ETN-FPI TS3 “Plenoptic Sensing”

Tutorial on Sparse Lightfield Capturing with RGB-D Multi-Camera Setups

Dr.-Ing. Sandro Esquivel

March 14, 2017

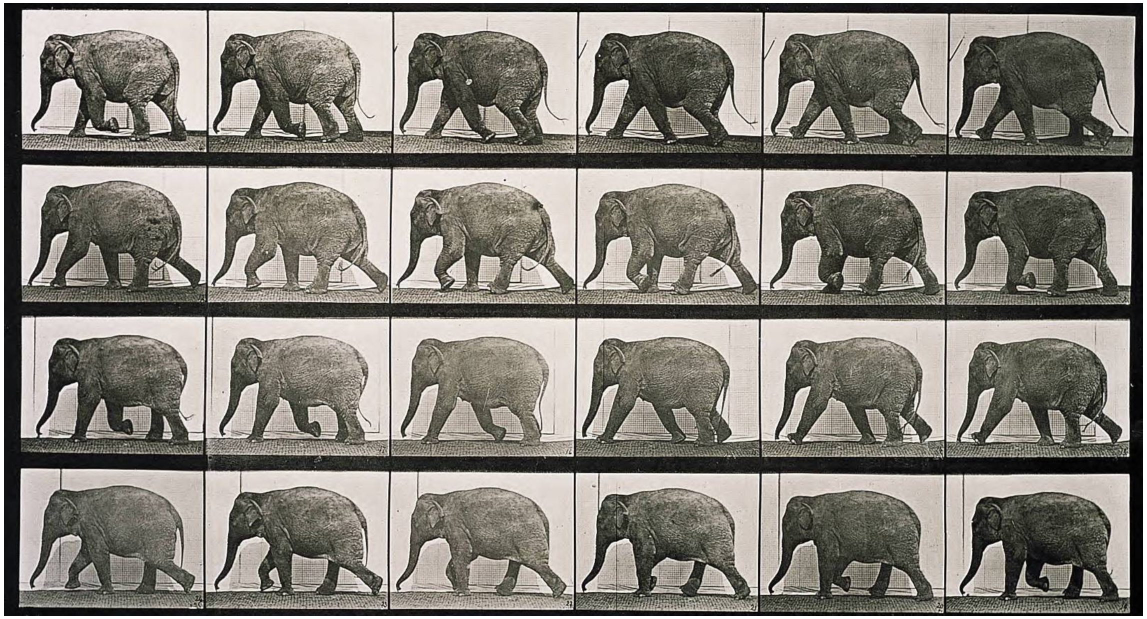
Multimedia Information Processing
Department of Computer Science
Christian-Albrechts-Universität zu Kiel

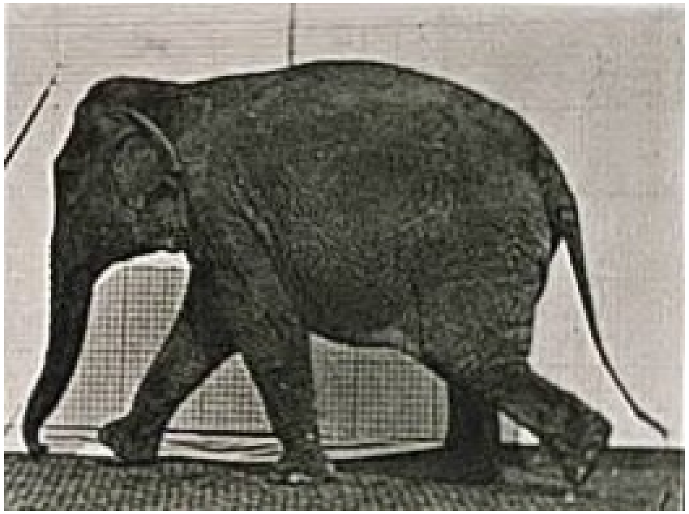
Outline

- **Introduction**
- Part I: Basics of Mathematical Optimization
  - Linear Least Squares
  - Nonlinear Optimization
- Part II: Basics of Computer Vision
  - Camera Model
  - Multi-Camera Model
  - Multi-Camera Calibration
- Part III: Depth Cameras
  - Passive Stereo
  - Structured Light Cameras
  - Time of Flight Cameras

Introduction

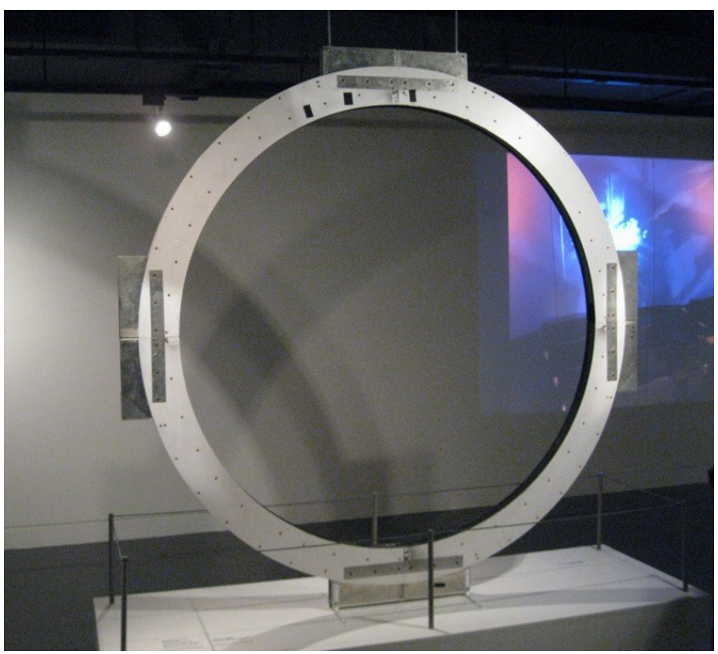
- Pioneers of light field capturing with multi-camera setups:
  - **Eadweard Muybridge (1870s/80s):**
    Capturing moving objects with array of time-delayed still cameras or simultaneously from different angles (1870s/80s)

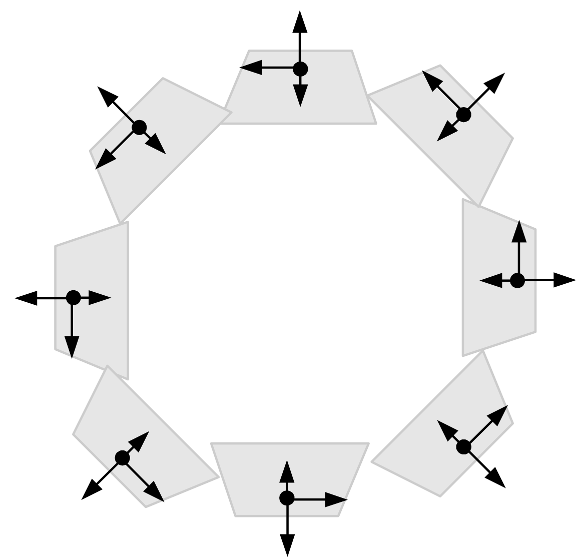




Introduction

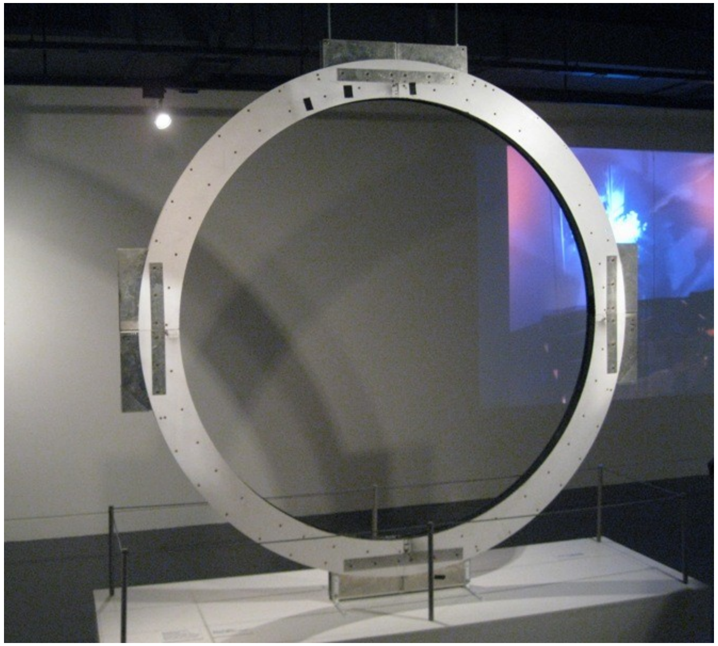
- Pioneers of light field capturing with multi-camera setups:
    - **Tim Macmillan, Dayton Taylor (1980s):** Time slice camera, capturing scene simultaneously with circular multi-camera setup, “frozen time” / “bullet time” effects





Introduction

- Pioneers of light field capturing with multi-camera setups:
  - **Tim Macmillan, Dayton Taylor (1980s):** Time slice camera, capturing scene simultaneously with circular multi-camera setup, “frozen time” / “bullet time” effects

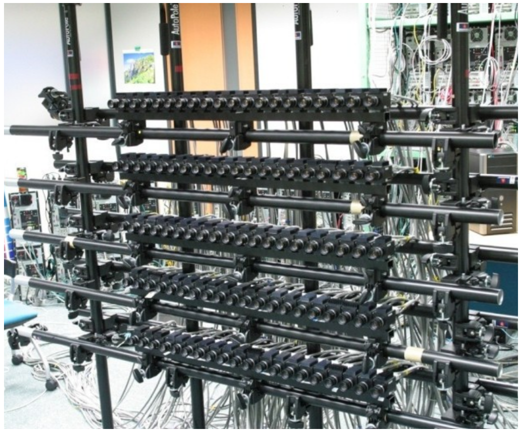


Introduction

- Devices for light field capturing (plenoptic sensing):
  - freely moving camera
  - multi-camera setup (or camera on gantry)
  - plenoptic camera (lens-based)
  - light field microscope

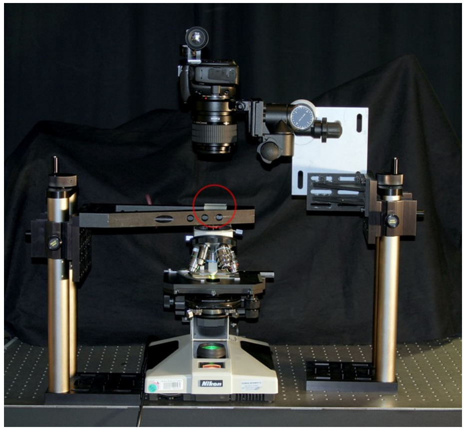
large-scale scene ↕ small-scale scene

Car-mounted camera (Google StreetView)



Multi-camera array (Tanimoto, 2010)

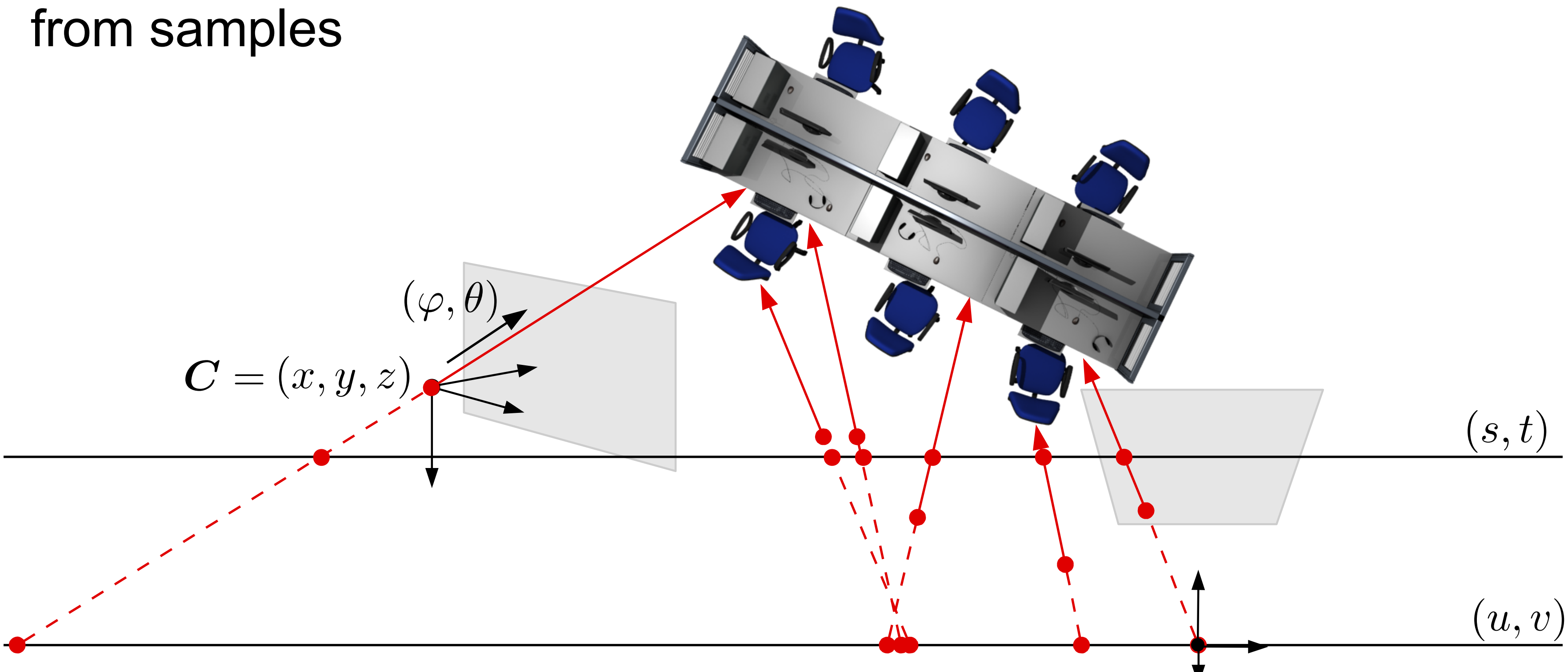
Plenoptic camera (Raytrix)



Light field microscope (Levoy, 2006)

Introduction

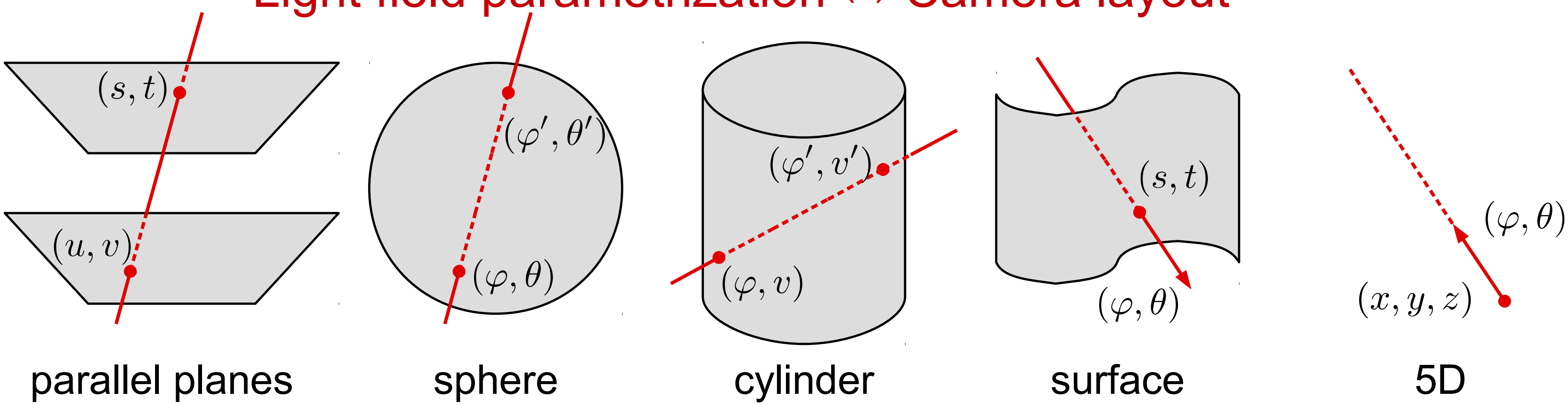
- Sample plenoptic function (light field) $L(x, y, z, \varphi, \theta)$ with many spatially distributed cameras
- Images sample light field with varying angles (φ, θ) and fixed (x, y, z)
- Reduce to 4D light field, *e. g.*, light slab representation $L(u, v, s, t)$
- **Aim:** Create new views (*e. g.*, different vantage point, focus change) from samples



Introduction

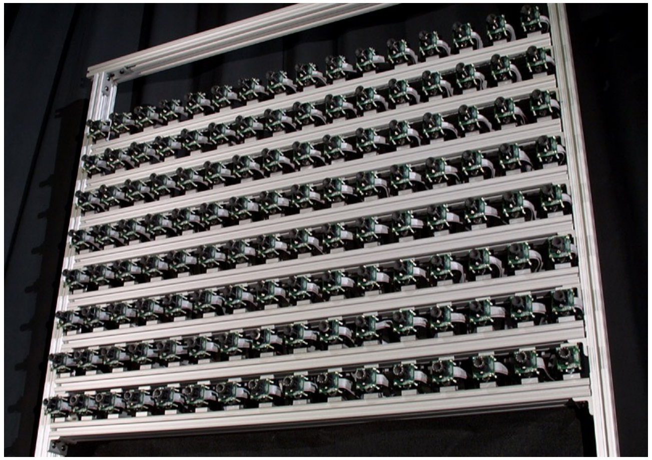
- Common ray parametrizations for light fields:
  - Intersection points with parallel planes (light slab, Lumigraph) or cube (Levoy & Hanrahan, 1996; Gortler et al., 1996)
  - Intersection points with sphere (Todt et al., 2007)
  - Intersection points with cylinder
  - 2D surface point and direction angles (Wood et al., 2000)
  - 3D point and direction angles (McMillan & Bishop, 1995)

Light field parametrization ↔ Camera layout

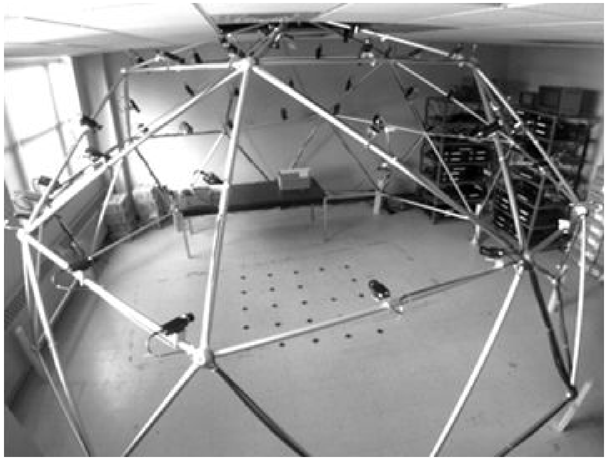


Introduction

- Camera layout depends on scene, application (*e. g.,* virtual motion range), and light field representation (*e. g.,* light slabs ↔ planar layout)
- Classical layouts: Grid layout (planar), Array layout (collinear), Dome layout (hemi-spherical)
- For moving camera: Layout equals motion range (*e. g.,* spherical gantry)
- Dense vs. sparse spatial distribution of cameras



Stanford Multi-Camera Array (Wilburn et al. 2004)



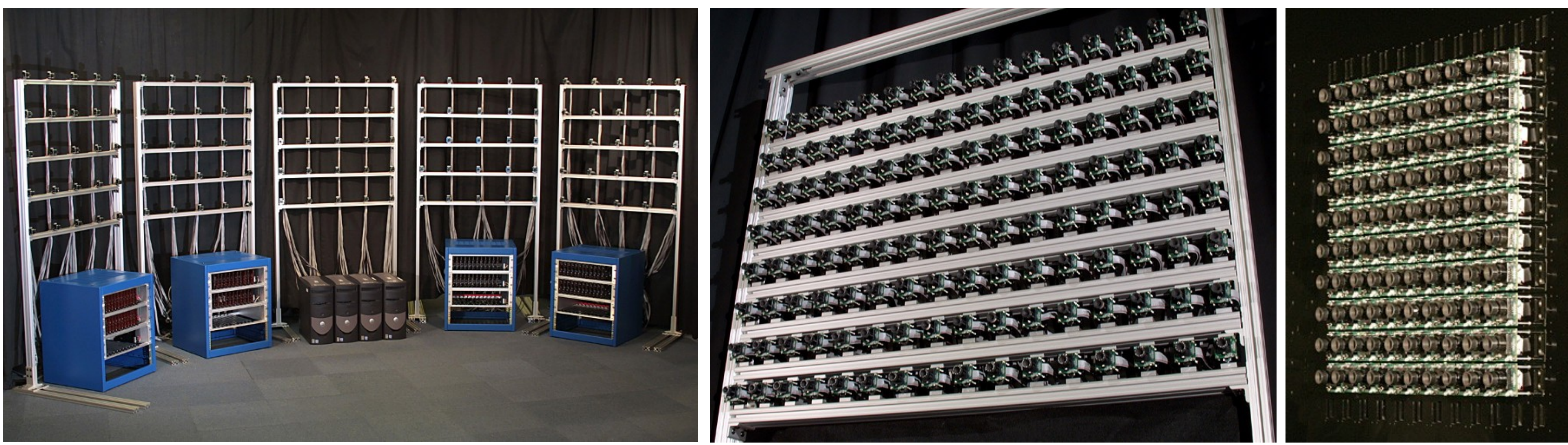
CMU "3D Room" VR Lab (Kanade et al. 1996)

OTOY Spherical Light Field Capturing System

Introduction

- Applications for multi-camera arrays for plenoptic sensing:
  - Widely spaced: Light field capturing for free-viewpoint video
  - Tightly packed: Synthetic aperture, refocusing, depth-of-field, light field capturing for small scenes and limited viewpoint video

wide ⟷ dense



Different versions of the Stanford Multi-Camera Array

Introduction

- Example applications for multi-camera arrays for plenoptic sensing:
  - Input data: Many images from different positions
  - 3D information from depth cameras (more about this later)

Introduction

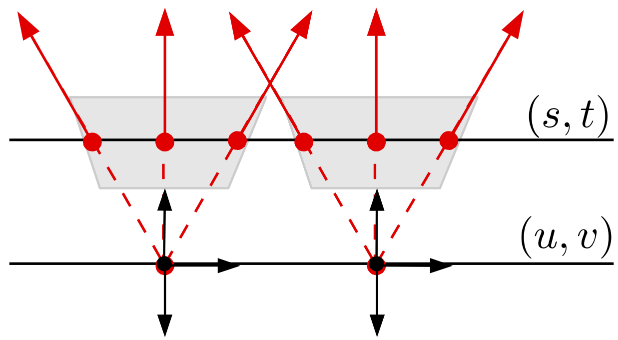
- Example applications for multi-camera arrays for plenoptic sensing:
  - Synthetic aperture (change of focus plane)

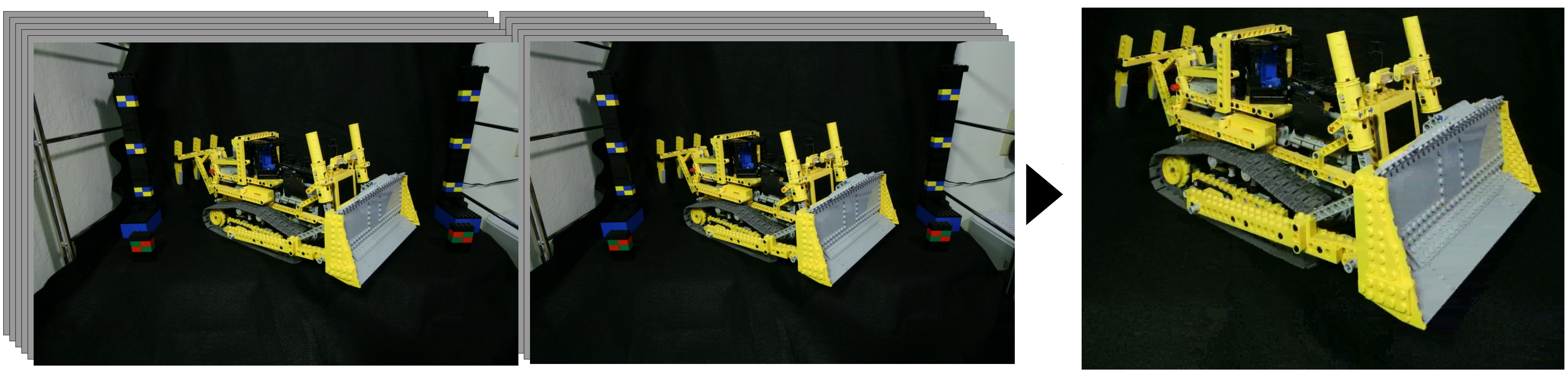
Introduction

- Example applications for multi-camera arrays for plenoptic sensing:
  - Synthetic aperture (change of focus plane)
  - Free-viewpoint rendering

Introduction

- **Example:** Synthetic aperture with planar camera layout (Stanford)
- Input: Rectified images, *i. e.,* mapped to light slab parametrization $L(u, v, s, t) = I_{u,v}(s, t)$, where (u, v) is camera position in plane
- Assumption: Camera centers are collinear, similar viewing directions
- Mapping to common plane must be known (or: full camera calibration)





Online viewer for light field datasets: http://lightfield.stanford.edu/lfs.html

Introduction

- **Task:** How to access the light field with multi-camera rigs? How to describe mapping from images $\mathbf{I}_1,\ldots,\mathbf{I}_N$ to light field $L(x, y, z, \varphi, \theta)$?

- **Aims** of this tutorial lecture:
  - Overview of Computer Vision problems involved
  - Tutorial on numeric methods (least squares, nonlinear optimization)
  - Details on geometric camera calibration
  - Working principles of depth cameras

- Topics beyond the scope of this lecture:
  - Technical issues (*e. g.*, synchronization, camera architecture)
  - Light field representation and compression
  - Light field rendering and video post-processing