Understanding camera trade-offs through a Bayesian analysis of light field projections

Anat Levin[1], Bill Freeman[1,2], Fredo Durand[1]

Computer Science and Artificial Intelligence Lab (CSAIL),

[1]Massachusetts Institute of Technology

and [2]Adobe Systems

Cameras, old and new

Traditional camera: Lens forms final 2D image

Cameras, old and new

Traditional camera: Lens forms final 2D image

Computational camera: Recorded data is not the final output.

- Visual array estimated from sensor measurements.
- Extra design degree of freedom.
  Beyond 2D images--acquisition of light field or depth.
  Post-exposure re-synthesis of image.

An explosion of cameras

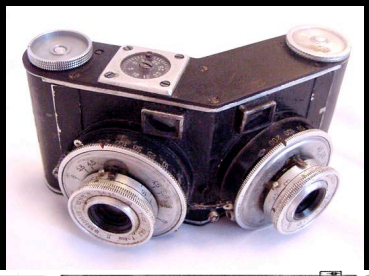
An explosion of cameras

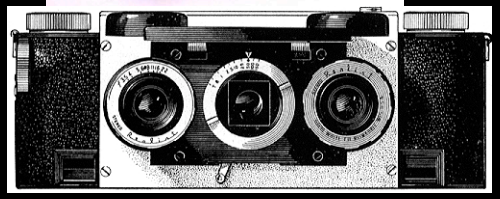
Conventional single-lens cameras

An explosion of cameras

Conventional single-lens cameras

Stereo and trinocular cameras

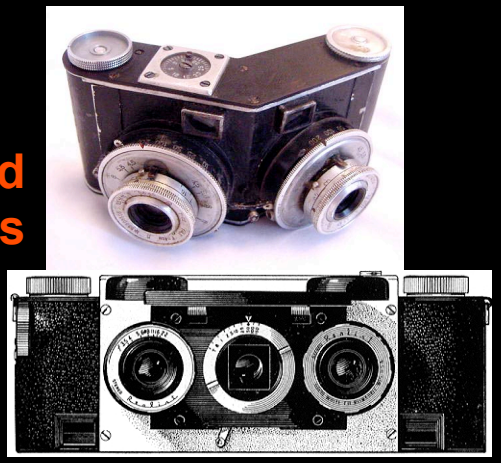




An explosion of cameras

Conventional single-lens cameras

Stereo and trinocular cameras

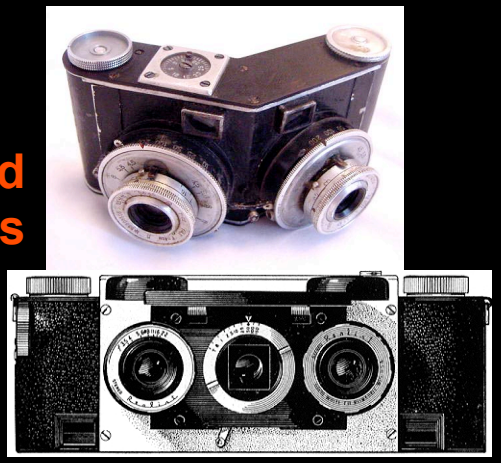


Coded aperture

An explosion of cameras

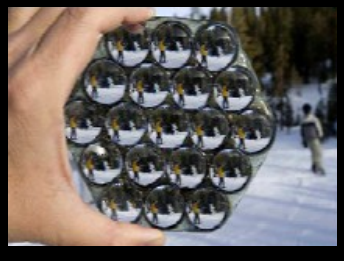
Conventional single-lens cameras

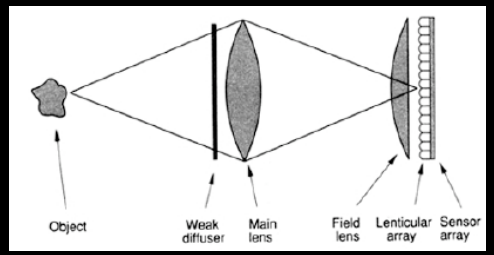
Stereo and trinocular cameras



Coded aperture

Plenoptic cameras

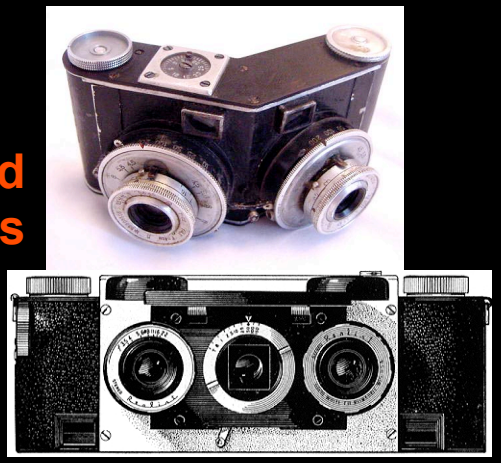




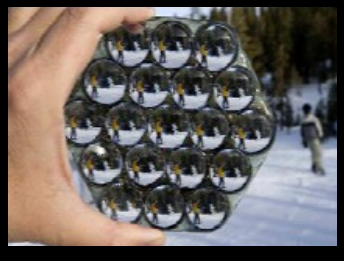
An explosion of cameras

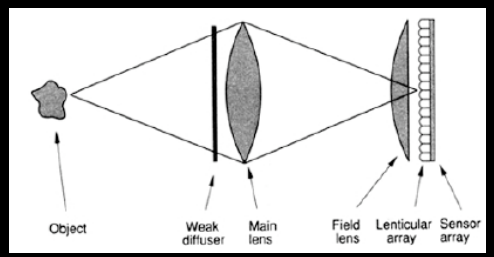
Conventional single-lens cameras

Stereo and trinocular cameras

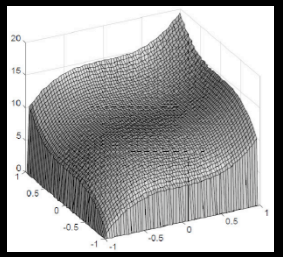


Plenoptic cameras





Coded aperture



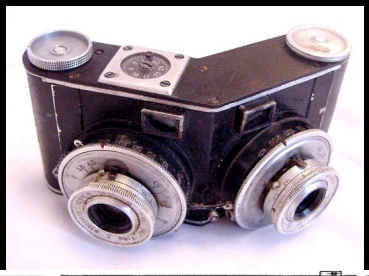
Wavefront coding

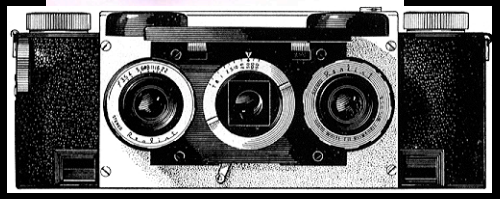
An explosion of cameras

- Best way to capture image and depth: Stereo? Plenoptic camera? Coded aperture? or...?
- What aspects of these cameras contribute to their performance?
- Can we design new cameras with improved reconstruction performance?

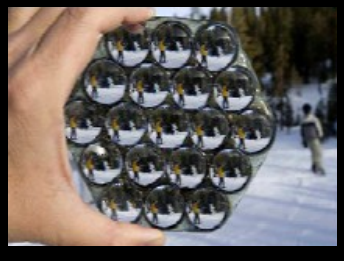
Conventional single-lens cameras

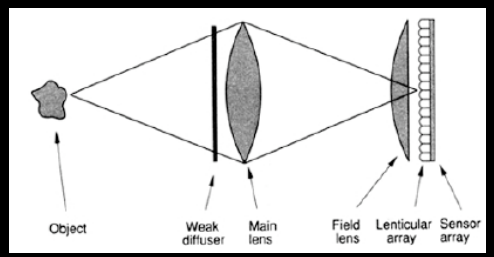
Stereo and trinocular cameras



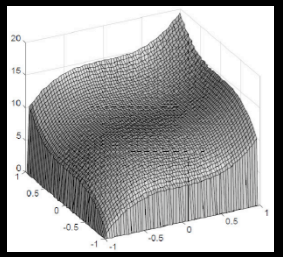


Plenoptic cameras





Coded aperture

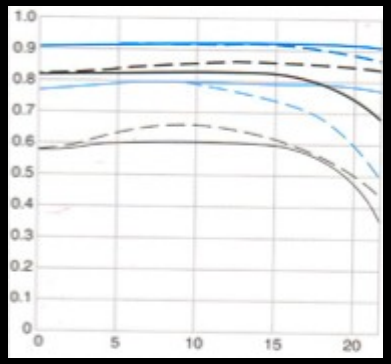


Wavefront coding

Camera evaluation, old and new

Traditional optics evaluation:

- **2D image sharpness (eg, Modulation Transfer Function)**

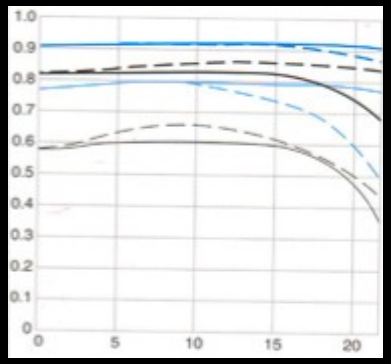


contrast vs. spatial frequency

Camera evaluation, old and new

Traditional optics evaluation:

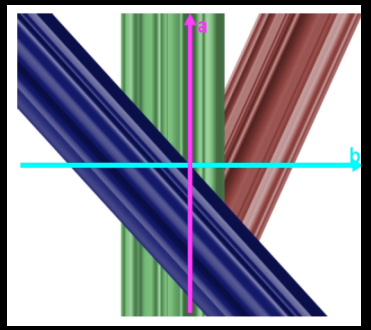
2D image sharpness (eg, Modulation Transfer Function)



contrast vs. spatial frequency

Our modern camera evaluation:

How well does the recorded data allow us to <u>estimate</u> the visual world - the lightfield?



lightfield reconstruction

Computational photography camera evaluation: an estimation problem

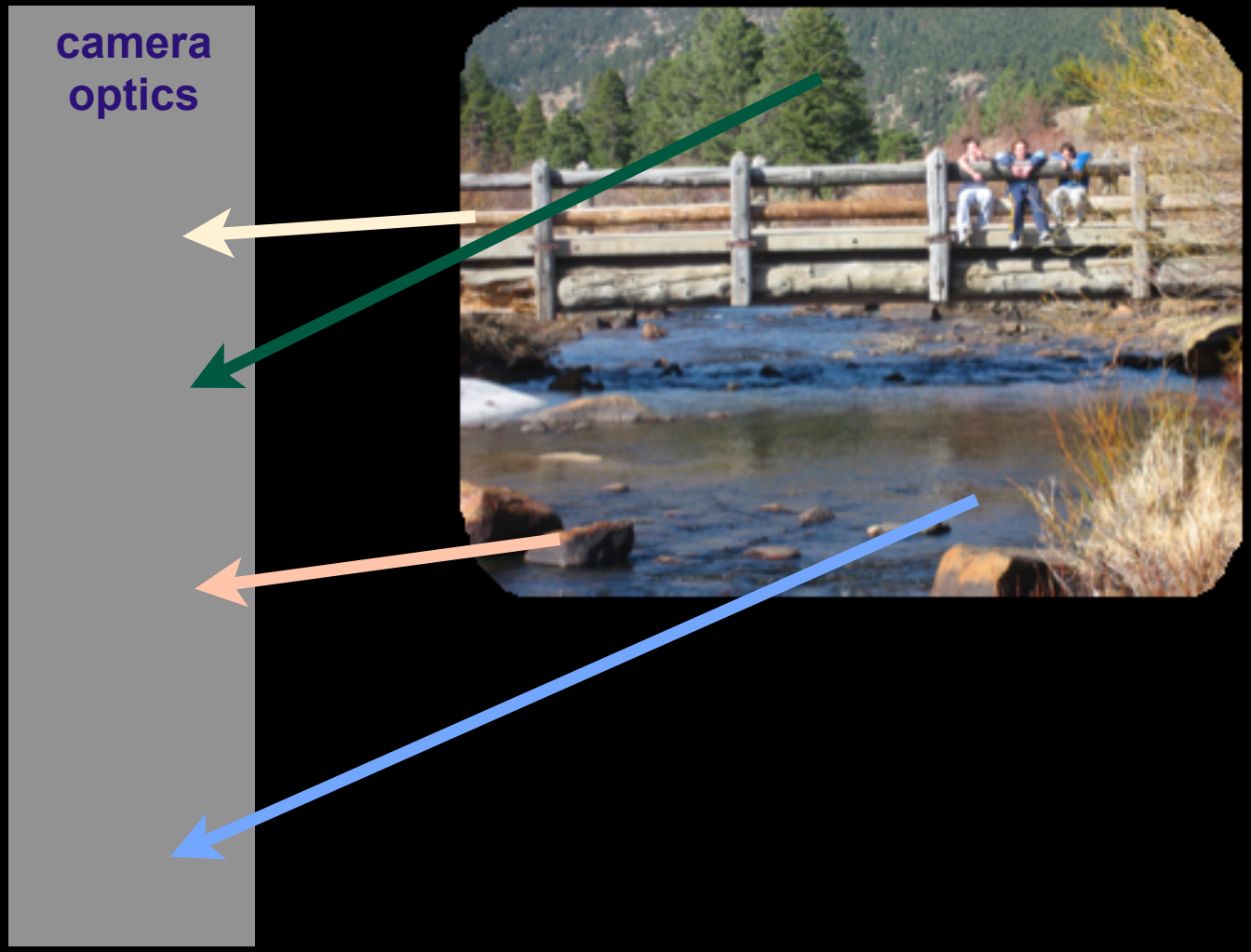
- Characteristics of the signal to be estimated.
- Projection functions of various cameras.
- Bayesian lightfield analysis
  - Reconstructing the lightfield from camera data.
  - Comparing performance tradeoffs of different cameras.

Computational photography camera evaluation: an estimation problem

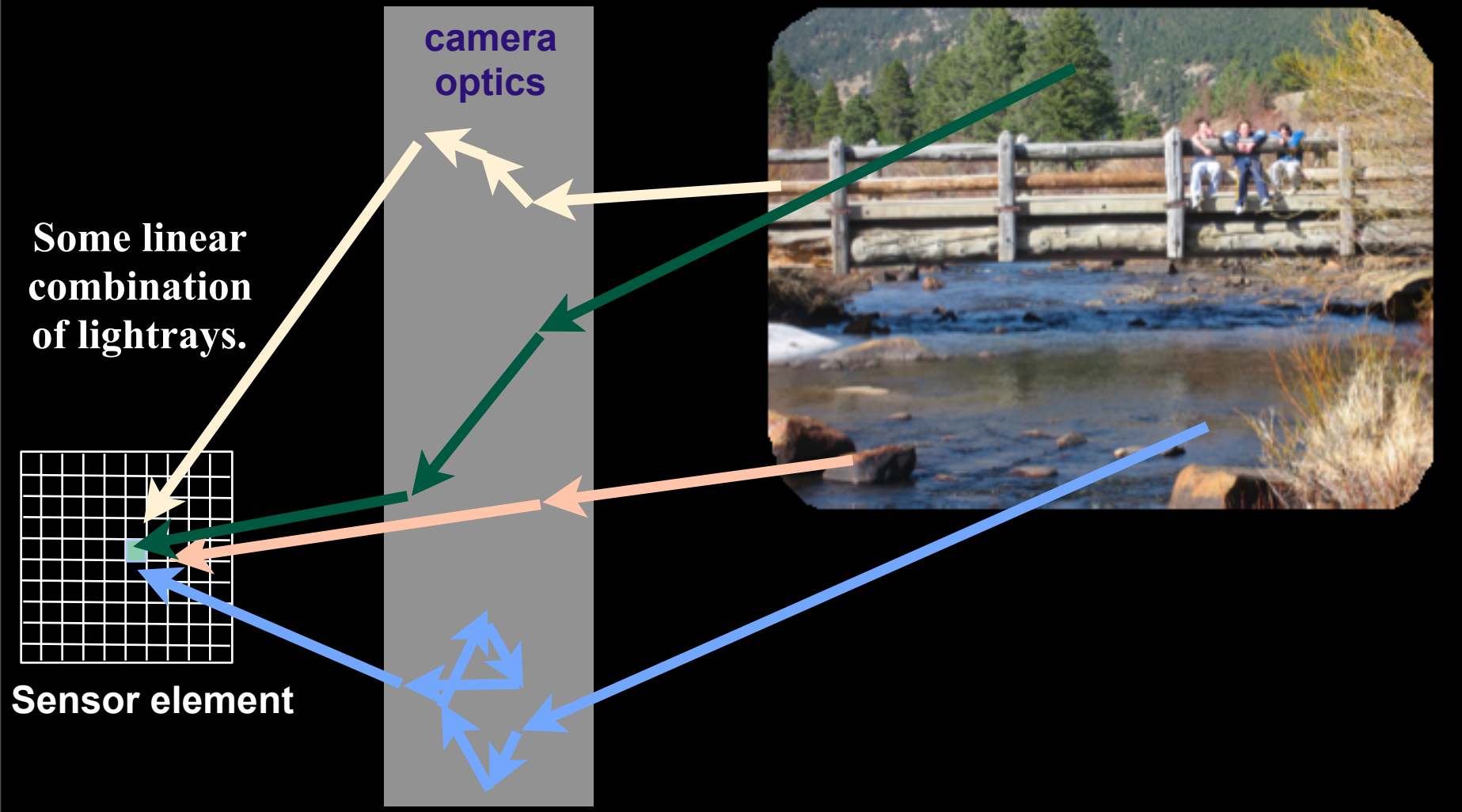
- **Characteristics of the signal to be estimated.**
- **Projection functions of various cameras.**
- **Bayesian lightfield analysis**
  - **Reconstructing the lightfield from camera data.**
  - **Comparing performance tradeoffs of different cameras.**

so let's talk about lightfields and cameras

What does a camera sensor element record?

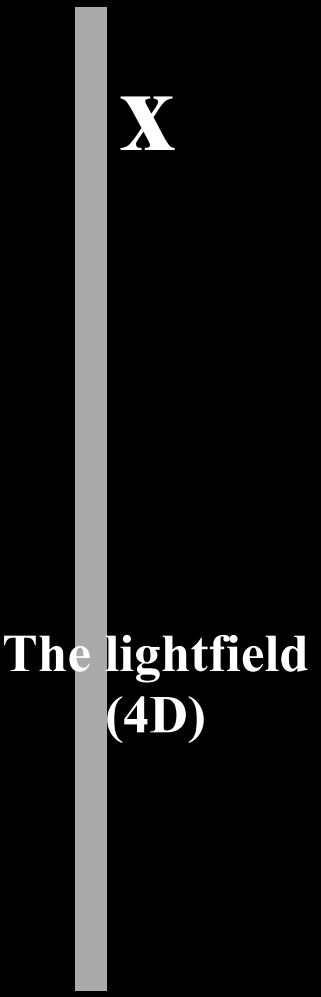


What does a camera sensor element record?

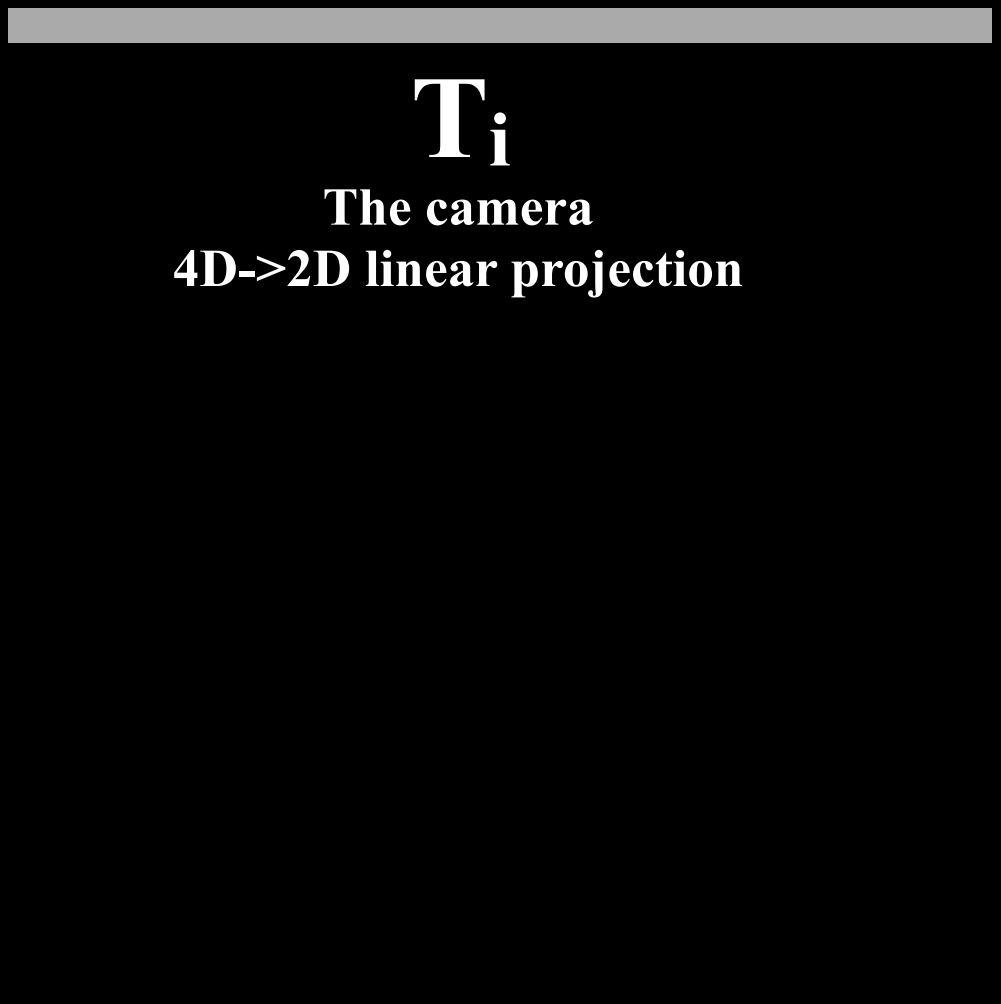


Sensor element data

Sensor element data



Sensor element data



Sensor element data

$$y_i = T_i \, x$$

datum

The camera
4D->2D linear projection

The lightfield
(4D)

Sensor element data

$$y_i = T_i \, x + n_i$$

y_i: datum

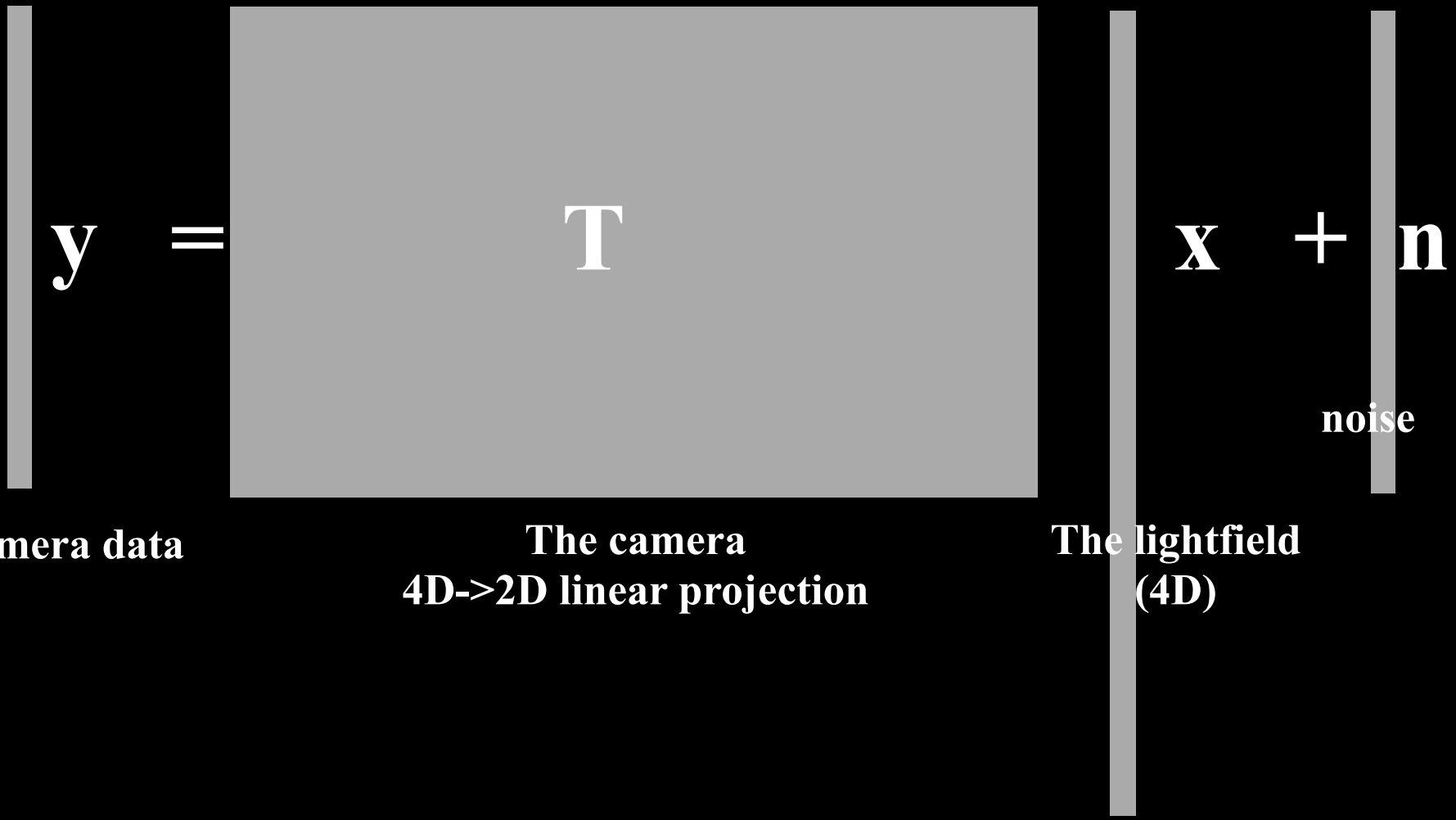
T_i: The camera
4D->2D linear projection

x: The lightfield
(4D)

n_i: noise

What is a camera?

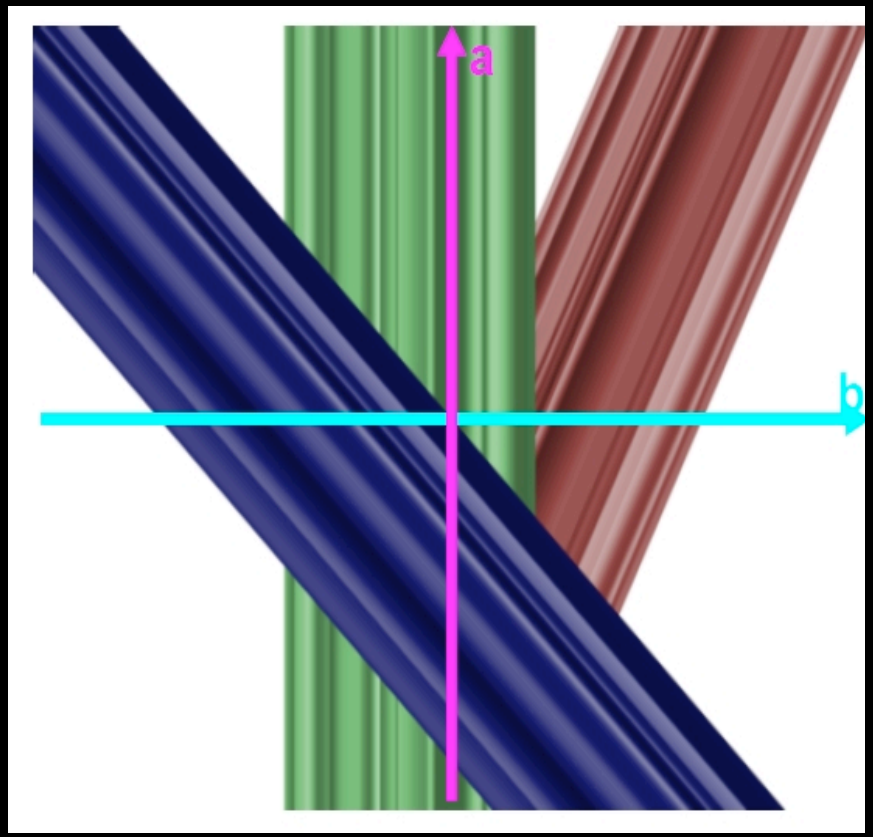
Camera: all-positive linear projection of a 4D lightfield



A more revealing parameterization of the lightfield

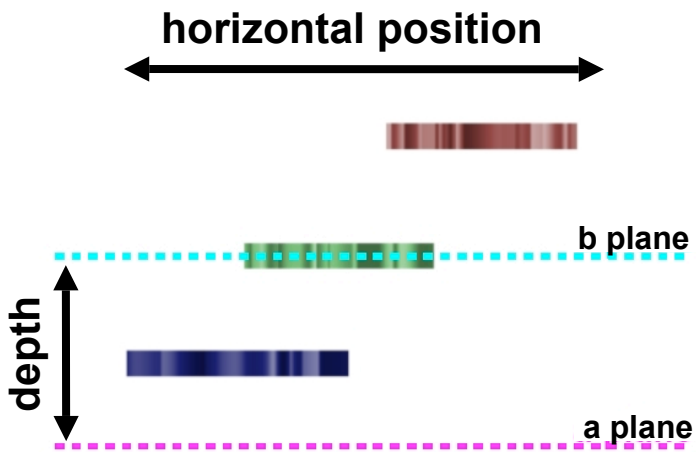
Light field: parameterization of the 4D space of light rays in the world

Provides a convenient way to model different lenses and cameras designs

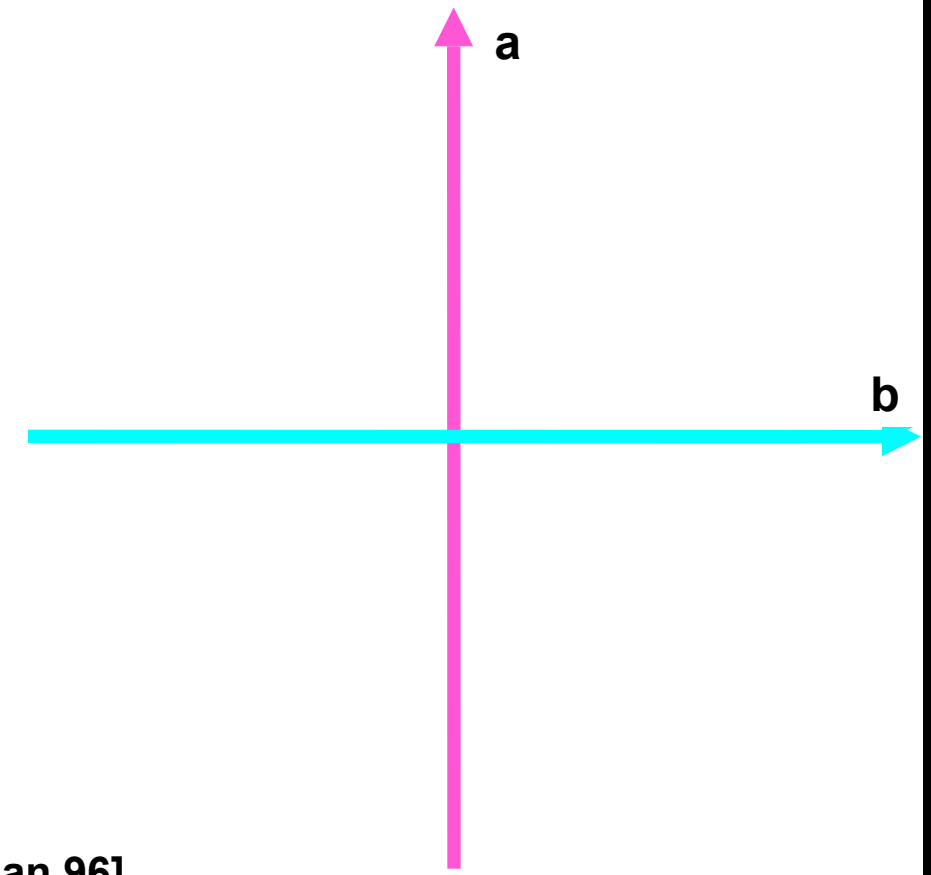


Lightfield tutorial

flatworld 1D scene



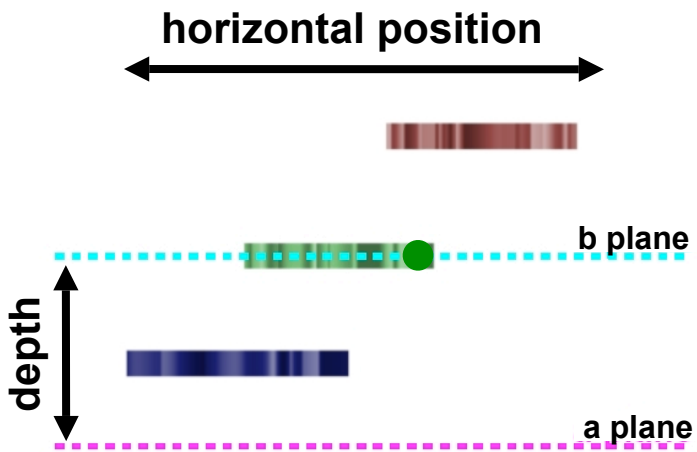
2D lightfield



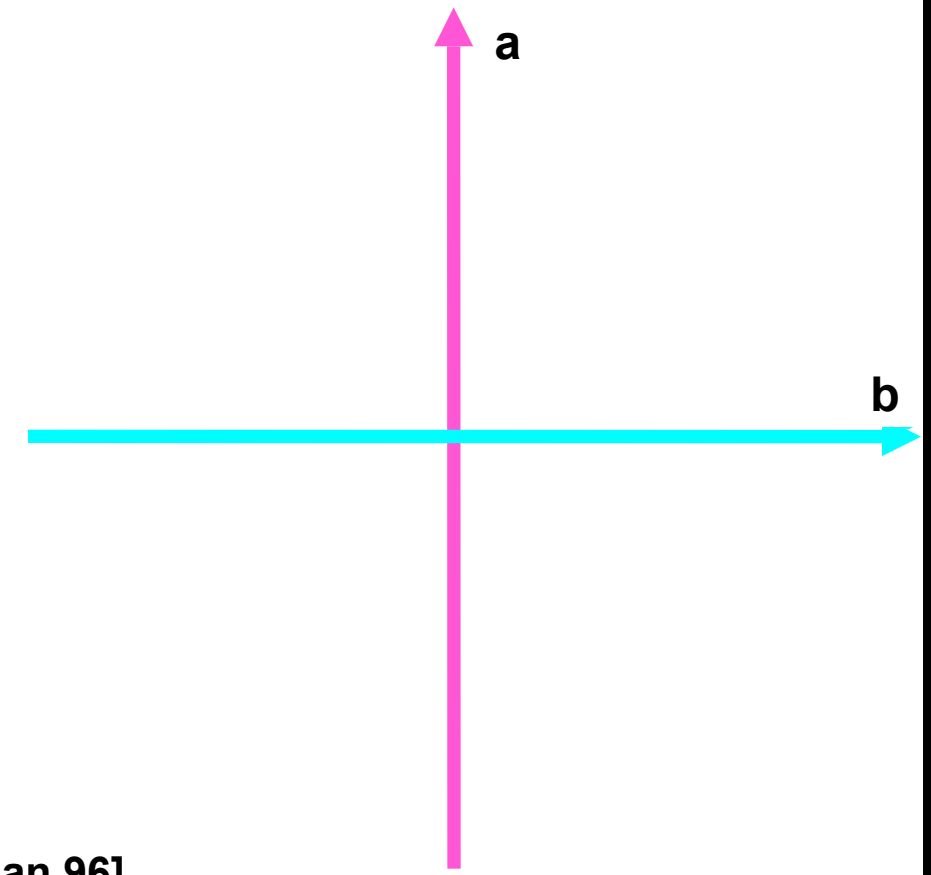
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



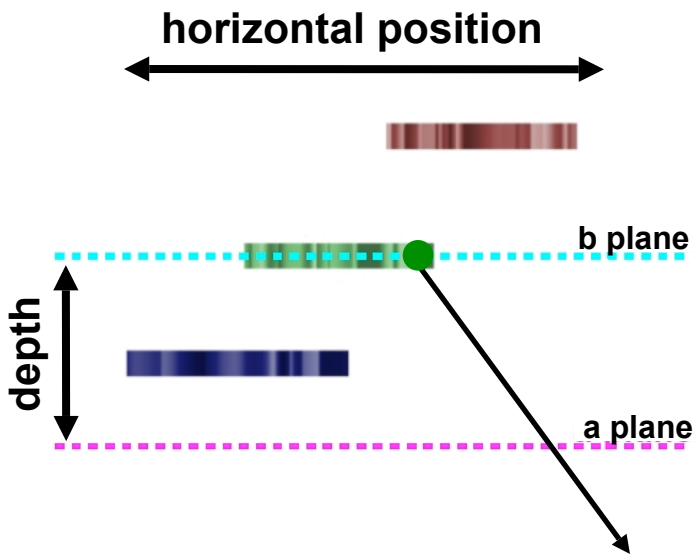
2D lightfield



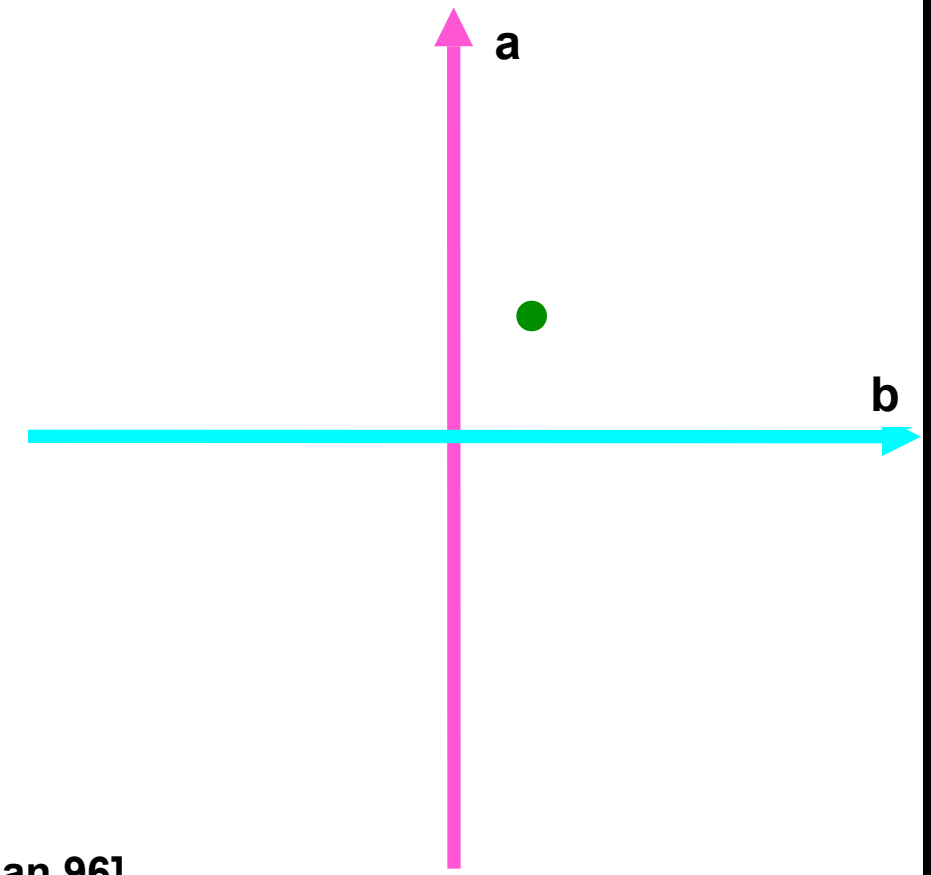
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



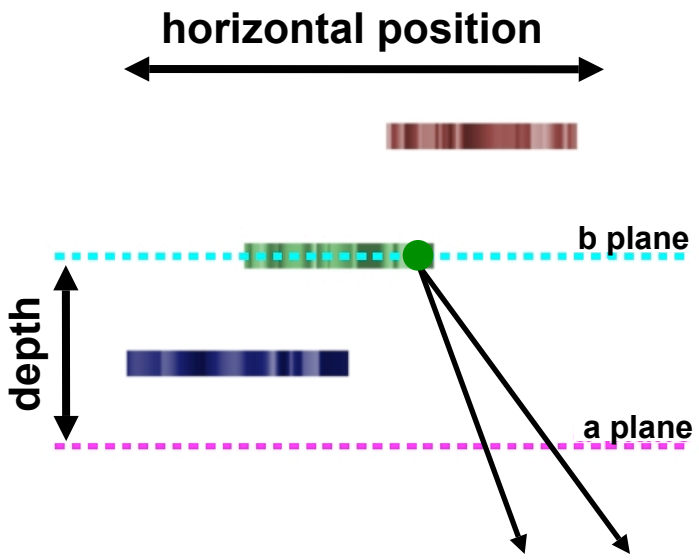
2D lightfield



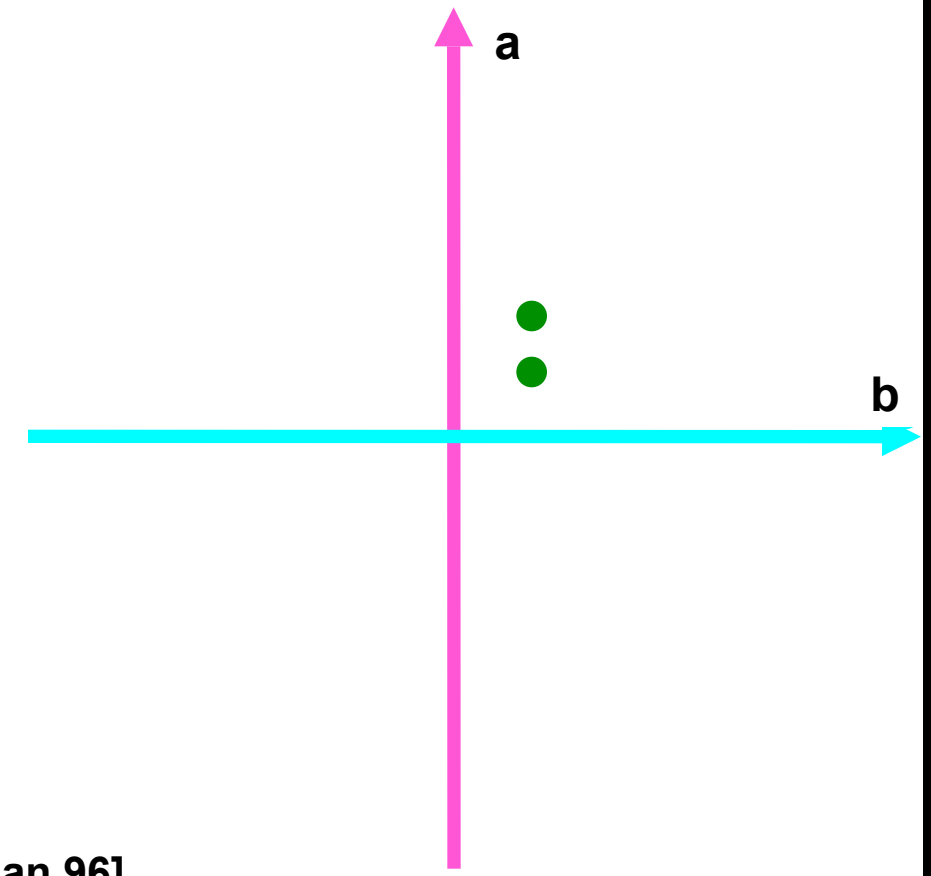
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



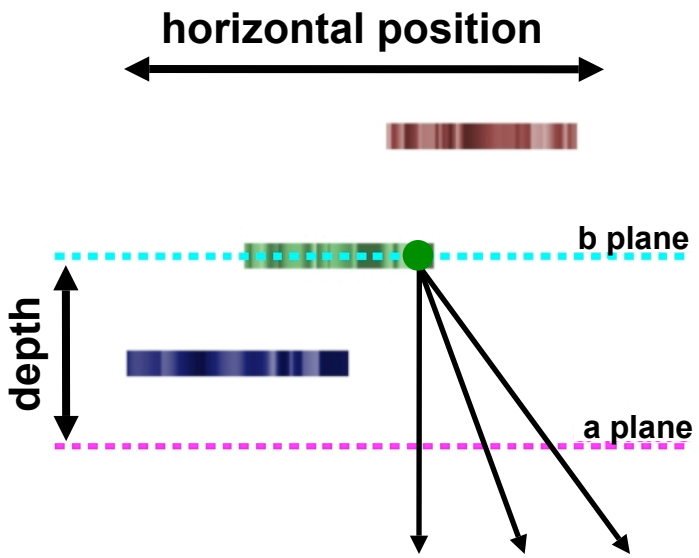
2D lightfield



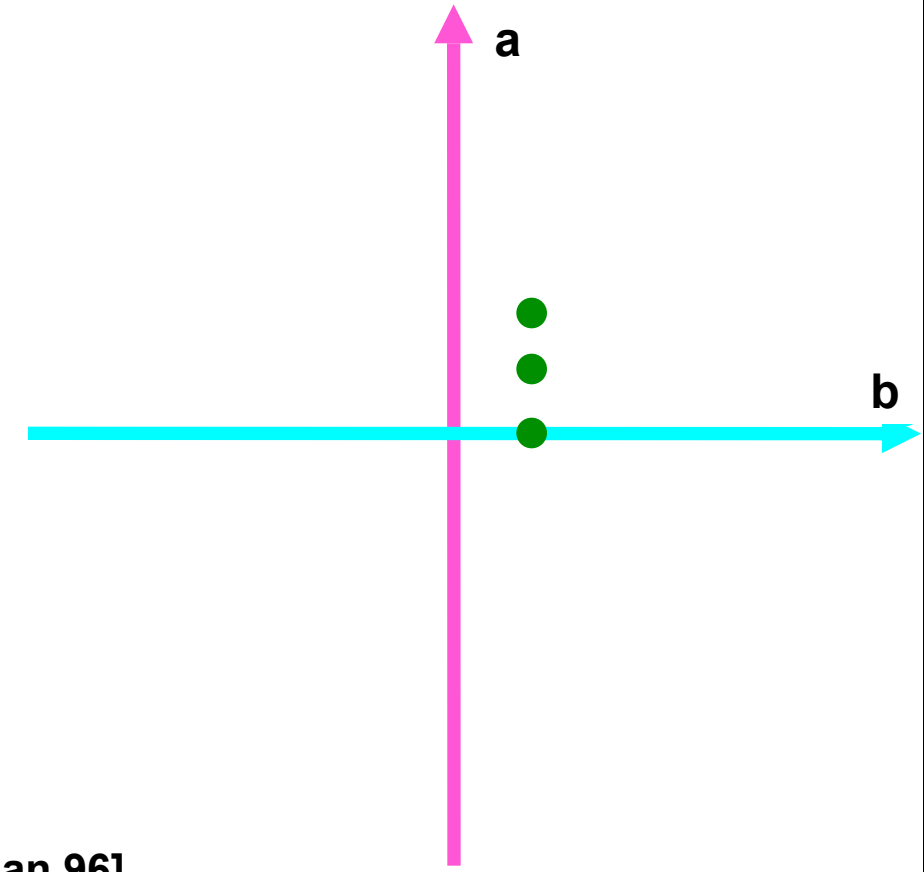
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



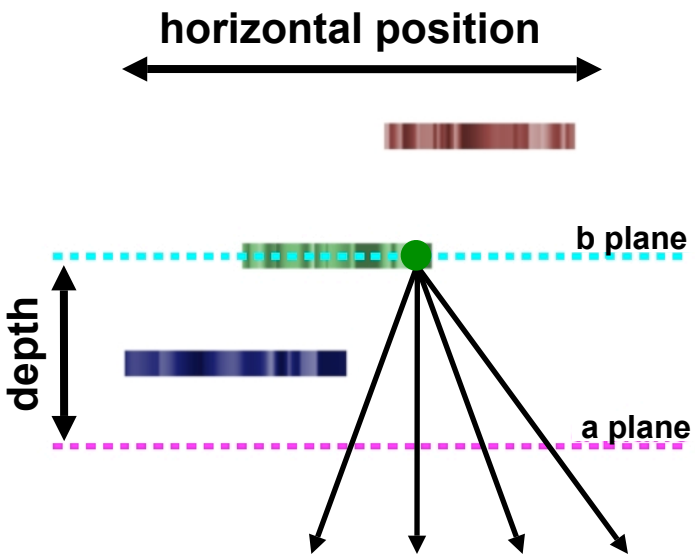
2D lightfield



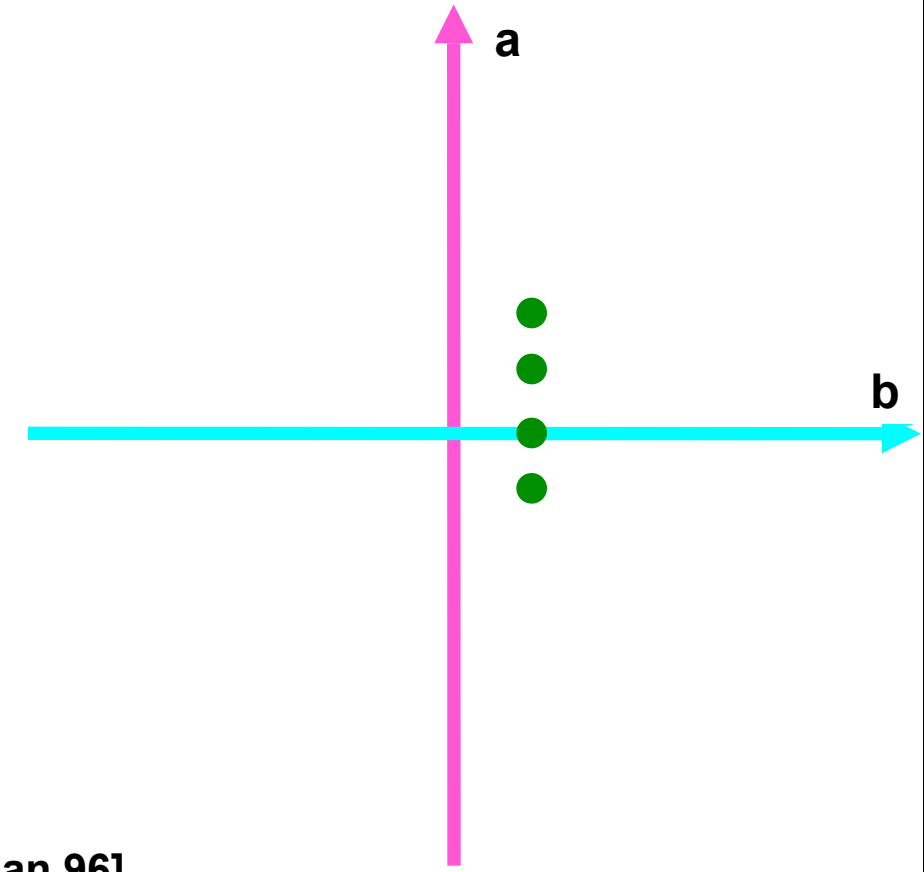
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



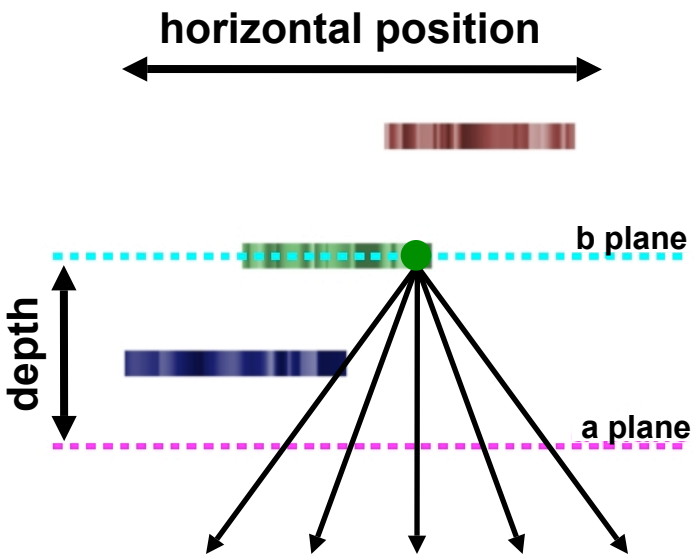
2D lightfield



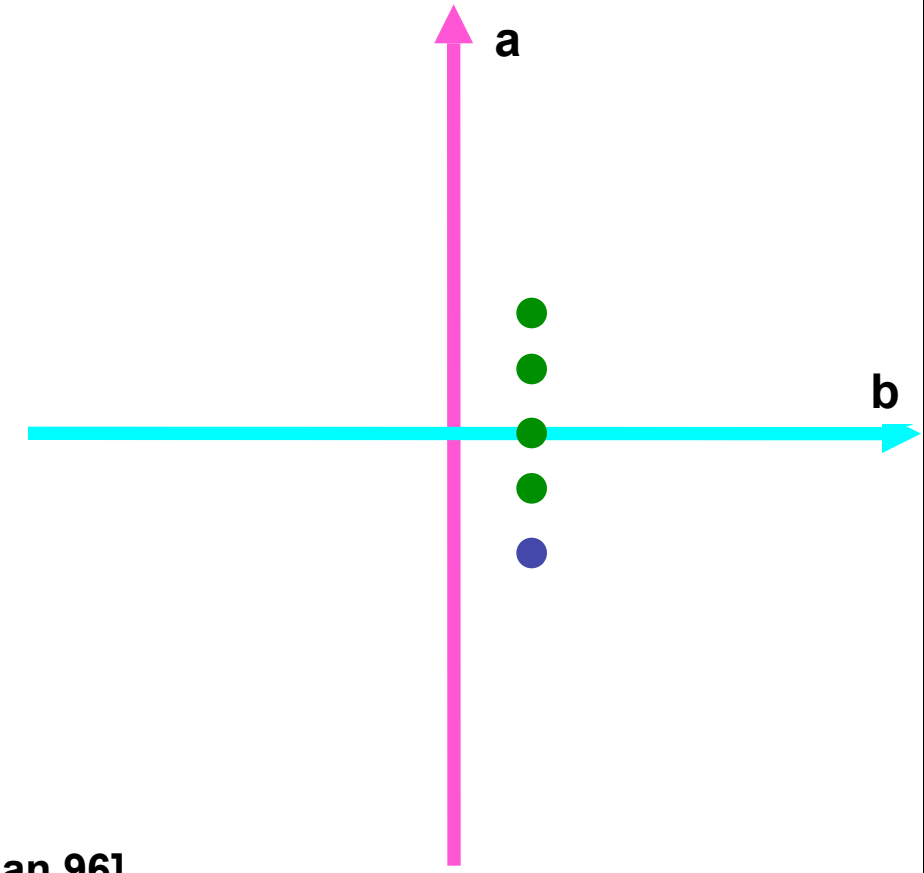
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



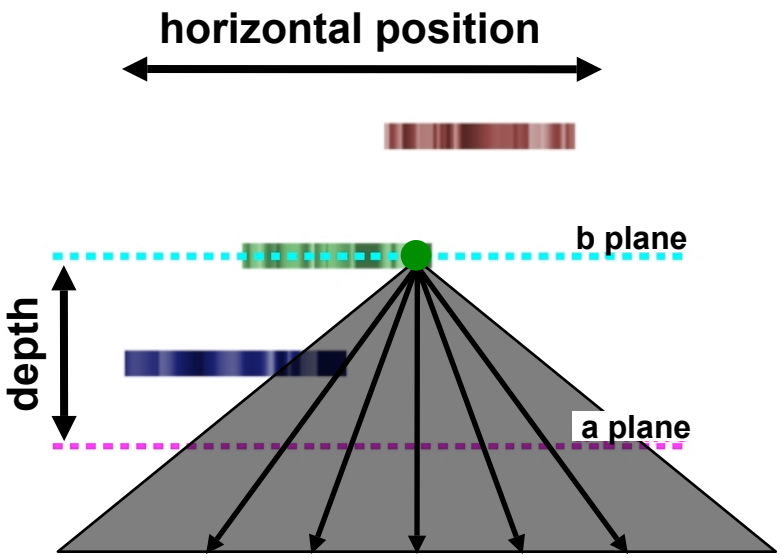
2D lightfield



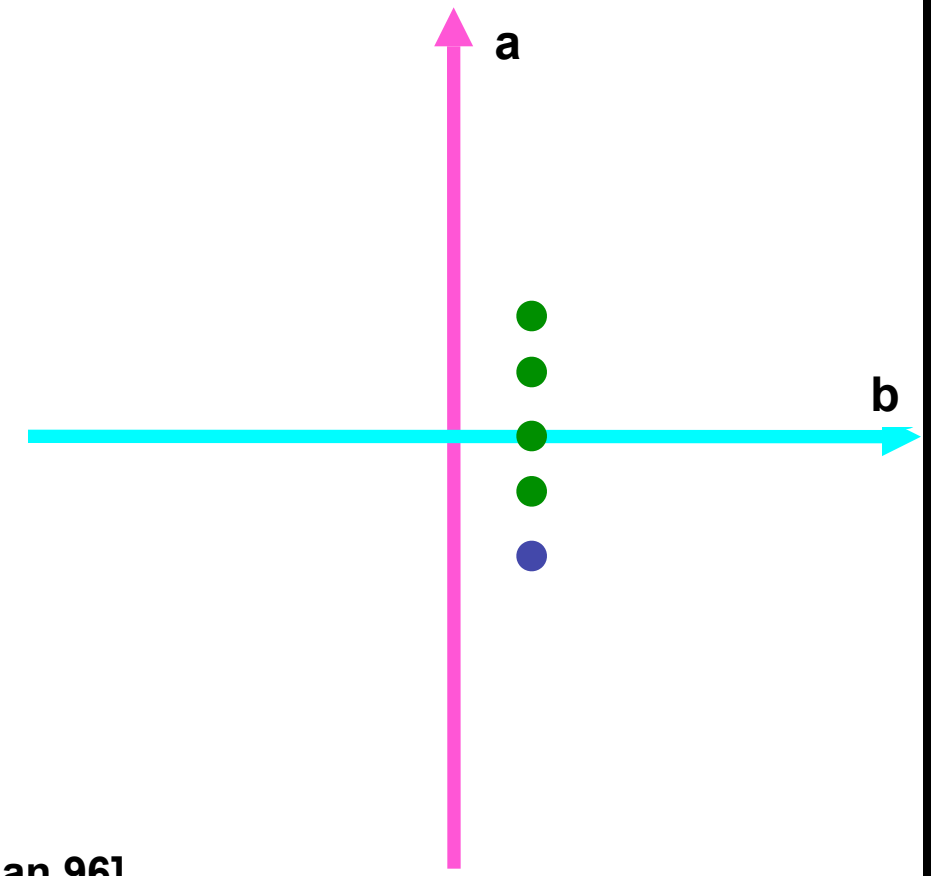
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



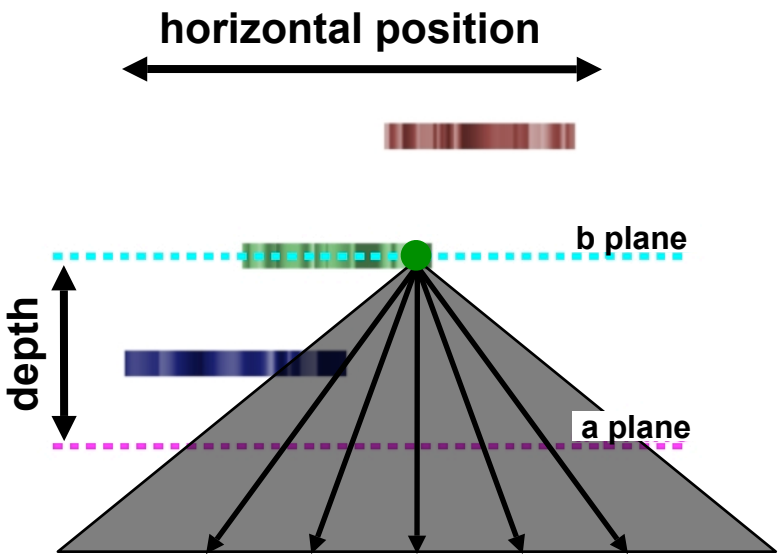
2D lightfield



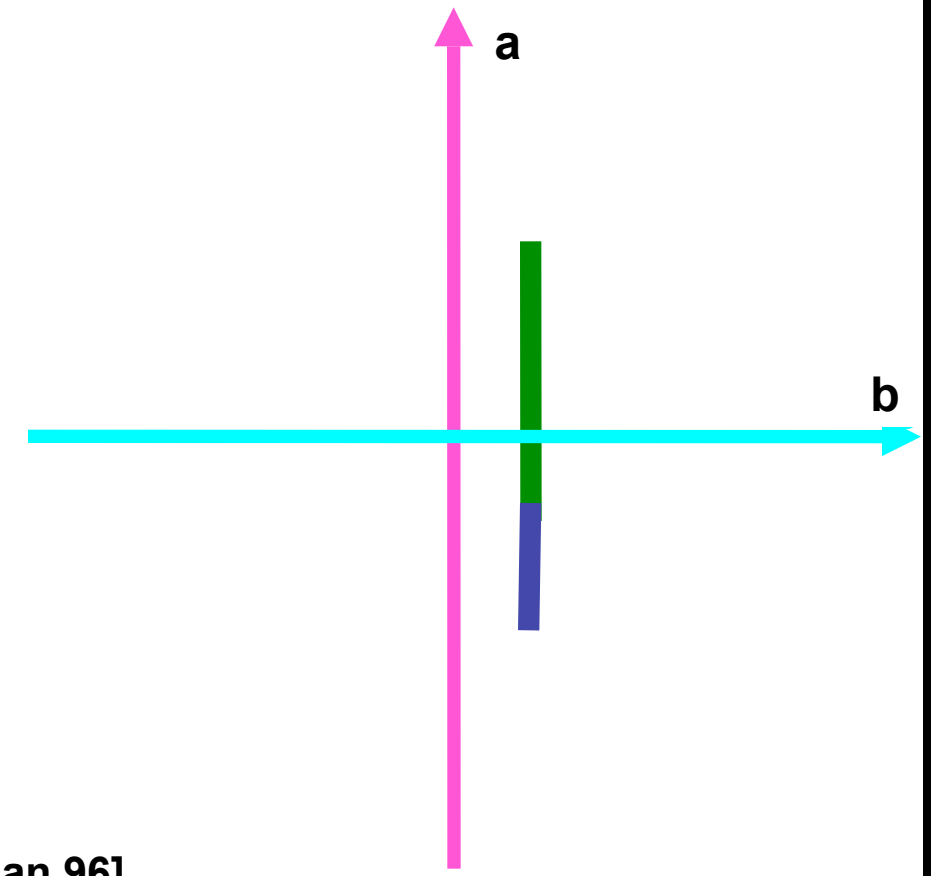
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



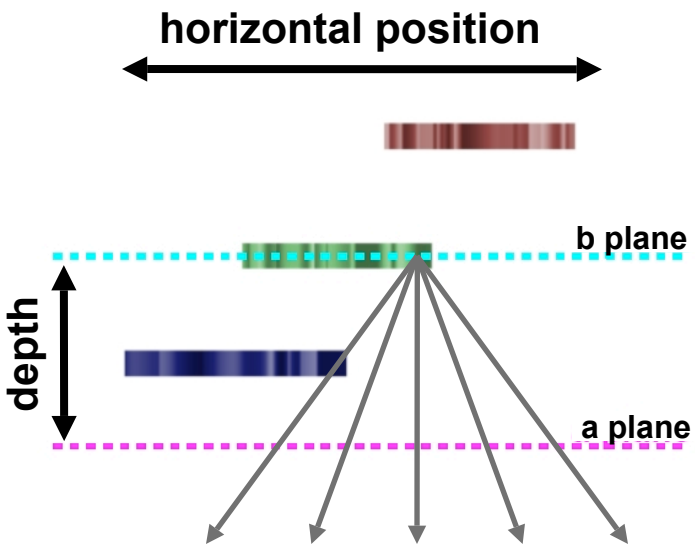
2D lightfield



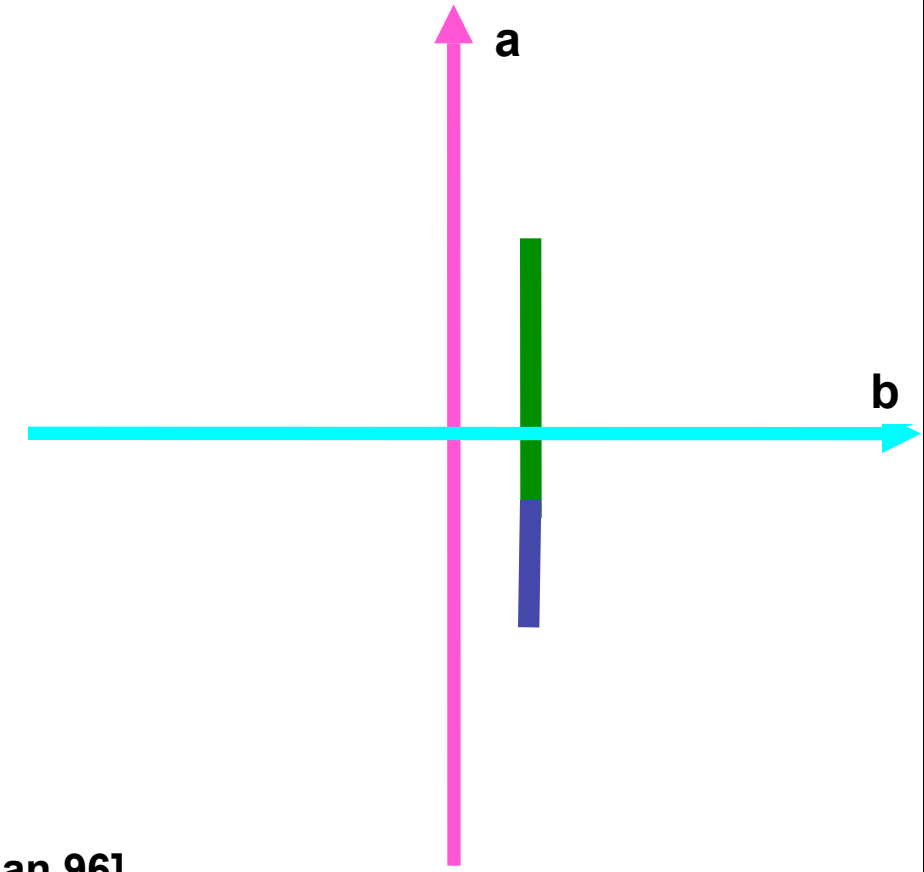
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



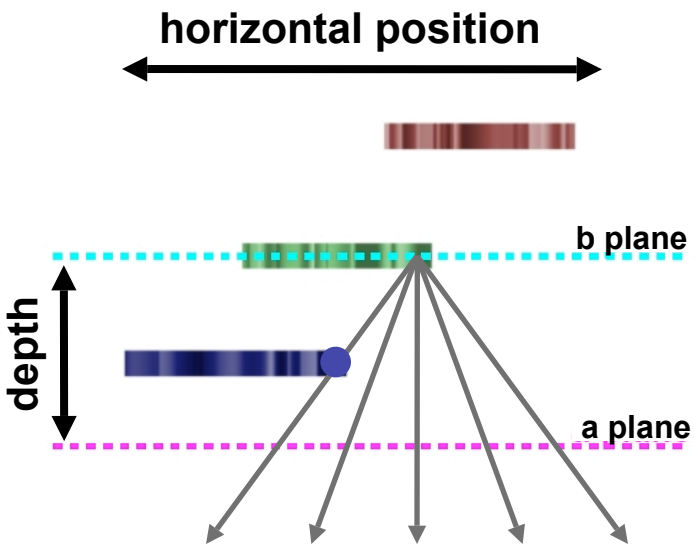
2D lightfield



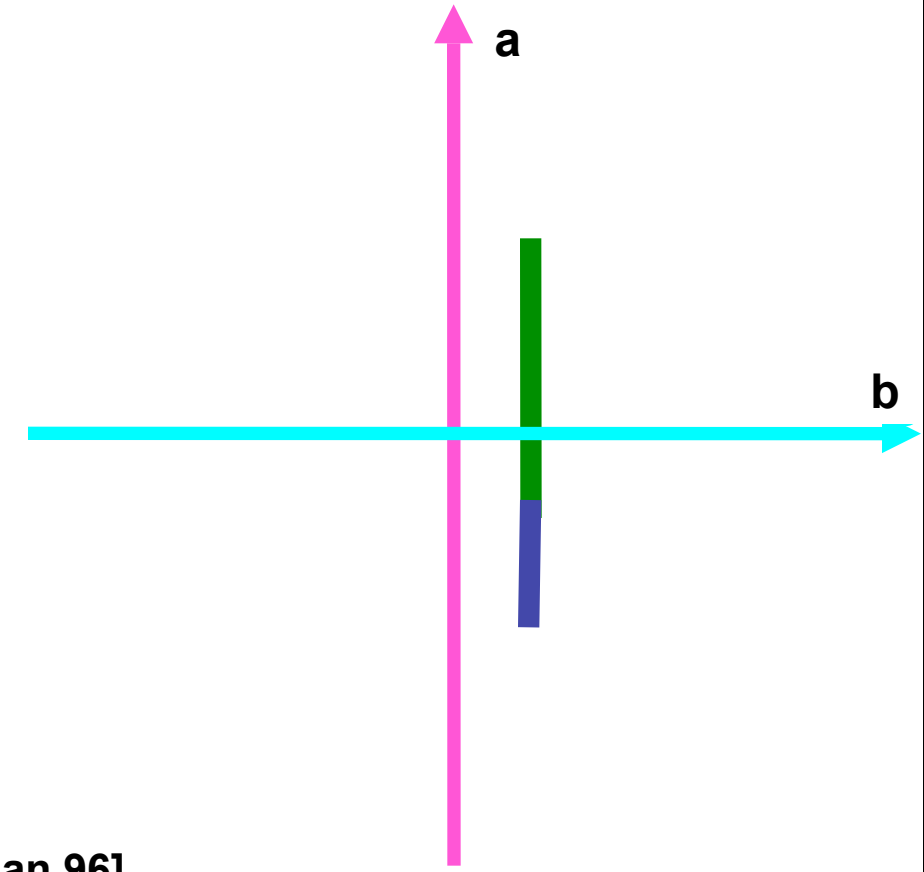
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



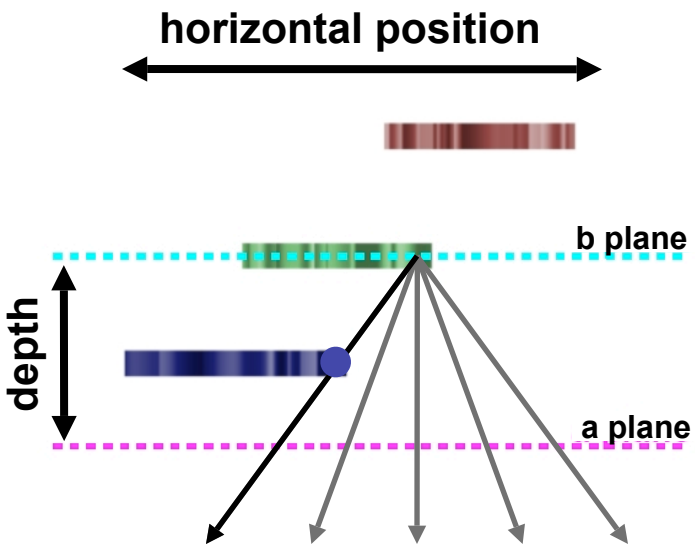
2D lightfield



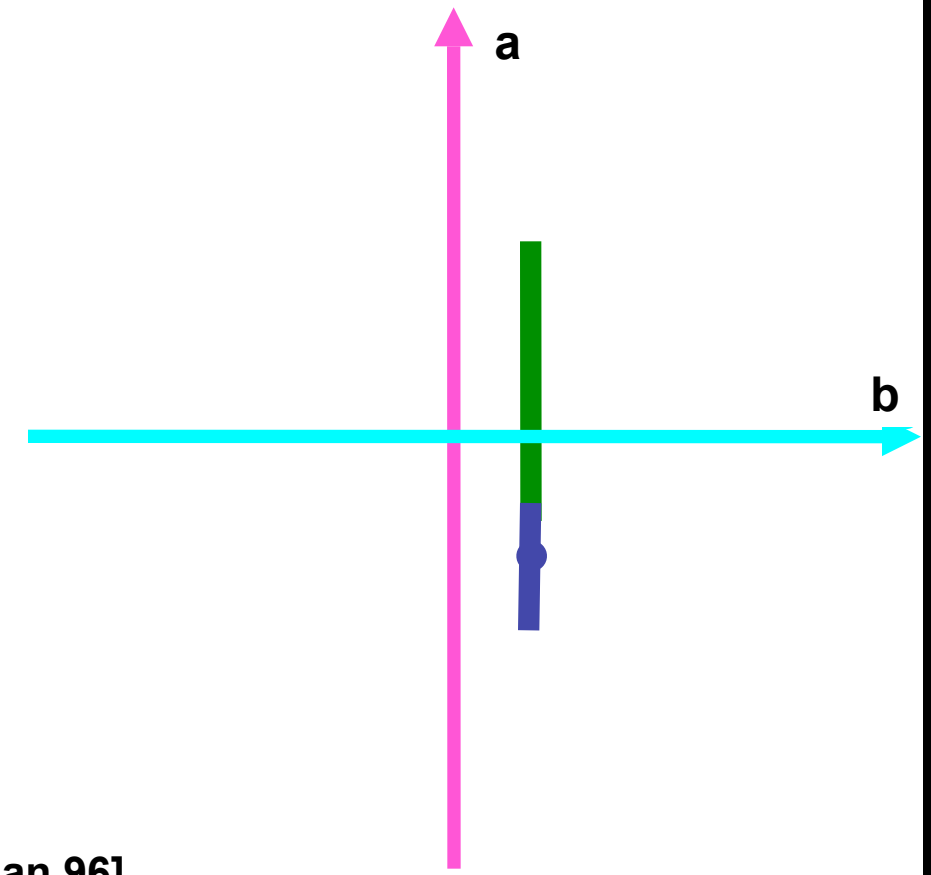
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



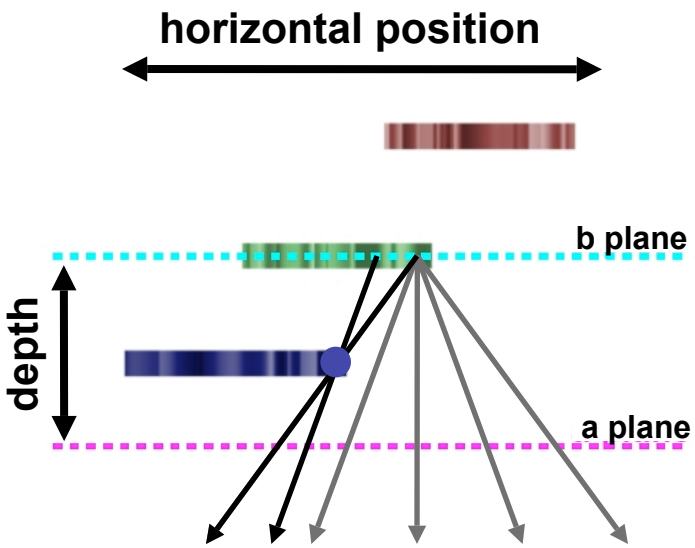
2D lightfield



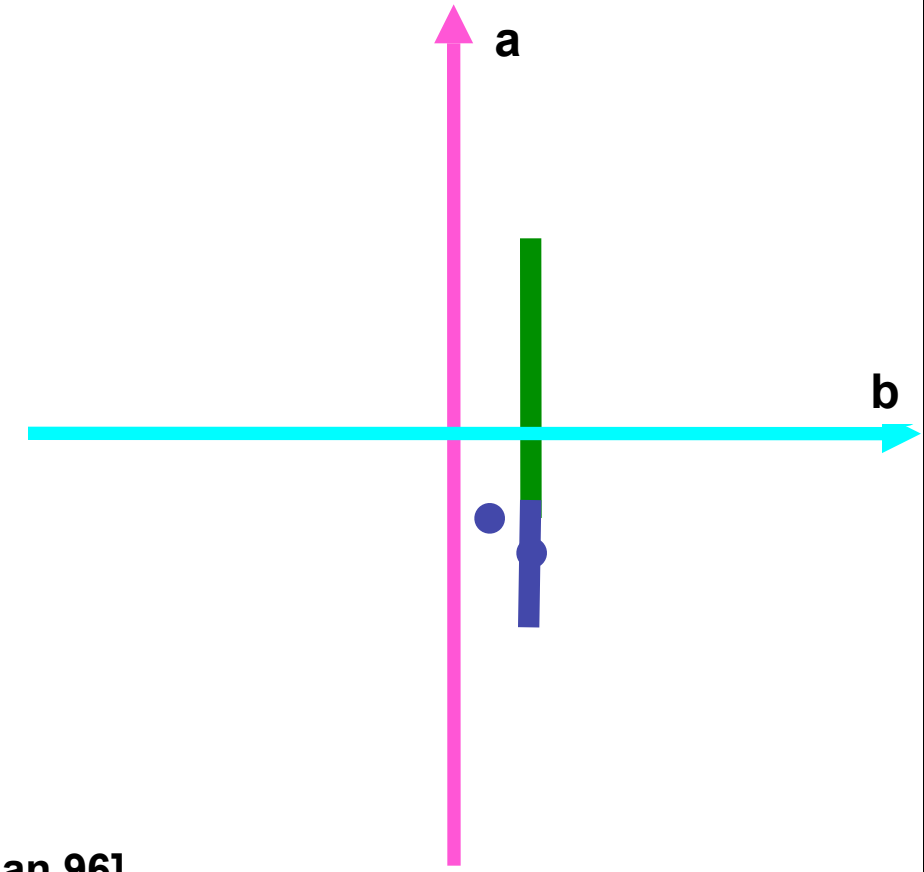
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



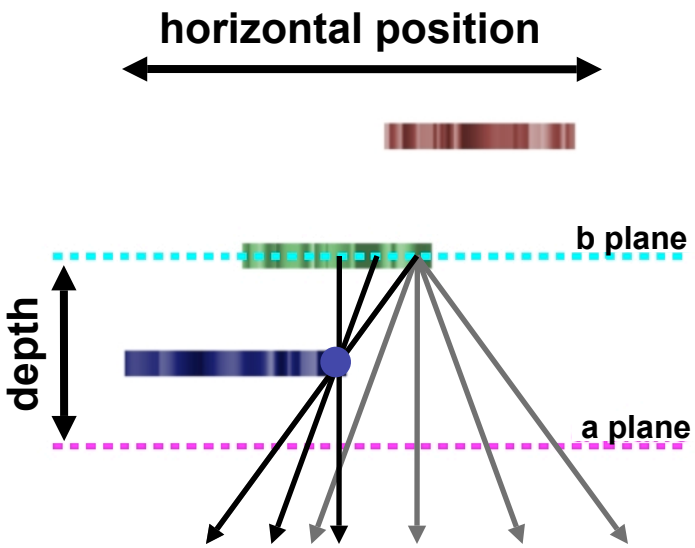
2D lightfield



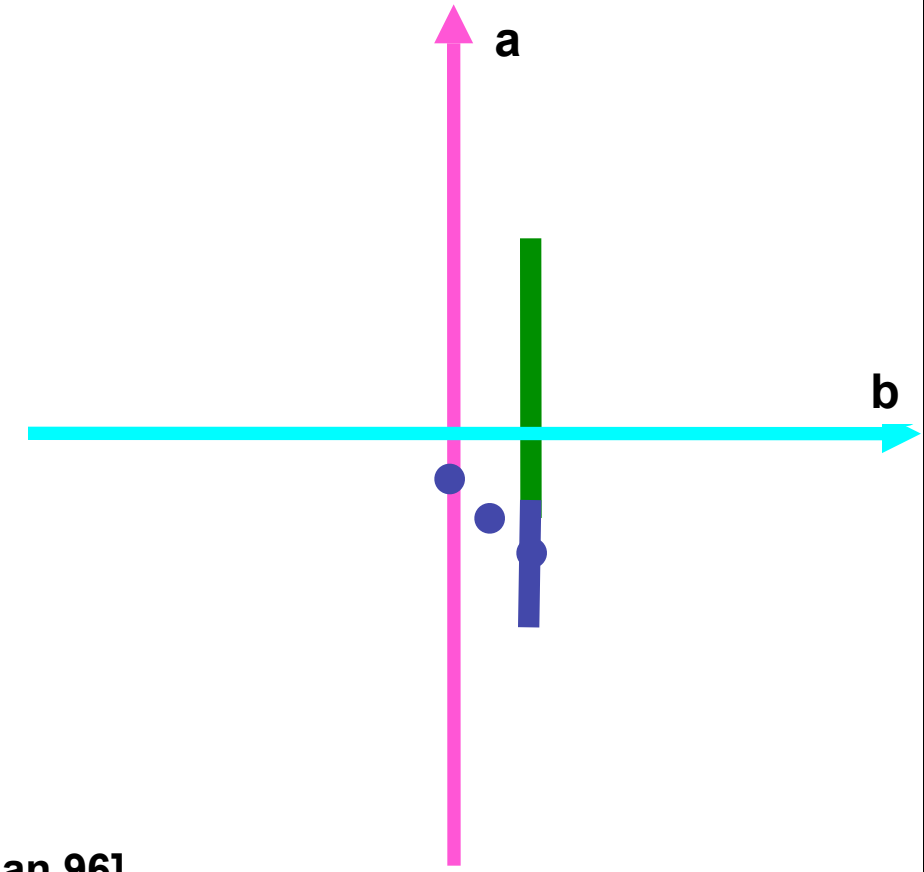
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



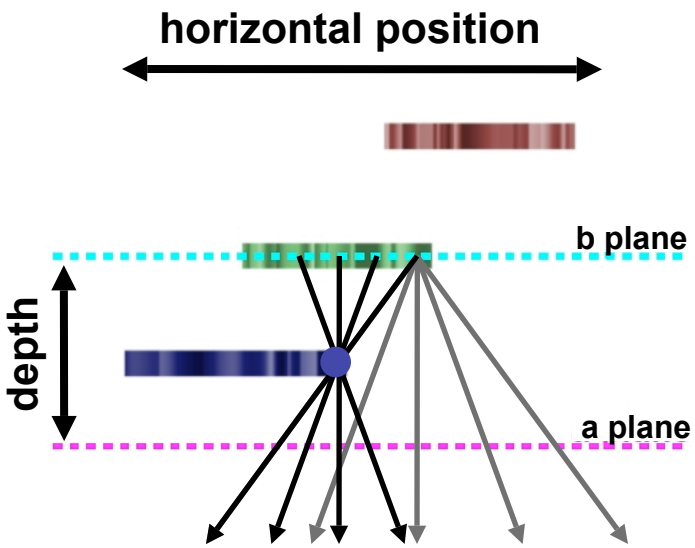
2D lightfield



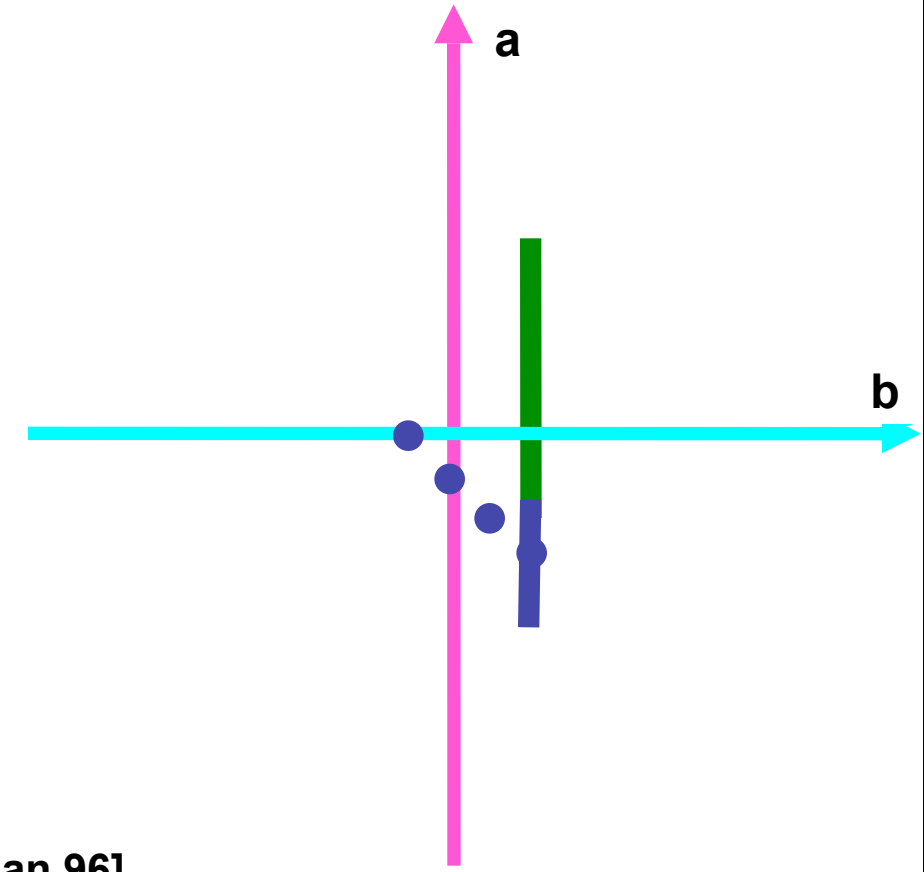
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



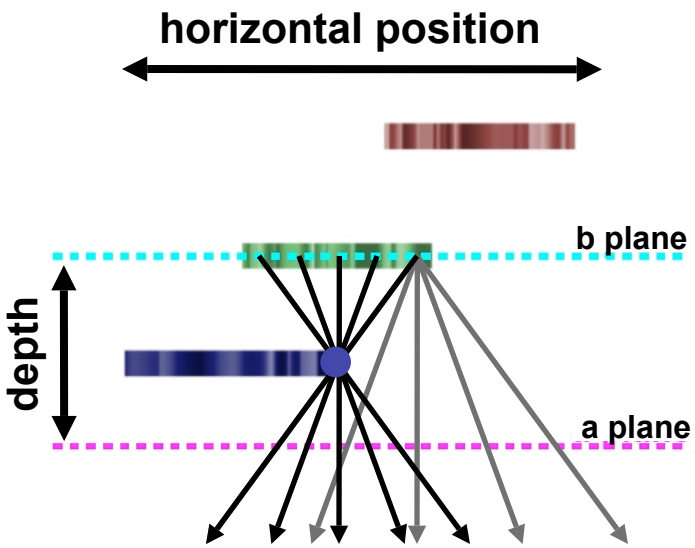
2D lightfield



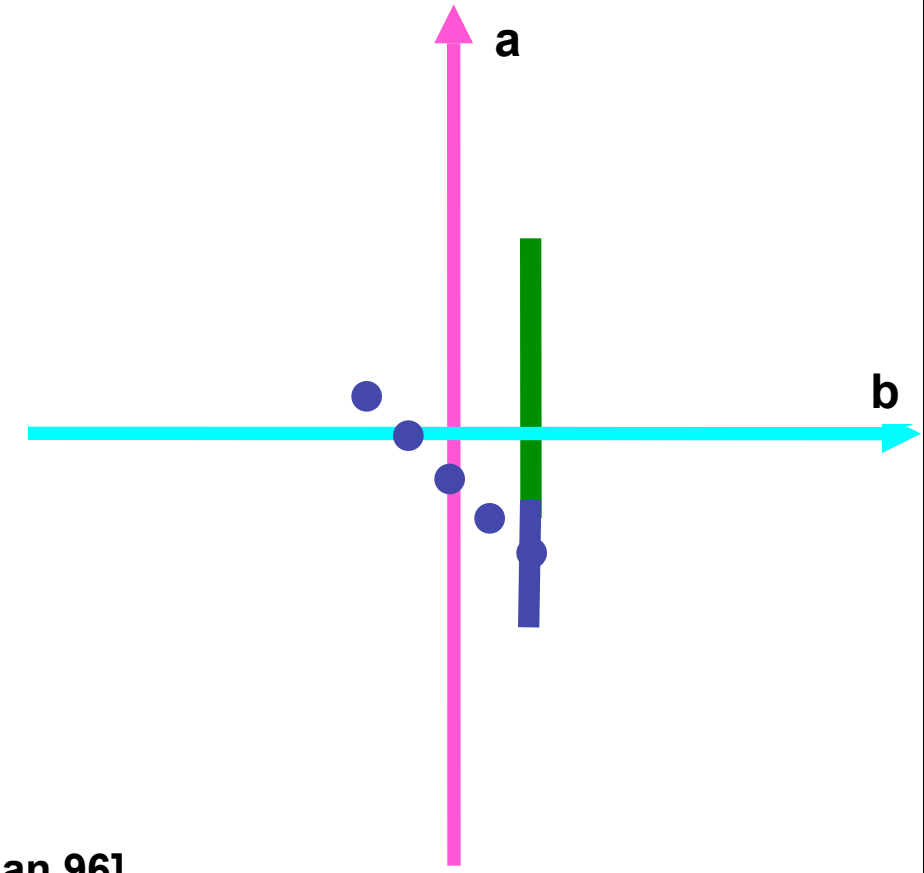
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



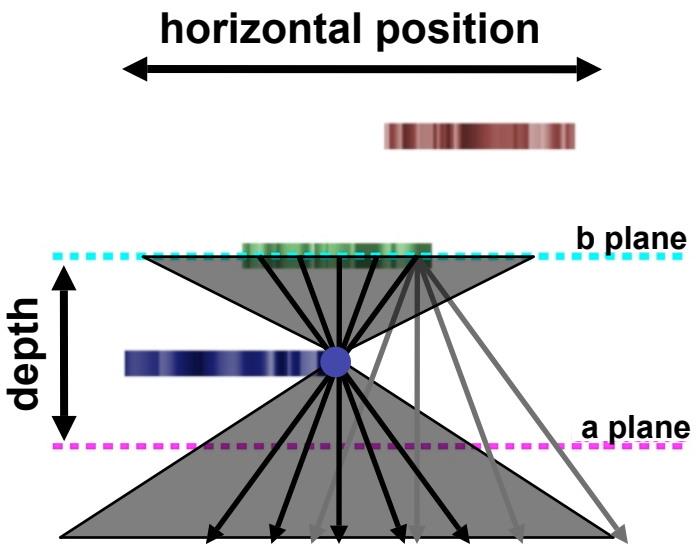
2D lightfield



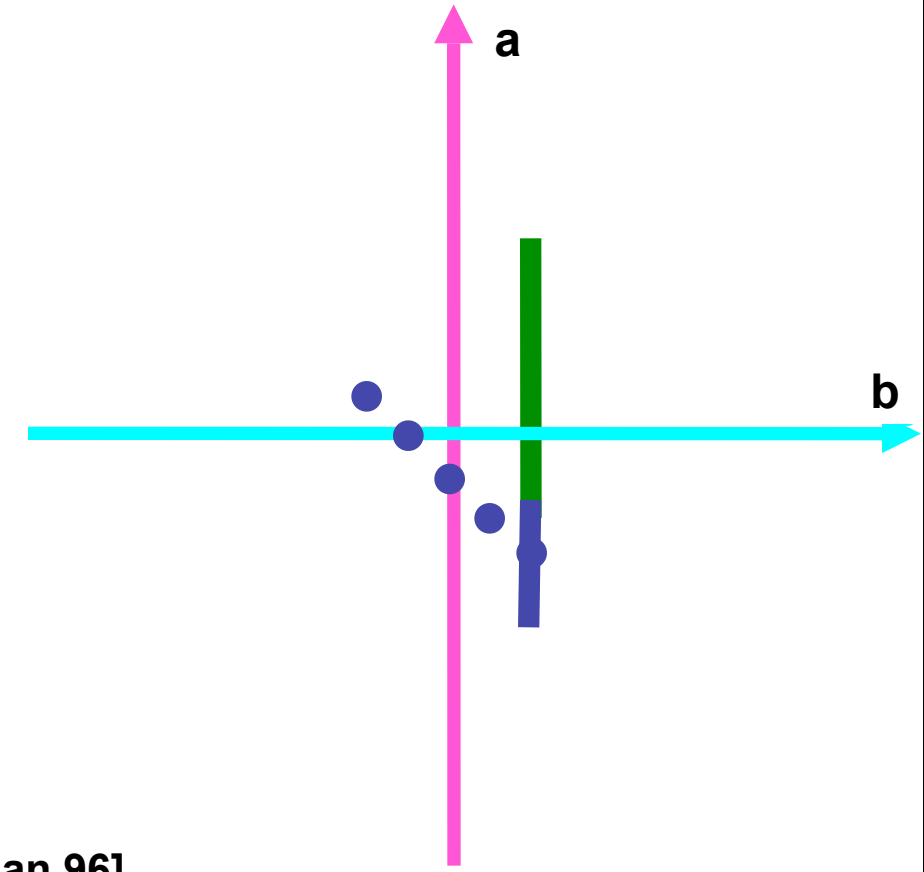
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



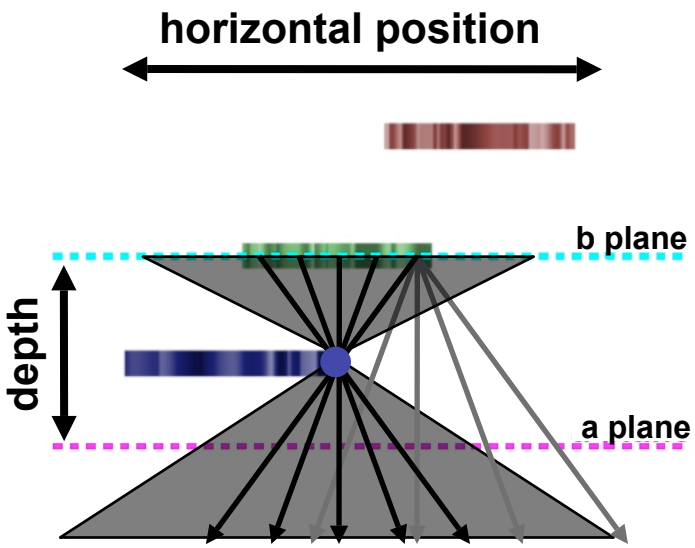
2D lightfield



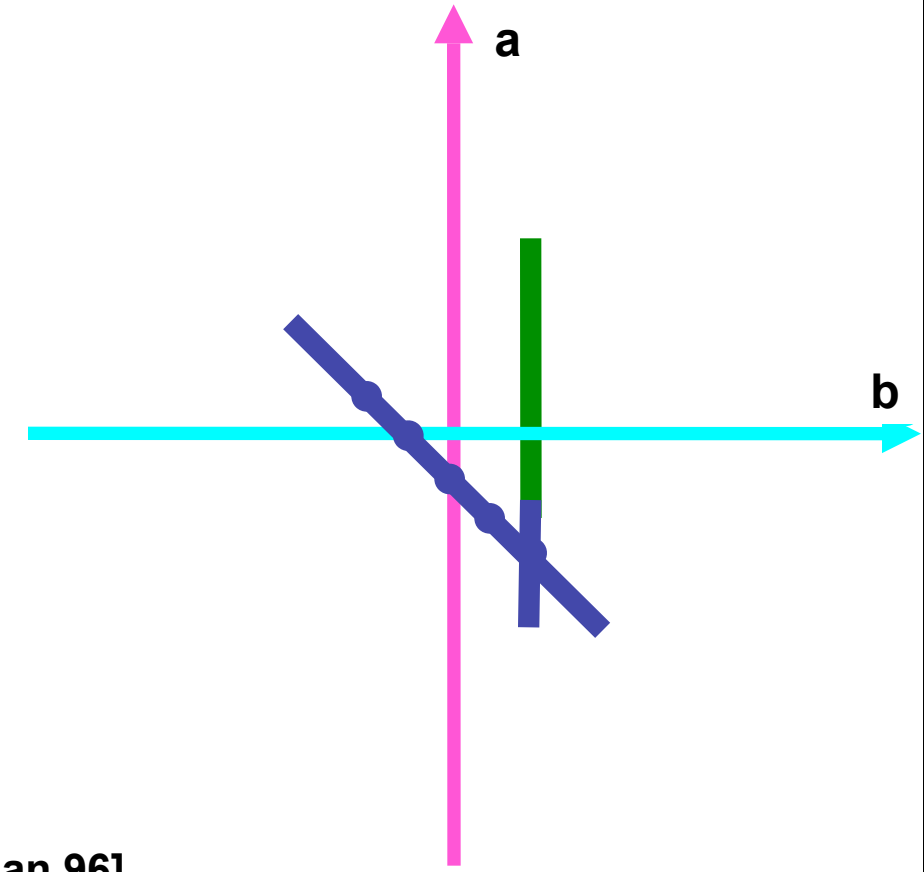
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



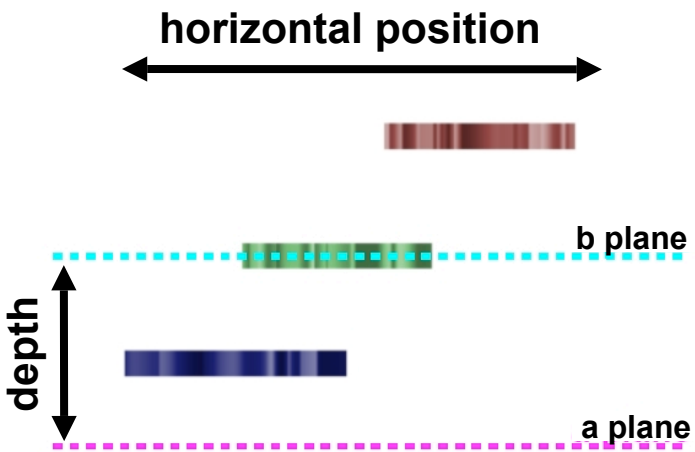
2D lightfield



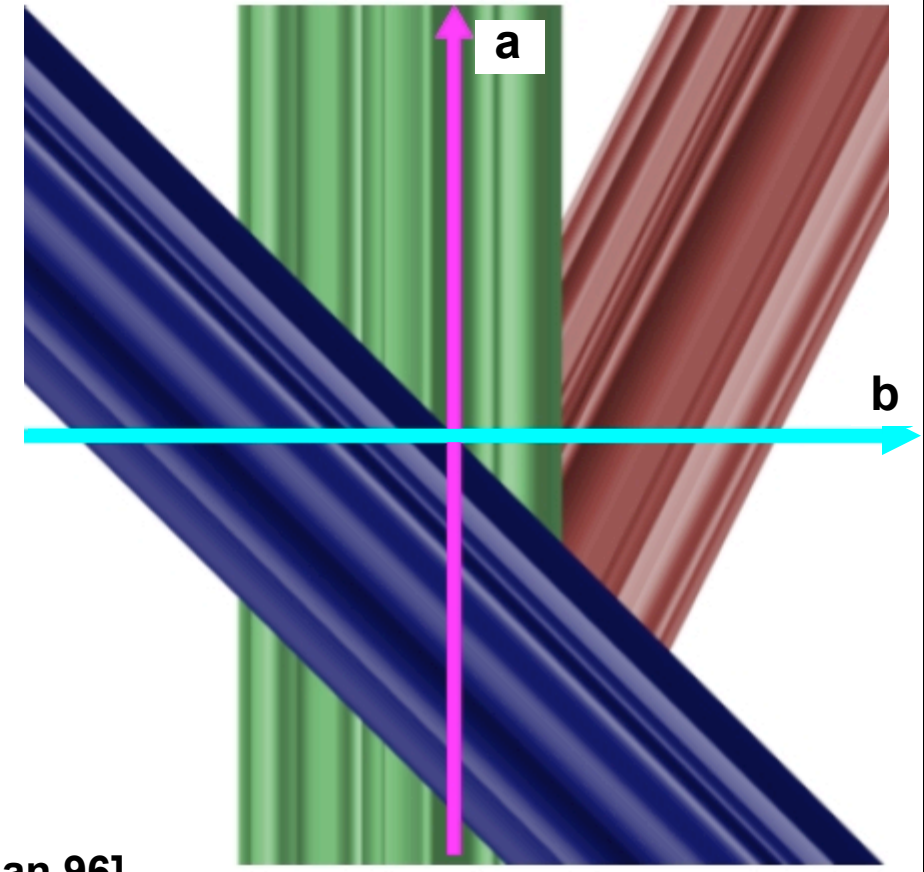
2 plane parameterization [Levoy and Hanrahan 96]

Lightfield tutorial

flatworld 1D scene



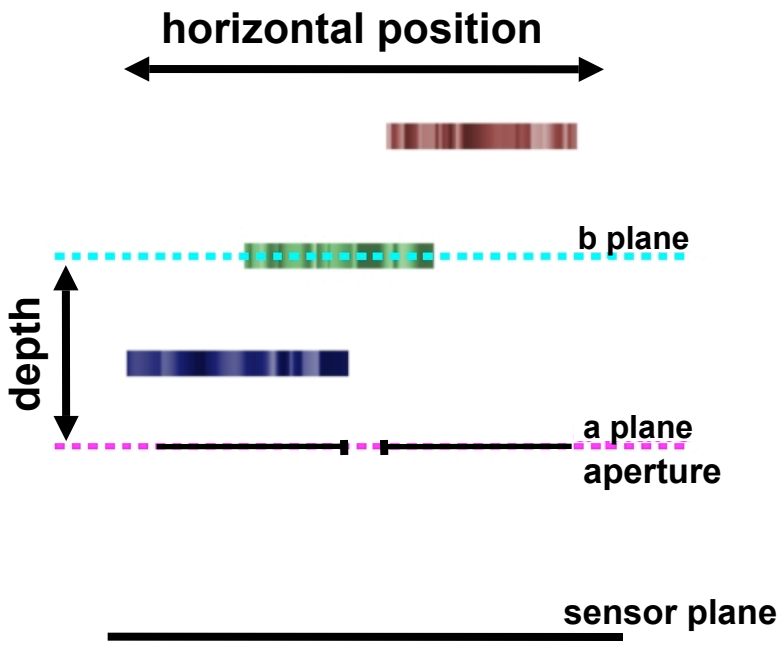
2D lightfield



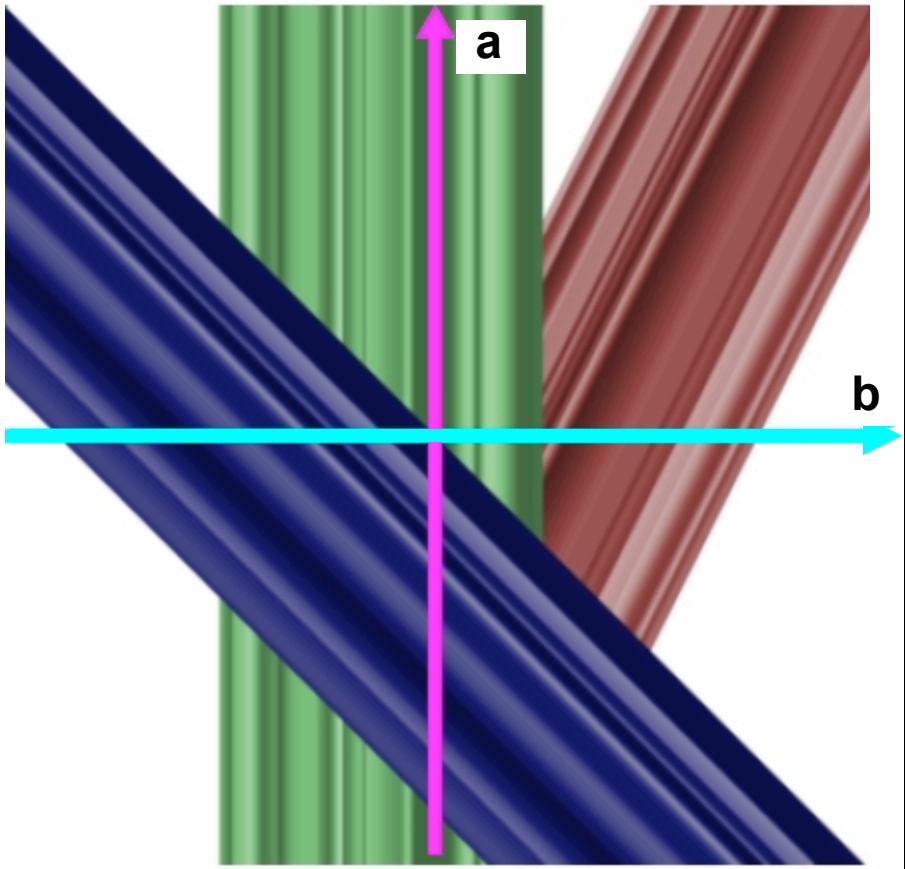
2 plane parameterization [Levoy and Hanrahan 96]

Pinhole camera

flatworld 1D scene

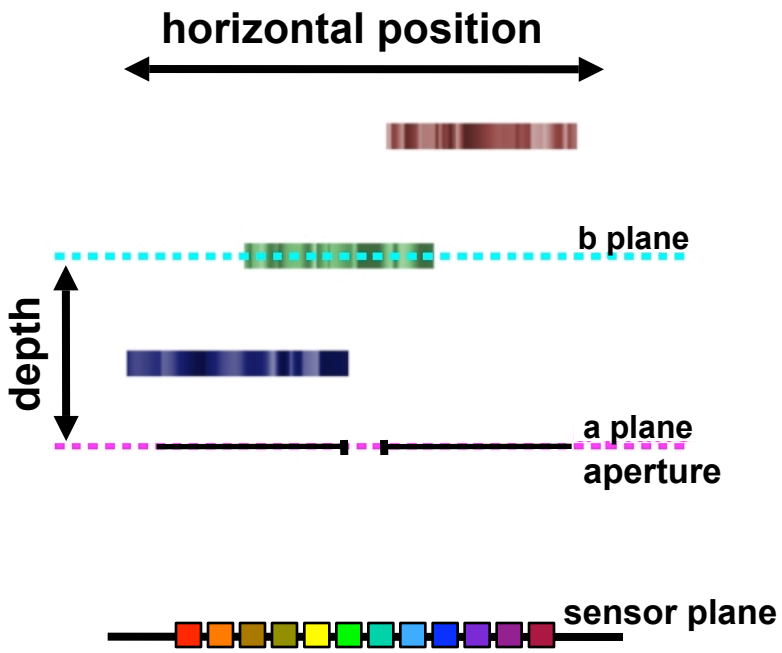


2D lightfield

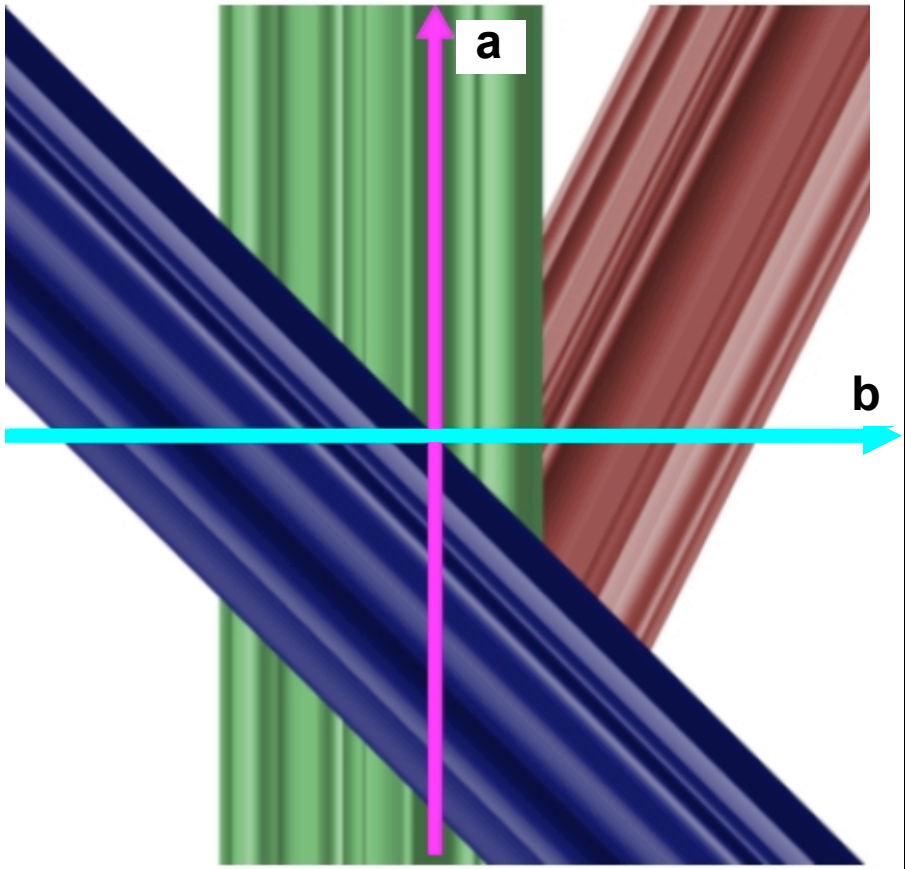


Pinhole camera

flatworld 1D scene

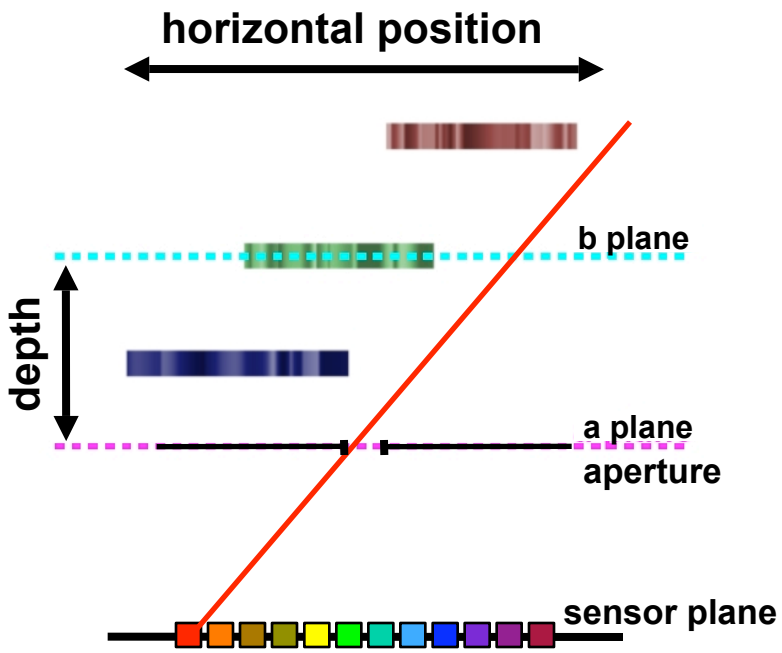


2D lightfield

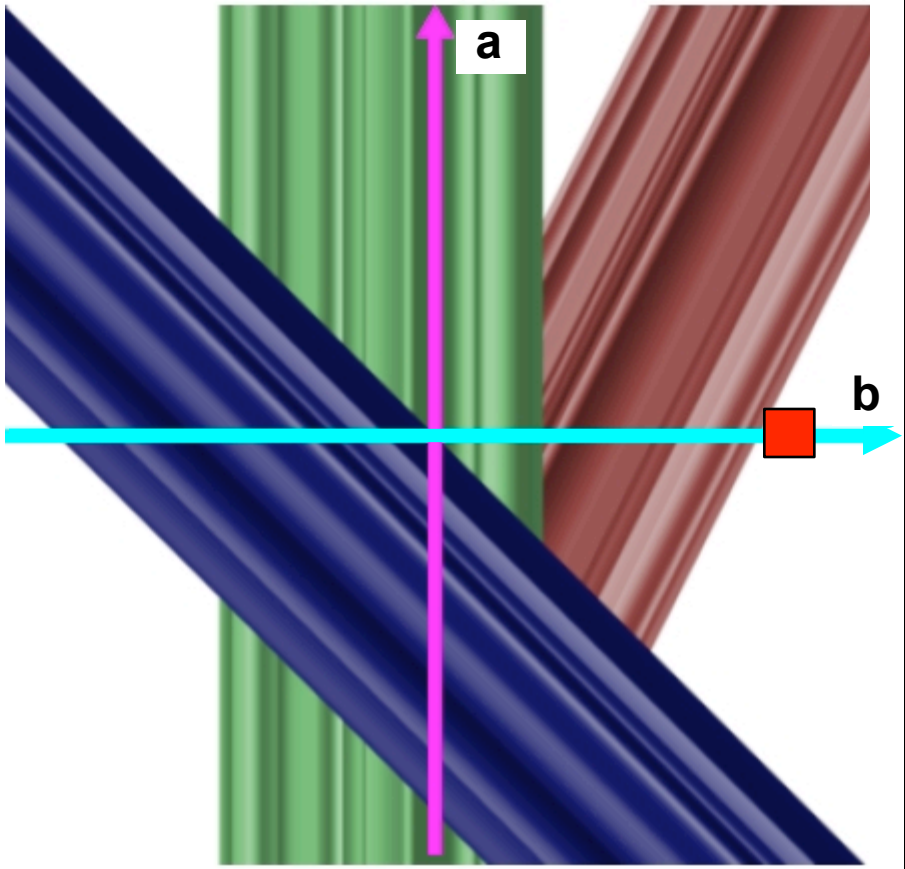


Pinhole camera

flatworld 1D scene

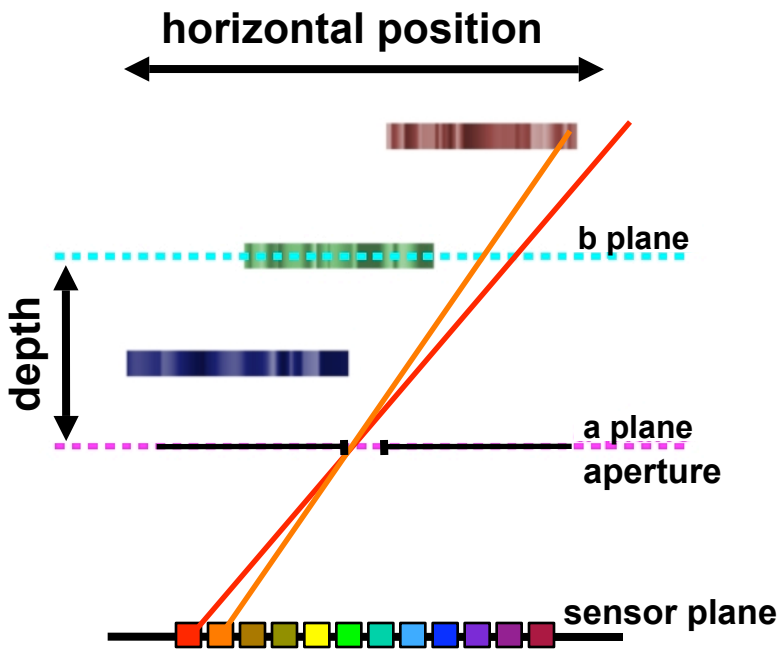


2D lightfield

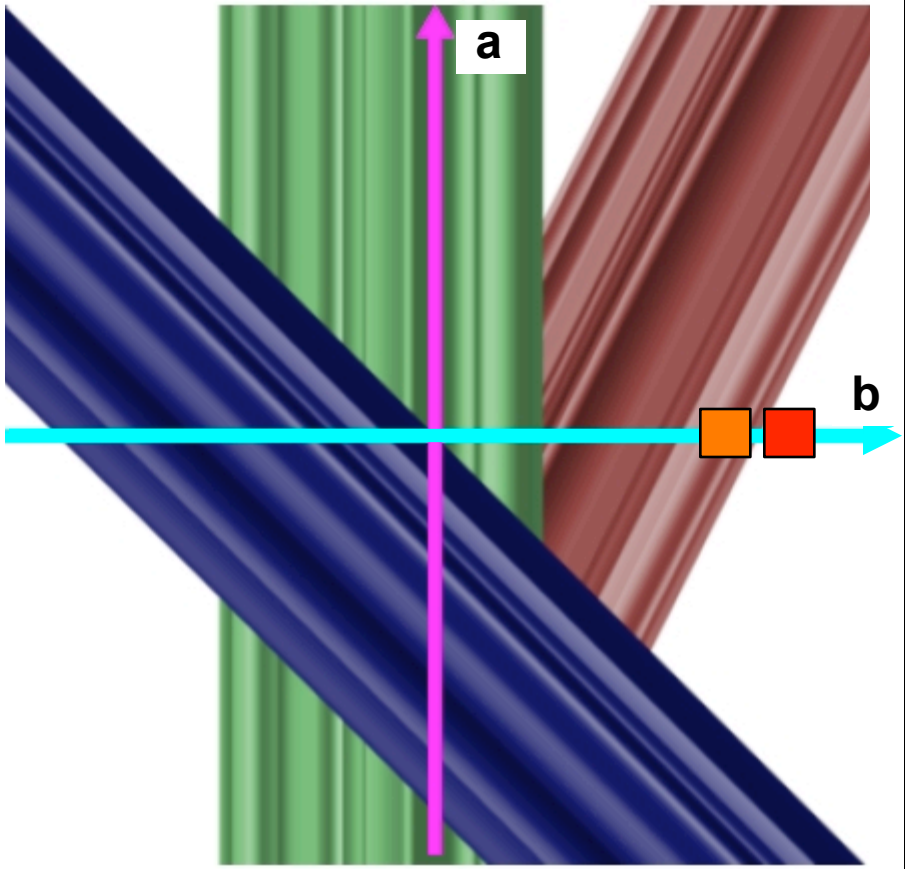


Pinhole camera

flatworld 1D scene

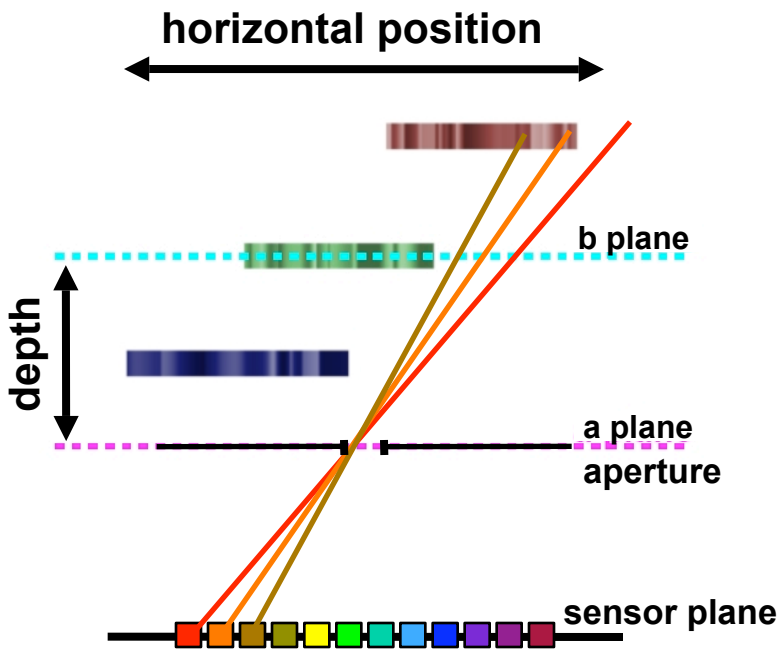


2D lightfield

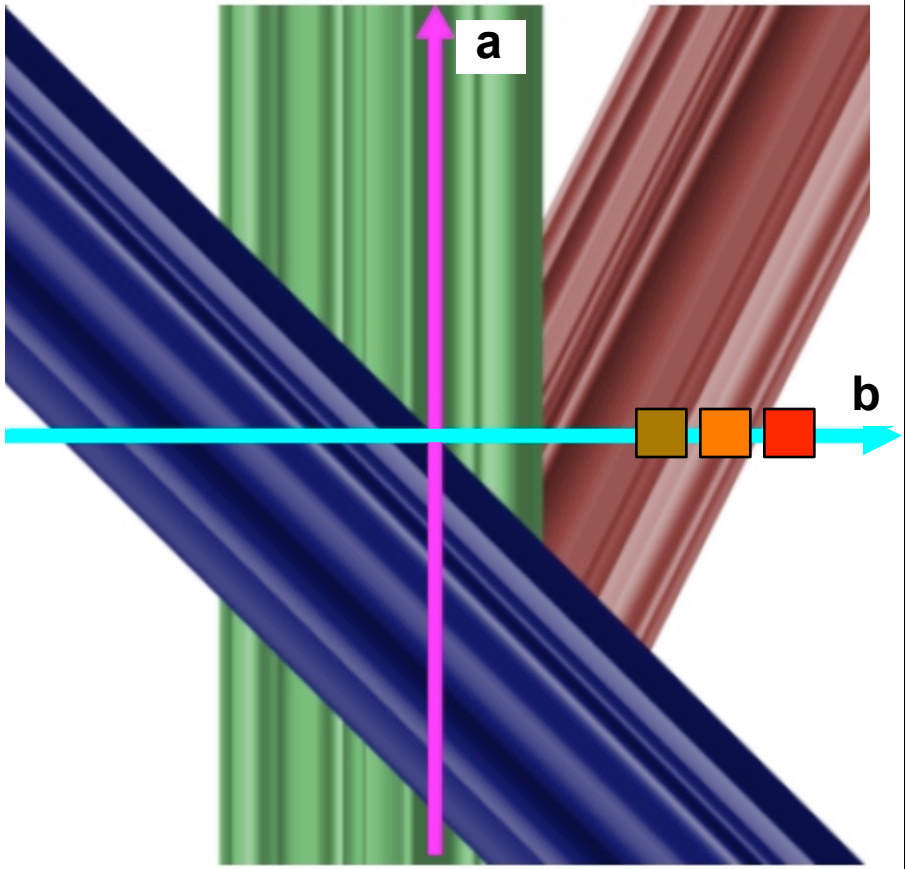


Pinhole camera

flatworld 1D scene

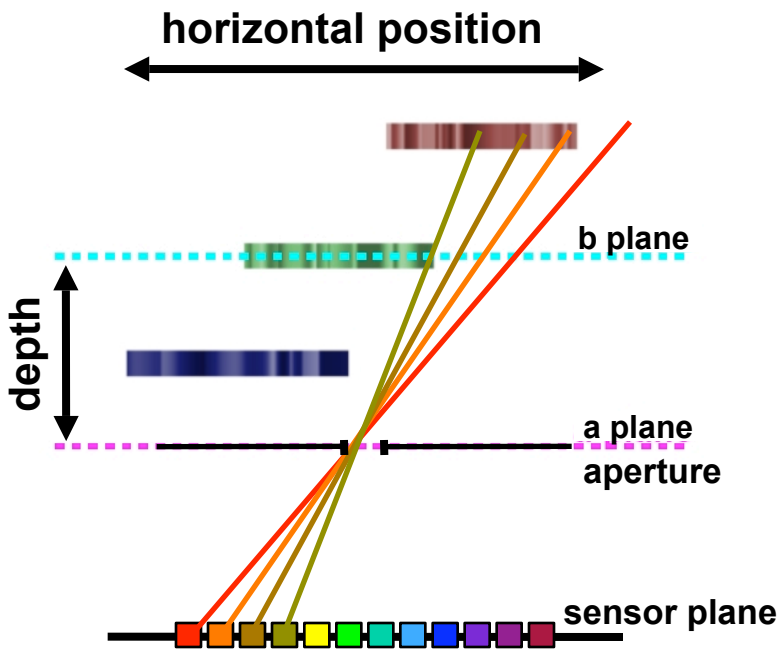


2D lightfield

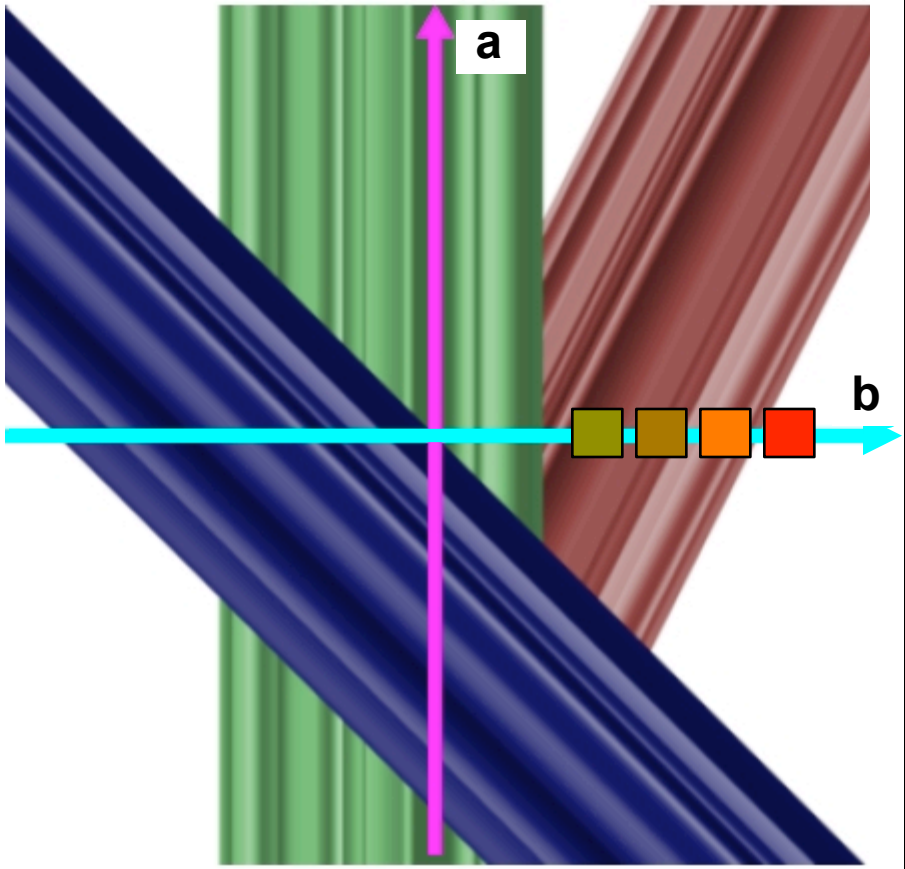


Pinhole camera

flatworld 1D scene

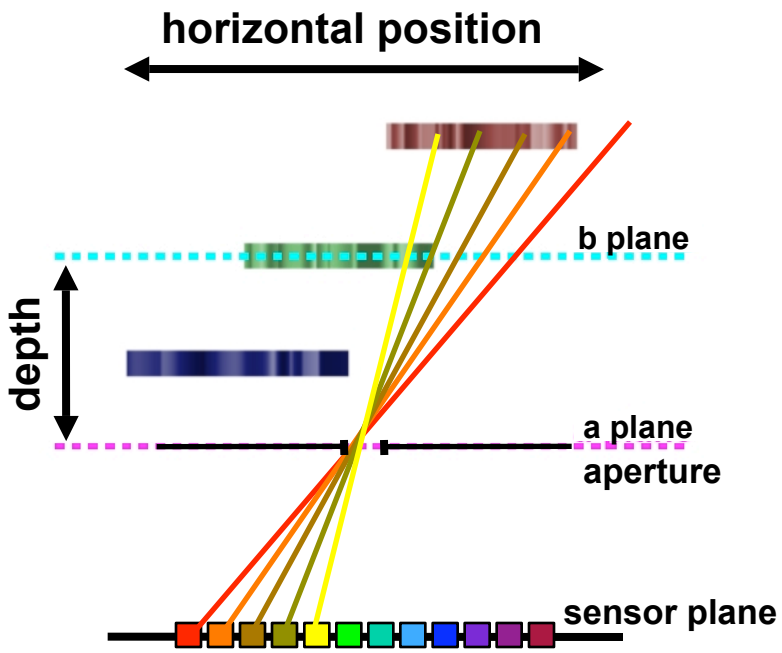


2D lightfield

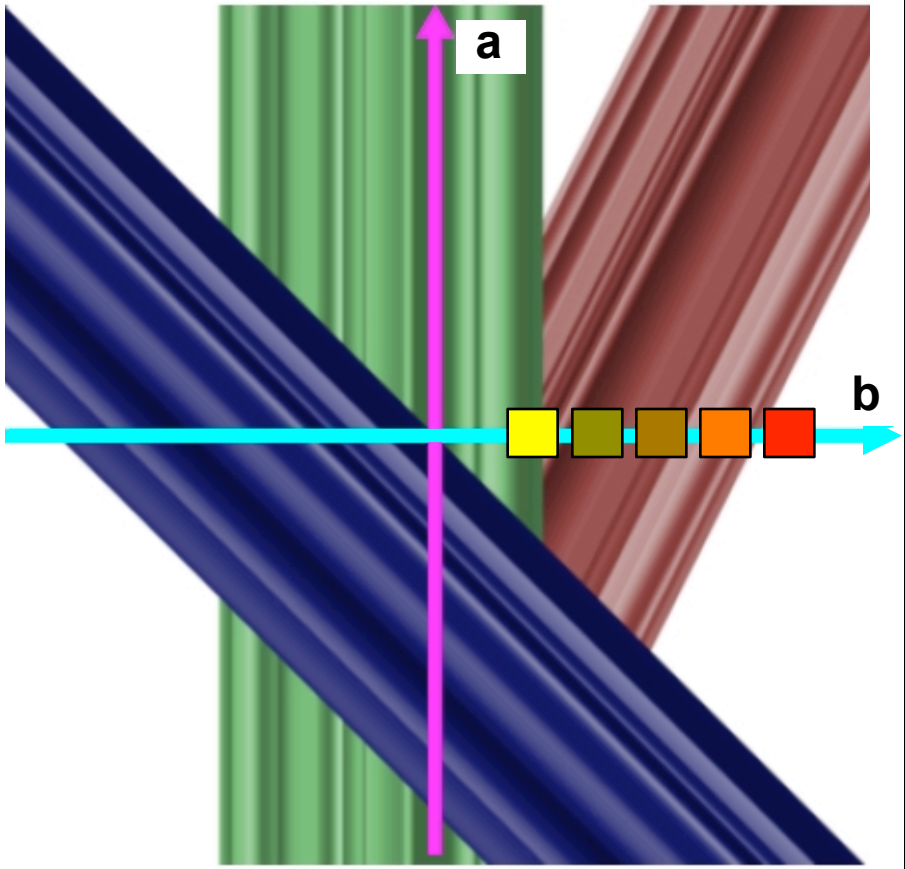


Pinhole camera

flatworld 1D scene

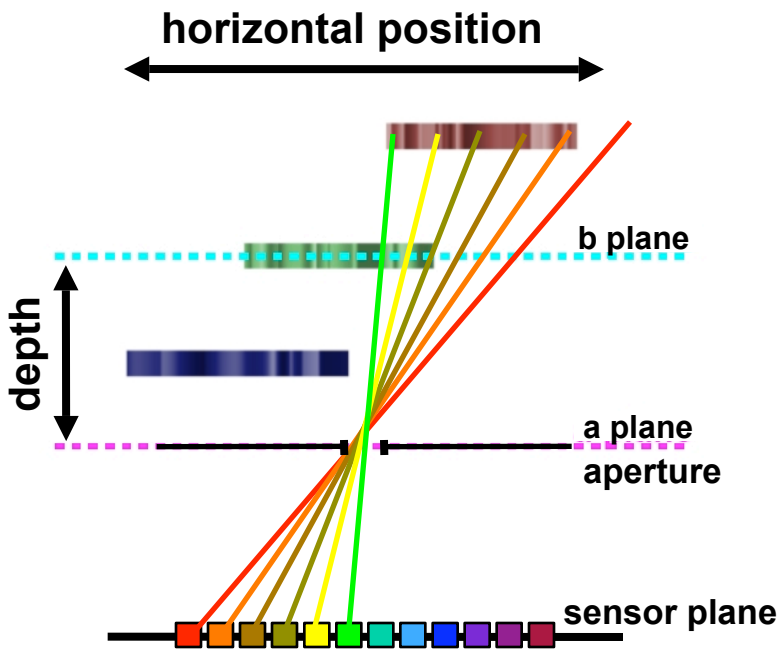


2D lightfield

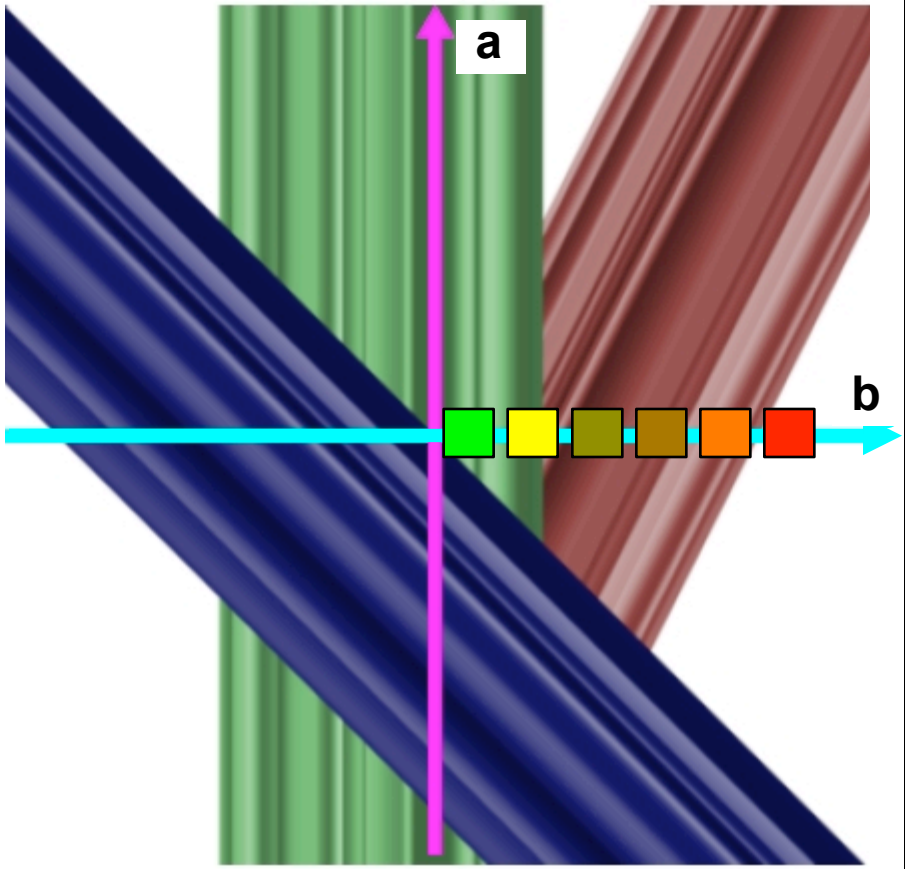


Pinhole camera

flatworld 1D scene

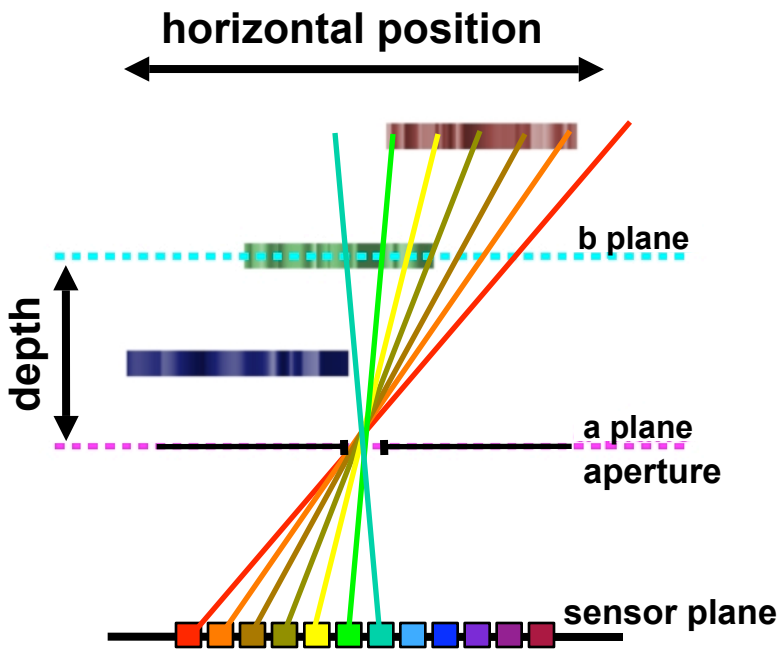


2D lightfield

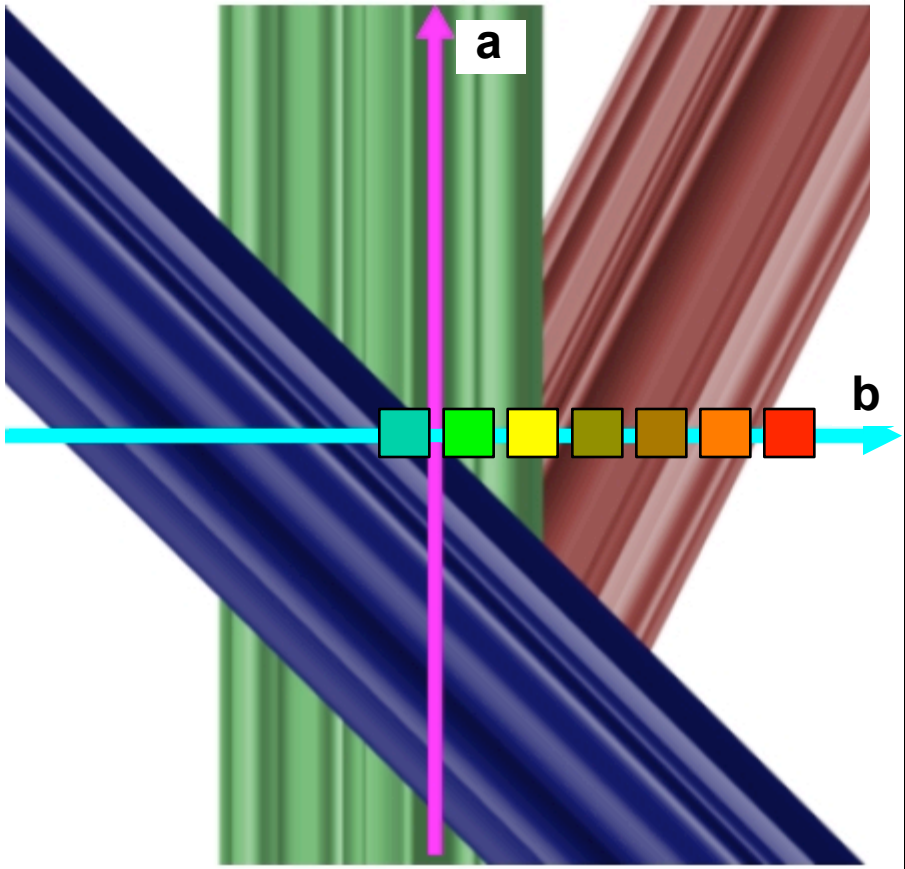


Pinhole camera

flatworld 1D scene

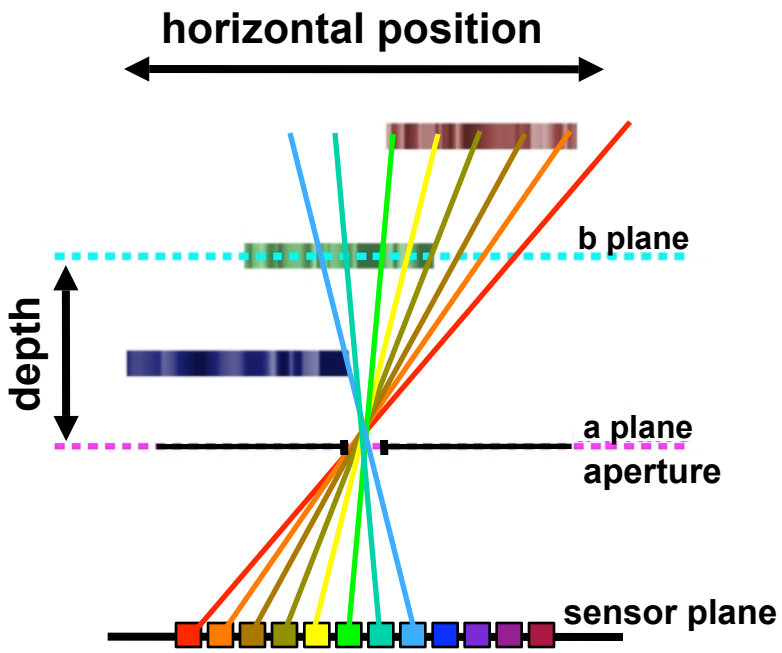


2D lightfield

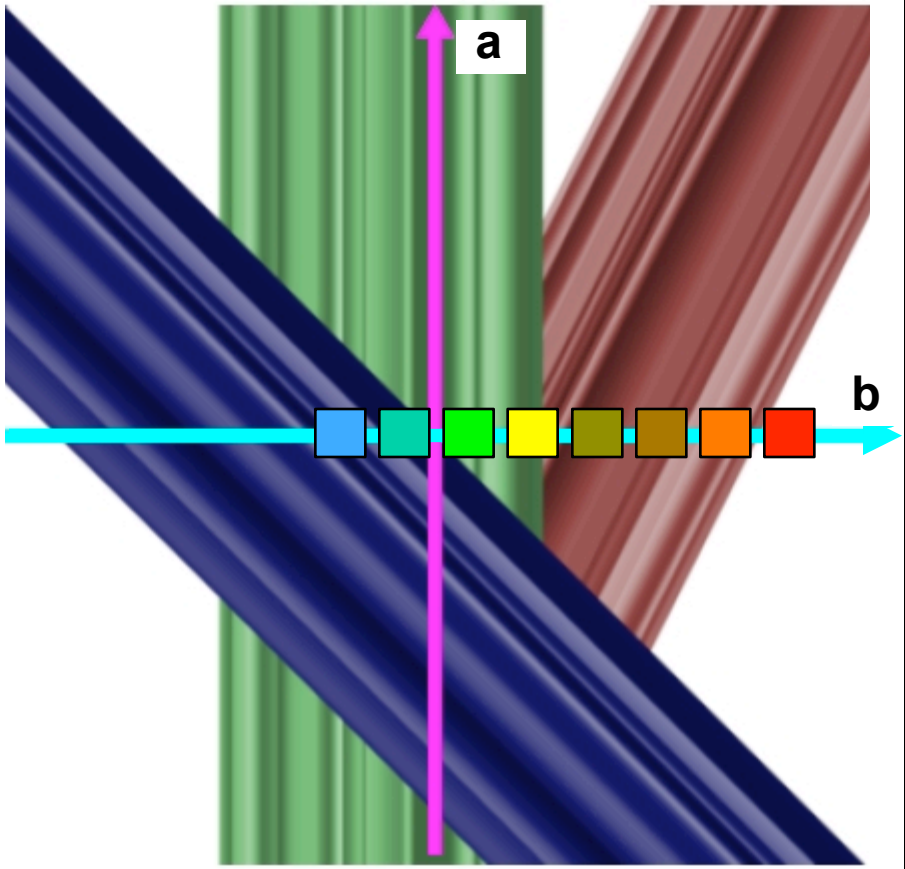


Pinhole camera

flatworld 1D scene

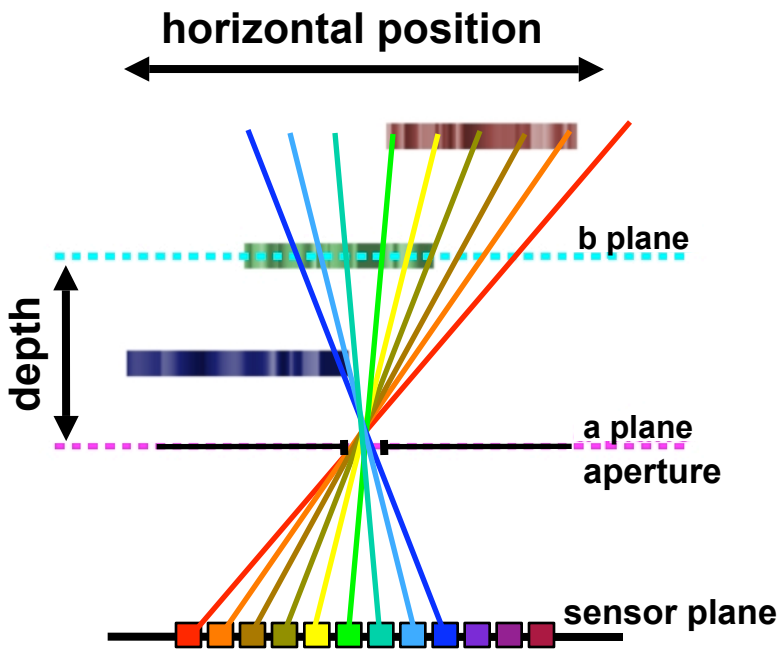


2D lightfield

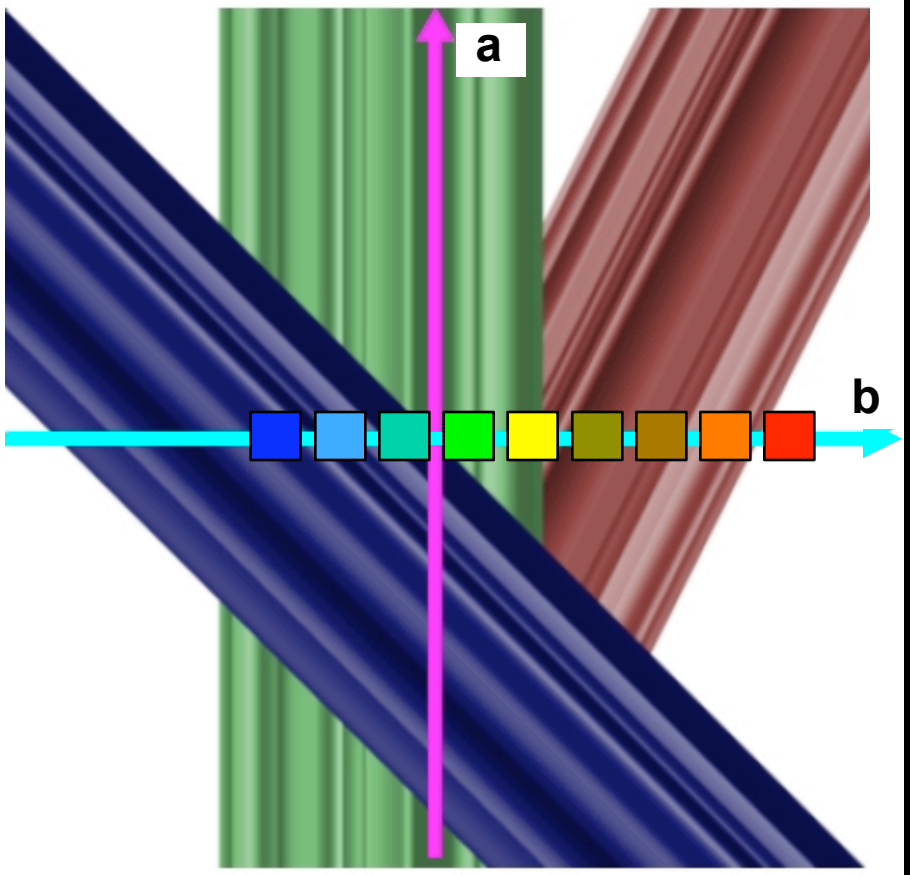


Pinhole camera

flatworld 1D scene

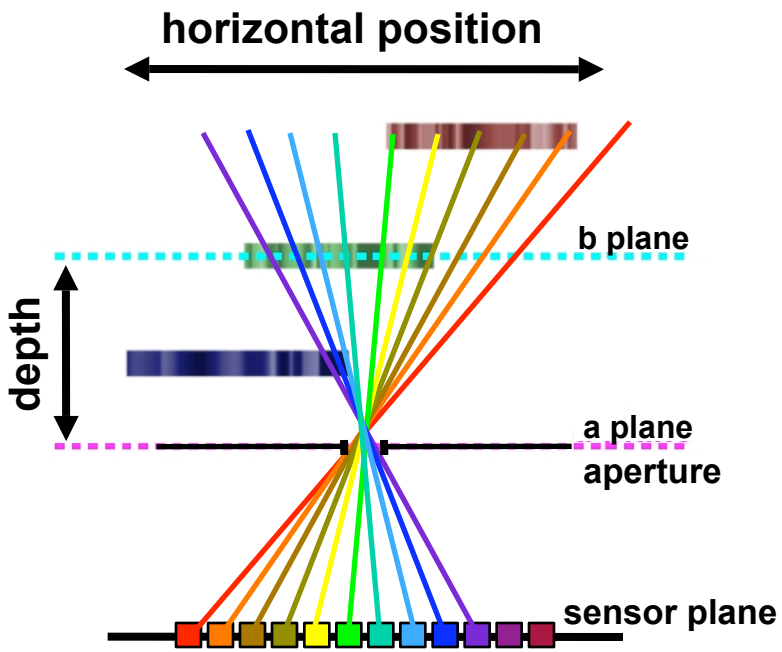


2D lightfield

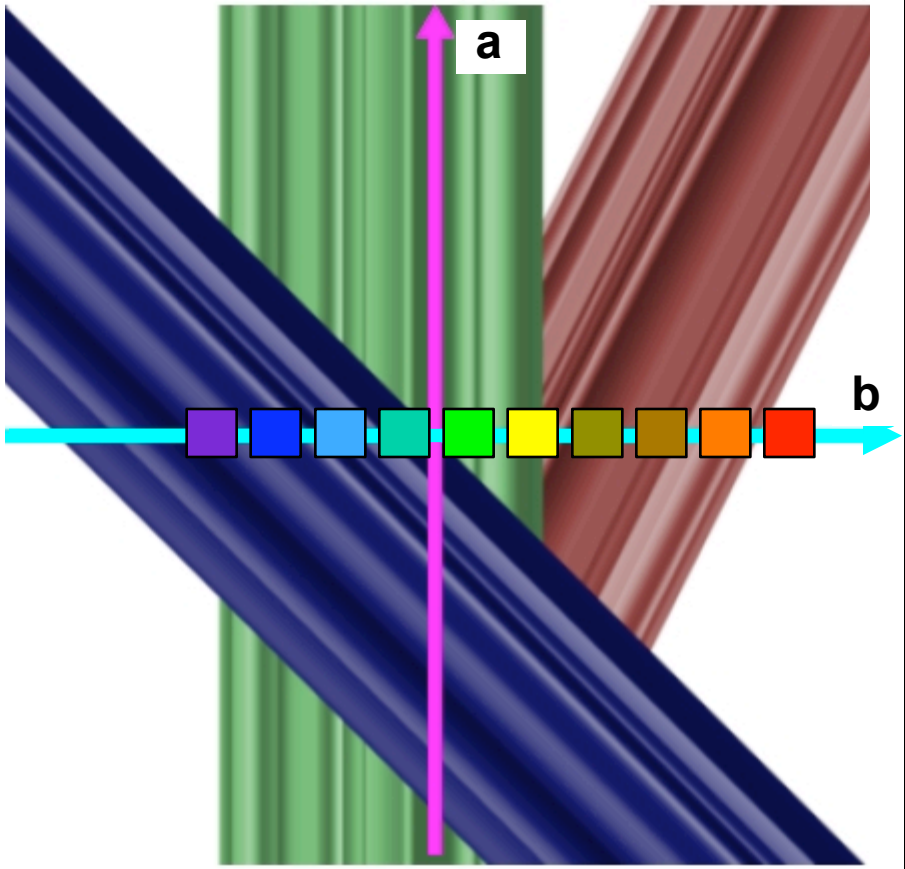


Pinhole camera

flatworld 1D scene

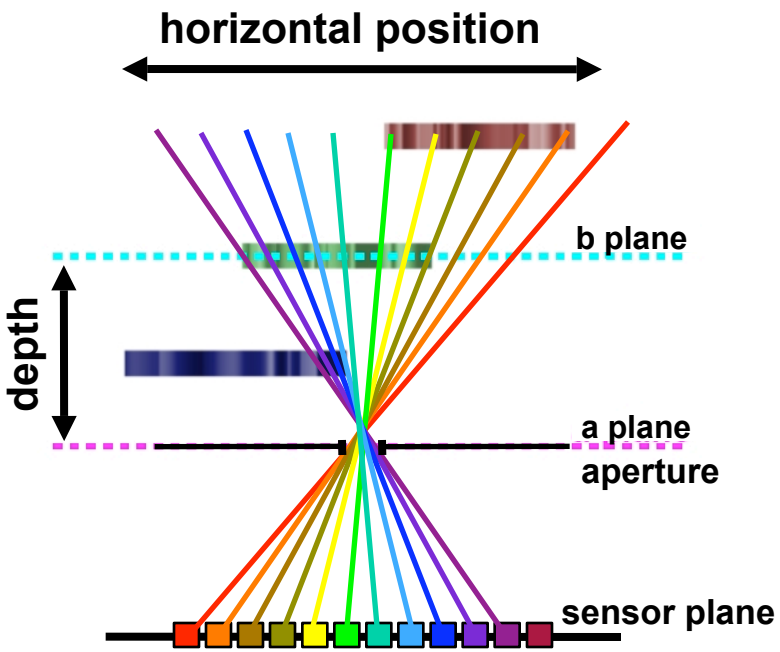


2D lightfield

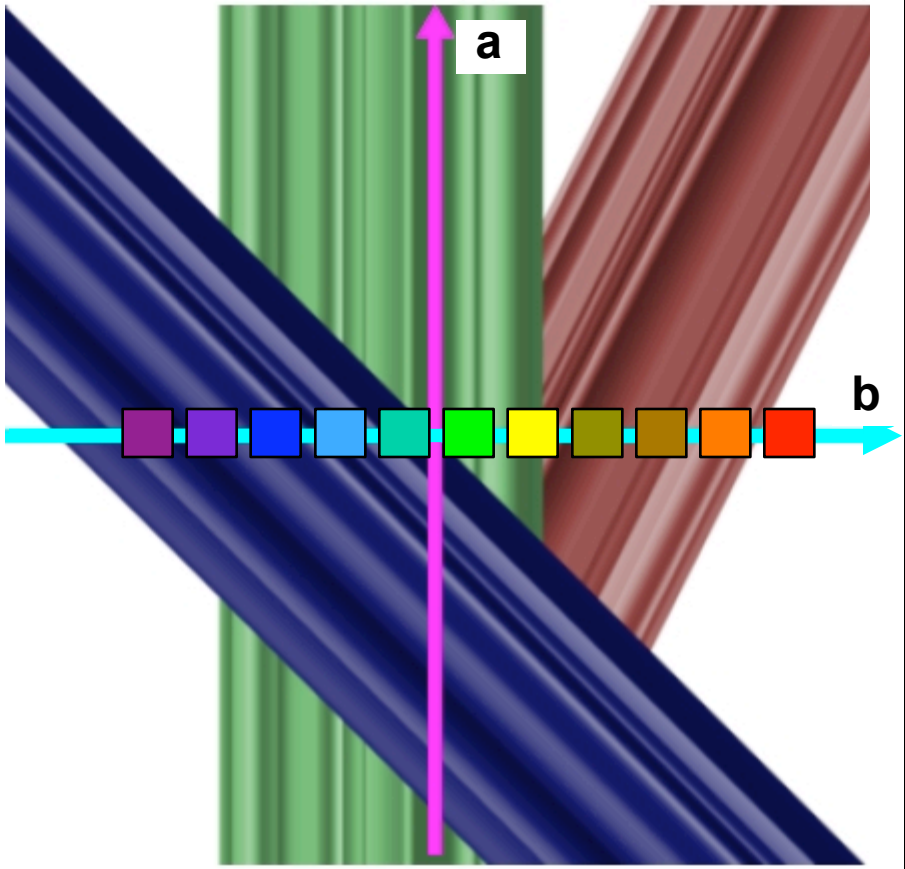


Pinhole camera

flatworld 1D scene

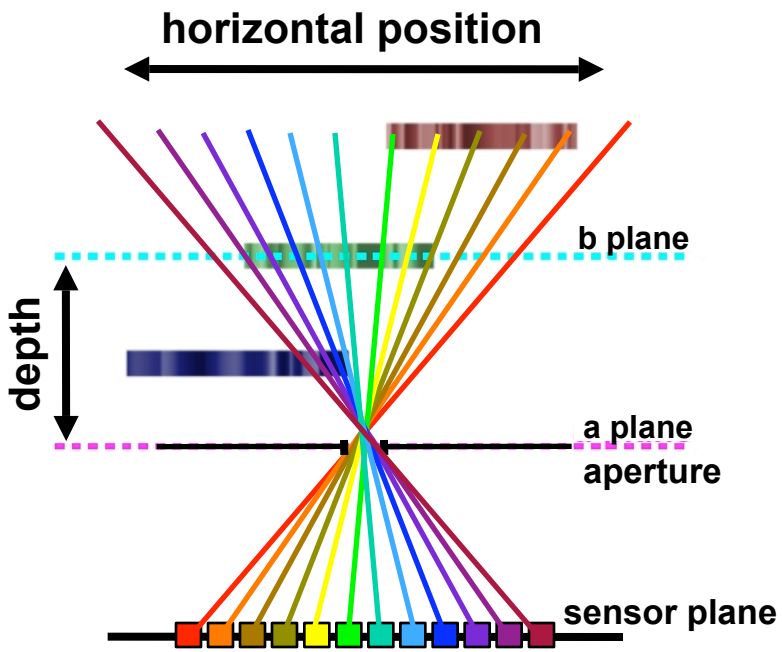


2D lightfield

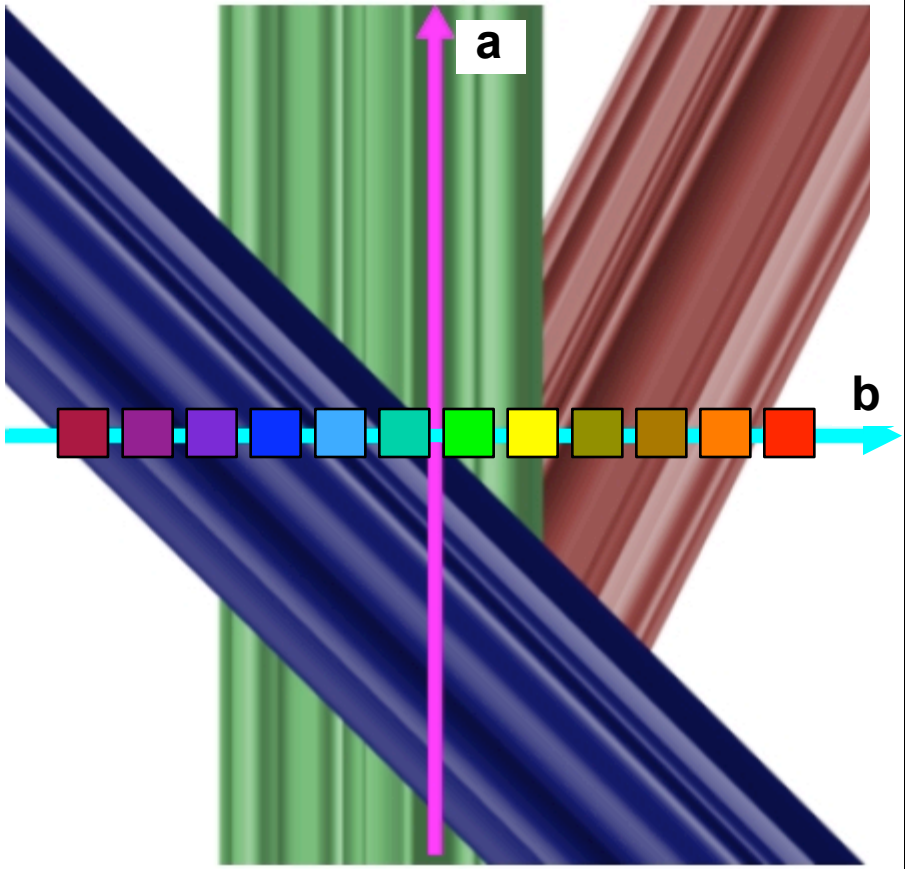


Pinhole camera

flatworld 1D scene

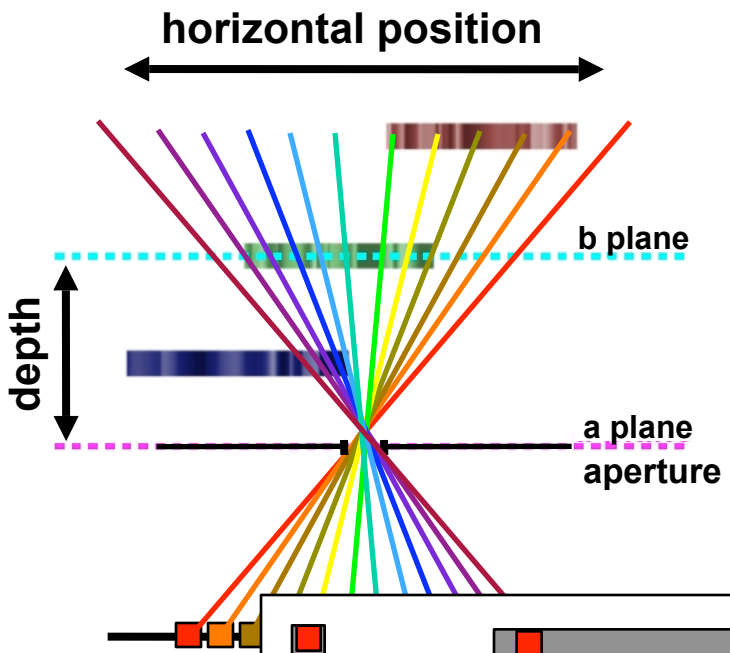


2D lightfield

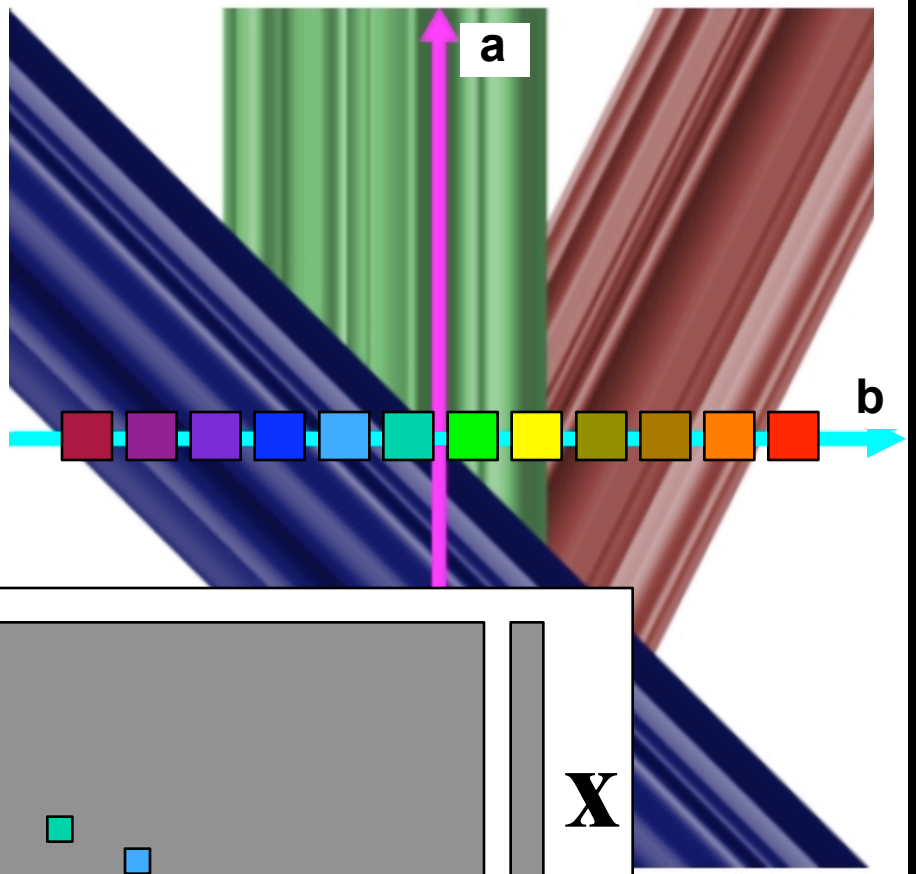


Pinhole camera

flatworld 1D scene



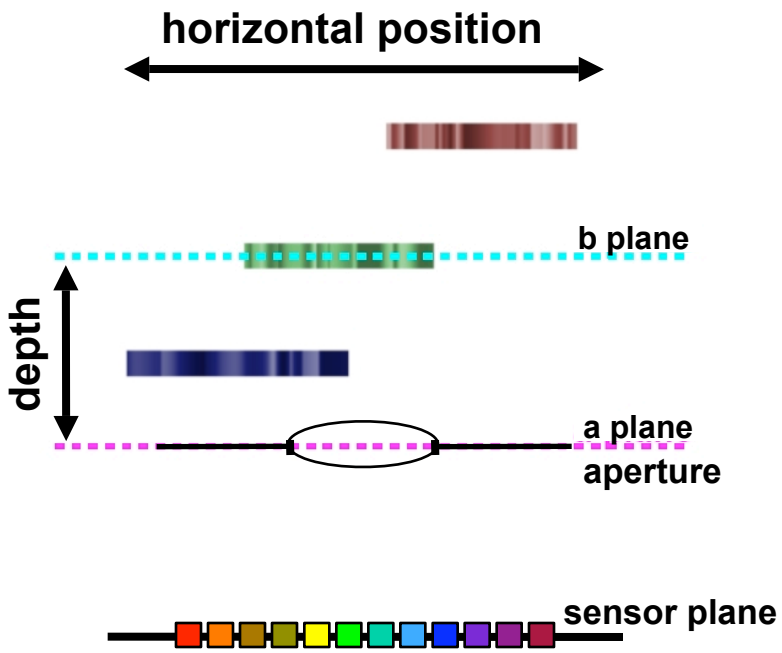
2D lightfield



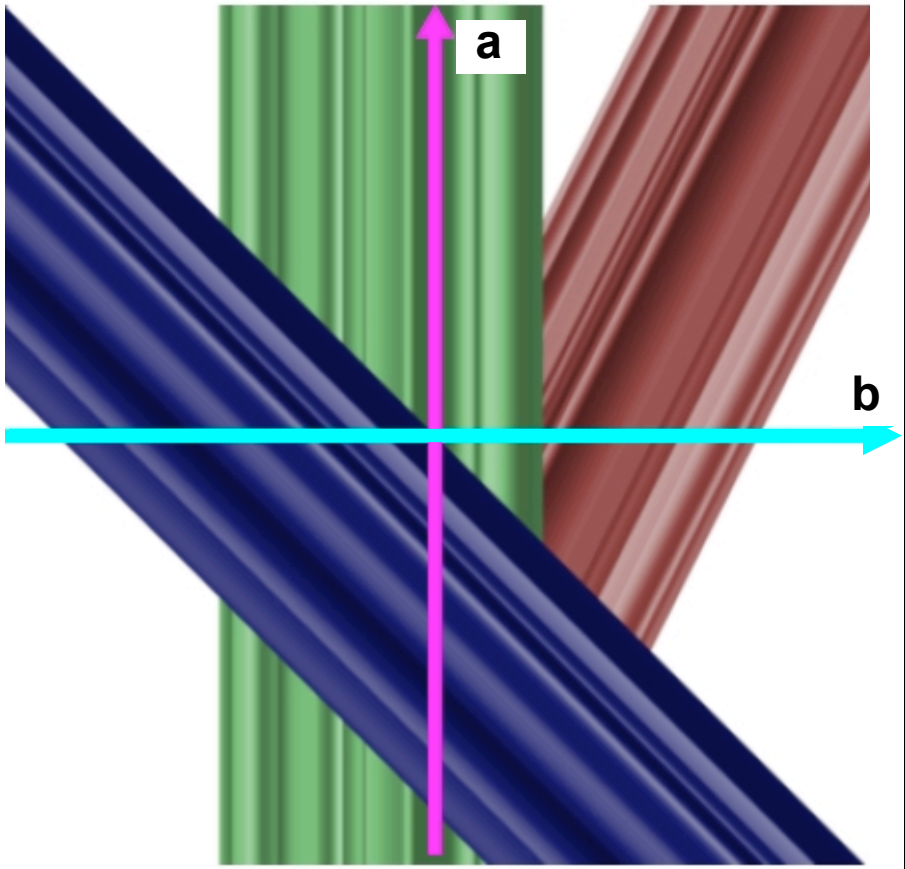
$$y = T x$$

Lens, focused at green object

flatworld 1D scene

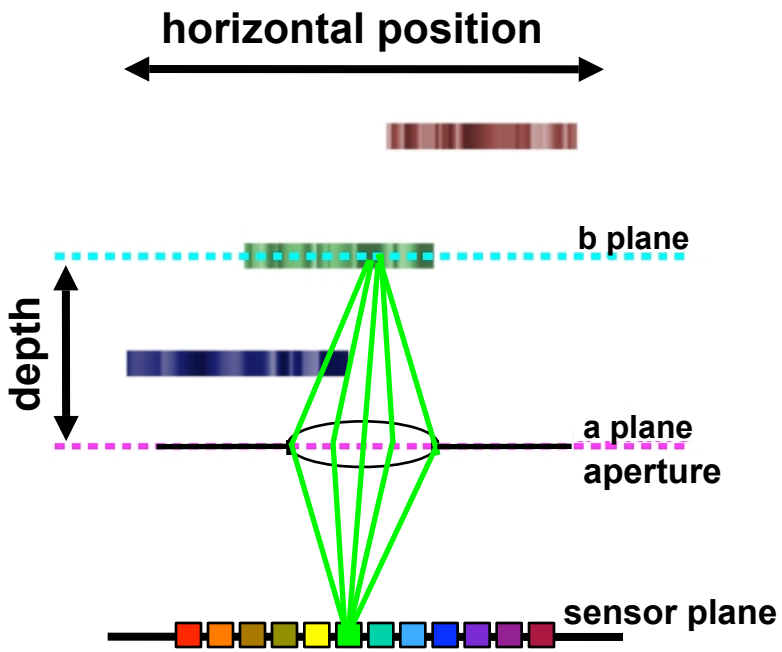


2D lightfield

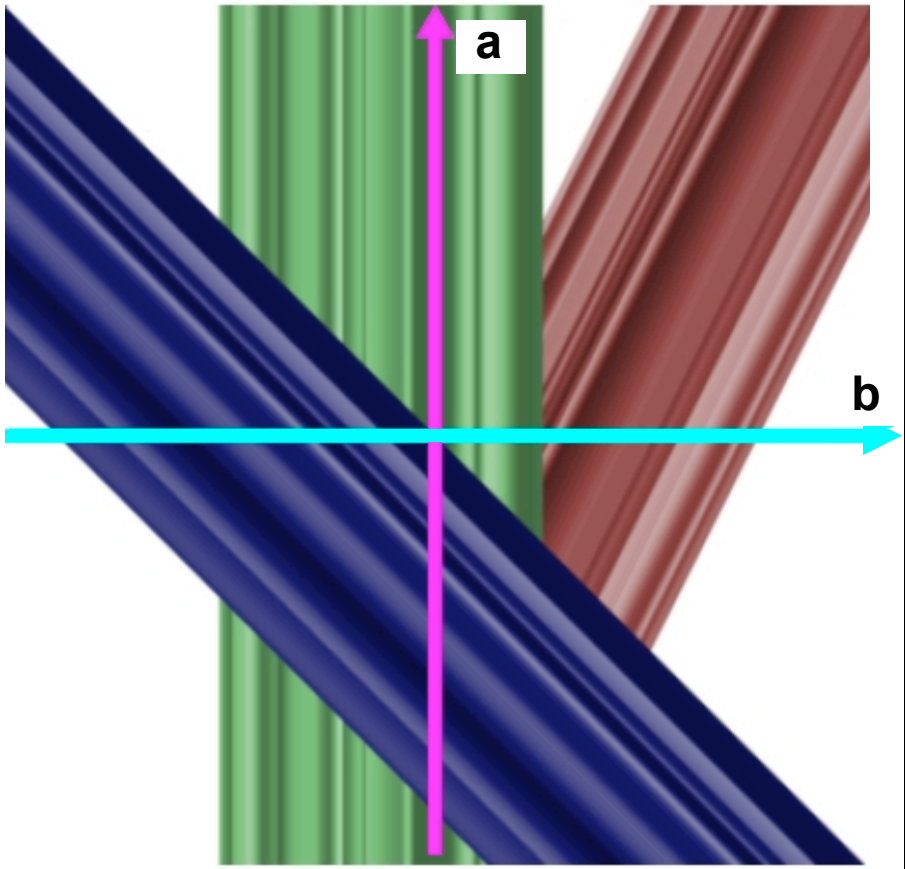


Lens, focused at green object

flatworld 1D scene

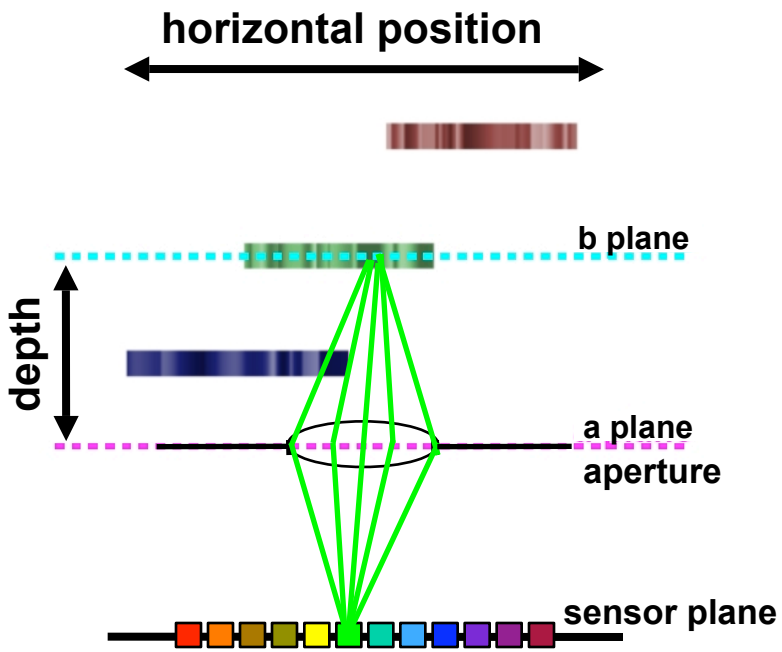


2D lightfield

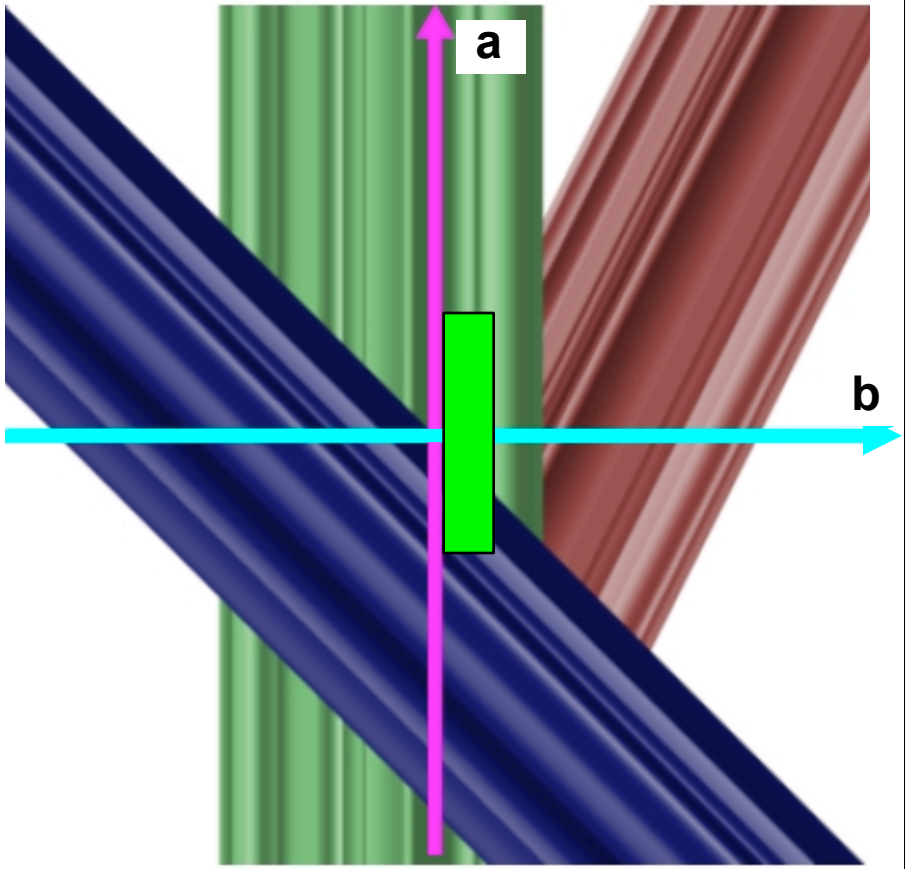


Lens, focused at green object

flatworld 1D scene

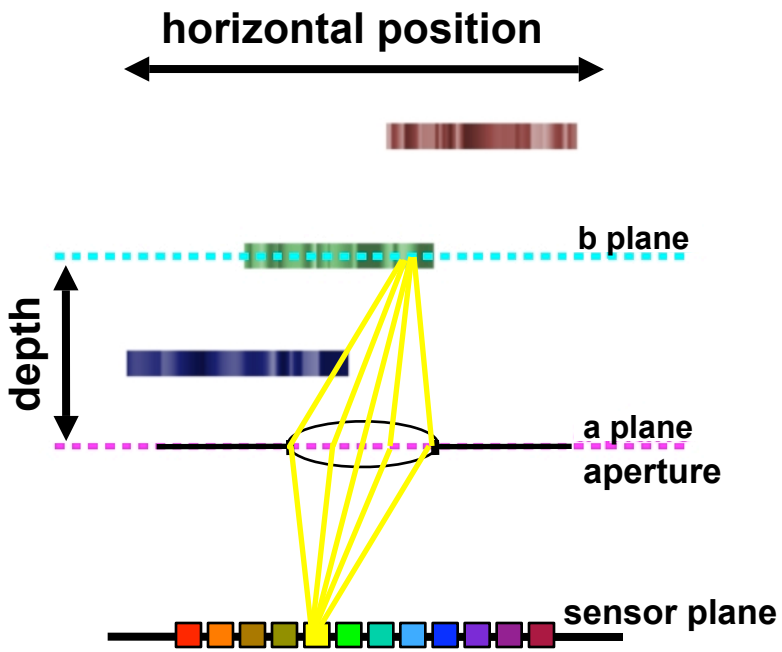


2D lightfield

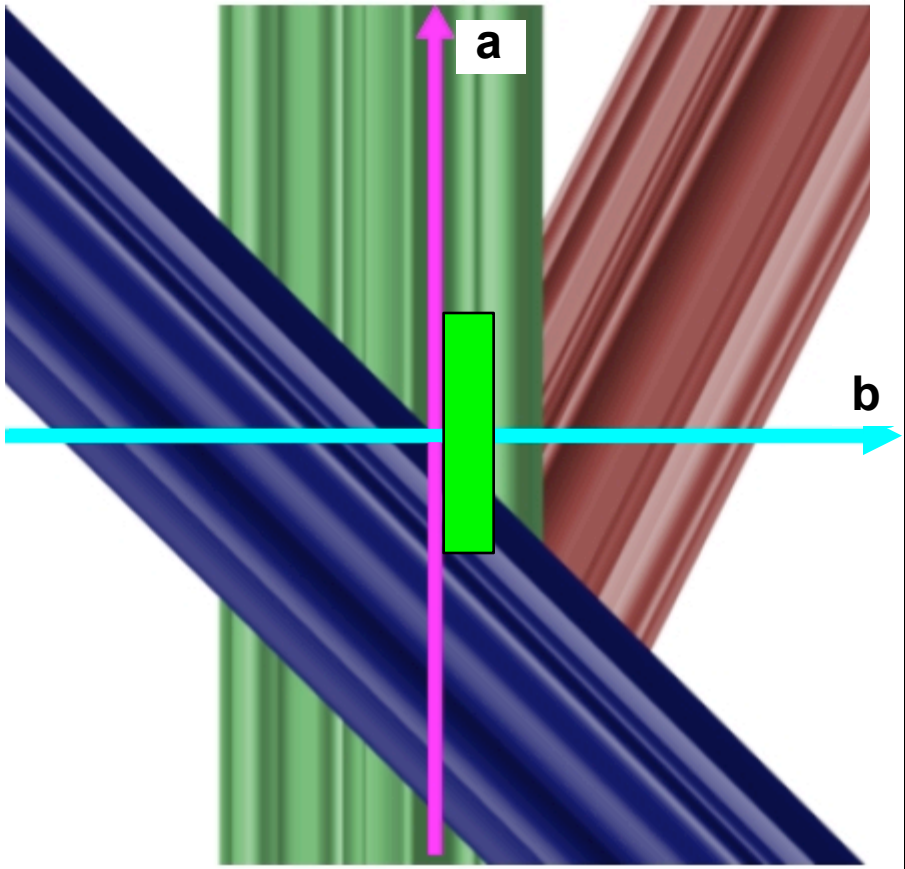


Lens, focused at green object

flatworld 1D scene

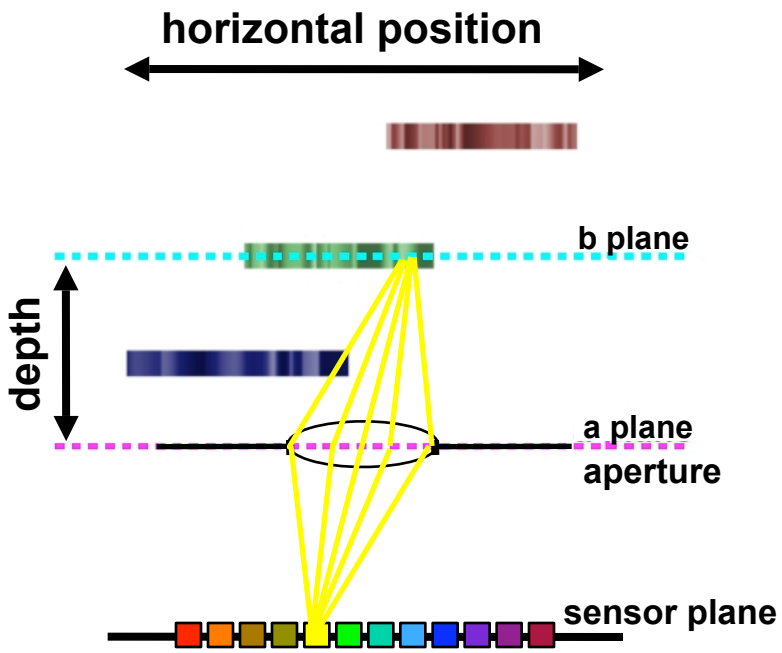


2D lightfield

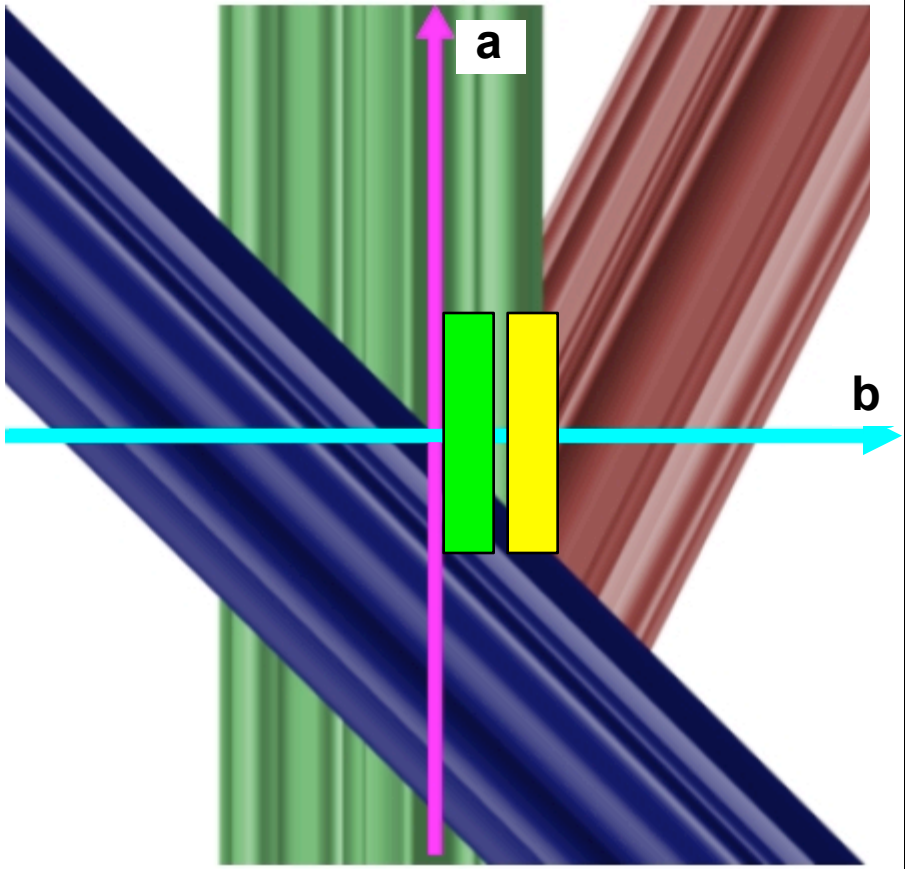


Lens, focused at green object

flatworld 1D scene

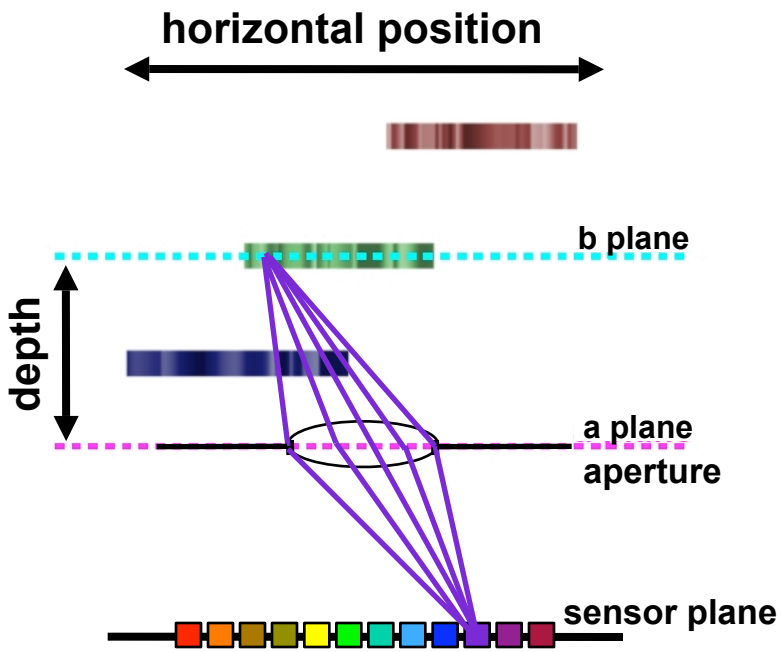


2D lightfield

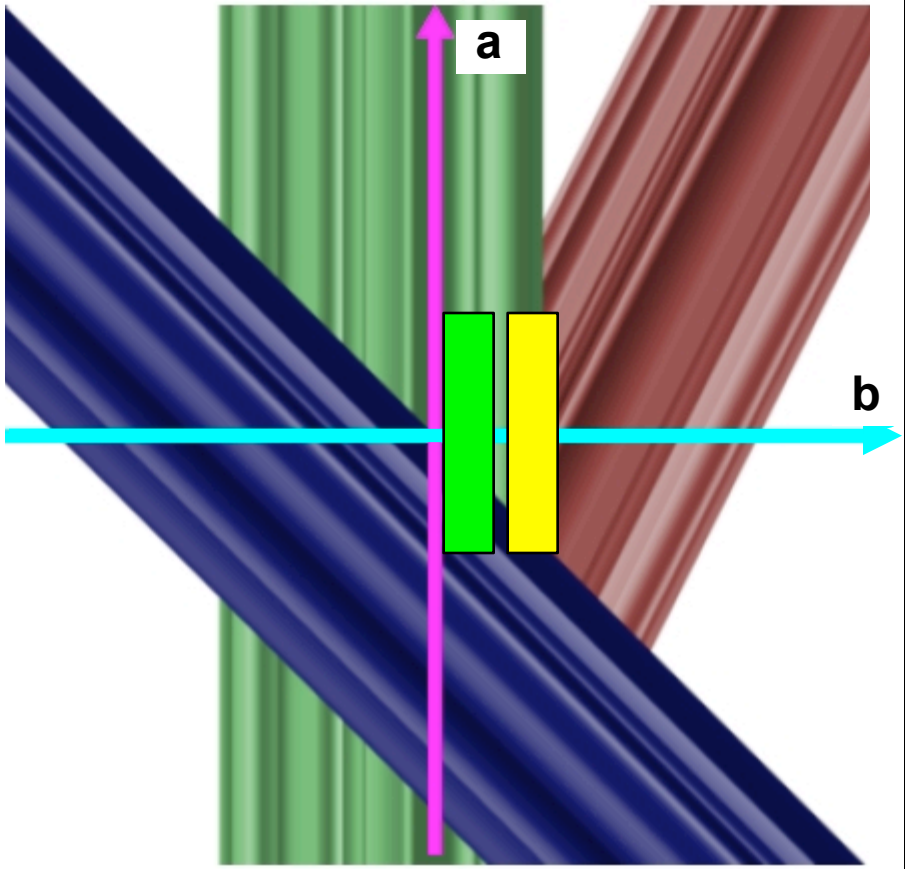


Lens, focused at green object

flatworld 1D scene

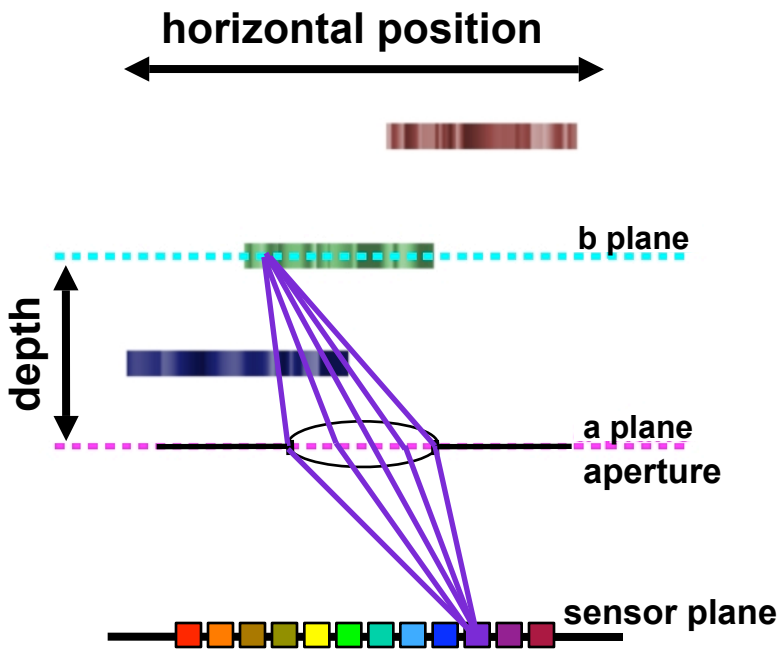


2D lightfield

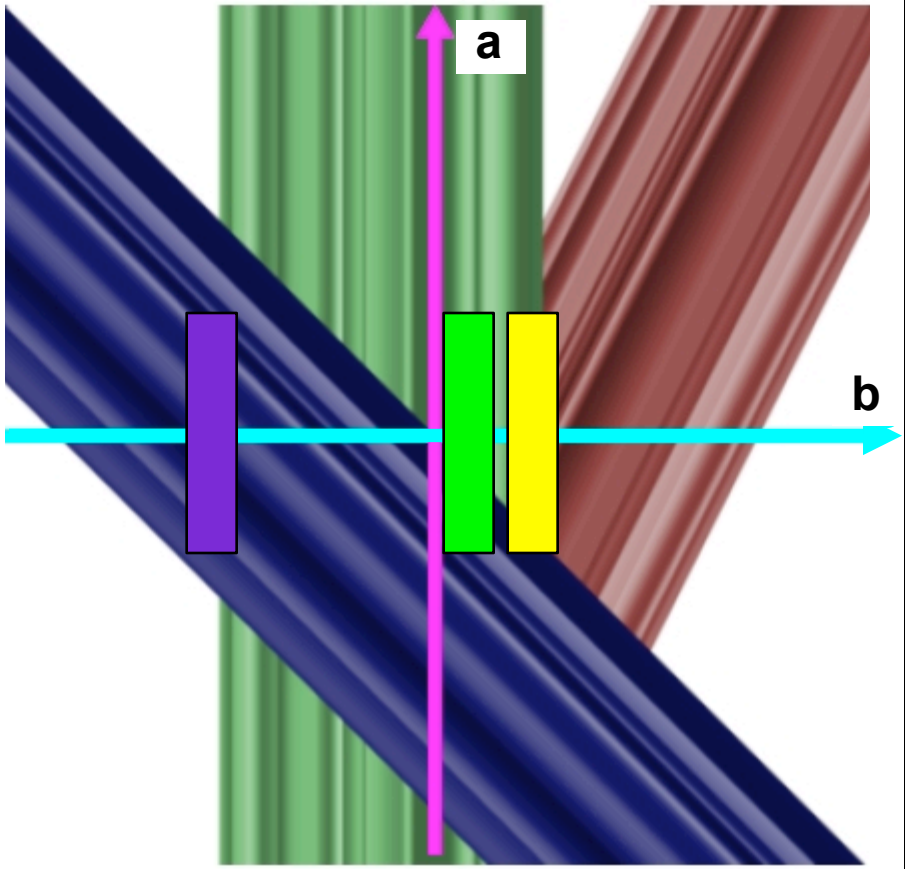


Lens, focused at green object

flatworld 1D scene

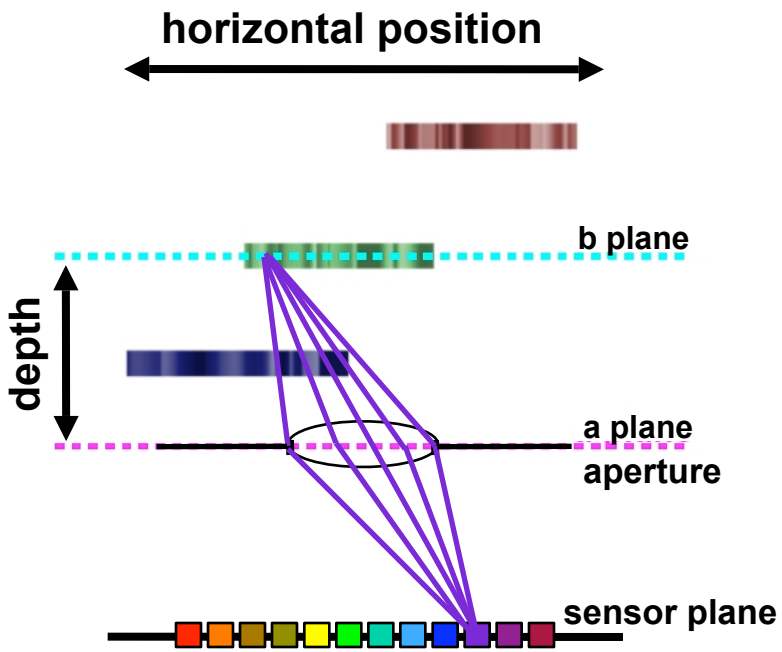


2D lightfield

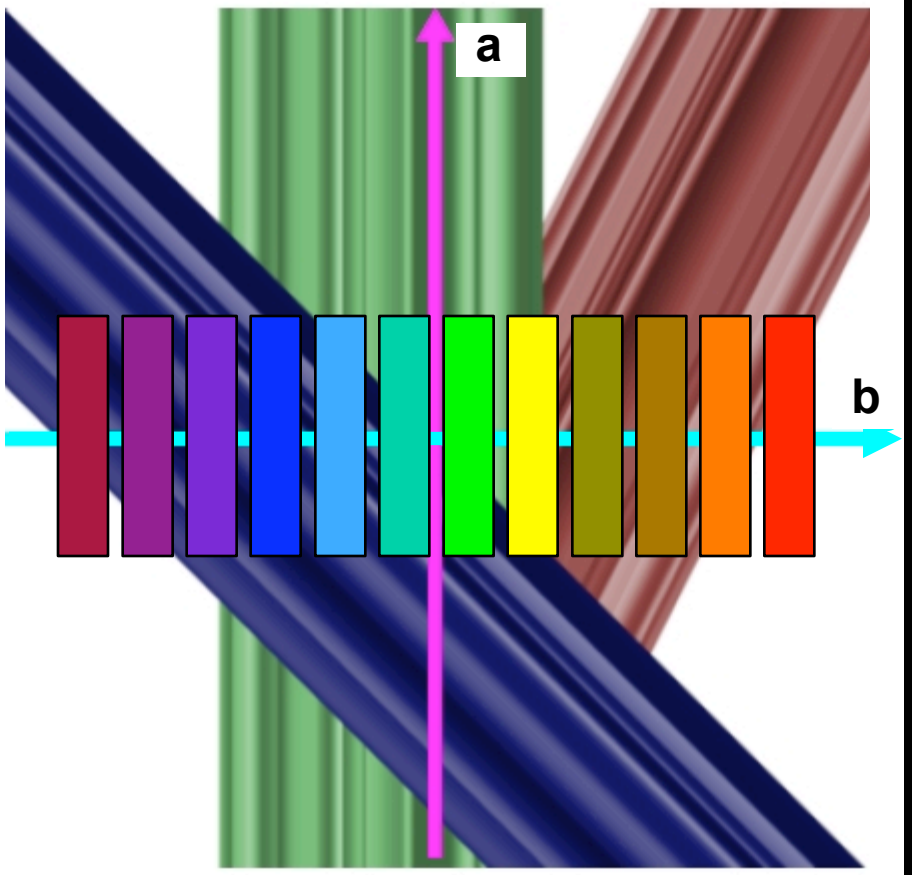


Lens, focused at green object

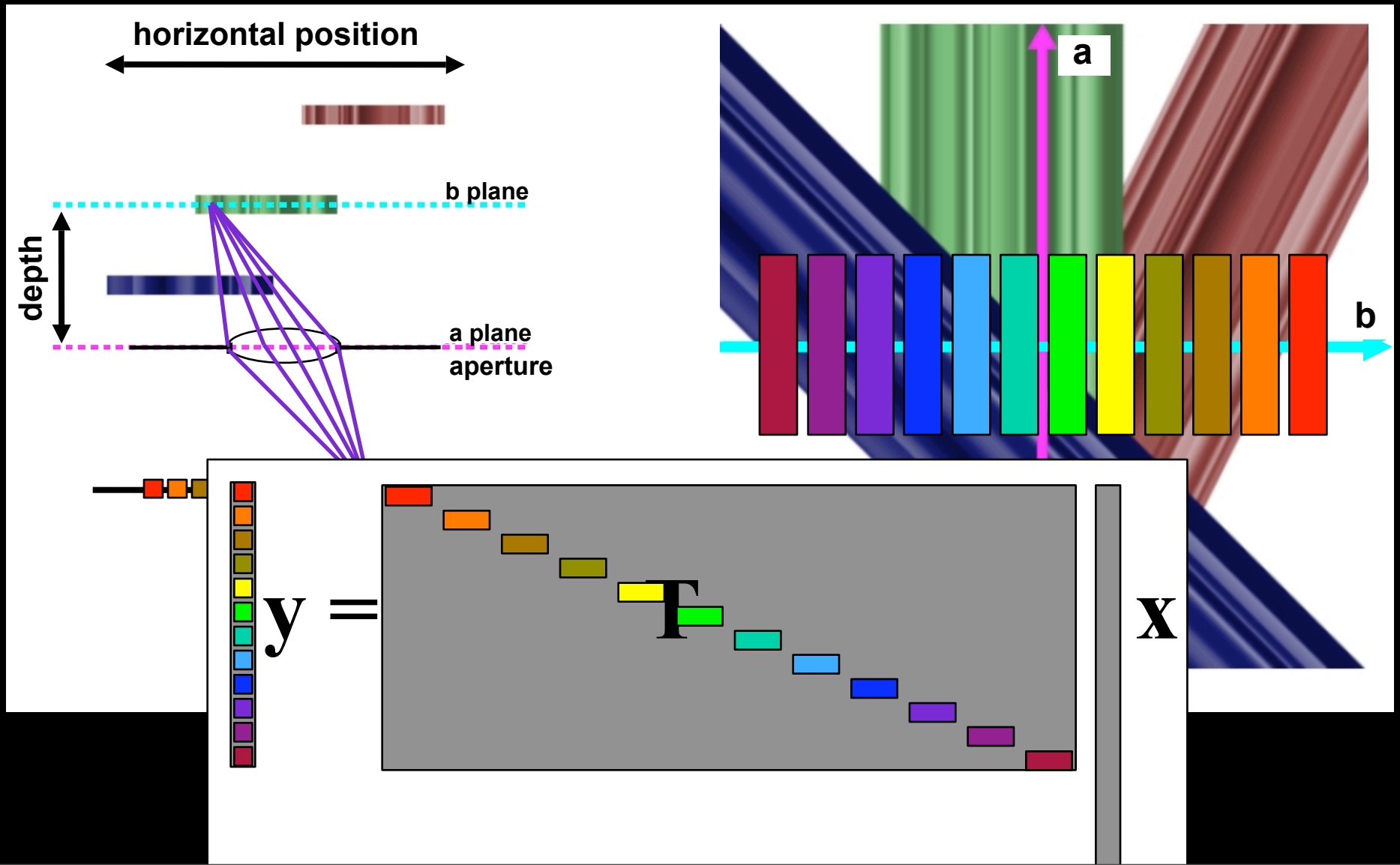
flatworld 1D scene



2D lightfield

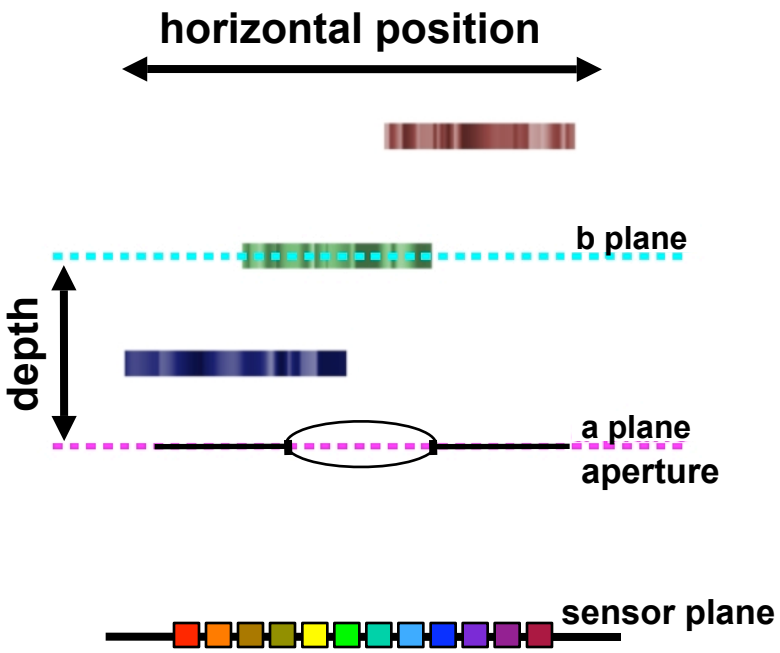


Lens, focused at green object

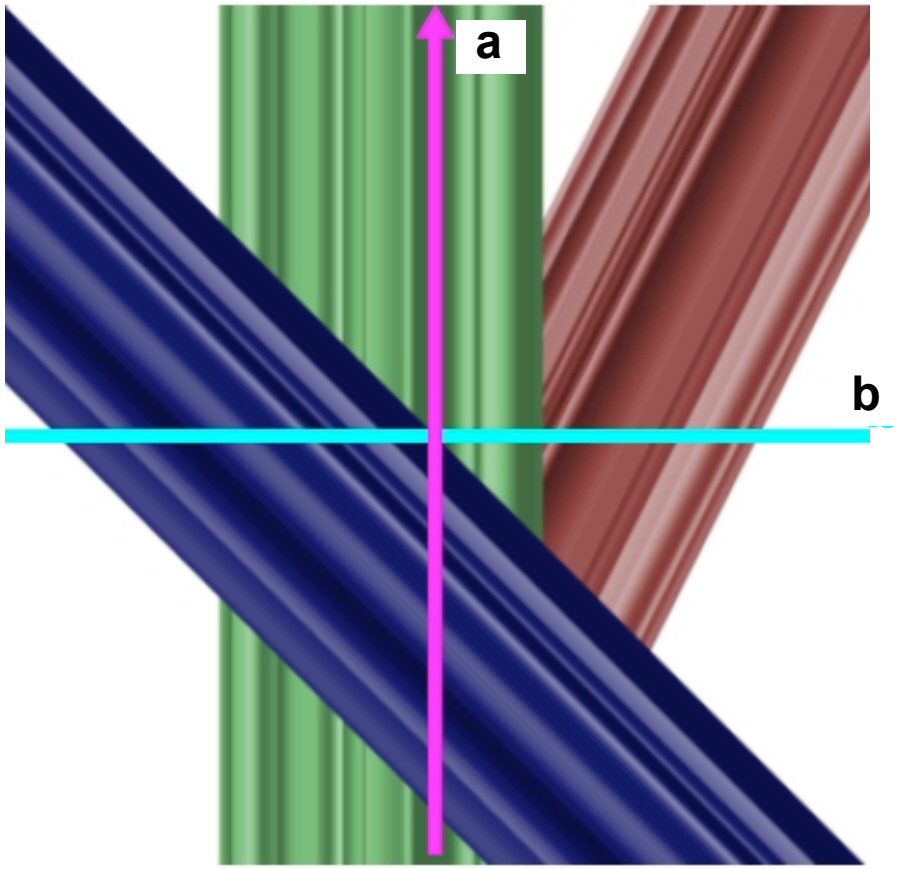


Lens, focused at blue object

flatworld 1D scene

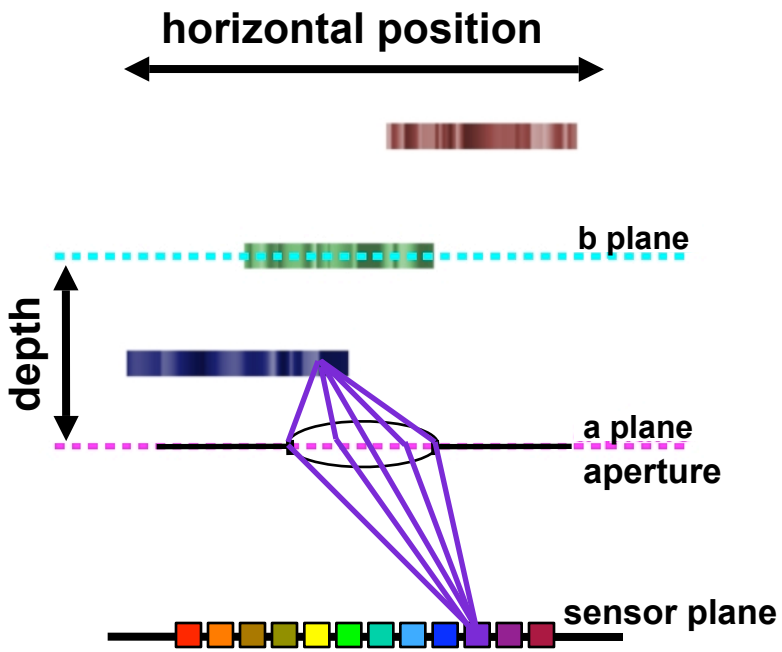


2D lightfield

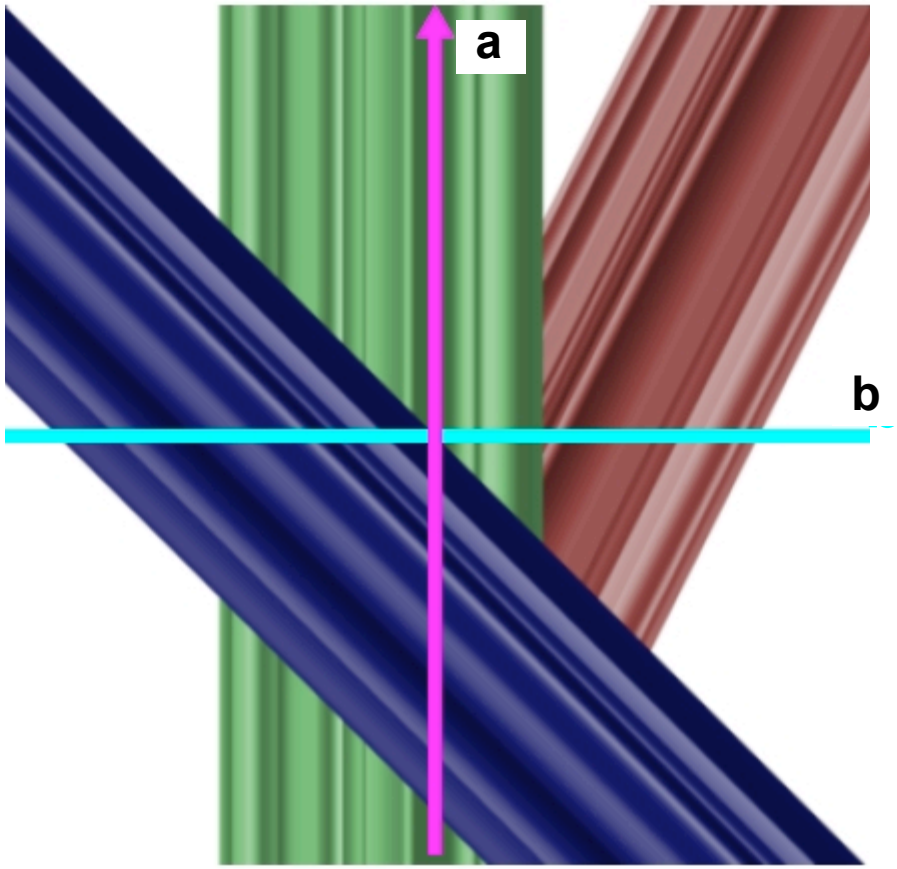


Lens, focused at blue object

flatworld 1D scene

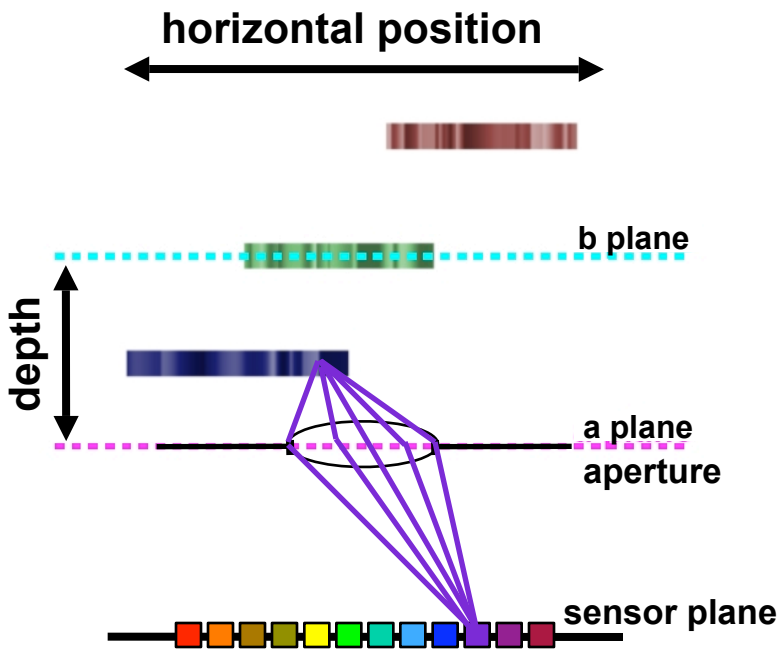


2D lightfield

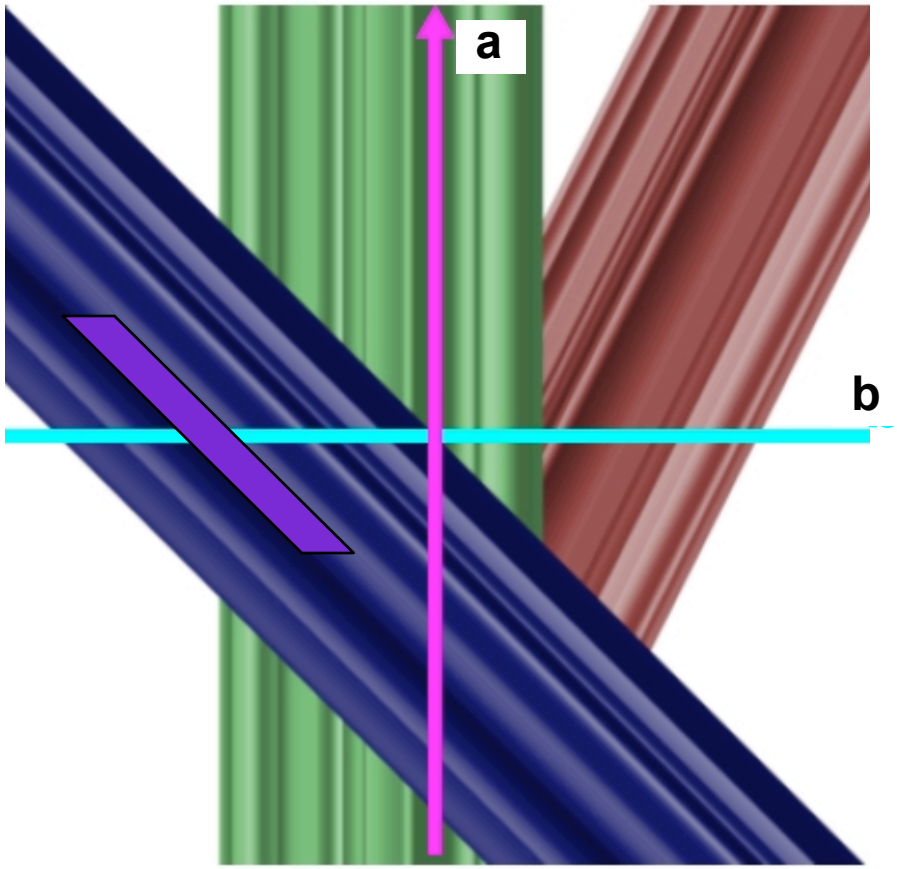


Lens, focused at blue object

flatworld 1D scene

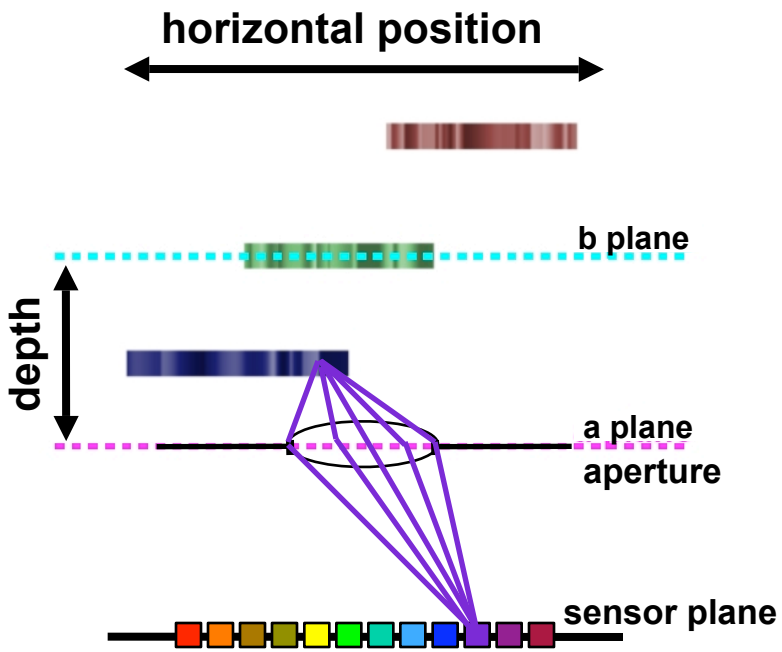


2D lightfield

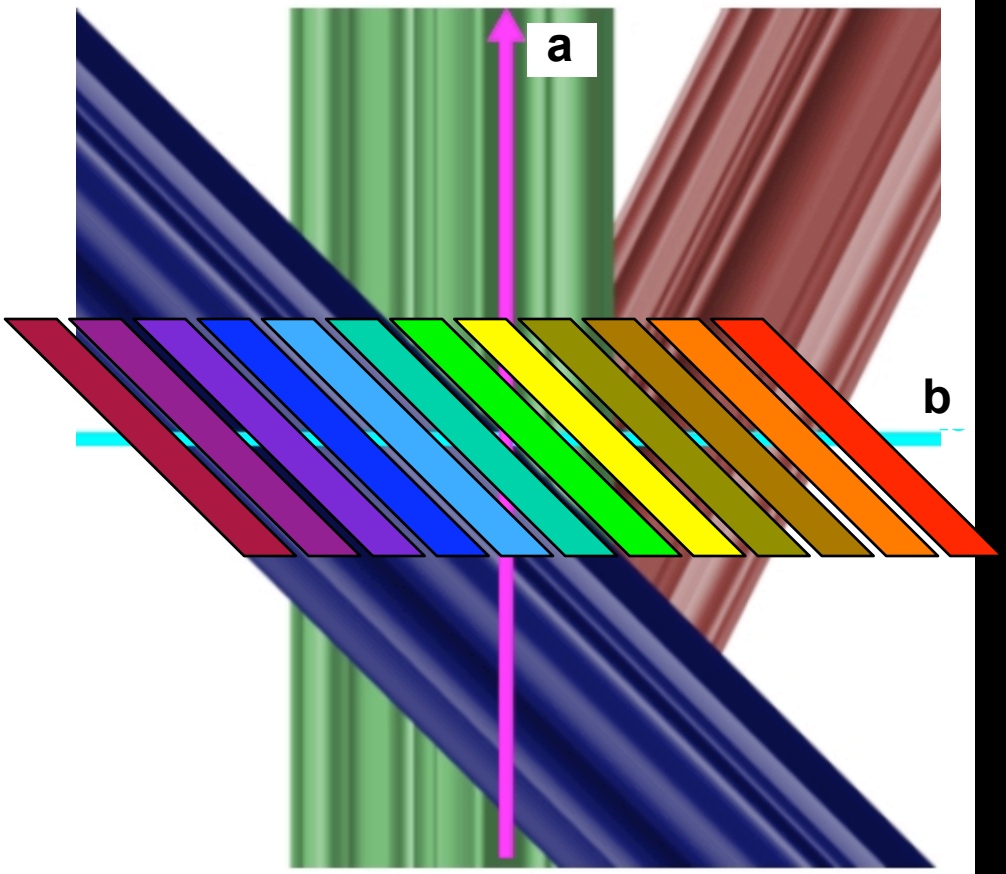


Lens, focused at blue object

flatworld 1D scene

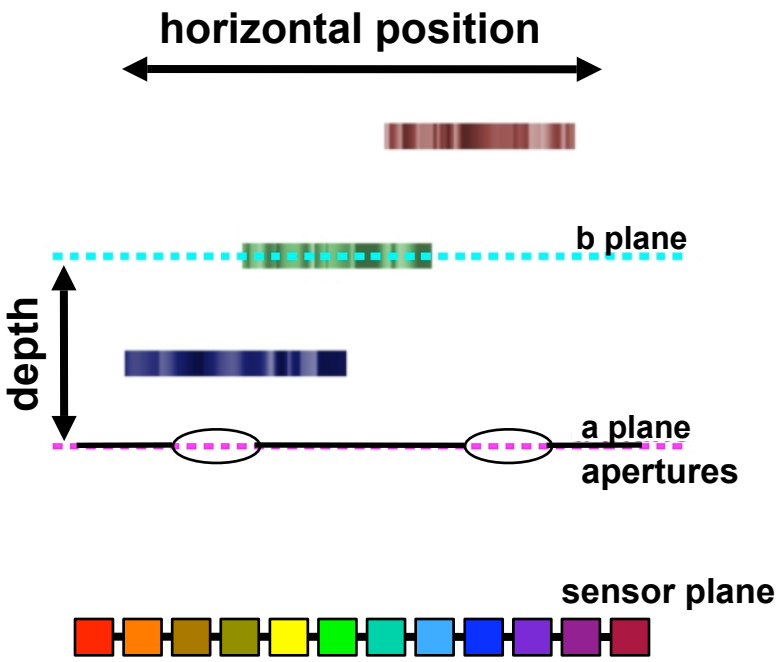


2D lightfield

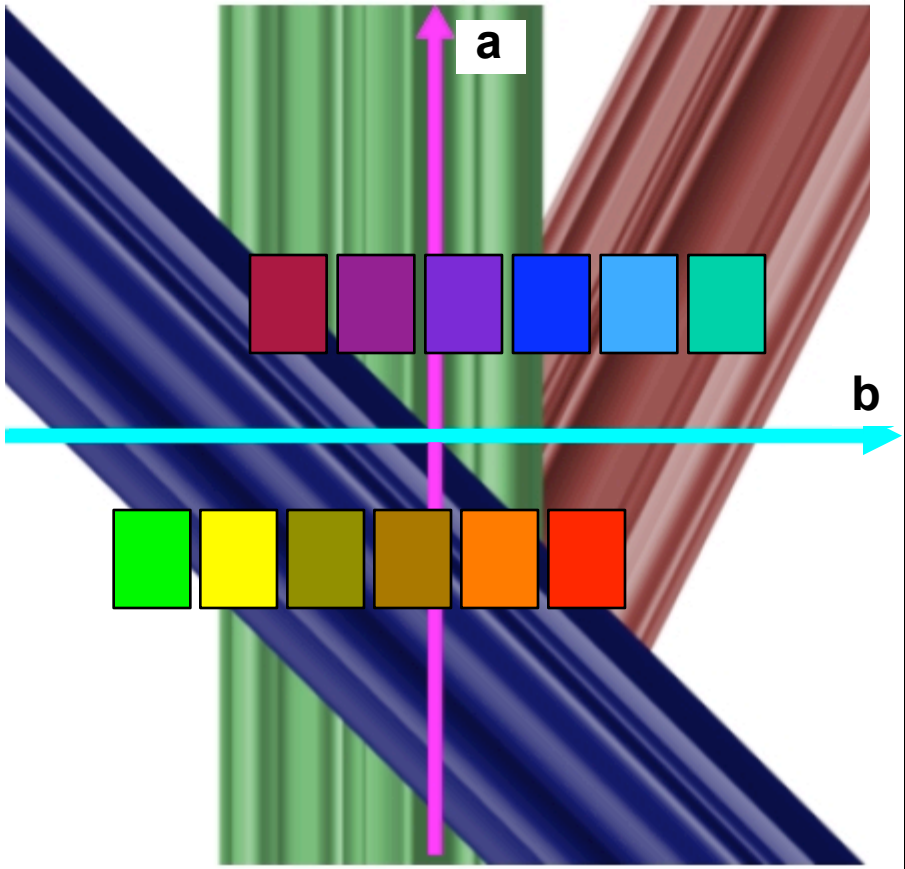


Stereo

flatworld 1D scene

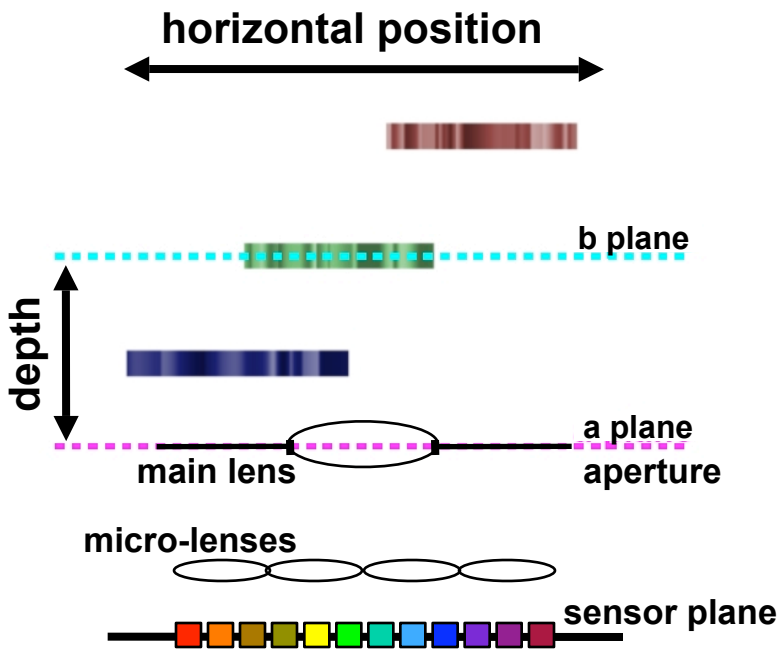


2D lightfield

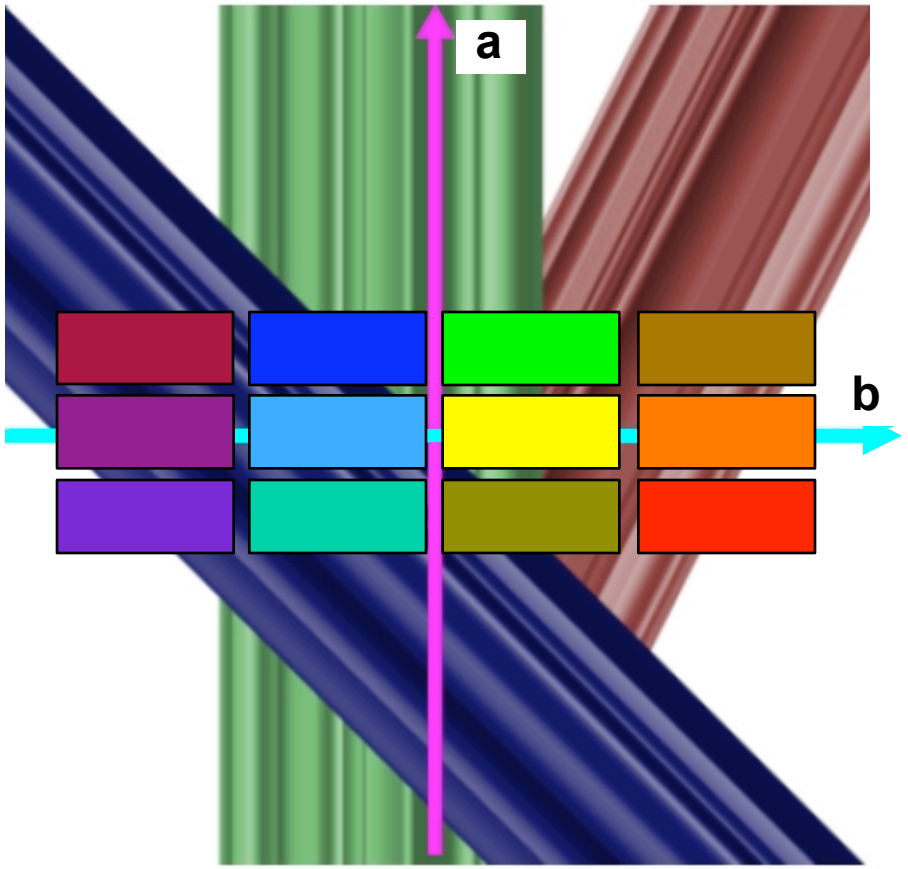


Plenoptic camera

flatworld 1D scene



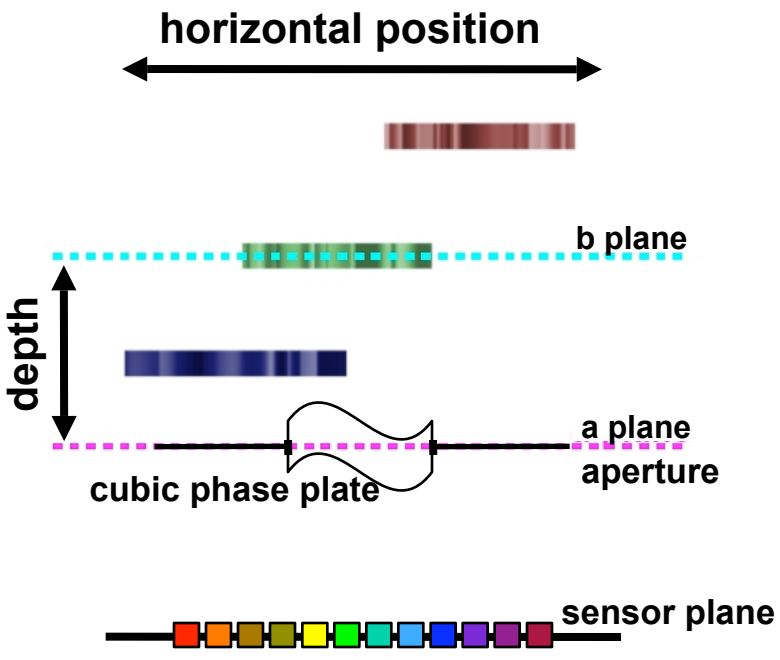
2D lightfield



Adelson and Wang 92, Ng et al 05

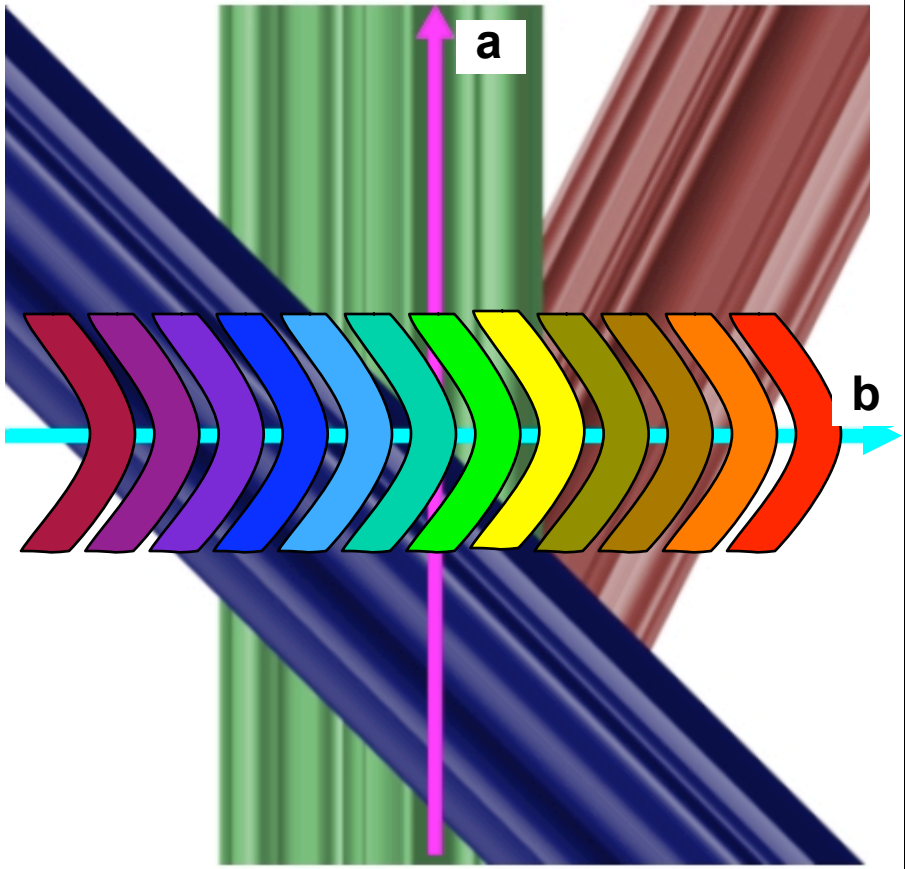
Wavefront coding

flatworld 1D scene



Dowski and Cathey,94

2D lightfield



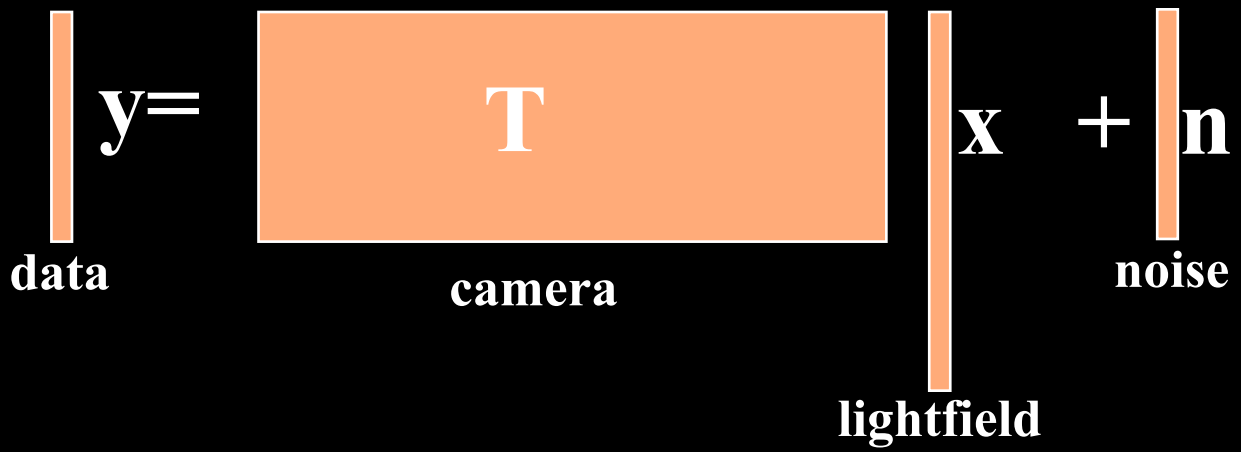
Computational imaging

Camera: Rank deficient projection of a 4D lightfield.

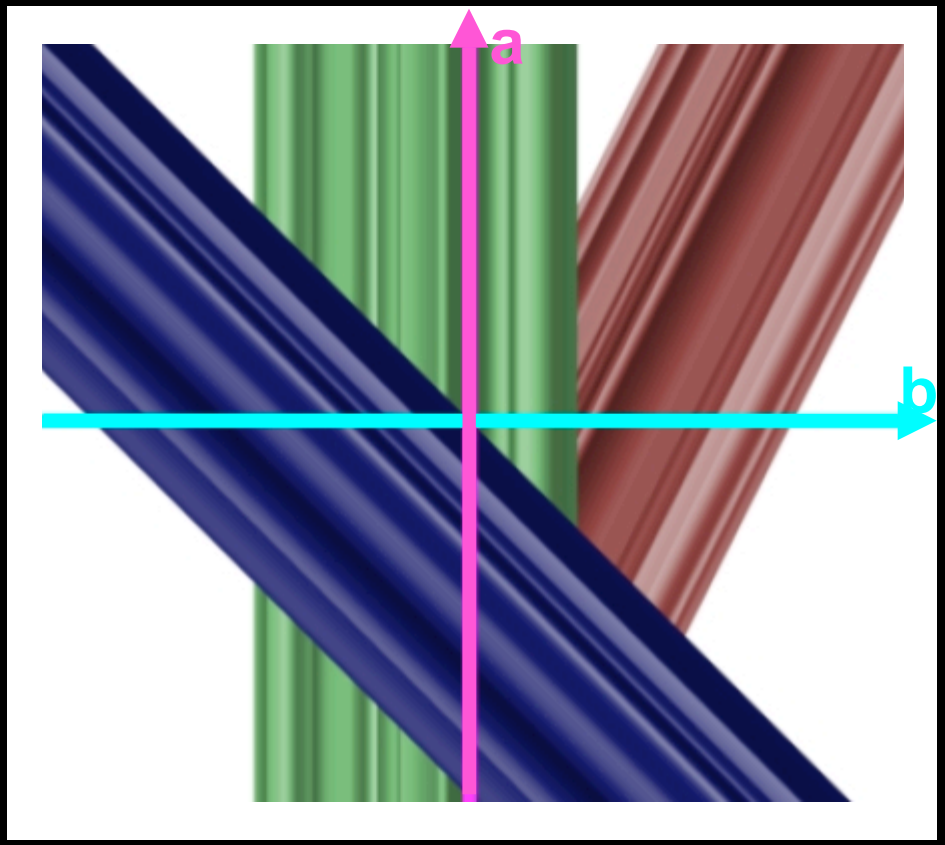
Decoding: ill-posed inversion, need prior on lightfield signals.

Camera evaluation: How well can recover the lightfield from projection?

$$y = Tx + n$$

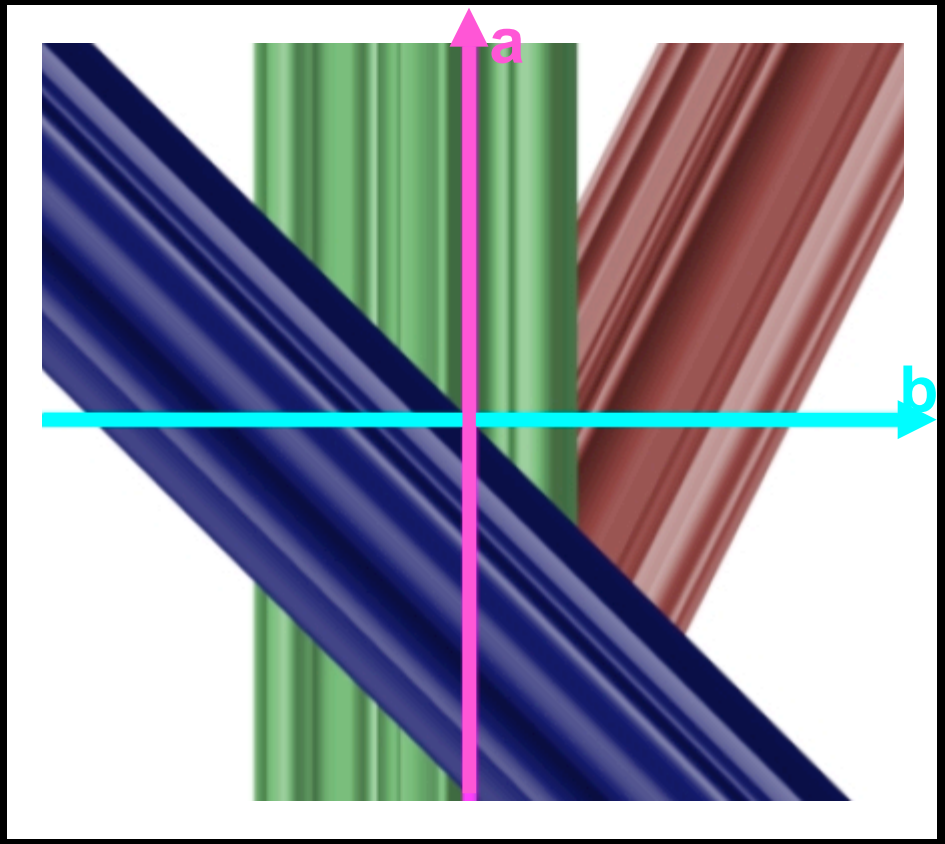


Varying imaging goals by weighted lightfield reconstruction



Varying imaging goals by weighted lightfield reconstruction

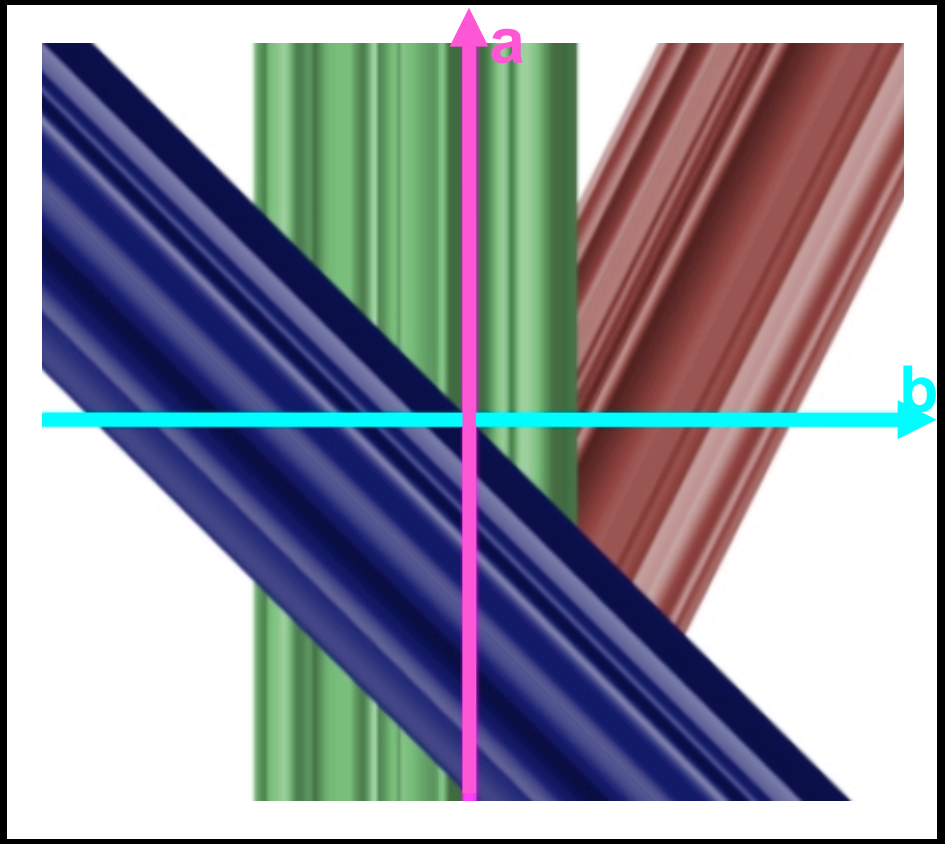
Weigh reconstruction error differently in different light field entries



Varying imaging goals by weighted lightfield reconstruction

- Full light field reconstruction (potentially image&depth)

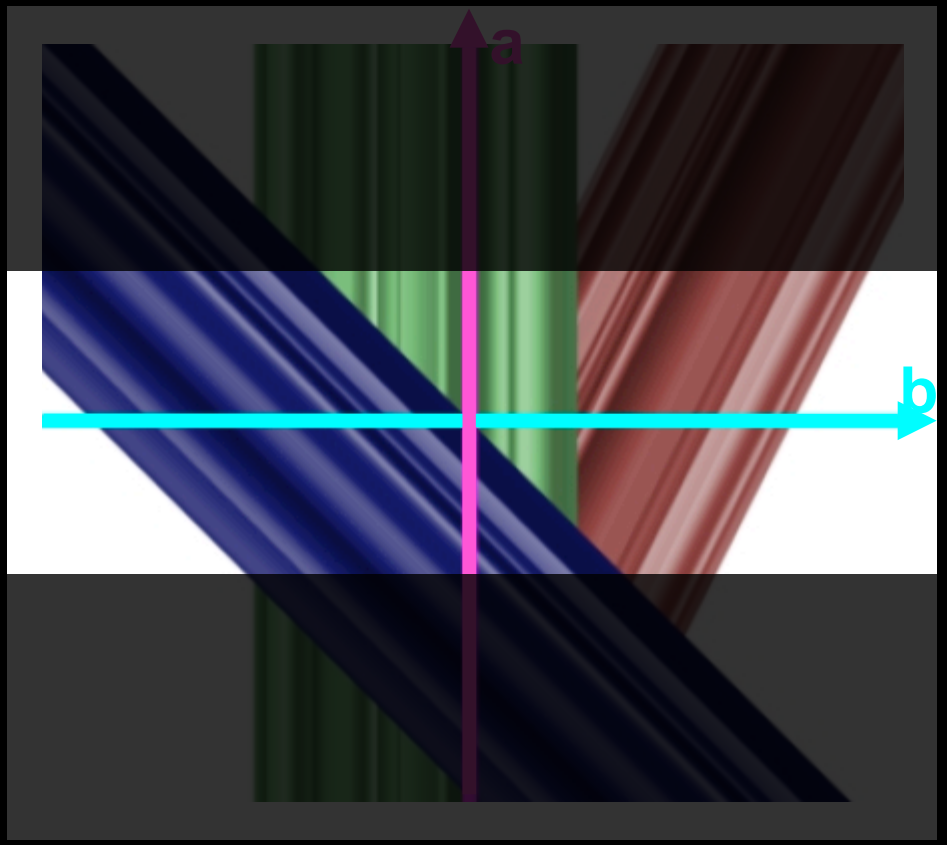
Weigh reconstruction error differently in different light field entries



Varying imaging goals by weighted lightfield reconstruction

- Full light field reconstruction (potentially image&depth)
- Reconstruct a bounded view range

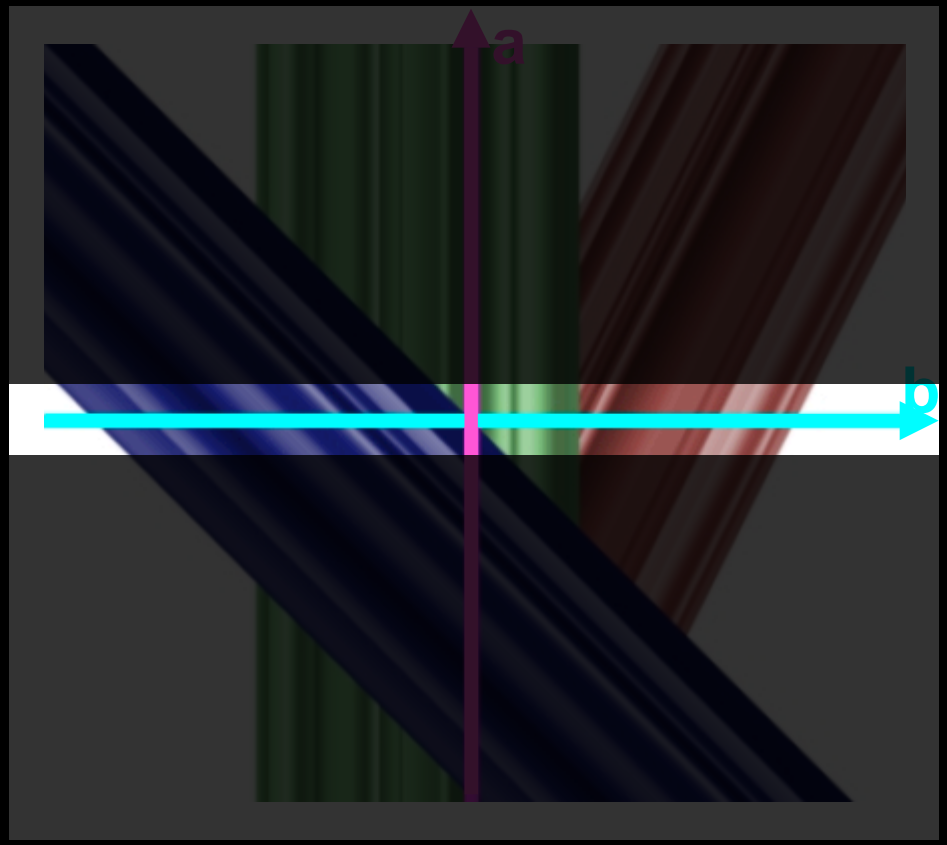
Weigh reconstruction error differently in different light field entries



Varying imaging goals by weighted lightfield reconstruction

- Full light field reconstruction (potentially image&depth)
- Reconstruct a bounded view range
- Single row light field reconstruction (pinhole all focused image)

Weigh reconstruction error differently in different light field entries



Bayesian lightfield imaging – Outline

- Specify lightfield reconstruction goals
  - Full lightfield / Single, all-focus view /…
- Specify lightfield prior
- Imaging with one computational camera
  - Specify camera projection matrix
  - Camera decoding - Bayesian inference
- Comparing computational cameras
  - Specify camera projection matrices
  - Evaluate expected error in lightfield reconstruction

Bayesian lightfield imaging – Outline

- Specify lightfield reconstruction goals
  - Full lightfield / Single, all-focus view /…
- **Specify lightfield prior**
- Imaging with one computational camera
  - Specify camera projection matrix
  - Camera decoding - Bayesian inference
- Comparing computational cameras
  - Specify camera projection matrices
  - Evaluate expected error in lightfield reconstruction

Our light field prior:
a mixture of signals at different slopes

Hidden variable S modeling local slope

Conditioning on slope:

small variance along slope direction

high variance along spatial direction

Our light field prior: a mixture of signals at different slopes

Hidden variable S modeling local slope

Conditioning on slope:

small variance along slope direction

high variance along spatial direction

Light field prior is a mixture of oriented Gaussians (MOG):

$$P(x) = \int_S P(S)P(x \mid S)$$

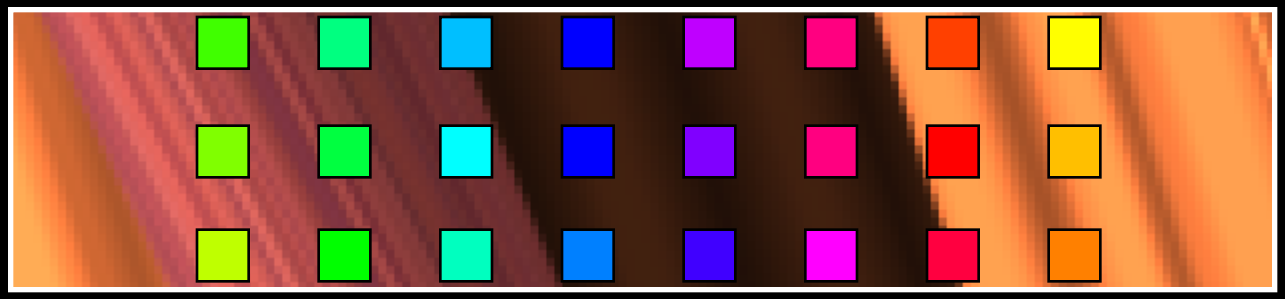
Piecewise smooth prior on slopes

Given slope, lightfield prior is Gaussian and simple

Bayesian lightfield imaging – Outline

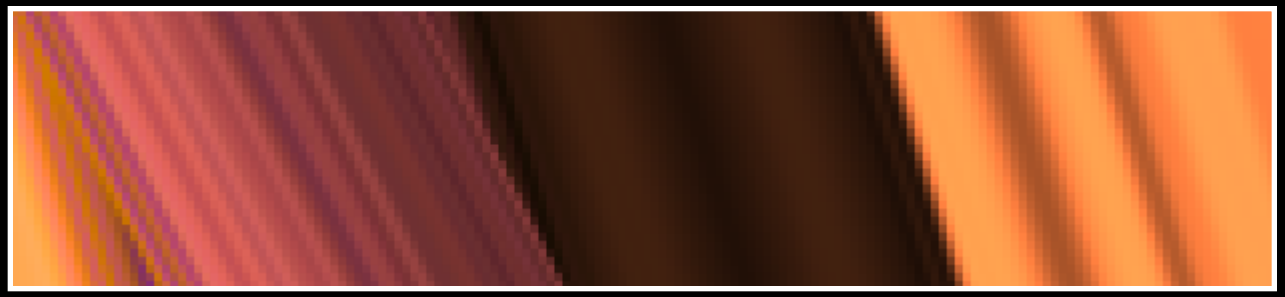
- Specify lightfield reconstruction goals
  - Full lightfield / Single, all-focus view /…
- Specify lightfield prior
- **Imaging with one computational camera**
  - **Specify camera projection matrix**
  - **Camera decoding - Bayesian inference**
- Comparing computational cameras
  - Specify camera projection matrices
  - Evaluate expected error in lightfield reconstruction

Prior effect on reconstruction



Band-limited reconstruction to account for unknown depth

See paper for inference details



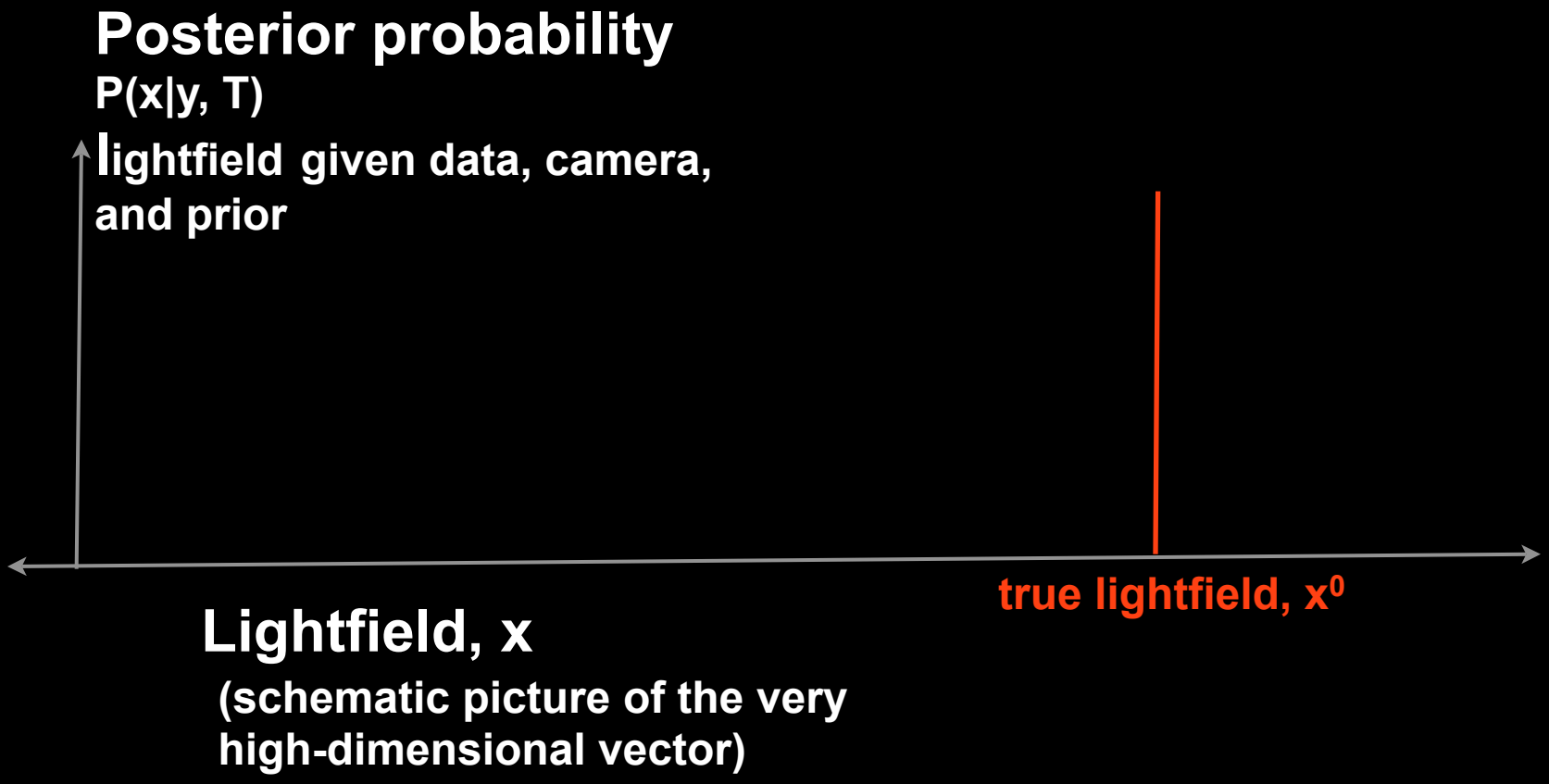
Reconstruction using light field prior

Bayesian lightfield imaging – Outline

- Specify lightfield reconstruction goals
  - Full lightfield / Single, all-focus view /…
- Specify lightfield prior
- Imaging with one computational camera
  - Specify camera projection matrix
  - Camera decoding - Bayesian inference
- Comparing computational cameras
  - Specify camera projection matrices
  - Evaluate expected error in lightfield reconstruction

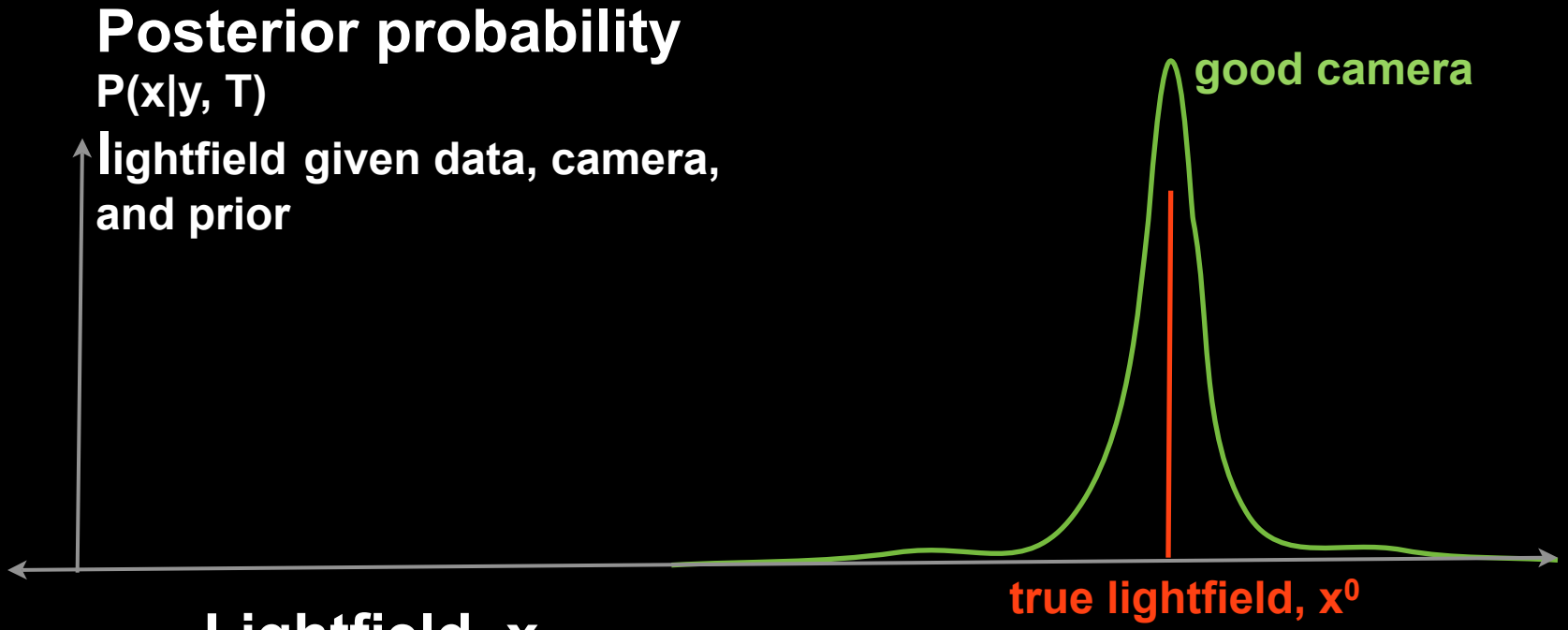
Camera evaluation

Goal: evaluate inherent ambiguity of a camera projection, independent of inference algorithm



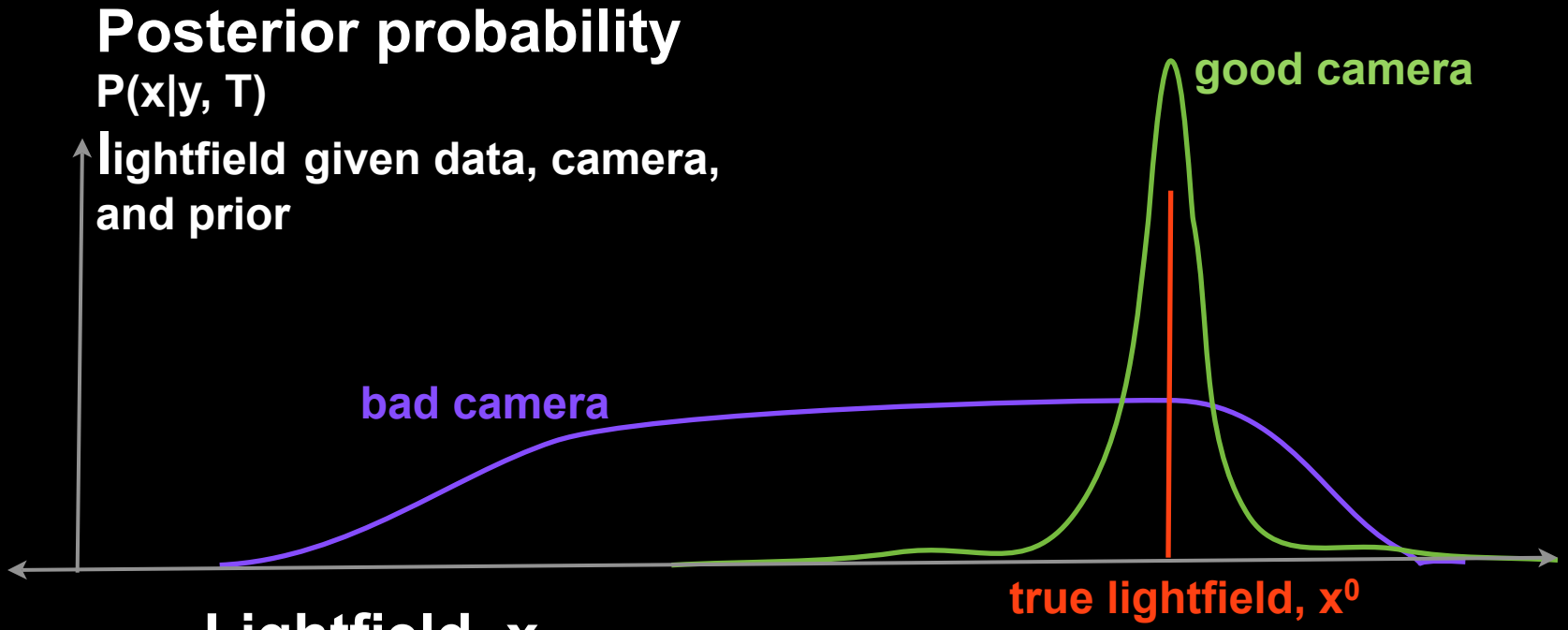
Camera evaluation

Goal: evaluate inherent ambiguity of a camera projection, independent of inference algorithm



Camera evaluation

Goal: evaluate inherent ambiguity of a camera projection, independent of inference algorithm



Camera evaluation function: expected squared error

$$E_{P(x|y;T)}\left[\| x - x^0 \|^2\right] = \int_x P(x \mid y;T) \| x - x^0 \|^2$$

Camera evaluation function: expected squared error

$$E_{P(x|y;T)}\left[\| x - x^0 \|^2\right] = \int_x P(x \mid y;T) \| x - x^0 \|^2$$

With our mixture model prior, conditioned on the lightfield slopes S, everything is Gaussian and analytic. So let's write the posterior as:

$$P(x \mid y;T) = \int_S P(S \mid y;T) P(x \mid y, S;T)$$

Camera evaluation function: expected squared error

$$E_{P(x|y;T)}\left[\| x - x^0 \|^2\right] = \int_x P(x \mid y;T) \| x - x^0 \|^2$$

With our mixture model prior, conditioned on the lightfield slopes S, everything is Gaussian and analytic. So let's write the posterior as:

$$P(x \mid y;T) = \int_S P(S \mid y;T) P(x \mid y, S;T)$$

Then our expected squared error becomes an integral over all slope fields:

$$E_{P(x|y;T)}\left[\| x - x^0 \|^2\right] = \int_S P(S \mid y;T) E_{P(x|y,S;T)}\left[\| x - x^0 \|^2\right]$$

Camera evaluation function: expected squared error

$$E_{P(x|y;T)}\left[\| x - x^0 \|^2\right] = \int_x P(x \mid y;T) \| x - x^0 \|^2$$

With our mixture model prior, conditioned on the lightfield slopes S, everything is Gaussian and analytic. So let's write the posterior as:

$$P(x \mid y;T) = \int_S P(S \mid y;T) P(x \mid y, S;T)$$

Then our expected squared error becomes an integral over all slope fields:

$$E_{P(x|y;T)}\left[\| x - x^0 \|^2\right] = \int_S P(S \mid y;T) E_{P(x|y,S;T)}\left[\| x - x^0 \|^2\right]$$

Approximate by Monte Carlo sampling near the true slope field:

$$E_{P(x|y;T)}\left[\| x - x^0 \|^2\right] \approx \sum_{S_i} P(S_i \mid y;T) E_{P(x|y,S_i;T)}\left[\| x - x^0 \|^2\right]$$

Bayesian camera evaluation tool

Input parameters:

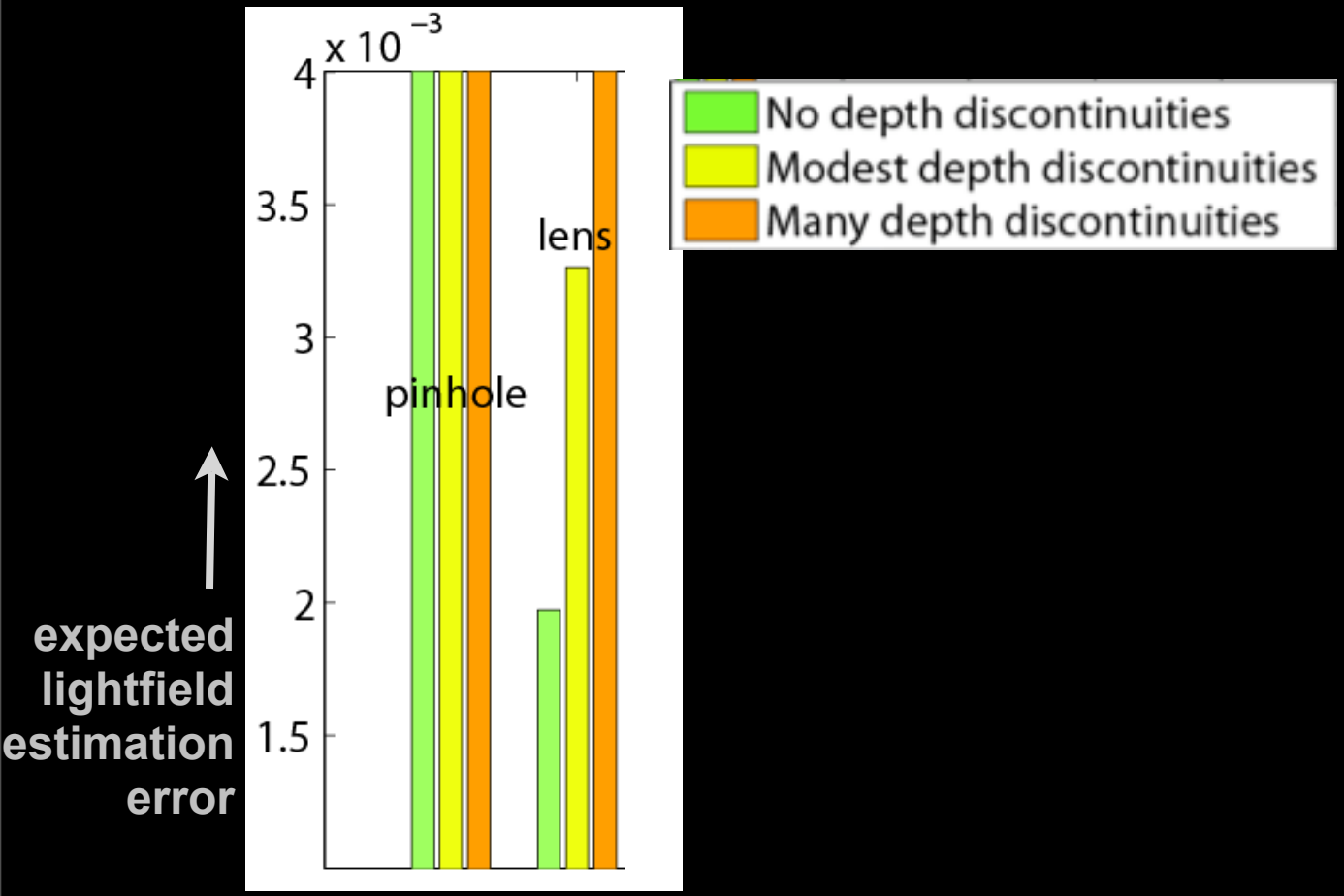
- Reconstruction goals (weight on light field entries)
- Camera matrix
- Noise level
- Spatial and depth resolution

Output: expected reconstruction error

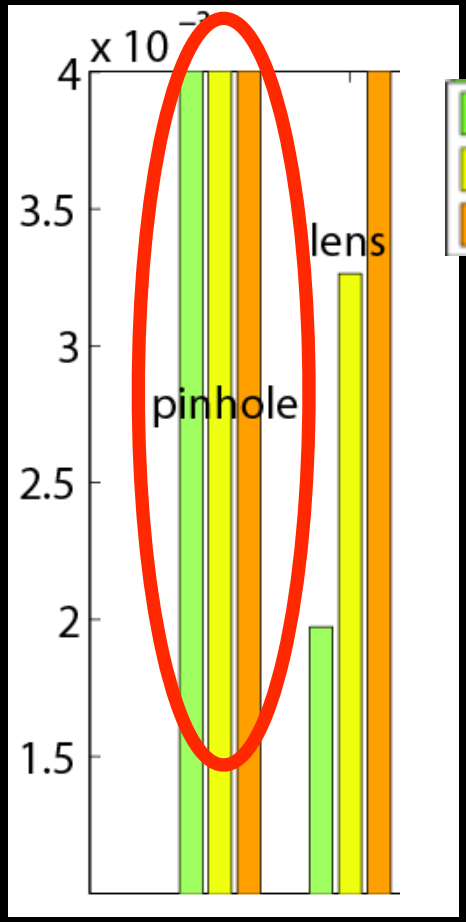
Matlab software online:

people.csail.mit.edu/alevin/papers/lightfields-Code-Levin-Freeman-Durand-08.zip

1D camera evaluation- full light field reconstruction



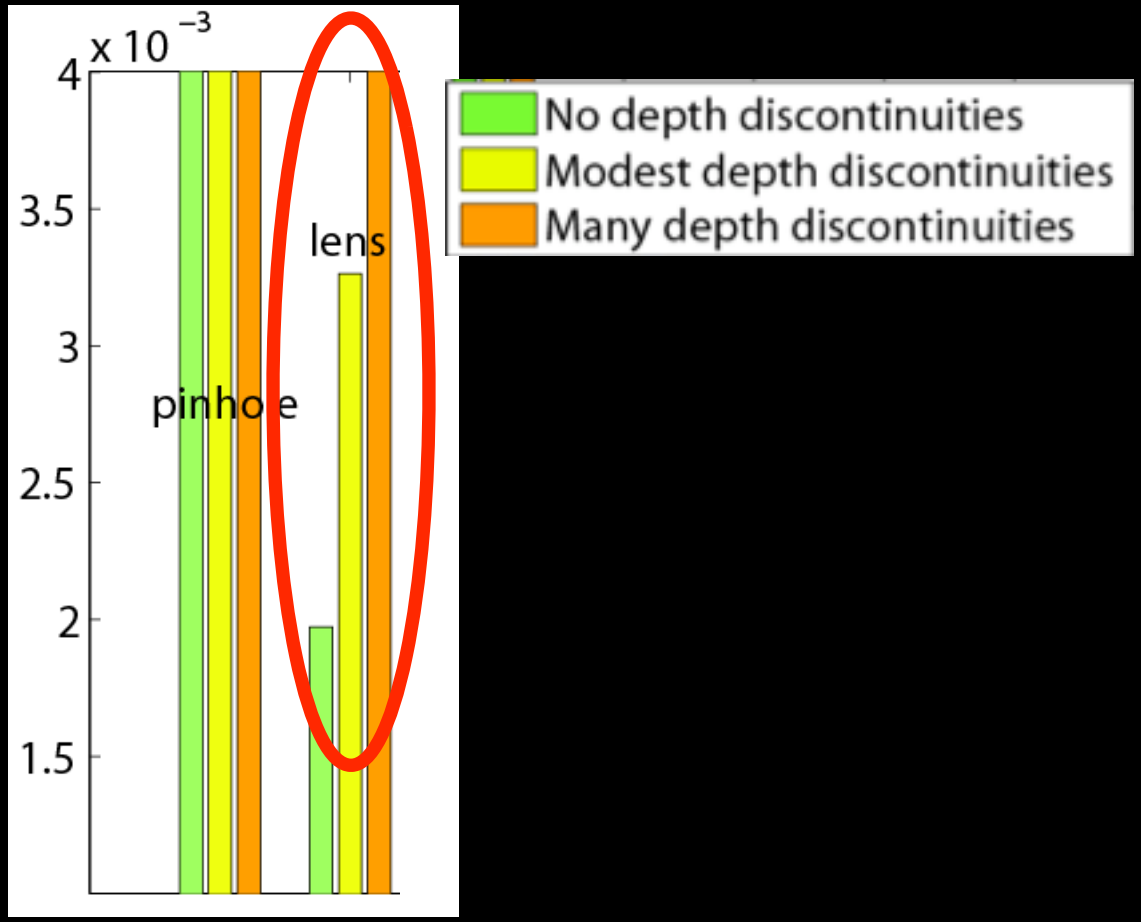
1D camera evaluation- full light field reconstruction



Observation:

As expected, a pinhole camera doesn't estimate the lightfield well

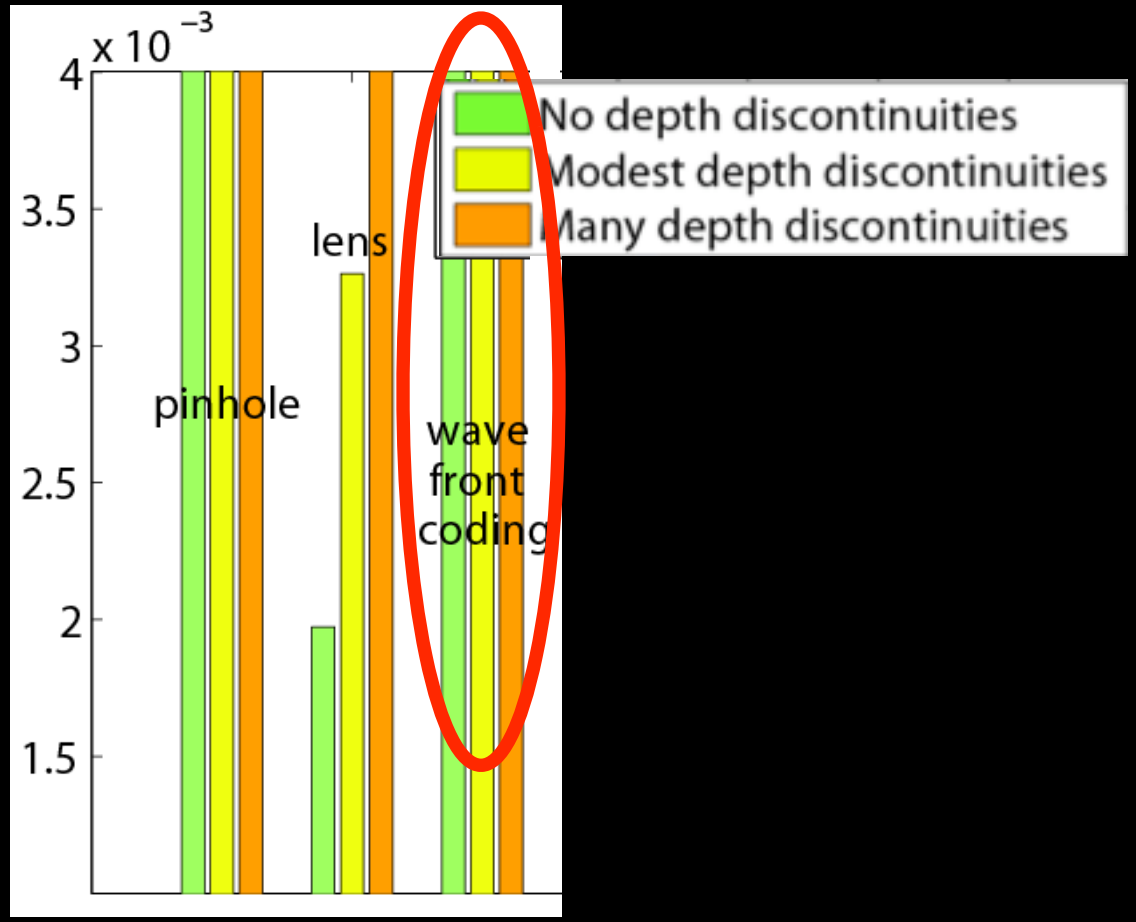
1D camera evaluation- full light field reconstruction



Observation:

When depth variation is limited, some depth from defocus exist in a single monocular view from a standard lens

1D camera evaluation- full light field reconstruction

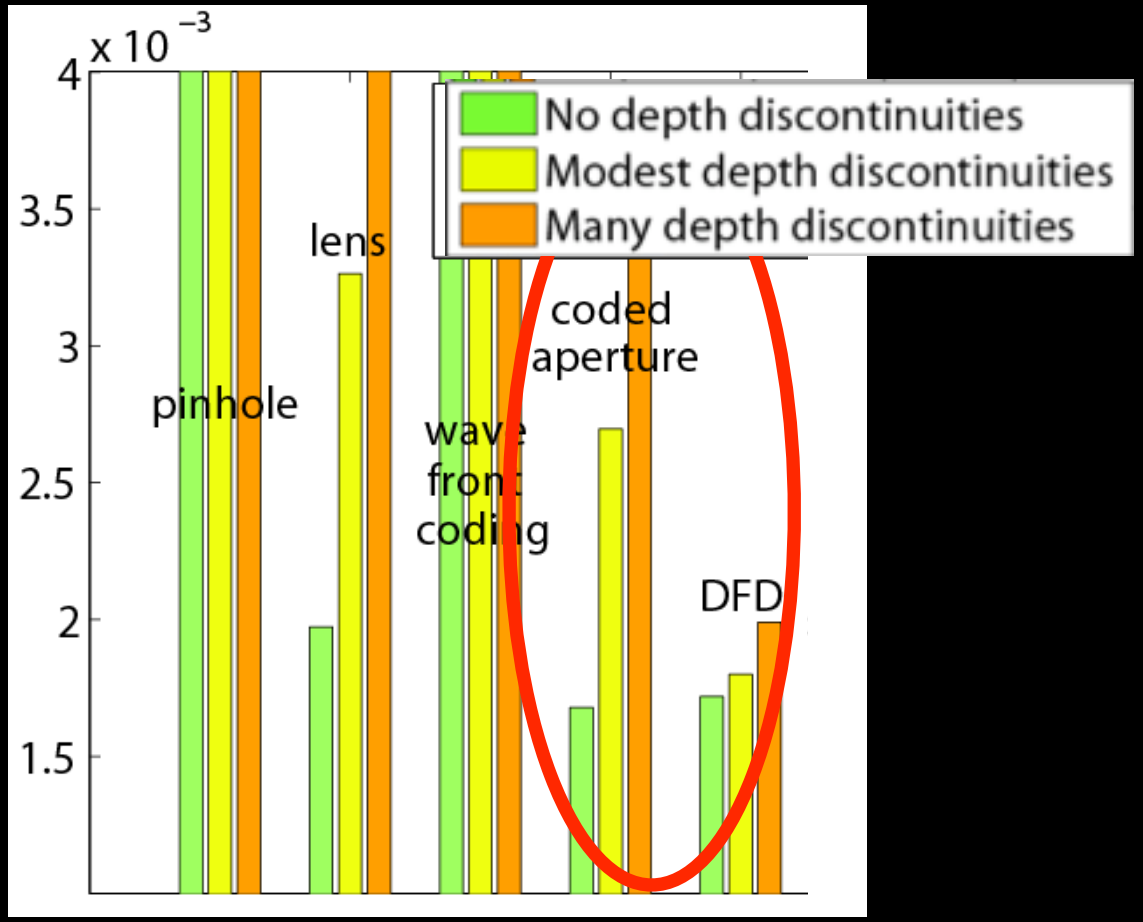


expected lightfield estimation error

Observation:

Wavefront coding, not designed to estimate the lightfield, doesn't.

1D camera evaluation- full light field reconstruction

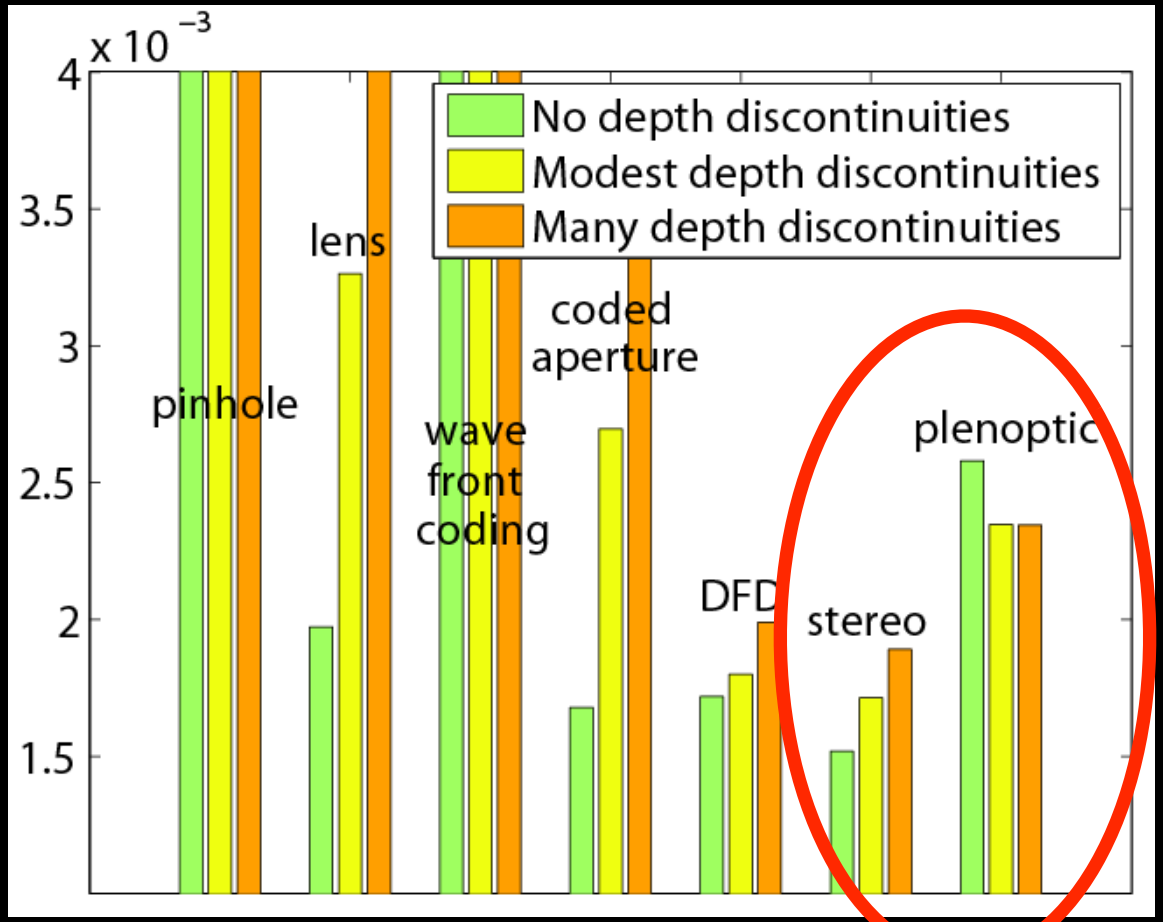


expected lightfield estimation error

Observation:

Depth-from-defocus (DFD) outperforms the coded aperture at these settings

1D camera evaluation- full light field reconstruction

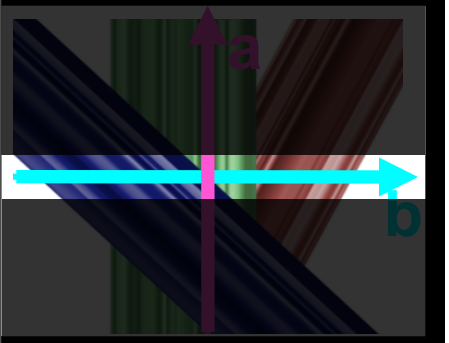


expected lightfield estimation error

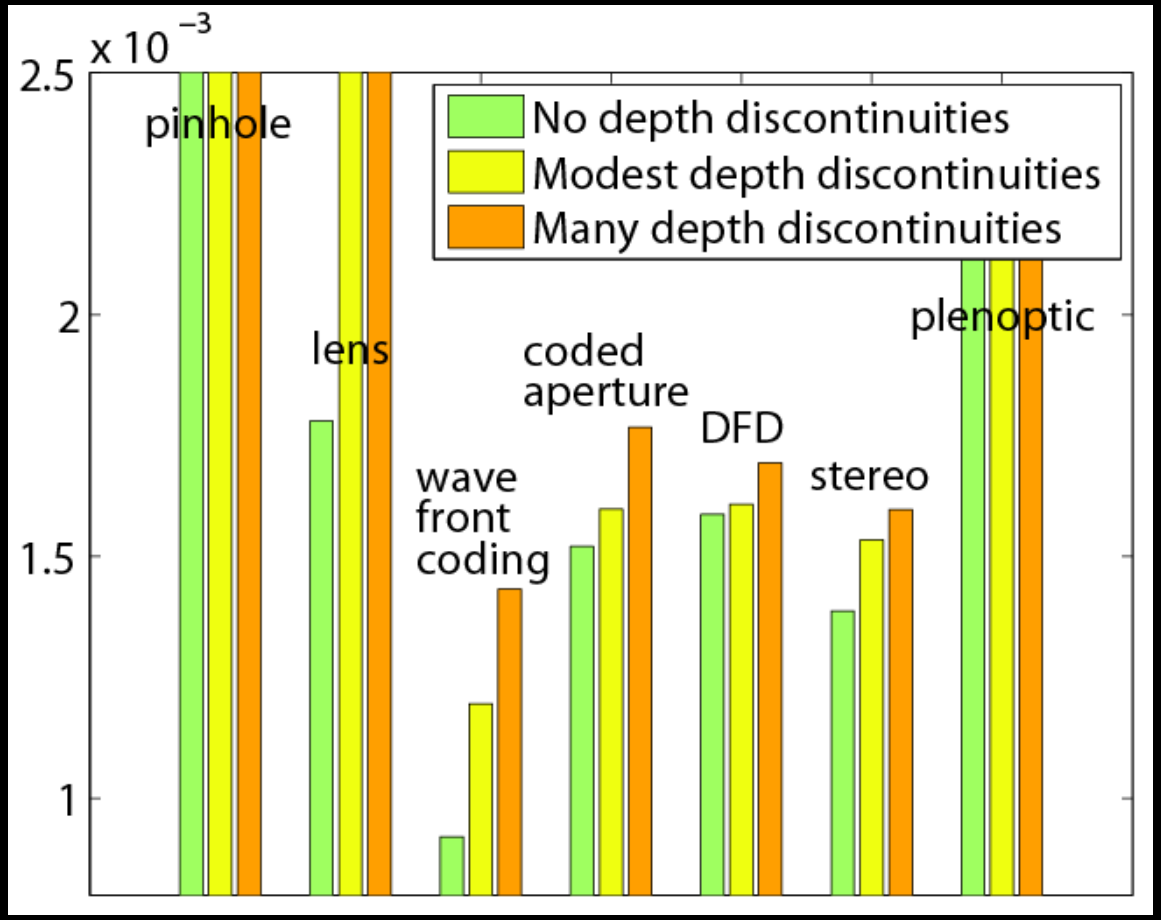
Observation: Stereo error is less than Plenoptic

Since depth variation is smaller than texture variation, no need to sacrifice so much spatial resolution to capture directional information

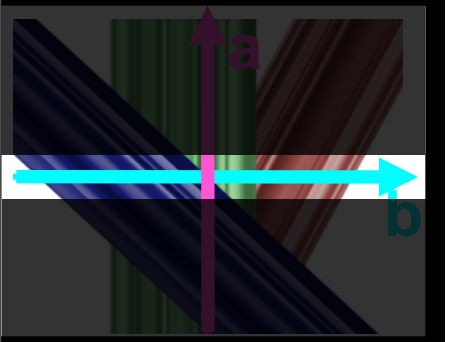
1D camera evaluation- single row reconstruction

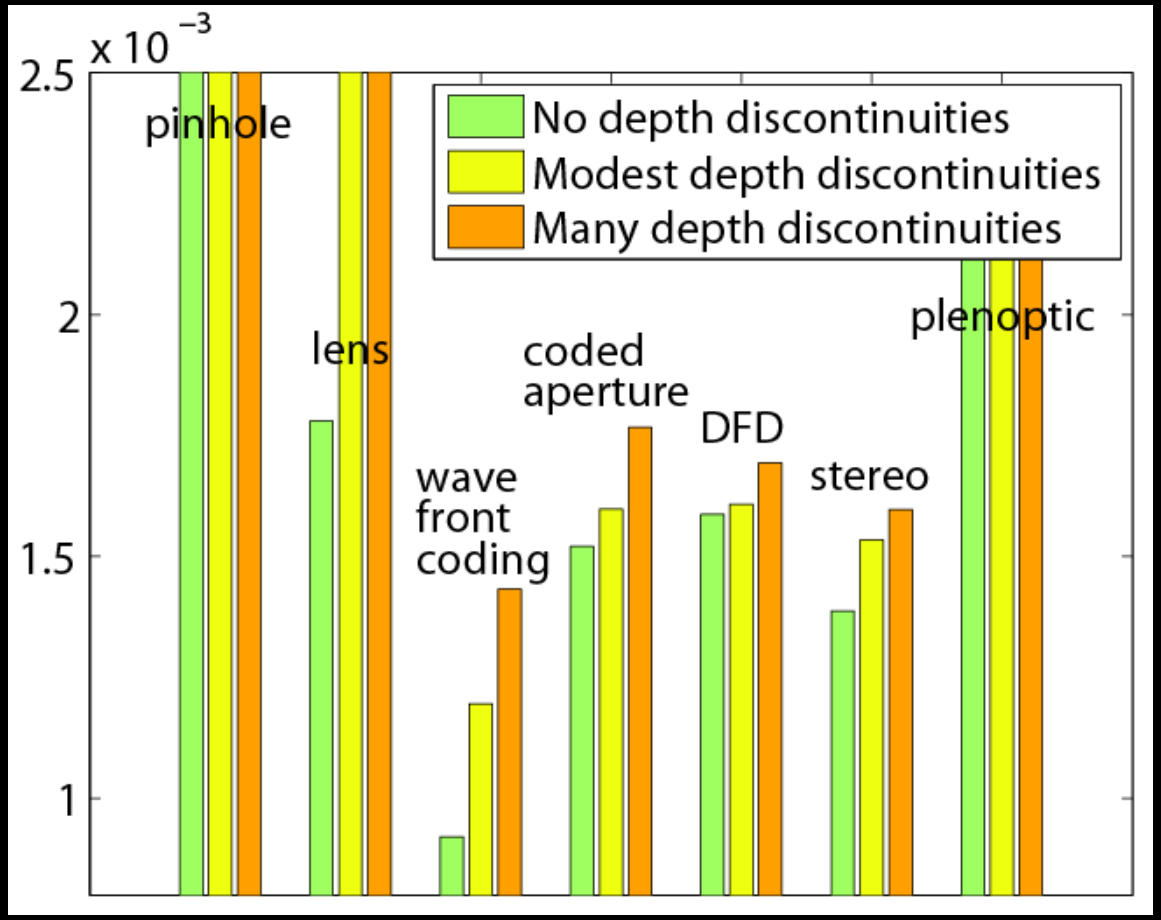


expected lightfield estimation error



1D camera evaluation- single row reconstruction

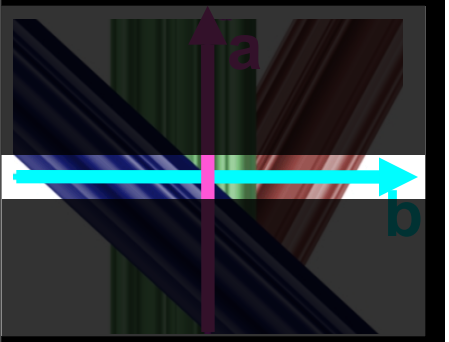


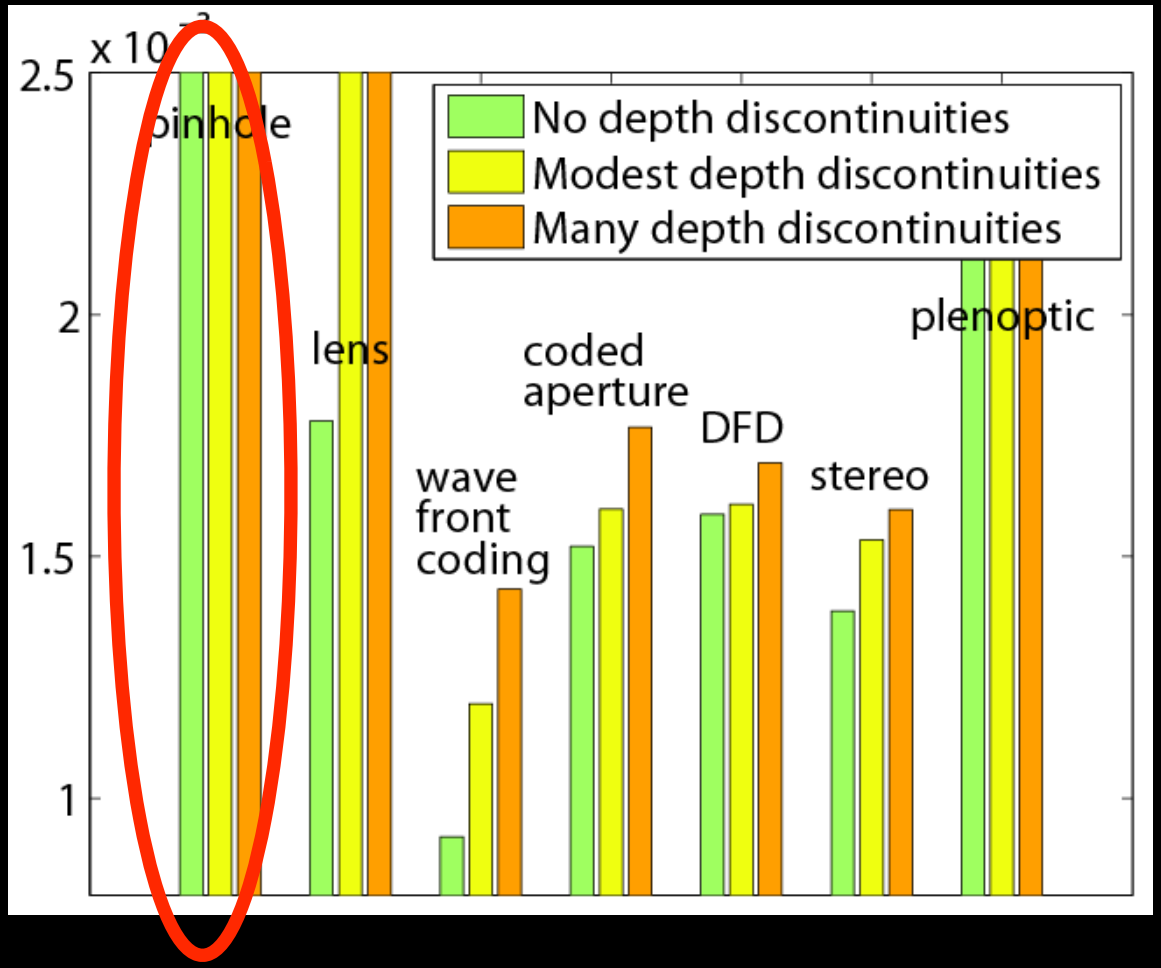


expected lightfield estimation error

Observations:

1D camera evaluation- single row reconstruction



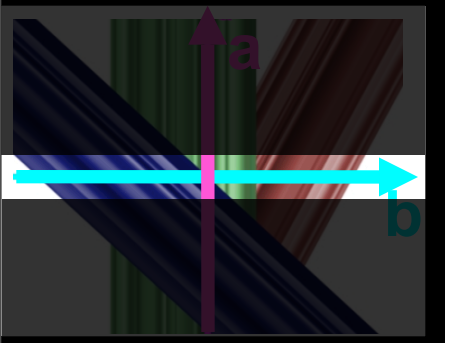


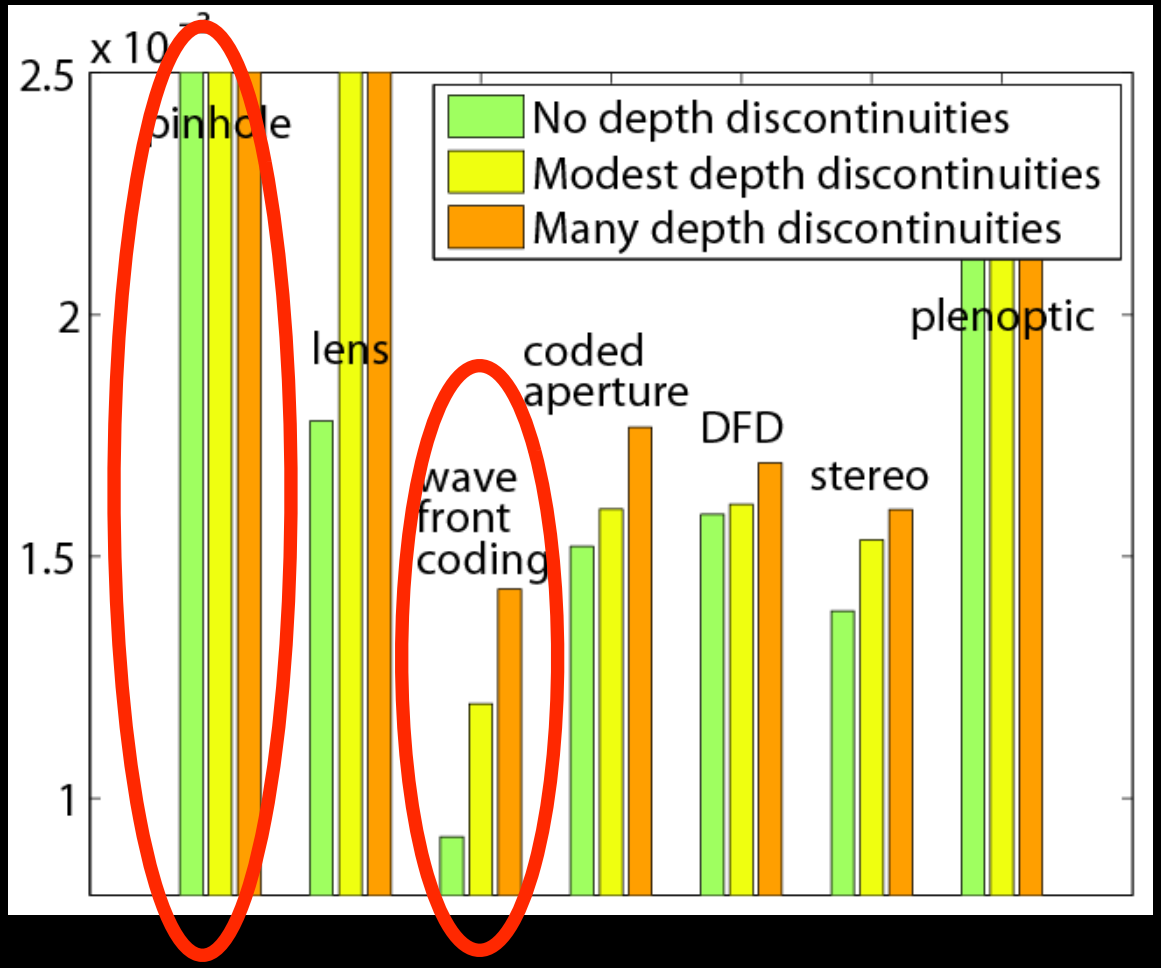
expected lightfield estimation error

Observations:

Pinhole camera- poor estimation due to noise

1D camera evaluation- single row reconstruction





expected lightfield estimation error

Observations:

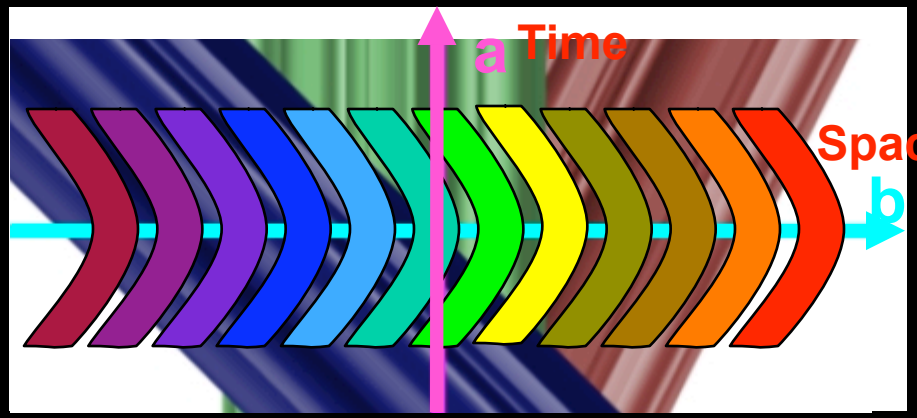
Pinhole camera- poor estimation due to noise

Wavefront coding- no depth information, but accurate reconst for a single view

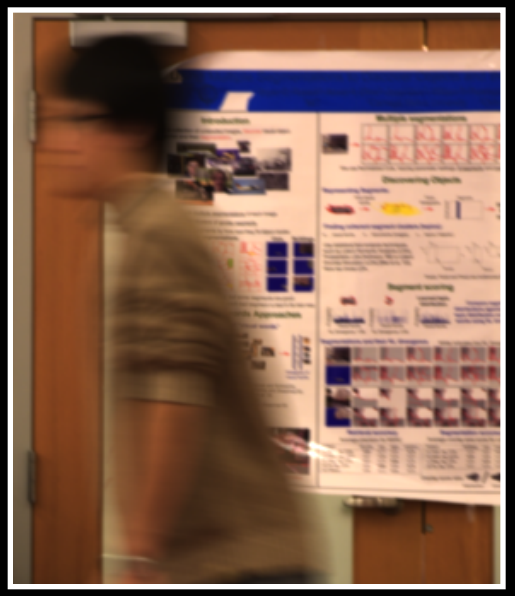
Application: motion invariant photography

Depth invariant integration

Motion invariant integration



SIGGRAPH 2008, Levin et al.



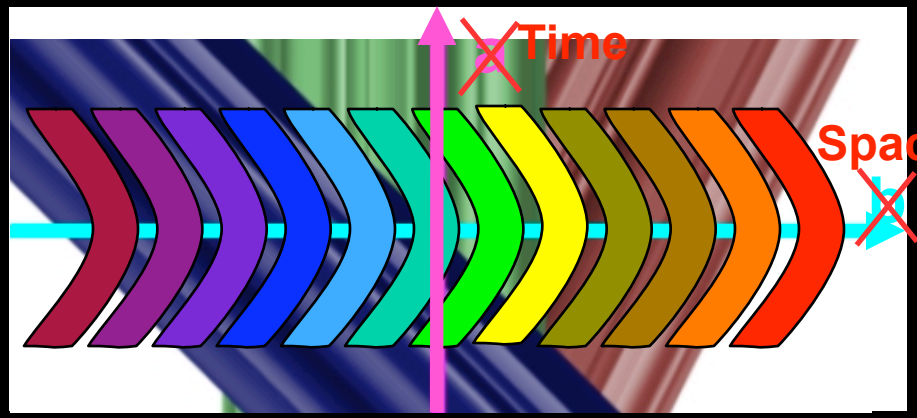
Static camera

speed-invariant blur allows non-blind deconvolution

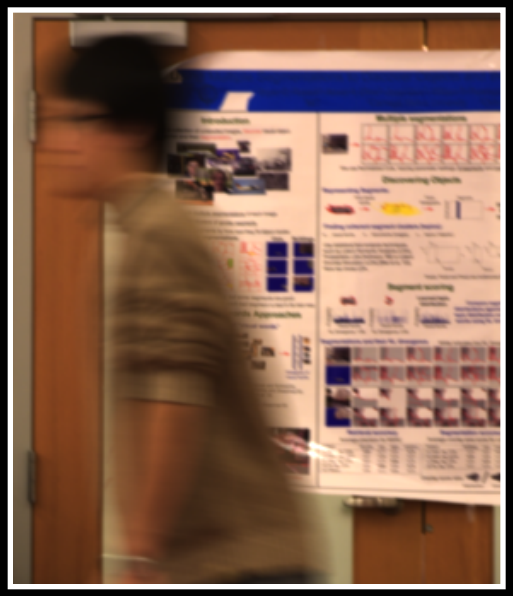
Application: motion invariant photography

Depth invariant integration

Motion invariant integration



SIGGRAPH 2008, Levin et al.



Static camera

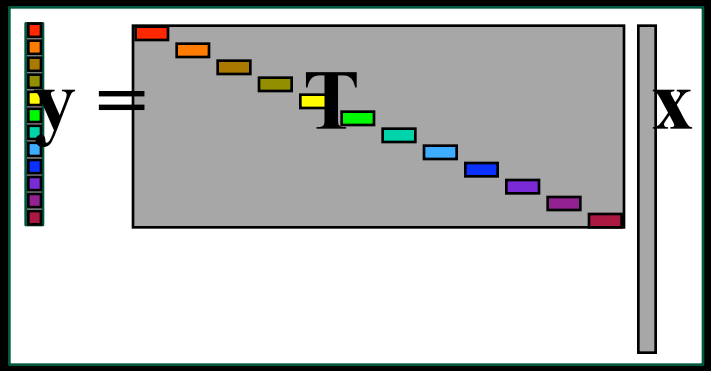
motion invariant input

output after deblurring

speed-invariant blur allows non-blind deconvolution

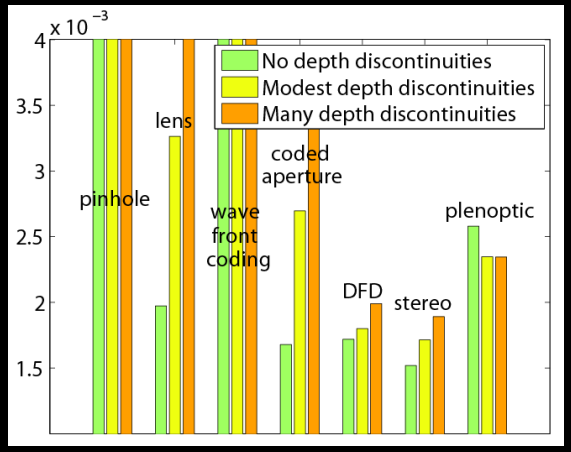
Summary: Bayesian lightfield imaging

- Model imaging as linear light field projection



- New prior on light field signals

- Camera decoding expressed as a Bayesian inference problem



- Framework and software for comparison across camera configurations, by evaluating uncertainty in light field reconstruction

- Principled novel camera design