

## A Low Patch-Rank Interpretation of Texture\*

Hayden Schaeffer<sup>†</sup> and Stanley Osher<sup>†</sup>

**Abstract.** We propose a novel cartoon-texture separation model using a sparse low-rank decomposition. Our texture model connects the separate ideas of robust principal component analysis (PCA) [E. J. Candès, X. Li, Y. Ma, and J. Wright, *J. ACM*, 58 (2011), 11], nonlocal methods [A. Buades, B. Coll, and J.-M. Morel, *Multiscale Model. Simul.*, 4 (2005), pp. 490–530], [A. Buades, B. Coll, and J.-M. Morel, *Numer. Math.*, 105 (2006), pp. 1–34], [G. Gilboa and S. Osher, *Multiscale Model. Simul.*, 6 (2007), pp. 595–630], [G. Gilboa and S. Osher, *Multiscale Model. Simul.*, 7 (2008), pp. 1005–1028], and cartoon-texture decompositions in an interesting way, taking advantage of each of these methodologies. We define our texture norm using the nuclear norm applied to patches in the image, interpreting the texture patches to be low-rank. In particular, this norm is easier to implement than many of the weak function space norms in the literature and is computationally faster than nonlocal methods since there is no explicit weight function to compute. This norm is used as an additional regularizer in several image recovery models. Using total variation as the cartoon norm and our new texture norm, we solve the proposed variational problems using the split Bregman algorithm [T. Goldstein and S. Osher, *SIAM J. Imaging Sci.*, 2 (2009), pp. 323–343]. Since both of our regularizers are of  $L^1$  type, a double splitting provides a fast algorithm that is simple to implement. Based on experimental results, we demonstrate our algorithm’s success on a wide range of textures. Also, our particular cartoon-texture decomposition model has the advantage of separating noise from texture. Our proposed texture norm is shown to better reconstruct texture for other applications such as denoising, deblurring, sparse reconstruction, and pattern regularization.

**Key words.** texture, low-rank, optimization, sparse reconstruction

**AMS subject classifications.** 65K05, 49M27

**DOI.** 10.1137/110854989

**1. Introduction.** One of the most important problems in image processing is the recovery of a corrupted image,  $f$ , which may be degraded by noise, blur, missing data, etc. The goal is to reconstruct important structural features of the original image, such as large scale objects (smooth regions), edges (discontinuities), textures (patterned small scale details), and noise (random and of mean zero). This problem is typically written as an inverse problem: given  $f$ , find a  $u$  which is a smooth approximation of  $f$  in some sense [1]. Specifically, given an  $f$ , decompose  $f = u + v$ , where  $u$  is the recovered image and  $v$  is the residual assumed to be

\*Received by the editors November 11, 2011; accepted for publication (in revised form) September 13, 2012; published electronically February 12, 2013. This research was supported by grants ONR N00014-08-1-1119, ONR N00014-11-0719, NSF DMS-1118971, NSF DMS-0914561, and an ARO MURI subcontract from Rice University, and was supported by the Department of Defense (DoD) through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program.

<http://www.siam.org/journals/siims/6-1/85498.html>

<sup>†</sup>Department of Mathematics, University of California at Los Angeles, Los Angeles, CA 90095 ([hschaeffer@ucla.edu](mailto:hschaeffer@ucla.edu), [sjo@math.ucla.edu](mailto:sjo@math.ucla.edu)).

noise. For such two component decompositions, the general formulation is

$$\inf_{(u,v) \in X_1 \times X_2} \mathcal{E}(u,v) = \{\mu\|u\|_{X_1} + \|v\|_{X_2}\}$$

subject to (s.t.)  $f = u + v,$

where  $\mu > 0$  is a tuning parameter. The recovered image is assumed to reside in the space  $X_1$ , which contains functions of positive differentiability (i.e., the “derivative” is well behaved in some sense), and the noise is taken to be in  $X_2 = L^2$ . In this class of image recovery methods, notable models include those by Rudin, Osher, and Fatemi (ROF) [32] and Chambolle and Lions [11]. In particular, the ROF model is

$$(1.1) \quad \inf_{(u,v) \in BV \times L^2} \mathcal{E}_{ROF}(u,v) = \left\{ \mu\|u\|_{TV} + \frac{1}{2}\|v\|_{L^2}^2 \right\}$$

s.t.  $f = u + v,$

which is known to reconstruct piecewise constant solutions very well, thereby recovering both large scale features and edges. Recall that a function  $u$  is of bounded variation ( $BV$ ) if and only if  $u \in L^1$  and there exists a finite Radon measure  $Du$  such that total variation ( $TV$ )

$$(1.2) \quad \|u\|_{TV} := \int_{\Omega} |Du| := \left\{ \int_{\Omega} u \operatorname{div}(\phi) dx \mid \phi \in C_c^1(\Omega, (R)^n), \|\phi\|_{L^\infty} \leq 1 \right\}$$

is finite. When the function  $u \in W^{1,1}$  (defined as the space of functions  $u$  such that  $u, \nabla u \in L^1$ ), the  $TV$  seminorm becomes  $\|u\|_{TV} = \int_{\Omega} |\nabla u| dx$ , where  $\nabla u$  is the weak derivative. With the introduction of the Bregman technique [29] and the split Bregman method [20],  $TV$  regularized problems can be solved quickly and efficiently while also remedying defects such as loss of contrast.

On the other hand, the texture is considered to be highly oscillatory, which is not well captured by the  $X_1$  norms. Thus in many of these reconstruction models,  $v$  also contains some textures and edges. In order to further decompose the texture and noise, texture regularized models have appeared in the literature, beginning with the work of Meyer [28], who first proposed recovering texture using spaces that are weaker than  $L^2$ . These weak spaces encourage oscillatory behavior, since their norms decrease as the number of oscillations increases. In this framework, the image is decomposed into the cartoon component and the texture component, while also removing any noise or residual. This becomes a three component decomposition: given  $f$ , decompose  $f = u + v + \rho$  by solving

$$(1.3) \quad \inf_{(u,v) \in BV \times T} \mathcal{E}_{C-T}(u,v) = \mu\|u\|_{TV} + \gamma\|v\|_T + \frac{1}{2}\|u + v - f\|_{L^2}^2.$$

The cartoon component  $u$  is appropriately modeled by the space  $BV$ , using the  $TV$  seminorm in the energy (see (1.2)). The recovered image is now  $u + v$ , the sum of the cartoon and texture components. The residual,  $\rho$ , is assumed to reside in  $L^2$ , although other  $L^p$  norms can be used. Meyer proposed several texture norms in his work; most notably, the  $G$ -norm (the predual of  $BV$ , i.e.,  $G^* = BV$ ) has seen much success in applications including color

image restoration [5], road detection [19], denoising [18], inpainting [6], image classification [4], and anomaly removal and pattern regularization [17]. The space  $G$  is defined as follows.

**Definition 1.1.** *The space  $G$  consists of all distributions  $v$  which can be written as  $v = \operatorname{div}(\vec{g})$ , where  $\vec{g} = (g_1, g_2)$  and  $g_1, g_2 \in L^\infty$ . The norm on this space is defined as*

$$(1.4) \quad \|v\|_G := \inf \left\{ \left\| \sqrt{g_1^2 + g_2^2} \right\|_\infty \mid v = \operatorname{div}(\vec{g}) \right\}.$$

This texture space is particularly elegant because of its symmetry (specifically duality) with the cartoon space  $BV$ . However, this space is difficult to handle numerically. In [33], Vese and Osher proposed a method to approximate the  $G$ -norm by using the Sobolev spaces of negative differentiability, defined as  $W^{-1,p} := \{v = \operatorname{div}\vec{g} \mid g_1, g_2 \in L^p\}$ , and sending  $p \rightarrow \infty$ . Rather than working with the texture  $v$  directly, this method works with the vector-field  $\vec{g}$ :

$$(1.5) \quad \inf_{(u,g) \in BV \times (L^p)^2} \mathcal{E}_{OV}(u, g) = \mu \|u\|_{TV} + \|\vec{g}\|_{L^p} + \frac{\lambda}{2} \|u + \operatorname{div}\vec{g} - f\|_{L^2}^2.$$

This formulation yields appropriate Euler–Lagrange equations for all  $p \geq 1$ , which is easy to compute numerically. The results give satisfactory decompositions, particularly in the case of  $p = 1$ . Osher, Solé, and Vese [30] proposed an alternative norm for the case of  $p = 2$ , which coincides with the space  $H^{-1}$  (the dual of the Hilbert space  $H^1$ ). This dual space has an explicit norm defined as  $\|v\|_{H^{-1}} = \|\nabla \Delta^{-1} v\|_{L^2}$ , and the resulting Euler–Lagrange equation of the model reduces to a fourth order nonlinear partial differential equation. Later, Aujol et al. [3] used projections to solve the original  $BV$ – $G$  decomposition model and applied it to denoising of cartoon images (where the  $G$ -norm captures the noise) and to cartoon–texture separation. These weak spaces inspired many other texture norms, mainly in the form of dual spaces. In Lieu and Vese [24], the negative Hilbert spaces  $H^{-s}$  for  $s > 0$ , which are dual to the Hilbert–Sobolev spaces  $(H^s)^*$ , were used for texture extraction and noise removal. Later, in Kim and Vese [22], the negative exponential Sobolev spaces  $W^{\alpha,p}$  for  $-2 \leq \alpha < 0$ , which are dual to the Sobolev spaces with pseudoderivatives, were used for texture reconstruction in the presence of blur.

There are, in fact, an almost infinite number of texture norms that have been proposed, and this brief discussion just introduces this active field. In much of this work, the texture norms provide extra regularization on the recovered image, better recovering many of the features from the original image. However, these norms tend to be particularly difficult to implement in practice and do not deal well with noise.

Rather than considering the cartoon and texture components separately, the nonlocal methods simultaneously reconstruct these components. The nonlocal (NL) methods were first proposed by Buades, Coll, and Morel [7, 8] as a nonlocal filter and were later formulated in a variational framework by Gilboa and Osher [16, 17]. The general framework involves replacing local derivatives by their nonlocal counterpart:

$$(\nabla_w u)(x, y) := (u(y) - u(x)) \sqrt{w(x, y)}$$

for all  $x, y \in \Omega$ , where  $w(x, y) = e^{-d(u(x), u(y))}$ ,  $d(u(x), u(y)) = \int_\Omega G(t)|u(x+t) - u(y+t)|^2 dt$ , and  $G(t)$  is a Gaussian with appropriate parameters. In the discrete version, each pixel is associated

with a patch, and local differences are replaced by differences between pixels that have similar patches. Using these patch-based differences in the energy encourages repetitive behavior in the reconstructed image, thereby recovering the cartoon and texture while removing random noise. In particular, the nonlocal extension of the classical ROF model,

$$(1.6) \quad \inf_u \mathcal{E}(u) = \|\nabla_w u\|_{L^1} + \lambda \|u - f\|_{L^2}^2,$$

has been shown to recover the texture well. Although these models are useful, due to the global nature of images, they tend to be very slow in practice because of the frequent recomputation of the weight function  $w(x, y)$ , especially for tasks such as deblurring and sparse reconstruction.

In this paper, we will combine cartoon-texture decomposition with the powerful patch-based methods using a texture norm motivated by recent work called robust principal component analysis (PCA). Robust PCA was proposed by Candès et al. [10] in order to recover the sparse and low-rank parts of a given matrix  $f$ . This is also formulated as a decomposition problem: decompose  $f = u + v$ , where  $u$  is sparse and  $v$  is low-rank, by minimizing

$$\begin{aligned} \inf_{u,v} \mathcal{E}_{PCA}(u,v) &= \|u\|_{L^1} + \lambda \|v\|_* \\ \text{s.t. } f &= u + v, \end{aligned}$$

where  $\|\cdot\|_*$  is the nuclear (or trace) norm, which is the sum of the singular values. Because this method almost exactly recovers the original sparse and low-rank components, it is becoming increasingly popular in practice. In the past few years, robust PCA has been applied to video surveillance [10], face recognition [10], video denoising [21], alignment [31], and low matrix rank textures [35]. This variational model was later extended by Gao et al., who replaced the  $L^1$  norm with a tight frame regularization [14] and a  $TV$  regularization [15]. The PCA models also benefit from the split Bregman method, making their implementation both efficient and fast.

Similar sparse decomposition methods exist in the class of dictionary learning techniques. The main idea of this technique is to find a sparse representation for an image with respect to a (learned) redundant dictionary. Elad and Aharon [13] and Zhou et al. [36] both applied (different) Bayesian approaches in order to construct dictionaries composed of subparts of the given image. In [27], Mairal, Elad, and Sapiro extended the idea of sparse representation over a dictionary to color image restoration, in particular, denoising, inpainting, and demosaicing. The idea of sparse coding was later combined with the nonlocal methods in the work of Mairal et al. [26]. Similarly, in [12], the idea of structural clustering and dictionary learning was proposed and used for highly textured image restoration. Works such as [25] have also provided more efficient and less costly ways to implement these dictionary learning methods.

At the time of this work, the PCA-based methods have not been combined with cartoon-texture models, although they are of a very similar nature. Both PCA methods and cartoon-texture models decompose a given  $f$  into two main components, where one is “sparse” in some sense and the other is “patterned.” In this paper, we will connect the cartoon-texture models, robust PCA, and patch-based methods. The paper is divided as follows. In section 2, we derive and discuss our proposed texture norm and the proposed model. In section 3,

some theoretical remarks are made on the model, with characterization of the minimizers (for proofs, see Appendix B). In section 4, we describe the numerical implementation based on the split Bregman algorithm. Finally, in section 5, we detail the results.

**2. Description of the model.** The proposed model will be of a form similar to that of the classical cartoon-texture decomposition models. Given an  $f$ , we decompose  $f = A(u + v) + \rho$  by minimizing the following energy:

$$\inf_{u,v} \mathcal{E}(u,v) = \mu \|u\|_{TV} + \gamma \|v\|_{Texture} + \frac{\lambda}{2} \|A(u+v) - f\|_2^2,$$

where  $A$  is some degradation operator, for example, a reduction operator for missing data. In practice, we solve this minimization using a double Bregman splitting, which can enforce  $f = A(u + v)$  exactly or can be relaxed in the presence of noise. The cartoon norm we use is the (discrete) total variation defined as

$$\|u\|_{TV} := \|Du\|_1 = \sum_{i,j} |Du_{i,j}|,$$

where  $D = [D_x, D_y]$  is a differencing operator and  $|Du_{i,j}| = \sqrt{(D_x u_{i,j})^2 + (D_y u_{i,j})^2}$ . For the rest of this work, we take  $D$  to be the forward differences. Before going into more detail, we first define our texture norm.

**2.1. The texture norm.** From an intuitive perspective, the texture component is a global and well-patterned structure within a given image. The distinct patterns that make up the entire texture are called *base textures*. We expect the number of base textures to be low, since an image may exhibit only a few individual patterns. In particular, each patch (i.e., subblock) should be composed of a combination of these base textures. Therefore, the overall collection of patches can be spanned by a small set of base patches. If the patches are written as vectors, then the collection of *patch-vectors* are (highly) linearly dependent and thus have low rank. This is the key to our definition of texture and its norm. Using this idea we have the following definition for the collection of patches and the texture norm.

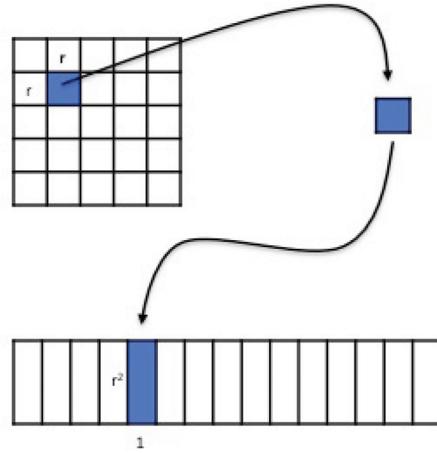
**Definition 2.1.** *The patch map  $\mathcal{P} : \mathbb{M}^{n,m} \rightarrow \mathbb{M}^{r^2, \frac{nm}{r^2}}$  is defined by the following: for  $v \in \mathbb{M}^{n,m}$ , partition  $v$  into  $r$  by  $r$  (nonoverlapping) submatrices, labeled  $\{B_i\}_{i=1}^{\frac{nm}{r^2}}$ . Next, transform each of the  $B_i$ 's into a column vector of length  $r^2$ , called  $w_i$ . Then augment the vectors together to form the new matrix*

$$\mathcal{P}v := \begin{bmatrix} w_1 & \cdots & w_{\frac{nm}{r^2}} \end{bmatrix}.$$

The specific ordering which maps the patch to a patch-vector is not important, as long as it is consistent (see Lemma 3.7). For the results in this paper, a raster scan (row by row) was used to arrange the patches, and a second raster scan was used to rewrite each patch as a vector.

Note that the patch map does not increase the number of terms since the patches are nonoverlapping. Also, the collection of texture patches is found by applying the patch map to the texture. Figure 1 depicts the patch map applied to a matrix. Using Definition 2.1 and the behavior we expect on the texture, the natural norm on  $\mathcal{P}v$  would be

$$\|v\|_T := \text{rank}(\mathcal{P}v).$$



**Figure 1.** Patch map.

However, this energy is nonconvex and difficult to use in practice. Furthermore, *rank* is not a norm in the mathematical sense. Using the ideas from robust PCA [10], we can replace the rank by the nuclear/trace norm, which is the convex envelope to the *rank* and is in fact a norm. Thus, the suggested texture norm can be relaxed to the following:

**Definition 2.2.** *A function  $v \in T$  if it is mean zero and if the following quantity is finite:*

$$\|v\|_T := \|\mathcal{P}v\|_*,$$

where  $\|\cdot\|_*$  is the nuclear/trace norm, i.e., the sum of the singular values. For a discrete function  $v$ , the norm  $\|v\|_T$  is always finite, but we use this definition for a more general  $v$  which will be addressed in future work.

Unlike the nonlocal methods, this norm does not explicitly calculate the weights between patches, but rather compares the patches implicitly. In this way, our nonlocal measure is computationally more efficient, while still being easy to compute and simple to minimize. To better understand Definition 2.2, here is an example of a texture and its norm.

**Example 2.3.** Let the texture be the zero mean vertical stripe pattern defined as

$$v_{i,j} = \begin{cases} 1, & j \text{ is even}, \\ -1 & \text{otherwise} \end{cases}$$

of size  $N$  by  $N$ . Then after applying the patch map with 2 by 2 patches, we have the following:

$$\mathcal{P}_{2 \times 2} v = \begin{bmatrix} -1 & -1 \\ 1 & 1 & \dots \\ -1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Similarly, after applying the patch map with 3 by 3 patches we have the following:

$$\mathcal{P}_{3 \times 3} v = \begin{bmatrix} -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & -1 & \cdots \\ -1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

In both cases the resulting matrices are of rank 1. It is easy to show that there is only one nonzero singular value equal to  $rN$ . Consequently,  $\|\mathcal{P}_{r \times r} v\|_* = rN$ . Of course,  $v$  has the same value in all  $L^p$  norms,  $\|v\|_{L^p} = N^2$ , and its total variation is  $\|v\|_{TV} = 2N^2$ . Since  $N$  is much larger than  $r$ , the norms are well ordered:  $\|\mathcal{P}_{r \times r} v\|_* \ll \|v\|_{TV} \ll \|v\|_{L^p}$ . This type of behavior is desired in cartoon-texture models.

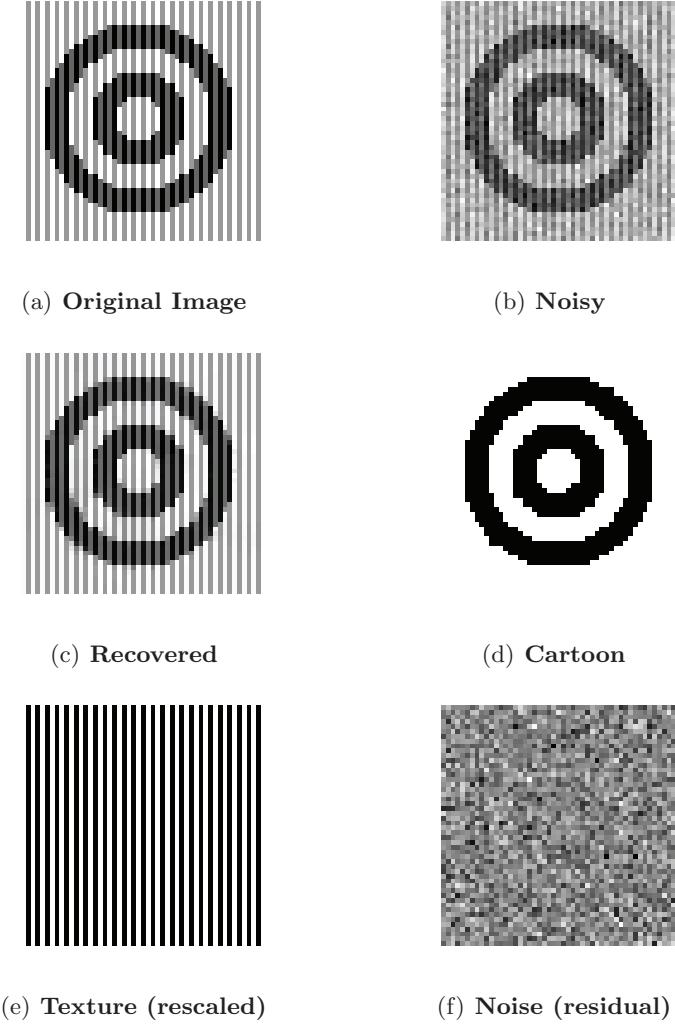
**2.2. Proposed model.** With Definition 2.2, our proposed model is as follows: given an  $f$ , we decompose  $f = A(u + v) + \rho$  by minimizing

$$(2.1) \quad \mathcal{E}(u, v) = \mu\|Du\|_1 + \gamma\|\mathcal{P}v\|_* + \frac{\lambda}{2}\|A(u + v) - f\|_2^2$$

for all  $(u, v) \in \mathcal{A}$ , where the admissible set is  $\mathcal{A} := \{(u, v) \mid \sum u = \sum f, \sum v = 0\}$ . In practice, the minimizers remain in this admissible set without any formal constraints. The operator  $A$  is assumed to be linear. For denoising,  $A$  is the identity; for deblurring, it is a convolution with a blur kernel; and for inpainting or sparse reconstruction, it is a reduction operator.

Recall that minimizing this energy ensures that the cartoon will be in discrete  $BV$ , thereby being piecewise smooth with sharp edges. The texture norm ensures a low patch-rank collection of textures and thus a small amount of repetitive textures. The residual (or noise) term  $\rho := A(u + v) - f$  remains in  $L^2$  since it is assumed to have no particular structure. In our examples, the noise is Gaussian; however, the residual term  $\frac{\lambda}{2}\|\rho\|_2^2$  can be replaced with other norms depending on the type of corruption. For example, impulse noise and blind inpainting (of small regions) is better captured by using the  $L^1$  norm, i.e.,  $\lambda\|\rho\|_1$ .

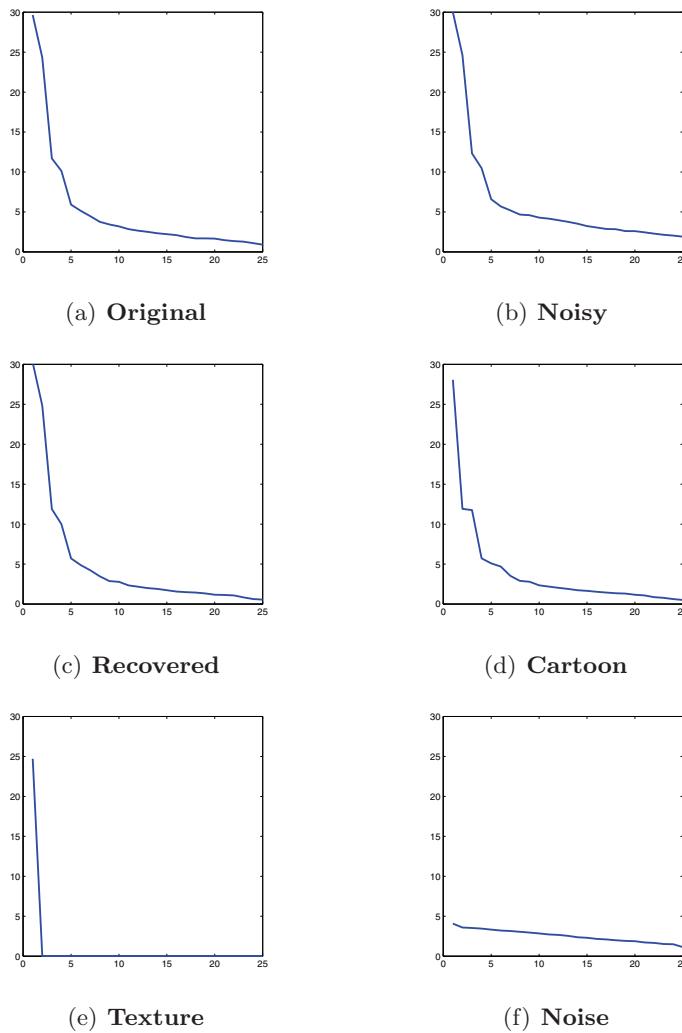
In Figure 2, we provide a simple decomposition example. The cartoon consists of concentric annuli, the texture is a repetitive stripe pattern, and the noise is random. In this example, the rank of the texture is 1, while the cartoon and noise components have full rank. It is worth noting that this decomposition is not equivalent to thresholding the input images' singular values. Figure 3 plots the singular values in descending order of each of the images from Figure 2 (after applying the patch map). The texture has only one singular value since it is composed of only one pattern. The cartoon has many patch patterns based on the various alignments of edges between the homogeneous regions, all with varying degrees of importance, with the most important being the constant patch. The noise has almost no coherent structures, which can be seen in its singular values. From Figure 3, it is clear that the texture is the only component to have many zero singular values. Even if the “smaller” singular values



**Figure 2.** Decomposition of a synthetic example image.

were thresholded, the texture component would still have the smallest singular value support. In Figure 4, we see that even seemingly random textures have a small patch-rank.

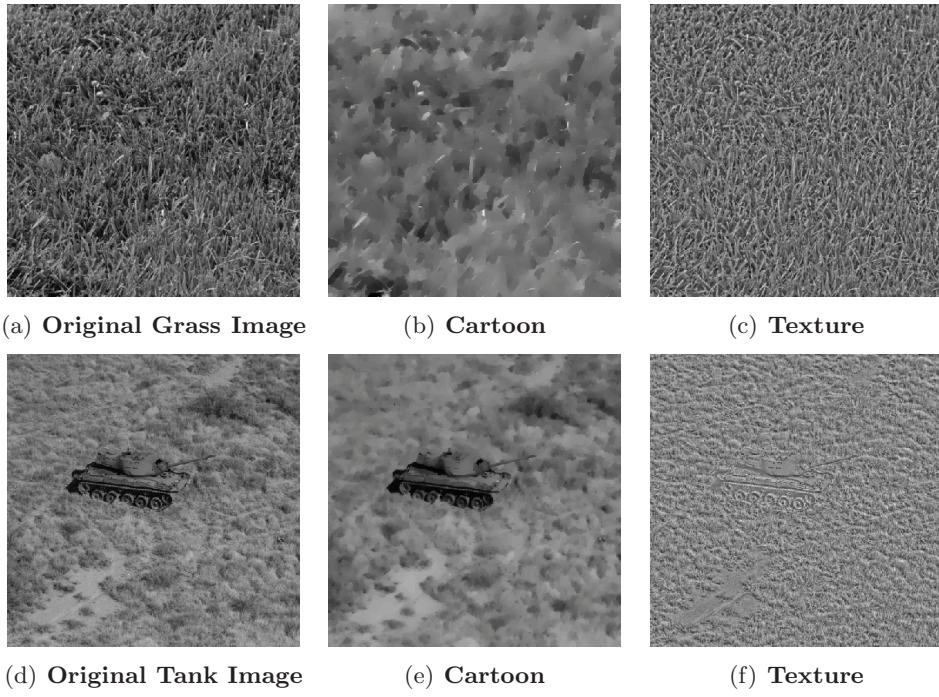
*Remark 2.4.* Like the cartoon-texture symmetry from the *BV-G* model, our cartoon and texture norms are related in a curious way. For the cartoon  $u$ , the operator  $D$  is applied and the result is measured in  $L^1$ . For the texture  $v$ , the patch-texture is decomposed by the singular value decomposition; consequently, there exist two unitary matrices  $U$  and  $V$  such that  $\mathcal{P}v = U\Sigma V^*$ , where  $\Sigma$  is a diagonal matrix that contains the singular values of  $\mathcal{P}v$ . This is equivalent to the  $L^1$  norm on  $\Sigma$ , i.e.,  $\|U^*\mathcal{P}vV\|_1$ . While the operator  $D$  looks at *local* comparisons in pixel values, the operation  $v \mapsto U^*\mathcal{P}vV$  compares patches *nonlocally*. The  $L^1$  norm is used to measure sparsity in jumps for the case of the cartoon, and sparsity in pattern for the case of the texture.



**Figure 3.** Singular values in descending order.

*Remark 2.5.* There are many ways to view our texture norm.

1. From the perspective of nonlocal methods, our norm measures the similarity between patches based on linear dependence rather than elementwise differences. In these methods, the typical measure of similarity is the exponential of the  $L^2$  distance between patches. This can cause problems when two patches agree exactly in texture but have different means. Specifically, if the difference in means is large, then the patches are considered to be dissimilar, which can lead to improper comparisons between patches. In practice, this leads to problems in contrast (see Figure 19).
  2. From a dictionary approach, our texture norm creates a basis  $B_i$  for  $i = 1, \dots, \text{rank}(\mathcal{P}v)$  composed of the singular vectors of  $\mathcal{P}v$ . All patches are thus composed of linear combinations from this basis, creating an implicit dictionary based on the image itself.
  3. From a cartoon-texture point of view, our texture norm induces a space  $(T, \|\cdot\|_*)$



**Figure 4.** Examples of image decomposition. The patch-rank of the texture components for the Tank and Grass images is 51.0% and 19.1%, respectively. Even seemingly random textures have a small patch-rank. Notice that spiking (or point structures) in the original images (a) and (d) are found in the cartoon components (b) and (e) and not the texture components (c) and (f).

which is used in the energy. As with the functional spaces, the proposed texture norm decreases as the pattern becomes more repetitive.

**3. Theoretical and analytical remarks.** In this section we will examine the behavior of the cartoon-texture separation model, with the constraint that the input data  $f$  is of mean zero. This is not a restriction in practice since it is simply a rescaling of the data.

**3.1. Characterization of minimizers by duality.** In the continuous framework, the predual of  $BV$  is  $G$ , the space of generalized functions. As in Definition 1.1, the space  $G$  is equivalent to the space of functions which are the divergence of  $L^\infty$  vector fields. The space and norm first appeared in the characterization of minimizers for the ROF model and then in cartoon-texture models [28]. For our discrete model, we would like an analogous space which has a similar duality to our discrete  $BV$  space. In [2], a discrete  $G$  is given as follows.

**Definition 3.1.** *The dual of (discrete)  $BV$  is the (discrete) space  $G$ , which has the following norm:*

$$\|v\|_G := \inf_{v=\text{div} \vec{g}} \left\| \sqrt{(g_{i,j}^1)^2 + (g_{i,j}^2)^2} \right\|_{L^\infty},$$

where  $\vec{g}_{i,j} = (g_{i,j}^1, g_{i,j}^2)$ .

We will consider the discrete divergence operator to be  $\text{div} = (D_x^-, D_y^-) \cdot$ , where  $D_i^-$  are the backward differences. It can be shown that this definition of divergence yields  $\text{div} = -D^*$ ,

where  $D^*$  is the adjoint of the operator from the definition of discrete TV. This is necessary to ensure a discrete duality principle, or, in other words, for any  $u$  and  $v$ ,  $|\langle u, v \rangle| \leq \|Du\|_1\|v\|_G$ . A similar duality is needed with respect to the texture norm, which follows from the following properties of the patch map.

**Lemma 3.2.** *If  $\mathcal{P}$  is a patch map, then the following hold:*

1. *It is a bijective linear operator.*
2.  *$\mathcal{P}$  is an isometry with respect to all elementwise norms.*
3.  $\|\mathcal{P}v\|_* \leq \sqrt{n}\|v\|_2$ , where  $n^2$  is the size of the matrix  $v$ .
4. *The dual norm of  $\|\mathcal{P} \cdot\|_*$  is  $\|\mathcal{P} \cdot\|_s$ , where  $\|\cdot\|_s$  is the spectral norm.*

There are sharper bounds for condition 3 above; however, they depend on the patch size and the texture component. The proofs for each of the above are easy to show. From these properties, the texture duality can be shown to be  $|\langle u, v \rangle| \leq \|\mathcal{P}u\|_s\|\mathcal{P}v\|_*$ . Based on these dualities, there are four characterization theorems given below. The characterization is centered around the pure decomposition version of the model (when  $A = I$ ):

$$(3.1) \quad \mathcal{E}(u, v) = \mu\|Du\|_1 + \gamma\|\mathcal{P}v\|_* + \frac{\lambda}{2}\|u + v - f\|_2^2.$$

The theorems give a relationship between the parameters  $(\mu, \gamma, \lambda)$  and the quantities  $\|f\|_G$  and  $\|\mathcal{P}f\|_s = \sigma_{max}(\mathcal{P}f)$  (which is the maximum singular value of the patch-form of the image). The first two theorems address the two trivial decompositions where no cartoon appears and the optimal solutions vary between the texture and residual terms. The proofs of these theorems can be found in Appendix B.

**Theorem 3.3.** *If  $0 < \gamma < \frac{2\mu}{n}$ , then the minimizer  $(u, v, \rho)$  must have  $u = 0$ . If, in addition,  $\|\mathcal{P}f\|_s \leq \frac{\gamma}{\lambda}$ , then the minimizer  $(u, v, \rho)$  yields  $u = \rho = 0$  and  $v = f$ .*

**Theorem 3.4.** *If  $\|f\|_G \leq \frac{\mu}{\lambda}$  and  $\|\mathcal{P}f\|_s \leq \frac{\gamma}{\lambda}$ , then the minimizer  $(u, v, \rho)$  must yield  $u = v = 0$  and  $\rho = f$ .*

The last two theorems are the more interesting cases. By choosing the parameters accordingly, the resulting optimal solutions will provide nontrivial decompositions.

**Theorem 3.5.** *If  $\|f\|_G > \frac{\mu}{\lambda}$  and  $\|\mathcal{P}f\|_s > \frac{\gamma}{\lambda}$ , then the minimizer yields  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ ,  $\langle \rho, u \rangle = \frac{\mu}{\lambda}\|Du\|_1$ , and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda}\|\mathcal{P}v\|_*$ .*

**Theorem 3.6.** *If  $\|f\|_G \leq \frac{\mu}{\lambda}$  and  $\|\mathcal{P}f\|_s > \frac{\gamma}{\lambda}$ , then three optimal cases hold:*

- (1)  $u = 0$ ,  $\|\rho\|_G < \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s \leq \frac{\gamma}{\lambda}$ , and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda}\|\mathcal{P}v\|_*$ .
- (2)  $v = 0$ ,  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s < \frac{\gamma}{\lambda}$ , and  $\langle \rho, u \rangle = \frac{\mu}{\lambda}\|Du\|_1$ .
- (3)  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ ,  $\langle \rho, v \rangle = \frac{\gamma}{\lambda}\|\mathcal{P}v\|_*$ , and  $\langle \rho, u \rangle = \frac{\mu}{\lambda}\|Du\|_1$ .

These theorems provide some insight into choosing coefficients to obtain particular behaviors in the minimizers.

**3.2. Characterization of texture based on  $\mathcal{P}$ .** We can further characterize solutions based on our patch map. The operator  $\mathcal{P}$  is implicitly dependent on both the way in which the elements are reassigned and the patch size. First, to address the “reshaping” we have the following theorem.

**Lemma 3.7.** *Let  $\mathcal{P}_1$  and  $\mathcal{P}_2$  be two patch maps which are identical except for the order in which they map the subblocks of  $v$  into the columns of  $\mathcal{P}_i v$ , and the elements within the subblock of  $v$  into the rows of  $\mathcal{P}_i v$ . Then we have that, for all  $v$ ,  $\|\mathcal{P}_1 v\|_* = \|\mathcal{P}_2 v\|_*$ .*

By this lemma we see that our method is independent of the ordering used to reshape the texture matrix into the patch-form.

Next, let us address the dependence of the texture component on the patch size  $r$  (for now assume that the patches are square). Like the tuning parameters in the energy, this parameter also determines certain characteristics of the minimizers. For example, the patch size has a subtle relationship with the various “texture frequencies.” Consider the following example.

*Example 3.8.* Let  $f(x, y)$  be a mean zero function that has an oscillation of period  $T$  only in the  $x$ -direction (and constant along the  $y$ -direction). Let  $r$  be the length of the square patches.

1. If  $T < r$ , then the patch-rank must be larger than or equal to 1. If there are small scale symmetries during one period, then the patch-rank can be equal to 1. For example, take  $r = \frac{T}{2}$ . If  $f(x, y) = -f(x + \frac{T}{2}, y)$ , then the patch on  $[1, \frac{T}{2}] \times [1, \frac{T}{2}]$  is equal to the negative of the patch on  $[\frac{T}{2} + 1, T] \times [\frac{T}{2} + 1, T]$ . It follows that there is only one underlying base patch, so the rank is 1.
2. If  $T > r$ , then the patch-rank must be larger than or equal to 1. This can be shown by using the previous argument with  $r = \frac{3T}{2}$ .
3. If  $T = r$ , then the patch-rank will be 1.

From this example, it is clear that to minimize the number of patches needed to describe a given texture, the parameter  $r$  should be as close to the pattern period as possible.

Last, the following lemma address the partitioning of the texture component.

**Lemma 3.9.** *Let  $\mathcal{P}$  be a patch map, and let  $\mathcal{S}$  be a index shifting operator (with periodic boundary conditions); then for any matrix  $v$ ,  $\|\mathcal{P}\mathcal{S}v\|_* = \|\mathcal{P}v\|_*$  holds.*

The lemma shows that the patch map is invariant under uniform translations of the indices; therefore our proposed method is invariant of the partitioning grid. This is easy to show, since shifting the grid is equivalent to reordering the indices but does not change the relationship between indices, as in Lemma 3.7.

**4. Numerical method.** The cartoon and texture are both defined with respect to fine (lower-dimensional) structures: codimensional 1 edges in the cartoon and codimensional 1 and 2 patterns in the texture. They are also both measured by  $L^1$ -type norms, which are efficiently solved by splitting methods — in our case the split Bregman method [20]. The added advantage of these splitting methods is that lower-dimensional structures get enhanced. This is normally referred to as *contrast enhancement*, i.e., the sharpening of edges. An analogous effect seems to sharpen the “texture contrast,” specifically linear features and isolated point structures. It is interesting to note that in practice, the Bregman iteration removes any block effects from the texture component, which occur from the nonoverlapping patch structure.

Below is an outline of the split Bregman technique applied to our model. We use a double Bregman, since both terms  $u$  and  $v$  are split. First introduce the auxiliary variables  $d_1 = Du$  and  $d_2 = \mathcal{P}v$ :

$$\min_{d_1=Du, d_2=\mathcal{P}v} \mu\|d_1\|_1 + \gamma\|d_2\|_* + \frac{\lambda}{2}\|A(u + v) - f\|_2^2.$$

Next add back the constraints  $d_1 = Du$  and  $d_2 = \mathcal{P}v$ , enforcing them with the Bregman

variables  $b_1, b_2$ :

$$\min_{d_1, d_2, u, v} \mu \|d_1\|_1 + \gamma \|d_2\|_* + \frac{\lambda}{2} \|A(u + v) - f\|_2^2 + \frac{\lambda_1}{2} \|d_1 - Du + b_1\|_2^2 + \frac{\lambda_2}{2} \|d_2 - \mathcal{P}v + b_2\|_2^2.$$

This splitting decouples the original equation into the following system:

$$(4.1) \quad (u^{n+1}, v^{n+1}) = \operatorname{argmin}_{u^n, v^n} \frac{\lambda}{2} \|A(u^n + v^n) - f + f^n\|_2^2 + \frac{\lambda_1}{2} \|d_1^n - Du^n - b_1^n\|_2^2 + \frac{\lambda_2}{2} \|d_2^n - \mathcal{P}v^n - b_2^n\|_2^2,$$

$$(4.2) \quad d_1^{n+1} = \operatorname{argmin}_{d_1^n} \mu \|d_1^n\|_1 + \frac{\lambda_1}{2} \|d_1^n - Du^{n+1} - b_1^n\|_2^2,$$

$$(4.3) \quad d_2^{n+1} = \operatorname{argmin}_{d_2^n} \gamma \|d_2^n\|_* + \frac{\lambda_2}{2} \|d_2^n - \mathcal{P}v^{n+1} - b_2^n\|_2^2,$$

$$(4.4) \quad b_1^{n+1} = b_1^n + Du^{n+1} - d_1^{n+1},$$

$$(4.5) \quad b_2^{n+1} = b_2^n + \mathcal{P}v^{n+1} - d_2^{n+1}.$$

In (4.1), we have also included the Bregman variable  $f^n$ , which is used when one wants to enforce the constraint  $A(u + v) = f$ . Each of the subproblems above can be easily solved, as follows. For the first subproblem, from Lemma 3.7 and the tensor discussion in Appendix A, we can rewrite (4.1) as

$$(u^{n+1}, v^{n+1}) = \operatorname{argmin}_{u^n, v^n} \frac{\lambda}{2} \|A(u^n + v^n) - f + f^n\|_2^2 + \frac{\lambda_1}{2} \|d_1^n - Du^n - b_1^n\|_2^2 + \frac{\lambda_2}{2} \|\mathcal{P}^{-1}d_2^n - v^n - \mathcal{P}^{-1}b_2^n\|_2^2.$$

Since this problem is differentiable, taking the first variation yields the following linear system:

$$(4.6) \quad \begin{cases} (\lambda A^* A - \lambda_1 \Delta) u^{n+1} + \lambda A^* A v^{n+1} = F_1, \\ \lambda A^* A u^{n+1} + (\lambda A^* A + \lambda_2) v^{n+1} = F_2, \end{cases}$$

where  $F_1 = \lambda A^*(f - f^n) - \lambda_1 D^*(d_1^n - b_1^n)$  and  $F_2 = \lambda A^*(f - f^n) + \lambda_2 \mathcal{P}^{-1}(d_2^n - b_2^n)$ . This equation can be solved completely (in the Fourier domain) or approximated with a few iterations of a Gauss–Seidel (GS) sweep. Next, (4.2) can be written out explicitly as a simple shrink:

$$d_1^{n+1} = \operatorname{shrink} \left( Du^{n+1} + b_1^n, \frac{\mu}{\lambda_1} \right),$$

where the *shrink* function above is defined pointwise for two-dimensional vectors  $x$  as  $\operatorname{shrink}(x, \tau) := \max(|x|_2 - \tau, 0) \frac{x}{|x|_2}$ , and where  $|\cdot|_2$  is the vector 2-norm and  $\tau \in \mathbb{R}$ . Last, (4.3) can be written explicitly as

$$d_2^{n+1} = \operatorname{SVT} \left( \mathcal{P}v^{n+1} + b_2^n, \frac{\gamma}{\lambda_2} \right),$$

where  $SVT$  is singular value thresholding, which is defined as follows: for a matrix  $M$  whose singular value decomposition (SVD) is given by  $M = U\Sigma V^*$ , the singular value thresholding function is defined as  $SVT(M, \tau) := U \max(\Sigma - \tau I, 0) V^*$ , where the max is taken elementwise. There are methods for computing the  $SVT$  without using the SVD (see [9]), which can speed up the computations for large matrices. Using these formulas, the algorithm is presented in the next section.

**4.1. The algorithm.** The splitting from the previous section reduces the problem from a difficult nonlinear one to a sequence of simple linear (or explicit) subproblems. The algorithm involves two main loops: the inner loop, which solves each minimization, and the outer loop, which adds back the error and re-solves the minimization. The outer loop's termination is dependent on the problem we are solving, specifically on the amount of noise. Given a tolerance  $tol$ , for pure decomposition with no noise or for inpainting with no noise, the outer loop is iterated until  $\|f - A(u^{n+1} + v^{n+1})\|_2 \leq tol$ . In the presence of noise with standard deviation  $\sigma_{noise}$ , the stopping criterion becomes  $\|f - A(u^{n+1} + v^{n+1})\|_2 \approx \sigma_{noise}$ . The number of outer loops also determines the amount of texture-noise separation that occurs.

---

### Algorithm

```

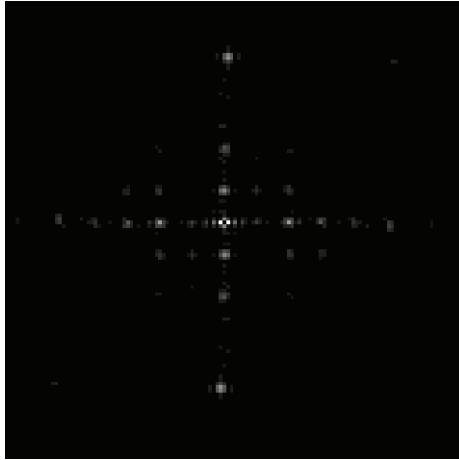
Initialize  $u^0 = f$ ,  $v^0 = 0$ ,  $d_{1,x}^0 = d_{1,y}^0 = b_{1,x}^0 = b_{1,y}^0 = 0_{n,n}$ , and  $d_2^0 = b_2 = 0_{r^2, \frac{n^2}{r^2}}$ 
while Outer Iteration do
    while Inner Iteration do
         $(u^{n+1}, v^{n+1}) = GS^n$ 
         $d_1^{n+1} = shrink \left( Du^{n+1} + b_1^n, \frac{\mu}{\lambda_1} \right),$ 
         $d_2^{n+1} = SVT \left( \mathcal{P}v^{n+1} + b_2^n, \frac{\gamma}{\lambda_2} \right),$ 
         $b_1^{n+1} = b_1^n + Du^{n+1} - d_1^{n+1},$ 
         $b_2^{n+1} = b_2^n + \mathcal{P}v^{n+1} - d_2^{n+1}$ 
    end while
     $f^{n+1} = f^n + f - (u^{n+1} + v^{n+1})$ 
end while

```

---

In the algorithm,  $GS^n$  is the application of Gauss–Seidel sweeping to (4.6). Typically, only a few sweeps are necessary, since only partial convergence is needed.

**5. Results.** In the previous section, methods for choosing appropriate parameters were given in the examples and theorems. Before discussing the numerical results, a summary with one more parameter bound will be provided. Informally, an upper bound on the number of expected textures can be predicted, thereby providing a bound on  $\text{rank}(\mathcal{P}v)$ . As previously mentioned, the texture is composed of highly oscillatory functions. If we model them as sums of sines and cosines, then the Fourier transform yields sums of Dirac delta functions at those frequencies. In practice, the textures appear as pairs of spikes (of various amplitudes) in the Fourier domain (see Figure 5). By counting the number of spikes in the Fourier domain,



**Figure 5.** Fourier transform of Brodatz.

one can estimate an upper bound for the patch-rank (it is only an upper bound since jump discontinuities can contribute to the spikes).

Let  $r \times r$  be the patch size of  $\mathcal{P}v$ ; then we have the following:

1. From the argument above, take  $r^2$  to be close to half the number of “large” spikes in  $|\mathcal{F}(v)|$ .
2. From Example 3.8, if the largest texture period is  $T$ , then take  $r$  to be as close to  $T$  as possible.
3. Choose  $(\mu, \gamma, \lambda)$  based on Theorems 3.3–3.6. Also, the ratio  $\frac{\mu}{\lambda_1}$  determines the amount of cartoon, while the ratio  $\frac{\gamma}{\lambda_2}$  determines the amount of texture.
4. Normalizing  $f$  such that  $\sum |f_{i,j}| = 1$  (or  $\max |f_{i,j}| = 1$ ) and  $\sum f_{i,j} = 0$  helps when choosing  $(\mu, \gamma, \lambda)$ .
5. Setting  $\lambda = \lambda_1 = \lambda_2$  gives appropriate results and removes two parameters from the model.
6. In regard to the algorithm, the number of outer loops determines the “amount of edges and texture” that is added back. If the original image is noisy, this parameter must be tuned in order to avoid adding back the noise. As mentioned in [20], the number of GS sweeps can be fixed to a value of 2–5 (only partially converging), and the method will give satisfactory results.

For the results that follow we take 5 by 5 patches (unless otherwise stated),  $\lambda = 1$ ,  $\mu \in [.75, 2]$ , and  $\gamma \approx \sigma_{\max}(\mathcal{P}v)$ .

**5.1. Decomposition.** For the pure cartoon-texture decomposition problem, we would like to remove the texture component without removing other key features such as edges and shading. Take, for example, the Barbara image (Figure 6). We estimated approximately 25 spikes in the Fourier domain, so a patch size of 5 by 5 is more than sufficient. In Figure 7, our method is compared to the standard  $TV$ - $G$  model [3] and the  $TV$ - $L^2$  model (i.e., ROF model). For  $TV$ - $L^2$ , we use the split Bregman approach from [20]. For all models, the parameters are chosen in order to have the same  $L^2$  norm on the texture component. As seen in Figure 7,



**Figure 6.** *Barbara.*

our method provides results similar to those of the classical  $TV-G$  separation while removing fewer edges, for example, Barbara's hair and the background. Unlike our method, the  $TV-L^2$  model does not remove the texture evenly.

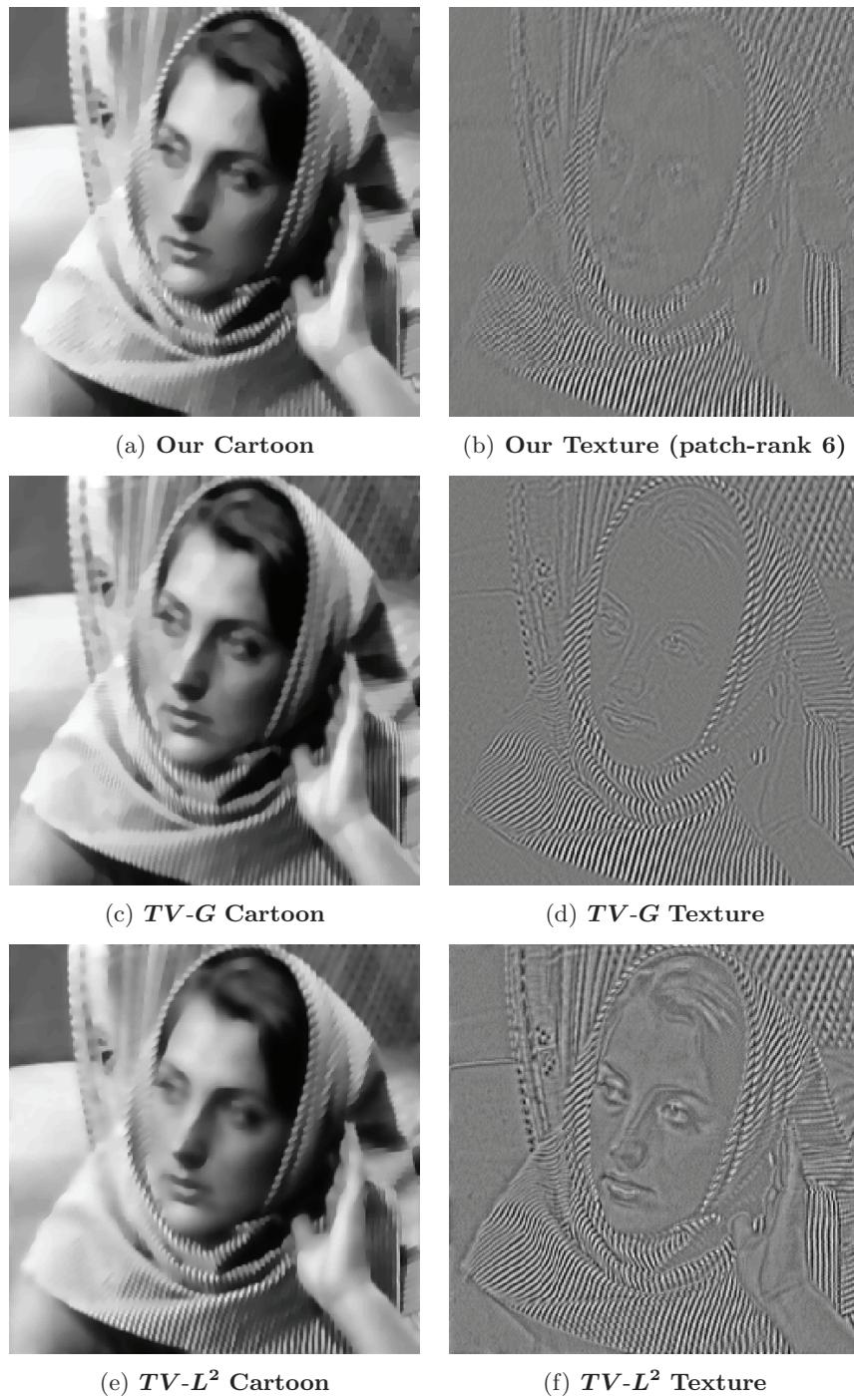
In Figure 8, the texture component's dependence on the patch size is examined. The similarity between the textures suggests some flexibility in determining the patch size, although the smaller patch size yields less texture, while the large patch size removes nontexture features. In Figure 9, the texture component's dependence on  $\gamma$  (the texture norm's coefficient in the energy) is investigated by fixing  $\mu$  and  $\lambda$ . Increasing  $\gamma$  gives smaller patch-ranks and removes fewer nontexture features, while decreasing  $\gamma$  removes more details from the cartoon (see Theorem 3.3).

Fixing the texture from Figure 8(d) and decomposing the patch-form by the SVD yields  $\mathcal{P}v = U\Sigma V^*$ , i.e.,  $v = \mathcal{P}^{-1}(U\Sigma V^*)$ , where the diagonal matrix  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$  contains the  $r$  singular values in descending order. In Figure 10, this texture is then reconstructed using only some singular vectors (thresholding  $\Sigma$ ). Note that the singular values capture the relative significances of the base textures.

Last, Figure 4 shows the wide range of textures that can be captured by our model. The Tank and Grass images have visually random patterns yet have small patch-ranks (51.0% and 19.1%, respectively). Section 5.5.1 will discuss in more detail the types of textures that can be well represented by our model.

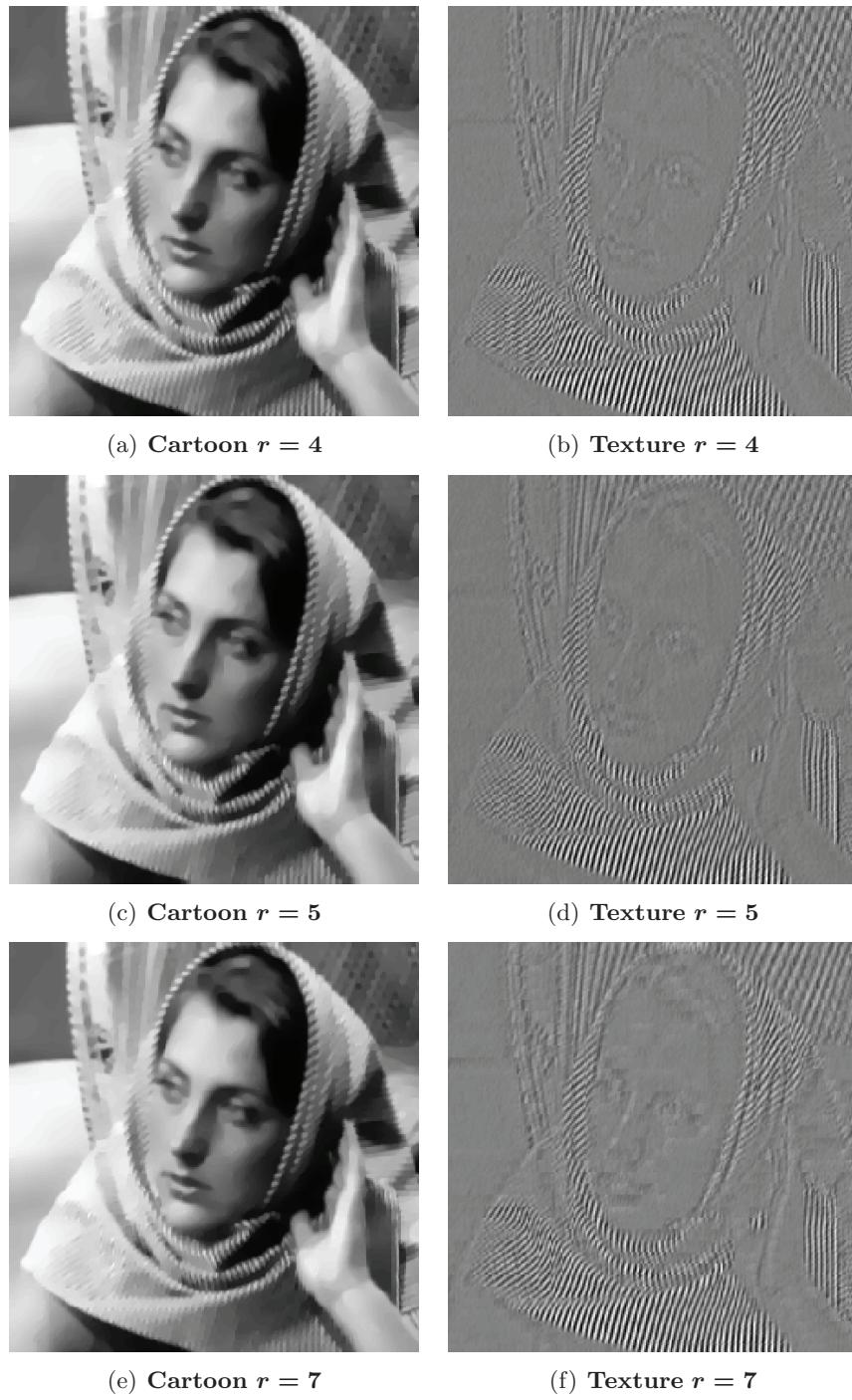
**5.2. Pattern regularization.** Since our texture norm penalizes nonuniform behavior between patches, our decomposition can be applied to image regularization. By regularizing the patterns, irregularities in the image are removed. In Figures 11 and 12, we regularize the highly textured Brodazt image using (3.1). The image  $f$  is decomposed into  $f = u + v + \rho$ , where  $u + v$  is the regularized image and  $\rho$  contains the irregularities. In Figure 12, it can be seen that, in our method, the essential vertical pattern is preserved, while the unwanted point structures are removed. Furthermore, our method also retains more texture than the standard median filter.

**5.3. Denoising.** In Figure 2, we saw the nearly perfect recovery of a noisy synthetic image. Texture-noise separation is difficult for most cartoon-texture decomposition methods; see Figure 13. In Figures 14 and 15, the Brodazt image is degraded by Gaussian noise of zero mean and recovered by  $TV-L^2$ ,  $NLT-V-L^2$ , and our method. For  $NLT-V-L^2$ , to make the



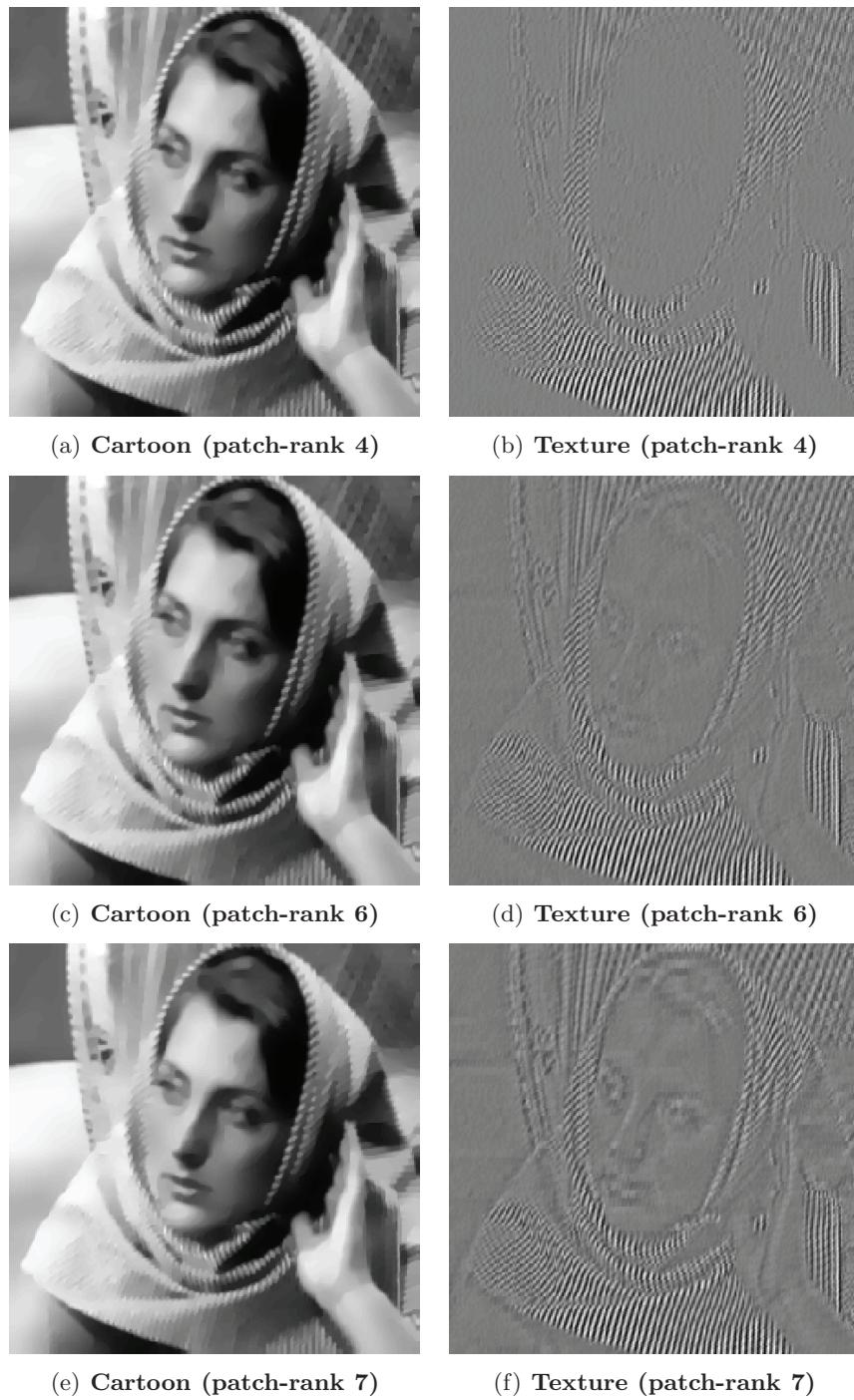
**Figure 7.** *Decomposing Barbara.*

comparison as similar as possible, we use a semilocal version with the same patch size as in our method [16, 17]. In both the  $TV-L^2$  and the  $NLT\!V-L^2$  models, the noise is removed from



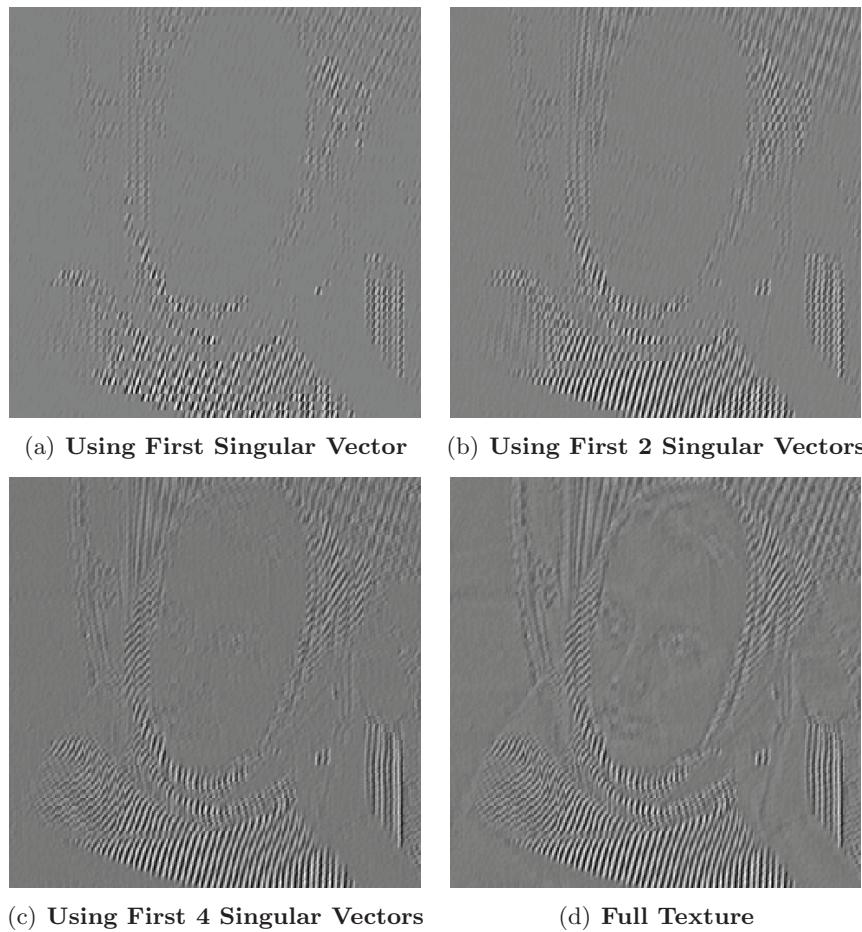
**Figure 8.** Decomposing *Barbara* with various patch sizes.

the top left quadrant. However, in more textured quadrants, more texture loss is seen (see Figure 15). For  $TV-L^2$  (peak signal-to-noise ratio (PSNR) = 30.2), although some texture

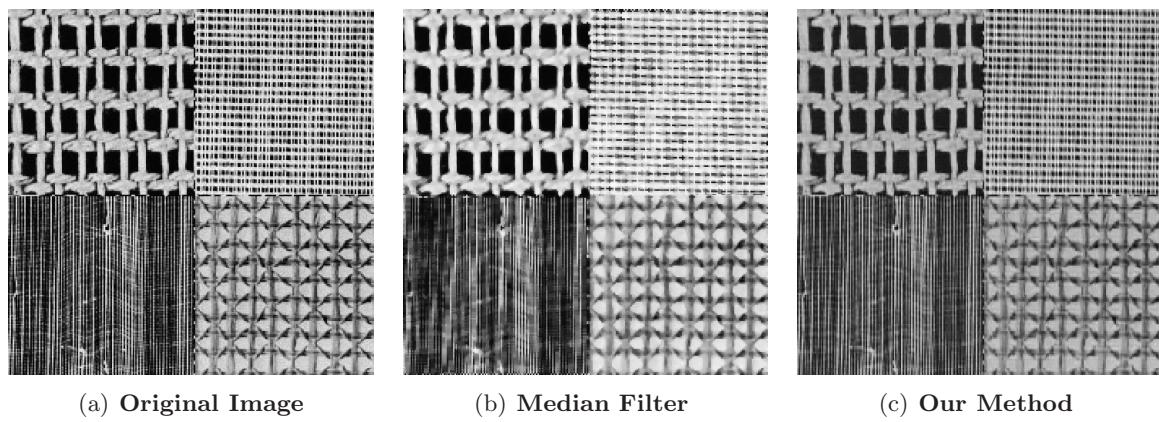


**Figure 9.** Decompositions with different patch-ranks (fixed  $\mu$  and  $\lambda$ ;  $\gamma$  variable).

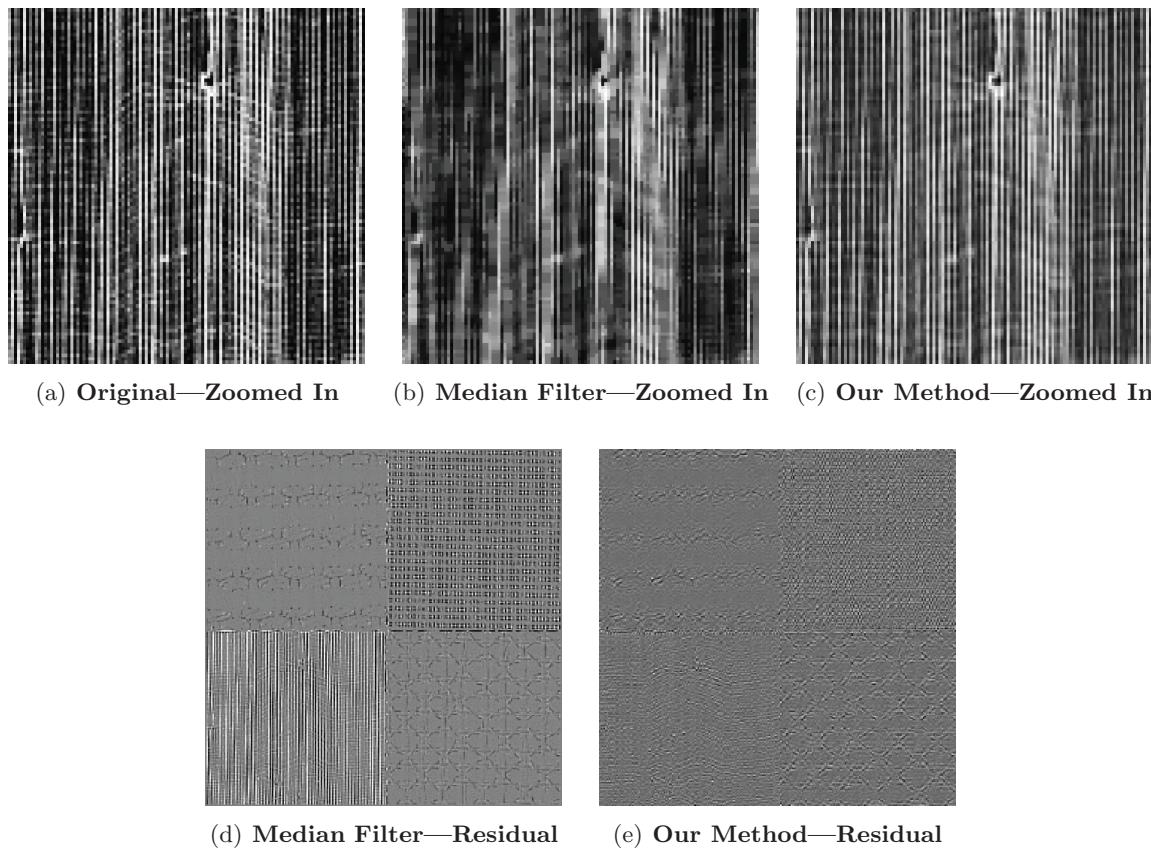
is recovered, only the cartoon-like regions are recovered well. The  $NLT V-L^2$  (PSNR = 31.8) recovers both the cartoon and texture well, but it loses some texture in order to obtain a



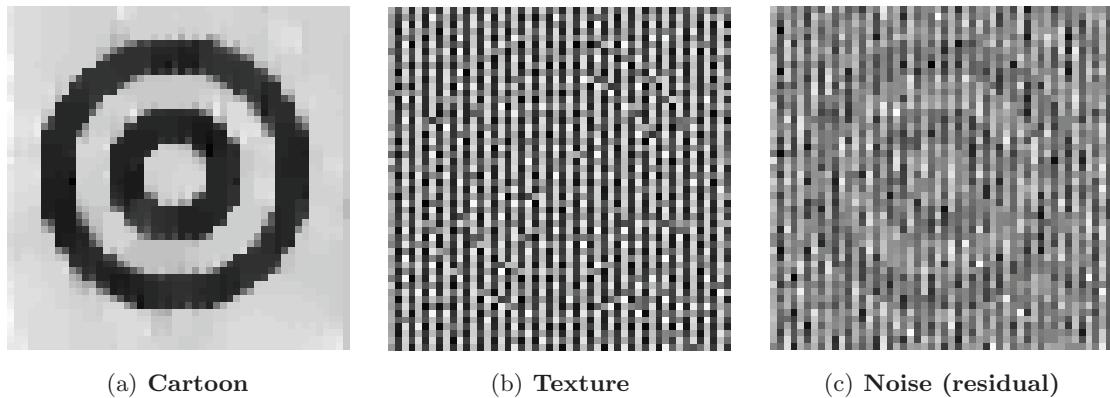
**Figure 10.** Various singular vectors for a fixed texture.



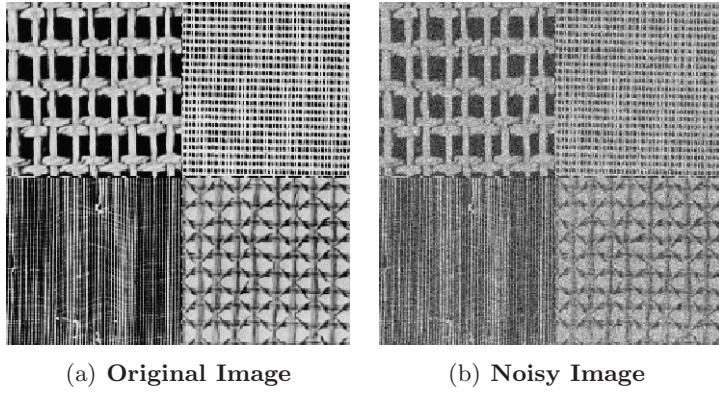
**Figure 11.** Image regularization.



**Figure 12.** *Image regularization.*



**Figure 13.** *Decomposition of a synthetic example image using  $TV\text{-}G$ .*

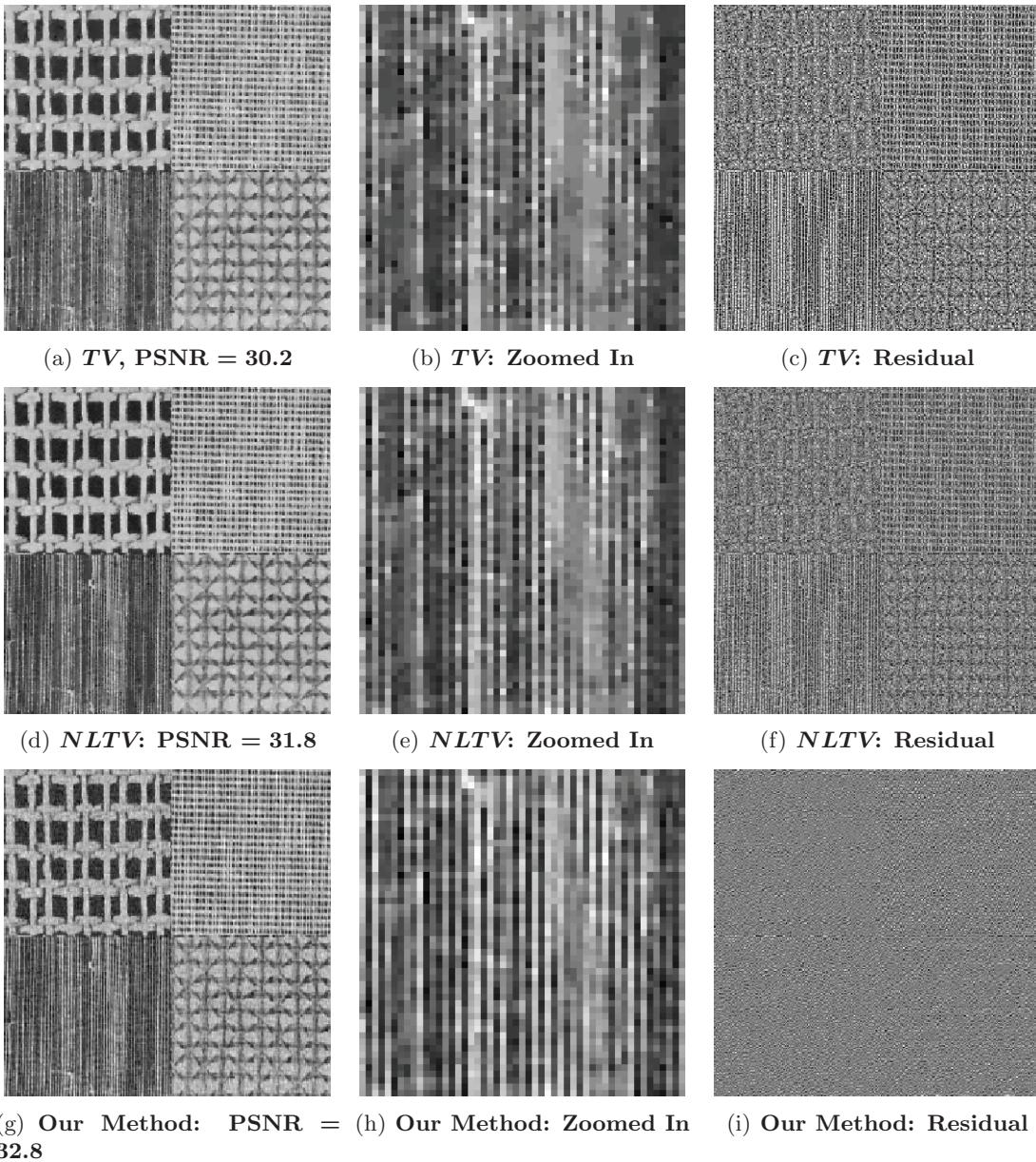


**Figure 14.** Decomposition of a synthetic example image.

particular level of smoothness. Our method recovers both parts well, with the highest PSNR of 32.8, while being faster than the semilocal method (restricting the window size to 11 by 11). The slight oscillatory pattern that appears in the top left quadrant is an artifact of the nonlocality of SVD.

**5.3.1. Denoising: Quantitative comparisons.** In general, there are no exact metrics in which to compare textures; however, there are some features which one prefers. In particular, a “good texture” component is noise-free and contains many sharp small scale details. Since noise is dense and has many false spikes, we can measure the noise level in the texture component by computing the sparsity of the texture component (as a percent of total pixels) and the percent of pixels which are considered edges (after applying a gradient-based edge detector). The smaller the value of both the edge and the intensity sparsity, the better the texture. To measure the amount of noise that is removed by a method, we compute the entropy of the noise component and compare it to the known entropy of the added noise. The entropy is also applied to the error term to measure the amount of structure and texture that is lost to the noise component. A small error entropy means that the recovered image is closer to the original in terms of small scale features (since oscillations may appear as larger entropy). Last, for reference, the patch-rank is computed for each method. We compare our proposed texture method to standard ones using various texture and noise metrics.

In Table 1, the synthetic image from Figure 2 is denoised by various methods. Each method has the same variance on its corresponding noise component. Since the true cartoon and true texture are known, we compare only the SNR of each component. Note that both of the entropy metrics order the methods in a way that is similar (but not exactly equivalent) to that of the SNR. In Table 2, a noisy version of the Brodatz–Wood image is denoised. Our method does as well as the nonlocal methods in terms of SNR, but does better in terms of removing the appropriate noise distribution (measured by entropy). Compared to the other standard cartoon-texture methods, our proposed model outputs a less noisy texture component. In Tables 3 and 4, we compare our method only with other cartoon-texture decomposition methods. In both cases, it is clear that our method better handles texture-noise separation. Using these results, we can partially conclude that this pattern-based interpretation is a more appropriate



**Figure 15.** Denoising of Brodatz.

definition for texture than those that are based on oscillations.

**5.4. Deblurring.** Blurry images tend to have severe texture loss and require methods which reconstruct the texture well. In Figure 16, the Barbara image is blurred by a Gaussian kernel with a standard deviation of 1.1. We compare our recovered image with the  $TV-L^2$  and  $NLTv-L^2$  [34] deblurring methods. In terms of PSNR, our method better reconstructs the image with the  $NLTv$  method getting very close results. Visually, the textures on the

**Table 1**

Denoising of the synthetic image. Each method has the same noise variance ( $L^2$  norm). The entropy of the true noise is 4.53.

Method	Recovered SNR	Cartoon SNR	Texture SNR	Noise entropy	Error entropy
Ours	<b>19.3</b>	<b>19.0</b>	<b>17.3</b>	<b>4.43</b>	<b>3.34</b>
$TV-L^2$	8.6	N/A	N/A	3.48	4.29
$NLT V-L^2$	19.1	N/A	N/A	4.42	3.50
$TV-G-L^2$	11.2	13.4	5.9	4.36	4.31
$TV-H^{-1}-L^2$	10.8	11.6	6.6	4.35	4.37

**Table 2**

Denoising of a noisy Brodatz-Wood image. Each method has the same noise variance ( $L^2$  norm) and same  $L^2$  norm on the texture component. The entropy of the true noise is 3.81. The patch-ranks for  $TV-G-L^2$  and  $TV-H^{-1}-L^2$  are computed after thresholding the smaller singular values.

Method	SNR	Edges	Sparsity	Patch-rank	Noise entropy	Error entropy
Ours	<b>15.14</b>	<b>10.9%</b>	<b>61.8%</b>	<b>15.1%</b>	<b>3.81</b>	<b>3.72</b>
$TV-L^2$	14.37	N/A	N/A	N/A	3.75	3.78
$NLT V-L^2$	<b>15.14</b>	N/A	N/A	N/A	3.80	3.79
$TV-G-L^2$	11.13	16.4%	63.3%	83.9%	4.00	4.15
$TV-H^{-1}-L^2$	11.50	11.6%	64.4%	82.6%	4.05	4.14

**Table 3**

Denoising of the noisy Grass image. Each method has the same noise variance ( $L^2$  norm) and same  $L^2$  norm on the texture component. The entropy of the true noise is 3.80. The patch-ranks for  $TV-G-L^2$  and  $TV-H^{-1}-L^2$  are computed after thresholding the smaller singular values.

Method	SNR	Edges	Sparsity	Patch-rank	Noise entropy	Error entropy
Ours	<b>7.83</b>	<b>8.8%</b>	<b>66.5%</b>	<b>15.6%</b>	<b>3.79</b>	<b>3.28</b>
$TV-G-L^2$	4.26	10.6%	66.6%	98.1%	3.78	3.94
$TV-H^{-1}-L^2$	5.84	13.0%	67.3%	100%	3.78	3.78

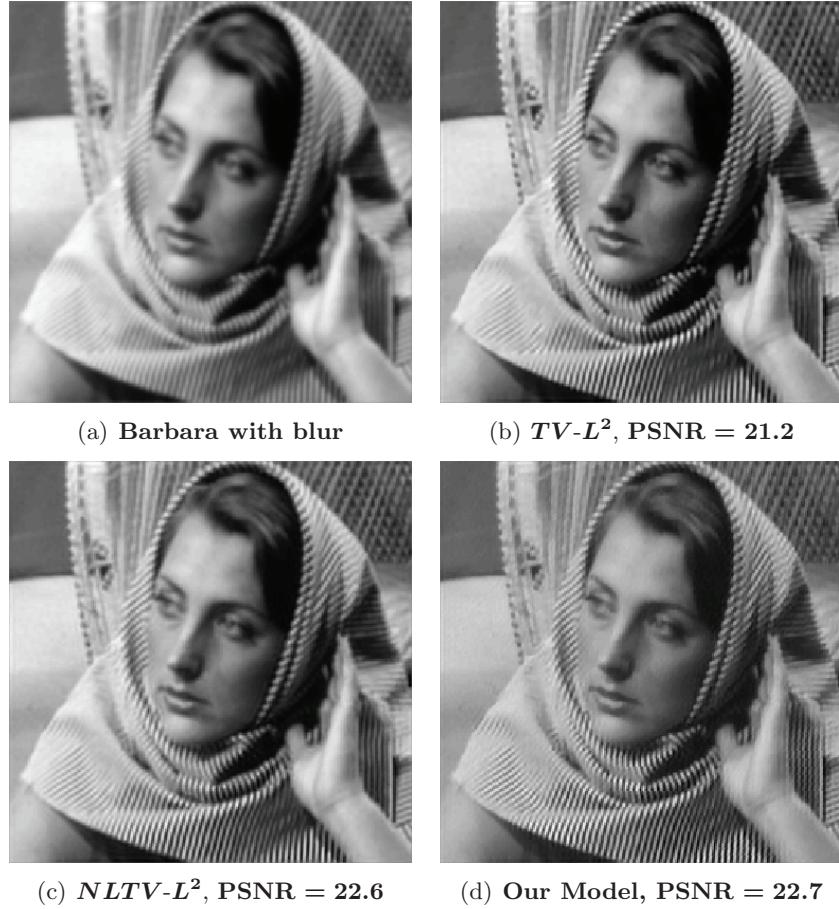
**Table 4**

Denoising of the noisy Tank image. Each method has the same noise variance ( $L^2$  norm) and the exact same cartoon component. The entropy of the true noise is 3.28. The patch-ranks for  $TV-G-L^2$  and  $TV-H^{-1}-L^2$  are computed after thresholding the smaller singular values.

Method	SNR	Edges	Sparsity	Patch-rank	Noise entropy	Error entropy
Ours	<b>11.86</b>	<b>7.2%</b>	<b>65.6%</b>	<b>40%</b>	<b>3.29</b>	<b>3.26</b>
$TV-G-L^2$	9.53	30.2%	67.6%	100%	3.35	3.49
$TV-H^{-1}-L^2$	9.57	0.3% <sup>1</sup>	67.8%	100%	3.34	3.49

<sup>1</sup>The  $TV-H^{-1}-L^2$  decomposition has a texture component that is mostly low amplitude noise, so the gradient detector outputs almost no edges.

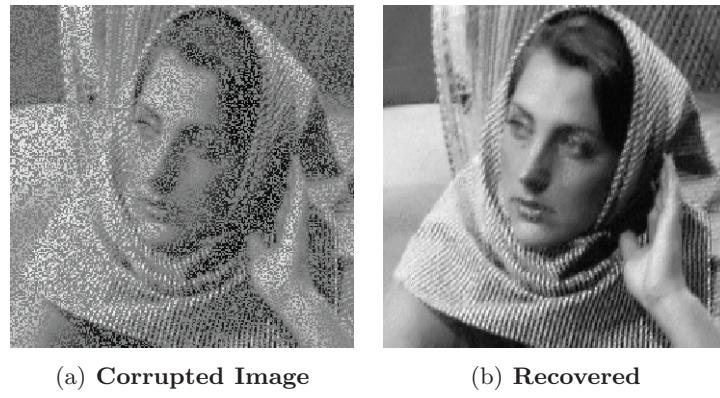
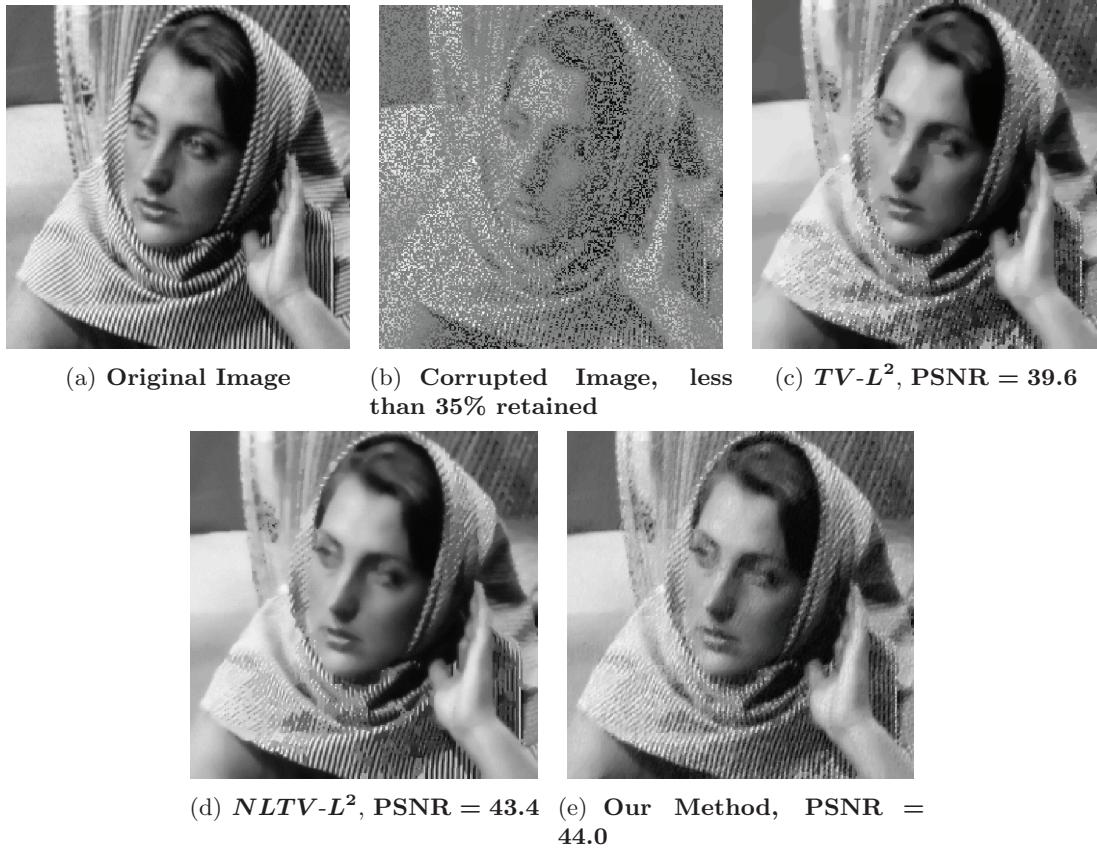
left and right ends of the scarf are sharper in our recovered image than in the others. Since the production of sharp edges is the key to deblurring methods, we can also quantitatively compare the results by measuring the percent of edge pixels in the image (using a gradient-based edge detector). The blurred image has 15.4% edge pixels, the  $TV$  method has 16.3%



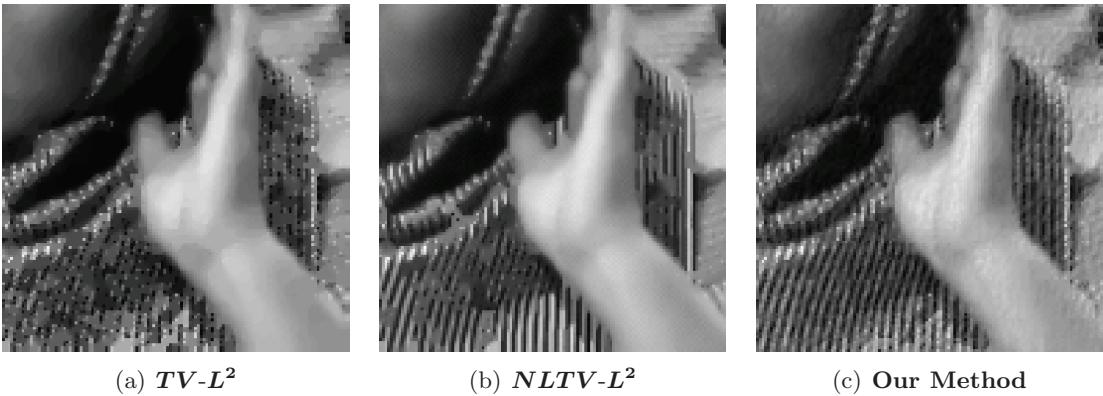
**Figure 16.** Deblurring comparison.

edge pixels, and the  $NLTВ$  method has 18.6% edge pixels, while ours has the most at 19.2% edge pixels. Using this metric, we can conclude that our method better produces features of sharp contrast. On a side note, in practice the texture regularized images do not seem to have the typical ringing effect associated with deconvolution problems. This leads us to believe that texture regularized models are more stable for deblurring.

**5.5. Inpainting: Sparse reconstruction.** In this and the next section, we discuss texture-regularized inpainting. There are two main types of inpainting: missing regions and sparse reconstruction. For inpainting missing regions, the image is first separated into its components, with the cartoon part recovered using a “structure” based inpainting, while the texture part is inpainted by texture-based techniques (for more on this methodology, see [6]). The numerical results presented here focus on sparse reconstruction. As an example, our method recovers an image almost perfectly (root mean square error (RMSE) is less than 0.08) with more than 50% of the pixels randomly removed; see Figure 17. In Figures 18 and 19, we compare our method to  $TV-L^2$  and  $NLTВ-L^2$  inpainting on an image with more than 65% of the pixels randomly removed. Our method is comparable in speed to  $TV-L^2$  and faster than  $NLTВ-L^2$ .

**Figure 17.** *Inpainting example.***Figure 18.** *Inpainting comparison.*

Our method recovers the image better ( $PSNR = 44.0$ ) than the  $TV-L^2$  ( $PSNR = 39.6$ ), while being comparable to  $NLT\mathcal{V}-L^2$  ( $PSNR = 43.4$ ).



**Figure 19.** *Inpainting comparison: zoomed in.*

**5.5.1. Sparse reconstruction and denoising: Quantitative comparisons.** Last, we investigate the type of textures which can be well represented within this framework. To do so, we generate a set of corrupt highly textured images by adding a fixed amount of noise to 100 images from [23] and then sparse sampling these images (removing half of the number of pixels). Our algorithm was applied to all of the corrupt images with parameters set at around  $\mu = 1$ ,  $\lambda = \lambda_1 = \lambda_2 = 1$ , 3 GS sweeps, 5 inner iterations, and 15 outer iterations. The patch size varied from 10 by 10 to 20 by 20 to match the texture scale of each individual image. The parameters were chosen to yield patch-ranks under 50. Table 5 displays statistics on the RMSE between the recovered image and the original, the RMSE between the original and the corrupt image, the percentage decrease in RMSE after the recovery, and the patch-rank (which is normalized out of 100). By considering the RMSE, the amount the RMSE changed, the patch-rank, and visual metrics, we can evaluate which textures are well represented by our method.

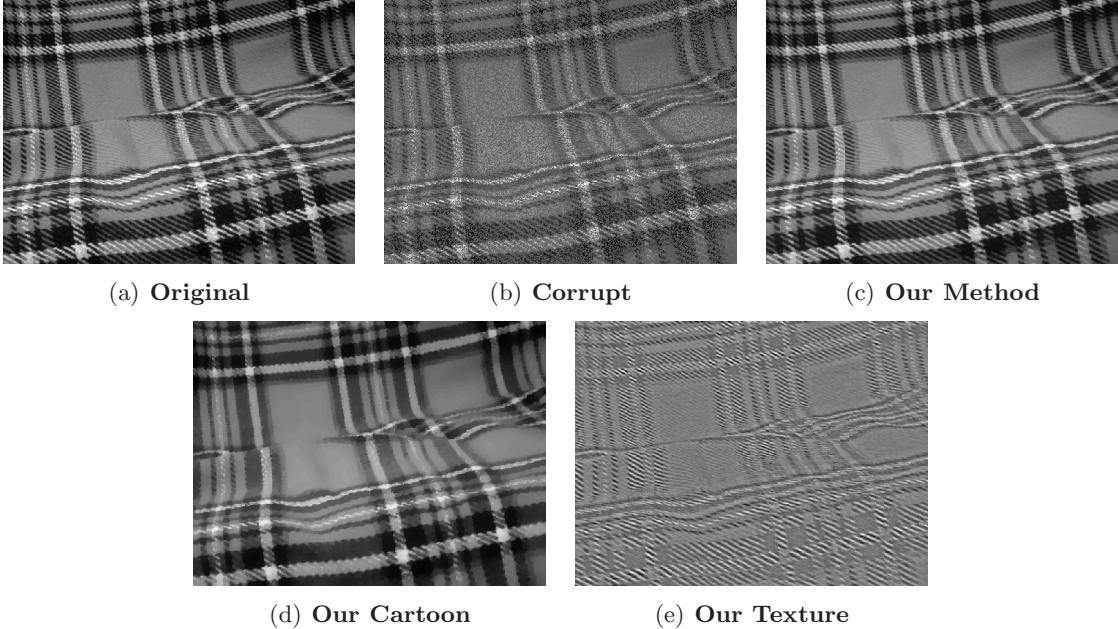
Although the patches are formed on a grid, the resulting textures do not have to be grid-like in structure or in pattern. Based on this experiment, our method is not sensitive to the angle or directionality of the texture pattern, the texture plane (i.e., frontal versus nonfrontal textures), the geometry of the pattern, or the geometry of the texture plane. For example, in Figure 20, the texture resides on a warped or nonuniform geometry, which does not align with the patch grid. However, our method performs well on this textured image and others like it. Our method has difficulty in capturing texture which does not have a regular pattern or has highly deviating structures. An example of this can be found in Figure 21, where the patterns have many irregular shapes over many scales. In this example, the texture component does not have a geometrically uniform pattern, and we can see loss of texture in the recovered image.

**6. Conclusion.** In this paper, we have presented a different way to model image texture or patterns in data. From the matrix perspective, we proposed measuring the patch-based singular values as a norm for the texture space. The new texture norm was used as a regularizer in models for decomposing, pattern regularizing, denoising, deblurring, and sparse reconstruction. Given a corrupt image  $f$ , the recovered image is reconstructed componentwise as  $u + v$ , where  $u$  is the cartoon (piecewise smooth) part, while  $v$  is the oscillatory/patterned texture

**Table 5**

Statistics on our algorithm applied to 100 images from the database in [23].

Statistic	Recovered RMSE	Original RMSE	RMSE decrease	Patch-rank
Minimum	0.035	0.091	-33.4%	5.5
Maximum	0.159	0.333	-83.0%	50.5
Mean	0.088	0.228	-61.0%	29.4
Median	0.085	0.231	-59.3%	30.9



**Figure 20.** Joint sparse reconstruction and denoising: warped/nonuniform geometry.

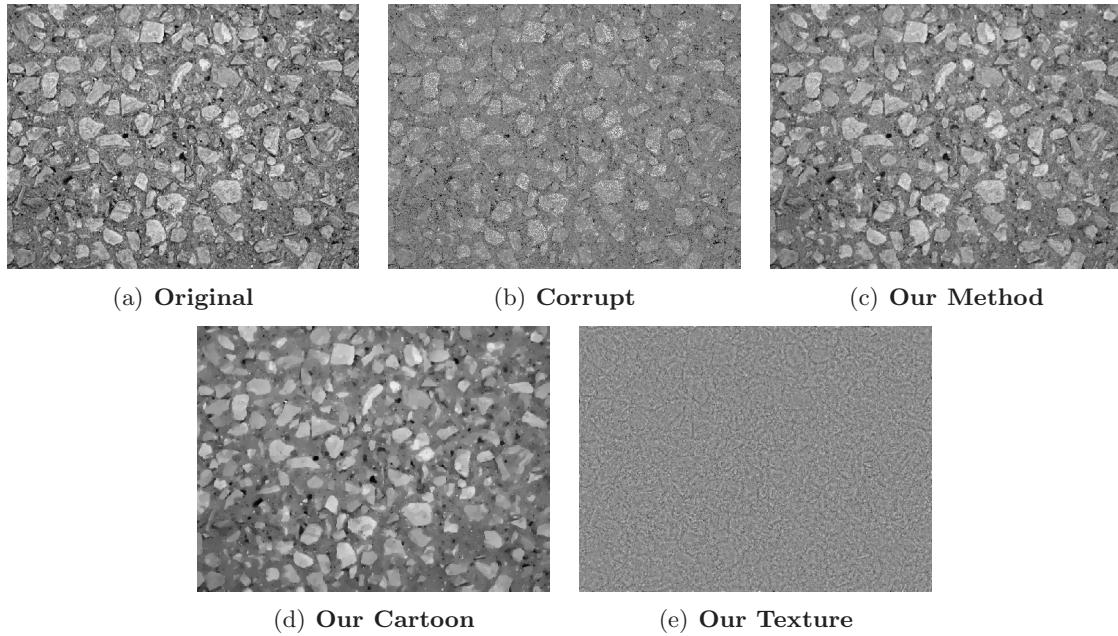
part. The various results support the benefit of including a texture regularizer rather than only reconstructing the smooth component. While other techniques for texture reconstruction exist, our method has the advantage of separating texture from noise, as well as its simplicity of construction, ease of implementation, and speed.

**Appendix A. A tensor interpretation.** As a small remark, since the patch map is a linear map between matrices, it can be associated with a tensor. The map  $\mathcal{P}$  is a rank 4 tensor (not to be confused with matrix rank) which maps  $\mathbb{R}^{n,n}$  into  $\mathbb{R}^{r^2, \frac{n^2}{r^2}}$ . Associating the map  $\mathcal{P}$  with the tensor  $P = [P_{i,j}^{k,l}]$ , the operation  $\mathcal{P} : v \mapsto w$  is defined by the following (using Einstein notation, where repeated indices are summed):

$$(A.1) \quad w_{i,j} = P_{i,j}^{k,l} v_{k,l},$$

where the elements of the tensor are defined as

$$P_{i,j}^{k,l} = \begin{cases} 1 & \text{if } (i,j) = \left((m(k)-1)r + m(l), 1 + \frac{n(k-m(k))}{r^2} + \frac{l-m(l)}{r}\right), \\ 0 & \text{otherwise} \end{cases}$$



**Figure 21.** Joint sparse reconstruction and denoising: multiscale with irregularities.

and  $m(\cdot) := \text{mod}(\cdot - 1, r) + 1$ .

The tensor is bijective since it is a unique invertible mapping. Furthermore, the tensor is a unitary transformation with respect to the elementwise inner product since for all  $v_1$  and  $v_2$ ,  $\langle \mathcal{P}v_1, \mathcal{P}v_2 \rangle = \langle v_1, v_2 \rangle$  holds.

**Appendix B. Proofs of theorems.** The proofs below are a generalization of the theory from Gilles and Meyer's work [19]. The main difference is that extra care must be given since there is no duality between our cartoon and texture norms.

**Lemma B.1.** *Take any Banach space with norm  $\|\cdot\|$  and dual norm  $\|\cdot\|'$ . Given an  $f$  that is decomposed into  $f = u + v$  by minimizing*

$$\|u\| + \frac{\lambda}{2}\|v\|_2^2,$$

*the following hold:*

1. *If  $\|f\|' \leq \frac{1}{\lambda}$ , then  $u = 0$  and  $v = f$ .*
2. *If  $\|f\|' > \frac{1}{\lambda}$ , then  $\|v\|' = \frac{1}{\lambda}$  and  $\langle u, v \rangle = \frac{1}{\lambda}\|u\|$ .*

Lemma B.1 will be used in many of the arguments in the various proofs. As in [19], the cartoon, texture, and error norms have a proper ordering.

**Lemma B.2.** *For all  $n$  by  $n$  matrices  $g$  (with mean zero), we have the following:*

$$\|\mathcal{P}g\|_* \leq \sqrt{n}\|g\|_2 \leq \frac{n}{2}\|Dg\|_1.$$

The lemma above provides a comparison between the terms in the energy and allows for the generalization of duality. More precisely, Lemma B.3 describes the dual pairing relationships.

**Lemma B.3.** For all  $n$  by  $n$  matrices  $u$  and  $v$ , we have the following:

$$\begin{aligned} |\langle u, v \rangle| &\leq \|Du\|_1 \|v\|_G, \\ |\langle u, v \rangle| &\leq \|\mathcal{P}u\|_s \|\mathcal{P}v\|_*. \end{aligned}$$

*Proof.* For the first inequality, take any  $(u, v)$  and let there exist a  $g$  such that  $v = \text{div}g$ , and thus

$$\begin{aligned} |\langle u, v \rangle| &= |\langle u, \text{div}g \rangle| \\ &= |\langle Du, g \rangle| \\ &\leq \|Du\|_1 \|g\|_\infty. \end{aligned}$$

Then taking the infimum with respect to all possible  $g$  yields the inequality. For the second inequality, we use the properties of  $\mathcal{P}$ :

$$\begin{aligned} |\langle u, v \rangle| &= |\langle \mathcal{P}u, \mathcal{P}v \rangle| \\ &\leq \|\mathcal{P}u\|_s \|\mathcal{P}v\|_*. \end{aligned}$$

Note that the spectral and trace norms are dual. ■

In order to apply Lemma B.1 to our model, we look at the pair  $(u, v)$  simultaneously. To do so, define  $w = u + v$  to be the reconstructed image composed of both components with the induced norm  $\|w\| = \inf \{\mu\|Du\|_1 + \gamma\|\mathcal{P}v\|_*\}$ . The dual norm is defined as  $\|g\|' = \sup \left\{ \frac{1}{\mu}\|g\|_G, \frac{1}{\gamma}\|\mathcal{P}g\|_s \right\}$ . Writing the energy in terms of  $w$  yields

$$\mathcal{E}(w) = \|w\| + \frac{\lambda}{2} \|f - w\|_2^2.$$

We will consider this in many of the proofs that follow.

Throughout Appendix B, the energy  $\mathcal{E}$  will be a functional with arguments  $u$ ,  $v$ , and  $\rho$  depending on the particular proof; however, all are equivalent by the relationship  $u + v + \rho = f$  with  $f$  given.

**Theorem B.4.** If  $0 < \gamma < \frac{2\mu}{n}$ , then the optimal decomposition yields  $u = 0$ .

*Proof.* Examining the energy with respect to the cartoon component  $u$  and the residual  $\rho$ , the energy can be bounded below by using Lemma B.2 and positivity of the norms:

$$\begin{aligned} \mathcal{E}(u, \rho) &= \mu\|Du\|_1 + \gamma\|\mathcal{P}(f - u - \rho)\|_* + \frac{\lambda}{2}\|\rho\|_2^2 \\ &\geq \frac{2\mu}{n}\|\mathcal{P}u\|_* + \gamma\|\mathcal{P}(f - u - \rho)\|_* + \frac{\lambda}{2}\|\rho\|_2^2 \\ &\geq \gamma\{\|\mathcal{P}u\|_* + \|\mathcal{P}(f - u - \rho)\|_*\} + \frac{\lambda}{2}\|\rho\|_2^2 \\ &\geq \gamma\|\mathcal{P}(f - \rho)\|_* + \frac{\lambda}{2}\|\rho\|_2^2 \\ &= \mathcal{E}(0, \rho). \end{aligned}$$

For all  $u \neq 0$  we have  $\|\mathcal{P}u\|_* > 0$ ; therefore  $u = 0$  is the minimizer. ■

Next, using Theorem B.4 with an additional constraint on  $f$  yields another trivial minimizer.

**Theorem B.5.** *If  $0 < \gamma < \frac{2\mu}{n}$  and  $\|\mathcal{P}f\|_s \leq \frac{\gamma}{\lambda}$ , then the optimal decomposition yields  $u = \rho = 0$  and  $v = f$ .*

*Proof.* From Theorem B.4, since  $0 < \gamma < \frac{2\mu}{n}$ , the energy is equivalent to the following:

$$\mathcal{E}(0, \rho) = \gamma \|\mathcal{P}(f - \rho)\|_* + \frac{\lambda}{2} \|\rho\|_2^2.$$

By Lemma B.1, since  $\|\mathcal{P}f\|_s \leq \frac{\gamma}{\lambda}$ , then  $\rho$  is identically 0.  $\blacksquare$

The theorems above provide the conditions in which the texture and/or residual component contains all of the information. The theorem below provides the final trivial case.

**Theorem B.6.** *If  $\|f\|_G \leq \frac{\mu}{\lambda}$  and  $\|\mathcal{P}f\|_s \leq \frac{\gamma}{\lambda}$ , then the optimal decomposition yields  $u = v = 0$  and  $\rho = f$ .*

*Proof.* Consider the simultaneous energy (with respect to  $w$ )

$$\mathcal{E}(w) = \|w\| + \frac{\lambda}{2} \|f - w\|_2^2.$$

The value of  $f$  in the dual norm can be calculated by using the assumptions

$$\begin{aligned} \|f\|' &= \sup \left\{ \frac{1}{\mu} \|f\|_G, \frac{1}{\gamma} \|\mathcal{P}f\|_s \right\} \\ &\leq \sup \left\{ \frac{1}{\lambda}, \frac{1}{\lambda} \right\} \\ &= \frac{1}{\lambda}. \end{aligned}$$

Applying Lemma B.1 yields  $w = 0$ , which implies both  $u = 0$  and  $v = 0$ .  $\blacksquare$

For the first nontrivial decomposition, Theorem B.7 characterizes minimizers when  $f$  is large with respect to particular norms.

**Theorem B.7.** *If  $\|f\|_G > \frac{\mu}{\lambda}$  and  $\|\mathcal{P}f\|_s > \frac{\gamma}{\lambda}$ , then the following holds for all minimizers:  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ ,  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$ , and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$ .*

*Proof.* Calculating the simultaneous dual norm yields the lower bound  $\|f\|' \geq \frac{1}{\lambda}$ , and therefore by Lemma B.1  $\|\rho\|' = \frac{1}{\lambda}$  and  $\langle w, \rho \rangle = \frac{1}{\lambda} \|w\|$ . Since the simultaneous dual norm is defined as  $\|\rho\|' = \sup \left\{ \frac{1}{\mu} \|\rho\|_G, \frac{1}{\gamma} \|\mathcal{P}\rho\|_s \right\}$ , we have the following bounds on the residual  $\rho$ :

$$\begin{aligned} \|\rho\|_G &\leq \frac{\mu}{\lambda}, \\ \|\mathcal{P}\rho\|_s &\leq \frac{\gamma}{\lambda}. \end{aligned}$$

By the duality principles from Lemma B.3, the following also hold:

$$\begin{aligned} \langle \rho, u \rangle &\leq \|Du\|_1 \|\rho\|_G = \frac{\mu}{\lambda} \|Du\|_1, \\ \langle \rho, v \rangle &\leq \|\mathcal{P}v\|_* \|\mathcal{P}\rho\|_* = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*. \end{aligned}$$

Together, these inequalities give

$$\langle u + v, \rho \rangle \leq \frac{1}{\lambda} \{ \mu \|Du\|_1 + \gamma \|\mathcal{P}v\|_* \}.$$

The inequality above is equivalent to  $\langle w, \rho \rangle \leq \frac{1}{\lambda} \|w\|$  for  $w = u + v$ . However, by Lemma B.1  $\langle w, \rho \rangle = \frac{1}{\lambda} \|w\|$ ; thus equality holds for all related inequalities:  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ ,  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$ , and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$ . ■

Last, when  $f$  is small in one dual norm and large in the other, many cases occur. The following theorem provides the various optimal solutions and their characterizations.

**Theorem B.8.** *If  $\|f\|_G \leq \frac{\mu}{\lambda}$  and  $\|\mathcal{P}f\|_s > \frac{\gamma}{\lambda}$ , then three case hold:*

- (1)  $u = 0$ ,  $\|\rho\|_G < \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ , and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$ .
- (2)  $v = 0$ ,  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s < \frac{\gamma}{\lambda}$ , and  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$ .
- (3)  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ ,  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$ , and  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$ .

All of the cases above are also the optimal decompositions for the energy.

*Proof.* Consider the simultaneous energy in terms of  $w$ :

$$\mathcal{E}(w) = \|w\| + \frac{\lambda}{2} \|f - w\|_2^2.$$

Under the assumptions of this theorem, the simultaneous dual norm of  $f$  is bounded below, i.e.,  $\|f\|' > \frac{1}{\lambda}$ , and by Lemma B.1,  $\|\rho\|' = \frac{1}{\lambda}$ , and  $\langle \rho, w \rangle = \frac{1}{\lambda} \|w\|$ . This can occur in three ways. The proof is divided into several subproofs for each of the cases.

1. *First assume that  $\|\rho\|_G < \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ .*

Since  $w = u + v$ , the result of Lemma B.1 is equivalent to

$$\langle \rho, u + v \rangle = \frac{1}{\lambda} \{ \mu \|Du\|_1 + \gamma \|\mathcal{P}v\|_* \}.$$

However, by the duality principles and the assumptions of this case, the following inequalities hold:

$$\begin{aligned} \langle \rho, u \rangle &\leq \|Du\|_1 \|\rho\|_G < \frac{\mu}{\lambda} \|Du\|_1, \\ \langle \rho, v \rangle &\leq \|\mathcal{P}v\|_* \|\mathcal{P}\rho\|_* = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*. \end{aligned}$$

Combining these statements produces the contradictory strict inequality

$$\langle \rho, u + v \rangle < \frac{1}{\lambda} \{ \mu \|Du\|_1 + \gamma \|\mathcal{P}v\|_* \}.$$

Thus  $u = 0$  and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$  to avoid the contradiction.

2. *Next, assume that  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s < \frac{\gamma}{\lambda}$ .*

By repeating the argument above,  $v = 0$   $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$  must hold.

3. *Assume that  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ ; then  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$ , and  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$  for the optimal solution.*

Similarly to the other two cases, by Lemma B.1, we have

$$\langle \rho, u + v \rangle = \frac{1}{\lambda} \{ \mu \|Du\|_1 + \gamma \|\mathcal{P}v\|_* \}.$$

The duality principles (Lemma B.3) and the assumptions of this case yield the following inequalities:

$$\begin{aligned}\langle \rho, u \rangle &\leq \|\rho\|_G \|Du\|_1 = \frac{\mu}{\lambda} \|Du\|_1, \\ \langle \rho, v \rangle &\leq \|\mathcal{P}\rho\|_s \|\mathcal{P}v\|_* = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*.\end{aligned}$$

Thus the equalities  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$  and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$  must hold.

4. For the next three cases, we show that the solutions are optimal. Assume  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s \leq \frac{\gamma}{\lambda}$ ,  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$ , and  $f = u + \rho$ ; then for all  $U$  and  $v$  we have

$$(B.1) \quad \mu \|D(u + U)\|_1 + \gamma \|\mathcal{P}v\|_* + \frac{\lambda}{2} \|\rho - U - v\|_2^2 \geq \mu \|Du\|_1 + \frac{\lambda}{2} \|\rho\|_2^2.$$

To show this, first rescale (B.1) by dividing by  $\lambda = \mu \|\rho\|_G^{-1}$  to get

$$\begin{aligned}& \|D(u + U)\|_1 \|\rho\|_G + \frac{\gamma}{\mu} \|\mathcal{P}v\|_* \|\rho\|_G + \frac{1}{2} \|\rho - U - v\|_2^2 \\& \geq \langle u + U, \rho \rangle + \frac{\gamma}{\mu} \|\mathcal{P}v\|_* \|\rho\|_G + \frac{1}{2} \|\rho\|_2^2 + \frac{1}{2} \|U + v\|_2^2 - \langle \rho, U \rangle - \langle v, \rho \rangle \\& \geq \langle u, \rho \rangle + \frac{\gamma}{\mu} \|\mathcal{P}v\|_* \|\rho\|_G + \frac{1}{2} \|\rho\|_2^2 + \frac{1}{2} \|U + v\|_2^2 - \langle v, \rho \rangle \\& = \frac{\mu}{\lambda} \|Du\|_1 + \frac{\gamma}{\mu} \|\mathcal{P}v\|_* \|\rho\|_G + \frac{1}{2} \|\rho\|_2^2 + \frac{1}{2} \|U + v\|_2^2 - \langle v, \rho \rangle \\& \geq \frac{\mu}{\lambda} \|Du\|_1 + \frac{1}{2} \|\rho\|_2^2.\end{aligned} \tag{B.2}$$

By the texture duality principle, i.e.,  $\langle v, \rho \rangle \leq \|\mathcal{P}\rho\|_s \|\mathcal{P}v\|_*$ , and the assumptions of this case, we have  $\langle v, \rho \rangle \leq \frac{\gamma}{\mu} \|\rho\|_G \|\mathcal{P}v\|_*$ , which is used in the inequalities above.

If equality holds, then  $U = -v$  and  $\langle v, \rho \rangle = \frac{\gamma}{\mu} \|\rho\|_G \|\mathcal{P}v\|_*$ . Also  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$  and  $\langle v, \rho \rangle = \|\mathcal{P}\rho\|_s \|\mathcal{P}v\|_*$  would hold. Returning to (B.1),

$$\mu \|D(u - v)\|_1 + \gamma \|\mathcal{P}v\|_* + \frac{\lambda}{2} \|\rho\|_2^2 \geq \mu \|Du\|_1 + \frac{\lambda}{2} \|\rho\|_2^2.$$

By canceling terms, we get

$$\mu \|D(u - v)\|_1 + \gamma \|\mathcal{P}v\|_* \geq \mu \|Du\|_1.$$

Using the equalities we found and the fact that  $\gamma = \lambda \|\mathcal{P}\rho\|_s$ , the following inequalities are equivalent:

$$\begin{aligned}\mu \|D(u - v)\|_1 + \lambda \|\mathcal{P}\rho\|_s \|\mathcal{P}v\|_* &\geq \mu \|Du\|_1, \\ \frac{\mu}{\lambda} \|D(u - v)\|_1 + \langle v, \rho \rangle &\geq \frac{\mu}{\lambda} \|Du\|_1, \\ \|\rho\|_G \|D(u - v)\|_1 + \langle v, \rho \rangle &\geq \frac{\mu}{\lambda} \|Du\|_1.\end{aligned}$$

However, by Lemma B.3,

$$\|\rho\|_G \|D(u - v)\|_1 \geq \langle u - v, \rho \rangle = \langle u, \rho \rangle - \langle v, \rho \rangle,$$

which implies  $\langle u, \rho \rangle = \frac{\mu}{\lambda} \|Du\|_1$ , so if we have equality, then  $\|\rho\|_G \|D(u - v)\|_1 = \langle u - v, \rho \rangle$ .

5. Assume that  $f = v + \rho$ ,  $\|\rho\|_G < \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ , and  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$ ; then the decomposition  $f = v + \rho$  is optimal.

As before, perturb  $v$  by  $V$ , with the corresponding energy:

$$(B.3) \quad \mathcal{E}(u, v + V) = \mu \|Du\|_1 + \gamma \|\mathcal{P}(v + V)\|_* + \frac{\lambda}{2} \|\rho - u - V\|_2^2.$$

By Lemma B.3 and the assumptions of this case, we have the following inequality:

$$\begin{aligned} \|\mathcal{P}(v + V)\|_* \|\mathcal{P}\rho\|_s &\geq \langle v + V, \rho \rangle \\ &\geq \langle v, \rho \rangle + \langle V, \rho \rangle \\ &= \frac{\gamma}{\lambda} \|\mathcal{P}v\|_* + \langle V, \rho \rangle. \end{aligned}$$

Since we assume  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ , the previous inequality is equivalent to

$$\|\mathcal{P}(v + V)\|_* \geq \|\mathcal{P}v\|_* + \frac{\lambda}{\gamma} \langle V, \rho \rangle.$$

Note that the decompositions of  $f$ , namely  $f = u + v + V + \rho$  and  $f = \rho + v$ , yield

$$\|u + V\|_2^2 = \|f - \rho - v\|_2^2 = 0.$$

Using this fact and expanding the  $L^2$  norm in (B.3) gives

$$(B.4) \quad \|\rho - u - V\|_2^2 = \|\rho\|_2^2 - 2 \langle u, \rho \rangle - 2 \langle V, \rho \rangle.$$

The second term on the right-hand side of (B.4) can be bounded by using Lemma B.3 and the assumption  $\|\rho\|_G < \frac{\mu}{\lambda}$ :

$$\lambda |\langle u, \rho \rangle| \leq \lambda \|Du\|_1 \|\rho\|_G \leq \mu \|Du\|_1.$$

Returning to (B.3), we have the following lower bound:

$$\begin{aligned} \mathcal{E}(u, v + V) &\geq \lambda |\langle u, \rho \rangle| + \gamma \|\mathcal{P}v\|_* + \lambda \langle V, \rho \rangle + \frac{\lambda}{2} \|\rho\|_2^2 - \lambda \langle u, \rho \rangle - \lambda \langle V, \rho \rangle \\ &\geq \gamma \|\mathcal{P}v\|_* + \frac{\lambda}{2} \|\rho\|_2^2 \\ &= \mathcal{E}(0, v). \end{aligned}$$

Thus the decomposition  $f = v + \rho$  is optimal.

6. Last, assume that  $\|\rho\|_G = \frac{\mu}{\lambda}$ ,  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ ,  $\langle \rho, v \rangle = \frac{\gamma}{\lambda} \|\mathcal{P}v\|_*$ , and  $\langle \rho, u \rangle = \frac{\mu}{\lambda} \|Du\|_1$  hold; then the decomposition is optimal.

Perturb the functions  $u$  by  $U$  and  $\rho$  by  $P$  with the perturbed energy:

$$(B.5) \quad \mathcal{E}(u + U, \rho + P) = \mu \|Du\|_1 + \gamma \|\mathcal{P}(v - U - P)\|_* + \frac{\lambda}{2} \|\rho + P\|_2^2.$$

Next, by combining the various assumptions in this case, we get  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda} > \|\mathcal{P}f\|_s$  and  $\|\rho\|_G = \frac{\mu}{\lambda} \geq \|f\|_G$ .

For the texture norm, Lemma B.3 yields

$$\|\mathcal{P}\rho\|_s \|\mathcal{P}(v - U - P)\|_* \geq \langle \rho, v - U - P \rangle.$$

By using this inequality and the assumption that  $\|\mathcal{P}\rho\|_s = \frac{\gamma}{\lambda}$ , the following holds:

$$(B.6) \quad \gamma \|\mathcal{P}(v - U - P)\|_* \geq \lambda \langle \rho, v \rangle - \lambda \langle \rho, U \rangle - \lambda \langle \rho, P \rangle.$$

With respect to the cartoon norm, Lemma B.3 yields

$$(B.7) \quad \|D(u + U)\|_1 \|\rho\|_G \geq \langle u + U, \rho \rangle.$$

By the assumption  $\|\rho\|_G = \frac{\mu}{\lambda}$ , (B.7) is equivalent to

$$(B.8) \quad \mu \|D(u + U)\|_1 \geq \lambda \langle u, \rho \rangle + \lambda \langle U, \rho \rangle.$$

Combining (B.7) and (B.8) with (B.5) and canceling terms gives the following lower bound for the energy:

$$\begin{aligned} \mathcal{E}(u + U, \rho + P) &= \mu \|Du\|_1 + \gamma \|\mathcal{P}(v - U - P)\|_* + \frac{\lambda}{2} \|\rho + P\|_2^2 \\ &\geq \lambda \left( \langle u, \rho \rangle + \langle U, \rho \rangle + \langle \rho, v \rangle - \langle \rho, U \rangle - \langle \rho, P \rangle + \frac{1}{2} \|\rho\|_2^2 + \frac{1}{2} \|P\|_2^2 + \langle P, \rho \rangle \right) \\ &\geq \lambda \left( \langle u + v, \rho \rangle + \frac{1}{2} \|\rho\|_2^2 + \frac{1}{2} \|P\|_2^2 \right) \\ &= \mu \|Du\|_1 + \gamma \|\mathcal{P}v\|_* + \lambda \left( \frac{1}{2} \|\rho\|_2^2 + \frac{1}{2} \|P\|_2^2 \right) \\ &\geq \mathcal{E}(u, \rho). \end{aligned}$$

This implies  $U = 0$  and  $P = 0$ ; therefore  $(u, \rho)$  is the minimizer. ■

**Acknowledgments.** We would like to thank Professor Luminita Vese for her help and time discussing cartoon-texture models. We would also like to thank the reviewers and the editors for their constructive comments.

## REFERENCES

- [1] G. AUBERT AND P. KORNPROBST, *Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations*, 2nd ed., Appl. Math. Sci. 147, Springer, New York, 2006.

- [2] J.-F. AUJOL, *Contribution à l'analyse de textures en traitement d'images par méthodes variationnelles et équations aux dérivées partielles*, Ph.D. thesis, Université de Nice Sophia Antipolis, Nice, France, 2004.
- [3] J.-F. AUJOL, G. AUBERT, L. BLANC-FÉRAUD, AND A. CHAMBOLLE, *Image decomposition into a bounded variation component and an oscillating component*, J. Math. Imaging Vision, 22 (2005), pp. 71–88.
- [4] J.-F. AUJOL AND T. F. CHAN, *Combining geometrical and textured information to perform image classification*, J. Vis. Commun. Image Rep., 17 (2006), pp. 1004–1023.
- [5] J.-F. AUJOL AND S. H. KANG, *Color image decomposition and restoration*, J. Vis. Commun. Image Rep., 17 (2006), pp. 916–928.
- [6] M. BERTALMIO, L. VESE, G. SAPIRO, AND S. OSHER, *Simultaneous structure and texture image inpainting*, IEEE Trans. Image Process., 12 (2003), pp. 882–889.
- [7] A. BUADES, B. COLL, AND J.-M. MOREL, *A review of image denoising algorithms, with a new one*, Multiscale Model. Simul., 4 (2005), pp. 490–530.
- [8] A. BUADES, B. COLL, AND J.-M. MOREL, *Neighborhood filters and PDE's*, Numer. Math., 105 (2006), pp. 1–34.
- [9] J.-F. CAI AND S. OSHER, *Fast Singular Value Thresholding without Singular Value Decomposition*, UCLA CAM Report 10-24, UCLA, Los Angeles, CA, 2010.
- [10] E. J. CANDÈS, X. LI, Y. MA, AND J. WRIGHT, *Robust principal component analysis?*, J. ACM, 58 (2011), 11.
- [11] A. CHAMBOLLE AND P.-L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
- [12] W. DONG, X. LI, L. ZHANG, AND G. SHI, *Sparsity-based image denoising via dictionary learning and structural clustering*, in Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011, pp. 457–464.
- [13] M. ELAD AND M. AHARON, *Image denoising via sparse and redundant representations over learned dictionaries*, IEEE Trans. Image Process., 15 (2006), pp. 3736–3745.
- [14] H. GAO, J.-F. CAI, Z. SHEN, AND H. ZHAO, *Robust principal component analysis-based four-dimensional computed tomography*, Phys. Med. Biol., 56 (2011), pp. 3181–3198.
- [15] H. GAO, H. YU, S. OSHER, AND G. WANG, *Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM)*, Inverse Problems, 27 (2011), 115012.
- [16] G. GILBOA AND S. OSHER, *Nonlocal linear image regularization and supervised segmentation*, Multiscale Model. Simul., 6 (2007), pp. 595–630.
- [17] G. GILBOA AND S. OSHER, *Nonlocal operators with applications to image processing*, Multiscale Model. Simul., 7 (2008), pp. 1005–1028.
- [18] J. GILLES, *Noisy image decomposition: A new structure, texture and noise model based on local adaptivity*, J. Math. Imaging Vision, 28 (2007), pp. 285–295.
- [19] J. GILLES AND Y. MEYER, *Properties of BV–G structures + textures decomposition models. Application to road detection in satellite images*, IEEE Trans. Image Process., 19 (2010), pp. 2793–2800.
- [20] T. GOLDSTEIN AND S. OSHER, *The split Bregman method for L1-regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.
- [21] H. JI, C. LIU, Z. SHEN, AND Y. XU, *Robust video denoising using low rank matrix completion*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010, pp. 1791–1798.
- [22] Y. KIM AND L. VESE, *Image recovery using functions of bounded variation and Sobolev spaces of negative differentiability*, Inverse Probl. Imaging, 3 (2009), pp. 43–68.
- [23] S. LAZEBNIK, C. SCHMID, AND J. PONCE, *A sparse texture representation using local affine regions*, IEEE Trans. Pattern Anal. Mach. Intell., 27 (2005), pp. 1265–1278.
- [24] L. H. LIEU AND L. A. VESE, *Image restoration and decomposition via bounded total variation and negative Hilbert–Sobolev spaces*, Appl. Math. Optim., 58 (2008), pp. 167–193.
- [25] J. MAIRAL, F. BACH, J. PONCE, AND G. SAPIRO, *Online learning for matrix factorization and sparse coding*, J. Mach. Learn. Res., 11 (2010), pp. 19–60.
- [26] J. MAIRAL, F. BACH, J. PONCE, G. SAPIRO, AND A. ZISSERMAN, *Non-local sparse models for image restoration*, in Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, 2009, pp. 2272–2279.

- [27] J. MAIRAL, M. ELAD, AND G. SAPIRO, *Sparse representation for color image restoration*, IEEE Trans. Image Process., 17 (2008), pp. 53–69.
- [28] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, AMS, Providence, RI, 2001.
- [29] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, AND W. YIN, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.
- [30] S. OSHER, A. SOLÉ, AND L. VESE, *Image decomposition and restoration using total variation minimization and the  $H^{-1}$  norm*, Multiscale Model. Simul., 1 (2003), pp. 349–370.
- [31] Y. PENG, A. GANESH, J. WRIGHT, W. XU, AND Y. MA, *RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images*, in Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010, pp. 763–770.
- [32] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
- [33] L. A. VESE AND S. J. OSHER, *Modeling textures with total variation minimization and oscillating patterns in image processing*, J. Sci. Comput., 19 (2002), pp. 553–572.
- [34] X. ZHANG, M. BURGER, X. BRESSON, AND S. OSHER, *Bregmanized nonlocal regularization for deconvolution and sparse reconstruction*, SIAM J. Imaging Sci., 3 (2010), pp. 253–276.
- [35] Z. ZHANG, A. GANESH, X. LIANG, AND Y. MA, *TILT: Transform invariant low-rank textures*, Int. J. Comput. Vision, 99 (2012), pp. 1–24.
- [36] M. ZHOU, H. CHEN, J. PAISLEY, L. REN, G. SAPIRO, AND L. CARIN, *Non-parametric Bayesian dictionary learning for sparse image representations*, in Proceedings of the 23rd Annual Conference on Neural Information Processing Systems, 2009, pp. 2295–2303.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.