Bits From Photons: Oversampled Binary Image Acquisition

Feng Yang
Audiovisual Communications Laboratory
École Polytechnique Fédérale de Lausanne

Thesis supervisor: Prof. Martin Vetterli
Thesis co-supervisor: Dr. Luciano Sbaiz

February 20, 2012

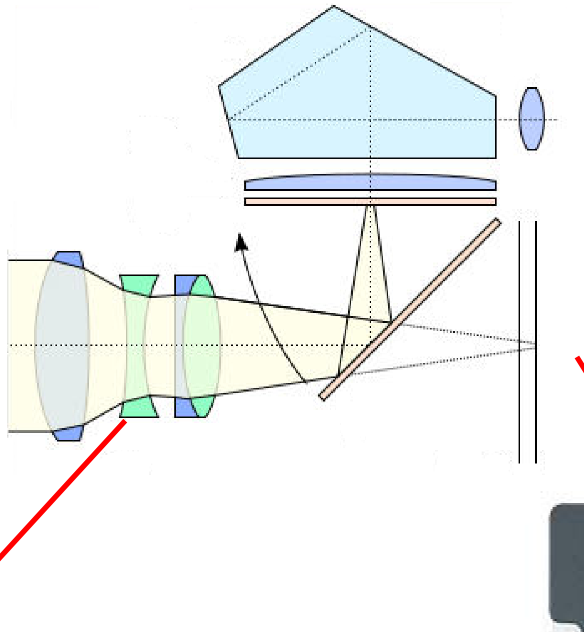
Outline

- Motivation
- Binary imaging
- Binary noisy imaging
- Threshold and optimal pattern design
- Generalized piecewise-constant model
- Conclusions and future research

Outline

- Motivation
- Binary imaging
- Binary noisy imaging
- Threshold and optimal pattern design
- Generalized piecewise-constant model
- Conclusions and future research

Traditional camera

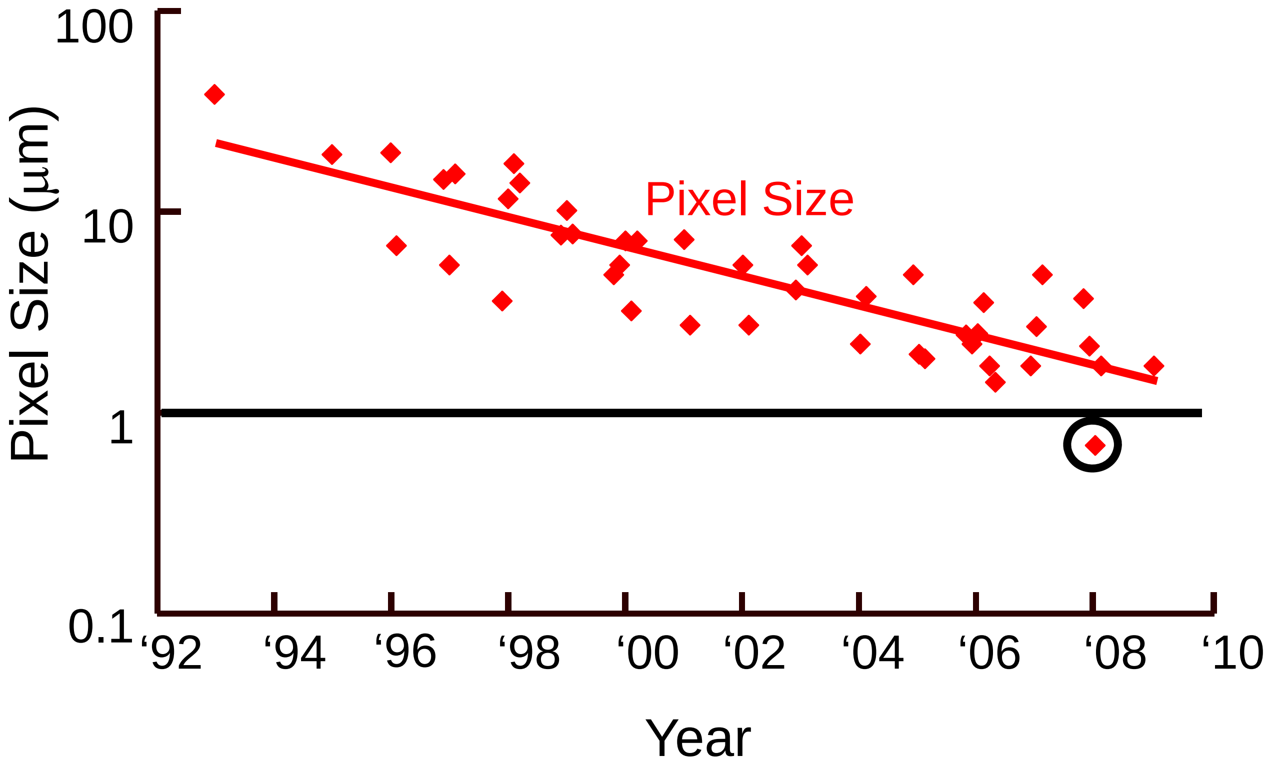


Optical Lens

CMOS sensor

Resolution

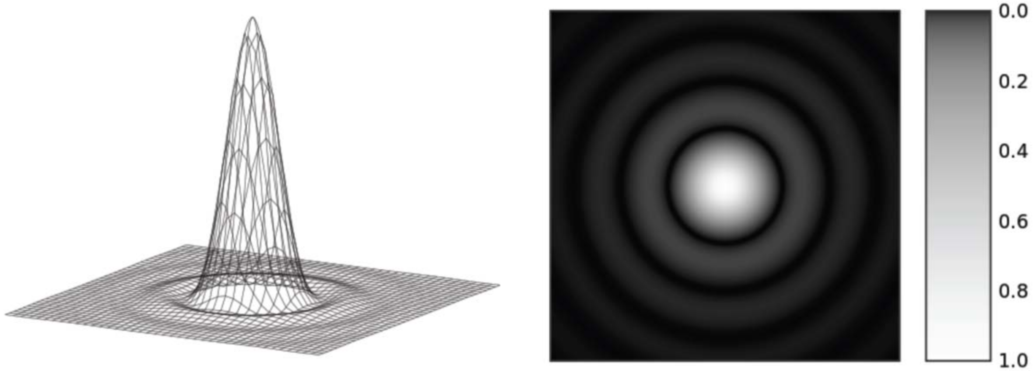
The evolution of pixel size in CMOS sensor



courtesy of Theuwissen [Theuwissen'08]

Do we need small pixels (1/2)

Shape of Airy disc on the sensor plane



Rayleigh Criterion: the minimum spatial resolution

$$\Delta l = 1.22 \lambda F\#$$

Example: $\lambda = 500nm,\ F\# = 8 \Rightarrow \Delta l \approx 5\mu m$

Do we need small pixels (2/2)

- Image sensor design
  - ~~High resolution~~
  - High sensitivity
  - Low noise
  - High dynamic range
- Small pixel
  - Low full well capacity
  - Low SNR
  - Low dynamic range
  - Oversampling device

- Similar to
  - Film (one-bit pixel)
  - Sigma-delta modulator (one-bit quantizer)

Oversampled binary image sensor

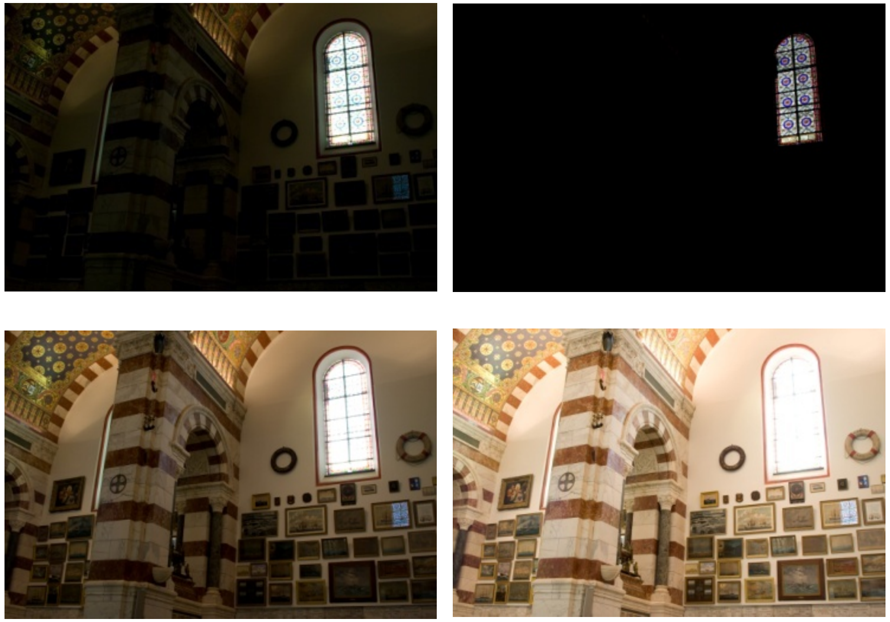
digital film [Fossum'05], gigavision camera [Sbaiz'09]

Pixel Size: > 1.25μm
Multi-photon response

Pixel Size: < 200nm
Binary response

High dynamic range imaging

Binary sensors with large oversampling factors achieve higher dynamic range than conventional sensors



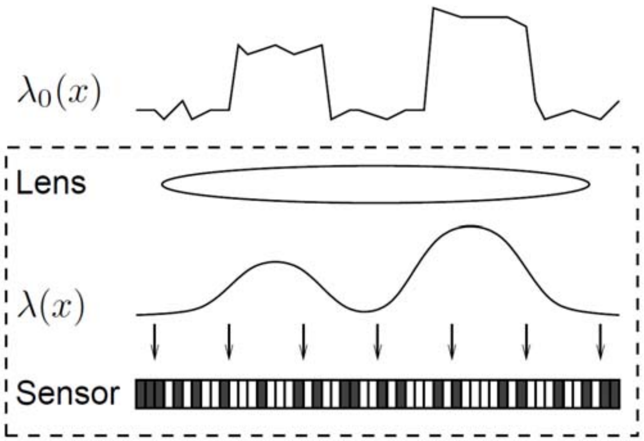
Conventional sensor: multiple exposures

Binary sensor: single exposure

Outline

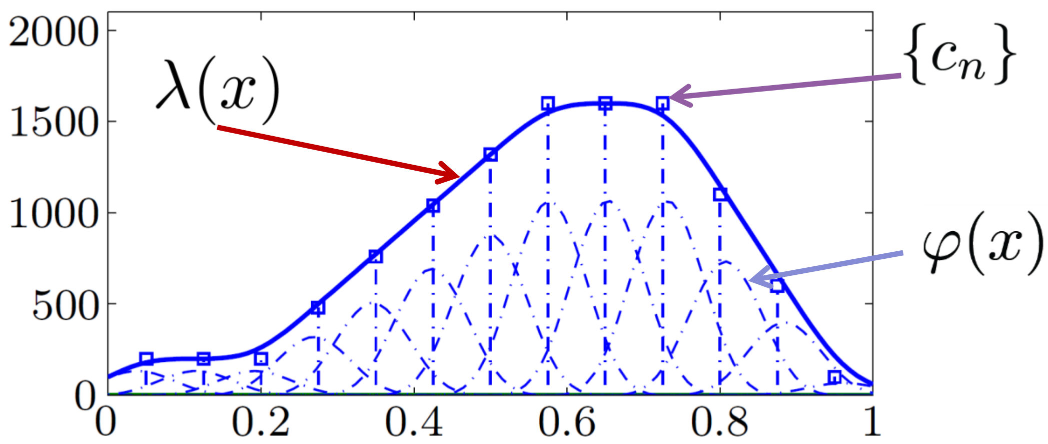
- Motivation
- Binary imaging
- Binary noisy imaging
- Threshold and optimal pattern design
- Generalized piecewise-constant model
- Conclusions and future research

Imaging model



sampled by *M* binary pixels

Light intensity field



The diffraction-limited light intensity field is modeled as

$$\lambda(x) = \frac{N}{\tau} \sum_{n=0}^{N-1} c_n \varphi(Nx - n)$$

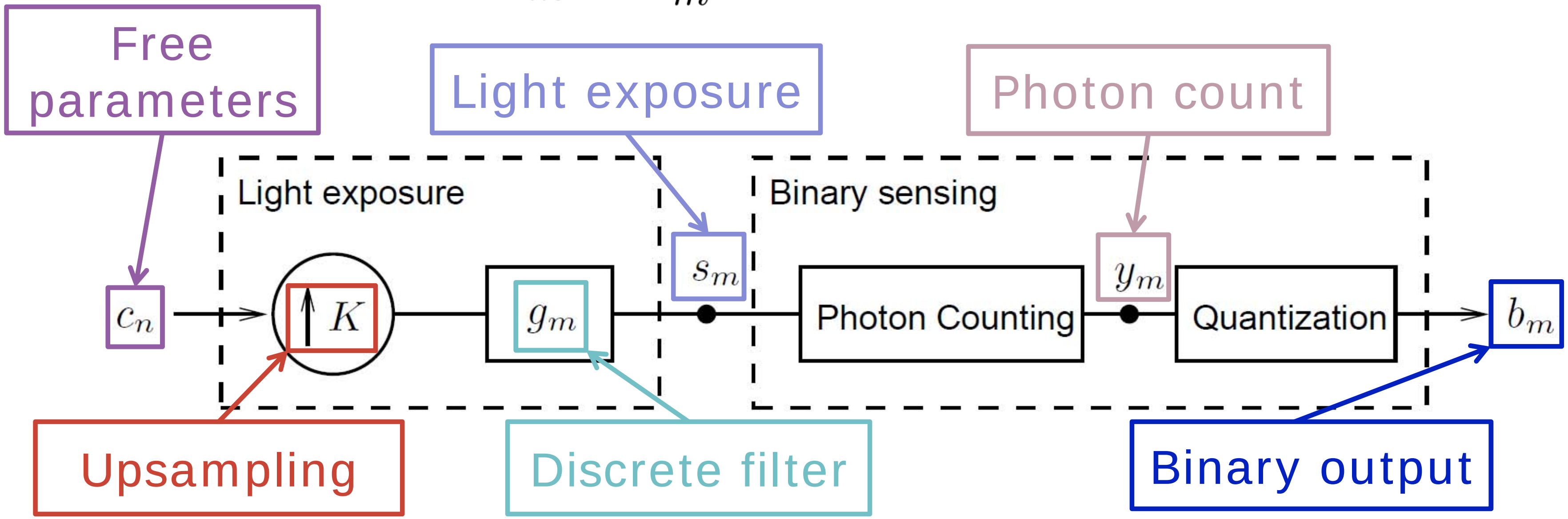
Free parameters (c_n)

Degrees of freedom (N)

Interpolation kernel (φ)

Signal processing model

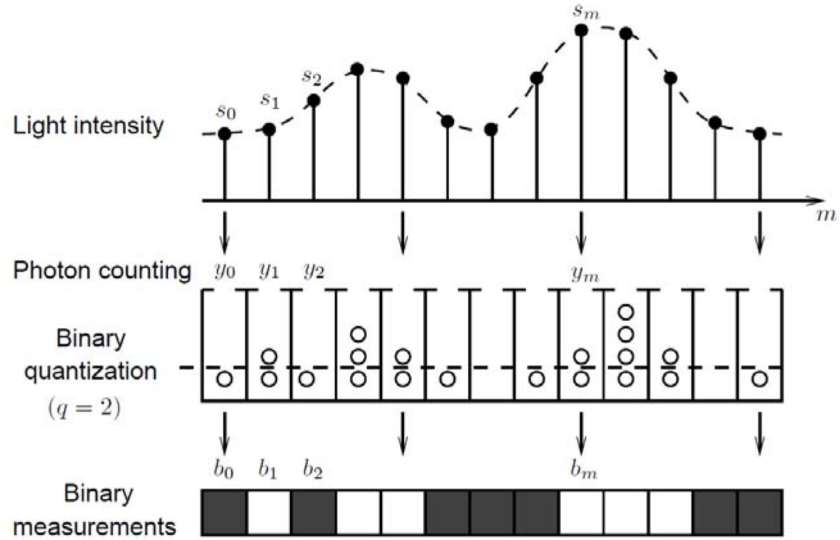
$$s_m = \boldsymbol{e}_m^T \boldsymbol{G} \boldsymbol{c}$$

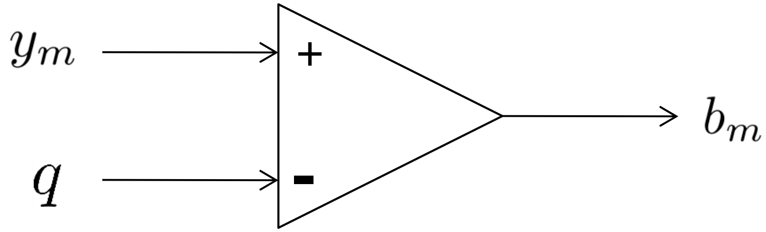


Spatial oversampling factor $K \stackrel{def}{=} \frac{M}{N}$

M: the number of binary pixels
N: the degrees of freedom of the light intensity field

Mathematical model of binary pixel





Photon count y_m: Poisson distribution

$$\mathbb{P}(Y_m = y_m; s_m) = \frac{s_m^{y_m}\, e^{-s_m}}{y_m!}, \qquad \text{for } y_m \in \mathbb{Z}^+ \cup \{0\}$$

Binary output b_m : Bernoulli distribution

$$\mathbb{P}(b_m = 0; s_m) = \sum_{k=0}^{q-1} \frac{s_m^k}{k!} e^{-s_m}$$

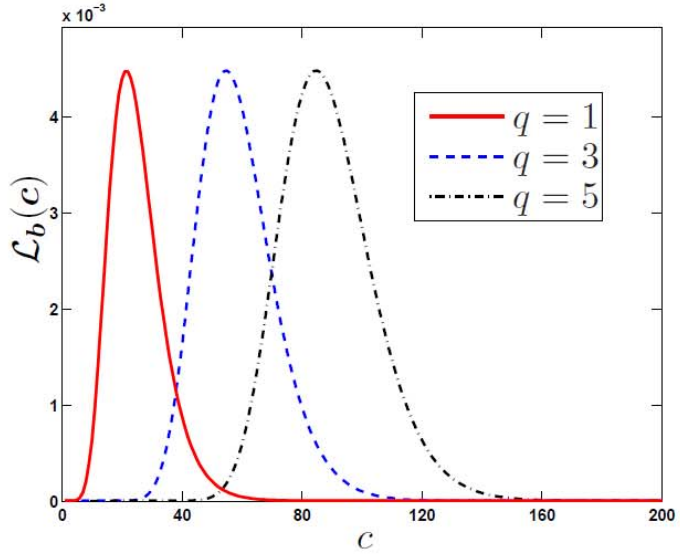
Reconstruction using maximum likelihood estimator

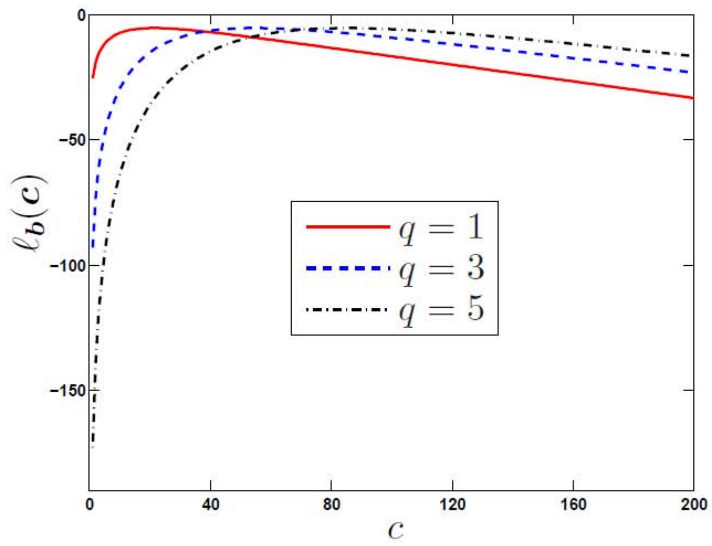
Likelihood function: $\mathcal{L}_{\boldsymbol{b}}(\boldsymbol{c}) \stackrel{\text{def}}{=} \prod_{m=0}^{M-1} \mathbb{P}(B_m = b_m; s_m) = \prod_{m=0}^{M-1} \mathbb{P}(B_m = b_m; \boldsymbol{e}_m^T \boldsymbol{G}\boldsymbol{c})$

Log-likelihood function: $\ell_{\boldsymbol{b}}(\boldsymbol{c}) \stackrel{\text{def}}{=} \log \mathcal{L}_{\boldsymbol{b}}(\boldsymbol{c})$

Maximum likelihood estimator: $\widehat{\boldsymbol{c}}_{\text{ML}}(\boldsymbol{b}) \stackrel{\text{def}}{=} \underset{\boldsymbol{c} \in [0,S]^N}{\arg\max}\ \mathcal{L}_{\boldsymbol{b}}(\boldsymbol{c}) = \underset{\boldsymbol{c} \in [0,S]^N}{\arg\max}\ \ell_{\boldsymbol{b}}(\boldsymbol{c})$

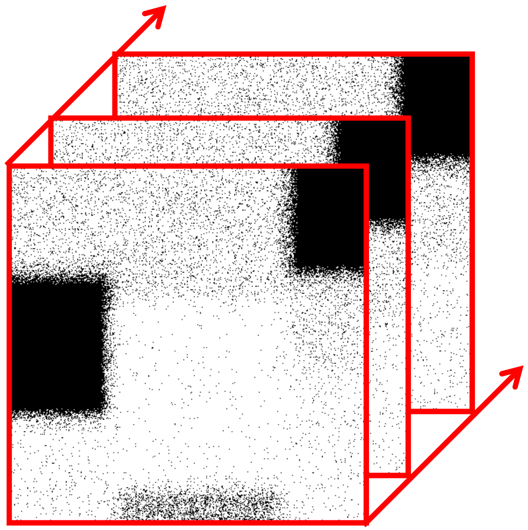
Theorem: The log-likelihood function is **concave**.





Extension: multiple exposures

- Temporal oversampling
- Equivalent to spatial oversampling, using box functions
- Log-likelihood function is concave



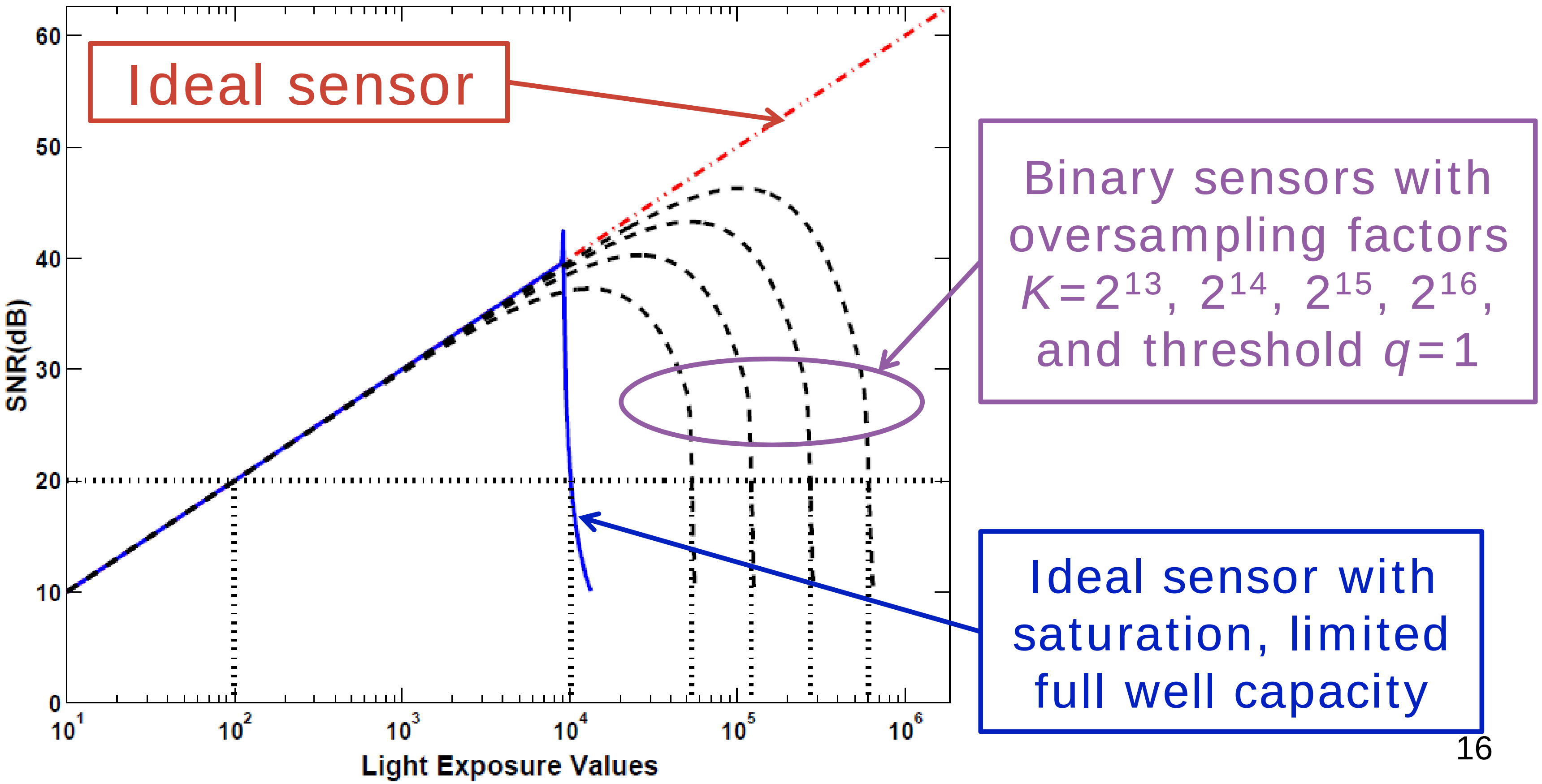
Exposure time for each frame τ/J

Total exposure time is τ

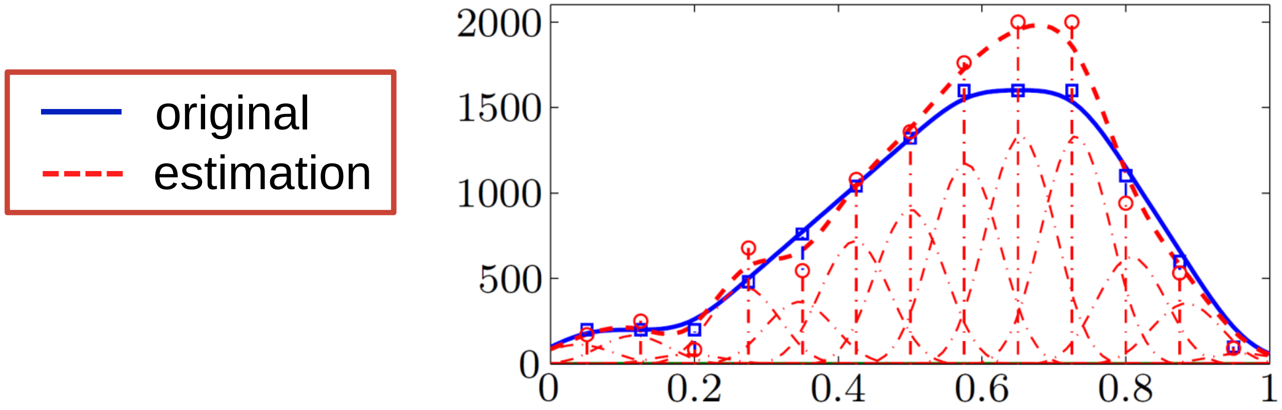
J binary images

Comparison with a conventional sensor

$$\mathrm{SNR} = 10\log_{10}\frac{c^2}{\mathbb{E}[(\widehat{c}-c)^2]}$$

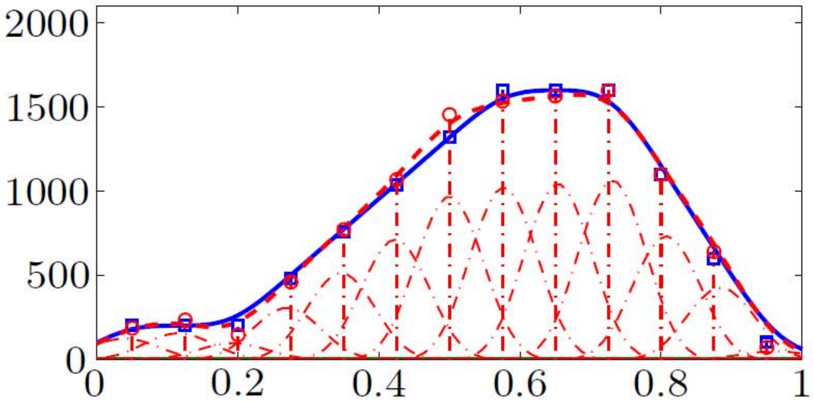


Numerical results: 1-D signals

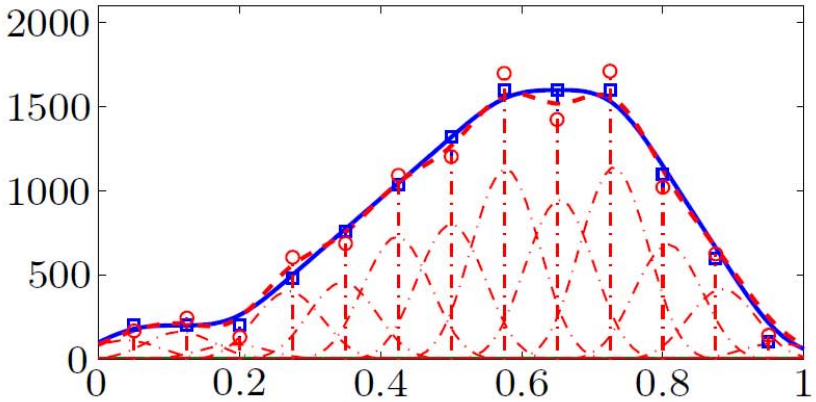


Threshold q=1

Spatial oversampling K=256

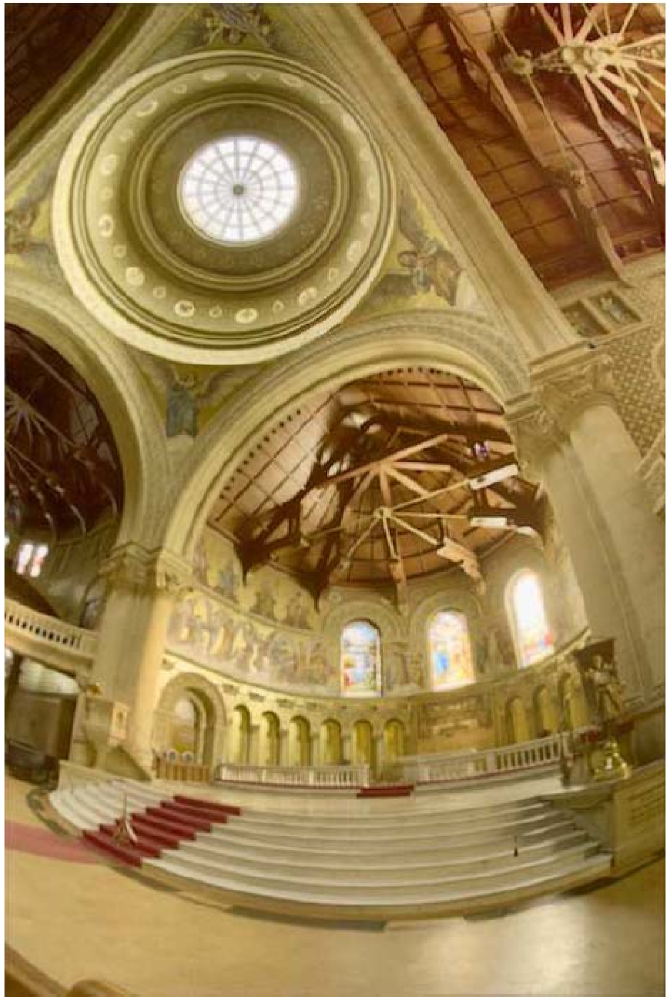


Spatial oversampling K=2048



Spatial oversampling K=256,
Temporal oversampling J=8

Numerical results: 2-D images



Threshold: q=1

Spatial oversampling: 32×32

Temporal oversampling: 256

Experimental results: real sensor

Single-photon avalanche diode (SPAD) camera

Resolution: 32×32
Pixel value: binary
Sensitivity: single photon

Experimental results: real sensor

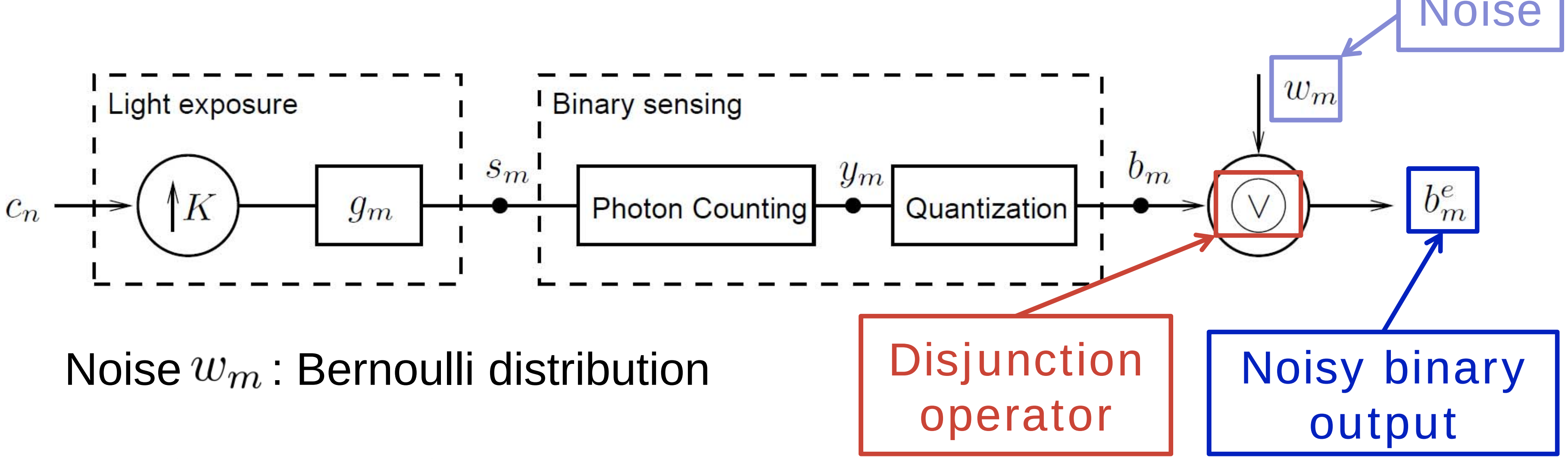
Resolution: 32×32, total images: 4096

Outline

- Motivation
- Binary imaging
- Binary noisy imaging
- Threshold and optimal pattern design
- Generalized piecewise-constant model
- Conclusions and future research

Signal processing model: noisy case

Noise source: dark current noise, threshold noise, etc.



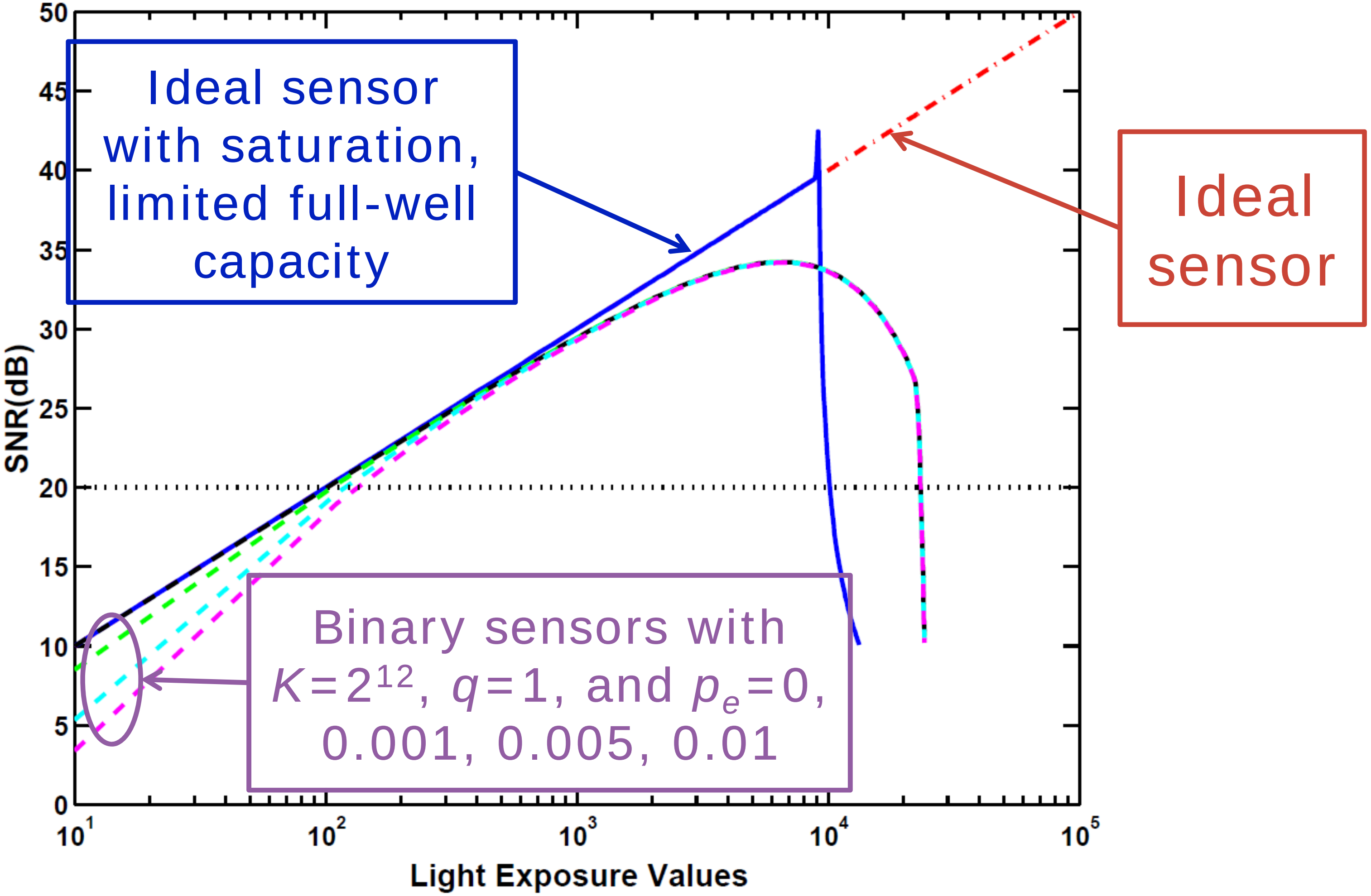
Noise w_m : Bernoulli distribution

$$\mathbb{P}(w_m = 1; p_e) = p_e \quad \text{(noise rate)}$$

Noisy binary output: $b_m^e \stackrel{\text{def}}{=} b_m \vee w_m$, Bernoulli distribution

$$\mathbb{P}(b_m^e = 0; s_m, p_e) = (1 - p_e) \sum_{k=0}^{q-1} \frac{s_m^k}{k!} e^{-s_m}$$

Influence of noise



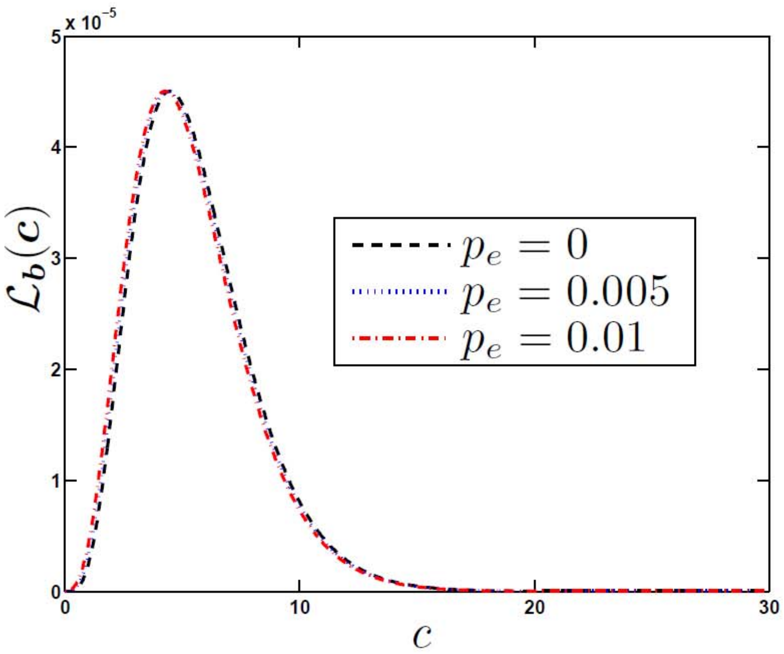
Robust to noise for large light exposure values

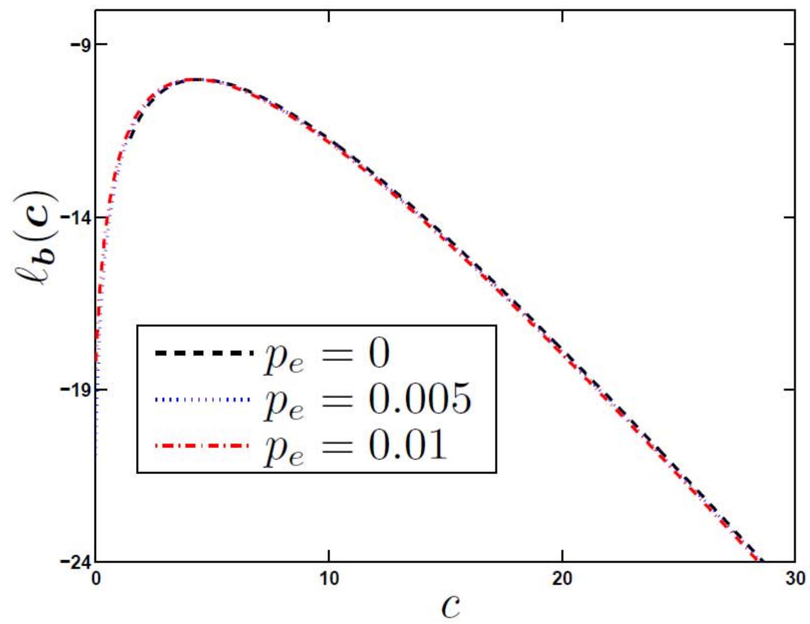
Reconstruction using MLE (1/3)

Maximum likelihood estimator

$$\widehat{\boldsymbol{c}}_{\mathrm{ML}}(\boldsymbol{b}^e) \stackrel{\mathrm{def}}{=} \underset{\boldsymbol{c}\in[0,S]^N}{\arg\max}\, \mathcal{L}^e_{\boldsymbol{b}}(\boldsymbol{c}) = \underset{\boldsymbol{c}\in[0,S]^N}{\arg\max}\, \ell^e_{\boldsymbol{b}}(\boldsymbol{c})$$

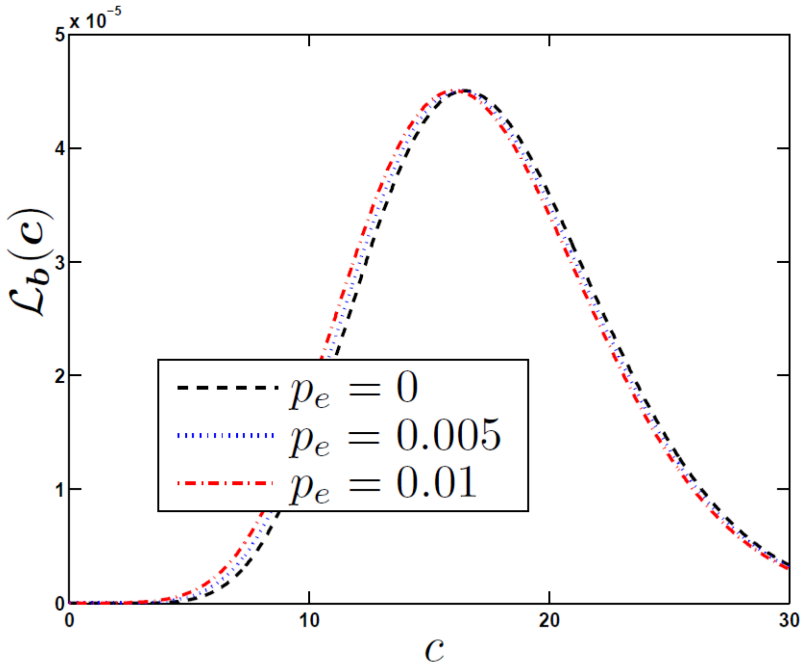
Theorem: constant light intensity field, and threshold q=1, the log-likelihood function is concave.

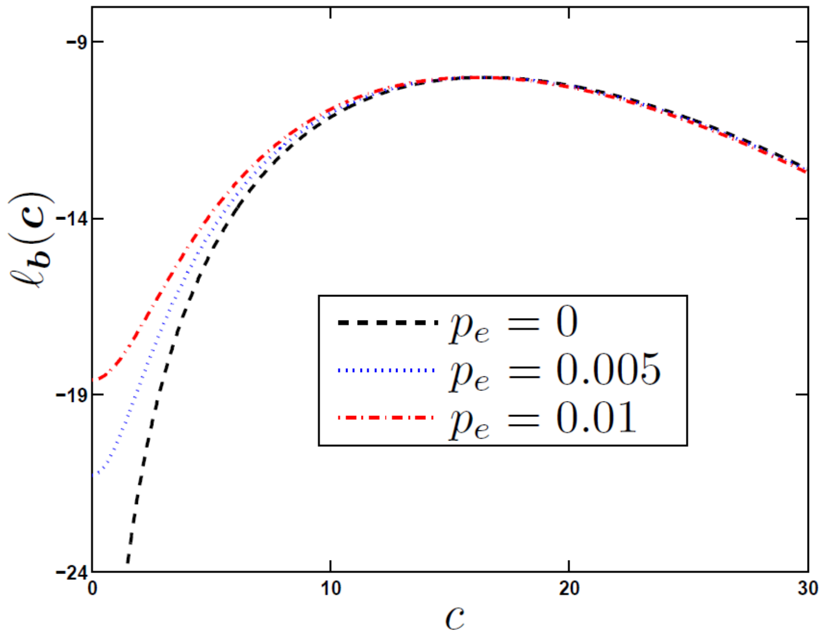




Reconstruction using MLE (2/3)

Theorem: constant light intensity field, arbitrary *q*, both the likelihood function and log-likelihood function are strictly pseudoconcave.

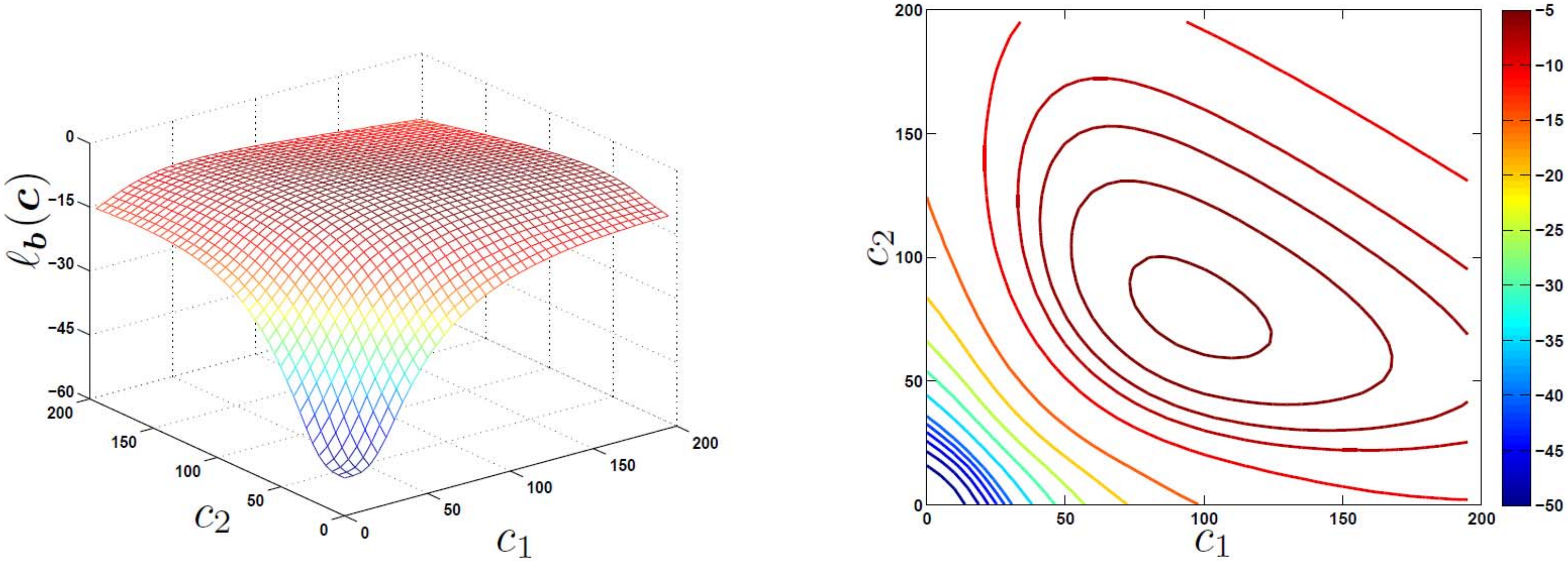




Piecewise-constant model, optimal solution can be achieved.

Reconstruction using MLE (3/3)

Log-likelihood function for general linear model

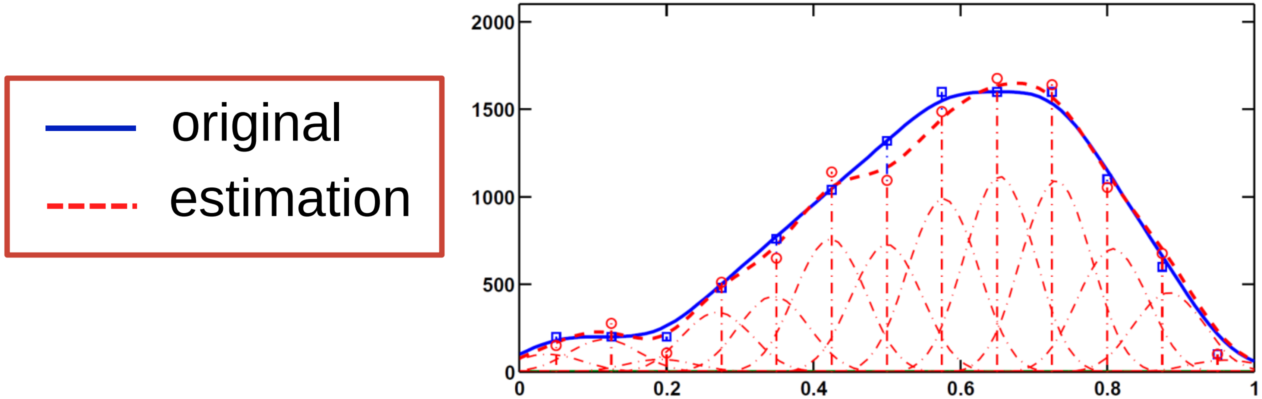


Not even quasi-concave, no guarantee for the optimal solution

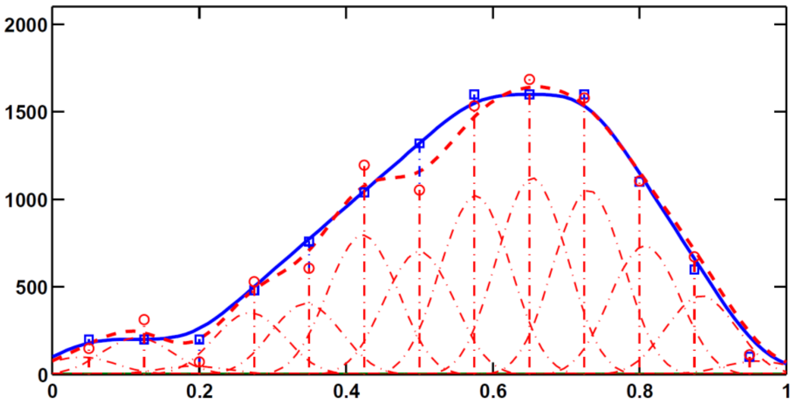
Reconstruction algorithm

1. Initial estimation, using piecewise-constant assumption
2. Refined estimation, using the iterative algorithm, i.e., Newton's method

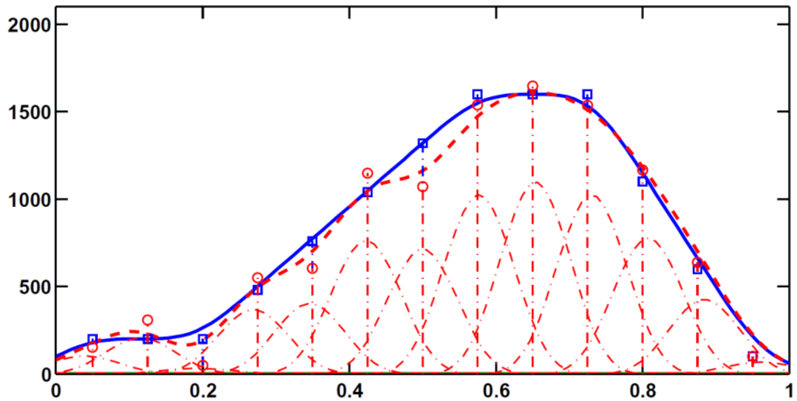
Numerical results: 1-D signals



Noise rate $p_e = 0$



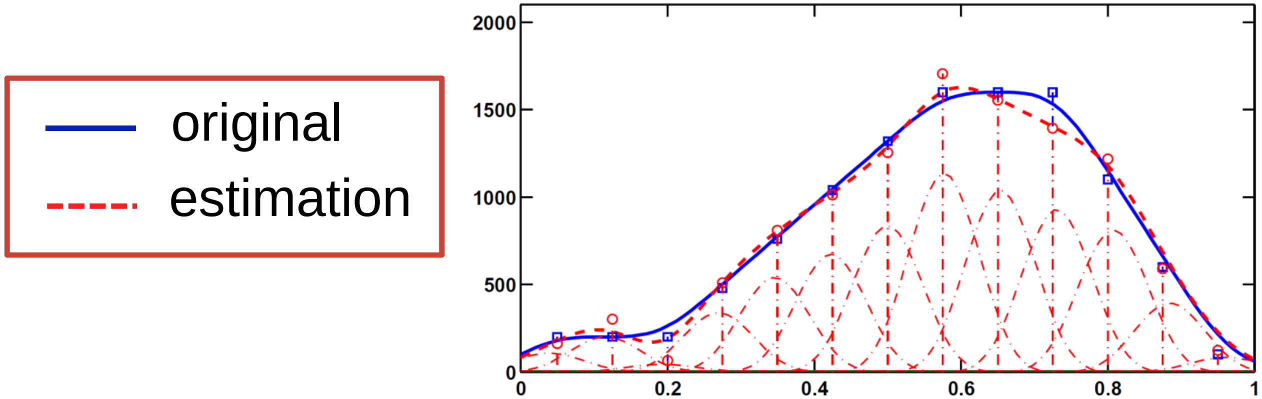
Noise rate $p_e = 0.1$



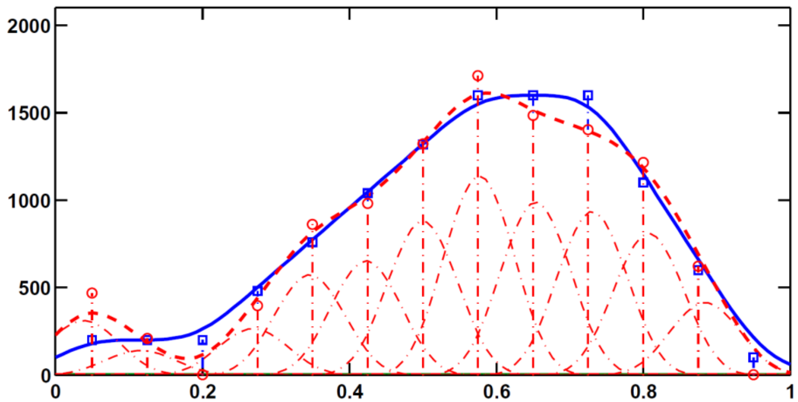
Noise rate $p_e = 0.2$

Spatial oversampling $K = 1024$, threshold $q = 1$

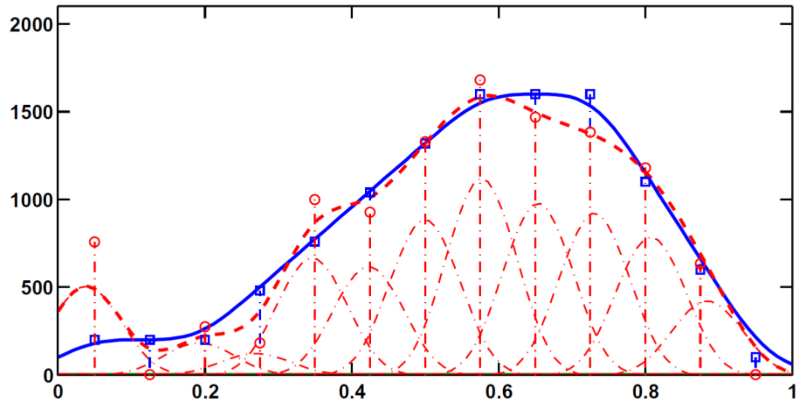
Numerical results: 1-D signals



Noise rate $p_e = 0$



Noise rate $p_e = 0.1$



Noise rate $p_e = 0.2$

Spatial oversampling $K = 512$, threshold $q = 3$

Numerical results: 2-D images

Threshold: *q*=1, spatial oversampling: 8×8, temporal oversampling: 128

Original

Noise rate $p_e = 0$

Noise rate $p_e = 0.1$

Noise rate $p_e = 0.2$

Numerical results: 2-D images

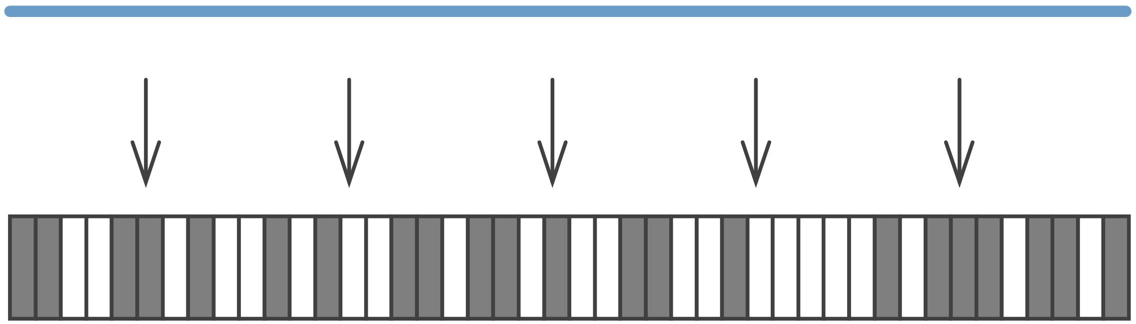
Noise rate $p_e = 0.2$

Outline

- Motivation
- Binary imaging
- Binary noisy imaging
- Threshold and optimal pattern design
- Generalized piecewise-constant model
- Conclusions and future research

The Cramér-Rao lower bound (CRLB)

Using *K* pixels to estimate a constant light exposure value *c*



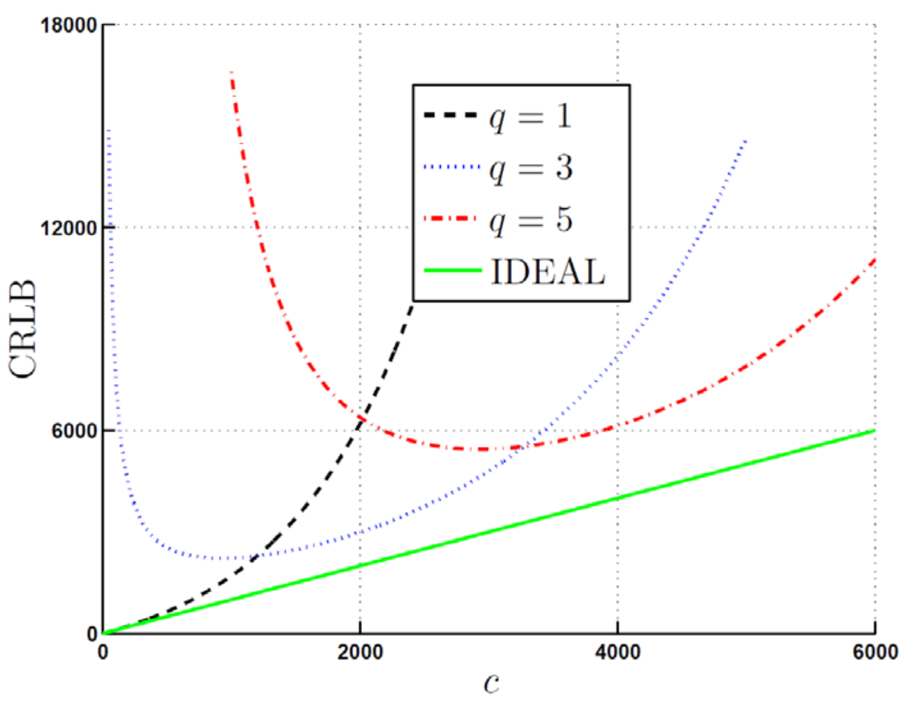
Ideal sensor: $\mathrm{CRLB}_{\mathrm{ideal}} = c$

Binary sensor:

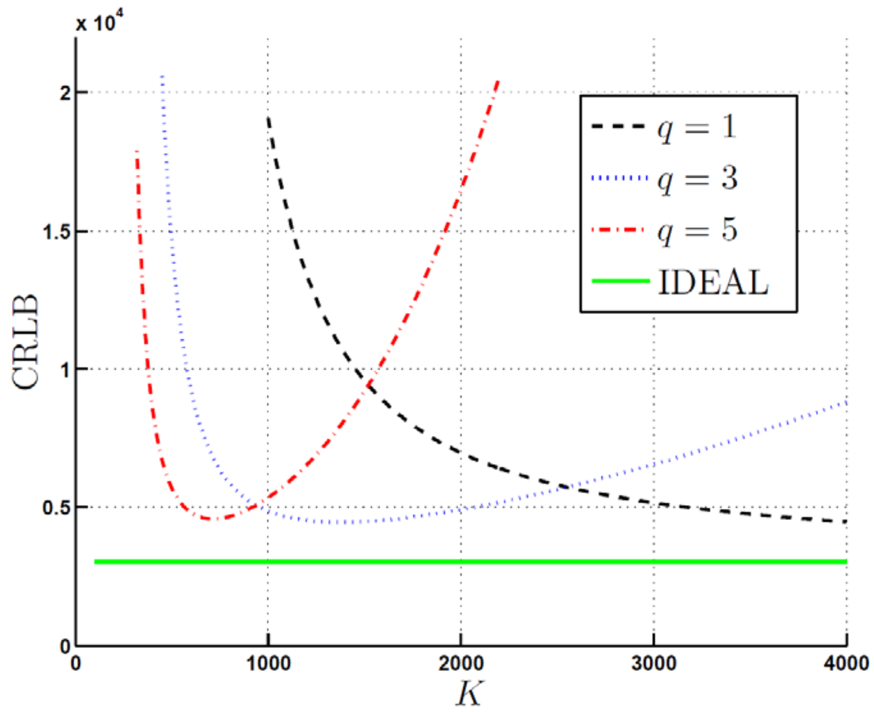
$$\mathrm{CRLB}_{\mathrm{bin}}(K, q) = c \left(\sum_{j=0}^{q-1} \frac{(q-1)!(c/K)^{-j}}{(q-1-j)!} \right) \left(\sum_{j=0}^{\infty} \frac{(q-1)!(c/K)^{j}}{(q+j)!} \right)$$

The Cramér-Rao lower bound (CRLB)

CRLB vs. light exposure value *c*



CRLB vs. oversampling factor *K*



Proposition:

For threshold *q*=1, $\lim_{K\to\infty} \mathrm{CRLB}_{\mathrm{bin}}(K,q) = \mathrm{CRLB}_{\mathrm{ideal}}$

For *q*>1, $\lim_{K\to\infty} \mathrm{CRLB}_{\mathrm{bin}}(K,q)/\mathrm{CRLB}_{\mathrm{ideal}} = \infty$

Optimal threshold pattern and reconstruction

2-D sensor with two interleaved thresholds

q_1	q_2	q_1	q_2
q_2	q_1	q_2	q_1
q_1	q_2	q_1	q_2
q_2	q_1	q_2	q_1

Optimal Criterion: Find the minimum average *CRLB*

$$(q_{1,\text{opt}}, q_{2,\text{opt}}) = \underset{1 \le q_1, q_2 \le q_{\max}}{\arg\min} \int_{c_{\min}}^{c_{\max}} \text{CRLB}_{\text{bin2}}(K/2, K/2, q_1, q_2, c)\, \mathrm{d}c$$

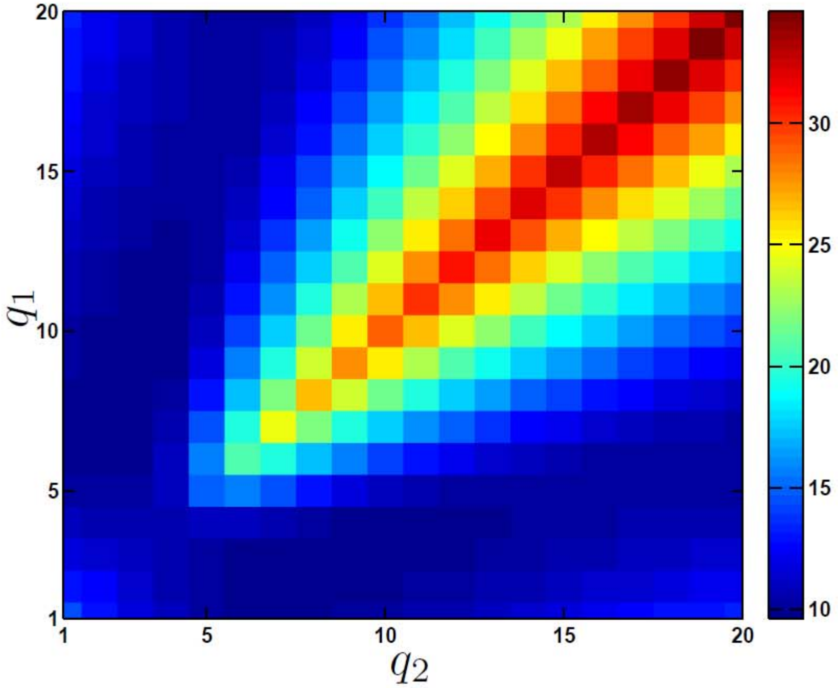
Maximum likelihood estimator

$$\widehat{\boldsymbol{c}}_{\text{ML}}(\boldsymbol{b}) \overset{\text{def}}{=} \underset{\boldsymbol{c} \in [0,S]^N}{\arg\max}\ \mathcal{L}_{\boldsymbol{b}}(\boldsymbol{c}) = \underset{\boldsymbol{c} \in [0,S]^N}{\arg\max}\ \ell_{\boldsymbol{b}}(\boldsymbol{c})$$

Proposition: The log-likelihood function is concave.

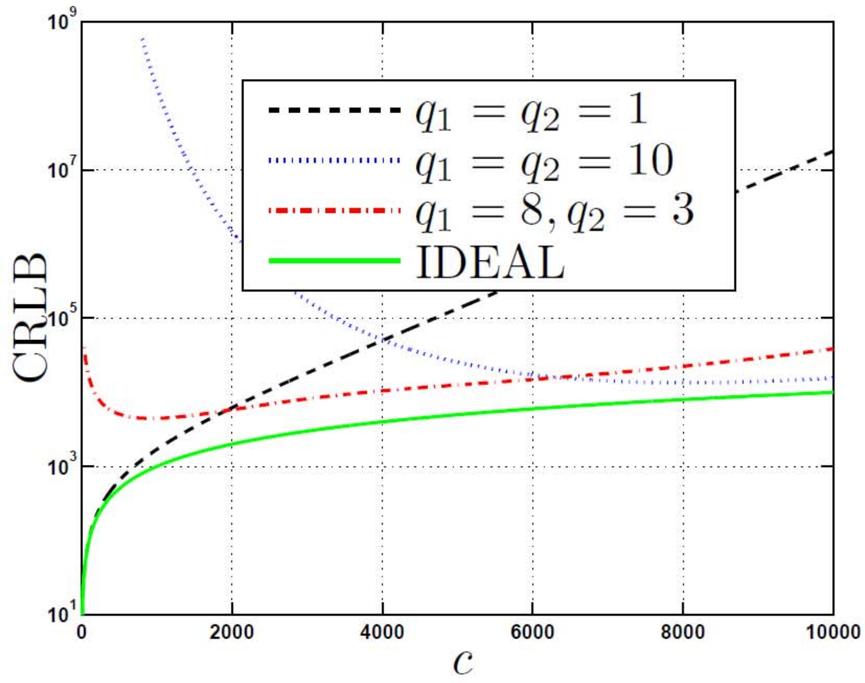
Design example

Optimal threshold pattern, when $K = 1024$, $c \in [10, 10^4]$



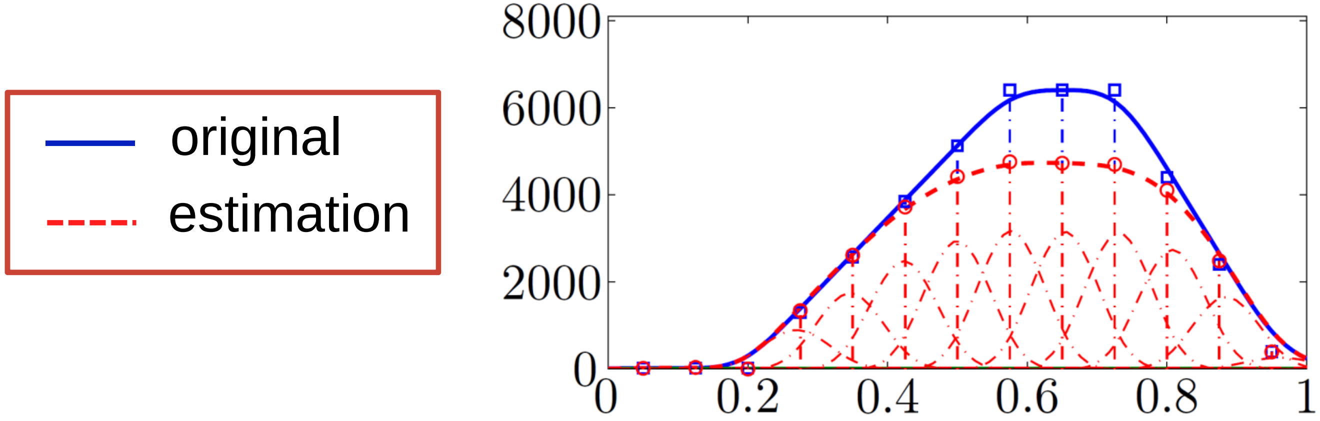
Average *CRLB* for different threshold patterns

Optimal pattern q_1=8, q_2=3

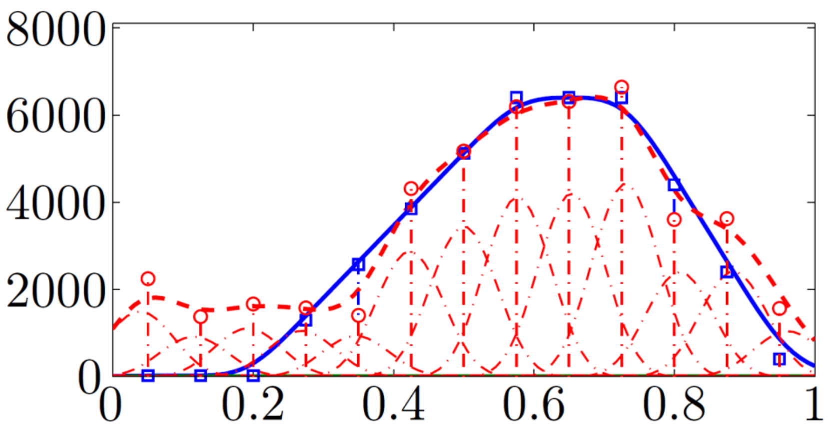


CRLB for different threshold patterns

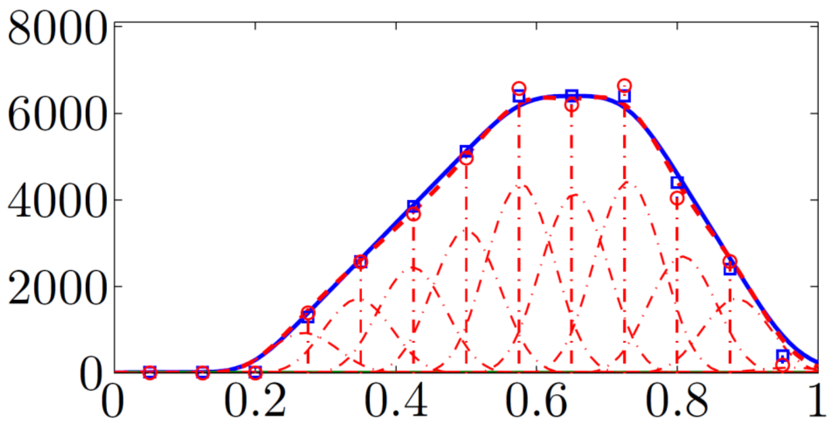
Numerical results: 1-D signals



Threshold pattern $q_1=q_2=1$



Threshold pattern $q_1=q_2=10$



Threshold pattern $q_1=8$, $q_2=3$

Numerical results – synthetic images

Threshold: q=1, spatial oversampling: 8×8, temporal oversampling: 16

Original

q_1=1
q_2=1

q_1=10
q_2=10

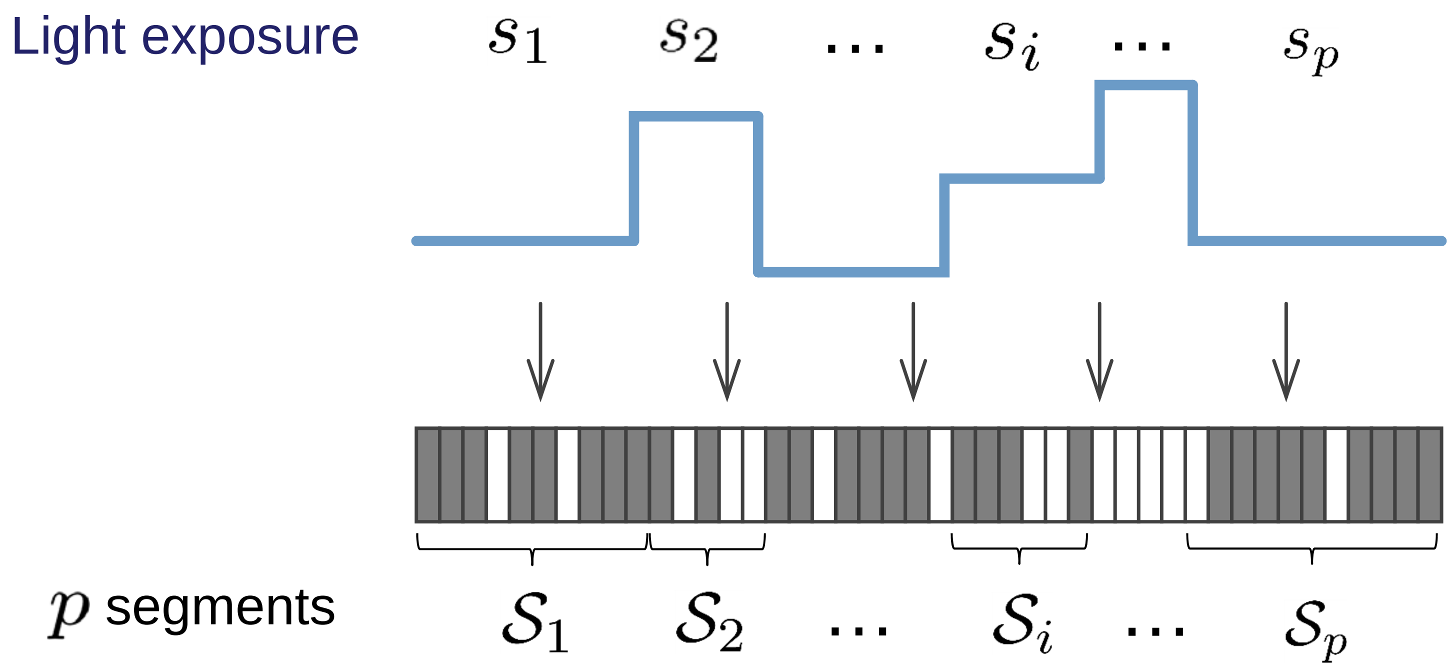
q_1=8
q_2=3

Outline

- Motivation
- Binary imaging
- Binary noisy imaging
- Threshold and optimal pattern design
- Generalized piecewise-constant model
- Conclusions and future research

Generalized piecewise-constant model

Estimate the light exposure values $\{s_m\}$ using *M* binary measurements



Estimate the light exposure values changed to reconstruct

$$\mathcal{P} \stackrel{\text{def}}{=} \{(\mathcal{S}_1, s_1), (\mathcal{S}_2, s_2), \ldots, (\mathcal{S}_i, s_i), \ldots, (\mathcal{S}_p, s_p)\}$$

Reconstruction using MLE (1/2)

Likelihood function

$$\mathcal{L}_{\boldsymbol{b}}(\mathcal{P}) = \prod_{i=0}^{p} \mathbb{P}(B_m = b_m, m \in \mathcal{S}_i; (\mathcal{S}_i, s_i))\, \theta^{p-1}(1-\theta)$$

Probability for segment i with light exposure value s_i

Probability when there are p segments

Log-likelihood function

$$\ell_{\boldsymbol{b}}(\mathcal{P}) \stackrel{\text{def}}{=} \log \mathcal{L}_{\boldsymbol{b}}(\mathcal{P})$$

Maximum likelihood estimator

$$\widehat{\mathcal{P}} \stackrel{\text{def}}{=} \arg\max_{\mathcal{P}, s_i \in [0,S]} \ell_{\boldsymbol{b}}(\mathcal{P})$$

$$= \arg\max_{\mathcal{P}, s_i \in [0,S]} \sum_{i=1}^{p} \mathbb{P}(B_m = b_m, m \in \mathcal{S}_i; (\mathcal{S}_i, s_i)) - \gamma p$$

Data term

Penalization term

Reconstruction using MLE (2/2)

Iteratively solving two problems

Segments are known

1. Estimate light exposure values

$$\widehat{s}_i \stackrel{\text{def}}{=} \arg\max_{s_i \in [0,S]} \mathbb{P}(B_m = b_m, m \in \mathcal{S}_i; (\mathcal{S}_i, s_i))$$

Solution: bisection method

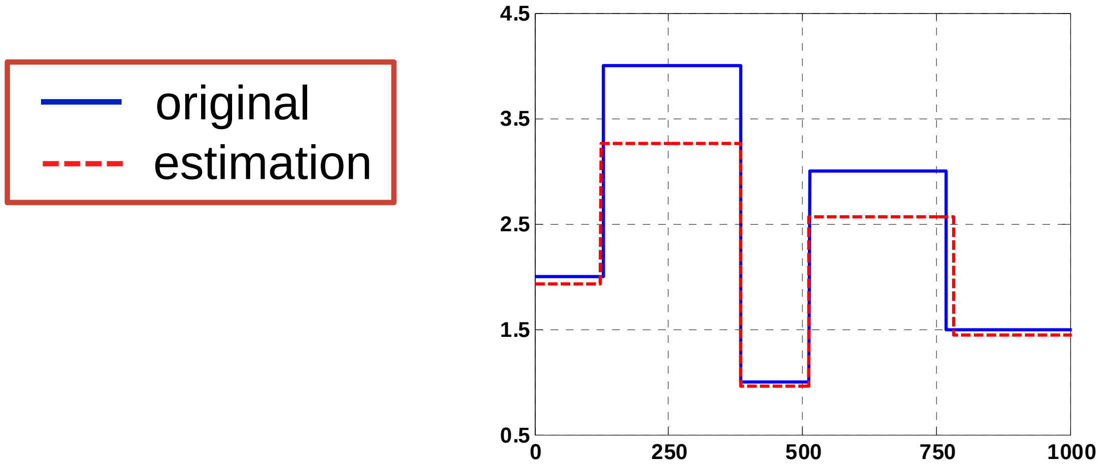
Light exposure values are known

2. Estimate segments

$$\widehat{\mathcal{G}} \stackrel{\text{def}}{=} \arg\max_{\mathcal{G}} \mathbb{P}(B_m = b_m, m \in \mathcal{S}_i; (\mathcal{S}_i, \widehat{s}_i)) - \gamma p$$

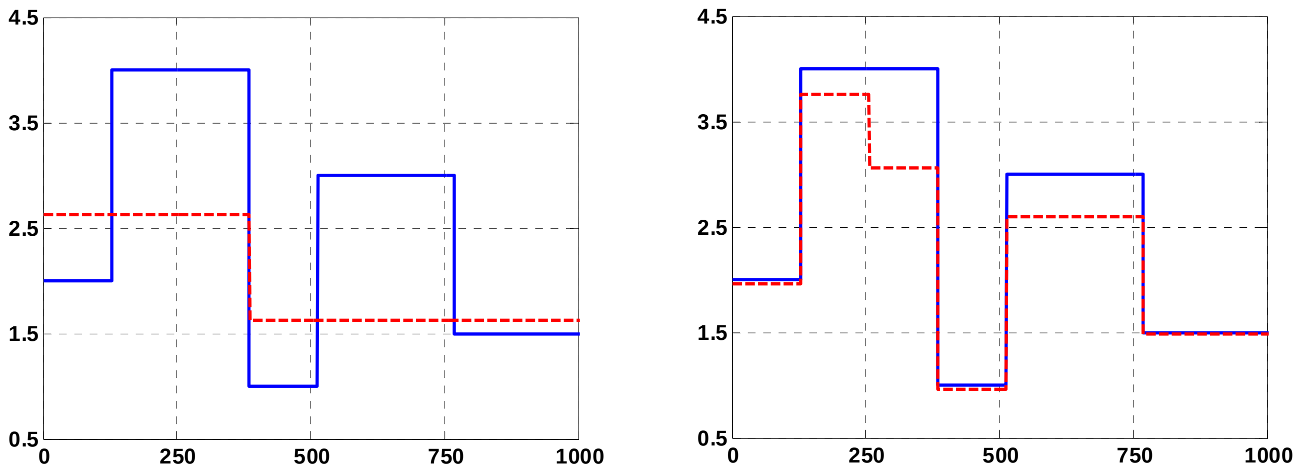
Solution: dynamic programming, greedy algorithm or pruning of binary trees (for 2-D case, quadtrees) .

Numerical results: 1-D signals



Dynamic programming

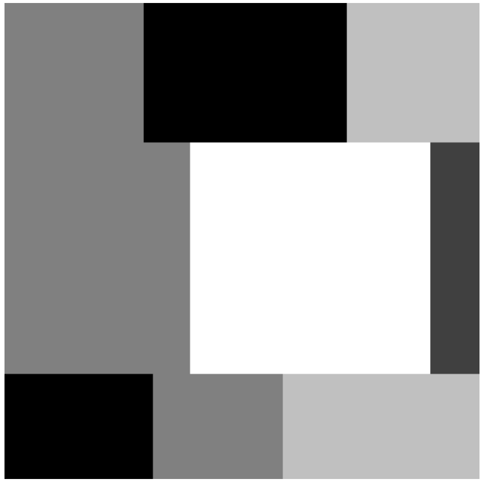
Greedy

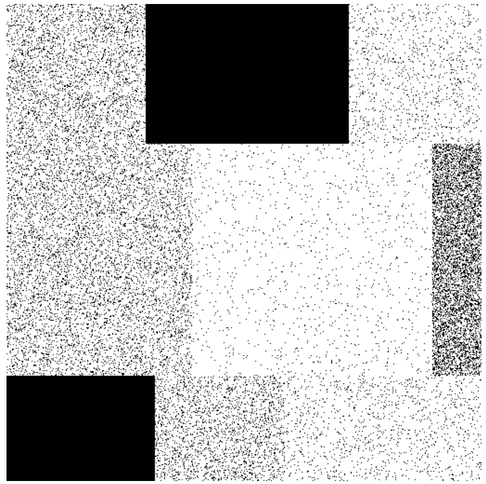


Pruning binary tree

Numerical results: synthetic images

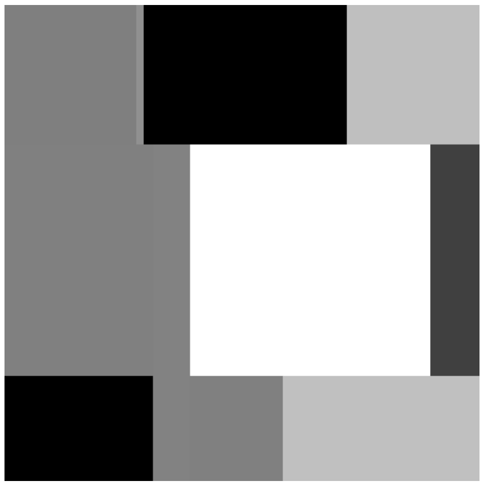
Original





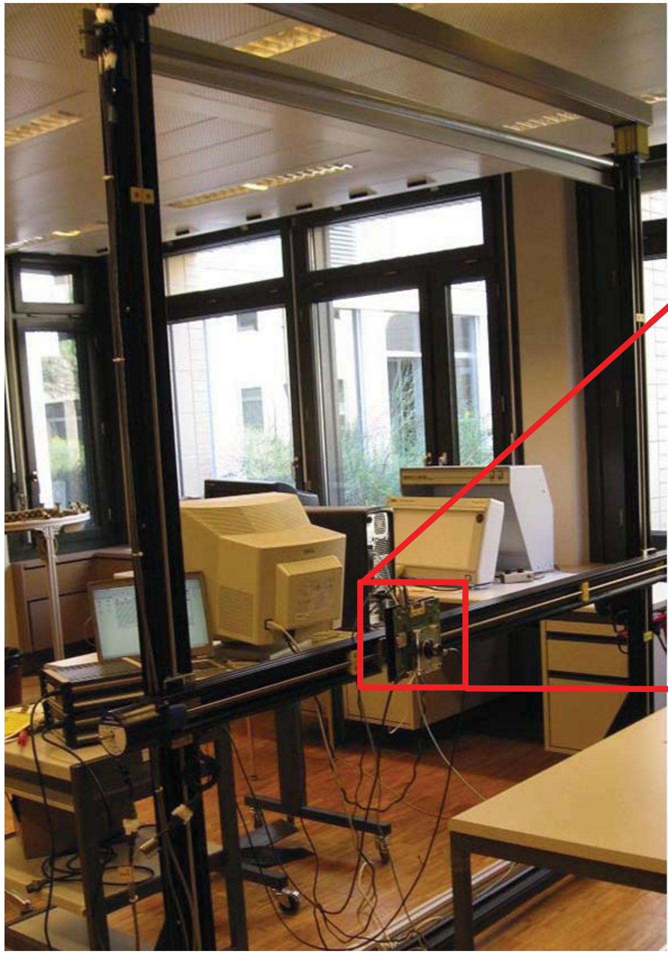
Binary image

Greedy



Pruning Quadtrees

Experimental results: real images (1/2)



SPAD camera
Resolution: 32×32

Experimental results: real images (2/2)

Binary image, 1024×1024

Greedy

128×128

Pruning Quadtrees

128×128

Outline

- Motivation
- Binary imaging
- Binary noisy imaging
- Threshold and optimal pattern design
- Generalized piecewise-constant model
- Conclusions and future research

Conclusions

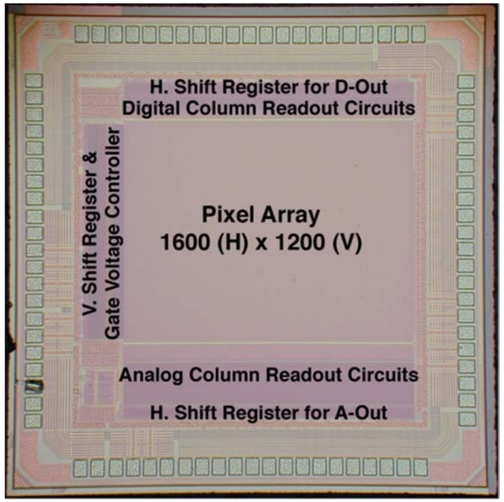
- Oversampled binary imaging
  - Diffraction limit and spatial oversampling
  - Binary pixel
  - Log-likelihood function is concave
- Noise performance
  - Robust to noise for large light intensity
  - Constant, log-likelihood function,
    - concave ($q=1$)
    - strictly pseudoconcave ($q>1$)

Conclusions

- Threshold and optimal pattern design
  - Asymptotic behavior
  - Large threshold strong light intensity, small threshold low light intensity
  - Optimal threshold pattern
  - Log-likelihood function is concave
- Generalized piecewise-constant model
  - Maximum likelihood estimator
  - Iteratively solving two problems

Future Research

- Sensor design



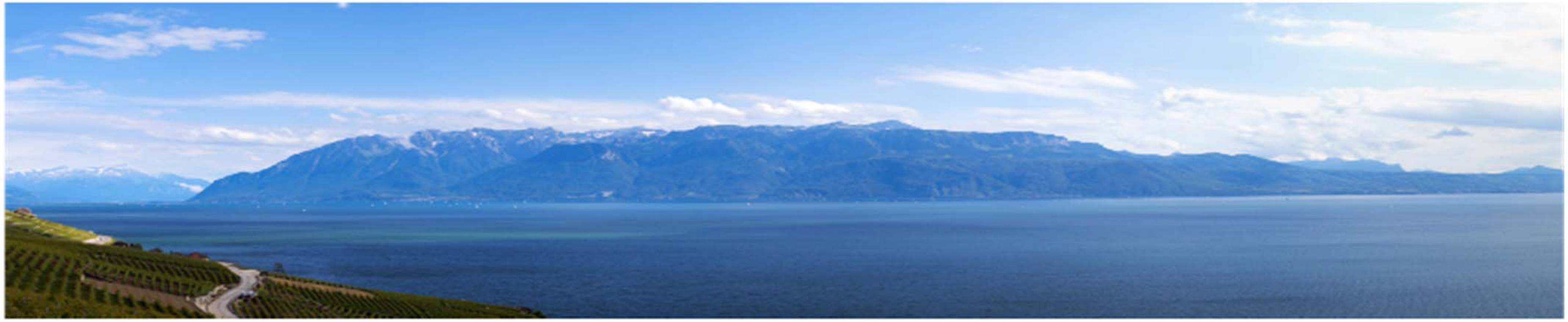
90nm technology
Pixel size: 0.75µm×0.75µm
Chip size: 2mm×2mm
Resolution: 1600×1200

Designed by Prof. Charbon's team

- Super resolution for binary images
- Color sensor

References

- A.J.P. Theuwissen. CMOS image sensors: State-of-the-art. In *Solid-State Electronics*, vol.52, Sep 2008, pp 1401-1406
- E. R. Fossum. Gigapixel digital film sensor. (invited) in *Nanospace Manipulation of Photons and Electrons for Nanovision Systems, The 7th Takayanagi Kenjiro Memorial Symposium and the 2nd International Symposium on Nanovision Science*, University of Shizuoka, Hamamatsu, Japan, October 25-26, 2005
- L. Sbaiz, **F. Yang**, E. Charbon, S. Süsstrunk, M. Vetterli. The gigavision camera. In *IEEE Conference on Acoustics, Speech and Signal Processing*, 2009
- **F. Yang**, L. Sbaiz, E. Charbon, S. Süsstrunk and M. Vetterli, Image reconstruction in the gigavision camera, *IEEE 12th International Conference on Computer Vision, Ninth Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras* (OMNIVIS 2009), pp. 2212-2219, 2009
- **F. Yang**, L. Sbaiz, E. Charbon, S. Süsstrunk and M. Vetterli, On pixel detection threshold in the gigavision camera, Proceedings of IS&T/SPIE Electronic Imaging, Digital Photography VI, Vol. 7537, 2010
- **F. Yang**, Y. Lu, L. Sbaiz and M. Vetterli, An optimal algorithm for reconstructing images from binary measurements, Proceedings of IS&T/SPIE Electronic Imaging, Computational Imaging VIII, 2010
- **F. Yang**, Y. Lu, L. Sbaiz and M. Vetterli, Bits from photons: Oversampled image acquisition using binary Poisson Statistics, IEEE Transactions on Image Processing, 2012, accepted



http://panorama.epfl.ch

Questions?