Report from Dagstuhl Seminar 13431

Real-World Visual Computing

Edited by

Oliver Grau[1], Marcus A. Magnor[2], Olga Sorkine-Hornung[3], and Christian Theobalt[4]

1 **Intel Visual Computing Institute, Saarbrücken, DE,** oliver.grau@intel.com
2 **TU Braunschweig, DE,** magnor@cg.cs.tu-bs.de
3 **ETH Zürich, CH,** hornung@disneyresearch.com
4 **MPI für Informatik – Saarbrücken, DE,** theobalt@mpi-inf.mpg.de

Abstract

Over the last decade, the tremendous increase in computational power of graphics hardware, in conjunction with equally improved rendering algorithms, have led to the situation today where real-time visual realism is computationally attainable on almost any PC, if only the digital models to be rendered were sufficiently detailed and realistic.

With rapidly advancing rendering capabilities, the modeling process has become the limiting factor in realistic computer graphics applications. Following the traditional rendering paradigm, higher visual realism can be attained only by providing more detailed and accurate scene descriptions. However, building realistic digital scene descriptions consisting of 3D geometry and object texture, surface reflectance characteristics and scene illumination, character motion and emotion is a highly labor-intensive, tedious process.

Goal of this seminar is to find new ways to overcome the looming stalemate in realistic rendering caused by traditional, time-consuming modeling. One promising alternative consists of creating digital models from real-world examples if ways can be found how to endow reconstructed models with the flexibility customary in computer graphics. The trend towards model capture from real-world examples is bolstered by new sensor technologies becoming available at mass-market prices, such as Microsoft's Kinect and time-of-flight 2D depth imagers, or Lytro's Light Field camera. Also, the pervasiveness of smart-phones containing camera, GPS and orientation sensors allows for developing new capturing paradigms of real-world events based on a swarm of networked smart-phones. With the advent of these exciting new acquisition technologies, investigating how to best integrate these novel capture modalities into the digital modeling pipeline or how to alter traditional modeling to make optimal use of new capture technologies, has become a top priority in visual computing research.

To address these challenges, interdisciplinary approaches are called for that encompass computer graphics, computer vision, and visual media production. The overall goal of the seminar is to form a lasting, interdisciplinary research community which jointly identifies and addresses the challenges in modeling from the real world and determines which research avenues will be the most promising ones to pursue over the course of the next years.

Seminar 20.–25. October, 2013 – www.dagstuhl.de/13431

1998 ACM Subject Classification I.2.10 Vision and Scene Understanding, I.3.3 Picture/Image Generation, I.3.5 Object Modeling, I.3.7 Three-Dimensional Graphics and Realism, I.4.1 Digitization and Image Capture, I.4.5 Reconstruction, I.4.8 Scene Analysis, I.4.10 Image Representation, I.4.9 Applications, I.3.1 Display devices

Keywords and phrases Image Aquisition, Scene Modeling/Rendering, Image/3D Sensors, Photorealism, Visual Effects, Motion Reconstruction, Animation

Digital Object Identifier 10.4230/DagRep.3.10.72

Edited in cooperation with Felix Klose

Real-World Visual Computing, *Dagstuhl Reports*, Vol. 3, Issue 10, pp. 72–91
Editors: Oliver Grau, Marcus A. Magnor, Olga Sorkine-Hornung, and Christian Theobalt
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Marcus Magnor
Oliver Grau
Olga Sorkine-Hornung
Christian Theobalt

Dagstuhl seminar 13431 "Real-World Visual Computing' took place October 20–25, 2013. 45 researchers from North America, Asia, and Europe discussed the state-of-the-art, contemporary challenges, and promising future research directions in the areas of acquiring, modeling, editing, and rendering of complex natural scenes and events. The seminar was encompassed an introductory and a closing session, 9 scientific presentation sessions, two book organizational sessions as well as one special session on the Uncanny Valley problem. The seminar brought together junior and senior researchers from computer graphics, computer vision, 3D animation and visual special effects, both from academia and industry, to address the challenges in real-world visual computing. Participants included international experts from Kyoto University, Tsinghua University, University of British Columbia, University of Alberta, University of North Carolina, University of Kentucky, Yale University, Technion – Haifa, Filmakademie Baden-Wuerttemberg, Hochschule der Medien Stuttgart, Disney Research Zurich, BBC Research & Development, Intel Visual Computing Institute, Nvidia Corp., Adobe Systems Inc., metaio GmbH as well as many more research institutions and high-tech companies.

Motivating this seminar was the observation that digital models of real-world entities have become an essential part of innumerous computer graphics applications today. With ever-increasing graphics hardware and software capabilities, however, so does the demand for more and more realistically detailed models. Because the traditional, labor-intensive process of digital model creation by hand threatens to stall further progress in computer graphics, conventional manual modeling approaches are giving way to new approaches that aim at capturing complex digital models directly from the real world. The seminar picked up on recent trends in acquisition hardware for real-world events (e.g., Microsoft Kinect, Lytro light field camera, swarm of smartphone sensors, ...) as well as in visual computing applications (e.g., 3D movies, Streetview, digital mock-ups, free-viewpoint systems, ...). It brought together experts from academia and industry working on contemporary challenges in image-based techniques, geometry modeling, computational photography and videography, BRDF acquisition, 3D reconstruction, 3D video, motion and performance capture etc. Collectively we fathomed the full potential of real world-based modeling approaches in computer graphics and visual computing.

Over the past decade, computer graphics has evolved into a mainstream area of computer science. Its economic impact and social pervasion range from professional training simulators to interactive entertainment, from movie production to trauma therapy, from geographic information systems to Google Earth. As a result, expectations on computer graphics performance are rising continuously. In fact, thanks to the progress in graphics hardware as well as rendering algorithms, visual realism is today within easy reach of off-the-shelf PCs, laptops, and even handheld devices. With rapidly advancing rendering capabilities, however, in many application areas of computer graphics the modeling process is becoming the limiting factor. Higher visual realism can be achieved only from more detailed and accurate scene descriptions. So far, however, digitally modeling 3D geometry and object texture, surface reflectance characteristics and scene illumination, motion and emotion is a labor-intensive, tedious process performed by highly trained animation specialists. The cost of conventionally creating models of sufficient complexity to engage the full potential of modern GPUs increasingly threatens to stall progress in computer graphics.

To overcome this bottleneck, an increasing number of researchers and engineers worldwide is investigating alternative approaches to create realistic digital models directly from real-world objects and scenes: Google and Microsoft already digitize entire cities using panorama video footage, 3D scanners, and GPS; RTT AG in Munich creates highly realistic digital mock-ups for the car industry from CAD data and measured surface reflectance characteristics of car paint; at Disney Research, algorithms are being developed to create stereoscopic movies from monocular input; and BBC R&D has developed various 3D sports visualization methods based on analyzing live-broadcast footage.

In recent years, special effects in movies and computer games have reached a new level of complexity. In their aim to construct convincing virtual environments or even virtual actors, VFX companies are more and more relying on techniques to capture models from the real world. Currently available reconstruction tools, however, are still in their infancy. A lot of time is still spent on manual post-processing and modeling. The research community has responded to this trend by investigating new image- and video-based scene reconstruction approaches that can capture richer and more complex models. An example are performance capture methods that estimate more detailed shape and motion models of dynamics scenes than do commercially available systems. Similar methods for reconstruction of entire sets are also currently investigated, but many algorithmic problems remain to be solved.

The trend towards model capture from real world-examples is additionally bolstered by new sensor technologies becoming available at mass-market prices, such as Microsoft's Kinect, time-of-flight 2D depth imagers, or Lytro's Light Field camera. Also the pervasiveness of smartphones containing a camera, GPS, and orientation sensors allows for developing new capturing paradigms of real-world events based on a swarm of networked handheld devices. With the advent of these exciting novel acquisition technologies, investigating how to best integrate these new capture modalities into the computer graphics modeling pipeline, or how to alter traditional modeling to make optimal use of the new capture approaches, has become a top priority in visual computing research.

Researchers working on all of these problems from different direction came together at the seminar to share their experiences and discuss the scientific challenges. Questions discussed were both theoretical and practical in nature. The seminar participants discussed the contemporary scientific challenges in modeling from the real world and determined which research avenues are likely to be the most promising and interesting ones to pursue over the course of the next years.

Among the questions and issues that have been addressed in the seminar are how to capitalize on new sensors for capture (computational cameras, light field cameras, Time-of-flight sensors, Kinect, omni-visual systems, ...), how to capture different object/scene aspects (geometry, reflectance, texture, material/fabric, illumination, dynamics, ...), how to digitally represent real-world objects/scenes (meshes, voxels, image-based, animation data, ...), how to convincingly & intuitively manipulate real-world models (relighting, motion editing, constrained manipulation, sketch-based, example-based, ...), how to realistically compose/augment and real-time render new scenes (F/X, movie post-production, games, perceptual issues, ...), and how to exploit the immense amount of community image and video data that are captured with handheld devices to build detailed models of the world (buildings, acting/dancing performances, sports events, fish tanks, ...). Also, the challenges arising from the large data sets of real-world models have been addressed. A special session on perceptional issues in animation (the Uncanny Valley problem) set out to identify the most important factors that are still unrealistic in computer animation. As the single most important area, facial animation was identified and some research directions for improvements were discussed.

The overall goal of the seminar to form a lasting, interdisciplinary research community was impressively underlined by the willingness of many seminar participants to work together on an edited book on the topic of the seminar. The book will be published with CRC Press. Completion of the manuscript is scheduled for August 2014.

2 Table of Contents

Executive Summary
Marcus Magnor, Oliver Grau, Olga Sorkine-Hornung, and Christian Theobalt . . . 73

Overview of Talks

Automatic Structures for Shape Modeling, Processing and Analysis
Tamy Boubekeur . . . 78

Dynamic Facial Processing and Capture: Basic Research to VFX
Darren Cosker . . . 78

BBC Research & Development and the work of the Immersive and Interactive Content team.
Robert Dawes . . . 78

Texturing from Images
Jean-Michel Dischler . . . 79

Image-based Analysis and Modelling
Peter Eisert . . . 79

New Perspectives on Uncalibrated Photometric Stereo
Paolo Favaro . . . 80

Applications of Rapid Manufacturing Tools to Real-World Visual Computing
Martin Fuchs . . . 81

Image-Based Approaches for Photo-Realistic Rendering of Clothing
Anna Hilsmann . . . 81

Accurate Binary Image Selection from Inaccurate User Input
Jan Kautz . . . 82

A Reconfigurable Camera Add-On for High Dynamic Range, Multispectral, Polarization, and Light-Field Imaging
Oliver Klehm . . . 83

Garment Replacement in Monocular Video Sequences
Felix Klose . . . 83

Online Scene Reconstruction using Point-Based Data Fusion
Andreas Kolb . . . 84

Capture Close Interacting Motions
Yebin Liu . . . 84

Multiview HDR systems
Celine Loscos . . . 85

Towards Perfect Rendering of an Imperfect (Virtual) World
Marcus A. Magnor . . . 85

3D Shape from Silhouettes in Water for Online Novel-View Synthesis
Shohei Nobuhara . . . 86

Photographic Time Machines: Taking Control over the Time of Day in Photos
Sylvain Paris . . . 86

Videos in Panoramic Contexts
Fabrizio Pece 86

Recent work on material models from captured data.
Holly E. Rushmeier 87

Accessing and Interacting with the Real World via the 3D-Web with XML3D
Philipp Slusallek 88

Challenges of high spatio-angular-temporal resolution image data
Alexander Sorkine-Hornung 89

Visualizing Point Sets
Ayellet Tal 90

Finite Element Based Tracking of Deforming Surfaces
Stefanie Wuhrer 90

Participants 91

3 Overview of Talks

3.1 Automatic Structures for Shape Modeling, Processing and Analysis

Tamy Boubekeur (Télécom Paris Tech, FR)

In this talk, I start by quickly reviewing the recent work of my group in structure-based shape analysis. I present how skeletons, cages and on- surface structures and descriptors can be computed automatically and efficiently for a wide range of applications, including interactive modeling, shape retrieval, surface simplification and animated data processing. In the second part of the talk, I focus on our most recent work on this topic, which relates to medial structures. In particular, I briefly introduce Sphere- Meshes and the Progressive Medial Axis Filtration method.

3.2 Dynamic Facial Processing and Capture: Basic Research to VFX

Darren Cosker (University of Bath, GB)

In this talk I will outline some of the challenges in creating facial performances and using facial models in visual effects. In particular, I will attempt to distinguish between academic challenges and industrial demands, and attempt to highlight some of the shared challenges.

I will also describe some of the work that myself and my group have been performing in the area of 4D facial processing. I will describe how this has led to us stepping back to focus on first solving more 'basic' (or fundamental) computer vision research problems – particularly in the area of optical flow, non-rigid tracking and shadow removal.

3.3 BBC Research & Development and the work of the Immersive and Interactive Content team.

Robert Dawes (BBC Research & Development – London, GB)

BBC R&D is the technical research department of the BBC. It is charged with researching and developing advanced and emerging media technologies for the benefit of the corporation, and wider UK and European media industries, and is also the technical design authority for a number of major technical infrastructure transformation projects for the UK broadcasting industry.

This presentation provides some background information about BBC R&D and its role and will describe the board range of work conducted by the department. It describes in detail the work of the Immersive and Interactive Content team and how it has used Visual Computing to enhance the BBC's output. This work includes virtual studio and sports telestration systems, multi-camera 3D reconstruction and metadata extraction. This work has had to deal with the constraints of broadcasting with particular emphasis on real-time operation on live video.

3.4 Texturing from Images

Jean-Michel Dischler (University of Strasbourg, FR)

Joint work of Vanhoey, Kenneth; Sauvage, Basile; Larue, Frédéric; Dischler, Jean-Michel
Main reference K. Vanhoey, B. Sauvage, F. Larue, J.-M. Dischler, "On-the-Fly Multi-Scale Infinite Texturing from Example," ACM Trans. Graph., 32(6):208:1–208:10, 2013.
URL http://dx.doi.org/10.1145/2508363.2508383

Providing efficient solutions for rendering detailed realistic environments in real-time applications, like games or flight/driving simulators, has always been a major focus in computer graphics. Details can be efficiently rendered using textures. But despite improvements of acquisition technology, graphics hardware, memory capacity and data streaming techniques, which allowed over the recent years for increased scene complexity, creating and rendering efficiently textures still remains a challenging issue. In this talk I'll present three techniques for texturing surfaces, the textures beening constructed from collections of images / photographs. The first method reconstructs noise-based textures from exemplars. It fits parameters based on a metric that computes texture descriptors. The second method manages multi-scale detail transitions by using one input image per represented scale. Undersampling artifacts are avoided by accounting for fine-scale features while colors are transferred between scales. Finally, a robust least-squares based method for fitting 2D parametric color functions on sparse and scattered data is presented. This technique allows us to reconstruct artefact-less surface lightfields over scanned 3D objects using sparse sets of photographs.

References

1 K. Vanhoey, B. Sauvage, F. Larue, J-M Dischler. *On-the-Fly Multi-Scale Infinite Texturing from Example*. ACM Trans. Graph., SIGGRAPH Asia, (2013)
2 K. Vanhoey, B. Sauvage, O. Génevaux, F. Larue, J-M Dischler. *Robust Fitting on Poorly Sampled Data for Surface Light Field Rendering and Image Relighting*. Comput. Graph. Forum 32(6): 101-112 (2013)
3 G. Gilet, J-M Dischler, D. Ghazanfarpour. *Multiple kernels noise for improved procedural texturing*. The Visual Computer 28(6-8): 679-689 (2012)
4 G. Gilet, J-M Dischler. *An Image-Based Approach for Stochastic Volumetric and Procedural Details*. Comput. Graph. Forum 29(4): 1411-1419 (2010)

3.5 Image-based Analysis and Modelling

Peter Eisert (Fraunhofer-Institut – Berlin, DE)

Joint work of Eisert, Peter; Blumenthal, David; Hilsmann, Anna

The presentation describes methods for the estimation of scene properties like geometry, motion, and deformation from multiple camera views. For 3D geometry reconstruction, a method is presented that estimates highly detailed mesh models from a pair of uncalibrated views from standard digital cameras [3, 1]. For better accuracy, lighting changes are explicitly considered in the optimization framework that jointly estimates vertex positions considering smoothness priors via Laplacians. In order to enhance the reconstruction of fine structures like hair, texture gradients are additionally exploited in the smoothness constraints. Comparisons with existing methods show better qualities and more fine details in the surface reconstruction. The method is for example used in the context of head reconstruction for virtual actor animation in the European project React. For the creation of a full head model, several partial models are fused and semantically labelled using a morphable head model as prior [4]. Another application presented is the modelling and tracking of the lower leg for navigated knee surgery. Based on a morphable leg model adapted to the reconstruction, motion and deformation is tracked over time with the optimization framework providing knowledge about bone movements. Finally, the method is applied for image-based rendering of complex objects like clothes [2] to create image-warps for view interpolation of deformable objects.

References

1 D. Blumenthal-Barby, P. Eisert. *High-Resolution Depth For Binocular Image-Based Modelling*. Computers & Graphics, accepted for publication, 2013.
2 A. Hilsmann, P. Fechteler, P. Eisert. *Pose Space Image-based Rendering*. Computer Graphics Forum (Proc. Eurographics 2013), May 2013.
3 D. C. Schneider, M. Kettern, A. Hilsmann, P. Eisert. *Deformable Image Alignment as a Source of Stereo Correspondences on Portraits*. CVPR'11 Workshop NORDIA, June 2011.
4 D. Schneider, P. Eisert. *Algorithms for Automatic and Robust Registration of 3D Head Scans*. Journal of Virtual Reality and Broadcasting, Oct. 2010.

3.6 New Perspectives on Uncalibrated Photometric Stereo

Paolo Favaro (Universität Bern, CH)

Joint work of Favaro, Paolo; Papadhimitri, Thoma
Main reference T. Papadhimitri, P. Favaro, "A Closed-Form, Consistent and Robust Solution to Uncalibrated Photometric Stereo Via Local Diffuse Reflectance Maxima," Int'l J. of Computer Vision, 16 pp., October 2013.
URL http://dx.doi.org/10.1007/s11263-013-0665-5

The Lambertian reflectance model is far from being a realistic model of light propagation on real objects. Yet, it remains one of the most widely adopted models for shading in computer vision and has led to very effective algorithms for 3D reconstruction. Surprisingly, this simple model has still more to offer. We illustrate this in the well-known uncalibrated photometric stereo problem, where one aims at reconstructing the normal map and the albedo of an object and the scene illumination given only a collection of images. At the core of uncalibrated photometric stereo is the integrability constraint, which allows to obtain a solution up to 3 free parameters, the so-called generalized bas-relief (GBR) ambiguity. Existing approaches are not able to fix these parameters as they employ heuristics that yield infinite solutions. In contrast, we show that a careful study of the Lambertian model yields two possible exact solutions: one exploiting locations of maximal diffuse brightness [2] and another exploiting the perspective effect in the camera projection model [1]. Both methods are validated on synthetic and real data and achieve state-of-the-art results.

References

1 Thoma Papadhimitri and Paolo Favaro. *A New Perspective on Uncalibrated Photometric Stereo*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013
2 Thoma Papadhimitri and Paolo Favaro. *A Closed-Form, Consistent and Robust Solution to Uncalibrated Photometric Stereo Via Local Diffuse Reflectance Maxima*. International Journal of Computer Vision, pp 1–16, 2013

3.7 Applications of Rapid Manufacturing Tools to Real-World Visual Computing

Martin Fuchs (Universität Stuttgart, DE)

Joint work of Fuchs, Martin; Hasan, Milos; Kächele, Markus; Lensch, Hendrik P. A.; Matusik, Wojciech; Pfister, Hanspeter; Raskar, Ramesh; Rusinkiewicz, Szymon; Seidel, Hans-Peter

In the last years, rapid manufacturing tools have begun to have an increasing impact on my work on the visual computing cycle. In this talk, I present several works I contributed to for which these tools were crucial in enabling success.

At SIGGRAPH 2008, we introduced passive reflectance field displays, which optically shape incident illumination into view-dependent pictures which can be programmed in up to six dimensions [1]. This requires precise alignment of numerous optical components, which was achieved by 3D printing enclosing structures with a fused deposition modeling process.

Two years later, at SIGGRAPH 2010, 3D printing stackings of selected base materials permitted us to manufacture objects with defined subsurface scattering properties [3], and recently [2], we performed CNC milling and laser cutting to inexpensively create precisely designed faceted mirrors which enabled light field acquisition with a catadioptric setup.

Open problems include simplified access to the internal representations of the manufacturing machines so as to bridge the gap between visual computing and mechanical engineering, and the research into a principled approach to assess the gamut of a fabrication pipeline for visual models.

References

1 Martin Fuchs, Ramesh Raskar, Hans-Peter Seidel and Hendrik P. A. Lensch. *Towards Passive 6D Reflectance Field Displays.* ACM Transactions on Graphics (Proc. SIGGRAPH special issue) 27(3) , 2008

2 Martin Fuchs, Markus Kächele and Szymon Rusinkiewicz. *Design and Fabrication of Faceted Mirror Arrays for Light Field Capture.* VMV 2012: Vision, Modeling & Visualization, 2012. An extended version is to appear at Computer Graphics Forum, DOI: 10.1111/cgf.12201.

3 Milos Hasan, Martin Fuchs, Wojciech Matusik, Hanspeter Pfister, and Szymon Rusinkiewicz. *Physical reproduction of materials with specified subsurface scattering.* ACM Transactions on Graphics (Proc. SIGGRAPH special issue) 29(4), 2010

3.8 Image-Based Approaches for Photo-Realistic Rendering of Clothing

Anna Hilsmann (HU Berlin, DE)

Joint work of Hilsmann, Anna; Fechteler Philipp; Eisert, Peter
Main reference A. Hilsmann, P. Fechteler, P. Eisert, “Pose Space Image Based Rendering,” Computer Graphics Forum, 32(2pt3):265–274, 2013.
URL http://dx.doi.org/10.1111/cgf.12046

One principal intention of computer graphics is the achievement of photorealism. Although modeling, animation and simulation tools for rendering of complex objects – e.g. human bodies, faces, or clothes – have been developed in the last decades, achieving photorealism by physically simulating material properties and illumination is still computationally demanding and time-consuming. Especially clothing is extremely difficult to simulate physically, because cloth deformation and drapery exhibit many degrees of freedom and wrinkles produce complex shading and texture deformations. Yet, these complex details are essential for a realistic appearance of the rendered clothes.

An alternative to synthesis and reconstruction is observation of appearance through a number of images. This is addressed by image-based rendering (IBR) approaches, which synthesize new images by appropriately interpolating and merging a database of prerecorded images. However, the database images typically show various viewpoints of a rigid and stationary object or scene and the synthesis of new images is limited to viewpoint changes. More complex animations are typically not possible.

For clothing that roughly follows the shape of a human body, it is a reasonable assumption that wrinkling depends on the articulated pose of a human body. Under this assumption, a new image-based rendering approach is proposed, which synthesizes new images of such types of clothing from a database of pre-recorded images based on pose information [2]. Image warps, i.e. transformation rules between the images, implicitly extract pose-dependent appearance and shading from the images. These warps are extracted both in the spatial as well as in the photometric domain [1]. For rendering, the images and warps are parametrized and interpolated in pose-space, i.e. the space of body poses, using scattered data interpolation methods that have already been successfully exploited in example-based animation methods. The high-dimensionality of the pose-space is handled by subdividing the space into subspaces, which we assume can be handled independently, assuming that wrinkling of tight-fitting clothes mainly depends on the nearest joints. This reduces the dimensionality of the interpolation domain and thereby allows for a larger variety of possible poses that can be synthesized.

Altogether, the presented approach shifts computational complexity from the rendering to an a-priori training phase. The use of real images and warp-based extraction of deformation and shading allow a photo-realistic visualization and modification of clothes, including fine details, without computationally demanding simulation of the underlying scene and object properties.

References

1. A. Hilsmann and D.C. Schneider and P. Eisert. *Realistic Cloth Augmentation in Single View under Occlusion*. Computers & Graphics 34(5), pp. 567–574, Oct. 2010.
2. A. Hilsmann and P.Fechteler and P. Eisert. *Pose Space Image Based Rendering*. Computer Graphics Forum 32(2) (Proc. of Eurographics 2013), pp. 265–274, May 2013.

3.9 Accurate Binary Image Selection from Inaccurate User Input

Jan Kautz (University College London, GB)

Joint work of Subr, Kartic; Paris, Sylvain; Soler, Cyril; Kautz, Jan
Main reference K. Subr, S. Paris, C. Soler, J. Kautz, "Accurate Binary Image Selection from Inaccurate User Input," Computer Graphics Forum, 32(2pt1):41–50, 2013.
URL http://dx.doi.org/10.1111/cgf.12024

Selections are central to image editing, e.g., they are the starting point of common operations such as copy-pasting and local edits. Creating them by hand is particularly tedious and scribble-based techniques have been introduced to assist the process. By interpolating a few strokes specified by users, these methods generate precise selections. However, most of the algorithms assume a 100% accurate input, and even small inaccuracies in the scribbles often degrade the selection quality, which imposes an additional burden on users. We propose a selection technique tolerant to input inaccuracies. We use a dense conditional random field (CRF) to robustly infer a selection from possibly inaccurate input. Further, we show that patch-based pixel similarity functions yield more precise selection than simple point-wise metrics. However, efficiently solving a dense CRF is only possible in low-dimensional Euclidean spaces, and the metrics that we use are high-dimensional and often non-Euclidean. We address this challenge by embedding pixels in a low-dimensional Euclidean space with a metric that approximates the desired similarity function. The results show that our approach performs better than previous techniques and that two options are sufficient to cover a variety of images depending on whether the objects are textured.

3.10 A Reconfigurable Camera Add-On for High Dynamic Range, Multispectral, Polarization, and Light-Field Imaging

Oliver Klehm (MPI für Informatik – Saarbrücken, DE)

Joint work of Manakov, Alkhazur; Restrepo, John F.; Klehm, Oliver; Hegedüs, Ramon; Eisemann, Elmar; Seidel, Hans-Peter; Ihrke, Ivo
Main reference A. Manakov, J. F. Restrepo, O. Klehm, R. Hegedüs, E. Eisemann, H.-P. Seidel, I. Ihrke, "A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging," ACM Trans. Graph., 32(4): 47:1–47:14, 2013..
URL http://dx.doi.org/10.1145/2461912.2461937

We propose a non-permanent add-on that enables plenoptic imaging with standard cameras. Our design is based on a physical copying mechanism that multiplies a sensor image into a number of identical copies that still carry the plenoptic information of interest. Via different optical filters, we can then recover the desired information. A minor modification of the design also allows for aperture sub-sampling and, hence, light-field imaging. As the filters in our design are exchangeable, a reconfiguration for different imaging purposes is possible. We show in a prototype setup that high dynamic range, multispectral, polarization, and light-field imaging can be achieved with our design.

3.11 Garment Replacement in Monocular Video Sequences

Felix Klose (TU Braunschweig, DE)

Joint work of Rogge, Lorenz; Klose, Felix; Stengel, Michael; Eisemann, Martin; Magnor, Marcus

In this talk a semi-automatic approach for replacing clothes in a monocular video sequences will be presented. From a monocular video sequence the actors body shape and pose is reconstructed using a parameterized body model. Using this animated body model as a mannequin arbitrary virtual garments can now be simulated. To obtain plausible renderings of this garment an additional scene lighting reconstruction is performed to replicate the original lighting situation. The remaining body shape and pose misalignments are then corrected purely in image space by computing and applying a pixel wise 2D image warp to the rendered image of the simulated garment. The results are then composed back onto the original video sequence.

3.12 Online Scene Reconstruction using Point-Based Data Fusion

Andreas Kolb (Universität Siegen, DE)

Joint work of Keller, Maik; Lefloch, Damien; Lambers, Martin; Izadi, Shahram; Weyrich, Tim; Kolb, Andreas
Main reference M. Keller, D. Lefloch, M. Lambers, S. Izadi, T. Weyrich, A. Kolb, "Real-time 3D Reconstruction in Dynamic Scenes using Point-based Fusion," in Proc. of the 2013 Int'l Conf. on 3D Vision (3DV'13), pp. 1–8, IEEE, 2013.
URL http://dx.doi.org/10.1109/3DV.2013.9

Online range image acquisition extremely evolved in the last view years due to new sensing devices like Time-of-Flight (ToF) camera or the Microsoft Kinect. Even though, this kind of sensors are well suited for accessing individual range images, using this kind of device for online scene acquisition is a challenging task. Processing paradigms that tackle this challenge need to deal with the high level of noise and the high frame rate / bandwidth of these sensors.

This talk discusses an approach based on KinectFusion using solely point-based representations. This approach allows for low-level handling of scene dynamics. Furthermore, open problems are discussed and more general lines of future research are sketched.

3.13 Capture Close Interacting Motions

Yebin Liu (Tsinghua University – Beijing, CN)

Joint work of Liu, Yebin; Wang, Yangang; Theobalt, Christian; Chai, Jinxiang; Dai, Qionghai; Gall, Juergen; Min, Jianyuan; Stoll, Carsten; Zhang, Jianjie; Xu, Feng; Seidel, Hans-Peter
Main reference Y. Liu, J. Gall, C. Stoll, Q. Dai, H.-P. Seidel, C. Theobalt, "Markerless Motion Capture of Multiple Characters Using Multi-view Image Segmentation," IEEE Trans. Pattern Anal. Mach. Intell., 35(11):2720–2735, 2013.
URL http://dx.doi.org/10.1109/TPAMI.2013.47
Main reference Y. Wang, J. Min, J. Zhang, Y. Liu, F. Xu, Q. Dai, J. Chai, "Video-based Hand Manipulation Capture Through Composite Motion Control," ACM Trans. Graph., 32(4):43:1–43:14, 2013.
URL http://dx.doi.org/10.1145/2461912.2462000

The talk presents marker-less based motion capture of close interacting motions. Close interacting motion is common in everyday life and there are important for graphics animation, user interaction, sport analysis, biomechanics and so on. The two typical cases are human body interacting with human body and hand interacting with object. Even for maker motion capture, it is challenging to have these motions, because of the serious occlusion, possible collision and the requirement for capture of subtle contact phenomena. We start the talk from traditional single character or single object motion capture. Given the observed image data, we usually try to search in the pose space, to find out the pose that minimize the difference between the observed images and the rendering result. With a sample in the pose space, we usually first skinning the pose to a 3D mesh and then possibly texture mapping on it to get the appearance. Such way is an analysis by synthesis approach with sampling then rendering and then measurement. In this talk, I will two works showing how analysis by synthesis approach can be used in reconstruction of close interacting motions. 1) Multi-view body labeling using Analysis by synthesis in multi- person motion capture; 2) Analysis by synthesis integrated with Motion Control for hand-object motion capture.

3.14 Multiview HDR systems

Celine Loscos (Universite de Reims, FR)

Joint work of Loscos, Celine; Bonnard, Jennifer; Valette, Gilles; Lucas, Laurent; Remion, Yannick; Ismael, Muhannad; Blache, Ludovic; Prevost, Stephanie; Nocent, Olivier
Main reference L. Lucas, C. Loscos, Y. Remion, (eds.), "3D Video: From Capture to Diffusion," 325 pp., ISBN: 978-1-84821-507-8, Wiley-ISTE, 2013.

Multiview systems are nowadays used in many applications, but they can serve different objectives: stereo acquisition for 3D visualization and multi-viewpoint acquisition for 3D reconstruction. In this presentation, I will go through the details of two different approaches, that are work in progress conducted by a team of researchers of CReSTIC-SIC at the University of Reims on: - 3D HDR video acquisition for auto-stereoscopic visualization, providing HDR pixel values to eight aligned views, - a 30-camera acquisition studio for reconstructing animated models of a moving actor, as well as a hybrid 4D modeling algorithm, associating visual hull with multiscopic approaches to gain precision.

3.15 Towards Perfect Rendering of an Imperfect (Virtual) World

Marcus A. Magnor (TU Braunschweig, DE)

Main reference M. Mustafa, L. Lindemann, M. Magnor, "EEG analysis of implicit human visual perception," in Proc. of the 2012 ACM SIGCHI Conf. on Human Factors in Computing Systems (CHI'12), pp. 513–516, ACM, 2012.
URL http://dx.doi.org/10.1145/2207676.2207746

Noise, outliers, and reconstruction errors will likely be accompanying real world-acquired digital models and scene representations for a long time to come. This talk presents strategies to assess the impact of model and rendering artifacts on visual perception based on electroencephalography (EEG).

The advantages of incorporating knowledge about our human visual system into graphics and visualization algorithms have been apparent for a long time. But while a considerable amount of work has been done on measuring cognitive processing in graphics and visualization through the use of external processes, e.g. surveys and direct observation, much less research has so far been devoted to take advantage of brain measurements. Due to recent hardware advances in the consumer market, primarily driven by applications to control games by thought (e.g. Emotiv's brain-computer interface), electroencephalography (EEG) has become an affordable technique to measure the visual processing in the brain.

It has long been understood that visual processing occurs without conscious perception or attention, and that conscious awareness of a visual stimulus is preceded by complex visual decision-making processes. Directly recording brain activity via EEG enables tapping into these implicit visual processes. A correctly designed EEG experiment is able to robustly and reproducibly disclose novel aspects of our visual perception, undeterred by decision-forming processes that necessarily precede any conscious answer to the questions in a conventional user study. On the other hand, successfully conducting an EEG experiment requires specialized knowledge of signal processing in the brain, in particular the strengths and limitations of the EEG technique. EEG measurements do not supplant traditional, question-based user studies, but offer access to new aspects of visual perception that are otherwise inaccessible.

3.16 3D Shape from Silhouettes in Water for Online Novel-View Synthesis

Shohei Nobuhara (Kyoto University, JP)

This paper is aimed at presenting a new algorithm for full 3D shape reconstruction and online free-viewpoint rendering of objects in water. The key contributions are (1) a new calibration model for the refractive projection, and (2) a new 3D shape reconstruction algorithm based on shape-from- silhouette (SfS) concept. We also propose an online free-viewpoint rendering system as a practical application.

3.17 Photographic Time Machines: Taking Control over the Time of Day in Photos

Sylvain Paris (Adobe Systems Inc. – Cambridge, US)

Joint work of Laffont, Pierre-Yves; Bousseau, Adrien; Drettakis, George; Shih, YiChang; Freeman, William; Durand, Frédo; Paris, Sylvain

Taking a picture in the middle of the day or during sunset makes a huge difference in the produced image. The changes of illumination throughout the day are a critical aspect of outdoor photography. These variations open creative opportunities but also constrain how outdoor photographers work. I will present two approaches to alter the illumination conditions in pictures of outdoor scenes to render them as if they had been taken at a different time of day. The first one relies on internet photo collections and partial 3D reconstructions, and the second on a dataset of time-lapse videos and scene matching. I will dedicate part of my talk to a discussion of the pros and cons of each approach. This is joint work with Pierre- Yves Laffont, Adrien Bousseau, George Drettakis, YiChang Shih, Bill Freeman, and Frédo Durand.

3.18 Videos in Panoramic Contexts

Fabrizio Pece (University College London, GB)

Joint work of Pece, Fabrizio; William, Steptoe; James, Tompkin; Rajvi, Shah; Simon, Julier; Fabian, Wanner; Tim, Weyrich; Sharham, Izadi; Christian, Theobalt; Jan, Kautz; Anthony, Steed
Main reference F. Pece, W. Steptoe, S. Julier, F. Wanner, T. Weyrich, J. Kautz, A. Steed, "PanoInserts: Mobile Spatial Teleconferencing," in Proc. of the 2013 ACM SIGCHI Conf. on Human Factors in Computing Systems (CHI'13) pp. 1319–1328, ACM, 2013.
URL http://dx.doi.org/10.1145/2470654.2466173

Panoramic images and video are now common, with the world quickly being mapped at street level by companies and tourists alike. On a spectrum between 3D virtual environments and 2D images, panoramas lie somewhere in between – a 360 degrees panorama can surround a user, but the scene has only spherical geometry and is effectively flat. For these reasons, panorama are an attractive basis for video- conferencing, as they provide a full 360 degree view of an environment in a single image, but they are also a convenient context to temporally and spatially relate videos within large collections.

In this talk I have presented two videos-in-panoramic-context systems, together with their evaluations.

Firstly, I have introduced PanoInserts (Pece at al. 2013): a novel tele- conferencing system that uses smartphone cameras to create a surround representation of meeting places. I have showed the results of a user study comparing our system with fully-panoramic video and conventional webcam video conferencing for two spatial reasoning tasks, and I have discussed the representational properties and usability of varying video presentations, exploring how they are perceived and how they influence users when performing spatial reasoning tasks.

I have then presented Vidicontexts (Tompkin et al.): a video-collections+context interface that embeds videos into a panorama. The system features a spatio-temporal index and tools for fast exploration of the space and time of the video collection. I have presented the results of our study on the effect of our video-collection+context system to spatio-temporal localization tasks.

Finally, I have showed various adaptations of VidiContexts to several display devices, discussing the implications of of each of them on videos-in-panoramic- context interfaces.

References

1 J. Tompkin, F. Pece, R. Shah, S. Izadi, J. Kautz and C. Theobalt. *Video Collections in Panoramic Contexts*. Symposium on User Interface Software and Technology (UIST 2013) St Andrews, UK, Oct. 8-11, 2013.

2 F. Pece, W. Steptoe, S. Julier, F. Wanner,T. Weyrich, J. Kautz, and A. Steed. *PanoInserts: Mobile Spatial Teleconferencing*. ACM CHI Human Factors in Computing Systems (CHI 2013) Paris, France, April 22-May 2 2013.

3.19 Recent work on material models from captured data.

Holly E. Rushmeier (Yale University, US)

Joint work of Rushmeier, Holly E.; Dorsey, Julie; Xue, Su; Lockerman, Yitzchak; Wu, Hongzhi
Main reference Y. D. Lockerman, S. Xue, J. Dorsey, H. E. Rushmeier, "Creating Texture Exemplars from Unconstrained Images," Report no. TR1483, Yale University, 2013.
URL http://graphics.cs.yale.edu/site/tr1483
Main reference H. Wu, J. Dorsey, H. E. Rushmeier, "Inverse Bi-Scale Material Design," ACM Trans. Graph., 32(6):163:1–163:10, 2013.
URL http://dx.doi.org/10.1145/2508363.2508394

We discuss the extraction of uniform textures from natural images. First, we present a method of clustering textures in an image using diffusion distances. We then modify this approach using simple user input for the scale and sample location of the desired texture to form an interactive technique for extracing the texture.

We also describe bi-scale reflectance modeling. First we show how the large scale BRDF can be computed from small scale BRDF and geometry by formulating this as the multiplication of large matrices which are then simplified using a randomization technique. Inverting this process is possible by using a large library of geometries and BRDF that are represented by small, approximate matrices.

3.20 Accessing and Interacting with the Real World via the 3D-Web with XML3D

Philipp Slusallek (DFKI – Saarbrücken, DE)

Joint work of Klein, Felix; Rubinstein, Dmitri; Sons, Kristian; Einabadi, Farshad; Herhut, Stephan; Slusallek, Philipp
Main reference F. Klein, D. Rubinstein, K. Sons, F. Einabadi, S. Herhut, P. Slusallek, "Declarative AR and Image Processing on the Web with Xflow," in Proc. of the 18th Int'l Conf. on Web 3D Technology (Web3D'13), pp. 157–165, ACM, 2013.
URL http://dx.doi.org/10.1145/2466533.2466544
URL https://graphics.cg.uni-saarland.de/2013/declarative-ar-and-image-processing-on-the-web-with-xflow/

A key element for Real World Visual Computing is the ability of non-experts to get access to what we come up with in our research. In the past this has been more than problematic mainly because of the large efforts of developing advanced interactive Visual Computing applications as well as the problems of deploying them across different devices, hardware architectures, and operating systems. The availability of the low-level WebGL interface in the latest browsers has not really changed this situation much.

With XML3D we provide a high-level interface to interactive 3D Graphics and Visual Computing. XML3D is an extension to HTML-5 that runs in all modern browsers via an implementation in JavaScript that efficiently simulates native support by the browser. We use HTML to describe the 3D scene, reusing existing Web technologies (like CSS, DOM events, JS, etc.) wherever possible. This makes it possible for millions of experienced Web developers to easily make use of and integrate interactive 3D graphics into their Web applications.

But XML3D goes significantly further than traditional scene graphs: Its novel generic data handling concept allows highly flexible and GPU-friendly operations while data can be asynchronously be retrieved via simple URLs from anywhere in the Internet (just like HTML images). Semantic annotations allow for intelligent processing (even AI) within the 3D scenes and linking the objects back to their origin, such that distributed Web services can easily be created. Real-time synchronization of 3D scenes among many users is currently being added to XML3D.

A declarative flow-graph approach integrates efficient animation, image processing, and AR tracking operators flexibly and directly into XML3D. This allows for portable and interactive AR application directly in any mobile device. Parts of the flow-graph can be efficiently and portably be mapped to the available hardware (e.g. via vertex or compute shaders). Our new material model "shade.js" provides portable material descriptions that allow XML3D scenes to be rendered equally well using rasterization (forward or deferred) and ray tracing.

In my presentation, I give an overview of the XML3D technology and highlight some of its core capabilities and applications that have already been realized. During the workshop I would like to explore opportunities how XML3D can be used more widely also in the research community – both as a tool for the research itself as well as for delivering the research results directly and interactively to users.

References

1 Felix Klein, Dmitri Rubinstein, Kristian Sons, Farshad Einabadi, Stephan Herhut, and Philipp Slusallek, *Declarative AR and Image Processing on the Web with Xflow*, Proceedings of the 18th International Conference on Web 3D Technology of Web3D'13 , pp. 157–165, June 2013

2 Jozef Dobos, Kristian Sons, Dmitri Rubinstein, Philipp Slusallek, and Anthony Steed, *XML3DRepo: A REST API for Version Controlled 3D Assets on the Web*, Proceedings of the 18th International Conference on 3D Web Technology of Web3D'13 , pp. 47–55, June 2013
3 Jacek Jankowski, Sandy Ressler, Kristian Sons, Yvonne Jung, Johannes Behr, and Philipp Slusallek, *Declarative Integration of Interactive 3D Graphics into the World-Wide Web: Principles, Current Approaches, and Research Agenda*, Proceedings of the 18th International Conference on 3D Web Technology of Web3D'13 , pp. 39–45, June 2013
4 Kristian Sons, Christian Schlinkmann, Felix Klein, Dmitri Rubinstein, and Philipp Slusallek, *xml3d.js: Architecture of a Polyfill Implementation of XML3D*, 6th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS), 2013
5 Felix Klein, Kristian Sons, Dmitri Rubinstein, Sergiy Byelozyorov, Stefan John, and Philipp Slusallek, *Xflow – Declarative Data Processing for the Web*, Proceedings of the 17th International Conference on Web 3D Technology , pp. 37–45, August 2012
6 Kristian Sons, Felix Klein, Dmitri Rubinstein, Sergiy Byelozyorov, and Philipp Slusallek, *XML3D – Interactive 3D Graphics for the Web*, Proceedings of the 15th International Conference on 3D Web Technology, pp. 175–184, July 2010

3.21 Challenges of high spatio-angular-temporal resolution image data

Alexander Sorkine-Hornung (Disney Research – Zürich, CH)

Joint work of Sorkine-Hornung, Alexander; Sorkine-Hornung, Olga; Kim, Changil; Zimmer, Henning; Pritch, Yael;Perazzi, Federico; Yücer, Kaan; Wang, Oliver; Gross, Markus
Main reference C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, M. Gross, "Scene Reconstruction from High Spatio-Angular Resolution Light Fields," ACM Trans. Graph., 32(4):73:1–73:12, 2013.
URL http://dx.doi.org/10.1145/2461912.2461926

Sensors capture the world at ever higher spatial, temporal, and angular resolution. For example, todays professional cinema cameras record 120fps at 4k and beyond. For the industry, one of the most obvious resulting problems is data management. However, an additional challenge arises, since many of our existing algorithms simply have not been designed to handle such densely sampled input. Besides the issue of increased computational complexity, in the past we have carefully tuned our existing methods to address issues of sparse sampling. For instance, standard strategies in image-based reconstruction and optical flow involve global regularization to handle ambiguous, incomplete, or noisy measurements. The corresponding benchmarks for stereo and flow rarely feature image resolutions over 1MP. In this presentation it is argued that for very densely sampled input, such as light fields or high frame rate video, existing strategies do not necessarily get the best out of the data, and that this type of data poses novel interesting research challenges beyond simply improving speed and scalability of existing techniques. As an example, a method for 3D scene reconstruction from "gigaray" light fields is proposed, which breaks with a number of established practices in image-based reconstruction in order to achieve improved reconstruction quality and speed.

3.22 Visualizing Point Sets

Ayellet Tal (Technion – Haifa, IL)

Joint work of Tal, Ayellet; Katz, Sagi; Basri, Ronen

If you try to draw a large point cloud, you my find that it is difficult to understand what the cloud represents. In this talk I will discuss the reduction of points, such that the visual comprehension of the set is improved. Two simple operators are utilized, the first computes the visible points from a given viewpoint and the second is its dual. Both operators work directly on the point cloud, skipping surface reconstruction.

References

1 Sagi Katz and Ayellet Tal and Ronen Basri. *Direct Visibility of Point Sets*. ACM Transactions on Graphics, Volume 26, Issue 3, 2007, 24/1-24/11
2 Sagi Katz and Ayellet Tal. *Improving the Visual Comprehension of Point Sets*. CVPR, 2013

3.23 Finite Element Based Tracking of Deforming Surfaces

Stefanie Wuhrer (Universität des Saarlandes, DE)

Joint work of Wuhrer, Stefanie; Lang, Jochen; Tekieh, Motahareh; Shu, Chang
Main reference S. Wuhrer, J. Lang, M. Tekieh, C. Shu, "Finite Element Based Tracking of Deforming Surfaces," arXiv:1306.4478v1 [cs.CV], 2013.
URL http://arxiv.org/abs/1306.4478v1

We present an approach to robustly track the geometry of an object that deforms over time from a set of input point clouds captured from a single viewpoint. The deformations we consider are caused by applying forces to known locations on the object's surface. Our method combines the use of prior information on the geometry of the object modeled by a smooth template and the use of a linear finite element method to predict the deformation. This allows the accurate reconstruction of both the observed and the unobserved sides of the object. We present tracking results for noisy low-quality point clouds acquired by a stereo camera and a Kinect sensor, and simulations with point clouds corrupted by different error terms, and show that our method is also applicable to large non-linear deformations.

Participants

- Philippe Bekaert
Hasselt Univ.. – Diepenbeek, BE
- Tamy Boubekeur
Télécom Paris Tech, FR
- Edmond Boyer
INRIA Rhône-Alpes, FR
- Gabriel Brostow
University College London, GB
- Darren Cosker
University of Bath, GB
- Carsten Dachsbacher
KIT – Karlsruhe Institute of Technology, DE
- Robert Dawes
BBC Research & Development – London, GB
- Jean-Michel Dischler
University of Strasbourg, FR
- Bernd Eberhardt
Hochschule der Medien – Stuttgart, DE
- Peter Eisert
Fraunhofer-Institut – Berlin, DE
- Paolo Favaro
Universität Bern, CH
- Dieter W. Fellner
TU Darmstadt, DE
- Jan-Michael Frahm
University of North Carolina – Chapel Hill, US
- Martin Fuchs
Universität Stuttgart, DE
- Bastian Goldlücke
Universität Heidelberg, DE
- Oliver Grau
Intel Visual Computing Inst. – Saarbrücken, DE
- Volker Helzle
Filmakademie Baden-Württemberg – Ludwigsburg, DE
- Anna Hilsmann
HU Berlin, DE
- Adrian Hilton
University of Surrey, GB
- Martin Jagersand
University of Alberta, CA
- Jan Kautz
University College London, GB
- Oliver Klehm
MPI für Informatik – Saarbrücken, DE
- Felix Klose
TU Braunschweig, DE
- Andreas Kolb
Universität Siegen, DE
- Hendrik P. A. Lensch
Universität Tübingen, DE
- Christian Lipski
metaio GmbH – München, DE
- Yebin Liu
Tsinghua Univ. – Beijing, CN
- Céline Loscos
Universite de Reims, FR
- Marcus A. Magnor
TU Braunschweig, DE
- Shohei Nobuhara
Kyoto University, JP
- Sylvain Paris
Adobe Systems Inc. – Cambridge, US
- Fabrizio Pece
University College London, GB
- Kari Pulli
NVIDIA Corporation – Santa Clara, US
- Bodo Rosenhahn
Leibniz Univ. Hannover, DE
- Holly E. Rushmeier
Yale University, US
- Hans-Peter Seidel
MPI für Informatik – Saarbrücken, DE
- Alla Sheffer
University of British Columbia – Vancouver, CA
- Philipp Slusallek
DFKI – Saarbrücken, DE
- Olga Sorkine-Hornung
ETH Zürich, CH
- Alexander Sorkine-Hornung
Disney Research – Zürich, CH
- Ayellet Tal
Technion – Haifa, IL
- Christian Theobalt
MPI für Informatik – Saarbrücken, DE
- Stefanie Wuhrer
Universität des Saarlandes, DE
- Ruigang Yang
University of Kentucky, US
- Remo Ziegler
LiberoVision AG – Zürich, CH