Capture4VR: From VR Photography to VR Video

Christian Richardt
University of Bath
christian@richardt.name

Peter Hedman
University College London
p.hedman@cs.ucl.ac.uk

Ryan S. Overbeck
Google LLC
rover@google.com

Brian Cabral
Facebook
bkc@fb.com

Robert Konrad
Stanford University
rkkonrad@stanford.edu

Steve Sullivan
Microsoft
steve.sullivan@microsoft.com

ABSTRACT

Virtual reality (VR) enables the display of dynamic visual content with unparalleled realism and immersion. However, VR is also still a relatively young medium that requires new ways to author content, particularly for visual content that is captured from the real world. This course, therefore, provides a comprehensive overview of the latest progress in bringing photographs and video into VR. Ultimately, the techniques, approaches and systems we discuss aim to faithfully capture the visual appearance and dynamics of the real world, and to bring it into virtual reality to create unparalleled realism and immersion by providing freedom of head motion and motion parallax, which is a vital depth cue for the human visual system. In this half-day course, we take the audience on a journey from VR photography to VR video that began more than a century ago but which has accelerated tremendously in the last five years. We discuss both commercial state-of-the-art systems by Facebook, Google and Microsoft, as well as the latest research techniques and prototypes. Course materials will be made available on our course website https://richardt.name/Capture4VR.

SIGGRAPH 2019 Courses, July 28 – August 1, 2019, Los Angeles, CA, USA

ACM ISBN 978-1-4503-6307-5/19/07.
https://doi.org/10.1145/3305366.3328028

ACM Reference Format:
Christian Richardt, Peter Hedman, Ryan S. Overbeck, Brian Cabral, Robert Konrad, and Steve Sullivan. 2019. Capture4VR: From VR Photography to VR Video. In *Proceedings of SIGGRAPH 2019 Courses.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3305366.3328028

COURSE OVERVIEW

We start with a brief introduction that welcomes the audience and covers the course's schedule and basic terminology. This is followed by a short review of history from the pre-computer era, including stereo photography and panoramic paintings. Each paragraph in this section corresponds to about 20 minutes of presentation.

We next cover traditional panorama stitching [Szeliski 2006], which has recently found its way into smartphones and consumer

360° cameras, omnidirectional stereo (ODS) panoramas [Richardt et al. 2013], which is the basis for many current VR photography techniques (e.g. Google Cardboard Camera), and recent approaches that reproduce motion parallax [Bertel et al. 2019; Luo et al. 2018].

Our focus then shifts towards VR photography approaches that reconstruct a textured 3D model of the scene from hand-held photos [Hedman et al. 2017; Hedman and Kopf 2018]. This makes exploration in VR more immersive, as it enables full 6 degrees-of-freedom (6-DoF) head motion, allowing users to look around freely in VR.

We conclude our coverage of VR photography approaches with a deep dive into the system that has shown the highest visual fidelity to date: Google's panoramic light fields [Overbeck et al. 2018]. This system captures around a thousand photos of a scene, processes them using structure-from-motion and multi-view stereo techniques [Shum et al. 2007], and compresses them to enable real-time rendering of novel views for 6-DoF VR with unprecedented visual quality, including soft specular reflections off shiny surfaces.

After a short break, the course moves on from VR photography to VR video. We start with today's most widespread VR video formats [Richardt et al. 2017]. 360° video can be captured with affordable consumer 360° cameras, but lacks depth. ODS video produces stereoscopic 360° video, but requires multi-camera rigs, such as the open-source Facebook Surround 360 design [Cabral 2016; Facebook 2016] or Google Jump [Anderson et al. 2016].

Existing ODS video approaches require computationally expensive off-line processing, which effectively prevents their use for live streaming. However, new imaging system designs with rotating sensors sidestep the computation and acquire ODS video directly [Konrad et al. 2017]. This enables live streaming of ODS video without computationally expensive preprocessing.

As in VR photography, there is currently a lot of research interest and excitement around enabling 6-DoF head motion within VR video [e.g. Huang et al. 2017], so that viewers can move their head around freely to explore a scene. We look at some of the latest camera rigs and techniques developed by Facebook, such as their x6, x24 and Manifold cameras [Facebook 2018], and the 6-DoF videos they can capture.

Finally, we turn cameras around from pointing outwards towards an environment to pointing inwards at objects of interests like people and animals to capture them volumetrically [Collet et al. 2015]. For this, we take a look at the technology behind Microsoft's state-of-the-art commercial Mixed Reality Capture Studios. The captured 'holograms' can be inserted into virtual environments for VR experiences, or they can be used to augment real scenes using mixed-reality headsets like the Microsoft HoloLens.

We conclude with a short summary, highlight the remaining challenges towards ubiquitous 6-DoF VR photography and video, and provide an outlook of things to come. This will be followed by questions and answers.

ACKNOWLEDGMENTS

This course is supported by RCUK grant CAMERA (EP/M023281/1) and an EPSRC-UKRI Innovation Fellowship (EP/S001050/1).

REFERENCES

Robert Anderson, David Gallup, Jonathan T. Barron, Janne Kontkanen, Noah Snavely, Carlos Hernandez, Sameer Agarwal, and Steven M. Seitz. 2016. Jump: Virtual Reality Video. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 35, 6 (November 2016), 198:1–13. https://doi.org/10.1145/2980179.2980257

Tobias Bertel, Neill D. F. Campbell, and Christian Richardt. 2019. MegaParallax: Casual 360° Panoramas with Motion Parallax. *IEEE Transactions on Visualization and Computer Graphics* 25, 5 (May 2019), 1828–1835. https://doi.org/10.1109/TVCG.2019.2898799

Brian Cabral. 2016. VR Capture: Designing And Building An Open Source 3D-360 Video Camera. SIGGRAPH Asia Keynote.

Alvaro Collet, Ming Chuang, Pat Sweeney, Don Gillett, Dennis Evseev, David Calabrese, Hugues Hoppe, Adam Kirk, and Steve Sullivan. 2015. High-quality Streamable Free-viewpoint Video. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 34, 4 (July 2015), 69:1–13. https://doi.org/10.1145/2766945

Facebook. 2016. Surround 360 Instruction Manual. GitHub. https://github.com/facebook/Surround360/blob/master/surround360_design/assembly_guide/Surround360_Manual.pdf

Facebook. 2018. Filming the Future with RED and Facebook 360. https://facebook360.fb.com/2018/09/26/film-the-future-with-red-and-facebook-360/

Peter Hedman, Suhib Alsisan, Richard Szeliski, and Johannes Kopf. 2017. Casual 3D Photography. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 36, 6 (November 2017), 234:1–15. https://doi.org/10.1145/3130800.3130828

Peter Hedman and Johannes Kopf. 2018. Instant 3D Photography. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 37, 4 (July 2018), 101:1–12. https://doi.org/10.1145/3197517.3201384

Jingwei Huang, Zhili Chen, Duygu Ceylan, and Hailin Jin. 2017. 6-DOF VR videos with a single 360-camera. In *Proceedings of IEEE Virtual Reality (VR)*. 37–44. https://doi.org/10.1109/VR.2017.7892229

Robert Konrad, Donald G. Dansereau, Aniq Masood, and Gordon Wetzstein. 2017. SpinVR: Towards Live-Streaming 3D Virtual Reality Video. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 36, 6 (November 2017), 209:1–12. https://doi.org/10.1145/3130800.3130836

Bicheng Luo, Feng Xu, Christian Richardt, and Jun-Hai Yong. 2018. Parallax360: Stereoscopic 360° Scene Representation for Head-Motion Parallax. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1545–1553. https://doi.org/10.1109/TVCG.2018.2794071

Ryan Styles Overbeck, Daniel Erickson, Daniel Evangelakos, Matt Pharr, and Paul Debevec. 2018. A System for Acquiring, Compressing, and Rendering Panoramic Light Field Stills for Virtual Reality. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 37, 6 (2018), 197:1–15. https://doi.org/10.1145/3272127.3275031

Christian Richardt, Yael Pritch, Henning Zimmer, and Alexander Sorkine-Hornung. 2013. Megastereo: Constructing High-Resolution Stereo Panoramas. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)*. 1256–1263. https://doi.org/10.1109/CVPR.2013.166

Christian Richardt, James Tompkin, Jordan Halsey, Aaron Hertzmann, Jonathan Starck, and Oliver Wang. 2017. Video for Virtual Reality. In *SIGGRAPH Courses*. https://doi.org/10.1145/3084873.3084894

Heung-Yeung Shum, Shing-Chow Chan, and Sing Bing Kang. 2007. *Image-Based Rendering*. Springer. https://doi.org/10.1007/978-0-387-32668-9

Richard Szeliski. 2006. Image alignment and stitching: a tutorial. *Foundations and Trends in Computer Graphics and Vision* 2, 1 (January 2006), 1–104. https://doi.org/10.1561/0600000009