The International Series in Video Computing

Saad Ali · Ko Nishino
Dinesh Manocha
Mubarak Shah *Editors*

Modeling, Simulation and Visual Analysis of Crowds

A Multidisciplinary Perspective

Springer

The International Series in Video Computing

Series Editor: Mubarak Shah, Ph.D
University of Central Florida
Orlando, Florida

For further volumes:
http://www.springer.com/series/6673

Saad Ali • Ko Nishino • Dinesh Manocha
Mubarak Shah
Editors

Modeling, Simulation and Visual Analysis of Crowds

A Multidisciplinary Perspective

Springer

Editors
Saad Ali
Center for Vision Technologies
SRI International
Princeton, NJ, USA

Ko Nishino
Department of Computer Science
Drexel University
Philadelphia, PA, USA

Dinesh Manocha
Department of Computer Science
University of North Carolina
Chapel Hill, NC, USA

Mubarak Shah
Center for Research in Computer Vision
University of Central Florida
Orlando, FL, USA

ISSN 1571-5205
ISBN 978-1-4614-8482-0 ISBN 978-1-4614-8483-7 (eBook)
DOI 10.1007/978-1-4614-8483-7
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013952951

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Accurate analysis and synthesis of human behavior in crowds, a large and dense group of people with varying characteristics and goals, is a common requirement across a wide range of domains. If the human behavior, including those of individuals, small groups of people, and even the crowd as a whole – can be interpreted and anticipated in arbitrary real-world situations, a repertoire of important applications, many of which are societally important, can be realized: For example, perpetrators disguised in a busy street corner will be easily spotted and tracked in a surveillance video feed; new buildings, public places and outdoor environments will be designed to optimize the space use with the dynamically changing flow of people in mind, while minimizing the time need for evacuation whenever necessary; and the social psychology of people can be studied based on large-scale, longitudinal observations, and many more.

The goal of this book is to provide the readers a comprehensive map of the current state of the art in distinct but related fields, mainly in computer vision, graphics, and evacuation dynamics, towards the common goal of better analyzing and synthesizing the pedestrian movement in dense, heterogeneous crowds. The monograph is organized into different parts that consolidate various aspects of research towards this common goal, namely the modeling, simulation, and visual analysis of crowds. Many of the chapters in these parts extend the works that were presented at the first workshop on the same topic at International Conference on Computer Vision, 2011, and collectively cover the diverse challenges involved in better understanding of human crowds. Our hope is, through this book, the readers will see the common ideas and vision as well as the different challenges and techniques for modeling, analyzing, and simulating crowds, that will stimulate novel approaches to getting us a step closer to fully grasping "crowds."

This book grew out of the first IEEE Workshop on Modeling, Simulation and Visual Analysis of Large Crowds, that was held in conjunction with International Conference of Computer Vision 2011. Therefore, first of all we would like to acknowledge the workshop program committee who worked tirelessly for the success of the workshop and authors that contributed their valuable pieces of work. We would also like to thank Prof. Jie Yang and National Science Foundation (NSF) funding based on grant IIS-1142382 to provide travel support for the workshop. We are also grateful to our host institutions (SRI International, Drexel University, University of North Carolina and University of Central Florida) for providing a highly stimulating research environment that enables pursuit of new research ideas and discoveries. Springer has provided excellent support throughout the preparation of the book, and we would like to specially thank their staff for their support and professionalism. Many people have helped proof reading draft material and providing comments and suggestions. We would like to thank all of them for their time and valuable contribution towards improving the quality of the book.

Princeton, NJ, USA Saad Ali
Philadelphia, PA, USA Ko Nishino
Chapel Hill, NC, USA Dinesh Manocha
Orlando, FL, USA Mubarak Shah

Contents

Part II Visual Analysis of Crowds

Contributors

Saad Ali Center for Vision Technologies, SRI International, 201 Washington Road, Princeton, NJ, USA

Norman I. Badler University of Pennsylvania, Philadelphia, USA

Stefania Bandini Department of Computer Science, Systems and Communication, Complex Systems and Artificial Intelligence (CSAI) Research Center, University of Milan – Bicocca, Milano, Italy

Maik Boltes Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, Jülich, Germany

Avishy Y. Carmi Department of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore

Antoni B. Chan Department of Computer Science, City University of Hong Kong, Hong Kong, China

Ke Chen Queen Mary University of London, London, UK

M. Chraibi Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany

M. Cristani Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia, Genova, Italy

Sean Curtis University of North Carolina at Chapel Hill, Chapel Hill, USA

A. Del Bue Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia, Genova, Italy

Funda Durupinar University of Pennsylvania, Philadelphia, USA

Simon J. Godsill Department of Engineering, University of Cambridge, Cambridge, UK

Shaogang Gong Queen Mary University of London, London, UK

Helena Grillon Centrale de Compensation, Geneva, Switzerland

Pini Gurfil Department of Aerospace Engineering, Technion Israel Institute of Technology, Kesalsaba, Israel

Stephen J. Guy University of North Carolina at Chapel Hill, Chapel Hill, USA

Mubbasir Kapadia Center for Human Modeling and Simulation, University of Pennsylvania, Philadelphia, USA

Louis Kratz Department of Computer Science, Drexel University, Philadelphia, USA

Laura Leal-Taixé Leibniz University Hannover, Hannover, Germany

Chen Change Loy Vision Semantics Limited, The Chinese University of Hong Kong, Hong Kong, China

Jonathan Maim Bengaluru Area, India

Dinesh Manocha Department of Computer Science, University of North Carolina, Chapel Hill, USA

Lyudmila Mihaylova School of Computing and Communications, Lancaster University, Lancaster, UK

V. Murino Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia, 16163 Genova, Italy

Ko Nishino Department of Computer Science, Drexel University, Philadelphia, USA

Sze Kim Pang DSO National Laboratories, Singapore

Jan-Frederik Pietschmann Center for Industrial and Applied Mathematics (CIAM), Royal Institute of Technology, Stockholm, Sweden

R. Raghavendra Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia, Genova, Italy

Bodo Rosenhahn Leibniz University Hannover, Hannover, Germany

E. Sangineto Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia, Genova, Italy

A. Schadschneider Institute for Theoretical Physics, Universität zu Köln, Köln, Germany

François Septier Signal Processing and Information Theory Group, TELECOM Lille 1, Villeneuve d'Ascq Cedex, France

Armin Seyfried Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich GmbH, Germany

Computer Simulation for Fire Safety and Pedestrian Traffic, Bergische Universität Wuppertal, Wuppertal, Germany

Mubarak Shah Center for Research in Computer Vision, Harris Corporation Engineering Center, University of Central Florida, Orlando, USA

Alexander Shoulson University of Pennsylvania, Philadelphia, USA

Daniel Thalmann Institute for Media Innovation, Nanyang Technological University, Singapore

Nuno Vasconcelos Department of Electrical and Computer Engineering, University of California, San Diego, USA

Giuseppe Vizzari Department of Computer Science, Systems and Communication, Complex Systems and Artificial Intelligence (CSAI) Research Center, University of Milan – Bicocca, Milano, Italy

Tao Xiang Queen Mary University of London, London, UK

Barbara Yersin Bengaluru Area, India

Basim Zafar Hajj Research Institute, Umm al-Qura University, Mecca, Saudi Arabia

Jun Zhang Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, Jülich, Germany

Chapter 1
Modeling, Simulation and Visual Analysis of Crowds: A Multidisciplinary Perspective

Saad Ali, Ko Nishino, Dinesh Manocha, and Mubarak Shah

Abstract Over the last several years there has been a growing interest in developing computational methodologies for modeling and analyzing movements and behaviors of 'crowds' of people. This interest spans several scientific areas that includes Computer Vision, Computer Graphics, and Pedestrian Evacuation Dynamics. Despite the fact that these different scientific fields are trying to model the same physical entity (i.e. crowd of people), research ideas have evolved independently. As a result each discipline has developed techniques and perspectives that are characteristically it's own. In this chapter we provide a brief overview of major research themes from these different scientific fields, discuss common challenges and point to problem areas that will benefit from common synthesis of perspectives from these fields. In addition we introduce various pieces of work that appear in this monograph as separate chapters.

S. Ali (✉)
Center for Vision Technologies, SRI International, 201 Washington Road, Princeton, NJ, USA
e-mail: saad.ali@sri.com

K. Nishino
Department of Computer Science, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA
e-mail: kon@drexel.edu

D. Manocha
University of North Carolina, Chapel Hill, NC, USA
e-mail: dm@cs.unc.edu

M. Shah
University of Central Florida, Orlando, FL, USA
e-mail: shah@crcv.ucf.edu

S. Ali et al. (eds.), *Modeling, Simulation and Visual Analysis of Crowds*, The International Series in Video Computing 11, DOI 10.1007/978-1-4614-8483-7_1,

1.1 Introduction

Over the last several years there has been a growing interest in developing computational methodologies for modeling and analyzing movements and behaviors of 'crowds' of people. This interest spans several scientific areas: in *Computer Vision* the need to carry out visual surveillance in crowded scenes is fueling research on topics related to visual representations of crowds [2], tracking of individuals and groups [3, 18, 32, 45], detection of normal and abnormal behaviors [44, 62], segmentation and classification of motion patterns [2, 63], and mathematical modeling of interactions among the pedestrians in the crowd [81]; in *Computer Graphics* the goal of modeling and simulating crowd behaviors in different real-world or synthetic environments, including models for homogeneous and heterogeneous crowd simulation [30, 66], is advancing the state of art in areas aggregate flow [30, 31, 91], agent-based motion simulation [49, 84, 96], motion planning for large scale crowds [36, 70, 71, 94], obstacle and collision avoidance [52, 82, 92], modeling group behaviors [57, 65], and representation of virtual humans at multiple levels of detail [16, 60, 61]; in *Evacuation Dynamics* a parallel effort is underway to develop motion, interaction and self-organization models for pedestrian simulation and evacuation analysis [24]. However, as opposed to computer graphics, the emphasis is more on empirical validation of simulated movements and collective behaviors in terms of fundamental diagrams (which captures relationship between crowd density and velocity) and flows [11, 82]. In addition to above mentioned scientific areas, the diversity of context in which 'crowds' are studied has a long history and includes studies from areas of Anthropology, Psychology, and Sociology [9, 53].

Despite the fact that these different scientific fields are trying to model the same physical entity (i.e. crowds of people), many research ideas have evolved independently, and as a result each discipline has developed techniques and perspectives that are characteristically it's own. However, we strongly believe that in order to make the next big leap in terms of solving the crowd modeling and related computational problems, there is a need to develop common insights and understanding of general principles that characterize various aspects of a crowd. This requires development of a common-platform for cross-disciplinary exchange of ideas and interaction that allows benefiting from each other's experience and scientific discoveries. Some of the recent research in computer vision [3, 62, 69, 99] points towards merits of such cross-disciplinary work, where pedestrian interaction pedestrian interaction models, originally developed in evacuation dynamics, have been successfully used to carry out visual tracking and abnormal behavior estimation. Similarly, recent research in data-driven crowd simulation in computer graphics makes use of crowd trajectories and behaviors that are extracted from videos using computer vision algorithms [52, 67, 72].

The central goal of this monograph is to facilitate a process of cross-disciplinary interaction among researchers from areas of compute vision, computer graphics and evacuation dynamics by providing a common platform. For this purpose, a number of peer-reviewed chapters from leading researchers in these fields are compiled.

These chapters provide an understanding of the state of the art and open problems related to crowd modeling in each scientific discipline.

The rest of the chapter is organized as follows: In Sect. 1.2 we discuss various aspects of crowd which make their modeling a challenging task. In Sect. 1.3 we introduce central themes of the book and provide an overview of the related literature. In Sect. 1.4 we provide an overview of the organization of the rest of the book.

1.2 Aspects of Crowds

In order to bring about a common understanding of concepts and approaches related to crowds of people, it is important to answer the question: how should we think about crowds? What are the particular characteristic of crowds which make its modeling a challenging task? As to the former, in the most basic sense, a crowd is any collection of individuals or pedestrians where behavior of one individual is influenced by the other. We believe the flexible nature of this definition makes it applicable to all scientific areas that are focus of this monograph. For instance, we have examples of computer vision techniques which represent this influence at a notional level through particle interactions [2] or through dynamic floor fields [3]. Similarly, in computer graphics and simulated environment, this influence is taken into consideration during design of algorithms for collision avoidance and local interaction [20, 92]. Finally, in evacuation dynamics, computational approaches for emergent behaviors tend to use force-based models and cellular automata [82]. Various chapters of the monograph will provide many more example of how influence among participants of crowds are represented in different settings.

Agreeing to this definition leads us to the next question. What are the particular aspects of crowds which make them really challenging to model. We list some of them next in no particular order:

- Human behavior is extremely complex and exhibit large variation based on situations and settings. It also depends on individual characteristics such as age, sex, height, and cultural background, to name the few. There is no existing mathematical model that can account of all these complexities in human behavior.
- Human behavior can vary drastically based on the given situation. For instance, transition from walking to panic can be instantaneous given a dangerous situation (e.g. stampede).
- When viewed from visual sensors, it is hard to discern individuals in dense crowds due to low resolution (i.e. few pixels per individual). This results in appearance ambiguity and severe occlusion which often results in breakdown of visual processing pipelines. For instance, detection of individual person in the crowd might not be possible.
- Inability to detect individuals in a crowds that are observed through visual sensors makes it difficult to explicitly model interactions among individuals.

- Crowds tend to be heterogeneous in nature and within the same scene some parts of the crowd may be behaving very differently. This makes it hard to represent the dynamics of the crowd using a single global model.
- Behavior and dynamics of an individual in a crowd are connected with other individuals, both at the level of structure as well as behavior. This means actions of individuals can not be modeled in isolation and the fact that crowd will react to it has to be taken into consideration.
- Lack of datasets representing the richness of crowd behaviors and associated ground truth (e.g. 3D scene layout, individual tracks, personal characteristics such as height, sex etc.) makes it difficult to verify and validate crowd modeling techniques.

1.3 Central Themes and Topics

With this set of ideas in mind, we now introduce some of the main themes and topics considered in this monograph and the ways in which they reinforce the underlying principles related to modeling crowds. We also provide pointers to the chapters that are related to each of these themes.

1.3.1 Visual Analysis of Crowds

Visual analysis of crowded scenes is an integral component of a wide array of applications that span a number of areas with direct social impact. For instance, the rising prevalence of video recording technology in crowded areas presents a dire need for automatic visual analysis that can operate on videos containing a large numbers of individuals. Due to a large number of pedestrians in close proximity, crowded areas are at a high risk for dangerous activities including crowd panic, stampedes, and accidents involving a large number of individuals. Crowded scenes, are one such scenario of high-density, cluttered scenes that contain a large number of individuals. The extent of activity within such scenes is difficult for even human observers to analyze, making crowded scenes perhaps in the most need of automatic video analysis. Analyzing the behavior of pedestrians in such crowded scenes is also essential to the understanding and prediction of human behavior in similar but different scene context. Video analysis of crowded scenes can thus directly serve as the means to obtain in situ measurements of human behavior for data-driven crowd simulation.

Despite these strong needs, crowded scenes pose unique challenges that severely impede the development of robust video analysis methods. The complexity of the scene, largely owing to the sheer number of people in the crowds, becomes a direct burden on the computational method for visually analyzing the scene. Such complexity manifests itself in frequent, partial or complete occlusions among the pedestrians; the fact that every individual is moving and also surrounded by other moving people blurs the boundary of foreground and background pixels in the scene; and the arbitrary directions pedestrians may take based on their personal goals, neighboring pedestrians, and the physical obstacles within the scene. These all combined result in a heterogeneous and dynamically evolving crowd motion that is often too complex to analyze with conventional computer vision methods.

Conventional video analysis methods learn the behavior of the scene in three steps: detecting objects, tracking objects, and compiling the tracked results into higher order models for individual or global crowd behavior modeling. The applicability of such object-centric methods is limited to scenes with relatively few objects. Discerning individuals in crowded scenes is difficult since they are typically surrounded by other moving pedestrians. Tracking is also difficult due to the frequent partial or complete occlusions in crowded scenes. Finally, such methods suffer from problems of scale: each new pedestrian that enters the scene increases the complexity of the model.

Next we briefly describe some popular approaches to modeling crowds and their behaviors in videos.

1.3.1.1 Object-Centric Visual Analysis

Conventional video analysis methods are mostly object-centric; they begin by analyzing each scene object. Such methods detect the scene objects (e.g., pedestrians or automobiles), track them, and then analyze the trajectories to model the behavior of the objects. These methods work well on scenes that are relatively sparse (roughly 5–20 pedestrians) and, as noted by Zhan et al. [102], are not appropriate for dense crowded scenes. We review related object-centric work that are designed for crowded scenes, but emphasize that there are many challenges in videos containing high density crowds.

Detecting the scene objects is often the first step in object-centric video analysis. Zhao et al. [103], for example, track pedestrians in videos of crowds by detecting each individual using a model of human shapes. Rodriguez and Shah [76] detect pedestrians using a voting scheme on the contours around each individual. The contours are computed by subtracting the background from each video frame. In high density crowded scenes, however, the background is rarely visible and pedestrians are often partially occluded, making the contours difficult to estimate. Leibe et al. [51] also segment pedestrians from the background, but use global image cues to add robustness to partial occlusions. Their method handles some partial occlusions well, but assumes that the torso of the pedestrian is visible. This is often not the case in near-view scenes where only the heads of most pedestrians are visible.

Other work detect pedestrians by assuming that they exhibit unique motion. Brostow and Cipolla [8] group short feature tracks (or “tracklets”) to identify similarly moving pedestrians. They assume that the subjects move in distinct directions and thus disregard possible local motion inconsistencies between different body parts.

As noted by Sugimura et al. [90], such inconsistencies cause a single pedestrian to be detected as multiple targets. In addition, pedestrians that are moving in the same direction are identified as a single group. Crowded scenes, especially when captured in relatively near-field views, as is often the case in video surveillance, necessitate a method that represents the multiple motions of a single individual or similar motions of different pedestrians.

After detection, the objects are tracked as they move through the scene. Data association methods, such as that of Betke et al. [7] or Gilbert and Bowden [19], track multiple targets in cluttered scenes by associating detection results of consecutive frames. These techniques assume that the detection is always reliable, and thus degrade in very crowded scenes. Wu and Nevatia [97] are able to track partially occluded pedestrians by detecting body parts, rather than the full pedestrian. The data association problem itself is NP-hard, and thus becomes less tractable in scenes with a large number of pedestrians. Often, approximation techniques are used to estimate a solution such as the Bayesian framework of Li et al. [54].

Other data association methods do not rely on detecting individuals. Khan et al. [40] model the interaction among detected interest points to improve the tracking of each object. Hue et al. [29] use a Markovian model on each tracked point to augment data association in generic domains. As noted by Khan et al. [41], however, a single point may be shared between multiple targets and can result in ambiguities. Shared points are often the result of motion boundaries or clutter, both of which occur frequently in videos of crowded scenes.

After tracking, the trajectories are used to characterize behaviors of objects within the scene. Wang et al. [95], for example, cluster trajectories to learn the common routes taken by pedestrians and automobiles. Dee and Hogg [14] use the tracking information to identify pedestrians that deviate from a goal-specific behavior. Hu et al. [26] learn global motion patterns (i.e., that describe motion over the entire frame) and use them to detect anomalies and predict future behaviors. Johnson and Hogg [34] estimate different distributions of trajectories, and attach semantics to each in order to identify specific events within the scene. Such methods not only depend on reliable detection and tracking, which may not be available in videos of crowded scenes, but also face problems of scale, in terms of handling large, crowded scenes. As more pedestrians enter the scene, the complexity of these methods increases and may become intractable with even moderately dense crowds.

1.3.1.2 Crowd Motion Patterns

To address the complexity of real-world scenes containing crowds, many researchers propose holistic techniques that characterize the scene as a collection of local motion estimates rather than a collection of objects. Often, holistic methods aim to identify behaviors within the scene that are part of the same physical process [27]. Mahadevan et al. [58] describe the typical dynamics of the crowd with a mixture of dynamic textures (previously used for segmentation by Chan and Vasconcelos [10]).

Using dynamic textures, however, retains appearance variations which can introduce noise into the model and degrade results. This approach is further elaborated in Chap. 11.

Moore et al. [64], Mehran et al. [63], and Ali and Shah [3] model crowds based on a hydrodynamics model that essentially treats each pedestrian as a particle in a fluid. As noted by Still [88], however, specific behaviors that occur in crowds, such as lane formations or clustering, do not occur in fluids. While particles are affected only by the external forces around them (such as other particles or the environment), the motion of pedestrians is a result of both external forces and their individual desires. Such differences between individual pedestrians form dynamic space-time structures in the crowd motion that can not be represented with a hydrodynamics model.

Other work assumes that the crowd flow is constant over the entire video. Ali and Shah [2] average the optical flow over a video clip, and use it to model a Finite Time Lyapunov Exponent field for segmenting the motion of the crowd. Similarly, Mehran et al. [62] measure the "social force" [25] by comparing the instantaneous optical flow to the optical flow averaged over the video clip. Raghavendra et al. [73] also estimate the social force, but do so using a particle swarm method that clusters similar motion vectors. In many crowded scenes, especially those with unconstrained environments, the motion of pedestrians can change dramatically in a short period of time as individuals move towards different goals.

Some researchers assume the crowd exhibits homogeneous motion in each area of the scene. Hu et al. [28], for example, identify global motion patterns (i.e., ones that take up the entire frame) in crowded scenes by clustering optical flow vectors in similar spatial regions. Similarly, Cheriyadat and Radke [13] detect dominant motions in crowds by clustering low-level tracked features. Such methods can not handle dynamically varying crowds or those with heterogeneous motions in local areas. A crosswalk, for example, naturally has pedestrians moving in two directions who emerge together as they pass each other.

Other methods capture the multi-modal nature of the crowd, but ignore the important temporal relationship between sequentially occurring motions exhibited by pedestrians. Rodriguez et al. [77] use a topical model (similar to the bag-of-word models) over quantized optical flow directions to describe the crowd motion. They later improve their tracking using a crowd density estimate [78]. Though they model the heterogeneous nature of the crowd, they does not encode the relationship between temporally co-occurring motions. By disregarding the temporal variations in the motions exhibited by pedestrians, these approaches cannot represent the underlying temporal pattern within the crowd motion.

Andrade et al. [4, 5] captures the temporal structure of the crowd by training hidden Markov models on optical flow vectors. They demonstrate that their method is a good indicator of emergency situations in simulated crowd flow data. Real-world crowded scenes, however, were not evaluated. Kratz and Nishino [44] modeled the local motion patterns with the distribution of the spatio-temporal gradients and derived a distribution-based hidden Markov model to encode their spatio-temporal variations. They successfully demonstrated the use of this crowd motion model for anomaly detection [44] and tracking of individuals [45, 47] in very crowded scenes. They further extended this model with directional statistics distributions to more faithfully encode the local motion patterns and introduced the use of pedestrian efficiency for modeling the magnitude of deviation of a person from the crowd motion to better detect and track unusual activities [46]. Saleemi et al. [79] developed a statistical representation of motion patterns of pedestrians in a scene observed by a static camera. Motion patterns are learned in an unsupervised manner directly from the salient patterns of optical using mixture model representation.

1.3.1.3 Particle-Based Representation of Crowds

To overcome the difficulty of detection individual objects in dense crowded scene, a particle based representation of crowd motion has been used by many researchers. Ali and Shah [2] introduced this representation where crowd motion is represented in terms of trajectories of a dense grid on particles. They used this representation to carry out segmentation dynamically distinct crowd flows (further elaborated in Chap. 8). In following years, particle based representation has been successfully used for tracking individual pedestrians in crowds [3], abnormal event detection [62, 98], and semantic scene understanding [55]. The work of Moore et al. [64] provides an expanded explanation of the intuition behind the particle based representation and its applicability to various problem areas.

1.3.1.4 Abnormal Event Detection

Various approaches have been proposed to perform abnormal event detection in dense crowded scenes. These can be characterized based on whether abnormality is detected locally or globally in the scene. The local abnormality detection methods mostly employ local 2D or 3D motion or appearance features with some added information to capture the local context information (e.g. using co-occurrence matrices). For instance, Adam et al. [1] measure the probability of optical flow in a local patch using histograms for detecting abnormal patterns. Kratz and Nishino [44] fit a Gaussian model to spatio-temporal gradients and then use HMM to detect abnormal events. Kim and Grauman [42] model local optical flow and enforce the consistency using Markov Random Field for detection of abnormal motion patterns. Mahadevan et al. [58] model the normal crowd behavior by mixtures of dynamic textures.

The global abnormality detection approaches label the motion in the entire scene as abnormal. This often happens in cases of panic situations such as stampede. In this direction, Mehran et al. [62] proposed a formulation of abnormal crowd behavior by adopting the social force model and then using Latent Dirichlet Allocation (LDA) to detect abnormality. In [98], chaotic invariants of particle trajectories are used for detecting abnormal motion.

Next we describe related approaches from area of crowd simulation and pedestrian evacuation dynamics.

1.3.2 Crowd Simulation and Behavior Modeling

Techniques to simulate and understand crowd behaviors and motion have been studied in computer graphics, virtual reality, social science, statistical physics, robotics, pedestrian and evacuation dynamics, and other areas of science and engineering. One of the major goals is to develop appropriate models that can be used to simulate and predict crowd behaviors in different real-world or synthetic environments. A key computation within these models at the *microscopic* level is to simulate the trajectory of each agent that avoids the static or dynamic obstacles and other human agents in the environment. The microscopic models also take into account the psychological and social behavior of each individual and how they respond to external events or stress. At the same time, we are also interested in *macroscopic* techniques that focus on the tendency and movement pattern of the entire crowd or the aggregate flow. These models seek to predict plausible or likely motion of human-like agents in terms of paths as well as other characteristics such as densities, speeds, and emerging behaviors. Overall, the design and formulation of such crowd simulation models is a challenging and multi-faceted problem. The range of trajectories that the humans follow, the pattern of crowd behaviors that we observe, and the variety of situations which the humans encounter are almost endless. As a result, crowd simulation remains an active area of research in many disciplines.

Any model for crowd simulation needs to take into account several components. This includes specifying a computer-based geometric model of the environment; computing the movement and trajectories of various agents, taking into account the interactions amongst the agents and with the environment; model external and dynamic events that affect crowd behavior. Some of the widely used models are based on multi-agent simulation. A crowd is composed of human-like agents (i.e., individuals in the crowds), with a collection of goals and a set of obstacles that constitute the environment. The individuals constituting a crowd may have similar or distinct goals. In heterogeneous crowd simulation, each individual in the crowd is assumed to have a physically embodied goal. The representation of this goal can vary based on the simulation scenario; for example, a goal may correspond to a specific position or a certain region in the environment. These goals may be dynamic and may change over time, say following some other individual or dynamic obstacle in the crowd. In addition to specifying the region and goals for each agent, the model must take into account the environment, which consists of walls, obstacles or other regions that may not be accessible to human-agents. These obstacles may be static (e.g. buildings) or dynamic (e.g. moving vehicles). Given the description of the agents and the environment, a key component of crowd simulation is the computation of trajectories for each agent that adhere to environmental factors, avoid collisions with the obstacles and other agents, and guide each agent towards its immediate goal. In particular, Reynolds refers to this process of intermediate-level planning as *steering behaviors* [75]. These steering behaviors are largely independent of the particulars of the agent's means of locomotion, but are used to navigate around the environment in a life-like and improvisational manner, and also results in collision avoidance. The combinations of such steering behaviors can be used to achieve some higher level goals, like following a walkay or a corridor, joining some other group of agents, etc.

1.3.2.1 Models for Crowd Motion

There is extensive literature on simulating crowd movements and dynamics. Many techniques have been proposed to compute the motion of individuals in crowds. Their application depends on the type of crowd patterns or behaviors that we want to simulate and the surrounding environment. An important issue with respect to crowd simulation is whether the crowds being simulated are homogeneous or heterogeneous. Homogeneous crowds correspond to instances where each agent has very similar behavior or goal. The study of heterogeneous crowds dates back to at least the work of Le Bon more than a century ago, who analyzed how members of a crowd can have different races, genders, intents, and backgrounds [50]. In heterogeneous crowds, every individual agent in the crowd maintains a distinct, observable identity. This identity is observed in the goals, desired speeds, aggressiveness, cooperation, and many other factors which affect the motion and trajectory of each agent. In contrast, homogeneous crowds are observed when a clear unity of action leads to a "disappearance of conscious personality", which results in a homogeneity of motion [50]. In terms of simulating homogeneous crowds, it may be possible to exploit the coherence in individual motion to accelerate the overall simulation. These include flow-based models [30, 31] that are governed by differential equations that uniformly dictate the flow of crowds across space. Other examples include models based on continuum crowds [91], which allow a small, fixed number of goals, and aggregate dynamics for dense crowd simulation [66].

1.3.2.2 Agent-Based Crowd Simulations

In contrast to continuum methods, agent-based simulation methods allow for true heterogeneity in simulating the motion and trajectory of each individual. In these simulations, each human-like character in the crowd is represented as a simulated agent. Since the motion of each agent is computed separately, it is possible to simulate crowds with varying characteristics and personalities for each agent. One of the most popular agent-based approaches was proposed by Reynolds in the Boids algorithm [74], which can generate steering behaviors that resemble flocking, herding, and school behaviors commonly observed in animal motion. This algorithm has been widely used in games and generating special effects in movies.

There is considerable literature in robotics and related areas on computing the motion and trajectories of multiple agents in a shared environment. The underlying motion-planning problem is known to have exponential complexity in terms of number of agents or the degrees of freedom [49]. At a broad level, prior work on motion planning can be classified into two kinds of approaches. The *centralized* approaches [48,49] consider the sum of all the robots or agents and treat the resulting system as a single composite system. In these methods, the configuration spaces of individual robots are combined (using the Cartesian product) in a composite space, and the resulting algorithm searches for a solution in this combined configuration space. In contrast, the *decoupled* planners proceed in a distributed manner, and coordination among them is often handled by exploring a *coordination space*, which represents the parameters along each specific path or are computed some kind of local rules. Decoupled approaches [84,96] are much faster than centralized methods, but may not be able to guarantee theoretical completeness. Some of the techniques from robot motion planning have been used to generate group behaviors [6, 37] and real-time navigation of large numbers of agents amongst obstacles [17, 89].

Most widely-used techniques for handling a large number of human-like agents are based on decentralized methods. This is a challenging task, particularly in densely-packed, crowded scenarios with several hundreds or thousands of agents. Each agent essentially has to navigate through an unknown dynamic environment; it has no prior knowledge of how other agents or the dynamic obstacles will move in the future. The standard approach to this class of problems is to let each agent run a continuous cycle of sensing and acting. During each cycle, the agent observes its surroundings, acquires information about the positions and velocities of other agents and obstacles in the synthetic environment, and computes a local path towards a goal that avoids collisions. If this cycle is executed at a high frequency, the agent is able to react in a timely way to changes in its surroundings. The computation of an agent's motion breaks down into two tasks: global and local navigation. Global navigation aims at computing a collision-free path towards a goal position that only takes into account the static obstacles, while local navigation techniques are used to avoid collisions with other agents and dynamic obstacles and steer each agent towards its goal position.

1.3.2.3 Global Navigation

Global path computation is typically performed using a global data structures, such as roadmaps or navigation meshes. A roadmap is a graph consisting of a set of vertices positioned in freespace (i.e. not inside an obstacle) and a set of edges connecting these vertices. Two vertices are connected by an edge if and only if the direct path between the two nodes is collision-free (i.e. no obstacles block the direct path). Such roadmaps can be constructed by an artist, or can be automatically generated using visibility graphs [56], probabilistic methods [39], or other techniques [48, 49]. Each agent can compute its global path using such roadmaps and performing graph search, such as Dijkstra's algorithm [15] or A* search [23]. In games and related applications, navigation meshes [36, 86, 94] have begun to supplant roadmaps. A navigation mesh is a decomposition of the freespace of the environment into a mesh consisting of convex polygons. As in roadmaps, the connectivity of the mesh is stored as a graph; however, navigation meshes have advantages over roadmaps in that all edges of a polygon are implicitly connected to each other, i.e. because of the convexity there is a straight-line path from any point in the polygon to any boundary. In addition, a single navigation mesh can encode clearance for arbitrarily sized agents. Computing a global path with a navigation mesh simply requires searching the connectivity graph for the shortest path between two nodes. The cost of a graph edge between two polygons depends on the length of the shared edge of those two polygons. If the edge is not large enough to accommodate the agent, the cost is infinite.

1.3.2.4 Local Navigation and Crowd Simulation

Several techniques have been proposed to animate or simulate large groups of autonomous agents or crowds. Most of these methods use a rather simple representation for each agent – for example, a circular shape in a 2D plane or a cylindrical object in the 3D space – and compute a collision-free trajectory for each agent. After computing the trajectory using a simple representation, these techniques use either footstep planning or walking synthesis methods to compute a human-like motion for each agent along the given trajectory.

Local navigation computation takes into account the motion of dynamic obstacles and other agents in the environment. At a broad level, prior methods for local collision avoidance and navigation can be classified into the following categories:

- *Potential-based methods:* These algorithms focus on modeling agents as particles with potentials and forces [24].
- *Boid-like methods:* These approaches, based on the seminal work of Reynolds, create simple rules for computing the velocities [74, 75].
- *Geometric and velocity methods:* These algorithms compute collision-free paths using sampling in the velocity space obstacles [92] or by using optimization methods [20, 85, 93].
- *Field based methods:* These algorithms compute fields for agents to follow [12, 33, 72, 100], or generate navigation fields for different agents [67].
- *Least effort crowds:* These methods for modeling the paths of crowds are based on the classic principle of Least Effort proposed by Zipf [104], many researchers have used that formulation to model the paths of crowds [35, 38, 80, 87]. Recently, it has been combined with multi-agent collision avoidance algorithms [93] and used to efficiently and automatically generate emergent behaviors for a large number of agents [21].
- *Data-driven methods:* These methods use real-world or data-driven techniques to simulate realistic crowd simulation as well as evaluate their accuracy [52, 72, 82].

In addition to these broad classifications, there are many other specific approaches designed to simulate crowd behavior based on cognitive modeling and behavior [83, 101], sociological or psychological factors [68], personality models [22], and dynamic behaviors [43].

1.4 Looking Ahead

Through the collection of chapters presented in this book, we hope to provide reader with an insight that will ultimately lead to addressing some of the following questions:

- What are the general principles that characterize complex crowd behavior of heterogeneous individuals?
- How can verifiable mathematical models of crowd motion and interaction can be developed based on these principles?
- How these general principles can be used to enhance performance of low level vision tasks such as object detection, tracking, and activity analysis in crowds?
- What are the possible problem areas in visual analysis of crowds that will benefit from crowd simulation and behavior models (e.g. tracking, target acquisition across sensor gaps, and sensor hand-off techniques etc.) and vice versa.

The rest of the book is organized into two parts. The first part presents a collection of chapters that focus on experimental validation of various pedestrian motion and interaction models (Chaps. 2 and 4), crowd simulation and behavior modeling (Chaps. 5–8), and relationship between micro and macroscopic models (Chap. 3). The second part of the book focus on approaches of visual analysis of crowded scenes. It covers topics of modeling crowd flows (Chaps. 9, 10, and 12), interaction among crowd participants (Chaps. 11 and 13), crowd counting (Chap. 14) and abnormal event detection (Chap. 15).

Acknowledgements We like to acknowledge Army Research Office, Nippon Telegraph and Telephone, National Science Foundation, and Office of Naval Research for providing support for writing this chapter. We also like to thank Louis Kratz for his contribution to this chapter.

References

1. Adam, A., Rivlin, E., Shimshoni, I., Reinitz, D.: Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Trans. Pattern Anal. Mach. Intell. **30**(3), 555–560 (2008)
2. Ali, S., Shah, M.: A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 17–22 June 2007, pp. 1–6. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4270002&isnumber=4269956
3. Ali, S., Shah, M.: Floor fields for tracking in high density crowd scenes. In: Proceedings of European Conference on Computer Vision. Lecture Notes in Computer Science, vol. 5303 (2008)
4. Andrade, E.L., Blunsden, S., Fisher, R.B.: Modelling crowd scenes for event detection. In: Proceedings of the 18th International Conference on Pattern Recognition, vol. 1, pp. 175–178 (2006). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1698931&isnumber=35817

5. Andrade, E.L., Blunsden, S., Fisher, R.B.: Hidden Markov models for optical flow analysis in crowds. In: Proceedings of the 18th International Conference on Pattern Recognition, vol. 1, pp. 460–463 (2006). . http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1698931&isnumber=35817
6. Bayazit, O.B., Lien, J.-M., Amato, N.M.: Better group behaviors in complex environments with global roadmaps. In: Standish, R.K., Bedau, M.A., Abbass, H.A. (eds.) International Conference on the Simulation and Synthesis of Living Systems (ICAL 2003) (Alife), pp. 362–370. MIT, Cambridge (2002)
7. Betke, M., Hirsh, D.E., Bagchi, A., Hristov, N.I., Makris, N.C., Kunz, T.H.: Tracking large variable numbers of objects in clutter. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '07), 17–22 June 2007, pp. 1–8 (2007). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4270019&isnumber=4269956
8. Brostow, G.J., Cipolla, R.: Unsupervised Bayesian detection of independent motion in crowds. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 594–601, 17–22 June 2006. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1640809&isnumber=34373
9. Cartwright, D., Zander, A.: Group Dynamics: Research and Theory, 3rd edn. Tavistock Publications, London (1968)
10. Chan, A.B., Vasconcelos, N.: Modeling, clustering, and segmenting video with mixtures of dynamic textures. IEEE Trans. Pattern Anal. Mach. Intell. **30**(5), 909–26 (2008)
11. Chattaraj, U., Seyfried, A., Chakroborty, P.: Comparison of pedestrian fundamental diagram across cultures. Adv. Complex Syst. **12**(03), 393–405 (2009)
12. Chenney, S.: Flow tiles. In: Proceedings 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '04), pp. 233–242. Eurographics Association, Aire-la-Ville (2004). http://dx.doi.org/10.1145/1028523.1028553
13. Cheriyadat, A., Radke, R.: Detecting dominant motions in dense crowds. IEEE J. Sel. Top. Signal Process. **2**(4), 568–581 (2008)
14. Dee, H., Hogg, D.: Detecting inexplicable behaviour. In: Proceedings of British Macine Vision Conference, The British Machine Vision Association, pp. 477–486 (2004)
15. Dijkstra, E.W.: A note on two problems in connexion with graphs. Numer. Math. **1**(1), 269–271 (1959)
16. Dobbyn, S., Hamill, J., O'Conor, K., O'Sullivan, C.: Geopostors: a realtime geometry/impostor crowd rendering system. In: Proceedings of the Symposium on Interactive 3D Graphics and Games, New York, pp. 95–102 (2005)
17. Gayle, R., Sud, A., Andersen, E., Guy, S., Lin, M., Manocha, D.: Interactive navigation of independent agents using adaptive roadmaps. IEEE Trans. Vis. Comput. Graph. **15**(1), 34–48 (2009)
18. Ge, W., Collins, R., Ruback, B.: Vision-based analysis of small groups in pedestrian crowds. IEEE Trans. Pattern Anal. Mach. Intell. **34**(5), 1003–1016 (2012)
19. Gilbert, A., Bowden, R.: Multi person tracking within crowded scenes. In: Elgammal, A., Rosenhahn, B., Klette R. (eds.) IEEE Workshop on Human Motion, pp. 166–179. Springer, Berlin/Heidelberg (2007)
20. Guy, S.J., Chhugani, J., Kim, C., Satish, N., Lin, M.C., Manocha, D., Dubey, P.: Clearpath: highly parallel collision avoidance for multi-agent simulation. In: Proceedings of ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 177–187 (2009)
21. Guy, S., Chuggani, J., Curtis, S., Dubey, P., Lin, M., Manocha, D.: Pledestrians: a least-effort approach to crowd simulation. In: Proceedings of Eurographics/ACM SIGGRAPH Symposium on Computer Animation, pp. 119–128 (2010)
22. Guy, S., Kim, S., Lin, M., Manocha, D.: Simulating heterogeneous crowd behaviors using personality trait theory. In: Proceedings of Eurographics/ACM SIGGRAPH Symposium on Computer Animation, pp. 43–52 (2011)
23. Hart, P.E., Nilsson, N.J., Raphael, B.: A formal basis for the heuristic determination of minimum cost paths. IEEE Trans. Syst. Sci. Cybern. **4**(2), 100–107 (1968)

24. Helbing, D., Molnar, P.: Social force model for pedestrian dynamics. Phys. Rev. E **51**, 4282 (1995)
25. Helbing, D., Molnar, P., Farkas, I.J., Bolay, K.: Self-organizing pedestrian movement. Environ. Plan. B Plan. Des. **28**(3), 361–383 (2001)
26. Hu, W., Xiao, X., Fu, Z., Xie, D., Tan, T., Maybank, S.: A system for learning statistical motion patterns. IEEE Trans. Pattern Anal. Mach. Intell. **28**(9), 1450–1464 (2006)
27. Hu, M., Ali, S., Shah, M.: Detecting global motion patterns in complex videos. In: Proceedings of International Conference on Pattern Recognition (ICPR 2008), Dec 2008, pp. 1–5. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4760950&isnumber=4760915
28. Hu, M., Ali, S., Shah, M.: Learning motion patterns in crowded scenes using motion flow field. In: Proceedings of International Conference on Pattern Recognition (ICPR 2008), pp. 1–5, 8–11 Dec (2008). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4761183&isnumber=4760915
29. Hue, C., Le Cadre, J.-P., Perez, P.: Posterior Cramer-Rao bounds for multi-target tracking. IEEE Trans. Aerosp. Electron. Syst. **42**(1), 37–49 (2006)
30. Hughes, R.: A continuum theory for the flow of pedestrians. Transp. Res. B Methodol. **36**(6), 507–535 (2002)
31. Hughes, R.L.: The flow of human crowds. Ann. Rev. Fluid Mech. **35**, 169–182 (2003)
32. Izadinia, H., Saleemi, I., Li, W., Shah, M.: (MP)2T: multiple people multiple parts tracker. In: Computer Vision – ECCV. Lecture Notes in Computer Science, vol. 7577, pp. 100–114 (2012)
33. Jin, X., Xu, J., Wang, C.C.L., Huang, S., Zhang, J.: Interactive control of large crowd navigation in virtual environment using vector field. IEEE Comput. Graph. and Appl. **28**(6):37–46 (2008). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4670099&isnumber=4670088
34. Johnson, N., Hogg, D.: Learning the distribution of object trajectories for event recognition. In: Proceedings of British Macine Vision Conference, pp. 583–592 (1995)
35. Kagarlis, M.: Method and apparatus of simulating movement of an autonomous entity through an environment. United States Patent No. US 7,188,056, Sept. 2002
36. Kallmann, M.: Shortest paths with arbitrary clearance from navigation meshes. In: Proceedings ACM SIGGRAPH Eurographics Symposium on Computer Animation (SCA '10), pp. 159–168. Eurographics Association, Aire-la-Ville (2010)
37. Kamphuis, A., Overmars, M.: Finding paths for coherent groups using clearance. In: Proceedings of ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '04), pp. 19–28. Eurographics Association, Aire-la-Ville (2004). http://dx.doi.org/10.1145/1028523.1028526
38. Karamouzas, I., Heil, P., Beek, P., Overmars, M.: A predictive collision avoidance model for pedestrian simulation. In: Proceedings of Motion in Games, pp. 41–52 (2009)
39. Kavraki, L., Svestka, P., Latombe, J., Overmars, M.: Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Trans. Robot. Autom. **12**(4), 566–580 (1996)
40. Khan, Z., Balch, T., Dellaert, F.: MCMC-based particle filtering for tracking a variable number of interacting targets. IEEE Trans. Pattern Anal. Mach. Intell. **27**(11), 1805–1819 (2005)
41. Khan, Z., Balch, T., Dellaert, F.: MCMC data association and sparse factorization updating for real time multitarget tracking with merged and multiple measurements. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 1960–1972 (2006)
42. Kim, J., Grauman, K.: Observe locally, infer globally, a spacetime MRF for detecting abnormal activities with incremental updates. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR2009), 20–25 June 2009, pp. 2921–2928. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5206569&isnumber=5206488
43. Kim, S., Guy, S., Manocha, D., Lin, M.C.: Interactive simulation of dynamic crowd behaviors using general adaptation syndrome theory. In: Proceedings of Interactive 3D Graphics Symposium (2012)

44. Kratz, L., Nishino, K.: Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition(CVPR 2009), 20–25 June 2009, pp. 1446–1453. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5206771&isnumber=5206488
45. Kratz, L., Nishino, K.: Tracking with local spatio-temporal motion patterns in extremely crowded scenes. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (2010). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5540149&isnumber=5539770
46. Kratz, L., Nishino, K.: Going with the flow: pedestrian efficiency in crowded scenes. In: Proceedings of European Conference on Computer Vision (ECCV 2012). Lecture Notes in Computer Science, vol. 7575, pp. 558–572 (2012)
47. Kratz, L., Nishino, K.: Tracking pedestrians using local spatio-temporal motion patterns in extremely crowded scenes. IEEE Trans. Pattern Anal. Mach. Intell. **34**(5), 987–1002 (2012)
48. Latombe, J.C.: Robot Motion Planning. Kluwer Academic Publishers, Boston (1991)
49. LaValle, S.M.: Planning Algorithms. Cambridge University Press (2006). Also available at http://msl.cs.uiuc.edu/planning/
50. Le Bon, G.: The Crowd: A Study of the Popular Mind Macmillan, New York (1896). Reprint available from Dover Publications
51. Leibe, B., Seemann, E., Schiele, B.: Pedestrian detection in crowded scenes. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR 2005), 20–25 June 2005, vol. 1. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1467359&isnumber=31472
52. Lerner, A., Chrysanthou, Y., Shamir, A., Cohen-Or, D.: Data driven evaluation of crowds. In: Proceedings of the 2nd International Workshop on Motion in Games. Lecture Notes in Computer Science vol. 5884, pp. 75–83 (2009)
53. Lewin, K.: In: Cartwright, D. (ed.) Field Theory in Social Science; Selected Theoretical Papers. Harper & Row, New York (1951)
54. Li, Y., Huang, C., Nevatia, R.: Learning to associate: hybridboosted multi-target tracker for crowded scene. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20–25 June 2009. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5206735&isnumber=5206488
55. Lin, D., et al.: Modeling and estimating persistent motion with geometric flows. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 13–18 June 2010, pp.1–8. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5539848&isnumber=5539770
56. Lozano-Pérez, T., Wesley, M.: An algorithm for planning collision-free paths among polyhedral obstacles. Commun. ACM **22**(10), 560–570 (1979)
57. Magnenat-Thalmann, N., Seo, H., Cordier, F.: Automatic modeling of virtual humans and body clothing. J. Comput. Sci. Technol. **19**(5), 575–584 (2004)
58. Mahadevan, V., Li, W., Bhalodia, V., Vasconcelos, N.: Anomaly detection in crowded scenes. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 13–18 June 2010, pp. 1975–1981 (2010). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5539872&isnumber=5539770
59. Mahadevan, V., Li, W., Bhalodia, V., Vasconcelos, N.: Anomaly detection in crowded scenes. In: CVPR (2010)
60. Maïm, J., Yersin, B., Pettré, J., Thalmann, D.: YaQ: an architecture for real-time navigation and rendering of varied crowds. IEEE Comput. Graph. Appl. **29**(4), 44–53 (2009)
61. McDonnell, R., Larkin, M., Dobbyn, S., Collins, S., O'Sullivan, C.: Clone attack! perception of crowd variety. ACM Trans. Graph. **27**(3), 1–8 (2008)
62. Mehran, R., Oyama, A., Shah, M.: Abnormal crowd behavior detection using social force model. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20–25 June 2009, pp. 935–942. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5206641&isnumber=5206488

63. Mehran, R., Moore, B.E., Shah, M.: A streakline representation of flow in crowded scenes. In: European Conference on Computer Vision (ECCV 2010). Lecture Notes in Computer Science, vol. 6313, pp. 439–452 (2010)
64. Moore, B.E., Ali, S., Mehran, R., Shah, M.: Visual crowd surveillance through a hydrodynamics lens. Commun. ACM **54**, 64–73 (2011)
65. Musse, S.R., Thalmann, D.: A hierarchical model for real time simulation of virtual human crowds. IEEE Trans. Vis. Comput. Graph. **7**(2), 152–164 (2001)
66. Narain, R., Golas, A., Curtis, S., Lin, M.C.: Aggregate dynamics for dense crowd simulation. ACM Trans. Graph. **28**(5), 1–8 (2009)
67. Patil, S., van den Berg, J., Curtis, S., Lin, M.C., Manocha, D.: Directing crowd simulations using navigation fields. IEEE Trans. Vis. Comput. Graph. **17**(2), 244–254 (2011)
68. Pelechano, N., Allbeck, J.M., Badler, N.I.: Controlling individual agents in high-density crowd simulation. In: Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '07), pp. 99–108. Eurographics Association, Aire-la-Ville (2007)
69. Pellegrini, S., Ess, A., Schindler, K., Van Gool, L.: You will never walk alone: modeling social behavior for multi-target tracking. In: Proceedings of IEEE International Conference on Computer Vision and Pattern, Sept. 29-Oct. 2 2009, pp. 261–268. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5459260&isnumber=5459144
70. Pettré, J., de Heras Ciechomski, P., Maïm, J., Yersin, B., Laumond, J.-P., Thal-mann, D.: Real-time navigating crowds: scalable simulation and rendering. J. Vis. Comput. Animat. **17**(3–4), 445–455 (2006)
71. Pettre, J., Grillon, H., Thalmann, D.: Crowds of moving objects: navigation planning and simulation. In: Proceedings of IEEE International Conference on Robotics and Automation, 10–14 April 2007, pp. 3062–3067. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4209555&isnumber=4209049
72. Pettre, J., Ondrej, J., Olivier, A., Cretual, A., Donikian, S.: Experiment-based modeling, simulation and validation of interactions between virtual walkers. In: Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '09), pp. 189–198. ACM (2009). http://doi.acm.org/10.1145/1599470.1599495
73. Raghavendra, R., Del Bue, A., Cristani, M., Murino, V.: Optimizing interaction force for global anomaly detection in crowded scenes. In: Proceedings of IEEE International Conference on Computer Vision (ICCV Workshops), 6–13 Nov 2011, pp. 136–143. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6130235&isnumber=6130192
74. Reynolds, C.W.: Flocks, herds and schools: a distributed behavioral model. Proc. ACM SIGGRAPH **21**, 25–34 (1987)
75. Reynolds, C.W.: Steering behaviors for autonomous characters. In: Game Developers Conference (1999)
76. Rodriguez, M.D., Shah, M.: Detecting and segmenting humans in crowded scenes. In: Proceedings of the 15th International Conference on Multimedia (MULTIMEDIA '07), pp. 353–356. ACM, New York (2007). http://doi.acm.org/10.1145/1291233.1291310
77. Rodriguez, M., Ali, S., Kanade, T.: Tracking in unstructured crowded scenes. In: Proceedings of IEEE International Conference on Computer Vision, 29 Sept-2 Oct 2009, pp. 1389–1396. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5459301&isnumber=5459144
78. Rodriguez, M., Sivic, J., Laptev, I., Audibert, J.: Density-aware person detection and tracking in crowds. In: Proceedings of IEEE International Conference on Computer Vision (ICCV), 6–13 Nov 2011, pp. 2423–2430. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6126526&isnumber=6126217
79. Saleemi, I., Hartung, L., Shah, M.: Scene understanding by statistical modeling of motion patterns. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 13–18 June 2010, pp. 2069–2076. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5539884&isnumber=5539770

80. Sarmady, S., Haron, F., Hj, A.Z.: Modeling groups of pedestrians in least effort crowd movements using cellular automata. In: Proceedings of 3rd Asia International Conference on Modeling and Simulation (AMS '09), pp. 520–525. IEEE Computer Society, Washington, DC (2009). http://dx.doi.org/10.1109/AMS.2009.16
81. Scovanner, P., Tappen, M.: Learning pedestrian dynamics from the real world. In: Proceedings of IEEE International Conference on Computer Vision, 29 Sept-2 Oct 2009, pp. 381–388. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5459224&isnumber=5459144
82. Seyfried, A., Boltes, M., Kähler, J., Klingsch, W., Portz, A., Rupprecht, T., Schadschneider, A., Steffen, B., Winkens, A.: Enhanced empirical data for the fundamental diagram and the flow through bottlenecks. In: Klingsch, W.W.F., Rogsch, C., Schadschneider, A., Schreckenberg, M. (eds.) Pedestrian and Evacuation Dynamics 2008, pp. 145–156. Springer, Berlin/Heidelberg (2010)
83. Shao, W., Terzopoulos, D.: Autonomous pedestrians. In: SCA '05: Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '05), pp. 19–28. ACM, New York (2005). http://doi.acm.org/10.1145/1073368.1073371
84. Simeon, T., Leroy, S., Laumond, J.: Path coordination for multiple mobile robots: a geometric algorithm. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), pp. 1118–1123 (1999)
85. Snape, J., van den Berg, J., Guy, S.J., Manocha, D.: Independent navigation of multiple mobile robots with hybrid reciprocal velocity obstacles. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, pp. 5917–5922 (2009)
86. Snook, G.: Simplified 3D movement and pathfinding using navigation meshes. In: DeLoura, M.A. (ed.) Game Programming Gems, pp. 288–304. Charles River, Hingham (2000). Chapter 3
87. Still, G.: Crowd dynamics. Ph.D. thesis, University of Warwick (2000)
88. Still, K.: Crowd dynamics. Ph.D. thesis, University of Warwick (2000)
89. Sud, A., Andersen, E., Curtis, S., Lin, M., Manocha, D.: Real-time path planning for virtual agents in dynamic environments. In: ACM SIGGRAPH 2008 classes (SIGGRAPH '08), Article 55 , 9pp. ACM, New York (2008)
90. Sugimura, D., Kitani, K., Okabe, T., Sato, Y., Sugimoto, A.: Using individuality to track individuals: clustering individual trajectories in crowds using local appearance and frequency trait. In: Proceedings of IEEE International Conference on Computer Vision, 29 Sept-2 Oct 2009, pp.1467–1474. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5459286&isnumber=5459144
91. Treuille, A., Cooper, S., Popovic, Z.: Continuum crowds. ACM Trans. Graph. **25**(3), 1160–1168 (2006)
92. van den Berg, J., Guy, S.J., Lin, M., Manocha, D.: Reciprocal n-body collision avoidance. In: International Symposium of Robotics Research. Robotics Research Springer Tracts in Advanced Robotics, vol. 70, pp. 3–19 (2009)
93. van den Berg, J., Seawall, J., Lin, M.C., Manocha, D.: Virtualized traffic: reconstructing traffic flows from discrete spatio-temporal data. Proc. IEEE Trans. Vis. Comput. Grap. **17**(1), 26–37 (2009). IEEE Computer Society. http://doi.ieeecomputersociety.org/10.1109/TVCG.2010.27
94. van Toll, W., Cook, A.F., Geraerts, R.: Navigation meshes for realistic multi-layered environments. In: Proceedings of IEEE RSJ International Conference on Intelligent Robots and Systems, 25–30 Sept 2011, pp. 3526–3532. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6094790&isnumber=6094399
95. Wang, X., Tieu, K., Grimson, E.: Learning semantic scene models by trajectory analysis. In: Proceedings of European Conference on Computer Vision, pp. 110–123 (2006)
96. Warren, C.W.: Multiple path coordination using artificial potential fields. In: Proceedings of IEEE Conference on Robotics and Automation, 13–18 May 1990, vol. 1, pp. 500–505. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=126028&isnumber=3534
97. Wu, B., Nevatia, R.: Tracking of multiple, partially occluded humans based on static body part detection. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 951–958 (2006)