Generalized Tensor Total Variation Minimization for Visual Data Recovery

Xiaojie Guo[1], Yi Ma[2]
[1]SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences. [2]School of Information Science and Technology, ShanghaiTech University.

In real data analysis applications, we often have to face handling dirty observations, say incomplete or noisy data. Recovering the missing or noise-free data from such observations thus becomes crucial to provide us more precise information to refer to. Besides, compared to 1-D vectors and 2-D matrices, the data encountered in real world applications is more likely to be high order, for instance a multi-spectral image is a 3-D tensor and a color video is a 4-D tensor. This work concentrates on the problem of visual data recovery, *i.e.* restoring tensors of visual data from polluted observations.

Suppose we have a matrix $A \in \mathbb{R}^{D_1 \times D_2}$, the response of A to a directional derivative-like filter f_{θ_*} can be computed by $f_{\theta_*} * A$. The traditional TV norm takes into account only the responses to derivative filters along fibers. In other words, it potentially ignores important details from other directions. One may wonder if the multi-directional response can be represented by the gradients. Indeed, for differentiable functions, the directional derivatives along some directions have an equivalent relationship with the gradients. However, for tensors of visual data, this relationship no longer holds as the differentiability is violated. Based on this fact, we here propose a definition of the responses of a matrix to m-directional derivative-like filters as follows:

Definition 1. *(RMDF: Response to Multi-directional Derivative-like Filters.) The response of a matrix to m derivative-like filters in θ_m directions with weights β_m is defined as $\mathfrak{R}(A, \beta) \in \mathbb{R}^{D_1 D_2 \times m} :=$*

$$\left[\beta_1 \operatorname{vec}(f_{\theta_1} * A) \middle| \beta_2 \operatorname{vec}(f_{\theta_2} * A) \middle| \cdots \middle| \beta_m \operatorname{vec}(f_{\theta_m} * A)\right],$$

where $\beta = [\beta_1, \beta_2, \cdots, \beta_m]$ with $\forall j \in [1, ..., m]$ $\beta_j \geq 0$ and $\sum_{j=1}^{m} \beta_j = 1$.

Another drawback of the traditional TV is its homogeneity to all elements in tensors, which favors piecewise constant solutions. This property would result in oversmoothing high-frequency signals and introducing staircase effects. Intuitively, in visual data, the high-frequency signals should be preserved to maintain the perceptual details, while the low-frequency ones could be smoothed to suppress noises. This intuition inspires us to differently treat the variations. Considering both the multi-directionality and the inhomogeneity gives the definition of generalized tensor TV as:

Definition 2. *(GTV: Generalized Tensor Total Variation Norm.) The GTV norm of an n-order tensor $\mathcal{A} \in \mathbb{R}^{D_1 \times \cdots \times D_n}$ is:*

$$\|\mathcal{A}\|_{GTV} := \sum_{k=1}^{n} \alpha^k \|W^k \odot \mathfrak{R}(A_{[k]}, \beta^k)\|_p,$$

where $\alpha = [\alpha^1, ..., \alpha^k, ..., \alpha^n]$ is the non-negative coefficient balancing the importance of k-mode matrices and satisfying $\sum_{k=1}^{n} \alpha^k = 1$, and p could be either 1 or 2,1 corresponding to the anisotropic total variation (ℓ^1) and the isotropic one ($\ell^{2,1}$), respectively. In addition, $W^k \in \mathbb{R}^{\prod_{i=1}^{n} D_i \times m}$ acts as the non-negative weight matrix, the elements of which correspond to those of $\mathfrak{R}(A_{[k]}, \beta^k)$.

It is apparent that GTV satisfies the properties that a norm should do, and the traditional TV norm is a specific case of GTV. With the definition of GTV, the visual data recovery problem can be naturally formulated as:

$$\begin{aligned} \underset{\mathcal{T},\mathcal{N}}{\operatorname{argmin}} \sum_{k=1}^{n} \alpha^k \|W^k \odot \mathfrak{R}(\mathcal{T}_{[k]}, \beta^k)\|_p + \lambda \Psi(\mathcal{N}) \\ \text{s.t. } \mathcal{P}_\Omega(\mathcal{O}) = \mathcal{P}_\Omega(\mathcal{T} + \mathcal{N}). \end{aligned} \tag{1}$$

To solve the associated optimization problem, we design an effective and efficient algorithm based on ALM-ADM strategy [3, 4]. The detailed algorithm and some important properties can be found in the paper.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

Figure 1: **Top row** shows the results with 30% information missed. **Middle** and **Bottom** correspond to those with 50% and 70% elements missed, respectively. **From Left to Right:** input frames, recoverd results by STDC, Ha LRTC and GTV, respectively.

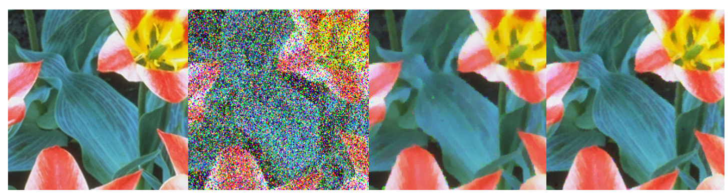

Figure 2: **Left:** Original image. **Mid-Left:** Polluted image by 40% Salt & Pepper noise (PSNR/SSIM: 8.71dB/0.1488). **Rest:** Recovered results by traditional TV (27.86dB/0.9144) and GTV (29.08dB/0.9234), respectively.

The proposed GTV can be widely applied to many visual data restoration tasks, such as visual data completion, denoising, and inpainting. We provide some experimental results here to show the advantages of GTV compared with the state-of-the-arts. Figure 1 shows an example on visual data completion, in terms of PSNR and SSIM, our method significantly outperforms STDC [2] and HaLRTC [5]. Figure 2 provides an example on image denoising to demonstrate the superior performance of GTV over the traditional TV [1, 6]. The middle-right picture in Fig. 2 is the best possible result of the traditional TV by tuning $\lambda \in \{0.1, 0.2, ..., 1.0\}$, while the right is automatically obtained by our method.

[1] S. Chan, R. Khoshabeh, K. Gibson, P. Gill, and T. Nguyen. An augmented lagrangian method for total variation video restoration. *TIP*, 20 (11):3097–3111, 2011.

[2] Y. Chen, C. Hsu, and H. Liao. Simultaneous tensor decomposition and completion using factor priors. *TPAMI*, 36(3):577–591, 2014.

[3] M. Hong and Z. Luo. On the linear convergence of the alternating direction method of multipliers. *arXiv preprint arXiv:1208.3922*, 2013.

[4] Z. Lin, R. Liu, and Z. Su. Linearized alternating direction method with adaptive penalty for low rank representation. In *NIPS*, pages 612–620, 2011.

[5] J. Liu, P. Musialski, P. Wonka, and J. Ye. Tensor completion for estimating missing values in visual data. *TPAMI*, 35(1):208–220, 2013.

[6] S. Yang, J. Wang, W. Fan, X. Zhang, P. Wonka, and J. Ye. An efficient admm algorithm for multimensional anisotropic toal variation regularization problems. In *SIGKDD*, pages 641–649, 2013.