Removing Motion Blur using Natural Image Statistics

Johannes Herwig, Timm Linder and Josef Pauli
Intelligent Systems Group, University of Duisburg-Essen, Bismarckstr. 90, 47057 Duisburg, Germany

Keywords: Deconvolution, Bayesian Inference, Sparse Gradient Priors, Two-color Model, Parameter Learning.

Abstract: We tackle deconvolution of motion blur in hand-held consumer photography with a Bayesian framework combining sparse gradient and color priors for regularization. We develop a closed-form optimization utilizing iterated re-weighted least squares (IRLS) with a Gaussian approximation of the regularization priors. The model parameters of the priors can be learned from a set of natural images which resemble common image statistics. We throughly evaluate and discuss the effect of different regularization factors and make suggestions for reasonable values. Both gradient and color priors are current state-of-the-art. In natural images the magnitude of gradients resembles a kurtotic hyper-Laplacian distribution, and the two-color model exploits the observation that locally any color is a linear approximation between some primary and secondary colors. Our contribution is integrating both priors into a single optimization framework and providing a more detailed derivation of their optimization functions. Our re-implementation reveals different model parameters than previously published, and the effectiveness of the color priors alone are explicitly examined. Finally, we propose a context-adaptive parameterization of the regularization factors in order to avoid over-smoothing the deconvolution result within highly textured areas.

1 INTRODUCTION

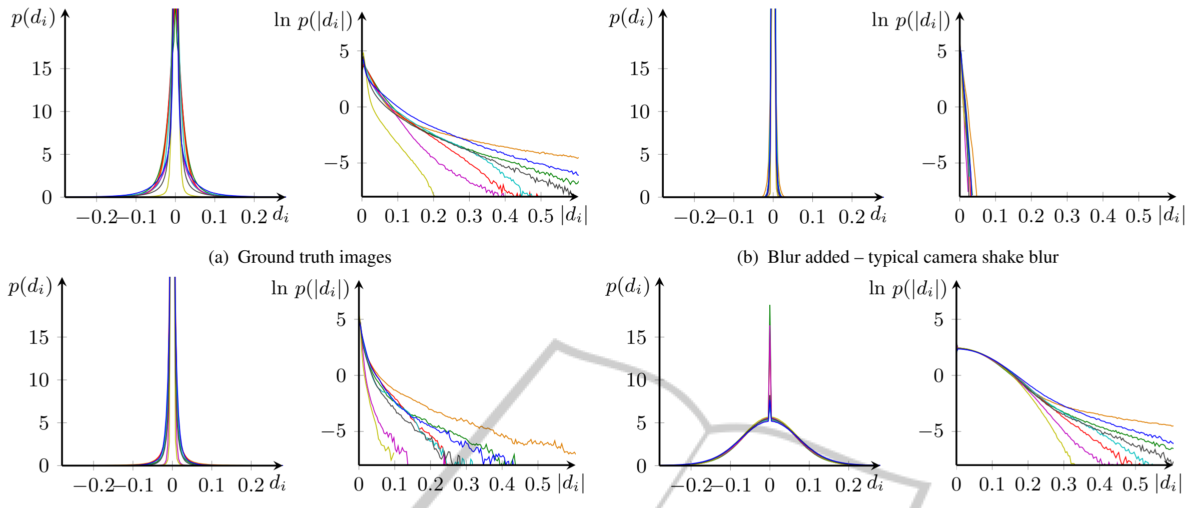
Removing motion blur due to camera shake is a special branch of the ill-posed deconvolution problem. Its specific challenges are the relatively large blur kernels and image noise which usually is stronger here, because camera shake is often caused by longer exposure times during low-light photography where sensor noise is inherently amplified due to higher analog gain and shot noise. Another characteristic property is that the blur kernels are not isotropic as with out-of-focus blur, but instead these point spread functions (PSFs) model the path of motion that a handheld camera undertakes during the exposure time of the photograph, and therefore the PSFs have a ridge-like and sparse appearance (Liu et al., 2008).

We here tackle the problem of non-blind deconvolution where the motion blur kernel (or PSF) is exactly known a priori. In the real world, the gyroscope of a mobile phone camera might give a good estimate of the blur kernel. It is however not straightforward to synchronize the gyroscope with start and end time of the exposure. If motion information is not available at all, then we talk about blind deconvolution where the blur kernel needs to be estimated solely with the help of the blurred image at hand (Shi et al., 2013; Dong et al., 2012a). Since this is rather difficult there are also some image fusion approaches, known as semi-blind deconvolution (Yuan et al., 2007; Ito et al., 2013; Wang et al., 2012). Thereby, multiple differently blurred or otherwise multimodal images are taken from the same scene with the same or different sensor which helps further constraining the blur kernel (Yuan et al., 2007; Ito et al., 2013; Wang et al., 2012). Our approach assumes a globally constant blur kernel (Schmidt et al., 2013), but in general image blur is space-varying (Sorel and Sroubek, 2012; Ji and Wang, 2012; Whyte et al., 2012; Gupta et al., 2010) because objects at different distances in the scene are blurred differently. Also, there are natural design constraints on the camera optics, so that an image is usually sharper in the center compared to its border. Additionally, there could be moving objects in the scene which overlay the movement of a hand-held camera (Cho et al., 2012). However, usually only static scenes are considered when there is only one image available.

1.1 Regularization

Most non-blind deconvolution approaches apply a regularization term to the gradients of the image, by

Herwig J., Linder T. and Pauli J..
Removing Motion Blur using Natural Image Statistics.
DOI: 10.5220/0004830201250136
In *Proceedings of the 3rd International Conference on Pattern Recognition Applications and Methods* (ICPRAM-2014), pages 125-136
ISBN: 978-989-758-018-5



(a) Ground truth images
(b) Blur added – typical camera shake blur
(c) Blur added – linear motion in y direction
(d) Noise added (no blur!) – $\sigma = 5\%$

Figure 1: Histograms of x-derivatives and x-derivative magnitudes. Colors correspond to Fig. 4.

penalizing steep gradients that could be indicative of noise. Regularization based upon the ℓ_2 norm (Gaussian prior) and the ℓ_1 norm (Laplacian prior, total variation (Chan and Shen, 2005)) tend to oversmooth the deconvolution results. The so-called sparse priors (Levin and Weiss, 2007; Levin et al., 2007b; Li et al., 2013) more adequately capture the observed hyper-Laplacian gradient distributions (Srivastava et al., 2003; Huang, 2000). Here, the color model-based regularization (Joshi et al., 2009) motivated by (Cecchi et al., 2010) imposes a two-color model upon locally smooth regions. Thereby, we concurrently make use of global and local sparseness (Dong et al., 2012b) alike by using gradient and color priors, respectively.

1.2 Sparse Gradient Prior

Most 'real' images resemble a common gradient distribution (Levin and Weiss, 2007; Levin et al., 2007b; Simoncelli, 1997). Under for example the ℓ_1-norm, the gradient magnitude for a pixel i is calculated by

$$\| (\nabla I)_i \|_1 = \sum_{k=1}^{n} |d_{k,i}|, \tag{1}$$

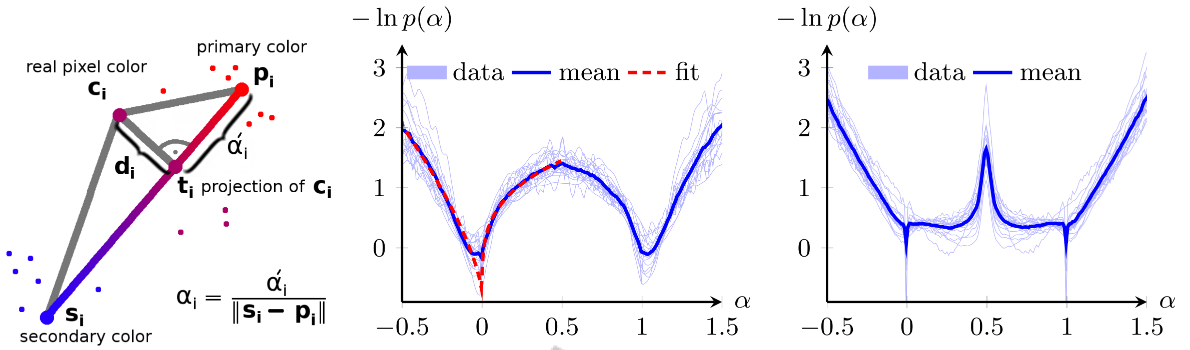
where $d_{k,i}$ represents the k-th partial derivative, $\vec{d}_k$, evaluated at pixel i of image I. Such a directional derivative $\vec{d}_k := \mathrm{vec}\,(I * G_k)$ can be determined by convolving the image I with derivative filter kernels G_k, like $\begin{pmatrix}1 & -1\end{pmatrix}$ and $\begin{pmatrix}1 & -1\end{pmatrix}^\top$ and the second-order derivatives $\frac{\partial I}{\partial x^2}$, $\frac{\partial I}{\partial y^2}$, $\frac{\partial I}{\partial xy}$.

In Fig. 1, we examine the gradient distributions of the images from Fig. 4 and compare them with unwanted deconvolution results. These histogram plots show that only the ground truth photographs exhibit the kurtotic hyper-Laplacian shape, but blurry and noisy images show totally different statistics. However, if the blur is linear and orthogonal to the direction of the derivative, then edges stay mostly intact – but still the kurtotic tail is lowered (compare Fig. 1(a) and 1(c)). Similarly to our analysis (Lin et al., 2011) shows gradient distributions of examplarily patches of motion blurred vs. sharp textures.

Instead of using gradients as a sparse prior, one could use any kind of filtering result that provides a sparse representation of the image. We also tried the learned filters approach within the Fields-of-Experts (FoE) framework. Thereby we modified the MATLAB code of (Weiss and Freeman, 2007) so that we obtained a kurtotic curve model. Then we learned two different sets of 5 and 25 filters of 15×15 pixels. As opposed to (Schmidt et al., 2011) we did not find an increase in performance, but our results were comparable to the sparse gradient prior.

1.3 Two-color Model

As in (Joshi et al., 2009), for each pixel $\vec{c}_i$ of a latent image estimate I, we define a pixel neighborhood – e. g. using a square 5×5 window – and determine the primary and secondary colors within this neighborhood. Thereby, an initial two-color model is obtained by k-means clustering (with $k = 2$). While k-means provides a good heuristic for finding an initial two-color model, the drawback is that one color sample can always only be assigned to exactly one cluster, and therefore noise is not appropriately han-



(a) Two-color model by Joshi (b) α distribution after EM clustering (c) α distribution after k-means

Figure 2: The two-color model, and the accompanying alpha distributions learned from real images.

dled. A fuzzy expectation-maximization (EM) algorithm based upon the method described in (Joshi et al., 2009) therefore refines the color clusters. Finally, the primary color $\vec{p}_i$ is assigned to the cluster whose center lies closest to the color of the center pixel $\vec{c}_i$ defining the neighborhood. The secondary color $\vec{s}_i$ is assigned to the other cluster. The two-color model as depicted in Fig. 2(a) now works upon the assumption that the color $\vec{c}_i$ can be represented by a linear interpolation between its associated primary and secondary colors $\vec{p}_i \in \mathbb{R}^3$ and $\vec{s}_i \in \mathbb{R}^3$, where $\alpha_i \in \mathbb{R}$ is the interpolation or mixing parameter (Joshi et al., 2009, see eq. 7):

$$\vec{c}_i \approxeq \vec{t}_i := \alpha_i \vec{s}_i + (1-\alpha_i)\vec{p}_i. \tag{2}$$

In contrast to (Joshi et al., 2009), $\vec{p}_i$ and $\vec{s}_i$ have been swapped with regard to α_i, such that α_i is minimal at $\vec{p}_i$ instead of $\vec{s}_i$, because this will simplify our IRLS optimization later on.

2 THE COLOR PRIOR

The so-called alpha prior introduced by (Joshi et al., 2009) penalizes α_i values which would put the estimated color $\vec{c}_i$ far away from either the primary or secondary color. The penalty is based upon the observed α_i distribution in natural images. The value of α_i for a pixel $\vec{c}_i$, given $\vec{p}_i$ and $\vec{s}_i$, is calculated using (Joshi et al., 2009, eq. 10):

$$\alpha_i = \frac{\alpha'_i}{\|\vec{s}_i - \vec{p}_i\|} = \underbrace{\left(\frac{(\vec{s}_i - \vec{p}_i)}{(\vec{s}_i - \vec{p}_i)^\top(\vec{s}_i - \vec{p}_i)}\right)^\top}_{=:\vec{\ell}_i \in \mathbb{R}^3} (\vec{c}_i - \vec{p}_i). \tag{3}$$

Typical distributions of α values determined from natural images using either the refined EM or bare k-means color clustering are shown in Fig. 2(b) and 2(c). As in Fig. 2(b), these negative log-likelihoods can be fit with a piecewise hyper-Laplacian prior term of the form

$$b|\alpha_i|^a = b\,|\,\vec{\ell}_i^{\,\top}\vec{c}_i - \vec{\ell}_i^{\,\top}\vec{p}_i\,|^a\,. \tag{4}$$

We calculated the two-color model from about 400 images of the Berkeley image segmentation database. For a custom set of fit parameters, the constructed color models were exported into Matlab and approximate probability densities have been estimated for each image using a Parzen window method. Then, a non-linear least-squares fit was performed using Matlab's `fminsearch()` method for the piecewise hyper-Laplacian function described above. In (Joshi et al., 2009) a 2 pieces fit is proposed, but we found that it is not sufficiently accurate for approaching the statistics of the relatively noise-free ground truth images. The resulting parameters for our 3 pieces fit are:

$$\begin{aligned} a &= 0.7153,\ b = 0.8066 && \text{for } \alpha_i < -0.5;\\ a &= 0.6448,\ b = 4.5318 && \text{for } -0.5 \le \alpha_i < 0;\\ a &= 0.2298,\ b = 2.7372 && \text{for } 0 \le \alpha_i. \end{aligned}$$

2.1 Optimization Techniques

For weighted least squares (WLS) (Faraway, 2002, p. 62), a weighting matrix $W \in \mathbb{R}^{s\times s}$ is introduced. The WLS objective function therefore is

$$\begin{aligned} \sum_k^s \sum_l^s r_k W_{k,l} r_l &= \|\vec{y} - T\vec{x}\|_W^2 \\ &= (\vec{y} - T\vec{x})^\top W (\vec{y} - T\vec{x}), \end{aligned} \tag{5}$$

where $\|\cdot\|_W$ is the Mahalanobis distance when $W = \Sigma^{-1}$. In order to minimize this function, we need to determine the gradient and set it equal to zero. The WLS derivative

$$\frac{\partial}{\partial \vec{x}} \|\vec{y} - T\vec{x}\|_W^2 = -2T^\top W\vec{y} + 2T^\top WT\vec{x} \qquad (6)$$

yields the system of the so-called normal equations of WLS (Gentle, 2007, p. 338)

$$\underbrace{(T^\top WT)}_{A}\,\vec{x} - \underbrace{T^\top W\vec{y}}_{\vec{\delta}} = \vec{0} \qquad (7)$$

which represents a linear equation system of the form $A\vec{x} - \vec{\delta} = \vec{0}$. Here, $A = T^\top WT$ is too large to be inverted in-place, and hence we use the CG (Conjugate Gradient) method.

The M-estimator (Meer, 2004, p. 47) applies a robust penalty or loss function ρ to the error residuals r_i. For $\rho(r_i) := |r_i|^p$, $p \neq 2$, the optimization becomes non-linear. However, the iteratively re-weighted least squares (IRLS) method (Scales et al., 1988; Scales and Gersztenkorn, 1988) approximates the solution by turning the problem into a series of WLS subproblems. A faster version of this algorithm for the problem at hand is discussed in (Krishnan and Fergus, 2009). In each IRLS iteration, a new set of weights is learned from the previous solution. For the first iteration, all weights can be initialized with a constant value. The weights w_i of the diagonal WLS weighting matrix W are ([1]: (Meer, 2004, p. 48); [2]: (Scales et al., 1988, p. 332)):

$$w_i^{(\tau+1)} = w(r_i^{(\tau)}) \overset{[1]}{=} \frac{1}{r_i^{(\tau)}} \frac{\mathrm{d}\rho(r_i^{(\tau)})}{\mathrm{d}r_i^{(\tau)}} \overset{[2]}{=} p\,|r_i^{(\tau)}|^{p-2}. \qquad (8)$$

2.2 Minimizing the Alpha Prior

As already shown in Fig. 2(b), the α_i distribution is bimodal since both $\alpha_i = 0$ and $\alpha_i = 1$ are minima and the distribution is symmetric at $\alpha_i = 0.5$. However, since we want to bias the observed color $\vec{c}_i$ to the primary color $\vec{p}_i$ at $\alpha_i = 0$, only the unimodal prior (represented by the red, dashed line) is used. The weights of the alpha prior in IRLS step $(\tau + 1)$ that follow by applying eqn. 8 to eqn. 4 are:

$$w_i^{(\tau+1)} = a \cdot b \cdot |\alpha_i^{(\tau)}|^{a-2}.$$

Note that the constant coefficient a is missing in this term given by (Joshi et al., 2009, eqn. 13). With these weights, the WLS can be performed with the RGB components of the latent image $I \in \mathbb{R}^{m\times n}$ as the parameter vector $\vec{x} := \left(\vec{c}_1{}^\top, \ldots, \vec{c}_s{}^\top\right)^\top = (I_{R,1}, I_{G,1}, I_{B,1}, \ldots, I_{R,s}, I_{G,s}, I_{B,s})^\top \in \mathbb{R}^{3s}$ of eqn. 5 with $s = mn$ the total amount of image pixels. Following the definition of α_i (eqn. 3) and splitting α_i into a variable and a constant part, the WLS coefficient matrix T is block diagonal:

$$T := \begin{pmatrix} -\vec{\ell}_1{}^\top & & 0 \\ & \ddots & \\ 0 & & -\vec{\ell}_s{}^\top \end{pmatrix} \in \mathbb{R}^{s\times 3s}.$$

The constant part $\vec{y}$ of the WLS objective function is then a vector

$$\vec{y} := \left(-\vec{\ell}_1{}^\top \vec{p}_1, \ldots, -\vec{\ell}_s{}^\top \vec{p}_s\right)^\top \in \mathbb{R}^s.$$

Due to the block-diagonal form of T, the WLS normal equations can be evaluated for the alpha prior individually per pixel. Inserting the above definitions and expanding eqn. 6 leads to the gradient in block matrix form $\frac{\partial}{\partial \vec{x}} \lambda_\alpha \|\vec{y} - T\vec{x}\|_W^2 =$

$$\underbrace{2\lambda_\alpha \begin{pmatrix} R_1\vec{c}_1 \\ \vdots \\ R_s\vec{c}_s \end{pmatrix}}_{A\vec{x}} - \underbrace{2\lambda_\alpha \begin{pmatrix} R_1\vec{p}_1 \\ \vdots \\ R_s\vec{p}_s \end{pmatrix}}_{\vec{\delta}} \in \mathbb{R}^{3s} \qquad (9)$$

with $R_i := w_i^{(\tau)} \cdot \vec{\ell}_i\vec{\ell}_i{}^\top \in \mathbb{R}^{3\times 3}$ where the 3×3 matrix R_i is called the re-weighting term by (Joshi et al., 2009, eqn. 13), and contains the weights $w_i^{(\tau)}$ learned from the previous IRLS iteration's deconvolution result. The outer product $\vec{\ell}_i\vec{\ell}_i{}^\top$ appears because of the matrix products $T^\top \cdot \ldots \cdot T$ and $T^\top \cdot \ldots \cdot \vec{y}$ in the term $2T^\top WT\vec{x} - 2T^\top W\vec{y}$. λ_α is a regularization factor of the alpha prior.

2.3 Penalty on the Distance d

Besides the prior on α_i values, another penalty term is introduced by (Joshi et al., 2009) that minimizes the squared distance d_i^2 (Fig. 2(a)). In contrast to the α_i prior, this penalty term is not based upon any observed probability distribution in real images. Instead, the d_i is simply minimized (Joshi et al., 2009, eqn. 8). Given $\vec{p}_i$ and $\vec{s}_i$, then $o_d(\vec{c}_i) := \lambda_d \cdot d_i^2 = \lambda_d\|\vec{c}_i - \vec{t}_i(\vec{c}_i)\|^2 = \lambda_d\|\vec{c}_i - [\alpha_i(\vec{c}_i) \cdot (\vec{s}_i - \vec{p}_i) + \vec{p}_i]\,\|^2$ whereby the regularization factor λ_d specifies the strength of this penalty term. In the above objective function, $\vec{c}_i \in \mathbb{R}^3$ represents the color of a single pixel i of the latent image I, and is thus a variable. α_i and hence $\vec{t}_i$ are functions of $\vec{c}_i$ (see eqn. 3). This is different from the alpha prior, where the calculated α_i was fixed during the CG optimization because the weights for the hyper-Laplacian alpha prior only get updated *between* IRLS iterations. $\vec{p}_i$ and $\vec{s}_i$, on the other hand, can be regarded as constants until a new color model is built.

The d penalty term is optimized by least-squares and its gradient is

$$\begin{aligned}\frac{\partial}{\partial \vec{c}_i}\, o_d(\vec{c}_i) &= \lambda_d \left[\vec{c}_i - \vec{t}_i(\vec{c}_i)\right]^\top \left[\vec{c}_i - \vec{t}_i(\vec{c}_i)\right] \\ &= 2\lambda_d \left[\mathrm{id}_3 - \frac{\partial}{\partial \vec{c}_i}\, \vec{t}_i(\vec{c}_i)\right]^\top \left[\vec{c}_i - \vec{t}_i(\vec{c}_i)\right]\end{aligned}$$

with id_3 being a 3×3 identity matrix. Further differentiation leads to

$$\begin{aligned}\frac{\partial}{\partial \vec{c}_i}\, \alpha_i(\vec{c}_i) &= \frac{\partial}{\partial \vec{c}_i}\, \vec{\ell}_i^{\,\top}(\vec{c}_i - \vec{p}_i) = \vec{\ell}_i, \\ \frac{\partial}{\partial \vec{c}_i}\, \vec{t}_i(\vec{c}_i) &= \frac{\partial}{\partial \vec{c}_i}\, \alpha_i(\vec{c}_i)(\vec{s}_i - \vec{p}_i) = \vec{\ell}_i(\vec{s}_i - \vec{p}_i)^\top,\end{aligned}$$

such that

$$\frac{\partial}{\partial \vec{c}_i}\, o_d(\vec{c}_i) = 2\lambda_d \left[\mathrm{id}_3 + \vec{\ell}_i(\vec{p}_i - \vec{s}_i)^\top\right] \left[\vec{c}_i - \vec{t}_i(\vec{c}_i)\right].$$

As $\alpha_i(\vec{c}_i)$ contains both a part that is dependent on $\vec{c}_i$ and one that is constant (namely $\vec{l}_i$), the gradient is split up for the CG method. The RGB blocks for the pixels $i = 1, \ldots, s$ of the vectors $(A\vec{x})_{3i-2,\ldots,3i}$ and $(\vec{\delta})_{3i-2,\ldots,3i} \in \mathbb{R}^{3s}$ are:

$$\begin{aligned}&2\lambda_d \left[\mathrm{id}_3 + \vec{\ell}_i(\vec{p}_i - \vec{s}_i)^\top\right] \left[\vec{c}_i + \vec{\ell}_i^{\,\top} \vec{c}_i(\vec{p}_i - \vec{s}_i)\right] \in \mathbb{R}^3, \\ &2\lambda_d \left[\mathrm{id}_3 + \vec{\ell}_i(\vec{p}_i - \vec{s}_i)^\top\right] \left[\vec{p}_i + \vec{\ell}_i^{\,\top} \vec{p}_i(\vec{p}_i - \vec{s}_i)\right] \in \mathbb{R}^3.\end{aligned} \tag{10}$$

3 SPARSE & COLOR PRIORS

For a closed-form expression of the linear system $A\vec{x} - \vec{\delta} = \vec{0}$, the gradients of the sparse prior, the data likelihood, the color prior α (eqn. 9) and the penalty term on d (eqn. 10) are summed up. Since the data likelihood and the sparse prior work on intensity images, the individual color channels $\in \mathbb{R}^s$ are extracted from the RGB vector $\vec{x} \in \mathbb{R}^{3s}$ and the blurry image $\vec{b} \in \mathbb{R}^{3s}$, and then combined again after the gradients of the penalty terms are applied as shown in Fig. 3.

Thereby, the binary operator $\mathrm{ext} : \{\mathrm{R},\mathrm{G},\mathrm{B}\} \times \mathbb{R}^{3s} \to \mathbb{R}^s$ extracts the color channel specified by the first argument from an image $\vec{v} \in \mathbb{R}^{3s}$ into a vector $\vec{u} \in \mathbb{R}^s$. The unary operator $\mathrm{join}\cdot$ merges a set $\{(\mathrm{R},\vec{u}_\mathrm{R}), (\mathrm{G},\vec{u}_\mathrm{G}), (\mathrm{B},\vec{u}_\mathrm{B})\}$ of 3 separate channels back into an RGB image $\vec{v}$.

The data likelihood term and the sparse prior are applied to the three color channels $\in \mathbb{R}^s$ of the current estimate $\vec{x} \in \mathbb{R}^{3s}$ and the blurry input image $\vec{b} \in \mathbb{R}^{3s}$ individually, where s is the total number of image pixels. The weights in the matrices W_k, R_i, as well as the primary and secondary colors of the two-color model, are recalculated after each IRLS iteration. In the first iteration, λ_α and λ_d are set to 0 and hence only the sparse prior is active then.

3.1 Regularization Parameters

First, we want to find a suitable range of parameter values with which reasonable deconvolution results can be achieved. Therefore, the blurred, noisy versions of the ground truth images from Fig. 4 have been deconvolved, using their accompanying PSFs as shown. For the sparse prior a hyper-Laplacian exponent of $\gamma = 0.5$ was used together with the default first- and second-order derivative filters (5 filters in total). The exponent $\gamma = 0.5$ was chosen because of $\gamma \in [0.5, 0.8]$ for the gradient distribution of most natural images (Huang, 2000, pp. 19–24). The influence $\lambda_{\nabla,k}$ of the second-order derivatives was set to a constant $\frac{1}{4}$, as done by (Levin et al., 2007a).

We used PSNR (peak signal-to-noise ratio) and MSSIM (multi-scale structural similarity index) (Wang et al., 2003) for evaluating the goodness of the deconvolution results. Thereby, MSSIM takes into account interdependencies of local pixel neighborhoods which otherwise get averaged out by the more traditional but established PSNR method. High-quality digital images have PSNRs between 30db and 50db, whereas 20db to 30db are still regarded as acceptable. With PSNR we have a contex-independent measure for sole signal quality, and MSSIM gives us the similarity between a ground truth and estimated texture without severely punishing correlated errors. There are metrics available that quantize the degree of image blur directly, but since these are more or less based on the same kurtotic model of the distribution of gradients (Yun-Fang, 2010; Liu et al., 2008) where our optimization model for natural images is built upon, we did not consider these further. Frequency-based methods to blur detection (Marichal et al., 1999) can only quantify the global blur of an image but do not cope with space-varying blur which is introduced by our non-linear and context-dependent regularization approach, and hence were not considered.

The diagrams in Fig. 5 show the mean MSSIM and PSNR (thick line) for various noise levels, averaged over the entire set of images and as a function of the regularization parameter λ_∇. The thin lines represent the maximum and minimum MSSIM and PSNR values of all 8 images, and the error bars denote the sample standard deviations. On average and also subjectively, best results were obtained for λ_∇ between 0.5 and 2.5 depending on the noise level.

The paper by (Joshi et al., 2009) suggests a reduced regularization factor, λ_∇, in the initialization phase of the sparse prior in order to preserve details. Then, λ_∇ can be increased, once the penalty terms based upon the two-color model become active. However, their proposed values are inconsis-

$$A\vec{x} = \mathrm{join}\Bigg\{ \Big(j,\, \big[\underbrace{C_K^\top \Sigma^{-1} C_K}_{\text{Data likelihood}} + \underbrace{\lambda_\nabla \sum_k \lambda_{\nabla,k} C_{G_k}^\top W_k C_{G_k}}_{\text{Sparse prior}}\big] \cdot \mathrm{ext}(j,\vec{x}) \Big) \Bigg|\, j \in \{\mathrm{R},\mathrm{G},\mathrm{B}\} \Bigg\}$$

$$+ \underbrace{2\lambda_\alpha \begin{pmatrix} R_1\vec{c}_1 \\ \vdots \\ R_s\vec{c}_s \end{pmatrix}}_{\text{Alpha prior}} + \underbrace{2\lambda_d \begin{pmatrix} \big[\mathrm{id}_3 + \vec{\ell}_1(\vec{p}_1 - \vec{s}_1)^\top\big]\big[\vec{c}_1 + \vec{\ell}_1^{\,\top}\vec{c}_1(\vec{p}_1 - \vec{s}_1)\big] \\ \vdots \\ \big[\mathrm{id}_3 + \vec{\ell}_s(\vec{p}_s - \vec{s}_s)^\top\big]\big[\vec{c}_s + \vec{\ell}_s^{\,\top}\vec{c}_s(\vec{p}_s - \vec{s}_s)\big] \end{pmatrix}}_{d \text{ penalty term}} \in \mathbb{R}^{3s}$$

$$\vec{\delta} = \mathrm{join}\Bigg\{ \Big(j,\, \underbrace{C_K^\top \Sigma^{-1} \cdot \mathrm{ext}(j,\vec{b})}_{\text{Data likelihood}} \Big) \Bigg|\, j \in \{\mathrm{R},\mathrm{G},\mathrm{B}\} \Bigg\}$$

$$+ \underbrace{2\lambda_\alpha \begin{pmatrix} R_1\vec{p}_1 \\ \vdots \\ R_s\vec{p}_s \end{pmatrix}}_{\text{Alpha prior}} + \underbrace{2\lambda_d \begin{pmatrix} \big[\mathrm{id}_3 + \vec{\ell}_1(\vec{p}_1 - \vec{s}_1)^\top\big]\big[\vec{p}_1 + \vec{\ell}_1^{\,\top}\vec{p}_1(\vec{p}_1 - \vec{s}_1)\big] \\ \vdots \\ \big[\mathrm{id}_3 + \vec{\ell}_s(\vec{p}_s - \vec{s}_s)^\top\big]\big[\vec{p}_s + \vec{\ell}_s^{\,\top}\vec{p}_s(\vec{p}_s - \vec{s}_s)\big] \end{pmatrix}}_{d \text{ penalty term}} \in \mathbb{R}^{3s}$$

Figure 3: Combining the color and sparse priors within the IRLS optimization framework.

tent: $\lambda_\nabla = 0.25$ followed by $\lambda_\nabla = 0.5$ is mentioned at one occasion, $\lambda_\nabla = 1$ at another. This approach can be problematic if the initial λ_∇ is chosen too low (e.g. $\lambda_\nabla = 0.8$). Details are preserved, but also artifacts within near-to homogeneous regions are introduced as can be deduced from Fig. 5. Here, the thin curves denoting the absolute minima of MSSIM values are significantly worse than their overall mean substracted by their standard deviation (whereas this gap is not observable for the maximum value curves; this observation is only true up until $\lambda_\nabla = 1.0$). On the other hand, a high regularization factor such as $\lambda_\nabla = 3.0$ over-smoothes the image. We therefore suggest a nearly constant regularization factor for the gradient prior. E.g., for a noise standard deviation of $\sigma = 2.5\%$, λ_∇ might initially be set to 1.5 and then be increased to 2. Note that $\lambda_\nabla = 2$ is slightly above the optimal value discovered for this noise level in Fig. 5; experience shows, though, that rather smooth images require a slightly higher λ_∇.

4 EVALUATION

First, we discuss the effects of the color prior and then we show some qualitative results.

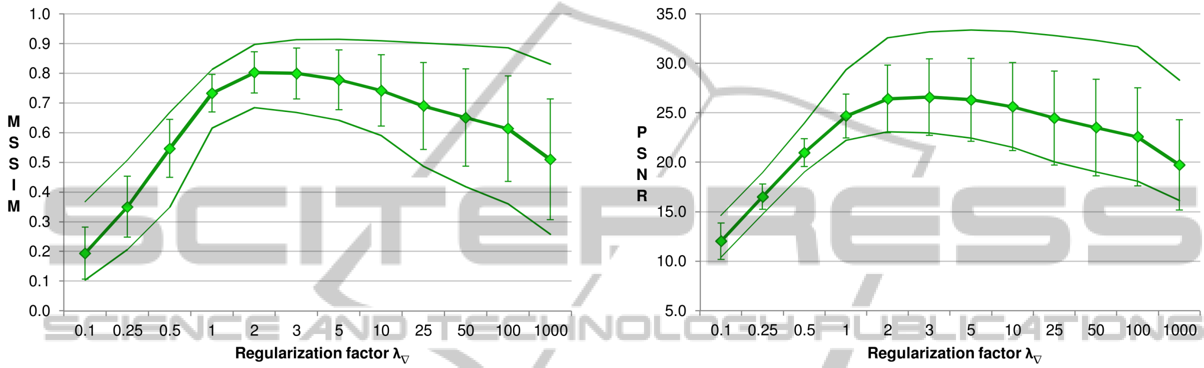
4.1 Understanding the Color Prior

In order to better understand the practical implications of the two-color model, we show some segmentation into primary and secondary colors in Fig. 6. The original image is decomposed by EM clustering of a 5×5 pixel neighborhood into a layer of primary colors (Fig. 6(b)) and secondary colors (Fig. 6(c)). The two-color model applies only at pixels where the color difference between both layers is large enough. Fig. 6(a) shows in black where the two-color model does apply, and in white where the priors derived from this model cannot be utilized. In these cases, a different kind of prior, e. g. a gradient prior, must be used. (Joshi et al., 2009) suggest to generally combine both a sparse gradient prior and the two-color model (where applicable) with a reduced regularization factor for the former.

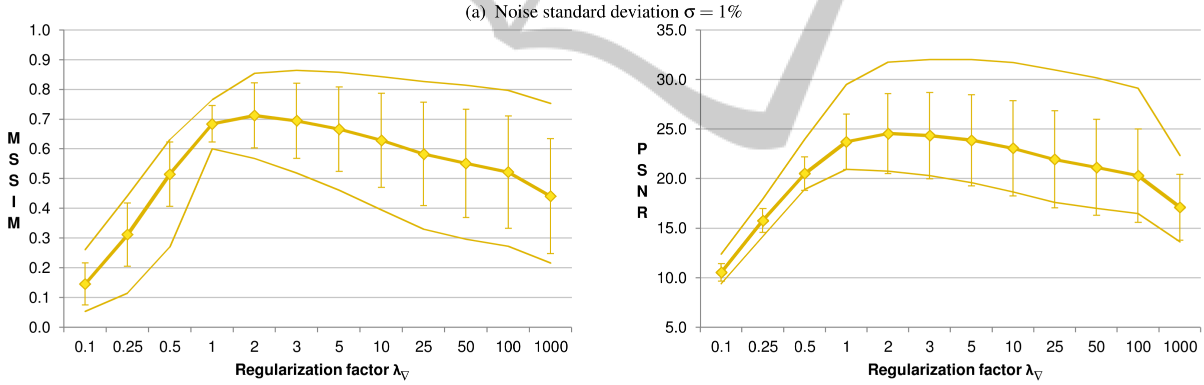
The histograms of the negative log-likelihoods in Fig. 7 illustrate the effect of the alpha prior penalty term on the distribution of alpha values. Both example images have been initially deconvolved with a sparse prior ($\lambda_\nabla = 2, \gamma = 0.5$) before enabling the two-color model ($\lambda_\alpha = 5$ for the first image, and $\lambda_\alpha = 100$ for the second which amplifies the effect for illustration purposes; $\lambda_d = 0$). The red line shows the distribution after the initial sparse prior deconvolution. The blue and green lines show the distribution after 1, respective 2, further IRLS iterations with the now

(a) (b) (c) (d) (e) (f) (g) (h)

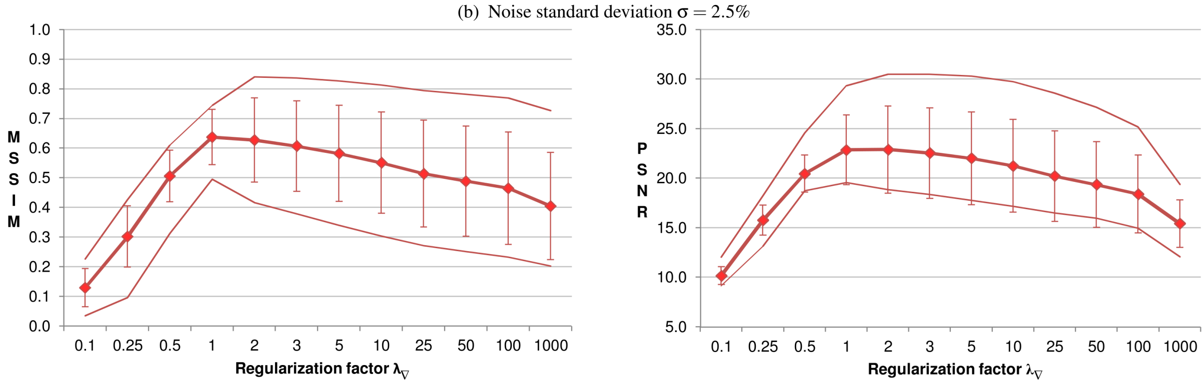
Figure 4: Ground truth pictures and their accompanying blur kernels. Image sizes are approx. 800×600 pixels and blur kernels are 27×27, 49×29, 31×31, 39×39, 27×27, 29×29, 51×45, 95×95, respectively.



(a) Noise standard deviation $\sigma = 1\%$

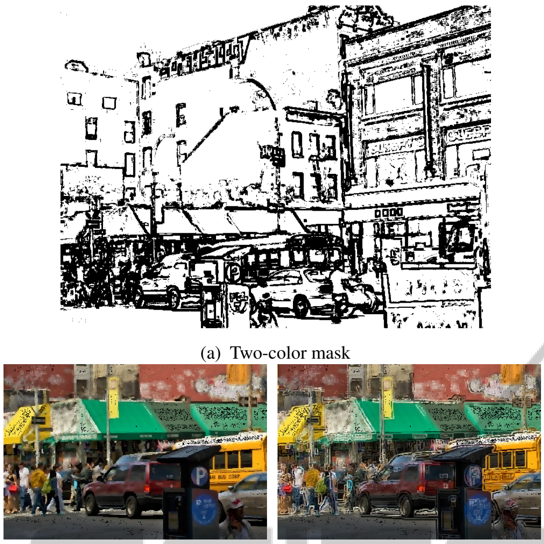


(b) Noise standard deviation $\sigma = 2.5\%$



(c) Noise standard deviation $\sigma = 5\%$

Figure 5: Average MSSIM and PSNR values for the evaluation image set with its paired blur kernels of Fig. 4 at three different noise levels σ as a function of the regularization parameter λ_∇.



(a) Two-color mask

(b) Primary colors (c) Secondary colors

Figure 6: Exemplary two-color model for the ground truth image of Fig. 4(c). The two-color mask shows where the color prior can be applied.

active alpha prior, while retaining sparse prior regularization. The grey line, in comparison, illustrates how the final distribution would have looked like if the alpha prior was never activated. Note how the shown distributions have a shape similar to the ones from Fig. 2(c), which is because the k-means only algorithm without EM refinement was used here to construct the color model. In comparison with the prior on distances d_i, the alpha prior is more effective. Both penalty terms require surprisingly large regularization factors, especially compared to the parameters mentioned by (Joshi et al., 2009).

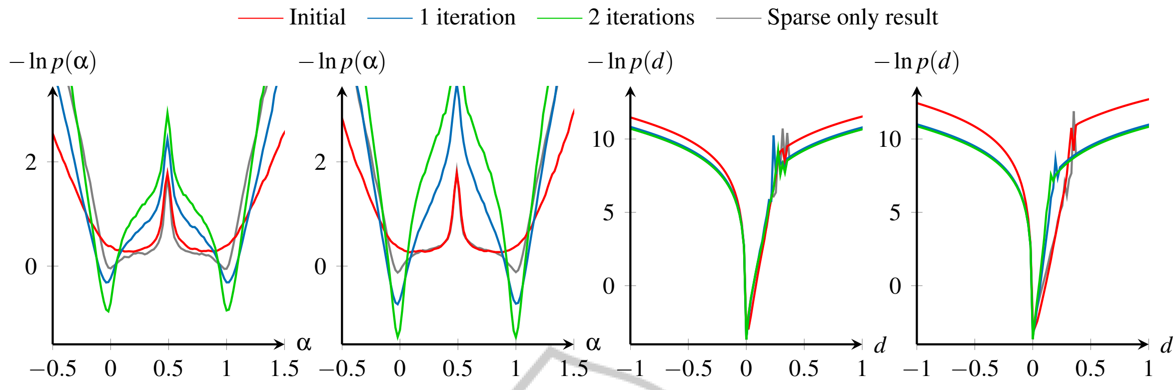
In Fig. 10, we show the effects of different amounts of regularization by λ_α. The color noise in the right Fig. 10(c) might indicate too few iterations with the sparse prior before the first color model was built by EM clustering. Hence, this result with stronger regularization is not necessarily worse: Some edges, e. g. at the perimeter of the blue parking meter sign, or the white graffiti at the building wall in the background, appear more clearly defined with larger λ_α. If λ_α becomes too large, the edges become jagged and the image more and more resembles the primary color layer. Strongly structured images with lots of edges seem to profit more from the prior on alpha values. Unlike the sparse prior, there is no recommendation for the choice of λ_α. Hence, some experimentation is required for each individual image.

4.2 Qualitative Evaluation

We show exemplarily qualitative results in Fig. 8 and Fig. 9, whereby the second example is an image of much less texture than the first image, and also it has much more noise added. Therefore the quality metrics show better values for the second example. Another reason for that can be found in the different blur kernels which are shown in Fig. 4. The second example is convolved with a PSF that has a weaker ridge along its motion path with only two achnor points, whereas the first example has a PSF with a stronger ridge that is equally thick along its whole motion path. Therefore, the first PSF mixes more pixels and it is more ill-posed to deconvolve. On the other hand, the second PSF mixes two locally aggregated clusters of pixels (due to its two main anchor points) which are seperated relatively far from each other. It can be seen in both cases that the Gaussian prior performs better than Richardson-Lucy, although it does not even conform with the real kurtotic model of the gradient distribution. The Gaussian prior was only justified because it is inexpensive to compute. But still its smoothing capabilities successfully reduce noise and hence outperform Richardson-Lucy. As expected, the Laplacian prior performs a little better but at the cost of much higher computation time. The sparse prior is in most cases an enhancement over the Laplacian, and as shown, even sub-optimal parameters tend to give good results. The color prior again adds more computational costs, but only minor improvements can be visually recognized, like some sharper edges and slightly reduced color noise, in the results of Fig. 8. The quantitative metrics are even a little worse when the color prior is enabled. The border effects in the deconvolution results are common artifacts which (Zhou et al., 2014) claims to reduce.

5 CONTEXT-ADAPTIVE PRIOR

Our experiments with the sparse priors suggest that these tend to oversmooth the result image if the chosen regularization factor λ_∇ is too large, and on the other hand produce a noisy result image if λ_∇ is too small. We therefore suggest a variable regularization that can adapt to local image structure. This would allow the user to have more control over the trade-off between regularization blur and noise, by choosing a stronger regularization in locally smooth image regions where blur does not cause so much trouble, and a weaker regularization in highly structured areas (at the cost of introducing noise at these locations). We experimentally study manual adaptation with user



(a) Fig. 4(g), $\lambda_\alpha = 5$ (b) Fig. 4(d), $\lambda_\alpha = 100$ (c) Fig. 4(g), $\lambda_d = 5$ (d) Fig. 4(d), $\lambda_d = 100$

Figure 7: Effects of the alpha and distance penalty terms on images of Fig. 4(g) and Fig. 4(d).

(a) Ground truth image

(b) Synthetically blurred image

(c) Richardson-Lucy MSSIM 0.455, PSNR 18.37 dB

(d) Gaussian prior: $\gamma = 2$, $\lambda_\nabla = 8$ MSSIM 0.560, PSNR 20.43 dB

(e) Laplacian prior: $\gamma = 1$, $\lambda_\nabla = 2$ MSSIM 0.605, PSNR 20.88 dB

(f) Sparse prior: $\gamma = 0.5$, $\lambda_\nabla = 1$ MSSIM 0.618, PSNR 20.87 dB

(g) Sparse prior: $\gamma = 0.5$, $\lambda_\nabla = 2$ MSSIM 0.585, PSNR 20.49 dB

(h) Color prior $\gamma = 0.5$, $\lambda_\nabla = 2$, $\lambda_\alpha = 5$, $\lambda_d = 5$ MSSIM 0.602, PSNR 20.79 dB

(i) Cropped details, $4\times$ enlarged

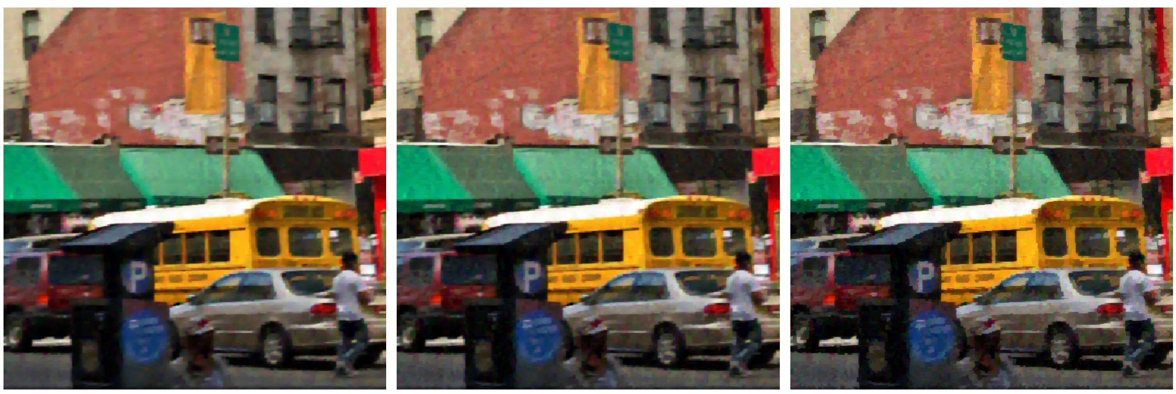
Figure 8: Deconvolution of the image of Fig. 4(c) with ground truth kernel and noise level $\sigma = 5\%$.

(a) Ground truth image (b) Synthetically blurred image (c) Richardson-Lucy MSSIM 0.548, PSNR 18.62 dB

(d) Gaussian prior: $\gamma = 2$, $\lambda_\nabla = 9$ MSSIM 0.599, PSNR 21.27 dB (e) Laplacian prior: $\gamma = 1$, $\lambda_\nabla = 3$ MSSIM 0.674, PSNR 23.48 dB (f) Sparse prior: $\gamma = 0.5$, $\lambda_\nabla = 1.5$ MSSIM 0.688, PSNR 23.72 dB

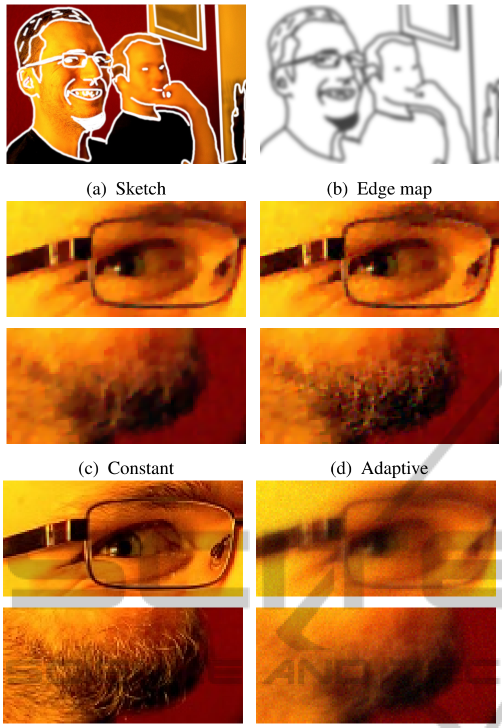
(g) Sparse prior: $\gamma = 0.8$, $\lambda_\nabla = 2.0$ MSSIM 0.683, PSNR 23.70 dB (h) Color prior $\gamma = 0.5$, $\lambda_\nabla = 2$, $\lambda_\alpha = 1$, $\lambda_d = 1$ MSSIM 0.666, PSNR 22.59 dB (i) Cropped details, $4\times$ enlarged

Figure 9: Deconvolution of the image of Fig. 4(h) with ground truth kernel and noise level $\sigma = 1\%$.



(a) $\lambda_\alpha = 0.1$ MSSIM 0.682, PSNR 22.15 dB (b) $\lambda_\alpha = 2$ MSSIM 0.685, PSNR 22.19 dB (c) $\lambda_\alpha = 100$ MSSIM 0.674, PSNR 21.86 dB

Figure 10: Deconvolution of the image of Fig. 4(c) with fixed $\sigma = 2.5\%$, $\lambda_\nabla = 2, \lambda_d = 0$ but varying λ_α.



(a) Sketch (b) Edge map (c) Constant (d) Adaptive (e) Ground truth (f) Blurry, $\sigma = 2.5\%$

Figure 11: Locally varying vs. constant regularization of the sparse prior, besides ground truth and blurry images.

intervention. The idea is to provide the user with either the blurry image or, if the blur is too strong to be able to recognize regions of salient structure, a rough estimate of the deblurred image from the first IRLS iteration. The user can then paint over the edges and other structured areas of the image to indicate weaker regularization, as illustrated in Fig. 11(a). The lines painted by the user can then be blurred slightly using a Gaussian filter to make the change in regularization less abrupt. The resulting image is then inverted and a threshold is introduced so that the sketched areas also experience a certain minimum amount of regularization (e. g. at least 25% regularization in comparison to areas where the user has indicated no important structure). This leads to an edge map like the one shown in Fig. 11(b). The intensities resulting from this edge map are then added as additional weights (multipliers) $\lambda_{\nabla,k,i}$ to the penalty terms $\rho_\nabla(d_{k,i})$. The results shown in Fig. 11 are encouraging. Recently, (Cui et al., 2014) proposed a similar regularization approach as an extension to Richardson-Lucy which is reported to successfully reduce ringing artifacts.

6 CONCLUSIONS

On the basis of the work by (Levin et al., 2007b) and their hyper-Laplacian penalty term, an extensible software framework for deconvolution using the IRLS method has been developed. Because regular photographs contain more than just intensity information, a further regularization approach based upon the two-color model proposed by (Joshi et al., 2009) has been re-implemented and integrated into our optimization framework. In the evaluation part, we proposed suitable regularization parameters for the presented penalty terms. Although enabling the additional color prior results in slightly sharper edges for some images (e.g. Fig. 8), its huge computational cost may not justify its general usage. Just using the sparse gradient prior even with the faulty Gaussian optimization model significantly performs better than Richardson-Lucy. With the additional computational cost when optimizing with a hyper-Laplacian exponent that better models the kurtotic shape of the sparse gradient distribution in an image, the deconvolution results take another significant leap forward. Finally, our experimental work showed that further context-adaptive regularization of gradient priors seems promising in avoiding over-smoothing. The presented deconvolution approach is robust in terms of image noise but performs poorly in case the blur kernel is not perfectly estimated (Zhong et al., 2013).

REFERENCES

Cecchi, G. A., Rao, A. R., Xiao, Y., and Kaplan, E. (2010). Statistics of natural scenes and cortical color processing. *Journal of Vision*, 10(11):1–13.

Chan, T. and Shen, J. (2005). *Image Processing And Analysis: Variational, Pde, Wavelet, And Stochastic Methods*. Siam.

Cho, S., Wang, J., and Lee, S. (2012). Video deblurring for hand-held cameras using patch-based synthesis. *ACM Trans. on Graphics*, 31(4):64:1–64:9.

Cui, G., Feng, H., Xu, Z., Li, Q., and Chen, Y. (2014). A modified Richardson–Lucy algorithm for a single image with adaptive reference maps. *Optics & Laser Technology*, 58:100–109.

Dong, W., Feng, H., Xu, Z., and Li, Q. (2012a). Blind image deconvolution using the fields of experts prior. *Optics Communications*, 285:5051–5061.

Dong, W., Shi, G., Li, X., Zhang, L., and Wu, X. (2012b). Image reconstruction with locally adaptive sparsity and nonlocal robust regularization. *Image Communication*, 27:1109–1122.

Faraway, J. J. (2002). Practical regression and ANOVA using R. Online. http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf.

Gentle, J. (2007). *Matrix algebra: theory, computations, and applications in statistics*. Springer texts in statistics. Springer, New York, NY.

Gupta, A., Joshi, N., Zitnick, L., Cohen, M., and Curless, B. (2010). Single image deblurring using motion density functions. In *ECCV '10: Proc. of the 10th Europ. Conf. on Comp. Vision*.

Huang, J. (2000). *Statistics of natural images and models*. PhD thesis, Brown Univ., Providence.

Ito, A., Sankaranarayanan, A. C., Veeraraghavan, A., and Baraniuk, R. G. (2013). Blurburst: Removing blur due to camera shake using multiple images. *ACM Trans. Graph.* Submitted.

Ji, H. and Wang, K. (2012). A two-stage approach to blind spatially-varying motion deblurring. In *IEEE Conf. on Comp. Vision and Pattern Recogn. (CVPR)*, pages 73–80. IEEE Comp. Soc.

Joshi, N., Zitnick, C., Szeliski, R., and Kriegman, D. (2009). Image deblurring and denoising using color priors. In *IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, pages 1550–1557. IEEE Comp. Soc.

Krishnan, D. and Fergus, R. (2009). Fast image deconvolution using hyper-laplacian priors. In *Neural Inform. Proc. Sys.*

Levin, A., Fergus, R., Durand, F., and Freeman, W. T. (2007a). Deconvolution using natural image priors. Technical report, MIT.

Levin, A., Fergus, R., Durand, F., and Freeman, W. T. (2007b). Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26.

Levin, A. and Weiss, Y. (2007). User assisted separation of reflections from a single image using a sparsity prior. *IEEE Trans. Pattern Anal. and Mach. Intell.*, 29(9):1647–1654.

Li, X., Pan, J., Lin, Y., and Su, Z. (2013). Fast blind deblurring via normalized sparsity prior. *Journal of Inform. & Comp. Sc.*, 10(16):5083–5091.

Lin, H. T., Tai, Y.-W., and Brown, M. S. (2011). Motion regularization for matting motion blurred objects. *IEEE Trans. Pattern Anal. and Mach. Intell.*, 33(11):2329–2336.

Liu, R., Li, Z., and Jia, J. (2008). Image partial blur detection and classification. In *IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE Comp. Soc.

Marichal, X., Ma, W.-Y., and Zhang, H.-J. (1999). Blur determination in the compressed domain using DCT information. In *Intern. Conf. on Image Proc. (ICIP)*, pages 386–390.

Meer, P. (2004). Robust techniques for computer vision. In Medioni, G. and Kang, S. B., editors, *Emerging Topics in Computer Vision*, chapter 4. Prentice Hall PTR, Upper Saddle River, NJ.

Scales, J. A. and Gersztenkorn, A. (1988). Robust methods in inverse theory. *Inverse Problems*, 4(4):1071.

Scales, J. A., Gersztenkorn, A., and Treitel, S. (1988). Fast l_p solution of large, sparse, linear systems: Application to seismic travel time tomography. *Journal of Computational Physics*, 75:314–333.

Schmidt, U., Rother, C., Nowozin, S., Jancsary, J., and Roth, S. (2013). Discriminative non-blind deblurring. In *IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, pages 604–611. IEEE Comp. Soc.

Schmidt, U., Schelten, K., and Roth, S. (2011). Bayesian deblurring with integrated noise estimation. In *IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, pages 2625–2632. IEEE Comp. Soc.

Shi, M.-z., Xu, T.-f., Feng, L., Liang, J., and Zhang, K. (2013). Single image deblurring using novel image prior constraints. *Optik - International Journal for Light and Electron Optics*, 124(20):4429–4434.

Simoncelli, E. (1997). Statistical models for images: Compression, restoration and synthesis. In *In 31st Asilomar Conf on Signals, Systems and Computers*, pages 673–678. IEEE Comp. Soc.

Sorel, M. and Sroubek, F. (2012). Restoration in the presence of unknown spatially varying blur. In Gunturk, B. and Li, X., editors, *Image Restoration: Fundamentals and Advances*, pages 63–88. CRC Press.

Srivastava, A., Lee, A. B., Simoncelli, E. P., and c. Zhu, S. (2003). On advances in statistical modeling of natural images. *Journal of Math. Imaging and Vision*, 18:17–33.

Wang, S., Hou, T., Border, J., Qin, H., and Miller, R. (2012). High-quality image deblurring with panchromatic pixels. *ACM Trans. Graph.*, 31(5).

Wang, Z., Simoncelli, E. P., and Bovik, A. C. (2003). Multiscale structural similarity for image quality assessment. In *Proc. of the 37th IEEE Asilomar Conf. on Signals, Systems and Computers*.

Weiss, Y. and Freeman, W. (2007). What makes a good model of natural images? In *IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE Comp. Soc.

Whyte, O., Sivic, J., Zisserman, A., and Ponce, J. (2012). Non-uniform deblurring for shaken images. *Intern. Journal of Comp. Vision*, 98(2):168–186.

Yuan, L., Sun, J., Quan, L., and Shum, H.-Y. (2007). Blurred/non-blurred image alignment using sparseness prior. In *Proc. of the 11th Intern. Conf. on Comp. Vision (ICCV)*, pages 1–8.

Yun-Fang, Z. (2010). Blur detection for surveillance video based on heavy-tailed distribution. In *Asia Pacific Conf. on Postgrad. Research in Microelectronics and Electronics*, pages 101–105.

Zhong, L., Cho, S., Metaxas, D., Paris, S., and Wang, J. (2013). Handling noise in single image deblurring using directional filters. In *IEEE Conf. on Comp. Vision and Pattern Recogn. (CVPR)*, pages 612–619. IEEE Comp. Soc.

Zhou, X., Zhou, F., Bai, X., and Xue, B. (2014). A boundary condition based deconvolution framework for image deblurring. *Journal of Computational and Applied Mathematics*, 261:14–19.