

Urban Mobility India Research Symposium 2022

Detecting social groups using low mounted camera in mass religious gatherings

Authors

Nipun Choubey¹, Karthika PS², Gangadhar Reddy³ and Ashish Verma²

¹RBCCPS, ²Department of Civil Engineering, IISc, Bangalore

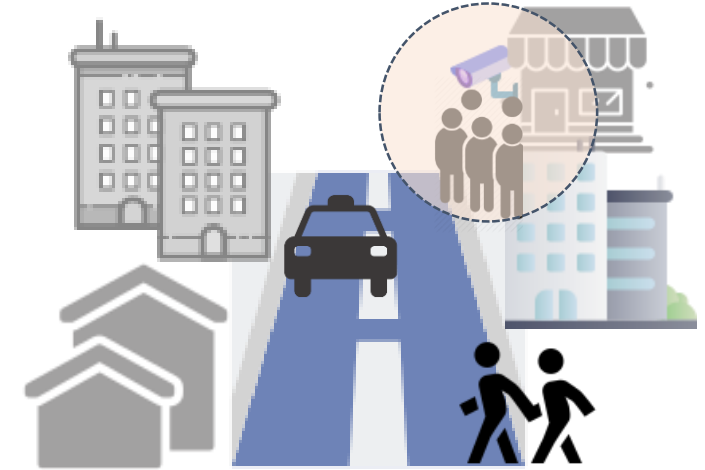
³Department of Civil Engineering, NIT Trichy

Outline

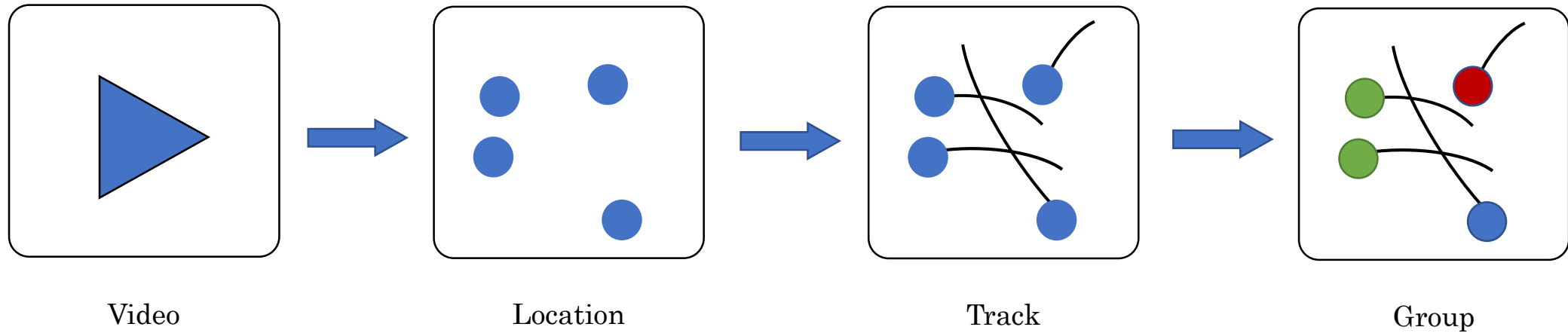
- Introduction
- Related Works
- Data Extraction
- Models
- Methodology
- Results
- Conclusion

Introduction

- In mass religious gatherings, people often come in group rather than individuals
- Event organisers use surveillance devices to get real time data
- Decisions are taken based on experience and varies from person to person
- Understanding groups pattern and movement could assist event organizers to plan things accordingly
- Detecting group is harder than detecting individual as the shape and size of groups changes with time.



Group detection



Identify the social connectedness based on visually observable parameters

Gaps

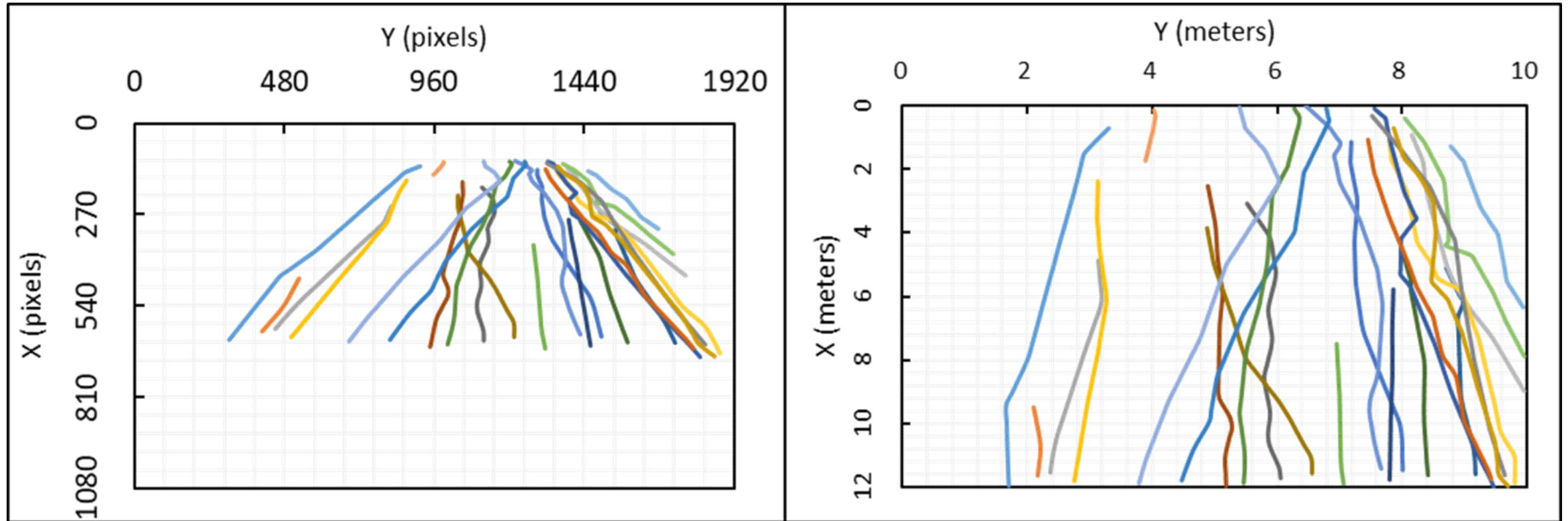
- Researchers generally try to mount camera higher and set angle close to normal from the ground.
- This allows them to avoid issues of perspective projection, occlusion etc.
- In this study, data was collected from a lower height and with lower angle from ground.
- Parameters such as distance/angle are calculated to classify whether the pair belongs to a group or not.

Data

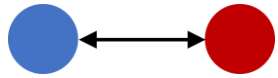


Clips	Length	# ped/frame
Clip1	101s	41
Clip2	147s	31

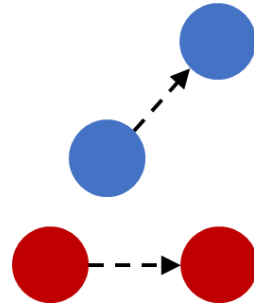
Pixel vs real-world trajectories



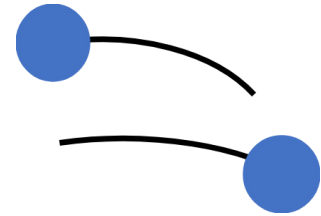
Parameters



Distance

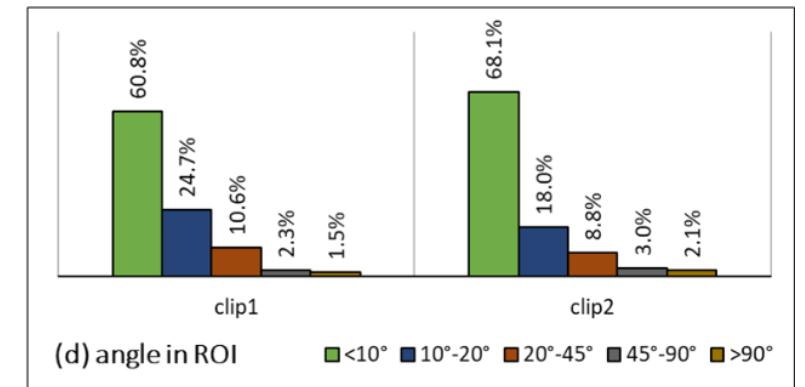
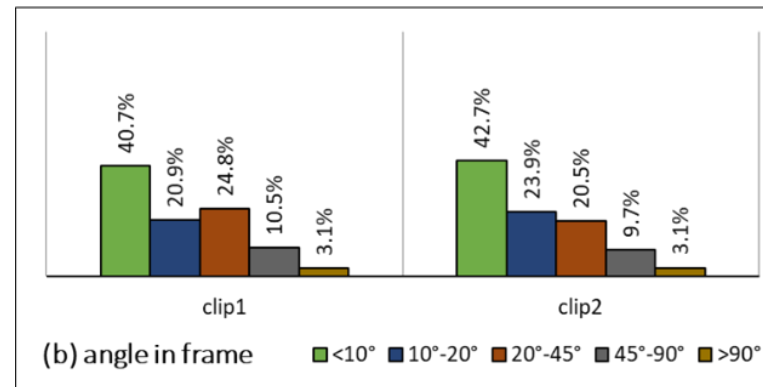
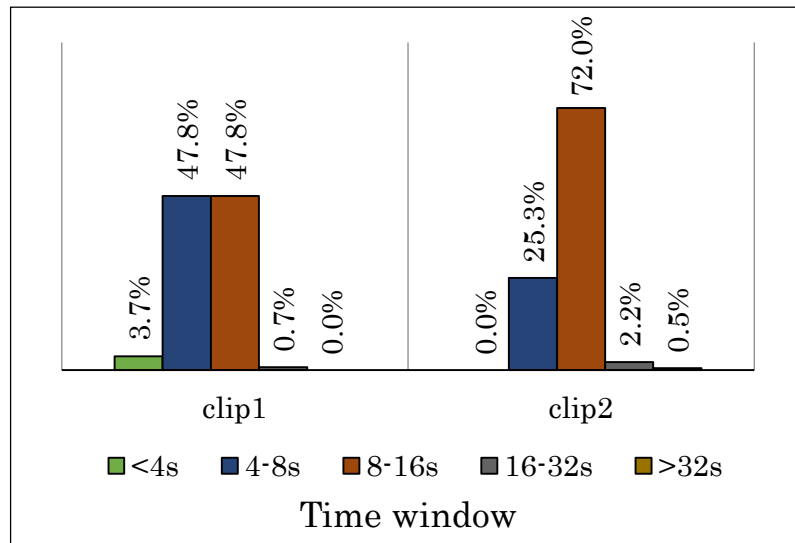
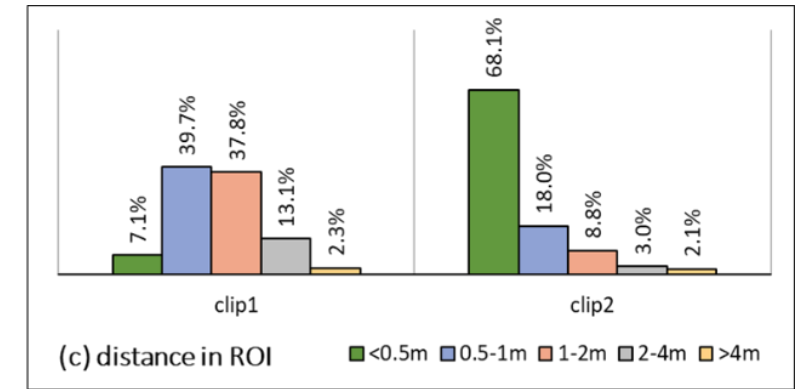
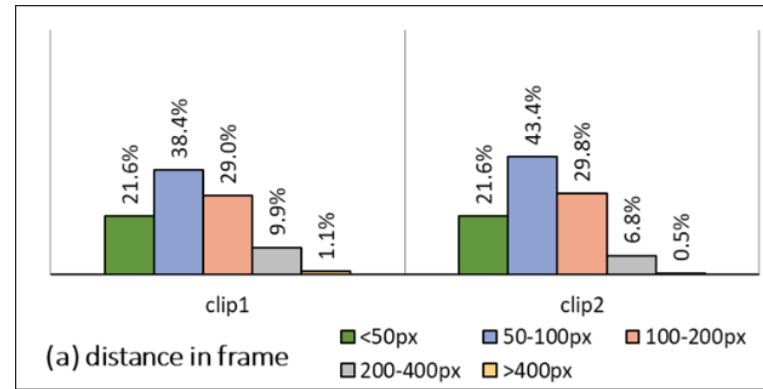
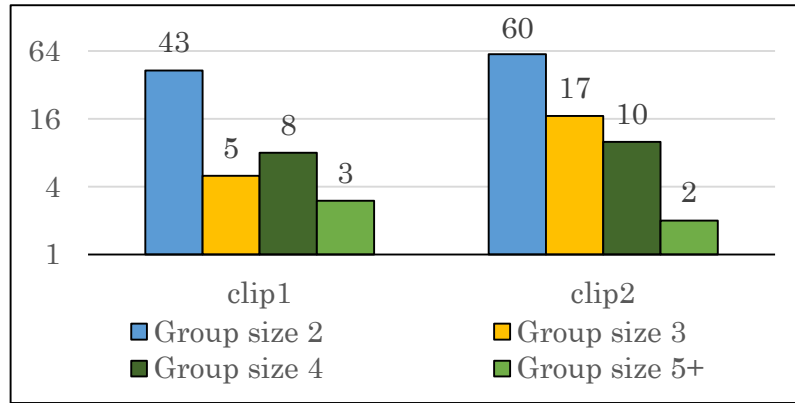


Angle

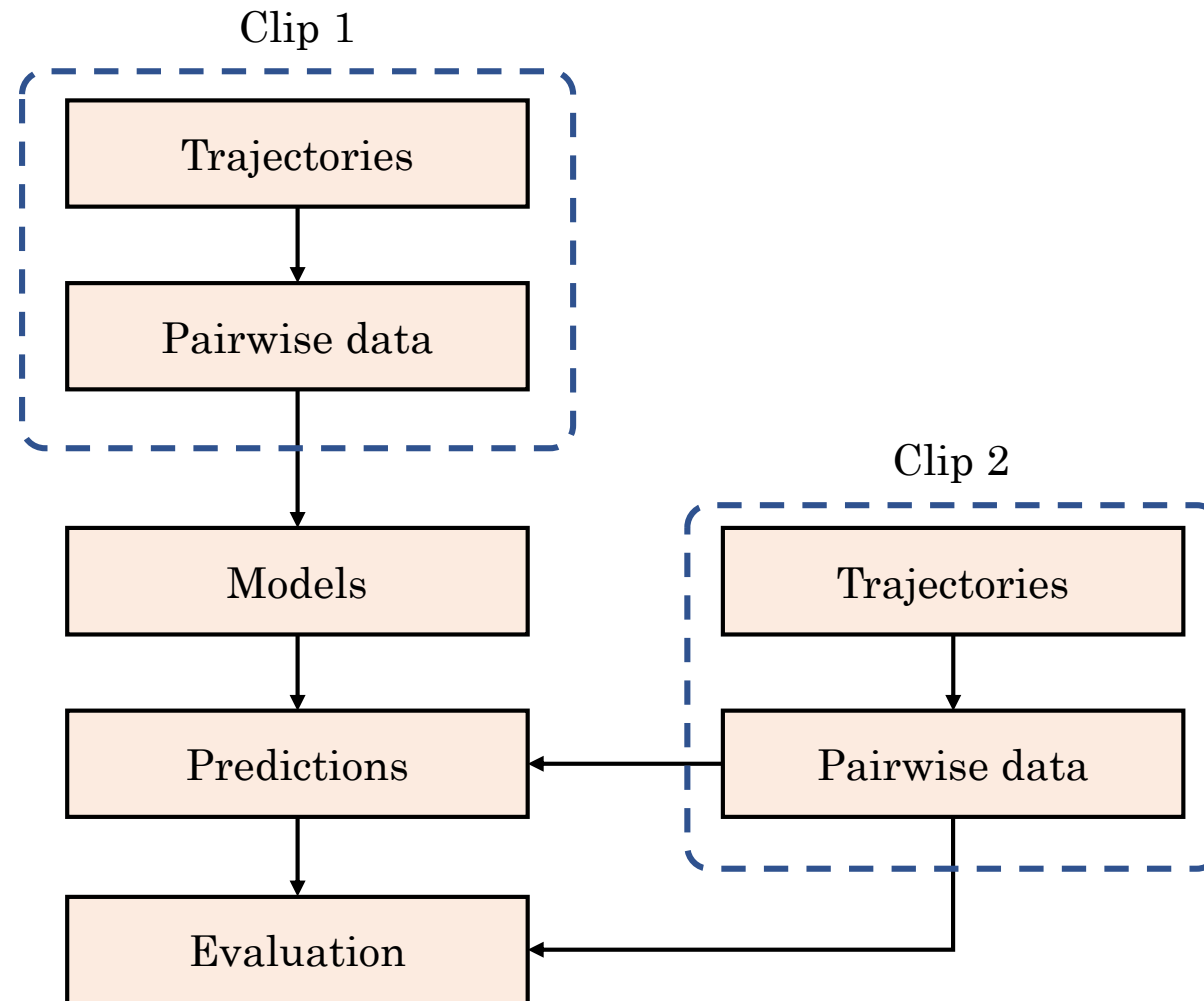


Tracklet info over
time

Parameter distribution



Methodology



Methodological flowchart

Group detection models

Threshold-based model

$$P_{t,d} = \begin{cases} 1, & \text{if } \bar{d}_{t,T} \leq D \\ 0, & \text{otherwise} \end{cases}$$

$$P_{t,a} = \begin{cases} 1, & \text{if } \bar{a}_{t,T} \leq A \\ 0, & \text{otherwise} \end{cases}$$

Linear Model

$$P_{t,d} = m_d \cdot \bar{d}_{t,T} + c_d$$

$$P_{t,a} = m_a \cdot \bar{a}_{t,T} + c_a$$

$$P_{t,g} = \left(1 + \exp \left(\frac{-\alpha P_t}{T} + \beta \right) \right)^{-1}$$

$$P_t = P_{t,d} \cdot P_{t,a}$$

Results

Clip	Model	Pixel based values					
		TP	FP	FN	TN	ACC	F1
Clip 1	Threshold	113	486	21	6431	0.928	0.308
	Linear	19	30	110	3812	0.965	0.213
Clip 2	Threshold	174	848	8	8624	0.911	0.289
	Linear	50	46	132	5581	0.969	0.360

Clip	Model	Real-world values					
		TP	FP	FN	TN	ACC	F1
Clip 1	Threshold	100	292	34	6625	0.954	0.380
	Linear	15	9	114	3833	0.969	0.196
Clip 2	Threshold	164	517	18	8955	0.945	0.380
	Linear	37	9	145	5618	0.973	0.325

Clip	Time window	Threshold model	
		ACC	F1
Clip1	T=0*	0.930	0.321
	T=1s	0.937	0.343
	T=2s	0.944	0.367
	T=4s	0.954	0.380
Clip2	T=0*	0.911	0.292
	T=1s	0.920	0.313
	T=2s	0.930	0.338
	T=4s	0.945	0.380

Conclusion

- Simple models have been proposed for group detection based on proximity and direction of movement.
- The real-world distance and angle information are found to be useful in identifying groups with low height/ angle mounted camera.
- Increase in accuracy is observed when time window is longer.
- Proposed models use reverse cumulative histogram to estimate model parameters. However alternate methods like joint distribution, Kalman filters, logistic regression can also be used.
- Further work
 - Consider all group members instead of closest neighbour.
 - Multi object tracking algorithm instead of using ground-truth
 - Visual features like gender and luggage

Thank you