

PROCEDURE ENABLING SIMULATION AND IN-DEPTH ANALYSIS OF OPTICAL EFFECTS IN CAMERA-BASED TIME-OF-FLIGHT SENSORS

M. Baumgart^{1,*}, N. Druml², C. Consani¹

¹ CTR Carinthian Tech Research AG, Europastr. 12, Villach, Austria - (Marcus.Baumgart, Cristina.Consani)@ctr.at

² Infineon Technologies Austria AG, Development Center Graz, Babenberger Str. 10, 8020 Graz, Austria -
 Norbert.Druml@infineon.com

Commission II, WG II/4

KEY WORDS: Time-of-Flight, ToF camera, optical simulation procedure, 3D reconstruction, ground truth, transient imaging

ABSTRACT:

This paper presents a simulation approach for Time-of-Flight cameras to estimate sensor performance and accuracy, as well as to help understanding experimentally discovered effects. The main scope is the detailed simulation of the optical signals. We use a raytracing-based approach and use the optical path length as the master parameter for depth calculations. The procedure is described in detail with references to our implementation in Zemax OpticStudio and Python. Our simulation approach supports multiple and extended light sources and allows accounting for all effects within the geometrical optics model. Especially multi-object reflection/scattering ray-paths, translucent objects, and aberration effects (e.g. distortion caused by the ToF lens) are supported. The optical path length approach also enables the implementation of different ToF sensor types and transient imaging evaluations. The main features are demonstrated on a simple 3D test scene.

1. INTRODUCTION

State-of-the-art Camera-based Time-of-Flight (ToF) Sensors enable a fast and convenient way to deliver valuable 3D information about the environment. The per-pixel distance information (3rd dimension) is calculated by directly or indirectly measuring the round trip time of actively emitted light from a pulsed or modulated source to the detector (Remondino and Stoppa, 2013). The round trip time is then converted into a geometric distance.

Typical camera-based ToF sensors are capable of measuring objects between a few centimeters and up to approximately 50 meters. The round trip time for such distances is very short (10 ns for 1.5 m distance). Thus, very short pulses and exposure times are needed for direct round trip time measurements (D-ToF) resulting in very low signal-to-noise-ratio (SNR) (Jarabo et al., 2017).

Indirect measurement methods allow the sensor to operate at lower frequencies (MHz range instead of GHz range), but need an additional correlation mapping of the signal to the depth information. This technique, also called correlation-based time-of-flight (C-ToF) imaging or phase-based time-of-flight (P-ToF) imaging or simply ToF imaging, modulates the light source signal and demodulates the received signal. The majority of ToF cameras available today use the Amplitude Modulated Continuous Wave (AMCW) principle, also called Continuous Wave Intensity Modulation (CWIM). The phase shift between the emitted and the received optical signal is measured. This measurement principle is most-widely implemented with a photon mixer device (PMD) per pixel (Schwarte et al., 1997). Each pixel additionally comprises integrated circuits for the demodulation based on the Lock-In principle (Lange, 2000). The optical sensor components and additional electronic controllers with their connections for such a ToF camera are depicted in Figure 1.

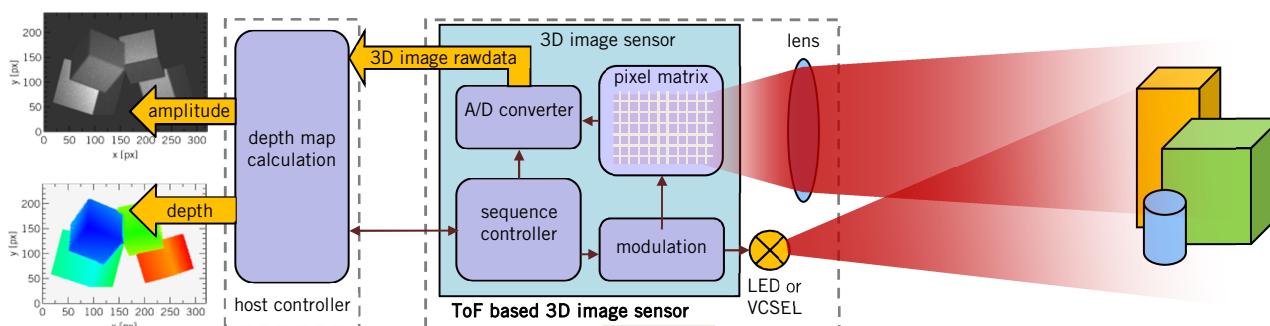


Figure 1. Measurement principle of a camera-based ToF sensor using the amplitude modulated continuous wave measurement technique (REAL3TM 3D Image Sensor, adapted from (Druml et al., 2015))

* Corresponding author

The simulation of a camera-based ToF sensor is essential to estimate sensor performance and accuracy, as well as to help identifying its limitations. Understanding these limits is important to improve the robustness of the distance measurements and the pattern recognition. ToF cameras typically operate at short integration times (up to several milliseconds), which puts high requirements regarding noise suppression on the sensors' electronic and optical components. Nonetheless a noticeable level of noise often affects real-world operation of ToF sensors. Additionally these sensors are affected by further sources of errors – some of them stem from the sensor's measurement principle and the others are due to technical limitations.

Figure 2 gives an overview of the most relevant error sources. A key for in-depth understanding of the sensor's performance is to investigate these error sources separately. This can be achieved with ToF simulation models. This paper presents a simulation pipeline for a detailed investigation of the optical signals and a reference implementation using Zemax OpticStudio and Python. With this approach we are able to account for all effects within the geometrical optics model and to support multiple and extended light sources. A major advantage of this simulation approach is that it allows quantitatively examining multi-object reflection/scattering ray-paths, translucent objects, and aberration effects (e.g. distortion caused by the ToF lens). Thus, it can be used for investigating global illumination and transient imaging effects plus ground truth comparisons.

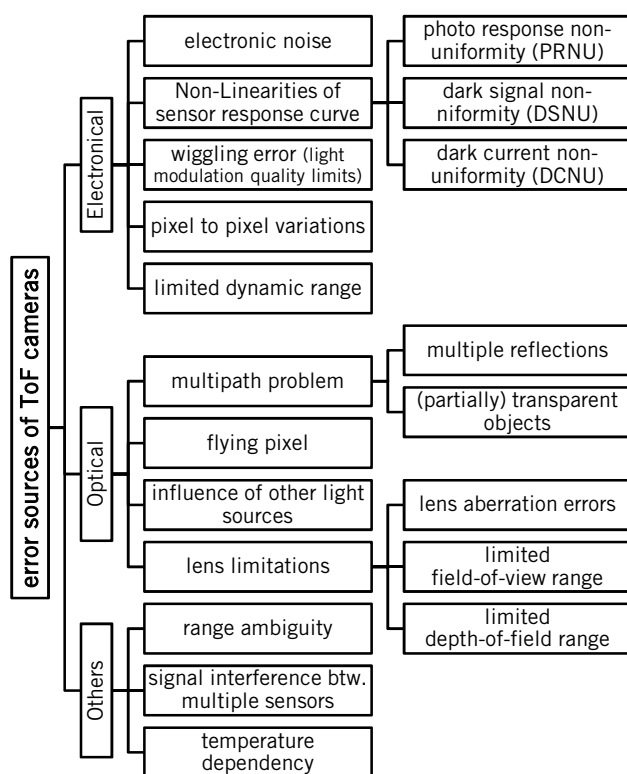


Figure 2. Important error sources in Camera-based Time-of-Flight Sensors (based on error investigations given by (Herbert and Krotkov, 1992), (Adams and Probert, 1996), (Schmidt, 2011), (Grzegorzec et al., 2013), (Meister, 2014))

2. SIMULATION PROCEDURE DESCRIPTION

2.1 Optical path length (OPL) as Master Parameter

Our approach of generating the sensor's optically "not biased"¹ intensity and depth signals is based on steady-state raytracing. The optical path length (OPL) is used as the master parameter. All subsequent time-dependent intensity distributions and derived depths of the scene are calculated using

$$t = s / c \quad (1)$$

where t = travel time of the ray
 s = optical path length of each relevant ray
 c = speed of light

This way the influence of certain ray paths, and frequency analysis for correlation-based ToF can be evaluated without re-tracing the scene. The complete procedure, as shown in Figure 3 and described in the following chapters, is valid for any raytracing-based simulation as long as the per ray optical path length information as part of the ray-tracing result can be stored or extracted.

The described procedure is based on an exemplary implementation using the raytracing-based simulation suite *Zemax OpticStudio Professional Edition (ZOS)* and the *Python* scripting language with the *numpy*, *struct* and *mmap* plugins. Thus, some information given below may be specific to these two Software products.

2.2 Choice of software for reference implementation

Our goal is to have a flexible solution, which can account for all effects within geometrical optics model. Some publicly available toolboxes for simulating ToF signals, namely PMDSim (PMDSim Development Team, 2017), BlenSor (Gschwandtner, 2011) and the Reference Implementation for Phasor Imaging (Gupta et al, 2015) already exist. Since much effort would be required to adapt these toolboxes to our needs, we developed here a different implementation. Our primary concern is to retain the full information on the optical path of each ray and thus to access and analyse the origin and the importance of different error sources on the image reconstruction.

ZOS is an optical system development software mainly designed for optical engineers and offers a user-friendly interface and good documentation of its functionality. It supports Monte Carlo methods for raytracing. Multiple light sources, 3D objects and detectors can be placed within the 3D scene and optical properties can be defined per object respective to its surface. Additionally, the raytracing kernel keeps track of the absolute energy values of each ray at any segment and can account for polarization (each ray is split into its electric and magnetic field vectors), which e.g. enables the calculation of the surface reflectance just from the object's index of refraction without any further assumptions. Rays hitting a certain object or detector can be stored with their complete ray path and additional information (Which objects were hit? Did splitting occur?) into a ray trace raw data file. Especially the powerful scene analysis tools and the in-built functionality of ray-segment fine data storage and access made *ZOS* ideal for OPL scene data generation. The choice of *Python* and its plugins as the programming language for post-processing the ray traced data was mostly because of its easy to understand high-level syntax, the free of charge *spyder* scientific development environment plugin and up to personal preference.

¹ within the geometrical optics model

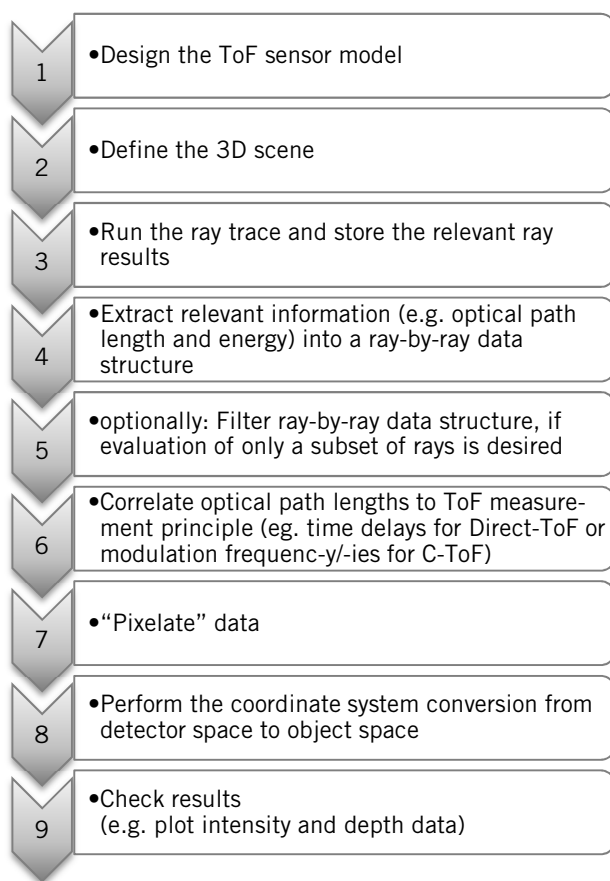


Figure 3. Flowchart of the simulation procedure of the Time-of-Flight sensor signals

2.3 ToF camera and 3D scenery creation

The ToF camera optics is modelled as a part of the scene and the pixel array is an object in the 3D scene that absorbs the incoming rays and stores their ray path parameters (intensity, coordinates, passed objects, object interaction types). *ZOS* supports numerous internal 3D object types like lenses, cubes and spheres, but can also operate on imported CAD scenes or a mixture of both. As a little drawback the light source(s) always have to be placed separately and cannot be imported as part of a CAD file. The ray parameters are typically created with a rigorous random or Sobol-based pseudo-random generation within the boundaries of the source definition. Alternatively, a regular grid of rays can be used.

2.4 Ray-tracing and list of rays containing detector coordinates, intensity and OPL

A ray trace raw data file including energy and optical path length information on a ray segment to ray segment level² is generated for each scene with the *ZOS* raytracing-based simulation suite. The ray segments information is then further analysed using scripts written in the *Python* programming language. Each ray's detector coordinates, its OPL and energy, as well as optionally further information, are extracted from the raw data file using a parser script.

² a ray segment is the space which does not alter the ray behaviour; e.g. the ray travel from one object to the next

2.5 Ray path filtering

The complete ray path from source to detector is stored in the raw data file. This information can be used to optionally filter or categorise the ray data. This helps examining how certain ray paths, wavelengths or object interaction types (refraction, reflection, scattering, ...) contribute to the overall intensity and depth signal.

It is possible to include ambient light sources within the 3D scenery. The separation of their contributions should be done in this filtering step, as they have to be processed differently in the following steps.

2.6 Correlating OPL to ToF measurement principle

The OPL is analysed in this state of the procedure. Depending on the ToF measurement principle to be simulated, the complexity of this method varies. The simplest ToF device class to be modelled is the one based on the D-ToF principle. No OPL conversion is necessary in this case. Only clipping for too far away objects has to be handled. C-ToF cameras operating at a single frequency can be implemented with a ray separation technique, which takes the ray's OPL and the sensor's modulation frequency into account. The modulation frequency defines the sensor's furthest away detectable distance as for longer distances phase-clipping occurs. Additionally, a phase delay discretisation has to be defined. This discretisation influences the accuracy of the depth reconstruction. At least 3 phase delay ranges are necessary. A higher number of ranges leads to more-precise depth reconstruction but also to longer measurement times per frame for a real sensor. The modulation frequency dependent phase delay of each ray is calculated and its intensity is distributed to the according phase delay ranges. The procedure details can be found in standard C-ToF literature (Lange, 2000). It is especially possible to account for the non-linearity of the arctangent behaviour (Luan, 2001) of C-ToF cameras with this implementation. A multi-frequency C-ToF sensor can overcome the limitation of the modulation-frequency bound measurement range and can also detect global illumination effects. Such a sensor is modelled via repeating the procedure of the single-frequency ToF for every frequency.

2.7 Pixelate data

Afterwards, the ray information for all rays within a certain area is combined (→ "pixelated"). The ray-based path length and intensity information is converted into a matrix-like pattern for a given number of horizontal and vertical sensor pixels in this discretization process. In the simplest case (no multi-frequency sensor, no direct vs. global illumination differentiation) this procedure results in 2 matrices (intensity and depth) with each element corresponding to one sensor pixel and containing the averaged path length and integrated intensity values. Different smoothing strategies, where the ray data also contributes to neighbouring pixels, exist. There are also approaches to calculate the standard deviation and confidence range for the depth value at each pixel.

2.8 Coordinate system conversion

Finally, a coordinate system conversion from the sensor size dimensions to object space is necessary. The camera lens behaviour has to be known to achieve this conversation. The simplest approach to define such a camera function is to use a global magnification factor for all x- and y-coordinates (pinhole camera). This gives a first approximation of the 3D scene. This approach does not take into account magnification changes due

to defocus and lens aberration effects for objects at different lateral and axial positions. A lot of more complex coordinate mapping and calibration approaches are presented and discussed in literature [7 – 10].

3. DEMONSTRATING THE CAPABILITIES OF OUR TOF SENSOR PROCEDURE ON A TEST SCENE

3.1 General considerations & Test scene setup

This chapter demonstrates some of the capabilities of our ToF simulation procedure on a simple test scene. Some features and objects are implemented quite detailed in the test scene while for others more simplified models were used. The idea is to show the broad range of use cases (e.g. detailed and rough models) for the procedure.

The test scene consists of 10 objects forming the ToF camera model and another 4 objects representing the actual scene resulting in a total number of only 14 objects.

The ToF camera model comprises a monochromatic point source at 830 nm with an angular Gaussian intensity distribution (18° divergence (full angle) @ $1/e^2$). These are typical value for a VCSEL light source. A petzval lens design was chosen for the receiver optics, because this lens type offers good light collection efficiency. An anti-reflective coating optimised for 830 nm was applied on the lens surfaces. The ToF sensor's pixel matrix is defined as an absorptive detector.

The scene itself consists of 4 cubes, different in size, but with the same surface properties. The surface absorbs 90% of the energy and scatters back the remaining 10%. The scattering is implemented as Gaussian scattering. The following Bi-Directional Scatter Distribution Function (BSDF) definition of Gaussian scattering is used:

$$BSDF(\vec{x}) = A e^{-\frac{|\vec{x}|^2}{\sigma^2}} \quad (2)$$

where A = normalizing constant

\vec{x} = (into surface plane projected vector of specular ray) - (into surface plane projected vector of scattered ray)

σ = width of the Gaussian distribution on the projected plane; set to 1 for all scattering surfaces

The complete scene is pictured in Figure 4A and 4B. Please note that all 3D plots shown in this publication use parallel projection for the 3D effect.

3.2 Scene reconstruction using raytracing data

The reconstruction results shown in Figure 4C – 4E were calculated with the D-ToF sensor implementation. Also the coordinate system conversion was done with the simplest mapping function by defining a global magnification factor for all x- and y-coordinates. This way, the lens aberration errors stemming from the petzval lens can be easily observed in the reconstructed 3D plot (Figure 4G). As the anti-reflective lens coating cannot suppress all reflections, the remaining reflections can be seen as noise in the outer areas of the imaged scene (Figure 4F). The magnitude of this noise is so low, that it has no significance for a real ToF camera. This is demonstrated with the 12 bit AD filtering of the intensity signals. The noise is completely filtered out (Figure 4E & 4G), which also shows high dynamic range achievable with our procedure.

3.3 Ray path based scene analysis

Light passing multiple objects is a problem for ToF sensors because this effect leads to a longer recalculated distance compared to the ground truth. An example of ray path based data filtering is given in Figure 4H – 4J. Here only the contributions of ray paths which hit at least two cubes are plotted. The colourbar of the distance plot validates the effect of the longer recalculated distance.

3.4 ToF sensor model influence on depth calculation

The majority of ToF cameras available today use the Amplitude Modulated Continuous Wave (AMCW) principle, which is mostly realised with a photon mixer device (PMD) per pixel. We implemented the C-ToF sensor based on the AMCW principle in the OPL correlation step in our procedure. The following AMCW single frequency sensor parameters were used to image the test scene. The modulation frequency was set to 25 MHz, which gives a maximum detectable distance of 3 m.

The phase shift was discretised into 4 ranges ($0^\circ - 90^\circ$, $90^\circ - 180^\circ$, $180^\circ - 270^\circ$, $270^\circ - 360^\circ$). The calculated depth from this ToF sensor type is shown in Figure 5A. The intensity distribution is the same like for the D-ToF sensor and, thus, not plotted again. The comparison of the calculated depths between the C-ToF and D-ToF sensor model show the nonlinear depth mapping of the C-ToF sensor (Figure 5B and 5C). This nonlinearity is typically corrected in an additional step in real C-ToF cameras.

3.5 Influence of glass window in front of cubes

The anti-reflective coated ToF camera lens in the test scene model reduced the lens reflections influence on depth calculation to a minimum. In real world scenes with transparent objects one does usually not find these objects anti-reflective coated at the ToF camera wavelength. To examine the influence of transparent objects, the test scene was modified by adding an uncoated glass window.

The window, modelled as BK7 quartz glass with 10 mm thickness, is positioned 1 m away from the ToF camera and covers only the horizontally left half of the scene (Figure 6A and 6B). Around 4% of the incoming light perpendicular to the glass surface is reflected as given by the Fresnel equations. Additionally a 1% Gaussian scattering (parameters as given in paragraph 3.1) was added to mimic surface roughness and glass impurities.

There are three main effects expected to influence the depth estimation. The first is caused by directly reflected or scattered light from the glass, which does not hit any of the cubes. As the glass window is in front of the cubes, these rays will give a too short distance estimation of the cubes. The second and third effect is caused by rays which come from a cube and hit the glass window afterwards, and are either reflected or scattered back towards the cubes (2nd effect) or undergo multiple inner-glass reflections or scatterings (3rd effect). The last two effects will both give a too long distance estimation of the cubes.

We switched back to the D-ToF sensor model to be able to directly compare the calculated depth estimations for the scene with and without glass window. The results are shown in Figure 6C – 6F. The 1st and 2nd effect mentioned above can be observed. In the region of the glass window, where no cubes are behind, the position of the window caused by directly reflected or scattered light can be noticed (effect 1). The cubes directly behind the glass are estimated too far away, which can be seen via the red coloured areas in Figure 6E. Here, the 2nd effect of light reflected or scattered back from the window to the scene

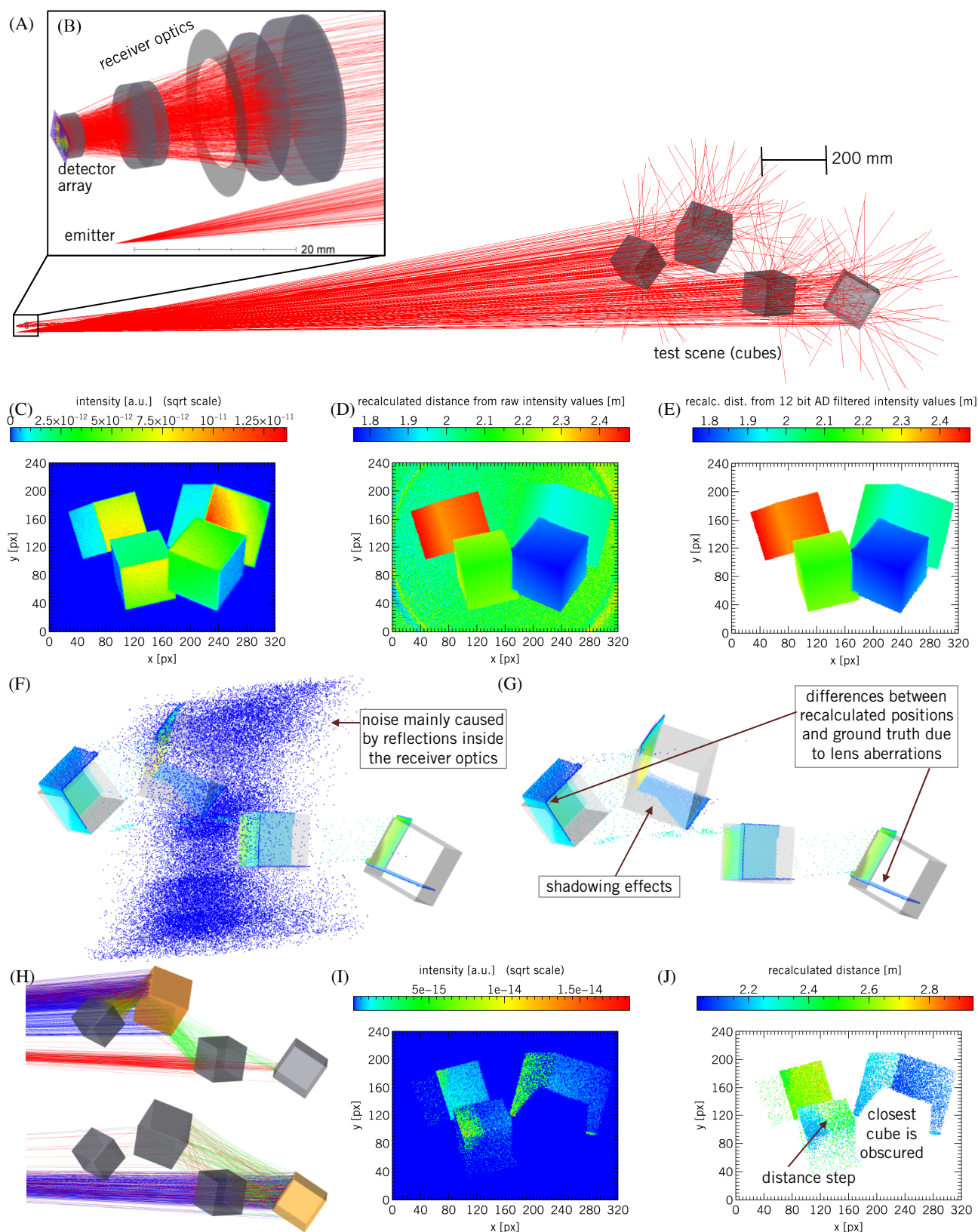


Figure 4. Simple test scene consisting of 4 cubes showing the feasibility of the proposed simulation procedure (D-ToF sensor model used). (A) Complete 3D scenery with ToF camera and cube objects. (B) Implementing the optically relevant ToF camera components comprising light source, receiver optics and detector array. (C) Intensity distribution of the received light. (D – E) Calculated depth from the optical path length using the global magnification factor correspondence (raw signal and 12 bit AD filtered). (F – G) 3D plot of the ground truth object positions overlaid with ToF sensor signals (raw signal and 12bit AD filtered). (H – J) Multiple cubes ray paths. (H) Two exemplary multiple cubes ray paths for rays first hitting the highlighted cube. (I) Intensity distribution of the ray paths which interacted with multiple cubes. (J) Calculated depth contribution from the multiple cubes ray paths.

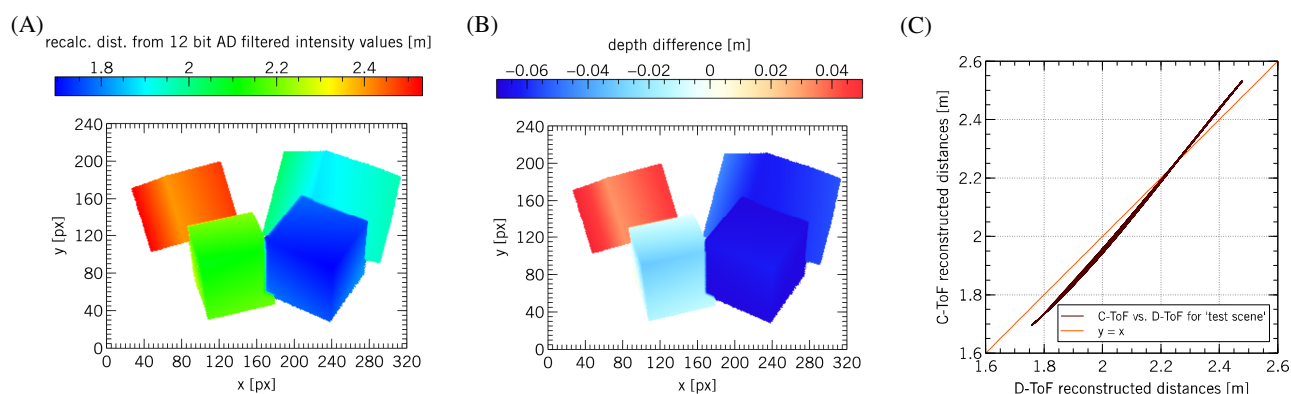


Figure 5. Test scene from Figure 4, but the depth is reconstructed with the single frequency PMD based C-ToF sensor model. (A) Calculated depth from the optical path length using the global magnification factor correspondence. (B) Difference between D-ToF and C-ToF depth calculation. (C) C-ToF vs. D-ToF emphasising the arctangent behaviour of the C-ToF reconstruction algorithm.

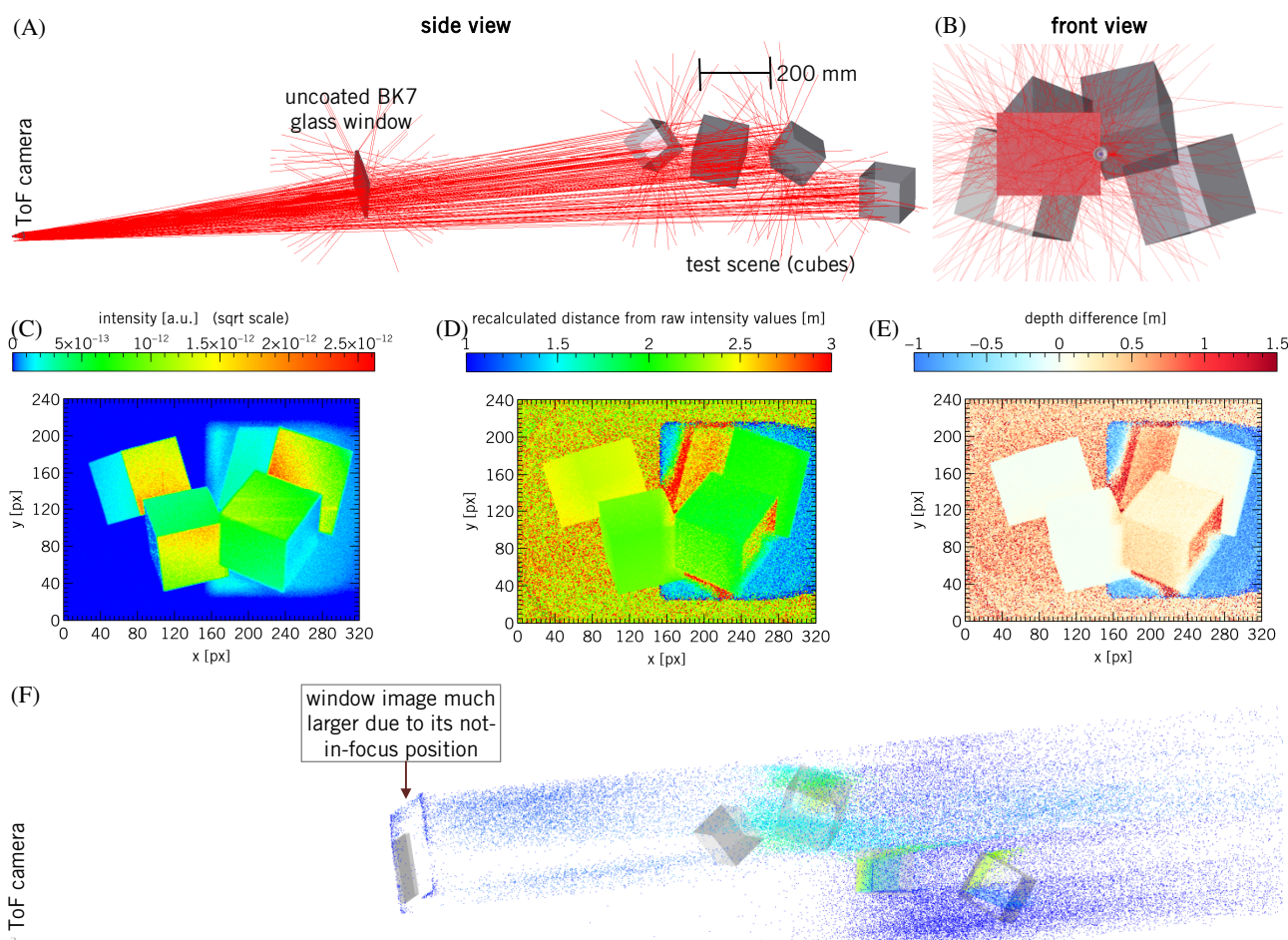


Figure 6. Test scene from Figure 4 after adding an uncoated BK7 window in front of the cubes. (A – B) Complete 3D scenery with ToF camera, cube objects and glass window. (C) Intensity distribution of the received light. (D) Calculated depth from the optical path length using the global magnification factor correspondence. (E) Depth difference between scene with and without glass window (referenced to scene without glass window). (F) 3D plot of the ground truth object positions overlaid with ToF sensor signals.

outweighs the first effect. The third effect cannot be observed in the calculated depth images, because the multiple inner-glass reflections and scatterings are overlaid with other ray paths.

4. CONCLUSIONS AND OUTLOOK

The described ray tracing based simulation procedure for ToF cameras could be successfully demonstrated. The main advantage is the flexibility of the presented approach. Single effects can be studied in detail while neglecting others. The central element of this flexibility is the creation of the scene's OPL master file, which holds all the relevant information, and serves as the starting point for any further analysis.

The choice of the ray tracing technique for the optical scene's evaluation allows considering all effects within the geometric optics model and the generation of a global illumination map. We demonstrated in chapter 3, that our procedure is capable to investigate numerous effects, like filtering and analysing certain ray paths, implementing different ToF sensor embodiments or examining ray interactions with transparent objects.

In the future we want to use the described procedure for multi-frequency ToF sensor analysis and consider implementing the option to add optical and electronical noise to the data. The current progress in GPU based raytracing implementations makes us optimistic to achieve a huge reduction of simulation time in the near future.

ACKNOWLEDGEMENTS

This project has received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 692480. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Saxony, Spain, Austria, Belgium, Slovakia. Austrian co-funding comes from the Federal Ministries of Transport, Innovation and Technology (BMVIT) and Digital and Economic Affairs (BMDW), managed on their behalf by the Austrian Research Promotion Agency (FFG). Additional financial support by the Austrian Provinces of Carinthia and Styria is also gratefully acknowledged.

REFERENCES

Adams, M.D., and Probert, P.J., 1996. The interpretation of phase and intensity data from AMCW light detection sensors for reliable ranging. *The International journal of robotics research*, 15(5), pp. 441-458.

Druml, N., Fleischmann, G., Heidenreich, C., Leitner, A., Martin, H., Herndl, T., & Holweg, G., 2015. Time-of-flight 3D imaging for mixed-critical systems. *2015 IEEE 13th International Conference on Industrial Informatics (INDIN)*, pp. 1432-1437.

Gupta, M., Nayar, S. K., Hullin, M. B., and Martin, J., 2015. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics (ToG)*, 34(5), Article 156.

Gschwandtner M., Kwitt R., Uhl A., and Pree W., 2011. BlenSor: Blender Sensor Simulation Toolbox. In: Bebis G. et al. (eds) *Advances in Visual Computing. ISVC 2011. Lecture Notes in Computer Science*, vol 6939. Springer, Berlin, Heidelberg.

Grzegorzec, M.; et al., eds., 2013. Time-of-Flight and Depth Imaging. Sensors, Algorithms and Applications: *Dagstuhl Seminar 2012 and GCPR Workshop on Imaging New Modalities*, Vol. 8200.

Hansard, M., Lee, S., Choi, O., and Horaud, R., 2012. Time-of-Flight Cameras: Principles, Methods, and Applications. *SpringerBriefs in Computer Science*, Springer Science & Business Media, Chapter 3.

Hebert, M., and Krotkov, E., 1992. 3D measurements from imaging laser radars: how good are they? *Image and vision computing*, 10(3), pp. 170-178.

Jarabo, A., Masia, B., Marco, J., and Gutierrez, D., 2017. Recent advances in transient imaging: A computer graphics and vision perspective. *Visual Informatics*, 1(1), pp. 65-79.

Lange, R., 2000. 3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology. Diss., University of Siegen, Germany.

Lindner, M., and Kolb, A., 2006. Lateral and depth calibration of PMD-distance sensors. *Advances in Visual Computing*, pp. 524-533.

Luan, X., 2001. Experimental investigation of photonic mixer device and development of TOF 3D ranging Systems based on PMD technology. Diss., Department of Electrical Engineering and Computer Science, University of Siegen, Germany, Chapter 6.

Meister, S.N.R., 2014. On Creating Reference Data for Performance Analysis in Image Processing. Diss., Ruperto-Carola University, Heidelberg, Germany.

PMDSim Development Team, 2017. Simulator for continuous-wave Time-of-Flight sensors, specifically PMD sensors. <https://github.com/marlam/pmdsim> (6 April 2018)

Remondino, Fabio, and David Stoppa, eds, 2013. TOF Range-Imaging Cameras. Springer Science & Business Media.

Schiller, I., Beder C., and Koch, R., 2008. Calibration of a PMD-camera using a planar calibration pattern together with a multi-camera setup. *The international archives of the photogrammetry, remote sensing and spatial information sciences*, 21, pp. 297-302.

Schmidt, M., 2011. Analysis, modeling and dynamic optimization of 3d time-of-flight imaging systems. Diss., Ruperto-Carola University, Heidelberg, Germany, Chapter 2.3.

Schwarte, R., Xu, Z., Heinol, H.G., Olk, J., Klein, R., Buxbaum, B., Fischer, B., and Schulte, J., 1997. New electro-optical mixing and correlating sensor: facilities and applications of the photonic mixer device (PMD). *Sensors, Sensor Systems, and Sensor Data Processing. Proc. SPIE 3100*, pp. 245-254.