Bayesian Inference for Neighborhood Filters with Application in Denoising Supplementary Material

Chao-Tsung Huang
National Tsing Hua University, Taiwan
chaotsung@ee.nthu.edu.tw

A. Proof of EM and KLD updates (22), (23), and (24)

For the function $\log N(\epsilon,\alpha)$, we have the partial derivatives

$$\frac{\partial \log N}{\partial \alpha} = \frac{\int_\epsilon^1 w(1-\log w)w^{-k/2}w^{-\alpha w}e^{\alpha w}\,dw}{N(\epsilon,\alpha)} = \mathbf{E}_{w;\alpha,\epsilon}[w(1-\log w)] = H(\alpha,\epsilon), \quad \text{(A-1)}$$

$$\frac{\partial \log N}{\partial \epsilon} = -\frac{\epsilon^{-k/2}\epsilon^{-\alpha\epsilon}e^{\alpha\epsilon}}{N(\epsilon,\alpha)}. \quad \text{(A-2)}$$

Then for the EM likelihood function

$$\mathcal{L}_{EM} = \log f_{s,w}(s,w;\hat\sigma,\epsilon,\hat\alpha) \quad \text{(A-3)}$$

$$= (k-1)\log s - k\log\hat\sigma - \frac{w}{w+1}\frac{s^2}{2\hat\sigma^2} - \hat\alpha w\log w + \hat\alpha w - \log N(\epsilon,\hat\alpha) - \frac{k}{2}\log(w+1) - \log 2^{\frac{k}{2}-1}\Gamma(k/2), \quad \text{(A-4)}$$

we have the partial derivatives

$$\frac{\partial \mathcal{L}_{EM}}{\partial \hat\sigma} = \frac{1}{\hat\sigma^3}\left(\frac{ws^2}{w+1} - k\hat\sigma^2\right), \quad \text{(A-5)}$$

$$\frac{\partial \mathcal{L}_{EM}}{\partial \hat\alpha} = w(1-\log w) - H(\hat\alpha,\epsilon). \quad \text{(A-6)}$$

Thus we can have the EM update (22) for $\hat\sigma$ by inserting (A-5) into

$$\frac{\partial Q}{\partial \hat\sigma} = \sum_j P(s_j)\frac{\partial q(s_j,\hat\sigma,\hat\alpha|\sigma,\alpha)}{\partial\hat\sigma}, = \sum_j P(s_j)\int_\epsilon^1 \frac{f_{s,w}(s_j,w)}{f_s(s_j)}\frac{\partial \mathcal{L}_{EM}}{\partial\hat\sigma}\,dw = 0. \quad \text{(A-7)}$$

Similarly, with $\frac{\partial Q}{\partial\hat\alpha} = 0$ and (A-6) we can have the EM update (23) for $\hat\alpha$.
As for the KLD update, we first let

$$h(s,w;\hat\sigma,\hat\alpha) = \frac{s^{k-1}\hat\sigma^{-k}}{(w+1)^{\frac{k}{2}}}e^{-\frac{w}{w+1}\frac{s^2}{2\hat\sigma^2}}w^{-\hat\alpha w}e^{\hat\alpha w} \quad \text{(A-8)}$$

such that

$$f_s(s;\hat\sigma,\epsilon,\hat\alpha) = \frac{\int_\epsilon^1 h(s,w;\hat\sigma,\hat\alpha)\,dw}{T(\epsilon,\hat\alpha)}. \quad \text{(A-9)}$$

With $\frac{\partial \log T}{\partial\epsilon} = \frac{\partial \log N}{\partial \epsilon}$, we can have

$$\frac{\partial \log f_s(s;\hat\sigma,\epsilon,\hat\alpha)}{\partial\epsilon} = \frac{-h(s,\epsilon;\hat\sigma,\hat\alpha)}{\int_\epsilon^1 h(s,\epsilon;\hat\sigma,\hat\alpha)\,dw} - \frac{\partial \log N}{\partial\epsilon}. \quad \text{(A-10)}$$

Then we can minimize $\mathcal{D}$ by $\frac{\partial\mathcal{D}}{\partial\epsilon} = 0$ and substitute (A-10), (A-9), and (A-2) to have:

$$\frac{\epsilon^{-k/2}\epsilon^{-\hat\alpha\epsilon}e^{\hat\alpha\epsilon}}{N(\epsilon,\hat\alpha)} = \sum_j P(s_j)\frac{h(s_j,\epsilon;\hat\sigma,\hat\alpha)}{f_s(s_j;\hat\sigma,\epsilon,\hat\alpha)T(\epsilon,\hat\alpha)}. \quad \text{(A-11)}$$

Thus the solution of ϵ becomes a fixed-point expression given the definition of $h(\cdot,\cdot)$ in (A-8)

$$\left(\frac{\epsilon+1}{\epsilon}\right)^{\frac{k}{2}} = \sum_j P(s_j)\cdot\frac{s_j^{k-1}\hat\sigma^{-k}e^{-\frac{\epsilon}{\epsilon+1}\frac{s_j^2}{2\hat\sigma^2}}}{2^{\frac{k}{2}-1}\Gamma(\frac{k}{2})f_s(s_j;\hat\sigma,\epsilon,\hat\alpha)}. \quad \text{(A-12)}$$

The update formulation for $\hat\epsilon$ in (24) is the first-iteration estimation of the optimal ϵ using the initial condition $\hat\epsilon^{(0)} = \epsilon$ on the right hand side.

■