ADAPTIVE PARAMETER SELECTION FOR GRADIENT-SPARSE + LOW PATCH-RANK RECOVERY: APPLICATION TO IMAGE DECOMPOSITION

Antoine Guennec, Jf Aujol, Yann Traonmilin

ADAPTIVE PARAMETER SELECTION FOR GRADIENT-SPARSE + LOW PATCH-RANK RECOVERY: APPLICATION TO IMAGE DECOMPOSITION

Antoine Guennec, Jean-François Aujol, Yann Traonmilin

University of Bordeaux, Bordeaux INP, CNRS, IMB, UMR 5251, F-33400 Talence, France
firstname.lastname@math.u-bordeaux.fr

ABSTRACT

In this work, we are interested in gradient sparse + low patch-rank signal recovery for image structure-texture decomposition. We locally model the structure as gradient-sparse and the texture as of low patch-rank. Moreover, we propose a rule based upon theoretical results of sparse + low-rank matrix recovery in order to automatically tune our model depending on the local content and we numerically validate this proposition.

***Index Terms*—** Image decomposition, texture, optimization, gradient-sparsity, low patch-rank

1. INTRODUCTION

The problem of decomposing an image into a structure/cartoon component and a texture component has been a longstanding area of research, with an extensive number of applications such as image restoration, pattern recognition, deraining and astromical imaging. The problem is typically given as follows: given an image f, find a decomposition $f = u + v$ such that u is a is piecewise-smooth approximation of the image, called the structure or cartoon component, and v is zero-mean, oscillating and locally patterned, called the texture component. The problem is notoriously ill-posed as there are twice more unknowns than known variables and it is difficult to define with precision what should be included in the structural component and in the texture component.

The problem of structure-texture decomposition (STD) is often formulated as an optimization problem of the form

$$\min_{\substack{u,v\\ f=u+v}} \mu R_1(u) + \gamma R_2(v), \tag{1}$$

where μ and γ are the tuning parameters. For the characterization of the structure component, most models use the total variation [1] since it characterizes well the piecewise smoothness of the structure component. It is defined by $R_1(u) = \|u\|_{TV} = \|\nabla u\|_1$. While the choice of the characterization of the structural component has remained relatively unchanged across most proposed STD models, many options exist when it comes to the texture component. Stemming from sparse decomposition methods that uses dictionary learning techniques, texture regularizers based upon the nuclear norm have gained some popularity [2, 3, 4].

Historically the first STD models, variational-based models typically use the total variation to characterize the structural component and use a functional space norm to constrain the textural component, such as the L2-norm [1], G-norm [5, 6], L1-norm [7] and $\mathcal{H}$-norm [8, 9]. While theoretically well-founded and able to capture the oscillating nature of texture, these norms are either difficult to implement or cannot capture textures with a small magnitude. In [4] the authors proposed a self-example and unsupervised learning approach where the STD functional is optimized using a neural network.

In the category of sparsity-prior and low-rank prior, the texture is considered to be sparsely represented in appropriate dictionary, which is either fixed or learned. One of the earliest approach on the subject was to consider that texture can be sparsely represented in a suitable given transformation (e.g discrete cosine transform (DCT), Gabor transform) [10, 11]. While very successful in some applications, the issue with this approach is that many textures that arise in practical applications cannot be modeled by DCT or other related dictionary. More recently, this approach was extended to use convolutional sparse coding instead [12], where convolutional filters are learned beforehand. However, this method is dependent on the resolution thus the learned convolutional filters should be trained accordingly.

The approach that we focus on in this paper is Schaeffer and Osher's low-patch rank (LPR) model [2] in which the texture is considered to be of low patch-rank. That is to say that given a patch map $\mathcal{P} : \mathbb{R}^{n\times m} \to \mathbb{R}^{p^2\times N}$ (with overlap), where $p \times p$ is the dimension of the patches and N the number of extracted patches from the image, rank$(\mathcal{P}(v))$ should be relatively small. Conceptually, this expresses the idea that patches of textures should reside in a common small vector space. However, as rank-minimization is well known to be a NP-hard problem, we minimize instead the nuclear norm, which is known to enforce low rank: $R_2(v) = \|\mathcal{P}(v)\|_* = \sum_i \sigma_i(\mathcal{P}(v))$. As such, the (LPR) model is written as

$$\min_{\substack{u,v\\ f=u+v}} \mu \|\nabla u\|_1 + \gamma \|\mathcal{P}(v)\|_*, \tag{2}$$

where μ and γ are the tuning parameters. This optimization problem can be solved using the *Alternating Direction Method of Multiplier* (ADMM) [13]. While the LPR model is more capable of extracting ideal textures from an image with well-paterned texture than previous models, the fact that it uses the nuclear norm to capture the low rank of patches of the texture globally is an issue, notably when the given image contains many different texture patterns. More recently, Ono et al. [3] proposed a blockwise low-rank texture model (BNN) to counteract against this issue with LPR.

However, while all the proposed methods up to now produce more or less acceptable results, they are all difficult to tune in order to achieve the required result since the tuning parameters can greatly vary between images in order to obtain the required STD.

Contributions: First, we present a localized version of the low patch-rank model. Second, using the fact that this model can be viewed as a gradient-sparse + low patch-rank interpretation of structure/texture decomposition, we propose a method to automatically tune model parameters based upon recent results from sparse+low-rank matrix recovery theory. Finally, we validate experimentally our proposed method.

2. LOCAL LPR DECOMPOSITION WITH AUTOMATIC PARAMETER SELECTION

2.1. Localized LPR

In order to add localization of texture into the LPR model, we introduce the following subdivision with overlap of the image

$$\mathcal{Q}^o_{(n_1,m_1)}(f) = \begin{bmatrix} \mathcal{Q}(f)_{1,1} & \cdots & \mathcal{Q}(f)_{1,q_m} \\ \vdots & \ddots & \vdots \\ \mathcal{Q}(f)_{q_n,1} & \cdots & (Q)_{q_n,q_m}, \end{bmatrix}, \qquad (3)$$

where o is the size of the overlap between adjacent blocks, (n_1, m_1) is the dimensions of the subdivision blocks $\mathcal{Q}(f)_{i,j}$. To simplify notations, we set $f_{i,j} = \mathcal{Q}(f)_{i,j}$ (and similarly $u_{i,j}$, $v_{i,j}$). Our model can be written as

$$\min_{\substack{u,v \\ f=u+v}} \sum_{i,j=1}^{q_1,q_2} \mu_{i,j} \|u_{i,j}\|_{TV} + \gamma_{i,j} \|\mathcal{P}(v_{i,j})\|_*, \qquad (4)$$

where $\{\mu\}_{i,j=1}^{q_1,q_2}$ and $\{\gamma\}_{i,j=1}^{q_1,q_2}$ are the regularization parameters of the model. Furthermore, we set $\mathcal{Q}^{-1}$ as the reconstruction mapping from the subdivision, where we use interpolation between frames to reconstruct overlapping regions.

2.2. Gradient-Sparse + Low Patch rank recovery

With their increase in number, it is not possible anymore to set manually the local regularization parameters. We propose a novel method to automatically adjust the regularization parameters, which adapts to the local content in the image. The method is largely inspired from the problem of *sparse + low-rank* recovery of compressive sensing. Given a s-sparse matrix $\bar{u} \in \mathbb{R}^n$ such that $\|\bar{u}\|_0 = \#\{i,j \mid \bar{u}_{i,j} \neq 0\} \leq s$, a low-rank matrix $\bar{v} \in \mathbb{R}^n$ such that $\text{rank}(\bar{v}) \leq r$ and a linear map $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$, the aim of *sparse + low-rank* recovery is to recover the couple $(\bar{u}, \bar{v})$ from a measurement $b = \mathcal{A}(\bar{u} + \bar{v})$. Recent works [14, 15] have shown that the couple $(\bar{u}, \bar{v})$ can be recovered under some conditions via the minimization problem:

$$\min_{\substack{u,v \\ b=\mathcal{A}(u+v)}} \mu \|u\|_1 + \gamma \|v\|_* \qquad (5)$$

when $\frac{\mu}{\gamma} = c\frac{\sqrt{r}}{\sqrt{s}}$, where c is a specified constant.

In [16], Chandrasekaran et al showed that the *sparse+low rank* matrix recovery problem could be solved with $\mathcal{A}{=}Id$ if the support space of the sparse component and the row+column space of the low-rank component are disjoint. Tanner and Vary's work [14] on the subject showed that if the incoherence between the sparse component and the low rank (LR) component (μ in the original paper) and the operator $\mathcal{A}$ verifies a restricted isometry property (i.e. $\mathcal{A}$ behaves almost like an isometry on sparse + LR objects with incoherence), then the couple $(\bar{u}, \bar{v})$ can be recovered from (5), with $\frac{\mu}{\gamma} = \sqrt{\frac{2r}{s}}$. In a more general context, when the unknown is the concatenation (u, v) of a sparse vector and a LR matrix, it was shown in [15] that, for operators having the restricted isometry property (without incoherence between components, thus mostly specializing to random observation operators), the choice of $\frac{\mu}{\gamma} = \sqrt{\frac{r}{s}}$ is optimal for the family of low rank+sparse problems. In Section 2.3, we propose to use this choice of parameters as a basis for automatic parameter selection in our localized LPR model. Indeed, in the LPR model the texture is interpreted as being of *low patch rank* and the structural component is constrained by the total variation which forces it to be *gradient-sparse*. Hence, the LPR model is a *gradient-sparse + low-patch rank* recovery problem, with $\mathcal{A} = Id$.

To test the validity of this prior, we conducted numerical experiments using synthetic 64×64 images where we control the gradient sparsity of the structure and the patch-rank of the texture (Fig. 1). The gradient sparsity is controlled by drawing a circle with a chosen radius, whereas the texture was synthesized using Fourier, Gaussian or Hadamard basis in the patch space. We then plot the relative error of reconstruction of (2) with $\frac{\mu}{\gamma} = c\sqrt{\frac{s}{r}}$, where s is the gradient-sparsity of the structure and r the patch-rank of the texture. We observe that for a given category of texture, a fixed c permits to minimize the recovery error.

2.3. Local LPR with adaptive parameter tuning

Since the gradient-sparsity of the structure component and the patch-rank of the texture component cannot be distinguished beforehand, we propose to approximate them during the iterative steps of the ADMM and update the tuning parameters as

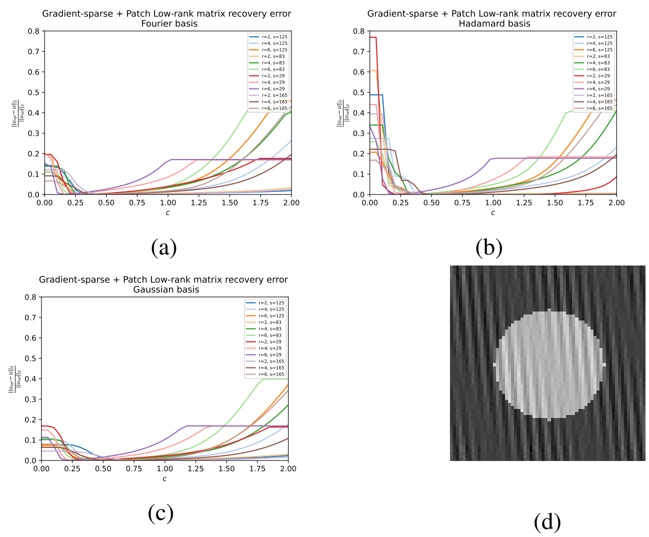

Fig. 1: Recovery relative error of the structure component with different gradient-sparsity s and patch-rank r via the LPR model with tuning parameters $\mu = \frac{c}{\sqrt{s}}$ and $\gamma = \frac{1}{\sqrt{r}}$. (a) Fourier texture, (b) Hadamard, (c) Gaussian basis, (d) Example of a synthetic test image.

follows:

$$\mu_{i,j} \approx \frac{c}{\sqrt{\|\nabla \tilde{u}_{i,j}\|_0}} \quad \text{and} \quad \gamma_{i,j} \approx \frac{1}{\sqrt{\text{rank}(\mathcal{P}(\tilde{v}_{i,j}))}}, \tag{6}$$

where $\tilde{u}$ and $\tilde{v}$ are an approximation of the structural component and textural components and c is a constant that we can set manually. In practice, the rank and gradient-sparsity are estimated by enumerating the top 90% gradients and patch-singular values, denoted by $\|\cdot\|_{0,est}$ and $rank(\cdot)_{est}$ respectively. While our experimental data (fig. 1) suggests setting $c \approx 0.4$ for the optimal decomposition, in practice we tend to overestimate the true underlying gradient-sparsity of the structure component and patch-rank of the texture component for natural image and thus c is set slightly higher in practice, e.g in the range $[0.6, 0.8]$. With this technique, we narrow down the number of required parameters to tune from $q_n \times q_m$ to a single parameter and each tuning parameter adapts to the local content in the image. In order to use the ADMM, we need to compute the proximal operator of the total variation, which can be quickly solved using FISTA [17] and the proximal operator of the nuclear norm, which is known as the singular value thresholding (SVT) operator [18]:

$$\text{prox}_{\|\cdot\|_*,\beta}(x) = SVT(x, \beta) = U \max(D - \beta I, 0) V^T, \tag{7}$$

where $x = UDV^T$ is the singular value decomposition of x [19] and the maximum is taken elementwise.

3. RESULTS AND DISCUSSION

We present some decomposition results (fig. 2), comparing our method to Low Patch Rank and Block Norm Normalization. The source code of our method can be found in the git

Algorithm 1 (Our proposed method)

$u^0 = f, v^0 = 0, y^0 = 0$
$\mu^0_{i,j} = \frac{c}{\sqrt{\|\nabla f_{i,j}\|_0}}, \gamma^0_{i,j} = \frac{1}{p}$
while not converged **do**
  $u^{k+1}_{i,j} = \text{prox}_{TV, \frac{\mu^k_{i,j}}{\rho}}(\mathcal{Q}(f - v^k - y^k)_{i,j})$
  $v^{k+1}_{i,j} = \mathcal{P}^{-1}(\text{SVT}(\mathcal{P}(f - u^{k+1} - y^k)_{i,j}, \frac{\gamma^k_{i,j}}{\rho}))$
  $y^{k+1} = y^k + (u^{k+1} + v^{k+1} - f)$
  if $k = 0$ mod M **then**
    $\mu^{k+1}_{i,j} = \frac{c}{\sqrt{\|\nabla u^{k+1}_{i,j}\|_{0,est}}}$
    $\gamma^{k+1}_{i,j} = \frac{1}{\sqrt{\text{rank}(v^{k+1}_{i,j})_{est}}}$
  end if
end while

repository [20]. In our experiment, the patch operator $\mathcal{P}$ was parameterized with a patch size $p = 5$ and the subdivision $\mathcal{Q}$ of the image was performed with $(n_1, m_1) = (64, 64)$ and an overlap $o = 16$. Finally, our tuning parameter was set to $c = 0.65$ for every image and in order to achieve comparative results with our method, we tuned the LPR and BNN models such that the output textural components are of similar magnitude by requiring $|\,\|v_{other}\|_2 - \|v_{proposed}\|_2\,| < 0.1$.

As seen in our results, our method achieves better than state of the art decompositions (the original LPR and BNN), i.e. for the same amount of texture, our structure component is sharper. While BNN captures well stripes-like textures, we observed that this tends to force a lot of structural details such as facial features into the textural component. Furthermore, compared with other methods, our method requires very little tuning since setting any $c \in [0.6, 0.8]$ achieves a good image decomposition for 512×512 images. To illustrate the robustness of our method: in Fig. 2, our method was performed with $c = 0.65$ whereas the other methods required changes of up to 60% and 53% in the parametrization for LPR and BNN, respectively. Furthermore, as our method is performed locally simultaneously, we can significantly accelerate the process by parallelizing the computing using a graphic processing unit.

However, there are still some unknowns which have yet to be investigated. From a theoretical point of view, the *gradient-sparse + low patch-rank* recovery problem has yet to be studied. Fundamentally, in the *sparse + low rank* recovery problem, sparse and LR matrices cannot be recovered by (5) when they are not sufficiently incoherent one to the other and it isn't yet clear how this translates on the type of structure and texture that can be recovered using the LPR model. Moreover, the convergence of the scheme we used to update our parameters should be explored further.

Up to now, tuning decomposition models has remained largely try-and-error and our tuning method could also benefit other sparsity/low rank prior based decomposition models such as BNN in order to reach an optimal decomposition.

Fig. 2: Comparison between different methods. From left to right: original image, LPR, BNN and our proposed method.

4. REFERENCES

[1] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[2] H. Schaeffer and S. Osher, "A low patch-rank interpretation of texture," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 226–262, 2013.

[3] S. Ono, T. Miyata, and I. Yamada, "Cartoon-texture image decomposition using blockwise low-rank texture characterization," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1128–1142, March 2014.

[4] F. Zhou, Q. Chen, B. Liu, and G. Qiu, "Structure and texture-aware image decomposition via training a neural network," *IEEE Transactions on Image Processing*, vol. 29, pp. 3458–3473, 2019.

[5] Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*, vol. 22, American Mathematical Soc., 2001.

[6] L. A. Vese and S. J. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *Journal of scientific computing*, vol. 19, pp. 553–572, 2003.

[7] S. Alliney, "A property of the minimum vectors of a regularizing functional defined by means of the absolute norm," *IEEE transactions on signal processing*, vol. 45, no. 4, pp. 913–917, 1997.

[8] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition—modeling, algorithms, and parameter selection," *International journal of computer vision*, vol. 67, pp. 111–136, 2006.

[9] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, "Image decomposition into a bounded variation component and an oscillating component," *Journal of Mathematical Imaging and Vision*, vol. 22, no. 1, pp. 71–88, 2005.

[10] J.-L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE transactions on image processing*, vol. 14, no. 10, pp. 1570–1582, 2005.

[11] M. J. Fadili, J.-L. Starck, J. Bobin, and Y. Moudden, "Image decomposition and separation using sparse representations: An overview," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 983–994, 2009.

[12] H. Zhang and V. M. Patel, "Convolutional sparse and low-rank coding-based image decomposition," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2121–2133, 2017.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[14] J. Tanner and S. Vary, "Compressed sensing of low-rank plus sparse matrices," *Applied and Computational Harmonic Analysis*, 2023.

[15] Y. Traonmilin, R. Gribonval, and S. Vaiter, "A theory of optimal convex regularization for low-dimensional recovery," *arXiv preprint arXiv:2112.03540*, 2022.

[16] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, "Rank-sparsity incoherence for matrix decomposition," *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 572–596, 2011.

[17] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, pp. 120–145, 2011.

[18] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[19] L. N. Trefethen and D. Bau, *Numerical linear algebra*, vol. 181, Siam, 2022.

[20] A. Guennec, J.-F. Aujol, and Y. Traonmilin, "Structure-texture-decomposition-gs-lpr," `https://github.com/aguennecjacq/Structure-Texture-Decomposition-GS-LPR`, 2023.