Multi target tracking in static background by Kalman Filter method

Yuehua Su, XJTU, Control Science and Engineering 7058, 3117038011

Abstract—The technology of target detection and tracking has a very wide application in real life. The traditional target detection algorithm can be relatively easy to detect simple background images, but it has a great limitation on the algorithm ability of complex scenes. Compared with other video detection tasks, the target recognition task applys to motion-based video is relatively simple because of the static background, and it pays attention to the multi target detection and tracking. The task is divided into two parts: detect each frame of motion objects; match the detected area in real time to the same object. The detection part uses the background subtraction method based on the Gaussian mixture model, and the tracking part uses Kalman Filter.

Index Terms—Object detection, Object tracking, Kalman Filter, Multi target tracking. Static background.

1 Introduction

As an important research area of computer vision and pattern recognition, target detection and tracking technology has been widely applied to image processing, machine learning and artificial intelligence, and has attracted wide attention at home and abroad in recent years, and has become a hot topic. It has been widely used in the field of intelligent traffic supervision, military and aerospace. The object image in a series of changes in the target detection is the study of how to quickly and accurately extract the interest, some common objects such as people, moving vehicles and other animal is our concern, has great application prospect. The target tracking is in the continuous video sequence, the supervision should track the position of the object, and get the complete motion track of the object. Given the location of the target coordinates of the first frame of the image, the exact location of the target in the next frame is calculated. During movement, objects may show some image changes, such as posture or shape change, scale change, background occlusion or brightness change.

The problem of motion-based object tracking can be divided into two parts: Detecting moving objects in each frame; Associating the detections corresponding to the same object over time.

In this paper, Gaussian mixture model[1] and Kalman Filter multi target detection and tracking[2] are used to analyze the task of learning target detection and tracking based Matlab.The detection of moving objects uses a background subtraction algorithm based on Gaussian mixture models. Morphological operations are applied to the resulting foreground mask to eliminate noise. Finally, blob analysis detects groups of connected pixels, which are likely to correspond to moving objects. The association of detections to the same object is based solely on motion. The motion of each track is estimated by a Kalman filter. The filter is used to predict the track's location in each frame, and determine the likelihood of each detection being assigned to each track.

2 Related work

Traditional target detection methods are based on Boosting framework: Haar/LBP/ integration, HOG/ACF feature + Adaboost, SVM:HOG+SVM or DPM, template matching (in special cases), region proposal and regression based deep learning CNN method. In this article, we apply Gaussian mixture method.

Classical target tracking methods, such as Meanshift, Particle Filter and Kalman Filter, and optical flow algorithms based on feature points. The Meanshift method is a tracking method based on probability density distribution, which makes the search of target always follow the direction of probability gradient, and converges to the local peak value of probability density distribution.

First, Meanshift[3] will model the target, such as the target's color distribution to describe the target, then calculate the probability distribution of the target on the next frame, so as to iteratively get the most densely located area. Meanshift is suitable for the color model of the target and the large background difference, and it is also used in the face tracking in the early stage. Because of the fast calculation of the Meanshift method, many of its improved methods have been applied to the present. The particle filter (Particle Filter) method is a method based on particle distribution statistics. Taking the tracking as an example, the tracking target is modeled first, and a similarity measure is defined to determine the degree of matching between the particle and the target. In the process of target search, it will scatter some particles according to a certain distribution (such as uniform distribution or Gauss distribution), count the similarity of these particles, and determine the possible location of the target. In these positions, the next frame adds more new particles to ensure that the target is tracked on a larger probability. Kalman Filter is often used to describe the motion model of a target. It does not aim at modeling the characteristics of the target, but instead models the moving model of the target, which is often used to estimate the location of the target in the next frame.

In addition, the classical tracking method also has optical flow tracking based on feature points, extracts some feature points on the target, and then calculates the optical flow matching points of these feature points in the next frame to get the location of the target statistically. In the process of tracking, we need to constantly supplement the new feature points and delete the feature points with poor confidence in order to adapt to the change of the shape of the target in motion. In essence, it can be considered that the optical flow tracking is a method of characterizing the target model by using a set of feature points.

The classical Calman filtering method has high robustness and accuracy. It has very important applications in tracking process, and is often regarded as an important auxiliary means.

3 Multi target tracking in static background by Kalman Filter method

3.1 Gaussian mixture method

The single Gaussian model is a pattern classification using the multidimensional Gaussian distribution probability. This is defined as:

$$N(x/C) = \frac{1}{\sqrt{(2\pi)|\sum|}} exp[-\frac{1}{2}(x-\mu)^T \sum^{-1}(x-\mu)]$$

Where, μ means value of the training sample, $\sum$ means sample variance and the sample vector of X is d dimension. The probability of class C is a positive or negative sample can be obtained by Gauss's probability formula. The Gaussian mixture model (GMM) is the result of data from a number of Gaussian distributions. Each GMM is linearly superimposed by the K Gaussian distribution.

$P(x) = \sum p(k) * p(x|k)$ is equivalent to weighting each Gaussian distribution (the greater the weight coefficient, the greater the probability that this data is belonging to this Gaussian distribution). In the actual process, we estimate the parameters of GMM under the premise of known data, specifically training a suitable GMM model for pictures.

In the foreground detection, we will take the static background (about 50 frames) to estimate the GMM parameters for the background modeling. The classification domain value official network obtains 0.7, the empirical value 0.7-0.75 can be adjusted. This step will separate the foreground and background and output the foreground two value mask. Then the morphological operation is carried out, and the centroids and bboxes of the moving area are returned through the function, and the foreground detection part is completed.

3.2 Kalman Filtering

The tracking part is used by the Kalman Filter. Kalman filter method is a linear estimation algorithm, which can establish the relationship of bboxes between frames[4].

The tracking is divided into 5 states:

- the new target appears
- the target match
- the target occlusion
- the target separation
- the target disappearing

Build a model:

- state equation:
  $X(k+1) = A(k+1,k)X(k) + w(k)$ where $X(k) = [x(k), y(k), w(k), h(k), v(k)]$
  x, y, w, h represents the vertical and horizontal coordinates of bboxs, length and width.
- measurement equation
  $Z(k) = H(k)X(k) + v(k)$ where $w(k), v(k)$ is unrelated Gaussian white noise.

After defining the observation equation and the state equation, the Kalman filter can be used to track the moving target. The following steps are as follows:

1) Calculate the characteristic information of the moving target (the motion center of mass, and the outer rectangle).
2) The Kalman filter is initialized with the obtained feature information, which can be initially 0 at the beginning.
3) The Kalman filter is used to predict the corresponding target area in the next frame, and the target match is carried out in the prediction area when the next frame arrives.
4) If the match is successful, the Kalman filter will be updated.

The Hungarian matching algorithm is used in the matching process, where the Hungarian matching algorithm matches the moving objects detected in the new frame to the corresponding trajectory. The matching process is achieved by minimizing the sum of the Euclidean distance between the centroid and the detected centroid by minimizing the Kalman prediction, which can be divided into two steps.

(1) The loss matrix is calculated to be [M N], where M is the number of trajectories, and N is the number of moving objects detected.

(2) Solving the loss matrix.

4 Experiment

The structure of experiment is representing in figure1.

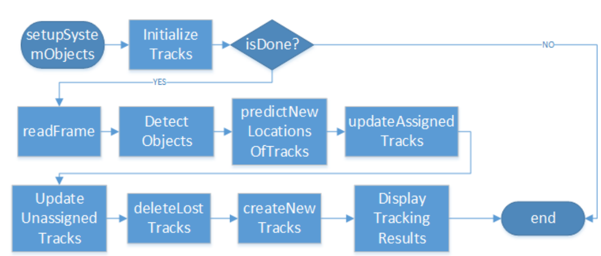

Fig. 1: Structure

The detailed steps:

- Initialize Video I/O
  Create System objects used for reading video, detecting moving objects We use video I/O create objects for reading a video from a file, drawing the tracked objects in each frame, and playing the video. The system create foreground detector and it will be

used to segment moving objects from background. It outputs a binary mask, where the pixel value of 1 corresponds to the foreground and the value of 0 corresponds to the background.

- Initialize tracks
  Create an array of tracks, where each track is a structure representing a moving object in the video. This structure maintains the state of a tracked object. The structure contains Kalman Filter, and this object used for motion-based tracking. After initialize tracks, system need to read a video frame by activate the read function.
- Detect Objects
  This function returns the centroids and the bounding boxes of the detected objects. As mentioned above, it also returns the binary mask, which has the same size as the input frame. The function performs motion segmentation using the foreground detector, and then performs morphological operations on the resulting binary mask to remove noisy pixels and fill the holes in the remaining blobs.
- Predict New Locations of Existing Tracks
  Here we use the Kalman filter to predict the centroid of each track in the current frame, and update its bounding box accordingly. Firstly it predicts the current location of the track according to the prior location information. Then adjust the bounding box in order to make its center and the predicted location. At last we get the real current location.
- Assign Detections to Tracks
  Assigning object detections in the current frame to existing tracks is done by minimizing cost. The cost is defined as the negative log-likelihood of a detection corresponding to a track. Using the distance method to compute the cost of assigning every detection to each track. It includes the confidence of the prediction by using the Euclidean distance. Kalman filter predicts the results and they are stored in an MxN matrix, where M is the number of tracks, and N is the number of detections.
  Then solve the assignment problem, it is represented by the cost matrix using the function. Update Assigned Tracks. This function updates each assigned track with the corresponding detection. Update Unassigned Tracks. Mark each unassigned track as invisible, and increase its age by 1.
- Delete Lost Tracks
  This function deletes tracks that have been invisible for too many consecutive frames. It deletes recently created tracks that have been invisible for too many frames overall.
- Create New Tracks
  Create new tracks from unassigned detections. Assume that any unassigned detection is a start of a new track. At last we display the result by drawing a bounding box and label ID for each track on the video frame and foreground mask.

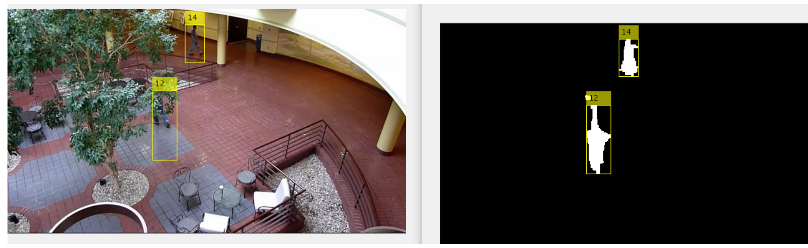

Fig. 2: Result 1

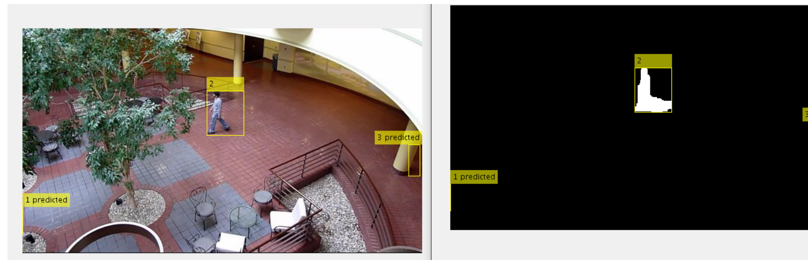

Fig. 3: Result 2

The results of experiment are representing in figure2,3.

5 Conclusion

In recent years, target detection and recognition in the field of artificial intelligence as an important branch of computer vision and pattern recognition, has attracted the attention of many scholars at home and abroad. And its application prospect has gradually entered into the field of vision. The traditional detection algorithm because of its simple algorithm and easy to realize has been widely used in a period of time. But there are many problems wait to be solved, such as light effects, and multi objects, scene and occlusion effects are unsatisfactory. The detection result is not conducive to further follow-up recognition. The use of Gaussian mixture method has good robustness on the environment, even in many scenes or in the presence of occlusions is good for detection. And Kalman Filter has a great help to predict the position through the current information and correct prediction of the next position. This will help to improve the accuracy of location information.

References

[1] Li H, Achim A, Bull D R. , *GMM-based efficient foreground detection with adaptive region update*, 365-370. IEEE International Conference on Image Processing. 2010:3181-3184.

[2] Wang M, Liu P, Li X, et al.*Multi-target tracking through occlusions using extended Kalman filter and network flows* IEEE International Conference on Computer and Communications. IEEE, 2017:2611-2617.

[3] Wang, Changyou and Zhang, Haiqiang and Yang, Fuping, *An Adaptive Kernel Bandwidth Mean-Shift Target Tracking Algorithm*,547-550,2012.

[4] Scott J, Pusateri M A, Cornish D, *Kalman filter based video background estimation*, 282-286. Applied Imagery Pattern Recognition Workshop. IEEE, 2009:1-7.