Chapter 2
State-of-the-Art

One of the largest areas where crowd behaviors have been modeled is the domain of safety science and architecture with the dominant application of crowd evacuation simulators. Such systems model movements of a large number of people in usually closed and well-defined spaces like inner areas of buildings [TM95a, BBM05], subways [Har00], ships [KMKWS00], or airplanes [OGLF98]. Their goal is to help designers to understand the relation between the organization of space and human behavior [OM93].

The most common use of evacuation simulators is the modeling of crowd behavior in case of forced evacuation from a confined environment due to some threat like fire or smoke. In such a situation, a number of people have to evacuate the given area, usually through a relatively small number of fixed exits. Simulations are trying to help answer questions like: Can the area be evacuated within a prescribed time? Where do the holdups in the flow of people occur? Where are the likely areas for a crowd surge to produce unacceptable crushing pressure [Rob99]? The most common modeling approach in this area is the use of cellular automata serving both as a representation of individuals and as a representation of the environment.

Simulex [TM95a, TM95b] is a computer model simulating the escape movement of persons through large, geometrically complex building spaces defined by 2D floor plans and connecting staircases. Each individual has attributes such as position, body size, angle of orientation, and walking speed. Various algorithms as distance mapping, way finding, overtaking, route deviation, and adjustment of individual speeds due to proximity of crowd members are used to compute egress simulation, where individual building occupants walk toward and through the exits.

G. Still developed a collection of programs named Legion for simulation and analysis of the crowd dynamics in evacuation from constrained and complex environments like stadiums [Sti00]. Dynamics of crowd motion is modeled by mobile cellular automata. Every person in the crowd is treated as an individual, calculating its position by scanning its local environment and choosing an appropriate action.

Helbing et al. [HM95, HFV00, WH03] proposed a model based on physics and sociopsychological forces in order to describe the human crowd behavior in panic situations. The model is set up by a particle system where each particle $i$ of mass $m_i$...
has a predefined speed $v^0_i$, i.e., the desired velocity, in a certain direction $e^0_i$ to which it tends to adapt its instantaneous velocity $v_i$ within a certain time interval $\tau$ (for 1st term of Eq. (2.1)). Simultaneously, the particles try to keep a velocity-dependent distance from other entities $j$ and walls $w$ controlled by interaction forces $f_{ij}$ and $f_{iw}$ (second and third terms of Eq. (2.1)), respectively. The change of velocity with time $t$ is given by the dynamical equation:

$$m_i \frac{dv_i}{dt} = F_i^{(H)} = m_i \frac{v^0_i e^0_i - v_i(t)}{\tau_i} + \sum_{j \neq i} f_{ij} + \sum_w f_{iw}$$ (2.1)

Braun et al. [BMB03, BBM05] extended the Helbing Model ($F_i^{(H)}$) in order to deal with different individuals and group behaviors, and also with complex environments. In this work, the agents’ population can be composed heterogeneously by individuals with different attributes.

This chapter presents several works on crowd domain, as crowd dynamics and simulation.

2.1 Crowd Dynamics

The behavior of real crowds was analyzed in [Hen71, Hen74, Fru71, Hel97, Sti00]; results of their analysis provide a useful reference for simulation and animation of crowds. Two important aspects that guide the motion of real people are: goal seeking, reflecting the target destination of each individual; and the least-effort strategy, reflecting the tendency of people to reach the goal along a path requiring the least effort [Sti00]. According to these strategies, people travel along smooth trajectories, since this requires less energy than frequent changes of direction or speed. In particular, adjustments of direction and speed, required to avoid collisions, are minimized. Further consequences of the least-effort strategy are the formation of lanes and the speed reduction effect. The first term refers to the tendency of people walking in the same directions to reduce their effort by closely following each other, while the second one refers to the reduction of speed in dense crowds.

The concept of personal space, the subject of study of interpersonal interactions in a spatial context (proxemics) [Hal59], also plays an important role in population dynamics. Personal space can be thought of as an area with invisible boundaries, surrounding each individual, which should not be penetrated by other individuals in order for interpersonal interactions to occur comfortably. The size of this zone depends on the environment as well as the people culture, and decreases as the crowd density gets higher. In the context of simulations, the personal space determines the minimum distance that should be maintained among the agents.

2.2 Sociological Models of Crowds

Despite being a field primarily interested in studying collective behavior, only a relatively small number of works on crowd simulations have been done in sociology.
2.3 Crowd Simulation

McPhail et al. [CPC92] studied individual and collective actions in temporary gatherings. Their model of the crowd is based on perception control theory [Pow73] where each separate individual is trying to control his or her experience in order to maintain a particular relationship to others: in this case it is a spatial relationship with others in a group. The simulation program GATHERING graphically shows movement, milling, and structural emergence in crowds. The same simulation system was later used by Schweingruber [Sch95] to study the effects of reference signals common to coordination of collective behavior and by Tucker et al. [TSM99] to study formation of arcs and rings in temporary gatherings.

Jager et al. [JPvdS01] modeled clustering and fighting in two-party crowds. A crowd is modeled by a multi agent simulation using cellular automata with rules defining approach–avoidance conflict. The simulation consists of two groups of agents of three different kinds: hardcore, hangers-on, and bystanders, the difference between them consisting in the frequency with which they scan their surroundings. The goal of the simulation was to study effects of group size, size symmetry, and group composition on clustering, and “fights”.

2.3 Crowd Simulation

Virtual crowds are usually modeled as collections of interacting agents, although treating a crowd as a continuum (for example, obeying laws of fluid dynamics) is also possible [TCP06]. In behavioral models, movements of a group of agents are an emergent property of individual agents, which are both influenced by and influencing their neighbors. These individual behaviors are defined using sets of simple goal-oriented rules, such as “move with the average speed of your neighbors” or “keep an optimal distance to your neighbors”. Behavioral animation was pioneered by Reynolds [Rey87], who simulated flocks of birds and schools of fish assuming that each agent has direct access to the motion characteristics (position and velocity) of other agents. Tu and Terzopoulos [TT94] improved the conceptual realism of this work by endowing artificial fish with synthetic vision and perception of the environment. Both the original results of Reynolds and the models of Tu and Terzopoulos were confined to relatively small, low-density groups of animals.

To control human crowds, Musse and Thalmann [MT01] proposed a hierarchical crowd organization with groups of different levels of autonomy. In a related work, Ulicny and Thalmann [UT01] proposed a model that provided agent control at the level of an individual, a group, and the crowd.

The rules governing the movements of agents in behavioral models may be viewed as an abstract representation of the “psychology” of modeled individuals. In contrast, in force-field models, interactions among agents (in this case, often referred to as particles) are based on analogies with physics. For example, Bouvier et al. [BCN97] modeled individuals in a high-density crowd as charges moving in electric fields. Helbing et al. [HM97, HFV00] introduced abstract attraction and repulsion forces to simulate groups of people in panic situations. Braun et al. [BMB03]
extended this model by endowing agents with individual characteristics and including the concept of groups, which improved the realism of simulations.

In data-driven models, the motion of real crowds is used to calibrate simulations. For example, Musse et al. [MJJB07] used computer vision techniques to track individual agents in images obtained using a video camera, then applied the resulting statistics to drive a physically-based simulator, while Lee et al. [LCHL07] used aerial images for a similar purpose. Lerner and collaborators [LCL07] proposed a model for collision avoidance using a manually-built database of videos of pedestrians. The goal of this work was to improve the realism of collisions treatment using real-world data.

Hybrid methods have also been proposed. For instance, the approaches presented by [PAB07] and [vdBPS*08] integrate behavioral and force-fields techniques in order to improve crowd control, aiming to minimize the drawbacks of both technologies. However, the negative aspect of these methods is the increase in complexity of the implementation.

Each category of models presents a tradeoff. Behavioral models are suited for an individualized specification of agents, but global crowd control is more difficult to achieve because of the emergent character of the motions. In contrast, force-field models offer good global crowd control in high-density situations, but tend to generate less realistic motions of individual characters, which reflect their simplistic physical basis. Finally, data-driven models make it possible to improve the realism of simulations, but the acquisition and interpretation of real-life data is often difficult. This sets the stage for our method, in which crowds of agents obeying simple behavioral rules can be globally controlled and relatively realistic motions can be obtained without tedious parametrization as emergent properties of the model.

2.4 Behavioral Animation of Groups and Crowds

Human beings are arguably the most complex known creatures, therefore they are also the most complex creatures to simulate. A behavioral animation of human (and humanoid) crowds is based on foundations of group simulations of much more simple entities, notably flocks of birds [Rey87, GA90] and schools of fish [TT94]. The first procedural animation of flocks of virtual birds was shown in the movie by Amkraut, Girard, and Karl called Eurhythmy, for which the first concept [AGK85] was presented at The Electronic Theater at SIGGRAPH in 1985 (final version was presented at Ars Electronica in 1989). The flock motion was achieved by a global vector force guiding a flow of flocks [GA90].

In his pioneering work, Reynolds [Rey87] described a distributed behavioral model for simulating aggregate motion of a flock of birds. The technical paper was accompanied by an animated short movie called “Stanley and Stella in: Breaking the Ice” shown at the Electronic Theater at SIGGRAPH’87. The revolutionary idea was that a complex behavior of a group of actors can be obtained by simple local rules for members of the group instead of some enforced global condition. The flock
is simulated as a complex particle system, with the simulated birds (called boids) being the particles. Each boid is implemented as an independent agent that navigates according to its local perception of the environment, the laws of simulated physics, and the set of behaviors. The boids try to avoid collisions with one another and with other objects in their environment, match velocities with nearby flock mates, and move toward a center of the flock. The aggregate motion of the simulated flock is the result of the interaction of these relatively simple behaviors of the individual simulated birds. Reynolds later extended his work by including various steering behaviors as goal seeking, obstacle avoidance, path following, or fleeing [Rey99], and introduced a simple finite-state machines behavior controller and spatial queries optimizations for real-time interaction with groups of characters [Rey00].

Tu and Terzopoulos proposed a framework for animation of artificial fishes [TT94]. Besides complex individual behaviors based on perception of the environment, virtual fishes have been exhibiting unscripted collective motions as schooling and predator evading behaviors analogous to flocking of boids. An approach similar to boids was used by Bouvier et al. [BG96, BCN97] to simulate human crowds. They used a combination of particle systems and transition networks to model crowds for the visualization of urban spaces. At the lower level, attractive and repulsive forces, analogous to physical electric ones, enable people to move around the environment. Goals generate attractive forces, obstacles generate repulsive force fields. Higher level behavior is modeled by transition networks with transitions depending on time, visiting of certain points, changes of local population densities, and global events.

Brogan and Hodgins [HB94, BH97] simulated group behaviors for systems with significant dynamics. Compared to boids, a more realistic motion is achieved by taking into account physical properties of motion, such as momentum or balance. Their algorithm for controlling the movements of creatures proceeds in two steps: first, a perception model determines the creatures and obstacles visible to each individual, and then a placement algorithm determines the desired position for each individual given the locations and velocities of perceived creatures and obstacles. Simulated systems included groups of one-legged robots, bicycle riders, and point-mass systems. Musse and Thalmann [Mus00, MT01] presented a hierarchical model for real-time simulation of virtual human crowds. Their model is based on groups, instead of individuals: groups are more intelligent structures, where individuals follow the groups specification. Groups can be controlled with different levels of autonomy: guided crowds follow orders (as go to a certain place or play a particular animation) given by the user in run-time; programmed crowds follow a scripted behavior; and autonomous crowds use events and reactions to create more complex behaviors. The environment comprises a set of interest points, which signify goals and way points; and a set of action points, which are goals that have some actions associated. Agents move between way points following Bezier curves.

Recently, another work was exploring group modeling based on hierarchies. Niederberger and Gross [NG03] proposed an architecture of hierarchical and heterogeneous agents for real-time applications. Behaviors are defined through specialization of existing behavior types and weighted multiple inheritance for creation
of new types. Groups are defined through recursive and modulo based patterns. The behavior engine allows for the specification of a maximal amount of time per run in order to guarantee a minimal and constant frame rate.

Ulicny and Thalmann [UT01, UT02] presented a crowd behavior simulation with a modular architecture for multiagent system allowing autonomous and scripted behavior of agents supporting variety. In their system, the behavior is computed in layers, where decisions are made by behavioral rules and execution is handled by hierarchical finite-state machines. Most recently, a real-time crowd model based on continuum dynamics has been proposed by [TCP06]. In their model, a dynamic potential field integrates global navigation with moving obstacles, efficiently solving for the motion of large crowds without the need for explicit collision avoidance. Perceived complexity of the crowd simulation can be increased by using levels of detail (LOD). O’Sullivan et al. [OCV*02] described a simulation of crowds and groups with level of details for geometry, motion, and behavior. At the geometrical level, subdivision techniques are used to achieve smooth rendering LOD changes. At the motion level, the movements are simulated using adaptive levels of detail. Animation subsystems with different complexities, as a keyframe player or a real-time reaching module, are activated and deactivated based on heuristics. For the behavior, LOD is employed to reduce the computational costs of updating the behavior of characters that are less important. More complex characters behave according to their motivations and roles, less complex ones just play random keyframes. The behavior of autonomous characters has been widely studied in the area of crowd simulation during the past few years. Most crowd simulation models obtain plausible macroscopic behaviors but have a limited ability to manage behavioral autonomy. Decision systems are generally applied to simple reactive behaviors such as collision avoidance because of the computational cost of implementing existing rational models with a crowd of virtual people. To address these challenges, Paris and Donikian [PD09] proposed a crowd simulation cognitive model that can be used to develop complex goal-oriented behaviors for numerous virtual people in real time. The model integrates a decision process that provides a full bidirectional link between four layers, biomechanical, reactive, cognitive, and rational (see Allen Newell’s Unified Theories of Cognition). Each layer informs the layer directly above it of specific information on imposed constraints and controls the layer directly underneath it. Each layer is built independently and exchanges only a set of identified data.

Shao and Terzopoulos artificial life approach [ST07] integrates motor, perceptive, behavioral, and cognitive components within a comprehensive model of pedestrians as individuals. They claimed that the model can yield results of unprecedented fidelity and complexity for fully autonomous multi-human simulation in a large urban environment. Following Tu and Terzopoulos [TT94], they adopted a bottom-up strategy that uses primitive reactive behaviors as building blocks that in turn support more complex motivational behaviors, all controlled by an action selection mechanism. The behavioral model consists of basic reactive behaviors, navigational behaviors, motivational behaviors, mental state and action selection. Realistic behavioral modeling, whose purpose is to link perception to appropriate actions, is a
big challenge in the case of autonomous virtual humans. Even for 3 pedestrians, the complexity of any substantive behavioral repertoire is high. Except computer graphics, many relevant studies in psychology, ethology, artificial intelligence, robotics, and artificial life are devoted to this subject. With these behavioral models, virtual humans can be interacted with in some situations. Here the behavioral models are limited to the applications of pedestrians. The cognitive and perceptual components are also helpful in improving the plausibility of crowd simulation, e.g. the problem of local collision can be avoided by behavioral modeling. However, complex behavioral models are usually too expensive to be used in real-time massive crowd simulation. For the interactive crowd simulation, a behavioral model is necessary, and we have to face the tradeoff between precision of the behavioral the model and the computing time.

2.5 Crowd Management Training Systems

The modeling of crowds has also been essential in police and military simulator systems used for training in how to deal with mass gatherings of people.

CACTUS [Wil95] is a system developed to assist in planning and training for public order incidents such as large demonstrations and marches. The software designs are based on a world model in which crowd groups and police units are placed on a digitized map and have probabilistic rules for their interactive behavior. The simulation model represents small groups of people as discrete objects. The behavioral descriptions are in the form of a directed graph where the nodes describe behavioral states (to which correspond actions and exhibited emotions) and transitions represent plausible changes between these states. The transitions depend on environmental conditions and probability weightings. The simulation runs as a decision making exercise that can include pre-event logistic planning, incident management, and debriefing evaluation.

Small Unit Leader Non-Lethal Training System [VSMA98] is a simulator for training U.S. Marines Corps in decision making with respect to the use of nonlethal munitions in peacekeeping and crowd control operations. Trainees learn rules of engagement, the procedures for dealing with crowds and mobs, and the ability to make decisions about the appropriate level of force needed to control, contain, or disperse crowds and mobs. Crowds move within a simulated urban environment along instructor-predefined pathways and respond both to actions of a trainee and to actions of other simulated crowds. Each crowd is characterized by a crowd profile—series of attributes like fanaticism, arousal state, prior experience with nonlethal munitions, or attitude toward Marines. During an exercise, the crowd behavior computer model operates in real time and responds to trainee actions (and inactions) with appropriate simulated behaviors such as loitering, celebrating, demonstrating, rioting, and dispersing according to a set of Boolean relationships defined by experts.
2.6 Group Behavior in Robotics and Artificial Life

Researchers working in the field of artificial life are interested in exploring how group behavior emerges from local behavioral rules [Gil95]. Software models and groups of robots were designed and experimented with in order to understand how complex behaviors can arise in systems guided by simple rules. The main source of inspiration is nature, where, for example, social insects efficiently solve problems such as finding food, building nests, or division of labor among nestmates by simple interacting individuals without an overseeing global controller. One of the important mechanisms contributing to a distributed control of the behavior is stigmergy, indirect interactions among individuals through modifications of the environment [BDT99].

Dorigo introduced ant systems inspired by behaviors of real ant colonies [Dor92]. Ant algorithms have been successfully used to solve a variety of discrete optimization problems including the traveling salesman problem, sequential ordering, graph coloring, or network routing [BDT00]. Besides insects, groups of more complex organisms such as flocks of birds, herds of animals, and schools of fish have been studied in order to understand principles of their organization. Recently, Couzin et al. presented a model of how animals that forage or travel in groups can make decisions even with a small number of informed individuals [CKFL05].

Principles from biological systems were also used to design behavior controllers for autonomous groups of robots. Mataric studied behavior-based control for a group of robots, experimenting with a herd of 20 robots whose behavioral repertoire included safe wandering, following, aggregation, dispersion, and homing [Mat97]. Molnar and Starke have been working on assignment of robotic units to targets in a manufacturing environment using a pattern formation inspired by pedestrian behavior [PJ01]. Martinoli applied swarm intelligence principles to autonomous collective robotics, performing experiments with robots that were gathering scattered objects and cooperating to pull sticks out of the ground [A.99]. Holland and Melhuish experimented with a group of robots doing sorting of objects based on ant behaviors where ants sort larvae and cocoons [HM99]. In an interesting work using a robot to control animal behavior, Vaughan et al. developed a mobile robot that gathers a flock of real ducks and maneuvers them safely to a specified goal position [VSH*00].

2.7 Environment Modeling for Crowds

2.7.1 Environment Models

Environment modeling is closely related to behavioral animation. The purpose of the models of the environment is to facilitate simulation of entities dwelling in their surrounding environments. Believability of virtual creatures can be greatly enhanced if they behave in accordance with their surroundings. On the contrary, the suspense of disbelief can be immediately destroyed if they perform something not expected or
not permitted in the real world, such as passing through the wall or walking on water. The greatest efforts have therefore been directed to representations and algorithms preventing forbidden behaviors from occurring; until quite recently the two major artificial intelligence issues concerning game development industry were collision avoidance and path-planning [Woo99, DeL00]. The majority of the population in the developed world lives in cities; it is there that most human activities take place nowadays. Accordingly, most of the research has been done for modeling of virtual cities. Farenc et al. [FRMS*99] introduced an informed environment dedicated to the simulation of virtual humans in the urban context. The informed environment is a database integrating semantic and geometrical information about a virtual city. It is based on a hierarchical decomposition of an urban scene into environment entities, like quarters, blocks, junctions, streets, and so on. Entities can contain a description of the behaviors that are appropriate for agents located on them; for example, a sidewalk tells that it should be walked on, or a bench tells that it should be sat on. Furthermore, the environment database can be used for a path-finding that is customized according to the type of client requesting the path, so that, for example, a pedestrian will get paths using sidewalks, but a car will get paths going through roads.

Another model of a virtual city for a behavioral animation was presented by Thomas and Donikian [TD00]. Their model is designed with the main emphasis on traffic simulation of vehicles and pedestrians. The environment database is split into two parts—a hierarchical structure containing a tree of polygonal regions, similar to the informed environment database; and a topological structure with a graph of a road network. Regions contain information on directions of circulation, including possible route changes at intersections. The agents then use the database to navigate through the city. In a recent work, Sung et al. [SGC04] presented a new approach to control the behavior of a crowd by storing behavioral information into the environment using structures called situations. Compared with previous approaches, environmental structures (situations) can overlap; behaviors corresponding to such overlapping situations are then composed using probability distributions. Behavior functions define probabilities of state transitions (triggering motion clips) depending on the state of the environment features or on the past state of the agent.

### 2.7.2 Path Planning

Path planning is an important and challenging task in crowd simulation, which helps each agent to find the path to its individual goal. The path planning problem has been widely explored by the robotics community. Although the multiple-agent path planning has been addressed for cooperative tasks of multiple robots, it is still a challenge to solve the path planning problem for large crowds in real time, especially for large-scale crowds. Because the methods used for robots are usually exponential in the number of robots, which are too expensive to be adopted in crowd simulation.
Benefit from motion planning algorithms in robotics, geometric representation of probabilistic roadmaps (PRM) can also be used for path planning in crowd simulation. PRM was applied to solve the problem of determining a collision-free path between a starting configuration of the robot and a goal configuration [KSLO96]. Arikan et al. [ACF01] used the visibility graph for the path planning for large numbers of virtual agents. The visibility graph connects together vertices of the environment if and only if they see each other. The PRM-based approaches were improved by being integrated with other techniques [BLA02a, BLA02b, SKG05]. Kallmann et al. [KBT03] proposed a fast path-planning algorithm based on a fully dynamic constrained Delaunay triangulation. Bayazit et al. [BLA02a] used global roadmaps to improve group behaviors in geometrically complex environments. Groups of creatures exhibited behaviors such as homing, goal searching, covering, or shepherding, by using rules embedded both in individual flock members and in roadmaps. Tang et al. [TWP03] used a modified A* algorithm working on a grid overlaid over a height-map generated terrain. Other approaches of geometric representation of environments have been explored specially for the path planning of multi-agent systems. Lamarche and Donikian [LD04] built an accurate hierarchical topological structure from geometric database of a virtual environment. They performed the following steps for the final navigation: spatial subdivision, topology abstraction, roadmap generation, and triangulation construction. It is reported that this approach can allow the navigation of several hundreds of agents in real time. Kamphuis and Overmars defined a walkable corridor that ensured sufficient clearance to allow a given group of units to pass [KO04]. The Voronoi diagram can be used to subdivide a free space based on a set of points, from which edges are generated to produce the roadmap. Sud et al. [SAC*08] proposed a new data structure based on Voronoi diagrams, which is used to perform path planning and proximity computations for each agent in real time. Inspired from Voronoi diagrams, Pettré et al. [PLT05, PdHCM*06, PGT08] presented a novel approach to automatically extract a topology from a scene geometry and handle path planning using a navigation graph. The environment is usually discretized into a fine regular grid in the potential field method. Kapadia et al. [KSHF09] introduced a discretization method of egocentric fields with variable resolution information representation. Helbing’s social force model [HM95] is one of the most influential models in agent-based motion planning. This model considers each agent as a particle subject to long-ranged forces induced by the social behavior of individuals. The movement of agents can be described with a main function which determines the physical and social forces, similar to Newtonian mechanics. The social force model is capable of describing the self-organization of several observed collective effects of pedestrian behavior. Nevertheless, due to lack of anticipation and pre-diction, the characters interact when they get sufficiently close. Consequently, the resulting motions tend to look unnatural and contain undesirable oscillations. The problem becomes more obvious in large and cluttered environments. This model was extended to achieve more realistic crowd behaviors [HBJW05, LKF05]. Karamouzas et al. [KHBO09] introduced the evasive force to improve the social force model. Their approach is based on the hypothesis that an individual adapts its route as early as possible, trying to minimize
the amount of interactions with others and the energy required to solve these interactions. With this model, the agents do not repel each other, but rather anticipate future situations avoiding collisions long in advance and with minimal effort. However, the applications of social force model are limited by the calculation efficiency because of its complex rules. Treuille et al. [TCP06] proposed realistic motion planning for crowds, making an analogy with potential fields. Their method produces a potential field from the addition of a static field (goal) and a dynamic field (modeling other people). Each pedestrian then moves against the gradient towards the next suitable position in space (a waypoint) and thus avoids all obstacles. Compared to agent-based approaches, these techniques allow simulating thousands of pedestrians in real time, and are also able to show emergent behaviors. However, they produce less believable results, because they require assumptions that prevent treating each pedestrian with individual characteristics. For instance, only a limited number of goals can be defined and assigned to sets of pedestrians. The resulting performance depends on the size of the grid cells and the number of sets. Hybrid architectures have also been explored to get advantages from different approaches. Pelechano et al. [PAB07] combined psychological, physiological, and geometrical rules with physical forces to simulate dense crowds of autonomous agents. Morini et al. [MYMT08] proposed a hybrid architecture to handle the path planning of thousands of pedestrians in real time, while ensuring dynamic collision avoidance. This method is detailed in Sect. 5.4.4. To address regions of varied interest, motion planning is ruled by different algorithms. Hybrid path planning should be deliberately designed to ensure agent motion continuity when switching between different algorithms.

### 2.7.3 Collision Avoidance

Except the topological model of the environment and path planning, collision avoidance is another challenging problem to be addressed. The collision avoidance techniques should be efficient enough to prevent a large number of agents from bumping into each other in real time. The greatest difficulty of collision avoidance is from the absence of other agents current velocities. Furthermore, the agents are not able to communicate to coordinate their navigation. A common solution to this problem is to assume that the other agents are dynamic obstacles whose future motions are predicted as linear extrapolations of their current velocities. The agent then selects a velocity that avoids collisions with the extrapolated trajectories of other agents. This is the idea of Velocity Obstacle. Considering the case in which each agent navigates independently without explicit communication with other agents, van den Berg et al. [vdBPS*08] propose a new concept the “Reciprocal Velocity Obstacle”, which takes into account the reactive behavior of the other agents by implicitly assuming that the other agents make a similar collision avoidance reasoning. This concept can be applied to navigation of hundreds of agents in densely populated environments containing both static and moving obstacles for real-time simulation.
Ondrej et al. [OPOD10] explored a novel vision-based approach of collision avoidance between walkers that fits the requirements of interactive crowd simulation. By simulating humans based on cognitive science results, visual stimuli is used to detect future collisions as well as the level of danger. The motor-response is twofold: a reorientation strategy prevents future collision, whereas a deceleration strategy prevents imminent collisions. Several simulation results show that the emergence of self-organized patterns of walkers is reinforced using this approach. The overall efficiency of the walkers’ traffic is improved and improbable locking situations is avoided.

2.8 Crowd Rendering

Real-time rendering of a large number of 3D characters is a considerable challenge; it is able to exhaust system resources quickly even for state-of-the-art systems with extensive memory resources, fast processors, and powerful graphic cards. “Brute-force” approaches that are feasible for a few characters do not scale up for hundreds, thousands, or more of them. Several works have been trying to circumvent such limitations by clever use of graphics accelerator capabilities, and by employing methods profiting from the fact that our perception of the scene as a whole is limited.

We can perceive in full detail only a relatively small part of a large collection of characters. A simple calculation shows that to treat every crowd member as equal is rather wasteful. Modern screens can display around 2 million pixels at the same time, where a fairly complex character can contain approximately 10,000 triangles. Even if assuming that every triangle were be projected to a single pixel, and that there would be no overlap of characters, the screen fully covered by a crowd would contain only around 200 simultaneously visible characters. Of course, in reality the number would be much smaller; a more reasonable estimate is around a few dozen fully visible characters, with the rest of the crowd either being hidden behind these prominent characters or taking significantly less screen space. Therefore, it makes sense to take full care only of the foremost agents, and to replace the others with some less complex approximations. Level of details techniques then switch visualizations according to position and orientation of the observer. In the recent work of Hamill et al. [HMDO05] they pursue psychophysics, a discipline to decide perceptual limitations to the human vision system for example. Doing tests on how motion affects the perception of a human represented by an impostor or by a geometric structure, they were able to define distances of least noticeable switching between models.

Billboarded impostors are one of the methods used to speed up crowd rendering. Impostors are partially transparent textured polygons that contain a snapshot of a full 3D character and are always facing the camera. Aubel et al. [ABT00] introduced dynamically generated impostors to render animated virtual humans. In their approach, an impostor creating process is running in parallel to full 3D simulations, taking snapshots of rendered 3D characters. These cached snapshots are then used
over several frames instead of the full geometry until a sufficient movement of either camera or a character will trigger another snapshot, refreshing the impostor texture.

In another major work using impostors, Tecchia et al. [TLC02a] proposed a method for real-time rendering of an animated crowd in a virtual city. Compared with the previous method, impostors are not computed dynamically, but are created in a preprocessing step. Snapshots are sampled from viewpoints distributed in the sphere around the character. This process is repeated for every frame of the animation. In run-time, images taken from viewpoints closest to the actual camera position are then used for texturing of the billboard. Additionally, the silhouettes of the impostors are used as shadows projected to a ground surface. Multitexturing is used to add variety by modulating colors of the impostors. In a later work they added lighting using normal maps [TLC02b]. Their method using precomputed impostors is faster than dynamical impostors, but it is very demanding on texture memory, which leads to lower image quality as size of textures per character and per animation frame have to be kept small.

A different possibility for a fast crowd display is to use point-based rendering techniques. Wand and Strasser [WS02] presented a multiresolution rendering approach which unifies image based and polygonal rendering. They create a view-dependent octree representation of every keyframe of animation, where nodes store either a polygon or a point. These representations are also able to interpolate linearly from one tree to another so that in-between frames can be calculated. When the viewer is at a long distance, the human is rendered using point rendering; when zoomed in, using polygonal techniques; and when in between, a mix of the two.

An approach that has been getting new life is that of geometry baking. By taking snapshots of vertex positions and normals, complete mesh descriptions are stored for each frame of animation as in the work of Ulicny et al. [UdHCT04]. Since current desktop PCs have large memories, many such frames can be stored and replayed. A hybrid approach of both baked geometry and billboard was presented at I3d, where only a few actors are fully geometrical while the vast number of actors are made up of billboards [DHO005]. A similar approach can be found in [CLM05]. A more recent approach to crowd rendering using geometry is through dynamic meshes as presented in the work of de Heras et al. [dHSM05], where dynamic meshes use systems of caches to reuse skeletal updates which are typically costly. A hybrid of dynamic and baked meshes is found in [YMdHC*05] where the graphics programming unit (GPU) is used to its fullest.

What is common to all approaches is instancing of template humans, by changing the texture or color, size, orientation, animation, animation style, and position. This is carefully taken care of to smoothly transition from one representation to another so as not to create pops in representation styles. In the billboard scenario this is done by applying different colors on entire zones such as torso, head, legs, and arms. This way the texture memory is used more efficiently as the templates are more flexible. For the geometrical approaches these kinds of differences are usually represented using entirely different textures as the humans are too close just to change basic color for an entire zone [UdHCT04].
2.9 Crowds in Non-real-time Productions

One of the domains with the fastest growth of crowd simulations in recent years is special effects. While only 10 years ago, there were no digital crowds at all, nowadays almost every blockbuster has some, with music videos, television series, and advertisements starting to follow. In comparison with crowds of real extras, virtual crowds allow one to significantly reduce costs of production of massively populated scenes and allow for bigger creative freedom because of their flexibility. Different techniques, as replications of real crowd video footage, particle systems, or behavioral animation, have been employed to add crowds of virtual extras to shots in a broad range of movies, from historical dramas,1,2,3 through fantasy and science fiction stories,4,5,6 to animated cartoons.7,8,9

The main factors determining the choice of techniques are the required visual quality and the production costs allowed for the project [Leh02]. It is common to use different techniques even in a single shot in order to achieve the best visuals; for example, characters in the front plane are usually real actors, with 3D characters taking secondary roles in the background.

Although a considerable amount of work was done on crowds in movies, only relatively little information is available, especially concerning more technical details. Most knowledge comes from disparate sources, for example, from “making-of” documentary features, interviews with special effects crew or industry journalist accounts. For big budget productions, the most common approach is in-house development of custom tools or suites of tools which are used for a particular movie. As the quality of the animation is paramount, large libraries of motion clips are usually used, produced mainly by motion capture of live performers. All production is centered around shots, most of the times only a few seconds long. In contrast to real-time simulations, there is little need for continuity of the simulation over longer periods of the time. It is common that different teams of people work on parts of the shots which are then composited in postprocessing.

The most advanced crowd animation system for non-real-time productions is Massive; used to create battle scenes for The Lord of the Rings movie trilogy.10 In Massive, every agent makes decisions about its actions depending on its sensory inputs using a brain composed of thousands of logic nodes [Koe02].

---

4 http://www.starwars.com/.
5 http://www.lordoftherings.net.
to the brain’s decision, the motion is selected from an extensive library of motion captured clips with precomputed transitions. For example, in the second part of the trilogy over 12 million motion captured frames (equivalent to 55 hours of animation) were used. Massive also uses rigid body dynamics, a physics-based approach to facilitating realistic stunt motion such as falling, or animation of accessories. For example, a combination of physics-based simulation and custom motion capture clips was used to create the scene of “The Flooding of Isengard” where orcs are fleeing from a wall of water and falling down the precipice [Sco03].

In comparison with real-time application, the quality of motion and visuals in non real-time productions is far superior, but it comes at a great cost. For example, for The Lord of the Rings: The Two Towers, rendering of all digital characters took 10 months of computations on a strong render farm with thousands of computers [Doy03].

2.10 Crowds in Games

In current computer games, virtual crowds are still relatively rare. The main reason is that crowds are inherently costly, both in terms of real-time resource requirements and for costs of a production. Nevertheless, the situation is starting to change, with the real-time strategy genre leading the way as increase of sizes of involved armies has a direct effect on gameplay [Rom04, The04].

The main concern for games is the speed of both rendering and behavior computation. In comparison with non-real-time productions, the quality of both motion and rendering is often sacrificed in a trade-off for fluidity. Similarly to movie production, computer games often inject realism into virtual worlds from the real world by using large libraries of animations, which are mostly motion captured. The rendering uses level-of-detail techniques, with some titles employing animated impostors [Med02].

To improve costs of behavior computations for games that involve a large number of simulated entities, simulation level-of-detail techniques have been employed [Bro02, Rep03]. In such techniques, behavior is computed only for characters that are visible or soon to be visible. Characters are created in a space around the player with parameters set according to some expected statistical distributions, the player lives in a “simulation bubble”. However, handling simulation LOD is much more complex than handling rendering LOD. It is perfectly correct not to compute visualization for agents that are not visible, but not computing behaviors for hidden agents can lead to an incoherent world. In some games it is common that the player causes some significant situation (for example, traffic jam), looks away, and then after looking back, the situation is changed in an unexpected way (a traffic jam is “magically” resolved).

In case the scenario deals with hundreds or thousands of entities, many times the selectable unit with distinct behavior is a formation of troops, not individual
soldiers. What appears to be many entities on the screen is indeed only one unit being rendered as several visually separated parts\textsuperscript{11} [Med02].\textsuperscript{12}

A special case are sport titles such as various football, basketball, or hockey simulations, where there is a large spectator crowd, but only of very low details. In most cases there is not even a single polygon for every crowd member (compared with individual impostors in strategy games). A majority of the crowd is just texture with transparency applied to stadium rows, or to a collection of rows, and only a few crowd members, close to the camera, are 3D models.

2.11 Crowd Scenario Authoring

Regardless of the quality of crowd rendering or the behavioral model, a virtual crowd simulation is not very useful if it is too difficult to produce content for it. The authoring possibilities are an important factor influencing the usability, of a crowd simulation system, especially when going beyond a limited number of “proof-of-concept” scenarios. When increasing the number of involved individuals, it becomes more difficult to create unique and varied content of scenarios with a large number of entities. Solving one set of problems for crowd simulation (such as fast rendering and behavior computation for large crowds) creates a new problem of how to create content for crowd scenarios in an efficient manner.

Only recently have researchers started to explore ways of how to author crowd scenes. Anderson et al. [AMC03] achieved interesting results for a particular case of flocking animation following constraints. Their method can be used, for instance, to create and animate flocks moving in shapes. Their algorithm generates a constrained flocking motion by iterating the simulation forwards and backwards. Nevertheless, their algorithm can get very costly when increasing the number of entities and simulation time.

Ulicny et al. [UdHCT04] proposed a method to create complex crowd scenes in an intuitive way using a Crowdbrush tool. By employing a brush metaphor, analogous to the tools used in image manipulation programs, the user can distribute, modify, and control crowd members in realtime with immediate visual feedback. This approach works well for creation and modification of spatial features; however, the authoring of temporal aspects of the scenario is limited.

Sung et al. [SGC04] used a situation-based distributed control mechanism that gives each agent in a crowd specific details about how to react at any given moment based on its local environment. A painting interface allows one to specify situations easily by drawing their regions on the environment directly like drawing a picture on the canvas. Compared with previous work where the user adds, modifies, and deletes crowd members, here the interface operates on the environment.

\textsuperscript{11}http://www.totalwar.com.
\textsuperscript{12}http://praetorians.pyrostudios.com.
Chenney [Che04] presented a novel technique for representing and designing velocity fields using flow tiles. He applied his method on a city model with tiles defining the flow of people through the city streets. Flow tiles drive the crowd using the velocity to define the direction of travel for each member. The use of divergence-free flows to define crowd motion ensures that, under reasonable conditions, the agents do not require any form of collision detection.

References


[Rey99] Reynolds C. W.: Steering behaviors for autonomous characters. In Game Developers Conference (San Jose, California, USA, 1999), pp. 763–782.


