Mobilizing Mocap, Motion Blending, and Mayhem: Rig Interoperability for Crowd Simulation on Incredibles 2

Paul Kanyuk
Pixar Animation Studios
pkanyuk@pixar.com

Patrick Coleman
Pixar Animation Studios
pcoleman@pixar.com

Jonah Laird
Pixar Animation Studios
bluestr@pixar.com

Figure 1: Crowd Scene from Incredibles 2 ©2018 Disney/Pixar. All rights reserved.

CCS CONCEPTS

• **Computing methodologies → Motion capture**; *Procedural animation*; Physical simulation; Agent / discrete models;

KEYWORDS

crowds, motion capture, rig simplification, pipeline

ACM Reference Format:
Paul Kanyuk, Patrick Coleman, and Jonah Laird. 2018. Mobilizing Mocap, Motion Blending, and Mayhem: Rig Interoperability for Crowd Simulation on Incredibles 2. In *Proceedings of SIGGRAPH '18 Talks.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3214745.3214803

The stylized world of *Incredibles 2* features large urban crowds both in everyday situations and in scenes of panicked mayhem. While Pixar's now academy award winning animation software, Presto, has allowed us to create expressive and nuanced rigs for our crowd characters, our proprietary approach has made it difficult to utilize animation from external sources, such as crowd simulations or from motion capture. In this talk, we discuss how we can automatically approximate our complex rigs with skinned skeletons, as well as how this has opened up our crowd pipeline to procedural look-ats, motion blending, ragdoll physics, and motion capture. In particular, the use of motion capture is novel for Pixar, and finding a way to integrate this workflow into our animator-centric pipeline and culture has been an ongoing effort. The system we designed allows us to capture motion data for multiple characters in the context of complex shots in Presto, and it facilitates choreography of nuanced and specifically timed crowd motions. Together with traditional hand animated motion cycles, our crowd choreography tools in Presto [Arumugam et al., 2013], and skeletal agent based simulation in SideFX's Houdini [SideFX, [n. d.]] via our MURE tools [Gustafson et al., 2016], the crowds team on *Incredibles 2* produced rich scenes of busy streets and urban panic.

SIGGRAPH '18 Talks, August 12-16, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5820-0/18/08.
https://doi.org/10.1145/3214745.3214803

1 RIG INTEROPERABILITY

First among the technical challenges to solve was "rig interoperability": that is, how to simplify our proprietary Presto rigs, which contain complex, semi-cyclic networks of deformers and weight objects, into a form that is usable in crowd simulation and motion capture: skeletons and linear blend skinning weights. While a skeleton can often be extracted from our characters by tagging specific rig transforms as bones, computing the skinning weights for these skeletons that approximates our full rig is non-trivial. We employ a system called *FrameBaker*, which uses training data to solve for these weights. As input, we run our character rig through a "calisthenic", in which every animation control is independently set to its minimum and maximum value; in practice, this is close to a thousand poses. Given the resulting mesh deformations and skeletal joint positions, FrameBaker solves a non-negative least square optimization problem to determine skinning weights that best match the input. Naturally, some deformations cannot be adequately represented by linear blend skinning weights, leaving residual error. Simply minimizing this error can lead to undesirable stretching and tearing, an artifact we notice happens on kinematically rigged garments, which we typically deform using cage based subdivision warps and relaxation deformers. Adding a smoothing term to the optimization, which penalizes sharp discontinuities in skinning weights, helps to reduce these tearing artifacts, even if the resulting residual error is numerically higher.

The approach as described works well for character bodies, and it was used for crowds on *Coco*, but determining skinning weights for facial articulation controls remained an open problem. One solution is to reduce the remaining error by running it through principal component analysis and selecting the most significant components for use as pose space deformation blend shapes. While effective, the size and complexity of the blend shapes reduced the potential for rig interoperability. Instead, we took inspiration from the motion capture community, which often utilizes tracking dots on actors' faces, and placed "fake bones" on our characters' faces to track key points. Then, by adding the facial animation controls to our training set, FrameBaker could automatically derive skinning weights for these bones, which works surprisingly well, despite the ad hoc nature of the fake bone placement.

2 MOTION CAPTURE

As a studio with deep roots in character animation, Pixar has traditionally viewed motion capture technology with a healthy dose of professional skepticism. Adding to that the previously mentioned challenges of rig interoperability, the technique historically did not receive much attention as an option. As explorations with camera capture and research into previzuation for the layout department progressed with success, interest in its use for animated characters increased. Our experimental project, *Smash and Grab* [Serritella et al., 2017], crossed the rubicon, and employed a small vicon [Vicon, [n. d.]] capture system as an animation blocking tool. The positive reception from animators to this process paved the way for the crowds department to use motion capture as a portion of our feature animation workflows on both *Coco* and *Incredibles 2*. The understanding was that animators would direct the motion capture shoots, thus empowering them instead of leaving them feeling replaced. In addition, we focus on body motion and do not record face and hand motion. While many studios have made excellent pipelines for facial capture, faces are the most nuanced and directed part of character performances and seemed better left to our animators to create as original content. This compromise also streamlines the amount of setup and calibration required for a motion capture shoot, which we found to be critical. Finally, focusing on crowds rather than the nuanced acting of the main characters, lead the animators to be far more comfortable with the approach.

Since we wish to interactively record and pose motion in a complex shot context with multiple characters, we have developed fast and scalable character rigs and an algorithm for real time motion retargeting from human actors onto our stylized animated characters. As the shot context we display for actors can include other characters, the set and props, and potentially preview effects, both the character rig that we use during recording sessions and the retargeting algorithm need to run with minimal computational cost. We use the skinning weights that we compute offline using FrameBaker to pose a reduced geometric complexity version of the character. The motion retargeting algorithm proceeds in two steps. First, we run a one time "sync" to the actor in which we map skeletal topology, rest pose transforms, and skinning weights between the actor state and the character rest pose, as these can differ significantly between the proportions of human actors and stylized characters. The second portion of the algorithm, which runs for every capture sample, fits the "size" of the actor's motion to the character and enforces ground contact. To do this, we scale the overall size of the motion based on differences in lower body proportions, adjust motion sample limb lengths to fit the character limb sizes, and use a closed form IK solver to account for any remaining ground contact error. With the combination of these approaches, our current bottleneck is hardware rendering overhead, rather than rig execution or motion retargeting.

3 CROWD SIMULATION

The goal of this new process and technology has been to produce an increased number of spectacular crowd simulations. *Incredibles 2* features two large sequences of urban crowd panic; one includes a giant tunnelling machine tearing through a main boulevard, and the other includes an audience witnessing a runaway train. The tunneller sequence was directed to illustrate unintended and costly damage caused by super hero work, but being a family film, had to be free of severe injuries. As a result, the crowds needed to be continuously escaping from harm at the last second to heighten the sense of danger. Traditional hand animated cycles could fill in some of the background, but last second escapes required specific reactions to the events in each scene. Given our new pipeline, when the shots arrived for production, we were able to capture performances in the context of a number of the representative shots, and we used those motion clips as the basis for the larger crowd motion. The skeletal pipeline facilitated lookats and clip blending, and opened the door for the use of rag-doll physics simulations. Inspired by Disney's use of Houdini's [SideFX, [n. d.]] wire solver for crowds on boats in *Moana*, the passengers in both a derailed monorail and runaway train had their skeletons simulated for dramatic physics. Our new ability to manipulate crowd skeletons and the ease of shooting context specific motion capture data now allows us to create some of Pixar's most dynamic crowd simulations to date.

4 FUTURE WORK AND CONCLUSIONS

While this talk focuses on crowd simulation, the potential to use FrameBaker for rig interoperability has opened up many new horizons across the pipeline. For a *Coco* related VR project, the crowds team aided our creative services department by exporting a Presto character into a maya compatible rig, a first for Pixar. The reliable presence of a skeletal representation also enables the use of motion controllers, a rising trend in the research community. Similarly, machine learning techniques are becoming viable now that we have a relatively low dimensional representation of our character poses.

While we now have a practical way to create interoperable rig representations, for future projects we hope to automate the placement of facial bones for FrameBaker. Also, the process of "promoting" one of our modified skeletons back to animation controls on the full quality rig is still a technique we are actively perfecting. That said, Pixar's crowds team has fully embraced the use of skeletal posing and motion capture as tools in our pipeline, and we look forward to their use in future films.

REFERENCES

Hemagiri Arumugam, Michael Frederickson, and J. D. Northrup. 2013. Crowds at Monsters University. In *ACM SIGGRAPH 2013 Talks (SIGGRAPH '13)*. ACM, New York, NY, USA, Article 51, 1 pages. https://doi.org/10.1145/2504459.2504523

Stephen Gustafson, Hemagiri Arumugam, Paul Kanyuk, and Michael Lorenzen. 2016. MURE: Fast Agent Based Crowd Simulation for VFX and Animation. In *ACM SIGGRAPH 2016 Talks (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 56, 2 pages. https://doi.org/10.1145/2897839.2927396

Vincent Serritella, David Lally, Brian Larsen, Farhez Rayani, Jason Kim, and Matt Silas. 2017. Smash and Grab: Off the Rails Filmmaking at Pixar. In *ACM SIGGRAPH 2017 Talks (SIGGRAPH '17)*. ACM, New York, NY, USA, Article 46, 2 pages. https://doi.org/10.1145/3084363.3085077

SideFX. [n. d.]. Company Website: http://sidefx.com.

Vicon. [n. d.]. Company Website: https://www.vicon.com/.