Precise multiple object identification and tracking using efficient visual attributes in dense crowded scene with regions of rational movement

Pushpa D.[1], Dr. H.S. Sheshadri[2]
[1]*Asst. Prof*: Dept of Information Science & Engg., Maharaja Institute of Technology
Mysore, Karnataka, India

[2]*Prof*. Dept. of Electronics & Communication Engg., PES College of Engg.
Mandya, Karnataka, India

Abstract— **The proposed model represents a unique technique for detection and tracking multiple objects from a dense cluttered area like crowd by deploying greed algorithm. Understanding the complexity of deploying various image attributes e.g. edge, color etc, the proposed system will illustrate cost effective and robust procedure of using low-level attributes which takes very less computational time in order to produce autonomous rational mobility region as resultant. The technique also considers various difficult real-time scenarios in the dense crowd in order to design a highly cost effective algorithm. 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- Object Detection, Dense Crowd Attribute Selection

I. 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 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. Majority of the video monitoring techniques are normally experimented with permanent image capturing device which is comparatively unproblematic in terms of real time application in order to accomplish image capturing parameters. Therefore, the current research work has been experimented with monocular adjusted image capturing device, which facilitates to design a distinctive connection between the subjects walking in real time and image plane. The phenomenon is considered for isometric view of the subjects walking. The concept of the experiment is chosen for isometric view-point has two factors; firstly, isometric view will capture the feeds with partial occlusion and secondly the feeds will be three dimensional in nature which will permit the algorithm for more accuracy in tracking the objects (or subject). As a result, the isometric view will signify the fundamental plane where the object related magnitudes as well as their attribute and patterns can be fused. The experiment will be free from dynamic camera calibration as the height of the camera, focal angle, and the angle of tilt with the subject and plane will be pre-assumed. Another prominent consideration is velocity of the movement of the object. The experiment is conducted with a video file where the movements of the objects are not uniform, i.e. each objects move at their own pace. This consideration is important as majority of the pedestrians in real time have slight variations in their speed of walking along with different actions sometimes. In Section 2, we review recent work on multi-object detection in video. In Section 3, fundamentals of object tracking in dense crowd is discussed followed by illustration of attribute selection on dense crowd in section 4. In section 5, various problem description is highlighted followed by proposed system in Section 6. Section 7 discussed about the proposed system followed by results and performance analysis on section 8. Section 9 finally concludes by summarizing the entire research work.

II. RELATED WORK

The work on object detection and tracking is not very new. There has been a dozens of research work done in past for the similar purpose, which unfortunately is not found with much effective and satisfying results in real time applications. In this context, the research work on pedestrian simulations conducted in past can be referred. An extensive work has been done to show the pedestrian movement detection where the majority of such work is targeted at building evacuation where the direction of object movement is quite predictable. Some of the prominent work done in pedestrian movement detection and tracking in terms of simulation are as follows:

- PEDSIM: It is application design for evaluating infinitesimal pedestrian and dense crowd simulation. This application uses social force framework used in Helbing and Moln´ar [1].
- NOMAD: It is another simulation tool designed by Transport & Planning department of the Delft University of Technology, which is designed on minuscule individuality of the moving objects like velocity of walking and size of objects. [2]
- EXODUS: It is another framework design in past for replicating the evacuation of great quantity of objects from predefined locations [3].
- SIMULEX: This is another model for replicating the section of IES Virtual Environment, which is a group of tools designed for assisting in the development and assessment of predefined location [4].
- EVAS: This is another object detection and tracking software especially designed for pedestrian based on Visibility Graph Analysis [5].
- EVACSIM: Designed in Java, this is another model developed for understanding patterns of evacuation by objects [6].

Therefore, it can be seen that there has been couple of simulators designed for understanding the object detection and tracking with respect to pedestrians, showing a higher scope in conducting research in the similar field. The data associated with the mobile objects has become one interesting point of research for evaluation of detection and tracking system in dense crowd scenario. Various types of parameters like shape, texture, and color are evaluated to have uniform visual stream when considered homogenous regions. The previous research has witnessed from [7] for evaluating rational movement of the groups of pixels of uniform color, where every image elements belongs to a defined group with specific feasibility and every group if categorized deploying non-deterministic framework of movement. Polana and Nelson [8] has adopted spatial-temporal discretization which is allocated to every region of image, accomplished normally by overlapping a grid on the image and its mean visual stream. The episodic movement prototype is estimated by performing Fourier analysis on the resulting attribute vector. The use of movement field by Shio and Sklansky [9] has shown the accomplishment by correlation technique over consecutive image frames, which results in identification of movement regions with uniform direction when smoothening is done along with quantization of image regions. Deployment of clusters of pixels has been done by Heisele et al. [10] to be attained by different grouping methods as fundamental constituents for tracking. The grouping analysis is performed by both spatial data and color information. The concept behind this is that appending spatial data generates grouping more steady than deploying only color data. There has been also abundant use of k-means algorithm while working on clustering techniques as evident in Cutler and Davis [11] for designing a change detection algorithm. The prototype also facilitates a map of the image elements signifying moving subjects along with precise thresholding. Use of wavelets transform has been seen in the work of Papageorgiou and Poggio [12] for characterizing a moving object shape and then examines the transform technique of the image to evaluate the movement patterns. Implementation of two dimensional contour shape analyses is seen in the work of Wren et al. [13] which tends to recognize different parts of body of the moving subjects using heuristics techniques. Cai and Aggarwal [14] proposed a model with a uncomplicated head-trunk replica to trail humans across multiple image capturing devices. Beymer and Konolige [15] have worked for fitting a trouble-free shape on the candidate image. Dimitrios Makris [16] 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 [17] 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 [18] 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 [19] 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. Table 1 shows all the prominent research work being done.

Table-1: Prior Research Work

Year	Authors	Problems Focused	Approaches Used	Results Obtained
2005 [20]	Ernesto Robert	Dense crowd detection problems	Pedestrian Path, Body, Motion Model	Satisfactory results but depended on costly training module.
2006 [21]	Ernesto Robert	Detection of abnormal events in crowds	Hidden Markov Model	Reliable Detection in crowd
2007 [22]	Sim Surendra	Reducing False Alarms in crowd	Perspective modeling	Only relies on the 2D image size
2009 [23]	Min Zhang Huang Tan	Detection of omega-shape attribute of people's head-shoulder parts.	Haar attribute, AdaBoost Classifier Histograms of Oriented Gradients	Rapidly detect and robustly track people's head-shoulders, not effective for isometric views.
2009 [24]	Chan Morrow Vasconcelos	Analysis of Crowded Scenes	Support vector machine Use of low-level attribute Perspective normalization Attribute extraction Gaussian process regression Event Classification	Good event detection results but highly depends on expensive training module
2009 [25]	Wang, Wu, Lien	Pedestrian Detection	Trilinear interpolation technique Gaussian-weighted window Cascaded AdaBoost detector.	Better detection performance Significant reduction in the computational time
2010 [26]	Dong, Liu, Li	Crowd Density Estimation	Gray-Gradient Dependence Matrix Sparse texture attribute	the technique of attribute extraction in this work only considers single images proposed approach is just tested indoors
2011 [27]	Rodriguez Laptev	Object detection in Dense crowd	energy minimization technique	Cannot jointly detect and track people in multiple frames.
2011 [28]	Sirmacek Reinartz	Tracking people from Airborne images	Kalman Filter Based attribute Analysis	Not enough precise results for detection or tracking.
2011 [29]	Xing Ai Liu Lao	Object detection in Dense crowd	Flow Guided Detection Responses Collection Pair-wise Association for Flow Growing	No consideration of occlusion
2011 [30]	Luvison Chateau	Automatic Detection of Unexpected Events in Dense Areas	supervised approach generic learning approach	temporal and spatial consistency not considered
2011 [31]	Riche Mancas	Abnormal motion selection in crowds	social signal processing, Bottom-Up Saliency	Precise tracking not possible
2011 [32]	Rodriguez Sivic	Data-driven Crowd Analysis	linear Kalman filter	More focused on crowd behaviour
2011 [33]	Sirmacek	automatic crowd analysis from very high resolution satellite mages	Probability theory, Shadow extraction, Road extraction	Still in infancy stage

III. OBJECT TRACKING IN DENSE CROWD

Object tracking is a prominent and complex issue which gives rise to various interest in research work towards computer vision with an purpose of accomplishing association between objects and object parts between consecutive frames of video. This phenomenon is considered as prominent evaluation in majority of the real-time monitoring system as it facilitates consistent temporal information related to mobile objects which are deployed both to improve lower level processing like segmentation of mobility and to facilitate greater score of extracting information like analysis of events as well as behaviour identification. However, it is very challenging situation to apply object tracking in dense crowded area due to imprecise object segmentation. Common issues in object tracking in crowded area explored are silhouette, incomplete and full occlusion of moving objects overlapping with one another (sometimes with stationary objects too). Therefore, mitigating silhouette at movement recognition stage and catering with partial or full-occlusion both at object segmentation stage and tracking phase is significant for strong tracking.

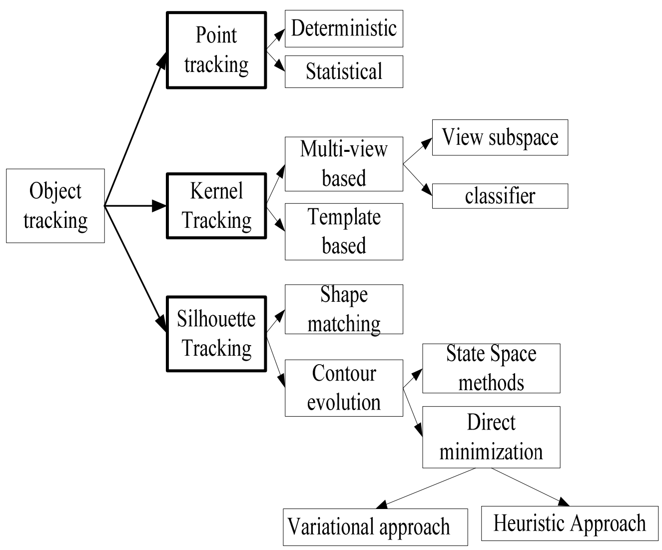


Figure 1: Classification of tracking methods

Figure 1 highlights the classification of main tracking techniques frequently used in the research work, which are further explained below:

- **Point Tracking**: Points are used for signifying objects identified in successive image frames and the connection of the points is based on the prior state of object possessing object location and movement information. This phenomenon often depends on the external technique to identify an object in every frame. Such types of procedure can be implemented either using probabilistic approach like Kalman filtering technique [34] or deterministic methods [35].
- Kernel Tracking: The shape of the object and its respective appearance are signified by Kernel. Various experiments uses kernel as elliptical shape or rectangular template in relation with histogram. The movements of the kernel in successive image frames are used for tracking the mobility of object in dense crowd area. Such types of mobility is normally in the representation of a parametric transformation like rotation, translation, rotation and affine. Mean-shift tracking technique is another suitable implementation [36].
- Silhouette Tracking: In this type of tracking technique, the mechanism is conducted by approximating the regions of object in every individual frame. The data embedded inside the region of object is deployed for such types of tracking. Data can be in representation of appearance density as well as shape models which are normally of the type of edge maps. Silhouettes are identified and followed by matching shapes or by evolution of contours with predefined model of object. Various active contour models [37] are used in such types of tracking.

IV. ATTRIBUTE SELECTION FOR TRACKING

The selection of the precise attribute plays a prominent role in tracking objects in dense crowd. Normally, the highly required characteristic of a visual attribute is its distinctiveness so that the objects (or subject) can be swiftly distinguished in the attribute space. The representation of an object is very much associated with the attribute selection. It can be seen that color attribute is considered as attribute for histogram based exterior representation and edges are considered as significant attribute for curve based representation. But in general, different object tracking protocols deploy a amalgamation of such attributes where some frequently used image attributes are as follows:

- Color: The perceptible color of an object in a crowd is highly affected by a) spectral energy distribution of illuminant and b) the plane reflectance charecteristics of an object. A wide invariants of color spaces have been deployed in tracking system in crowd.(e.g. RGB, HSV)
- Edges: It has been observed that a boundary of an object normally shows potential alterations in image intensities. One of the significant charecteristics of an edge is that they are minimally sensitive to intensity of light changes dynamically in comparison to color attributes. Protocols or a technique which track edges of an object normally deploys this as delegate attribute.
- Optical Flow: This is a impenetrable field of certain displacement vectors which illustrates the rendition of every pixels in a region. Such technique is frequently deployed as an attribute in majority of movement based segmentation and tracking application.
- Texture: This is another attribute which is an estimate of light strength deviation of a surface which measures charecteristics like efficiency and reliability. Certain efficient processing phase is highly demanded in accomplishing texture descriptors. This attribute is also

minimally sensitive to light intensity alteration in color like edge attribute.

Out of all the attributes discussed, color is one of the frequently deployed attribute for tracking. Unfortunately, majority of the bands of colors are highly perceptive to the light intensity deviation. Therefore in environment of experiments, where such phenomenon is unavoidable, other attributes are implemented for modelling object appearance.

V. 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 [38]. 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.

Figure 2: Overestimations (red) Vs false positives (Green)

The current research work will aim to design a tracking system in dense crowd which is able to count the appropriate quantity of objects targeted. Such phenomenon will also give rise to overestimation issue but not alike to false positive issue in terms of detection and tracking systems. Fundamentally a false positive is distinguished as a trailer located on region of an image which does not relate to object targeted, which can be a surroundings area or another subject which is not required to be identified. It was also seen that a target overestimation normally takes place when the object to be identified is implicated to a multiple detection which normally give rise to multiple trackers, and bad news is majority of them are not false positives at all. This is a very much challenging situation to overcome. In terms of examples of an object walking on the street and when an algorithm for tracking is implemented, the overestimation of the objects normally produces multiple trajectories for the same object. Figure 2 illustrates the description of overestimation.

VI. 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 attribute, 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 attribute. 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

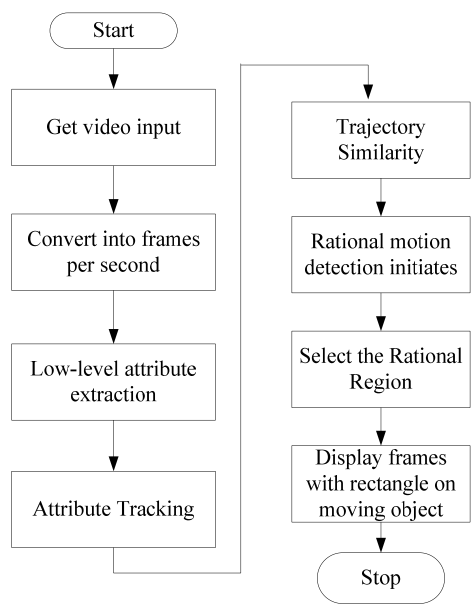


Fig 3. Initial phase Implementation

The application takes the input of the video file which is then converted to frames per second for ease in analyzation. Low level attribute point extraction takes place which is subsequently followed by attribute tracking. After the previous

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 2, March 2012
ISSN (Online): 1694-0814
www.IJCSI.org

step is accomplished trajectory similarity is evaluated and finally rational motion detection is estimated as shown in Fig 3.

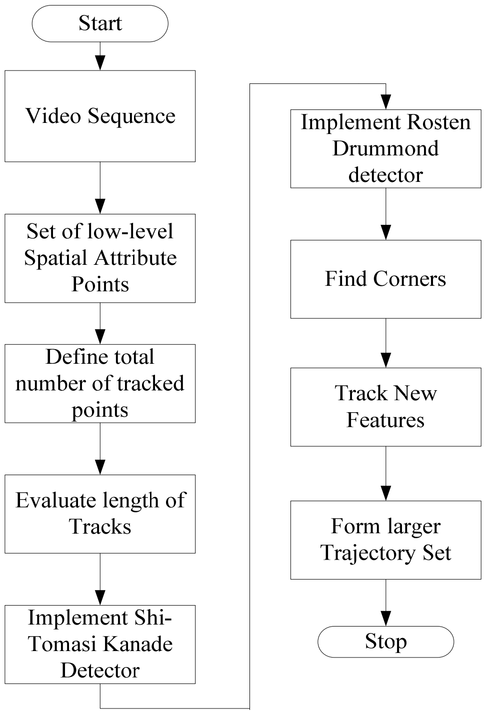


Fig.4. Mid Phase Implementation

In this implementation, the application first identify low-level attributes in the initial frame using the standard Shi-Tomasi-Kanade detector as well as the Rosten-Drummond detector, which is considered as a fast algorithm for finding corners. The low-level attributes are tracked over time using a hierarchical implementation of the Kanade-Lucas-Tomasi optical flow algorithm. The new attributes are tracked along with the existing point tracks to form a larger trajectory set. For trajectories that have initially stationary segments, it retains only the remaining part of the trajectory that shows significant temporal variations as shown in Fig 4.

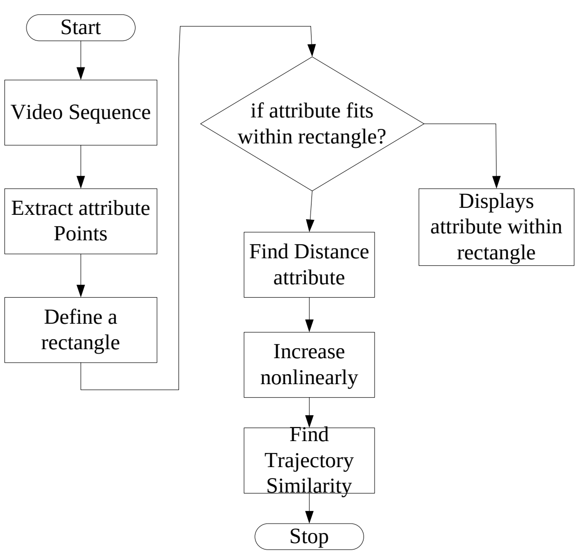


Fig.5. Final Phase Implementation

If the attributes do not fit within a rectangle, the distance between them is nonlinearly increased. The expectation is that attribute point tracks from the same underlying object are likely to have a low maximum distance as well as a low variance in distance over the region of overlap. Hence, it computes an overall trajectory similarity, 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 as shown in Fig 5.

VII. IMPLEMENTATION

One of the greatest challenging part of the proposed work is the design of an empirical model which is efficient and at the same time to enhance the low level task, e.g. object detection and tracking for variable speed of moving objects and which is also applicable to most dynamic real time situation too. 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 attribute traced over time captured using digital image capturing device. The traced locus i^{th} of the attribute is represented as L^i.

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

i=1,. . . , Z (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 attribute will be efficiently traced. The hierarchical-temporal approach of the proposed system will enable the tracing for locus of low level attributes of the input video frames. Based on the progress of the frames, all the upcoming mobile attributes 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

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 attributes (in pixels)*;

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

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

10 *Define threshold for the selection*;

11 *Define threshold for rejection of a point*;

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

13 *Estimate low level attribute 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 attribute*

19 *Maximize 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 attribute are found

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

26 *end*

Algorithm: Attribute 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 attribute selector.

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*

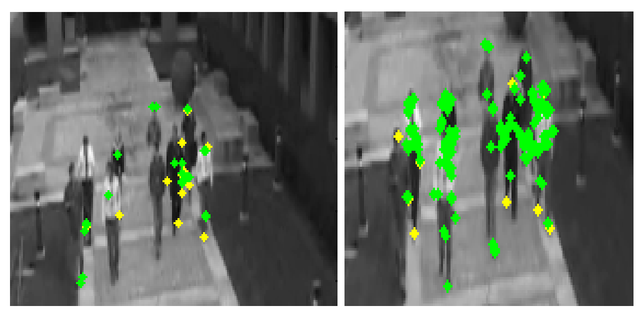
8 *Find the non -zero values, Get the column value, Get the Row value, Calculate the distance between X coordinates*

9 *Estimate the unique attribute*

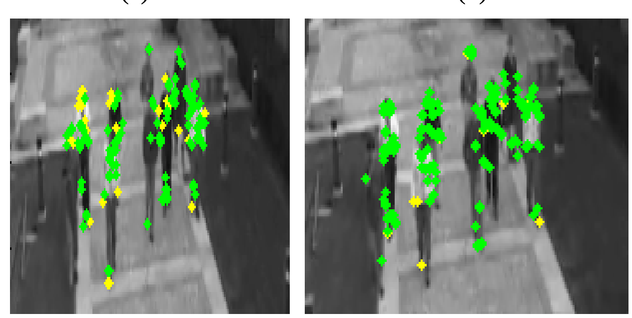
The proposed algorithm with spatial-temporal approach will be design in such a way that it should be independent of deployment of object-specific contour models as well as it will dependent on identification and integrating object specific components too. The proposed system is also free from calibration 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.

VIII. RESULT

The experiment has been performed using CAVIAR database. The initial part of video clips were recorded for the CAVIAR project [39] with a wide angle camera lens in the entrance lobby of the INRIA Labs at Grenoble, France. The resolution is half-resolution PAL standard (384 x 288 pixels, 25 frames per second) and compressed using MPEG2. The file sizes are mostly between 6 and 12 MB, a few up to 21 MB [39]. The video clip considered is 19.3 MB which is used for testing the algorithm. The sophistication of the video feed is very high as multiple number of objects are shown walking at the same time and walking in different speed. It also has instant of partial and full occlusion in highly dense crowd scene. As a result, the targeted object deformation introduced is maximized to test the efficiency of the algorithm. 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.



6(a) 6(b)



6(c) 6(d)

Fig 6. Traced Locus of attribute

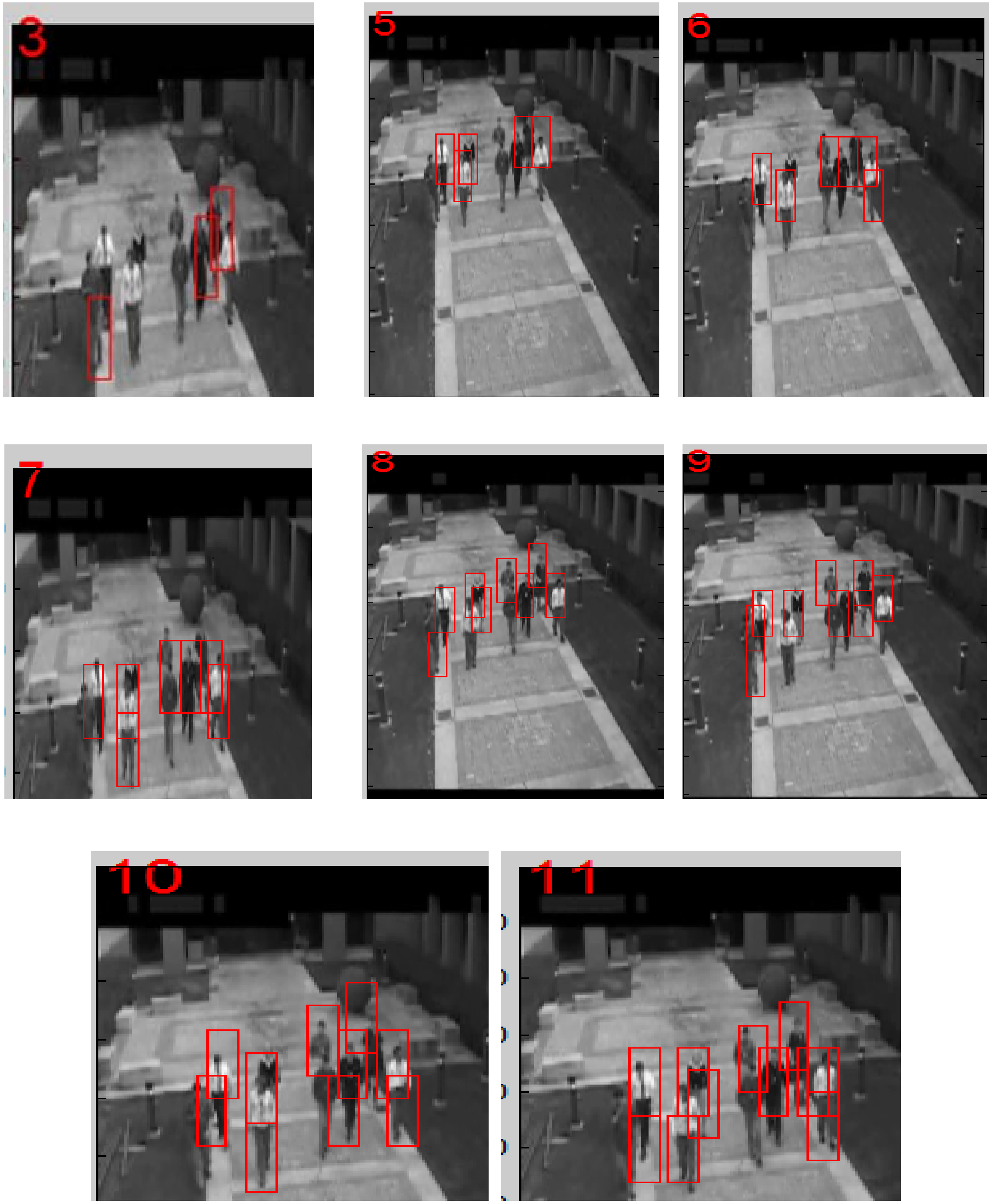


Fig 7. Object Detection and Tracking

The execution of the application highlights all the traced locus of attribute as shown in green dots in Fig 6. and Fig 7 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 7, although in initial results, only 3 people are detected leaving the remaining 6 people not detected. But with gradual spatial-temporal approach (as shown in Fig 8), it can also be seen that frame rate 7 is shown to detect 7 and final frame 11 shows to detect with 100% accuracy. The algorithm processes attribute 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.

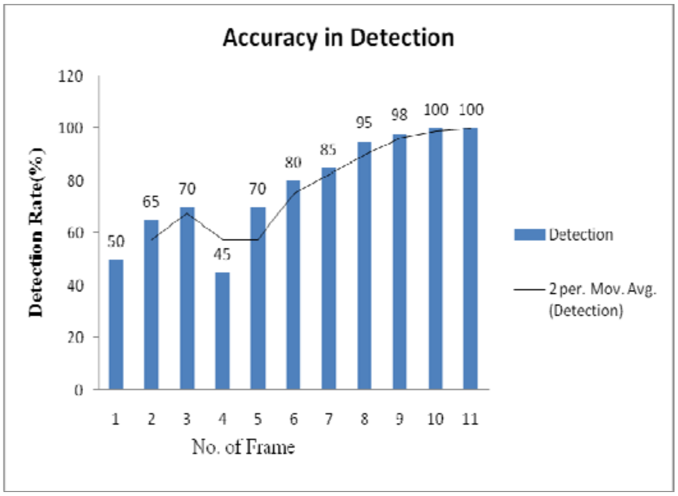


Fig 8. Accuracy in Detection (No. of Frames Vs. Detection Rate)

The various constraints of the project work include the following points:

- Although the input of the project work is basically a pre-captured video, but for the analysis, it will be a real troubleshooting if the quality of the video is poor and shaky. So, CCTV or any visual capturing mounted in the specific region should be real rock steady to capture the better input of video.
- The low-level attribute can be obtained using Shi-Tomasi-Kanade detector and Rosten-Drummond detectors.
- For initiating the project development stage, the video file is considered to be in AVI file format only.
- Probability of detection of dual heterogeneous object as one. For an example, if a video is captured for a person is pulling a cart, the actual input of object here is one person, but the application can probably predict as two objects including cart.

IX. 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 attribute 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. The future work will also focus on homography based detection and tracking for object in the dense crowd under challenging illumination condition using multiple cameras This may overshadow budget limitations when essential or sensitive areas are considered.

Reference

[1] Helbing, D. and Moln´ar, P., Social force model for pedestrian dynamics, Physical review E 51(5): 4282 – 4286, 1995
[2] http://citg.tudelft.nl/index.php?id=18769&L=1. Accessed on 10th Jan, 2012
[3] Gwynne, S., Galea, E. R., Lawrence, P., Owen, M. and Filippidis, L., Exodus. http://fseg.gre.ac.uk/exodus/., 1997
[4] Thompson, P. A. and Marchant, E. W., Simulex. http://www.ies4d.com/. 1994
[5] Turner and Penn , Evas. http://www.vr.ucl.ac.uk/research/evas/., 2002
[6] Kopp, N., EvacSim, an evacuation simulation model. www.nathan.kopp.com/EvacSim/evacsim.html., 1999
[7] Bregler, C., Learning and recognizing human dynamics in video sequences, CVPR ’97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR ’97), IEEE Computer Society, Washington, DC, USA, p. 568., 1997
[8] Polana, R. and Nelson, R. C., Detection and recognition of periodic, nonrigid motion, Int. J. Comput. Vision 23(3): 261 – 282., 1997
[9] Shio, A. and Sklansky, J. Segmentation of people in motion, IEEE Workshop on Visual Motion, pp. 325 – 332., 1997
[10] Heisele, B., Kressel, U. and Ritter, W. Tracking non-rigid, moving objects based on color cluster flow, Proc. of IEEE Conference on Computer Vision and Pattern Recognition, San Juan, pp. 257 – 260, 1997
[11] Cutler, R. and Davis, L. S., Robust real-time periodic motion detection, analysis and applications, IEEE Trans. Patt. An. Mach. Int. 22(8): 781 – 796., 2000
[12] Papageorgiou, C. and Poggio, T., A pattern classification approach to dynamical object detection, ICCV ’99: Proceedings of the International Conference on Computer Vision-Volume 2, IEEE Computer Society, Washington, DC, USA, p. 1223, 1999

[13] Wren, C., Azarbayejani, A., Darrell, T. and Pentland, A., Pfinder: Realtime tracking of the human body, IEEE Trans. Pattern Anal.Mach.Intell. 19(7): 780－785., 1997
[14] Cai, Q. and Aggarwal, J., Tracking human motion using multiple cameras, Proc. of International Conference on Pattern Recognition, Vienna, pp. 68－72., 1996
[15] Beymer, D. and Konolige, K., Real-time tracking of multiple people using continuous detection., 1999
[16] 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)
[17] 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
[18] 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
[19] 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
[20] Ernesto L Andrade, Robert B Fisher, Simulation of Crowd Problems for Computer Vision, Computer and Information Science, Volume: 3, Pages: 71-80, 2005
[21] Ernesto L. Andrade, Scott J. Blunsden, Robert B. Fisher, Performance Analysis of Event Detection Models in Crowded Scenes, IET International Conference on Visual Information Engineering VIE 2006, ISBN: 0863416713, Pages: 427-432
[22] Chern-Horng Sim and Surendra Ranganath, Reducing False Alarms for Detections in Crowd, ACCV'07 Workshop on Multi-dimensional and Multi-view Image Processing, Tokyo, Nov., 2007
[23] Min Li, Zhaoxiang Zhang, Kaiqi Huang and Tieniu Tan, Rapid and robust human detection and tracking based on omega-shape features,
[24] Antoni B. Chan Mulloy Morrow Nuno Vasconcelos, Analysis of Crowded Scenes using Holistic Properties, 11th IEEE Intl. Workshop on Performance Evaluation of Tracking and Surveillance (PETS 2009), Miami, 2009
[25] Chi-Chen Raxle Wang, Jin-Yi Wu, Jenn-Jier James Lien, Pedestrian detection system using cascaded boosting with invariance of oriented gradients, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 23, No. 4 (2009) 801–823
[26] Nan Dong, Fuqiang Liu, Zhipeng Li, Crowd Density Estimation Using Sparse Texture Features, Journal of Convergence Information Technology Volume 5, Number 6, August 2010
[27] Mikel Rodriguez, Ivan Laptev, Josef Sivic, Jean-Yves Audibert, Density-aware person detection and tracking in crowds, Proceedings of the IEEE International Conference on Computer Vision (2011)
[28] Beril Sirmacek, Peter Reinartz, Kalman filter based feature analysis for tracking people from airborne images, ISPRS Workshop on High-Resolution Earth Imaging for Geospatial Information, Germany, 2011
[29] Junliang Xing, Haizhou Ai, Liwei Liu, Shihong Lao, Robust crowd counting using detection flow, 18th IEEE International Conference on Image Processing, 2011
[30] Bertrand Luvison, Thierry Chateau, Jean-Thierry Lapreste, Patrick Sayd and Quoc Cuong Pham, Automatic Detection of Unexpected Events in Dense Areas for Video surveillance Applications, ISBN 978-953-307-436-8, 2011
[31] Matei Mancas, Nicolas Riche, Julien Leroy, Bernard Gosselin, Abnormal motion selection in crowds using bottom-up saliency, ICIP, 2011
[32] Mikel Rodriguez, Josef Sivic, Ivan Laptev, Jean-Yves Audibert, Data-driven Crowd Analysis in Videos, ICCV-2011
[33] Beril Sirmacek, Peter Reinartz, Automatic crowd analysis from very high resolution satellite images, Proceedings of 2nd International Conference on Recent Advances in Space Technologies, 2011
[34] T. Broida and R. Chellappa. Estimation of object motion parameters from noisy images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(1):90–99, 1986.
[35] J. T. Reindeers J. Veenman and E. Backner. Resolving motion correspondence for densely moving points. IEEE Transaction on Pattern Analysis and Machine Intelligence, 23(1), January 2001.
[36] D. Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5), May 2002.
[37] G. Sapiro M. Bertalmio and G. Randall. Morphing active contours. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(7), July 2000.
[38] Pabboju Sateesh Kumar, Multi-agent tracking under occlusion and 3D motion interpretation, Doctorial Thesis, Aug-2006
[39] http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/. Accessed on 10th Jan, 2012