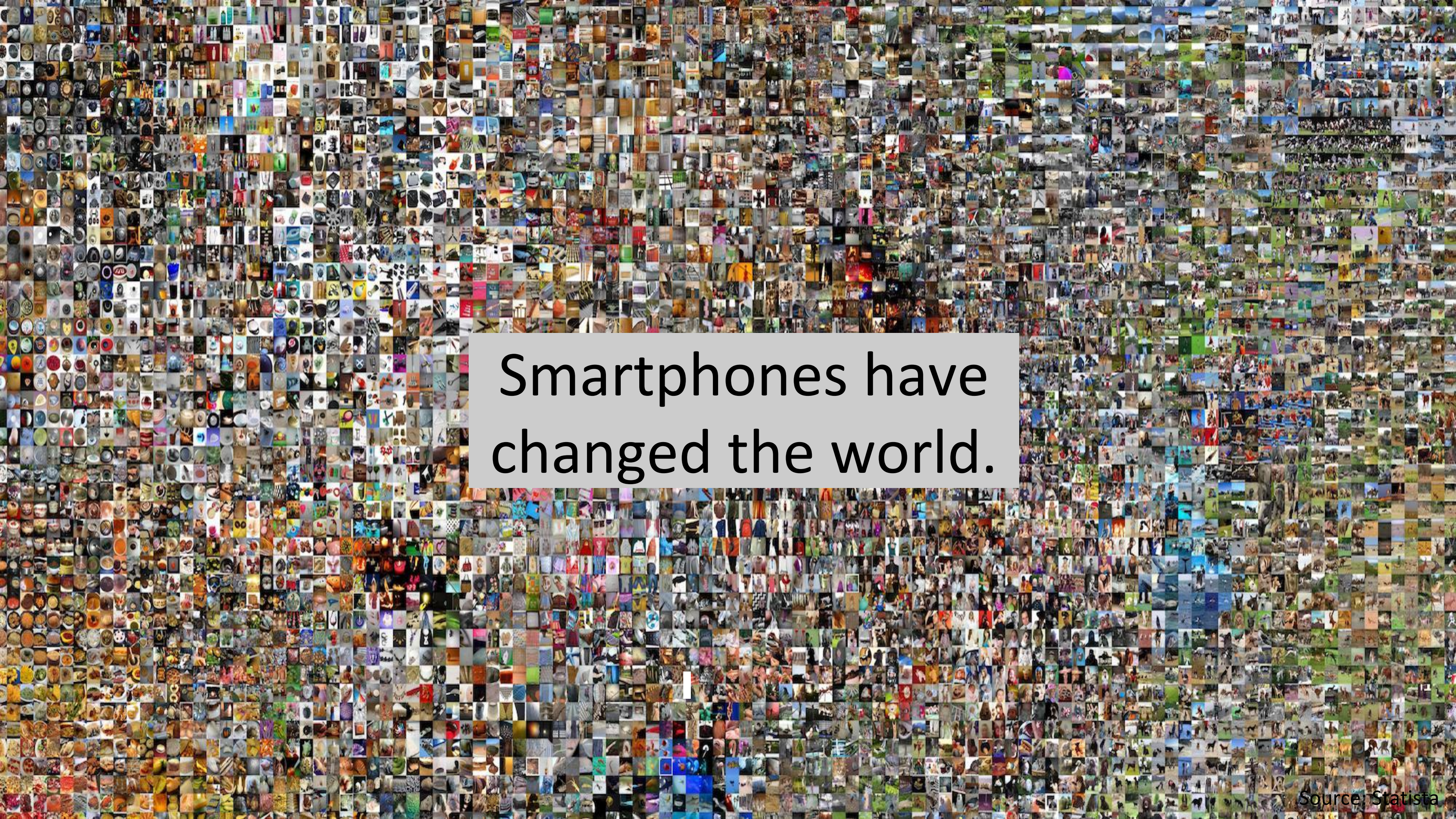
Computation + Photography

How the mobile phone became a camera

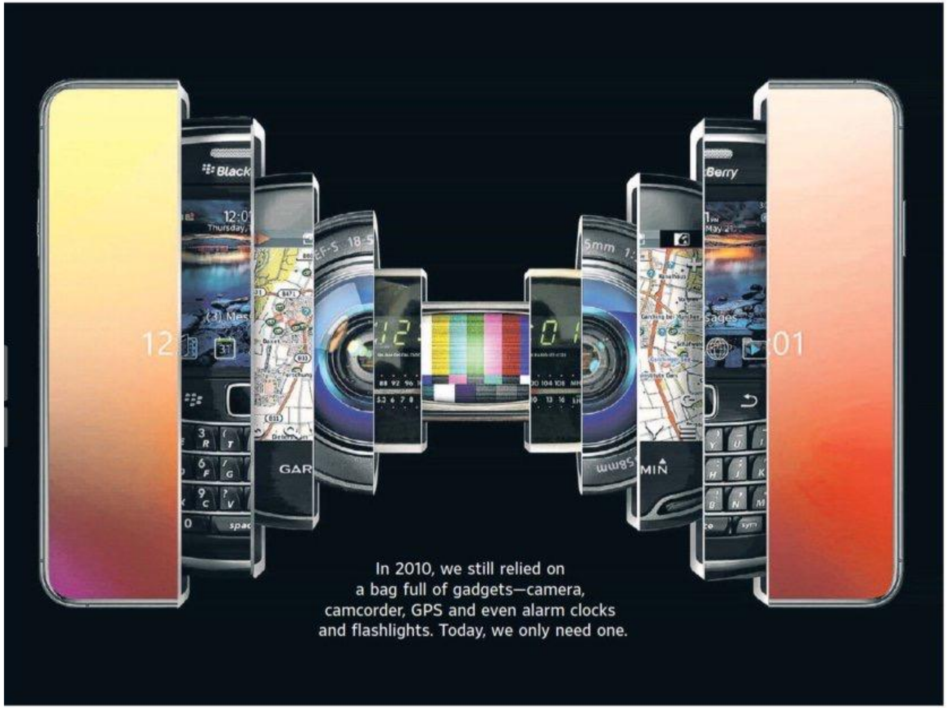
Part 1 : History

Peyman Milanfar

Google Research


Smartphones have changed the world.
Source: Statista

Smartphones have changed the world.



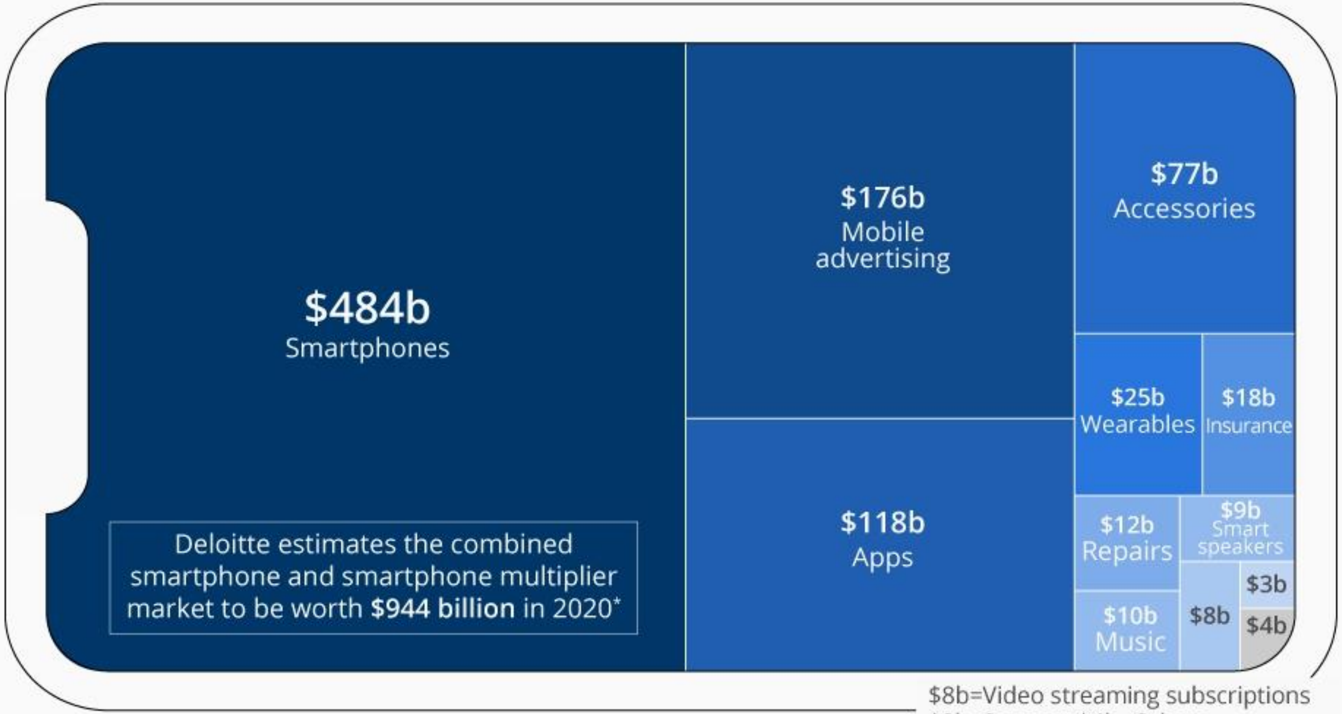
THE SMARTPHONE CHANGED. THEN IT CHANGED US.

THE WALL STREET JOURNAL.

JOANNA STERN

The Trillion-Dollar Smartphone Economy

Estimated sales of smartphones and related hardware, content and services in 2020



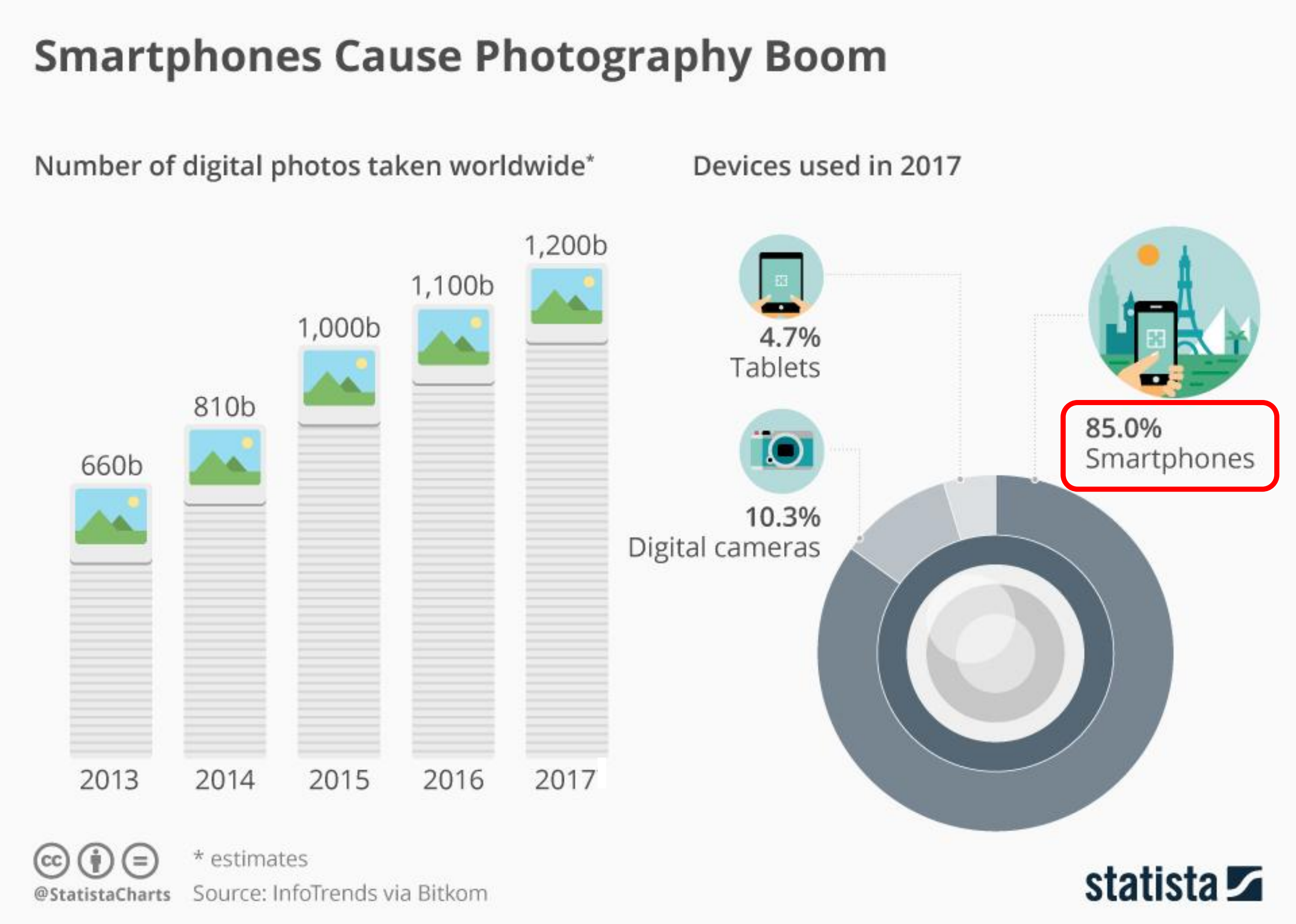
$8b=Video streaming subscriptions
$3b=Storage; $4b=Others

* Deloitte defines the "smartphone multiplier" as total sales of products and services depending on smartphone ownership.

Source: Deloitte analysis of data from App Annie, IFPI, Zenith and others

Vatican Square

Pope Benedict announcement



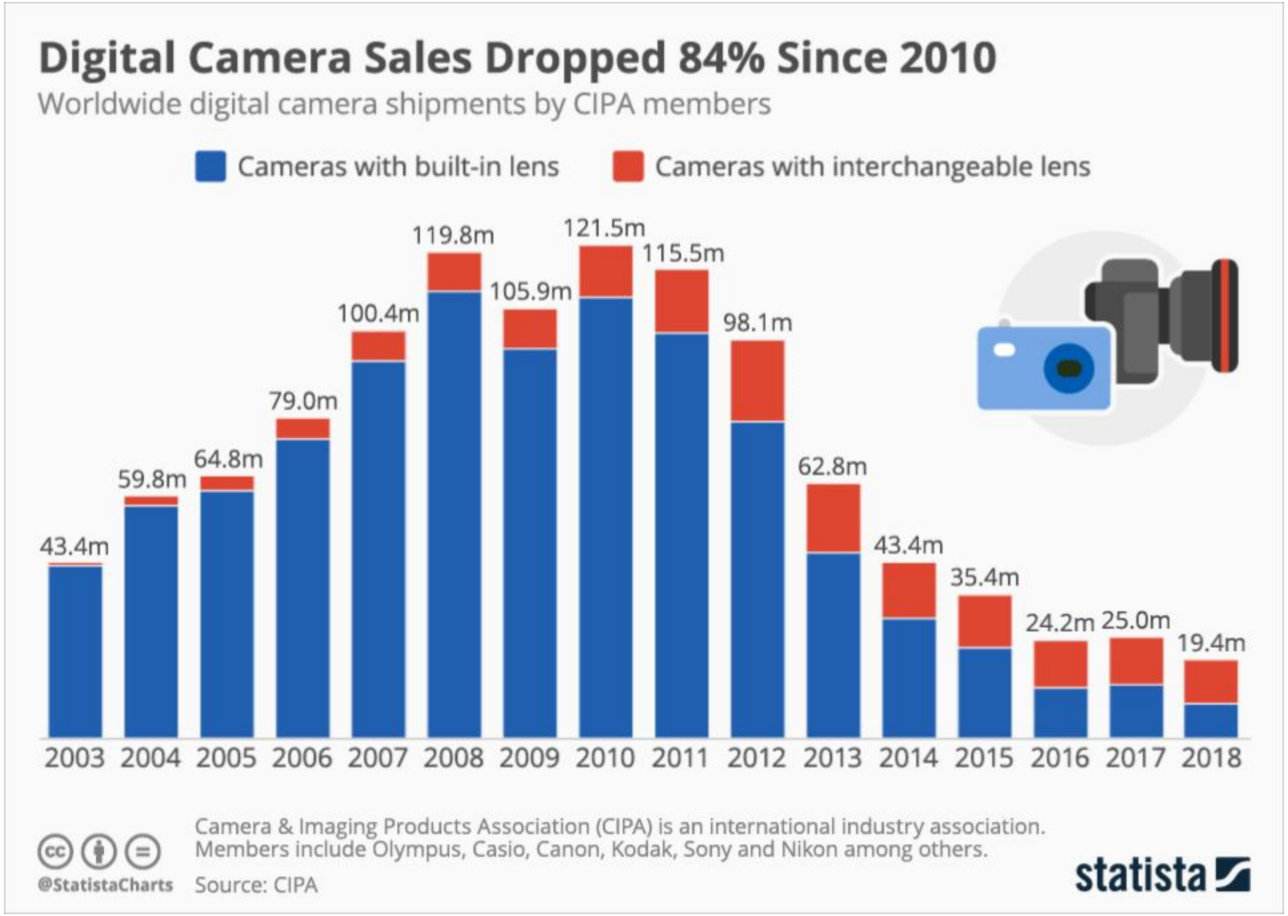
Smartphones Cause Photography Boom
Number of digital photos taken worldwide*
660b
810b
1,000b
1,100b
1,200b
2013
2014
2015
2016
2017
Devices used in 2017
4.7%
Tablets
10.3%
Digital cameras
85.0%
Smartphones
* estimates
Source: InfoTrends via Bitkom
@StatistaCharts
statista

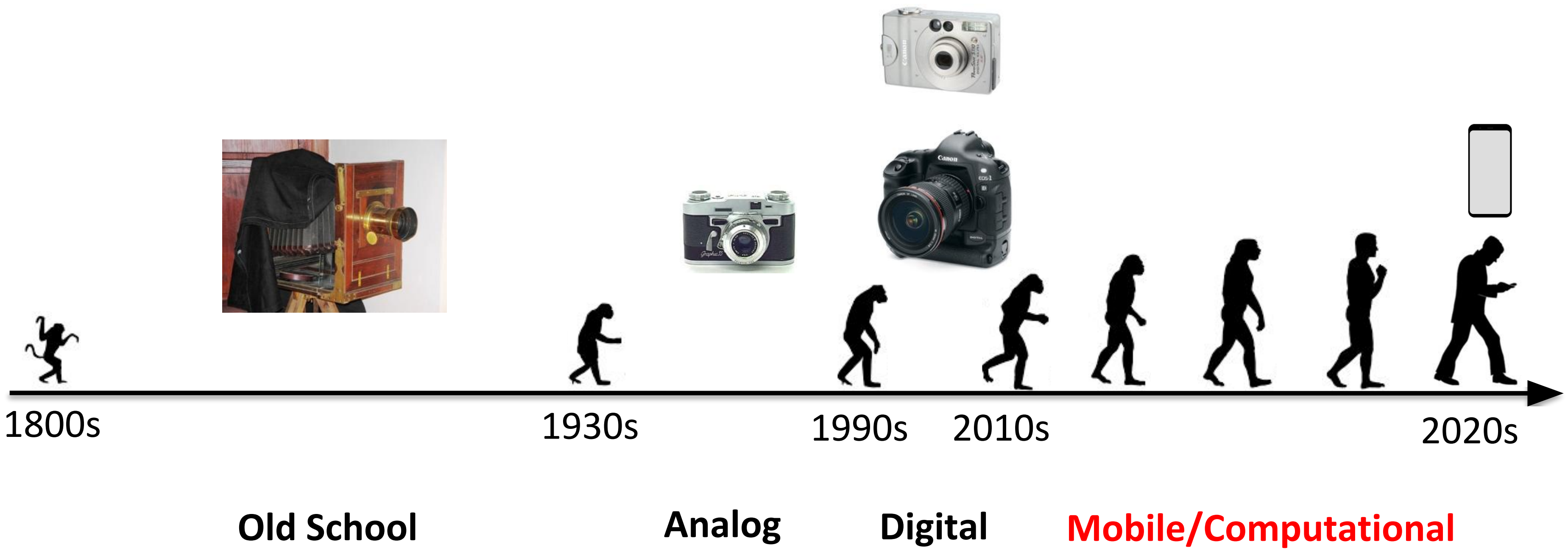
More than 2 billion photos shared on social media *per day*

Over 100 million are "selfies"

Reuters/CBS/TIME/KPCB

What Smartphones Have Done to the Camera Industry

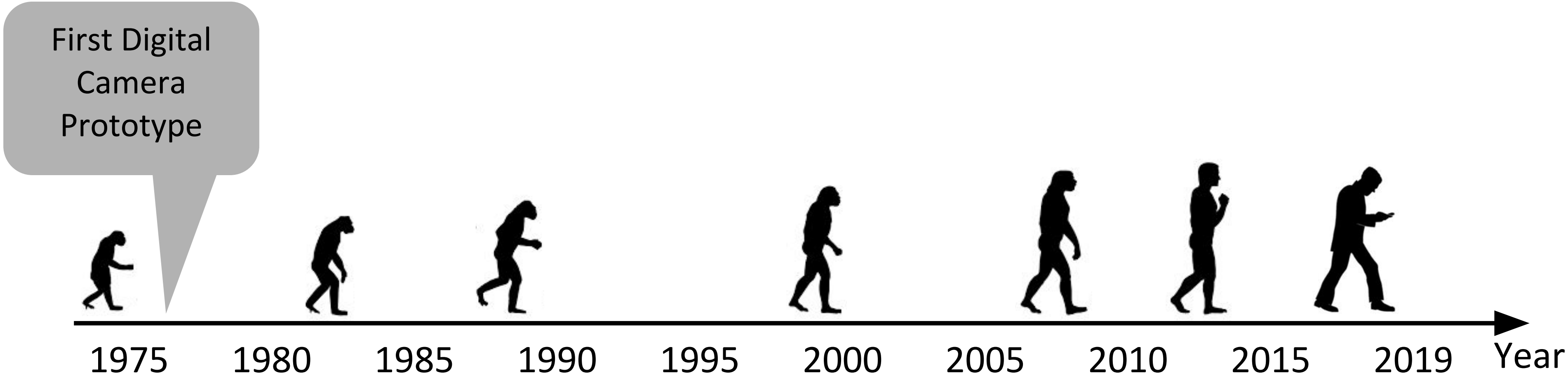



1800s
1930s
1990s
2010s
2020s
Old School
Analog
Digital
Mobile/Computational

Analog Film: 1935-1985

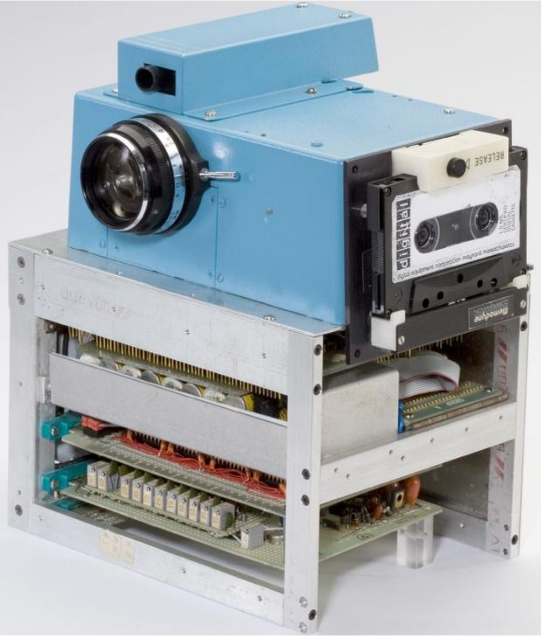
- Introduced in 1935, and dominant for about 50 years.
- Largely discontinued around 2005.

DIGITAL PHOTOGRAPHY



First Digital Camera Prototype
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year

First Digital Camera 1975



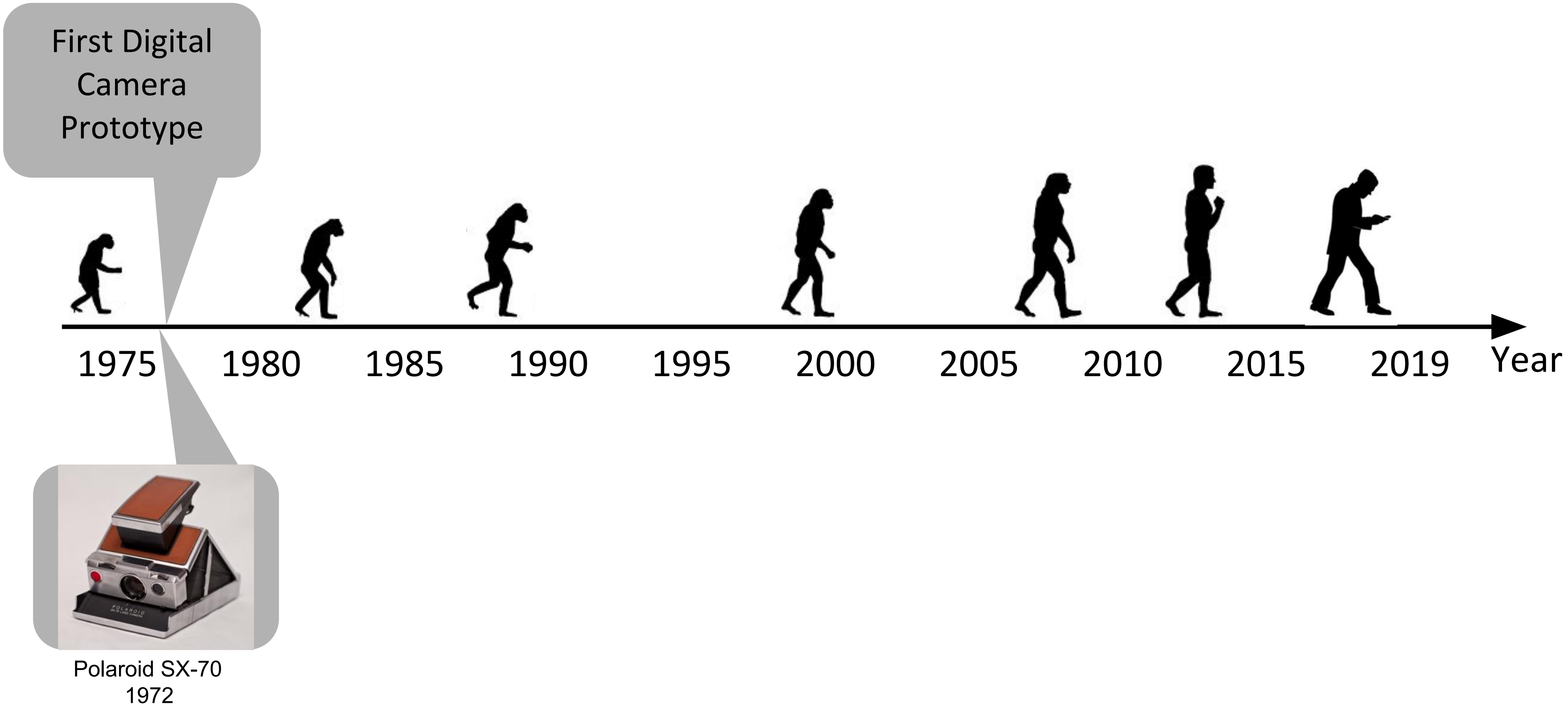
E-cam (Electronic Still Camera)
100x100 resolution (0.01Mpix)
Took 20 seconds to shoot a picture
Patented in 1978

"[Kodak executives] were convinced that no one would ever want to look at their pictures on a screen."
— *Steven Sasson*

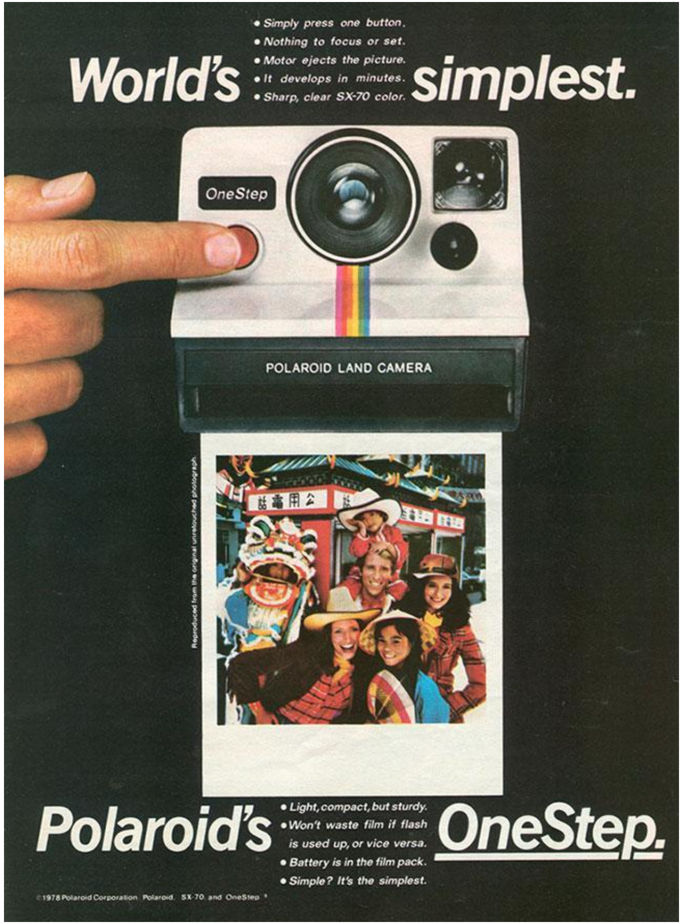
35 years later, Sasson got his due…

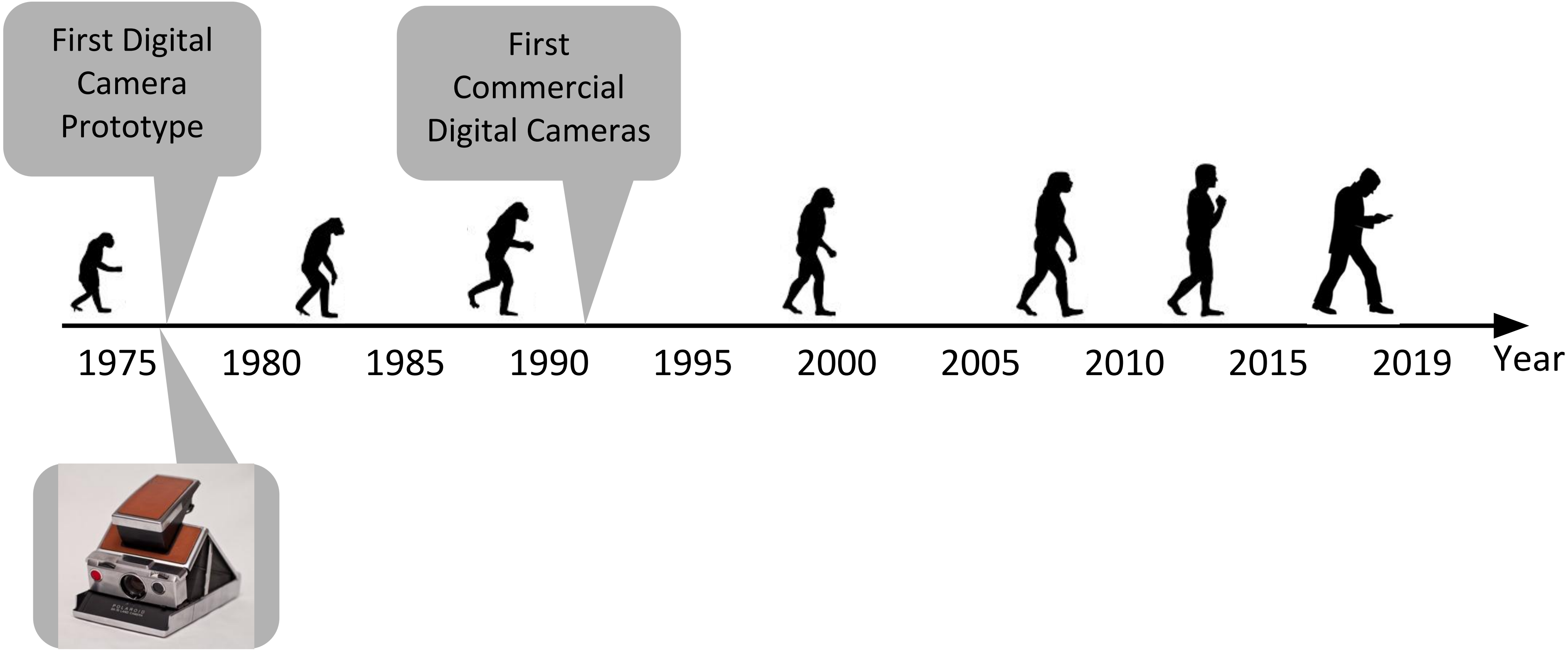
National Medal of Technology in 2009

“[Kodak executives] were convinced that no one would ever want to look at their pictures on a screen.”
— *Steven Sasson*


First Digital Camera Prototype
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
Polaroid SX-70
1972

Instant Gratification




First Digital Camera Prototype
First Commercial Digital Cameras
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
POLAROID

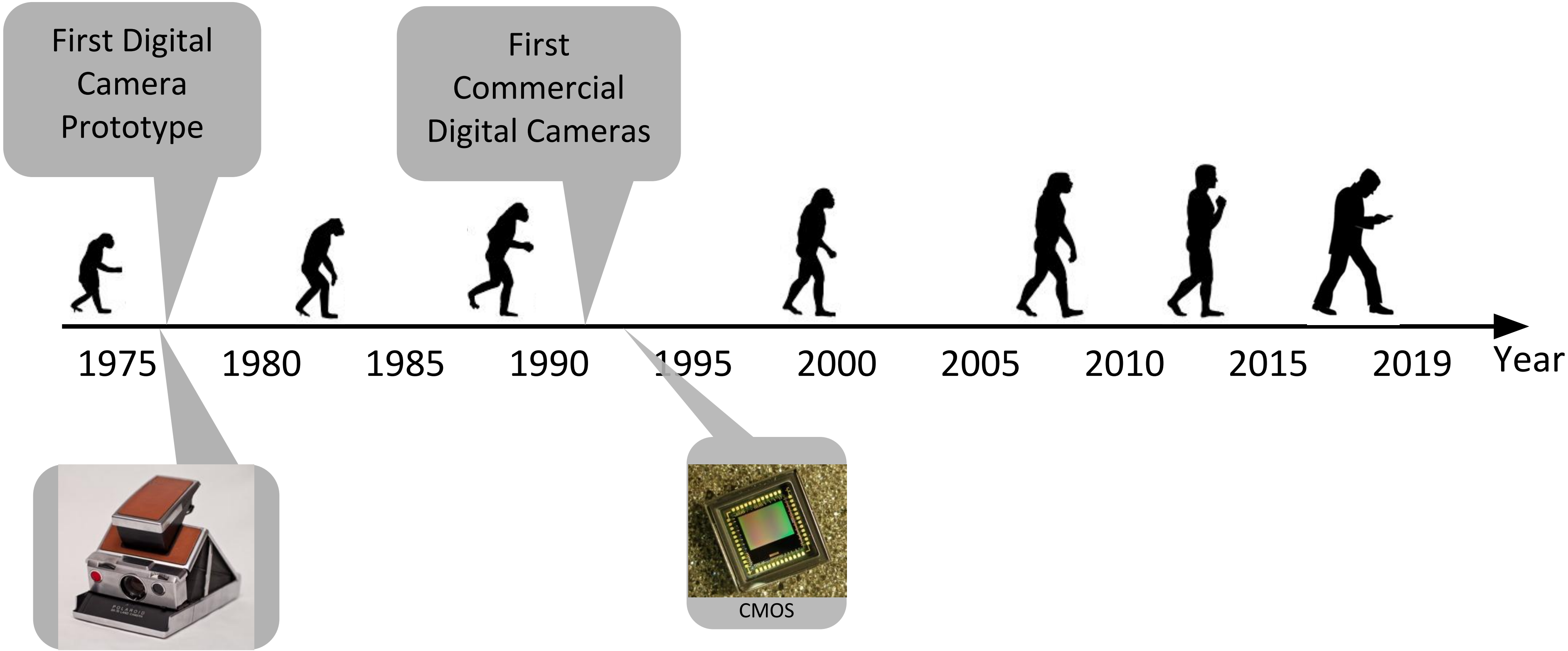
First Commercial Digital Camera 1990

Logitech Fotoman, 1990

376x284 resolution;
Black/white w/ 256 gray levels;
1Mb internal RAM;
Cost: $1000

Nikon bodies, Kodak sensors, 1992
First DSLR

1.5 Mpix resolution;
Tethered External Hard Disk
Cost: up to $20,000
Sold < 1000 units

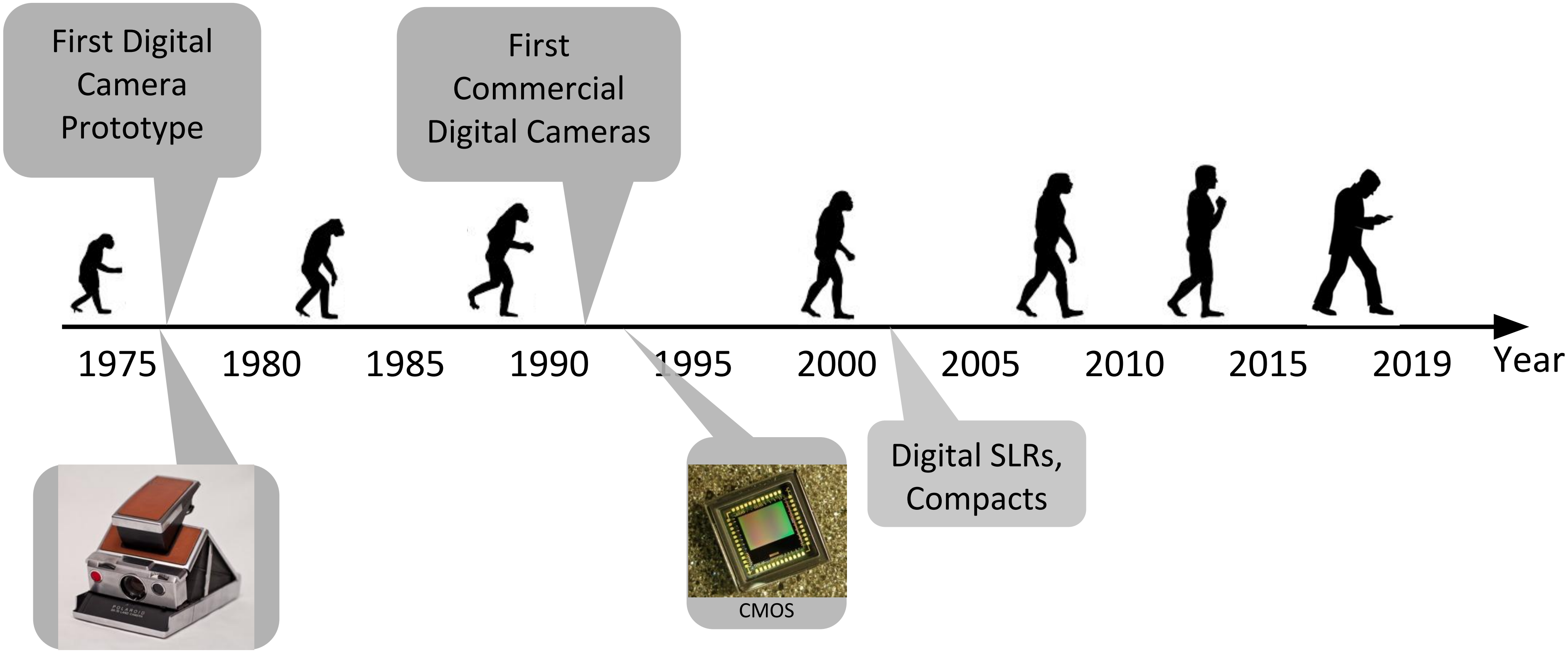

First Digital Camera Prototype
First Commercial Digital Cameras
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
CMOS

Invention of CMOS/Camera on a Chip

+ Cheaper, power efficient
- Noisier, rolling shutter readout

It would take another 10 years before CMOS systems would enable mass production of affordable (mobile) cameras

"Active Pixel Sensors: Are CCD's Dinosaurs?" Eric R. Fossum (1993), Proc. SPIE Vol. 1900, p. 2–14, in *Charge-Coupled Devices and Solid State Optical Sensors III*, Morley M. Blouke; Ed.


First Digital Camera Prototype
First Commercial Digital Cameras
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
CMOS
Digital SLRs, Compacts

Digital SLRs and Compacts (CCD)

Canon Powershot, 2000

1.5 Mpix resolution;
Cost: $500

Nikon D1, 2000

2-3 Mpix resolution;
Cost: $3 - 5K

Fast Forward to Today (CMOS)

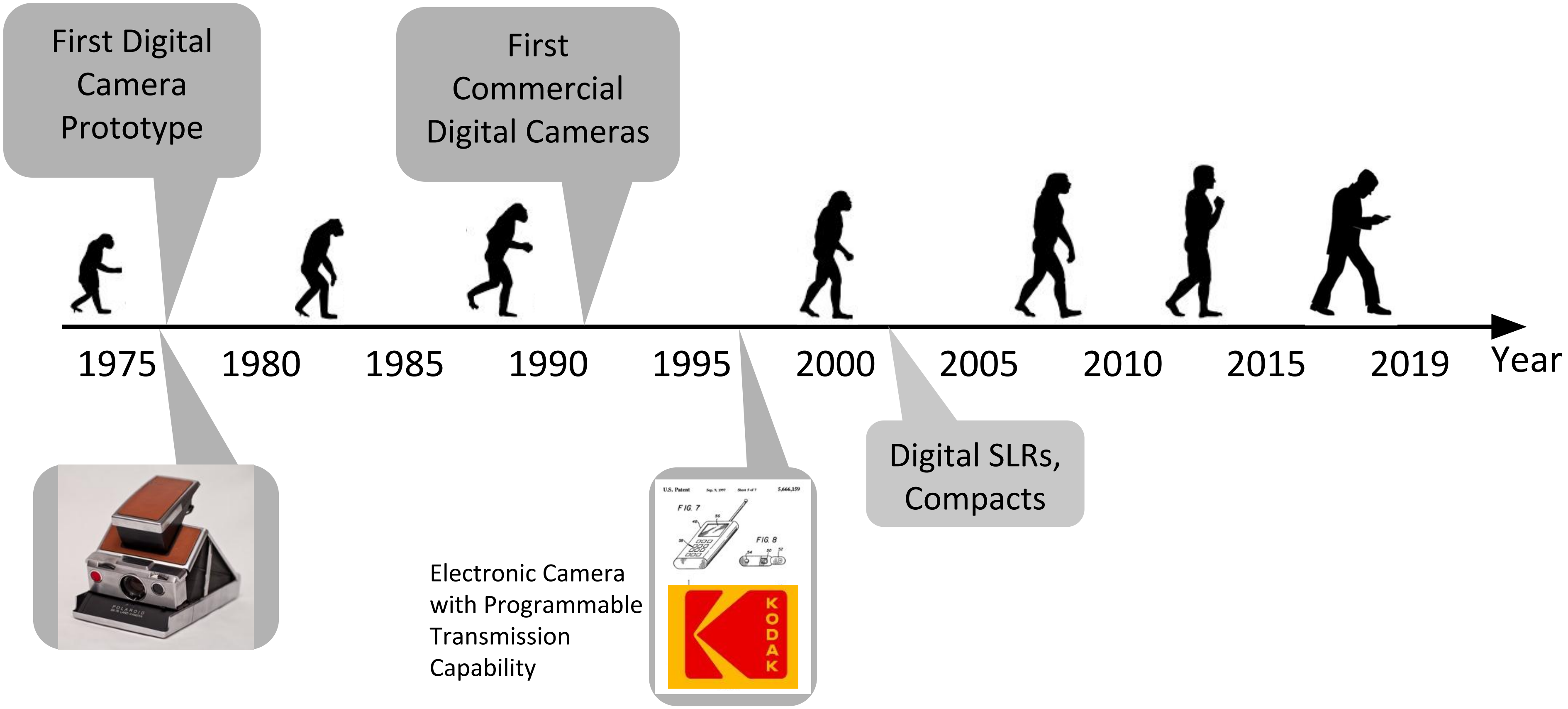
Sony RX100, 2019

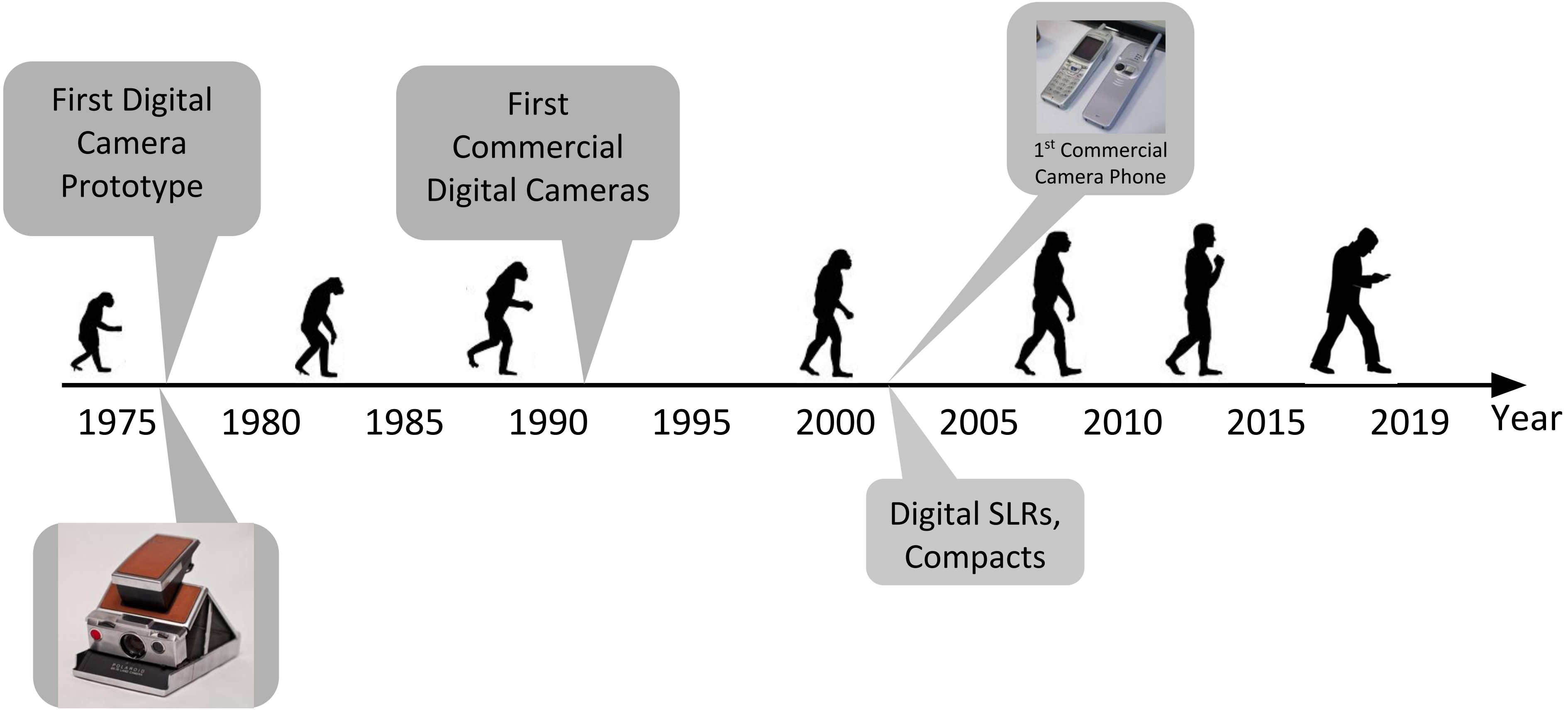
20 Mpix resolution;
Cost: $500

Nikon D810, 2019

36 Mpix resolution;
Cost: $3- 5K

MOBILE PHOTOGRAPHY


First Digital Camera Prototype
First Commercial Digital Cameras
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
Digital SLRs, Compacts
U.S. Patent
5,666,159
FIG. 7
FIG. 8
KODAK
Electronic Camera with Programmable Transmission Capability


First Digital Camera Prototype
First Commercial Digital Cameras
1st Commercial Camera Phone
Digital SLRs, Compacts
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year

J-Phone (Sharp), sold in Japan ‘00

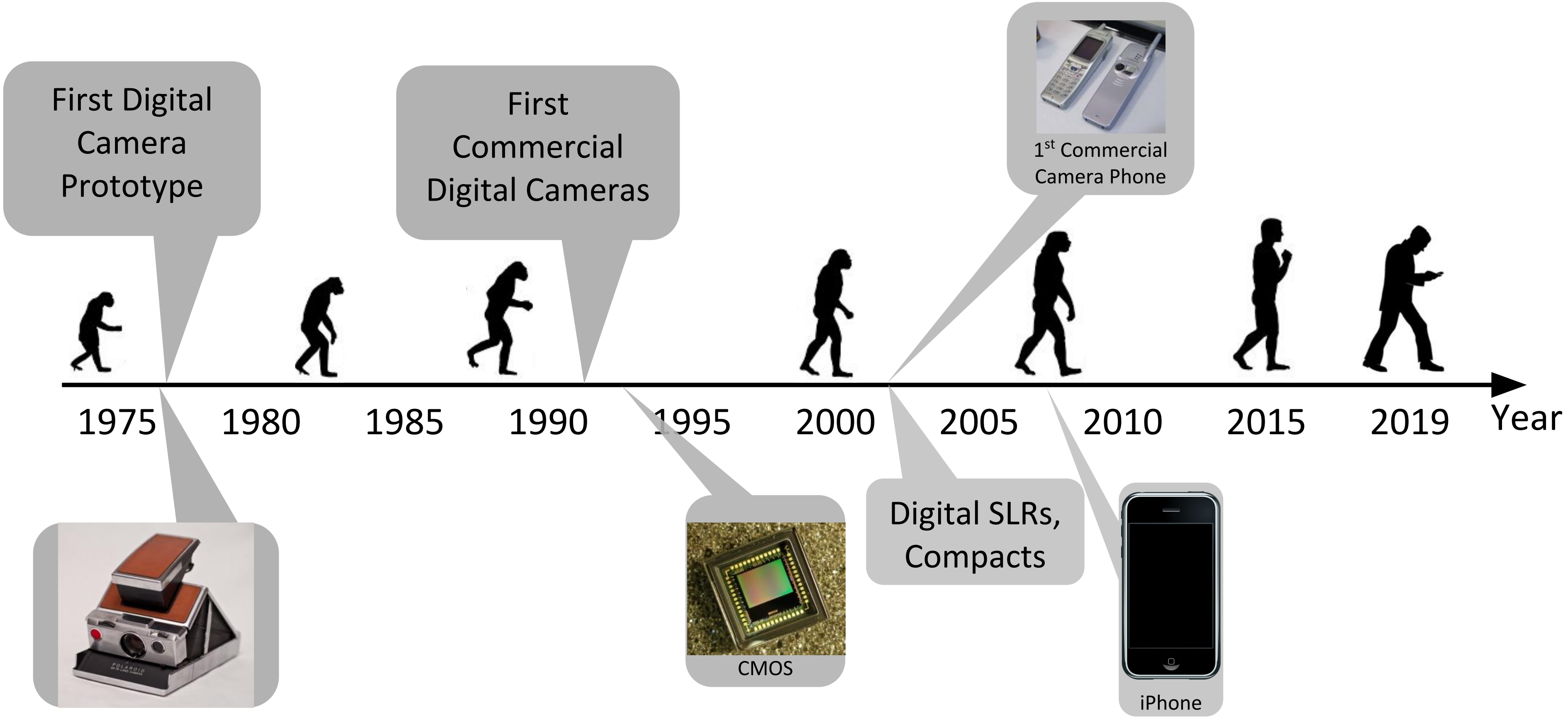
0.1 Mpix, CCD
256 color disp.
$500

$500

First camera (flip) phone in the US two years later in 2002

“Video Calls”

First phone with front camera a year later in 2003


First Digital Camera Prototype
First Commercial Digital Cameras
1st Commercial Camera Phone
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
CMOS
Digital SLRs, Compacts
iPhone

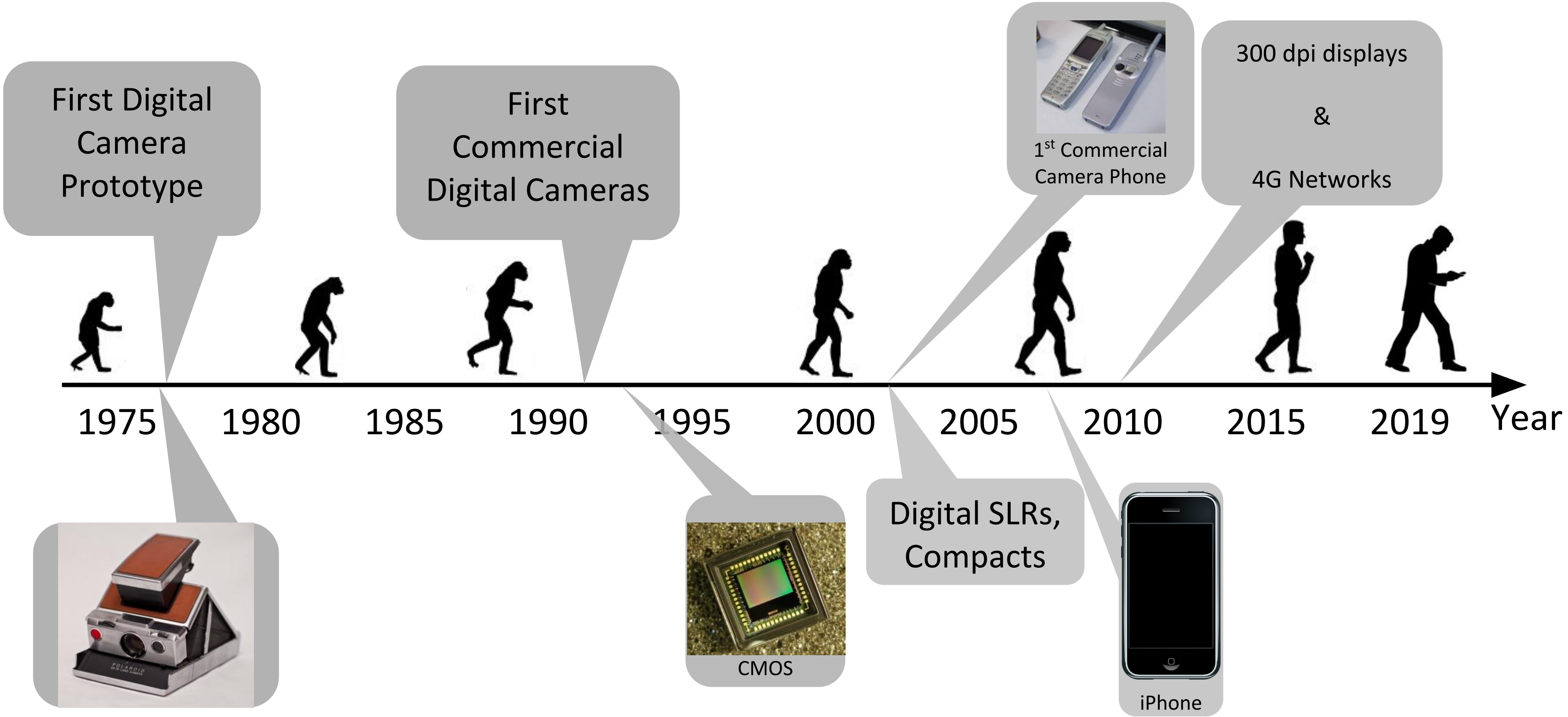
“Apple reinvents the phone” (but not the camera)

Display and UI were king.

“On the back, the biggest thing of note is we’ve got a two megapixel camera built right in.”

- Steve Jobs

Competition: Compact Cameras

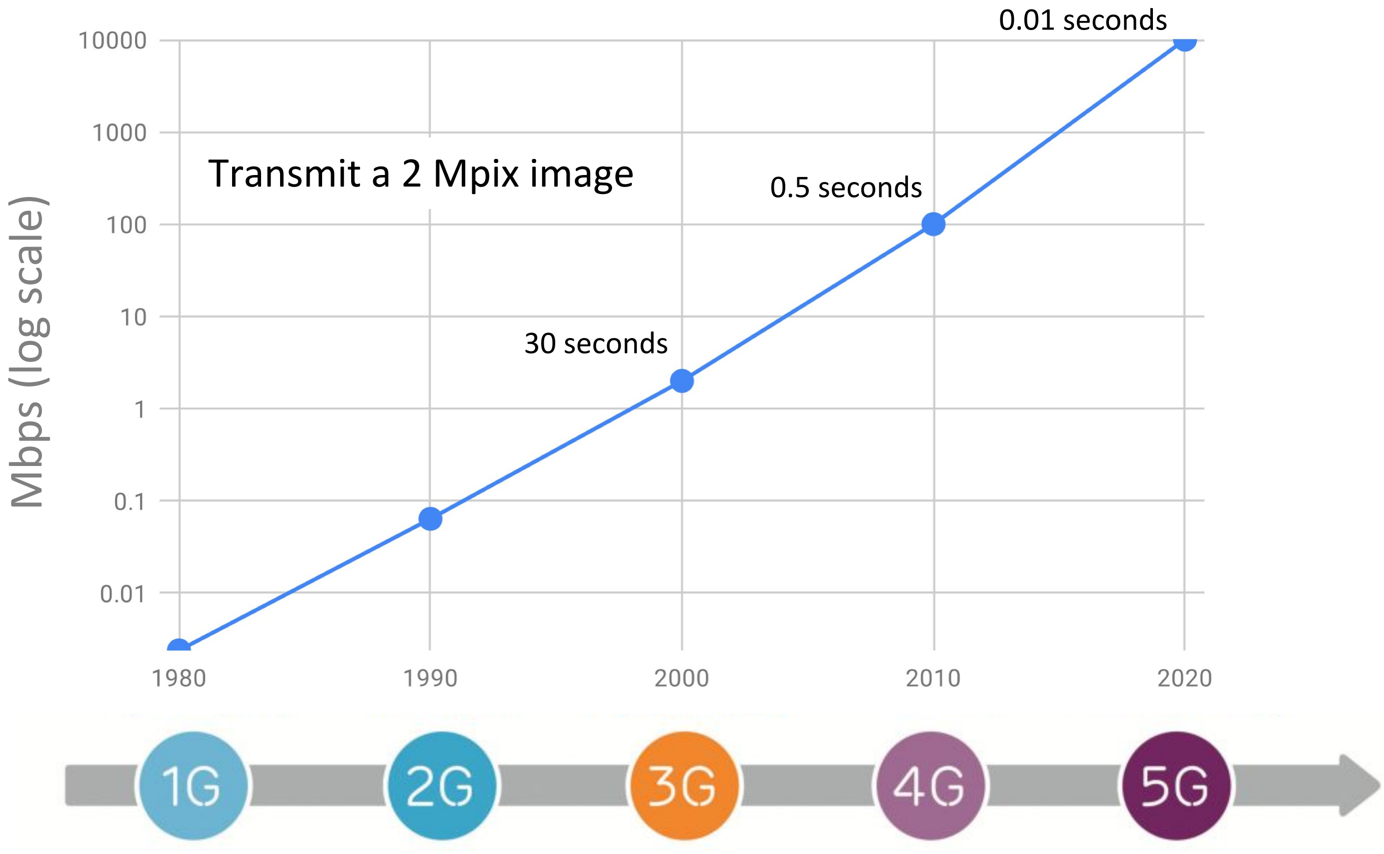

First Digital Camera Prototype
First Commercial Digital Cameras
1st Commercial Camera Phone
300 dpi displays
&
4G Networks
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
CMOS
Digital SLRs, Compacts
iPhone

Displays

1990 2010 2020

300dpi
displays

Wireless Network Speed



Source: Silika

2010 -

COMPUTATIONAL PHOTOGRAPHY

"The best camera is the one that's with you."

Computation + Photography
How the mobile phone became a camera

Part 2 : Modern Technology

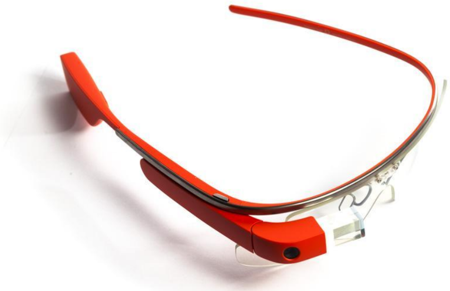
Peyman Milanfar
Google Research

2010 -

COMPUTATIONAL PHOTOGRAPHY

“The best camera is the one that’s with you.”

A Recent History at Google

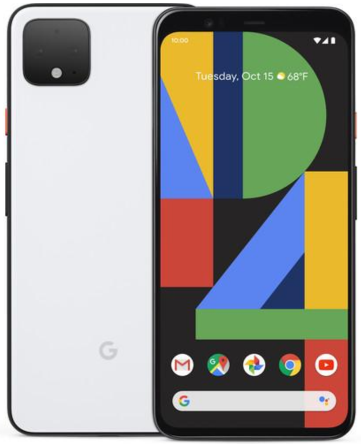


2012

2014

2016

2018



2019

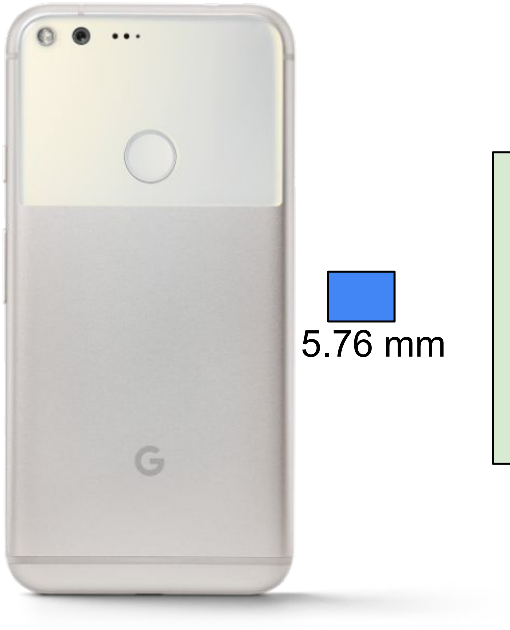
Can one be as good as the other?

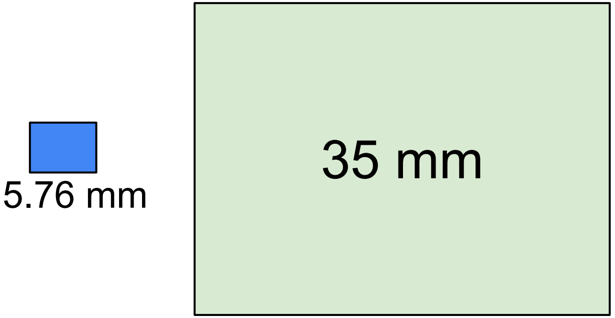
Can one be as good as the other?

$4.7\ mm^2$ **~ 300x =** $1360\ mm^2$

Credit Yael Pritch

Less light gets recorded





Compete with hardware!

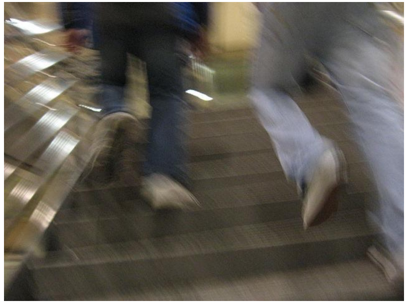
1 camera 2 cameras 3 camera 4 cameras 5 cameras

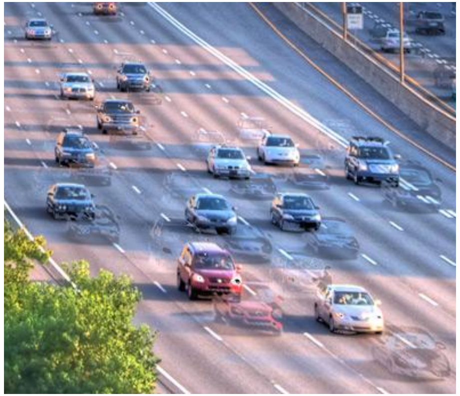
Yet most of the improvements are due to software.

Want: More light, dynamic range, resolution

Use a flash

Longer / bracketed exposures

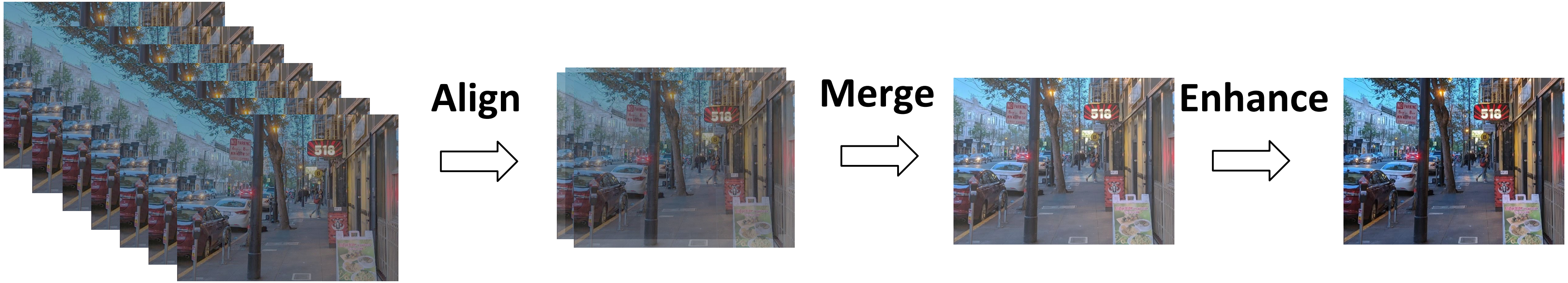




Capture a burst ✓

Modern Mobile Imaging: Burst Photography

Exposure control

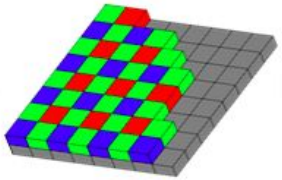


Align: Reliable Optical Flow – Scene is never stationary

Merge: Artifact-free Fusion – Alignment failures, occlusion, ...

Enhance: Denoise, Sharpen, Contrast enhancement, Dynamic Range

Classic Camera Image Processing Pipeline



Sensor with color filter array (CCD/CMOS)

Gain Control
A/D Converter
Possible LUT

White Balance

"Merge"

Demosaic

"Enhance"

Noise Reduction/ Sharpening

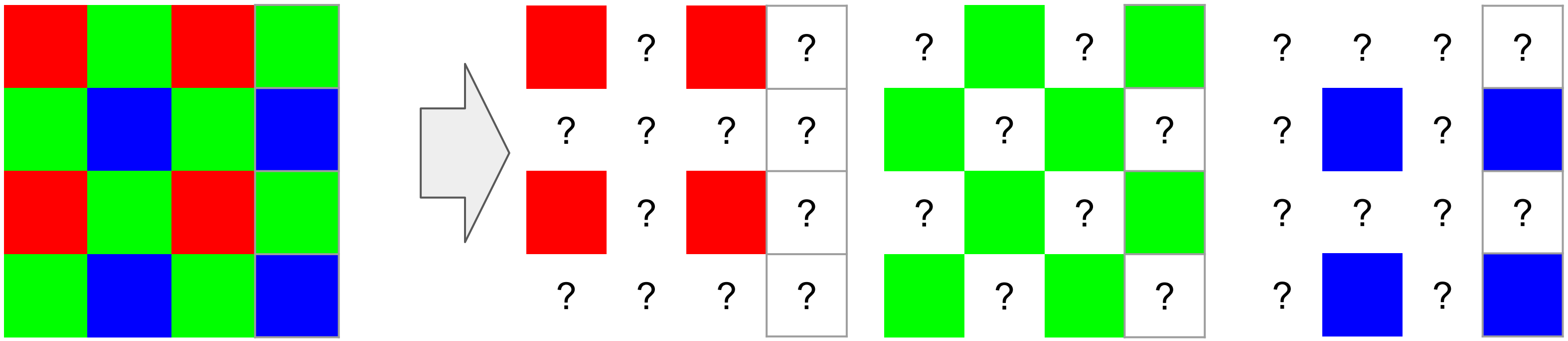
Color Space Transform + Color Preferences

Tone Reproduction

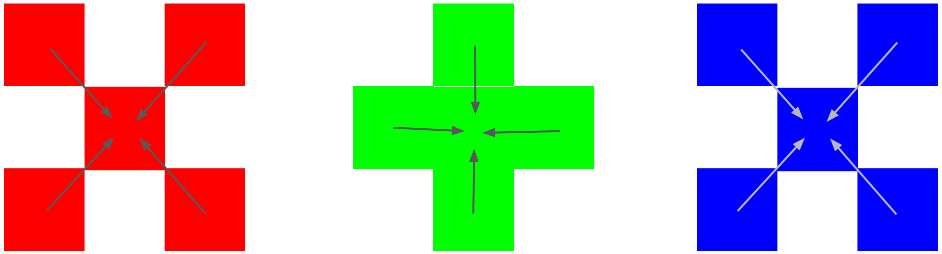
JPEG Compression

Michael Brown

Demosaicing : 12MP sensor ≠ 12 million RGB pixels



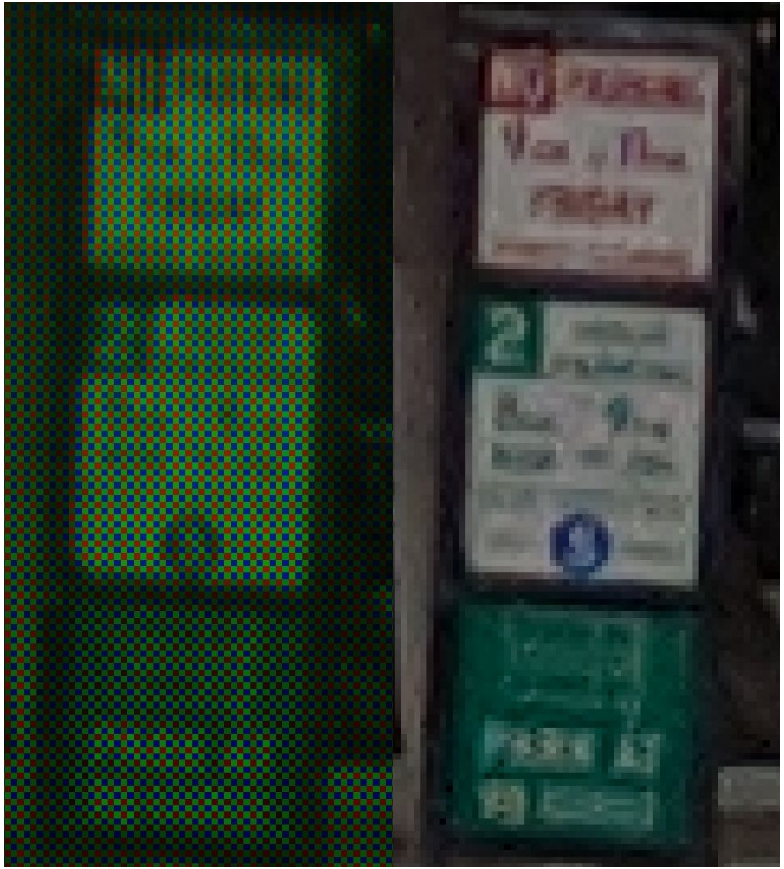
Missing information



Two-thirds of your picture is made-up!

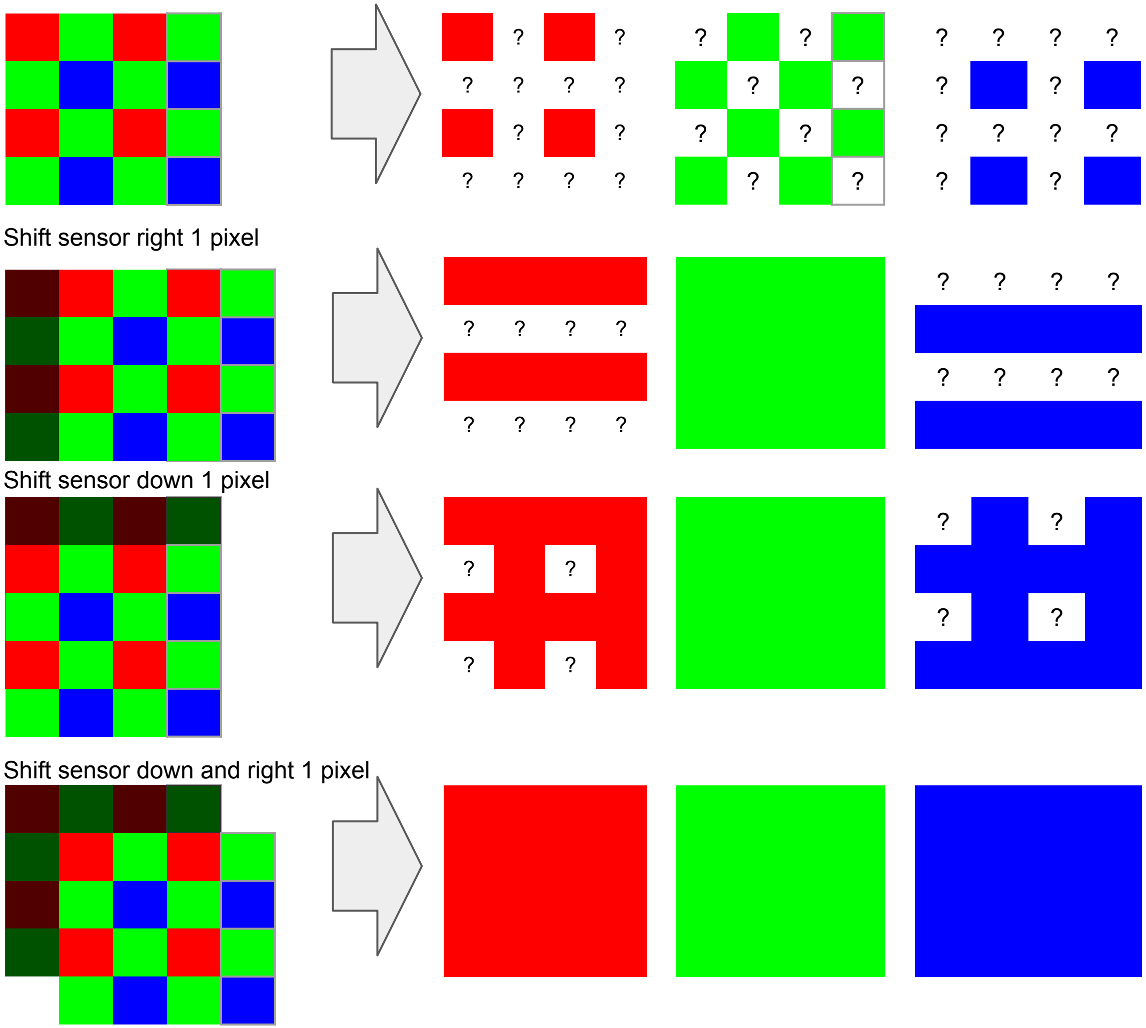
Demosaicing

Demosaicing …. Kills Details and Produces Artifacts

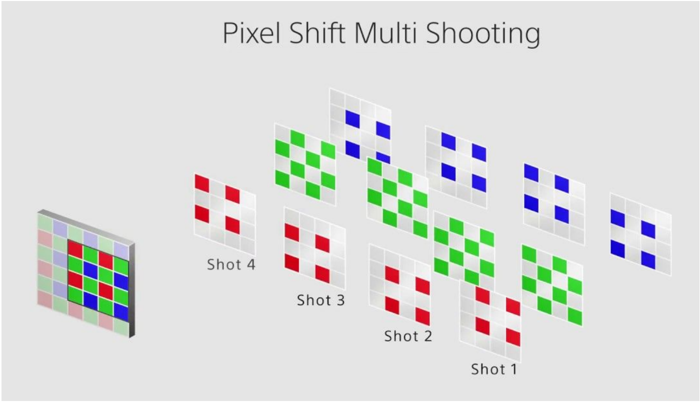


Instead ….. Replace demosaicing with multiple frames

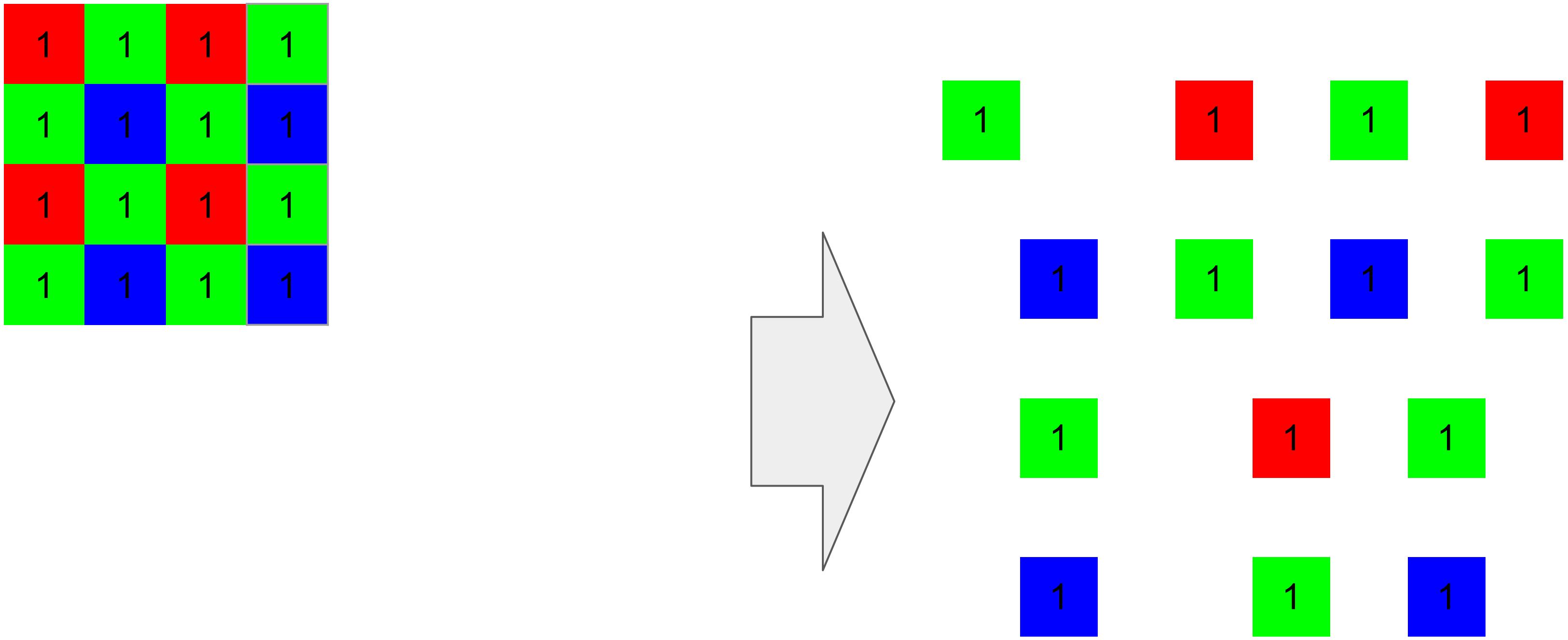
How: “Pixel-shifting”

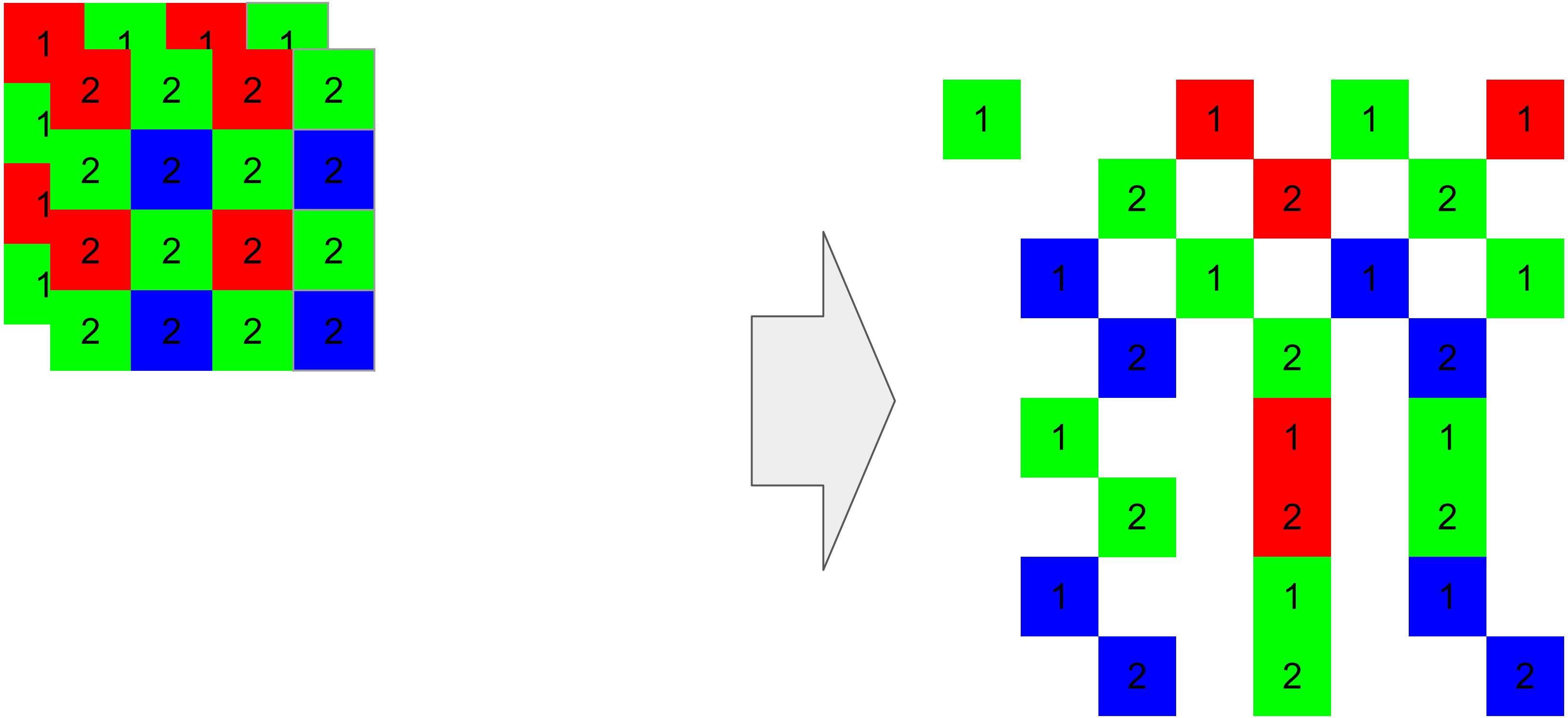


Some Mirrorless Cameras do “Pixel Shift Mode”

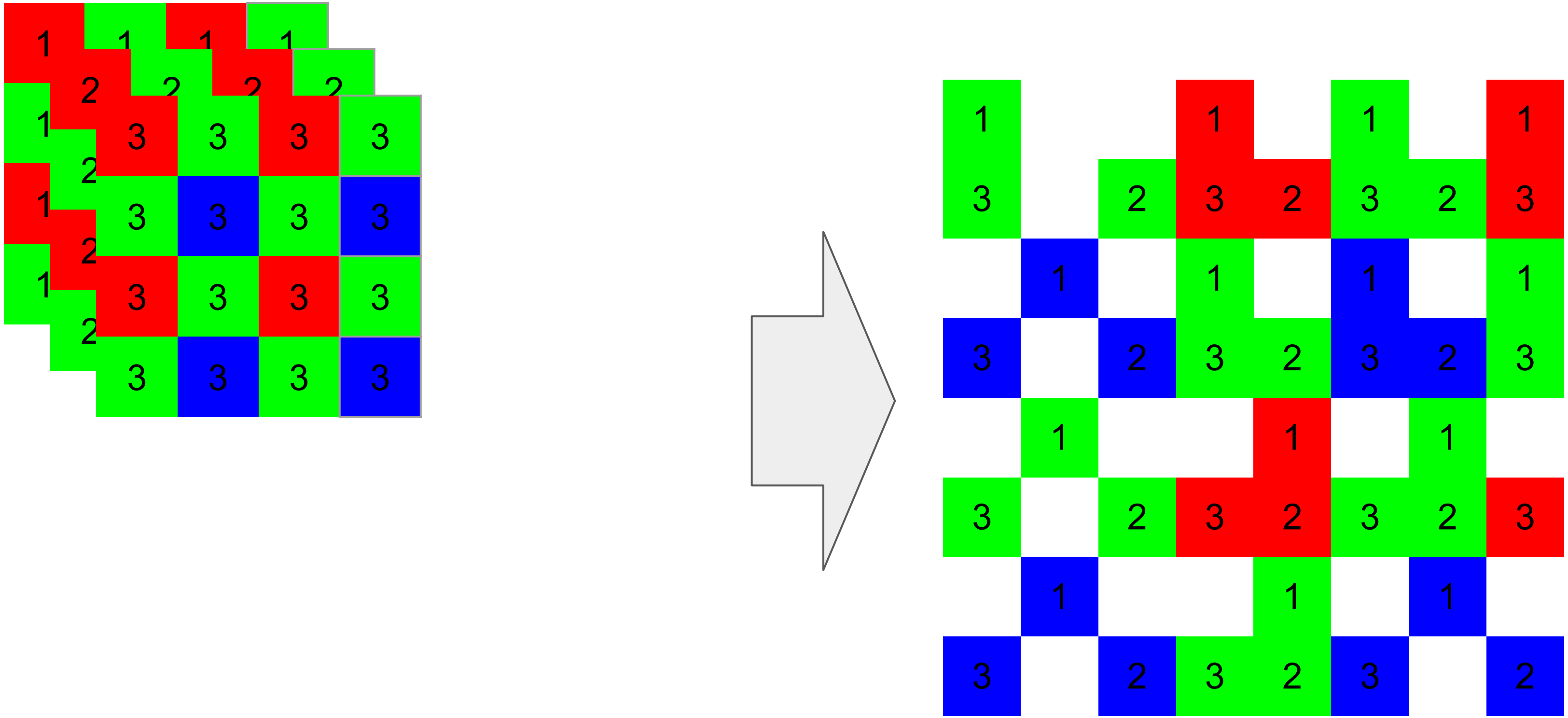


Life is not so simple.

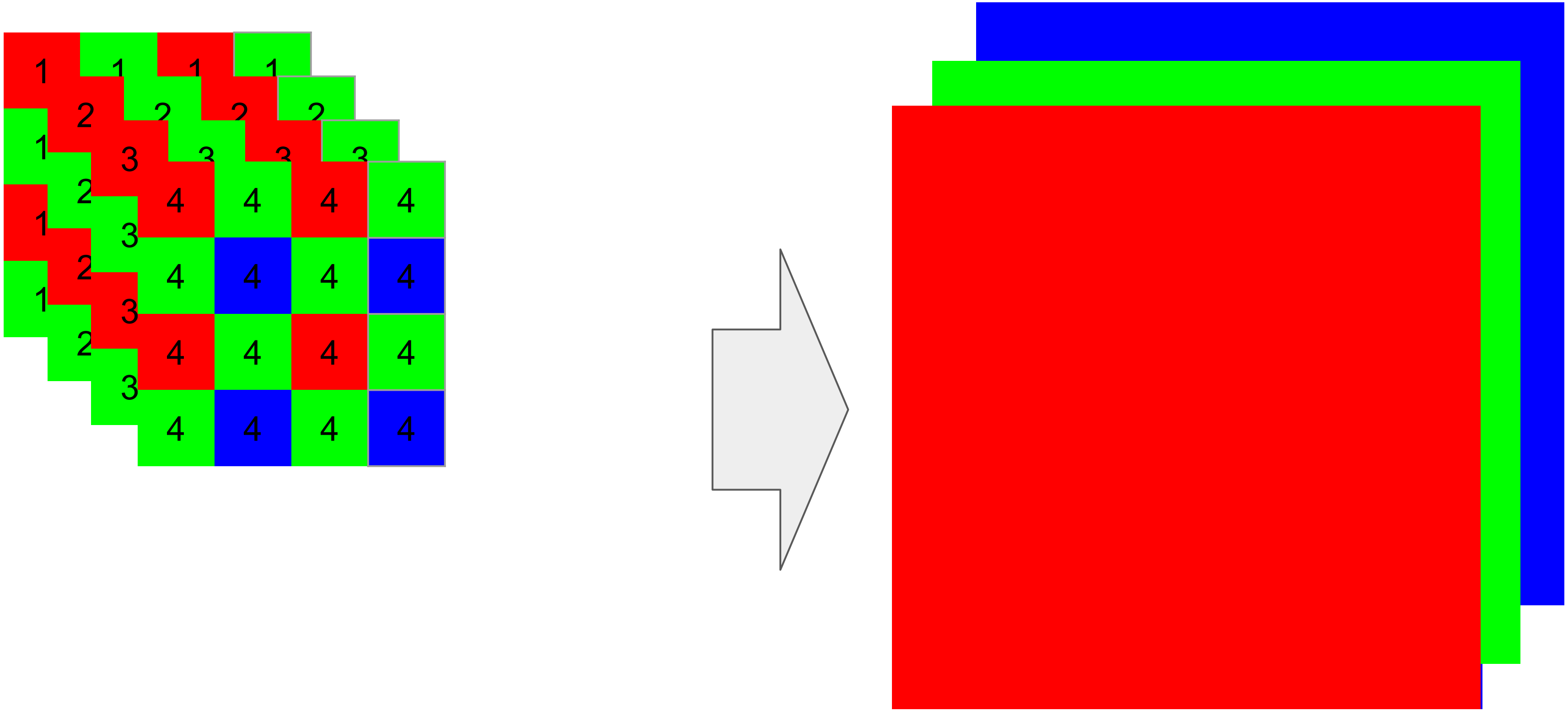




1 1 1 1
1 2 2 2 2
1 2 2 2 2
1 2 2 2 2
1 2 2 2 2
1 1 1 1
2 2 2
1 1 1 1
2 2 2
1 1 1
2 2 2
1 1 1
2 2 2



Multi-dimensional, non-uniform, interpolation

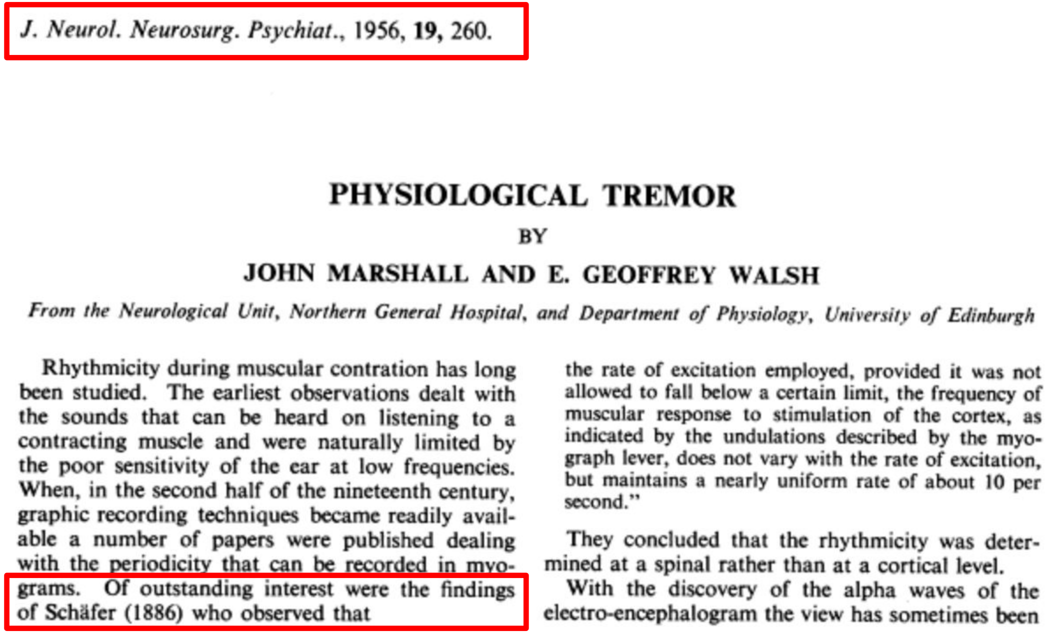


Source of **motion in mobile imaging**?

Handheld burst capture

After alignment: what's still moving?

(Natural) Physiological Tremor



J. Neurol. Neurosurg. Psychiat., 1956, **19**, 260.

PHYSIOLOGICAL TREMOR

BY

JOHN MARSHALL AND E. GEOFFREY WALSH

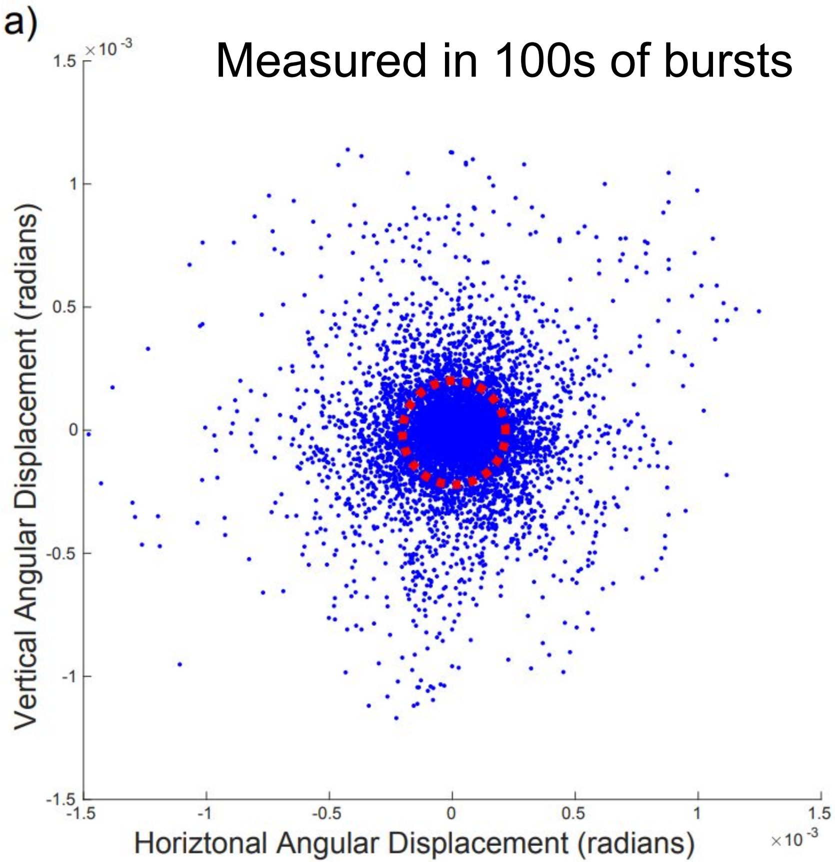
From the Neurological Unit, Northern General Hospital, and Department of Physiology, University of Edinburgh

Rhythmicity during muscular contration has long been studied. The earliest observations dealt with the sounds that can be heard on listening to a contracting muscle and were naturally limited by the poor sensitivity of the ear at low frequencies. When, in the second half of the nineteenth century, graphic recording techniques became readily available a number of papers were published dealing with the periodicity that can be recorded in myograms. Of outstanding interest were the findings of Schäfer (1886) who observed that

the rate of excitation employed, provided it was not allowed to fall below a certain limit, the frequency of muscular response to stimulation of the cortex, as indicated by the undulations described by the myograph lever, does not vary with the rate of excitation, but maintains a nearly uniform rate of about 10 per second."

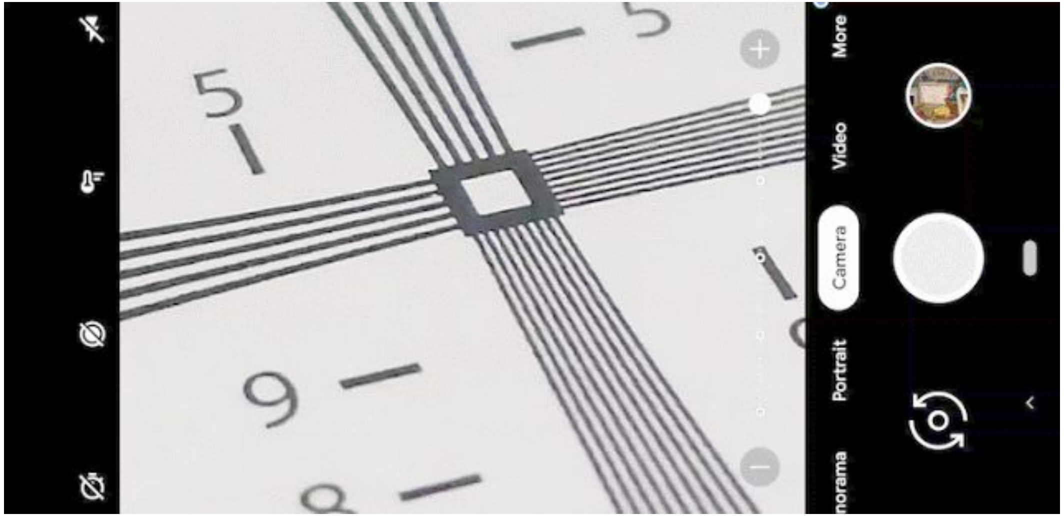
They concluded that the rhythmicity was determined at a spinal rather than at a cortical level.

With the discovery of the alpha waves of the electro-encephalogram the view has sometimes been



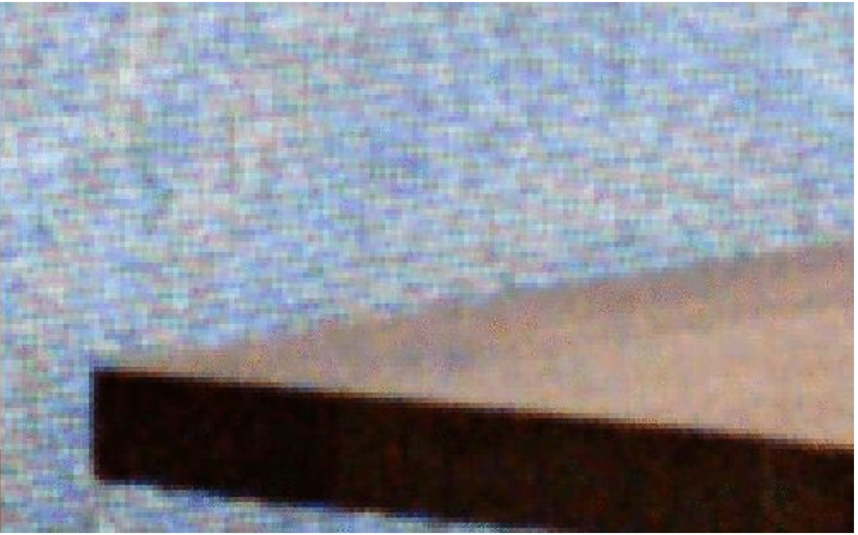
What if phone/camera is immobilized?

Simulated “tremor”



Motion : Phase Diversity

Aliasing + Phase diversity → Multi-frame Super-Res



Aliasing + Subpixel Motion

Super-res

The visual system appears to do super-resolution

(via micro-saccades)

Vol 447|14 June 2007|doi:10.1038/nature05866 nature

LETTERS

Miniature eye movements enhance fine spatial detail

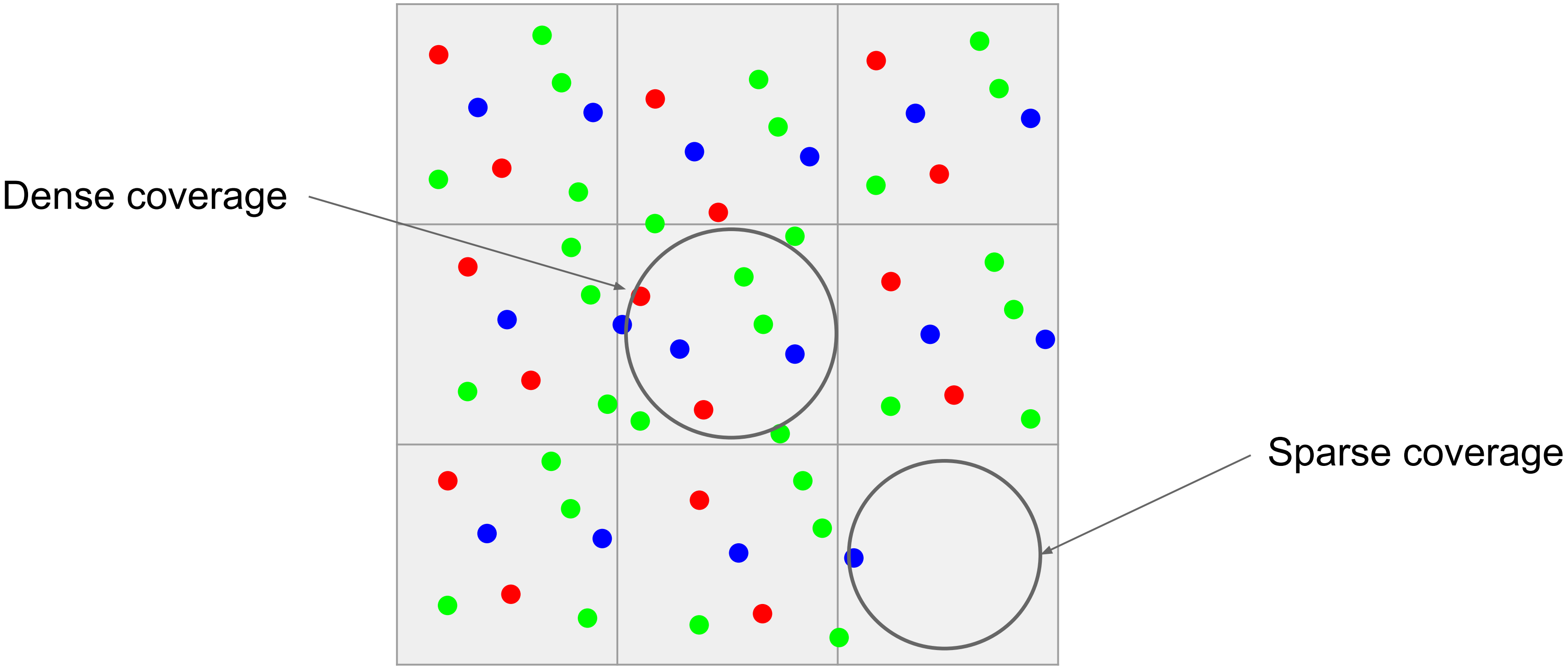
Michele Rucci[1], Ramon Iovin[1], Martina Poletti[1] & Fabrizio Santini[1]

Our eyes are constantly in motion. Even during visual fixation, small eye movements continually jitter the location of gaze[1–4]. It is known that visual percepts tend to fade when retinal image motion is eliminated in the laboratory[5–9]. However, it has long been debated whether, during natural viewing, fixational eye movements have functions in addition to preventing the visual scene from fading[10–17]. In this study, we analysed the influence in humans of fixational eye movements on the discrimination of gratings masked by noise that has a power spectrum similar to that of natural images. Using a new method of retinal image stabilization[18], we selectively eliminated the motion of the retinal image that normally occurs during the intersaccadic intervals of visual fixation. Here we show that fixational eye movements improve discrimination of high spatial frequency stimuli, but not of low spatial frequency stimuli. This improvement originates from the temporal modulations introduced by fixational eye movements in the visual input to the retina, which emphasize the high spatial frequency harmonics of the stimulus. In a natural visual world dominated by low spatial frequencies, fixational eye movements appear to constitute an effective sampling strategy by which the visual system enhances the processing of spatial detail.

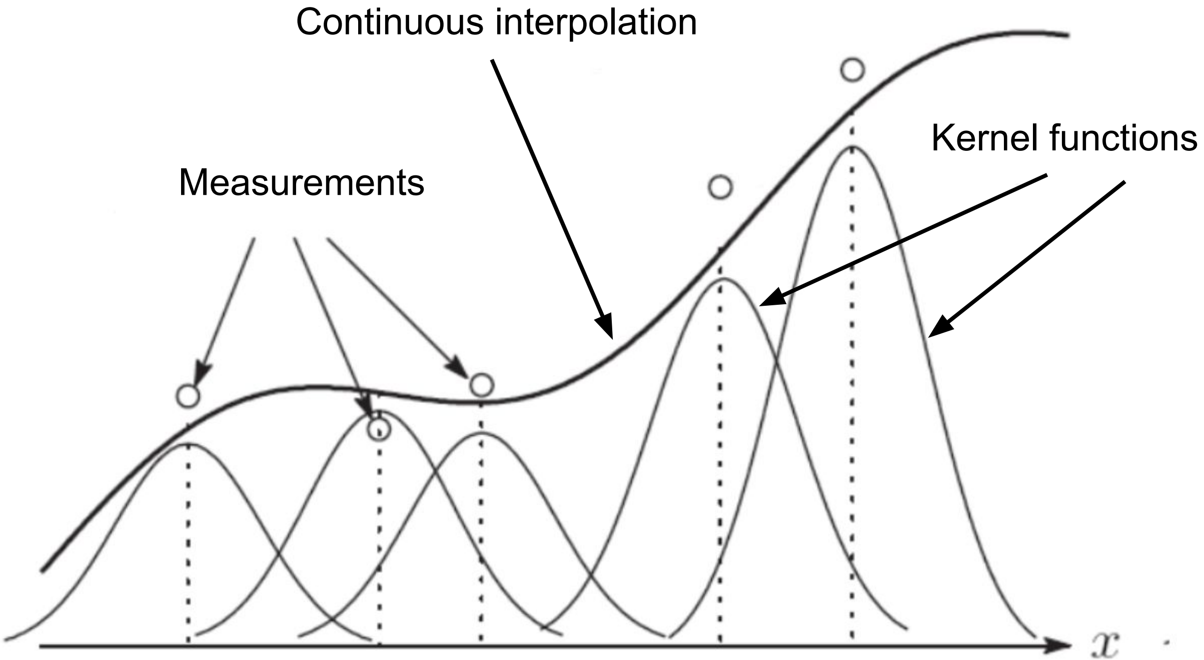
stabilization during periods of visual fixation between saccades, as would have been necessary to study fixational eye movements in their natural context[23–25]. Instead, all trials with stabilized vision had to be run in uninterrupted blocks while the subject maintained fixation—a highly unnatural condition that unavoidably led to visual fatigue and fading.

In this study, we examined the influence of fixational eye movements on the discrimination of targets at different spatial frequencies (grating spacings). We compared discrimination performances measured in two conditions: with and without the retinal image motion produced by fixational eye movements. To overcome the limitations of previous experiments, we developed a new retinal stabilization technique based on real-time processing of eye-movement signals[18]. Like previous stabilization methods, this technique does not guarantee perfect elimination of retinal image motion; however, unlike previous methods, it combines a high quality of stabilization with experimental flexibility (see Supplementary Information). This flexibility enabled us to display and selectively stabilize the stimulus after a saccade, a method that isolates the normal fixational motion of the eye. It also allowed us to randomly alternate between trials with retinal stabilization and trials with normal retinal motion, a procedure that

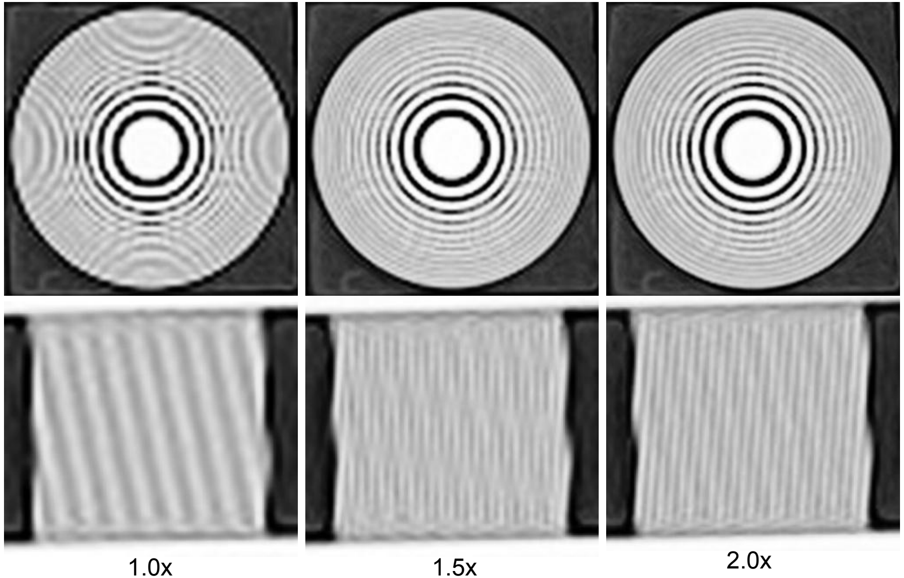
Non-uniform coverage



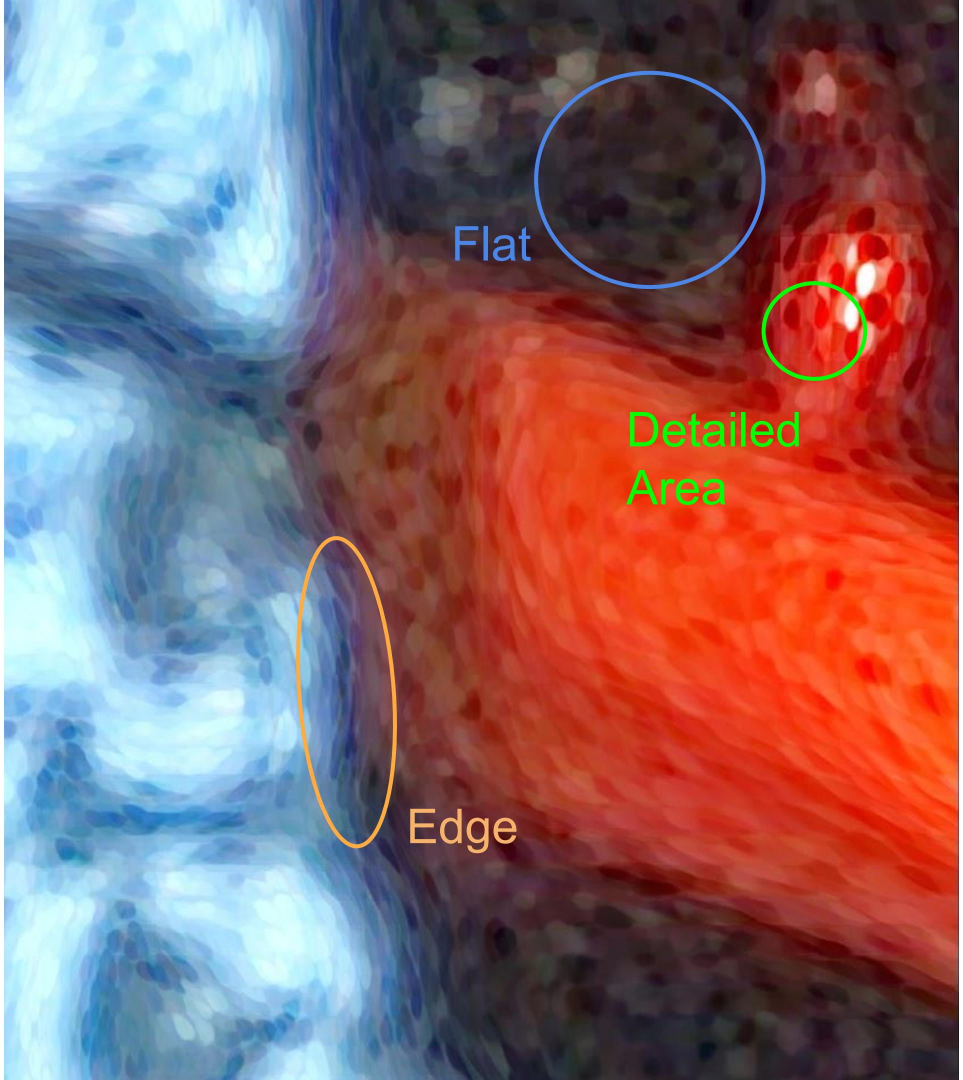
Merge: Nonlinear Kernel Regression



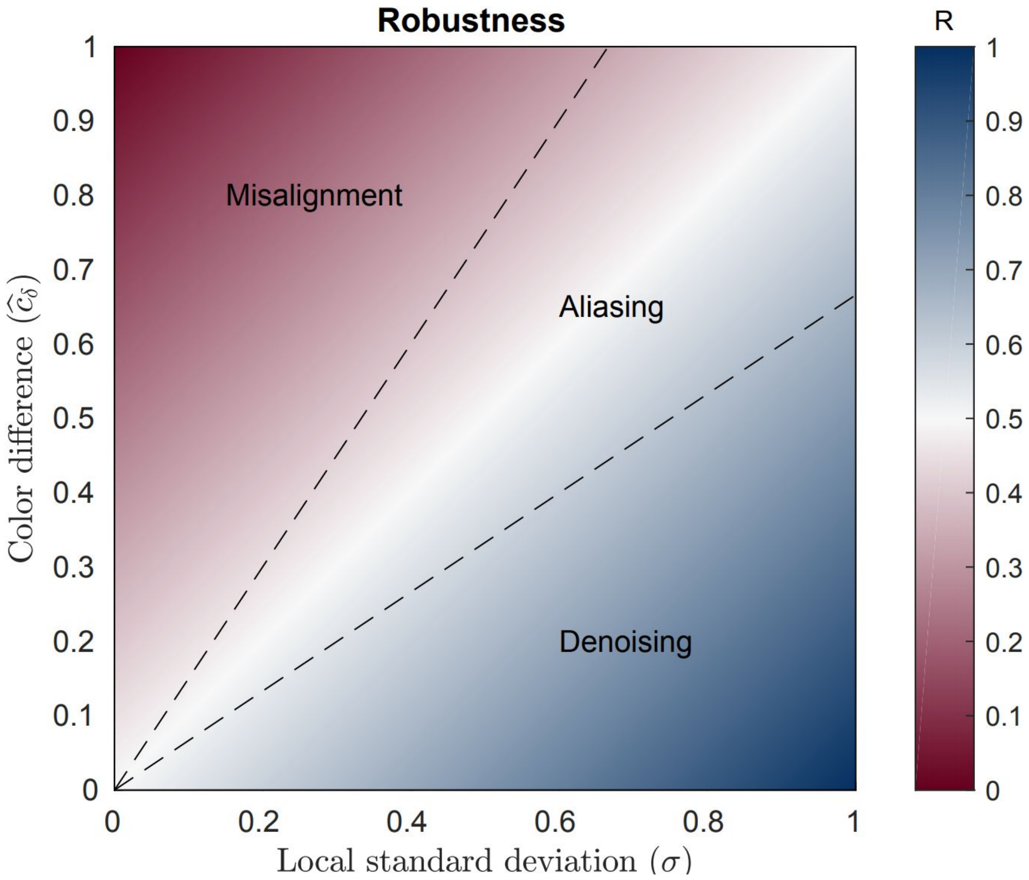
We can also merge onto higher-res grid



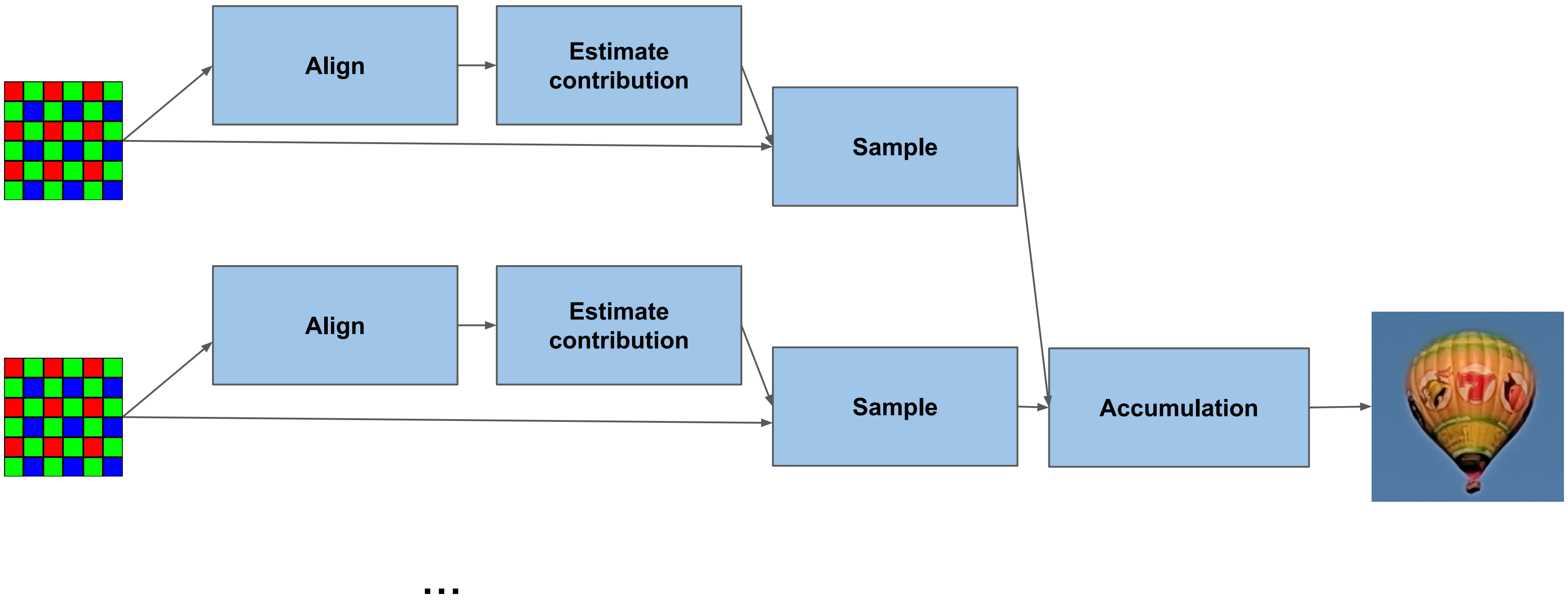
- This has its limits
  - depends on pixel/lens spot size tradeoff
  - for typical mobile sensors, limit is 2x


NO PARKING
518
B
Indochine Vegan Cuisine

Flat
Detailed
Area
Edge

Robustness model



Gather → Parallel Process



Crops

Full picture (reference)

Hasinoff et al. [2016]

Ours

Studio scene
Image comparison tool

DIGITAL PHOTOGRAPHY REVIEW
DPREVIEW

Lighting: Image size: FULL COMP PRINT

”The Pixel 3 is the first smartphone camera to rival cameras with Micro 4/3 sensors.”

6 mm < M4/3 17 mm < APS-C 21 mm < 35 mm

M4/3
17 mm

<

6 mm
+
Super-Resolution

<

APS-C
21 mm

<

35 mm

[SIGGRAPH 2019]

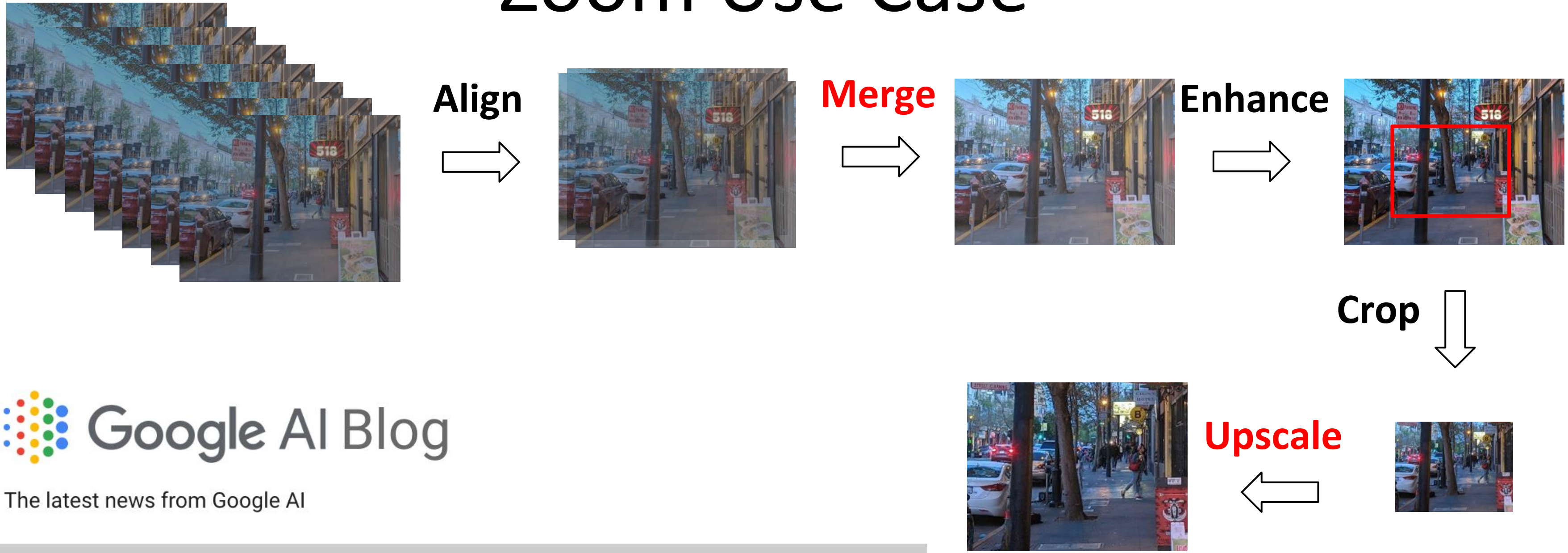
Handheld Multi-Frame Super-Resolution

BARTLOMIEJ WRONSKI, IGNACIO GARCIA-DORADO, MANFRED ERNST, DAMIEN KELLY, MICHAEL KRAININ, CHIA-KAI LIANG, MARC LEVOY, and PEYMAN MILANFAR, Google Inc.

Use Cases: Night Sight, Super-res Zoom


Night Sight

Zoom Use Case



Google AI Blog

The latest news from Google AI

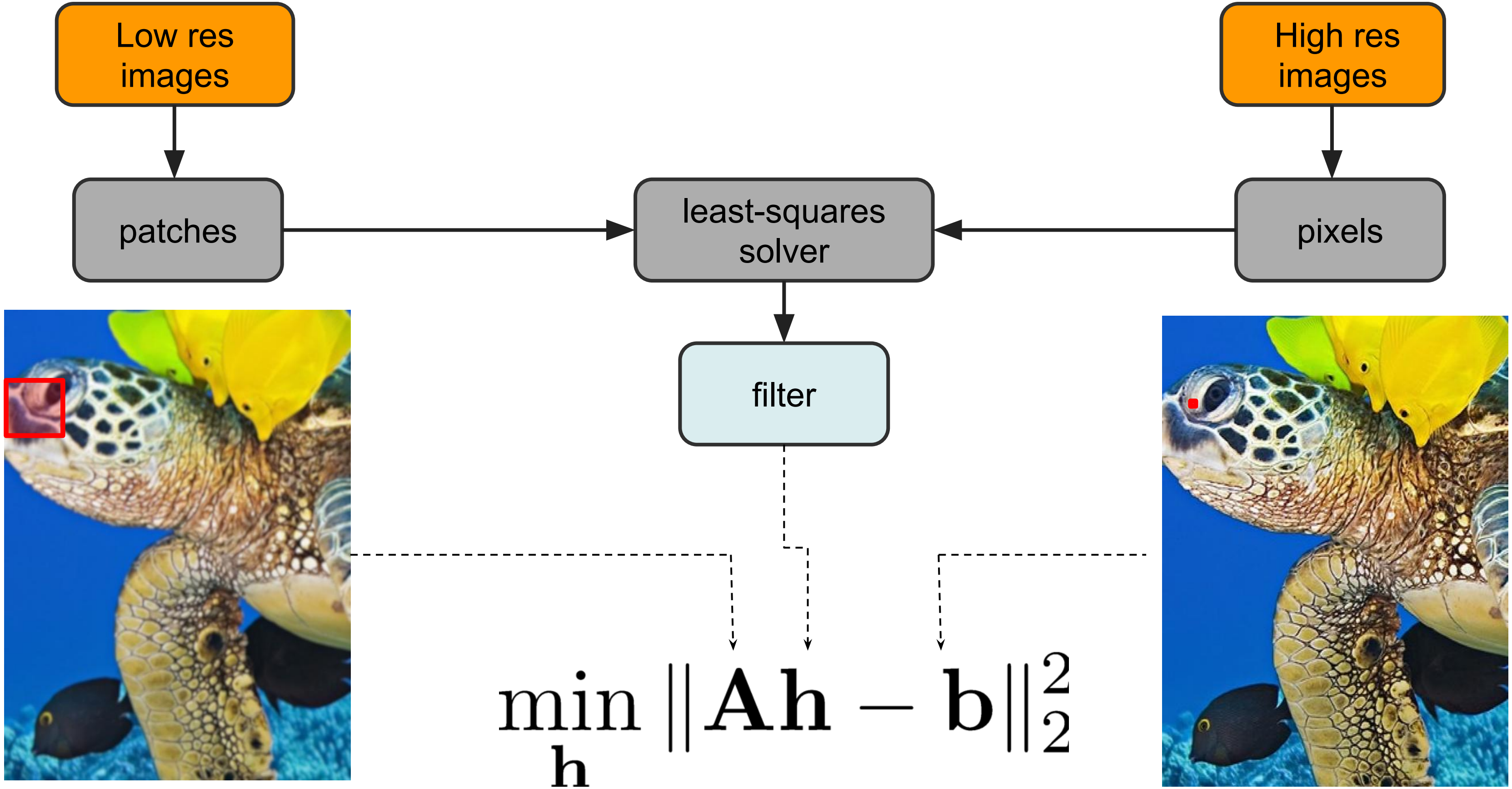
Enhance! RAISR Sharp Images with Machine Learning

Monday, November 14, 2016

Posted by Peyman Milanfar, Research Scientist

[Romano, Milanfar, Isidoro, Transactions on Computational Imaging, 2017]

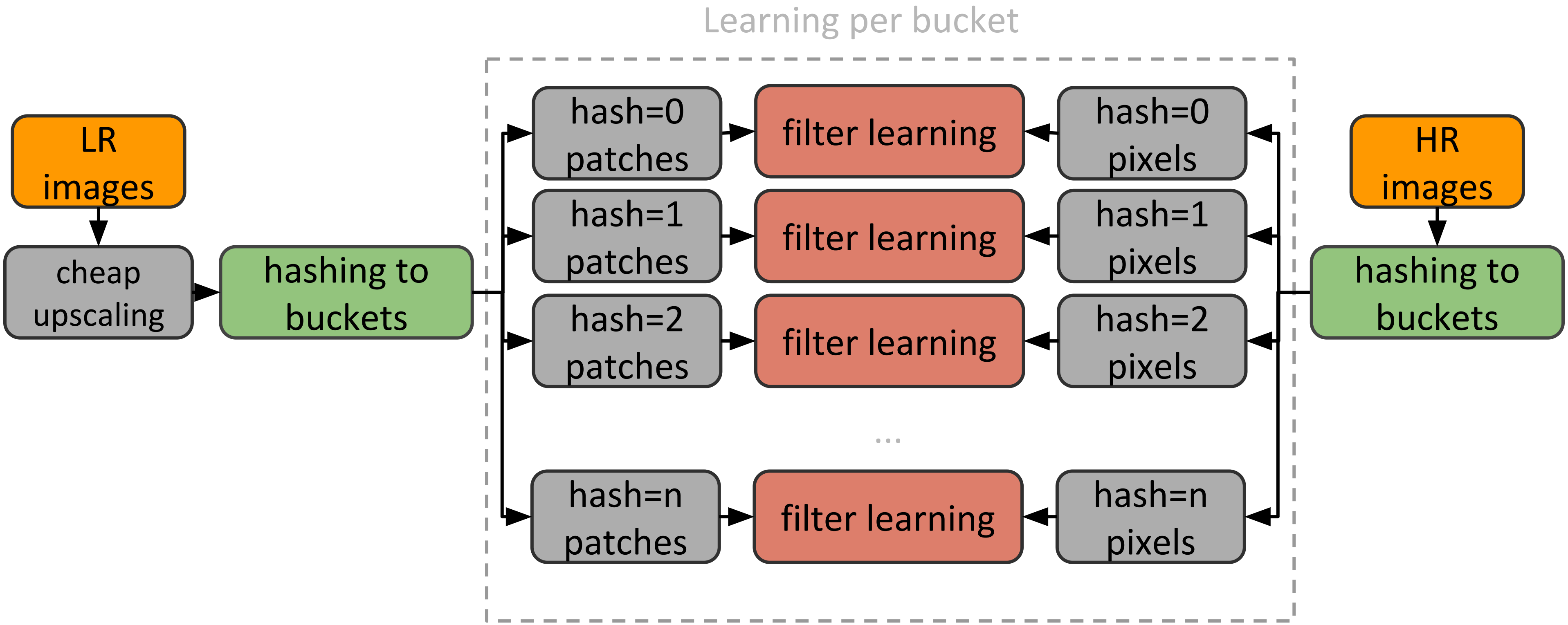
Filter Learning



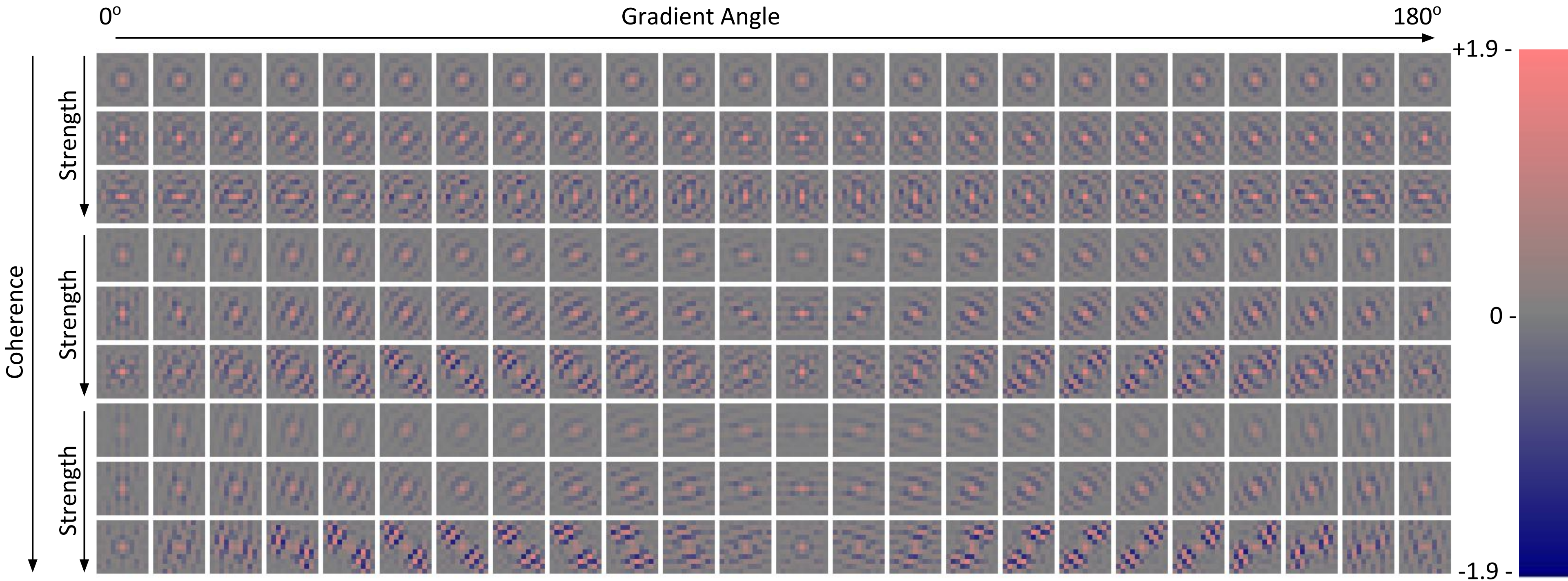
$$\min_{\mathbf{h}} \|\mathbf{A}\mathbf{h} - \mathbf{b}\|_2^2$$

We can do even better

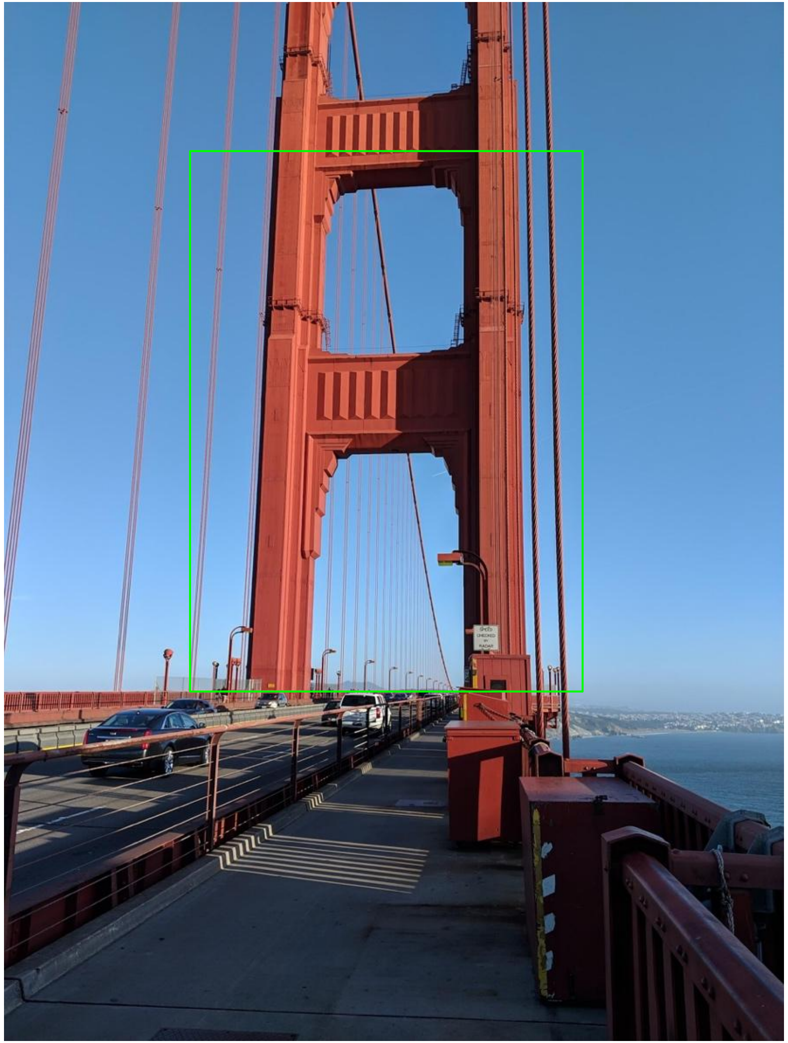
- Bucket similar patches together and train within buckets



Learned 2x Upscaling Filters



No zoom



(2x zoom)

(2x zoom crop)

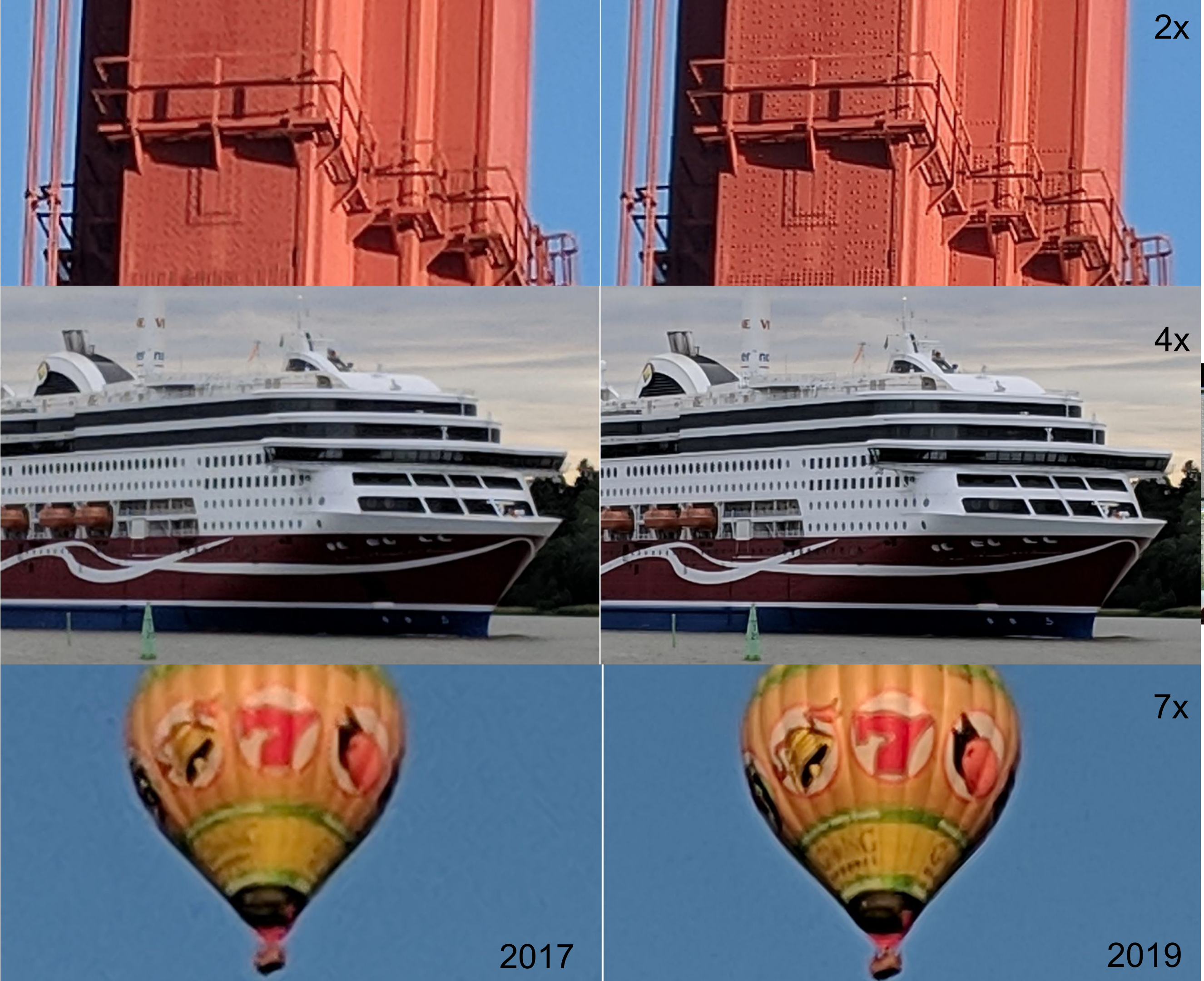
Standard Digital Zoom

(2x zoom crop)

Single-frame Super-res

(2x zoom crop)

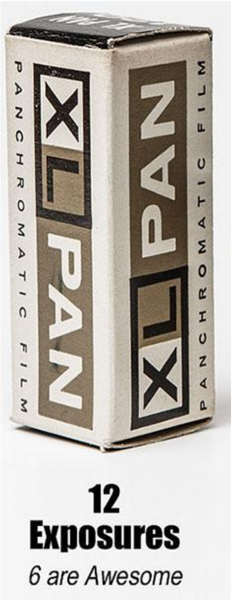
Multi-frame Super-res



"Best hybrid (optical/digital) zoom on the market"

OTHER CHALLENGES IN COMPUTATIONAL IMAGING

Curation



Google AI Blog

e latest news from Google AI

troducing NIMA: Neural Image Assessment

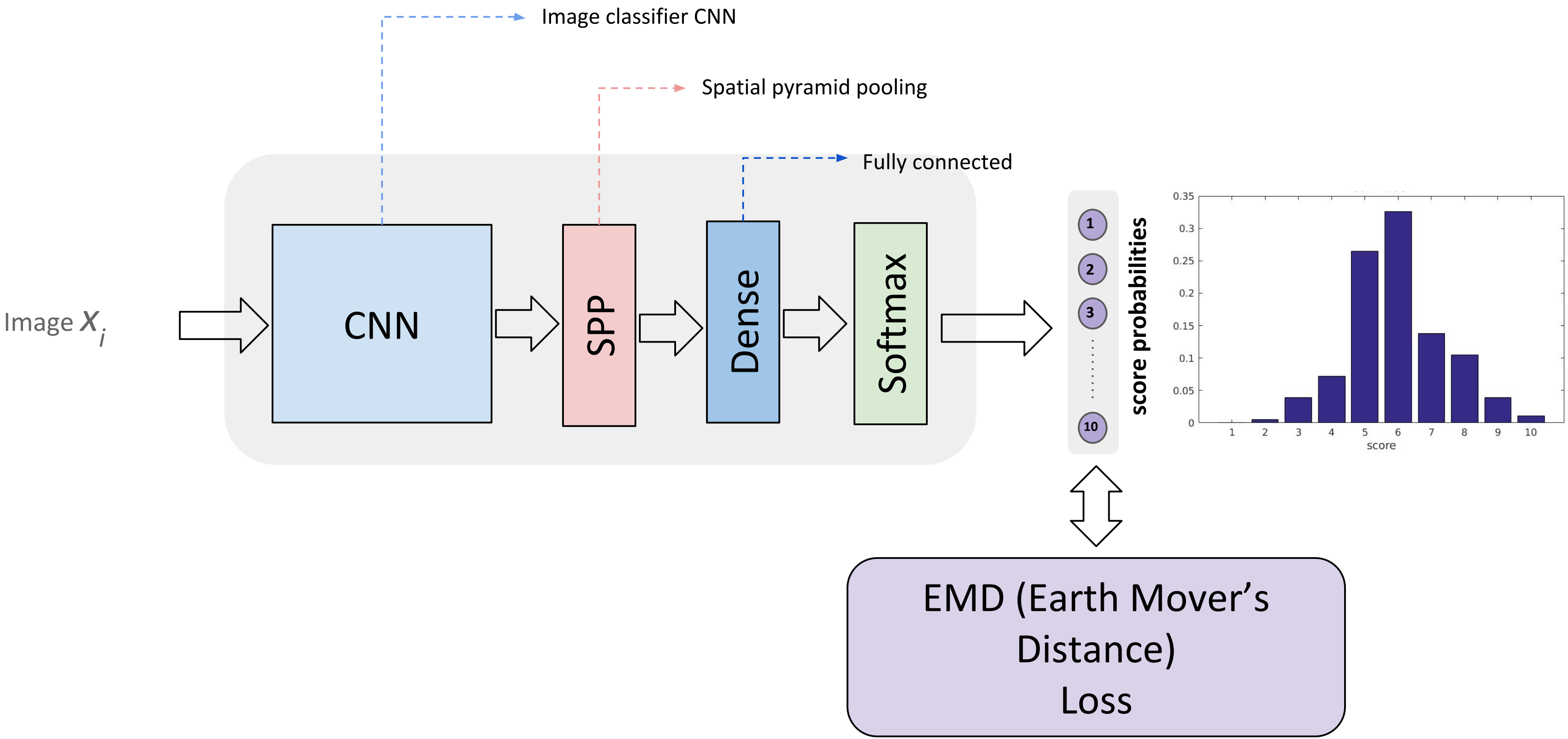
onday, December 18, 2017

sted by Hossein Talebi, Software Engineer and Peyman Milanfar Research Scientist, Machine
rception

antification of image quality and aesthetics has been a long-standing problem in image
ocessing and computer vision. While technical quality assessment deals with measuring pixel-
el degradations such as noise, blur, compression artifacts, etc., aesthetic assessment captures
mantic level characteristics associated with emotions and beauty in images. Recently, deep
nvolutional neural networks (CNNs) trained with human-labelled data have been used to address
subjective nature of image quality for specific classes of images, such as landscapes. However,
ese approaches can be limited in their scope, as they typically categorize images to two classes
low and high quality. Our proposed method predicts the distribution of ratings. This leads to a
more accurate quality prediction with higher correlation to the ground truth ratings, and is
applicable to general images.

H. Talebi and P. Milanfar, "NIMA: Neural Image Assessment", *IEEE Transactions on Image Processing 2018.*

NIMA: Neural Image Assessment



H. Talebi and P. Milanfar, "[NIMA: Neural Image Assessment]()", In *IEEE Transactions on Image Processing 2018.*

NIMA for Aesthetic Quality


6.229


6.225


5.729

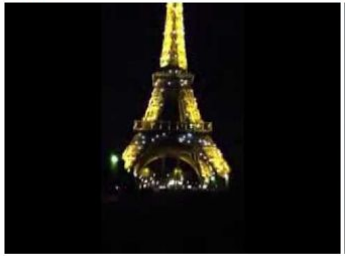

5.614


5.133


5.083

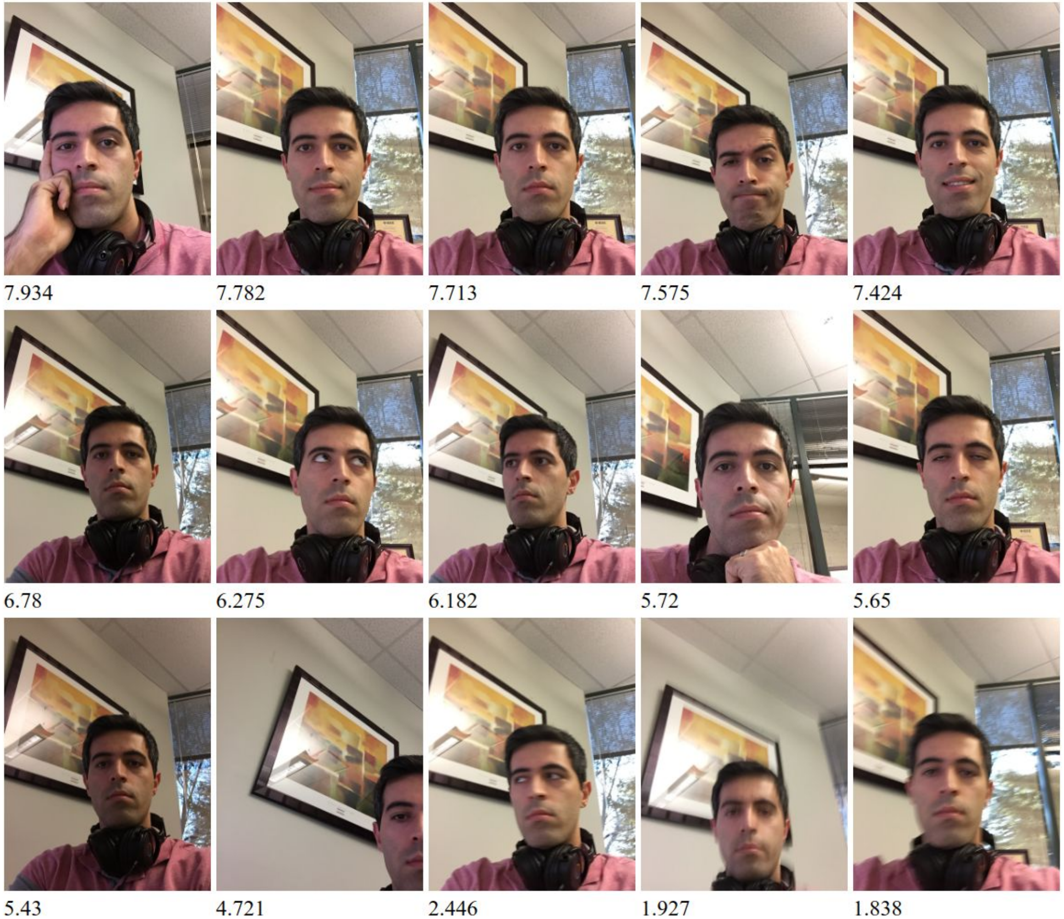

4.725


4.376


3.254


3.12

NIMA For Technical Quality



peyman.milanfar@gmail.com

http://www.milanfar.org