
Functional Crowds

Jan M. Allbeck

Abstract

Most crowd simulation research either focuses on navigating characters through an environment while avoiding collisions or on simulating very large crowds. Functional crowds research focuses on creating populations that inhabit a space as opposed to passing through it. Characters exhibit behaviors that are typical for their setting, including interactions with objects in the environment and each other. A key element of this work is ensuring that these large-scale simulations are easy to create and modify. Automating the inclusion of action and object semantics can increase the level at which instructions are given. To scale to large populations, behavior selection mechanisms must be kept relatively simple and, to demonstrate typical human behavior, must be based on sound psychological models. The creation of roles, groups, and demographics can also facilitate behavior selection. The simulation of functional crowds necessitates research in animation, artificial intelligence, psychology, and human-computer interaction (HCI). This chapter provides a brief introduction to each of these elements and their application to functional crowds.

Keywords

Crowd simulation • Virtual humans • Patterns of life • Computer animation • AI

Contents

Introduction	2234
State of the Art	2235
Animation to AI	2236
Heterogeneity	2238
Human-Computer Interaction	2238

J.M. Allbeck (✉)
George Mason University, Fairfax, VA, USA
e-mail: jallbeck@gmu.edu

Conclusions	2239
Cross-References	2240
References	2241

Introduction

Virtual humans can be used as stand-ins when using real humans would be too dangerous and cost-prohibitive or precise control is required. Virtual humans are often used as extras or background characters in movies and games (see Fig. 1). They are similarly used in virtual training scenarios for military personnel and first responders. They can also be used to analyze urban and architectural design as well as various policies and procedures. For many of these applications and others, the virtual humans must both reflect typical or normal human behavior and also be controllable or directable. Furthermore, in order to create sizeable crowds of virtual humans functioning in rich virtual environments, they must have relatively simple behavior selection mechanisms.

Functional crowds, in contrast to more typical crowd simulations, depict animated characters interacting with the environment in meaningful ways. They do not simply walk from one location to another avoiding obstacles. They perform the same behaviors we see from real humans every day, as well as not so typical behaviors that might be required for the application.

The first element needed to achieve functional crowds is animation. Traditional crowd simulations focus on walking animation clips, perhaps with a few idle behaviors or depending on the application some battle moves. There is little or no interaction with objects in the environment. Animating virtual humans manipulating objects can be quite challenging. It involves detection of collisions and fine motor movements. We will give an overview of some of these challenges and approaches for solving them in this chapter.



Fig. 1 Virtual characters in a scene in the unreal game engine

Another required element relates to providing the virtual humans with the knowledge of what actions can be performed and what objects are required to perform them. If we are going to eat, we need an object corresponding to food to eat. We may optionally need instruments such as utensils. A lot of this needed information could be considered commonsense, but unless explicitly supplied to the virtual humans, they lack it. This information is also needed as input to higher-level artificial intelligence mechanisms such as behavior selectors, planners, and narrative generators.

Functional crowds should also depict a heterogeneous population. In real life not everyone does the same thing, they do not have the same priorities, and they do not all perform tasks in exactly the same way. Some of these variations stem from prior observations and experiences. They are learned. Some stem from psychological states and traits, such as personalities and emotions.

Finally, many, if not all, applications of functional crowds require some human-computer interaction (HCI). This interaction may come during the authoring of the crowd behavior. The application or scenario may require some of the behaviors to be more tightly controlled or even scripted. The application may also require users (e.g., players, trainees, evaluators, etc.) to interact with the crowd during the simulation. These interactions may simply require the virtual humans to avoid collisions with the real human's avatar, or they may require communication and perhaps even cooperation between the real and virtual humans.

This chapter will present these various elements of functional crowds and discuss challenges and approaches to address them. We will start by providing a snapshot of the current state of the art in related research fields. Then we will in turn discuss issues related to AI and animation, psychological models, and HCI. Finally, we will conclude with a brief summary and potential future direction.

State of the Art

In the past decade or so, crowd simulations have made enormous progress in the number of characters that can be simulated and in creating more natural behaviors. More detailed analysis of crowd simulation research can be found in a number of published volumes (Kapadia et al. 2015; Pelechano et al. 2008, 2016; Thalmann et al. 2007). It is now possible to simulate over one million characters in real time in high-density crowds.

Crowd simulations can also be more heterogeneous. Not every character looks or behaves exactly the same. Certainly some variations stem from differences in appearance and motion clips (Feng et al. 2015; McDonnell et al. 2008). Others come from psychological models such as emotion and personality (Balint and Allbeck 2014; Durupinar et al. 2016; Li et al. 2012). Most crowd simulations assign, fairly randomly, starting positions and ending destinations for the characters in the simulation. While this appears fine for short durations at a distance, if a player follows a character for a period of time, it quickly appears false. Sunshine-Hill and

Badler have created a framework for generated plausible destinations for characters on the fly to provide reasonable “alibis” for them (Sunshine-Hill and Badler 2010).

Simulating functional crowds also requires other advanced computer graphics techniques. Commercial game engines, such as Unity[®] and Unreal[®], provide much of the technology necessary. In the past couple of years, they have both changed their licensing structure in ways that enable researchers to take advantage of and add to their capabilities. Other needed advancements come from the animation research community. A key feature of functional crowds is the ability of characters to interact with objects in their environment in meaningful ways. We require animations of characters sitting and eating food, getting in and out of vehicles, conversing with one another, displaying emotions, getting dressed, etc. (Bai et al. 2012; Clegg et al. 2015; Hyde et al. 2016; Shapiro 2011).

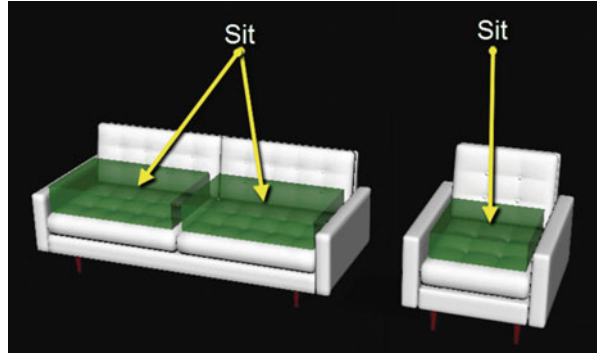
Animation to AI

To simulate a functional crowd, we need the characters to interact with their object-rich environments and with each other. While great work has been done in pathfinding, navigation, and path following, additional advancements are still needed (Kapadia et al. 2015; Pelechano et al. 2016). Characters still struggle to get through cluttered environments with narrow walkways. We need to give characters enhanced abilities to turn sideways, sidestep, and even back up in seamless natural motions.

Furthermore, characters need to be able to grab, carry, place, and use objects of different shapes and sizes and do so when the objects are placed at various locations in the world and approached from any direction. The core of motions for characters is generally generated in one of three ways: key framed, motion capture, or procedural. Artist created key-framed and motion-captured motions that tend to look natural and expressive, but lack the flexibility needed for most object interactions. Procedurally generated motions use algorithms such as inverse kinematics that work well to target object locations (e.g., for a reach and grab), but often lack a natural look and feel and require objects to be labeled with sites, regions, and directions referenced in the code. While progress continues in virtual human animation research, natural-looking functional crowds will require even more advancement to make the authoring and animating large populations of characters more feasible.

Once the characters can be animated interacting with objects in the environment, they need to possess an understanding of what can be done with objects and what objects are needed in order to perform various actions. In other words, they need to understand object and action semantics. This includes knowing what world state must exist prior to the start of an action (i.e., applicability conditions and preparatory specifications), what state must hold during the action, what the execution steps of the action are, and finally what the new world state is after the successful execution of the action. As indicated previously, there also needs to be information about the parts and various locations of the objects (e.g., grasp locations, regions to sit on (see Fig. 2), etc.) so that animation components can be effective. Representations, such as the Parameterized Action Representation (PAR), are designed to hold this

Fig. 2 Regions indicating places where characters could sit



information, but authoring them is time-consuming and error prone (Bindiganavale et al. 2000).

In order to scale to the level needed to simulate functional crowds in large, complex environments, the creation of action and object semantics needs to be automated. Automating action and object semantics would also help to ensure some consistency within and between scenarios, whereas ad hoc, handing authoring tends to be sloppy and error prone. Online lexical databases, such as WordNet, VerbNet, and FrameNet, have been shown to provide a viable foundation for action and object semantics for virtual worlds (Balint and Allbeck 2015; Pelkey and Allbeck 2014). Additional work is needed to ensure the information represented is what is needed for the applications in virtual worlds and to ensure that mechanisms for searching and retrieval are fast enough.

Given that characters have some basic understanding of the virtual world they are inhabiting, the next question is at any given time, how should characters select their behaviors? Planning and other sophisticated AI techniques can be computationally intensive and difficult to control. For functional crowds, it would be better to start with simple techniques both in authoring and execution (J. M. Allbeck 2009, 2010). Human behaviors stem from a variety of different impetuses. Some behaviors, such as going to work or school or attending a meeting, are *scheduled*. These actions provide some structure to our lives and the lives of our virtual counterparts. They are selected based on the simulated time of day. *Reactive* actions are responses to the world around us. They add life and variation to the behaviors of virtual characters. They are selected based on the objects, people, and events around the character. *Aleatoric* or stochastic actions include sub-actions with different distributions. For example, we may want a character to appear to be working in her office, but are not very concerned with the details. Our *WorkInOffice* action would include sub-actions like talking on the phone, filing papers, and using the computer. The character would switch between these actions for the specified period of time at the specified distribution, but what exact sub-action is being performed at any point in time would not need to be specified. *Need-based* actions add humanity to the virtual characters. Needs grow over time and are satisfied by performing certain actions with the necessary object participants (e.g., eat food). As a need grows, the priority

of selecting a behavior that would fulfill it also grows. These needs could correspond to established psychological models, such as Maslow's hierarchy of needs, or they could be specific to the scenario (e.g., drug addiction).

These are just a few examples of simple behavior selection mechanisms. Certainly, others are possible and may be more applicable to some scenarios. Practically speaking, it may be best to completely script the behaviors of some key characters in a scenario. Background characters could then have variations in their schedules, reactions, needs and distributions. More sophisticated AI techniques could be included when and where needed, as long as the overall framework remains fast enough for human interaction.

Heterogeneity

In real human populations, not everyone is doing the same thing at the same time. There are variations in behaviors that stem from different factors. The psychological research committee has spent decades positing numerous models of personality, emotions, roles, status, culture, and more. The virtual human research community has taken these models as inspiration for computational models for virtual human behaviors (Allbeck and Badler 2001; IVA 1998; Li and Allbeck 2011). Variations in behavior and behavior selection can also evolve as the characters learn about and from their environment and each other (Li and Allbeck 2012).

All of this research needs additional attention and revision. In particular, how these different traits are manifest in expressive animation needs continuing work, as does the interplay of psychological models. How does personality effect emotion and the display of emotion? How does a character's roles and changing role effect emotional displays? Certainly culture and its impacts are not well modeled in virtual humans. How do all of these psychological models influence a character's priorities? At any point in time, a character's behavior selection should reflect what is most important for them to achieve at that time. Their priorities can be influenced by any number of factors. For functional crowds, it is important that priorities be weighed quickly and behavior selection is not delayed by an overly complex psychological framework. An open question for most scenarios is what parts of human behavior are really important to model and what can be left out? It is possible that a fair amount of just random choices would suffice for the majority of the characters a lot of the time, but this depends on the duration of the simulation and how often the same character or characters are encountered by the viewer.

Human-Computer Interaction

Most applications of functional crowds require them to have some interaction with real humans either during the authoring process and/or while the simulation unfolds. Authoring the behavior of an entire population of characters from the ground up would be infeasible. Providing a layer of automatically generated common

understanding (i.e., action and object semantics) does help. Simple, yet robust, behavior selection mechanisms are also helpful. Furthermore, the action types described earlier can be linked to even higher-level constructs, such as groups and roles (Li and Allbeck 2011).

When authoring behaviors, it is important to balance autonomy and control. To accomplish the objectives of the scenario, authors need to have control over the characters and their behaviors. However, authoring every element of every behavior of every character would be overwhelming even for short-duration simulations of forty or fifty characters. The characters need to have some level of autonomy. They need to be able to decide what to do and how to do it on their own. Then, when and if they receive instruction from the simulation author, they need to suspend or preempt their current behaviors to follow those instructions. There may also be times when authors have an overall narrative in mind for the simulations, but are less concerned about some of the details of the characters' behaviors. This is one place where more sophisticated AI methods like partial planners may play a role (Kapadia et al. 2015).

HCI also comes into play as one or more humans interact with the functional crowds during the simulation. They may be using a standard keyboard, mouse, and monitor. They may be using a mobile device. They may be using a gaming console. Or they may be using more advanced virtual reality (VR) devices. VR devices can provide a higher fidelity and therefore enable the subjects to see the virtual world in more detail. Using head-mounted displays (HMD) or CAVE systems allows the subject to view the virtual characters in a life-size format. The movements of subjects can also be motion captured in real time and displayed on their avatar, providing more realistic interaction with the virtual characters. Hardware interfaces can impact the level of a subject's immersion into a virtual world and potentially their level of presence in the virtual world.

Another aspect of HCI with virtual characters and functional crowds is a kind of history. If a subject spends longer durations in the virtual world and/or has repeated exposure to it, he or she may become familiar with some of the characters and form expectations about them. Subjects may learn their personality and behavioral quirks. Subject will expect some consistency in these behaviors. They may also expect the virtual characters to have some level of memory of past interactions. While these expectations can be met, it is still a challenge to provide the virtual characters with techniques that make these memories compact, efficient, and plausibly imperfect (Li et al. 2013). More research is needed.

Conclusions

Functional crowds can increase the number of applications of crowd simulations and increase their accuracy, but as this chapter has discussed, there is additional research needed from character animation to AI to psychological models to HCI. Increased computing power will help, but is not an overall solution. Attempting to simulate realistic human behaviors is difficult. It is even more challenging at a large scale. When attempting to simulate realistic human behavior, we can end up losing focus.

One model or technique leads us to another and another until we have lost sight of our original goal. Too often researchers also design and implement a method and then go in search of a problem it might address. We might be better served to keep focused on an application or scenario and then determine what is and is not most critical to achieving its goals. Does the application really require a sophisticated planner or emotion model? How closely and for how long is the subject going to be observing the characters' behaviors? Also, do we really need to simulate 500,000 characters at a time? At ground level in the center of a village or even large city, how many people can be seen at one time? Are there existing tools, open source or commercial, that can be used or modified? Too often researchers feel they have to construct their own models from scratch, ignoring years of effort done by others. In terms of both human effort and computation, use available resources wisely and do not put a large amount of effort into areas that will have little impact on the application.

In this area of research, another question that is often asked is how do you *validate* your model. How can one validate human behavior? We could show videos of functional crowds to hundreds of people and ask them a variety of questions to try to determine if they think the character behaviors are realistic, reasonable, or even plausible, but we all have rather different ideas of what is reasonable behavior. Instead we choose to framework work in this area as the construction of a toolset to be used by subject matter experts to achieve their own goals. For example, an urban planner may wish to use functional crowds to analyze a proposed transportation system. Evaluate then becomes about whether or not the urban planner can use the functional crowds toolset to do the desired analysis. Does it have the parameters required? Is it usable by nonprogrammers? Can they increase or decrease fidelity relative to the input effort?

As a research area, functional crowds is a young, but promising direction. It sits at the overlap of several other research communities, namely, computer graphics and animation, artificial intelligence, human-computer interaction, and psychology. As advances are made in each one of these disciplines, functional crowds can benefit.

Cross-References

- ▶ [Biped Controller for Character Animation](#)
- ▶ [Blendshape Facial Animation](#)
- ▶ [Crowd Evacuation Simulation](#)
- ▶ [Crowd Formation Generation and Control](#)
- ▶ [Data-Driven Character Animation Synthesis](#)
- ▶ [Data-Driven Hand Animation Synthesis](#)
- ▶ [Depth Sensor-Based Facial and Body Animation Control](#)
- ▶ [Example-Based Skinning Animation](#)
- ▶ [Eye Animation](#)
- ▶ [Hand Gesture Synthesis for Conversational Characters](#)
- ▶ [Head Motion Generation](#)

- ▶ [Laughter Animation Generation](#)
- ▶ [Perceptual Study on Facial Expressions](#)
- ▶ [Physically Based Character Animation Synthesis](#)
- ▶ [Real-Time Full Body Motion Control](#)
- ▶ [Real-Time Full-Body Pose Synthesis and Editing](#)
- ▶ [Video-Based Performance Driven Facial Animation](#)
- ▶ [Visual Speech Animation](#)

References

- Allbeck JM (2009) Creating 3D animated human behaviors for virtual worlds. University of Pennsylvania, Philadelphia, PA
- Allbeck JM (2010) CAROSA: a tool for authoring NPCs. In: Presented at the international conference on motion in games. Springer, pp 182–193.
- Allbeck J, Badler NI (2001) Consistent Communication with Control. In: Workshop on non-verbal and Verbal Communicative Acts to achieve contextual embodied agents at autonomous agents.
- Bai Y, Siu K, Liu CK (2012) Synthesis of concurrent object manipulation tasks. *ACM Trans Graph* 31(6):156
- Balint T, Allbeck JM (2014) Is that how everyone really feels? emotional contagion with masking for virtual crowds. In: Presented at the international conference on intelligent virtual agents. Springer, pp 26–35
- Balint T, Allbeck JM (2015) Automated generation of plausible agent object interactions. In: Presented at the international conference on intelligent virtual agents. Springer, pp 295–309
- Bindiganavale R, Schuler W, Allbeck JM, Badler NI, Joshi AK, Palmer M (2000) Dynamically altering agent behaviors using natural language instructions. In: Proceedings of the fourth international conference on autonomous agents. ACM, New York, pp 293–300. <https://doi.org/10.1145/336595.337503>
- Clegg A, Tan J, Turk G, Liu CK (2015) Animating human dressing. *ACM Trans Graph* 34(4):116
- Durupinar F, Gudukbay U, Aman A, Badler N (2016) Psychological parameters for crowd simulation: from audiences to Mobs. *IEEE Trans Vis Comput Graph* 22(9):2145–2159
- Feng A, Casas D, Shapiro A (2015) Avatar reshaping and automatic rigging using a deformable model. In: Presented at the proceedings of the 8th ACM SIGGRAPH conference on motion in games. ACM, pp 57–64
- Hyde J, Carter E, Kiesler S, Hodgins J (2016) Evaluating animated characters: facial motion magnitude influences personality perceptions. *ACM Trans Appl Percept* 13(2):8:1–8:17
- IVA (1998) International conference on intelligent virtual humans. Springer, Berlin
- Kapadia M, Pelechano N, Allbeck J, Badler N (2015) Virtual crowds: steps toward behavioral realism. Morgan & Claypool Publishers, San Rafael, California
- Li W, Allbeck JM (2011) Populations with purpose. In: *Motion in games*. Springer, Berlin/Heidelberg, pp 132–143
- Li W, Allbeck JM (2012) Virtual humans: evolving with common sense. In: Presented at the international conference on motion in games. Springer, pp 182–193
- Li W, Di Z, Allbeck JM (2012) Crowd distribution and location preference. *Comput Anim Virtual Worlds* 23(3–4):343–351
- Li WP, Balint T, Allbeck JM (2013). Using a parameterized memory model to modulate NPC AI. In: Presented at the intelligent virtual agents: 13th international conference, IVA 2013, Edinburgh, August 29–31, 2013, Proceedings, vol 8108. Springer, p 1
- McDonnell R, Larkin M, Dobbyn S, Collins S, O’Sullivan C (2008) Clone attack! perception of crowd variety. In: Presented at the ACM Transactions on Graphics (TOG), vol 27. ACM, p 26

- Pelechano N, Allbeck JM, Badler NI (2008) Virtual crowds: methods, simulation, and control. *Synth Lect Comput Graph Animation* 3(1):1–176
- Pelechano N, Allbeck JM, Kapadia M, Badler NI (eds) (2016) *Simulating heterogeneous crowds with interactive behaviors*. CRC Press, Boca Raton, FL
- Pelkey CD, Allbeck JM (2014) Populating semantic virtual environments. *Computer Animation and Virtual Worlds* 25(3–4):403–410
- Shapiro A (2011) Building a character animation system. In: Presented at the international conference on motion in games. Springer, pp 98–109
- Sunshine-Hill B, Badler NI (2010) Perceptually realistic behavior through alibi generation. In *Proceedings of the Sixth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE'10)*, G. Michael Youngblood and Vadim Bulitko (Eds.). AAAI Press 83–88
- Thalmann D, Musse SR, Braun A (2007) *Crowd simulation*, vol 1. Springer, Berlin