

Depth of Field. Light field cameras

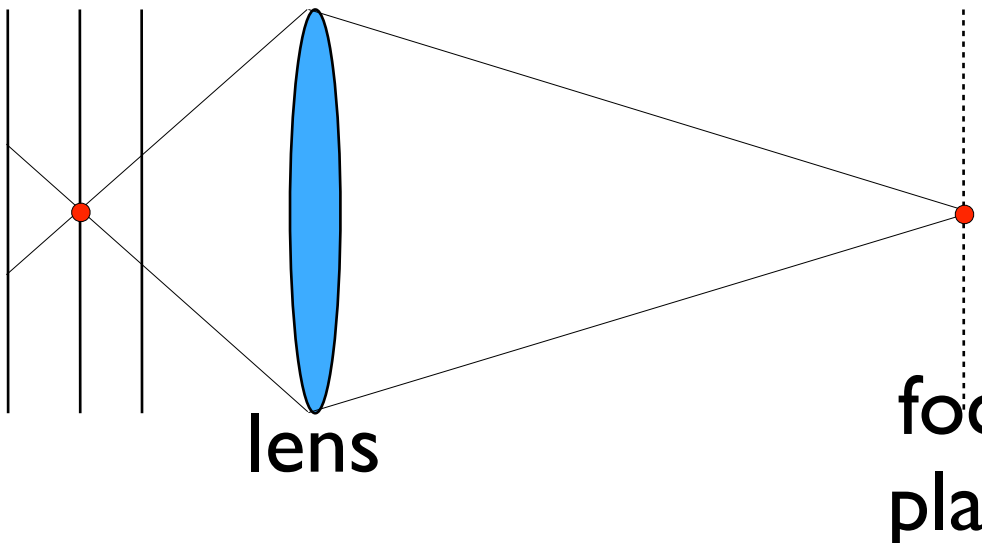
Frédo Durand

MIT EECS

Light fields and plenoptic cameras

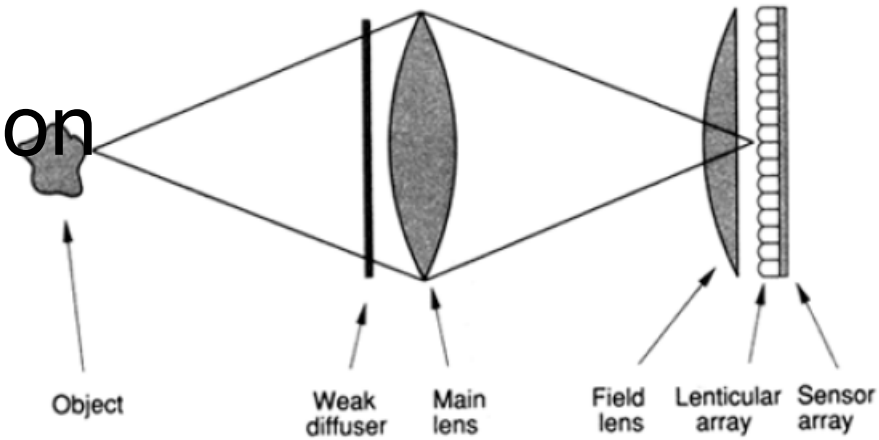
Focusing and defocus

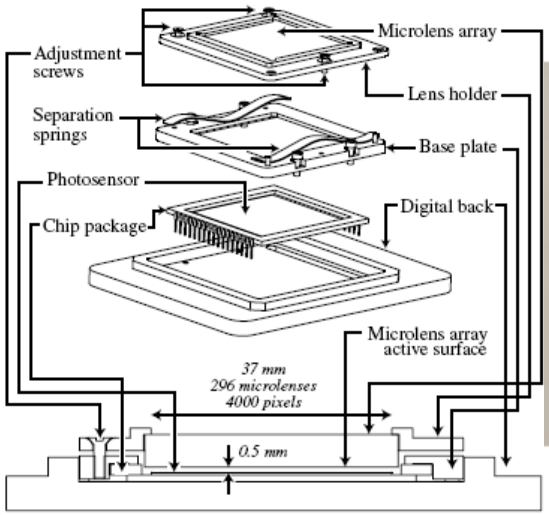
- Each pixel is the integral of many light rays across the aperture
- Which rays are integrated at which pixel changes the focus
- Idea: capture rays, not pixels



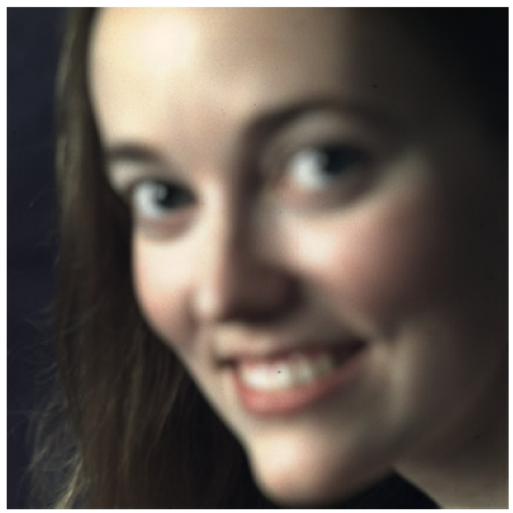
Plenoptic/light field cameras

- Different goals
  - reproduce 3D impression
  - enable refocusing
  - do computer vision





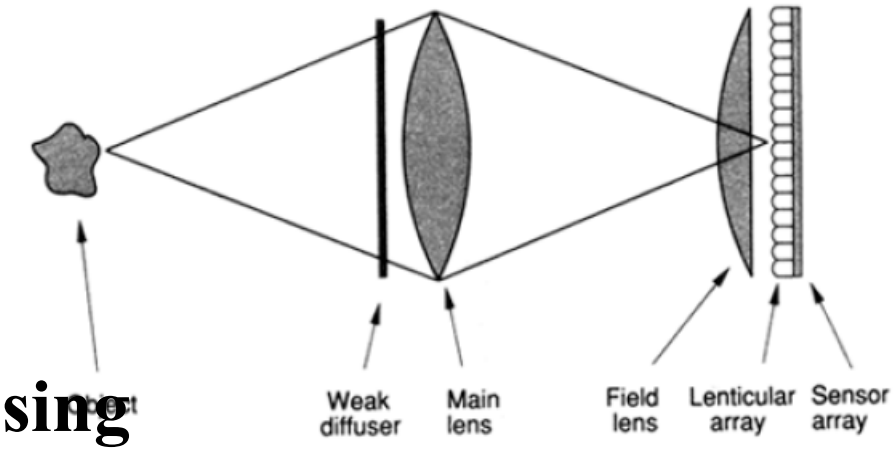


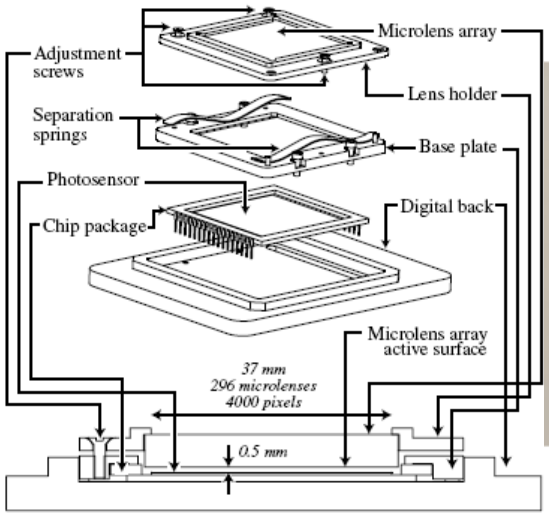




Plenoptic/light field cameras

- **Lipmann 1908**
  - "Window to the world"
- **Adelson and Wang, 1992**
  - Depth computation
- **Revisited by Ng et al. for refocusing**
  - turned into Lytro







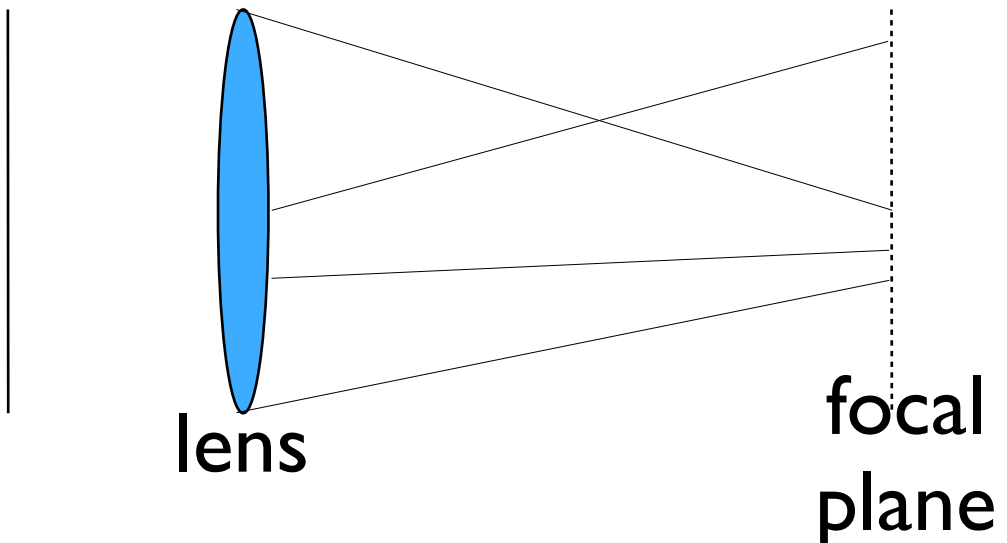




Light fields

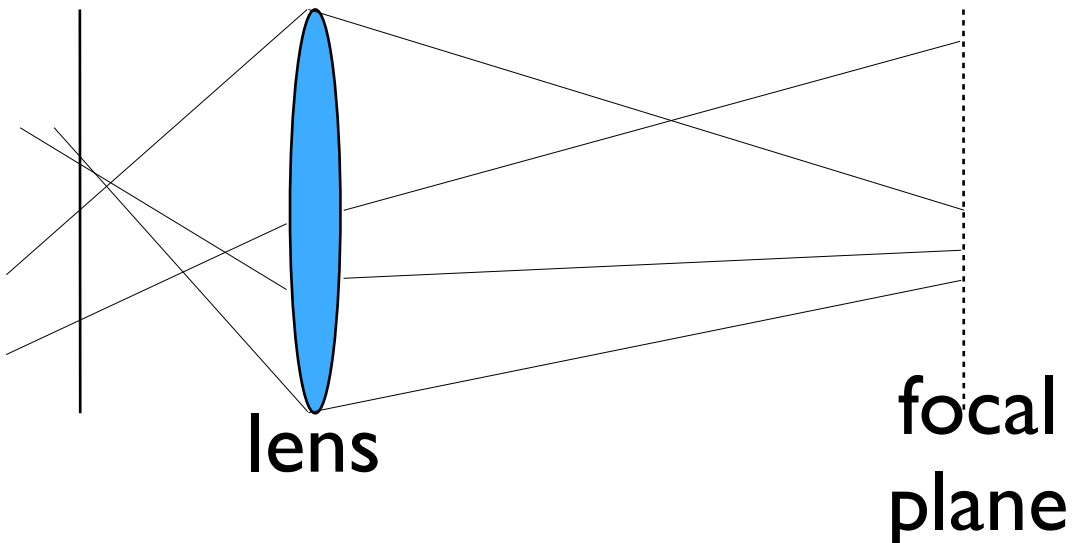
For a camera

- Capture all rays reaching the lens
- 4D:
  - 2D for lens aperture
  - 2D for focal plane (or any other plane)



For a camera: conjugate

- Same on the image side
- 4D:
  - 2D for lens aperture
  - 2D for sensor



Two-plane parameterization

- **Line parameterized by intersection with 2 planes**
  - Careful, there are different "isotopes" of such parameterization (slightly different meaning of stuv)

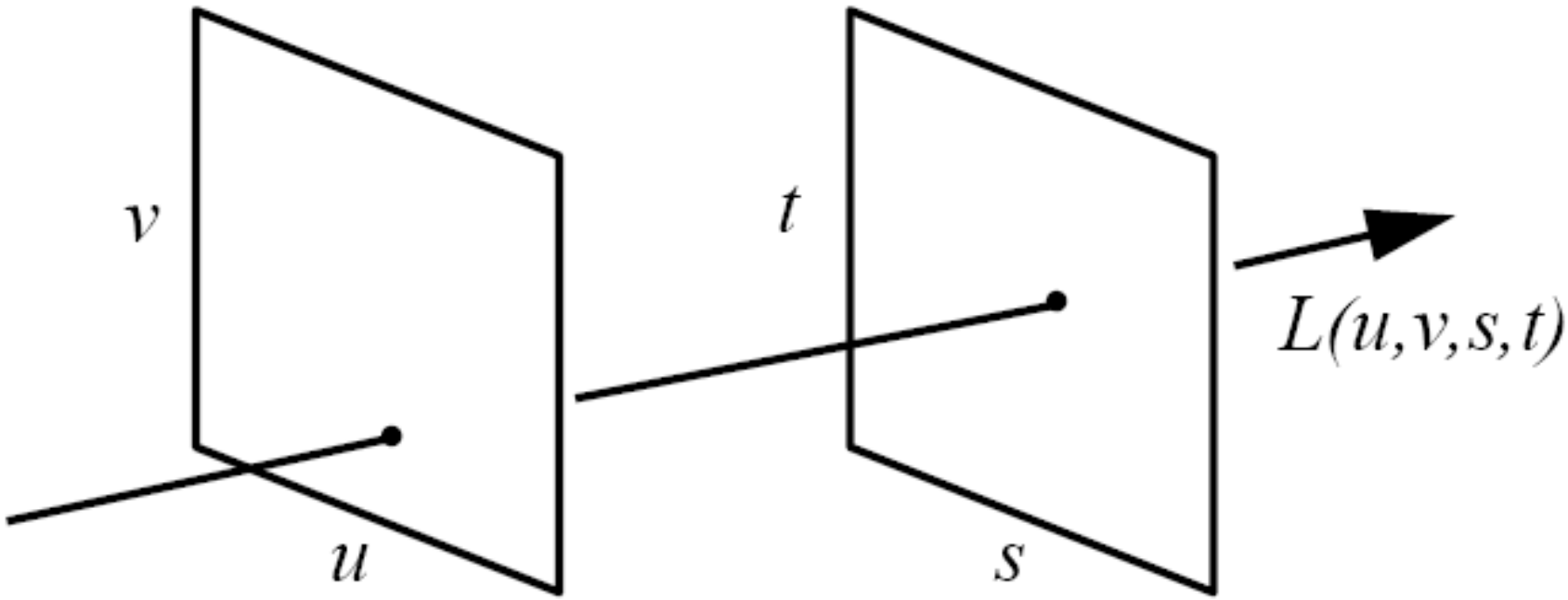
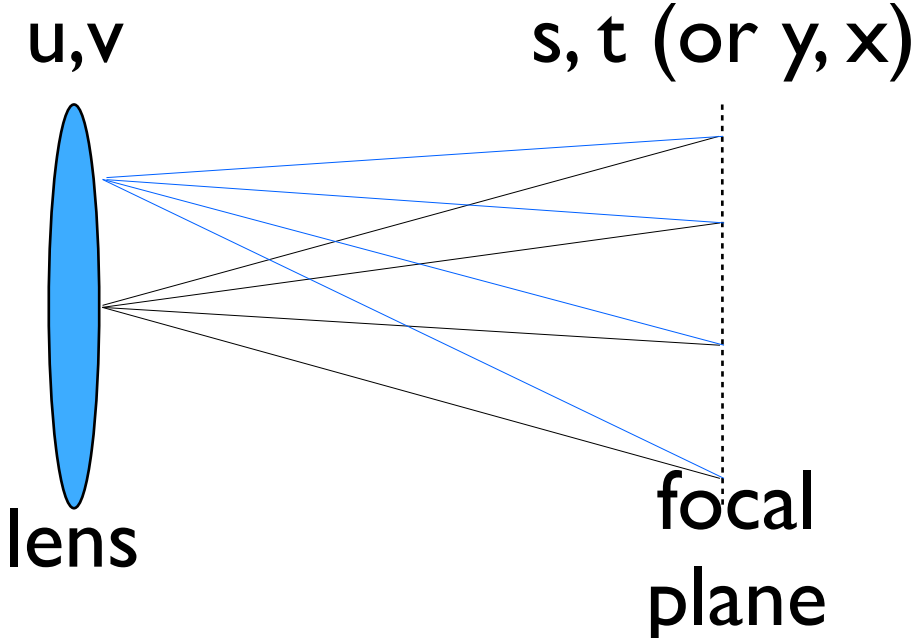
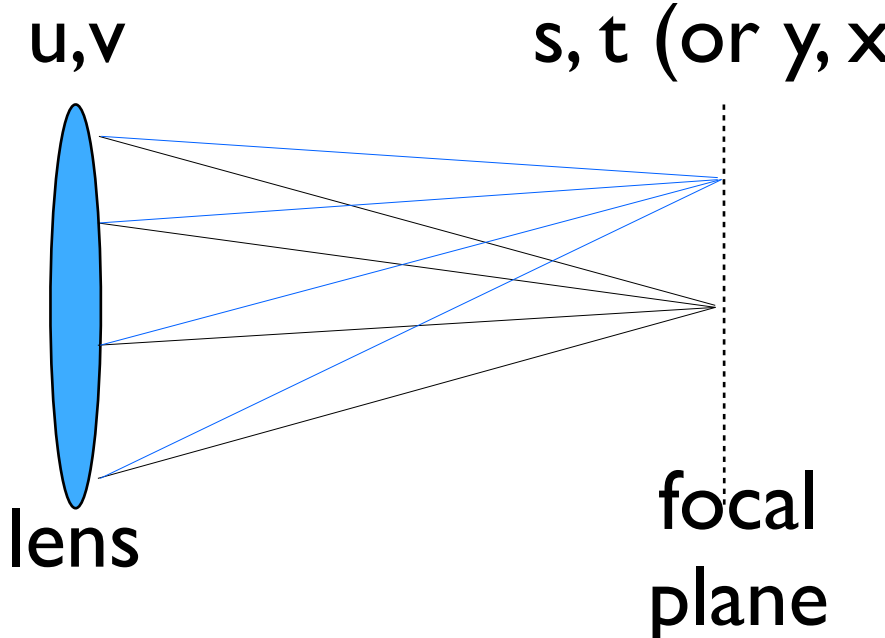


Figure 1: The light slab representation.

Understanding 4D

- 2D slices
  - constant u, v: regular pinhole images from different aperture points
  - constant x, y: aperture view
    - how much the color of a point on the focal plane varies with angle (but depends on





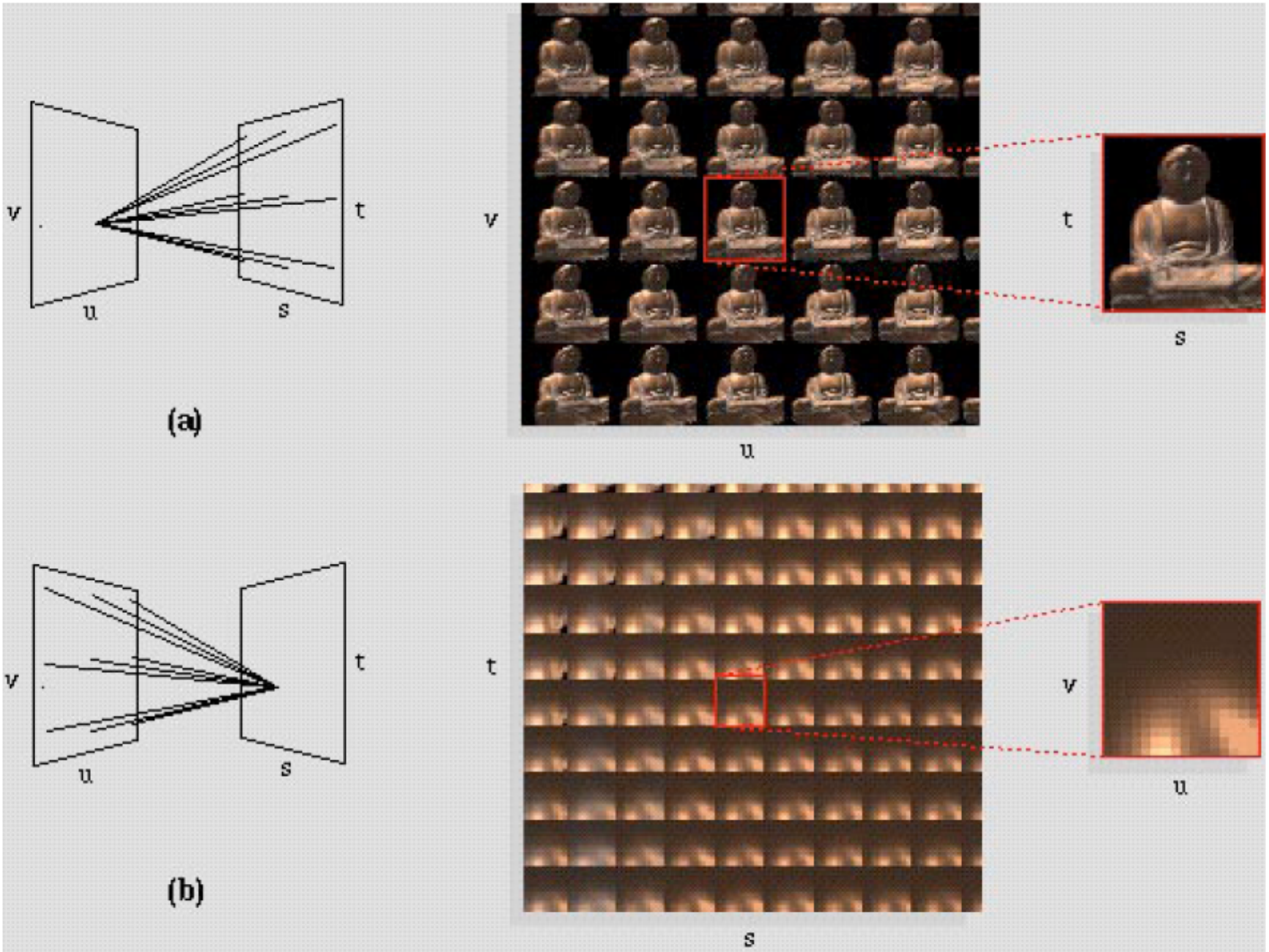


Figure 6: Two visualizations of a light field. (a) Each image in the array represents the rays arriving at one point on the uv plane from all points on the st plane, as shown at left. (b) Each image represents the rays leaving one point on the st plane bound for all points on the uv plane. The images in (a) are off-axis (i.e. sheared) perspective views of the scene, while the images in (b) look like reflectance maps. The latter occurs because the object has been placed astride the focal plane, making sets of rays leaving points on the focal plane similar in character to sets of rays leaving points on the object.

Cool visualization

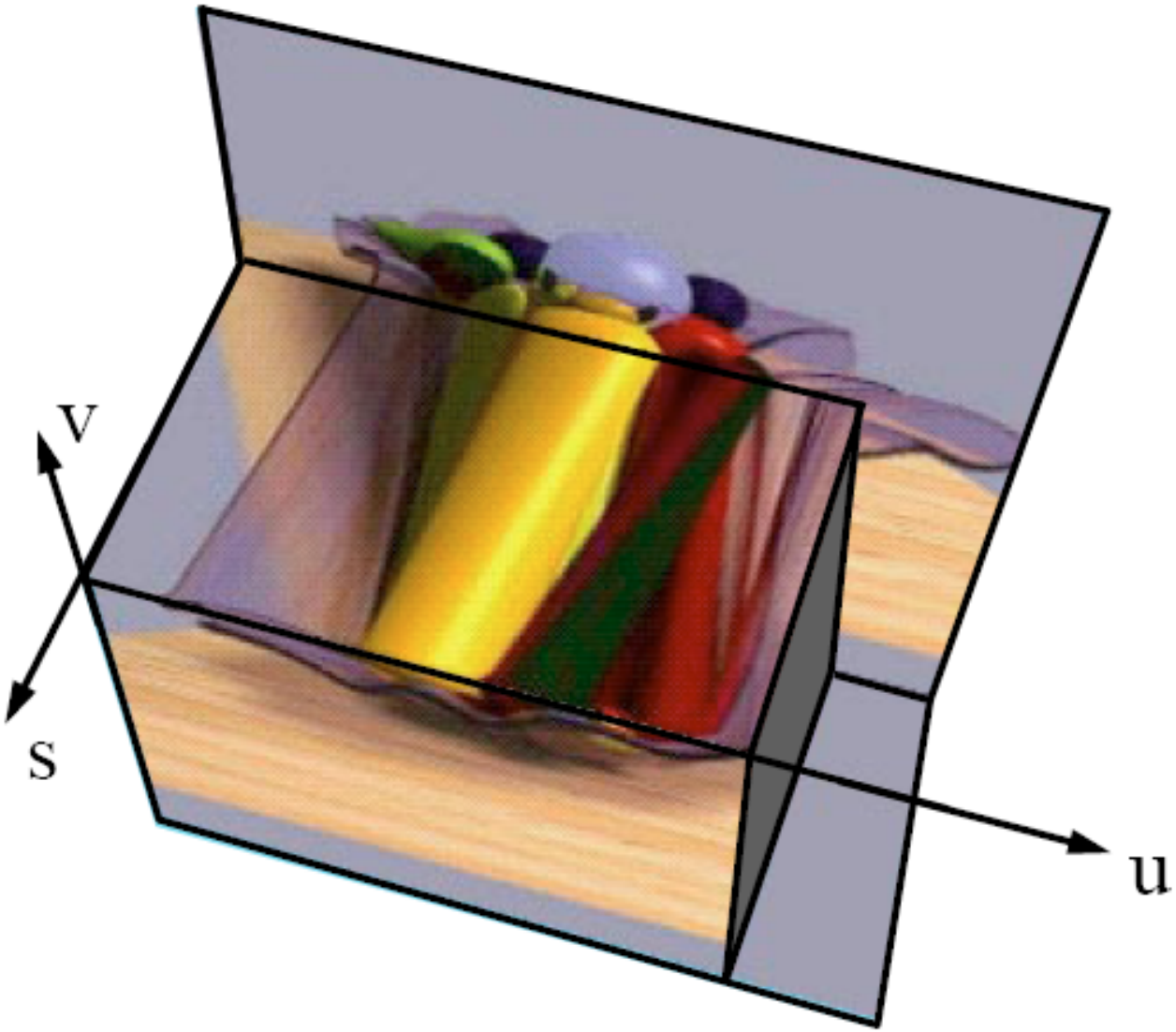
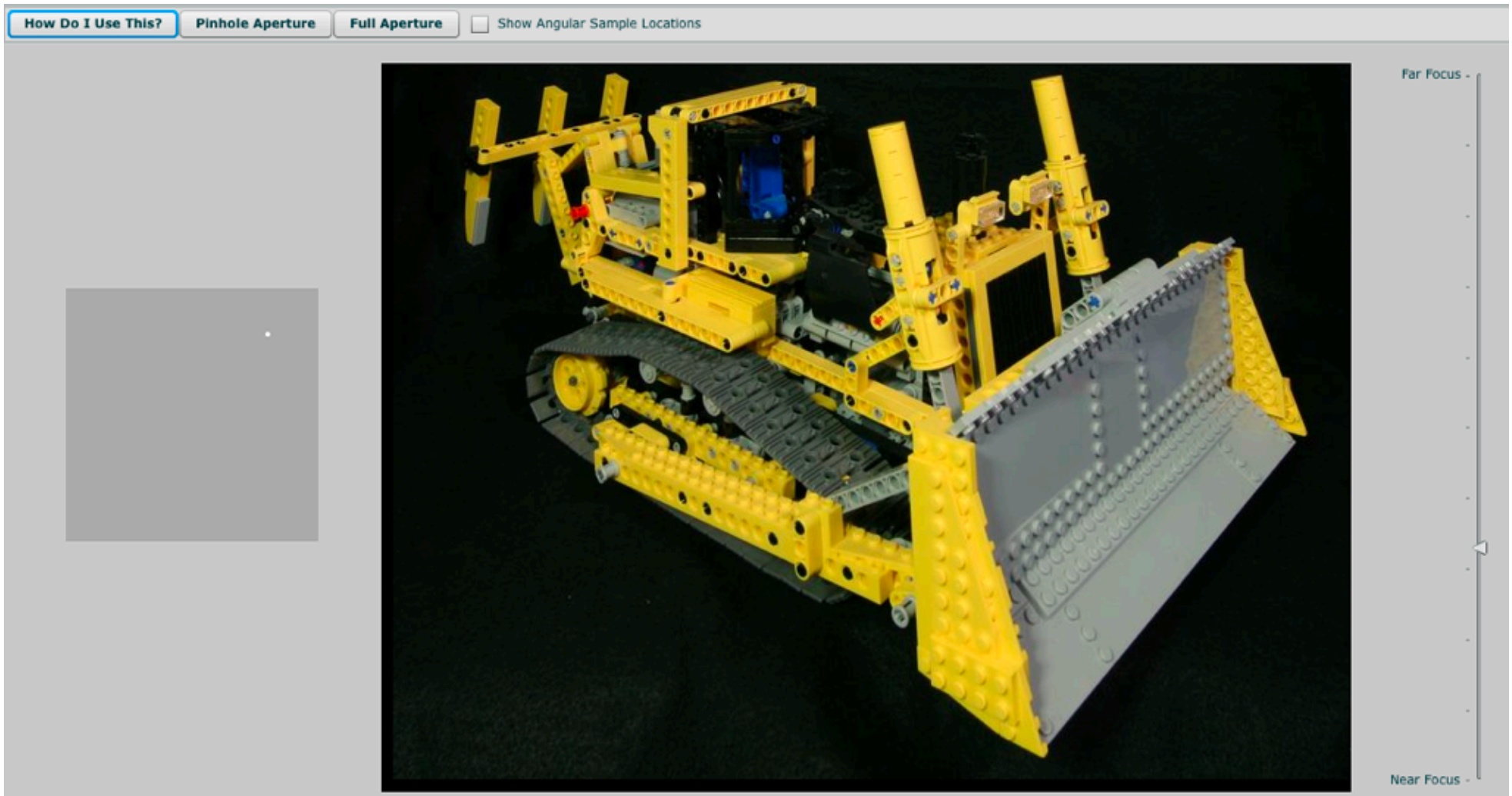


Figure 7: An (s, u, v) slice of a Lumigraph

From Gortler et al.

Demo light field viewer

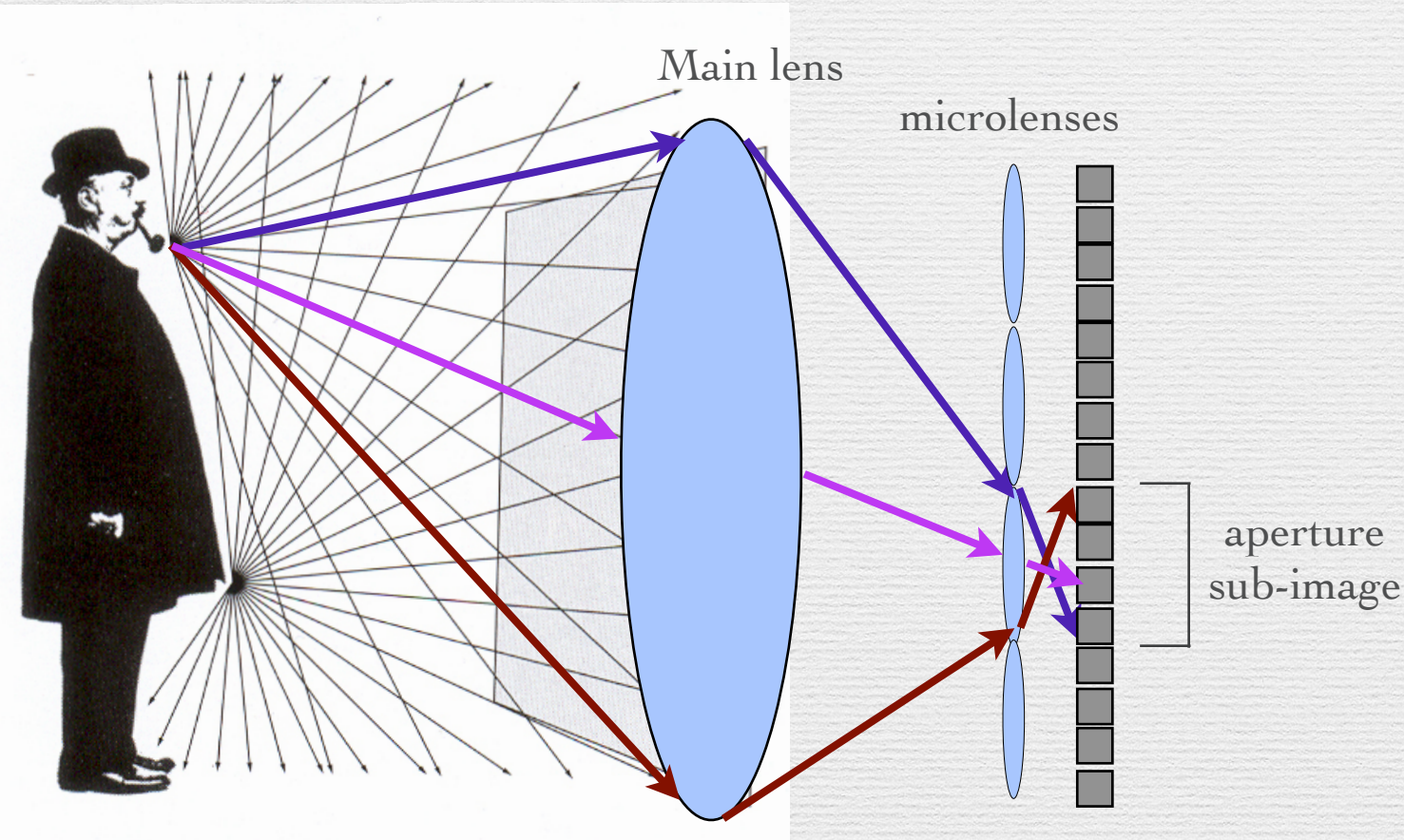
- **http://lightfield.stanford.edu/lfs.html**



Light field cameras

Plenoptic/light field cameras

- [Lipmann 1908, Adelson and Wang, 1992, Ng et al. 06]
- Record 4D set of rays
  - using microlens array in front of sensor



Lippmann 1908

- http://people.csail.mit.edu/fredo/PUBLI/Lippmann.pdf
- http://www.tgeorgiev.net/RadianceCameras/EpreuvesReversibles.pdf

PHOTOGRAPHIE. — *Épreuves réversibles. Photographies intégrales.* Note de M. G. LIPPMANN.

1. La plus parfaite des épreuves photographiques actuelles ne montre que l'un des aspects de la réalité ; elle se réduit à une image unique fixée dans un plan, comme le serait un dessin ou une peinture tracée à la main. La vue directe de la réalité offre, on le sait, infiniment plus de variété. On voit les objets dans l'espace, en vraie grandeur et en relief, et non dans un plan. De plus, leur aspect change avec les positions de l'observateur ; les différents plans de la vue se déplacent alors les uns par rapport aux autres ; la perspective se modifie ; les parties cachées ne restent pas les mêmes ; enfin, si le spectateur regarde le monde extérieur par une fenêtre, il est maître de voir les diverses parties d'un paysage venir s'encadrer successivement entre les bords de l'ouverture, si bien que dans ce cas ce sont des objets différents qui lui apparaissent successivement.

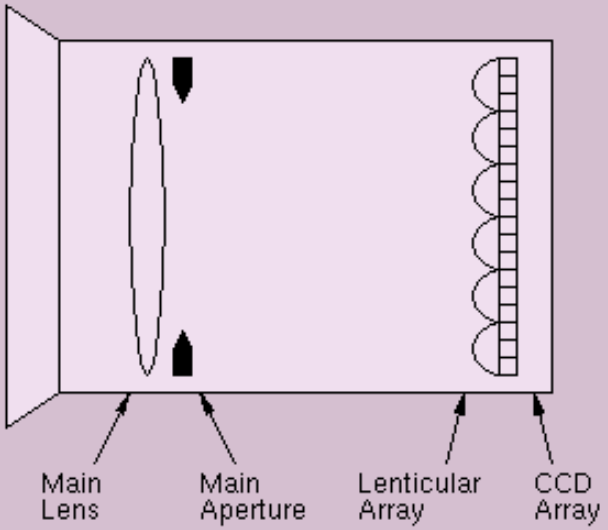
Peut-on demander à la Photographie de nous rendre toute cette variété

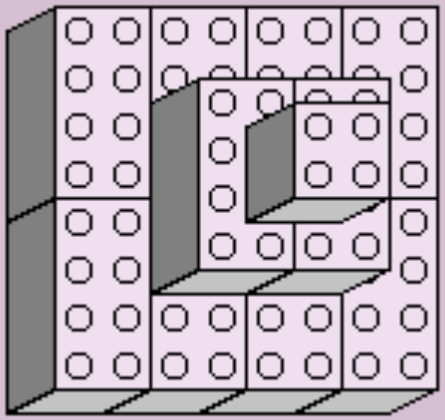
Fig. 1.

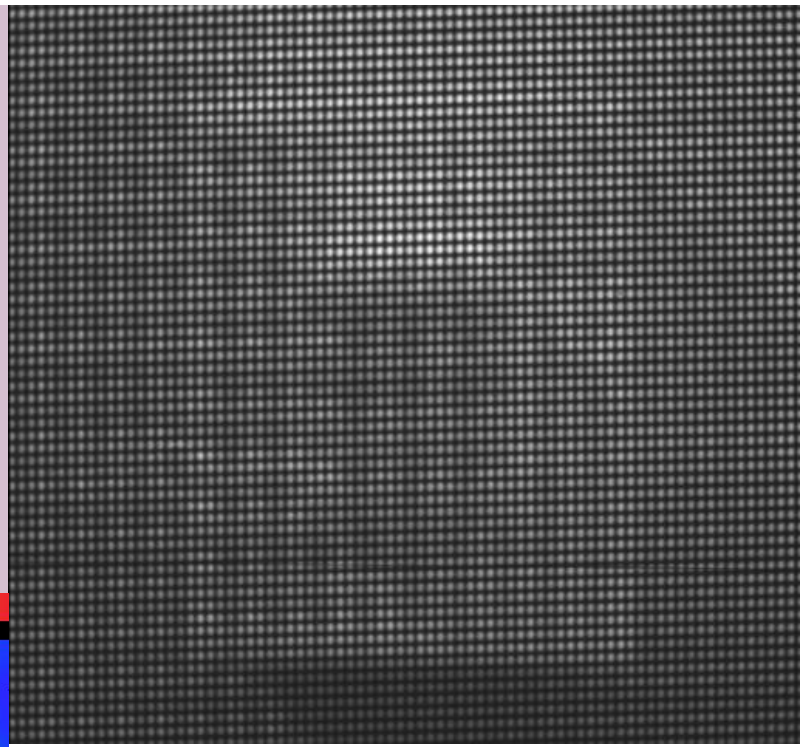


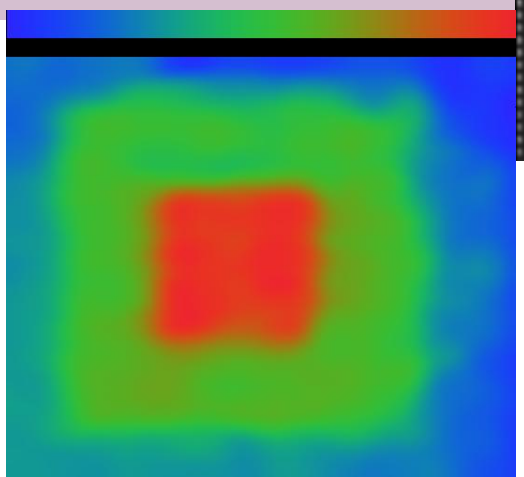
Plenoptic camera

- **For depth extraction**
- **Adelson & Wang 92**
  http://www-bcs.mit.edu/people/jyawang/demos/plenoptic/plenoptic.html



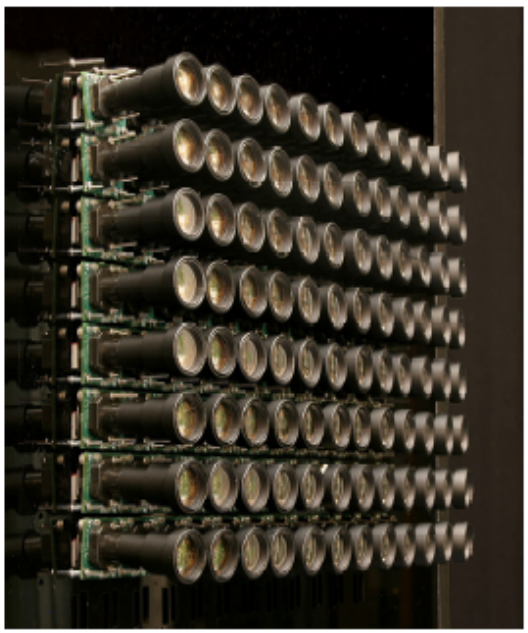


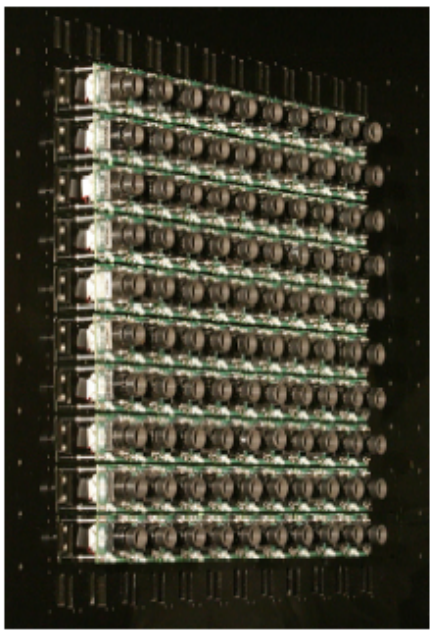




Camera array

- **Willburn et al.** http://graphics.stanford.edu/papers/CameraArray/







Camera arrays

- **http://graphics.stanford.edu/projects/array/**

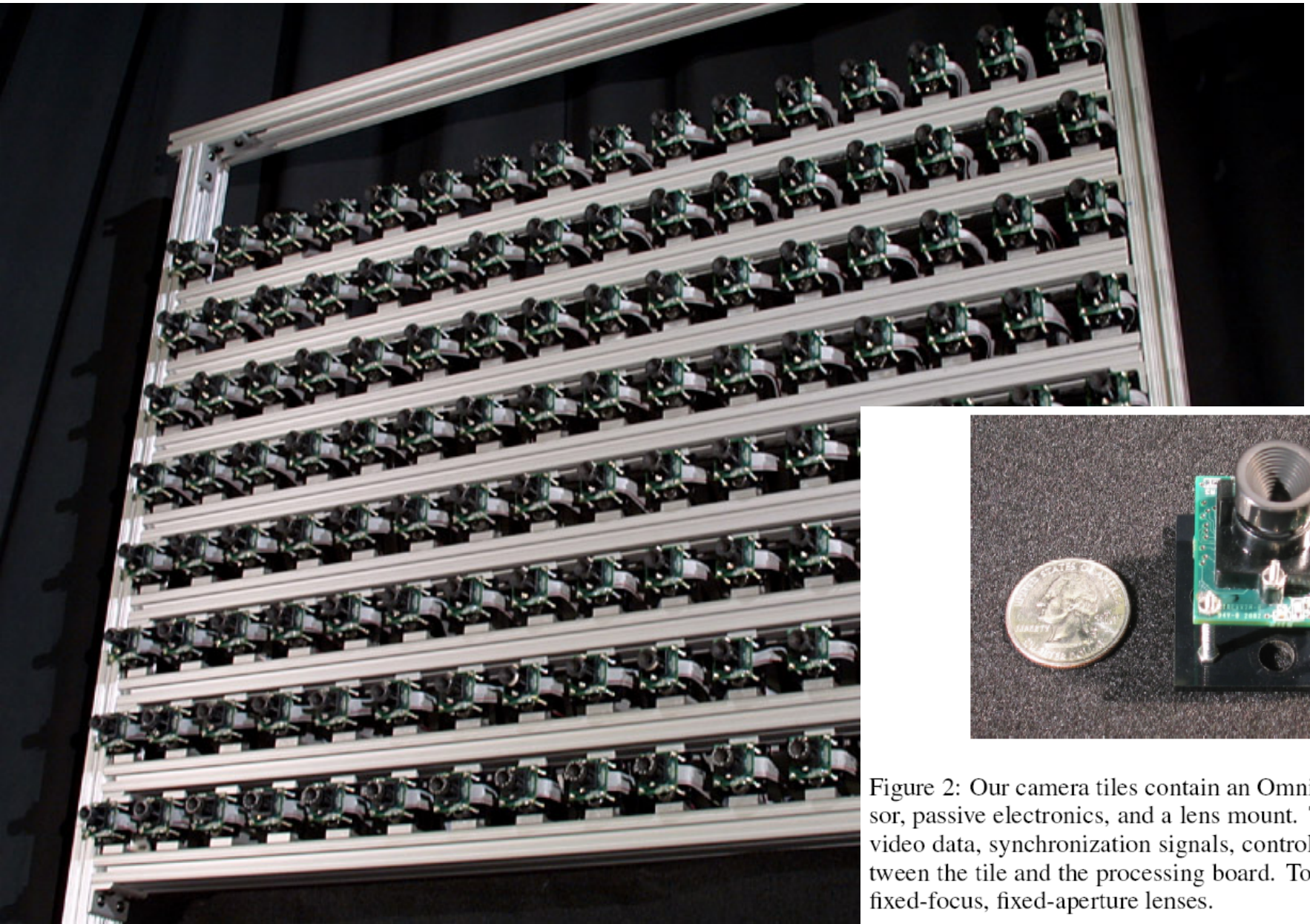


Figure 2: Our camera tiles contain an Omnivision 8610 image sensor, passive electronics, and a lens mount. The ribbon cables carry video data, synchronization signals, control signals, and power between the tile and the processing board. To keep costs low, we use fixed-focus, fixed-aperture lenses.

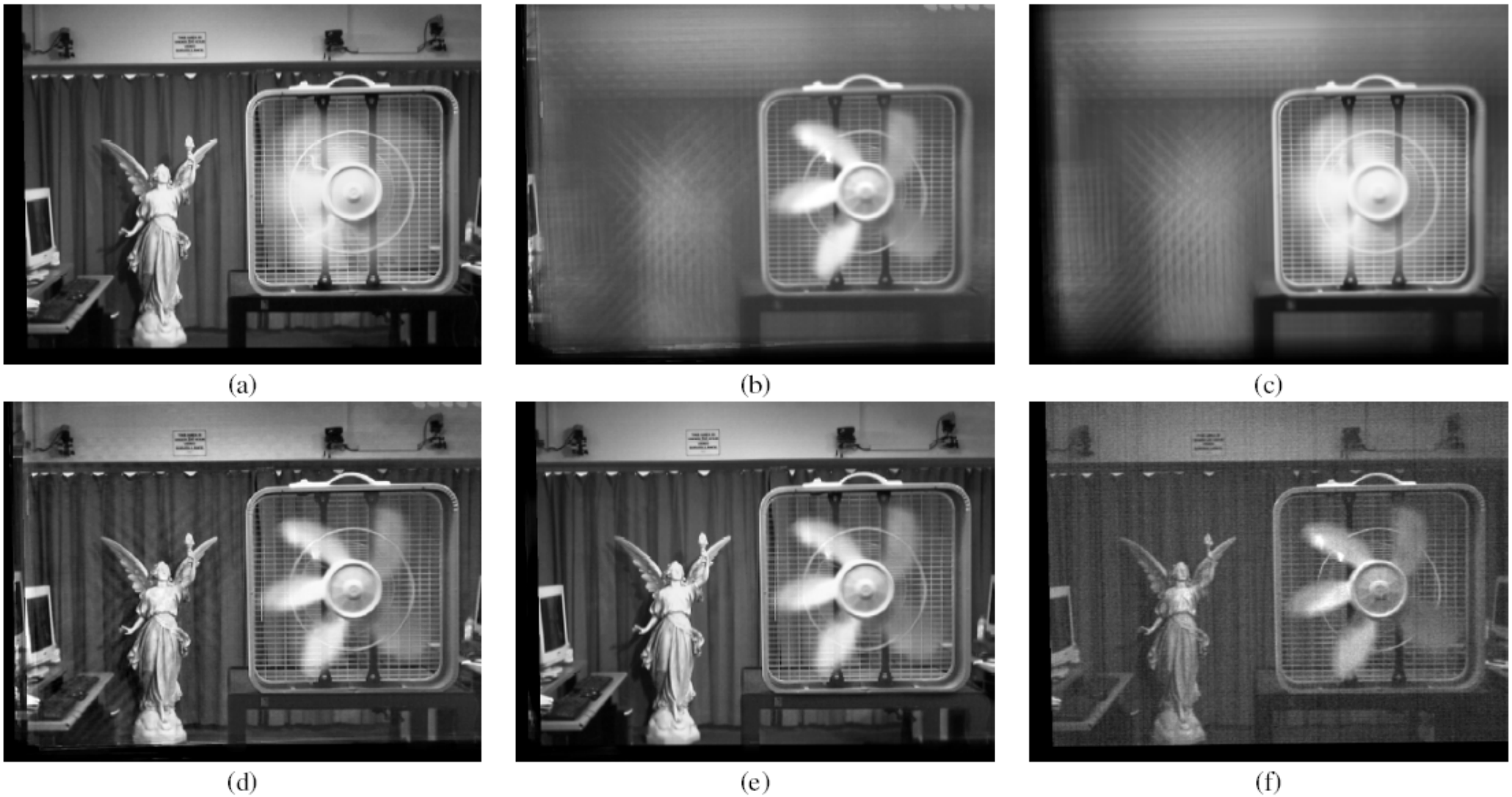
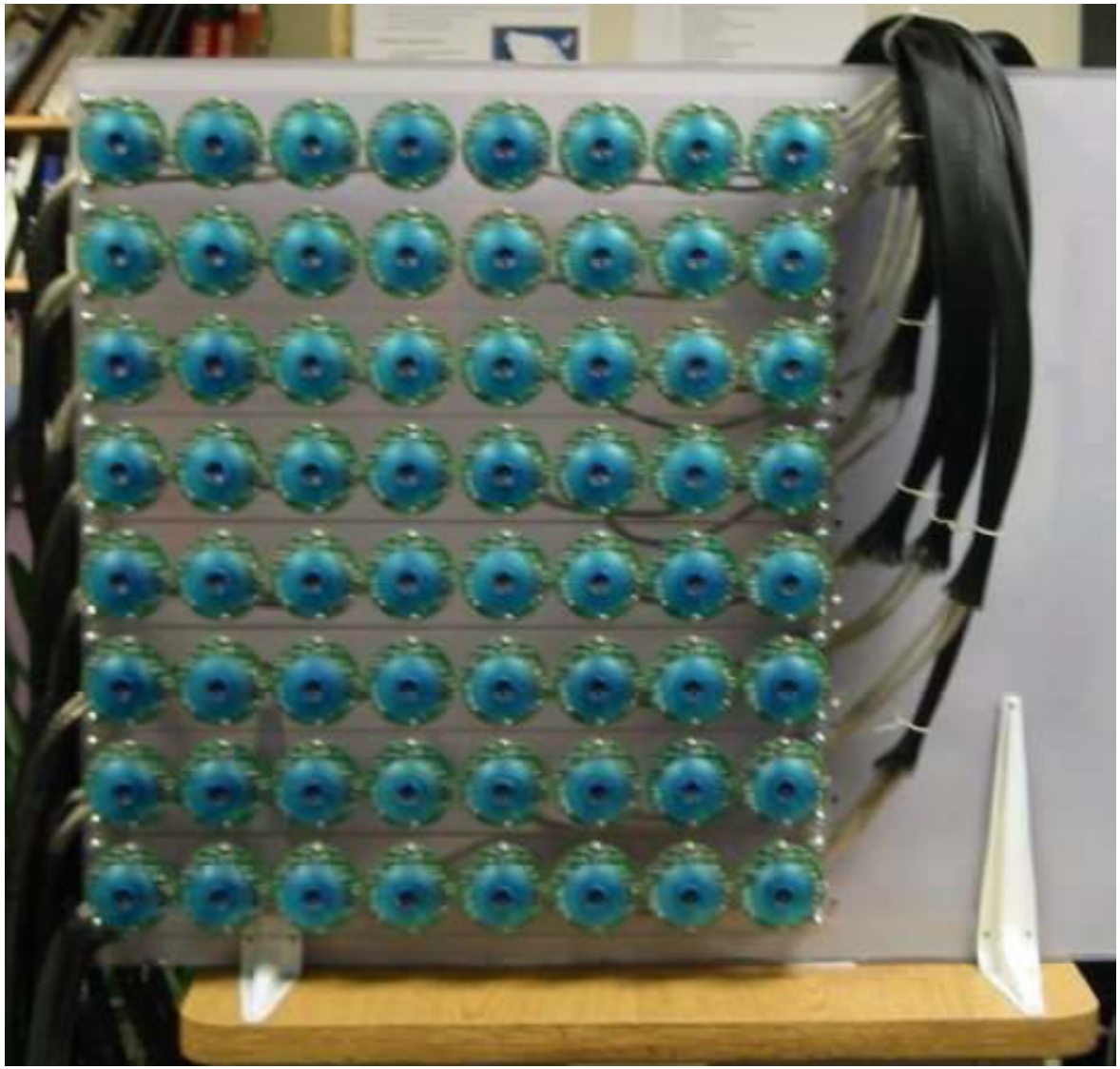


Figure 12: Hybrid synthetic aperture photography for combining high depth of field and low motion blur. (a-c) Images captured of a scene simultaneously through three different apertures: a single camera with a long exposure time (a), a large synthetic aperture with short exposure time (b), and a large synthetic aperture with a long exposure time. Computing (a+b-c) yields image (d), which has aliasing artifacts because the synthetic apertures are sampled sparsely from slightly different locations. Masking pixels not in focus in the synthetic aperture images before computing the difference (a + b - c) removes the aliasing (e). For comparison, image (f) shows the image taken with an aperture that is narrow in both space and time. The entire scene is in focus and the fan motion is frozen, but the image is much noisier.

MIT version

- **Jason Yang**



Bullet time

- **Time splice http://www.ruffy.com/frameset.htm**



Robotic Camera



Image Leonard McMillan



Image Levoy et al.

Legos

- http://lightfield.stanford.ed a/self_portrait_lf/preview.zip



Flatbed scanner camera

- **By Jason Yang**



Hand-Held Light Field Camera



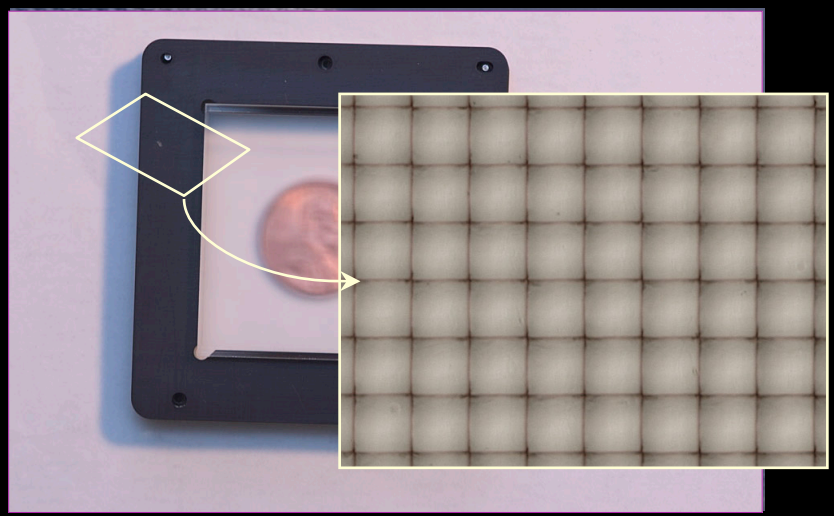
Medium format digital camera



Camera in-use



16 megapixel sensor



Microlens array

Slide by Ren Ng.

New version: Lytro

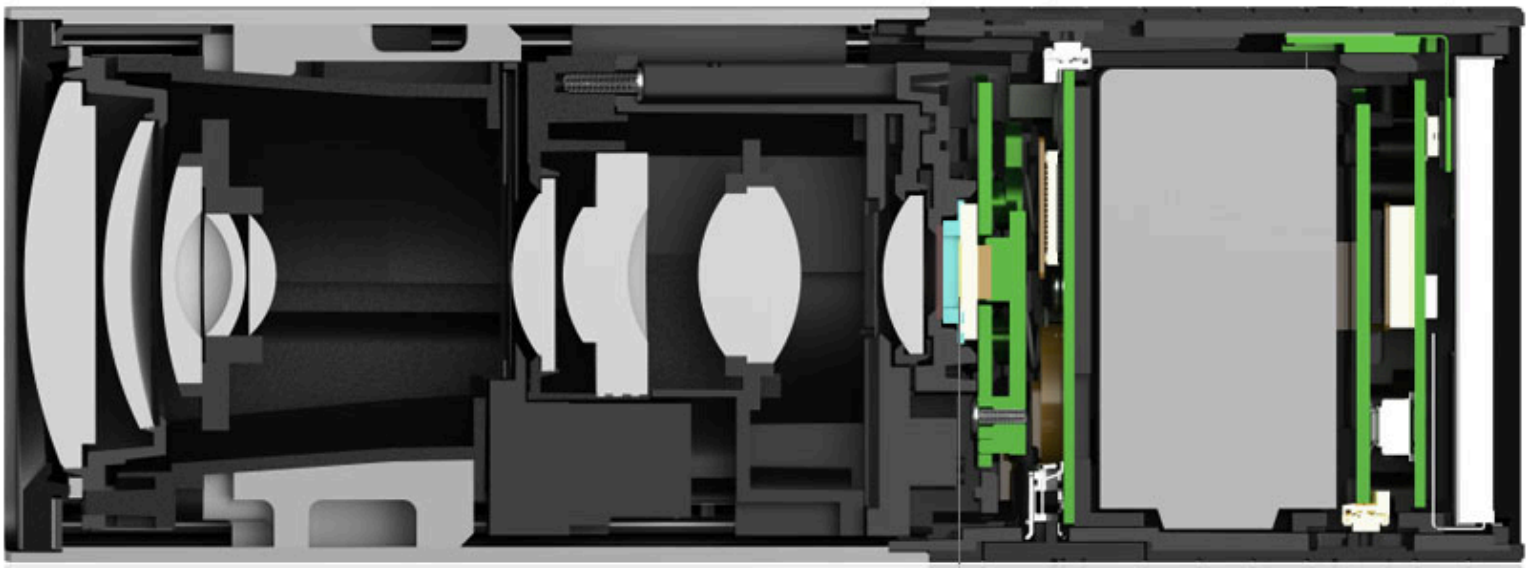


- **http://www.lytro.com/**
- **hexagonal microlens array**

Lens
The Lytro Light Field Camera starts with an 8X optical zoom, f/2 aperture lens. The aperture is constant across the zoom range allowing for unheard of light capture.

Light Field Engine 1.0
The Light Field Engine replaces the supercomputer from the lab and processes the light ray data captured by the sensor.

The Light Field Engine travels with every living picture as it is shared, letting you refocus pictures right on the camera, on your desktop and online.



Light Field Sensor
From a roomful of cameras to a micro-lens array specially adhered to a standard sensor, the Lytro's Light Field Sensor captures 11 million light rays.

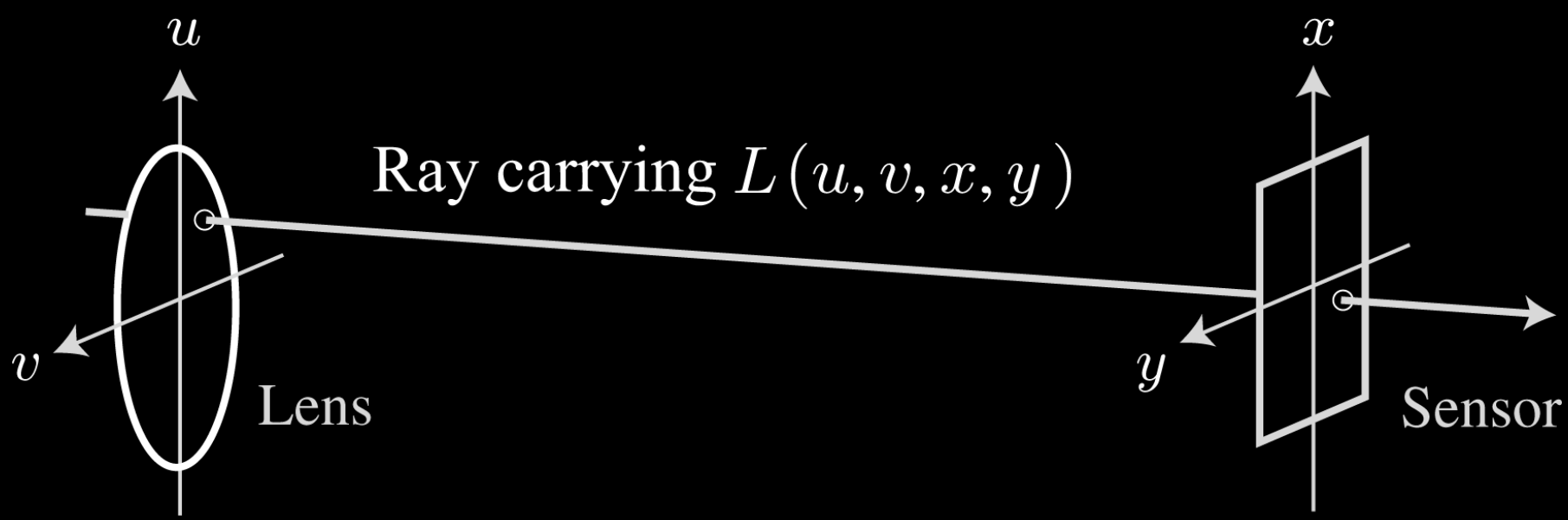
Plenoptic camera refocusing

Conventional Photograph



Slide by Ren Ng.

Light Field Photography



- **Capture the light field inside the camera body**

Slide by Ren Ng.

Hand-Held Light Field Camera



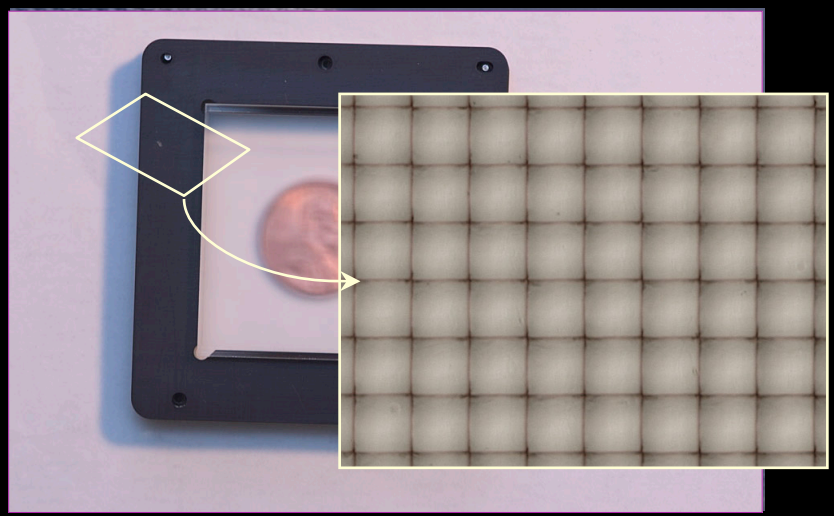
Medium format digital camera



Camera in-use



16 megapixel sensor



Microlens array

Slide by Ren Ng.

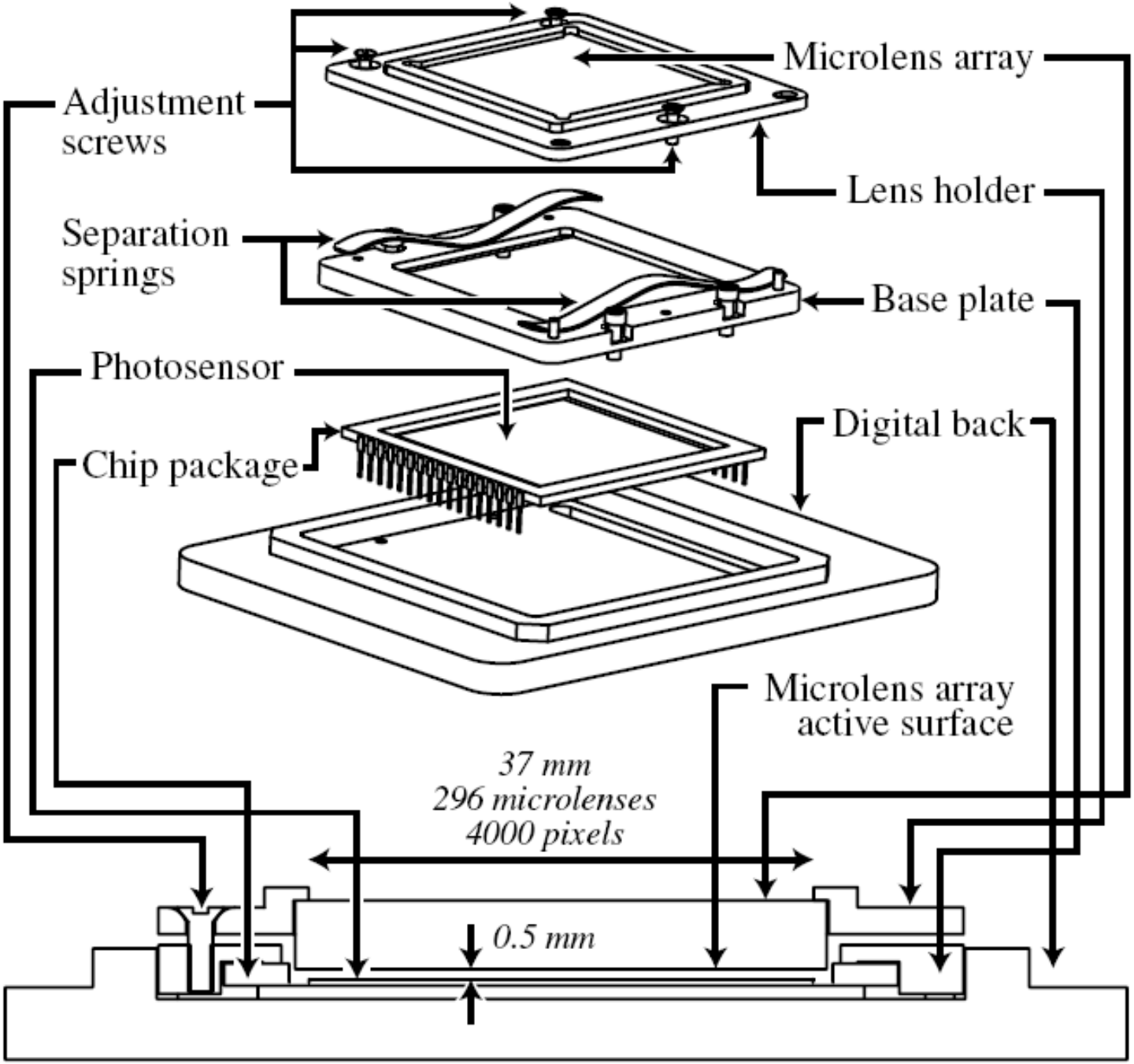


Figure 8: *Top:* Exploded view of assembly for attaching the microlens array to the digital back. *Bottom:* Cross-section through assembled parts.

New version: Lytro

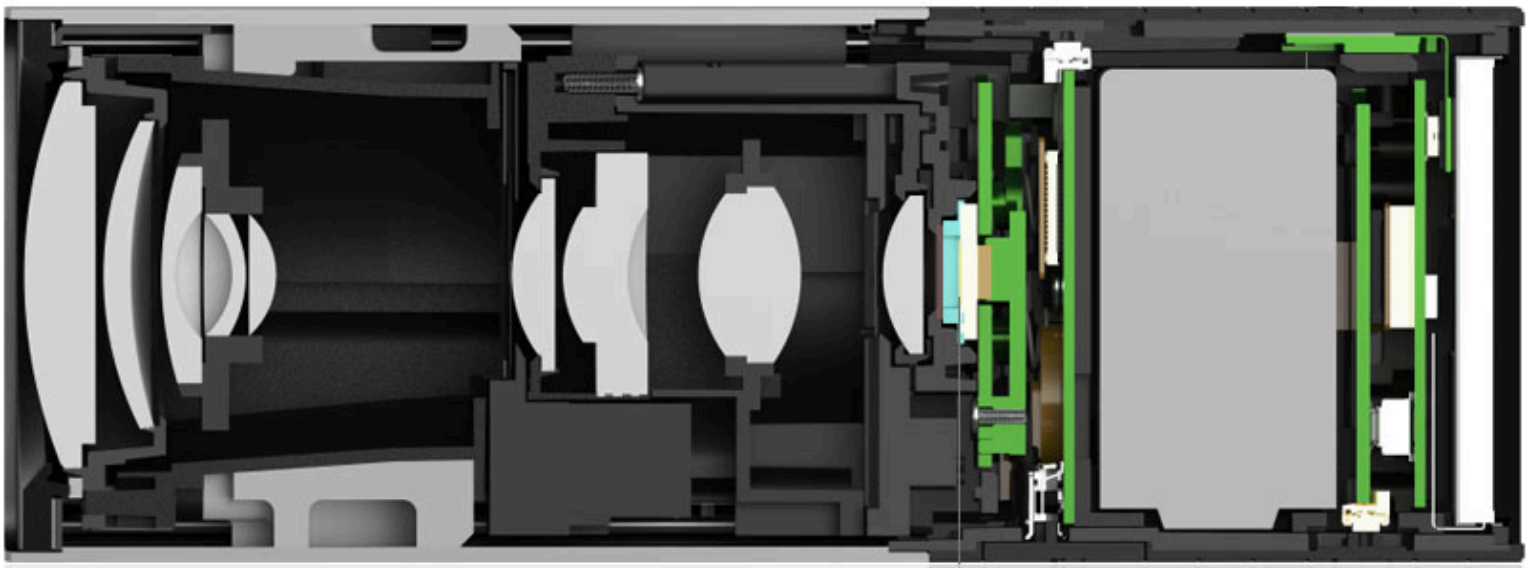


- **http://www.lytro.com/**
- **hexagonal microlens array**

Lens
The Lytro Light Field Camera starts with an 8X optical zoom, f/2 aperture lens. The aperture is constant across the zoom range allowing for unheard of light capture.

Light Field Engine 1.0
The Light Field Engine replaces the supercomputer from the lab and processes the light ray data captured by the sensor.

The Light Field Engine travels with every living picture as it is shared, letting you refocus pictures right on the camera, on your desktop and online.



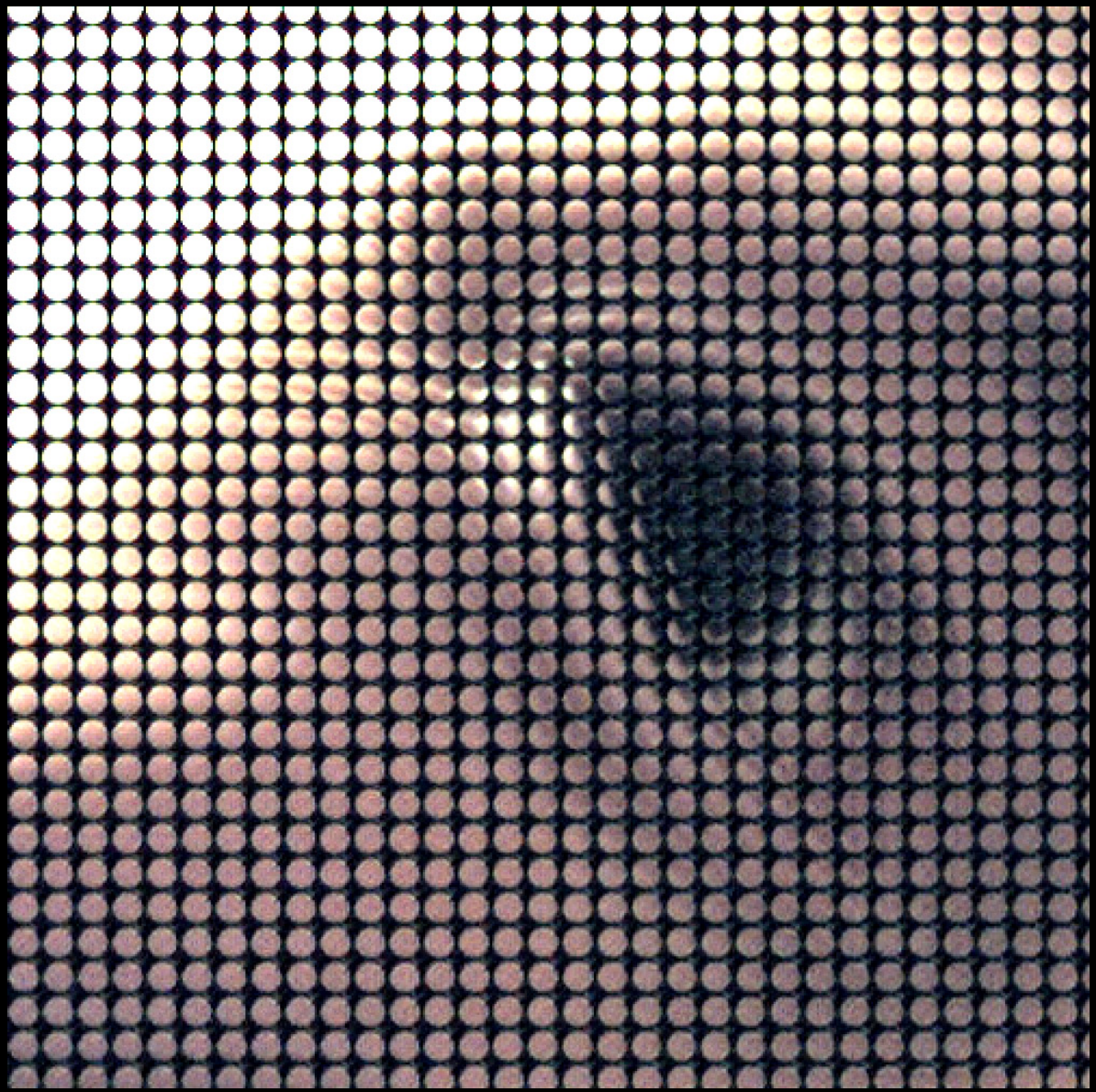
Light Field Sensor
From a roomful of cameras to a micro-lens array specially adhered to a standard sensor, the Lytro's Light Field Sensor captures 11 million light rays.

Light Field in a Single Exposure



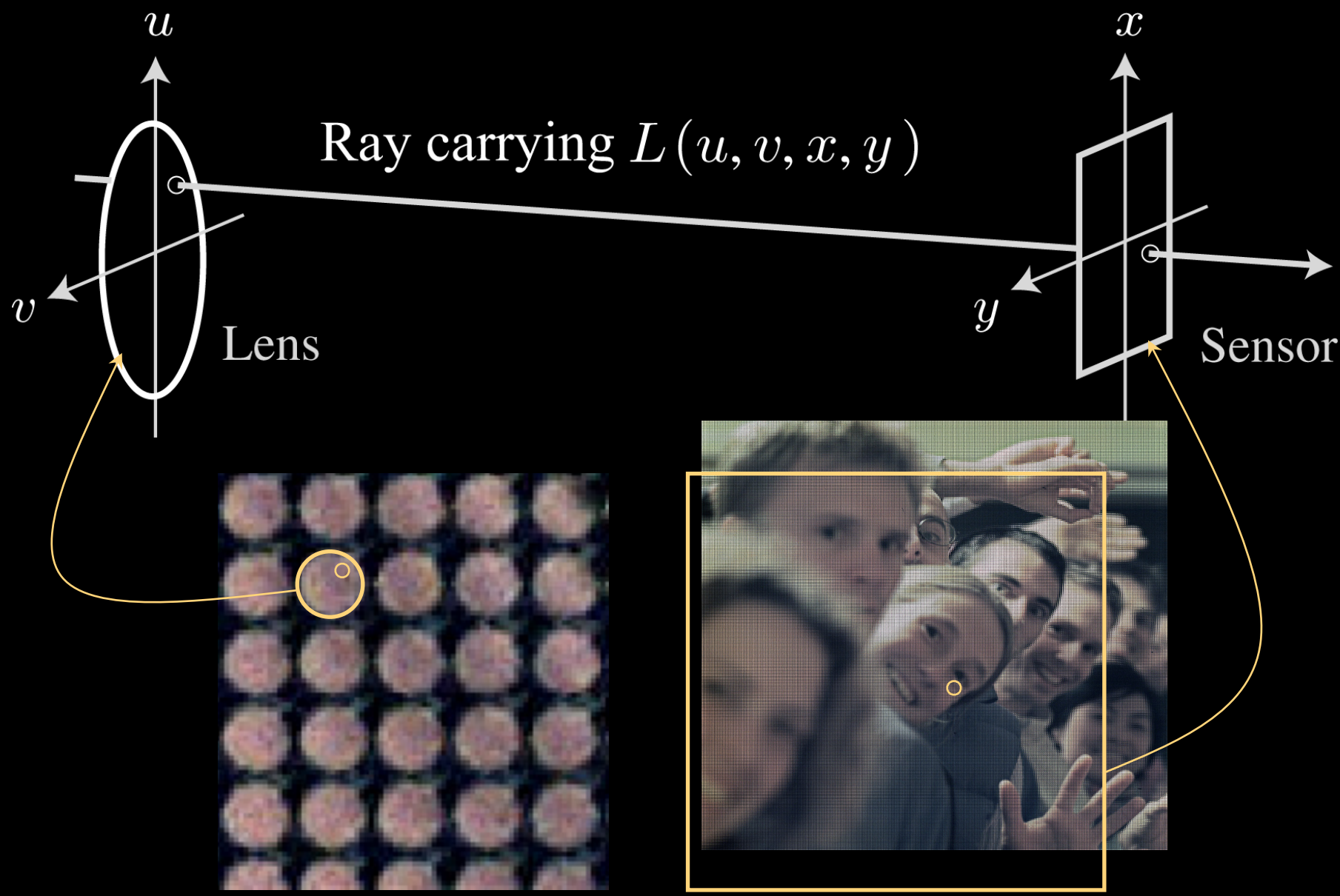
Slide by Ren Ng.

Light Field in a Single Exposure



Slide by Ren Ng.

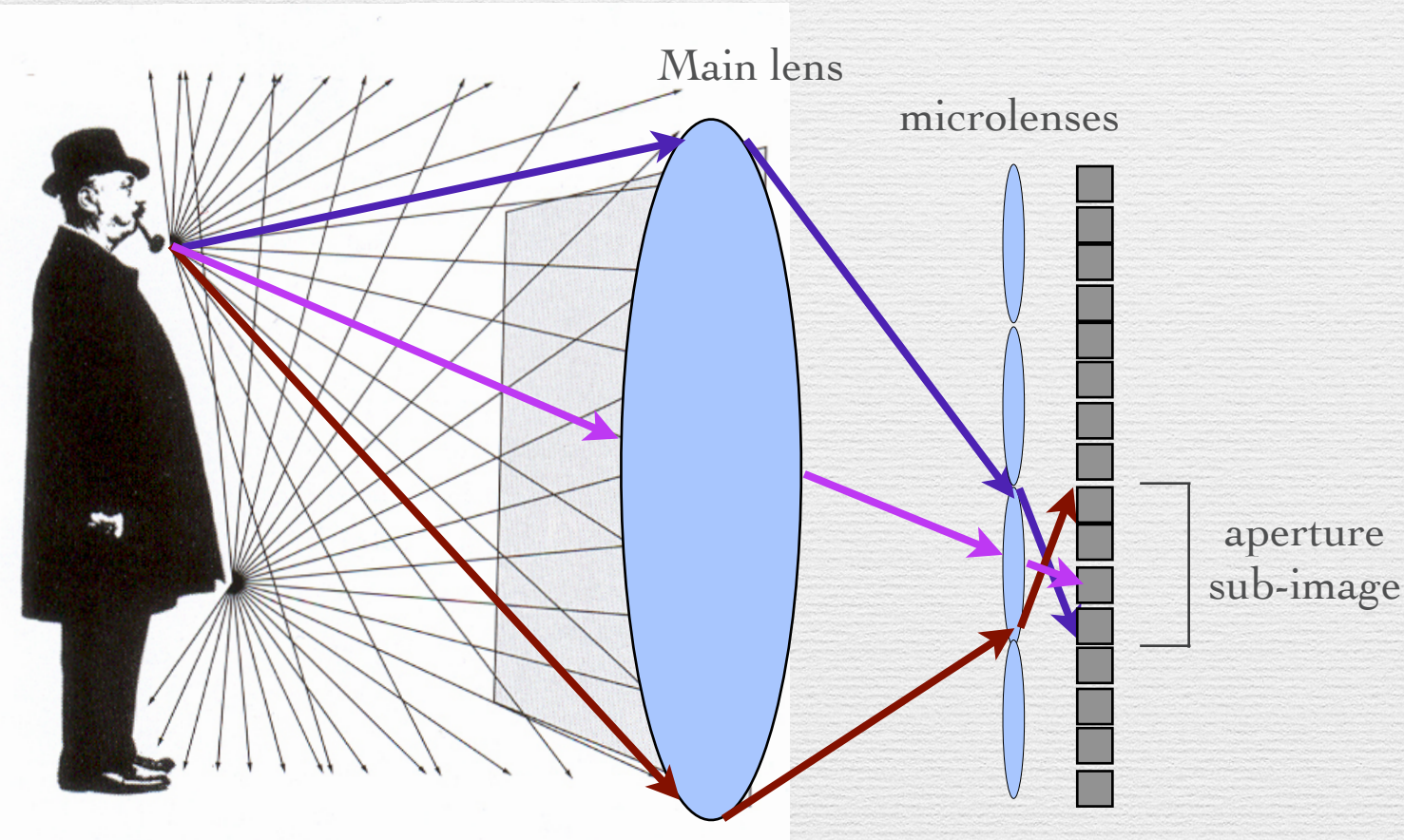
Light Field Inside the Camera Body



Slide by Ren Ng.

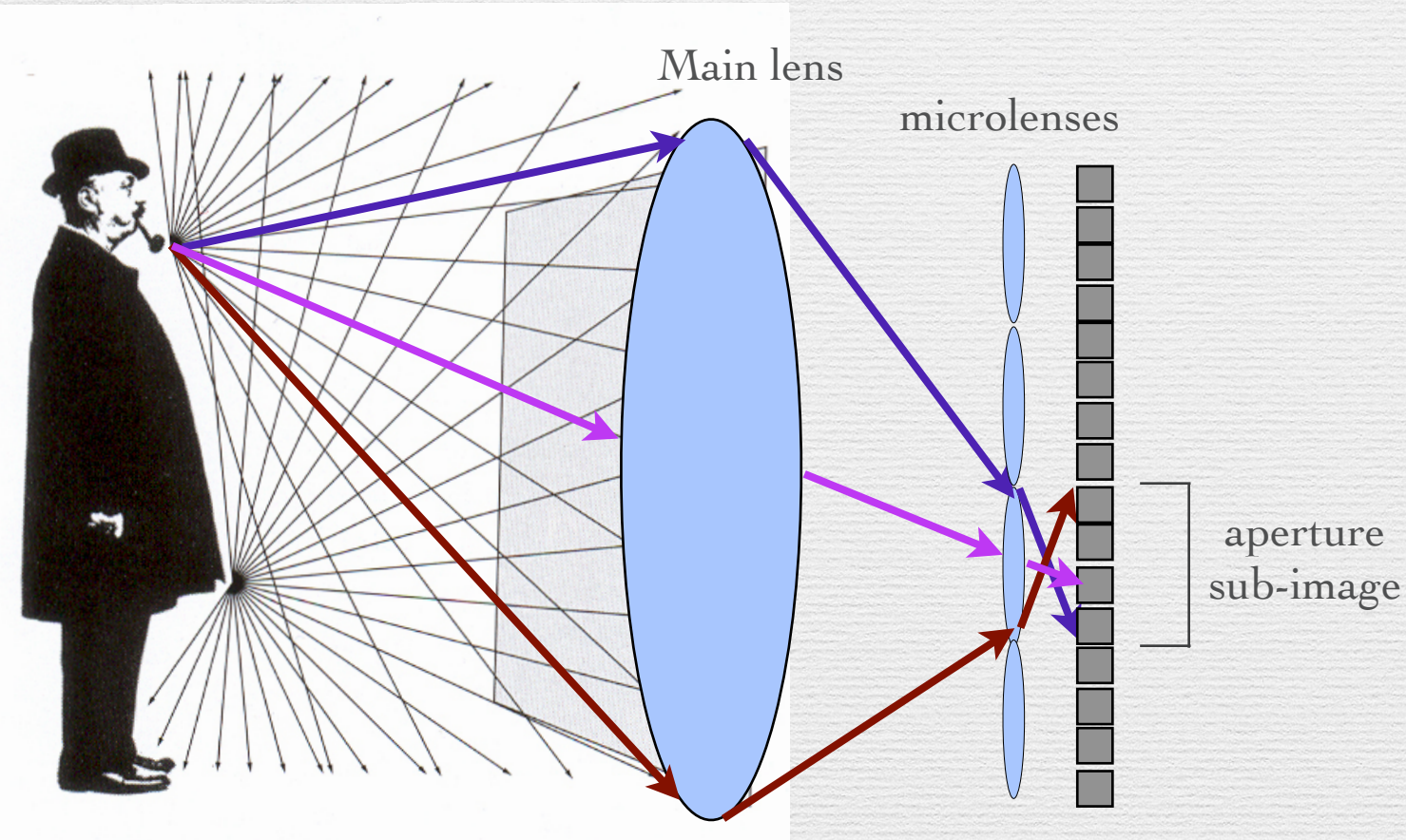
Recap & Questions?

- Record 4D set of rays
  - using microlens array in front of sensor



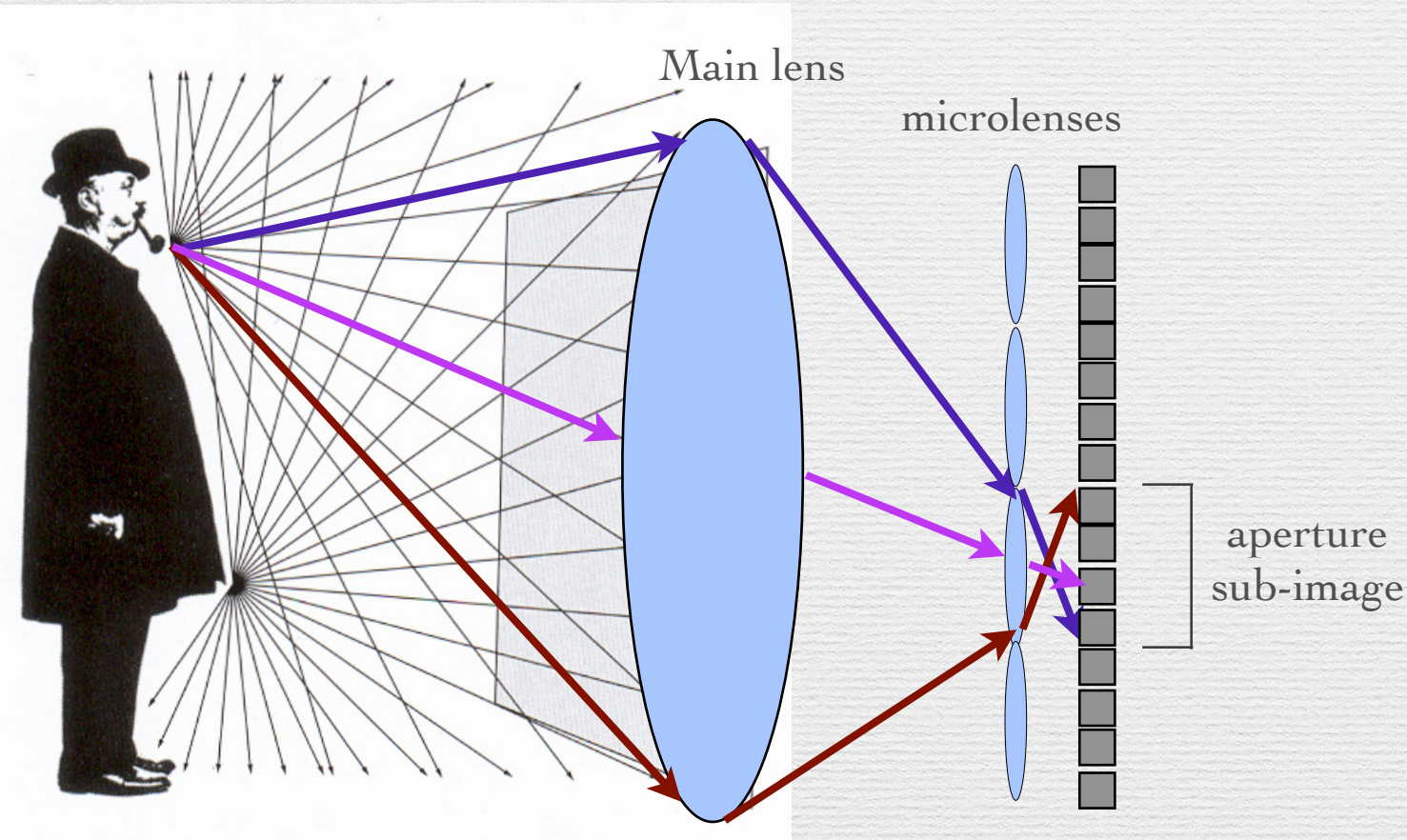
Reconstructing a normal image

- Just integrate all the pixels in an aperture sub-image

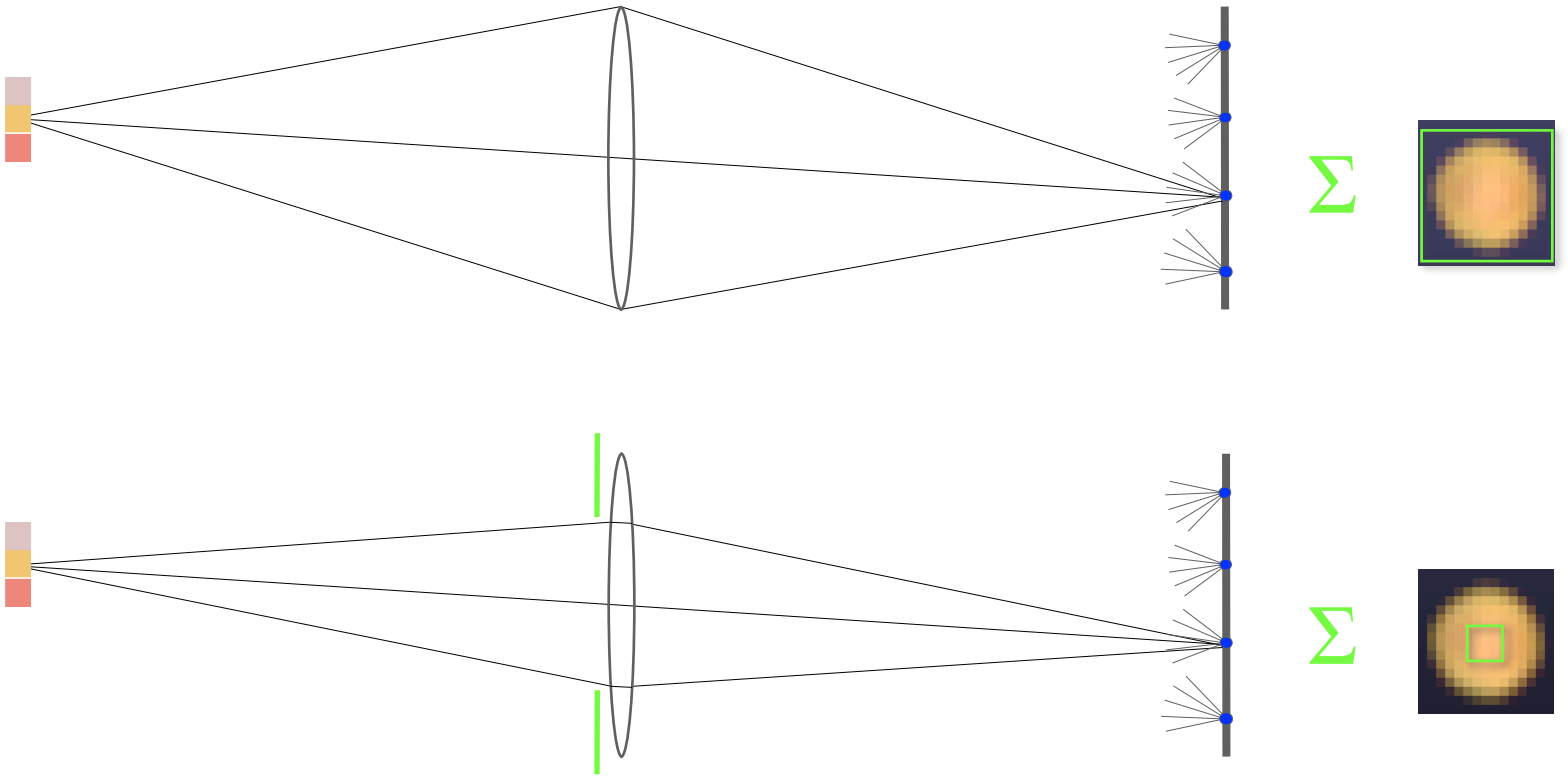


Stopping down the lens

- Just integrate the pixels only in the center of an aperture sub-image



Digitally stopping-down



stopping down = summing only the central portion of each microlens

Changing the viewpoint

- **Just take one given pixel in each aperture sub-image**
  - Viewpoint limited to the lens aperture

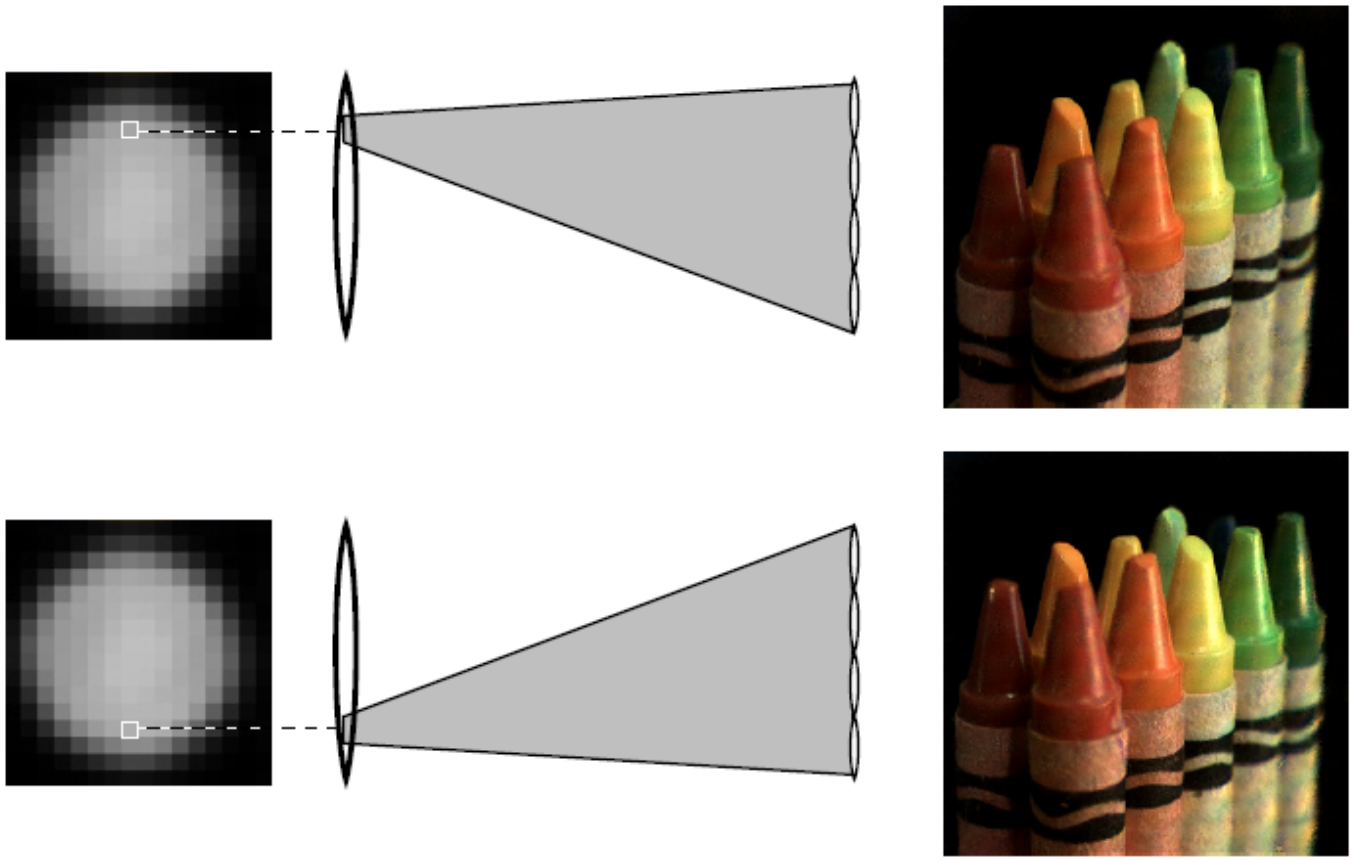
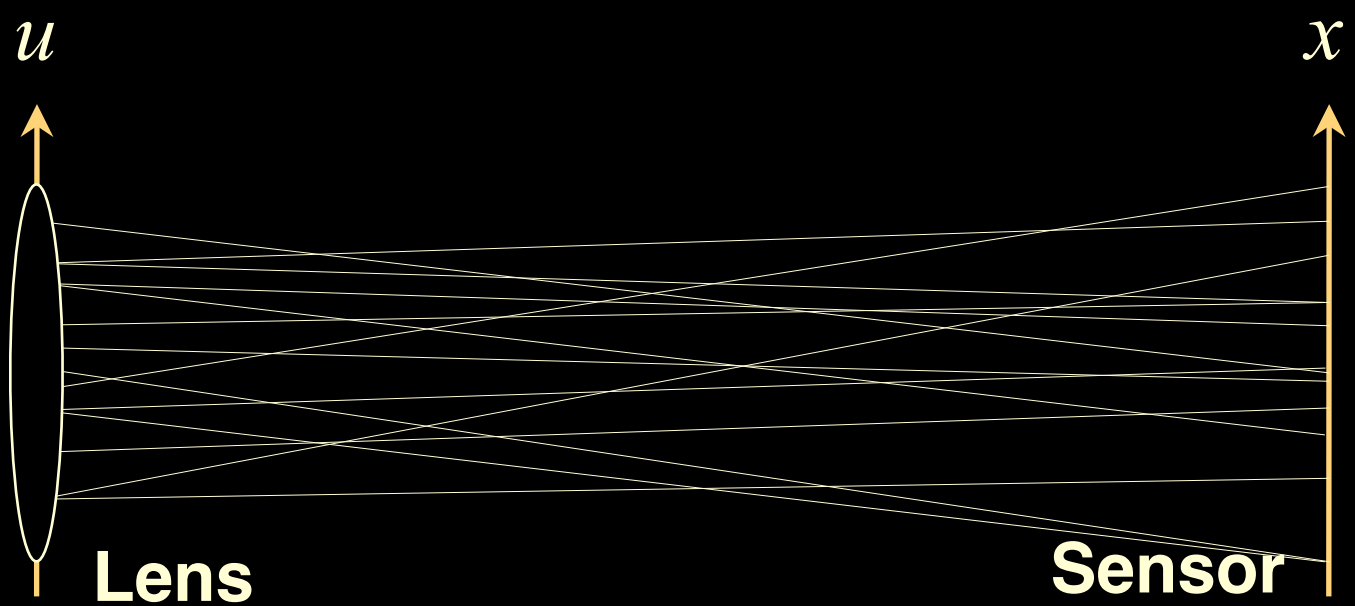


Figure 4: Two sub-aperture photographs obtained from a light field by extracting the shown pixel under each microlens (depicted on left). Note that the images are not the same, but exhibit vertical parallax.

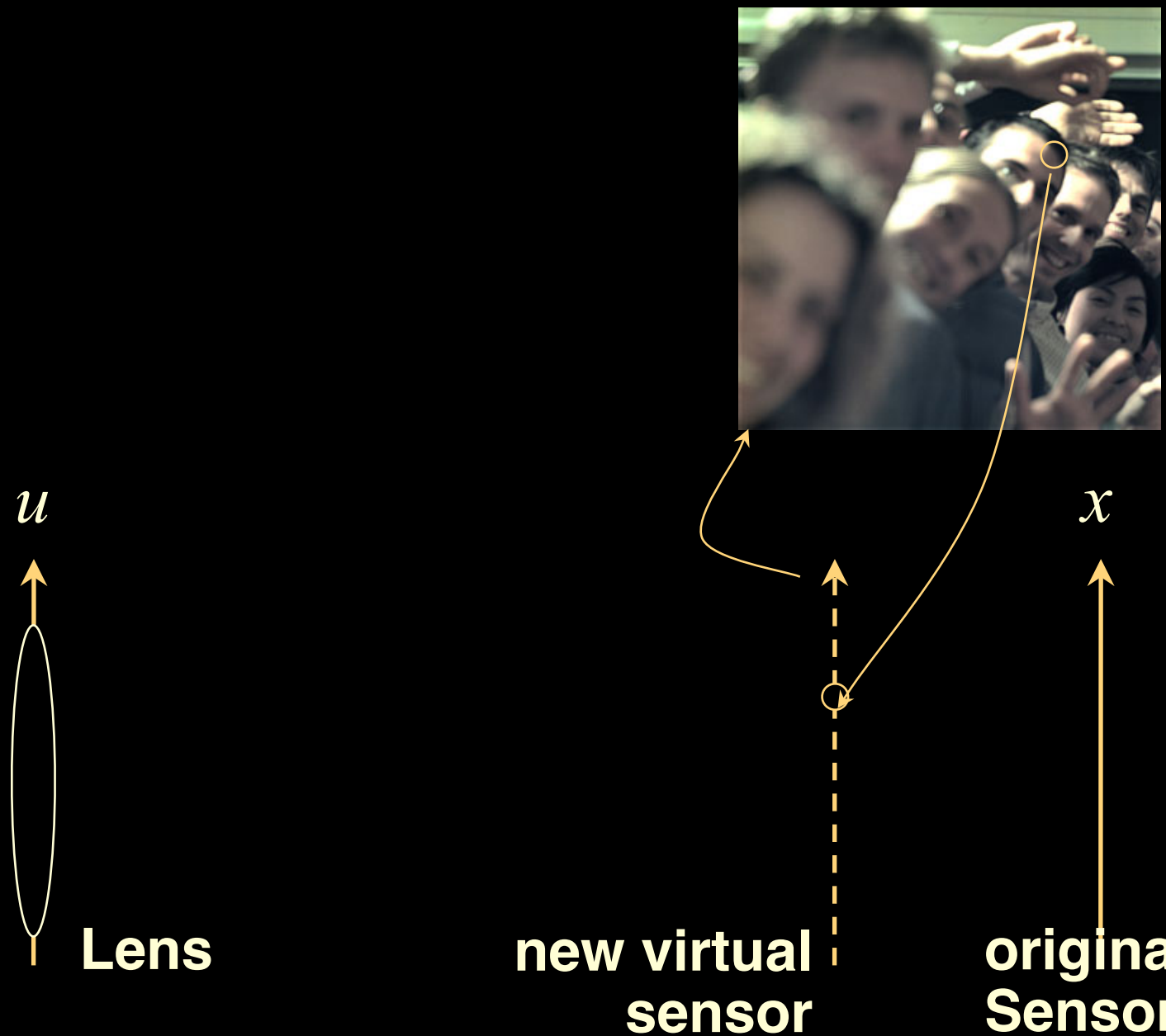
Digital Refocusing by Ray-Tracing

We have all the light rays
We can generate any virtual view



Slide by Ren Ng.

Digital Refocusing by Ray-Tracing

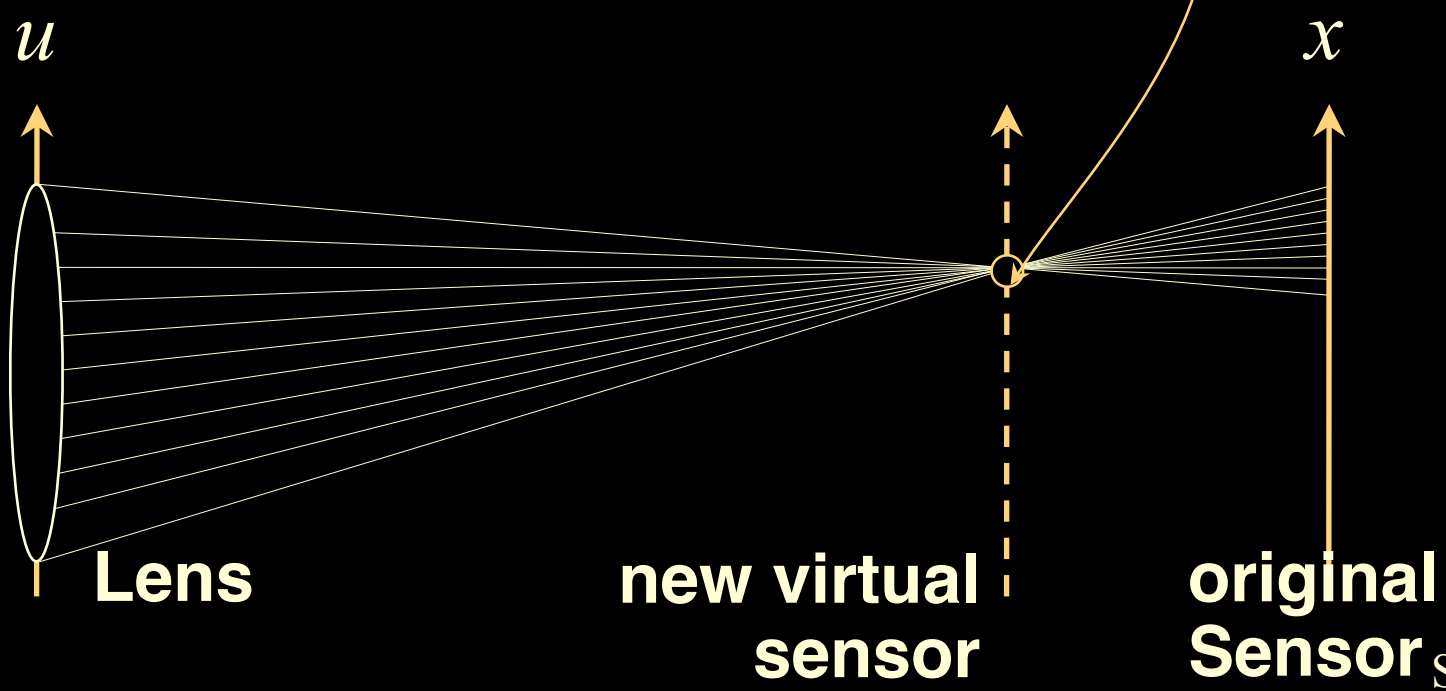


Slide by Ren Ng.

Digital Refocusing by Ray-Tracing

For each pixel of the new view
Integrate light rays
from different original
aperture images

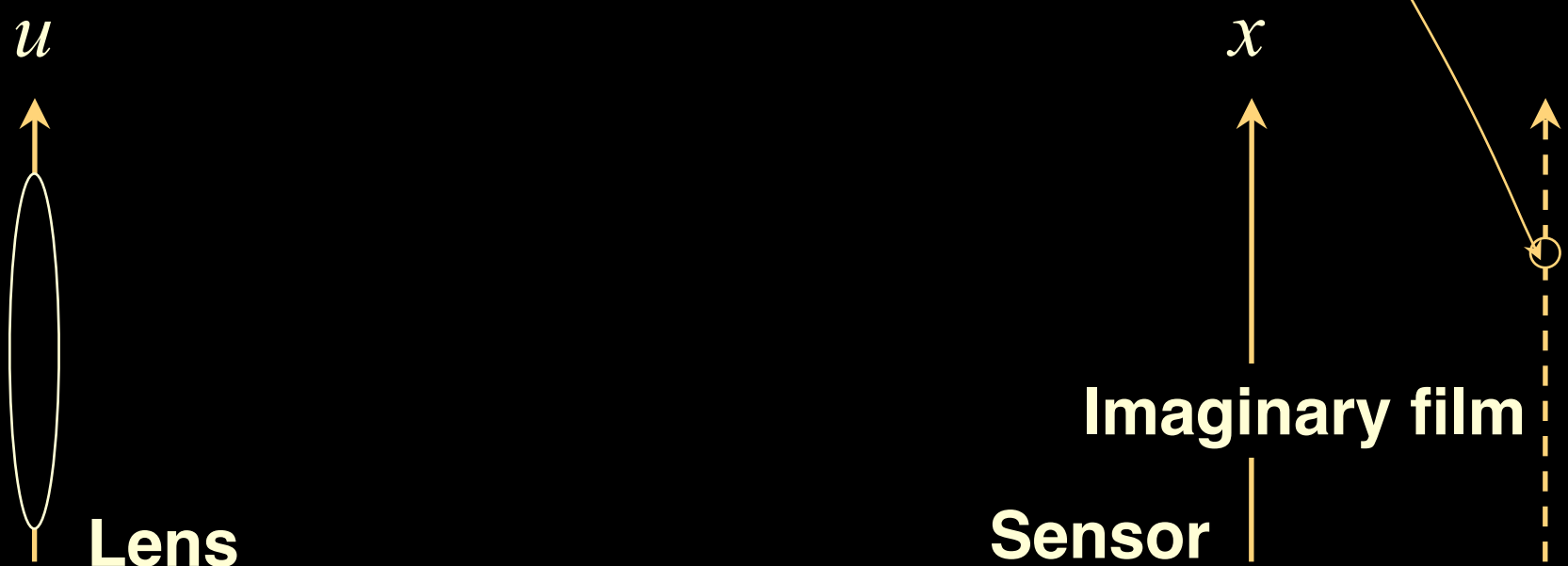




Slide by Ren Ng.

Digital Refocusing by Ray-Tracing

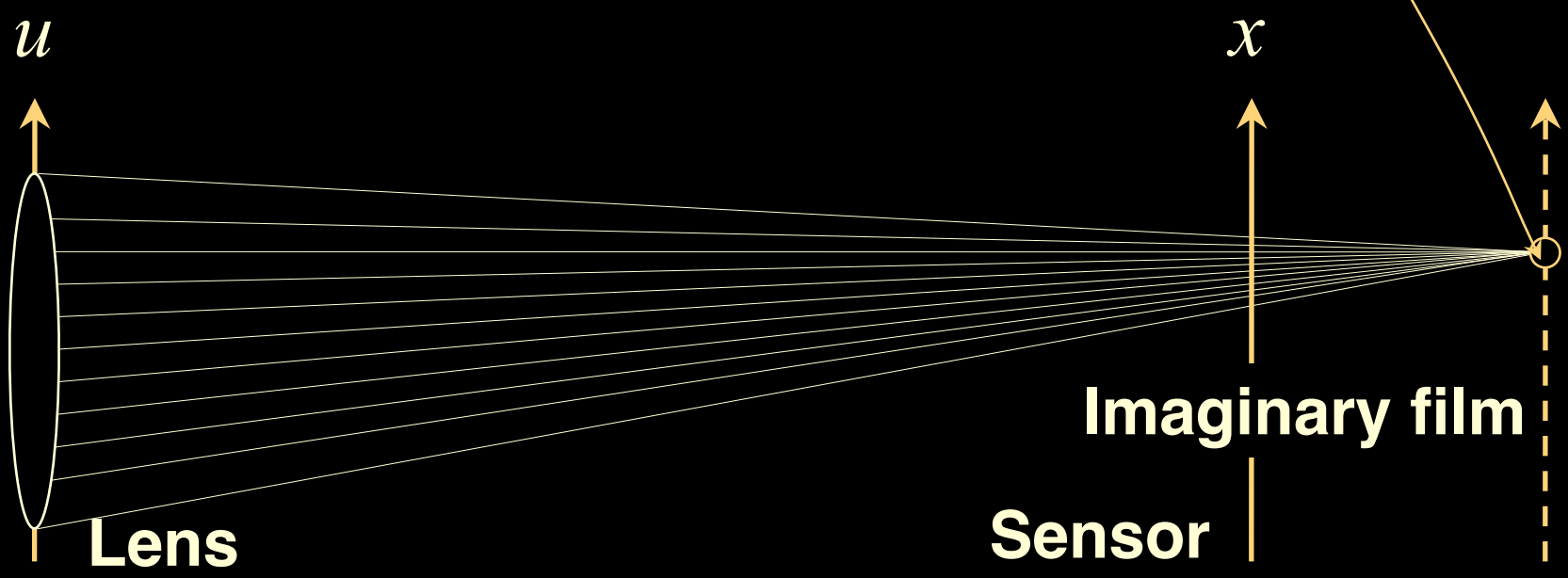




Slide by Ren Ng.

Digital Refocusing by Ray-Tracing





Slide by Ren Ng.

Digital Refocusing



Slide by Ren Ng.

Digital Refocusing



Slide by Ren Ng.

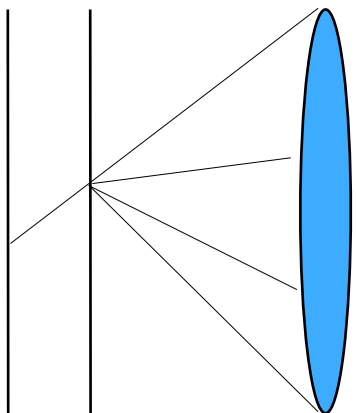
Web demo

- http://lytro.com/gallery/



Light Field Refocusing in Practice

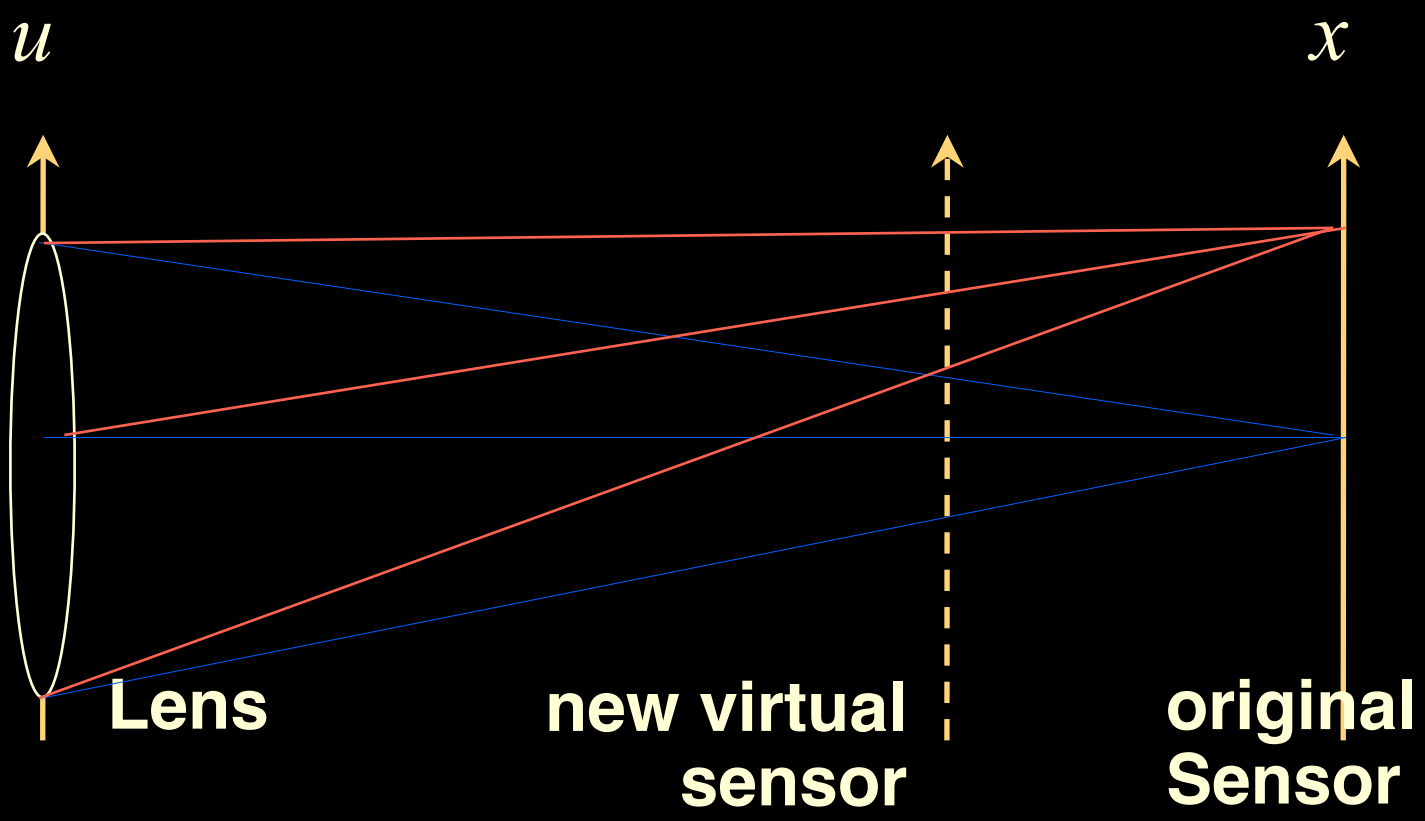
recorded



lens

Digital Refocusing by Ray-Tracing

Let's see where different aperture locations get reprojected



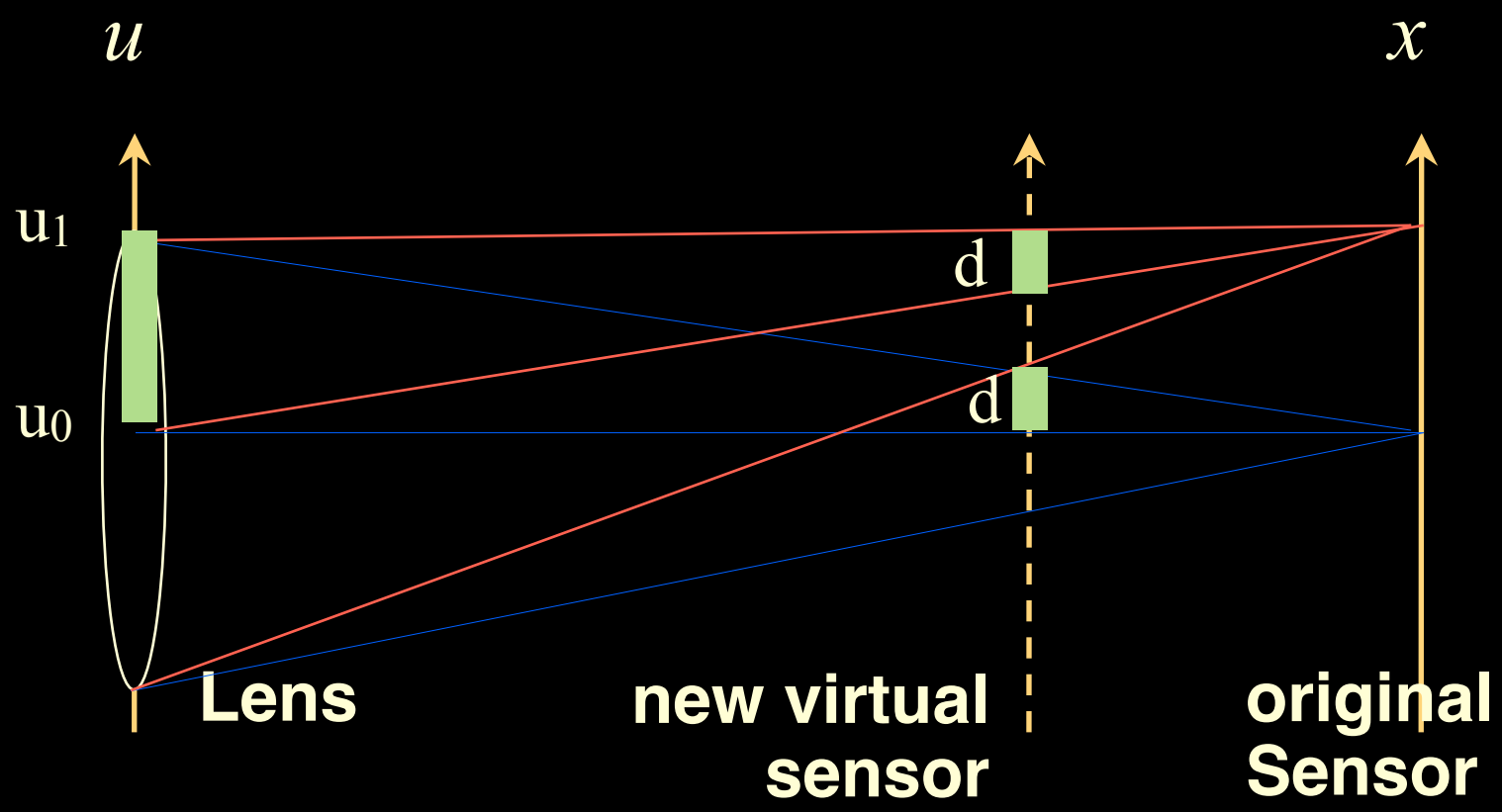
Digital Refocusing by Ray-Tracing

Let's see where different aperture locations get reprojected

Similar triangles:

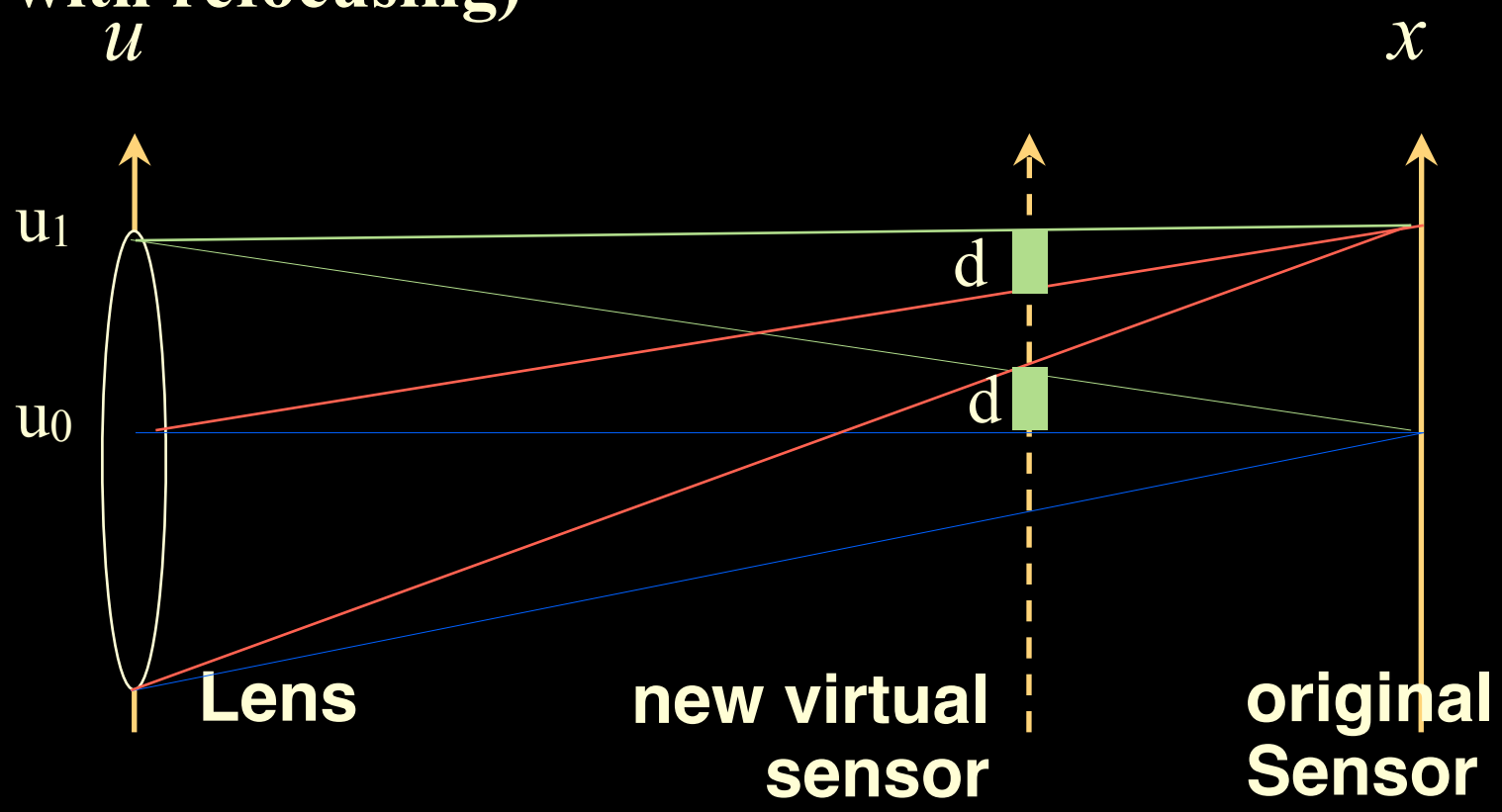
The rays from u_1 are shifted with respect to the rays from u_0 by the same amount

This amount is proportional to $\|u_1 u_0\|$



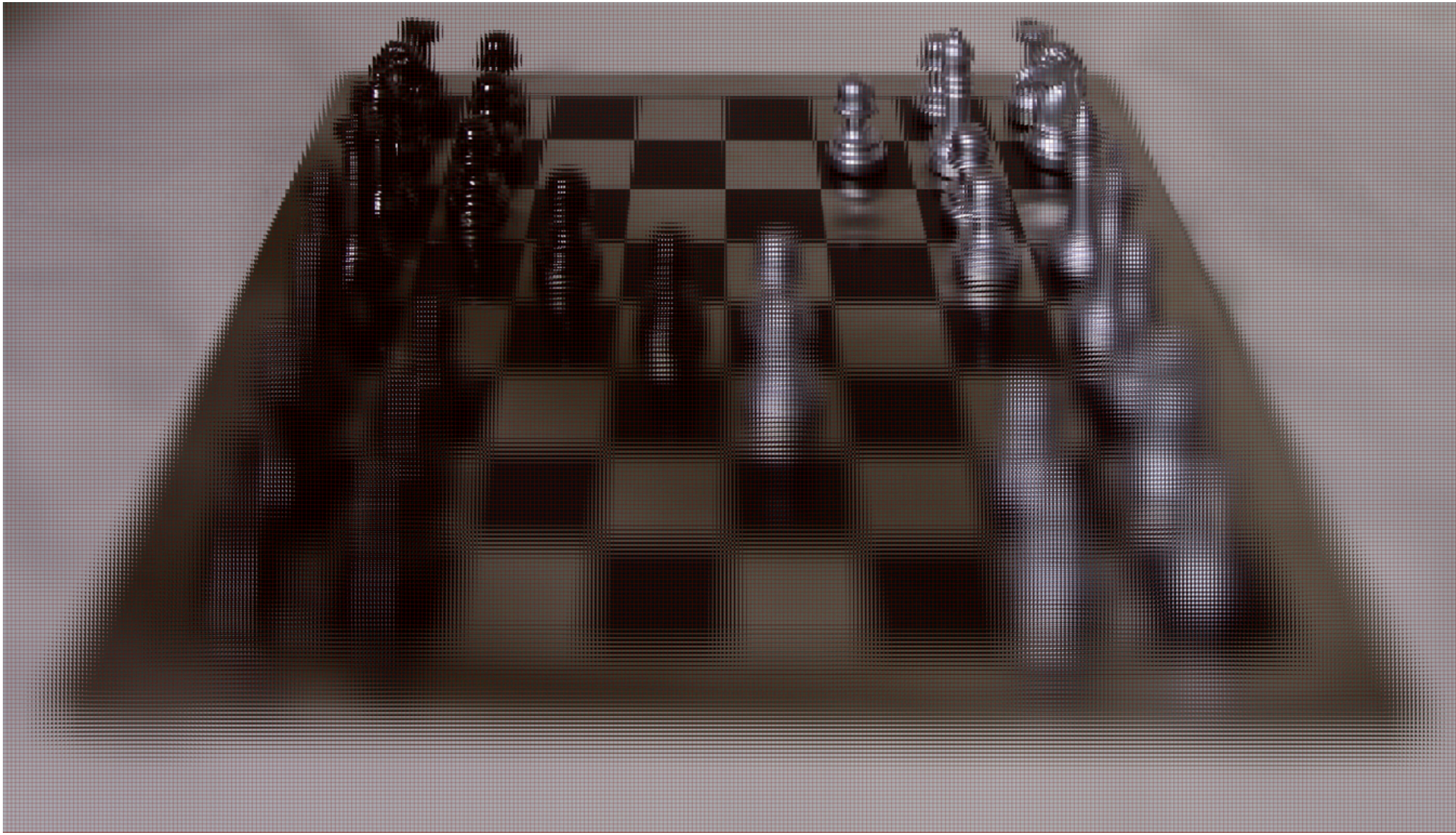
Algorithm

- **For all aperture locations u, v in [-1, 1]**
  - translate image along y, x by ku, kv
  - add to final image
- **Note: we ignore a scale factor (recall, field of view changes with refocusing)**

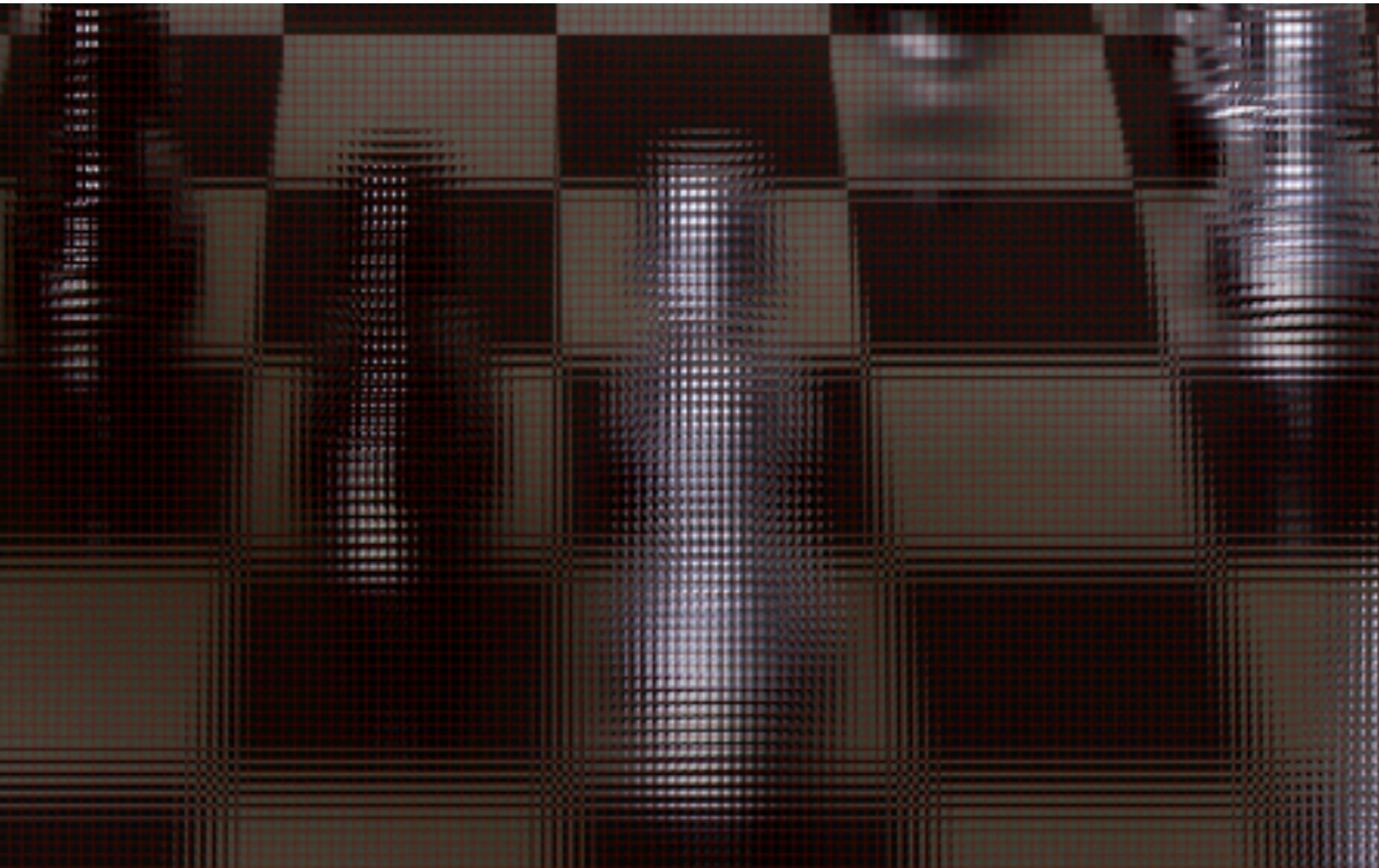


Input light field

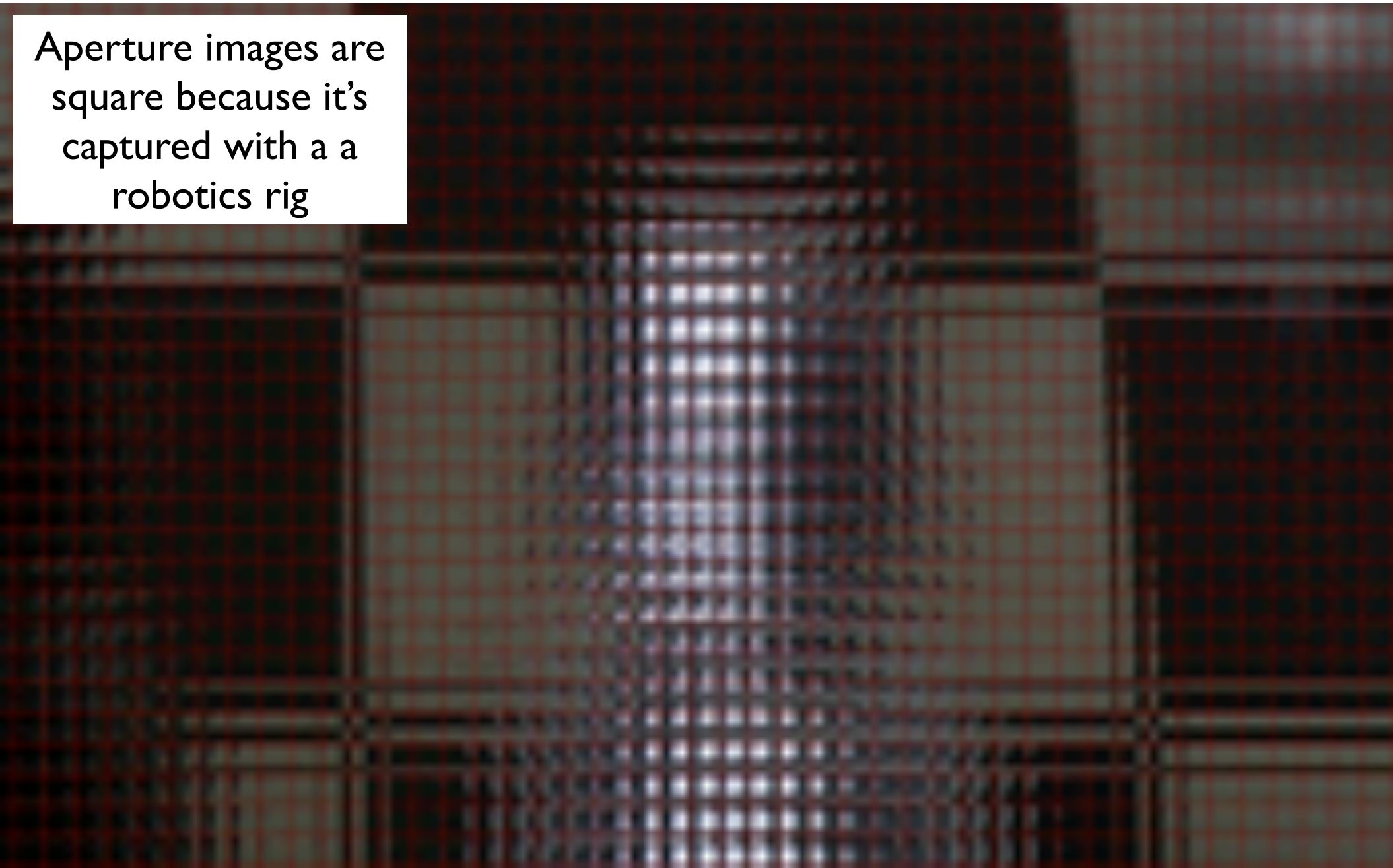
- Yes, the focal stack was synthetic!



Input light field

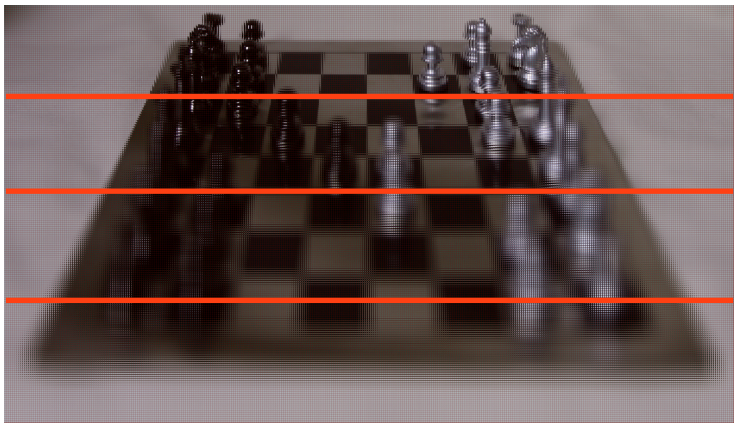


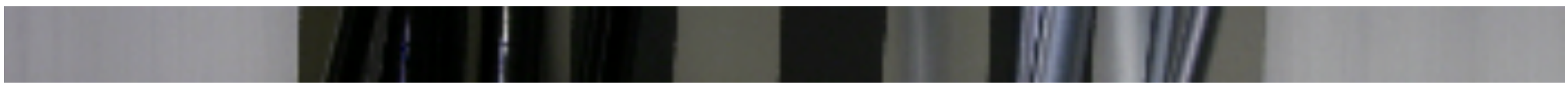
Input light field

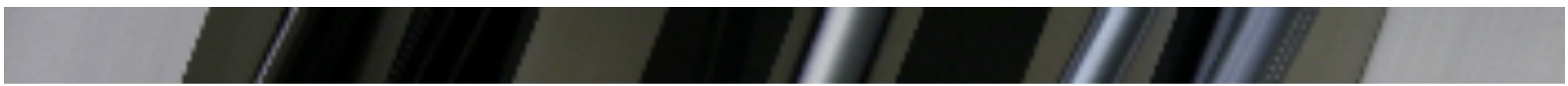


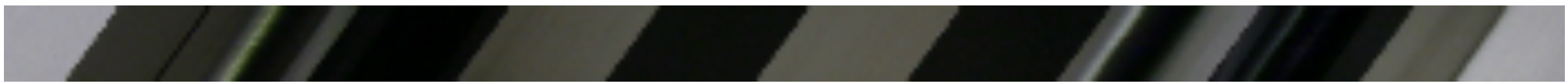
Epipolar slices

- y, v = constant









Results



Results



Results



Results



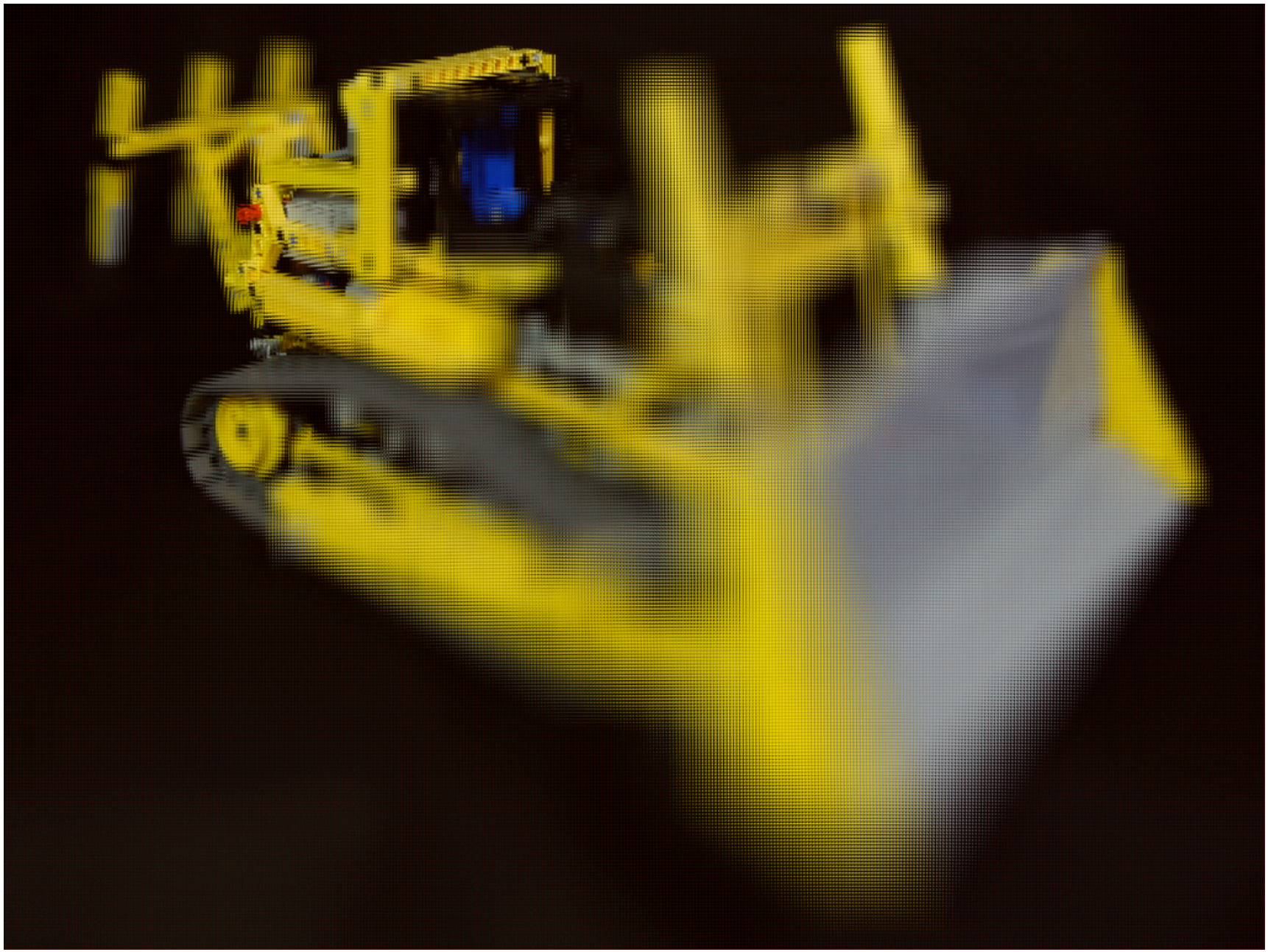
Results



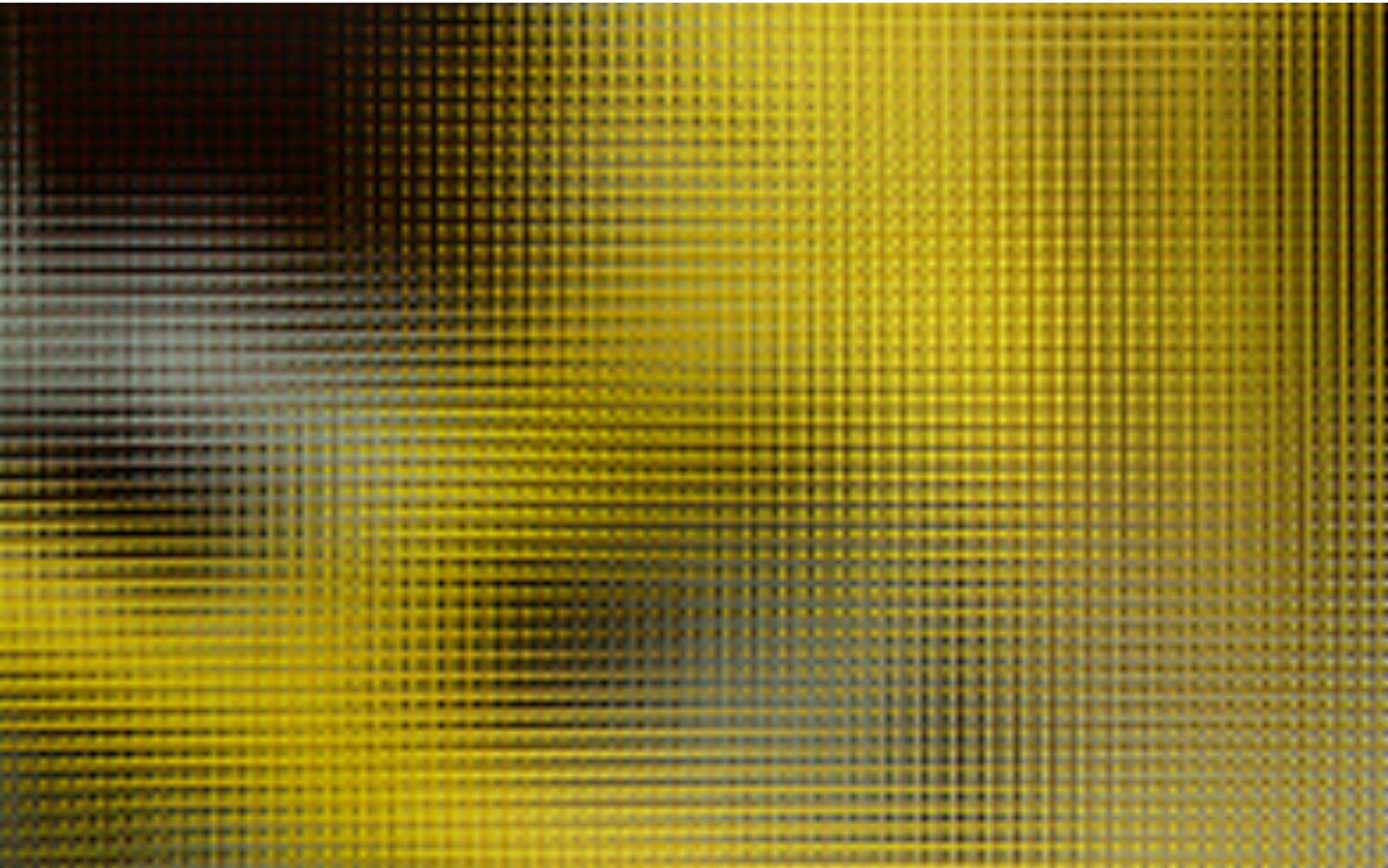
Results



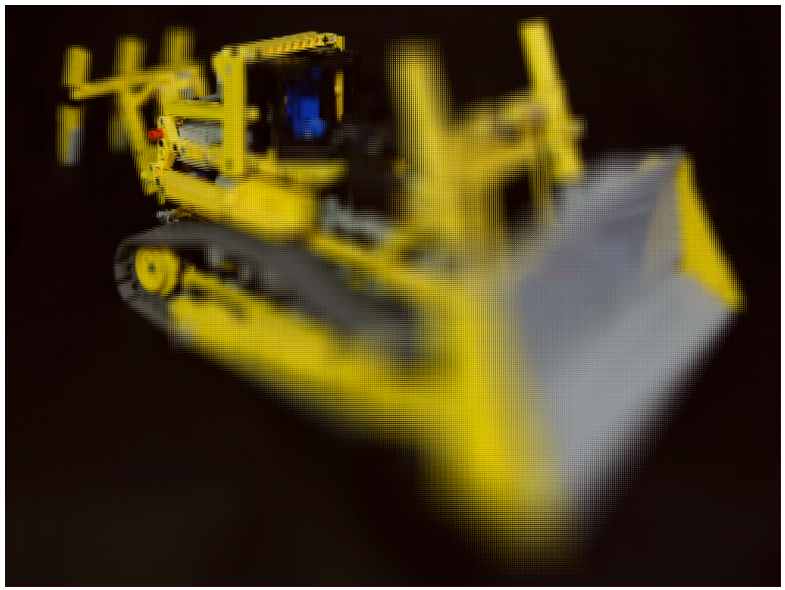
More results

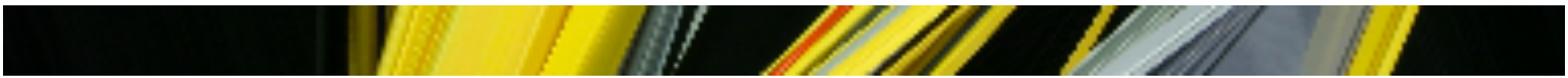


More results



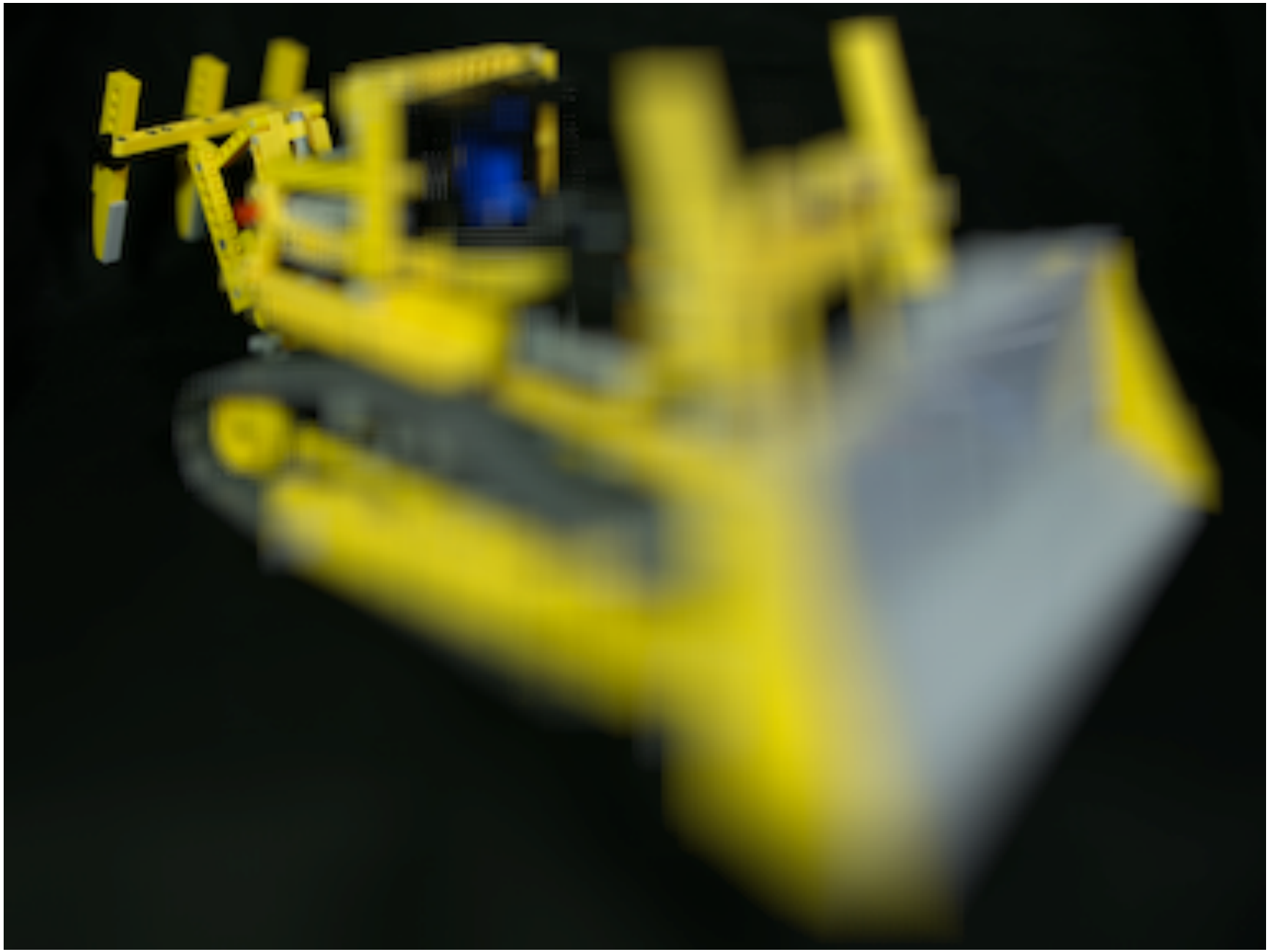
Epipolar slice



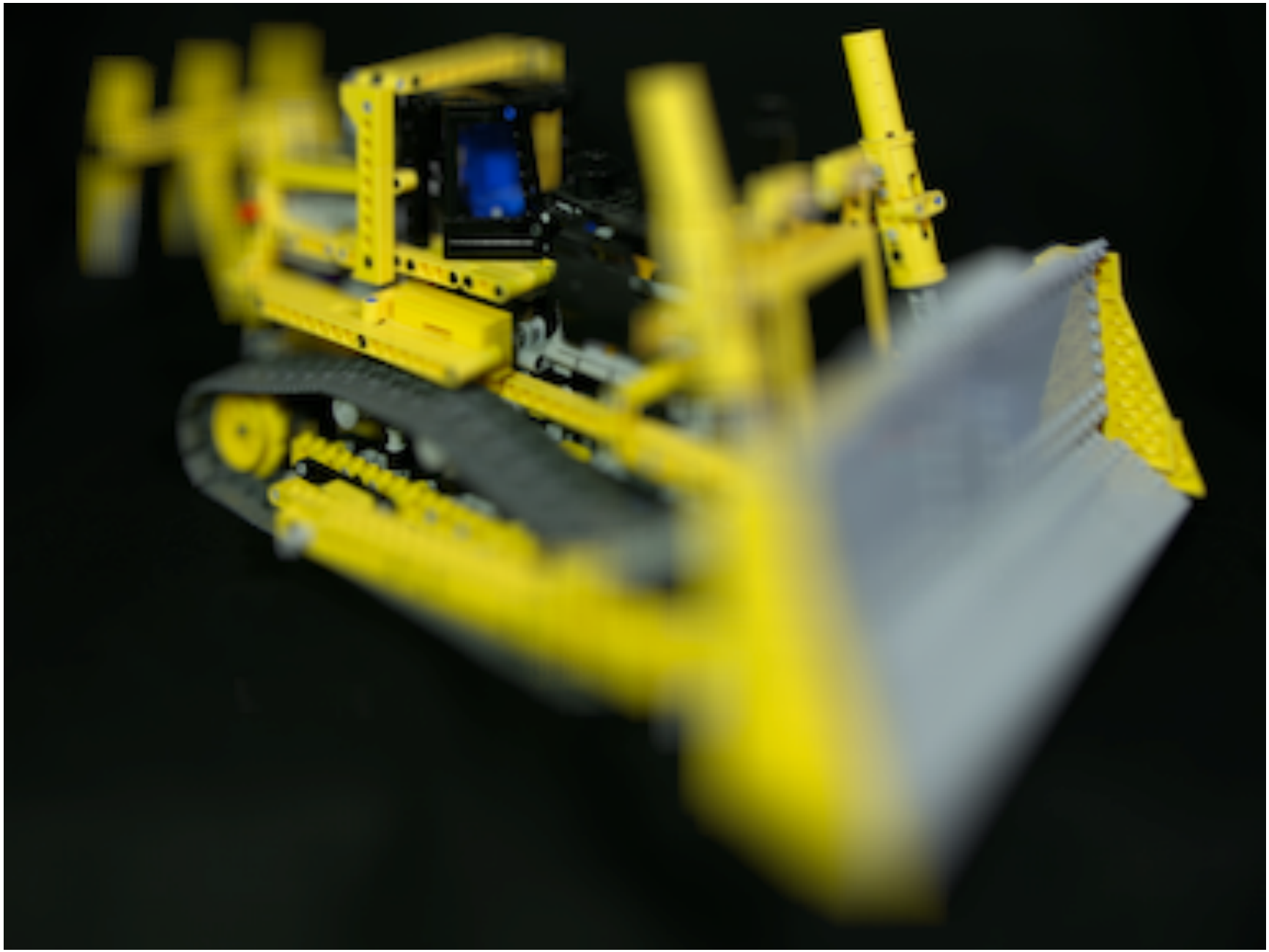




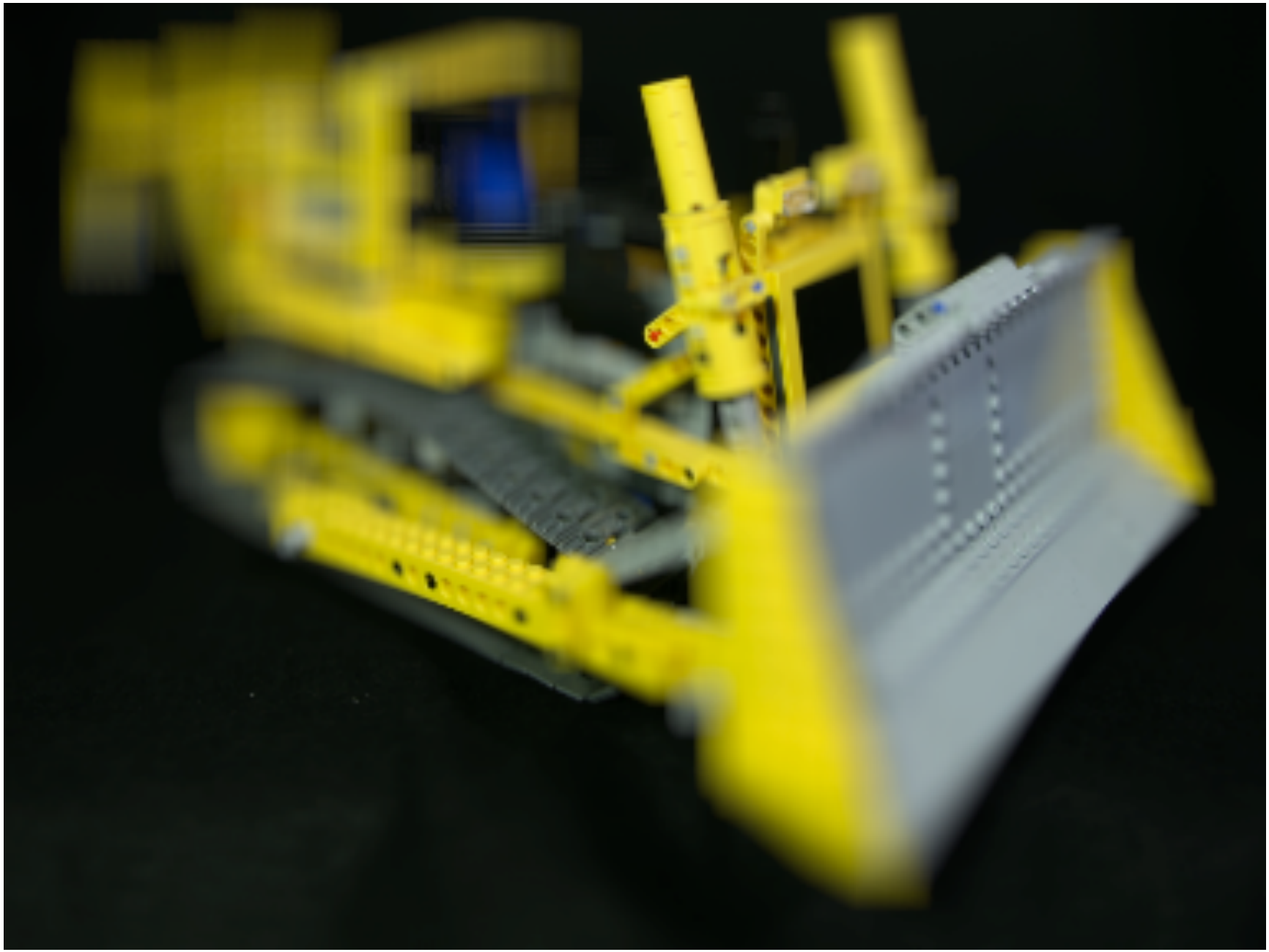
More results



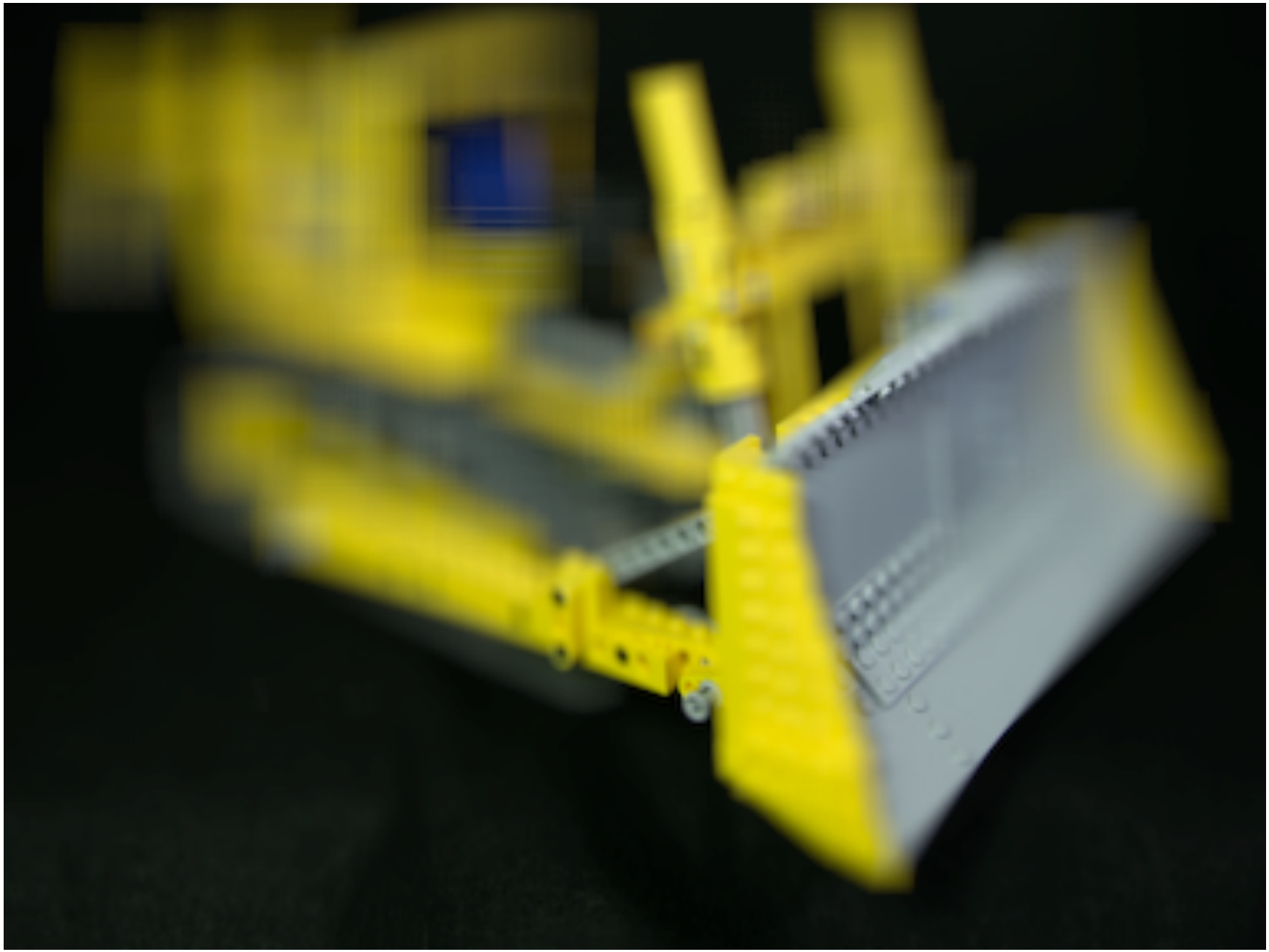
More results



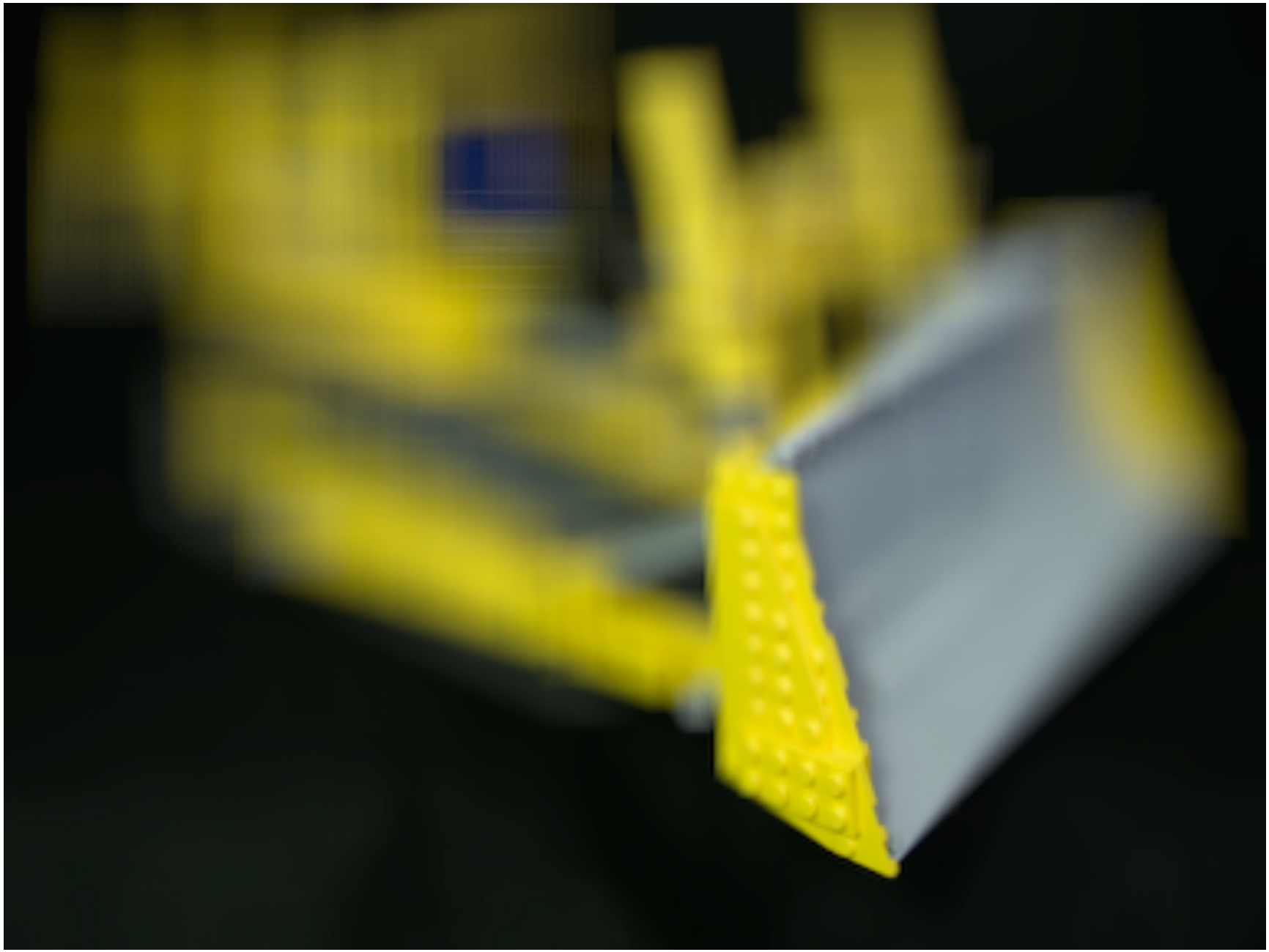
More results



More results



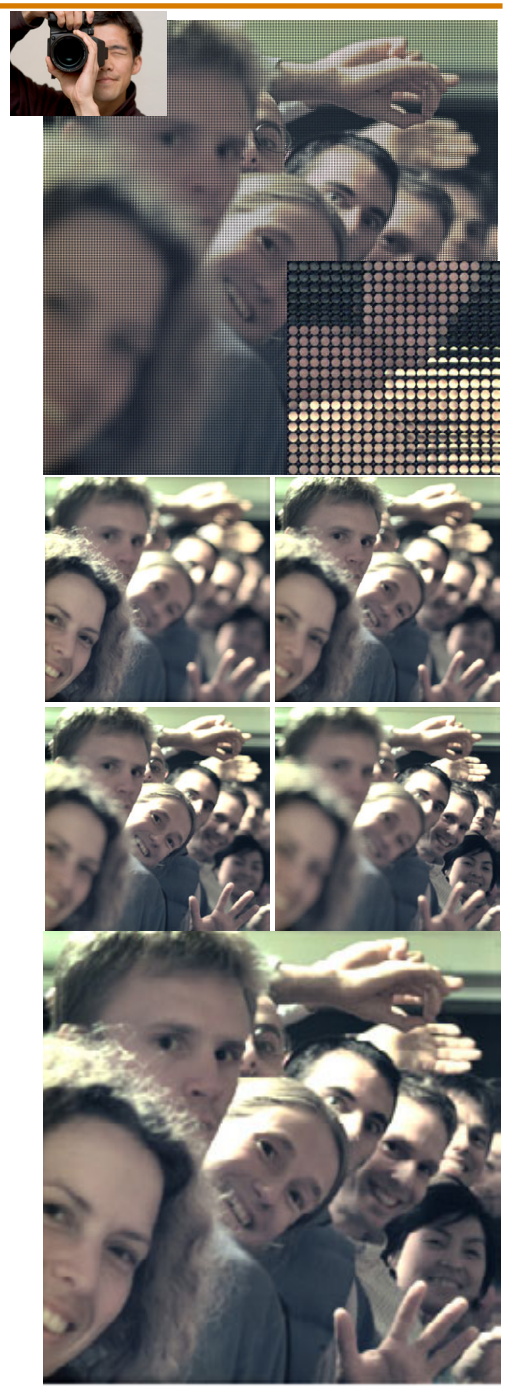
More results



Focal stack & plenoptic camera

Light Field Photography with a Hand-Held Plenoptic Camera, Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz, Pat Hanrahan

- **See next time**
- **Capture light field**
- **Refocus to create focal stack**
- **Use photomontage to generate all-focus image**



Focal stack & plenoptic camera



Figure 15: *Left*: Extended depth of field computed from a stack of photographs focused at different depths. *Right:* A single sub-aperture image, which has equal depth of field but is noisier.

From Ng et al. http://graphics.stanford.edu/papers/lfcamera/

Autofocus

Phase detection focusing

- **Used e.g. in SLRs**

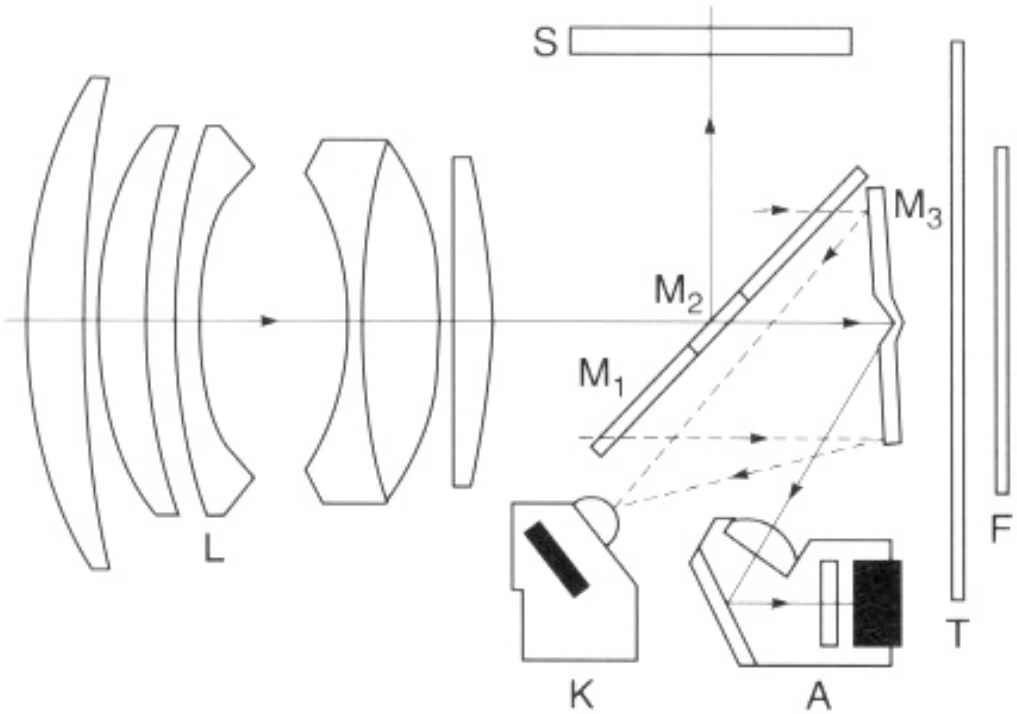
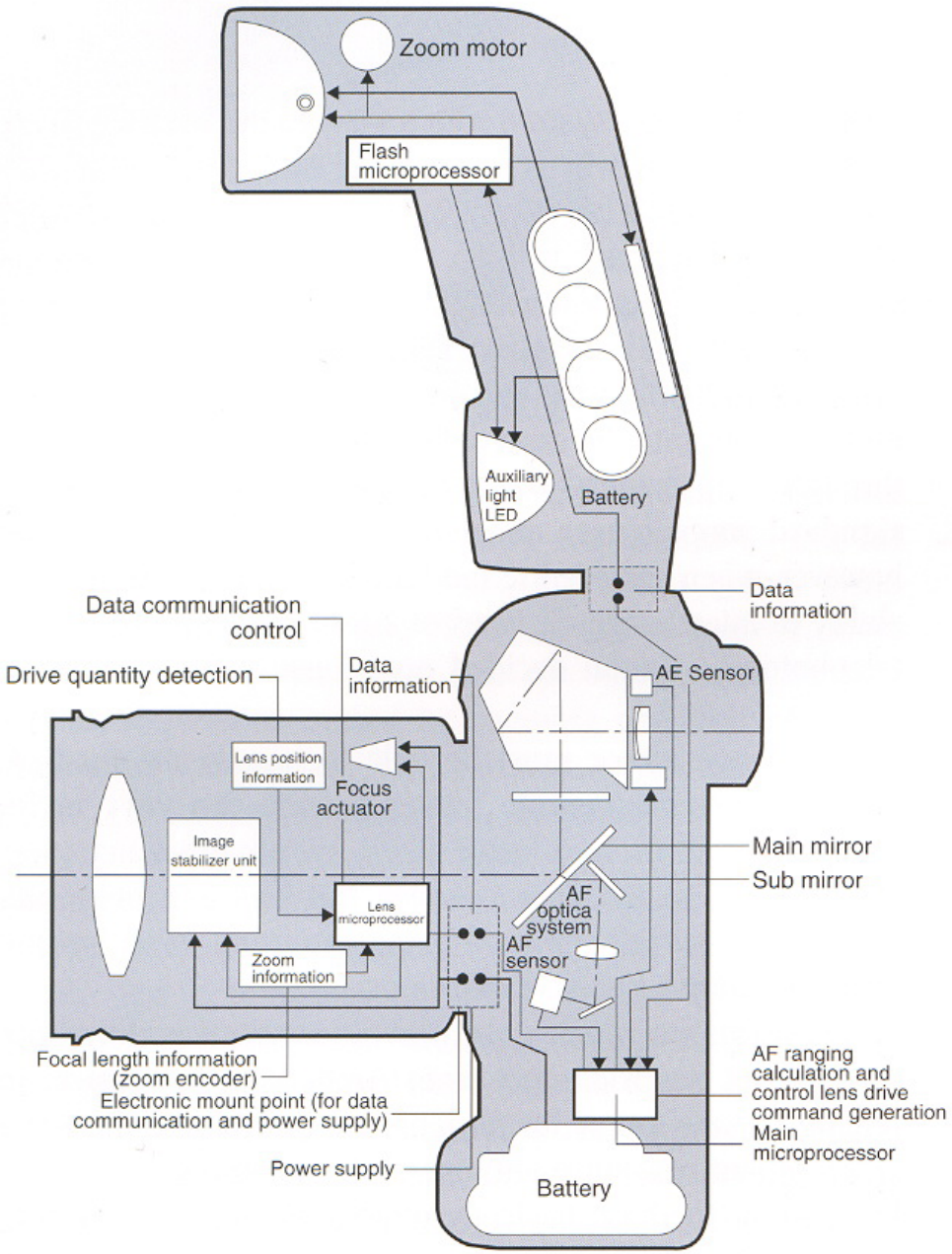


Figure 9.24 Location of autofocus and metering modules. L, camera lens; S, focusing screen; F, film in gate; M_1, reflex mirror with 30 per cent transmission; M_2, central region with 50 per cent transmission; M_3, secondary mirror with two focusing regions; A, autofocus module; K, metering module; spot or centre-weighted

From The Manual of Photography



From the Canon red book

Phase detection focusing

- **Stereo vision from two portions of the lens on the periphery**
- **Not at the equivalent film plane but farther ➔ can distinguish too far and too close**
- **Look at the phase difference between the two images**

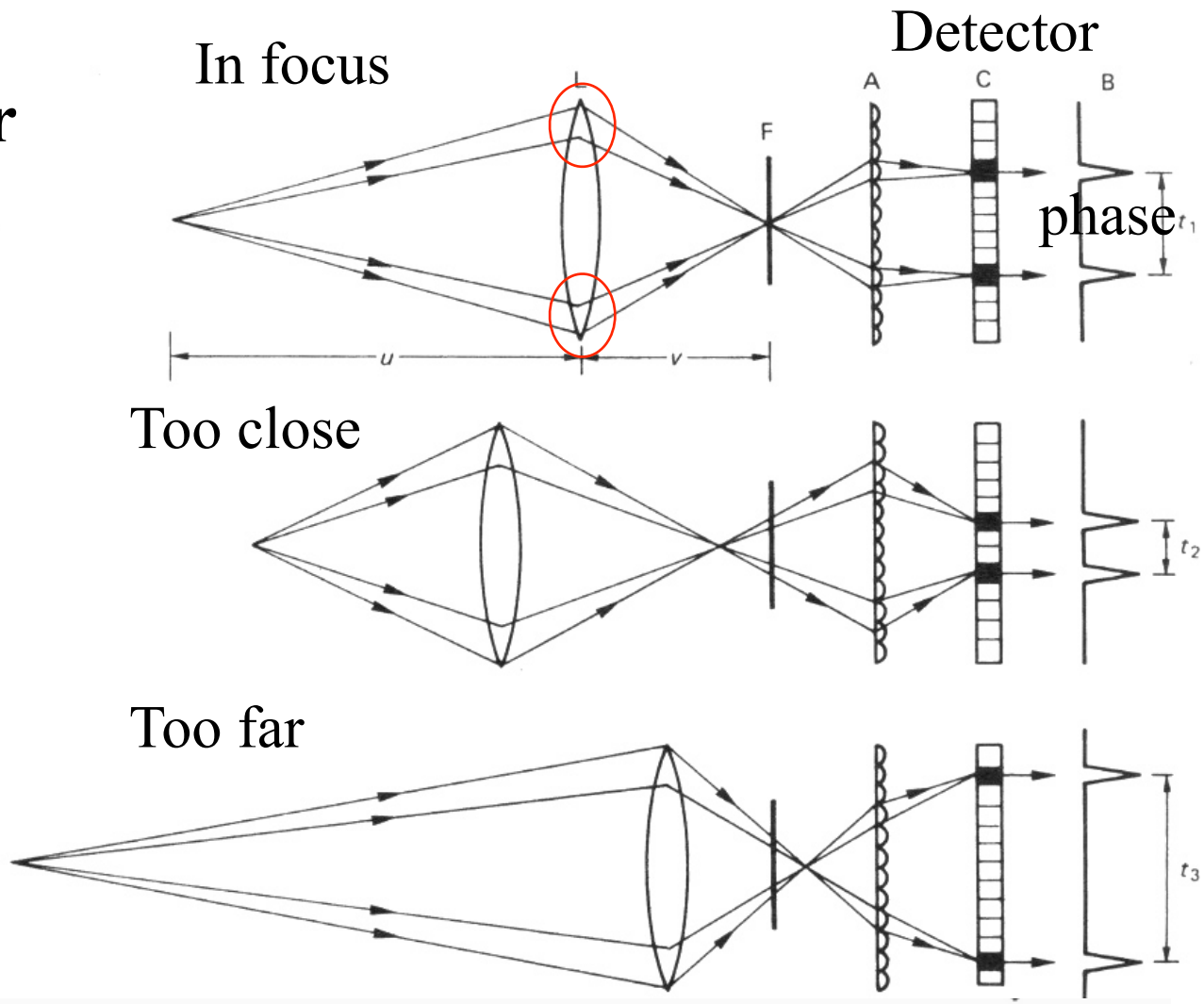
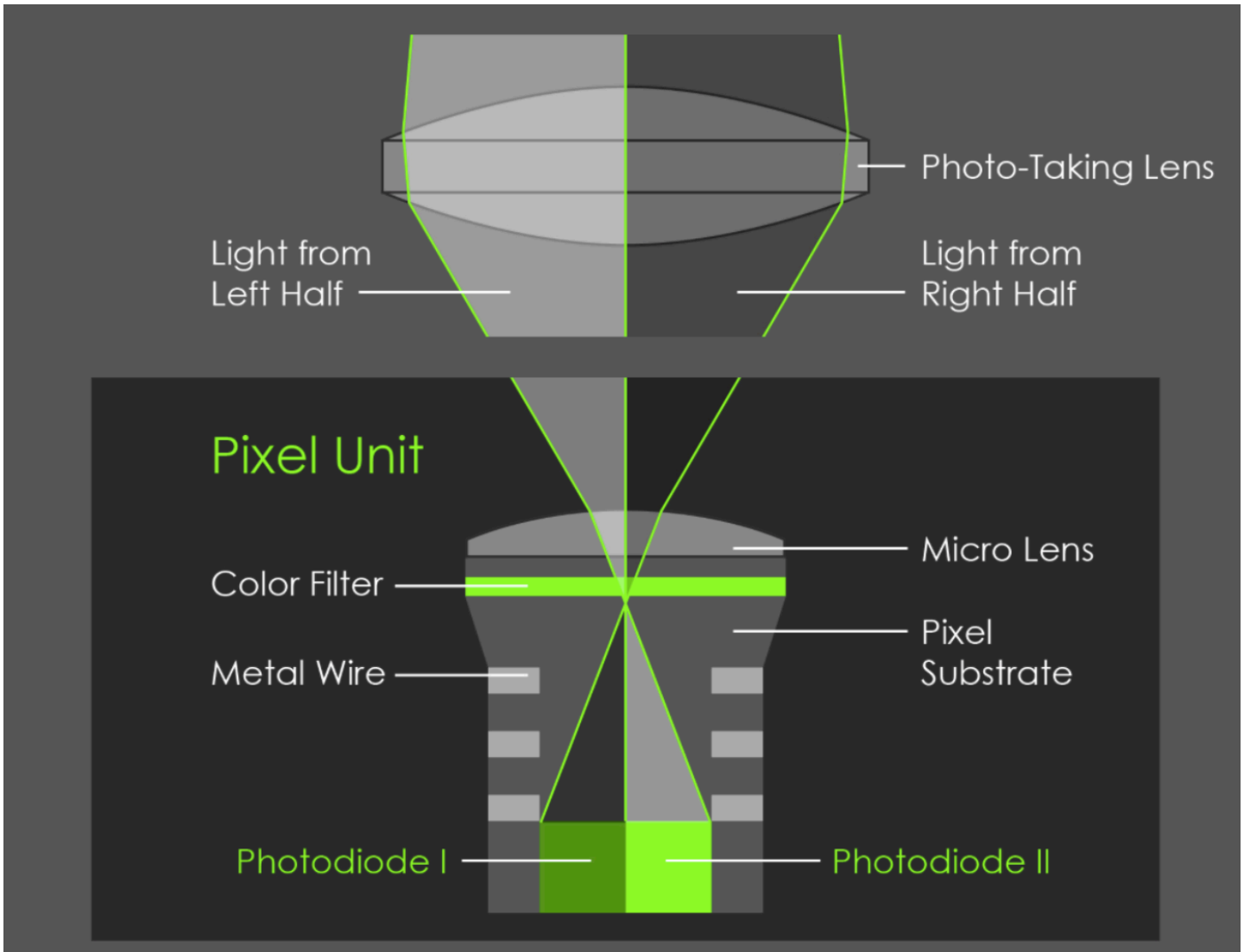


Figure 9.25 Principles of autofocus by phase detection. (a) Subject in focus. (b) Focus in front of subject. (c) Focus beyond subject. Key: L camera lens; F equivalent focal plane; A lenslet array; C CCD linear array; B output signals with time delay t_1 etc.

From The Manual of Photography

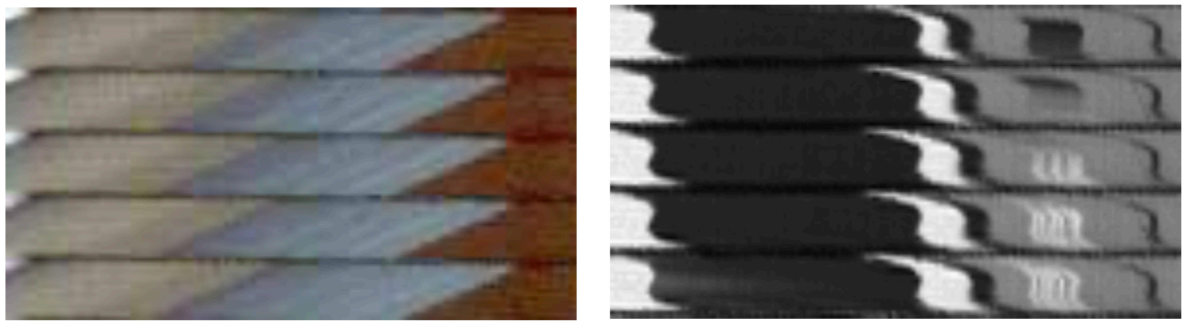
On-sensor phase autofocus

- **https://research.googleblog.com/2017/10/portrait-mode-on-pixel-2-and-pixel-2-xl.html**



Optical aberration correction

- **The less known (and possibly most important) chapter in Ren Ng's thesis.**
- **https://stanford.edu/class/ee367/reading/Ren%20Ng%20thesis%20Lytro.pdf**
- **https://www.spiedigitallibrary.org/conference-proceedings-of-spie/6342/63421E/Digital-correction-of-lens-aberrations-in-light-field-photography/10.1117/12.692290.full**



(A): Six-element lens (B): Single-element lens

Figure 7.3: Comparison of epipolar images with and without lens aberrations.

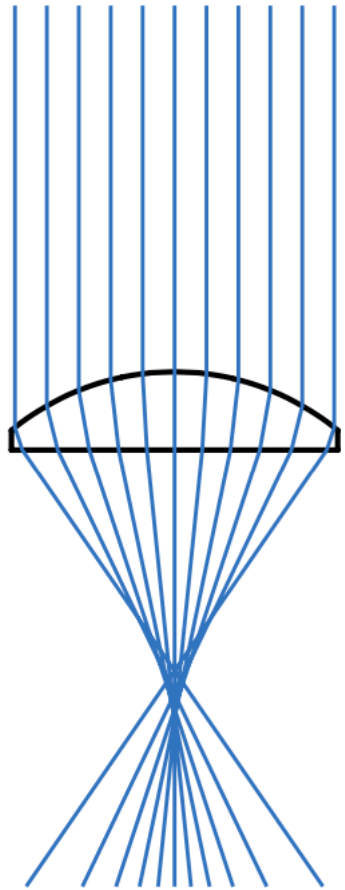


Figure 7.1: Spherical aberration.

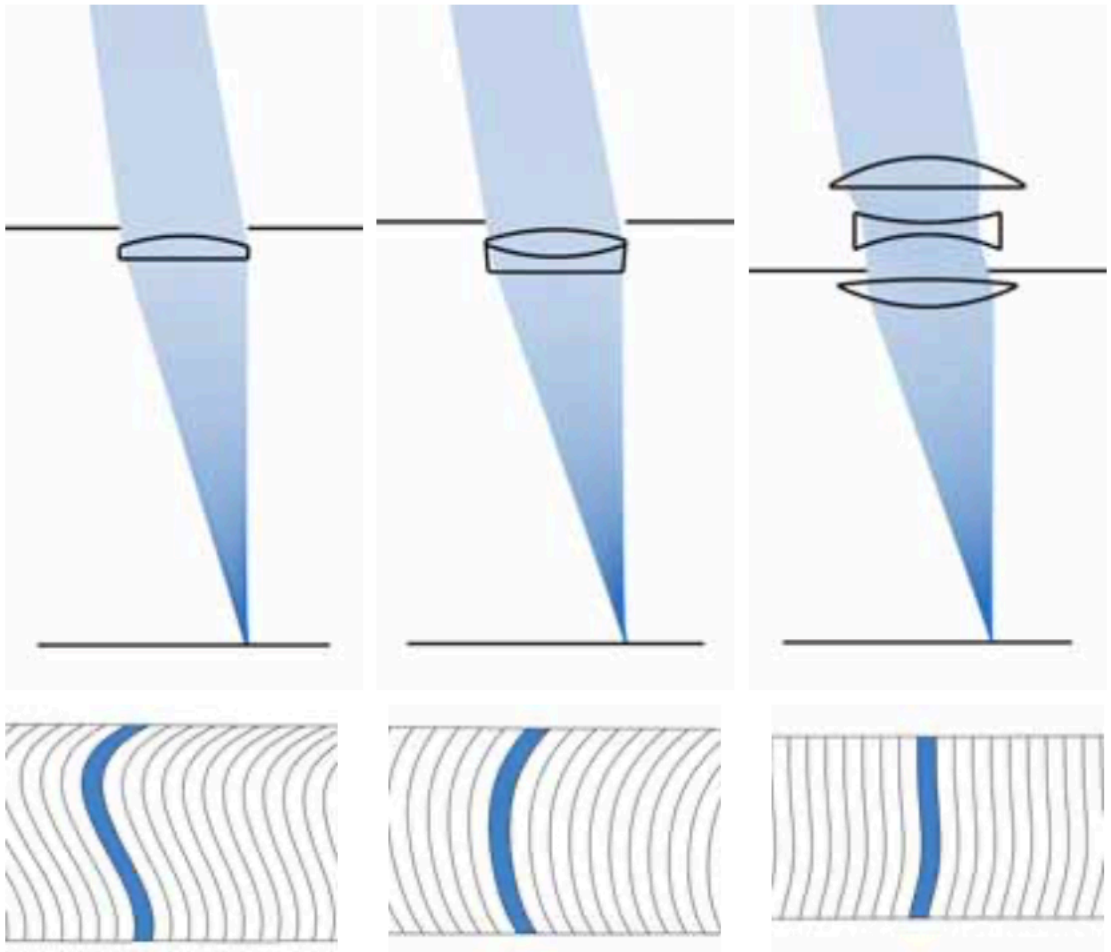


Figure 7.6: Classical reduction in aberrations by adding glass elements to the lens.

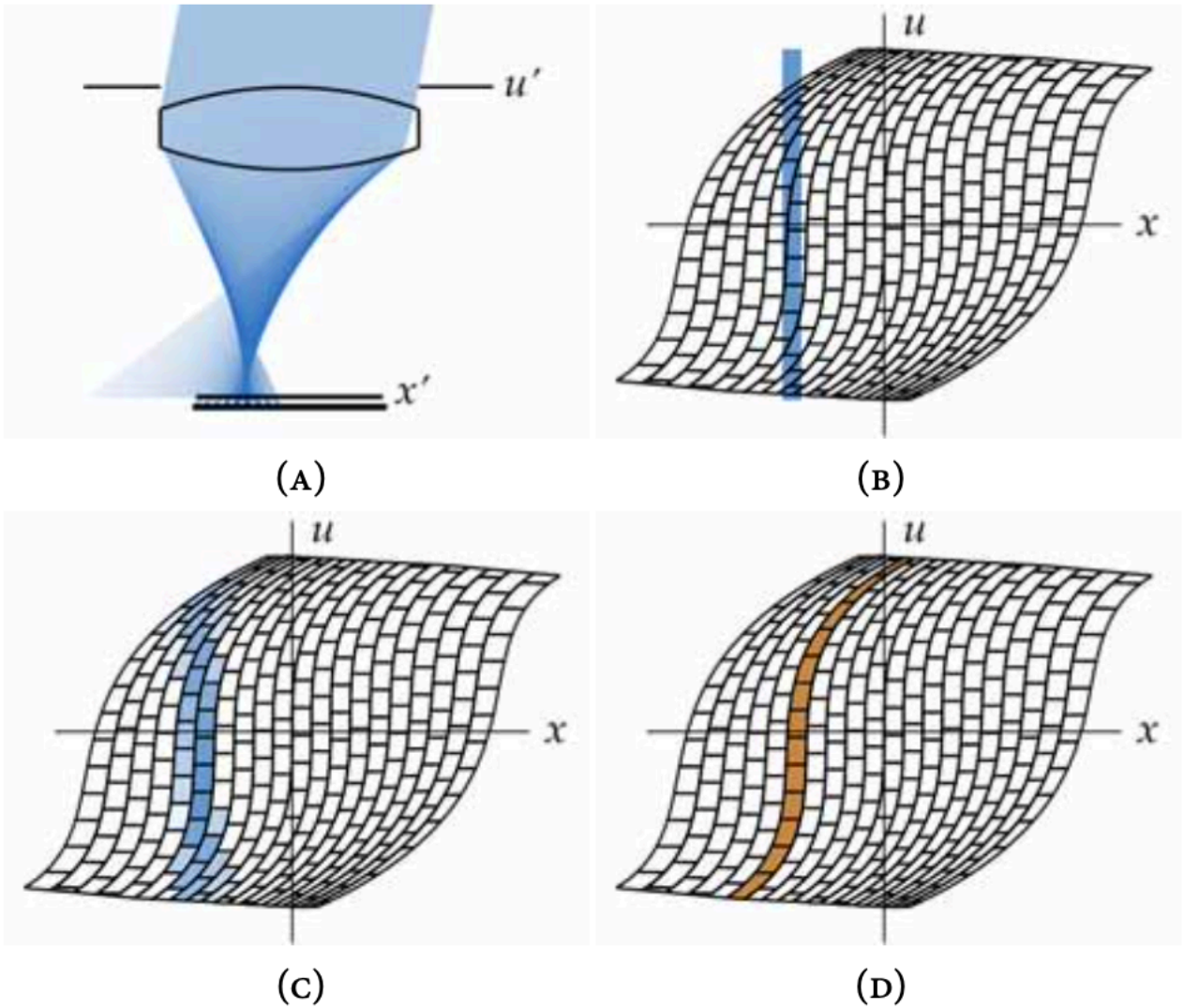
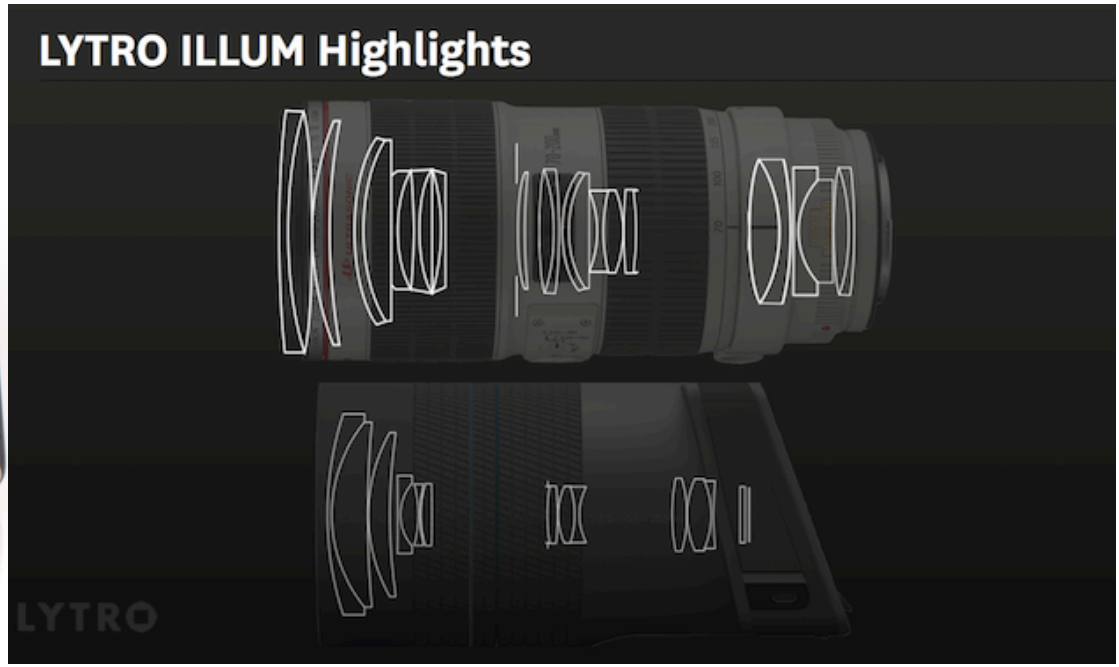


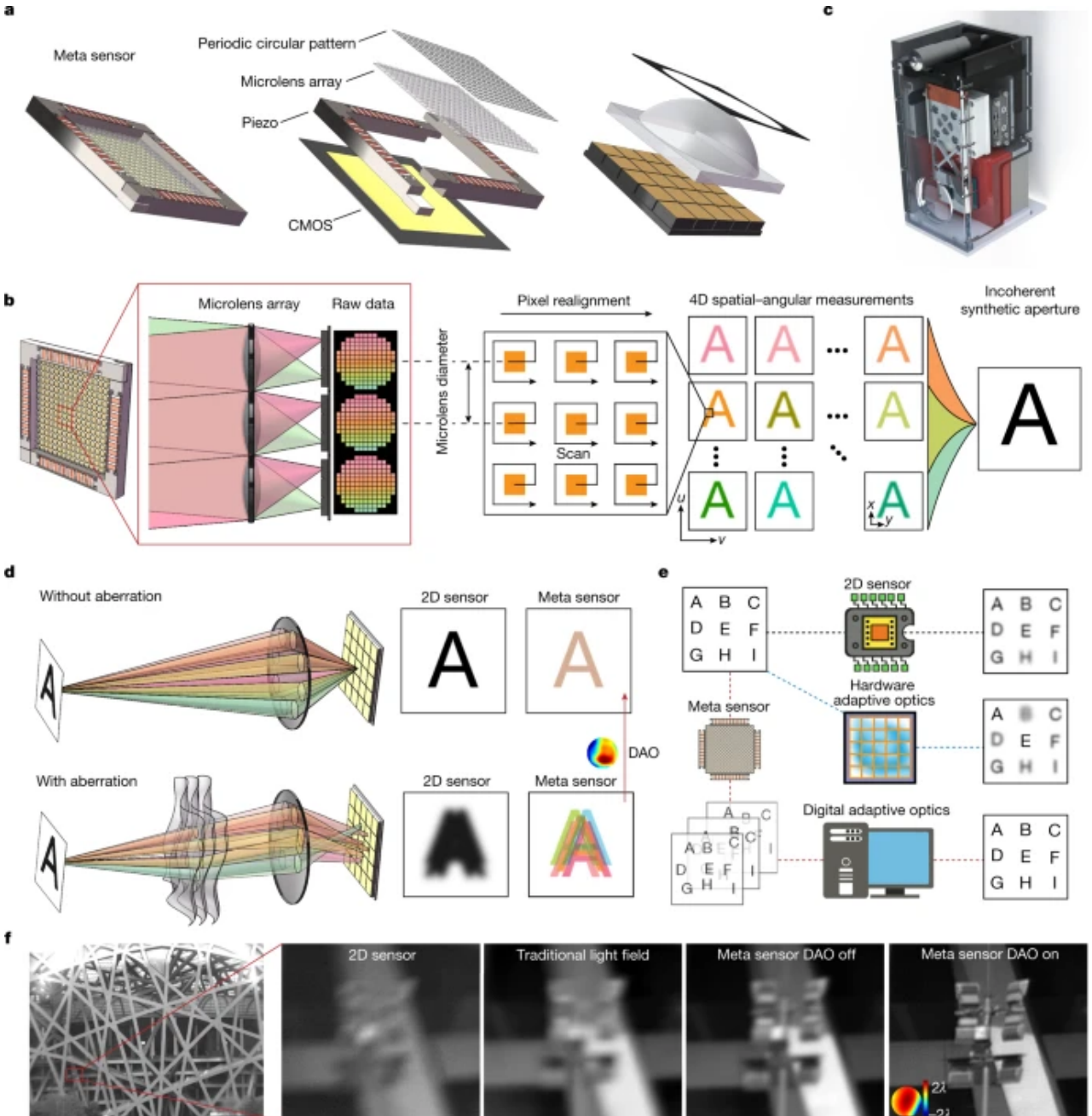
Figure 7.7: Ray-space illustration of digital correction of lens aberrations.




LYTRO ILLUM Highlights
LYTRO

Related work

- https://www.nature.com/articles/s41586-022-05306-8



Article

An integrated imaging sensor for aberration-corrected 3D photography

https://doi.org/10.1038/s41586-022-05306-8
Received: 9 November 2021
Accepted: 1 September 2022
Open access
Check for updates

Jiamin Wu[1,2,3,6], Yuduo Guo[4,5,6], Chao Deng[1,3,6], Anke Zhang[3,5], Hui Qiao[1,3], Zhi Lu[1,3], Jiachen Xie[1,3], Lu Fang[2,3,4,5] ✉ & Qionghai Dai[1,2,3] ✉

Planar digital image sensors facilitate broad applications in a wide range of areas[1–5], and the number of pixels has scaled up rapidly in recent years[2,6]. However, the practical performance of imaging systems is fundamentally limited by spatially nonuniform optical aberrations originating from imperfect lenses or environmental disturbances[7,8]. Here we propose an integrated scanning light-field imaging sensor, termed a meta-imaging sensor, to achieve high-speed aberration-corrected three-dimensional photography for universal applications without additional hardware modifications. Instead of directly detecting a two-dimensional intensity projection, the meta-imaging sensor captures extra-fine four-dimensional light-field distributions through a vibrating coded microlens array, enabling flexible and precise synthesis of complex-field-modulated images in post-processing. Using the sensor, we achieve high-performance photography up to a gigapixel with a single spherical lens without a data prior, leading to orders-of-magnitude reductions in system capacity and costs for optical imaging. Even in the presence of dynamic atmosphere turbulence, the meta-imaging sensor enables multisite aberration correction across 1,000 arcseconds on an 80-centimetre ground-based telescope without reducing the acquisition speed, paving the way for high-resolution synoptic sky surveys. Moreover, high-density accurate depth maps can be retrieved simultaneously, facilitating diverse applications from autonomous driving to industrial inspections.

Fredo's vision

Practical matters

Do you lose all that resolution?

- **It's complicated**

Can you do video

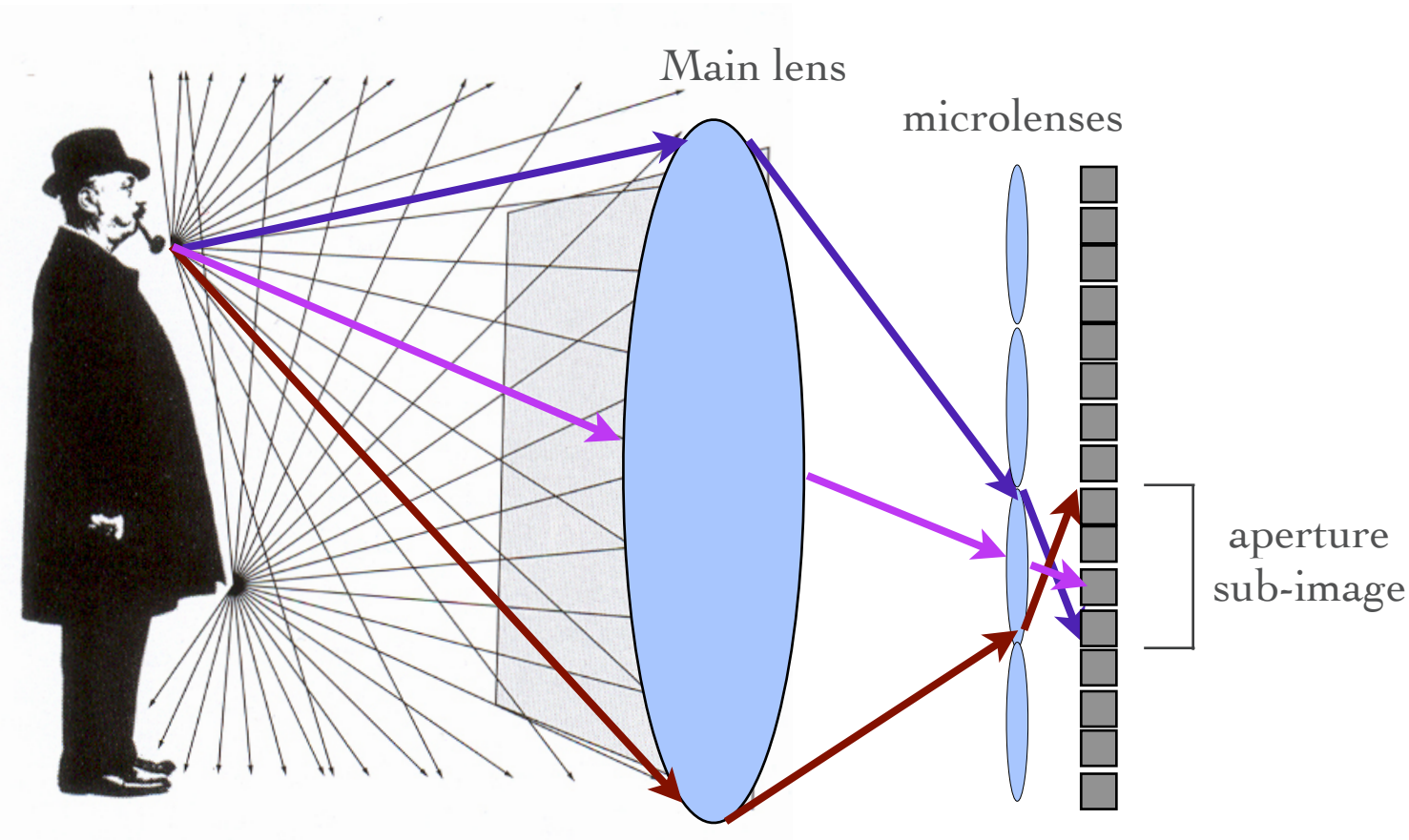
- **Bandwidth is the challenge**

Is it practical?

- **Full frame sensor with cell phone density**

- **24x36mm ->**
- **Sony a7riii has 4.5micron pitch**
- **Bandwidth of phantom high speed camera: 6,600 fps at 2048 x 1920 4MPixel. -> 660 at 40MPixels -> 30fps at 40*4.5*4.5 [they say 26GPixels/sec] 12 bits**
- **XQD: 400MB/s -> budget of 1/8bit per pixel at full burst**
- **JPEG vs RAW: 70MB vs 7MB; 1 to 10 -> gets us still 1 order of magnitude short**
- **Big challenge: quick, local light field compression. Can dimensionality gap help? Steer DCT?**
- **Alternative: 2 microns, factor of 4 obviously, 10MPixels with 4.5x4.5**

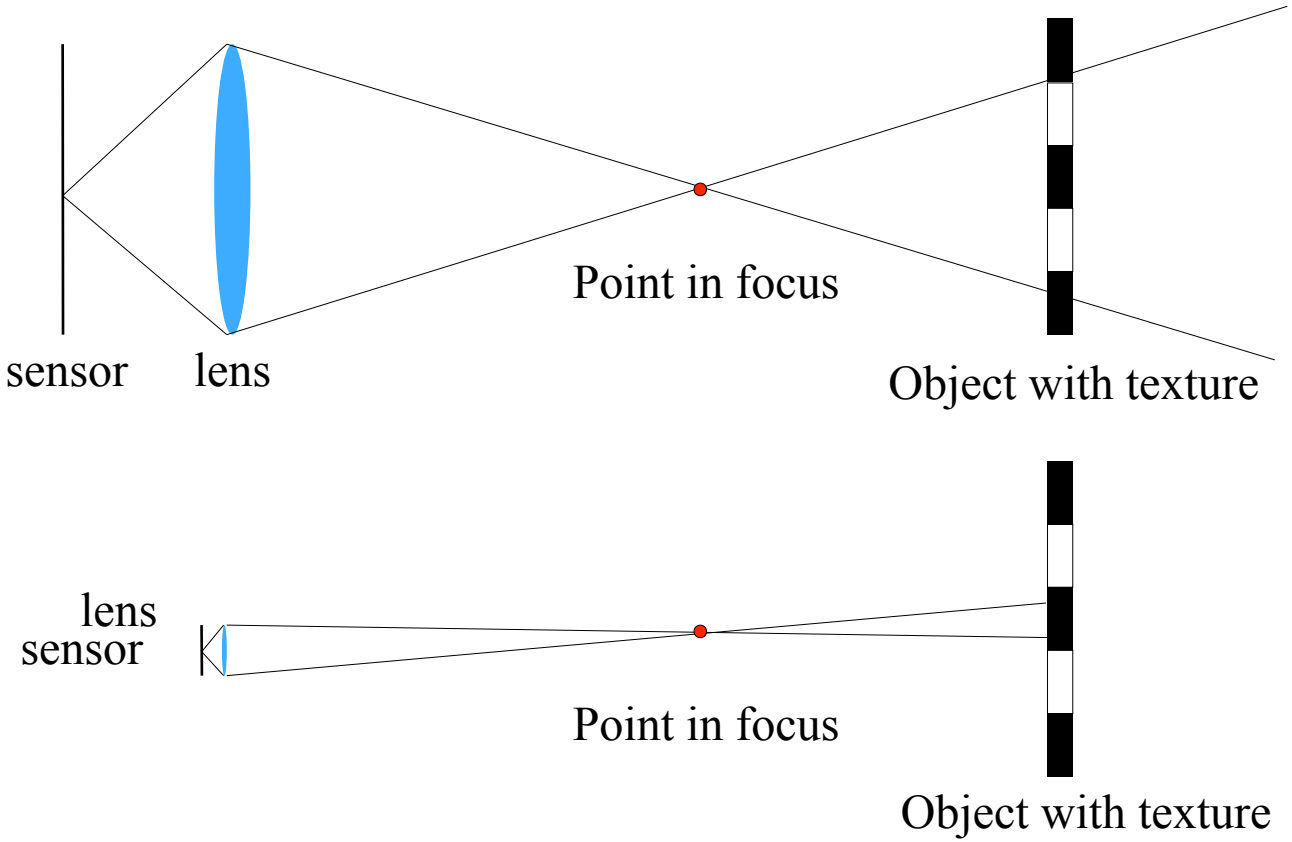
Diaphragm (or lack thereof)



Fake depth of field for cell phones (no light field)

Challenge

- **Shallow depth of field hard with small sensors**



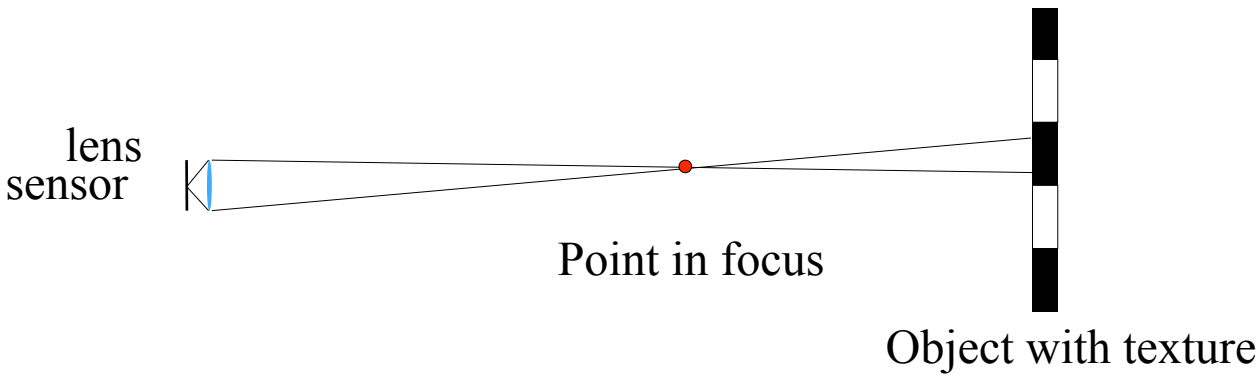
Challenge



(a) large sensor (22.2 x 14.8 mm)



(b) small sensor (7.18 x 5.32 mm)



- **https://people.csail.mit.edu/soonmin/dof/**

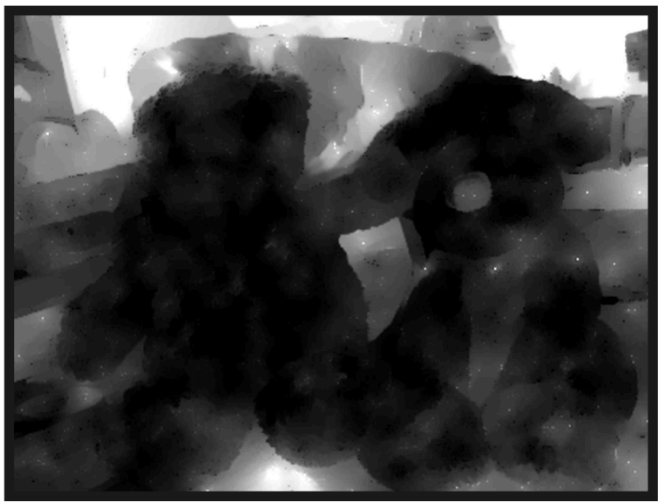
Defocus Magnification

Soonmin Bae Frédo Durand

Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology



(a) input



(b) defocus map



(c) our result with magnified defocus

Figure 1: *Our technique magnifies defocus given a single image. Our defocus map characterizes blurriness at edges. This enables shallow depth of field effects by magnifying existing defocus. The input photo was taken by a Canon PowerShot A80, a point-and-shoot camera with a sensor size of 7.18 × 5.32 mm, and a 7.8 mm lens at f/2.8.*

Portrait mode

- https://research.google/blog/portrait-mode-on-the-pixel-2-and-pixel-2-xl-smartphones/
- https://arxiv.org/pdf/1806.04171





- **Machine learning-based foreground-background segmentation**
- **Use dual pixels to get a depth map (or analyze local defocus)**
- **Predict desired per-pixel blur radius**
- **Render with non-linear blur (more later)**

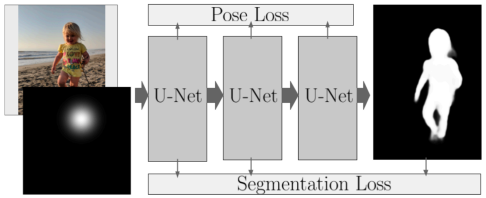
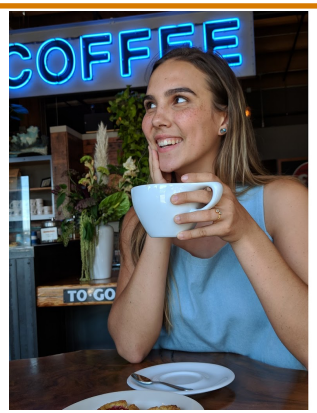
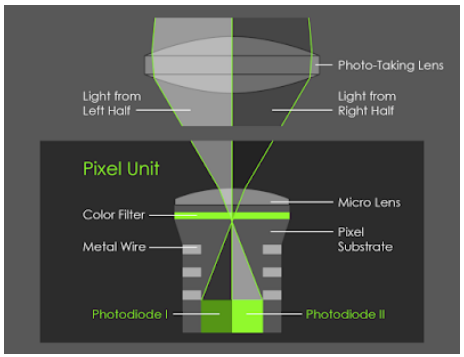


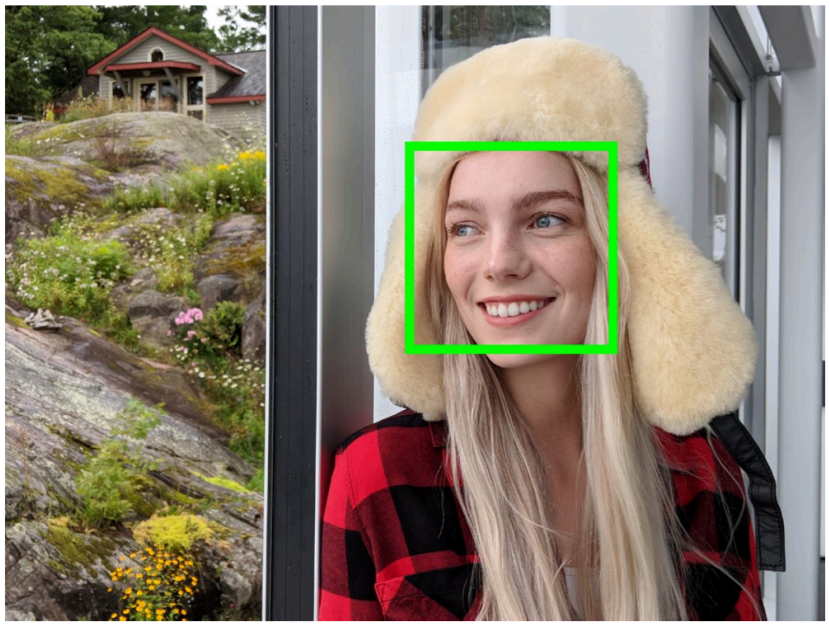
Fig. 3. **Person Segmentation Network.** RGB image (top-left) and face location (bottom-left) are the inputs to a three stage model with pose and segmentation losses after each stage.











(a) Input image with detected face



(b) Person segmentation mask



(c) Mask + disparity from DP



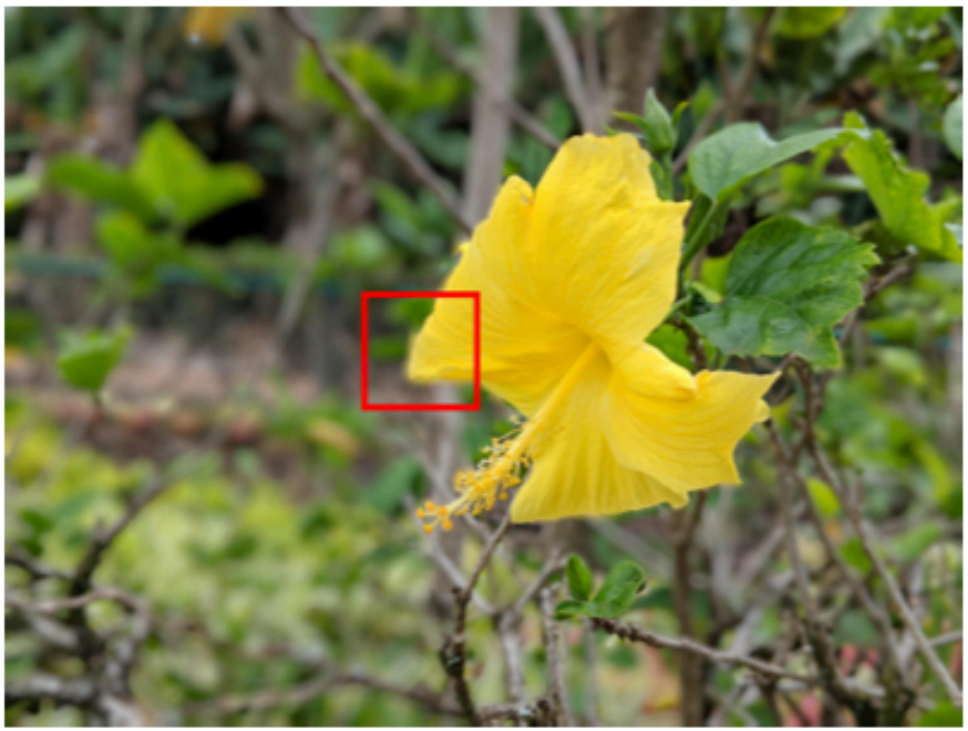
(d) Our output synthetic shallow depth-of-field image

Fig. 1. We present a system that uses a person segmentation mask (b) and a noisy depth map computed using the camera's dual-pixel (DP) auto-focus hardware (c) to produce a synthetic shallow depth-of-field image (d) with a depth-dependent blur on a mobile phone. Our system is marketed as "Portrait Mode" on several Google-branded phones.

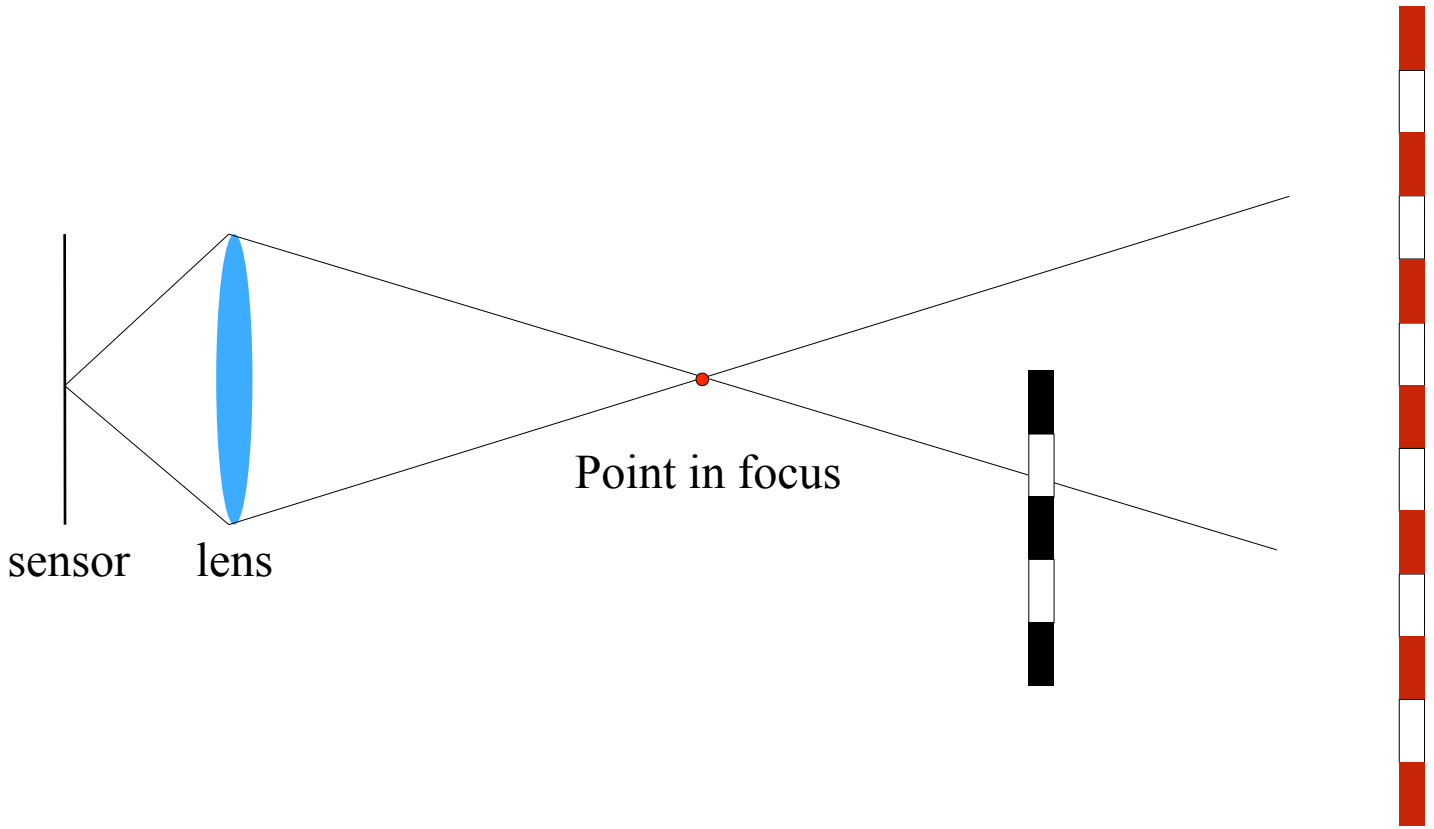
Challenge with defocus blur

- **Remember, not a convolution!**
- **Blur radius depends on input scene depth, not output pixel location**
- **Plus occlusion matters**

Naïve gather






Point in focus
sensor
lens

$$B_{scatter}(\mathbf{x}) = \sum_{\Delta \mathbf{x}} I(\mathbf{x} + \Delta \mathbf{x}) K_{\mathbf{x}+\Delta \mathbf{x}}(-\Delta \mathbf{x}).$$

Occlusion

- **Divide the image into k (=5) depth layers with opacity values**
- **Process each layer independently**
- **Composite back to front (with “over” operator)**


THC




COFFEE
TO GO


COFFEE
TO GO

Challenges

- **Masks are hard, especially around hair!**
- **Need notion of sub-pixel coverage**
-

Some more recent work

- https://arxiv.org/abs/1904.05822
- https://imaging.cs.cmu.edu/dual_pixels/

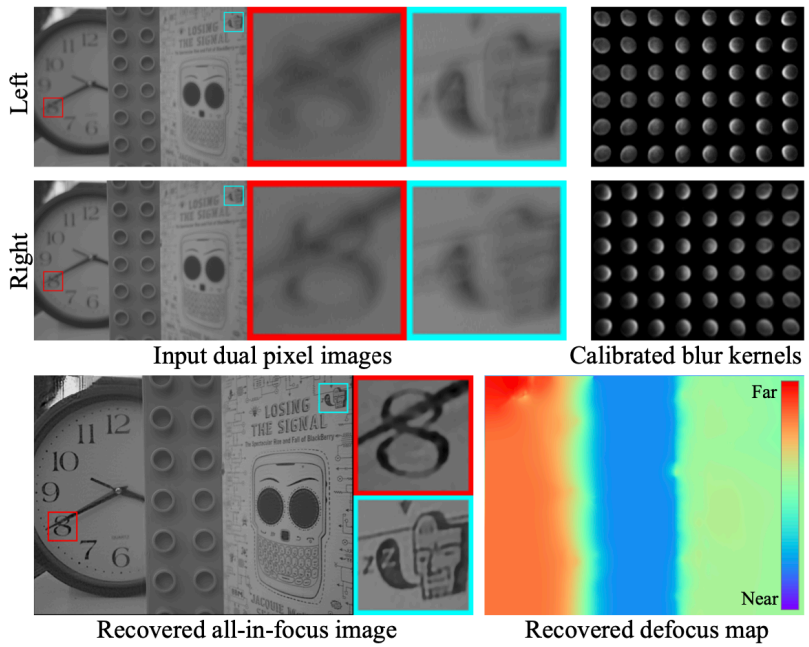


Figure 1: Given left and right dual-pixel (DP) images and corresponding spatially-varying blur kernels, our method jointly estimates an all-in-focus image and defocus map.