Bayesian image reconstruction

Clayton W. Seitz

June 12, 2022

Outline

References

Photon statistics of CMOS cameras

- Imaging noise consists of shot noise, thermal noise, and readout noise
- Shot noise is Poisson, thermal noise and readout noise are Gaussian

For a CMOS pixel n, the true signal S_n $[\mathrm{ADU}]$ is a Poisson process with rate parameter λ_n

$$S_n = \gamma g_n P_n(\lambda_n)$$

where γ $[e^{-}/p]$ is the quantum efficiency and g_n $[\mathrm{ADU}/e^{-}]$ is the pixel's gain

$$P(S_n) = \frac{\exp(-\lambda_n)\,\lambda_n^p}{p!}$$

But what is the distribution over the corrupted signal $P(\hat{S}_n)$?

Photon statistics of CMOS cameras

To find $P(\hat{S}_n)$, we first evaluate the joint density $P(S_n, \hat{S}_n)$

$$\begin{aligned} P(S_n, \hat{S}_n) &= P(\hat{S}_n | S_n = s) P(S_n = s) \\ &= \frac{1}{Z} \exp\left(-\frac{(\hat{S}_n - g_n s - \mu_n)^2}{\sigma_n^2}\right) \frac{\exp\left(-\lambda_n\right) \lambda_n^s}{s!} \end{aligned}$$

Marginalizing over S_n gives the desired distribution over $\hat{S}_n$

$$P(\hat{S}_n) = \frac{1}{Z} \sum_{s=0}^{\infty} \frac{\exp\left(-\lambda_n\right) \lambda_n^s}{s!} \exp\left(-\frac{(\hat{S}_n - g_n s - \mu_n)^2}{\sigma_n^2}\right)$$

Bayesian parameter inference for CMOS photon statistics

The parameters in our model $\theta = (\lambda_n, g_n, \mu_n, \sigma_n^2)$ are unknown apriori

$$P(\theta|\hat{S}_n) \propto P(\hat{S}_n|\theta)P(\theta)$$

We can just computed the likelihood $P(\hat{S}_n|\theta)$ on the last slide. Samples from the posterior can be found for example by MCMC or we could use MAP estimation

Either of these approaches only make sense for stationary statistics, which means the physical locations and photophysics of the sample remain unchanged in time

For example photostable fluorophores like quantum dots would be a good choice

Fisher Information and the Cramer-Rao Bound

Consider the general prescripton of maxmimum likelihood parameter estimation:

$$\mathcal{E}_{\mathrm{MLE}} : \theta^* = \underset{\theta}{\operatorname{argmax}}\, \ell(\mathcal{D}|\theta)$$

where $\ell = \log \mathcal{L}$ is the log-likelihood function

Question: can we derive a theoretical lower bound on our uncertainty in θ^* for an arbitrary estimator $\mathcal{E}$?

Start by defining the *score* of ℓ with respect to θ as

$$\mathcal{S} = \underset{x \sim p}{\mathbb{E}} \left[\frac{\partial}{\partial \theta} \ell(x|\theta) \right]$$

Since x is a continuous random variable, we have to consider the average score

Fisher Information for a single parameter

The Fisher Information $I(\theta)$ is defined as the variance of the score

$$I(\theta) = \mathop{\mathbb{E}}_{x \sim p}\left[\frac{\partial}{\partial\theta}\left(\ell(x|\theta)\right)\right]^2 = \mathop{\mathbb{E}}_{x \sim p}\left[\frac{\partial^2}{\partial\theta^2}\left(\ell(x|\theta)\right)\right]$$

for $x \in \mathcal{D}$. The variance takes this from because it can be shown that $\mathcal{S} = 0$

Intuitively, if the likelihood is insensitive changes in θ, then $\mathcal{D}$ does not provide very much information about θ

The Cramer-Rao Bound places a lower bound on the variance in our parameter estimate in iterms of $I(\theta)$:

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{I(\theta)}$$

Fisher Information for a multiple parameters

When there are many parameters, the Fisher Information (second moment of the score) is a covariance matrix

$$I_{ij}(\theta) = \mathop{\mathbb{E}}_{x\sim p}\left[\frac{\partial}{\partial\theta_i}\left(\ell(x|\theta)\right)\frac{\partial}{\partial\theta_j}\left(\ell(x|\theta)\right)\right]$$

We are going to consider the case where $\mathcal{L}$ is a multivariate Gaussian distribution:

$$\mathcal{L}(x|\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T\right)$$

We have $\theta = (\mu, \Sigma)$ so we need to compute $\frac{\partial\ell}{\partial\mu}$ and $\frac{\partial\ell}{\partial\Sigma}$

Fisher Information for a multiple parameters

Using our definition of the Fisher information:

$$\begin{aligned}\mathbf{I}(\theta) &= \mathop{\mathbb{E}}_{x\sim p}\begin{pmatrix}\frac{\partial^2\ell}{\partial\mu^2} & \frac{\partial\ell}{\partial\Sigma}\frac{\partial\ell}{\partial\mu}\\ \frac{\partial\ell}{\partial\mu}\frac{\partial\ell}{\partial\Sigma} & \frac{\partial^2\ell}{\partial\Sigma^2}\end{pmatrix}\\ &= \mathop{\mathbb{E}}_{x\sim p}\left[\mathbf{H}_\ell\,\ell(x|\mu,\Sigma)\right]\end{aligned}$$

where $\mathbf{H}_\ell$ is the Hessian of the log-likelihood. We note that $\mathbf{I}(\theta)$ is essentially the Hessian of the cross-entropy

Thus we need to evaluate $\mathbf{H}_\ell$ explicitly

Fisher Information for a multiple parameters

The following derivatives can be shown using matrix calculus:

$$
\begin{aligned}
\frac{\partial \ell}{\partial \mu} &= \Sigma^{-1}(x-\mu) \\
\frac{\partial \ell}{\partial \Sigma} &= -\frac{1}{2}\left(\Sigma^{-1} - \Sigma^{-1}(x-\mu)(x-\mu)^{T}\Sigma^{-1}\right)
\end{aligned}
$$

References I