

Real VR – Importing the Real World into Immersive VR and Optimizing the Perceptual Experience of Head-Mounted Displays

Edited by

Marcus A. Magnor¹ and Alexander Sorkine-Hornung²

¹ TU Braunschweig, DE, magnor@cg.cs.tu-bs.de

² Facebook Zürich, CH, alexsh@fb.com

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 19272 “Real VR – Importing the Real World into Immersive VR and Optimizing the Perceptual Experience of Head-Mounted Displays”.

Motivated by the advent of mass-market VR headsets, this Dagstuhl Seminar addresses the scientific and engineering challenges that need to be overcome in order to experience omni-directional video recordings of the real world with the sense of stereoscopic, full-parallax immersion as can be provided by today’s head-mounted displays.

Since the times of the Lumière brothers, the way we watch movies hasn’t fundamentally changed: Whether in movie theaters, on mobile devices, or on TV at home, we still experience movies as outside observers, watching the action through a “peephole” whose size is defined by the angular extent of the screen. As soon as we look away from the screen or turn around, we are immediately reminded that we are only “voyeurs”. With modern full-field-of-view, head-mounted and tracked VR displays, this outside-observer paradigm of visual entertainment is quickly giving way to a fully immersive experience. Now, the action fully encompasses the viewer, drawing us in much more than was possible before.

For the time being, however, current endeavors towards immersive visual entertainment are based almost entirely on 3D graphics-generated content, limiting application scenarios to purely digital, virtual worlds only. The reason is that in order to provide for stereo vision and ego-motion parallax, which are both essential for genuine visual immersion perception, the scene must be rendered in real-time from arbitrary vantage points. While this can be easily accomplished with 3D graphics via standard GPU rendering, it is not at all straight-forward to do the same from conventional video footage acquired of real-world events.

Another challenge is that consumer-grade VR headsets feature spatial resolutions that are still considerably below foveal acuity, yielding a pixelated, subpar immersive viewing experience. At the same time, the visual perception characteristics of our fovea are decidedly different from our peripheral vision (as regards spatial and temporal resolution, color, contrast, clutter disambiguation etc.). So far, computer graphics research has focused almost entirely on foveal perception, even though our peripheral vision accounts for 99% of our field of view. To optimize perceived visual quality of head-mounted immersive displays, and to make optimal use of available computational resources, advanced VR rendering algorithms need to simultaneously account for our foveal and peripheral vision characteristics.

The aim of the seminar was to collectively fathom what needs to be done to facilitate truly immersive viewing of real-world recordings and how to enhance the immersive viewing experience by taking perceptual aspects into account. The topic touches on research aspects from various fields, ranging from digital imaging, video processing, and computer vision to computer graphics, virtual reality, and visual perception. The seminar brought together scientists, engineers and practitioners from industry and academia to form a lasting, interdisciplinary research community who set out to jointly address the challenges of Real VR.



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 3.0 Unported license

Real VR – Importing the Real World into Immersive VR and Optimizing the Perceptual Experience of Head-Mounted Displays, *Dagstuhl Reports*, Vol. 9, Issue 6, pp. 143–156

Editors: Marcus A. Magnor and Alexander Sorkine-Hornung



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Seminar June 30 – July 3, 2019 – <http://www.dagstuhl.de/19272>

2012 ACM Subject Classification Computing methodologies → Computer graphicsComputing methodologies → Computer vision

Keywords and phrases immersive digital reality, perception in vr, real-world virtual reality


Digital Object Identifier 10.4230/DagRep.9.6.143

Edited in cooperation with Susana Castillo

1 Executive Summary

Marcus A. Magnor (TU Braunschweig, DE)

Alexander Sorkine-Hornung (Facebook Zürich, CH)

License  Creative Commons BY 3.0 Unported license
© Marcus A. Magnor and Alexander Sorkine-Hornung

The Dagstuhl seminar brought together 27 researchers and practitioners from academia and industry to discuss the state-of-the-art, current challenges, as well as promising future research directions in Real VR. Real VR, as defined by the seminar participants, pursues two overarching goals: facilitating the import of real-world scenes into head-mounted displays (HMDs), and attaining perceptual realism in HMDs. The vision of Real VR is enabling to experience movies, concerts, even live sports events in HMDs with the sense of immersion of really “being-there”, unattainable by today’s technologies.

In the welcome and overview session, the participants collectively decided on the seminar program for the following days. In total, the seminar program included the overview session, three research presentation sessions, two breakout sessions including a demo track, two sessions for one-on-one discussions and individual exchange, one session for writing up the results, plus the summary and closing session.

To kick off the seminar, Alexander Sorkine-Hornung from Oculus VR presented the latest developments from an industrial perspective. He gave insights from the development of the just-released Oculus Quest and Oculus Rift S HMDs. In the research presentation sessions, 21 participants gave talks on their work. Participants also met in smaller groups in the breakout sessions to discuss the specific challenges of these fields in more detail. In due course, it became apparent that Real VR concerns research challenges in a number of different fields:

- Capture
- Reconstruction & modeling
- Rendering & perception
- Display technologies
- Interaction & virtual avatars
- Production & applications

Some exemplary results of the seminar on these topics were:

The persistent lack of consumer-market, i.e. affordable, mid- to high-resolution 360-degree video cameras to **capture** dynamic real-world scenes omnidirectionally still hamper research and development in Real VR. So far, research groups largely build their own custom-designed omnidirectional video cameras. Prominent examples include the omnidirectional camera designs by the group of Philippe Bekaert from Hasselt University, Belgium, and the top-of-the-line Manifold camera presented by Brian Cabral from Facebook. Besides novel devices, also simpler recording methods are sought, e.g. by Tobias Bertel and Christian Richardt at Bath, in order to capture real-world content more casually.

On **scene reconstruction and representation**, the jury is still out whether omnidirectional video should be considered to represent sparse light field data with dense depth/disparity as side information, or whether panoramic footage should (and could) be processed to provide full 3D geometry representations of the scene. As pointed out by Atanas Gotchev from TU Tampere, Marco Volino from the University of Surrey, and Christian Richardt from the University of Bath, both forms of representation have their respective advantages and drawbacks, e.g. when aiming to augment the real scene with additional virtual content.. Memory requirements and real-time streaming bandwidth requirements are challenging in either case.

The form of scene representation also determines which **rendering** approaches are viable. For 3D rendering, Dieter Schmalstieg from Graz presented his Shading Atlas Streaming approach to efficiently divide shading and rendering computation between server and client. To make use of visual **perception** characteristics in wide field-of-view HMDs, on the other hand, foveated rendering approaches, e.g. based on hardware ray tracing and accelerated machine learning, as presented by Anjul Patney from NVidia, have great potential. As shown by Qi Sun from Adobe, perceptual methods like saccade-aware rendering can also be used to enable walking through huge virtual worlds while actually not leaving the confines of one's living room. To render from dense depth-annotated 360-deg video, in contrast, advanced image-based warping methods and hole-filling approaches are needed, as was convincingly outlined by Tobias Bertel from the University of Bath.

Gordon Wetzstein from Stanford University presented how future HMDs will become even more realistic by overcoming current limitations of near-eye **displays**, in particular the vergence-accommodation conflict. Along similar lines, Hansung Kim from the University of Surrey showed how spatial audio enhances perceived VR realism even more.

Social **interaction** in the virtual world requires having digital doubles available. The elaborate steps needed to create convincing human avatars from real-world people were outlined by Feng Xu from Tsinghua University, Darren Cosker from the University of Bath, Christian Theobalt from MPII, and Peter Eisert from TU Berlin, covering the full range of human face, hand and body capture, reconstruction, and modeling. To interact with objects in virtual space, on the other hand, Erroll Wood from Microsoft Cambridge described how hand motion and gestures can be reliably tracked and identified in real-time by the upcoming HoloLens 2 device. Also based on real-time tracking, Li-Yi Wei from Adobe presented a system that enables presenters to augment their live presentation by interacting with the shown content in real-time using mere hand gestures and body postures.

Regarding **content production and applications**, Christian Lipski from Apple presented the ARKit software framework developed for creating captivating augmented reality experiences. James Tompkin from Brown University presented work on multi-view camera editing of Real VR content during post-production. Johanna Pirker from TU Graz showed how virtual reality can be paired with human-computer interaction to enhance learning experiences in the physics classroom. Production aspects and cinematic VR experiences were also considered prominent drivers of contemporary Real VR research by other presenters, e.g. Marco Volino, Darren Cosker, Philippe Bekaert, Peter Eisert and Brian Cabral.

Practically experiencing the new, tetherless Oculus Quest brought along by Alexander Sorkine-Hornung in the **demonstration track** made impressively clear how free, unrestricted user motion extends the usability and acceptance of VR tremendously, made possible by the pass-through view feature of this HMD.

Finally, in the coming months, a number of seminar participants will compile an edited book volume on the state-of-the-art in Real VR that Springer has already agreed to publish as part of their well-known Lecture Notes on Computer Science (LNCS) Survey Series.

2 Table of Contents

Executive Summary

<i>Marcus A. Magnor and Alexander Sorkine-Hornung</i>	144
---	-----

Overview of Talks

Bridging the gap between VR and TV <i>Philippe Bekaert</i>	148
Casual Real VR <i>Tobias Bertel</i>	148
Full-360 Inside-out Volumetric Video Capture. Six Degrees of Freedom Video. <i>Brian Cabral</i>	149
Virtual Characters and Interaction in Immersive Worlds <i>Darren Cosker</i>	149
A Perceptual Graphics Approach to Non-Verbal Communication <i>Douglas Cunningham</i>	149
Hybrid Human Modeling <i>Peter Eisert</i>	150
3D Visual Scene Sensing, Light Field Processing and Immersive Visualization <i>Atanas Gotchev</i>	150
Spatial Audio Reproduction System for VR Using 360° Cameras <i>Hansung Kim</i>	150
Introducing Apple's ARKit 3: Real-Time Computer Vision on Mobile Devices <i>Christian Lipski</i>	151
From Reality to Immersive VR. What's missing in VR? <i>Marcus A. Magnor</i>	151
Towards Deep Real-time Rendering for Mixed Reality <i>Anjul Patney</i>	152
HCI meets VR: Learning through VR <i>Johanna Pirker</i>	152
6-DoF VR Video: Towards Immersive 360° VR Video with Motion Parallax <i>Christian Richardt</i>	152
Shading Atlas Streaming <i>Dieter Schmalstieg</i>	153
Human Learning: Understanding and Computing the Eyes and Brain in VR <i>Qi Sun</i>	153
Capturing the real world for interaction in VR and AR <i>Christian Theobalt</i>	153
Multi-view Camera Editing: A Few Notes <i>James Tompkin</i>	154
Light Fields for Immersive Virtual Experiences <i>Marco Volino</i>	154

Interactive Body-driven Graphics for Augmented Video Performance
Li-Yi Wei 154

Computational Near-eye Displays & Digital Eyeglasses with Focus Cues
Gordon Wetzstein 155

Hand Tracking for HoloLens 2
Erroll Wood 155


Light Weight 3D Dynamic Reconstruction
Feng Xu 155

Participants 156

3 Overview of Talks

3.1 Bridging the gap between VR and TV

Philippe Bekaert (Hasselt University – Diepenbeek, BE)

License  Creative Commons BY 3.0 Unported license
© Philippe Bekaert

In this talk, I present a brief overview of VR/AR related R&D work at the visual computing group of the Expertise center for Digital Media of Hasselt University in Diepenbeek, Belgium, and our young spin-off AZilPix. Topics covered include 360 degrees video camera rigs and stitching, free viewpoint video, and a range of applications we have been addressing: immersive collaboration, free viewpoint video for sports broadcasting, video capture for multi-scopic (“holographic”) video displays, researching the “grammar” of the new medium which immersive video implies (with the art performance company CREW, www.crewonline.org), investigating the use of 360 video as “the next best thing to being there” (e.g. at major rock festivals), an immersive video environment for medical surgery training (with another young company dlivemed, www.dlivemed.com), and the use of immersive video as a conventional TV broadcasting production tool (the topic of our AZilPix spin-off, www.azilpix.com).

AZilPix commercialises a multi-camera video capture and processing system offering virtual pan-tilt-zoom functionality in ultra high resolution (48 megapixels) cameras. It is rooted in our expertise with 360 degree video systems, and offers the ability to mix 360 video and regular broadcast TV video in one and the same system. It offers entirely stitching artifact free TV broadcast quality 360 degrees video in a very simple and convenient way: by means of an extreme fish eye lens on our ultra-high resolution camera.

3.2 Casual Real VR

Tobias Bertel (University of Bath, GB)


License  Creative Commons BY 3.0 Unported license
© Tobias Bertel

Joint work of Tobias Bertel, Neill D.F. Campbell, Christian Richardt

Creating Real VR experiences is a very challenging and multi-disciplinary task. Real-world capturing procedures as well as realistic scene representations are hard to obtain for general environments and are key for high-quality novel-view synthesis. My talk begins by motivating image-based rendering (IBR) techniques. I focus on the 360° representation and talk about its issues. My work, MegaParallax, which provides 360° environments with motion parallax, does a step forward to provide more immersive Real VR experiences in ego-centric environments. Concepts for the current state-of-the-art in traditional and learned novel-view synthesis, namely layered scene geometry, are motivated and illustrated. My talk concludes by pointing to encountered issues during my research dealing with 360° environments and sharing a set of ideas for the shiny future, which are mainly orbiting around reliable 3D reconstruction for ego-centric camera paths, which are crucial for casual Real VR and AR applications, and deep learning.

3.3 Full-360 Inside-out Volumetric Video Capture. Six Degrees of Freedom Video.

Brian Cabral (Facebook – Menlo Park, US)

License  Creative Commons BY 3.0 Unported license
© Brian Cabral

Using a single camera array to capture depth + RGB color remains a daunting challenge for a variety of reasons ranging from the physical creation of a reliable camera rig, to managing the large data sets, and then to the reconstruction of a final 360 image. As part of our research and development we built such a system based on cinematic quality cameras and sensors. We present the HW and SW architecture, key algorithms, and technical hurdles we overcame. The talk will conclude with a presentation of results and remaining challenges.

3.4 Virtual Characters and Interaction in Immersive Worlds

Darren Cosker (University of Bath, GB)

License  Creative Commons BY 3.0 Unported license
© Darren Cosker

We are now beginning to see high quality real time characters suitable for immersive experiences. However, these typically one-off exercises require a great deal of effort, cost and expertise to create. In this talk I will outline some recent exciting examples of work in this area – including recent CAMERA work which contributed to commercial VR and video game projects with the BBC and Aardman. I will draw out some of our and our commercial partners experiences in this area – applicable to many different aspects of virtual character creation and animation – which I hope will lead to interesting future challenges.

3.5 A Perceptual Graphics Approach to Non-Verbal Communication


Douglas Cunningham (BTU Cottbus-Senftenberg, DE)

License  Creative Commons BY 3.0 Unported license
© Douglas Cunningham

By integrating the knowledge and methods from the fields of Psychology and Informatics, it is possible to address research questions that were otherwise intractable. In particular, my lab combines low-level psychophysics with modern computer graphics to allow an ecological psychology approach to modeling and synthesizing behaviorally relevant, spatio-temporal information. In this talk, I will provide an overview of recent work on the perception and (artificial) synthesis of facial expressions as well as work on the relationship between low-level image statistics and perception.

3.6 Hybrid Human Modeling

Peter Eisert (Fraunhofer-Institut – Berlin, DE)

License  Creative Commons BY 3.0 Unported license
© Peter Eisert

Photorealistic modeling and rendering of humans is extremely important for VR environments, as the human body and face are highly complex and exhibit large shape variability but also especially as humans are extremely sensitive to looking at humans. Furthermore, in VR environments, interactivity plays an important role. While purely computer graphics modeling can achieve highly realistic human models, achieving real photorealism with these models is extremely computationally expensive. In our research we investigate hybrid representations for human bodies and faces, combining classical computer graphics models with image- and video-based as well as example-based approaches in order to combine interactivity with photorealism. In this talk, we will cover recent advances at HHI in 4D human body reconstruction and modeling as well as hybrid face modeling.

3.7 3D Visual Scene Sensing, Light Field Processing and Immersive Visualization

Atanas Gotchev (Tampere University of Technology, FI)

License  Creative Commons BY 3.0 Unported license
© Atanas Gotchev

This talk summarizes three topics of our research placed along the chain of 3D visual scene sensing, light processing and immersive visualization.

First, we discuss the problem of denoising 3D scene range measurements acquired by Time-of-flight (ToF) range sensors and composed in the form of 2D image-like depth maps. We address the specific case of ToF low-sensing environment, which is set by low-light sensing conditions, low-power hardware requirements, and low-reflectivity scenes. We present an elaborated analysis of noise properties of ToF data sensed in low-sensing conditions and a related non-local denoising approach working in complex domain.

Second, we discuss the problem of light field reconstruction through sparse modelling in shearlet domain. We argue that the generated densely sampled light field of a given 3D scene is suitable for all applications which require light field reconstruction.

Third, we discuss our practical experience in developing a wide super-multiview head-up display, which augment the vehicle windshield with images flying few meters in front of the driver and maintaining smooth parallax.

3.8 Spatial Audio Reproduction System for VR Using 360° Cameras

Hansung Kim (University of Surrey – Guildford, GB)

License  Creative Commons BY 3.0 Unported license
© Hansung Kim

Recent progresses in Virtual Reality (VR) and Augmented Reality (AR) allow us to experience various VR/AR applications in our daily life. In order to maximise the immersion of user in VR/AR environments, a plausible spatial audio reproduction synchronised with visual

information is essential. In this talk, we introduce a simple and efficient system to estimate room acoustic for plausible reproduction of spatial audio using 360° cameras for VR/AR applications. A pair of 360° images is used for room geometry and acoustic property estimation. A simplified 3D geometric model of the scene is estimated by depth estimation from captured images and semantic labelling using a convolutional neural network (CNN). The real environment acoustics are characterised by frequency-dependent acoustic predictions of the scene. Spatially synchronised audio is reproduced based on the estimated geometric and acoustic properties in the scene. The reconstructed scenes are rendered with synthesised spatial audio as VR/AR content.

3.9 Introducing Apple's ARKit 3: Real-Time Computer Vision on Mobile Devices

Christian Lipski (Apple Computer Inc. – Cupertino, US)

License © Creative Commons BY 3.0 Unported license
© Christian Lipski

ARKit is a cutting-edge software framework that makes it easy for developers to create captivating augmented reality experiences for iPhone and iPad. The ARKit team performs research in the field of computer vision/machine learning, on various topics such as visual-inertial slam, face tracking, scene reconstruction and scene understanding. I will present some of the latest advancements and show which tracking and scene reconstruction algorithms run today on our mobile devices. Furthermore, I will address the challenges of ensuring consistent, high quality results of these algorithms.

3.10 From Reality to Immersive VR. What's missing in VR?

Marcus A. Magnor (TU Braunschweig, DE)


License © Creative Commons BY 3.0 Unported license
© Marcus A. Magnor

Current endeavors towards immersive visual entertainment are still almost entirely based on 3D graphics-generated content, limiting application scenarios to virtual worlds only. The reason is that in order to provide for stereo vision and ego-motion parallax, two essential ingredients for genuine visual immersion perception, the scene must be rendered in real-time from arbitrary vantage points. While this can be easily accomplished with 3D graphics via standard GPU rendering, it is not at all straight-forward to do the same from conventional video footage acquired of real-world events. In my talk I will outline avenues of research toward enabling the immersive experience of real-world recordings and how to enhance the immersive viewing experience by taking perceptual issues into account.

I will report on the latest results of our project “Immersive Digital Reality” which is funded by the German Science Foundation (DFG MA2555/15-1).

3.11 Towards Deep Real-time Rendering for Mixed Reality

Anjul Patney (Facebook – Redmond, US)

License  Creative Commons BY 3.0 Unported license
© Anjul Patney

Advances in real-time graphics have enabled some of today’s most immersive visual experiences. However, recent disruptions have transformed the landscape of technology in this area. On one hand, increasing demands from mixed-reality and other displays have forced us to rethink how we render real-time graphics for high resolutions, framerates, and fields of view. On the other hand, hardware ray tracing and accelerated machine learning have provided two promising opportunities to address these challenges. The result is an exciting opportunity to shape the next generation of interactive computer graphics. In my talk I will present how the two directions fit into my vision of the future of mixed reality graphics.

3.12 HCI meets VR: Learning through VR

Johanna Pirker (TU Graz, AT)

License  Creative Commons BY 3.0 Unported license
© Johanna Pirker

3.13 6-DoF VR Video: Towards Immersive 360° VR Video with Motion Parallax

Christian Richardt (University of Bath, GB)

License  Creative Commons BY 3.0 Unported license
© Christian Richardt

Main reference Tobias Bertel, Neill D. F. Campbell, Christian Richardt: “MegaParallax: Casual 360° Panoramas with Motion Parallax”, IEEE Trans. Vis. Comput. Graph., Vol. 25(5), pp. 1828–1835, 2019.

URL <http://dx.doi.org/10.1109/TVCG.2019.2898799>

Virtual reality is a new kind of medium that requires new ways to author content. My goal is therefore to create a new form of immersive 360-degree VR video that overcomes the limitations of existing 360-degree VR video. This new form of VR content – 6-DoF VR video – will achieve unparalleled realism and immersion by providing freedom of head motion and motion parallax, which is a vital depth cue for the human visual system and entirely missing from existing 360-degree VR video. Specifically, my aim is to accurately and comprehensively capture real-world environments, including visual dynamics such as people and moving animals or plants, and to reproduce the captured environments and their dynamics in VR with photographic realism, correct motion parallax and overall depth perception. 6-DoF VR video is a significant virtual reality capability that will be a significant step forward for overall immersion, realism and quality of experience.

3.14 Shading Atlas Streaming

Dieter Schmalstieg (TU Graz, AT)

License © Creative Commons BY 3.0 Unported license
© Dieter Schmalstieg

Main reference Joerg H. Mueller, Philip Voglreiter, Mark Dokter, Thomas Neff, Mina Makar, Markus Steinberger, Dieter Schmalstieg: “Shading atlas streaming”, ACM Trans. Graph., Vol. 37(6), pp. 199:1–199:16, 2018.

URL <http://dx.doi.org/10.1145/3272127.3275087>

Streaming high quality rendering for virtual reality applications requires minimizing perceived latency. We introduce Shading Atlas Streaming (SAS), a novel object-space rendering framework suitable for streaming virtual reality content. SAS decouples server-side shading from client-side rendering, allowing the client to perform framerate upsampling and latency compensation autonomously for short periods of time. The shading information created by the server in object space is temporally coherent and can be efficiently compressed using standard MPEG encoding. Our results show that SAS compares favorably to previous methods for remote image-based rendering in terms of image quality and network bandwidth efficiency. SAS allows highly efficient parallel allocation in a virtualized-texture-like memory hierarchy, solving a common efficiency problem of object-space shading. With SAS, untethered virtual reality headsets can benefit from high quality rendering without paying in increased latency.

3.15 Human Learning: Understanding and Computing the Eyes and Brain in VR

Qi Sun (Adobe Inc. – San José, US)

License © Creative Commons BY 3.0 Unported license
© Qi Sun

Immersive content authoring and consumption are critical components for the next generation VR platforms. It includes real-time rendering/interface software, and display hardware. In this talk, I will present my previous research on understanding and (more importantly) leveraging human perceptual factors in the eye and brain for fundamental geometry/imaging/rendering algorithms and practical applications, such as VR walk-through and computationally foveated displays.

3.16 Capturing the real world for interaction in VR and AR


Christian Theobalt (MPI für Informatik – Saarbrücken, DE)

License © Creative Commons BY 3.0 Unported license
© Christian Theobalt

In this talk, I will briefly review some of our recent work on high quality human face, body and hand capture, as well as inverse rendering with a single color or depth camera.

3.17 Multi-view Camera Editing: A Few Notes


James Tompkin (Brown University – Providence, US)

License  Creative Commons BY 3.0 Unported license
© James Tompkin

Capturing the world realistically for VR reproduction is often only the start of the creative process: editing the data comes next. In my talk, I discussed some of the problems with editing multi-camera footage and presented some of my recent work in this area. These works attempt to build a set of tools, similar to how Adobe Photoshop can edit 2D data, but for multi-camera footage. For instance, I discussed how we can filter multi-view videos in a spatio-temporally consistent way. Then, I discussed how we can perform region selection in multi-view image data via view-consistent superpixels and scribble selection. I also discussed future display devices for displaying VR content in social settings using massive tiled displays, e.g., for virtual concerts.

3.18 Light Fields for Immersive Virtual Experiences


Marco Volino (University of Surrey, GB)

License  Creative Commons BY 3.0 Unported license
© Marco Volino

There is increasing interest in creating photo-realistic cinematic experiences for virtual and augmented reality (AR/VR). The use of games engines and computer generated imagery allows inter-activity, such as head movement, but do not achieve cinematic photo-realism. Conversely, 360 and stereo 360 video achieve photo-realism for a fixed location but do not allow head movement or realistic parallax. Light field capture offers a potential solution for capturing a scene with full photo-realism whilst allowing changes in viewpoint with correct parallax. This talk will discuss recent work for the creation of light field video assets for use in cinematic VR experiences.

3.19 Interactive Body-driven Graphics for Augmented Video Performance

Li-Yi Wei (Adobe Inc. – San José, US)


License  Creative Commons BY 3.0 Unported license
© Li-Yi Wei

Joint work of Nazmus Saquib, Rubaiat Habib Kazi, Li-Yi Wei, Wilmot Li
Main reference Nazmus Saquib, Rubaiat Habib Kazi, Li-Yi Wei, Wilmot Li: “Interactive Body-Driven Graphics for Augmented Video Performance”, in Proc. of the 2019 CHI Conference on Human Factors in Computing Systems, CHI 2019, Glasgow, Scotland, UK, May 04-09, 2019, p. 622, ACM, 2019.
URL <https://doi.org/10.1145/3290605.3300852>

We naturally use hand gestures and body postures when talking to one another, but often resort to pre-created slides or videos when presenting with synthetic or virtual effects. Is there a way to combine the spontaneity of natural gestures with the power of synthetic graphics to enrich our presentation and communication? We investigate a prototype system that augments live presentations with interactive graphics to create a rich and expressive storytelling environment. Using our system, the presenter interacts with the graphical elements in real-time with gestures and postures, thus leveraging our innate, everyday skills to enhance our communication capabilities with the audience.

3.20 Computational Near-eye Displays & Digital Eyeglasses with Focus Cues


Gordon Wetzstein (Stanford University, US)

License  Creative Commons BY 3.0 Unported license
© Gordon Wetzstein

Immersive visual and experiential computing systems, i.e. virtual and augmented reality (VR/AR), are entering the consumer market and have the potential to profoundly impact our society. Applications of these systems range from communication, entertainment, education, collaborative work, simulation and training to telesurgery, phobia treatment, and basic vision research. In every immersive experience, the primary interface between the user and the digital world is the near-eye display. Thus, developing near-eye display systems that provide a high-quality user experience is of the utmost importance. Many characteristics of near-eye displays that define the quality of an experience, such as resolution, refresh rate, contrast, and field of view, have been significantly improved over the last years. However, a significant source of visual discomfort prevails: the vergence-accommodation conflict (VAC). Further, natural focus cues are not supported by any existing near-eye display. In this talk, we discuss frontiers of engineering next-generation opto-computational near-eye display systems to increase visual comfort and provide realistic and effective visual experiences.

3.21 Hand Tracking for HoloLens 2


Erroll Wood (Microsoft Research – Cambridge, GB)

License  Creative Commons BY 3.0 Unported license
© Erroll Wood

Traditional computing devices require us to adapt ourselves to digital input and output mechanisms, mediating our interactions through flat screens, button presses and touch events. Mixed reality devices like HoloLens open up a new computing paradigm, where digital content integrates seamlessly with the physical world, and our devices understand the environments we inhabit and the people that inhabit them. I will describe how HoloLens 2 tracks the user's hands, allowing the user to interact with virtual objects by simply reaching out and touching them.

3.22 Light Weight 3D Dynamic Reconstruction

Feng Xu (Tsinghua University Beijing, CN)

License  Creative Commons BY 3.0 Unported license
© Feng Xu

3D dynamic reconstruction is an important topic in graphics and computer vision. This talk majorly focuses on the reconstruction of human face, body and hand, which have many applications in HCI, VR and AR. The techniques use one or two consumer level sensors to reconstruct the motions of the target in real time, which we believe may have a great potential to be used in the future mobile devices and applications for end users.

Participants

- Philippe Bekaert
Hasselt University –
Diepenbeek, BE
- Tobias Bertel
University of Bath, GB
- Brian Cabral
Facebook – Menlo Park, US
- Susana Castillo Alejandro
TU Braunschweig, DE
- Darren Cosker
University of Bath, GB
- Douglas Cunningham
BTU Cottbus-Senftenberg, DE
- Peter Eisert
Fraunhofer-Institut – Berlin, DE
- Atanas Gotechev
Tampere University of
Technology, FI
- Adrian Hilton
University of Surrey –
Guildford, GB
- Moritz Kappel
TU Braunschweig, DE
- Hansung Kim
University of Surrey –
Guildford, GB
- Christian Lipski
Apple Computer Inc. –
Cupertino, US
- Marcus A. Magnor
TU Braunschweig, DE
- Anjul Patney
Facebook – Redmond, US
- Johanna Pirker
TU Graz, AT
- Christian Richardt
University of Bath, GB
- Dieter Schmalstieg
TU Graz, AT
- Alexander Sorkine-Hornung
Facebook Zürich, CH
- Frank Steinicke
Universität Hamburg, DE
- Qi Sun
Adobe Inc. – San José, US
- Christian Theobalt
MPI für Informatik –
Saarbrücken, DE
- James Tompkin
Brown University –
Providence, US
- Marco Volino
University of Surrey, GB
- Li-Yi Wei
Adobe Inc. – San José, US
- Gordon Wetzstein
Stanford University, US
- Erroll Wood
Microsoft Research –
Cambridge, GB
- Feng Xu
Tsinghua University Beijing, CN

