Segmentation and Visualization of Crowd Flows in Videos using Hybrid Force Model

Shreetam Behera[1], Debi Prosad Dogra[1], Malay Kumar Bandyopadhyay[2] and Partha Pratim Roy[3]

[1]*School of Electrical Sciences, Indian Institute of Technology Bhubaneswar, Odisha, India*

[2]*School of Basic Sciences, Indian Institute of Technology Bhubaneswar, Odisha, India*

[3]*Department of Computer Science and Engineering, Indian Institute of Technology Roorkee, Uttarakhand, India*

Keywords: Crowd Flow, Smooth Particle Hydrodynamics, Langevin Equation, Crowd Phenomena, Crowd Behavior, Crowd Flow Segmentation.

Abstract: Understanding crowd phenomena is a challenging task. It can help to monitor crowds to prevent unwanted incidents. Crowd flow is one of the most important phenomena that describes the motion of people in crowded scenarios. Crowd flow analysis is popular among the computer vision researchers as this can be used to describe the behavior of the crowd. In this paper, a hybrid model is proposed to understand the flows in densely crowded videos. The proposed method uses the Smooth Particle Hydrodynamics (SPH)-based method guided by the Langevin-based force model for the segmentation of linear as well as non-linear flows in crowd gatherings. SPH-based model identifies the coherent motion groups. Their behavior is then analyzed using the Langevin equation guided force model to segment dominant flows. The proposed method, based on the hybrid force model, has been evaluated on public video datasets. It has been observed that the proposed hybrid scheme is able to segment linear as well as non-linear flows with accuracy as high as 91.23%, which is 4-5% better than existing crowd flow segmentation algorithms. Also, our proposed method's execution time is better than the existing techniques.

1 INTRODUCTION

In the last few decades, many researchers have put their interests in understanding collective motion behavior and analysis. Analyzing collective motion can help understand human behavior in groups in the context of visual surveillance. Understanding crowd behavior can help develop systems for monitoring and managing people in crowded scenarios. According to (Junior et al., 2010), computer vision-based techniques are highly popular for such surveillance purposes. Automation of visual surveillance can results in efficient crowd monitoring and management with higher accuracy, better information fusion, and, most importantly, reduction in human efforts to a greater extent. Automated visual surveillance-based systems can be used for identifying unusual behavior or activity in the crowd. This feature can help law-enforcing bodies to plan and take the necessary steps to avoid undesirable incidents.

In developing automatic crowd surveillance systems, flow detection, and segmentation are two instrumental elements in understanding crowd movements. According to (Zhang et al., 2018), crowd flows can be found using physical modeling-based methods or signal processing and machine learning-based methods. Physical modeling-based methods can be physics-based or particle-based techniques. These methods consider the crowd as fluid or system and process similar movement patterns contained within the flows.

In (Ali and Shah, 2007), Lagrange particle dynamics have been used to segment crowd flows. The authors have also done a stability analysis of the segmented crowd flows. However, their method is computationally intensive. The authors in (Ali and Shah, 2008) have proposed a force model based on static, dynamic, and boundary floor fields of the crowd. However, the method is sensitive to camera shake. Zhang et al. in (Zhang et al., 2017) have developed a streak-lines-based model to describe high-density crowd behaviors. In (Wu et al., 2017), bi-linear interaction of curl and divergence of the flows are analyzed for understanding crowd behaviors. The method proposed in (Lim et al., 2014) detects temporal flow variations in the crowd. In (Su et al., 2013), a spatio-

Behera, S., Dogra, D., Bandyopadhyay, M. and Roy, P.
Segmentation and Visualization of Crowd Flows in Videos using Hybrid Force Model.
DOI: 10.5220/0009328708610867
In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020) - Volume 4: VISAPP*, pages 861-867
ISBN: 978-989-758-402-2; ISSN: 2184-4321

temporal viscous fluid field-based technique has been proposed to recognize the large-scale crowd behavior from appearance and driven factor perspectives. An agent-based simulation method to monitor crowd density has been mentioned in (Basak and Gupta, 2017). Mehran et al. (Mehran et al., 2010) combined social force graph technique and streaklines to understand crowd flows in a video. In (Zhang et al., 2015), the social attributes-aware force model has been used for analyzing crowd movements. The social force model, in terms of attractive and repulsive forces, is used for crowd analysis in (Andrade and Fisher, 2005). A hybrid social influence model (HSIM) proposed in (Ullah et al., 2018) segments pedestrian motion in crowds. In (Lin et al., 2013), a heat map-based method has been proposed to understand group activities in crowd videos. Smooth Particle Hydrodynamics-based model, along with multi-layer spectral clustering, has been proposed to understand and detect coherent regions in density varying crowd scenarios in (Ullah et al., 2017). However, the method doesn't perform finer-level crowd flow segmentation.

None of the aforementioned addresses flow segmentation in terms of randomness in the crowd. Also, there is no such method that addresses the crowd in terms of hydrodynamics and random particles. This has been the primary motivation behind the work proposed in this paper. In this paper, a hybrid model is proposed for crowd flow segmentation. This hybrid model consists of two parts. The first part presents a Smooth Particle Hydrodynamics-based model that segments out the initial coherent regions, followed by a Langevin-theory guided force model that segments crowd flows irrespective of the randomness present in the system.

The paper is organized as follows. In Section 2, we explain the proposed hybrid force model and the relevant foundations. In Section 3, we present the outputs of the proposed model evaluated on two video datasets. Conclusion and Future directions are presented in Section 4.

2 PROPOSED HYBRID MODEL

In this section, we present the proposed hybrid model to segment motion flows in dense crowd videos. Crowd flow movements in videos can be represented analogously to fluid particles moving with different velocities in the fluid. While their movement, the particles experience different forces due to viscosity of the fluid, hydrodynamic interactions, and random events occurring within the fluid. In (Ullah et al., 2017), the authors have modeled coherency regions in the crowd. However, the model doesn't address the randomness in the crowd and only accounts for interactions similar to hydrodynamics interactions. The proposed hybrid model is a combination of two models, as illustrated in Figure 1. The first part of the model is based on Smooth Particle Hydrodynamics, which is used for coherence motion detection in videos. The coherent regions provide initial segmentation information that is then fed to the Langevin-based force model. This forms the second part of the hybrid system for performing crowd flow segmentation and accounts for the randomness in the crowd. As a result, both hydrodynamics and random behavior of motion particles are maintained in the proposed hybrid model.

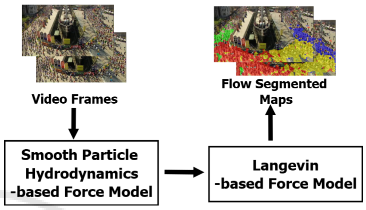

Figure 1: Block diagram representation of the proposed hybrid crowd flow segmentation model.

2.1 Smooth Particle Hydrodynamics Guided Grouping

Smooth-particle hydrodynamics (SPH) is a computational fluid dynamics method used for understanding fluid flows (Monaghan, 1992). It is a mesh-free method that does not need a mesh for spatial derivative calculations. A set of ordinary differential equations can describe the equations of momentum and energy of the fluid particles. According to the Navier-Stokes equation, the movement of fluid particles is guided by the force model presented in (1),

$$\frac{dv}{dt} = -\nabla P + \rho g + \mu \nabla^2 V \tag{1}$$

where ∇, P, ρ, g, μ, and V are derivative operator, pressure, density, acceleration, viscosity, and velocity, respectively.

The first term of (1) represents pressure force ($F_{pressure}$), the second term ($F_{external}$) represents external force due to other particles in the fluid and the third term represents viscosity force ($F_{viscosity}$). In a more generalization form, the equation (1) can be written as (2),

$$\frac{dv}{dt} = F_{pressure} + F_{external} + F_{viscosity} \tag{2}$$

where

$$F_{pressure} = m \sum_{i,N} K(r_c - r_i, \lambda) \tag{3}$$

$$F_{external} = m\sum_{i,N} \nabla K(r_c - r_i, \lambda) \quad (4)$$

$$F_{viscosity} = m\sum_{i,N} \gamma v \nabla^2 K(r_c - r_i, \lambda) \quad (5)$$

In the above equations, m represents the mass of the particle, v is the average velocity of the particle with respect to its neighboring particles, N represents the neighboring particles (r_1, ..., r_N) surrounding the central particle r_c and γ is the viscosity coefficient. $K(r_c - r_i, \lambda)$ is the smoothed kernel function that describes the physical properties of the central particle (r_c) by considering the relevant properties of the particles (r_i) falling in the range of the kernel. Here, the distance and density are considered as the relevant properties of the particles. Mathematically, the kernel function is represented using (6),

$$K(r_c - r_i, \lambda) = |K_t(r_c - r_i, \lambda) - K_{t-1}(r_c - r_i, \lambda)| \quad (6)$$

where λ is the set of neighboring particles influencing r_c. The term $K_t(r_c - r_i, \lambda)$ signifies the influence of the center particle r_c with the neighboring particles at time instant t.

$$K_t(r_c - r_i, \lambda) = \frac{315}{64\pi(\lambda)^9}((\lambda)^2 - \|r_c - r_i\|^2)^3 \quad (7)$$

By rearranging (2), a hydrodynamics-based magnitude represented in (8) can be obtained which is used for segregating the fluid particles with similar properties representing the group coherency among the particles.

$$Hydrodynamics_{mag} = \left(\frac{F_{external}}{F_{pressure}} + \frac{F_{viscosity}}{F_{pressure}}\right)K \quad (8)$$

2.2 Langevin Theory Guided Force Model

In physics, the Langevin equation is used to describe the dynamics of a Brownian particle (Langevin, 1908; Coffey and Kalmykov, 2004). According to (Schweitzer, 2007), the motion of the Brownian particle can be described as a combination of forces that are represented in (9).

$$m_i \ddot{\overrightarrow{r_i}} = -\gamma\dot{\overrightarrow{r_i}} + \overrightarrow{F_i} + \overrightarrow{R_i}(t) \quad (9)$$

where m_i is the mass of the particle, $\overrightarrow{r_i}$ is the position of the particle, $\dot{\overrightarrow{r_i}}$ represents the velocity of the particle, $\ddot{\overrightarrow{r_i}}$ is the acceleration acting upon the particle, $-\gamma\dot{\overrightarrow{r_i}}$ is the drag force responsible for removing force caused due to friction with γ as the viscosity coefficient (satisfying $\frac{K_BT}{D}$), F_i is the force exerted on the i^{th} particle, and $R_i(t)$ is the random force describes the randomness of the particle and it should satisfy (10). The above equation (9) holds good for passive systems with Brownian particles,

$$\langle\overrightarrow{R_i}(t)\rangle = 0,\ \langle\overrightarrow{R_i}(t)\overrightarrow{R_j}(t')\rangle = 6K_BT\gamma\delta_{ij}\delta(t-t') \quad (10)$$

where $\langle\overrightarrow{R_i}(t)\rangle$ represents an average value considered with respect to the distribution of the realizations of the variable $\overrightarrow{R_i}(t)$, and K_B is Boltzmann's constant, T is the temperature, D is the dissipation coefficient and δ is the delta function. Upon these particles, (9) can thus be applied to estimate the dynamics of the particle in terms of position and velocity.

2.3 Crowd Flow Segmentation

This section discusses the implementation of a hybrid flow segmentation model. The proposed crowd flow model works on a windowing scheme. The windowing scheme ensures periodic re-initialization of the process. This reduces the load of tracking of the flow changes happening in the temporal domain. For a given window W, the first two consecutive frames are used for dense optical flow calculation to generate magnitude and orientation maps.

For a certain magnitude threshold, the orientation map is quantized into four different bins within the range of 0-2π. This step is useful in two ways. Firstly, magnitude-based thresholding removes motion noise. Secondly, the orientation is generalized into fixed values for these points.

Then, the hydrodynamics-based magnitude is calculated, as mentioned in (8). The motion points flowing in the same quantized directions and over a particular $Hydrodynamics_{mag}$ threshold are then retained and grouped based on the connected component analysis. This step identifies the coherent regions in the crowd. The coherent points are then fed to the Langevin-based force model for performing segmentation in the remaining frames in the window. The parameters like γ and λ have been set to 0.3 and 5 during empirical evaluation. The magnitude and $Hydrodynamics_{mag}$ thresholds are set to be 0.4 and 0.1 based on empirical study. F is basically a group force which is computed as the cumulative sum of acceleration of the particles within a specific neighborhood as mentioned in (11),

$$F_{drift/confinement} = m\sum_i \frac{dv_i}{dt} \quad (11)$$

where v represents the velocity of the particle, and m is initialized to unity throughout the experiment. In this work, we calculate F using (11) as drift force causing the particle to drift along x direction, and F is a confinement force causing the particle to be confined along y direction. We assume drift movement to

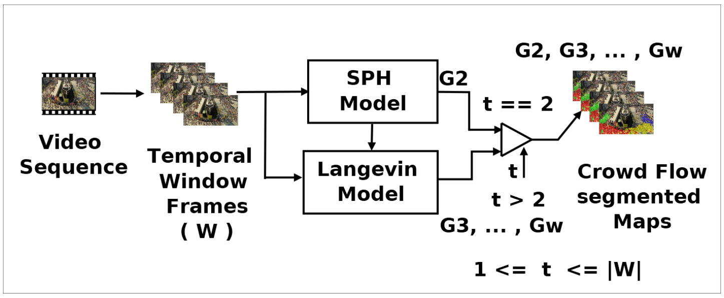

Figure 2: Proposed Implementation of the hybrid crowd flow segmentation model.

be a positive phenomenon and particle confinement to be a negative phenomenon, thus assigning the polarity of forces accordingly. The random force is generated within $(0-1)$. The block diagram of the hybrid model is presented in Figure 2.

3 RESULTS AND DISCUSSIONS

The evaluation of the proposed model has been carried out on two video datasets. One is a publicly available dataset, i.e., Marathon Video used in (Shao et al., 2014) and the other dataset contains more than ten hours of video recording of a popular event known as Puri Rath Yatra that happens at Puri, Odisha every year in India during the month of July. The Intersection over Union (IoU) metric presented in (12) is used to quantify the percent overlap between the ground truth and segmentation output.

The Marathon dataset is a dense crowd video, where people are running in an elliptical path. The path formed is non-linear in different directions. The output of the proposed method can be seen as segmented flows of four different directions represented by four different colors, indicating the four different directional flows, similar to the ground truths, as displayed in Figure 3.

The Rath Yatra video is a semi-dense crowd video with a certain degree of randomness. In this video, both structured and unstructured motions can be observed. People can be seen pulling the car (Rath) that can be considered as a structured linear motion. Along the sides of the linear flow, people can be seen moving towards different directions, thus forming unstructured motion. The proposed method is able to segment such linear flow to a greater extent, matching closely with the ground truths, as evident in Figure 4. Accuracy has been calculated using (12).

$$Accuracy = \frac{Area(S_w \cap G_T)}{Area(S_w \cup G_T)} \qquad (12)$$

where S_w is the segmented image, and G_T is the ground truth image. It has been found that the proposed method outperforms the basic SPH method proposed in (Ullah et al., 2017). The proposed hybrid method shows good improvement in both accuracy and execution time which is represented in Table 1, Figure 5 and Figure 6, respectively. This happens because the proposed method estimates the position and velocity of the particles using the Langevin-based model. Therefore, there is no need to apply optical flow in every frame. As a result, a significant amount of computation time can be saved.

Table 1: Comparison of the proposed hybrid model with basic SPH method proposed in (Ullah et al., 2017). The comparison has been done with respect to accuracy and average time taken per frame in seconds.

#Videos	Accuracy (in %)		Time Taken for execution per frame (in seconds)	
	SPH Method (Ullah et al., 2017)	Proposed Method	SPH Method (Ullah et al., 2017)	Proposed Method
Marathon	90.37	91.23	3.89	2.32
Rath Yatra	77.12	81.29	2.69	1.97

The proposed hybrid force model can segment both linear and non-linear crowd flows. The hybrid force model detects the coherent motion regions in the video frames. The Langevin-based force model is able to trace the segmented flows within each window. This approach makes the proposed model faster than the method proposed in (Ullah et al., 2017) since it is not necessary to apply the Smooth Particle Hydrodynamics on every frame. The reliability of the Langevin-based force model can be observed in the temporal segmented maps presented in Figure 3 and Figure 4, respectively. The segmented regions can be used as input data for machine learning models in order to detect abnormal activities in the crowd.

4 CONCLUSION AND FUTURE DIRECTIONS

Crowd flows are essential components in crowd analysis to understand crowd movements. Thus, understanding crowd movement behavior can help law-

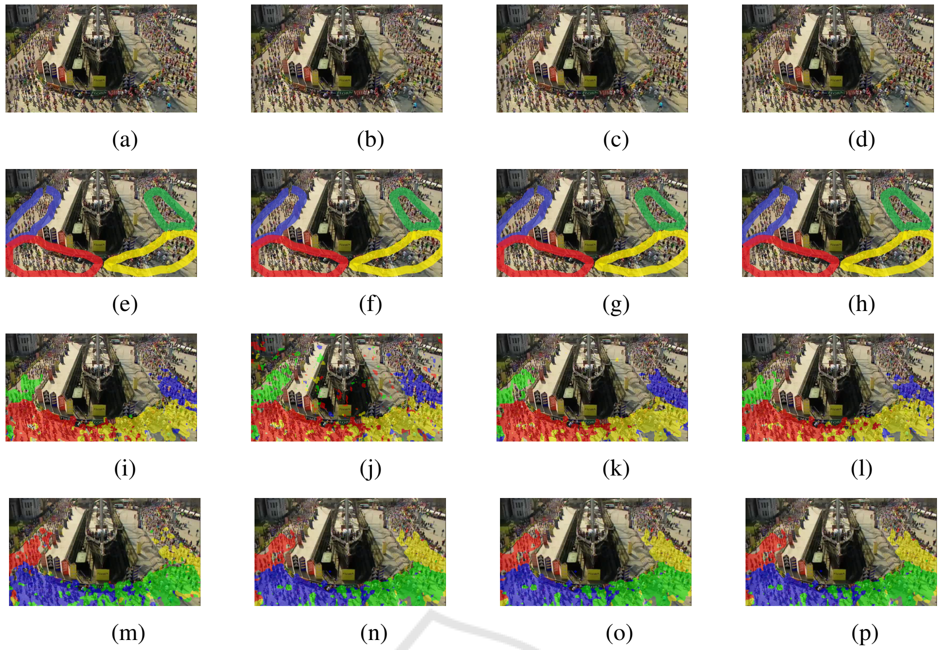

Figure 3: (a-d) Original Frames (361-364) of the Marathon video (e-h) Ground Truth Frames, (i-l) represent outputs of method proposed in (Ullah et al., 2017), (m-p) represent outputs of the proposed segmentation method, respectively. (Best viewed in color).

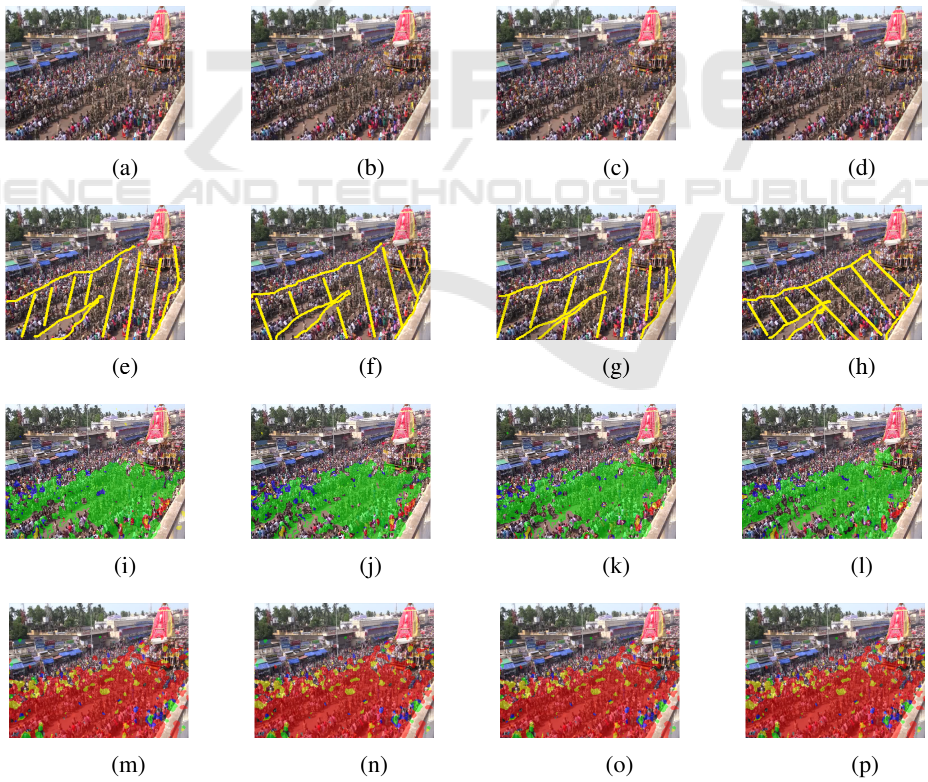

Figure 4: (a-d) Original Frames (31-34) of the Rath Yatra video (e-h) Ground Truth Frames, (i-l) represent outputs of method proposed in (Ullah et al., 2017), (m-p) represents the output of the proposed method, respectively. (Best viewed in color).

enforcing agencies to take necessary actions to avoid unwanted accidents. Therefore, it is essential to segment the crowd flows efficiently. The proposed segmentation method based on a hybrid method can seg-

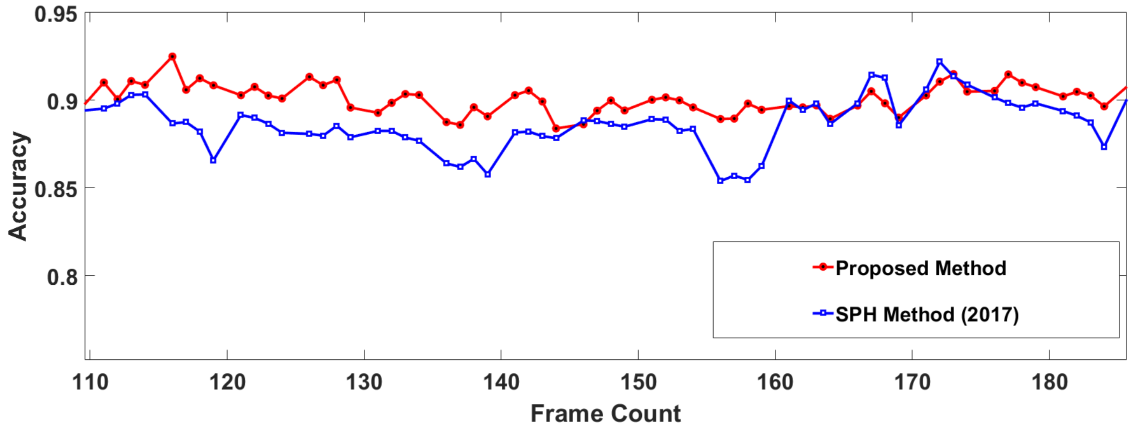

Figure 5: Segmentation Accuracy for the proposed method (red) and the SPH method (Ullah et al., 2017) (blue) for Marathon Video.

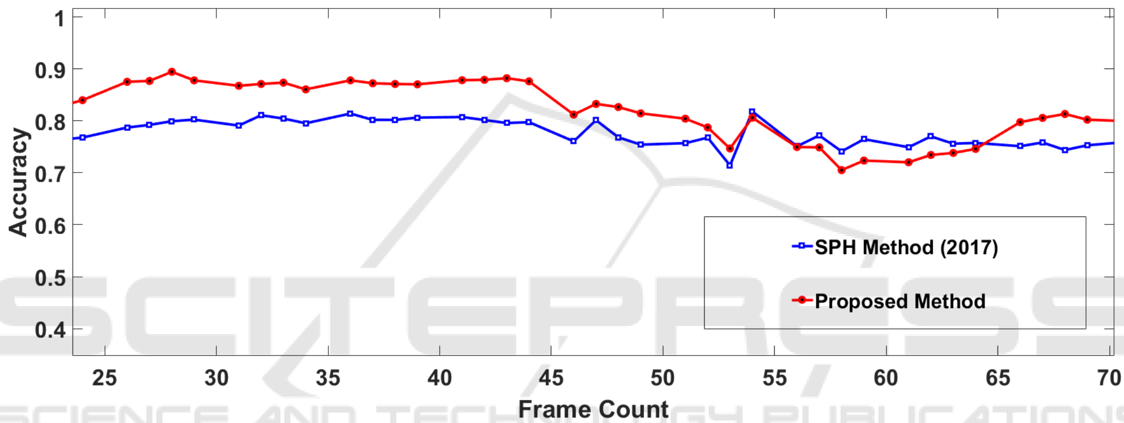

Figure 6: Segmentation Accuracy for the proposed method (red) and the SPH method (Ullah et al., 2017) (blue) for Rath Yatra Video.

ment both linear, and non-linear motion flows with considerable accuracy. The proposed hybrid model comprises of particle-based and physics-based force models. The particle-based force model segments the coherent regions effectively. The physics-based force model then uses these regions for flow segmentation in the successive frames. As a result, the proposed model is better in terms of accuracy and speed when compared to existing SPH-based methods.

The model also ensures robustness in the segmentation of the flows, even in the presence of a high degree of randomness in the crowd. There are a few possible extensions of the present work. For example, the SPH model of our proposed system can be replaced with advanced mesh-free methods to increase the robustness and efficiency of the flow segmentation algorithm. Moreover, the method can be amalgamated with machine learning techniques for anomaly detection and classification.

ACKNOWLEDGEMENTS

The authors would like to thank the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India for funding this research work through the grant YSS/2014/000046.

REFERENCES

Ali, S. and Shah, M. (2007). A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6.

Ali, S. and Shah, M. (2008). Floor fields for tracking in high density crowd scenes. In *European Conference on Computer Vision*, pages 1–14. Springer.

Andrade, E. L. and Fisher, R. B. (2005). Simulation of crowd problems for computer vision. In *First Inter-*

national Workshop on Crowd Simulation, volume 3, pages 71–80.

Basak, B. and Gupta, S. (2017). Developing an agent-based model for pilgrim evacuation using visual intelligence: A case study of ratha yatra at puri. *Computers, Environment and Urban Systems*, 64:118–131.

Coffey, W. T. and Kalmykov, Y. P. (2004). *The Langevin equation: with applications to stochastic problems in physics, chemistry and electrical engineering*. World Scientific.

Junior, J. C. S. J., Musse, S. R., and Jung, C. R. (2010). Crowd analysis using computer vision techniques. *IEEE Signal Processing Magazine*, 27(5):66–77.

Langevin, P. (1908). Sur la théorie du mouvement brownien. *CR Acad. Sci. Paris*, 146:530–533.

Lim, M. K., Chan, C. S., Monekosso, D., and Remagnino, P. (2014). Detection of salient regions in crowded scenes. *Electronics Letters*, 50(5):363–365.

Lin, W., Chu, H., Wu, J., Sheng, B., and Chen, Z. (2013). A heat-map-based algorithm for recognizing group activities in videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(11):1980–1992.

Mehran, R., Moore, B. E., and Shah, M. (2010). A streakline representation of flow in crowded scenes. In *European Conference on Computer Vision*, pages 439–452. Springer.

Monaghan, J. J. (1992). Smoothed particle hydrodynamics. *Annual review of astronomy and astrophysics*, 30(1):543–574.

Schweitzer, F. (2007). *Brownian agents and active particles: collective dynamics in the natural and social sciences*. Springer.

Shao, J., Change Loy, C., and Wang, X. (2014). Scene-independent group profiling in crowd. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2219–2226.

Su, H., Yang, H., Zheng, S., Fan, Y., and Wei, S. (2013). The large-scale crowd behavior perception based on spatio-temporal viscous fluid field. *IEEE Transactions on Information Forensics and Security*, 8(10):1575–1589.

Ullah, H., Ullah, M., and Uzair, M. (2018). A hybrid social influence model for pedestrian motion segmentation. *Neural Computing and Applications*, pages 1–17.

Ullah, H., Uzair, M., Ullah, M., Khan, A., Ahmad, A., and Khan, W. (2017). Density independent hydrodynamics model for crowd coherency detection. *Neurocomputing*, 242:28–39.

Wu, S., Su, H., Yang, H., Zheng, S., Fan, Y., and Zhou, Q. (2017). Bilinear dynamics for crowd video analysis. *Journal of Visual Communication and Image Representation*, 48:461–470.

Zhang, D., Xu, J., Sun, M., and Xiang, Z. (2017). High-density crowd behaviors segmentation based on dynamical systems. *Multimedia Systems*, 23(5):599–606.

Zhang, X., Yu, Q., and Yu, H. (2018). Physics inspired methods for crowd video surveillance and analysis: a survey. *IEEE Access*.

Zhang, Y., Qin, L., Ji, R., Yao, H., and Huang, Q. (2015). Social attribute-aware force model: exploiting richness of interaction for abnormal crowd detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(7):1231–1245.