CASA 2019: Capturing Real-world Data for Characters, Sometimes with Machine Learning

Full Day Tutorial

Eric K. Patterson
Clemson University
Clemson, SC, 29634 USA
ekp@clemson.edu

Dan Casas
Universidad Rey Juan Carlos
Madrid, Spain
dan.casas@urjc.es

Daljit Singh J. Dhillon
Clemson University
Clemson, SC, 29634 USA
djsingh@clemson.edu

1 Overview

The theme of this tutorial is to present methods for capturing real-world data across the pipeline from acquisition to animation for generating an avatar or other character automatically, with an emphasis on photo-realistic characters and movement, along with discussing some of the areas where machine learning (ML) is entering both research and workflows already. We will present concepts, a survey of related work and research, work-in-progress, and discuss practical aspects. Some particular emphasis will be given to the expertise of the organizers including capturing faces, skin, and body movement along with a variety of theoretical and practical considerations.

2 Capturing Appearance

2.1 Geometry

Methods such as laser-scan, structured-light stereography, and photogrammetry have been devised for acquiring detailed meshes of the face. Commercial systems include such as marketed by Cyberscan, 3DMD, Arctec, and others. Photogrammetric solver packages such as PhotoScan by Agisoft and CaptureReality by Reality Capture are also in growing use, as well as open-source solutions such as Meshroom based on AliceVision, and studios have built large systems of camera arrays to capture the images needed. Simpler commodity systems have also been used for inexpensive capture. Concepts as well as practical examples related to face and character-maquette capture will be presented, including commercial and open-source workflows with a survey of equipment and methods. Recent work in application of machine-learning to mesh acquisition may be surveyed given time.

2.2 Material

A variety of means have been researched for acquiring real-world, light-scattering data from materials such as clothing, skin, and other surfaces over the last several decades. Most include controlled illumination and viewing. Material maps are often constructed as outputs from these means to control varying parameters during rendering. Methods for extracting these from both research and a practical sense will be discussed as well as ongoing developments. Machine-learning approaches to material capture have have recently been directed at simplifying methods for capturing and appying real-world data [AWL15] [DAD+18].

2.2.1 Skin

One of the most important materials for characters, of course, is skin. Human skin is a complex yet mostly well layered composition of fibrous biomass. It embodies heterogeneous distribution of choromophores such as melanin and hemoglobin, a labrinyth of blood vessels, hair follicles with nerve endings and sweat as well as sebaceous glands. To top it all, human skin secretes a fine layer of oil on its rough micro-geometric exterior. Optically, we can model this complex composition well as a semi-transparent, multilayered material whose absorption, scattering and transmission characteristics vary with each layer. For graphical renderings, we aim to realistically reproduce human skin's rich variety in color tones, spatial variations and nuances such as freckles, moles, blemishes, under-grown hair and general impressions of paleness or cold/neutral/warm undertones, preferably at interactive speeds.

This part of the tutorial will begin with a short introduction to human skin's composition, its optical characterization. Next, the fundamentals for bidirectional scattering-surface reflectance distribution functions (BSSRDFs) and Kubelka-Munk theory [Nob85] for diffusion based sub-scattering models will be discussed. It will include dipole and multipole methods for evaluating such diffusion based, multi-layered BSSRDF models. This will be followed by a detailed explanation on the most basic, physically-based, two-layered, chromophoral model (epidermis with melanin and dermis with hemoglobin) by Donner and Jensen [DJ06] which incorporates four intuitive parameters to reproduce a variety of homogeneous skin tones. It will then cover the variations of this Donner and Jensen model to cover additional parameters for improving the gamut of skin coloration.

Lastly, regarding skin, heterogeneous reflectance models for reproducing realistic appearances with blemishes, freckles and even tattoos will be discussed [DWd+08]. Thereafter, advanced, light-stage based, pure data-driven models used in digital media production will be discussed [ARL+10, FJA+14]. Finally, research in acquisition for user-controllable parametric variations of physically based models for accurate reconstruction of an individual's facial skin in rendering will be presented.

2.3 Hair, Clothes, and Other

A brief survey will be included of recent work in these areas to capture other real-world aspects related to characters given time.

3 Motion and Animation

3.1 Capturing Expressions and Poses

Poses and expressions are particularly important for animation and thus planning well for capture is crucial. Some interdisciplinary discussion will be included regarding what should be captured in various contexts for automating creation of characters including aspects such as facial dynamics.

3.2 Establishing Correspondence

Acquired meshes fall short of being ready for direct use as an avatar, character, or any application needing animation. A similarly shaped mesh with well formed topology needs to be constructed and is typically done through a manual or assisted "re-topology" process. Firstly, the new mesh needs to follow the musculature of the human face, assisting in creating a character rig that can deform believably during animation. Secondly, conforming meshes facilitates other goals such as uniform layout of material maps used during rendering, blending expressions, or even matching mesh format across different subjects. Thirdly, an overly dense mesh may be reduced to a lower polygon count so that it can perform in a reasonable manner. Commercial software exists for manual re-topology as well as for assisted correspondence once a template mesh is established. Some systems such as Vuvuzela, developed at ICT, began as fairly manual correspondence systems created to align meshes of varying expressions and have progressed to more automated means, using 2D landmarking techniques, for aligning meshes and per-expression texture maps for a single individual[IAD13][CFA+16]. Work for 3D morphable models have attempted various correspondence systems for faces across subjects, often also using 2D texture space, but these usually vary little in expression nor have been targeted specifically for generating a mesh that can be animated as is needed for many applications. Work continues to examine mechanisms for automatically aligning meshes both from a single individual and across subjects[PB18][PBS18]. We will present work in progress for using recent deep-learning based landmarking methods to align and find correspondence among scanned meshes and texture maps for creating animatable meshes in a common, well-designed topology.

3.3 Rigging

A survey of attempts for preparing rigging for animation automatically will be included given time.

3.4 Bodies and Clothing

ML methods have been recently introduced by the authors of this proposal to animate bodies, hands, and clothing. In particular, Casas and Otaduy [CO18] proposed a method to learn soft-tissue dynamics as a function of body shape and motion. Similarly, Santesteban *et al.* [SOC19] presented a pose-dependent data-driven method to animate garments. Importantly, these ML approaches enable to skip the computationally expensive physical simulation step at run time and, once the model is trained, only require a highly efficient forward pass through the regression algorithm to predict the next state of an animation.

4 Organizers

Eric K. Patterson He is Associate Professor in Visual Computing and also Associate Director for Digital Production Arts in the School of Computing at Clemson University in Clemson, SC, USA. He was previously Professor of Computer Science at University of North Carolina Wilmington where he co-founded the I3S Institute and Face Aging Lab; introduced a Digital Arts program; and assisted in the founding of multiple graduate and undergraduate programs as well as the establishment of Film Studies as a Department. Among other efforts, he has been an IPAX Faculty Fellow at Sony Pictures Imageworks in Culver City, CA and worked with the Technical Art team of Red Storm Ubisoft in Cary, NC, where he particularly focused on photogrammetry and material-capture efforts. His current interests include facial analysis along with material capture and sometimes machine learning.

Dan Casas. He is Assistant Professor at the Universidad Rey Juan Carlos (URJC), Spain. Previously, he was Marie Skłodowska-Curie Individual Fellow (2016–2018) the URJC, postdoc (2015–2016) at the Max Planck Institute in Saarbrucken, Germany, Research Fellow (2014–2015) at the University of Southern California's Institute for Creative Technology (CA,USA). Dan received his PhD in Computer Graphics in 2014 from the University of Surrey (UK), supervised by Prof. Adrian Hilton. Dan's dissertation introduced novel methods for character animation from multi-camera capture that allow the synthesis of video-realistic interactive 3D characters. His current interest include Machine Learning methods for capturing and animating human performances, including bodies and clothing.

Daljit Singh Dhillon He is Assistant Professor in Visual Computing in the School of Computing at Clemson University, Clemson, SC, USA. Prior to joining Clemson, he worked as Research Associate (2016-2018) with SNSF post-doc fellowship (2016-2017) at Realistic Graphics and Imaging group, Imperial College London, UK under Dr. Abhijeet Ghosh's supervision. At Imperial, he worked on image based methods for acquisition and modeling of complex material appearances involving diffraction effects and sub-surface scattering. Daljit Singh received his PhD in Computer Science in 2015 from University of Bern, Switzerland. His interdisciplinary doctoral thesis dealt with modeling and simulating reptile skin colouration; in collaboration with Prof. Michel Milinkovitch (biologist) and his team at LANE lab, University of Geneva, Switzerland. Presently, he is interested in surface appearances and phenomena involving wave-optics.

References

[ARL+10] Oleg Alexander, Mike Rogers, William Lambeth, Jen-Yuan Chiang, Wan-Chun Ma, Chuan-Chang Wang, and Paul Debevec. The digital emily project: Achieving a photorealistic digital actor. *IEEE Computer Graphics and Applications*, 30(4):20–31, 2010.

[AWL15] Miika Aittala, Tim Weyrich, and Jaakko Lehtinen. Two-shot svbrdf capture for stationary materials. *ACM Trans. Graph.*, 34(4):110:1–110:13, July 2015.

[CFA+16] Dan Casas, Andrew Feng, Oleg Alexander, Graham Fyffe, Paul Debevec, Ryosuke Ichikari, Hao Li, Kyle Olszewski, Evan Suma, and Ari Shapiro. Rapid photorealistic blendshape modeling from rgb-d sensors. In *Proceedings of the 29th International Conference on Computer Animation and Social Agents*, pages 121–129, 2016.

[CO18] Dan Casas and Miguel A. Otaduy. Learning nonlinear soft-tissue dynamics for interactive avatars. *Proc. ACM Comput. Graph. Interact. Tech.*, 1(1), July 2018.

[DAD+18] Valentin Deschaintre, Miika Aittala, Frédo Durand, George Drettakis, and Adrien Bousseau. Single-image svbrdf capture with a rendering-aware deep network. *ACM Transactions on Graphics (SIGGRAPH Conference Proceedings)*, 37(128):15, aug 2018.

[DJ06] Craig Donner and Henrik Wann Jensen. A spectral bssrdf for shading human skin. *Rendering techniques*, 2006:409–418, 2006.

[DWd+08] Craig Donner, Tim Weyrich, Eugene d'Eon, Ravi Ramamoorthi, and Szymon Rusinkiewicz. A layered, heterogeneous reflectance model for acquiring and rendering human skin. In *ACM Transactions on Graphics (TOG)*, volume 27, page 140. ACM, 2008.

[FJA+14] Graham Fyffe, Andrew Jones, Oleg Alexander, Ryosuke Ichikari, and Paul Debevec. Driving high-resolution facial scans with video performance capture. *ACM Transactions on Graphics (TOG)*, 34(1):8, 2014.

[IAD13] Ryosuke Ichikari, Oleg Alexander, and Paul Debevec. Vuvuzela: a facial scan correspondence tool. In *ACM SIGGRAPH 2013 Posters*, page 89. ACM, 2013.

[Nob85] James H Nobbs. Kubelka—munk theory and the prediction of reflectance. *Review of Progress in Coloration and Related Topics*, 15(1):66–75, 1985.

[PB18] Eric Patterson and Jessica Baron. Automated re-topology of polygonal meshes of the human face. In *ACM I3D 2018 Posters*. ACM, 2018.

[PBS18] Eric Patterson, Jessica Baron, and Devin Simpson. Landmark-based re-topology of stereo-pair acquired face meshes. In Leszek J. Chmielewski, Ryszard Kozera, Arkadiusz Orłowski, Konrad Wojciechowski, Alfred M. Bruckstein, and Nicolai Petkov, editors, *Computer Vision and Graphics*, pages 425–437, Cham, 2018. Springer International Publishing.

[SOC19] Igor Santesteban, Miguel A. Otaduy, and Dan Casas. Learning-Based Animation of Clothing for Virtual Try-On. *Computer Graphics Forum (Proc. Eurographics)*, 38(2), 2019.