CLIM Workshop

Technical Presentations

Retinal light-field 3D display

Alexander Kvasov, CREAL 3D, Ecublens, Switzerland.

Monocular depth perception is entirely missing in today's Virtual and Mixed Reality (VR/MR) headsets which show 3D image depth only by using binocular cues (stereoscopic vision), while each eye is forced to focus on the flat image at a certain distance. This causes visual conflicts, including vergence-accommodation conflict, which in turn cause an unpleasant eye-strain and drastically reduces the depth immersion. This problem is even more pronounced in MR applications where an eye needs to focus on both real and virtual objects at the same time. To resolve vergence-accommodation conflict, near-eye displays (NEDs) providing monocular depth cues were developed. NEDs developed by Magic Leap [1] and Avegant [2] only provide a finite number of depth planes (light-field imitation), while digital holography approach by Microsoft [3] is too complex and computationally intensive to be implemented in today's devices. NVIDIA's NEDs use defocused pinlights [4] or micro-lens array covering OLED display [5] to generate light-field image, however, suffer from relatively low image resolution or alternatively large size. We propose true light-field NEDs [6] with high image resolution, wide field of view, and continuous accommodation. Our NED has a robust design, does not have high computational requirements enabling real-time rendering, and can be scaled down to the small form factor compatible with fashion eye-wear. Moreover, CREAL3D's NED uses an array of perspective views as input and is compatible with light-field cameras allowing to display captured 3D images and video.

[1] Virtual and augmented reality systems and methods, patent US20170276948

[2] System, apparatus, and method for displaying an image using focal modulation, patent US20160295202

[3] Andrew Maimone, Andreas Georgiou, Joel S. Kollin, Holographic near-eye displays for virtual and augmented reality, ACM Transactions on Graphics (TOG), Volume 36, Issue 4, July 2017

[4] Andrew Maimone, Douglas Lanman, Kishore Rathinavel, Kurtis Keller, David Luebke, Henry Fuchs, Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources, ACM Transactions on Graphics (TOG), Volume 33 Issue 4, July 2014

[5] Douglas Lanman, David Luebke, Near-eye light field displays, ACM Transactions on Graphics (TOG), Volume 32 Issue 6, November 2013

[6] Near-eye sequential light-field projector with correct monocular depth cues, PCT application WO2018091984

Extended depth of field and broadband imaging with diffractive optical elements,

Erdem Sahin, Tampere Technical Univ., Finland.

In this talk, we discuss the extended depth of field (EDoF) and broadband imaging problems from the perspective of computational imaging. In a computational imaging setup, the optics of a conventional camera is modified or it is designed from scratch. When combined with a tailored (post-processing) computational imaging algorithm, the resulting device can be better suited for the problem at hand. In the EDoF problem, we aim at increasing the DoF of a conventional camera by employing a diffractive optical element (DoE) at the aperture plane. Such elements provide efficient solution to wavefront (phase) modulation: they are easy to fabricate and have thin form-factor (e.g., 1-2mm). We set the EDoF problem as joint optimization of the DoE with the corresponding computational imaging (i.e., deblurring) algorithm. On the other hand, in the broadband imaging problem, we discuss imaging in the visible spectrum via diffractive lenses, which are again implemented by DoEs. The main motivation of such a setup is the thin form-factor optics, and the main challenge is to correct the inherent chromatic aberrations (due to dispersion) in diffractive optics.

Development of the K|Lens – from research to market

Matthias Schmitz, K-lens, Saarbrücken, Germany

The presentation will describe the technology in principle and walk through our experience in the process of transferring research results to a market-ready product. The focus will be on hardware and software engineering challenges and potential solutions based on recent research in lightfield and 3D imaging and the compromise of reaching highest possible quality at acceptable development cost. We will also address open issues, where we still see need for further academic research and development efforts.

Multi-camera imaging with applications

Jonas Unger, Linkoping University, Sweden

This talk will give an overview of research activities around multi-sensor and light field imaging, ranging from High Dynamic Range (HDR) video capture and its use in photo-realistic computer graphics to glasses free 3D displays and recording of beating hearts during surgery. The common denominators of these research activities are multiple imaging sensors, large data sets and the requirement of efficient processing and rendering.

Computer-generated hologram synthesis using geometrical models and multi-perspective images

Jani Mäkkinen Tampere University, Finland.

Computer-generated holography enables visualization of artificially generated scenes as well as real 3D scenes recorded under white light illumination by simulating the physical wave propagation phenomenon during recording. In this talk, an overview of the methods for the synthesis of computer-generated holograms (CGH) will be presented. The methods are divided into two categories: wavefront-based and ray-based. In the wavefront-based methods geometrical representations of the scene, such as point clouds or polygons, are utilized, whereas the ray-based methods build the hologram from incoherently captured 2D images of the 3D scene. The differences between the hologram synthesis methods will be discussed from the perspective of imaging quality and computational efficiency. Moreover, the synthesis and numerical reconstruction of CGHs via simulating the viewing process of the human eye will be illustrated during the talk.

Sparse representations for compression and compressed sensing of visual data

Ehsan miandji, Linkoping University, Sweden

The ongoing advances in computational photography have introduced a range of new imaging techniques for capturing multidimensional visual data such as light fields, BRDFs, BTFs, and more. A key challenge inherent to such imaging techniques is the large amount of high dimensional visual data that is produced, often requiring GBs, or even TBs, of storage. Moreover, the utilization of these datasets in real time applications poses many difficulties due to the large memory footprint. Furthermore, the acquisition of large-scale visual data is very challenging and expensive in most cases. In this presentation, I will present the main results of my research during my PhD studies with regards to acquisition, compression, and real time rendering of high dimensional visual data in the context of computer graphics and imaging applications.

Intrinsic light field decomposition,
Anna Alperovich, Konstanz Univ., Germany.

A pipeline for lenslet light field quality enhancement"
Mikael Le Pendu, Trinity College Dublin, Ireland.

We present in this talk an advanced pipeline that improve the quality of light fields captured with handheld lenslet cameras. Such cameras are known to suffer from various visual artefacts and inconsistencies which are not corrected by current available tools used to decode RAW lenslet data. After extracting sub-aperture images from the RAW images with our demultiplexing method, we perform three correction steps. We first remove hot pixel artefacts, then correct colour inconsistencies between views using a colour transfer method, and finally we apply a state of the art light field denoising technique to ensure a high image quality."

Real-time View Synthesis for Virtual Reality and Light Field displays
Gauthier Lafruit, Univ. Libre de Bruxelles (ULB), Brussels, Belgium

Recently, MPEG-I has accepted the Reference View Synthesizer (RVS) using Depth Image-Based Rendering (DIBR) for creating any virtual view to a scene using a sparse set of camera views at fixed positions in space. We present an OpenGL real-time version - TRAVIS: To Real-time Accelerated VIew Synthesizer - based on RVS 3.1, rendering stereoscopic views on a Head Mounted Display (HMD) for Virtual Reality Free Navigation (VR-FN) in a limited volume (aka 3DoF+ in MPEG-I). We also showcase its use for Light Field displays (even for holographic stereograms which are a particular implementation of Light Field displays).

A Lightweight Neural Network for Monocular View Generation with Occlusion Handling,
Simon Evain, Inria, France

This talk will present a very lightweight neural network architecture, trained on stereo data pairs, which performs view synthesis from one single image. With the growing success of multi-view formats, this problem is increasingly relevant. The network returns a prediction built from disparity estimation, which fills in wrongly predicted regions using a occlusion handling technique. To do so, during training, the network learns to estimate the left-right consistency structural constraint on the pair of stereo input images, to be able to replicate it at test time from one single image. At test time, the approach can generate a left-side and a right-side view from one input image, as well as a depth map and a pixelwise confidence measure in the prediction. The work outperforms visually and metric-wise state-of-the-art approaches on the challenging KITTI dataset, all while reducing by a very significant order of magnitude the required number of parameters (6.5 M).

Signal processing challenges for digital holographic video display systems
Colas Schretter, Vrije Universitat Brussels, Belgium

Holography is considered to be the ultimate display technology since it can account for all human visual cues such as stereopsis and eye focusing. Aside from hardware constraints for building holographic displays, there are still many research challenges regarding holographic signal processing that need to be tackled. In this overview, we delineate the steps needed to realize an end-to-end chain from digital content acquisition to display, involving the efficient generation, representation, coding and quality assessment of digital holograms. We discuss the current state-of-the-art and what hurdles remain to be taken to pave the way towards realistic visualization of dynamic holographic content.

Scalable compression of digital holograms
Patrick Gioia, Orange lab. / IRT-BCOM

Among the emerging immersive media, digital holography is considered as the most promising technology for a natural, comfortable and authentic three-dimensional visualization. In this talk, we will present and discuss an approach exhibiting a duality between 'light rays' representation of 3D scenes and holograms space / frequency analysis. Such parallel allows considering a new type of scalability based on viewpoint degradation that might be useful in many practical applications, among which Holographic Augmented Reality Displays or real-time transmission of holographic videos. In a first part we will explain the basics of digital holography, discuss the motivation for hologram compression and emphasize the need for a better understanding of underlying holographic information. Then, while commenting recent results in the domain of Gabor wavelets-based hologram compression, we will introduce viewpoint scalability and show how such approach allows efficient degradation of data in concrete use cases.

Light field denoising and super-resolution using the LFBM5D filter
Martin Alain, Trinity College Dublin

In this talk, the LFBM5D filter is presented, which extend the concept of the state-of-the-art BM3D denoising filter for single image denoising to light field denoising. The core idea of the LFBM5D filter is to exploit redundancies over the light field angular and spatial dimensions, as well as self-similarities occurring in natural images. For that purpose, 5D patches are built from similar 2D patches, and filtered in the 5D transform domain. An extension of the LFBM5D denoising filter to spatial super-resolution is also introduced, which rely on the sparsity of the spectrum in the 5D transform domain.

Light Field denoising using using 4D Anisotropic Diffusion
Pierre Allain, INRIA

In this talk, we will present a novel light field denoising algorithm using a vector-valued regularization operating in the 4D ray space. The method performs a PDE-based anisotropic diffusion along directions defined by local structures in the 4D ray space. It does not require prior estimation of disparity maps. The local structures in the 4D light field are extracted using a 4D tensor structure. Experimental results show that the proposed denoising algorithm performs well compared to state of the art methods while keeping tractable complexity

Posters and Demos

High Dynamic Range Light Fields via Weighted Low Rank Approximation"

Mikael Le Pendu, Trinity College Dublin, Ireland.

In this poster, we will present a method for capturing High Dynamic Range (HDR) light fields with dense viewpoint sampling. Analogously to the traditional HDR acquisition process, several light fields are captured at varying exposures with a plenoptic camera. The RAW data is de-multiplexed to retrieve all light field viewpoints for each exposure and perform a soft detection of saturated pixels. Considering a matrix which concatenates all the vectorized views, we formulate the problem of recovering saturated areas as a Weighted Low Rank Approximation (WLRA) where the weights are defined from the soft saturation detection. We show that our algorithm successfully recovers the parallax in the over-exposed areas while the Truncated Nuclear Norm (TNN) minimization, traditionally used for single view HDR imaging, does not generalize to light fields. Advantages of our weighted approach as well as the simultaneous processing of all the viewpoints are also demonstrated in our experiments.

A Learning based depth estimation framework for 4D densely and sparsely sampled light field

Xiaoran Jiang, Inria, France

In this poster, we will present a learning based solution to disparity (depth) estimation for either densely or sparsely sampled light fields. Disparity between stereo pairs among a sparse subset of anchor views is first estimated by a fine-tuned FlowNet 2.0 network adapted to disparity prediction task. These coarse estimates are fused by exploiting the photo-consistency warping error, and refined by a Multi-view Stereo Refinement Network (MSRNet). The propagation of disparity from anchor viewpoints towards other viewpoints is performed by an occlusion-aware soft 3D reconstruction method. The experiments show that, both for dense and sparse light fields, our algorithm outperforms significantly the state-of-the-art algorithms, especially for subpixel accuracy.

Graph-based compression of light fields

Mira Rizkallah, Inria, France

We will present geometry aware non separable and separable local graoh transforms for energy compaction of dense 4D light fields. Local graphs are constructed on super-rays that can be seen as a grouping of spatially and geometry-dependent angularly correlated pixels. In the case of separable transforms, we show that when the shape of corresponding super-pixels in the different views is not isometric, the basis functions of the spatial transforms are not coherent, resulting in decreased correlation between spatial transform coefficients. We hence propose a novel transform optimization method that aims at preserving angular correlation even when the shapes of the super-pixels are not isometric. Experimental results show the benefit of the approach in terms of energy compaction. A coding scheme is also described to assess the rate-distortion perfomances of the proposed transforms and is compared to state of the art encoders namely HEVC and JPEG Pleno VM 1.

Multiview 360° video and video light fields datasets

Thomas Maugey, Inria and Laurent Guillo, CNRS, France

This demo will first show some downloadable video light field data sets captured with a Raytrix 2.0 plenoptic camera and will illustrate the potential of dense video light fields for 3D scene modeling, thanks to a robust depth estimation method. We will also present a new dataset in order to serve as a support for researches in Free Viewpoint Television (FTV), 6 degrees-of-freedom (6DoF) immersive communication, and light field estimation. This dataset relies on a novel acquisition procedure consisting in a synchronized capture of a scene by 40 omnidirectional cameras. We have also developed a calibration solution that estimates the position and orientation of each camera with respect to a same reference. This solution relies on a regular calibration of each individual camera, and a graph-based synchronization of all these parameters. These videos and the calibration solution are made publicly available.