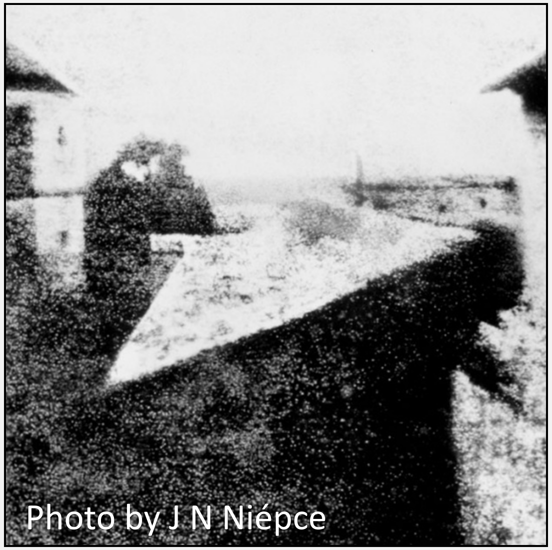
Department of Computer Science
University of Bristol

COMS30030 - Image Processing and Computer Vision



Video Lecture 03

Image Acquisition & Representation

Majid Mirmehdi | majid@cs.bris.ac.uk

So what does the human visual system see?

Or what should an automatic *computer vision system* see?

slide by Fei Fei, Fergus & Torralba

Detection: are there cars?

slide by Fei Fei, Fergus & Torralba

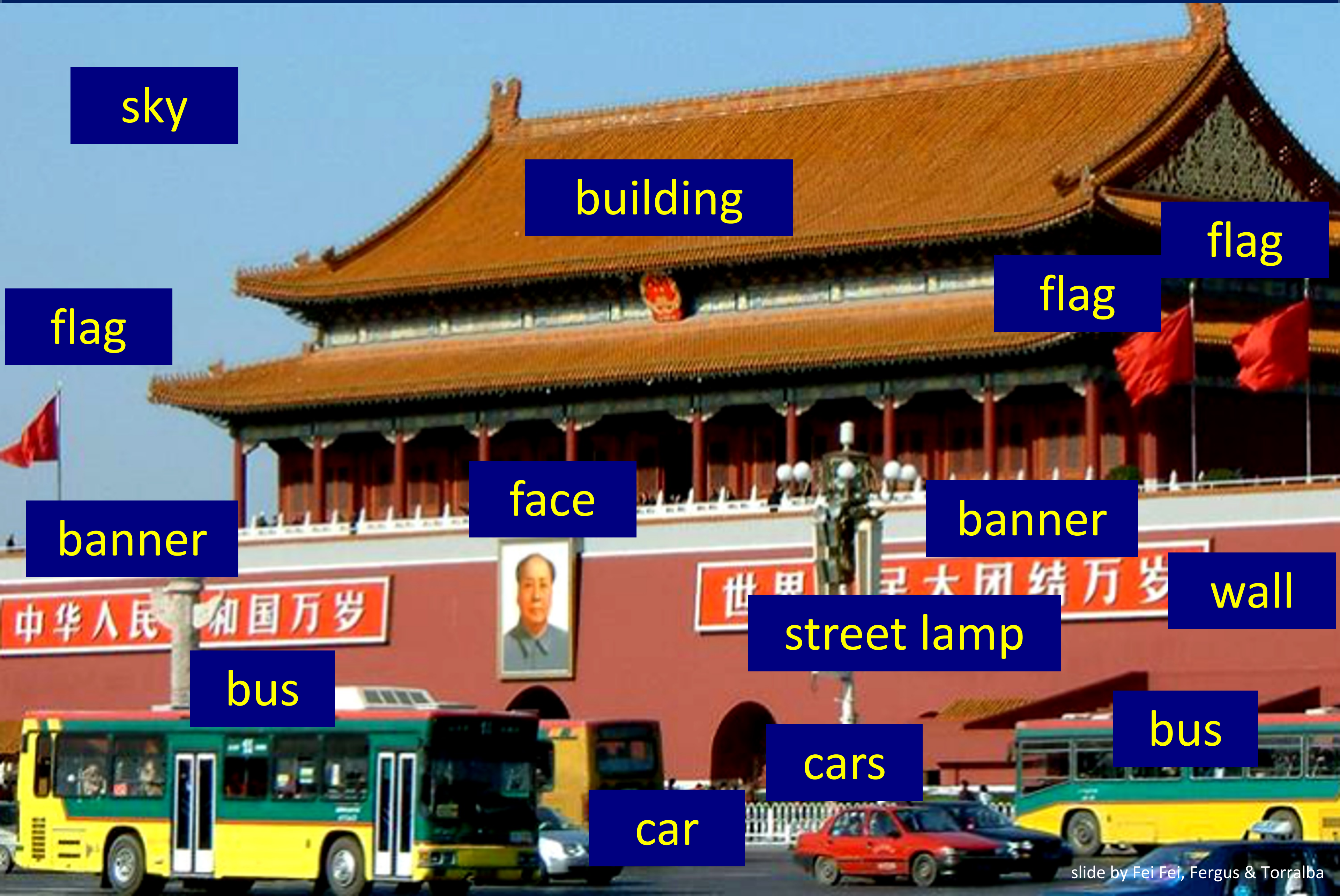
Verification: is that a bus?

slide by Fei Fei, Fergus & Torralba

Identification: is that a picture of Mao?

slide by Fei Fei, Fergus & Torralba

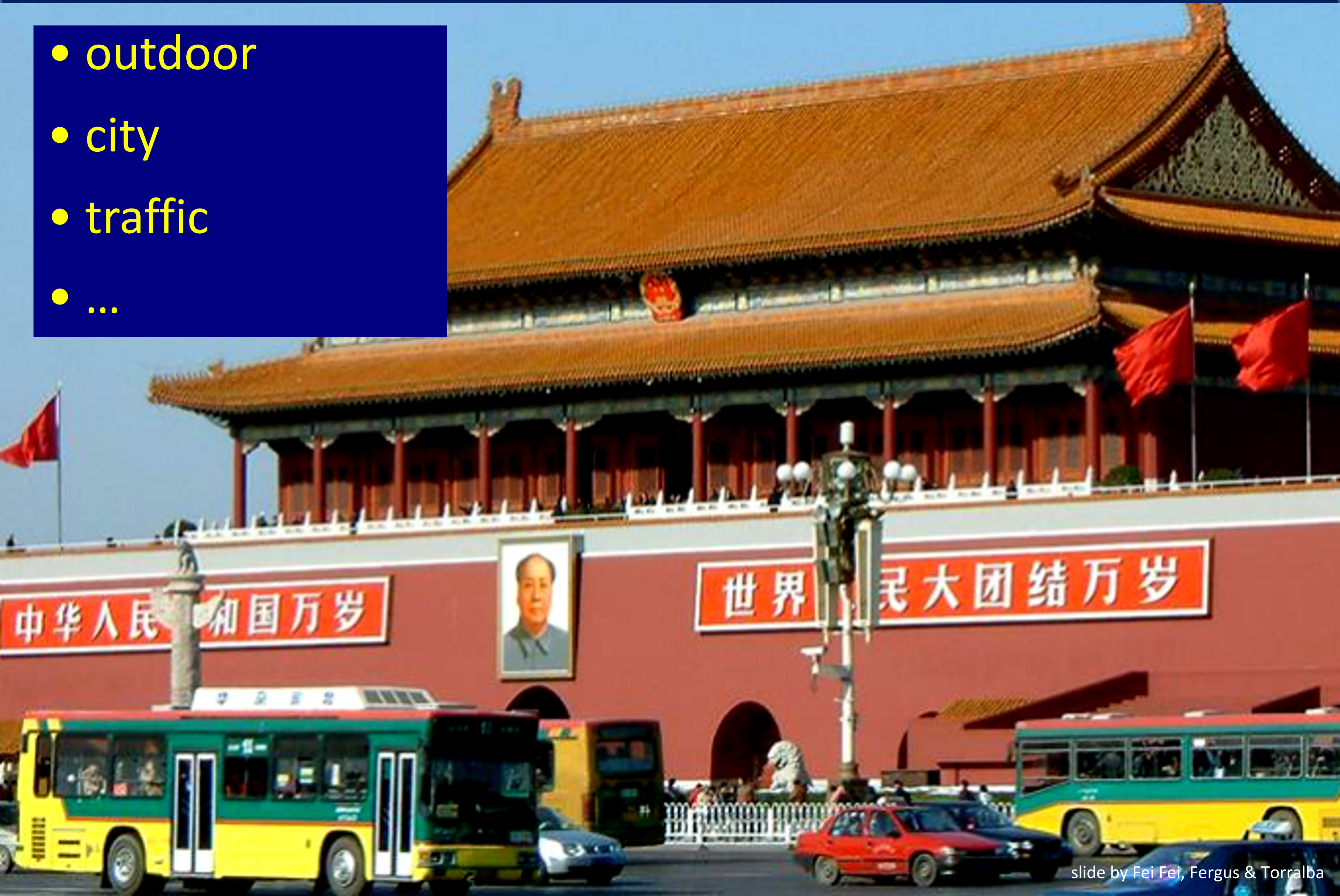
Object categorization



slide by Fei Fei, Fergus & Torralba

Scene and context categorization

- outdoor
- city
- traffic
- …



slide by Fei Fei, Fergus & Torralba

Rough 3D layout, depth ordering

slide by Fei Fei, Fergus & Torralba

Challenges

So what are the challenges faced by a *computer vision system*?

9 challenges coming up, in no particular order…

Challenge 1: view-point variation

Michelangelo 1475-1564

original Challenge slides by Efros, Ullman, and others

Challenge 2: illumination

original Challenge slides by Efros, Ullman, and others

Challenge 3: occlusion

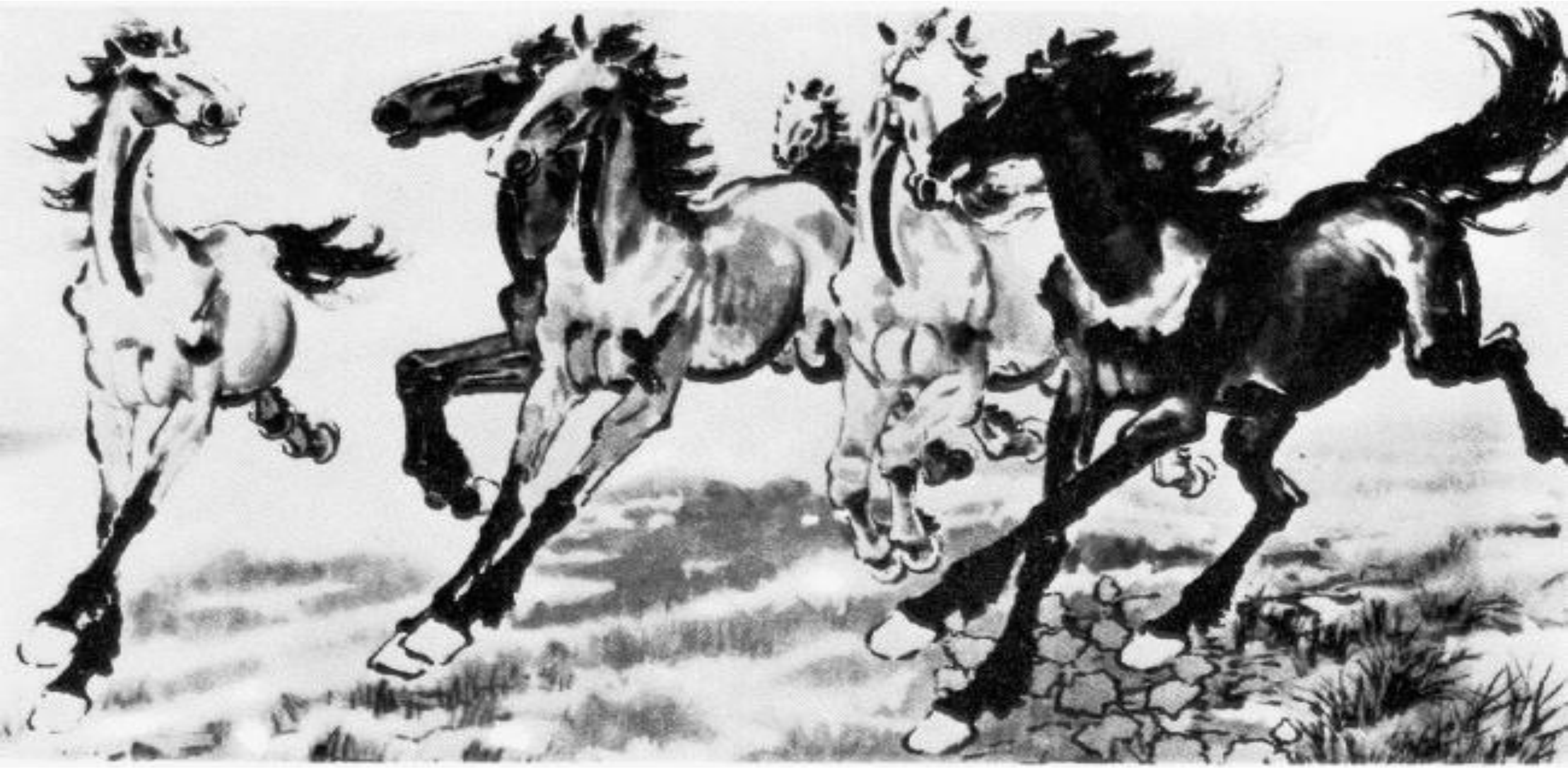
Magritte, 1957

original Challenge slides by Efros, Ullman, and others

Challenge 4: scale

original Challenge slides by Efros, Ullman, and others

Challenge 5: deformation



Xu, Beihong 1943

original Challenge slides by Efros, Ullman, and others

Challenge 6: background clutter

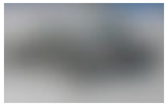
Klimt, 1913

original Challenge slides by Efros, Ullman, and others

Challenge 7: object intra-class variation

original Challenge slides by Efros, Ullman, and others

Challenge 8: local ambiguity



original Challenge slides by Efros, Ullman, and others

Challenge 9: the world behind the image

The Basics of Image Acquisition and Representation

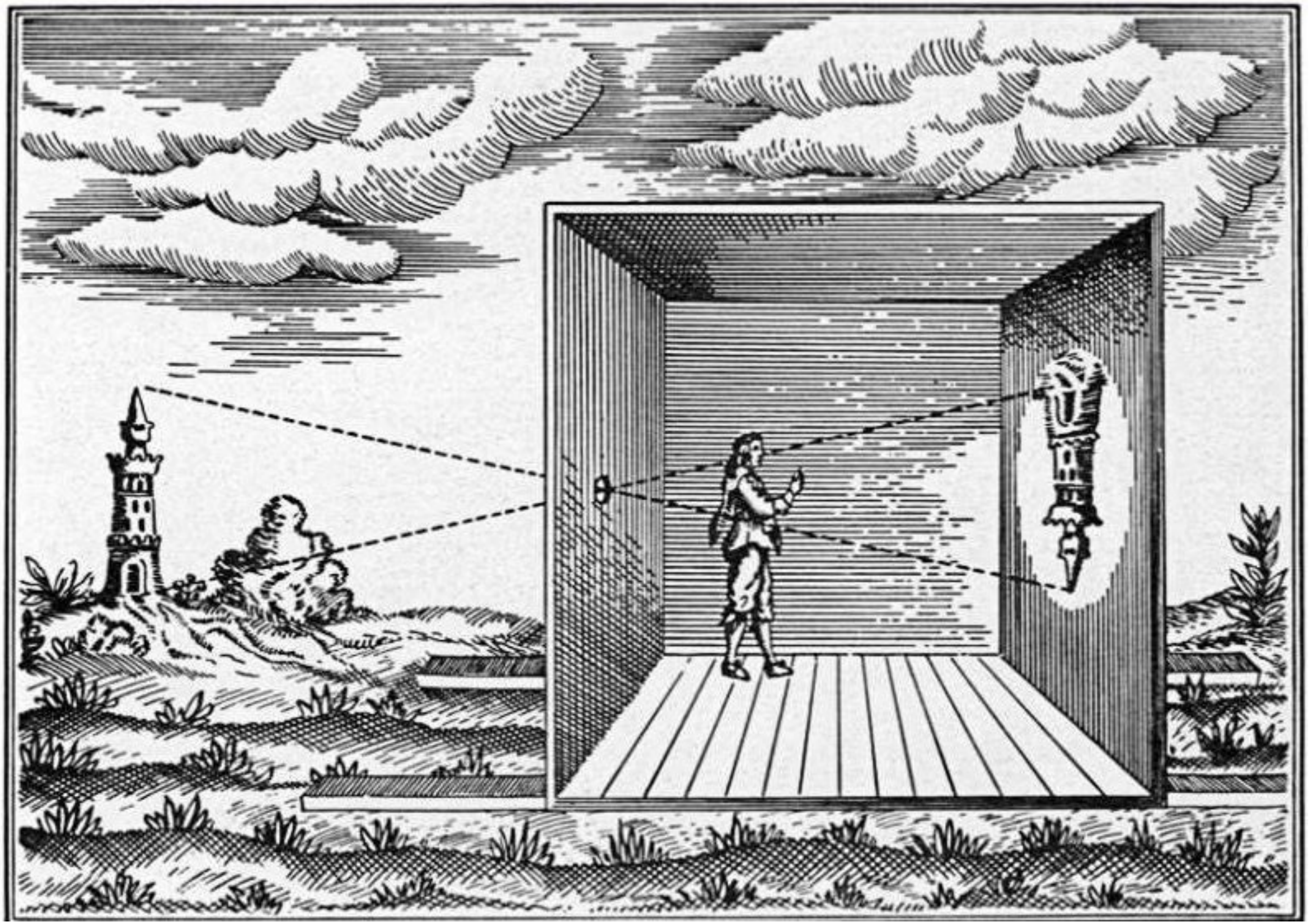
Images as Sensory Data

- How are images acquired?
- Which processes influence digital image formation?

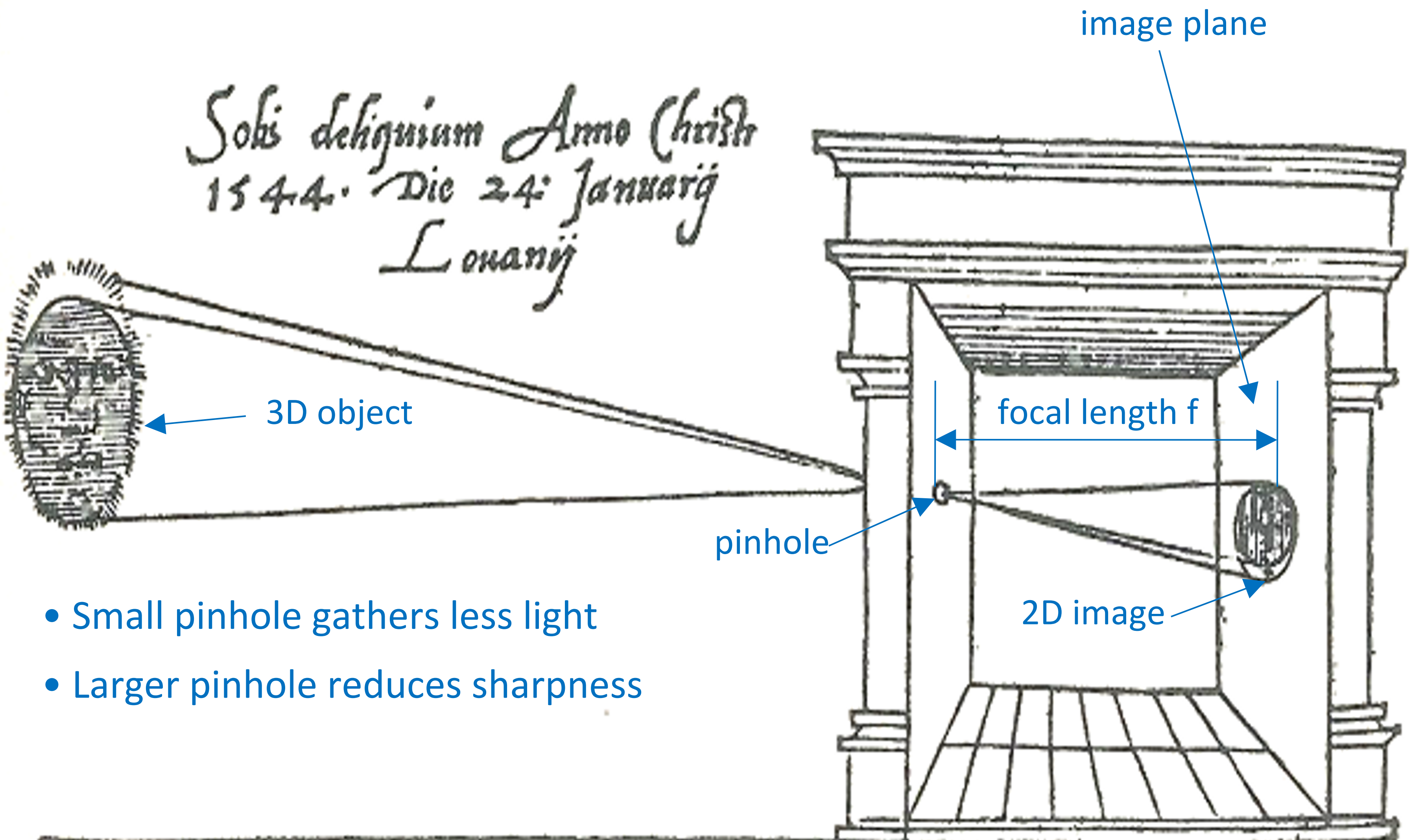
Images as Structured Data

- How can digital images be represented?

The Camera Obscura (Pinhole Camera)



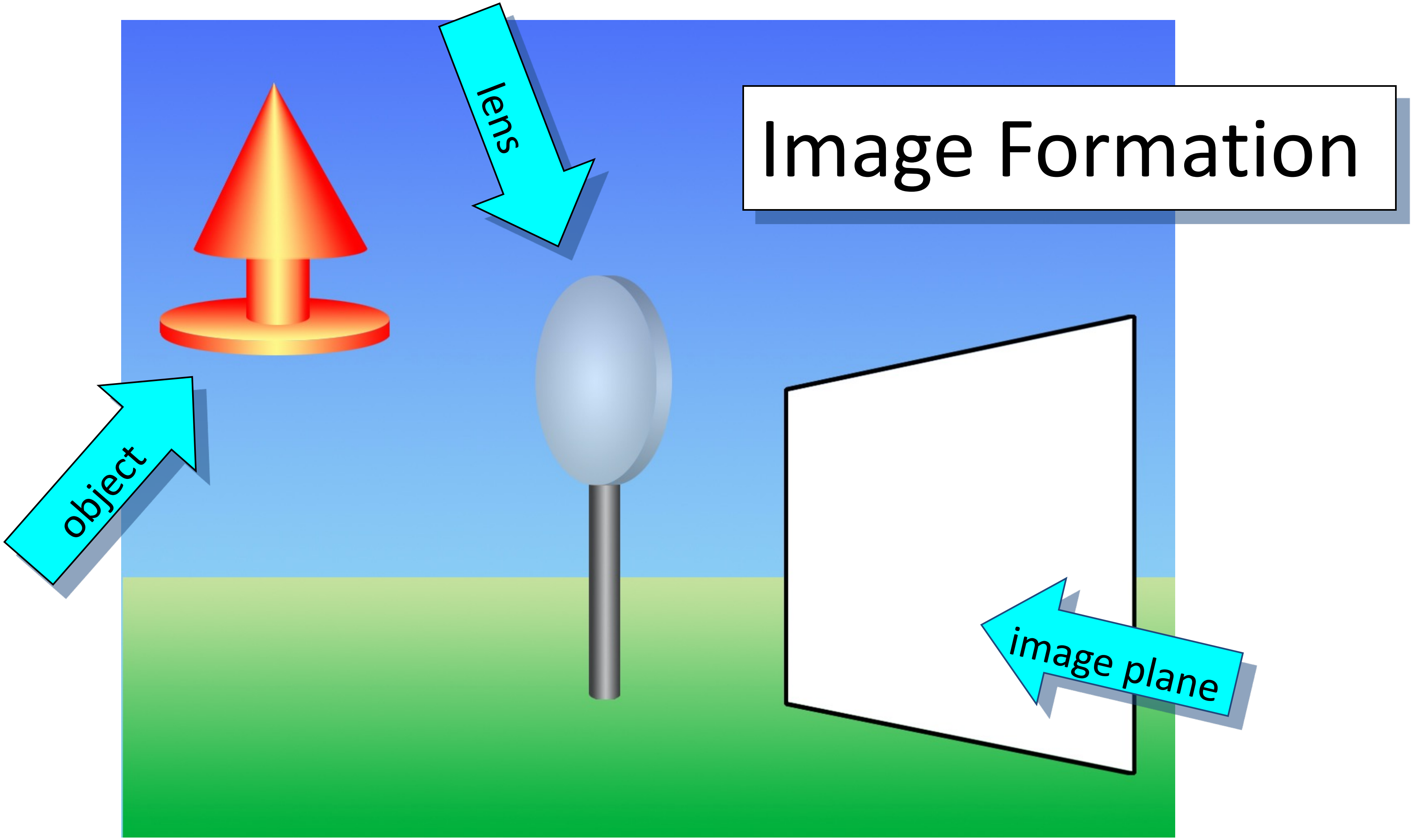
The Camera Obscura (Pinhole Camera)



- Small pinhole gathers less light
- Larger pinhole reduces sharpness

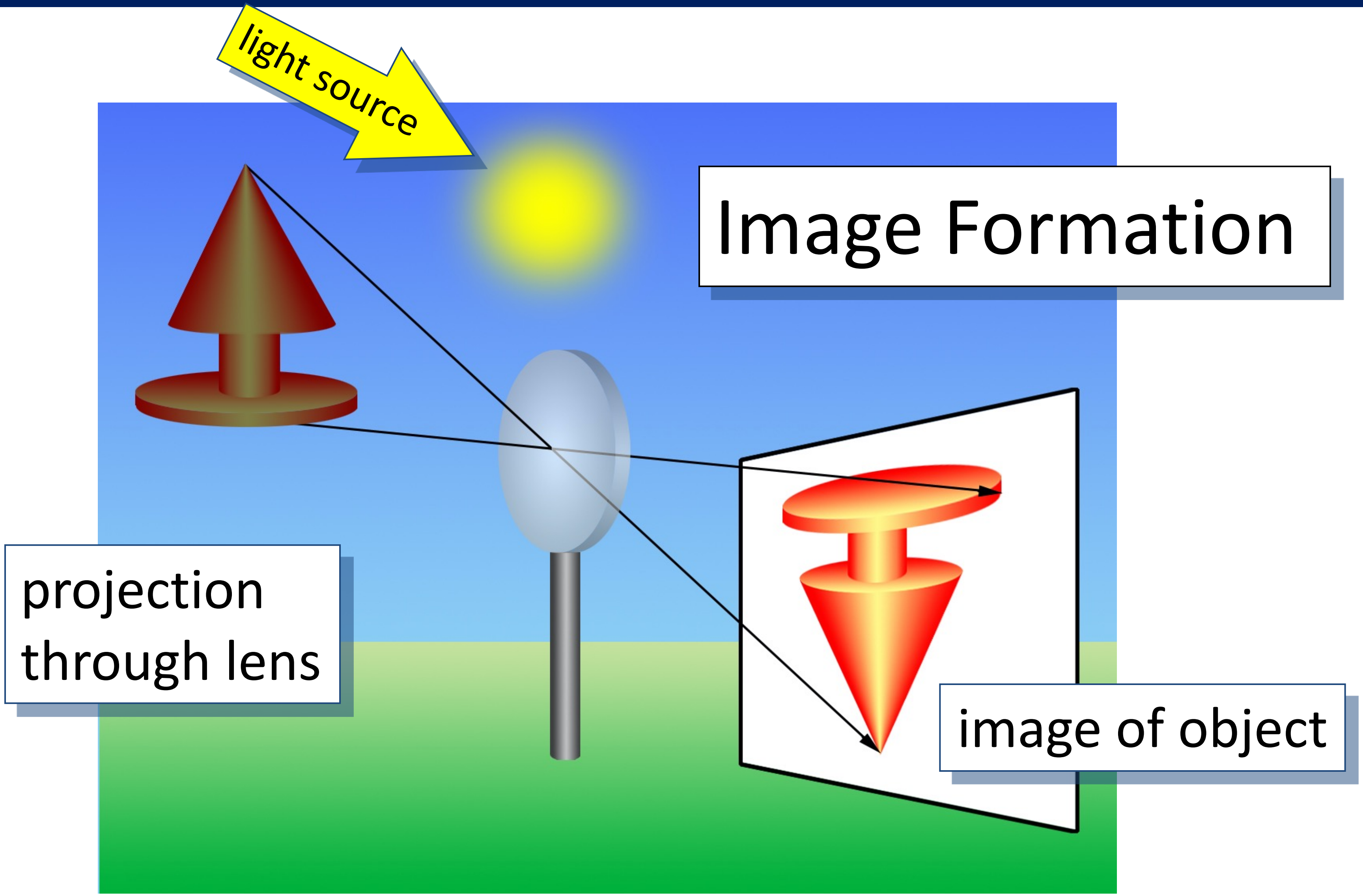
First published picture of camera obscura in Gemma Frisius' 1545 book *De Radio Astronomica et Geometrica*

Digital Image Acquisition



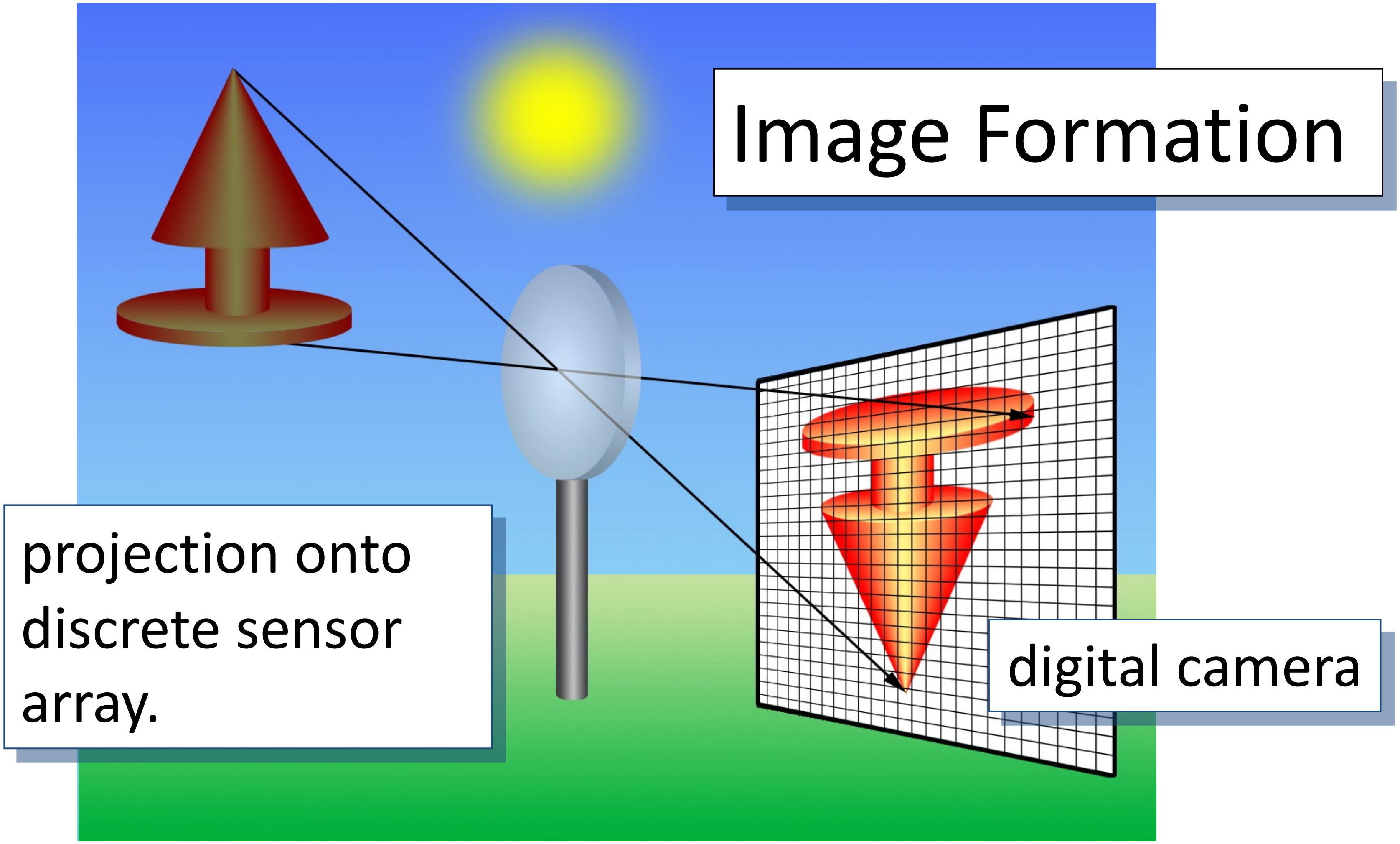
Slide by R A Peters

Digital Image Acquisition



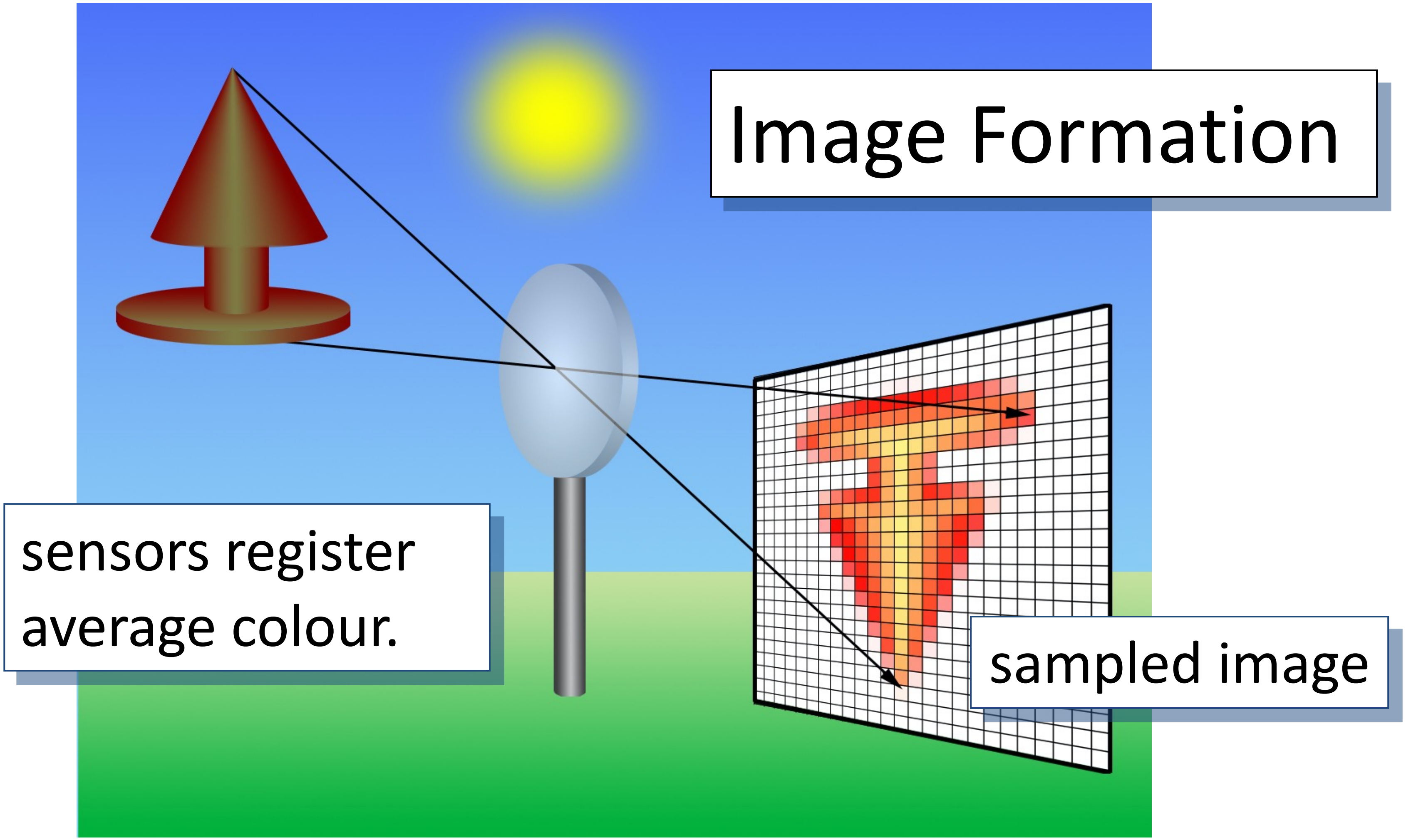
Slide by R A Peters

Digital Image Acquisition



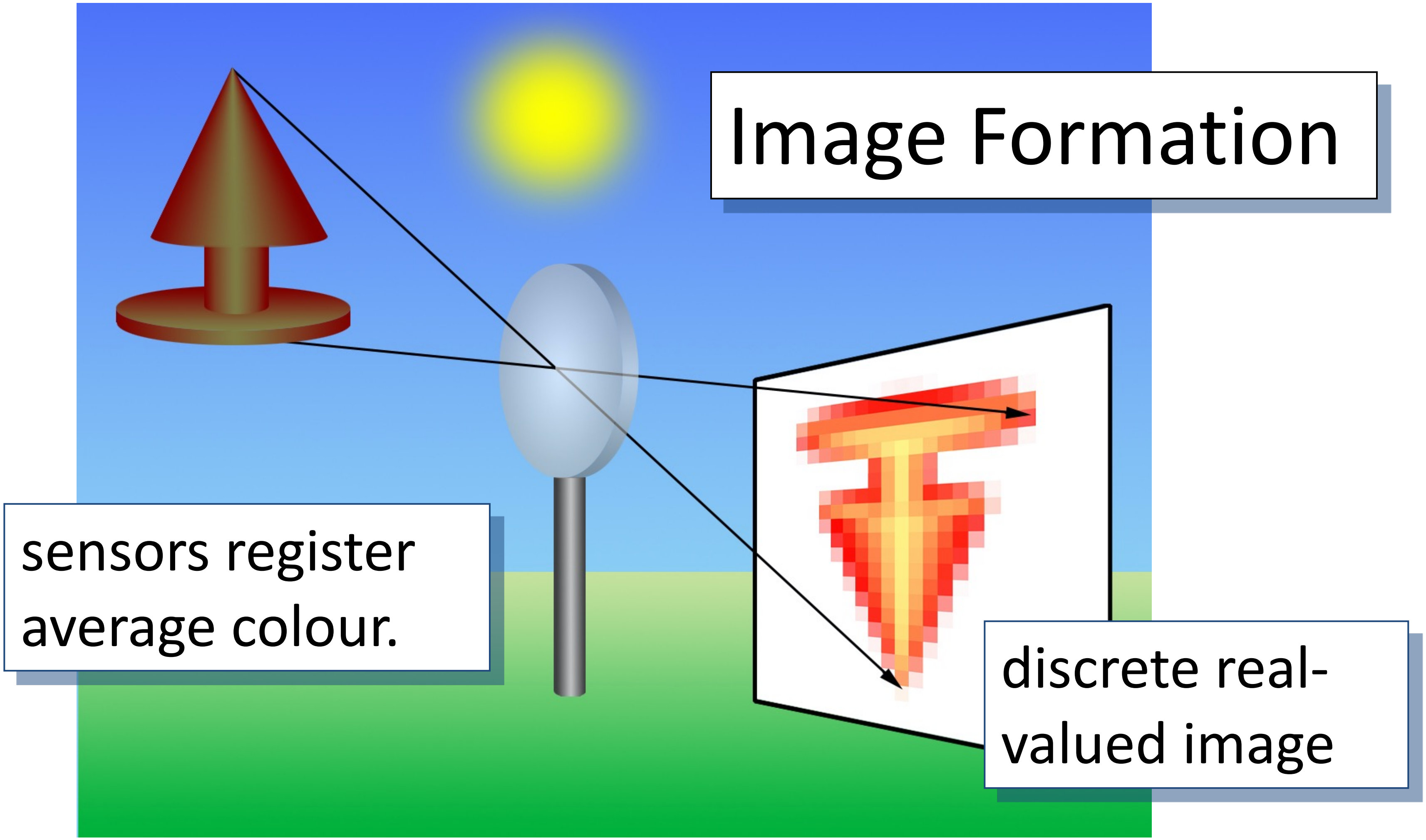
Slide by R A Peters

Digital Image Acquisition



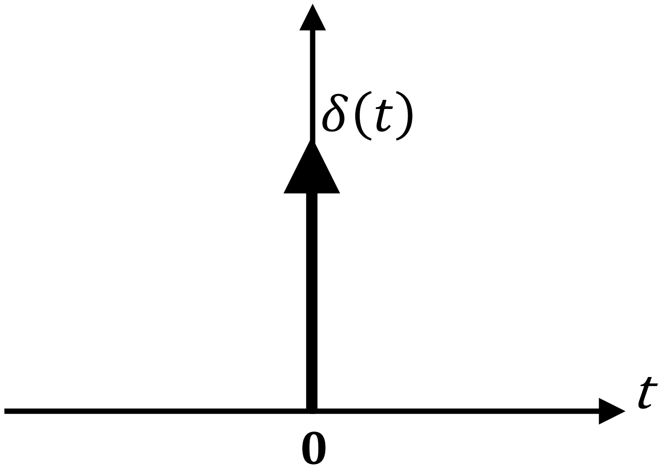
Slide by R A Peters

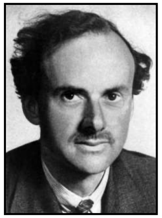
Digital Image Acquisition



Slide by R A Peters

Modelling a Spatial Brightness Pulse - Dirac Delta-Function

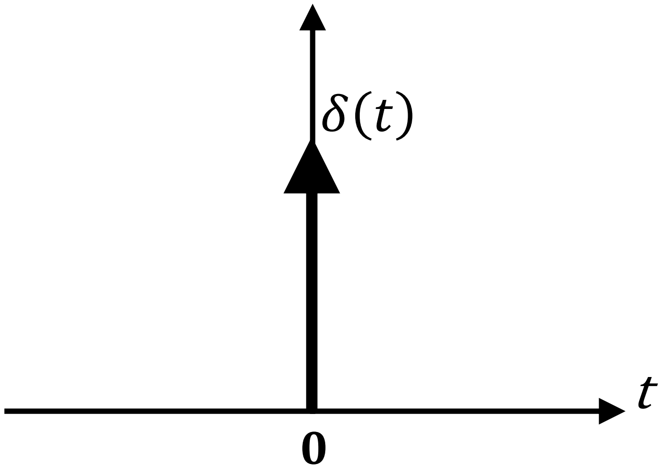


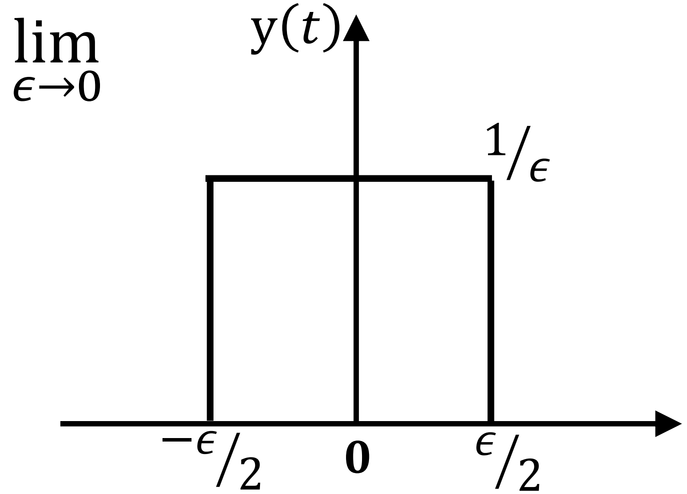

Paul Dirac

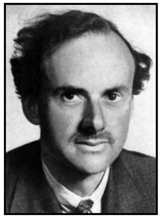
Definition:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

Modelling a Spatial Brightness Pulse - Dirac Delta-Function






Paul Dirac

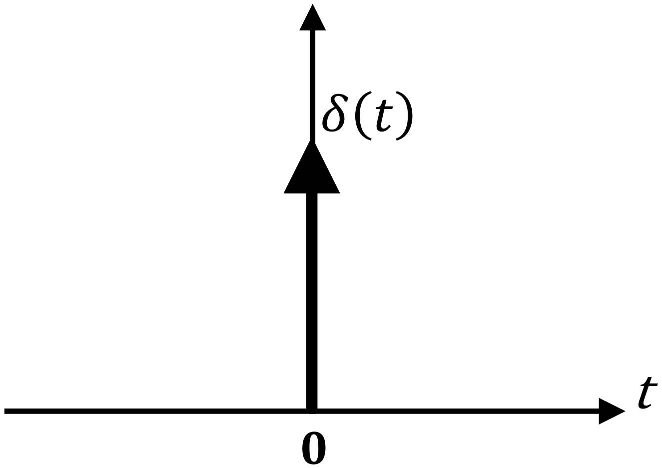
Definition:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

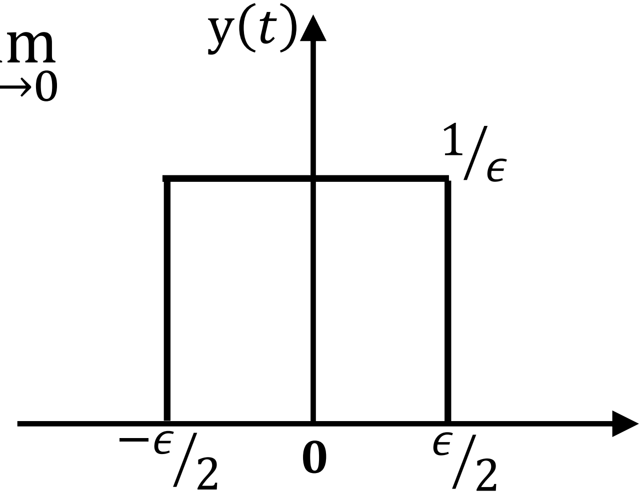
Intuitively:

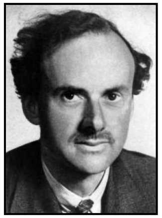
$$\delta(t) = \lim_{\epsilon \to 0}[y_\epsilon(t)]$$

Modelling a Spatial Brightness Pulse - Dirac Delta-Function



$\lim_{\epsilon \to 0}$





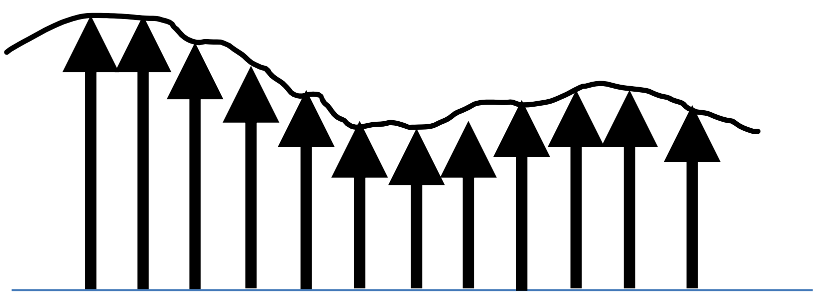
Paul Dirac

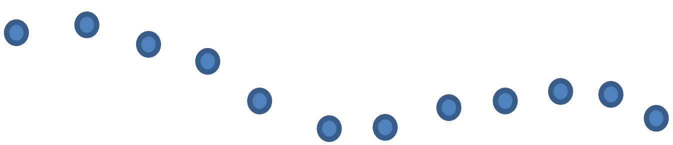
Definition:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

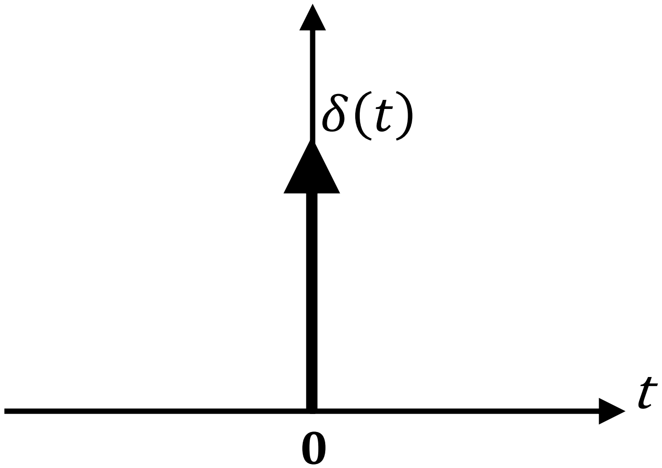
Intuitively:

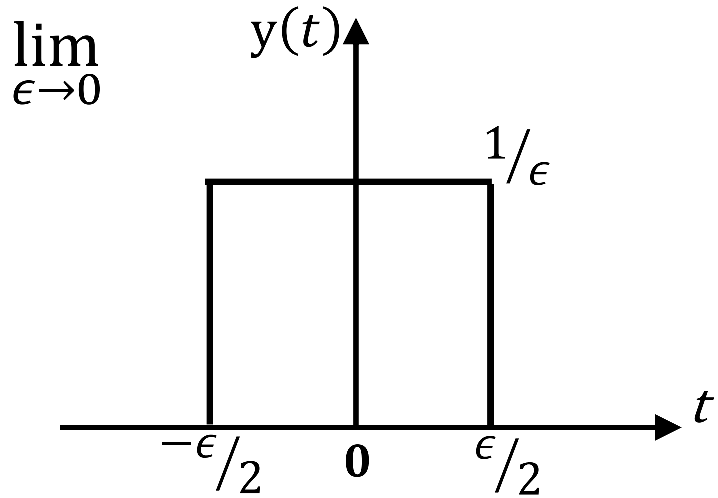
$$\delta(t) = \lim_{\epsilon \to 0}[y_\epsilon(t)]$$

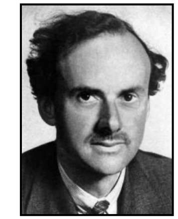




Modelling a Spatial Brightness Pulse - Dirac Delta-Function






Paul Dirac

Definition:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

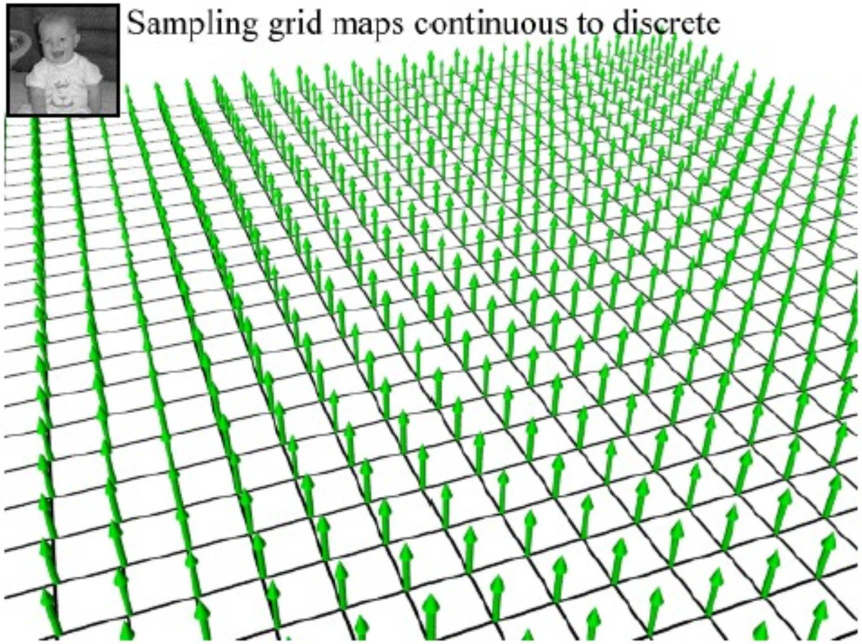
Intuitively:

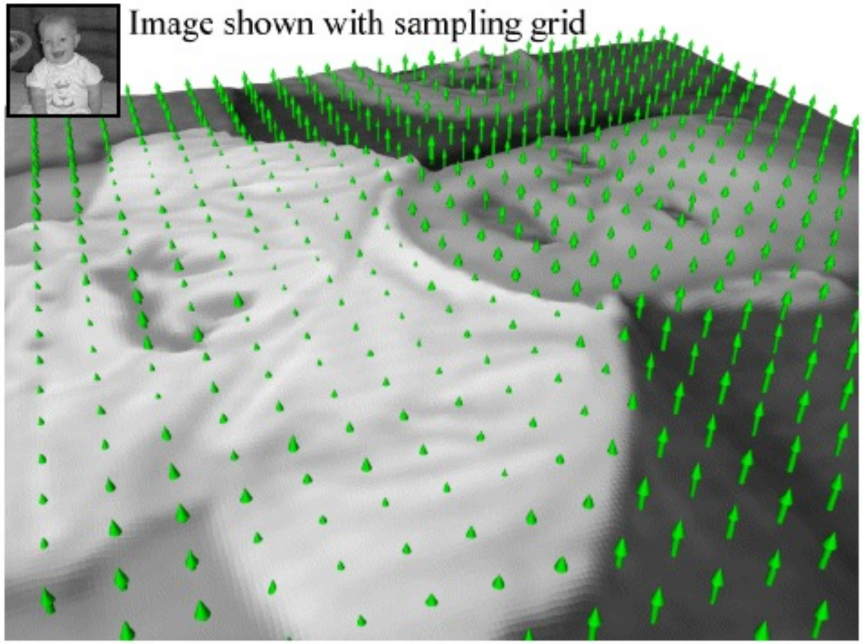
$$\delta(t) = \lim_{\epsilon \to 0}[y_\epsilon(t)]$$

Sifting Property:

$$\int_{-\infty}^{\infty} f(t)\delta(t)dt = f(0)$$

Sampling in 2D To Obtain An Image





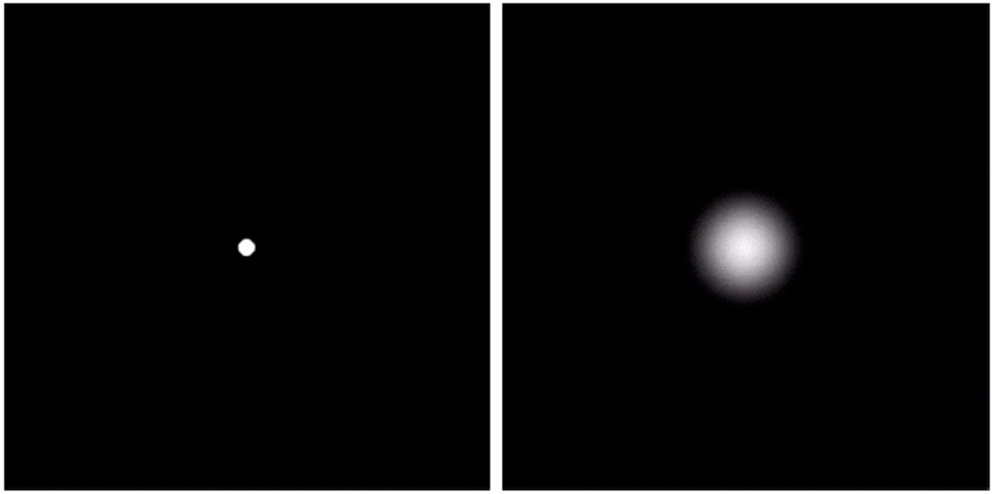
Images from Durand, Cutler and Hanrahan

The sifting property can be used to express a 2D 'image function' as a linear combination of 2D Dirac pulses located at points (a,b) that cover the whole image plane:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a,b)\, \delta(a-x, b-y)\, da\, db = f(x,y)$$

The Point Spread Function

scene → Optical System → image



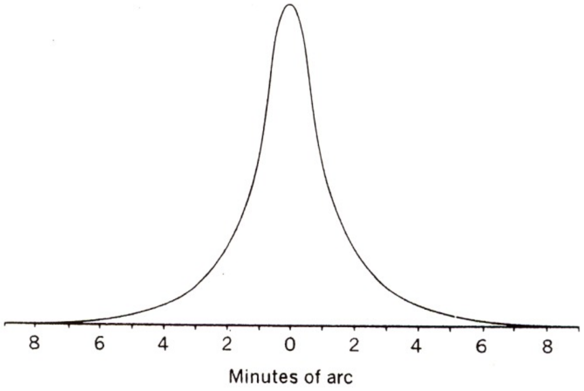
- Ideally, the optical system should be mapping point information to points again.

- However, optical systems are never ideal.

$\delta(x)$ → Optical System → $PSF(x)$

point source → point spread function

- Superposition Principle:
  An image is the sum of the PSF of all its points.



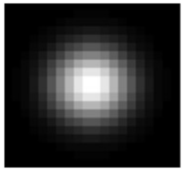
- Point spread function of Human Eyes

PSF Example

$$g(x, y) = f(x, y) * h(x, y)$$

$f(x, y)$ Original

$h(x, y)$



Blurry outcome $g(x, y)$

$h(x, y)$ is the PSF of the imaging device.

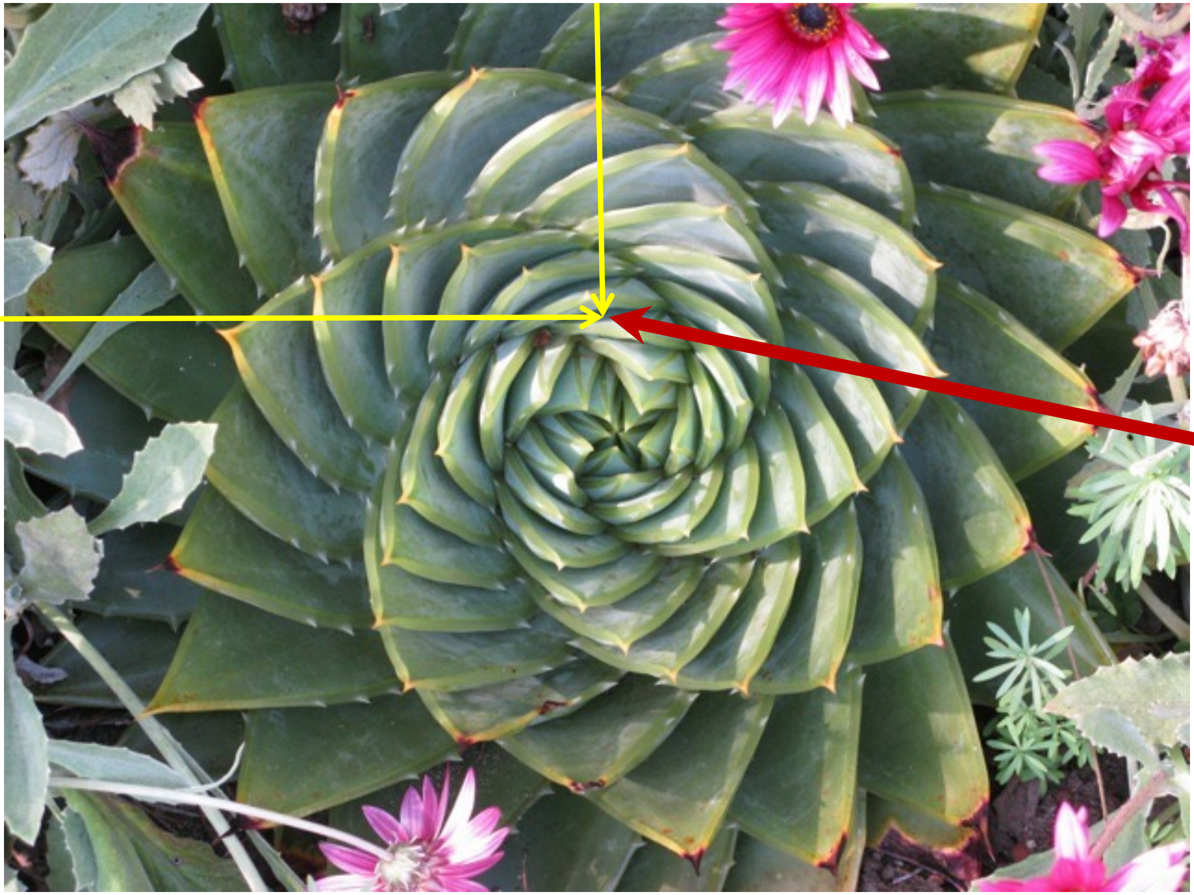
Adapted from a slide by A. Zisserman

How to model an image?

Addressing

y_j

x_i



Colour

(c_R, c_G, c_B)

How to represent a digital image?

$$f[x,y] = \begin{bmatrix} f[0,0] & f[0,1] & \cdots & f[0,N-1] \\ f[1,0] & f[1,1] & \cdots & f[1,N-1] \\ \vdots & \vdots & \ddots & \vdots \\ f[M-1,0] & f[M-1,1] & \cdots & f[M-1,N-1] \end{bmatrix}$$

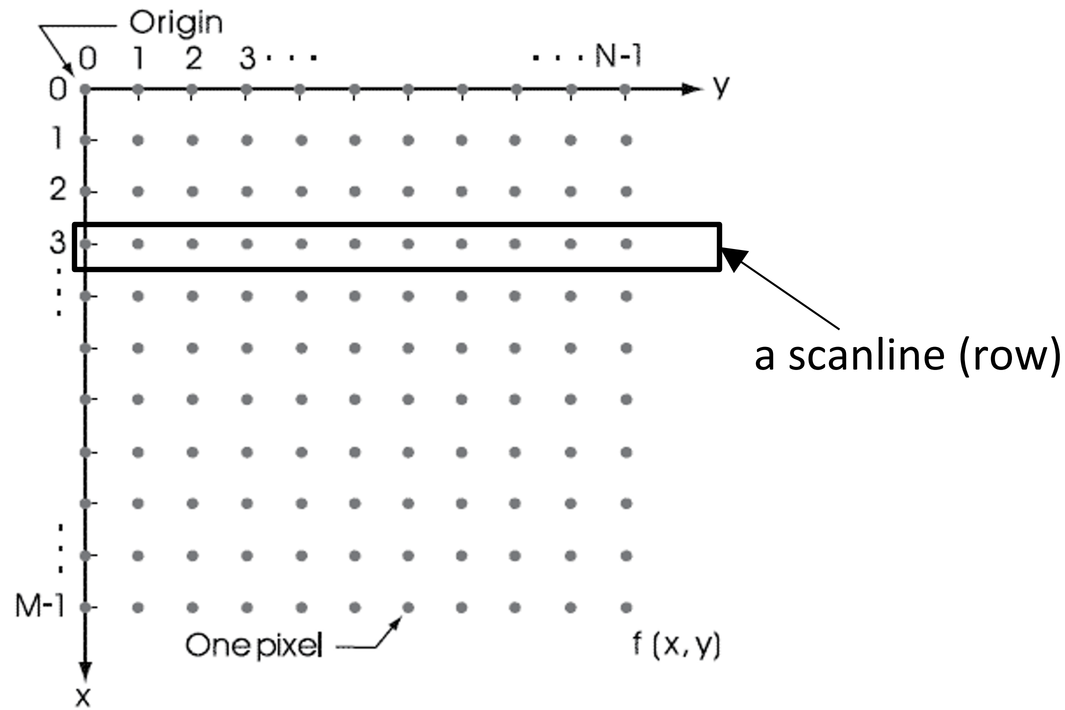


Image Representation: Colour Spaces

$$f(x, y) = (R, G, B)$$

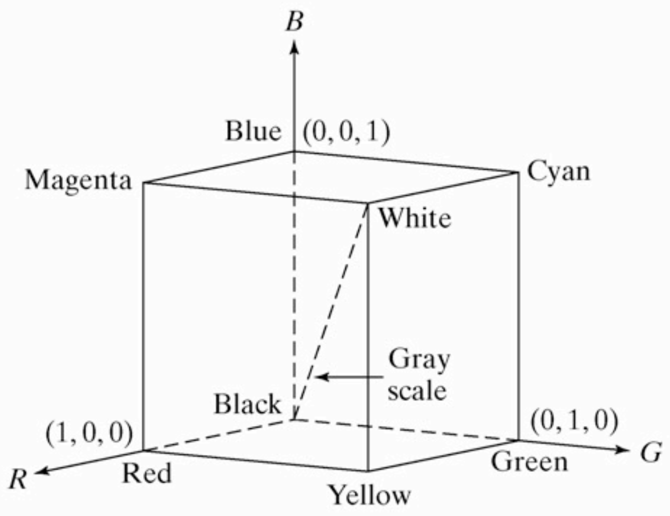
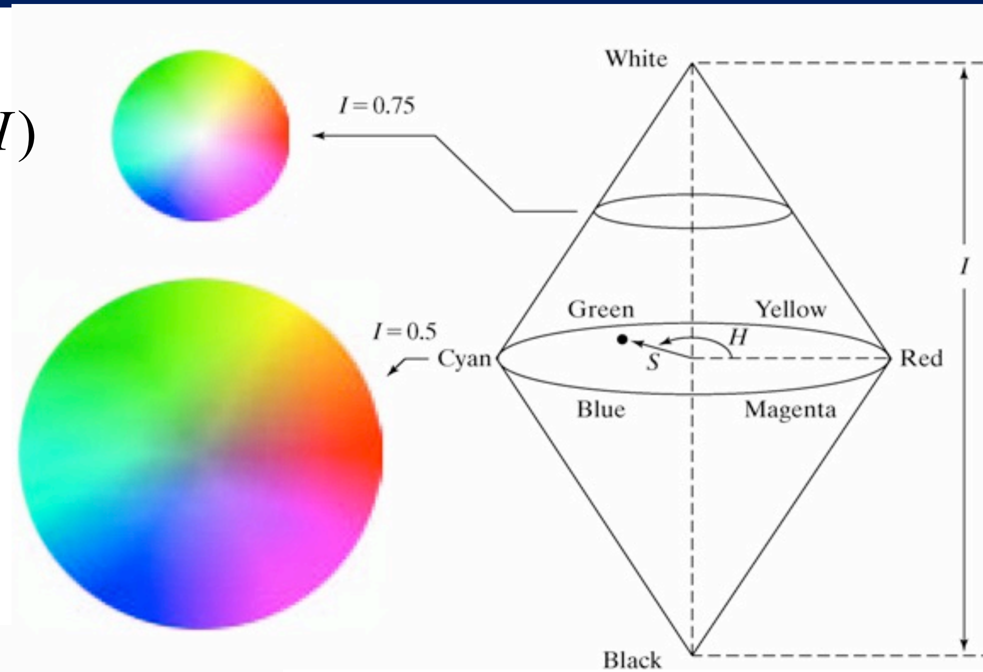
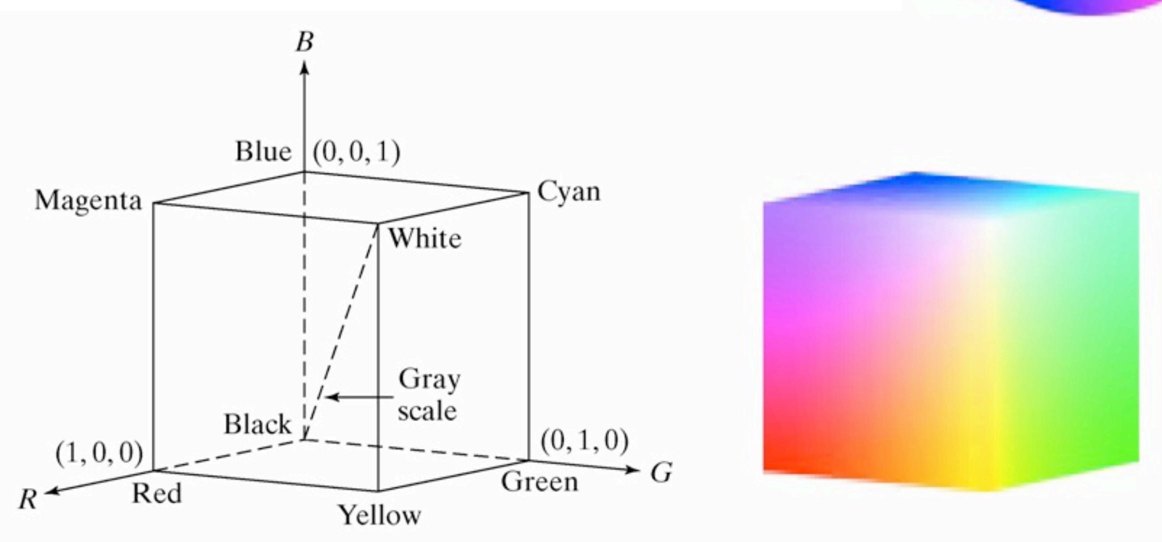


Image Representation: Colour Spaces

$$f(x,y) = (H,S,I)$$



$$f(x,y) = (R,G,B)$$



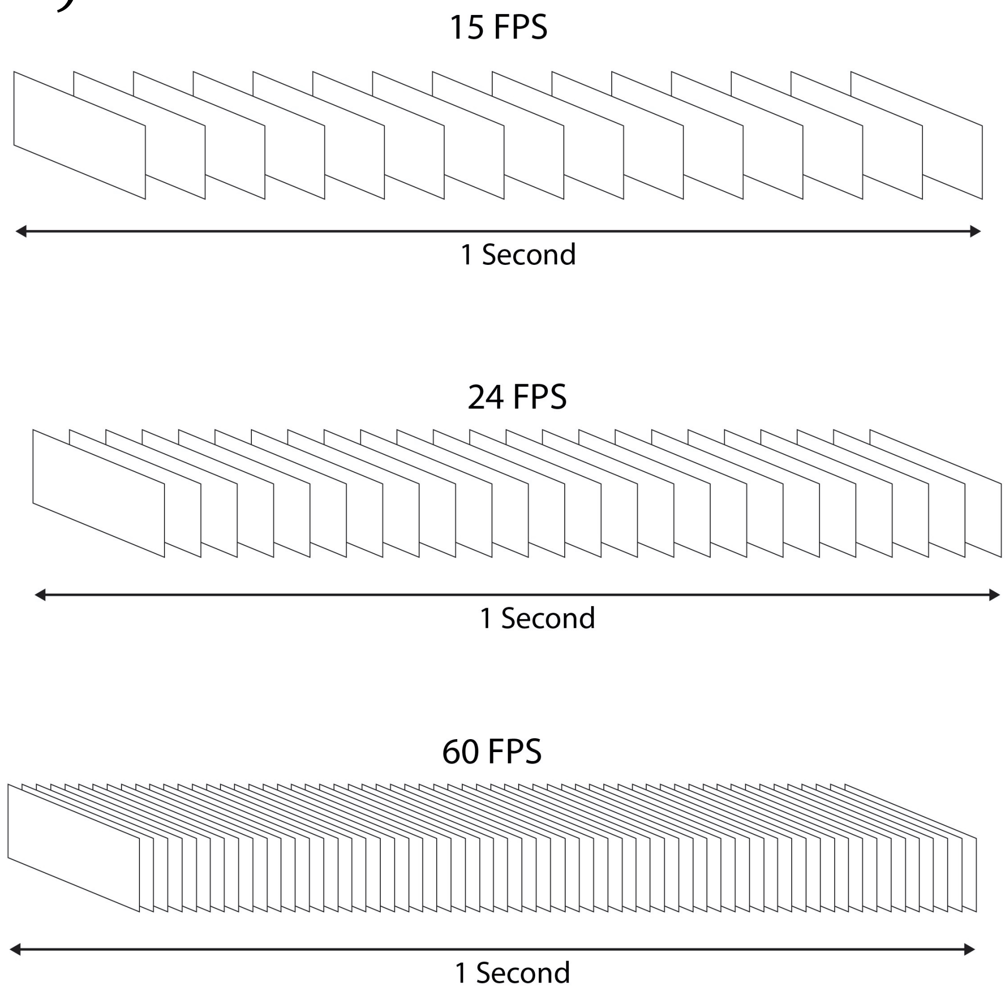
Other examples are:

$$f(x,y) = (L,a,b)$$
$$f(x,y) = (Y,U,V)$$

…

How to represent a video?

$$f[x, y, t] = (R, G, B)$$



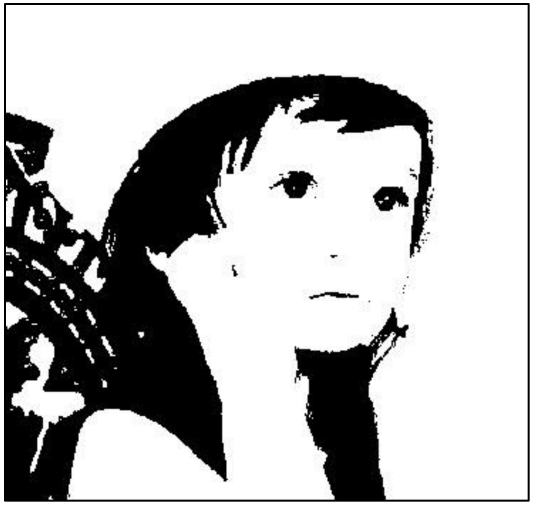
From quora.com

Quantization of the Image Function

Representing a continuously varying single channel image function $f(x, y)$ with a discrete one using quantization levels:

16 levels

6 levels



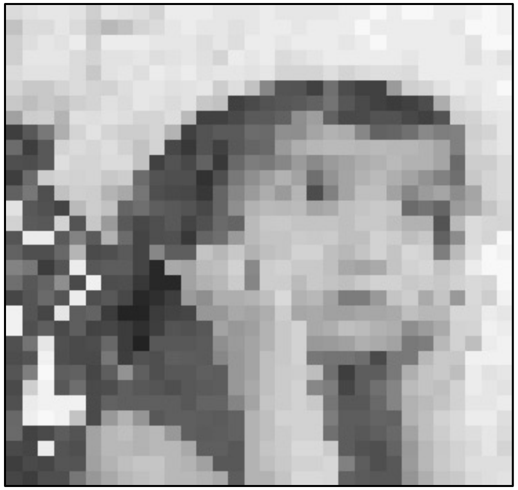
2 levels

Spatial Sampling in Practice

The effect of very sparse sampling … is often ALIASING

256 x256

64x64



32x32

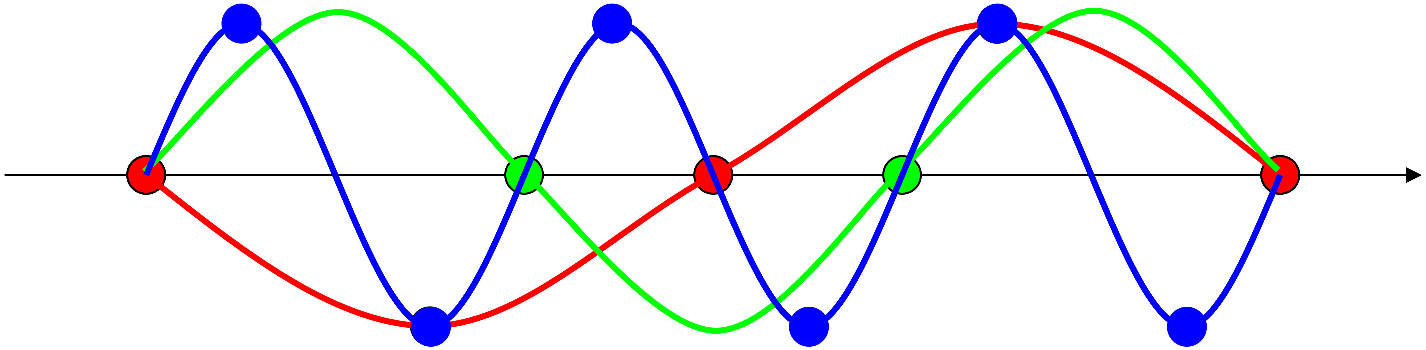
Anti-aliasing can be achieved by removing all spatial frequencies above a critical limit (so-called Shannon-Nyquist Limit).

Shannon's Sampling Theorem

"An analogue signal containing components up to some maximum frequency ***u*** may be completely reconstructed by regularly spread samples, provided the sampling rate is above 2***u*** samples per second."

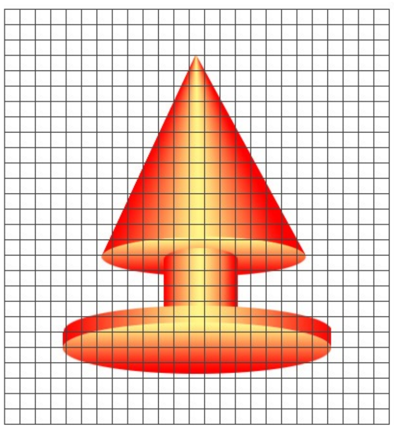
Claude Shannon

Also referred to as the Shannon-Nyquist criterion: Sampling **must** be performed above twice the highest (spatial) frequency of the signal to be lossless.



Sampling and Quantization Summary

pixel grid



real image

sampled

quantized

sampled & quantized

Slide by R A Peters

Next Week's Lectures

Convolution

The Fourier Domain