First-Photon Imaging and Other Imaging with Few Photons

*Modeling at the right scale*

*Inverse-problem mentality => denoising mentality*

Vivek K Goyal
Boston University
Electrical and Computer Engineering
Reflectivity and depth from 1 detected photon per pixel
Half from active source, half from background light and dark counts

Key idea: Image formation that integrates physical modeling of acquisition and scene modeling can provide dramatic improvements

### Photon-efficient depth+reflectivity imaging: Variations

<table>
<thead>
<tr>
<th>Study</th>
<th>Deterministic acquisition time</th>
<th>Exploit pulse shape</th>
<th>Exploit transverse smoothness</th>
<th>Uncalibrated background</th>
<th>Estimate multiple layers</th>
<th>Compensate for array properties</th>
</tr>
</thead>
</table>

In prep: fluorescence lifetime imaging, transverse super-resolution, unambiguous range extension
Time-of-flight depth imaging
Conventional active optical depth imaging

pulsed light source

pulsed illumination

time-resolved sensor
Conventional active optical depth imaging

- pulsed light source
- pulsed illumination
- time-resolved sensor
- back-reflected light
Conventional active optical depth imaging

pulsed light source

time-resolved sensor

transmitted

received
Conventional active optical depth imaging

Pixelwise measurement
- distance
Conventional active optical depth imaging

Pixelwise measurement
- distance
- ambient light

Pulsed light source

Time-resolved sensor

Time of flight

Background
Conventional active optical depth imaging

Pixelwise measurement
- distance
- ambient light
- reflectivity
Conventional active optical depth imaging

- Pulsed light source

- Reflectivity image

- Depth image

- Time-resolved sensor
Photon-efficient implementation
Detector sensitive to individual photons

- Micro Photon Devices single-photon avalanche diode: 35% quantum efficiency, 100 μm x 100 μm

Finite-resolution time tagging

- PicoQuant HydraHarp time-correlated single-photon counting module: 8 ps

Histogram as proxy for waveform

A: brighter, nearer
B: darker, farther
Classical noise models
Detector noise models for optical imaging systems (ground truth)

Reflectivity \( \{\alpha_{ij}\} \)

Range (depth) \( \{Z_{ij}\} \)

Photon flux

\( n_{ij} \)

no. of photons

pixel index

time
$\gtrapprox 10^4$ detected photons/pixel

error in photon count
$\approx$ Gaussian

error in ML depth est.
$\approx$ Gaussian
$\leq 10^4$ detected photons/pixel

 photon count = Poisson r.v.

 error in ML depth est. $\approx$ Gaussian

 no. of detected photons/pixel [no background]
≈1 photon/pixel

(half signal, half noise ...)
Depth imaging of two scenes (few photons per pixel)

<table>
<thead>
<tr>
<th>Ground truth depth</th>
<th>Log-matched filter</th>
<th>Median filter</th>
<th>BM3D</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Ground truth depth" /></td>
<td><img src="image2.png" alt="Log-matched filter" /></td>
<td><img src="image3.png" alt="Median filter" /></td>
<td><img src="image4.png" alt="BM3D" /></td>
</tr>
<tr>
<td>RMSE = 392.6 cm</td>
<td>RMSE = 43.6 cm</td>
<td>RMSE = 36.9 cm</td>
<td>RMSE = 20.1 cm</td>
</tr>
</tbody>
</table>
Imaging from only first photon detection
Conventional image formation (one detected photon/pixel)

Featureless image

\[ \text{var}(t_{ij}) \propto \text{mean-square pulse width} \]
Conventional image formation with background noise (one detected photon/pixel)

Featureless image

\[ \text{var}(t_{ij}) \approx (\text{pulse-period})^2 / 12 \]
Quantum nature of photon detection
raster-scanned, pulsed light source illuminates patch \((i,j)\)
raster-scanned, pulsed light source

single-photon avalanche detector (SPAD)
Poisson photo-detection statistics

\[ \lambda_{ij}(t) \]

Backscattered signal photons detections are arrivals in an inhomogeneous Poisson process

\[ \alpha_{ij} \]

\[ \frac{2Z_{ij}}{c} \]

[Saleh et al.]
Background and dark count detections are arrivals in a homogeneous Poisson process.
Photon detections are arrivals in a merged Poisson process.
Low-light level photo-detection

\[ \lambda_{ij}(t) \]

Not every incident light pulse generates a photon detection

\[ n_{ij} = 1 \]

\[ T_r \]

---

total response
Low-light level photo-detection

\[ \lambda_{ij}(t) \]

\[ n_{ij} = 1 \quad n_{ij} = 2 \]

total response
Low-light level photo-detection

\[ \lambda_{ij}(t) \]

\[ n_{ij} = 1 \quad n_{ij} = 2 \quad n_{ij} = 3 \]

- scene response
- background
- total response

photon detected!
Low-light level photo-detection

\[ \lambda_{ij}(t) \]

\[ n_{ij} = 1 \quad n_{ij} = 2 \quad n_{ij} = 3 \]

- scene response
- background
- total response

photon detected!
Low-light level photo-detection

\[ \lambda_{ij}(t) \]

\[ n_{ij} = 1 \quad n_{ij} = 2 \quad n_{ij} = 3 \]

- scene response
- background
- total response

photon detected!
Reinterpret what was observed
Two key random variables

\( n_{ij} \) = number of pulses transmitted before first photon detection

\( t_{ij} \) = detection time relative to time of last pulse emission

Roles of these variables

\( n_{ij} \) encodes reflectivity \( \alpha_{ij} \) via geometric distribution

\( t_{ij} \) encodes depth \( z_{ij} \) via normalized pulse shape distribution
Aside: **Fixed** number of pulses

**Key random variables**

\[ k_{ij} = \text{number of photon detections} \]

\[ t_{ij1}, t_{ij2}, ..., t_{ijk_i} = \text{detection times relative to times of last pulse emission} \]

**Roles of these variables**

\[ k_{ij} \text{ encodes reflectivity } \alpha_{ij} \text{ via binomial distribution} \]

\[ t_{ijk_i}'s \text{ encode depth } z_{ij} \text{ via normalized pulse shape distribution} \]
Quantitative acquisition modeling

\[ n_{ij} = 1 \]

\[ n_{ij} = 2 \]

\[ n_{ij} = 3 \]

Pr[no photodetection in one period] = \( e^{-(\alpha_{ij} + b\lambda)T} \)

Pr[\( n_{ij} = k; \alpha_{ij} \)] = \( e^{-(\alpha_{ij} + b\lambda)T(k-1)}[1 - e^{-(\alpha_{ij} + b\lambda)T}] \)

\( f_{t_{ij}}(\tau | \text{detected signal photon}; z_{ij}) \propto s(\tau - 2z_{ij}/c) \)

Pr[background photon detection] = \( \frac{b\lambda}{\alpha_{ij} + b\lambda} \)

= 1 – Pr[signal photon detection]

\( f_{t_{ij}}(\tau | \text{detected background photon}; z_{ij}) = \frac{1}{T_r} \)

In our experiment Pr[background photon detection] = 0.5
Raster scanning is used to collect first-photon data for each image pixel.
First-photon reflectivity data

Number of pulses transmitted before first photon detection

\[ n_{ij} \]

\[
Pr[n_{ij} = k; \alpha_{ij}] = e^{-(\alpha_{ij} + b_2 T)(k-1)} \left[ 1 - e^{-(\alpha_{ij} + b_2 T)} \right]
\]
First-photon reflectivity data

First-photon time-of-flight data

Number of pulses transmitted before first photon detection

\[ n_{ij} \]

\[
\Pr[n_{ij} = k; \alpha_{ij}] = e^{-(\alpha_{ij} + b \lambda T)(k-1)} \left[ 1 - e^{-(\alpha_{ij} + b \lambda T)} \right]
\]

First-photon detection time relative to last transmitted pulse

\[ t_{ij} \]

\[
f_{t_{ij}} (\tau; z_{ij} \mid \text{detected signal photon}) \propto s(\tau - 2z_{ij}/c)
\]
Pointwise estimation
Reflectivity information overlaid on 3D data
Novel image formation
Combining first-photon physics with spatial correlations
Combining first-photon physics with spatial correlations

Image formation method

Step 1: Estimate reflectivity from elapsed pulse data $\{n_{ij}\}$

Step 2: Censor background noise photons using ROAD filtering

Step 3: Estimate depth from uncensored time-of-arrival data $\{t_{ij}\}$
Step 1: Reflectivity estimation using regularized maximum likelihood estimation

**data likelihood**

\[ \Pr[n_{ij} = k; \alpha_{ij}] = e^{-(\alpha_{ij} + b\lambda)T(k-1)}[1 - e^{-(\alpha_{ij} + b\lambda)T}]} \]

**regularized ML estimation**

\[ \arg\min_{A = \{\alpha_{11}, \ldots, \alpha_{NN}\}} \sum_{i} \sum_{j} - \log \Pr[n_{ij}; a_{ij}] + \beta \| \Phi_{\alpha} A \|_1 \]

- data fidelity term
- parameter
- analysis with sparsity-promoting basis (wavelet)
Step 1

Reflectivity Reconstruction

(Mannequin Dataset)
Combining first-photon physics with spatial correlations

Image reconstruction method

Step 1: Estimate reflectivity from elapsed pulse data \( \{n_{ij}\} \)

Step 2: Censor background noise photons using ROAD filtering

Step 3: Estimate depth from uncensored time-of-arrival data \( \{t_{ij}\} \)
Step 2: Background photon censoring

Pr[background photon detection] = \frac{b_\lambda}{\alpha_{ij} + b_\lambda}

f_{t_{ij}}(\tau | \text{detected signal photon } ; z_{ij}) \propto s(\tau - 2z_{ij}/c)

f_{t_{ij}}(\tau | \text{detected ambient photon } ; z_{ij}) = \frac{1}{T_r}

\begin{align*}
|t_{ij} - t_1| \\
|t_{ij} - t_2| \\
|t_{ij} - t_3| \\
& \vdots \\
|t_{ij} - t_8| \\
\end{align*}

= \text{ROAD}_{ij}
Step 2: Background photon censoring

ROAD filtering

\[
\text{if } \text{ROAD}_{ij} \geq 4T_p \frac{b_\lambda}{\alpha_{ij} + b_\lambda} \quad \text{then } (i, j) \text{ is censored.}
\]

RMS pulse-width

estimated reflectivity

background level
Step 2

Background Noise Censoring

(Mannequin Dataset)
Combining first-photon physics with spatial correlations

Image reconstruction method

Step 1: Estimate reflectivity from elapsed pulse data \( \{n_{ij}\} \)

Step 2: Censor background noise photons using ROAD filtering

Step 3: Estimate depth from uncensored time-of-arrival data \( \{t_{ij}\} \)
Step 3: Depth estimation using regularized maximum likelihood estimation

Data likelihood:

\[ f_{t_{ij}} (\tau | \text{detected signal photon} ; z_{ij}) \propto s(\tau - 2z_{ij}/c) \]

Regularized ML estimation:

\[
\arg\min_{D = \{d_{ij}\}} \sum_{\text{uncensored locations}} - \log f_{t_{ij}} (\tau | \text{signal} ; d_{ij}) + \beta \parallel \Phi \alpha D \parallel_1
\]

- Data likelihood
- Regularized ML estimation
- Parameter
- Data fidelity term at uncensored pixels
- Sparsity promoting bases (wavelets)
Step 3

3D form Reconstruction

(Mannequin Dataset)
Reflectivity information overlaid on 3D data

Conventional pixelwise maximum likelihood estimates

First-photon imaging
Experiments and evaluation
Experimental setup

- Incandescent lamp
- Optical filter: 2 nm bw, 49%
- Geiger-mode APD: (100 μm)$^2$, 35%
- Scanning galvo
- Pulsed light source: 640 nm laser diode, 226 ps RMS duration, 10 MHz repetition rate, 0.6 mW average power
- Scene: ≈ 2 m
- Control and processing
- Sync and timing
- Correlator: PicoQuant HydraHarp, 8 ps resolution
Experimental setup
Limitations

- High error at lateral surfaces, edges, and corners
- High optical flux (almost every pulse leads to detection)
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Journal/Conference</th>
<th>Year</th>
<th>Deterministic acquisition time</th>
<th>Exploit pulse shape</th>
<th>Exploit transverse smoothness</th>
<th>Uncalibrated background</th>
<th>Estimate multiple layers</th>
<th>Compensate for array properties</th>
</tr>
</thead>
</table>

In prep: fluorescence lifetime imaging, transverse super-resolution, unambiguous range extension
Related work with fixed dwell time

- Parallelizable (detector array)
- For low photon count on average, many pixels have 0 detections
- Performance can be even better than FPI

<table>
<thead>
<tr>
<th></th>
<th>Random dwell</th>
<th>Fixed dwell</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mannequin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean $T_a$</td>
<td>244 $\mu$s</td>
<td>244 $\mu$s</td>
</tr>
<tr>
<td>Mean $k_{i,j}$</td>
<td>1 ppp</td>
<td>2.7 ppp</td>
</tr>
<tr>
<td>Pixels missing data</td>
<td>0%</td>
<td>33%</td>
</tr>
<tr>
<td>PSNR</td>
<td>35 dB</td>
<td>37 dB</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.4 cm</td>
<td>0.3 cm</td>
</tr>
<tr>
<td><strong>Sunflower</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean $T_a$</td>
<td>15 $\mu$s</td>
<td>15 $\mu$s</td>
</tr>
<tr>
<td>Mean $k_{i,j}$</td>
<td>1 ppp</td>
<td>8.7 ppp</td>
</tr>
<tr>
<td>Pixels missing data</td>
<td>0%</td>
<td>18%</td>
</tr>
<tr>
<td>PSNR</td>
<td>15 dB</td>
<td>16 dB</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.8 cm</td>
<td>0.5 cm</td>
</tr>
<tr>
<td><strong>Basketball-and-can</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean $T_a$</td>
<td>181 $\mu$s</td>
<td>181 $\mu$s</td>
</tr>
<tr>
<td>Mean $k_{i,j}$</td>
<td>1 ppp</td>
<td>1.7 ppp</td>
</tr>
<tr>
<td>Pixels missing data</td>
<td>0%</td>
<td>24%</td>
</tr>
<tr>
<td>PSNR</td>
<td>40 dB</td>
<td>40 dB</td>
</tr>
<tr>
<td>RMSE</td>
<td>1.1 cm</td>
<td>1.1 cm</td>
</tr>
<tr>
<td><strong>Reflectivity chart</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean $T_a$</td>
<td>120 $\mu$s</td>
<td>120 $\mu$s</td>
</tr>
<tr>
<td>Mean $k_{i,j}$</td>
<td>1 ppp</td>
<td>1.7 ppp</td>
</tr>
<tr>
<td>Pixels missing data</td>
<td>0%</td>
<td>27%</td>
</tr>
<tr>
<td>PSNR</td>
<td>40 dB</td>
<td>42 dB</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.4 cm</td>
<td>0.4 cm</td>
</tr>
<tr>
<td><strong>Depth chart</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean $T_a$</td>
<td>6.2 $\mu$s</td>
<td>6.2 $\mu$s</td>
</tr>
<tr>
<td>Mean $k_{i,j}$</td>
<td>1 ppp</td>
<td>1.1 ppp</td>
</tr>
<tr>
<td>Pixels missing data</td>
<td>0%</td>
<td>35%</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.4 cm</td>
<td>0.4 cm</td>
</tr>
<tr>
<td>Authors</td>
<td>Deterministic</td>
<td>Exploit pulse shape</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------</td>
<td>---------------------</td>
</tr>
</tbody>
</table>

In prep: fluorescence lifetime imaging, transverse super-resolution, unambiguous range extension.
Unknown background and no transverse regularization

- Exploit union-of-subspaces model for each pixel separately
- CoSaMP-inspired efficient algorithm
- Compared to log-matched filter, MAE reduction factor of 6 using 15 photons per pixel
<table>
<thead>
<tr>
<th>Article</th>
<th>Deterministic acquisition time</th>
<th>Exploit pulse shape</th>
<th>Exploit transverse smoothness</th>
<th>Uncalibrated background</th>
<th>Estimate multiple layers</th>
<th>Compensate for array properties</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shin, Xu, Wong, Shapiro, Goyal, <em>Optics Express</em>, 24(3):1873-1888, 2016</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

In prep: fluorescence lifetime imaging, transverse super-resolution, unambiguous range extension
Multiple depths per pixel and no transverse regularization

- Exploit longitudinal sparsity for each pixel separately
- ISTA-inspired efficient algorithm for convex relaxation of problem
- Compare to mixture of Gaussians fit with EM (shown at 19 ppp)
<table>
<thead>
<tr>
<th>Authors</th>
<th>Deterministic acquisition time</th>
<th>Exploit pulse shape</th>
<th>Exploit transverse smoothness</th>
<th>Uncalibrated background</th>
<th>Estimate multiple layers</th>
<th>Compensate for array properties</th>
</tr>
</thead>
</table>

In prep: fluorescence lifetime imaging, transverse super-resolution, unambiguous range extension
Mitigate challenges of SPAD array

- Much coarser time resolution, hot pixels
- Developed improved noise censoring, longitudinal super-resolution
- RMS error at one-third RMS pulse width with 1 signal photon/pixel
First-Photon Imaging and Other Imaging with Few Photons

Image formation that integrates physical modeling of acquisition and scene modeling can provide dramatic improvements.

Model at the right scale.

Apply an inverse problem mentality.
First-Photon Imaging and Other Imaging with Few Photons