UniHIST: A Unified Framework for Image Restoration With Marginal Histogram Constraints

Xing Mei[1,2] Weiming Dong[2] Bao-Gang Hu[2] Siwei Lyu[1]
[1]Computer Science Department, University at Albany, SUNY
[2]NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China
xmei@albany.edu, {wmdong,hubg}@nlpr.ia.ac.cn, slyu@albany.edu

Abstract

Marginal histograms provide valuable information for various computer vision problems. However, current image restoration methods do not fully exploit the potential of marginal histograms, in particular, their role as ensemble constraints on the marginal statistics of the restored image. In this paper, we introduce a new framework, UniHIST, to incorporate marginal histogram constraints into image restoration. The key idea of UniHIST is to minimize the discrepancy between the marginal histograms of the restored image and the reference histograms in pixel or gradient domains using the quadratic Wasserstein (W_2) distance. The W_2 distance can be computed directly from data without resorting to density estimation. It provides a differentiable metric between marginal histograms and allows easy integration with existing image restoration methods. We demonstrate the effectiveness of UniHIST through denoising of pattern images and non-blind deconvolution of natural images. We show that UniHIST enhances restoration performance and leads to visual and quantitative improvements over existing state-of-the-art methods.

1. Introduction

A histogram is a discrete approximation of the probability distribution of a continuous random variable. For high dimensional signals such as images, when one treats individual data points as independent samples from a random source, the corresponding histogram is known as a *marginal histogram* [11]. In image processing and computer vision, marginal histograms in pixel or gradient domains provide a simple and effective summary of image statistics. For instance, intensity and color histograms capture the contrast and tone variations of the corresponding image, which are widely used as features in image enhancement [15] and retrieval [16]. Important image feature descriptors such as SIFT [23] and HOG [9] are in the form of histograms of oriented gradients. Matching histograms of band-pass filter responses is the essential step in Heeger and Bergen's influential method for texture analysis and synthesis [17]. Marginal histograms of image gradients and wavelet coefficients show sharp peaks around zero and heavy tails, which inspired many parametric image models such as Gaussian scale mixtures [31, 33] and hyper-Laplacian [13, 19]. These models have achieved state-of-the-art performance for various image restoration tasks [20, 21, 38].

For most vision applications, a marginal histogram is interpreted as an approximate description of the underlying distribution, specifying the probability of the data points taking a particular value. However, a marginal histogram collected from an image provides more information as it represents an *ensemble* constraint on the image, specifying the proportion of the pixels taking each value. This introduces a negative correlation between the data points - if the marginal histogram specifies that 10% of the pixels take a certain value, this means that the remaining 90% pixels can only take other values. Therefore a marginal histogram provides a constraint on the marginal statistics of the output image. Similar is true for marginal histograms in other transform domains of the images. Several recent studies have shown that image restoration tasks can benefit from such ensemble constraints, but incorporating these constraints with existing restoration methods in a numerically stable manner remains a challenging problem [7, 30, 35, 39].

In this work, we emphasize the aspect of marginal histograms as ensemble constraints and introduce a unified framework, UniHIST, to incorporate such constraints in image restoration problems. In UniHIST, we use the *quadratic Wasserstein* (W_2) metric [32] (also known as Cramer-von Mises distance) to measure the statistical distance between the marginal histograms of the output image and the reference histogram. The W_2 distance can be computed directly from data without relying on density estimation, providing a smooth and differentiable form to measure the dissimilarity between histograms. By including this term in an optimization framework, UniHIST can readily work with most

existing restoration methods. We demonstrate the benefits of marginal histograms and the effectiveness of UniHIST through two applications: denoising of pattern images and non-blind deconvolution of natural images. We show that UniHIST enhanced restoration algorithms lead to improved restoration quality over existing state-of-the-art methods.

2. Related Works

The task of image restoration is to recover an image from its degraded observation [1]. A complete survey of the image restoration techniques is beyond the scope of this paper, and we only review methods that enforce constraints on the marginal histograms of the output in the restoration process.

In the context of Markov Random Field (MRF), Zhu *et al*. proposed a FRAME model which adapts a Gibbs sampler to generate texture images with similar marginal histograms as the example image [36]. Schmidt *et al*. employed a more efficient Gibbs sampler and achieved state-of-the-art denoising performance with this generative model [30]. These sampling-based methods require multiple samples to predict an output, which might be computationally expensive in practice. Woodford *et al*. showed that MRF inherently encourages delta function marginal histograms, which might introduce statistical bias with weak data fidelity terms [35]. They extended MRF into a marginal probability field (MPF) model, which enforces marginal histogram constraints for a range of applications, including segmentation, denoising and texture synthesis. However, their model is defined in the discrete label space, which can not be easily extended to handle image restoration tasks with continuous intensity values.

Cho *et al*. proposed to preserve the heavy-tailed marginal gradient distributions in the deconvolution process, which greatly enhances the visual details in the restored image [7]. However, their method minimizes the KL divergence between the output image gradients and the reference distribution, which requires parametric density fitting and a complex reweighting optimization scheme. Recently, Zuo *et al*. proposed a texture-enhanced denoising method, which combines a non-local patch-based image model with a heuristic gradient penalty term [39]. UniHIST differs from these methods in several aspects: first, our method minimizes the histogram discrepancy using the W_2 distance, which leads to consistent objective functions with provable convergence and efficient numerical solutions; second, our method provides a unified procedure to integrate marginal histogram constraints with a wide range of existing denoising and deconvolution algorithms; finally, our method works directly with nonparametric histograms and does not need to fit them with parametric models.

The Wasserstein distance and its variants have been used to find perceptually similar contents in image retrieval [29]. Recently, Peyré and his collaborators proposed the sliced Wasserstein distance for multi-dimensional signals, and applied this measure in texture mixing [3] and color transfer [27], but no previous studies have explored this metric for general image restoration tasks.

3. Method

In this section, we describe a unified approach to utilize marginal histogram constraints for various image restoration problems. The basic intuition is that if we have knowledge of the uncorrupted image in certain domains in the form of marginal histograms, the desirable output image from the restoration algorithm should respect such constraints. Following this intuition, we first introduce the quadratic Wasserstein (W_2) distance as our dissimilarity measure between probability distributions or histograms (Section 3.1). We then present the UniHIST framework in Section 3.2. Section 3.3 briefly discusses several ways to obtain the corresponding reference marginal histograms.

3.1. Quadratic Wasserstein Distance

Given two probability measures p, q over the real line $\mathcal{R}$ with finite variances, their quadratic Wasserstein (W_2) distance is defined as the variational solution to the following Monge problem [32]:

$$W_2^2(p,q) = \inf_{\phi} \int_{-\infty}^{\infty} (x - \phi(x))^2 p(x) dx \qquad (1)$$

where the infimum is over all deterministic functions $\phi : \mathcal{R} \mapsto \mathcal{R}$ that transfer an arbitrary random variable x with distribution p to a new variable $\phi(x)$ with distribution q. In other words, the W_2 distance is the minimum ℓ_2 disturbance caused by any mapping functions between the two given distributions. Unlike the Kullback-Leibler divergence, the W_2 distance is a true metric between probability distributions, and becomes zero if and only if p and q are identical.

For a given probability distribution p over the real line, we define its *cumulative distribution function* as $F_p(x) = \int_{-\infty}^{x} p(\tau) d\tau$, and its *percentile function* (the pseudo-inverse of F_p) as $F_p^{-1}(t) = \inf(x : F_p(x) > t\}$. A fundamental result in the optimal transport theory[1] is that the optimal ϕ minimizing Eq.(1) has a closed-form solution [32]:

$$\phi_{p \to q}(x) = F_q^{-1}(F_p(x)) \,. \qquad (2)$$

For our framework, we need to measure the statistical distance between the marginal histogram $h_{\mathbf{x}}$ of an image $\mathbf{x}$ with n pixels and a given marginal histogram h_q. Most histogram distances (Chi-Square, KL divergence, *etc*.) directly compare two histogram vectors [5]. However, the process of extracting $h_{\mathbf{x}}$ from $\mathbf{x}$ is a non-differentiable operation on $\mathbf{x}$, which hinders their usage in image restoration. Instead,

[1] An accessible and comprehensive introduction of optimal transport theory can be found in [32].

we use the equivalent definition of the W_2 distance, by treating $\mathbf{x} = (x_1, \cdots, x_n)^T$ as n independent samples drawn from a distribution p, and h_q as the discrete approximation of another distribution q. Following Eq.(1), we introduce an empirical $\hat{W}_2$ measure as follows:

$$\hat{W}_2^2(p, q) = \hat{W}_2^2(h_{\mathbf{x}}, h_q) = \min_{\phi} \frac{1}{n} \sum_{i=1}^{n} (x_i - \xi_i)^2 \quad (3)$$

where function ϕ maps x_i to $\xi_i = \phi(x_i)$, such that the transformed samples $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)^T$ satisfy the marginal histogram h_q. Analogous to the continuous case, the optimal $\hat{\phi}$ for Eq.(3) is given as [32]:

$$\xi_i = \hat{\phi}_{h_{\mathbf{x}} \to h_q}(x_i) = F_{h_q}^{-1}(F_{h_{\mathbf{x}}}(x_i)) \quad (4)$$

where $F_{h_q}^{-1}$ and $F_{h_{\mathbf{x}}}$ are the percentile function and the cumulative distribution function constructed from h_q and $h_{\mathbf{x}}$ respectively. A close look at Eq.(4) shows that, for image $\mathbf{x}$, $\hat{\phi}$ actually corresponds to the "histogram matching" operation used to improve the image contrast [15]. A notable property of the empirical W_2 measure is that it can be evaluated in the quadratic form of $\mathbf{x}$. On the other hand, for other measures such as the KL divergence, the direct comparison between two histograms is a non-smooth operation on $\mathbf{x}$, which makes them difficult to work with existing restoration algorithms. Equations (3) and (4) form the basis of the coming UniHIST framework.

3.2. The UniHIST Framework

We use vectorized notations for images and adopt the image formation model as:

$$\mathbf{y} = A\mathbf{x} + \mathbf{n} \quad (5)$$

where $\mathbf{y}$ is the observed image, $\mathbf{x}$ is the original uncorrupted image, $\mathbf{n}$ is the additive Gaussian noise with known variance, and the matrix A subsumes the linear transforms in the imaging process such as blurring and downsampling. We are mainly interested in problems that can be cast into the following form:

$$\min_{\mathbf{x}} \frac{\lambda}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \Gamma(\mathbf{x}) \quad (6)$$

where the first term (the data fidelity term) represents the ℓ_2 difference between the observation $\mathbf{y}$ and the image formation model $A\mathbf{x}$, and $\Gamma(\mathbf{x})$ (the regularization term) represents our prior knowledge about the original image $\mathbf{x}$. Popular $\Gamma(\mathbf{x})$ for image restoration include heavy-tail probabilistic models in the gradient domain such as Gaussian scale mixtures and generalized Laplacian models [20, 21] and probabilistic models of image patches [10, 37]. Parameter λ controls the balance of the two terms. Eq.(6) covers a wide range of restoration tasks, such as demosaicing, denoising, non-blind deconvolution, and super-resolution [18].

Suppose we have knowledge about the clean image $\mathbf{x}$, specifying the marginal histogram of $\mathbf{x}$ in a linearly transformed domain such as gradients or wavelet coefficients. We would like to use this information in the restoration method. To this end, we add a marginal histogram constraint to Eq.(6) and solve the following constrained minimization problem:

$$\min_{\mathbf{x}} \frac{\lambda}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \Gamma(\mathbf{x}) \quad \text{s.t.} \quad h_{B\mathbf{x}} = h_r \quad (7)$$

where h_r is the reference histogram, and the matrix B denotes the corresponding linear transformation. In Section 4, we will discuss different choices of such domains for specific applications. Constraints on marginal histograms in more than one transformation domains can be incorporated similarly, but for simplicity, we only discuss the case when one domain (a single B) is used.

Using the $\hat{W}_2$ measure and the *penalty function* method [2], we convert Eq.(7) to the following unconstrained minimization problem:

$$\min_{\mathbf{x}} \frac{\lambda}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \Gamma(\mathbf{x}) + \frac{\beta}{2} \hat{W}_2^2(h_{B\mathbf{x}}, h_r) \quad (8)$$

With the penalty weight $\beta \to \infty$, the local optimal solution of Eq.(8) reaches a local optimal solution of Eq.(7) [2].

Following the definition in (3), we introduce an auxiliary vector $\boldsymbol{\xi}$ (of the same dimension as $B\mathbf{x}$) and minimize the following objective function:

$$\min_{\mathbf{x}, \boldsymbol{\xi}} \frac{\lambda}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \frac{\beta}{2} \|\boldsymbol{\xi} - B\mathbf{x}\|_2^2 + \Gamma(\mathbf{x}) \text{ s.t. } h_{\boldsymbol{\xi}} = h_r \quad (9)$$

This problem, as the final form of UniHIST, can be readily solved with block coordinate descent by alternatively minimizing Eq.(9) with regards to $\boldsymbol{\xi}$ and $\mathbf{x}$ respectively. This algorithm is guaranteed to converge to a local minimum.

Fixing $\boldsymbol{\xi}$, the $\mathbf{x}$ problem comes with two quadratic terms and a regularization term on $\mathbf{x}$:

$$\min_{\mathbf{x}} \frac{\lambda}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \frac{\beta}{2} \|\boldsymbol{\xi} - B\mathbf{x}\|_2^2 + \Gamma(\mathbf{x}) \quad (10)$$

If $\Gamma(\mathbf{x})$ is also a quadratic term, Eq. (10) reduces to an unconstrained least squares problem, which can be solved with conjugate gradient or FFT-based methods [20, 21]. If $\Gamma(\mathbf{x})$ is a non-smooth term, which is often the case for image restoration, we can further decompose Eq. (10) into subproblems with half-quadratic splitting [14] or proximal algorithms [24]. For these subproblems, we can reuse some well-developed solutions for specialized $\Gamma(\mathbf{x})$ models. This is the key component to incorporate UniHIST into existing image restoration methods. We show such an example with patch-based models in Sec. 4.2.

Fixing $\mathbf{x}$, minimizing $\boldsymbol{\xi}$ reduces to

$$\min_{\boldsymbol{\xi}} \frac{\beta}{2} \|\boldsymbol{\xi} - B\mathbf{x}\|_2^2, \text{ s.t. } h_{\boldsymbol{\xi}} = h_r \quad (11)$$

Using the results in (4), the optimal $\boldsymbol{\xi}$ is given by the histogram matching operation as follows:

$$\xi_i = \hat{\phi}_{h_{B\mathbf{x}} \to h_r}((B\mathbf{x})_i) \tag{12}$$

where i denotes the i-th element. To sum up, UniHIST iteratively performs two main steps until convergence: update $\boldsymbol{\xi}$ through Eq.(12) and then update $\mathbf{x}$ through Eq.(10). Note that the penalty weight β may vary in the iterative computation process: if we want to enforce a strict marginal histogram constraint, we can start with a small β value and increase it by a scaling factor after each iteration.

3.3. Reference Marginal Histogram Generation

Although UniHIST provides a unified approach to enforce marginal histogram constraints in image restoration tasks, the generation methods of the reference histograms might vary from one application to another. Here we consider several different methods for marginal histogram generation: (1) Marginal histograms can be specified with external sources. Woodfood *et al*. proposed to build a histogram database with natural images and use the closest match for image denoising [35]. (2) Marginal histograms can be specified through user interaction and parametric models. It is well known that the marginal gradient histograms of natural images follow a hyper-Laplacian model. By tuning the parameters of the model, it is possible to get a good estimation of the original histograms [6]. (3) For some very specific problems, it is possible to directly estimate the marginal histogram from the degraded observation. For images corrupted with small Gaussian noise, the gradient histograms can be roughly estimated as a 1D deconvolution process [39]. Finally, histograms might be provided along with the observation. For example, marginal histograms can be embedded in the image, for instance as a "fingerprint" of the authenticity or quality of the image [34].

4. Applications

In this section, we describe two applications of the proposed UniHIST framework: denoising of pattern images and non-blind deconvolution of natural images. We show that UniHIST can be integrated with some well-developed restoration methods to improve the output image quality.

4.1. Pattern Image Denoising

Image denoising has been extensively studied in the computer vision and image processing community, with many state-of-the-art techniques, including Non-Local Means [4], BM3D [8] and KSVD [10]. Natural images exhibit some common properties such as sparsity and nonlocal self-similarity, which have been heavily exploited by existing denoising methods. However, less attention has been paid to "pattern images" with regular or near-regular structures, which can be ubiquitously found in man-made objects and artistic designs. Some real-world pattern image examples are presented in Fig. 1a. An important property of the pattern images is that their marginal intensity histograms usually center around selected intensity levels (see the blue curves in Fig. 1e), which can be estimated from similar pattern images or directly specified in the design process. We validate this observation with a simple test: we randomly sample square patches from high resolution pattern images, and for each example, all patches share very close intensity histograms when patch size goes beyond a certain limit. For this application, we assume the marginal intensity histogram of the pattern is provided along with the noisy observation. UniHIST provides a natural way to exploit this extra information for pattern image denoising.

Model We make two assumptions: the observation image $\mathbf{y}$ is corrupted by an additive white Gaussian noise with variance σ^2; the intensity distribution of the latent image follows a reference histogram h_r. Following the definition of UniHIST (Eq.(9)), we introduce an auxiliary vector $\boldsymbol{\xi}$ and formulate the denoising problem as follows:

$$\min_{\mathbf{x},\boldsymbol{\xi}} \tfrac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{x}\|_2^2 + \tfrac{\beta}{2}\|\boldsymbol{\xi}-\mathbf{x}\|_2^2 + \Gamma(\mathbf{x}) \text{ s.t. } h_{\boldsymbol{\xi}} = h_r \tag{13}$$

where Γ is a selected denoising regularization term.

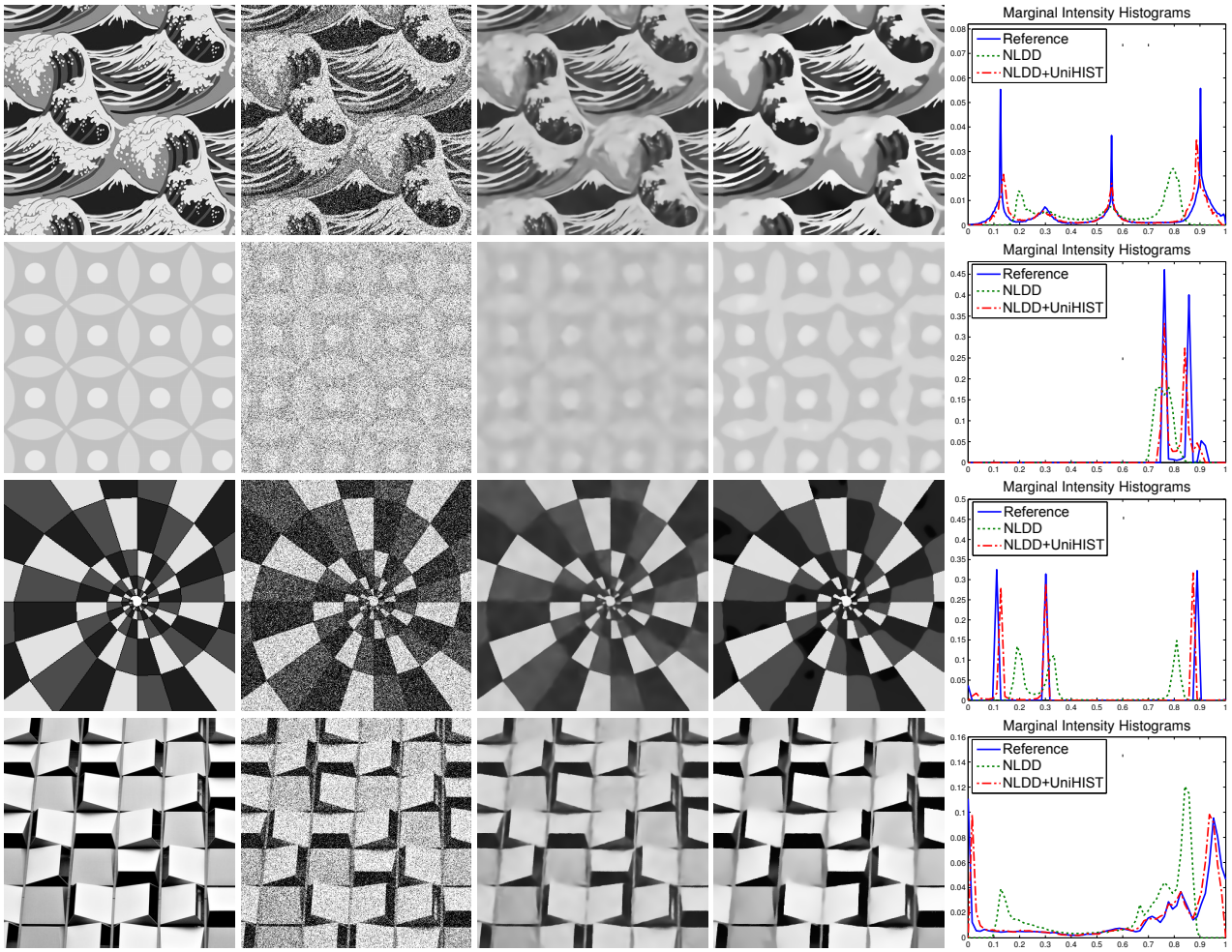
Solution As presented in Sec. 3.2, we solve this problem with block coordinate descent. The $\boldsymbol{\xi}$ step is solved with a histogram matching operation; while the $\mathbf{x}$ step reduces to the following problem:

$$\min_{\mathbf{x}} \tfrac{1}{2\hat{\sigma}^2}\|\hat{\mathbf{y}}-\mathbf{x}\|_2^2 + \Gamma(\mathbf{x}) \tag{14}$$

where $\hat{\mathbf{y}} = \frac{\mathbf{y}+\beta\sigma^2\boldsymbol{\xi}}{1+\beta\sigma^2}$ is a weighted interpolation between $\mathbf{y}$ and $\boldsymbol{\xi}$, and $\hat{\sigma} = \frac{\sigma}{\sqrt{1+\beta\sigma^2}}$. Considering $\hat{y}$ as the observation and $\hat{\sigma}^2$ as the noise variance, Eq. (14) is a standard Gaussian denoising problem which can be readily solved with any existing denoisers. For this application, we choose the Non-Local Dual Denoising method (NLDD), which has been recently reported with state-of-the-art denoising performance and fast processing speed [25]. After each iteration, β is scaled by a factor of 2.

Results We evaluate our algorithm on eight pattern images collected from the internet. For each example, we scale the image to the range $[0, 1]$ and degrade the clean image with a strong Gaussian noise ($\sigma^2 = 0.1$), such that the quality of the output results relies heavily on prior knowledge Γ and h_r. For h_r, we use a 256-bin histogram for all the tests.

Four examples and their output results with the original NLDD method and our method are presented in Fig. 1. Please refer to the supplemental material for the complete results. Compared to NLDD, our method enforces the output marginal intensity histogram to be more consistent with the reference histogram (see Fig. 1e), which in turn improves the visual quality in several aspects: first, the refer-



(a) Pattern Image (b) Noisy Observation (c) NLDD [25] (d) NLDD+UniHIST (e) Intensity Histograms

Figure 1: The output results of NLDD and our method (NLDD+UniHIST) on four pattern image examples. Our method enforces the output marginal intensity histogram to be more consistent with the reference histogram (see the last column), and improves the denoising quality in various regions. Details are best viewed in electronic form.

ence histogram provides the correct locations of the intensity modes, which helps to preserve the contrast of the pattern and recover bright and dark regions better than NLDD; second, the reference histogram provides an ensemble constraint on the proportion of pixels at each intensity level, which helps to suppress the outliers locating at wrong intensity levels in noisy environment; finally, in some special cases where the intensity modes are close to each other (see the second example), NLDD might blur image structures, while our method recovers better region boundaries.

We provide the quantitative comparison of the two methods in terms of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index in Fig. 2. As these results show, our method outperforms NLDD in both metrics. Following the same framework, we replaced NLDD with the classic BM3D denoiser [8] and repeated the experiments with the same settings. Further results (included in the supplemental material) confirm similar performance improvement with marginal histogram constraints.

4.2. Non-Blind Deconvolution

Assuming a known blurring kernel $\mathbf{k}$ and noise levels, non-blind image deconvolution tries to recover a latent image $\mathbf{x}$ from a blurry and noisy observation $\mathbf{y} = \mathbf{x} \otimes \mathbf{k} + \mathbf{n}$. This is inherently a more ill-posed problem than image denoising due to the coexistence of $\mathbf{k}$ and $\mathbf{n}$. Most state-of-the-art methods rely on parametric or semi-parametric models learned for natural images, such as the hyper-Laplacian gradient model [21], Field-of-Experts [28] and EPLL [37]. It has been noticed that most parametric models, when incorporated in a maximum-a-posteriori (MAP) framework, usually favor smooth solutions with low energy

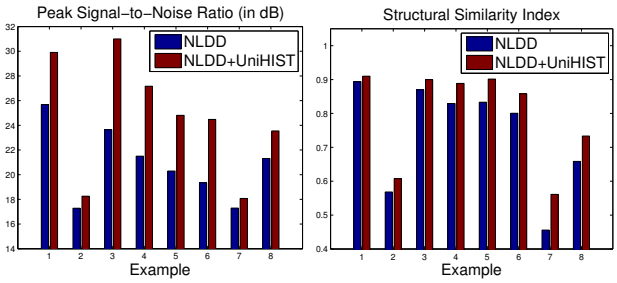


Figure 2: Quantitative comparison of NLDD and our method (NLDD+UniHIST) in terms of PSNR and SSIM on eight pattern image examples. Compared to NLDD, our method improves the average PSNR and the average SSIM index by 3.86dB and 0.057 respectively.

settings [7, 22, 30, 35], leading to statistical bias and loss of mid-level details in the output results. Zuo *et al*. proposed an effective method to enhance the textures of the restored image, but their solution is specialized for image denoising [39]. Cho *et al*. proposed a detailed-enhanced non-blind deconvolution method, which significantly enhances the visual details by preserving the marginal gradient distribution of the latent image in the deconvolution process [7]. However, their iterative distribution reweighting method (denoted as IDR) directly minimizes the KL divergence between the output image gradients and the reference distribution, which requires a complex optimization scheme. UniHIST provides a much simpler way for detailed-enhanced non-blind deconvolution.

Gradient Histogram Estimation We estimate the marginal gradient histograms of the latent image using a similar method as in [6, 7]: we deconvolve the corrupted image with a weak regularization term, then down-sample the restored image, calculate gradient components and extract marginal gradient histograms for horizontal and vertical components respectively.

Model We make two assumptions: the observation image $\mathbf{y}$ is corrupted by a known spatially-invariant kernel $\mathbf{k}$ and an additive white Gaussian noise with variance σ^2; the horizontal and vertical gradient components of the latent image should follow two reference gradient histograms h_1, h_2 respectively. Following the UniHIST framework, we introduce two auxiliary vectors $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ and formulate the deconvolution problem as follows:

$$\begin{aligned} &\min_{\mathbf{x},\boldsymbol{\xi}_1,\boldsymbol{\xi}_2} \tfrac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{k}\otimes\mathbf{x}\|_2^2+\tfrac{\beta}{2}\sum_{j=1}^{2}\|\boldsymbol{\xi}_j-f_j\otimes\mathbf{x}\|_2^2+\Gamma(\mathbf{x}) \\ &\text{s.t.} \quad h_{\boldsymbol{\xi}_j}=h_j, j=1,2 \end{aligned} \tag{15}$$

where $f_1=[1,-1], f_2=[1,-1]^T$ are gradient filters along horizontal and vertical directions respectively, and Γ represents an existing image model. For this application, we employ the popular EPLL patch model [37]:

$$\Gamma(\mathbf{x}) = -\sum_i log(p(P_i\mathbf{x})) \tag{16}$$

where the matrix P_i extracts the i-th patch from $\mathbf{x}$ out of all overlapping patches, and p represents a Gaussian mixture model (GMM) learned from natural image patches. The patch size is fixed to 8×8 for all the examples.

Solution We iteratively update $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ and $\mathbf{x}$ in two main steps. Given fixed $\mathbf{x}$, we update $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ with the reference histograms h_1, h_2:

$$(\boldsymbol{\xi}_j)_i = \hat{\phi}_{h_{f_j\otimes\mathbf{x}}\to h_j}((f_j\otimes\mathbf{x})_i), j=1,2 \tag{17}$$

where i denotes the i-th element. The $\mathbf{x}$ problem involves two quadratic terms on $\mathbf{x}$ and a regularization term on overlapping patches $P_i\mathbf{x}$:

$$\min_{\mathbf{x}} \tfrac{1}{2\sigma^2}\|\mathbf{y}-K\mathbf{x}\|_2^2+\tfrac{\beta}{2}\sum_{j=1}^{2}\|\boldsymbol{\xi}_j-F_j\mathbf{x}\|_2^2-\sum_i log(p(P_i\mathbf{x})) \tag{18}$$

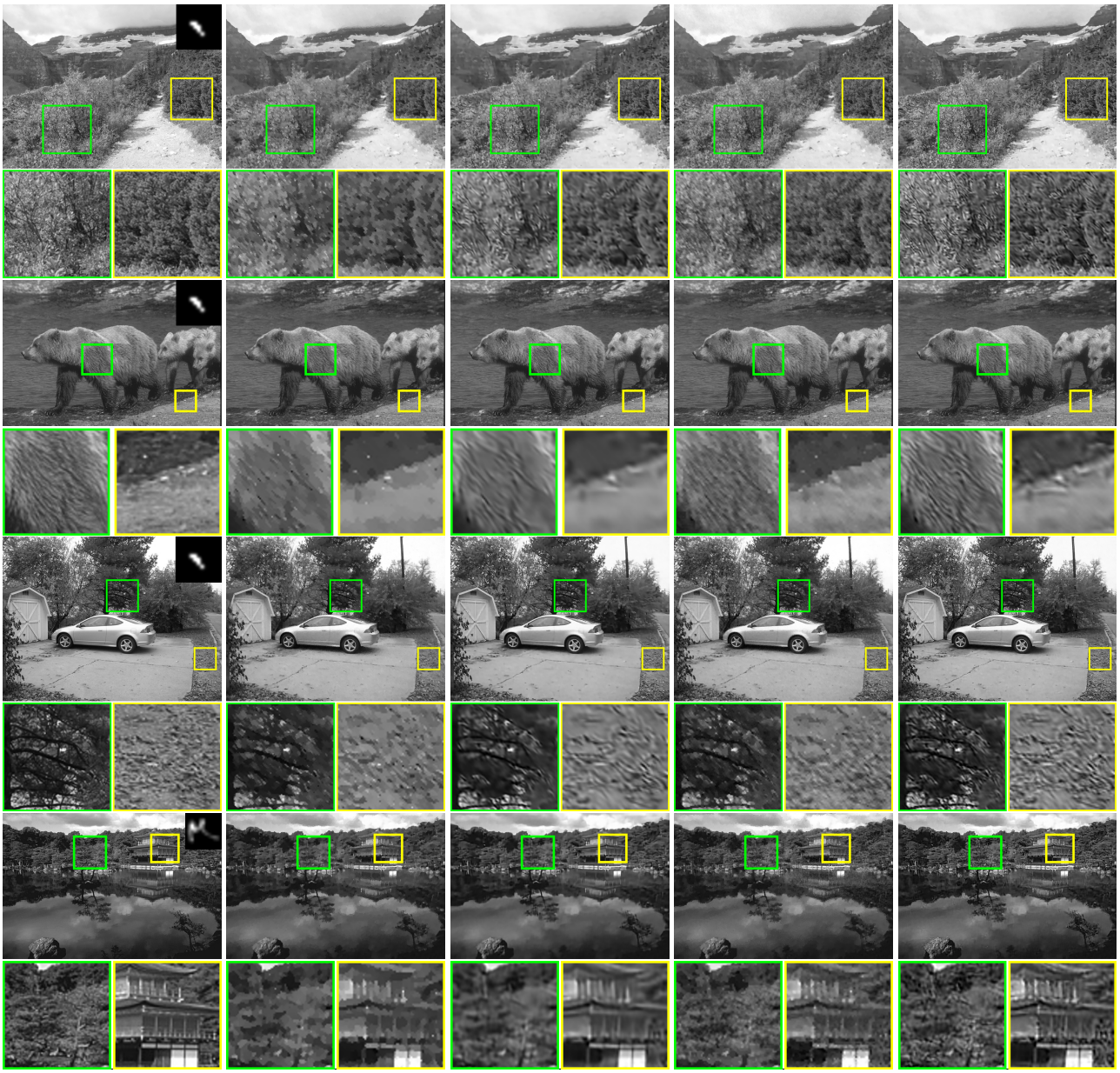
where K, F_1, F_2 are the block Toeplitz matrix representations of $\mathbf{k}, f_1, f_2$ respectively. We then solve the $\mathbf{x}$ problem with a half-quadratic splitting scheme [14, 37]. We introduce a set of patches $\mathbf{z}_i$, one for each patch $P_i\mathbf{x}$, and minimize the following objective function:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z}_i} \quad & \tfrac{1}{2\sigma^2}\|\mathbf{y}-K\mathbf{x}\|_2^2+\tfrac{\beta}{2}\sum_{j=1}^{2}\|\boldsymbol{\xi}_j-F_j\mathbf{x}\|_2^2 \\ & +\tfrac{\gamma}{2}\sum_i\|\mathbf{z}_i-P_i\mathbf{x}\|_2^2-\sum_i log(p(\mathbf{z}_i)) \end{aligned} \tag{19}$$

where γ gradually increases in an iterative optimization process. Now the $\mathbf{x}$ subproblem is a least squares problem, which can be solved with conjugate gradient, while the $\mathbf{z}_i$ subproblem can be solved with an approximate MAP estimation scheme from the original EPLL method [37].

Results We test our algorithm on six examples from [7]. The first two example images are degraded with a simple blurring kernel and 7% Gaussian noise; examples 3 and 4 are degraded with the same kernel and 5% noise; while for the last two examples, the images are degraded with a complicated kernel and 2% noise.

We present the deconvolution results of the EPLL method and our method for examples 1, 2, 3 and 6 in Fig. 3. For comparison, we also include the results generated by a gradient-based hyper-Laplacian prior and the IDR method from [7]. Please refer to the supplemental material for the complete results. As can be seen from all the examples, our method significantly enhances the visual details over the original EPLL method in various regions, while IDR improves over the original hyper-Laplacian model. Both IDR and our method successfully retrieve some stochastic texture details in regions covered by trees, sand and animal furs; while our method seems to perform better in regions



(a) Original Image (b) hyper-Laplacian (c) EPLL [37] (d) IDR [7] (e) EPLL+UniHIST

Figure 3: The deconvolution results of four methods (hyper-Laplacian, EPLL, IDR and our method) for examples 1, 2, 3, 6. Our method significantly enhances the visual details over the original EPLL method in various regions. See close-up regions for visual detail comparison. This figure is best viewed in electronic form.

with fine directional structures (see close-up regions of examples 2 and 6).

We evaluate the symmetric KL divergence of the marginal gradient histograms between the restored image and the original image for the six examples in Fig. 4: for all the examples, our method provides the lowest KL divergence for both horizontal and vertical gradients, which demonstrates the effectiveness of the UniHIST framework; IDR also gives lower KL divergence than hyper-Laplacian and EPLL. The evaluation with the empirical W_2 metric also reports similar quantitative results. We provide the marginal gradient histograms (in logarithmic scale, only the positive half) of example 6 in Fig. 5, which shows our method enforces the marginal histogram constraints better than the other methods, and therefore preserves more details in the deconvolution process.

We provide the PSNR and SSIM values of the six examples in Fig. 6, which shows some interesting properties

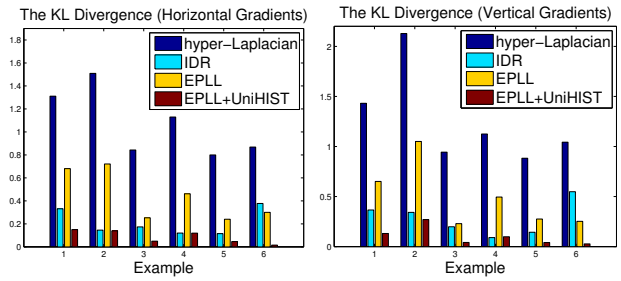


Figure 4: The symmetric KL divergence of the horizontal and vertical gradient histograms between the restored image and the original image for the six examples. For horizontal gradient histograms, the average KL divergence of hyper-Laplacian, IDR, EPLL and our method are $1.08, 0.21, 0.44, 0.09$ respectively; while for vertical gradient histograms, the average KL divergence of the four methods are $1.26, 0.28, 0.49, 0.10$ respectively.

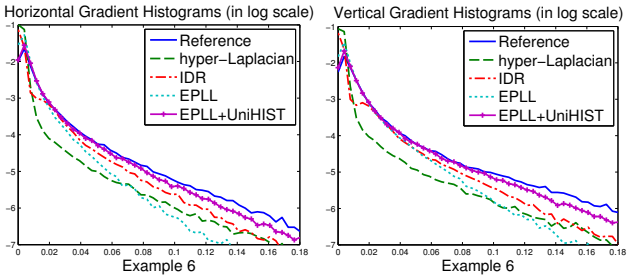


Figure 5: The marginal gradient histograms (in logarithmic scale) of example 6 produced by different methods. Considering the rough symmetry of the gradient distribution, we only show the curves along the positive real line.

of our algorithm. It can be expected that EPLL and our method show better quantitative performance than hyper-Laplacian and IDR, since we employ a patch-based image model which is more powerful than the gradient sparsity model. Compared to EPLL, our method gives lower PSNR values for all the examples but higher SSIM values for five examples. This is mostly due to the fact that for texture regions, UniHIST "hallucinates" textures perceptually similar to the original textures, while EPLL and other sparsity-based algorithms try to "reconstruct" the original textures. While our method can produce more visually appealing restoration results (reflected by the SSIM values of five examples), when evaluated by mean squared error based metrics such as PSNR, it does not show significant advantages. Compared to the hyper-Laplacian based method, IDR performs better on the first example, but gives lower quantitative results on the other examples. Although IDR enhances the visual details, this improvement might not be well evaluated with mean squared difference based

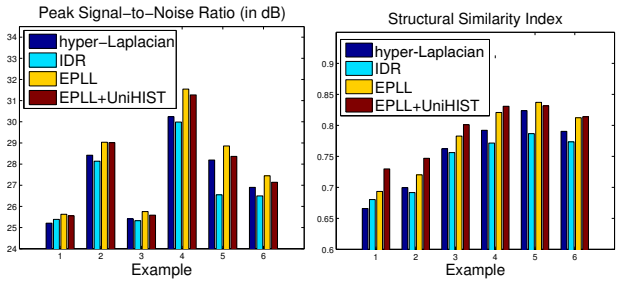


Figure 6: Quantitative results of the four methods in terms of PSNR and SSIM for the six examples. The average PSNR values of hyper-Laplacian, IDR, EPLL and our method (EPLL+UniHIST) are 27.40dB, 26.98dB, 28.04dB and 27.82dB respectively; while the average SSIM values of the four methods are $0.76, 0.74, 0.77, 0.79$ respectively.

image metrics such as PSNR. The same performance difference has been observed in Cho *et al.*'s work [7].

5. Conclusions and Future Work

In this paper, we have presented a unified approach to integrate marginal histogram constraints with image restoration methods. We emphasize the aspect of marginal histograms as ensemble constraints that specify the proportions of pixel or gradient values for the restored image. Specifically, we incorporate marginal histogram constraints into image restoration algorithms with the UniHIST framework, which measures and minimizes the discrepancy between the marginal histograms of the restored image and the reference histograms with the quadratic Wasserstein (W_2) distance. We show the effectiveness of UniHIST through denoising of pattern images and non-blind deconvolution of natural images, where it shows visual and quantitative improvements over state-of-the-art methods.

There are a few things we would like to further improve the current work. First, we would like to study the use of gradient histogram constraints in *blind* deconvolution. Previous studies [12, 22] have shown that the sparsity regularized framework is often not sufficient to recover satisfactory blurring kernel, and we hope that marginal gradient histograms will provide more useful information to this end. Also, enforcing consistency of histograms over large number of band-pass filters has been shown as a successful approach to texture synthesis [17]. As such, we also expect to further improve the deconvolution performance by using other types of band-pass filters (e.g., steerable filters) beyond the simple gradient filters. Last, like marginal histograms, joint histograms (e.g., color histograms) can also be regarded as a type of ensemble constraints. Following recent advances in multi-dimensional distribution matching [3, 26, 27], we are interested in developing a similar framework with joint histograms.

Acknowledgements

This work is supported by National Science Foundation (Grant Nos. IIS-0953373, CCF-1319800), National Institute of Justice (Grant No. 2013-IJ-CX-K010), Natural Science Foundation of China (Grant Nos. 61271430, 61372184, and 61332017) and CASIA-Tencent BestImage Joint Research Project.

References

[1] M. R. Banham and A. K. Katsaggelos. Digital image restoration. *IEEE Signal Process. Mag.*, 14(2):24–41, 1997.

[2] D. P. Bertsekas. *Constrained optimization and Lagrange multiplier methods,*. Athena Scientific, 1996.

[3] N. Bonneel, J. Rabin, G. Peyré, and H. Pfister. Sliced and radon wasserstein barycenters of measures. *JMIV*, 51(1):22–45, 2015.

[4] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *CVPR*, pages 60–65, 2005.

[5] S.-H. Cha and S. N. Srihari. On measuring the distance between histograms. *Pattern Recognition*, 36(6):1355–1370, 2002.

[6] T. S. Cho, N. Joshi, C. L. Zitnick, S. B. Kang, R. Szeliski, and W. T. Freeman. A content-aware image prior. In *CVPR*, pages 169–176, 2010.

[7] T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R. Szeliski, and W. T. Freeman. Image restoration by matching gradient distributions. *TPAMI*, 34(4):683–694, 2012.

[8] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *TIP*, 16(8):2080–2095, 2007.

[9] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR*, pages 886–893, 2005.

[10] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *TIP*, 15(12):3736–3745, 2006.

[11] W. Feller. *An introduction to probability theory and its applications, 3rd edition*, volume 2. Wiley, 1968.

[12] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *TOG*, 25(3):787–794, 2006.

[13] D. Field. What is the goal of sensory coding? *Neural Computation*, 6(4):559–601, 1994.

[14] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. *TIP*, 4(7):932–946, 1995.

[15] R. C. Gonzalez and R. E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., 2006.

[16] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar. Multiresolution histograms and their use for recognition. *TPAMI*, 26(7):831–847, 2004.

[17] D. J. Heeger and J. R. Bergen. Pyramid-based texture analysis/synthesis. In *ACM SIGGRAPH*, pages 229–238, 1995.

[18] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pajak, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian, J. Kautz, and K. Pulli. FlexISP: A flexible camera image processing framework. In *ACM SIGGRAPH Asia*, 2014.

[19] J. Huang and D. Mumford. Statistics of natural images and models. In *CVPR*, pages 1541–1547, 1999.

[20] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-Laplacian priors. In *NIPS*, pages 1033–1041, 2009.

[21] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *TOG*, 26(3), 2007.

[22] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, pages 1964–1971, 2009.

[23] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.

[24] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):123–231, 2013.

[25] N. Pierazzo, M. Lebrun, M. E. Rais, J. M. Morel, and G. Facciolo. Non-local dual denoising. In *ICIP*, 2014.

[26] F. Pitié, A. C. Kokaram, and R. Dahyot. N-dimensional probability density function transfer and its application to color transfer. In *ICCV*, pages 1434–1439, 2005.

[27] J. Rabin and G. Peyré. Wasserstein regularization of imaging problems. In *ICIP*, pages 1541–1544, 2011.

[28] S. Roth and M. Black. Field of experts. *IJCV*, 82(2):205–229, 2009.

[29] Y. Rubner, C. Tomasi, and L. J. Guibas. The earth mover's distance as a metric for image retrieval. *IJCV*, 40(2):99–121, 2000.

[30] U. Schmidt, Q. Gao, and S. Roth. A generative perspective on MRFs in low-level vision. In *CVPR*, pages 1751–1758, 2010.

[31] E. Simoncelli and E. Adelson. Noise removal via Bayesian wavelet coring. In *ICIP*, pages 379–382, 1996.

[32] C. Villani. *Optimal Transport: Old and New*, volume 338. Grundlehren der mathematischen Wissenschaften, 2009.

[33] M. J. Wainwright and E. Simoncelli. Scale mixtures of Gaussians and the statistics of natural images. In *NIPS*, volume 12, pages 855–861, 2000.

[34] Z. Wang, G. Wu, H. R. Sheikh, and E. P. Simoncelli. Quality-aware images. *TIP*, 15(6):1680–1689, 2006.

[35] O. J. Woodford, C. Rother, and V. Kolmogorov. A global perspective on MAP inference for low-level vision. In *ICCV*, pages 2319–2326, 2009.

[36] S.-C. Zhu, Y. Wu, and D. Mumford. Filters, random fields and maximum entropy (FRAME): towards a unified theory for texture modeling. *IJCV*, 27(2):107–126, 1998.

[37] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *ICCV*, pages 479–486, 2011.

[38] W. Zuo, D. Meng, L. Zhang, X. Feng, and D. Zhang. A generalized iterated shrinkage algorithm for non-convex sparse coding. In *ICCV*, pages 217–224, 2013.

[39] W. Zuo, L. Zhang, C. Song, D. Zhang, and H. Gao. Gradient histogram estimation and preservation for texture enhanced image denoising. *TIP*, 23(6):2459–2472, 2014.