TSBB21, Lecture: Specialized Cameras

- Dynamic range extension (HDR)
  - External HDR
  - Internal HDR
    - Log-intensity
    - Dual diode
    - Piecewise linear response
- Seeing in the dark, photomultipliers
  - I) Photomultiplier with only one accelerating plate
  - II) Photomultiplier tube
  - III) Microchannel
- Line cameras
  - General
  - Pushbroom
- Smart cameras

  for machine vision applications
- Event cameras
- Computational cameras
  - Light-field cameras
  - Omni-directional cameras
  - Coded aperture
  - Flexible depth-of-field
- Thanks to:
  - Klas Nordberg. Many slides are similar to his slides.
  - Robert Forchheimer. Many slides are similar to his slides.
  - Hannes Övrén. Some slides on Event cameras are taken from his Guest lecture slides.

Dynamic range extension (HDR)

- A standard camera takes images that can be …
  - Overexposed if we want to see details in the dark areas. Bright areas become “too bright”.
  - Underexposed if we want to see details in the bright areas. Dark areas become “too dark”.
- This becomes a problem if the same scene contains both bright and dark parts.

Dynamic range extension (HDR)

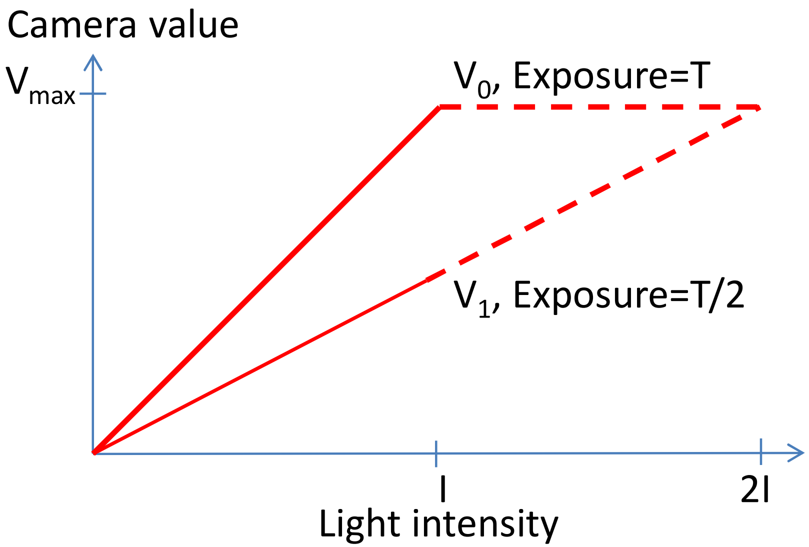
- Several techniques can be used to produce images of higher dynamic range (HDR) than provided by the basic sensor technology
- External HDR (outside the sensor chip)
- Internal HDR (inside the sensor chip)

External HDR

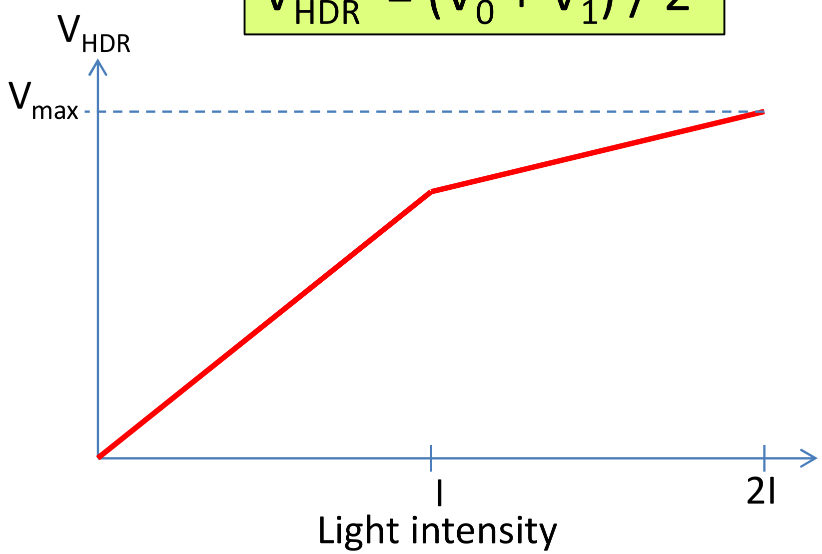
- Take two or more images, one immediately after the other, with different exposure time
- Read out each image as normal
- Produce an HDR image outside the sensor chip by, for example, at each pixel:
  - Combine the measured intensity values from the different exposures, after a suitable normalization.
  - The simplest way to generate a high dynamic range image is to take two images with different exposure times and take the mean of them, see next slide.
- Assumes fast exposure of all images
  - Works fine with CMOS
  - Or static scene

External HDR

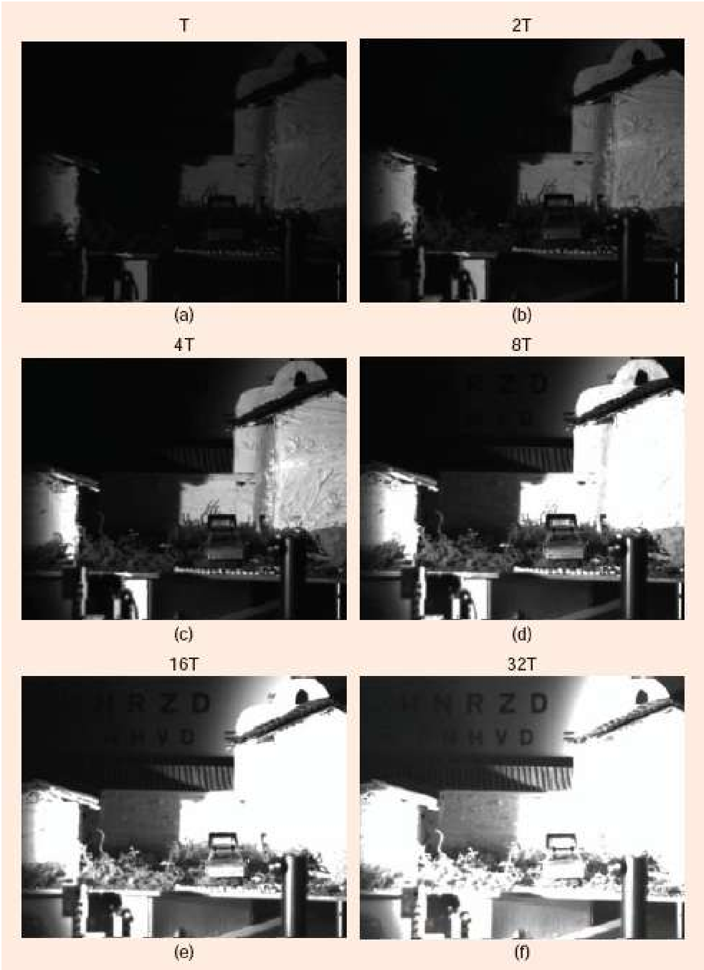
- Simple reconstruction: take the mean of the images.



$$V_{HDR} = (V_0 + V_1) / 2$$



External HDR

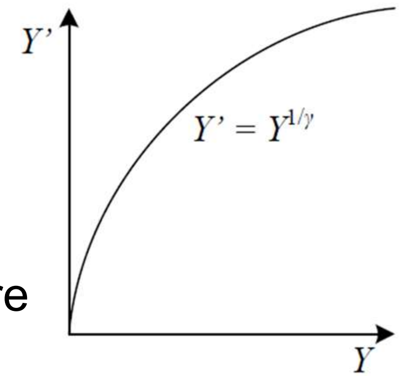


The same principle can be extended to more than two images.
Here, is an HDR image produced from 6 images with different exposure times.

External HDR, Advanced procedure

From:
Richard Szeliski: Computer Vision: Algorithms and Applications, 2nd Edition

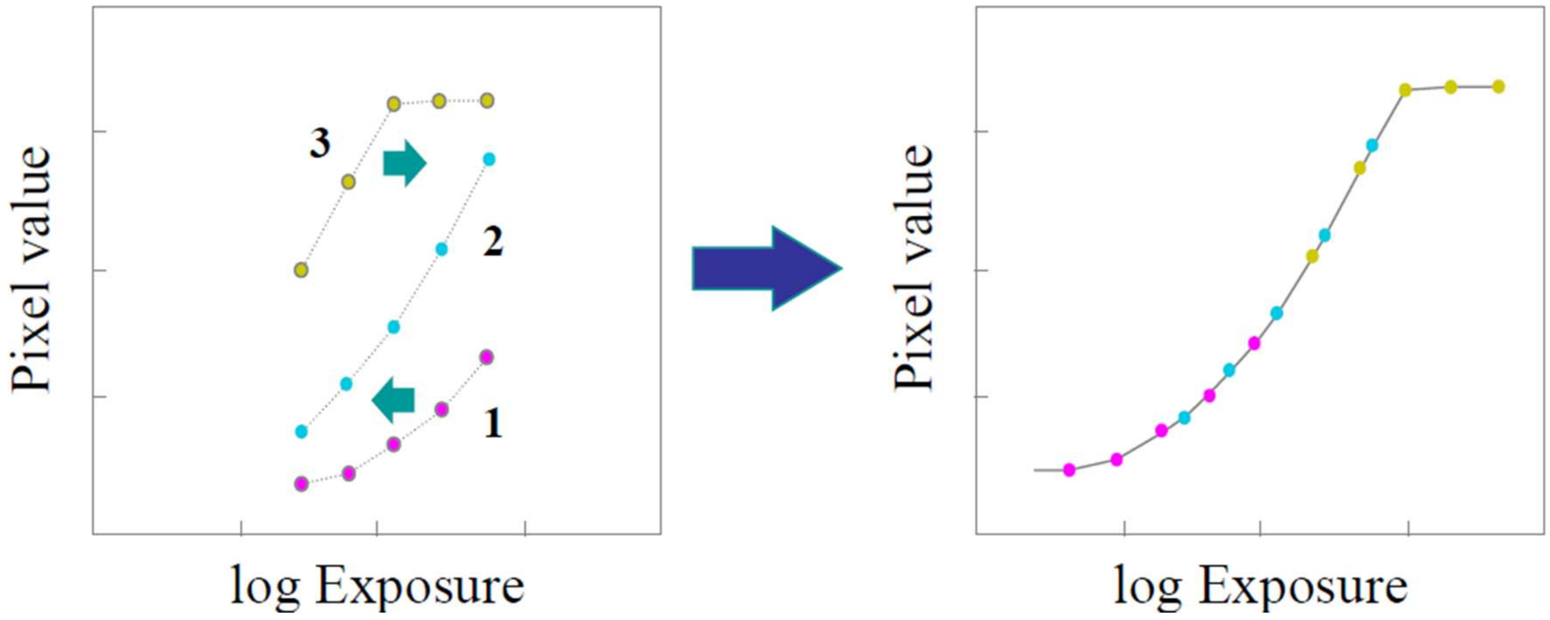
- While it is possible to combine pixels from different exposures directly into a final composite by e.g. taking the mean, this approach risks creating contrast reversals and halos.
- A more advanced approach is this:
  - 1. Estimate the radiometric response function from the aligned images.
  - 2. Estimate a radiance map by selecting or blending pixels from different exposures. Store in 32-bit float.
  - 3. Tone map the resulting high dynamic range (HDR) image back into a displayable range. Use e.g. gamma compression (from Y to Y'). Store in 8-bit unsigned integer.



External HDR, Advanced procedure

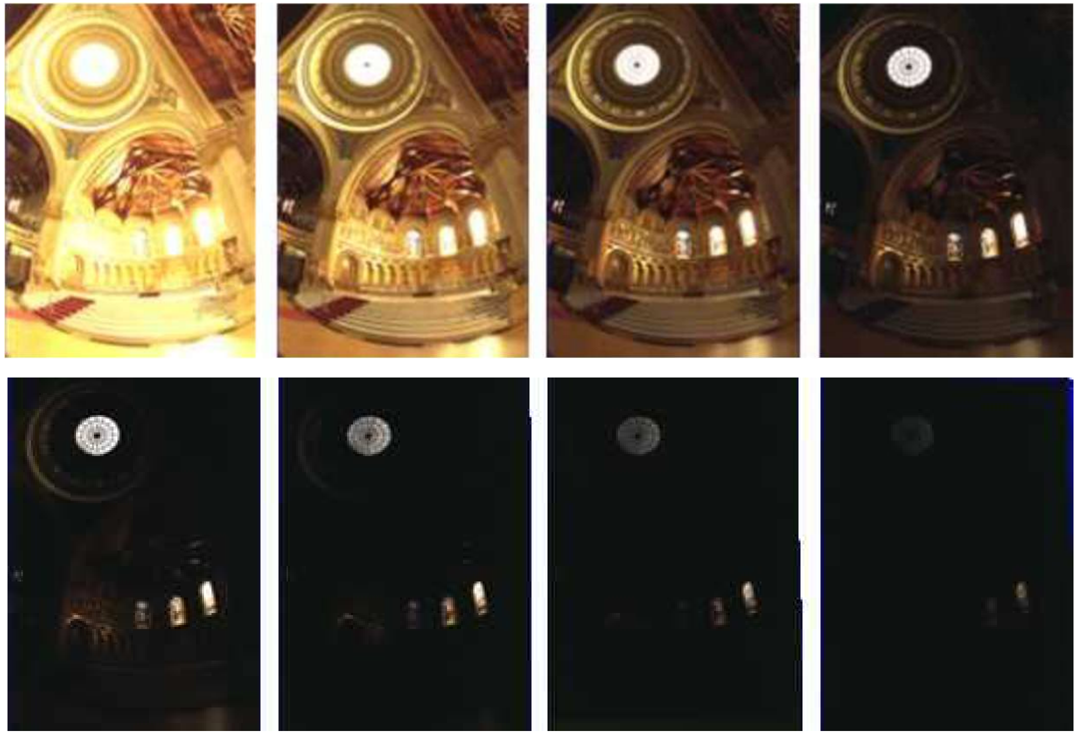
From:
Richard Szeliski: Computer Vision: Algorithms and Applications, 2nd Edition

- Radiometric calibration using 3 exposures.
- Corresponding pixel values are plotted as functions of log exposures (irradiance). The curves on the left are shifted to account for each pixel's unknown radiance until they all line up into a single smooth curve.

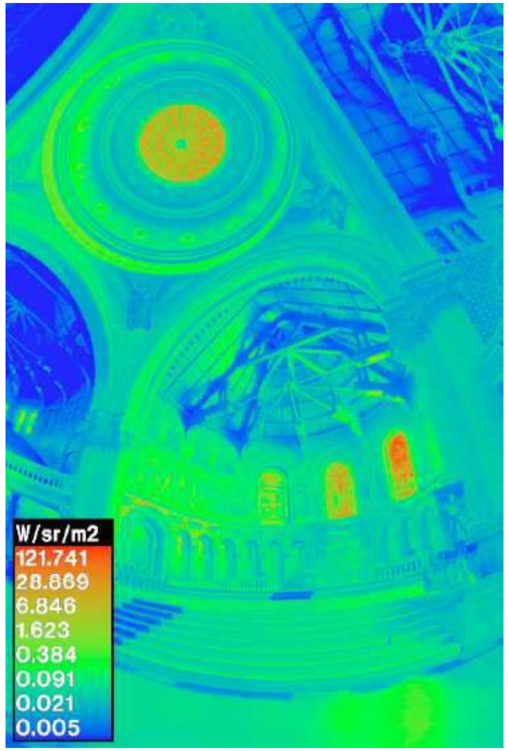


External HDR, Advanced procedure

From:
Richard Szeliski: Computer Vision: Algorithms and Applications, 2nd Edition



8 exposures



Resulting radiance image displayed in pseudocolor (3 are needed: R,G,B)

Internal HDR

- HDR can also be obtained through internal means in the sensor.
- 3 examples:
  - Log-intensity cameras
  - Dual diode
  - Piecewise linear response

Log-intensity cameras

- CMOS technology can be used to achieve a logarithmic dependence between the absorbed light and the resulting photo-voltage
- Enables a camera with very high dynamic range

  (>120 dB = 6 decades)
- Linear and logarithmic mode can be combined in the same camera
- In fact quite simple, since a diode has an exponential current vs voltage characteristic

Logarithmic pixel

- 3T logarithmic pixel
- Transistor M_{rst} acts as a non-linear resistor ("diode")
- V_{pd} will be a logarithmic function of the light intensity (continuous sensing, no integration).

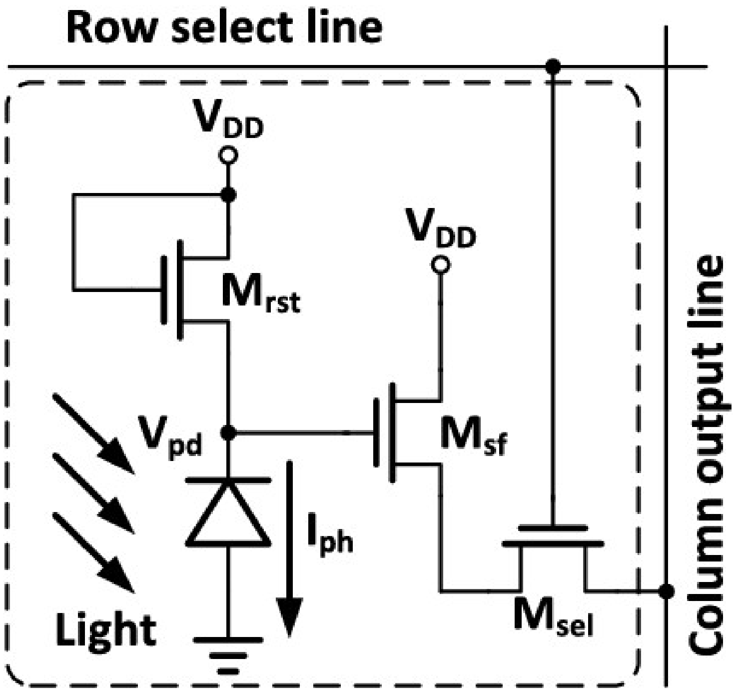


Image: Wei-Fan Chou et al. IEEE Sensors Journal 2014.

Log-intensity cameras

- Example: PhotonFocus AG Lin-log camera

Linear mode

Logaritmic mode

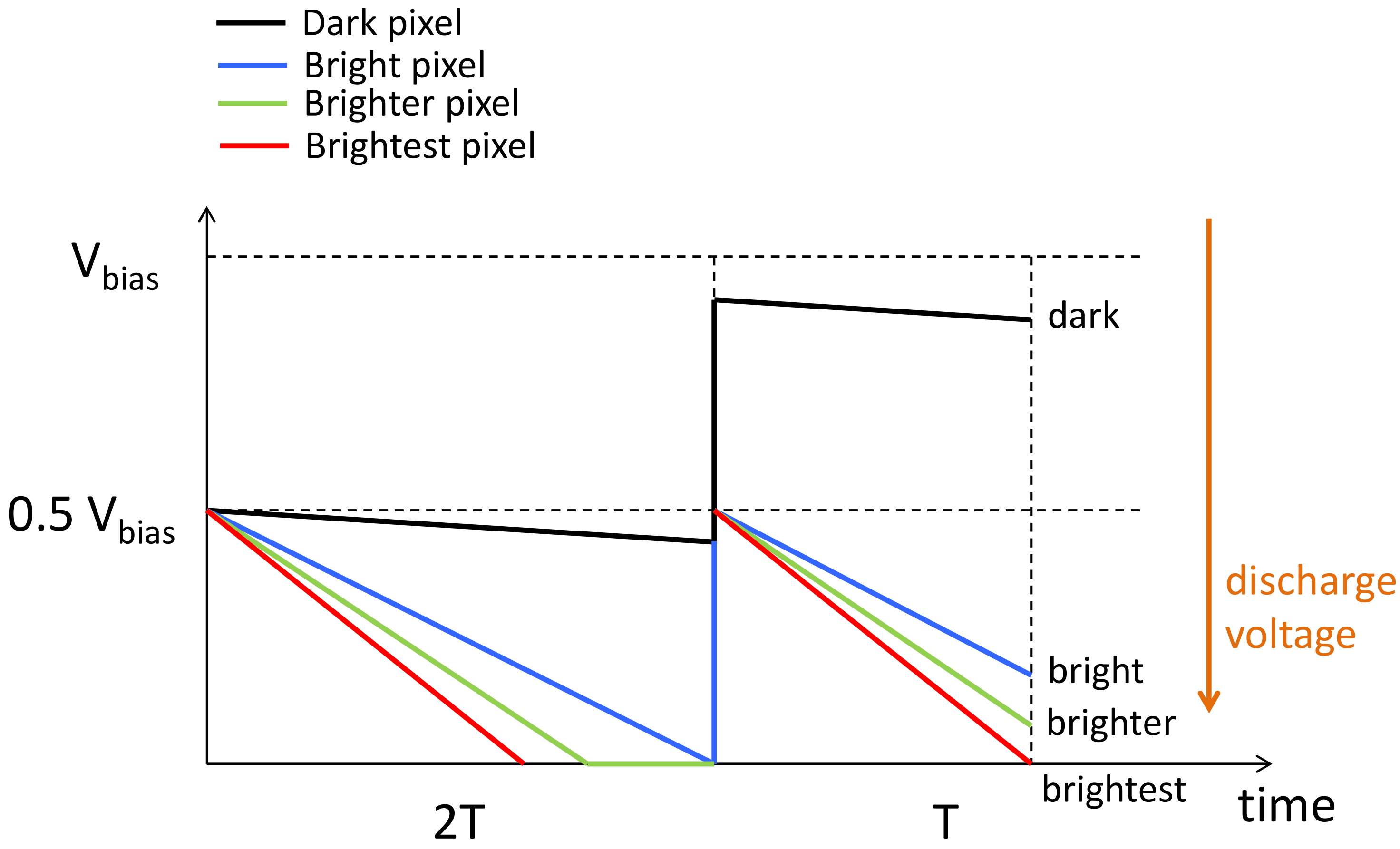
Dual diode

- CMOS
- Each pixel contains two photo diodes with different sensitivities to light
- A high-sensitivity diode
  - Operates well in dark areas
- A low-sensitivity diode
  - Operates well in bright areas

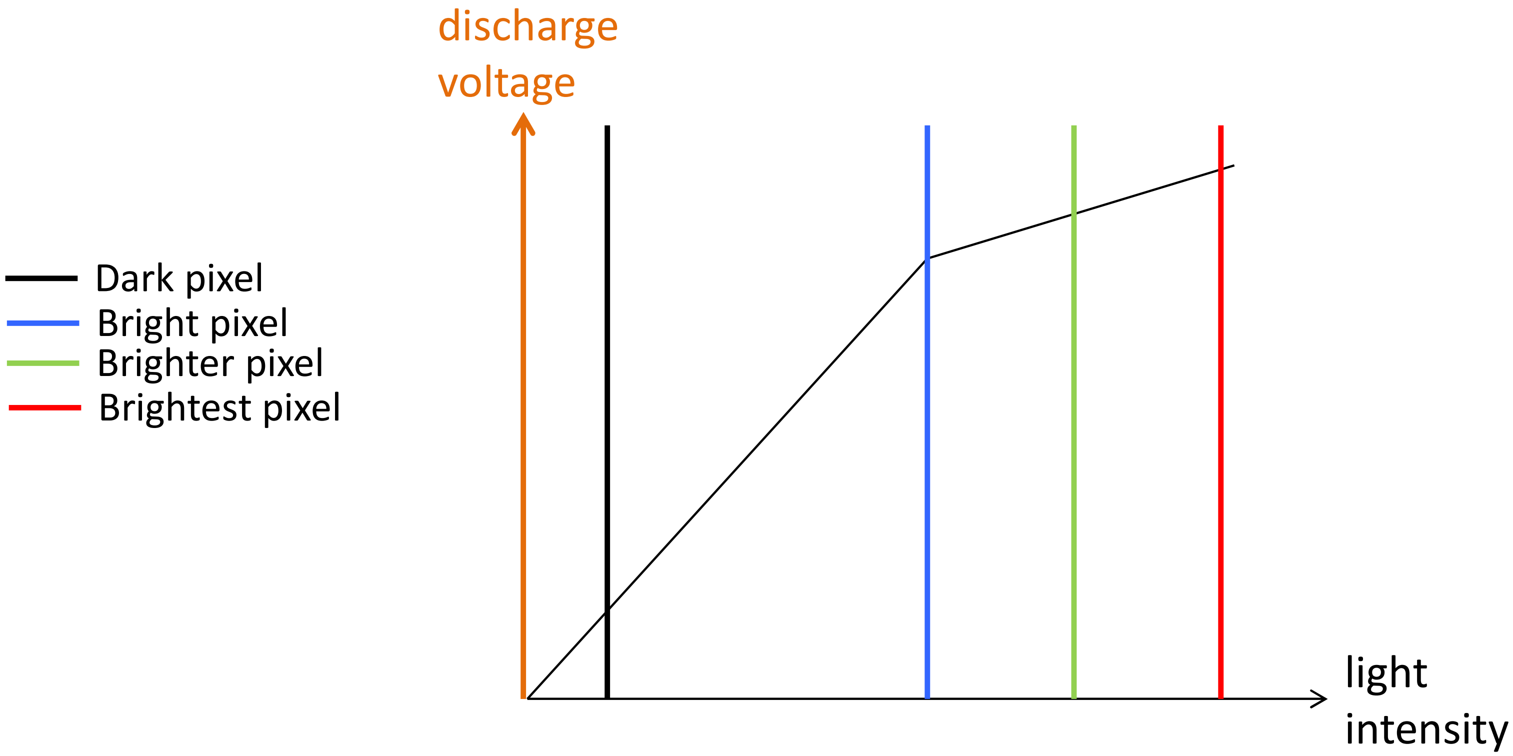
Piecewise linear response

- One example:
  - The initial precharge is only applied to 50% *)
  - Discharge the diode during an initial and longer exposure time
  - Dark areas: The diode never becomes fully discharged
  - Bright areas: The diode becomes fully discharged (saturated)
  - Add the missing 50% precharge (keeping any remaining charge from the first exposure). This requires a modified pixel circuit, though.
  - Continue discharging during a second and shorter exposure time

Piecewise linear response



Piecewise linear response



Seeing in the dark

- However, if it is very dark, long exposure time or logarithmic pixels are not enough.
- Then it is necessary to amplify the image itself.
- Night goggles perform this.

Photo: DSA

Photo-electric detectors

- Basic idea
  - Each electron that has been excited by a photon is made to leave the material and is accelerated using an electric field
  - Further on, the electron enters a photo-multiplier …

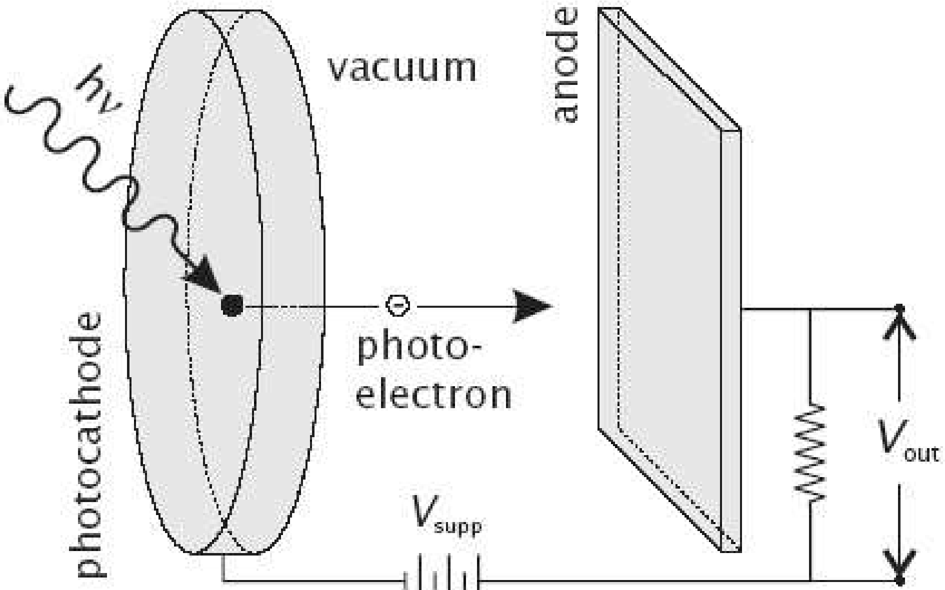
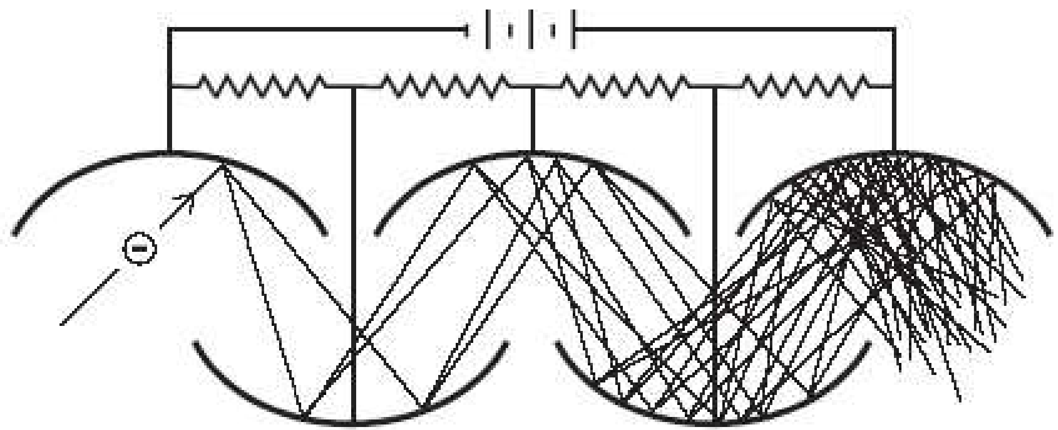


Photo-multiplier

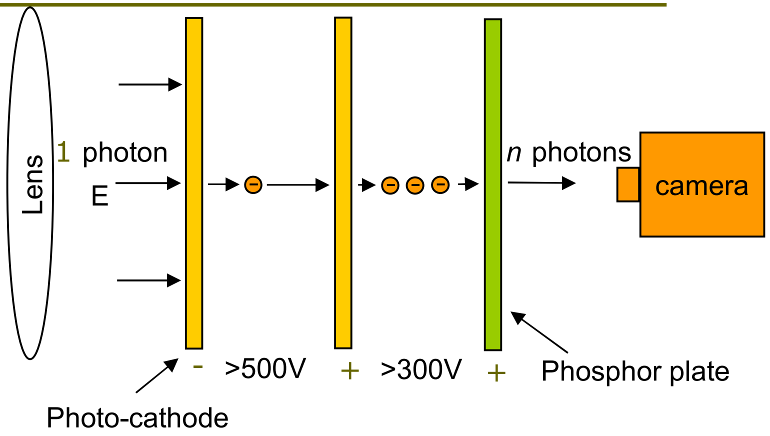
- In a *photo-multiplier*, the field is **strong** enough to make the electron, on impact, knock out two or more electrons
- These, in turn, knock out **several electrons** and in the end, an amplification of > 10^6 can be accomplished

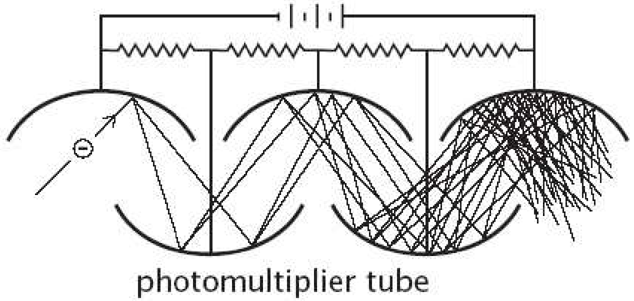


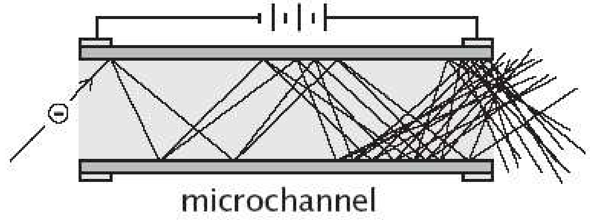
photomultiplier tube

3 different types of photo-multipliers

- **Type 1** consists of a first metal layer that performs the conversion from photons to electrons. An intermediate metal plate accelerates the electrons and multiplies them. The last layer converts electrons back into photons through a phosphorous material. Since there is only one accelerating plate, the amplification is quite modest.
- **Type 2** represents a traditional photo-multiplier tube used to multiply a few photo-generated electrons into a large number of electrons. The number of electrons grows exponentially with the number of accelerating plates. It cannot be used to amplify a complete image.
- **Type 3**, the microchannel plate photo-multiplier, can be used to amplify a complete image.







Micro-channel plate (MCP)

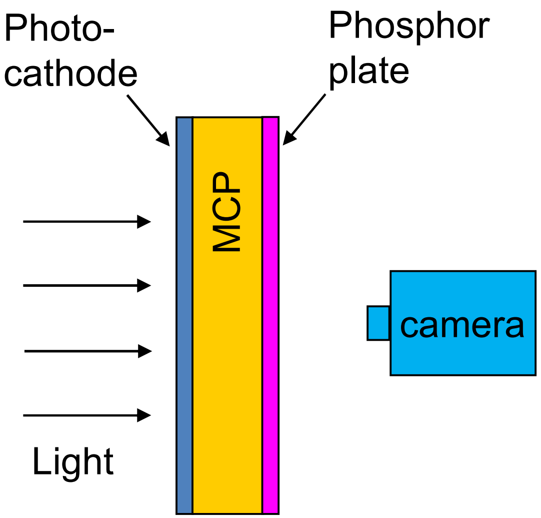
- A micro-channel is a photomultiplier in the form of a tube
  - Can be as small as 10 μm in diameter and a few mm long
  - Electron gain $>10^4$
- A micro-channel plate consists of an array of such micro-channels stacked side by side

Micro-channel plate (MCP)

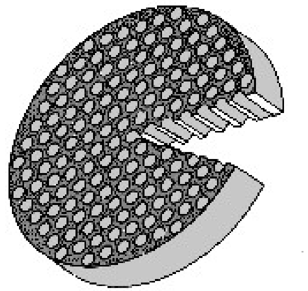
- Applications
  - Photo-multiplication of visual light

    Combined with a phosphor plate, the microchannel plate can be used for very light-sensitive night goggles.
  - X-ray detectors

    It can also be used to convert X-rays to a visible image.



50 mm diameter
>3000 channels across



Tectra GmbH, MCP

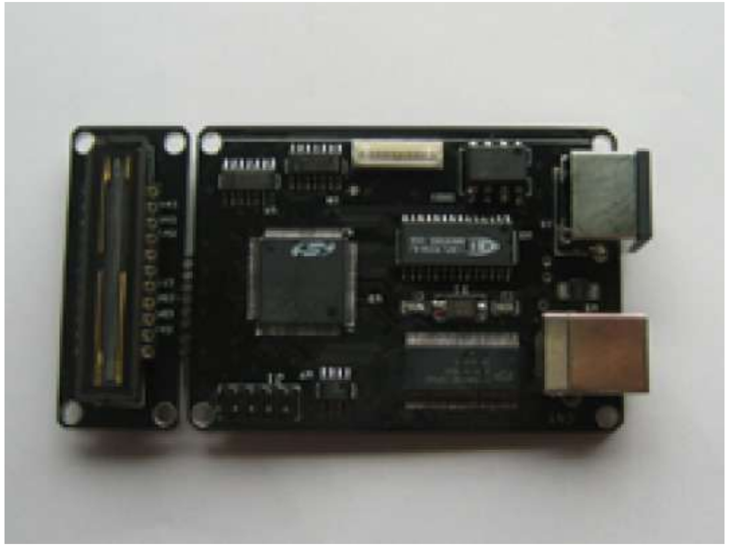
Line camera

- Are used where very high 1-dimensional resolution is required, and where fast read-out is required.
- Can produce 2D images by
  - Translating the scene relative to the camera
  - Translating the camera relative to the scene
  - Rotating the camera relative to the scene
- Also known as
  - Push-broom camera (PBC): It “paints” the 2D image by moving the camera.

    A push-broom camera normally implies that the camera is moving relative to the scene to produce a 2D image

Line camera

- Why a line camera?
  - High resolution along one axis (>10 kpixels)
  - High resolution in bits/pixel (>16 bits)
  - High scanning rate (>10.000 scans/sec)
  - Allows integration on the chip of a processing unit per pixel (smart camera)

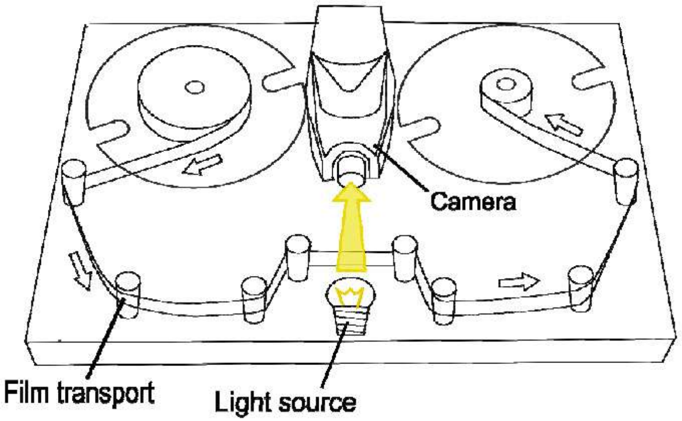


Mightex 3648-pixel line camera

Line camera, applications

- Applications:
  - High-resolution inspection
  - Photo finish
  - Fax, copy machines
  - Film scanners
  - ...

This is **not** an image of a 3D scene at some point in time. It is an image of a **line** at different points in time.



ImageSystems GoldenEye film scanner

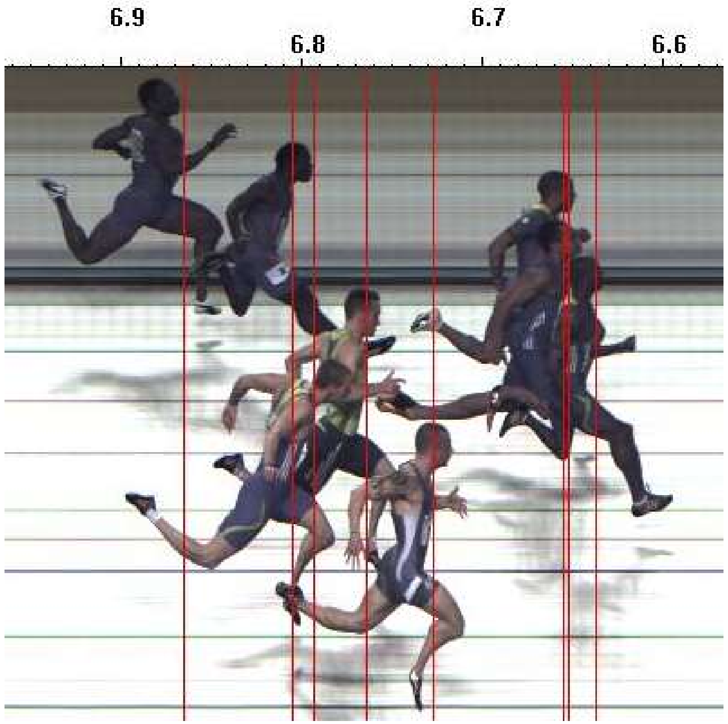
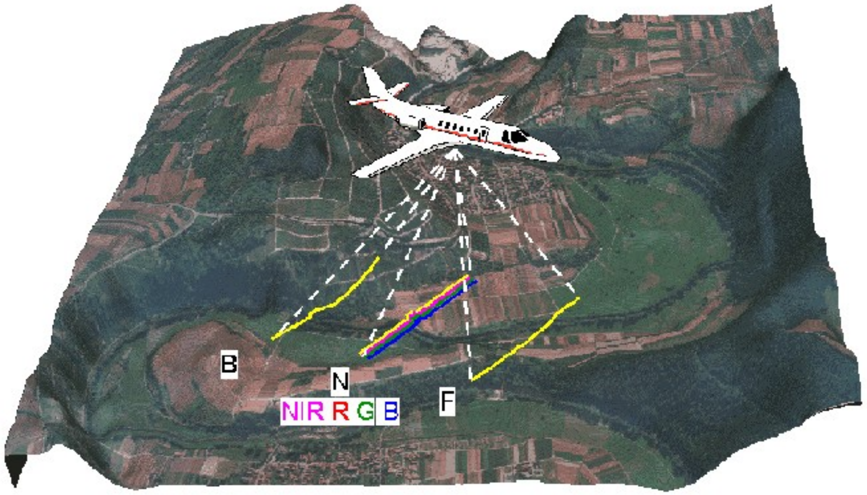
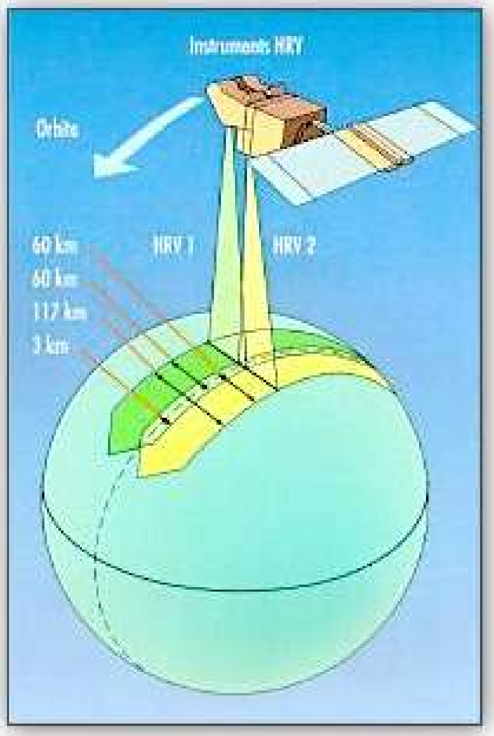


Photo finish

Line camera, push-broom variant, applications



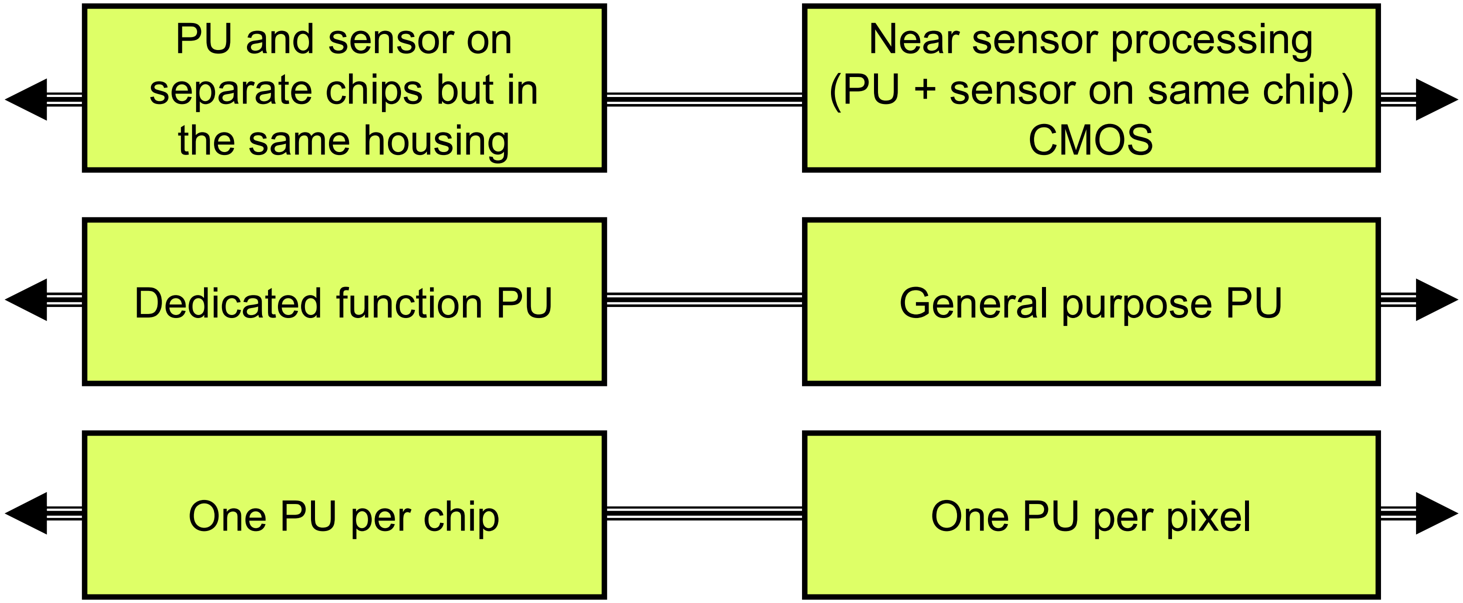
Ground mapping from air



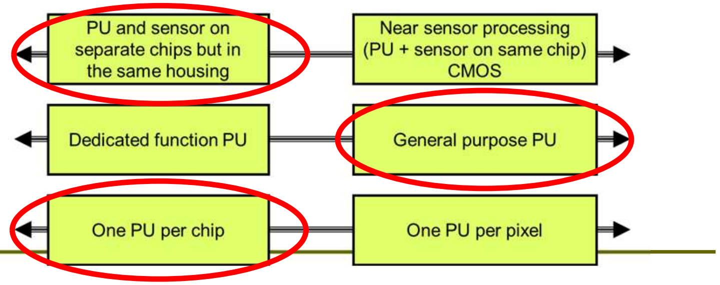
Ground mapping from satellite

Smart cameras
for machine vision applications

- Integrates processing unit(s) (PU) with a camera chip. Several different solutions exists:



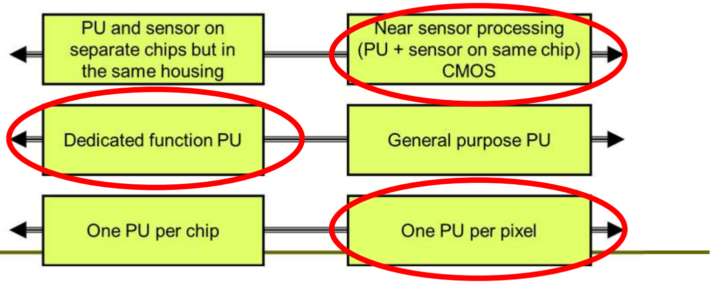
Smart cameras



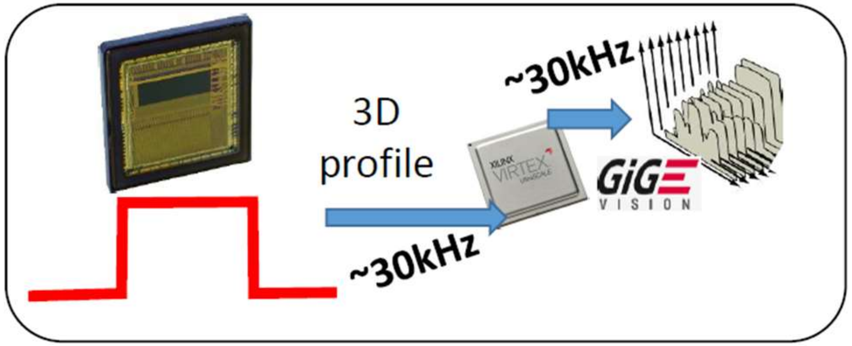
- Some cameras come with a completely integrated development environment (IDE)
  - Processing is defined in a GUI on a standard PC
  - Processing code is downloaded into the smart camera and executed
- Some cameras are integrated with illumination, e.g.
  - IR
  - visual
- Most cameras have a simple interface to a network, e.g.
  - TCP/IP protocol over Ethernet
  - Integrated web server

Cognex Checker 252

Smart Camera: SICK Ruler3000



- The smart cameras from the company SICK-IVP are quite different.
  - The sensor and image processor are integrated into the same chip.
  - The processor is programmable but not general purpose.
  - There is a dedicated processing unit per pixel.
  - These cameras are aimed at very fast industrial and robotic applic:s.
  - SICK claims that they produce the fastest and most compact laser triangulation systems in the world.



The camera contains a custom CMOS image sensor plus a “medium” FPGA

Smart cameras

- Used where a low-cost solution is required, or alternatively, extremely high performance.

- Applications
  - Range imaging
  - Bar code reading / data matrix / OCR
  - Event/motion detection
  - Counting objects/people passing by
  - Surveillance
  - Gaze measurement
  - …

Event cameras

- Previously, in the course, we had a guest lecture on Event cameras by Hannes Övrén from the Swedish Defence Research Agency (FOI).
  - Hannes Ovrén, Scientist (PhD)
    Cyber Defence and C2 Technology
    Sensor Data Analysis
    hannes.ovren@foi.se
- There is a video at FOIs website, where some of Hannes' movies are included: https://foi.se/nyheter-och-press/nyheter/2024-09-03-eventkameror---nasta-generations-spaningssensorer.html
- Exjobb
  - Exjobbskatalogen 2025, Online list
  - Applied Research
  - Defence / Public Security
  - Swedish citizenship needed

foi.se/exjobb

Event cameras

- **Short definition**: An event camera only senses motion in the scene.

- **Long definition**: An event camera is an imaging sensor that responds to local changes in brightness. Event cameras do not capture images using a shutter as conventional cameras do. Instead, each pixel inside an event camera operates independently and asynchronously, reporting changes in brightness as they occur, and staying silent otherwise. (Wikipedia)

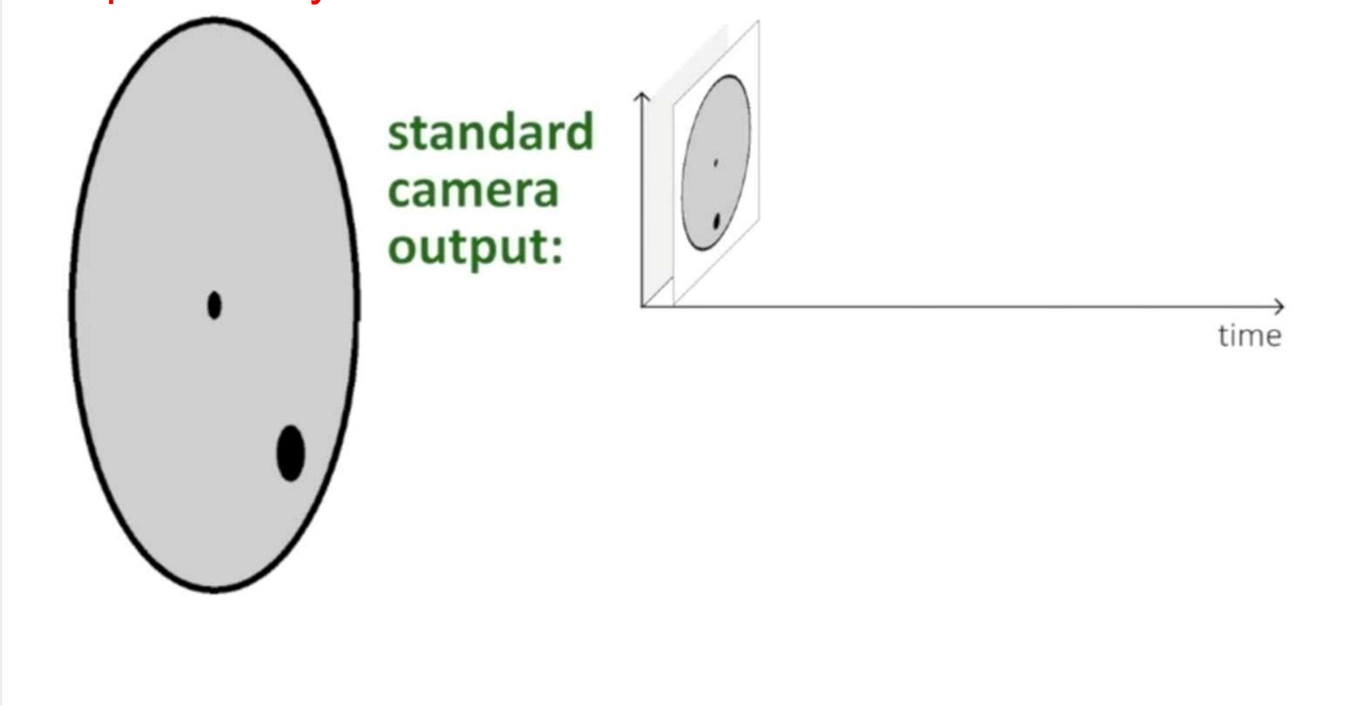
Event cameras

- Event Camera Properties
  - High dynamic range
  - Asynchronous
  - (Almost) no motion blur
  - Hightemporal resolution
  - Low power
  - Low bandwidth
- Lots of applications (hot research topic)

From Hannes' Guest lecture

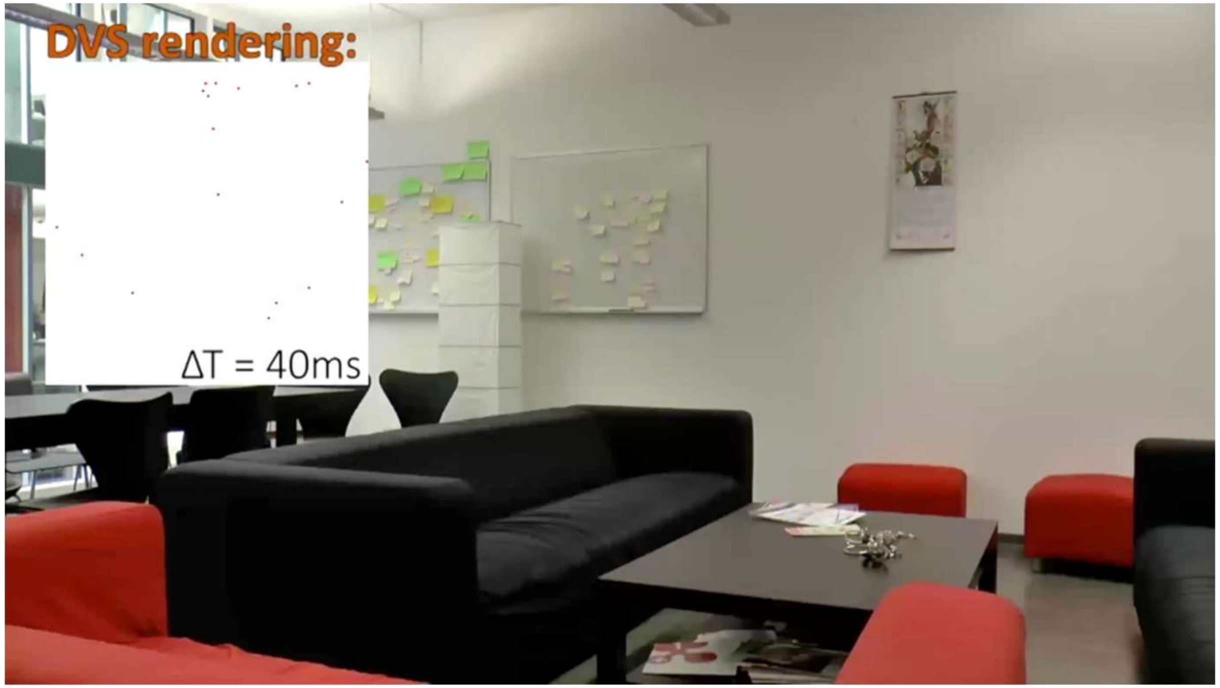
Event cameras

- Illustration of the difference between an event camera (Samsung DVS) and a standard camera.
  - For the full video, see: https://www.youtube.com/watch?v=LauQ6LWTkxM



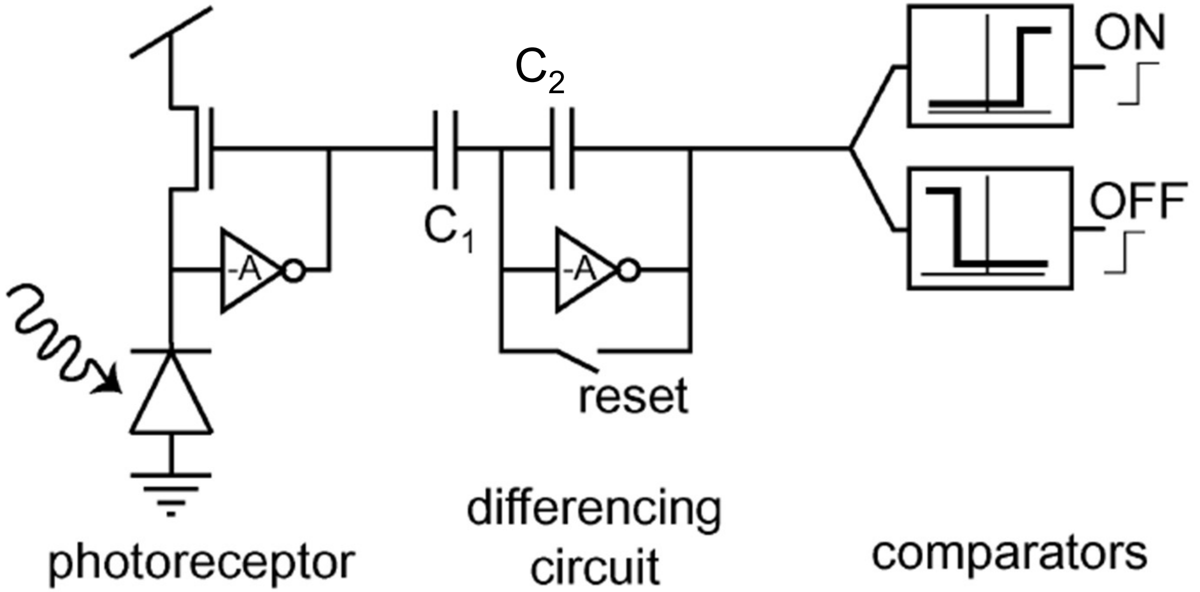
Event cameras

- Motion in a natural scene.
- The white screen shows the output from the event camera.
  - Full video: https://www.youtube.com/watch?v=LauQ6LWTkxM



Event camera

The Dynamic Vision Sensor (DVS) Pixel



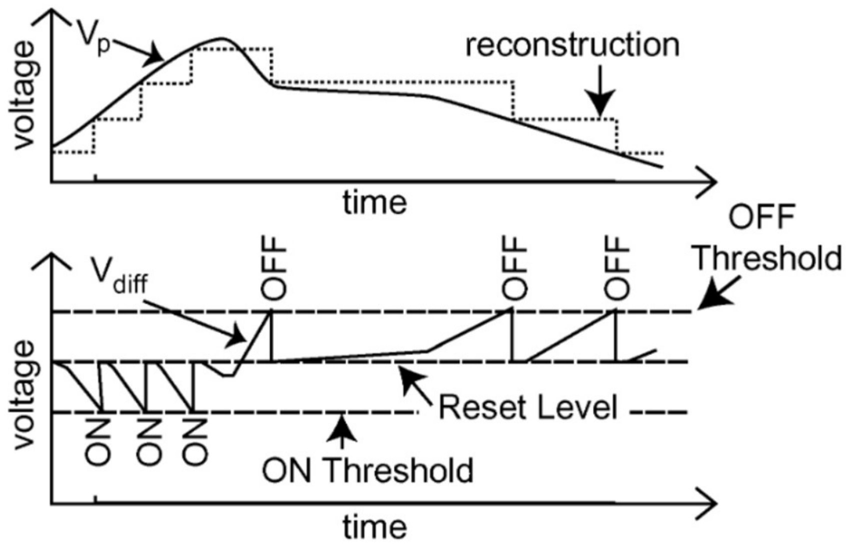
From Robert Forchheimer, Hannes' Guest lecture, and Lichtsteiner 2008

1) Photoreceptor measures continuously (no integration).
2) Reset ON: C_1 stores current light intensity as a charge.
   Any charge in C_2 is empied.
3) Reset OFF: if the charge in C_1 changes, a corresponding change will happen in C_2 and an ON or OFF event is generated (with some threshold)
4) The event is read-out and a new reset is done (typically within μs)

Event camera

The Dynamic Vision Sensor (DVS) Pixel

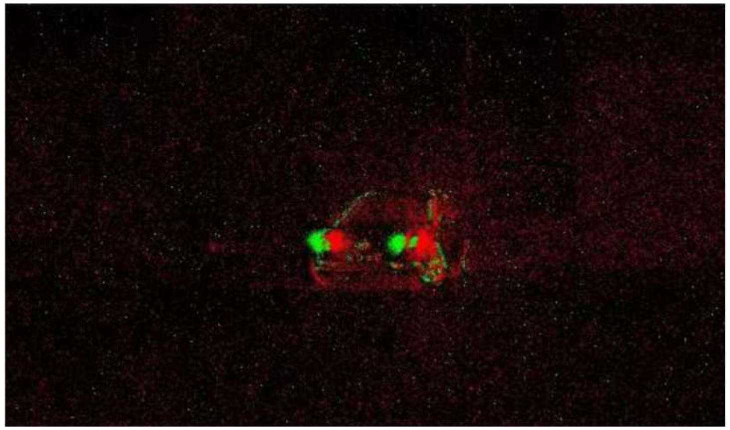
- Bottom right: The output (ON and OFF pulses) from an event camera.
- Top right: The integration (with some ”adjustments and adaptations”) gives a common image (almost).



From Hannes' Guest lecture, and Lichtsteiner 2008

Event camera, example Car with headlights

- The car moves.
- Top right: Image from a normal camera.
- Bottom right: Image from an event camera.

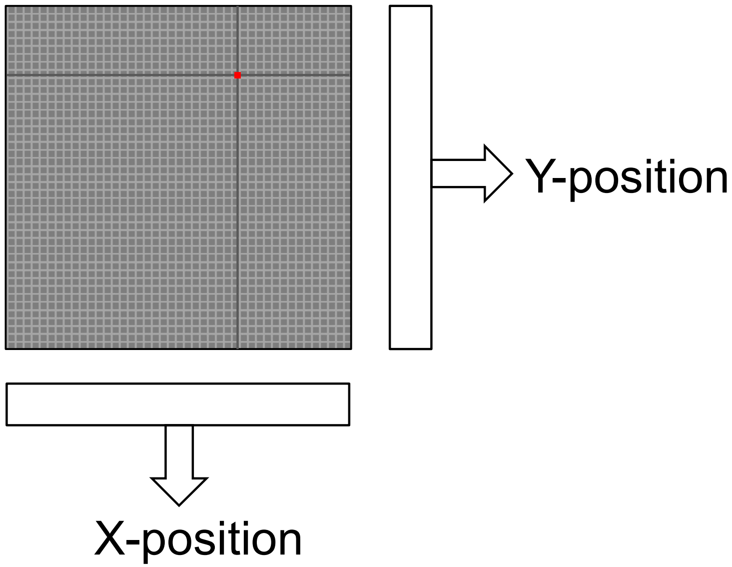


From Hannes' Guest lecture,

Event cameras

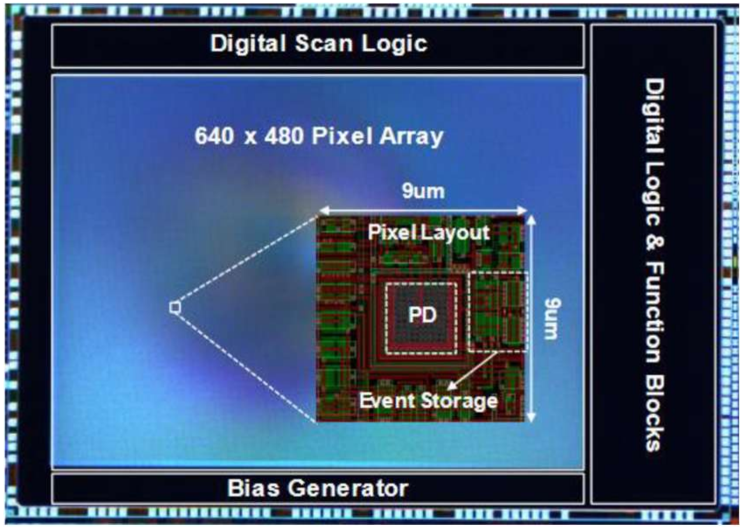
Event address representation (AER)

- An event (ON or OFF) triggers a readout of the pixel address.
- For every event, its X and Y position is delivered from the sensor.
- A modern event camera can deliver millions of events per second.



Event camera
Can I Buy One?

- Three main suppliers
  - iniVation (Switzerland), cooperates with Samsung
  - Prophesee (France), cooperation with Sony
  - CelePixel (China)
- Sensor sizes
  - Today: 640 x 480
  - Soon: 1280 x 720 and
    1920 x 1280
- Cost: From 3,000 €



From Hannes'
Guest lecture,

Computational cameras

- Cameras that capture optically coded images and require computations to produce “real images”

- Examples
  - Light-field cameras
  - Omni-directional cameras
    - Multiple pinhole cameras
    - Fish-eye lens
    - Catadioptric imaging
  - Coded aperture
  - holographic imaging (not included)
  - flexible depth-of-field
  - …

Light field

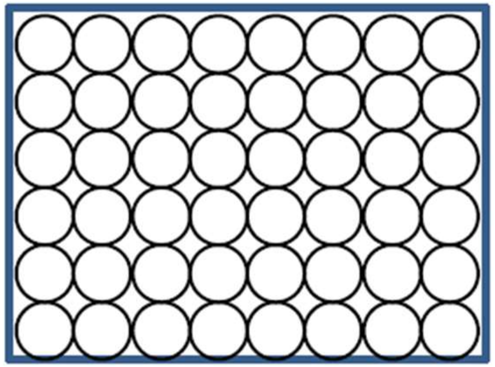
- Repetition from lecture 1:
  - At each point **x** in 3D space,
    in each direction **n**,
    there is an amount of light passing through **x**
  - The plenoptic function I(**x**, **n**)
    also known as Light field

Light field camera

- A light field camera makes a denser sampling of I(**x**, **n**) than a standard camera
  - Ideally all **x** and all **n** (not practically possible)
- Practical implementation: an array of pin-hole cameras, approximated by lens cameras

Light field camera

- A more practical implementation:
  - Use a large sensor chip: lots of pixels
  - Divide the chip into several small "cameras"
  - Use a sophisticated lens system,"lenslet array", to emulate an array of pin-hole camera projections onto the chip.



lenslet array

Light field camera, examples

Lytro

Stanford plenoptic camera

Raytrix

Light field camera, applications

- A point in the scene is viewed from multiple directions:
  - 3D reconstruction possible from “single image”
  - Extended depth of field
  - Adjustable object plane (and thus focus) after exposure

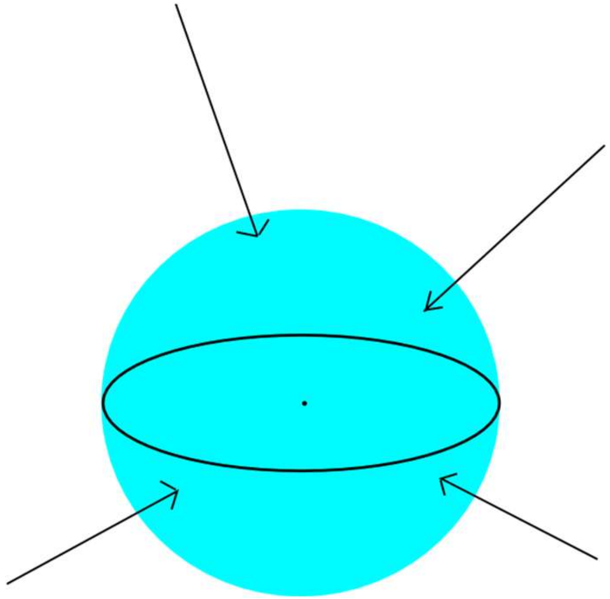
Images from Lytro

Omni-directional cameras

- Omni-directional cameras represent another class of light-field cameras. Here, it is the ability to capture many different directions that is of importance.
- Omni: Latin for every, all
  - In theory: a camera that sees in all directions
  - Often in practice: not all directions, but a much larger field of view compared to a standard camera
- There are several design approaches to omni-directional cameras
  - Multiple pinhole cameras
  - Fish-eye lens
  - Catadioptric camera

Omni-directional cameras

- The image is best represented as a sphere instead of a plane
  - Can be mapped to a plane image but with severe distortion
  - Ideally: all light rays intersect at a single point
  - Cannot be achieved in practice, but we will look into some approximations.



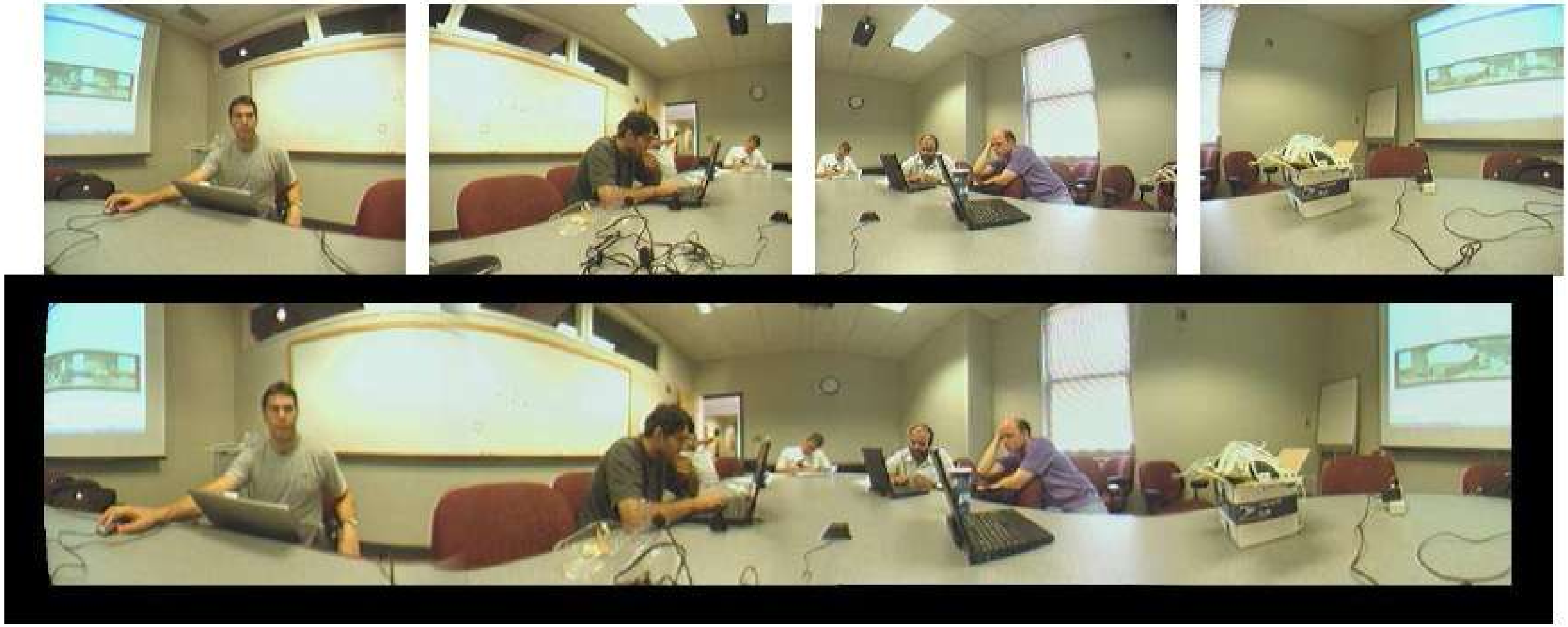
Multiple pinhole cameras

- Set up multiple pinhole cameras to cover the desired set of directions.
- Use image stitching to produce a representation of the image sphere (or parts thereof).
- We did it partially in the panorama lab!

Ladybug 3 from
Point Gray Research Inc.

Multiple pinhole cameras

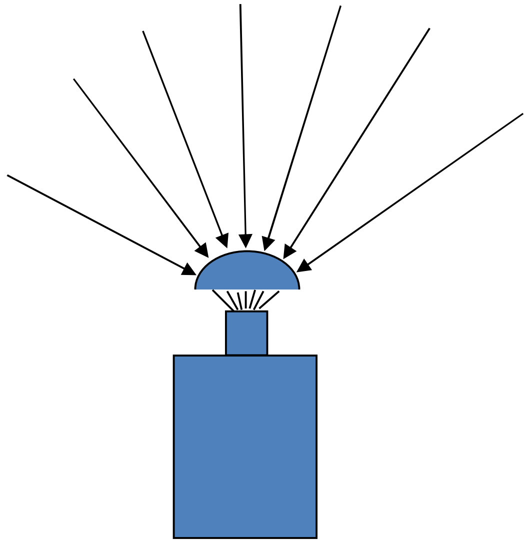
- The multiple pinhole camera can produce a wide field-of-view by stitching several images together.



From: de la Torres, et al,
Learning to Track Multiple People in Omnidirectional Video

Fish-eye lens

- A single camera with a fish-eye lens can cover approximately a hemi-sphere.

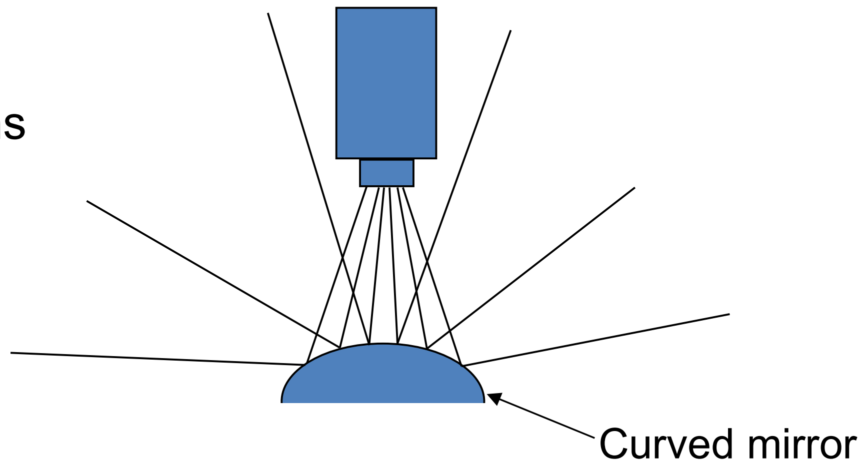


Fish-eye lens

Image: Dan Slater

Catadioptric cameras

- Mirrors and/or lenses re-project the light rays into a single camera lens
  - Special case: fish-eye lens
- The mirror can be
  - Spherical
  - Hyperbolic
  - Conic
  - …



Catadioptic cameras are described in the video by **Shree Nayar**, Time 2:00-11:30 Link on the course webpage.

Catadioptric cameras

Kaidan Inc. catadioptric objective

Image: Tomas Pajdla

Catadioptric vs. fish-eye lens

- Fish-eye lens: cheap and simple
  - For example, a door peep-hole: 50 SEK
  - Can give approx 180^{o} field of view
- Catadioptric camera system
  - Exact control of how the plenoptic function is sampled by choosing the curvature of the mirror
  - User specified curved mirrors: expensive
  - Single optic center can be accomplished
  - Can also capture a larger field of view, nearly 360 degrees, with camera occlusion

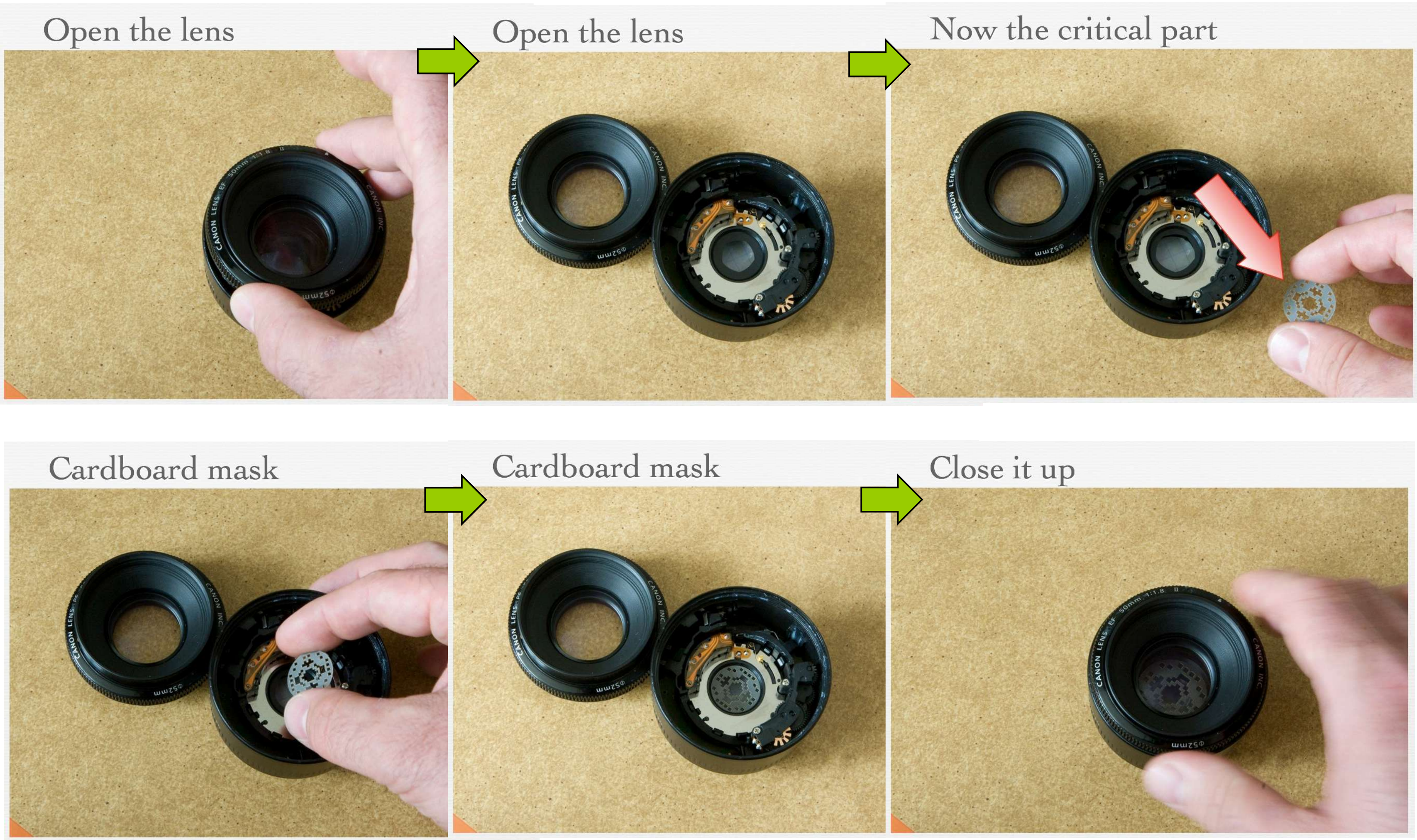
Omni-directional cameras

- Applications
  - Video conferences
  - Surveillance
  - Environment mapping
  - …

Coded aperture

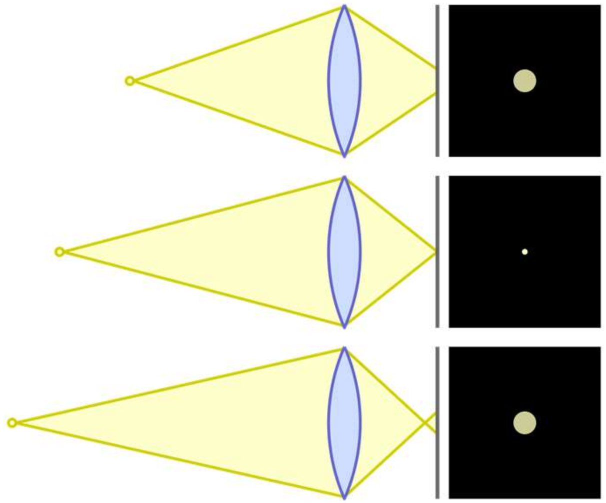
- **Aimed at extending the depth of field, with the possibility to obtain an All-focus image!**
- A computational algorithm is needed to reconstruct the image
- A simple coded aperture: the pinhole camera with a circular aperture
- More complex apertures can be used to improve the reconstructed image
- The next two slides are from the previous lecture: Image Formation, Lenses

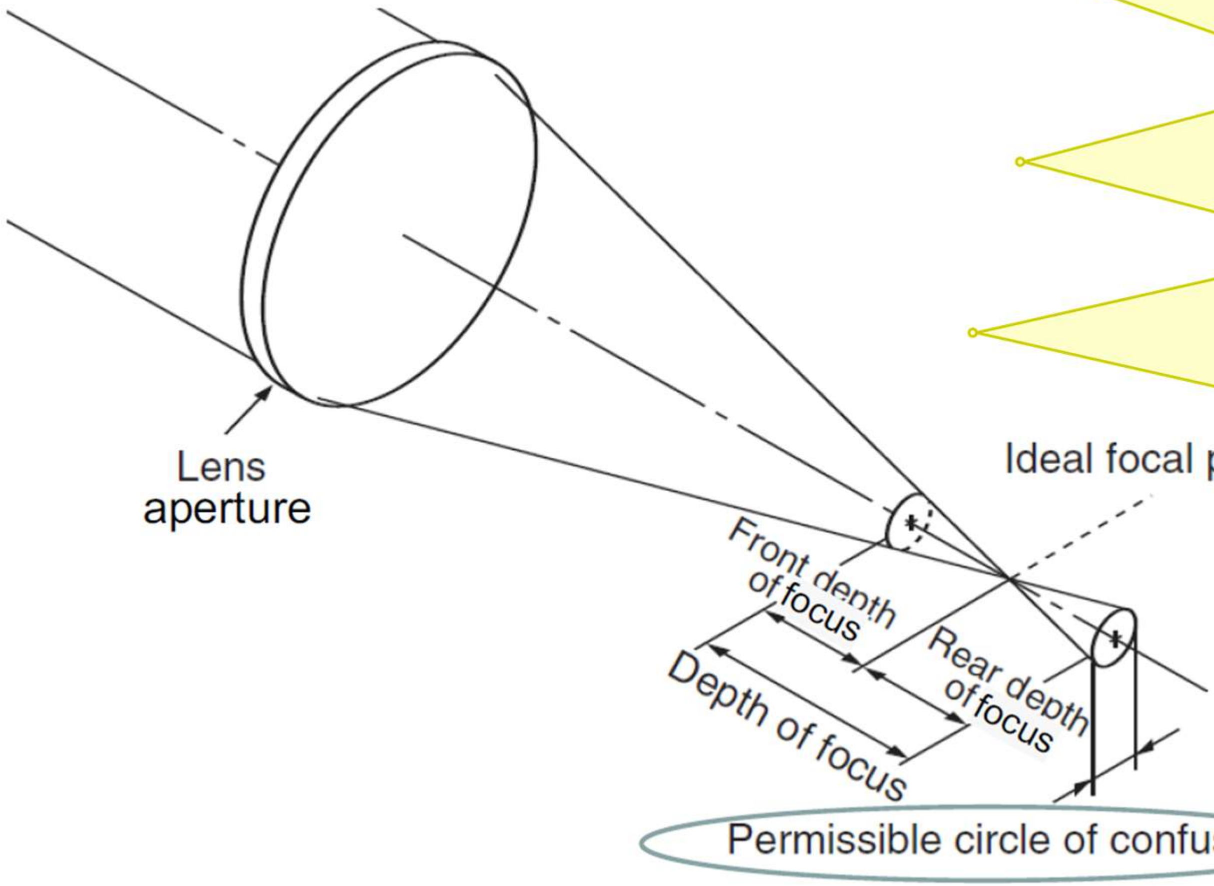
How to modify an ordinary camera to use a coded aperture



Depth of field, Depth of focus, Circle of confusion

From the previous lecture:
Image Formation, Lenses

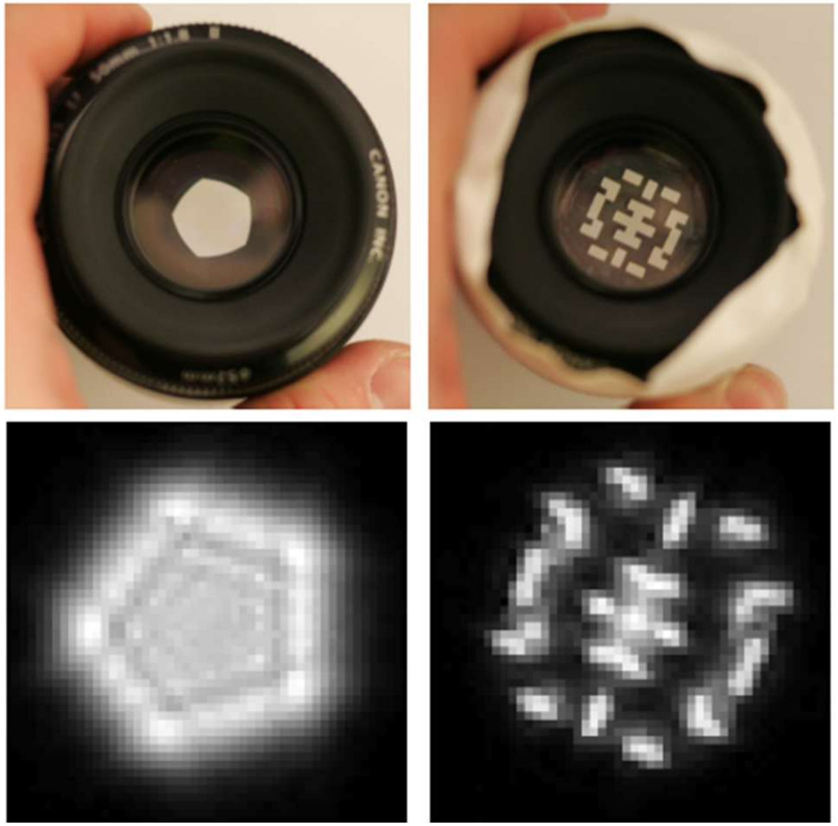




Examples of defocused images of a point source from other apertures than a circle

From the previous lecture: Image Formation, Lenses

- Top left: a standard Canon 50mm f/1.8 lens with the aperture partially closed.
- Bottom left: the resulting blur pattern. The intersecting aperture blades give the pentagonal shape, while the small ripples are due to diffraction.
- Top right: the same model of lens but with a filter inserted into the aperture.
- Bottom right: the resulting blur pattern.
- From Levin et al: Image and Depth from a Conventional Camera with a Coded Aperture



More explanation to the blur pattern

- Consider a point light source that is out of focus. The resulting point spread function will be strongly influenced by the shape of the aperture.
- If the image plane is in focus, we know that the point spread function is the Fourier transform of the aperture - followed by absolute value and squared.
- On the other hand, if the image plane is put close to the lens, then the lens would have little effect and the image would simply look like a silhouette of the aperture.
- Obviously, any location between these two extremes will produce something in between ”Fourier transform of the aperture - followed by absolute value and squared” and ”silhouette of the aperture”. This will contain both spatial properties such as the basic shape of the aperture as well as frequency properties such as diffraction patterns. A similar mix will appear also when the image plane is moved beyond the focal point.

Solution 1: Lens with pentagonal (almost circular) shape aperture

Aperture

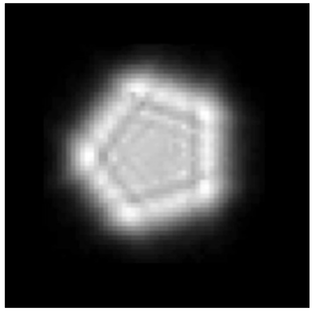
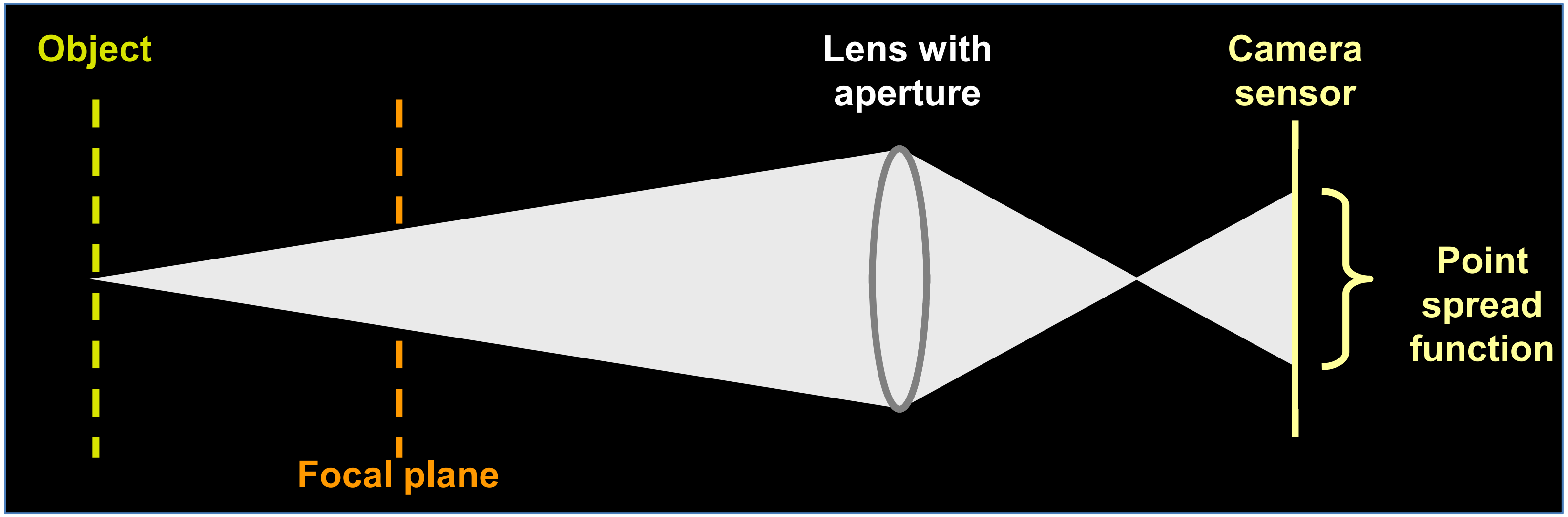


Image of a defocused point light source



Levin et.al. SIGGRAPH07

Solution 2: Lens with coded aperture

Coded aperture

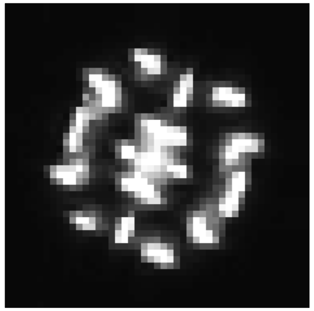
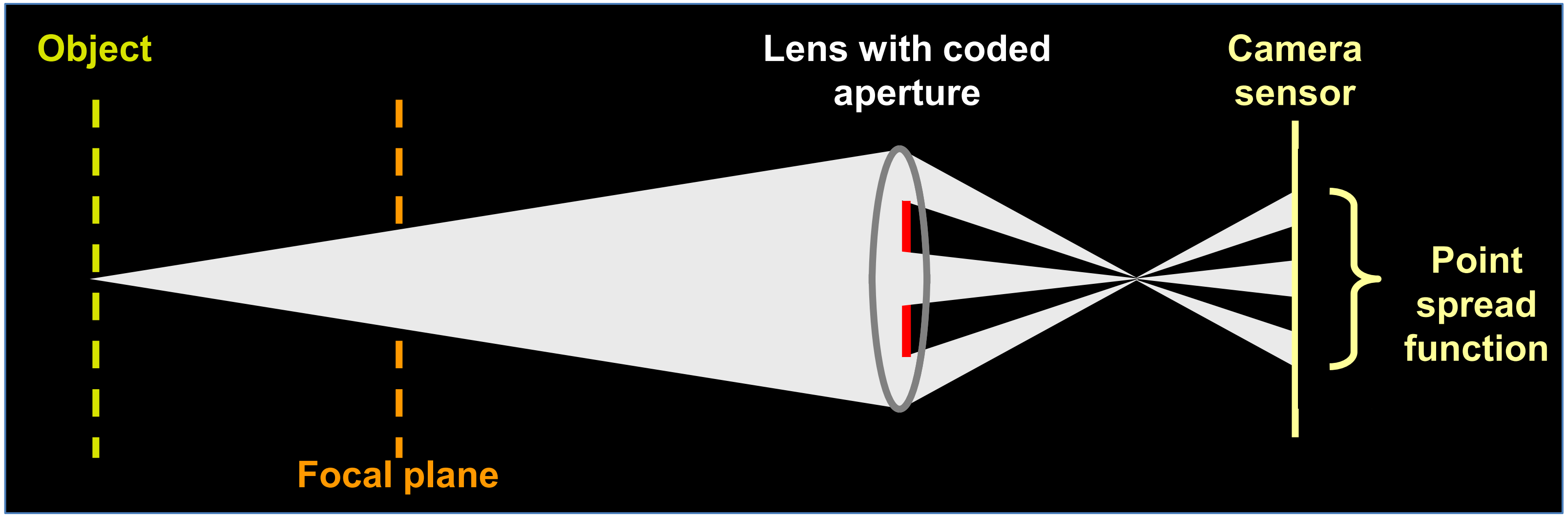


Image of a defocused point light source



Levin et.al. SIGGRAPH07

Solution 2: Lens with coded aperture. Move the object closer to the focal plane.

Coded aperture

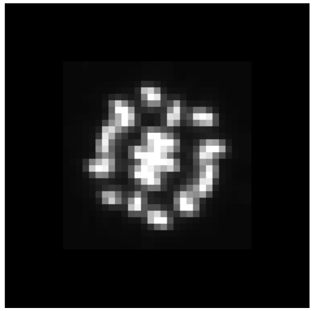
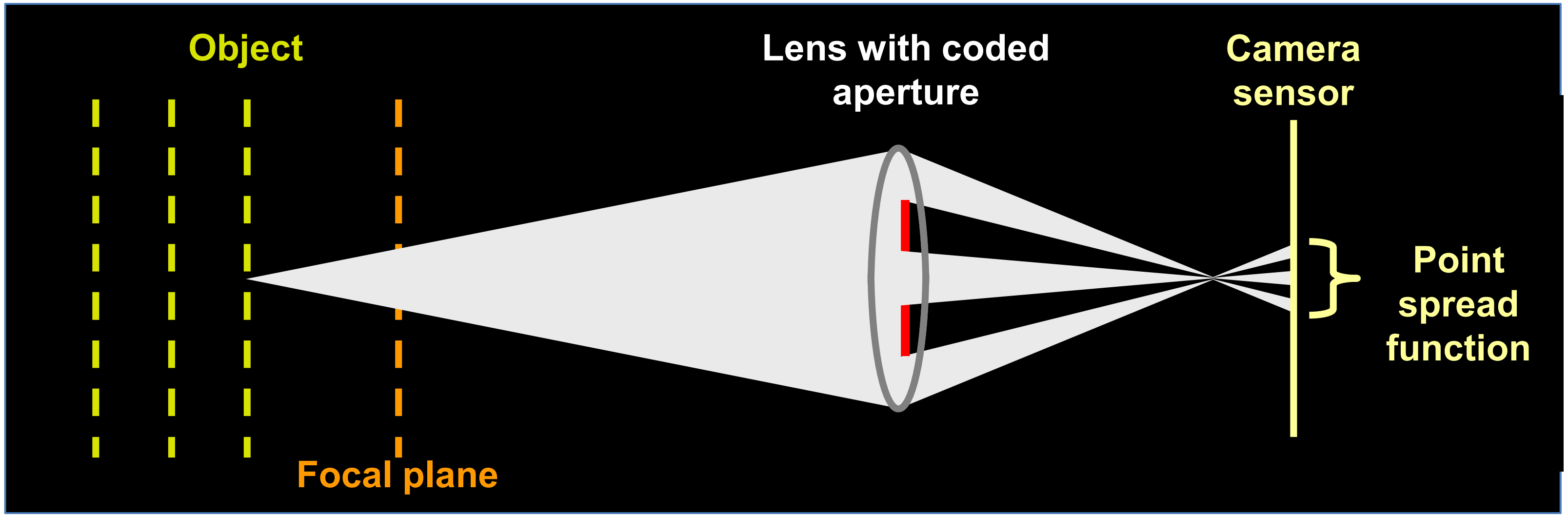


Image of a defocused point light source



Levin et.al. SIGGRAPH07

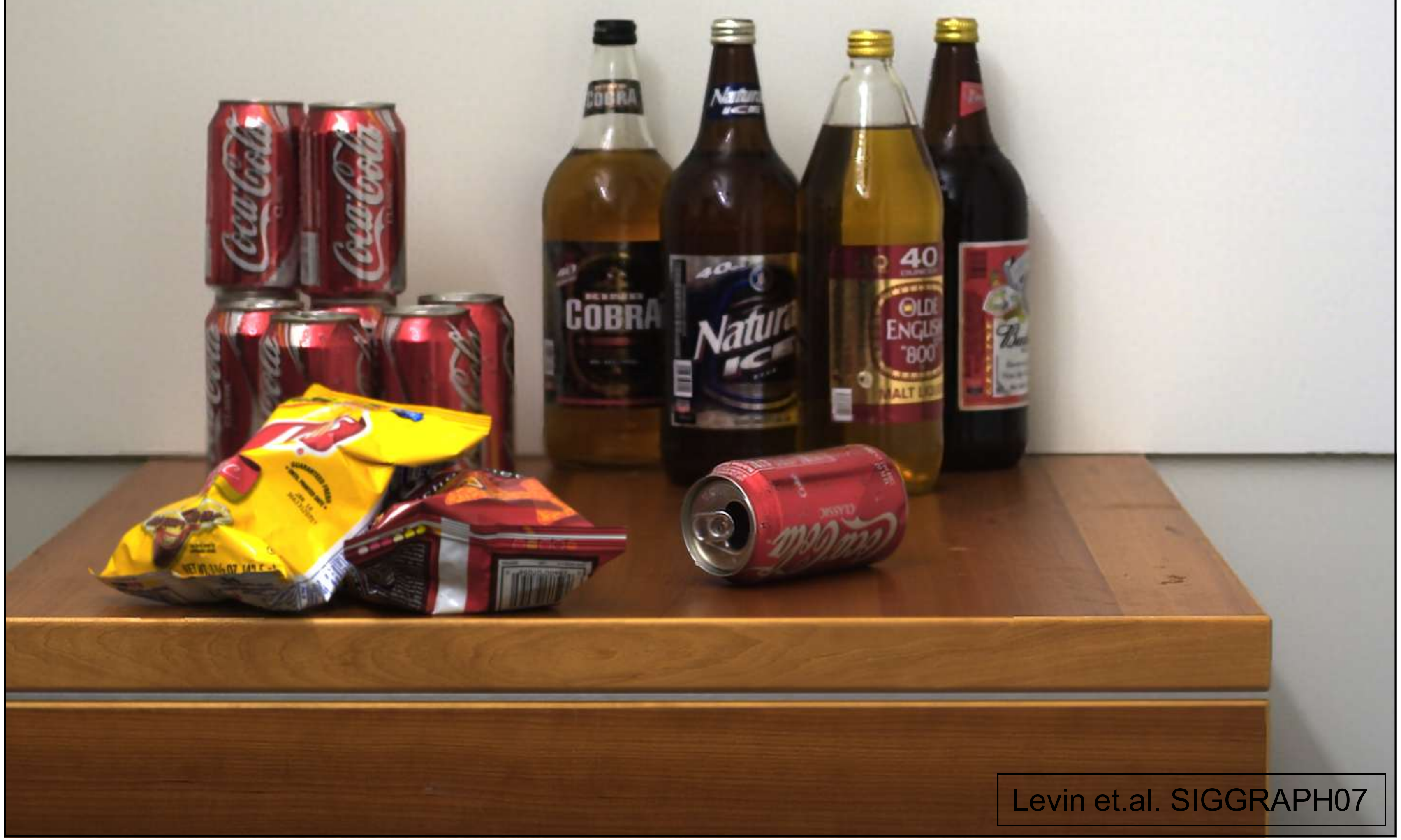
Reconstruction

- The image is formed by convolution of "the point spread function" with "a sharp image of the object": $y = f_k * x$
- f_k is the blurred point spread function, where k indicates that the point spread function depends on the depth
- Deblur locally by minimizing $|f_k * x - y|^2$
- Search for the best depths (k) in the local areas
- Reconstruct all parts of the image

Example: The captured picture before applying the deconvolution

Levin et.al. SIGGRAPH07

Example: The captured picture after applying the deconvolution



Levin et.al. SIGGRAPH07

Close-up

Original image

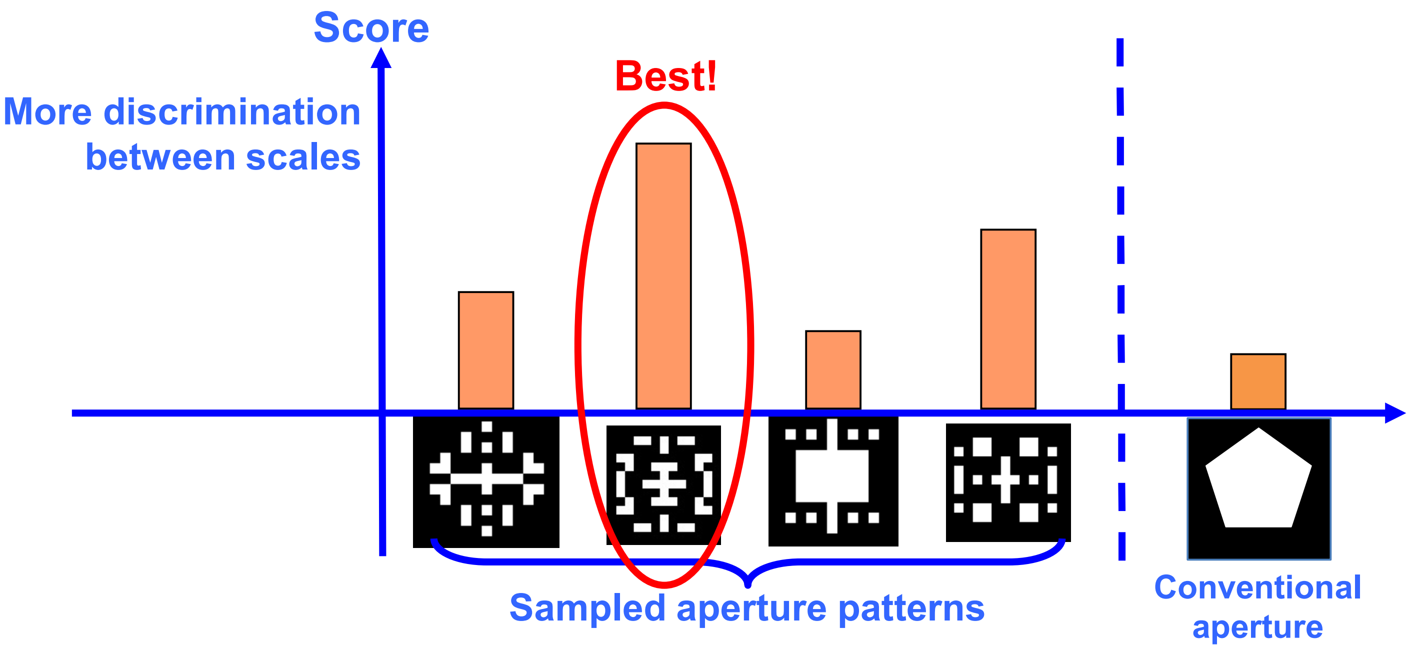
All-focus image

Levin et.al. SIGGRAPH07

Reconstruction quality depends on the particular aperture

- If the spectrum of the point spread function, also called the optical transfer function, is close to zero at certain frequencies, it is not possible to find an exact inverse.
- A circular aperture is essentially a low-pass filter and will not let high frequencies go through. Such frequencies can thus not be restored afterward.
- A much better aperture function is shown on the next slide. It will generate a point-spread function that includes also high frequencies, which makes it possible to design a good inverse filter for deconvolution.

Reconstruction quality depends on the particular aperture



Example: Note that the person to the left is out of focus

Example: After restoration

Levin et.al. SIGGRAPH07

Close-up

Original image

All-focus image

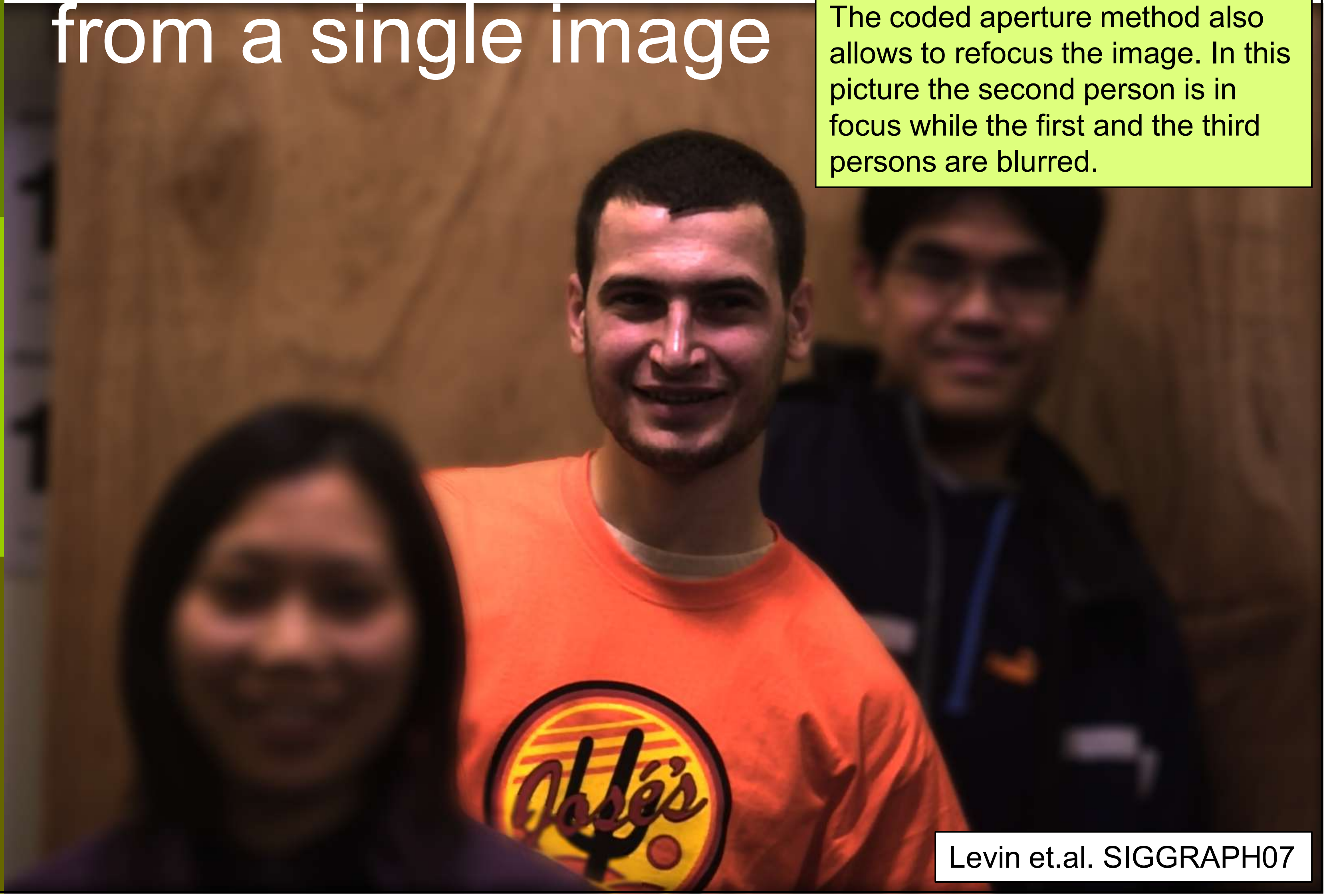
With a conventional sharpening filter

This filter performs worse because it lacks information about the coded aperture

Levin et.al. SIGGRAPH07

Application: Digital refocusing from a single image

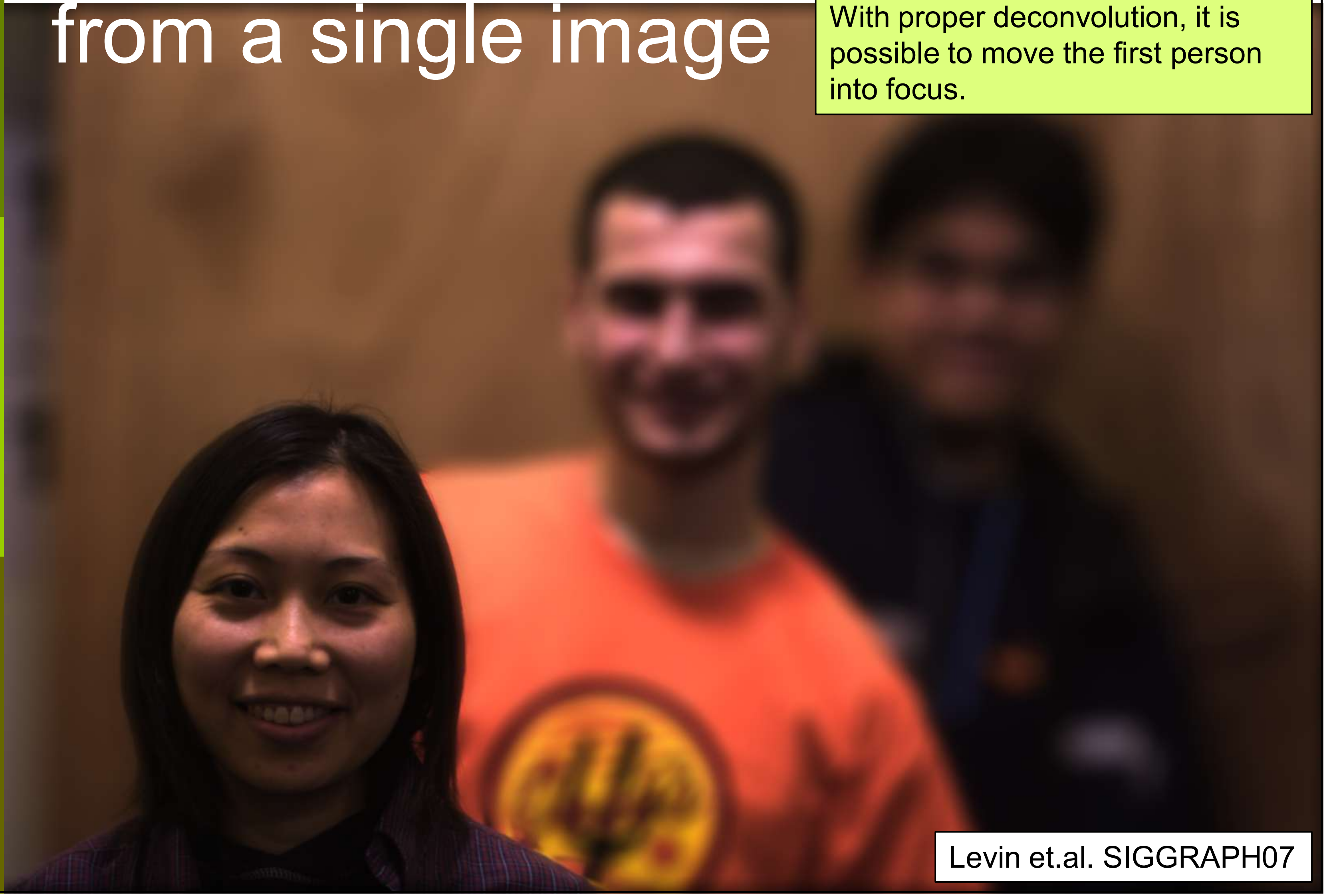
The coded aperture method also allows to refocus the image. In this picture the second person is in focus while the first and the third persons are blurred.



Levin et.al. SIGGRAPH07

Application: Digital refocusing from a single image

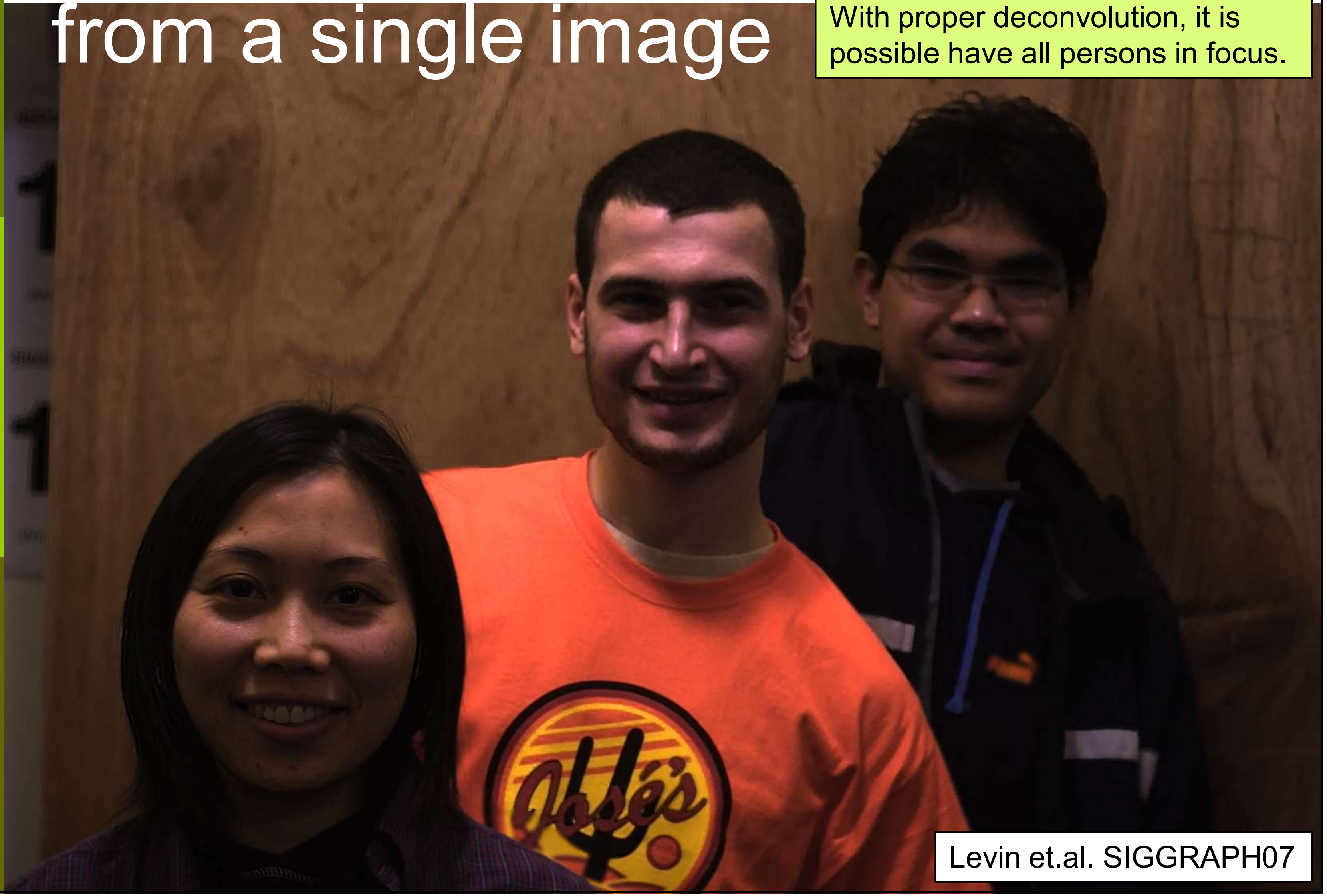
With proper deconvolution, it is possible to move the first person into focus.



Levin et.al. SIGGRAPH07

Application: Digital refocusing from a single image

With proper deconvolution, it is possible have all persons in focus.



Levin et.al. SIGGRAPH07

Added bonus: depth estimation

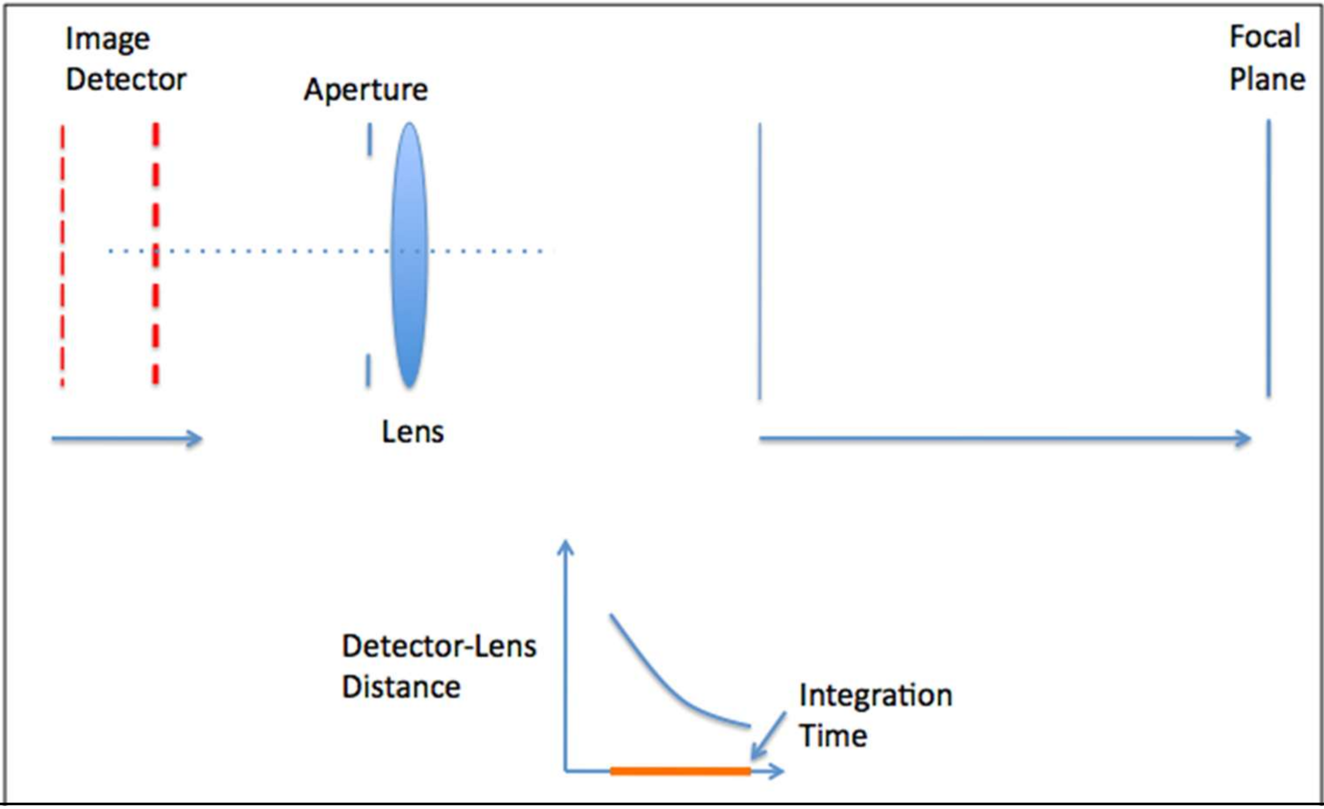
Levin et.al. SIGGRAPH07

The tradeoff between Depth-of-Field and Noise

- In a normal camera, a large aperture gives a small depth-of-field, much light and low noise
- In a normal camera, a small aperture gives a large depth-of-field, less light and high noise
- The flexible depth of field (EDOF) is a method to get a large depth-of-field and low noise at the same time
- Described in: "Flexible Depth of Field Photography," H. Nagahara, S. Kuthirummal, C. Zhou, and S.K. Nayar, European Conference on Computer Vision (ECCV), 2008.
- Also described in the video by Shree Nayar, Time 31:50-40.00. Link on the course webpage.

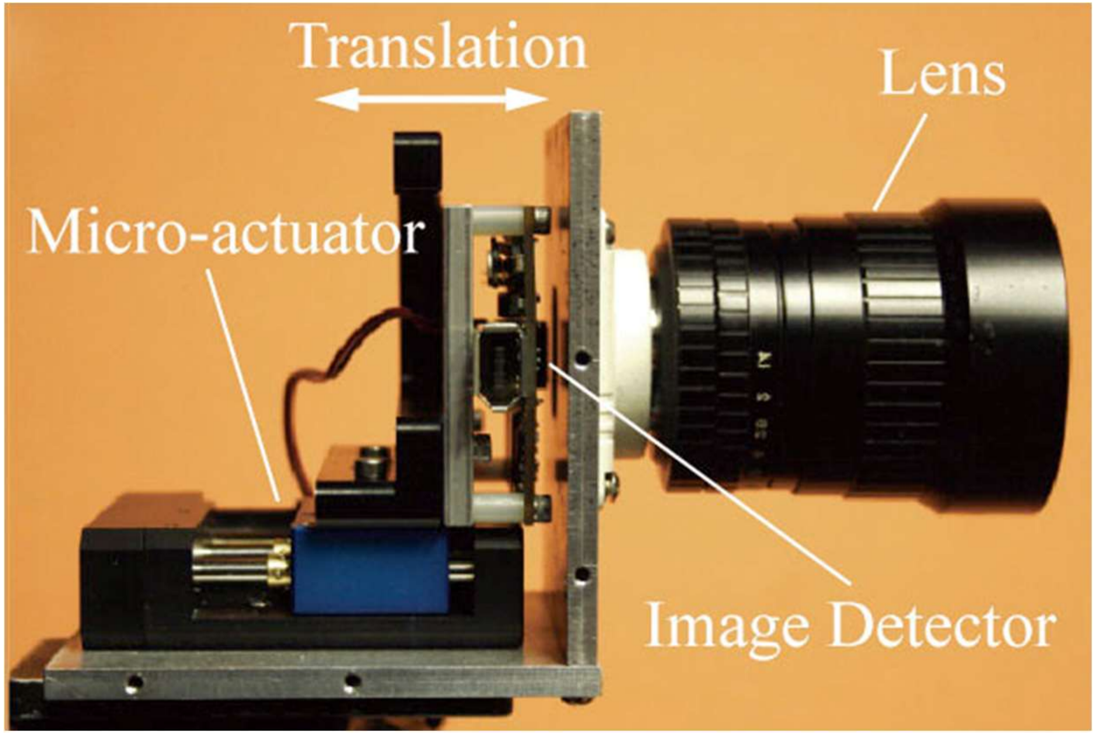
Flexible depth of field (EDOF)

- Aimed at extending depth of field
- Based on moving the sensor during exposure



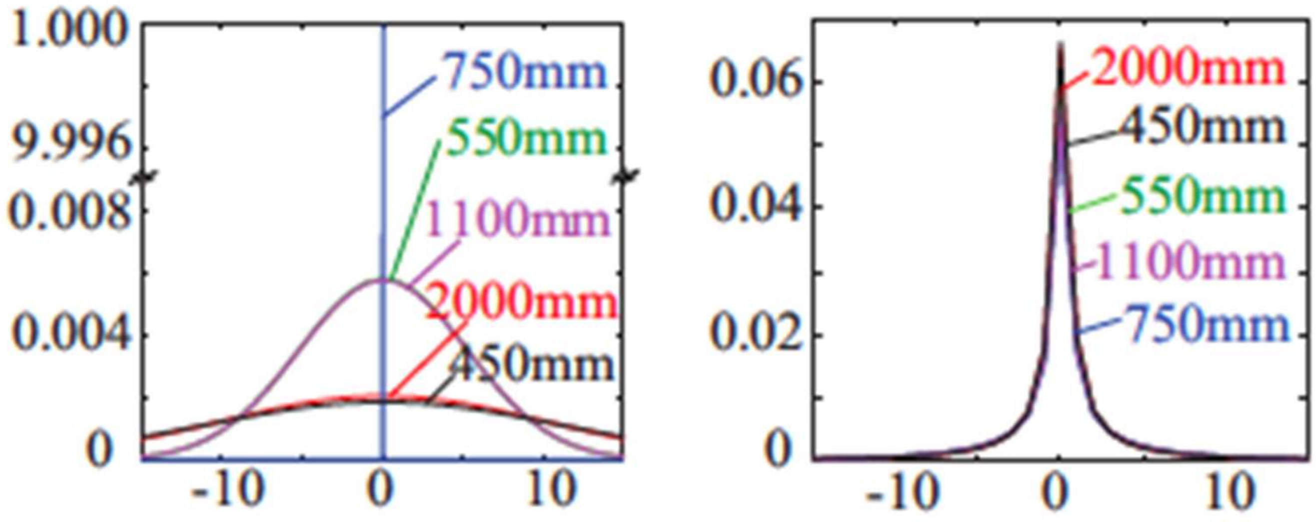
Flexible depth of field (EDOF)

- Here is shown a prototype setup.
- The sensor is being translated mechanically with a micro-actuator.



The PSF becomes independent on object distance

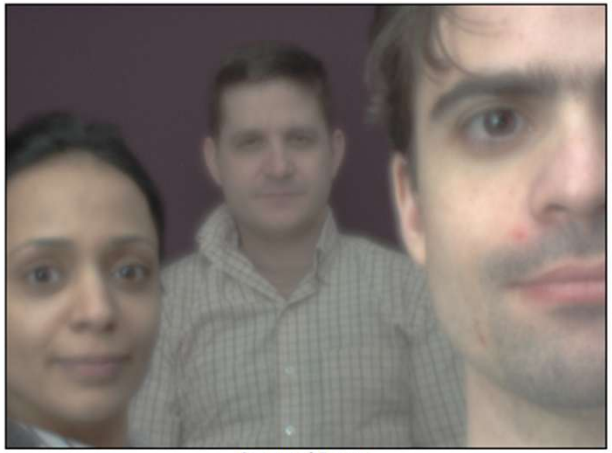
- Thus, deblurring can be performed with a fixed filter for all image parts!



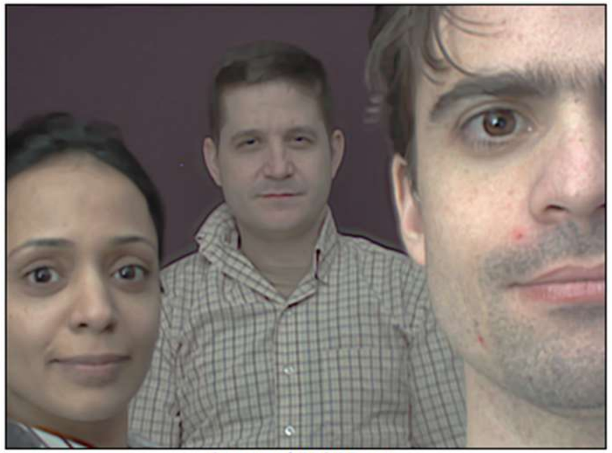
(c) Normal Camera PSF (Gaussian) (d) EDOF Camera IPSF

Imaging with a flexible depth of field (EDOF)

EDOF captured image



Captured Image
(f/1.4, T=0.36sec)



Computed EDOF Image

EDOF computed image

Normal camera, large aperture => small depth-of-field, much light

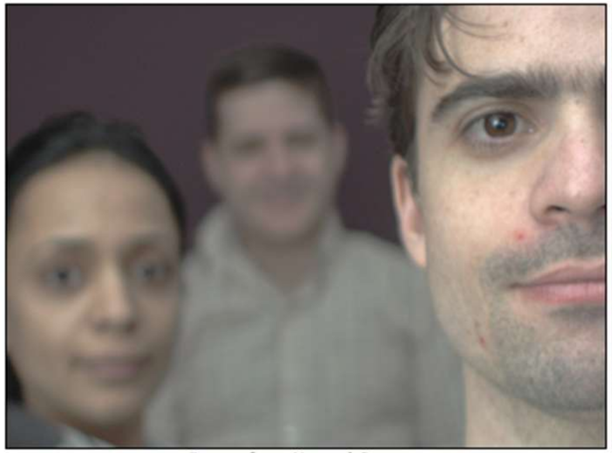


Image from Normal Camera
(f/1.4, T=0.36sec, Near Focus)

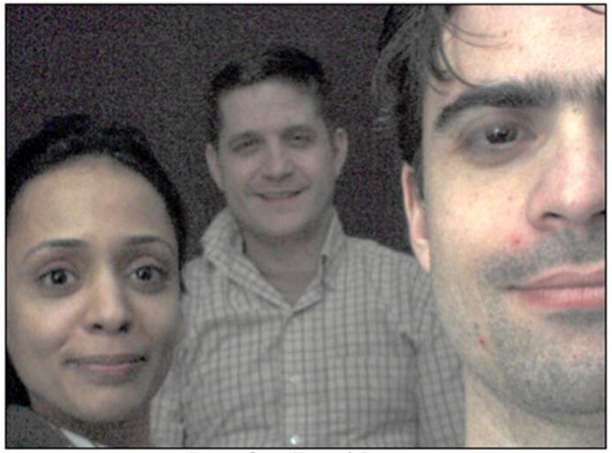


Image from Normal Camera
(f/8, T=0.36sec, Near Focus) with Scaling

Normal camera, small aperture => large depth-of-field, less light and more noise