ILM Facial Performance Capture

Brian Cantwell*
Industrial Light & Magic

Paige Warner†
Industrial Light & Magic

Michael Koperwas ‡
Industrial Light & Magic

Kiran Bhat §
LoomAi

Figure 1: *Anna Galvin as Draka in* Warcraft *and Noel Fisher as Mikey in* Teenage Mutant Ninja Turtles: Out of the Shadows.

Abstract

Industrial Light & Magic utilized facial capture technology on a massive scale to bring more than two dozen computer-generated characters to life in the new Warcraft, Teenage Mutant Ninja Turtles, and Star Wars films. Here we present our flexible proprietary facial capture system which was used to faithfully translate all the subtle nuances, lip movements, and saccades of over 1200 facial performances onto editable, artist-friendly animation rigs, and ultimately the big screen.

Keywords: Facial Capture, Animation, Human, Creature

Concepts: •Computing methodologies → Computer graphics; Animation; Rendering; Texturing;

1 Facial Rigging

Everything starts with the facial rigs. The actor's facial rig provides a series of controls to animate a digital version of a specific actor's face. This rig will be solved to match the actor's captured facial performance. When the solve is complete, the actor rig animation will be copied to the character rig.

We start with a facial scanning session with the actor, leveraging Disney Research Zurich's MEDUSA system [Beeler et al. 2011] to capture the actor's face through a wide range of facial expressions. Scans processed from each expression are used to create a blendshape-based facial animation rig. These blendshape controls are passed into our facial solving setup later in the process. Using deformation transfer, the actor blendshape rig is transferred to initialize the character facial rig. Subsequently, these shapes are further sculpted and refined to create the final character blendshapes that are used for both facial capture and keyframe animation. Correspondence between actor rig and character rig is very important as this directly affects the results of our facial retargets. A smile on one rig should look like a smile on the other, jaw motion should be proportional, and lip motion must match.

2 Capturing

The actor's face is marked with a set of carefully placed dots. The actor's performance is then captured with a head mounted camera (HMC) rig that uses either 2 or 4 cameras depending upon the project. The use of multiple cameras provides us with the necessary information to generate an accurate 3d reconstruction of the actor's performance. After the shoot is complete, the selected takes go through a 2d tracking process of all of the dots, as well as additional features such as the irises and teeth. Additionally, roto spline contours are created to follow the inner lip lines and eye lids. The 2d tracks and splines are passed as input into our facial solving system [Bhat et al. 2013].

3 Solving

The core of our solving system is a flexible, multi-scale spatio-temporal framework that uses the tracked dots and roto spline contours to solve for 3d reconstruction, rigid registration to the actor model, and the controls on our actor rig. We also leverage our proprietary SnapSolver [Bhat et al. 2013], a per-vertex/per-frame mesh deformer capable of capturing temporal high frequency out-of-model deformations. Using the same 2d tracks and roto spline contours, the SnapSolver's per-frame deformation captures tiny twitches, sticky lips, and the secondary dynamics on the face caused by body movement such as the bounce and jiggle from walking and running, or the sagging due to gravity when lying down.

4 Editing

Depending upon the situation, an actor's performance may need to be enhanced with secondary animation, retimed or blended between multiple takes. It also may need to transition into and out of purely

*e-mail: cantwell@ilm.com
†e-mail: paigewarner@ilm.com
‡e-mail: mkoper@ilm.com
§e-mail: kiran.s.bhat@gmail.com

SIGGRAPH '16, July 24-28, 2016, Anaheim, CA,
ISBN: 978-1-4503-4282-7/16/07
DOI: http://dx.doi.org/10.1145/2897839.2927438

keyframed moments. While much of the retargetted performance may be on animation curves on the rig, the per-frame deformations are edited using an additional layer of filters and deformers on the rig, similar to those used during the retarget.

Acknowledgements

The authors would like to thank Jeff White, Jason Smith, Hal Hickel, Nigel Sumner, Duncan Jones, Bill Westenhofer, Pablo Helman, Robert Weaver, Kevin Martel, Shawn Kelly, Paul Giacoppo, Jung-Seung Hong, Owen Calouro, Kate Lee, Andy Ritchie, Sonia Contreras, Sheila SanSantos, Clint Spillers, Alex Aponte, Matt Rank, Glenn Derry, the talented artists and engineers at ILM, and all of our actors, whose great performances drive everything we do.

References

BEELER, T., HAHN, F., BRADLEY, D., BICKEL, B., BEARDSLEY, P., GOTSMAN, C., SUMNER, R. W., AND GROSS, M. 2011. High-quality passive facial performance capture using anchor frames. *ACM Trans. Graph. 30*, 4 (July).

BHAT, K. S., GOLDENTHAL, R., YE, Y., MALLET, R., AND KOPERWAS, M. 2013. High fidelity facial animation capture and retargeting with contours. In *Proceedings of the 12th ACM SIGGRAPH/eurographics symposium on computer animation*, 7–14.