Expectation Propagation for Scalable Inverse Problems in Imaging

Dan Yao

Heriot-Watt University

19, June, 2023

This talk is about Expectation Propagation (EP)

- **Problems and Challenges**

  image inverse problems

  Bayesian estimation strategy

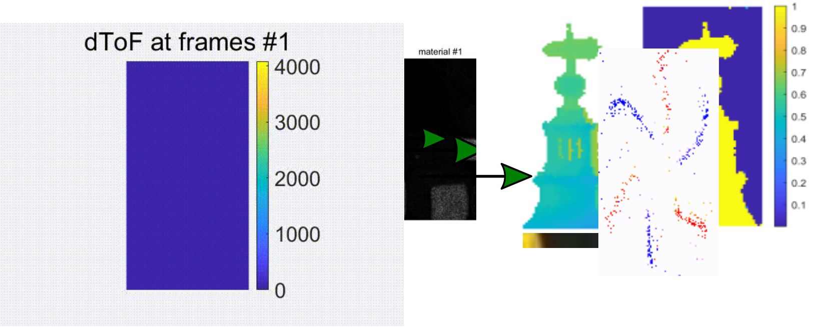
  challenges

- **Approximate Solution by using EP**

  basic idea of EP

  key steps for practical implementation

- **Real-world Applications of EP**



- **Summary**

This talk is about Expectation Propagation (EP)

- **Problems and Challenges**

  image inverse problems

  Bayesian estimation strategy

  challenges

- **Approximate Solution by using EP**

  basic idea of EP

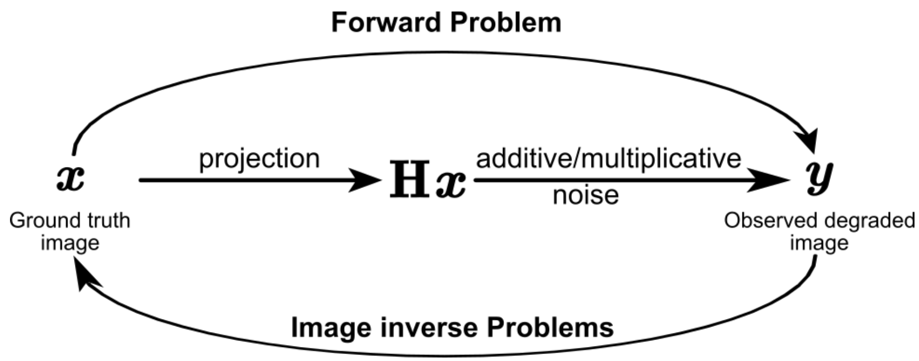
  key steps for practical implementation

- **Real-world Applications of EP**

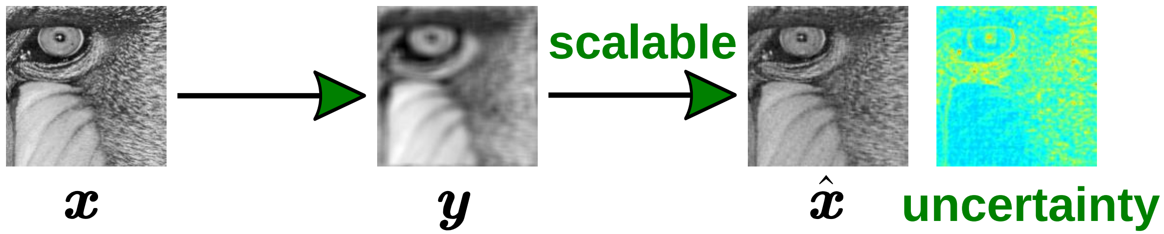
- **Summary**

Problems and Challenges

- **Image inverse problems**



goal: y —scalable→ $\hat{x}$ + uncertainty quantification



Problems and Challenges

- **Bayesian estimation strategy**

likelihood for forward model: $f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})$

prior knowledge: $f_x(\boldsymbol{x}|\boldsymbol{\theta})$

posterior inference: $p(\boldsymbol{x}|\boldsymbol{y}) = \dfrac{f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})\,f_x(\boldsymbol{x}|\boldsymbol{\theta})}{\int f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})\,f_x(\boldsymbol{x}|\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{x}}$

$\hat{\boldsymbol{x}}$: $\hat{\boldsymbol{x}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{x}} \; p(\boldsymbol{x}|\boldsymbol{y})$ or $\hat{\boldsymbol{x}}_{\mathrm{MMSE}} = \mathbb{E}_p[\boldsymbol{x}]$

uncertainty: $\mathrm{Cov}_p(\boldsymbol{x}) = \mathbb{E}_p\left[(\boldsymbol{x} - \mathbb{E}_p[\boldsymbol{x}])(\boldsymbol{x} - \mathbb{E}_p[\boldsymbol{x}])^T\right]$

scalable

goal: $\boldsymbol{y}$ → $\mathbb{E}_p[\boldsymbol{x}] + \mathrm{Cov}_p[\boldsymbol{x}]$

$\hat{\boldsymbol{x}}$ **+ uncertainty quantification**

Problems and Challenges

- **Challenges: high-dimensional unknowns in $\boldsymbol{x}$**

e.g. $\boldsymbol{x} = [x_1, \ldots, x_{10000}]^T$

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})\, f_x(\boldsymbol{x}|\boldsymbol{\theta})}{\int f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})\, f_x(\boldsymbol{x}|\boldsymbol{\theta})\, \mathrm{d}[x_1, \ldots, x_{10000}]^T}$$

$$\mathbb{E}_p[\boldsymbol{x}] = \int \boldsymbol{x}\, p(\boldsymbol{x}|\boldsymbol{y})\, \mathrm{d}[x_1, \ldots, x_{10000}]^T$$

$$\mathrm{Cov}_p(\boldsymbol{x}) = \int (\boldsymbol{x} - \mathbb{E}_p[\boldsymbol{x}])(\boldsymbol{x} - \mathbb{E}_p[\boldsymbol{x}])^T p(\boldsymbol{x}|\boldsymbol{y})\, \mathrm{d}[x_1, \ldots, x_{10000}]^T$$

This talk is about Expectation Propagation (EP)

- **Problems and Challenges**

  image inverse problems

  Bayesian estimation strategy

  challenges

- **Approximate Solution by using EP**

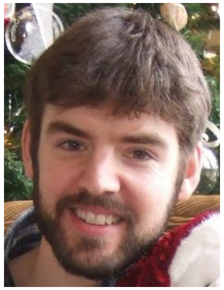
  basic idea of EP

  key steps for practical implementation

- **Real-world Applications of EP**

- **Summary**

Approximate Solution by using EP



362 MINKA UAI 2001

Expectation Propagation for Approximate Bayesian Inference

Thomas P. Minka
Statistics Dept.
Carnegie Mellon University
Pittsburgh, PA 15213

- **Basic idea:** $q(\boldsymbol{x})$ $\boldsymbol{Kullback-Leibler}$ divergence minimization $\approx$ $p(\boldsymbol{x})$

➢ $q(\boldsymbol{x})$ is easier to compute: $p(\boldsymbol{x}|\boldsymbol{y}) = \dfrac{f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x}) f_x(\boldsymbol{x}|\boldsymbol{\theta})}{\int f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x}) f_x(\boldsymbol{x}|\boldsymbol{\theta}) \mathrm{d}\boldsymbol{x}}$

➢ Computing $\mathbb{E}_q[\boldsymbol{x}], \mathrm{Cov}_q(\boldsymbol{x})$ could be scalable:

$$\mathbb{E}_{q(\boldsymbol{x})}[\boldsymbol{x}] \overset{\text{scalable}}{\approx} \mathbb{E}_p[\boldsymbol{x}] \qquad \mathrm{Cov}_q(\boldsymbol{x}) \overset{\text{scalable}}{\approx} \mathrm{Cov}_p(\boldsymbol{x})$$

Approximate Solution by using EP

KL divergence: $q(\boldsymbol{x}) = \arg\min_{q(\boldsymbol{x}) \in \mathcal{Q}} KL(p(\boldsymbol{x}) || q(\boldsymbol{x}))$

Exponential family: $\mathcal{Q} = \{q(\boldsymbol{x}) = \boldsymbol{e}^{T(\boldsymbol{x})^T \boldsymbol{\eta} - A(\boldsymbol{\eta}) + \boldsymbol{B}(\boldsymbol{x})}\}$

moment matching: $\frac{\partial A(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \mathbb{E}_p[T(\boldsymbol{x})]$

! $\mathbb{E}_p[T(\boldsymbol{x})]$ **intractable**

- **Key steps for practical implementation: factorization**

$$p(\boldsymbol{x}) = p_1(\boldsymbol{x}) p_2(\boldsymbol{x}) \ldots p_i(\boldsymbol{x}) \ldots$$

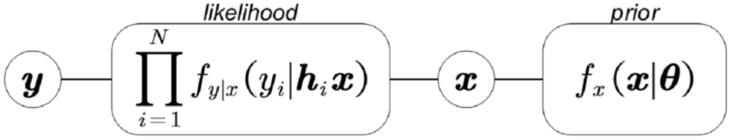
$$\wr\wr \qquad \wr\wr \qquad \wr\wr$$

$$q_1(\boldsymbol{x})\, q_2(\boldsymbol{x}) \ldots q_i(\boldsymbol{x}) \ldots = q(\boldsymbol{x}) \qquad \textbf{difficult} \gg \{\textbf{simple} + \ldots + \textbf{simple}\}$$

$$KL\left(p_1(\boldsymbol{x}) q_2(\boldsymbol{x}) \ldots q_i(\boldsymbol{x}) \ldots q_N(\boldsymbol{x}) || \underline{q_1(\boldsymbol{x})} q_2(\boldsymbol{x}) \ldots q_i(\boldsymbol{x}) \ldots q_N(\boldsymbol{x})\right) \Longrightarrow \frac{\partial A(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \mathbb{E}_{pq}[T(\boldsymbol{x})]$$

Approximate Solution by using EP

factor graph representation



$$\prod_{i=1}^{N} q_i(\boldsymbol{x}) \qquad q_0(\boldsymbol{x})$$

sequential KL divergence minimization

$$\begin{cases} q_0(\boldsymbol{x}) = \arg\min_{q_0(\boldsymbol{x})\in\mathcal{Q}} KL\left(f(\boldsymbol{x}|\boldsymbol{\theta})q_{\backslash 0}(\boldsymbol{x})||q_0(\boldsymbol{x})q_{\backslash 0}(\boldsymbol{x})\right) \\ q_1(\boldsymbol{x}) = \arg\min_{q_1(\boldsymbol{x})\in\mathcal{Q}} KL\left(f(y_1|\boldsymbol{h}_1\boldsymbol{x})q_{\backslash 1}(\boldsymbol{x})||q_1(\boldsymbol{x})q_{\backslash 1}(\boldsymbol{x})\right) \\ q_2(\boldsymbol{x}) = \arg\min_{q_2(\boldsymbol{x})\in\mathcal{Q}} KL\left(f(y_2|\boldsymbol{h}_2\boldsymbol{x})q_{\backslash 2}(\boldsymbol{x})||q_2(\boldsymbol{x})q_{\backslash 2}(\boldsymbol{x})\right) \\ \vdots \\ q_N(\boldsymbol{x}) = \arg\min_{q_N(\boldsymbol{x})\in\mathcal{Q}} KL\left(f(y_N|\boldsymbol{h}_N\boldsymbol{x})q_{\backslash N}(\boldsymbol{x})||q_N(\boldsymbol{x})q_{\backslash N}(\boldsymbol{x})\right) \end{cases}$$

…until convergence $q_0(\boldsymbol{x})\prod_{i=1}^{N} q_i(\boldsymbol{x}) = q(\boldsymbol{x}) \approx p(\boldsymbol{x})$

This talk is about Expectation Propagation (EP)

- **Problems and Challenges**
  image inverse problems
  Bayesian estimation strategy
  challenges

- **Approximate Solution by using EP**
  basic idea of EP
  key steps for practical implementation

- **Real-world Applications of EP**

- **Summary**

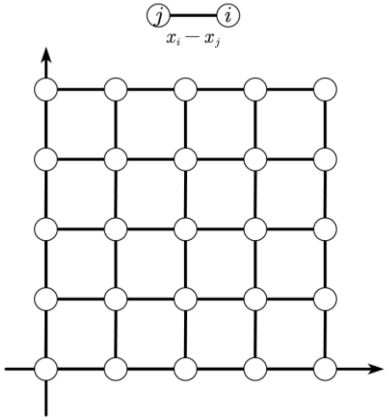
Real-world Applications of EP

- **application 1** - classical image estimation and uncertainty quantification

blurry observation $\boldsymbol{y}$
grayscale, 165×165 pixels

Gaussian likelihood: $f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y};\mathbf{H}\boldsymbol{x},\sigma^2)$

Total Variation prior: $f_x(\boldsymbol{x}|\boldsymbol{\theta}) \propto e^{-\lambda\, TV(\boldsymbol{x})}$

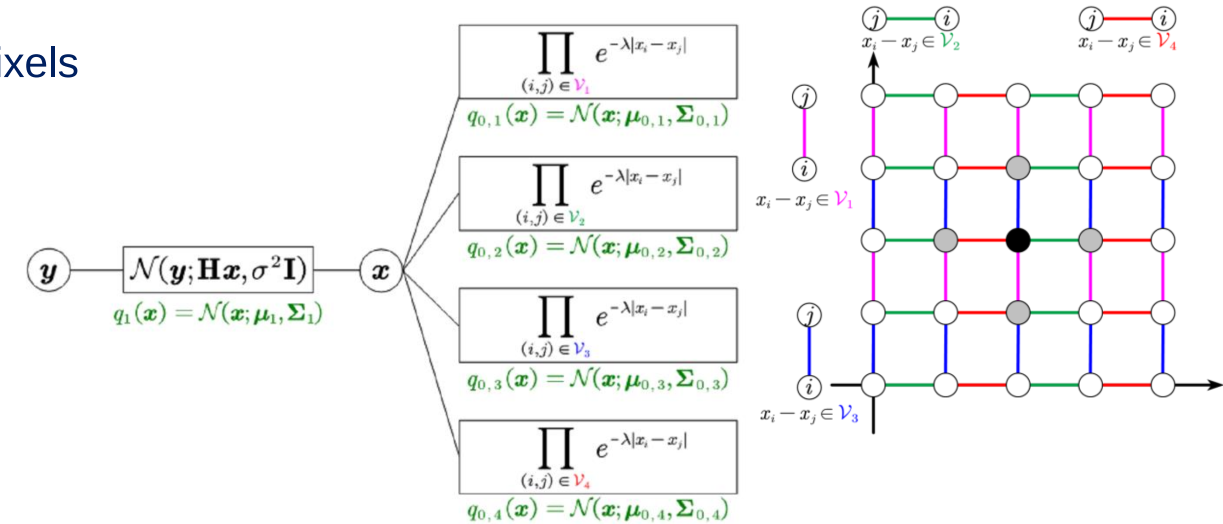


$$TV(\boldsymbol{x}) = \sum_{i,j} |x_{i+1,j} - x_{i,j}| + |x_{i,j+1} - x_{i,j}|$$

Real-world Applications of EP

- **application 1** - classical image estimation and uncertainty quantification

blurry observation $\boldsymbol{y}$
grayscale, 165×165 pixels



EP approximate posterior distribution of the deblurred image $\boldsymbol{x}$:

$$q(\boldsymbol{x}) \propto q_1(x) q_{0,1}(\boldsymbol{x}) q_{0,2}(\boldsymbol{x}) q_{0,3}(\boldsymbol{x}) q_{0,4}(\boldsymbol{x}) \propto \mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Real-world Applications of EP

- **application 1** - classical image estimation and uncertainty quantification

observation $\boldsymbol{y}$
PSNR: 18.4 dB

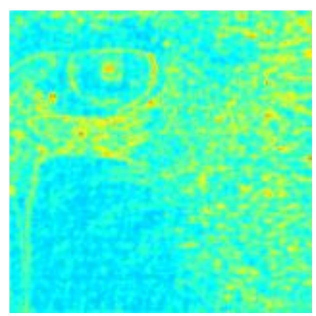
$\hat{\boldsymbol{x}}$

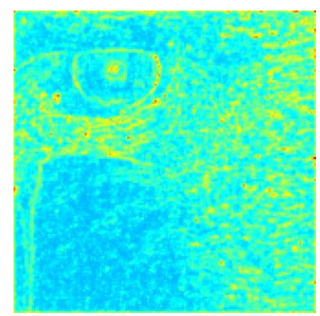
EP
PSNR: 23.0 dB

comparison: MCMC
PSNR: 23.0 dB

ground truth $\boldsymbol{x}$

uncertainty:





CPU computational time: 40 seconds 1 hour

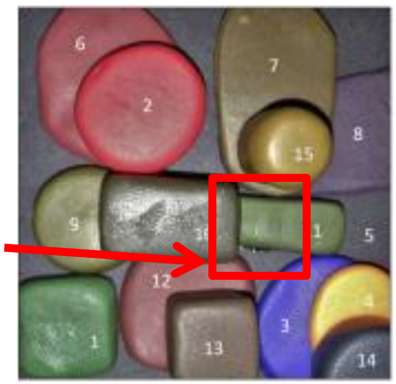
Dan Yao, Stephen McLaughlin, and Yoann Altmann. "*Fast Scalable Image Restoration using Total Variation Priors and Expectation Propagation,*" IEEE Transactions on Image Processing, 2022, 31: 5762-5773.

Real-world Applications of EP

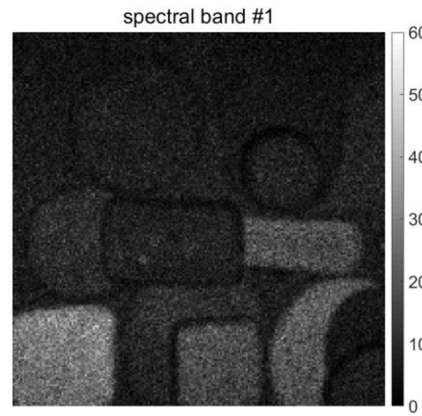
- **application 2** - robust multispectral unmixing from single-photon Lidar imaging

reference of the scene

anomalies (glue)



multi-spectral Lidar data observation $\boldsymbol{y}$
(190×190 pixels)×33 spectral bands



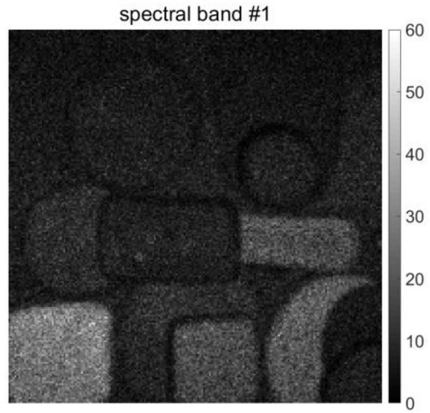
$$\text{Poisson likelihood: } f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x}) = \mathcal{P}((1-\boldsymbol{z}) \odot \mathbf{H}\boldsymbol{\lambda} + \boldsymbol{z} \odot \boldsymbol{r})$$

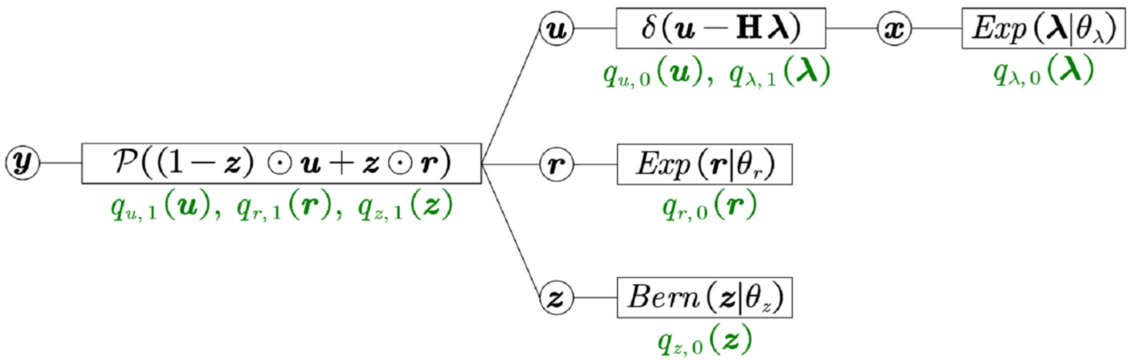
$$\text{positive \& sparse \& Bernoulli priors: } f_\lambda(\boldsymbol{\lambda}|\theta_\lambda),\ f_r(\boldsymbol{r}|\theta_r),\ f_z(\boldsymbol{z}|\theta_z)$$

Real-world Applications of EP

- **application 2** - robust multispectral unmixing from single-photon Lidar imaging

multi-spectral Lidar data observation $\boldsymbol{y}$
(190 × 190 pixels) × 33 spectral bands





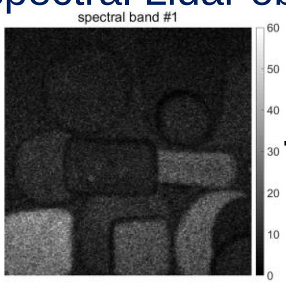
→ EP approximate posterior distributions of:

- signal spectral channels: $q(\boldsymbol{u}) \propto q_{u,1}(\boldsymbol{u})q_{u,0}(\boldsymbol{u}) \propto \mathcal{N}(\boldsymbol{u}; \boldsymbol{\mu}_u, \boldsymbol{\Sigma}_u)$
- anomaly amplitude: $q(\boldsymbol{r}) \propto q_{r,1}(\boldsymbol{r})q_{r,0}(\boldsymbol{r}) \propto \mathcal{N}(\boldsymbol{r}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r)$
- anomaly appearance : $q(\boldsymbol{z}) \propto q_{r,1}(\boldsymbol{z})q_{r,0}(\boldsymbol{z}) \propto \boldsymbol{Bern}(\boldsymbol{z}; \boldsymbol{p}_r)$

Real-world Applications of EP

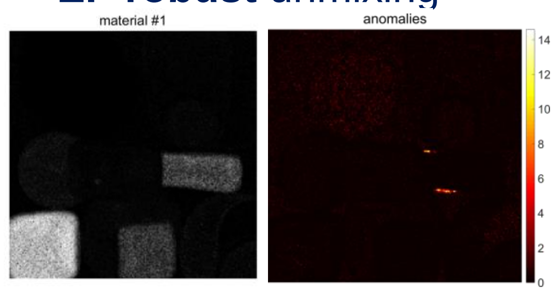
- **application 2** - robust multispectral unmixing from single-photon Lidar imaging

multi-spectral Lidar observation

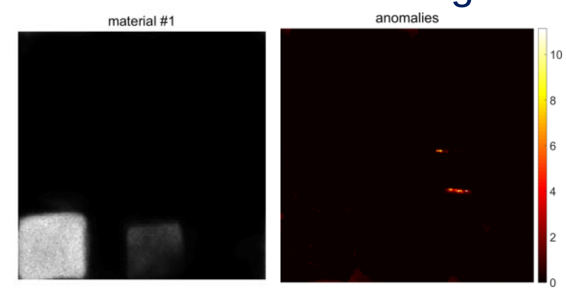


$\hat{\boldsymbol{x}}$

EP robust unmixing

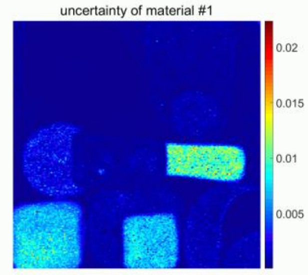


MCMC robust unmixing



reference

uncertainty:



CPU computational time: 0.3 seconds (per pixel) 14.3 hours

Dan Yao, Stephen McLaughlin, Yoann Altmann, Michael E Davies. "*Joint Robust Linear Regression and Anomaly Detection in Poisson noise using Expectation-Propagation*", 28th European Signal Processing Conference (EUSIPCO). 2021. pp. 2463-2467.

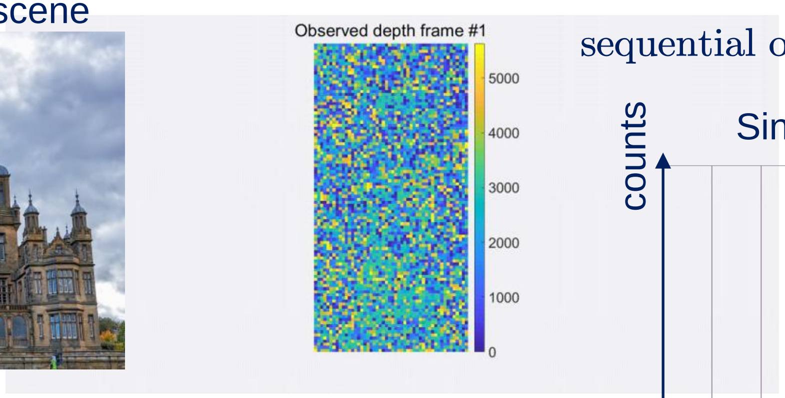
Real-world Applications of EP

- **application 3** - joint surface detection and depth estimation from single-photon Lidar imaging

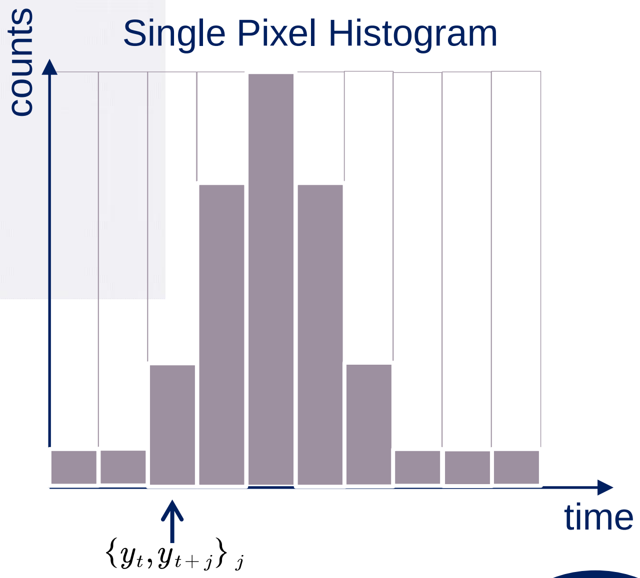
single-photon Lidar Time-of-Flight observations $\boldsymbol{y}$

(100 × 50 pixels) × T frames

reference of the scene



sequential observations: $\boldsymbol{y} = [y_1, y_2, \ldots, y_T]$



Real-world Applications of EP

- **application 3** - joint surface detection and depth estimation from single-photon Lidar imaging

Bayesian model for a single pixel:

$$\text{likelihood: } f_{y|x}(y_t|d,w) = w\mathcal{N}(y_t; d, \mathrm{s}^2) + (1-w)\mathcal{U}(y_t)$$

$$\text{priors: } f_d(d_t) = \mathcal{N}(d_t; m_t, s_t),\ f_w(w_t) = \mathcal{B}eta(\omega_t; \alpha_t, \beta_t)$$

$$\text{exact posterior: } p(d, w|y_1, \ldots, y_T)$$

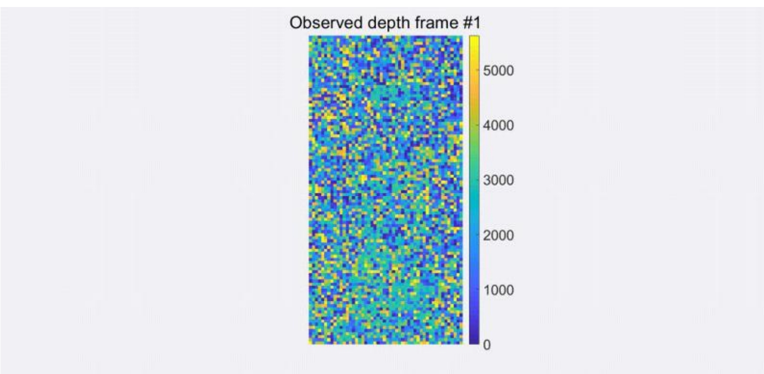
EP online posterior approximation: $q(d_t, \omega_t|y_t)$

Real-world Applications of EP

- **application 3** - joint surface detection and depth estimation from single-photon Lidar imaging

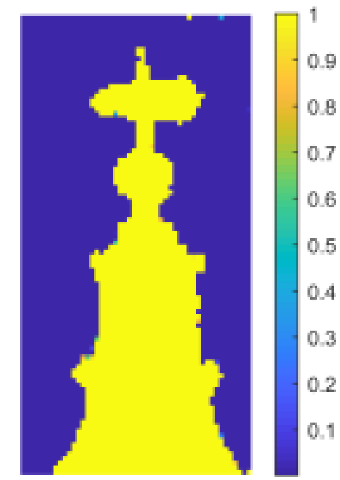
single-photon Lidar Time-of-Flight observations $\boldsymbol{y}$

(100 $\times$ 50 pixels) $\times$ T frames



$\hat{\boldsymbol{d}}, \hat{\boldsymbol{w}}$

depth estimate



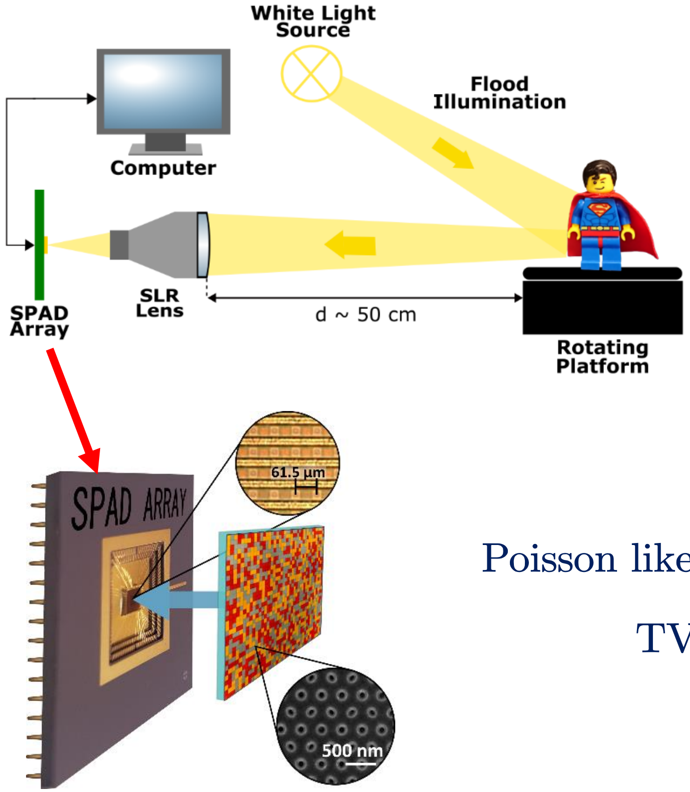
detected surface

K. Drummond, D. Yao, A. Pawlikowska, R. Lamb, S. McLaughlin and Y. Altmann. “Efficient joint surface detection and depth estimation of single-photon Lidar data using assumed density filtering“, *2022 Sensor Signal Processing for Defence Conference (SSPD)*, London, United Kingdom, 2022, pp. 1-5.

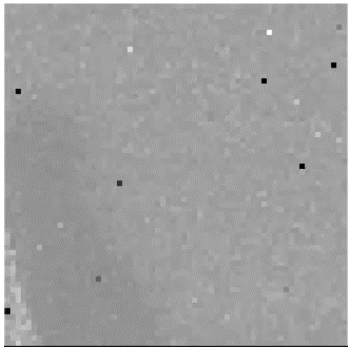
Real-world Applications of EP

- **application 4** - rapid color imaging of moving objects under low-lignt-level condition using a single-photon avalanche detector (SPAD) array

lab experimental setup



sequence of SPAD frames observations $\boldsymbol{y}$

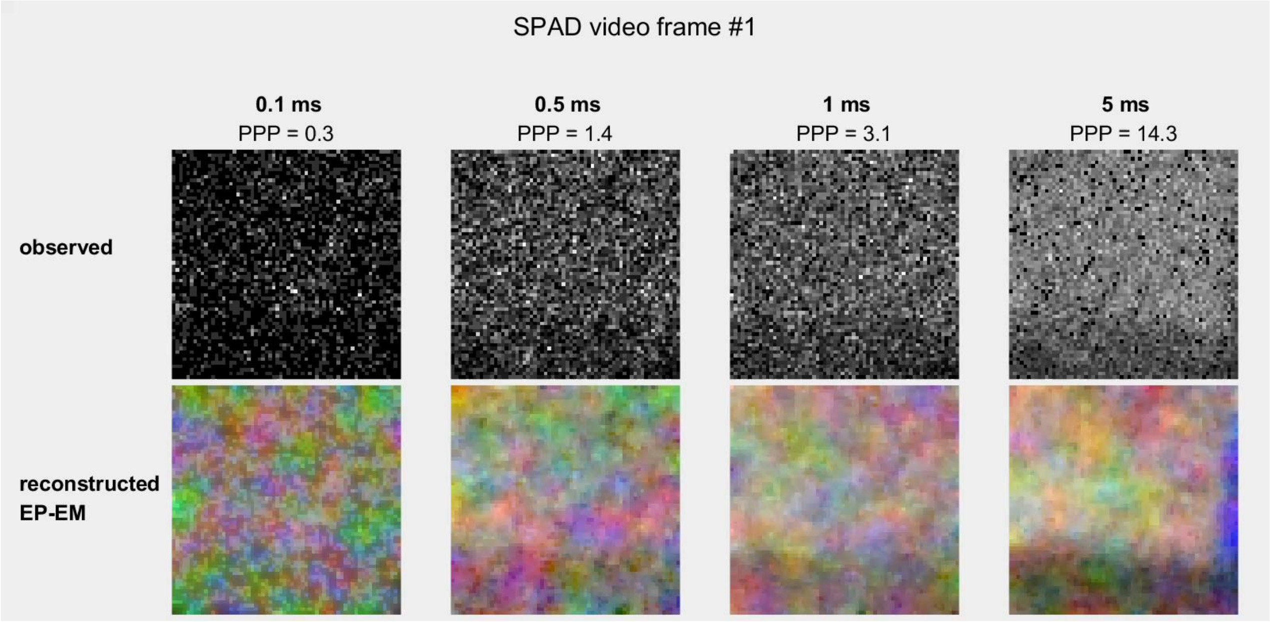


Poisson likelihood: $f_{y|x}(\boldsymbol{y}_t|\mathbf{H}\boldsymbol{x}_t) = \mathcal{P}(\mathbf{H}_R\boldsymbol{x}_{R,t} + \mathbf{H}_G\boldsymbol{x}_{G,t} + \mathbf{H}_B\boldsymbol{x}_{B,t} + \mathbf{b})$

TV prior: $f_x(\boldsymbol{x}_{R,t}),\ f_x(\boldsymbol{x}_{G,t}),\ f_z(\boldsymbol{x}_{B,t})$

Real-world Applications of EP

- **application 4** - rapid low-level-light of moving objects from single-photon color imaging

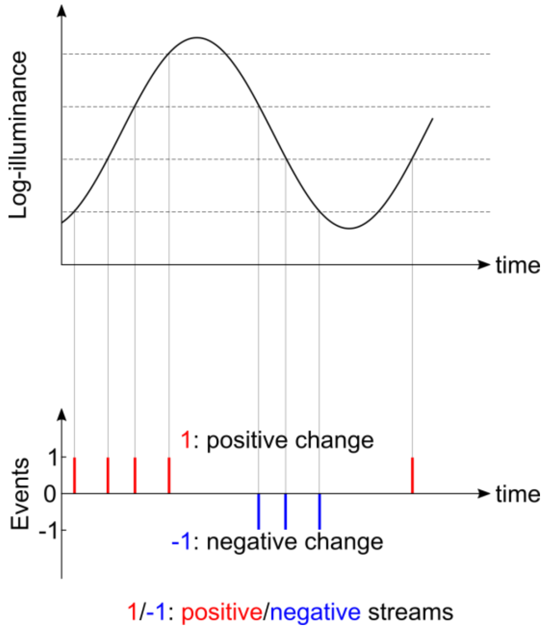


Dan Yao, Petet W. R. Connolly, Arran J. Sykesy, Yash D. Shan, Claudio Accarino, James Grant, David R. S. Cumming, Gerald S. Buller, Stephen McLaughlin, Yoann Altmann. "*Rapid Single-Photon Color Imaging of Moving Objects*", Optics Express (Under Review).

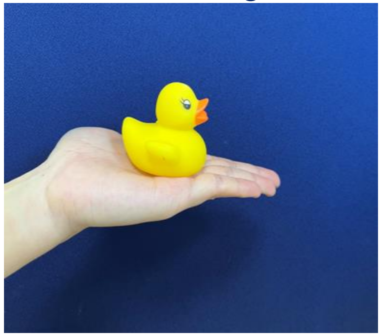
Real-world Applications of EP

- **application 5** - turning SPAD arrays into neuromorphic cameras

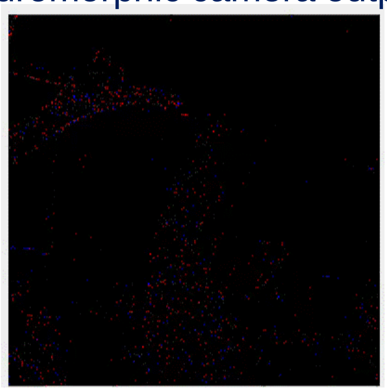
bio-inspired imaging and sensing by neuromorphic cameras



scene-moving duck



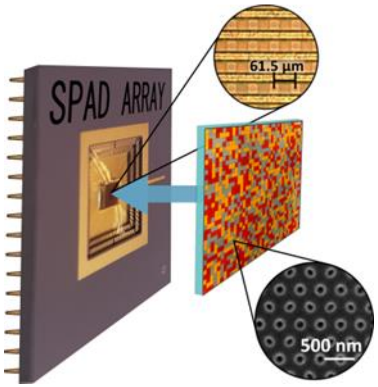
neuromorphic camera output



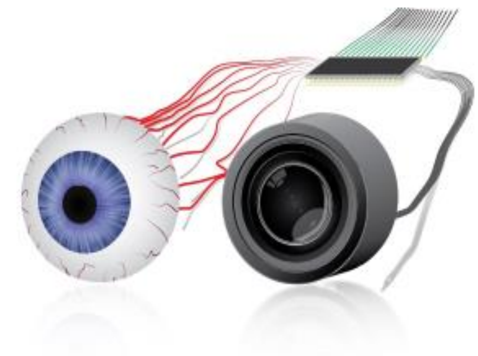
- **high temporal resolution**
- **high dynamic range**
- **low latency**

Real-world Applications of EP

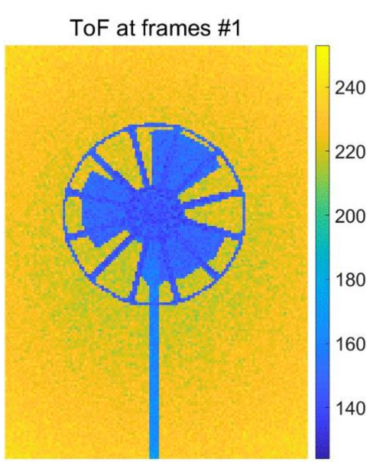
- **application 5** - turning SPAD arrays into neuromorphic cameras



SPAD array



neuromorphic camera
https://nusneuromorphic.github.io/

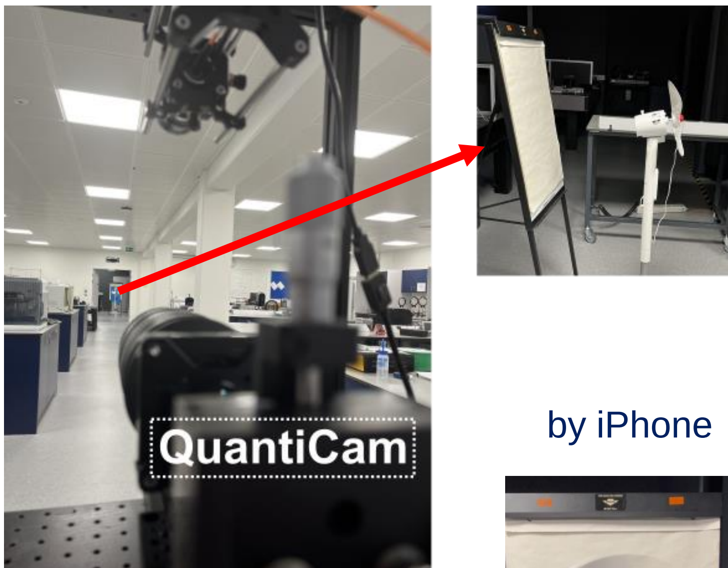


(**depth-based events**)

Real-world Applications of EP

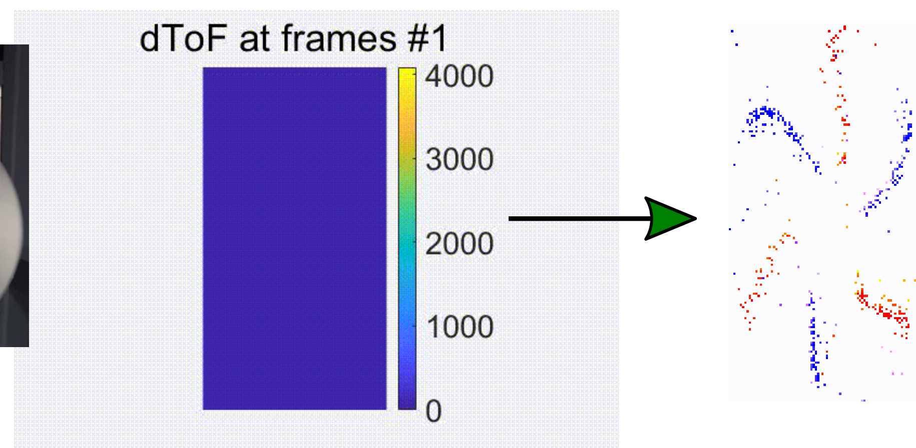
- **application 5** - turning SPAD arrays into neuromorphic cameras

lab experimental setup



by iPhone

by SPAD array



This talk is about Expectation Propagation (EP)

- **Problems and Challenges**

  image inverse problems

  Bayesian estimation strategy

  challenges

- **Approximate Solution by using EP**

  basic idea of EP

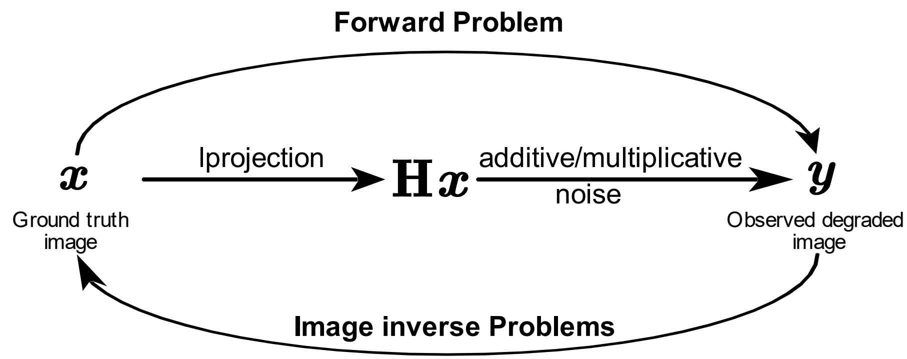
  key steps for practical implementation

- **Real-world Applications of EP**

- **Summary**

Summary

- ✓ image inverse problems

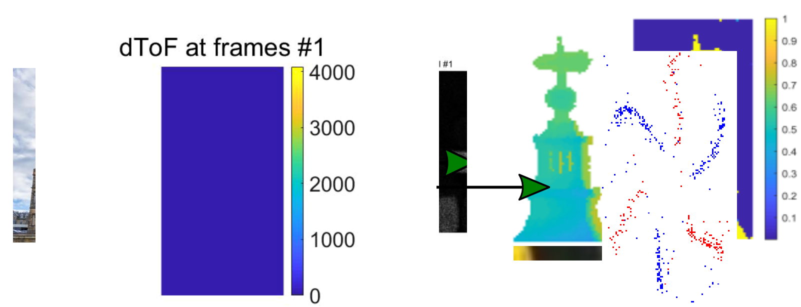


goal: $\boldsymbol{y} \longrightarrow \widehat{\boldsymbol{x}}$ + uncertainty quantification

- ✓ Approximate solution by using Expectation Propagation

$$\mathbb{E}_{q(\boldsymbol{x})}[\boldsymbol{x}] \overset{\text{scalable}}{\approx} \mathbb{E}_{p}[\boldsymbol{x}] \qquad \mathrm{Cov}_q(\boldsymbol{x}) \overset{\text{scalable}}{\approx} \mathrm{Cov}_p(\boldsymbol{x})$$

- ✓ Real-world applications



Some EP references

- **EP tutorial videos:**

- 1. Thomas Minka: **Approximate Inference** http://videolectures.net/mlss09uk_minka_ai/
- 2. Simon Barthelmé: **The Expectation-Propagation algorithm: a tutorial - Part 1** https://youtu.be/0tomU1q3AdY

- **Homepages:**

- 1. Thomas Minka: **A roadmap to research on EP** https://tminka.github.io/papers/ep/roadmap.html
- 2. Matt Wand:**Statistics Methodology and Theory** http://matt-wand.utsacademics.info/statsPapers.html
- 3. José Miguel Hernández-Lobato: **Scalable methods for approximate inference** https://jmhl.org/publications/
- 4. Matthias Seeger, Young-Jun Ko: **Scalable variational approximate inference algorithms** https://mseeger.github.io/

Thanks for you attention !

Expectation Propagation for Scalable Inverse Problems in Imaging

Dan Yao

Heriot-Watt University

D.Yao@hw.ac.uk

19, June, 2023