Video for Virtual Reality

SIGGRAPH 2017 Course

Christian Richardt
University of Bath

James Tompkin
Brown University

Jordan Halsey
VR Playhouse

Aaron Hertzmann
Adobe Research

Jonathan Starck
Foundry

Oliver Wang
Adobe Research

Course website with slides:
http://richardt.name/pub/Video4VR/

SIGGRAPH 2017 Courses, July 30 – August 3, 2017, Los Angeles, CA, USA
ACM 978-1-4503-5014-3/17/07
https://doi.org/10.1145/3084873.3084894

Abstract

Video can capture the dynamic appearance of the real world in a way no other technology does; virtual reality technology, on the other hand, enables the display of dynamic visual content with unparalleled realism and immersion. The fusion of these two technologies—*video for virtual reality (VR)*—promises to enable many exciting photo-realistic experiences.

Over half a day, this course will provide an overview of three aspects of this exciting medium: the technical foundations, current systems in practice, and the potential for future systems of VR video.

In the first section, we will explore the geometric and optical problems underpinning VR video. Then, we will introduce both 360 degree video and stereoscopic video, including how 360 video is captured, analyzed, and stitched, including the mathematics behind how stereo 360 video can be captured.

This background material provides the prerequisites for understanding current systems in use. In the middle hour of the course, we explain how state-of-the-art stereo 360 video is produced from camera systems and computational processing. Then, we will consider the art of storytelling in VR, and how new tools for editing VR video can aid in the craft of this art production. Finally, this section provides an industry perspective covering current production and post-production choices and practice, including CG integration.

The final part of our course focuses on the next generation of video for VR, where we move to 6 degrees-of-freedom (6DoF) experiences. We introduce the basics and challenges behind light field cameras, processing and displays, and see how they can enable 6DoF experiences. This will be followed by another industry perspective on how light field camera arrays have been used to create cutting-edge experiences integrating volumetric live-action elements. To conclude the course, we will see how far we still must go toward the ideal system, in hopes of inspiring the attendees to push the boundary farther to reach it.

We hope this course is useful to a broad audience—at SIGGRAPH and beyond—as we cover the academic, artistic, and production sides of VR video.

Course history and relevant existing courses

This is a new course on a topic that has so far not been covered at SIGGRAPH. While significant advances in video and virtual reality have been made in the past decade, less has been specifically written about this common use of VR technology. We hope to address this with our course.

Some aspects related to our course topics have been covered in previous SIGGRAPH (Asia) courses:

– Human-centered design for VR interactions[1] by Jason Jerald and Richard Marks (SIGGRAPH 2016) — Considered the perception of VR, including motion sickness and how it can be reduced in interactive VR experiences. While we will mention these issues, they will not be the focus of our course. Instead, we focus on the broader technologies, including (post-)production issues.

– Augmented reality: principles and practice[2] by Dieter Schmalstieg and Tobias Höllerer (SIGGRAPH 2016) — We will consider mostly VR technology. While the last 'future' part of our course will talk about volumetric live-action elements, which could be used in AR, we will focus more on techniques for production and post-production rather than viewing device fundamentals.

– User-centric computational videography[3] by Christian Richardt, James Tompkin, Jiamin Bai, and Christian Theobalt (SIGGRAPH 2015) — This course covers mostly 2D video, and does not cover VR.

– Video-based rendering[4] by Marcus Magnor and Marc Pollefeys (SIGGRAPH 2005) — Video-based rendering is used in video exploration systems, and has greatly improved in rendering quality in the 12 years since this course.

There is also a second set of courses which, while having some overlap, are less relevant to the topics covered in our course:

– Put on your 3D glasses now: the past, present, and future of virtual and augmented reality[5] by Douglas Lanman, Henry Fuchs, Mark Mine, Ian McDowall, and Michael Abrash (SIGGRAPH 2014) — A comprehensive survey of VR display technology, with a strong focus on head-mounted displays. We will focus on content creation and processing, and only briefly discuss displays.

– Computational plenoptic imaging[6] by Gordon Wetzstein, Ivo Ihrke, Douglas Lanman, Wolfgang Heidrich, Ramesh Raskar, and Kurt Akeley (SIGGRAPH 2012) — The course contains segments on light field capture, and "inside the Lytro camera", but did not discuss light field video, or take a post-production viewpoint.

– Lightfield photography: theory and methods[7] by Todor Georgiev and Andrew Lumsdaine (SIGGRAPH Asia 2009–2010) — The course presents the mathematical foundations of light field cameras. While this may be a useful reference for attendees of our course, and while we will cover some of the basics, we will focus more on the production aspects of video for VR.

[1] https://doi.org/10.1145/2897826.2927320
[2] https://doi.org/10.1145/2897826.2927365
[3] https://doi.org/10.1145/2776880.2792705
[4] https://doi.org/10.1145/1198555.1198678
[5] https://doi.org/10.1145/2614028.2628332
[6] https://doi.org/10.1145/2343483.2343494
[7] https://doi.org/10.1145/1900520.1900527

Course scope

In our course, we focus on the video aspect of virtual reality experiences, i.e. the visual component, and will tend not to discuss the audio component. As spatial audio is an important part of immersive VR experience, we feel it would be best discussed in its own right.

Intended audience

In short: visual computing researchers, video producers, enthusiastic video users.

As VR and video cover a wide range of technologies, our course is targeted at a broad audience and contains information for many parties: enthusiastic video users and video professionals will discover cutting-edge VR video production techniques in the industry perspectives; academics will discover new research work (CHI 2017, EG 2017) and an overview of future 6DoF challenges; video producers will gain insight into the same techniques which may, in the near future, make it into production tools; and students will benefit from the background to current state-of-the-art systems and the problems currently being tackled to bring future VR video systems to use.

Prerequisites

A basic knowledge of video processing, computer graphics, and computer vision is useful. However, we expect people without this experience to still find the course engaging—the section on storytelling and tools, for example, will be accessible to all audiences.

Level of difficulty

Intermediate.

Course notes

We will make annotated slides available on our course website `http://richardt.name/pub/Video4VR/`.

Course presenter information

Christian Richardt — *University of Bath*
christian@richardt.name • http://richardt.name

Christian Richardt is a Lecturer (=assistant professor) at the University of Bath. He received a BA and PhD in Computer Science from the University of Cambridge in 2007 and 2012, respectively. He was previously a postdoctoral researcher at Inria Sophia Antipolis, Max Planck Institute for Informatics and the Intel Visual Computing Institute. His research combines insights from vision, graphics and perception to extract and reconstruct visual information from images and videos, to create high-quality visual results and experiences, with a focus on video processing for 360 degree videos, light fields, and for user-centric applications. He has co-organized the 2015 SIGGRAPH course on User-Centric Computational Videography.

James Tompkin — *Brown University*
james_tompkin@brown.edu • http://www.jamestompkin.com

James Tompkin is an assistant professor in Computer Science at Brown University. He received an MSci degree in Computer Science from King's College London in 2006, and an EngD degree in Virtual Environments, Imaging, and Visualization from University College London in 2012. He was a postdoctoral researcher at the Max-Planck-Institute for Informatics (2012–2014), and at the Harvard Paulson School of Engineering and Applied Science (2014–2016). His research applies vision, graphics, machine learning, and interaction to create new visual computing tools and experiences. He has previously co-organized the 2015 SIGGRAPH course on User-Centric Computational Videography.

Jordan Halsey — *VR Playhouse*
jordan@vrplayhouse.com • http://vrplayhouse.com/

Jordan Reece Halsey is a Los Angeles based video artist and the technical director for VR Playhouse. At VR Playhouse, he has designed and created virtual reality experiences for Time Life, Honda, Boost Mobile, Four Seasons, and many others. Jordan is an expert in projection mapping and has worked at the forefront of emerging technologies for many years, collaborating with companies like Google, Disney, and Ford; and with artists like Madonna, Miley Cyrus and A$AP Rocky. Other notable projects include Amon Tobin's seminal Isam 2.0 tour, "Mothership 2.0" by Skrillex, and Michael Jackson's final tour and film, "This Is It".

Aaron Hertzmann — *Adobe Systems Incorporated*
hertzman@adobe.com • http://www.dgp.toronto.edu/~hertzman/

Aaron Hertzmann is a Principal Scientist at Adobe. He received a BA in Computer Science and Art/Art History from Rice University in 1996, and a PhD in Computer Science from New York University in 2001, respectively. He was a Professor at University of Toronto for ten years, and has also worked at Pixar Animation Studios, University of Washington, Microsoft Research, Mitsubishi Electric Research Lab, Interval Research Corporation and NEC Research. He is an associate editor for ACM Transactions on Graphics. His awards include the MIT TR100 (2004), a Sloan Foundation Fellowship (2006), a Microsoft New Faculty Fellowship (2006), the CACS/AIC Outstanding Young CS Researcher Award (2010), and the Steacie Prize for Natural Sciences (2010). He has spoken at four SIGGRAPH courses (most recently, the 2009 course on Human Body Movement), including one he organized (2004 course on Machine Learning).

Jonathan Starck — *Foundry*
jonathan.starck@foundry.com • https://www.foundry.com/

Jonathan Starck is Head of Research at Foundry, a global developer of computer graphics, visual effects and 3D modelling software for the design, visualization, and entertainment industries. He received a MEng in Engineering from the University of Cambridge in 1996 and a PhD in Computer Vision from University of Surrey in 2003. His research focuses on the development of image and video processing tools to accelerate artist workflows in high-end content creation. These technologies have created industry standard tools for movie post-production (NukeX and Ocula). Foundry has recently released a new toolset (CaraVR) to assist artists in the creation of high quality live-action VR content in post production, and are also exploring the video processing pipelines for immersive 6DoF VR experiences.

Oliver Wang — *Adobe Systems Incorporated*
owang@adobe.com • http://www.oliverwang.info

Oliver Wang is a Senior Research Scientist at Adobe. He received his PhD in Computer Science in 2010 from the University of California, Santa Cruz. Since then, he has worked in a number of research positions at HP Labs, Industrial Light and Magic and the Max Planck Institut Informatik, and a six year stint at Disney Research in Zurich. During this time his work led to a number of patents, and has been integrated into the production process of several mainstream movies. His research broadly spans the areas of image and video processing, computer vision, and machine learning.

Schedule

—— Introduction and Background (about 15 minutes) ——

1. **Welcome and Introduction**
   James Tompkin, Brown University, 15 minutes
   - Motivation: why care about video in VR?
   - Overview: outline of course topics.
   - Brief primer on human audio/visual perception.
   - Brief history of VR video.

—— Technical Foundations (about 55 minutes) ——

2. **360 Spherical Video**
   Oliver Wang, Adobe, 20 minutes
   - 360 video formats / spherical projection choices.
   - Content creation: Fisheye/multicam setups.
   - Stitching: Structured/unstructured (panoramic video, handheld), mesh warping (Rich360).
   - Analysis: Stabilization, recinematography, pano2vid, saliency (where do people look?), highlight detection.

3. **Stereoscopic 3D Videos and Panoramas**
   Christian Richardt, University of Bath, 20 minutes
   - Capturing and displaying stereoscopic 3D videos.
   - Viewing comfort: accommodation–vergence conflict & comfort zone.
   - Editing stereo 3D videos: StereoBrush, de-anaglyph etc.
   - Megastereo: creating high-quality stereoscopic panoramas.

4. **Q&A followed by a short break**
   15 minutes

—— Current Practice (about 75 minutes) ——

5. **Industry Perspective: 360 Video Cameras in Production**
   Jonathan Starck, Foundry, 20 minutes
   - Introduction to 360 video production: stories and experience.
   - Camera rigs: mono, stereo, Nokia OZO, Google Jump and Facebook Surround 360.
   - Challenges in production and the impact on stitching in post-production.

6. **Art, Storytelling, and Tools**
   Aaron Hertzmann, Adobe, 20 minutes
   - History of visual storytelling media.
   - VR video as an art form: special issues, needs, and artistic considerations.
   - Editing VR footage paper highlight:
     Vremiere: In-headset Virtual Reality Video Editing[8] (CHI 2017)

7. **Industry Perspective: Post-production Pipelines for 360 Video**
   Jonathan Starck, Foundry, 20 minutes
   - Introduction to post-production: node-based compositing.
   - 360 stitching pipelines: setting up rigs and stitching.
   - Challenges in post: clean up and CG integration.

8. **Q&A followed by a short break**
   15 minutes

—— The Future (about 40 minutes) ——

9. **Light Field Video Basics and Challenges**
   James Tompkin, Brown University, 20 minutes
   - Introduction to light fields: What, why, and how?
   - Capturing light fields: state-of-the-art cameras and future potential.
   - Processing light fields: filtering[9] (EG 2017), editing, challenges ahead.
   - Viewing light fields with virtual and augmented reality displays.

10. **Industry Perspective: Light Field Production and Post Production**
    Jordan Halsey, VR Playhouse, 20 minutes
    - Developing immersive experiences using light field camera arrays.
    - Connecting Google Jump to post-production workflows.
    - Real-time experiences using volumetric live-action elements.

—— Conclusion (10 minutes) ——

11. **Closing and Q&A**
    All, 10 minutes
    - Summary & outlook
    - How far to go to the ideal system? (*James Tompkin*)
    - Questions & answers

= Total time of 3 hours and 15 minutes

[8] http://web.cecs.pdx.edu/~fliu/project/vremiere/
[9] http://liris.cnrs.fr/~nbonneel/cameraarrays