CENIIT 11.01 - High Dynamic Range Video

Final project report 2018
Jonas Unger <jonas.unger@liu.se>
March 17, 2018

The project CENIIT 11.01 High Dynamic Range Video with Applications started in 2011 with a planning grant and ended June 2017. This report gives an overview of the project results and synergies created by the project.

Executive summary

To summarize the project results we would like to highlight the following items:

- **Scientific output** - This CENIIT project has directly contributed to 10 journal publications, 24 papers and tutorials at leading conferences such as SIGGRAPH, EUROGRAPHICS, and ICCP, 3 book chapters, and 1 patent.

- **Ph.D. and M.Sc. theses** - This CENIIT project has contributed two 2 Ph.D. theses and 6 M.Sc. theses.

- **Industrial collaborations** - This project has over its life-time collaborated with companies including: IKEA Communications AB, Stiller Studios, RaySpace AB, SpheronVR (Germany), VOLVO Cars, GoHDR (UK), and TP-Vision (Netherlands), Tandemlaunch (Canada), 7DLabs (USA), and IrysTec (Canada). Jonas Unger is co-founder and member of IrysTec’s and 7D Lab’s scientific advisory boards.

- **Startup companies** - Results from this project have contributed to the launch of three startup-companies, RaySpace AB developing methods, algorithms and models for measurement and representation of material properties (color, reflectance, texture) for photo-realistic image synthesis, IrysTec (Canada) developing algorithms for perceptually based image and video enhancement for displays on mobile platforms and displays in automotive applications, and 7D Labs Inc. (USA) developing software for simulation of automotive computer vision systems and generation of synthetic training data for machine learning applications.

- **Open source software** - This project has made available two software libraries under open source licenses: 1) Depends, [C15, C18], is a workflow management system for structuring and reusing compute components and fusion of large scale multi-modal computer vision data sets (e.g. laser scanner point clouds, HDR-video sequences, multi-spectral color and material measurements ) in computational imaging applications such as scene reconstruction, see
2) LumaHDRv, [C23] is a library with a (C/C++) API and a set of supporting softwares for perceptually based HDR-video encoding and decoding, see http://www.lumahdrv.org.

- **Open data** - This project has made available unique HDR-video data for research and educational purposes in video processing as well as recorded lighting environments for photo-realistic image synthesis. These data sets can be downloaded from: http://www.hdrv.org.

- **Graduate and undergraduate teaching** - Results from this CENIIT project are directly used in two undergraduate courses (TNCG13 SFX - Tricks of the trade, and TNM089 Imaging technology), and two graduate course (CADICS summer course and WASP Introduction to Autonomous Systems course).

**Project overview**

High dynamic range (HDR) imaging is rapidly becoming a core technology in a range of imaging applications. With image synthesis and visual effects/film production as the early adopters, HDR technology is now becoming adopted in many fields including automotive applications, surveillance, computer vision, film, video post processing, and everyday photography. In this project, we have taken part in and driven the development of both theory, algorithms and methodology in all steps of the HDR-video pipeline [C19]:

- **HDR-video capture** - we have developed a new statistically based framework for image reconstruction based on input from multiple image sensors or cameras. High quality image reconstruction is ensured by modelling the noise characteristics of the sensors and taking this into account in the reconstruction process [C8, J4]. We have also extended the framework to enable HDR-image reconstruction from a single RAW input image where the per pixel gain is varying over the sensor [C11, J5, B4]. This type of single image HDR capture was also implemented for off-the-shelf Canon cameras.

- **HDR-video processing and display** - An important problem in HDR imaging and video is to map the dynamic range of the HDR image/video to the (usually) much smaller dynamic range of the display device, [J9]. While an HDR image captured in a high contrast real scene often exhibit a dynamic range in the order of 7-8 log10 units, a conventional display system is limited to a dynamic range in the order of 2-4 log10 units. The mapping of pixel values from an HDR image or video sequence to the display system is called tone mapping, and is carried out using a tone mapping operator (TMO). In this project, we have developed new tone mapping algorithms which rely on

---

1HDR-video imaging is a rapidly growing paradigm in imaging. The dynamic range of an image refers to the ratio between the maximum and minimum intensities that can be measured by the imaging system. The imaging systems developed and used within this project exhibit a dynamic range in the order of 10,000,000 : 1. This should be compared to the dynamic range found in most conventional camera systems, which is in the order of 1,000 : 1
models of the human visual system to optimize the input HDR-video sequence the display by taking into account the display characteristics (e.g. black level and peak luminance), the noise properties of the input video, and the ambient lighting in the viewing environment [J3, J8, B3].

- **HDR-video compression** - In order to represent the full dynamic range and color gamut in real scenes, each frame in an HDR-video sequence is usually stored using a 32bit floating point representation. This leads to very large memory footprints for storage and requires high bandwidth during data transmission, e.g. 3 second of uncompressed 25 frames per second HDR-video at 1080p HD resolution amounts to almost 2GB of data. In this project we have used models of the human visual system to derive mappings that can be used to convert the floating point HDR-video frames to 11bit quantized integer representations that can be compressed using standard video codecs such as H.265 codecs or Google VP9. While keeping coding artifacts below the just noticeable difference (JND) threshold, our compression algorithms, [C21, C23], allows for average compression ratios in the order of 1400 : 1.

- **Photo-realistic image synthesis** - Within the project, we have developed a pipeline for capture and reconstruction of real scenes based on HDR-video and laser scan point clouds as input. By re-projecting the calibrated HDR-video data onto the recovered geometry it is possible to build a digital model of the real scene that describes both its geometry and radiometric properties. This makes it possible to place virtual objects e.g. furniture into the digital scene model and create ultra-realistic computer graphics images so that virtual objects cannot be distinguished from real, [C9, J2, J6]. We refer to this as the Virtual Photo Set (VPS). A key aspect of creating photo-realistic images is accurate simulation of scattering at surfaces in the scene. Within this project, we use HDR-imaging techniques to measure multi-spectral surface color and reflectance properties [C18]. We have also developed new models for accurate representation of Bi-Directional Reflectance Distribution Functions (BRDF) see [C3, J1], and methods for real-time image synthesis [C5, C6, J7].

**Industrial collaborations**

The industrial collaborations reflect in both joint research projects with co-authored publications as well as efforts oriented towards commercial products. On the HDR-video capture side, we collaborate with the German camera manufacturer SpheronVR AG. Together with SpheronVR, the project has developed an HDR-video camera and novel HDR reconstruction algorithms. Together with IKEA Communications AB, we have worked on joint projects in scene capture and photo-realistic image synthesis, the so called Virtual Photo Sets. IKEA have made their photo studios and 3D model database available to the project. Together with IKEA, we also have developed hardware systems for measuring reflectance properties to build digital copies of physical material samples. Together with TP Vision (formerly Philips Corporate Research in the Netherlands), we have performed a study of tone mapping operators and methods for displaying HDR-video on HDR displays. GoHDR is a company in
UK, with whom we have developed software for HDR-video compression. Together with the Canadian company IrysTec, we are building products around tone mapping and algorithms for image enhancement running directly on the display devices.

**Direct uptake of research results:** Results from the research on tone mapping described in [J3, J8] form the basis for the patent [P1] governing the core technology of the Canadian company IrysTec. As of today, IrysTec has 28 employees and is building the (world-wide) market in perceptual display technology with customers in the automotive industry (displays in cars) and manufacturers of hand-held devices such as cell phones and tablets. 7D Labs Inc. (USA) develops systems for generating synthetic training data for automotive computer vision applications such as semantic segmentation and object detection. Direct uptake of results from this CENIIT project include models for accurate simulation of camera sensors [J4, B4] and methods and algorithms for photo-realistic image synthesis [J2]. Together with 7D Labs, we have filed several provisional patent applications. Ray Space AB is using the HDR imaging technology developed within this project to develop digitization techniques for the vast number of material samples used in virtually all manufacturing industries for product development, quality control and product visualization.

**Research group buildup**

This CENIIT project has contributed to the buildup of the Computer Graphics and Image Processing group at the division Media and Information Technology at the Department of Science and Technology. As of today, the group consists of 13 co-workers and has attracted funding from most major funding agencies including the Swedish Science Council (VR), the Swedish Foundation for Strategic Research (SSF), Vinnova, and the Wallenberg Autonomous Systemts and Software Program (WASP). With a strong foundation in the theoretically oriented research, the group is active within a number of industrial and academic collaborations directed towards development of state-of-the-art applications ranging from 3D reconstruction of scenes, photorealistic image synthesis and digitization of optical material properties to computer vision for heart surgery, perceptual display algorithms and software for generating training data for autonomous systems applications such as self driving cars and robot navigation. The group is active at the forefront of the field(s) and has a strong track record of publications at internationally leading venues such as SIGGRAPH, Eurographics, and ACM Transactions on Graphics.

**Project costs and staff**

This CENIIT project has been supporting: Joel Kronander (PhD student), Gabriel Eilertsen (PhD student) and Jonas Unger (project leader). Apart from the CENIIT grant, Kronander, Eilertsen and Unger have been partially funded through other sources: GU, the CADICS Linnaeus Centre, Foundation For Strategic Research (SSF) through grant IIS11-0081, the Swedish Science Council (VR), Vinnova grants, and WASP.
Publications

Journal Publications


Conference Publications


Book chapters and Ph.D. theses


Patents