Plenoptic Modeling and Lightfields

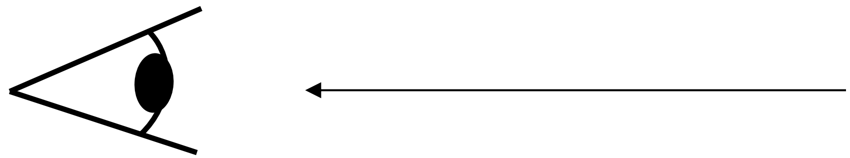
Daniel G. Aliaga

CS635

(slides based on Richard Szeliski, Michael Cohen, Marc Levoy, Leonard McMillan, Jian Huang, Jin-Xiang Chai, and Heung-Yeung Shum)

Recall: Light is…

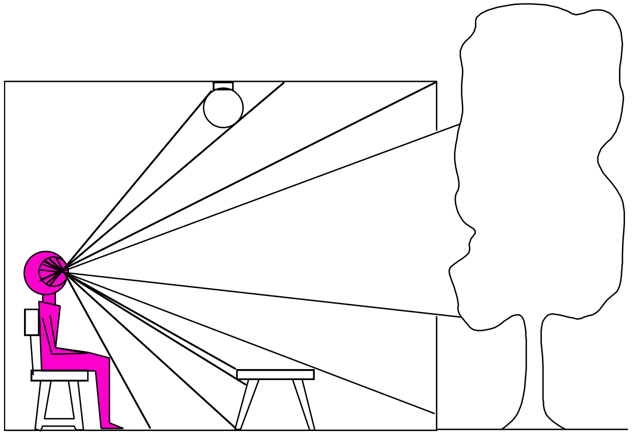
- Electromagnetic radiation (EMR) moving along rays in space
  - R(λ) is EMR, measured in units of power (watts)
    - λ is wavelength



- Useful things:
- Light travels in straight lines
- In vacuum, radiance emitted = radiance arriving
  - i.e. there is no transmission loss

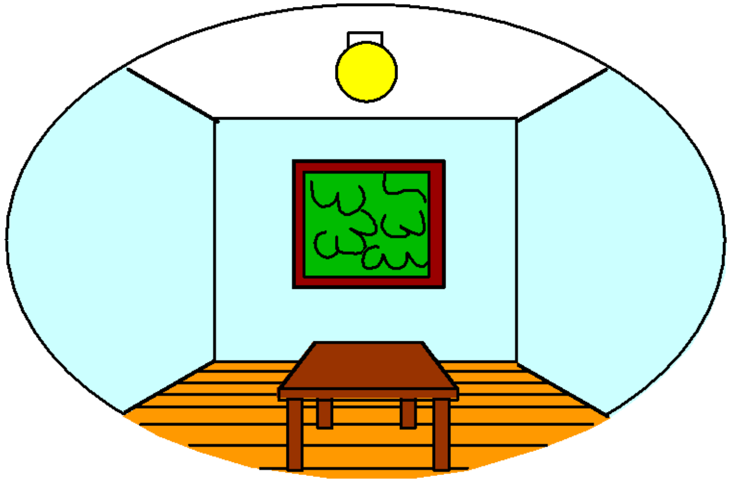
What do we see?

3D world



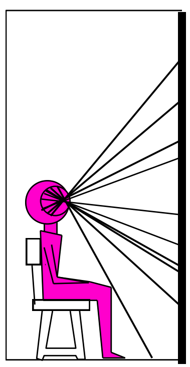
Point of observation

2D image



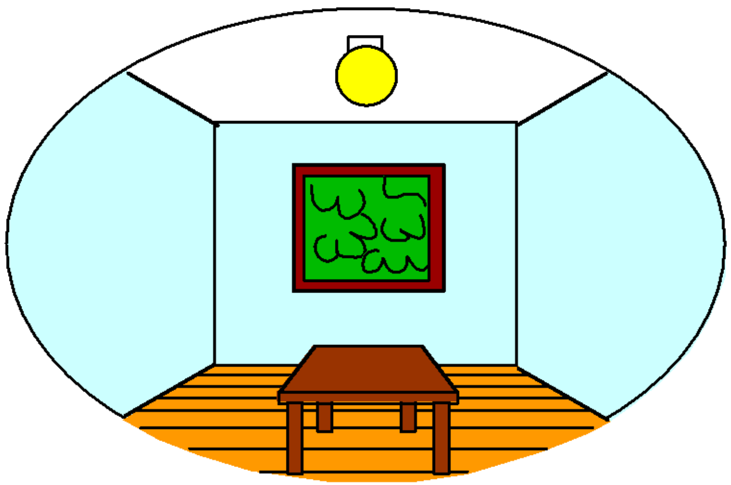
What do we see?

3D world



Painted backdrop

2D image



The Plenoptic Function

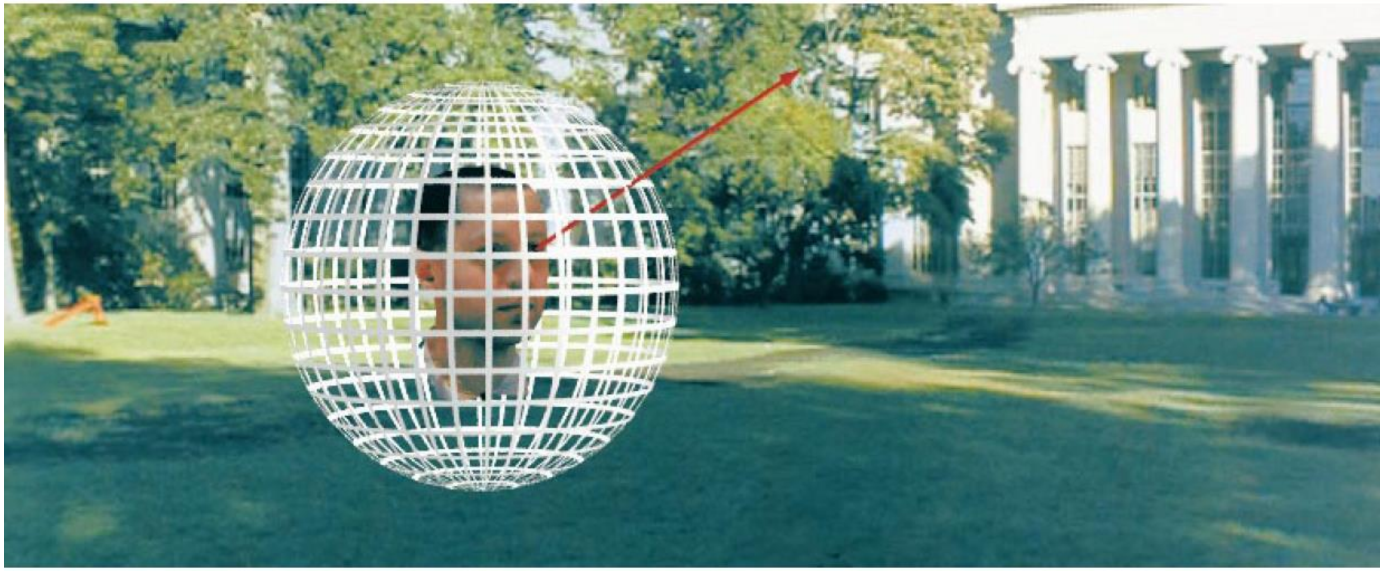


Figure by Leonard McMillan

- Q: What is the set of all things that we can ever see?
- A: The Plenoptic Function (Adelson & Bergen)

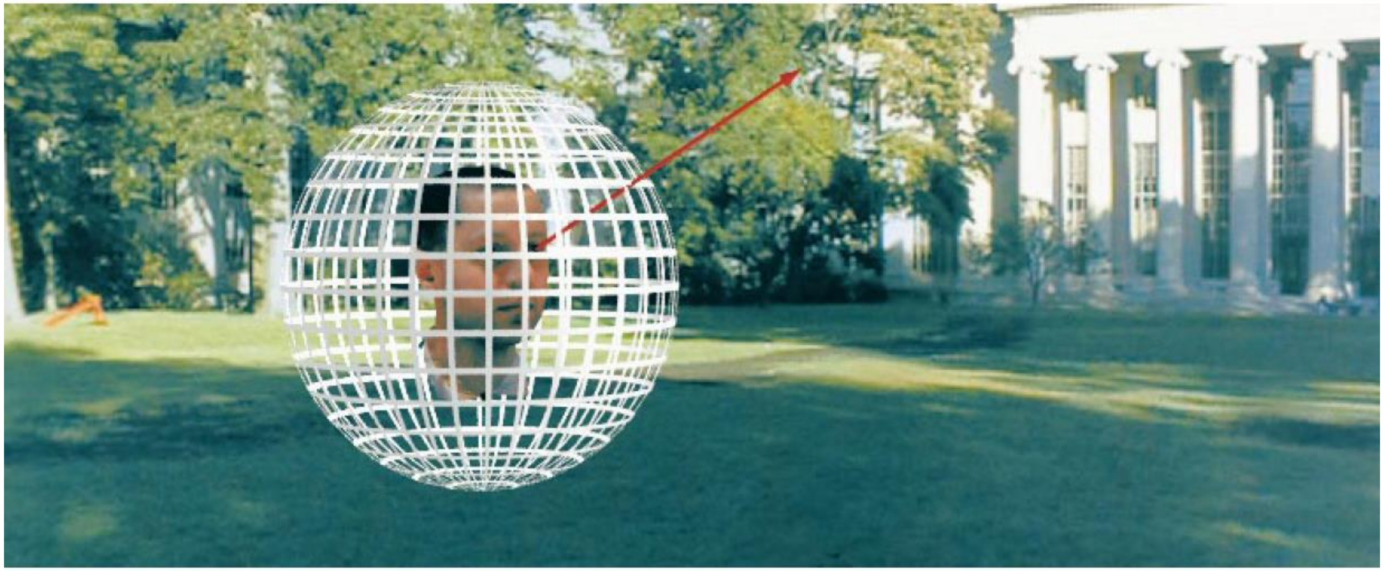
- Let’s start with a stationary person and try to parameterize <u>everything</u> that she can see…

Grayscale snapshot

$$P(\theta,\phi)$$

- is intensity of light
  - Seen from a single view point
  - At a single time
  - Averaged over the wavelengths of the visible spectrum
- (can also do *P(x,y),* but spherical coordinate are nicer)

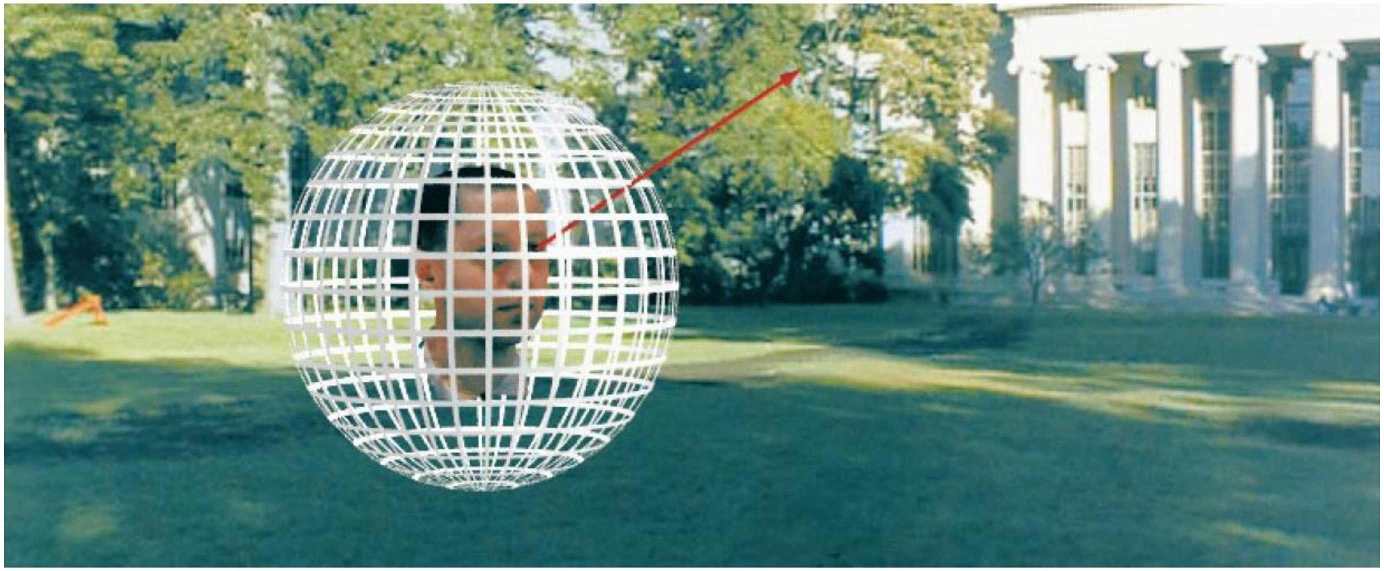
Color snapshot



$$P(\theta,\phi,\lambda)$$

- is intensity of light
  - Seen from a single view point
  - At a single time
  - As a function of wavelength

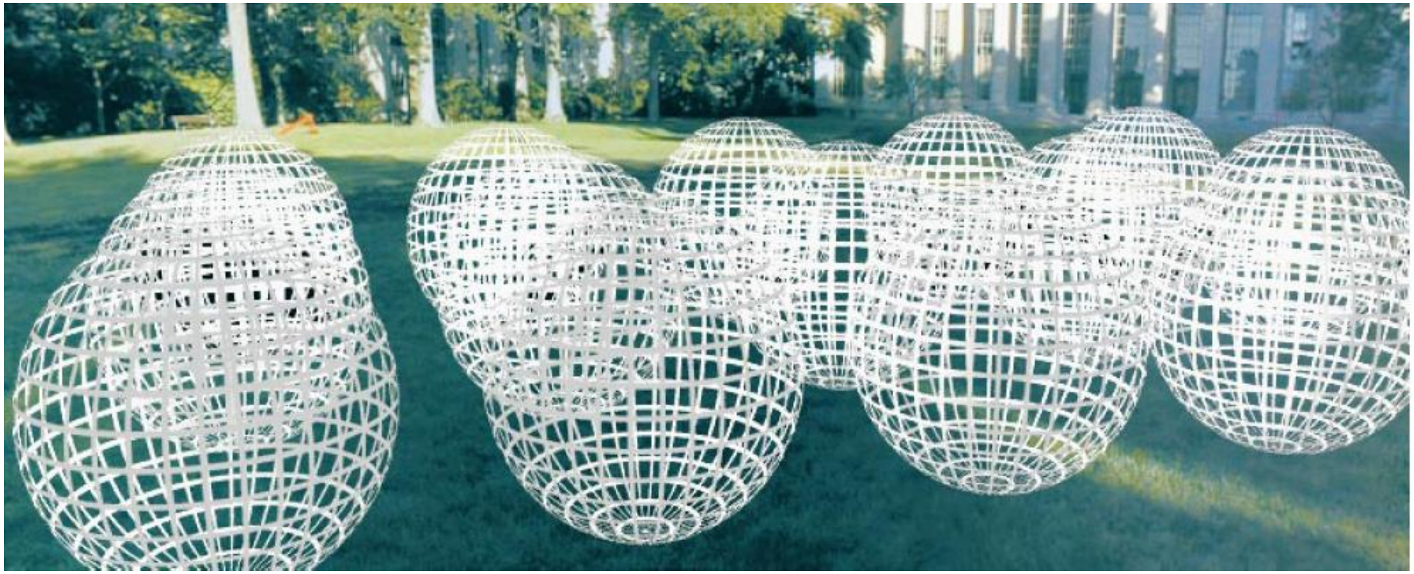
A movie



$$P(\theta,\phi,\lambda,t)$$

- is intensity of light
  - Seen from a single view point
  - Over time
  - As a function of wavelength

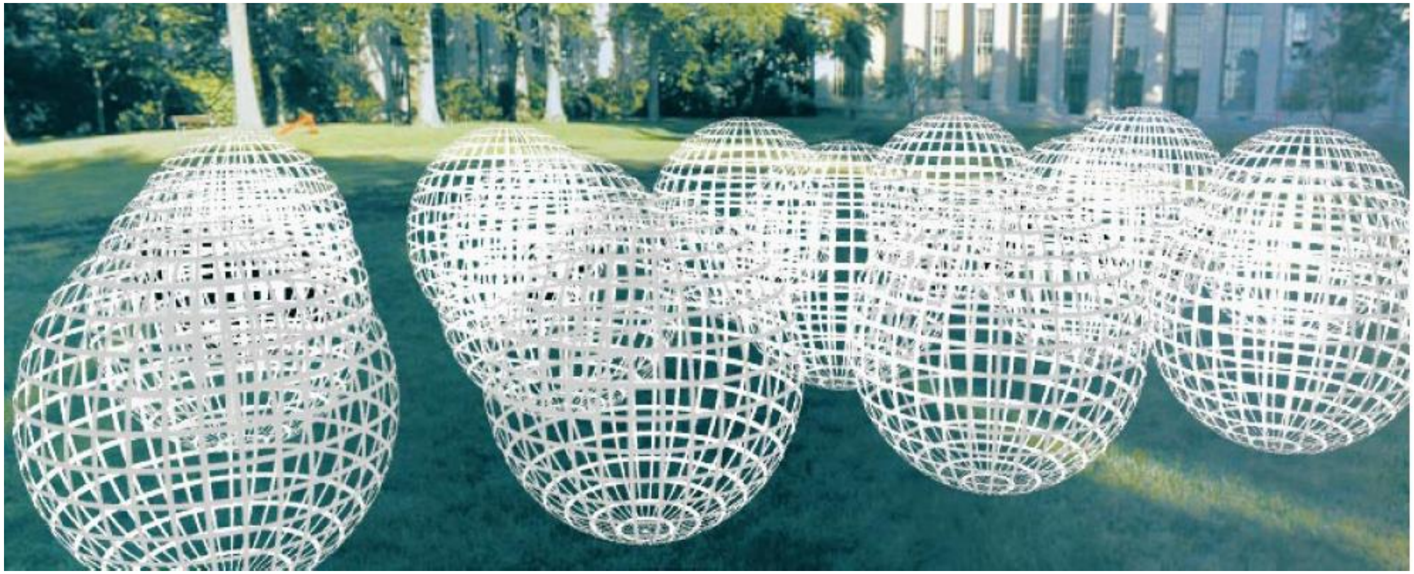
Holographic movie



$$P(\theta,\phi,\lambda,t,V_X,V_Y,V_Z)$$

- is intensity of light
  - Seen from ANY viewpoint
  - Over time
  - As a function of wavelength

The Plenoptic Function

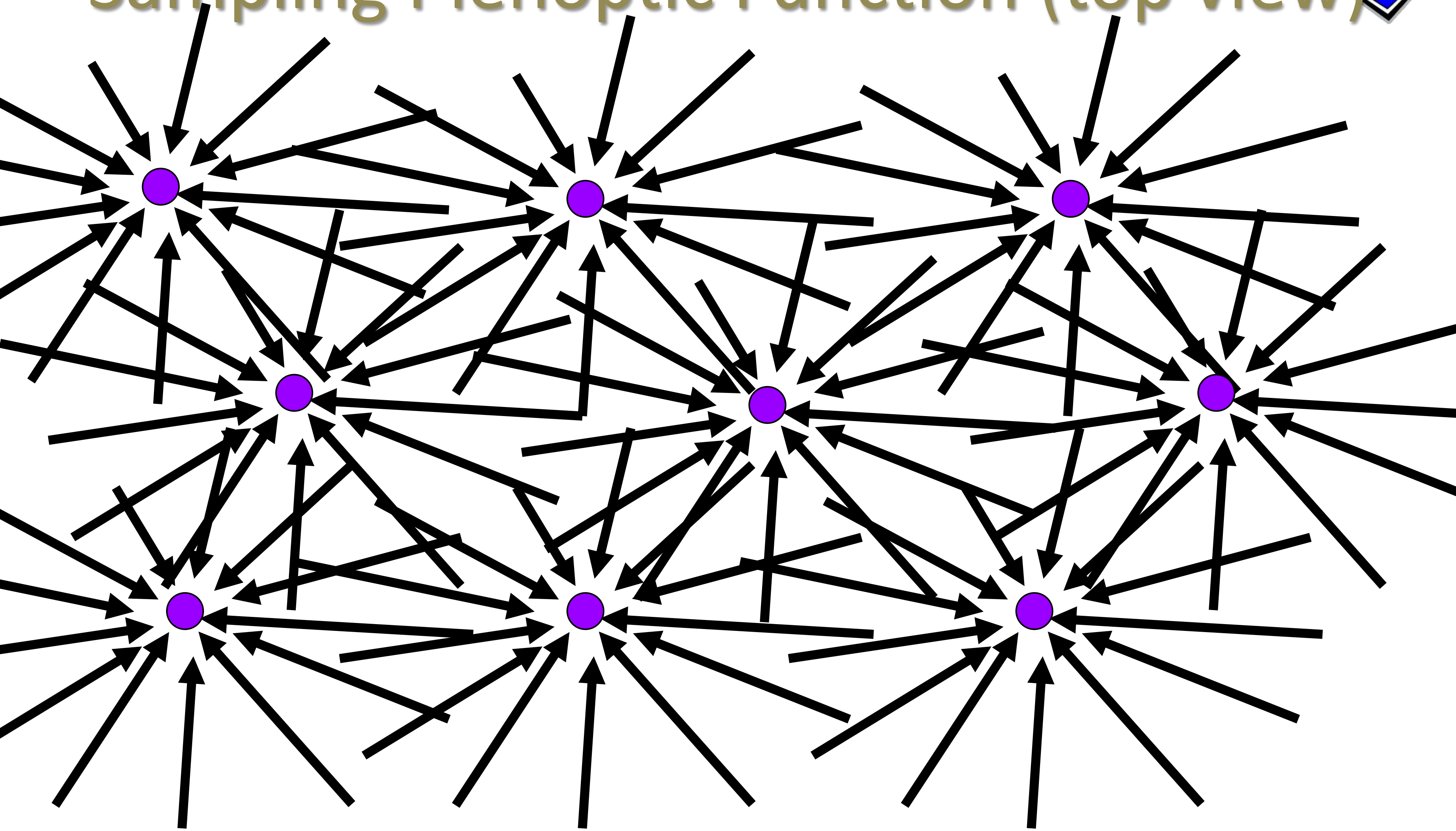


$$P(\theta,\phi,\lambda,t,V_X,V_Y,V_Z)$$

– Can reconstruct every possible view, at every moment, from every position, at every wavelength

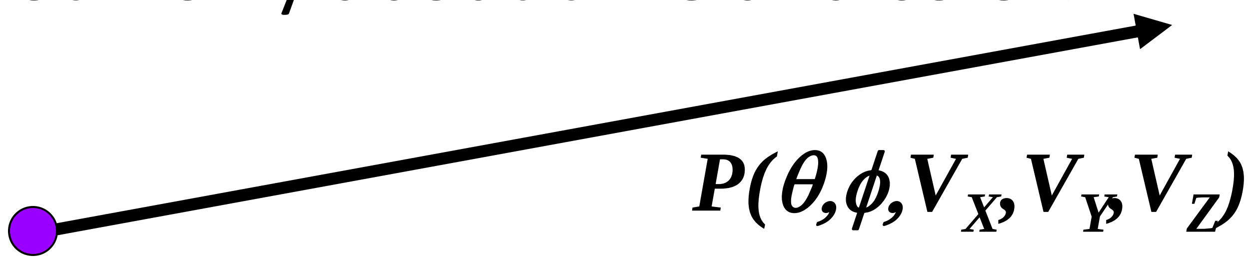
– Contains every photograph, every movie, everything that anyone has ever seen.

Sampling Plenoptic Function (top view)



Ray

- Let’s not worry about time and color:



- 5D
  - 3D position
  - 2D direction

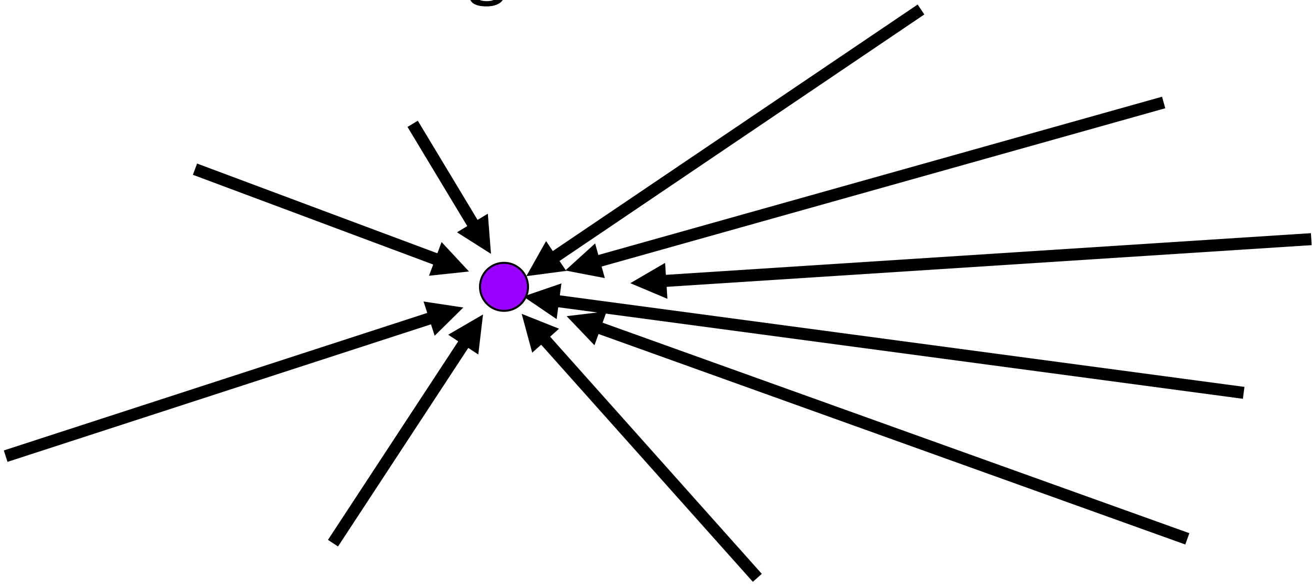
Slide by Rick Szeliski and Michael Cohen

Plenoptic Function

- “Holodeck” (Star Trek)
- Layered Depth Images [Shade98]
- 3D Image Warping [Max95, McMillan95, …]
- View Interpolation [Chen93]
- Sea of Images [Aliaga01]
- Lightfield/Lumigraph [Levoy96, Gortler96]
- Plenoptic Stitching [Aliaga99]
- Concentric Mosaics [Shum99]
- Panoramic Images [Szeliski97, …]

2D: Image

- What is an image?

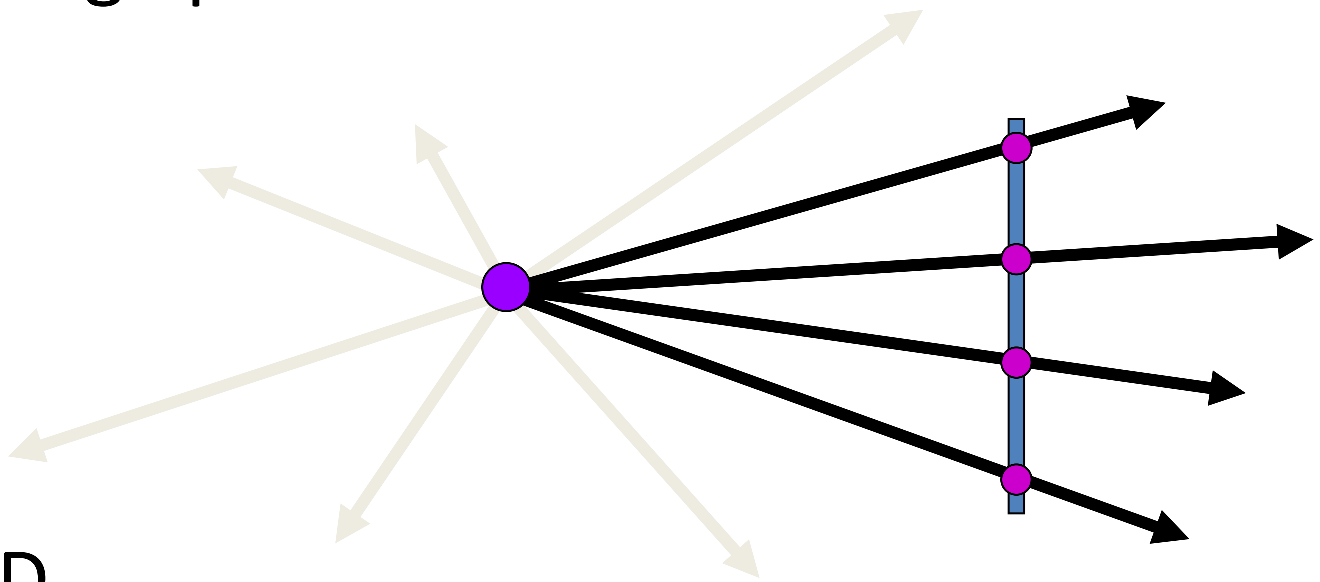


- All rays through a point

Slide by Rick Szeliski and Michael Cohen

2D: Image

- Image plane



- 2D
  - position

2D: Spherical Panorama

See also: 2003 New Years Eve
http://www.panoramas.dk/fullscreen3/f1.html

- All light rays through a point form a panorama
- Totally captured in a 2D array -- $P(\theta,\phi)$

3D: Space-Time

- Moving in time:
  - Spatio-temporal volume: $P(\theta,\phi,t)$
  - Useful to study temporal changes
  - Long an interest of artists:

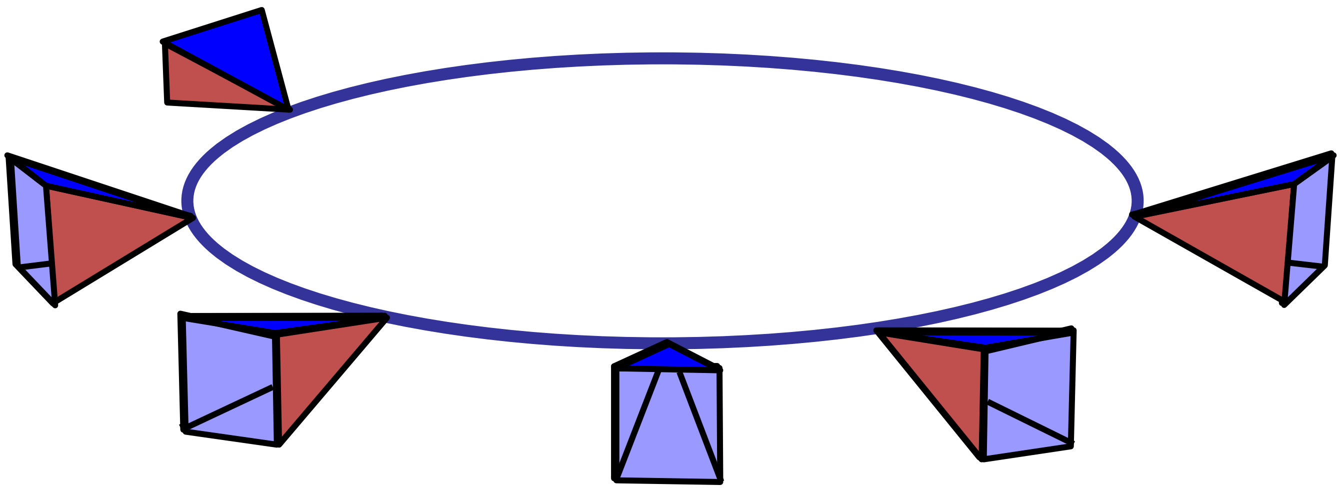
Claude Monet, Haystacks studies

3D: Space-Time

Other ways to slice the plenoptic function…

3.5D: Concentric Mosaics

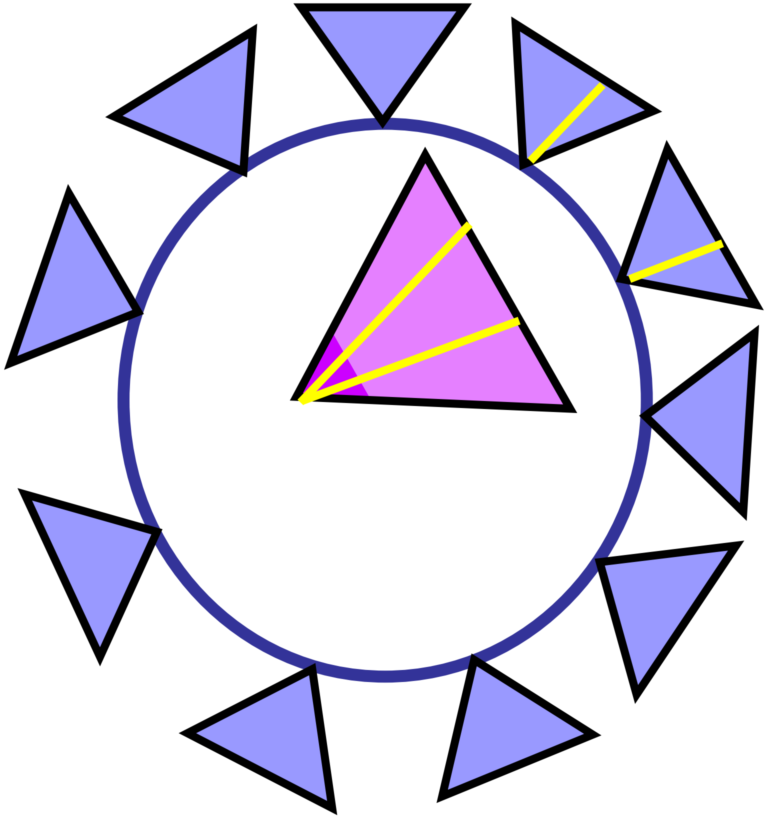
- Replace "row" with "circle" of images



3.5D: Concentric Mosaics

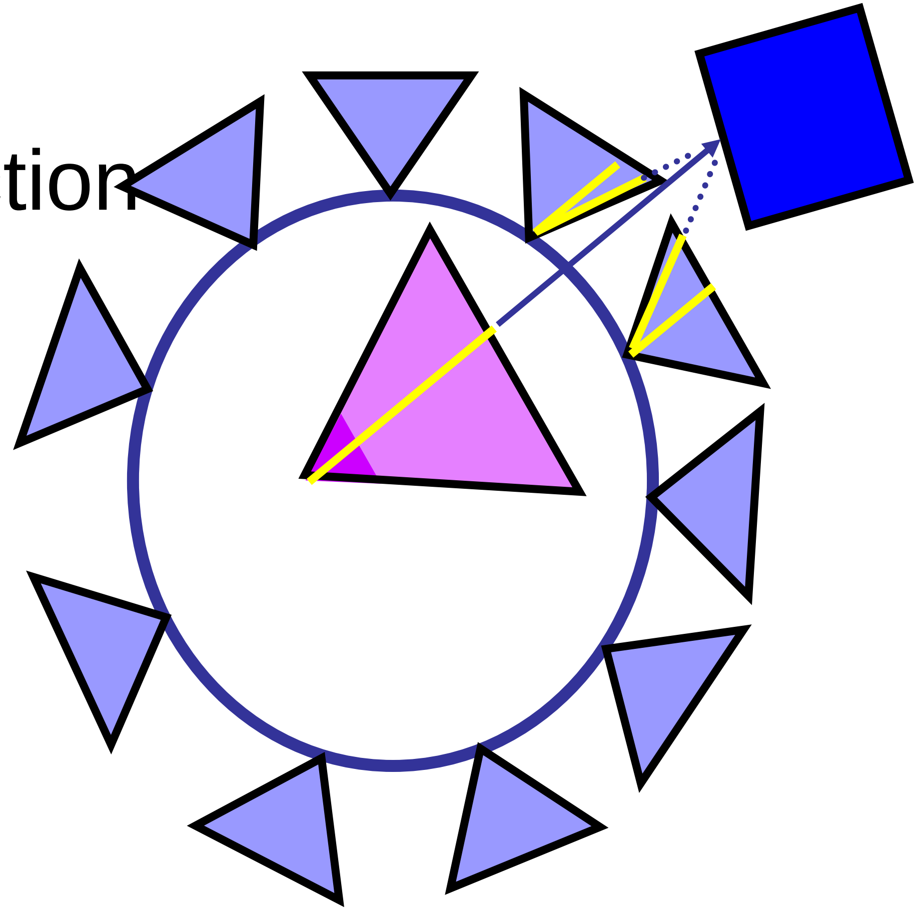
3.5D: Concentric Mosaics

- From above



Concentric Mosaics

Depth correction

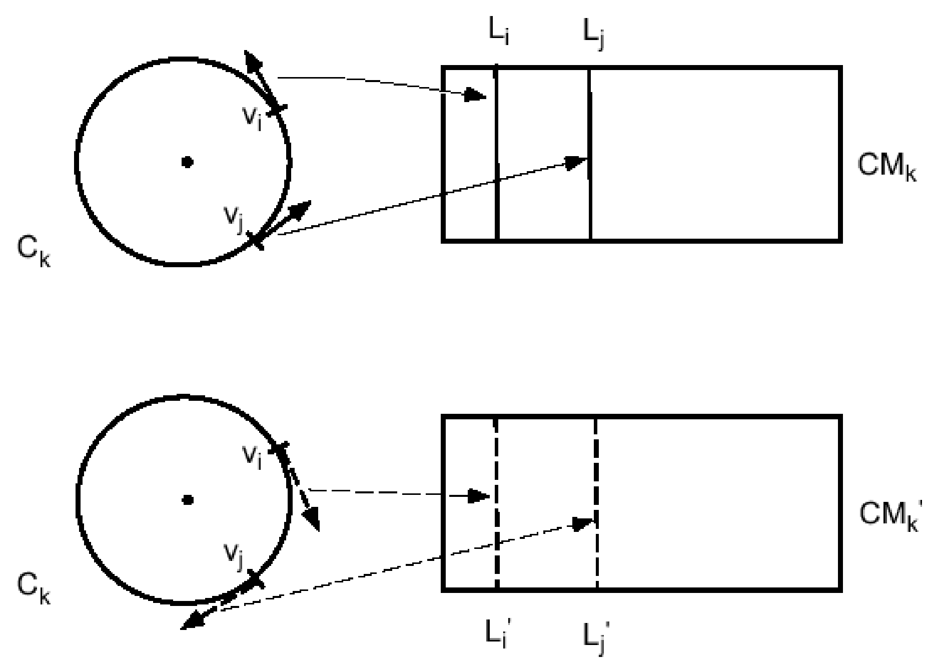


Concentric Mosaics

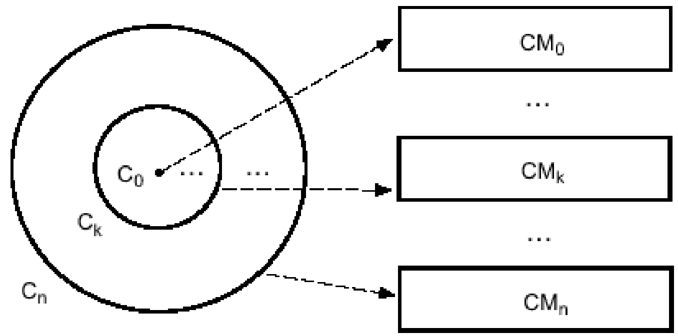
- Panorama

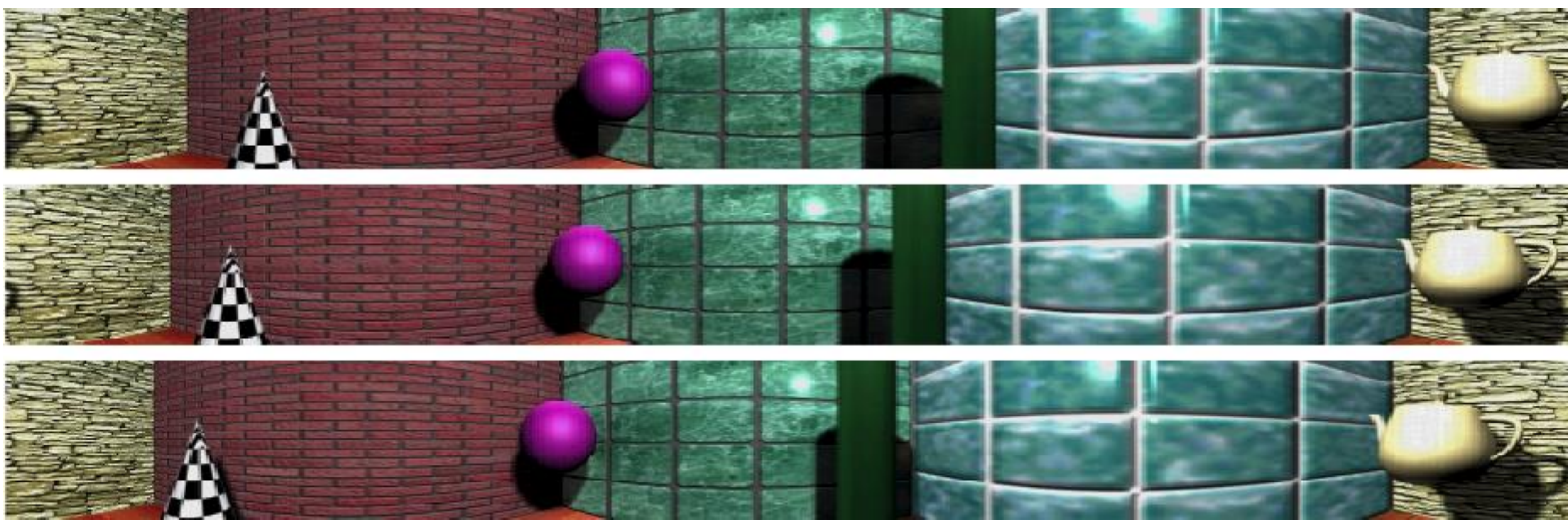
Concentric Mosaics

- A set of manifold mosaics constructed from slit images taken by cameras rotating on concentric circles

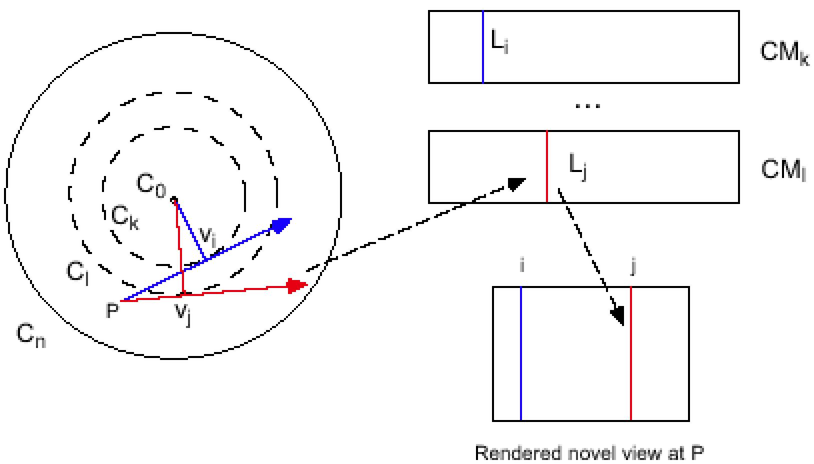


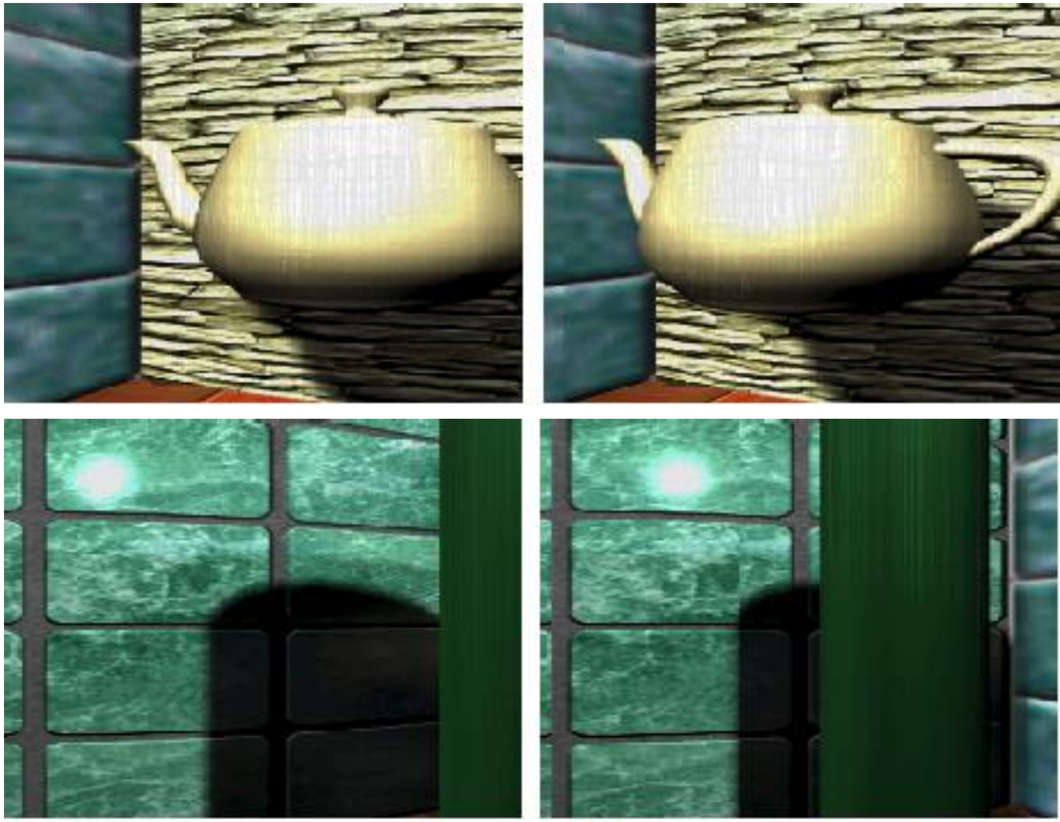
Sample Images





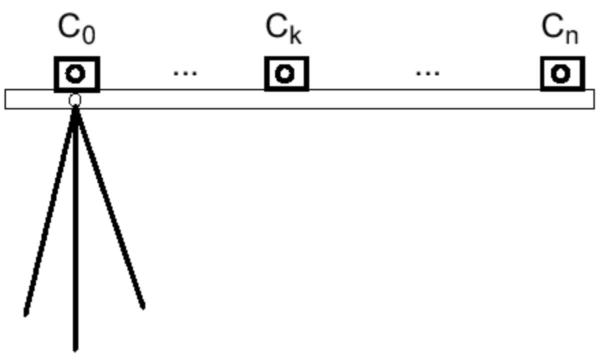
Rendering a Novel View





Construction of Concentric Mosaics (2)

- Real scenes



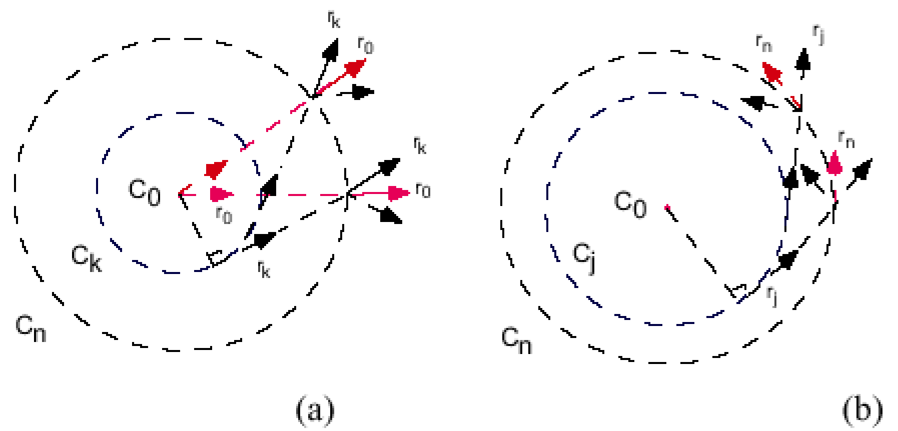


Figure 10: Construction of concentric mosaics from one circle: camera along (a) normal direction; (b) tangential direction.

Bulky, costly

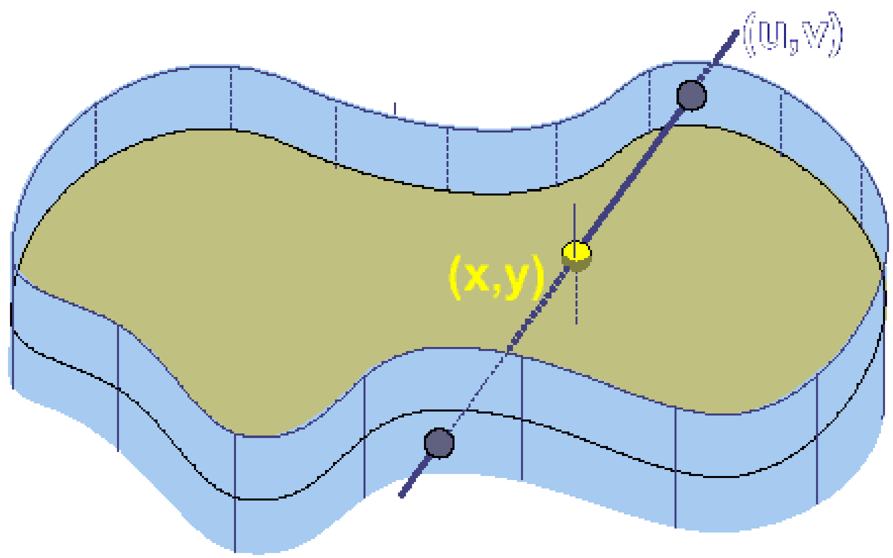
Cheaper, easier

Results

Issues with Concentric Mosaics

- **Only has horizontal parallax effects**
- Limited horizontal fov
- Non-uniform spatial horizontal resolution

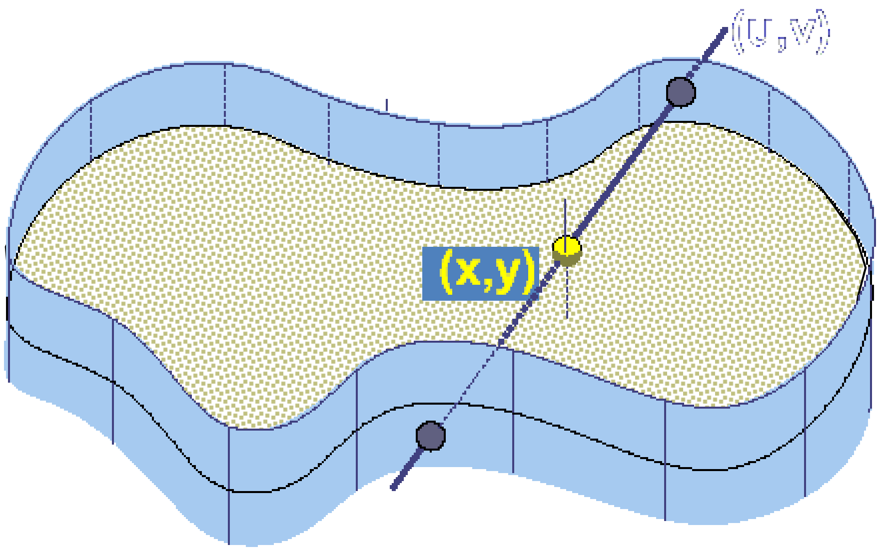
3.5D: Plenoptic Stitching



Observer plane at eye height

- 4D parameterization of the plenoptic function suitable for walkthrough applications

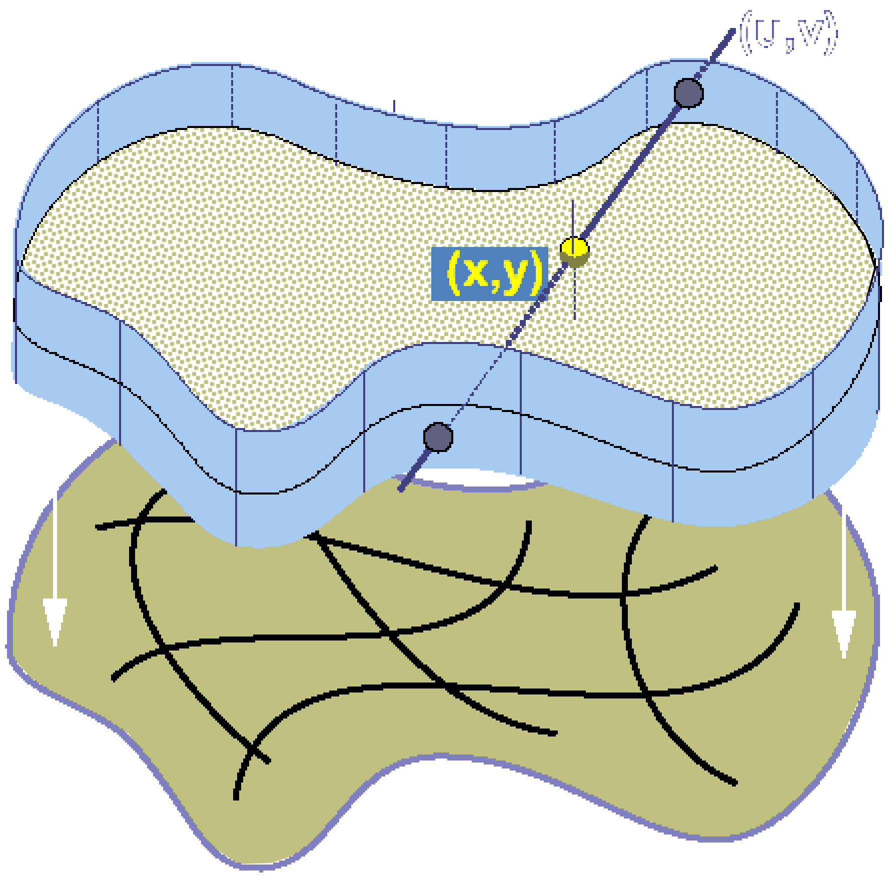
3.5D: Plenoptic Stitching



Ideally:

Dense Sampling

3.5D: Plenoptic Stitching



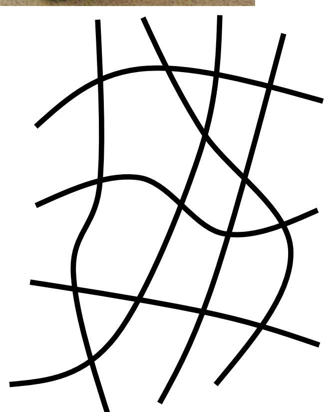
<u>Ideally</u>:

Dense Sampling

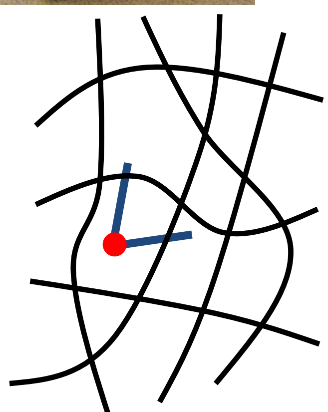
<u>Instead</u>:

Image Loop Sampling

3.5D: Plenoptic Stitching



3.5D: Plenoptic Stitching



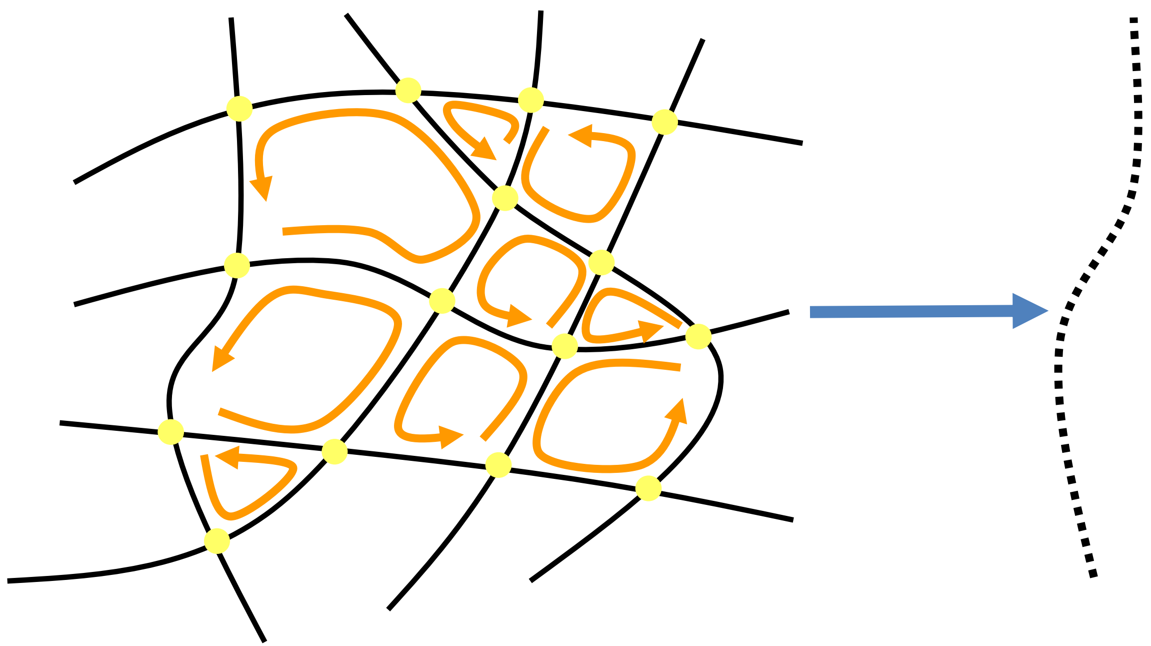
3.5D: Plenoptic Stitching

- Advantages:
  - **Gives horizontal and vertical parallax!**

- Problems:
  - How do we sample the environment with image loops?
  - How do we reconstruct the environment for a viewpoint within an image loop?

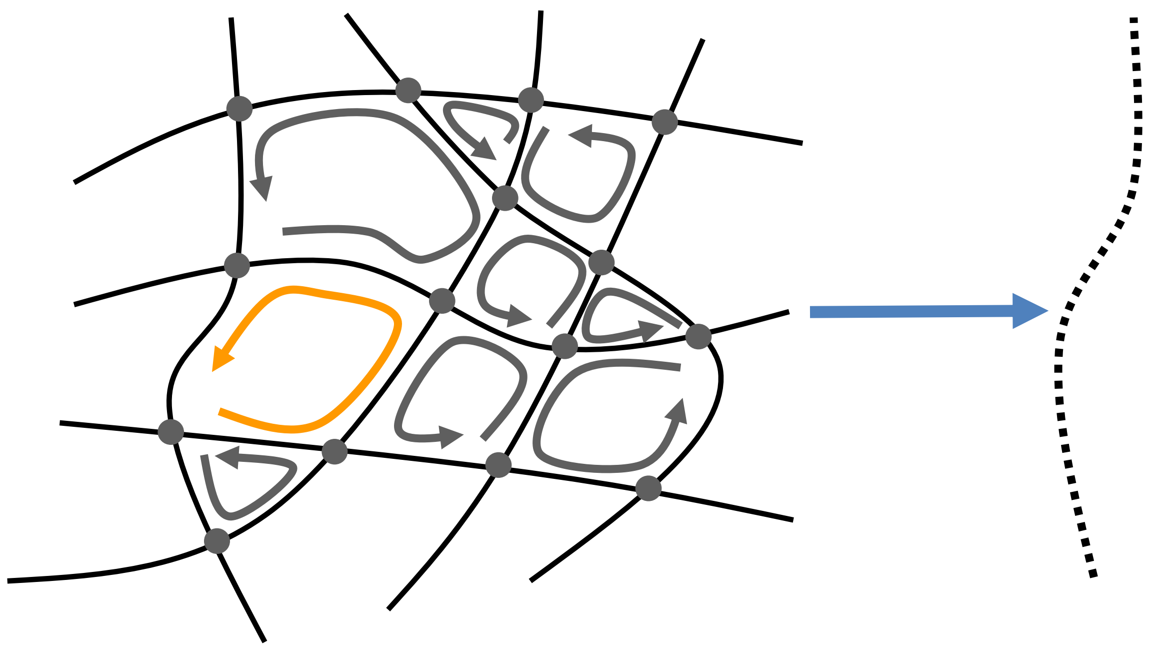
Capture

- Image loop creation
  - Each path is simply a sequence of images
  - We intersect the paths and determine image loops



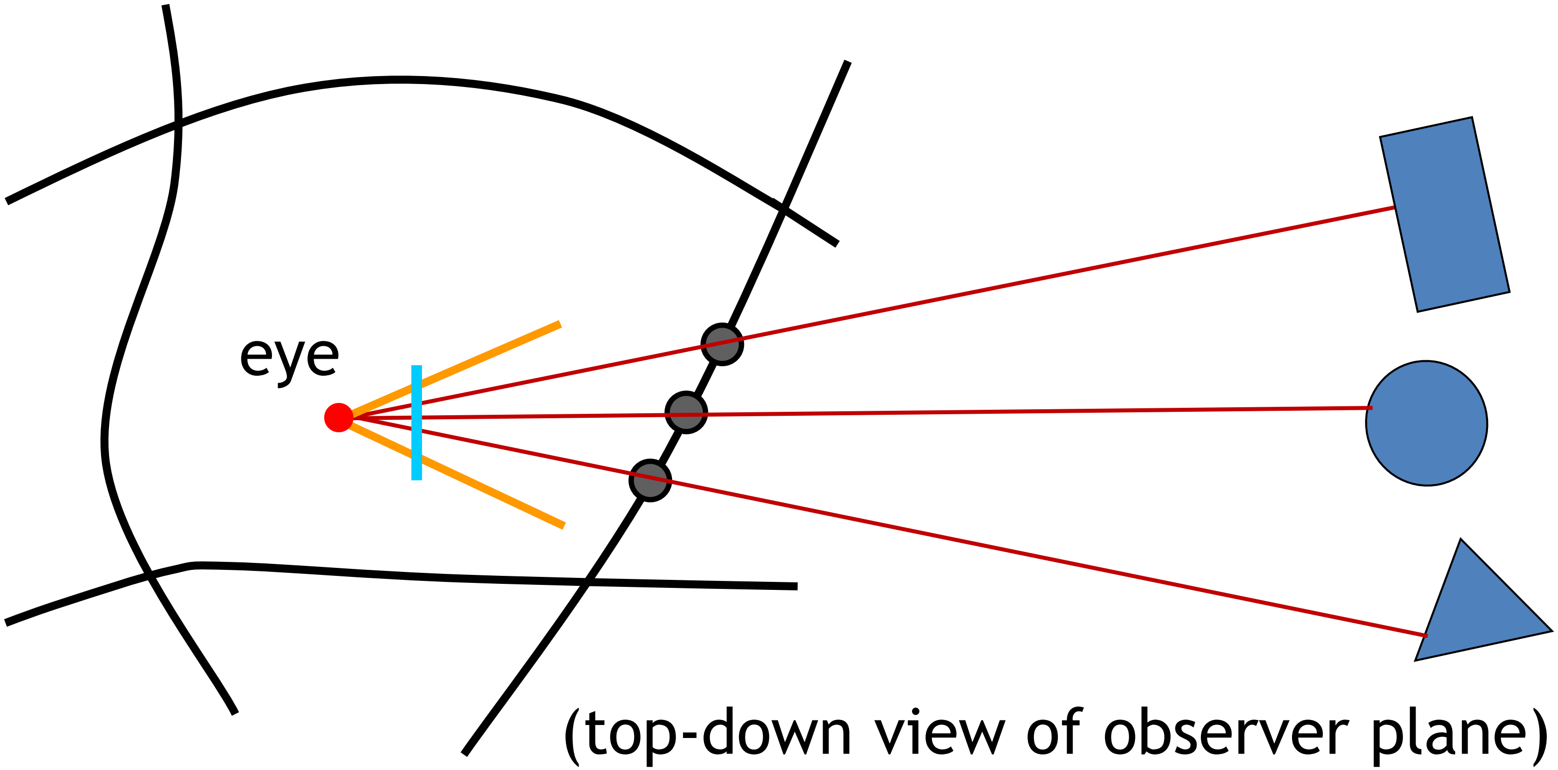
Capture

- Image loop creation
  - Each path is simply a sequence of images
  - We intersect the paths and determine image loops



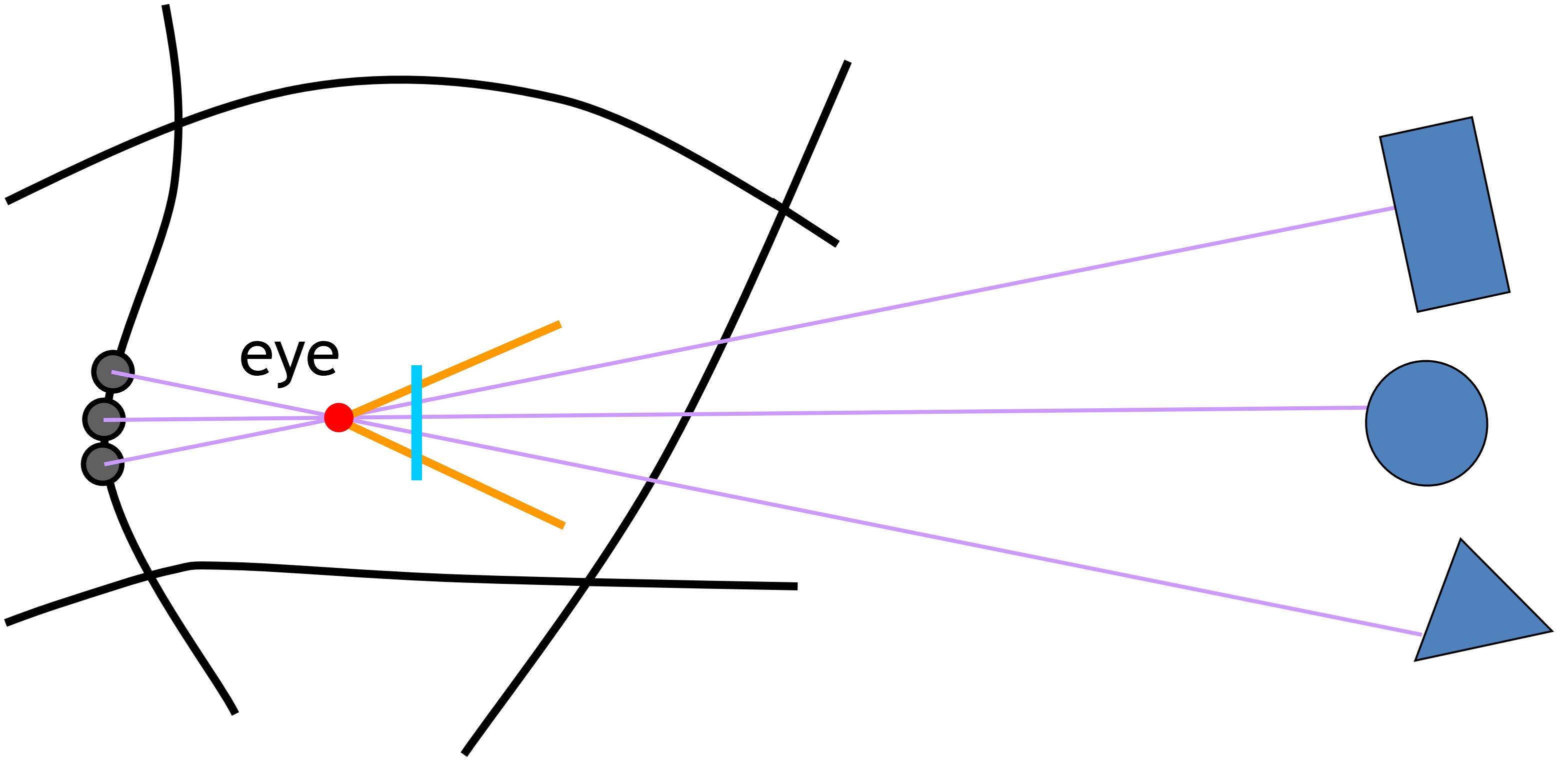
2D Reconstruction

- Extract from the front omnidirectional images the desired light rays



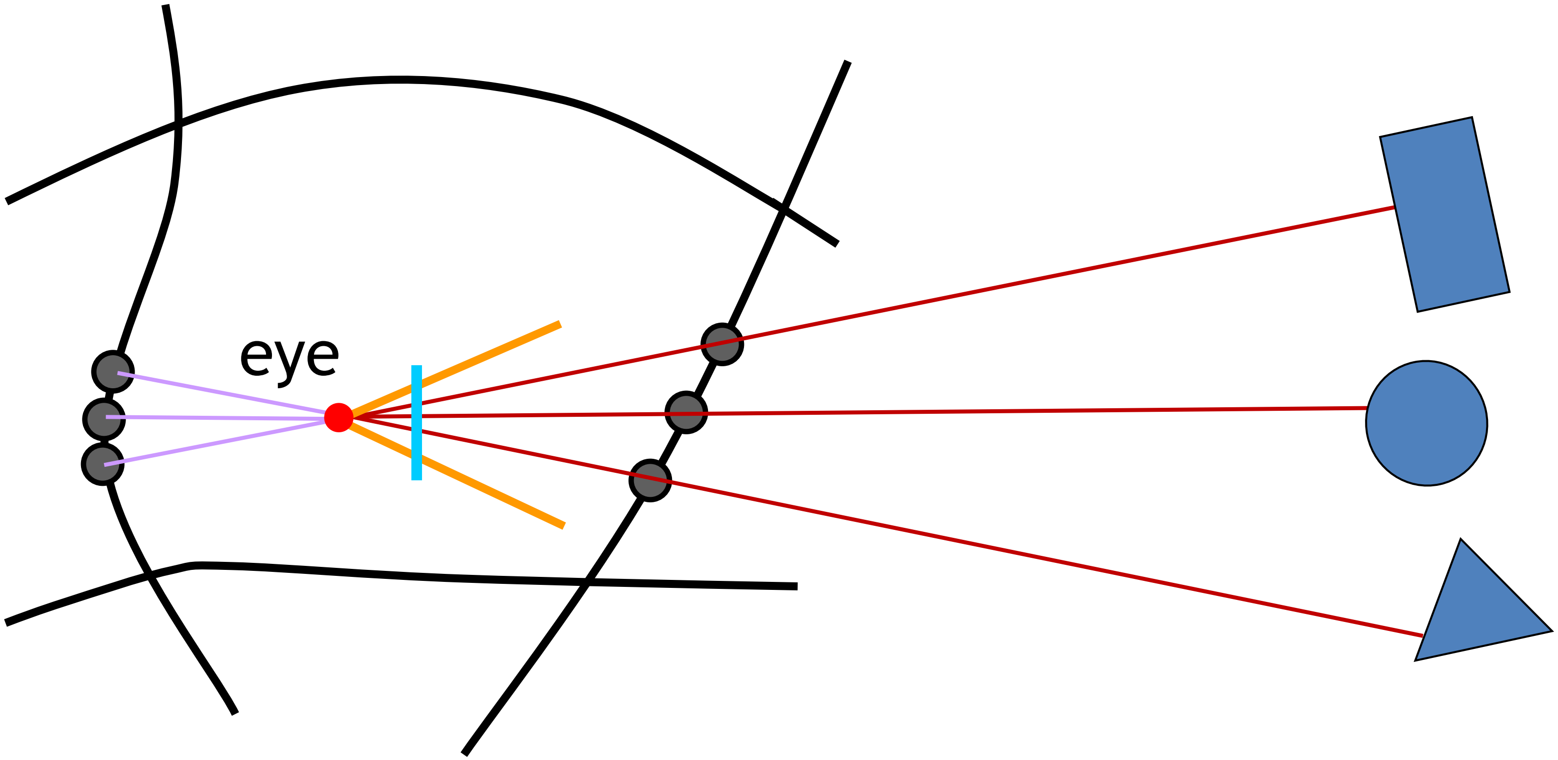
2D Reconstruction

- Extract from the back omnidirectional images the desired light rays



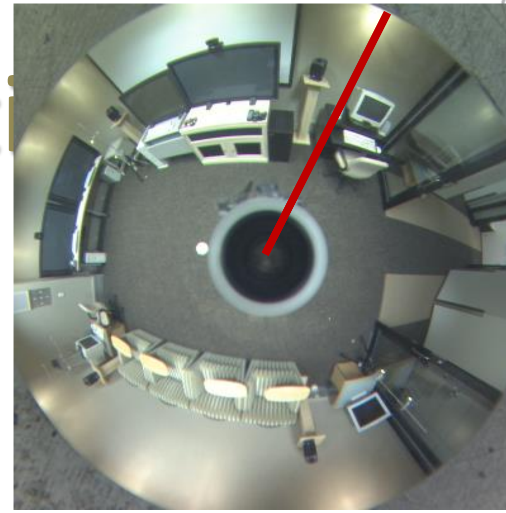
2D Reconstruction

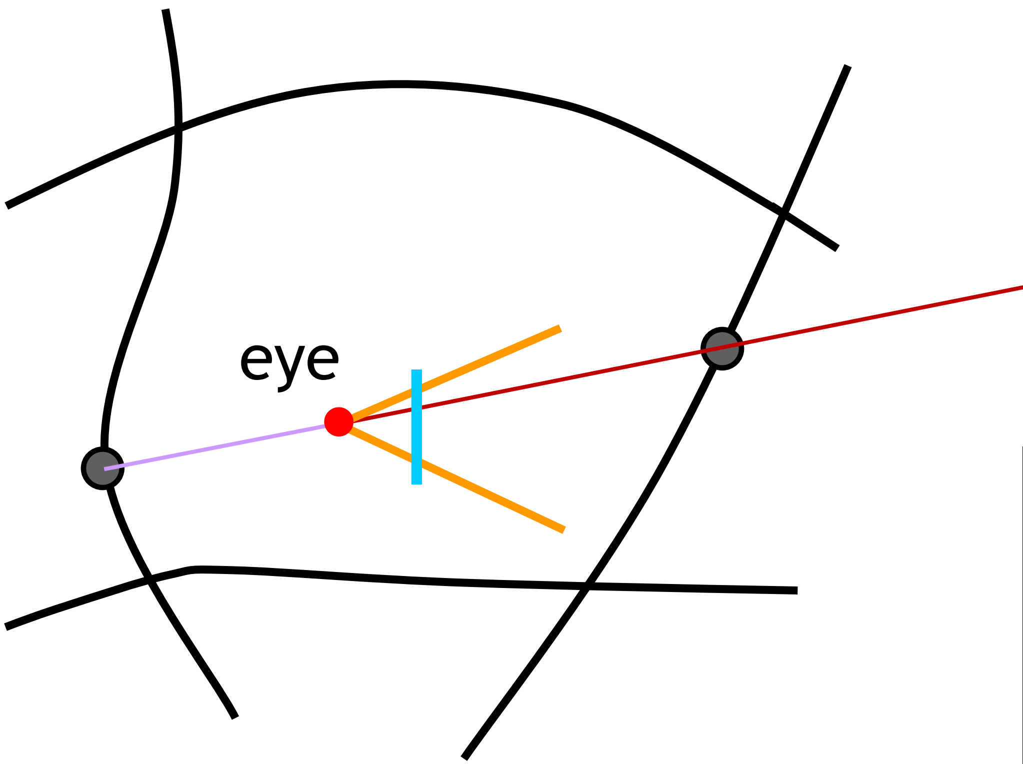
- Extract from the front and back omnidirectional images an average of the desired light rays

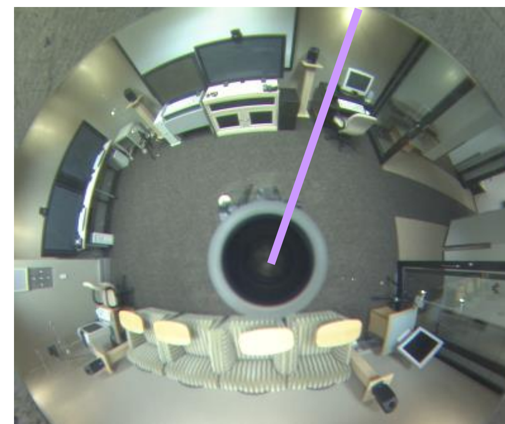


3D Reconstructi

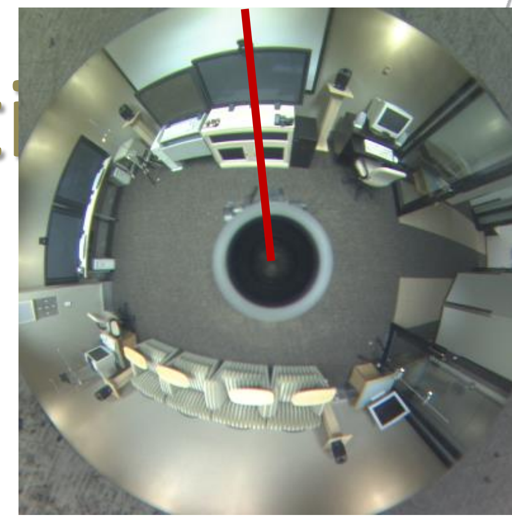
- Extract from the omnidirectional images an average of the **columns** (or radial lines) of light rays



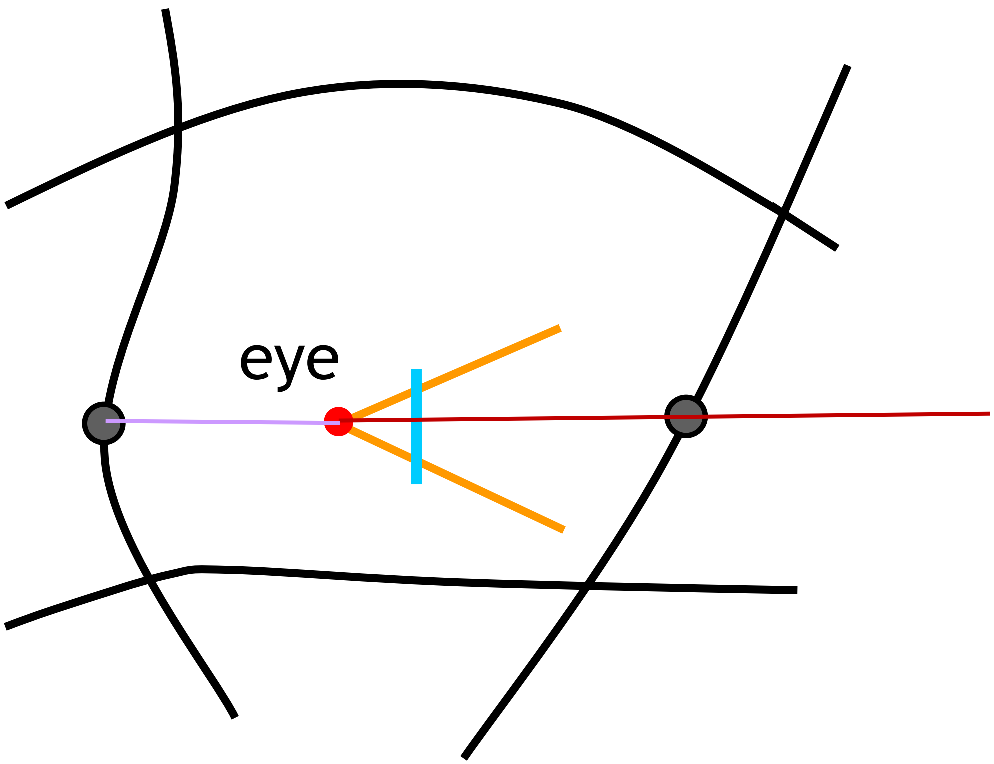


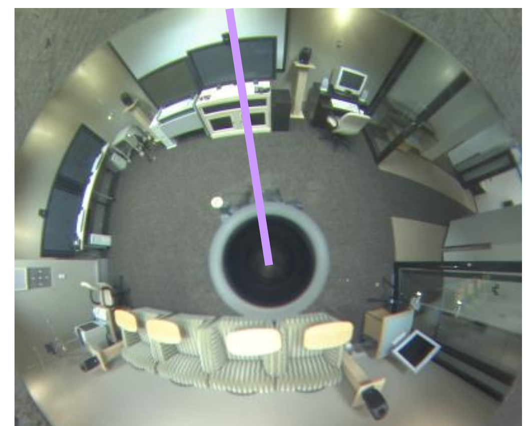


3D Reconstructi



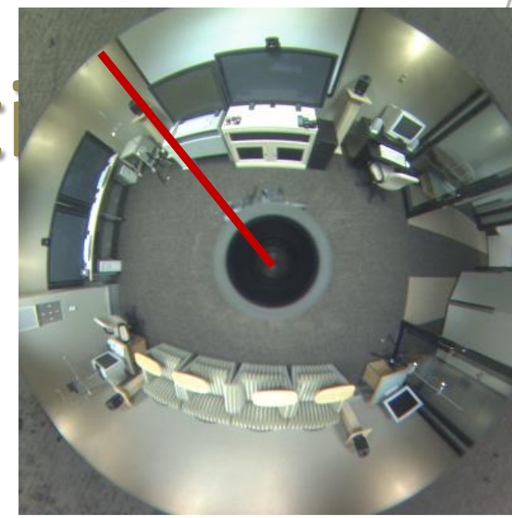
- Extract from the omnidirectional images an average of the **columns** (or radial lines) of light rays

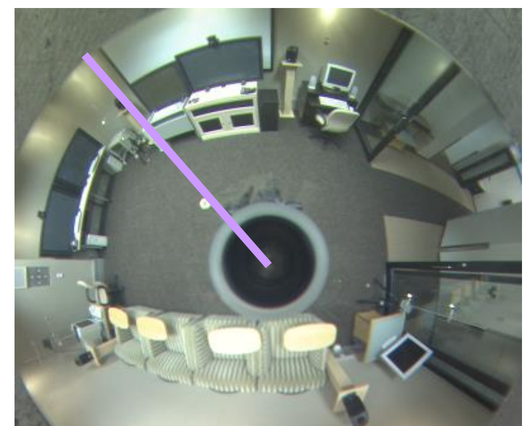


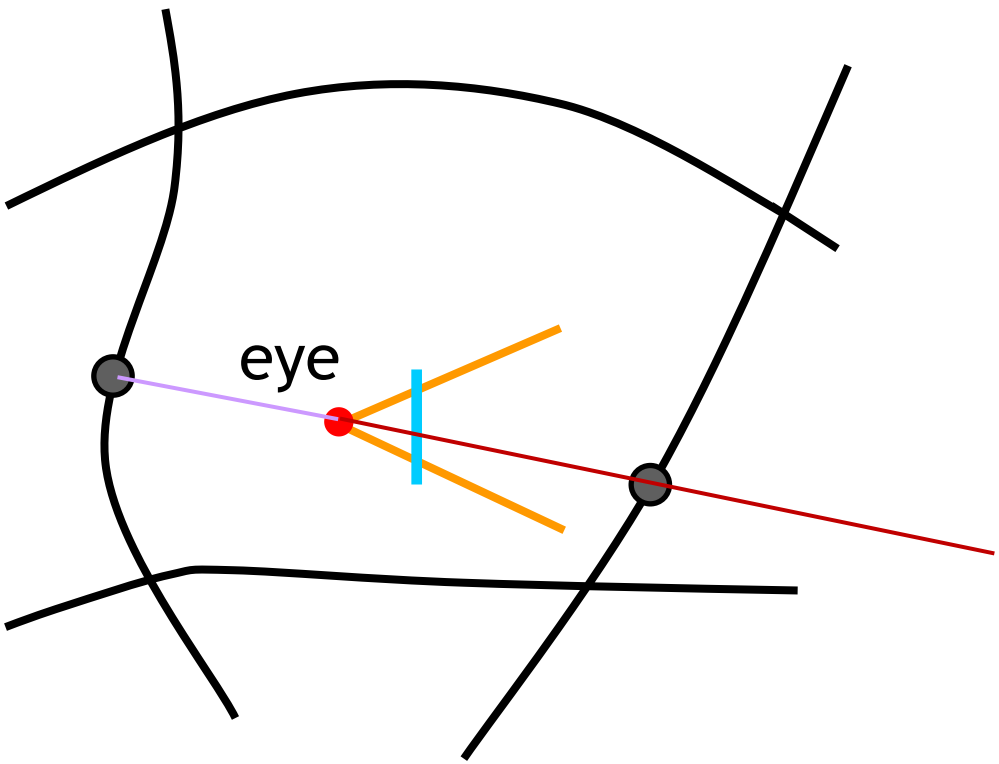


3D Reconstructi

- Extract from the omnidirectional images an average of the **columns** (or radial lines) of light rays

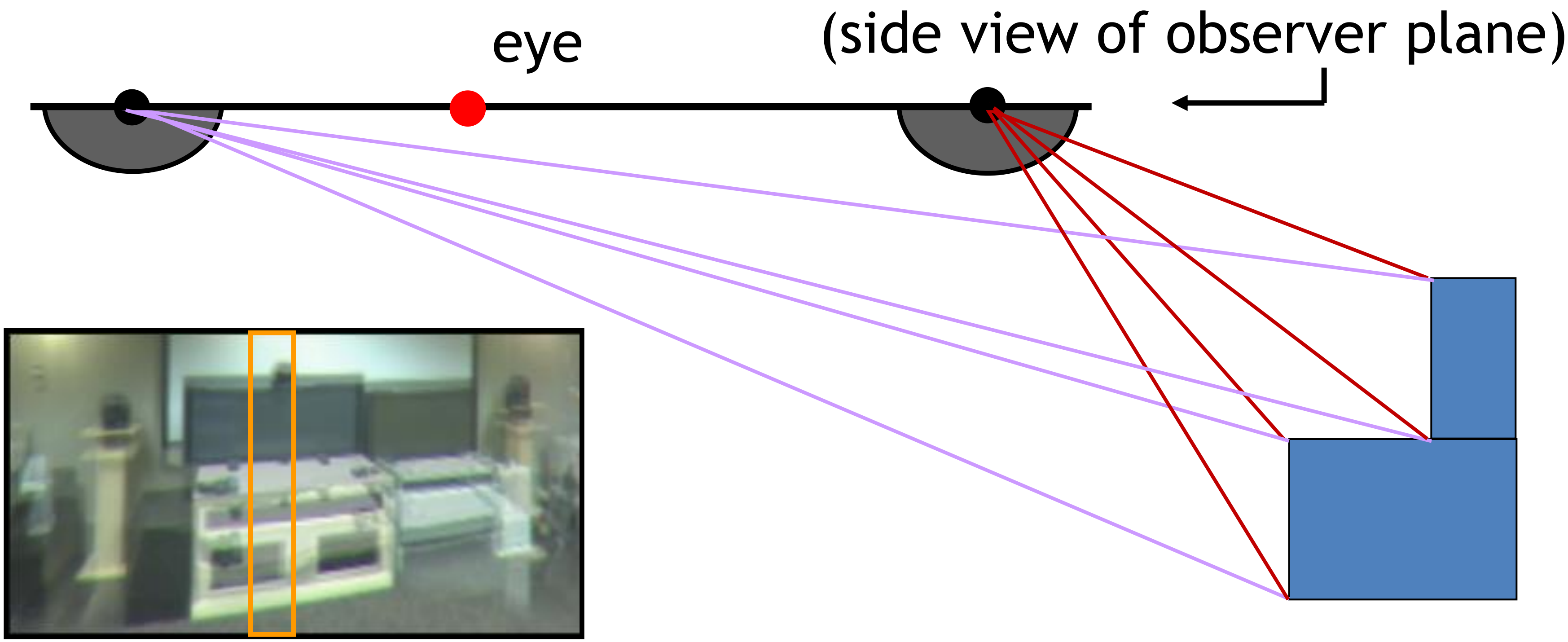






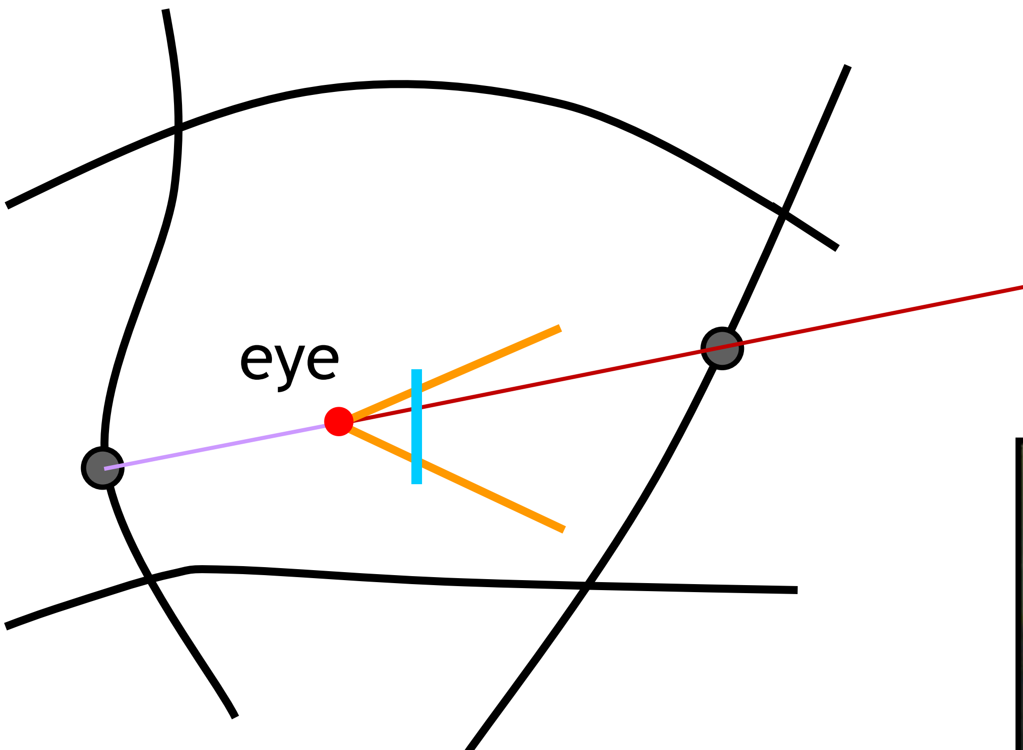
3D Reconstruction

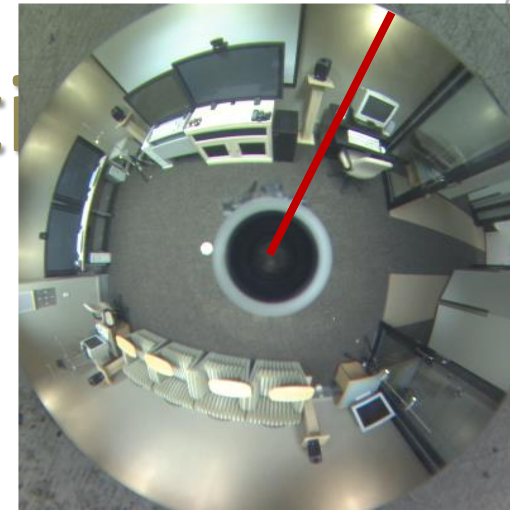
- Why does this not work?
  - Because we do not have samples of the desired **column** of light rays

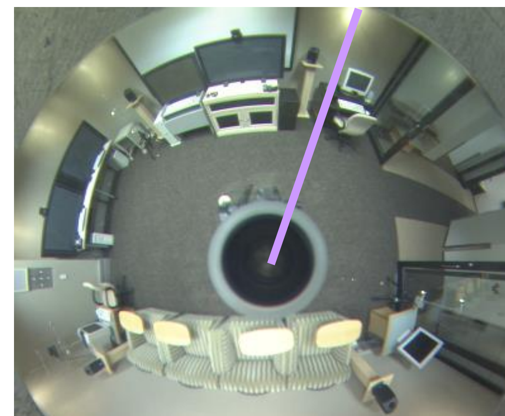


3D Reconstructi

- We perform a nonlinear interpolation between columns (or radial lines) of light rays

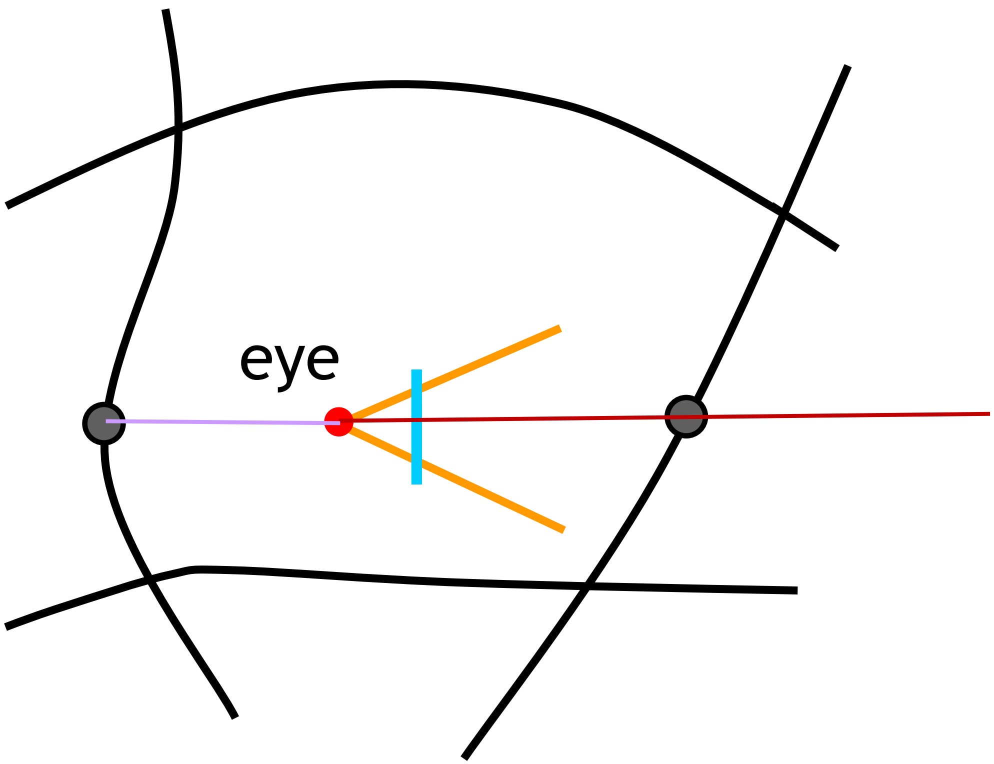


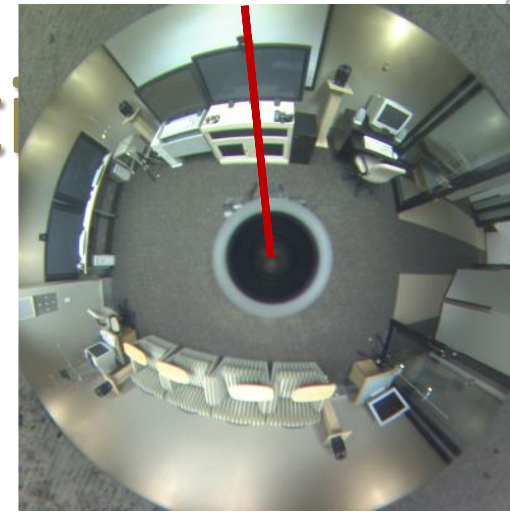


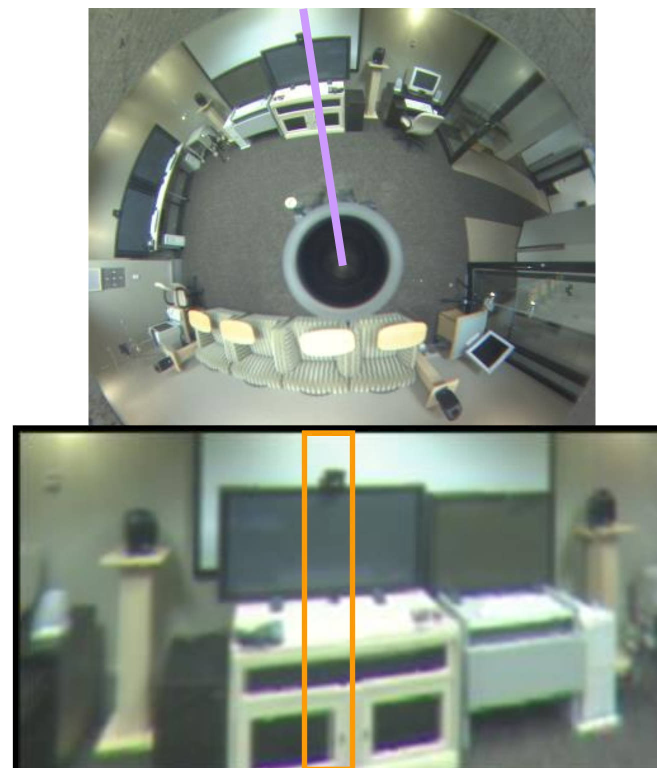


3D Reconstructi

- We perform a nonlinear interpolation between columns (or radial lines) of light rays

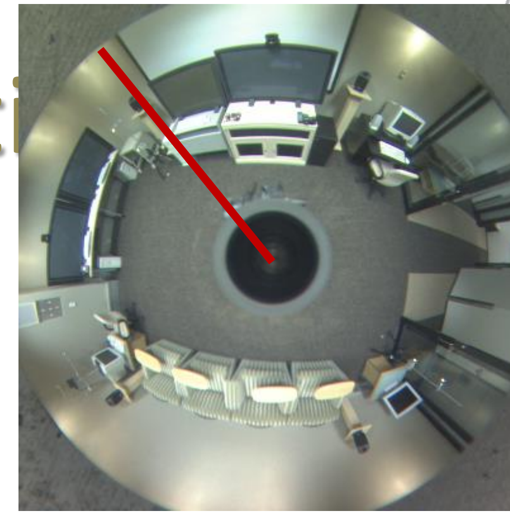


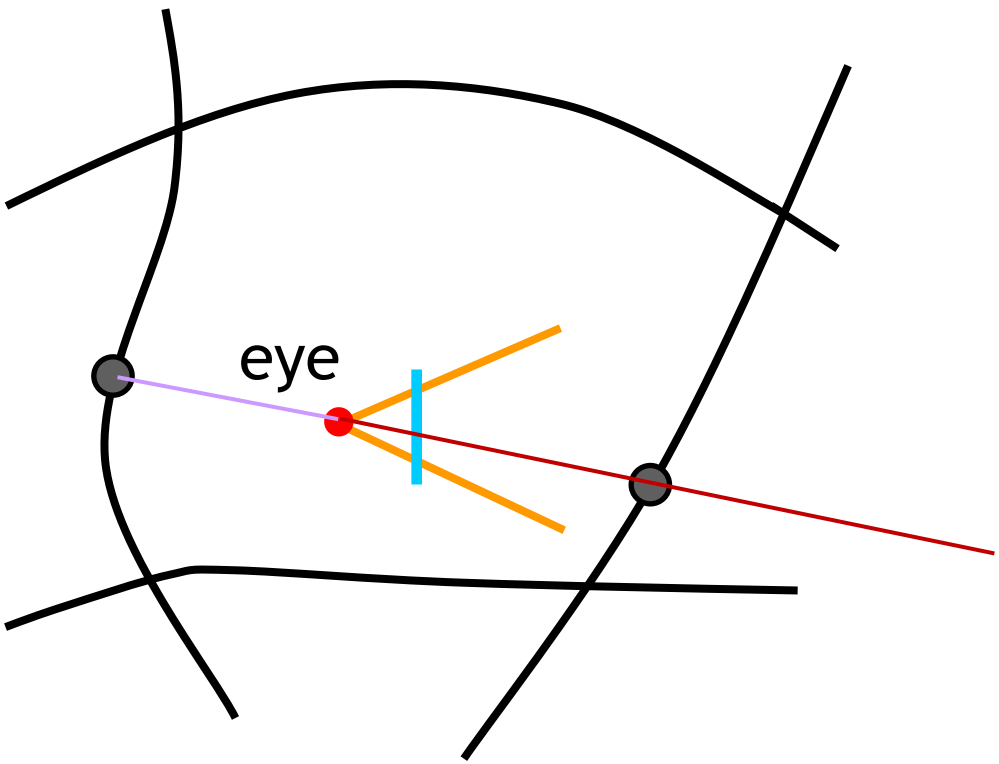


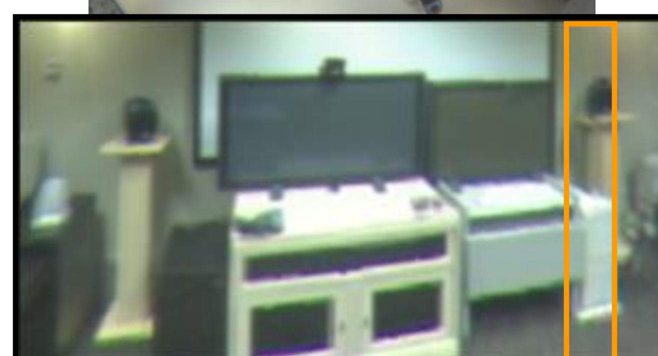


3D Reconstructi

- We perform a nonlinear interpolation between columns (or radial lines) of light rays

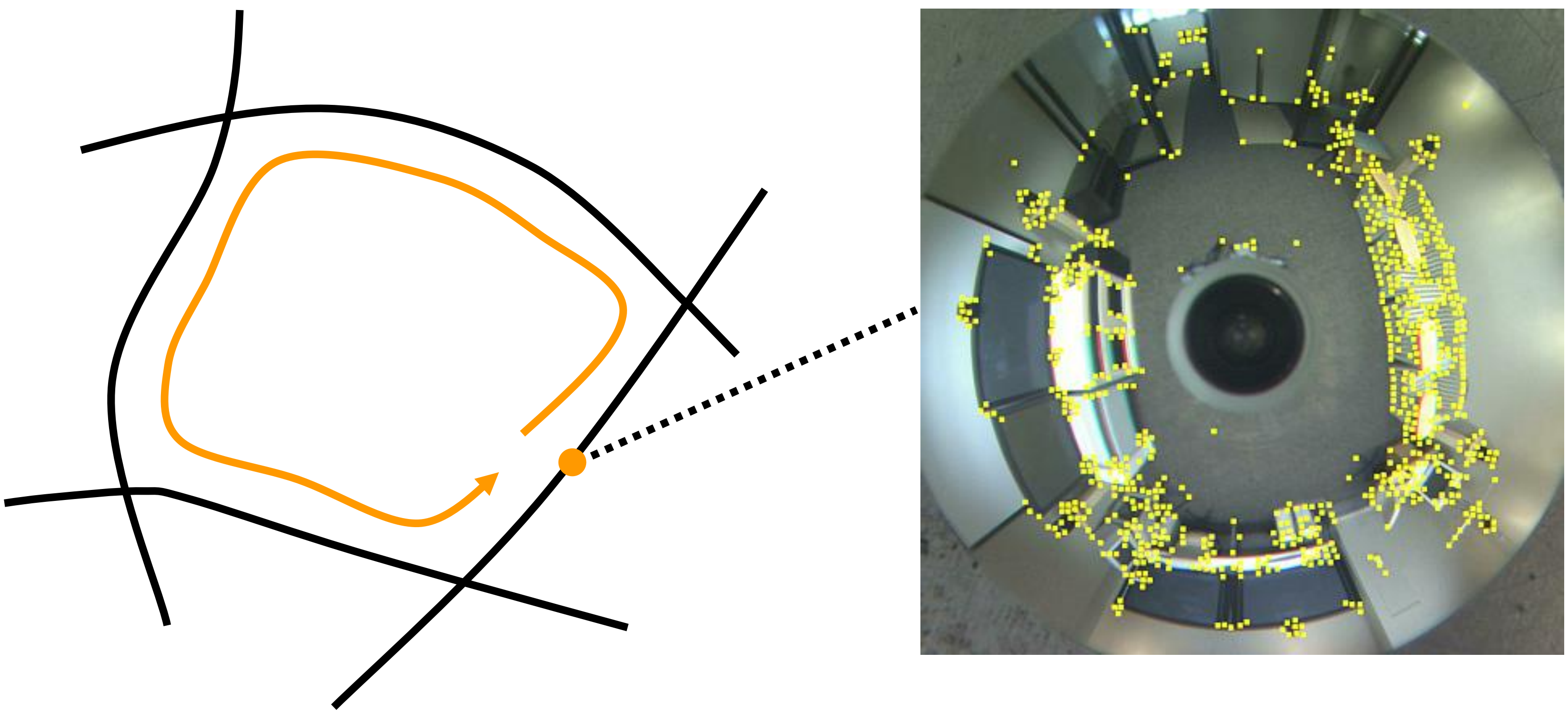






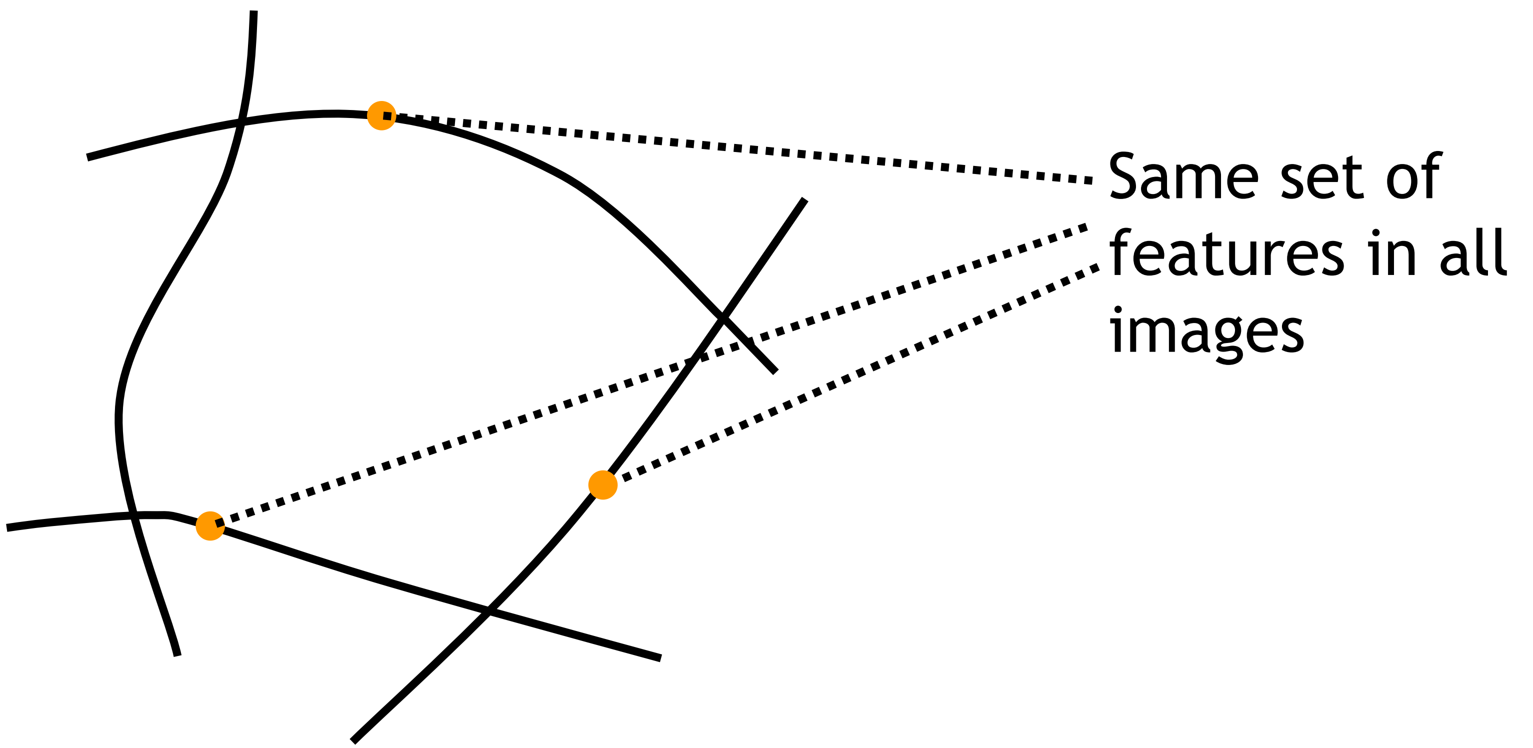
Reconstruction

- To interpolate, we first track features from frame-to-frame around the loop

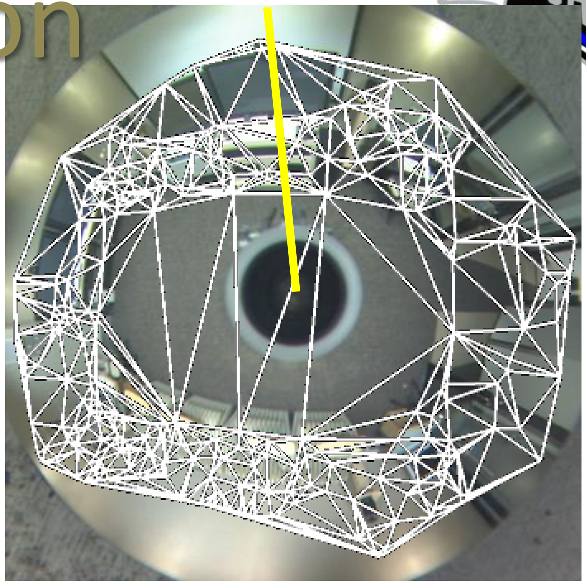


Reconstruction

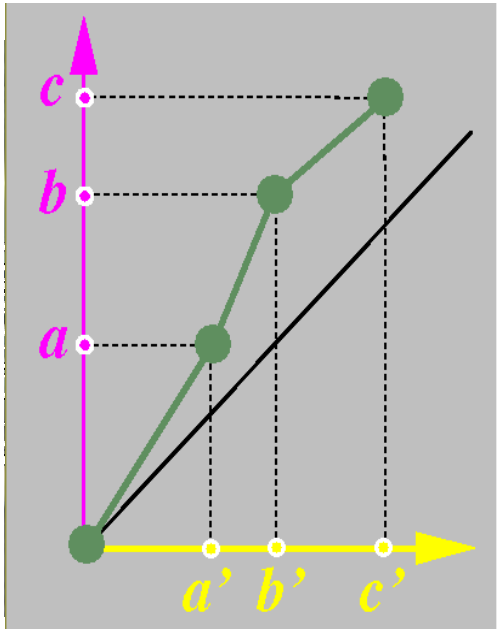
- Keeping only those features successfully tracked all the way around the loop



Reconstruction



- For each omnidirectional image pair, **triangulate** tracked features
- **Intersect** the triangulation edges with the radial lines
- Establish a **mapping** between segments of the radial line pairs and use this to warp pixels to the intermediate viewpoint



Implementation Issues

- Optimizations
  - Compensate for horizontal (angular) misalignment
  - Compensate for vertical misalignment
- Reconstruction Acceleration
  - Pre-compute radial line mappings
  - Optionally warp pixels using lower-resolution information
- Caches and Compression
  - Data compressed using JPEG and Lempel-Ziv
  - Three LRU caches: (1) compressed bitstreams, (2) uncompressed image subsets, (3) uncompressed mappings

Results: Examples

Reconstructions for novel viewpoints near the middle of several image loops within various environments (generated at 5-10 fps)

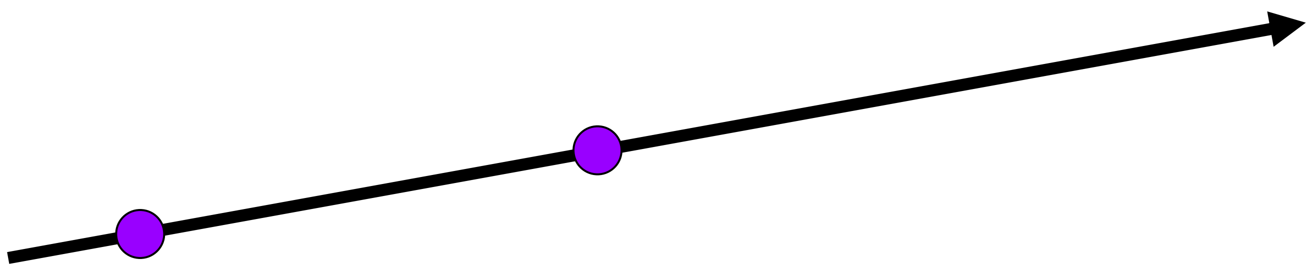
Results: Image Comparison

Reconstructed image for a viewpoint near the middle of an image loop (where reconstruction is most difficult)

Captured image from approximately the same viewpoint

4D: Rays in a Vacuum

- Infinite line
  - Assume light is constant (vacuum)



- 4D
  - 2D direction
  - 2D position
  - non-dispersive medium

Slide by Rick Szeliski and Michael Cohen

Only need plenoptic surface

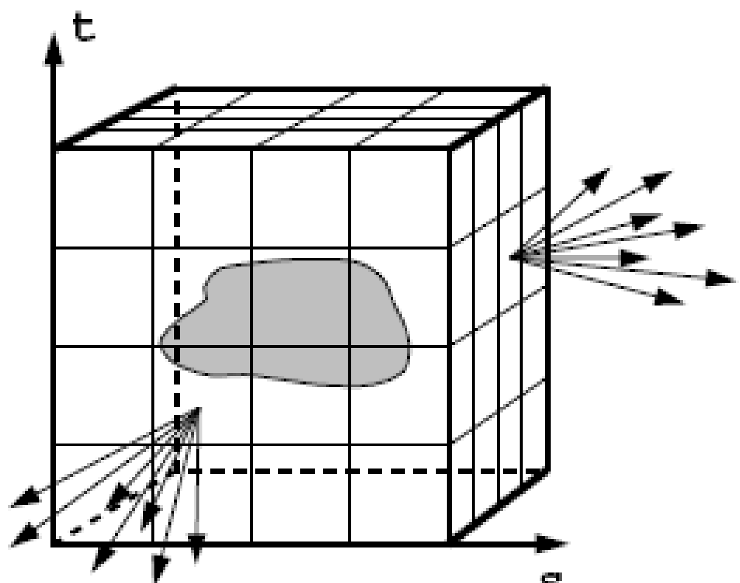
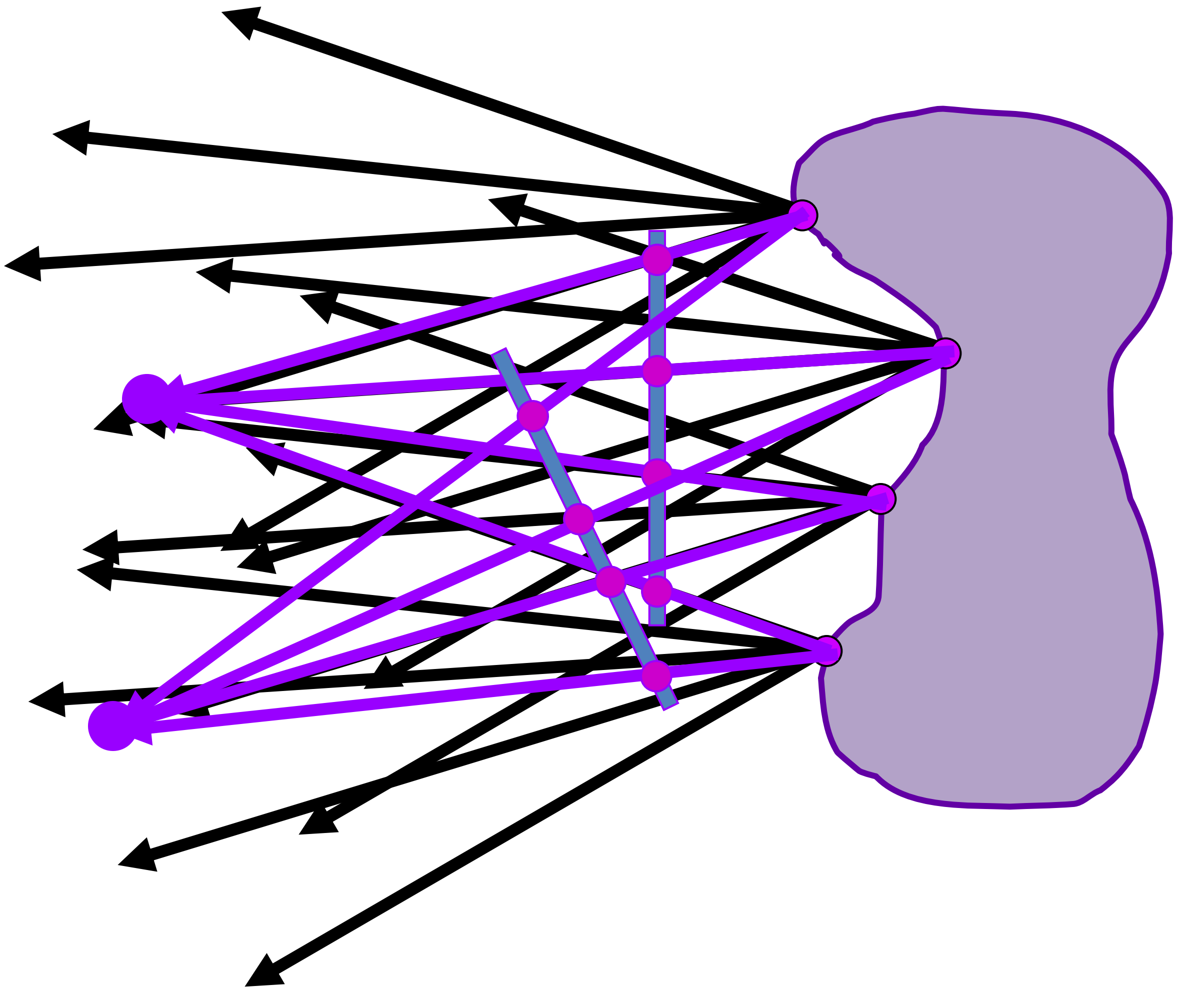


Figure 1: The surface of a cube holds all the radiance information due to the enclosed object.

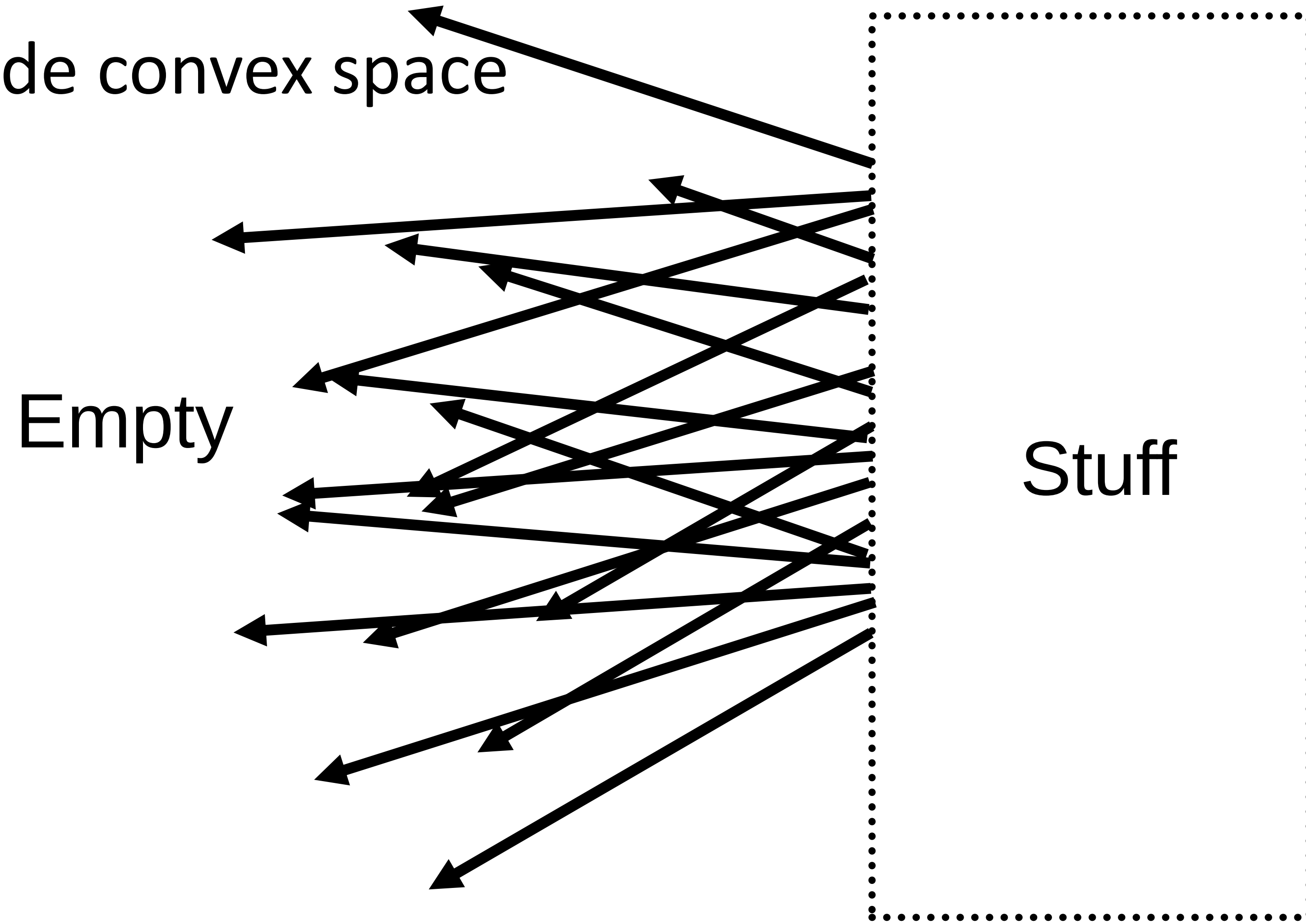
Synthesizing novel views



Slide by Rick Szeliski and Michael Cohen

Lumigraph / Lightfield

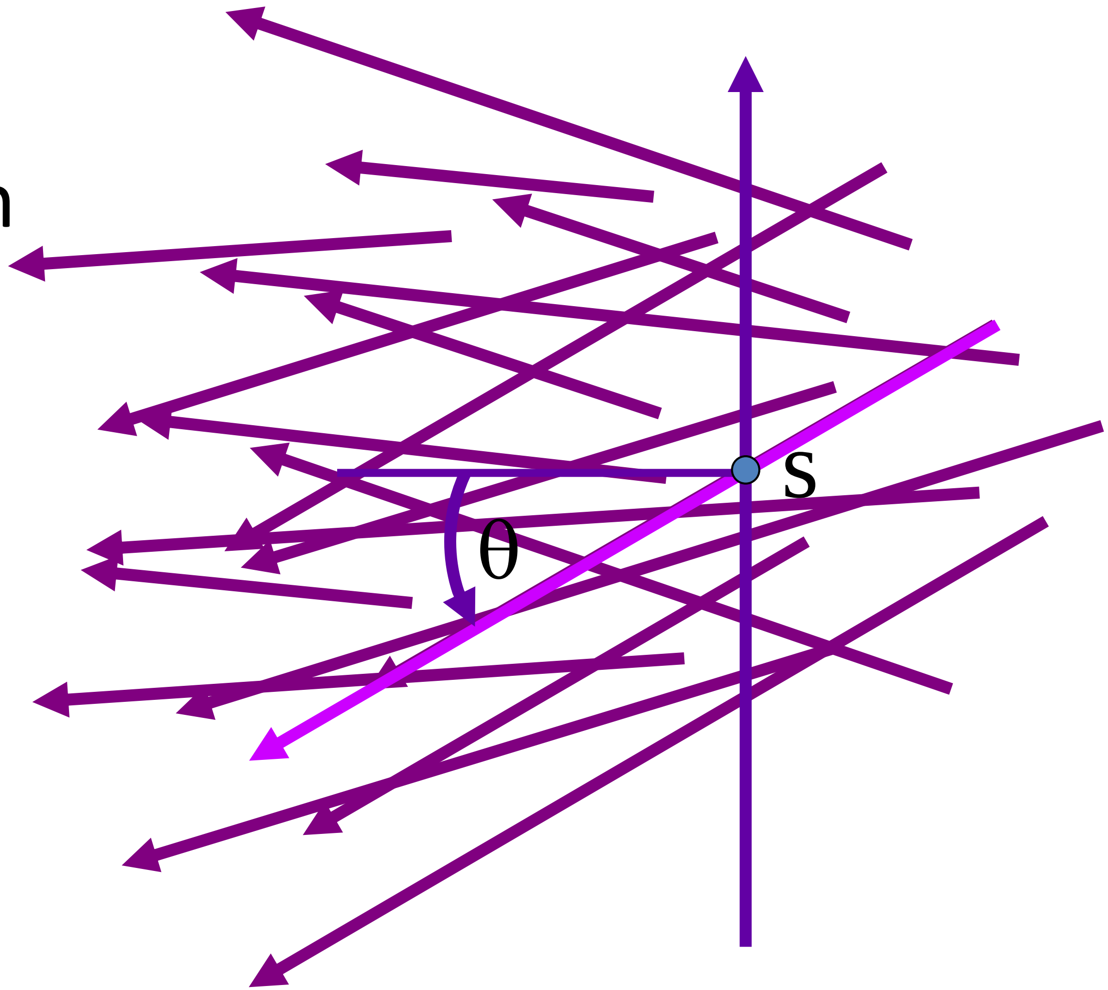
- Outside convex space
- 4D



Slide by Rick Szeliski and Michael Cohen

Lumigraph / Lightfield

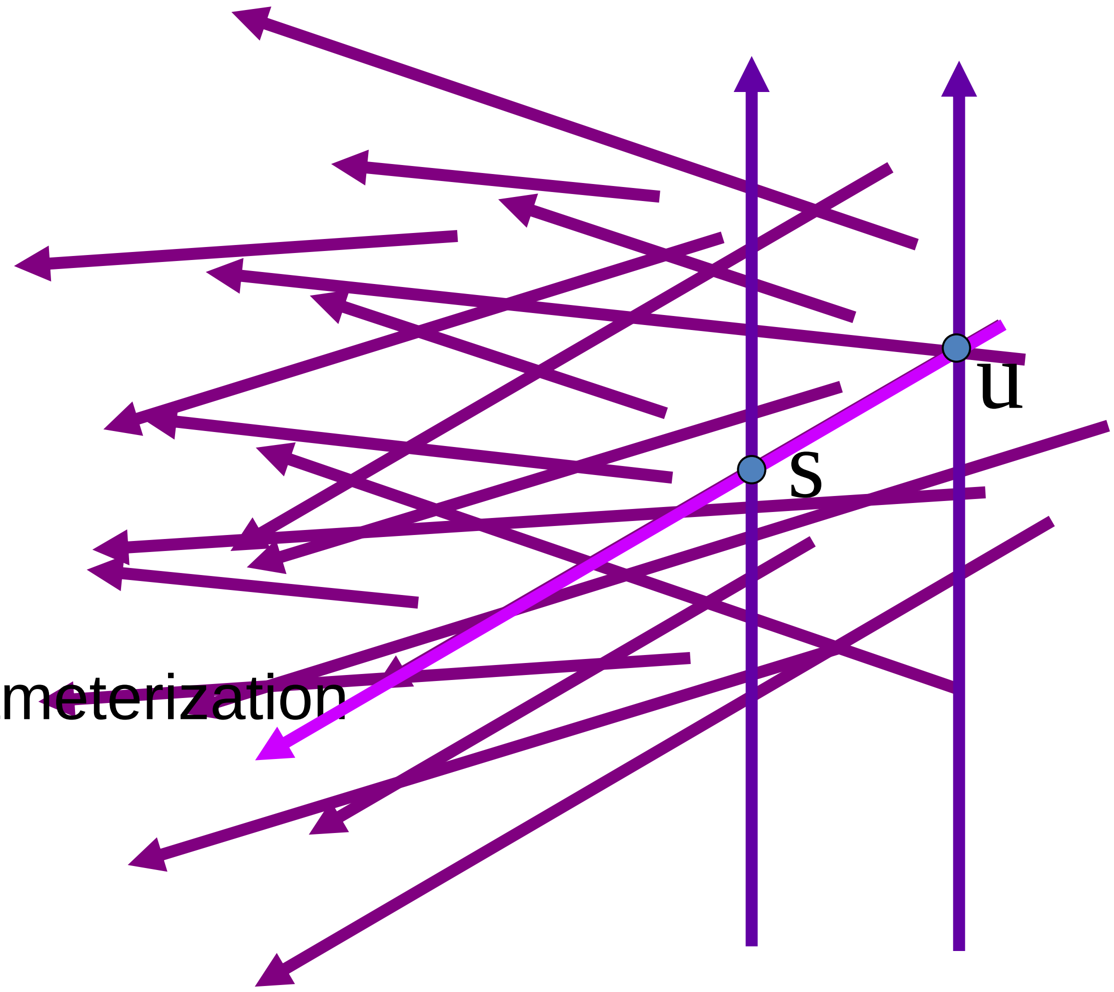
- 2D position
- 2D direction



Slide by Rick Szeliski and Michael Cohen

Lumigraph / Lightfield

- 2D position
- 2D position

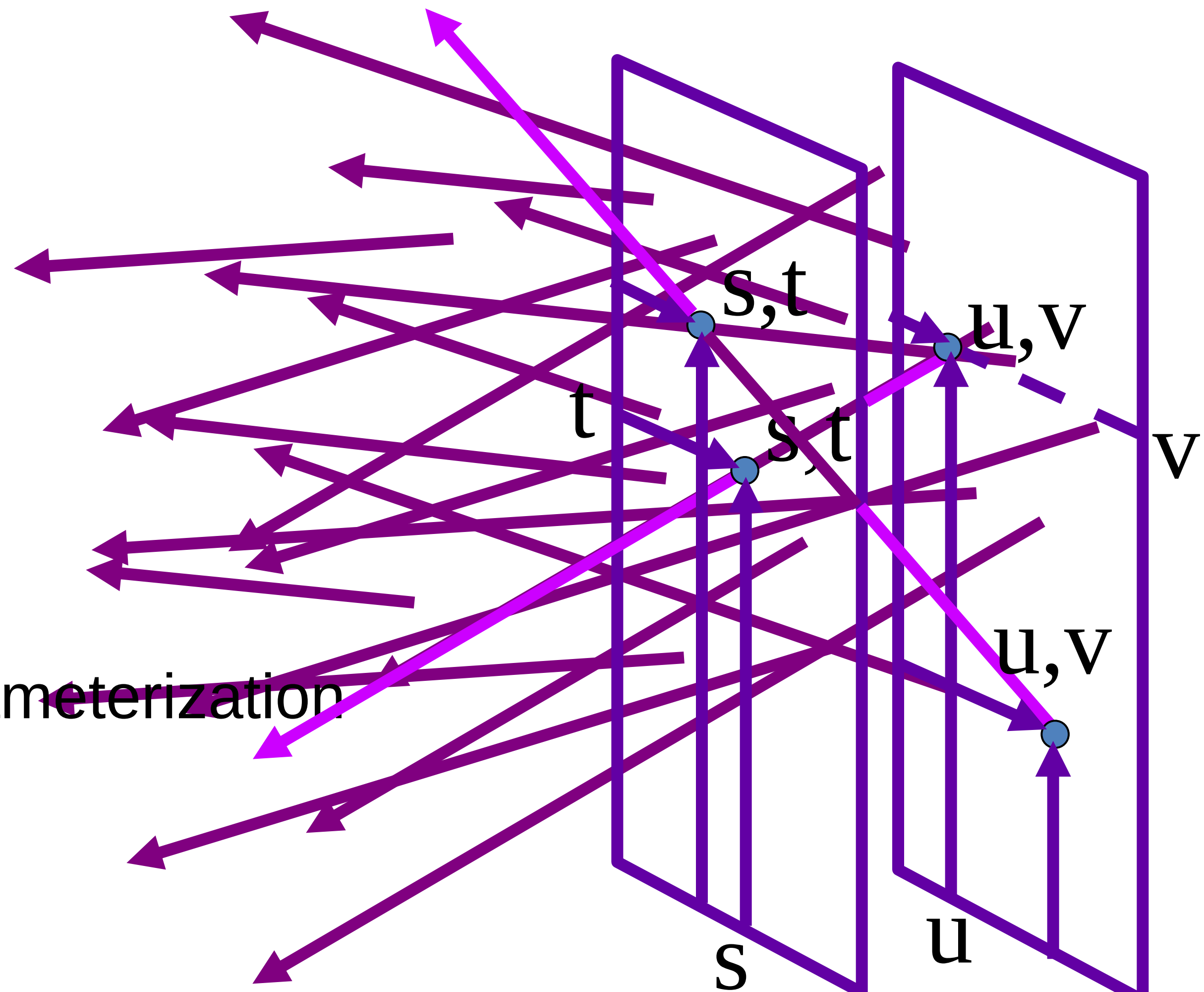


- 2 plane parameterization

Slide by Rick Szeliski and Michael Cohen

Lumigraph / Lightfield

- 2D position
- 2D position

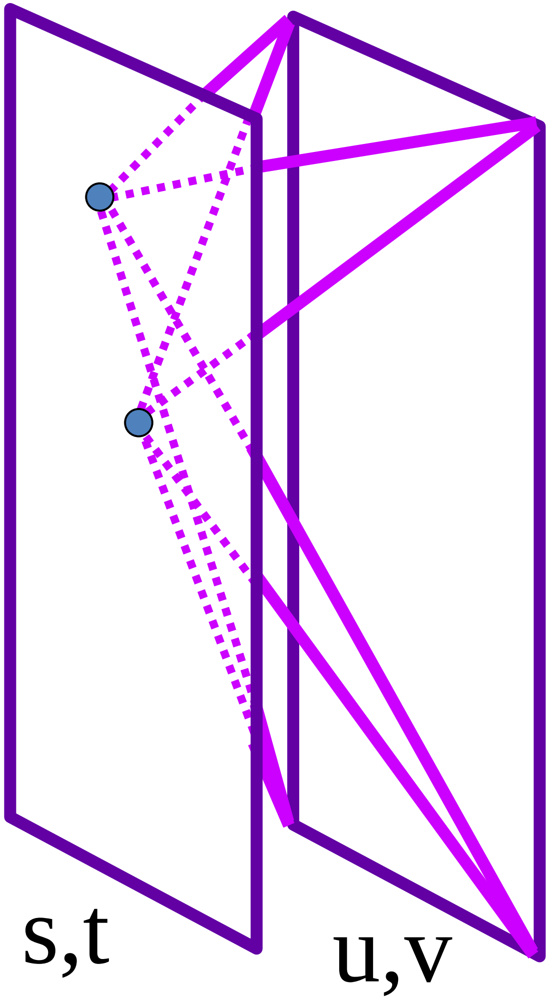


- 2 plane parameterization

Slide by Rick Szeliski and Michael Cohen

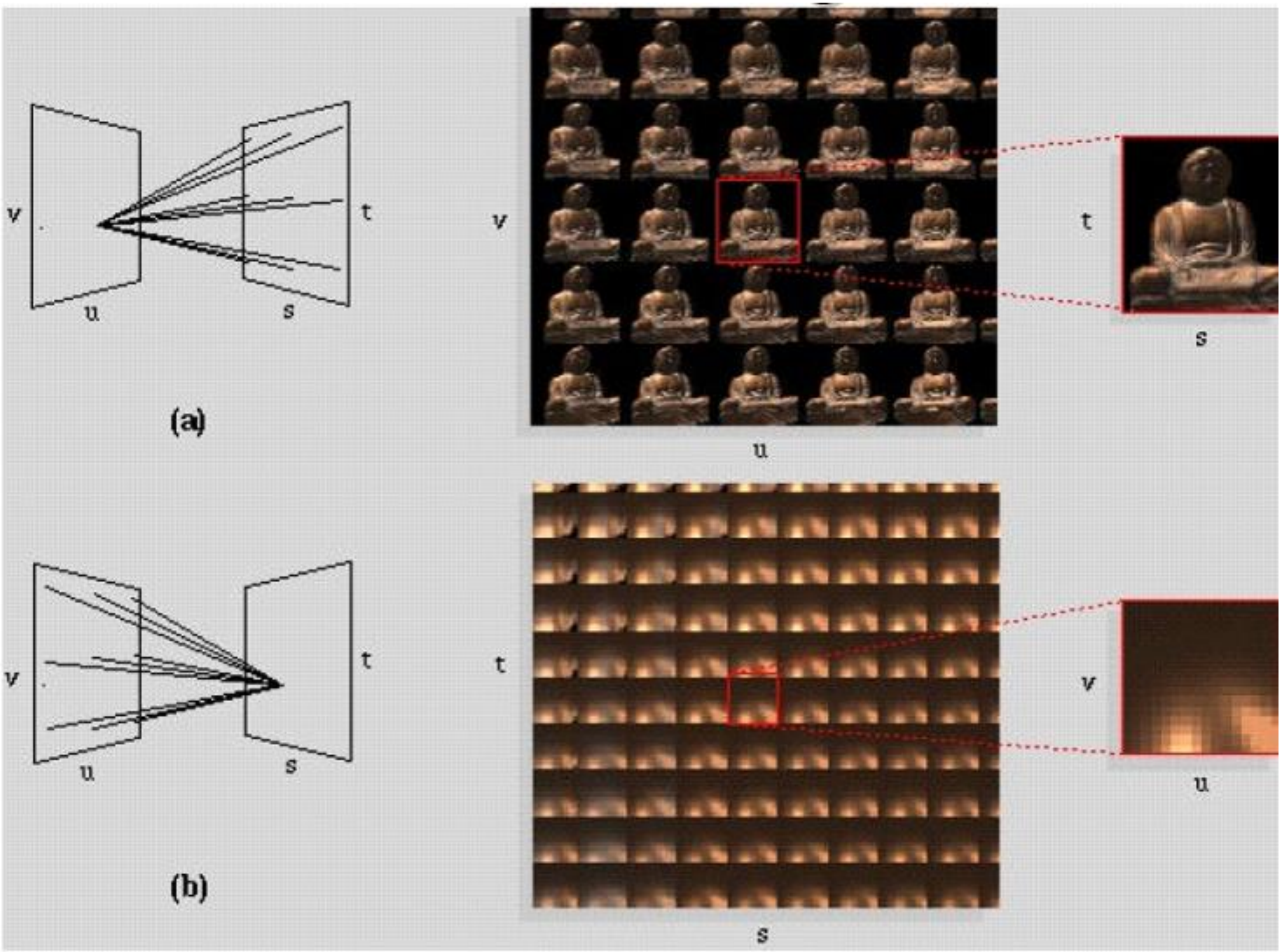
Lumigraph / Lightfield

- Hold s,t constant
- Let u,v vary
- An image



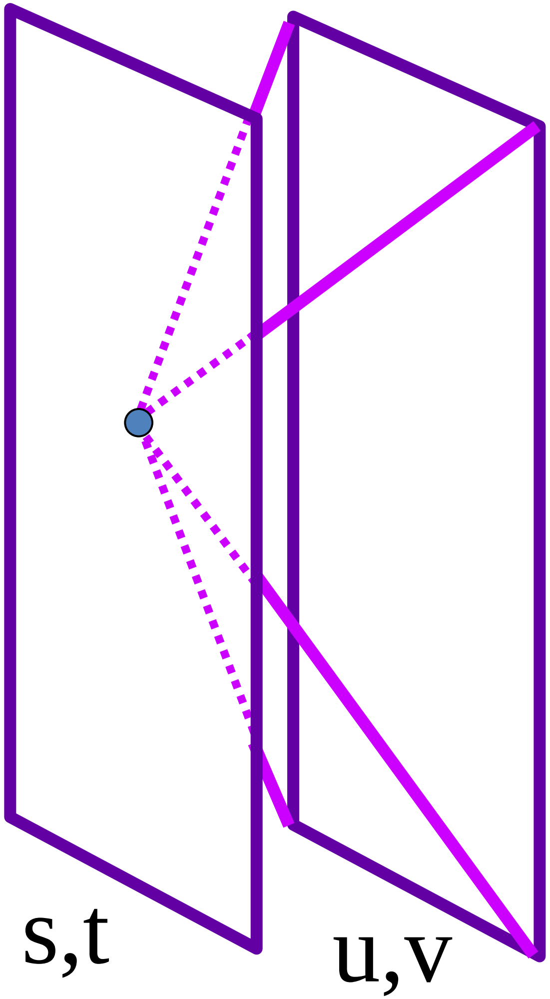
Slide by Rick Szeliski and Michael Cohen

Lumigraph / Lightfield



Lightfield - Capture

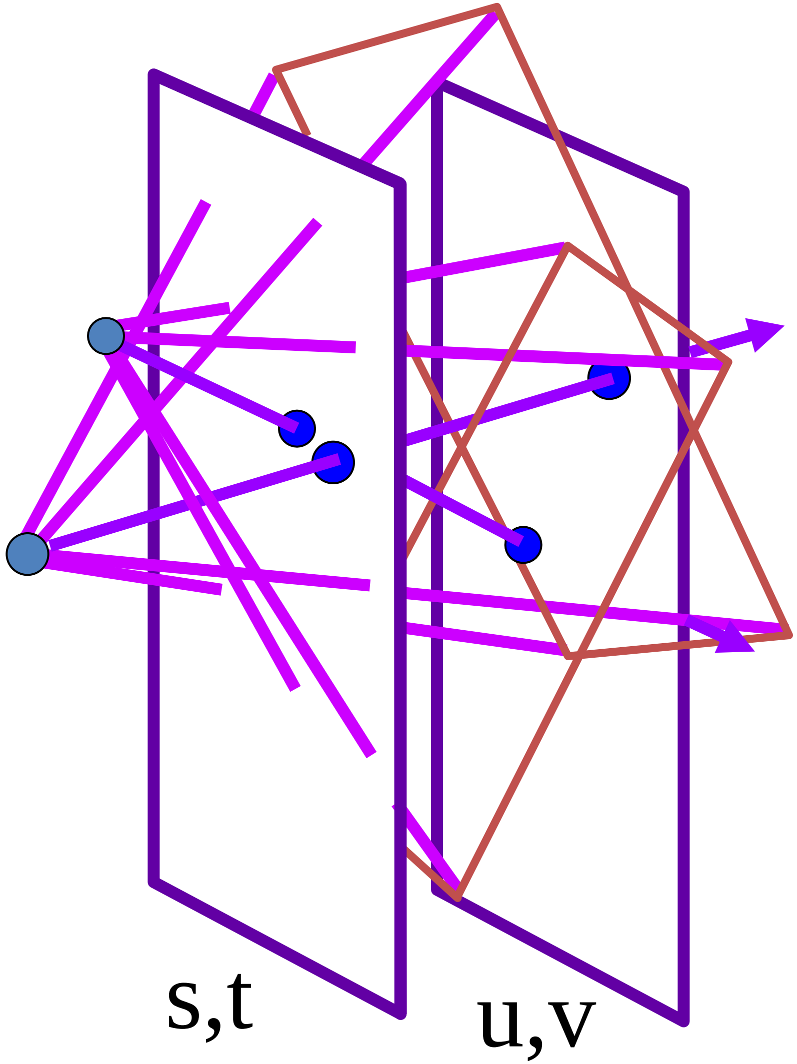
- Idea 1
  - Move camera carefully over s,t plane
  - Gantry
    - = Lightfield paper



Slide by Rick Szeliski and Michael Cohen

Lumigraph - Capture

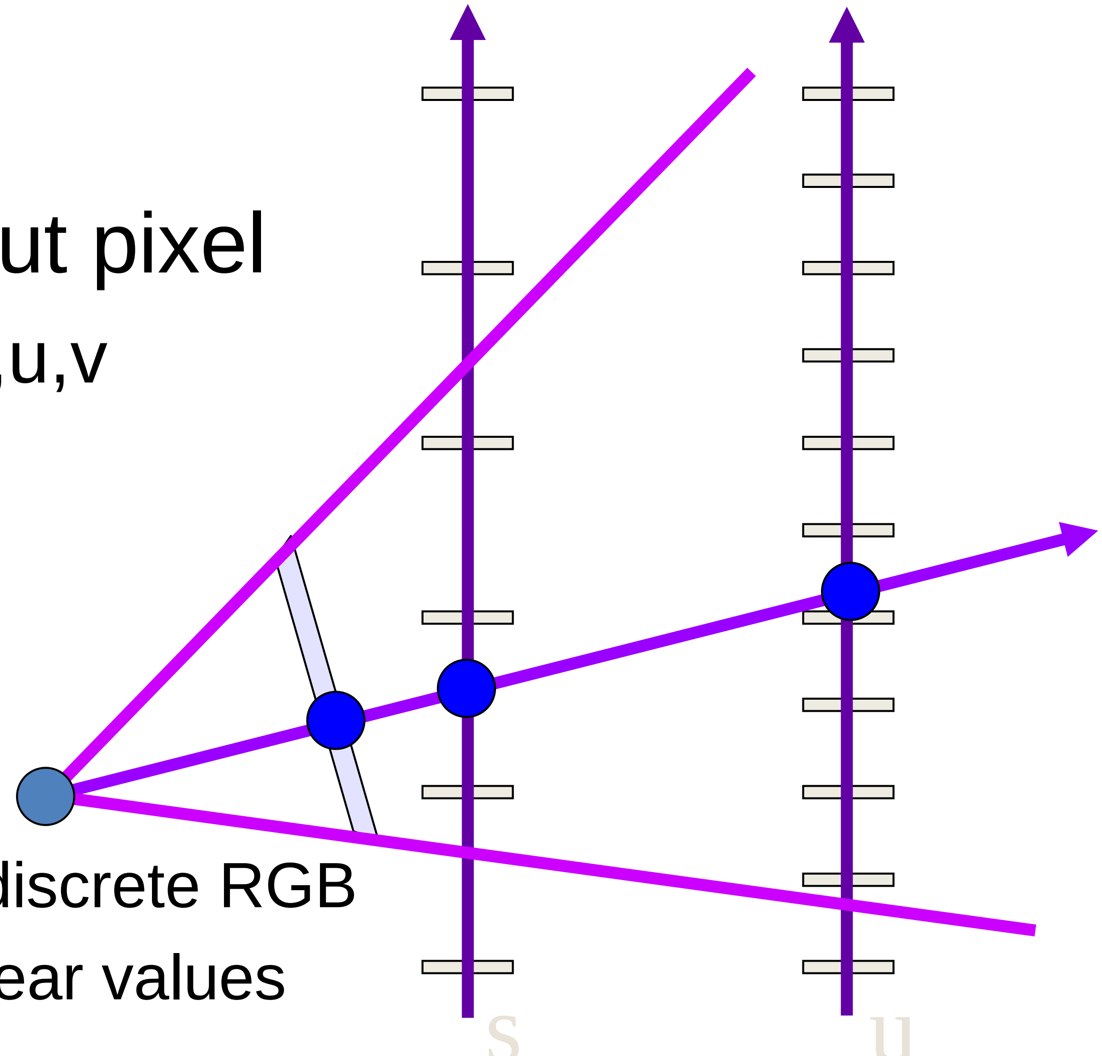
- Idea 2
  - Move camera anywhere
  - Rebinning
    - = Lumigraph paper



Slide by Rick Szeliski and Michael Cohen

Rendering

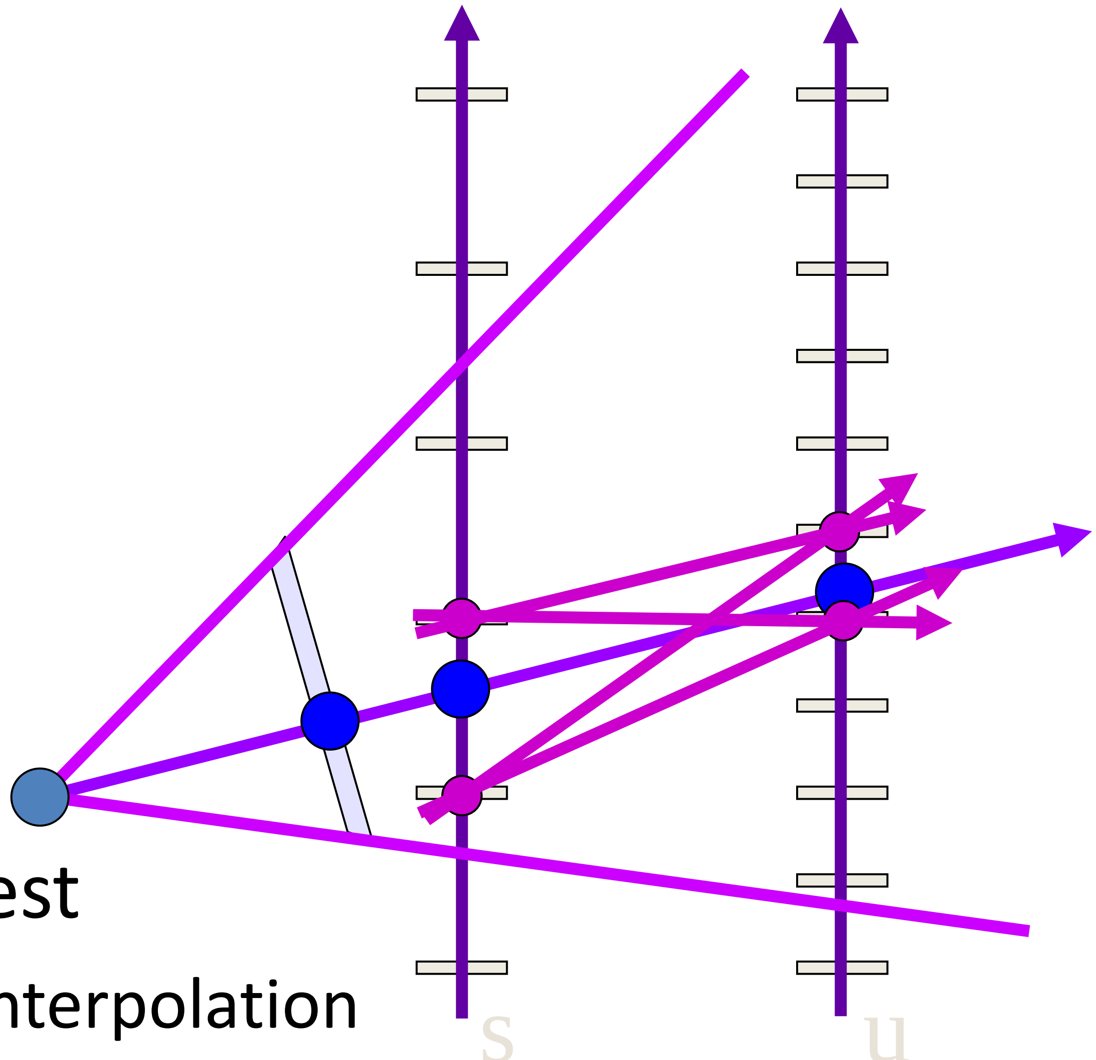
- For each output pixel
  - determine s,t,u,v
  - either
    - use closest discrete RGB
    - interpolate near values



Slide by Rick Szeliski and Michael Cohen

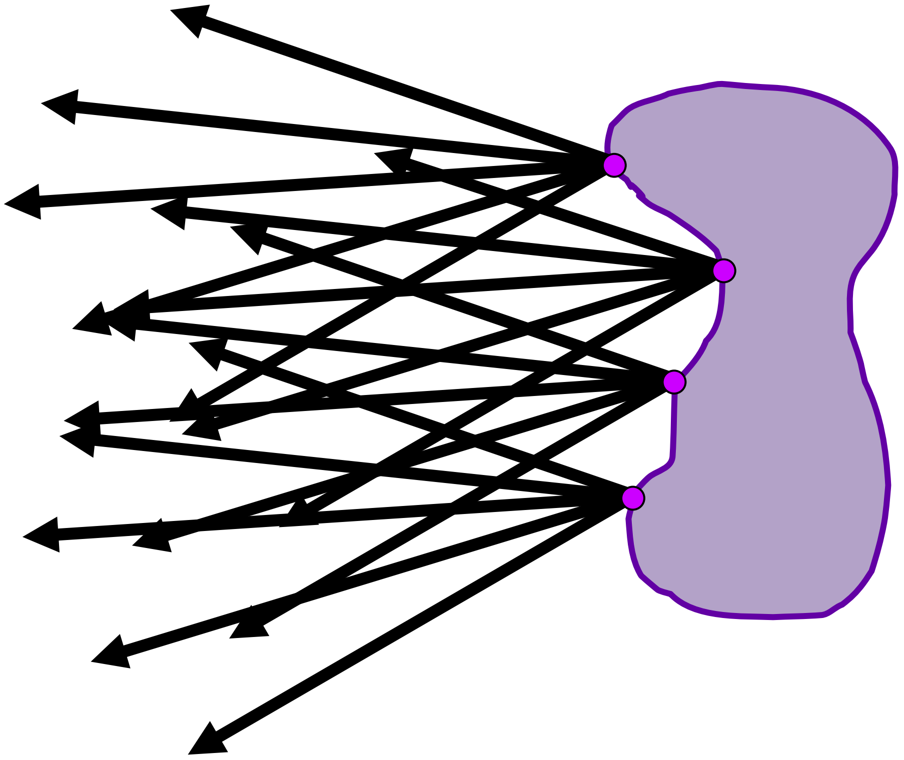
Rendering

- Nearest
  - closest s
  - closest u
  - draw it
- Blend 16 nearest
  - quadrilinear interpolation



4D: Surface Lightfields

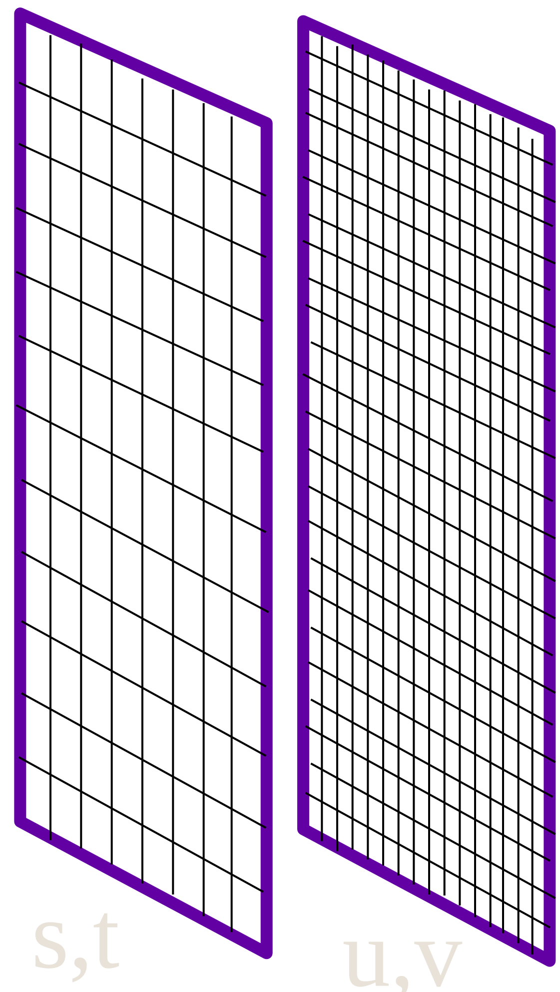
- Turn 4D parameterization around
- Leverage coherence



4D: Surface Lightfields

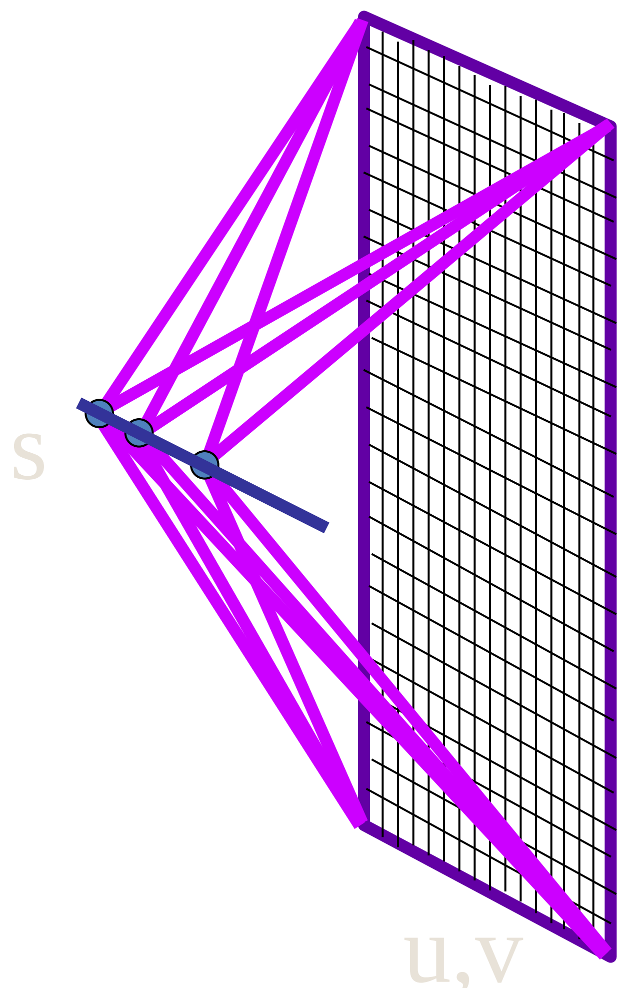
(3D: Lumigraph/Lightfield)

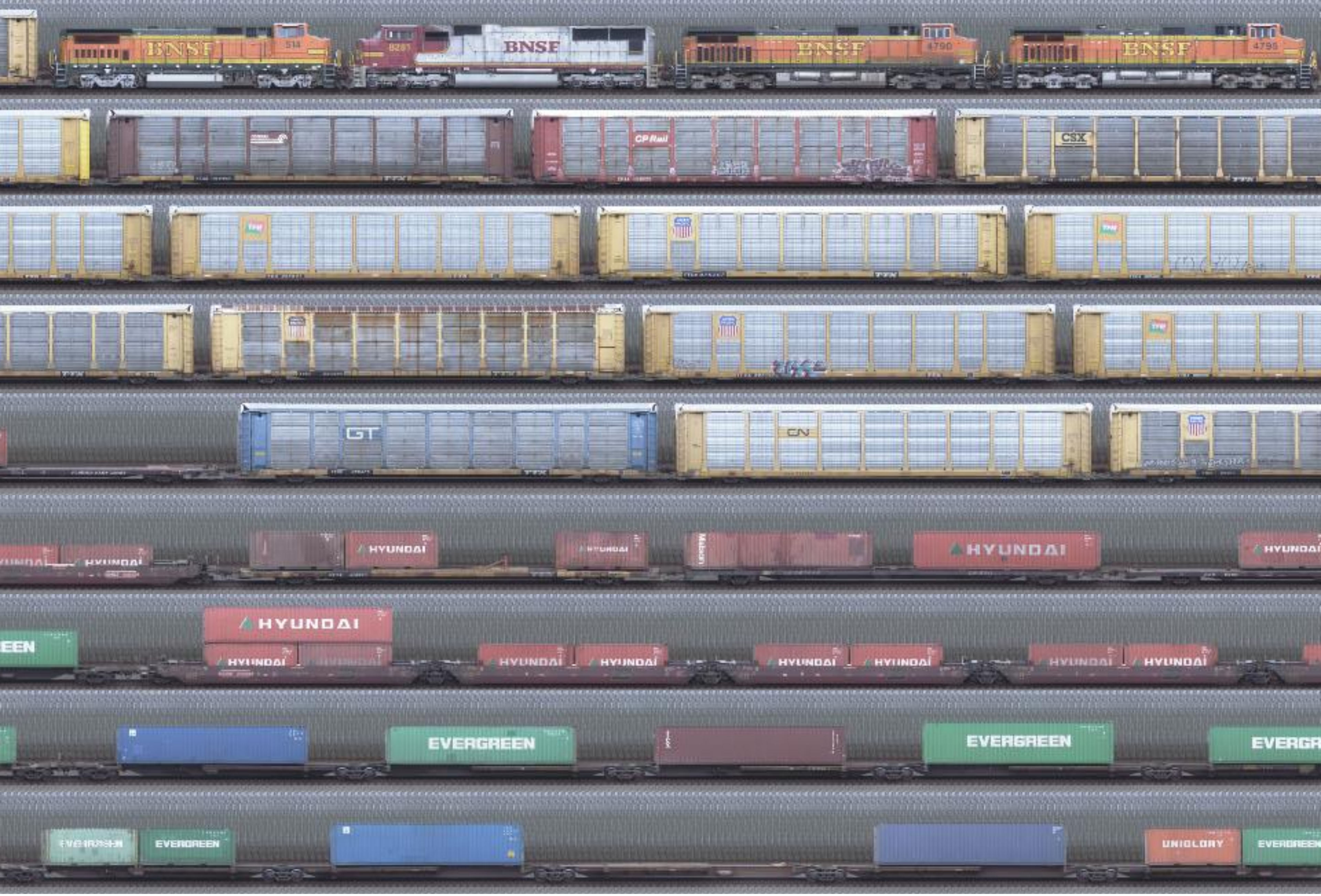
- One row of s,t plane
  - i.e., hold t constant



(3D: Lumigraph/Lightfield)

- One row of s,t plane
  - i.e., hold t constant
  - thus s,u,v
  - a "row of images"

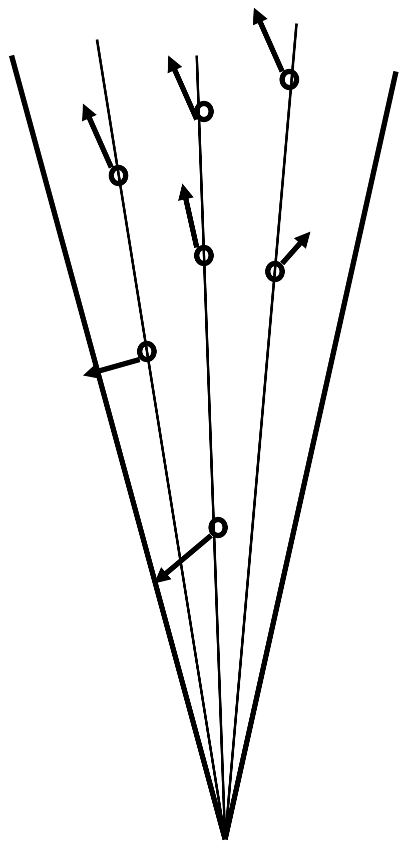




P(x,t) by David Dewey

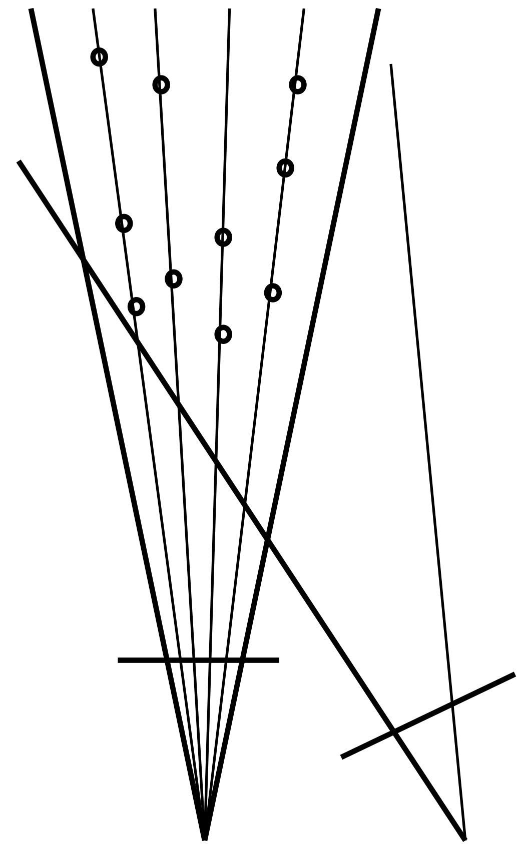
(3-5D: Layered Depth Images)

- Idea:
  - Handle disocclusion
  - Store invisible geometry in depth images
- Data structure:
  - Per pixel list of depth samples
  - Per depth sample:
    - RGBA
    - Z
    - Encoded: Normal direction, distance
  - Pack into cache lines



(3-5D: Layered Depth Images)

- Computation:
  - Incremental warping computation
  - Implicit ordering information
    - Process in up to four quadrant
  - Splat size computation
    - Table lookup
    - Fixed splat templates
  - Clipping of LDIs

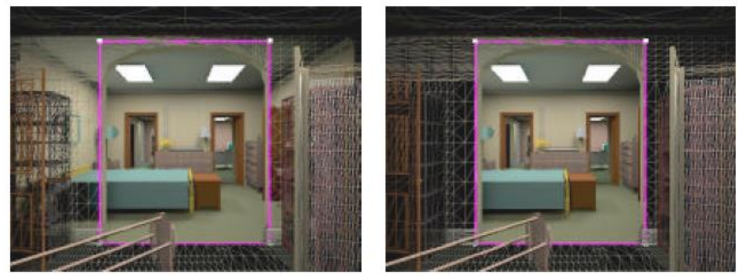


(3-5D: Layered Depth Images)

(3-5D: Layered Depth Images)

Plate 2. The first two layers of the LDI. Pixels without values are shown in red.

Plate 3. Savings due to recursive clipping of the LDI before warping.



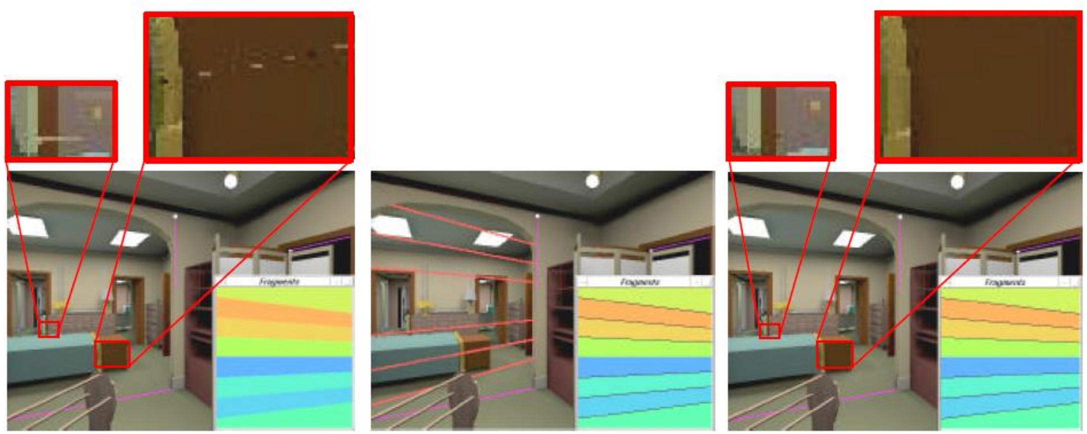


Plate 4. Images are rendered in parallel, one fragment per processor. a) Artifacts (highlighted in red) appear at the borders between fragments because of incorrect occlusion compatible traversal. b) Buffer zones between fragments are rendered during a second pass. c) Image rendered using two-pass traversal does not exhibit the artifacts of image (a).

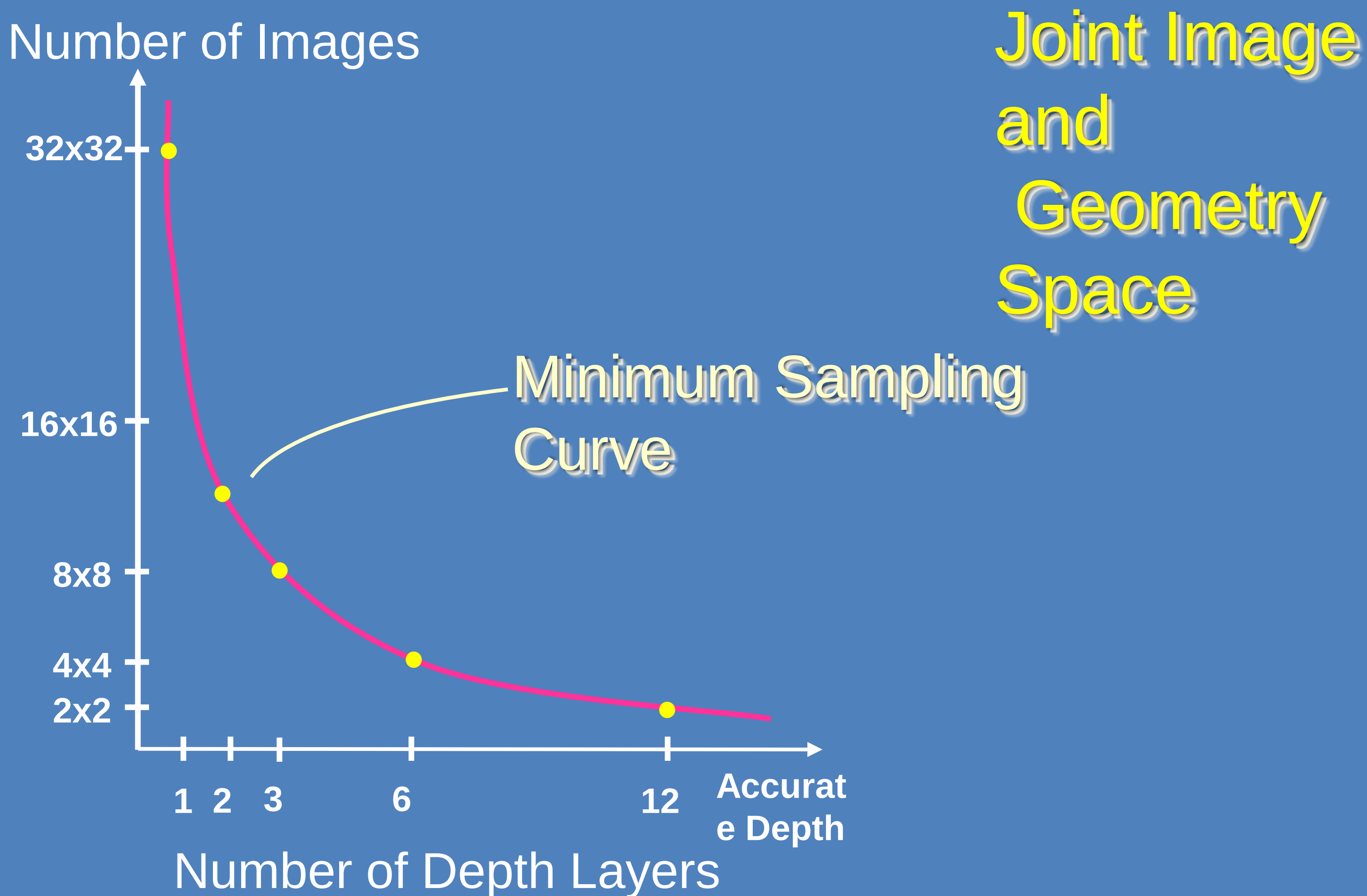
(3-5D: LDI Tree)

- "LDI Tree: A sampling-rate Preserving Hierarchical Representation for Image-based Rendering" by Chang et al.

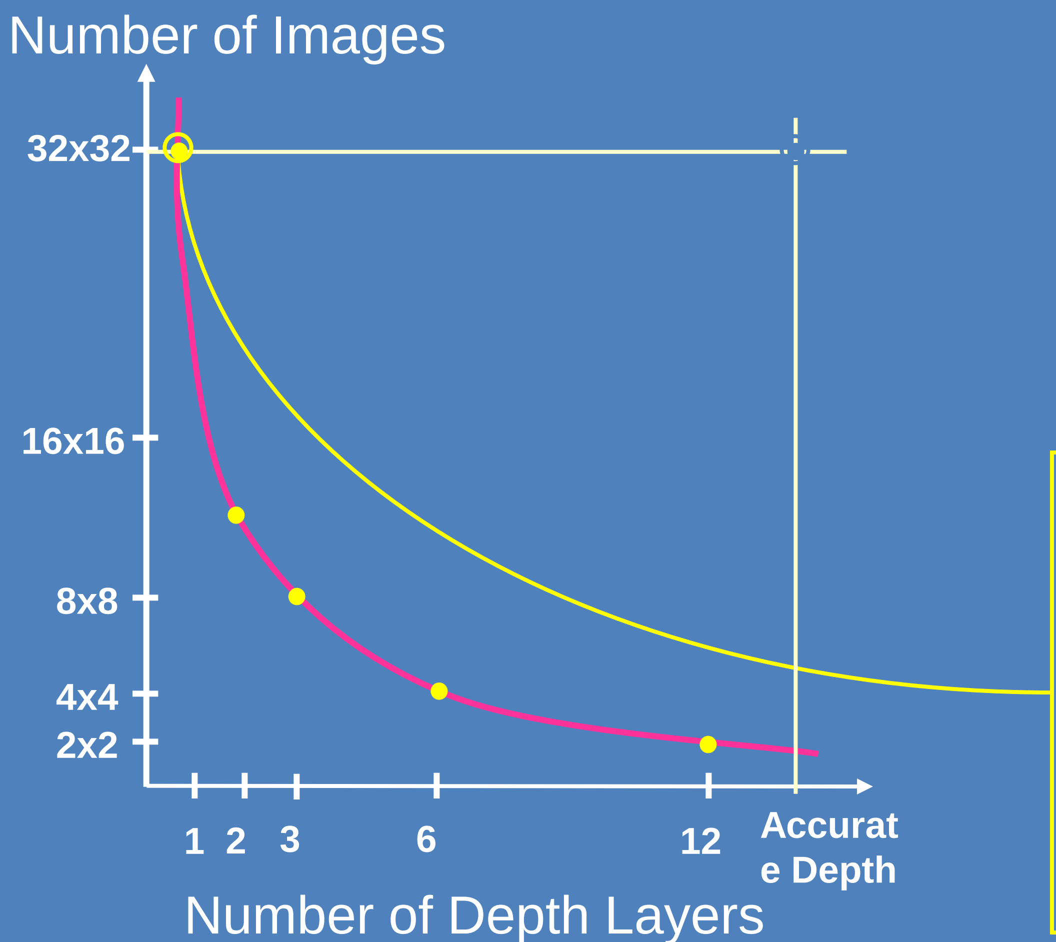
(~5D: LDC Cube)

- 3 orthogonal LDIs (or LDI Trees)
  - Part of "Surfels" paper

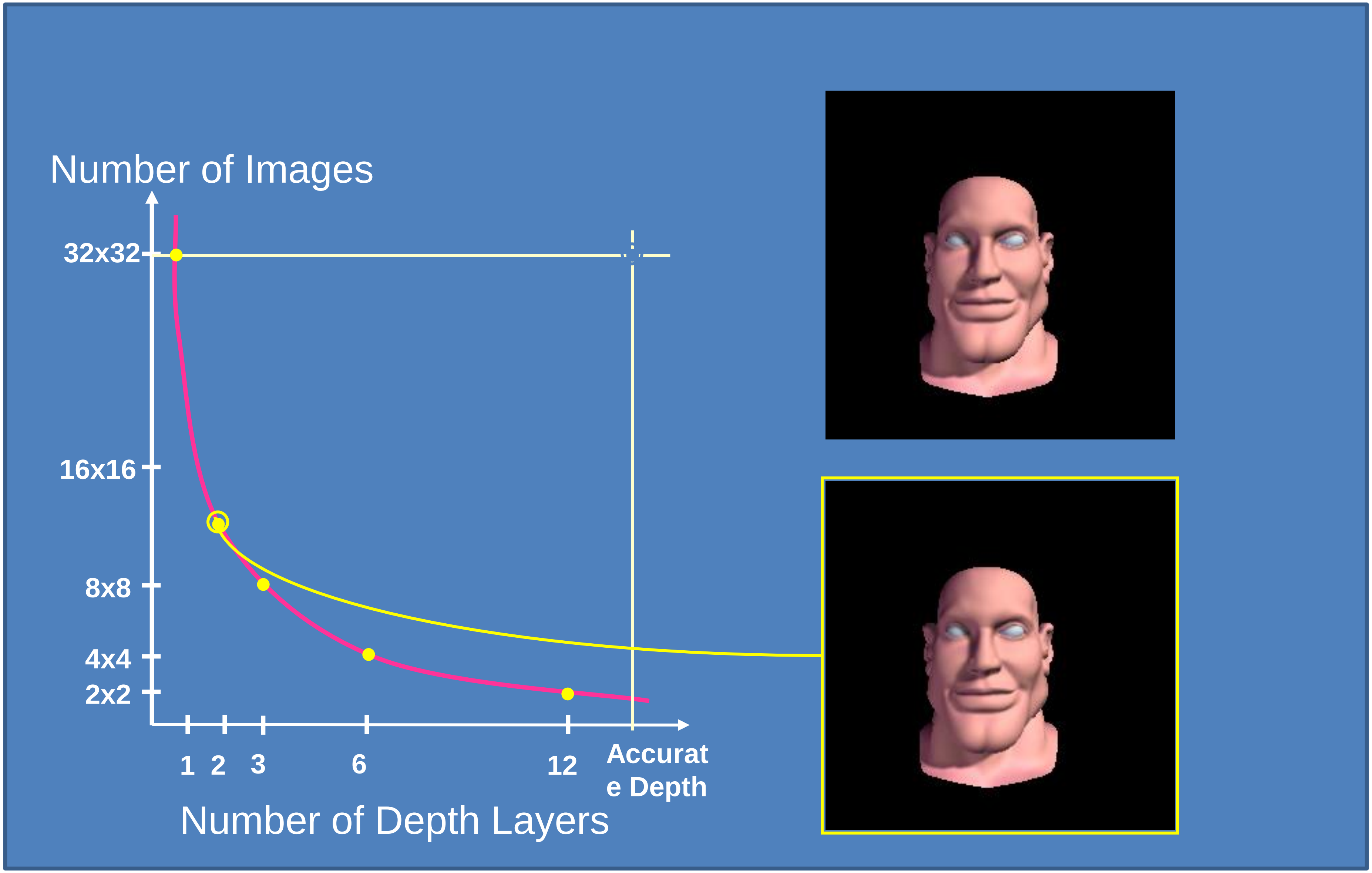
Minimum Sampling Curve



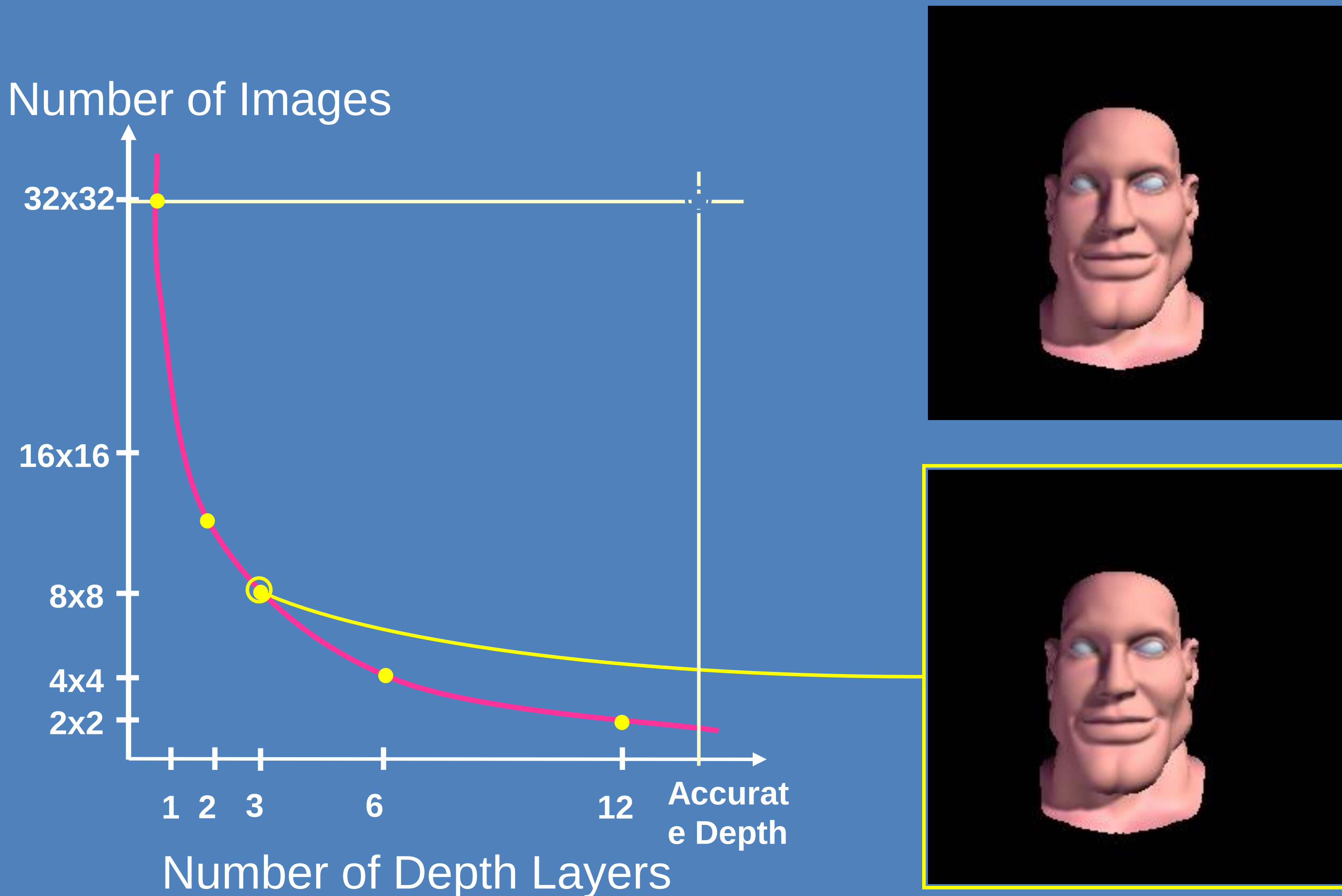
Minimum Sampling Curve



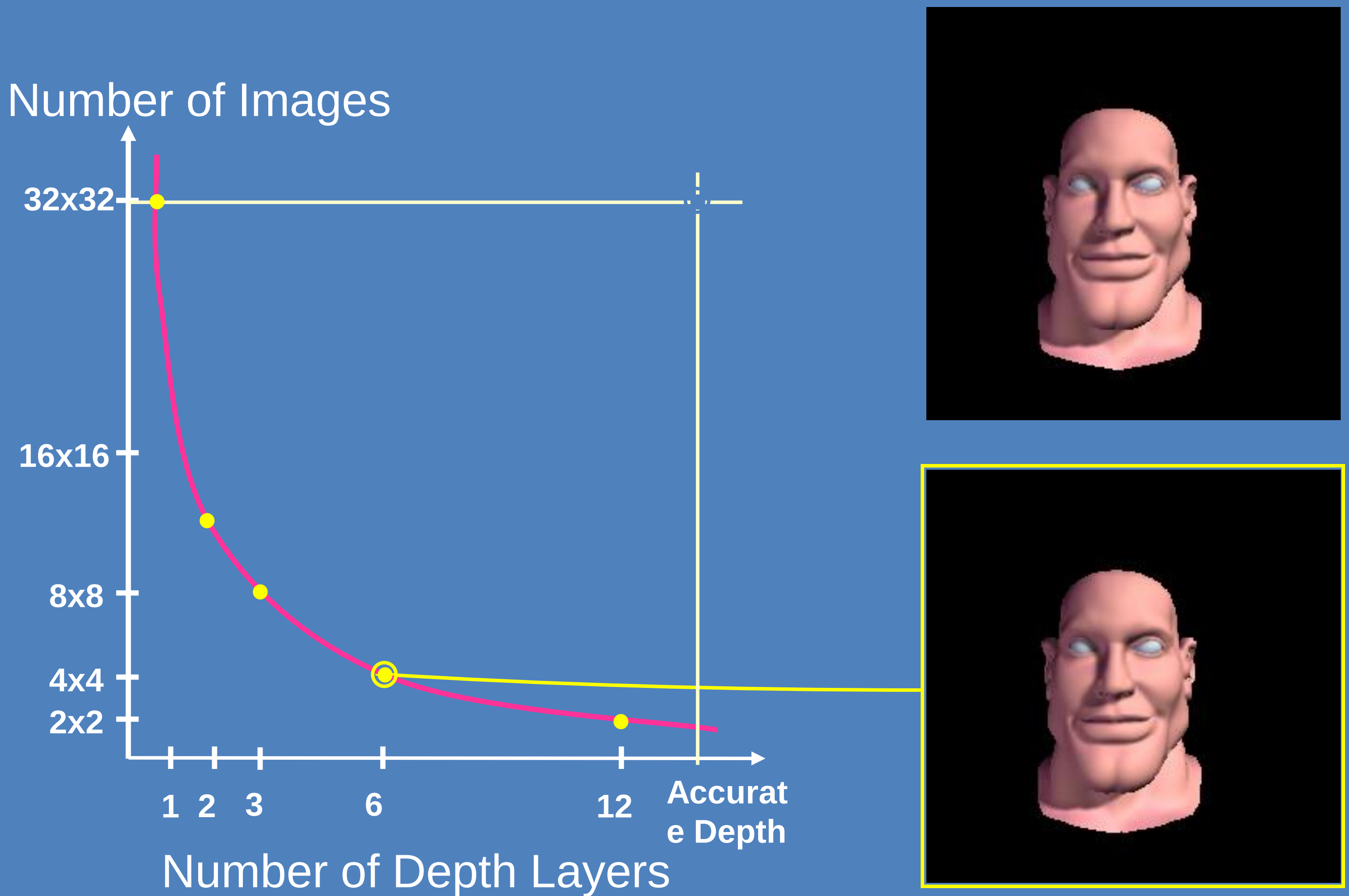
Minimum Sampling Curve



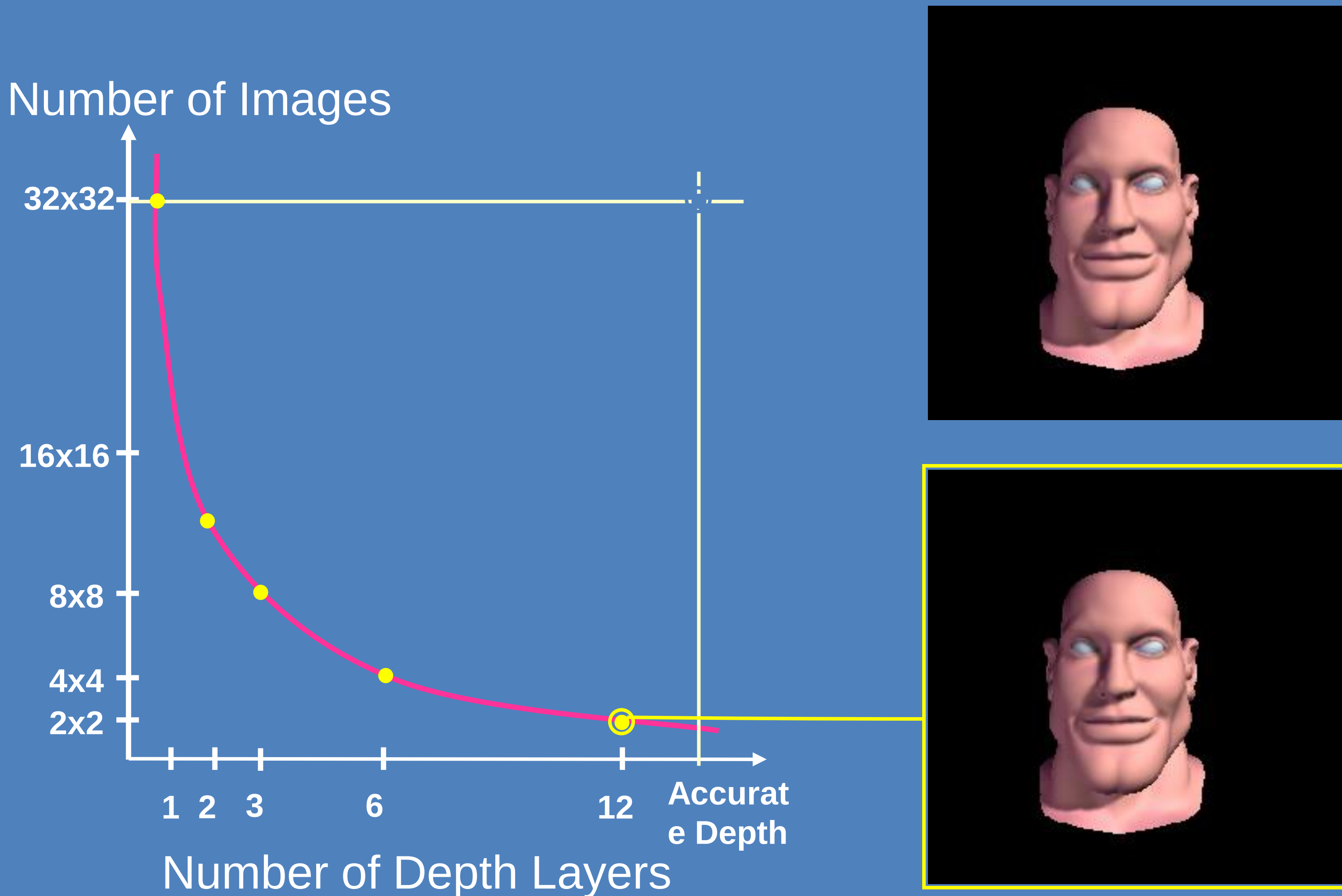
Minimum Sampling Curve



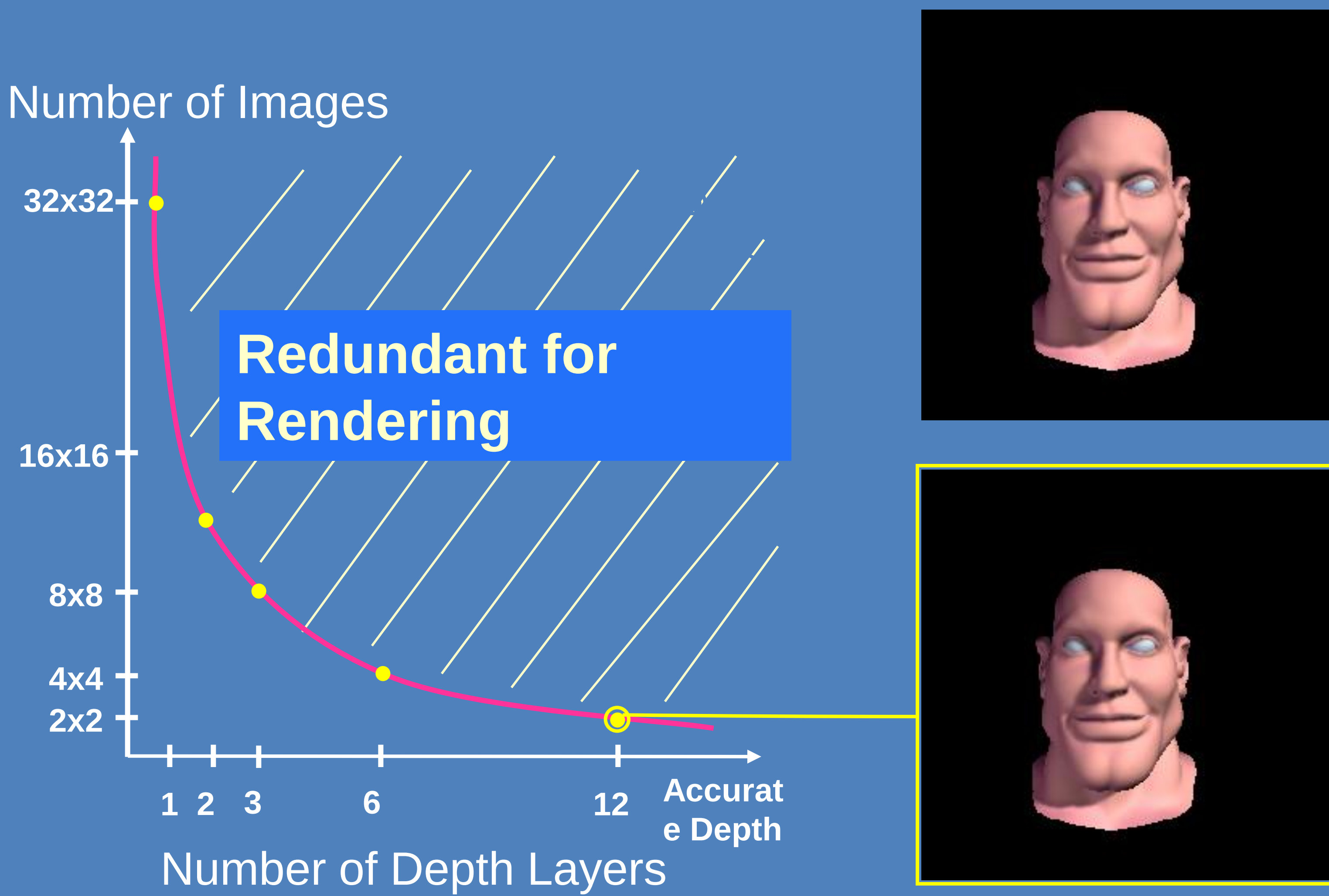
Minimum Sampling Curve



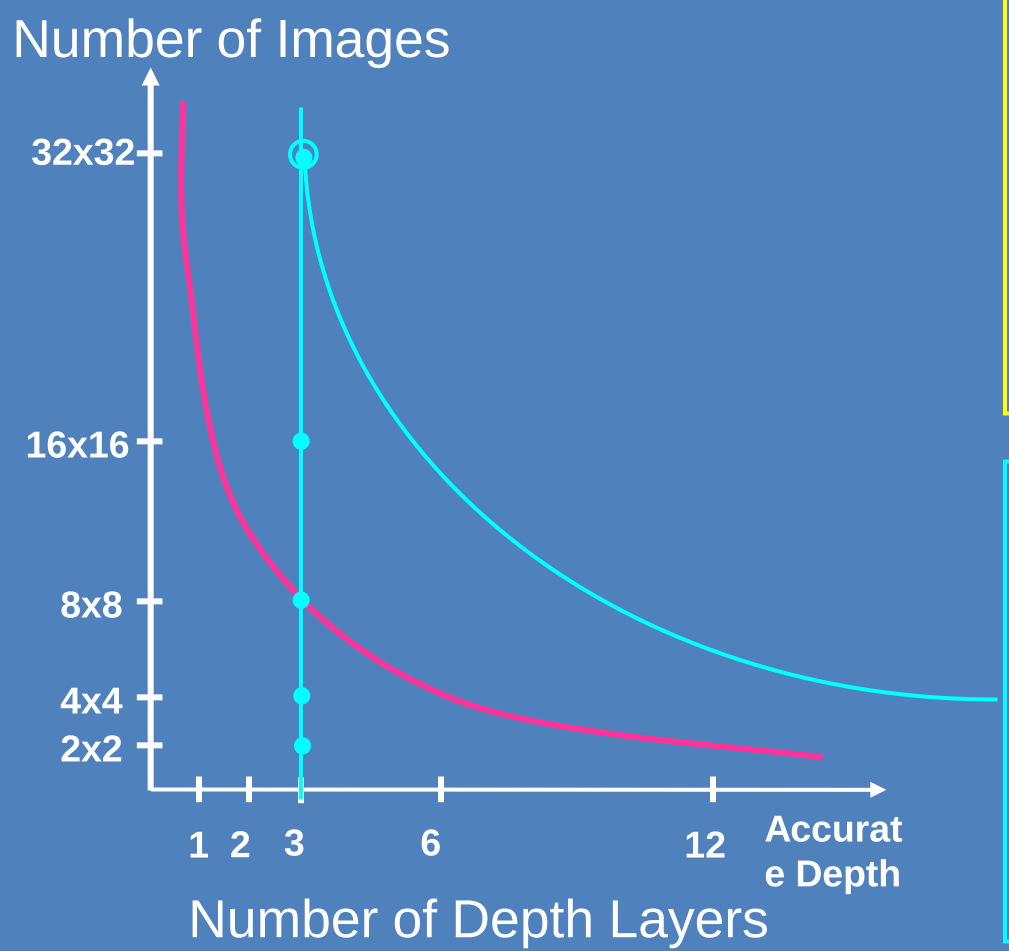
Minimum Sampling Curve

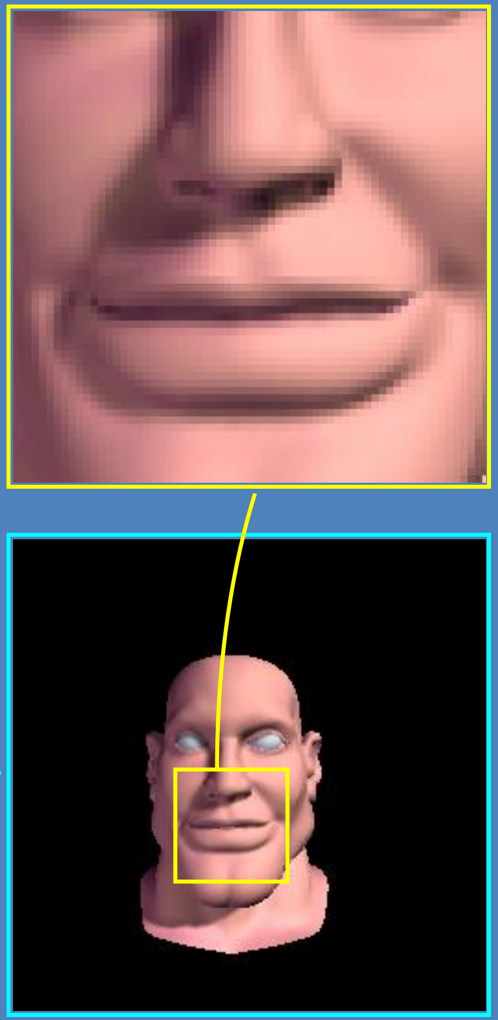


Minimum Sampling Curve

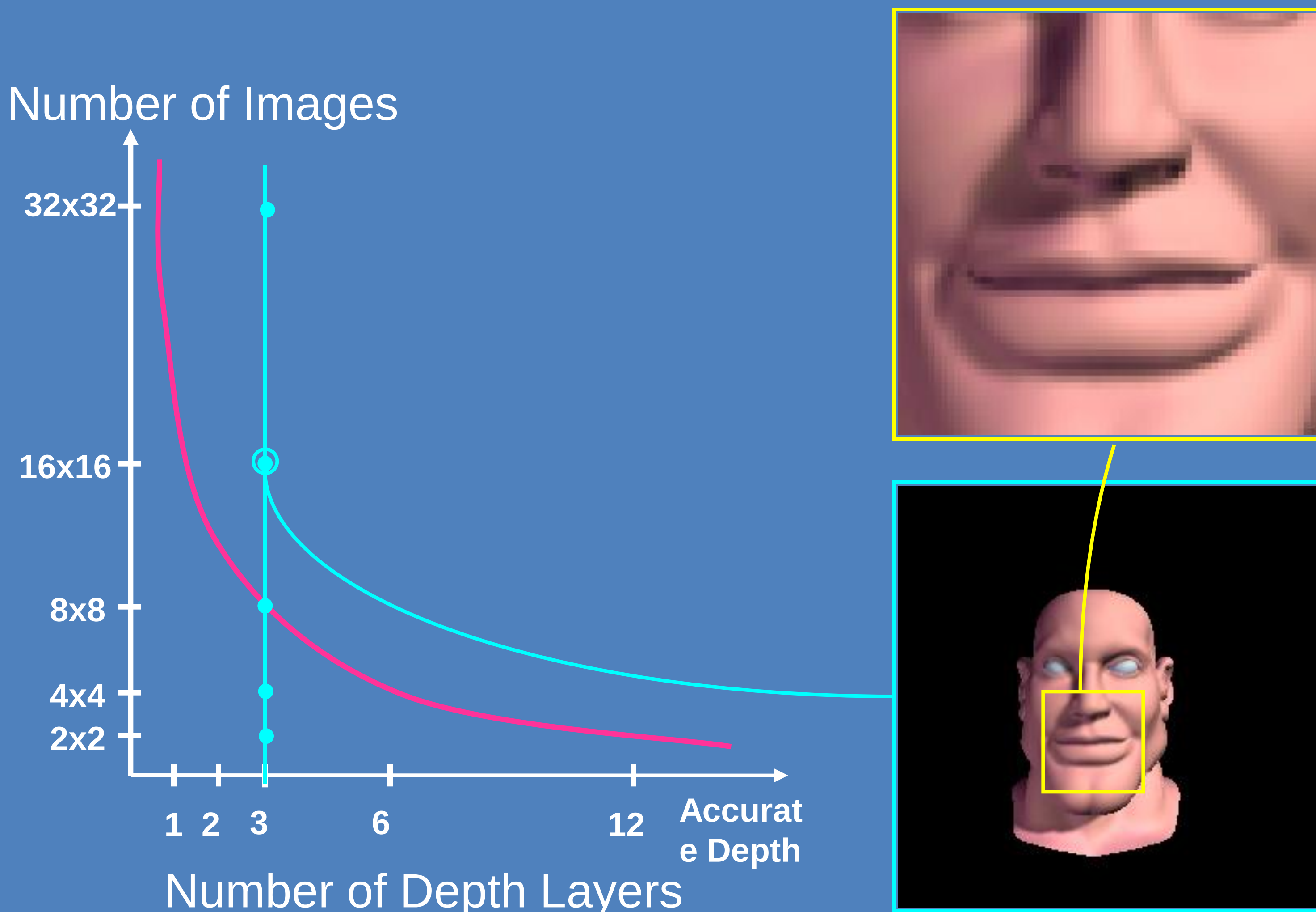


More Geometry: 3 Layers

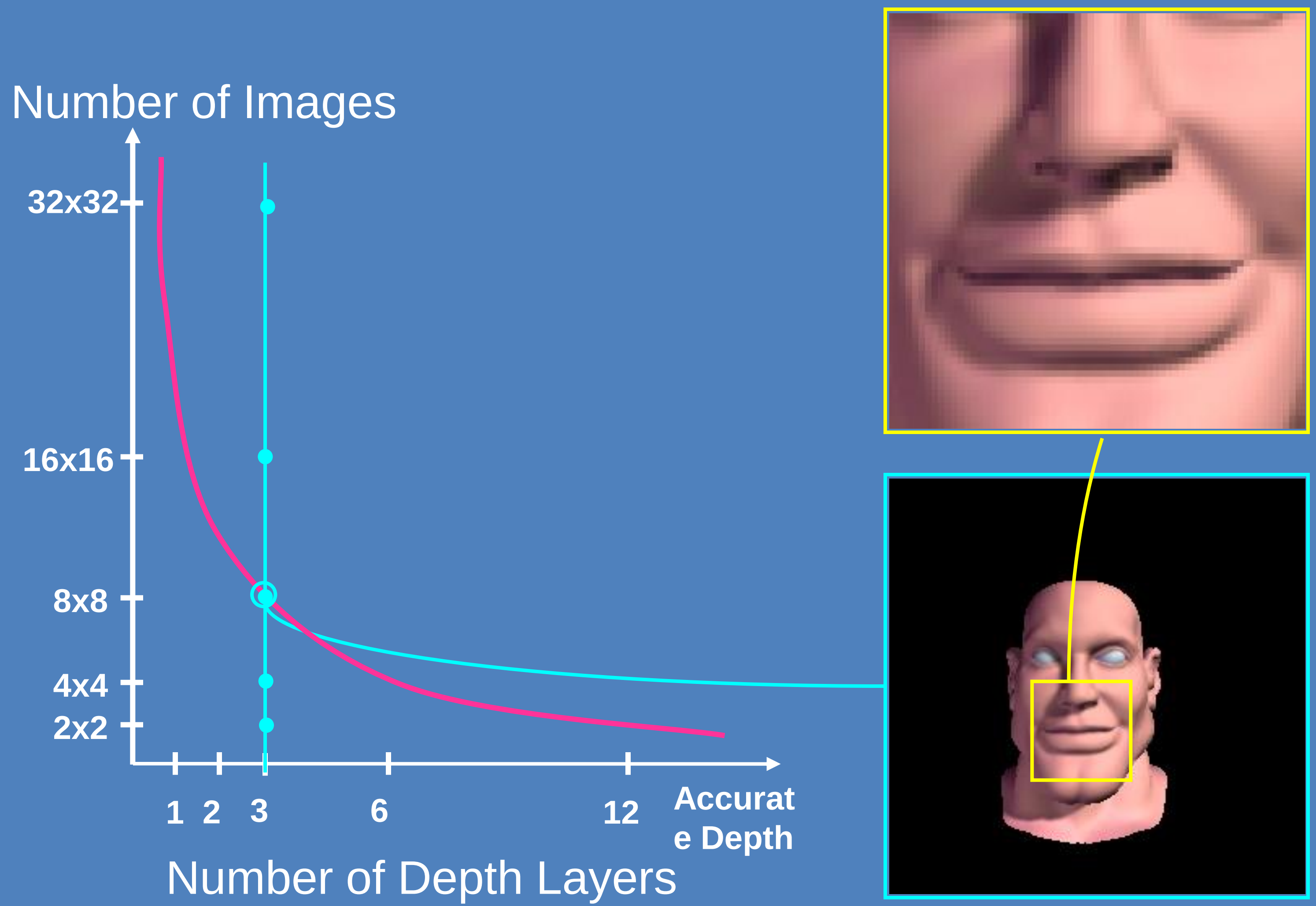




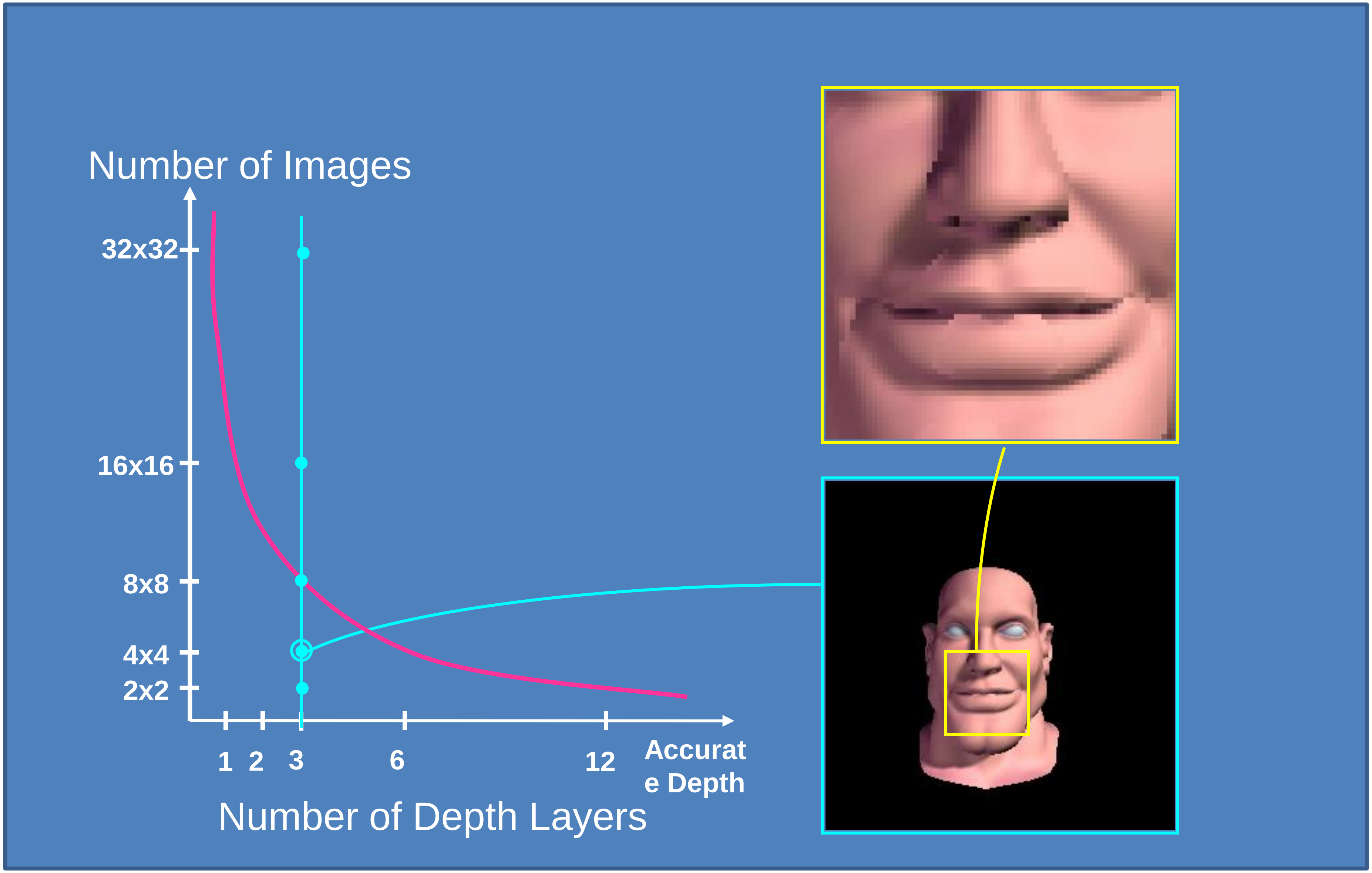
3 Layers


Number of Images
32x32
16x16
8x8
4x4
2x2
1
2
3
6
12
Accurat
e Depth
Number of Depth Layers

3 Layers

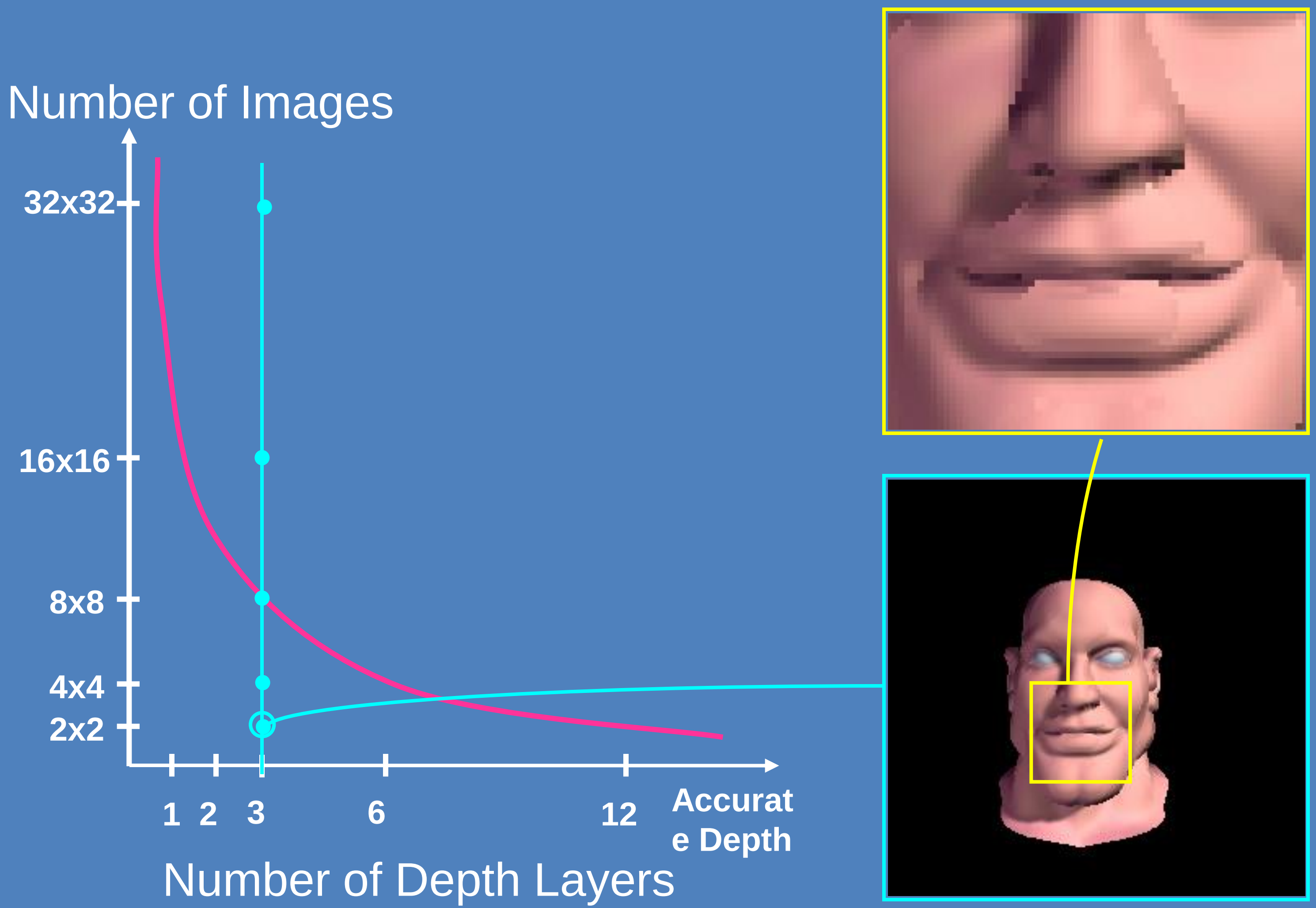

Number of Images
32x32
16x16
8x8
4x4
2x2
1
2
3
6
12
Accurat
e Depth
Number of Depth Layers

3 Layers



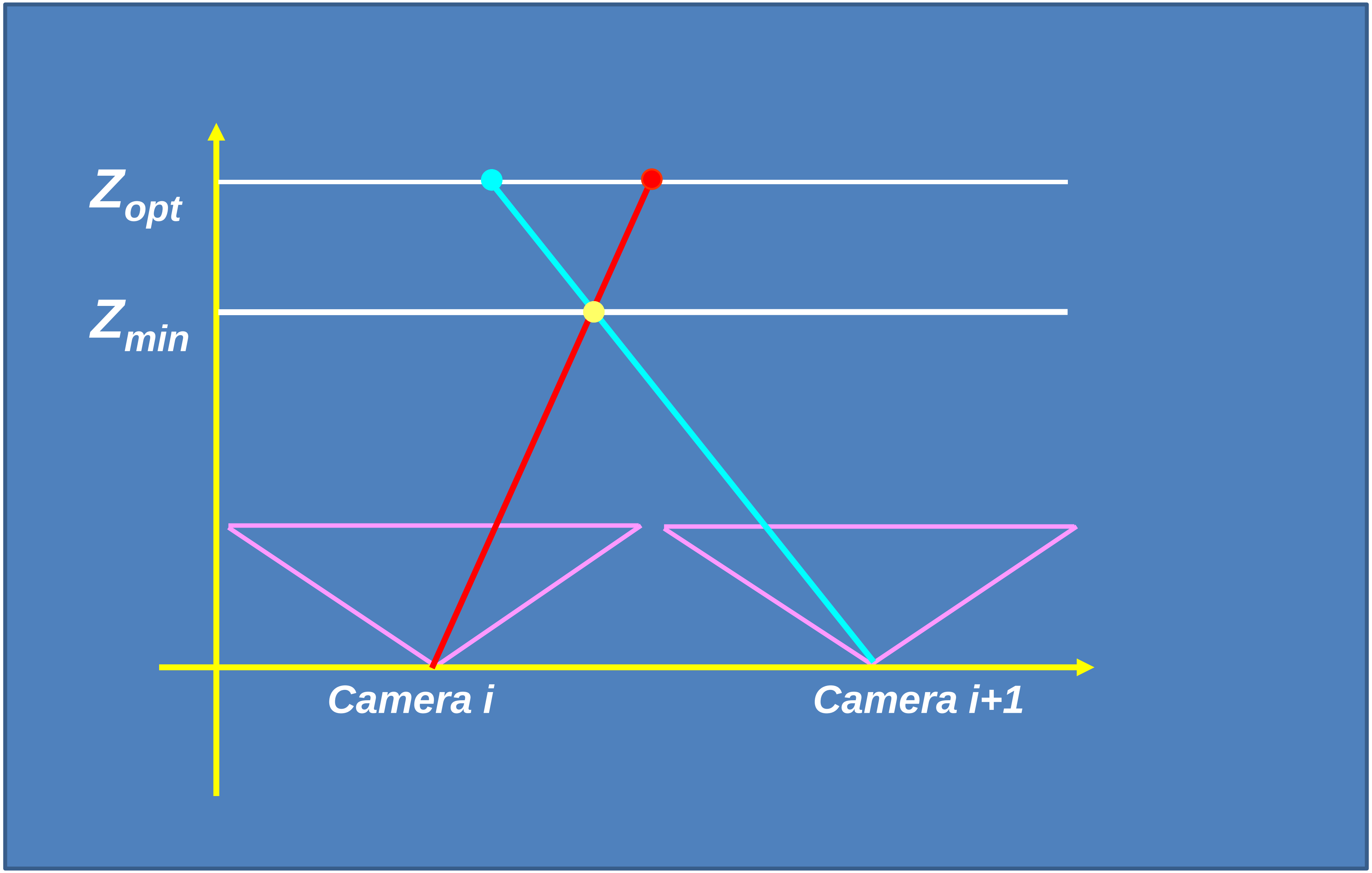
Number of Images
32x32
16x16
8x8
4x4
2x2
1
2
3
6
12
Accurat
e Depth
Number of Depth Layers

3 Layers

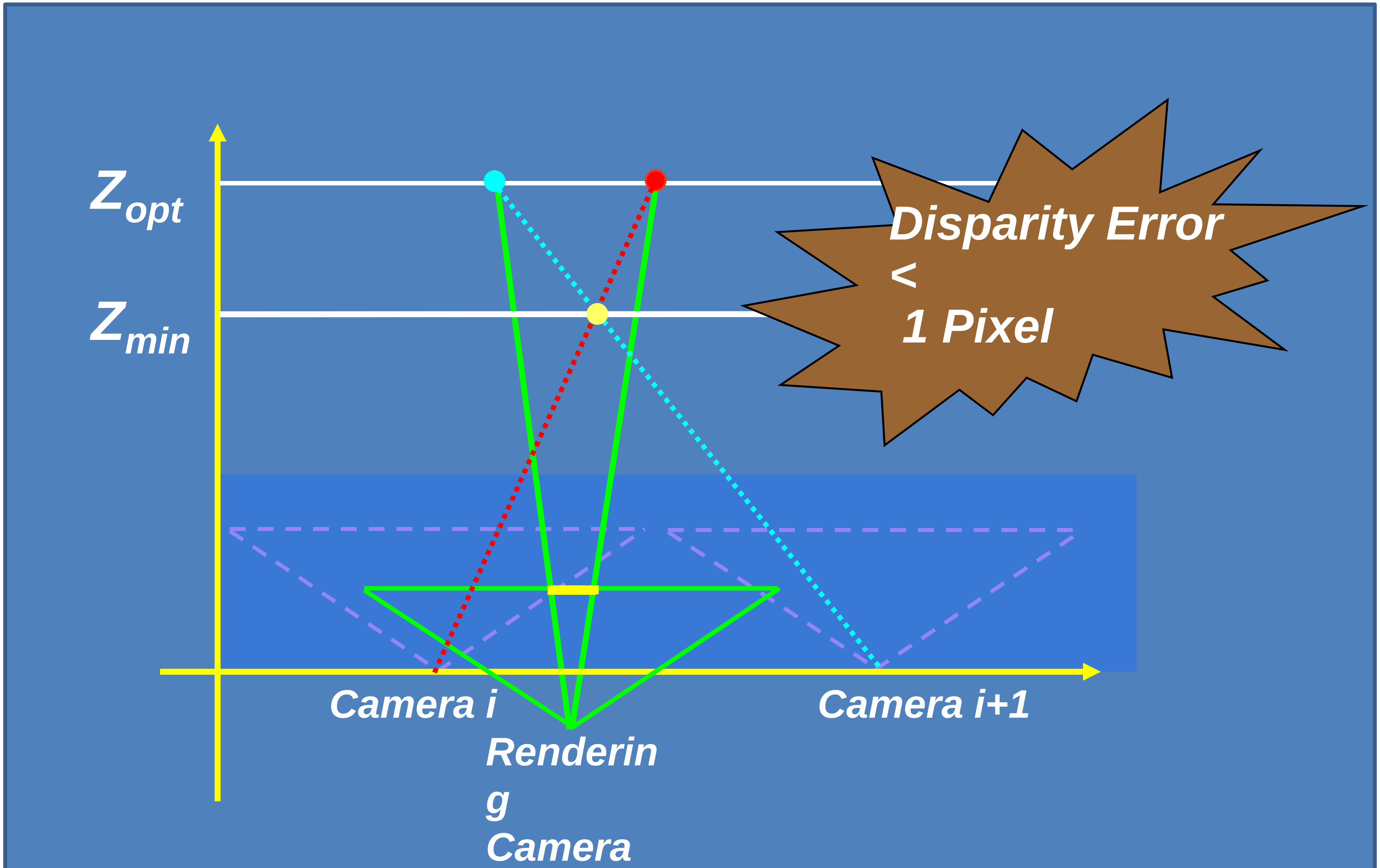


Number of Images
32x32
16x16
8x8
4x4
2x2
1
2
3
6
12
Accurat
e Depth
Number of Depth Layers

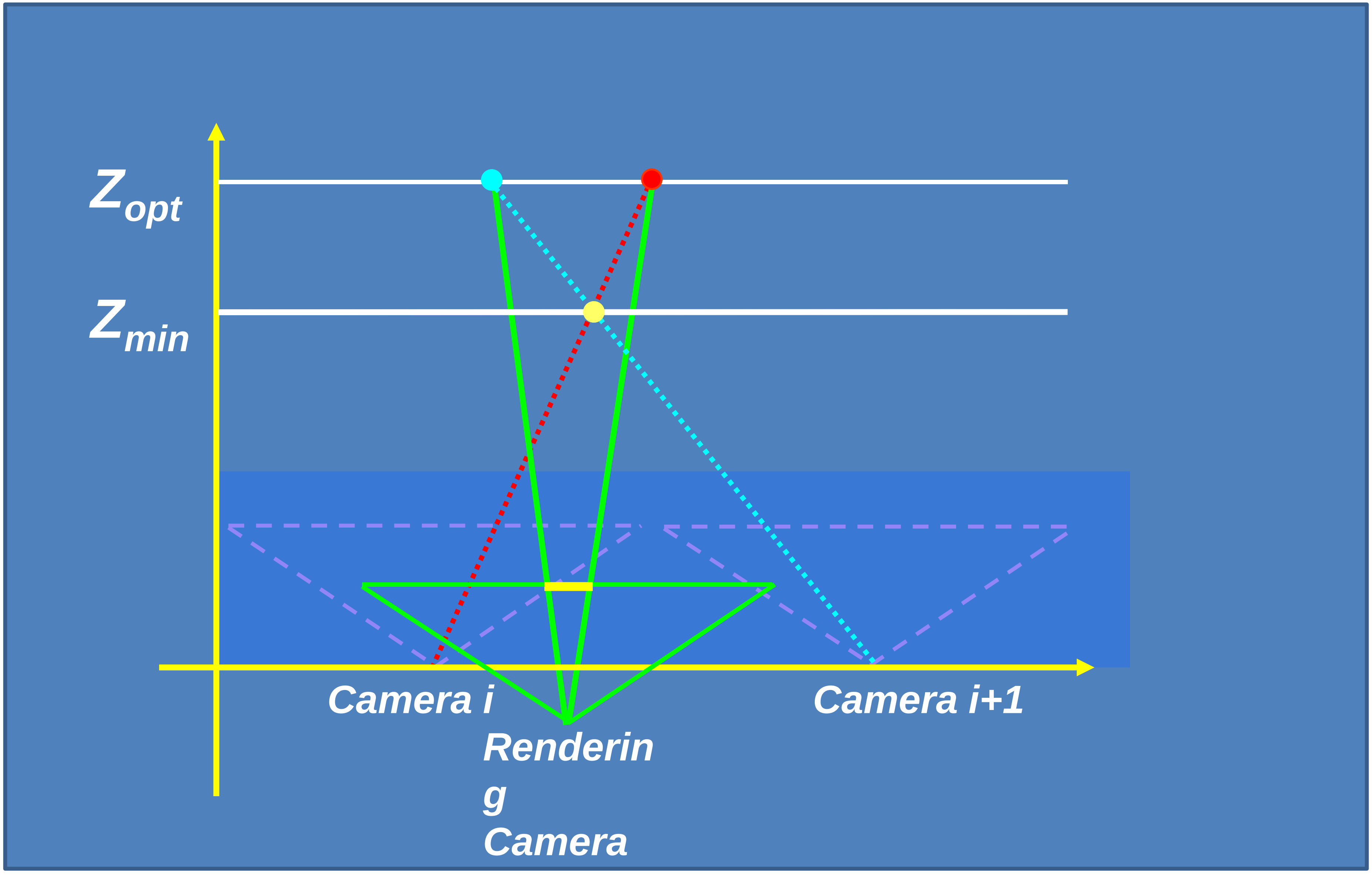
A Geometrical Intuition



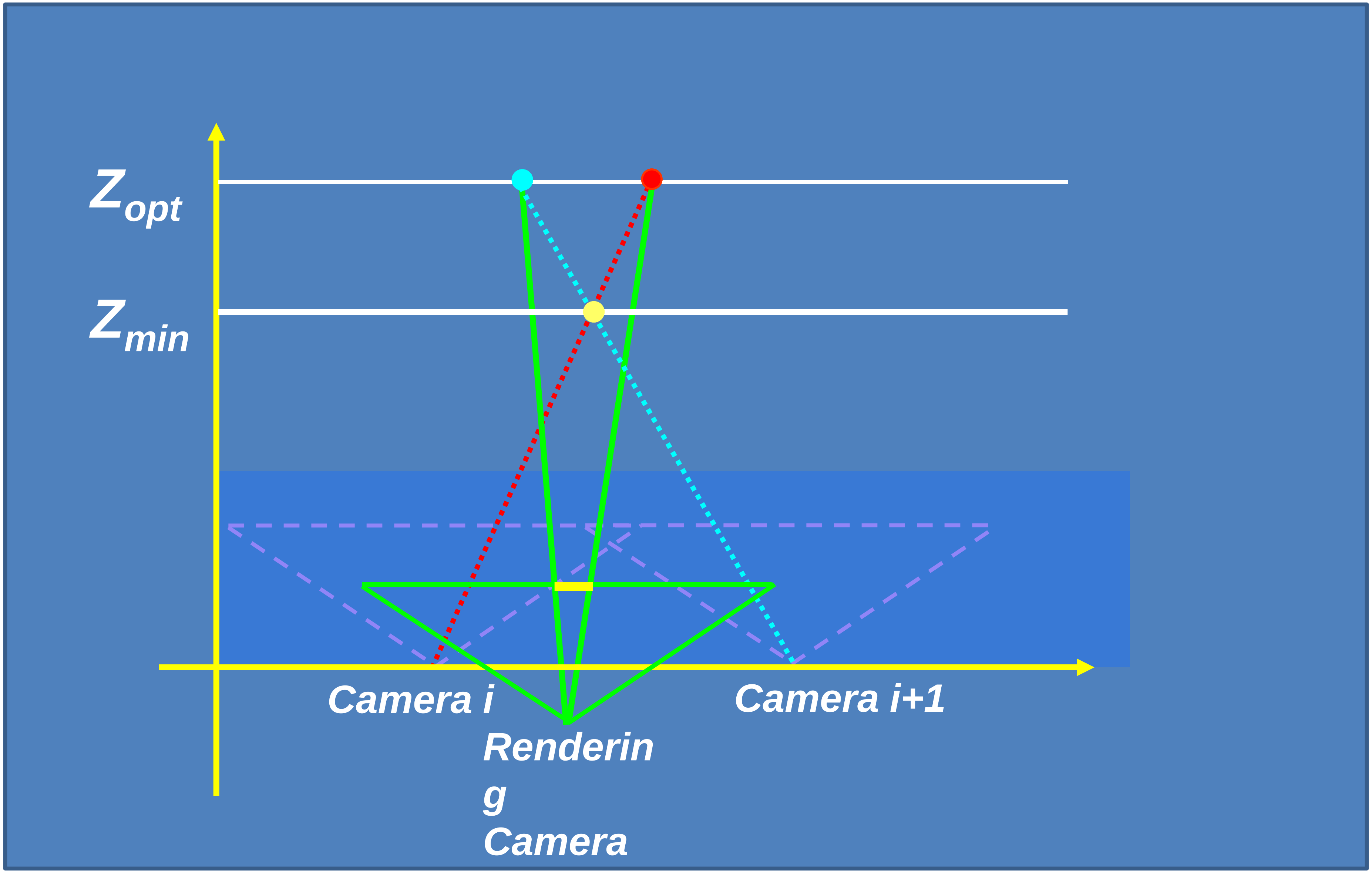
A Geometrical Intuition



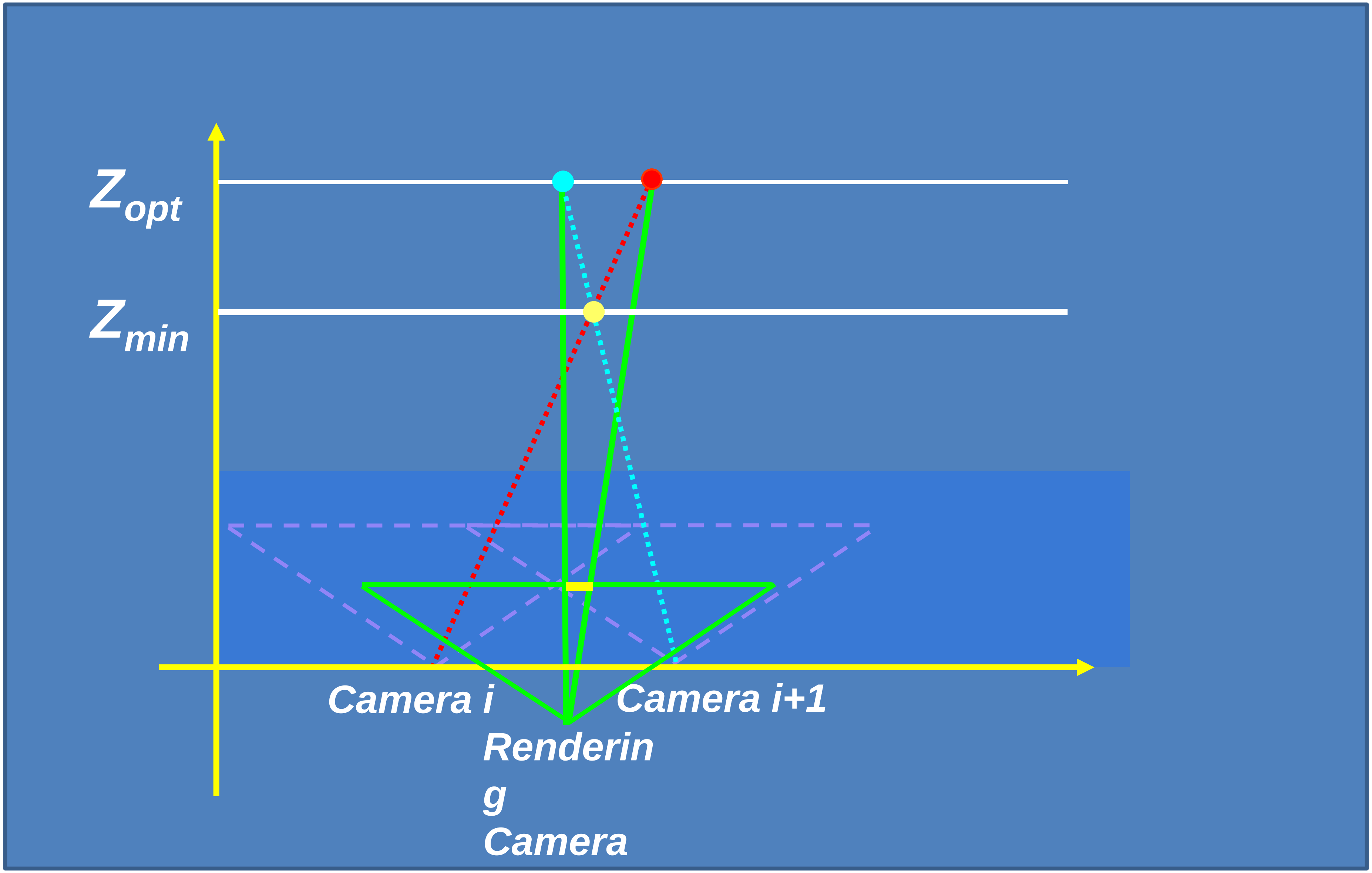
A Geometrical Intuition



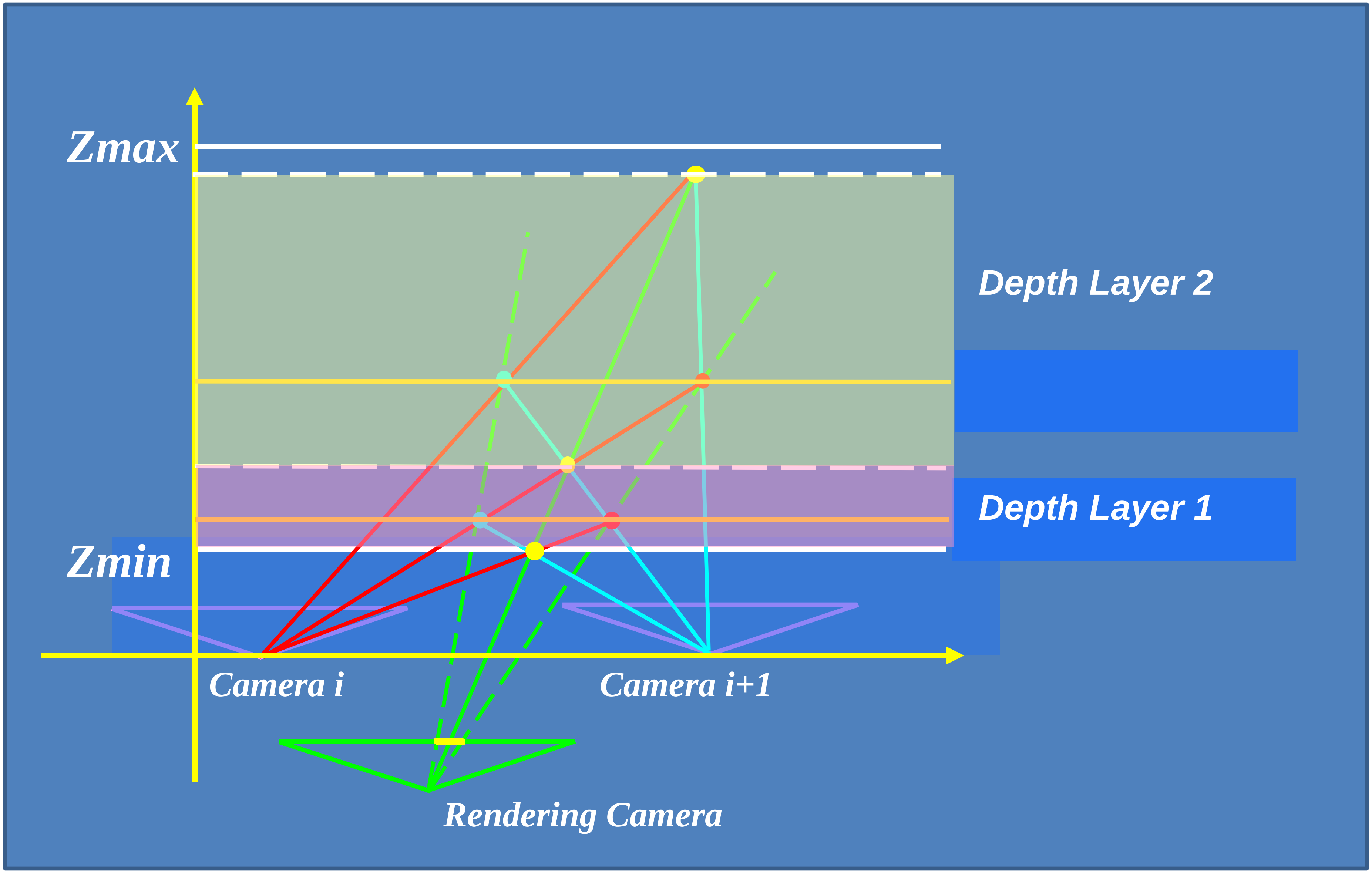
A Geometrical Intuition



A Geometrical Intuition



A Geometrical Intuition



Plenoptic Sampling

48X48 Images
No Depth

16X16 Images
3Bits Depth

Plenoptic Sampling

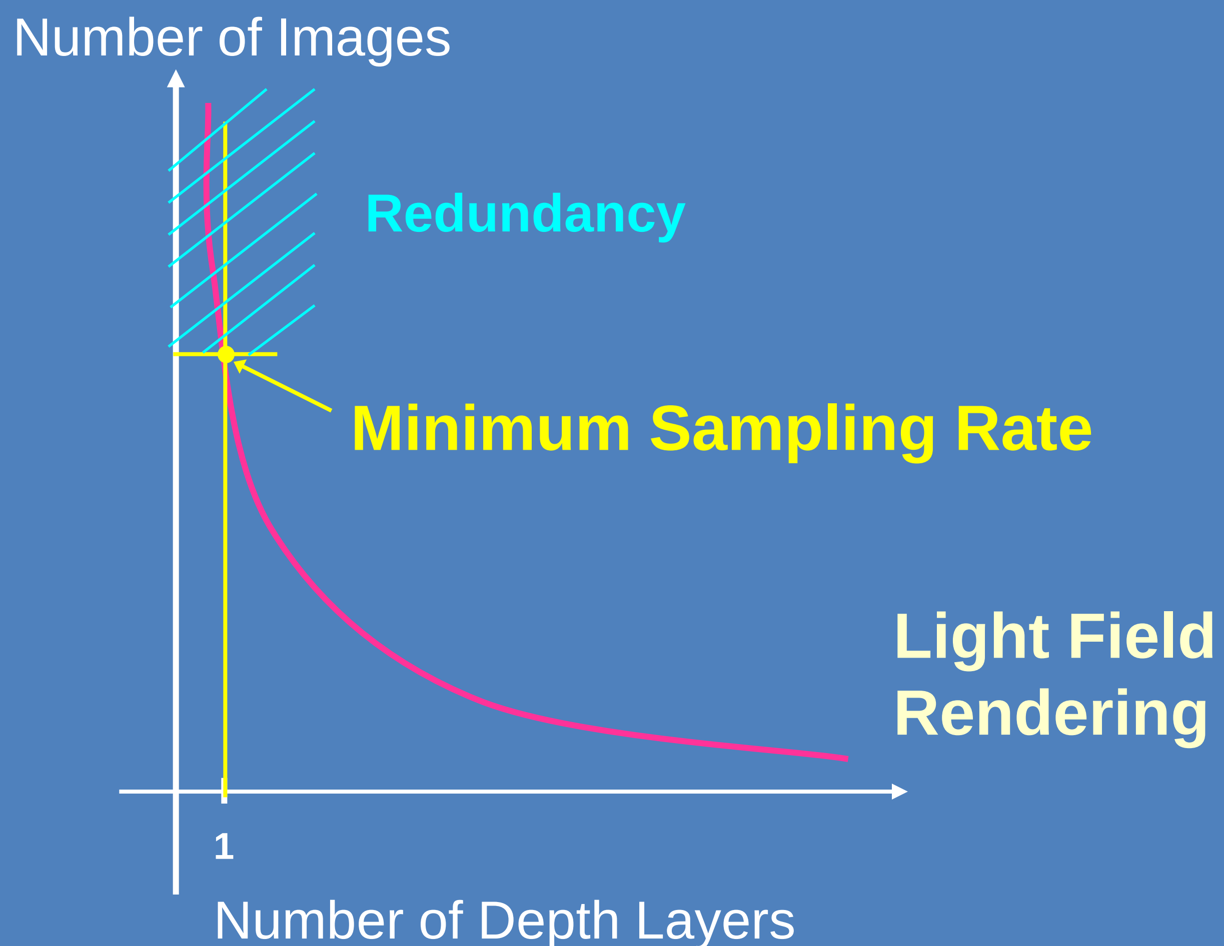
48X48 Images without Depth

Antialiasing Rendering Needs
2930X2930 images = **5,000GB**

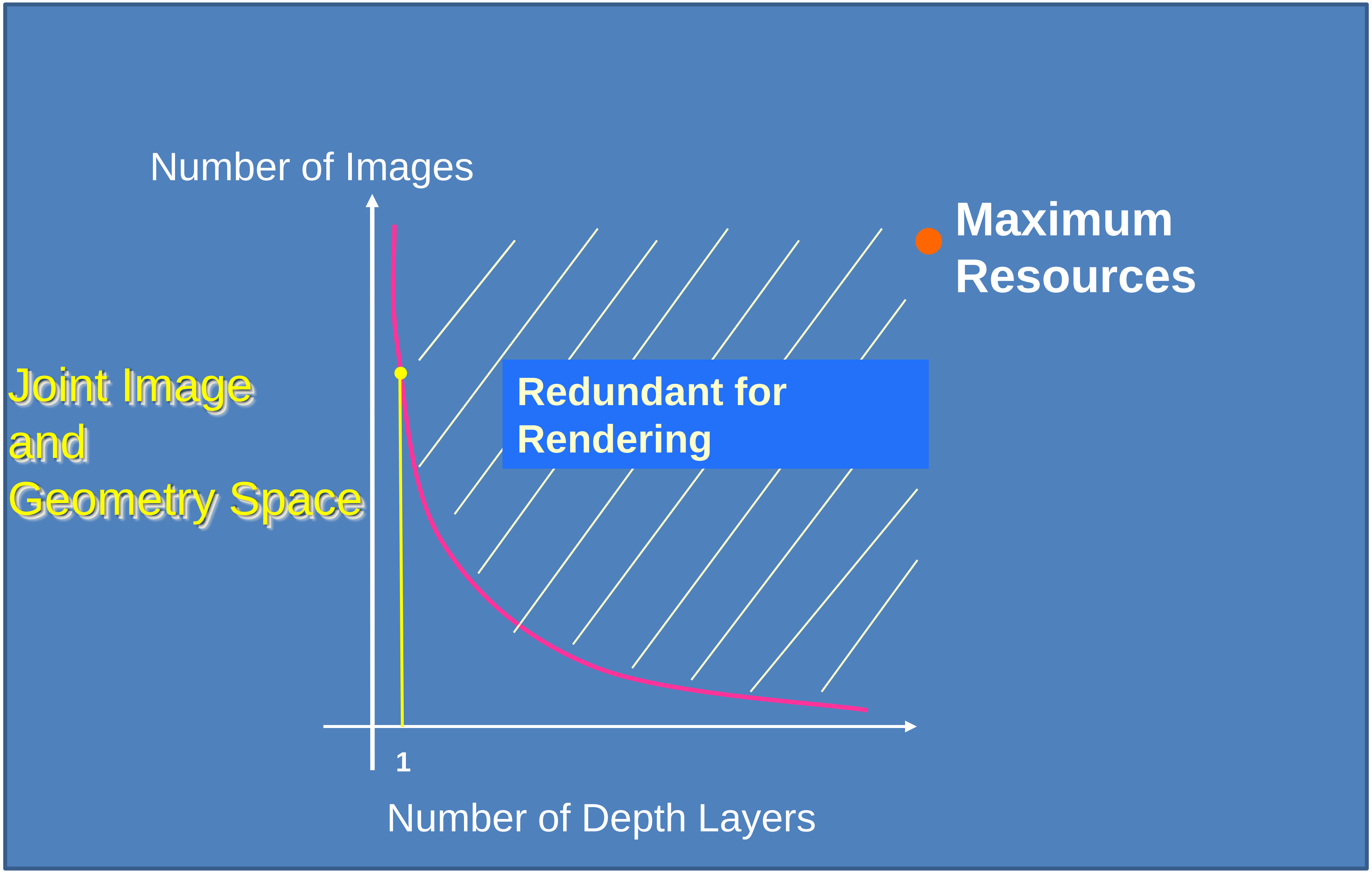
24X24 Images with 7Bits Depth

Antialiasing Rendering Needs
24X24 RGBD images = **0.5GB**

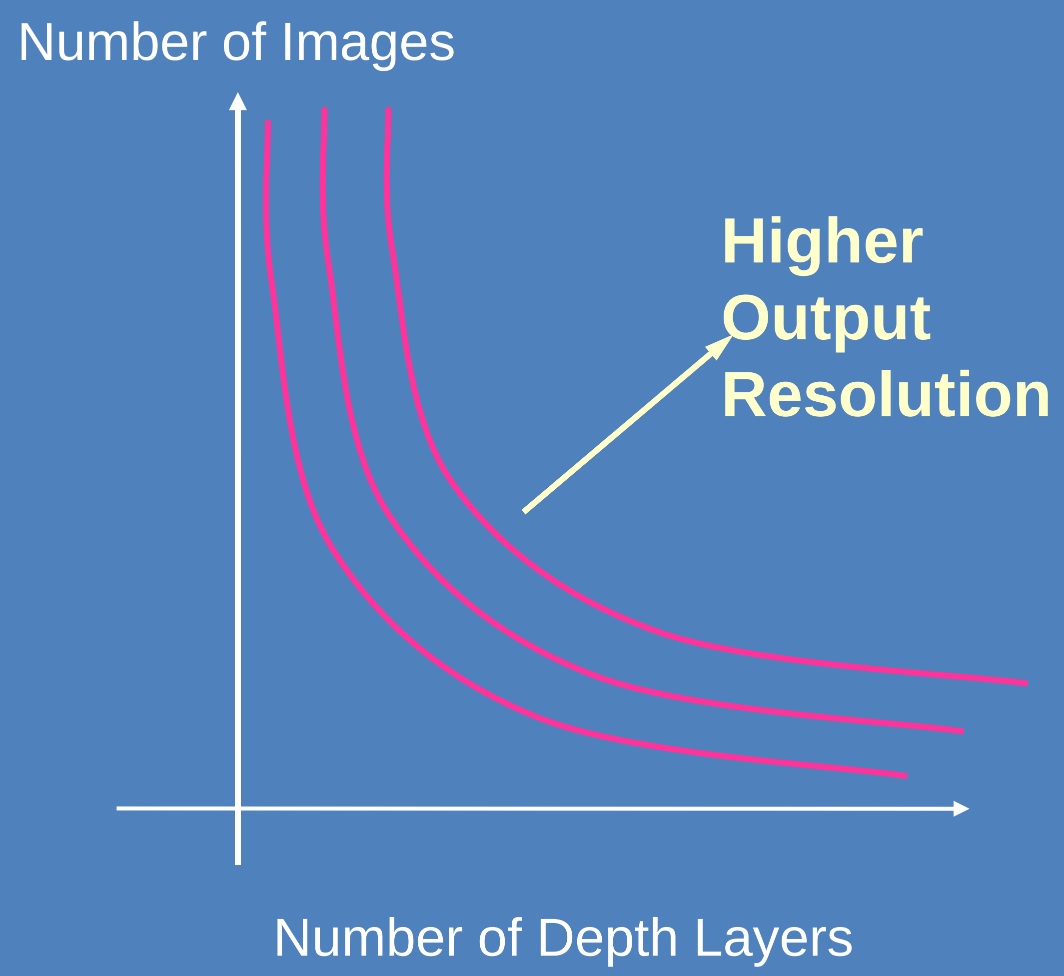
Plenoptic Sampling



Plenoptic Sampling



Plenoptic Sampling

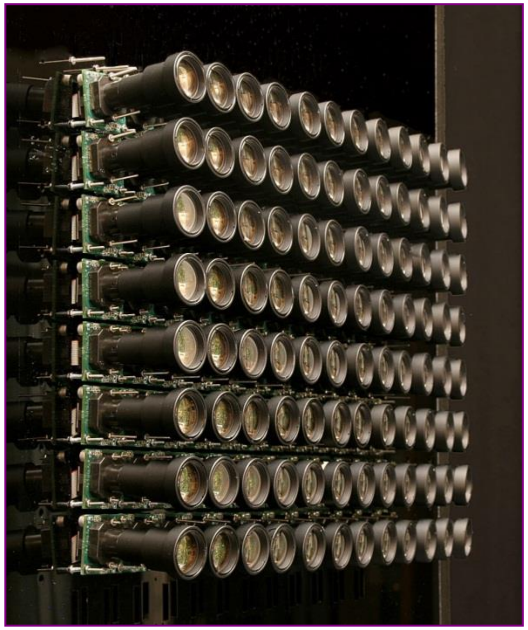


List of projects

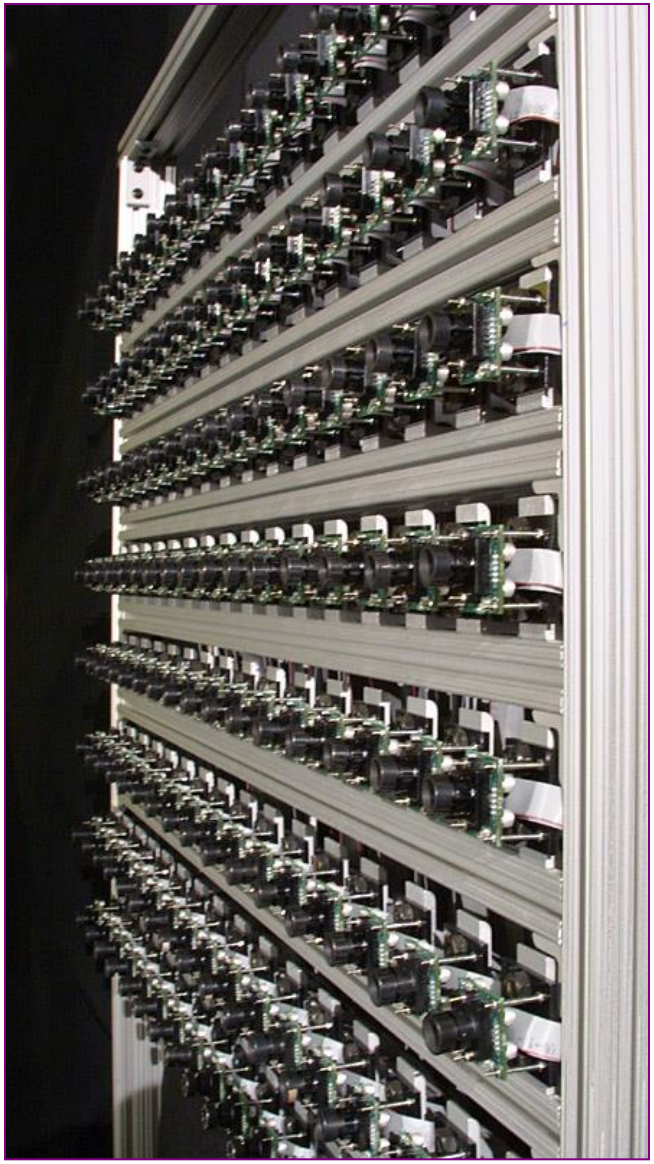
- high performance imaging
  using large camera arrays
- light field photography
  using a handheld plenoptic camera
- dual photography

High performance imaging using large camera arrays

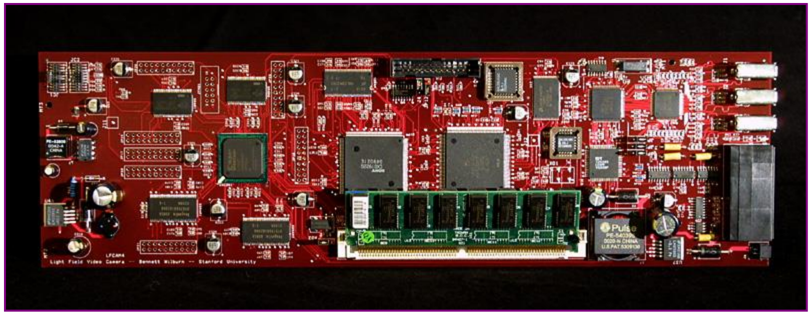
Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez,
Adam Barth, Andrew Adams, Mark Horowitz, Marc Levoy



Stanford multi-camera array

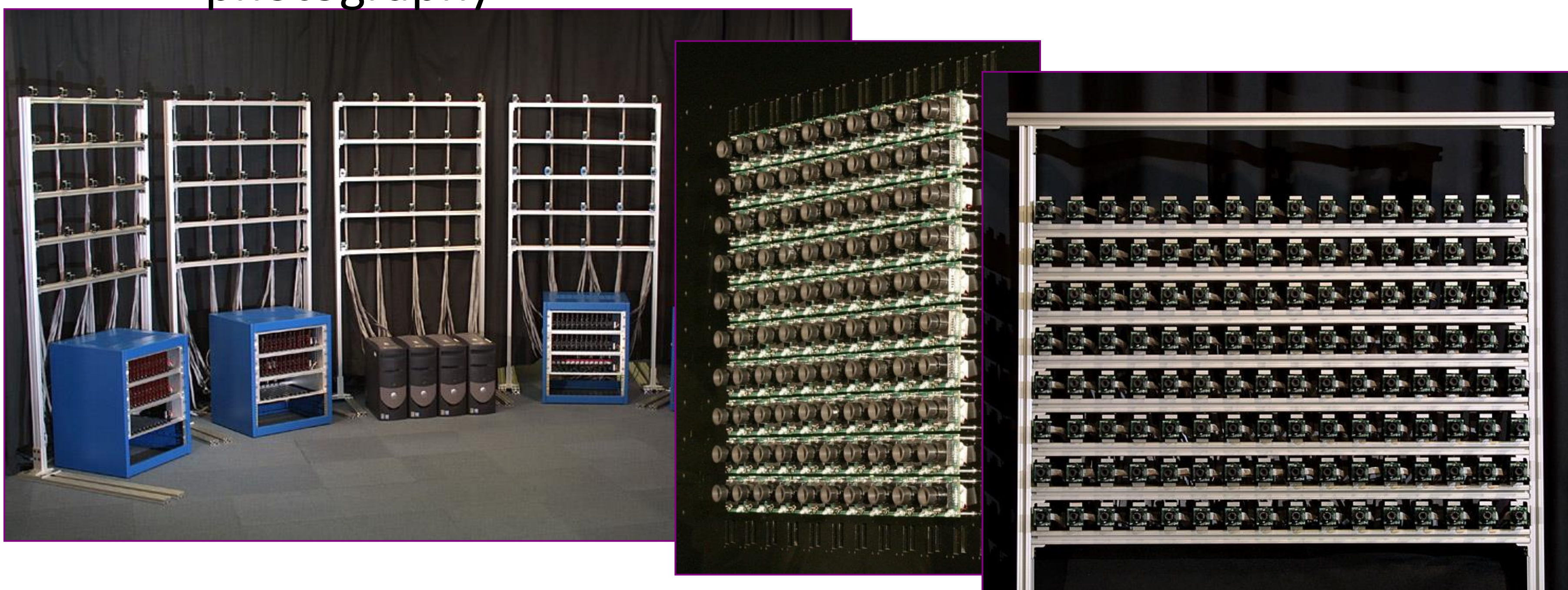


- 640 × 480 pixels × 30 fps × 128 cameras
- synchronized timing
- continuous streaming
- flexible arrangement

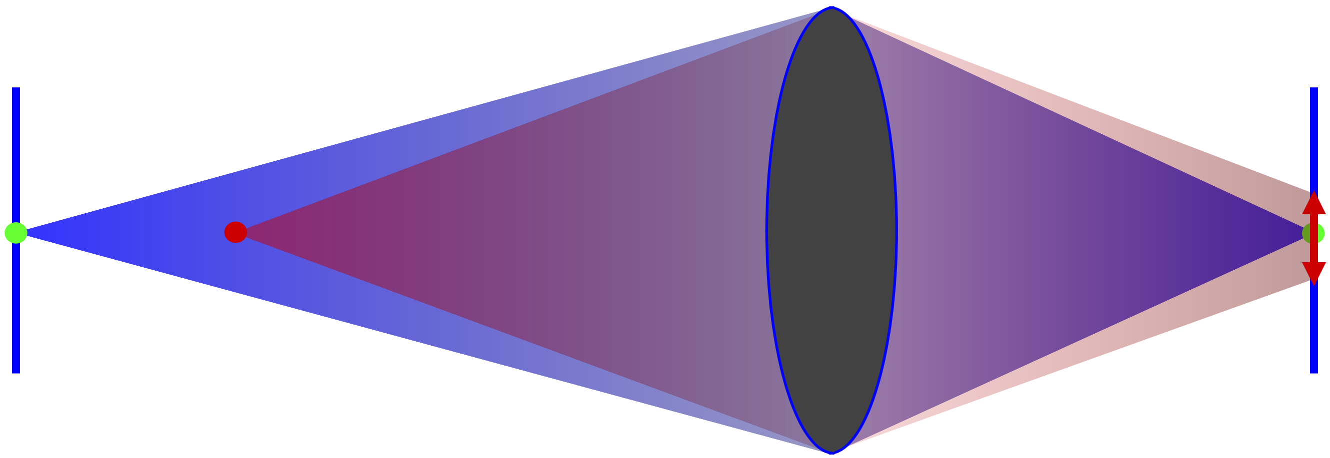


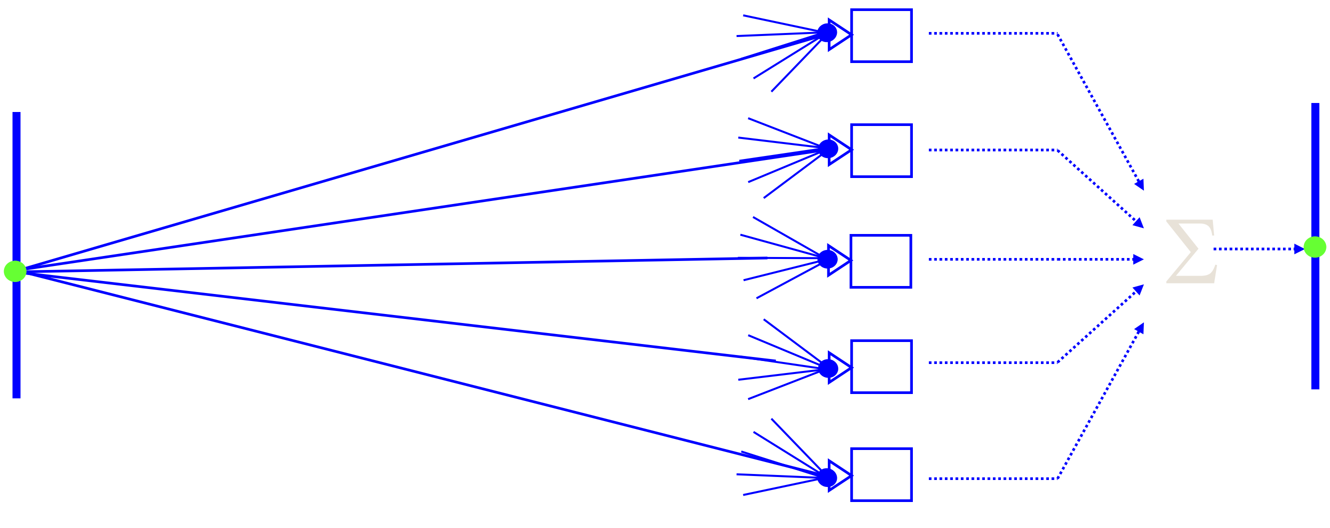
Ways to use large camera arrays

- widely spaced → light field capture
- tightly packed → high-performance imaging
- intermediate spacing → synthetic aperture photography



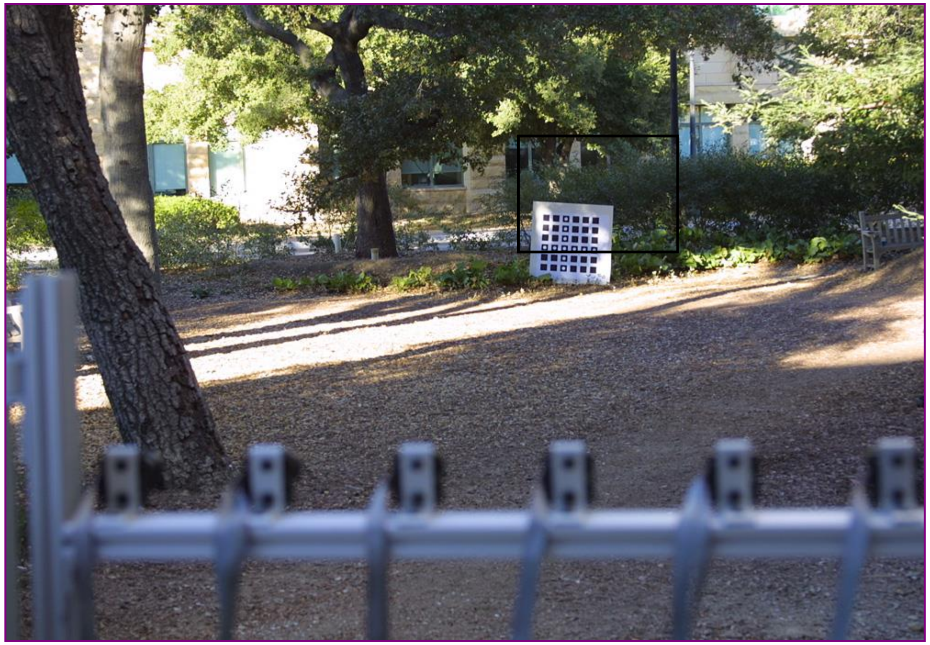
Intermediate camera spacing: synthetic aperture photography





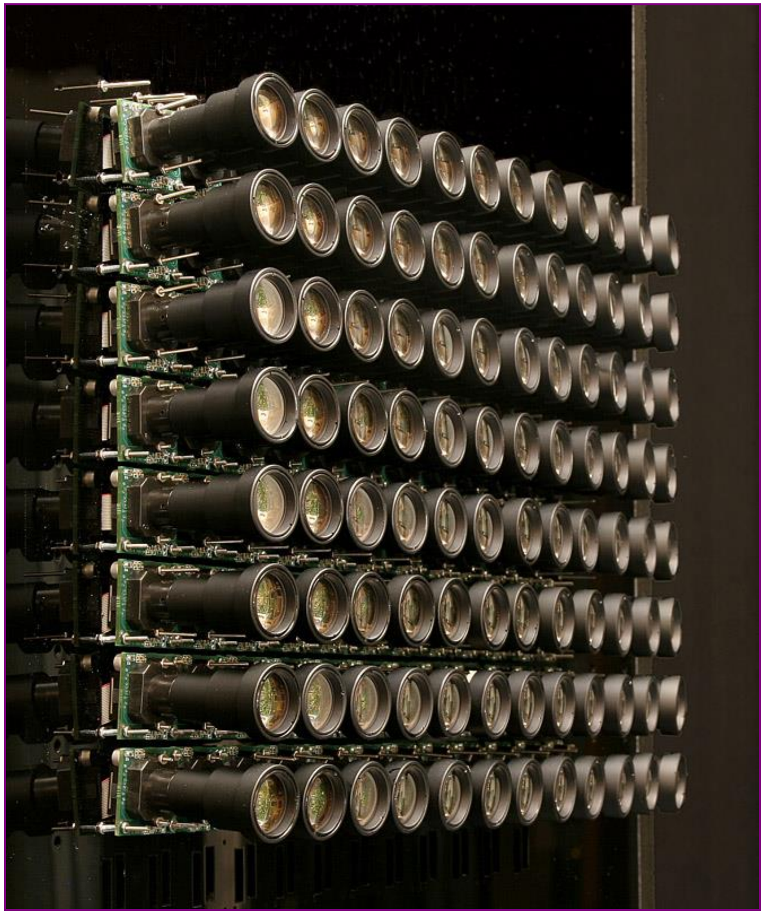
Example using 45 cameras

[Vaish CVPR]



Tiled camera array

Can we match the image quality of a cinema camera?



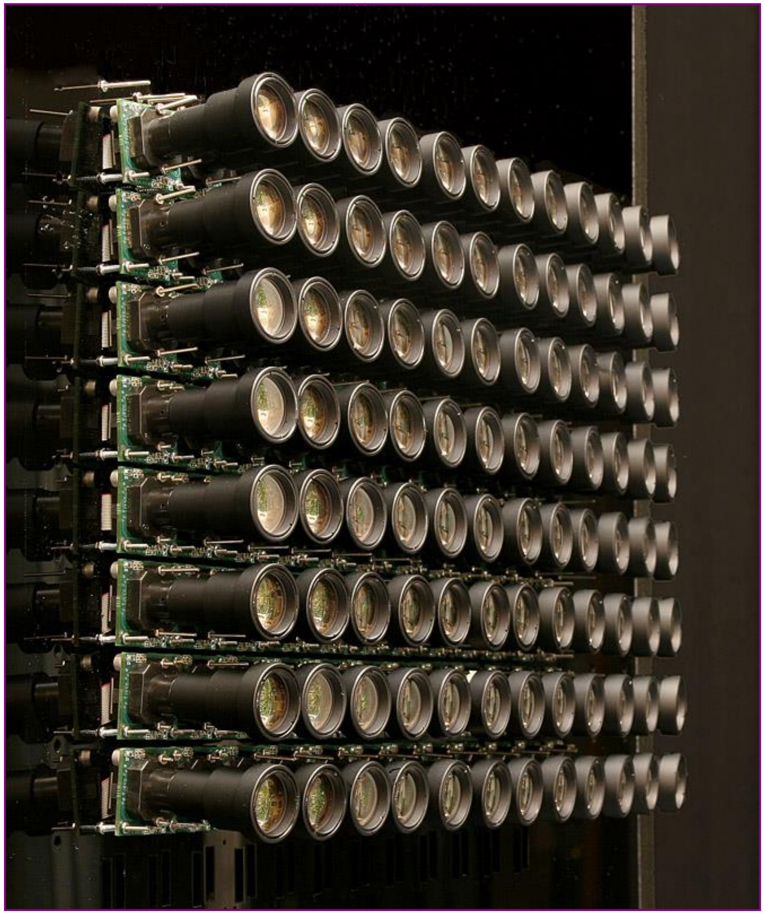
- world's largest video camera
- no parallax for distant objects
- poor lenses limit image quality
- seamless mosaicing isn't hard

Tiled panoramic image (before geometric or color calibration)

Tiled panoramic image (after calibration and blending)

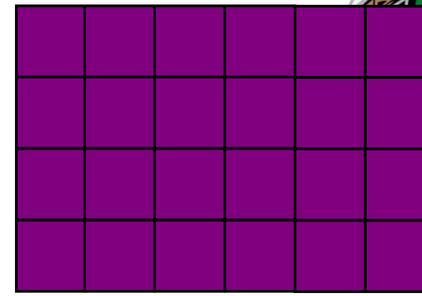
Tiled camera array

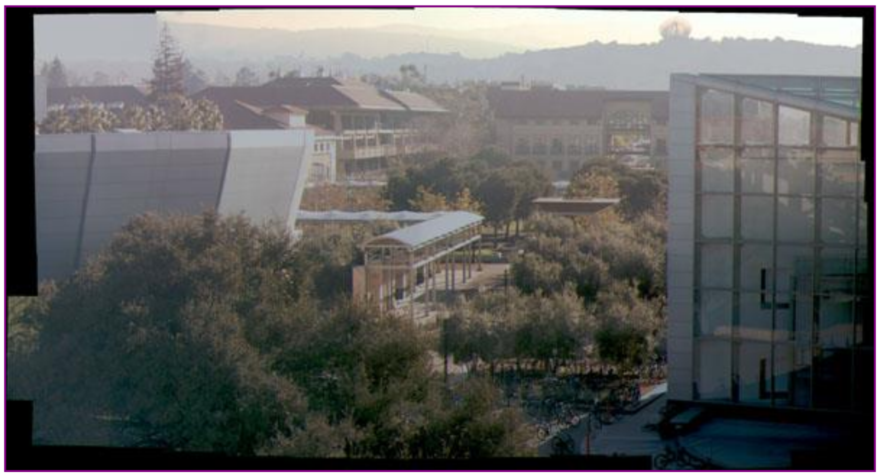
Can we match the image quality of a cinema camera?



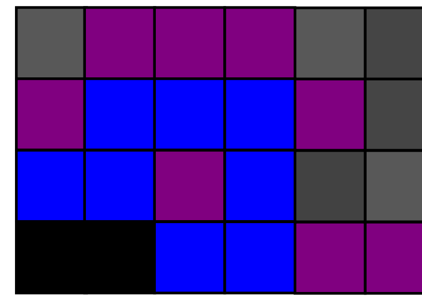
- world's largest video camera
- no parallax for distant objects
- poor lenses limit image quality
- seamless mosaicing isn't hard
- per-camera exposure

same
exposure in
all cameras

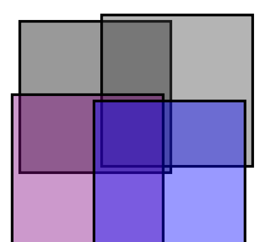


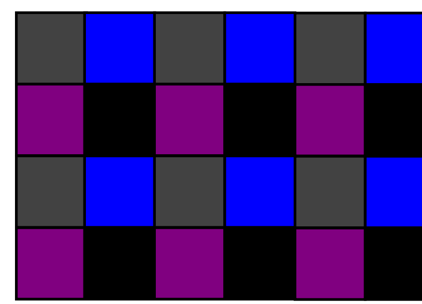


individually
metered



checkerboard
of exposures

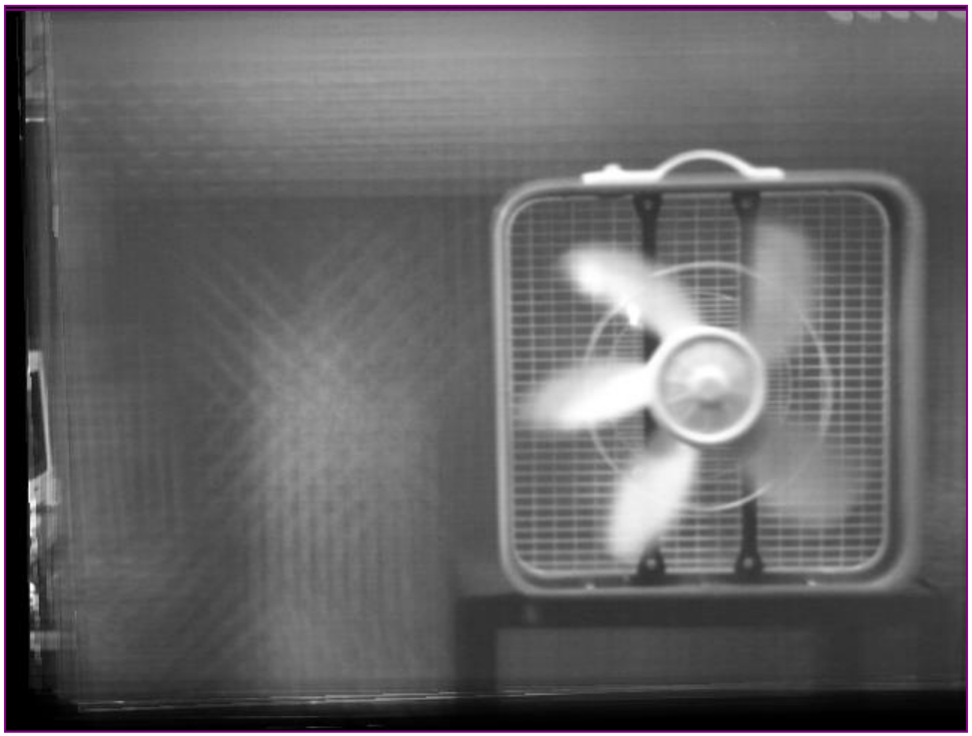




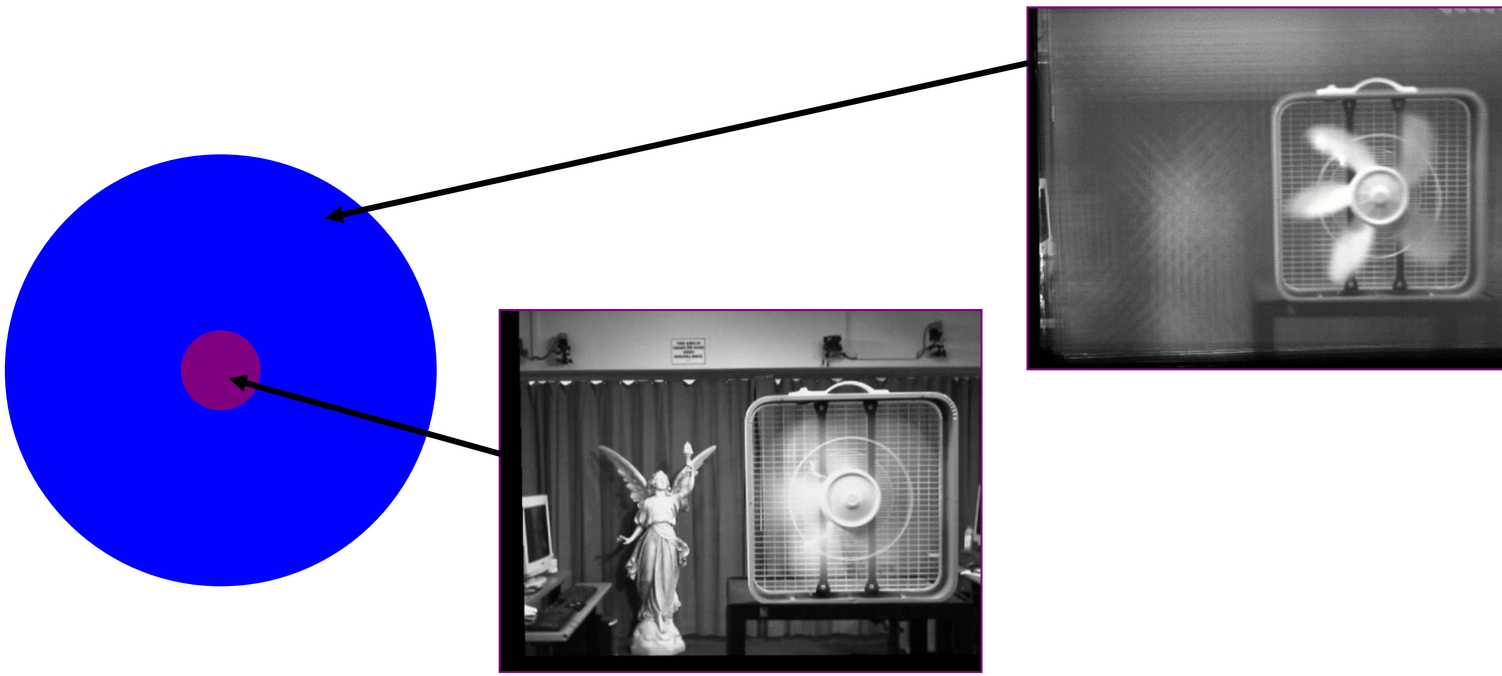
High-performance photography as multi-dimensional sampling

- spatial resolution
- field of view
- frame rate
- dynamic range
- bits of precision
- depth of field
- focus setting
- color sensitivity

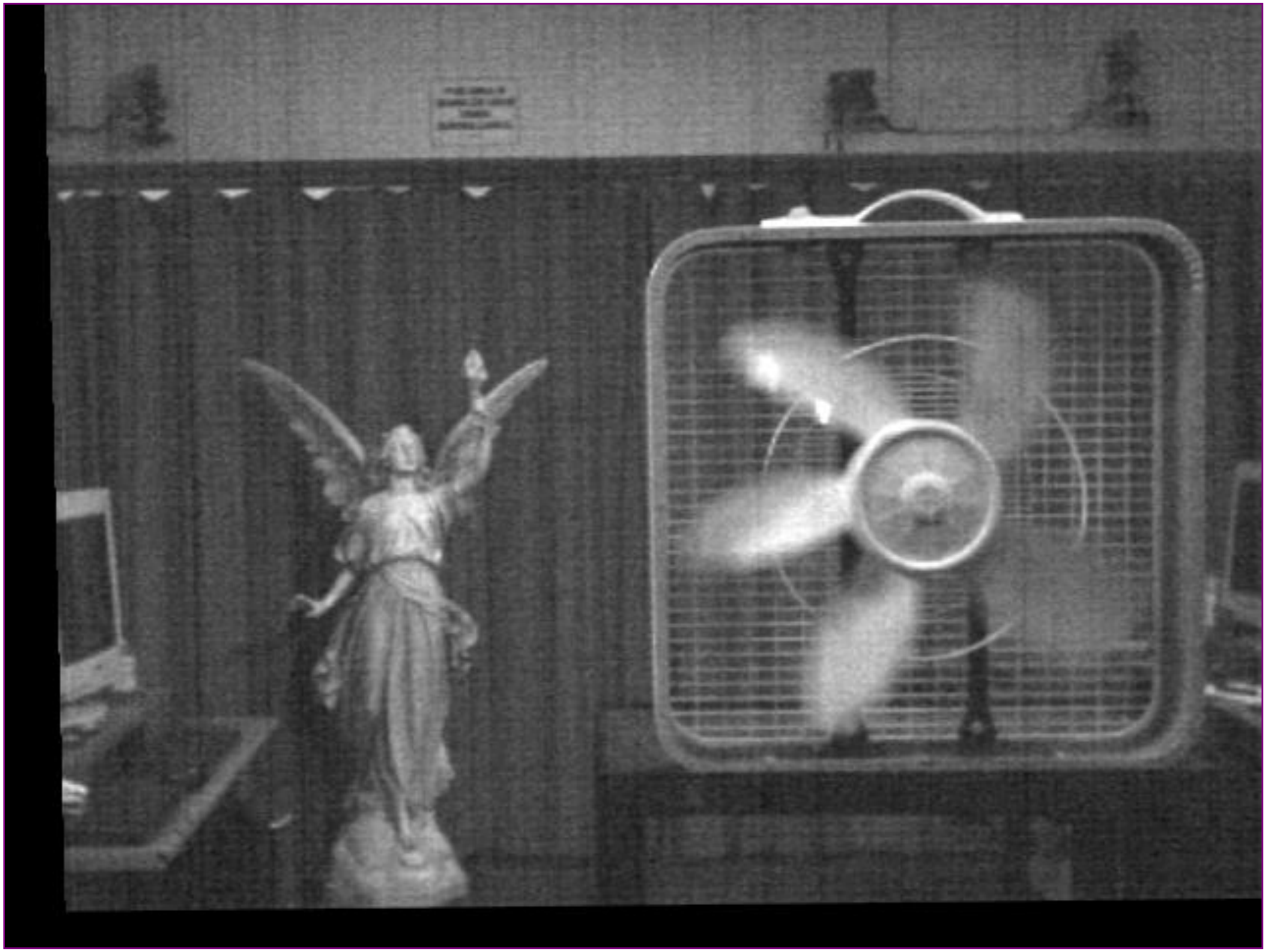
Spacetime aperture shaping



- shorten exposure time to freeze motion → dark
- stretch contrast to restore level → noisy
- increase (synthetic) aperture to capture more light → decreases depth of field

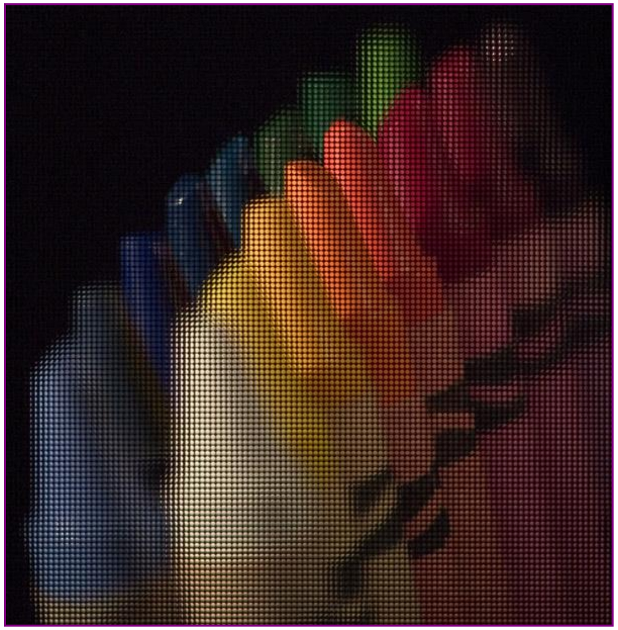


- center of aperture: few cameras, long exposure →
  high depth of field, low noise, but action is blurred
- periphery of aperture: many cameras, short exposure →
  freezes action, low noise, but low depth of field



Light field photography using a handheld plenoptic camera

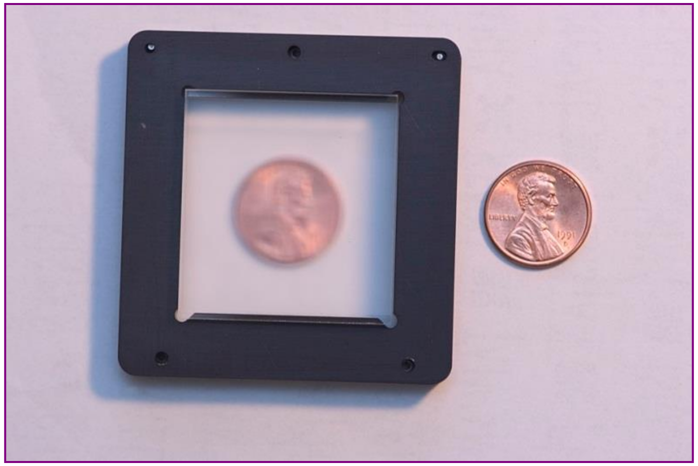
Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz and Pat Hanrahan



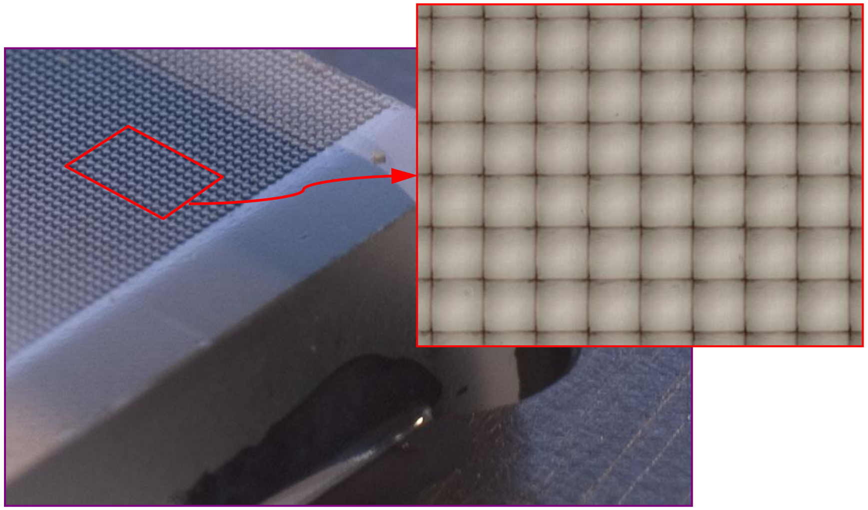
Prototype camera

Contax medium format camera

Kodak 16-megapixel sensor



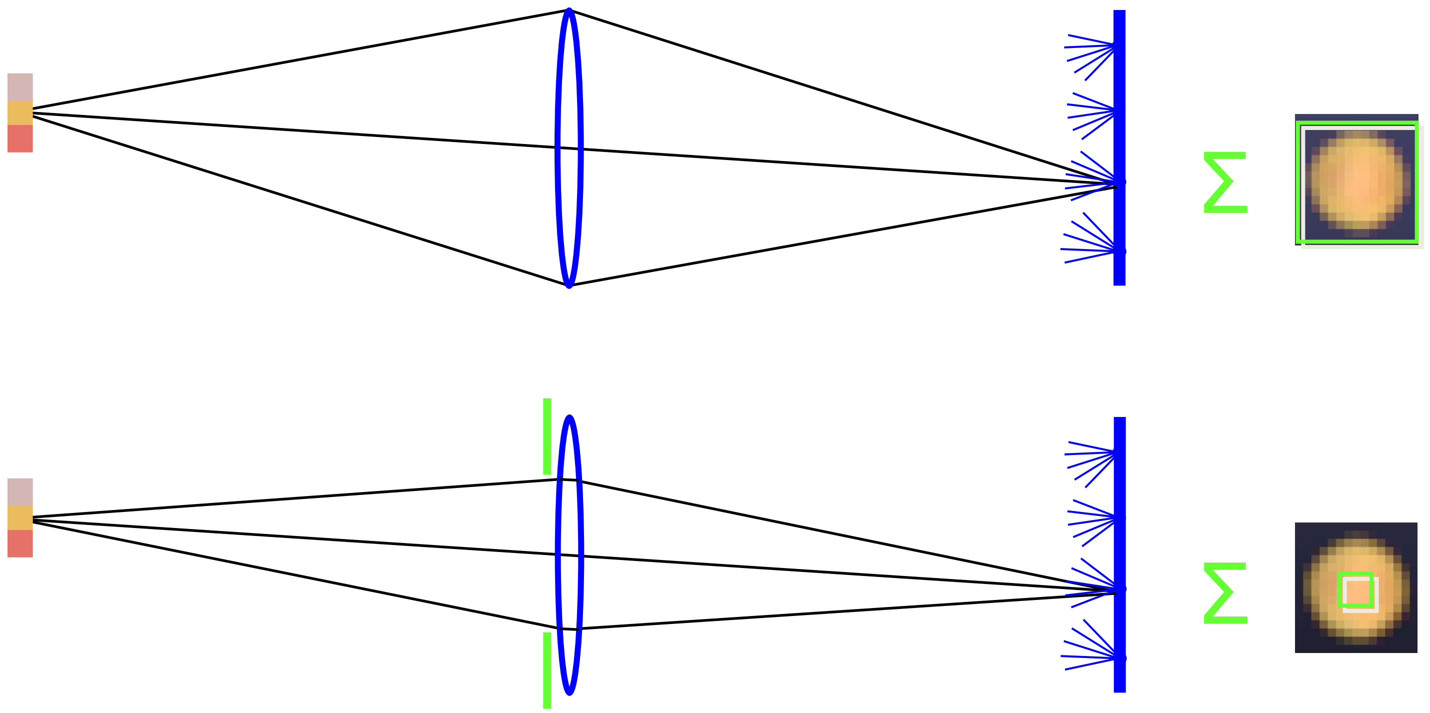
Adaptive Optics microlens array



125μ square-sided microlenses

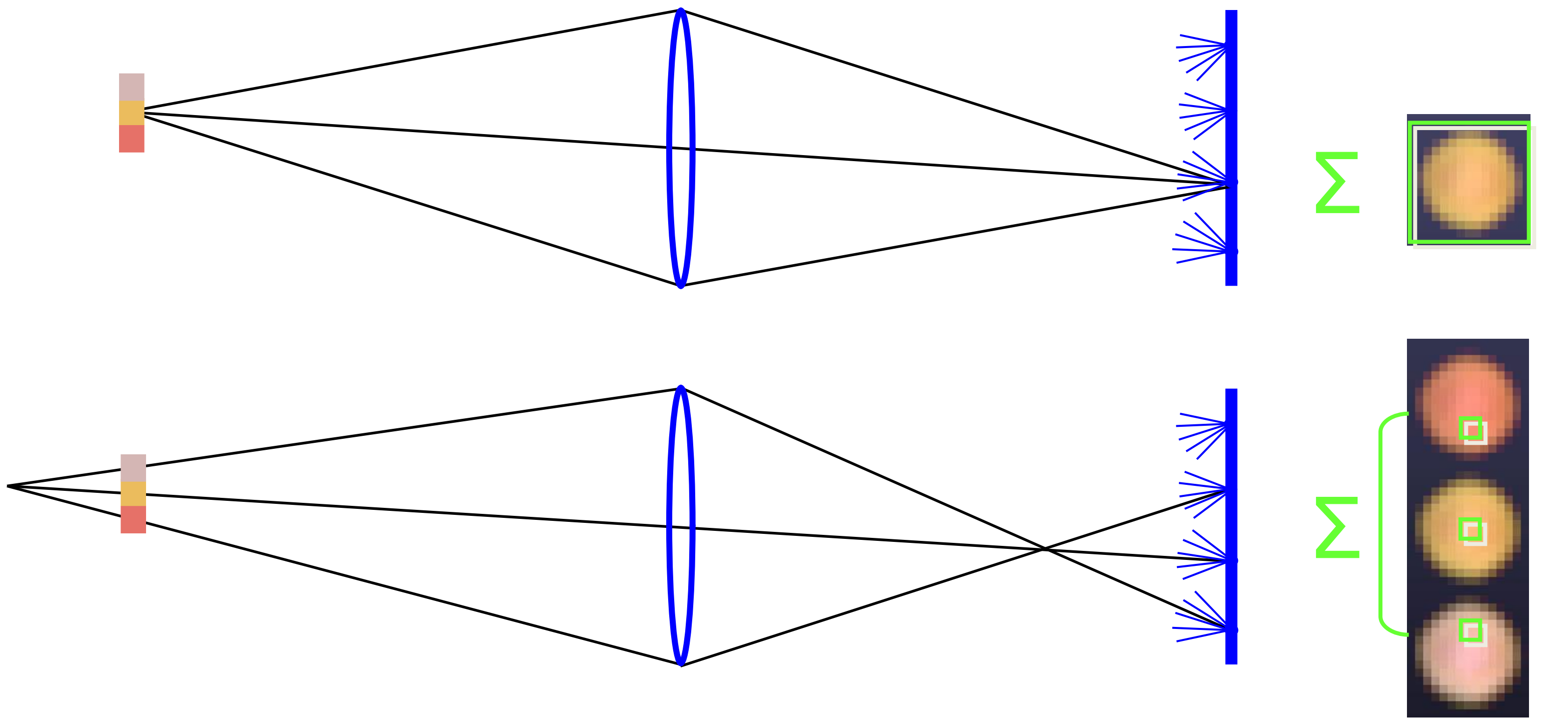
4000 × 4000 pixels ÷ 292 × 292 lenses = 14 × 14 pixels per lens

Digitally stopping-down



- stopping down = summing only the central portion of each microlens

Digital refocusing



- refocusing = summing windows extracted from several microlenses

A digital refocusing theorem

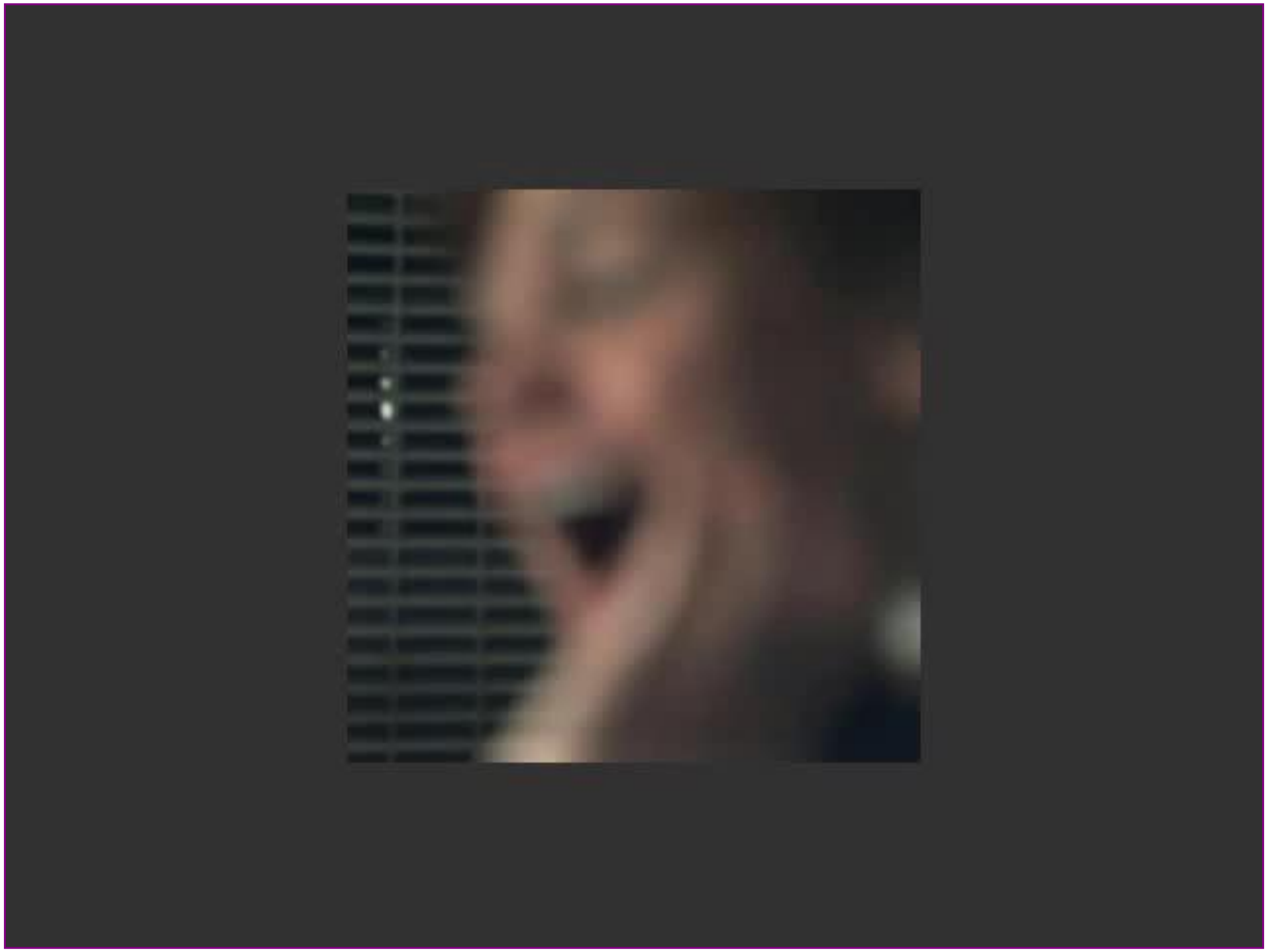
- an f / N light field camera, with P × P pixels under each microlens, can produce views as sharp as an f / (N × P) conventional camera

  – or –

- it can produce views with a shallow depth of field (f / N) focused anywhere within the depth of field of an f / (N × P) camera

Example of digital refocusing

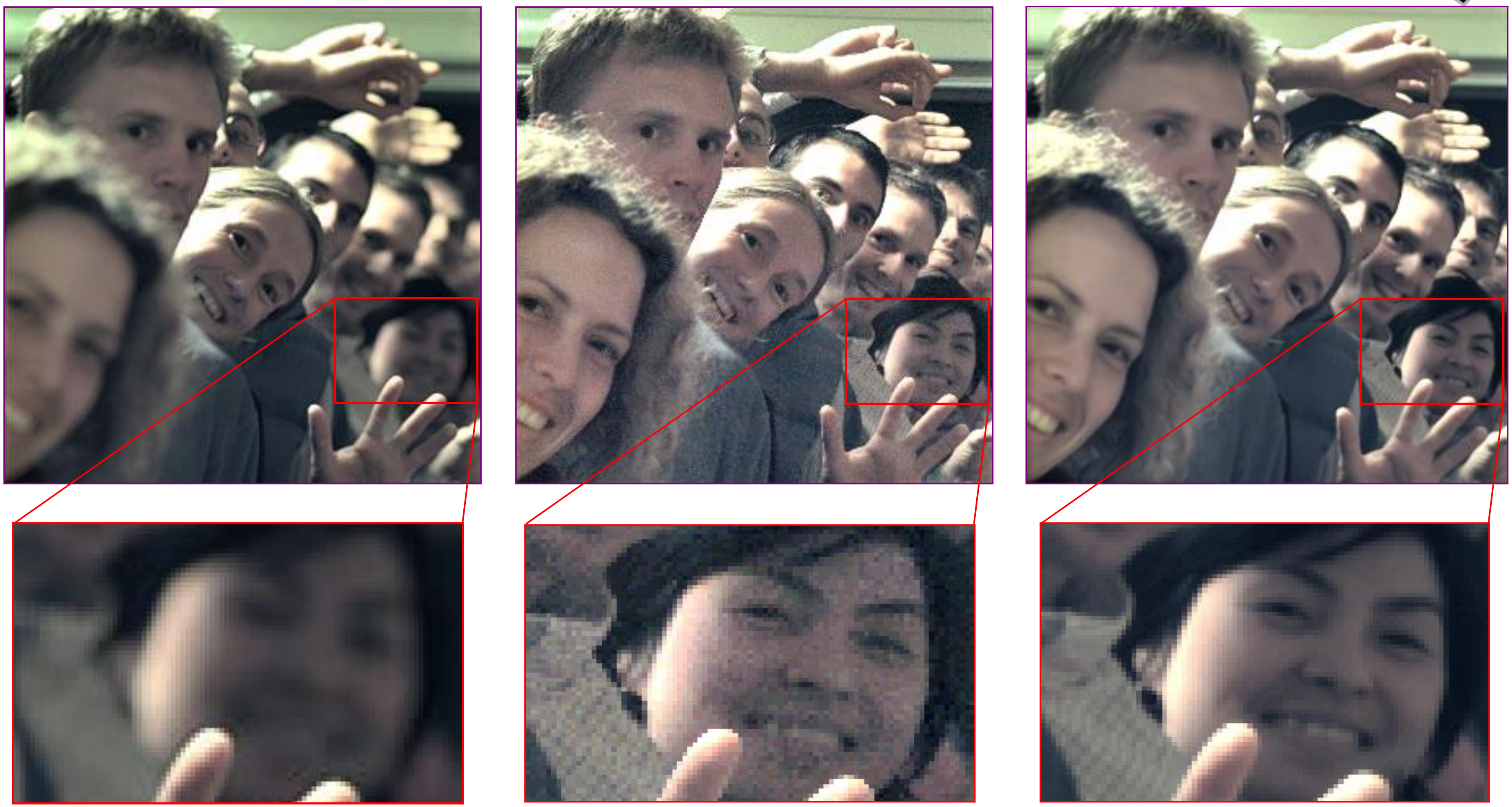
Refocusing portraits



Action photography

Focusing through a splash of water

Extending the depth of field



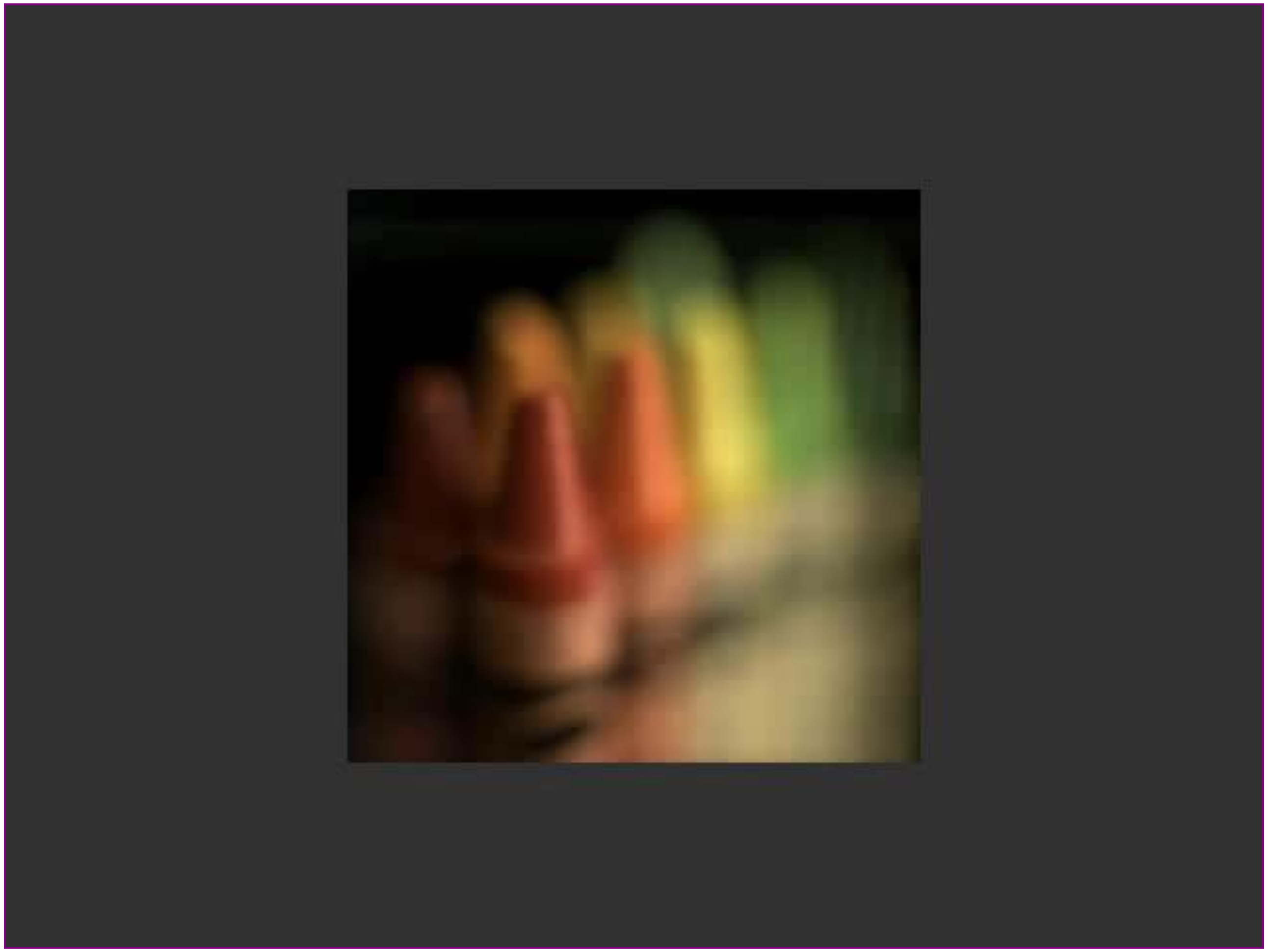
conventional photograph, main lens at f / 4

conventional photograph, main lens at f / 22

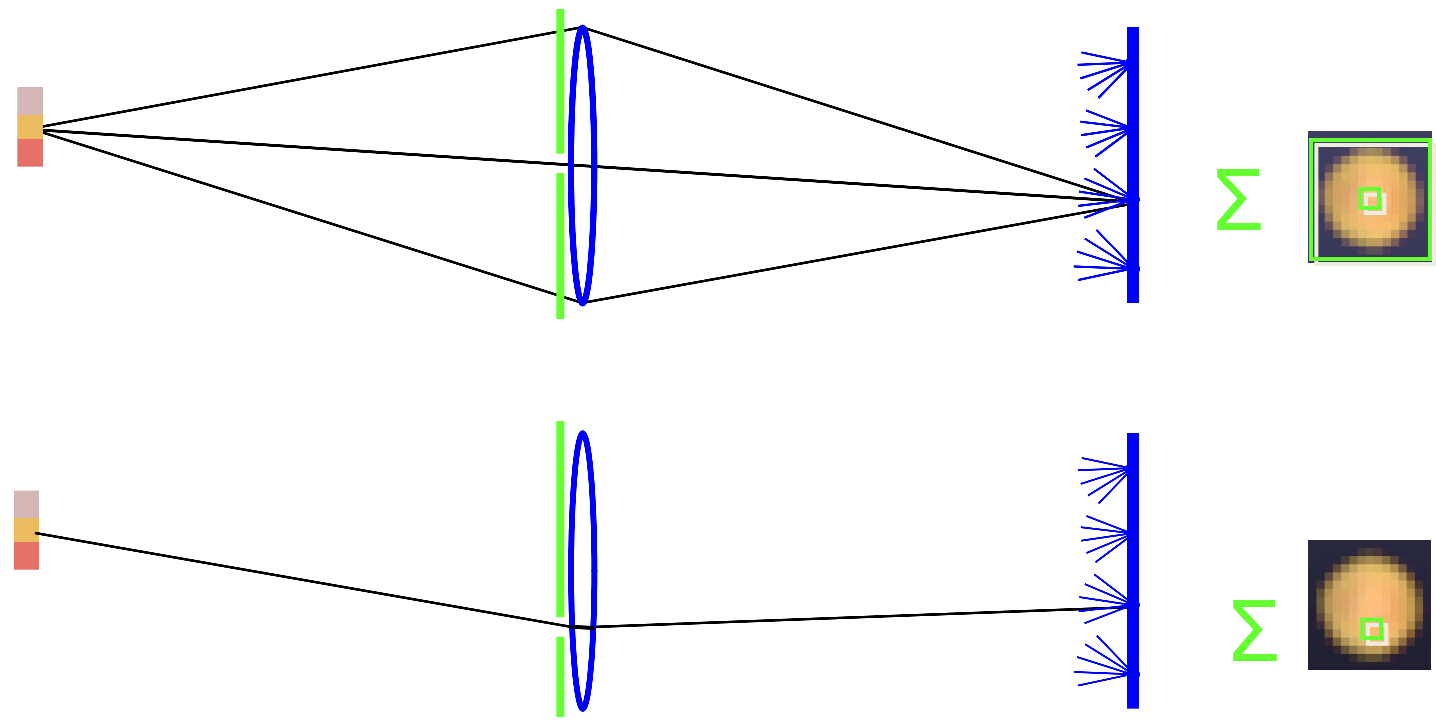
light field, main lens at f / after all-focus algorithm

[Agarwala 2004]

Macrophotography

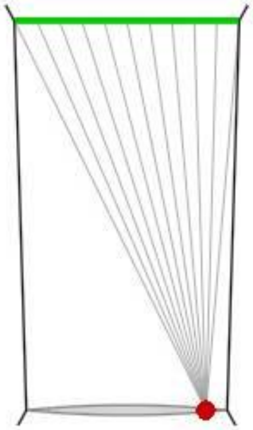


Digitally moving the observer

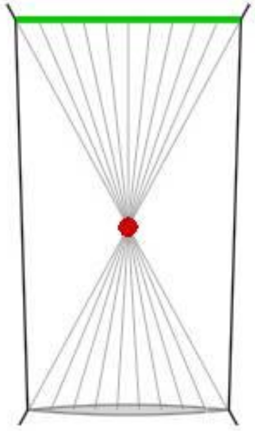


- moving the observer = moving the window we extract from the microlenses

Example of moving the observer

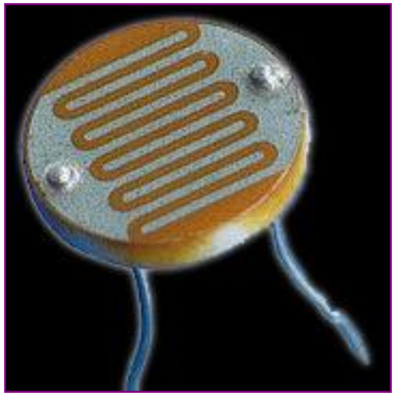


Moving backward and forward

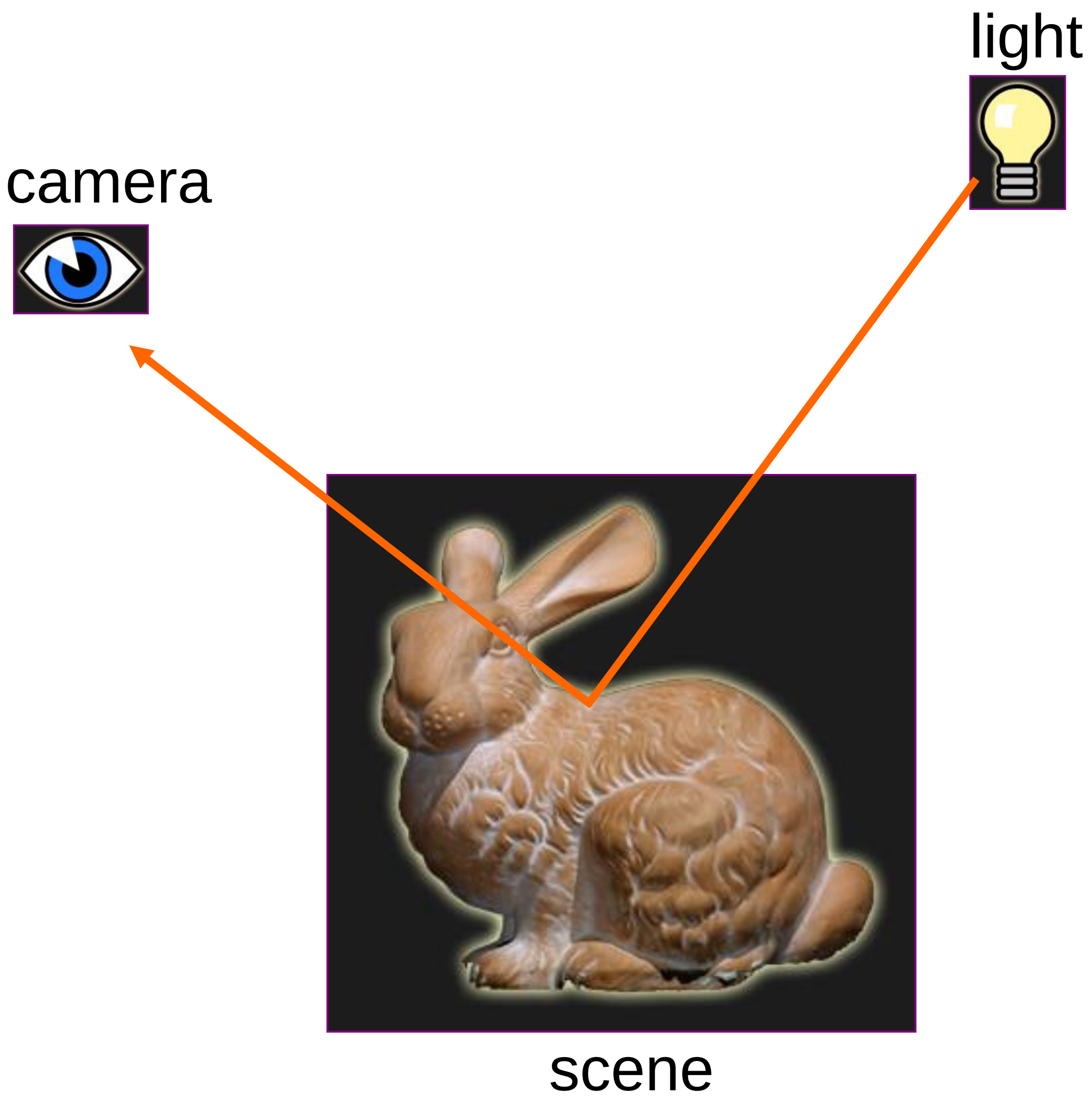


Dual Photography

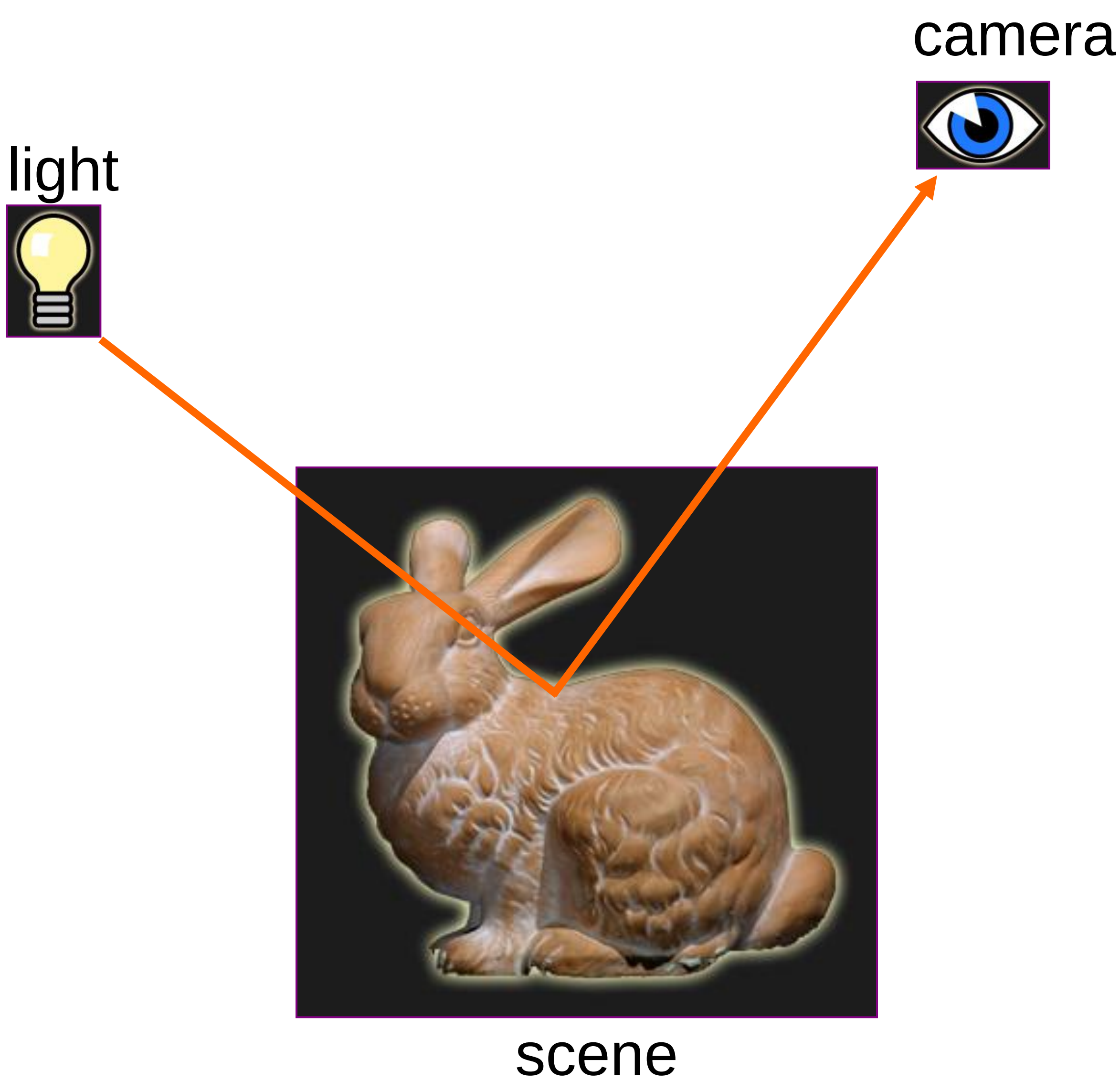
Pradeep Sen, Billy Chen, Gaurav Garg, Steve Marschner,
Mark Horowitz, Marc Levoy, Hendrik Lensch



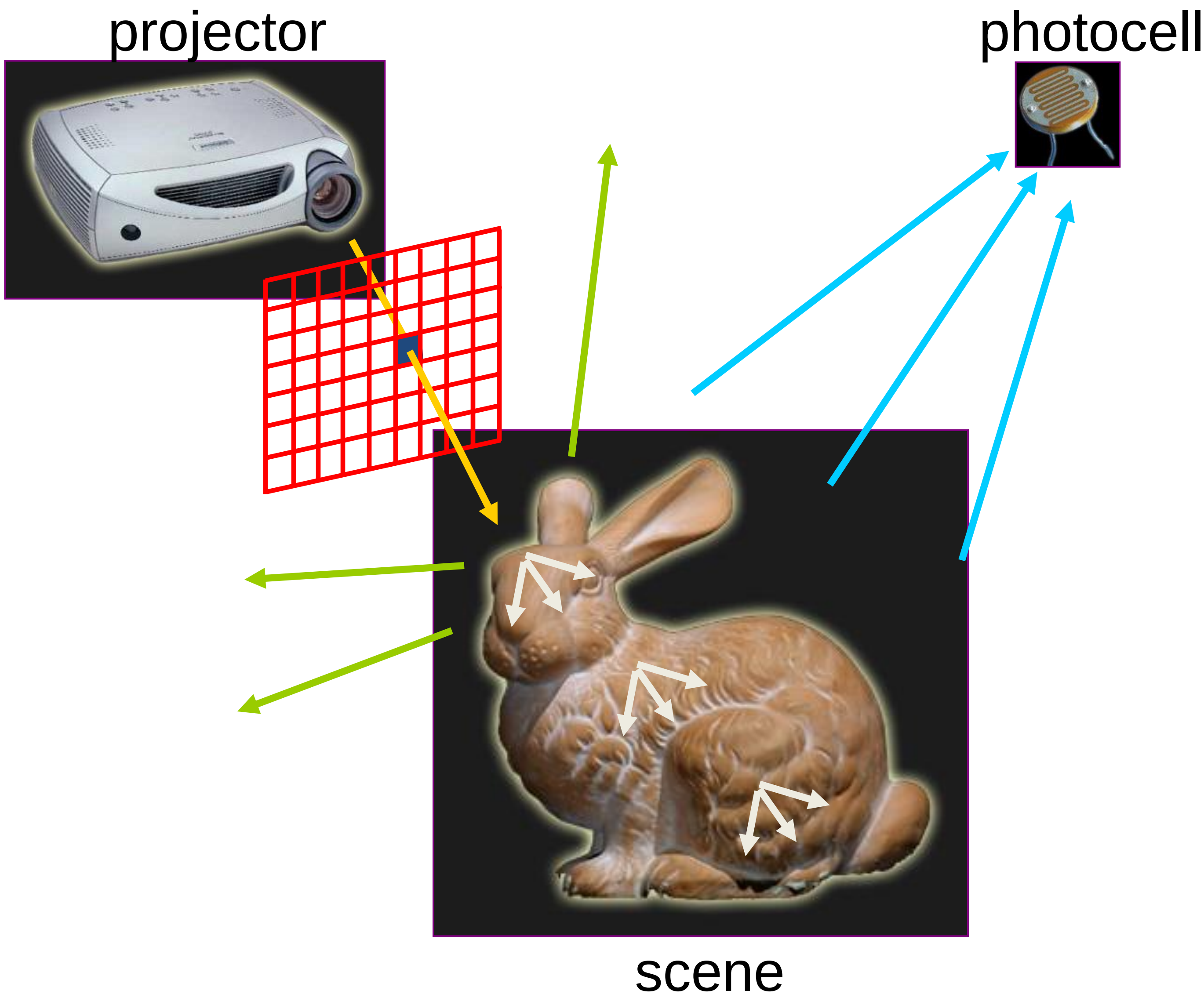
Helmholtz reciprocity



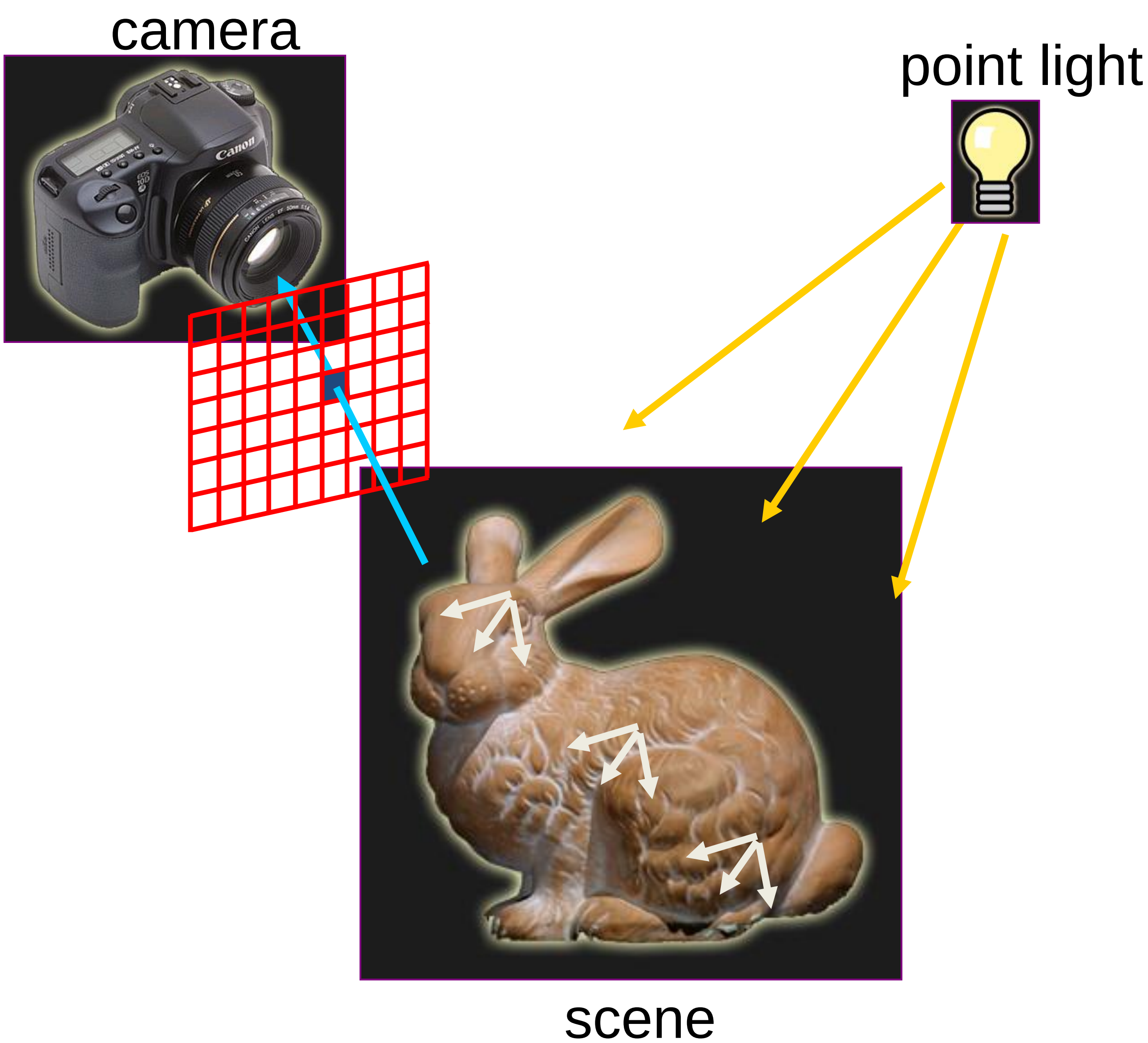
Helmholtz reciprocity



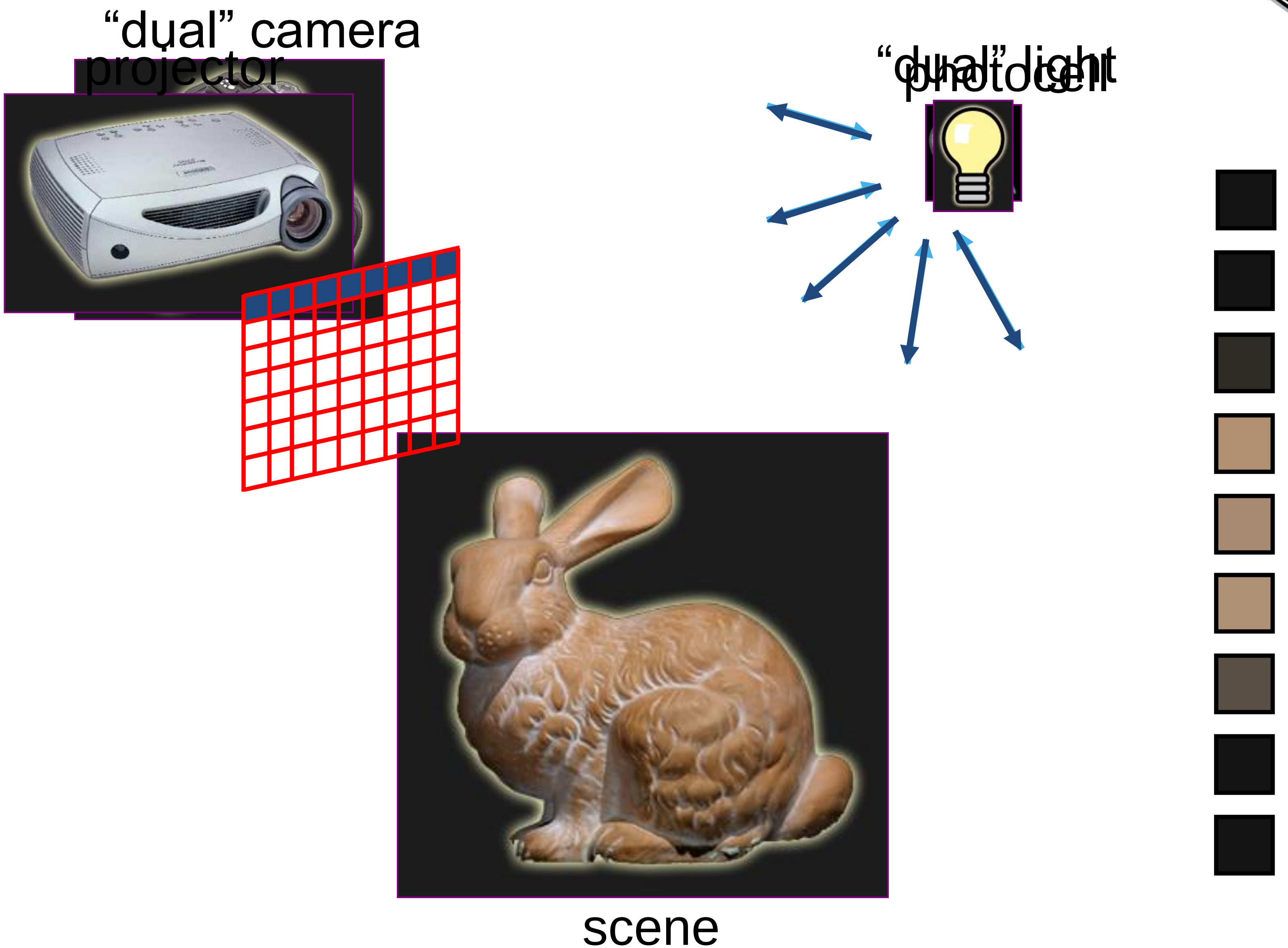
Measuring transport along a set of paths



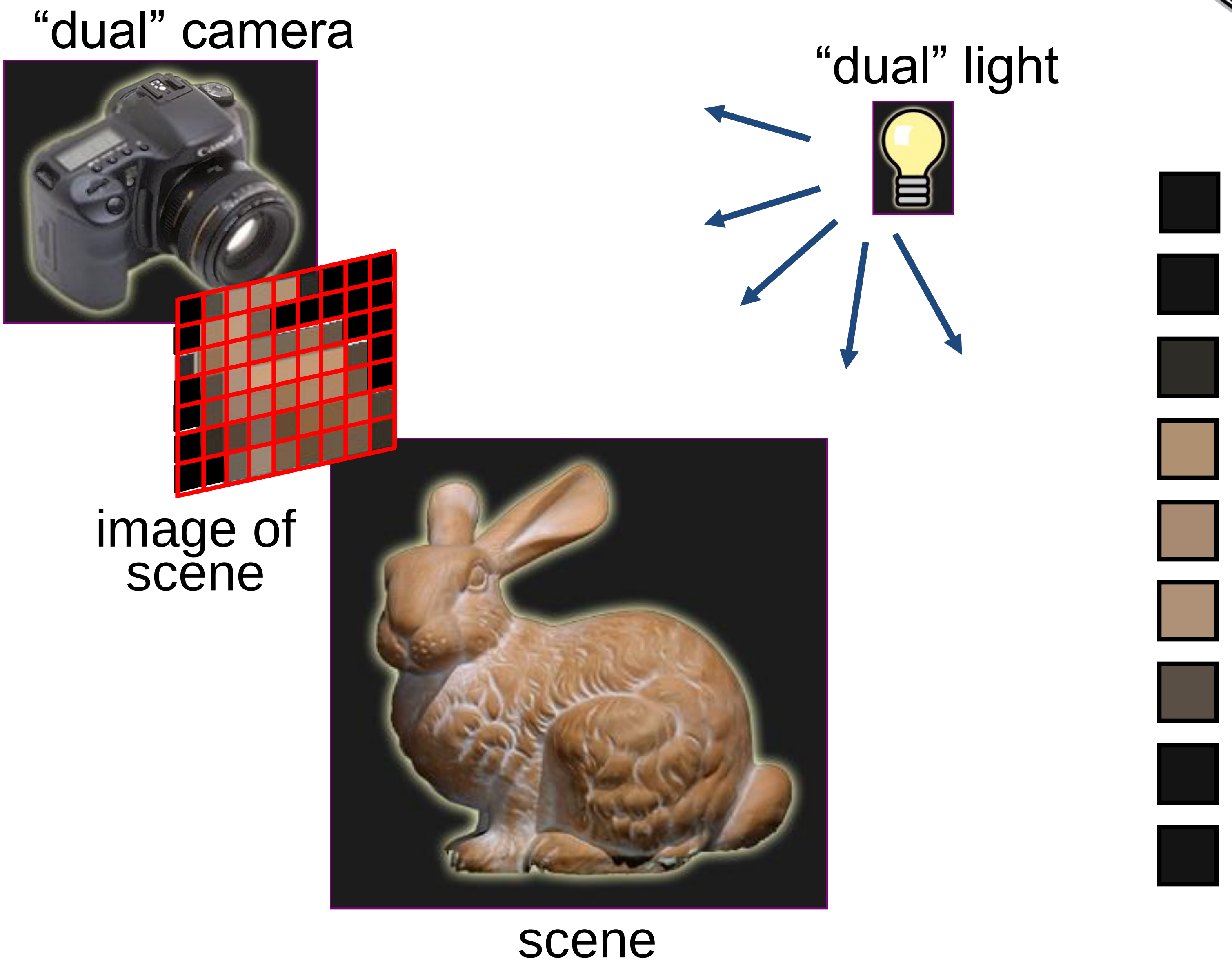
Reversing the paths



Forming a dual photograph



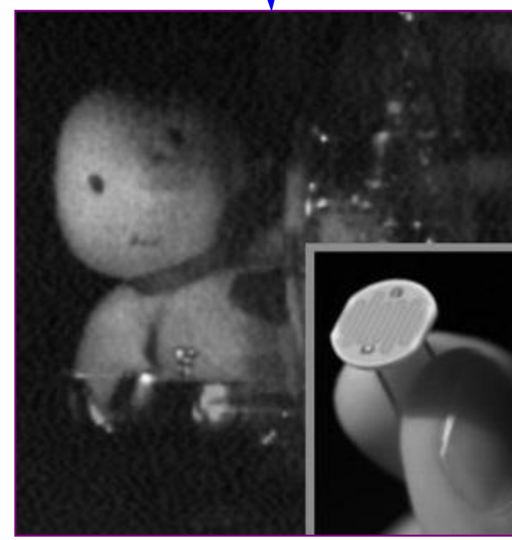
Forming a dual photograph



Physical demonstration

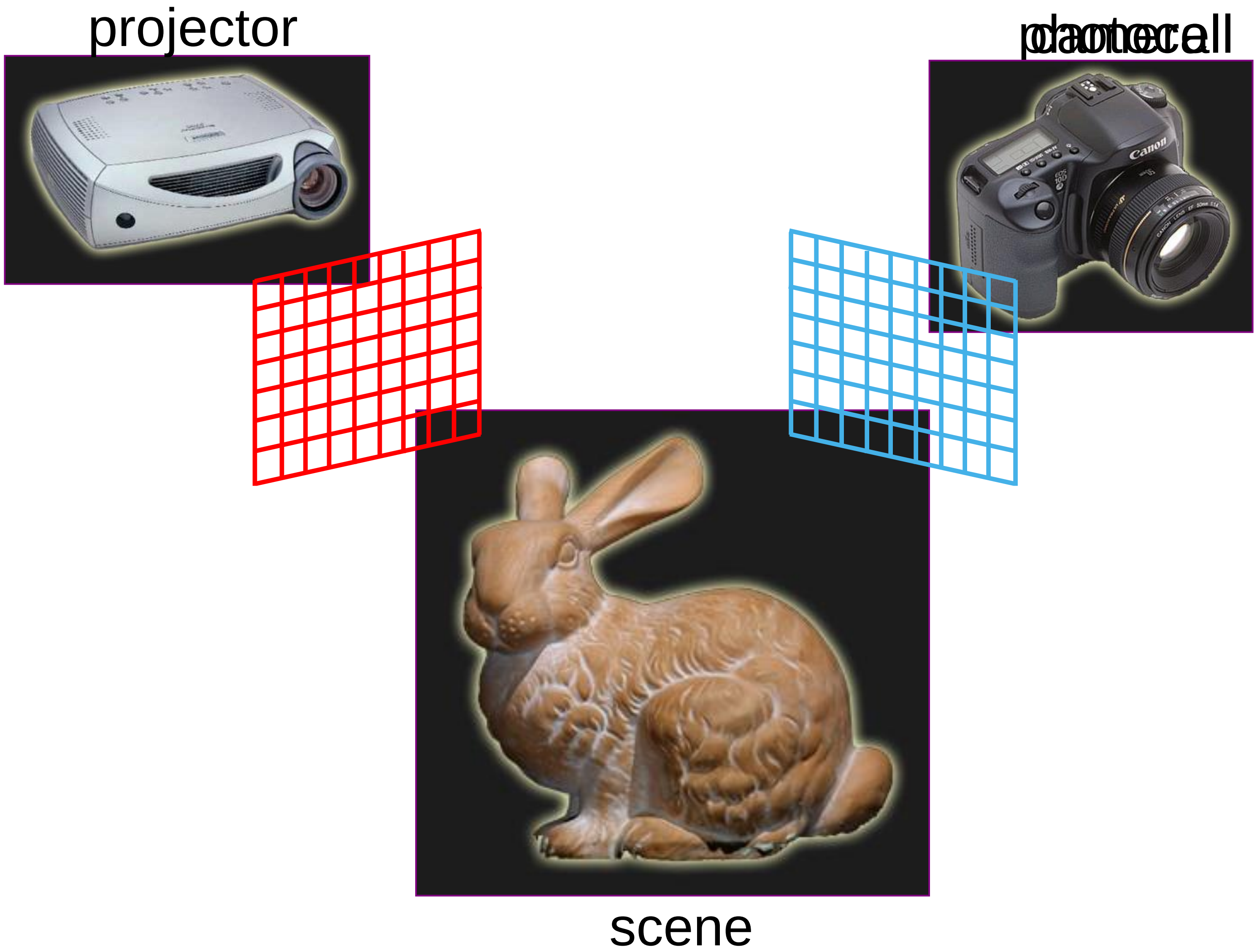
- light replaced with projector
- camera replaced with photocell
- projector scanned across the scene

conventional photograph,
with light coming from right

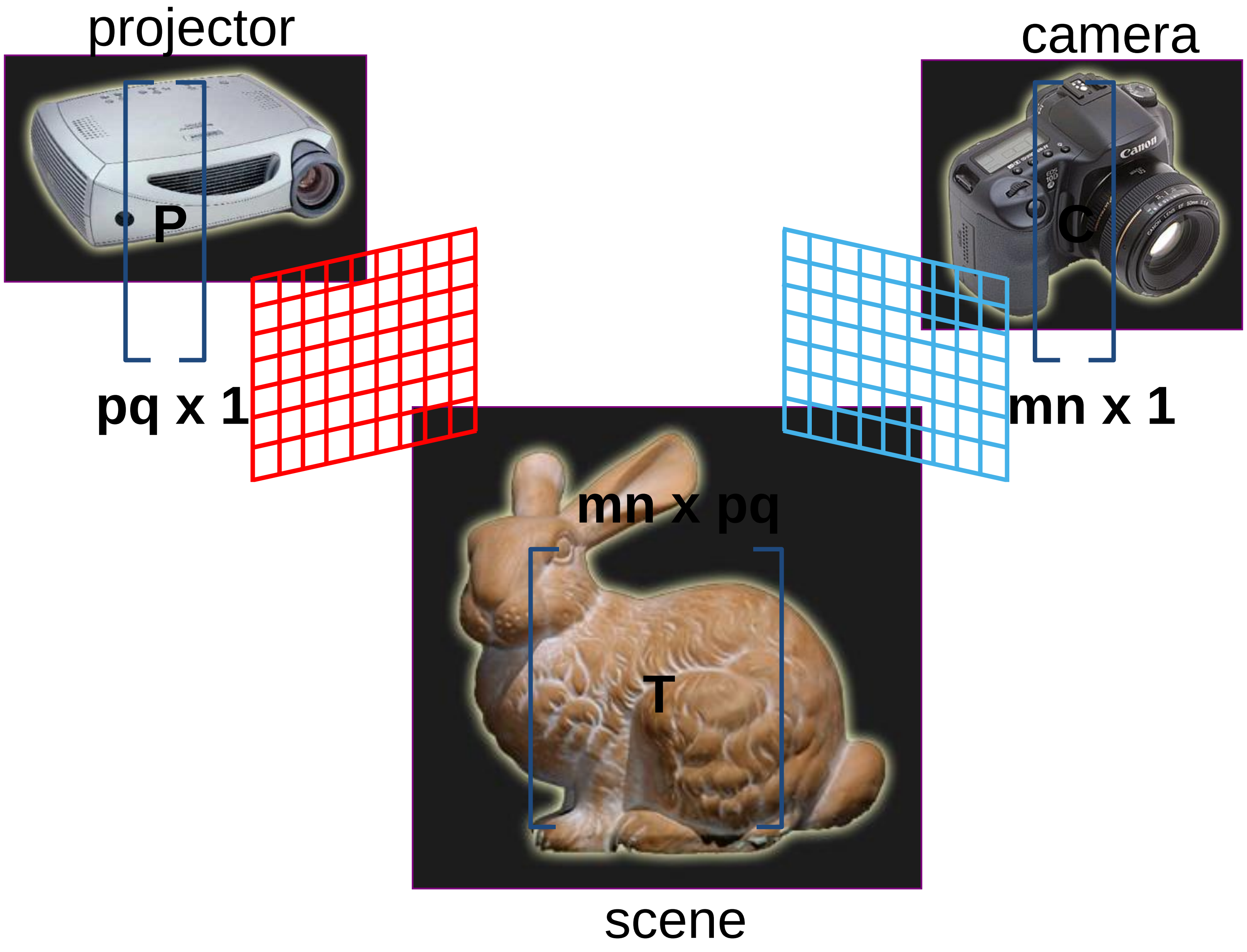


dual photograph,
as seen from projector's position
and as illuminated from photocell's position

The 4D transport matrix



The 4D transport matrix



The 4D transport matrix

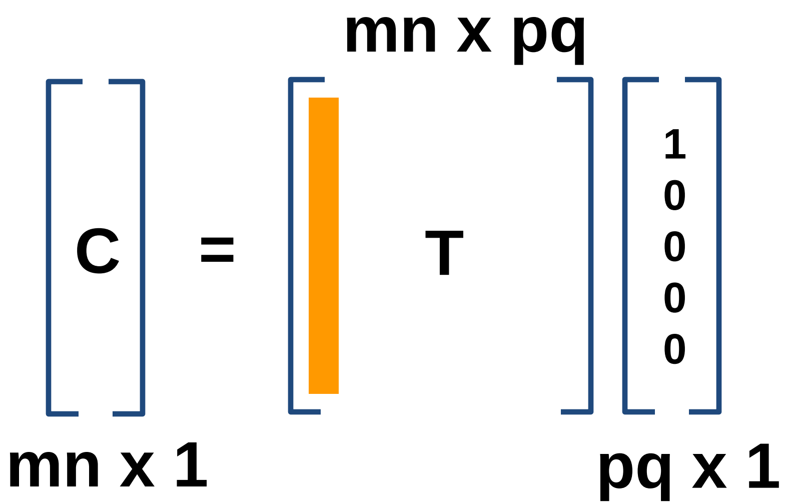
mn x pq

$$\begin{bmatrix} C \end{bmatrix} = \begin{bmatrix} T \end{bmatrix} \begin{bmatrix} P \end{bmatrix}$$

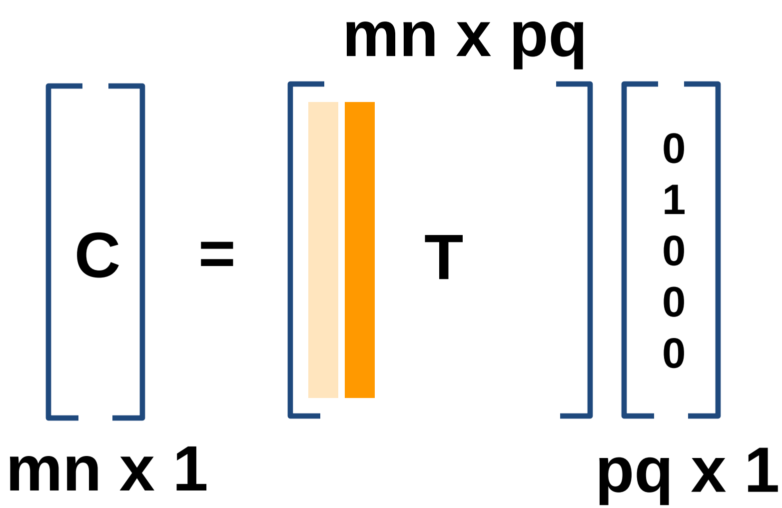
mn x 1

pq x 1

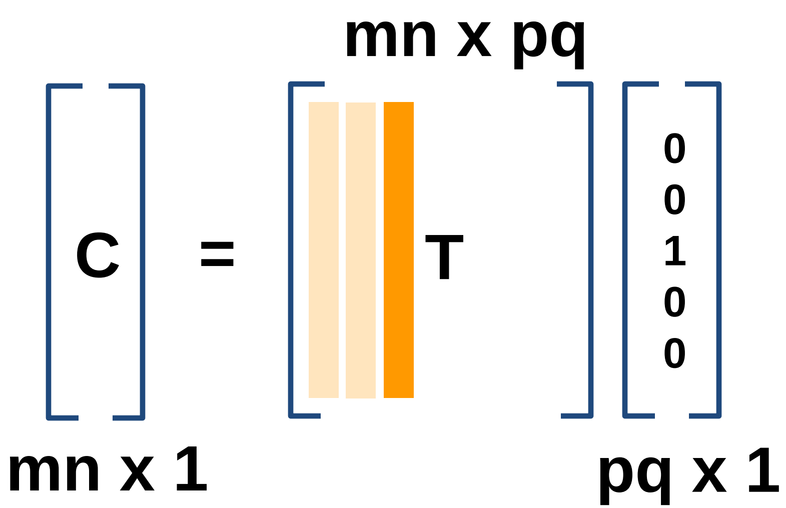
The 4D transport matrix



The 4D transport matrix



The 4D transport matrix



The 4D transport matrix

$$\underbrace{\begin{bmatrix} C \end{bmatrix}}_{mn \times 1} = \overbrace{\begin{bmatrix} \quad T \quad \end{bmatrix}}^{mn \times pq} \underbrace{\begin{bmatrix} P \end{bmatrix}}_{pq \times 1}$$

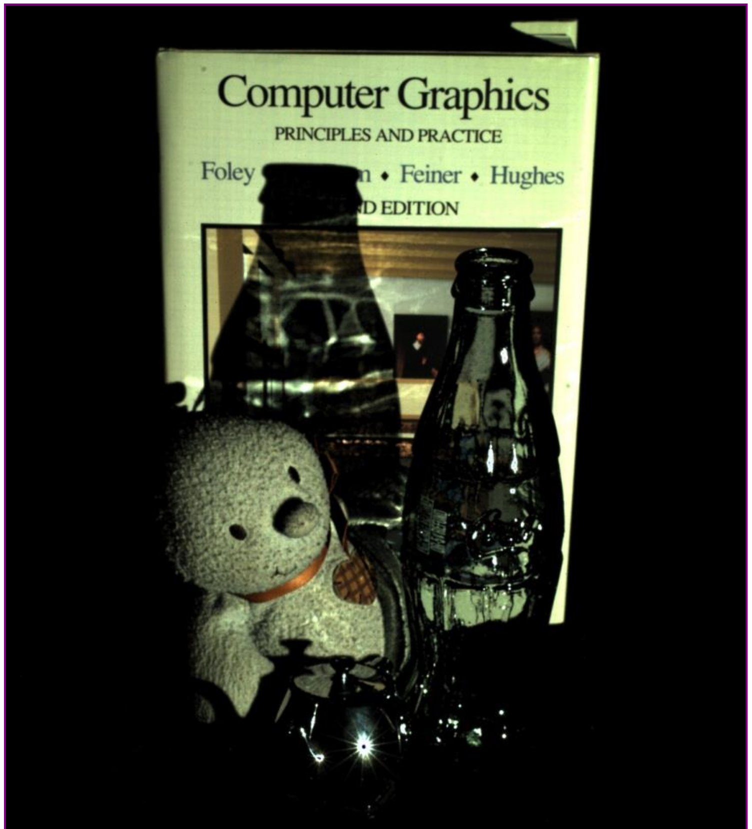
The 4D transport matrix

$$\underset{mn \times 1}{\begin{bmatrix} C \end{bmatrix}} = \underset{mn \times pq}{\begin{bmatrix} \quad T \quad \end{bmatrix}} \underset{pq \times 1}{\begin{bmatrix} P \end{bmatrix}}$$

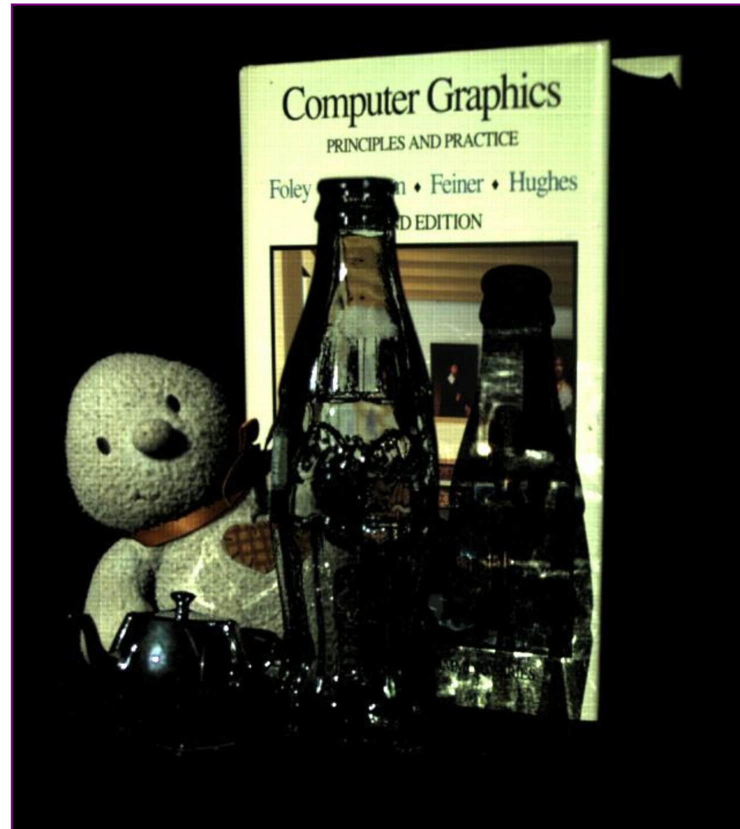
applying Helmholtz reciprocity...

$$\underset{pq \times 1}{\begin{bmatrix} C' \end{bmatrix}} = \underset{pq \times mn}{\begin{bmatrix} \quad T^T \quad \end{bmatrix}} \underset{mn \times 1}{\begin{bmatrix} P' \end{bmatrix}}$$

Example



conventional photograph
with light coming from right



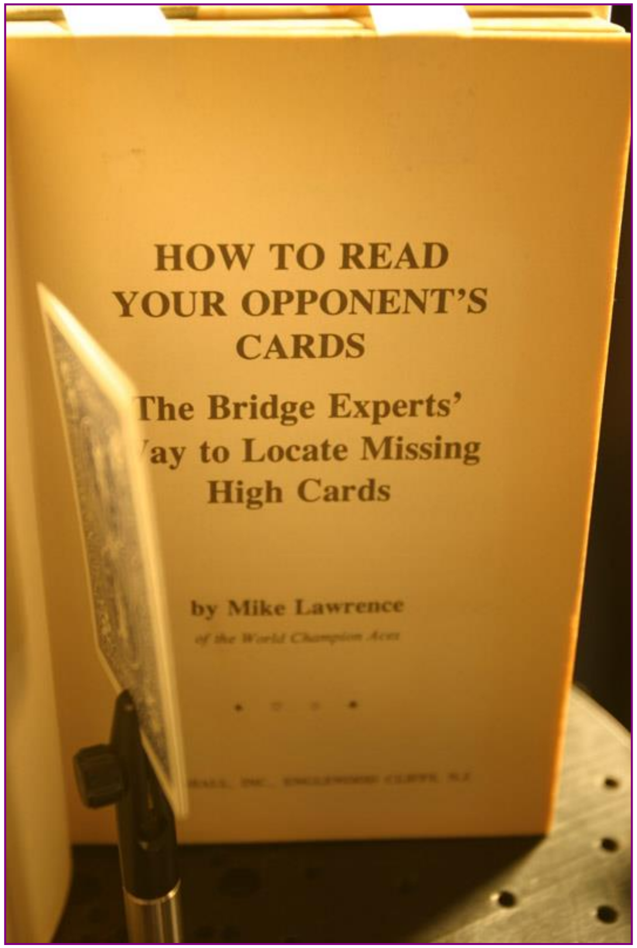
dual photograph
as seen from projector's position

Properties of the transport matrix

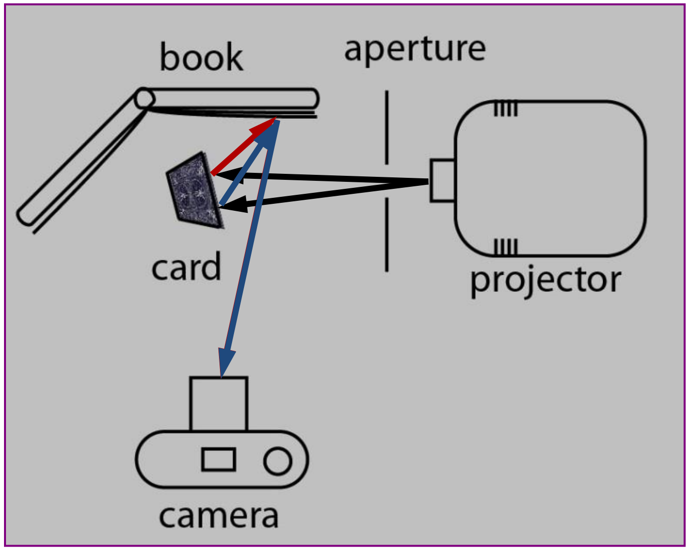
- little interreflection → sparse matrix
- many interreflections → dense matrix
- convex object → diagonal matrix
- concave object → full matrix

Can we create a dual photograph entirely from diffuse reflections?

Dual photography from diffuse reflections



the camera's view

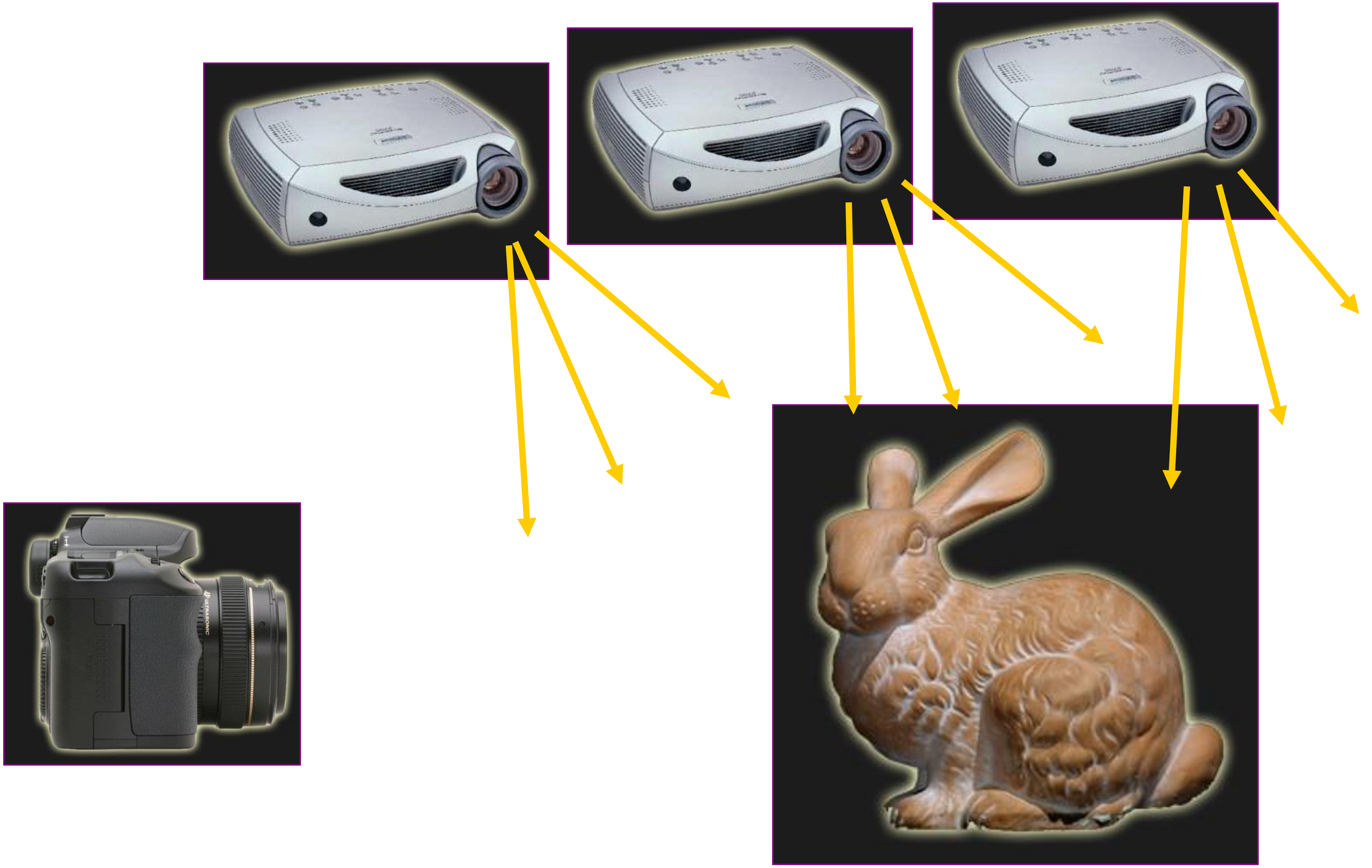


The relighting problem

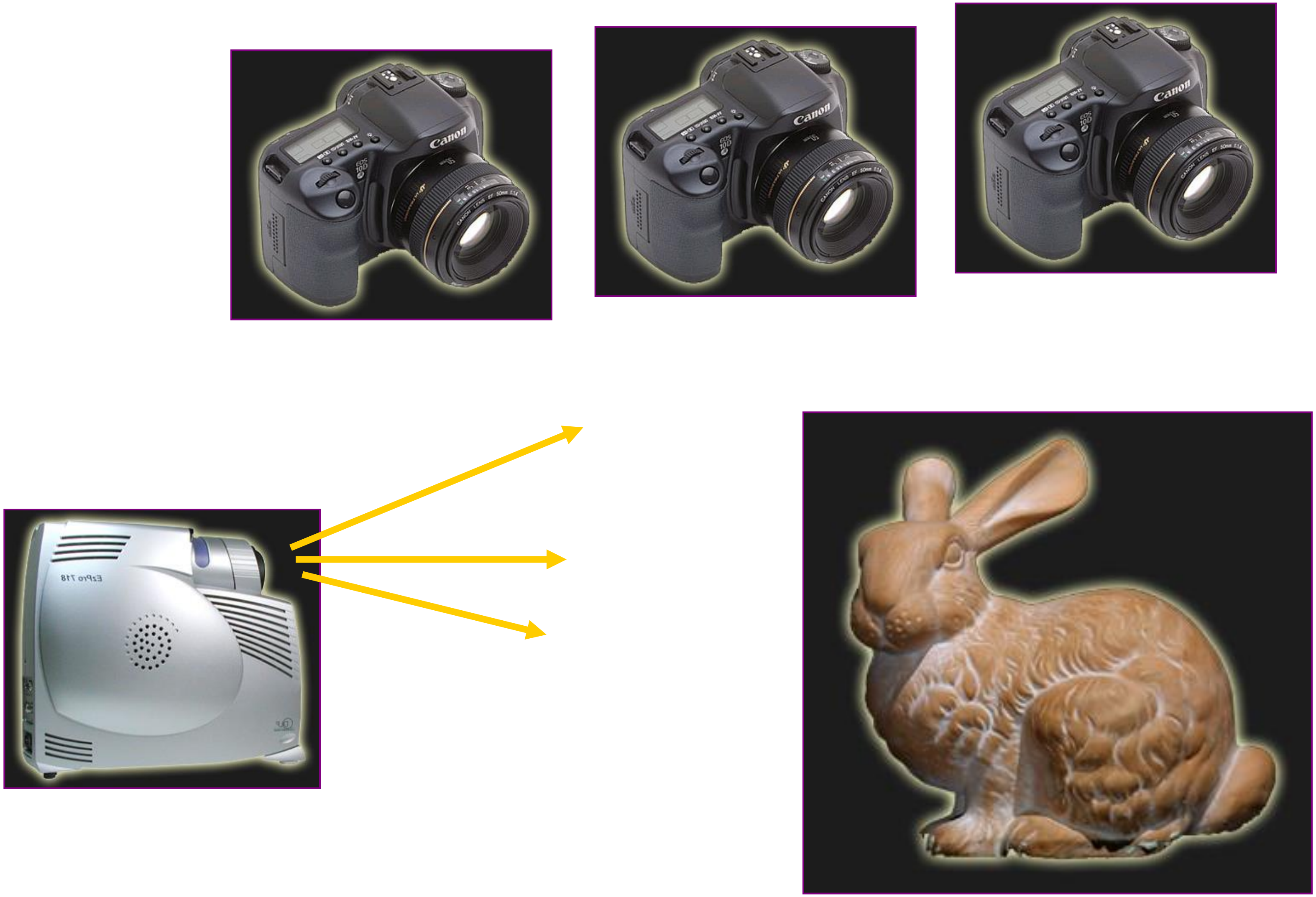
Paul Debevec's
Light Stage 3

- subject captured under multiple lights
- one light at a time, so subject must hold still
- point lights are used, so can't relight with cast shadows

The 6D transport matrix



The 6D transport matrix



The advantage of dual photography

- capture of a scene as illuminated by different lights cannot be parallelized

- capture of a scene as viewed by different cameras <u>can</u> be parallelized

Measuring the 6D transport matrix

projector

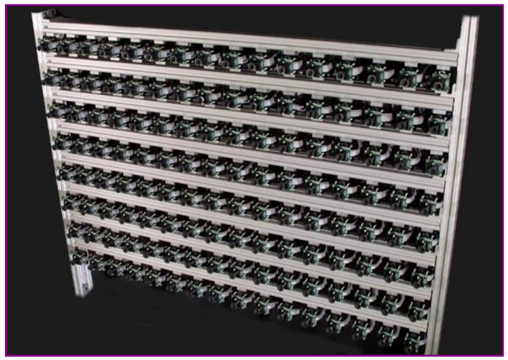
camera array

scene

Relighting with complex illumination

projector

camera array



scene

pq x mn x uv

$$C' = T^T P'$$

pq x 1 mn x uv x 1

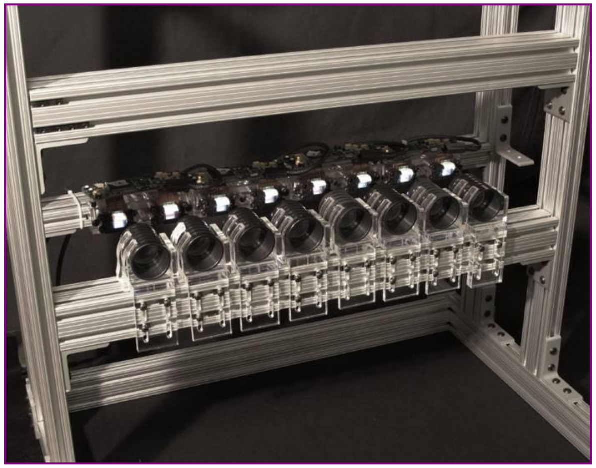
- step 1: measure 6D transport matrix T
- step 2: capture a 4D light field
- step 3: relight scene using captured light field

Running time

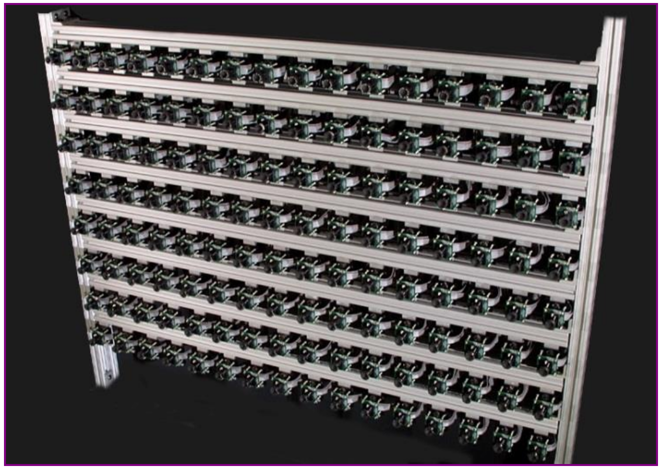
- the different rays within a projector can in fact be parallelized to some extent
- this parallelism can be discovered using a coarse-to-fine adaptive scan
- can measure a 6D transport matrix in 5 minutes

Can we measure an 8D transport matrix?

projector array



camera array



scene

Update

- Google, AR/VR, and Lightfields:
  - https://www.youtube.com/watch?v=lRK0MtIyj0U
- Seeing through things with lightfields:
  - http://graphics.stanford.edu/papers/plane+parallax_calib/
- Microscope Lightfields
  - http://graphics.stanford.edu/projects/lfmicroscope/
- Stanford New Lightfield Archive
  - http://graphics.stanford.edu/data/LF/lfs.html
    - e.g.,"http://graphics.stanford.edu/data/LF/chess_lf/preview.zip&zoom=1"
  - Old: http://graphics.stanford.edu/software/lightpack/lifs.html

Deep Learning Lightfields

- Learning-Based View Synthesis for Light Field Cameras
  - https://www.youtube.com/watch?v=RCD2B5o1K8U

- Light Field Video Capture Using a Learning-Based Hybrid Imaging System
  - https://www.youtube.com/watch?v=TqVKcssYfAo