Time-of-Flight Imaging: Basics, State, and Perspectives

Bernd Jähne

Heidelberg Collaboratory for Image Processing (HCI)
at the Interdisciplinary Center for Scientific Computing
Heidelberg University
Bernd.Jaehne@iwr.uni-heidelberg.de

73. Heidelberger Bildverarbeitungsforum, Stuttgart, 01. Oktober 2019

This material is only thought for personal use of the participants of the 73th Heidelberg Image Processing Forum.
It must not be passed on to third parties without explicit written authorization by the author.

Outline

I **Basics**

II **Characterization**

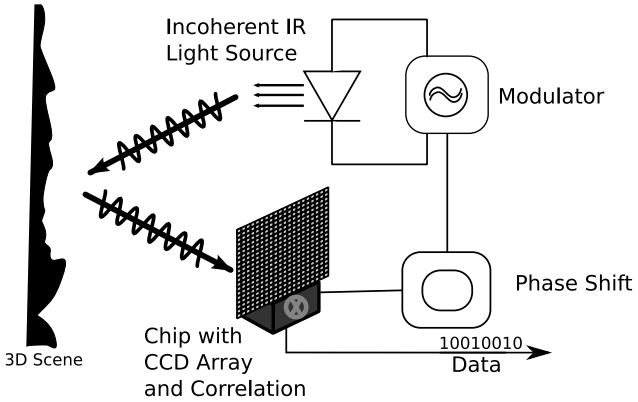
III **Limitations and Artefacts**

IV **Conclusions and Perspectives**

I Basics

Time-of-Flight Imaging

Basic Time-of-Flight Principle I



Depth z from delay Δt:

$$\Delta t = \frac{2z}{c} \quad \rightsquigarrow \quad z = \frac{c\Delta t}{2} \tag{1}$$

Basic Time-of-Flight Principle II

Principle advantages

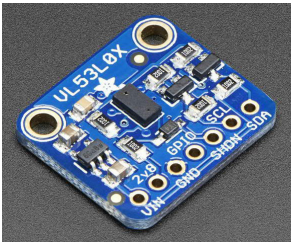
- Direct, calibration free distance/depth measurement
- Principally more compact than any triangulation-based system (e. g., stereo)
- Much less occlusion than with triangulation

Challenges

- Delays short: 1 μs for $z = 150$ m, 1 ns for $z = 150$ mm requiring picoseconds resolution to achieve mm precision
- Illumination not a problem: laser pulses in the range of ps common technology
- Problem related to image sensors: typical exposure times for image sensors $> 1\,\mu$s, clock rates on image sensors ≤ 100 MHz (10 ns period)
- Whole scene is illuminated: indirect illumination via third object make objects appear to be more distant

Common Device: Single Point Laser Distance Meter

Source Würth Electronics



Source Adafruit

As handheld devices and miniature modules with mm accuracy

Scanning Laser Distance Meters

Source Velodyne

Still limited resolution compared to imaging devices

- 900 Points horizontally $\times$ 128 lines $\times$ 20 fps $\rightsquigarrow$ 2.4 Mpixel/s
- compared to 357 Mpixel/s for IMX174 (Sony Pregius 1st generation) 2-D image sensor
- compared to 9,2 Mpixel/s for IMX556PLR (Sony ToF 3-D sensor)

Techniques to measure short delays in imaging devices

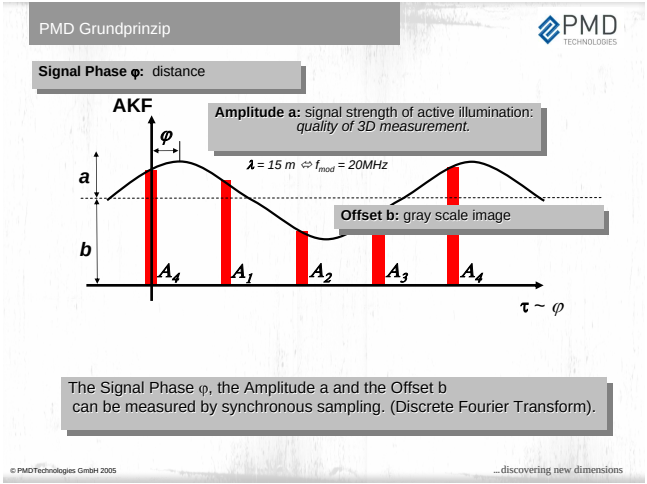
Direct

- Directly measure delay, when first photons are returned
- Required single photon avalance photodiode arrays (SPAD) and very high temporal resolution
- Coincidence measurements allow effective background light suppression (Beer, 2018)

Indirect with much less requirements on temporal resolution and precision limited only by signal-to-noise ratio

- Measuring phase shifts instead of time delay using periodically modulated light
- Gated acquisition with pulsed light (not further considered, mathematically very similar)

Signal Sampling with Periodic Light Modulation



Source PMDTec, 2006, Lynkeus Project

Signal Model for Phase Measurements

Signal $I = a_0 + a_m \cos(\alpha + \Delta\phi)$ mit

- Offset a_0 (DC part of modulated light + background light),
- Amplitude a_m (of reflected modulated light),
- Phase shift $\Delta\varphi = 2\pi\nu\Delta t$

Using Eq. (1), relation to distance/depth results in

$$z = \frac{c\Delta\varphi}{4\pi\nu} \tag{2}$$

Three unknowns $\rightsquigarrow \geq 3$ samples required,
determined by N equidistant samples at phases $\alpha = 2n\pi/N$

Phase Measurement: General Solution

Compute mean and amplitude and phase of smallest frequency of Fourier series = basic known modulation frequency

- DC part a_0 = mean of all samples:

$$a_0 = \frac{1}{N}\sum_{n=0}^{N-1} I_n \tag{3}$$

- Modulation amplitude a_m = magnitude of basic frequency:

$$a_m = \frac{2}{N}\left|\sum_{n=0}^{N-1} I_n \exp\left(\frac{2\pi \mathrm{i} n}{N}\right)\right| \tag{4}$$

- Phase shift $\Delta\varphi$ = phase of basic frequency:

$$\Delta\varphi = \arg\left[\sum_{n=0}^{N-1} I_n \exp\left(\frac{2\pi \mathrm{i} n}{N}\right)\right] \tag{5}$$

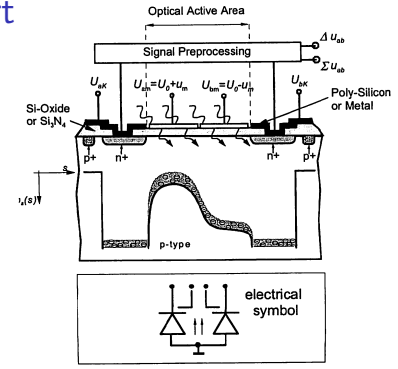
Simplest Solution with Four Samples

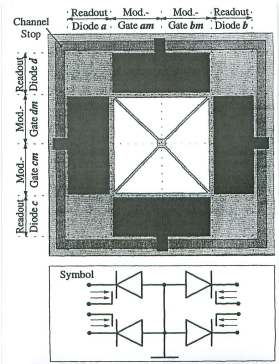
Samples at phases α 0, 90, 180 und 270 degree:

$$\begin{aligned} \text{Offset} \quad a_0 &= \frac{I_0 + I_1 + I_2 + I_3}{4} \\ \text{Amplitude} \quad a_m &= \frac{\left[(I_0 - I_2)^2 + (I_1 - I_3)^2\right]^{1/2}}{2} \\ \text{Phase shift} \quad \Delta\varphi &= \arctan \frac{I_3 - I_1}{I_2 - I_0} \end{aligned} \tag{6}$$

Early example of computational embedded vision, where four channels are combined to compute depth images

State of the Art





- So far only two-tap sensors are realized. Depth images require two consecutive image (one with 0 and 180 degree sampling, one with 90 and 270 degree sampling).
- Four-tap-**P**hoton **M**ixing **D**evice (PMD) image sensor was already proposed Schwarte et al. (1997); realization?

II Characterization

Time-of-Flight Imaging

Temporal Noise of Phase

Noise model for *ideal* sensor noise variance σ_I^2 (Rapp, 2007):

Offset a_0	$\sigma(a_0) = \frac{\sigma_I}{2}$
Amplitude a_1	$\sigma(a_1) = \frac{\sigma_I}{\sqrt{2}}$
Phase shift $\Delta\phi$	$\sigma(\Delta\phi) = \frac{\sigma_I}{\sqrt{2}a_1} = r_I$ (r_I relative amplitude error)
Depth z	$\sigma(z) = \frac{c}{4\pi\nu}\sigma(\Delta\phi) = \frac{c}{4\pi\nu} r_I$
Depth range Δz	$\Delta z = \frac{c}{4\pi\nu} 2\pi = \frac{c}{2\nu}$ phase range $[0, 2\pi]$)
Relative depth error	$\frac{\sigma(z)}{\Delta z} = \frac{r_I}{2\pi}$

Temporal Noise of Phase *ideal* ToF-Sensor

- Only Photon noise (no dark noise): $\sigma_I^2 = I$ (I in number of electrons)
- Best possible amplitude modulation, two tap, sinusodial : $a_1 = \dfrac{I}{2\sqrt{2}}$
- Depth error: $\sigma_z = \dfrac{c}{2\pi n}\dfrac{\sigma_I}{I}$

Example: Sony IMX556PLR

- Modulation frequency $\nu = 100\,\mathrm{MHz}$
- Full amplitude for one tap: $I \approx 135\,000\ \mathrm{e}^-$ according to measurements of Chang (2018)
- $\rightsquigarrow$ Minimum depth error $\sigma_z = 1.3\,\mathrm{mm}$

Temporal Noise Measurements (Chang, 2018)

TEMPORAL PRECISION

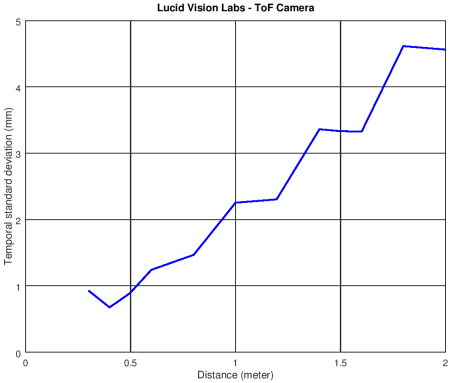
PRELIMINARY DATA

Test Conditions

- 4 lasers, at 1W optical power each
- Room light off
- Exposure time = 1000 us for all tested distances except for 30 cm
- 11x11 pixels in image center
- Standard deviation over 100 images

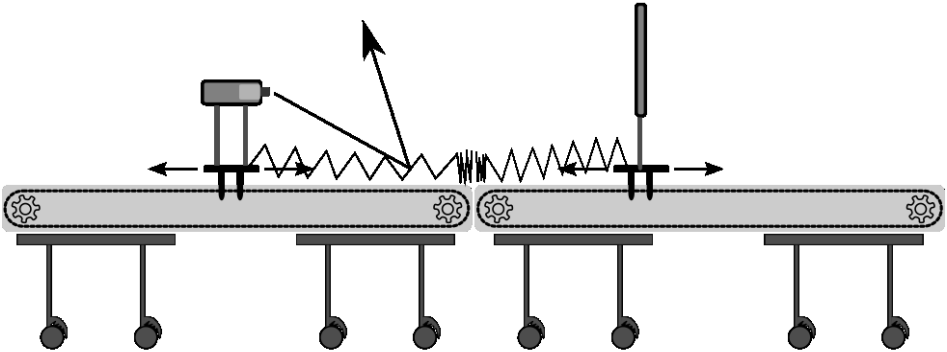
Results

- σ = 2.25 mm at 1 meter
- σ = 4.56 mm at 2 meter



III Limitations and Artefacts

ToF Calibration Set-up

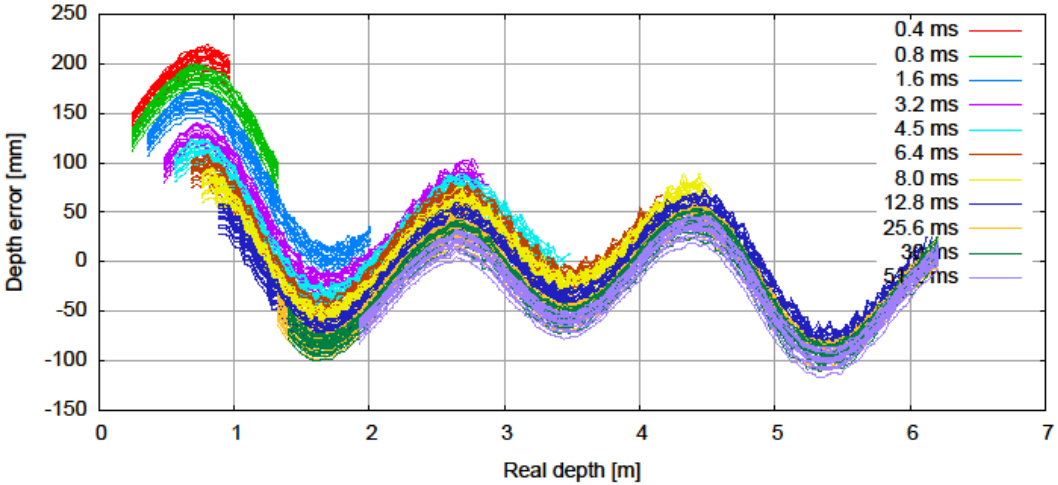


Sketch of calibration setup for ToF-cameras

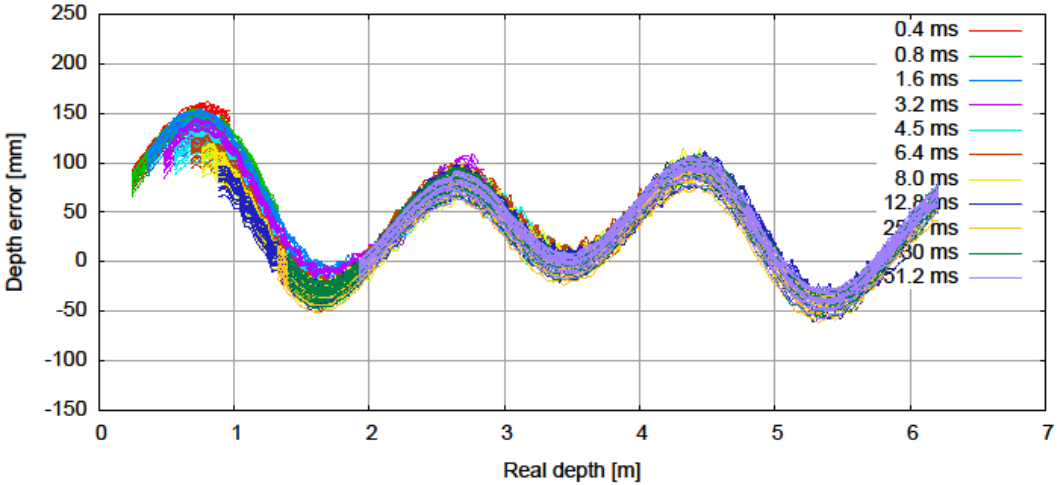
Multiple paths from light source to object

- Cannot be detected with a single frequency: summation of periodic signals of same frequency result in a periodic signal
- Multiple frequencies can help, but not really sufficient
- For conditions of good recovery of multiple reflections and/or objects in different depthes, see Jähne (2020)

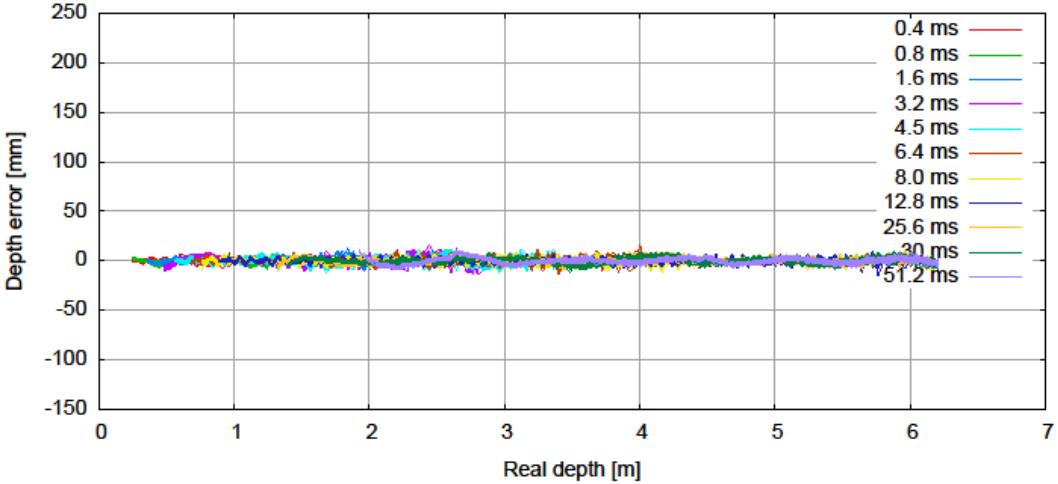
Depth Error Raw Data



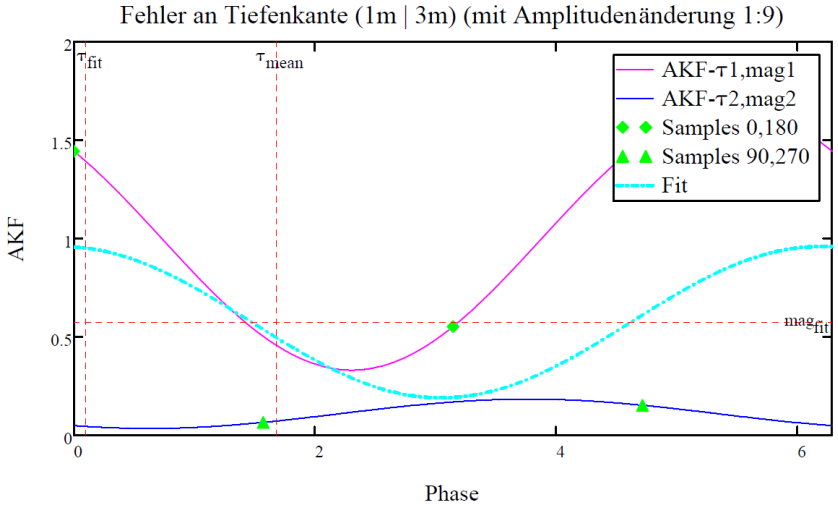
Depth Error after Correction of Amplitude Dependency



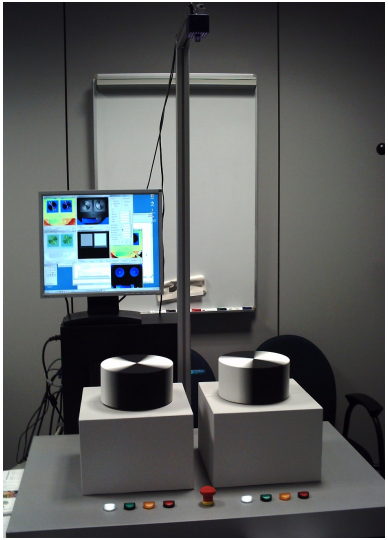
Depth Error after Correction of "Depth Wiggling"



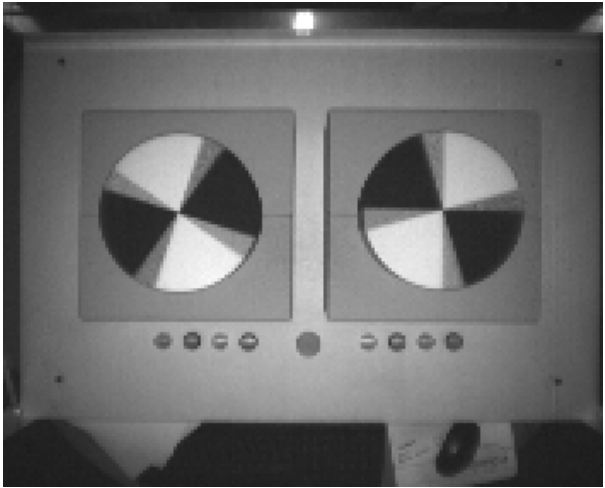
Motion-induced Depth Error



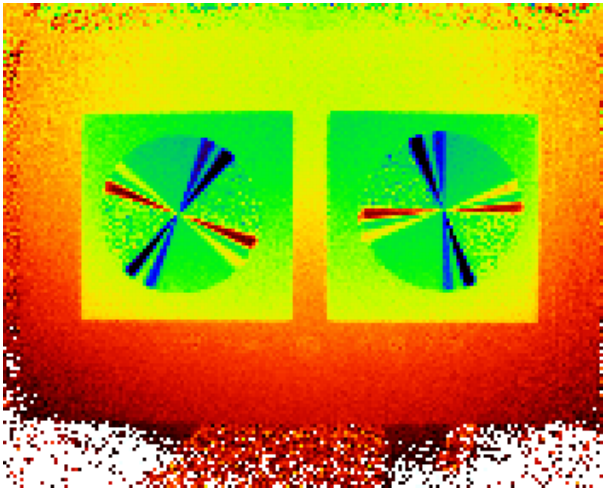
Demo:Rotating Cylinders with Black/White Pattern



Rotating Cylinders: Amplitude Image



Rotating Cylinders: Depth Image



Limited Depth Dynamics

- Basic problem related to square distance law in far field: $a_1 \propto 1/Z^2$
- Temporal noise increases with depth
- *Ideal* ToF-Sensor, photon noise limited and no background light:

  $\sigma(Z) \propto 1/\sqrt{a_1} \propto Z$

- Dominating environmental light:

  $\sigma(Z) \propto 1/a_1 \propto Z^2$

- As bad as any triangulation technique (Stereo, see Jähne, 2012):

  $\sigma(Z) \propto Z^2$

Temporal Noise Measurements (Chang, 2018)

TEMPORAL PRECISION

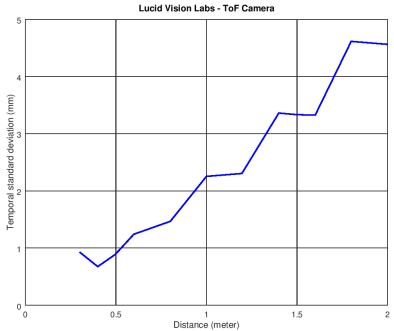
PRELIMINARY DATA

Test Conditions

- 4 lasers, at 1W optical power each
- Room light off
- Exposure time = 1000 us for all tested distances except for 30 cm
- 11x11 pixels in image center
- Standard deviation over 100 images

Results

- σ = 2.25 mm at 1 meter
- σ = 4.56 mm at 2 meter



LUCID VISION LABS, INC.

LUCID VISION LABS

Photon-noise limited linear increase of temporal noise with depth possible in the dark

IV Conclusions and Perspective

Conclusions and Perspective

- ToF imaging is more than 20 years old but still in its infancy; can have significant advantages over triangulation, but to be used with care considering current limitations
- Most serious limitations are low dynamic depth range, motion artefacts, and bad performance in cluttered scenes
- ToF imaging has bright future with rapid progress of solid state images
- Most important are multi-tap (at least three taps, better much nore) ToF image sensors and sensors with pixels individually adapting to brightness level for higher dynamic range
- Also required: lenses with much less internal scatter
- LIDAR systems remain biggest competition; the race is open, may coexist because of different application areas

Literatur I

B. Jähne, H. Haußecker und P. Geißler (eds.).
Handbook of Computer Vision and Applications, Vol. 1 Sensors and Imaging.
Academic Press, San Diego, 1999.

B. Jähne und H. Haußecker (eds.).
Computer Vision and Applications, A Guide for Students and Practitioners.
Academic Press, San Diego, 2000,
http://www.sciencedirect.com/science/book/9780123797773.

B. Jähne.
Digitale Bildverarbeitung und Bildgewinnung.
Springer, Berlin, 7. Auflage, 2012, http://dx.doi.org/10.1007/978-3-642-04952-1.

Literatur II

B. Jähne.
Digitale Bildverarbeitung und Bildgewinnung.
Springer, Berlin, 8. Auflage, 2020, in preparation.

R. Schwarte, Z. Xu, H. Heinol, J. Olk, and B. Buxbaum.
New optical four-quadrant phase-detector integrated into a photogate array for small and precise 3D-cameras.
SPIE Proc. Vol. 3023, p. 119, 1997, http://dx.doi.org/10.1117/12.269749.

H. Rapp.
Experimental and Theoretical Investigation of Correlating TOF-Camera Systems.
Diplomarbeit, Universität Heidelberg, 2007,
http://www.ub.uni-heidelberg.de/archiv/7666

Literatur III

M. Beer.
SPAD-basierte Sensoren für die laufzeitbasierte Distanzmessung bei hoher Hintergrundlichtintensität.
Diissertation, Universität Duisburg-Essen, 2018,
https://duepublico2.uni-due.de/receive/duepublico_mods_00047708.

J. Chang.
New Sony DepthSense(TM) ToF Technology.
https://ibv.vdma.org/documents/256550/27019077/2018-11-07_Stage1_1630_LUCID_Advanced+Material+Handling+with+Sony+Depthsense.pdf/.