From Noise Modeling to Blind Image Denoising

Cheng Guan

July 10, 2018

1. LR-MoG Filter

In this section, the authors introduces the LR-MoG filter to recover low-rank signals from noisy signals contaminated by MoG noise. Let $X = [x_1, x_2, \cdots]$ be a $d \times n$ matrix and $x_i \in \mathbb{R}^d, i \in (1, 2, \cdots, n)$, be a noisy signal. X can be decomposed as shown in Eq. 1

$$X = \hat{X} + E \tag{1}$$

Here,$\hat{X} = \left[x_1, \hat{\cdots}, \hat{x_n}\right]$, where $\hat{x}_i$ is low-rank representable.$E = [e_1 \cdots e_n]$, where e_i is the noise on x_i and follows MoG distribution.LR-MoG is used to recover $\hat{X}$ from obesered X.In order to get the signal from noise, the authors introduce Bayesian approach by introduction priors on $\hat{X}$ and E. Then ths signal is recovered via inference.

1.1. Modeling Low-rank Component

Firstly, the authors model the low-rank component $\hat{X}$. Decomposing each vector $\hat{x_i}$, where $i \in \{1, \cdots, n\}$, with $\hat{x_i} = A\hat{y_i} + u$.A is a $d \times d$ matrix and u is the mean vector of $\{\hat{x_i}\}_{i=1}^n$.In the filed of dimensional reduction [3], some similar decomposition methods can be also found. The matrix $x\hat{X}$ can be decomposed as shown in Eq.2

$$\hat{X} = A\hat{Y}^{\mathrm{T}} + U \tag{2}$$

where $\hat{Y} = [\hat{y}_1 \cdots \hat{y}_n]^{\mathrm{T}}$ $U = [u \cdots u]$ and the column vectors is matrix $A\hat{Y}$ are zero mean. The rank(M) is the rank of a matrix M. Introduce the Gaussian prior on u as shown in Eq.3

$$u \sim \mathcal{N}(\cdot|u_0, K_0) \tag{3}$$

Here, $\mathcal{N}(\cdot|u_0, \sum)$ is a normal distribution with mean μ and covariance $\sum$. In order to model the low-rank property of $\hat{X}$, let matrix $A\hat{Y}^{\mathrm{T}}$ be low-rank representable, because rank(U) = 1 and rank$\left(\hat{X}\right)$ = rank$\left(A\hat{Y}^{\mathrm{T}} + U\right)$ $\leq$ rank $\left(A\hat{Y} + 1\right)$.

To model the low-rank property of $A\hat{Y}^{\mathrm{T}}$, the authors introduce the trace-norm prior [1, 2]and it can be well approximated by the ARD[27]. And the authors adopt the soft trace-norm prior on $A\hat{Y}^{\mathrm{T}}$ as shown in Eq.4 and Eq.5

$$p(A) \propto \exp\left(-\frac{1}{2}\mathrm{tr}\left(AC_A^{-1}A^{\mathrm{T}}\right)\right) \tag{4}$$

$$p\left(\hat{Y}\right) \propto \exp\left(-\frac{1}{2}\mathrm{tr}\left(\hat{Y}C_{\hat{Y}}^{-1}A^{\mathrm{T}}\right)\right) \tag{5}$$

where $\mathrm{tr}(\cdot)$ represents the trace of a matrix. C_A and $C_{\hat{Y}}$ are set to be diagonal positive semidefinite with Eq.6 and Eq.7

$$C_A = diag_d\left\{c_{a_1}^2, \cdots, c_{a_d}^2\right\} \tag{6}$$

$$C_{\hat{Y}} = diag_d\left\{c_{y_1}^2, \cdots, c_{y_d}^2\right\} \tag{7}$$

Here, diag $diag_d\left\{c_{a_1}^2, \cdots, c_{a_d}^2\right\}$ represents a $d \times d$ diagonal matrix with diagonal items $c_1^2, \cdots, c_n^2$. Let $\hat{a}_j$ be the jth row vector of matrix A, and derive the specific formula of Eq.5 with Eq.8

$$\begin{aligned} p(A) &\propto \exp\left(-\frac{1}{2}\mathrm{tr}\left(AC_A^{-1}A^{\mathrm{T}}\right)\right) \\ &\propto \prod_{j=1}^{d} \exp\left(-\frac{1}{2c_{a_j}^2}\hat{a}_j^{\mathrm{T}}\hat{a}_j\right) \\ &\propto \prod_{j=1}^{d} \mathcal{N}\left(\hat{a}_j|0, c_{a_j}^2 I\right) \end{aligned} \tag{8}$$

2. Modeling MoG Noise

To model the complex noise on practical noisy images, the authors use MoG distribution for noise modeling. Because the number of components is not provided, tehy use Bayesian nonparametric technique and introduce the Dirichlet pro- cess prior to the MoG. Each Gaussian component is repre- sented by $\mathcal{N}(\cdot|\mu_i, \sum_i)$ with mean vector μ_i and covariancematrix $\sum_i$.

References

[1] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM*, 2011. 1

[2] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 2009. 1

[3] N. M. Nasrabadi. Pattern recognition and machine learning. *Journal of Electronic Imaging*, 2007. 1