Immersive Imaging Technologies: from Capture to Display

Martin Alain
Emin Zerman
alainm@tcd.ie
zermane@tcd.ie
Trinity College Dublin
Dublin, Ireland

Cagri Ozcinar
Samsung R&D Institute
United Kingdom
cagriozcinar@gmail.com

ABSTRACT

New immersive imaging technologies enable creating multimedia systems that would increase the viewer presence and provide an immersive experience. This half-day tutorial aims to give an overview of these new immersive imaging systems and help the participants understand the content creation and delivery pipeline for the immersive imaging technologies. The tutorial will go over the full imaging pipeline, from camera setup for content capture, through content compression / streaming, to content display and related perceptual studies.

CCS CONCEPTS

• **General and reference** → **Surveys and overviews**; • **Computing methodologies** → *Image processing*; *Computer graphics.*

KEYWORDS

immersive imaging, omnidirectional imaging, light field imaging, point clouds, textured meshes

1 INTRODUCTION

The advances in imaging technologies in the last decade brought a number of alternatives to the way we acquire and display visual information. These new imaging technologies are immersive as they provide the viewer with more information which either surrounds the viewer or helps the viewer to be immersed in this augmented representation. The immersive imaging technologies include light fields [6], omnidirectional images and videos [12], and volumetric (also known as free-viewpoint) videos [1, 11]. These different modalities cover the full spectrum of immersive imaging, from 3 degrees of freedom (DoF) to 6DoF, and can be used for virtual reality (VR) as well as augmented reality (AR) [9]. Applications of immersive imaging notably include education & cultural

heritage [8], entertainment [3], remote collaboration [4], and communication [5]. In this tutorial, we cover all stages of the immersive imaging technologies from content capture to display. The main concepts of immersive imaging are first introduced, and creative experiments based on immersive imaging are presented as a specific illustration of these technologies [7]. Next, content acquisition based on single or multiple camera systems is presented, along with the corresponding data formats. Content coding is then discussed, notably ongoing standardisation efforts [2, 10], followed by adaptive streaming strategies. Immersive imaging displays are then presented, as they play a crucial role in the user's sense of immersion. Image rendering algorithms related to such displays are also explained. Finally, visual perception and quality evaluation of immersive imaging are presented.

2 TARGET AUDIENCE

This tutorial is suited for researchers, entrepreneurs, and individuals who are interested in immersive imaging who wish to expand their knowledge of immersive imaging technologies, either from a low level (e.g. content capture, camera technologies) or a high level (e.g. applications, use cases, AR/VR) perspective.

3 OUTLINE

The main outline of this tutorial will be as follows. The tutorial will be delivered in the form of pre-recorded videos and interactive online Q&A sessions will take place in between the main parts.

- Part I: Immersive Imaging Technologies (∼20 minutes)
  - Immersion and Tele-Immersion
  - Imaging Modalities
  - Applications
- Part II: Acquisition and Data Format (∼40 minutes)
  - Single-camera systems
  - Multi-camera systems
  - Data formats
- Part III: Content Delivery (∼40 minutes)
  - Coding
  - Adaptive Streaming
- Part IV: Rendering and Display Technologies (∼30 minutes)
  - Rendering algorithms
  - Traditional planar displays
  - Head-mounted displays
- Part V: Perception and Quality Evaluation (∼50 minutes)
  - Visual Perception
  - Quality Assessment
  - Visual Attention

4 MATERIALS

The tutorial slides, links to additional resources, and references, will be available on the tutorial website: v-sense.scss.tcd.ie/lectures/tutorial-on-immersive-imaging-technologies/.

5 ORGANIZER BIOS

Dr. Martin Alain received the Master's degree in electrical engineering from the Bordeaux Graduate School of Engineering (ENSEIRB-MATMECA), Bordeaux, France in 2012 and the PhD degree in signal processing and telecommunications from University of Rennes 1, Rennes, France in 2016. As a PhD student working in Technicolor and INRIA in Rennes, France, he explored novel image and video compression algorithms.

Since September 2016, he is a postdoctoral researcher in the V-SENSE project at the School of Computer Science and Statistics in Trinity College Dublin, Ireland. His research interests lie at the intersection of signal and image processing, computer vision, and computer graphics. His current topic involves light field imaging, with a focus on denoising, super-resolution, compression, scene reconstruction, and rendering.

Martin is a reviewer for the Irish Machine Vision and Image Processing conference, IEEE International Conference on Image Processing, IEEE Transactions on Image Processing, IEEE Transactions on Circuits and Systems I, IEEE Transactions on Circuits and Systems for Video Technology, and IEEE Transactions on Multimedia. He is co-organizer of the special session on Recent Advances in Immersive Imaging Technology held at EUSIPCO 2018 in Rome, ICIP 2019 in Taipei, ICME 2020 in London, and MMSP 2020 in Tampere. He co-organized the tutorial "Immersive Imaging Technologies: from Capture to Display" at ICME 2020.

Dr. Cagri Ozcinar is a senior engineer at Samsung R&D Institute UK. Previously, he was a research fellow within the V-SENSE project at Trinity College Dublin, Ireland. He was also a postdoctoral fellow in the Multimedia group at Institut Mines-Telecom Telecom ParisTech, Paris, France.

Cagri received the M.Sc. (Hons.) and the Ph.D. degrees in electronic engineering from the University of Surrey, UK, in 2010 and 2015, respectively. In his Ph.D., he was working on coding and peer-to-peer delivery solutions for multi-view video. After that, he explored coding and streaming solutions for high dynamic range imaging, 360-degree image/video, and light-fields His current research interests include machine learning, visual attention (saliency), and computer vision techniques for immersive multimedia technologies.

Cagri has been serving as a reviewer for a number of journal and conference proceedings, such as IEEE TIP, IEEE TCSVT, IEEE TMM, IEEE Journal of STSP, CVPR, IEEE ICASSP, IEEE ICIP, IEEE QoMEX, IEEE MMSP, EUSIPCO, BMVC, and WACV. Cagri has been involved in organizing workshops, challenges, and special sessions in EUSIPCO, ICIP, ICME, and MMSP. He co-organized the tutorial "Immersive Imaging Technologies: from Capture to Display" at ICME 2020.

Dr. Emin Zerman is a postdoctoral research fellow in V-SENSE project at the School of Computer Science and Statistics, Trinity College Dublin, Ireland since February 2018. He received his Ph.D. degree (2018) in Signals and Images from Télécom ParisTech, France,

and his M.Sc. degree (2013) and B.Sc. degree (2011) in Electrical and Electronics Engineering from the Middle East Technical University, Turkey. His research interests include image and video processing, immersive multimedia applications, human visual perception, high dynamic range imaging, and multimedia quality assessment.

Emin is a member of Institute of Electrical and Electronics Engineers (IEEE) and IEEE Signal Processing Society. He has been acting as a reviewer for several conferences and peer-reviewed journals, including Signal Processing: Image Communications, IEEE TTCSVT, IEEE TIP, IEEE TMM, MDPI Journal of Imaging, MDPI Applied Sciences, IEEE MMSP, IEEE ICASSP, European Signal Processing Conference (EUSIPCO), and IEEE ICIP. He is one of the special session organisers at the ICME 2020 in London and MMSP 2020 in Tampere. He co-organized the tutorial "Immersive Imaging Technologies: from Capture to Display" at ICME 2020.

ACKNOWLEDGMENTS

REFERENCES

[1] 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 (ToG)* 34, 4 (2015), 1–13.

[2] Marek Domański, Olgierd Stankiewicz, Krzysztof Wegner, and Tomasz Grajek. 2017. Immersive visual media - MPEG-I: 360 video, virtual navigation and beyond. In *International Conference on Systems, Signals and Image Processing (IWSSIP)*.

[3] Peter Graham. 2018. Cinematic Experience AWAKE: Episode One Now Available For HTC Vive. https://www.vrfocus.com/2018/12/cinematic-experience-awake-episode-one-now-available-for-htc-vive/. https://www.vrfocus.com/ - [Accessed: 2020-08-21].

[4] Simon N.B. Gunkel, Hans M. Stokking, Rick Hindriks, and Tom de Koninck. 2019. VR Conferencing: communicating and collaborating in photo-realistic social immersive environments. In *Virtual Reality International Conference (VRIC)*.

[5] Simon N. B. Gunkel, Hans M. Stokking, Martin J. Prins, Nanda van der Stap, Frank B. ter Haar, and Omar A. Niamut. 2018. Virtual Reality Conferencing: Multi-User Immersive VR Experiences on the Web. In *Proceedings of the 9th ACM Multimedia Systems Conference* (Amsterdam, Netherlands) *(MMSys '18)*. Association for Computing Machinery, New York, NY, USA, 498–501. https://doi.org/10.1145/3204949.3208115

[6] Marc Levoy and Pat Hanrahan. 1996. Light Field Rendering. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*. Association for Computing Machinery, New York, NY, USA, 31–42. https://doi.org/10.1145/237170.237199

[7] Néill O'Dwyer and Nicholas Johnson. 2019. Exploring volumetric video and narrative through Samuel Beckett's Play. *International Journal of Performance Arts and Digital Media* 15, 1 (2019), 53–69. https://doi.org/10.1080/14794713.2019.1567243 arXiv:https://doi.org/10.1080/14794713.2019.1567243

[8] Néill O'Dwyer, Jan Ondrej, Rafael Pagés, Konstantinos Amplianitis, and Aljoscha Smolic. 2018. Jonathan Swift: Augmented Reality Application for Trinity Library's Long Room. In *ICIDS*.

[9] R. Pagés, K. Amplianitis, D. Monaghan, J. Ondřej, and A. Smolić. 2018. Affordable content creation for free-viewpoint video and VR/AR applications. *Journal of Visual Communication and Image Representation* 53 (2018), 192 – 201. https://doi.org/10.1016/j.jvcir.2018.03.012

[10] Peter Schelkens, Touradj Ebrahimi, Antonin Gilles, Patrick Gioia, Kwan-Jung Oh, Fernando Pereira, Cristian Perra, and Antonio M. G. Pinheiro. 2019. JPEG Pleno: Providing representation interoperability for holographic applications and devices. *ETRI Journal* 41, 1 (2019), 93–108. https://doi.org/10.4218/etrij.2018-0509

[11] Aljoscha Smolic. 2011. 3D video and free viewpoint video—From capture to display. *Pattern Recognition* 44, 9 (2011), 1958 – 1968. https://doi.org/10.1016/j.patcog.2010.09.005 Computer Analysis of Images and Patterns.

[12] Mai Xu, Chen Li, Shanyi Zhang, and Patrick Le Callet. 2020. State-of-the-art in 360 video/image processing: Perception, assessment and compression. *IEEE Journal of Selected Topics in Signal Processing* 14, 1 (2020), 5–26.