PostDoc Journal
Vol. 1 No. 1, January 2013

Journal of Postdoctoral Research
www.postdocjournal.com

A Review of Physics-based Methods for Group and Crowd Analysis in Computer Vision

Hyunjoon Jo*, Kabir Chug†, and Ricky J. Sethi‡
USC Information Sciences Institute
Marina del Rey, CA 90292 USA
*hyunjoon@usc.edu, †kabir.chug@gmail.com, ‡rickys@sethi.org

***Abstract*—Crowd analysis is a popular topic in computer vision, with important applications to video surveillance, social media analysis, and multimedia retrieval, to name just a few areas. In this paper, we review some of the physics-based methods for group and crowd analysis in computer vision. In particular, we examine approaches for physics-based analysis of groups, crowds, and the simulation of crowds. The purpose of this review is to categorize and delineate the various physics-based and physics-inspired approaches that have been applied to the examination of groups and crowds in video.**

I. INTRODUCTION

The dynamic nature of video makes activity search and recognition from video databases a very difficult problem [28, 1, 2, 18, 26, 27]. This is especially apparent when analyzing groups and crowds in video. Given the importance of such analyses for video surveillance, social media analysis, multimedia retrieval, and a host of other applications, many researchers have started turning to established physics concepts in order to combat this arduous task. Some of these methods are inspired by physics while others are inherently rooted in fundamental physics ideas.

In this paper, we give an overview of some of the more pertinent physics-based and physics-inspired methods for tackling the important problem of group and crowd analysis in video. We start by examining how researchers have used methods inspired by basic physics concepts to determine the transition from individuals to groups and crowds in Section II. Then, in Section III, we review some of the rich body of work in physics-based crowd analysis. Finally, in Section IV, we turn to the problem of crowd simulations, which is rooted in physics principles.

II. PHYSICS-BASED GROUP ANALYSIS

Before we can analyze crowds, we have to make the determination that individuals have come together to actually form a group or a crowd. Most researchers, like [14, 13, 31], simply assume a group is any collection of individuals; others [30] narrowly define a group as being determined by the social relationships amongst pedestrians. [12] goes further and talks about group formation in terms of personal space, proxemics, and spatial relationships amongst people. In particular, they use Perceived Personal Space and People Comfort as the measures to detect grouping as either voluntary or involuntary. [17] also posits a qualitative definition from sociology of a group as a collection of individuals with a common goal and defines a crowd as a large group of individuals in the same environment sharing a common goal.

Although there are a few such *qualitative* models of groups and crowds in computer vision, [23] is inspired by fluid dynamics and formulates a *quantitative* model for the transition of individuals to groups and crowds. Specifically, [23, 24, 25] model the transition of Individuals $\Rightarrow$ Groups $\Rightarrow$ Crowds analogously to the transition of Individual Particles $\Rightarrow$N-Body $\Rightarrow$Fluids in fluid dynamics, as shown in Figure 1. They thus define the *Group Transition Ratio,* G_{tr}, as:

$$G_{tr} = \frac{L}{\lambda} \tag{1}$$

where λ is the mean free path and L is the characteristic length. The G_{tr} allows a quantitative characterization of the formation and dispersion of groups, as well as identification of so-called Atomic Group Actions by examining the time variation of the G_{tr}.

III. PHYSICS-BASED CROWD ANALYSIS

Once a group or crowd is formed, we can employ any of a plethora of crowd-analysis methods based on physics to analyze the situation. [11] notes that crowds can be categorized using the image space domain, the sociological domain, the level of services, or the computer graphics domain. The image space domain dynamics inform about crowd formation when the density of people gets too large to do individual tracking. Crowd density estimation models are currently based on three prevalent model types: pixel-level analysis, texture analysis, and object-level analysis. The sociological domain deals with crowd mentality, in which crowds form a collective psychology in response to different triggers; the most obvious amongst these are: least effort hypothesis, lane formation, and the bottleneck effect. The level of services provides different crowd conditions in terms of the density of people and it's temporal evolution. Finally, the computer graphics domain deals with crowd simulation on different levels to achieve realistic models for crowd behaviour.

Many approaches, like [10, 3, 29, 15, 4, 22, 16], have used physics-based approaches for analyzing the image space domain. Although [15] uses joint models of appearance and dynamics, called Dynamic Textures, to model complex dynamic scenes and crowd abnormalities, [4] deals with the important problem of tracking in high density crowds. This

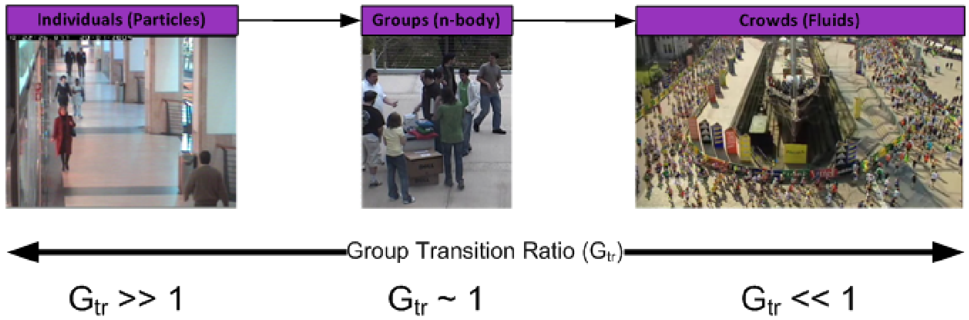

Fig. 1. Physics-inspired model of transitions from individuals to crowds: modeling individuals as free particles, groups as an n-body, and crowds as fluids in [23].

is challenging since, as the density of objects increases, the number of pixels on each individual object decreases. The constant interaction amongst the individuals and constant occlusions by inter-object interactions also makes this task difficult. Finally, the goal-directed dynamics and psychological characteristics of a crowd influences how individuals in the crowd behave. Their approach is based on the observation that the motions of individuals in crowded scenes depends upon the space-time interactions of individuals amongst themselves as well as with the scene layout. This information about natural crowd flow and scene constraints can be used as priors to impose high-level direction for tracking purposes. The crowd flow information and the scene constraints are encapsulated by the idea of floor fields, which model the interactions between individuals and their preferred direction of movement by transforming long range forces into local ones that affect the instantaneous probability of those local moves. The transition probability of a tracked person then depends on the strength of the floor field in his/her neighborhood. The concept of a floor field is inspired by the field of evacuation dynamics, where floor fields are manually designed to simulate behaviors of pedestrians in panic situations and they compute three floor fields from the visual data: a “Static Floor Field”, a “Boundary Floor Field”, and a “Dynamic Floor Field”, which can be picturized in Figure 2.

In [3], they handle crowded scenes by treating crowds as a dynamic system in a specified field by using Lagrangian Particle Dynamics for the segmentation of high density crowds and detection of flow instabilities. They start by representing moving crowds as an aperiodic dynamical system that is manifested by a time-dependent flow field. A grid of particles is then overlaid on this flow field and then advected using numerical integration. The evolution of the grid of particles through this flow field is tracked using a Flow Map; their maximum eigenvalue is used to construct a Finite Time Lyapunov Exponent field which, in turn, reveals the Lagrangian Coherent Structures in the underlying flow. The existence of these coherent structures is key for their theory as they divide the flow into regions of qualitatively different dynamics.

[16] works from the perspective that people in crowds, in some ways, behave as particles in fluids. As they point out, crowds where there are few interactions between people behave like gases and can be modelled using aerodynamics. However, crowds where interaction forces tend to dominate the motion of people can be modelled as a liquid using hydrodynamics. They do this at both the macroscopic scale for crowd segmentation and the mesoscopic scale for behaviour detection. All their methods rely upon optical flow and the associated particle advection from a Lagrangian approach to fluid dynamics and builds upon the work of [10, 3, 29, 4, 22].

IV. Physics-based Crowd Simulations

One of the main limitations in crowd analysis, however, is the dearth of datasets consisting of videos of crowds. Thus, many researchers rely upon robust, increasingly realistic simulations of crowds as described in [19], which gives a comprehensive survey of the techniques and requirements for simulating large-scale virtual human populations ranging from computational crowd models to functional models of human behaviour. [17] uses a distributed random behavioural model but uses simple vector analysis to avoid collisions. [20] addresses the problem of real-time virtual environments and develops a control model for local motions and a global path planning algorithm in such circumstances. They use both psychological attributes and geometrical rules like distance, areas of influence, and relative angles, to eliminate artifacts. Others have gone further; e.g., [7, 8] used particle systems and dynamics for modeling the motion of groups with physics. Specifically, using a point-mass system model, their group behaviour algorithm computed goal positions for individuals based on the current positions and velocities of their neighbours, obstacles, and the global desired group velocity, as shown in Figure 3. They went further and reproduced movements of legged robots, bicycle riders and point-mass systems based on basic dynamics.

[5] modeled people as a set of interactive particles and then adapted the use of particle systems for studying crowd movements. Their model for the motion of individuals utilizes Newtonian forces but also incorporates human goals and decisions. Their concept of the decision charge of a person interacting with a surrounding decision field was modelled after the way an electric charge is influenced by an electric field. Their model for crowd simulation in immersive space management used particle systems as a generic model for simulations of dynamic systems.

[9] describes methods to simulate the movement of pedestrians based on a social force model which is a microscopic

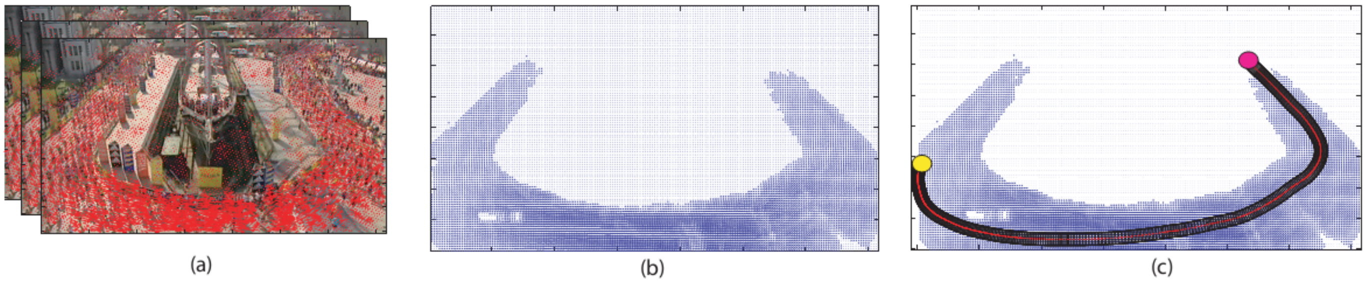

Fig. 2. (a) Dense Optical Flow; (b) Computed Point Flow Field; and (c) Sink Seeking Process from [4].

(personal) approach for simulating pedestrian motion. In their approach, they solve Newton's equation for each individual and consider repulsive interactions, friction forces, dissipation and fluctuations. [6] builds upon the Helbing model [9] and generalizes it to deal with different individualities for agent and group behaviours. The fundamental dynamical equation they utilize for varying velocity, v, with forces, f, and direction, e, over a certain time interval, τ, is:

$$m_i \frac{dv_i}{dt} = m_i \frac{v_i(t)e_i(t) - v_i(t)}{\tau_i} + \sum_{i \neq j} f_{ij} + \sum f_{iw} \quad (2)$$

The left side of this equation computes the net force (following Newton's Third Law of Motion) as mass (m_i) times acceleration, where the acceleration, $a = \frac{dv_i}{dt}$, which is used to computer the change in velocity, v. The velocity, in turn, can be used to computer the change of position, $r(t)$, as $v(t) = \frac{dr}{dt}$.

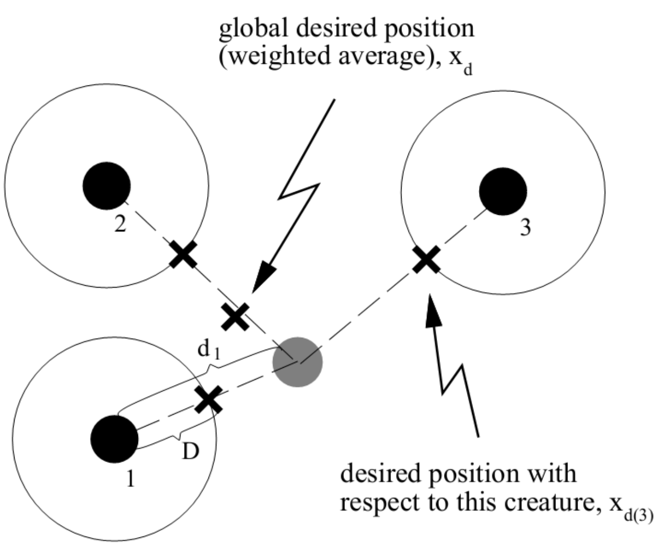

Fig. 3. Positions of the other visible members of the group (called creatures in their simulations) is used to compute global desired positions for each individual in [7].

The Helbing model is based on physics and also uses socio-physiological forces to describe human behavior in panic situations using particle systems. The use of particle systems, however, harkens back to [21], who first developed a model for simulating a school of fish and a flock of birds using a particle systems method. Particle systems were used to represent simulated birds as particles and their aggregate motion was created by a behavioural model where each individual bird navigates according to its own local dynamic environment and dictated by the simulated physics of motion.

V. Conclusion

In this paper we have reviewed some of the more pertinent work in group and crowd analysis that utilize physics methods to varying extent. We have classified these different approaches as dealing with group analysis, crowd analysis, or crowd simulation. In our review, we identified a large corpus of research and selected some of the more representative papers. Although this paper reveals important progress made in the field of physics-based methods for group and crowd analysis, we intend to expand this review into a more comprehensive survey in the future.

Important issues that still need to be addressed in future work are identifying significant datasets for group and crowd analysis, elaboration of abnormality detection, and surveying the social force and behavioural models that are also utilized in crowd analysis. One of the most significant problems facing group and crowd analysis, however, is the dearth of widely-available, realistic datasets. There is a glaring need for datasets that address specific actions like group formation and fundamental, atomic group actions, as well as realistic footage of surveillance, sports actions, movies, and other such video data from the Internet. Identifying, creating, and analyzing such datasets will no doubt challenge the current state-of-the-art in group and crowd analysis.

VI. Acknowledgment

This work was partially supported by the National Science Foundation under Grant #1019343 to the Computing Research Association for the CIFellows Project.

References

[1] J. Aggarwal and Q. Cai. Human motion analysis: A review. *CVIU*, 1999. I

[2] J. Aggarwal and M. Ryoo. Human activity analysis: A review. *ACM Computing Surveys*, 2012. I

[3] S. Ali and M. Shah. A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis. In *CVPR*, pages 1–6. Ieee, June 2007. III, III

[4] S. Ali and M. Shah. Floor fields for tracking in high density crowd scenes. In *ECCV*, pages 1–14, 2008. III, III, 2

[5] C. E. . N. L. Bouvier, E. From crowd simulation to airbag deployment: particle systems, a new paradigm of simulation. *Journal of Electronic Imaging*, 1997. IV

[6] A. Braun, S. Musse, L. de Oliveira, and B. Bodmann. Modeling individual behaviors in crowd simulation. *Proceedings 11th IEEE International Workshop on Program Comprehension*, (6):143–148, 2003. IV

[7] D. Brogan and J. Hodgins. Group behaviors for systems with significant dynamics. *Autonomous Robots*, 153, 1997. IV, 3

[8] D. C. Brogan, R. A. Metoyer, and J. K. Hodgins. Dynamically simulated characters in virtual environments. *Computer Graphics and Applications*, 1998. IV

[9] D. Helbing, I. Farkas, and T. Vicsek. Simulating dynamical features of escape panic. *Nature*, pages 1–16, 2000. IV

[10] M. Hu and S. Ali. Detecting global motion patterns in complex videos. In *ICPR*, pages 1–5. Ieee, Dec. 2008. III, III

[11] J. C. S. Jacques, S. R. Musse, and C. R. Jung. Crowd analysis using computer vision techniques. *Signal Processing*, (September):66–77, 2010. III

[12] J. C. S. Jacques, Jr., A. Braun, J. Soldera, S. R. Musse, and C. R. Jung. Understanding people motion in video sequences using voronoi diagrams: Detecting and classifying groups. *Pattern Anal. Appl.*, 10(4):321–332, Oct. 2007. II

[13] T. Lan, Y. Wang, G. Mori, and S. N. Robinovitch. Retrieving Actions in Group Contexts. In *International Workshop on Sign Gesture Activity*, 2010. II

[14] T. Lan, Y. Wang, W. Yang, S. N. Robinovitch, and G. Mori. Discriminative Latent Models for Recognizing Contextual Group Activities. *IEEE transactions on pattern analysis and machine intelligence*, 1(1):1–14, Nov. 2011. II

[15] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos. Anomaly Detection in Crowded Scenes. In *CVPR*, 2010. III

[16] B. E. Moore, S. Ali, R. Mehran, and M. Shah. Visual crowd surveillance through a hydrodynamics lens. *Communications of the ACM*, 54:64, 2011. III, III

[17] S. Musse and D. Thalmann. A model of human crowd behavior: Group inter-relationship and collision detection analysis. In *Computer Animation and Simulation*, volume 97, pages 39–51. Citeseer, 1997. II, IV

[18] N. M. Nayak, R. J. Sethi, B. Song, and A. K. Roy-Chowdhury. Motion Pattern Analysis for Event and Behavior Recognition. In T. B. Moeslund, L. Sigal, V. Kruger, and A. Hilton, editors, *Visual Analysis of Humans*, pages 289–309. Springer-Verlag, 2011. I

[19] N. Pelechano, J. M. Allbeck, and N. I. Badler. *Virtual Crowds: Methods, Simulation, and Control*, volume 3. Jan. 2008. IV

[20] Y. Rao, L. Chen, Q. Liu, W. Lin, Y. Li, and J. Zhou. Real-time control of individual agents for crowd simulation. *Multimedia Tools and Applications*, 54(2):397–414, May 2010. IV

[21] C. W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 1987. IV

[22] M. Rodriguez, S. Ali, and T. Kanade. Tracking in unstructured crowded scenes. In *ICCV*, number September, 2009. III, III

[23] R. J. Sethi and A. K. Roy-Chowdhury. Modeling and Recognition of Complex Multi-Person Interactions in Video. In *ACM MM MPVA*, pages 0–3, 2010. II, 1

[24] R. J. Sethi and A. K. Roy-Chowdhury. Physics-based Activity Modelling in Phase Space. In *ICVGIP*, 2010. II

[25] R. J. Sethi, A. K. Roy-Chowdhury, and S. Ali. Activity Recognition by Integrating the Physics of Motion with a Neuromorphic Model of Perception. In *WMVC*. WMVC, 2009. II

[26] N. Shroff, P. Turaga, and R. Chellappa. Moving vistas: Exploiting motion for describing scenes. CVPR, 2010. I

[27] B. Song, R. J. Sethi, and A. K. Roy-Chowdhury. Robust Wide Area Tracking in Single and Multiple Views. In T. B. Moeslund, L. Sigal, V. Kruger, and A. Hilton, editors, *Visual Analysis of Humans*, pages 1–18. Springer-Verlag, 2011. I

[28] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea. Machine recognition of human activities: A survey. *CSVT*, 2008. I

[29] S. Wu, B. Moore, and M. Shah. Chaotic Invariants of Lagrangian Particle Trajectories for Anomaly Detection in Crowded Scenes. In *CVPR*, 2010. III, III

[30] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg. Who are you with and where are you going? In *CVPR*, pages 1345–1352. Ieee, June 2011. II

[31] D. Zhang, D. Gatica-Perez, and S. Bengio. Modeling individual and group actions in meetings with layered HMMs. *IEEE Transactions on Multimedia*, 8(3):509–520, 2006. II