INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN: 2278 – 0211 (Online)

Detecting Number Of Drift Objects In Packed Surroundings Using Coherent Motion Region

Miss.Rajashree T. Gadhve
ME (Comp. Engg), MEM's CET, University of Mumbai, India
Prof. Vidya Chitre
H.O.D. (Comp. Dept.), Bharti Vidyapeeth's COE, Navi Mumbai, India
Prof. Vijay Bhosale
Computer Dept.,MGM's CET, University of Mumbai, India

Abstract:
The main approach of an object detection system is to find all the possible coherent motion regions to track the moving objects which takes the locations of low-level tracked feature points as input, and produces a set of independent coherent motion regions as output. In past approaches only pedestrian is detected using a source of information in a single detector but in this, the system finds the regions based on the tracked set of features using a likelihood functions which is parameterized on locations of potential individual. For the identified coherent motion region assign a point track to at most one region to extract the subset that maximises an overall likelihood function. In case of multi-object motion, many possible coherent motion regions can be constructed around the set of all feature point tracks. The approach is robust to partial occlusion, shadows, clutter, and can operate over a large range of challenging view angles. This approach gives semantically correct detections and count of similar objects moving through crowded scenes from selected coherent regions using greedy algorithm.

Key words: *Greedy Algorithm, Coherent Motion, Moving Object detection, Low Level Feature, Multi-object detection, Trajectory, Feature Tracking*

1.Introduction

Detection of individuals in the crowded environment is comparatively little attention in vision because it gives the problem of segmentation, recognition and tracking. There are some techniques which handle larger numbers of entities, but generally struggle is initialized when the crowed is dense. There are some sophisticated person detectors such as [1] has been developed, which assumed that people are well separated.

The others [2][3] used mechanism such as head detectors to segment crowds. Failure is occurred when these features is not get observed. One technique is used called model based segmentation in which it extracts a large set of image features and partition into individuals using explicit feature grouping. The one described [4] in experiment that the ability of humans to distinguish activities and count independent motions simply form 2D projections of a sparse set of feature points manually identified on human joints. This is applied on the video segments it automatically extracted features that is bright dots against a dark background, from which human observer can easily able to detect and classify the moving objects. This helps to develop an object detection system which can detect and count independently moving objects based on feature point trajectories alone. A coherent motion region contains a group of point tracks; but a single moving object corresponds to a single coherent motion region. Based on others [5][6] likelihood function they define a system which identifies whether a given set of features should be grouped together or not. Using the greedy search the segmentation is achieved were the most likely grouping is first identified and then removed. They used a variant of expectation-maximization algorithm for the estimation of shape parameters from image observations via hidden assignment vectors of features to cliques. The global optimization is desirable when the likelihood functions do not consider for all images feature simultaneously, because of high ambiguity the difficulties can occur with the local context. This can occur in regions such as the centre of dense crowd.

To detect drift object in that to detect human is difficult, because people can move fast. One uses the dynamic model and configuration in the current frame to predict the next configuration; all predictions are then being refined using image data. This model is identify the humans on the basis of not to change the “appearance” from frame to frame. It describes a people tracker which builds models of the people to be tracked from video sequence then tracks them. It has the advantage that knowing the appearances model of each body part greatly constrains our search and so simplifies data association. Other is it prevent drift, recover from occlusion relatively easily, and count individuals. When the

Parameter driven generative model is concerned, the various parameters are used to perform segmentation. The parameters may include the number of people, their location and their shape. Based on the set of given parameters the grouping of all features are evaluated. The global optimization is achieved by searching for maximum likelihood estimates for the model parameters. But the issues with the parameters are, optimization requires good initial estimates and expensive search techniques. By assuming some prior knowledge [7] regarding the number of people and their appearance the optimization is achieved by exhaustive local search [8].

Figure 1 : Different view angle samples shows partial occlusion

The next approach is based on power generative models with the simplicity associated with the feature grouping methods is needed. In this there is no assumption is required about the number of people in the scene, needs only trivial initialization and does not require random search. The main approach is based on the Expectation Maximization (EM) formulation which has shape parameters for all potential individuals and treats feature assignments as hidden variables [9].

Other technique in which detecting multiple humans in crowded environment is based on the model based approach, it interpret the image observations by multiple partially occluded human hypotheses in a Bayesian framework. The differ features like knowledge of various aspects, human shape, camera model, and image cues are all integrated in one framework. The tracking of humans is obtained by using an efficient sampling method, data-driven Markov Chain Monte Carlo (DDMCMC), which uses image observations for proposal probabilities. It uses the direct image feature from bottom-up image analysis to improve the computational efficiency as importance proposal probabilities to guide the moves of the Markov chain. The work includes A 3D part based human body model which enables the segmentation and tracking of humans in 3D and the interference of inter object occlusion naturally, A Bayesian framework that integrates segmentation and tracking based on a joint likelihood for the appearance of multiple objects, the Markov dynamics, directed by proposal probabilities based on image cues, and the incorporation of a colour-based background model in a mean-shift tracking step.

Based on human detectors it detects appearance or shape-based patterns of humans directly [10][11][12][13]. The learning base method need a large number of training samples and may be sensitive to imaging viewpoint variations. Besides motion and shape, face and skin colour are also useful cues for human detection, but these cues utilized in environment could be limited. This method is limited to blob tracking. The model based tracking can solve the problem of blob merge and split problems by enforcing a global shape constraint. The shape model could be parametric, ellipsoid [14] or nonparametric, edge template [13], and either 2D or in 3D [14] [15].

Parametric models are usually generative and of high dimensionality while nonparametric models are usually learned from real samples. For 2D model makes detailed hypothesis but in 3D models are more natural for occlusion reasoning. For human tracking we do not need to capture the detailed body articulation, a rough body model such as the generic cylinder in [16], the ellipsoid in [14], and the multiple rectangles in [15] suffices. For tracking multiple objects require matching hypotheses.

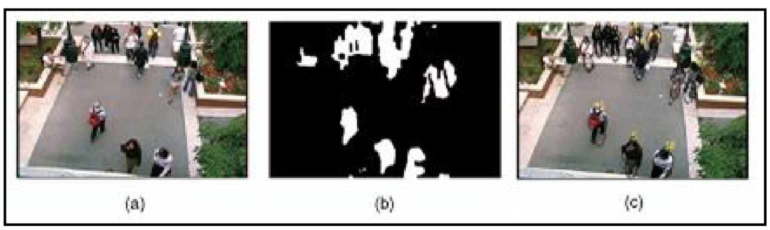

Figure 2 : a) Sample Frame b) Motion blobs c) tracked people for crowed situation

When objects are highly inter occluded, their image observation are from independent, hence a joint likelihood for multiple objects is necessary [15] [17] [16][19] [20] [21]. The data driven MCMC method used for various applications, for multi object tracking [22] estimating articulated structures. The main difference between MCMC and other method is it uses 3D perspective effect in a typical camera setting, while the ant tracking problem described I [22] is almost a 2D problem. It utilize the acquired appearance where each object is of different appearance, while ants in are assumed to have the same appearance.

2.The System Architecture

The system architecture is shown in figure. It is focused on an effective algorithm for counting coherent moving objects in dense crowd. The architecture includes different block set such as video acquisition, feature extraction, tracking trajectory similarity, coherent motion detection. It counts the total number of moving objects in a video from different video streams with dense crowed.

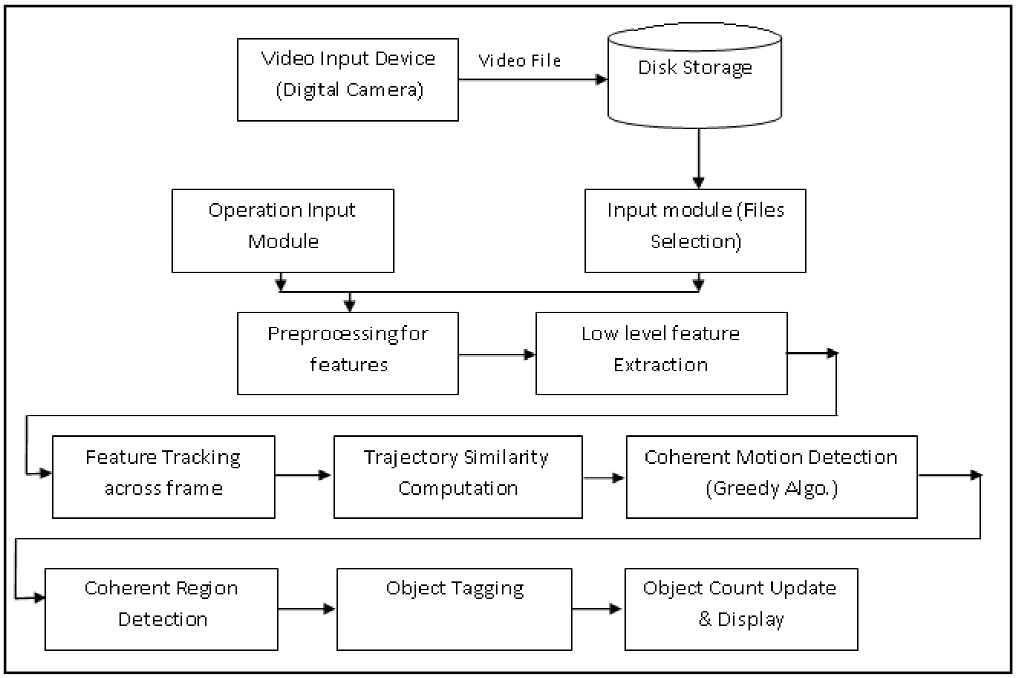

Figure 3: Project Architecture

3. Detailed System Description

3.1.Pre-processing

A set of basic pre-processing algorithms from Image processing are applied to individual frames of the video considered for object tracking and counting. These algorithm can be proposed for situation like– objects with internal motion (object shape variation is high), Objects with low internal motion (the shape variation is low), Handling of variation in lighting conditions, corner detection.

3.2.Low Level Feature Extraction

A set of low-level spatial feature points tracked over time through a video sequence. define the i^{th} feature point track by X^i:

$X^i = \{(x_t^i, y_t^i), t = T_{init}^i, \ldots\ldots\ldots\ldots, T_{final}^i\},$

$i = 1, \ldots\ldots\ldots\ldots, Z$

where, Z represents the total number of point tracks. The lengths of the tracks vary depending on the durations for which corresponding feature points are successfully tracked we first identify low-level features in the initial frame detector, a fast algorithm for finding corners. The low level features are tracked over time using a hierarchical implementation of the Kanade-Lucas-Tomasi optical flow algorithm.

3.3.Feature Tracking

The new features are tracked along with the existing point tracks to form a larger trajectory set. For trajectories that have initially stationary segments, we retain only the remaining part of the trajectory that shows significant temporal variations. The user is also required to sketch a single rectangle that matches the rough dimensions of the objects to be detected in the sequence. Let the dimensions of this rectangle be w x h.

3.4.Trajectories Similarities

It requires a measure of similarity between two feature point tracks. If both X^i and X^j exist at time t, we define $d_t^x(i,j) = (x_t^i - x_t^i) \left(1 + \max\left(0, \frac{|x_t^i - x_t^j| - w}{w}\right)\right)$

$d_t^y(i,j) = (y_t^i - y_t^j) \left(1 + \max\left(0, \frac{|y_t^i - y_h^j| - h}{h}\right)\right)$

$D_t(i,j) = \sqrt{d_t^x(i,j)^2 + d_t^y(i,j)^2}$

That is, if the features do not fit within a w x h rectangle, the distance between them is nonlinearly increased. Our expectation is that feature point tracks from the same underlying object are likely to have a low maximum D_t as well as a low variance in D_t over the region of overlap. Hence, we compute an overall trajectory similarity as

$S(i,j) = \exp\{-\alpha\ (\max(D_t(i,j)) + var(D_t(i,j))\}$

where the maximum and variance are taken over the temporal region where both trajectories exist. For those trajectories where there is no overlap the similarity value is set to zero. The pair wise similarities are collected into a Z x Z matrix S. In our experiments below, we set α as 0.025.

3.5.Coherent Motion Region

A coherent motion region is a spatiotemporal sub volume that fully contains a set of associated feature point tracks. In our case, each coherent motion region is a contiguous chunk of (x, y, t) space that completely spans the point tracks associated with it. Note that all the feature point tracks inside this region might not have complete temporal overlap. The set of all coherent motion regions can be represented by a binary Z x M matrix A indicating which point tracks are associated with each coherent motion region.

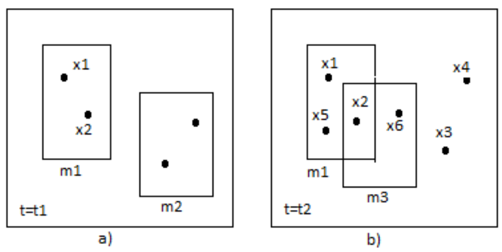

Fig 4: a) Candidates coherent motion regions for a given set of tracks

There is an M-vector L with the set of coherent motion regions, indicating a "strength" related to the overall likelihood that the coherent motion region corresponds to a single object, created by

$L(j) = A(j)^T SA(j)$

Where A(j) is the j^{th} column of A.

3.6.Selecting the Coherent Motion Region Subset

Select a subset V of coherent motion regions that maximizes the sum of the strengths in V, subject to the constraint that a point track can belong to at most one selected coherent motion region. It uses a greedy algorithm to estimate a good subset V. This approach differs from a soft-assign approach of assigning tracks to cliques. The greedy algorithm explicitly enforces the constraint that selected coherent motion regions be disjoint on per iteration basis. In our approach it starts with highly connected graph in which edges are progressively removed and the sum of edge weights updated. The objects which are visible for some time for that object it is easy to find the best coherent motion regions. From each frame it detects new feature points; a longer-duration object has a greater opportunity to gather feature tracks. Selecting a high likelihood coherent motion region in the greedy algorithm results in lowering the likelihood values of other coherent motion regions containing tracks that are part of the selected coherent motion regions.

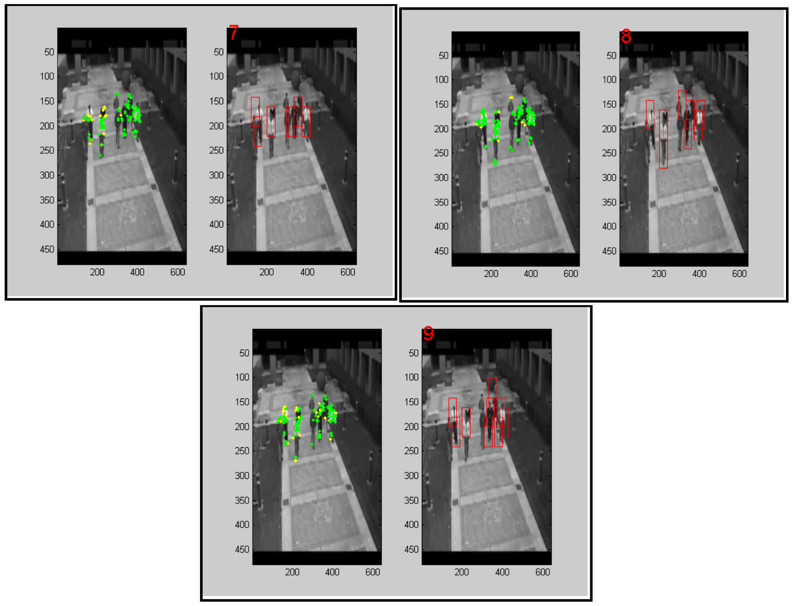

Figure 5: Sample result for a video sequence

The objects which are visible for some time for that object it is easy to find the best coherent motion regions. From each frame it detects new feature points; a longer-duration object has a greater opportunity to gather feature tracks. Selecting a high likelihood coherent motion region in the greedy algorithm results in lowering the likelihood values of other coherent motion regions containing tracks that are part of the selected coherent motion regions.

4. Conclusion

The object detection system uses approach based on coherent motion region detection for counting and locating objects in the presence of high object density and inter-object occlusions. The system tracks low-level features to construct all possible coherent motion regions, and choose a good disjoint set of coherent motion regions representing individual objects using greedy algorithm. Any object with the similar shape to a person may be misclassified. This problem is handles by investigating using appearance and motion features to select disjoint coherent region. The feature point trajectory generation is fast and comparable with the frame rate of video sequence.

It does not require any complex shape or appearance models for objects. Tracking, operating in a 2-frame interval, has a local view therefore ambiguities inevitably exits. It occurs in case of tracking multiple close-by or overlapping objects. The analysis in the level of trajectories resolves the local ambiguities. The analysis considered into the account the prior knowledge on the valid object trajectories including their starting and ending point. It is effective technique for counting coherent moving objects over a variety of scales and challenging camera angles in the dense crowd.

5.Reference

1. D. Gavrila. Pedestrian detection from a moving vehicle. In Proc. 6th European Conf. Computer Vision, Dublin, Ireland, volume 2, pages 37–49, 2000.
2. I. Haritaoglu, D. Harwood, and L. S. Davis. HYDRA: Multiplepeople detection and tracking using silhouettes. In In IEEE International Workshop on Visual Surveillance, pages 6–13, 1999.
3. T. Zhao and R. R. Nevatia. Bayesian human segmentation in crowded situations. In IEEE Computer Vision and Pattern Recognition, Madison, Wisconsin, volume 2, pages 459–466,2003
4. G. Johansson. Visual motion perception. Scientific American, 14:76–78, 1975.
5. C. Schmid K. Mikolajczyk and A. Zisserman. Human detection based on a probabilistic assembly of robust part detectors. In Proc. 8th European Conf. Computer Vision, Prague, Czech Republic, volume 1, pages 69–82, 2004.
6. Y. Song, L. Goncalves, and P. Perona. Monocuolar perception of biological motion - clutter and partial occlusion. In Proc. 6th European Conf. Computer Vision, Dublin, Ireland, volume 2, pages 719–733, 2000.
7. A. Elgammal and L. Davis. Probabilistic framework for segmenting people under occlusion. In Proc. 8th Int. Conf. on Computer Vision, Vancouver, BC, volume 2, pages 145–152, 2001.
8. T. Zhao and R. R. Nevatia. Bayesian human segmentation in crowded situations. In IEEE Computer Vision and Pattern Recognition, Madison, Wisconsin, volume 2, pages 459–466, 2003.
9. Z. Tu and A. Yuille. Shape matching and recognition – using generative models and informative features. In Proc. 8th European Conf. Computer Vision, Prague, Czech Republic, volume 3, pages 195–209, 2004.
10. B. Wu and R. Nevatia, “Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors,” Proc. 10th Int’l Conf. Computer Vision, vol. 1, pp. 90-97, 2005.
11. C. Papageorgiou, T. Evgeniou, and T. Poggio, “A Trainable Pedestrian Detection System,” Proc. IEEE Intelligent Vehicles Symp., pp. 241-246, 1998.
12. L. Davis, V. Philomin, and R. Duraiswami, “Tracking Humans from a Moving Platform,” Proc. Int’l Conf. Pattern Recognition, vol. 4, pp. 171-178, 2000.
13. D. Gavrila and V. Philomin, “Real-Time Object Detection for “Smart” Vehicles,” Proc. Seventh Int’l Conf. Computer Vision, vol. 1, pp. 87-93, 1999.

14. T. Zhao and R. Nevatia, “Tracking Multiple Humans in Complex Situations,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 9, pp. 1208-1221, Sept. 2004.
15. H. Tao, H. Sawhney, and R. Kumar, “A Sampling Algorithm for Tracking Multiple Objects,” Proc. Workshop Vision Algorithms, 1999.
16. M. Isard and J. MacCormick, “Bramble: A Bayesian Multiple-Blob Tracker,” Proc. Eighth Int’l Conf. Computer Vision, vol. 2, pp. 34-41, 2001.
17. J. MacCormick and A. Blake, “A Probabilistic Exclusion Principle for Tracking Multiple Objects,” Proc. Seventh Int’l Conf. Computer Vision, vol. 1, pp. 572-578, 1999.
18. C. Rasmussen and G.D. Hager, “Probabilistic Data Association Methods for Tracking Complex Visual Objects,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 560-576, June 2001.
19. K. Okuma, A. Taleghani, N. de Freitas, J. Little, and D. Lowe, “A Boosted Particle Filter: Multitarget Detection and Tracking,” Proc. Eighth European Conf. Computer Vision, vol. 1, pp. 28-39, 2004.
20. T. Yu and Y. Wu, “Collaborative Tracking of Multiple Targets,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 834-841, 2004.
21. Z. Khan, T. Balch, and F. Dellaert, “MCMC-Based Particle Filtering for Tracking a Variable Number of Interacting Targets,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 11, pp. 1805-1819, Nov. 2005.
22. M.-W. Lee and I. Cohen, “A Model-Based Approach for Estimating Human 3D Poses in Static Images,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 6, pp. 905-916, June 2006.