Multiple Object Detection and Tracking in Cluttered Region with Rational Mobile Area

Pushpa D.
Asst. Prof: Dept of Information Science & Engg.
Maharaja Institute of Technology,
Mysore, Indian

H. S. Sheshadri
Prof.: Dept. of Electronics & Communication Engg
PES College of Engg,
Mandya, India

ABSTRACT

The proposed system highlights a novel approach of detecting and tracking multiple objects in the cluttered area like crowd using greedy algorithm. The proposed framework uses position of traced low-level feature points to generate a group of autonomous rational mobility region as resultant. Various challenging factors towards the accuracy of detection rate for multiple objects are considered. The proposed approach has detected all the feasible rational mobile regions and extorts the sub-group which increases a total likelihood function along with assignment of each traced locus to one mobile region. Performance analysis is carried out with different set of video sequences to find that proposed system has gradual robust detection rate as well as highly cost-effective computationally

Keywords- Component, object detection, greedy algorithm, feature selection

1. INTRODUCTION

Object tracking as well as object detection is one of the prime importances in the field of visual surveillance. Although with the course of prior research work, it has become almost a standard to treat object detection and tracking as separate procedure. Object detection can be represented as motion segmentation as well as object detection. Motion detection is utilized for segmentation of independent mobile regions which can be considered as computationally cost efficient solution, specifically for static mounted camera. Unfortunately, for the cluttered scenario like crowd, various challenging factors arises e.g. noisy background modelling, motion blobs based representation of observations, which cannot correspond to targets, causes severe challenges in tracking. The detection of an object is basically based on analytical framework or pattern of target's shape or morphology. It is considered that pattern based object detection is more efficient in comparison to illumination variation as well as camera mobility. But, rotation as well as scale invariant objects detection is normally computational demanding for tracking purpose.

If the object tracking can be represented as a sequences of shapes which are produced in the spatio-temporal space by the target, the crucial constraint will be in tracing the mobility smoothness, appearance and framework data of such a trace over a period of time. Therefore, the proposed system highlights the formulation of multiple object tracking in order to estimate a logical meaning of the forefront region with better smoothness. In the crowd, there are various problems which pose a serious development issues. Problems like inter-object occlusion, different object moving at all different speed, walking in arbitrary direction etc. will pose a huge challenge in the development as well as design of multiple object detection as well as tracking from the crowded scene.

In Section 2, we review recent work on multi-object detection in video. In Section 3, various problem descriptions is highlighted followed by proposed system in Section 4. Section 5 discussed about the implementation of the proposed system followed by results and performance analysis on section 6. Section 7 finally concludes by summarizing the entire research work.

2. RELATED WORK

Dimitrios Makris [1] with the non-trivial problem of performance evaluation of motion tracking. We propose a rich set of metrics to assess different aspects of performance of motion tracking. We 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. Max e.t. al [2] describes performance results from a real-time system for detecting, localizing, and tracking pedestrians from a moving vehicle. It achieves results comparable with alternative approaches with other sensors, but offers the potential for long-term scalability to higher spatial resolution, smaller size, and lower cost than other sensors. But issue of this work is it cannot segment people or objects in close contact. Jifeng [3] A novel object tracking algorithm is presented in this paper by using the joint color texture histogram to represent a target and then applying it to the mean shift framework. Niels Willems [4] present an integrated and multi-disciplinary approach for analyzing behavior of moving objects. This approach allows us to visually detect patterns in trajectory data by means of exploring the trajectory's attributes and additional attributes obtained by various web sources and reasoning.

3. PROBLEM DESCRIPTION

The situation of dynamic and messy visual scenes is one of the crucial problems in detection and tracking of multiple moving objects. It is specifically of highest importance in the field of visual surveillance, where the frame should capture all unforeseen event in crowded very fast and efficiently. Environment is another challenge factor. For mounted-automobile camera, it will be extremely tough to get a satisfactory background subtraction. The second challenging factor will be tracking agent in non-trivial scenario e.g. occlusion, which requires robust frame of agent. An agent can be deformable due to change of shape (human), orientation from camera view and scale change due to distance variation from camera [5]. The third issue will be the occlusion which will cause missing of agents making tracking with high false positive rate. Therefore it can be seen that although an extensive research has been conducted in object detection and tracking from cluttered visual scene, but a number of technical challenges still remain in the case of real world object detection and tracking.

4. PROPOSED SYSTEM

The main aim of the proposed work is to design a framework application for detection and tracking the object with highest accuracy. The system uses position of traced locus with low-level feature, which then yields results of cluster of regions with autonomous balanced motion. Analytically, the balanced motion is a spatial-temporal approach consisting of group of connected traced locus of feature. The problem focused in the proposed system is represented with a model of selecting a robust group of disjoint rational motional regions representing the unique moving object. The choice depends on a traced equivalency estimation which computes the probability that the entire generated curve within a rational motion region created from a unique object in the crowd using greedy algorithm. The proposed approach can be accepted as efficient algorithm which can be used for detection as well as counting huge quantity of equivalent mobile objects in the cluttered scenes.

5. IMPLEMENTATION

The proposed work is carried out in Matlab IDE with Windows OS of 32 bit, 1.86GHz and dual core processor. A video of .wmv format in the cluttered crowded location is considered as an input for this proposed system with spatial-temporal approach. The proposed algorithm initiates with a group of locus of low level of spatial feature traced over time captured using digital image capturing device. The traced locus ith of the feature is represented as Li.

$$L^i = \{(x_t^i, y_t^i), t = T_{init}^i, ..., T_{final}^i\}, \quad \text{i=1,. , Z} \quad (1)$$

Where, Z represents the total quantity of the traced locus. The extent of the locus traced will vary which depends on the time duration for which corresponding locus of features will be efficiently traced. The hierarchical-temporal approach of the proposed system will enable the tracing for locus of low level features of the input video frames. Based on the progress of the frames, all the upcoming mobile features of the object will be traced along with the previous traced locus to form a greater set of curve. The curve with significant temporal difference will only be preserved for the curve which has previous static frame fragment. A rough rectangle (w x h) is designed with an average match of the shape of the object in the crowd. The algorithm details are as below:

Algorithm: Object Detection and Tracking

Input: A video frame of finite length

Output: Object detection and tracking

Steps:

1 *Input a video and read the frames*;

2 *Read the size of frame;*

3. *[m n L]=size(vidFrames)*; // m=rows, n=Columns, L= Total No. of frames

4 *Initialize parameter for tracking in pixels;*

5 *Define selection window sizes;*

6 *Define tracking window sizes;*

7 *Define minimum spacing between two features (in pixels);*

8 *Define pixels around the screen (selection);*

9   *Define Minimum spacing between two features while appending;*

10   *Define threshold for the selection;*

11   *Define threshold for rejection of a point;*

12 *Define Minimum feature storage, Variation threshold, Height and width of rough dimension of moving object;*

13 *Estimate low level features and Tracking;*

14   *Create space for data collection;*

15   *Take first frame from gray scale dataset;*

16   *Read the size of image;*

17 *Store the size of the image in SampleSize;*

18   *Find Low level features*

19   *Increase the length of track to store it in space data collection*

20 *Save the first image for use in dissimilarity calculations*

21 *Display the down_sampled images*

22   *Compute downsampled images*

23   *Overwrite the points on image*

24   *if (size(coord,1) > 0)*   // If features are found

25   *plot(coord(2,:), coord(1,:), 'g.');*

26   *end*

Algorithm: Feature Selection

Input: Grayscale image, Selection window size X-direction, selection window size Y-direction, threshold ratio multiplier for the selection of pixels greater than the threshold*(maximum of the minimum eigenvalues for each pixel)

Output: The trackable points selected by the feature selector.

Steps:

1 *Take the 2-D gradient of the image (gx, gy)*

2 *Perform 2D convolution on (gx,gy)*

3 *Apply 2-D median smoothing mask*

4 *Compute the eigenvalues of each pixel using the matrix*

5   *Find the minimum eigenvalue at every pixel of the image*

6 *From those pixels, retain the local max pixels within the 3x3 neighborhood*

7 *Keep the subset of these pixels whos minimum distance between any pair of pixels*

8   *Find the non -zero values, Get the column value, Get the Row value, Calculate the distance between X coordinates*

9 *Estimate the unique features*

The proposed algorithm with spatial-temporal approach will be design in such a way that it should be independent of deployment of object-specific contour models as well as it will dependent on identification and integrating object specific components too. The proposed system is also free from calibration of image capturing device. The proposed system can be found to be computationally efficient and appropriate for camera network applications which require large number of expensive visual sensors. The proposed technique at every iteration has explicit use of greedy algorithm which compels the constraints that chosen rational mobility locations be a disjoint.

6. RESULTS & PERFORMANCE ANALYSIS

The performance analyzation of the proposed system has been conducted using different frames of videos which comprises of 9 people walking on street with everyone following different speed of walking.
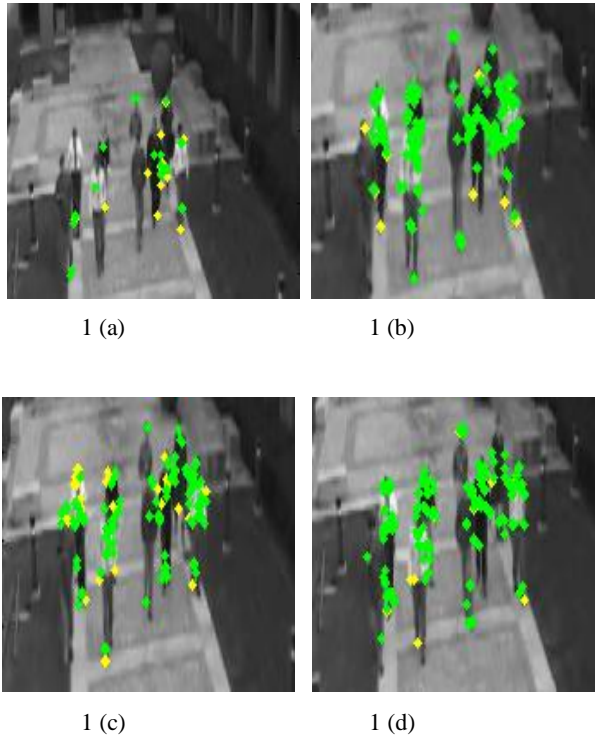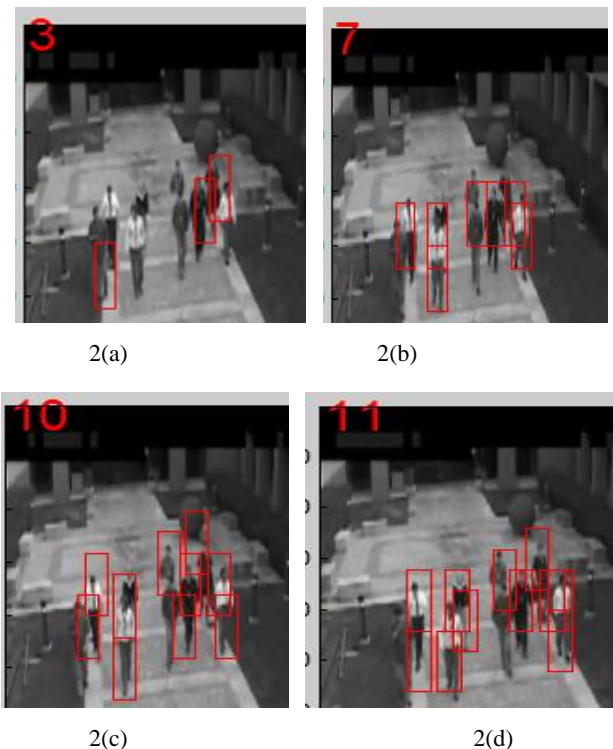

Fig 1. Traced Locus of Features

Fig 2. Object Detection and Tracking

The execution of the application highlights all the traced locus of features as shown in green dots in Fig 1. and Fig 2 represents detection and tracking of the same experimented objects. The detection and tracking efficient of the algorithm is analyzed when some of the object will start moving at different pace. Therefore, it can be seen in Fig 2, although in 2(a), only 3 people are detected leaving the remaining 6 people not detected. But with gradual spatial-temporal approach (as shown in Fig 3), it can also be seen that frame rate 7 is shown to detect 7 and final frame 11 shows to detect with 100% accuracy.
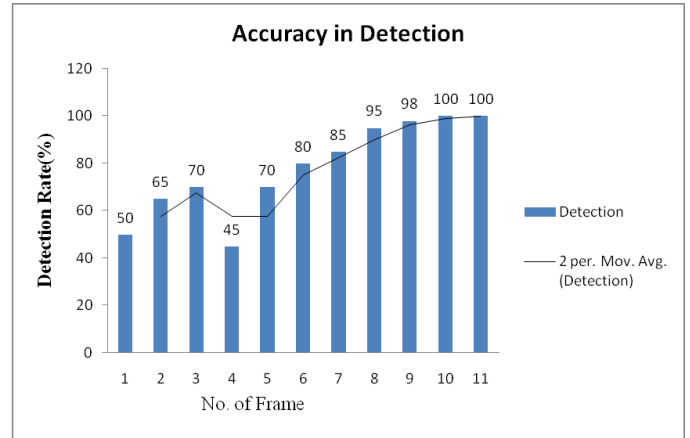

Fig 3. Accuracy in Detection (No. of Frames Vs. Detection Rate)

The algorithm processes feature locus curve generation in fast speed and is also as good as captured frame rates of the input video. The source code is designed in Matlab without using any external optimization tools where the run time of the each of the experimented video is seen to start in less than 10 seconds.

7. CONCLUSION

The proposed system introduces a cost efficient technique of distinctive detection as well as tracking of multiple object in the cluttered area like crowd. The proposed work could also be used in most challenging scenarios of inter-object occlusion for object-counting as well as object localization . The system captures the data originated from low level features in order to design all feasible rational mobility oriented scene and then select a robust disjoint group of rational mobility scene exhibiting unique object by deploying greedy algorithm. For checking the efficiency of the algorithm, various sets of experimented video sequences were used to show the detection rate with 100% accuracy. However, certain limitation also exist like, the model recognizes two-conjoined objects as one. E.g. A mother carrying kid on lap may be counted as one. However, the future work could focus on removing such minor false positive in object detection system.

8. REFERENCES

[1] Fei Yin, Dimitrios Makris, Sergio Velastin, Performance Evaluation of Object Tracking Algorithms, In 10th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS2007), Rio de Janeiro, Brazil (October 2007)

[2] Max Bajracharya, Baback Moghaddam, Andrew Howard, Shane Brennan, Larry H. Matthies, Results from a Real-time Stereo-based Pedestrian Detection System on a Moving Vehicle, Proceedings of the IEEE ICRA 2009 Workshop on People Detection and Tracking Kobe, Japan, May 2009

[3] Jifeng Ning, Lei Zhang, David Zhang, Chengke Wu, Robust object tracking using joint color-texture histogram, International Journal of Pattern Recognition and Artificial Intelligence Vol. 23, No. 7 (2009)

[4] Niels Willemsa, Willem Robert van Hageb, Gerben de Vriesc, Jeroen H.M. Janssensd, Veronique Malaiseb, An integrated approach for visual analysis of a multi-source moving objects knowledge base, International Journal of Geographical Information Science Vol. 24, No. 9, September 2010, 1–16

[5] Pabboju Sateesh Kumar, Multi-agent tracking under occlusion and 3D motion interpretation, Doctorial Thesis, Aug-2006