MORGAN & CLAYPOOL PUBLISHERS

Virtual Crowds

Methods, Simulation, and Control

Nuria Pelechano
Jan Allbeck
Norman Badler

SYNTHESIS LECTURES ON
COMPUTER GRAPHICS AND ANIMATION

Brian A. Barsky, *Series Editor*

Virtual Crowds: Methods, Simulation, and Control

Synthesis Lectures on Computer Graphics and Animation

Interactive Shape Design
Marie-Paule Cani, Takeo Igarashi and Geoff Wyvill
2008

Real-Time Massive Model Rendering
Sung-eui Yoon, Enrico Gobbetti, David Kasik, and Dinesh Manocha
2008

Virtual Crowds: Methods, Simulation and Control
Nuria Pelechano, Jan. Allbeck, and Norman I. Badler
2008

High Dynamic Range Video
Karol Myszkowski, Rafal Mantiuk, and Grzegorz Krawczyk
2008

GPU-Based Techniques For Global Illumination Effects
L'aszl'o Szirmay-Kalos, L'aszl'o Sz'ecsi and Mateu Sbert
2008

High Dynamic Range Imaging Reconstruction
Asla Sa, Paulo Carvalho, and Luiz Velho IMPA, Brazil
2007

High Fidelity Haptic Rendering
Miguel A. Otaduy, Ming C. Lin
2006

A Blossoming Development of Splines
Stephen Mann
2006

Virtual Crowds: Methods, Simulation, and Control
Nuria Pelechano, Jan M. Allbeck, and Norman I. Badler
www.morganclaypool.com

ISBN: 9781598296419 paperback

ISBN: 9781598296426 ebook

DOI: 10.2200/S00123ED1V01Y200808CGR008

A Publication in the Morgan & Claypool Publishers series

SYNTHESIS LECTURES ON COMPUTER GRAPHICS AND ANIMATION #8

Lecture #8

Series Editor and Affiliation: Brian A. Barsky, University of California–Berkeley

Series ISSN

ISSN 1933-8996 print

ISSN 1933-9003 electronic

Virtual Crowds: Methods, Simulation, and Control

Nuria Pelechano
Universitat Politècnica de Catalunya

Jan M. Allbeck and Norman I. Badler
University of Pennsylvania

SYNTHESIS LECTURES ON COMPUTER GRAPHICS AND ANIMATION #8

MORGAN & CLAYPOOL PUBLISHERS

ABSTRACT

There are many applications of computer animation and simulation where it is necessary to model virtual crowds of autonomous agents. Some of these applications include site planning, education, entertainment, training, and human factors analysis for building evacuation. Other applications include simulations of scenarios where masses of people gather, flow, and disperse, such as transportation centers, sporting events, and concerts. Most crowd simulations include only basic locomotive behaviors possibly coupled with a few stochastic actions. Our goal in this survey is to establish a baseline of techniques and requirements for simulating large-scale virtual human populations. Sometimes, these populations might be mutually engaged in a common activity such as evacuation from a building or area; other times they may be going about their individual and personal agenda of work, play, leisure, travel, or spectator. Computational methods to model one set of requirements may not mesh well with good approaches to another. By including both crowd and individual goals and constraints into a comprehensive computational model, we expect to simulate the visual texture and contextual behaviors of groups of seemingly sentient beings.

KEYWORDS

Animated characters, autonomous agents, CAROSA, collision avoidance, computer animation, crowd simulation, evacuation studies, HiDAC, human behaviors, navigation planning, parameterized actions, pedestrians, presence, psychological factors, roles, social forces, virtual environments

Dedication

To our families

Acknowledgments

We thank the many individuals and organizations that have supported this work with their donations, intellect, talent, and funding. We are grateful to Autodesk, nVIDIA, and Cal3D for their continued development of software and hardware that enrich computer graphics in general and the Center for Human Modeling and Simulation (HMS) at the University of Pennsylvania in particular. We thank all of the members of HMS for their support of these efforts and individually recognize Jeff Wajcs, Grace Fong, Funda Durupinar, and Catherine Stocker for their contributions. Additionally, we appreciate the guidance provided by Joan Pelechano, Mel Slater, and Ali Malkawi. The perceptive comments of the book reviewers are greatly appreciated, which we hope are reflected in better content, coverage, and structure of this work. Thanks also to our publisher, Michael Morgan, who encouraged us to undertake this project and whose enthusiasm kept us close to schedule.

The authors' research described in this volume was partially supported by the National Science Foundation grant IIS-0200983, the Office of Naval Research Virtual Technologies and Environments grant N0001 4-04-1-0259, the U.S. Army Research Office grants N61339-05-C-0081 and MURI W911NF-07-1-0216, the T. C. Chan Center Building Simulation and Energy Studies, a Fulbright scholarship, a Spanish Government grant TIN2007-67982-C02-01, and the School of Engineering and Applied Science at the University of Pennsylvania. Any opinions expressed are entirely those of the authors and do not reflect any official positions of the sponsors.

Contents

CHAPTER 1

Introduction

As we journey through our day, our lives intersect with other people. We see people leaving for work, waiting for trains, meeting with friends, working at their jobs, and engaging in numerous other activities. People create a rich tapestry of activity throughout the day, a *human texture*. We may not always be conscious of this texture, but we would definitely notice if it were missing, and it *is* missing in many computer graphics simulations of 3D environments populated by collections of animated virtual humans.

There are many applications of computer animation and simulation where it is necessary to model virtual crowds of autonomous agents. Some of these applications include site planning, education, entertainment, training, and human factors analysis for building evacuation. Other applications include simulations of scenarios where masses of people gather, flow, and disperse, such as transportation centers, sporting events, and concerts. Most crowd simulations include only basic locomotive behaviors possibly coupled with a few stochastic actions. Our goal in this survey is to establish a baseline of techniques and requirements for simulating large-scale virtual human populations. Sometimes these populations might be mutually engaged in a common activity, such as evacuation from a building or area; other times they may be going about their individual and personal agenda of work, play, leisure, travel, or spectator. Computational methods for modeling one set of requirements may not mesh well with good approaches to another. By including both crowd and individual goals and constraints into a comprehensive computational model, we expect to simulate the visual texture and contextual behaviors of groups of seemingly sentient beings.

The structure of this exposition includes surveys of existing computational crowd motion models, descriptive models originating in data analysis and urban planning, and functional models of human behavior. We stop short of exploring the details of individual human motion animation techniques, preferring instead to focus on the collective structure, motion, and control of groups and the differentiation and individualism of roles within groups. Within this matrix, however, one can readily embed various computational methods for animating more personal details of individuals adapted to the spatial context and task execution desired.

1.1 TERMINOLOGY

A wide variety of terms appear in the literature that refer to humans "inhabiting" virtual worlds. *Avatars* are characters that represent and are controlled directly by a real person. The term *avatar* is often confused with characters that take their motivations and behaviors from computer programs and simulators. These computer-driven characters or *virtual humans* may also be called *digital humans*, *autonomous agents*, or *humanoids*. We use these terms interchangeably. Although finer distinctions are plausible, they are not meaningful here: sometimes "autonomous" is equated to "undirected" behaviors, but we prefer to consider all action choices as under some sort of computational control. Different virtual human control mechanisms exist, and we will differentiate their qualities and characteristics shortly.

Existing terminology used to describe multiple beings can be even more confusing. *Artificial life*, "*boids*," and *multiagent systems* simulate more than one (autonomous) character. Other terms used include *crowds*, *pedestrians*, *groups*, and *populations*. It is not always clear what is meant by these terms or the functionality of the individuals that these systems represent. Here we will briefly examine some of the terminology associated with crowds.

WordNet definitions (Fellbaum 1998):

- *crowd*: a large number of things or people considered together
- *pedestrian*: a person who travels by foot
- *populace*: people in general considered as a whole
- *population*: the people who inhabit a territory or state.

Crowd, *populace*, and *population* all inherit from the *group* hypernym. However, *crowd* is also a *gathering* and a *social group*, whereas *populace* and *population* are people. Hyponyms for *crowd* include:

- *army*: a large number of people united for some specific purpose
- *crush*, *jam*, *press*: a dense crowd of people
- *drove*, *horde*, *swarm*: a moving crowd
- *huddle*: a disorganized and densely packed crowd
- *mob*, *rabble*, *rout*: a disorderly crowd of people
- *lynch mob*: a mob that kills a person for some presumed offense without legal authority
- *phalanx*: any closely ranked crowd of people
- *troop*, *flock*: an orderly crowd.

We are not claiming that WordNet is necessarily the best source for clarification, but we found it to be a useful jumping off point. Looking at the psychosocial literature on crowds yields, for example, Figure 1.1, showing a taxonomy of crowds as described by Brown (1954). Here Brown first makes a division into active (mobs) and passive (audiences) crowds. He then further breaks down these divisions according to their purpose or feeling. For example, an intentional, recreational audience may be watching a basketball game.

From a computational perspective, the question is: do all crowd simulators really simulate crowds? Systems that animate more than a few agents seem to be regarded as crowd simulators by the research community, but not all are. Terzopoulos and colleagues (Shao and Terzopoulos 2005; Yu and Terzopoulos 2007) have called their work simulations of "autonomous pedestrians". As they simulate commuters in a train station, this seems an apt term. One of the simulations shown by Treuille et al. (2006), "continuum crowds," depicts agents on city streets; these might also be deemed pedestrians. The seminal work by Reynolds (1987) on "flocks, herds, and schools" — generically called boids — seems to be very appropriately named. Much of the work in crowd simulations has been for evacuation (Helbing et al. 2000; Pelechano and Badler 2006). Figure 1.1 shows that these events would be categorized as escapes and may be further broken into events with and without panic. Here escapes are a type of mob which, in turn, is a type of crowd. For the remainder of this exposition, we will use the common word *crowd* to describe simulations of modest numbers of characters, though for some of the discussions, terms such as pedestrians or populace may be more appropriate.

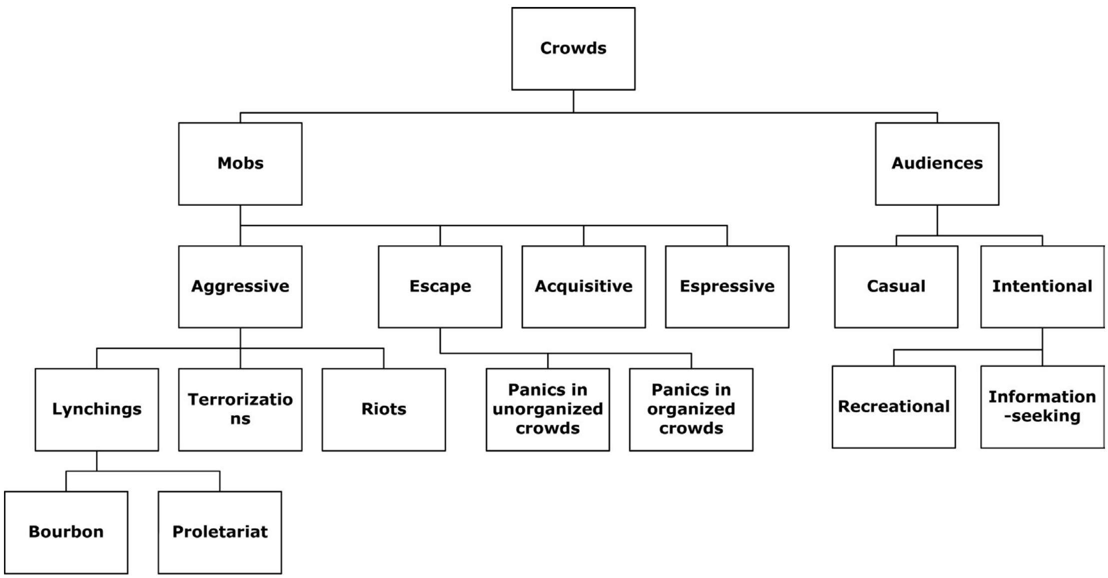

FIGURE 1.1: Mass phenomena from Brown (1954).

FIGURE 1.2: Examples of crowd behavior simulated with our system. On the left is a drill evacuation simulation and on the right is a cocktail party simulation, where virtual agents perform several actions to interact with other agents (Pelechano, Stocker et al. 2008).

To provide an overall structure for the presentation, we use the context of three of our own simulation systems that have been joined into a single framework. HiDAC was constructed for simulating high-density crowds, i.e., *crushes* or *presses*, but is parameterizable and can also simulate *huddles*, *mobs*, *escapes*, and *flocks*, for example. HiDAC combines a social forces model with a rule-based model to yield interesting emergent behaviors (Section 5.5). MACES builds on HiDAC to simulate additional behaviors, including evacuation, wayfinding with interagent communication and certain personal roles (Chapter 6). CAROSA builds on the foundations of HiDAC and MACES, but moves the focus away from *escapes* by incorporating a richer role and action representation that allows the depiction of functional heterogeneous crowds (Chapter 7). CAROSA might be used to simulate *audiences*, but would better be described as a system for simulating a populace or population of individuals such as the daily activities of inhabits of an office building or neighborhood, or perhaps even an entire city (Figure 1.2).

1.2 OVERVIEW

Our intention here is to provide insight into crowd simulation software design choices and their resultant characteristics and features. We begin by providing a survey of crowd simulation research.

It can be difficult to compare virtual human technologies. Table 1.1 shows an attempt to create a few features and points along a scale for each of the features. *Appearance* fairly straightforwardly describes aspects of the visual qualities of characters. *Function* reflects the capabilities of the virtual humans. What movements and behaviors are possible and how accurate are they? Are the behaviors a visualization of the state of the character in terms of injuries and psychology, for example? Ultimately, can the behaviors of the virtual human coordinate with other virtual humans to create teams? *Time* reflects efficiency in computation. Is the computation too heavy for even interactive manipulations? Can the motions be synthesized in real time? Can more than one character be simulated in real time, and can members of the crowd be coordinated at a viable frame rate? *Autonomy* indicates the level at which the character can control itself. Does the character creator have to specify every moment of the character's behavior, or can it make its own decisions? Can the autonomous character aid in the decision making of other virtual humans and, therefore, act as a leader? *Individuality* indicates to what level characters can be differentiated from one another in terms of the other features. Do all of the characters look and behave the same? Can observers recognize cultural distinctions and personality types? Ultimately, can specific individuals be recognized?

Table 1.1 was originally created with one or a few virtual humans in mind. Here we are more interested in larger numbers. In terms of *appearance*, we would certainly desire variability for realistic scenes. This is also true of *functionality*. We would not want to see every character in a scene walking identically, for example. Recent work by McDonnell et al. (2007, 2008) (Figure 1.3) has explored the amount of differentiation needed in these features for crowds. The *time* feature already

TABLE 1.1: Comparative virtual humans originally published by Allbeck and Badler (2002)

Appearance	2D drawings > 3D wireframe > 3D polyhedra > curved surfaces > freeform deformations > accurate surfaces > muscles, fat > biomechanics > clothing, equipment > physiological effects (perspiration, irritation, injury)
Function	cartoon > jointed skeleton > joint limits > strength limits > fatigue > hazards > injury > skills > effects of loads and stressors > psychological models > cognitive models > roles > teaming
Time (time to create movement at the next frame)	off-line animation > interactive manipulation > real-time motion playback > parameterized motion synthesis > multiple agents > crowds > coordinated teams
Autonomy	drawing > scripting > interacting > reacting > making decisions > communicating > intending > taking initiative > leading
Individuality	generic character > hand-crafted character > cultural distinctions > sex and age > personality > psychological-physiological profiles > specific individual

included how many characters can be simulated per frame. *Autonomy* may be extended to include mechanisms for characters to coordinate, collaborate, and even compete with other agents. The characters would themselves decide when a task requires additional personnel and schedule appropriately. *Individuality* makes crowd scenes more interesting and realistic. Also important is the formation of groups. People have affiliations with others that change over time and even during the course of a day. In scenarios with larger numbers of virtual humans, these dynamic changes become important for realism.

Crowd simulation researchers tend to focus on furthering development in one or perhaps a couple of these areas. McDonnell and colleagues have focused mainly on overall appearance (McDonnell et al. 2007, 2008), but this relates to time in that frame rate is affected by the resolution of the characters (McDonnell et al. 2005). Ahn et al. (2006) examined using not just the level of detail in character models but also the level of detail in motion to increase the number of characters that can be simulated in real time. Many research groups have looked to increase the level of function of crowds particularly in the area of navigation and collision avoidance (Reynolds 1999; Helbing et al. 2005; Treuille et al. 2006; Pelechano et al. 2007; Sud, Andersen, et al. 2007). Collaboration and coordination within crowd simulations has been explored, but much more work is needed (Shao and Terzopoulos 2005; Pelechano and Badler 2006; Yu and Terzopoulos 2007). Individuality has been the focus not only of many autonomous agents researchers, but also of a few crowd simulation research groups (Musse and Thalmann 2000; Pelechano et al. 2005; Pelechano and Badler 2006; Durupinar et al. 2008).

Animating virtual crowds is often mediated by local rules (Musse and Thalmann 2001), forces (Helbing et al. 2000), or flows (Chenney 2004). The goal is usually either to achieve real-time simulation for very large crowds, where each individual's behavior is not important as long as the overall crowd movement looks realistic, or to focus on individual behaviors using complex cognitive models (but achieving real time only for smaller crowds). Much effort has been put into improving

FIGURE 1.3: Perception of crowd variety (McDonnell et al. 2008) C 2008 ACM, Inc. Reprinted by permission.

the behavioral realism of each of these approaches; however, none of those models can realistically animate the complexity and function of population movements in a large building or a city. Current work in crowd motion has focused on realistically animating moderate and high-density crowds for real-time applications, where agents are endowed with psychological elements that will drive not only their high-level decision making, but also their reactive behavior (pushing, moving faster, being impatient, etc.) (Pelechano et al. 2007). The task of specifying and animating large collections of functional individuals is just beginning to be addressed.

On the global navigation level, most approaches either deal with simple environments or assume that agents have complete knowledge of the environment and move toward their goal as individuals (without interacting with other agents). Other work has focused on realistically simulating how communication affects the behavior of autonomous agents (Cassell et al. 1999) and on how combining different personalities and situations affects an agents' navigation and the way it interacts with other agents and the environment (Pelechano and Badler 2006).

To have autonomous agents navigating a virtual environment, it is necessary to endow them with a wayfinding algorithm to obtain a cognitive map of a building or environment. Wayfinding is the process of determining and following a route to some destination (Golledge 1999). It deals with the cognitive component of navigation and, therefore, with the knowledge and the information processing required to move (locomote) from an initial position to a goal position. Initially, the individuals of the crowd may only have partial information about the internal building structure, but as they explore it and communicate with other individuals in the crowd, they will be able to find paths toward their goals or exits.

Wayfinding is defined as a spatial problem-solving process with three subprocesses: decision making, decision execution, and information processing. To carry out wayfinding, each agent would need:

- a cognitive map: a mental model of the space
- an orientation: its current position within the cognitive map
- exploration: processes to learn the features of the space (doors, walls, hazards, etc).
- navigation: the process of making its way through the environment.

There have been several cognitive agent architectures proposed to generate human-like behavior. Cognitive architecture features include a broad set of possible requirements, such as perception, memory, attention, planning, reasoning, problem solving, learning, emotions, and mood. Agent architectures motivated from a cognitive perspective generally consist of a knowledge representation, algorithms that learn, and modules that plan actions based on that knowledge (Wray et al. 1999; Yu and Terzopoulos 2007). Tu and colleagues have worked on behavioral animation for creating artificial life, where virtual agents are endowed with synthetic vision and perception of the environment (Tu