Exploring new avenues in motion picture production: A groundbreaking RGB+Z camera demands cutting-edge TOF technology

Johannes Steurer

Ein Gedi, 10 March 2014

A Filmmaker's Vision

Capture depth channel together with image stream

Derive mask signals from the depth map for post production

- Substitute green screens,
- Artificial shallow depth of field

Build moving 3D model of the real scene for visual effects

- e.g. shadow casting, interactions between real and virtual characters
- Modify real captured scenes as easily as CGI models: Substitute textures, warp shapes, etc.
- "Object oriented image processing"

Filmmaker‘s reality

Today

VFX Production Today

VFX Production Today

VFX Production Today

Film Producers Vision

With RGB+Z camera

With RGB+Z Camera – Depth Map

With RGB+Z Camera

With RGB+Z Camera

With RGB+Z Camera

With RGB+Z Camera

VFX Applications

- **Selective Color Grading** – with depth map as segmentation mask
- Modifying optical characteristics like **shallow depth of field** after the fact – with depth map as key for blurriness (objects farther from a distinct z-plane are more blurry)
- **Depth-Keying** – depth map as mask for foreground – background distinction
- **3D-Interaction between real and virtual actors** and other scene elements, like casting shadows – generate 3D models from depth map
- **Virtual lighting** of real scenes – depth map generated 3D models casting shadows on real and virtual backgrounds
- **2D → 3D Conversion** – measured depth maps to start with rather than a hand-painted (invented) depth map
- **Auto-stereoscopic displays** - calculate intermediate camera perspectives by means of depth based warping

Camera Applications

- **Focus Assistance** – measure the distances of actors and objects to support the focus puller

- **Follow Focus** – measure the distance of objects, follow the objects by image tracking adjust focus of lens

- **Self localization** of the camera – position and orientation in the set environment

Requirements for depth imaging apps.

Application	Requirements				
	Z - Range	Z - Accuracy	XY-Resolution	Field of View	Latency
Color Grading	medium	medium	medium	Full FOV	n.a.
2D → 3D conv	medium	medium	medium	Full FOV	n.a.
Depth of Field	medium	high	medium	Full FOV	n.a.
Depth keying	medium	medium	high	Full FOV	n.a.
3D composite	high	high	high	Full FOV	n.a.
Focus Assist	high	high	medium	Reg of Interest	low
Follow Focus	high	high	medium	Reg of Interest	realtime

EU-funded Research Project SCENE

SCENE: **Novel scene representations for richer networked media**

Duration: 11/2011 – 10/2014 (3 years)

Goal: **Novel Scene representation (SRA)**

- *Manipulate real captured content like CGI*
- *Give CGI content natural emotion like real capture*

Consortium:

ibbt connect.innovate.create

Centre d'innovació Barcelona Media

Fraunhofer Heinrich Hertz Institute

3DLIZED

BRAINSTORM

ARRI

intel VISUAL COMPUTING INSTITUTE

technicolor

UNIVERSITY OF SURREY

SEVENTH FRAMEWORK PROGRAMME

Motion Scene Camera (MSC)

- Based on an ARRI ALEXA Studio camera
- Extended by depth sensing capabilities using a Time-of-Flight module
- Synchronous capture of RGB and Z image data
- Complemented by VFX meta data like lens data (aperture, focus, focal length)

RGB

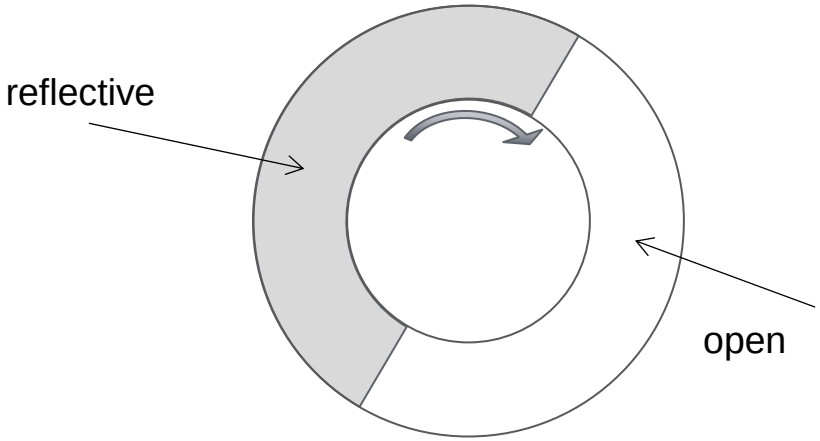
ARRI ALEXA Studio

+

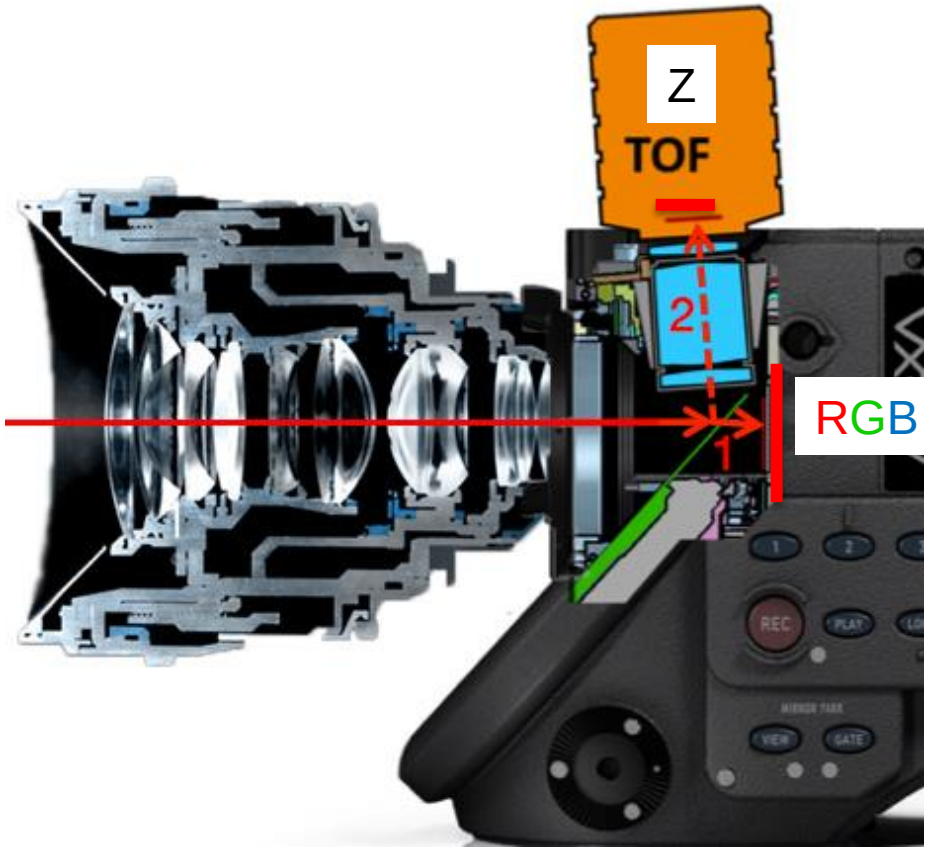
Depth → Z

Time-of-Flight Depth Camera

Opto-mechanical Integration



Rotating mirror shutter



Alexa Studio

Image resolution	2880 x 2160 pixel
Active area	23.76 x 17.82 mm
Frame rate	1 – 60 fps
Sensitivity	EI 160 to EI 3200
Exposure latitude	14.5 stops
Rotating mirror shutter	

Time-of-Flight Camera

Active illumination & CMOS sensor

→ “Photo mixer device – PMD”

- NIR-LED radiates light pulse
- CMOS sensor measures round travel time of light
- 1 nsec ≙ 70 mm

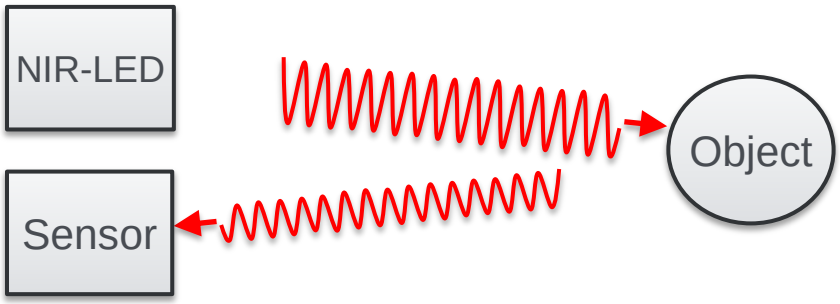


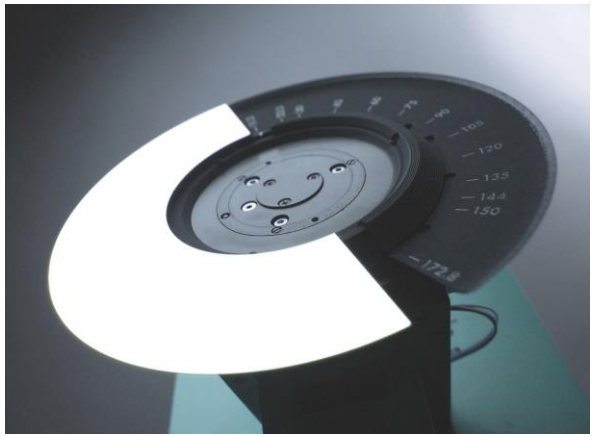
Image resolution	352 x 288 pixel
Illumination Units	870 nm (NIR)
Depth range	7 m
Depth noise	~80 mm

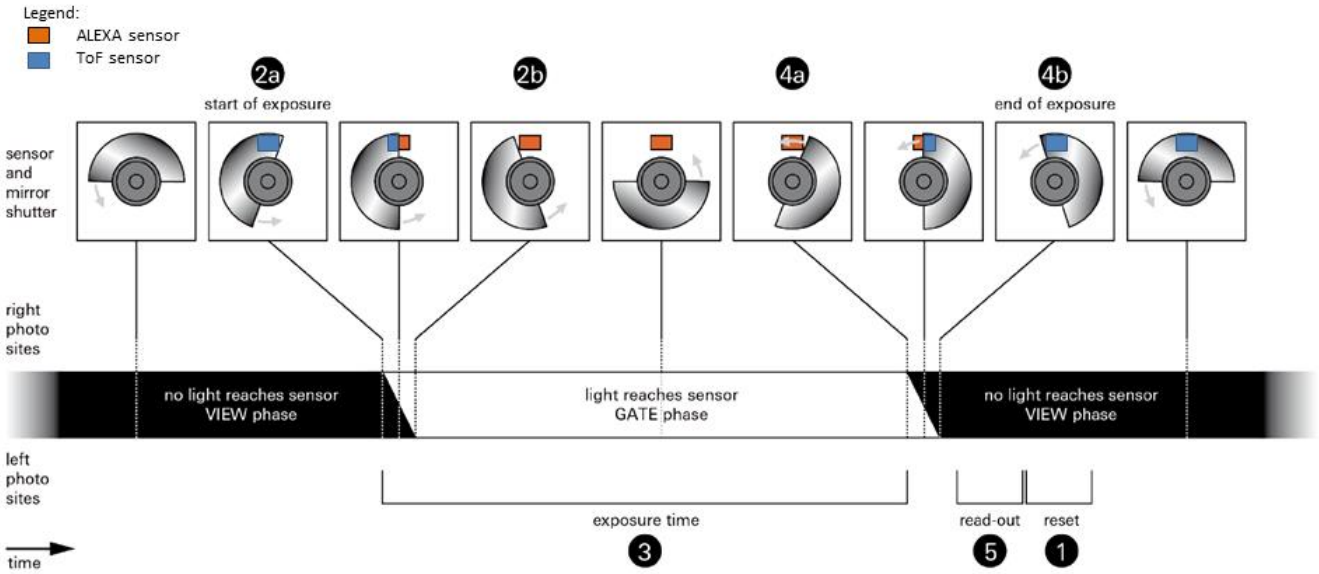
Motivation

Why using a Time-of-Flight camera for depth measurement?

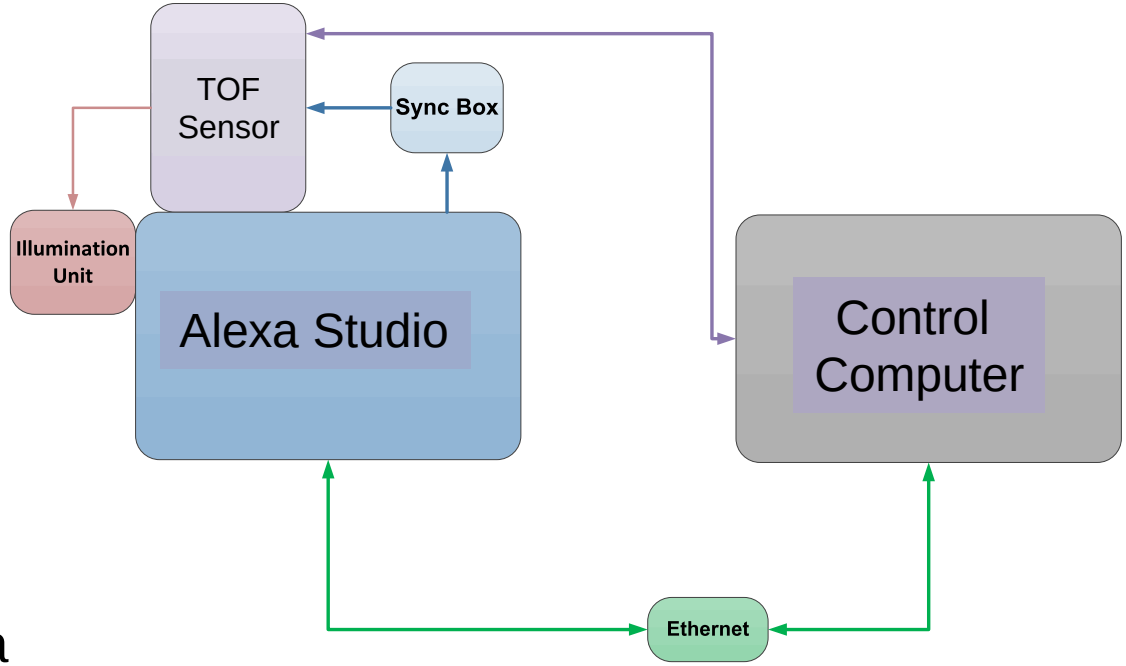
- Independent depth information for each (ToF) sensor pixel
- Technology suitable for real-time applications
- Low computational load for depth measurement
- High frame rates
- Competes with triangulation approaches due to contrary requirements (e.g. no contrast needed)
- Systematic depth error: ToF ($\Delta Z \sim$ const.) vs. Stereovision ($\Delta Z \sim Z^2$)
- Small size and light-weight

Mirror Shutter Details





Hardware Integration



Important system elements:

- ALEXA Studio RGB camera
- PMD Time-of-Flight camera
- NIR illumination units
- Synchronization module (Sync Box)
- Control computer for TOF Sensor

  Sensor-control, preview, processing, storage

Motion “Scene“ Camera - Prototype

- ✓ ALEXA camera enabled to capture additional depth information
- ✓ Occlusion-free RGB+Z using the same entrance pupil
- ✓ Synchronous capture
- ✓ Prototype used in field-testing
- ✓ Early stage of research

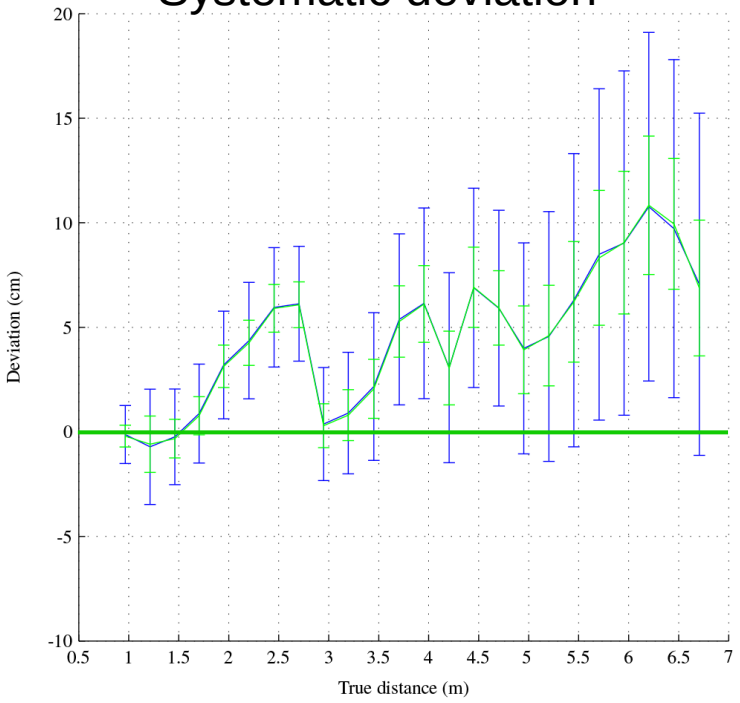
Ongoing research

- ➢ timing and frame rate
- ➢ light distribution of active illumination
- ➢ depth range and resolution
- ➢ denoising and upsampling of low-res depth maps

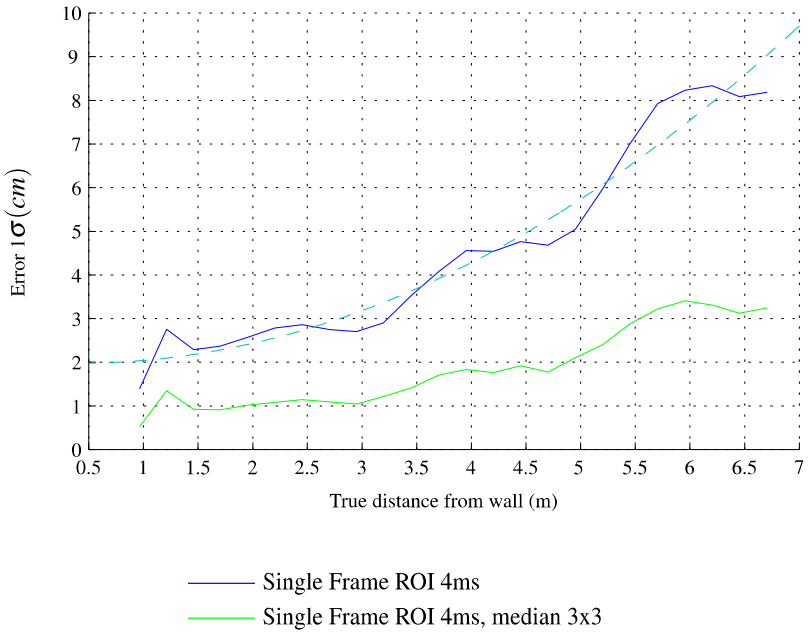
Measurement Errors

- Systematic “wiggling“ around real distance
- 8 cm within the range of 7.5 m (high NIR reflectance)
- Statistical error due to:
  - Sensor noise
  - Decreasing active light intensity with increasing distance an object

Systematic deviation

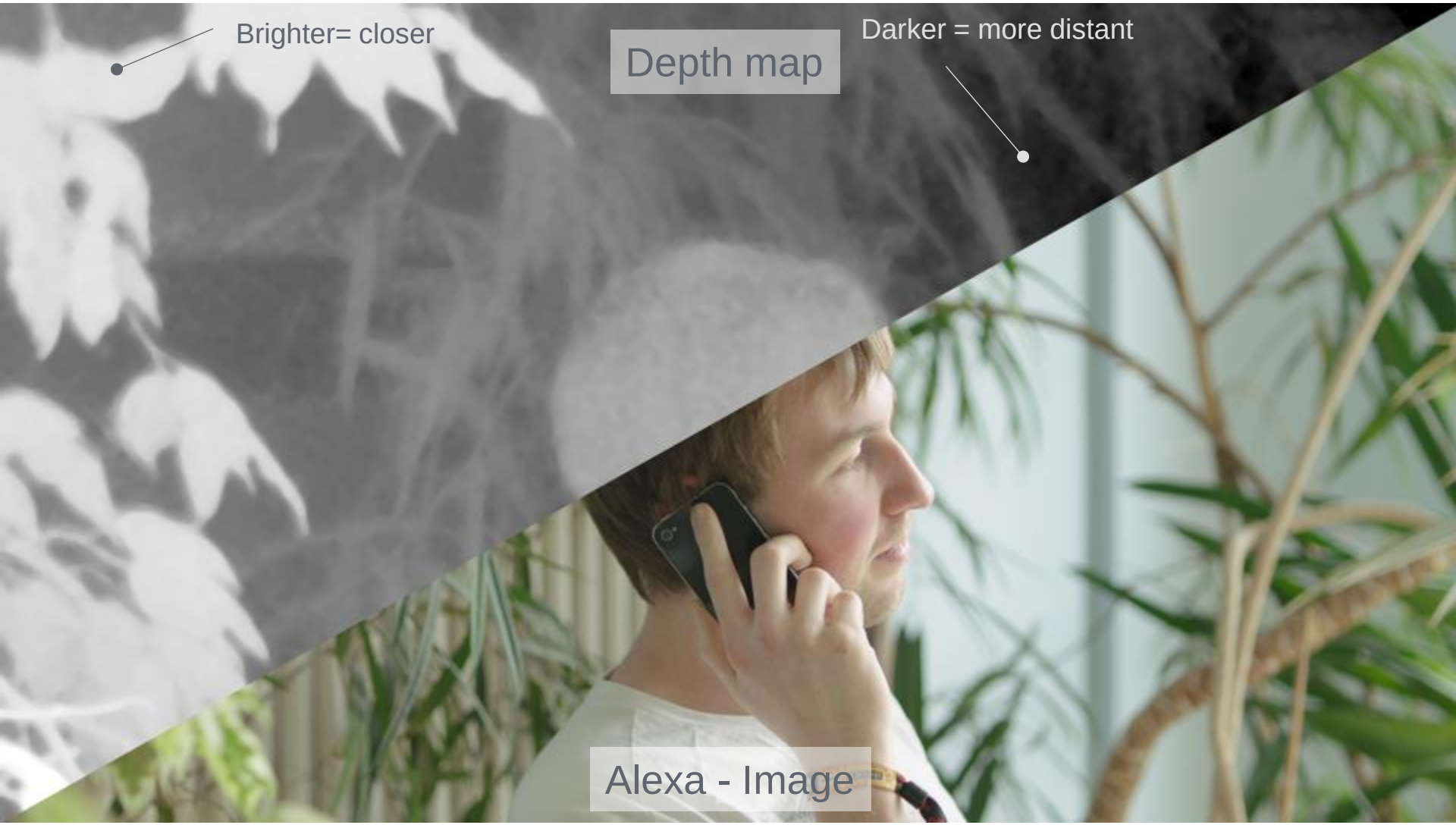


Depth noise

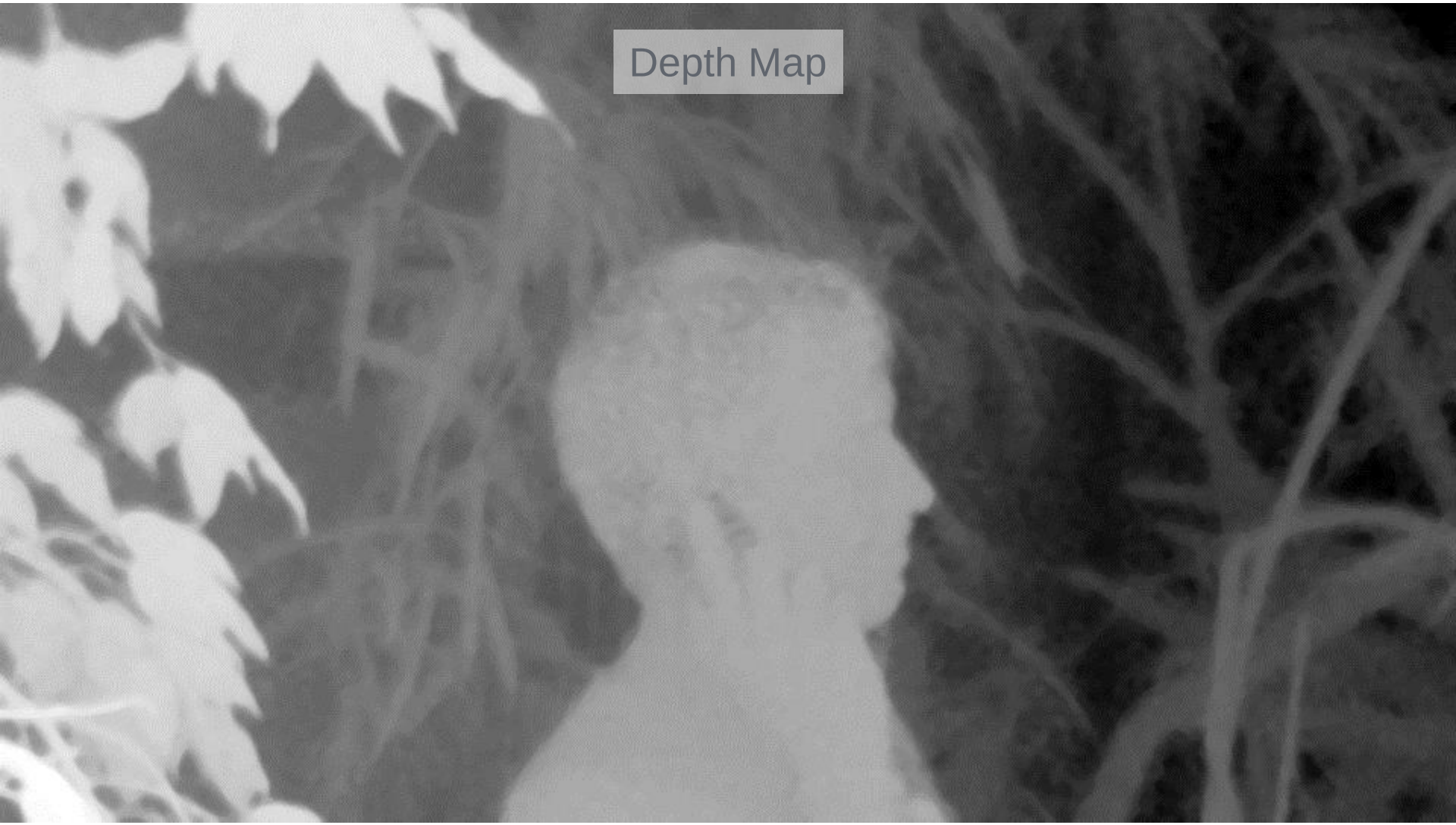


Sample 1

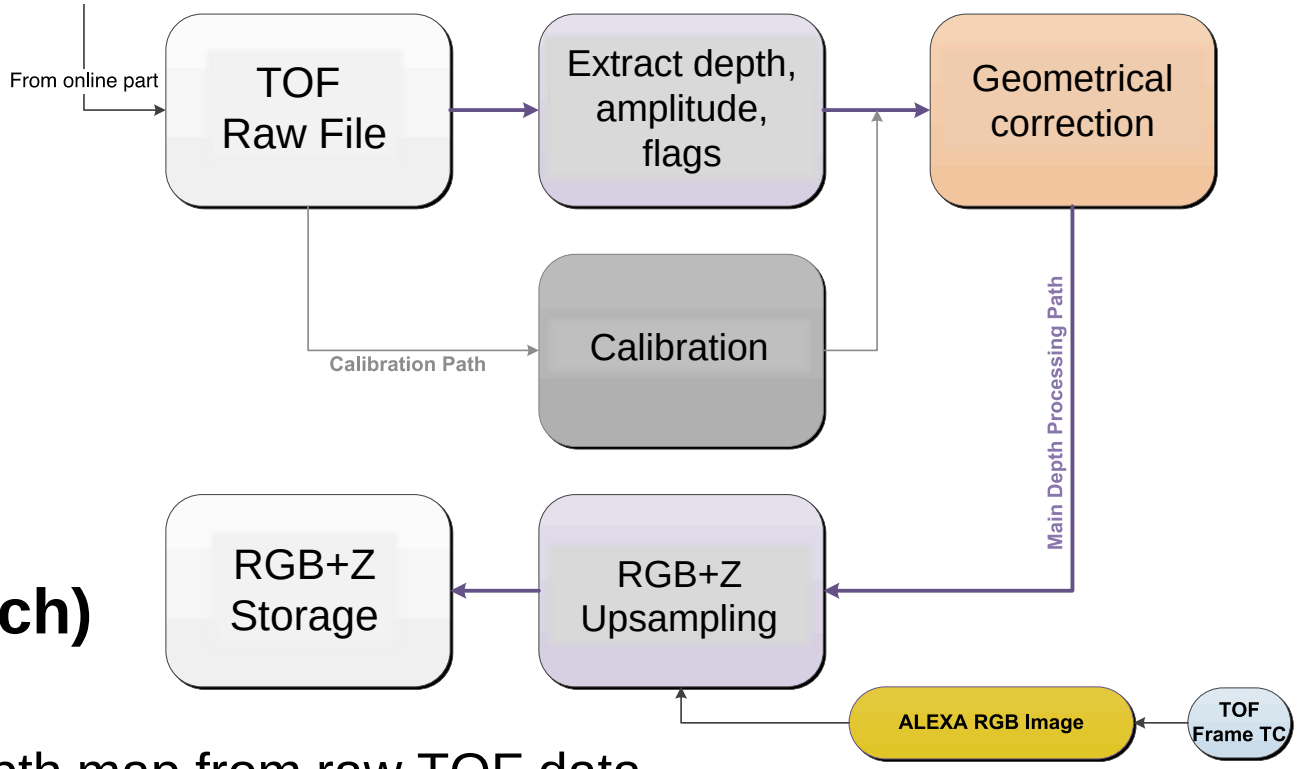
Sample 1



Sample 1



Signal Processing and Data Fusion

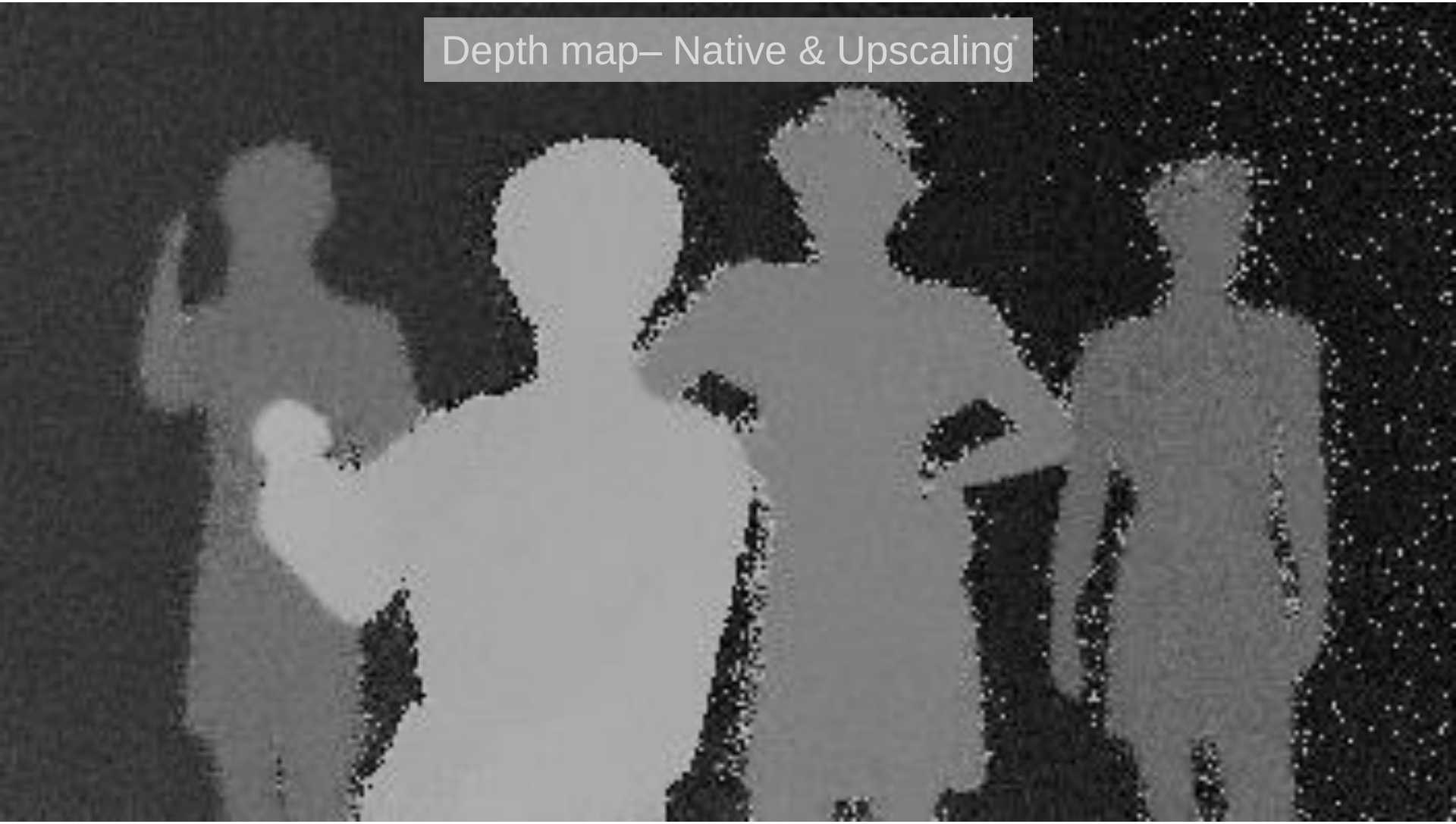


R &D platform (algorithm research)

- Extraction of the depth map from raw TOF data
- Geometrical calibration of depth map
- De-noising and refining of depth
- RGB-aided up-scaling of depth
- Data storage in a suitable container (2-layer-representation in OpenEXR)

Sample 2 Mannequins

Sample 2 Mannequins



Sample 2 Mannequins

Sample 2 Mannequins

Sample 2 Selective Color Grading

Sample 2 Selective Color Grading

Sample 2 Selective Color Grading

Sample 2 Selective Color Grading

Sample 2 Selective Color Grading

Challenges for depth sensing technologies

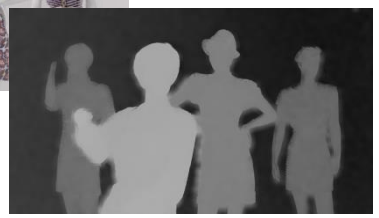
Current limitations and required parameters

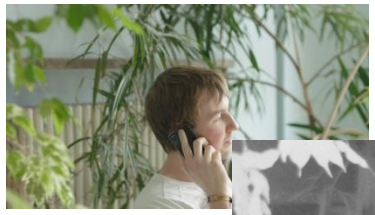
Parameter	Current	Production ready	Optimum
Pixel resolution	350 x 290	1k x 0,5k	4k x 2k
Depth accuracy	0,1m	0,02m	<0,01m
Depth noise	0,1m	0,02m	0,01m
Range	<5m	10m	>10m
Frame rate	8fps	60 fps	200 fps
Multiple cameras on set	1	5	20

➔ **More, faster, better depth-pixel, please!**

Opportunities

Sensor size is not a limit – could be larger
Modulation type – handle interferences, and non ambiguity range
Joint filtering of RGB + Z to mutually improve signals





Acknowledgement
Part of the research leading to these results has received funding from the European Commission's Seventh Framework Programme under the project SCENE.

Thank You!