Thoughts on Computational Photography

Frédo Durand
MIT CSAIL

Computational Photography

Computation is a central component of image formation

The final image is not a mere projection of light onto a sensor but results from computation

I love how interdisciplinary computational photography is

Algorithms

Signal Processing

Machine Learning

Medicine

Theory

Mechanical Engineering

Arts

Architecture

Design

Optics

Human Perception

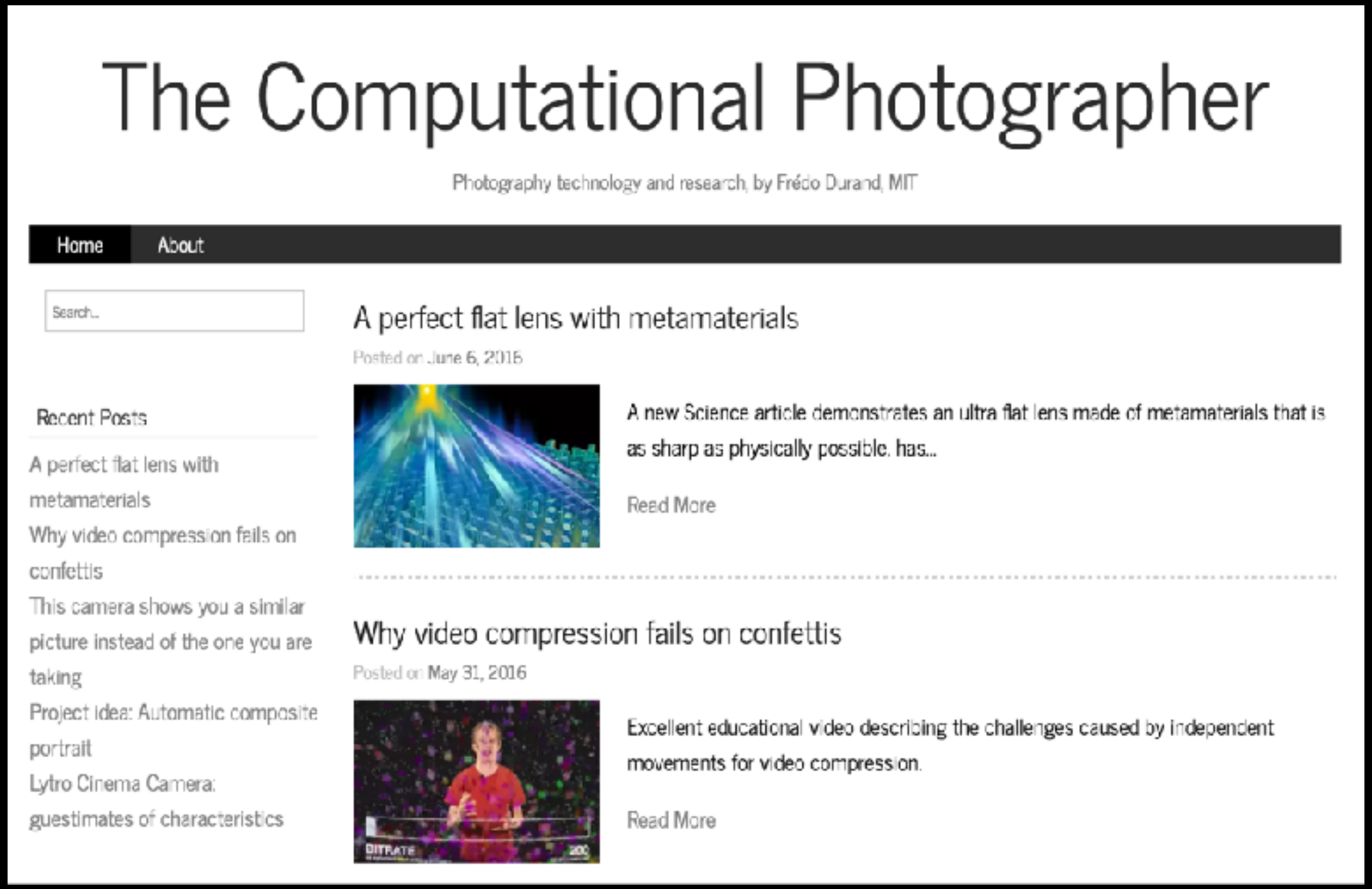
Systems

Hardware Design

Mathematics

Blog!

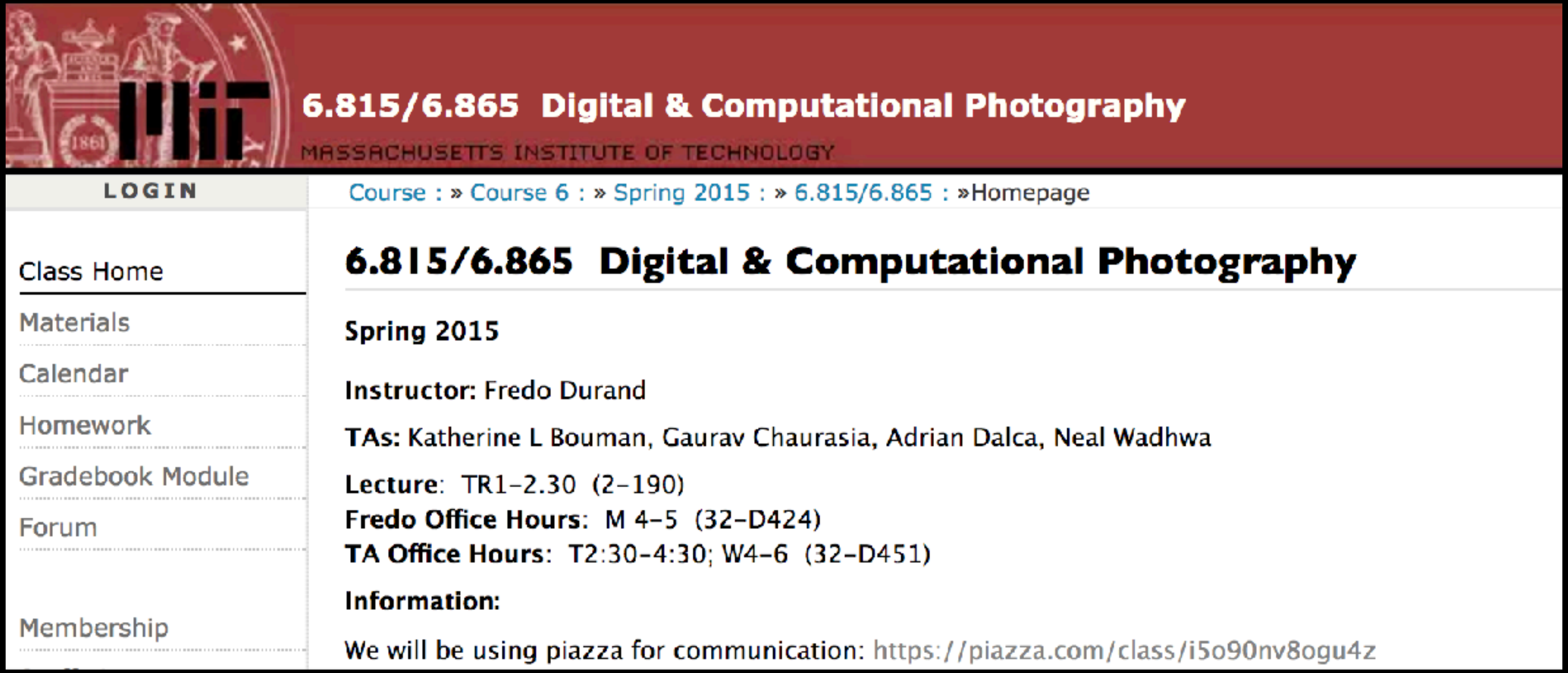
- http://thecomputationalphotographer.com
- I will post these slides there



To get started, follow my course

Accessible from http://people.csail.mit.edu/fredo/

Lots of simple programming assignments that cover the field



Plan for this talk

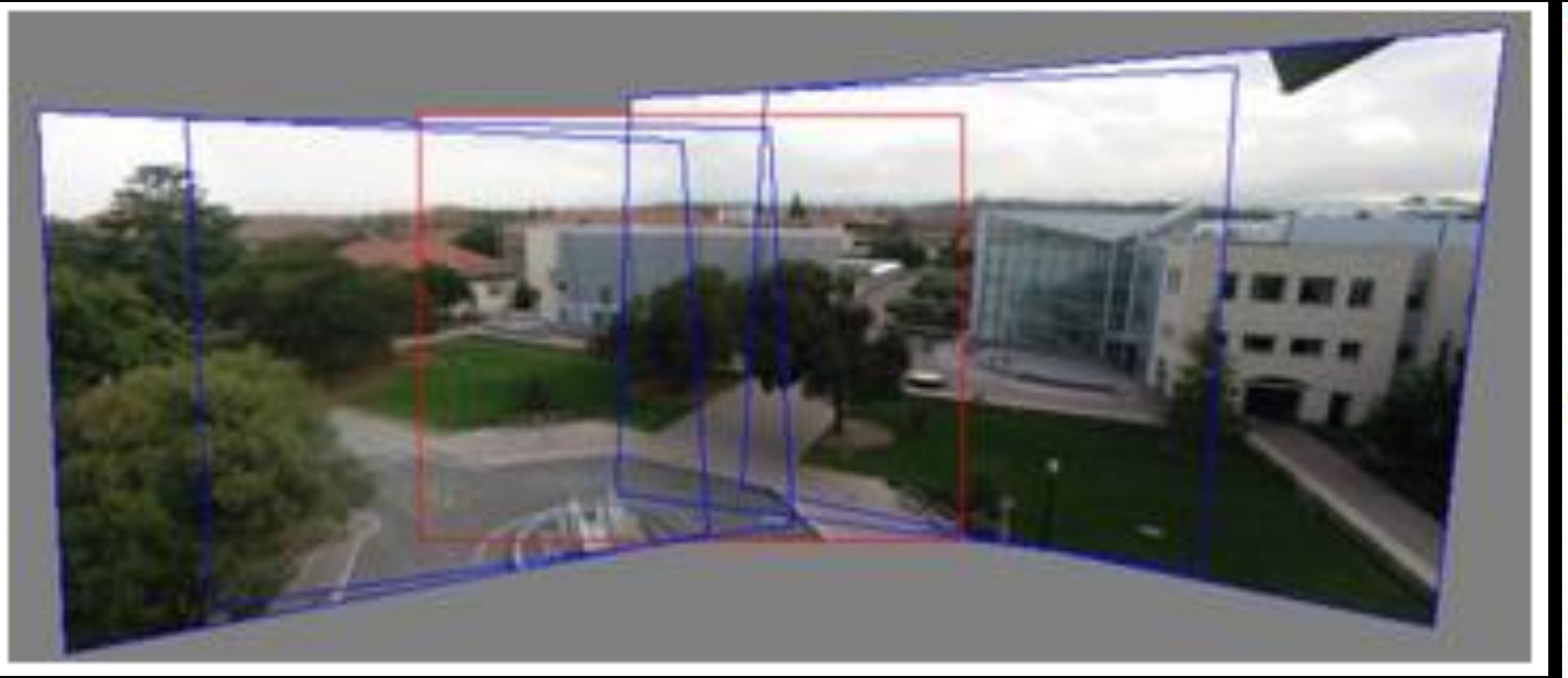
Big Ideas in Computational Photography

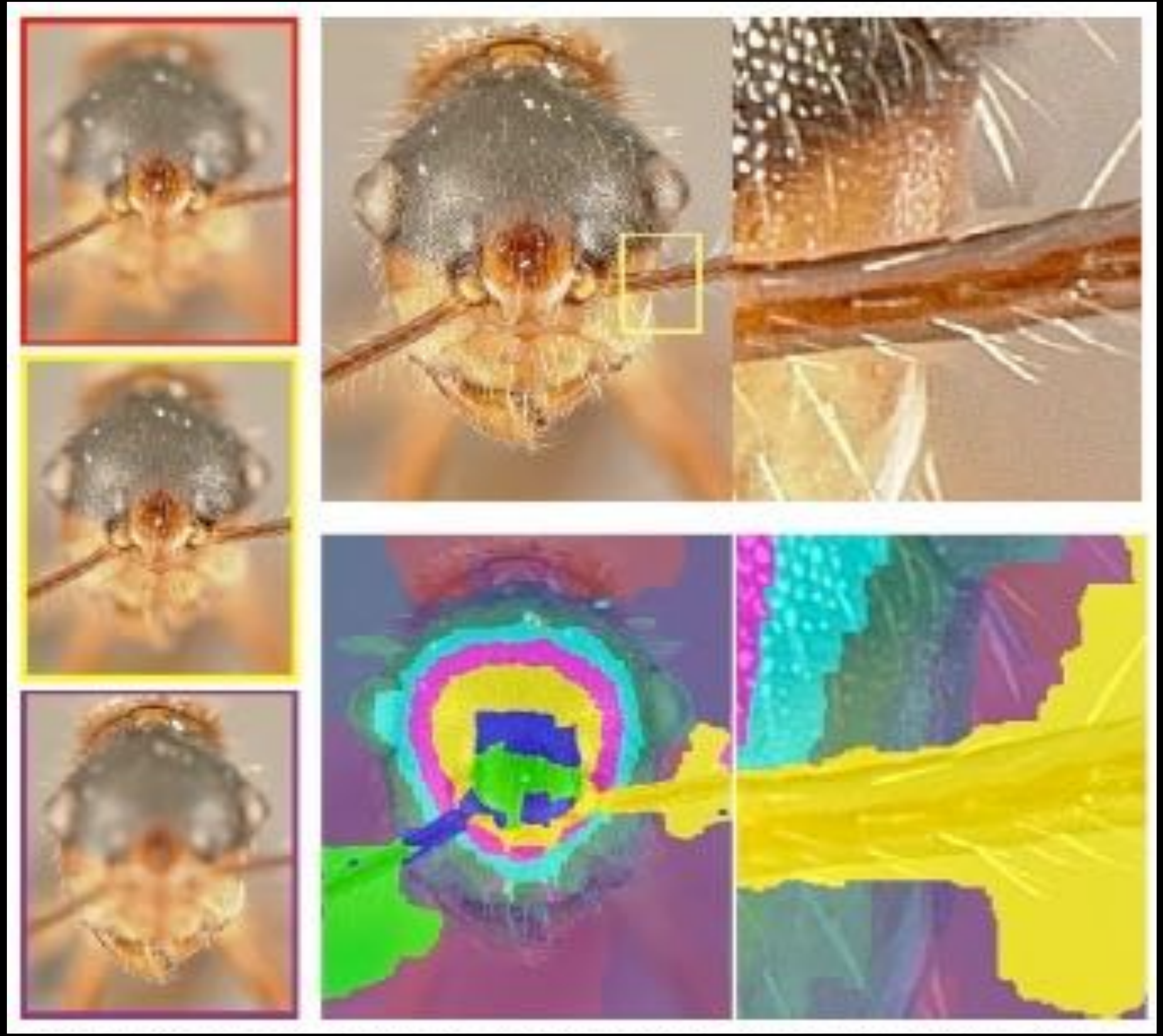
Assessment of the field

Research suggestions

Multiple exposures expand capabilities

Field of view, dynamic range, depth of field, superresolution.



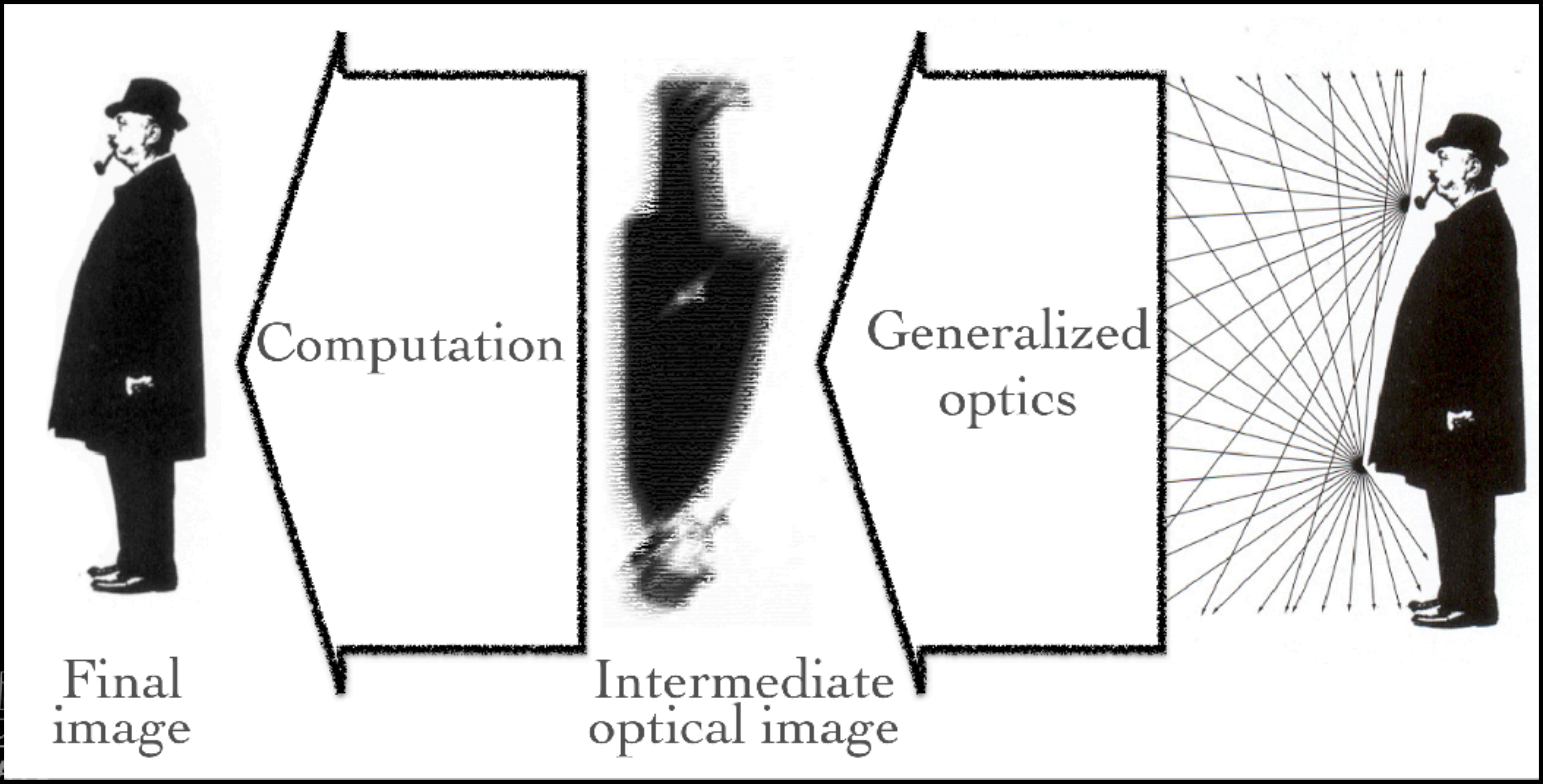


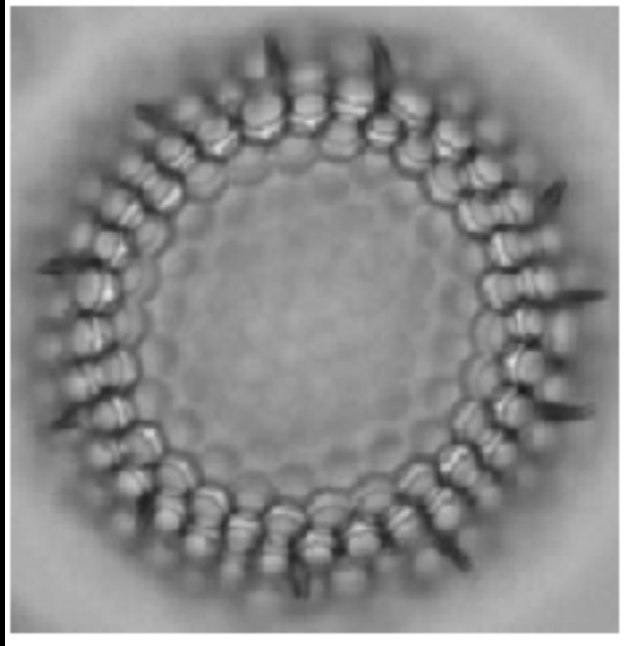
Coded imaging optimizes information gathering

Optics encodes, computation decodes

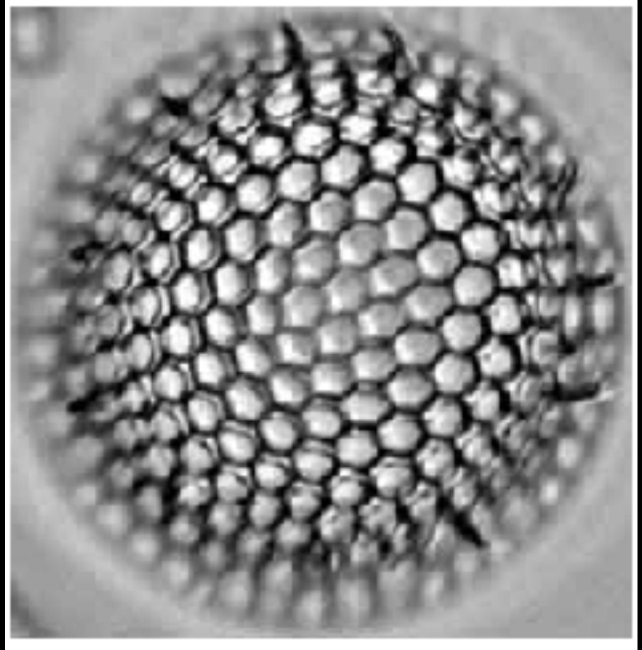
Two goals: better information transmission, flexibility after shooting

e.g. coded aperture, wavefront coding, blur avoidance





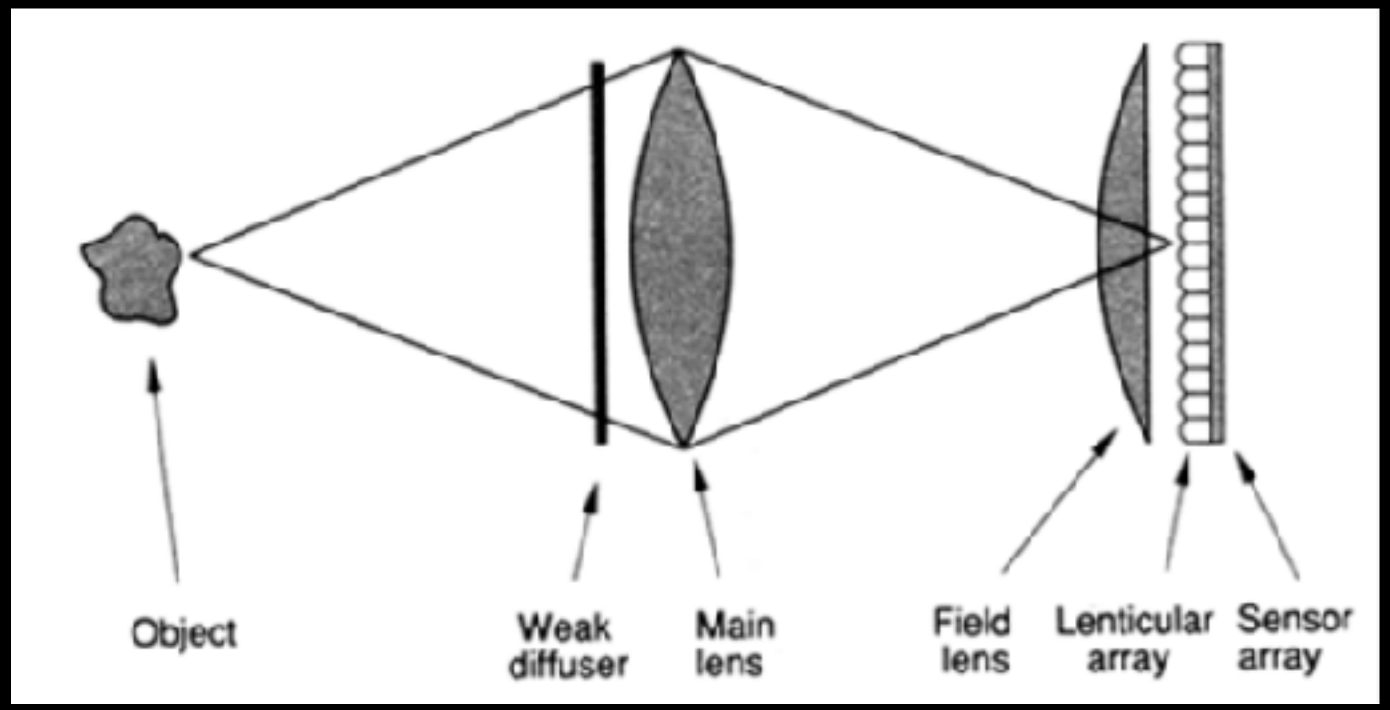
Single cell algae imaged without wavefront coding.

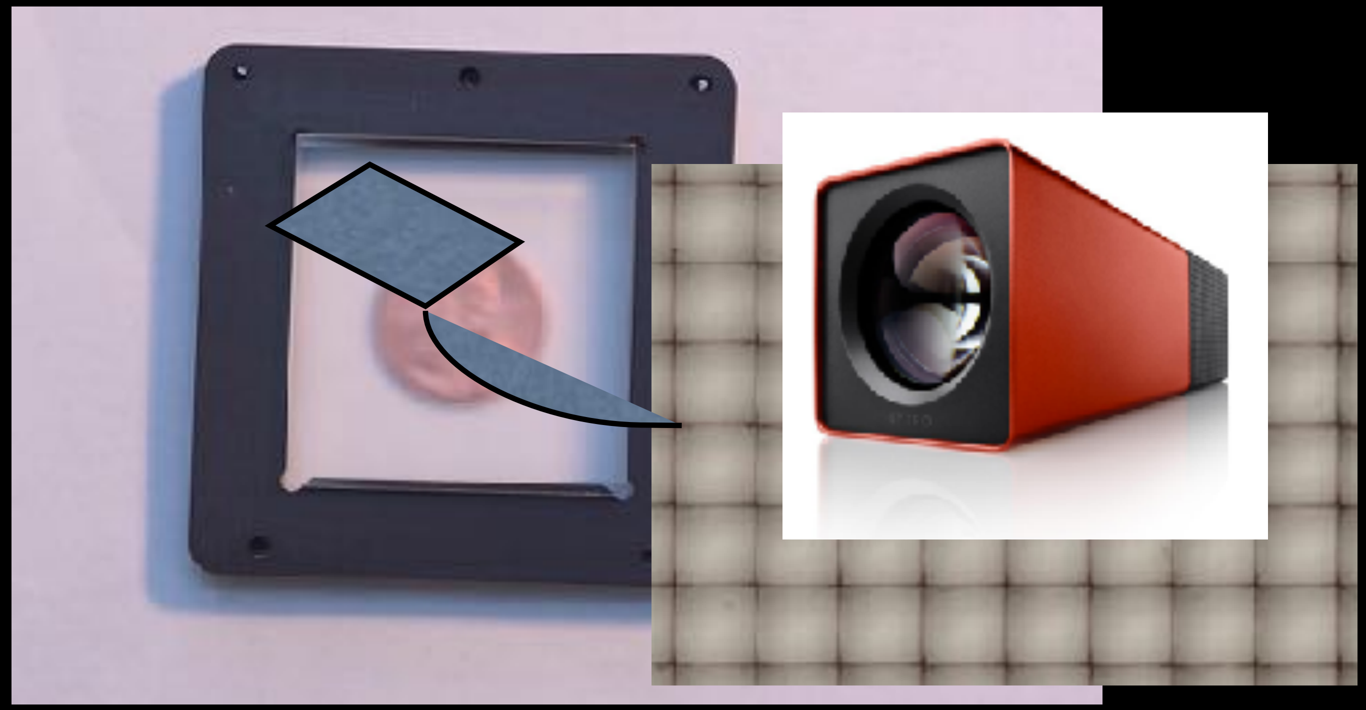


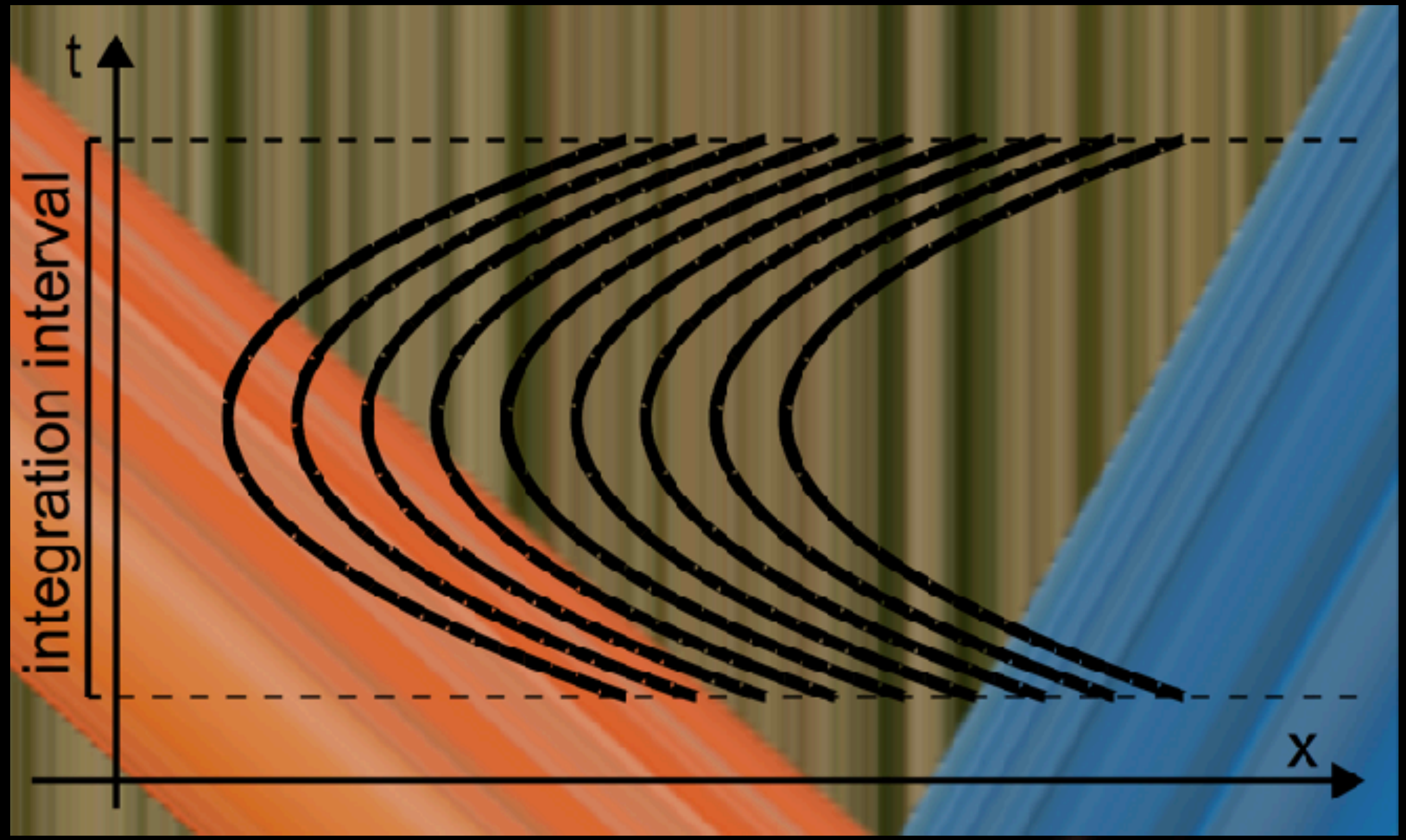
Single-cell algae imaged with wavefront coding.

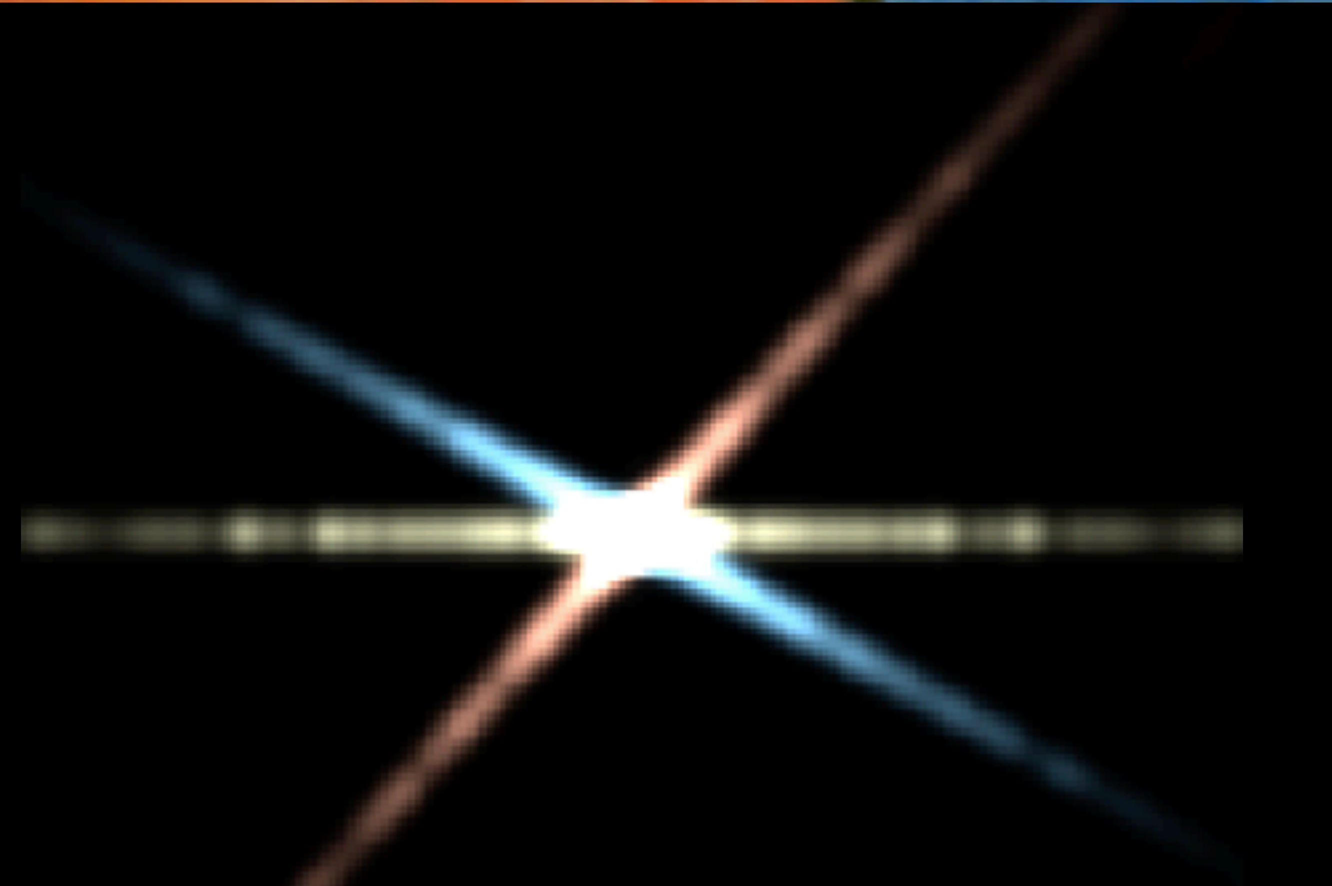
The real raw is high dimensional

Light fields, space time, + Fourier for analysis



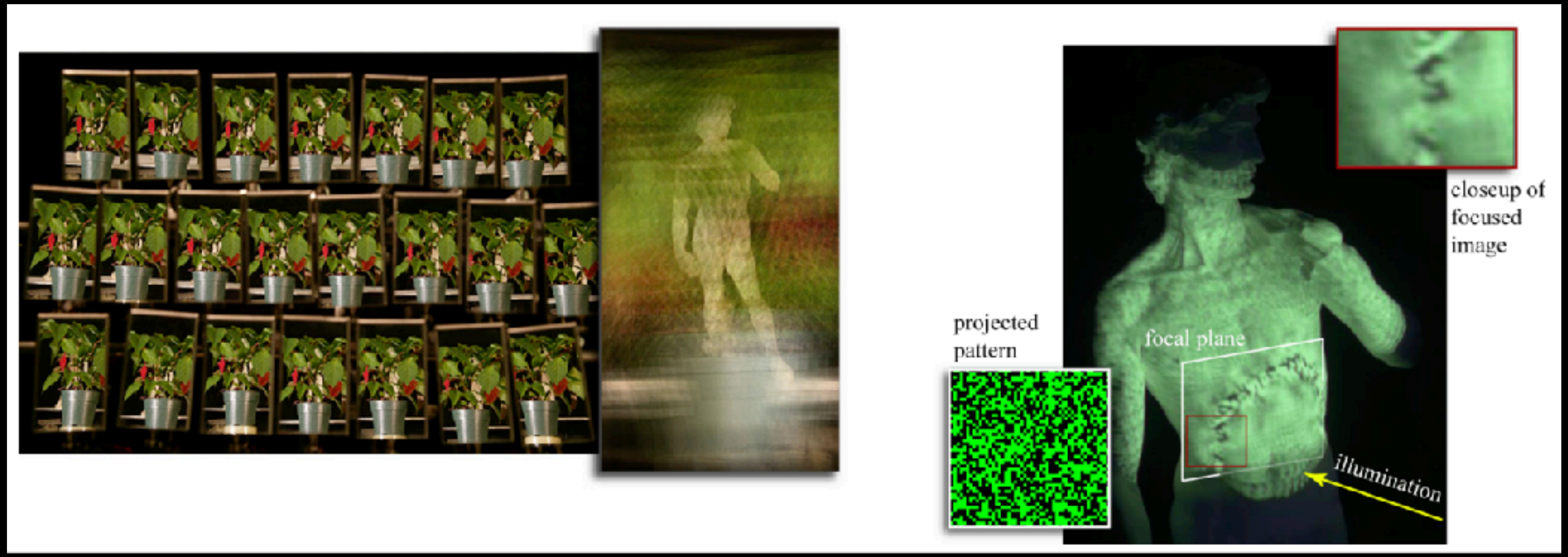






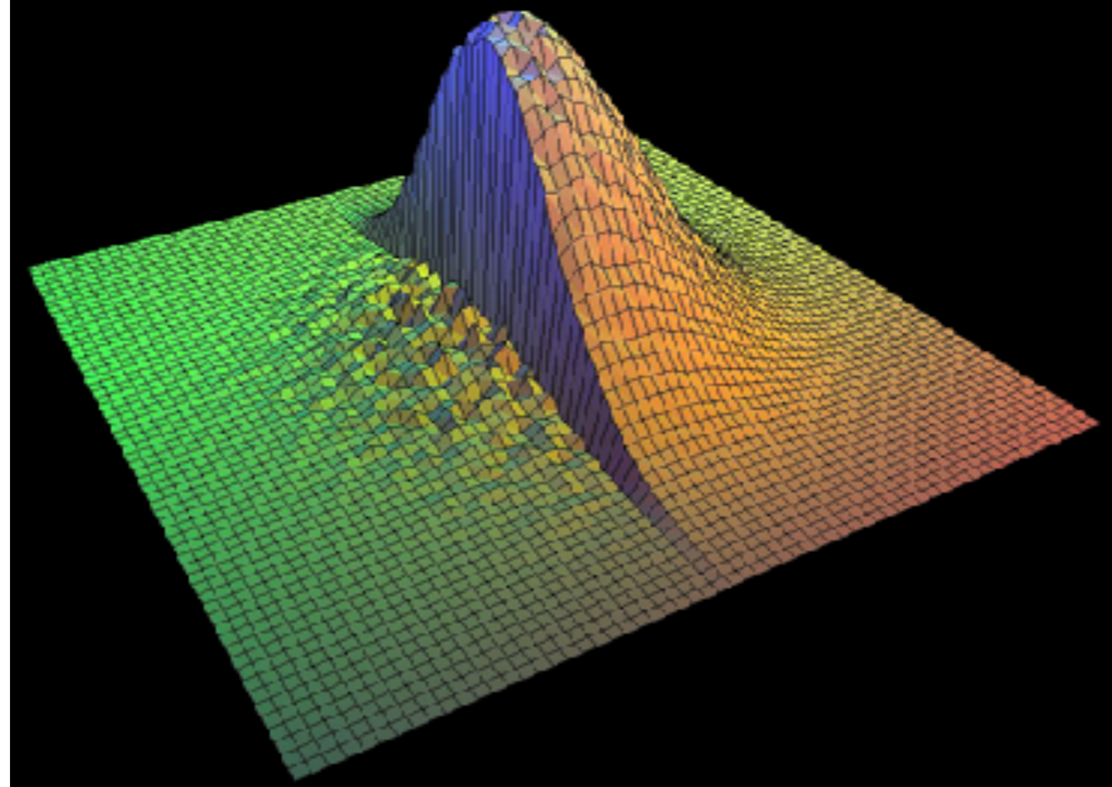
Photography means writing with light

Flash-no-flash, dual photography, confocal light field, light stages



Edges should be treated carefully

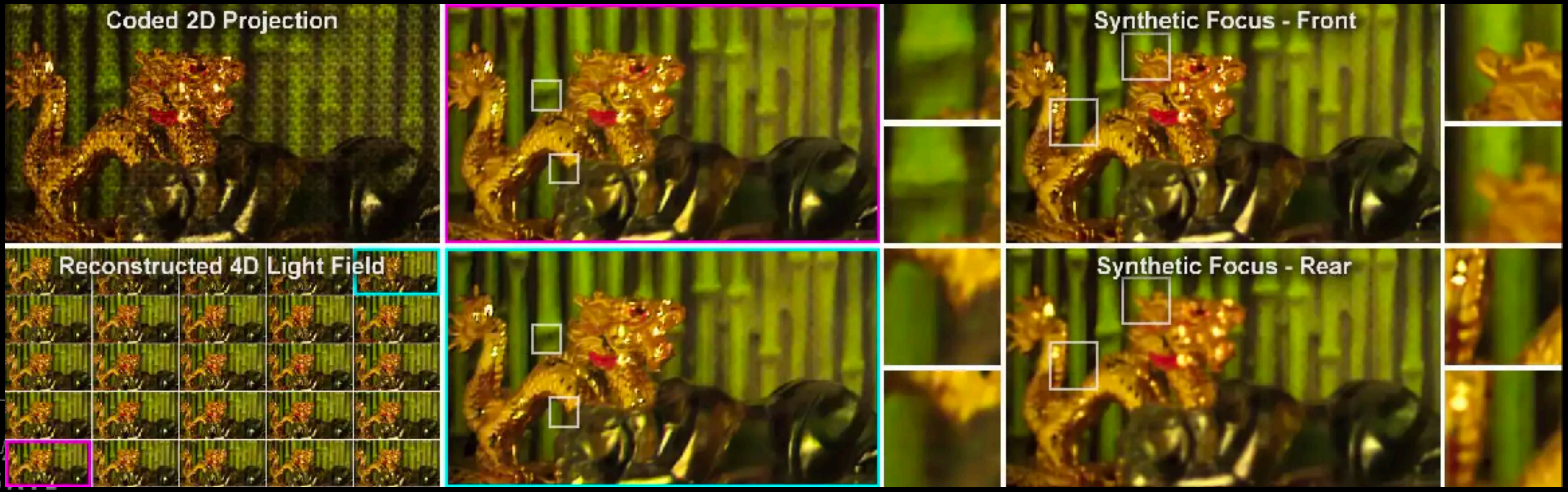
Gradient domain, bilateral filter, PDE, geodesic transform, non-uniform least squares, matting Laplacian



Priors: Tell me something I don’t know

Prior: something we know about the output before we even get the input.

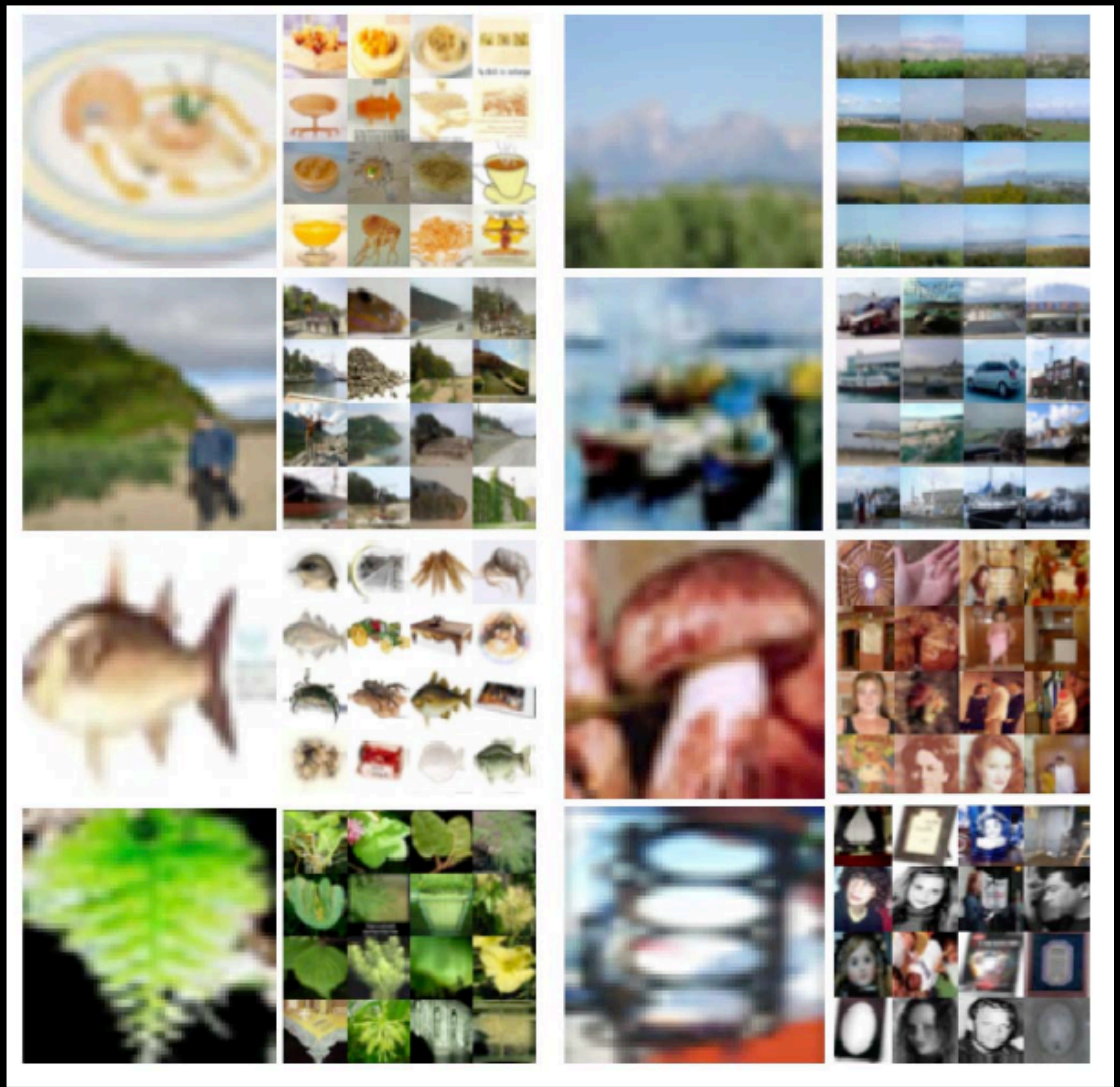
Compressive sensing, deblurring. Powerful with coded imaging.



The embarrassing power of data

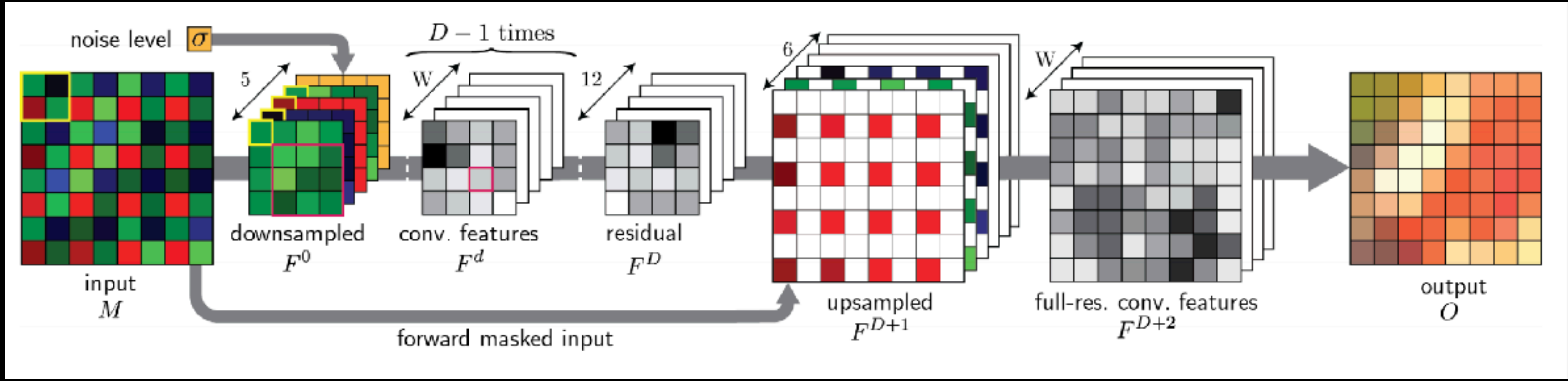
Big collection: Image completion, photo tourism, colorization

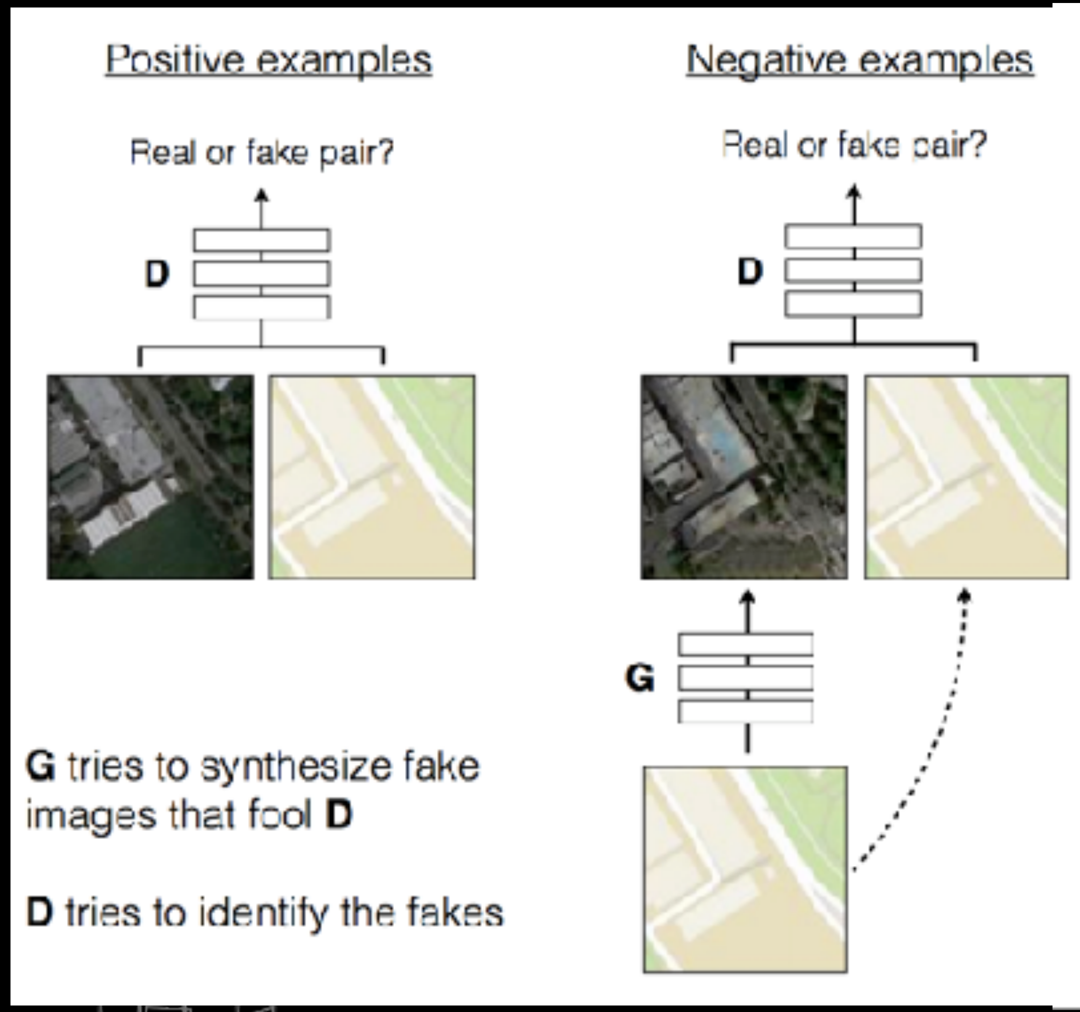
Within an image: patch match, NL-means

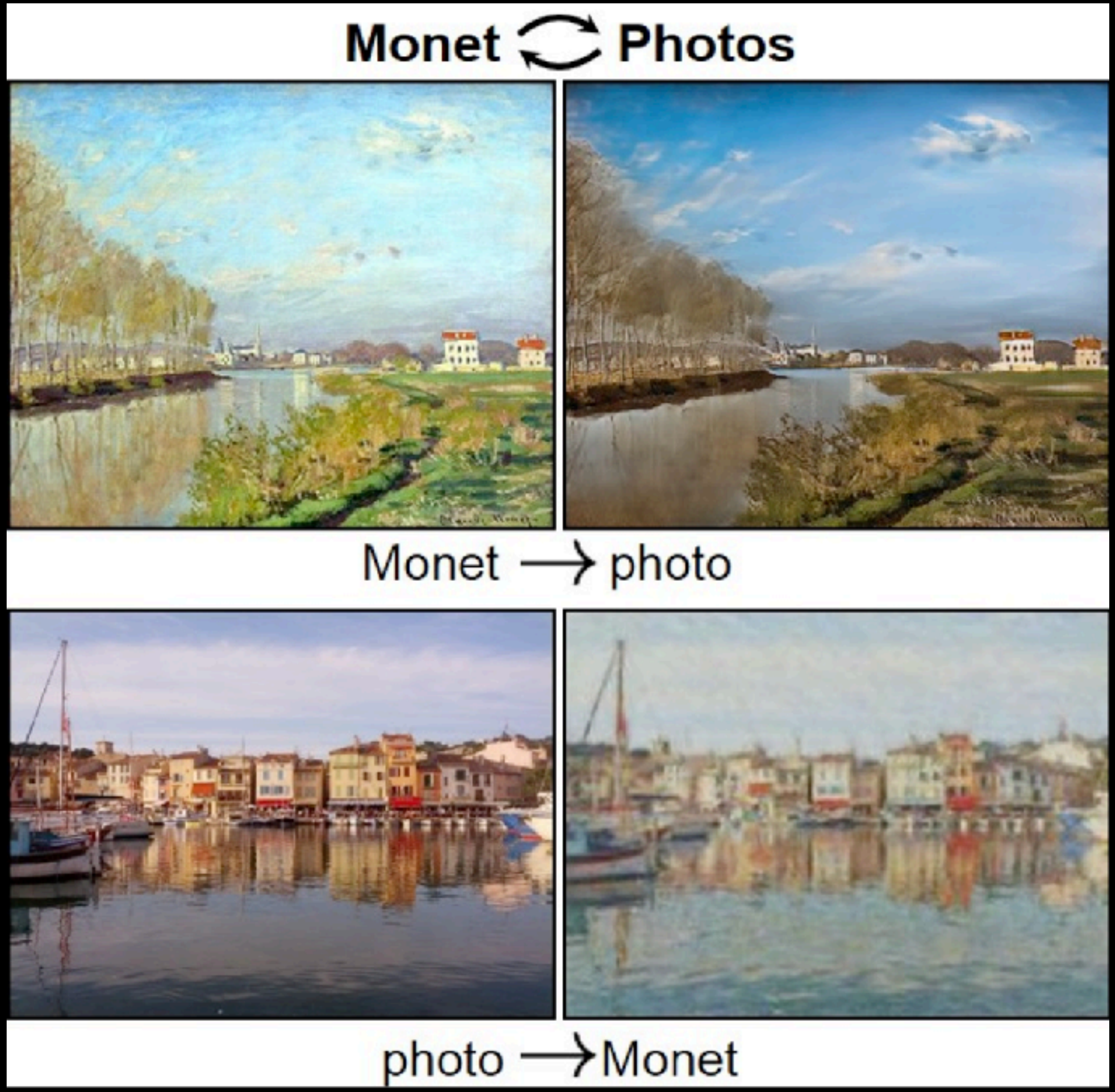


How I learned to stop worrying about overfitting and love deep learning

Regression, Generative Adversarial Networks, creative training





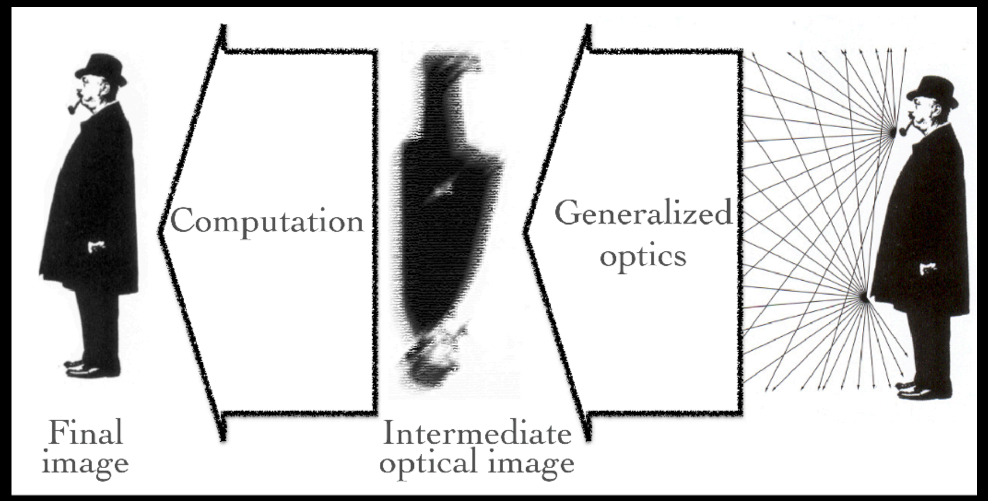


Fredo Durand, MIT CSAIL

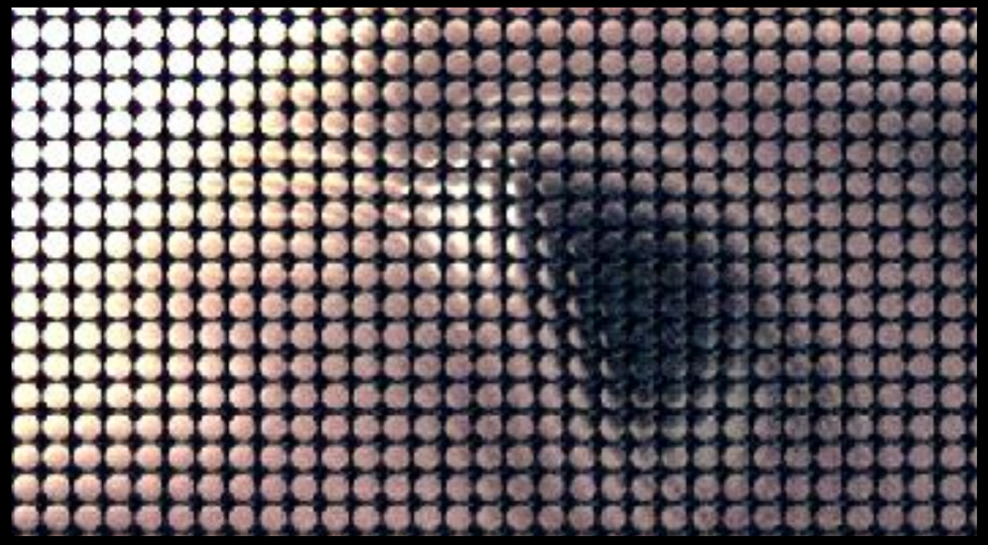
Big ideas recap

Multiple exposure

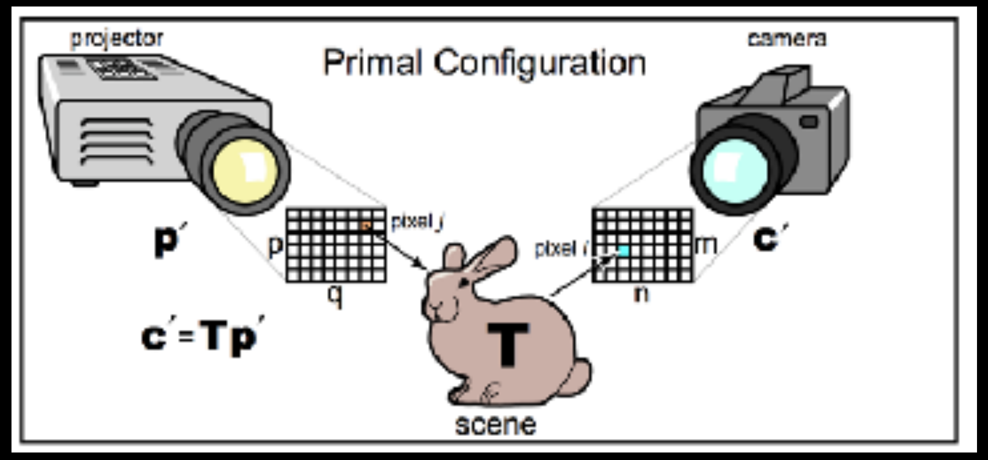
Coded imaging



High-dimensional Raw



Control illumination

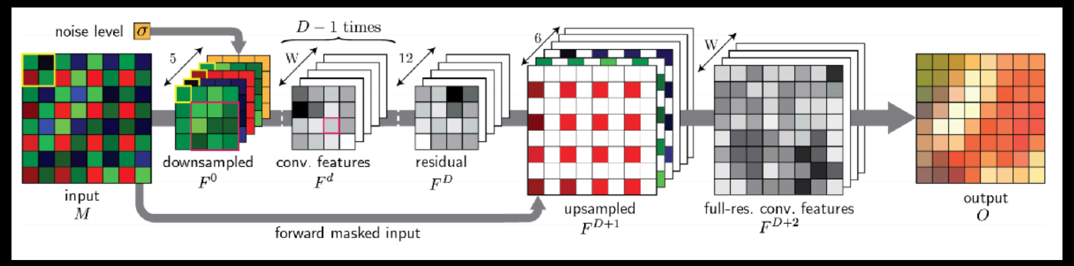


Edges matter

Priors

Big data

Deep learning



Fredo Durand, MIT CSAIL

Plan for this talk

Big Ideas in Computational Photography

Assessment of the field

Research suggestions

Computational photography has great successes

Basic camera pipeline
(demosaicking, denoising, compression)

Panorama stitching

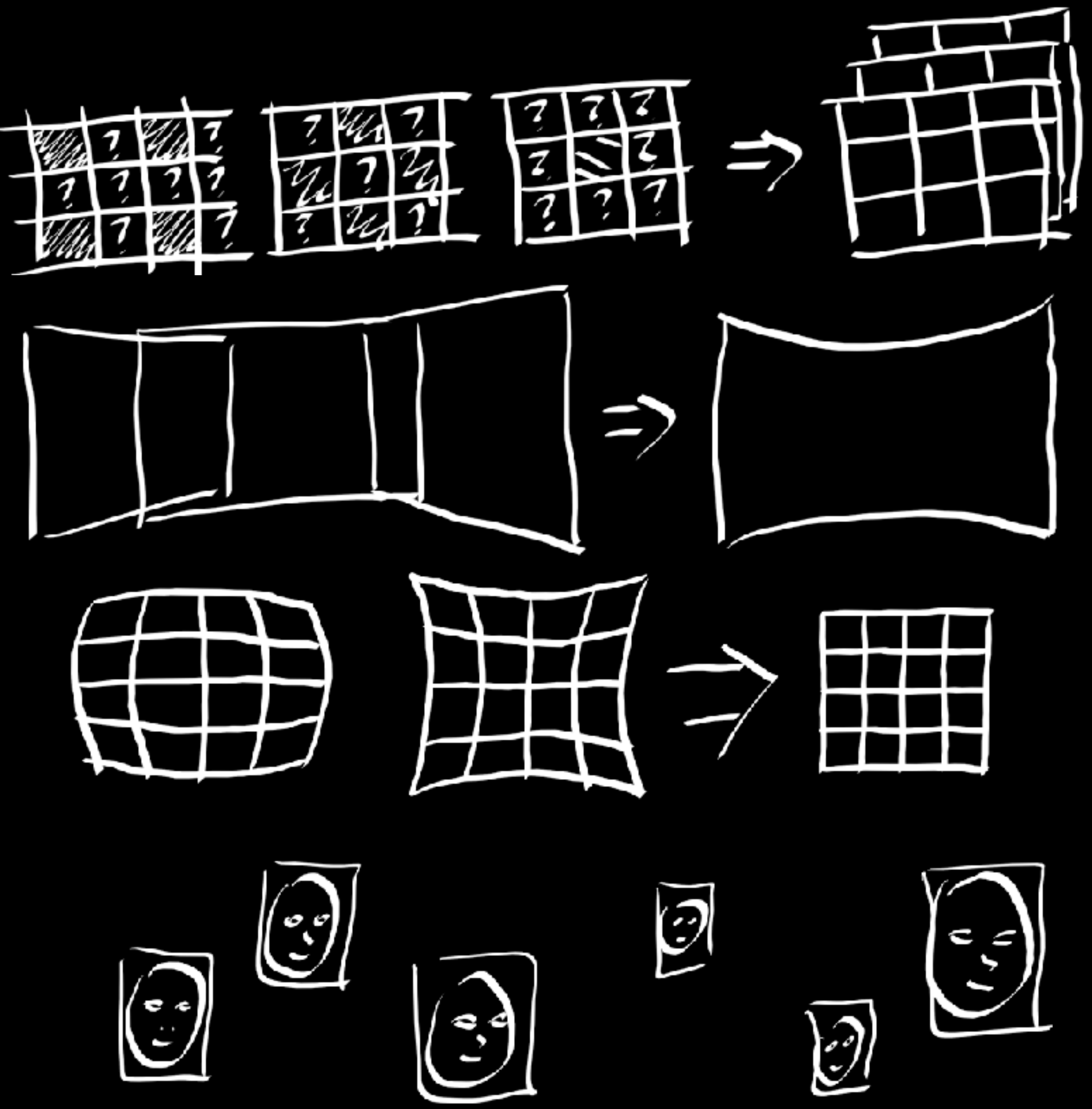
HDR, tone mapping

Correction of optical aberrations

Face detection & recognition

Video stabilization

Gradient editing, patch match, warps



But some ideas have remained mostly academic

Computational optics

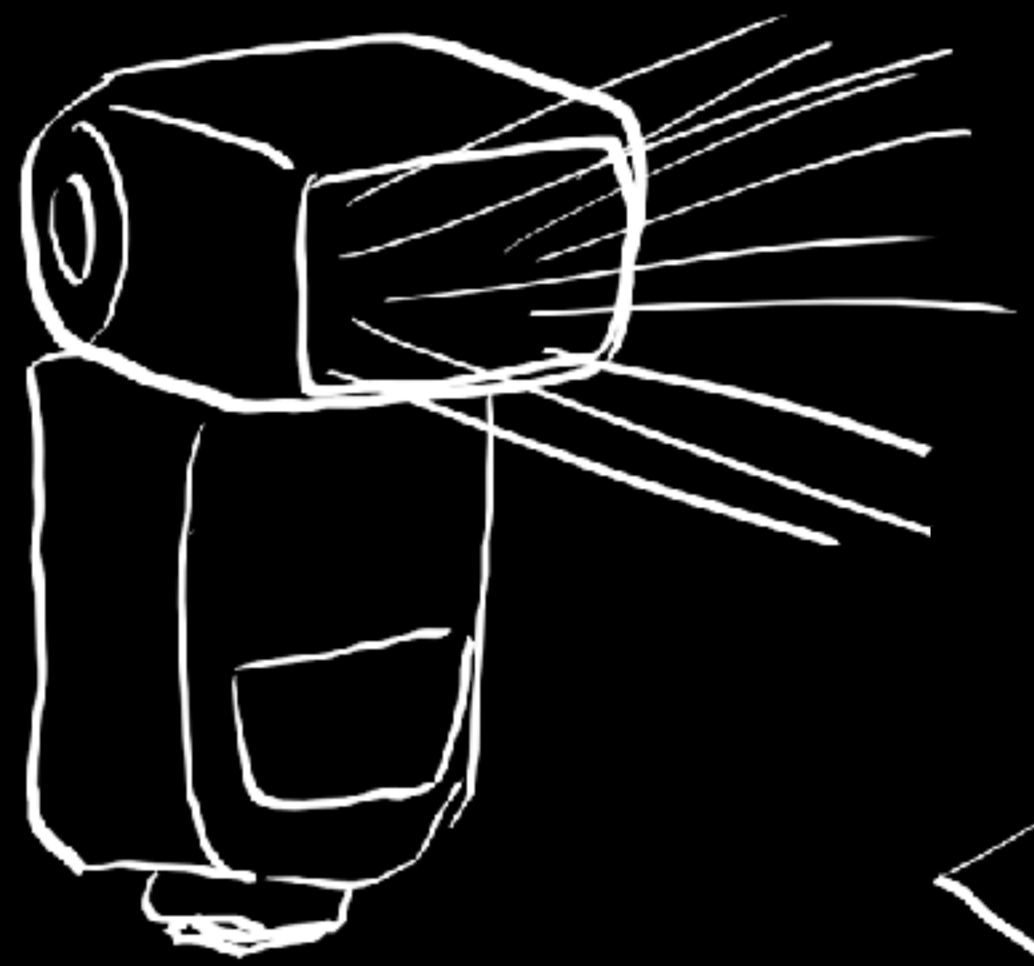
Flash no-flash, multi-flash,

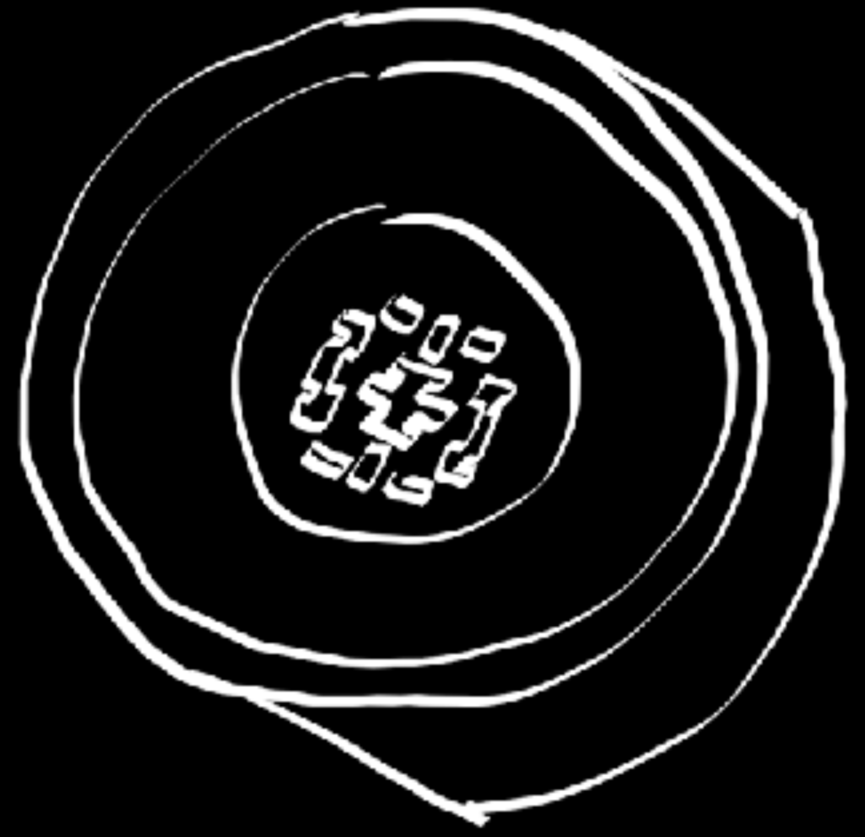
Light fields

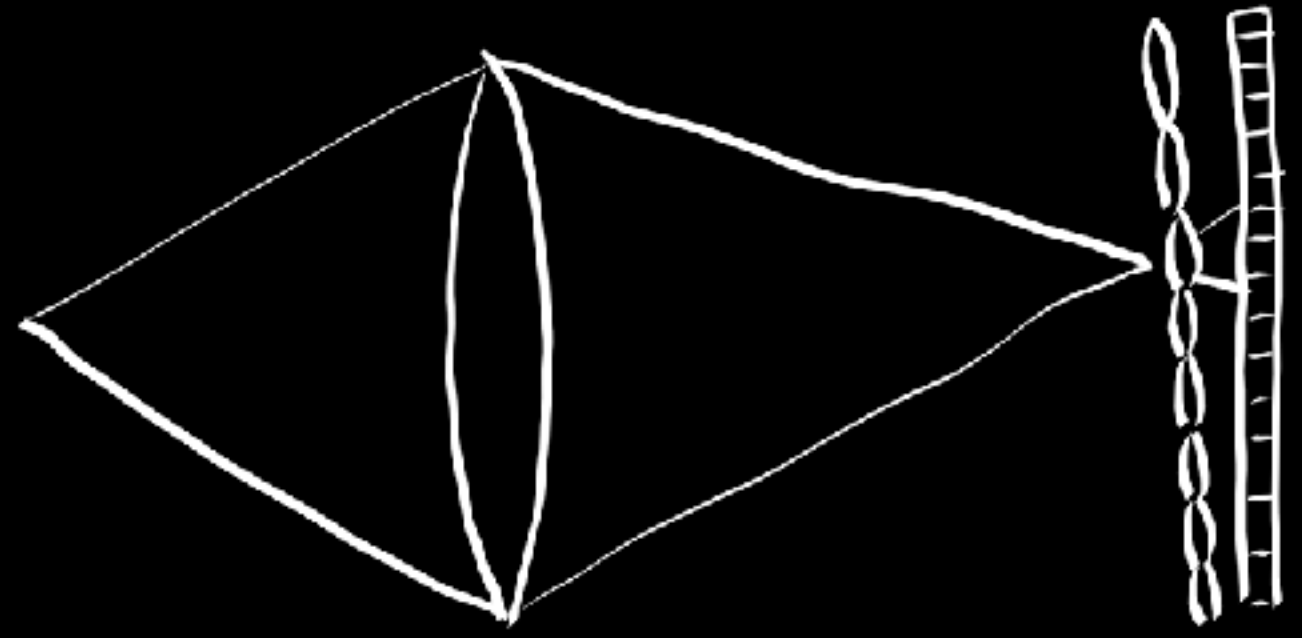
Deblurring

Superresolution

Most work on image collections







Limited impact doesn't mean limited idea

Some ideas have not been refined enough yet

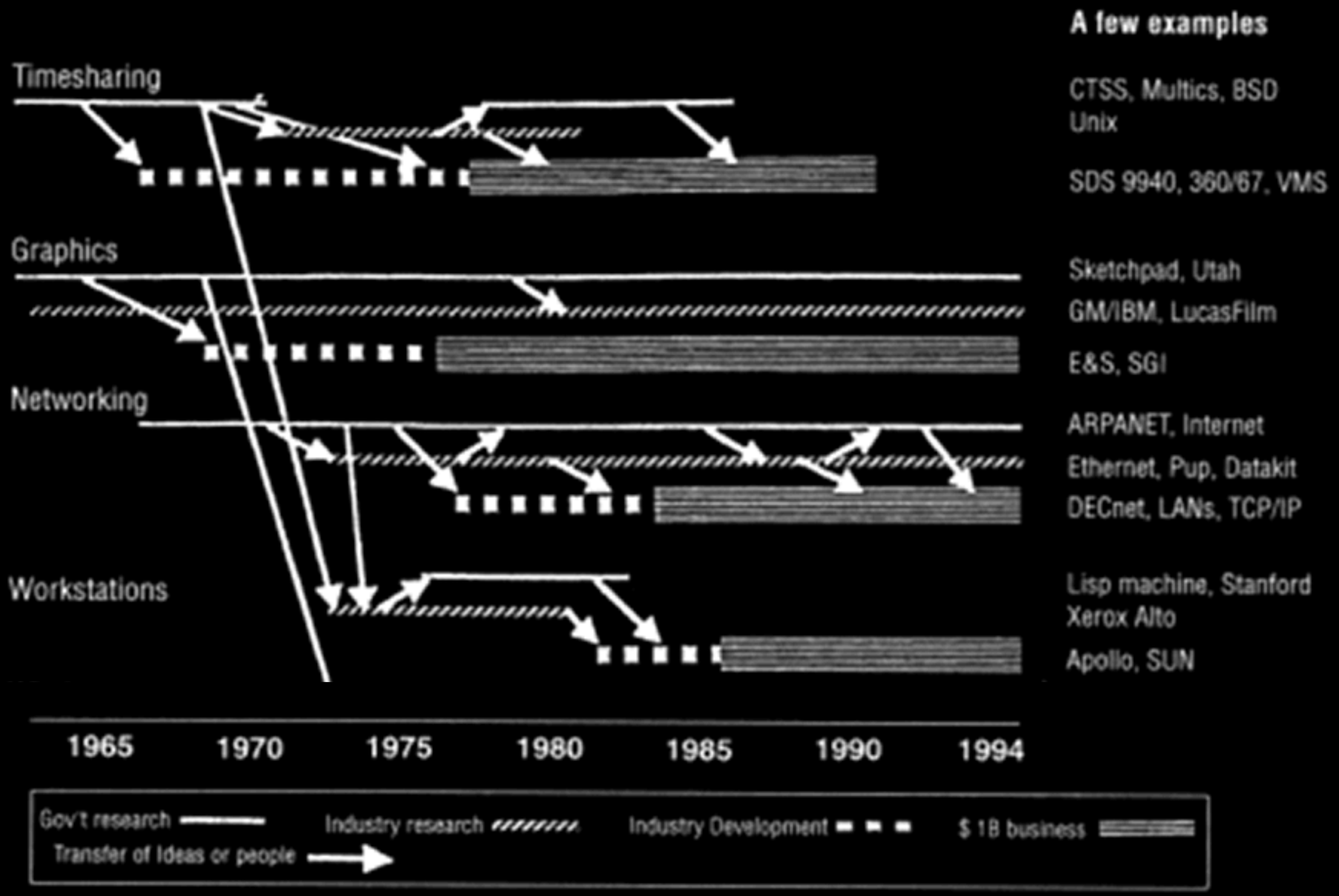
Some good ideas may be applied to the wrong application

We should keep plowing forward and solve engineering challenges

Always assess what the biggest obstacle is and tackle it

Be open-minded to new applications

From idea to market takes time



Fredo Durand, MIT CSAIL

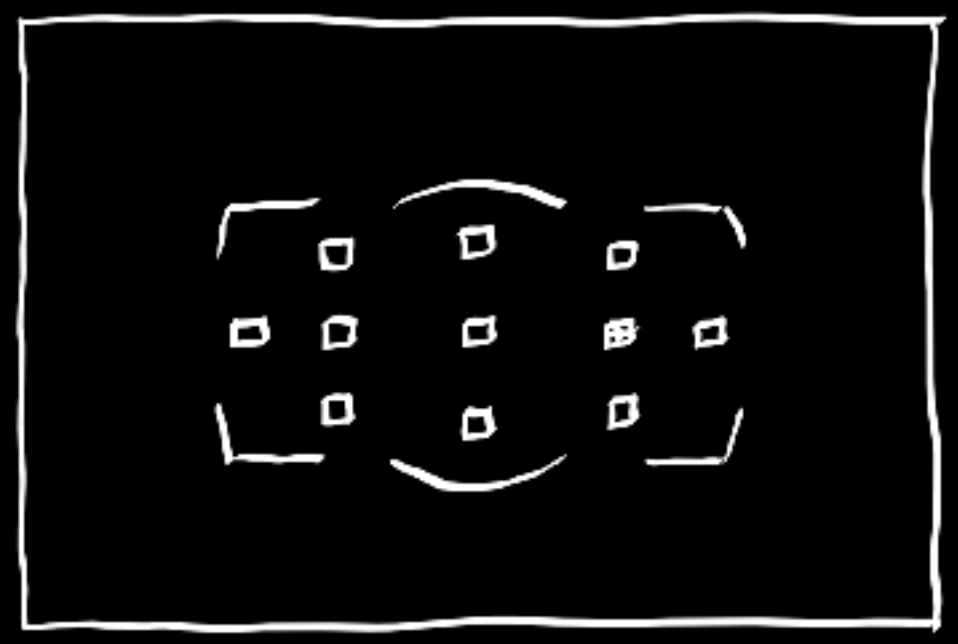
from : Funding a Revolution: Government Support for Computing Research (1999)

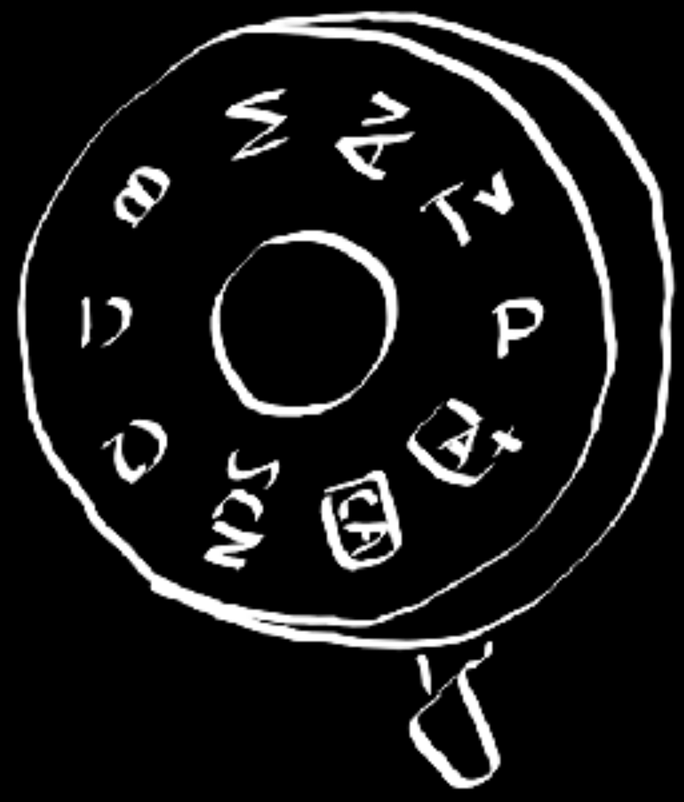
Some basic camera functions are still open problems

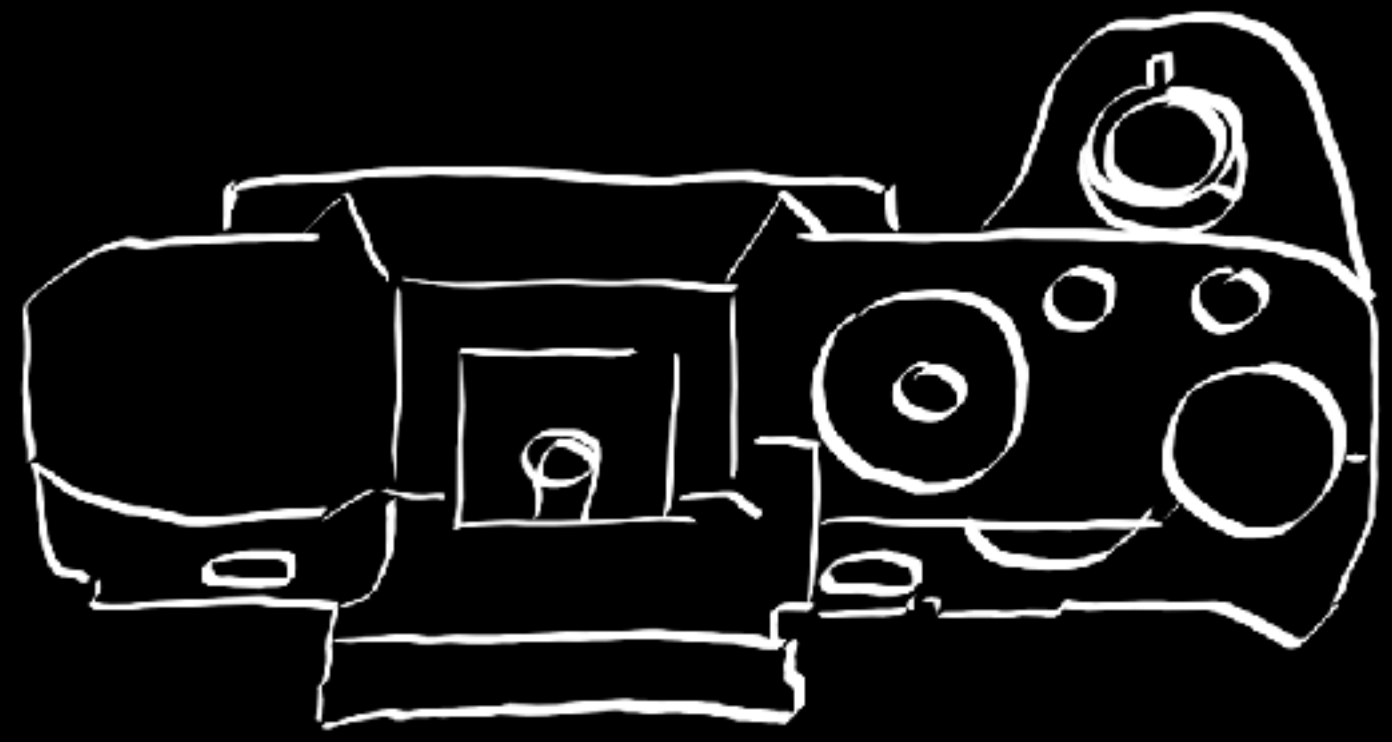
White balance (especially with mixed lighting)

Focusing , auto-exposure (especially for video)

User Interface







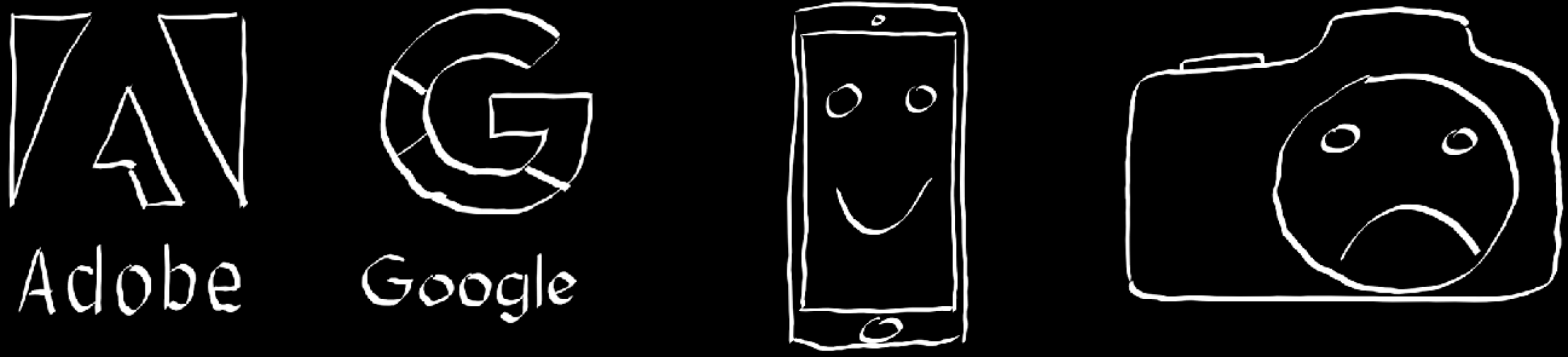
We do not have enough synergy with industry

Adobe and Google do great tech transfer

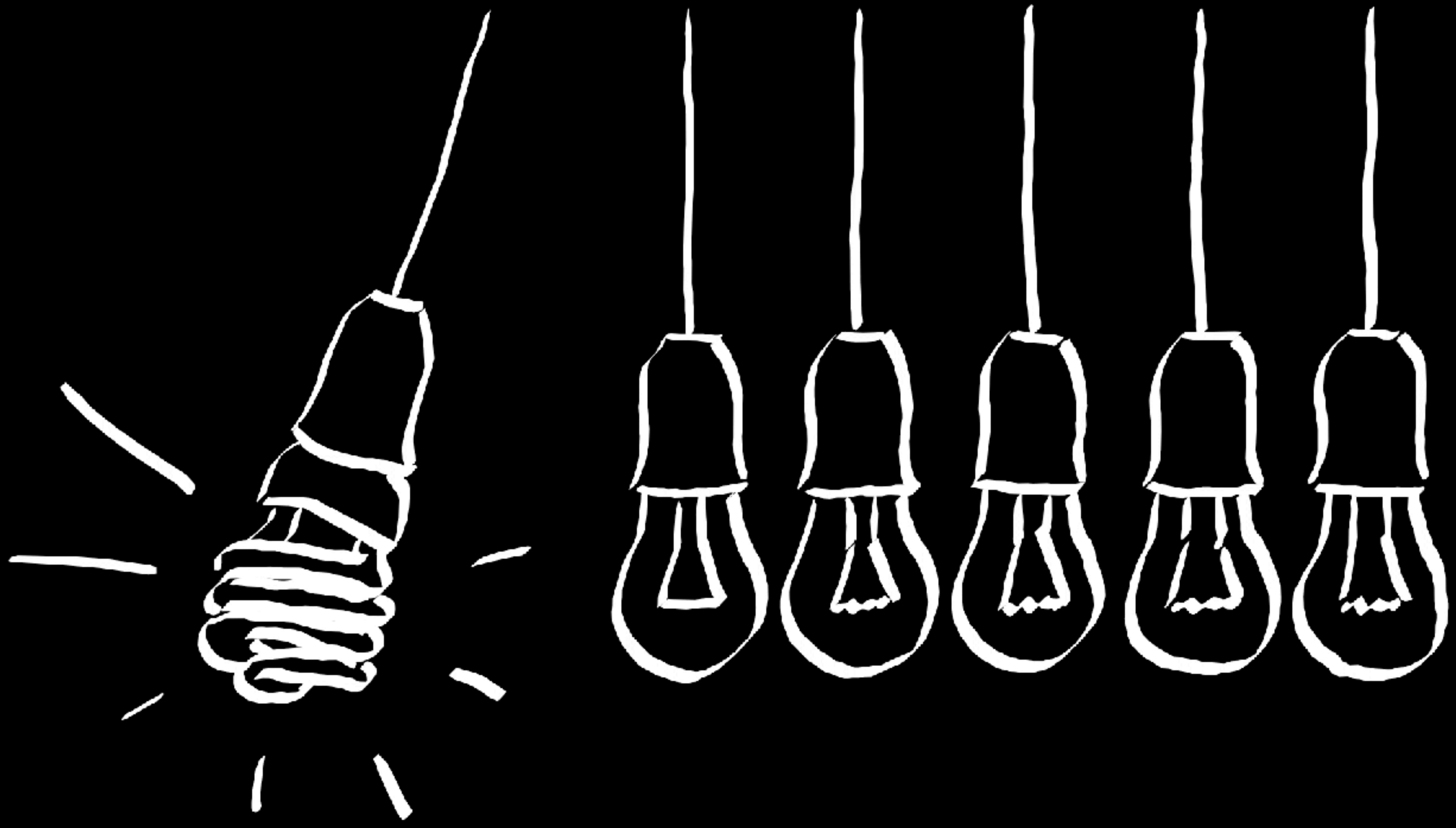
Other software companies do great research but transfer is not as strong

Cell phones are an open platforms for computational photography

Camera manufacturers are not thinking computationally



Can we sustain and improve our impact?



After https://researchimpactnetwork.wordpress.com/

Plan for this talk

Big Ideas in Computational Photography

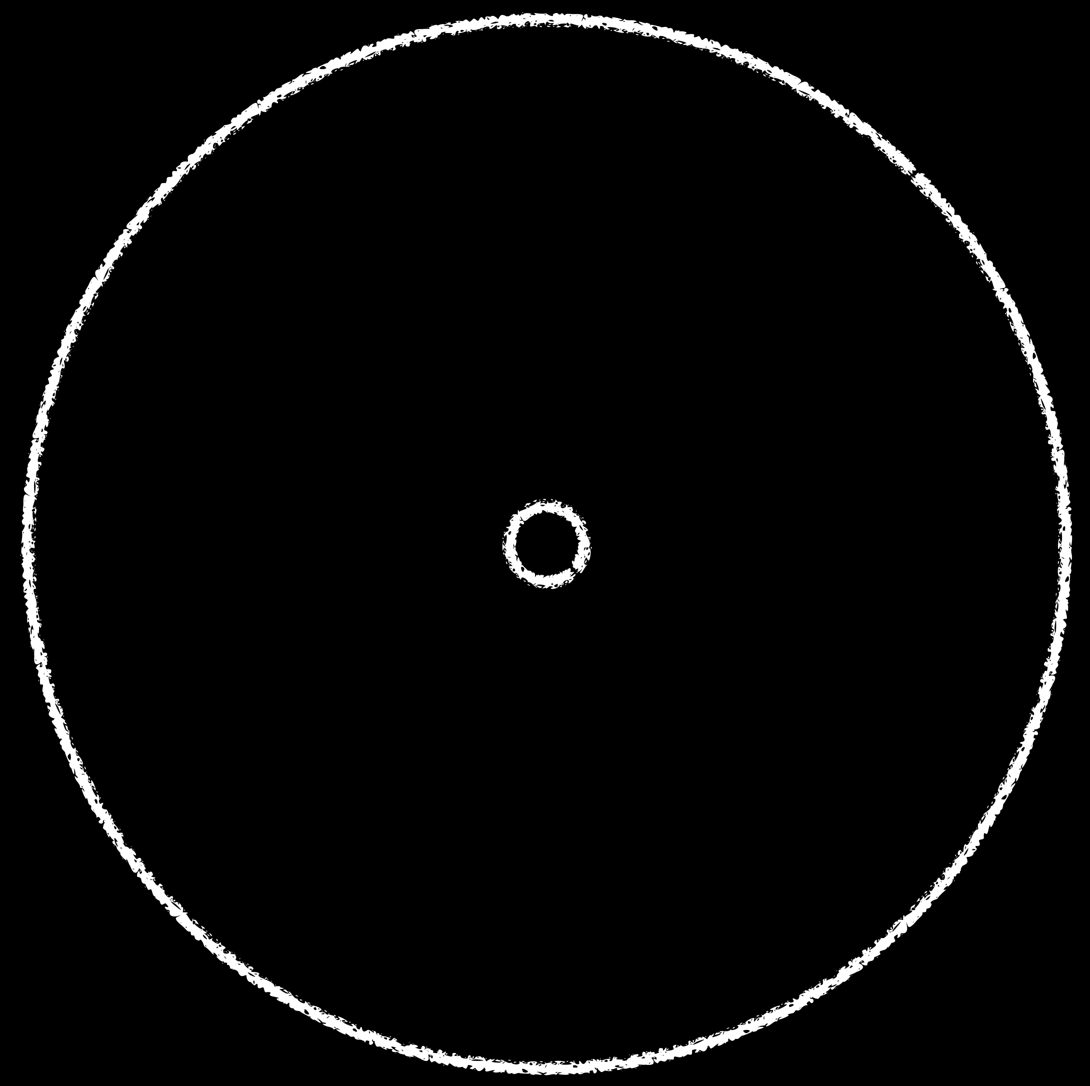
Assessment of the field

Research suggestions

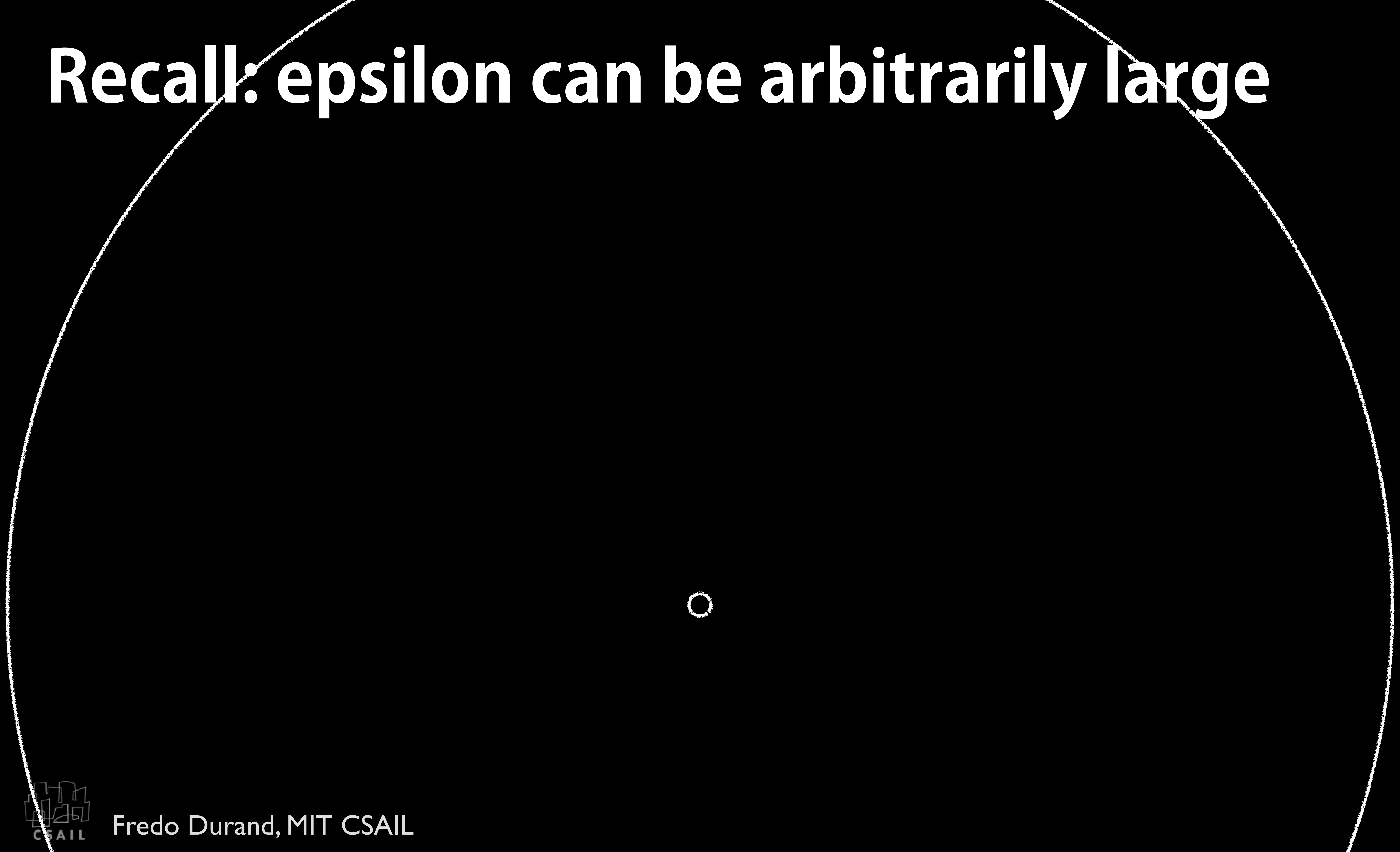
How to use research suggestions

The strong student starts doing what the advisor has asked,
sees that it doesn't work,
looks around within some epsilon ball of the original proposal
to find what does work.

told by Yair Weiss



Recall: epsilon can be arbitrarily large



Mobile devices are the main cameras

Mobile devices are the main cameras

Fredo Durand, MIT CSAIL

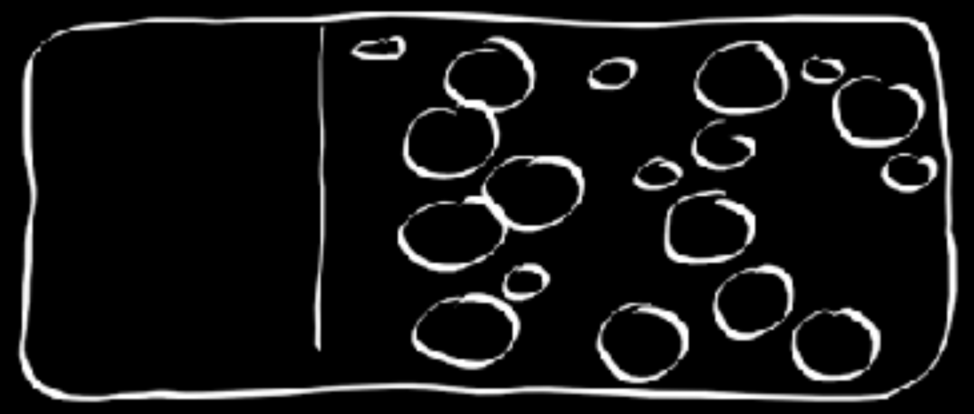
Mobile devices are the main cameras

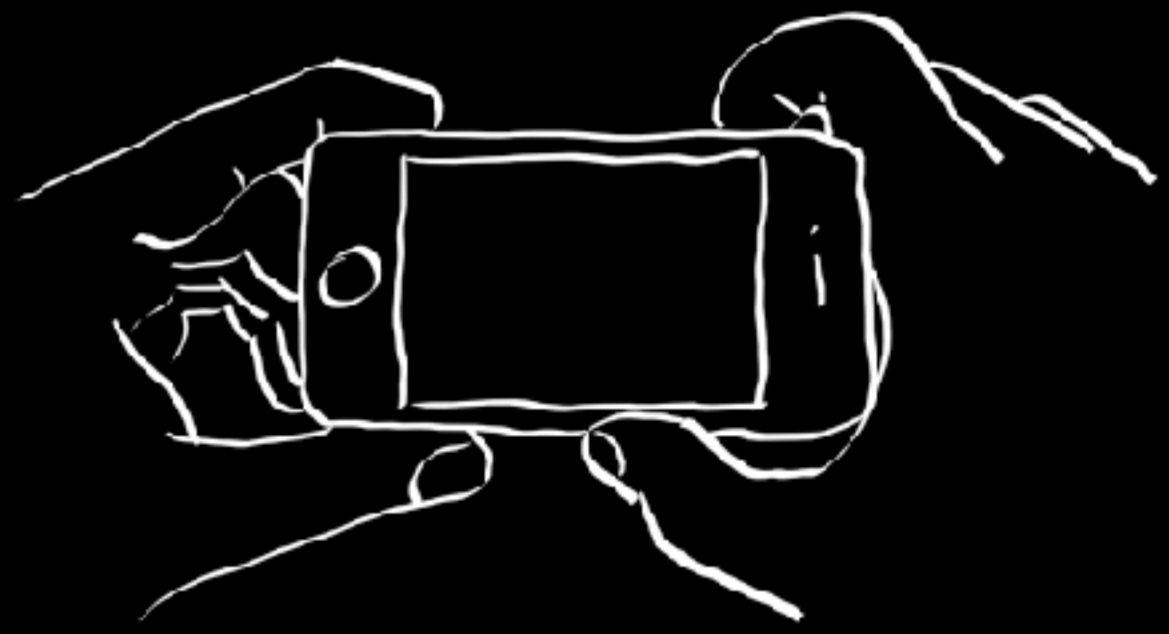
Form factor create harsh constraints (power, optics, computation, HCI)

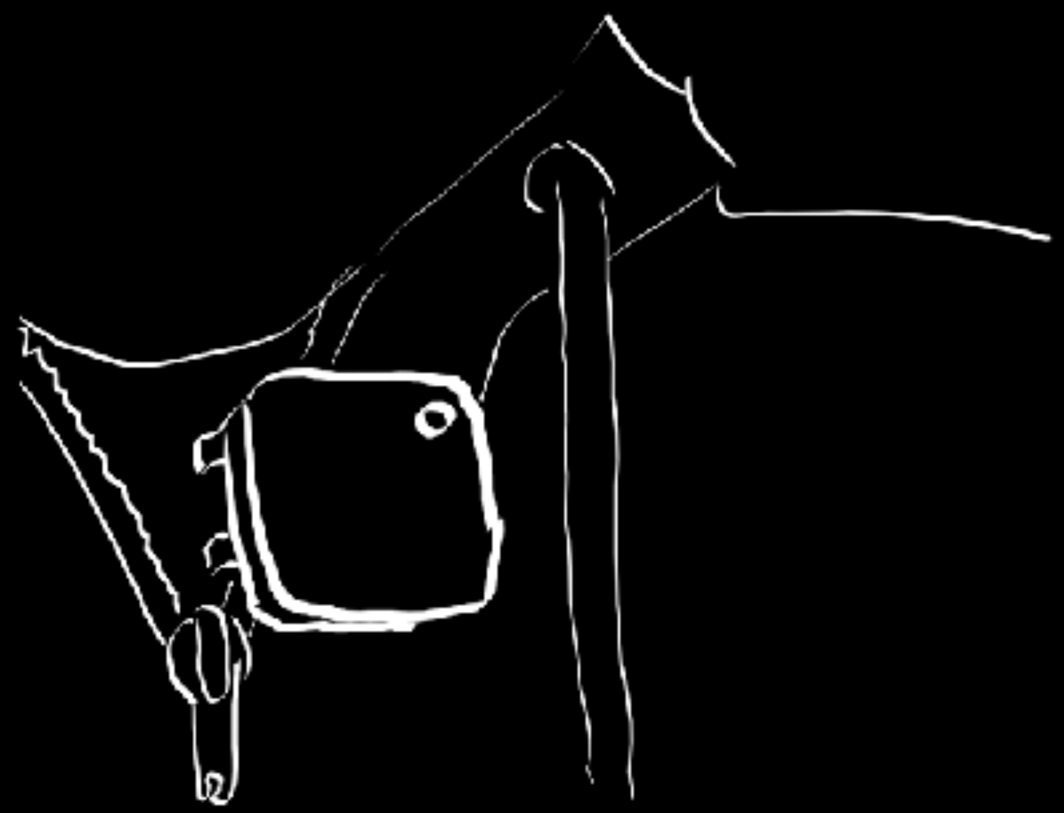
Connectivity offers opportunity and challenges (power, time)

People expect all their data on all devices with similar software

Wearable imaging and life logging compounds these issues







PhD topics

Image enhancement compression for cloud computing

Multiresolution methods for same performance on mobile/desktop

Computational multi-imagers: optics, algorithms, theory

HCI for wearable cameras (speech, eye tracking, etc.)

Low-power systems to decide when a life-logging camera must record

Lots of systems challenges, see later

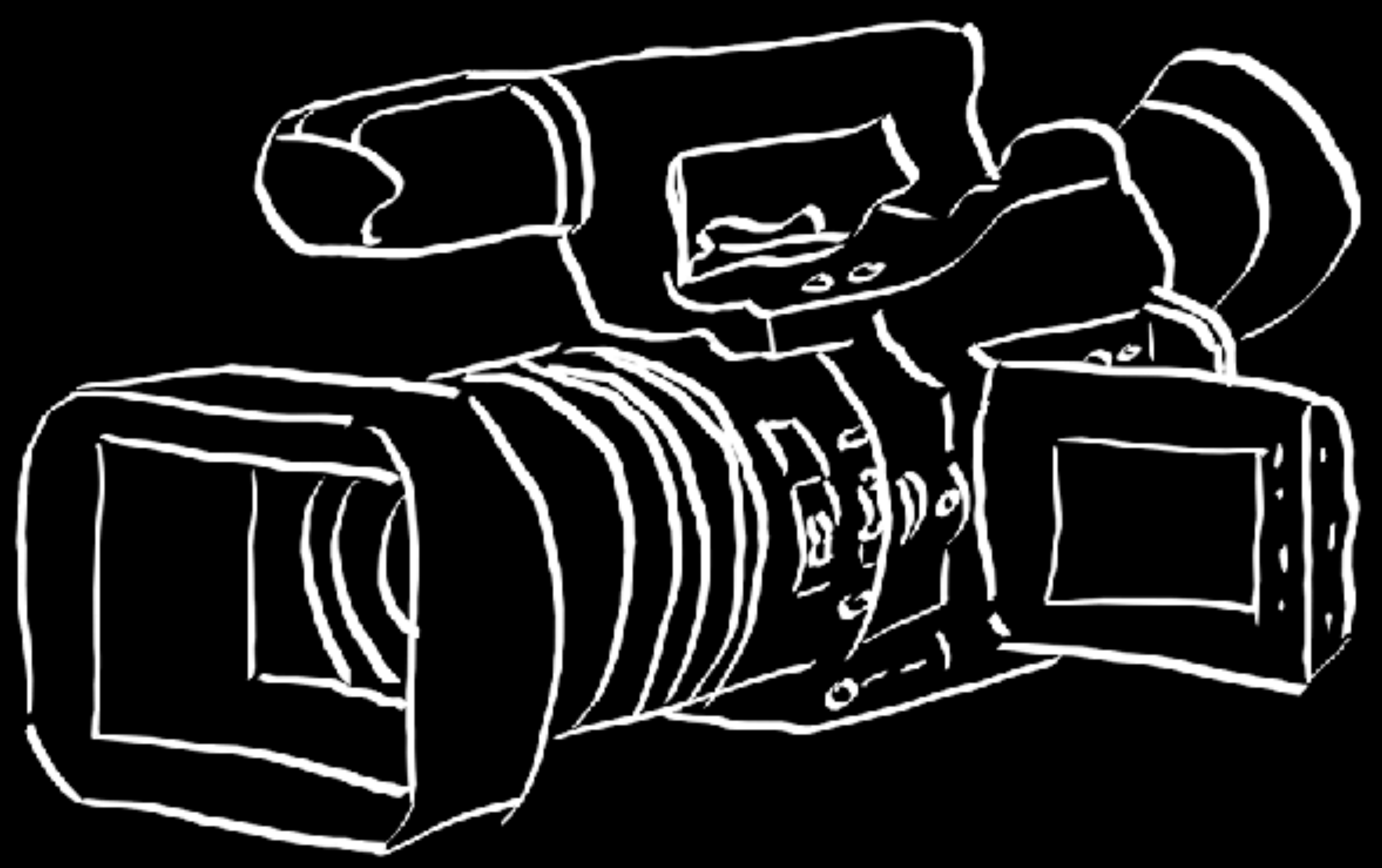
Video is still in its infancy

The gap between amateurs and pros is much larger than for still

Focus, autoexposure are hard and must be temporally coherent

Good and stable framing is a challenge

Editing and selection are painstaking and tools are complex



PhD topics

Hybrid hardware/software video stabilization

Predictive and smart autofocus

Plenoptic cameras to smooth focusing errors after the fact

Robotic second shooter

Editing while shooting via vocal or touch tagging

Auto editing (deep learning? training data?)

Automatic editing for life logging cameras

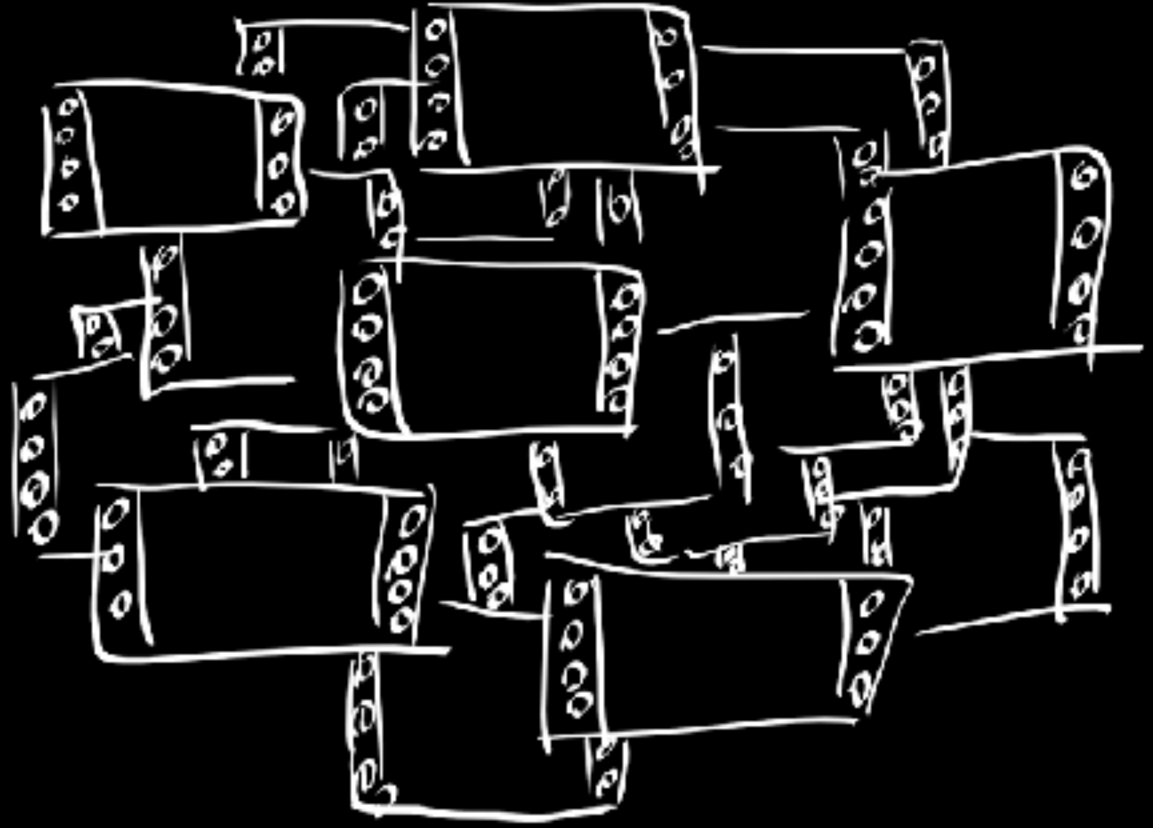
Scale is an opportunity and a challenge

How can we find the needle in gazillion photos and videos?

We need search and selection algorithms, interfaces, visualizations.

How can we get enough data to train deep learning?

We have systems challenges with scale, especially on mobiles and across devices



PhD topics

Automatic selection or rating of personal photos

AI+UI for selection and rating

AI+UI for personal image search : show me the photos where….

Compression for life logging cameras

Storage, retrieval, organization of large collections

Deep learning should be used and extended

We need new extensive data and be creative in using existing one

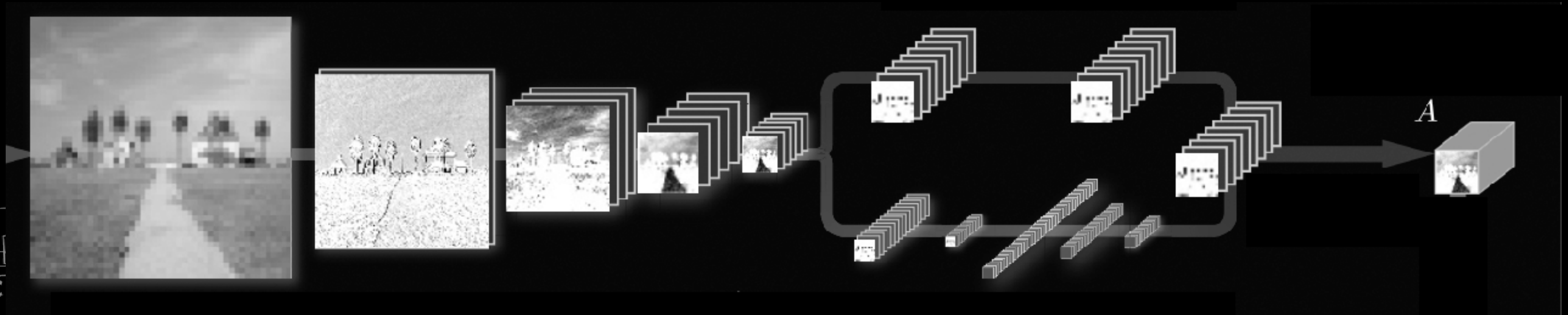
Learn from smaller less labeled datasets

New network architectures must fit problems to better allocate capacity

We must develop creative training strategy (e.g. cycleGAN)

Performance matters, both at training and evaluation

Neural networks should become modular



PhD topics

8 bit learning & automatic performance optimization

Architectures inspired by image processing

Learn or fine tune from very small data

Learn modular image prior (e.g. real NVP)

Create dataset and learn photo style

Optimal transport applied to deep learning

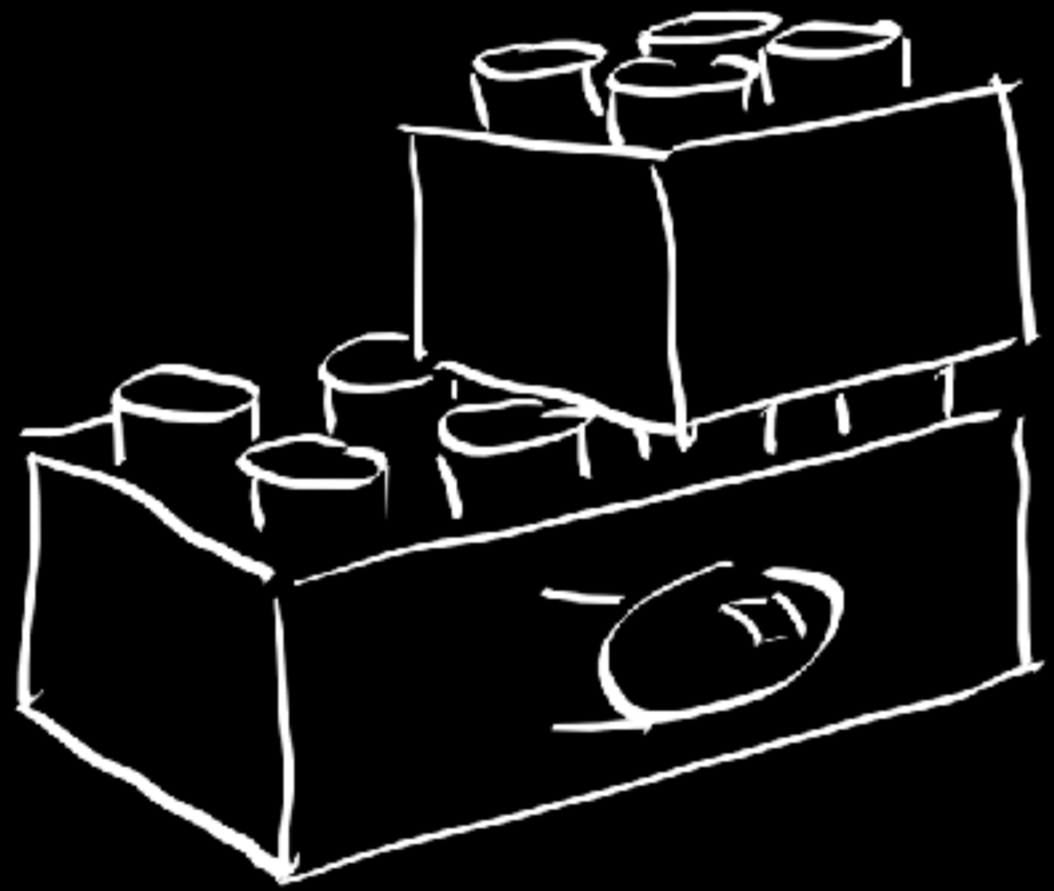
We need more Systems research

Users expect their data everywhere on heterogenous devices

New compilers and hardware are needed for performance

Software and hardware should be modular

What are data structures for metadata and new types of data?



PhD topic

Google doc photoshop

More Halide (variable datarate, better autotune, numerical algorithms, auto LOD/preview)

Metadata

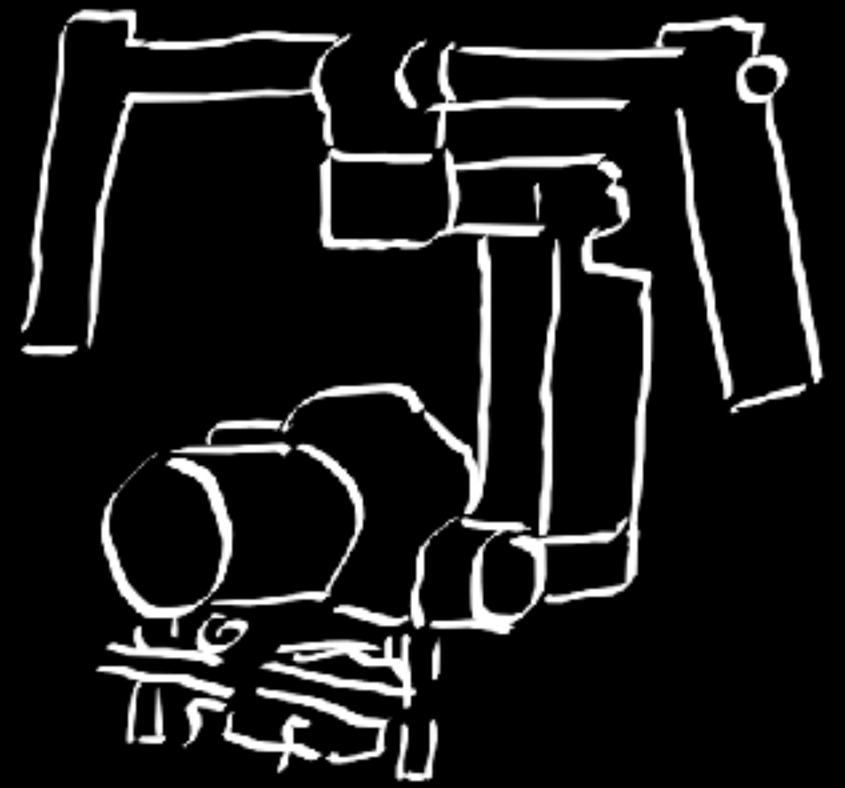
OpenGL comp photo

Modular OS for cameras

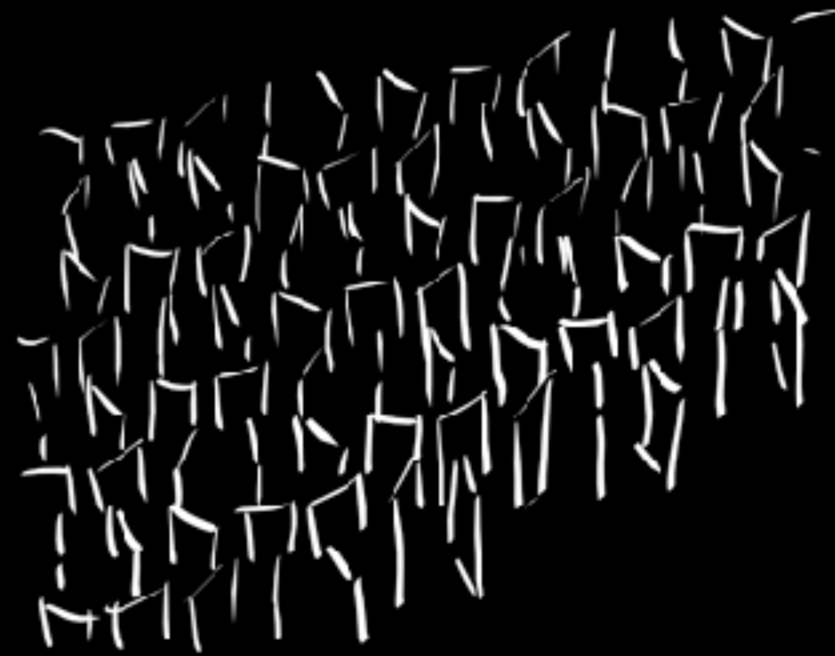
We need more research on hardware

Sensors should evolve for light fields, depth, other information, and on-board processing

Actuators and drones can move or stabilize cameras and lights



Computational fabrication and metamaterials can customize optics, sensors and lighting



PhD topics

Light field sensor with built-in compression

On-sensor computational stabilization

Leverage other sensors: gyro, accelerometer, nIR

Drone cinematography

Robotic secondary shooter

Smart gimbal

Inverted light stage with orientable flash

Custom waveguides

Metalenses for depth of fiels optimization

Metalens optimization technique

We need more research on human factors

New UI is needed for traditional and computational photography

Visual perception must be studied and incorporated

We need better image metrics and priors

PhD topic

UI studies of photography practices, goals

Modular UI

HCI for computational cameras, e.g. light field

Versatile visual difference predictor

Perception of style

We need more theory

We must derive fundamental limits and scaling laws

We should unify time of flight and linear imaging

Light field and computational illumination need more theory

$$SNR(\Phi^2) = \frac{\Phi^2}{var(\Phi)} = \sum_{ij} m_{ij} \frac{\Phi^2 t_i^2 [I_{ij} < I_{max}]}{\Phi t_i + \sigma_r^2 + \sigma_{adc}^2 g^2}$$

$$\frac{1}{\alpha^2 F^2} \iint L_F\left(u(1-1/\alpha) + x/\alpha,\ v(1-1/\alpha) + y/\alpha,\ u, v\right) du\, dv$$

$$\int\!\!\int\!\!\int\!\!\int\!\!\int k(x,y,v,v)\, e^{-2i\pi(\omega_x x_0 + \omega_y y_0)}\, dx\, dy\, dv \Big| du \leq A^3$$

PhD topics

DoF bound topic 1: come up with better design

DoF bound topic 2: better bound

Light field aliasing, diffraction, reconstruction, prior

Unify time of flight and radiance

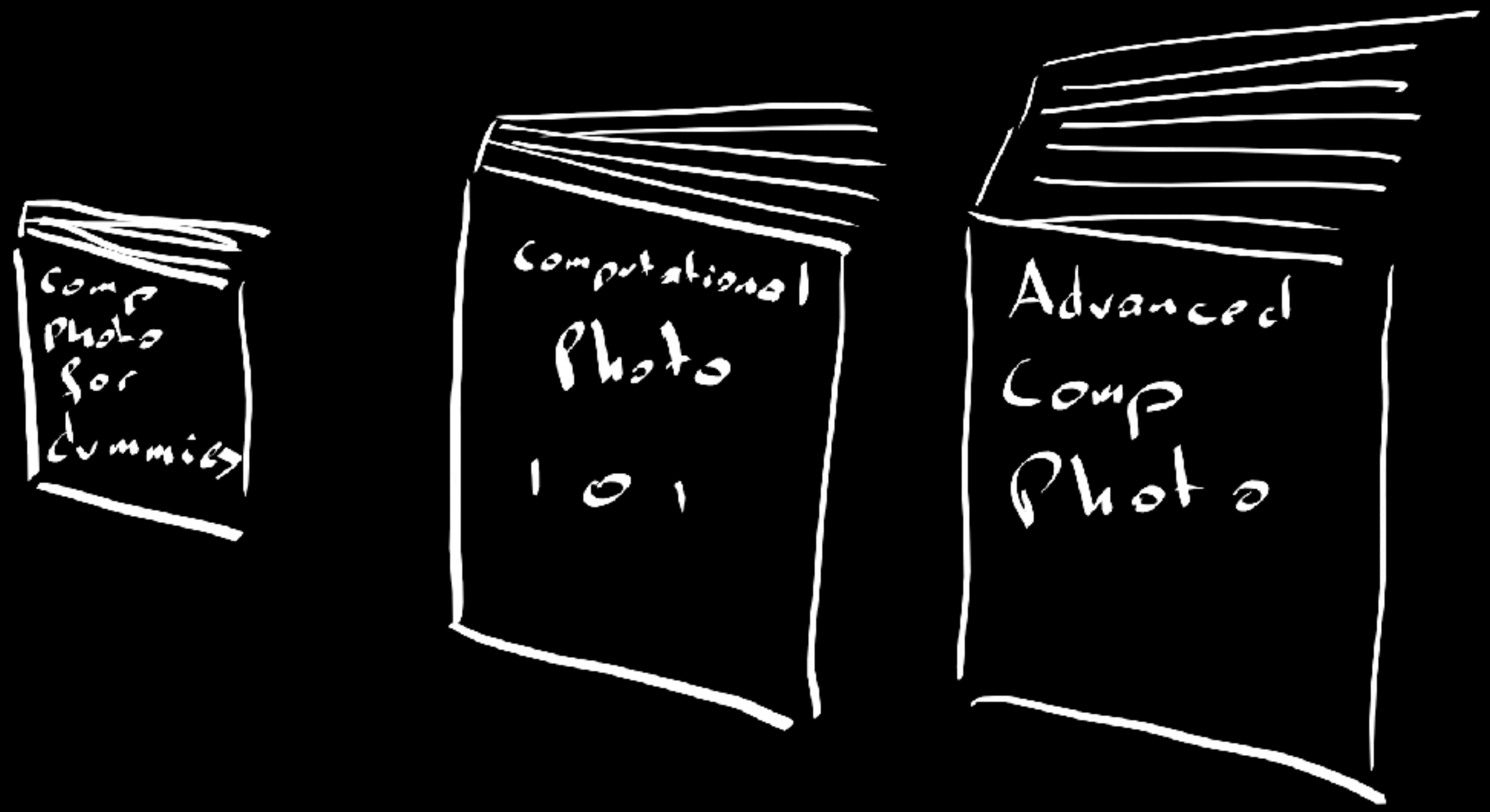
Dual photography

Unification of edge preservation

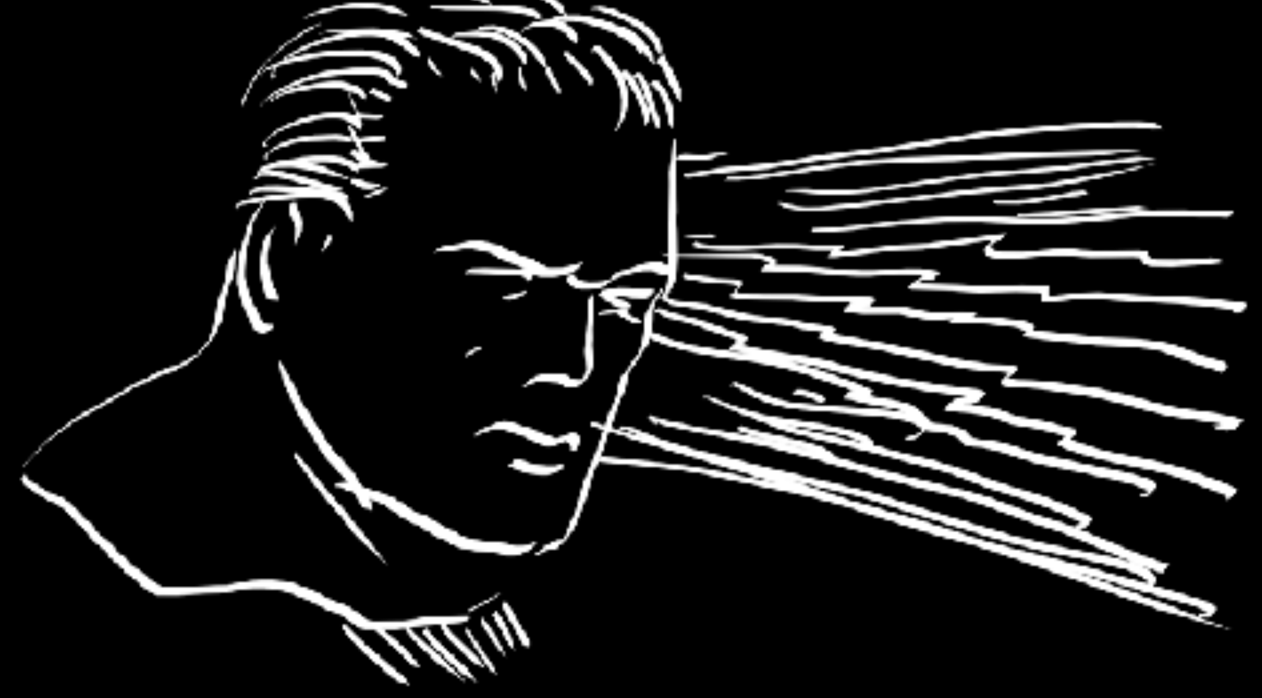
We need textbooks

We must put all we do into a coherent intellectual framework.

Systems and theory help with this too.

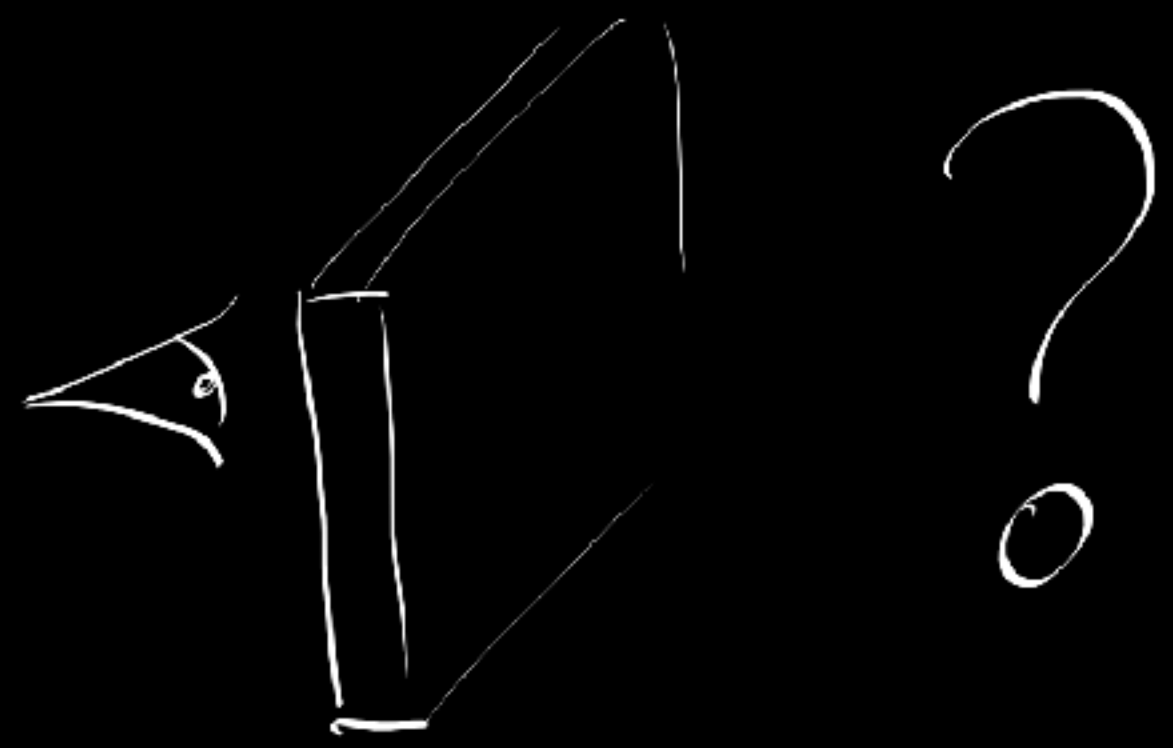


We should Reveal the invisible



a.k.a. super-hero vision

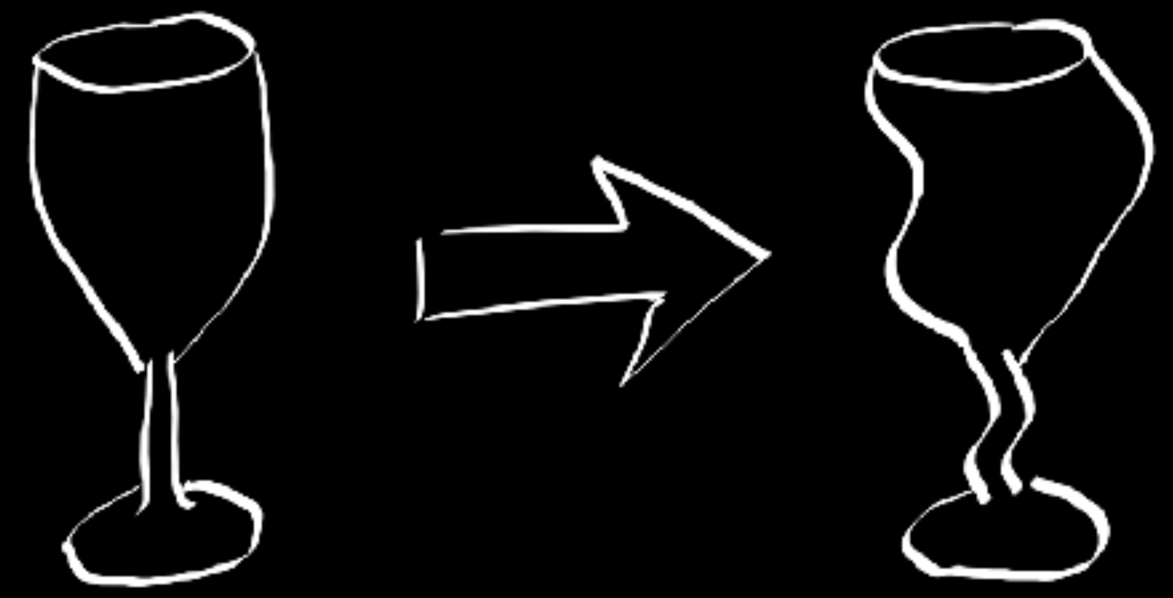
See beyond occlusion



Video magnification and video comparison

Reveal non-visible properties and effects

Leverage augmented reality to overlay new info



cf. my slides on the topic

PhD topics

Selection of what motion must be magnified

Reveal small differences between large motions

Seeing beyond corner without time of flight

Theory of seeing beyond corners, fundamental limits

Reveal non-optical properties: temperature, air flow, wifi signal

Reveal internal parts of objects

We must explore applications beyond photo

Computer Vision, inverse rendering

Virtual Reality, Augmented Reality

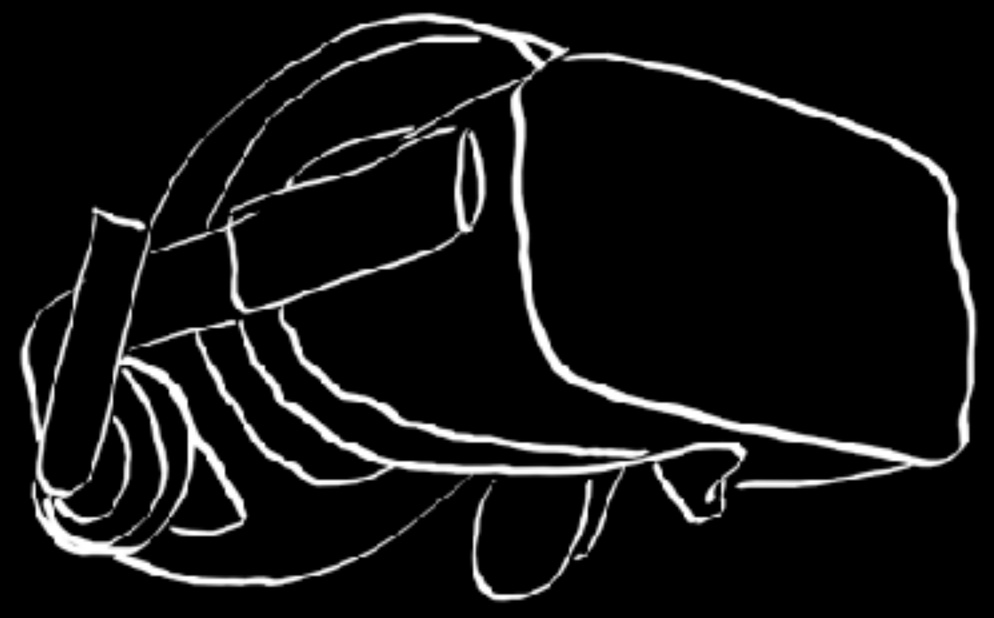
Automotive, robotics

Scientific & medical imaging

Various inputs, e.g. barcode reader

HCI, depth sensors, Kinect

Wireless networks and communication



PhD topics

VR/AR tracking

AR acquisition

Computer vision and deep learning from computational optics

Computational optics for driving

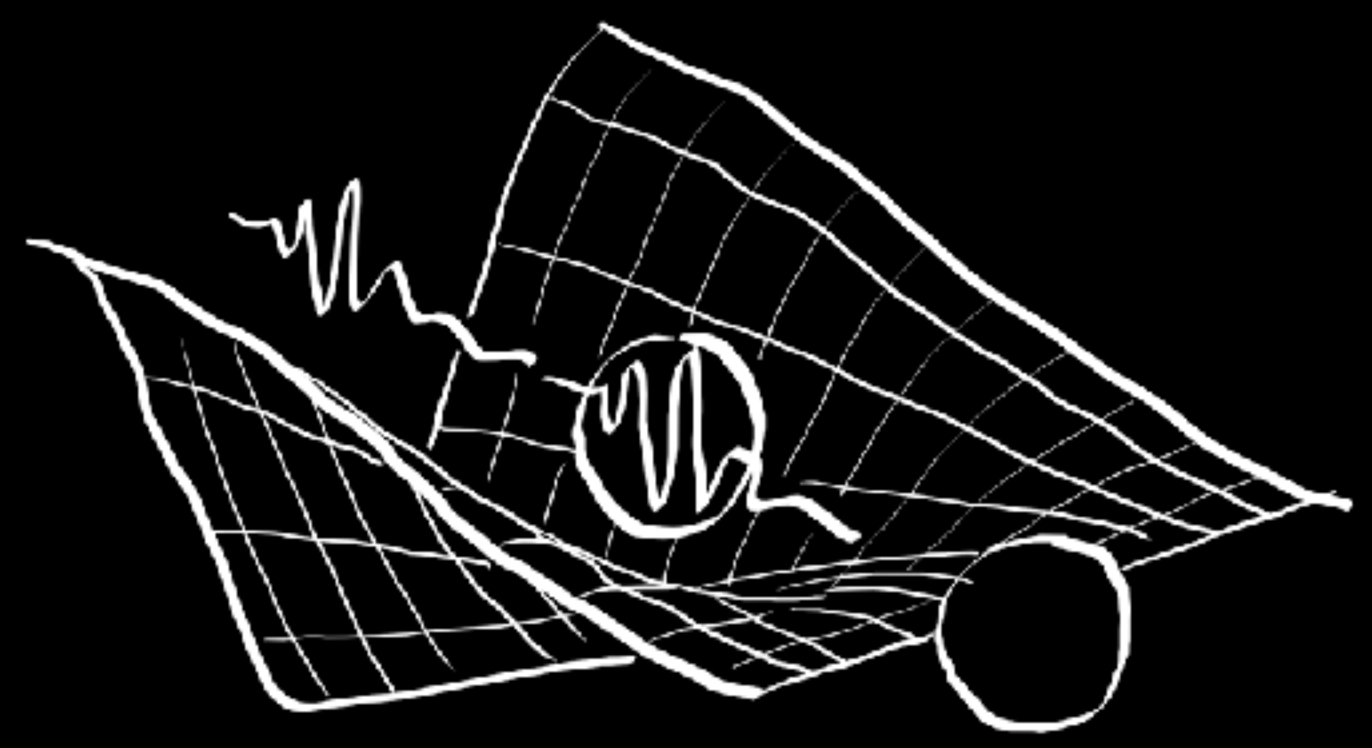
Wireless with optical wavelengths

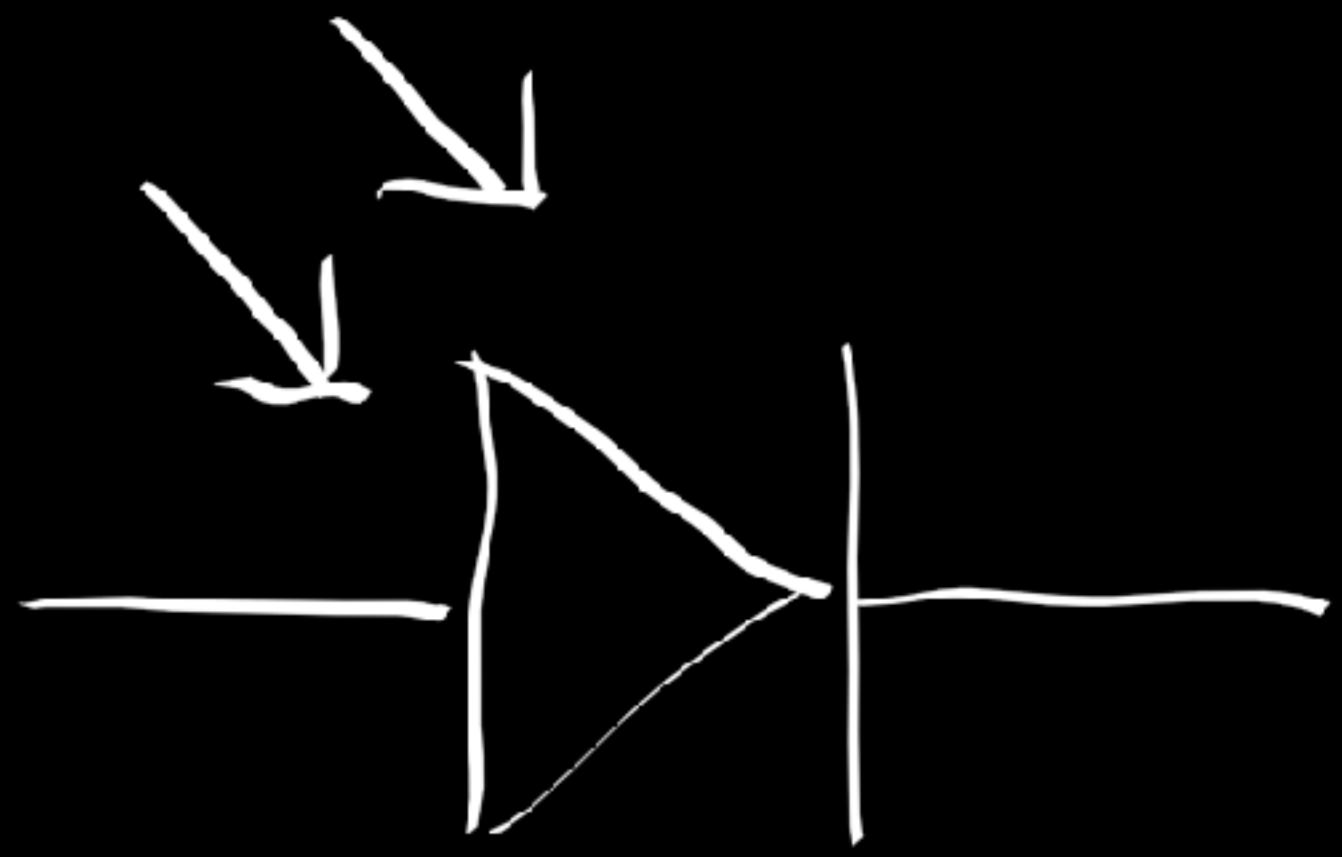
Computational ultrasound

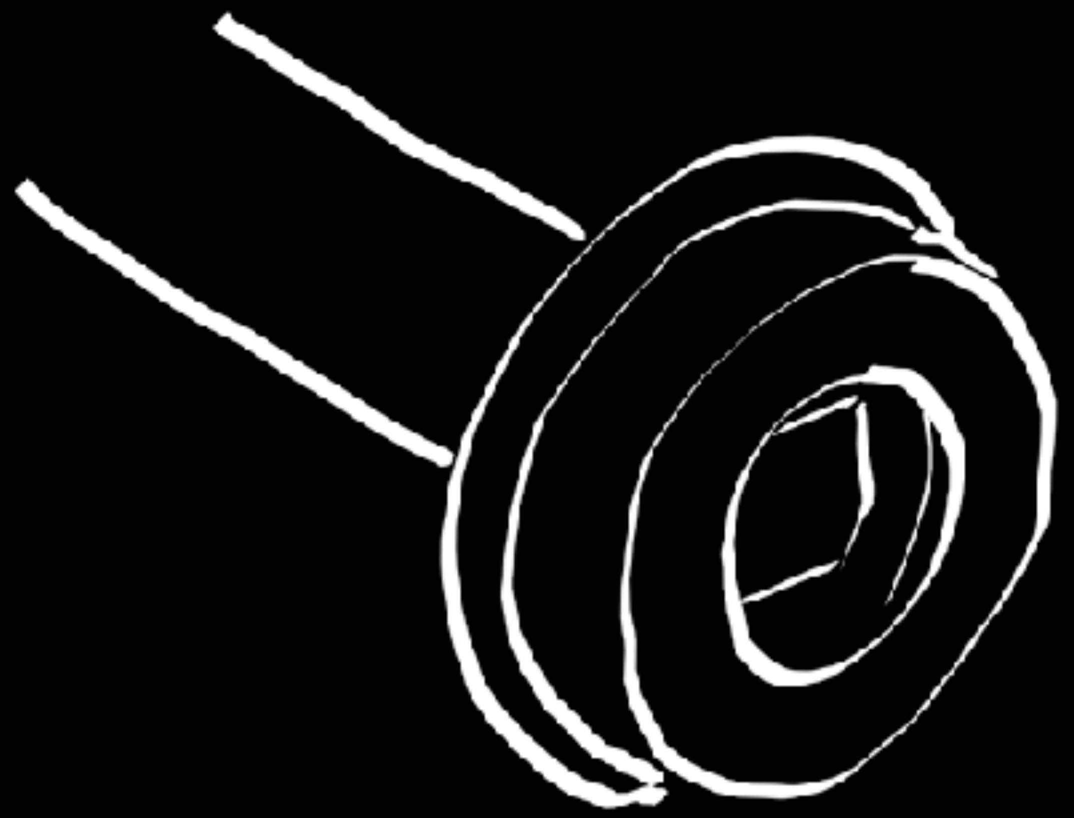
Most sensing is optical

We mostly know how to measure photons and electrons

Whatever you are measuring (bio, chemistry, medical, etc.), make it have a visible effect (fluorescent dye, diffracting bead, etc.)



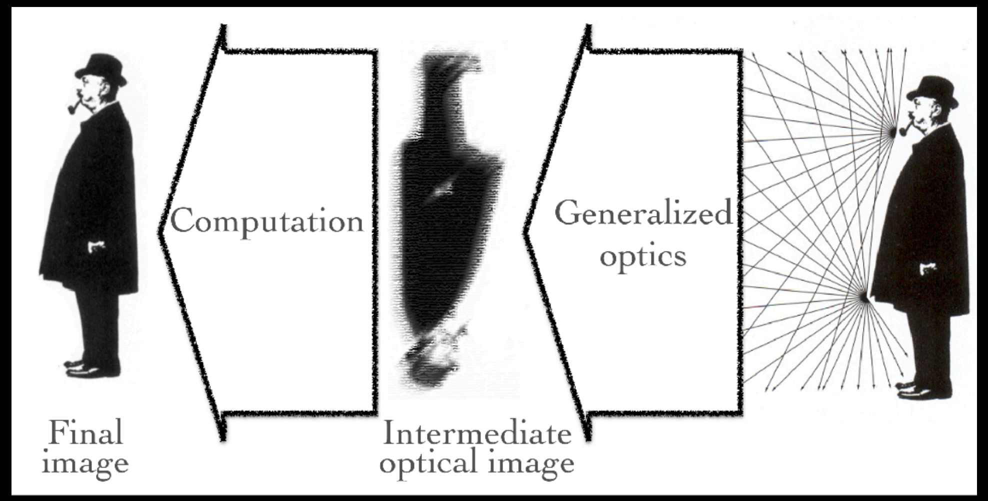




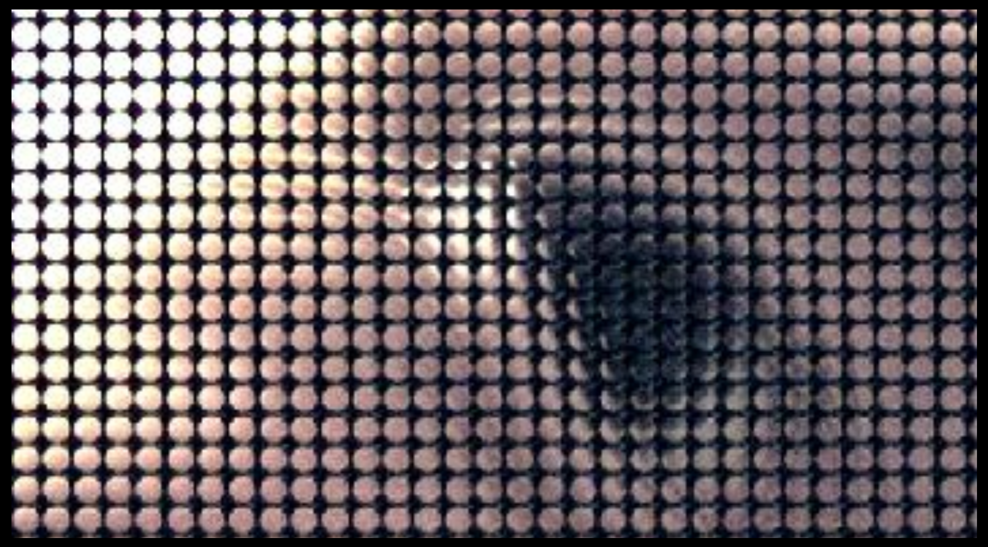
Big ideas recap

Multiple exposure

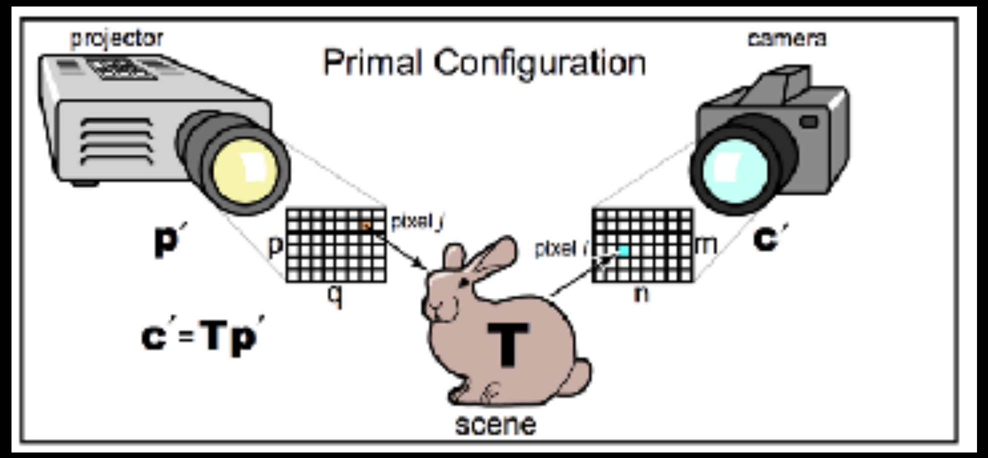
Coded imaging



High-dimensional Raw



Control illumination

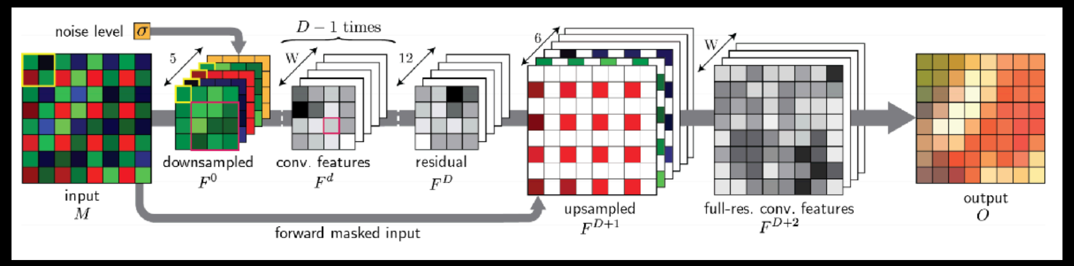


Edges matter

Priors

Big data

Deep learning



Research suggestions recap

Mobile rules

Video is still in its infancy

Scale

Systems

Deep learning

Theory

Hardware

Reveal the invisible

Applications in other fields

Thank You

Computation has revolutionized photography

Challenges and opportunity still abound

Computational photography should revolutionize other fields as well

Blog: thecomputationalphotographer.com

Thank you

My PhD advisors: George Drettakis, Claude Puech

Students, post-docs, collaborators: Jonathan Ragan-Kelley, Abe Davis, Zoya Bylinskii, YiChang Shih, Hijung (Valentina) Shin, Michael Gharbi, Ronnachai (Tiam) Jaroensri, Guha Balakrishnan, Amy Zhao, Neal Wadhwa, Michael Rubinstein, Emily Whiting, Hao-Yu Wu, Alec Rivers, Tilke Judd, Jaakko Lehtinen, Andrew Adams, Vladimir Bychkovsky, Jennifer Chan, Jiawen "Kevin" Chen, Forrester Cole, Anat Levin, Wojciech Matusik, Tom Annen, Thouis Ray Jones, Elmar Eisemann, Florent Duguet, Xavier Décoret, Alexis Charignon, Max Chen, Mok Oh, Yann Semet, Eric Chan, Barb Cutler, Stéphane Grabli, Emmanuel Turquin, Katie Bouman, Matthias Zwicker, Jan Kautz, Addy Ngan, Tom Mertens, Sylvain Paris, Will Hasenplaugh, Mike Doggett, Kari Puli, Adrien Bousseau, Taeg Sang Cho, Nicolas Bonneel, Kevin Egan, Aner Ben-Artzi, Timo Aila, Samuli Laine, Kartic Subr, Paul Green, Sara Su, Soonmin Bae, Marco da Silva, Sam Hasinoff, Bill Freeman, Ted Adelson, Antonio Torralba, Nicolas Holzschuch, Francois Sillion, Eugene Hsu, Henrik Wann Jensen, Shai Avidan, Victor Ostromoukhov, Cyril Soler, Manohar Srikanth, Boaz Nadler, Charlie Kilpatrick, Christophe Hery, Doug Epps, Hanspeter Pfister, Connely Barnes, Saman Amarasinghe, Peter Sand, Ce Liu, Carsten Dachsbacher, John Guttag, Ravi Ramamoorthi, Julie Dorsey, Yair Weiss, Marc Levoy, John Ochsendorf, Kavita Bala