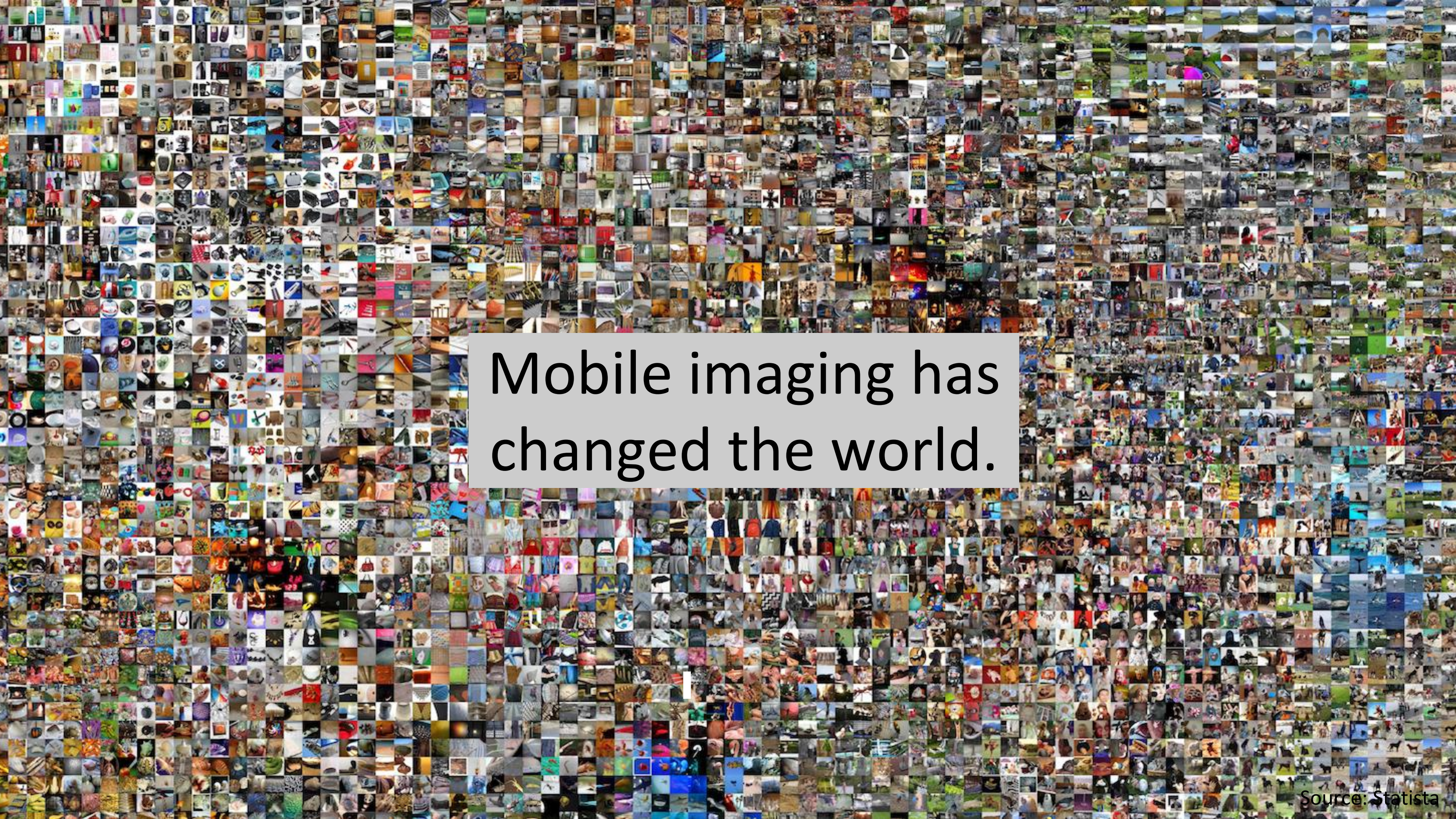
Computation + Photography

How the mobile phone became a camera

Peyman Milanfar

Google AI

Mobile imaging has changed the world.



Source: Statista

Vatican Square

Pope Benedict announcement

Pope Francis announcement

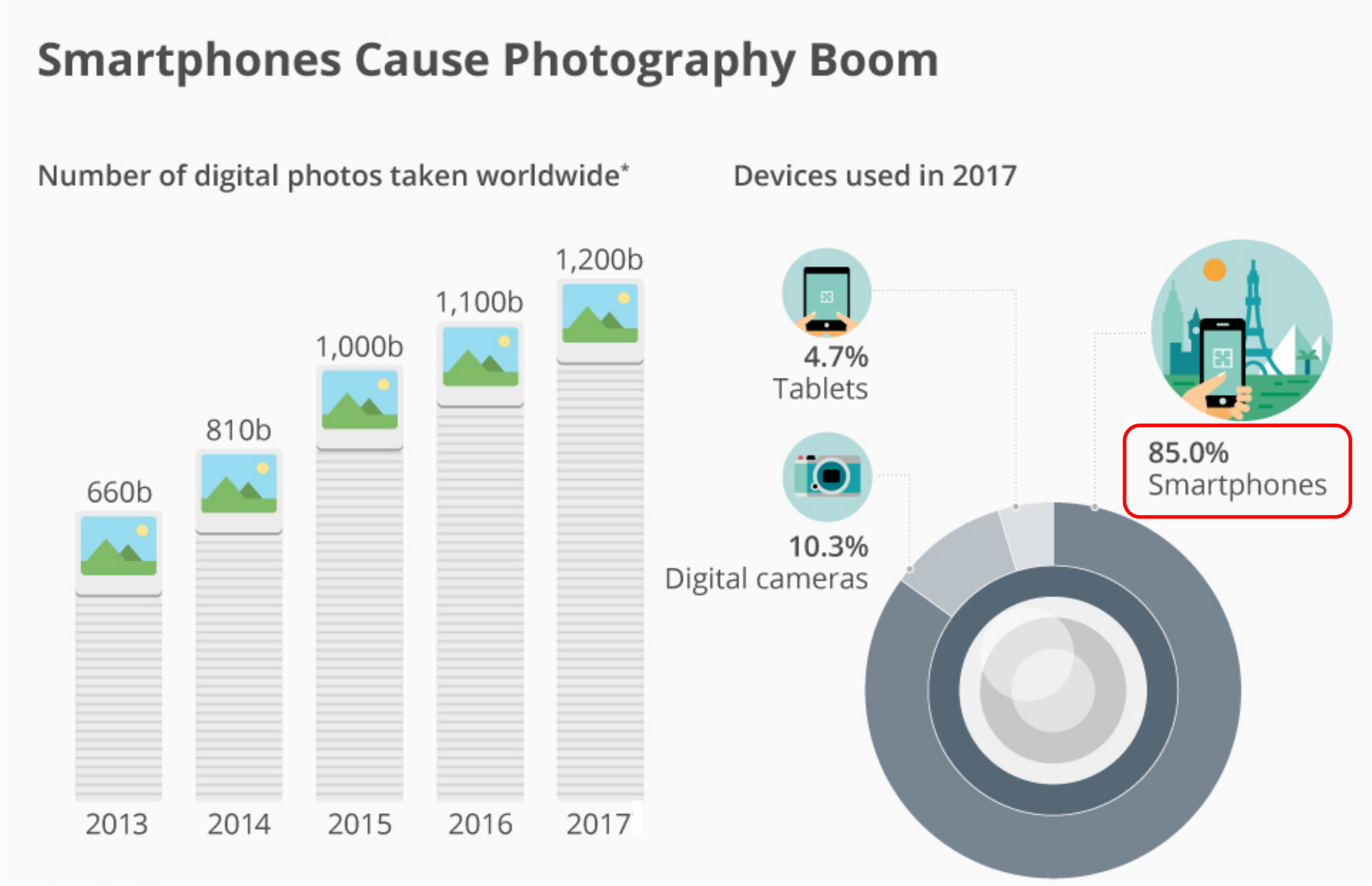
More than 2 billion photos shared on social media *per day*

That's 23,000 frames/sec

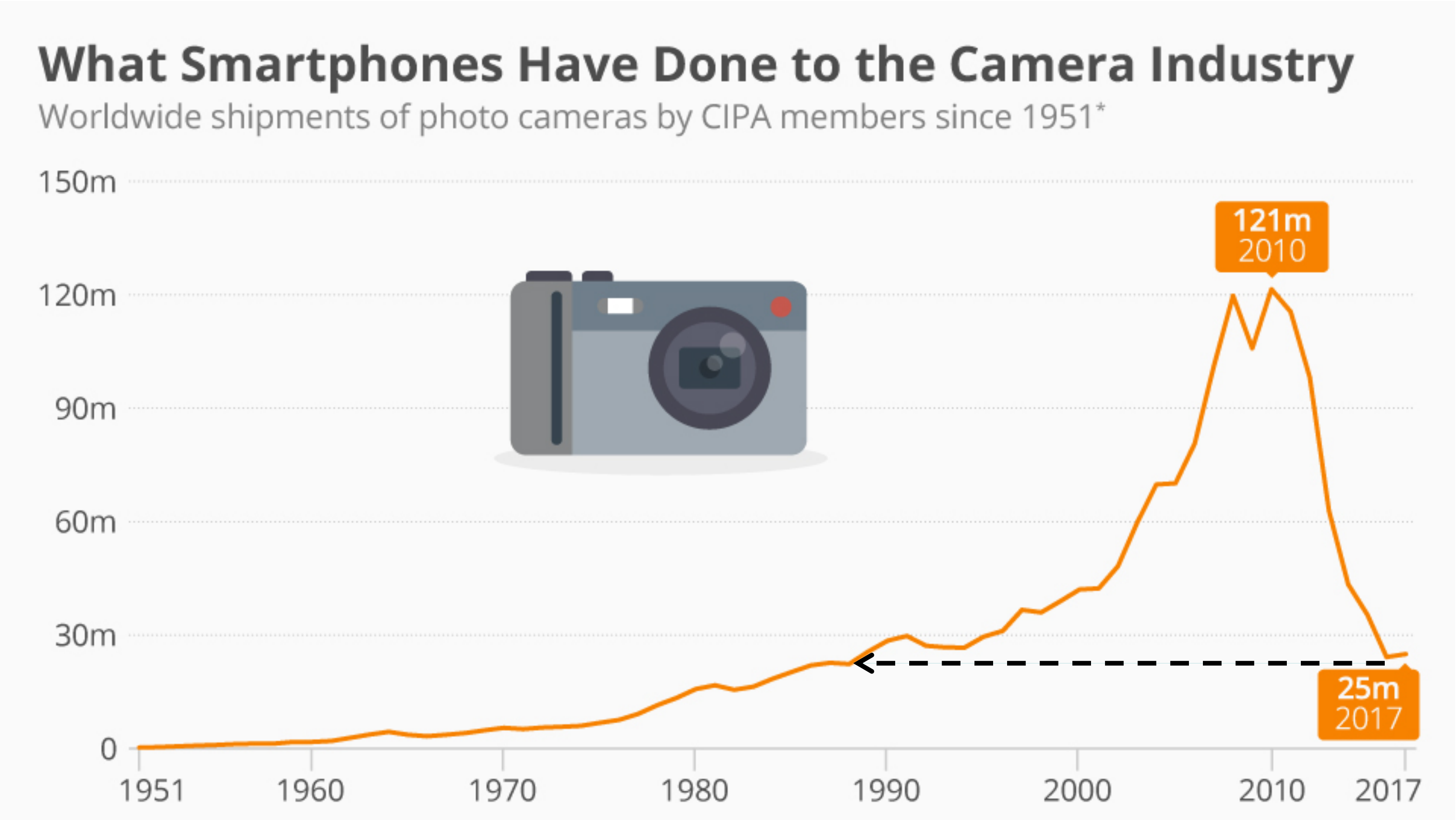
Over 100 million are "selfies"

That's 1,200 frames/sec

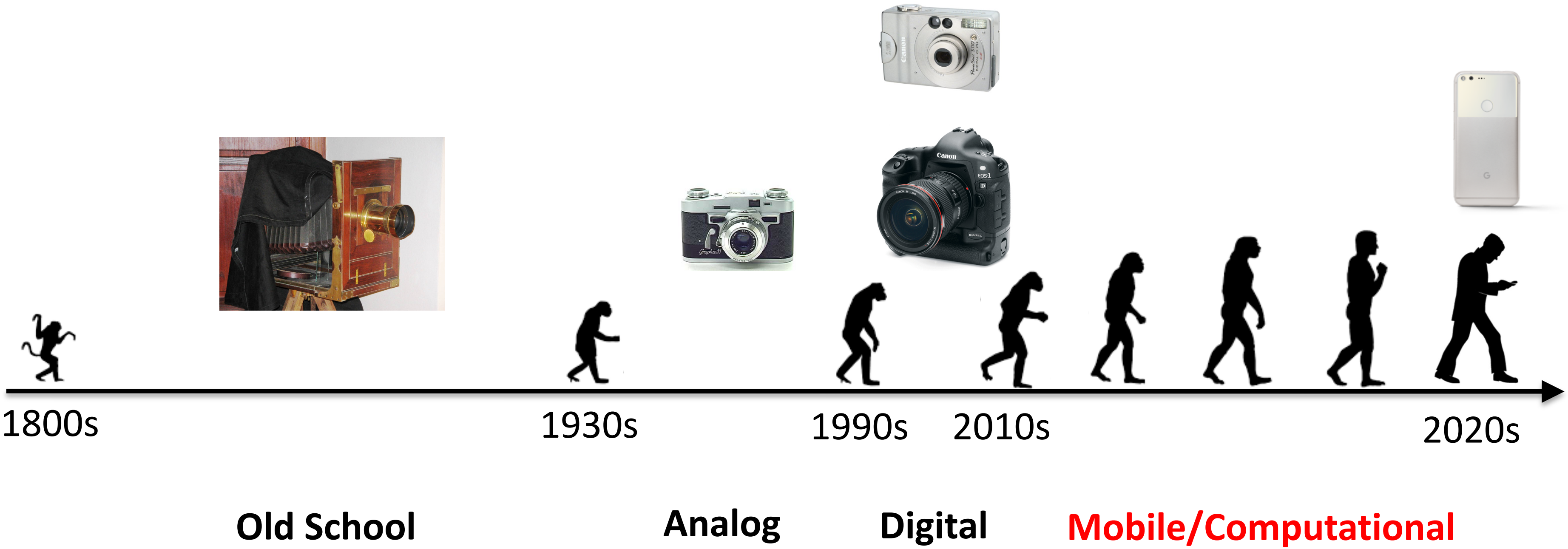
Reuters/CBS/TIME/KPCB

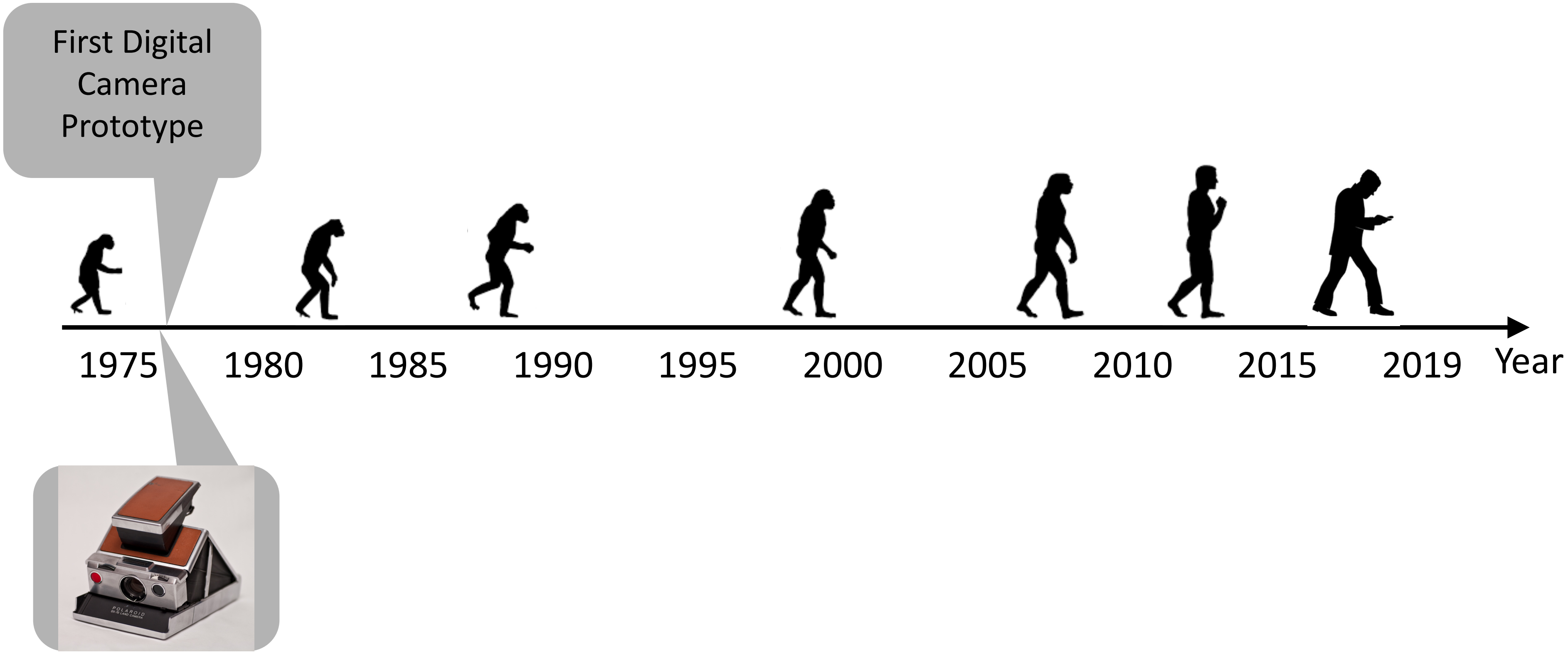


Source: Statista



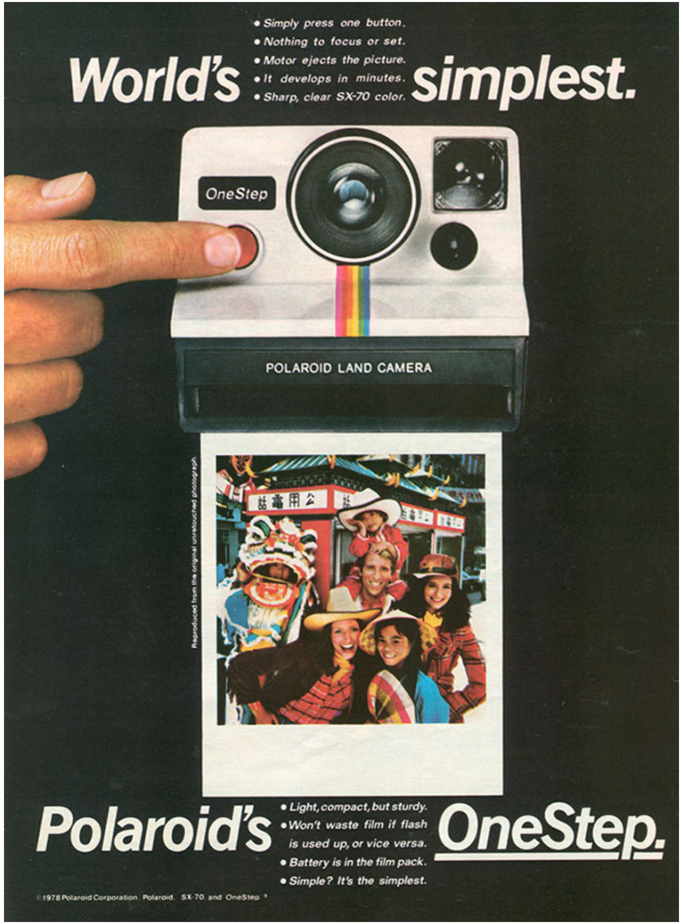
Source: Statista


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

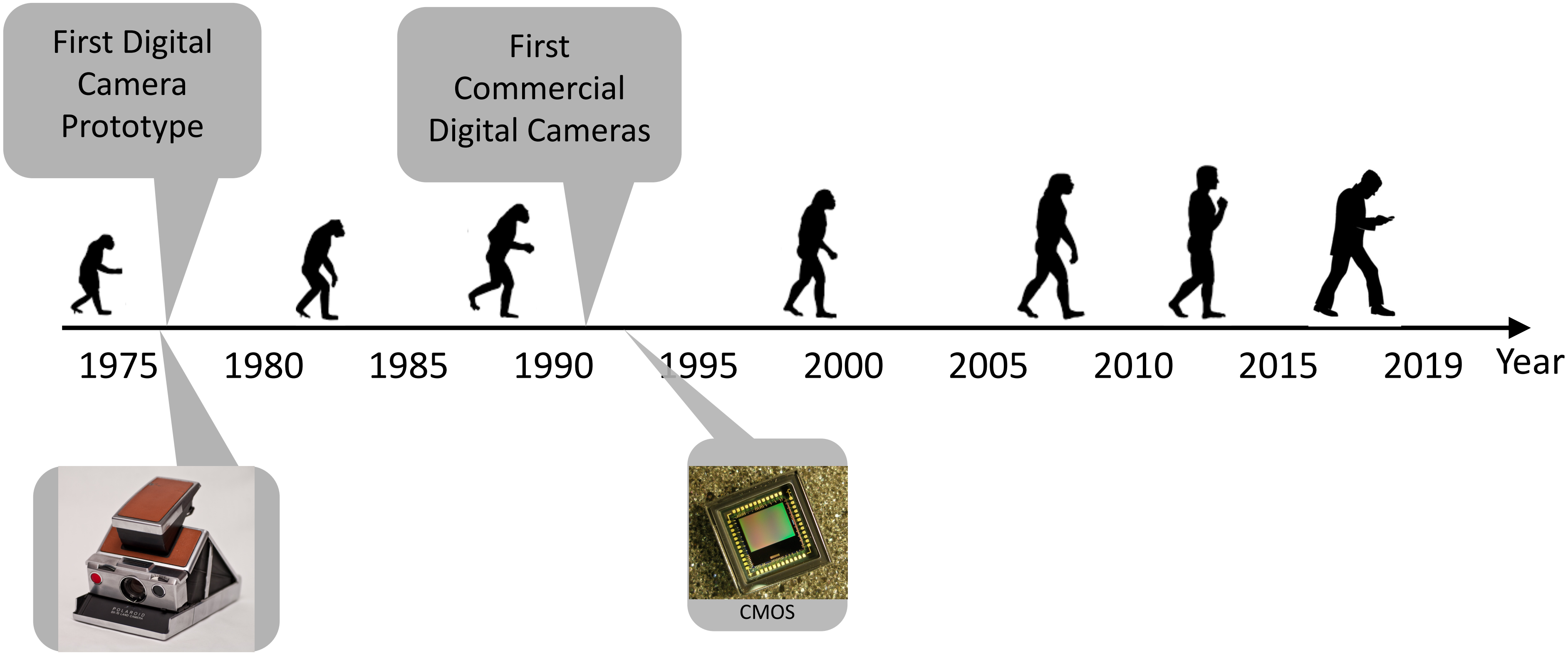


Polaroid SX-70
1972

Instant Gratification



Circa 1978


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

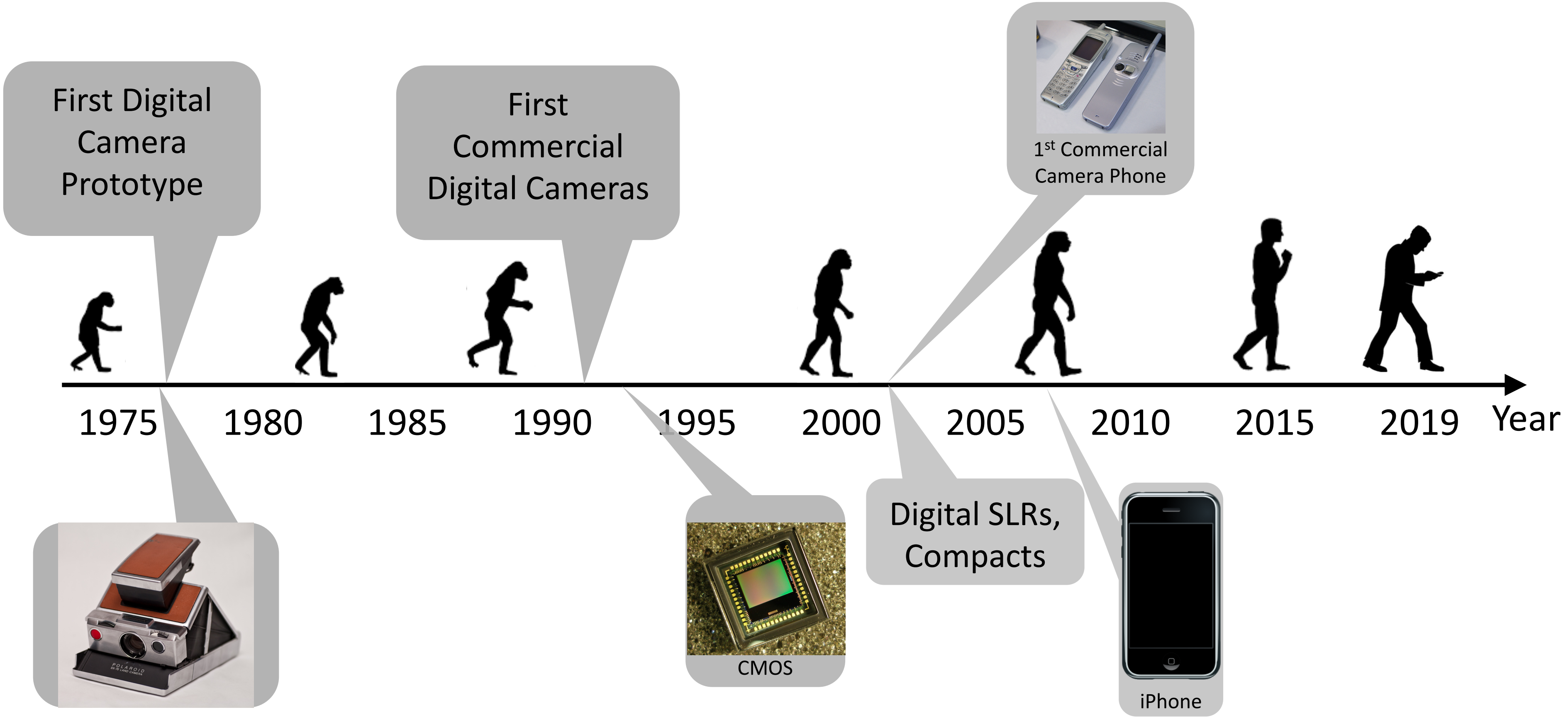
Invention of CMOS/Camera on a Chip

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

+ Cheaper, power efficient
- Noisier, rolling shutter readout

"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.

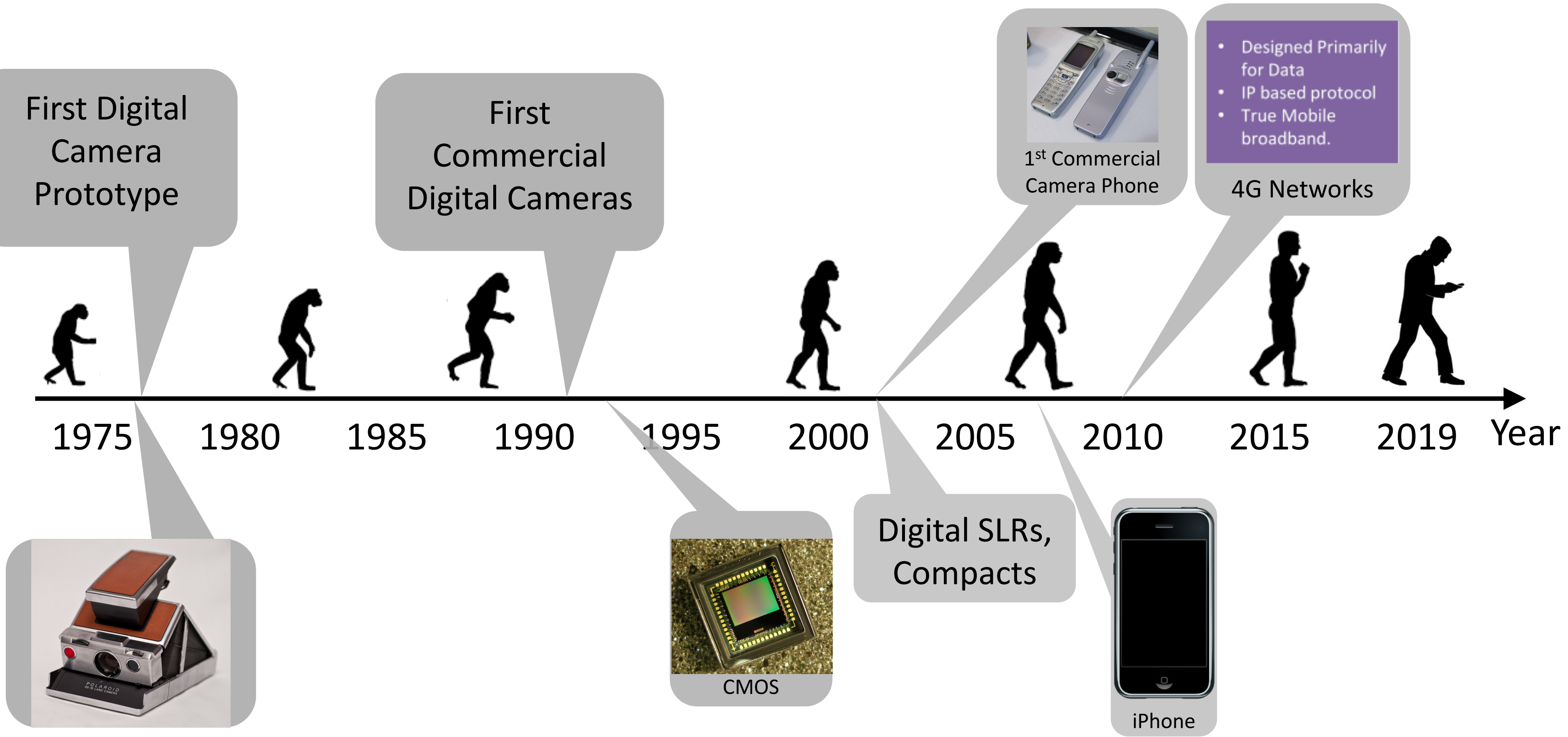
MOBILE PHOTOGRAPHY


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

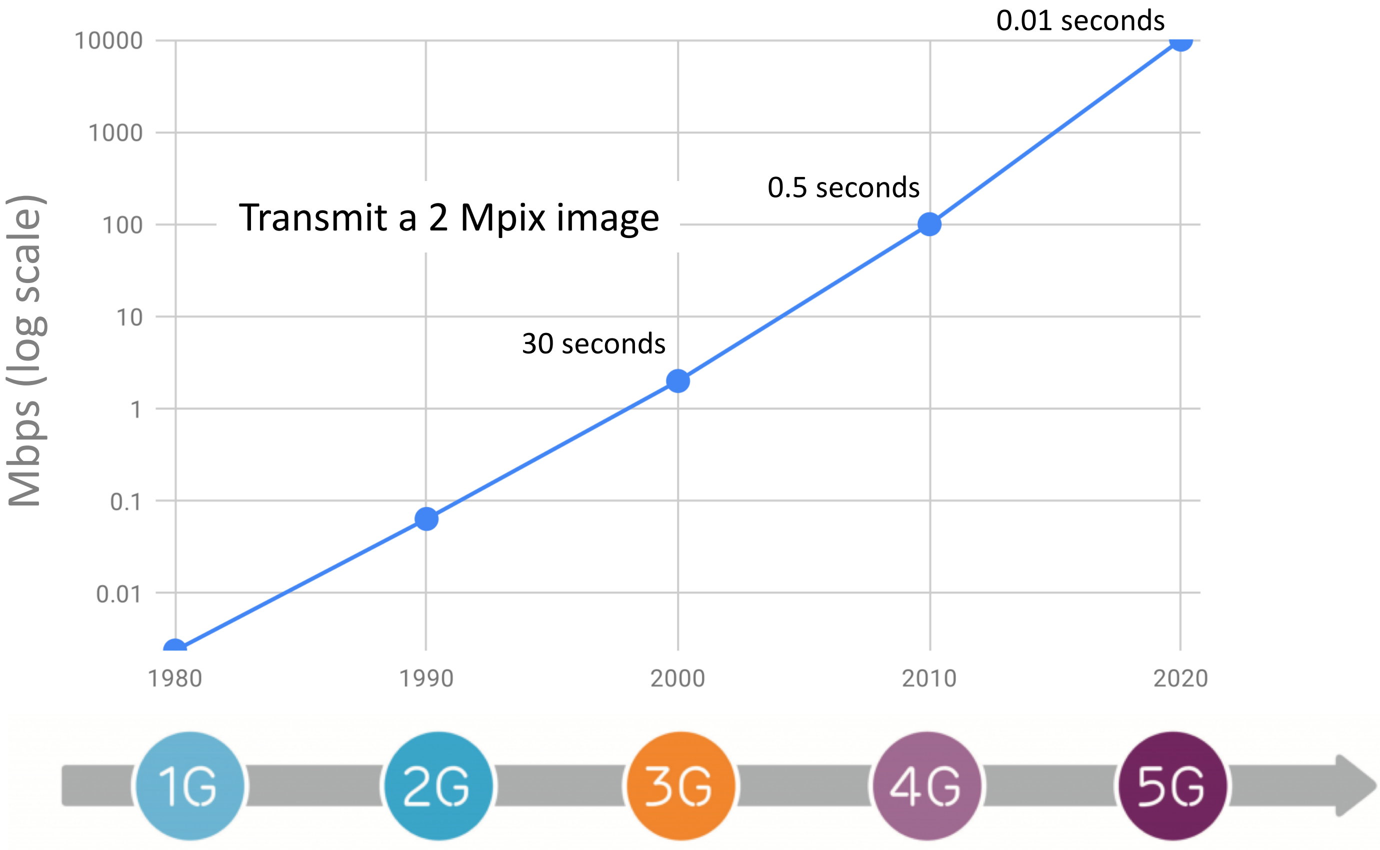
Displays

1990 2010 2020

300dpi displays


First Digital Camera Prototype
First Commercial Digital Cameras
1st Commercial Camera Phone
Designed Primarily for Data
IP based protocol
True Mobile broadband.
4G Networks
1975
1980
1985
1990
1995
2000
2005
2010
2015
2019
Year
CMOS
Digital SLRs, Compacts
iPhone

Wireless Network Speed



Source: Silika

2010 -

COMPUTATIONAL PHOTOGRAPHY

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

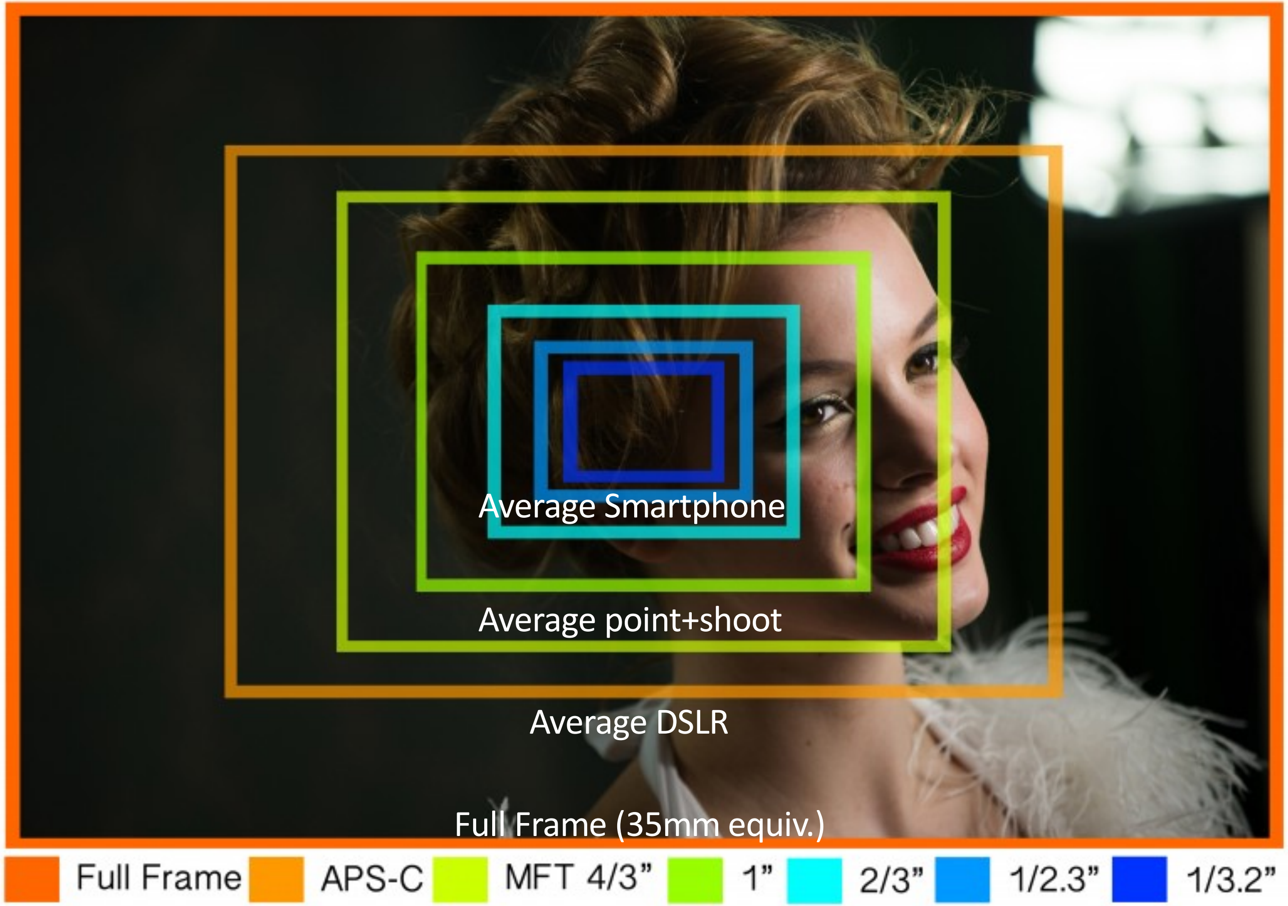
Can one be as good as the other?

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.

Modern Mobile Imaging: Burst Photography

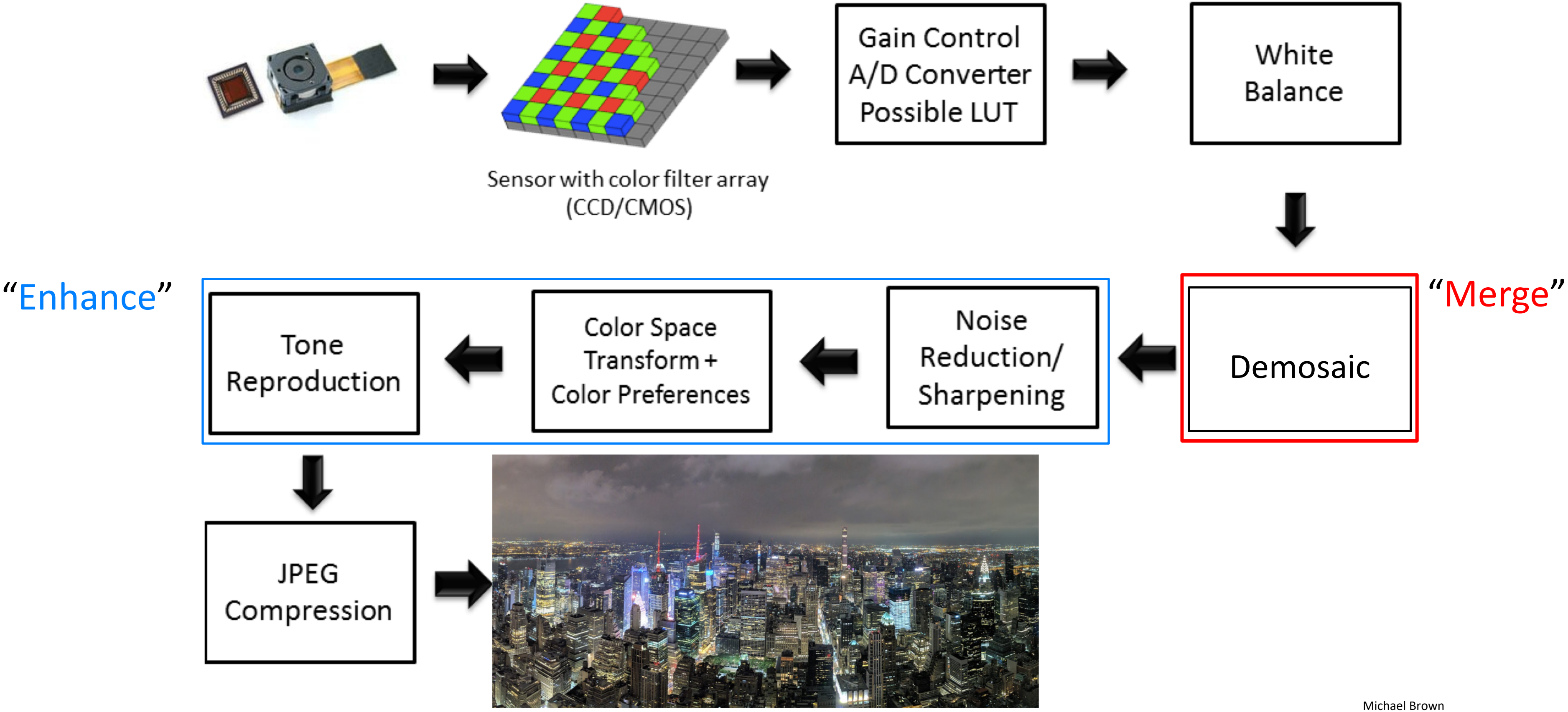
Exposure control

Align: Reliable Optical Flow – Scene is never stationary

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

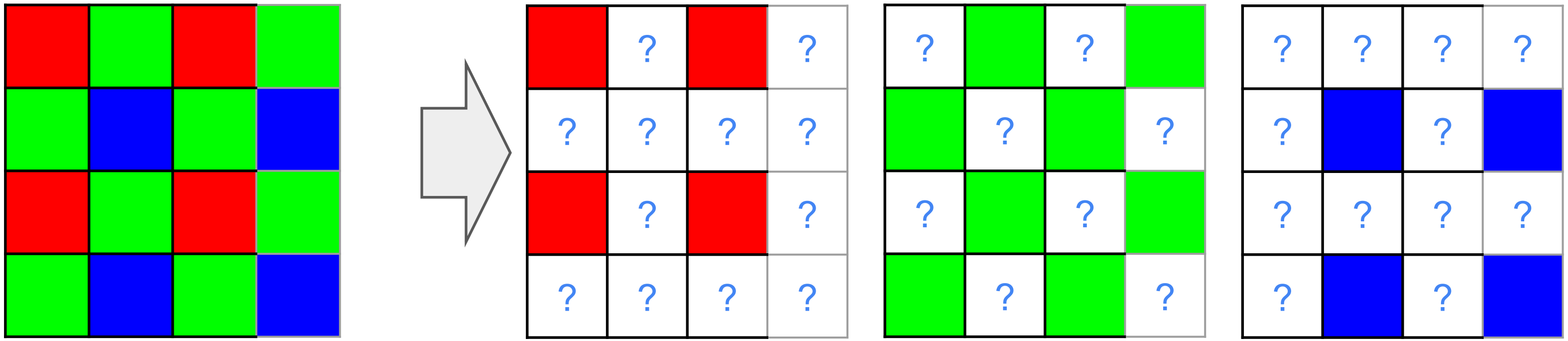
Enhance: Denoise, Sharpen, Contrast, Dynamic Range

Classic Camera Image Processing Pipeline

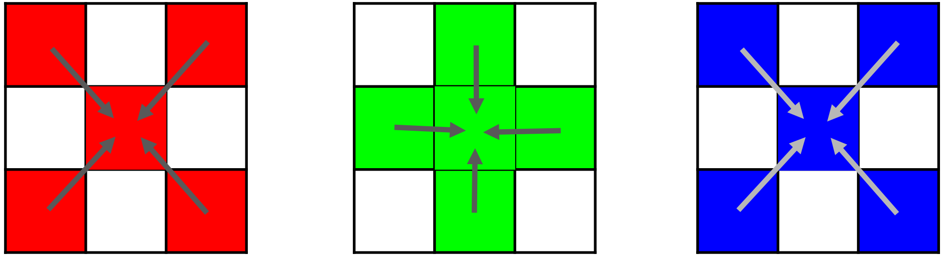


Michael Brown

Classic camera pipeline - demosaicing



Missing information



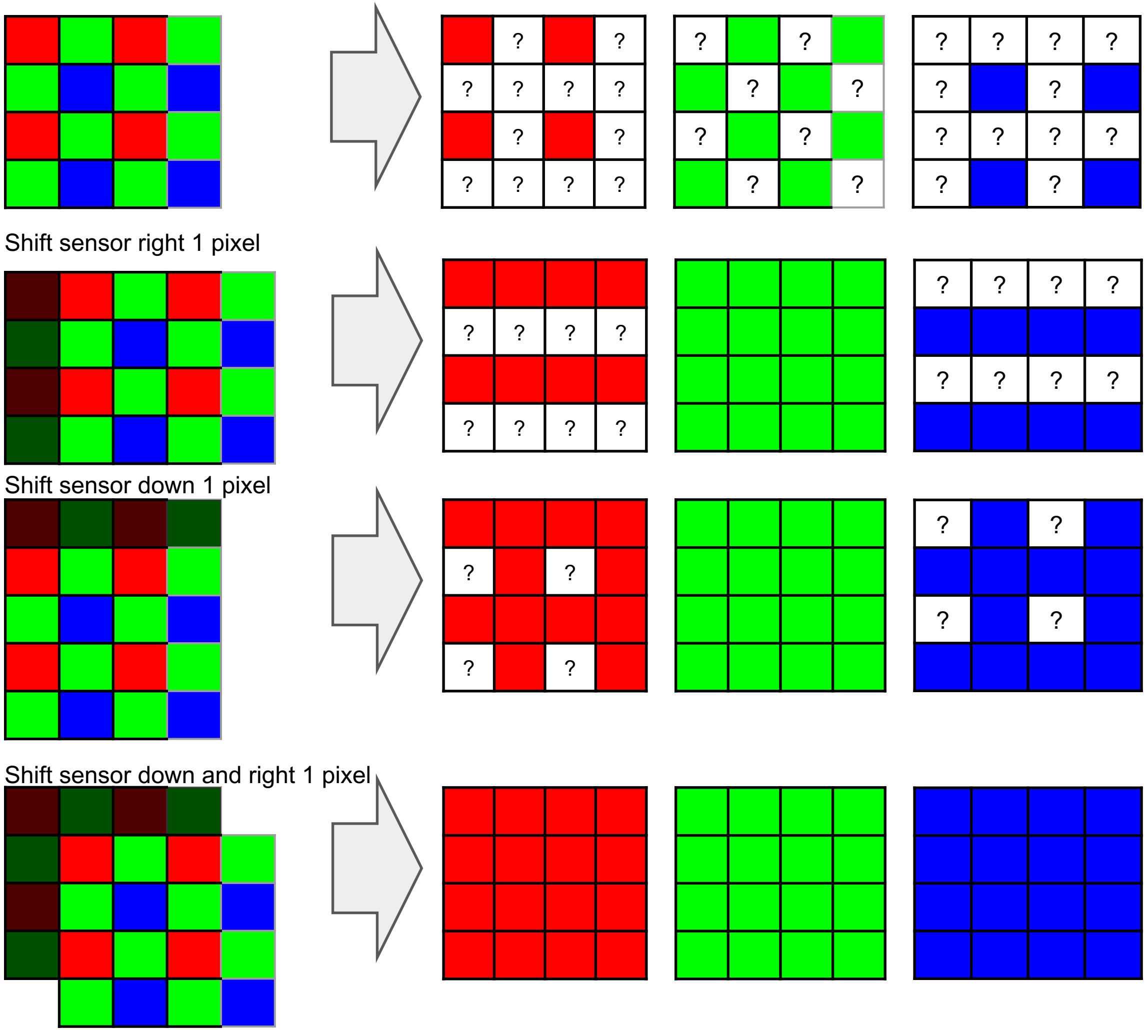
Two-thirds of your picture is made-up!

Demosaicing

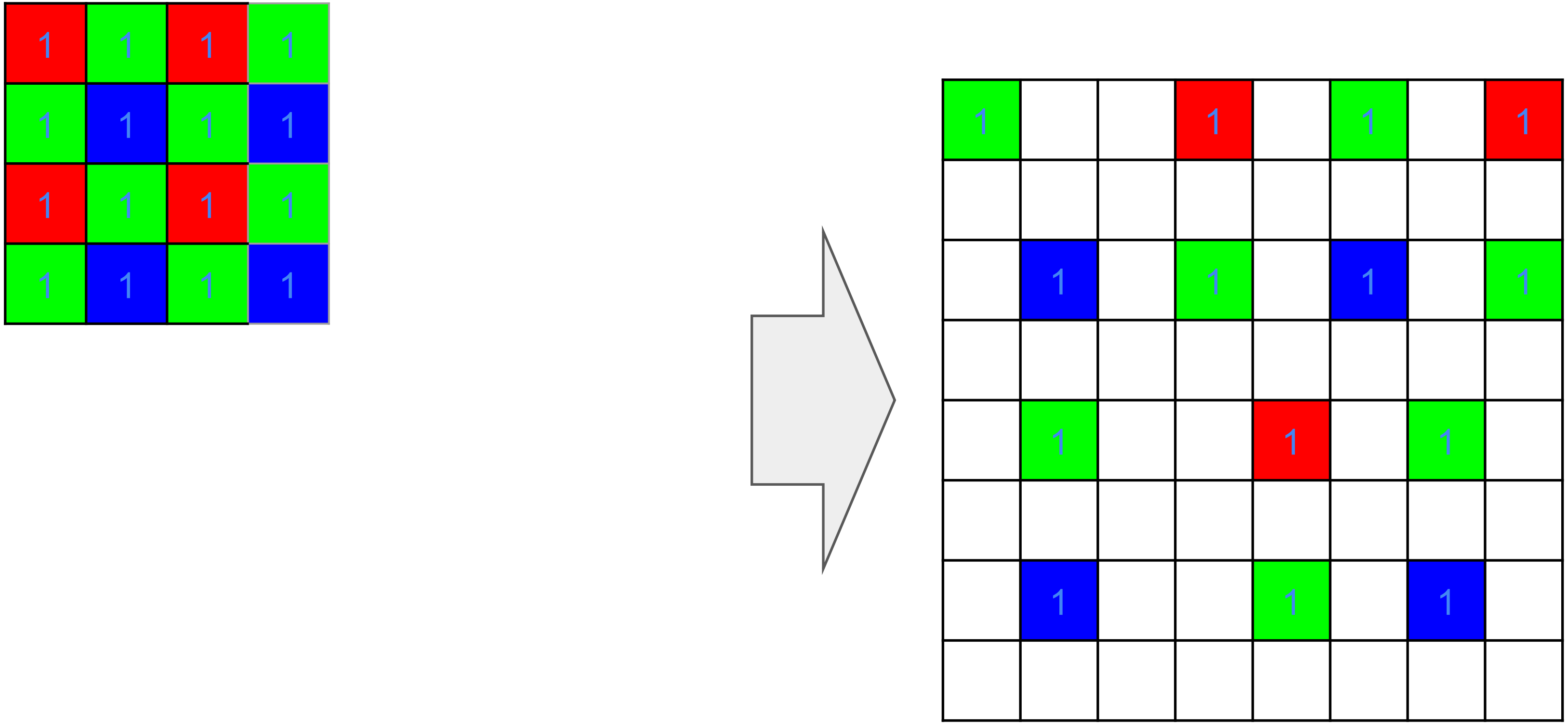
Demosaicing .... Kills Details and Produces Artifacts

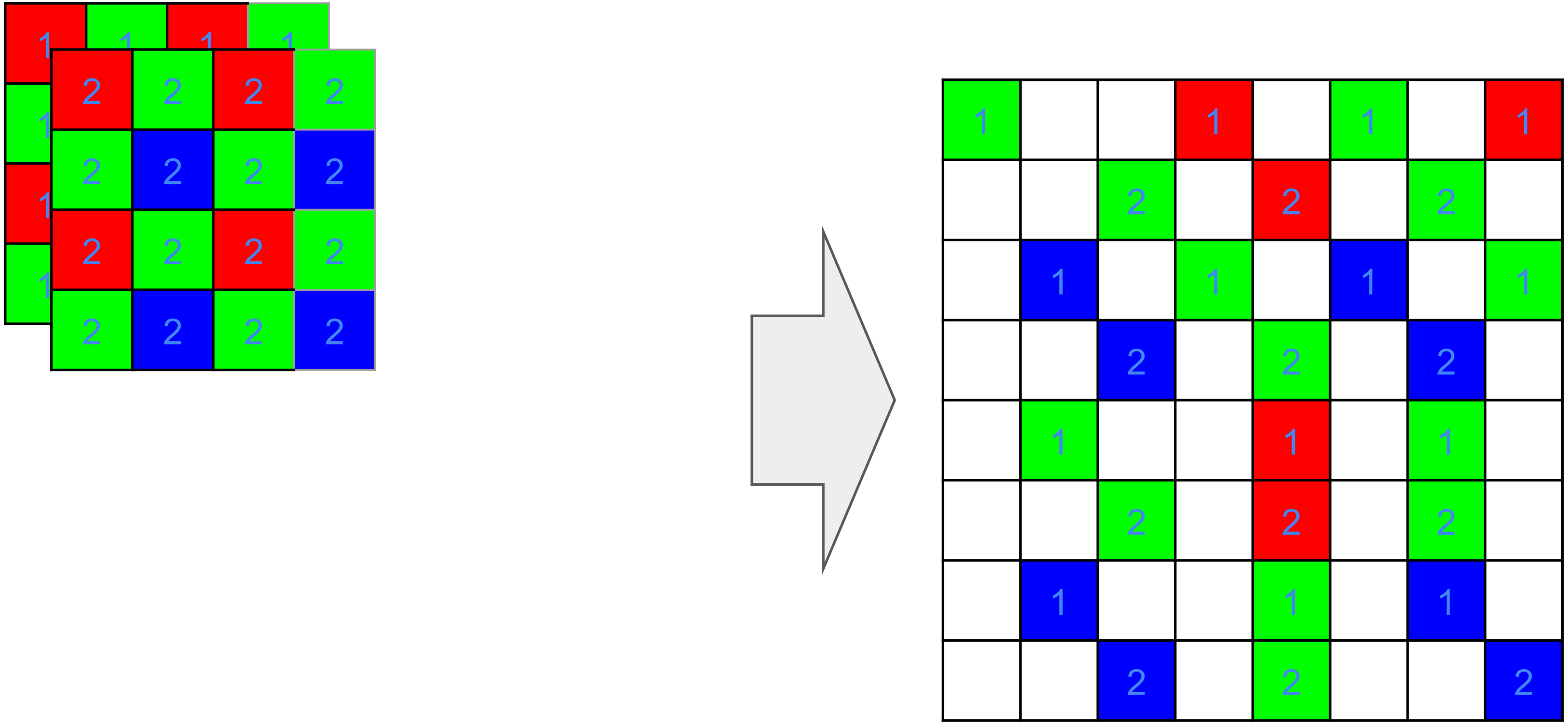
Instead ..... Replace demosaicing with multiple frames

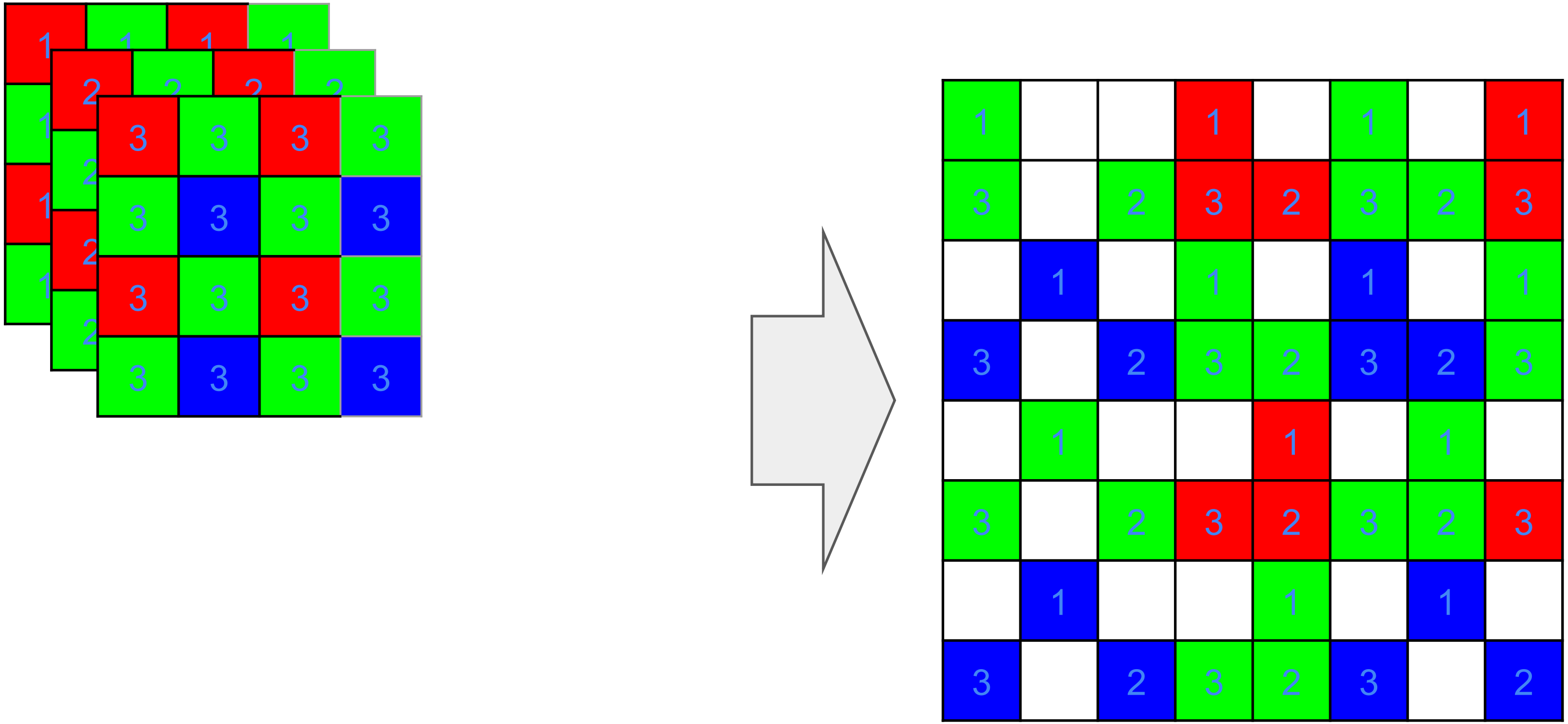
How:
“Pixel-shifting”



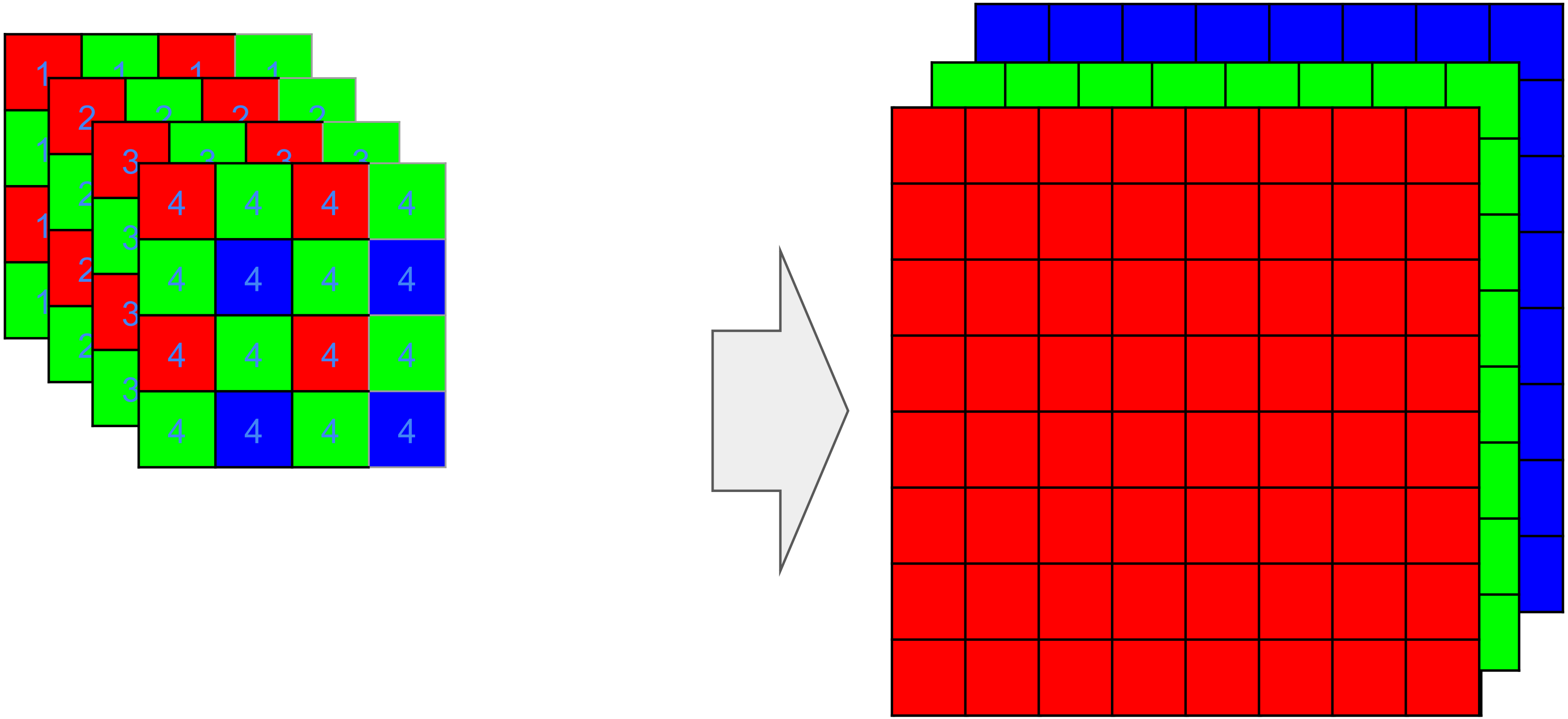
Life is not so simple.



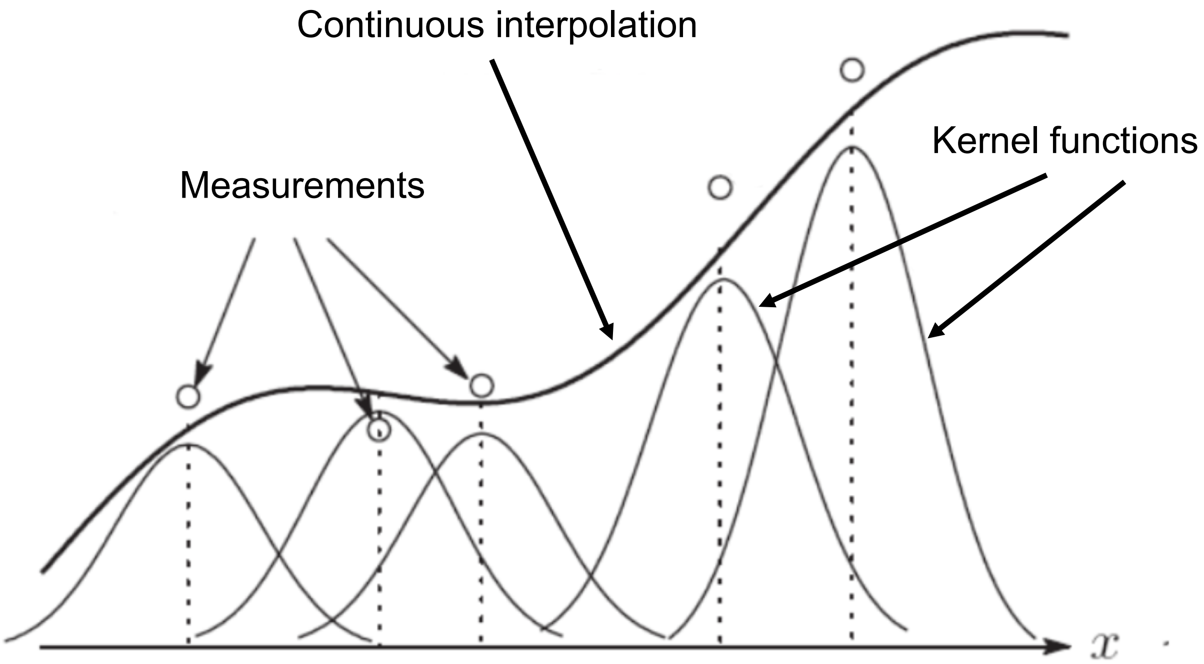

2 2 2 2
2 2 2 2
2 2 2 2
2 2 2 2
1 1 1 1
2 2 2
1 1 1 1
2 2 2
1 1 1
2 2 2
1 1 1
2 2 2

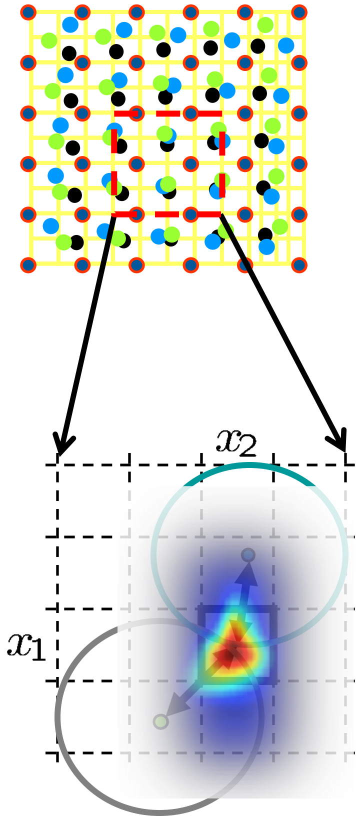


Multi-dimensional, non-uniform, interpolation

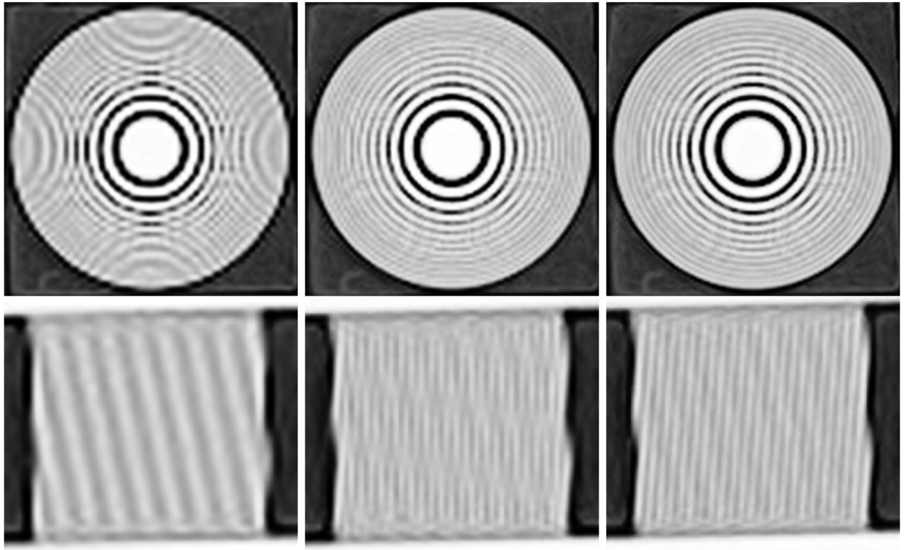


Merge: Nonlinear Kernel Regression





We can also merge onto higher-res grid



1.0x 1.5x 2.0x

- This has its limits
  - depending on the pixel size / lens spot size tradeoff
  - for typical mobile sensors, up to ~2x is possible

Source of **motion in mobile imaging**?

Handheld burst capture

(Natural) Physiological Tremor

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

PHYSIOLOGICAL TREMOR

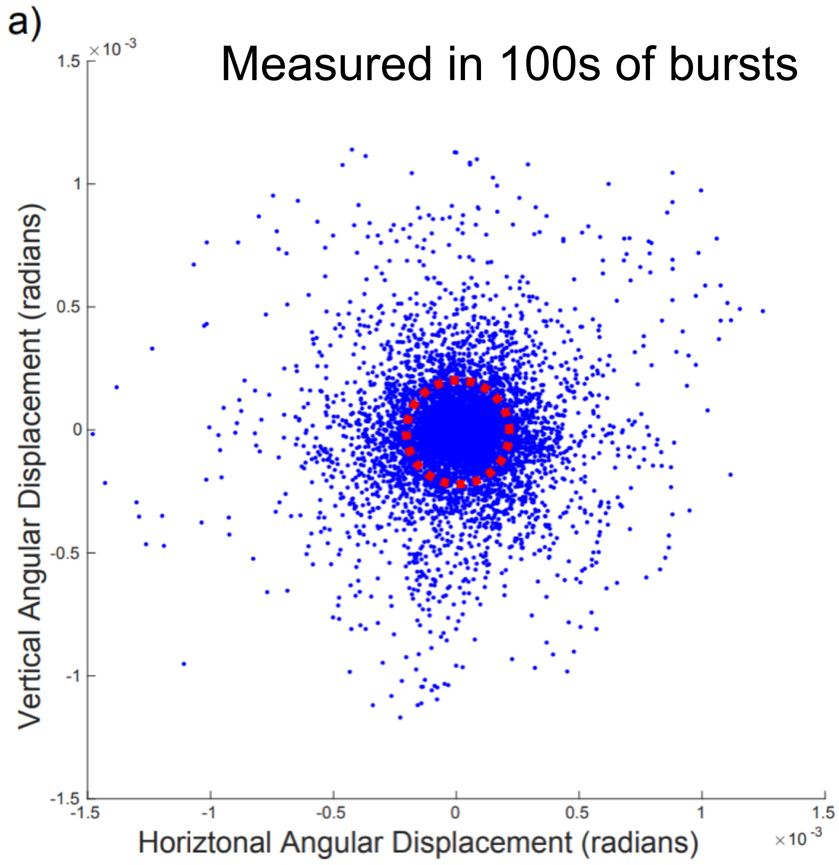
BY

JOHN MARSHALL AND E. GEOFFREY WALSH

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

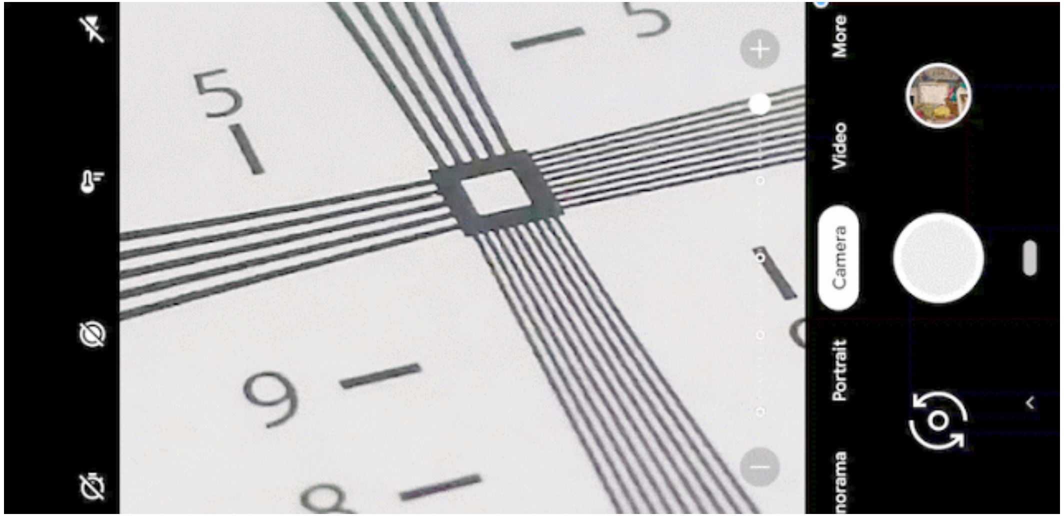
Rhythmicity during muscular contration has long been studied. The earliest observations dealt with the sounds that can be heard on listening to a contracting muscle and were naturally limited by the poor sensitivity of the ear at low frequencies. When, in the second half of the nineteenth century, graphic 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?

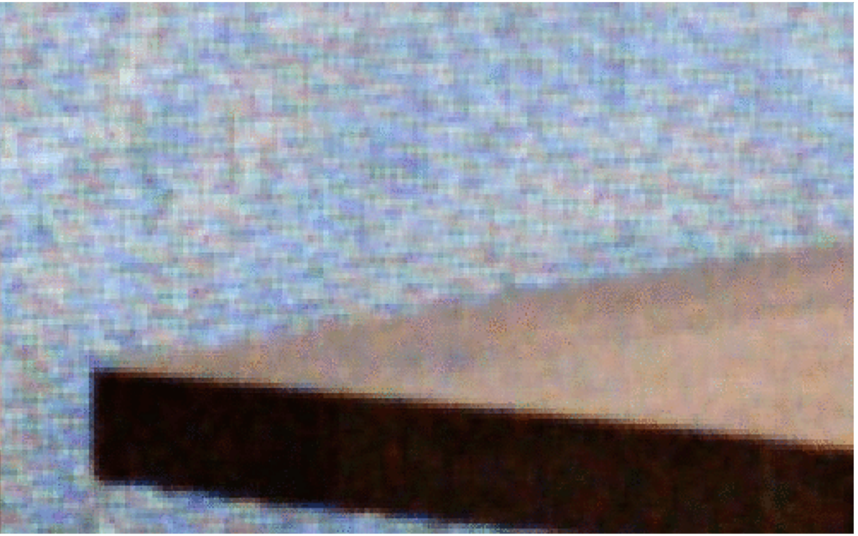
Simulated “tremor”




cintas

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

Crops

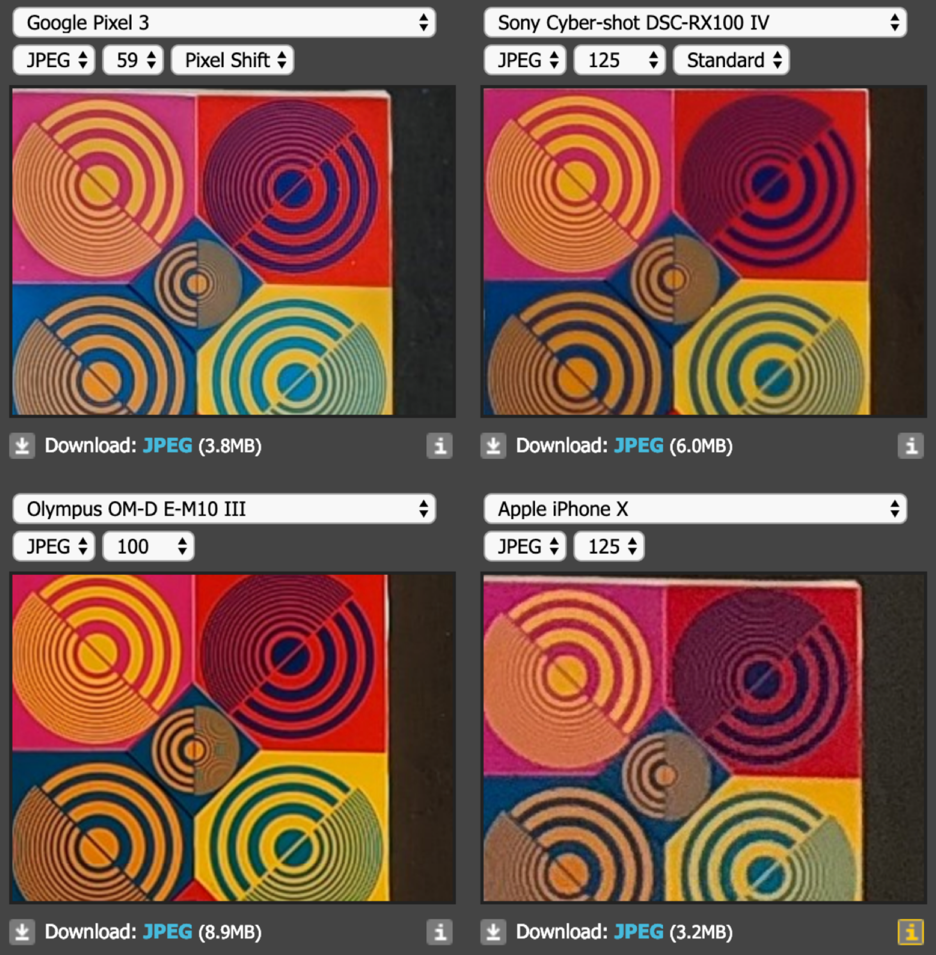
Hasinoff et al. [2016] | Ours

Full picture (reference)

Studio scene
Image comparison tool

DIGITAL PHOTOGRAPHY REVIEW
DPREVIEW

Lighting: Image size: FULL COMP PRINT



"The Pixel 3 is the first smartphone camera to truly challenge traditional cameras from an image quality standpoint, . . . rivaling cameras with Micro 4/3 sensors in [super-res] mode."

[SIGGRAPH 2019]

Handheld Multi-Frame Super-Resolution

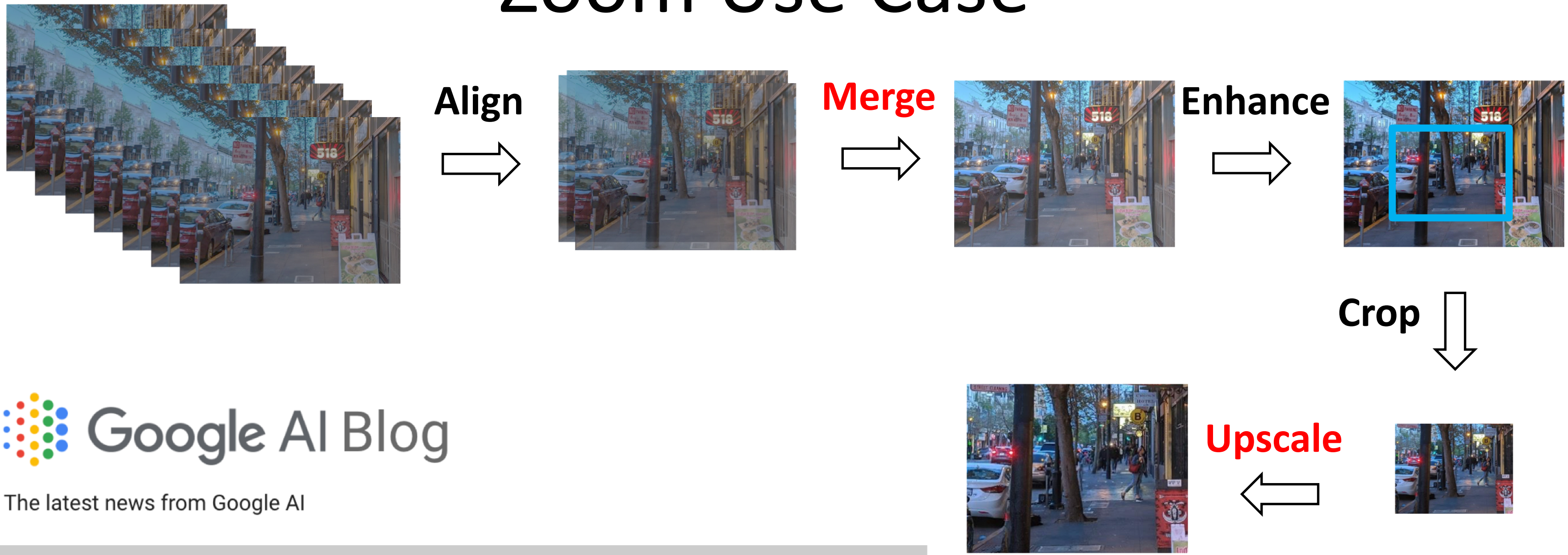
BARTLOMIEJ WRONSKI, IGNACIO GARCIA-DORADO, MANFRED ERNST, DAMIEN KELLY, MICHAEL KRAININ, CHIA-KAI LIANG, MARC LEVOY, and PEYMAN MILANFAR, Google Inc.


Night Sight


Phone X


Pixel 3 Night Mode

Zoom Use Case



Google AI Blog

The latest news from Google AI

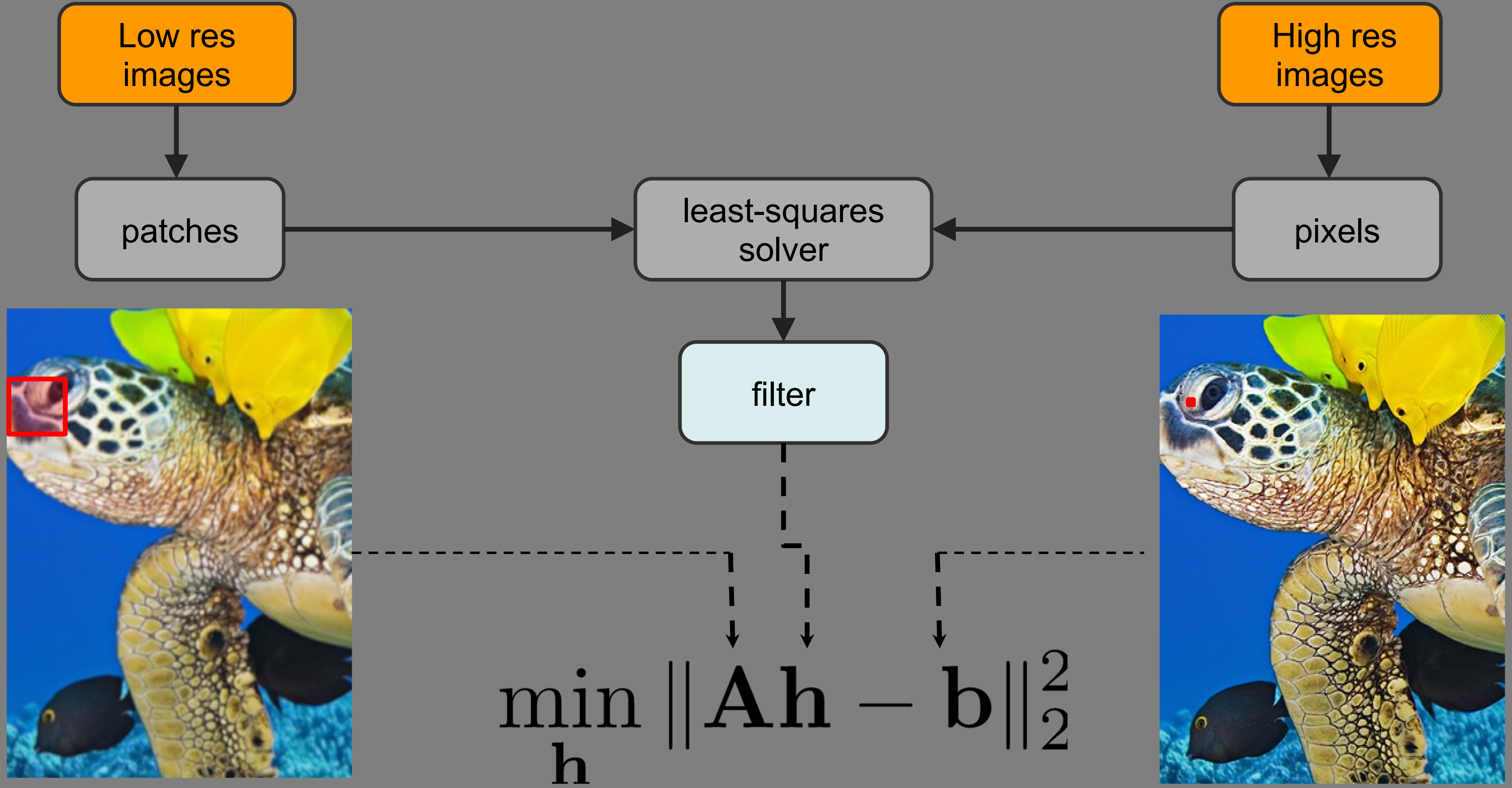
Enhance! RAISR Sharp Images with Machine Learning

Monday, November 14, 2016

Posted by Peyman Milanfar, Research Scientist

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

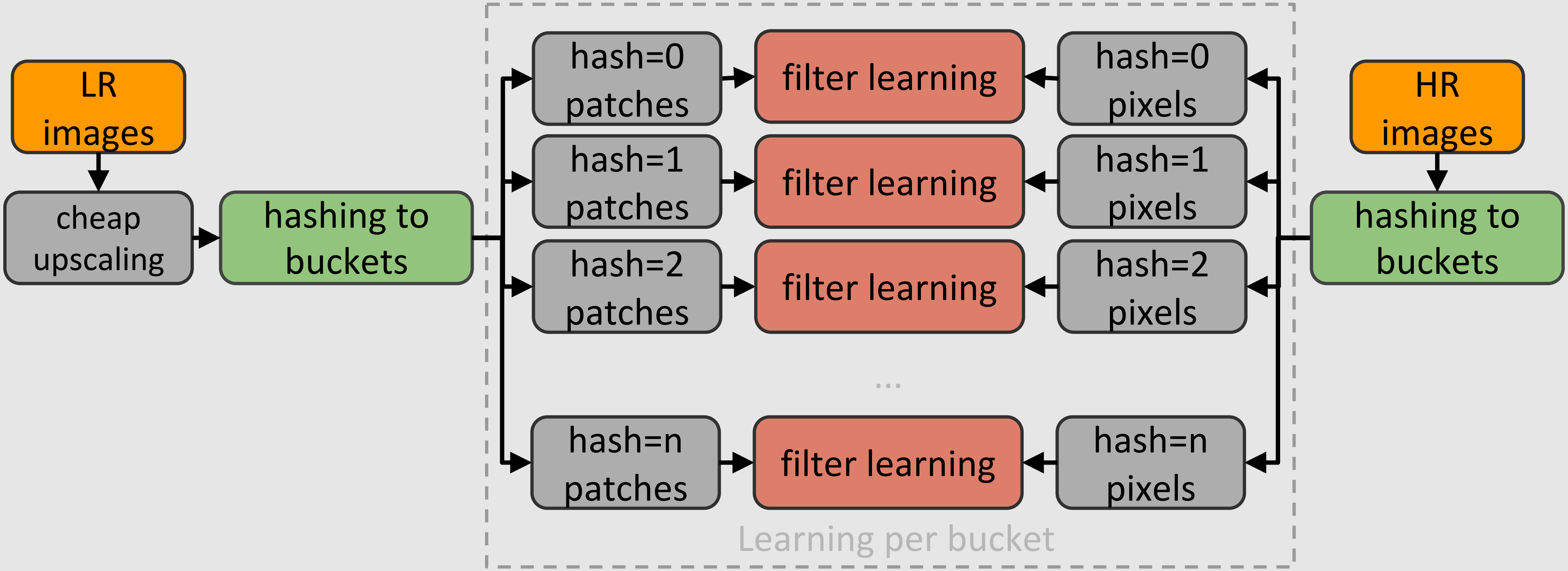
Filter Learning



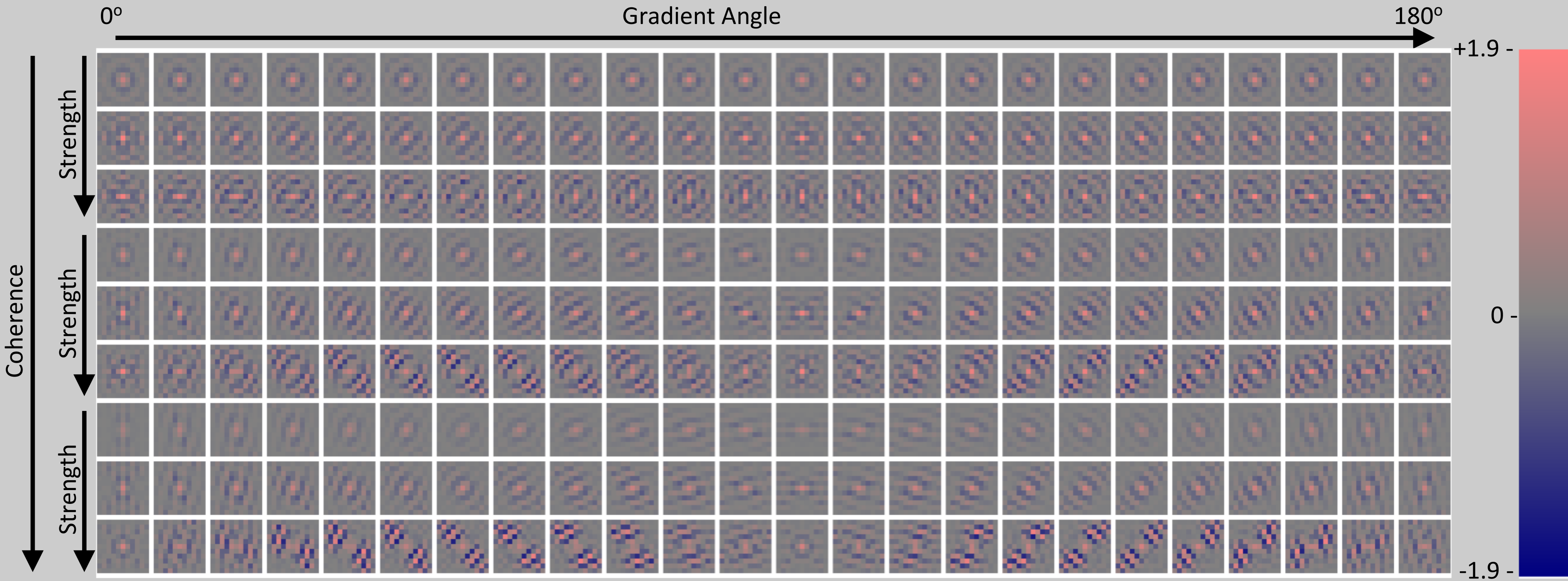
$$\min_{\mathbf{h}} \|\mathbf{A}\mathbf{h} - \mathbf{b}\|_2^2$$

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)

2017 Single-frame Super-res

(2x zoom crop)

2018 Multi-frame Super-res

85% of optical zoom resolution at 2x

“Best digital zoom on the market”

OTHER CHALLENGES IN COMPUTATIONAL IMAGING

Curation

Google AI Blog

The latest news from Google AI

[Talebi & Milanfar, IEEE Transactions on Image Processing 2018]

Introducing NIMA: Neural Image Assessment

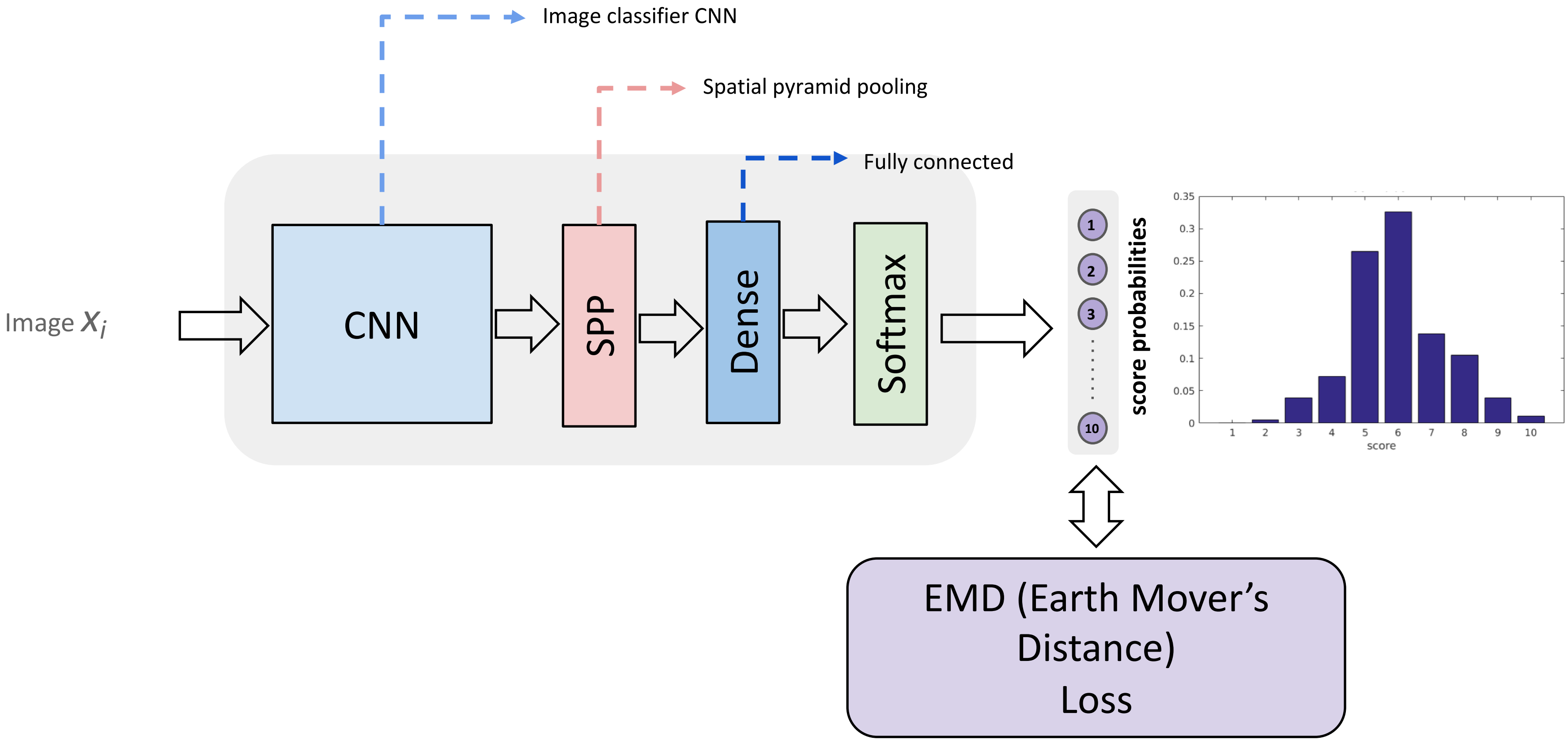
Monday, December 18, 2017

Posted by Hossein Talebi, Software Engineer and Peyman Milanfar Research Scientist, Machine Perception

Quantification of image quality and aesthetics has been a long-standing problem in image processing and computer vision. While technical quality assessment deals with measuring pixel-level degradations such as noise, blur, compression artifacts, etc., aesthetic assessment captures semantic level characteristics associated with emotions and beauty in images. Recently, deep convolutional neural networks (CNNs) trained with human-labelled data have been used to address the subjective nature of image quality for specific classes of images, such as landscapes. However, these approaches can be limited in their scope, as they typically categorize images to two classes of 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.

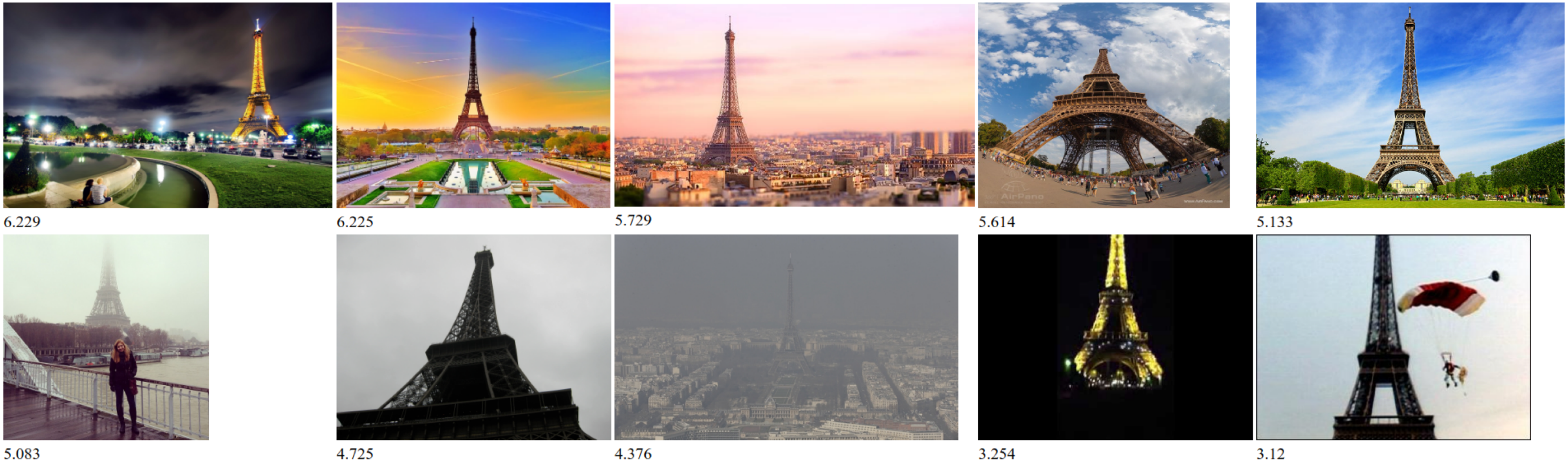
H. Talebi and P. Milanfar, "NIMA: Neural Image Assessment", In *IEEE Transactions on Image Processing 2018.*

NIMA: Neural Image Assessment

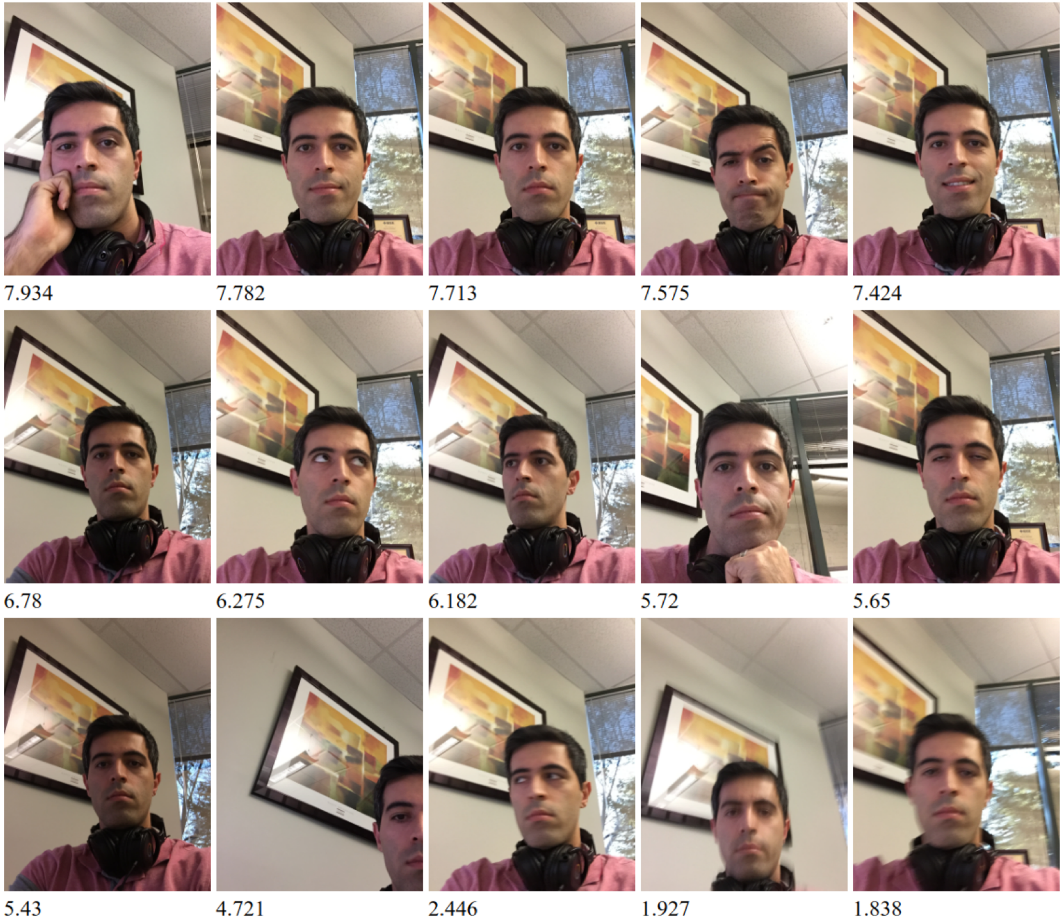


H. Talebi and P. Milanfar, "NIMA: Neural Image Assessment", In *IEEE Transactions on Image Processing 2018.*

NIMA for Aesthetic Quality



NIMA For Technical Quality




Shot on Pixel 3
With Super-res
Night Sight mode