https://doi.org/10.2352/ISSN.2470-1173.2021.6.IRIACV-301

High-speed Inline Computational Imaging for Area Scan Cameras

Bernhard Blaschitz, Simon Breuss, Lukas Traxler, Laurin Ginner; AIT Austrian Institute of Technology; Vienna, Austria; Svorad Štolc; Photoneo; Bratislava, Slovakia

Abstract

For quality inspection in different industries, where objects may be transported at several m/s, acquisition and computation speed for 2d and 3d imaging even at resolutions in the micrometer (μm) scale is essential. AIT's well-established Inline Computational Imaging (ICI) system has until now used standard multi-linescan cameras to build a linear light field stack. Unfortunately, this image readout mode is only supported by few camera manufacturers thus effectively limiting the application of ICI software. However, industrial grade area scan cameras now offer frame rates of several hundred FPS, so a novel method has been developed that can match previous speed requirements while upholding and eventually surpassing previous 3D reconstruction results even for challenging objects. AIT's new area scan ICI can be used with most standard industrial cameras and many different light sources. Nevertheless, AIT has also developed its own light source to illuminate a scene by high-frequency strobing tailored to this application. The new algorithms employ several consistency checks for a range of base lines and feature channels and give robust confidence values that ultimately improve subsequent 3D reconstruction results. Its lean output is well-suited for real-time applications while holding information from four different illumination direction. Qualitative comparisons with our previous method in terms of 3d reconstruction, speed and confidence are shown at a typical sampling of $22\mu m/pixel$. In the future, this fast and robust inline inspection scheme will be extended to microscopic resolutions and to several orthogonal axes of transport.

Introduction

Quality and process control in production facilities e.g. for automotive, electronics print or packaging inspection place high demands for fast optical inline inspection systems, which often means continuously moving objects at speeds up to several meters per second. Typical criteria are high throughput, identification of small defects, both glossy and dark surfaces and a necessity for precise 3D measurements. Typical inline acquisition systems are laser line triangulation, structured light projection, stereo linescan cameras or light field imaging. There the requirements are met by capturing a vast number of data, in particular information from multiple illumination directions and various observation directions. The market is well served with new camera developments, providing high frame rates and high pixel counts, not only for linescan or multi-linescan, but also for area scan cameras. To make those applicable for ultra fast inline inspection, appropriate algorithms must be developed. High data throughput is required to handle the online processing and it is essential that these algorithms can be scaled to available state of the art GPUs.

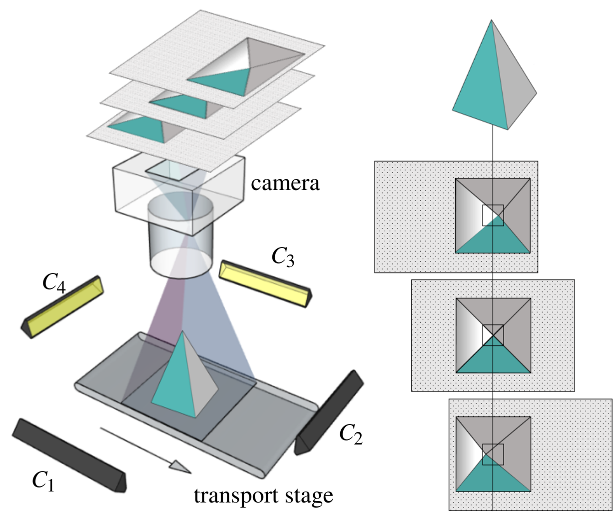

Figure 1. left: *AIT's novel Inline Computational Imaging (ICI) system consists of an industry grade camera, four illuminations $C_1, \cdots, C_4$ and a transport stage that moves an object (here: a pyramid) in front of the camera, usually at several hundred mm/s. Meanwhile, the illuminations are strobed sequentially and an image sequence is acquired.* right: *For each illumination, an image stack consisting of several area scan images is created (only one of four shown here); the transport shift is known and can be used to simplify finding of corresponding features in the stack. Details on the feature matching in Fig. 2.*

Inline Computational Imaging (ICI)

For challenging ultra fast inline inspection tasks we propose a single sensor acquisition system as depicted in Fig. 1. The system consists of a high-speed area scan camera and typically four custom-built high power LED light sources for four different illumination directions. The object is captured periodically during a continuous movement. This results in a light-field image stack showing the inspected object from various observation angles along the transport directions and illuminated from different directions. The four LED modules are strobed sequentially, thus every single image contains the information from one out of four illumination directions. We demonstrate the acquisition process with a Bonito CL-400 camera running at a frame rate of 200 FPS at a resolution of 2320×1728 pixels, but have run the same software also on cameras from Mikrotron as well as JAI or Basler, thus showing the general usability of the software.

Area scan instead of multi-linescan ICI

Having used ICI systems based on multi-linescan acquisition for several years [2], [5] and tested it under different industrial environments, the advent of high-speed area scan cameras called for a paradigm shift in the acquisition mode towards area scan. There are also good reasons to switch to area scan acquisition from an illuminance point of view, more on this in the section on *light utilization*.

Four instead of two lights

Up to now we have been using diffused line light sources, which are great for multi-linescan acquisitions but are not designed for illuminating a whole area at once. Furthermore, experiments showed that an illumination not only from front and back (w.r.t. the transport direction), but also orthogonal to that significantly improves the 3d reconstruction results and their confidences as occlusions and complex geometries can be more robustly calculated (see Fig. 5). Not having found a suitable and affordable light source on the market, we have developed our own.

Overview of this paper

We will show how the paradigm shift from multi-linescan to area scan acquisition is done algorithmically in the next section on the *Algorithmic Framework* and explain all the necessary computational steps. In *Resulting 3d Reconstructions* section, we will show actual 3D reconstructions and how they improve with the new concept. We conclude with a *Summary*, where we also give an outlook to future work.

Algorithmic Framework

The main difference between multi-linescan ICI and area scan ICI is that the latter works with area images (the whole readout of an area scan camera), which are fed into the pipeline sequentially and on which direct stereo matching between views are performed. The overall 3D matching process for an image I_k and its consecutive matching partner image I_{k+b}, where b denotes the baseline (see Fig. 2) can be described as follows:

1. Multi-view matching for m feature channels
2. Fusion of the disparity maps
3. Generate the 3D model
4. Regularization of the final model using total generalized variation (TGV)

Details of the algorithm

While an object moves in front of the "sensor head" (camera plus light sources, see Fig. 1), it is sequentially illuminated from the four light sources $C_s, s = 1 \cdots, 4$, so for each of the light sources, one separate image stack consisting of all images that show the object illuminated from one particular direction is acquired. The images of this stack $I_{\cdot,s}$ are fed into the following pipeline.

1. Feature Calculation & Multi-View Matching for feature channels For a chosen baseline b and a number m of feature channels, each pixel of the image I_k is compared with the same feature in images I_{k-b} and I_{k+b} (see Fig. 2), leading to m different disparity values for every pixel of I_k and analogously for I_{k-b} and I_{k+b}. We also introduce a confidence measure for every such matching. So, for each image I_k, m *disparity maps* and m *confidence maps* are generated.

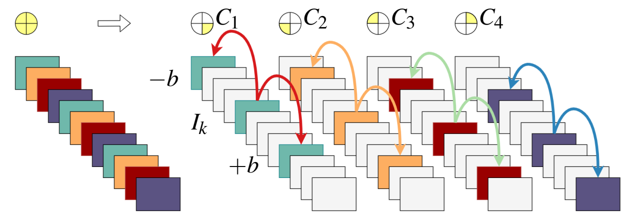

Figure 2. *During the sequential image acquisition, the four LED light sources C_s alternately illuminate the scene (for illustration purposes shown in different colors but actually always with white light). The full image stack is reshuffled into four images stacks $I_{\cdot,1}$ to $I_{\cdot,4}$, one for each light source C_s. For a chosen baseline b, each pixel of the image I_k is compared with images I_{k-b} and I_{k+b} using m different features sizes, leading to m disparity and confidence values for every pixel of I_k. For robustness, the baseline is varied and results are compared.*

2. Fusion of the disparity maps into one single result We take a mean over the m confidence measures and a weighted mean over the m disparities. In the next step, we use the deviation of the m disparities to run an additional consistency check but differently from e.g. [8] by only *reducing* the confidence value in case of discrepancies, not setting it to NaN. This leaves us with exactly one disparity map and one confidence map per image I_k and baseline b. We then rerun this process with a different baseline b and aggregate over disparities and confidences for robustness.

3. Generate the 3D model Having employed a rectification method [2] for our inline system, we can easily integrate the different images $I_{k,s}$ into a scene illuminated from s and save texture, disparity, and confidence for each pixel of the scene. Because of this calibration, we can also overlay the different image stacks stemming from the four different illuminations and thus construct a disparity/confidence tensor which is as wide and long as the acquired scene and has "depth" 4 (because of 4 light sources) or less.

This means that different from other methods that construct a depth volume, whose "depth" is the number of tested disparity steps, we have a very lean and memory friendly data structured that actually contains far more information, namely depth and confidence from 4 different illumination and (possibly) view directions, while using far less memory. This leads to a significant performance gain by approx. a factor of 4 compared to our former approach [5]. The resulting unfiltered 3D model can be see in Fig. 5*a*.

4. Regularization & Denoising After we updated our 3D model using all of the depth, confidence, and texture information form all of our multi-view matches, the final model is regularized. The algorithm used for denoising is an iterative TGV solver. For more details about the Primal Dual algorithm used to solve the TGV problem, see [1]. The results of the denoising can be see in Fig. 5(*c and h*).

Resulting 3D Reconstructions

To demonstrate the performance of the novel area based 3D reconstruction method an artificial scene with various challenging objects such as a coin, an black anodized aluminium screw and part of a drill shaft was arranged. The same scene was acquired with AIT's multi-linescan ICI as well as our novel area scan ICI and compared qualitatively. Fig. 3 shows this scene and

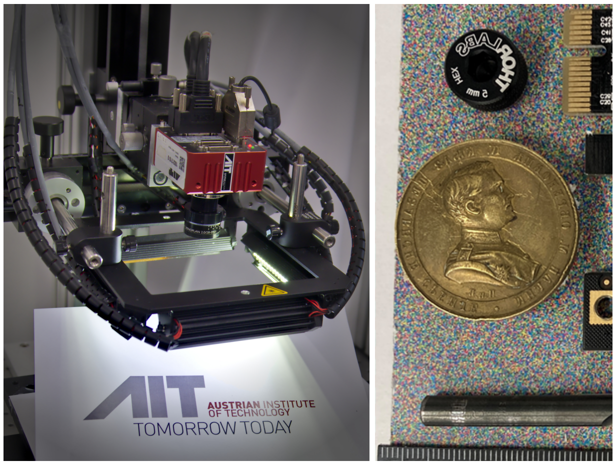

Figure 3. *Real world scene with custom built-light source;* left: *experimental setup, depicting AIT's four prototype light sources that allow for fast strobing of an area underneath the used camera @200FPS,* right: *artificial scene with various surface properties and a size of approximately 45x85mm.*

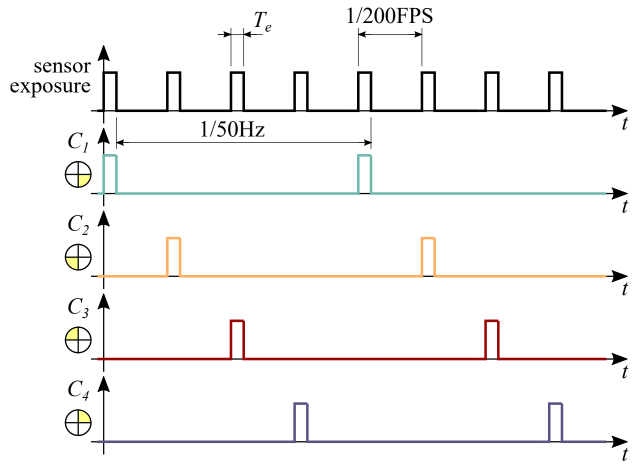

Figure 4. *Strobing pattern for the light-field acquisition system: the area-sensor is used at its maximum frame rate of 200 FPS. To avoid motion blur a short exposure time $T_e = 300\mu s$ (not in scale) is used. Each light source C_1 to C_4 only strobes every fourth cycle. This results in a very short duty-cycle of 1.5% allowing for high peak currents without thermal overload of the LEDs.*

the used experimental setup. For a detailed description of quantitative comparisons please refer to [7].

Multi-linescan vs. Area scan ICI

There are several reasons for the paradigm shift in our Inline Computational Imaging system:

Camera availability. Although multi-linescan ICI [5] is very good for continuous movements and covers a wider range of illumination angles, only few camera manufacturers support this particular mode of readout. If they don't, we had to read out the whole sensor information and crop the multiple lines in memory, which is wasting bandwidth and time, but viable for our first steps in microscopy [6]. Furthermore, the frame rates of area scan cameras are increasing every year and the traditional distinction in performance between (multi-)linescan and area scan cameras is starting to vanish.

Algorithmic robustness. Furthermore, feature matching in area scan images is more robust than in multi-linescan light field stacks, which reduces the dependency on a highly precise linear stage. With future applications on gantry systems and robotic platforms in mind, this new mode was explored.

Light utilization. It is virtually impossible to construct a light source that would optimally direct light into those narrow regions of space that are captured by the multi-linescan approach. This is even truer for the case of photometric stereo [1], where the light should come from specifically defined directions as well. Moreover, the multi-linescan acquisition strongly limits the amount of light which can be utilized: acquisitions during fast continuous movements require very short exposure times to avoid motion blur. The illuminance of the light sources is a crucial limiting factor for the maximal acquisition speed. In the multi-linescan operation only several sparsely distributed lines (e.g. 43 lines for 43 independent views in the light-field) are captured, in contrast in area scan operation mode all 1728 sensor lines read during the illumination flash. This implies that in area scan operation the utilization of light is by factor 40 better than in multi-linescan operation.

Novel 4 light Illumination for ICI

In contrast to other common in-line acquisition systems (e.g. laser line triangulation or stereo line-scan cameras), which only require the illumination of a single line, our novel ICI system requires the illumination of the whole area observed by the camera. To serve industrial applications it is necessary that fast moving objects must not be stopped for inspection. Considering a spatial sampling of $22\mu m$/pixel and a target acquisition speed of up to $73mm/s$, exposure times above $300\mu s$ would result in motion blur deteriorating the 3D reconstruction quality. A considerably high illuminance is required for such short exposure times. In contrast to equivalent line illuminations the required total luminous flux is by factor 100 larger for the area scan illumination. Additionally, fast strobing is required for the sequential illumination from different directions. In total these requirements far exceed available machine vision light sources.

Consequently, a custom illumination had to be developed. Each bar illumination $(C_1, C_2, C_3$ and $C_4)$ consist of 160 white LEDs. The LEDs are specified with a luminous flux of $247lm$ with a power consumption of $2W$ at $700mA$ forward current in constant operation. Using a high speed camera facilitating a maximum frame-rate of $200FPS$, the sequential strobing of C_1 to C_4 (see Fig. 4) results in a flash rate of $50Hz$ for each bar. Given the exposure time of only $300\mu s$ the light sources are only used with an duty cycle of 1.5% during full speed operation. Thus, a peak current of up to $20A$ can be used without exceeding the thermal limits. The LEDs suffer from a drop in efficiently at high peak currents, still the overdrive in strobed operation allows to facilitate around $600klm$ for each light bar.

Discussion of the results

The acquisitions of the artificial scene (see Fig. 3, *right*) was performed with a lateral sampling of $22\mu m$/pixel, qualitative comparison of the two modalities, (multi-linescan and area scan) was performed depicting approximately the same region of interest. As Fig. 5 shows, the overall 3D reconstruction of our

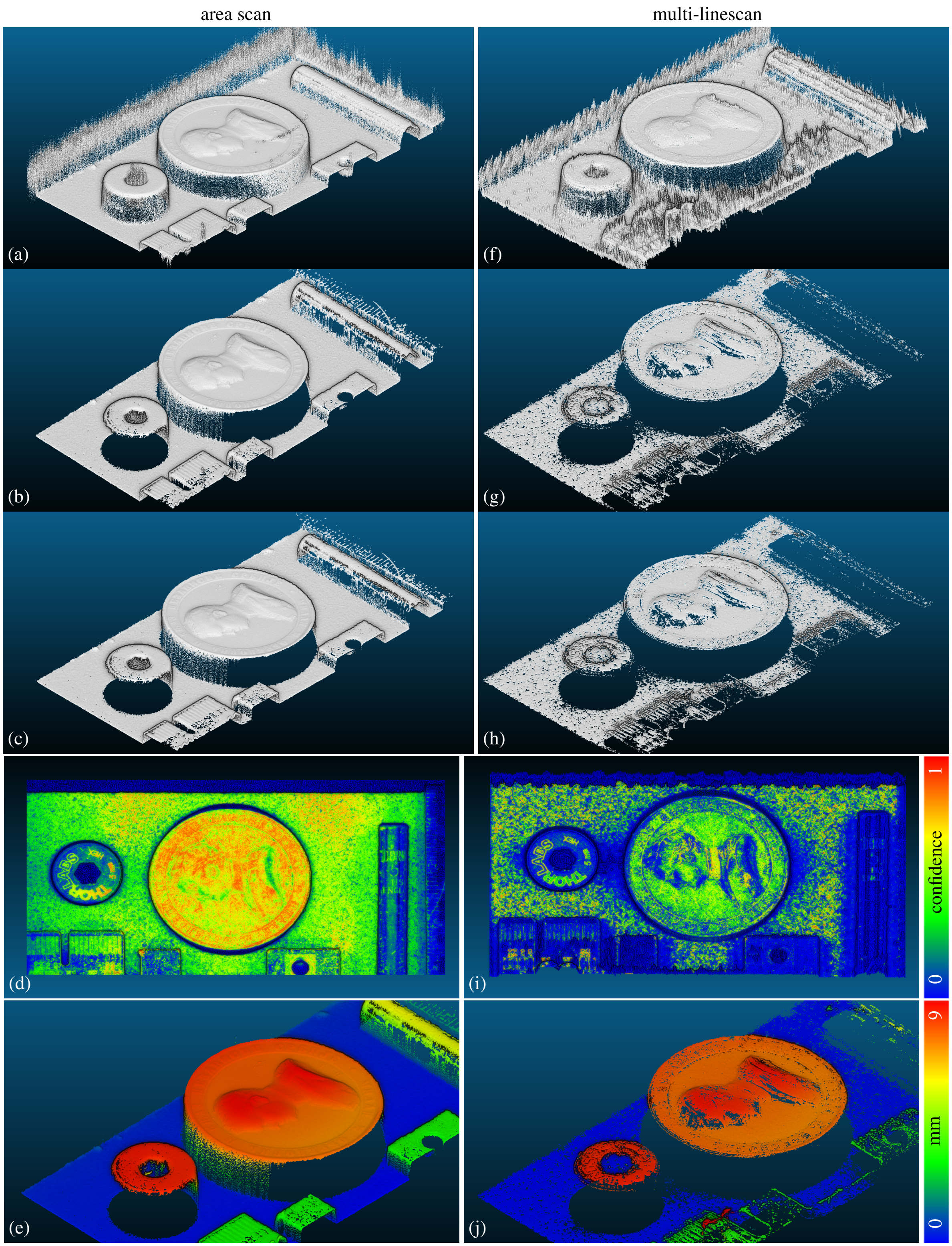

Figure 5. *Qualitative comparison:* (a-e) *depict the results form the novel area scan ICI,* (a) *gives the unfiltered reconstruction of the point cloud from the scene in Fig. 3,* (b) *is the filtered version with a confidence threshold of 0.1 to remove low confident values,* (c) *shows result using a denoising algorithm.* (d)*: The confidence map is given with the corresponding color-table and in* (e) *the color-coded depth map of* (c) *is shown in the depth range of [0-9mm]. In* (f-j) *the corresponding results of our former used multi-linescan ICI are given. It can be seen that the results after the first step (*a *vs.* f*) is already much better and that additional lighting directions help boost the confidence along ridges (*b *vs.* g*). The new methods therefore increases the overall confidence for the validity of depth results (*d *vs.* i*) approx. by a factor of 2.*

novel area scanning approach improves compared to the previous multi-linescan approach because of a better matching scheme, higher information throughput thanks to the lean memory usage and a better illumination configuration. The confidence values are now significantly higher, especially on difficult regions such as the black aluminium screw or the drill. This effect can be seen in the filtered version (b) and (g), where a confidence filter of 0.1 was used to dispose low confident points of the point cloud. Here, regions that very previously partly or completely in the shadow can now be illuminated better, which significantly improves feature matching. When using the denoising algorithm (TGV) during processing, the overall results improve in smoothness, which can of course also lead to errors in the depth estimation, as can be seen in (h) at the drill shaft for example. Confidence measure images (d) and (i) give additional information at which location high or low confident depth values can be calculated. The color-coded depth values can be found in (e and j), which are the color-coded denoised point clouds from (c and h). These images highlight the fact that ICI is a metrically calibrated system that can be used for inline measurement and error quantifications in industrial processes. As can be seen the novel area scanning approach gives significantly higher confidence values as more matching points can be used for the same structural features. The overall confidence of the whole scene is 0.38 for the area scan method and 0.18 for the multi-linescan method. Note that these values are not directly quantitatively comparable as the imaging area is slightly different for the two modalities, nevertheless the confidence values for the area scan method is approximately by a factor of 2 higher. For definitions of the used metrics and more quantitative results, please refer to [7].

Summary

This work shows a novel 3D reconstruction method, which improves our previous multi-linescan ICI, based on a sequential acquisition of area scan images. A new disparity calculation method is presented that surpasses the standard way of calculating depth volumes because of a very efficient memory usage scheme by approximately a factor of 4. Moreover, every pixel in the scene is saved with several depth hypotheses which come from four different light field stacks and accompanied with a matching confidence value, thus allowing in the following 3D reconstruction (possible for multiple depth hypothesis, see [4]) to pick disparity values only matching certain confidence values. We also introduce a LED bar light source that was constructed at AIT in order to fulfil the speed and high illumination requirements of our setup. In conclusion, the new ICI framework is computer vision software that turns standard machine vision components into an inline 3d measurement device.

Future Work

A more extensive analysis of the performance in comparison to other 3D reconstruction methods will be presented in [7].

In the future, we want to apply the new area ICI algorithms also in microscopy, building on [6]. As feature matching in area scan images is more robust than in multi-linescan imaging, it reduces the dependency on a highly precise linear stage. As a next step, we will work with a gantry system used in manufacturing and ultimately want to try how well ICI performs on a robotic arm.

References

[1] Antensteiner, Doris. Fusion of Light Field with Photometric Stereo. Technische Universität Graz, Institute of Computer Graphics and Vision (ICG), 2018, PhD thesis.

[2] Blaschitz, Bernhard, Štolc, Svorad and Antensteiner, Doris. Geometric calibration and image rectification of a multi-line scan camera for accurate 3D reconstruction. Electronic Imaging 2018 (2018): 240-1.

[3] Diebold, Maximilian. Light-Field Imaging and Heterogeneous Light Fields. Universität Heidelberg, 2016, PhD thesis.

[4] Kopf, Christian, Pock, Thomas, Blaschitz, Bernhard and Štolc, Svorad. Inline Double Layer Depth Estimation with Transparent Materials, German Conference on Pattern Recognition, submitted.

[5] Štolc, Svorad, Soukup, Daniel, Holländer, Branislav and Huber-Mörk, Reinhold. Depth and all-in-focus imaging by a multi-line-scan light field camera, Journal of Electronic Imaging, 23(5), 2014.

[6] Traxler, Lukas and Štolc, Svorad. 3D microscopic imaging using Structure-from-Motion, Electronic Imaging 2019.16 (2019): 3-1.

[7] Traxler, Lukas, Ginner, Laurin, Breuss, Simon and Blaschitz, Bernhard. Experimental Comparison of Optical In-line 3D Measurement and InspectionSystems. submitted.

[8] Zinner, Christian, Humenberger, Martin, Ambrosch, Kristian and Kubinger, Wilfried. An optimized software-based implementation of a census-based stereo matching algorithm. In International Symposium on Visual Computing (pp. 216-227). Springer, 2008.

Author Biography

Dr. Bernhard Blaschitz earned his master's degree in Mathematics from the University of Vienna in 2008 and a PhD degree in Applied Geometry from Technical University of Vienna, Austria in 2014. He joined the AIT Austrian Institute of Technology in 2015 and works as a scientist at the Center for Vision, Automation & Control. His main research areas are computational imaging and calibration methods.

DI Simon Breuss earned his master's degree in computer science at the Alpen-Adria University of Klagenfurt in 2008, where he also worked as a member of the System Security Research Group. In 2009 he joined the AIT Austrian Institute of Technology. He is working in the field of computational imaging and 3d reconstruction, and is focused on the development and implementation of algorithms, models, and their simulation.

FH-Prof. DI Dr. Lukas Traxler earned his master's degree in Biomedical Engineering in 2014 and a PhD degree in Technical Physics in 2018 both at the Technical University of Vienna, Austria. He joined the AIT Austrian Institute of Technology as scientist in 2017. Since 2014 he is lecturer at the University of Applied Sciences Technikum Wien. His main research areas are technical optics and computational imaging.

DI Laurin Ginner, PhD wrote his doctoral thesis on the Medical University of Vienna (2019) after completing his master thesis in Physical Energy and Measurement Engineering on the Technical University of Vienna. In 2020 he joined the AIT Austrian Institute of Technology as a Scientists in the topic's technical optics and computational imaging.

Dr. Svorad Štolc is the Head of R&D in company Photoneo. He is an expert in machine vision, artificial intelligence and parallel computing. In 2001 he gained a masters degree at Comenius University in Bratislava and PhD degree in 2009 at Slovak Academy of Sciences in Bratislava and at Technical University of Košice. He took the role of senior scientist at AIT Austrian Institute of Technology GmbH in Vienna for several years. During this time he published multiple internationally awarded scientific articles.