A Novel approach for Evaluation of video tracking Under Real world conditions

[1]M. Jeeva, [2]Mr. J.R. Balakrishnan

PG Student, Professor and Director
Department of Computer Science & Engineering Anand Institute of Higher Technology, kazhipattur Chennai
603 103

 Abstract: *Tracking object in videos has applications in video surveillance and other main applications. The absence of a commonly adopted performance evaluation Framework is hampering advances in the design of effective video trackers. Tracking an object in a video is a challenging process. The proposed algorithm to track the multiple objects in a video even in conditions such as diversity of targets (head, vehicle and person) and challenges (occlusions, background clutter, pose changes and scale changes).The objects can be tracked based on the compression domain and pixel domain. Using background subtraction method the objects in the fore ground are extracted. The objects in the frame are then detected and then the detected objects are labelled. Performance analysis is done and the tracking accuracy and tracking failures are measured. The proposed system is used to improve the tracking accuracy and decrease the tracking failures in the video tracking process. Finally Performance analysis shows that for different type of trackers our algorithm shows best performance.*
 Index Terms: Evaluation measure, Performance evaluation, video trackers.

I.   Introduction

Video tracking is the process of locating a moving object (or multiple objects) over time using a camera. Video tracking can be a time consuming process due to the amount of data that is contained in video. The objective of video tracking is to associate target objects in consecutive video frames. The association can be especially difficult when the objects are moving fast relative to the frame rate. A video tracker is to measure position as a function of time for an object. The performance measure evaluates tracking accuracy and tracking failure, and my algorithm shows best performance.

A number of efforts have been made toward performance evaluation of trackers in the form of evaluation campaigns and small-scale evaluation frameworks. The performance of trackers is still tested using different evaluation criteria and varying datasets, thus hindering an effective evaluation and comparison. Performance evaluation may involve the computation of the discrepancy between the estimated and the ground-truth position and size of the target.

The discrepancy is computed based on a distance-based criterion or an overlap based Criterion. Distance-based criteria use the concept of distance minimization between estimation and ground truth to evaluate performance. Naive distance-based evaluation may not include target size variations in the evaluation procedure and may not effectively reflect instances of tracking failure that refers to the case of no-overlap between estimated and ground-truth states. Overlap-based criteria compute the amount of overlap between estimation and ground truth.

Trackers operate under various challenges in real-world applications and therefore these challenges should be considered when evaluating and comparing performance. Tracking challenges include initialization errors caused by a detector; sensor noise; latency due to the transmission of video data over a channel or due to the delayed generation of results by the tracker; changing illumination in the scene; and compression of the video data. Then evaluate performance score of tracker and protocol by evaluating and comparing state-of-the-art trackers.

MS (Mean Shift) consistently tracks more accurately in the presence of initialization errors than other trackers. Semi Boost has the smallest variation in performance compared with other trackers. MS handles better initialization errors and outperforms other trackers with compressed videos and resolution reductions. Boost copes well with noise, with frame dropping and with changing illumination. The trackers performance is measured by using single target tracking performance evaluation.

The proposed system is Multi-target tracking performance evaluation.Tracking an object in a video is a challenging process. Tracking objects in a video has applications in video surveillance and other some applications. The objects can be tracked based on the compression domain and pixel domain. The proposed method is used to track object in a real world environment. This uses Background subtraction method to extract the object then the object is labeled and finally performance analysis is done. The goal is to achieve robustness

and high levels of accuracy when tracking multiple objects. To make the multi-target tracking algorithm is robust to false detections.
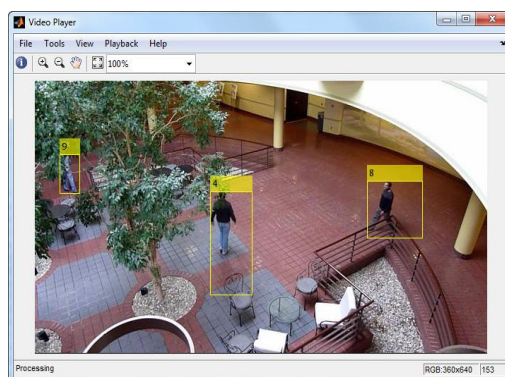

Fig 1.1 Sample tracking results generated from video

II. Related Work

1) J. Black, T. Ellis, and P. Rosin says a novel method for video tracking performance evaluation (2003).This paper considers a methodology for evaluating the performance of video surveillance tracking systems. The performance evaluation framework automatically selects ground truth tracks from a surveillance database in order to construct pseudo synthetic video sequences. They have compiled a comprehensive set of metrics, which can be used to measure the quality of the ground truth tracks, as well as characterise tracking performance. They recognise that the pseudo synthetic video have a degree of bias to the motion detection algorithm used to capture the original data. The main strength is generating a variety of different testing datasets.

2) L. M. Brown, A. W. Senior, Y.-L. Tian , J. Connell, A. Hampapur, C. f. Shu, H. Merkl, and M. Lu says performance evaluation of surveillance systems under varying Conditions (2005)**.** This paper measure the performance of our system under several different conditions including: indoor/outdoor, different weather conditions (precipitation, wind, and brightness), different cameras/viewpoints. They evaluate the performance of the background subtraction alone and with a simple tracking system using two different sets of metrics. Performance evaluation systems have been developed to analyze the two primary levels of processing: background subtraction and tracking. They proposed a new two-pass matching scheme. In phase 1, each system (S) track is allowed to match to many ground-truth (GT) tracks. A GT track is matched to the system track if there is both temporal overlap and spatial overlap. In phase 2, each GT track is matched to many system tracks.

3) Faisal Bashir, Fatih Porikli says Performance evaluation of object detection and tracking systems (2006). Set of metrics and algorithms is used for performance evaluation of object tracking system. They proposed metrics from statistical detection and estimation theory tailored to object detection and tracking tasks using frame-based as well as object-based evaluation paradigms. These metrics are evaluated after establishing correspondences between ground truth and tracker result objects. Object correspondences between multiple ground truth objects to multiple tracker result objects are established from a correspondence matrix. The many-to-many correspondence matrix is built using three different methods of distance computation between trajectories.

4) F. Yin, D. Makris, and S. A. Velastin says Performance evaluation of object tracking algorithms (2007). The proposed a rich set of metrics to assess different aspects of performance of motion tracking. This approach use six different video sequences that represent a variety of challenges to illustrate the practical value of the proposed metrics by evaluating and comparing two motion tracking algorithms. Such an approach is justified when the objective of performance evaluation is object detection. In object tracking, measuring True Positive, False Positive and False Negative in terms of tracks rather than frames is a natural choice that is consistent to the expectations of the end-user.

5) Hanxi Li, Chunhua Shen, and Qinfeng Shi says Real-time visual tracking using Sparse Representation (2011). They proposed a Real Time Compressed Sensing Tracking (RTCST) by exploiting the signal recovery power of Compressed Sensing (CS). Dimensionality reduction and a customized Orthogonal Matching Pursuit (OMP) algorithm are adopted to accelerate the CS tracking. A stationary camera defined tracker is designed by integrating a CS-based background model (CSBM). This CSBM equipped tracker coined as RTCST-B, outperforms most state of-the-arts with respect to both accuracy and robustness. The experimental results on various video sequences, which are verified by a new metric is Tracking Success Probability (TSP), show the excellence of the proposed algorithms.

6) Tim Ellis says Performance Metrics and Methods for Tracking in Surveillance (2002). Performance evaluation has become an increasingly important feature of video surveillance systems, as researchers attempt to assess the reliability and robustness of their operation. In order to satisfy the requirements of a real video

surveillance task, the algorithms will need to be quantitatively assessed over a wide range of conditions. In the first part of this paper to identify the principal sources of disturbing influences, circumscribed bound box which can be considered as a form of noise or error. In the second part, it attempt to categorise metrics that have been used to characterise the performance of both detection and tracking algorithms.

7) D. Scharstein and R. Szeliski says a taxonomy and evaluation of dense two frame stereo correspondence Algorithms (2002). A taxonomy of dense, two-frame stereo methods designed to assess the different components and design decisions made in individual stereo algorithms.Their goals are two-fold: 1.To provide a taxonomy of existing stereo algorithms that allows the dissection and comparison of individual algorithm components design decisions, 2. To provide a test bed for the quantitative evaluation of stereo algorithms. Towards this end, they are placing sample implementations of correspondence algorithms along with test data and results on the Web. A comparative evaluation is necessary to assess the performance of both established and new algorithms and to gauge the progress of the field.

8) R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, J. Garofolo says Framework for performance evaluation of face, text, and vehicle Detection and tracking in video: Data, metrics, and protocol (2009). Common benchmark data sets, standardized performance metrics, and baseline algorithms have demonstrated considerable impact on research and development in a variety of application domains. These resources provide both consumers and developers of technology with a common framework to objectively compare the performance of different algorithms and algorithmic improvements. In this paper, they presented such a framework for evaluating object detection and tracking in video: specifically for face, text, and vehicle objects. This framework includes the source video data, ground-truth annotations (along with guidelines for annotation), performance metrics, evaluation protocols, and tools including scoring software and baseline algorithms.

9) P. Pan, F. Porikli, and D. Schonfeld says Recurrent tracking using multifold consistency (2009). An adaptive object tracking algorithm that is based on a novel consistency measurement computed recursively over a multifold of forward and backward frames is proposed. They proposed multifold consistency to assess the quality of the tracking. They introduced a multifold consistency score computed for each tracked object. This enables estimating the instantaneous condition of the tracking and discovering potential failures for each individual object. The multifold consistency is utilized as the basic principle to automatically adapt parameters as well as switch among different algorithms. The tracking approach can identify occlusions and determine whether an object left the scene or not.

10) D. Comaniciu, V. Ramesh says Kernel-Based Object Tracking (2003). The main contribution of the paper is to introduce a new framework for efficient tracking of nonrigid objects. They show that by spatially masking the target with an isotropic kernel, a spatially-smooth similarity function can be defined and the target localization problem is then reduced to a search in the basin of attraction of this function. The smoothness of the similarity function allows application of a gradient optimization method which yields much faster target localization compared with the (optimized) exhaustive search. The similarity between the target model and the target candidates in the next frame is measured using the metric derived from the Bhattacharyya coefficient. The Bhattacharyya coefficient has the meaning of a correlation score.

III.   Conclusion Of Related Work

The above papers are evaluating performance of tracking algorithms and measure tracking accuracy, but in my approach tracking accuracy is much faster than the above approaches. In the above methods tracking can be done only using background subtraction method, but in my proposed system labelling method and measure tracking failure are used along with background subtraction method in order to increase accuracy. In the above approaches metrics is used for performance evaluation. In my proposed system metrics not required.

IV.   Motivation Of Research

The motivation of my research is to detect objects from video and measure tracking accuracy and tracking failure by using tracking algorithm.

- How to convert video into frames?
- How to remove noises from frames?
- How to perform tracking the target objects from frames?
- How to measure performance of different trackers?

V.   Proposed Work

The project deals with a method to track object in a real world environment. The proposed algorithm to track the objects in a video even in conditions such as diversity of targets (head, vehicle and person) and challenges (occlusions, background clutter, pose changes and scale changes).

In preprocessing unnecessary noises in the video are eliminated. An unnecessary noise refers to the unwanted pixels in the frames. Precious bits are wasted on noisy video. Color correction can mask noise. It's used to improve the overall image quality. Pre-processing methods use a small neighborhood of a pixel in an input image to get a new brightness value in output image. Such pre-processing operations are also called filtration. Deinterlacing, scaling, and color space down sampling are most commonly used video preprocessing functions. Gamma correction is also an important step in the preprocessor because PC monitors tend to be darker than the television sets. Local pre-processing methods can be divided into the two groups according to the goal of the processing. Gaussians filter to remove noise from the frame. Gaussian filter removes the noise from color image directly. In the image each pixel's new value is set to a weighted average of that pixel's neighborhood. The original pixel's value receives the heaviest weight and neighboring pixels receive smaller weights as their distance to the original pixel increases. This will preserve the information in the image more clearly.
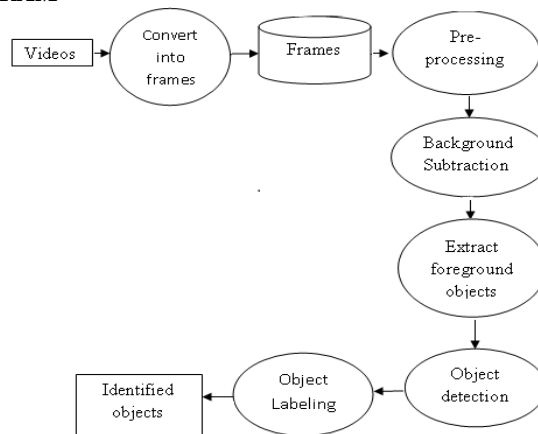
ARCHITECTURAL DIAGRAM

Fig 5.1 Overall design

Background subtraction is a widely used approach for detecting moving objects in videos from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called "background image", or "background model". The background from the frame is subtracted using absolute difference of the frame with a empty background. The difference between the image at a time t and the image at a time t-1 removes the background. Then obtained image pixels are compared with the threshold. The pixels that are greater than the threshold are taken and then the pixels are made into 1 the remaining pixels made into 0. The result will be the foreground. The obtained foreground represents the objects in the video.

After background subtraction the foreground object is detected. Object detection is a process for identifying a specific object in a digital image or video. Object detection algorithms rely on matching or learning algorithms using appearance-based or feature-based techniques. The appearance based mechanism to detect the objects. The object is detected by measuring the co-ordinates of the objects. The centroid position, Area, Height, Width of the blobs obtained are calculated using regionprops () function.

Finally the detected objects are labeled. Object labeling is done using pattern matching. Object in the video is labeled using the previously calculated measurements of the blob. The height and breadth of the each blob is calculated using the x and y coordinates which is calculated previously. Then a rectangle is plotted around the blob. The height and the breadth of the blob becomes the height and breadth of the rectangle and then the centre value of the rectangle is the centroid position of the blob. The area of the blob becomes the area of the rectangle. The blob area, height, width is selected in the given frame which gives the object in the video.

Performance evaluation may involve the computation of the discrepancy between the estimated and the ground-truth position and size of the target. The proposed measure quantifies how accurately and how long a target is tracked across a sequence. Tracking accuracy and the tracking failure are measured and using the values performance analysis is done Performance analysis compare performance different type of trackers. This method enables evaluating the robustness of the performance of tracking algorithms under such a wide variety of real-world conditions.

ALGORITHM DESIGN
Tracking Algorithm
Step 1: The input can be logical or a nonsparse numeric array of any class and dimension. The resulting array has the same size and class as the input.
Step 2: imfilter() function is used to remove noises from multidimensional array.

Step 3: The computation is performed using double precision floating point numbers.
Step 4: Frame difference (absolute) at time t + 1.
Step 5: The background is assumed to be the frame at time t.
Step 6: This difference image would only show some intensity for the pixel locations which have changed in the two frames.
Step 7: Then removed the background, this approach will only work for cases where all foreground pixels are moving and all background pixels are static.
Step 8: Compute threshold "Th" is put on this difference image to improve the subtraction.
Step 9: Identify accuracy of this approach is dependent on speed of movement in the scene. Faster movements may require higher thresholds.

VI. Conclusion

Thus the proposed system detecting target objects in video. The following methods are used to tracking objects from frames and analyze performance of different trackers , they are: (i) Pre-processing (ii) Background subtraction (iii)Object detection (iv) Labelling (v) performance analysis. These methods are giving support for video tracking.

References

[1]   B. Ristic, D. Clark, and B.-T Vo (2011), 'A metric for performance evaluation of multi-target tracking algorithms', IEEE Transactions on Signal Processing, pp. 3452–3457.
[2]   D. Comaniciu, V. Ramesh, and P. Meer (2003), 'Kernel-based object tracking', IEEE Transactions, pp. 564–577.
[3]   D. Scharstein and R. Szeliski (2002), 'A taxonomy and evaluation of dense two-frame stereo correspondence algorithms', International Journal of Computer Vision, pp. 7–42.
[4]   F. Bashir and F. Porikli (2006), 'Performance evaluation of object detection and tracking systems', Proceeding of IEEE International Workshop, pp. 7–14.
[5]   F. Yin, D. Makris, and S. A. Velastin (2007), 'Performance evaluation of object tracking algorithms', Proceeding of IEEE International Workshop.
[6]   H. Li, C. Shen, and Q. Shi (2011), 'Real-time visual tracking using sparse Representation', Proceeding of IEEE Conference, pp. 1305–13125.
[7]   J. Black, T. Ellis, and P. Rosin (2003) , 'A novel method for video tracking performance evaluation' , Proceeding of IEEE International Workshop.
[8]   L. M. Brown, A. W. Senior, Y.-L. Tian, J. Connell, A. Hampapur, C. f. Shu, H. Merkl, and M. Lu (2005), 'Performance evaluation of surveillance systems under varying conditions', Proceeding of IEEE International Workshop , pp. 1–8.
[9]   P. Pan, F. Porikli, and D. Schonfeld (2009), 'Recurrent tracking using multifold consistency',Proceeding International Workshop, pp. 1–10.
[10]   Pingali, S. and Segen, J. (1996), 'Performance Evaluation of People Tracking Systems', IEEE Workshop on Applications of Computer Vision, p33-38.
[11]   R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, J. Garofolo, R. Bowers, M. Boonstra, V. Korzhova, and J. Zhang (2009), 'Framework for performance evaluation of face, text, and vehicle detection and tracking in video: Data, metrics, and protocol', IEEE Transactions Pattern , pp. 319–336.
[12]   T. Nawaz and A. Cavallaro (2011), 'PFT: A protocol for evaluating video Trackers', Proceeding of IEEE International Conference on Image Processing.
[13]   T.J.Ellis (2002), 'Performance Metrics and Methods for Tracking in Surveillance', Proceedings of the Third International Workshop on Performance Evaluation of Tracking and Surveillance (PETS'2002).