Optimizing Large Scale Crowds in *Ralph Breaks the Internet*

Josh Richards
Walt Disney Animation Studios
josh.richards@disneyanimation.com

Le Joyce Tong
Walt Disney Animation Studios
joyce.tong@disneyanimation.com

Moe El-Ali
Walt Disney Animation Studios
mohammad.el-ali@disneyanimation.com

Tuan Nguyen
Walt Disney Animation Studios
tuan.nguyen@disneyanimation.com

Figure 1: Crowds Complexity: Left) Concept art. Right) Example shot with multiple layers of crowds.

ABSTRACT

In Walt Disney Animation Studio's 57th animated feature *Ralph Breaks the Internet*, the vastness of the internet is imagined as a bustling city where websites are buildings, *Netizen* characters represent algorithms, and *Net Users* travel from site to site. The enormous scope of bringing the world of the internet to life required the Crowds department to rethink how we go about populating our scenes. We extended the Zootopia Crowd Pipeline [El-Ali et al. 2016] to support pose reuse based on level-of-detail, and developed a procedural workflow to populate the world with millions of agents and efficiently render only those visible to camera.

CCS CONCEPTS

• **Computing methodologies** → **Procedural animation**.

KEYWORDS

Crowd, Camera Based LOD, Occlusion Culling, Optimization, Procedural, Crowd Instancing

ACM Reference Format:

Josh Richards, Le Joyce Tong, Moe El-Ali, and Tuan Nguyen . 2019. Optimizing Large Scale Crowds in *Ralph Breaks the Internet. In Proceedings of SIGGRAPH '19 Talks.* ACM, New York, NY, USA, 2 pages. *https://doi.org/10.1145/3306307.3328185*

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGGRAPH '19 Talks, July 28 - August 01, 2019, Los Angeles, CA, USA

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6317-4/19/07.
https://doi.org/10.1145/3306307.3328185

1 INTRODUCTION

The large scale of the internet required the Crowds department to animate characters and vehicles at a wide range of visible depths, from front-and-center to as far back as set extension. However, our current pipeline was not built to support the large number of set elements requiring crowd characters. Our new approach was to break up the scene per set element, procedurally overfill it using character formulas for distribution, movement and interactions, followed by a cull and optimization pass.

2 BUILDING BLOCKS

2.1 Per Building Simulation

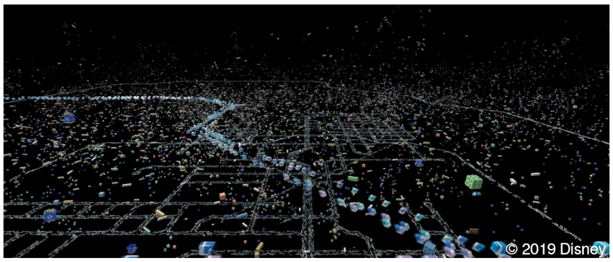
The enormity of the world depicted in the internet required an efficient way to populate all buildings without making it seem like a tiled approach. We created simulations for each unique building with a procedural way to randomly pick the simulation start time. The look of each Net User visiting the websites represented by the building was also varied randomly. Since there were scenes with over a thousand buildings, we developed a system to automatically place the pre-constructed setups in the correct location based only on the root transform without loading any geometry.

2.2 Recursive Spawning

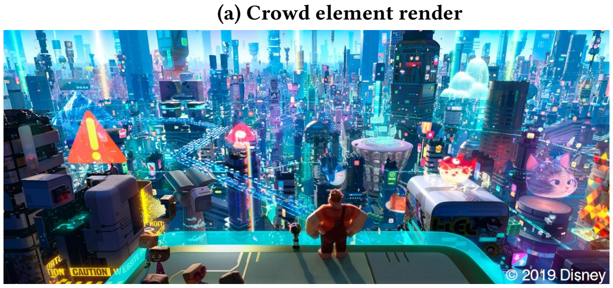
In addition to the per building simulations, location-based simulations were done for set dressing pieces like cafes, seating arrangements, advertisements, etc. These smaller simulations required integration with the larger per-building simulations. Similar tools were used to generate smaller crowd setups recursively from these simulations. While most of these setups are spawned from root transforms, some of the setups used an additional step of computing specific character joint transforms. This was mostly used for objects like props, although spawning a set of new characters at

this point was not unheard of. An example of this, which was used extensively in the film, are the speech bubbles which spawn above the Net Users' heads. The ability to limit certain parameters or prevent further generation based on level of detail was also used to overcome infinitely recursive spawning.

2.3 Net User Traffic System



(a) Crowd element render



(b) Final render

Figure 2: Traffic system example

Two different traffic systems were created to achieve the look desired for the film. The first system was for the Net Users to travel around the internet in pre-defined lanes and was based on the Intelligent Driver Model [Treiber et al. 2000]. This car-following simulation represents vehicles as points and moves them along curves in u-space. We set speed multipliers based on the curvature and the vehicles sampled these values as they traveled. The desired speed of each vehicle was updated each time step, allowing for realistic behavior like slowing down before turns and speeding up in straight sections. Vehicle LOD, passengers and secondary dynamics were automatically added post-simulation based on closest distance to camera.

The second traffic system represented the vehicles as automated processes of the internet like e-mail delivery, garbage removal and advertisements. These vehicles were allowed to flow freely around the buildings and were intended by the Art Director to resemble fish swimming around a reef. To accomplish this effect, we first created a convex hull for each unique building that was shared across building instances. This resulted in a light-weight polygonal representation of the entire scene. Vehicles were spawned based on the height of the individual buildings; the taller the building, the more vehicles it would generate. Most vehicles randomly chose goal points and attempted to reach them by maintaining a minimum distance from the buildings; garbage trucks kept their starting height and traced the contour of the nearest building. Only the main piece of each vehicle was simulated–this allowed us to generate variation and set additional vehicle components like cube trails, advertisements and trailers post-simulation.

3 OPTIMIZATIONS

3.1 Camera Based Level-of-Detail (LOD)

The first optimization step was to base the character's LOD on a screen size metric. This was calculated using the maximum 2D projected bounding box of the joints that make up the character for each frame of the shot. Using the joints' bounding box allowed us to quickly calculate the required LOD prior to skinning. As a result, only characters that are close to camera have high-resolution geometry, hair, and facial expressions;characters in the distance are represented with minimal poly counts, baked proxy hair and textures.

3.2 Pose Reuse

Once we determined the level of detail of each in-camera character, we performed a pass to share similar character poses. We had a total of 6752 unique crowd geometry variants to choose from on the show, so it was imperative to limit the number of unique poses. The closest characters were allowed to be fully-unique while those farther away were simplified in regards to geometry and animation. We removed modifiers, such as additional procedural animation, and aligned the animation clips by adjusting their playback speed and frame offsets so their models could be re-used (instanced) by the renderer.

3.3 Occlusion Cull

The final optimization pass removed any crowd characters obstructed by objects in the scene including other crowd characters. This was done pre-render in GL by hiding semi-transparent objects, counting the number of frames a character was drawn, and removing those with a count of zero.

4 CONCLUSIONS

To put it in perspective, Ralph Breaks the Internet broke many records for Walt Disney Animation Studios. The shots with crowds nearly tripled that of *Big Hero 6* and *Moana*, and doubled that of *Zootopia*.The unique character count was 6752. This is a large increase over the previous productions which peaked at 740. With our new approach, we were be able to simulate and render large-scale complex shots, realizing the concept of a vast internet world.

REFERENCES

Moe El-Ali, Le Tong, Josh Richards, Tuan Nguyen, Alberto Luceno Ros, and Norman Moses Joseph. 2016. Zootopia Crowd Pipeline. In *ACM SIGGRAPH 2016 Talks (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 59, 2 pages. https://doi.org/10.1145/2897839.2927467

Martin Treiber, Ansgar Hennecke, and Dirk Helbing. 2000. Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E* 62 (Aug 2000), 1805–1824. Issue 2. https://doi.org/10.1103/PhysRevE.62.1805