Report from Dagstuhl Seminar 12431

Time-of-Flight Imaging: Algorithms, Sensors and Applications

Edited by

James Davis[1], Bernd Jähne[2], Andreas Kolb[3], Ramesh Raskar[4], and Christian Theobalt[5]

1 **University of California – Santa Cruz, US,** `davis@cs.ucsc.edu`
2 **Universität Heidelberg, DE**
3 **Universität Siegen, DE,** `andreas.kolb@uni-siegen.de`
4 **MIT – Cambridge, US,** `raskar@media.mit.edu`
5 **MPI für Informatik – Saarbrücken, DE,** `theobalt@mpi-inf.mpg.de`

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 12431 "Time-of-Flight Imaging: Algorithms, Sensors and Applications". The seminar brought together researchers with diverse background from both academia and industry to discuss various aspects of Time-of-Flight imaging and general depth sensors. The executive summary and abstracts of the talks given during the seminar as well as the outcome of several working groups on specific research topics are presented in this report.

Seminar 21.–26. October, 2012 – www.dagstuhl.de/12431
1998 ACM Subject Classification I.2.10 Vision and Scene Understanding, I.4.1 Digitization and Image Capture, H.5.2 User Interfaces
Keywords and phrases Time-of-Flight, Kinect^{TM}, depth sensor
Digital Object Identifier 10.4230/DagRep.2.10.79
Edited in cooperation with Kwang In Kim

1 Executive Summary

James Davis
Bernd Jähne
Andreas Kolb
Ramesh Raskar
Christian Theobalt

In recent years, Time-of-Flight (ToF) depth imaging technology has seen immense progress. Time-of-Flight imaging is based on measuring the time that light, emitted by an illumination unit, requires to travel to an object and back to a detector. From this time, scene depth and possibly additional information that can not be measured by traditional intensity imaging, is inferred. While early ToF cameras were merely lab prototypes to prove a concept, recent sensor designs are at the edge of becoming operative products for mass market applications. A wide range of research disciplines is able to benefit from reliable and fast depth imaging technology, such as computer vision, computer graphics, medical engineering, robotics and computational photography, to name a few. Easy availability of affordable depth cameras will

Time-of-Flight Imaging: Algorithms, Sensors and Applications, *Dagstuhl Reports*, Vol. 2, Issue 10, pp. 79–104
Editors: James Davis, Bernd Jähne, Andreas Kolb, Ramesh Raskar, and Christian Theobalt
DAGSTUHL REPORTS Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

open the door for many new applications. The commercial success of the Microsoft™ Kinect device – a depth sensor based on an alternative measurement principle – gives a first impression on this.

Currently, manufacturers of ToF systems mainly focus on sensor technology and on the design of cameras. Sensor design has seen great advancements, but the data delivered by the cameras remain challenging and are affected by many types of systematic distortions and difficult scene dependencies. ToF data are thus hardly usable out-of-the-box and it takes proper mathematical modeling and algorithmic processing to apply the data in practical imaging and reconstruction scenarios. Algorithm design for ToF imagers, however, is still in its early days and many challenges remain. In this seminar, we plan to discuss and extend the state of the art in ToF imaging algorithms and applications with leading researchers in the field.

Also, currently, there is little dialogue between researchers developing ToF algorithms and sensor designers. Therefore, the seminar also strongly supported the manufacturers in getting up to date with all relevant research results and, even more importantly, it offered the opportunity to establish long-term partnerships and research collaborations. We also believe that this stronger interaction will lead to more advanced sensor designs, as well as more powerful algorithmic solutions at the same time.

Description of the Seminar: Topics, Goals and Achievements

General Motivation

Time-of-Flight technology is based on measuring the time that light, emitted by an illumination unit, requires to travel to an object and back to a detector. This allows to measure distances with high precision and high speed. Recently, this principle has been the basis for the development of new range-sensing devices realized in standard CMOS and CCD technology and which are called ToF cameras, as well as in the context of photogrammetry Range Imaging (RIM) sensors. Unlike other 3D systems, the ToF camera is a very compact device. It has the potential of being one of the first low-price, off-the-shelf system to provide full-range distance information in at video rate.

Today the community using Time-of-Flight technology is scattered over many research disciplines without intense communication across research areas. Such communication is necessary, however, to fuse results from sensor technology, low–level ToF data processing and high–level image processing. Each of the above research disciplines that employs time-of-flight imaging has to develop algorithmic solutions to these very same core problem areas. Additionally, there are new hot topics that currently do not make use of this new technology but might benefit from it in the future, which further underlines the importance of ToF algorithm design.

In this seminar, we exploited this multi-disciplinarity, and brought researchers from computer vision, computer graphics, computational photography, image processing and engineering disciplines together that work with ToF imagers. Together, we defined the current state of the art in core algorithmic questions that ToF imaging researchers are confronted with (additionally to the seminar by an edited book on the main results). We also contributed advancing the field by identifying current limitations, important future research directions, and by enabling a closer dialogue between algorithm and hardware designers to discuss future sensor designs.

Topics

Time-of-Flight imaging devices can measure scene depth largely independently of scene appearance and are generally based on extensions of standard video intensity camera hardware. ToF sensors can thus be used for static and dynamic scene capture. However, the data of these sensors suffer from a variety of deficiencies, such as low resolution, strong random noise, and non-trivial systematic distortions. These challenges have to be algorithmically addressed before ToF cameras become mainstream in any field of application. The main topic of this seminar was the definition and extension of the state of the art in ToF imaging problems in three core areas of algorithm and technology development that are described in the following.

Low level data processing, calibration and characterization

Researchers in computer vision, computer graphics and image processing only just started to mathematically model the measurement characteristics of ToF sensors [29]. This is a fundamental prerequisite for calibration [6], as well as for well-founded design of low level ToF data processing.

The phase-based Time-of-Flight technology suffers from some specific problems that cause systematic calibration errors and parameter correlation issues. Due to the physical realization of light modulation in the emitting LEDs, the ideal sine-waveform light emittance is approximated by a band-limited rectangular waveform. This causes nonlinear depth distortions, called *wiggling errors*. In addition, there are several non-linear effects depending on multi-path light propagation, for example in the optical system or due to multiple reflections in the scene. Some effects are well-understood, but there are still open issues in depth calibration [20]. In addition, the calibration of external camera parameters suffers from strong correlation, since typically the cameras have limited field of view and low image resolution. Solutions to this problem can be found if a synchronous multi-camera calibration with rigidly coupled color and range camera rigs are investigated [28]. Coupling of high-resolution color video cameras with ToF cameras is hence an issue of further investigation. Latest ideas on sensor calibration will be reviewed and augmented in the seminar.

The knowledge gained through sensor calibration can also be exploited to create sensor simulations of high fidelity in software. This will be an invaluable tool test new algorithms. Proper sensor modeling also enables detailed sensor comparison and evaluation, and eventually even certification. A couple of research initiatives are underway to build in-depth mathematical sensor model of ToF imagers which will be discussed at the seminar [12].

Low level sensor calibration and sensor modeling enables more efficient and effective design of algorithms for low-level TOF processing. For instance, first low-level ToF filtering [35, 2] and ToF 3D superresolution approaches have been proposed [30, 31]. Most of these approaches have already demonstrated that a proper sensor model can be exploited for higher quality processing. In the seminar, we reviewed latest low-level processing techniques, and evaluated how new and better filtering and data enhancement techniques can be developed, also for rarely considered depth camera artifacts, such as ToF motion blur. We also discussed how such techniques can be integrated on the sensor and how the gained understanding of sensor characteristics can benefit the design of future sensors.

High level data processing for 3D reconstruction, understanding and recognition

Low-level ToF imaging builds the foundations for the higher-level processing tasks that researchers and practitioners from many disciplines are confronted with. In most cases, such higher level processing aims to recover high-quality 3D models of static and/or dynamic scenes that should be displayed, analyzed, interactively modified, or used for recognition and scene understanding.

One major field of research using higher-level ToF image processing is computer graphics. Here, efficient acquisition of geometric models of static and dynamic scenes is of tremendous importance, and has many applications in interactive 3D rendering, geometric modeling and product design, 3D human computer interaction, cultural heritage, as well as professional media and game productions. ToF sensors can be an important asset here in order to replace the costly, highly specialized, complex and often intrusive acquisition technology currently used for such tasks. Static scene acquisition is mostly performed based on active scanners, using structured light or laser-based triangulation. Dynamic scene capture can also be achieved with structured light devices, and specialized optical systems that track fiducial markers exist for capturing motion. Being able to solve similar reconstruction tasks with only ToF cameras would be a big step ahead and eventually make 3D acquisition technology available to a wider range of users.

For a long time computer vision researchers have successfully developed 3D reconstruction approaches from single or multiple cameras that exploit certain photometric or radiometric cues. Many of them have in common that they are computationally expensive and that they only succeed under certain scene conditions, such as if scenes are sufficiently textured. An enormous potential lies in the fusion of ToF sensors with standard sensors for computer vision and robotics problems. Most areas in computer vision benefit from depth or range information; however, due to the difficulty in reconstruction of robust depth maps in real–world environments — especially in real–time applications — most state of the art solutions in areas like object recognition, gesture and action recognition in man-machine communication, pedestrian detection, and low–level tasks like segmentation just rely on 2D intensity information. Available depth and shape cues in real–time together with intensity information will open new possibilities to improve quality and robustness of algorithms and systems in such areas [25, 10, 13, 11, 24, 21]. In this context, there are several open problems, which were discussed during this seminar: do we need to define new features to be extracted from Time-of-Flight data and which feature will lead to a gain in quality compared to nowadays state of the art solutions? How can we deal with resolution and noise level of such cameras to complement normal 2D intensity information? Will we need to fuse Time-of-Flight information with 2D intensity data of standard CCD cameras, or are there applications, that can benefit from Time-of-Flight cameras by itself?

Another area in which ToF imaging will play a major role in future, is video processing, in particular 3D video and 3D TV. The analysis of dynamic 3D scenes from video requires the simultaneous processing of color video and range data. While traditional approaches using multi-view video are already quite successful, the advent of ToF range technology allows novel insights, novel applications and ease of acquisition. Traditional multi-view depth reconstruction requires sufficiently textured scenes, which might not be the case for arbitrary scenes, especially in indoor environments. This might lead to incomplete reconstruction results. ToF range acquisition has the potential for handling range data in dynamic video, but still many issues need to be solved and discussed by experts: in particular the challenging noise, uncertainty in the measurements, and low resolution of current ToF cameras represent a challenge. First applications handling video-rate HD-TV depth processing can be found in systems for 3D-Television capture [7] or in general computer graphics applications [17].

In many other areas, for instance computational photography, computational videography and medical engineering, researchers are facing similar reconstruction problems and can benefit from ToF sensors. For instance, in medical engineering, ToF cameras have been used to detect patient position [26] and respiratory motion in radiotherapy [27, 23].

The above list of examples shows that the algorithmic problems to be solved for making ToF sensors usable for high level reconstruction in different areas are very similar. The main challenge will be to enable high quality reconstruction despite strongly distorted and low-resolution raw ToF sensor output. Several strategies have been explored to attack these problems: Sensor fusion approaches combine depth and intensity cameras, spatio-temporal reconstruction approaches recover higher detail by accumulating and aligning measurements over time, superresolution and alignment can be combined to enable high-quality 3D reconstruction. Given such better quality reconstructions, the captured data can be employed as scene models ore further analyzed for capturing motion and gestures, for recognizing activities, for recognizing objects, or for analyzing the environment in a navigation scenario. The seminar therefore reviewed latest algorithms for static and spatio-temporal 3D reconstruction from ToF data. We have also discussed how they need to be tuned for specific applications, such as motion capture and recognition. Finally, we discussed ways to better integrate low-level and high-level processing.

Sensor technology and new depth sensor designs

While algorithm design for low-level and high-level TOF imaging were the main focus of this seminar, we also initiated to enable a dialogue between hardware manufacturers and algorithm designers. On the one hand this familiarized hardware designers with the state-of-the-art in ToF data processing, and sensibilized them for the existing challenges and specific application requirements. In return, algorithm designers deepened their knowledge about the fundamental physical principles of ToF imaging and gain a better understanding for the physical origins of sensor characteristics.

It is possible that relatively simple changes to the ToF hardware would result in the possibility of new sensor designs. ToF cameras make use of a CMOS sensor that is an enhanced version of a normal camera with extra circuitry at each pixel, and a structured IR illuminator. A great deal of prior research exists on using “normal” CMOS cameras together with triangulation based structured light to recover depth. The structured illuminator in these two research areas makes use of different principles, and the internal frame rate of the ToF camera is much higher, but the hardware components are broadly similar, suggesting that sharing of ideas might be fruitful.

Importantly, ToF and triangulation have complementary error characteristics, strengths, and weaknesses. For example, ToF sensors tend to perform better at a distance, and triangulation tends to perform better at close range. This leaves open the possibility of new sensor designs that make use of ideas from both ToF and structured light, with greatly improved robustness and accuracy. For example: chips could be designed with both “normal” and “ToF” pixels, the ToF light source could have a focusing lens and spatial pattern, the modulated light used with the ToF sensor could be similar to structured light patterns, the data from ToF could be used as a rough guess to disambiguate phase/depth in structured light when there are not enough patterns.

We are convinced that through a dialogue between hardware and algorithm designers, both sides can benefit. An example for a related research area in which such a dialogue has already resulted in great advancements is the area of computational photography. There, algorithm designers and hardware manufacturers have worked together on new designs for optical systems and processing algorithms that open up new ways of digital imaging, e.g., through high dynamic range imaging, wave front coding etc. We believe that the advent of ToF depth imaging technology is a further boost to this development, as it was already shown by new ideas on space-time imaging [16]. We also believe that ToF designs can have a similar impact in the emerging field of computational videography where future video sensors and processing paradigms are developed. We believe that the seminar served as a platform to initiate such developments by bringing together key players in the field. In this context, the pros and cons of alternative depth measurement sensors, such as IR-based active stereo cameras, have also been discussed.

Goals and Achievements of the Seminar

The overall goal of this seminar was to bring together researchers from several TOF-related disciplines, review the state-of-the-art in ToF imaging on both the algorithmic and hardware side, and develop new concepts for algorithms and devices that advance the field as a whole. The seminar was not intended to be a classical workshop or conference where mostly finished research is presented. We wanted the seminar to be a platform for identifying and discussing the big open research questions and challenges. More specifically, the following is a list of challenges that have been discussed at the seminar, since they form the basis of low-level and applied research with Time-of-Flight cameras:

- Low-level processing
  - Basic mathematical modeling of ToF cameras: image formation model, noise modeling, calibration of the sensor and optics.
  - Low-level image processing problems: resolution enhancement through superresolution and sensor fusion, data filtering, feature extraction under random and systematic distortions.
- High-level processing
  - Static shape scanning: high-quality geometry scanning, 3d superresolution, alignment approaches, probabilistic methods for reconstruction and alignment under noise.
  - Dynamic shape scanning: Spatio-temporal filtering, multi-sensor fusion approaches, model-based dynamic scene reconstruction, unsupervised dynamic scene reconstruction (joint model-building and motion reconstruction), marker-less motion and performance capture, 3d video.
- Improvements of sensor design: pixel design, light source design and arrangement, Time-of-Flight measurement principles: amplitude modulation vs. shutter. In this context we will also discuss standardization questions.

The seminar was very successful with respect to the set goals and initiated great interaction between researchers from different domains which had never happened in this way at other conferences or workshops.

In order to best initiate this interaction, we decided to organized a multi-faceted scientific programme. It consisted of a variety of different presentation formats. In particular, we had a series of research talks on the different research problems which we wanted to address in the seminar. When selecting the research talks, we planned for having a mix of presentations by junior and senior researchers, as well as balance of different topics. Presenters dedicated at least half the presentation time to address open research problems in order to spawn new research projects and collaborations. In order to further initiate discussion between researchers with different backgrounds, and in order to very practically identify potential research projects, we also organized working groups in which small teams discussed certain focus topics. Finally, the seminar participants organized very informal evening sessions in which special cross-disciplinary research topics were discussed in a very informal way. Finally, a demo session enables researchers and hardware specialists to showcase their latest results.

As an outcome of this, a very lively discussion and interaction was started between participants, and many concrete research projects were defined. Most fruitful discussions started on the topics of: 1) how to better exploit existing hardware and software systems; 2) the limitation of existing sensors and how to break them; 3) new combinations of existing (heterogeneous) sensors; 4) technical and economical limitations of hardwares.

To achieve sustainability beyond the seminar the organizers will edit a book summarizing the main methods, applications, and challenges in the context of ToF technology based on the presentations and discussions during the seminar. Such a book is currently missing in the community and the seminar itself shall also act as catalyzer for such a project. For more rapid dissemination of ideas and results, the organizers also created Wiki[1] which will be eventually relocated and maintained permanently.

---

[1] http://www.dagstuhl.de/wiki/index.php?title=12431

2 Table of Contents

Working Groups

3 Overview of Talks

Research talks addressed specific algorithmic problems in Time-of-Flight imaging. Each presenter dedicated a lot of his presentation time to talk about big open research challenges that the community should look into.

3.1 Benchmarking Time-of-Flight Data for Specific Application Demands

Michael Balda (Metrilus GmbH – Erlangen, DE)

Joint work of Balda, Michael; Schaller, Christian; Placht, Simon

Depth data acquired with Time-of-Flight (ToF) sensors suffers from many typical measurement artifacts such as motion artifacts, intensity related depth error, flying pixels, temperature drift, interference between multiple ToF cameras or effects caused by multi-path reflections and many more. Some of these issues can be addressed on chip or illumination level, others can be reduced with proper post-processing methods or, of course, hybrid approaches.

In this talk we outline the influence of these artifacts in practical medical and industrial ToF-applications such as robotics and gesture interaction.From these practical experiences we derive some of the requirements on countermeasures for specific scenarios. When developing new algorithms that process ToF-data and deal with ToF-artifacts it is of course necessary to quantify their performance in specific scenarios in a reproducible way. We suggest some selected benchmarks to evaluate the efficiency of existing countermeasures in standardized scenarios and give a short overview on how the currently available sensors perform in these benchmarks.

Furthermore, we would like to discuss ideas for benchmarking ToF in general and identify possible drawbacks and improvements for benchmarks and find necessary prerequisites to ensure a fair comparison of different sensors and algorithms. These prerequisites could for instance cover the influence of background light or reasonable camera resp. illumination warm-up periods.

3.2 Capturing and Visualizing Light in Motion

Christopher Barsi (MIT – Cambridge, US)

We show a technique to capture ultrafast movies of light in motion and synthesize physically valid visualizations. The effective exposure time for each frame is under two picoseconds. Capturing a 2D video with such a high time resolution is highly challenging, given the extremely low SNR associated with a picosecond exposure time, as well as the absence of 2D cameras that can provide such a shutter speed. We re-purpose modern imaging hardware to record an ensemble average of repeatable events that are synchronized to a streak tube, and we introduce reconstruction methods to visualize propagation of light pulses through macroscopic scenes. Capturing two-dimensional movies with picosecond resolution, we observe many interesting and complex light transport effects, including multi bounce, delayed mirror reflection, and sub-surface scattering. We notice that the time instances recorded by the camera, i.e. "Camera time" is different from the time of the events as they happen locally at the scene location, i.e. "world time". We introduce a notion of time warp between the two space-time coordinate system, and rewarp the space-time movie for a different perspective. This technique offers support for image-based rendering of relativistic events.

3.3 Frequency Analysis of Transient Light Transport with Applications in Bare Sensor Imaging

Christopher Barsi (MIT – Cambridge, US)

Light transport has been extensively analyzed in both the spatial and the frequency domain; the latter allows for intuitive interpretations of effects introduced by propagation through free space and optical elements, as well as for optimal designs of computational cameras capturing specific visual information. We relax the common assumption that the speed of light is infinite and analyze free space propagation in the frequency domain considering spatial, temporal, and angular light variation. Using this analysis, we derive analytic expressions for cross-dimensional information transfer and show how this can be exploited for designing a new, time-resolved bare sensor imaging system.

3.4 Gesture-based interaction with ToF cameras

Erhardt Barth (Universität Lübeck, DE)

I will base my talk on the experience with our startup company gestigon (www.gestigon.de). But do not worry, this will not be any kind of commercial. Rather, the startup endeavor tells you in a clear way, which can hurt, what the current limitations of the technology are and what future developments are required. If you need to track all degrees of freedom of two hands with very little hardware, if you need to do that through windows, in bright sunlight, and in cars, you are faced with a number of challenges, for example to (i) invent simple and robust algorithms, (ii) talk to the users and define use cases, (iii) talk to the camera people and jointly optimize hardware and algorithm design. I will briefly sketch our approach to hand- and body-skeleton tracking and discuss the main challenges. In addition to those mentioned above, one of the main technical challenges is to obtain a tight coupling between the user and the application. On top of that you need to understand user intent and a basic alphabet of gestures. It seems worth discussing how such an alphabet could be defined. It would be nice if we could set up an interdisciplinary, international interest group for TOF-based gesture technology.

3.5 Mitigating common distortion sources, and exploring alternative applications, for Time-of-Flight cameras

Adrian Dorrington (University of Waikato, NZ)

The Chronoptics research group has more then ten years experience in the field of Time-of-Flight (ToF) imaging. We have developed several technologies to improve the quality of ToF cameras (see more detail at http://www.chronoptics.com). Some of these technologies are well developed, but others suffer from practical limitations. This talk will introduce the following techniques, and call for collaborators to help address their limitations.

We have developed a Mixed Pixel Separation technique that can resolve multiple returns detected by a single pixel. This is useful for correcting edge effects such as mixed, or so-called "flying", pixel distortion, and for rejecting multi-path distortion. Published and unpublished results demonstrate the efficacy of this algorithm when the phase difference between the multiple returns is large and the signal-to-noise ratio is high, but the algorithm fails with small phase differences or when the measured depth precision is poor.

The combination of "Fluttered shutter" and optical flow techniques have allowed us to detect and quantify motion independently for multiple objects and to correct local motion blur. Although directly quantifying individual object velocity and direction is of interest in many fields, this technique requires further work to improve computational complexity and to automatically enumerate moving objects.

ToF sensors have potential for applications other than distance measurement. For example, the process of Diffuse Optical Tomography (DOT) data acquisition for internal medical imaging is very similar to ToF distance measurement, only the illumination and detection is in contact with the subject's skin and the data is processed in a very different way. We have demonstrated proof-of-principle showing that ToF cameras can be used for non-contact DOT using the NIRFast software package. Due to the multidisciplinary nature of this project, it would benefit significantly from a collaborative approach.

3.6 Difficulties and novel applications in a low-cost multi-view depth camera setting

Martin Eisemann (TU Braunschweig, DE)

Joint work of Eisemann, Martin; Berger, Kai; Ruhl, Kai; Guthe Stefan; Klose, Felix; Lipski, Christian; Hell, Benjamin; Magnor, Marcus
URL http://www.cg.cs.tu-bs.de/publications/

High-quality ToF cameras are still expensive nowadays, ranging from only a few to several thousand dollars. While less expensive solutions exist, these come at the cost of a higher signal to noise ratio and lower resolution. In this talk we will take a look at the difficulties, applications and our solutions for such low-cost depth cameras in a multi-view setting.

In this setting we treat the problem of calibration using mirrors, gas reconstruction using depth-images, super-resolution for IR sensors and integrating approximate depth data into dense image correspondence estimation.

Hopefully, these examples give rise to some fruitful discussion on new application fields for depth cameras besides the typical 3D scene reconstruction in the later part of the talk.

3.7 Will Depth Cameras Have a Long-term Impact on Computer Vision Research?

Juergen Gall (MPI für Intelligente Systeme – Tübingen, DE)

Depth cameras have become a commercial success and their popularity in the research community increased with the drop of sensor prices. Since many approaches focus on applying techniques that are well-known from 2D image/video analysis or stereo vision, it is time to discuss if depth sensors will open new research directions in computer vision that will have a long-term impact. To this end, I will review recent publications that appeared at computer vision workshops or conferences and made use of depth cameras for high-level computer vision tasks. Finally, I would like to start a discussion of the future of depth cameras in high-level computer vision research.

3.8 Capturing and Visualizing Light in Motion

Diego Gutierrez (University of Zaragoza, ES)

We show a technique to capture ultrafast movies of light in motion and to synthesize physically valid visualizations. The effective exposure time for each frame is under two picoseconds (ps). Capturing a 2D video with this time resolution is highly challenging, given the low signal-to-noise ratio (SNR) associated with a picosecond exposure time, as well as the absence of 2D cameras that can provide such a shutter speed. We re-purpose modern imaging hardware to record an ensemble average of repeatable events that are synchronized to a streak tube, and we introduce reconstruction methods to visualize both propagation of light pulses through macroscopic scenes, as well as relativistic effects of moving bodies.

Capturing two-dimensional movies with picosecond resolution, we observe many interesting and complex light transport effects, including multibounce scattering, delayed mirror reflections, and subsurface scattering. We notice that the time instances recorded by the camera, i.e., “camera times” are different from the time of the events as they happen locally at the scene location, i.e., “world times”. We introduce the notion of time unwarping between the two space-time coordinate systems.

3.9 Open questions in full-body motion estimation with depth cameras

Thomas Helten (MPI für Informatik – Saarbrücken, DE)

The analysis and recording of full-body human motion data is an important strand of research with applications to movie and game productions, sport sciences, and human computer interaction.In the recent years, the availability of cheap range sensors, such as the Microsoft Kinect has boosted the research on tracking human motions from monocular depth images. Despite the promising approaches in this field there are still unsolved challenges such as (self) occlusions or ambiguities. In this talk, I want to elaborate on the reasons for these challenges and ideas to approach them.

3.10 Can we reconstruct the shape of a mirror-room from multi-bounce ToF measurements?

Ivo Ihrke (Universität des Saarlandes, DE)

I will discuss our recent progress in answering this question. We assume that we are given the measurements of a tempo-angularly resolved receiver recording the response of the room to a spherical pulse source. The receiver and the source are inside the room but at separate locations. I will show initial positive results for convex polyhedral rooms in two dimensions. Our method can deal with limited angular and/or temporal data. Also, it is not necessary to know the bounce order of a received pulse. It is however, difficult to establish performance bounds and reconstruction guarantees.

3.11 Deformable Object Detection in Underwater ToF Videos

Slobodan Ilic (TU München, DE)

The aquaculture industry has been continuously thriving since the 1980s. As the fish farming grows, it becomes important to develop a remote monitoring system to estimate the biomass of a large number of fishes bred in cages. Since around 80% of all sales of farmed fish are arranged pre-harvest, the profit on the sale directly depends on correct estimations of weight, size distribution and totalbiomass. Therefore, the goal of this research is to build an automated and relatively affordable tools for biomass estimation.

Here we will rely on ToF camera images acquired underwater, that are supposed to film fishes in the cage for certain period of time. In order to estimate the biomass the volume of the fish has to be estimated. This can be achieved by first detecting the fishes in every range image of the incoming video stream and then fitting a 3D model to these detections. To find the algorithm that is in line with our problem, we need to understand the challenges in detecting fishes. They include the motion of the fish which makes the object of interest deformable, the location of the fish with respect to the camera and occlusions caused by having multiple fishes in every available frame.

In this talk I will present our approach to this problem mainly addressing high-level processing task summarized in developed algorithm for deformable object detection. In addition I would like to briefly introduce the technical challenges related to data acquisition using ToF camera underwater and get opinions of ToF experts about constructing reliable acquisition device underwater.

3.12 Efficient Deformation Reconstruction from Depth and Color Images using Analysis by Synthesis

Andreas Jordt (Universität Kiel, DE)

The reconstruction of deformations has always been a difficult problem due to the lack of a generic deformation model. Hence, 3d reconstruction of deforming objects ever since can be separated into two classes: Those using an explicit deformation model, e.g. skeletal models, – and those heavily relying on feature movement or optical flow.

Depth sensors like ToF cameras and or the Kinect depth sensor provide valuable scene information but do not provide a stable base for optical flow or feature movement calculation. Approaches associating these depth values with optical flow or feature movement from color images try to circumvent this problem but suffer from the fact that color features are often generated at edges and depth discontinuities, areas in which depth sensors deliver inherently unstable data. This talk introduces how the full potential of depth and color can be tapped by direct methods such as analysis by synthesis, utilizing the complete image data directly to calculate the result. A set of generic and specialized deformation models are introduced as well as an efficient way to synthesize and to optimize high dimensional models. The resulting reconstruction algorithms range from real-time deformation reconstruction methods to very accurate deformation retrieval using models of 100 dimensions and more.

3.13 Efficient Gaussian Process Regression-based Image Enhancement

Kwang In Kim (MPI für Informatik – Saarbrücken, DE)

Many computer vision and computational photography applications essentially solve an image enhancement problem. The image has been deteriorated by a specific noise process that we would like to remove, such as aberrations from camera optics and compression artifacts. In this talk, we discuss an algorithm for learning-based image enhancement. At the core of the algorithm lies a generic regularization framework that comprises a prior on natural images, as well as an application-specific conditional model based on Gaussian processes (GPs). To overcome the high computational complexity of GPs, an efficient approximation scheme of large-scale GPs are presented. This facilitates instantly learning task-specific degradation models from sample images. The efficiency and effectiveness of our approach is demonstrated by applying it to an example enhancement application: single-image super-resolution.

3.14 Real Time Handling of Depth Data

Andreas Kolb (Universität Siegen, DE)

The availability of depth data at interactive frame rates poses the challenge of handling a large amount of 3D data at nearly the same speed in order to realize interactive applications.

This talk presents recent results in handling large sets of streamed range data, taking into account the questions of how to reduce the amount of data and how to efficiently store range data. Here, also the quest of handling scene dynamics and of varying range data quality plays a role.

3.15 Automated classification of therapeutical face exercises using the Kinect

Cornelia Lanz (TU Ilmenau, DE)

The presentation is going to propose an approach for the topic of therapeutical facial exercise recognition using depth images recorded with the Kinect. In cooperation with speech-language therapists, we determined nine exercises that are beneficial for the therapy of patients suffering from dysfunction of facial movements. Extracted depth features comprise the curvature of the face surface and characteristic profiles that are derived using distinctive landmark points. The presentation will focus on the evaluation of the features. This comprises:

- their discriminative power concerning the classification of nine therapeutical exercises.
- their suitability for a fully automated real-world scenario. This requires features that are robust with respect to slightly varying feature extraction regions.

3.16 Enhancing ToF measurements: current work, evaluation with ground truth and open problems

Frank Lenzen (Universität Heidelberg, DE)

In 2010, the research project 'Algorithms for low cost depth imaging' started at the Heidelberg Collaboratory for Image Processing(HCI), co-financed by the Intel Visual Computing Institute (IVCI) in Saarbrücken. We report on the ongoing work within this project. In particular we consider the topic of denoising ToF data. In order to evaluate the quality of our approaches, we use a ToF data set which was created within this project and is annotated with ground truth. This dataset will be made publicly available.

Moreover, we discuss open problems we encountered during our research and which are of interest for the community.

3.17 Patch Based Synthesis for Single Depth Image Super-Resolution

Oisin Mac Aodha (University College London, GB)

Joint work of Mac Aodha, Oisin; Campbell, Neill; Nair, Arun; Brostow, Gabriel J.
Main reference O. Mac Aodha, N. Campbell, A. Nair, G.J. Brostow, "Patch Based Synthesis for Single Depth Image Super-Resolution," ECCV (3), 2012, 71–84.
URL http://dx.doi.org/10.1007/978-3-642-33712-3_6

We present an algorithm to synthetically increase the resolution of a solitary depth image using only a generic database of local patches. Modern range sensors measure depths with non-Gaussian noise and at lower starting resolutions than typical visible-light cameras. While patch based approaches for upsampling intensity images continue to improve, this is the first exploration of patching for depth images.

We match against the field of each low resolution input depth patch, and search our database for a list of appropriate high resolution candidate patches. Selecting the right candidate at each location in the depth image is then posed as a Markov random field labeling problem. Our experiments also show how important further depth-specific processing, such as noise removal and correct patch normalization, dramatically improves our results. Perhaps surprisingly, even better results are achieved on a variety of real test scenes by providing our algorithm with only synthetic training depth data.

3.18 Can ToF Cameras Enable Dynamic Interactive Ubiquitous Displays?

Aditi Majumder (University of California – Irvine, US)

Large dynamically changing surface geometry lighted seamlessly at a very high resolution by multiple projectors allowing interaction from multiple users makes displays truly ubiquitous–available to anyone anywhere. It is a dream harbored by many communities including computer graphics and vision; human computer interaction; and virtual, mixed and augmented reality. These can have tremendous applications in education, entertainment, simulation and training.

Several inroads have been made in the past in this direction, including a big body of work from our lab at UCI. Multiple aspects of this problem have been studied including fast and accurate surface reconstruction from multiple sensors (potentially uncalibrated); cross-validation across multiple devices to achieve robust calibration; fast and accurate 3D gesture recognition of multiple users from multiple sensors; centralized and distributed paradigms to achieve modularity in system design and improvement in efficiency, performance and ease in deployment; and efficient data management to handle large data sets.

However, SAR on very large and dynamically changing surfaces which people can interact with is still limited. Capturing depth from multiple cameras is still not fast enough and limits dynamism and interactivity. The inaccuracies in the estimated depth, especially in the presence of textures, limit the seamless registration of projected imagery on the surface. Both of these can be significantly alleviated by ToF Cameras. However, using multiple time of flight (ToF) cameras in the same system bring in a different set of challenges in terms of interference with one another and background noise.

In this talk I will first present the large amount of work done in our lab at UCI in making large dynamic and multi-user interactive display systems a possibility. Then I will briefly discuss the motivation and challenges in using ToF cameras in this setting. It seems that many of these challenges overlap with those faced by other communities and I hope to make important connections so that we can work together and reuse findings in multiple domains to make such ubiquitous display systems a reality of the future.

3.19 TOF Ground Truth Generation

Rahul Nair (Universität Heidelberg, DE)

With the eve of low cost depth imaging techniques such as Kinect and TOF the generation of ground truth for vision applications should not be reserved to those who can afford expensive equipment. For large parts of computer vision ground truth generation means the measurement of geometric (and radiometric) properties of the scene. With these new possibilities to provide cheap ground truth with lower accuracy it is crucial to start specifying GT accuracy. Otherwise we cannot benchmark methods against such sequences. Although there are physically correct TOF noise models in literature, many computer vision researchers prefer Kinect rather than TOF cameras because of its systematic errors such as multipath effects caused by interreflections. If we could overcome these errors TOF imaging has the same capability of reaching the mass markets and being used by vision researchers worldwide. We will describe how starting from a precise sensor characterization and physical model we try to tackle both these tasks. By combining this noise model with state of art techniques from computer graphics we are able to simulate the raw image acquisition process in the camera. This enables us to simulate systematic errors such as multipath for further studies. We also show how we produce ground truth data using TOF combined with other modalities with a confidence in each pixel such that this "weak" ground truth may still be used to test other vision algorithms.

3.20 Real-world 3D video production with ToF cameras

Shohei Nobuhara (Kyoto University, JP)

Main reference T. Matsuyama, S. Nobuhara, T. Takai and T. Tung, "3D Video and Its Applications," Springer-Verlag, 2012.
URL http://dx.doi.org/10.1007/978-1-4471-4120-4

This talk first introduces essential design factors on building 3D video production systems with multiple-cameras for lab environment. It covers camera arrangement, background design, illumination, etc for robust silhouette extraction, 3D shape reconstruction, and texture generation. Then it presents some half-baked ideas on how to utilize ToF cameras to improve the production pipeline. It covers simultaneous multi-view silhouette extraction in real environment, and high-fidelity rendering with view-dependent 3D shape optimization.

3.21 Time-of-Flight cameras for computer-assisted interventions: opportunities and challenges

Alexander Seitel (DKFZ – Heidelberg, DE)

Minimally-invasive procedures are increasingly gaining in importance for cancer diagnosis and treatment. To date, computer-based assistance systems mainly rely on external or internal markers and optical or electromagnetic tracking systems for assessment of patient position and movement. Often, these approaches are combined with intra-operative imaging that exposes the patient to additional radiation. With the Time-of-Flight (ToF) camera technique, a markerless and radiation-free approach for transferring planning data acquired before the intervention to the situation at the patient during the intervention is possible.

This talk will present challenges and opportunities of the application of ToF cameras for computer-assisted medical interventions. It will firstly summarize methods for calibration of ToF cameras for use in the medical environment and registration of pre- and intra-operatively acquired surfaces and point out difficulties in correctly evaluating those algorithms due to the lack of adequate ground truth. Furthermore, the feasibility of applying a ToF camera for intra-operative imaging is shown and exemplary applied in an application for navigated percutaneous needle insertions. Lastly, there will be an overview on potential future medical applications, e.g. in the field of ToF-endoscopy, with focus on requirements and potential difficulties for the use of ToF cameras for ToF to be used in the medical environment.

3.22 SoftKinetic DepthSensing and 3D gesture recognition technologies

Julien Thollot (SoftKinetic-Brussels, BE)

What Do SoftKinetic Do?

SoftKinetict's vision is to give everyone the freedom to control, explore and enjoy the entire digital world through the most natural and intuitive user interfaces and machine interactions. SoftKinetic is the leading provider of gesture-based platforms for the consumer electronics and professional markets. The company offers a complete family of 3D imaging and gesture recognition solutions, including patented 3D CMOS time-of-flight sensors and cameras(DepthSense®, formerly Optrima), multi-platform and multi-camera 3D gesture recognition middleware and tools (iisu®as well as games and applications from SoftKinetic Studios.

With over 10 years of R&D on both hardware and software, SoftKinetic solutions have already been successfully used in the field of interactive digital entertainment, consumer electronics, health care and other professional markets such as digital signage and medical systems.

For more information on any of our products and services, please contact us at: sales@softkinetic.com

- SoftKinetic Solutions link:
  http://www.softkinetic.com/Solutions/iisuSDK.aspx

- SoftKinetic Depthsensor chip link:
  http://www.softkinetic.com/en-us/solutions/depthsensesensors.aspx
- SoftKinetic DepthSense ToF Cameras link:
  RGBZ ToF cameras (image registration included)
  DS311 : 160*120 Z, 720p RGB sensor + audio, Close interaction (0.1-1m)
  DS325 : 320*240 Z, 720p RGB sensor + audio, Close&far (0.1-1&0.5-5m)
  http://www.softkinetic.com/en-us/solutions/depthsensecameras.aspx
- SoftKinetic iisu Middleware link (gesture recognition and user or hand feature)
  http://www.softkinetic.com/en-us/solutions/iisusdk.aspx
- Video and demo:
  - iisu middleware + DS311 close interaction mode + arduino:
    http://www.youtube.com/user/Softkinetic
  - iisu 3.5 (full body skeleton tracking + close range hand interactions):
    http://www.youtube.com/watch?v=5LvhdFudp50&list=UUS7kIRSSm_cXBvnszuegUoA&index=2&feature=plcp
  - Corporate video:
    http://www.youtube.com/watch?v=Xfz_uRoJGjE&feature=relmfu
- Perceptual computing SDK solutions by intel (SoftKinetic embedded) including DS325 ToF camera custom build (Creative branded) at 149$:
  http://software.intel.com/en-us/vcsource/tools/perceptual-computing-sdk

3.23 Frequency Analysis of Transient Light Transport with Applications in Bare Sensor Imaging

Gordon Wetzstein (MIT – Cambridge, US)

Light transport has been extensively analyzed in both the spatial and the frequency domain; the latter allows for intuitive interpretations of effects introduced by propagation in free space and optical elements as well as for optimal designs of computational cameras capturing specific visual information. We relax the common assumption that the speed of light is infinite and analyze free space propagation in the frequency domain considering spatial, temporal, and angular light variation. Using this analysis, we derive analytic expressions for cross-dimensional information transfer and show how this can be exploited for designing a new, time-resolved bare sensor imaging system.

3.24 3D Modeling and Motion Analysis from a Single Depth Camera

Ruigang Yang (University of Kentucky, US)

3D has become a hot topic recently, partly due to two recent technical innovations: 3D TVs and commodity depth cameras–the Kinect camera from Microsoft. In this talk, I will first present an approach to create a complete dynamic 4D (space + time) model from a RGB+depth video sequence. Unlike traditional Structure from motion or point cloud merging algorithm, our approach can deal with deformable subjects. Then I will talk about an approach that estimates skeleton motion using a single depth camera. Trading speed for accuracy, our approach reduces the average motion estimation error from 50 mm to be less than 10mm. Finally I will present a sensor-fusion approach that combines photometric stereo with active stereo (e.g., Kinect) to significantly improve the quality of the depth map. Unlike previous fusion approaches, we model depth discontinuity and occlusion explicitly.

4 Working Groups

4.1 Non-standard usage of ToF hardware – Brainstorming

James Davis (University of California– Santa Cruz, US)

At this workshop we are studying Time-Of-Flight Cameras, which implies the relatively narrow goal of obtaining depth using a measurement of time. However the hardware itself is a device which obtains multiple photonic measurements in rapid succession, and perhaps the device could be used for a purpose other than obtaining depth, or perhaps depth could be obtained through some principle other than strict time estimates. This alternative sessions goal is to brainstorm in these areas:

- alternative applications for the basic hardware
- modifications or flexibility that software researchers would like to see in the hardware
- modifications that hardware researchers would be happy to make, but don't yet know a suitable application

- 15 min – Intro of brainstorming session and example ideas:
  - Use ToF hardware with 4 time slots for triangulation structured light instead of depth phase estimation
  - Currently using only phase, joint coding with structured light source to use amplitude also?
  - "Sophisticated" image processing on the raw data, rather than the "simple" already computed depth
  - Sweep laser and get intersect time as intensity, using modified subpixel activation timing
  - I could build this thing with 8-tap super-pixels, does anyone want that?
- 30 min – Groups of 5 people brainstorm at least one non-standard purpose/method to abuse the hardware.
- 10 min – Sketch some slides/diagrams on overhead transparencies for the best brainstorm idea(s) in each group.
- 45 min – Groups report on their best ideas (About 8 groups with 5 minutes each).
- 5 min – Session wrap up and plan for continuity (Wiki to have areas for ideas and contact info for people potentially interested)

4.2 Time of Flight Cameras vs. Kinect

Seungkyu Lee (SAIT, South Korea)

Recently, many researchers have used either ToF sensor or Kinect for various applications. Many of them ask which type of depth sensor is best for their own applications. Actually, this is one of the most frequent and critical question for beginners of these sensors but the answer is not simple. In principle, both ToF and Kinect sensors have respective pros and cons. In this session, inputs from diverse experienced researchers and practitioners can be collected and shared. They can share what was the problem in using ToF (or Kinect) sensor in their applications and how about if they replace the sensor by Kinect (or ToF) sensor. We may not conclude which one is better than the other, but we can get better understanding and comparison on both sensing principles for future use.

- 15min Opening– start with a short intro./instruction for this topic including basic knowledge on both principles.
- 45min Brainstorming– divide the group into 3 4 along with their experiences or interests, such as interaction, imaging or 3D reconstruction, low level processing etc. ToF users and Kinect users (if available) can be mingled in each group. Let them share their experiences and thoughts on the sensors; what is good, what is bad, what is main obstacle for their app., and why is that. And they can discuss; 1) which characteristic of the sensor has to be improved and 2) what if they replace their ToF sensor by Kinect or vice versa. Each group makes a list/slides of pros-cons of ToF or Kinect for specific application.
- 40min Discussion – Each group report their lists and issues raised and discuss. If someone can advise on a point of the list, that can be added at each list and refine them.
- 5 min – Session wrap-up.

4.3 Real-world 3D video production with ToF cameras

Shohei Nobuhara (Kyoto University, Japan)

Conventional 3D video (or 4D modeling of dynamic 3D surface) production systems utilize multi-view 2D cameras, and reconstruct 3D surfaces based on shape-from-silhouettes (SFS) and wide-baseline stereo (WBS). SFS stably provides a rough but full 3D geometry (visual hull), and WBS refines it where WBS can be confident by prominent textures. This strategy is known to work well (sub-centimeter resolution of 3D human in motion) for controlled environments such as green/blue-backgrounds, but it is not directly applicable for real-world / outdoor scenes because its stability depends on the accuracy of the multi-view silhouettes which is not easily available for such environments. The motivation of this alternative session is to discuss about how ToF cameras can help 3D video production be robust in real-world.

Plan:

- 15 min– introduction of this session and example of ideas.
  - Accurate 2D multi-view silhouette acquisition with ToF cameras,
  - Direct full 3D reconstruction mainly by color cameras but with help of ToF cameras,

  - Direct full 3D reconstruction by ToF cameras.
- 30 min– Groups of 5 people brainstorm at least one non-standard purpose/method to abuse the hardware.
- 10 min– Sketch some slides/diagrams on overhead transparencies for the best brainstorm idea(s) in each group.
- 45 min– Groups report on their best ideas (About 8 groups with 5 minutes each)
- 5 min– Session wrap up and plan for continuity (Wiki to have areas for ideas and contact info for people potentially interested)

References

1 Naveed Ahmed, Christian Theobalt, Christian Roessl, Sebastian Thrun, and Hans-Peter Seidel. Dense correspondence finding for parametrization-free animation reconstruction from video. In *Proc. of IEEE Intl. Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008.

2 Derek Chan, Hylke Buisman, Christian Theobalt, and Sebastian Thrun. A noise-aware filter for real-time depth upsampling. In *Proc. of ECCV Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications*, 2008.

3 Ryan Crabb, Colin Tracey, Akshaya Puranik, and James Davis. Real-time foreground segmentation via range and color imaging. *Computer Vision and Pattern Recognition Workshop*, 0:1–5, 2008.

4 Yan Cui, Sebastian Schuon, Derek Chan, Sebastian Thrun, and Christian Theobalt. 3d shape scanning with a time-of-flight camera. In *Proc. of IEEE Intl. Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2010.

5 James Davis and Hector Gonzalez-banos. Enhanced shape recovery with shuttered pulses of light. In *IEEE Workshop on Projector-Camera Systems*, 2003.

6 M. Erz and B. Jähne. Radiometric and Spectrometric Calibrations, and Distance Noise Measurement of TOF Cameras. In R. Koch and A. Kolb, editors, *3rd Workshop on Dynamic 3-D Imaging*, volume 5742 of *Lecture Notes in Computer Science*, pages 28–41. Springer, 2009.

7 A Frick, B Bartczack, and R Koch. 3d-tv ldv content generation with a hybrid tof-multicamera rig. *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2010*, pages 1 – 4, 2010.

8 Hector Gonzalez-Banos and James Davis. Computing depth under ambient illumination using multi-shuttered light. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:234–241, 2004.

9 Doaa Hegazy and Joachim Denzler. Combining appearance and range based information for multi-class generic object recognition. In *CIARP*, pages 741–748, 2009.

10 Doaa Hegazy and Joachim Denzler. Generic 3d object recognition from time-of-flight images using boosted combined shape features. In *Proceedings of International Conference on Computer Vision, Theory and Applications (VISAPP 09)*, 2009.

11 Olaf Kähler, Erik Rodner, and Joachim Denzler. On fusion of range and intensity information using graph-cut for planar patch segmentation. *International Journal of Intelligent Systems Technologies and Applications*, 5(3/4):365–373, 2008.

12 M. Keller and A. Kolb. Real-time simulation of time-of-flight sensors. *J. Simulation Practice and Theory*, 17:967–978, 2009.

13 Michael Kemmler, Erik Rodner, and Joachim Denzler. Global context extraction for object recognition using a combination of range and visual features. In R. Koch and A. Kolb, editors, *Proceedings of the Dynamic 3D Imaging Workshop*, volume 5742 of *Lecture Notes in Computer Science*, pages 96–109. Springer, 2009.

14 Young Min Kim, Derek Chan, Christian Theobalt, and Sebastian Thrun. Design and calibration of a multi-view tof sensor fusion system. In *IEEE CVPR Workshop on Time-of-flight Computer Vision*, 2008.

15 Young-Min Kim, Christian Theobalt, James Diebel, Jana Kosecka, Branislav Micusik, and Sebastian Thrun. Multi-view image and tof sensor fusion for dense 3d reconstruction. In *Proc. of 3DIM*, 2009.

16 Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using transient imaging. In *ICCV*, pages 159–166, 2009.

17 A. Kolb, E. Barth, R. Koch, and R. Larsen. Time-of-flight cameras in computer graphics. *COMPUTER GRAPHICS forum, vol. 29, no. 1*, pages 141 – 159, 2010.

18 M. Lindner and A. Kolb. Lateral and depth calibration of pmd-distance sensors. In *Proc. Int. Symp. on Visual Computing*, LNCS, pages 524–533. Springer, 2006.

19 M. Lindner, A. Kolb, and K. Hartmann. Data-fusion of PMD-based distance-information and high-resolution RGB-images. In *Int. Sym. on Signals Circuits & Systems (ISSCS), session on Algorithms for 3D TOF-cameras*, pages 121–124. IEEE, 2007.

20 Marvin Lindner, Ingo Schiller, Andreas Kolb, and Reinhard Koch. Time-of-flight sensor calibration for accurate range sensing. *Computer Vision and Image Understanding*, 114(12):1318 – 1328, 2010. Special issue on Time-of-Flight Camera Based Computer Vision.

21 Christoph Munkelt, Michael Trummer, Peter Kuehmstedt, Gunther Notni, and Joachim Denzler. View planning for 3d reconstruction using time-of-flight camera data. In J. Denzler, G. Notni, and H. Suesse, editors, *DAGM 2009*, volume 5748 of *352–361*. Springer, 2009.

22 Jochen Penne, Kurt Höller, Michael Stürmer, Thomas Schrauder, Armin Schneider, Rainer Engelbrecht, Hubert Feußner, Bernhard Schmauss, and Joachim Hornegger. Time-of-Flight 3-D Endoscopy. In G.-Z. Yang et al., editor, *MICCAI 2009, Part I, LNCS 5761*, volume 5761, pages 467–474, Berlin / Heidelberg, 2009.

23 Jochen Penne, Christian Schaller, Joachim Hornegger, and Thorsten Kuwert. Robust Real-Time 3D Respiratory Motion Detection Using Time-of-Flight Cameras. *Computer Assisted Radiology and Surgery 2008*, 3(5):427–431, 2008.

24 Martin Rapus, Stefan Munder, Gregory Baratoff, and Joachim Denzler. Pedestrian recognition using combined low-resolution depth and intensity images. In Martin Rapus, editor, *IEEE Intelligent Vehicles Symposium*, pages 632–636, Eindhoven University of Technology Eindhoven, The Netherlands, June 2008.

25 Erik Rodner, Doaa Hegazy, and Joachim Denzler. Multiple kernel gaussian process classification for generic 3d object recognition from time-of-flight images. In *Proceedings of the International Conference on Image and Vision Computing*, 2010.

26 Christian Schaller, Andre Adelt, Jochen Penne, and Joachim Hornegger. Time-of-flight sensor for patient positioning. In Ehsan Samei and Jiang Hsieh, editors, *Proceedings of SPIE*, volume 7258, 2009.

27 Christian Schaller, Jochen Penne, and Joachim Hornegger. Time-of-Flight Sensor for Respiratory Motion Gating. *Medical Physics*, 35(7):3090–3093, 2008.

28 Ingo Schiller, Christian Beder, and Reinhard Koch. Calibration of a pmd camera using a planar calibration object together with a multi-camera setup. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVII. Part B3a, pages 297–302, Beijing, China, 2008. XXI. ISPRS Congress.

29 M. Schmidt and B. Jähne. A physical model of time-of-flight 3d imaging systems, including suppression of ambient light. In R. Koch and A. Kolb, editors, *3rd Workshop on Dynamic 3-D Imaging*, volume 5742 of *Lecture Notes in Computer Science*, pages 1–15. Springer, 2009.

30 Sebastian Schuon, Christian Theobalt, James Davis, and Sebastian Thrun. High-quality scanning using time-of-flight depth superresolution. In *IEEE CVPR Workshop on Time-of-flight Computer Vision*, 2008.

31 Sebastian Schuon, Christian Theobalt, James Davis, and Sebastian Thrun. Lidarboost: Depth superresolution for tof 3d shape scanning. In *In Proc. of IEEE Intl. Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2009.

32 A. Seitel, T. Santos, S. Mersmann, J. Penne, R. Tetzlaff, and H.-P. Meinzer. Time-of-flight kameras für die intraoperative oberflächenerfassung. In *Proc BVM*, pages 11–15, 2010.

33 Savil Srivastava, Ashutosh Saxena, Christian Theobalt, Sebastian Thrun, and Andrew Ng. i23 - rapid interactive 3d reconstruction from a single image. In *Proc. of VMV*, 2009.

34 Oliver Wang, Jonathan Finger, Qingxiong Yang, James Davis, and Ruigang Yang. Automatic natural video matting with depth. In *Proceedings of the 15th Pacific Conference on Computer Graphics and Applications*, pages 469–472, Washington, DC, USA, 2007. IEEE Computer Society.

35 Qingxiong Yang, Ruigang Yang, James Davis, and David Nister. Spatial-depth super resolution for range images. *Proc. CVPR*, pages 1–8, 2007.

36 Jiejie Zhu, Liang Wang, Ruigang Yang, James E. Davis, and Zhigeng pan. Reliability fusion of time-of-flight depth and stereo for high quality depth maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99(PrePrints), 2010.

37 J.J. Zhu, L. Wang, R.G. Yang, and J. Davis. Fusion of time-of-flight depth and stereo for high accuracy depth maps. In *Proc. CVPR*, pages 1–8, 2008.

Participants

- Michael Balda
Metrilus GmbH – Erlangen, DE
- Christopher Barsi
MIT – Cambridge, US
- Erhardt Barth
Universität Lübeck, DE
- Sebastian Bauer
Univ. Erlangen-Nürnberg, DE
- James E. Davis
University of California – Santa Cruz, US
- Adrian Dorrington
University of Waikato, NZ
- Martin Eisemann
TU Braunschweig, DE
- Peter Eisert
Fraunhofer-Institut – Berlin, DE
- Jürgen Gall
MPI für Intelligente Systeme – Tübingen, DE
- Marcin Grzegorzek
Universität Siegen, DE
- Diego Gutierrez
University of Zaragoza, ES
- Uwe Hahne
SICK AG – Waldkirch, DE
- Thomas Helten
MPI für Informatik – Saarbrücken, DE
- Ivo Ihrke
Universität des Saarlandes, DE
- Slobodan Ilic
TU München, DE
- Shahram Izadi
Microsoft Research UK – Cambridge, GB
- Bernd Jähne
Universität Heidelberg, DE
- Andreas Jordt
Universität Kiel, DE
- Thomas Kilgus
DKFZ – Heidelberg, DE
- Kwang In Kim
MPI für Informatik – Saarbrücken, DE
- Reinhard Klein
Universität Bonn, DE
- Andreas Kolb
Universität Siegen, DE
- Daniel Kondermann
Universität Heidelberg, DE
- Jens Kubacki
Mesa Imaging AG – Zürich, CH
- Cornelia Lanz
TU Ilmenau, DE
- Seungkyu Lee
SAIT – South Korea, KR
- Damien Lefloch
Universität Siegen, DE
- Frank Lenzen
Universität Heidelberg, DE
- Oisin Mac Aodha
University College London, GB
- Aditi Majumder
Univ. of California – Irvine, US
- Rahul Nair
Universität Heidelberg, DE
- P. J. Narayanan
IIIT – Hyderabad, IN
- Shohei Nobuhara
Kyoto University, JP
- Martin Profittlich
PMD Technologies – Siegen, DE
- Ramesh Raskar
MIT – Cambridge, US
- Christian Schaller
Metrilus GmbH – Erlangen, DE
- Andreas Schilling
Universität Tübingen, DE
- Alexander Seitel
DKFZ – Heidelberg, DE
- Christian Theobalt
MPI für Informatik – Saarbrücken, DE
- Julien Thollot
SoftKinetic-Brussels, BE
- Gordon Wetzstein
MIT – Cambridge, US
- Giora Yahav
Microsoft – Haifa, IL
- Ruigang Yang
University of Kentucky, US