ARPN Journal of Engineering and Applied Sciences




# CROWD DENSITY ANALYSIS AND TRACKING

P. V. V. Kishore, M. Nanda Kishore, R.Rahul and D. Prudhvi Raj
Department of Electronics and Communication Engineering, K. L. University, Green Fields, Vaddeswaram, Guntur D. T., Andhra Pradesh, India
E-Mail: pvvkishore@kluniversity.in

## ABSTRACT

Crowd Density Analysis (CDA) aims to compute concentration of crowd in surveillance videos. The central theme of this paper is to estimate the crowd density using crowd feature tracking with optical flow. Features for Accelerated Segment Test (FAST) algorithm extracts local features for each of the surveillance video frame. Optical flow tracks the extracted local features between frames. This process identifies the crowd features in consecutive frames. Kernel density estimator computes the crowed density in each successive frame. Finally individual people are tracked using estimated flows. The drawback of this method is similar to suffered by most of the estimation methods in this class that is reliability. Hence testing with three popular optical flow models is initiated to find the best optical flow. Three methods are Horn-Schunck (HSOF), Lukas-Kanade (LKOF) and Correlation optical flow (COF). Five features extraction methods were tested along with the three optical flow methods. FAST features with horn-schunck estimates crowed density better than the remaining methods. People tracking application with this algorithm gives good tracks compared to other methods.

**Keywords:** crowd density analysis, FAST features, optical flow feature tracking, crowd density maps.

## INTRODUCTION

Many recent works emphasize on motion detection of face and facial features but few concentrate on motion detection of people in crowd. The reason, motion pattern of crowd is not always uniform and it is also not necessary that the objects in motion in crowd are always humans. These challenges remain unaddressed even in this paper.

This work accepts that crowds are dynamic and humans are moving objects in the crowd [1]-[2]. The aim is to design a semi-automated system to observe the crowd density at railway stations, bus stands, shopping malls, public meetings, temples and many more. India by its population size is by default a crowded country.

The crowd monitoring is vital towards avoidance of crowd disasters. For example in 2014 during Dussehra celebrations at the Gandhi Maidan in India around 32 people died in the stampede. Every year there are around two to three incidents of this kind in Indian subcontinent. All these disasters call for an intelligent monitoring system at crowd areas that can map crowd density and track people giving rise to early warning systems.

This research aims to estimate the crowd density from surveillance videos [3]. The database of surveillance video collection includes various locations in and around Vijayawada city and from Vijayawada railway station. Different algorithms test the local feature extraction on video frames by specifying the interest points [4]. The feature extraction algorithms tested are, Features from Accelerated Segment Test (FAST) [5], Scale Invariant Feature Transform (SIFT) [6], Speeded up Robust Features (SURF) [7], Maximally Stable Extremal Regions (MSER) [8] and MinEigen features (MEiF).

The extracted local features in consecutive frames forms input to tracking algorithm. Optical flow estimation is the tracking algorithm for feature tracking. Three variations of optical flow constitute tracking phase. They are Horn-Schunck Optical Flow (HSOF) [9], Lucas Kanade Optical Flow (LKOF) [10] and Correlation Optical Flow (COF) [11].

All of the above methods display a trade-off between feature vectors and quality of the motion estimates and the optimum method satisfying both these domains is used [12]-[13].

Finally, for analysis crowd density maps are generated with Kernel density estimate (KDE) by using the distance between motions tracks obtained in the previous step comparing with ground truth results, the proposed method's strength is assessed.

### Feature extraction

This is the foremost step in the project. Features are the distinctive attributes present in an image. Features are extracted in order to get these different attributes. Which are stored as a vector and same extraction algorithm is defined for every frame of the video. The vectors of the consecutive frames are compared and the differences are obtained. Various types of feature extraction methods are present. Broadly there are five types of methods (i) Low level Feature Extraction[14] (ii) Curvature Feature Extraction (iii)Image motion (iv) Shape based Feature Extraction (v) Flexible Methods[15]. More about the Feature Extraction methods and algorithms are presented in [1], which gives the information as well as comparison among various types of features present.

VOL. 10, NO 20, NOVEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



**FAST features**

Features from Accelerated Segment Test are a corner detection algorithm. In an image, even though it states that corners are detected but more than corners the algorithm detects interest points. There are many corner detection algorithms like Morvec corner detection algorithm, SUSAN corner detection algorithm, Multi-Scale Harris operator algorithm and etc., among all those the FAST algorithm is known as the best one.

In the FAST algorithm a bresenham circle of radius 3 is selected. The intensity comparison of pixels lying on circumference is done, and based on the threshold value the decision is made whether to pick the pixel as interest point or not. More about the FAST algorithm is given in [5].

**Optical flow**

The cost functions in classical optical flow in spatially discrete setup between two video frames $F_1^{xy}$ and $F_2^{xy}$ is

$$\begin{aligned} C\left(v_x, v_y\right) = \sum_{x,y \in N} & \left(F_1^{xy} - F_2^{(x+v_x, y+v_y)}\right) \\ & + \alpha\left[\left(v_x - v_{x+1}\right) + \left(v_x - v_{y+1}\right)\right] \\ & + \left[\left(v_y - v_{x+1}\right) + \left(v_y - v_{y+1}\right)\right] \end{aligned} \tag{1}$$

Where $v_x$ and $v_y$ are the optical flow fields to be estimated from the consecutive frames of the surveillance video $F_1^{xy}$ and $F_2^{xy}$ with $\alpha$ as a smoothing parameter. Sometimes a penalizing function is multiplied to the velocity expressions in (1) as a catalyst for convergence. In this case we have chosen differential variant of L1 norm, $\sqrt{x^2 + \varepsilon^2}$ which is the most robust convex function. Where $\varepsilon$ the 2nd order error between the two frames is

$$|\varepsilon| \le \frac{d^2\left|F_1^{x''}\right|}{2\left|F_1^{x'}\right|} + O\left(\varepsilon^3\right) \tag{2}$$

With appropriate initial estimate eq'n (1) might converge towards a iterative estimation, by decreasing the position error in each iteration. This whole process wraps one video frame $F_1^{xy}$ on to another $F_1^{xy}$ by estimating the shift between the two frames.

There are three standard models for computing optical flow. They are Horn-Schunck Optical Flow (HSOF), Lucas Kanade Optical Flow (LKOF) and Correlation Optical Flow (COF).

**Crowd density estimation**

The process of crowed density estimation is formulated in sequential steps. First, the local features are extracted for each frame in the video. Local features of objects in the video

**Figure-1.** FAST local features in 4 consecutive frames of a surveillance video at K.L. University during a tech fest.

frames are obtained with Features form accelerated segment test (FAST) feature extraction algorithm. Figure-1 shows the effect of FAST algorithm on 4 consecutive frames of a surveillance video at K.L.University campus. Most of the database videos used for crowed analysis consists of only moving crowds. Where it happens to be a constraint on judging the performance of the system.

Hence in this research the videos are captured with moving and static people along with other natural moving backgrounds. Thus this is quite a challenge. Tracking local features from consecutive frames of a video estimates the crowd movement.

Optical flow algorithm uses pixel intensities of the local features as cues for tracking. Directly computing optical flow estimate incurs two basic hurdles. One is direct optical flow estimate on the pixel values of video frames induces increased computation time for long video sequences. Two, it will also help to classify crowd and non-crowd portions in the video. Our algorithm assumes the brightness variation between two consecutive frames is constant.

VOL. 10, NO 20, NOVEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences




**Figure-2.** FAST Feature tracking using Horn Shunk (HS) optical flow in the 4 consecutive frames in Figure-1.

Figure-2 gives an output of optical flow horn schunck algorithm. The advantage offered by tracking features is clearly visible in Figure-2. This happened only because of feature tracking. The non-crowd features in Figure-1 are missed in tracking due to their static nature. The potential tracks of people are generated by grouping the obtained motion vectors.

The computed motion vectors in each frame will have velocity in x and y directions between consecutive frames. Density estimation is accomplished by applying these motion vectors to kernel density estimator. Density is the measure of concentration of motion vectors between frames. If these motion vectors come close during the motion estimation, their concentration increases and therefore the density. Kernel density estimator calculates the probability density function (pdf) of a Gaussian kernel. For k$^{th}$ frame $F_k^{xy}$ and (k-1)$^{th}$ frame $F_{k-1}^{xy}$ having motion vectors $m_k^{xy}$ estimated over respective pixel locations (x,y), the pdf gives the density map defied by

$$p_k^{xy} = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{i=1}^{m_k^{xy}} e^{-\left(\frac{\left(v_x - v_x^i\right)^2 + \left(v_y - v_y^i\right)^2}{2\sigma^2}\right)} \tag{3}$$

Where $\sigma$ is deviation in Gaussian kernel.

This results in a density map indicating the concentration of pedestrians at locations in the frames. Crowed density maps for the test images in figure 1 are in Figure-3.

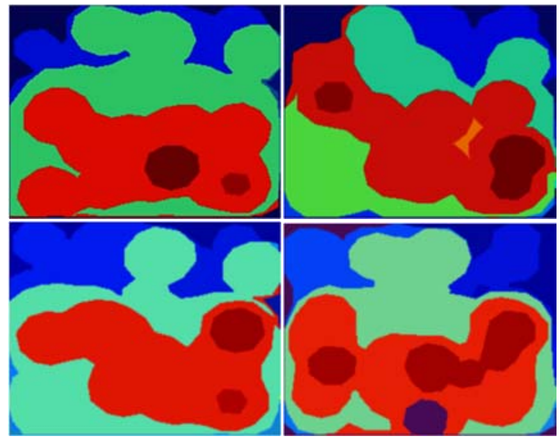


**Figure-3.** Crowd density maps of video frames in Figure-1.

The dark red and medium red portions show that there is a maximum concentration of people in that part of the frame. Crowd analysis can be applied to a variety of problems of crowd based image processing such as people tracking, crowd behavior analysis and privacy protection.

In this research we have taken up the problem of crowd tracking or people tracking in crowd videos. The foremost problem is localization of humans in the crowd to track them individually. But in cowed videos it a biggest challenge for image processing engineering. At the basic level we tried to use spatial domain morphology to separate out individual motion vectors. The erosion tool is used two to three times with line structuring element to separate overlapping points.

Next frame differencing of motion vectors between two successive frames will locate the changing motion vectors. Next, changing motion vectors are treated with dilation operator having diamond structuring element in an 8 neighborhood. Bounding box is inserted on the connected motion vectors between consecutive frames. Following Figure-4 reveals the bounding boxes on frames in Figure-1.

From Figure-4, the red bounding boxes are not always consistent between successive frames. This is due to overlapping people in crowd videos that makes individual people tracking a challenging task for engineers.

VOL. 10, NO 20, NOVEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



**Figure-4.** People tracking application on video frames in figure 1 using the model generated for crowd analysis.

Testing of the proposed method for other popular features is initiated to understand the importance of feature vector in crowd analysis applications. This research uses SIFT, SURF, MSER and MEiF apart from FAST features. Figure-5 compares all the five features on a tracked frame of crowd with horn-schunck optical flow.

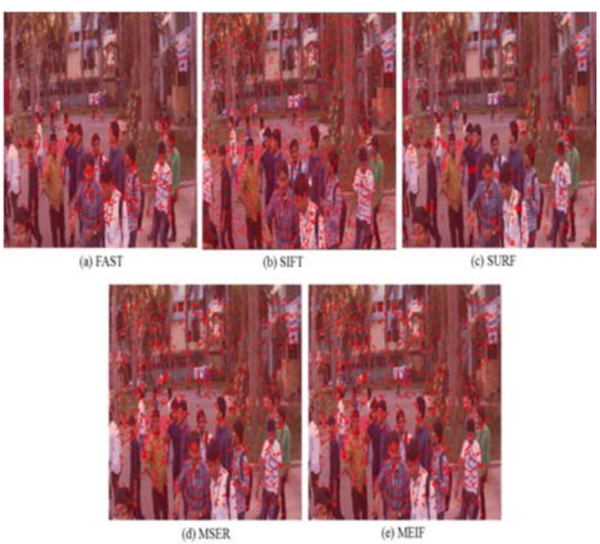



**Figure-5.** Comparison of five different feature vectors during tracking using HS optical flow.

From Figure-5, visual observation gives the superiority of FAST features used in tracking in comparison with other features. The unwanted motion vectors calls for false density estimates, thereby reducing the efficiency of the algorithm.

Computing kernel density estimates with the tracked motion vectors from the features extracted for five feature vectors is shown in Figure-6.

Overlapping the density map on the original frame will

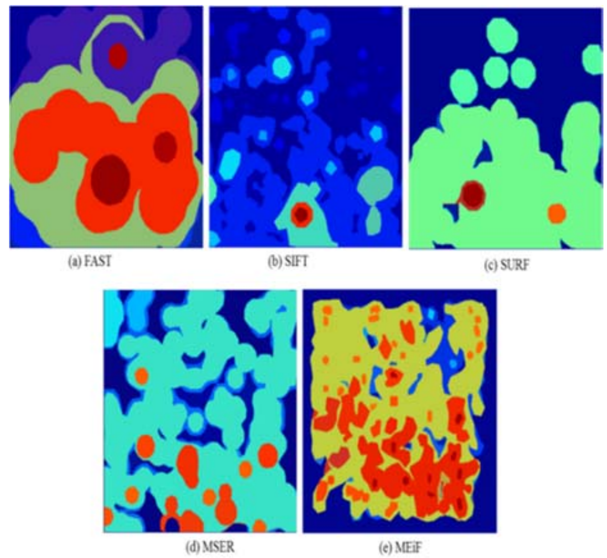
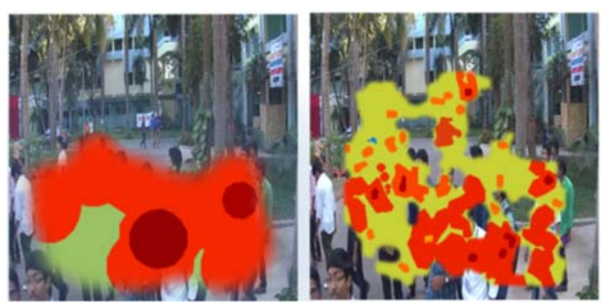



**Figure-6.** Comparison of five different feature vectors during tracking using HS optical flow.

show the potentiality of the FAST based HSOF compared to MEiF features in Figure-7.

Finally ground truth video is used to test the worth of the algorithm. Figure-8 shows the results obtained with two



(a) FAST DensityMap (b) MEiF Density Map

**Figure-7.** Comparison of density maps between FAST and MEiF.

consecutive frames in FAST and HS optical flow. The crowd density map in figure 8 is the ground truth density

VOL. 10, NO 20, NOVEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



map. In Figure-9, the ground truth density map and obtained density map are compared visually. It is observable that both the frames display almost identical crowd density maps. The mean square error between the density maps is 0.1823.

To understand the use of HS optical flow (HSOF) extensively for crowed analysis, it is compared with LKOF and COF. Figure-10 compares the density maps of the three methods using FAST features with ground truth in figure 9. This shows the authenticity of HSOF for the crowd analysis compared to other two optical flow methods.

To prove that HSOF with FAST feature tracking will work effectively, the algorithm is applied to a different set of videos obtained for a different source. The frames in Figure-11 show the surveillance footage at Vijayawada railway station in INDIA. The data is obtained requesting the railway authorities by submitting letters that these videos will only be used for research.

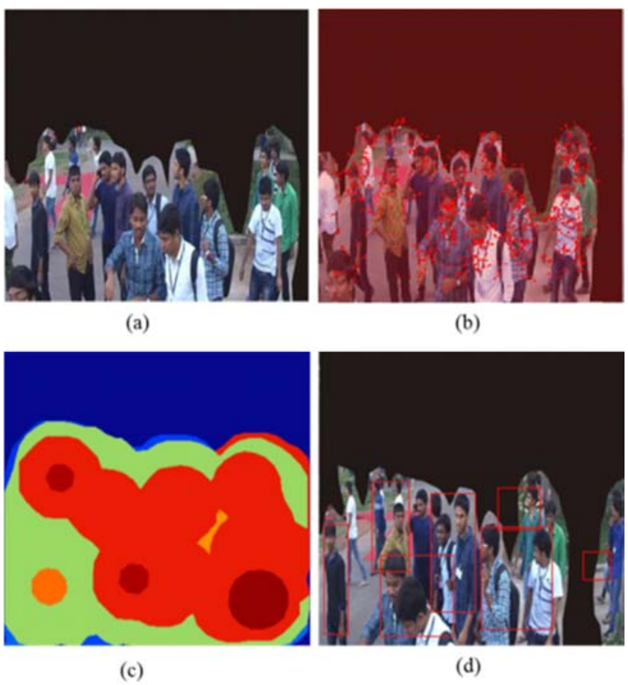



**Figure-8.** Ground truth video frame and its results. (a) Ground truth frame (b) FAST features (c) Density map (d) People tracking.

**Figure-9.** Ground truth video frame and its results. (a) Ground truth frame (b) FAST features (c) Density map (d) People tracking.

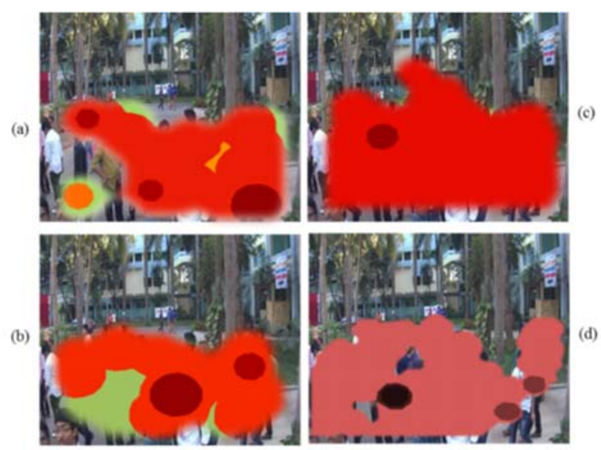



**Figure-10.** Ground truth video frame compared with various optical flow methods.(a) HSOF (b) Ground truth frame (c) LKOF (d) COF.

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences


www.arpnjournals.com

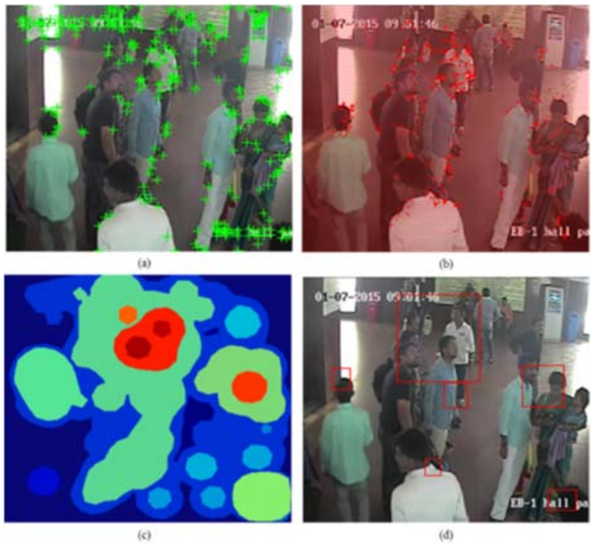


**Figure-11.** Ground truth video frame and its results. (a) Ground truth frame (b) FAST features (c) Density map (d) People tracking.

## CONCLUSIONS

The solution for the crowd analysis is provided in this paper from a density measurement point of view. Feature tracking with horn-schunck optical flow is computed to estimate the motion vectors between two consecutive frames. Five types of feature vectors are tested for extracting features. Three types of optical flow based algorithms are using as motion estimators for feature tracking. Density maps constructed on tracked motion vectors. An application of people tracking is initiated and solved in this work with maximum accuracy which was authenticated by ground truth density map. The density map generated during the implementation of proposed method has around 91%similar to ground truth density map. People tracking are also showing the similar results when compared to ground truth. Different scenarios where tested for the proposed algorithm and it always gave convincing results.

## REFERENCES


[1] H. Fradi, J.-L. Dugelay. 2014, Sept. Towards crowd density-aware video surveillance applications. ELSEVIER. Information fusion. [Online]. 24(2015), pp. 3-15. Available: http://dx.doi.org/10.1016/j.inffus.2014.09.005.

[2] Rainer Könnecke and Volker Schneider. 2014. On the Safety of large scale events - Crowd flow modeling of ingress and egress scenarios. ELSEVIER. Transportation research Procedia. [Online]. 2, pp. 501-506. Available: http://creativecommons.org/licenses/by-nc-nd/3.0/.

[3] Yuan Yuan, Jianwu Fang, and Qi Wang. 2015 Mar. On the Online Anomaly Detection in Crowd Scenes via Structure Analysis. IEEE TRANS. CYBERNETICS. [Online]. 45(3): 562-575. Available: http://www.ieee.org/publications_standards/publications/rights/index.html.

[4] Sudipto Mukherjee, Debdipta Goswami, and Sarthak Chatterjee. 2013. Analyzing Motion Patterns in Crowded Scenes via Automatic Tracklets Clustering. IEEE TRANS. SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS. [Online]. (99), pp. 1-12. Available: http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

[5] Jie Xu, Hua-Wen Chang, Shuo Yang, Minghui Wang. 2012. Fast feature-based video stabilization without accumulative global motion estimation. IEEE TRANS. Consumer Electronics. [Online]. 58(3): 993-999.

[6] Zafer Arıcan, Pascal Frossard. 2012. Scale-Invariant Features and Polar Descriptors in Omnidirectional Imaging. IEEE TRANS. Image Processing. [Online]. 21(5), pp. 2412-2423. Available: http://ieeexplore.ieee.org.

[7] Chen-Chien Hsu; Po-Ting Huang; Zhong-Han Cai; Ming-Chih Lu; Yin-Yu Lu. 2014. Depth measurement based on pixel number variation and Speeded up Robust Features. IEEE Conference. Consumer electronics.

[8] Donoser M.; Riemenschneider H.; Bischof H. 2010. Shape Guided Maximally Stable Extremal Region (MSER) Tracking. IEEE Conference, Pattern Recognition.

[9] Omer, O.A. 2012. Region-based Horn-Schunck optical flow estimation. IEEE Conference, JEC-ECC.

[10] Fan Zhang, Yang Gao. Bakos, J.D. 2014. Lucas-Kanade Optical Flow estimation on the TI C66x digital signal processor. IEEE Conference, High Performance Extreme Computing.

[11] Wei Chen. 2015. Current Motion Tracking from Satellite Image Sequence with Global Similarity Optimization Model. IEEE TRANS. Geoscience and Remote Sensing. [Online]. 53(2): 1008-1015. Available: http://ieeexplore.ieee.org.

[12] Kishore P. V. V. and P. Rajesh Kumar. 2012. A Model For Real Time Sign Language recognition System. International Journal of Advanced Research in Computer Science and Software Engineering. 2(6), ISSN 2277-6451.

[13] Kishore P. V. V. and P. Rajesh Kumar. 2012. Segment, Track, Extract, Recognize and Convert Sign Language Videos to Voice/Text. International Journal of Advanced Computer Science and Applications (IJACSA) ISSN (Print)-2156 5570, 3(6). (DOI): 10.14569/IJACSA.2012.030608.

[14] Kishore, PVV; Kishore, SRC; Prasad, MVD. 2013. Conglomeration of Hand Shapes and Texture Information for Recognizing Gestures of Indian Sign Language Using Feed forward Neural Networks International Journal of engineering and Technology (IJET), Vol.5, No.5, pp.3742-3756.

[15] Kishore, PVV; Sastry, ASCS; Kartheek, A. 2014. Visual-verbal machine interpreter for sign language recognition under versatile video backgrounds Networks & Soft Computing (ICNSC), 2014 First International Conference on, pp.135-140, IEEE.