Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

Integration of image processing and crowd simulation for understanding crowd dynamics in Kumbh Mela

P S Karthika[a], Ashish Verma[b*], S Pal[c]

[a] *Research Scholar, Department of Civil Engineering, Indian Institute of Science (IISc), Bangalore - 560012, India*

[b] *Associate professor, Department of Civil Engineering and Robert Bosch Centre for Cyber Physical Systems, Indian Institute of Science (IISc), Bangalore-560012, Karnataka, India*

[c]*Bachelor of Technology, Indian Institute of Technology, Kharagpur- 721302, India*

Abstract

The paper highlights the need to study crowd dynamics in mass religious gatherings. Kumbh Mela, the largest peaceful mass gathering in the world is considered for this purpose. The aim of the study was to evaluate the performance of an image processing module when combined with a crowd simulator. The pedestrian inflow as obtained from the image processing module is automatically fed into the crowd simulator. The pedestrian movements are simulated by using the social force model and a demo heat map is generated for assessing the crowd risk situations. The image processing module has an average accuracy of 73.45%. The effect of varying border threshold on the pedestrian movement is also explored. With a more accurate and realistic heat map in future, it can be used by the concerned authorities to make timely decisions on ensuring crowd safety in large gatherings.

Keywords: Image processing; crowd risk;crowd simulation; mass religious gatherings

1. Introduction

Religious gatherings are one of the prominent crowd pullers apart from political rallies, tournaments, concerts, processions etc. Historical trends from India and other countries suggest that the stampedes in mass gatherings,

* Ashish Verma. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000
E-mail address: ashishv@iisc.ac.in

especially in religious events occur frequently. These highlight the importance of studying the crowd behavior more scientifically and implementation of proper crowd management strategies, which can prevent future stampedes. An Early Warning System [EWS] must be developed which can predict crowd risk in any mass gatherings. The predictions from EWS can help the decision makers to control crowd density and prevent any kind of stampede incidents by implementing extra safety measures. Here, data obtained from Kumbh Mela held in Ujjain, 2016 is used for simulating the crowd risk.

Nomenclature

CRI	Crowd Risk Index
t_int	Updation time
Imp_t	Impatience threshold
$density_t$	Pedestrian density threshold
Imp_{wt}	Impatience factor
$density\_{}_{wt}$	Pedestrian density factor

Kumbh mela is the largest peaceful religious gathering in the world, where pilgrims flock to take dip in the sacred river. This congregation of saints (sadhus), visitors, and foreigners is held cyclically at the banks of river Godavari in Nashik, Triveni Sangam (confluence of Ganga, Yamuna, and mythical Saraswati) in Prayag/Allahabad, river Ganga in Haridwar and river Shipra in Ujjain, in about every three years. It is believed that one dip in these sacred rivers cleanses all of their sins and liberates them. The festival encapsulates almost all possible religious rituals in practice in the country including listening to spiritual discourses from sadhus (Satsang's), participating in holy prayers, taking Shahi Snan (holy dip), visiting temples, participating in long processions etc. Millions of pilgrim's flocks to the mela area with bare minimum facilities, creating a unique platform to study crowd behavior in its essence (See Fig.1).

a b

Fig. 1. Pilgrims taking part in Holy dip in the ghat† regions

Essentially crowd patterns emerge as a result of pedestrian interactions between themselves and with the environment. The pedestrians normally feels increasingly uncomfortable as they get closer to a strange person, who may react in an aggressive way. This results in repulsive effects of other pedestrians (Helbing et. al, 1995). The need to conduct studies focusing on real data collected from mass religious gatherings have been pointed out by Hariharan

† River banks

et al., (2017). To analyze the crowd patterns, pedestrian trajectories needs to be extracted from the video data collected in field. Abbas et al. (2013) proposed a technique based on computing the traffic load by comparing two images, the reference image (empty road) and the live traffic image. To improve the visibility of the vehicles, the difference image is converted to a binary image based on a threshold value. The vehicles are marked first and then their numbers are counted. The goal of this study was to reduce the traffic congestion by calculating the traffic density in a direction of the road and set green time of traffic signal accordingly. Idrees et al. (2013) presented an approach to count the number of individuals in extremely dense crowds. The repetitive nature of the crowd is exploited, and the repetitions were captured by Fourier Transform, where the periodic occurrence of heads shows as peaks in the frequency domain. A low-pass filter is applied to remove very high frequency content. The number of local maxima’s in the reconstructed image gave the people count. The image processing technique has been predominantly used in vehicular traffic to estimate the traffic density. However, there are very few studies which integrates both the image processing and crowd simulation.

Therefore, a major challenge in the analysis of crowd from real data in mass gatherings is crowd counting. Also, accounting for the heterogeneity of the crowd in such gatherings is a challenge by itself. Here, the first limitation is addressed by incorporating an image processing module to process the video data. The image-processing module will automatically detect people in the video and store the entry time of each people. This is then fed as an input to the crowd simulator to generate crowd risk.

The following section describes the details of the data collected. The subsequent section explains, in brief the methodology adopted. The next section highlights the results of the study. The last section concludes the paper with discussion on the future aspects of the study.

2. Data Collection

The data for the study was collected during Kumbh Mela held in Ujjain in 2016 from 22/04/16 to 21/05/16. About 75 million people participated in this Kumbh Mela, showing a significant increase in the pilgrims from 40 million in Allahabad Kumbh (2001), and 70 million in Allahabad Kumbh (2007). The data was collected using video cameras, smart phones, drones etc. In this study, the video recordings from Harsiddhi Chauraha, one of the major activity centers in Kumbh Mela is used as an input to the image processing module. Fig 2 shows a few snapshots of the different locations in Kumbh Mela.

Fig. 2. (a) Pilgrims taking part in Panchkroshi yatra; (b) Pilgrims moving along Gao Ghat, Kumbh Mela, Ujjain

3. Methodology

The study methodology adopted is shown in the flowchart (Fig. 3). The different modules are explained in brief.

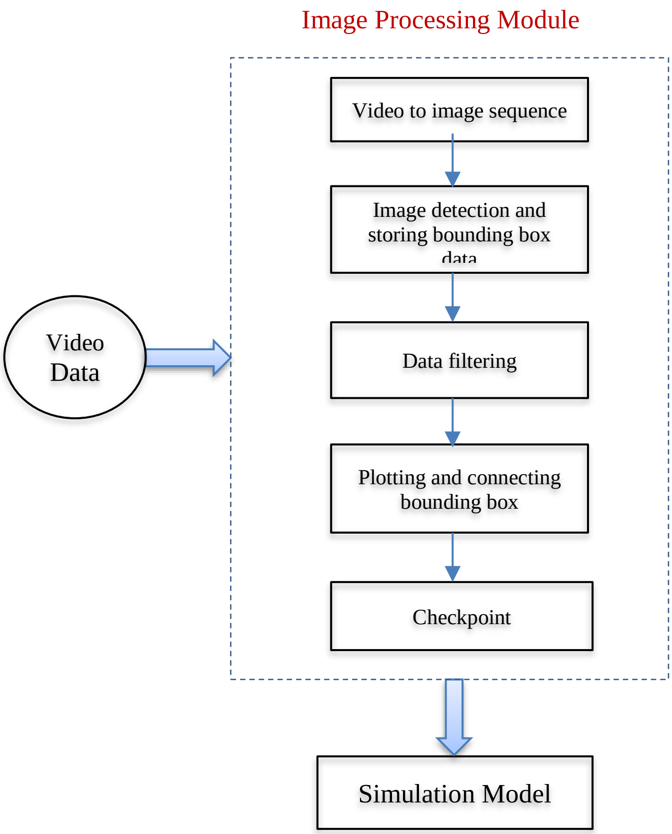

Fig. 3. Flowchart of study methodology

3.1. Image Processing Module

The video recorded from the Kumbh Mela Experiment Project is used here. The data is transferred to the Image Processing Module. Firstly, all the frames in a video are converted into images. Open Source Computer Vision Library (OpenCV), an open source computer vision and machine learning software library is used for this purpose. Next step is to detect people in images and store the bounding box data. You only look once (YOLO) is used for the detection. It is a state-of-the-art, real-time object detection system based on Darknet and Darknet is an open source Convolutional Neural Network framework written in C. The detected people are indicated by a square box. The detection confidence, frame number, detection class, detection box's centroid position, height, and width are stored in an excel file. A pre-trained weight file is used in YOLO. The detailed explanation of these steps is given in the following subsections.

3.1.1. Video to Image Processing

Open Source Computer Vision Library (OpenCV), an open source computer vision and machine learning software library is used to convert every frame of the video data file into a JPEG image.

3.1.2. Image detection and storing bounding box data

The detected people are indicated by a square box. The detection confidence, frame number, detection class, detection box's centroid position, height, and width are stored in an excel file.

3.1.3. Data filtering

Different types of filtering method are applied to the collected data.

- The data regarding classes other than people are deleted by checking the class id number. For detection class ‘person’ the class id is ‘0’.
- Some detections have either a comparatively low confidence score or very big bounding box. Therefore, data which have less than 0.8 confidence score or have size greater than the half of image’s width are deleted.

After doing these filtering, there were almost thirty-nine thousand data for only three thousand frames. To reduce the number of data points an upper and lower boundary is applied. The data outside of these two boundaries are deleted. This technique reduces the number of data points to twelve thousand.

3.1.4. Plotting bounding boxes

Python Image Library (PIL) is used to plot the bounding boxes for each person in every image. The algorithm takes the filtered bounding box data, appropriate image from user-defined image directory and draw the bounding box for every person in that image and store it in a user-specified destination folder. Fig.4 shows the pedestrian detection from the video data at different locations within the study area.

Fig. 4. (a) Pedestrians detected using IPM at Mahakal Dharmshala; (b) Pedestrians detected using IPM in Gao Ghat, Kumbh Mela, Ujjain

3.1.5. Bounding box connectivity

The next issue was to connect all the bounding boxes for a distinct person across different frames. To solve this problem, a simple search algorithm is used and a new parameter ‘searchRadius’ is defined. The algorithm will store centroid position of a pedestrian in a certain frame as a temporary variable and search for a point in next frame with previous centroid position as a centre and ‘searchRadius’ as the radius of the circle. If it can find a point in the next frame then it stores it in the matrix, update the temporary variables and follow the same steps until it can’t find any point. The rules in the search algorithm are as follows –

- Search in immediate frame using centroid coordinate.
- Search in immediate frame using top-left corner coordinate.
- Search in immediate frame using top-right corner coordinate.
- Search in immediate frame using bottom-left corner coordinate.
- Search in immediate frame using bottom-right corner coordinate.
- Search in next to next frame using centroid coordinate with 1.5 times ‘searchRadius’.

3.1.6. Using checkpoints to input the pedestrian inflow

This function detects when a pedestrian enters the simulation model. A straight-line equation is used. The algorithm takes two consecutive centroid point of a pedestrian trajectory and checks whether both are on the opposite side of the straight-line or not. If yes, then it stores the time in second, frame number and the point where the pedestrian crosses the straight line in an excel file. Fig. 5 shows the checkpoint (green line) and the pedestrians detected using the image processing module (IPM).

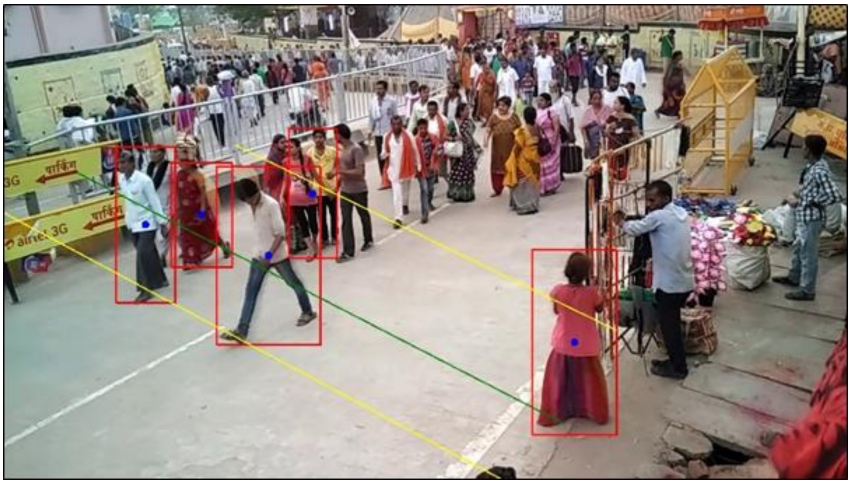

Fig.5. (a) Checkpoint of pedestrian inflow as indicated by the green line

3.2. Crowd Simulator

The pedestrian movement in the study section is simulated using a Social Force model (Helbing et al., 2005). The social force model is based on the premise that any pedestrian would be subject to three basic social force, namely; driving force which motivates a person to walk in his/her desired direction, pedestrian repulsion force which delineates a territorial kind of space around a person, and a border repulsion force which allows people to distance themselves from walls and obstacles. The net effect of these forces propels the movement of pedestrians.

The social force model function takes the list containing all the pedestrian instances as argument and updates the pedestrian kinematic values with the ones after the time interval, 't_int'. The checkpoints indicate the boundary at which the pedestrians enter the study area. The position and velocity of the pedestrians are calculated for each updation time based on the forces acting on them. The loop iterates through all the pedestrians and calculates the force on the pedestrians and calculates the three force functions and obtains the net force of each pedestrian in the x and y directions. It then updates the positions and velocities of the pedestrians using kinematic equations. Finally, the properties unique to every pedestrian is saved in a property matrix.

The simulation assumes a maximum desired speed to be 1.4m/s for a pedestrian with a circular radius of 0.3m. The lane under consideration was 3m wide and the updation time, t_int was taken to be 0.1s. The effect of varying the border threshold i.e., the distance beyond which the walls, do not influence was studied. It is observed that pedestrians tend to organize themselves into lanes as the border threshold increase. This is discussed in detail in the upcoming section.

A demo heat map based on a crowd risk index (CRI) can be developed for crowd risk assessment. Here, for demonstration purpose, the CRI was calculated based on an impatience factor, Imp_{wt} and the pedestrian density, $density\_{wt}$. The threshold values of various factors influencing CRI, beyond which, it is treated as a possibly danger situation is:

- $Imp_{wt} = 2$
- $density\_{wt}=1$

The value of Crowd risk index (CRI) is governed as follows:

$$\text{CRI} = f(Imp_t, Imp_{wt}, density_t, density_{wt})$$

The variations in CRI can be visualized using the heat map. A snapshot of the simulation and the resultant crowd risk is given in Fig. 6.

4. Results and Discussions

This section discusses the results of the study including the efficiency of the image processing module (IPM) and the crowd simulation of the study region. Fig. 6 shows the movement of a detected person and the trajectory of the centroid of his bounding box.

4.1. Simulation results

The Fig. 6 shows the simulation output for the study area under consideration.

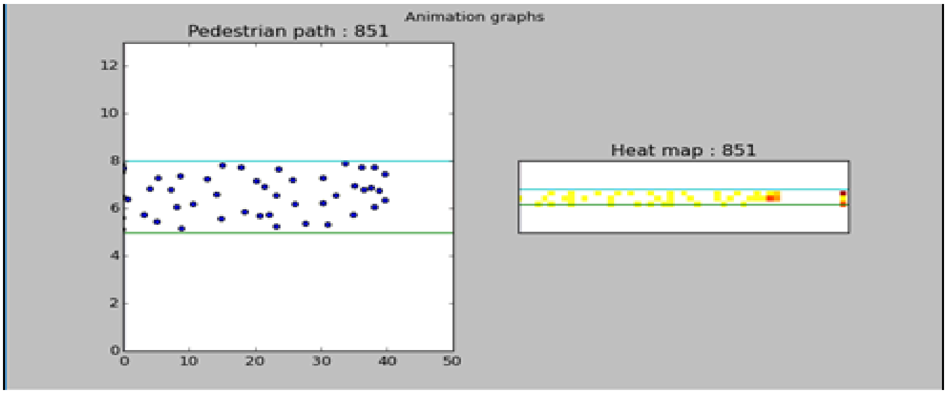

Fig.6. Pedestrian simulation of the study section and the associated crowd risk

4.2. Validating the Image Processing Module

The Table.1 gives a comparison between number of people crossing the checkpoint line detected by developed image processing module and manual count.

Table 1. Performance comparison between manual count and count by IPM

Start time(sec)	End time(sec)	Manual count	Count by IPM	Accuracy (%)
0	15	11	9	81.82
15	30	34	20	58.82
30	45	33	22	66.67
45	60	22	15	68.18
60	75	8	8	100.00
90	105	44	21	47.73
105	120	22	20	90.91

Fig.6. Tracked pedestrian and the trajectory of the centroid of the bounding box

The average accuracy of the Image Processing module is 73.45%. The IPM offers higher accuracy when the number of people is less. In time interval 0 to 15 sec and 60 to 75 sec the number of people is less than 11. The model is not able to identify people to a significant extent from videos recorded in the night-time. Therefore, a proper lighting condition while video recording can increase the level of detection. It is also noted that when the number of incoming people is greater than 30, the accuracy of IPM goes down to around 66%. The reason behind the accuracy drop is when large number of people are coming towards the checkpoint line; some of them are not visible due to other pedestrians in front of them. The problem can be solved partially by changing video camera angle such that almost every pedestrian will be visible. Another solution will be to train the IPM such that it can detect people based on their head.

4.3. Effect of varying border threshold

The crowd management authorities use a variety of methods to mange crowds in large gatherings such as Kumbh Mela. Temporary barricades such as iron railings, ropes are used to delineate the crowd moving in different directions. The behaviour of pedestrians when they encounter fixed walls or boundaries, temporary barricades, and porous boundaries such as road markings are different. This has been cited in many literatures (Gulhare et.al, 2018). Here, an attempt is made to understand the effect of varying the boundary threshold on pedestrian movement. Fig.7 shows a significant difference in the crowd movement. It could be seen that as the border threshold increases, pedestrians tend to organize themselves into lanes.

a b

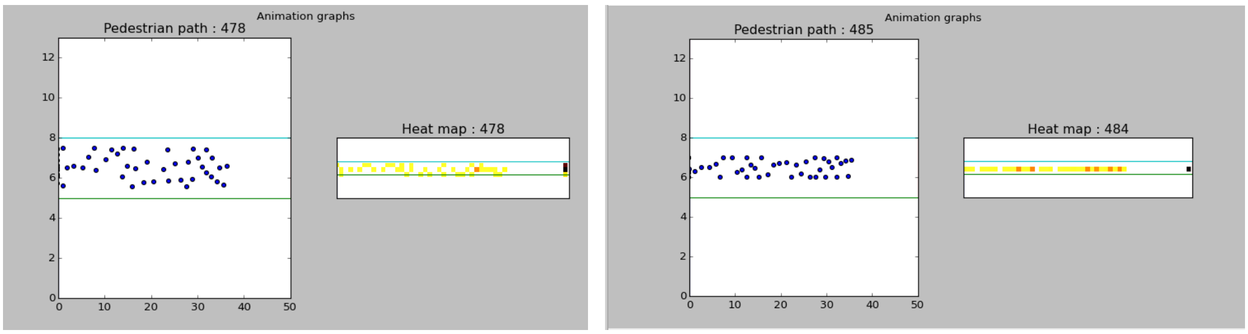

Fig.7. (a) Pedestrian simulation of the study section with border threshold of 0.5m; (b) Pedestrian simulation of the study section with border threshold of 1m

5. Conclusion

The main objective of this study was to build a system, which can detect incoming pedestrian automatically from the video camera data. For the pedestrian detection, YOLO, based on Darknet, an open-source Convolution Neural Networks is used. YOLO gives promising result without any manual training. It stores the detected pedestrian's bounding box height, width and centroid position in an excel file. Then a data filtering process runs through on the obtained data to erase unnecessary class detections, low confidence detections, false detections and data outside of the study region. A connection algorithm is applied for each person to connect all the bounding box associated with a distinct person in different frames. To integrate the Image Processing module with simulation model, a straight line is used as checkpoint. The time is stored when a pedestrian crosses the checkpoint and in the simulation on that time stamp, that pedestrian will be spawned. The whole model is tested on a Kumbh mela video data. The result obtained from the IP module is compared with the manual count. The IP module work well when number of people is less and gives an overall average accuracy of 73.45% without any manual training. However, extracting pedestrian trajectories in high density situations, and from low-resolution cameras needs to be explored further.

The crowd simulation model can be applied to different crowd situations by modifying the boundary conditions. However, to incorporate behavioral realism, the heterogeneity in the crowd needs to be accounted for. The Kumbh Mela, one of the largest mass gatherings celebrated in India, bringing together people of varying shades of faith, offers a unique opportunity to witness crowd dynamics on a mass scale. An early warning system which can predict the possible crowd risk situations beforehand would benefit the authorities to make tactical decisions.

Acknowledgements

The work reported in this paper is part of the project titled “The Kumbh Mela Experiment: Measuring and Understanding the Dynamics of Mankind’s largest crowd,” funded by the Ministry of Electronics and IT Ministry of Communication and Information Technology, Government of India (MITO-0105), Netherlands Organization for Scientific Research, NWO (Project No. 629.002.202), and Robert Bosch Center for Cyber Physical Systems, Indian Institute of Science, Bangalore. The authors also express their gratitude towards Kumbh Mela administration and government of Madhya Pradesh, India for providing constant support and official permissions to carry out research work and establish Indo-Dutch collaborative research camp at Kumbh Mela 2016.

References

Abbas, N., Tayyab, M., & Qadri, M. T.2013. Real Time Traffic Density Count using Image Processing. International Journal of Computer Applications, 83(9), 16-19. doi:10.5120/14476-2736

Dehghan, A., Idrees, H., Zamir, A. R., & Shah, M. 2013. Automatic Detection and Tracking of Pedestrians in Videos with Various Crowd Densities. Pedestrian and Evacuation Dynamics 2012,3-19. doi:10.1007/978-3-319-02447-9_1

Gayathri, H., Aparna, P., & Verma, A. 2017. A review of studies on understanding crowd dynamics in the context of crowd safety in mass religious gatherings. International Journal of Disaster Risk Reduction, 25, 82-91. doi: 10.1016/j.ijdrr.2017.07.017

Gulhare, S., Verma, A., Chakroborty, P. 2018. Comparison of Pedestrian Data of Single File Movement Collected from Controlled Pedestrian Experiment and from Field in Mass Religious Gathering. Collective Dynamics, 3. doi:10.17815/cd.2018.16

Helbing, D., Buzna, L., Johansson, A., & Werner, T. 2005. Self-Organized Pedestrian Crowd Dynamics: Experiments, Simulations, and Design Solutions. Transportation Science, 39(1), 1-24. doi:10.1287/trsc.1040.0108

Helbing, D., & Molnár, P. 1995. Social force model for pedestrian dynamics. Physical Review E,51(5), 4282-4286. doi:10.1103/physreve.51.4282

OpenCV library. (n.d.). https://opencv.org/&p=DevEx.LB.1,5069.1

Redmon, J. (n.d.). Retrieved from https://pjreddie.com/darknet/yolo/

Pillow¶. (n.d.). Retrieved from https://pillow.readthedocs.io/