A Framework for Analysis of Computational Imaging Systems

Kaushik Mitra, Oliver Cossairt, Ashok Veeraghavan

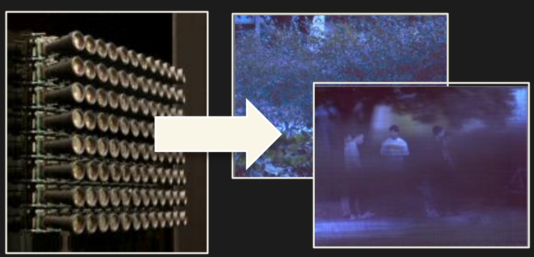
Rice University

Northwestern University

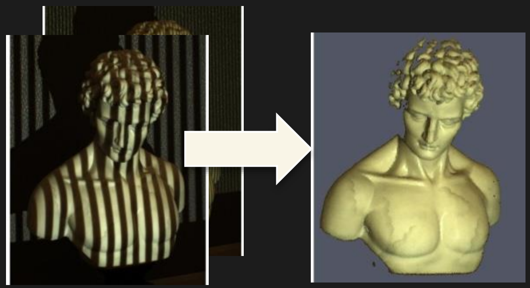
Computational imaging

CI systems that adds new functionality

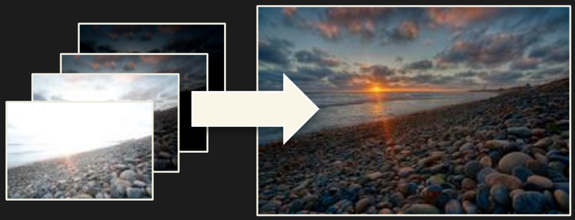
Light Field Capture



Structured Lighting

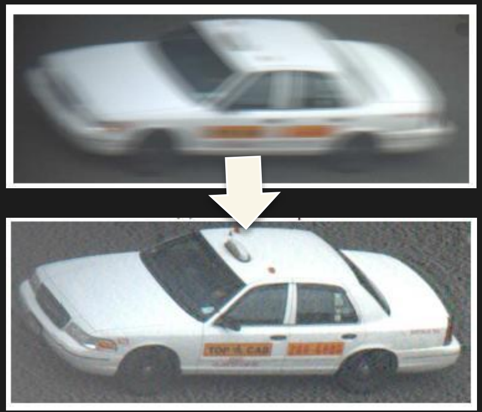


HDR Imaging

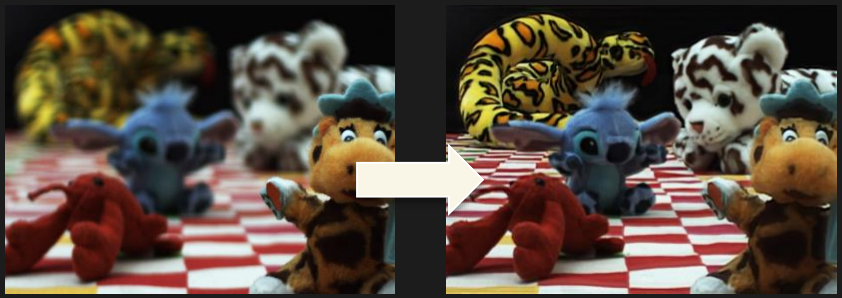


CI systems that improves performance

Motion deblurring system



Extended depth of field



Others: Multiplexed
- Light field
- Illumination
- Spectography

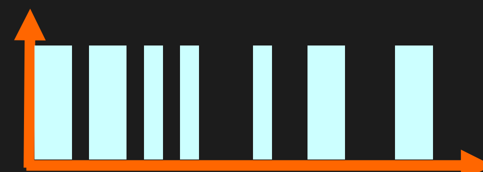
How does CI improve performance?

Increased light throughput

Short exposure

Sharp, but noisy

Flutter Shutter



Increased light throughput but blurry

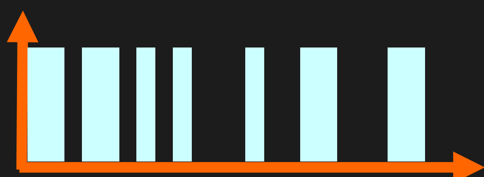
Deblurred image

Slide coutesy Amit Agarwal

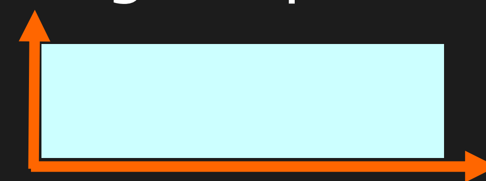
How does CI improve performance?

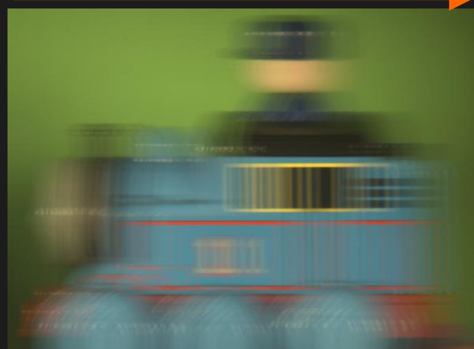
Well conditioned optical coding

Flutter Shutter

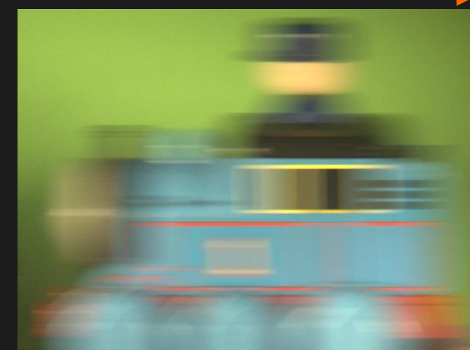


Large exposure

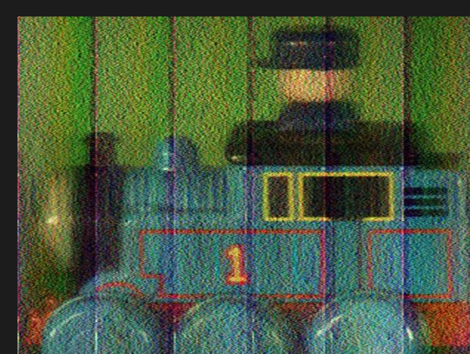




Captured image

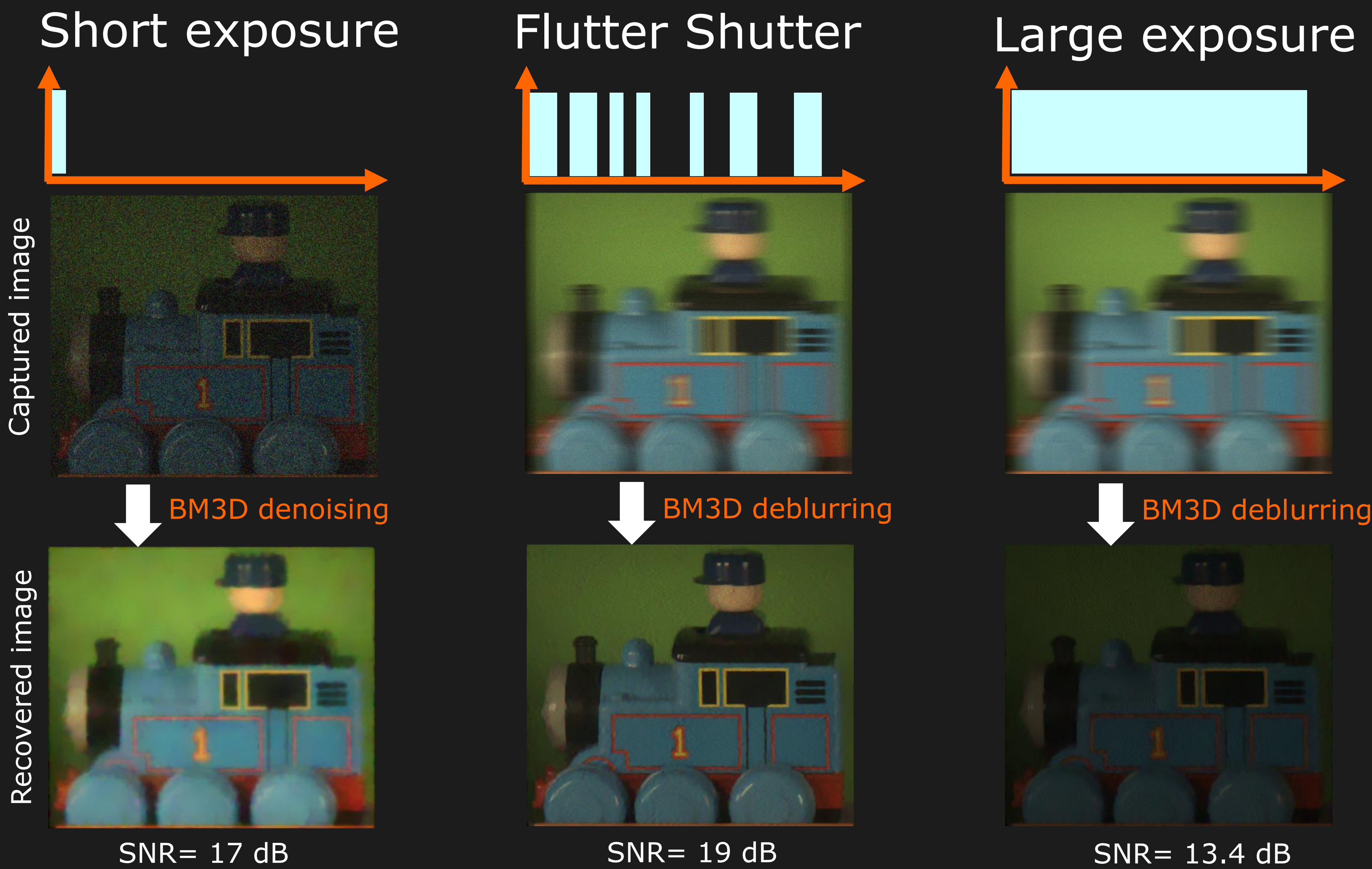


Deblurred image



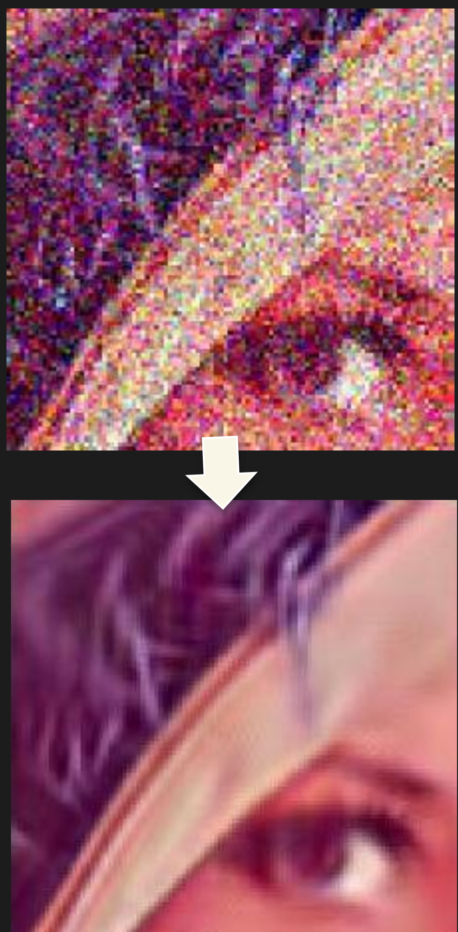
Slide coutesy Amit Agarwal

One key flaw: Signal prior has not been taken into account



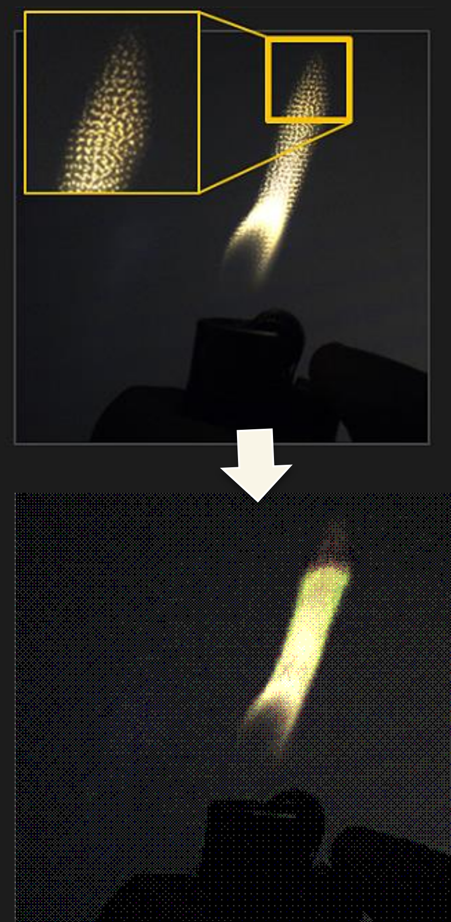
State-of-the-art systems use signal prior

Denoising using BM3D



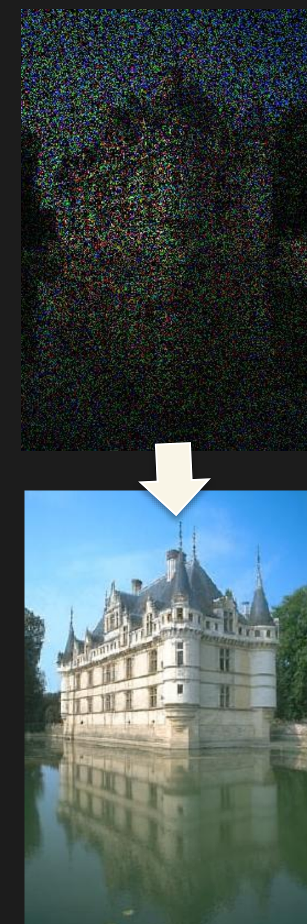
Dabov et al., 2011

Coded exposure video using dictionary learning



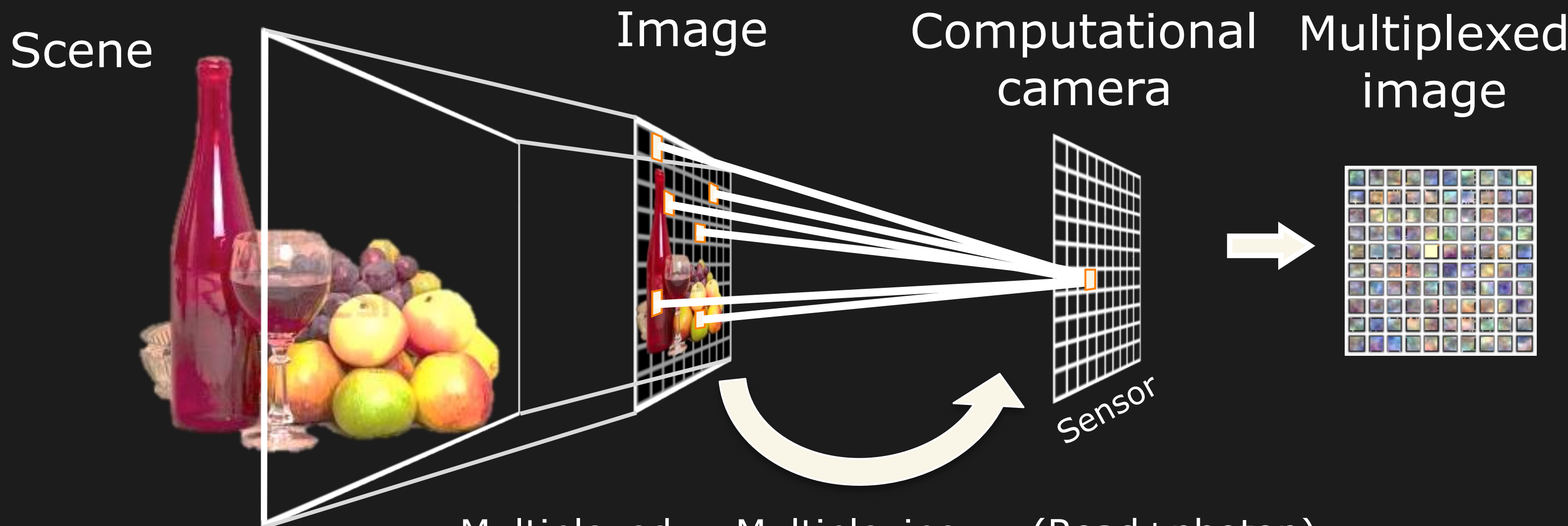
Hitomi et al., 2011

Inpainting using GMM



Yu et al., 2011

Our goal: A comprehensive analysis



Multiplexed image — Multiplexing matrix — (Read+photon) Noise

$$y = Hx + n$$

Signal prior P(x)

Our analysis takes into account:

- Signal prior
- Multiplexing matrix
- Noise characteristics

Figure courtesy Oliver Cossairt

Prior Work: Analysis of CI systems

1. Analysis under read noise withour prior

$$y = Hx + n_r$$

Harwitt et al. 1979

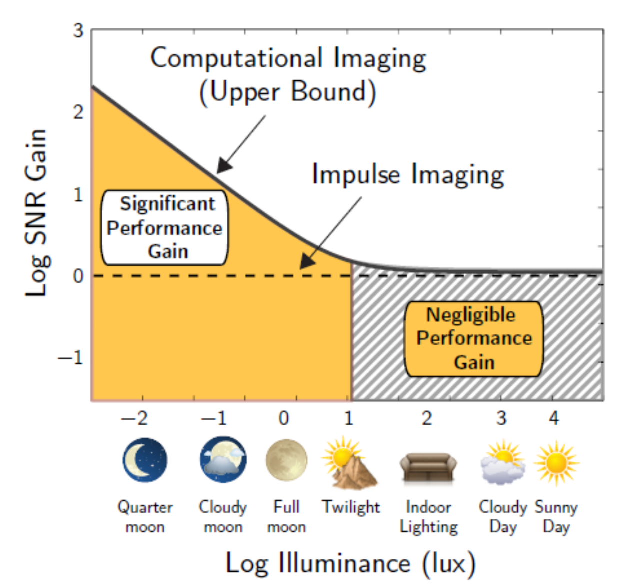
2. Analysis under affine noise withour prior

$$y = Hx + n_r + n_p$$

Ratner et al. 2007, Wuttig 2007, Hasinoff et al. 2008, Ihrke et al. 2010, Cossairt et al. 2011

3. Relates performance to practical considerations such as illumination, sensor characteristics, etc.

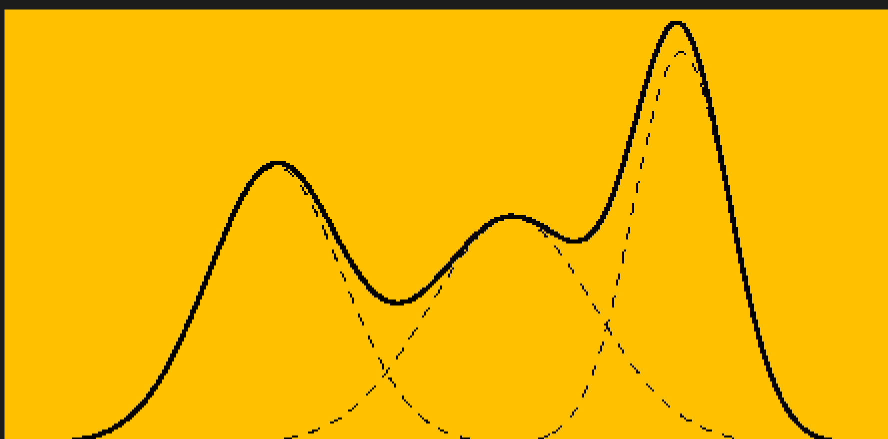
Cossairt et al. 2012



Our analysis framework: GMM as signal prior

Advantages of GMM

1.Universal approximation property



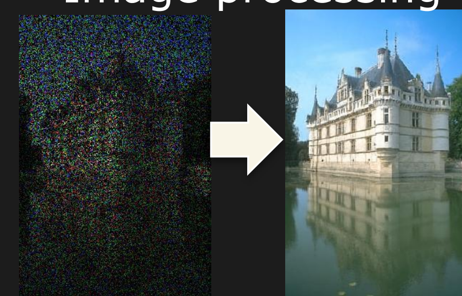
Sorenson et al., 1971

2. Analytically tractable

A special case is Gaussian prior, whose MMSE can be computed analytically

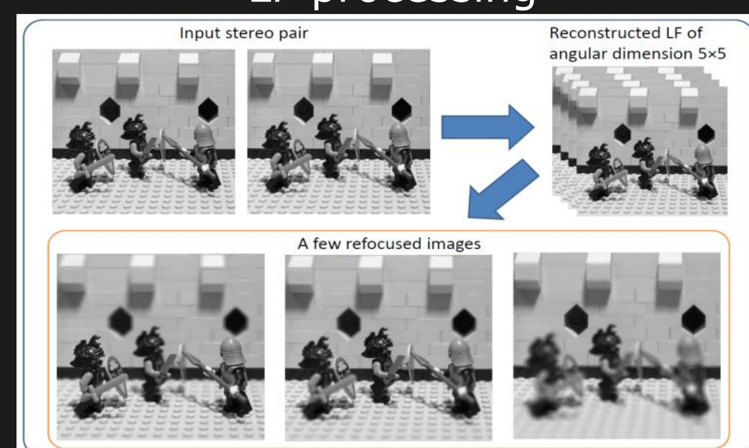
3. State-of-the-art results

Image processing



Yu et al. 2010

LF processing



Mitra et al. 2012

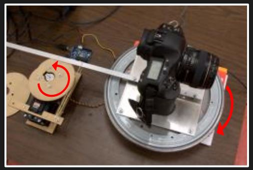
Our analysis framework: Linear system

Multiplexed image — Multiplexing matrix — Noise

$$y = Hx + n$$

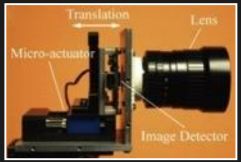
Motion blur

[Raskar '06]

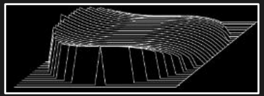


[Levin '08]
[Cho '10]

Defocus blur



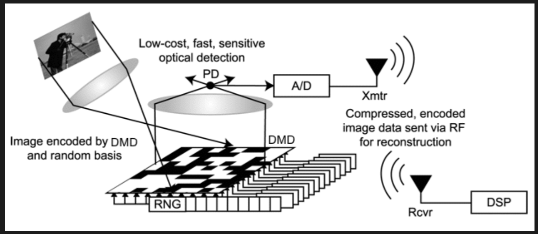
[Hausler '72]
[Nagahara '08]



[Dowski, Cathey '96]

[Levin et al. '07]
[Zhou, Nayar '08]

Single pixel camera



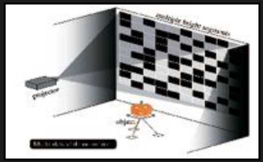
[Wakin et al., 2006]

Light Field Capture

[Lanman '08]

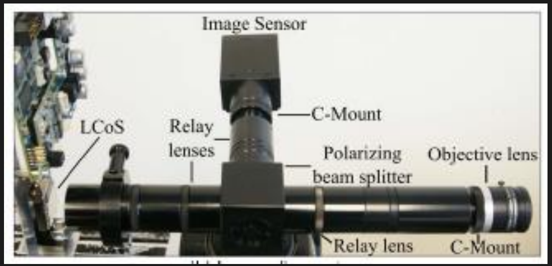
[Veeraraghavan '07]
[Liang '08]

Reflectance



[Schechner '03]
[Ratner '07]
[Ratner '08]

High speed video



[Hitomi et al. 2011][Veera et al., 2011]

Our analysis framework: Affine noise model

Noise Variance at ith Pixel:

$$\sigma_i^2 = J_i + \sigma_r^2$$

- photon noise (J_i): aperture, lighting, pixel size
- read noise (σ_r^2): electronics, ADC's, quantization

- J_i: ith pixel intensity
- Signal dependent / independent noise
- Ignore Dark current, fixed pattern

Noise PDF:

$$n \sim N(0, \sigma^2 I)$$

$$\bar{J} = \sum_i J_i / K \quad \sigma^2 = \bar{J} + \sigma_r^2$$

- Photon noise modeled as Gaussian (good approx. if #photons > 10)
- Photon noise spatially averaged

Slide courtesy Oliver Cossairt

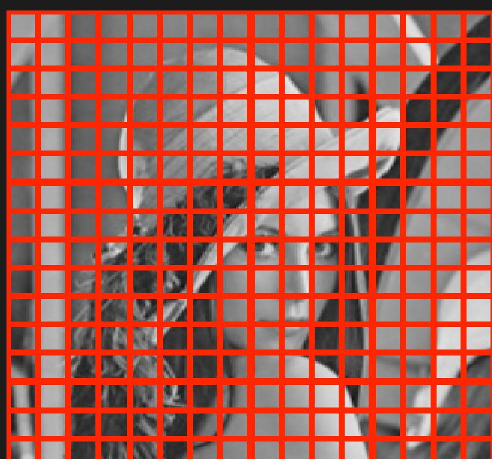
Complete specification of the framework

Multiplexed measurement — Multiplexing matrix — Noise

$$y = Hx + n \;\Rightarrow\; P(n) \sim N(0, \sigma^2 I)$$

$$\sigma^2 = \bar{J} + \sigma_r^2$$

Learn patch-based GMM prior



GMM Cluster 1: mean and PCA components

GMM Cluster 2: mean and PCA components

⋮

$$P(x) = \sum_{i=1}^{K} \alpha_i N(m_i, \Sigma_i)$$

GMM patch prior

- α_i → Cluster weight
- m_i → Cluster mean
- Σ_i → Cluster covariance

MMSE as a performance metric

Mean Squared Error (MSE) of an estimator $\hat{x}$ is defined as:

$$MSE = Tr(E[(x-\hat{x})(x-\hat{x})^T])$$

MMSE estimator:

- Defined as the estimator that achieves the minimum MSE
- Given by the posterior mean ➡ $\hat{x}_{MMSE} = E[x \mid y]$
- MMSE is the corresponding MSE error

MMSE: a scalar that characterizes the performance of a system *H*

Computation of MMSE estimator

The posterior PDF is also a GMM:

$$P(x \mid y) = \sum_{k=1}^{K} \tilde{\alpha}_k N(\tilde{m}_k, \tilde{\Sigma}_k)$$

with new weights $\tilde{\alpha}_k = \alpha_k \left(\frac{P_k(y)}{\sum_i \alpha_i P_i(y)} \right)$

old weight

Probability of y coming from kth cluster

and with new mean and covariance:

$$\tilde{m}_k = m_k + \Sigma_k H^T (H \Sigma_k H^T + \Sigma_n)^{-1} (y - H m_k)$$

$$\tilde{\Sigma}_k = \Sigma_k - \Sigma_k H^T (H \Sigma_k H^T + \Sigma_n)^{-1} H \Sigma_k$$

The MMSE estimator (posterior mean): $\hat{x}_{mmse} = \sum_{k=1}^{K} \tilde{\alpha}_k \tilde{m}_k$

Interpretation of MMSE

$$mmse(H) = E_{x,y} \| x - \hat{x}_{mmse}(y) \|^2$$

$$mmse(H) = \underbrace{\sum_{k=1}^{K} \alpha_k Tr(\widetilde{\Sigma}_k)}_{\text{Intra-cluster error, can be computed analytically}} + \underbrace{\sum_{k=1}^{K} \alpha_k \int_y \| x_{mmse}(y) - \widetilde{m}_k \|^2 P_k(y) dy}_{\text{Inter-cluster error needs MC simulations}}$$

Intra-cluster error, can be computed analytically

Inter-cluster error needs MC simulations

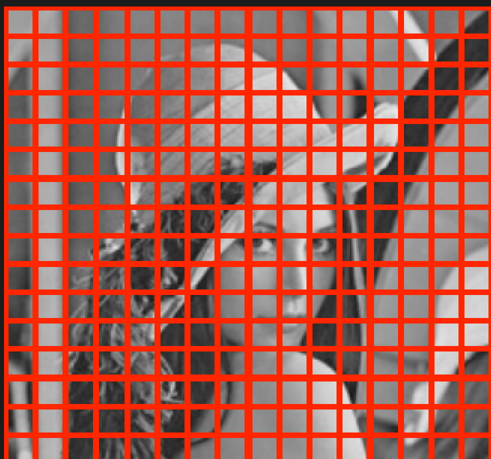
We have an analytical lower bound for the MMSE:

$$mmse(H) \geq \sum_{k=1}^{K} \alpha_k Tr(\widetilde{\Sigma}_k)$$

Tight bound for fully-determined system H

Limitations of analysis

Patch-based GMM prior

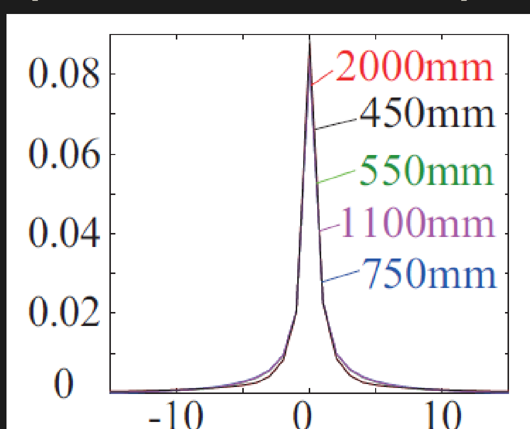


GMM Cluster 1

GMM Cluster 2

→ Local multiplexing

Shift invariant blur (motion and focus)



[Nagahara 08] [Dowsky 96] [Levin 08]

Other assumptions:

- Linear systems
- Affine noise

Practical implications of the analysis

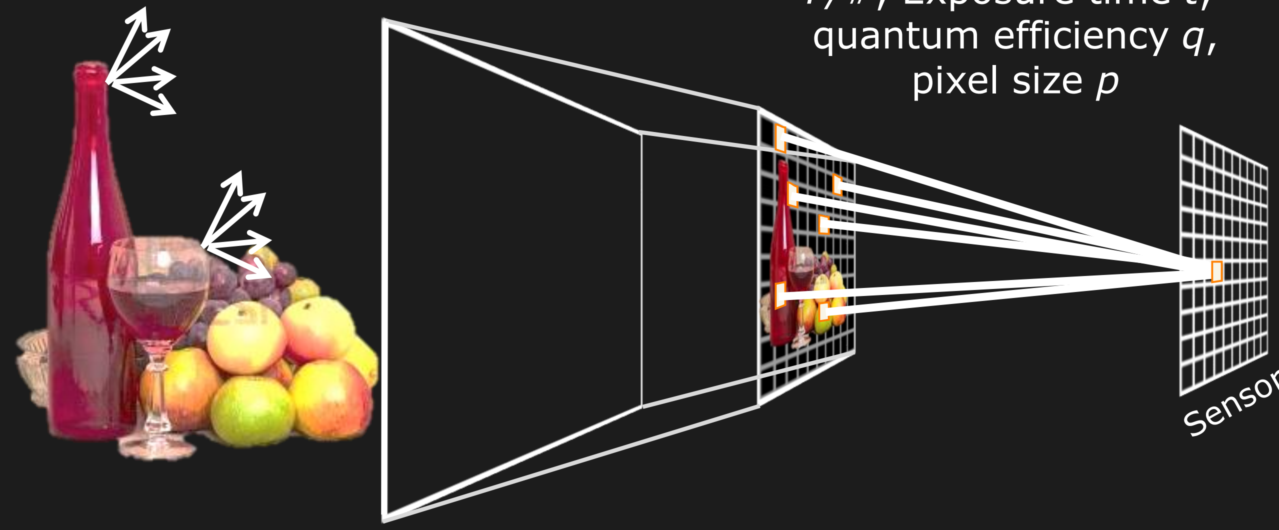
Practical system performance depends on

1. Illumination condition

2. Scene reflectivity

3. Camera parameters

F/#, Exposure time *t*, quantum efficiency *q*, pixel size *p*



Average signal-level is given by:

$$\bar{J} \approx 10^{15} \cdot I_{src} \cdot R \cdot (F/\#)^{-2} \cdot t \cdot q \cdot p^{2}$$

Average Signal (e^{-}) | Illumination (lux) | Reflectivity | Aperture | Exposure Time (s) | Quantum Efficiency | Pixel Size (m)

Average signal level for three form factors

SLR camera
Pixel size $p_{SLR} = 8\ \mu m$

Machine vision camera (MVC)
$p_{MVC} = 2.5\mu m$

Smartphone camera (SPC)
$p_{SPC} = 1\mu m$

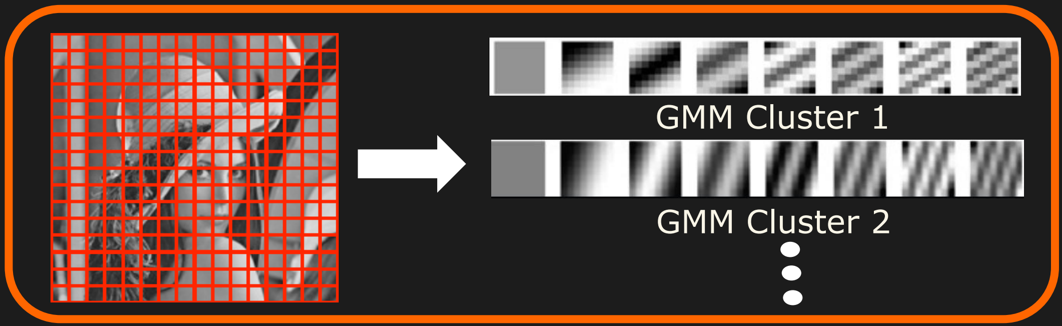
Typical values of average signal level J for different illumination levels

	Quarter moon	Full moon	Twilight	Indoor lighting	Cloudy day	Sunny day
I_{src} (lux)	10^{-2}	1	10	10^2	10^3	10^4
$J_{SLR}(e^-)$	8×10^{-3}	0.8	8.1	81.4	814.3	8143
$J_{MVC}(e^-)$	7.9×10^{-4}	7.9×10^{-2}	0.79	7.9	79.5	7952
$J_{SPC}(e^-)$	1.3×10^{-4}	1.3×10^{-2}	0.13	1.27	12.7	127

Other parameters:
$q=.5$,
$R = .5$,
$F/11$,
$t = 6ms$,
$\sigma_r=4$

Common analysis and simulation framework

Learn GMM prior of patch size 16×16



Analytical computations:

Without prior: $mse(H) = Tr(H^{-1}\Sigma_n H^{-T})$

Under GMM prior: $mmse(H) = \sum_{k=1}^{K} \alpha_k Tr(\widetilde{\Sigma}_k) + \sum_{k=1}^{K} \alpha_k \int_y \| x_{mmse}(y) - \widetilde{m}_k \|^2 \, P_k(y) dy$

Simulation computations:

Perform per-patch reconstruction. Let y be the observed patch.

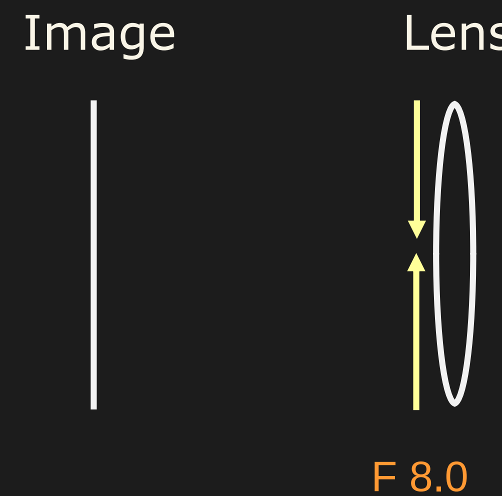
Without prior: $\hat{x} = H^{-1} y$

Under GMM prior: $\hat{x} = \sum_{k=1}^{K} \widetilde{\alpha}_k \widetilde{m}_k(y)$

Performance measure: SNR gain w.r.t impulse imaging $G(H) = 10\log\left(\frac{mse(I)}{mse(H)}\right)$

Analysis of Extended DOF systems

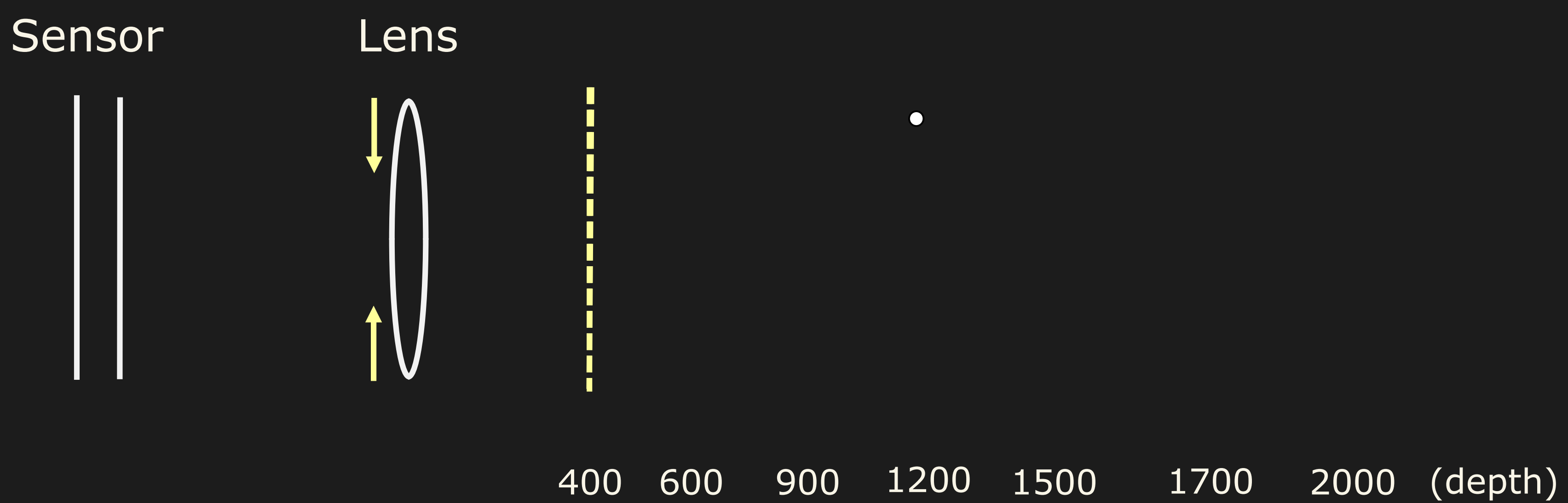
Depth of Field and SNR



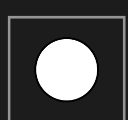
Small apertures have large depth of field and low SNR

Slide courtesy Oliver Cossairt

Focal Sweep: An example EDOF system



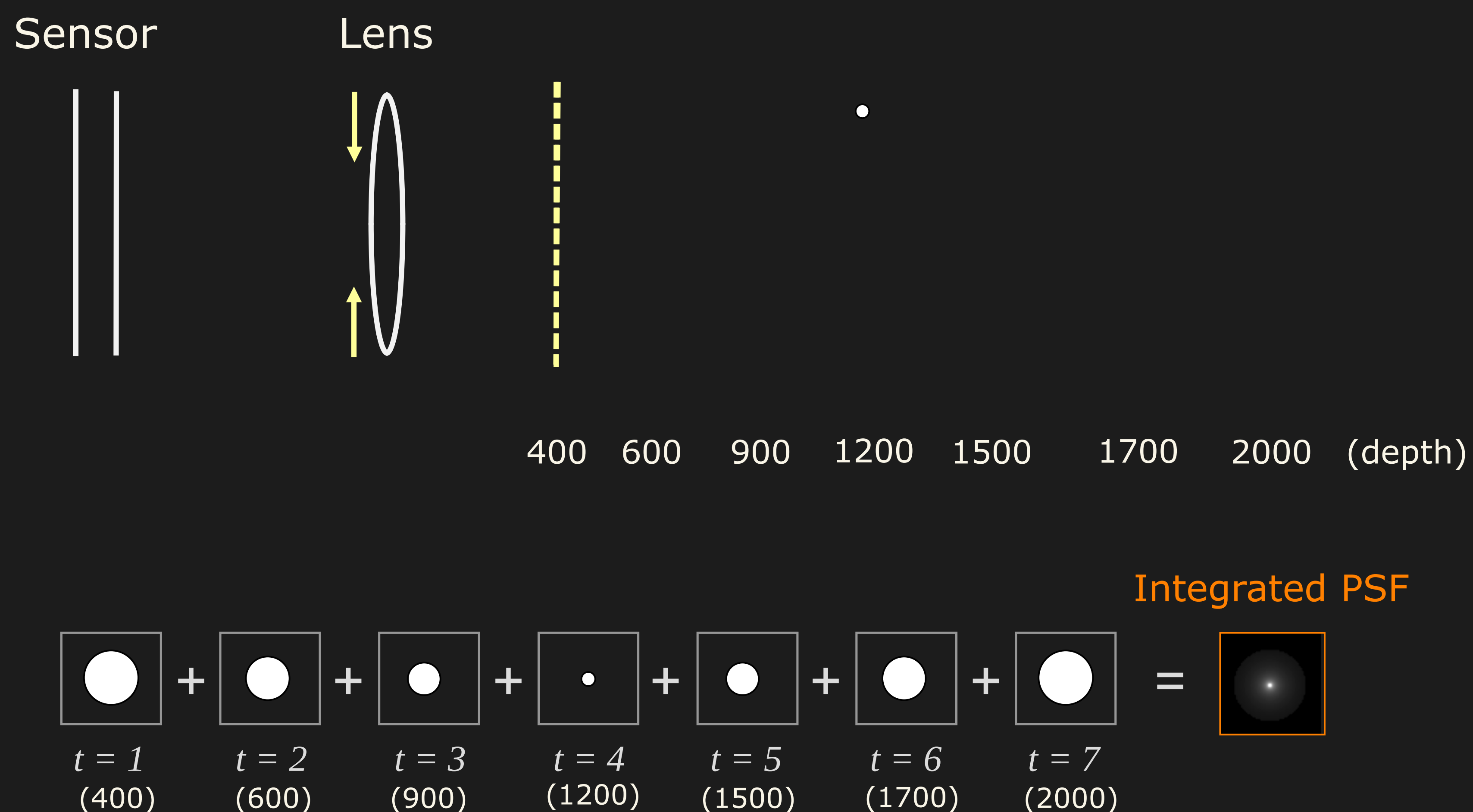
Point Spread Function (PSF)



[Hausler '72, Nagahara et al. '08]

Slide courtesy Oliver Cossairt

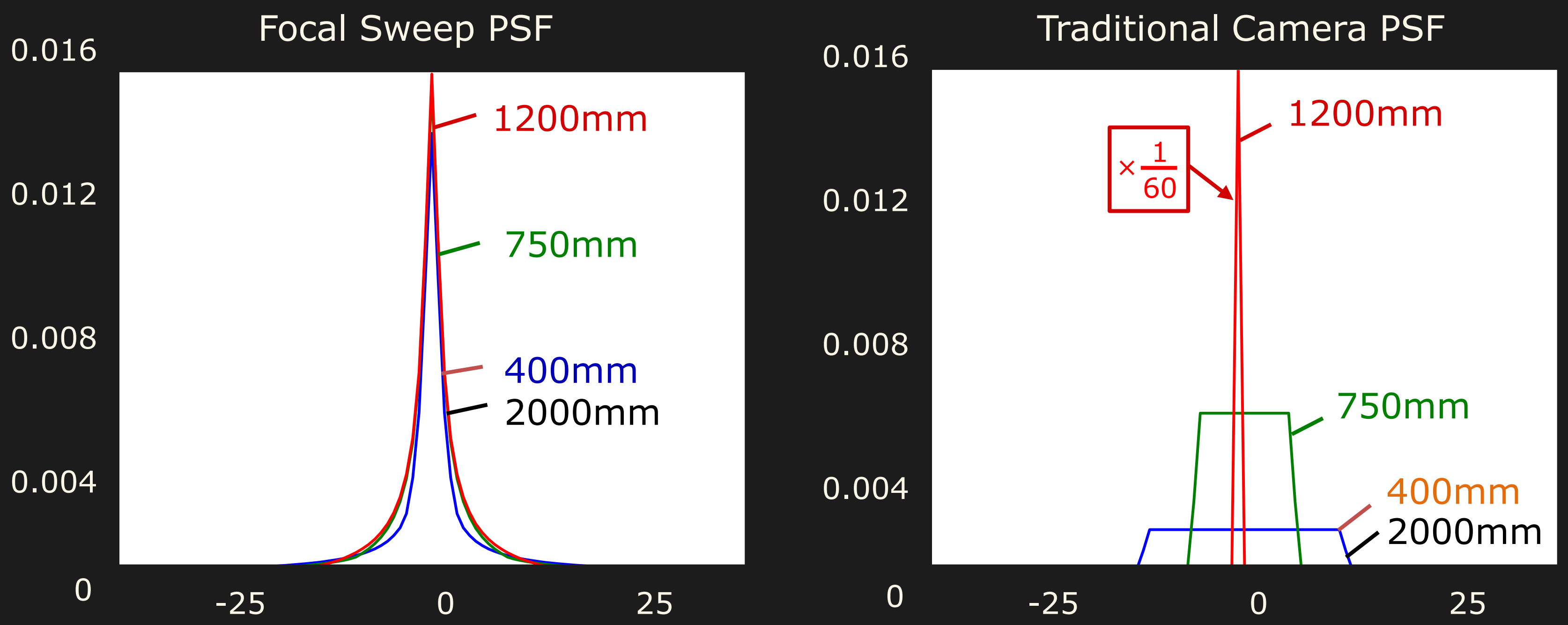
Focal Sweep: An example EDOF system



[Hausler '72, Nagahara et al. '08]

Slide courtesy Oliver Cossairt

Depth Invariant PSF



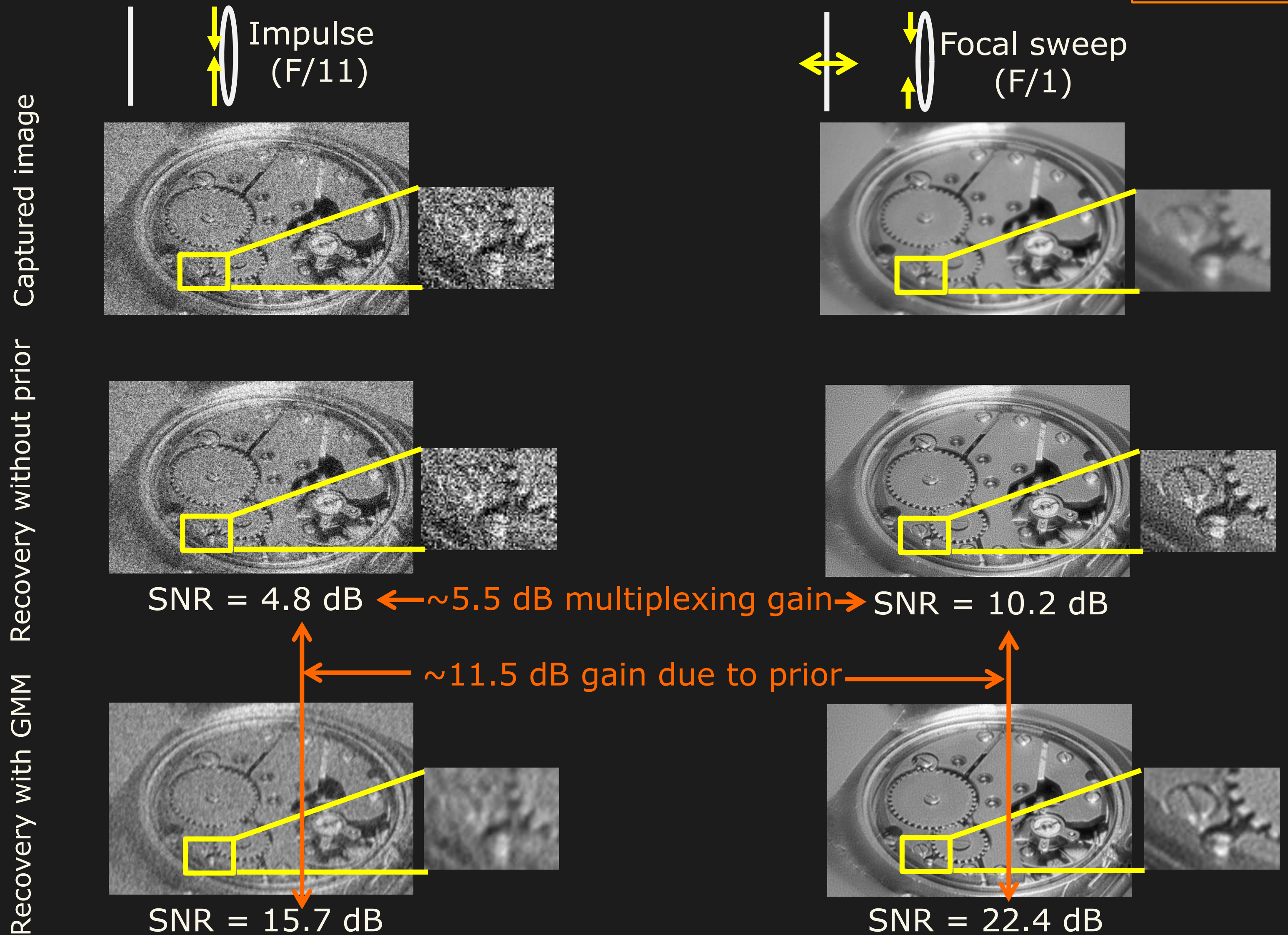
Extended depth of field with a single deconvolution

Slide courtesy Oliver Cossairt

Simulation performance

Low light condition (10 lux)

Pixel size = 8 um
Exp time = 6 ms

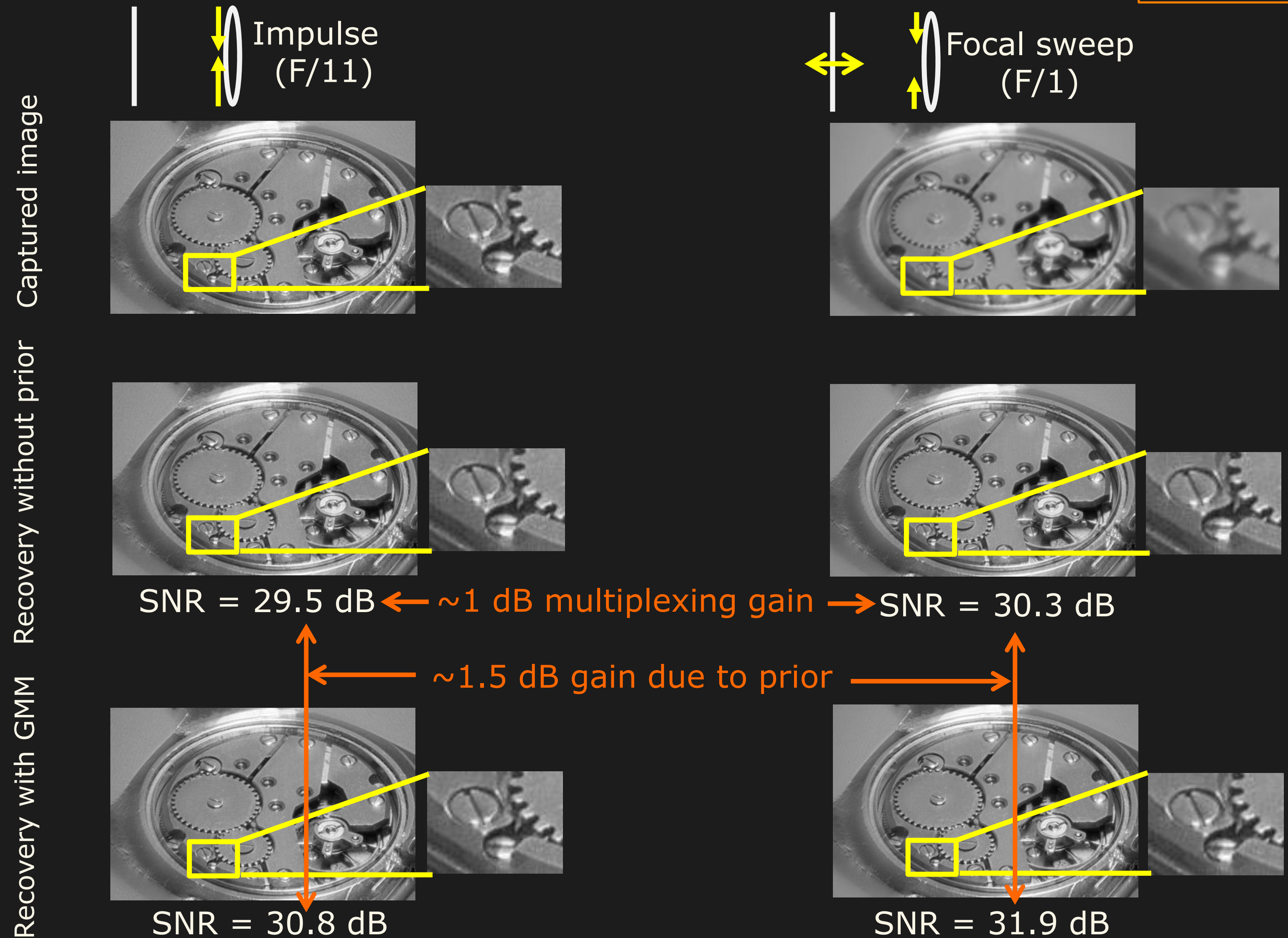


Gain due to prior is much greater than gain due to multiplexing

Simulation performance

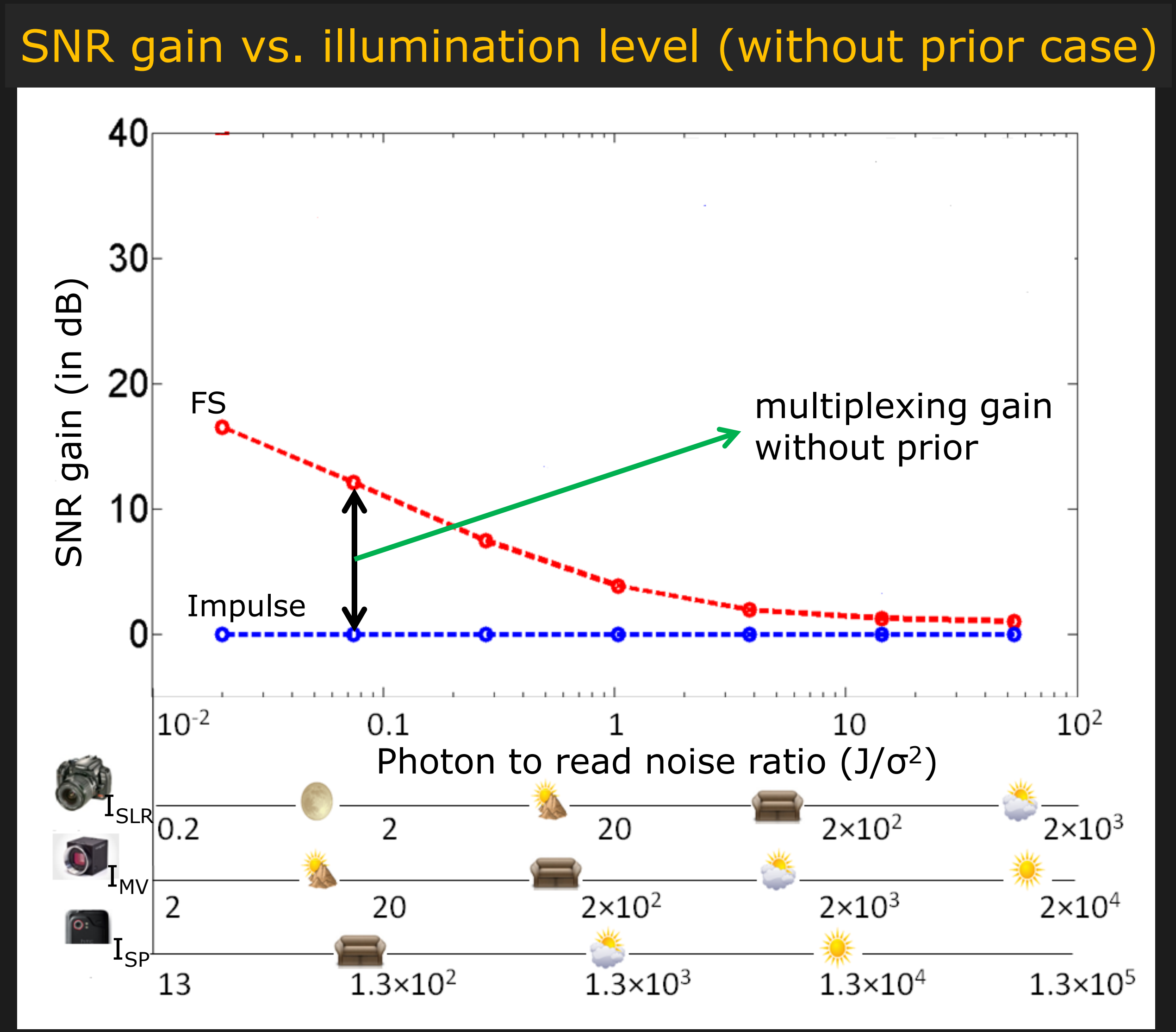
High light condition (1000 lux)

Pixel size = 8 um
Exp time = 6 ms



At high light condition, gain due to both prior and multiplexing is negligible.

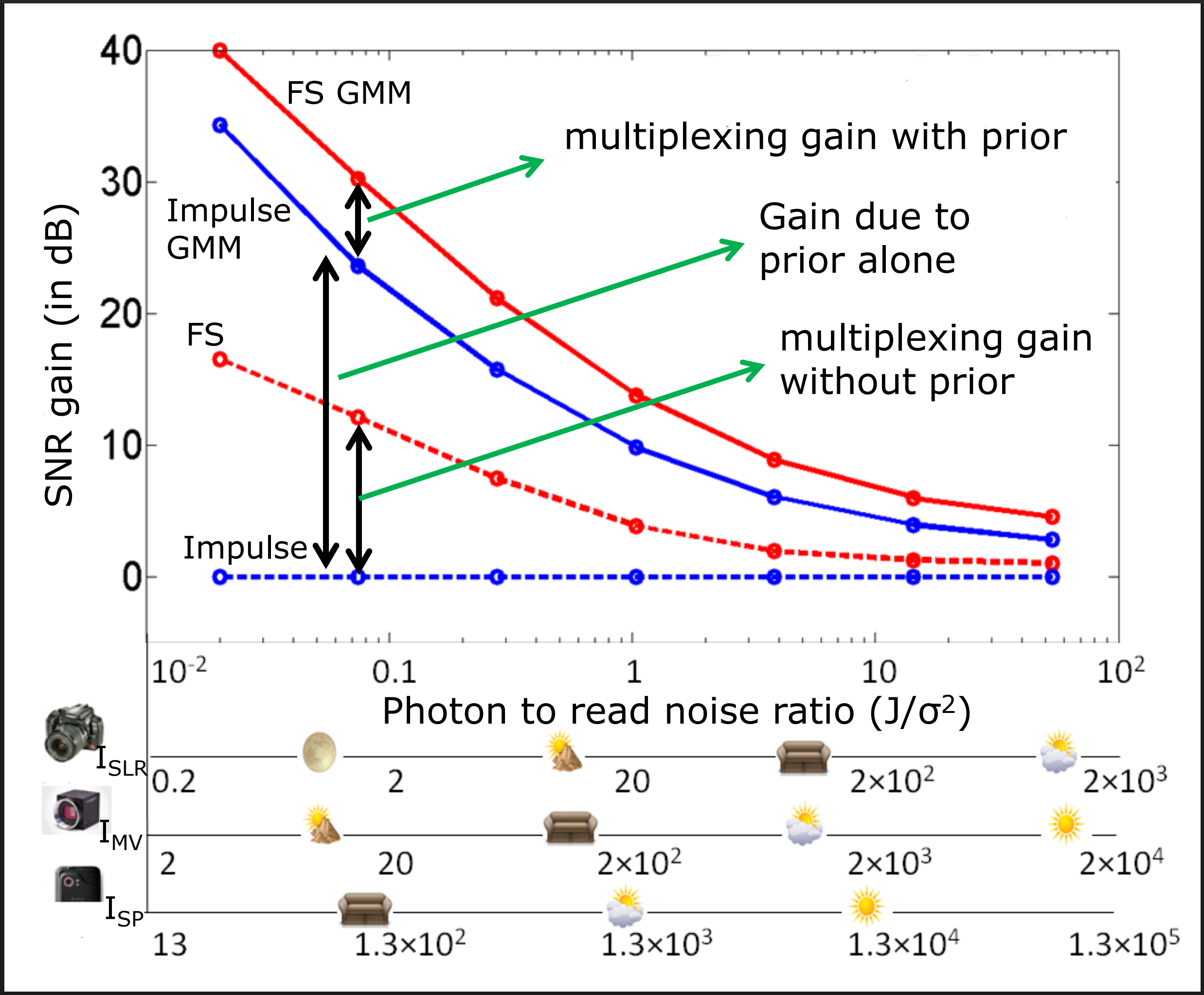
Analytic performance:



Huge multiplexing gain at low light levels

Analytic performance:

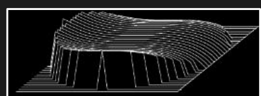
SNR gain vs. illumination level (with prior case)



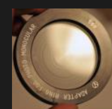
Under signal prior moderate multiplexing gain at low light levels

Other EDOF systems

Depth invariant PSF systems



[Dowski, Cathey '96]

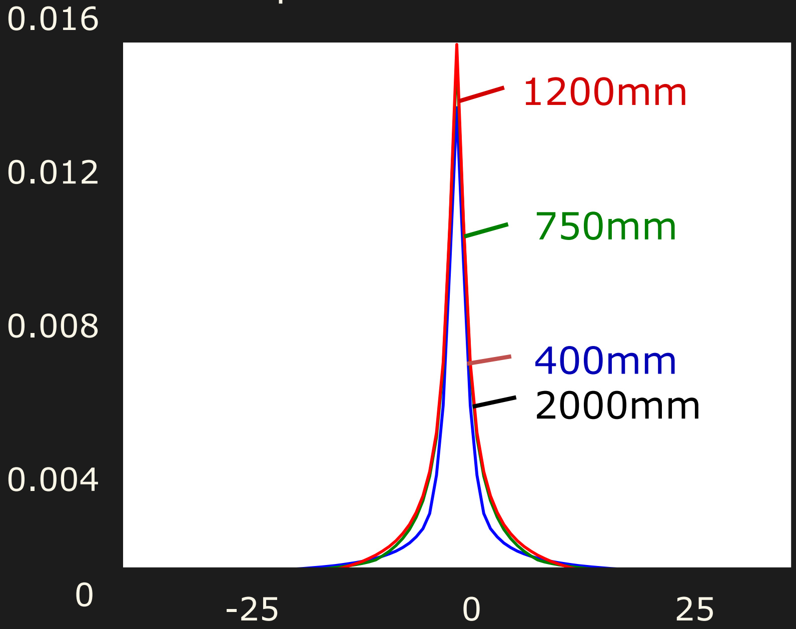


[Cossairt et al. '10]

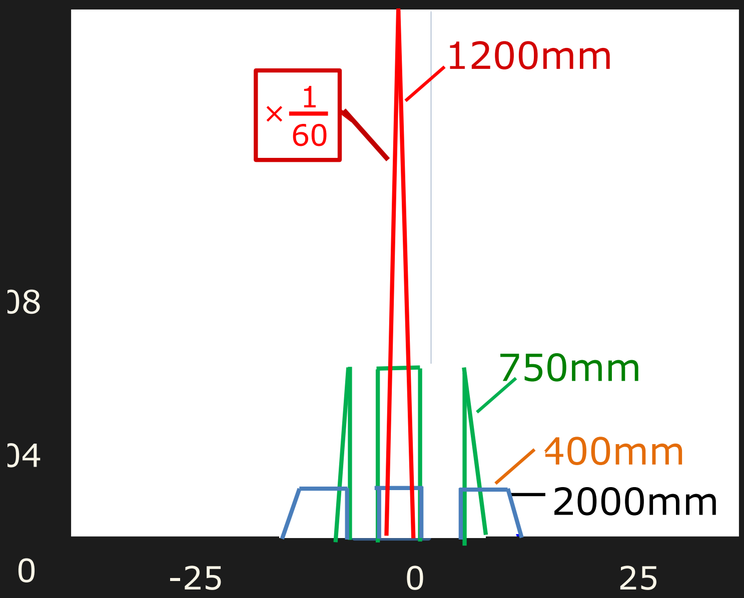
Coded aperture systems

[Levin et al. '07]
[Zhou et al. '08]
[Veeraraghavan et al. '07]

Depth invariant PSF

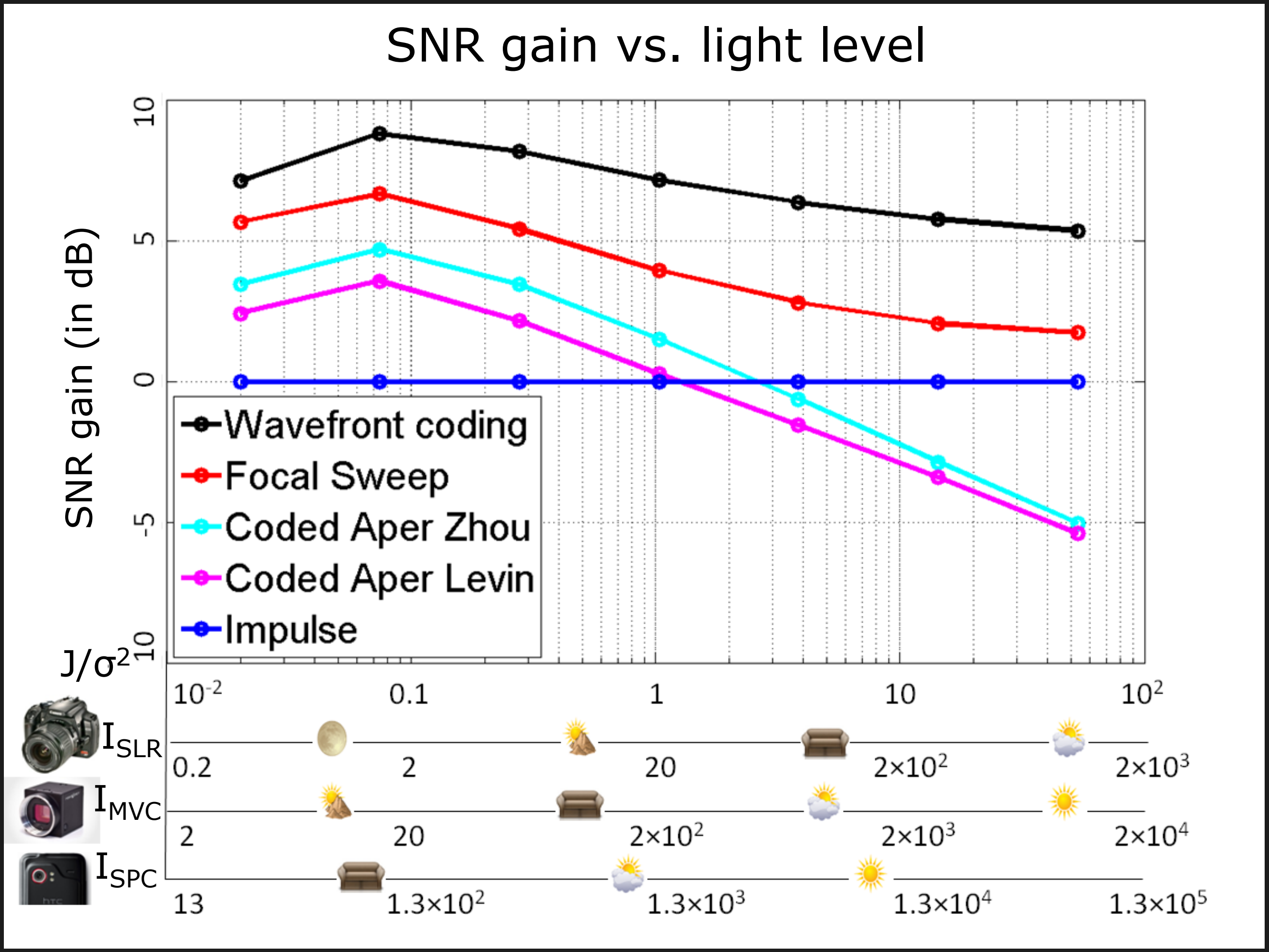


Coded Camera PSF



Analytic performance with Prior

Impulse camera: F/11
Other cameras: F/1



Good EDOF systems perform 9 dB better than impulse imaging

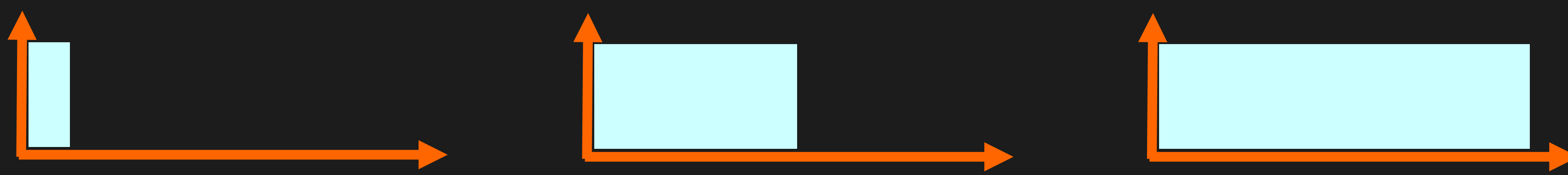
Analysis of motion deblurring systems

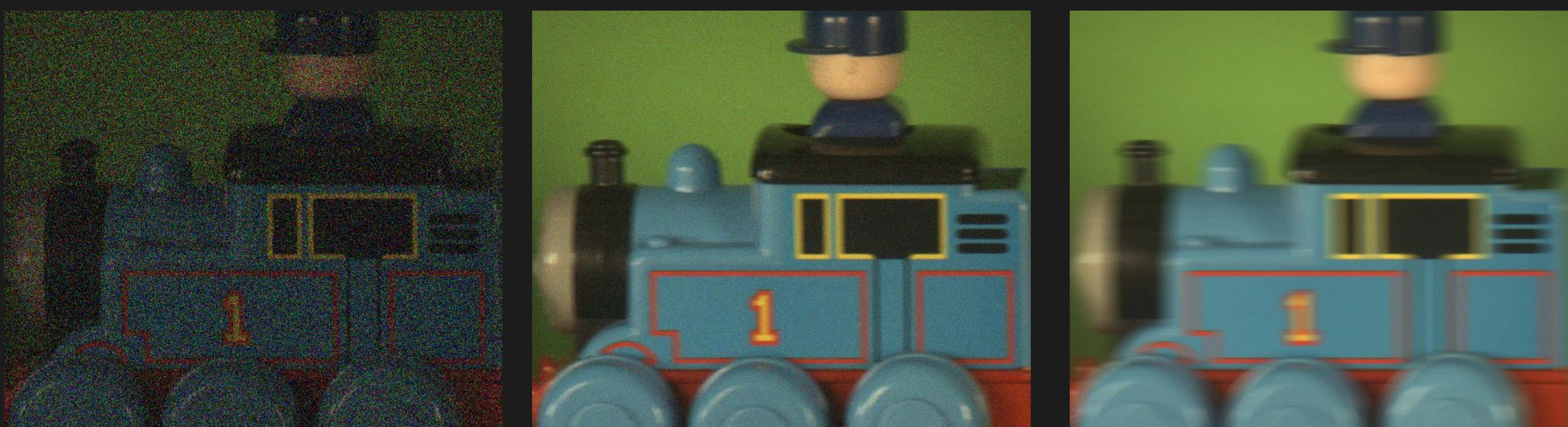
Light throughput vs. motion blur

Increasing exposure time

→

Noise decreases but motion blur increases

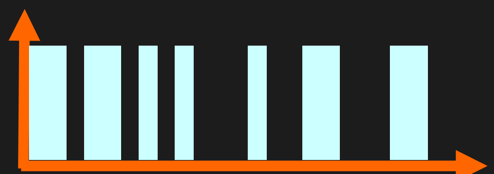




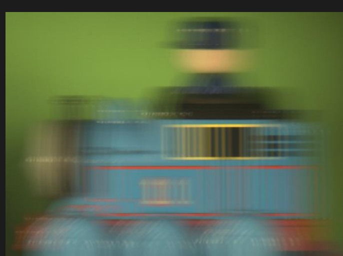
Motion deblurring CI systems

Coded exposure (Flutter shutter)

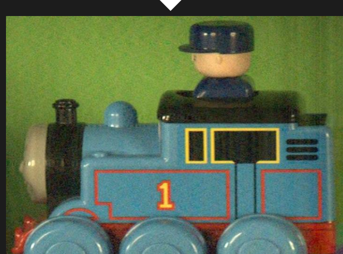
[Raskar ’06]



Increased light throughput and inversion better conditioned



Captured image



Deblurred image

Motion invariant photography

[Levin ‘08] [Cho ’10]

Captured image has same motion blur for different motions

Whole image deblurred using a single blur kernel

Simulation Performance under signal prior

$t_{CI} = 33 \times t_{impulse}$

Low light condition (10 lux)

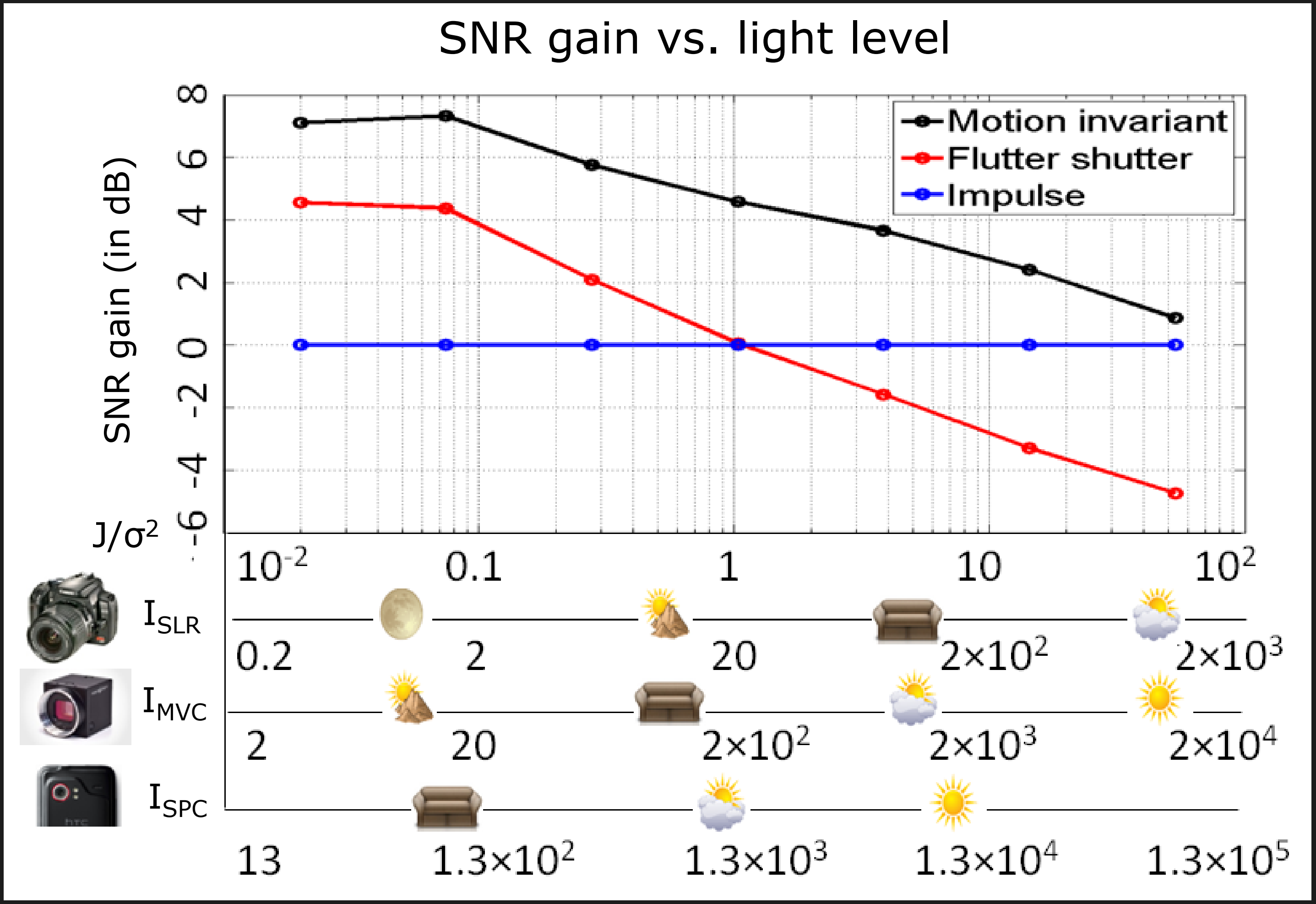
Motion invariant 7 dB better than impulse

High light condition (1000 lux)

Motion invariant 1.2 dB better than impulse

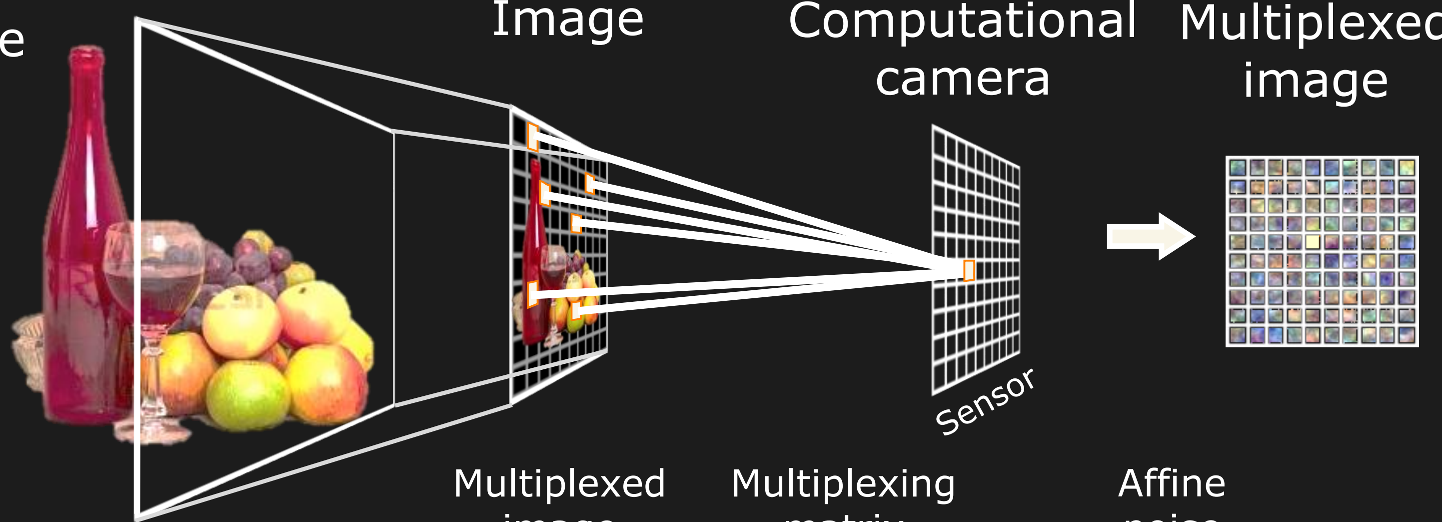
Analytic Performance under signal prior

$t_{CI}=33\times t_{impulse}$



Motion invariant camera achieves a peak SNR gain of 7.5 dB

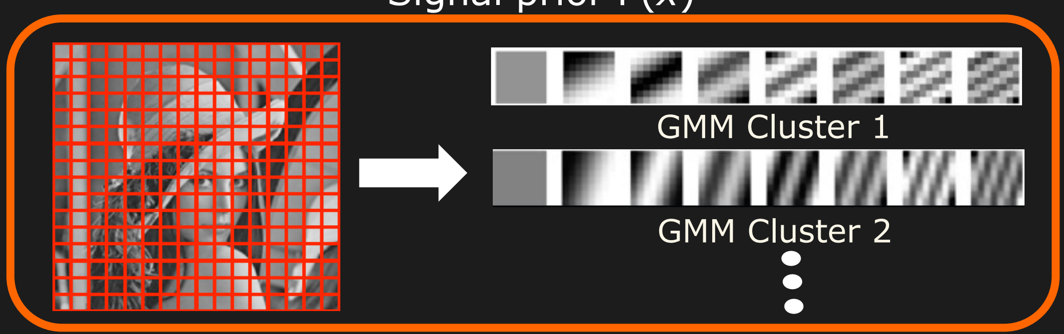
Conclusion:
Comprehensive analysis framework of CI



Multiplexed image | Multiplexing matrix | Affine noise

$$y = Hx + n$$

Signal prior P(x)

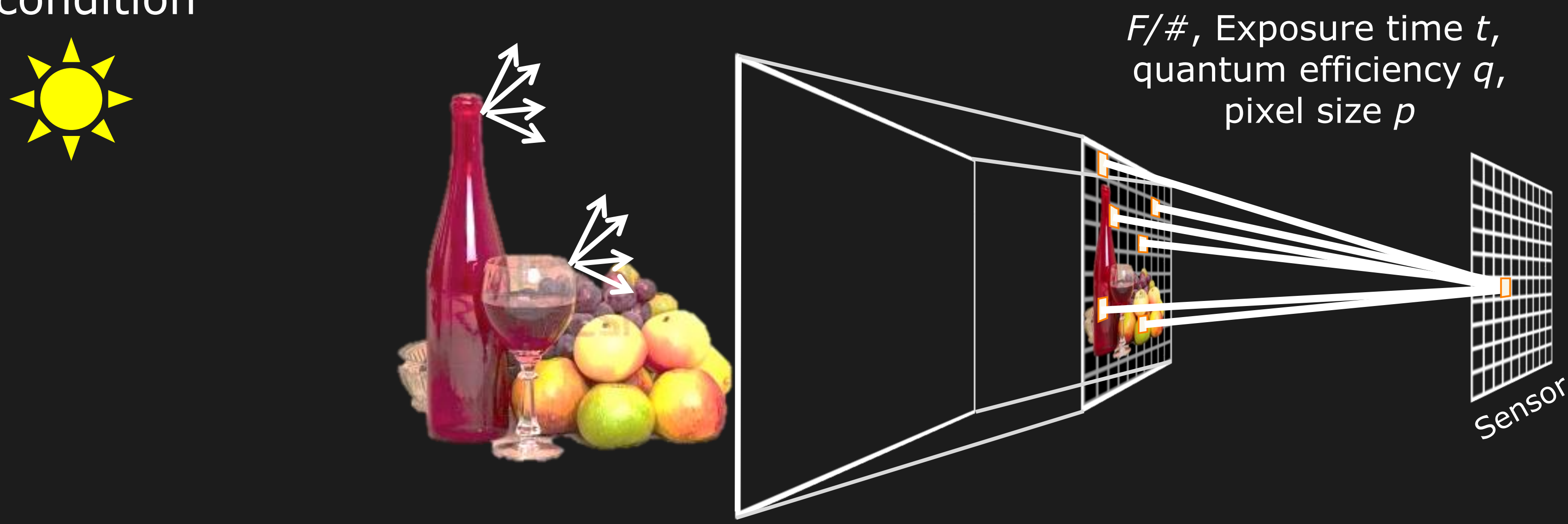


Our analysis accounts for

- Signal prior (GMM)
- Optical coding (H)
- Noise (affine)

Conclusion: Practical implications

1. Illumination condition
2. Scene reflectivity
3. Camera parameters



We analyzed EDOF and motion deblurring systems for typical values of:

- Illumination conditions
- Scene characteristics
- Camera parameters

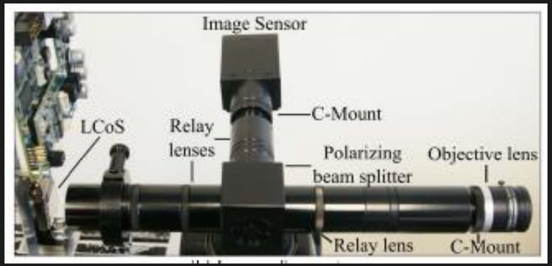
Conclusion: Our observations

- More gain due to prior than multiplexing
- Gain due to multiplexing modest when prior is taken into account
- CI systems provide significant advantage over impulse imaging under various illumination and camera parameters
- EDOF systems provides on average 7 dB gain over impulse imaging
- Motion deblurring systems provides on average 4.5 dB gain over impulse imaging

Future Work

Analyze compressive systems:

High speed video

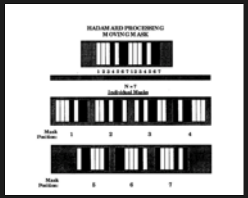


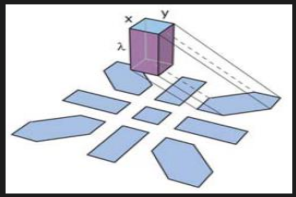
[Hitomi et al. 2011][Veera et al., 2011]

Light Field Capture

[Lanman '08] [Veeraraghavan '07]
[Liang '08]

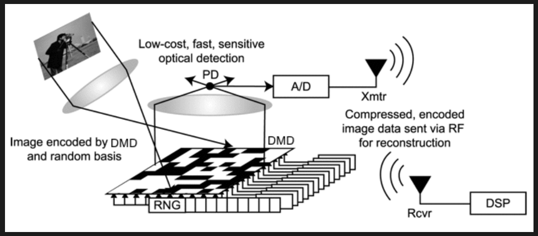
Multi/Hyper-Spectral





[Sloane '79] [Hanley '99]
[Baer '99]
[Wetzstein et al., '12]

Single pixel camera



[Wakin et al., 2006]

Design optimal CI systems

$$\arg\min_{H} \sum_{k=1}^{K} \alpha_k Tr(\tilde{\Sigma}_k) + \sum_{k=1}^{K} \alpha_k \int_y \| x_{mmse}(y) - \tilde{m}_k \|^2 P_k(y) dy$$