

# User-Centric Computational Videography

**Christian Richardt**

*Intel Visual Computing Institute  
Max Planck Institute for Informatics*

**James Tompkin**

*Harvard School of Engineering and Applied Sciences*

**Jiamin Bai**

*Light, Co.*

**Christian Theobalt**

*Max Planck Institute for Informatics*

Course website:

[http://gvv.mpi-inf.mpg.de/teaching/uccv\\_course\\_2015/](http://gvv.mpi-inf.mpg.de/teaching/uccv_course_2015/)

## ABSTRACT

Digital video is ubiquitous: virtually every mobile device comes with at least one high-resolution video camera, and users often upload video to community websites – 300 hours of video are uploaded to YouTube alone every minute<sup>1</sup>. Yet, many commercial solutions for capturing, editing and browsing videos are difficult to use for end users, and hence constrict user creativity.

Video capture requires framing techniques and shot planning to effectively convey the intended message and be comfortable to watch. Handheld capture with mobile devices in particular often results in shaky and wobbly footage. In addition, traditional video editing tools are not keeping pace with the proliferation of video content, and some provide little more than an image-editing interface with a timeline. Unfortunately, this rather trivial addition of the temporal axis to images does not enable users to perform complex editing tasks that change the content of videos, like adding or removing objects in a video. Furthermore, existing community video collections typically treat videos like photos by having users navigate static thumbnails, instead of visualising the spatial or temporal overlaps between videos.

*User-Centric Computational Videography* aims to improve the quality and flexibility of capturing, editing, and exploring consumer videos. In this course, we discuss recent techniques in computer vision and graphics, and analyze how they have advanced towards this goal. By finding and exploiting inter- and intra-video content connections, these techniques make videos easier for amateur users, for example by enabling dynamic object removal in videos, and provide new empowering video experiences like content-based video browsing. We will take stock of the progress made so far on this topic, discuss current trends in the software industry as well as in research, and propose directions for future research.

Our course is targeted at a broad audience: enthusiastic video users will discover cutting-edge video processing techniques that may soon find their way into consumer applications, video editors will learn about powerful approaches that break from the norm of timeline editing, and researchers will benefit from a high-level overview and analysis of user-centric video techniques. We consider the first and last quarters of our course to be most suitable for beginners, as they provide the background on existing video tools and timeline editing, and discuss user interfaces for video exploration. The middle half of our course covers more advanced video editing techniques, which are also of interest in video production.

---

<sup>1</sup><https://www.youtube.com/yt/press/statistics.html>

## COURSE HISTORY

This is a new course on a topic that has so far not been covered at SIGGRAPH. The most related courses from the last 10 years are:

- [Visual algorithms for post production](#)<sup>2</sup> by Simon Robinson, Anil Kokaram and Mike Seymour (SIGGRAPH 2009) — This course covers mostly low-level video-processing techniques that we will only discuss briefly.
- [Computational photography](#)<sup>3</sup> by Paul Debevec, Ramesh Raskar and Jack Tumblin (SIGGRAPH 2005–2008) — Some of these techniques can be applied to video processing, but they are largely orthogonal to it.
- [Video-based rendering](#)<sup>4</sup> 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 10 years since this course.

Even the much broader area of video processing has not been the subject of a course in recent years, although significant advances have occurred during the past decade. We hope to address this imbalance with our course.

## COURSE NOTES

We make annotated slides available on our course website:

[http://gvv.mpi-inf.mpg.de/teaching/uccv\\_course\\_2015/](http://gvv.mpi-inf.mpg.de/teaching/uccv_course_2015/).

## INTENDED AUDIENCE

Computer graphics/vision researchers, video editors and enthusiastic video users.

## PREREQUISITES

A basic knowledge of image and video processing, computer graphics and vision is useful.

## LEVEL OF DIFFICULTY

Intermediate

## ACKNOWLEDGMENTS

We wish to thank Maneesh Agrawala for his contribution to the conception and preparation of this course.

---

<sup>2</sup><http://dx.doi.org/10.1145/1667239.1667262>

<sup>3</sup><http://dx.doi.org/10.1145/1401132.1401162>

<sup>4</sup><http://dx.doi.org/10.1145/1198555.1198678>

## COURSE PRESENTER INFORMATION

**Christian Richardt** — *Intel Visual Computing Institute, MPI for Informatics*  
[christian@richardt.name](mailto:christian@richardt.name) • <http://richardt.name>

Christian Richardt is a postdoctoral researcher at the Intel Visual Computing Institute at Saarland University and also in the Graphics, Vision and Video group at Max-Planck-Institute for Informatics in Saarbrücken, Germany. He was previously a postdoctoral fellow in the REVES team at Inria Sophia Antipolis, France, and graduated with a PhD and BA from the University of Cambridge, England, in 2012 and 2007, respectively. 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.

**James Tompkin** — *Harvard School of Engineering and Applied Sciences*  
[jtompkin@seas.harvard.edu](mailto:jtompkin@seas.harvard.edu) • <http://www.jamestompkin.com>

James Tompkin 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 from 2012 to 2014, and is currently a post-doctoral researcher at Harvard University. His research applies vision, graphics, machine learning, and interaction to create new visual computing tools and experiences.

**Jiamin Bai** — *Light, Co.*  
[jiamin@light.co](mailto:jiamin@light.co) • <https://www.linkedin.com/in/bjiamin>

Jiamin Bai is a research scientist at Light, Co. in Palo Alto. He received his PhD in Computer Science with a minor in Visual Studies from the University of California, Berkeley in 2014. He got interested in photography when he was pursuing his BSc in Electrical Engineering with a minor in Photography in Carnegie Mellon University in 2008. His research interests are user-centric video editing, computational photography and light transport.

**Christian Theobalt** — *Max Planck Institute for Informatics*  
[theobalt@mpi-inf.mpg.de](mailto:theobalt@mpi-inf.mpg.de) • <http://gvv.mpi-inf.mpg.de>

Christian Theobalt is a Professor of Computer Science at the Max-Planck-Institute for Informatics and Saarland University in Saarbrücken, Germany. Most of his research deals with algorithmic problems that lie on the boundary between the fields of Computer Vision and Computer Graphics, such as dynamic 3D scene reconstruction and marker-less motion capture, computer animation, appearance and reflectance modelling, machine learning for graphics and vision, new sensors for 3D acquisition, advanced video processing, as well as image- and physically-based rendering. He received the Otto Hahn Medal of the Max Planck Society in 2007, the EUROGRAPHICS Young Researcher Award in 2009, the German Pattern Recognition Award in 2012, and an ERC Starting Grant in 2013.

## SCHEDULE

### 1. Introduction and welcome

*Richardt, 10 minutes*

- Topic: what we are covering in our course
- Motivation: why is this area relevant and useful?
- Overview: outline of course topics

### 2. Background: state-of-the-art video tools

*Richardt, 20 minutes*

- Brief overview of techniques available in consumer and professional tools
- Video stabilization, camera tracking (SfM), image-based warping
- Segmentation, matting, compositing, transitions, color grading
- Authoring videos from photos: photo tourism, photobios

### 3. Timeline editing

*Bai, 20 minutes*

- Temporal synchronization & alignment of videos (e.g. VideoSnapping)
- Cutting of interview videos, social cameras
- Video digests

### 4. Model-free video editing

*Theobalt, 20 minutes*

- Video cut & paste: segmentation & keying
- Video inpainting for object removal
- Intrinsic videos & video relighting

### 5. Model-based video editing

*Theobalt, 20 minutes*

- Manipulation of 3D objects and environments
- Human bodies and faces
- Inverse rendering

### 6. Break

*15 minutes*

### 7. Spatiotemporal video editing and processing

*Richardt, 20 minutes*

- Advanced video transitions (multi-perspective, DuctTake, video warping)
- Motion visualization & magnification
- Virtual cinematography, hyperlapses, stereo conversion
- Editing & improving videos using photos (unwrap mosaics)

## **8. Motion editing in videos: Cinemagraphs & Cliplets**

*Bai, 20 minutes*

- Video looping, cliplets
- Selectively de-animating video
- Automatic cinemagraph portraits

## **9. Exploring videos**

*Tompkin, 20 minutes*

- Video summarization and timelapse
- Video tapestries and panopticons
- Augmenting panoramic video
- Multi-dimensional metadata exploration: sports, music, news videos

## **10. Exploring videos in contexts**

*Tompkin, 20 minutes*

- Brief history of video in context
- Integrating information (smartphones): temporal, spatial
- Navigating collections (community cameras, action cameras)

## **11. Closing and Q&A**

*all, 10 minutes*

- Summary & outlook
- Questions & answers