Modeling Social Group Interactions for Realistic Crowd Behaviors

Seung In Park

Dissertation submitted to the faculty of
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Computer Science and Applications

Francis Quek, Chair
Yong Cao, Chair
A. Lynn Abbott
Doug Bowman
Yang Cao

February 6, 2013
Blacksburg, Virginia

Modeling Social Group Interactions for Realistic Crowd Behaviors

Seung In Park

(ABSTRACT)

In the simulation of human crowd behavior including evacuation planning, transportation management, and safety engineering in architecture design, the development of pedestrian model for higher behavior fidelity is an important task. To construct plausible facsimiles of real crowd movements, simulations should exhibit human behaviors for navigation, pedestrian decision-making, and social behaviors such as grouping and crowding. The research field is quite mature in some sense, with a large number of approaches that have been proposed to path finding, collision avoidance, and visually pleasing steering behaviors of virtual humans. However, there is still a clear disparity between the variety of approaches and the quality of crowd behaviors in simulations.

Many social science field studies inform us that crowds are typically composed of multiple social groups (James, 1953; Coleman and James, 1961; Aveni, 1977). These observations indicate that one component of the complexity of crowd dynamics emerges from the presence of various patterns of social interactions within small groups that make up the crowd. Hence, realism in a crowd simulation may be enhanced when virtual characters are organized in multiple social groups, and exhibit human-like coordination behaviors.

Motivated by the need for modeling groups in a crowd, we present a multi-agent model for large crowd simulations that incorporates socially plausible group behaviors. A computational model for multi-agent coordination and interaction informed by well-established Common Ground theory (Clark, 1996; Clark and Brennan, 1991) is proposed. In our approach, the task of navigation in a group is viewed as performing a joint activity which requires maintaining a state of common ground among group members regarding walking strategies and route choices. That is, group members communicate with, and adapt their behaviors to each other in order to maintain group cohesiveness while walking. In the course of interaction, an agent may present gestures or other behavioral cues according to its communicative purpose. It also considers the spatiotemporal conditions of the agent-group's environment in which the agent interacts when selecting a kind of motions.

With the incorporation of our agent model, we provide a unified framework for crowd simulation and animation which accommodates high-level socially-aware behavioral realism of animated characters. The communicative purpose and motion selection of agents are consistently carried through from simulation to animation, and a resulted sequence of animated character behaviors forms not merely a chain of reactive or random gestures but a socially meaningful interactions.

We conducted several experiments in order to investigate the impact of our social group

interaction model in crowd simulation and animation. By showing that group communicative behaviors have a substantial influence on the overall distribution of a crowd, we demonstrate the importance of incorporating a model of social group interaction into multi-agent simulations of large crowd behaviors. With a series of perceptual user studies, we show that our model produces more believable behaviors of animated characters from the viewpoint of human observers.

Acknowledgments

I would like to thank to the following people for their help, support and advice in this dissertation. First and foremost, I am greatly indebted to my parents and grandmother for their unconditional love, understanding, support, and patience during the entire period of my study. My special thanks go to my fiance, Kaeyoung Shin. He has given me constant encouragement, dedicated support, and love throughout my pursuit of Ph.D. Without him, I doubt I would have finished this dissertation. My brother Seung Yeon also deserves thanks for his encouragement and support.

I would like to express my deepest gratitude to my advisor, Dr. Francis Quek. He helped me select the right topic, and has given me motivation, insightful comments, and invaluable suggestions throughout the course of this work. Without his continued support and advice, this work would not have been accomplished. I would also like to extend my deepest gratitude to my another advisor, Dr. Yong Cao, for his knowledgeable guidance, constructive comments, and time on my work. He even sat with me to find bugs in my code. I would like to acknowledge my committee members, Dr. Doug Bowman, Dr. Yang Cao, and Dr. Lynn Abbott, for their helpful remarks and guidance.

My colleagues in the VisLab and Animation Lab at Virginia Tech have been a great help in discussing research ideas and preparing my presentation. In particular, I want to thank Ji-Sun Kim for her helpful comments and discussion in my user study design and conduction. I would also like to thank Chao Peng for his contribution in developing the crowd rendering engine. Blake Sawyer brought the idea of using a cover story in the perceptual user evaluation. I would also like to acknowledge Rongrong Wang, Bing Fang, Yannick Verdie, Chreston Miller, Sharon Chu, Junpeng Wang, Mi Peng, Yasmin Elglaly, and Xiao Lin for their help in different aspects along the way. It has been invaluable to have such good colleagues in the lab. Finally, thanks to Jing Huang and all Korean friends those who have made my stay at Virginia Tech enriching and enjoyable.

Table of Contents

List of Figures

List of Tables

Glossary

Before entering into the main part of this dissertation, we introduce key terms used throughout this document. The following definitions are intended to help readers understand the terms used in this dissertation and the context in which they are used.

Agent is an autonomous entity capable of perceiving its virtual environment and interacting with other objects. It moves towards a desired goal position while avoiding collisions with obstacles in the environment and coordinating its behaviors to other agents.

Virtual environment is a computer-simulated environment in which agents navigate around and interact with objects in it.

Goal is a specific and definable geographic point in a given virtual environment. Agents traverse a virtual environment to reach their goals.

Group is any number of people or agents considered as a social unit. Group members maintain group cohesiveness by communicating and adapting their behaviors to each other.

Chapter 1

Introduction

1.1 Motivation

In the simulation of human crowd behavior including evacuation planning, transportation management, and safety engineering in architecture design, the development of pedestrian model for higher behavioral realism is an important and challenging task. To construct plausible facsimiles of real crowd movements, simulations should exhibit human behaviors for navigation, pedestrian decision-making, and social behaviors such as grouping and crowding. The research field is quite mature in some sense, with a large number of approaches that have been proposed to path finding, collision avoidance, and visually compelling steering behaviors of virtual humans. However, there is still a clear disparity between the variety of approaches and the quality of crowd behaviors in simulations.

Social science field studies inform us that crowds are not made up of a mass of isolated individuals. Up to 70% of observed pedestrians are walking in groups, and most groups are composed of a small number of members (James, 1953; Coleman and James, 1961; Aveni, 1977; Sommer, 1979). In an earlier field study in Oregon, James (James, 1953) reported that 74% of groups contains only 2 individuals, 20% contains 3, 4% contains 4, and 1% includes 5 and more. These observations indicate that one component of the complexity of crowd dynamics emerges from the presence of various patterns of social interactions within small groups that make up the crowd. Therefore pedestrian behaviors should be modeled by taking a perspective of mindful entity as a social group member to achieve the enhanced realism in crowd simulations.

Typical approaches to crowd modeling tend to treat a crowd either as an aggregated whole or as a collection of isolated individuals. A widely used method in the former approaches involves the application of fluid dynamics (Hughes, 2002; Narain et al., 2009). To the degree that humans are physical entities, and exhibit some particle-like behavior in crowds, fluid dynamics-based approaches are able to model such behavior. However, humans are autonomous deciders, and these decisions result in complexity that simple fluid dynamics modeling cannot address. In the latter approaches, an agent-based modeling method is commonly used. Towards the creation of intelligent and autonomous agent which reacts in a believable way to the surroundings, the aspects of cognitive, psychological, emotional, and cultural functioning of human are incorporated into the behavior model of agent (Rao and Georgeff, 1995; Hoogendoorn and Bovy, 2004; Antonini et al., 2006; Endrass et al., 2011). However, what is overlooked in these approaches is that the intelligence of groups differs in a sense from individual intelligence because it is an emergent phenomena driven by the need of the groups to maintain cohesive activity.

The absence of considering the social group dynamics and its impact on human crowd behavior should be addressed to achieve enhanced realism and valid simulation results in crowd simulation and animation. Our research goal in this dissertation is to present a crowd model that accommodates high-level socially-aware behavioral realism, where individuals within the crowd belong to different social groups, and where the behavior responds to the surroundings in the simulation environment.

1.2 Problem Statement

Several aspects need to be addressed in the design and development of crowd simulation applications that consider group organization. First, in order to produce plausible aggregate group behaviors, it should be taken into account that a group is more than a simple collection of individuals, and individuals in a group interact and coordinate with each other with respect to common motives and goals. Second, while the results of crowd simulations may be presented in different ways depending on the objective of applications, visualizing the simulation output with 3D animated characters is necessary to allow the analysis of crowd behaviors. Therefore the simulation model should facilitate crowd animation. Third, the outcome of the simulation model should be validated with regard to if the model captures real crowd effects and creates more correct group behavior patterns.

1.2.1 Socio-psychological Approach to Behavior Modeling

Human behaviors are affected by a wide variety of social and psychological factors, such as social norms, culture, and emotion. With the integration of agent-based modeling techniques, there have been a number of approaches to incorporate some existing social and psychological theories into the model of human behaviors (Pelechano, 2005; Hoogendoorn and Soumokil, 2010; Bouchet and Sansonnet, 2011; Kim et al., 2012b). Most of these approaches focus on improving the intelligence and reactivity of individual agent with the information from the theories.

However, a group is a social unit comprising several members who stand in status and relationships with one another. Behaviors of individual members are regulated in matters of consequence to the group. In group activities, people perform actions such as body movements, gestures, and eye gazes as means of participating with others in the group. Therefore, in order to better model agent-agent interaction for plausible social group behaviors, it is essential to be informed by models of human group interaction and coordination from the social sciences. A crowd model employing such an agent behavior model of high-level sociality will be capable of generating the crowd movement that is affected by agents being part of a social group, and thus will result in more realistic crowd simulation.

1.2.2 Groups within Crowds

Generating believable group behaviors within a crowd has been a focus of some research efforts in crowd simulation recently. In several studies, attentions have been put to reproduce walking patterns and spatial organizations of small groups in a crowd simulation (Peters and Ennis, 2009; Moussaïd et al., 2010; Karamouzas and Overmars, 2010; Lemercier et al., 2012). In these approaches, empirical data of pedestrian crowds are collected using video recordings, and the observed group formations are analyzed. The desired formations are represented as reference points in a local coordinate system of a group, and used to guide each group member's relative position. Maintaining a desired formation while walking at each cycle is formulated as a collective optimization problem for group members, and the individual movement is manipulated at a reactive motion planning level.

However, walking in a group is not just a matter of how to maneuver to reach a desired position at a low level. People communicate with other members, and trade off certain action, path, and location according to the particular situation of a group. It is hard to mechanistically construct such higher-level behavioral activity completely from bottom up. For a simulation of human-like group behaviors, we need to consider social interactions not only at a reactive motion level but also at a higher level for making decisions over time.

A behavioral aspect in group dynamics has been also considered in some research. The effect of actions and gestures, such as interacting distance, orientation, and synchrony of visual and aural cues of actions, on the plausibility of conversing groups have been identified and applied to the simulation of groups of conversing characters in (Ennis et al., 2011; Ennis and O'Sullivan, 2012). However, in these approaches, the selection of stance, movement, and motion is not tightly coupled to the underlying simulation model, and the sequence of character gestures does not draw a socially meaningful story. Also, their approaches have focused on generating static conversing characters, thereby limiting applications in a crowd simulation.

1.2.3 Crowd Animation

Although crowd modeling and visualization has been the subject of many previous research efforts, not many of them have been concerned with animating the resulted crowd simulation using 3D fully articulated virtual humans and their behavioral realism and correctness.

In many approaches, abstract representation of characters such as point particles are used (Helbing and Molnar, 1995; Reynolds, 1999; Helbing et al., 2001). Other work employed human-like figures (Pelechano et al., 2007; Kapadia et al., 2009; Guy et al., 2010), but much of this work has focused on presenting a realistic macroscopic pedestrian flow, and the detailed movement and gesture of individual characters were overlooked.

Animating more complex interacting behaviors among multiple characters has been studied recently. Data-driven approaches using motion capture data to create a dense crowd of virtual humans with expressive gestures have been proposed in (Lai et al., 2005; Lee et al., 2006; Kim et al., 2012a; Ju et al., 2010). They collected motion episodes of multiple characters in which the characters interact with each other for a short amount of time. Then a larger crowd of interacting characters for an extended period of time is generated by weaving the episodes. However, these approaches are expensive techniques because of the computational complexity to create connecting transitions between the motion data segments. They have been also suffer from the lack of flexibility in that the motion production is always limited by the given set of pre-collected capture data, and individual motion control is not allowed. More importantly, each motion episode is fragmentary, therefore simply stitching the episodes cannot form a meaningful long-term crowd phenomena.

In these days research in crowd simulation, it is still an open question that how simulation and animation can be unified in terms of some kind of 'behavioral intent' so that socially plausible animation behaviors are produced in a crowd simulation.

1.2.4 Crowd Model Evaluation

Another important however often overlooked aspect of research in crowd simulation is evaluation and validation of a proposed crowd model. Compared to the large number of crowd simulation models that have been developed, very limited progress has been made in validating the crowd models. In a few number of studies, simulated crowd movements have been compared to real world data to demonstrate that their models are able to bring the real crowd phenomena into simulations (Helbing et al., 2000; Seyfried et al., 2007; Guy et al., 2010). Another common approach to model validation has been conducting user studies to assess if the desired quality of the simulation has been obtained (Peters and Ennis, 2009; Castellano et al., 2012).

However, there exists yet no standard method to objectively evaluate and validate the behaviors of agents and crowds. This is because the expected characteristics of crowd behavior largely depend on target scenarios and applications of the simulation model. Hence we have to find out how to examine the impact of a proposed model, and prove the importance of considering such a model in crowd simulations. For example, if we demonstrate that different crowd distribution patterns are observed in simulations with and without our model of group behaviors, this indicates that crowd models that do not consider the cost of social interaction may fail to capture the real crowd effects.

Demonstrating some differences resulted in simulations due to a certain crowd model, though, does not mean that the model produces more correct or believable simulations. To investigate how a proposed crowd model affects the realism, believability, and interpretability of a simulation, we have to validate the outcomes of simulation with human judgement.

1.3 Research Questions

The problem statement above introduces a range of research opportunities for crowd simulation with particular attention to social group behaviors. We defined five research questions that we addressed through the research presented in this dissertation, in order to improve the behavioral realism of computer generated agents in crowd simulations.

1. **How to design a plausible crowd model where individuals within the crowd belong to different social groups, and where the behavior responds to the surroundings in the simulation environment?**

   In this research question, we were concerned with the need for incorporating the model of higher-level human group behaviors into a crowd model. As mentioned in the problem statement, walking in a group is not only a low-level locomotor activity but also a higher-level joint activity which requires group members to act in coordination with each other. Therefore it is necessary to integrate social approaches to physicalist approaches to address both higher cognitive/communicative processes for coordinated behaviors and low-level reactive motions for collision-free and cohesive movement of individuals in a crowd.

2. **How to operationalize the conceptual crowd model for large crowd simulations?**

   In order to bring the concept to reality, we should provide a formal and explicit representation of a conceptual crowd model, and operationalize the model with a concrete implementation design for target computational platforms. In addition to that, this operationalization must be computationally tractable concerning the simulation execution time, and scalable with respect to the size of crowd.

   Another problem with many existing crowd models is in that the applications of these models are highly limited by domain specific settings and constraints. We need a crowd model which allows a flexible adaptation of the model to different simulation scenarios and situations.

3. **How to map the simulation data into animated virtual characters and their motions? Can we integrate behavioral intent across simulation and animation layers?**

   In order to generate more correct and believable behaviors of animated characters for social group interaction, the behavioral purpose of simulation entities should be consistently maintained across mass simulation, individual simulation, and animation layers.

One of our research goals was to provide a complete integration framework of crowd simulation and animation so that the motion selection information generated by communicative intents of individuals in the simulation model is carried to and used in a rendering engine for animating virtual human characters and their actions.

4. **What is the effect of modeling social groups in a crowd simulation?**

   Given that there is no standard way to evaluate crowd behaviors in crowd modeling and simulation, we sought for what aspect of simulation needs to be measured and would provide useful insight regarding the impact of a proposed model in the resulted simulations.

5. **If there is an effect of our model, does it produce more correct simulated crowd behaviors, and how do we determine this?**

   If our fourth research question is answered in the affirmative (i.e., that the model produces results different than existing approaches), how do we know that these results are more correct? This requires that we derive ways to measure this correctness, and to apply these to our results to determine the correctness of our model. This will tell us whether the derived crowd model ultimately produces better crowd simulation, and whether our operationalization of the theory from social sciences produces realistic crowd models. The question may be reformulated by asking whether our model produces plausible animations with human judgement.

1.4 Approach

1.4.1 Model of Social Group Interaction

We advance a computational model informed by Clark's Common Ground (CG) theory to incorporate the impact of social group dynamics into crowd behaviors. According to Clark, people engage in a joint activity when they act in coordination with others to pursue a common goal. CG is negotiated among participants in a joint activity, and has to be maintained for the coordination to be successful. It considers the mutual knowledge, beliefs, and assumptions among collaborating individuals.

Clark's CG model has two distinct strengths. First, it is scientifically validated model on human behaviors. Second, it consists of logically separable components and thus lends itself to a direct operationalization of coordination process. The CG concept as the basis of joint activity has seen many applications in AI agents (Klein et al., 2005), human-robot collaboration (Kirby et al., 2009), and Computer-Supported Cooperative Work (Monk, 2003; Neale et al., 2004; Convertino et al., 2008).

In our CG-based crowd model, the task of navigation in a group is viewed as performing a joint activity with group members, and members of a group have to maintain a state of CG regarding route choices and walking strategies for group cohesiveness. We shall call our model the Common Ground-based Crowd Simulation (CGCS) model throughout this disseratation.

1.4.2 Operationalization

The agent-based approach may be the most natural way to model and simulate a crowd because it allows modelers considerable freedom to incorporate various behavioral details into the crowd model. Hence, we operationalize our CGCS model for multi-agent systems using agent-based modeling method, where individual agents are capable of perceiving and responding to their immediate surroundings, and are organized into groups or 'co-travellers'. Group members maintain group cohesiveness by communicating and adapting their behaviors to each other while walking.

The CG model is framed as micro- and macro-coordination strategies, which allows the multi-agent system to adapt to varying domains. Macro-coordination strategies relate to overall action plans to accomplish the groups' goals, and is dependent on the domain (e.g., a group of soldiers may select a particular strategy influenced by training and doctrine, while a group of friends at a sporting activity may decide how to meet at a predetermined location). Micro-coordination strategies relate to the dynamics of human communicative behavior, and is determined to a degree by the constraints of human perception, the physics of sound in voice communication, and cultural concerns. Micro-coordination strategies are also influenced by the state of the immediate environment (e.g., voice communication may be suppressed in a movie theater, or visual displays and uptake may be constrained in a very dense crowd).

Our approach focuses on simulating the coordination process established by CG theory rather than modeling the process of social interaction at a lower level of communicative intent to reasoning. We believe that this intermediate level simulation is viable for two reasons. First, it is computationally more tractable than trying to simulate the micro processes of communication and negotiation of joint action itself. Second, CG theory is a well-vetted scientific theory, and already takes the effect of real human micro-level negotiation of joint activity into consideration. In fact, because the theory is well established, any bottom-up micro-negotiation model will be hard pressed to match the realism of a CG model.

1.4.3 Unified Framework for Crowd Simulation and Animation

While the CG theory primarily addresses the coordination process among participants in a conversation, CG maintenance is not solely a linguistic activity. Behavior may come into play whereby agents 'display their intentions' visually, and mutual behavior awareness plays an important role. That is, in our CGCS model, an agent may present gestures or other behavioral cues according to its communicative purpose in order to establish the state of common ground in the course of interaction. It also considers temporal and spatial factors of the situation in which the agent operates.

We provide a unified crowd modeling framework, where the simulation and rendering engines are tightly coupled so that the behavior selection information generated in the simulation model is used to produce more believable animation behaviors of characters in the rendering stage. Specifically, in a simulation engine, the behavior selection information of each agent is represented as a set of values for position, orientation, velocity, and kind of gestures and its activation time at each frame. This information is formed into an action script. The action script is transmitted to the rendering engine, and employed to produce animated characters. Since character behaviors in animation are governed by the underlying simulation model in which the dynamic social and environmental context of the agents is considered, a sequence of character behaviors is not merely a chain of reactive gestures, but form socially meaningful interactions. We believe that this will produce more realism both in the overall simulation and individual animations of the agents.

1.4.4 Crowd Model Evaluation

To examine the influence of our CGCS model in crowd simulations, we performed several different experiments. First, we showed that communicative behavior within individual groups can impact on the distribution of the simulated crowd as a whole. Specifically, the dynamic congestion distribution was measured for simulated data where we varied whether our CGCS model was included in the simulation. From the comparison of the resulted simulation data, it was shown that notable differences on the dynamic congestion level and crowd circulation pattern were made depending on the extent to the group communicative behaviors within a crowd.

Second, we conducted a series of perceptual user studies to determine if our CGCS model enhances the overall plausibility of crowd scenes. In narrative psychology, Bruner advanced an idea that people understand and interpret intentional behaviors by taking them into narrative structures (Bruner, 1991). That is, one's belief concerning the truth of a phenomenon is dependent on one's ability to explain and rationalize the phenomenon. We employed this concept to test our operationalization of the CGCS model by determining whether the model yields purposive interpretations, hence believability, of the resulting animation by observers.

Third, we performed a scalability comparison among several simulation models to prove the computational efficiency of our approach.

1.5 Contributions

This dissertation provides seven contributions related to crowd modeling, simulation, visualization, and evaluation.

Crowd Model and Operationalization

1. We advanced a model of social group interactions for crowd simulation. Because our model operationalizes a well-vetted Common Ground theory for the model of human interaction, it inherits the social realism provided by the theory.

2. Our formulation for the model of social group behavior consists of macro- and micro-coordination strategies. The two-level approach allows a flexible adaptation of the model into different application domains. Users can selectively override or specialize the rule for joint goal adaptation and constraints by their needs for macro-coordination strategy. The set of human communicative actions remains consistent across application domains, therefore the component of micro-coordination strategy is reusable.

3. We operationalized the conceptual model of social group behaviors for a multi-agent system. The operationalization is computationally efficient and can be used in large scale crowd simulations.

Unified Framework for Crowd Simulation and Animation

1. We provided a unified crowd simulation framework to produce socially plausible behaviors of animated characters. In our framework, the motion selection of animated character derives from the underlying simulation model, where the dynamic social and environmental context in which the agent functions is accounted. The behavioral and communicative purpose of agent is carried through from simulation to animation consistently, and allows purposive interpretations of the resulting animation for human observers.

Crowd Model Evaluation

1. By showing the impact of social group interaction model on the dynamic congestion distribution in simulations, we demonstrated the importance of considering communicative behavior within individual groups in large crowd simulations.

2. Through a series of perceptual user evaluation studies, we demonstrated that the believability of animation is affected by communicative and social interactions among characters.

3. By revealing that the comprehensibility is essential to believable agents, our study results informed that the model of character behaviors should be designed to provide interpretability, hence rationality, to external observers for achieving the enhanced realism.

1.6 Organization of Dissertation

The remainder of this dissertation is organized as follows. Before discussing our CGCS model and its evaluation, we review related research on crowd simulation and animation in Chapter 2. In Chapter 3 we present our preparatory work on simulating a large number of small groups using continuum-dynamics based approach. In Chapter 4, our CGCS model is described. The conceptual model established by information from CG theory is first presented, and details on its operationalization for multi-agent system is followed. Chapter 5 describes our unified crowd modeling framework for crowd simulation and animation. Chapter 6 presents a series of studies to demonstrate the impact of our CGCS model on crowd behaviors in a simulation, and validate the simulation results. We draw conclusions and discuss possible future research directions in Chapter 7.

Chapter 2

Background and Related Work

2.1 Overview

In this chapter, we provide a review of the literature that pertains to crowd modeling and simulation. Existing modeling approaches to crowd simulations may be categorized into two classes, including the agent-based modeling approach and fluid dynamics-based approach. Agent-based modeling approach has been often used to simulate small-to-medium scale crowds with some various behavior controls on the simulation entities (Section 2.2), while the fluid dynamics-based approach has been mostly employed to generate large scale crowd simulations with a computational efficiency (Section 2.3). Section 2.4 reviews several techniques on global path planning, which is a necessary component to build an autonomous agent in pedestrian simulations. Section 2.5 provides reviews on character and crowd animation techniques. Section 2.6 discusses some of efforts that have been put to crowd model evaluation and validation.

2.2 Small-to-Medium Scale Crowd Simulation

2.2.1 Agent-based Modeling

A commonly used approach to simulate the locomotion of a number of entities is the agent-based method. In earlier work using agent-based modeling, individuals have been often modeled as a set of homogeneous entities driven by physics-based rules. In the seminal work of Reynolds on boids, simple physics-based rules for local collision avoidance, cohesion, separation, and alignment are used to produce visually compelling flocking behaviors of birds or fishes (Reynolds, 1987). Later he extended the boids model to simulate more complex steering behaviors by supporting additional set of rules (Reynolds, 1999). Since Reynolds work, there have been many other approaches introducing various physics-based rules to agent-based modeling. These include Helbing's social force model and its extensions, in which repulsion and tangential forces are formulated to describe the pedestrians movements (Helbing and Molnar, 1995; Helbing et al., 2001; Pelechano et al., 2007). Brogan and Hodgins have used particle dynamics for modeling the motion of groups (Brogan and Hodgins, 1997). Velocity obstacle models have formulated the local navigation with collision avoidance as a geometric constraint problem (Fiorini and Shillert, 1998; van den Berg et al., 2008a; Guy et al., 2009, 2011).

In many agent-based methods, an agent is modeled as an intelligent and autonomous entity with cognitive and reasoning capability. Funge et al. (Funge et al., 1999) have equipped an agent with the ability not only to respond to the environmental stimulus but also to learn and make decisions with the perceived information. This model was later extended by Shao et al. to build an autonomous pedestrian with path planning (Shao and Terzopoulos, 2007). To address the challenge in modeling the human decision making process in the development of intelligent agents, a BDI (Belief-Desire-Intention) model (Rao and Georgeff, 1995) has been often integrated into an agent-based model (Cho et al., 2008; Koh and Zhou, 2011; Zuckerman et al., 2012). In addition to that, many other attempts were made to extend the agent-based model by addressing the effects of cognitive (Kapadia et al., 2009), psychological (Pelechano, 2005; Hoogendoorn and Soumokil, 2010; Yeh et al., 2008), emotional (Luo et al., 2010), and cultural (Tsai et al., 2011; Endrass et al., 2011) factors on behaviors to achieve various human-like behavioral details in crowd simulations.

Agent based approaches can be thought of as being group-size-agnostic. They can be applied to small or large groups, and the behavior of the entire group is an emergent phenomena that arises from the interaction of the individuals. However, in many agent-based approaches to crowd simulation, all agents are homogeneous (i.e., they do not model small group membership), in effect making the crowd the interaction of isolated individuals alone. Furthermore, agent-based models extended to support more complex and deliberative behaviors cannot be used to simulate large crowds because of the computational cost of such simulations.

2.2.2 Social Group Modeling

To provide the impression of pedestrian sub-group structures in simulations, some approaches introduced parametric models to represent various group interrelationships and simulate cohesive movements of agents according to given group characteristics (Musse and Thalmann, 1997; Qiu and Hu, 2010). Lee et al. presented data-driven methods to simulating various interaction patterns in groups (Lee et al., 2007).

In some studies, walking patterns and spatial organizations of small groups are analyzed from collected video recordings (Peters and Ennis, 2009; Moussaïd et al., 2010; Lemercier et al., 2012). The observed formations are represented as reference templates in a local coordinate system of a group and used to guide each group member's relative position (Peters and Ennis, 2009). Numerical models to include such formation influences are proposed in (Moussaïd et al., 2010; Lemercier et al., 2012; Qiu and Hu, 2010). Maintaining the desired formations while walking is formulated as a collective optimization problem for group members in (Karamouzas and Overmars, 2012).

A behavioral aspect in group dynamics is also addressed by several reserachers. The effect of actions and gestures, such as interacting distance, orientation, and synchrony of visual and aural cues of actions, on the plausibility of conversing groups have been identified and applied to the simulation of groups of conversing characters in (Ennis et al., 2011; Ennis and O'Sullivan, 2012).

Other researchers have provided methods to control the group behaviors at the global level.

Takahashi et al. presented an approach to manipulate group formation using the spectral graph theory (Takahashi et al., 2009). A path planning technique for coherent group navigation is presented in (Kamphuis and Overmars, 2004).

Recently several researchers have incorporated theories from social science into a crowd model for bringing realistic group dynamics. They have been mostly focused on developing applications for evacuation scenarios. For example, Kaminka and other researchers have proposed a cognitive model based on Social Comparison Theory, which was proposed by Festinger (Festinger, 1954), for simulating crowd behaviors in evacuation applications (Kaminka and Fridman, 2006; Tsai et al., 2011; Fridman et al., 2012). Lou et al. have applied Social Attachment Theory (Bowlby, 1982) into their agent model for emergency evacuation simulations (Luo et al., 2008). Kim et al. proposed a model for simulating pedestrian flow by using the concept of pedestrian mental stress based on Adaptation Syndrome Theory (Kim et al., 2012b).

2.3 Large Scale Crowd Simulation

One common approach to the simulation of large crowds involves the application of fluid dynamics. In the fluid dynamics-based approaches, the movement of a crowd is typically described by some partial differential equations. These approaches are computationally efficient because they treat individuals in a crowd as particles in a single flow field that can be computed over the simulation space together.

Kachroo has presented a mathematical model on conservation laws for governing multi-directional flow of pedestrians. He demonstrated the efficacy and flexibility of the model by applying it into a range of control models for various simulations (Kachroo, 2008, 2009).

Hughes proposed a crowd model using continuum dynamics (CD) (Hughes, 2002) and Treuille et al. brought the model into a discretized particle representation (Treuille et al., 2006). Crowds are represented as density fields, and individuals in a crowd are guided toward goal by an evolving potential function on the density field. Narain et al. showed that CD based approach is viable to simulate densely packed crowds (Narain et al., 2009).

Yersin et al. advanced a fine-coarse approach to boost computational efficiency at the expense of 'simulation correctness' (Yersin et al., 2008). Observing that for such applications as games, the virtual camera focuses in only on certain area of interest, they divided a scene into multiple regions of varying interest. The full CD computation is performed only for the most interesting part that is seen by the camera. They employed simpler and faster solutions that do not handle such details as collision avoidance in the low-interest areas to support real time simulation. Local reaction is calculated with additional computation process. Continuity maintenance is needed to provide smooth trajectories of agents when they transition across simulation zones.

To the degree that humans are physical entities, and behave like particles, e.g., forming vortices as they move through an aperture (Treuille et al., 2006), these flow-based models have been able to account for certain crowd characteristics. However, humans are more than particles, and physics-only flow models are incapable for accounting for human volitional behavior such as individual interest, social activity maintenance, and individual and group preferences.

Another limitation is that, the CD approach does not scale well with respect to the number of groups with different goals and preferences. The CD algorithm needs to be run for each group separately, and we have to account for the interaction across groups. For example, if 100% of 1000 people are in pairs, 500 CD field simulations for the 500 pairs are required. At each time step, each of these 500 simulations would not know the result of the other 499 simulations. Hence for each step, all simulations would have to report their results and receive results from other simulations. Each of the 500 continuum dynamics simulations would have to be rerun for each time step.

2.4 Path Planning

Modeling of reactive behaviors alone cannot create realistic crowds where each agent needs to steer towards its goal. Several global planning approaches have been proposed to solve this problem. A popular method to navigate agents to their goal positions is the A* algorithm (Hart et al., 1968), and several derivatives are also proposed (Stentz and Mellon, 1993; Sun et al., 2008; Trovato and Dorst, 2002).

The roadmap approach represents the connectivity of the free configuration space (Shao and et al., 2005; van den Berg et al., 2008b) that may be searched using different algorithms. In the various 'potential fields' approaches, a global field is used to represent the entire landscape with obstacles, and crowd elements move under the influence of these fields (Arkin, 1990; Chenney, 2004; Jin et al., 2008; Warren, 1989).

2.5 Character and Crowd Animation

Animation is to visualize the results of a crowd simulation. There have been extensive research on generating realistic character animations. In the research field of embodied conversational agents (ECAs), creating believable non-verbal conversational behaviors of animated characters has been studied (Cassell et al., 2001; Jan et al., 2007; Pelachaud, 2005; Stone et al., 2004). ECAs combine synthesized speech with hand gestures, head movements and facial expressions to create believable social avatars in conversational setting. However, most of these works do not address dynamic movement and positioning of characters, hence restrict the scope of applications.

Numerous approaches to motion synthesis and editing have been proposed to provide physically realistic motions for virtual characters (Li et al., 2002; Liu et al., 2006). To simulate sophisticated character motions, character behavior adaptation techniques accounting for the environmental objects and other characters have been proposed in (Lau and Kuffner, 2005; Pettré et al., 2003; Thomas and Donikian, 2000; Wampler et al., 2010). The main limitation of these approaches is that tremendous computational cost for searching suitable and smooth transition points from one pose to the next at each frame limits the application of the methods to large scale crowd simulations.

Rendering 3D fully articulated virtual characters for a natural looking and large scale crowd simulation is challenging work. To increase the motion variety of a crowd, Gu and his colleagues provided a method to control motion styles of agents dynamically (Gu and Deng, 2011). In their method, primitive motions are extracted and stylized based on their kinetic energy characteristics first. At runtime, the motion style of each agent at every time step is dynamically selected using the information of its local neighbors, current style, and the global style distribution.

Some work proposed a way to create a dense crowd of virtual humans with expressive gestures (Lai et al., 2005; Lee et al., 2006; Kim et al., 2012a; Ju et al., 2010). They collected motion episodes of multiple characters in which the characters interact with each other for a short period of time. A crowd of interacting characters for an extended period of time is generated by weaving the episodes. The control of individual motion is not allowed, and that limits the application of the approaches. Moreover, the motion selection is not tightly coupled to the underlying simulation model, thus the sequence of character gestures does not form socially meaningful interactions.

2.6 Evaluation of Crowd Simulation

Crowd model evaluation is one of the most challenging issues in crowd modeling and simulation. We can classify existing approaches to model evaluation into three categories: human data-based analysis, statistical analysis, and human perception evaluation. The detailed discussions on each approach follow.

2.6.1 Human Data-based Analysis

The most fundamental approach to evaluating crowd models may be to compare the simulation result of the model with the real world crowd data. In the work of Kretz, a crowd model based on cellular automata has been introduced. They applied their model to the simulation of evacuation scenario, and compared the results of the simulation (e.g., egress rate, time) against human data collected from a controlled experimental environment (Kretz et al., 2008). In other work, a real pedestrian movement around an area of interest was captured by cameras, and the data was used to validate the simulated crowd movement by visual analysis (Seyfried et al., 2007; Guy et al., 2010; Karamouzas and Overmars, 2012).

2.6.2 Statistical Analysis

In many studies, comparisons between different crowd models have been performed on the collected data of travel distance, time, and trajectory of agents (Brogan and Johnson, 2003; Pettré et al., 2009; Guy et al., 2012). To analyze the impact of cultural differences on crowd dynamics in evacuation situations, data including evacuation time, level of fear, group connectivity, and mean speed of agents on varying simulations settings have been collected and compared (Fridman et al., 2012).

As an effort to provide a standard way to evaluate a crowd simulation, Kapadia and his collegues proposed a statistical method to characterize coverage, quality, and failure of steering algorithms of a given crowd model (Kapadia et al., 2011). For evaluating pedestrian steering behaviors, a benchmark suite that consists of 38 testing scenarios with different level of complexity has been proposed in (Singh et al., 2009).

2.6.3 Human Perception of Crowds

In recent years, human perceptual evaluation of virtual characters and associated animations has received attention as an important means to develop and validate the crowd model. Much research has been conducted regarding the perception of animation and motion of individual characters. For example, McDonnell et al. have investigated human subjects' perception of variety in crowds, and found that people are more sensitive to appearance clones than to motion clones (McDonnell et al., 2008). They also have identified that the head and upper torso of characters attract the majority of first fixation of human observers. By exploiting this knowledge, they proposed a selective variation method to generate realistic varied crowd animation (McDonnell et al., 2009).

In relating to crowd simulation, several researchers have examined the effects of position and orientation on the plausibility of pedestrian formations in static crowd scene (Peters et al., 2008; Ennis et al., 2008). They found that people consistently discern between the real and artificial scenes depends on the context of the scene (e.g., open location, corridor location), and are more sensitive to the positioning of characters in a constrained zone. The study has been further explored the effects of position, orientation, and camera viewpoint in combination on the plausibility of pedestrian formations (Ennis et al., 2011).

It has been also found that adding plausible groups, where appropriate proportions of single pedestrians, pairs, and trios are distributed, to a pedestrian crowd scene plays an important role for an increased sense of realism in crowd scens (Peters and Ennis, 2009).

Regarding the behavioral fidelity, Ennis et al. have investigated the importance of matching audio to body motion in virtual conversing character animations through a series of perceptual experiments (Ennis et al., 2010). It is shown that people are more sensitive to visual desynchronization of body motions, than to mismatches between the characters gestures and their voices from their results. Jarabo et al. have explored how important accurate illumination is for the perceived fidelity of dynamic crowd scenes, and showed that errors in illumination could be masked by the aggregate characteristics of crowds (Jarabo et al., 2012).

Chapter 3

Preparatory Study

3.1 Introduction

A preparatory study we performed prior to developing the CGCS model is presented in this chapter. In this study, we proposed an efficient method using continuum-dynamics to simulate a crowd which is composed of a large number of small heterogeneous groups.

Studies of spatial factors in small groups have informed us that small interacting groups occupy an interaction space that is conceived of as bounded and protected by others (hence *social territory*) (Knowles, 1973; Knowles et al., 1976). Non-group members are responsive to a group's social territory and avoid walking through it. Another aspect to be considered in the walking process is the local perception of the environment. People respond only to the immediate surroundings but not to the entire world. It is only when they perceive direct evidence of a threat or opportunity that they react.

These observations inspired our approach where the behavior of small groups is modeled effectively and efficiently by having group members interact and react only within some locally defined region. We call the locally defined space *Local Interaction Field* (LIF). For each group, the LIF is a distinctive space that reflects the immediate space-time information of the surroundings of the group. Our approach is computationally efficient because a group only deals with its LIF to plan the navigation. Group members plan their immediate paths (actions) independently using the LIF information and shared goals. Hence, there is variation in individual paths as group cohesion is maintained through shared goals and LIF.

3.2 Overview and Notation

In this section, we provide the system overview and introduce the notation. The term 'group' is used here in its sociological sense. A group in our approach is formulated to provide social ties, such as friends or family members, among agents.

3.2.1 Simulation Overview

Our crowd model for group-based movement uses a two-level approach. At a higher level, a global path information is generated to guide a group into a collision free route among static obstacles (section 3.3.1). At a lower level, a local behavior planning for cohesive group movement and reactive steering is derived (section 3.3.2).

Figure 3.1 shows a main flow of simulation. In a precomputed step, a global path is planned per group. At run time, all agent behavior are generated using a continuum flow approach.

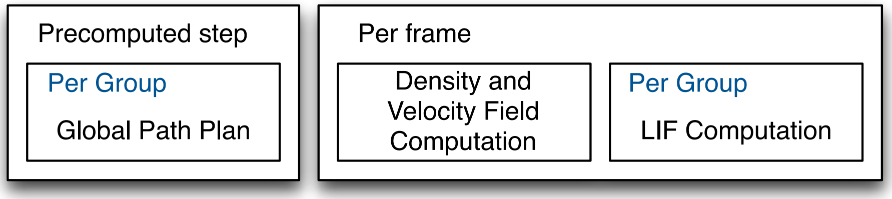


Figure 3.1: An overview of our simulation

At each frame, density and velocity fields are updated from the simulation result of the previous time step. The group's LIF is updated using crowd density and velocity information from the global crowd representation, and a CD field is computed using this LIF. Group members use this LIF-based CD field to compute their own paths and velocities.

3.2.2 Crowd, Groups, and Individuals

We handle individuals in our simulation by permitting groups with a single member. Suppose we have small groups of $[1, 2, ..., k]$ members with a mean of $\bar{s}$. Crowd simulation would have M groups or equivalent $M \times \bar{s}$ individuals. Each agent is assigned its group membership G_i. All members of group G_i have the common goal g_i and LIF$_i$. A leader agent l_i is selected per group.

3.2.3 Discrete Crowds

We represent a crowd in terms of density and flow velocity on a regular grid, as is typical for modeling pedestrian flow in CD-based crowd simulations (Hughes, 2002; Treuille et al., 2006; Narain et al., 2009). To aid comprehension, we summarize the representation and our adaptations of it here. In the standard CD approach, each node occupies a region that may contain several individuals, and this occupancy is used to determine the density value at that node. From this a velocity field value is computed for each node. We scale our grid such that each node corresponds to the occupancy of a single individual. This obviates the need to handle multiple agents occupying the same grid node. We address the computation of the node density ρ_n and field velocity $\bar{v}_n$ at (x_n, y_n) as a sums of 'influences' of nodes in its neighborhood:

$$\rho_n = \Sigma \rho_m \;\text{ and }\; \bar{v}_n = \frac{\Sigma \rho_m v_m}{\rho(n)}.$$

The influence of each node m at (x_m, y_m) on its neighboring cells at (x_n, y_n) is given by the 2D Gaussian function:

$$\rho_m = \Omega_m K e^{-(\frac{(x_n - x_m)^2}{2\sigma_x^2} + \frac{(y_n - y_m)^2}{2\sigma_y^2})}$$

where K is the amplitude of the Gaussian, and σ_x, σ_y are the x and y spreads of the influence function. K, σ_x and σ_y can be controlled depending on a cell size. Ω_m is the binary occupancy of the node with a value of 1 if an agent is present, and 0 otherwise. For this study, we used $K = 3, \sigma_x = \sigma_y = 2$ with a cell size 1.

ρ and $\bar{v}$ on a grid is then used to build each group's LIF as detailed in Section 3.3.2.

3.3 Group Based Navigation

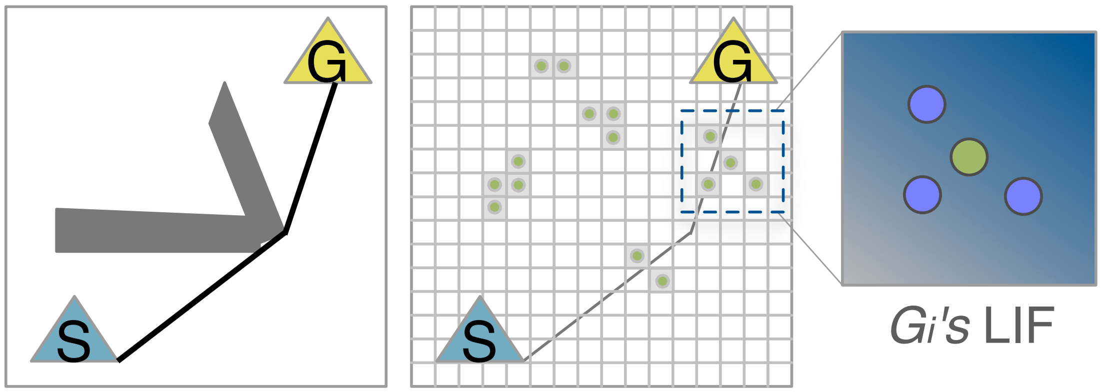
3.3.1 Global Path Planning

To allow groups to navigate in a complex virtual environment, our approach employes a global path planner. Since it is used to determine only a preferred direction to proceed, a global path computed using coarse connectivity information is sufficient. Any type of global path planner can be used as long it generates a collision-free route. In our system, a virtual simulation space containing static obstacles is represented as a roadmap in which overall connectivity information is represented as a graph.

Given a goal position g_i, the global path is extracted using A* algorithm on the roadmap. Figure 3.2(a) shows a global path represented by line segments between start and goal positions, which are marked with blue and yellow triangles respectively.

3.3.2 Local Planning

Along the guidance of global path, the space opens up to a local grid for each group. We call this grid the *local interaction field (LIF)*. This space is local since a group has to be concerned only with the navigation to the edge of the grid where the grid then is blended with the global path of the higher level as illustrated in Figure 3.2(b). Each cell value in the LIF represents the traversability of the cell, and an agent in the field selects its velocity which will bring it to a region of good traversability.



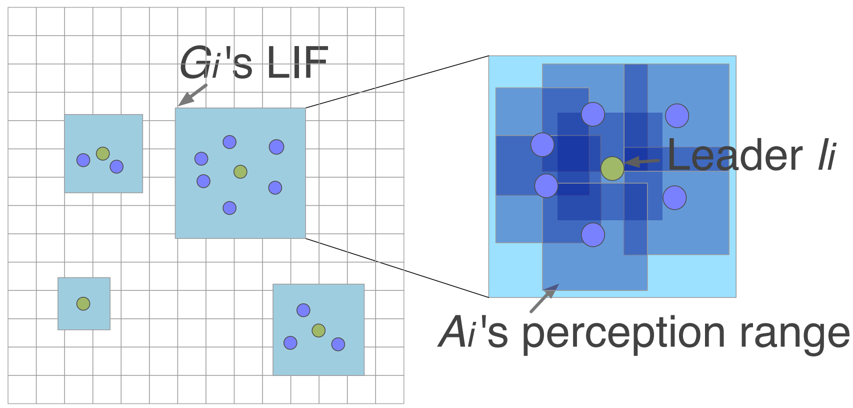
(a) (b)

Figure 3.2: (a) Global path (b) LIF combined with a global path

Local interaction field

As a group traverses, the group's LIF is carried along with it to reflect the immediate surroundings. The LIF is centered at the group leader's position though the center of the LIF can be varied, e.g., to be at the center of mass of the group, if desired. Each member contributes to fill the LIF of its group by collecting the traversability within its perceivable range. We assume all the agent have the same omnidirectional perception range to collect information. If some cells in the LIF is covered by several agents, only one measurement is taken to prevent under-estimation of traversability. The effect of each agent's perceptual field on the group LIF is illustrated in Figure 3.3(b). The perception field of each agent is shown as a dark blue square centered at the agent. Areas covered by more than one agent is shown in darker blue. Regions in the LIF that has not been seen by any agent is in light blue. Since those areas have not been seen, they are treated as impassable.

The size of the LIF depends on the number of group members as is illustrated in Figure 3.3(a). The reasons for this variation in LIF size are twofold. 1. It is reported that the extension of the boundary around a group increases with the size of the group and is reflected in the actual behavior of pedestrians (Knowles et al., 1976). 2. We can relate the size of local field to the collective perceptive range of those agents in a group. This design choice (representing 'group perception' as a single shared LIF) is a trade-off between computational concerns and conceptual correctness. Although in reality, a group member does not see everything that other members of the group sees, it can nonetheless be argued that the group does have a cohesive shared knowledge that members of other groups do not share. Employing a single shared LIF avoids the overhead of representing how different knowledge may pass among members of a group, and supports reasonable realism in the simulation outcome.



(a) (b)

Figure 3.3: (a) Various sizes of LIF (b) Each agent contributes to fill the LIF

Traversabiltiy

For each agent, the traversabiltiy is measured by taking into account space-time factors including isolation from a group, discomfort arising by the proximity of non-member agents, path length, estimated travel time, and static objects.

Compared to solitary individuals, agents who belong to a social group tend to stay together throughout the navigation. A group aggregation factor for the cohesive group movement is introduced. A center of mass μ for a group G_i is calculated and serves as a reference point to check the level of isolation from the group. As an agent goes far away from μ, it costs more on the path from its current position x to a cell at p.

The discomfort $\mathcal{D}$ for an agent A_i of G_i needs to be computed for its anticipated position after some number of time steps τ (τ depends on the maximum speed of agents in the simulation). Hence, we need to compute a predicted density grid depending on the anticipated locations of all other agents $A_j, j \neq i$. For computational efficiency, an agent has to consider only changes in its perceptual range. Hence for each agent A_i, we do the prediction only for all non-member agents A_j within A_i's perception range (we do not consider members of group G_i because each agent occupies a single grid, and we assume that within-group agents can occupy adjacent nodes with no discomfort). For each A_j we compute its predicted cell location (x_j, y_j) and update the influence function ρ_j. The discomfort for agent A_i is then computed as: $\mathcal{D}_i = \Sigma_{t=k}^{k+\tau} \Sigma_j \rho_j$.

Isolation, discomfort, path length and travel time are linearly combined to constitute a traversability cost function. Then the occupancy of a cell for static objects is checked. The traversability cost $C_{x \rightarrow p}$ of a cell at p measured by an agent at x is given by the function:

$$C = \max(\underbrace{\alpha \int_{\mu}^{p} 1ds}_{Isolation} + \underbrace{\beta \int_{p} \mathcal{D}dt}_{Discomfort} + \underbrace{\gamma \int_{p} 1ds}_{Path\ length} + \underbrace{\epsilon \int_{p} 1dt}_{Travel\ time},\ \delta), \tag{3.1}$$

where α, β, γ and ϵ are weights for isolation, discomfort felt, path length, and travel time, respectively. The weights for the linear combination represent the preferences or tolerance of each group for the factors that make up the traversability cost function. δ is set to infinity for unavailable cells due to the static obstacles, otherwise set to 0.

A potential function ϕ to determine the optimal path given the cost C is the Eikonal equation: $C = \|\nabla\phi(x)\|$. $\phi = 0$ in a goal, and ϕ values for the rest of grid cell are approximated by solving the potential function. By definition, pedestrians prefer terrain with the low cost while moving towards a destination. Therefore agents move in the direction perpendicular to the estimated cost. Velocity v of an agent is:

$$v = -\bar{v} \cdot \frac{\nabla\phi(x)}{\nabla\|\phi(x)\|}.$$

Local goal

We use a large scale of high resolution environment and the LIF covers only small portion of it. Therefore the goal position g_i is likely to be outside the LIF_i for group G_i. To resolve this problem, a group has to identify a local goal within its LIF area. Each group leader locates the local goal for its group in a process illustrated in Figure 3.4. We call the connection point on the global path a *waypoint*. The current waypoint W_i is set to the nearest waypoint towards g_i that lies on the global path and is outside LIF. The local goal is found as the intersection of the projector L_1 from the leader's position to W_i and the appropriate edge, L_2 of the LIF.

An advantage of our approach is that it only solves the potential function on a small fixed-size LIF grid for each group, the performance is not subject to the resolution and size of the environment.

3.3.3 Group Behaviors

In our model, group members plan their paths independently. Group cohesion emerges because members of a group share a LIF and a joint goal, differing only in their individual starting locations.

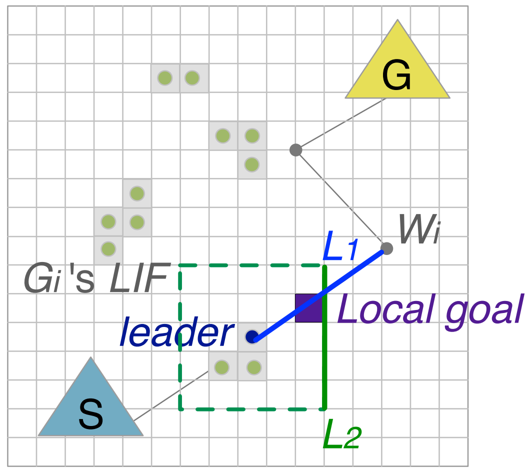


Figure 3.4: A local goal is identified using the leader's position and a waypoint

Cohesion recovery

It is possible that some of agents fall outside its group's LIF. For example, local obstacles or interference by members of other groups may force one member of G_i outside LIF$_i$. If an agent is determined to be out of the boundary of its LIF, a leader notices this and modifies the group's local goal towards where the lost agent is. This results in efficiency advantages over other crowd simulation methods, most of which need to compute the distance of every member to the rest of the group to detect lost agents.

Social space

The phenomenon of bounded social territory is an emergent effect of our approach. Since the approach requires that members of a group remain in its LIF, this increases the density of the area occupied by the group. Members of some group G_i will have a built-in preference to be in LIF$_i$. Members of some other group $G_j, j \neq i$ do not have any preference to be in LIF$_i$, and will in fact avoid the area because of its increased density. Hence non-G_i group members will avoid the social territory of G_i even though they are not privy to the existence or location of the LIF per se. The degree to which the social space boundary protects the group from intrusion can vary as a function of the size of the group, the location of its members, and the distance maintained between group members. Interestingly, if a group is scattered or not small, it might actually incur intrusions as agents not in the group evaluate the cost of going through the dense area to be lower than the cost of going around it. We note that this also happens in the real world.

Path preference

If dynamic factors such as crowd density influence path choice, pre-determined global path can be updated as needed. If a waypoint is determined to be in a congested area using the density field information, that area may be marked as blocked on the global path. A group then relaunch the global path planner with the updated information. This is more similar to how human react to the world while navigating; it is impossible to know all the condition of the world in advance as it was assumed in the original CD approach. Note that members of different groups may also have different path preferences. For example, members of some groups may have greater aversion to congestion by having a high β value in Equation 3.1, while members of others may plan for the shortest path to the goal at the expense of other valuations for traversability, such as tolerance for going through a congested area.

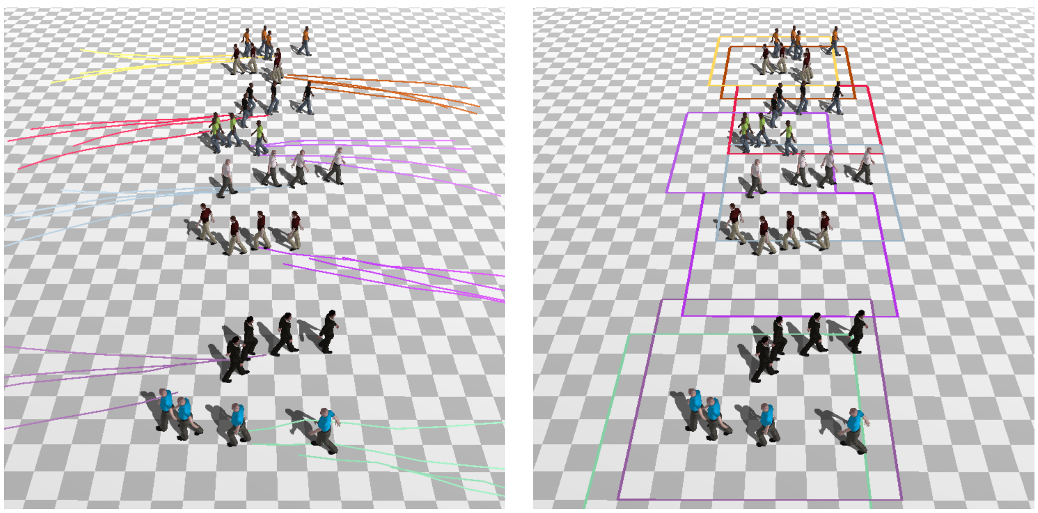
3.4 Results and Discussion

3.4.1 Simulation evaluation

Group based movement

Figure 3.5: A crowd of 1,156 groups is simulated on a $1,024 \times 1,024$ grid

We create a simulation of a crowd of 1,156 groups, made up of 300 individuals, 510 groups of 2 individuals, 240 groups of 3, 50 groups of 4, 30 groups of 5, 22 groups of 10, and 4 groups of 32 individuals to evaluate the behavioral realism of our simulation approach. This configuration approximately follows the distribution of pedestrians reported in (James, 1953). This simulation is challenging because the departure and destination points of the groups are randomly assigned, and the groups may cross each other at many different angles. Figure 3.5 shows that agents belonging to the same group stay and move together. A $1,024 \times 1,024$ grid is used in this example, the ground checker board shows the resolution of the simulation grid, which is agent-sized.



(a) (b)

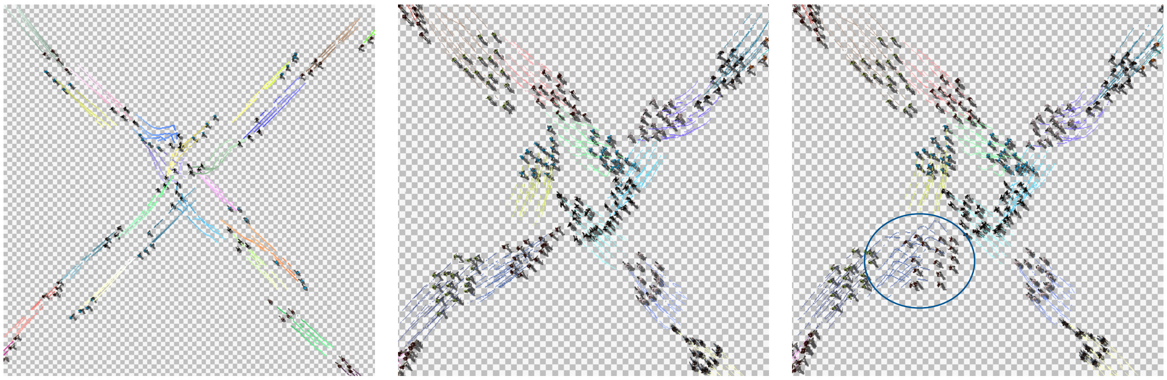
Figure 3.6: (a) and (b) 8 groups move as forming lanes

Figure 3.6 (a) shows a simulation of eight groups crossing to reach opposite sites, and naturally forming lanes. Smooth motions as shown with trajectories of the agents demonstrate that maintaing the potential field only at the local regions to the group does not result in path and motion discontinuities. Color-coded LIFs of the groups are shown in Figure 3.6 (b).

In Figure 3.7 (a), we show a simulation of 64 groups of 4 agents. 16 groups start from each corner and navigate to the opposite corner of the environment. Group members form a social territory as walking in tighter physical configurations, and prefer not to move through the area occupied by agents of other groups. At the same time, a vortex and lanes emerge as the groups cross paths.

Path update

Figure 3.7 (b) and (c) show the results for the different behavior patterns of groups. In Figure 3.7 (b), all groups prefer the shortest path by setting α to a high value in Equation 3.1, therefore they surge into the congested center area. In Figure 3.7 (c), a group marked with a blue circle is set to avoid congestion, and agents of the groups circumnavigate the crowd



(a) (b) (c)

Figure 3.7: (a) and (b) 64 groups move as forming a vortex. (c) A group marked with a blue circle circumnavigate the crowd.

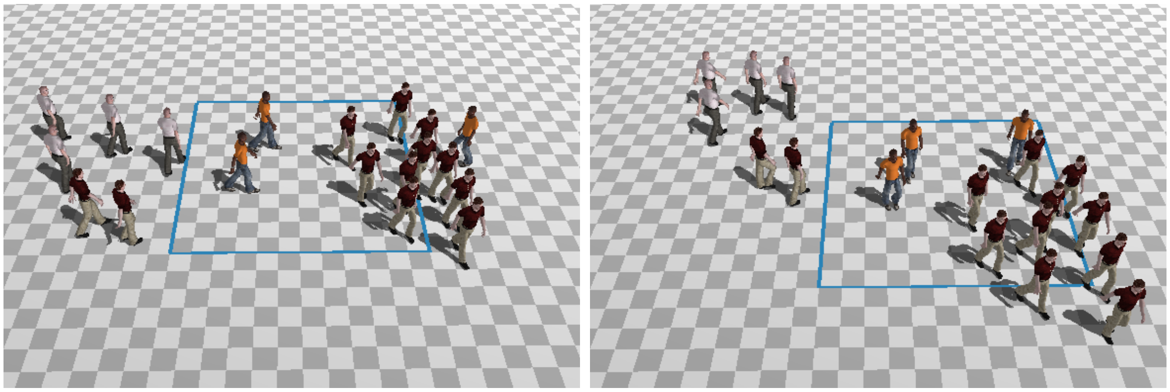
while still moving toward its goal.

Cohesion recovery

Figure 3.8 shows how a group that has lost cohesion may recover. An individual of group composed of 4 people is left outside due to the interference by members of other groups as in Figure 3.8 (a). Bright green cells represent the area occupied by agents of other groups. Since agents can only move in regions that they have seen (i.e., is in the group's LIF), the agent outside the LIF will stop moving. The group leader notices that an agent has left the LIF and moves the group to recover the lost agent. After recovering group cohesion as in Figure 3.8 (b), the group resumes its original movement toward its destination.

3.4.2 Performance Comparison

Our simulation system was built for parallel computation on NVIDIA's CUDA programming framework (Nickolls et al., 2008). The simulation was tested on a desktop with an Intel i7 3.20 GHz CPU, 4GB system memory, and a NVIDIA GTX 580 graphics card. We use OpenGL-based rendering to visualize the animated characters driven by our simulation engine. Each character associated with a simulated dot updates its 3D poses according to the per-frame simulation result. To improve rendering quality, shading effects with shadow mapping was used in our demonstration.



(a) (b)

Figure 3.8: (a) An individual of a group is left outside the LIF (b) The group moves to the lost agent to recover the cohesion

Parallel implementation

CPU-based computation is used to set up initial and goal positions of the agents, and the topographical information of the environment. The global path planning for each group is performed on the CPU. All the information is passed to the GPU, and a number of CUDA *threads* are created to compute the density and velocity fields in parallel. Each thread takes n cells on the simulation grid where n is determined by the width and height of the simulation grid, and the number of total threads available in the GPU. Once the global environment information is computed, the simulation is partitioned by agent group and *blocks* of threads are assigned to build the LIF, solve the related potential function and update the positions of agents.

Large numbers of small groups

We evaluated the scalability of our approach with respect to the number of groups. We ran a comparison of our method and the original CD approach of Treuille et al. Treuille et al. (2006). The same base code was used for both implementations with one exception; the original CD had been set to use the entire environment grid for the potential function computation while our method computed potential fields for the LIF of each group. Six different group configurations were tested on a 256×256 grid. A relatively small environment grid had to be used because the original CD cannot handle larger grids for the potential field computation. For our approach, a 8×8 grid is used as a basic perception range for an agent in the experiment. Then the LIF size is determined dynamically by the number of members and their distribution. We augmented the heuristic approach for the decision of LIF size. For example, a 8×8 grid of LIF is used for a group with a single member, a 16×16 grid is used for groups of 2-8 members, a 24×24 grid is used for groups of 9-32 members, etc. This model is ad-hoc, but efficient for the parallel implementation and optimization, and provides a visually plausible simulation for group dynamics.

We fixed the total number of agents to be 2,048 and varied crowd organization for our 6 group configurations. An extreme case is 2,048 agents of a single group, and the other end is 32 groups of 64 members per group. Again, the number of groups in this experiment is limited by the original CD, and 32 was the maximum number of groups that could run with the original CD approach in our target machine.

Table 3.1: The performance comparison of the original CD method and our approach.

Number of	CD		LIF	
groups	fps	time(ms)	fps	time(ms)
1	63	13.21	91	8.21
2	48	18.18	86	9.16
4	42	21.72	77	10.02
8	30	31.95	67	13.67
16	19	49.56	59	12.51
32	11	85.69	45	14.34

The results are shown in Table 3.1. For a single group of 2,048 agents, the original CD code achieves 63 fps while our LIF approach runs at up to 91 fps while concurrently rendering the simulation scene. The fps values in the table include both the computation and rendering time. The time measurement only includes the computation time. The performance of the original CD dropped in more rapid fashion than for our approach. At 32 groups of 64 members, the original CD performance falls to 11 fps (or $\approx 1/6$ of the single group performance), while for the LIF version the performance falls only to 45 fps (or $\approx 1/2$ of the single group performance). In both approaches, the evaluation of discomfort arising by the proximity of agents of other groups causes the performance degradation with respect to the number of groups.

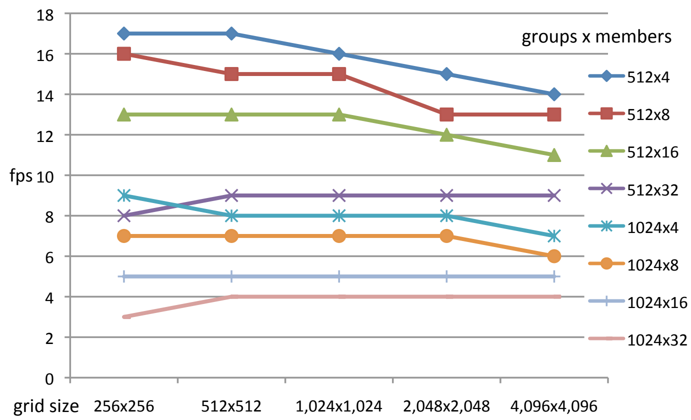


Figure 3.9: The performance measured in fps: scalability

Large scale environment simulation

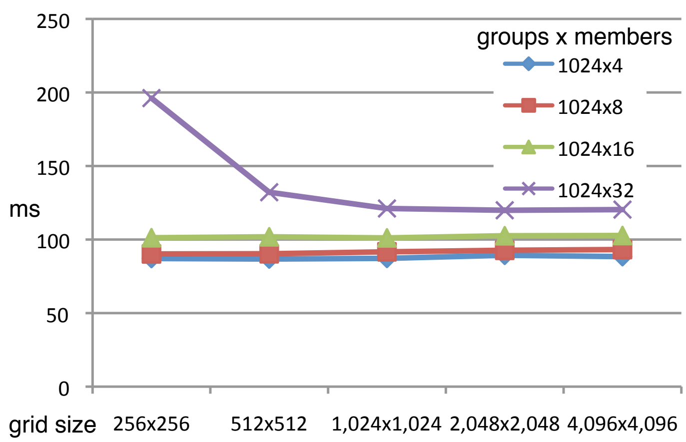


Figure 3.10: The performance measured in fps: LIF execution time

The results of our scalability test with respect to environment size follows. As shown in Figures 3.9, 3.10, and 3.11, various environment sizes were tested with various numbers of agents ranging from 2,048 to 32,768 that were arranged in two configurations. In the first configuration, the total agents were divided into 512 groups. The second set has 1,024 groups of agents. Taking 15 fps as the minimum interactive rate, Figure 3.9 shows that our model achieves this rate for configurations up to a $4,096 \times 4,096$ grid with 512 groups of 4 members. This running time includes a rendering time for the visualization.

The size of the LIF increases with respect to the number of members in a group, and the fps

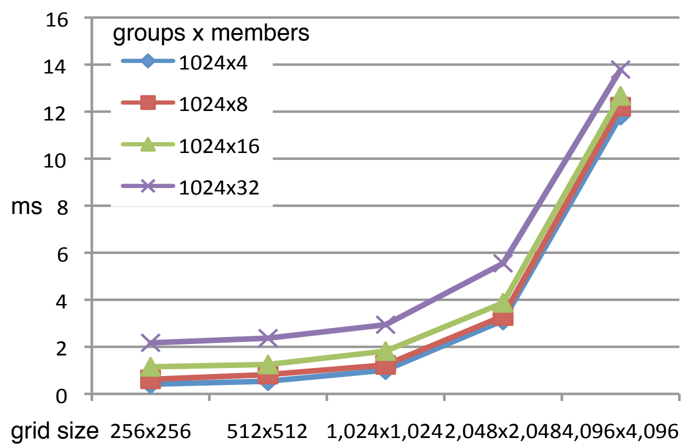


Figure 3.11: The performance measured in fps: Density and velocity fields execution time

values have an inverse relation to it. The size of LIF is fixed for those groups of the same number of members, time for the potential field computation is independent to the size of the environment. However, the execution time to populate the density and velocity values for the environment increases and becomes a bottleneck to the simulation, and is therefore the limit to scalability. Figures 3.10 and 3.11 clearly show that the LIF execution time remains constant while the execution time for the environment grid increases exponentially with respect to the size of grid with the simulations of 1,024 groups.

A slight increase in execution time was observed in the case of a very dense crowd simulation, e.g., $1,024 \times 32$ agents on a 256×256 grid. This is explained by our use of CUDA *atomic functions* in environment information computation. When a number of threads access the same cell for the accumulation of density or velocity value due to the several agents' being in a neighboring area, the summation is linearized to prevent data corruption. If many agents are confined to a small area, the probability of the agents being in the same vicinity increases, resulting in performance degradation due to the increase of atomic operation execution.

3.5 Summary

We presented an efficient approach to simulate many number of small groups which leverages the benefits of the continuum based crowd simulation method. By modeling space-time information of the surrounding environment of each group as a locally defined interaction space, group dynamics in a large crowd was effectively reproduced. In addition to that, an emergent group behavior of forming social territory, which can be observed in a real crowd, was produced. From the performance experiments, it was shown that our approach scales well with respect to the number of groups and the size of environment.

We designed LIF to be a collective perceptive range of those agents in a group. This design choice of representing 'group perception as a single shared LIF is a trade-off between computational concerns and conceptual correctness. In reality, a group member does not see everything that other members of the group sees, hence the approach is not effective if we aim to be cognitively correct for the modeling. However the main concern here was to achieve the behavioral realism with reasonable computational resource. Therefore, we can say that a cohesive group share a 'group consciousness' that other groups do not at some level of abstraction.

The main limitation of this approach was in that the control of each agent was not allowed because of the use of continuum dynamics-based model, and hence any human volitional behavior such as individual interest and social activity maintenance could not be generated. Specifically, an agent in our approach considers only physical factors including proximity to others, degree of density at a location, travel time and length for a give destination, to make its movement and group cohesion. However, when people walk in a group, their decision for movement are determined not only by physical constraints induced by other people around and the environment, but also by communicative interactions among the their accompanied members. This problem in lack of social functionality of agents motivated our research goal for modeling social group behaviors with coordination process.

Chapter 4

Modeling Social Group Interaction using Common Ground Theory

4.1 Introduction

In order to address the need for socially plausible aggregate behavior in a crowd simulation, we propose a model of social group interaction. We employ Herbert Clark's Common Ground (CG) theory (Clark, 1996) from social-psychology and linguistics, and operationalize the model for multi-agent systems.

In Clark's CG model, people in coordination must negotiate and maintain a state of CG as a precondition of joint activity. We view a task of group navigation as a kind of joint activities, hence group members have to maintain a state of CG throughout their navigation. Although the CG theory was established to explain the coordination process among conversing people, CG maintenance is not solely a linguistic activity. Behavior may come into play whereby participants display their intentions visually (e.g., using gestures), and mutual behavior awareness plays an important role. In the course of establishing CG regarding the group navigation plan, agents in our model may exchange behavioral cues with their group members.

In the rest of this chapter, we present how CG theory relates to group and crowd behaviors, the design of our multi-agent system for crowd simulations, the model of social group interaction, and implementation details to transform the conceptual CGCS model into a computational framework.

4.2 Common Ground Theory

In Clark's model, people engage in a joint activity when they act in coordination with others to pursue a common goal. The concept *common ground* is established as a mechanism by which participants involved in a joint activity coordinate their actions. It considers the mutual knowledge, beliefs, and assumptions among individuals in a collaborative process. According to Clark, p is common ground for members of group $\mathcal{G}$ if and only if (Clark, 1996):

1. *members of $\mathcal{G}$ know that p;*
2. *members of $\mathcal{G}$ know that members of $\mathcal{G}$ know that p;*
3. *members of $\mathcal{G}$ know that members of $\mathcal{G}$ know that members of $\mathcal{G}$ know that p.*

Suppose that A and B walk in an airport terminal. As they pass a schedule board, A thinks that they should check the departure flight information, and informs B of her plan to go to the board and to return to their current location, x. We denote the plan to divide and reunite at x as $\mathcal{P}$. For the plan to succeed, A needs to know that B knows the plan $\mathcal{P}$, and vice versa. This, however, is insufficient for coordination. B needs to know that A knows that he is privy to $\mathcal{P}$, otherwise he might not be convinced that A will return to x. Furthermore, if the agreement ends here, A may not know that B knows that she knows the plan, and may, therefore not be confident to execute the plan. Hence, A needs to know that B knows that she knows the plan.

The CG may be arrived at verbally, or may be enacted through action. For example, A may signal her intention by pointing toward the schedule board and pointing to their current location x. This requires that B be within the range of sight and be looking at A. A needs to see that B is looking at her, and has signaled agreement (e.g., by nodding). B needs to see that A sees his nodding. Finally A needs to see that B sees that she has seen and acknowledged the plan.

4.3 Multi-Agent System

We operationalize our model for multi-agent systems using agent-based modeling approach. In our approach, the multi-agent system is structured as a layered architecture. In the lower layer, an individual agent acquires understanding of its immediate surroundings or situation that includes the state of member of the group to which it belongs and the nearby environment. With the information the agent selects proper behaviors for successful coordination with group members. In the upper layer, group level attributes for different social groups are maintained.

4.3.1 Agent Model

An agent with its personal identifier i is denoted as A_i. In order to function with group members, an agent should be able to interpret status and intentional signals of the members and adapt its behaviors. The agent also needs to understand its surrounding situation and respond to the environment conditions. Hence, we model an agent as having sensory capabilities for speech, vision, and touch. Figure 4.1 shows an agent's sensory model with a perception geometry. In our model, touch can be sensed within range of agent radius r, hearing can be omnidirectional with range limitation d_h, and vision is directional and is effective up to a range, d_v, along its gaze direction (for simplicity, body orientation is synonymous to gaze direction in our simulation) and within a field of view defined by an angle, α. An agent can also sense the level of congestion in surrounding unit area. In a crowded area, d_h and d_v might decrease with respect to the level of congestion.

Our agent model is able to handle 'goal interrupts' through the introduction of new intermediate sub-goal determined either stochastically or through interaction with the

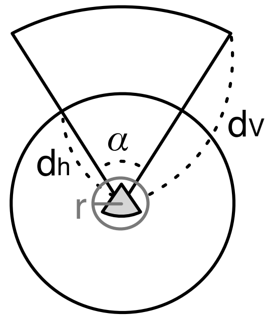


Figure 4.1: Agent perception geometry

environment. For example, a members of groups may have to visit the restroom (stochastically generated sub-goal), or a member of a group may see a store that sells something he is interested in (sub-goal generated through interaction with the environment). To determine whether an agent is attracted by certain places during navigation, an agent A_i maintains a list of interests $I_{i,\alpha}$, $\alpha = 1, \ldots, K$ and corresponding propensity-to-visit values ranging from 0 to 1.0. When passing by some point of interest, an agent compares its propensity-to-visit value to the attraction intensity of the place, and selects potential sub-goals, s_h, accordingly. To prevent an agent proposing the same attraction point as a sub-goal repeatedly, we endow an agent with ability to keep the list of visited places in memory.

An agent is associated with a specific role $\mathcal{R}$ in the context of the current coordination mode. For example, an agent takes a role of proponent if it proposes a new sub-goal, and other members become respondents to the agent. These roles are used to determine which communicative and joint actions are taken by whom, and in what orders in micro- and macro-coordination strategies.

4.3.2 Group Model

Our model assumes that the group memberships and goal of groups are known *a priori* and the relationships of agents are not subject to change throughout the simulation. Note that a group in our approach is formulated to provide social ties among agents, such as friends or family members. Group members tend to walk together and maintain the group cohesiveness. We model a group k as

$$\mathcal{G}_k = [k, L_k, S_k, p_k^0, g_k, Pr_k]$$

such that $\mathcal{G}_k$ has a leader agent L_k, a set of members S_k, initial position p_k^0, final goal g_k and a roughly planned global path from p_k^0 to g_k. Each group has a pre-assigned leader agent whose velocity vector is used as a base in a velocity matching process (see section 4.5.5).

To add a variability in group characteristics, each group may have different preference probability Pr_k for macro-coordination strategies (see section 4.4). We handle individuals in our simulation by permitting groups with a single member.

4.4 Model of Social Group Interaction

At each time step, a group $\mathcal{G}_k$ travels toward a final goal g_k by following a preplanned global path. A group walks in a clustered way by minimizing the distance between members while avoiding collisions to each other. When a sub-goal is triggered, the group evokes a set of coordination behaviors.

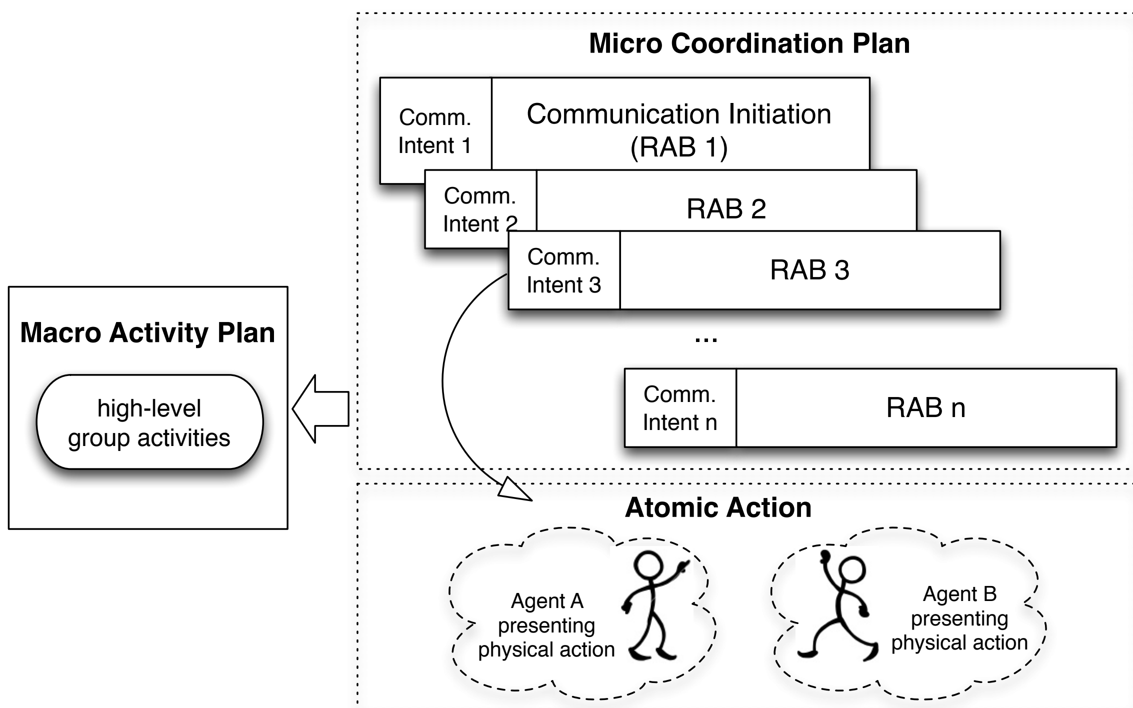


Figure 4.2: Coordination behavior decomposition in the CGCS model

Figure 4.2 illustrates the coordination behavior decomposition in our model into: 1) macro-coordination, 2) micro-coordination, and 3) atomic action units. Macro-coordination relates to the overall high-level activity determining the spatial movements of group members over time to accomplish a navigation goal and sub-goals of a group. The plan of divide and wait in the airport scenario in section 4.2 is an example of a macro-coordination behavior. Micro-coordination simulates the negotiation of CG among group members to decide on a macro-coordination plan given a new sub-goal. A micro-coordination plan may be further decomposed into a set of 'purposive' action blocks of reciprocal actions among the groups (hence Reciprocating Action Block, or RAB). RAB may specify that an agent needs to gain the attention of its group members, indicate the location of a sub-goal, or specify a meeting point for the group after the sub-goal is completed. Atomic actions are behavior pieces that may be animated, and can be used to build the RAB or the actions needed to accomplish a macro-coordination plan.

4.4.1 Macro-Coordination Strategy

When a sub-goal is triggered, a group may select a macro-coordination strategy (macroCS) from a predefined set of possible plans. This set of plans are designed to satisfy the needs of particular simulation/animation requirements.

For instance, in an emergency scenario, a set of macroCSs may be to abandon an original goal and find the nearest exit, to follow an authority figure, or to find a missing member. In a military scenario, a macroCS may be to send out a reconnaissance and wait, to divide and proceed to different goals, or to spread out and head to original goal position. The design of macroCSs may be based on common sense, typical domain-specific strategies, or from ethnographic observations of a population of people whose strategies are being modeled (Tsai et al., 2011; Endrass et al., 2011).

The plan selection is based on a probabilistic preference function of a group (hence Pr_k for $\mathcal{G}_k$), and members of a group share the chosen macroCS by each doing their participatory actions in particular roles. The selection of a macroCS results in a set of values for heading direction, desired position, and velocity for those agents involved in a group activity.

Because our interest is on generating social interaction behaviors of agents in a pedestrian simulation, we provide four macroCSs for the most common navigation strategies. The four macroCSs are 'detour-together, 'divide-and-wait', 'divide-and-meet', and 'divide-and-proceed'.

If a 'detour-together' plan is selected, the entire group detours together when a group member has to go to some point of interest. This plan reflects the follow-the-leader behavior, which is a commonly adopted approach for simulating group behaviors in other work (Lemercier et al., 2012; Reynolds, 1999). In the 'divide-and-wait' plan, an agent heads for a sub-goal by itself while the rest of a group members stay at the current location. After it achieves the sub-goal, the divided agent returns to where it left the group members. If different sub-goals are simultaneously generated for multiple agents, the ‘divide-and-meet’ behavior allows for all agents or sub-groups to go and execute their sub-goals and return to the point of separation. This plan can be thought as analogous to a temporary sub-group generation observed in real group movements (Peters and Ennis, 2009). Once all parties have accomplished their sub-goals, they return to the previous location where they divided up and resume the original navigation when the group is reconstituted. If the ‘divide-and-proceed’ plan is selected, a member that received the sub-goal trigger detours to visit a sub-goal while the rest of members proceeds with their original navigation plan. They reunite at the final goal location.

4.4.2 Micro-Coordination Strategy

In our model, micro-coordination relates to the simulation of CG negotiation. That is, a micro-coordination strategy (microCS) specifies how group members get each other to understand what they intend, so that some particular macroCS may be initiated. We call this a micro-coordination because it is always situated and local to the current group configuration. In order to act accordingly to the local situation, an agent constantly observes the surroundings and updates its awareness of the situation.

Situation Assessment

In a normal navigation mode, an agent checks the presence of points of interest or proximity to the final goal point. When it finds an interest place, it proposes the place as a sub-goal to group members by choosing and initiating an appropriate microCS according to its group's physical configuration.

To select a microCS, an initiator agent $A_i \in \mathcal{G}_k$ collects the following sensory inputs with respect to its group members that we model as the situational relationships (sit_{ij}) of A_i with all agents $A_j \in \mathcal{G}_k, \forall j \neq i$.

$$sit_{ij} : [d_{ij}, gd_j, V_{ji}|\alpha_i, E_{\mathcal{G}_k}]; \quad A_j \in \mathcal{G}_k, \forall j \neq i, \tag{4.1}$$

where d_{ij} is a distance between A_i and A_j, gd_j is a gaze direction of A_j, $V_{ji}|\alpha_i$ is the visibility of A_j to A_i subject to A_i's field of view α_i, and $E_{\mathcal{G}_k}$ describes the local environment of $\mathcal{G}_k$ as the level of congestion around the group $\mathcal{G}_k$. The effect of $E_{\mathcal{G}_k}$ is to determine the value of α_i and the visibility and audibility thresholds for d_{ij}. In a sparse crowd, longer-range gestures and speeches may be allowed while agents must be much closer together to communicate in a dense crowd. Furthermore, each agent may have its own evaluation function as a property of the agent (to determine its tolerance for crowdedness, for example). Then it selects an available microCS depending on the sit_{ij}; $\forall A_j \in \mathcal{G}_k, j \neq i$.

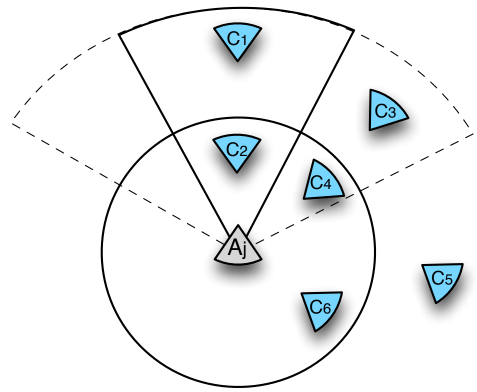


Figure 4.3: Six possible relations of group members

Figure 4.3 shows the six possible relations some agent A_j may have to A_i that will influence the microCS choices made by A_i. In each case, we assume that A_i is already oriented toward A_j (i.e. V_{ji} is TRUE), otherwise no microCS initiation is possible because the initiator would not know where the recipient is. A_i could already be in the visual field of A_j (conditions c_1 and c_2). A_i could be near to the visual field of A_j (conditions c_3 and c_4). We denote the near visual field condition as V'_{ij}. A_i could be outside V'_{ij} (conditions c_5 and c_6). For notation, we call this state $\overline{V}_{ij}$. Additionally we model the audibility range of A_j to A_i (meaning if A_i can gain the attention of A_j by using a vocal signal). Conditions c_2, c_4, and c_6 are in the audibility range, which we denote $\mathcal{H}_{ij}$, and c_1, c_3, and c_5 are in the $\overline{\mathcal{H}}_{ij}$ condition. In summary, the condition to state mapping are: $c_1 \leftarrow (V_{ij}, \overline{\mathcal{H}}_{ij})$, $c_2 \leftarrow (V_{ij}, \mathcal{H}_{ij})$, $c_3 \leftarrow (V'_{ij}, \overline{\mathcal{H}}_{ij})$, $c_4 \leftarrow (V'_{ij}, \mathcal{H}_{ij})$, $c_5 \leftarrow (\overline{V}_{ij}, \overline{\mathcal{H}}_{ij})$, and $c_6 \leftarrow (\overline{V}_{ij}, \mathcal{H}_{ij})$. Hence, Equation 4.1 can be evaluated for each A_j to assign it with one of our six configuration labels.

Selection of Micro-Coordination Strategy

Consider the situation where agent A_i receives a sub-goal event trigger (e.g., to visit the restroom). Depending on the evaluation of Equation 4.1, it will select an appropriate microCS, μ_γ, to perform with the closest A_j. For example, if Equation 4.1 evaluates to $c_3 \leftarrow (V'_{ij}, \overline{\mathcal{H}}_{ij})$, it may select the microCS described in Table 4.1. For our example, suppose the macroCS, $\mathcal{P}_\beta$, in step 3 is a 'divide-and-wait' strategy, and A_j will be tasked to communicate an 'inform-to-wait' activity with all other group members. This may require A_j to propagate the message to other group members by performing a microCS action with each one to initiate the 'inform-and-wait' macroCS. The macro-coordination activity of A_i, in this case, is to proceed to its sub-goal and return.

In our c_3 example, a second microCS may have been selected to move into hearing range

Table 4.1: An execution of a micro-coordination, μ_γ. $(V'_{ij}, \overline{\mathcal{H}}_{ij})$ condition

step	Action Description
1	A_i moves to satisfy V_{ij}
2	A_i performs a signaling action, s_i.
	A_j gives attention to A_i.
3	A_i proposes a macro-coordination, $\mathcal{P}_\beta$ (i.e. select $\mathcal{P}_\beta$)
	A_j signifies acknowledgement for $\mathcal{P}_\beta$.
4	A_j accepts A_i's acknowledgement
	A_i finalize the agreement on using $\mathcal{P}_\beta$.
	If μ_γ is successful,
	Return TRUE (i.e. execute $\mathcal{P}_\beta$),
	Else Return FALSE (coordination failed)

(satisfying $\mathcal{H}_{ij}$) and simulating a speech interaction. Yet another c_3 microCS may be to have A_i walk within hearing range, and signal A_j to look at her, and then to proceed with a visual/gestural interaction. In our model each of these microCSs are coded separately as alternatives from which A_i may choose. This range of possible choices picked either randomly or based on some agent-specific preference may serve to make the simulation seem less mechanistic or fully deterministic, thereby adding to realism. The key is that the design of the microCS behaviors conform to the rules for CG negotiation outlined in Section 4.2. Similarly, each configuration $c_1, \ldots, c_6$ can initiate a set of appropriate microCSs.

The failure or success of a microCS is dependent on two factors. The first is that the respondent agent A_j is somehow unable to comply with the signaling request of A_i (e.g., because it is simultaneously attending to another microCS from some other agent in the group), or if any step in the microCS script is unattainable (e.g., A_i's way is blocked to make it visible to A_j). The second is that the final random evaluation function of the microCS returns a FALSE value. We do this to add some variability to the simulation. The probability of this random function returning FALSE is generally set to a very low number, meaning that most microCS negotiations result in a TRUE result. The associated macroCS script is activated only if the microCS is successful. The resulting action when a microCS fails is dependent on domain of simulation, and intended action. A_i could decide to abandon the sub-goal, or it may retry with a different microCS.

Reciprocating Action Block (RAB)

A microCS may be further decomposed into a set of RABs. RAB consists of a set of actions to simulate the execution of a specific unit of communicative intent (e.g., an agent getting the attention of its interlocutors). An RAB typically involves dual actions that need to be executed together or consecutively (e.g., the interlocutor nods when the first agent waves in its field of view). The utility of RAB is thus to create cohesive atomic behaviors within a coordination sequence. Table 4.2 illustrates such an action block in our airport scenario. A_i may signal her intention $\mathcal{S}$ of heading to a schedule board by pointing toward it. This is followed by A_j signaling acknowledgement by nodding at A_i.

Table 4.2: RAB

Role	Action Description
proponent	A_i presents s_i to A_j intending that $\mathcal{S}$
respondent	A_j takes up s_i by presenting s_j

Initially, an agent who receives a subgoal trigger is assigned with a proponent role in an RAB. As the communication proceeds, the roles may be interchanged. For example, at the proposal of using the 'divide-and-stay' macro plan from A_i, A_j understands the intention of A_i but may suggest to use the 'divide-and-meet' plan by pointing where he wants to drop by. Then, in the next chunk of RAB, A_j takes a proponent role and A_i becomes a respondent.

Communication Initiation

The first RAB in a microCS is always the preparatory action needed to ensure effective communication. For an agent who needs to initiate communication, it must first identify the other member agent and the spatial relation between them.

The agent has to assess the state of the respondent agents with respect to their perception model described in Figure 4.1 and the situational relationships described in Figure 4.3, and perform a necessary action to meet the condition for establishing communication.

Figure 4.4 revisits possible spatial relations of any two agents, A_i and A_j to detail the communication initiation. In this case, A_i is the initiator of communication. If A_i determines it is in the c_3 position, it is outside the immediate perception of A_j and has to move into a position where one of her means of communication is possible. The first RAB may then be selected to move within A_j's field of view (for example, position x_i). The second action block is for A_i to get A_j's attention (e.g., by waving). An alternative RAB may be to have A_i walk into hearing range (e.g., position x_j), before calling out to A_j to get his attention.

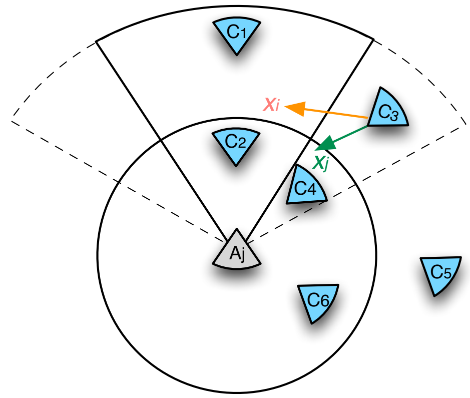


Figure 4.4: Communication initiation and spatial relation of group members

Chains of Coordination

An example of microCS composed of a chains of RABs between two agents is shown in Table 4.3 (Note that a microCS may be extended to include any number of participants). A_i moves to be within A_j's range of view in order to initiate communication. In step 2, A_i may wave at A_j to get his attention, and A_j gives attention to A_i by turning at A_i. In step 3, A_i may point to the schedule board for indicating that she wants to check the departure time, and A_j looks at where she points as a response to her signal. However, A_j wants to go to a restroom, so he points towards a nearby restroom and then to their current location. A_i understands what he means, so nods at him, in step 4. As B sees A_i's nodding at him, A_j nods back to her to indicate that he knows that she got the plan. A_i finalizes that they are on the same plan, and both take the movements.

4.4.3 Atomic Action

RAB satisfies a particular communicative intent, and specifies the reciprocating atomic actions within it. More than one RAB may satisfy a communicative intent, and different compositions of RABs result in a variety of microCSs. We show examples of the high-level communication intents and corresponding RABs (as atomic action pairs) in Table 4.4.

Table 4.3: Micro-coordinations plan, μ_θ

RAB	Action Description
RAB1	A_i moves to be within A_j's view range
RAB2	A_i performs a signaling action, s_i. A_j gives attention to A_i.
RAB3	A_i proposes a macro plan, $\mathcal{P}_\alpha$ (i.e. select $\mathcal{P}_\alpha$) A_j signals his understanding of A_i's intention
RAB4	A_j proposes a macro plan, $\mathcal{P}_\beta$ (i.e. select $\mathcal{P}_\beta$) A_i signifies acknowledgement for $\mathcal{P}_\beta$.
RAB5	A_j accepts A_i's acknowledgement A_i finalize the agreement on using $\mathcal{P}_\beta$.
	If μ_γ is successful, Return TRUE (i.e. execute $\mathcal{P}_\beta$), Else Return FALSE (coordination failed)

Table 4.4: Examples of communication intents and corresponding atomic-actions in RAB

Communication Intent	RAB Selected
Initiate communication	A_i moves into A_j's view; A_j turns toward A_i
Request attention	A_i waves in direction of A_j; A_j looks at A_i & nods
Suggest macro-behavior	A_i points toward sub-goal; A_j nods at A_i

4.5 Implementation

In this section we discuss the development of CGCS model for crowd simulations. We used C++ to implement code that runs on the CPU. The system generates low-level collision-free motion steering decisions using the velocity obstacle model (van den Berg et al., 2008a; Fiorini and Shillert, 1998), employing a publicly available RVO2 Library (van den Berg et al., 2011). OpenGL is adopted to render the simulation results in 2D and 3D graphics.

4.5.1 Simulation Overview

We distinguish pedestrian navigation tasks at the following three levels: 1) global path planning, 2) coordinated activity planning, and 3) reactive local planning. Figure 4.5 shows the relationship among the three navigation levels.

In the step of global path planning, a collision free route among static obstacles toward g_k is precomputed for each $\mathcal{G}_k$. This produces a sequence of waypoints from p_k^0 to g_k. From these, a set of velocities are generated to connect the waypoints. We shall denote these coarse global-level velocities as *guiding velocities* to distinguish them from other velocity values in this model. Note that a global path does not specify how exactly an agent should move from one waypoint to the next one along the planned path.

While adhering to the global path plan of its group from a long-term perspective, an agent determines its short-term activity plans by interacting with other group members at the coordinated activity planning step. As a result, a set of *preferred velocities* are derived.

Then an agent considers the maintenance of appropriate proximity to its group members as well as the presence of other agents and moving obstacles to determine its immediate steering in the reactive local planning step.

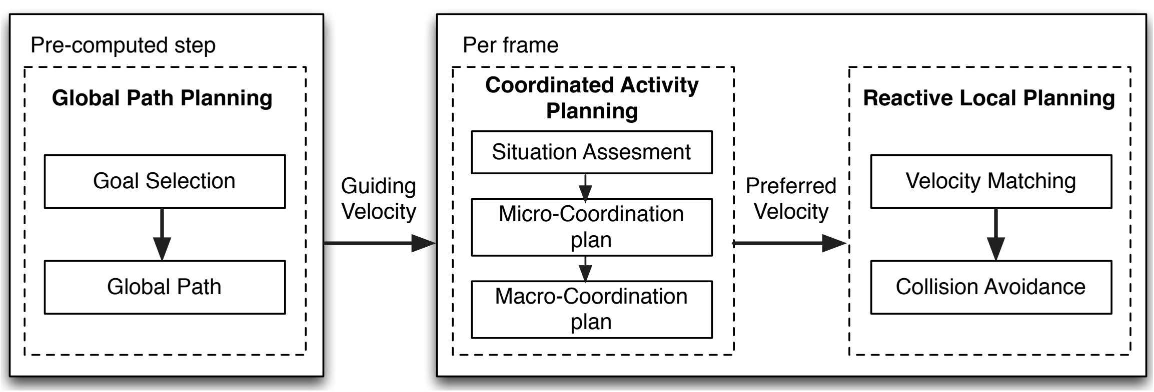


Figure 4.5: Three steps of navigation planning in our pedestrian model.

4.5.2 Agent Instance

For simplicity, we apply the perception model described in Figure 4.1 in the same way to all agents. The distance parameters d_h and d_v are subject to an environmental influence function, which is modeled as a sigmoid function:

$$
\begin{aligned}
d'_\epsilon &= d_\epsilon, \quad x \le t, \\
d'_\epsilon &= d_\epsilon\left(\frac{1}{1+e^{\frac{(x-t)}{w_\epsilon}}}\right), \quad x > t,
\end{aligned} \tag{4.2}
$$

where x is a number of agents in the unit cell, t is a tolerance for the crowdedness, w_ϵ determines the width of the sigmoid for each perceptual mode.

4.5.3 Environment Model and Global Path Planning

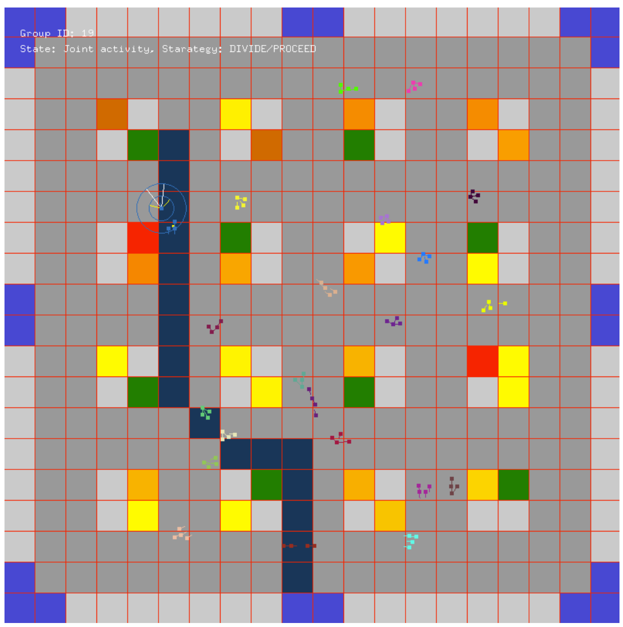


Figure 4.6: A snapshot of the shopping mall environment

Our virtual environment is represented as a 2D grid. An example of our environment model is shown in Figure 4.6. An environment may contain several points of interest as potential sub-goals. For example, restrooms, stores, or other kinds of places may be added depending on the target scenario of simulations.

We employed the A* algorithm to generate a global path for each $\mathcal{G}_k$, given the initial position, p_i^0, and final goal, g_k. The global path of $\mathcal{G}_{19}$ is drawn in navy blue on top of Figure 4.6. Any type of global path planners can replace the use of A* algorithm as long as it generates a collision-free route around static obstacles.

4.5.4 Coordinated Activity Plannning

Micro-Coordination Strategy

A series of communicative actions to simulate the microCS is defined for each situational relationship, $c_1, \ldots, c_6$, which are shown in Figure 4.3. Multiple options of microCS are available at each case, and the initiating agent must select one of them. For example, in c_1, a visual gestural coordination strategy may be activated in some cases, while in others, the agent may move to meet the audibility range constraint and communicate verbally.

Table 4.5: A list of microCS in our simulation

Action Part 1	
Sit.	Action Description
c_1	Immediate Comm. Visual Gesture
	Move to satisfy $\mathcal{H}_{ij}$ & Comm. Verbal
c_2	Immediate Comm. Visual Gesture
	Immediate Comm. Verbal
c_3	A_i moves to satisfy V_{ij} & Comm. Visual Gesture
	A_i moves to satisfy $\mathcal{H}_{ij}$ & Comm. Verbal
c_4	Immediate Comm. Verbal
	A_i moves to satisfy V_{ij} & Comm. Visual Gesture
c_5	A_i moves to satisfy V_{ij} & Comm. Visual Gesture
	A_i moves to satisfy $\mathcal{H}_{ij}$ & Comm. Verbal
c_6	A_j calls A_j, A_j turns around & Comm. Verbal
	A_i approaches and taps A_j, A_j turns around & Comm. Verbal
Action Part 2	
A_i performs a signaling action, s. A_j gives attention to A_i.	
A_i proposes a macroCS, $\mathcal{P}$ (i.e. select $\mathcal{P}$). A_j signifies acknowledgement for $\mathcal{P}$.	
If μ_γ is successful, Return TRUE (i.e. execute $\mathcal{P}$), Else Return FALSE (coordination failed).	

As previously shown in Table 4.1, a microCS consists of several steps of action blocks (i.e., RABs). Actions in step 1 in the table are pertinent to a specific situational relationship of agents. Actions described in step 2 through step 4 in the table are general and commonly applicable to any microCS.

Thus, we frame a microCS specification with two parts. Action part 1 contains RABs composed of set of particular actions subject to a kind of microCS, and action part 2 is the generalizable part of microCS.

Once agents complete the actions specified in action part 1, they perform the actions stated in the action part 2. A signaling action s implements the communication mode selected in action part 1. Table 4.5 summarizes a list of available microCSs in our simulation, in which $A_i; (A_i \in \mathcal{G}_k)$ is an initiator, and $A_j; (A_j \in \mathcal{G}_k, \forall j \neq i)$ is a respondent agent.

Macro-Coordination Strategy

Table 4.6: A list of macroCSs in our simulation

Macro-plan	Action Description
$\mathcal{M}_1$: detour-together	A_j leads $\mathcal{G}_k$ to a goal x_p All $A_j \in \mathcal{G}_k$ follow A_i If the goal is accomplished, $\forall A \in \mathcal{G}_k$ proceed to a final destination x_r
$\mathcal{M}_2$: divide-and-wait	A_i proceeds to a goal x_p If the goal is accomplished, A_j returns to its previous location
	A_j stays and waits until A_i returns
	If A_j returns, $\forall A \in \mathcal{G}_k$ proceed to a final destination x_r
$\mathcal{M}_3$: divide-and-meet	A_i proceeds $\mathcal{G}_k$ to a goal x_p If the goal of A_i is accomplished, A_i returns to their previous location
	A_j proceeds $\mathcal{G}_k$ to a goal x_q If the goal of A_j is accomplished, A_j returns to their previous location
	If A_i and A_j reunited, $\forall A \in \mathcal{G}_k$ proceed to a final destination x_r
$\mathcal{M}_4$: divide-and-proceed	A_i proceeds to a goal x_p, If the goal is accomplished, proceed to a next goal x_q
	A_j proceeds to a final destination x_r

As discussed in Chapter 4.4.1, we support four macroCSs, 'detour-together', 'divide-and-wait', 'divide-and-meet' and 'divide-and-proceed',which are listed in Table 4.6. In the table, A_i is a member of $\mathcal{G}_k$ and initiator, and $A_j; (A_j \in \mathcal{G}_k, \forall j \neq i)$ is a respondent agent. In our simulation, an agent choose a macroCS using the group preference characteristic, Pr_k.

4.5.5 Reactive Local Planning

A low-level reactive steering for collision-free and cohesive motion of agents are generated based on a Velocity Obstacle (VO) model (Fiorini and Shillert, 1998), using a publicly available RVO2 Library (van den Berg et al., 2011). A VO of an agent A_i is a set of all velocities that will eventually cause collisions with others, assuming that the other agents keep their current velocities. Hence, it is guaranteed that no collision occurs if A_i chooses a velocity outside the *VO*. The *VO* of agent A_i induced by agent A_j with velocity v_{A_j} is defined as:

$$VO_{A_i}^{A_j}(v_{A_j}) = \{v_{A_i} | \exists t > 0 : (v_{A_i} - v_{A_j})t \in D(x_{A_j} - x_{A_i}, r_{A_i} + r_{A_j})\},$$

where A_i and A_j have position x_{A_i} and x_{A_j}, and radius r_{A_i} and r_{A_j}, respectively. $D(x, r)$ represents a disc of radius r centered at x. The RVO2 Library computes velocities for each agent taking into account the reactive behavior of the other agents by implicitly assuming that the other agents draw a similar collision avoidance reasoning. Therefore, it chooses a new velocity that is the average of its current velocity and a velocity that lies outside the other agents VO (van den Berg et al., 2008a; Guy et al., 2012).

VO models do not provide support for group-based movement. In order to simulate a cohesive movement pattern among group members such as walking together, we incorporate the velocity matching process into the step of final velocity calculation.

We use a group leader's forward velocity vector V_{leader} as a base for the velocity matching. For each group member A_i of $\mathcal{G}_k$, a distance to leader L_k projected on V_{leader} is computed:

$$V_{leader} \cdot distance(L_k, A_i).$$

A distance greater than a higher threshold (1.5m in our simulation) implicates A_i is far behind L_k, therefore A_i accelerates while L_k decelerates. A distance less than a lower threshold (-1.5m in our simulation) means A_i is far ahead L_k, therefore L_k accelerates and A_i decelerates.

4.6 Summary

We presented a computational model informed by common ground theory to incorporate the impact of social interaction in multi-agent simulations. In our approach the CG model is framed as micro- and macro- coordination plans, which allows the multi-agent system to adapt to varying domains.

Our approach focuses on simulating the coordination process established by CG theory rather than modeling the process of social interaction at a lower level of communicative intent to reasoning. We believe that this intermediate level simulation is viable for two reasons. First, it is computationally more tractable than trying to simulate the micro processes of communication and negotiation of joint action itself. Second, CG theory is a well-vetted scientific theory, and already takes the effect of real human micro-level negotiation of joint activity into consideration. In fact, because the theory is well established, any bottom-up micro-negotiation model will be hard pressed to match the realism of a CG model.

Chapter 5

Unified Crowd Modeling Framework for Socially Plausible Animation Behaviors

5.1 Introduction

Producing realistic simulations of multiple virtual human characters is a challenging task. It is not enough to provide animated characters merely with visually pleasing motion styles and variety of appearance in a simulated crowd. Delivering a higher behavioral plausibility plays a vital role in the simulation. When multiple characters are involved in a simulation, realism is enhanced when the virtual characters interact and coordinate behaviors with other characters in a socially plausible human-like manner.

In this chapter we present a unified framework for crowd modeling that produces socially plausible animation behaviors of characters, where individuals within the crowd belong to different social groups, and where the behavior responds to events and scenarios in the simulation environment.

Our CGCS model presented in chapter 4 is capable of producing appropriate motion information that includes synchronized body-orientation and gesture for individual agents within the simulation. In our unified crowd modeling framework, the motion information generated by the simulation model is carried to and used in a rendering engine for animating virtual human characters. Because our model operationalizes a well-founded social-linguistic CG theory of human interaction, the behavior chains generated by the CGCS model form socially meaningful interactions among the characters.

5.2 System Overview

Figure 5.1 shows an architecture of our crowd modeling framework, composed of simulation engine and animation rendering engine. A simulation engine employs the multi-agent system with our social group interaction model presented in sections 4.3 and 4.4, and includes a goal generator, situation assessment module and coordination planner. As previously discussed, in our simulation each character is motivated by stochastically generated and/or environmentally driven goals. A random event generator triggers characters to have these goals. For example, a members of groups may have to visit the restroom (stochastically generated goal), or a member of a group may see a store that sells something she is interested in (environmentally driven goal).

When a character receives a goal event trigger, a situation assessment module evaluates the spatial relation among group members and the state of immediate environment for the character. A coordination planner collects the result of situation evaluation and selects appropriate micro- and macro-behaviors from databases in which each of micro- and macro-coordination strategies are coded separately as alternatives.

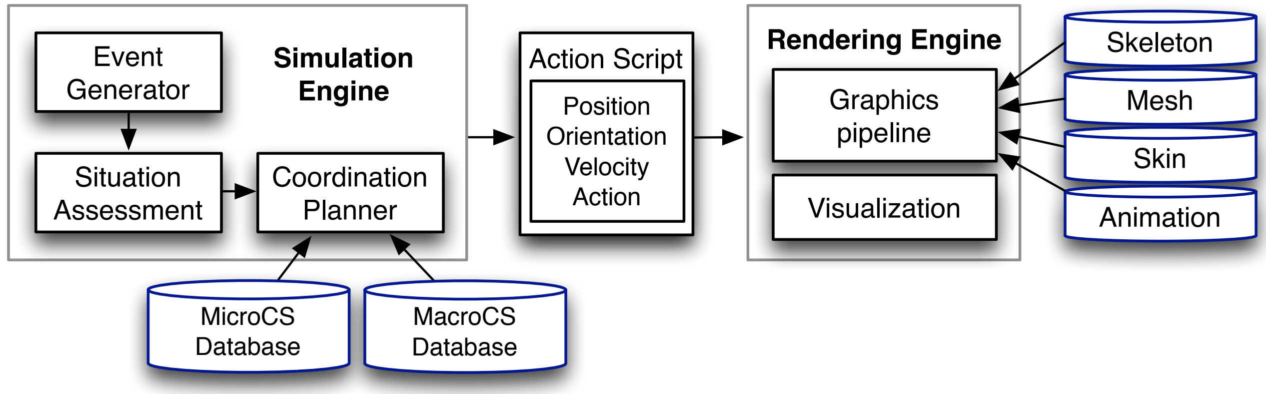


Figure 5.1: System architecture

As a result, information of position, orientation, velocity, and action specification defining kind of behavior and activation time is generated for each character at each frame. The information forms an action script and is transmitted to a rendering engine.

An animation rendering engine works in accordance with databases of characters' skeleton, mesh, skin, and pre-canned animation cycles to render out the animated actors given the input of action script. The set of animation sequences has been designed to support the repertoire of motions defined in the micro- and macro-coordination databases.

5.3 Animation Rendering Engine

As described in section 4.4.2, a behavior chain of participants is generated over the course of micro- and macro-coordination strategy executions, and that information can be associated with character animations. We define two kinds of behaviors; micro-behavior and macro-behavior. Micro-behavior relates to a micro-coordination strategy execution, hence corresponds to the atomic actions defined in RABs. Macro-behavior refers to the atomic actions determining the spatial movements of members to accomplish a macro-coordination specification. A's proceeding to a schedule board and returning to x, and B's waiting at x are examples of macro-behaviors in our airport terminal scenario in section 4.2.

5.3.1 Micro-behavior

We show examples of the high-level communication intents and how they are represented as bodily behaviors in Table 5.1. Available micro-behaviors are not limited to the given list, and

Table 5.1: Examples of communication intents and corresponding micro-behaviors

Communication Intent	Micro-behavior
Initiate communication	Call out, move toward, turn toward
Request attention	Call out, look at, wave hand(s)
Give acknowledgement	Look at, nod at
Suggest a macro-behavior	Point towards direction, wave over

any action reflecting what human carries out during the course of interaction can be added.

Figure 5.2 shows samples of 3D animation primitives that we produced for representing micro-behaviors defined in Table 5.1. We have two kinds of 'waving' gestures for motion variety, and a specific waving motion is selected by considering the distance between the proponent and respondent agents. For the 'pointing toward direction' motion, a specific pointing motion is selected by considering the angle between the character's body orientation and its sub-goal location.

Suppose any two agents, A_i and A_j, are in need of microCS execution with a spatial relation illustrated in Figure 5.3. A_i is outside the immediate perception of A_j, therefore it has to move into a position where one of her means of communication is possible. A micro-behavior may then be selected to move within A_j's field of view, x_i, and wave at A_j to get his attention. An alternative micro-behavior choice may be to have A_i walk into hearing range, x_j, and calling out to A_j to get his attention.

If the former micro-behavior is selected, A_j may respond by looking at A_i and nodding his head in acknowledgement. Then A_i may suggest a macro-behavior plan $\mathcal{P}_\beta$ by pointing toward the store and then to the ground. A_j will have to look at where A is pointing and turn back to A_i to signal that he has seen what she is pointing at and acknowledges her plan by nodding. Finally A_i has to show that she understands that A_j is in on the plan by nodding back.

Spatial movement parameters (i.e., desired body orientation and position) in micro-behaviors are derived by considering goal locations, character positions, and orientations. Given the length of animated motion, we empirically determined an appropriate synchronized time lag between a character's micro-behavior and the response of the others. The time lag for walking and turning motions are adjusted by considering the distance to travel and angle to rotate, respectively.

Table 5.2 shows a resulting action script specifying a sequence of micro-behaviors for A and B in the example just cited. This action script is applicable to the scenario we described in

(a) call out (b) wave over (c) wave one hand (d) wave both hands

(e) point to right($\angle 90$) (f) point to right($\angle 60$) (g) point forward (h) point ground

Figure 5.2: Micro-behavior animation primitives

the beginning of this section. x_a and x_b are positions of A and B respectively, x_i is some point within B's field of view, x_s is a store location, and x'_s is where A is pointing at. t_0 is a start time of the interaction. A length of an animated motion is denoted as $f^{actor\_id}_{motion\_name}$, in which *motion_name* represents a kind of motion, and *actor_id* is an operator of the motion.

5.3.2 Macro-behavior

Given the ability to signal to other group members to adapt their actions, actors select and perform a specific macro-coordination to accomplish the joint navigation. A a series of macro-behaviors compose a macro-coordination specification, and participants in the macro-coordination execution have to carry out their participatory actions. As

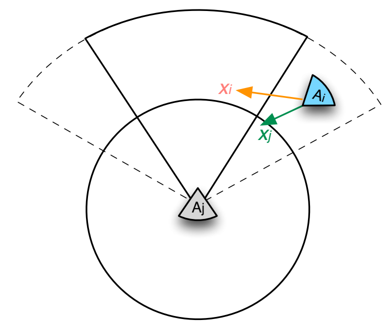


Figure 5.3: A_i and A_j in need of micro-behavior selections

Table 5.2: An example of micro-behavior action script for A_i and A_j, c_3 condition

microCS	Micro-behavior and time stamp			
step	A_i, motion and activation time		A_j, motion and activation time	
1	Walk to x_i	t_0		
2	Wave at x_b	$t_1 : t_0 + f^{A_i}_{walk\_to\_x_i}$		
3			Turn towards x_a	$t_2 : t_1 + f^{A_i}_{wave}$
4a	Point toward x_s	$t_3 : t_2 + f^{A_j}_{turn}$		
5a			Turn towards x_s and nod	$t_4 : t_3 + f^{A_i}_{point}$
4b	Point down	$t_5 : t_4 + f^{A_j}_{turn'} + f^{A_j}_{nod}$		
5b			Nod head	$t_6 : t_5 + f^{A_i}_{point'}$
6	Nod head	$t_7 : t_6 + f^{A_j}_{nod'}$		

previously discussed in section 4.5, we support four macro-coordination plans for pedestrian simulations, 'divide-and-proceed', 'divide-and-wait', 'detour-together', and 'divide-and-meet' (see Table 4.6).

Suppose the 'divide-and-proceed' macroCS plan is chosen after communication between $A_i; (A_i \in \mathcal{G}_k)$ and $A_j; (A_j \in \mathcal{G}_k, \forall j \neq i)$. A_i is an initiator of the macroCS, and A_j is a respondent actor. A_i leaves its group and proceeds to its own goal x_p while A_j keeps moving towards the group's original destination x_r. In the selection of the 'divide-and-wait' strategy, A_i heads for its own goal x_p by itself while the rest of group members stay at the current location. Once it achieves the goal, A_i returns to where it left the group members. When all of the group members reunite, they resume the original navigation.

If the 'detour-together' plan has been selected instead, A_i leads all the members together to a goal x_p. When the goal is accomplished, they resume the original navigation from the goal location. In essence, during the detour, the group is in a 'follow-the-leader' macroCS plan with A_i in the role of the leader.

The selection of a macroCS results in a set of desired position, velocity, and orientation values for those characters involved in a joint navigation. Table 5.3 shows an example of the macro-behavior outputs for A_i and A_j in our shopping mall scenario. $\mathcal{P}_\beta$ is selected after the coordination in Table 5.2, and actors start executing the macro-behaviors at time t_8. We empirically set an appropriate constant time as a duration of visit in a shop.

Table 5.3: An example of macro-behavior action script for A_i and A_j, $\mathcal{P}_\beta$ selection

macroCS	Macro-behavior and time stamp			
step	A, motion and activation time		B, motion and activation time	
1	Walk to x_s	$t_8 : t_7 + f^A_{nod}$	Wait (Idle)	$t_8 : t_7 + f^A_{nod}$
2	Stay at x_s	$t_9 : t_8 + f^A_{walk\_to\_x_s}$		
3	Walk to x_b	$t_{10} : t_9 + f^A_{stay}$		
4	Walk	$t_{11} : t_{10} + f^A_{walk\_to\_x_b}$	Walk	$t_{11} : t_{10} + f^A_{walk\_to\_x_b}$

5.4 Implementation and Results

Autodesk 3ds Max is used to develop 3D modeling and animation of virtual characters and their motions (see Figure 5.2). The Open Graphics Library (OpenGL) is used for visualizing simulated data with the animation primitives. Our crowd framework is tested on a desktop equipped with an Intel Core i7 2.67 GHz processor, 12GB of RAM, and a Nvidia Quadro 5000 graphics card with 2.5GB device memory.

5.4.1 Effects in Character Behaviors

We present resulted animation crowd scenes with our CGCS framework in this section. The simulation takes place in a generic shopping mall setting. Figure 5.4 gives an overview of our shopping mall model. The virtual shopping mall contains 6 restrooms and 82 stores as potential goals. We populated the shopping mall with 1000 characters, of which 800 are arranged into 200 social groups, and the remaining 200 characters are individual actors in the crowd. A number of members in a group ranges from 2 to 6, and is arbitrarily selected for each group.

Figure 5.5 shows characters communicating with other group members and presenting gestures. Depending on the spatial and environmental conditions, actors select different kinds of micro-behavior. Figure 5.5(a) shows a female character (labeled A) calling out to a

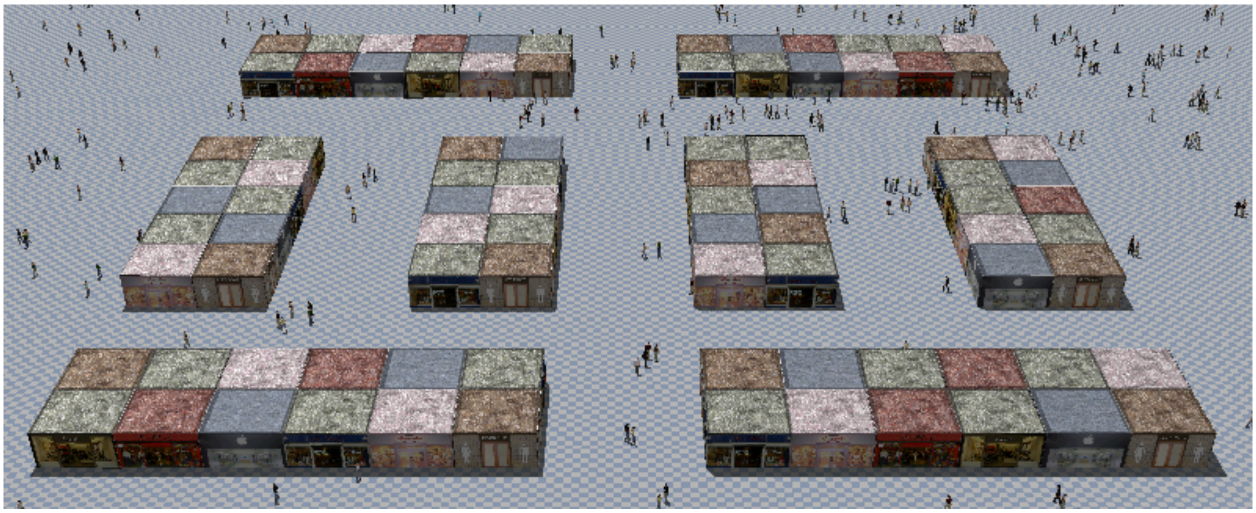


Figure 5.4: A snapshot of the shopping mall model

friend to initiate communication. A man character (labeled B) is waving his hand to bring a colleague's attention in Figure 5.5(b). On the left side of him, a group of people (labeled C) is having a conversation.

Figure 5.6(a) shows several individuals (D, E, and F) and a couple (G) standing near a restroom, and waiting for their friends or family members. Individual characters are in idle poses. The couple is talking to each other, and the female character is presenting a hand gesture to the male character. In Figure 5.6(b), it is observed that a clustered movement emerges, because people in a group walk together and maintain a cohesive movement.

5.5 Summary

In this chapter, we presented a crowd simulation framework to accommodate high-level socially-aware behavioral realism of animated characters. The framework completely integrate the simulation engine and animation engine, and the character behaviors for animations are determined at the point of simulation by considering the context of interactions and coordination among agents. Therefore a sequence of animated character behaviors is not merely a chain of reactive gestures, but form socially meaningful interactions.

There are several possible improvements for our crowd modeling framework. On a presentation side, the actions of the characters were repetitive and not always smooth enough because of the lack of variety in motions in the database. It is desirable to add a richer set of communicative gestures and movements for micro-behaviors (e.g., 'looking at' by head

(a) (b)

Figure 5.5: (a) A female character calling out to a friend (b) A male character waving hand

(a) (b)

Figure 5.6: (a) Individuals and a couple waiting for their members near a restroom (b) Group based movement

movement) that characters can display. In addition to that, motion synthesis techniques could be incorporated into the framework to improve the rendering engine for achieving good connectivity and smooth transitions in motions.

Visualizing a large number of animated characters is a heavy computational process. To increase the rendering efficiency, a dynamic level-of-detail approach to reduce the workload for the stages of graphics pipeline could be employed.

Chapter 6

Crowd Model Evaluation

6.1 Introduction

In order to investigate the impact of our social group interaction model in crowd simulation and animation, we designed and conducted several experiments. First, we compared simulation results with and without the incorporation of the CGCS model to examine the effect of considering communication costs in the pedestrian simulation. Second, a series of perceptual user evaluation studies was performed to investigate whether the CG model is applicable for crowd simulation, and whether our operationalization of the theory produces realistic crowd behaviors. Third, a set of experiments to demonstrate the computational efficiency and scalability of our CGCS model was performed.

6.2 Impact of Coordination Costs in Crowd Simulation

We performed two experiment sessions to show the quantitative differences in simulation results with and without incorporation of the CGCS model. In the first session, congestion levels at key points were measured. In the second session, the overall distribution of a crowd across the entire simulation space was analyzed. Our hypothesis for these experiments was that the pragmatic need to maintain common ground in individual groups incurs costs at the level of the entire simulation.

6.2.1 Congestion Levels at Key Points

Virtual Environment

A virtual shopping mall model shown in Figure 6.1 was used in this experiment. The shopping mall model contains 8 restrooms (green squares) and 25 shops (yellow to red squares) as potential sub-goals. The color intensity of shops represents the level of attraction. For example, shops in red are the points of highest attraction and those in yellow are the least attractive points. To make analysis results across conditions comparable in the current study, we assume all agents have homogeneous interests. Eight exits (light purple squares) are generated as possible final goals for agents. Each cell in the 2D grid describes a 5m$\times$5m area. The global path of $\mathcal{G}_{19}$ is drawn in navy blue in top of Figure 6.1.

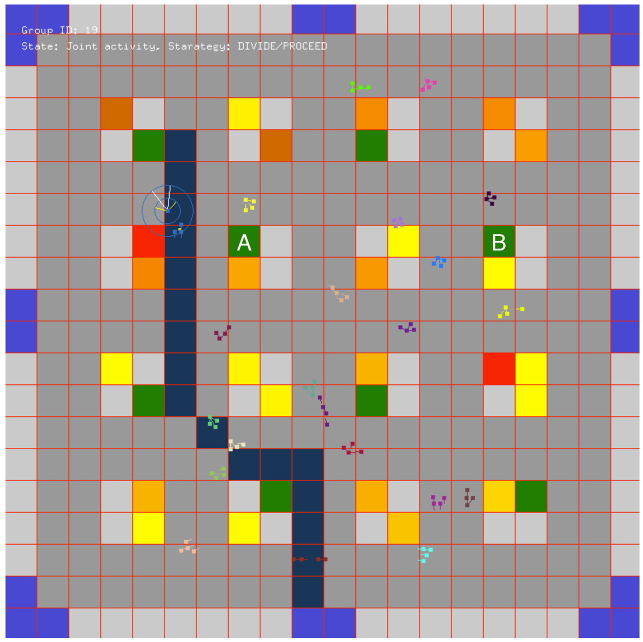


Figure 6.1: A snapshot of the shopping mall environment

Simulation Scenario and Condition

We set four simulation configurations and compared the congestion levels at key points to show the impact of the 'coordination overhead' in the simulation. Given the total number of agents n, the agents are organized as:

1. n of solely individual agents
2. $n/4$ simple groups of 4 members without considering the CG coordination model
3. $n/4$ groups of 4 members which instantly choose a macroCS without processing micro-coordination
4. $n/4$ groups of 4 members with incorporating the CG coordination model.

We shall call each of the four configurations CI (for Individuals), CDT (for always 'Detour-Together'), CM (for Macro only), and CμM (for both micro and Macro CSs), respectively.

To simulate simple groups in CDT, we set the groups to always choose 'detour-together' (i.e., all group members will satisfy all sub-goals together, before proceeding to the final goal). A simulation under this condition represents a commonly used approach of simulating a 'group-based movement' in current crowd simulation research (Reynolds, 1999; Lemercier et al., 2012). To make analysis results across conditions comparable in the current study, we assume all agents have homogeneous interests.

For all the individual agents in CI and agents of groups in CDT, CM and CμM, one of the eight exits in the virtual airport terminal (see Figure 6.1) is selected as a final goal g_k at random. Initial positions of agents are randomly distributed in a shopping mall, taking care that members within groups are collocated. Though agents in CI are all individuals, we initialized the simulation with agents clustered in groups of 4 to have the same initial conditions as CDT, CM, and CμM.

Starting from initial positions, agents walk around the shopping mall and eventually proceed to the exit. A random event generator triggers agents to visit the nearest restroom. When agents pass by shops, they may be attracted to some shop within a range (e.g., 15m$\times$15m in our implementation).

We used 1000, 2000, and 3000 for the total number of agents n, and hence they are organized as 250, 500, and 750 groups, respectively. Default perceptual geometry of agent is set to: $[d_v, d_h, \alpha] = [15\text{m}, 5\text{m}, \angle 60]$. For the environmental influence introduced in Equation 4.2, $t = 12$ and $w = 35$ are used (for simplicity, we use the same w for d_h and d_v). w is determined assuming that the maximum agent capacity of a unit cell of 5m$\times$5m being around 50.

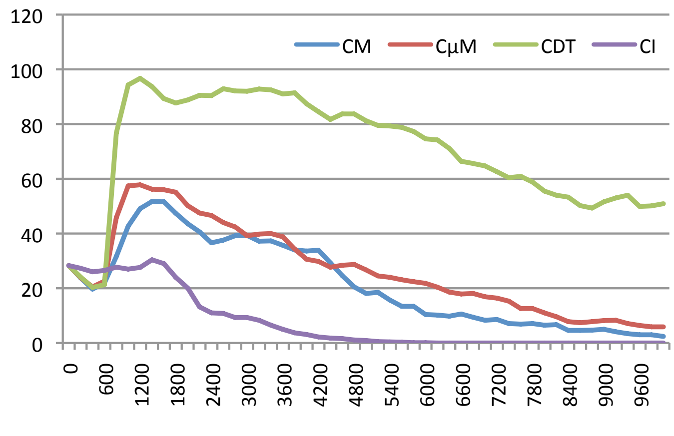
Procedure

In order to compare the congestion levels among the four simulation configurations, we counted the number of agents in 3$\times$3 unit cells (hence 15m$\times$15m) centered at key points such as shops or restrooms per frame.

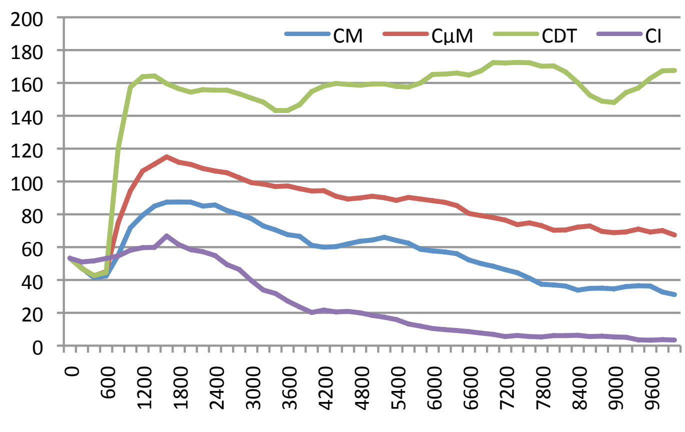
Results and Discussion

We present the results measured in area A and B (marked in Figure 6.1) in Figure 6.2 and 6.3. All the results are averaged over 10 independent simulations. Stochastic events invoking agents to visit a restroom are indicated with gray vertical lines. Figures 6.2(a)-(c) are the level of congestion measured in area A for 250, 500, and 750 groups, respectively. Figures 6.3(a)-(c) are the results for area B. The x-axis in each graph represents time and the y-axis measures the level of congestion in # of agents per unit area.

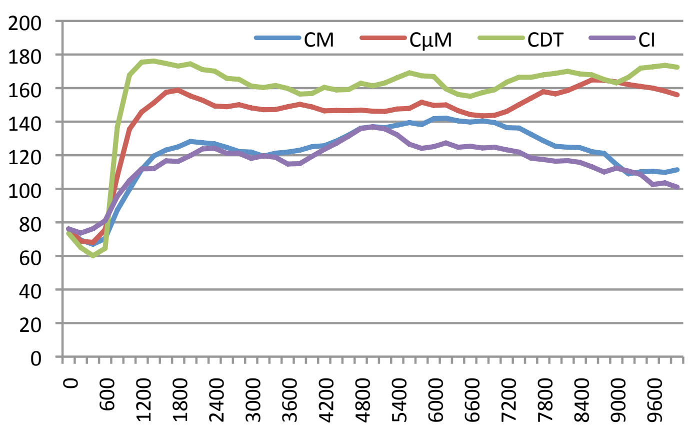
The results presented in the graphs indicates the emergent effects of individual agent and small agent group behavior on the overall crowd simulation. From the graphs, it can be seen that the CI case almost always results in the lowest congestion because there is no coordination or joint action overhead in the individual action plans. CDT consistently results in the highest congestion across crowd size simply because the total distances traversed by all agents is the highest in this condition. The interesting differences come in the CM and



(a) 1000 agents in 250 groups

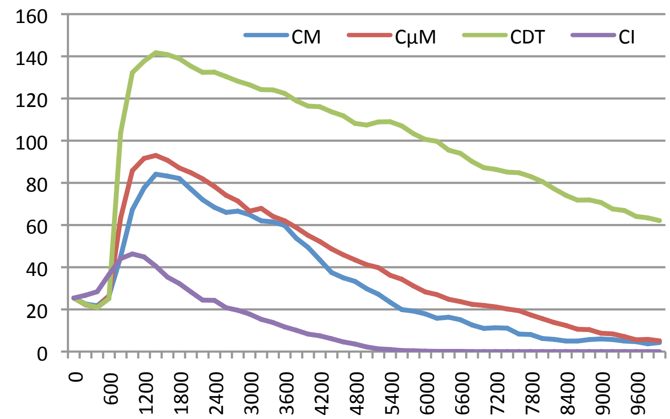


(b) 2000 agents in 500 groups

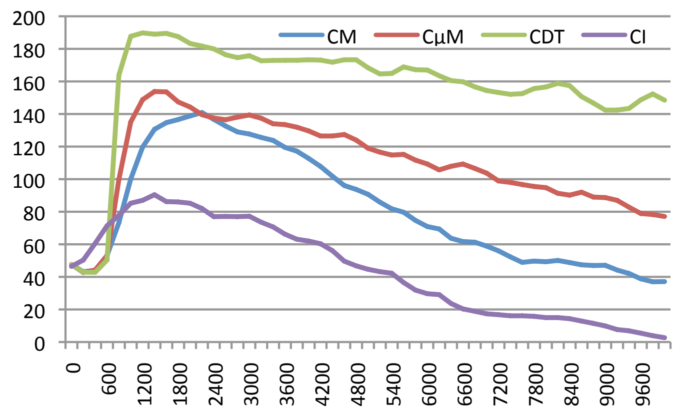


(c) 3000 agents in 750 groups

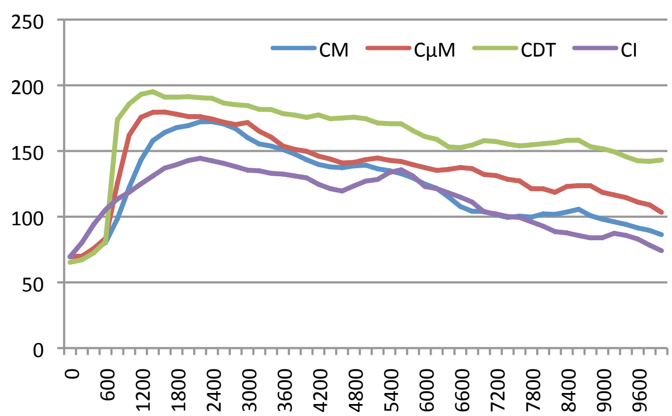
Figure 6.2: Congestion levels measured as the number of agents in unit areas in area A for CI, CDT, CM, and CμM conditions



(a) 1000 agents in 250 groups



(b) 2000 agents in 500 groups



(c) 3000 agents in 750 groups

Figure 6.3: Congestion levels measured as the number of agents in unit areas in area B for CI, CDT, CM, and CμM conditions

CμM conditions as these vary depending on the overall crowd size. Notice that across all crowd sizes, congestion tends to fall over time as some agents and agent groups leave the mall, leading to overall lower congestion everywhere. Since the sub-goal events are generated stochastically, not all groups in the CDT, CμM, and CM cases will have members that receive sub-goal events at any particular sub-goal event cycle.

When we compare the conditions with respect to overall crowd size, we notice that apart from CDT and CI, CM and CμM are quite similar when the crowd is sparse (1000 agents in 250 groups), with CμM resulting in slightly higher congestion followed by CM and then CI. This shows that the joint action and the micro-coordination overheads have some effect at this crowd level.

This order of difference becomes accentuated in the middle level of crowd size (2000 agents in 500 groups). The general difference between locations A and B in the simulation is that A is near the middle of the environment where the effects across conditions are accentuated, and B is near the exit where congestion tends to fall more quickly as agents exit the environment over time. In both areas A and B, the micro-coordination overhead causes higher congestion than the macroCS-only (CM) condition. This shows that the behavioral cost of performing the necessary micro-coordination activity cannot be ignored in overall crowd simulation with multiple agents.

When the overall crowd density is very large (3000 agents in 750 groups), the congestion across the CDT, CμM, CM, and CI conditions become relatively similar because we have arrived at a ceiling effect of overall over-congestion. One way to think of this is that when the crowd density gets too high, no strategy will work effectively. However, even in this condition the order of congestion levels across groups (CDT, CμM, CM, CI) is generally maintained, and the micro-coordination process adds higher congestion than the macroCS-only condition. In area A, especially the congestion of CμM even approaches that of CDT (which is an unrealistic condition in real life), while the CM congestion level falls to near CI levels after 8800 time steps as agent groups exit the environment.

We found that a non-CG simulation to always 'out-perform' a CG-based model with regard to level of congestion. This indicates that models that do not consider the cost of coordination may fail to capture the real crowd effects. Such underestimations of crowd complexity may, for example, cause designers to be overly optimistic about the size of corridors or of evacuation rates in an emergency.

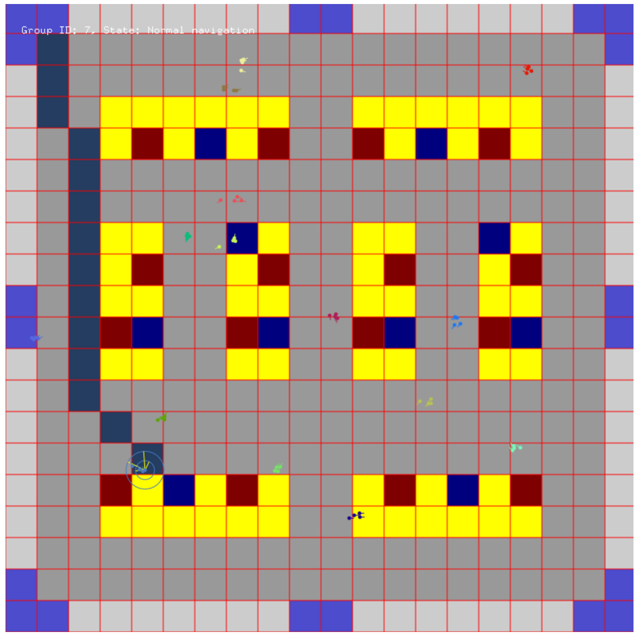


Figure 6.4: A snapshot of the airport terminal model

6.2.2 Overall Crowd Distribution

Virtual Environment

In this experiment we used an airport terminal environment shown in Figure 6.4. The airport contains 10 restrooms (blue squares), 16 flight schedule boards (red squares), and 62 stores (yellow squares) as potential sub-goals. Eight gates (light purple squares) are generated as possible final goals for agents. Each cell in the 2D grid describes a 5m×5m area. The global path of $\mathcal{G}_7$ is drawn in navy blue on top of Figure 6.4.

Simulation Scenario and Condition

The four simulation conditions, CI (for Individuals), CDT (for always 'Detour-Together'), CM (for Macro only), and CμM (for both micro and Macro CSs), that we used in the first experiment (see section 6.2.1) are also employed in this experiment. We used 2000 for the total number of agents n, and hence they are organized as 500 groups. Parameters for the perceptual geometry of agents and environmental influence were also set to the same as the experiment in section 6.2.1.

For all the individual agents in CI and agents of groups in CDT, CM and CμM, one of the eight gates in the airport terminal is selected as a final goal g_k at random. Initial positions of agents are randomly distributed in an airport terminal, taking care that members within groups are collocated. Agents in CI are clustered in groups of 4 to have the same initial conditions as CDT, CM, and CμM. Starting from initial positions, agents walk around the airport and eventually proceed to the gate. A random event generator triggers agents to visit the nearest restroom. When agents pass by shops, they may be attracted to some shop within a range (e.g., 15m$\times$15m in our implementation).

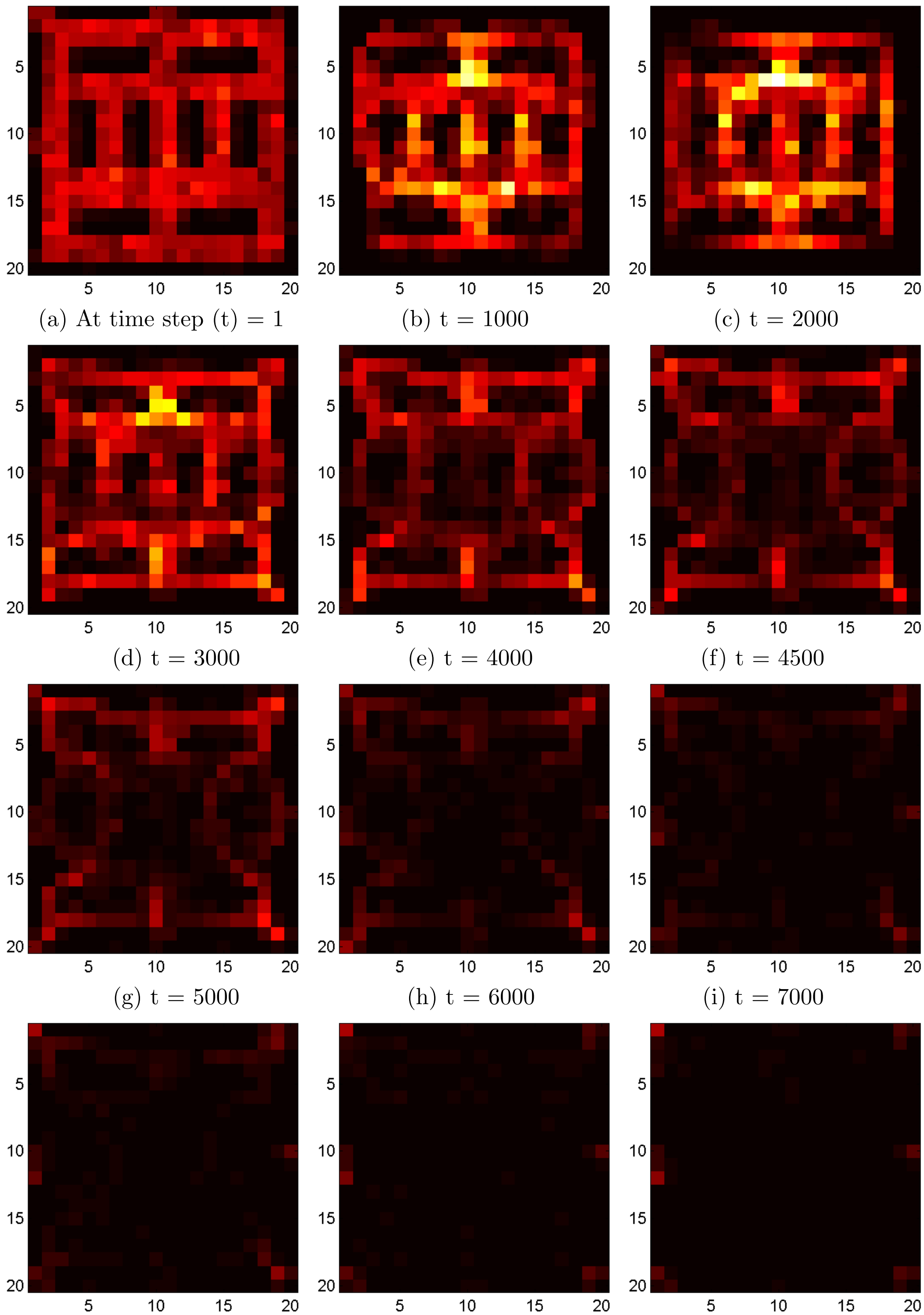
Procedure

In order to compare the congestion distribution among the four simulation configurations, we counted the number of agents in a 1$\times$1 unit cell (hence 5m$\times$5m) across entire simulation space per frame. We visualize the dynamic crowd distribution as a color map where color varies smoothly from black through shades of red, orange, and yellow, to white. We normalized the range of color mapping into [0, 30], where 0 mapped to black and 30 mapped to white. If any values exceed 30, white color was mapped.

Results and Discussion

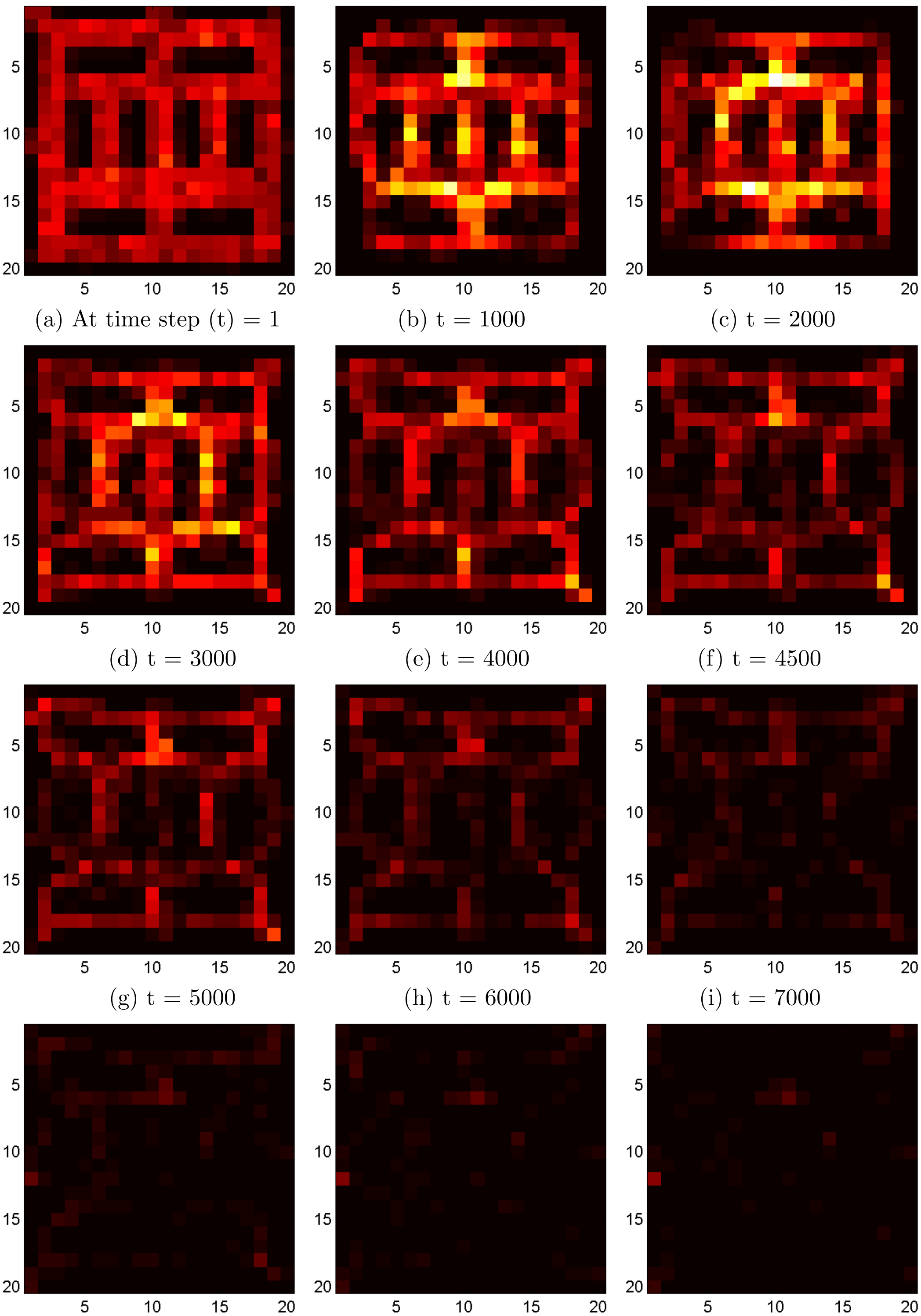
Figures 6.5, 6.6, and 6.7, and 6.8 show the crowd distributions over time in CI, CM, CμM, and CDT simulations, respectively. All the results are averaged over 10 independent simulations. In the color map, the x- and y-axes are aligned to the coordinates of the virtual environment. Note that we set $(0, 0)$ in the color map at the top left corner so that corresponds to the OpenGL window.

The results presented in the color maps indicates the effects of different simulation models on the congestion distribution over the entire simulation space. Overall, the differences in the congestion distribution over the four conditions are similar to those observed in the first experiment. At the beginning of the simulation, the crowd distributions are similar across all conditions (see Figures 6.5(a), 6.6(a), 6.7(a), and 6.8(a)). Comparisons of congestion distributions shown in (g)-(l) of Figures 6.5, 6.6, 6.7, and 6.8 confirm that the congestion tends to fall more quickly in CI case than the other conditions. This is because there is no micro- and macro-coordination overheads in the individual action plans. The CDT simulation results in the congestion lasting for longer time among other conditions. In addition to that, we notice that the congestion is more spread throughout the entire space in the CDT case. This is because the total distances traversed by all agents is the longest in this condition. When compares CμM vs CM, the CμM simulation records higher density in several spots while the overall distributions appear similar (see Figures 6.6(d)-(f), and 6.7(d)-(f)).



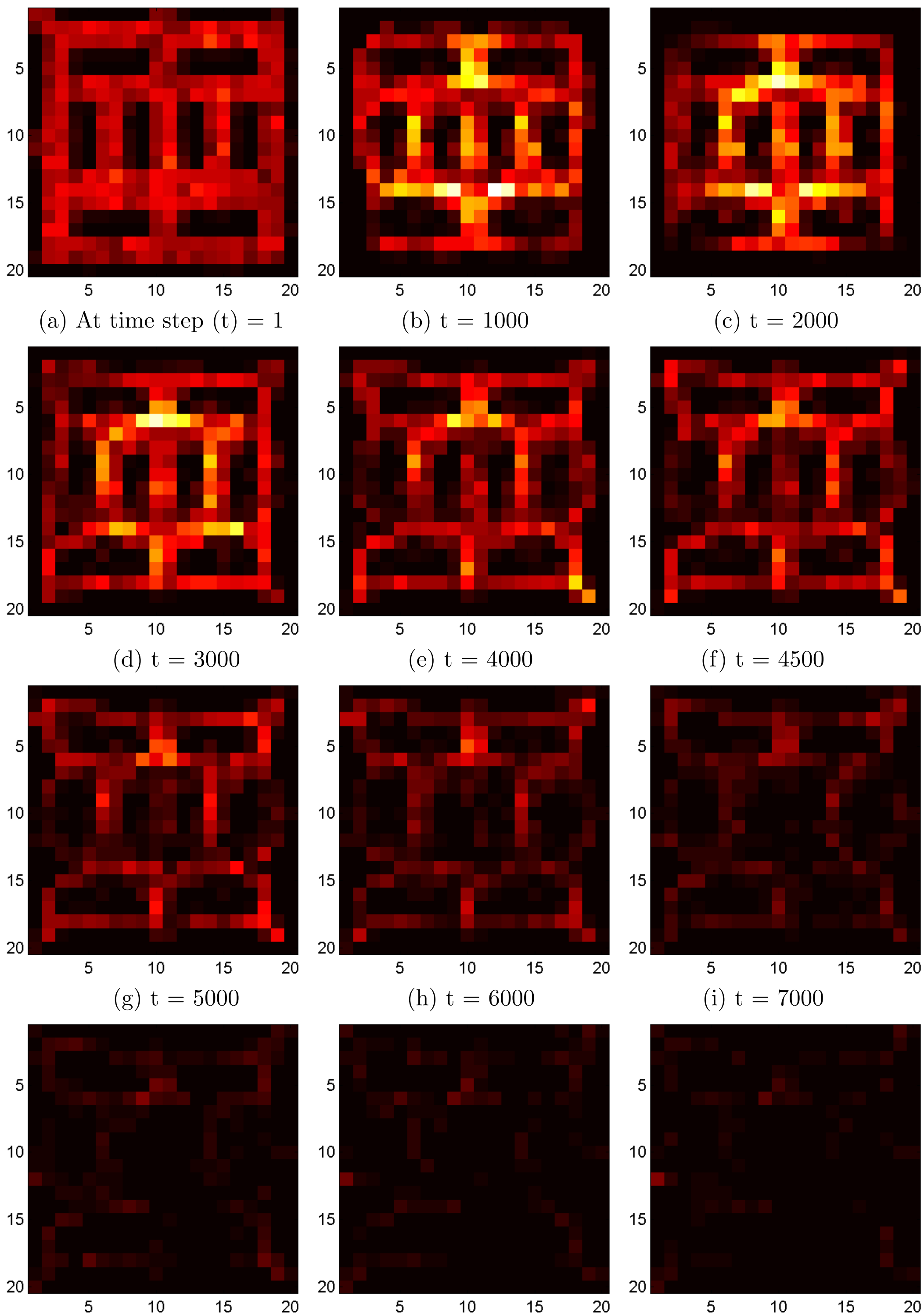
(a) At time step (t) = 1 (b) t = 1000 (c) t = 2000
(d) t = 3000 (e) t = 4000 (f) t = 4500
(g) t = 5000 (h) t = 6000 (i) t = 7000
(j) t = 8000 (k) t = 9000 (l) t = 10000

Figure 6.5: Crowd distribution in the CI condition, 2000 agents of 500 groups



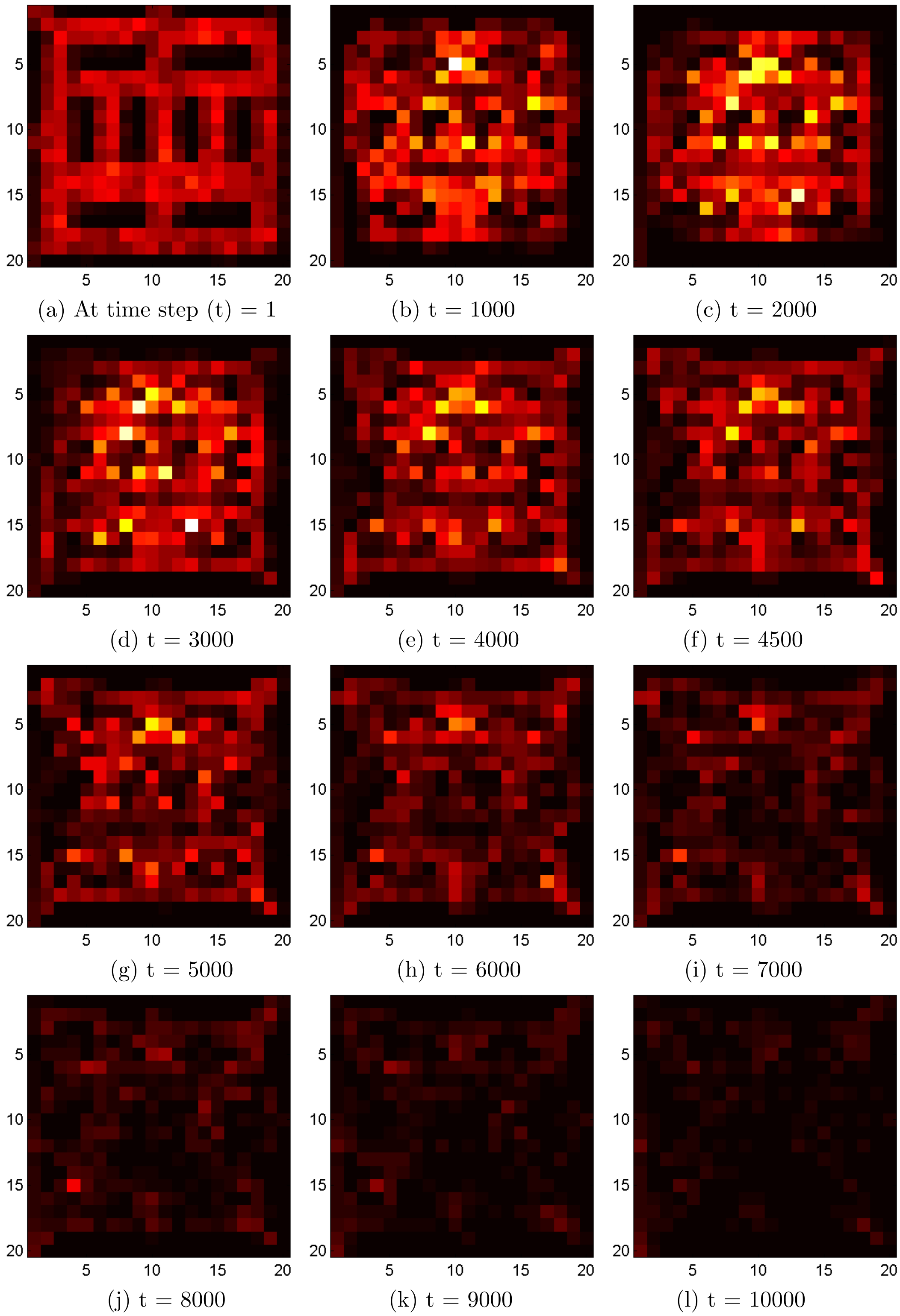
(a) At time step (t) = 1
(b) t = 1000
(c) t = 2000
(d) t = 3000
(e) t = 4000
(f) t = 4500
(g) t = 5000
(h) t = 6000
(i) t = 7000
(j) t = 8000
(k) t = 9000
(l) t = 10000

Figure 6.6: Crowd distribution in the CM condition, 2000 agents of 500 groups



(a) At time step (t) = 1
(b) t = 1000
(c) t = 2000
(d) t = 3000
(e) t = 4000
(f) t = 4500
(g) t = 5000
(h) t = 6000
(i) t = 7000
(j) t = 8000
(k) t = 9000
(l) t = 10000

Figure 6.7: Crowd distribution in the CμM condition, 2000 agents of 500 groups


(a) At time step (t) = 1 (b) t = 1000 (c) t = 2000

(d) t = 3000 (e) t = 4000 (f) t = 4500

(g) t = 5000 (h) t = 6000 (i) t = 7000

(j) t = 8000 (k) t = 9000 (l) t = 10000

Figure 6.8: Crowd distribution in the CDT condition, 2000 agents of 500 groups

6.3 Perceptual Evaluation

Our CGCS model derives from CG theory with a basis in extensive observational science, and provides a means to simulate purposive behavior of people groups in interacting. The open research questions are whether this model is suitable for crowd simulation, and whether our operationalization of the theory produces realistic crowd models. The question may be reformulated by asking whether our model produces plausible animations from the viewpoint of human observers in a series of user studies.

Since our model decomposes the overall coordinated behavior of groups into Macro- and μicro-coordination components, we designed a set of pairwise-comparative studies to investigate the efficacy of the approach. We shall discuss the virtual setting in which our tests are conducted, the study conditions tested, the design of the studies, and our study results.

6.3.1 Virtual Environment and Scenario

In these studies, we employed an airport terminal scenario because there would be higher chances to use divide-and-meet and divide-and-proceed walking strategies for travelers in an airport terminal than for shoppers in a shopping mall.

Our virtual airport terminal environment is shown in Figures 6.9 and 6.10. The airport terminal model contains 10 restrooms (blue squares), 16 flight schedule boards (red squares), and 62 stores (yellow squares) as potential sub-goals. Eight gates (light purple squares) are generated as possible final goals for agents. Same as in the previous experiments, each cell in the 2D grid describes a 5m×5m area.

A crowd in each animation was made up of 30 individuals, 50 groups of 2 individuals, 15 groups of 3, 5 groups of 4, thus 100 groups in total. This distribution of pedestrian was determined by approximately following the information reported in (James, 1953). However, to make the gestures of characters easily observable to viewers, the camera was set to focus at a few number of groups, but not at the overall scene.

For all of the individuals and agents of groups, one of the eight gates is selected as a final goal at random. Starting from initial positions, agents walk around the terminal and eventually proceed to the gate. A random event generator may trigger agents to visit the nearest restroom. When agents pass by schedule boards and shops, they may be attracted to some of the places.

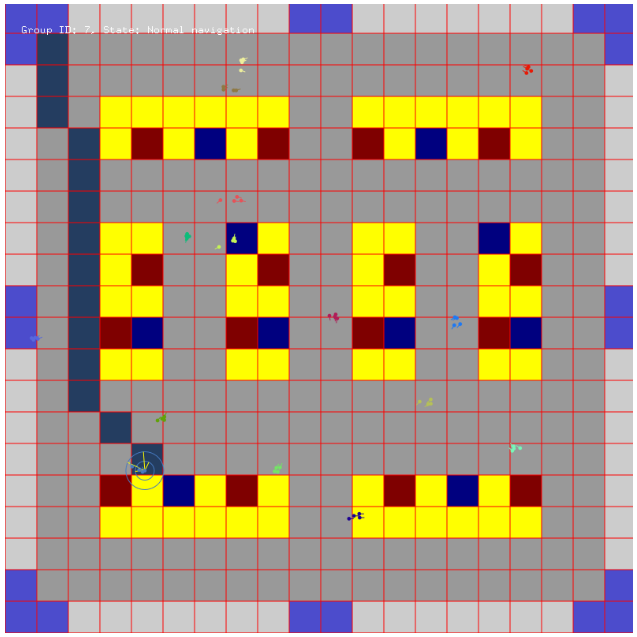
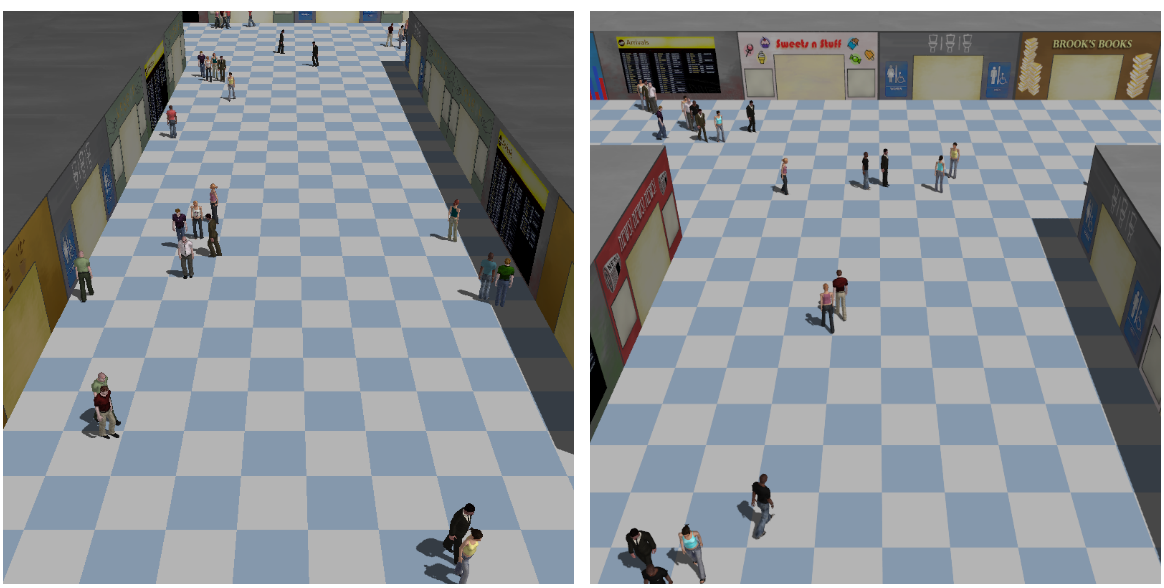


Figure 6.9: A snapshot of 2D airport terminal environment



(a) (b)

Figure 6.10: (a) and (b) 3D representation of the virtual airport environment from different camera angles

6.3.2 Simulation Condition

To test the degree of realism afforded by our model, we generated a number of 30-second animations in the four conditions summarized in Table 6.1. We varied whether μ(micro)-coordination was included in the simulation, and the kind of M(macro)-coordination strategies employed. When no μ-coordination is used, the groups just proceeded to the M-coordination plan once a sub-goal is introduced. In CDT and CμDT, the groups always chose commonly used detour-together strategy (Lemercier et al., 2012; Reynolds, 1999), and in the CM and CμM condition, our four Macro-coordination plans described in Section 4.4.1 were employed.

Table 6.1: Study Conditions

Condition	μ-Coordination	M-Coordination
CDT	No	Detour-Together
CμDT	Yes	Detour-Together
CM	No	Varied
CμM	Yes	Varied

6.3.3 Study Design

We tested our study conditions using a pairwise comparison design as in (Peters and Ennis, 2009). To determine the effect of micro-coordination on human perceptions of the crowd behaviors produced by our model, we compare CDT vs CμDT, and CM vs CμM. The first comparison tests the effectiveness of introducing CG to the common detour-together strategy (Lemercier et al., 2012; Reynolds, 1999), and the second comparison tests the effectiveness of adding CG to a more varied set of macro-coordination strategies. We use two measures as our dependent variable. The first, $\mathcal{M}_{\mathcal{R}}$, measures the participant's estimation of the realism of a simulation, and the second $\mathcal{M}_{\mathcal{P}}$ measures the participant's estimation of the plausibility of a simulation.

For each pair of model comparisons, the participant were shown several animation pairs that were generated using the two models. That is, each participant was shown 10 pairs of different CDT and CμDT, and 11 pairs of different CM and CμM animations. The order of the presentations were randomized. We followed a within-subjects design, therefore all of the participants were shown the 21 pairs of animations.

Our study consists with three tasks. The first two tasks are for the realism and plausibility measures. The third task is for investigating participants' understanding of character behaviors.

For our realism measure, we employ a cover story to avoid demand characteristic biases. For each pair of simulations presented, the participant was told that one animation was derived from tracking data from a real crowd, and the other was synthetically generated. The participant was given a forced choice task of determining which was 'Real' and which was synthetic. $\mathcal{M}_{\mathcal{R}}$ measures the realism estimate for a simulation condition as the fraction of the number of times a simulation in that condition is rated as 'Real' across multiple exposures. For example, if a participant judges 7 of the 10 CμM simulations as being from 'real data', then CμM has a $\mathcal{M}_{\mathcal{R}}$ score of 0.7 (and CM has a $\mathcal{M}_{\mathcal{R}}$ score of 0.3).

For our plausibility measure, participants were asked to say if the behaviors of the groups in a particular simulation are plausible on a 7-point likert scale. Hence for our 10 presentations, each simulation model will have 10 likert scores. The plausibility measure $\mathcal{M}_{\mathcal{P}}$ of the model is the average of the 10 likert scores.

To obtain a better understanding of the criteria used by our participants to judge plausibility, an additional pair of CM and CμM simulations were shown to the participants where three members of a group select the 'divide-and-stay' plan. This time, we highlighted a particular group of agents in each simulation (with a white circle). Participants were asked to describe their impression and understanding of character behaviors in the animations they saw. The rationale for this third study condition is the notion of 'narrative intelligence' whereby one's belief concerning the truth of a phenomenon is dependent on one's ability to explain the phenomenon (Bruner, 1991; Mateas and Sengers, 2003). At the end of our 3-part study, the participants were given a semi-structured interview to gain better insight for how they judged the realism and plausibility of the simulations.

Figure 6.11 shows some animation scenes from the pair of CM and CμM which was used for this task. Characters start from the same initial positions (Figure 6.11 (a)) and the focused groups select the divide-and-stay plan (Figure 6.11 (c)). Before a split occurs, the characters exchange communicative actions in the CμM condition (bottom of Figure 6.11 (b)) while they simply leave each other in the CM condition (top of Figure 6.11 (b)).

6.3.4 Procedure

2 volunteers (28 females, 14 males), aged 18 to 38, were recruited for the study. At the beginning of the study, we showed them a demo video of our virtual airport terminal with a large number of virtual characters.

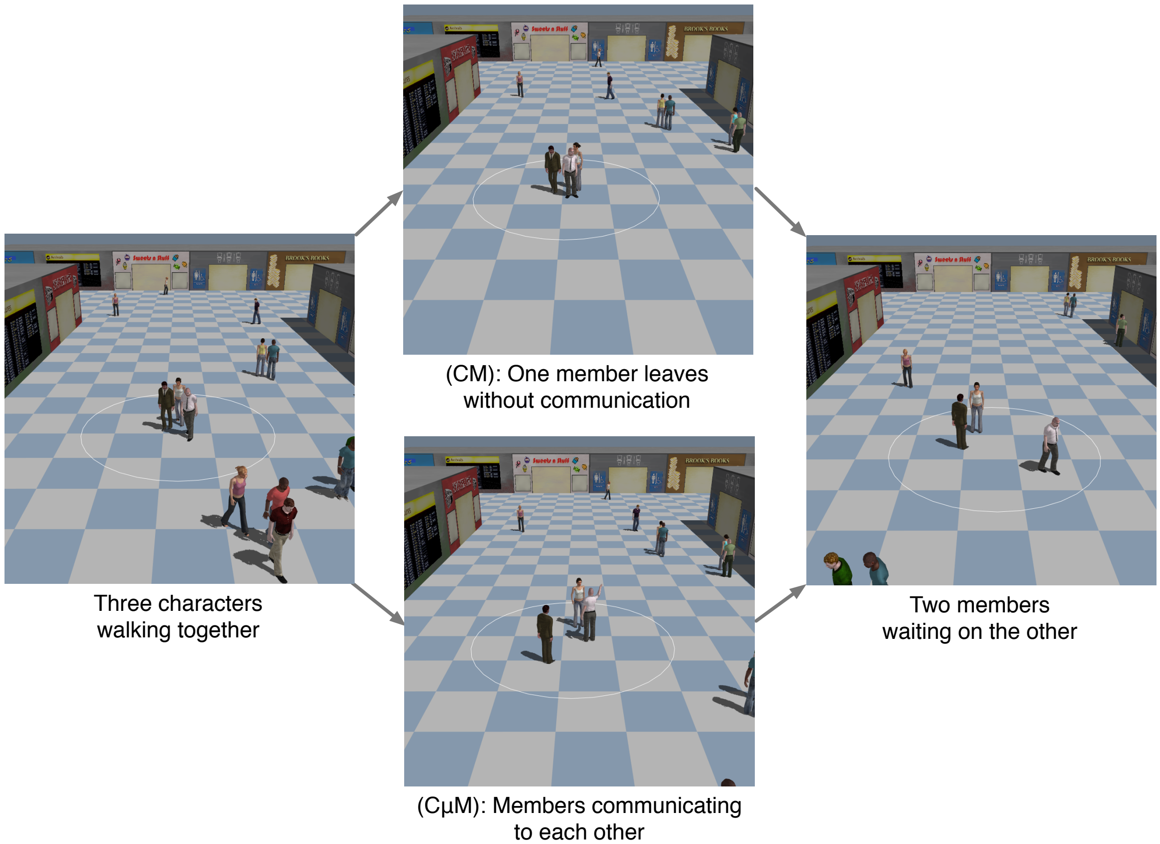


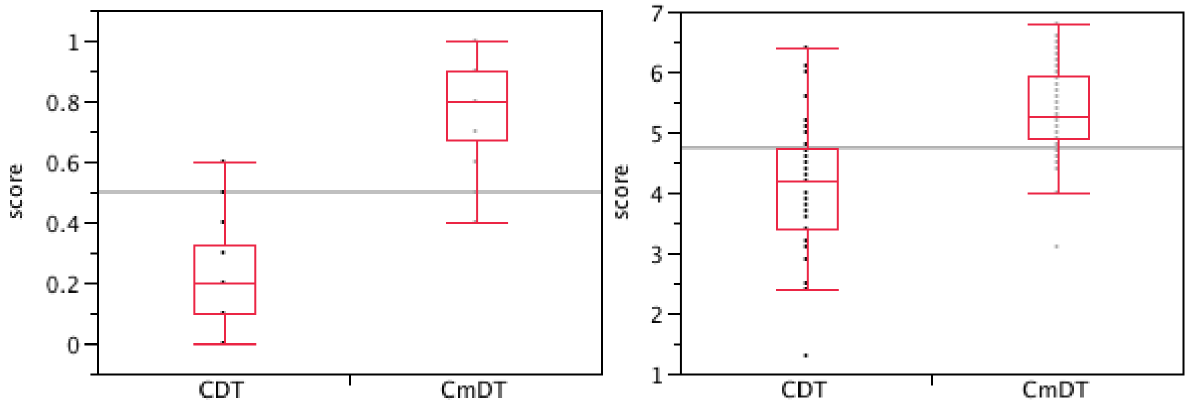
Figure 6.11: A paired animation of CM and CμM conditions

6.3.5 Results

Quantitative Analysis

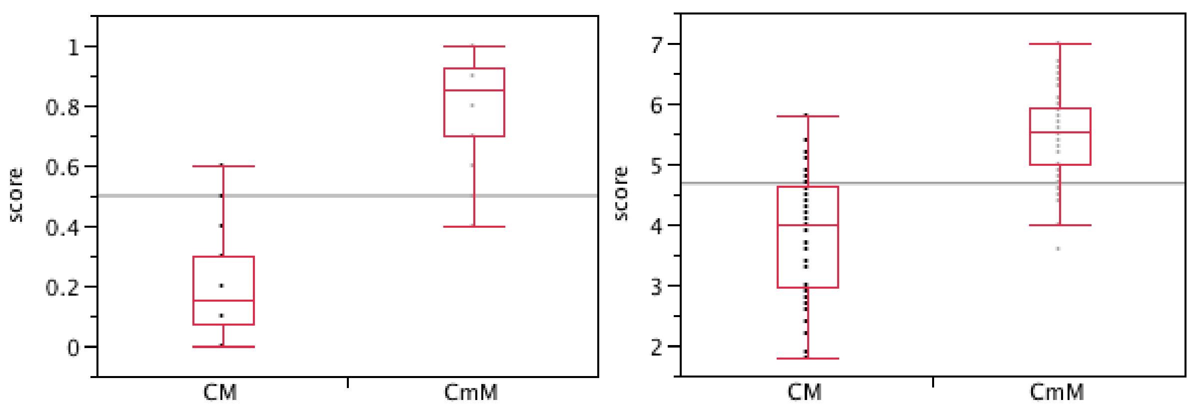
We hypothesized that groups employing μ-coordination would appear more realistic and believable than the groups without the μ-coordination. Specifically, we hypothesized that groups in CμDT and CμM simulations would score higher $\mathcal{M}_{\mathcal{R}}$ and $\mathcal{M}_{\mathcal{P}}$ than those in CDT and CM simulations, respectively. Paired t-tests were conducted to compare $\mathcal{M}_{\mathcal{R}}$ values for CDT and CμDT, and CM and CμM animations, and $\mathcal{M}_{\mathcal{P}}$ values for CDT and CμDT, and CM and CμM animations.

We found a significant effect of incorporating social behaviors of coordination in the participants' responses on the crowd animations. Figures 6.12 (a) and (b) show that participants chose the CμDT groups as real more often than the CDT groups ($p < 0.01$), and CμDT groups as more plausible than the CDT groups ($p < 0.01$). The analysis results in Figures 6.13 (a) and (b) also confirm that participants rated the CμM groups as real more often than the CM groups ($p < 0.01$), and CμM groups as more plausible than the CM groups ($p < 0.01$).



(a) $\mathcal{M}_\mathcal{R}$ (b) $\mathcal{M}_\mathcal{P}$

Figure 6.12: CDT vs CμDT



(a) $\mathcal{M}_\mathcal{R}$ (b) $\mathcal{M}_\mathcal{P}$

Figure 6.13: CM vs CμM

Qualitative Analysis

We employed two qualitative approaches to analyze our qualitative data. First, we analyzed the participants' responses for objective statements of belief concerning each model in the third task. This allows us to determine why one model was judged as more believable than another. For the CM animation, 10 participants stated that they were not sure on what was going on in the animation (e.g., "it seems strange, the people randomly stop and one person leaves"). 24 participants provided just a factual description of what they witnessed without providing any reasons for what they saw, for example, "one guy in a suit walked away while others are standing." 10 among these 34 indicated that they thought the characters are not with together (e.g., "I think the first two are traveling together, and the guy at the end wasn't with them"). 8 subjects made an interpretation in which they assumed the group members communicated before they split up.

In contrast, for the CμM animation, 40 subjects subjects indicated that they had a better understanding of the character behaviors, and interpreted the split as resulting from negotiation (e.g., "one of the character actually told the other two characters to wait on him"). 10 of the 40 subjects explicitly stated that the CμM animation was much more clear, and it was due to the exhibition of communicative acts of characters (e.g., "there was obvious communication between the members of the group so it was very direct and I didn't have to assume what was going on"), and 11 of them added more stories into their description (e.g., "he might say something like 'do you know where we have to go,' so he checks..."). 2 provided the similar factual description to what they had for the CM animation.

In the semi-structured post interview, we asked participants what criteria they used to evaluate animations. To obtain categories for the rationale for the participants' beliefs, we employed an open coding method (Strauss and Corbin, 1990). We performed two passes through the data. In the first pass, we collected categories of responses concerning belief, and in the second, we employed these categories to group the participants' responses.

Through our analysis, we were able to determine three categories for reasons that the participants rated the CG model as more believable. The three categories of rationale are: *Interaction*, *Formation*, and *Cohesion*. In the *Interaction* category, participants were attentive to whether there was evidence of interaction among characters before stopping or changing direction while walking (e.g., exchange of gestures, body alignments to talk). Subjects answered that the characters with the μ-coordination behaved in a way that allows them to structure the sequence of behaviors into a narrative whole, and makes the animation be more comprehensible and believable. For the *Formation* category, spatial patterns of groups such as side-by-side walking and linear walking formation were considered. In the *Cohesion* category, subjects looked for whether group members maintained appropriate proximity and/or respected group integrity (i.e. circumnavigated other groups rather than just cutting through them). As shown in Figure 6.14, 34 participants employed *Interaction*,

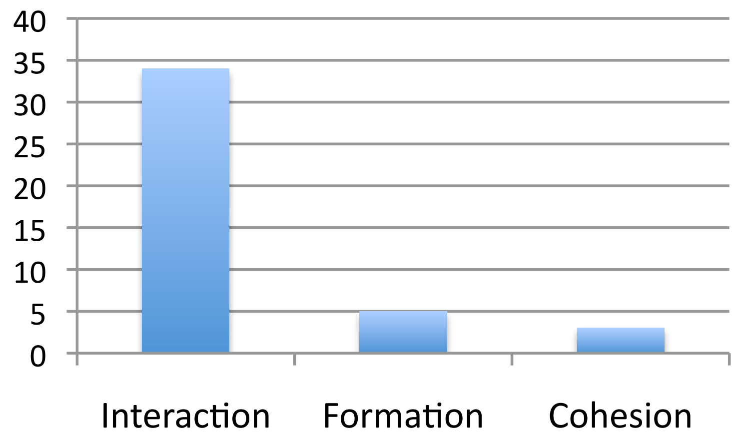


Figure 6.14: Criteria used to evaluate animations

5 used *Formation*, and 3 used *Cohesion* as their criteria to make their evaluations.

Our results demonstrate that the believability of an animation is affected by communicative and social interactions among characters. People are attentive to not only what the characters do but also why, because they try to understand the chains of character behaviors by constructing a coherent story (Bruner, 1991). People rated a given animation realistic and more plausible when they thought a specific walking strategy (e.g., divide-and-stay) of a group was made as a result of communication among the characters. This indicates that the meaning of behaviors of individuals in group activities is not decidable in isolation but people relate the behavior to the behavior of interacting entities to understand them. It is shown that the comprehensibility may be essential to believable agents, and this suggests that the model of character behaviors should be designed to provide interpretability, or rationality, to external observers in order for achieving the enhanced realism.

6.3.6 Discussion

We validated the outcomes of our CGCS model with human judgement with respect to the realism, plausibility, and interpretability of animation. Although our method provided a means to evaluate the character behaviors, it is limited by several drawbacks. In order to use our approach, it requires to represent the simulation results with animated characters. However, depending on the objective of simulations, it could be more effective and efficient to employ simplified representations and evaluate some other aspects of the simulation results. For instance, if we need to examine the model effects on the simulation of evacuation planning scenario, it would be sufficient to employ point particles for the data representation, and evaluate the result by comparing the trajectory of the simulated evacuees with that of real human crowds from a video recording as in (Seyfried et al., 2007; Guy et al., 2010; Karamouzas and Overmars, 2012). Or, simply statistical measures such as egress rate and time could be compared to those of the real world data without visualization.

Another limitation of our approach is that, the human visual perception and evaluation of animation may be influenced by the fidelity and quality of animated characters. Humans are sensitive to artifacts and anomalies in animated characters and their behaviors as discussed in (Skrba and O'Sullivan, 2009; McDonnell et al., 2009; Larkin and O'Sullivan, 2011). If some gesture animation primitives appear unnatural in a rendering engine side, it may be interpreted in an unintended way and result in a negative evaluation of the simulation and animation.

6.4 Scalability Evaluation

We evaluated the scalability of our approach with respect to the number of groups and total agents. The simulation was tested on a desktop with an Intel i7 3.20 GHz CPU, and 4GB system memory. Frames per second are measured in the CI (for Individuals), CDT (for always 'Detour-Together'), CM (for Macro only), and CμM (for both micro and Macro CSs) configurations.

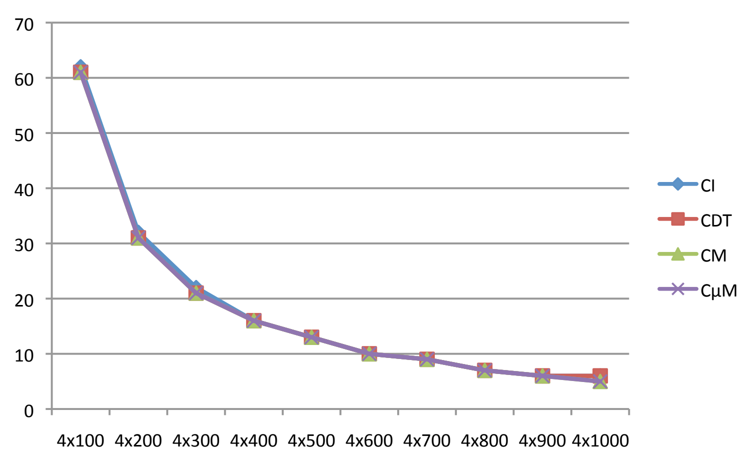


Figure 6.15: Frames per second measured in the CI, CDT, CM, and CμM conditions

As shown in Figure 6.15, the computational overhead for group coordination simulation in CμM is not significant. Taking 15 fps as the minimum interactive rate, the simulation with our group coordination model runs at an interactive rate up to 2,000 of agents of 500 groups.

6.5 Summary

We conducted a series of studies and showed notable differences resulted in the simulations quantitatively and qualitatively when our common ground model is enabled.

Especially, we found that a non-CG simulation to always 'out-perform' a CG-based model with regard to level of congestion. The difference of the congestion level among the different crowd models in our experiment demonstrates the significance of incorporating realistic human social interactions into an agent design to achieve valid simulation results. For example, suppose that we were designing a mall with certain target number and placement of shops for a target number of visitors. Underestimations of crowd complexity may cause designers to be overly optimistic about the size of corridors or of evacuation rates in an emergency.

The CGCS model enabled our agents to present communicative behavioral cues in coordination with other agents in a group. Because of the coherence of the behaviors emerged among the group members, human observers were able to construct a coherent story about the group interaction, and that contributed to their judgement of the animated scene as 'real'. We conducted a series of user studies in which the efficacy of our CGCS model was examined. The study results showed that the communicative purpose in our model can be consistently carried through from simulation to animation, and it produces more believable behaviors of animated characters from the viewpoint of human observers. Our study results inform that the 'interpretability' and 'explainability' should be concerned in the development of crowd models in order to achieve enhanced realism in crowd simulation and animation.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

Incorporating the sense of social intelligence for virtual characters is important to achieve valid simulation data and plausible aggregate behaviors in crowd simulation and animation. The goal of this dissertation was to address the need for modeling social group interactions in order to reproduce realistic crowd behaviors in a simulation. Overall, the broader contributions of this research include the computational model informed by CG theory for simulating social group behaviors, unified crowd simulation framework for accommodating high-level socially-aware behavioral realism of animated characters, and understanding how communicative behavior within individual groups can impact the distribution of the simulated crowd as a whole and perceived realism of the animation by the observers.

7.2 Crowd Model and Operationalization

We presented a crowd model founded on CG theory to incorporate the impact of social interaction among group members in the crowd simulation. In particular, consideration of how each agent evaluates a particular action, path, and location depending on the ongoing communication and coordination among group members allowed us to reproduce realistic group dynamics in a pedestrian simulation. The conceptual crowd model was operationalized for a multi-agent system, and it was shown that our operationalization is efficient and scalable for large numbers of groups and individuals.

Our model includes micro- and macro-coordination by which behaviors of groups are embedded into a larger crowd simulation in a space rich with interaction possibilities for the agents. Micro-coordination strategies relate to CG theory, and hence to the dynamics of human communicative behavior. Macro-coordination strategies relate to simulation contexts. This two-level approach allowed our model to adapt to varying application scenarios and domains.

We focused on simulating the coordination process established by CG theory rather than modeling the process of social interaction at a lower level of communicative intent to reasoning. This intermediate level simulation made the computation tractable while still maintaining realism in the simulation. Since CG theory is a well-researched scientific theory, and already takes the effect of real human micro-level negotiation of joint activity into consideration, any bottom-up micro-negotiation model was hard pressed to match the realism of a CG model.

7.3 Unified Framework for Crowd Simulation and Animation

We presented a unified crowd simulation framework that can produce socially plausible animation behaviors of virtual human characters. In our simulation model agents were designed to present the kind of behavioral evidences according to their communicative purpose, and maintain group cohesiveness by coordinating and adapting their behaviors to each other. This model provided a means to simulate purposive behavior of human groups in interacting in the animation. By integrating simulation and rendering engines into a unified crowd modeling framework, the motion selection information generated in the simulation model was transferred to the rendering engine at each frame, and resulted sequence of character behaviors formed a socially meaningful, hence plausible interactions.

7.4 Evaluation of Crowd Model

We conducted a series of experiments for investigating the impact of our social group interaction model on a simulated crowd. We found that a CG-based model always causes higher congestion than a non-CG simulation. For agents of CG-based model, the pragmatic need to maintain common ground incurred costs at the level of the entire simulation. This result indicates that the behavioral cost of performing the necessary micro-coordination activity (i.e., communicative activity) cannot be ignored in overall crowd simulation with multiple agents. For example, if we were designing a mall with certain target number and placement of shops for a target number of visitors, such underestimations of crowd complexity may cause designers to be overly optimistic about the size of corridors or of evacuation rates in an emergency.

In narrative psychology, the notion of 'narrative intelligence' was advanced whereby one's belief concerning the truth of a phenomenon is dependent on one's ability to explain the phenomenon (Bruner, 1991; Mateas and Sengers, 2003). We employed this concept to test our operationalization of the CG-based crowd simulation model by determining whether the model yields purposive interpretations of the resulting animation. Through a series of perceptual user evaluation studies, we demonstrated the interpretability of the character behaviors is associated with the believability of the animation by observers. People were attentive to not only what the characters do but also why, because they tried to understand the chains of character behaviors by constructing a coherent story (Bruner, 1991). People rated a given animation realistic and more plausible when they thought a specific walking strategy (e.g., divide-and-stay) of a group was made as a result of communication among the characters. This indicates that the meaning of behaviors of individuals in group activities is not decidable in isolation but people relate the behavior to the behavior of interacting entities to understand them. It was shown that the comprehensibility may be essential to believable agents, and this suggests that the model of character behaviors should be designed to provide interpretability, or rationality, to external observers in order for achieving the enhanced realism.

7.5 Future Work

Currently our model does not handle coordination failures and recovery scenarios. However, perfection in communication and coordination among people is impossible to attain in real life situations. For example, group members may fail to communicate with each other because they have been blocked by other people in a highly congested environment. In another situation, people may be distracted by an unanticipated event which attracts their immediate attention, and may fail to address the signals of their group members. In these cases, people who need to initiate communication, request attention, or deliver their intention may retry the coordination with their group members by using an alternative means of communication. For instance, at a failure of receiving an acknowledgment from friends in a noisy and crowded place, one may move towards and tap them on the shoulder instead of beckoning or calling out. Another example of coordination failure may be that people miscommunicate the location to meet up or visit as a sub-goal. In such a case, some member may get lost and the rest of people in the group will need to look for the lost one. While looking for the missing person, the remaining members have to ensure not to get separated and lost each other again, hence it will require a sub-coordination plan among them. In order to account these coordination failure situations, additional set of activity rules can be designed and added to the current set of micro- and macro-coordination strategies.

Another future research direction includes extending our model to handle subgroups and various types of relationships. In the real world, individuals are embedded in different social structures simultaneously, such as subgroups (e.g., parents, siblings), groups (e.g., family), and organization (e.g., pedestrians). Also, intergroup ties could vary (e.g., pedestrians-pedestrians, pedestrians-authority figures). The different social relationships may have an impact on the use of micro-coordination and macro-coordination strategies. For example, pedestrian-pedestrian coordination will require the construction of ad hoc proximal groups with different strategies. Such extension will provide interesting challenges to extensions of our model. It is also promising to apply our model to different simulation contexts by providing a pertinent repertoire of macro-behaviors to a target application, for example, a military coordination training scenario.

There are several possible improvements for unified our crowd modeling framework on a presentation side. First, the actions of the characters were repetitive and not always smooth enough because of the lack of variety in motions in the database. It is desirable to add a richer set of communicative gestures and movements for micro-coordination behaviors (e.g., ‘looking at’ by head movement) that characters can display. Second, motion synthesis techniques could be applied to the rendering engine for achieving good connectivity and smooth transitions in characters motions. Third, visualizing a large number of animated characters is a heavy computational process, therefore a dynamic level-of-detail approach to reduce the workload for the stages of graphics pipeline could be employed to increase the rendering efficiency.

Bibliography

Antonini, G., Bierlaire, M., and Weber, M. (2006). Discrete choice models of pedestrian walking behavior. *Transportation Research Part B: Methodological*, 40(8):667 – 687.

Arkin, R. C. (1990). Integrating behavioral, perceptual, and world knowledge in reactive navigation. *Robot. Auton. Syst.*, 6:105–122.

Aveni, A. F. (1977). The Not-So-Lonely Crowd: Friendship Groups in Collective Behavior. *Sociometry*, 40(1):96–99.

Bouchet, F. and Sansonnet, J. (2011). Influence of personality traits on the rational process of cognitive agents. In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2011 IEEE/WIC/ACM International Conference on*, volume 2, pages 81 –88.

Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry*, 52(4):664–678.

Brogan, D. C. and Hodgins, J. K. (1997). Group behaviors for systems with significant dynamics. *Auton. Robots*, 4:137–153.

Brogan, D. C. and Johnson, N. L. (2003). Realistic human walking paths. In *Proc. of the 16th Intl. Conf. on Computer Animation and Social Agents (CASA 2003)*, CASA '03, pages 94–101. IEEE Computer Society.

Bruner, J. (1991). The narrative construction of reality. *Critical inquiry*, 18(1):1–21.

Cassell, J., Vilhjálmsson, H. H., and Bickmore, T. (2001). Beat: the behavior expression animation toolkit. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '01, pages 477–486, New York, NY, USA. ACM.

Castellano, G., Mancini, M., Peters, C., and McOwan, P. (2012). Expressive copying behavior for social agents: A perceptual analysis. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 42(3):776 –783.

Chenney, S. (2004). Flow tiles. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '04, pages 233–242, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Cho, K., Iketani, N., Kikuchi, M., Nishimura, K., Hayashi, H., and Hattori, M. (2008). Bdi model-based crowd simulation. In *Proceedings of the 8th international conference on Intelligent Virtual Agents*, IVA '08, pages 364–371, Berlin, Heidelberg. Springer-Verlag.

Clark, H. H. (1996). *Using Language*. Cambridge University Press.

Clark, H. H. and Brennan, S. E. (1991). Grounding in communication. In Resnick, L. B., Levine, J. M., and Teasley, S. D., editors, *Perspectives on Socially Shared Cognition*, pages 127–149. American Psychological Association, Washington DC.

Coleman, J. S. and James, J. (1961). The equilibrium size distribution of freely-forming groups. *Sociometry*, 24(1):pp. 36–45.

Convertino, G., Mentis, H. M., Rosson, M. B., Carroll, J. M., Slavkovic, A., and Ganoe, C. H. (2008). Articulating common ground in cooperative work: content and process. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1637–1646, New York, NY, USA. ACM.

Endrass, B., André, E., Rehm, M., Lipi, A. A., and Nakano, Y. (2011). Culture-related differences in aspects of behavior for virtual characters across germany and japan. In *The 10th Intl. Conf. on Autonomous Agents and Multiagent Systems*, volume 2 of *AAMAS '11*, pages 441–448.

Ennis, C., McDonnell, R., and O'Sullivan, C. (2010). Seeing is believing: body motion dominates in multisensory conversations. *ACM Trans. Graph.*, 29:1–9.

Ennis, C. and O'Sullivan, C. (2012). Perceptually plausible formations for virtual conversers. *Computer Animation and Virtual Worlds*, 23(3-4):321–329.

Ennis, C., Peters, C., and O'Sullivan, C. (2008). Perceptual evaluation of position and orientation context rules for pedestrian formations. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV '08, pages 75–82, New York, NY, USA. ACM.

Ennis, C., Peters, C., and O'Sullivan, C. (2011). Perceptual effects of scene context and viewpoint for virtual pedestrian crowds. *ACM Trans. Appl. Percept.*, 8(2):10:1–10:22.

Festinger, L. (1954). *A Theory of Social Comparison Processes*. Bobbs-Merrill reprint series in the social sciences. Bobbs-Merrill.

Fiorini, P. and Shillert, Z. (1998). Motion planning in dynamic environments using velocity obstacles. *International Journal of Robotics Research*, 17:760–772.

Fridman, N., Zilka, A., and Kaminka, G. A. (2012). The impact of cultural differences on crowd dynamics in pedestrian and evacuation domains. In *In Proceedings of the Eleventh International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-12)*, page An Extended Abstract.

Funge, J., Tu, X., and Terzopoulos, D. (1999). Cognitive modeling: knowledge, reasoning and planning for intelligent characters. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '99, pages 29–38, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.

Gu, Q. and Deng, Z. (2011). Context-aware motion diversification for crowd simulation. *IEEE Computer Graphics and Applications*, 31:54–65.

Guy, S. J., Chhugani, J., Kim, C., Satish, N., Lin, M., Manocha, D., and Dubey, P. (2009). Clearpath: highly parallel collision avoidance for multi-agent simulation. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '09, pages 177–187, New York, NY, USA. ACM.

Guy, S. J., Curtis, S., Lin, M. C., and Manocha, D. (2010). Pledestrians - a least-effort approach to crowd simulation. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation 2010*.

Guy, S. J., Curtis, S., Lin, M. C., and Manocha, D. (2012). Least-effort trajectories lead to emergent crowd behaviors. *Phys. Rev. E*, 85:016110.

Guy, S. J., Kim, S., Lin, M. C., and Manocha, D. (2011). Simulating heterogeneous crowd behaviors using personality trait theory. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '11, pages 43–52, New York, NY, USA. ACM.

Hart, P., Nilsson, N., and Raphael, B. (1968). A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107.

Helbing, D., Farkas, I., and Vicsek, T. (2000). Simulating dynamical features of escape panic. *Nature*, 407(6803):487–490.

Helbing, D., Moln, P., Farkas, I. J., and Bolay, K. (2001). Self-organizing pedestrian movement. *Environment and Planning B: Planning and Design*, 28(3):361–383.

Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics. *PHYSICAL REVIEW E*, 51:4282.

Hoogendoorn, M. and Soumokil, J. (2010). Evaluation of virtual agents utilizing theory of mind in a real time action game. In van der Hoek, Kaminka, Lespérance, Luck, and Sen, editors, *Proceedings of the 9th Intl. Conf. on Autonomous Agents and Multiagent Systems*, volume 1 of *AAMAS '10*, pages 59–66, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Hoogendoorn, S. P. and Bovy, P. H. L. (2004). Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169 – 190.

Hughes, R. L. (2002). A continuum theory for the flow of pedestrians. *Transportation Research Part B: Methodological*, 36(6):507 – 535.

James, J. (1953). The Distribution of Free-Forming Small Group Size. *American Sociological Review*, 18(5):569–570.

Jan, D., Herrera, D., Martinovski, B., Novick, D., and Traum, D. (2007). A computational model of culture-specific conversational behavior. In *Proceedings of the 7th international conference on Intelligent Virtual Agents*, IVA '07, pages 45–56, Berlin, Heidelberg. Springer-Verlag.

Jarabo, A., Eyck, T. V., Sundstedt, V., Bala, K., Gutierrez, D., and O'Sullivan, C. (2012). Crowd light: Evaluating the perceived fidelity of illuminated dynamic scenes. *Comp. Graph. Forum*, 31(2pt4):565–574.

Jin, X., Xu, J., Wang, C. C. L., Huang, S., and Zhang, J. (2008). Interactive control of large-crowd navigation in virtual environments using vector fields. *IEEE Comput. Graph. Appl.*, 28:37–46.

Ju, E., Choi, M. G., Park, M., Lee, J., Lee, K. H., and Takahashi, S. (2010). Morphable crowds. *ACM Trans. Graph.*, 29(6):140:1–140:10.

Kachroo, P. (2008). *Pedestrian Dynamics*. Understanding Complex Systems. Springer-Verlag Berlin Heidelberg.

Kachroo, P. (2009). *Pedestrian Dynamics: Mathematical Theory and Evacuation Control*. CRC Press.

Kaminka, G. and Fridman, N. (2006). A cognitive model of crowd behavior based on social comparison theory. In *Proceedings of the AAAI-2006 workshop on cognitive modeling*.

Kamphuis, A. and Overmars, M. H. (2004). Finding paths for coherent groups using clearance. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '04, pages 19–28, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Kapadia, M., Singh, S., Hewlett, W., and Faloutsos, P. (2009). Egocentric affordance fields in pedestrian steering. In *Proceedings of the 2009 symposium on Interactive 3D graphics and games*, I3D '09, pages 215–223, New York, NY, USA. ACM.

Kapadia, M., Wang, M., Singh, S., Reinman, G., and Faloutsos, P. (2011). Scenario space: characterizing coverage, quality, and failure of steering algorithms. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '11, pages 53–62, New York, NY, USA. ACM.

Karamouzas, I. and Overmars, M. (2010). Simulating the local behaviour of small pedestrian groups. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, VRST '10, pages 183–190, New York, NY, USA. ACM.

Karamouzas, I. and Overmars, M. (2012). Simulating and evaluating the local behavior of small pedestrian groups. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):394–406.

Kim, M., Hwang, Y., Hyun, K., and Lee, J. (2012a). Tiling motion patches. In *Proceedings of the 2012 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '12, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Kim, S., Guy, S. J., Manocha, D., and Lin, M. C. (2012b). Interactive simulation of dynamic crowd behaviors using general adaptation syndrome theory. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, I3D '12, pages 55–62, New York, NY, USA. ACM.

Kirby, R., Simmons, R., and Forlizzi, J. (2009). Companion: A constraint-optimizing method for person-acceptable navigation. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 607 –612.

Klein, G., Feltovich, P. J., Bradshaw, J. M., and Woods, D. D. (2005). Common ground and coordination in joint activity. In Rouse, W. B. and Boff, K. R., editors, *Organizational Simulation*, pages 139–184. John Wiley and Sons, Inc.

Knowles, E. S. (1973). Boundaries around group interaction: The effect of group size and member status on boundary permeability. *Journal of Personality and Social Psychology*, 26(3):327 – 331.

Knowles, E. S., Kreuser, B., Haas, S., and Hyde, M. (1976). Group size and the extension of social space boundaries. *Journal of Personality and Social Psychology*, 33(5):647 – 654.

Koh, W. L. and Zhou, S. (2011). Modeling and simulation of pedestrian behaviors in crowded places. *ACM Trans. Model. Comput. Simul.*, 21(3):20:1–20:23.

Kretz, T., Kaufman, M., and Schreckenberg, M. (2008). Counterflow extension for the f.a.s.t.-model. In Umeo, H., Morishita, S., Nishinari, K., Komatsuzaki, T., and Bandini, S., editors, *Cellular Automata*, volume 5191 of *Lecture Notes in Computer Science*, pages 555–558. Springer Berlin Heidelberg.

Lai, Y.-C., Chenney, S., and Fan, S. (2005). Group motion graphs. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '05, pages 281–290, New York, NY, USA. ACM.

Larkin, M. and O'Sullivan, C. (2011). Perception of simplification artifacts for animated characters. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, APGV '11, pages 93–100, New York, NY, USA. ACM.

Lau, M. and Kuffner, J. J. (2005). Behavior planning for character animation. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '05, pages 271–280, New York, NY, USA. ACM.

Lee, K. H., Choi, M. G., Hong, Q., and Lee, J. (2007). Group behavior from video: a data-driven approach to crowd simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '07, pages 109–118, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Lee, K. H., Choi, M. G., and Lee, J. (2006). Motion patches: building blocks for virtual environments annotated with motion data. *ACM Trans. Graph.*, 25(3):898–906.

Lemercier, S., Jelic, A., Kulpa, R., Hua, J., Fehrenbach, J., Degond, P., Appert-Rolland, C., Donikian, S., and Pettre, J. (2012). Realistic following behaviors for crowd simulation. *Comp. Graph. Forum*, 31(2pt2):489–498.

Li, Y., Wang, T., and Shum, H.-Y. (2002). Motion texture: a two-level statistical model for character motion synthesis. *ACM Trans. Graph.*, 21(3):465–472.

Liu, C. K., Hertzmann, A., and Popović, Z. (2006). Composition of complex optimal multi-character motions. In *Proceedings of the 2006 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '06, pages 215–222, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Luo, L., Zhou, S., Cai, W., Lees, M., and Low, M. Y. H. (2010). Modeling human-like decision making for virtual agents in time-critical situations. In *Proceedings of the 2010 International Conference on Cyberworlds*, CW '10, pages 360–367, Washington, DC, USA. IEEE Computer Society.

Luo, L., Zhou, S., Cai, W., Low, M. Y. H., Tian, F., Wang, Y., Xiao, X., and Chen, D. (2008). Agent-based human behavior modeling for crowd simulation. *Computer Animation and Virtual Worlds*, 19(3-4):271–281.

Mateas, M. and Sengers, P. (2003). *Narrative Intelligence*. Advances in Consciousness Research Series. John Benjamins Pub.

McDonnell, R., Larkin, M., Dobbyn, S., Collins, S., and O'Sullivan, C. (2008). Clone attack! perception of crowd variety. *ACM Trans. Graph.*, 27(3):26:1–26:8.

McDonnell, R., Larkin, M., Hernández, B., Rudomin, I., and O'Sullivan, C. (2009). Eye-catching crowds: saliency based selective variation. *ACM Trans. Graph.*, 28(3):55:1–55:10.

Monk, A. (2003). Common Ground in Electronically Mediated Communication: Clark's Theory of Language Use. In Carroll, J., editor, *HCI Models, Theories and Frameworks: Toward a Multi-Disciplinary Science*, chapter 10, pages 265–290. Morgan Kaufmann.

Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., and Theraulaz, G. (2010). The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics. *PLoS ONE*, 5(4):e10047+.

Musse, S. R. and Thalmann, D. (1997). A Model of Human Crowd Behavior: Group Inter-Relationship and Collision Detection Analysis. In *Workshop Computer Animation and Simulation of Eurographics*, pages 39–52.

Narain, R., Golas, A., Curtis, S., and Lin, M. C. (2009). Aggregate dynamics for dense crowd simulation. *ACM Trans. Graph.*, 28:122:1–122:8.

Neale, D. C., Carroll, J. M., and Rosson, M. B. (2004). Evaluating computer-supported cooperative work: models and frameworks. In Herbsleb, J. D. and Olson, G. M., editors, *Proc. of the 2004 ACM conf. on Computer supported cooperative work*, CSCW '04, pages 112–121, New York, NY, USA. ACM.

Nickolls, J., Buck, I., Garland, M., and Skadron, K. (2008). Scalable parallel programming with cuda. *Queue*, 6:40–53.

Pelachaud, C. (2005). Multimodal expressive embodied conversational agents. In *Proceedings of the 13th annual ACM international conference on Multimedia*, MULTIMEDIA '05, pages 683–689, New York, NY, USA. ACM.

Pelechano, N. (2005). Crowd simulation incorporating agent psychological models, roles and communication. In *First International Workshop on Crowd Simulation*, pages 21–30.

Pelechano, N., Allbeck, J. M., and Badler, N. I. (2007). Controlling individual agents in high-density crowd simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '07, pages 99–108, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Peters, C. and Ennis, C. (2009). Modeling groups of plausible virtual pedestrians. *IEEE Comput. Graph. Appl.*, 29(4):54–63.

Peters, C., Ennis, C., McDonnell, R., and O'Sullivan, C. (2008). Crowds in Context: Evaluating the Perceptual Plausibility of Pedestrian Orientations. pages 33–36.

Pettré, J., Laumond, J.-P., and Siméon, T. (2003). A 2-stages locomotion planner for digital actors. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '03, pages 258–264, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Pettré, J., Ondřej, J., Olivier, A.-H., Cretual, A., and Donikian, S. (2009). 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, pages 189–198, New York, NY, USA. ACM.

Qiu, F. and Hu, X. (2010). Modeling group structures in pedestrian crowd simulation. *Simulation Modelling Practice and Theory*, 18(2):190 – 205.

Rao, A. S. and Georgeff, M. P. (1995). Bdi-agents: from theory to practice. In Lesser, V. R. and Gasser, L., editors, *Proceedings of the First Intl. Conference on Multiagent Systems*, pages 312–319, San Francisco. The MIT Press.

Reynolds, C. (1999). Steering Behaviors for Autonomous Characters. In *Game Developers Conference 1999*.

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.*, 21:25–34.

Seyfried, A., Steffen, B., Klingsch, W., Lippert, T., and Boltes, M. (2007). The fundamental diagram of pedestrian movement revisited — empirical results and modelling. In Schadschneider, A., Pöschel, T., Kühne, R., Schreckenberg, M., and Wolf, D., editors, *Traffic and Granular Flow'05*, pages 305–314. Springer Berlin Heidelberg.

Shao, W. and et al. (2005). Environmental modeling for autonomous virtual pedestrians.

Shao, W. and Terzopoulos, D. (2007). Autonomous pedestrians. *Graphical Models*, 69(5-6):246 – 274. Special Issue on SCA 2005.

Singh, S., Kapadia, M., Faloutsos, P., and Reinman, G. (2009). SteerBench: a benchmark suite for evaluating steering behaviors. *Computer Animation and Virtual Worlds*, 9999(9999):n/a+.

Skrba, L. and O'Sullivan, C. (2009). Human perception of quadruped motion. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, APGV '09, pages 130–130, New York, NY, USA. ACM.

Sommer, R. (1979). *Personal space*. Prentice Hall.

Stentz, A. and Mellon, I. C. (1993). Optimal and efficient path planning for unknown and dynamic environments. *International Journal of Robotics and Automation*, 10:89–100.

Stone, M., DeCarlo, D., Oh, I., Rodriguez, C., Stere, A., Lees, A., and Bregler, C. (2004). Speaking with hands: creating animated conversational characters from recordings of human performance. *ACM Trans. Graph.*, 23(3):506–513.

Strauss, A. and Corbin, J. (1990). *Basics of qualitative research: grounded theory procedures and techniques*. Sage Publications.

Sun, X., Koenig, S., and Yeoh, W. (2008). Generalized adaptive a*. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 1*, AAMAS '08, pages 469–476, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Takahashi, S., Yoshida, K., Kwon, T., Lee, K. H., Lee, J., and Shin, S. Y. (2009). Spectral-based group formation control. *Computer Graphics Forum*, 28(2):639–648.

Thomas, G. and Donikian, S. (2000). Virtual humans animation in informed urban environments. In *Proceedings of the Computer Animation*, CA '00, pages 112–119, Washington, DC, USA. IEEE Computer Society.

Treuille, A., Cooper, S., and Popović, Z. (2006). Continuum crowds. *ACM Trans. Graph.*, 25(3):1160–1168.

Trovato, K. I. and Dorst, L. (2002). Differential a*. *IEEE Trans. on Knowl. and Data Eng.*, 14:1218–1229.

Tsai, J., Fridman, N., Bowring, E., Brown, M., Epstein, S., Kaminka, G., Marsella, S., Ogden, A., Rika, I., Sheel, A., Taylor, M. E., Wang, X., Zilka, A., and Tambe, M. (2011). Escapes: evacuation simulation with children, authorities, parents, emotions, and social comparison. In *The 10th Intl. Conf. on Autonomous Agents and Multiagent Systems*, volume 2 of *AAMAS '11*, pages 457–464.

van den Berg, J., Guy, S. J., Snape, J., Lin, M. C., and Manocha, D. (2011). Rvo2 library: Reciprocal collision avoidance for real-time multi-agent simulation. `http://gamma.cs.unc.edu/RVO2/`.

van den Berg, J., Lin, M., and Manocha, D. (2008a). Reciprocal velocity obstacles for real-time multi-agent navigation. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pages 1928 –1935.

van den Berg, J., Patil, S., Sewall, J., Manocha, D., and Lin, M. (2008b). Interactive navigation of multiple agents in crowded environments. In *Proceedings of the 2008 symposium on Interactive 3D graphics and games*, number 9 in I3D '08, pages 139–147. ACM.

Wampler, K., Andersen, E., Herbst, E., Lee, Y., and Popović, Z. (2010). Character animation in two-player adversarial games. *ACM Trans. Graph.*, 29(3):26:1–26:13.

Warren, C. (1989). Global path planning using artificial potential fields. In *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on*, pages 316 –321 vol.1.

Yeh, H., Curtis, S., Patil, S., van den Berg, J., Manocha, D., and Lin, M. (2008). Composite agents. In *Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '08, pages 39–47, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Yersin, B., Maïm, J., Morini, F., and Thalmann, D. (2008). Real-time crowd motion planning: Scalable avoidance and group behavior. *Vis. Comput.*, 24:859–870.

Zuckerman, I., Kraus, S., and Rosenschein, J. S. (2012). The adversarial activity model for bounded rational agents. *Autonomous Agents and Multi-Agent Systems*, 24(3):374–409.

Appendix A: Experiment Document for Perceptual Study of Modeling Social Groups for Virtual Crowds

Office of Research Compliance
Institutational Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, VA 24060
540/231-4606 Fax 540/231-0959
email irb@vt.edu
website http://www.irb.vt.edu

MEMORANDUM

DATE: November 6, 2012

TO: Francis Quek, Seung In Park

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires May 31, 2014)

PROTOCOL TITLE: Perceptual Effects of Modeling Social Groups for Virtual Crowds

IRB NUMBER: **12-959**

Effective November 6, 2012, the Virginia Tech Institution Review Board (IRB) Administrator, Carmen T Green, approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As:	**Expedited, under 45 CFR 46.110 category(ies) 7**
Protocol Approval Date:	**November 6, 2012**
Protocol Expiration Date:	**November 5, 2013**
Continuing Review Due Date*:	**October 22, 2013**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Date*	OSP Number	Sponsor	Grant Comparison Conducted?

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the IRB office (irbadmin@vt.edu) immediately.

PROJECT TITLE: Perceptual Effects of Modeling Social Groups for Virtual Crowds

INVESTIGATORS: Seung In Park, Francis Quek, Ph.D.

I. PURPOSE OF THE STUDY: The purpose of the study is to gather some information about how **small-group activity** in a crowd may affect realism and believability in a crowd simulation.

II. PROCEDURES:

Prior to the start of the study, you will fill a pre-questionnaire in which you will be asked simple questions on your experience with crowd simulations and computer animations. Then you will be informed about what crowd simulation is, and given a sample virtual-crowd video.

In the tasks that follow, you will see several pairs of animations in which virtual human characters walk around in a generic airport terminal setting. As you know, behavior in an airport may include going to destinations, taking detours to check flight information, going to the bathroom etc. The behavior may include where individuals are facing and pointing and nodding behavior. Our animation model does not include luggage graphics for simplicity.

The pairs of animation compares behavior collected from real data against behavior that is completely synthesized by the computer. In the real-data condition, we collected crowd motion data including the paths taken by individuals and groups and behaviors (like pointing and body orientation) from a large set of crowd videos. The paths and behaviors of the individuals in the simulation mirrors the data we collected. In the other condition, all paths and behaviors were computer-generated.

For each pair of animations we want to know if you can tell the animation produced from real natural data from the synthetic one. You may base your evaluation on how plausible the behavior is in a real airport scene. For some animations, we want to know your interpretation of the character behaviors in the observed scenes. Please note that the characters in either condition may appear 'wooden' because our animation may not be smooth enough. This woodeness should not affect your judgment. We are interested in overall behavior.

There is no "right" or "wrong" answer to any response. We just want to know what you think. The whole process will take around 1 hour.

III. RISKS: There is no potential risk or harm of any form to study participants.

IV. BENEFITS: In the simulation of human crowd behavior including evacuation planning, transportation management, and safety engineering in architecture design, the development of crowd model for higher behavior fidelity is an important issue. The findings of this study will contribute to increase the understanding of how a crowd model should be designed to capture the real crowd effects. In addition to that, study participants may be able to experience current crowd simulation and computer animation technologies that they normally do not have access to.

V. CONFIDENTIALITY: You will be identified only with an identifier code on all documents related to the study, and we will treat all data as confidential. Your personal details such as name and your identifier code will be kept in locked cabinets and destroyed 5 years after the end of the study.

VI. COMPENSATION: You will be given one extra credit for the research participation if you are enrolled in a class that is accepting extra credit and have signed up for this experiment using the Sona system.

VII. FREEDOM TO WITHDRAW: Participation in this study is entirely voluntary. You may withdraw without any penalty at any time throughout the study period should you feel the need to do so.

VIII. APPROVAL OF RESEARCH: This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects of Virginia Polytechnic Institute and State University.

IX. PARTICIPANT'S CONSENT: I have read and understood the Information Sheet and conditions of this project. I hereby acknowledge the above and give my voluntary consent for participation in this project by signing this form.

NAME (Print): ________________________________
Signature: ___________________________________
Date: _______________________________________
Email: ______________________________________
Phone: ______________________________________

Should I have any pertinent questions about this research or its conduct, and research subjects' rights, and whom to contact in the event of a research-related injury to the subject, I may contact:

Investigator: Seung In Park, spark80@vt.edu
Faculty Advisor: Dr. Francis Quek: (540) 231-8453, quek@cs.vt.edu
Human Research Protections: Dr. David Moore: (540) 231-4991,
moored@vt.edu, 2000 Kraft Drive, Suite 2000. Blacksburg, VA 24060

Experimental Instructions

We are interested in gathering some information about how **small-group activity** in a crowd may affect realism/believability in a crowd simulation. In the tasks that follow, you will see several pairs of animations in which virtual human characters walk around in a generic airport terminal setting. As you know, behavior in an airport may include going to destinations, taking detours to check flight information, going to the restroom etc. The behavior may include where individuals are facing and pointing and nodding behavior.

The pairs of animation compares behavior collected from **real data** against behavior that is completely synthesized by the computer. In the real-data condition, we collected crowd motion data including the paths taken by individuals and groups and behaviors (like pointing and body orientation) from a large set of crowd videos. The paths and behaviors of the individuals in the simulation mirrors the data we collected. In the other condition, all paths and behaviors were **computer-generated**.

For each pair of animations we want to know if you can tell the animation produced from real natural data from the synthetic one. You may base your evaluation on how plausible the behavior is in a real airport scene. For some animations, we want to know your interpretation of the character behaviors in the observed scenes. Please note that the characters in either condition may appear 'wooden' because our animation may not be smooth enough. This woodeness should not affect your judgment. We are interested in overall behavior.

There is no "right" or "wrong" answer to any response. We just want to know what you think. The whole process will take around 1 hour.

Part 1. For each pair of the animations,

A. Please select which animation you think is produced from **real natural data**.

B. Please rate each animation on the level of plausibility regarding **overall small-group activities in a crowd** in Likert scale from 1 to 7.

Set 1:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 2:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 3:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 4:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 5:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 6:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 7:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 8:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 9:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 10:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Part 2. For each pair of the animations,

A. Please select which animation you think is produced from **real natural data**.

B. Please rate each animation on the level of plausibility regarding **overall small-group activities in a crowd** in Likert scale from 1 to 7.

Set 1:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 2:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 3:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 4:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 5:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 6:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 7:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 8:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 9:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Set 10:
A. a) b)

B. a) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

b) 1: not at all ----- 2 ----- 3 ----- 4: neutral ----- 5----- 6 ----- 7: very much

Part 3. Can you interpret what is happening among the highlighted characters in the animation? If so, please describe your interpretation. If not, what is the reason for your answer?

a)

b)