Computational photography
&
cinematography

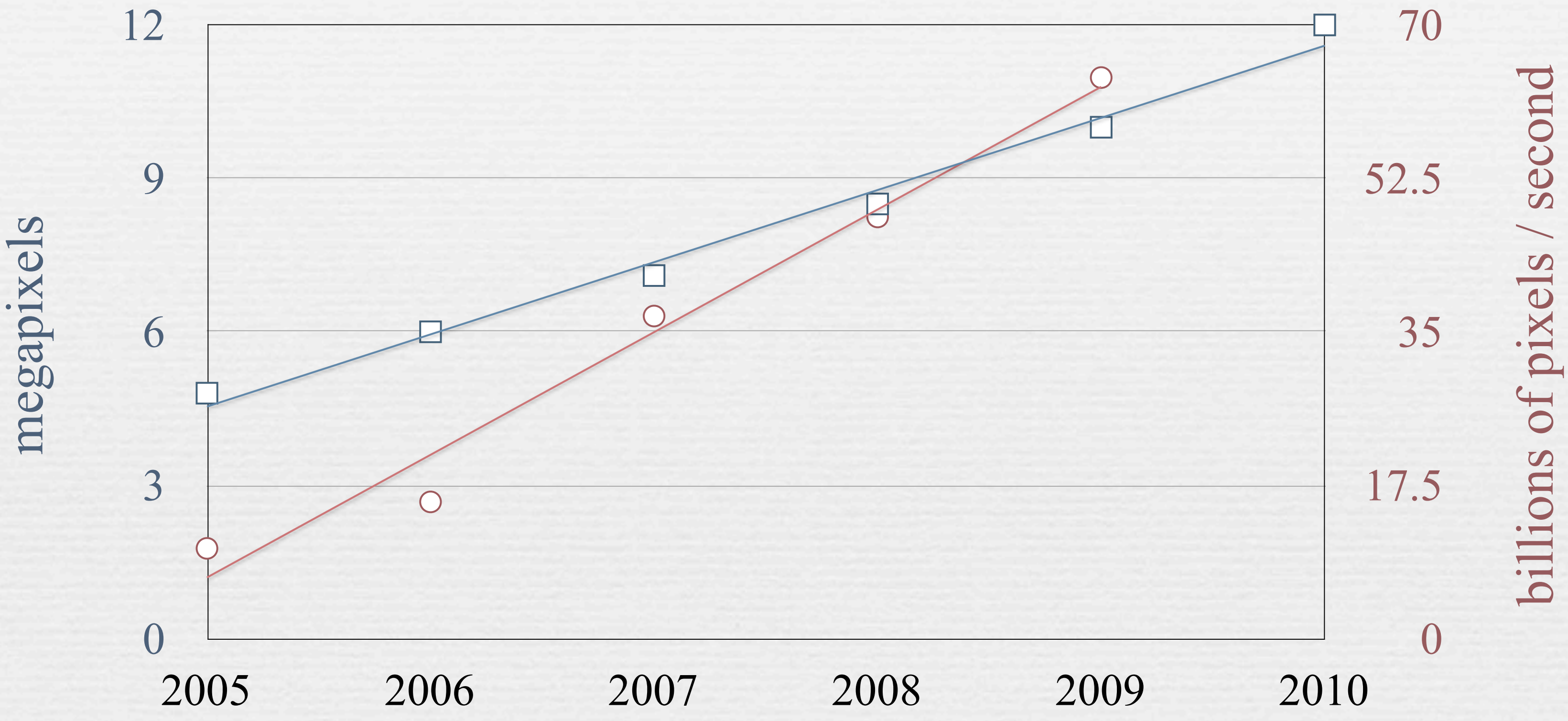
(for a lecture specifically devoted to the Stanford Frankencamera,
see http://graphics.stanford.edu/talks/camera20-public-may10-150dpi.pdf)

Marc Levoy
Computer Science Department
Stanford University

The future of digital photography

- the megapixel wars are over (and it's about time)
- computational photography is the next battleground in the camera industry (it's already starting)

Premise: available computing power in cameras is rising faster than megapixels



□ Avg megapixels in new cameras, CAGR = 1.2
○ NVIDIA GTX texture fill rate, CAGR = 1.8
(CAGR for Moore’s law = 1.5)

- this “headroom” permits more computation per pixel, or more frames per second, or less custom hardware

The future of digital photography

- the megapixel wars are over (long overdue)
- computational photography is the next battleground in the camera industry (it's already starting)

- how will these features appear to consumers?
  - standard and invisible
  - standard and visible (and disable-able)
  - aftermarket plugins and apps for your camera

The future of digital photography

- ✦ the megapixel wars are over (long overdue)
- ✦ computational photography is the next battleground in the camera industry (it's already starting)
- ✦ how will these features appear to consumers?
  - standard and invisible
  - standard and visible (and disable-able)
  - aftermarket plugins and apps for your camera
- ✦ traditional camera makers won't get it right
  - they'll bury it on page 120 of the manual (like Scene Modes)
  - the mobile industry will get it right (indie developers will help)

SynthCam

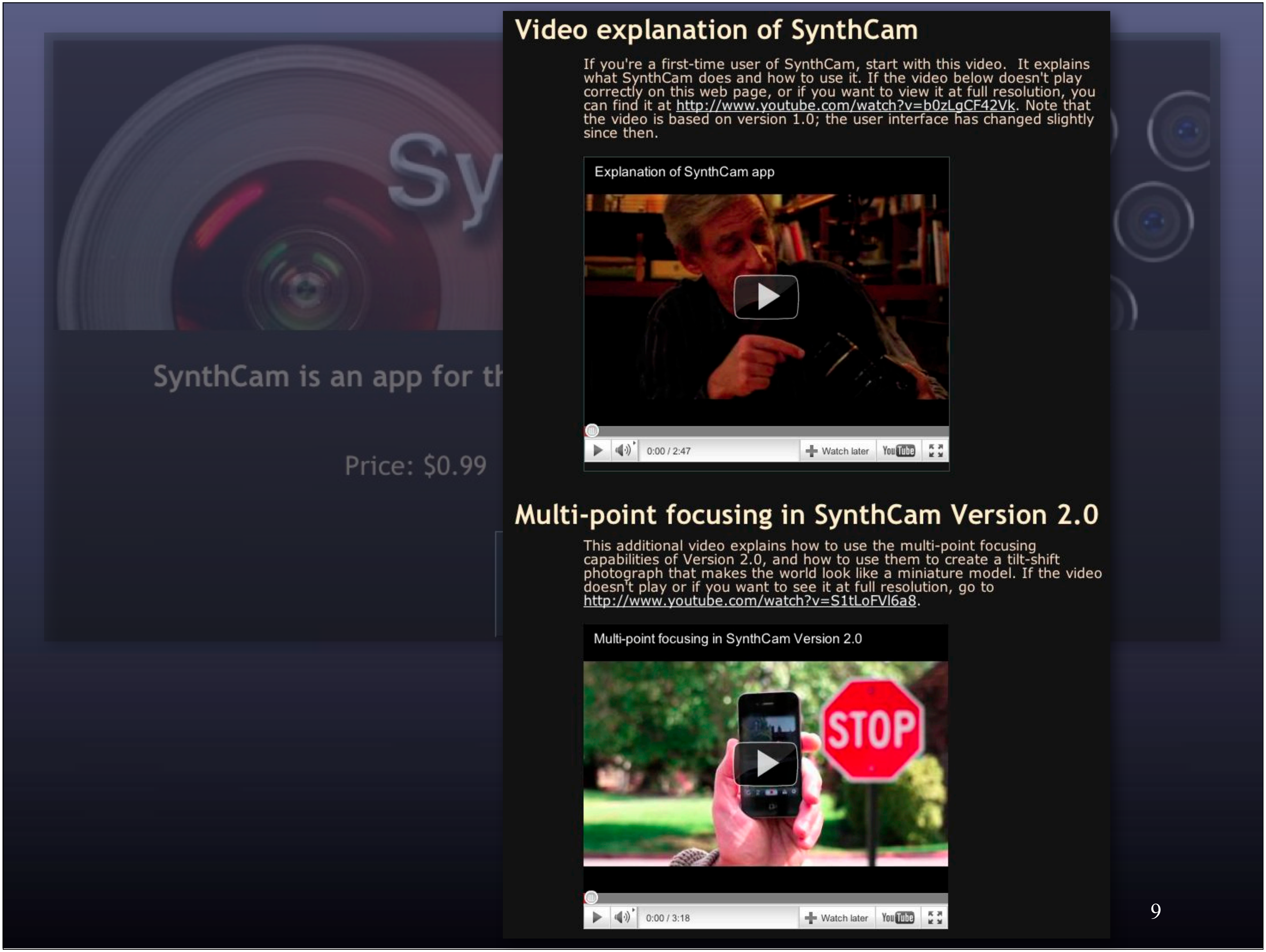
for the iPhone 4,

0.99

Available on the App St


SynthCa
2


SynthCam
d2



Video explanation of SynthCam

If you're a first-time user of SynthCam, start with this video. It explains what SynthCam does and how to use it. If the video below doesn't play correctly on this web page, or if you want to view it at full resolution, you can find it at http://www.youtube.com/watch?v=b0zLgCF42Vk. Note that the video is based on version 1.0; the user interface has changed slightly since then.

Multi-point focusing in SynthCam Version 2.0

This additional video explains how to use the multi-point focusing capabilities of Version 2.0, and how to use them to create a tilt-shift photograph that makes the world look like a miniature model. If the video doesn't play or if you want to see it at full resolution, go to http://www.youtube.com/watch?v=S1tLoFVl6a8.

Computational Photography

Film-like Photography with bits

Digital Photography

Image processing applied to captured images to produce better images.

Examples:
Interpolation, Filtering, Enhancement, Dynamic Range Compression, Color Management, Morphing, Hole Filling, Artistic Image Effects, Image Compression, Watermarking.

Computational Camera

Computational Processing

Processing of a set of captured images to create new images.

Examples:
Mosaicing, Matting, Super-Resolution, Multi-Exposure HDR, Light Field from Mutiple View, Structure from Motion, Shape from X.

Computational Imaging/Optics

Capture of optically coded images and computational decoding to produce new images.

Examples:
Coded Aperture, Optical Tomography, Diaphanography, SA Microscopy, Integral Imaging, Assorted Pixels, Catadioptric Imaging, Holographic Imaging.

Computational Sensor

Detectors that combine sensing and processing to create smart pixels.

Examples:
Artificial Retina, Retinex Sensors, Adaptive Dynamic Range Sensors, Edge Detect Chips, Focus of Expansion Chips, Motion Sensors.

Smart Light

Computational Illumination

Adapting and Controlling Illumination to Create revealing image

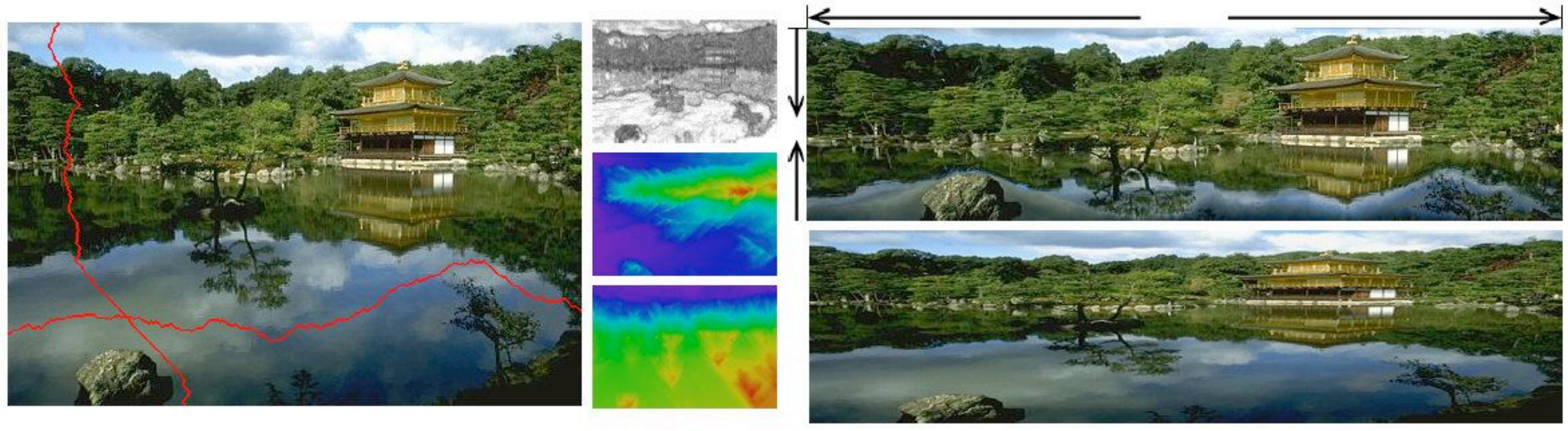
Examples:
Flash/no flash, Lighting domes, Multi-flash for depth edges, Dual Photos, Polynomial texture Maps, 4D light source

[Nayar, Tumblin]

Content-aware image resizing

[Avidan SIGGRAPH 2007]

- <u>to expand</u>: insert pixels along seams that, if removed in order, would yield the original image



Content-aware image resizing
[Avidan SIGGRAPH 2007]

- to compress: remove pixels along lowest-energy seams, ordered using dynamic programming
- to expand: insert pixe
  in order, v
- application to object removal

NOW AVAILABLE IN PHOTOSHOP !!

- extendable to video

Film-like Photography with bits	Computational Photography			
	Computational Camera			Smart Light
Digital Photography	Computational Processing	Computational Imaging/Optics	Computational Sensor	Computational Illumination
Image processing applied to captured images to produce better images.	Processing of a set of captured images to create new images.	Capture of optically coded images and computational decoding to produce new images.	Detectors that combine sensing and processing to create smart pixels.	Adapting and Controlling Illumination to Create revealing image
Examples: Interpolation, Filtering, Enhancement, Dynamic Range Compression, Color Management, Morphing, Hole Filling, Artistic Image Effects, Image Compression, Watermarking.	Examples: Mosaicing, Matting, Super-Resolution, Multi-Exposure HDR, Light Field from Mutiple View, Structure from Motion, Shape from X.	Examples: Coded Aperture, Optical Tomography, Diaphanography, SA Microscopy, Integral Imaging, Assorted Pixels, Catadioptric Imaging, Holographic Imaging.	Examples: Artificial Retina, Retinex Sensors, Adaptive Dynamic Range Sensors, Edge Detect Chips, Focus of Expansion Chips, Motion Sensors.	Examples: Flash/no flash, Lighting domes, Multi-flash for depth edges, Dual Photos, Polynomial texture Maps, 4D light source

[Nayar, Tumblin]

High dynamic range (HDR) imaging

Too dark

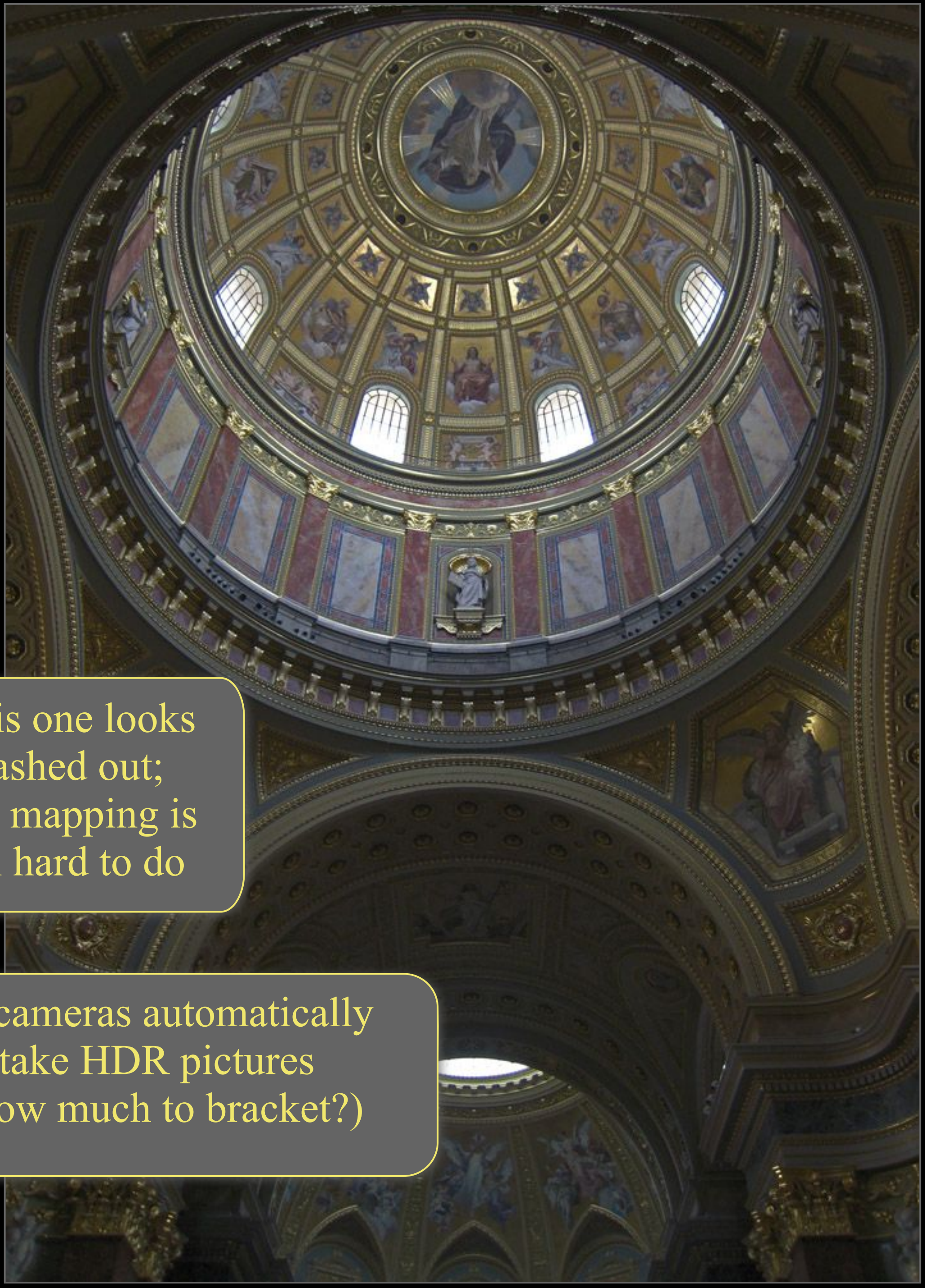
High dynamic range (HDR) imaging

Too bright

High dynamic range (HDR) imaging

this example worked well, but...

Tone mapped combination



...this one looks washed out; tone mapping is still hard to do

no cameras automatically take HDR pictures (How much to bracket?)

Aligning a burst of short-exposure, high-ISO shots using the Casio EX-F1

1/3 sec

Aligning a burst of short-exposure, high-ISO shots using the Casio EX-F1

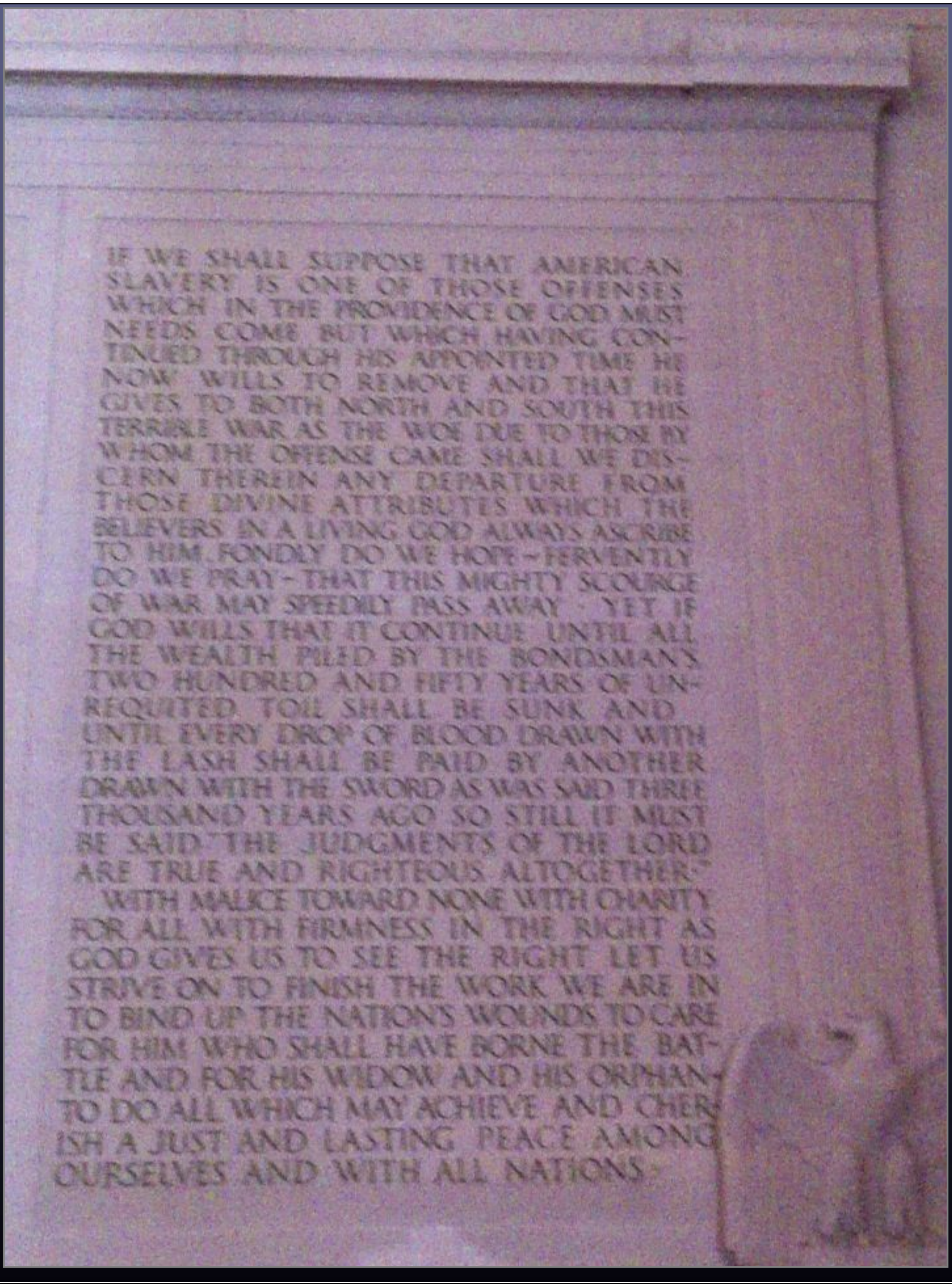
burst
at 60fps

Aligning a burst of short-exposure, high-ISO shots using the Casio EX-F1

1/3 sec

burst
at 60fps

iPhone 4,
single HD
video frame

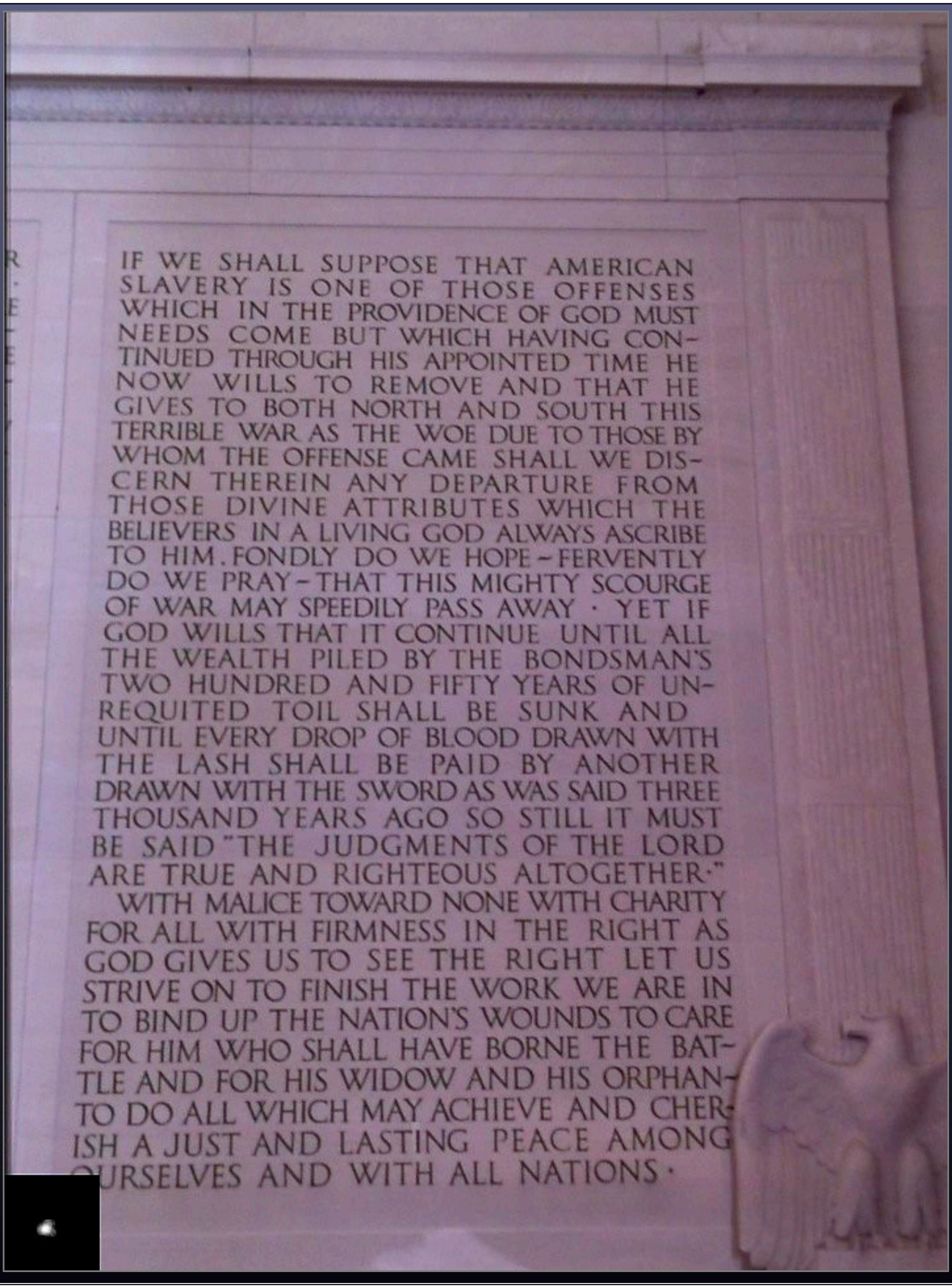


IF WE SHALL SUPPOSE THAT AMERICAN SLAVERY IS ONE OF THOSE OFFENSES WHICH IN THE PROVIDENCE OF GOD MUST NEEDS COME BUT WHICH HAVING CONTINUED THROUGH HIS APPOINTED TIME HE NOW WILLS TO REMOVE AND THAT HE GIVES TO BOTH NORTH AND SOUTH THIS TERRIBLE WAR AS THE WOE DUE TO THOSE BY WHOM THE OFFENSE CAME SHALL WE DISCERN THEREIN ANY DEPARTURE FROM THOSE DIVINE ATTRIBUTES WHICH THE BELIEVERS IN A LIVING GOD ALWAYS ASCRIBE TO HIM. FONDLY DO WE HOPE – FERVENTLY DO WE PRAY – THAT THIS MIGHTY SCOURGE OF WAR MAY SPEEDILY PASS AWAY. YET IF GOD WILLS THAT IT CONTINUE UNTIL ALL THE WEALTH PILED BY THE BONDSMAN'S TWO HUNDRED AND FIFTY YEARS OF UNREQUITED TOIL SHALL BE SUNK AND UNTIL EVERY DROP OF BLOOD DRAWN WITH THE LASH SHALL BE PAID BY ANOTHER DRAWN WITH THE SWORD AS WAS SAID THREE THOUSAND YEARS AGO SO STILL IT MUST BE SAID "THE JUDGMENTS OF THE LORD ARE TRUE AND RIGHTEOUS ALTOGETHER."

WITH MALICE TOWARD NONE WITH CHARITY FOR ALL WITH FIRMNESS IN THE RIGHT AS GOD GIVES US TO SEE THE RIGHT LET US STRIVE ON TO FINISH THE WORK WE ARE IN TO BIND UP THE NATION'S WOUNDS TO CARE FOR HIM WHO SHALL HAVE BORNE THE BATTLE AND FOR HIS WIDOW AND HIS ORPHAN – TO DO ALL WHICH MAY ACHIEVE AND CHERISH A JUST AND LASTING PEACE AMONG OURSELVES AND WITH ALL NATIONS.

SynthCam,
align & average
~30 frames

SNR increases as sqrt(# of frames)



IF WE SHALL SUPPOSE THAT AMERICAN SLAVERY IS ONE OF THOSE OFFENSES WHICH IN THE PROVIDENCE OF GOD MUST NEEDS COME BUT WHICH HAVING CONTINUED THROUGH HIS APPOINTED TIME HE NOW WILLS TO REMOVE AND THAT HE GIVES TO BOTH NORTH AND SOUTH THIS TERRIBLE WAR AS THE WOE DUE TO THOSE BY WHOM THE OFFENSE CAME SHALL WE DISCERN THEREIN ANY DEPARTURE FROM THOSE DIVINE ATTRIBUTES WHICH THE BELIEVERS IN A LIVING GOD ALWAYS ASCRIBE TO HIM. FONDLY DO WE HOPE – FERVENTLY DO WE PRAY – THAT THIS MIGHTY SCOURGE OF WAR MAY SPEEDILY PASS AWAY · YET IF GOD WILLS THAT IT CONTINUE UNTIL ALL THE WEALTH PILED BY THE BONDSMAN'S TWO HUNDRED AND FIFTY YEARS OF UNREQUITED TOIL SHALL BE SUNK AND UNTIL EVERY DROP OF BLOOD DRAWN WITH THE LASH SHALL BE PAID BY ANOTHER DRAWN WITH THE SWORD AS WAS SAID THREE THOUSAND YEARS AGO SO STILL IT MUST BE SAID "THE JUDGMENTS OF THE LORD ARE TRUE AND RIGHTEOUS ALTOGETHER·"

WITH MALICE TOWARD NONE WITH CHARITY FOR ALL WITH FIRMNESS IN THE RIGHT AS GOD GIVES US TO SEE THE RIGHT LET US STRIVE ON TO FINISH THE WORK WE ARE IN TO BIND UP THE NATION'S WOUNDS TO CARE FOR HIM WHO SHALL HAVE BORNE THE BATTLE AND FOR HIS WIDOW AND HIS ORPHAN – TO DO ALL WHICH MAY ACHIEVE AND CHERISH A JUST AND LASTING PEACE AMONG OURSELVES AND WITH ALL NATIONS ·

Removing foreground objects by translating the camera

- align the shots
- match histograms
- apply median filter

Computational Photography

Film-like Photography with bits	Computational Camera			Smart Light
Digital Photography	Computational Processing	Computational Imaging/Optics	Computational Sensor	Computational Illumination
Image processing applied to captured images to produce better images.	Processing of a set of captured images to create new images.	Capture of optically coded images and computational decoding to produce new images.	Detectors that combine sensing and processing to create smart pixels.	Adapting and Controlling Illumination to Create revealing image
Examples: Interpolation, Filtering, Enhancement, Dynamic Range Compression, Color Management, Morphing, Hole Filling, Artistic Image Effects, Image Compression, Watermarking.	Examples: Mosaicing, Matting, Super-Resolution, Multi-Exposure HDR, Light Field from Mutiple View, Structure from Motion, Shape from X.	Examples: Coded Aperture, Optical Tomography, Diaphanography, SA Microscopy, Integral Imaging, Assorted Pixels, Catadioptric Imaging, Holographic Imaging.	Examples: Artificial Retina, Retinex Sensors, Adaptive Dynamic Range Sensors, Edge Detect Chips, Focus of Expansion Chips, Motion Sensors.	Examples: Flash/no flash, Lighting domes, Multi-flash for depth edges, Dual Photos, Polynomial texture Maps, 4D light source

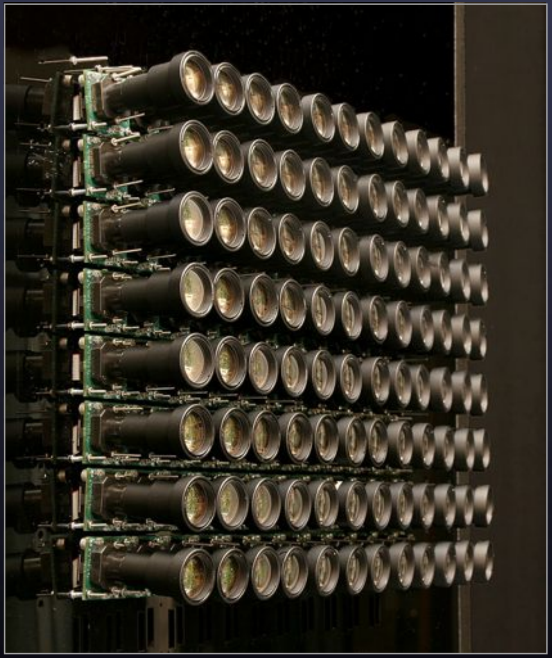
[Nayar, Tumblin]

Stanford Multi-Camera Array

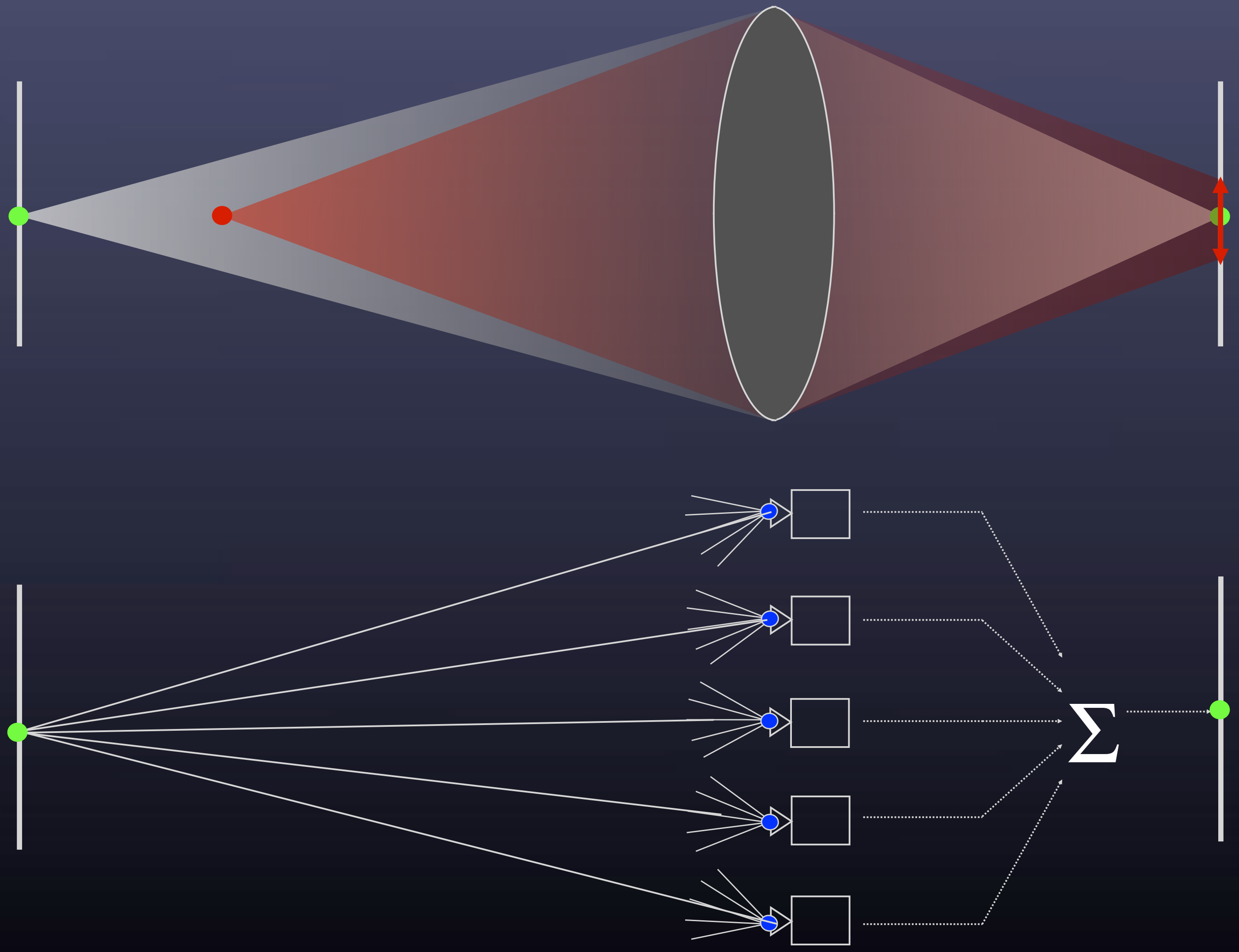
[Wilburn SIGGRAPH 2005]

- 640 × 480 pixels × 30 fps × 128 cameras

- synchronized timing
- continuous streaming
- flexible arrangement

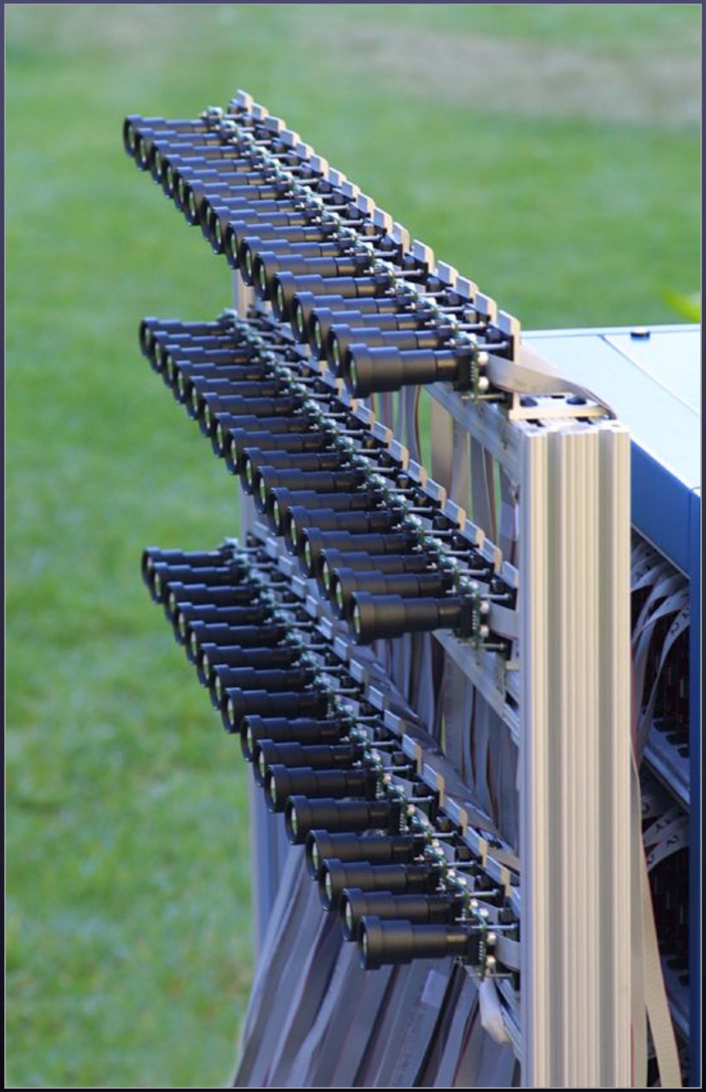


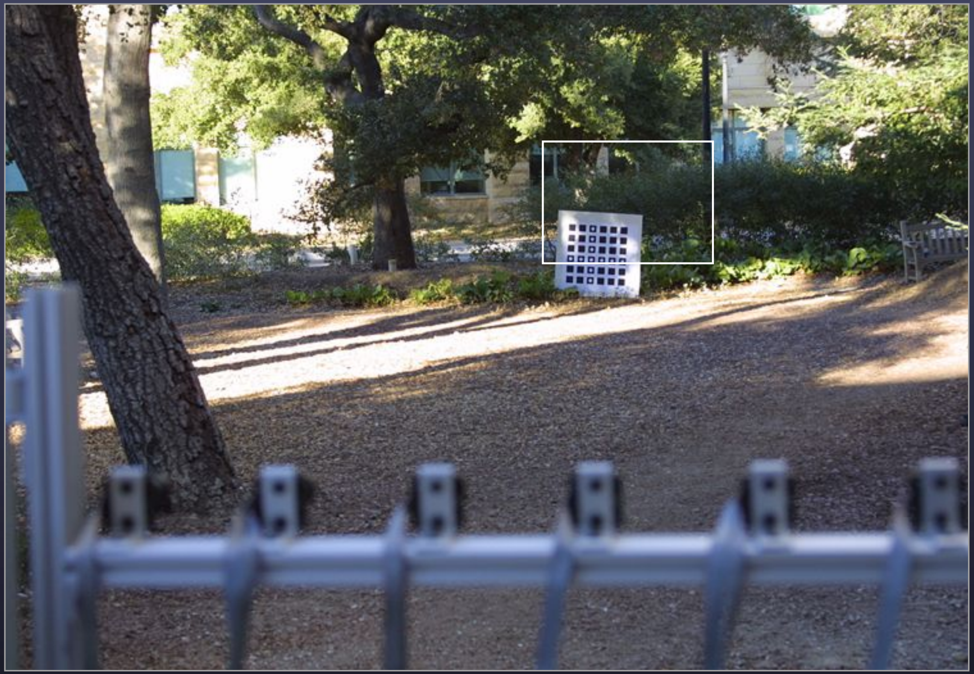
Synthetic aperture photography

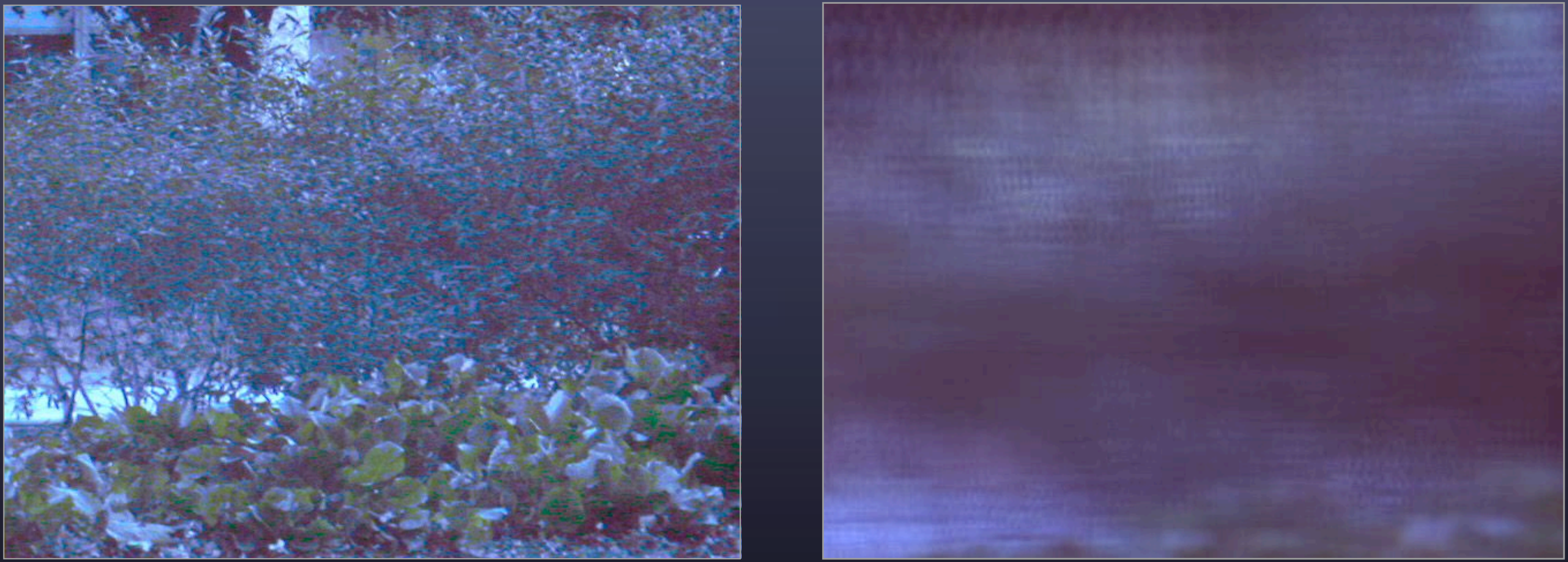


Example using 45 cameras

[Vaish CVPR 2004]





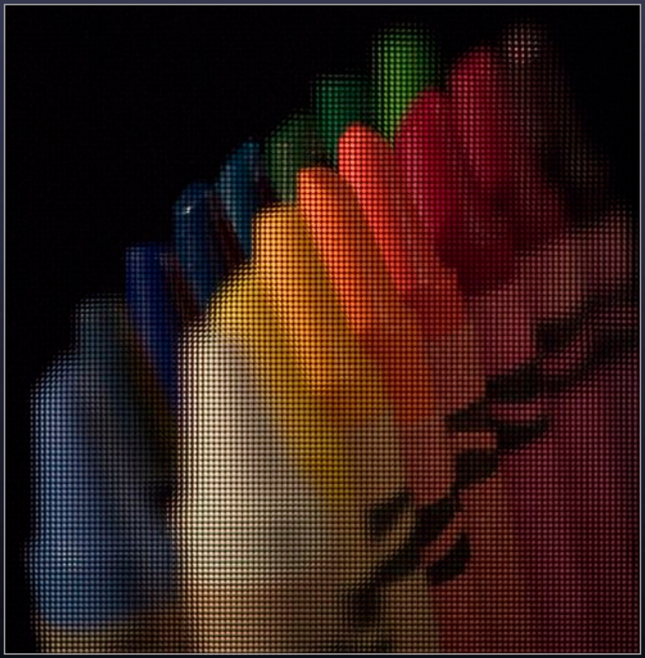


(movie is available at http://graphics.stanford.edu/projects/array)

Light field photography using a handheld plenoptic camera

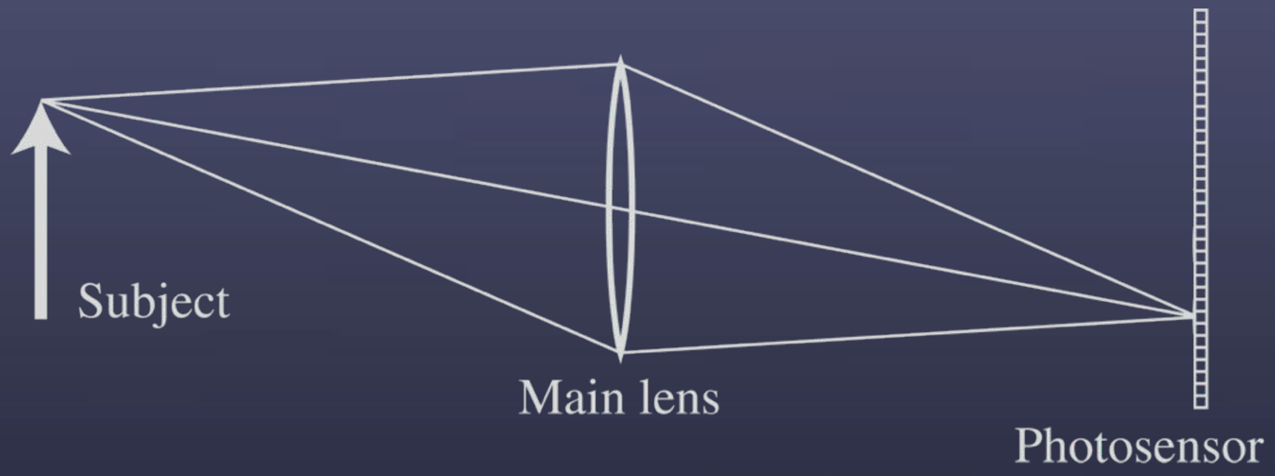
Ren Ng, Marc Levoy, Mathieu Brédif,
Gene Duval, Mark Horowitz and Pat Hanrahan

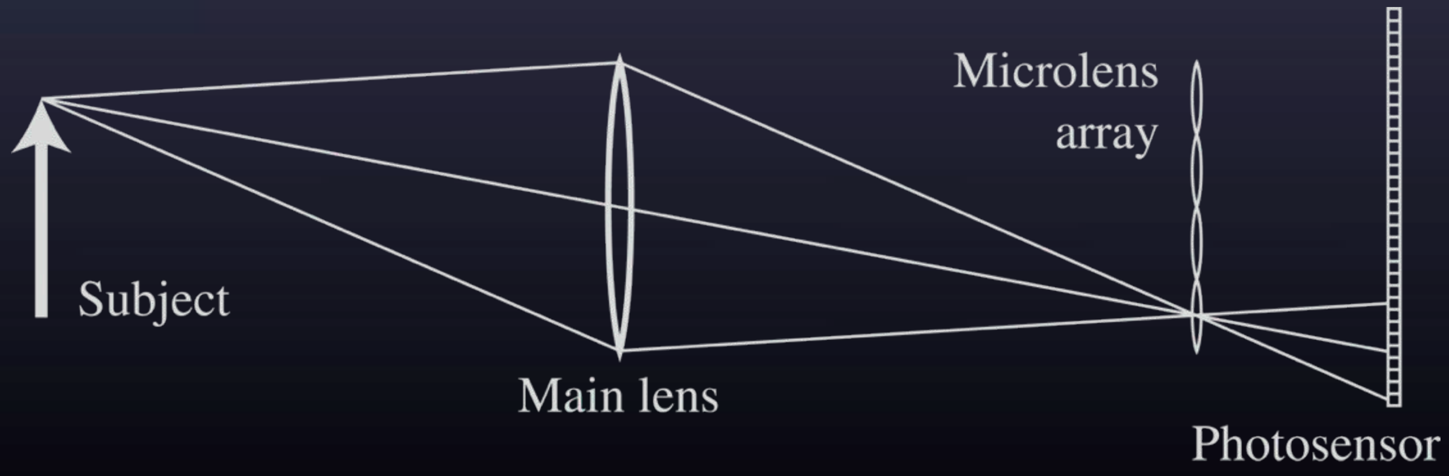
(Proc. SIGGRAPH 2005
and TR 2005-02)



Light field photography

[Ng SIGGRAPH 2005]

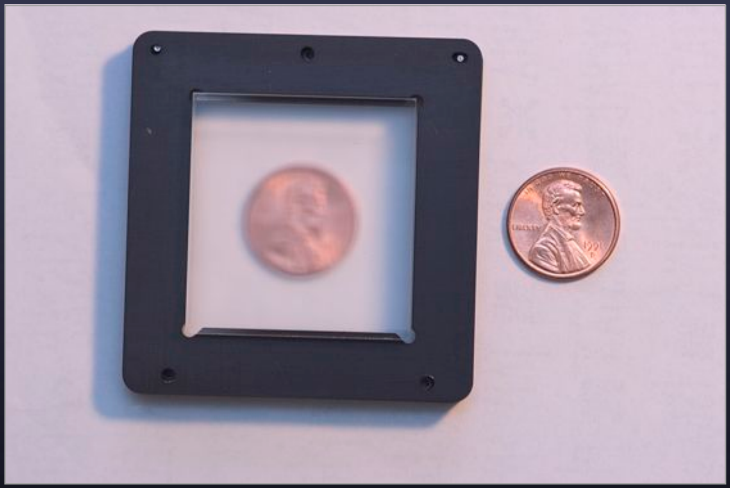




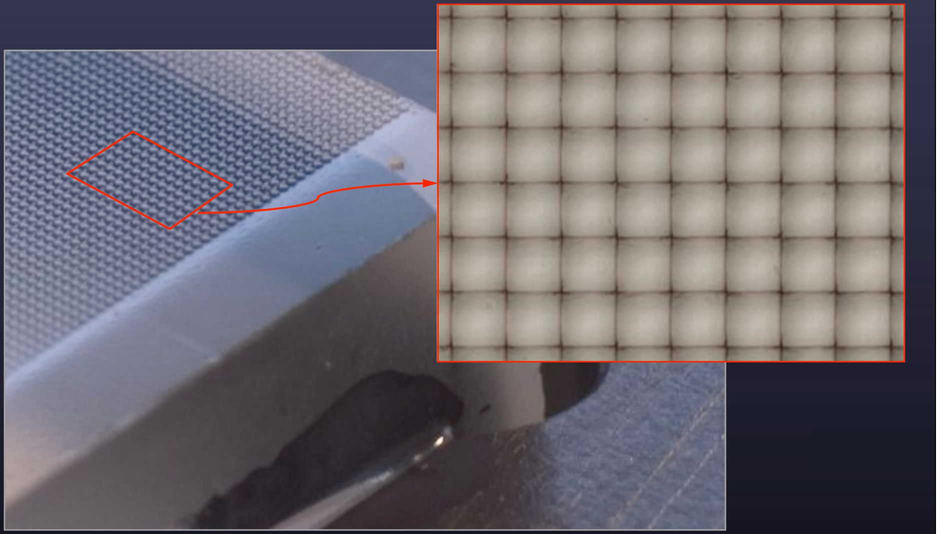
Prototype camera

Contax medium format camera

Kodak 16-megapixel sensor



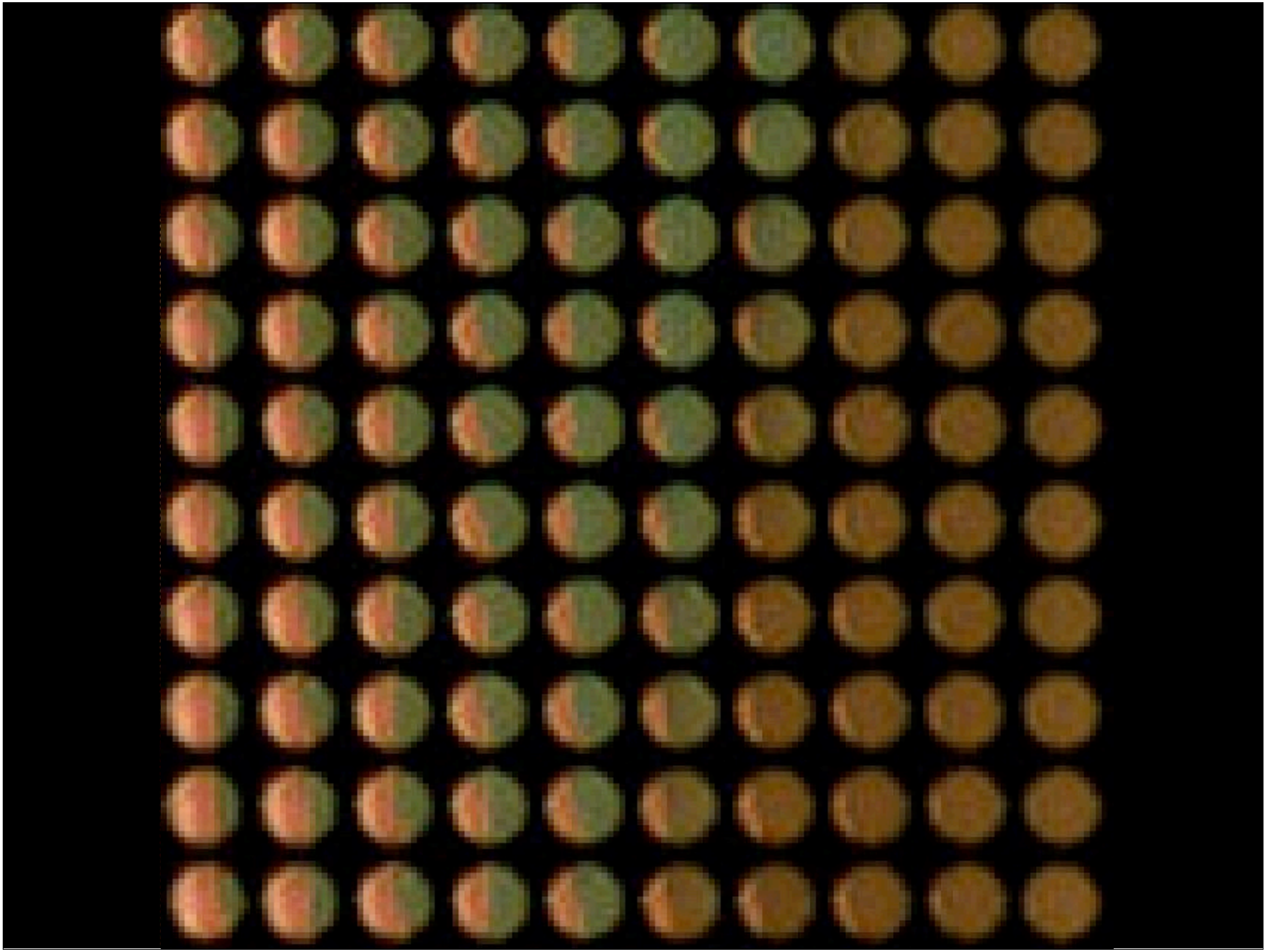
Adaptive Optics microlens array



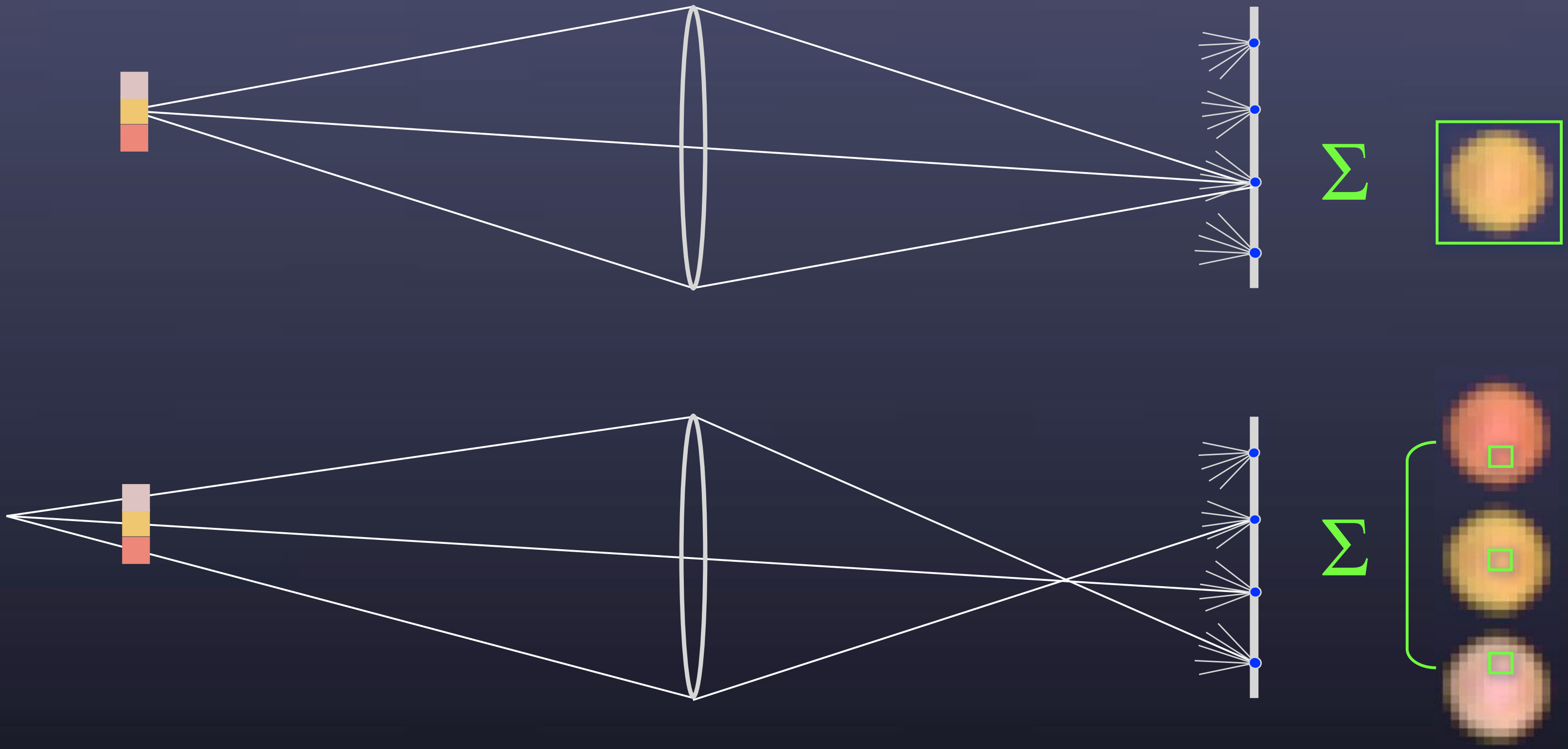
125µ square-sided microlenses

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

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



Digital refocusing



Example of digital refocusing

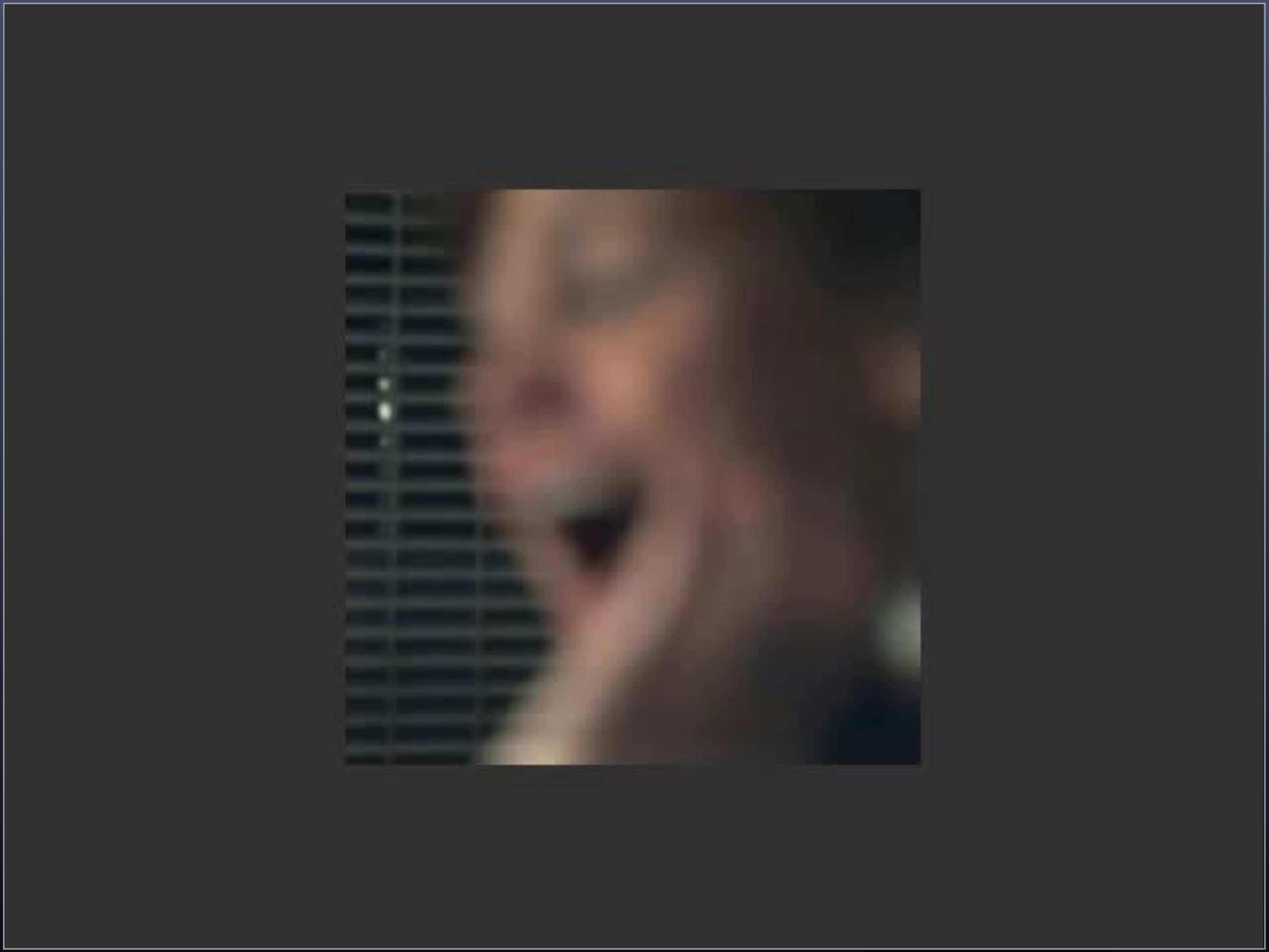
Example of digital refocusing

Example of digital refocusing

Example of digital refocusing

Example of digital refocusing

Refocusing portraits



(movie is available at http://refocusimaging.com)

Application to sports photography

Application to sports photography

Application to sports photography

Film-like Photography with bits	Computational Photography			
	Computational Camera			Smart Light
Digital Photography	Computational Processing	Computational Imaging/Optics	Computational Sensor	Computational Illumination
Image processing applied to captured images to produce better images.	Processing of a set of captured images to create new images.	Capture of optically coded images and computational decoding to produce new images.	Detectors that combine sensing and processing to create smart pixels.	Adapting and Controlling Illumination to Create revealing image
Examples: Interpolation, Filtering, Enhancement, Dynamic Range Compression, Color Management, Morphing, Hole Filling, Artistic Image Effects, Image Compression, Watermarking.	Examples: Mosaicing, Matting, Super-Resolution, Multi-Exposure HDR, Light Field from Mutiple View, Structure from Motion, Shape from X.	Examples: Coded Aperture, Optical Tomography, Diaphanography, SA Microscopy, Integral Imaging, Assorted Pixels, Catadioptric Imaging, Holographic Imaging.	Examples: Artificial Retina, Retinex Sensors, Adaptive Dynamic Range Sensors, Edge Detect Chips, Focus of Expansion Chips, Motion Sensors.	Examples: Flash/no flash, Lighting domes, Multi-flash for depth edges, Dual Photos, Polynomial texture Maps, 4D light source

[Nayar Tumblin]

Flash-noflash photography

[Agrawal SIGGRAPH 2005]

- compute ambient + flash – features in sum that don't appear in ambient alone (as determined from image gradients) (except where ambient image is nearly black)

What's wrong with this picture?

- many of these techniques require modifying the camera
  - digital refocusing
- some of these techniques could use help from the camera
  - metering for HDR
- none of these ideas are finding their way into consumer cameras...

Why have traditional camera makers been so slow to embrace computational photography?

(soapbox mode ON)

- the camera industry is secretive
  - no flow of workers between companies and universities
  - few publications, no open source software community

- camera companies sell hardware, not software
  - many are not comfortable with Internet ecosystems

- some computational techniques are still not robust
  - partly because researchers can't test them in the field

(soapbox mode OFF)

Camera 2.0

Marc Levoy
Computer Science Department
Stanford University

NOKIA
Connecting People

Adobe

Kodak

hp
invent

Camera 2.0

Marc Levoy
Computer Science Department
Stanford University

SONY
NVIDIA
intel
TEXAS INSTRUMENTS
The Walt Disney Company
Google
NOKIA
Connecting People
Adobe
Kodak
hp
invent

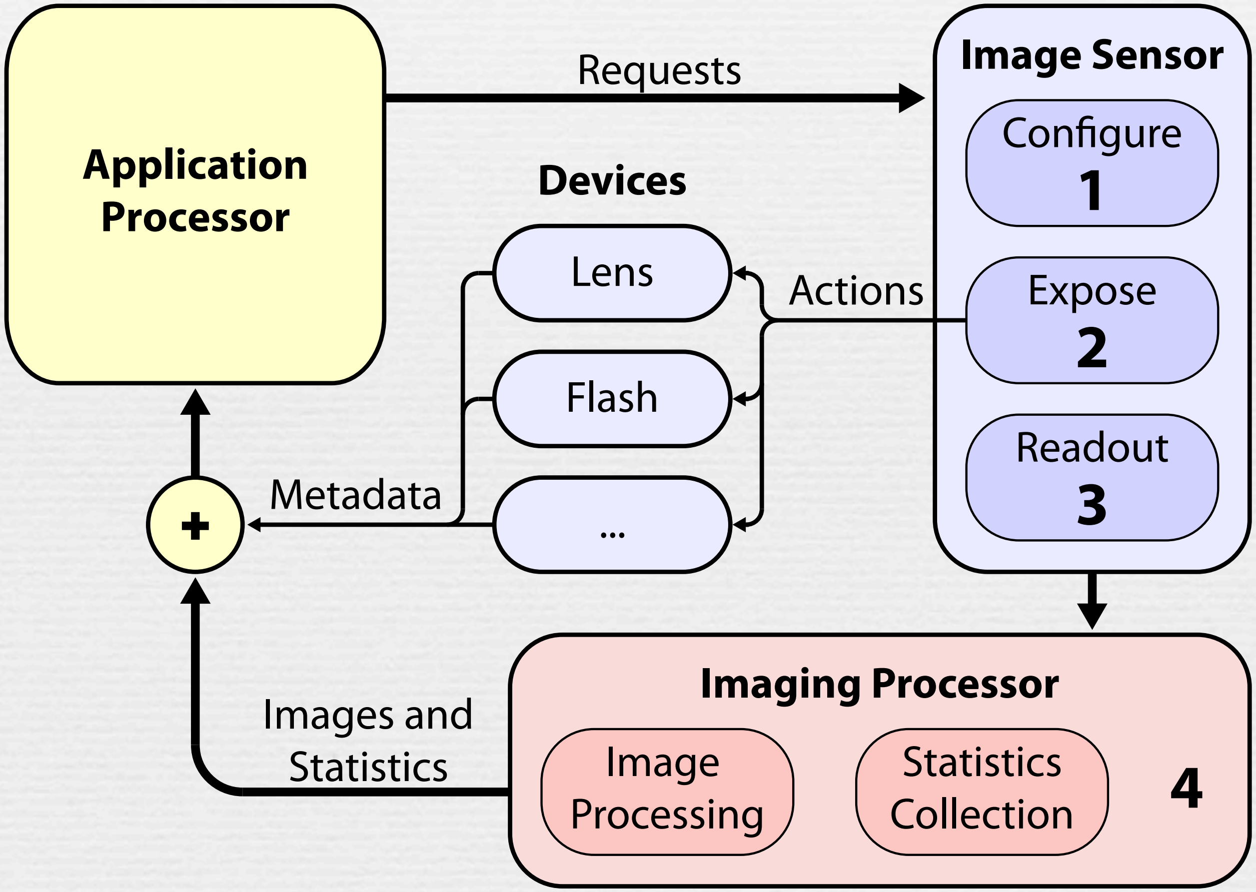
The Stanford Frankencameras

Frankencamera F2

Nokia N900 “F”

- facilitate research in experimental computational photography
- for students in computational photography courses worldwide
- proving ground for plugins and apps for future cameras

Frankencamera architecture



Frankencamera software: the `FCAM` API

```
Sensor sensor;
Flash flash;
vector<Shot> burst(2);

burst[0].exposure = 1/200.;
burst[1].exposure = 1/30.;

Flash::FireAction fire(&flash);
fire.time = burst[0].exposure/2;
burst[0].actions.insert(fire);

sensor.stream(burst);

while (1) {
  Frame flashFrame =
    sensor.getFrame();
  Frame noflashFrame =
    sensor.getFrame();
}
```

Demonstration applications

- Canon 430EX (smaller flash) strobed continuously
- Canon 580EX (larger flash) fired once at end of exposure

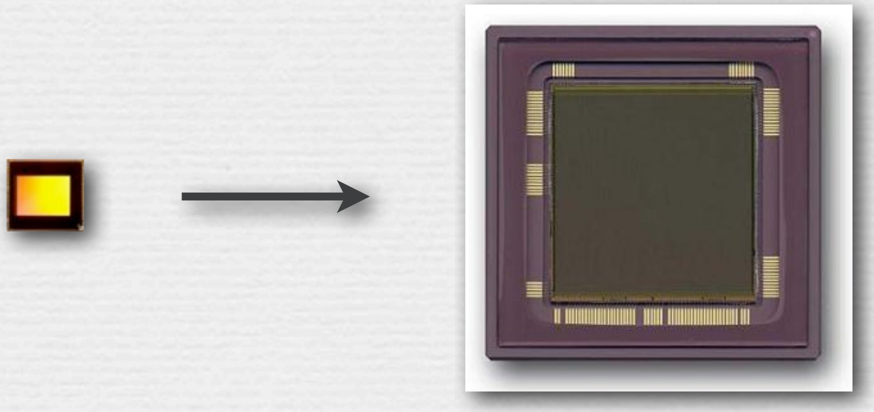
Short-term roadmap

- distribution to hobbyists, 3rd party developers
  - probably only N900s or equiv.
  - plugins and apps

Short-term roadmap

- distribution to hobbyists, 3rd party developers
  - probably only N900s or equiv.
  - plugins and apps

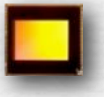
- distribution to researchers and students
  - Frankencamera F3 + N900s + courseware
  - bootstrap open-source community

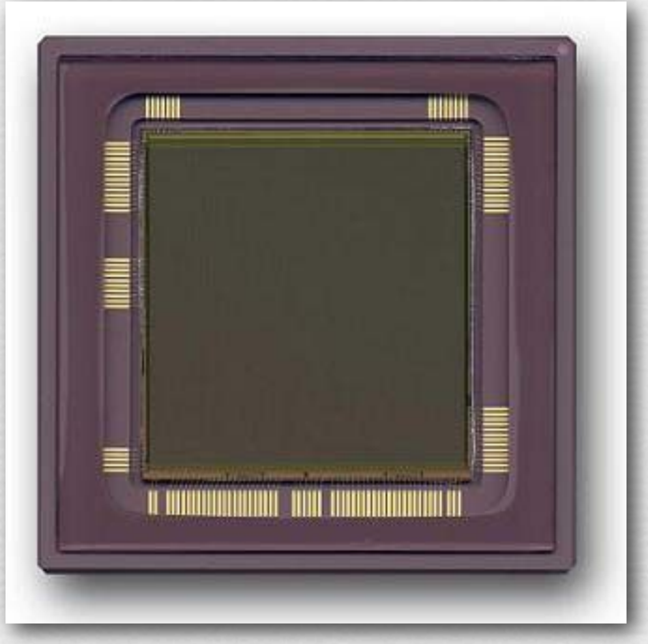


Unretouched pictures from Nokia N95

(5 megapixels, Zeiss lens, auto-focus)

Sensors for our Frankencameras



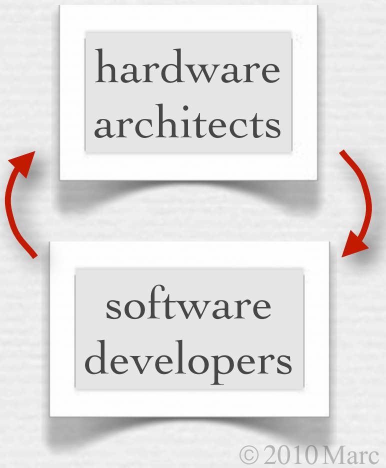


Micron MT9001
- 5 megapixel
- cell phone quality
- $150

Cypress LUPA 4000
- $1500
- DSLR quality
- arbitrary ROIs and non-destructive readout

Short-term roadmap

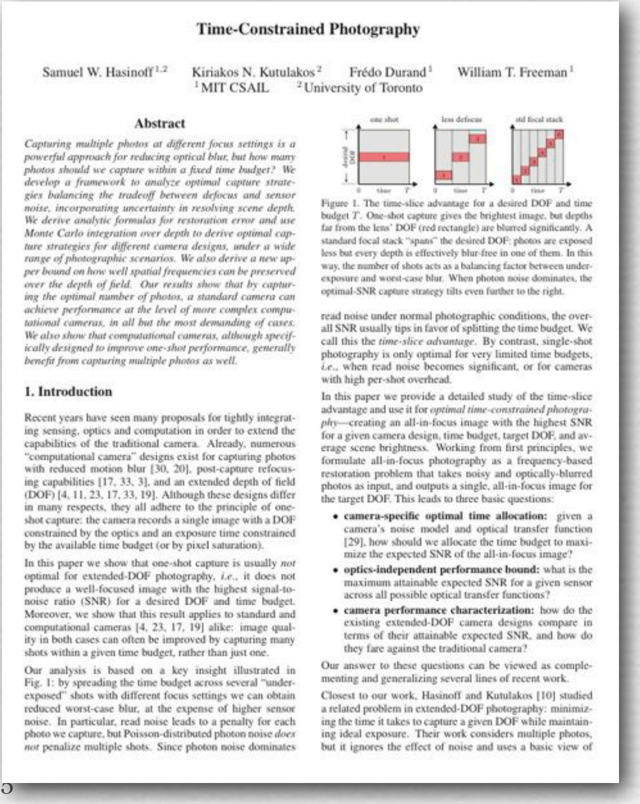
- distribution to hobbyists, 3rd party developers
  - probably only N900s or equiv.
  - plugins and apps
- distribution to researchers and students
  - courseware + Frankencameras/N900s
  - bootstrap open-source community
- wish list for makers of camera hardware
  - per-frame resolution switching at video rate
  - fast path into GPU texture memory
  - hardware feature detector


hardware architects

software developers

Long-term prospects

- high-speed burst-mode photography
    - all still cameras should capture at (up to) 500 fps
    - capture while aiming the camera - no shutter half-press
    - frameless photography - ROIs, MOIs (“M” = Moment)



Time-Constrained Photography

Samuel W. Hasinoff [1,2]   Kiriakos N. Kutulakos [2]   Frédo Durand [1]   William T. Freeman [1]
[1] MIT CSAIL   [2] University of Toronto

Abstract

Capturing multiple photos at different focus settings is a powerful approach for reducing optical blur, but how many photos should we capture within a fixed time budget? We develop a framework to analyze optimal capture strategies balancing the tradeoff between defocus and sensor noise, incorporating uncertainty in resolving scene depth. We derive analytic formulas for restoration error and use Monte Carlo integration over depth to derive optimal capture strategies for different camera designs, under a wide range of photographic scenarios. We also derive a new upper bound on how well spatial frequencies can be preserved over the depth of field. Our results show that by capturing the optimal number of photos, a standard camera can achieve performance at the level of more complex computational cameras, in all but the most demanding of cases. We also show that computational cameras, although specifically designed to improve one-shot performance, generally benefit from capturing multiple photos as well.

1. Introduction

Recent years have seen many proposals for tightly integrating sensing, optics and computation in order to extend the capabilities of the traditional camera. Already, numerous “computational camera” designs exist for capturing photos with reduced motion blur [30, 20], post-capture refocusing capabilities [17, 33, 3], and an extended depth of field (DOF) [4, 11, 23, 17, 33, 19]. Although these designs differ in many respects, they all adhere to the principle of one-shot capture: the camera records a single image with a DOF constrained by the optics and an exposure time constrained by the available time budget (or by pixel saturation).

In this paper we show that one-shot capture is usually *not* optimal for extended-DOF photography, *i.e.*, it does not produce a well-focused image with the highest signal-to-noise ratio (SNR) for a desired DOF and time budget. Moreover, we show that this result applies to standard and computational cameras [4, 23, 17, 19] alike: image quality in both cases can often be improved by capturing many shots within a given time budget, rather than just one.

Our analysis is based on a key insight illustrated in Fig. 1: by spreading the time budget across several "under-exposed" shots with different focus settings we can obtain reduced worst-case blur, at the expense of higher sensor noise. In particular, read noise leads to a penalty for each photo we capture, but Poisson-distributed photon noise *does not* penalize multiple shots. Since photon noise dominates read noise under normal photographic conditions, the overall SNR usually tips in favor of splitting the time budget. We call this the *time-slice advantage*. By contrast, single-shot photography is only optimal for very limited time budgets, *i.e.*, when read noise becomes significant, or for cameras with high per-shot overhead.

Figure 1. The time-slice advantage for a desired DOF and time budget T. One-shot capture gives the brightest image, but depths far from the lens' DOF (red rectangle) are blurred significantly. A standard focal stack “spans” the desired DOF: photos are exposed less but every depth is effectively blur-free in one of them. In this way, the number of shots acts as a balancing factor between under-exposure and worst-case blur. When photon noise dominates, the optimal-SNR capture strategy tilts even further to the right.

In this paper we provide a detailed study of the time-slice advantage and use it for *optimal time-constrained photography*—creating an all-in-focus image with the highest SNR for a given camera design, time budget, target DOF, and average scene brightness. Working from first principles, we formulate all-in-focus photography as a frequency-based restoration problem that takes noisy and optically-blurred photos as input, and outputs a single, all-in-focus image for the target DOF. This leads to three basic questions:

- **camera-specific optimal time allocation:** given a camera's noise model and optical transfer function [29], how should we allocate the time budget to maximize the expected SNR of the all-in-focus image?
- **optics-independent performance bound:** what is the maximum attainable expected SNR for a given sensor across all possible optical transfer functions?
- **camera performance characterization:** how do the existing extended-DOF camera designs compare in terms of their attainable expected SNR, and how do they fare against the traditional camera?

Our answer to these questions can be viewed as complementing and generalizing several lines of recent work.

Closest to our work, Hasinoff and Kutulakos [10] studied a related problem in extended-DOF photography: minimizing the time it takes to capture a given DOF while maintaining ideal exposure. Their work considers multiple photos, but it ignores the effect of noise and uses a basic view of

Noise-Optimal Capture for High Dynamic Range Photography

Samuel W. Hasinoff   Frédo Durand   William T. Freeman
Massachusetts Institute of Technology
Computer Science and Artificial Intelligence Laboratory

Abstract

Taking multiple exposures is a well-established approach both for capturing high dynamic range (HDR) scenes and for noise reduction. But what is the optimal set of photos to capture? The typical approach to HDR capture uses a set of photos with geometrically-spaced exposure times, at a fixed ISO setting (typically ISO 100 or 200). By contrast, we show that the capture sequence with optimal worst-case performance, in general, uses much higher and variable ISO settings, and spends longer capturing the dark parts of the scene. Based on a detailed model of noise, we show that optimal capture can be formulated as a mixed integer programming problem. Compared to typical HDR capture, our method lets us achieve higher worst-case SNR in the same capture time (for some cameras, up to 19 dB improvement in the darkest regions), or much faster capture for the same minimum acceptable level of SNR. Our experiments demonstrate this advantage for both real and synthetic scenes.

1. Introduction

Taking multiple exposures is an effective solution to extend dynamic range and reduce noise in photographs. However, it raises a basic question: what should the set of exposures be? Most users rely on a geometric progression where the exposure times are spaced by factors of 2 or 4 with the number of images set to cover the range. The camera sensitivity (ISO) is usually fixed to the nominal value (typically 100 or 200) to minimize noise. Given that noise is the main factor that limits dynamic range in the dark range of values, it is critical to understand how noise can be minimized in high dynamic range (HDR) imaging. In this paper, we undertake a systematic study of noise and reconstruction in HDR imaging and compute the optimal exposure sequence as a function of camera and scene characteristics.

We present a model that predicts signal-to-noise ratio at all intensity levels and allows us to optimize the set of exposures to minimize worst-case SNR given a time budget, or to achieve a given minimum SNR in the fastest time. To do this, we use a detailed model of camera noise that takes into account photon noise, as well as additive noise before and after the ISO gain. This allows us to optimize all parameters of an exposure sequence, and we show that this reduces to solving a mixed integer programming problem. In particular, we show that, contrary to suggested practice (*e.g.*, [5]), using high ISO values is desirable and can enable significant gains in signal-to-noise ratio.

The most important feature of our noise model is its explicit decomposition of additive noise into pre- and post-amplifier sources (Fig. 1), which constitutes the basis for the high ISO advantage. The same model has been used in several unpublished studies characterizing the noise performance of digital SLR cameras [7, 20], supported by extensive empirical validation. Although all the components in our model are well-established, previous treatments of noise in the vision literature [13, 18] do not model the dependence of noise on ISO setting (*i.e.*, sensor gain).

To the best of our knowledge, varying the ISO setting has not previously been exploited to optimize SNR for high dynamic range capture. However, in the much simpler context of single-shot photography, the *expose to the right* technique [25, 20] considers the ISO setting to optimize SNR. This technique advocates using the *lowest* ISO setting possible, but increasing ISO when the exposure time is tightly constrained. Another related idea is the dual-amplifier sensor proposed by Martinec, which would capture exposures at ISO 100 and 1600 simultaneously and then combine them to extend dynamic range [20]. Our method can be thought of as formalizing these ideas, generalizing them to a multi-shot setting, and showing how to optimize the capture sequence for a given camera and scene.

Most previous work in HDR imaging has focused on calibrating the response curve of the sensor [8, 22], merging the input images [16, 3], and tone mapping the merged HDR result [17, 9]. Surprisingly little attention, however, has been paid to the capture strategy itself, which is the focus of this paper. One notable exception is a method that computes the optimal set of exposure times to reduce quantization in the merged HDR result [10]. This works by effectively dithering the exposure levels, but assumes that exposure times can be controlled arbitrarily, and does not incorporate a detailed model of noise. Another recent method [4] showed how to minimize the number of photos spanning a given dynamic range, but takes a simplified geometric view of dynamic range, without any noise model.

Denoising vs. Deblurring: HDR Imaging Techniques Using Moving Cameras

Li Zhang   Alok Deshpande   Xin Chen
University of Wisconsin, Madison   ACM
`http://www.cs.wisc.edu/~lizhang/projects/motionhdr/`

Abstract

New cameras such as the Canon EOS 7D and Pointgrey Grasshopper have 14-bit sensors. We present a theoretical analysis and a practical approach that exploit these new cameras with high-resolution quantization for reliable HDR imaging from a moving camera. Specifically, we propose a unified probabilistic formulation that allows us to analytically compare two HDR imaging alternatives: (1) deblurring a single blurry but clean image and (2) denoising a sequence of sharp but noisy images. By analyzing the uncertainty in the estimation of the HDR image, we conclude that multi-image denoising offers a more reliable solution. Our theoretical analysis assumes translational motion and spatially-invariant blur. For practice, we propose an approach that combines optical flow and image denoising algorithms for HDR imaging, which enables capturing sharp HDR images using handheld cameras for complex scenes with large depth variation. Quantitative evaluation on both synthetic and real images is presented.

1. Introduction

High Dynamic Range (HDR) Imaging has been an active topic in vision and graphics in the last decade. Debevec and Malik [14] developed the widely-used approach that combines multiple photos with different exposure to create an HDR image. This approach is well suited to early digital cameras, which often have 8-bit Analog-to-Digial conversion (ADC). Today, many consumer SLRs or machine vision cameras have higher resolution ADC; for example, Canon EOS 7D and Point Grey Grasshopper have 14-bit ADC, and many others have at least 12-bit ADC. In this paper, we present an effective approach that exploits new cameras with high-resolution ADC to widen the operating range of HDR imaging.

The inconvenient requirement of [14] is that the camera must remain still during the image acquisition and the scene must be static. The requirements of a still camera and scene are due to the need for long-exposure shots to record dark image regions accurately. Any motion of the camera or of the scene will introduce blur in the image. This requirement will not be simply relieved by using a 14-bit sensor, because the lower bits of each pixel only encode the noise accurately.

To capture a good HDR image in a flexible setting, without assuming stationary scenes or cameras, we have to either accumulate more photons using a long exposure and later remove the motion blur, or accumulate less photons using a short exposure and later remove the noise. Since the second approach takes less time, within a fixed time budget, we can take more images for better noise reduction. In this paper, we present a probabilistic formulation that allows us to compare which of denoising and deblurring can produce better HDR images. Specifically, we compare the following HDR imaging choices:

- Deblurring a single blurry but clean image captured with a long exposure time Δ and a low ISO setting;
- Denoising a series of sharp but noisy images, each captured with a high ISO, together captured within time Δ.

We note that a high-resolution ADC is *essential* for both the procedures to succeed, in particular for denoising, because the noise must be digitized accurately to be averaged out among the multiple frames. Our contributions include:

- We propose a novel probability formulation that unifies both single-image deblurring and multi-image denoising. These two problems are formulated differently in the literature; comparing their solutions analytically is difficult.
- Using variational inference with motion as hidden variables, we derive the approximate uncertainty in the estimation of HDR images analytically for both imaging procedures. Our conclusion is that denoising is a better approach for HDR imaging.
- To put our analytical insight to practical use, we present a novel approach that combines existing optical flow and image denoising techniques for HDR imaging. This approach enables capturing sharp HDR images using handheld cameras for complex scenes with large depth variation. Such scenes cause spatially-varying motion blur for handheld cameras, which cannot be handled by the latest HDR imaging method [22].

Large depth-of-field, high dynamic range, and small motion blur are three of the major goals of computational camera research. Our work shows that, if a camera has high-resolution ADC, high frame rate, and high ISO, it is possible to achieve all the three goals through computation without resorting to specialized optical designs. This feature makes our approach suitable to micro-cameras with simple optics, such as those found in cellphones or used in performing surgeries.

2. Related Work

Our work is related to the recent research combining multiple images of different exposure to produce a sharp and clean image. Yuan *et al.* [27] and Tico and Vehvilainen [24] combined a noisy and blurry image pair, and Agrawal *et al.* [3] combined multiple blurry image with different exposure; all this research is limited to spatially-invariant blur.

One approach to address this limitation is to use video denoising techniques on multiple noisy images. In particular, our work is inspired by Boracchi and Foi [6], who combined a state-of-the-art video denoising method, VBM3D [12], and homography-based alignment for multi-frame denoising. They compared debluring a noisy and blurry image pair and

1

Long-term prospects

- high-speed burst-mode photography
  - all still cameras should capture at 500 fps
  - capture while aiming the camera - no shutter half-press
  - frameless photography - ROIs, MOIs ("M" = Moment)
- computational videography & cinematography
  - stereo, view interpolation, free-viewpoint video
  - stabilization

3D video stabilization

[Agarwala 2011]

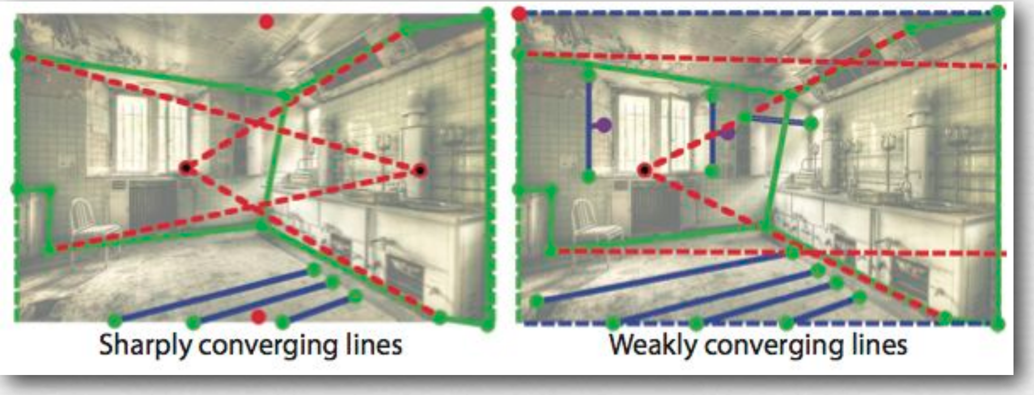
(http://web.cecs.pdx.edu/~fliu/project/subspace_stabilization/demo.mp4)

Long-term prospects

- high-speed burst-mode photography
  - all still cameras should capture at 500 fps
  - capture while aiming the camera - no shutter half-press
  - frameless photography - ROIs, MOIs (“M” = Moment)
- computational videography & cinematography
  - stereo, view interpolation, free-viewpoint video
  - stabilization
  - extending computational photography to video
    - HDR, EDoF, plenoptic refocusing
    - retargeting (a.k.a. content-aware image resizing)
    - perspective manipulation

Image warps for Artistic Perspective Manipulation

[Carroll SIGGRAPH 2010]



Sharply converging lines

Weakly converging lines

extendable to video?

Input

Result: Sharply converging lines

Result: Weakly converging lines

- specify vanishing points and line constraints manually
- image is warped to optimally satisfy all constraints
- resulting image is not a correct linear perspective

Long-term prospects

- high-speed burst-mode photography
    - all still cameras should capture at 500 fps
    - capture while aiming the camera - no shutter half-press
    - frameless photography - ROIs, MOIs (“M” = Moment)
- computational videography & cinematography
    - stereo, view interpolation, free-viewpoint video
    - stabilization
    - extending computational photography to video
        - HDR, EDoF, plenoptic refocusing
        - retargeting (a.k.a. content-aware image resizing)
        - perspective manipulation
    - style transfer, non-photorealistic video

The HDR "look"

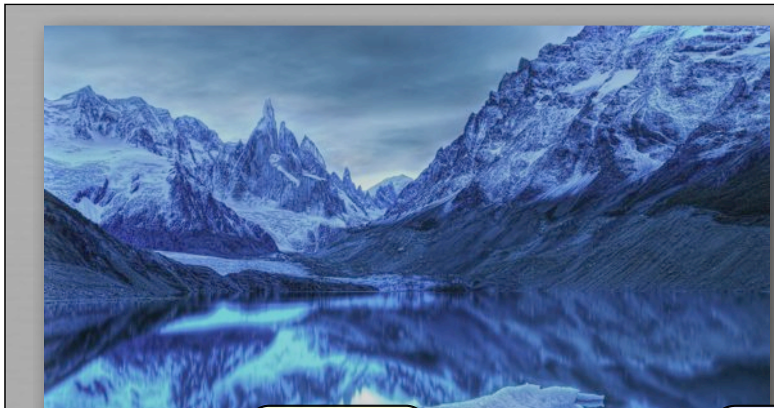
(Trey Ratcliff, http://www.stuckincustoms.com)

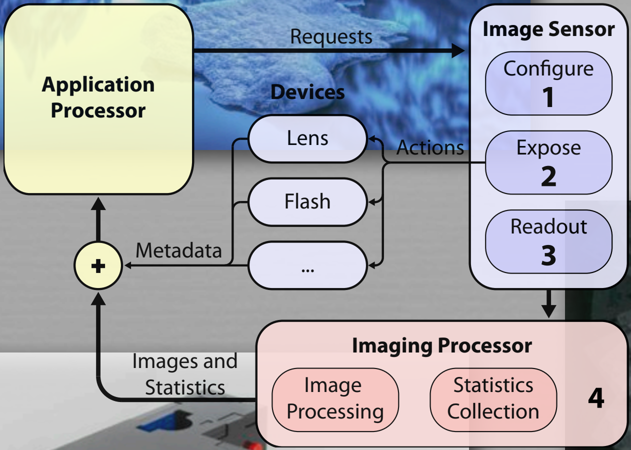
The HDR "look"

(Trey Ratcliff, http://www.stuckincustoms.com)

The HDR "look"

(Trey Ratcliff, http://www.stuckincustoms.com)





Application Processor

Requests

Image Sensor

Configure 1

Expose 2

Readout 3

Devices

Lens

Flash

...

Actions

Metadata

+

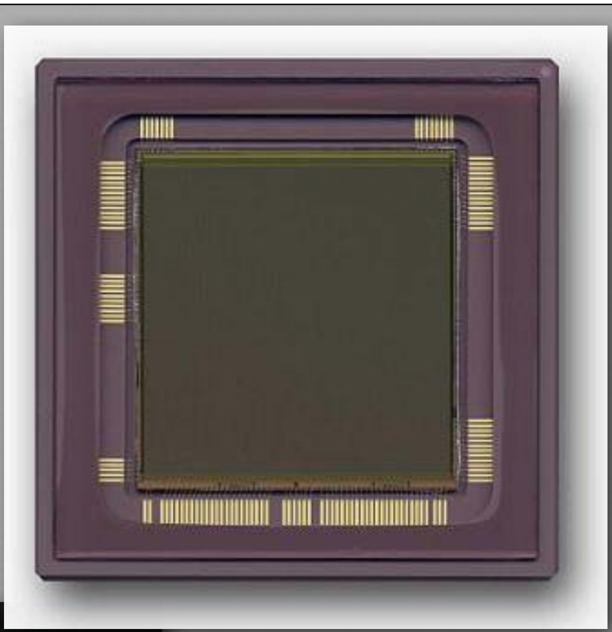
Imaging Processor

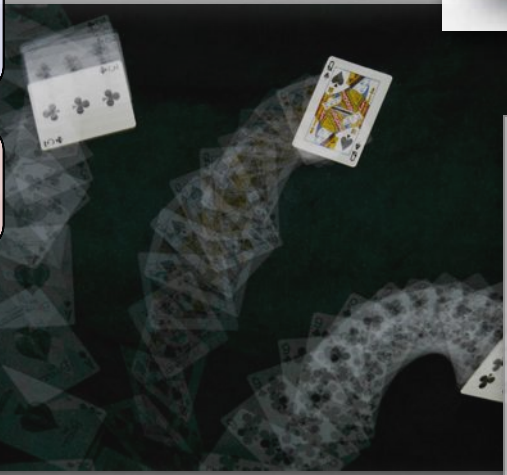
Image Processing

Statistics Collection

4

Images and Statistics





```
Sensor sensor;
Shot low, med, high;

low.exposure = 1/80.;
med.exposure = 1/20.;
high.exposure = 1/5.;

sensor.capture(low);
sensor.capture(med);
sensor.capture(high);

Frame frames[3];
frames[0] = sensor.getFrame();
frames[1] = sensor.getFrame();
frames[2] = sensor.getFrame();

fused = mergeHDR(frames);
```

http://graphics.stanford.edu/projects/camera-2.0/