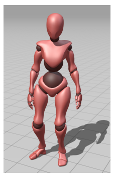
I work on **real-time optimization** algorithms for **virtual media systems**, such as **Virtual, Augmented and Mixed Reality**. My research results are in **interactive**, **human-centered design** algorithms for large virtual environments[10,11], and **simulation** of virtual **AI agents** and crowds in these environments[4,12,13]. My work has been recognized by the ACM SIGGRAPH conference on Motion in Games, where I received the **Best Paper Award**[12]. In addition, I was a **finalist** in the **ACM SIGGRAPH Thesis Fast Forward** 2018, and a **finalist** in the **ACM SIGGRAPH Asia Doctoral Consortium** 2018. I also have **industry research** experience, in **Amazon**, **Autodesk Research** and **Wayfair**. In the following sections, I summarize my current research, and describe my long-term research goals.

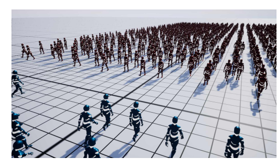
1. Simulating the Motion of Multi-Agent Crowds



Agent

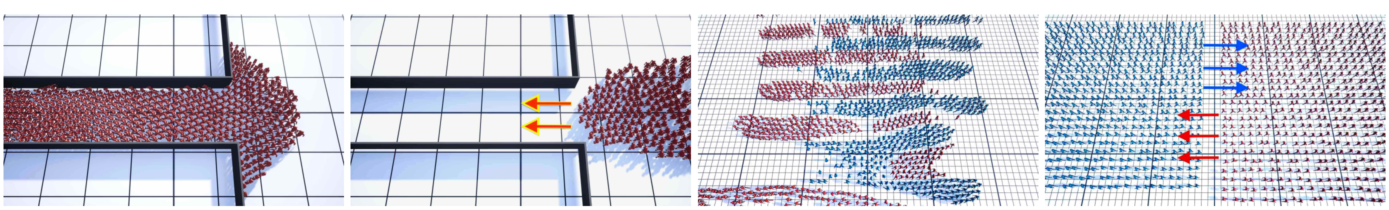
Crowd simulation is an essential part of multiple domains, from interactive media systems, pedestrian analysis in evacuation scenarios, urban planning, to autonomous multi-agent path planning and coordination. A **crowd** is a **collection of independent, self-actuated agents**. Each agent has individual navigational goals in this shared environment. Since agents share the same environment, they can interact and collide with each other. Agent movement is controlled by a navigation algorithm, which needs to ensure that an **agent progresses towards its goal, while avoiding collisions**. However, computing collision-free agent motion is difficult, due to the complexity of such dynamic interactions. Simulation allows quick insight into the dynamics and performance of a navigation algorithm, for a given experiment. Visualizing experiments allows us to capture high-level details, as well as group behaviors that might not be otherwise noticeable. Performing such experiments is more efficient and less costly than real-world experiments, and allows us to easily test "what-if" scenarios.

Simulations are composed of a series of discrete time steps. Agents run through a continuous cycle of sensing and acting, where each cycle correlates to a time step. At the beginning of each cycle, each agent independently computes a trajectory to its goal, while avoiding collisions with other agents or obstacles. In addition to collision-free movement, a crowd simulation method should **capture both individual and group behavior** observed in real crowds, while being computationally interactive.



Multi-Agent Crowds

Despite 30+ years in crowd simulation research, simulation methods have subspecialized, and are **computationally effective for either sparse or dense crowds, but not for both**[5]. For example, methods for dense crowds smooth out individual agent motions, while methods for sparse crowds cannot computationally cope with dense situations. Since most simulation scenarios contain varying crowd densities, neither of the existing approaches are robust, scalable, or uniform.



(a) A group of agents passing through a narrow corridor. **(b)** Two groups exchanging positions.

Figure 1: Emergent phenomena in crowds. My result[12], is able to recreate phenomena found in real crowds, without any scripting or other user-directed control. These phenomena include: (a) clogging and arching near bottlenecks[2], (b) groups self-organizing into lanes[1], with stable and real-time performance for 100,000+ individuals. This result is enabled by a modeling the motion of crowds as a constraint optimization problem, which is then efficiently solved with a GPU.

My work is the first method[12] that allows real-time simulation of **both dense and sparse crowds**. This result is made possible by reframing agent motion as a **constrained optimization problem**. Motion is controlled by numerical constraints on agent positions. For example, a constraint for collision avoidance extrapolates each agent's trajectory to detect future collisions. In case of a collision, agents adjust their trajectory according to the direction that **minimizes the constraint violation**, with the magnitude of the change depending on the amount of time until an impending collision[3]. Otherwise, agents locomote directly to their goal.

All computations are processed **efficiently with a GPU**, allowing **real-time motion for 100,000+ agents**. This is achievable since agent motion constraints are solved independently, and therefore are easily parallelizable (Fig. 1). For this result, I received the **Best Paper Award** in the ACM SIGGRAPH conference on Motion in Games 2017[12].

Aside from the immediate practical implications for the gaming and visual effects industry, my results inspire future research on large-scale simulation of complex behavior, that would allow to tackle current challenges in Crowd Simulation[7], including heterogeneous crowds that exhibit different gaits, and locomotion styles.

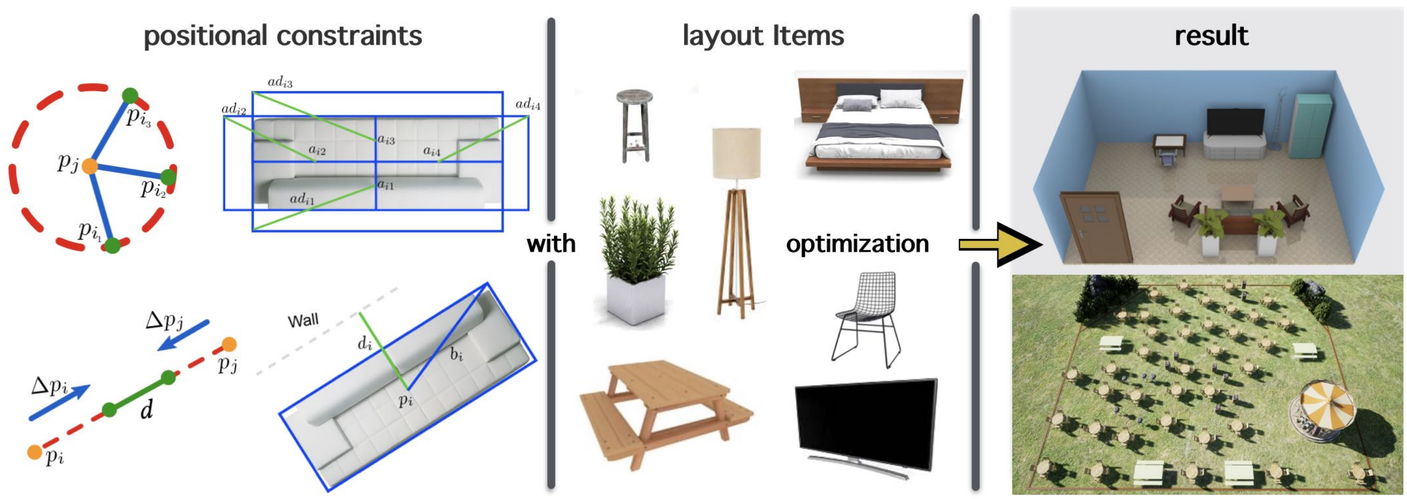


Figure 2: My result[10] in 3D scene synthesis is faster by an order of magnitude than previous probabilistic methods, and is the first to allow a **real-time, interactive, synthesis of scenes which were previously intractable**. I show how to reformulate layout synthesis in terms of a constrained optimization problem, which is then solved by a fast, physics-inspired algorithm I developed. Given input positional constraints and selected layout items, the algorithm rapidly outputs a variety of synthesized layouts.

2. Computational Design of Virtual and Real Worlds

Virtual worlds are growing in terms of complexity and interactivity, and are **challenging to construct efficiently and realistically**. Interior scenes are an intrinsic part of virtual worlds, and are prevalent in interactive media. They are also needed for non-interactive applications, such as furniture marketing, residential showcasing, and many other similar commercial applications.

Currently, the aesthetic, and creative process of designing interiors, real or virtual, is mostly manual, and done by professionals with extensive training in 3D modeling software. To accelerate this process, researchers have proposed several methods for interior scene synthesis. These methods usually adopt a stochastic optimization scheme, which samples from a distribution of layout configurations. Unfortunately, such methods are slow and inefficient, and not suitable for interactive design. The novelty of my work[10] is in enabling **real-time, interactive synthesis of layouts**, even for **large scale** scenes, which **were previously intractable**[6]. This is achieved by viewing layout synthesis from the prism of simulating deformable bodies. Both layouts and deformable bodies can be described by geometric constraints, which

Figure 3: My result in fast layout synthesis[14] enables interactive furnishing of real spaces. Given an input image of an interior or exterior space, a general user specification of the desired furnishings, and layout constraints, my method automatically furnishes a scene and displays it to the user by augmenting the original image.

should be satisfied for a layout to be realistic, or for the movement of a body to be plausible. Layout constraints describe distances, orientations, and formations of layout items, while in deformable bodies, constraints describe the shape and points of the body. Both can be tackled using continuous optimization procedures. A layout synthesis solution is then an arrangement that satisfies such geometric constraints (Fig. 2).

There are other benefits stemming from this research breakthrough. For example, we can rapidly synthesize new datasets that can be used for training computer vision systems. Another application is to **interactively visualize furniture inside images** of empty interior, and exterior spaces (Fig. 3). Such interactivity allows rapid visual assessment of various layout designs, a process which would typically be time consuming if done manually, or with previous algorithmic approaches. This is especially **valuable for users** with mobile phones and other camera-enabled devices. A future step in this direction would be to develop an Augmented Reality tool for interactive decoration of spaces[8].

In addition to synthesizing interiors, I developed software for synthesizing **large scale architecture**. During my tenure at **Autodesk Research** labs, I contributed to the development of an architectectual floor plan synthesis system. The input to the system is a user sketch, and the output is a floor plan, which the user can further interact, edit, and optimize.

What other research problems do these 3D scenes allow us to explore? Currently, such scenes do not embed possible camera positions or light conditions, but such qualities affect the **mood and perception of a scene**. Another research question is why how do we perceptually rank layouts? To address these questions, I will employ **crowd-sourced learning** to rank, and perceptually understand layouts. Such data, combined with statistical and machine learning, can lead to **new insights** on interior scene synthesis.

Architectectual synthesis system I developed at Autodesk Research.

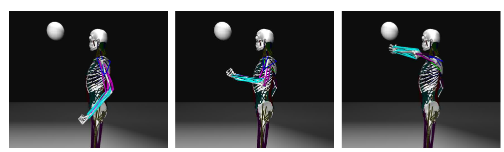
3. Research Agenda

Learning multimodel scene representations. The current state of the art[6] in scene synthesis is based on a combination of **manual user input** and **annotated datasets**, but such input can be **ill-defined**, and data can be **incorrect or insufficient**. For example, if a dataset consists purely of living rooms, we would never be able to synthesize a kitchen, because the dissimilar contents of these rooms serve separate human functions. These shortcomings point to a higher level of abstraction — for example, a studio apartment has elements of both. Such observation begs the question, **what level of granularity or modality of data should we consider**?

The majority of scene synthesis research is focused on purely visual elements in terms of 3D geometries, yet there are **other modalities associated with scenes**, such as light, sound, smell, airflow, temperature, textual description, and so on. For example, a train station, and a restaurant kitchen have particular aspects that cannot be captured only with 3D geometry – there is a variety of activities abound, sounds, heat and so on. A scene synthesis model should **capture** such qualities. My goal is twofold: first, to find such multimodal scene representations, so an end user should only have to provide a high-level description for creating scenes. Second, explore how can we manipulate scenes with an abstract instruction, such as — "clean and organize the apartment". The main difficulties would be with data collection, scene segmentation, and in defining a coupling relationship between each of the modalities, and objects embedded in a scene.

Future Workforce Training. As they continue to improve, **VR/AR** technologies will be **useful** for **diverse training** purposes. For example, **technical training** for the **electronic and mechanical industries**, **emergency or other medical services**, **interior design**, and **architecture**, among others. **Virtual environments**, as well as **simulation of workplace activities** within these environments, can be generated for each of the training scenarios listed above. Once the training requirements, expectations, and action space are defined, I intend to propose computational methods for creating believable, dynamic training environments. Collaboration with experts from each domain would allow enough granularity in terms of the requirements, tasks, and possible outcomes for each training scenario, which would allow to **procedurally generate and simulate** the specific training environments.

Human-centered simulation and perception. Currently, augmented, virtual, and mixed reality systems are concerned with purely visual elements. However, the **physical world** contains other **modalities**, such as sound, smell, touch, temperature, and so on. Incorporating these modalities into a simulation model can create an immersive, **realistic experience**. For example, if an area is designated to be cold, then simulated virtual agents will either try to huddle together, or move towards a warmer area. Such perception capabilities can advance our understanding in how to devise virtual human-like intelligent agents. Additionally, if we generalize, the same simulation model can be used to **train autonomous systems** that interact with the real world. One of the main difficulties in accomplishing this goal will in defining the granularity of such simulation models, and corresponding virtual agents.



Recently, I, along with collaborators, took a step along this research direction by using machine learning to train a **biomechanically-driven virtual human** to move and partake in interactions with the environment[4]. The agent moves are based on realistic neural networks that activate simulated muscles. Simulating such a virtual human-like agent is a challenging problem. This theme, along analysis of informed algorithms for simulation and human-centered design[8,9,10], will constitute the core of my first **NSF proposal** submission.

4. Funding Opportunities

Industry. My research theme is directly applicable to the *gaming and entertainment* industries. For example, my results in procedural content generation (Sec. 2) accelerate the development of gaming content. My real-time simulation results (Sec. 1) can be directly incorporated into current game engine frameworks. I will also build on my industry experience and contacts (Amazon, Autodesk Research, and Wayfair) for future industry collaboration opportunities. For example, learning and ranking user preferences in the context of products (such as interiors), and interactive design tools for stylists and architects.

MURI Air Force Office of Scientific Research. In my multi-agent crowd simulation work (Sec. 1), I formulated the motion of autonomous agents as a constrained optimization problem. Numerical constraints control agent interactions. These constraints are inspired by real-world data of human-motion, and the method I developed to solve the constraint system allows scalable, stable, and efficient simulation of system dynamics, with agent collision-avoidance guarantees. My multi-agent dynamics results address the call for *"physically viable learning for control of autonomous dynamical systems"*.

NSF Cyber-Physical Systems (CPS). My results in user-centric computational design synthesis (Sec. 2), relate to CPS research, by bridging the gap between real and virtual environments. Strengthening my research with an immersive, augmented and virtual reality interface, will address the proposal for the *"human-in-the-loop"* interest area, and accelerate CPS convergence into the real world.

References

[1] Helbing, D., Buzna, L., Johansson, A., and Werner, T. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation science 39*, 1 (2005), 1–24.

[2] Helbing, D., Farkas, I., and Vicsek, T. Simulating dynamical features of escape panic. *Nature 407*, 6803 (2000), 487.

[3] Karamouzas, I., Skinner, B., and Guy, S. Universal power law governing pedestrian interactions. *Phys Rev Lett 113* (Dec 2014), 238701.

[4] M.Nakada, H.Chen, T.Zhao, T.Weiss, and D.Terzopoulos. Deep learning of neuromuscular control for pointing tasks in a biomechanical human model. In *ACM Transactions on Graphics (SIGGRAPH)* (2018).

[5] Pelechano, N., Allbeck, J., Kapadia, M., and Badler, N. *Simulating Heterogeneous Crowds with Interactive Behaviors*. CRC Press, 2016.

[6] Smelik, R. M., Tutenel, T., Bidarra, R., and Benes, B. A survey on procedural modelling for virtual worlds. In *Computer Graphics Forum* (2014), vol. 33, Wiley Online Library, pp. 31–50.

[7] Thalmann, D., Grillon, H., Maim, J., and Yersin, B. Challenges in crowd simulation. In *CyberWorlds. CW'09. International Conference on* (2009), IEEE, pp. 1–12.

[8] Weiss, T., et al. Decorating spaces in augmented reality. *In preparation, to be submitted to ACM Symposium on User Interface Software and Technology (UIST)* (2019).

[9] Weiss, T., Klar, G., and Terzopoulos, D. Optimizing design of physical objects for fabrication. *UCLA Computer Science Technical Report*, 170005 (2017).

[10] Weiss, T., Litteneker, A., Duncan, N., Jiang, C., Yu, L.-F., and Terzopoulos, D. Fast and scalable position-based layout synthesis. *IEEE Transactions on Visualization and Computer Graphics (TVCG)* (2018).

[11] Weiss, T., Litteneker, A., Duncan, N., Jiang, C., Yu, L.-F., and Terzopoulos, D. Fast, scalable layout synthesis. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2018), IEEE. Extended abstract.

[12] Weiss, T., Litteneker, A., Jiang, C., and Terzopoulos, D. Position-based multi-agent dynamics for real-time crowd simulation. In *ACM SIGGRAPH Conference on Motion in Games* (2017), MIG '17, pp. 10:1–10:8. **Best Paper Award**.

[13] Weiss, T., Litteneker, A., Jiang, C., and Terzopoulos, D. Position-based real-time simulation of large crowds. *Computers & Graphics* (2018).

[14] Weiss, T., Nakada, M., and Terzopoulos, D. Automated layout synthesis and visualization from images of interior or exterior spaces. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2017), IEEE, pp. 41–47.