
3. T. Brox and J. Malik, "Object Segmentation by Long Term Analysis of Point Trajectories", Proc.European Conf. Computer Vision 2010
4. R. Mazumder, T. Hastie, and R. Tibshirani, "Spectral Regularization Algorithms for Learning Large Incomplete Matrices,"J. Machine Learning Research, vol. 11, 2010.
5. Y. Sheikh, O. Javed, and T. Kanade, "Background Subtraction for Freely Moving Cameras," Conf. Computer Vision,2009.
6. C. Papageorgiou, M. Oren, and T. Poggio, "A General Framework for Object Detection," Proc. IEEE Int'l Conf. Computer Vision,1998.
7. N. Oliver, B. Rosario, and A. Pentland, "A Bayesian Computer Vision System for Modeling Human Interactions," IEEE Trans.Pattern Analysis and Machine Intelligence, Aug. 2000.
8. E. Candes, X. Li, Y. Ma, and J. Wright, "Robust PrincipalComponent Analysis?" J. ACM, article 11, 2011.
9. Y. Sheikh, O. Javed, and T. Kanade, "Background Subtraction for Freely Moving Cameras," Proc. IEEE Int'l Conf. Computer Vision, 2009