Avengers: Capturing Thanos's Complex Face

Darren Hendler
Digital Domain

Lucio Moser
Digital Domain

Rishabh Battulwar
Digital Domain

David Corral
Digital Domain

Phil Cramer
Digital Domain

Ron Miller
Digital Domain

Rickey Cloudsdale
Digital Domain

Doug Roble
Digital Domain

Figure 1: Image's of Thanos from Marvel's *Avengers: Infinity War*

ABSTRACT

In Marvel's *Avengers: Infinity War*, Thanos (played by actor Josh Brolin,) is entirely CG and is one of the the main characters in this live action movie. The plot depends on the emotional performances of this digital creature and it was imperative that Thanos's facial performance convey the actor's performances faithfully. Digital Domain's performance capture process, *Direct Drive* is a major departure from traditional blendshape solver techniques and was used to create Thanos's performances. We will present an overview of our updated multistage facial retargeting process. We have removed the reliance on high-resolution, per-shot facial capture and refined the process of training the system. This system is faster to set up, needs far less artist input, and preserves elements of the performance that were previously lost using traditional facial capture techniques.

CCS CONCEPTS

• **Computing methodologies → Motion capture**;

KEYWORDS

Facial motion capture, head-mounted cameras

ACM Reference Format:

Darren Hendler, Lucio Moser, Rishabh Battulwar, David Corral, Phil Cramer, Ron Miller, Rickey Cloudsdale, and Doug Roble. 2018. Avengers: Capturing Thanos's Complex Face . In *Proceedings of SIGGRAPH '18 Talks.* ACM, New York, NY, USA, 2 pages. https://doi.org/http://dx.doi.org/10.1145/3084363.3085086

SIGGRAPH '18 Talks, July 30 - August 03, 2017, Vancouver, BC, USA

ACM ISBN 978-1-4503-5008-2/17/07.
https://doi.org/http://dx.doi.org/10.1145/3084363.3085086

1 OVERVIEW

Facial capture is the process in which a digital character's face is driven by an actor's captured facial performance. Typically, markers on the actor's face are tracked using footage from a helmet-mounted camera (HMC) rig and then solved to an actor blend-shape face rig [Lewis et al. 2014]. Unfortunately trying to represent an actor's complex facial motion as a series of weighted FACS shapes results in significant loss of the original performance. It also needs a large number of high-quality face shapes, requiring a long development cycle (months) between initial capture and first realistic performances.

To speed up turnaround and improve the quality of the character's facial transfer, Digital Domain developed the Direct Drive system [Hendler et al. 2017] which directly transformed an actor's seated facial capture data into a high-resolution character performance. This process greatly improved quality and sped up turnaround but was limited as it required the actor to deliver all performances in seated capture environment and was expensive to process.

For *Avengers: Infinity War* we propose a novel combination of approaches and new refinements to facilitate the transfer of an actor's mobile helmet camera performance quickly and accurately to a character's face *with limited per shot manipulation.*

2 OUR APPROACH

For Avengers it was important, that all actors portraying digital characters deliver their lines, live on set, interacting with one-another using HMC capture. To achieve this we used Masquerade (see [Moser et al. 2017]) which produces a high-resolution detailed actor mesh from a sparse marker set, in-conjunction with Direct Drive. This unique combination produced final character facial performances from on-set actor HMC footage removing the reliance on per shot seated capture.

In addition, we made improvements to the Masquerade and Direct Drive processes to improve both the quality and speed over previous actor-to-character face transfers.

2.1 Masquerade - Actor Mesh Creation

In the early stages of *Avengers: Infinity War* we captured several performances in Disney's seated facial capture system, Medusa [1]. This system captures thousands of discrete virtual points to create a consistent high-resolution mesh which produced 2500 frames of 3D facial training data for our actor.

In a new step, we enhanced this dense captured data using 20-30 refinement face shapes, created using a combination of modeling and higher resolution scans. This 2nd step is achieved using an example-based correction workflow [Lewis et al. 2016] [Bickel et al. 2008] which recovers missing areas, corrects scan issues and introduces missing fine scale details.

During the principal shoot, the actor is captured live, on-set using an HMC rig. The actor has 200 dots applied to their face in a set pattern, these dots are then tracked through multiple camera views and are triangulated in 3D space. This produces a sparse mask of points representing consistent 3D positions on the actor's face. This sparse tracked marker set is used in conjunction with the high-resolution refined training data in Masquerade (see [Moser et al. 2017]). The result is a per-frame, high-resolution actor mesh with the same fine-scale details as the training data but synced to the actor's on-set HMC performance. This technique can produce a high-resolution, accurate actor facial performance from an HMC capture in 1-2 weeks instead of months and does not suffer the same performance loss as the traditional-solver/blendshape techniques.

For secondary characters we developed a new process, where we used Direct Drive to genericize and re-target one actor's training data to another actor's face. This transfered training data was used in conjunction with the new actor's tracked HMC markers in Masquerade to produce a high-resolution actor facial performance. The fine scale facial details in this performance, did not match the new actor exactly but produced a believable, realistic performance when targeted to their driven character, without the need for seated capture.

2.2 Direct Drive Facial Transfer

In the next stage, we transfer the actor's Masquerade generated 3D face data directly to the character's modeled face. First, the actor's face mesh is non-rigidly deformed to match the creature's face in a neutral pose [Tam et al. 2013]. To make sure that the key facial features on the deformed actor's mesh align with the same semantic features on the creature's face, we help the deformation process by adding some manually placed correspondences.

The Direct Drive [Hendler et al. 2017] system uses this mapping of correspondences and the animated facial performance to replicate the actor's expressions, frame by frame, on the creature's face. A transfer rig is built with a deformation stack that uses gradient-based deformation transfer [Sumner and Popović 2004] and general purpose mesh deformers to constrain points on the mesh, resolve skin to bone collisions, and to control the rigidity on the face [Moser et al. 2016].

[1] http://medusa.disneyresearch.com/

To account for any expression and fine-scale differences between actor and character we use an example-based correction workflow [Lewis et al. 2016] [Bickel et al. 2008]. This automatically becomes part of the Direct Drive process allows riggers and artists to make art-directed changes easily and efficiently to control the deformation transfer from actor to character.

The final character performance is decomposed onto a multi-layer facial rig that allows animation to adjust or refine the character's face as needed while still preserving every detail from the original transfer.

3 CONCLUSIONS

Previously, when using helmet cameras to capture an actor's performance we were limited to using a FACS shape rig and facial solves. This new process allows us to capture a far greater range of the actor's performance and allows us to leverage our previous Direct Drive process which was previously limited to expensive per-shot seated capture. Using this system we have the freedom and cost-effectiveness of HMC capture without having to sacrifice mobility to get a high quality realistic facial transfer.

This combination of techniques and new refinements have resulted in a more realistic facial performance in far less time than either previous system (solving / Direct Drive) and has sped up our turnaround on creating fully digital talking creatures. Our animation team has embraced this process as they can concentrate on the creative modifications rather than the missing foundation. This system has reduced our facial animation time by half and reduced our initial development cycle from months to weeks. This system is currently being used on *Avengers: Infinity War* to deliver over 300 shots of hero facial performances.

REFERENCES

Bernd Bickel, Manuel Lang, Mario Botsch, Miguel A. Otaduy, and Markus Gross. 2008. Pose-space Animation and Transfer of Facial Details. In *Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '08)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 57–66. http://dl.acm.org/citation.cfm?id=1632592.1632602

Darren Hendler, Steve Preeg, David Corral, Carolyn Wong, Ron Miller, Lucio Moser, and Rishabh Battulwar. 2017. Capturing a Face. (2017). DigiPro.

J. Lewis, Victoria University of Wellington. School of Engineering, and Computer Science. 2016. *Learning from the Artist: Theory and Practice of Example-Based Character Deformation : a Thesis Submitted to the Victoria University of Wellington in Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Computer Science.* Victoria University of Wellington. https://books.google.com/books?id=HCKsnQAACAAJ

J. P. Lewis, Ken Anjyo, Taehyun Rhee, Mengjie Zhang, Fred Pighin, and Zhigang Deng. 2014. Practice and Theory of Blendshape Facial Models. In *Eurographics 2014 - State of the Art Reports*, Sylvain Lefebvre and Michela Spagnuolo (Eds.). The Eurographics Association. https://doi.org/10.2312/egst.20141042

Lucio Moser, David Corral, and Doug Roble. 2016. As-rigid-as-possible Deformation Transfer for Facial Animation. In *ACM SIGGRAPH 2016 Talks (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 29, 2 pages. https://doi.org/10.1145/2897839.2927411

Lucio Moser, Darren Hendler, and Doug Roble. 2017. Masquerade: Fine-scale Details for Head-mounted Camera Motion Capture Data. In *ACM SIGGRAPH 2017 Talks (SIGGRAPH '17)*. ACM, New York, NY, USA, Article 18, 2 pages. https://doi.org/10.1145/3084363.3085086

Robert W. Sumner and Jovan Popović. 2004. Deformation Transfer for Triangle Meshes. In *ACM SIGGRAPH 2004 Papers (SIGGRAPH '04)*. ACM, New York, NY, USA, 399–405. https://doi.org/10.1145/1186562.1015736

Gary K. Tam, Zhi-Quan Cheng, Yu-Kun Lai, Frank Langbein, Yonghuai Liu, A. David Marshall, Ralph Martin, Xianfang Sun, and Paul Rosin. 2013. Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Nonrigid. *IEEE Transactions on Visualization and Computer Graphics* 19, 7 (July 2013), 1199–1217. https://doi.org/10.1109/TVCG.2012.310