Deep K-SVD Denoising

Scetbon Elad Milanfar

Presented by Yao Zhang

The guy is a populace

https://zhims.github.io/

June 23, 2020

Notation

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} & \cdots & \cdots & d_{1m} \\ d_{21} & d_{22} & d_{23} & \cdots & \cdots & d_{2m} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ d_{p1} & d_{p2} & d_{p3} & \cdots & \cdots & d_{pm} \end{bmatrix} \in \mathbb{R}^{p \times m} \tag{1}$$

Assume that $Rank(D) = s \ll p$ in other word, D is redundant.

Notation

x stands for a patch,

$$x \in \mathbb{R}^{\sqrt{p} \times \sqrt{p}} \overset{reshape\ by\ column}{\rightarrow} x \in \mathbb{R}^{p} \quad (2)$$

For example, *if* $p = 625$ *then* $\sqrt{p} = 25$

$$x = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,25} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,25} \\ \vdots & \vdots & \ddots & \vdots \\ x_{25,1} & x_{25,2} & \cdots & x_{25,25} \end{bmatrix} \in \mathbb{R}^{25 \times 25} \overset{reshape\ by\ column}{\rightarrow} \begin{bmatrix} x_{1,1} \\ \vdots \\ x_{25,1} \\ x_{1,2} \\ \vdots \\ x_{25,2} \\ \vdots \\ x_{1,25} \\ \vdots \\ x_{25,25} \end{bmatrix} \in \mathbb{R}^{625 \times 1} \quad (3)$$

Notation

X: clean image, Y: noisy image.

$$X, Y \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}} \overset{reshape\ by\ column}{\rightarrow} X, Y \in \mathbb{R}^{N\times 1} \tag{4}$$

K-SVD Model

$$\min_{\{\alpha_k\},X,D} \frac{\mu}{2}\|X-Y\|_2^2 + \sum_k \left(\lambda_k\|\alpha_k\|_0 + \frac{1}{2}\|D\alpha_k - R_k X\|_2^2\right) \qquad (5)$$

where R_k self-representation coefficient matrix.

1. Sparse coding

2. Diction learning

Sparse Coding (just a patch)

Assume that

$$x = D\alpha \tag{6}$$

where $x \in \mathbb{R}^{p\times 1}$, $D \in \mathbb{R}^{p\times m}$ and $\alpha \in \mathbb{R}^{m\times 1}$.

$$\alpha^* = \arg\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \|D\alpha - y\|_2^2 \leqslant p\sigma^2 \quad where \;\; y \in \mathbb{R}^{p\times 1} \tag{7}$$

Eq.7 can be written as

$$\alpha^* = \arg\min_{\alpha} \lambda\|\alpha\|_0 + \frac{1}{2} \|D\alpha - y\|_2^2 \tag{8}$$

Sparse Coding (just a patch)

Eq.8 can be soft-convex as

$$\alpha^* = \arg\min_{\alpha} \lambda\|\alpha\|_1 + \frac{1}{2}\|D\alpha - y\|_2^2 \tag{9}$$

Eq.9 can be solved via

1. BP
2. OMP
3. SVT

For example,

$$\min \ \lambda\|\alpha\|_1 + \frac{1}{2}\|D\beta - x\|_2^2 + \langle Y_1, \alpha - \beta\rangle + \frac{\gamma}{2}\|\alpha - \beta\|_2^2 \tag{10}$$

where $Y \in \mathbb{R}^{m\times 1}$

Sparse Coding (all patches)

It is easily by addition.

Update X

$$X^* = \arg\min_X \frac{\mu}{2}\|X - Y\|_2^2 + \frac{1}{2}\sum_k \|D\alpha_k^* - R_k X\|_2^2 \tag{11}$$

Eq.11 can be easily obtain via

$$X^* = \left(\sum_k (R_k^T R_k + \mu I)\right)^{-1}\left(\mu Y + \sum_k R_k^T D\alpha_k^*\right) \tag{12}$$

Remark 1

The form of formula (12) is very important.

Dictionary Learning

$$D^* = \arg\min_{D} \sum_k \|D\alpha_k - R_k X\|_2^2 \tag{13}$$

Eq 13 can solved by SVD, but the sparsity constraint is not enforced via SVD. Therefore, Elad et al. presented the K-SVD (2006) method for Eq 13.

https://sites.fas.harvard.edu/~cs278/papers/ksvd.pdf

https://en.wikipedia.org/wiki/K-SVD#:~:text=In%20applied%20mathematics%2C%20K-SVD,a%20singular%20value%20decomposition%20approach.

Deep K-SVD

1. Patch denoising

2. Reconstruct the full image

Sparse Coding

In Eq.9

$$\alpha^* = \arg\min_{\alpha} \lambda \|\alpha\|_1 + \frac{1}{2} \|D\alpha - y\|_2^2$$

λ, D is fixed in the traditional methods,respectively.

But λ, D are evaluated by MLP with 3 hidden layers here.

The solution to Eq.9 is

$$\alpha_{t+1} = \mathcal{S}_{\lambda/c}\left(\alpha_t - \frac{1}{c} D^T (D\alpha_t - y)\right) \tag{14}$$

where c is the square spectral norm of D.

The Loss Function

The loss function:

$$\|x_k - D\alpha_k\|_2^2 \tag{15}$$

By Eq.14,

$$\left\| x_k - D\mathcal{S}_{\lambda_k/c}\left[\frac{1}{c}D^T\left(D\alpha_k - y_k\right)\right] \right\|_2^2 \tag{16}$$

Therefore, our goal is

$$\underset{D,c,\lambda_k}{\arg\min} \sum_k \left\| x_k - \underbrace{D\mathcal{S}_{\lambda_k/c}\left[\frac{1}{c}D^T\left(D\alpha_k - y_k\right)\right]}_{\textit{informally denote } F_\theta(y_k)} \right\|_2^2 \tag{17}$$

Reconstruct the Full Image

$$X = \frac{\sum_k R_k^T (w \odot (x_k))}{\sum_k R_k^T w} \tag{18}$$

Remark 2

w is computed directly depend on λ_k and D, not from the network.

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.64.2767&rep=rep1&type=pdf

Architecture of Deep K-SVD

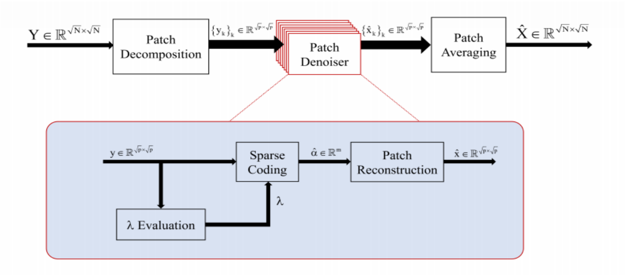

Figure 1: Architecture of the Deep K-SVD

https://elad.cs.technion.ac.il/wp-content/uploads/2019/09/Deep-KSVD.pdf

In the Future

...

References

M. Scetbon, M. Elad and P. Milanfar(2019)
Deep K-SVD Denoising
submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence

M. Aharon, M. Elad, and A. Bruckstein (2006)
The K-SVD: An Algorithm for Designing of Overcomplete Dictionaries for Sparse Representation
IEEE Transactions On Signal Processing Vol54(11), 4311–4322

O. Guleryuz (2007)
Weighted Averaging for Denoising With Overcomplete Dictionaries
IEEE Transactions on Image Processing Vol.16(12), 3020 – 3034

M. Elad(2010)
Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing
Springer