ETN-FPI TS3 “Plenoptic Sensing”

3D-TV LDV Content Generation with a Hybrid ToF-Multi-Camera Rig

Anatol Frick, Bogumil Bartczak and Reinhard Koch

Multimedia Information Processing
Institute of Computer Science
Christian-Albrechts-University of Kiel

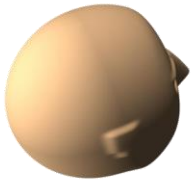
Outline

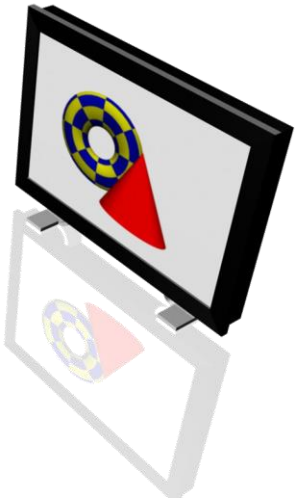
- Introduction
- Content Representation (Depth Estimation, LDV)
- Camera Setup and System Overview
- LDV Generation
- Conclusions

Introduction to 3D-TV

- **3D ?**
  - Minimum of 2 different views
  - Separate Views (left / right eye)

One view

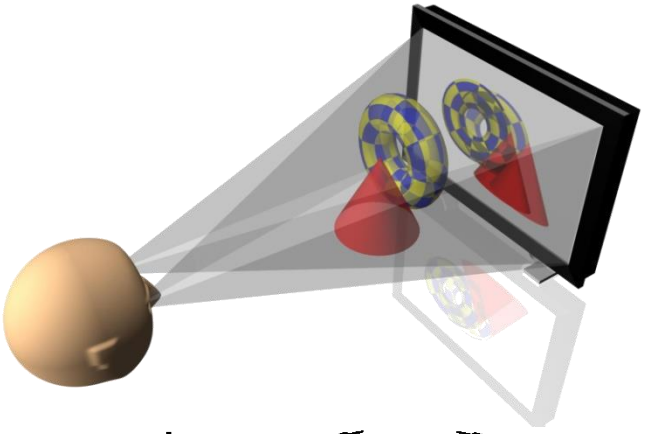




2D

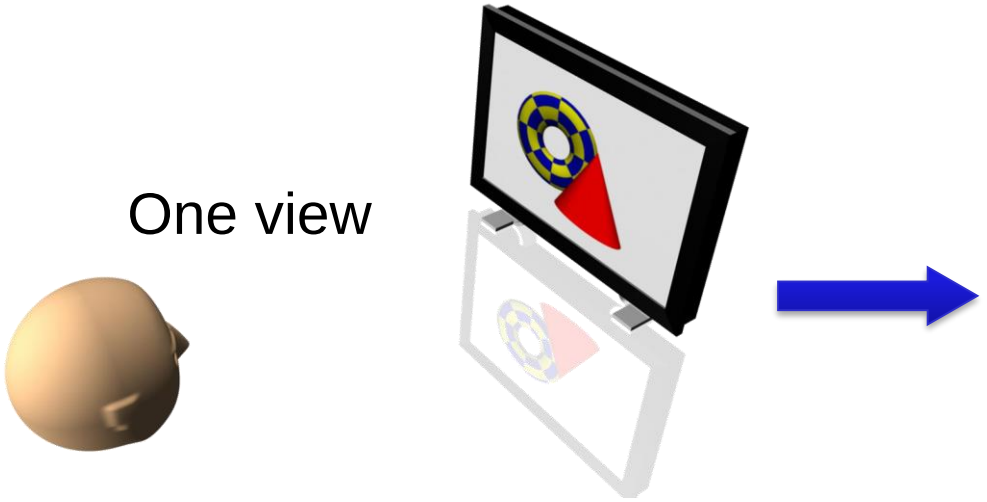
Introduction to 3D-TV

- **3D ?**
  - Minimum of 2 different views
  - Separate Views (left / right eye)
  - Stereoscopic (two views)



Two views

One view

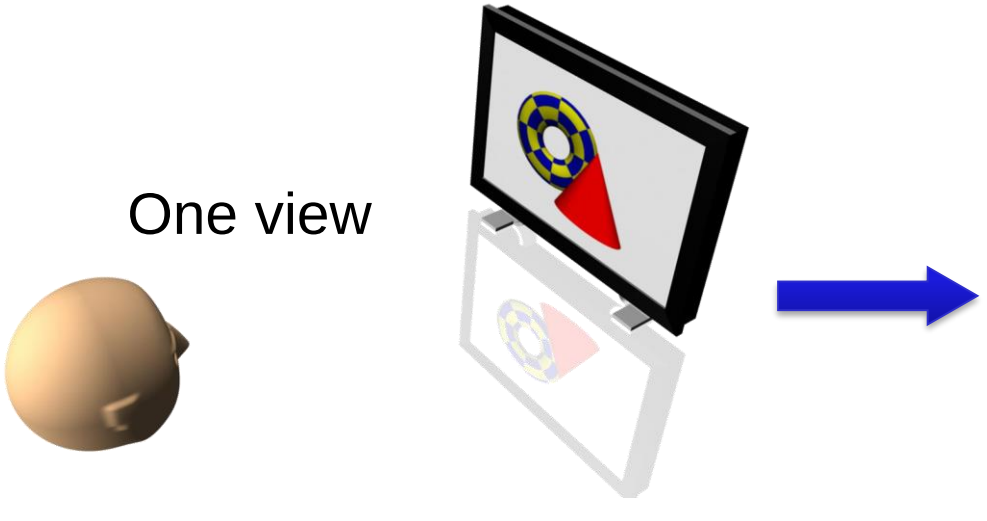


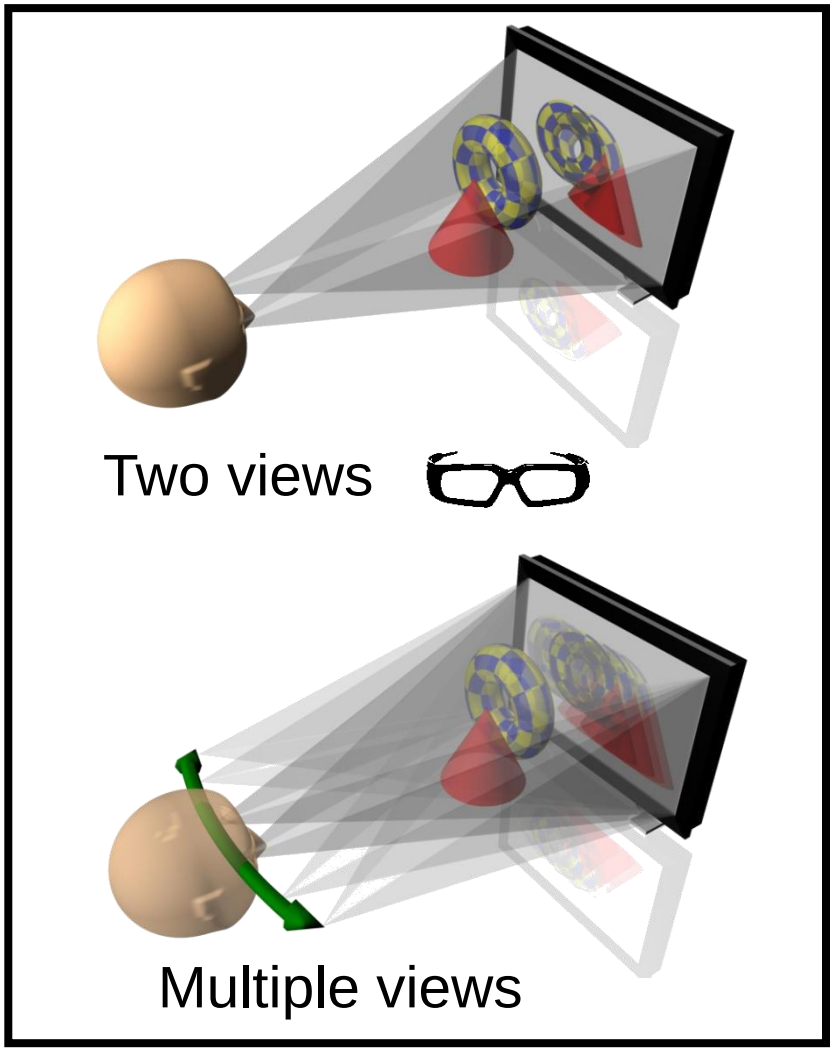
2D

3D

Introduction to 3D-TV

- **3D ?**
  - Minimum of 2 different views
  - Separate Views (left / right eye)
  - Stereoscopic (two views)
  - Auto-stereoscopic (multiple views)





Introduction to 3D-TV

- **Goal : Multiview Displays**
  - No glasses
  - Look-around capability
  - Content can also be shown on standard stereo displays

Introduction to 3D-TV

- **Goal : Multiview Displays**
  - No glasses
  - Look-around capability
  - Content can also be shown on standard stereo displays

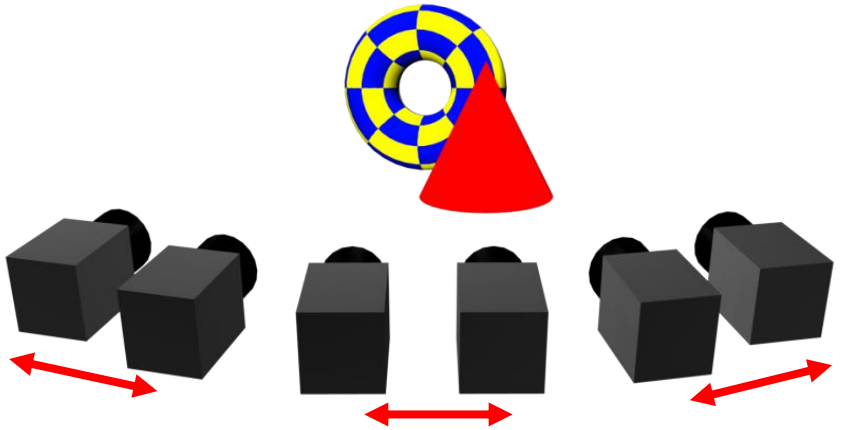
- **Required :**
  - Multiple cameras
  - Suitable content representation

Content Representation

- **Image Based**
- **Depth Image Based**

Content Representation

- **Image Based**
  - Store image for each camera (n cameras, n video streams)
- Shortcomings:
  - 3D effect depends on camera arrangement
  - Number of views depends on the number of cameras
  - Much data (storing and transmitting ?)
  - Inflexible format (display geometry, post processing)



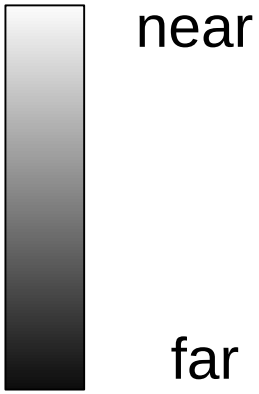
Baseline ?

Content Representation

- **Depth Image Based**
- Additional information describing geometry (depth image)

Color image

Depth image



- Number of views independent of the number of cameras
- 3D effect is largely independent of camera arrangement
- Flexible format (content can be modified in post processing)

Content Representation

- **Depth Image Based**
- Additional information describing geometry (depth image)

Content creation challenging !

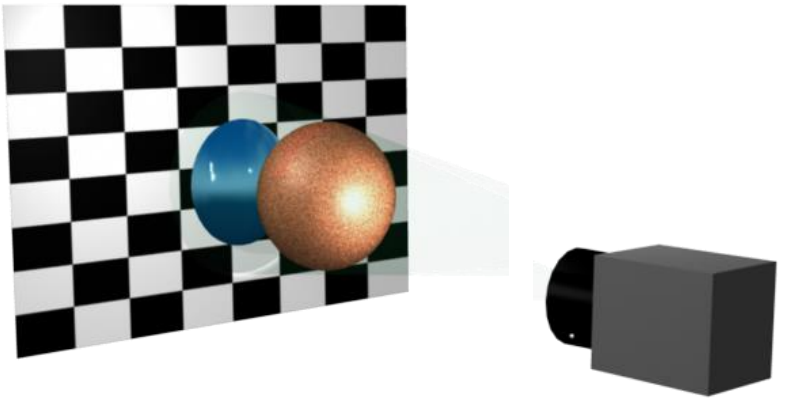
Color image

Depth image

- Number of views independent of the number of cameras
- 3D effect is largely independent of camera arrangement
- Flexible format (content can be modified in post processing)

LDV Format

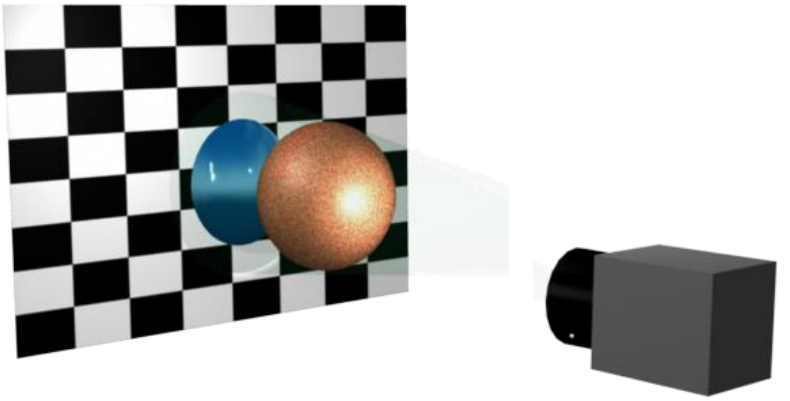
- **Layered Depth Video (LDV)**
  - Reference camera



Reference camera

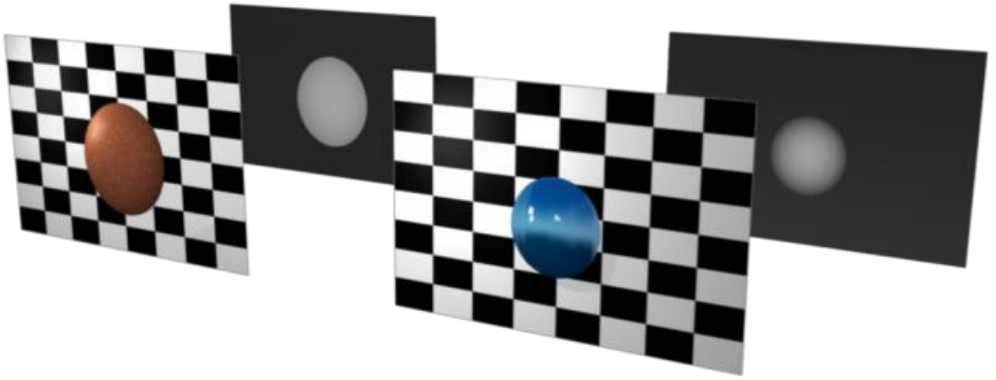
LDV Format

- **Layered Depth Video (LDV)**
  - Reference camera
  - Foreground Layer (scene seen from reference camera)
  - Occlusion Layers (scene not seen from reference camera)



Reference camera

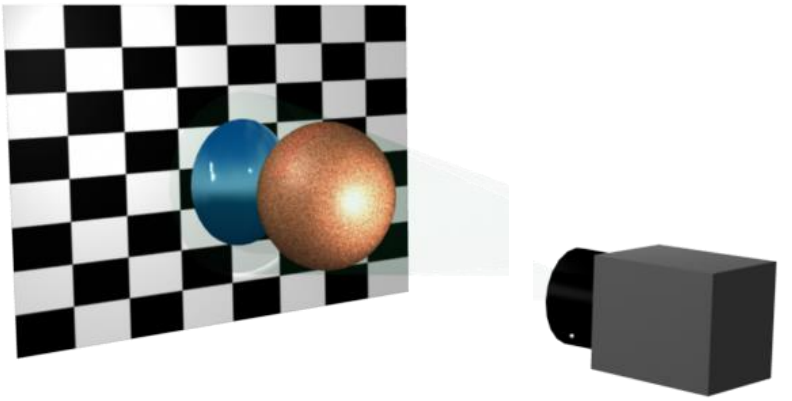
Foreground Layer Occlusion Layer



LDV-Frame (2 Layers)

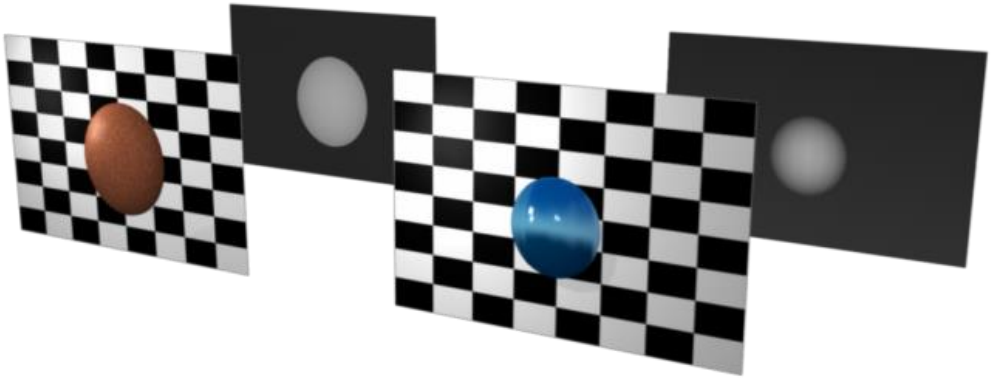
LDV Format

- **Layered Depth Video (LDV)**
  - Reference camera
  - Foreground Layer (scene seen from reference camera)
  - Occlusion Layers (scene not seen from reference camera)

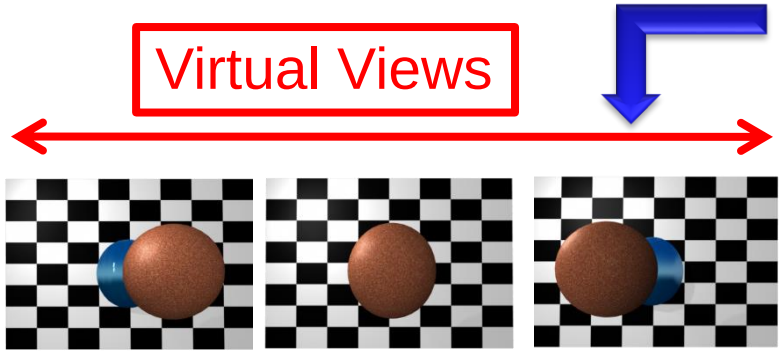


Reference camera

Foreground Layer Occlusion Layer



LDV-Frame (2 Layers)



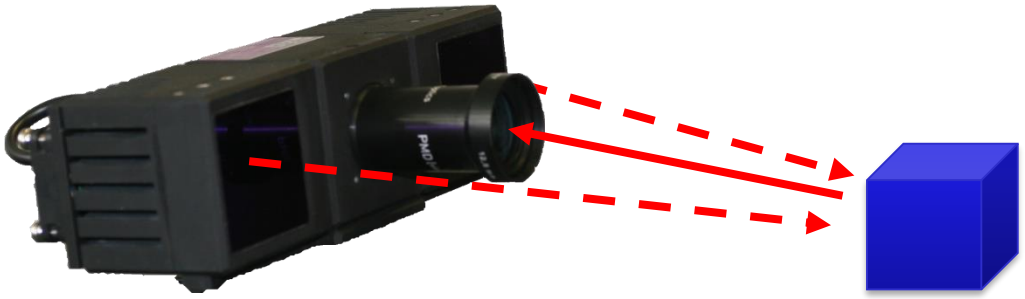
Left Central Right

Depth Estimation

- **Stereo Matching**
  - \- Slow for acceptable quality
  - \- Not reliable in texture-less regions or in presence of repetitive patterns
  - \+ High resolution 1920x1080 (fine details)
- **Time of Flight Cameras (ToF)**
  - \+ Active depth measurement (does not rely on texture)
  - \+ Fast (25 Frames per second)
  - \- Low resolution max. 204x204 (missing fine details)
  - \- Limited range : 1m – 7.5m (indoor scenarios)
  - \- Bad signal to noise ratio

Depth Estimation

- **Time of Flight Camera (ToF)**



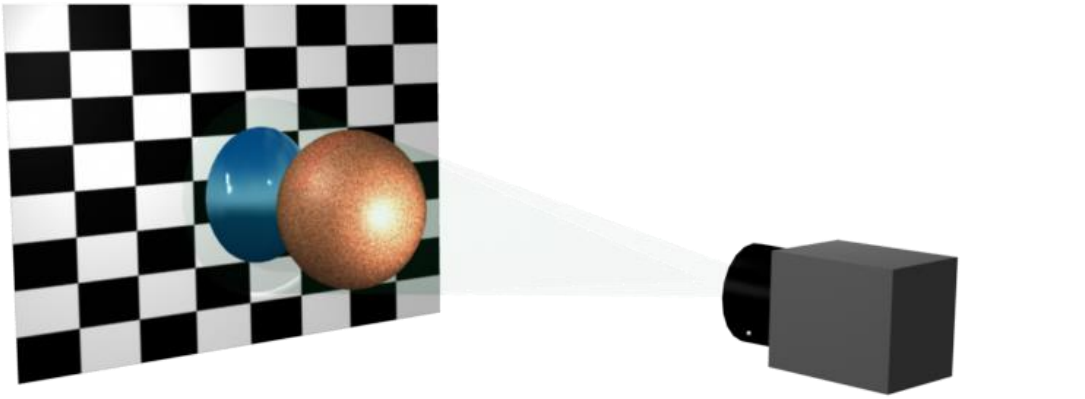
PMD CamCube

- Emits modulated infrared light (18, 19, 20, 21 MHz)
- Phase difference (per pixel) between incoming and out coming light gives distance to object

204x204 px

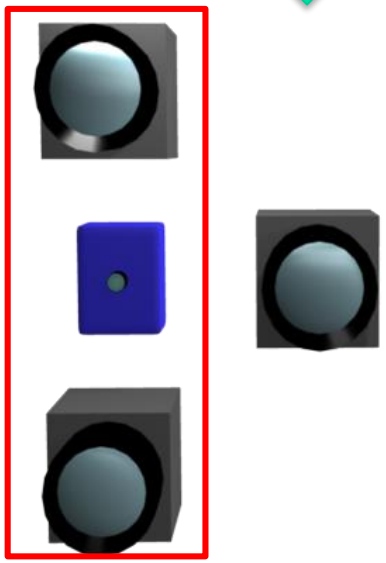
Camera Setup

- **Camera Setup for LDV-Generation**
  - Additional Cameras
  - ToF Cameras



Reference camera

Module Left



Module Right

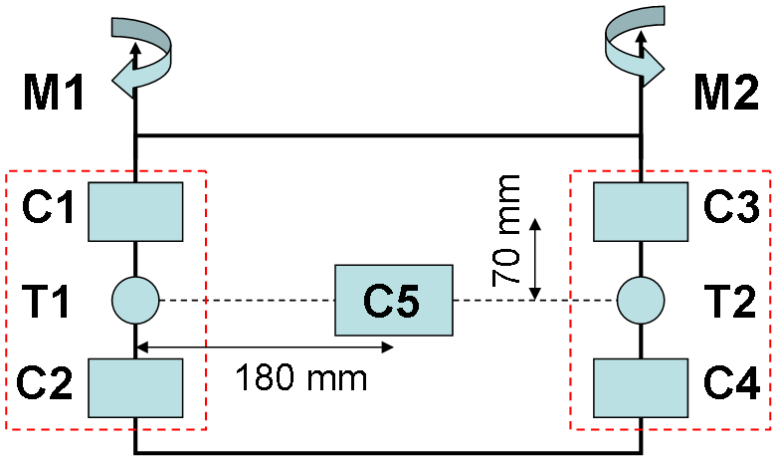
Color Camera Top

Time of Flight Camera

Color Camera Bottom

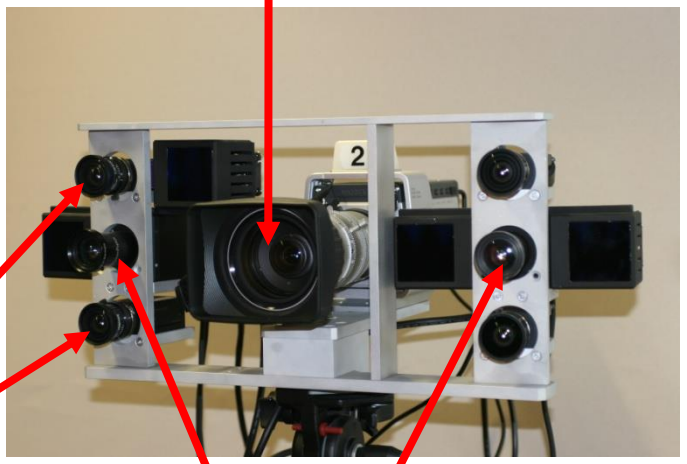
Camera Setup

- **Proposed camera setup (5+2)**
  - Target Format: LDV for central camera
  - Modular (Modules: M1 and M2; Central Camera)
  - Adjustable (Modules M1 and M2 are rotational)



Schematic representation

Central HD-Camera



Additional Color Cameras

Time of Flight Cameras

Camera Setup

- **Proposed camera setup (5+2)**
  - Optimal central view coverage (Module rotation)

(16:9) (1920x1080)

(1:1)

(1:1)

PMD CamCube (204 x 204)

System Overview

- **Image Processing Chain**

(25 F / s)

(1) Capturing

- 1920x1080 YUV
- (4) 1600x1200 Bayer Pattern
- (2) Depth Images (PMD)

System Overview

- **Image Processing Chain**

(25 F / s)

(410 MB / s)

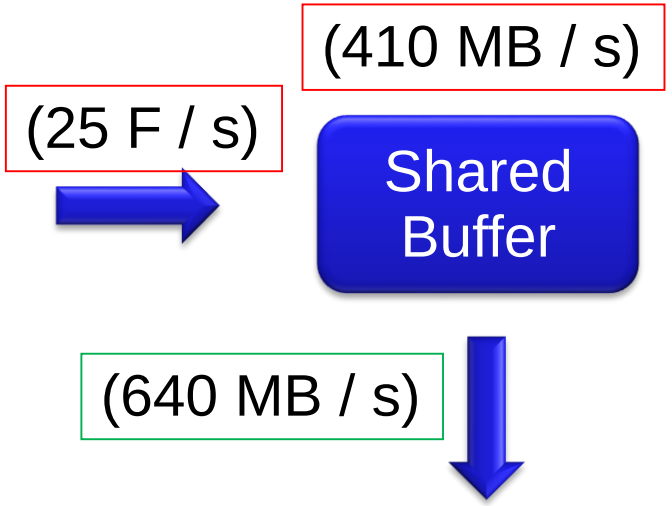
Shared Buffer

(1) Capturing

- 1920x1080 YUV
- (4) 1600x1200 Bayer Pattern
- (2) Depth Images (PMD)

System Overview

- **Image Processing Chain**



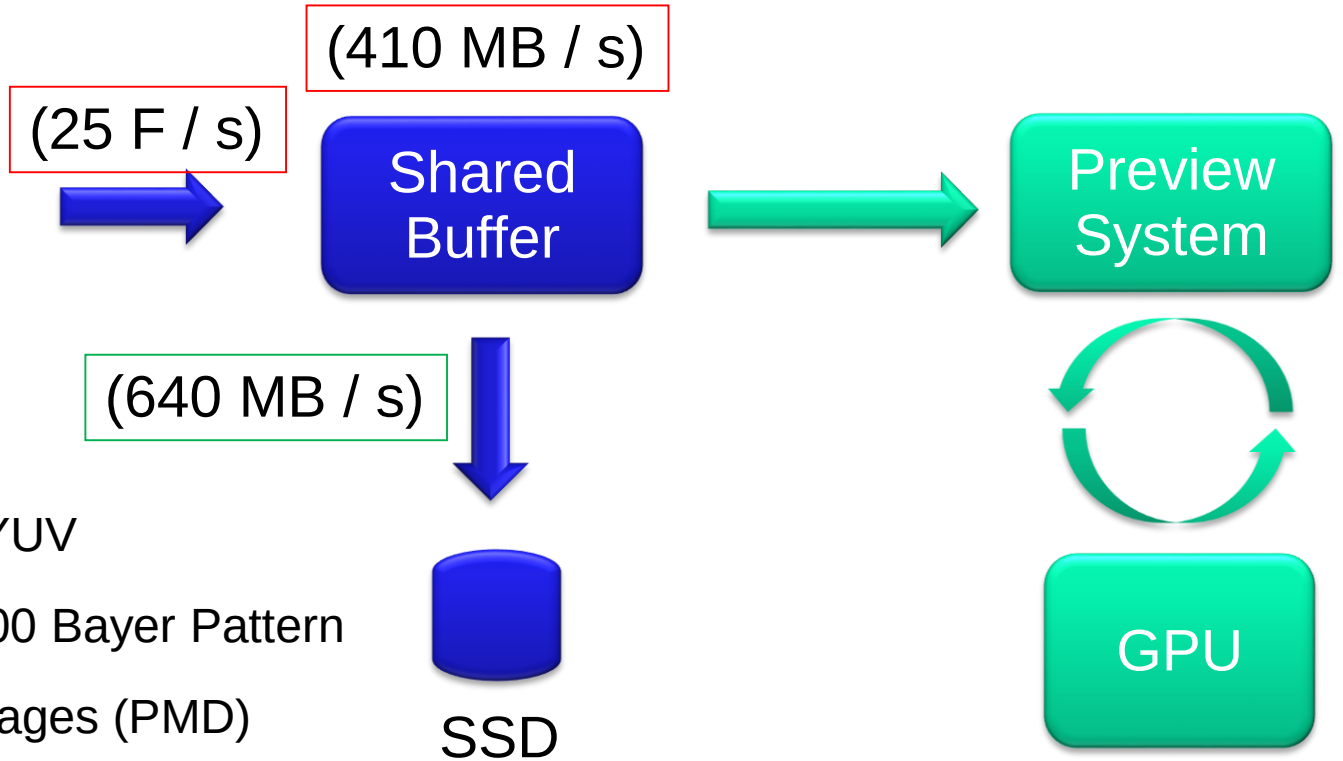
(1) Capturing

- 1920x1080 YUV
- (4) 1600x1200 Bayer Pattern
- (2) Depth Images (PMD)

(2) Store to disk

System Overview

- **Image Processing Chain**



(1) Capturing

- 1920x1080 YUV
- (4) 1600x1200 Bayer Pattern
- (2) Depth Images (PMD)

(2) Store to disk / Preview

Camera calibration parameters

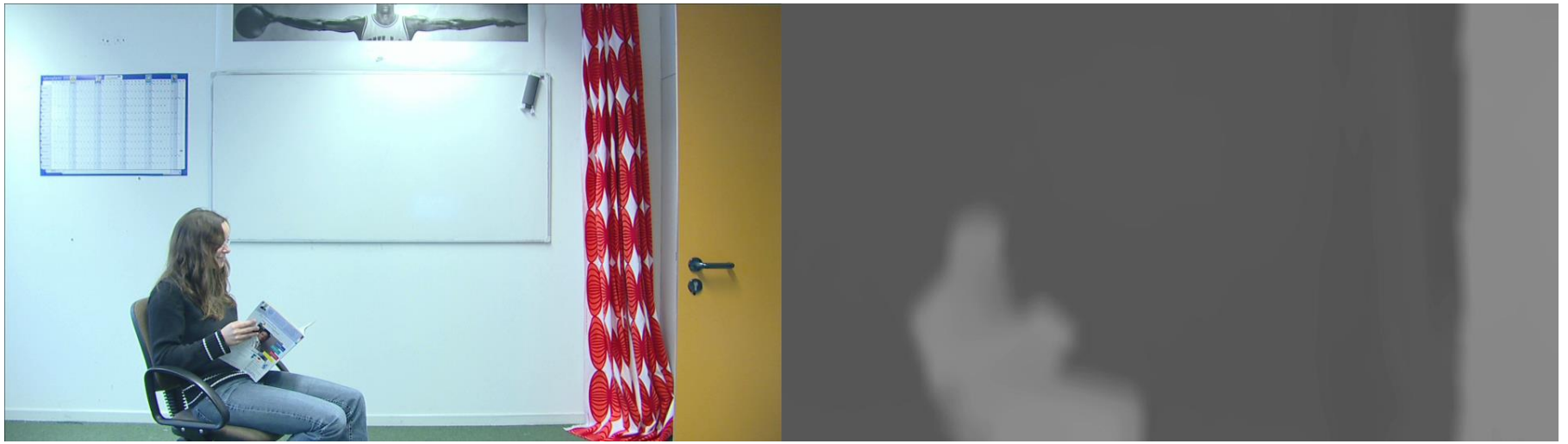
Real - Time Preview

- **Real-Time LDV Generation**
  - Only Foreground Layer, no occlusion information
  - Reference image and 2 ToF images at 25 Frames / s
  - 3D Warping to Reference View
  - Filtering

3D-Warping to the Reference View

Real - Time Preview

- **Real-Time LDV Generation**
  - Only Foreground Layer, no occlusion information
  - Preview on auto-stereoscopic display from Philips (9 views)
  - 25 Frames / s



LDV – Foreground Layer

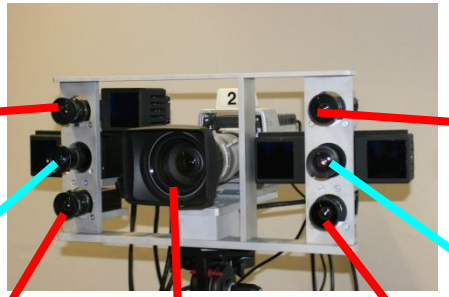
Real – Time Preview

- **Live Demo**

3D - Preview

LDV Generation

Calibrated Camera Setup

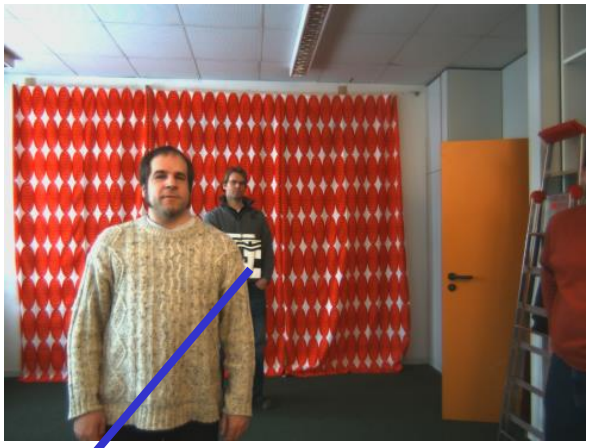




Reference view

LDV Generation

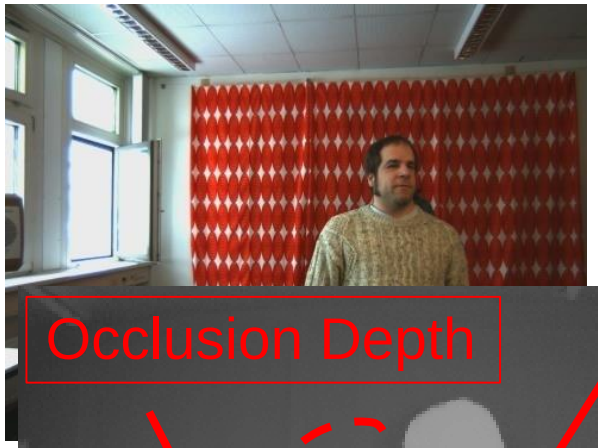
Calibrated Camera Setup

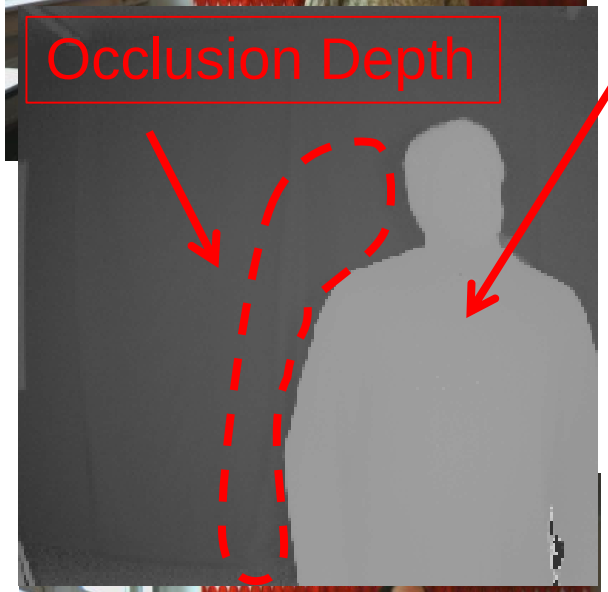


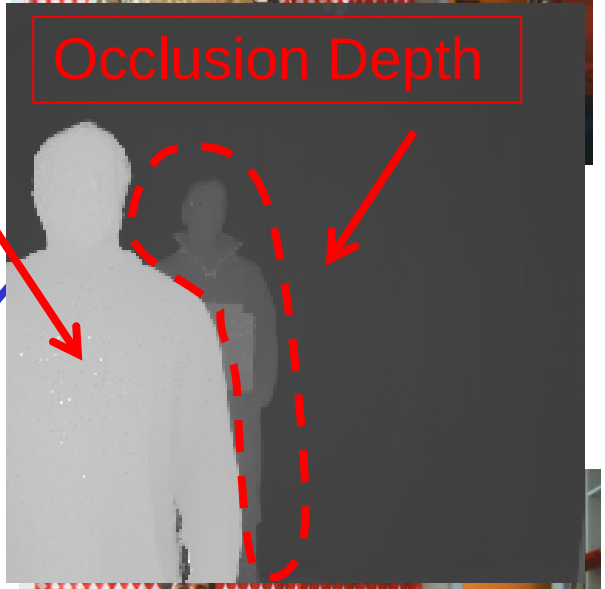
Occlusion
Color

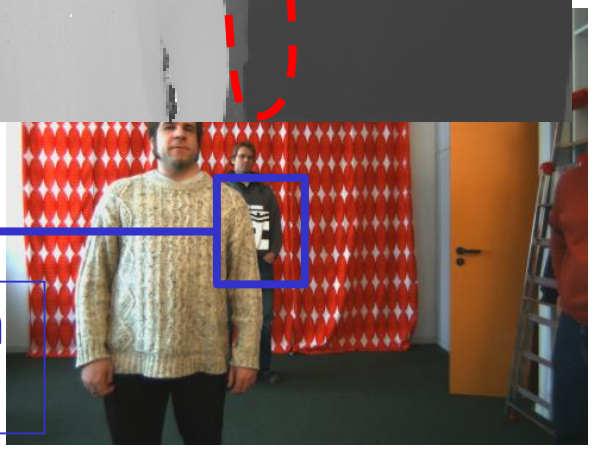
LDV Generation

Calibrated Camera Setup

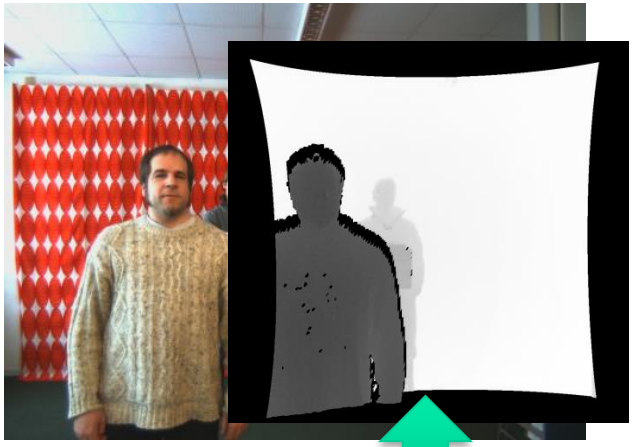






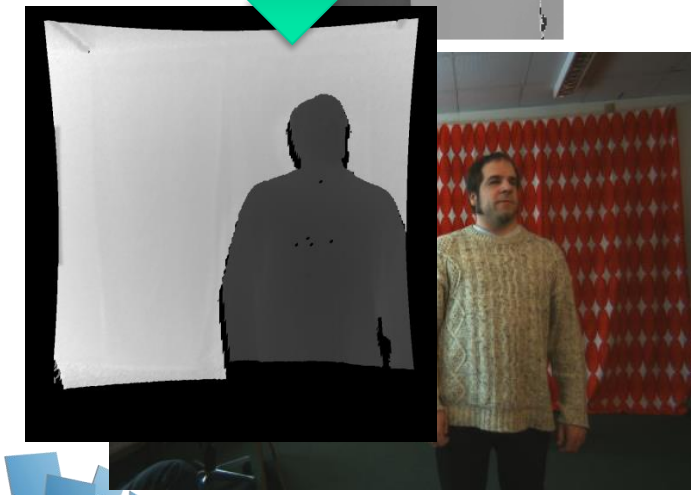


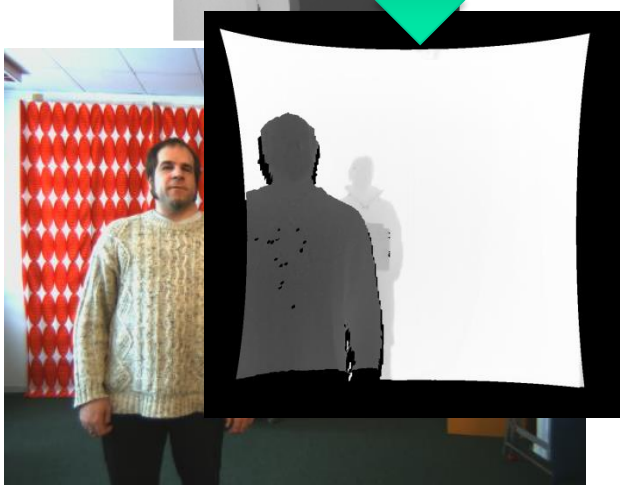
LDV Generation



3D Warping

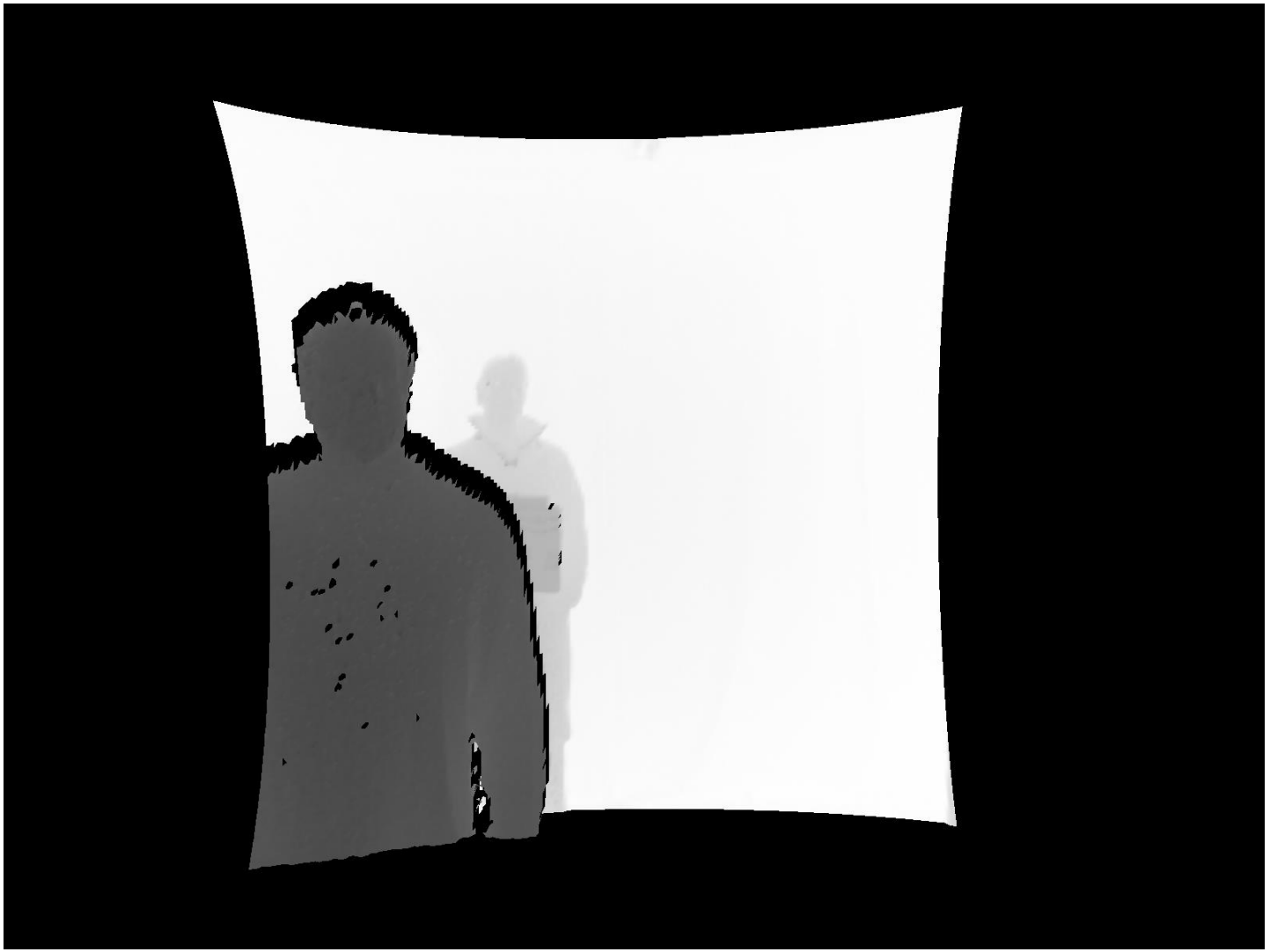
3D Warping





First: Warp to corners

LDV Generation



Right ToF warped to Right Top Camera

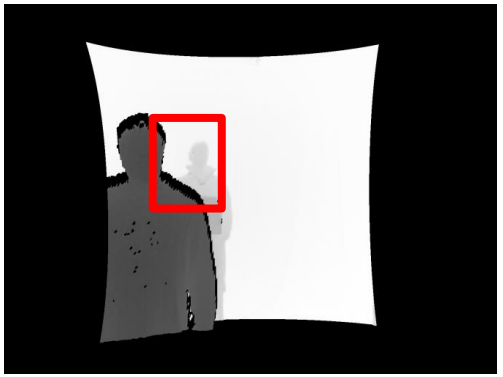
Problems (Warping Artifacts):

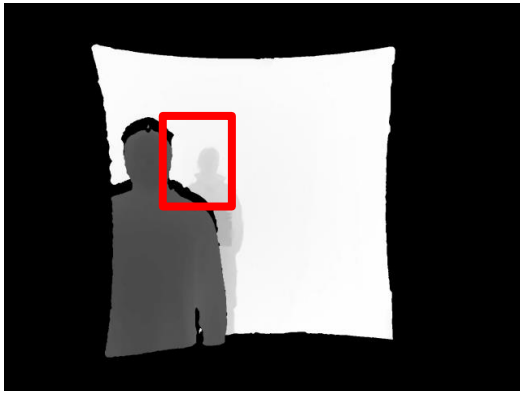
- due to view point change, noise, occlusions
- and resolution difference (1600 x 1200 vs. 204 x 204)

LDV Generation

Color Alignment

- Cost volume (distance costs)
- Iterative refinement (based on Bilateral Filtering)

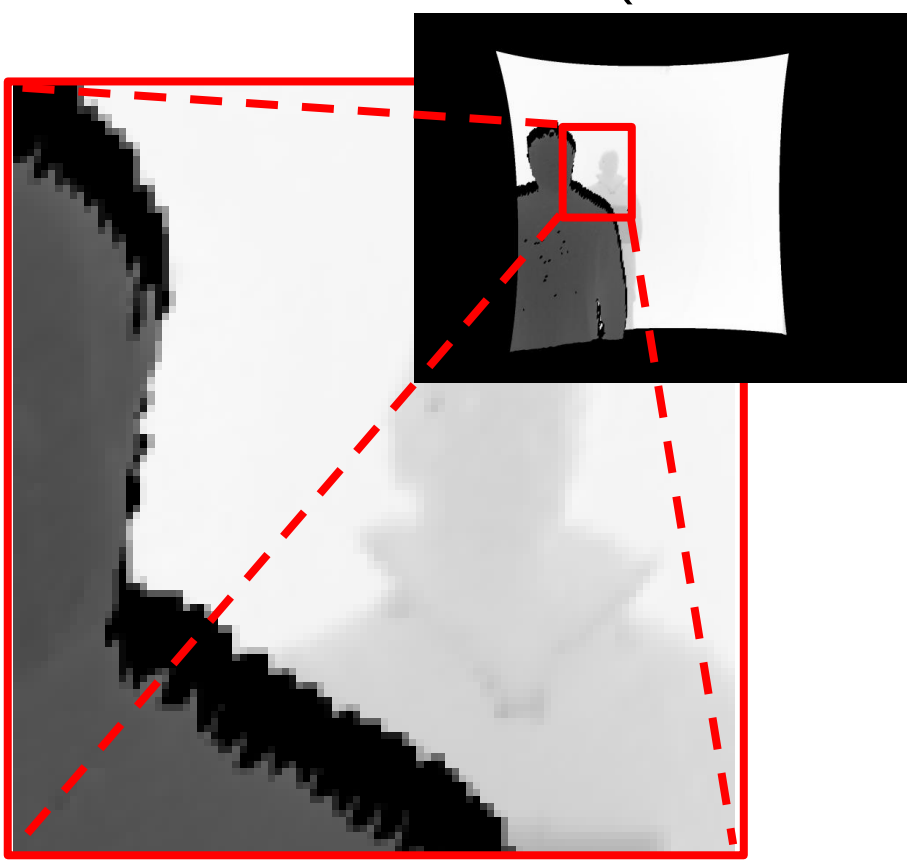


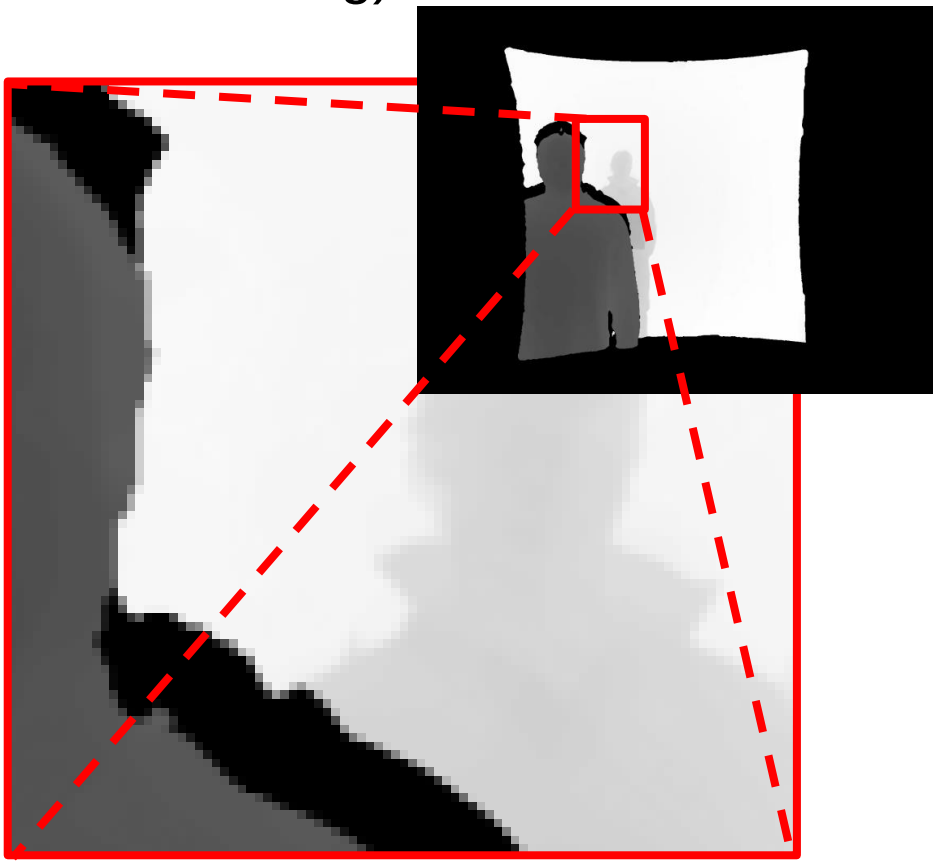


LDV Generation

Color Alignment

- Cost volume (distance costs)
- Iterative refinement (based on Bilateral Filtering)



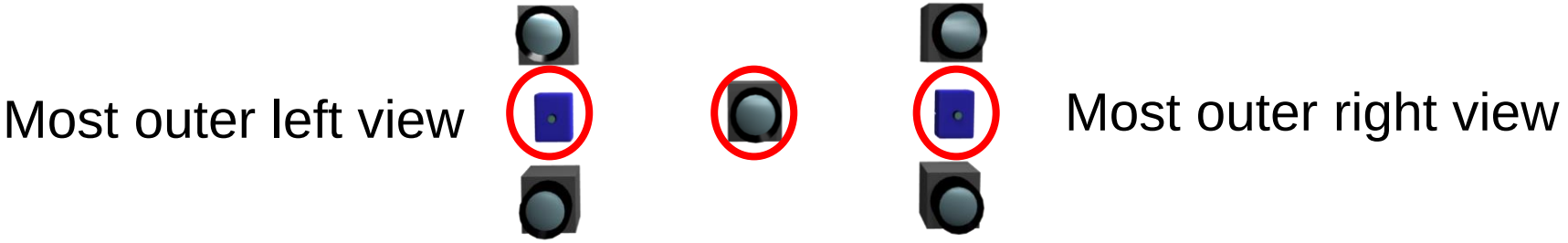


LDV Generation

- Generation of three views on a baseline

LDV Generation

- Generation of three views on a baseline

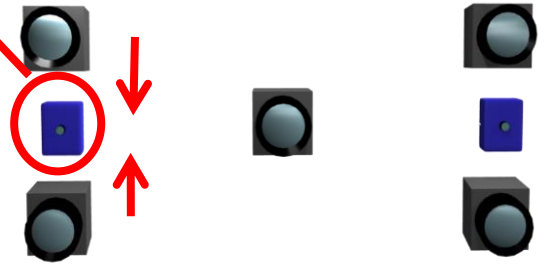


LDV Generation

- Generation of three views on a baseline

  Left virtual

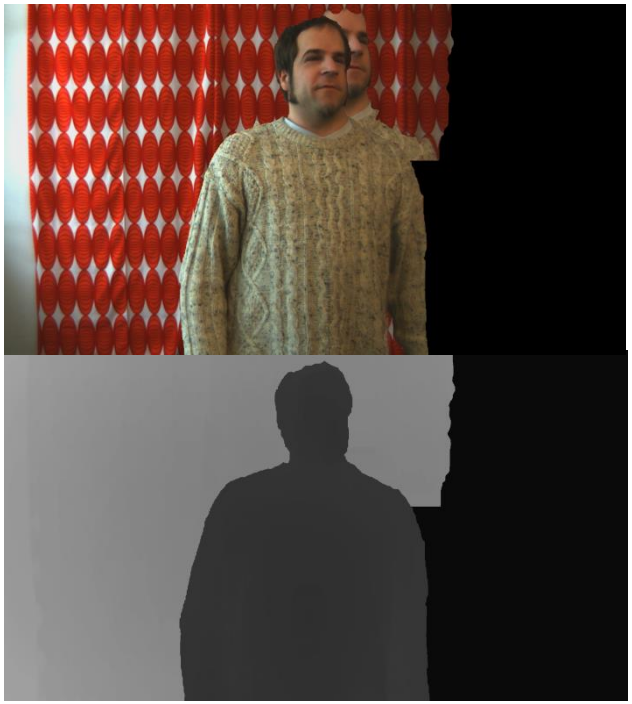
16 : 9



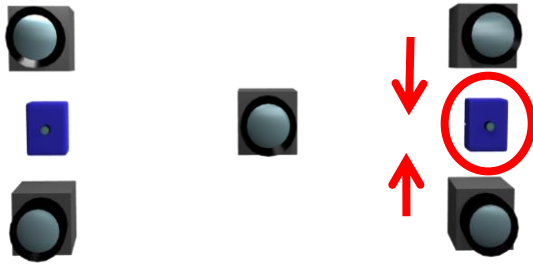
LDV Generation

- Generation of three views on a baseline

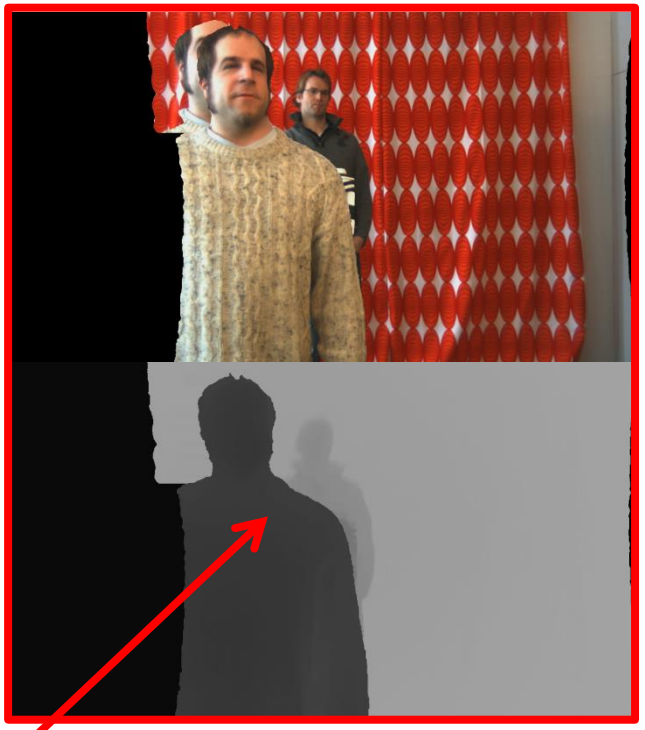
Left virtual



16 : 9



Right virtual

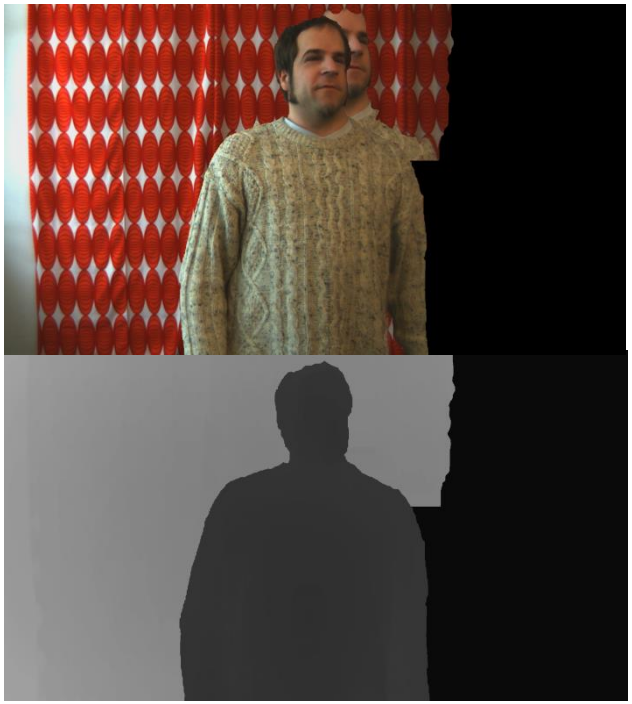


16 : 9

LDV Generation

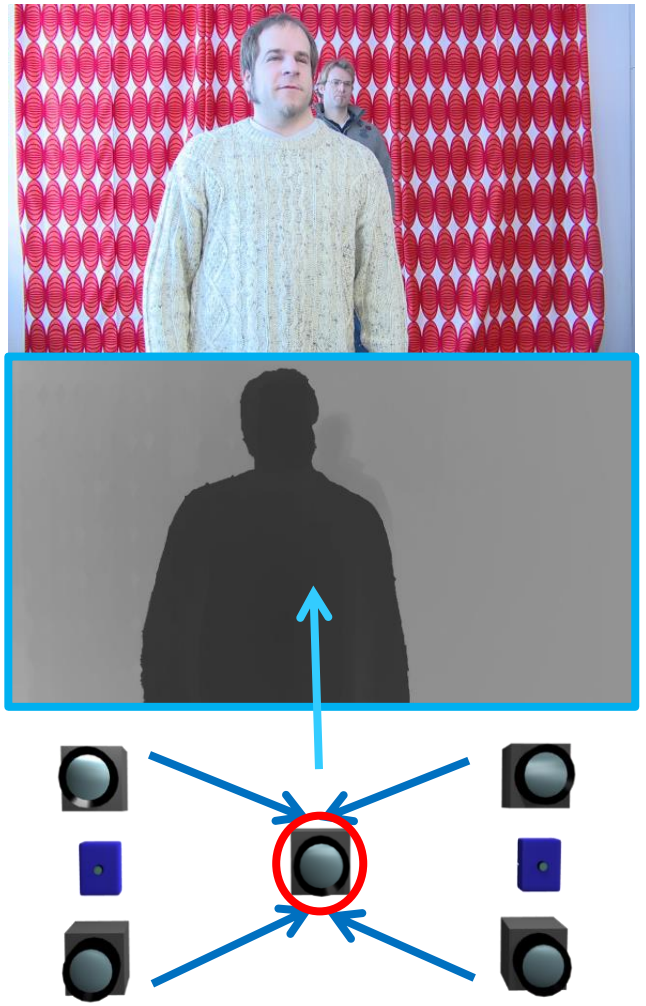
- Generation of three views on a baseline

Left virtual

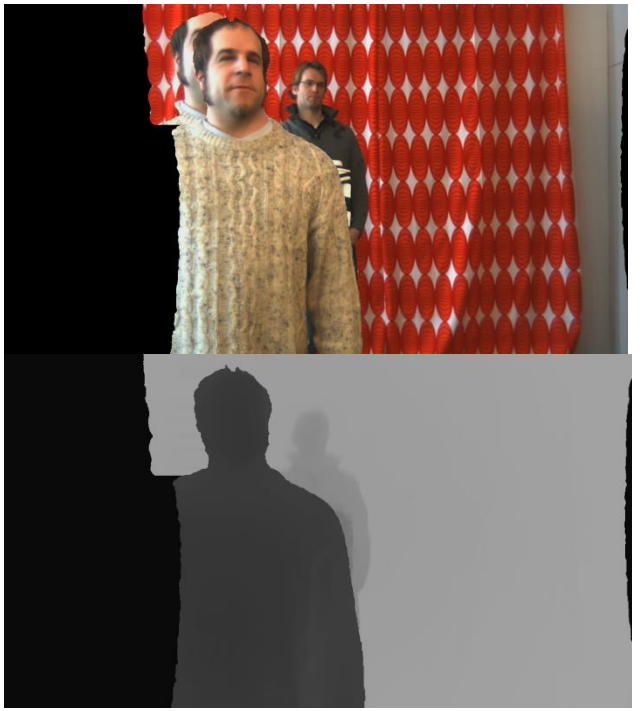


16 : 9

Original color (depth 16:9)



Right virtual

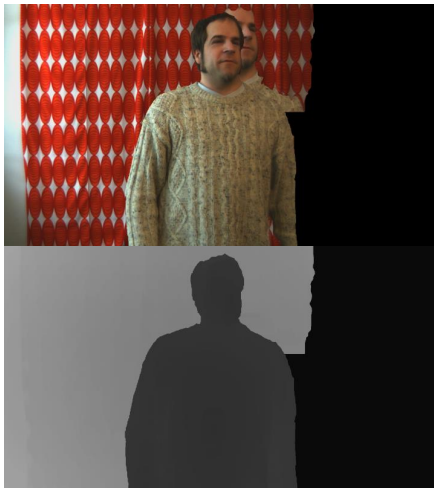


16 : 9

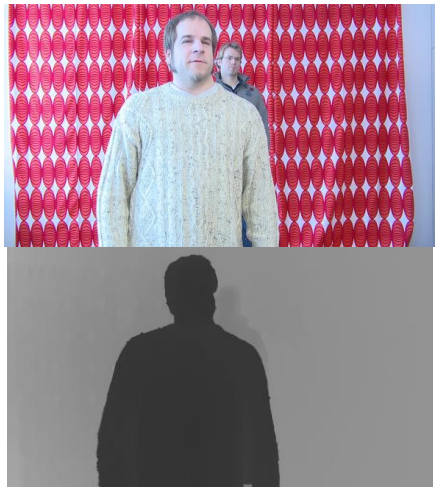
LDV Generation

- Generation of Occlusion Layer

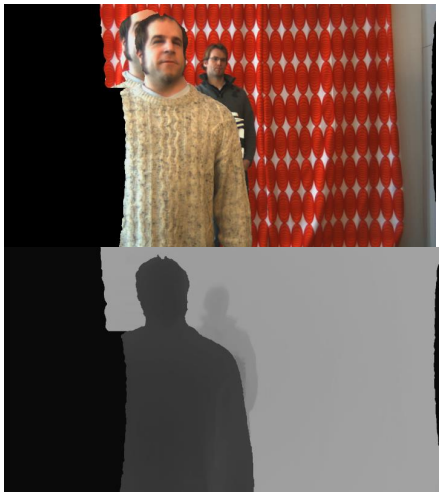
Left virtual



Original



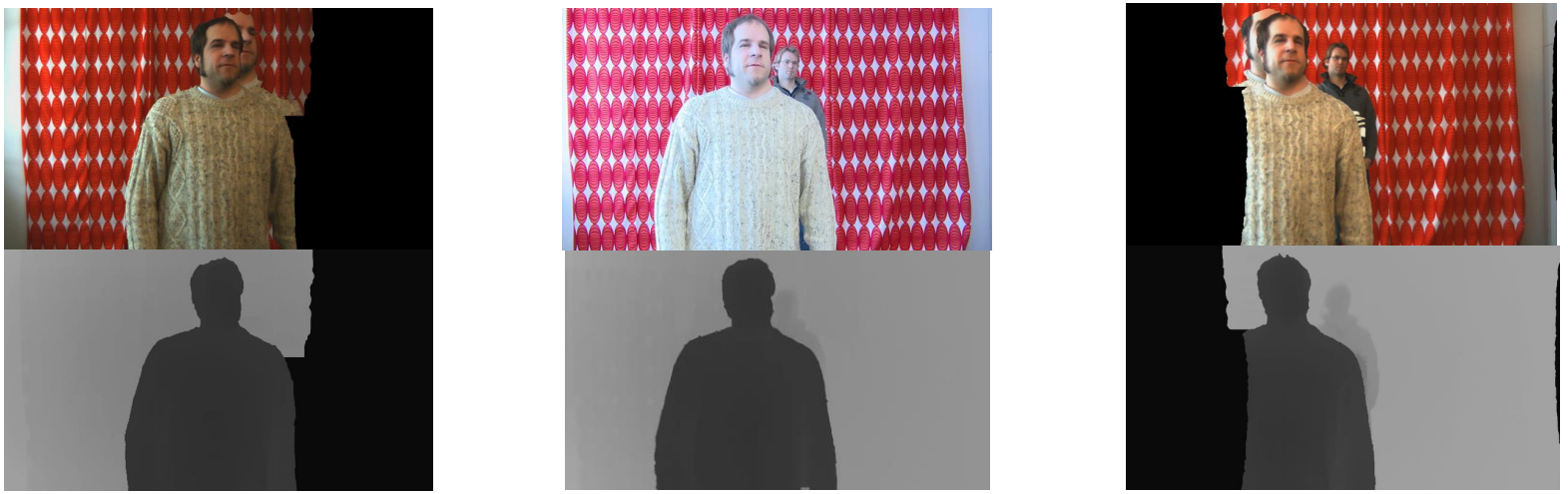
Right virtual

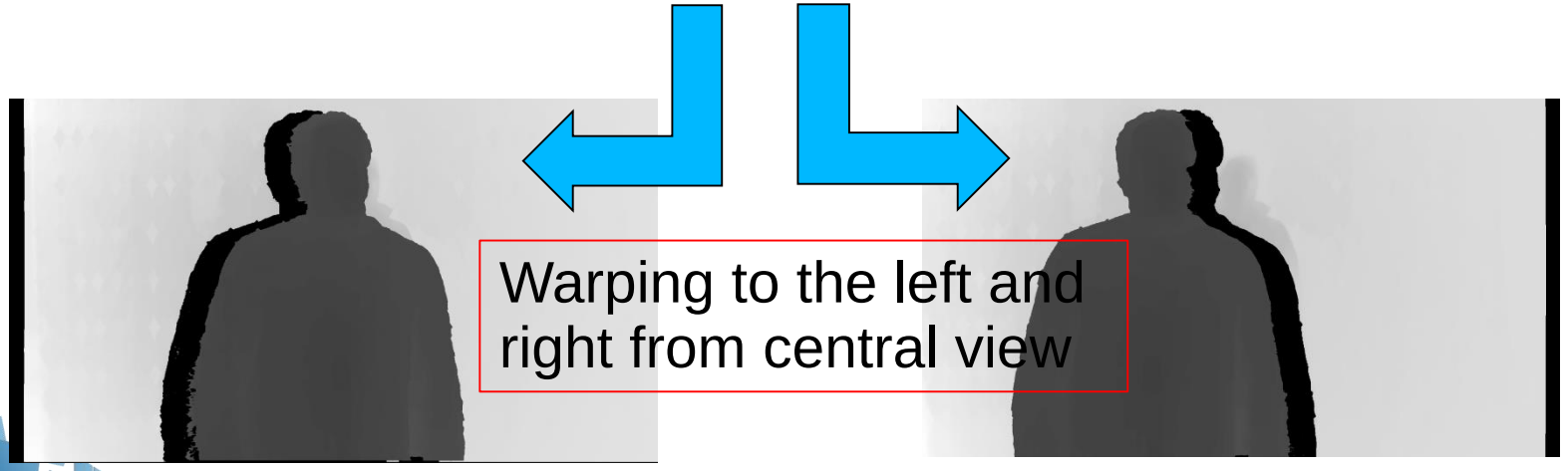


LDV Generation

- Generation of Occlusion Layer

Left virtual | Original | Right virtual

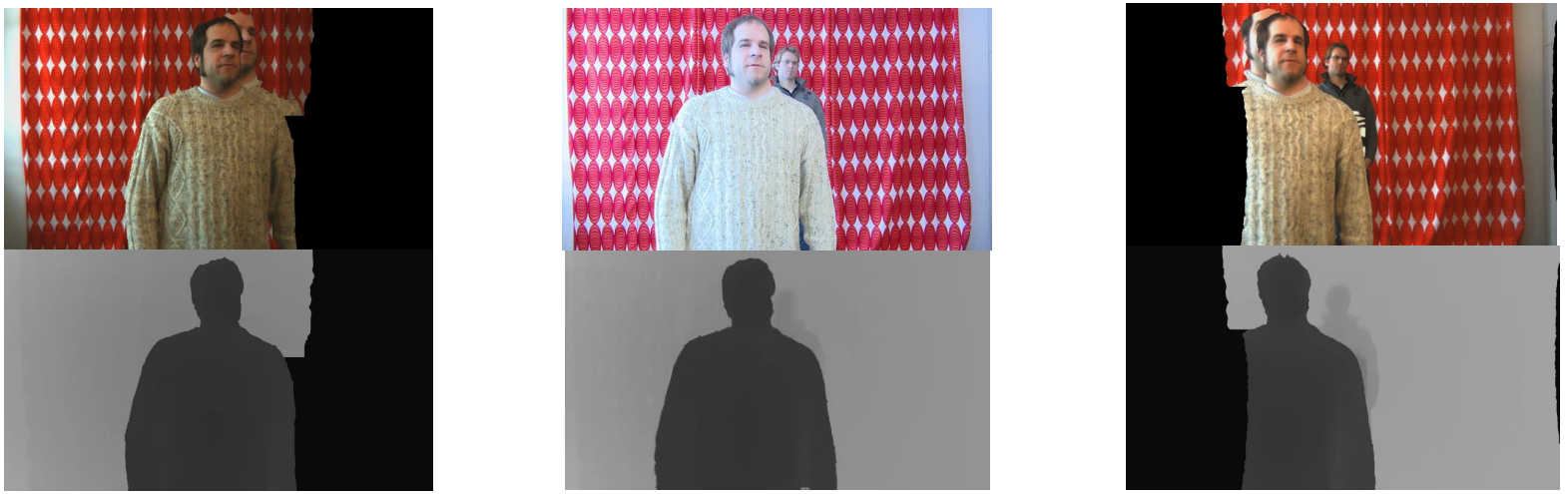


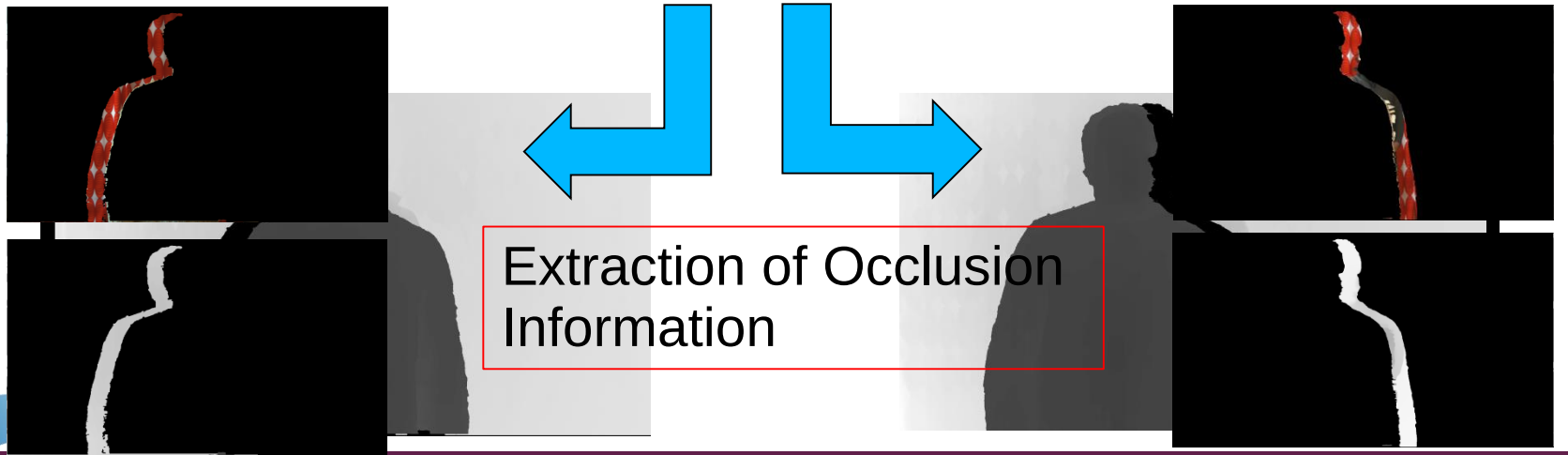


LDV Generation

- Generation of Occlusion Layer

Left virtual | Original | Right virtual

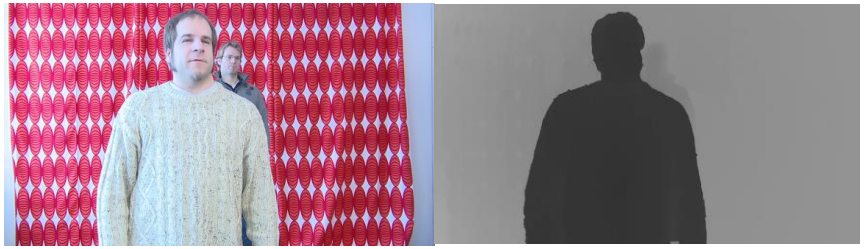




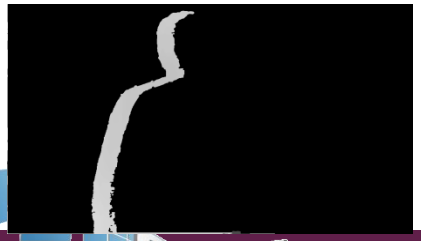
LDV Generation

- Generation of Occlusion Layer

Foreground Layer

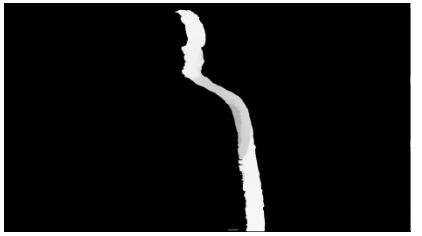


Occlusion Layer

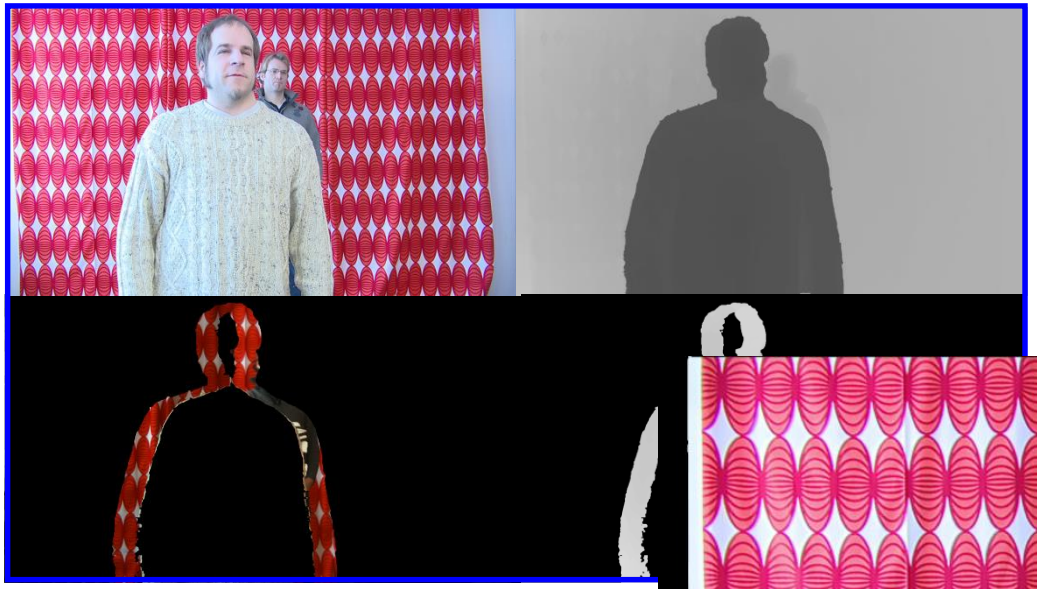


Warping to the central view

Merging to occlusion layer



Rendering from LDV



LDV Frame (2 Layers)

Rendered Virtual Views

Conclusions

- **Summary**
  - Hybrid ToF-Multi-Camera rig (5+2)
  - LDV Generation with occlusion layer (offline)
  - LDV Generation without occlusion layer (online)

- **Future work**
  - Merging ToF and Stereo Matching
  - Color adjustment between Occlusion and Foreground Layer