MULTIPLE OBJECT TRACKING METHOD USING KALMAN FILTER

Ajay Meena

Guide: Alper Erdogan

ABSTRACT

In Image processing, tracking algorithms have been greatly researched. This is due to the increasing interest to maintain the identity of multiple targets while tracking them in some applications such as person identification, behaviour understanding, video surveillance and in self driving cars. Also novel techniques has been developed to answer the challenges of inter-object occlusion, occlusion of the objects by background obstacles, splits and merges, which are observed when objects are being tracked in real-time. In this report, an algorithm using Kalman filter and basic Image processing techniques for tracking multiple objects is discussed and implemented. The framework is completely automatic and requires no manual contribution. The framework results demonstrate that the proposed strategy can recognize and track multiple moving objects with linear as well as nonlinear motion trajectory and occlusions.

***Index Terms*—** Multi-Object Detection, Kalman Filter, Object detection,Computer Vision

1. INTRODUCTION

Techniques for extracting significant information from still images and videos captured in an constrained environment, are being contemplated for a very long while to improve images or to build automated systems. One of the problem is to detect multiple moving objects in video. It is a challenging problem to a great extent because of camera noise, shifting light condition, object occlusion, and the fluctuating size and direction of various objects. The problem of object detection, although, can be moderately simple to achieve by many existing strategies, for example, background subtraction [1],which is quite popular technique in computer vision. For instance, block matching is one of the background subtraction strategies ordinarily utilized for moving object detection. Moving object tracking of video image sequences is one of the most important subjects in computer vision. It has already been applied in many computer vision field, such as video surveillance, artificial intelligence, military guidance, safety detection, and robot navigation, medical and biological application etc. This report discusses about an algorithm implemented for detecting and tracking multiple objects in a video. The implemented algorithm exploit a simple yet practical idea to track multiple objects with occlusion. Although the objects overlap with each other, different objects move with different directions and/or speeds, which do not change much during several frames. Considering the continuity of the movement, we can predict the next position of each object based on its estimated velocity when they depart. By comparing to the previously detected objects, we can match the objects after overlapping based on their predicted velocities.
The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to predict the state of a process, in a way that minimizes the mean of the squared error. The Kalman filter is a recursive estimator. This means that only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state. In contrast to batch estimation techniques, no history of observations and/or estimates is required which makes it more efficient. The Kalman filter has two distinct phases: Predict and Update. The predict phase uses the state estimate from the previous time step to produce an estimate of the state at the current time step. This predicted state estimate is also known as the a priori state estimate because, although it is an estimate of the state at the current time step, it does not include observation information from the current time step. In the update phase, the current a priori prediction is combined with current observation information to refine the state estimate. This improved estimate is termed the a posteriori state estimate

To Summarize, the report addresses the problem of tracking multi-object in video is given below:

1. Detection of visual motion object in the given video is initially done by the background subtraction method to extract the moving object.

2. Preprocessing is the second stage where noise is removed with the help of Morphological operations like erosion and dilation.

3. Prediction of new location of the object is done by the kalman filters.

4. Finally, Hungarian algorithm is used for assigning the prediction to tracks.

2. OBJECT DETECTION

2.1. Background Subtraction

As the name suggests, background subtraction is the process of separating out foreground objects from the background in a sequence of video frames. Background subtraction is a widely used approach for detecting moving objects from static cameras. But this approach sometimes fail when the background image is not fixed, so the algorithm should adapt to Illumination changes.
The problem with background subtraction techniques is sometimes it also detect unwanted objects which can be due slight change in background of subsequent frames, this problem is dealt in this work using morphological operations [2]. Morphological operators often take a binary image and a structuring element as input and combine them using a set operator (intersection, union, inclusion, complement). They process objects in the input image based on characteristics of its shape, which are encoded in the structuring element.

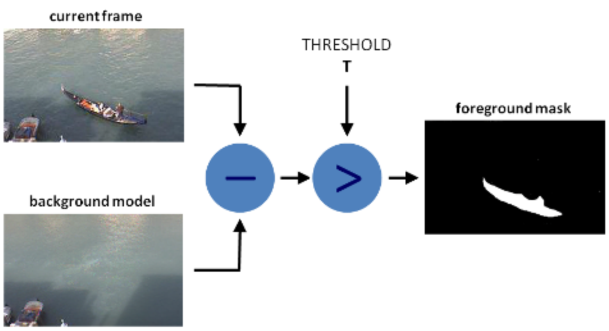


Fig. 1: Background Subtraction for object segmentation

3. OBJECT TRACKING

3.1. Typical Kalman Filter

Mathematically, Kalman filter is an estimator that predicts and corrects the states of wide range of linear processes. It is not only efficient practically but attractive theoretically as well. Precisely, the optimal state is found with smallest possible variance error, recursively. However, an accurate model is an essential requirement.

In Kalman filer, we consider a tracking system where x_k is the state vector which represents the dynamic behaviour of the object, where subscript k indicate the discrete time. The objective is to estimate x_k from the measurement z_k . Following is the mathematical description of Kalman filer, which for understanding we have sectioned into four phases.

1. *Process equation*

$$x_k = Ax_{k-1} + w_{k-1}$$

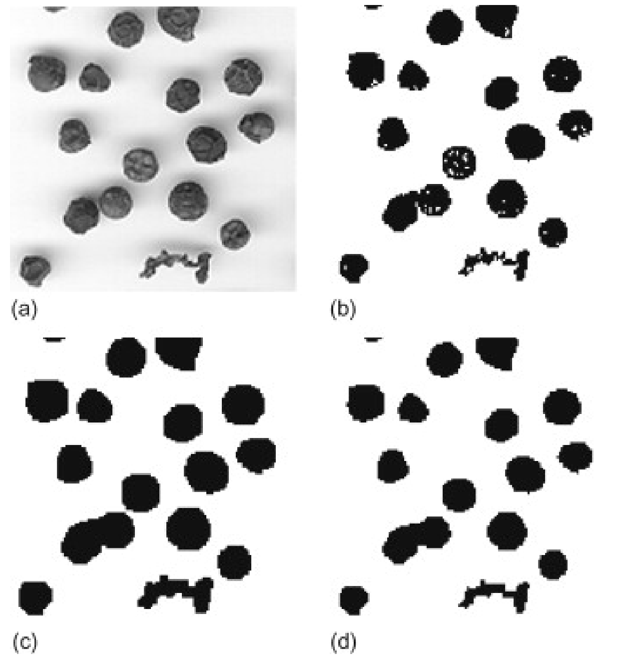


Fig. 2: Processing Image using Morphology, the operation removes the noisy gaps present in (b), (c) is dilated image of (b), (d) is eroded image of (c)

Where **A** represent transition matrix and vector w_{k-1} represents Gaussian process noise $\sim N(0, Q)$.

2. *Measurement equation*

$$z_k = Hx_k + v_k$$

Where **H** is measurement matrix and z_k is measured data and v_k is the Gaussian measurement noise $\sim N(0, R)$.

3. *Time update equations*

Above equations formulate the linear model at time k. Apriori estimate of state $\hat{x}_k^-$ and covariance error x_k estimate is obtained for the next time step k. Then:

$$\hat{x}_k^- = A\hat{x}_{k-1} + w_k$$

$$P_k^- = AP_{k-1}A^T + Q$$

4. *Measurement update equations*

For estimating $\hat{x}_k$ which is a linear combination of the apriori estimate and the new measurement z_k.

$$K_k = P_k^- H^T (HP_k^- H_T + R)^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$P_k = (1 - K_k H)P_k^-$$

Where K_k is Kalman gain, P_k^- is estimated aposterior error i.e. $P_k^- = E[e_k e_k^T]$ $where$ $e_k = x_k - \hat{x}_k^-$. Using above equation the apriori estimate is calculated from previous apriori estimate, hence estimating x_k from z_k.The time and measurement equations are calculated recursively with previous aposterior estimates to predict new aprior estimate. This recursive behaviour of estimating the states is one of the highlights of the Kalman filter.

3.2. Kalman Filter for Multi-Object Tracking

Kalman filter used for tracking is defined in terms of its states, motion model, and measurement equations. The proposed system consists of a multi-tracker where each tracker is associated with every moving object entering in the video.Describing the objects geometric features can include location, shape and center of mass (centroid). Therefore, we choose centroid and size of tracking window as the feature value to describe moving object. This method is implemented by first detecting object in video using background subtraction then by finding the centre coordinates and radius of largest contour that could fit in segmented object. Kalman filter tracking model can be divided into three sub modules which are: motion estimation model, assignment problem using Hungarian Algorithm, Model update.

1. *Motion Estimation Model*

   Kalman filter used for tracking is defined in terms of its states, motion model, and measurement equations matrix x_k is an two-dimensional system state vector, which can be expressed as:

   $$x_k = [x_{0,k}, y_{0,k}]$$

   Where $x_{0,k}, y_{0,k}$ are horizontal and vertical coordinates of centroid of contour found after background subtraction.

   The measurement vector of the system adopts the following from:

   $$z_k = [x_{0,k}, y_{0,k}]$$

   .

   In the following section A is the transition matrix and H is the measurement matrix of our tracking system along with the Gaussian process w_k and noise v_k.

   $$A = \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix}$$

   $$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

   In this implementation we have taken following assumption for process noise matrix Q and observation noise matrix R.

   $$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

   $$R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

   After the state equation and measurement equation of motion model are defined, in the next frame, Kalman filter can be used to estimate the objects location to gain trajectories of moving objects.

2. *Assignment problem using Hungarian Algorithm*

   Assigning detections to tracks in the process of tracking the multi-objects is done by Hungarian algorithm. It also determines which, are all the tracks that were missing and which detection should begin a new track. This method returns the indices of assigned tracks. Let euclidean distance between i_{th} predicted coordinate and j_{th} detected coordinate be C_{ij}, the Hungarian algorithm basically try to assign detected coordinates to predicted coordinates such that total cost get minimized. This scheme also allows us to re-label the objects which are lost by the high occlusions.
   The following discuss the assignment problem briefly. In the case where the number of detections is equal to the number of tracked objects, this is a case of the assignment problem, in which the goal is to find an optimal matching between two sets of object. Commonly, one set is called agents and the other tasks. There is a cost C_{ij} associated with assigning a task j to an agent i, and the objective is to find an assignment with minimum total cost, such that no agent is assigned more than one task, and no task is assigned to more than one agent. Formally, if A is the set of agents, T the set of tasks, and x_{ij} represents the assignment, taking value 1 if task j is assigned to agent i and 0 otherwise, then the objective is to minimize

   $$\sum_{i \in A} \sum_{j \in T} C_{ij} x_{ij}$$

   Subject to

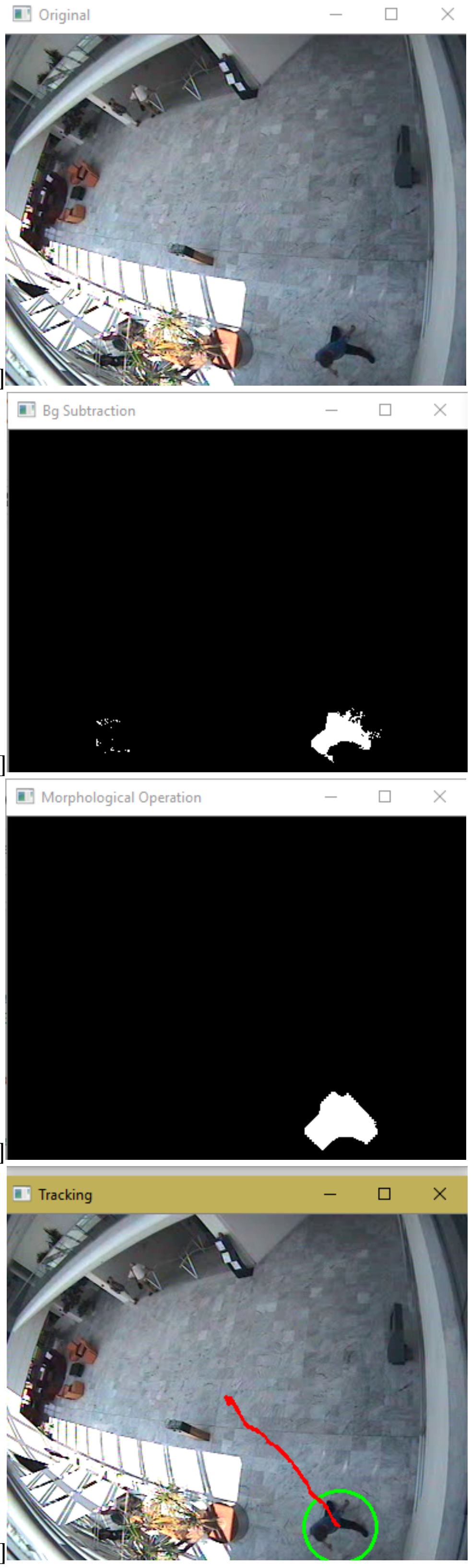
   $$\sum_{i \in A} x_{ij} = 1 \;\; j \in T$$

   ,

   $$\sum_{j \in T} x_{ij} = 1 \;\; i \in A$$

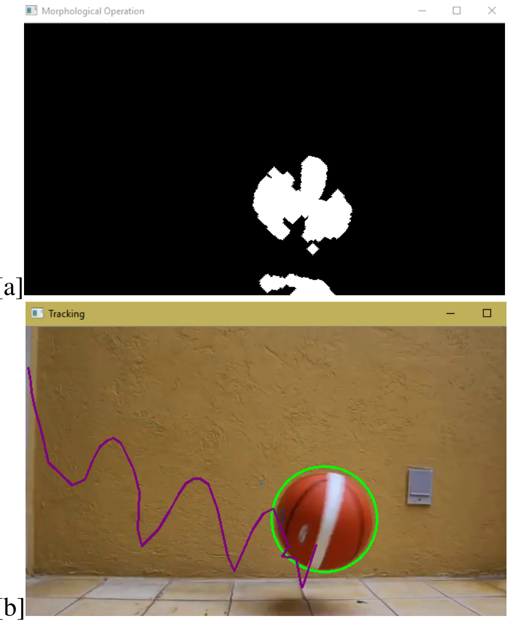
   $$x_{ij} \geq 0 \;\; j \in T, i \in A$$

4. EXPERIMENTAL RESULTS



[a] [b] [c] [d]

Fig. 3: a is original frame. b is frame after background subtraction. c is frame after applying morphological operation on b. d is final frame with track



[a] [b]

Fig. 4: a is output after morphological operation. b is final frame with track.

5. CONCLUSION

In this report, we have proposed and implemented an effective approach for multiple moving object tracking. It can be observed that object tracking has many useful applications in the robotics and computer vision fields. Several researchers have explored and implemented different approaches for tracking. The success of a particular approach depends largely on the problem domain. In other words, a method that is successful in robot navigation may not be equally successful in automated surveillance.

After estimation of model parameters using Kalman filter, measurements are used for updating the model of moving objects. Image processing techniques are used to reduce noise from extracted background. It is observed and tested that processing of each frame is not necessary which could save the computational time and increase efficiency in real time Multi-Object tracking.

6. REFERENCES

[1] Rudrika Kalsotra, Sakshi Arora. Morphological based moving object detection with background subtraction method.

[2] Jonathan Owensa, Andrew Hunter. A Fast Model-Free Morphology Based Object Tracking Algorithm.

[3] Xi Chen, Xiao Wang and Jianhua Xuan. Tracking Multiple Moving Objects Using Unscented Kalman Filtering Techniques.

[4] Xin Li Kejun Wang,Wei Wang and Yang Li. A Multiple Object Tracking Method Using Kalman Filter.

[5] Sumit Kumar Pal, Sohan Ghorai. Moving Object Tracking System In Video With Kalman Filter.

[6] Hui-Lee Ooi, Guillaume-Alexandre Bilodeau, Nicolas Saunier, and David-Alexandre Beaupre.Multiple Object Tracking in Urban Traffic Scenes with a Multiclass Object Detector.

[7] http://studentdavestutorials.weebly.com/multi-bugobject-tracking.html