COMPUTATIONAL IMAGING

L3. Future of Imagers.

Dmitry Dylov

Associate Professor

Skoltech

Week 1 – Understand Imaging Math

Week 2 – Image Processing and Computational Illumination

Week 3 – Specialized cameras, Future of Imaging, Case studies

- Recap from the Labs

L2. Outline

CONSUMER-GRADE "FUTURE" CAMERAS:

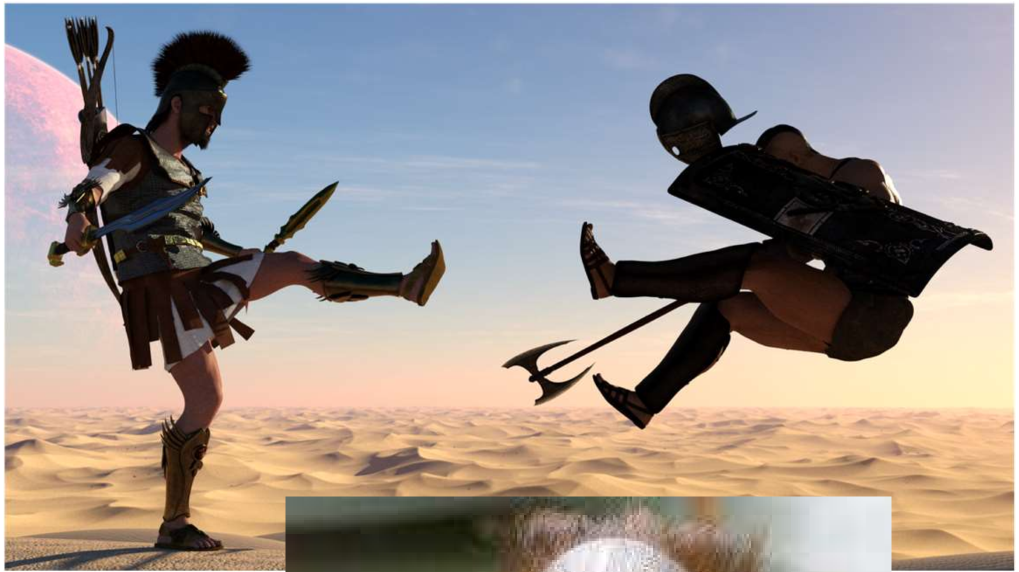
- ε-Photography
- Hyperspectral imaging
- Plenoptic cameras

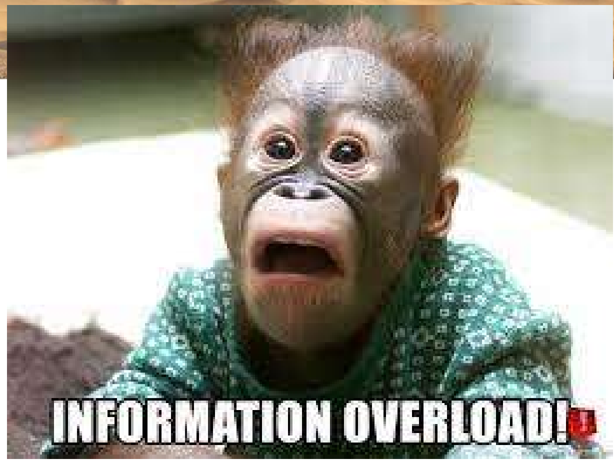
SCIENTIFIC-GRADE "FUTURE" CAMERAS:

- Single pixel camera (what??)
- Femto-photography
- Event-based vision

CASE STUDIES:

- Nerve Labeling for Da-Vinci robot
- Laser Speckle Imaging for Septic patients
- NIR Imaging for phlebotomy

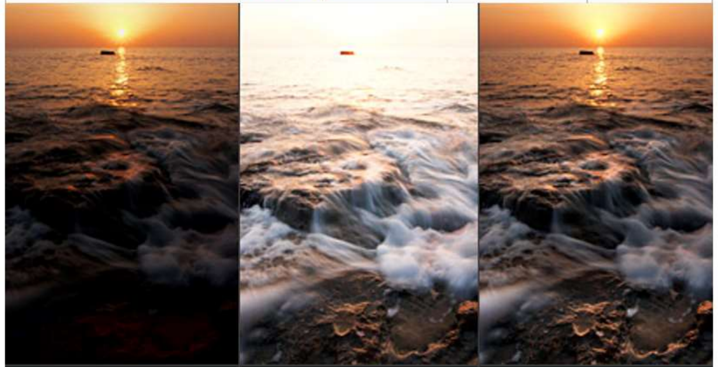




Human eye vs camera?

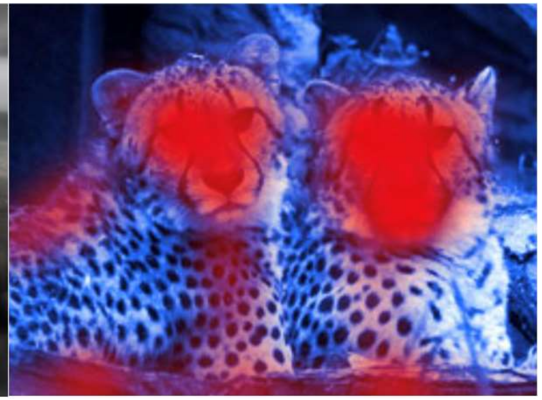
Dynamic ranges of common devices

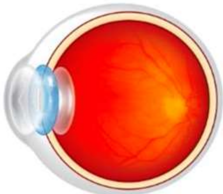
Device	Stops	Contrast Ratio
Single exposure		
Human eye: close objects	7.5	150...200
Human eye: 4° angular separation	13	8000...10000
Human eye (static)	10...14 [7]	1000...15000
Negative film (Kodak VISION3)	13 [8]	8000
best 1/1.7" camera (Nikon Coolpix P340)	11.9[citation needed]	3800
best 1" camera (Canon PowerShot G7 X)	12.7[citation needed]	6600
best Four-Thirds DSLR camera (Panasonic Lumix DC-GH5)	13.0[citation needed]	8200
best APS DSLR camera (Nikon D7200)	14.6 [9]	24800
best Full Frame DSLR camera (Nikon D810)	14.8 [9]	28500

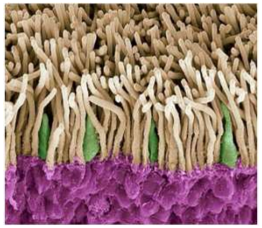


Eye Focuses on Background | Eye Focuses on Foreground | Our Mental Image

cambridgeincolour.com/tutorials/cameras-vs-human-eye.htm



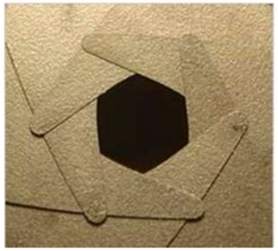


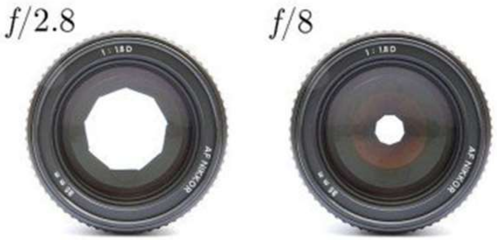


Human retina:

- ➢ 130 million photoreceptors
- ➢ But only 2 million axons!

Human eye vs camera?

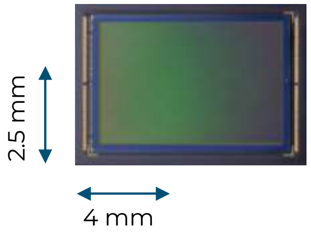




Aperture (size)

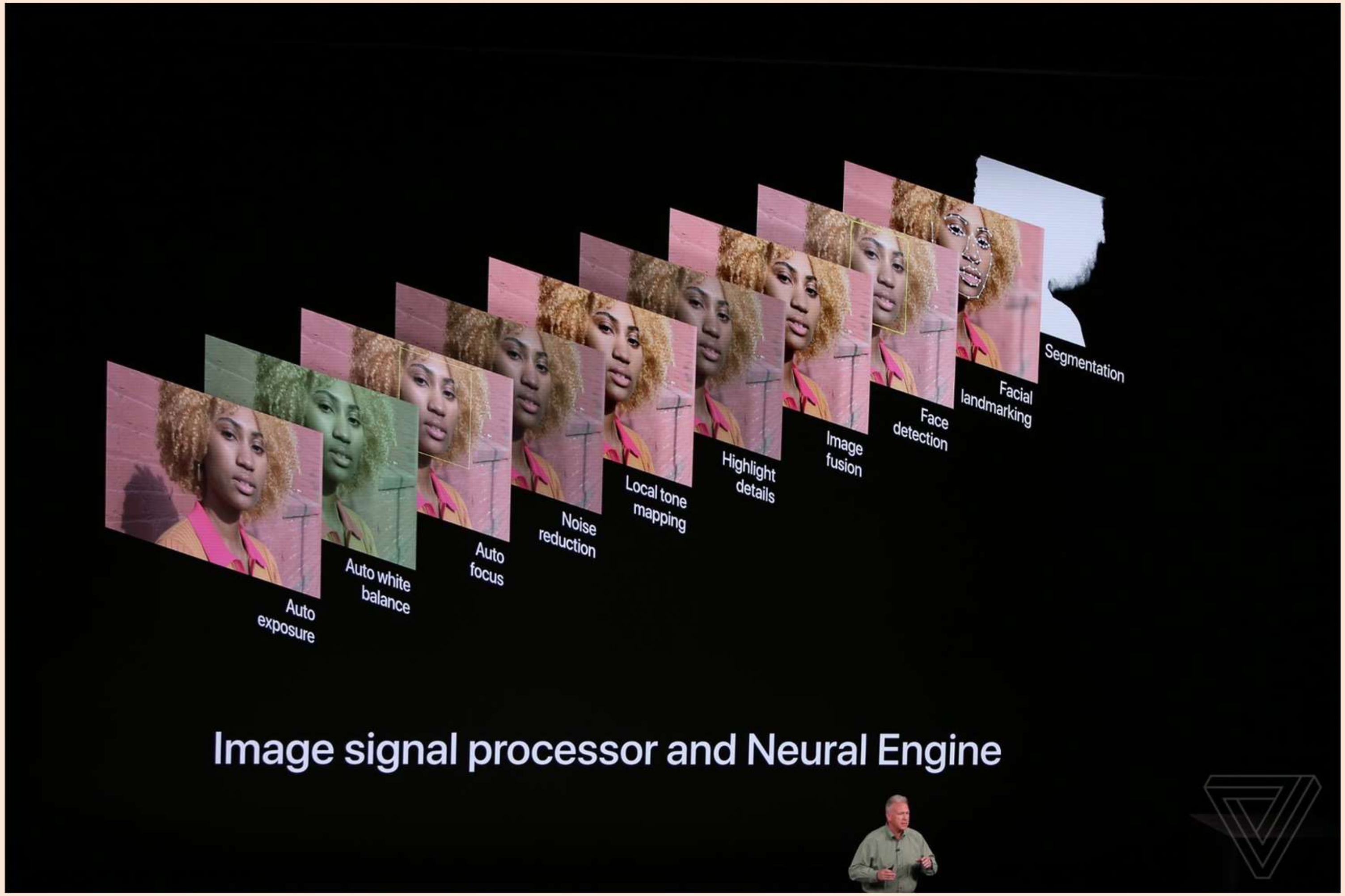
Exposure (time)

10 megapixel sensor

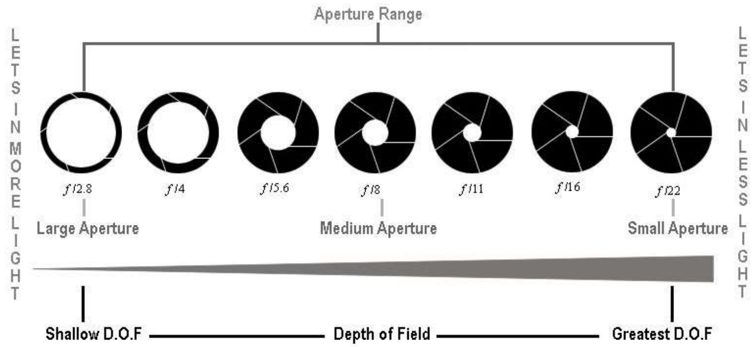


ISO (sensitivity)

ε-photography: process stream of photos on the fly

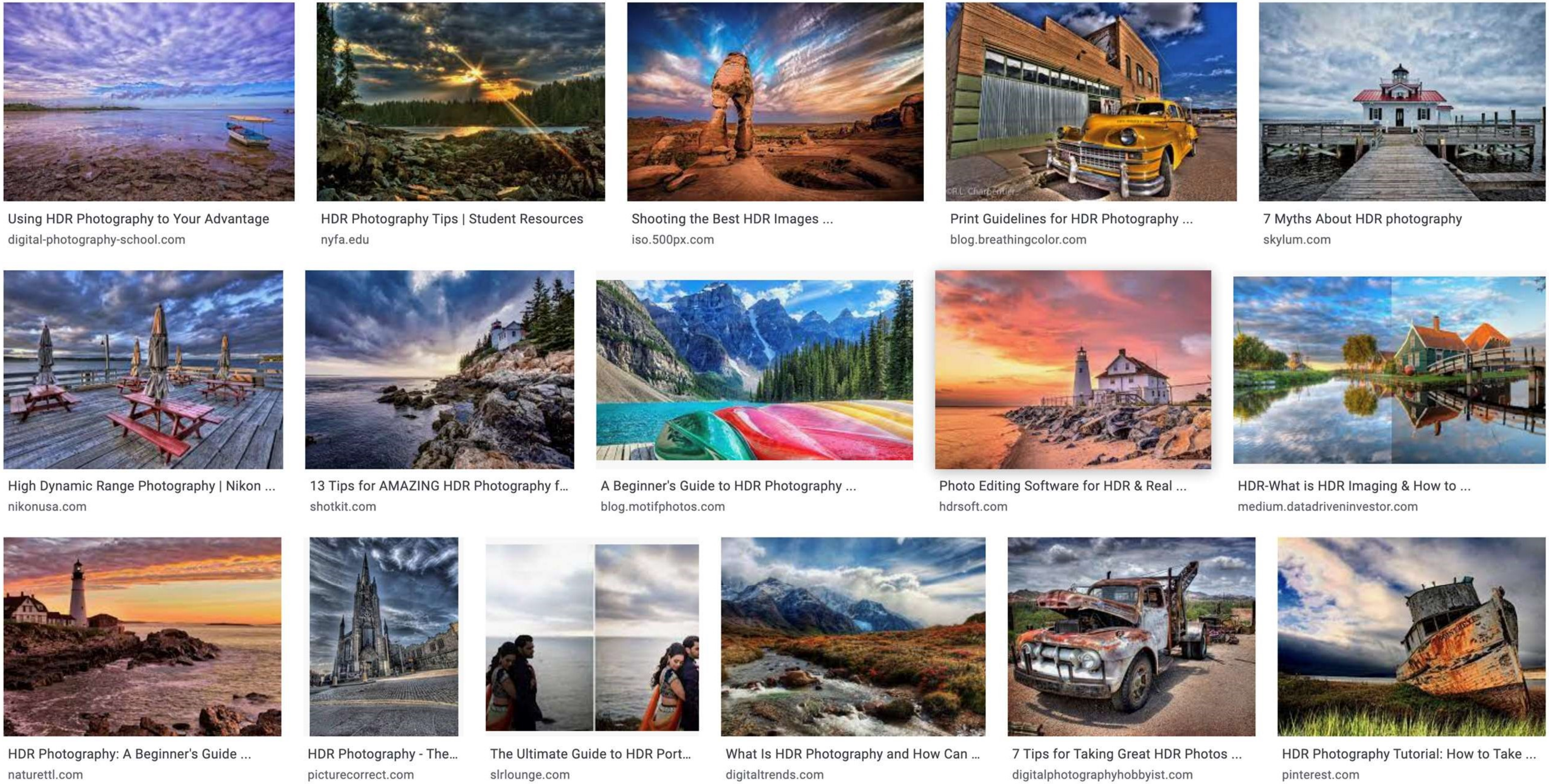


ε-photography: Z-stacking for DOF reconstruction



ε-photography: panorama stitching

cv2.createStitcher

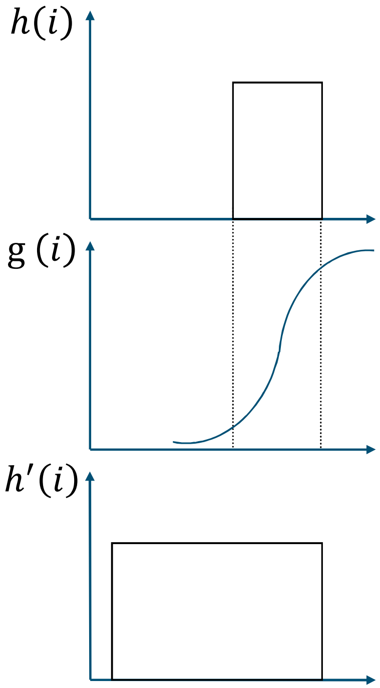

Using HDR Photography to Your Advantage
digital-photography-school.com
HDR Photography Tips | Student Resources
nyfa.edu
Shooting the Best HDR Images ...
iso.500px.com
Print Guidelines for HDR Photography ...
blog.breathingcolor.com
7 Myths About HDR photography
skylum.com
High Dynamic Range Photography | Nikon ...
nikonusa.com
13 Tips for AMAZING HDR Photography f...
shotkit.com
A Beginner's Guide to HDR Photography ...
blog.motifphotos.com
Photo Editing Software for HDR & Real ...
hdrsoft.com
HDR-What is HDR Imaging & How to ...
medium.datadriveninvestor.com
HDR Photography: A Beginner's Guide ...
naturettl.com
HDR Photography - The...
picturecorrect.com
The Ultimate Guide to HDR Port...
slrlounge.com
What Is HDR Photography and How Can ...
digitaltrends.com
7 Tips for Taking Great HDR Photos ...
digitalphotographyhobbyist.com
HDR Photography Tutorial: How to Take ...
pinterest.com

Histogram equalisation

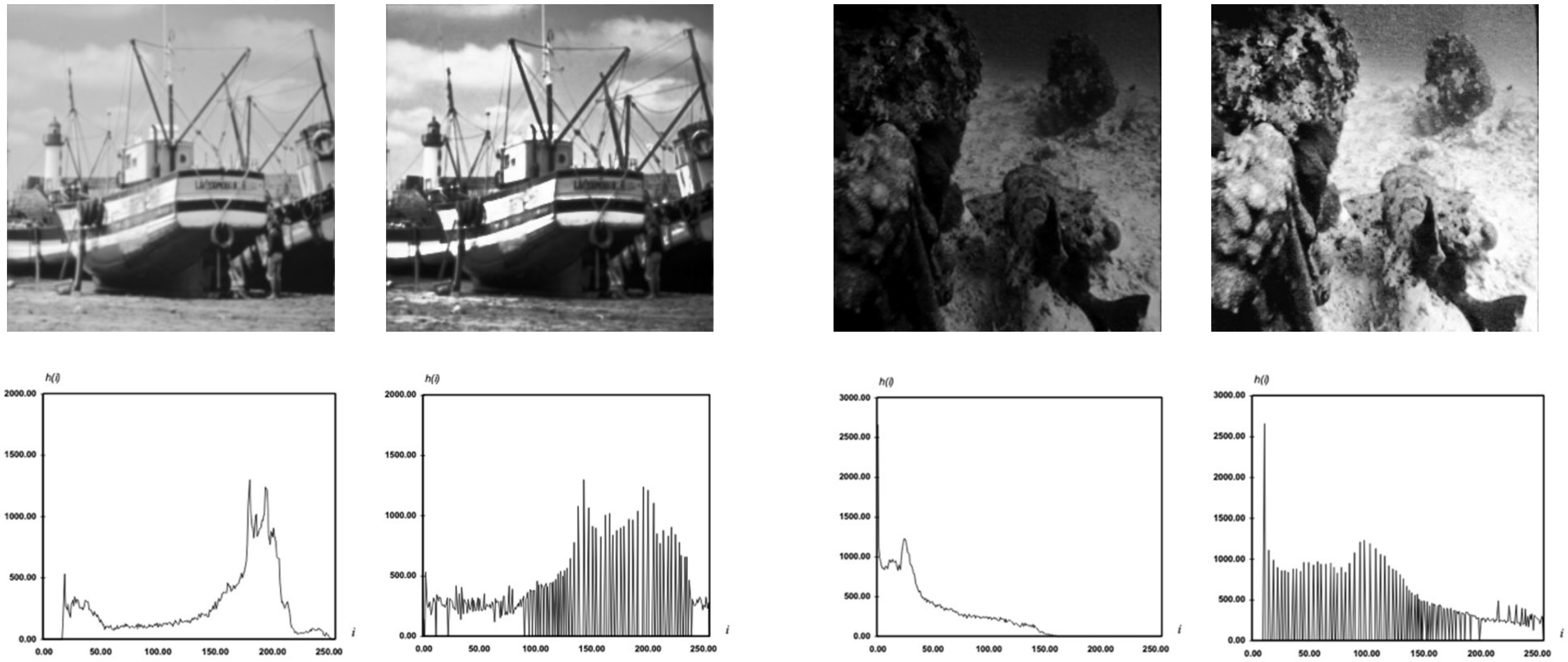
- In an image of low contrast, the image has grey levels concentrated in a narrow band
- Define the grey level histogram of an image $h(i)$ where :
- $h(i)$ = number of pixels with grey level = i
- For a low contrast image, the histogram will be concentrated in a narrow band
- The full greylevel dynamic range is not used

$$h'(i) = h\left(g^{-1}(i)\right)$$

$$g(i) = \frac{1}{1 + exp\left(\lambda(0 - i)\right)}$$

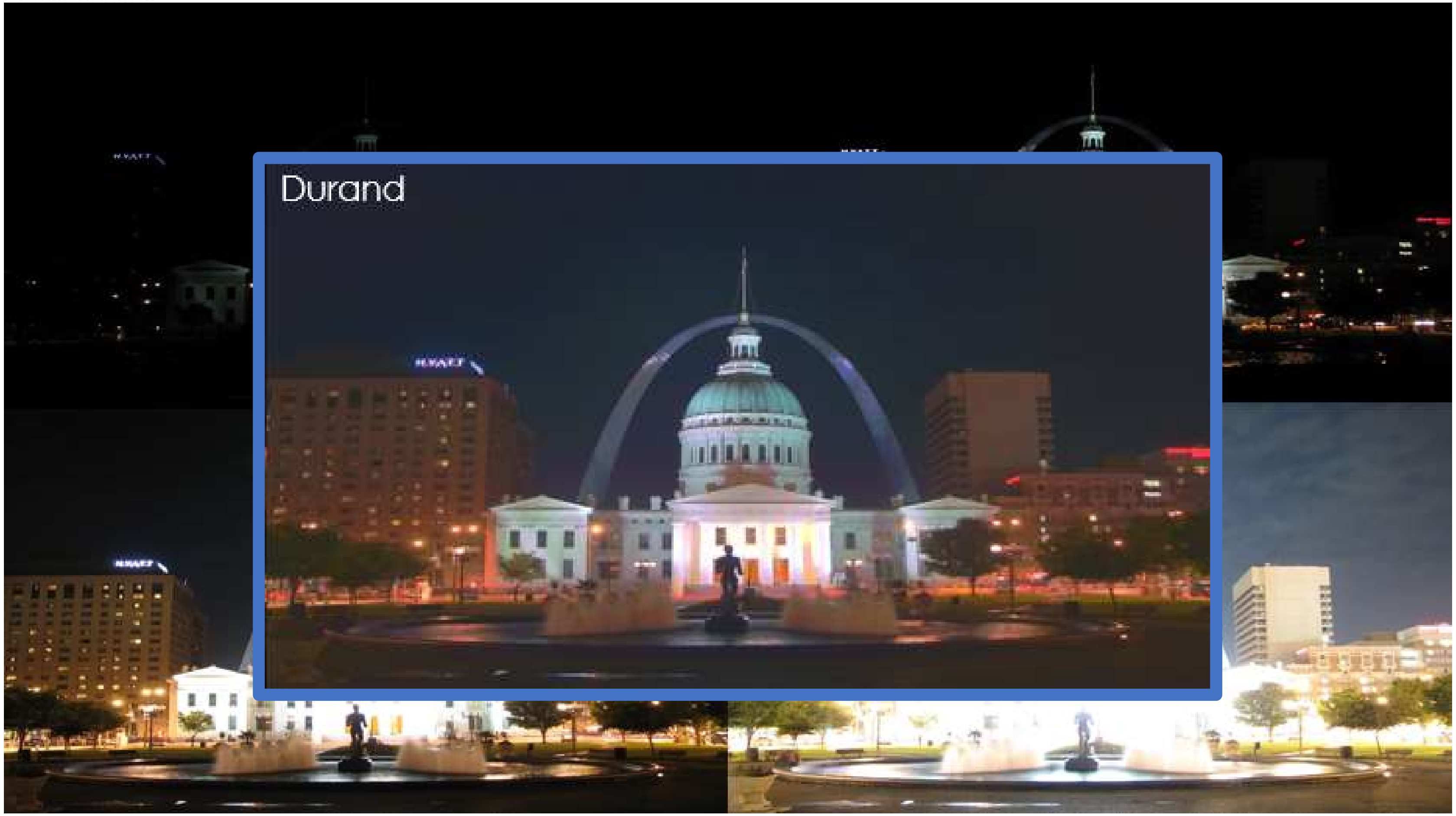


Histogram equalisation

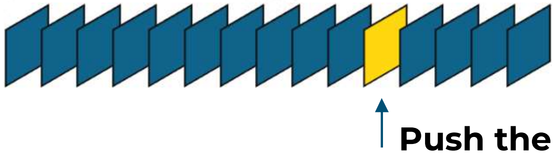


HDR eliminates these artifacts

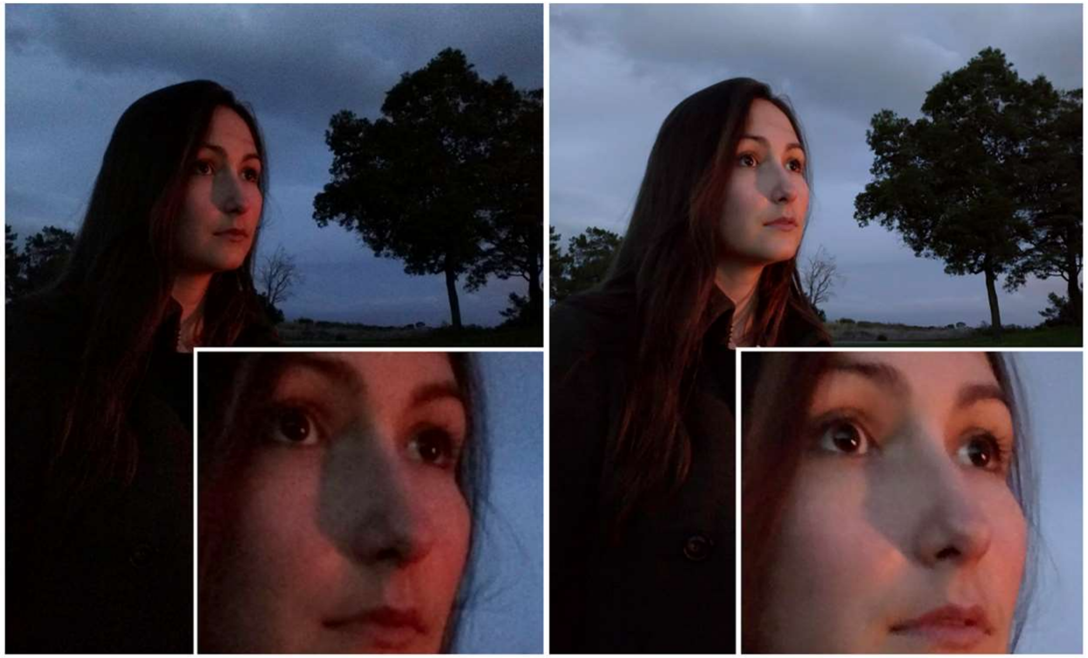
Take pictures at different exposure times



HDR+



Push the button



Poorly lit photos:

- Long exposure time = blur
- Large ISO = grain noise
- Optical image stabilization = helpless if really dark

HDR+:

- take a burst of shots with short exposure times
- align them algorithmically
- Averaging multiple shots reduces noise, using short exposures reduces blur.

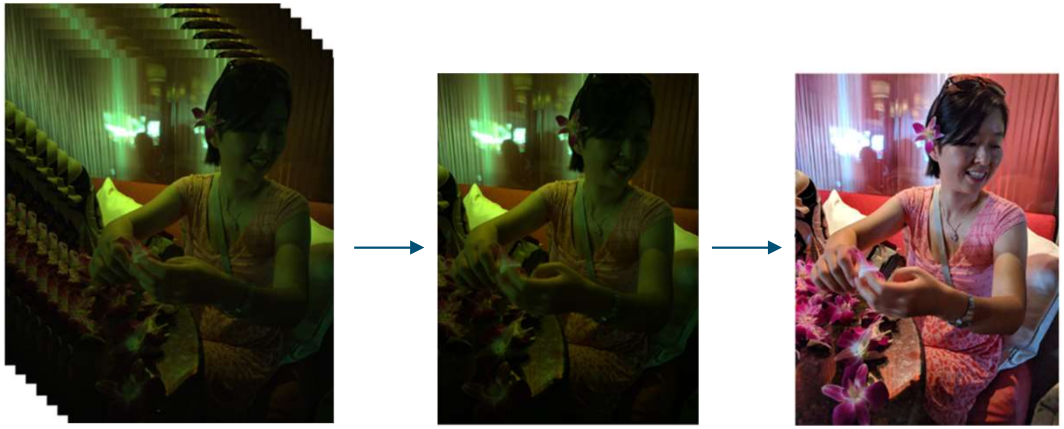
HDR+



- No ghost artefacts
- Natural color due to averaging

Introducing the HDR+ Burst Photography Dataset

Monday, February 12, 2018 Posted by Same Hasinoff, Software Engineer, Machine Perception

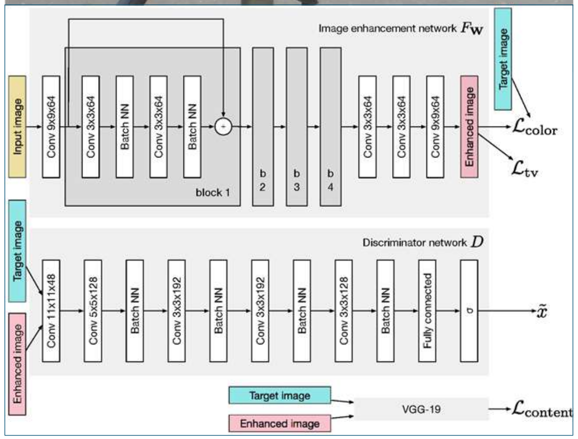


Burst of raw frames **Merged raw images** **Final high-quality result**

Yes, this is the time for ANNs!

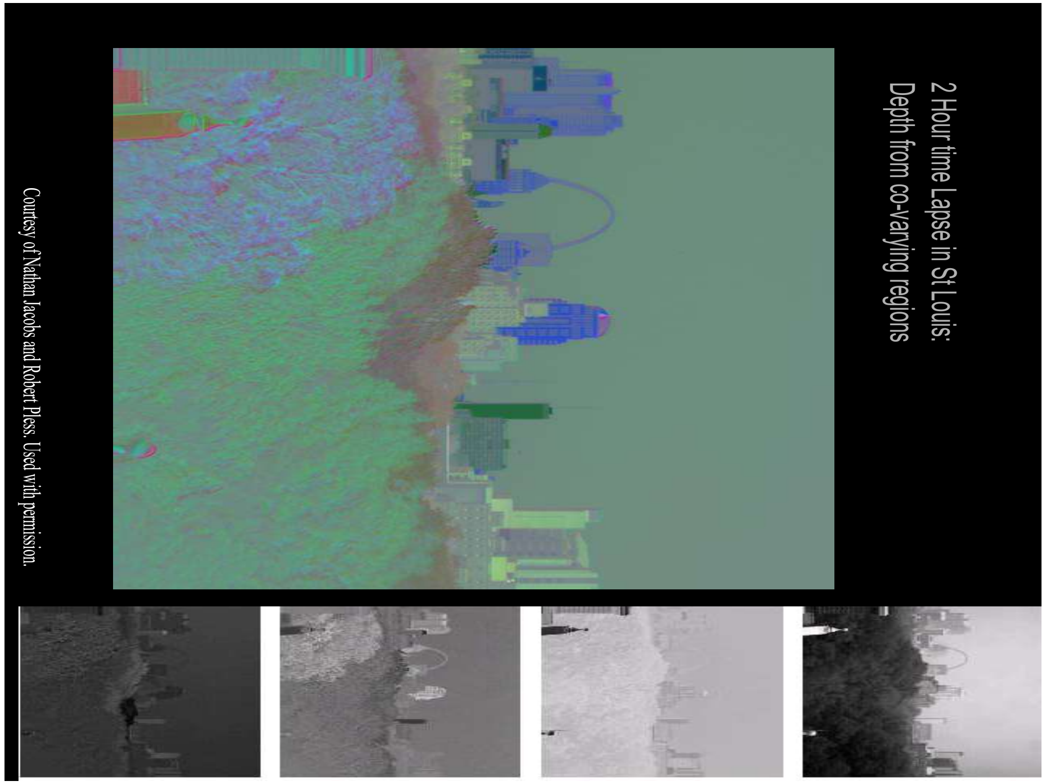
Tone-mapping is a very hard nonlinear problem, prone to errors in the high-scatter areas

DSLR-Quality Photos on Mobile Devices with Deep Convolutional Networks



ε-photography: Intensity Flicker

Wait long enough,
collect many images,
and reconstruct Depth



ε-photography: short-exposure image as prior

Long exposure (blurry) | **Short exposure (dark)** | **Same, scaled up (noisy)** | **Joint deconvolution**

Hyperspectral Cameras

Human Vision is Trichromatic

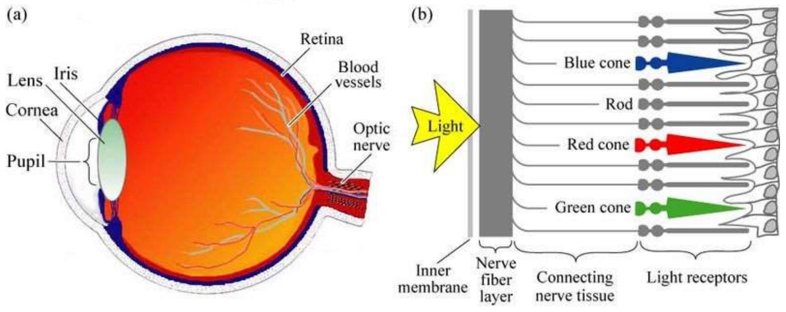
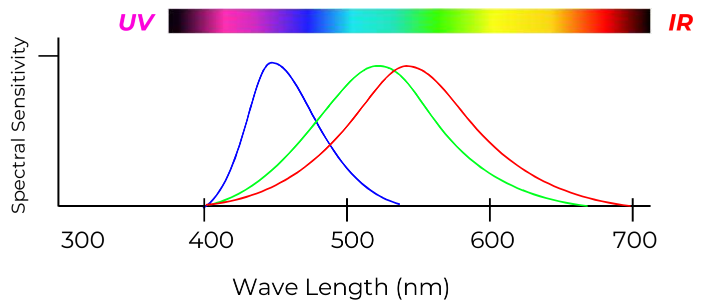


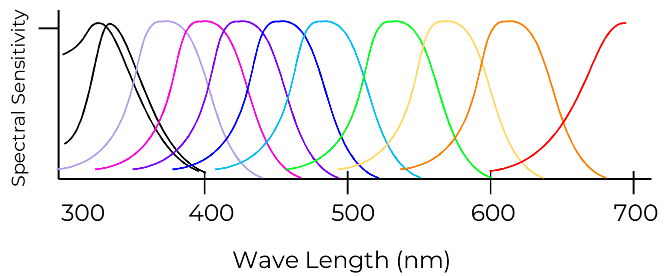
Fig. 16.1. (a) Cross section through a human eye. (b) Schematic view of the retina including rod and cone light receptors (adapted from Encyclopedia Britannica, 1994).

E. F. Schubert
Light-Emitting Diodes (Cambridge Univ. Press)
www.LightEmittingDiodes.org

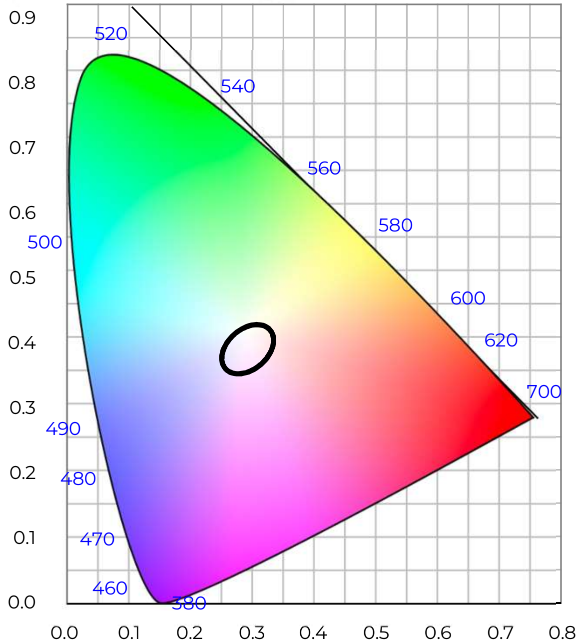


Shrimp is a better color sensor!

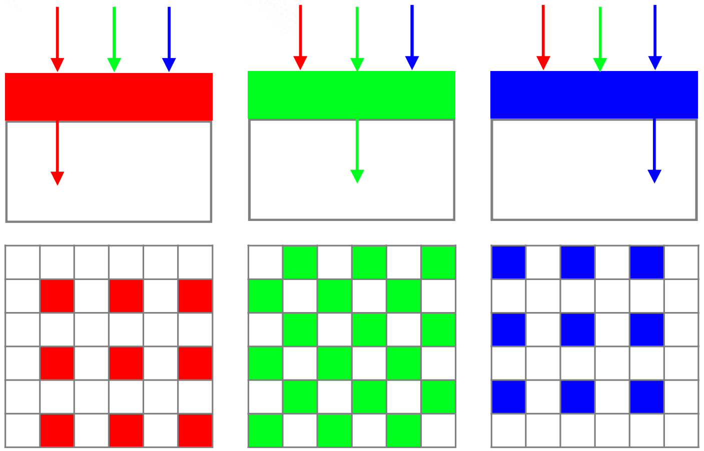
Neogonodactylus oestedii



CIE Color classification



Color sensor

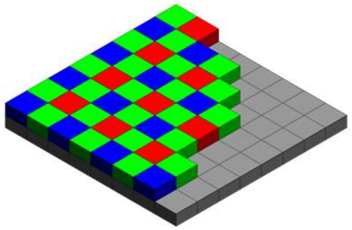


Incoming light

Filter layer

Sensor array

Resulting pattern



Bayer Filter

Demosaicing

Incoming light

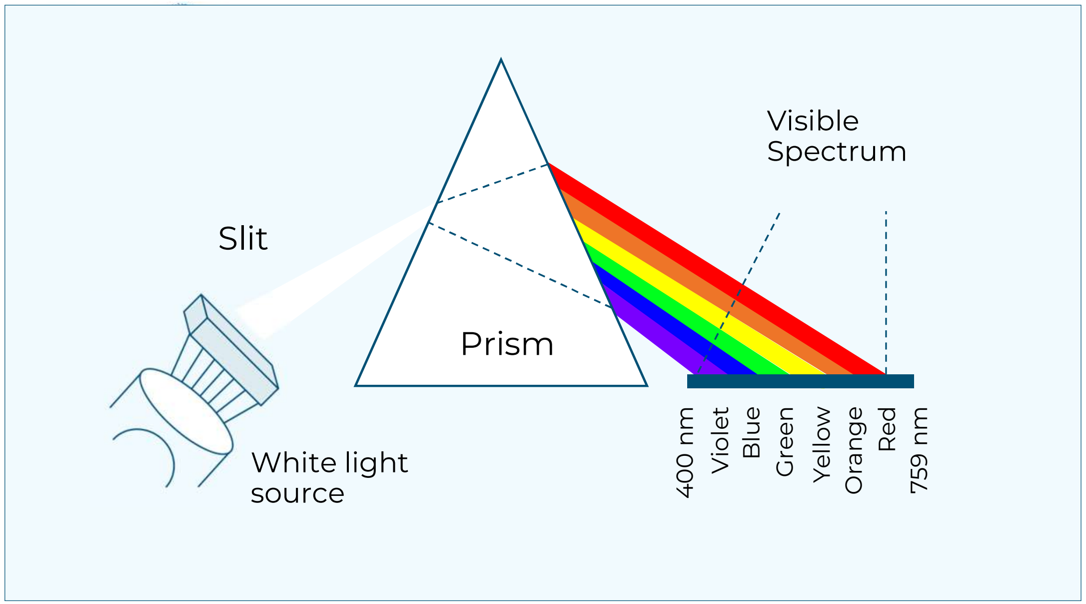
Filter layer

Sensor array

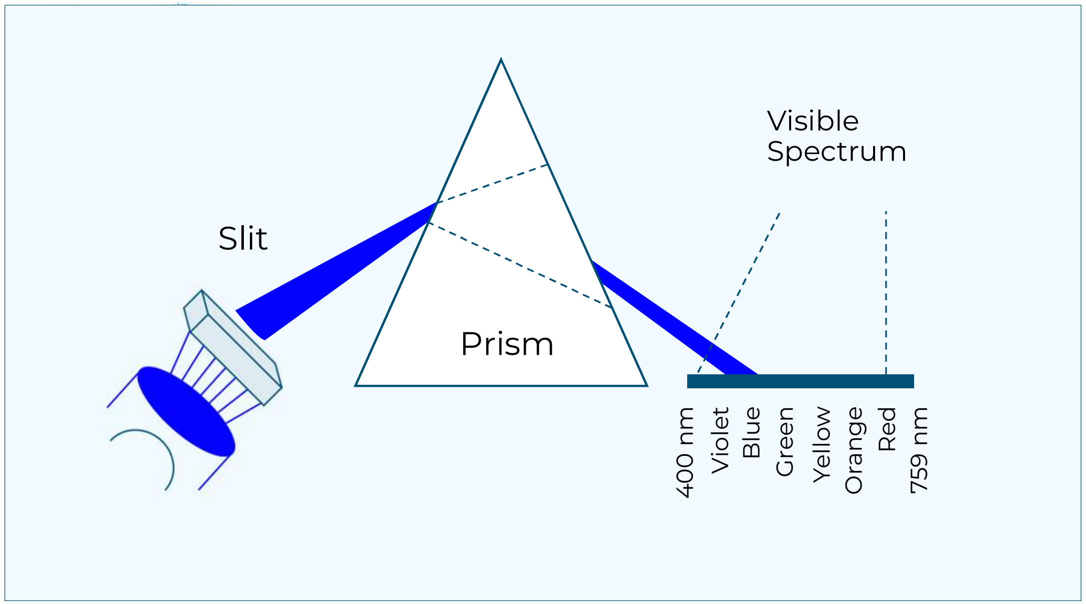
Resulting pattern

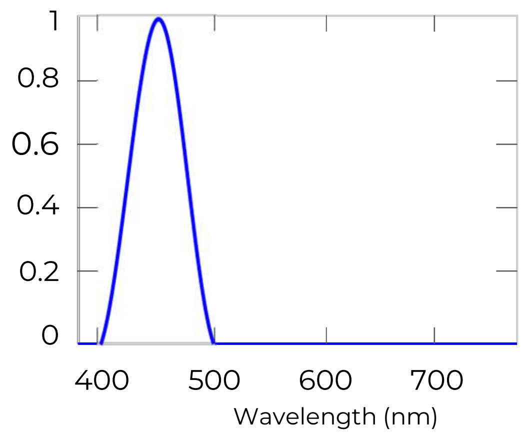
Interpolate missing colors!

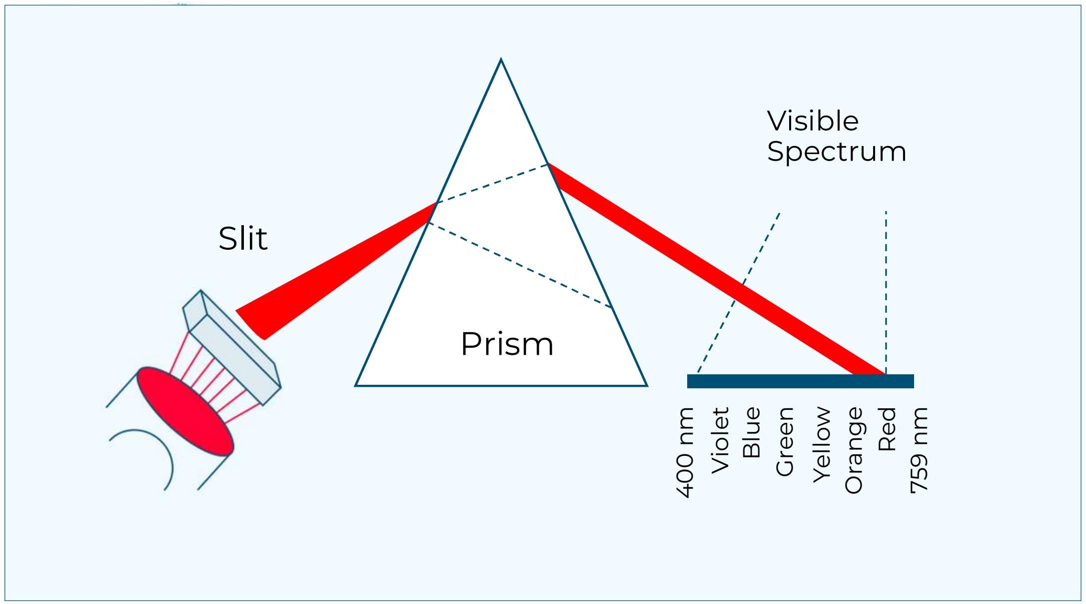
Noise? -- Average

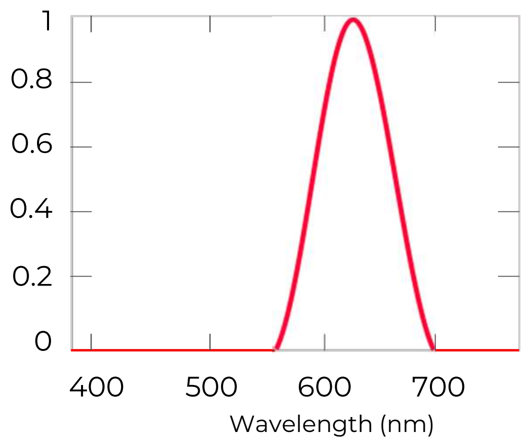


Prism Splitting White Light into Different Colors

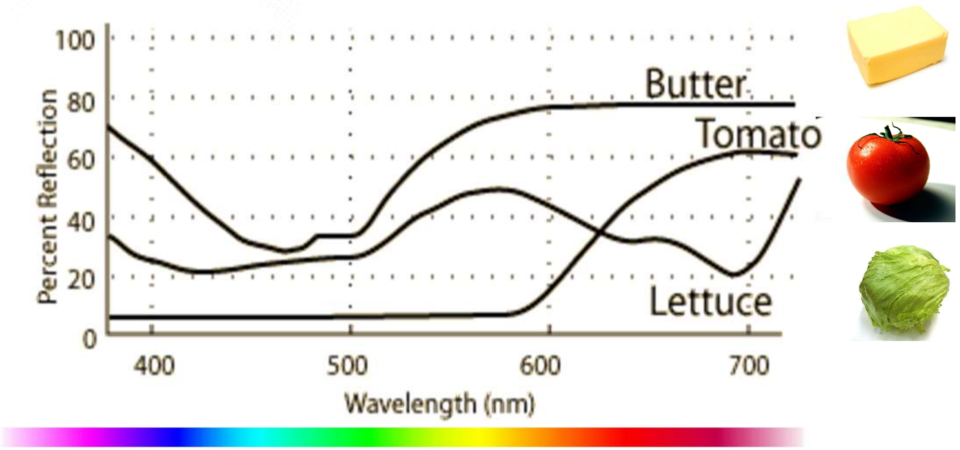

Visible
Spectrum
Slit
Prism
400 nm
Violet
Blue
Green
Yellow
Orange
Red
759 nm


1
0.8
0.6
0.4
0.2
0
400
500
600
700
Wavelength (nm)

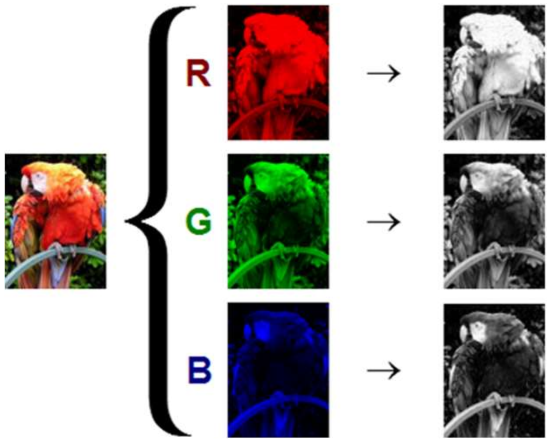

Visible
Spectrum
Slit
Prism
400 nm
Violet
Blue
Green
Yellow
Orange
Red
759 nm


1
0.8
0.6
0.4
0.2
0
400
500
600
700
Wavelength (nm)

Color Images

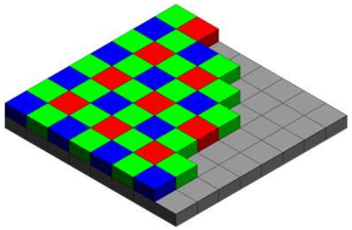


Need spectrum over wavelengths to described color



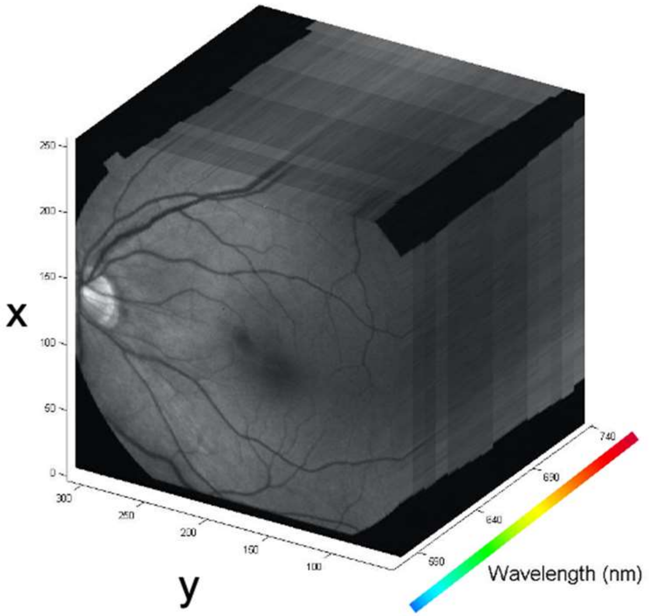
Typical RGB Color Image (Only 3 Bands)

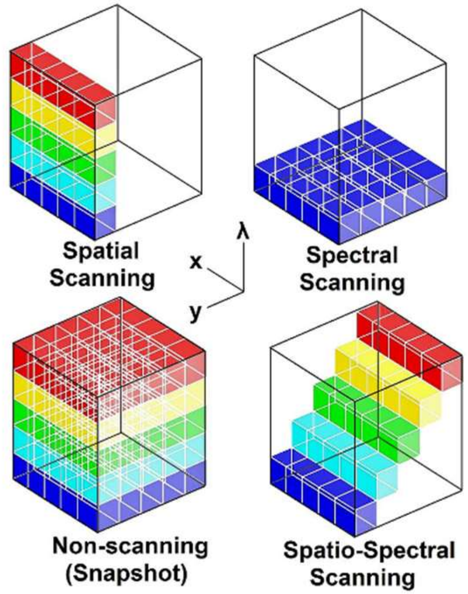
Why do we only have three colors in images?



Bayer Filter

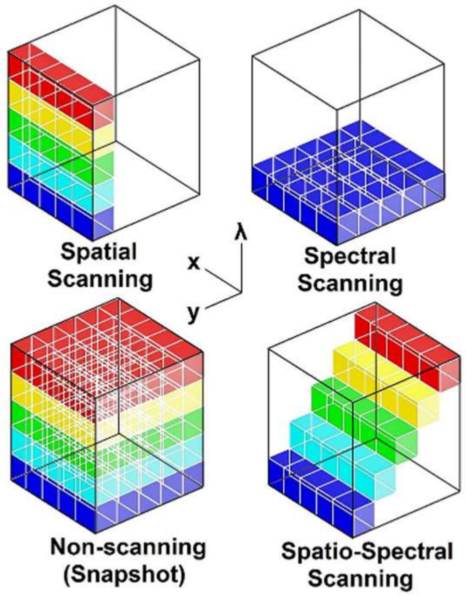
How to Capture?





How to Capture?

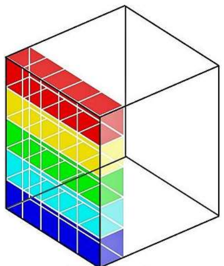
Standard Methods (Slow and noisy.)



Fastest but expensive (May have lower resolution.)

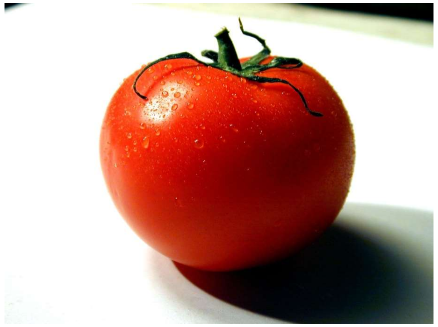
Prototype introduced in 2014. (Needs moving of imaging equipment to cover the entire cube.)

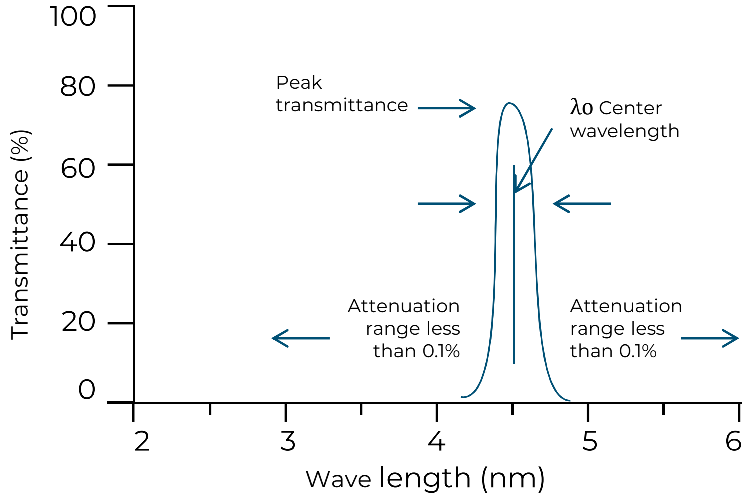
Before: Bulky Hyperspectral Cameras $50k



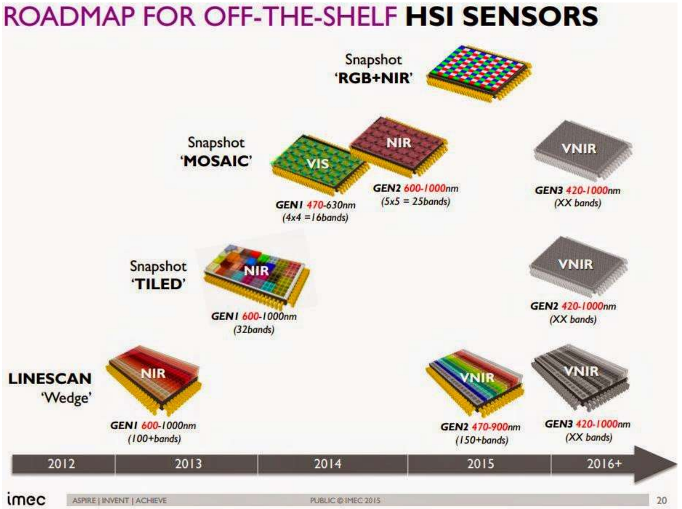
Spatial Scanning

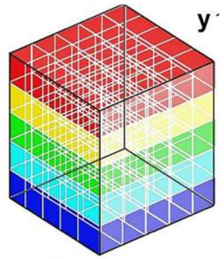
Narrowband Filter to Only Let Certain Wavelengths In



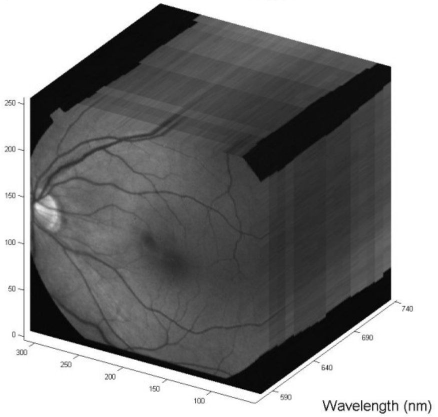


Now: Snapshot Hyperspectral Imaging

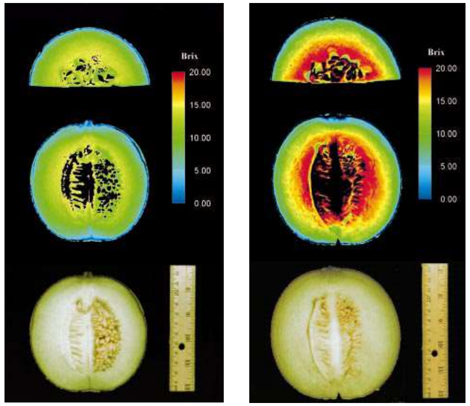




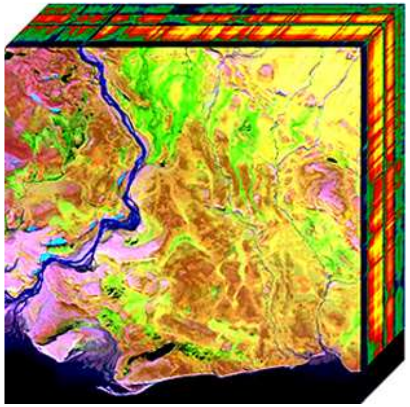
Hyperspectral Imaging Applications



Medical Diagnostics



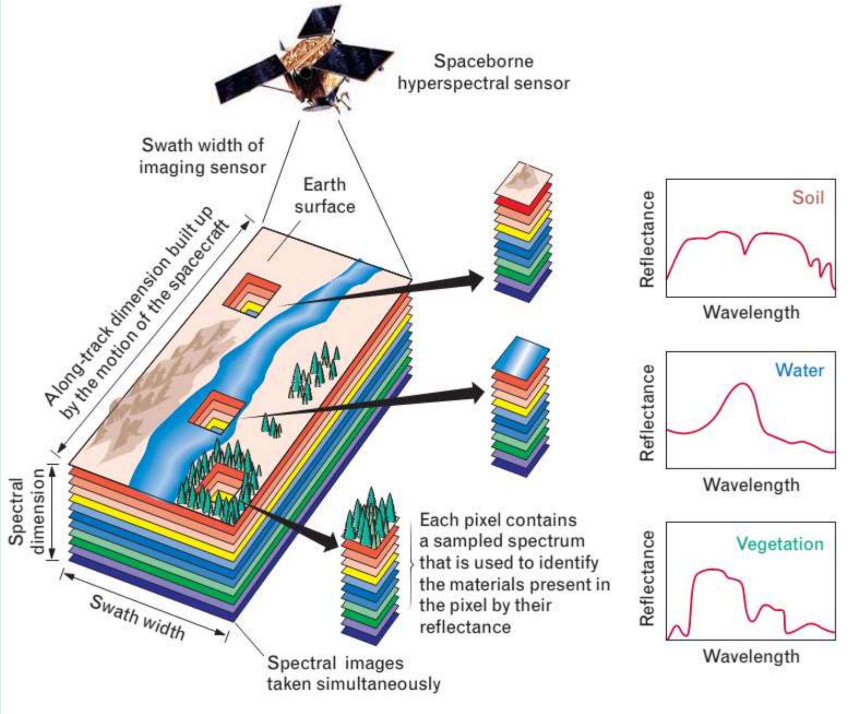
Food Quality Inspection



Satellite Imaging

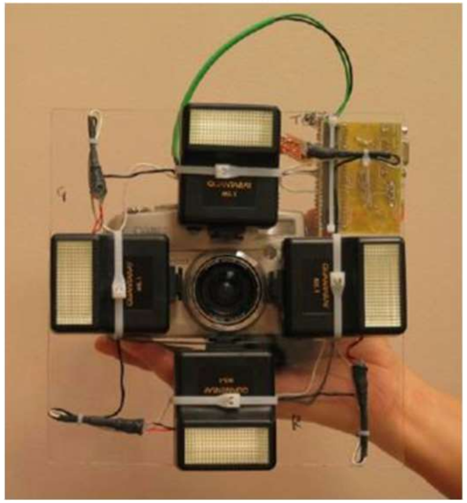
...and more!

Application: Remote Sensing



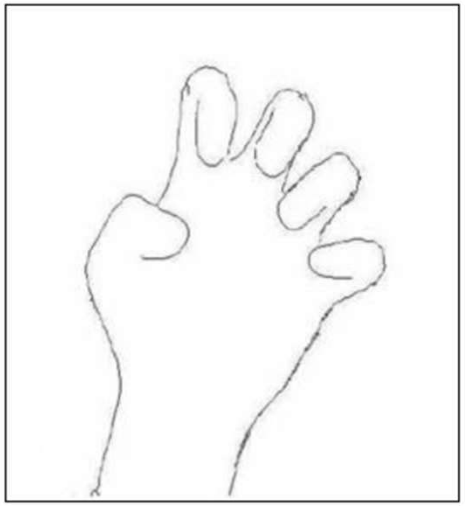
Shaw and Burke (Lincoln Lab Journal 2003)

Example of computational illumination: Illuminate 4 times, then reconstruct



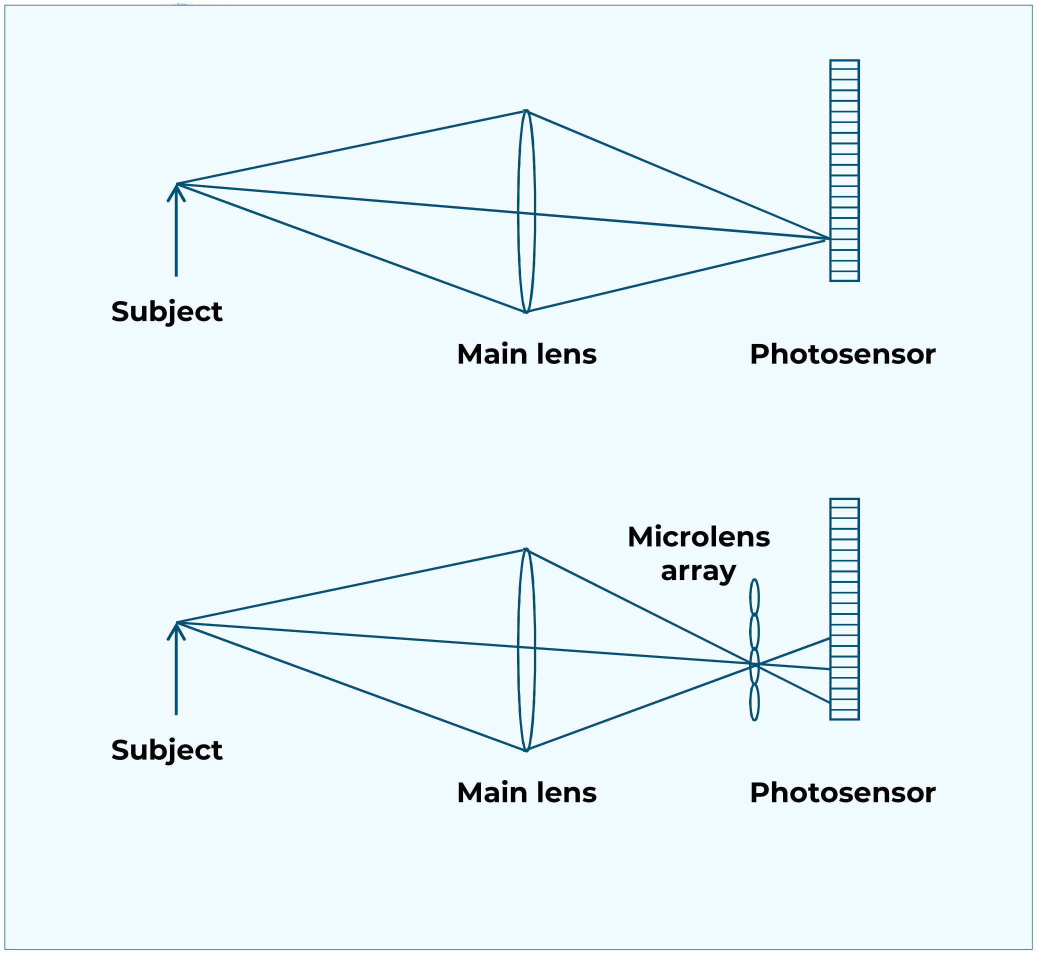
Input Photo

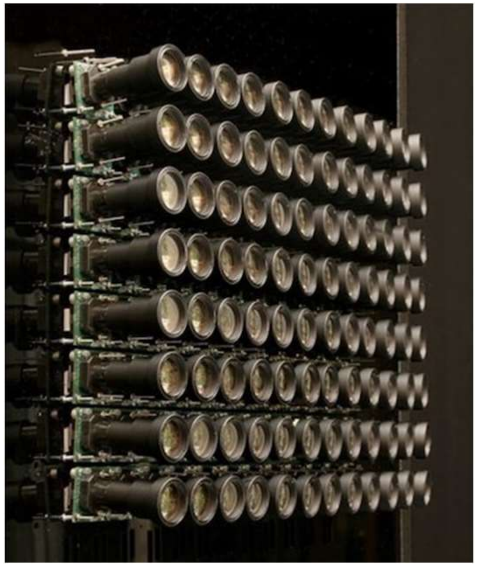
Canny Edges



Multi-flash camera image

Plenoptic cameras

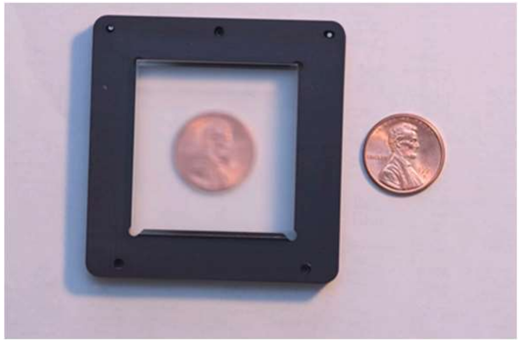

Subject
Main lens
Photosensor
Microlens
array
Subject
Main lens
Photosensor



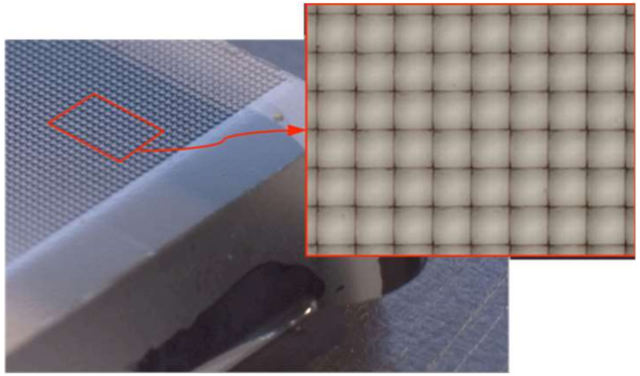
Plenoptic Camera

Contax medium format camera

Kodak 16-megapixel sensor



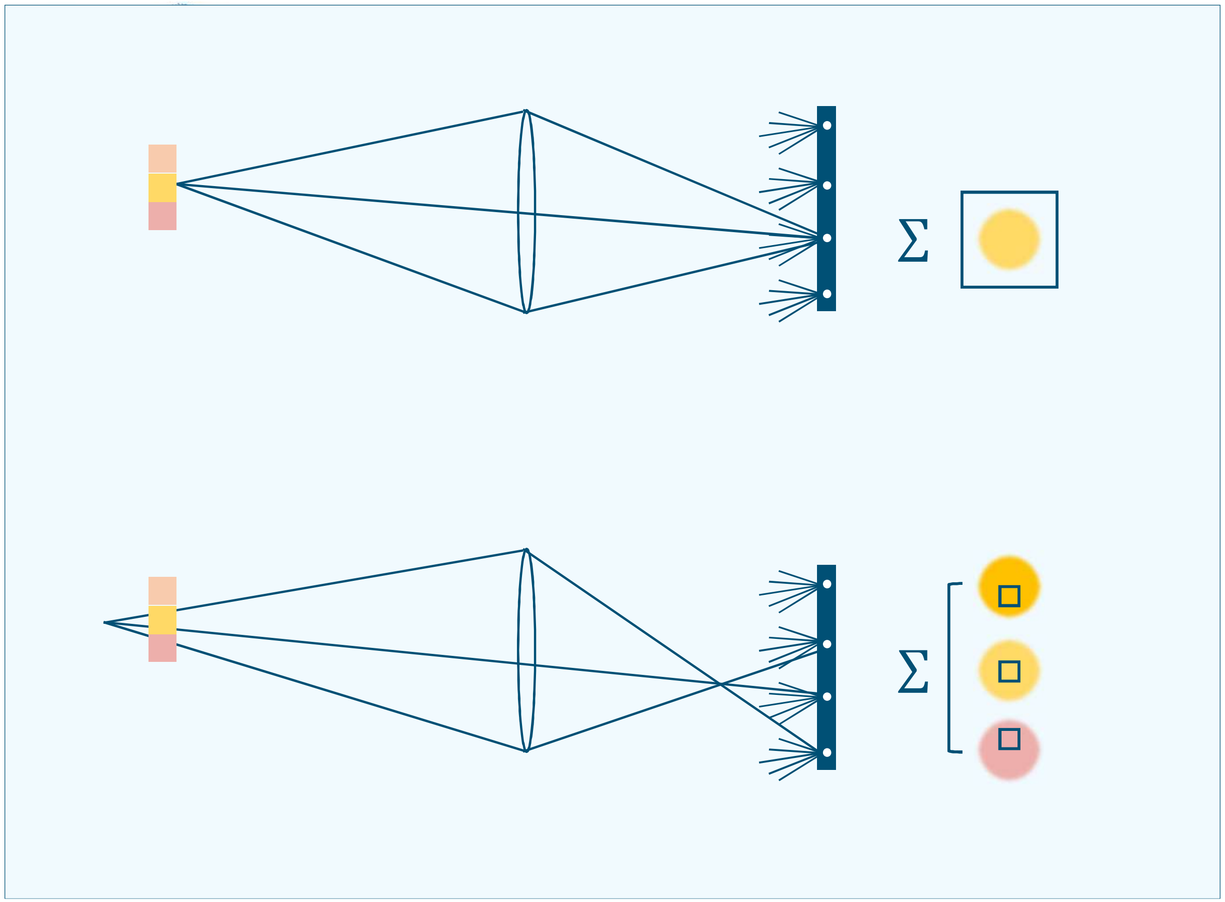
Adaptive Optics micro lens array



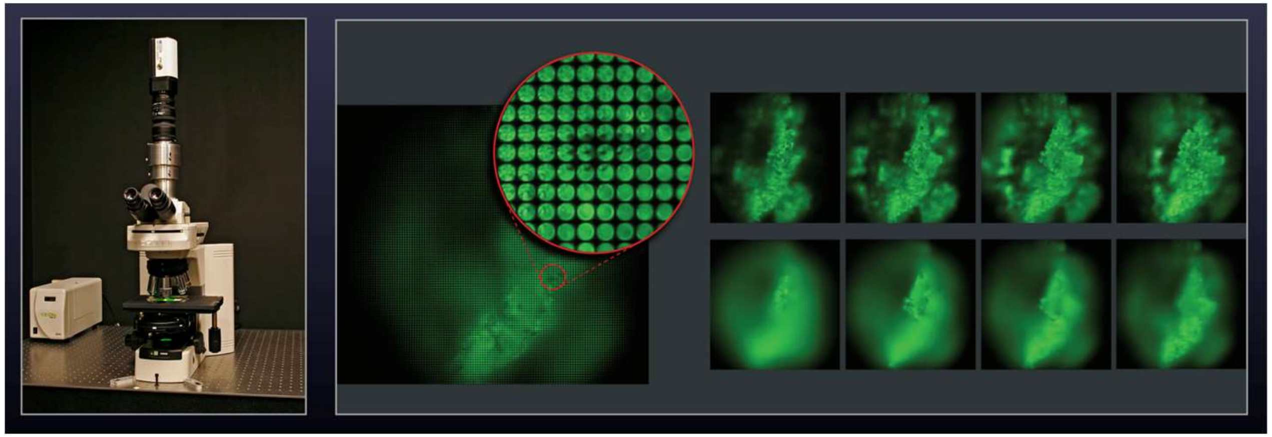
125μ square-sided micro lenses

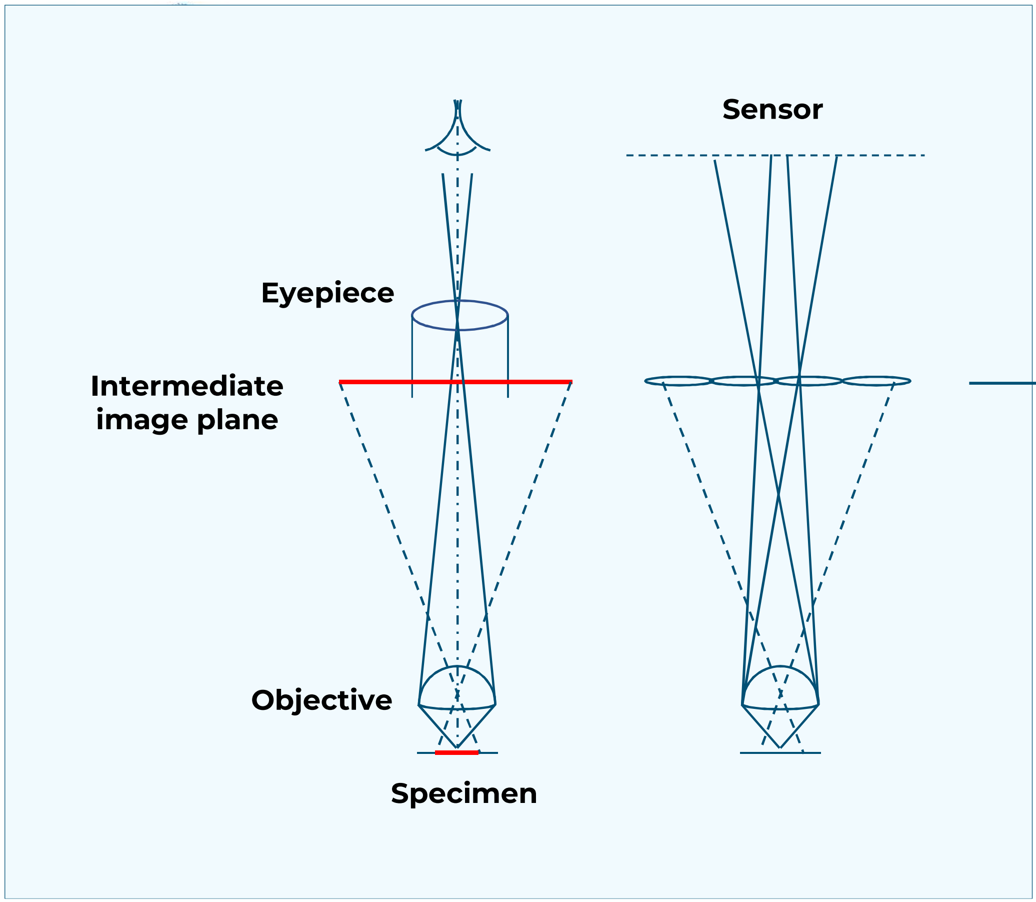
4000×4000 pixels ÷ 292×292 lenses = 14×14 pixels per lens

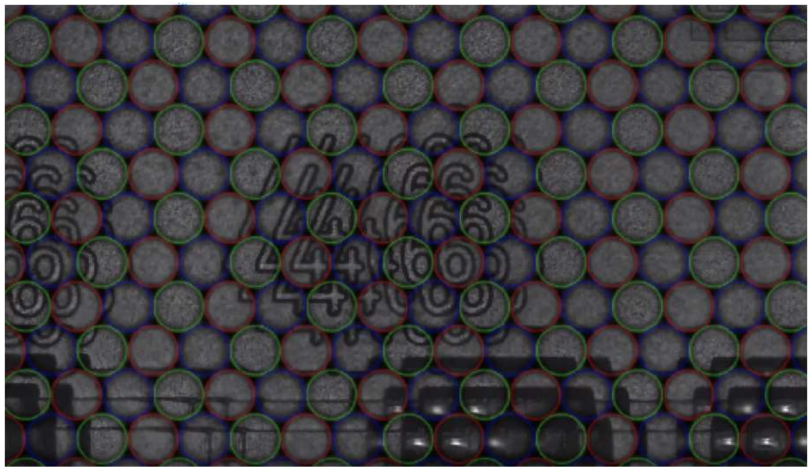
Typical image captured by camera (show here at low res)


Σ
Σ

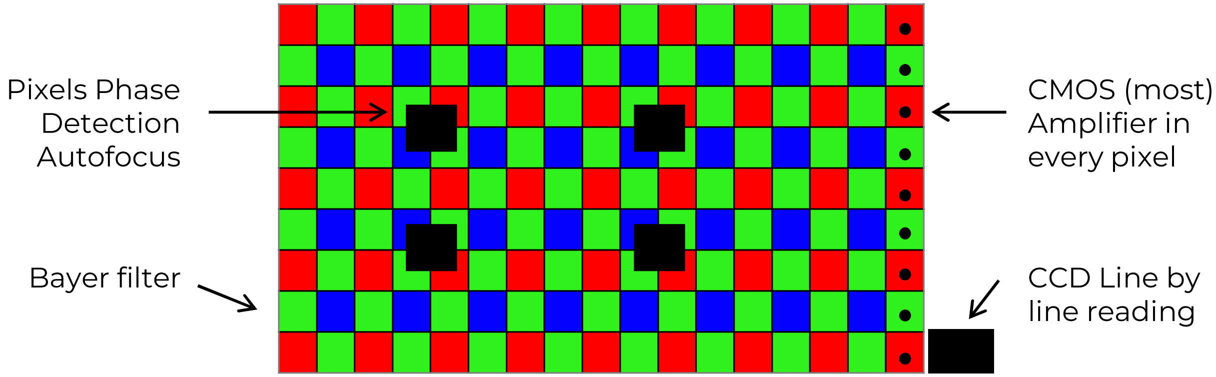



Sensor
Eyepiece
Intermediate image plane
Objective
Specimen

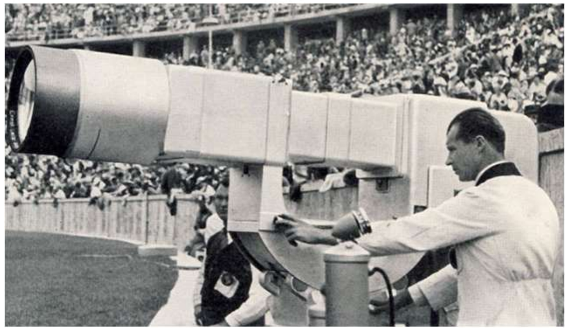


https://vas3k.ru/blog/computational_photography/

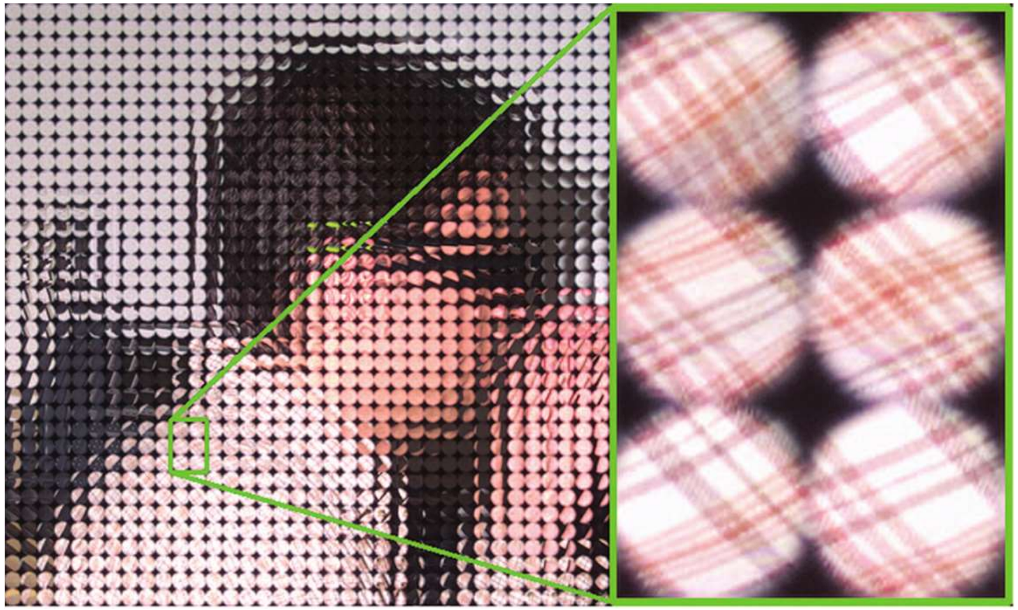


SONY

Plenoptic video cameras: back to historic standpoint



Resolution: 750+Mpx

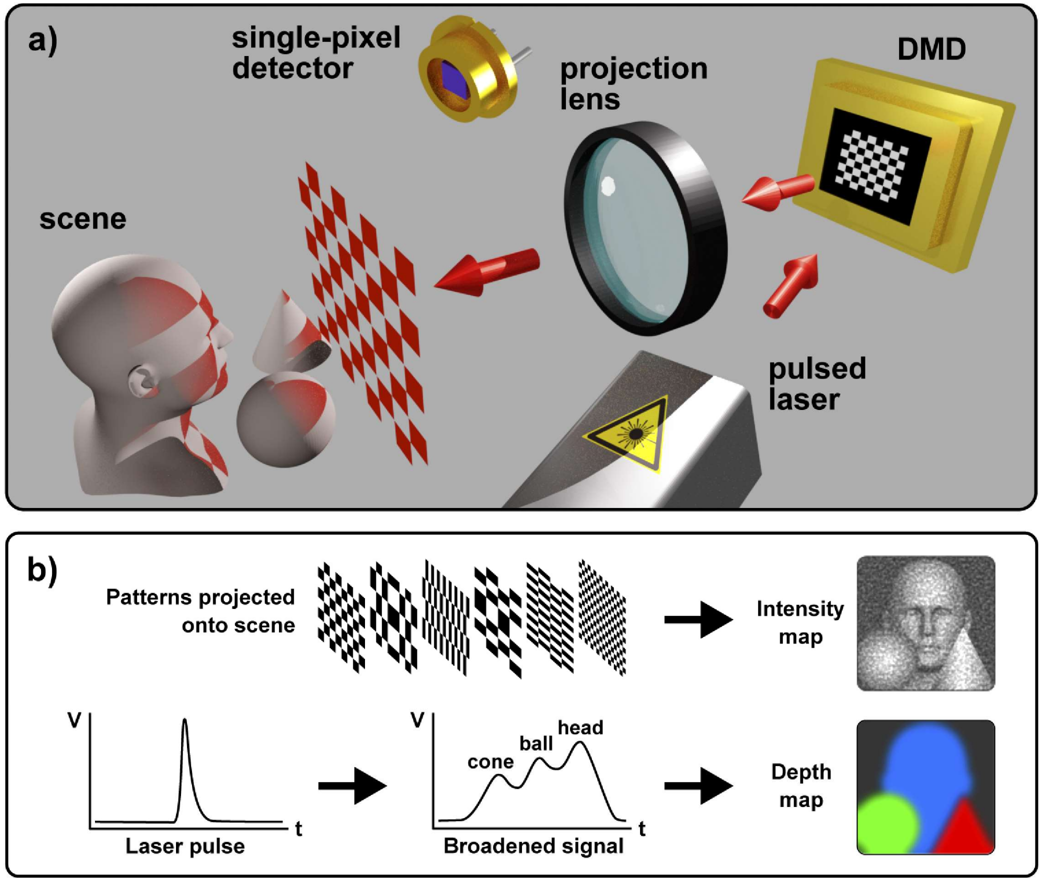


- Survey (in Rus): https://m.habr.com/ru/post/440652/
- Video of image reconstruction (Lytra): https://www.youtube.com/watch?time_continue=987&v=4qXE4sA-hLQ

Single-pixel cameras

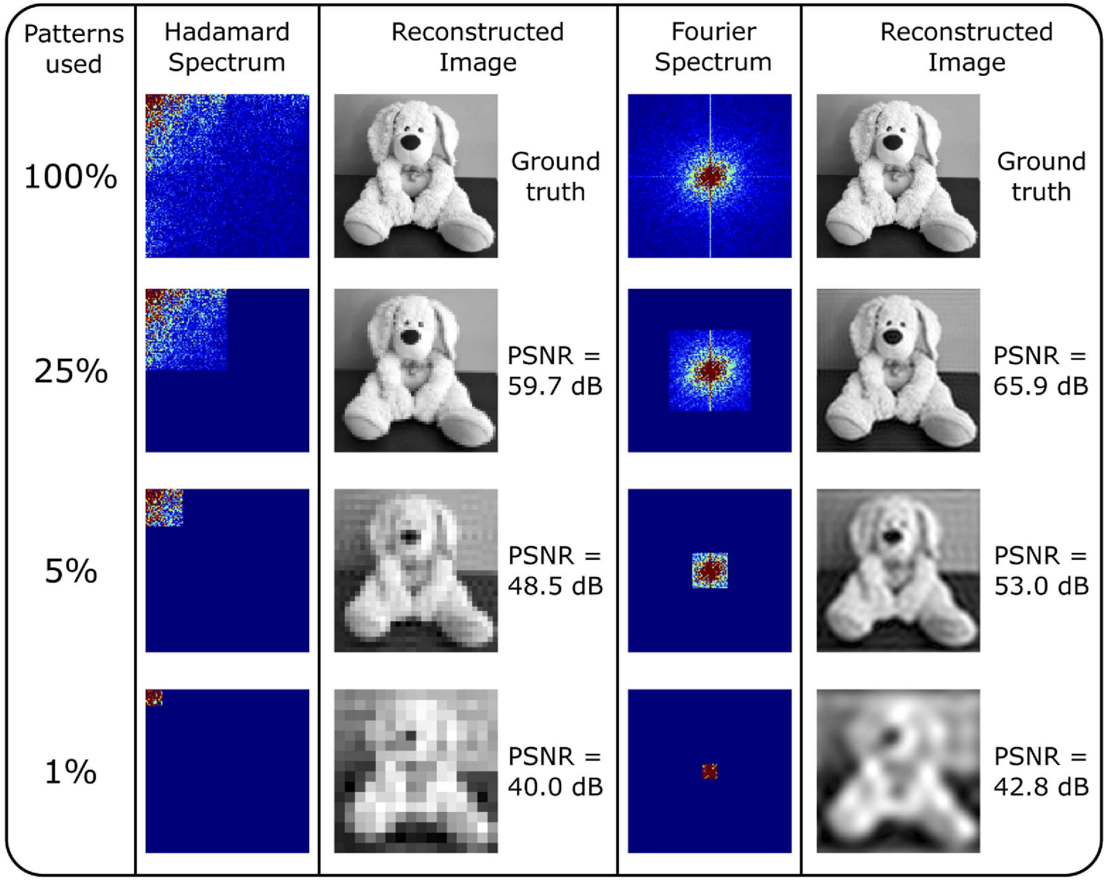
Any ideas how?

Single pixel imaging



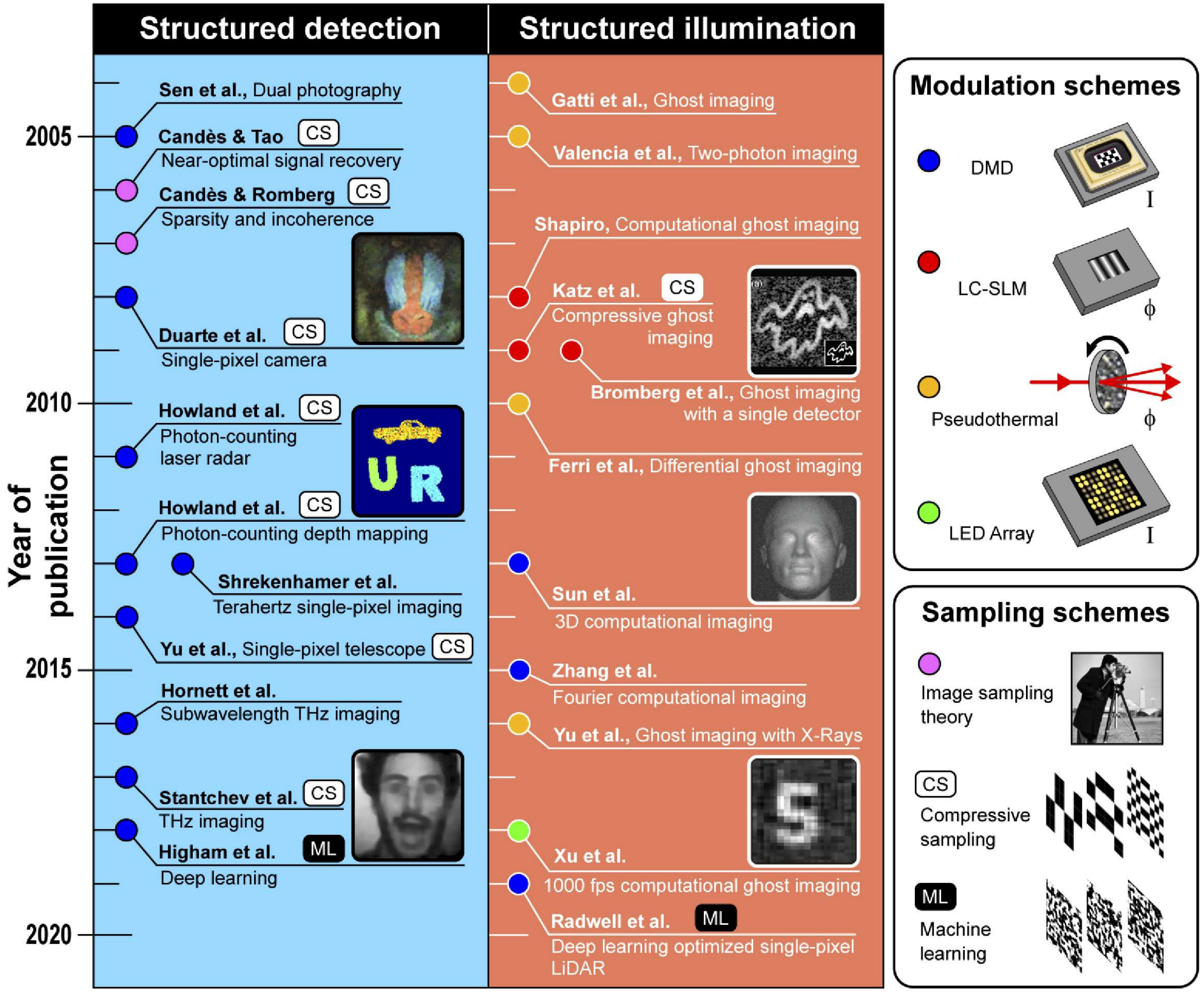
Graham M. Gibson, Steven D. Johnson, Miles J. Padgett, "Single-pixel imaging 12 years on: a review," Opt. Express **28,** 28190-28208 (2020); https://www.osapublishing.org/oe/abstract.cfm?uri=oe-28-19-28190

Single pixel imaging



Graham M. Gibson, Steven D. Johnson, Miles J. Padgett, "Single-pixel imaging 12 years on: a review," Opt. Express **28,** 28190-28208 (2020);
https://www.osapublishing.org/oe/abstract.cfm?uri=oe-28-19-28190

Single pixel imaging



Graham M. Gibson, Steven D. Johnson, Miles J. Padgett, "Single-pixel imaging 12 years on: a review," Opt. Express **28,** 28190-28208 (2020); https://www.osapublishing.org/oe/abstract.cfm?uri=oe-28-19-28190

Computational Photography Gurus

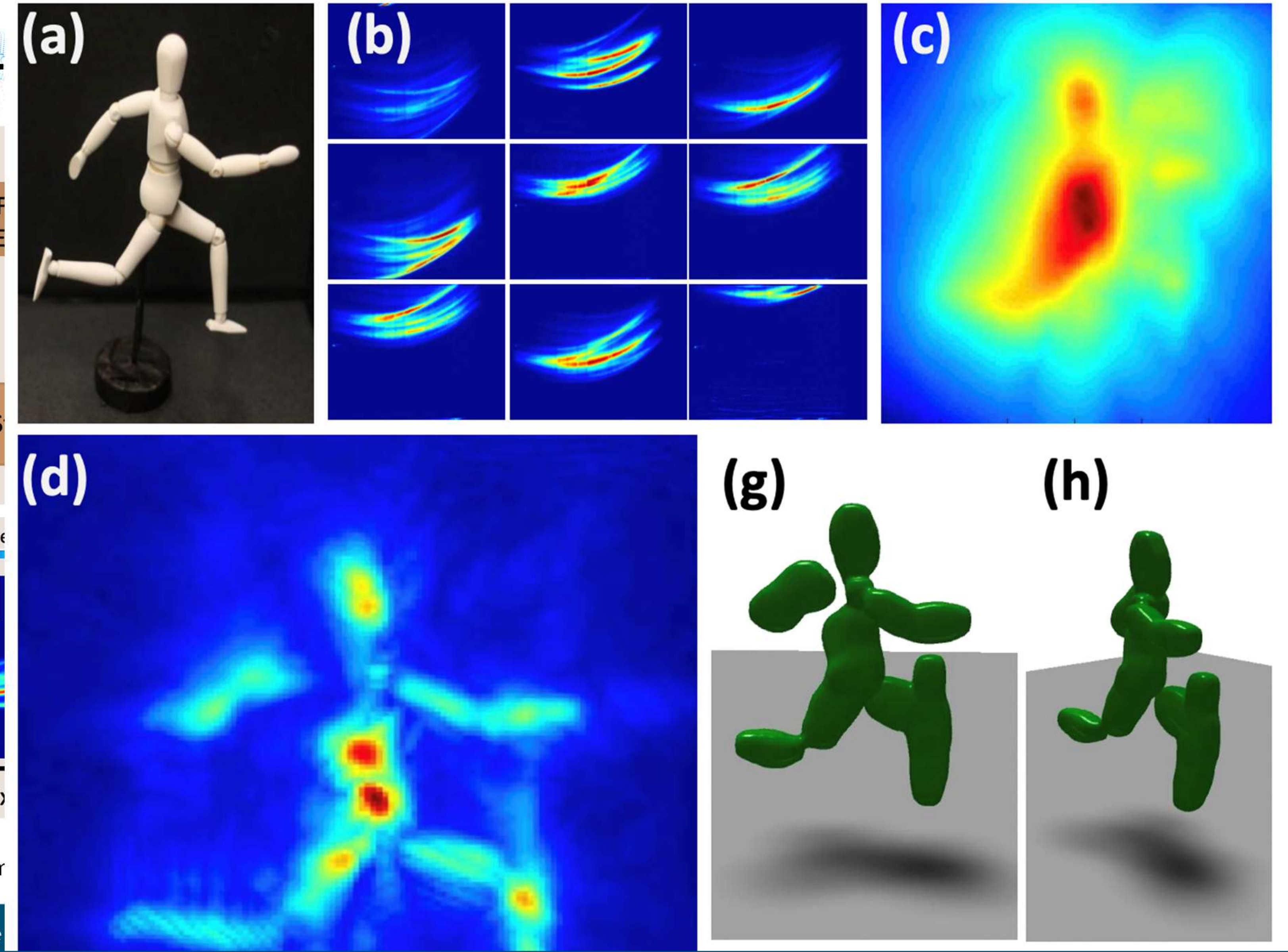
Marc Levoy

http://graphics.stanford.edu/~levoy/

Ramesh Raskar

http://web.media.mit.edu/~raskar/

Femto-

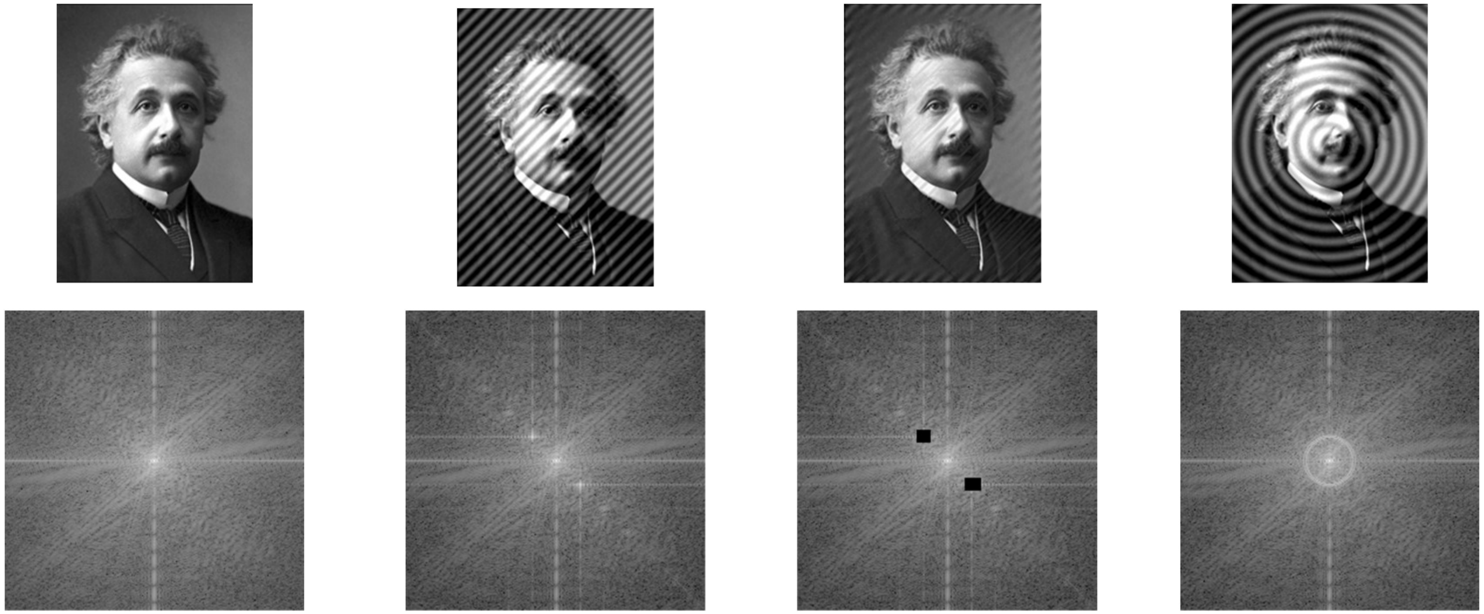


https://dam-prod.m

Case studies

Hope, you enjoyed the lab!

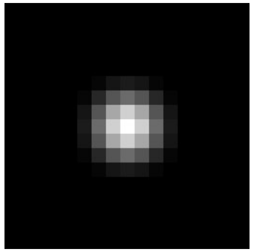
- Fourier representation and filtering/sampling of particular frequencies



The deconvolution problem

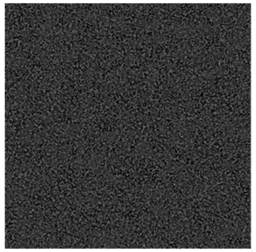
Underlying image

$\otimes$



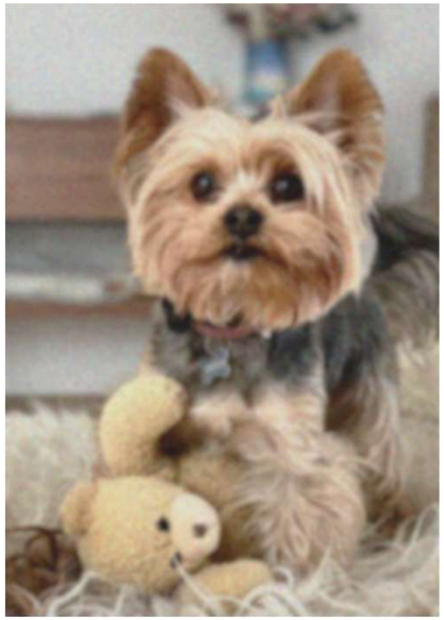
Blurring kernel

$+$



Noise (e.g., Gaussian)

$=$



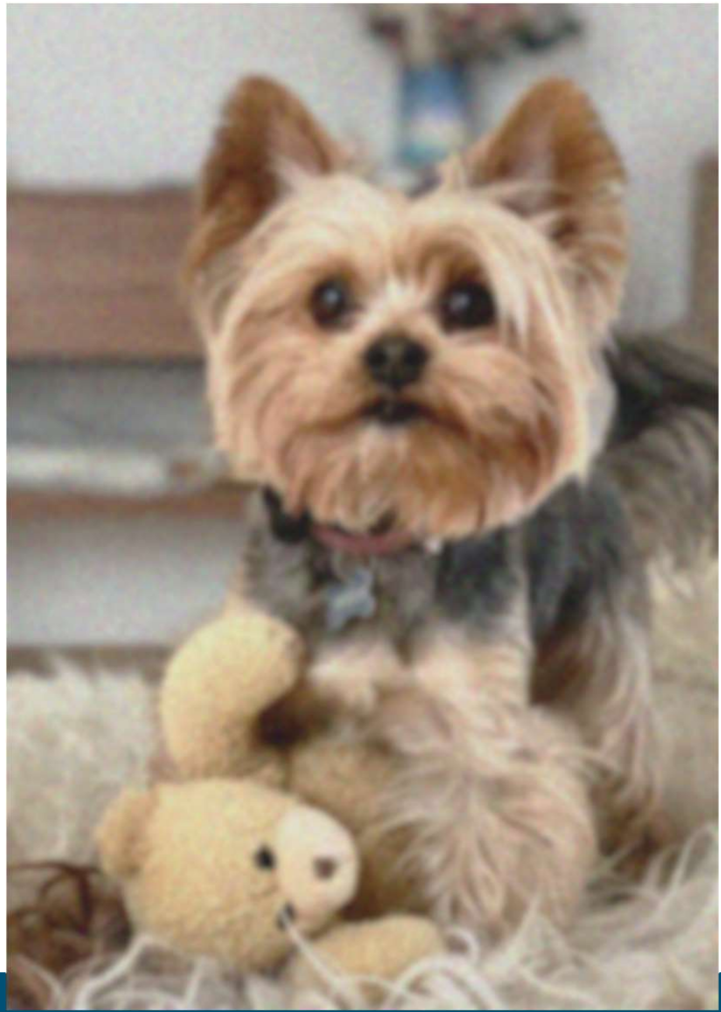
Distorted image

$$x \otimes K + n = y$$

Solution of the optimization problem is obtained by minimizing the objective function

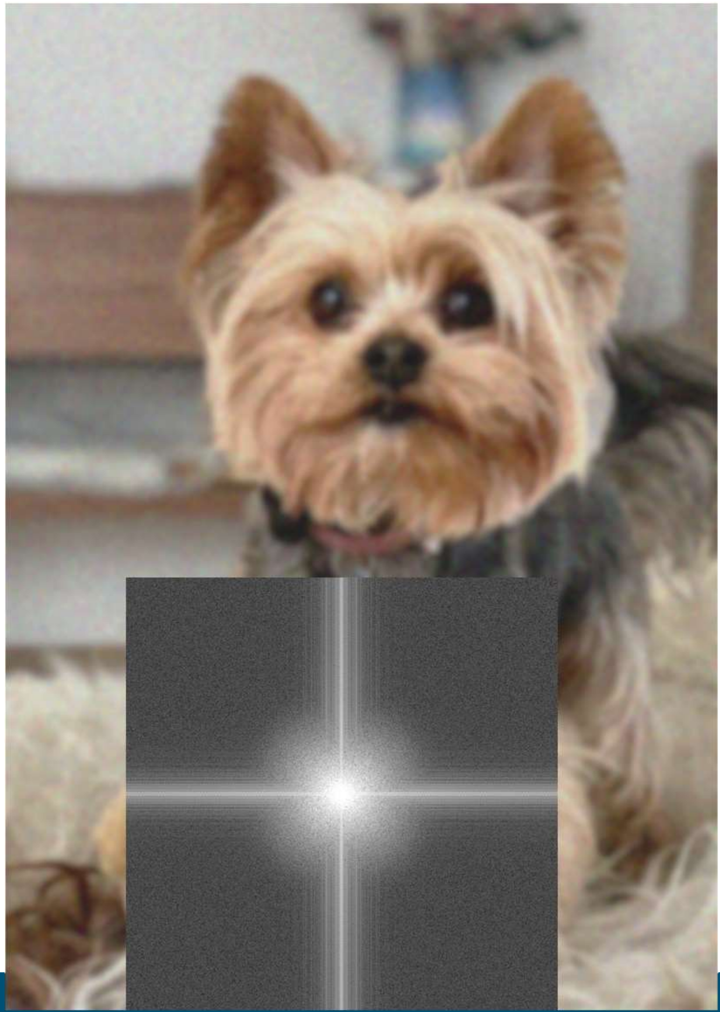
$$\hat{\mathbf{x}} = \underset{x}{\operatorname{argmin}} \frac{1}{2} ||\mathbf{y} - \mathbf{K}\mathbf{x}||_2^2$$

What is the difference, really?

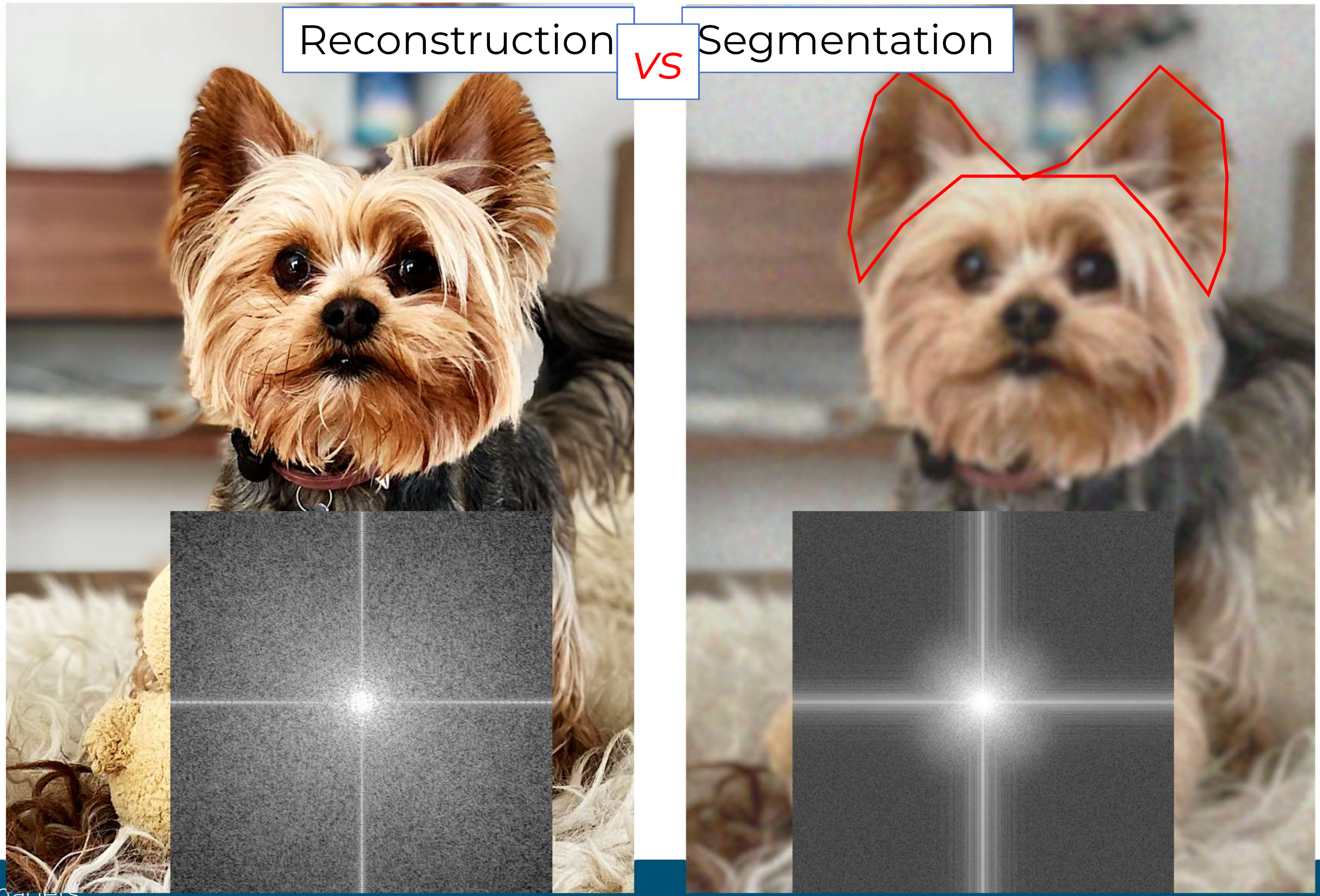


What is the difference, really?

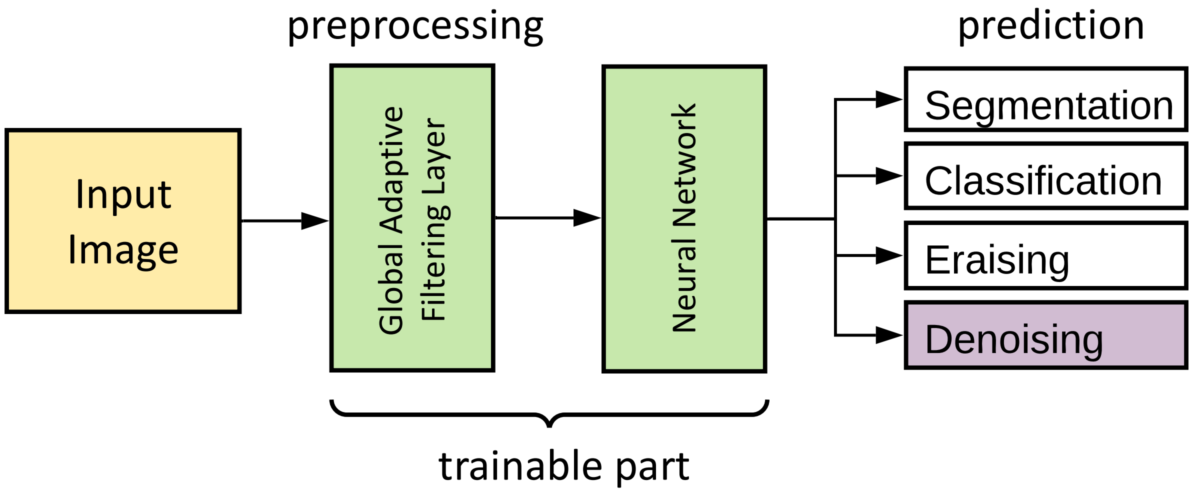
Fourier Space



How to search for the optimal frequencies?



Global Adaptive Filtering Layer[1]



Fourier Transform

$$\mathcal{F}I(\mathrm{u},\mathrm{v}) = \sum_{\mathrm{x}=0}^{n-1}\sum_{\mathrm{y}=0}^{m-1} \frac{I(\mathrm{x},\mathrm{y})}{nm} \exp\left\{-\frac{2\pi i \mathrm{x}\mathrm{u}}{n} - \frac{2\pi i \mathrm{y}\mathrm{v}}{m}\right\}$$

Global Adaptive Filtering Layer (GAFL)

Input: I – Initial image.

$\mathcal{F}$ – Fast Fourier Transform operator.

1: $W_1, W_2, B_1, B_2 = ReLU(\ W_1, W_2, B_1, B_2)$;

2: $F = \mathcal{F} I$;

3: $S = W_2 * \sigma\big(W_1 * |F| + B_1\big) + B_2$;

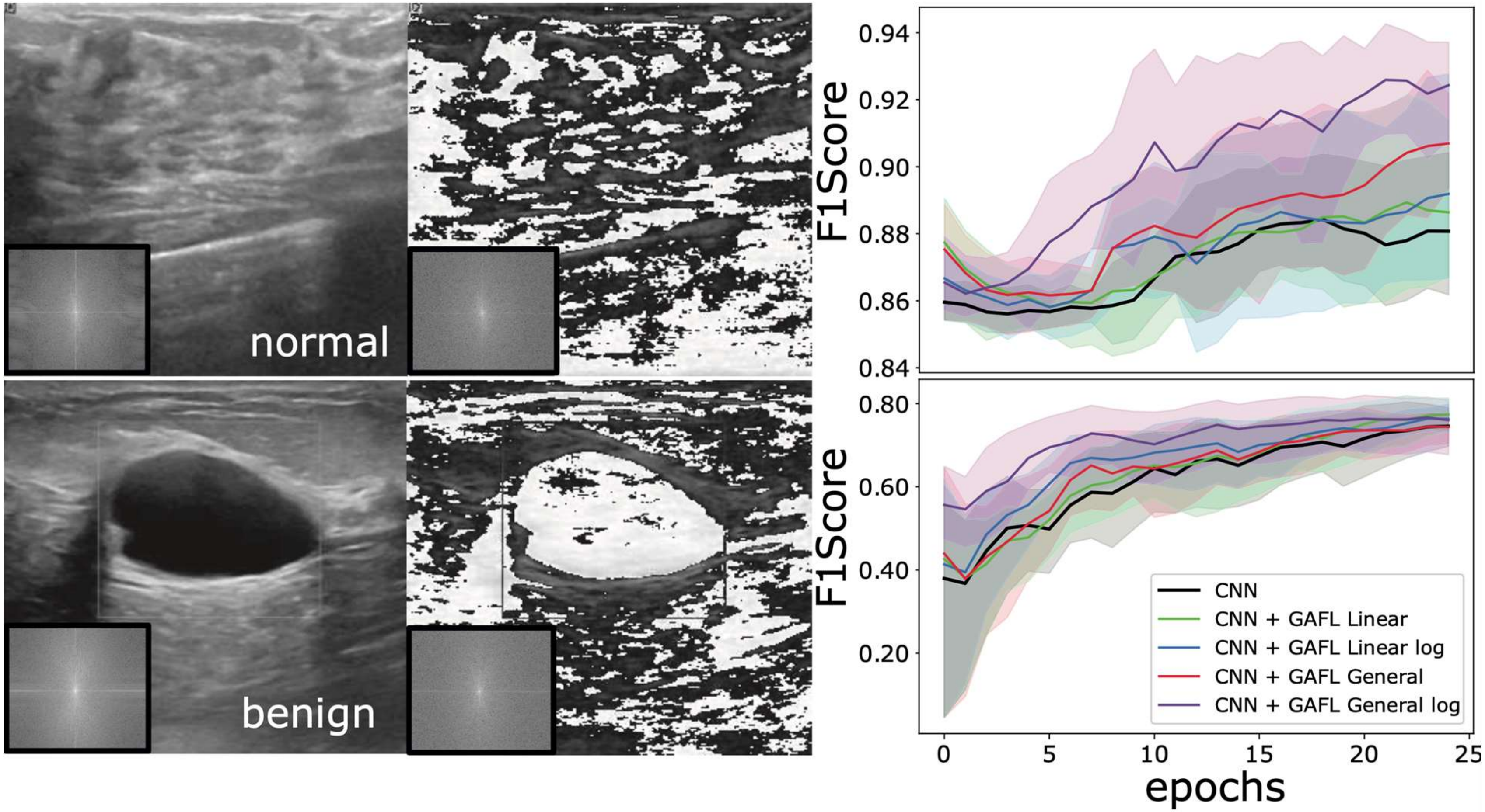
4: $S = S * F / |F|$;

5: $I' = \mathcal{F}^{-1} S$;

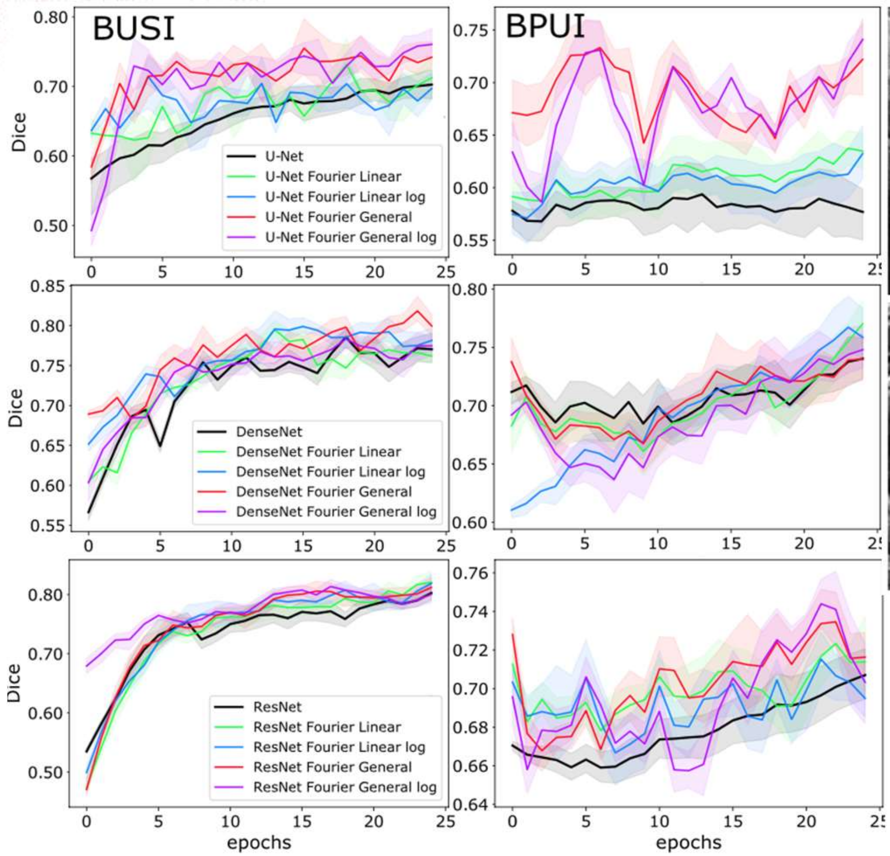
Output: I' – Image after global frequency filtering.

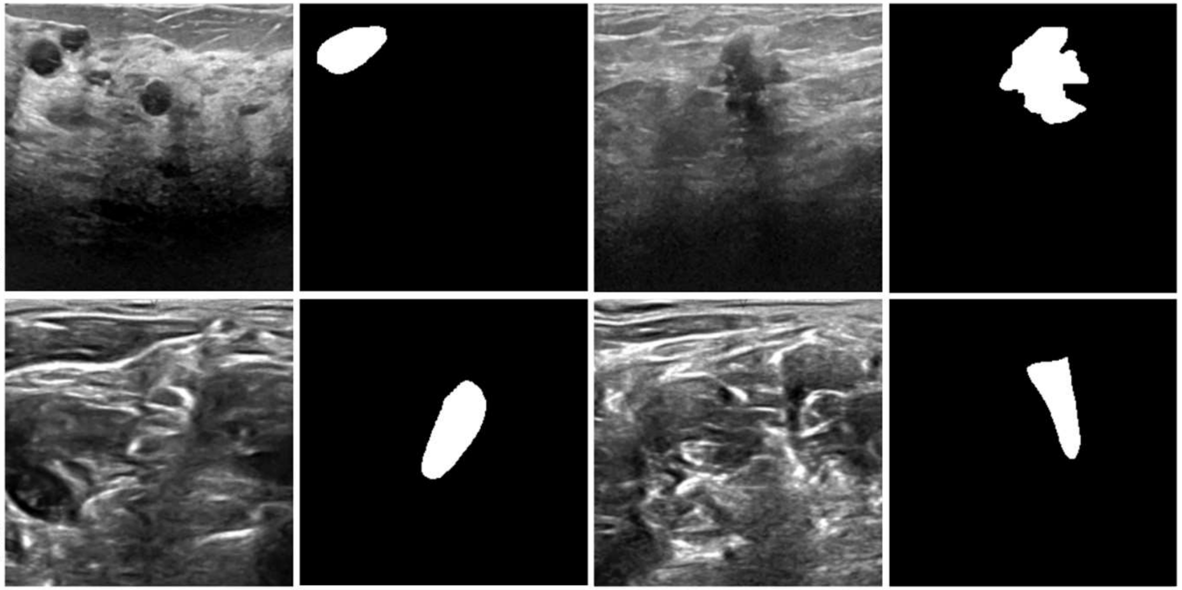
[1] Shipitsin, Bespalov, Dylov, «GAFL: Global Adaptive Filtering Layer» arXiv: 2010.01177, 2021.

Classification results



Segmentation results

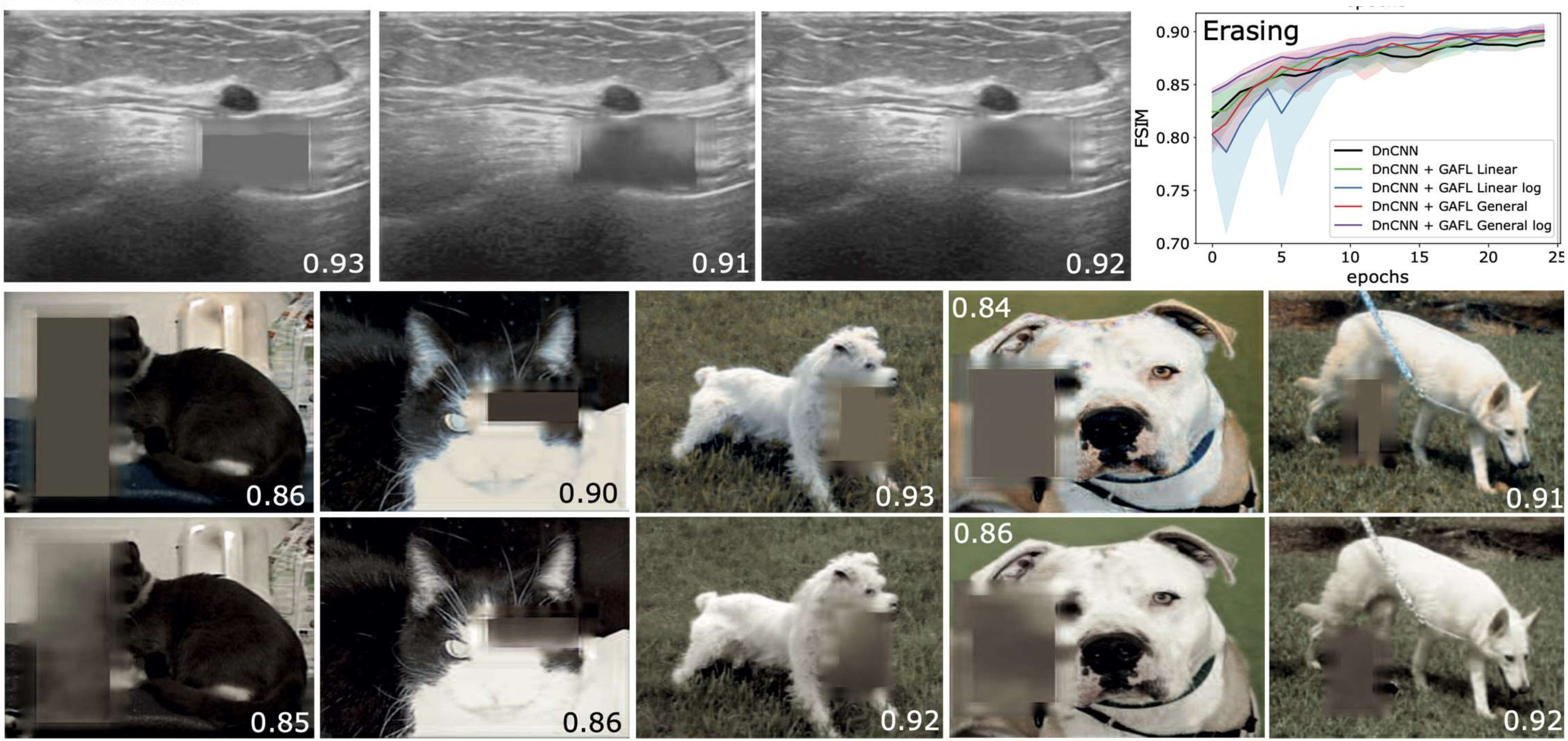




first raw: BUSI dataset (breast cancer)[1].
second raw: BPUI (branchial plexus).[2]

[1] Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Aly Fahmy. Dataset of breast ultrasound images.
[2] https://www.kaggle.com/c/ ultrasound-nerve-segmentation/data.

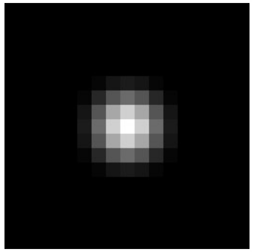
Denoising/Erasing results



The Imaging Equation

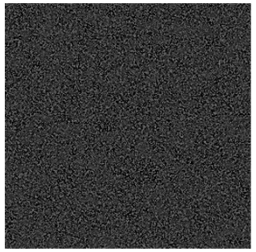
Underlying image

$\otimes$



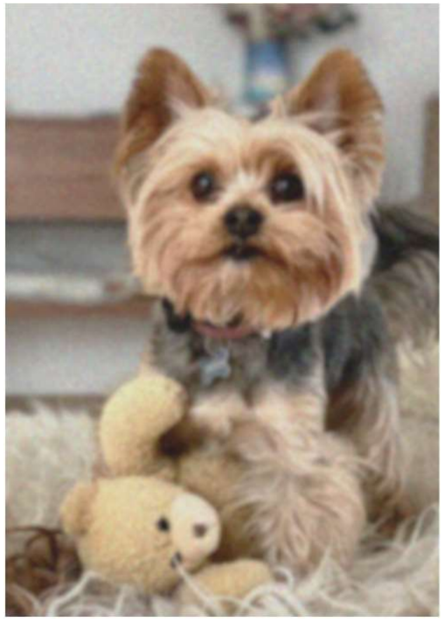
Blurring kernel

$+$

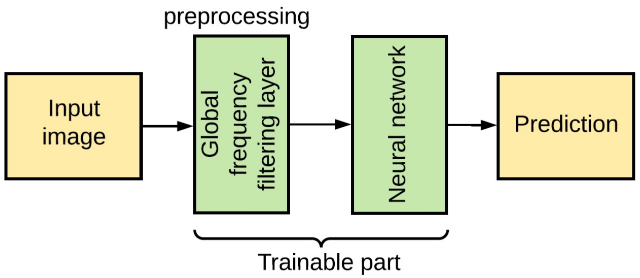


Noise (e.g., Gaussian)

$=$

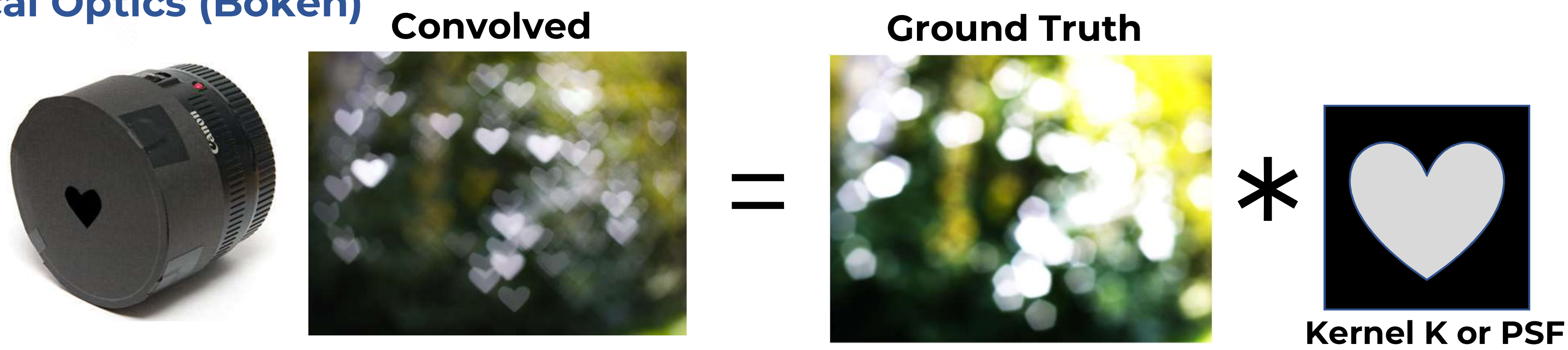


Distorted image



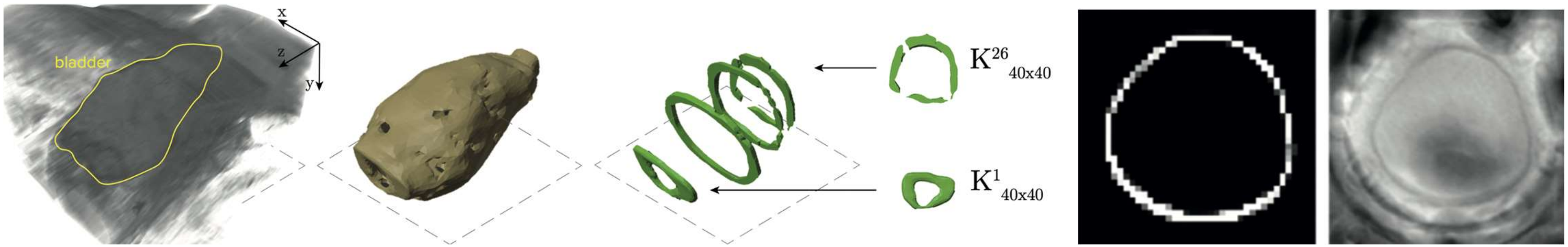
Motivation: analogies between *physical* and *digital* worlds

Physical Optics (Bokeh)

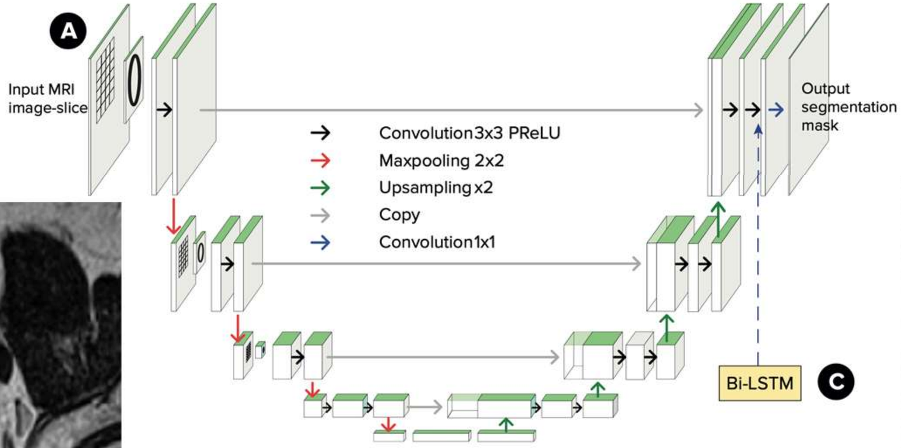


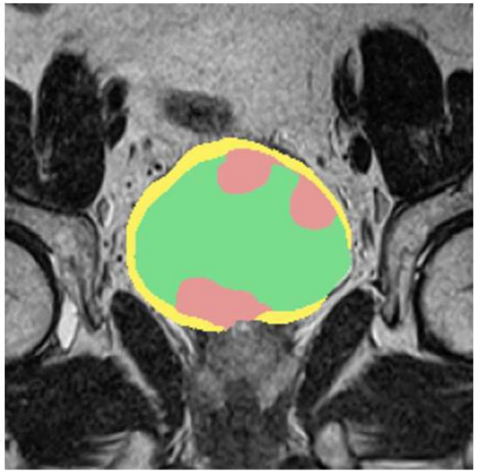
Source: http://lullaby.homepage.dk/diy-camera/bokeh.html

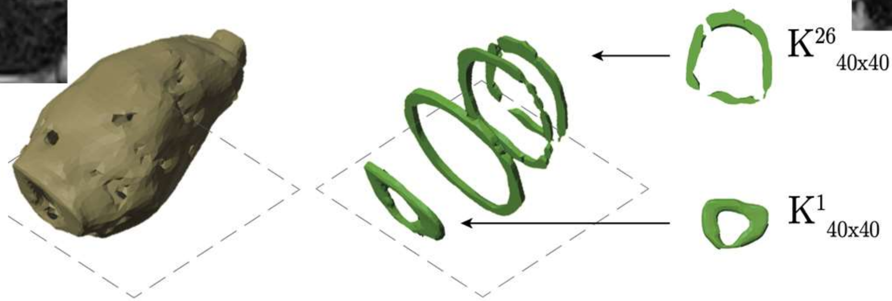
Medical Computer Vision (Segmentation task)



Bladder segmentation problem:
U-Net and new **object-resembling kernels** for it



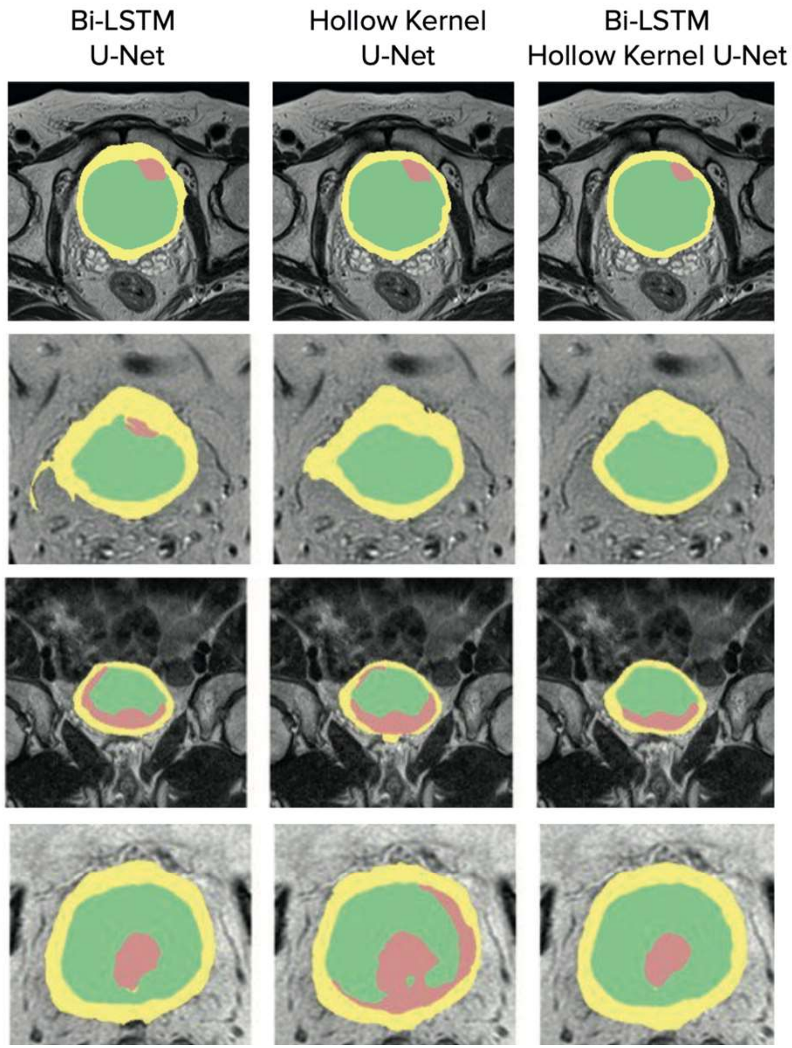




New State-of-the-Art for Bladder

Table 1: Mean Dice coefficient values on test data.

Architecture	Inner Wall	Outer Wall	Tumor
U-Net [8]	0.911 ± 0.017	0.699 ± 0.023	0.587 ± 0.049
U-Net Baseline [10]	0.897 ± 0.022	0.699 ± 0.024	0.638 ± 0.048
U-Net Dilated [10]	0.921 ± 0.013	0.718 ± 0.022	0.663 ± 0.052
U-Net Progressive Dilated [10]	0.914 ± 0.016	0.713 ± 0.022	0.656 ± 0.050
E-Net [9]	0.913 ± 0.015	0.712 ± 0.018	0.674 ± 0.045
Bi-LSTM U-Net [11]	0.914 ± 0.019	0.715 ± 0.019	0.642 ± 0.051
Temporal U-Net	0.917 ± 0.016	0.710 ± 0.017	0.646 ± 0.043
Temporal U-Net Dilated	0.910 ± 0.017	0.701 ± 0.021	0.617 ± 0.058
U-Net models with hollow kernels:			
A1 Config. 1.1 L1	0.905 ± 0.020	0.698 ± 0.024	0.623 ± 0.066
A1 Config. 1.1 L2	0.908 ± 0.020	0.693 ± 0.023	0.605 ± 0.063
A1 Config. 1.1 L3	0.894 ± 0.028	0.714 ± 0.023	0.652 ± 0.052
A2 Config. 1.1	0.918 ± 0.016	**0.725** ± 0.020	0.662 ± 0.052
A2 Config. 2.1	**0.923** ± 0.013	0.724 ± 0.022	0.673 ± 0.052
A2 Config. 2.2	0.922 ± 0.014	0.714 ± 0.022	0.642 ± 0.054
A2 Config. 2.3	0.917 ± 0.018	0.721 ± 0.023	**0.675** ± 0.052
A2 Config. 3.1	0.914 ± 0.018	0.709 ± 0.020	0.635 ± 0.050
LORCK (best)	**0.936** ± 0.011	**0.736** ± 0.018	**0.712** ± 0.037



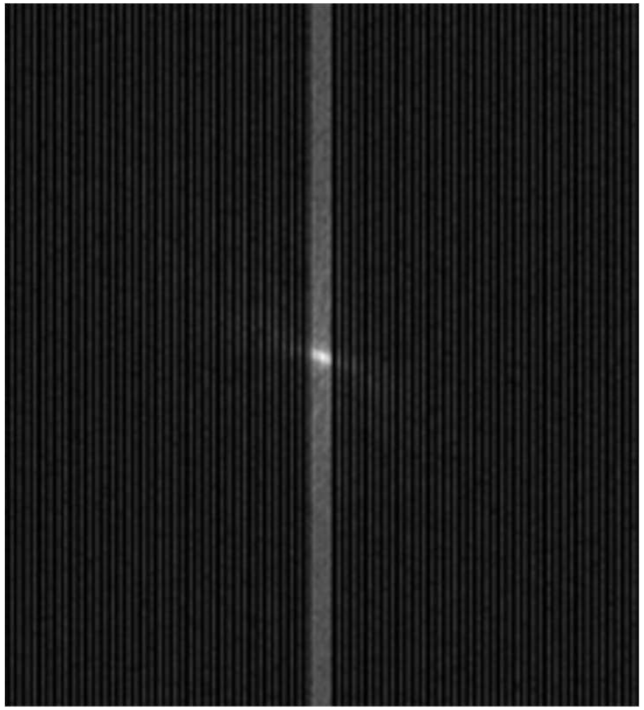
“Learnable Object-Resembling Convolution Kernels”
Elizaveta Lazareva, Oleg Rogov, Olga Shegai, Denis Larionov, Dmitry V. Dylov, arXiv 2007.05103 (2020).

Magnetic Resonance Imaging

Data acquisition takes **a lot of time**: typically, from 20 to 60 min

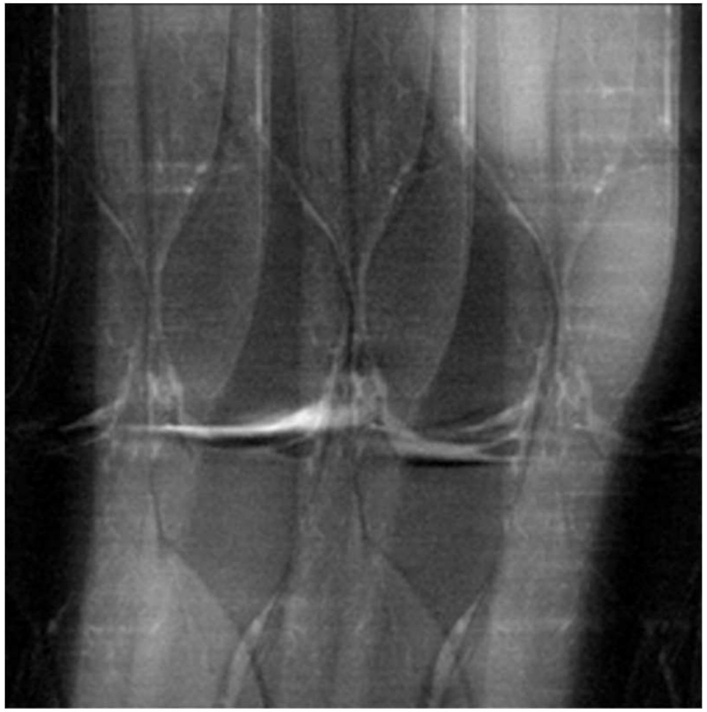
Standard sequential Cartesian filling of k-space (Courtesy of Brian Hargreaves)

Problem statement



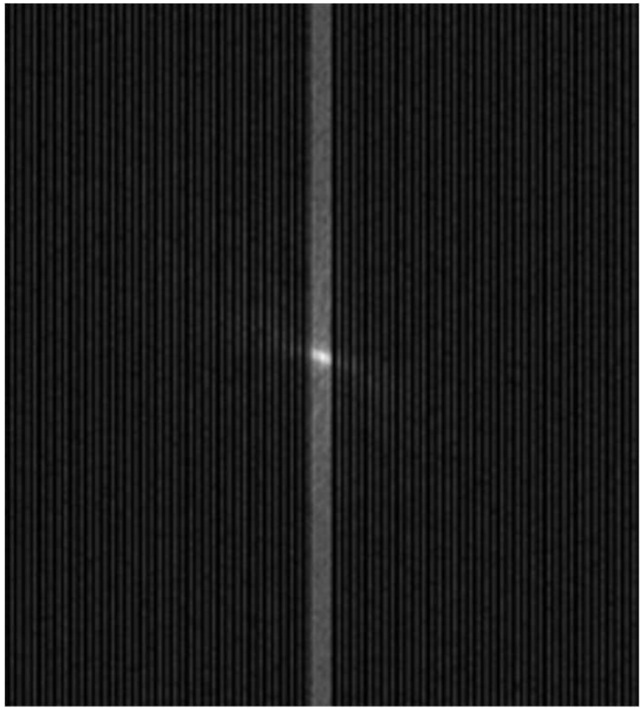
k-space

Standard reconstruction



Reconstructed image

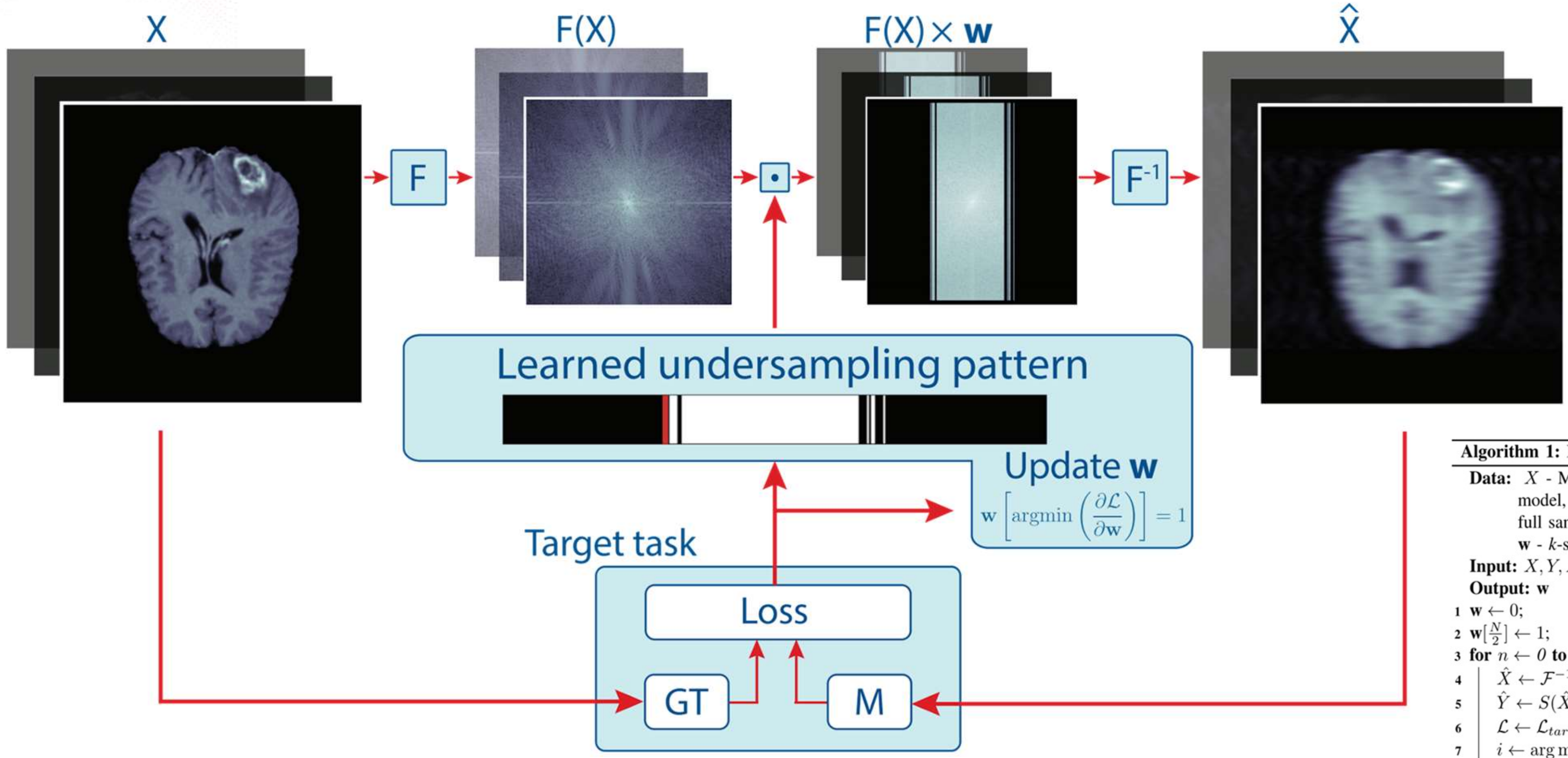
Problem statement



k-space

Reconstructed image

Learn the Fourier mask iteratively

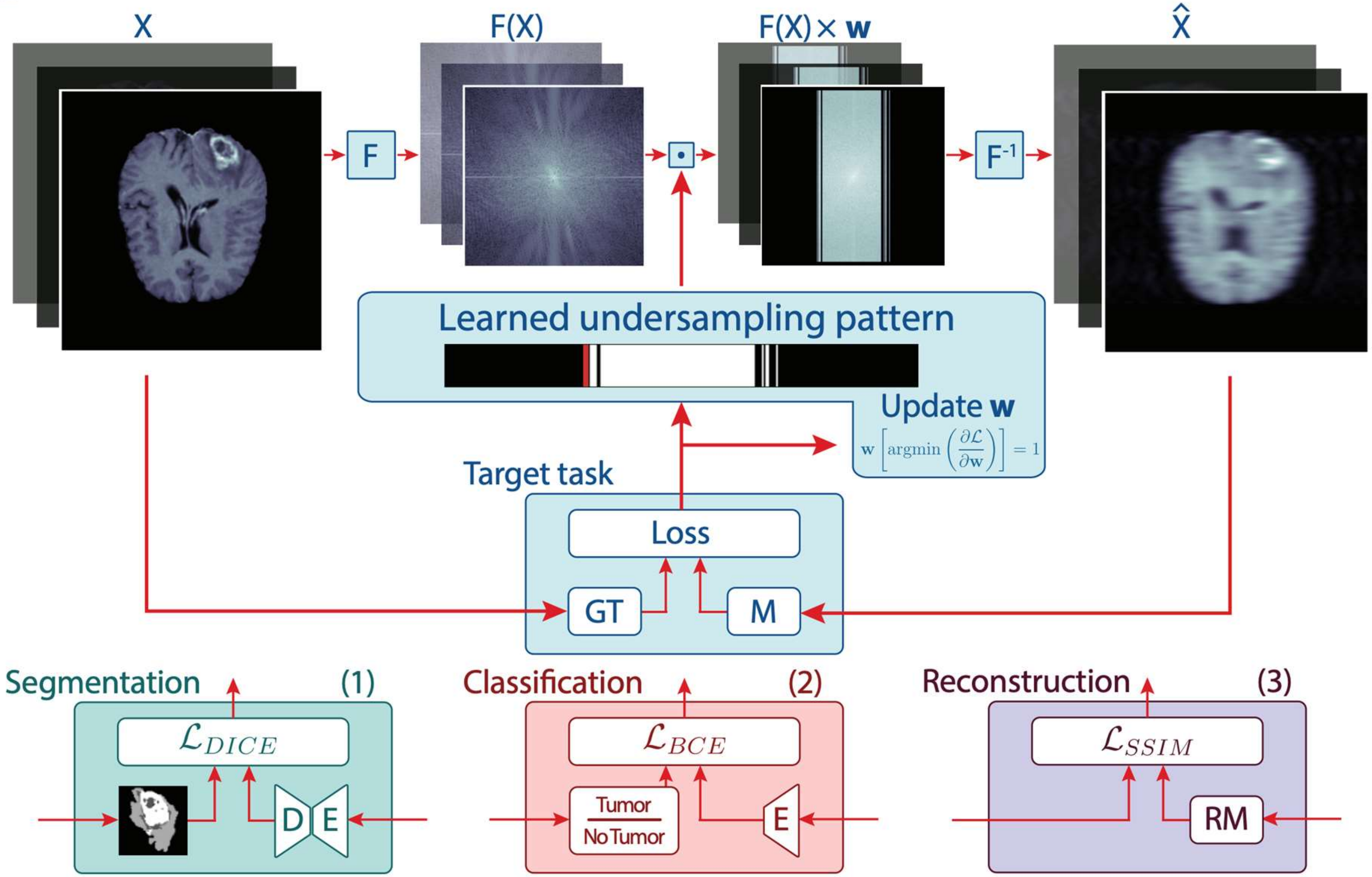


Algorithm 1: Iterative gradients sampling

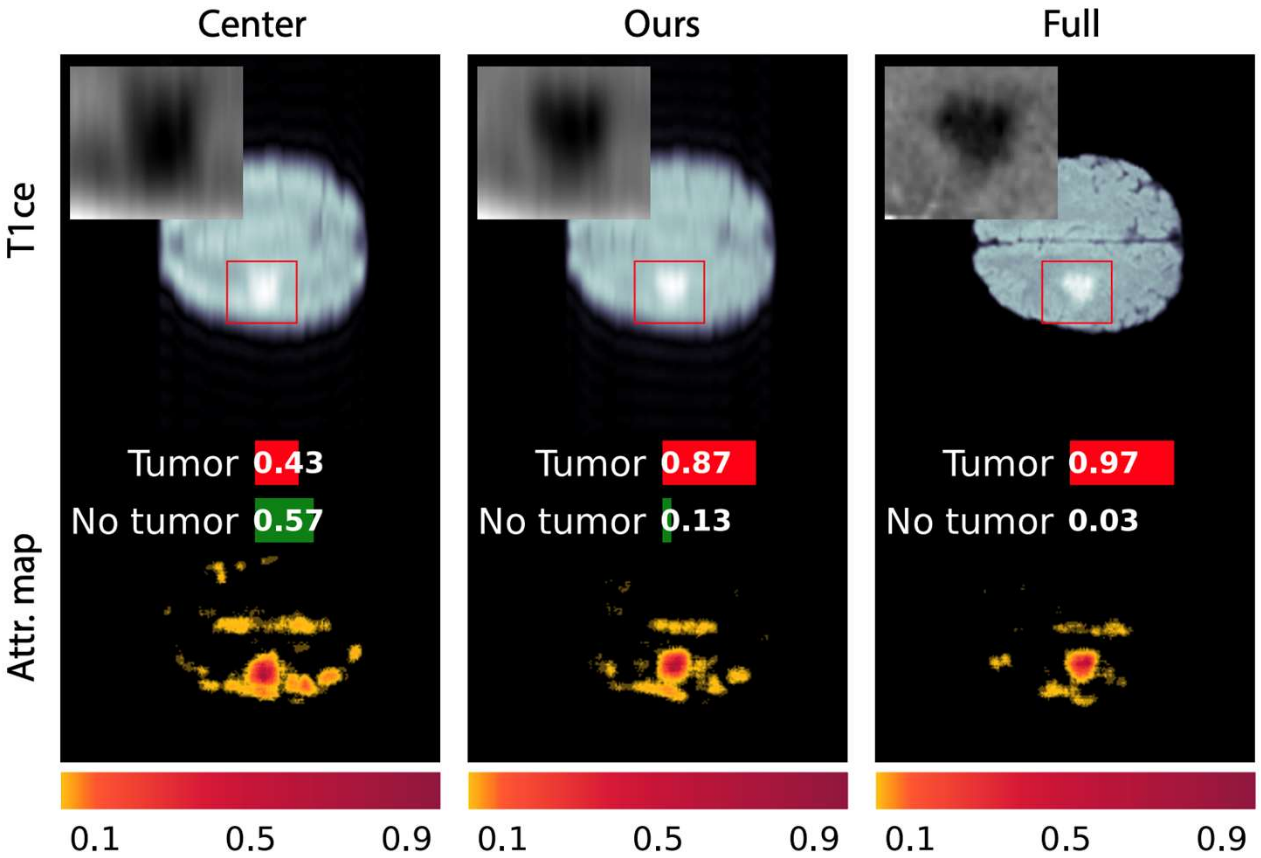
Data: X - MR images, Y - ground truth, $S(\cdot)$ - CNN model, $\mathcal{L}_{target}(\cdot,\cdot)$ - target loss function, N - full sampling size, N_s - partial sampling size, $\mathbf{w}$ - k-space undersampling pattern.
Input: X, Y, N, N_s
Output: $\mathbf{w}$

```
w ← 0;
w[N/2] ← 1;
for n ← 0 to N_s do
    X̂ ← F^{-1}(F(X) · w);
    Ŷ ← S(X̂);
    L ← L_target(Y, Ŷ);
    i ← arg min_{i | w[i]=0} (∂L/∂w);
    w[i] ← 1
end
```

Result: Optimized sampling pattern $\mathbf{w}$ for acceleration factor $\alpha = N_s/N$ with respect to model $S(\cdot)$.


X
F(X)
F(X) × w
X̂
F
F-1
Learned undersampling pattern
Update w
$\mathbf{w}\left[\text{argmin}\left(\frac{\partial \mathcal{L}}{\partial \mathbf{w}}\right)\right] = 1$
Target task
Loss
GT
M
Segmentation
(1)
$\mathcal{L}_{DICE}$
D
E
Classification
(2)
$\mathcal{L}_{BCE}$
Tumor
No Tumor
E
Reconstruction
(3)
$\mathcal{L}_{SSIM}$
RM

Classification results



Correct Fourier mask allows for more accurate classification at fixed acceleration (16X)

Segmentation results

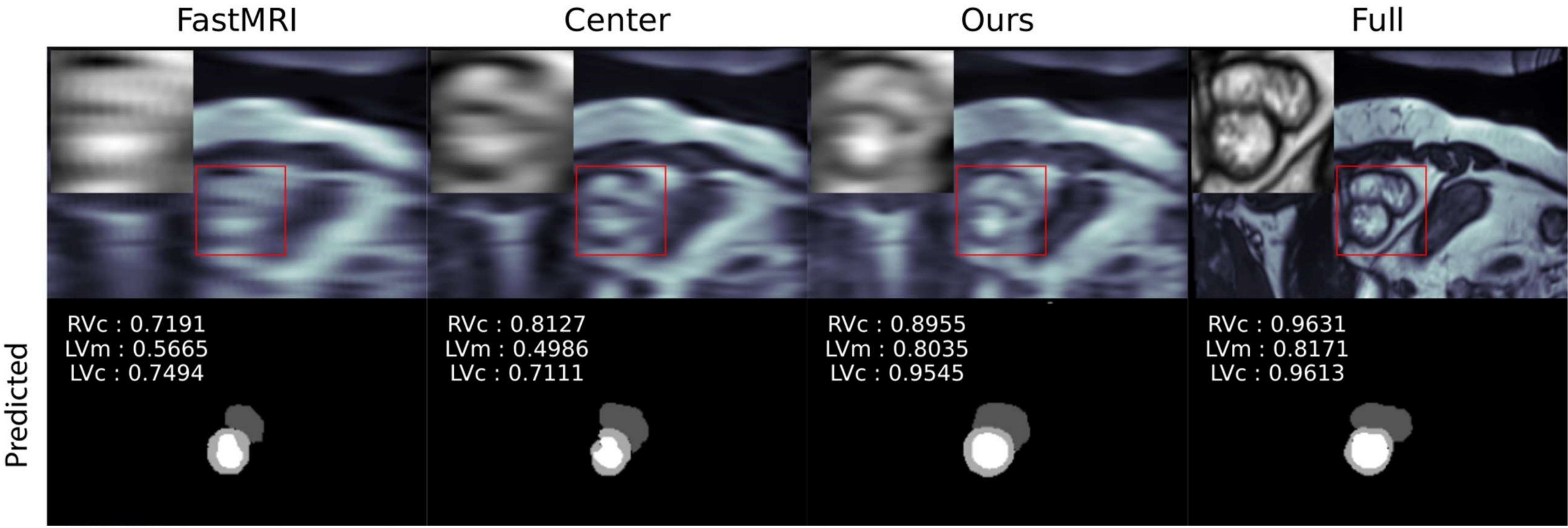
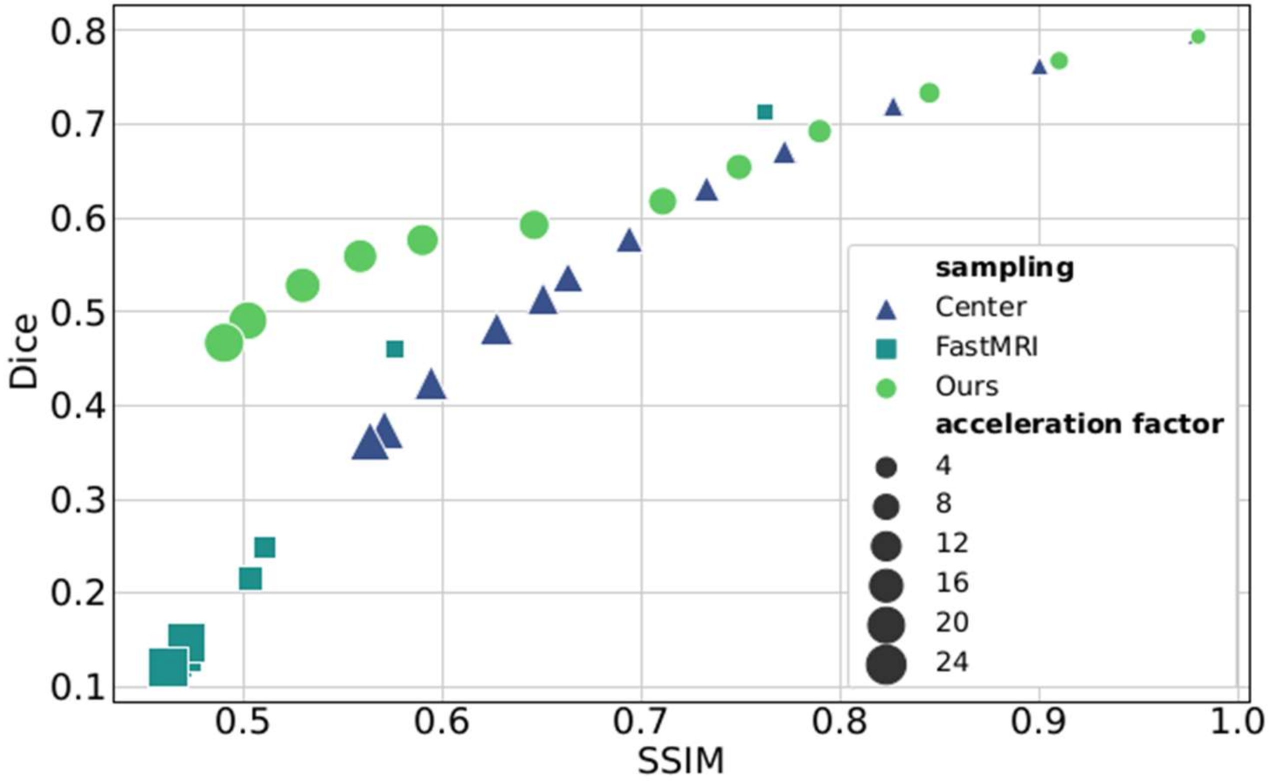


TABLE IV: **Segmentation** results. Dice scores on BraTS dataset at ×16 acceleration.

	FastMRI		Center		Ours		Full	
	U-Net	U-Net 3D	U-Net	U-Net 3D	U-Net	U-Net 3D	U-Net	U-Net 3D
WT	0.808	0.739	0.821	0.739	**0.835**	0.787	0.873	0.888
TC	0.640	0.650	0.673	0.623	0.690	**0.698**	0.726	0.792
ET	0.387	0.448	0.442	0.390	0.468	**0.481**	0.575	0.623

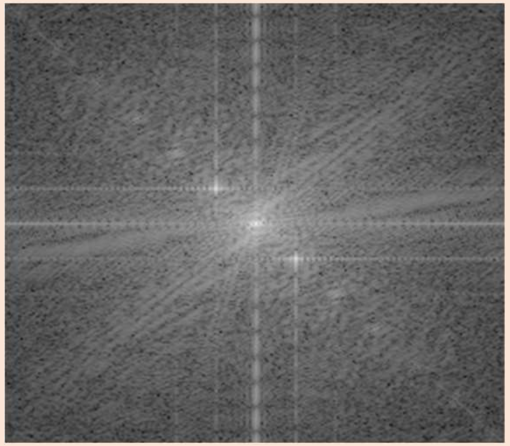
Segmentation vs Reconstruction



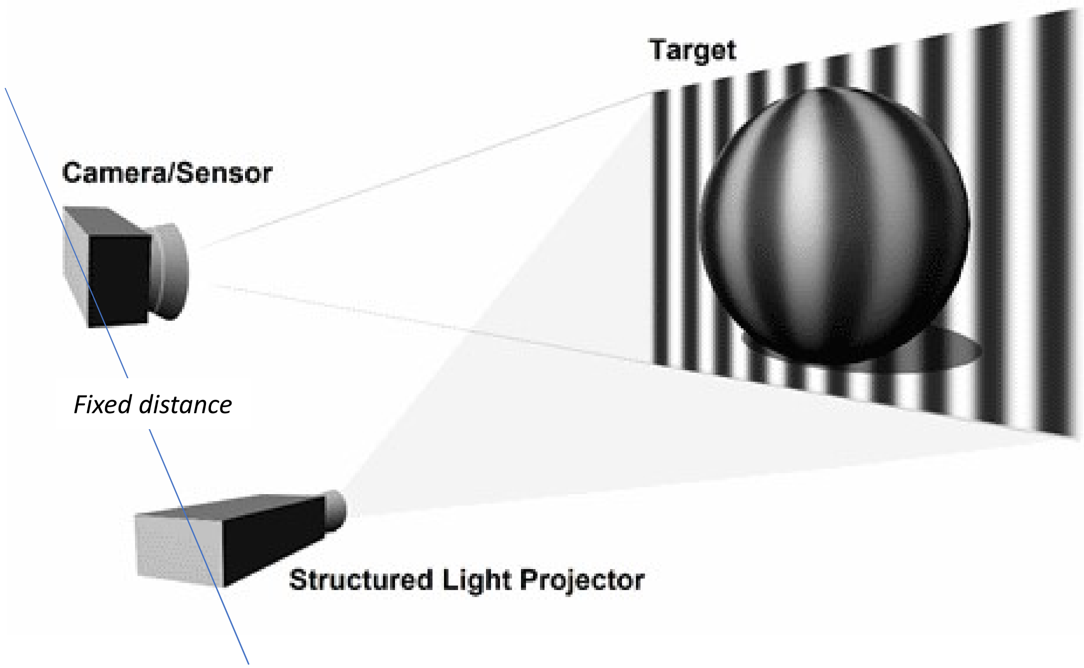
Optimization for the ultimate benefit of the downstream task!

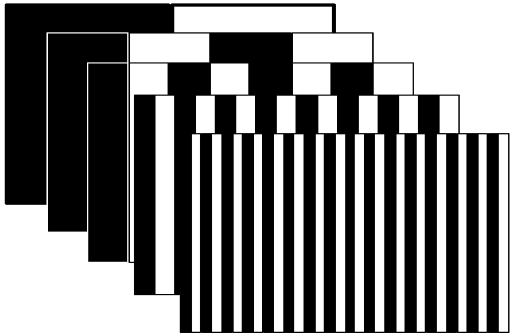
Razumov, Rogov, Dylov,
MICCAI 2022. arxiv:2108.04914

Your lab #3



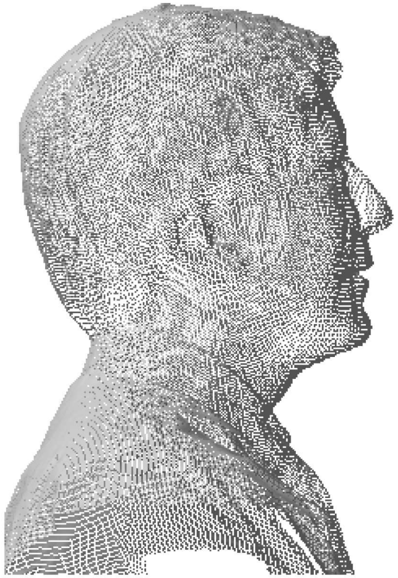
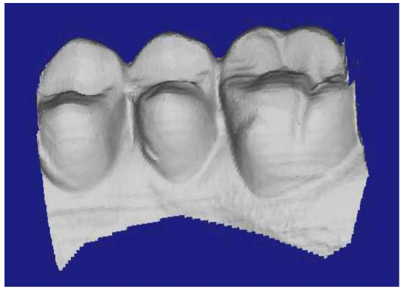
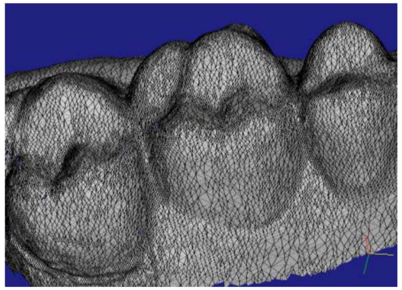
Structured computational illumination





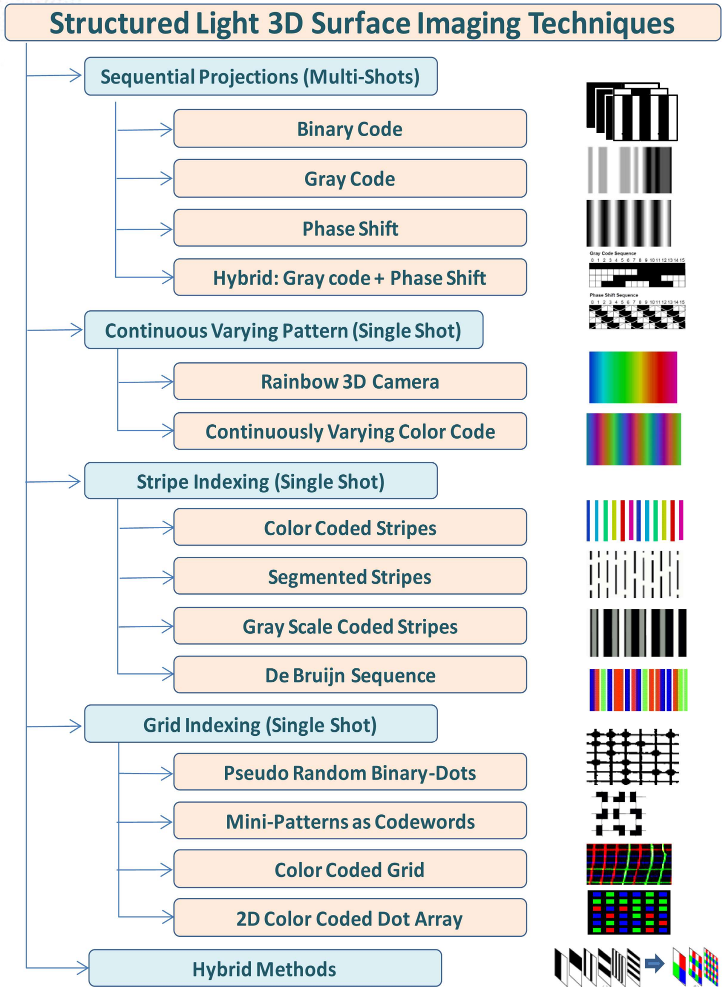
Jason Geng, "Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photon. **3,** 128-160 (2011); https://www.osapublishing.org/aop/abstract.cfm?uri=aop-3-2-128

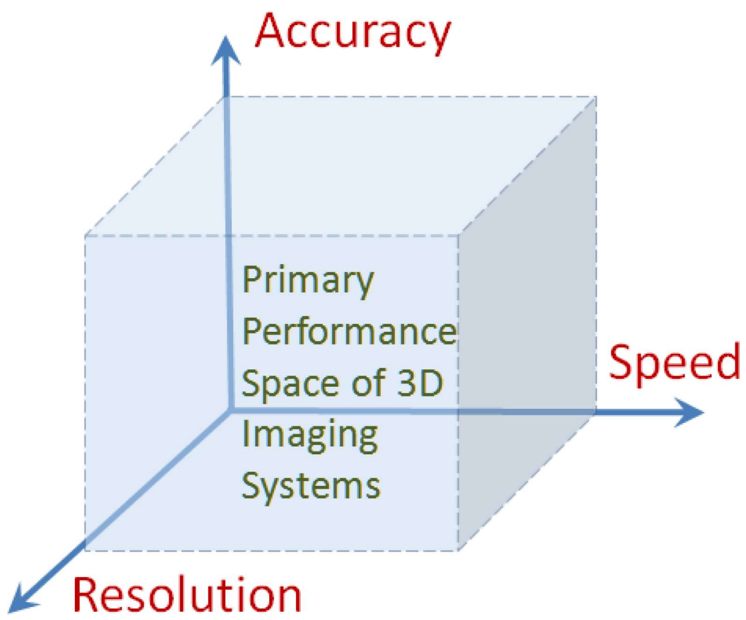
Structured illumination allows to reconstruct 3D surface precisely


Jason Geng, "Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photon. **3,** 128-160 (2011);
https://www.osapublishing.org/aop/abstract.cfm?uri=aop-3-2-128

Structured illumination: hundreds of ways to encode known patterns

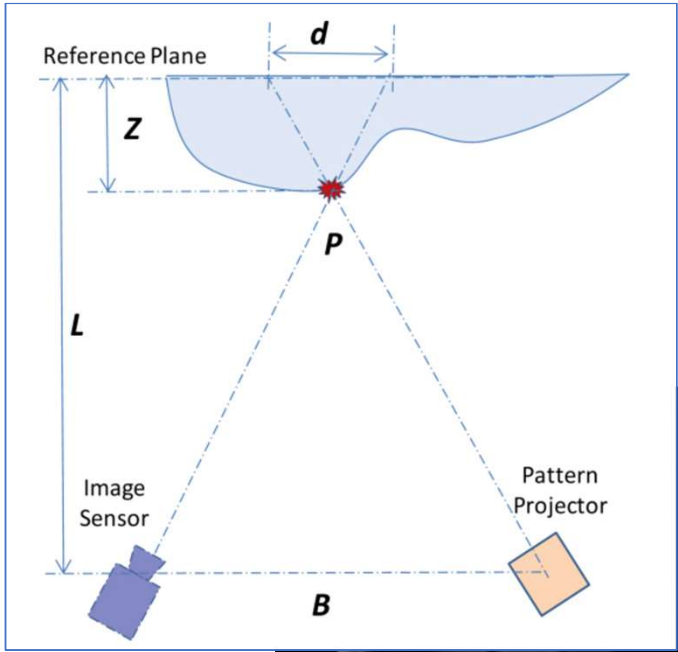


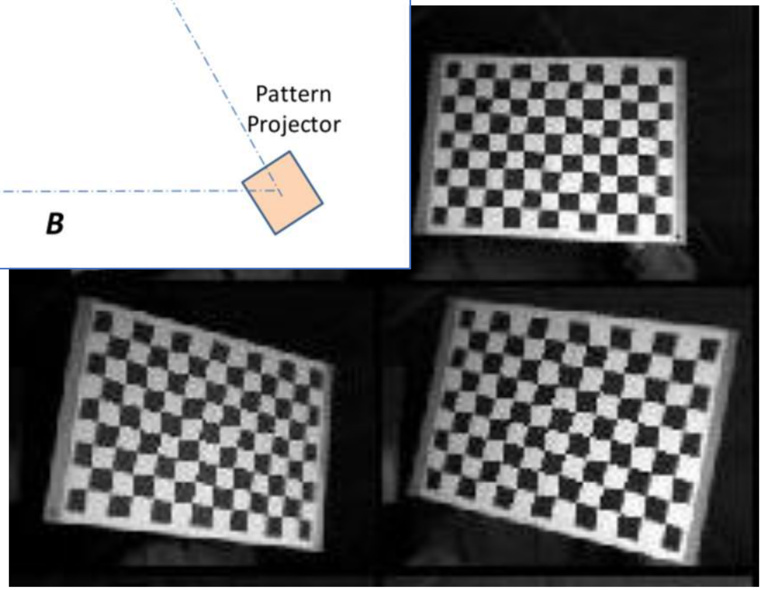


Jason Geng, "Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photon. **3,** 128-160 (2011); https://www.osapublishing.org/aop/abstract.cfm?uri=aop-3-2-128

Advice for Lab#3

- Read the paper!
- 4 teams -- find your mates before the lab.
- Most important: Camera and Projector Calibration
  - both must be calibrated simultaneously
  - find opencv realization of calibration functions
  - once finished calibrating, do not touch the setup!





Resources

- Richard Szeliski, “Computer Vision” ISBN-13: 978-1848829343
- Rastislav Lukac, “Computtional Photography” ISBN-13: 978-1439817490
- Rick Nolthenius, Noah Snavely, George Bebis, S. Seitz, L. Zhang, D. Lowe, L. Fei-Fei, Antony Lam
- Pics: Wiki commons & Khan academy
- http://cvcl.mit.edu/hybridimage.htm
- MIT media lab materials, MERL edge detection
- https://vas3k.ru/blog/computational_photography/
- https://www.baslerweb.com/en/vision-campus/vision-systems-and-components/find-the-right-lens/
- http://www.pptsing.com
- https://www.cambridgeincolour.com/tutorials/cameras-vs-human-eye.htm
- https://ai.googleblog.com/2014/10/hdr-low-light-and-high-dynamic-range.html
- http://web.media.mit.edu/~raskar/Talks/ETCVparis08/raskarCompPhotoEpsilonCodedETVC08paper.pdf
- https://graphics.stanford.edu/talks/compphot-publictalk-may08.pdf
- https://www.cambridgeincolour.com/tutorials/cameras-vs-human-eye.htm

© Credits

- **Slide |3|** E F Schubert Light Emitting Diodes (Cambridge Univ Press
- **Slide |18|** Ramesh
- **Slide |30|** Raipur, Chhattisgarh
- **Slide |31|** Survey (in Rus): https://m.habr.com/ru/post/440652/ Video of image reconstruction (Lytra): https://www.youtube.com/watch?time_continue=987&v=4qXE4sA-hLQ
- https://www.media.mit.edu/projects/looking-around-corners/overview/

VIDEOS:

https://www.youtube.com/watch?v=0q6ap_OSBAk (Event-based)

https://www.youtube.com/watch?v=sbJAi6SXOQw (Event-based)

https://www.youtube.com/watch?v=4qXE4sA-hLQ&t=987s (Plenoptic)

https://www.youtube.com/watch?v=0JbERFPWNyU (Plenoptic)

https://www.youtube.com/watch?v=-fSqFWcb4rE (Femto-photography)

Practical Machine Vision HOW TO's

- **Camera selection:**

https://www.baslerweb.com/en/vision-campus/vision-systems-and-components/camera-selection/

- **Lens & Lighting Selection:**

https://www.baslerweb.com/en/vision-campus/vision-systems-and-components/find-the-right-lens/