An Automated System for Detecting Congestion in Huge Gatherings

Sunil Kumar Aithal S
Computer Science and Engineering
NMAMIT, Nitte
Karkala, Karnataka, India

Reeja S R
Computer Science and Engineering
NMAMIT, Nitte
Karkala, Karnataka, India

ABSTRACT
Recently, a huge number of people travel together and may meet in a particular location for a specific purpose such activity is termed as mass gatherings. Rally of politics, gathering at fairs, religious occasions, gathering at railways are the events of mass gatherings. Such gatherings pose severe threats for crowds, because huge number of people moving in small areas can promote the spread of dangerous diseases and enhance the threat of injury which eventually results in crowd stampedes. Crowd can rely on factors, such as individual's emotion, elderliness, and consumption of liquor, which influence the severity of harm. Such factors can make crowds to be more violent. Thus in such places several issues should be addressed well for avoiding unexpected situations.

Nowadays, many researchers are showing lots of interest in computer vision area. Study of human movement topic in that area has been popular for modern technical decades. The popularity is due to a huge expansion of applications in video vigilance and crowd dynamics. Because the activity is of immense scientific concern, it offers different computational challenges and because of an expeditious upsurge in video vigilance technology deployed in both private and public areas.

In this paper, we present a system for tracking and provide early information of hazardous locations in huge gatherings. It is based on optic flow estimations and detects sequences of crowd motion that are characteristic for devastating congestions. For optic flow computation, Lucas - Kanade method is employed to determine the optical flow vectors for the gathered video. Video sequences are segmented and optic flow is determined for respective segments. A threshold optic flow is chosen in such a way that the tracking of congested area in video is easily done by comparing it with respective segment's determined optic flow values. Finally, we present the location of crowd congestion which helps in taking further protective measures to handle unusual events.

Keywords
Computer vision, congestion detection, large gatherings, optical flow estimation approach, segmenting frames, video processing.

1. INTRODUCTION
Around the globe extensive gatherings are most frequent and very popular in socio-activities. Currently, typical examples include sporting events, gatherings at airways, railway stations, famous spots, fairs, or recitals. They attract uncontrollable large number of visitors and major precautious actions are always necessary. In spite of everything, the use of current vision field topic such as video surveillance; lethal human stampedes and crowd tragedy still happen. (refer Table 1).

Many research reports conclude that during large gatherings, the density of people rapidly increases; reports show that on an average ten people per square meter density has been recorded [1]. Immense density of crowd naturally occurs from varying patterns of crowd behavior such as stop-and-go waves or turbulence in crowd [2]. In order to avoid frequent collision in crowd, they have tendency to transit from a steady flow to a stop-and-go manner, as a result individuals move with reduced footsteps distance and finally stops moving at any point of time. This kind of situation where people are packed densely results in crowd turbulence. Intermittent forward pedestrian movement and backward free space propagation are the characteristic of stop-and-go wave patterns. They are caused when the obstructions and densities of pedestrian are uncontrollable. During such uncontrollable situation involuntary pedestrian movement will cause rapid motion of other neighboring people. Thus in such scenario individuals may walk falsely and finally collapse on ground. Subsequently, because of lethal pressure of up to 4500 N/m on pedestrian chests and also deprivation in oxygen leads to loss of life [2].

The motivating factor is to implement a system that is capable of identifying and track crowd motion, which is a major interesting research topic in vision analysis, security, crowd dynamics, and visual surveillance fields. The system must be able to track people movement in the subsequent video frames and should provide early information to the concerned authority about the location of overcrowding. The developed system might find real-time application in visual surveillance, crowd analysis, identifying suspicious people activities in crowd. Thus, developed system can be enhanced to address security issue in immense crowding.

In this paper, we present a system for the movement detection of crowd gathered in video sequences. System incorporates estimating optic flow approach and is designed to track crowd motion pattern which is a feature for harmful congestion. It follows assumption that, sequences of ordered frames of video allows the computation of motion as either discrete image or frame displacements or instantaneous image or frame velocities. For optic flow, Lucas- Kanade approach is embodied to estimate the optical flow values for the gathered

Table 1. Statistics on recent crowd stampedes.

Year	Location	Deaths Occurred
2013	Kumbh Mela disaster at Allahabad, Uttar Pradesh, India	more than 30
2012	A huge riot at Egyptian city of Port Said	more than 70
2011	Stampede at Kerala Festival, India	more than 100
2006	Stampede during hajj near Jamarat Bridge	nearly 350
2006	Philippine Stadium stampede	within 80

sample video. Video frames are segmented and for respective segments optical flow vectors are calculated. A threshold optic flow values are chosen in such a way that the tracking of congested region in video is easily performed by comparing it with respective segment's estimated optic flow values. Finally, we present the location of overcrowding which helps in taking further protective measures to handle unusual hazardous incidents.

2. LITERATURE REVIEW

Nowadays, research on crowd dynamics topic in computer vision field is gaining huge popularity due to its significant use in various applications that detects and provides early indications of densely crowded area and also supports emergency security aspects [3]. Computer vision is the core technology which deals with automated vision analysis. Thus, from several recent decades visual systems are focusing on modeling and interpreting various sorts of pedestrian activities including the individual's behavior in crowds. Pedestrian movement analysis is an active research topic from previous two decades. Earlier, dense pedestrian group and their self-behaving characteristics in crowd were studied using simulations [12]. For identifying people behavior in densely packed crowd different physical models exist which depicts pedestrian motion as a model. They are modeled based on the correspondence to fluid dynamics. Later various popular models came into existence termed as socio-force model [5] and cellular automaton [4], which models pedestrian motion on a microscopic level.

From past two decades, various vision analysis and crowd analysis techniques have been developed specifically for tracking and recognizing unusual human activities in crowd[9]. Presently modeling of efficient and accurate vision-based systems embeds different models of simulation that are capable of tracking individual movement in immense gatherings. Cellular automata model admired by Ali and Shah took the initiative to present a framework for detecting human in the crowd [4]. It is an automated model composed with the knowledge on human activities in crowd, locations of obstructing objects and entry or exit door's position and mainly used for computation of flow vectors. Ali and Shah also presented a framework that deals with segmentation of the pedestrian flow that helps in tracking varying patterns among crowd scenes [3]. To track unusual events in mass gatherings Mehran et al. considered socio-force model and then estimated the force interactions [5]. Further, various research and experiments are performed in crowd dynamics to understand hazardous pedestrian behavior for enhancing the existing physical methods. Numerous attributes like crowd flow, pedestrian velocity, pressure, people density are calculated either ancient known methods [6] or by means of recent automated digital visual processing techniques [7]. Usually, vision analysis experiments are performed on videos captured using top and front-view cameras mainly to avoid object obstructions and for supporting automated vision analysis.

Common approaches to automated video surveillance tracks and recognizes single person or object in video. Nevertheless, many works have reported the infeasibility of those approaches and thus they are not efficient for video surveillance of densely packed crowd, because huge gatherings video captured from any source will have numerous persons moving which necessitates complex calculations. Therefore, to ascertain crowd motion patterns and their directions computation of dense optic flow vectors is considered as efficient. By utilizing optical flow estimation approach not only accuracy of complex calculations is met but also the confidentiality of human identity is guaranteed. An observable movement of any object, areas or regions in a video sequences that occurs due to the instantaneous movement between perceiver and the scene is referred as an optical flow. They can also be termed as an optic flow. American psychologist James J. Gibson during the year 1940 demonstrated the concept of optical flow in order to present the idea of visual stimulus those were experienced by variety of animals roaming in any place. Further, the importance of optic flow is also reported by him for explaining the concept of affordance perception, which helps in recognizing various observable activities occurring in an environment. Recently, the concept of optical flow finds immense usage in the robotics field which incorporates various combined methodologies from video or image processing as well as from navigation control, like object's motion tracking, segmenting video frames and surface or edge detection in videos. Movement detection or estimation, video analysis or compressions are also a proactive research topic in a crowd dynamics field. Many works concludes optical flow computation knowledge is considered to be very useful in balancing the various operations of micro air vehicles [11].

3. DETECTION OF OVERCROWDED REGION

In huge gatherings, people moves with decreased velocity which results in oscillatory motion. Thus, tracking their velocity and motion direction becomes much difficult because its people tendency to change directions more frequently while moving in such gatherings. Liu et al. work shows that people moves in crowd with oscillatory motion and with differing velocities from 0.26 m/s up to 1.72 m/s [6]. For study of vision dynamics various authors have reported the existence of linearity between the velocity and the amplitude also between the velocity and the frequency. We estimate dense optical flow vectors using the Lucas-Kanade approach which is a differential technique. The perceptible movement of brightness pattern in a video sequences is called as an optical flow. Video sequences consist of ordered collection of frames or images that allow the estimation of flow vectors as either instantaneous frame velocities or discrete displacement of frames or images. Various methods exist for the calculation of optical flow vectors that detects motion between two subsequent or different frame sequences that are captured at times T and $T+\Delta T$ at every volumetric pixel location. Optic flow technique used is called differential because they consider only partial derivatives of spatio-temporal co-ordinates and also approximations of frame or image sequences are done based on Taylor series assumptions.

3.1 Optical Flow Estimation Approach

The embodied Lucas–Kanade approach is a popular and robust differential technique for optic flow vectors estimation. A major assumption made is that the flow does not vary much surrounding the region of the considered volumetric pixel and solves the basic optic flow equations for every pixels in that surrounding region by making use of the least square criteria. The approach also assumes that the motion observed between subsequent frames is negligible and remains nearly same in a neighborhood of the considered point say ***v***. Therefore the standard optic flow equation is expected to satisfy for every pixel belonging to an assumed window where ***v*** is its center. Thus, the local image or frame velocity vectors [m, n] should satisfy

$$I_x(r_1)m + I_y(r_1)n = -I_t(r_1)$$
$$I_x(r_2)m + I_y(r_2)n = -I_t(r_2)$$
$$\ldots\ldots$$
$$I_x(r_n)m + I_y(r_n)n = -I_t(r_n)$$

where r_1, r_2, and r_n are the pixels within the window and $I_x(r_i), I_y(r_i), I_t(r_i)$ are the derivatives which are considered to be partial for the image intensity *I* taken at x, y coordinates and time t near the point r_i . To solve the optical flow constraint equation for m and n, the above said approach divides the original image or frame into smaller sections and assumes a constant velocity in every section. Then, it performs a weighted least-square fit of the optical flow constraint equation and minimizes following equation:

$$\sum_{x \in \propto} W^2[I_x m + I_y n + I_t]^2$$

Here, W is a window function that emphasizes the constraints at the center of each section and $\propto$ represents corresponding section.

The stated approach solves velocity vectors m and n as follows:

1. Compute I_x and I_y using [-1 8 0 -8 1]/12 as convolution kernel, for every pixels in the first and second image.
2. Compute I_t using [-1 1] as convolution kernel, for every pixels in the first and second image.
3. Smoothing of I_x, I_y, I_t components are done using the kernel [1 4 6 4 1]/16.
4. For every pixel solving the 2-by-2 linear equations are done using the following approach:

$$If\ A = \begin{bmatrix} p & q \\ r & s \end{bmatrix},$$

$$where\ \ p = \sum W^2 I_x^2, \quad q = \sum W^2 I_x I_y,$$

$$r = \sum W^2 I_y I_x, \quad s = \sum W^2 I_y^2.$$

$$Then\ the\ eigenvalues\ of\ A\ are\ \lambda_i = X \pm Y, where\ X = (p+s)/2\ and\ \ Y = (\sqrt{4q^2 + (p-s)^2})/2\ , for\ i = 1,2 \ldots$$

5. Compare eigenvalues with the chosen threshold, τ, final results fall into one of the following cases:

   Case 1: $\lambda_1 \geq \tau\ and\ \lambda_2 \geq \tau$

   A is nonsingular; solve equations by Cramer's rule.

   Case 2: $\lambda_1 \geq \tau\ and\ \lambda_2 < \tau$

   A is singular; normalizes the gradient flow to calculate m and n.

   Case 3: $\lambda_1 < \tau\ and\ \lambda_2 < \tau$

   The optical flow, m and n, is 0.

In video analysis, a convolution kernel, or mask is a matrix used for sharpening, smoothing, edge-identification, etc. This is accomplished by means of convolution between an image and a kernel.

3.2 Segmenting Video Sequences or Frames

Gathered crowd video sequences or frames are segmented for easy and efficient processing of video sequences. We decompose sequence of frames from collected crowd video and each frame or image is segmented into several parts or segments. For each segment the above mentioned optical flow computation approach is applied. In an overcrowded area, the people either move very slowly or stops moving itself. If the scenario is analyzed for top-view portion of video the estimation of crowd velocity can be done by assuming only the magnitude of optical flow vectors. So for every segment flow vectors are deduced by just depending on its magnitude values. Further, a threshold optical flow vectors are chosen for comparing each segment’s already estimated optical flow fields with them. Therefore, if the optical velocity values in any segment are less than the chosen threshold then there is no crowd movement in that region and thus congestion exist. Else, if the optical flow values in any segment are more than the threshold then we conclude people in that area moves freely and thus no congestion can exist. Finally, the area is reported where there exists congestion for taking further preventive measures. This technique of estimating optical flow vectors can also be applied to real time crowd video footage for detecting congestion.

4. RESULTS

The technique of movement detection is performed for human movement tracking and congestion detection in immense gatherings. We take a sample video footage of gatherings near railways and extract frames from collected video. Individual frames are segmented and for each segment optical flow is computed. A threshold optic flow is chosen in such a way that the tracking of congested area in video is easily done by comparing it with respective segment’s determined optic flow values. Final system should be able to track congested region among crowd gatherings as shown in following figures where Figure 1 and Figure 2 are the output frame sequences in gathered video and rectangular white box in frame sequences depicts congested area near railways.

Fig. 1: Congestion detection at railway gatherings

Fig. 2: Congestion detection at railway gatherings

5. CONCLUSION & FUTURE WORK

The proposed system automatically detects an overcrowded area using above said optical flow estimation technique and segmenting frame sequences. It can be enhanced to track and inform such congested regions occurring in real-time video sequences. The video footage assumed can be crowding at railways or overcrowding at political rally etc. Such useful information can be used to take necessary steps in controlling hazardous incidents occurring in mass gatherings. The System can also be enhanced to track individuals and analyze their suspicious activities in mass gatherings and inform the concerned authorities to take necessary actions.

6. REFERENCES

[1] *Crowd Dynamics*. K. Still. PhD thesis, Univ. of Warwick, 2000.

[2] D. Helbing and A. Johansson. Pedestrian, crowd and evacuation dynamics. In R. A. Meyers, editor, Encyclopedia of Complexity and Systems Science. Springer, 2009.

[3] S. Ali and M. Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In CVPR, 2007.

[4] A. Kirchner and A. Schadschneider. Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics. Physica A, 312(1-2):260–276, 2002.

[5] R. Mehran, A. Oyama, and M. Shah. Abnormal crowd behavior detection using social force model. In CVPR, 2009.

[6] X. Liu, W. Song, and J. Zhang. Extraction and quantitative analysis of microscopic evacuation characteristics based on digital image processing. Physica A, 388(13):2717–2726, 2009.

[7] A. Johansson, D. Helbing, H. Z. Al-Abideen, and S. Al-Bosta. From crowd dynamics to crowd safety: A video-based analysis. Advances in Complex Systems, 11(04):497–527, 2008.

[8] B. Zhan, D. Monekosso, P. Remagnino, S. Velastin, and L. Xu. Crowd analysis: a survey on Machine Visions and Applications, 19(5–6):345–357, 2008.

[9] T. Zhao and R. Nevatia. Tracking multiple humans in crowded evnironment. In CVPR, 2004.

[10] David J. Fleet and Yair Weiss (2006). "Optical Flow Estimation". In Paragios et al. *Handbook of Mathematical Models in Computer Vision*. Springer. ISBN 0-387-26371-3.

[11] OpticalFlow<http://en.wikipedia.org/wiki/Optical_flow>

[12] CROSS: Modelling Crowd Behaviour with Social Cognitive Agents *Journal of Artificial Societies and Social Simulation* **16** (4) 1 <http://jasss.soc.surrey.ac.uk/16/4/1.html>.

[13] Lucas, B., and Kanade, T. (1981) An Iterative Image Registration Technique with an Application to Stereo Vision**,** *Proc. of 7th International Joint Conference on Artificial Intelligence (IJCAI),* pp. 674-679.