High-Quality Computational Imaging Through Simple Lenses

Felix Heide[1], Mushfiqur Rouf[1], Matthias B. Hullin[1], Björn Labitzke[2], Wolfgang Heidrich[1], Andreas Kolb[2]

[1]University of British Columbia, [2]University of Siegen

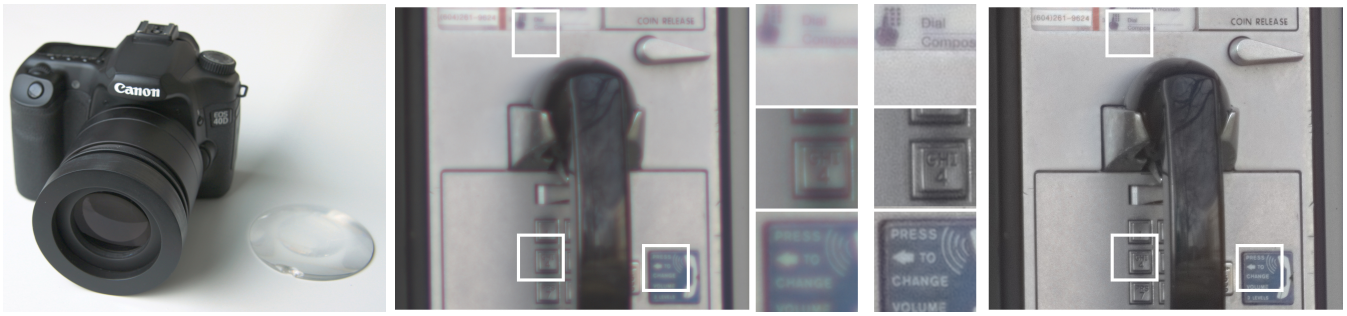

Fig. 1. Our system reliably estimates point spread functions of a given optical system, enabling the capture of high-quality imagery through poorly performing lenses. From left to right: Camera with our lens system containing only a single glass element (the plano-convex lens lying next to the camera in the left image), unprocessed input image, deblurred result.

Modern imaging optics are highly complex systems consisting of up to two dozen individual optical elements. This complexity is required in order to compensate for the geometric and chromatic aberrations of a single lens, including geometric distortion, field curvature, wavelength-dependent blur, and color fringing.

In this paper, we propose a set of computational photography techniques that remove these artifacts, and thus allow for post-capture correction of images captured through uncompensated, simple optics which are lighter and significantly less expensive. Specifically, we estimate per-channel, spatially-varying point spread functions, and perform non-blind deconvolution with a novel cross-channel term that is designed to specifically eliminate color fringing.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Picture/Image Generation—*Digitizing and scanning*

Additional Key Words and Phrases: Computational Photography, Deconvolution, Optimization, Optics

1. INTRODUCTION

Over the past decades, camera optics have become increasingly complex. The lenses of modern single lens reflex (SLR) cameras may contain a dozen or more individual lens elements, which are used to optimize light efficiency of the optical system while minimizing *aberrations*, i.e., non-linear deviations from an idealized thin lens model.

Optical aberrations include effects such as geometric distortions, chromatic aberration (wavelength-dependent focal length), spherical aberration (focal length depends on the distance from the optical axis), and coma (angular dependence on focus) [Mahajan 1991]. All single lens elements with spherical surfaces suffer from these artifacts, and as a result cannot be used in high-resolution, high-quality photography. Instead, modern optical systems feature a combination of different lens elements with the intent of canceling out aberrations. For example, an achromatic doublet is a compound lens made from two glass types of different dispersion, i.e., their refractive indices depend on the wavelength of light differently. The result is a lens that is (in the first order) compensated for chromatic aberration, but still suffers from the other artifacts mentioned above.

Despite their better geometric imaging properties, modern lens designs are not without disadvantages, including a significant impact on the cost and weight of camera objectives, as well as increased lens flare.

In this paper, we propose an alternative approach to high-quality photography: instead of ever more complex optics, we propose to revisit much simpler optics used for hundreds of years (see, e.g., [Rashed 1990]), while correcting for the ensuing aberrations computationally. While this idea is not entirely new (see, e.g., [Schuler et al. 2011]), our approach, which exploits cross-channel information, is significantly more robust than other methods that have been proposed. It is therefore able to handle much larger and more dense blur kernels, such as disk-shaped kernels with diameters of 50 pixels and more, which occur frequently in uncorrected optics unless the apertures are stopped down to an impractical size. A key component to achieve this robustness is the development of a convex solver with guaranteed convergence properties that can minimize the resulting cross-channel deconvolution problem. In fact, our experiments show that a failure of current methods to arrive at a global optimum of the objective function limits the quality of state-of-the-art deconvolution methods. The specific technical contributions that enable the use of simple lens designs for high-quality photography are

—a new cross-channel prior for color images,

—a deconvolution method and convex solver that can efficiently incorporate this prior and is guaranteed to converge to a global optimum, and

—a robust approach for per-channel spatially-varying PSF estimation using a total variation (TV) prior based on the same optimization framework.

We demonstrate our method by producing high-quality photographs on modern 12 megapixel digital SLRs using only single lens elements such as plano-convex or biconvex lenses, and achromatic doublets. We achieve high quality results comparable to conventional lenses, with apertures around $f/4.5$. Finally, we show that our formulation outperforms competing deconvolution approaches in terms of both quality and computational performance. Using our techniques, simplified camera optics become viable alternatives to conventional camera objectives without the disadvantages outlined above.

2. RELATED WORK

Only a few years after Gauss introduced the linear model of optics [1841] that gives rise to the thin lens model, Seidel [1857] introduced aberration theory, i.e., the study of non-linear effects in optical systems. The common way to correct for these effects in optical systems is to design increasingly complex systems with larger and larger numbers of individual lens elements [Mahajan 1991]. An alternative way to deal with aberrations is to remove the resulting blur after the capture in a deconvolution step.

Image Deconvolution.
The most basic deconvolution approaches include frequency-space division and the Wiener filter [Wiener 1964], which however deal poorly with frequencies that are suppressed by the blur kernel. A classical iterative method for image deconvolution is Richardson-Lucy [Richardson 1972; Lucy 1974], which was developed for use in astronomy to compensate for blur in optical systems that are insufficiently corrected for aberrations. The basic Richardson-Lucy algorithm has been extended in many ways, including residual deconvolution [Yuan et al. 2007] and Bilateral Richardson-Lucy [Yuan et al. 2008].

Recently, there has been a lot of work on incorporating image priors into the deconvolution process, as an attempt to better cope with very broad blur kernels and the complete loss of certain frequencies. Such priors include convex terms such as total variation (TV, e.g., [Chan et al. 2011]) and total curvature [Goldluecke and Cremers 2011], which can be optimized with specialized solvers that are guaranteed to converge to the globally optimal solution. Other authors have used non-convex regularization terms that are optimized using techniques such as Iteratively Reweighted Least Squares (IRLS) [Levin et al. 2007; Joshi et al. 2009]. These methods are not guaranteed to converge to a global optimum, but produce state-of-the-art results in practice. Krishnan and Fergus [2009] introduce a prior based on the Hyper-Laplacian, and an efficient solver to implement it. Like IRLS, this new solver is not guaranteed to converge to a globally optimal solution. We show in Appendix B that this limitation is in fact practically relevant, and that a TV solver with guaranteed global convergence will often produce sparser gradient distributions than these Hyper-Laplacian solvers.

PSF Estimation.
Although blind deconvolution methods exist, the best image quality is achieved with known blur kernels. Several methods for estimating the kernels exist, including spectral methods [Gibson and Bovik 2000] and estimation from edge-based [Joshi et al. 2008] or noise-based patterns [Brauers et al. 2010] or image pairs [Yuan et al. 2007].

Deconvolution for Aberration Correction.
As mentioned above, the use of deconvolution algorithms for correcting for aberrations dates back at least to the original development of the Richardson-Lucy algorithm. More recently, there has been renewed interest in solving such problems with modern techniques, including deconvolution for spatially varying PSFs [Kee et al. 2011] and the removal of chromatic aberrations [Chung et al. 2009; Joshi et al. 2009], both for images taken with complex optical systems. We note that in both cases the point spread functions are significantly smaller than with the simple lenses we use. Joshi et al.'s color prior [2009] encourages color gradients that are linear blends of two base colors. This is not a good prior for removing color fringing from chromatic aberration since large-area color fringing is typically composed of several segments that are linear mixtures of different base colors.

Kang [2007] proposes a method specifically for removing color fringing in images. However, this method is based on edge detection, which is feasible for images taken with partially corrected optics, but is impossible in the presence of very large blurs that result from the use of uncorrected optics.

Cho et al. [2010; 2012] propose a deconvolution method that uses locally learned gradient statistics for deconvolution. Although this has not been attempted in the literature, one could imagine using this method to transfer gradient information from one channel to another. Instead, our approach is to directly match point-wise gradients in the different channels, which provides better localization and therefore suppression of color fringing than an approach based on area statistics.

Schuler et al. [2011] presented the work most similar to ours. They solve the demosaicing problem as well as the deconvolution problem at the same time by working in Yuv color space. While working in this color space is a strong-enough prior to avoid color artifacts stemming from PSFs with relatively sparse features and thin structures, this approach is not sufficient to avoid artifacts in the presence of large disk-like blurs such as the ones produced by spherical and chromatic aberrations (Section 7). Schuler et al.'s method also suffers from the lack of global convergence stemming from their use of a Hyper-Laplacian prior.

In contrast to all these methods, our approach uses a *convex* cross-channel prior that can be implemented efficiently and with guaranteed global convergence. As a result, it can produce excellent image quality even in the presence of very large blur kernels. We emphasize that our contribution is not only to introduce a color prior, but also to find a convex optimization framework for this prior, which is a key component in achieving excellent image quality in the presence of very large blur kernels.

Finally, there has been work on using chromatic aberrations to increase the depth of field of the imaging system [Cossairt and Nayar 2010; Guichard et al. 2009]. This problem is somewhat orthogonal to our goal, and could be added to our method as well. We also note that Cossairt and Nayar [2010] still require an optical system in which all aberrations *except* for chromatic ones are minimized.

3. OVERVIEW

The goal of our work is to replace complex optical systems used in modern camera objectives with very simple lens designs such as plano-convex lenses, while achieving image quality comparable to modern cameras at the full resolution of current image sensors (Fig. 2).

The primary challenge in achieving these goals is that simple lenses with spherical interfaces exhibit *aberrations*, i.e., higher-order deviations from the ideal linear thin lens model [Seidel 1857]. These aberrations cause rays from object points to focus imperfectly onto a single image point, thus creating complicated point spread functions that vary over the image plane, which need to be

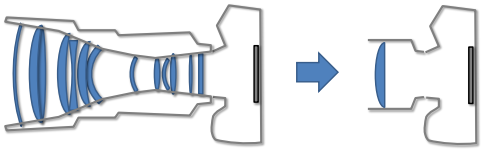

Fig. 2. In this paper we propose to replace complex lens systems by a single lens element plus a set of computational photography techniques.

removed through deconvolution. The effect gets more pronounced at larger apertures where more off-axis rays contribute to the image formation. Figure 3 shows the variation of the PSF over the image plane for two simple lens elements. We can make several observations from these images:

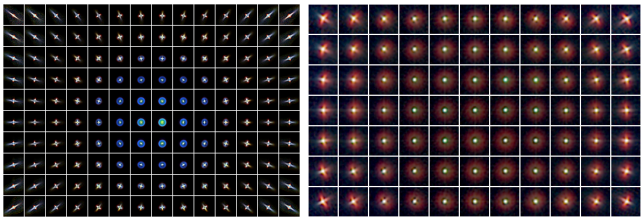

Fig. 3. Patch-wise estimated PSFs for two simple lenses. The PSF estimation and non-blind deblurring in our method is done in patches to account for the PSFs' spatial variance. Left: PSF of a single biconvex lens at $f/2.0$. Right: PSF of a single plano-convex lens at $f/4.5$.

—The blur is highly spatially varying, ranging from disc-like structures (spherical aberration) with diameters of 50 pixels and more to elongated streaks (coma and astigmatism). We can address this problem by subdividing the image into tiles over which we assume a constant point spread function (see, e.g., [Levin et al. 2007]).

—The blur is highly wavelength dependent (chromatic aberration). This results in objectionable color fringing in the image. At the same time, the PSF of at least one of the color channels often contains more energy in high spatial frequencies than the others (one channel is usually focused significantly better than the others); note here that we do not require it to be perfectly in focus. This suggests that we may be able to utilize cross-channel information and reconstruct spatial frequencies that are preserved in at least one of the channels.

The rest of the paper is structured as follows. In Section 5 we describe a cross-channel prior, and an efficient convex optimization algorithm for solving the deconvolution problem. Section 6 describes a PSF estimation algorithm for recovering tile-based PSFs such as the ones shown in Figure 3. Both the deconvolution algorithm and the PSF estimation use the convex optimization framework by Chambolle and Pock [2011], which we summarize in Section 4. We conclude with results and a discussion in Sections 7 and 8.

4. REVIEW OF OPTIMAL FIRST-ORDER PRIMAL-DUAL CONVEX OPTIMIZATION

To solve both the image deconvolution and the PSF estimation problem for working with simple lenses, we derive optimization methods based on the optimal first-order primal-dual framework by Chambolle and Pock [2011]. In this section we present a short overview of this optimization framework. We refer the reader to the original work by Chambolle and Pock for an in-depth discussion.

The optimization framework considers general problems of the form

$$\mathbf{x}_{\text{opt}} = \underset{\mathbf{x}}{\operatorname{argmin}} \ \mathbf{F}(\mathbf{K}\mathbf{x}) + \mathbf{G}(\mathbf{x}). \tag{1}$$

For example, in an inverse problem with a TV regularizer, the first term in Eq. (1) is the regularizer (that is $\mathbf{K}(\mathbf{x}) = \nabla\mathbf{x}$, $\mathbf{F}(\mathbf{y}) = \|\mathbf{y}\|_1$ for TV), while the second term is the data fitting term (some residual norm).

Let X, Y be finite-dimensional real vector spaces for the primal and dual space, respectively. The operators from Eq. (1) are then formally defined as:

$$\begin{array}{lll} \mathbf{K} & : X \to Y & \text{is a linear operator from } X \text{ to } Y \\ \mathbf{G} & : X \to [0,+\infty) & \text{is a proper, convex, (l.s.c.) function.} \\ \mathbf{F} & : Y \to [0,+\infty) & \text{is a proper, convex, (l.s.c.) function.} \end{array} \tag{2}$$

where l.s.c. stands for lower-semicontinuous. The dual problem of Eq. (1) is given as

$$\mathbf{y}_{\text{opt}} = \underset{\mathbf{y}}{\operatorname{argmax}} \ -\left(\mathbf{G}^*\left(-\mathbf{K}^*\mathbf{y}\right) + \mathbf{F}^*(\mathbf{y})\right), \tag{3}$$

where the * denotes the convex conjugate. To solve the above (primal and dual) problem, the following algorithm is proposed by Chambolle and Pock [2011]:

ALGORITHM 1. (First-order primal-dual algorithm)

—*Initialization: Choose* $\tau, \sigma \in \mathbb{R}^+$, $\theta = 1$ *with* $\tau\sigma L^2 < 1$
Choose initial iterates $(\mathbf{x}_0, \mathbf{y}_0) \in X \times Y$, $\bar{\mathbf{x}}_0 = \mathbf{x}_0$

—*Iterations* $(n \geq 0)$*: Update* $\mathbf{x}_n, \mathbf{y}_n, \bar{\mathbf{x}}_n$ *as following:*

$$\begin{cases} \mathbf{y}_{n+1} &= \mathbf{prox}_{\sigma\mathbf{F}^*}(\mathbf{y}_n + \sigma\mathbf{K}\bar{\mathbf{x}}_n) \\ \mathbf{x}_{n+1} &= \mathbf{prox}_{\tau\mathbf{G}}(\mathbf{x}_n + \tau\mathbf{K}^*\mathbf{y}_{n+1}) \\ \bar{\mathbf{x}}_{n+1} &= \mathbf{x}_{n+1} + \theta\left(\mathbf{x}_{n+1} - \mathbf{x}_n\right) \end{cases}$$

The *resolvent* or *proximal* operator with respect to $\mathbf{G}$ is defined as:

$$\begin{aligned} \mathbf{prox}_{\tau G}(\tilde{\mathbf{x}}) \ &:= \ \left(\mathbb{I} + \tau\partial\mathbf{G}\right)^{-1}(\tilde{\mathbf{x}}) \\ &= \ \underset{\mathbf{x}}{\operatorname{argmin}} \ \frac{1}{2\tau}\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2 + \mathbf{G}(\mathbf{x}) \end{aligned} \tag{4}$$

and analogously for $\mathbf{prox}_{\sigma\mathbf{F}^*} := \left(\mathbb{I} + \sigma\partial\mathbf{F}^*\right)^{-1}$. It is assumed that these operators are easy to compute, for example in closed form or using a Newton-like method. The parameter L, which is necessary to compute valid σ, τ, is defined as the operator norm $L = \|\mathbf{K}\|_2$.

Note that Algorithm 1 never minimizes functions including both $\mathbf{G}$ and $\mathbf{F}^*$ at the same time. This splitting of the minimizations for $\mathbf{G}$ and $\mathbf{F}$ from Eq. (1) and the alternation between those minimizations is the high-level idea behind all forward-backward splitting methods [Combettes and Pesquet 2011], [Friedlander 2011]. In each iteration, these methods first perform a forward step to minimize $\mathbf{G}$ in the objective function and then minimize $\mathbf{F}$, where both minimizations are coupled by terms that ensure that the solution is close to the previous solution (note the first term in the minimization in Eq. (4)). A basic forward-backward method that illustrates this idea is the projected gradient method. Consider the example of solving Eq. (1) with $\mathbf{K}(\mathbf{x}) = \mathbf{x}, \mathbf{F}(\mathbf{y}) = \delta_{\mathcal{C}}(\mathbf{y})$ and arbitrary $\mathbf{G}$, where $\delta_{\mathcal{C}}(\mathbf{y})$ is the indicator function for a set $\mathcal{C}$ that is 1 if $\mathbf{y} \in \mathcal{C}$, else ∞. Thus, in this example $\mathbf{G}$ is to be minimized with the hard constraint $\mathbf{x} \in \mathcal{C}$. The projected gradient method would be the following iteration over $n \geq 0$:

$$\mathbf{x}_{n+1} = \underbrace{\mathbf{prox}_{\sigma\mathbf{F}}}_{\text{backward step}} \underbrace{(\mathbf{x}_n - \sigma\nabla\mathbf{G}(\mathbf{x}_n))}_{\text{forward step}}, \tag{5}$$

where σ is a step size. The iteration is illustrated in Fig. 4. We can see that in each iteration the projected gradient method first performs a gradient descent step and then a prox-step for $\mathbf{F}$. For the considered problem this is a Euclidean projection on the hard constraint set $\mathcal{C}$. Consequently the resolvent operators from Eq. (4) can be thought of as a generalization of the projection.

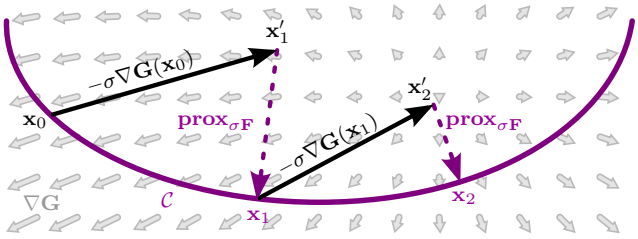

Fig. 4. Projected gradient method that illustrates the high-level idea behind forward-backward splitting methods: Alternating sequence of gradient descent (forward) steps and projection (backward) steps.

Many known sparse optimization methods fall into the class of forward-backward splitting methods, illustrated by the projected gradient example above. Other forward-backward splitting methods include ISTA, FISTA, Split Bregman, ADMM [Friedlander 2011]. A distinguishing property of Algorithm 1 is that the authors show that it converges to a global minimum with the rate $O(1/N)$, which is *optimal* for the general problem from Eq. (1) (see again [Chambolle and Pock 2011]).

5. DEBLURRING

Having summarized the general optimization framework, we now derive a specific and efficient deconvolution algorithm based on this framework. We start by introducing an image formation model, and motivating a novel cross-channel prior for deconvolving multi-channel images, then define an optimization function for deconvolution with this term, and cast it into the framework from the previous section. Next, we discuss an improvement of the basic method to deal with dark image regions, and finally we describe how the developed core method can be integrated into an efficient scale-space deconvolution method.

5.1 Image Formation Model

Consider a grayscale image tile captured with $n \times m$ resolution. Let $J, I, N \in \mathbb{R}^{n\times m}$ be the observed blurred image, the underlying sharp image and additive image noise, respectively. The formation of the blurred observation with the blur kernel B can then be formulated as

$$J = B \otimes I + N \tag{6}$$
$$\mathbf{j} = \mathbf{B}\mathbf{i} + \mathbf{n}, \tag{7}$$

In the second form, $\mathbf{B}, \mathbf{j}, \mathbf{i}$, and $\mathbf{n}$ are the corresponding quantities in matrix-vector form. As mentioned in the overview, an actual camera image will be composed of many tiles, each with a PSF that is assumed constant over the tile.

5.2 Deconvolution with Cross-Channel Information

Real optical systems suffer from dispersion in the lens elements, leading to a wavelength dependency of the PSF known as *chromatic aberration*. While complex modern lens assemblies are designed to minimize these artifacts through the use of multiple lens elements that compensate for each others' aberrations, it is worth pointing out that even very good lenses still have a residual amount of chromatic aberration. For the simple lenses we aim to use, the chromatic aberrations are very severe – one color channel is focused significantly better (although never perfectly in focus) than the other channels, which are blurred beyond recognition (of course excluding achromatic lenses which compensate for chromatic aberrations).

Given individual PSFs $B_{\{1\ldots3\}}$ for each color channel $J_{\{1\ldots3\}}$ of an image J one might attempt to independently deconvolve each color channel. As Figure 5 demonstrates, this approach does not in general produce acceptable results, since frequencies in some of the channels may be distorted beyond recovery. Note the severe ringing in the top center and strong residual blur in the bottom center.

Fig. 5. Effect of using the cross-channel priors (with the addition from Sec. 5.5) for two image patches (full images in results section). Left: Original blurred capture. Middle: Independently deconvolved results for each channel using the method of Levin et al. [2007]. Right: Deconvolution result with our cross channel prior.

We propose to share information between the deconvolution processes of the different channels, so that frequencies preserved in one channel can be used to help the reconstruction in another channel. Our cross-channel prior is based on the assumption that edges in the image appear in the same place in all channels, and that hue changes are sparse throughout the image (also see Figure 6). These assumptions lead to the following prior for a pair of channels l, k

$$\begin{aligned} &\nabla\mathbf{i}_k/\mathbf{i}_k \approx \nabla\mathbf{i}_l/\mathbf{i}_l \\ \Leftrightarrow\ &\nabla\mathbf{i}_k \cdot \mathbf{i}_l \approx \nabla\mathbf{i}_l \cdot \mathbf{i}_k, \end{aligned} \tag{8}$$

which we enforce in a sparse (ℓ_1) fashion. Note that the division and multiplication $/, \cdot$ are pixel-wise operators.

5.3 Minimization Problem

Using this cross-channel prior, we formulate the problem of jointly deconvolving all channels as the optimization problem

$$\begin{aligned} (\mathbf{i}_{1\ldots3})_{\text{opt}} = \underset{\mathbf{i}_{1\ldots3}}{\operatorname{argmin}} \ &\sum_{c=1}^{3} \|\mathbf{B}_c\mathbf{i}_c - \mathbf{j}_c\|_2^2 + \lambda_c \sum_{a=1}^{5} \|\mathbf{H}_a\mathbf{i}_c\|_1 \\ &+ \sum_{l\neq c} \beta_{cl} \sum_{a=1}^{2} \|\mathbf{H}_a\mathbf{i}_c \cdot \mathbf{i}_l - \mathbf{H}_a\mathbf{i}_l \cdot \mathbf{i}_c\|_1, \end{aligned} \tag{9}$$

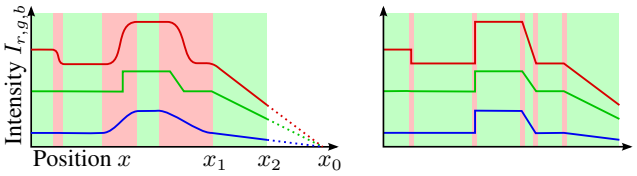

Fig. 6. Blurred scanline on the left with different PSFs in each channel and a sharp green channel. Reconstructed scanline on the right. Our cross-channel regularizer enforces gradient consistency between the color channels and allows sparse hue changes. Constant regions and pure luma gradients have a low regularizer energy (regions shaded in light green), but changes in chroma are penalized (red regions). For the region between x_1 and x_2, the extrapolated intensities of all channels intersect the x-axis in the same point x_0, which visualizes Equation 8. By using an ℓ_1 norm, sparse chroma changes occurring exactly at edges (right) are preferred over wider regions (left) that are typically perceived as color fringing.

where the first term is a standard least-squares data fitting term, and the second term enforces a heavy-tailed distribution for both gradients and curvatures. The convolution matrices $\mathbf{H}_{\{1,2\}}$, implement the first derivatives, while $\mathbf{H}_{\{3\dots5\}}$ correspond to the second derivatives. We use the same kernels as Levin et al. [2007] but employ an ℓ_1 norm in our method rather than a fractional norm. This ensures that our problem is convex. The last term of Eq. (9) implements our cross-channel prior, again with an ℓ_1 norm. $\lambda_c, \beta_{cl} \in \mathbb{R}$ with $c, l \in \{1\dots3\}$ are weights for the image prior and cross-channel prior terms, respectively.

5.4 Deconvolution Algorithm

The minimization from Eq. (9) is implemented by alternately minimizing with respect to one channel while fixing all the other channels. To optimize for this single channel $\mathbf{x} = \mathbf{i}_c$ we derive a first-order primal-dual algorithm adopting the framework described in Sec. 4. First, the optimization is rewritten in matrix-vector form as

$$
\begin{aligned}
(\mathbf{i}_c)_{\text{opt}} &= \underset{\mathbf{x}}{\text{argmin}} \left\| \begin{bmatrix} \lambda_c \mathbf{H}_1 \\ \vdots \\ \lambda_c \mathbf{H}_5 \\ \beta_{cl}\left(\mathbf{D}_{\mathbf{i}_l}\mathbf{H}_1 - \mathbf{D}_{\mathbf{H}_1\mathbf{i}_l}\right) \\ \vdots \\ \beta_{cl}\left(\mathbf{D}_{\mathbf{i}_l}\mathbf{H}_2 - \mathbf{D}_{\mathbf{H}_2\mathbf{i}_l}\right) \\ \vdots \end{bmatrix} \mathbf{x} \right\|_1 + \|\mathbf{B}_c\mathbf{x} - \mathbf{j}_c\|_2^2 \\
&= \underset{\mathbf{x}}{\text{argmin}} \ \|\mathbf{S}\mathbf{x}\|_1 + \|\mathbf{B}_c\mathbf{x} - \mathbf{j}_c\|_2^2
\end{aligned}
\tag{10}
$$

where here $\mathbf{D}$ denotes the diagonal matrix with the diagonal taken from the subscript vector. See Appendix A for a detailed derivation. $\mathbf{S}$ is a matrix consisting of the sequence of all $t = 5 + 2(3-1)$ matrices coming from the ℓ_1 minimization terms in Eq. (9). By comparison with Eq. (1), we can now define

$$
\begin{aligned}
\mathbf{K}(\mathbf{x}) &= \mathbf{S}\mathbf{x} \\
\mathbf{F}(\mathbf{y}) &= \|\mathbf{y}\|_1 \\
\mathbf{G}(\mathbf{x}) &= \|\mathbf{B}_c\mathbf{x} - \mathbf{j}_c\|_2^2 \ .
\end{aligned}
\tag{11}
$$

Given this structure, the following resolvent operators necessary to apply Algorithm 1 are already known [Chambolle and Pock 2011]:

$$
\mathbf{y} = \mathbf{prox}_{\sigma\mathbf{F}^*}(\tilde{\mathbf{y}}) \Leftrightarrow \mathbf{y}_i = \frac{\tilde{\mathbf{y}}_i}{\max(1, |\tilde{\mathbf{y}}_i|)} \tag{12}
$$

$$
\begin{aligned}
\mathbf{x} &= \mathbf{prox}_{\tau\mathbf{G}}(\tilde{\mathbf{x}}) \\
\Leftrightarrow 0 &= \frac{1}{\tau}(\mathbf{x} - \tilde{\mathbf{x}}) + 2\left(\mathbf{B}_c^T\mathbf{B}_c\mathbf{x} - \mathbf{B}_c^T\mathbf{j}\right) \\
\Leftrightarrow \mathbf{x} &= \mathcal{F}^{-1}\left(\frac{\tau 2\mathcal{F}\left(\mathbf{B}_c\right)^*\mathcal{F}\left(\mathbf{j}_c\right) + \mathcal{F}\left(\tilde{\mathbf{x}}\right)}{\tau 2\left|\mathcal{F}\left(\mathbf{B}_c\right)\right|^2 + 1}\right),
\end{aligned}
\tag{13}
$$

where $\mathcal{F}(.)$ in the last line denotes here the Fourier transform and $\tilde{\mathbf{y}}$, $\tilde{\mathbf{x}}$ are the function variables of the proximal operators as defined in Eq. (4). The first proximal operator is just a per-pixel projection operator (for detailed derivation we refer the reader to [Chambolle and Pock 2011]). The second proximal operator is the solution to a linear system as shown in the second line. Since the system matrix is composed of convolution matrices with a large support we can efficiently solve this linear system in the Fourier domain (last line).

What still needs to be determined is the convex conjugate $\mathbf{K}^*$ of the linear operator $\mathbf{K}$, which is given as follows:

$$
\begin{aligned}
\mathbf{K}^* = S^T &= \left[\mathbf{S}_1^T \cdots \mathbf{S}_t^T\right] \\
&= \left[\lambda_c\mathbf{H}_1^T \cdots \lambda_c\mathbf{H}_5^T \quad \beta_{cl}\left(\mathbf{H}_1^T\mathbf{D}_{\mathbf{i}_l} - \mathbf{D}_{\mathbf{H}_1\mathbf{i}_l}\right)\cdots\right],
\end{aligned}
\tag{14}
$$

where t is the number of matrices S is composed of. In summary, the matrix-vector multiplication $\mathbf{K}^*(\mathbf{y})$ in Algorithm 1 can be expressed as the following sum

$$
\mathbf{K}^*(\mathbf{y}) = \mathbf{S}^T\mathbf{y} = \sum_{k=1}^{t} \mathbf{S}_k^T\mathbf{y}_{[(k-1)\cdot nm,\dots,k\cdot nm-1]}, \tag{15}
$$

where each $\mathbf{S}_k^T$ is just a sum of filtering operations and point-wise multiplications. Likewise, the resolvent operators given above can be implemented using small filtering operators or the FFT for larger filters.

Parameter Selection

Algorithm 1 converges to the global optimum of the convex functional if $\theta = 1, \tau\sigma L^2 < 1$ with $\tau, \sigma > 0$ and $L = \|\mathbf{K}\|_2$ (see [Chambolle and Pock 2011] for the proof). We use $\theta = 1, \sigma = 10$ and $\tau = \frac{0.9}{\sigma L^2}$ for the deconvolution algorithm proposed in this paper. That only leaves open how to compute the operator norm L.

Since we have $\mathbf{K}(\mathbf{x}) = \mathbf{S}\mathbf{x}$ where $\mathbf{S}$ was a matrix, $\|\mathbf{K}\|_2$ is the square root of the largest eigenvalue of the symmetric matrix $\mathbf{S}^T\mathbf{S}$ (see, e.g., [Rowland 2012]). We find the value L by using the power iteration where again all matrix-vector-multiplications with $\mathbf{S}^T\mathbf{S}$ can be decomposed into filtering operations (other eigenvalue methods like the Arnoldi iteration are consequently also suited for computing L).

5.5 Regularization for Low-Intensity Areas

In this section, we propose a modification of the basic cross-channel prior that produces improved results in dark regions, i.e., for pixels where all channel values approach zero. In these regions, the prior from Eq. (8) is not effective, since the hue normalization reduces the term to zero. As a result, significant color artifacts (such as color ringing) can remain in dark regions, see Figure 7. Note that by allowing luma gradients in the original cross-prior this is an inherent design problem of this prior and not an optimization issue.

In these regions, we therefore propose to match absolute (rather than relative) gradient strengths between color channels. The operator $\mathbf{G}$ from (11) is modified as

Fig. 7. Chromatic artifacts in low-intensity areas. Left: Original blurred patch. Center: Reconstruction using Eq. (9). Notice the chromatic artifacts in all low-intensity areas and the correct reconstruction in the other areas. Right: Regularization for low-intensity areas added as discussed below.

$$\mathbf{G}(\mathbf{x}) = \|\mathbf{B}_c\mathbf{x} - \mathbf{j}_c\|_2^2 + \lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \|\mathbf{D}_\mathbf{w}\left(\mathbf{H}_a\mathbf{x} - \mathbf{H}_a\mathbf{i}_l\right)\|_2^2, \quad (16)$$

where $\mathbf{D}_\mathbf{w}$ is a spatial mask that selects dark regions below a threshold ϵ. The mask is blurred slightly with a Gaussian kernel K_σ to avoid spatial discontinuities at the borders of regions affected by the additional regularization term:

$$\mathbf{w} = \left(1 - \frac{\sum_l \beta_{cl}\mathbf{T}\left(\mathbf{i}_l\right)}{\sum_l \beta_{cl}}\right) \otimes K_\sigma \quad \text{with} \quad \mathbf{T}\left(\mathbf{i}\right)_k = \begin{cases} 1 & ;\mathbf{i}_k \leq \epsilon \\ 0 & ;\text{else} \end{cases},$$

and $\epsilon = 0.05$ and $\sigma = 3$ in our implementation.

The resolvent operator with respect to $\mathbf{G}$ from Eq. (13) is replaced by the following derivation (see Appendix A for details):

$$\begin{aligned} &\mathbf{u}_{\text{opt}} = \mathbf{prox}_{\tau\mathbf{G}}(\tilde{\mathbf{u}}) \\ \Leftrightarrow &\left[2\tau\mathbf{B}_c^T\mathbf{B}_c + \mathbb{I} + 2\tau\lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \mathbf{H}_a^T\mathbf{D}_\mathbf{w}^2\mathbf{H}_a\right]\mathbf{u}_{\text{opt}} \\ &= 2\tau\mathbf{B}_c^T\mathbf{j}_c + \tilde{\mathbf{u}} + 2\tau\lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \mathbf{H}_a^T\mathbf{D}_\mathbf{w}^2\mathbf{H}_a\mathbf{i}_l \\ \Leftrightarrow &\mathbf{A}\mathbf{u}_{\text{opt}} = \mathbf{b} \end{aligned} \quad (17)$$

This expresses the solution of the resolvent operator as the matrix inversion problem $\mathbf{A}\mathbf{u}_{\text{opt}} = \mathbf{b}$. Since blur kernel sizes of the order of magnitude of $10^2 \times 10^2$ can be expected for practical applications, $\mathbf{A}$ is very large and impractical to invert explicitly. The system is solved using the Conjugate Gradient (CG) algorithm. This allows us to express the matrix-vector multiplication in the CG-algorithm as a sequence of filtering operations as before.

5.6 Residual and Scale-Space Deconvolution

The basic deconvolution approach described so far can be accelerated and further improved by adopting a scale-space approach [Yuan et al. 2008] as well as residual deconvolution [Yuan et al. 2007].

The idea behind residual deconvolution is to iteratively deconvolve the residual image starting with a large regularization weight which is progressively decreased. Since the residual image has a significantly reduced amplitude, its deconvolved reconstruction contains less ringing which is proportional to the amplitude. The iteration is shown in the center of Fig. 8.

Our method handles saturation in the blurred image by removing the rows where $\mathbf{j}$ is saturated from the data fitting term. This is done by pre-multiplying the residual $\mathbf{B}_c\mathbf{x} - \mathbf{j}_c$ with a diagonal weighting matrix whose diagonal is 0 for saturated rows and 1 else; the derivation from Eq. (17) is changed straightforwardly.

To increase the performance of the algorithm by using good starting points for Eq. (10), the method is performed in scale space.

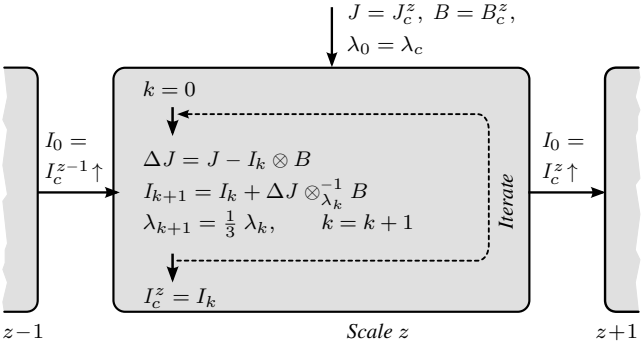

Fig. 8. Iterative deconvolution of the residual image which eliminates ringing artifacts, center. The operator $\otimes_{\lambda_k}^{-1}$ is the deconvolution of the image to the left using the kernel to the right and a regularizer-weight λ_k. For performance reasons the algorithm is implemented in scale space. This figure shows the deconvolution on a level z in the pyramid. We set $I_c^{-1} = 0$.

See again Fig. 8, where $\uparrow$ is an upsampling operator to the next finer scale. We use nearest neighbor since it preserves edges. The pyramids $\{J^z\}_{z=0}^Z$, $\{B^z\}_{z=0}^Z$ of the blurred image/kernel pairs are computed by bicubic downsampling of J, B with the scale factor $\frac{1}{2}$. The reconstruction pyramid $\{I^z\}_{z=0}^Z$ is progressively recovered from coarse (scale 0) to fine, where at each scale the initial iterate is the upsampled result of the next coarser level. Note, that our scale space implementation purely serves as a performance improvement. In particular we do not need to overcome local minima in our minimization problem since it is convex.

Contrary to [Yuan et al. 2007], we are also not using any information from coarser scales in the deblurring at a considered scale. Since the reconstructions can contain significantly less detail we found that guiding fine scale deblurring with coarser scale information is problematic in many cases.

6. PSF ESTIMATION

The previous section assumes that the PSF of the optical system is given for each image tile. While the PSF can be obtained by any existing technique, we use a calibration-pattern-based approach. In contrast to methods that directly measure the PSF like Schuler et al. [2011], no pinhole-light source and a dark-room lab is necessary. Instead we just use a consumer laser-printer to make our targets. To estimate the PSFs from the target images, it is natural to apply the same optimization framework that was used for deblurring also for the PSF estimation step. This method is detailed below. Note again that any PSF estimation method could be used here and the whole estimation process is a calibration procedure, which only needs to be performed once for each lens.

6.1 PSF Estimation as Deconvolution

The PSF estimation problem can be posed as a deconvolution problem, where both a blurred image and a sharp image of the same scene are given. The blurred image is simply the scene imaged through the simple lens, with the aperture open, while the sharp image can be obtained by stopping the lens down to a small, almost pinhole aperture, where the lens aberrations no longer have an effect. By acquiring a sharp image this way (as opposed to a synthetic sharp image) we avoid both geometric and radiometric calibration issues in the sharp reference image.

Let J be an image patch in a considered blurred channel, I the corresponding sharp pinhole-aperture patch. We estimate a PSF

B_{opt} describing the blur in J by solving the minimization problem

$$\mathbf{b}_{\text{opt}} = \underset{\mathbf{b}}{\operatorname{argmin}} \ \|\mathbf{I}\mathbf{b} - s \cdot \mathbf{j}\|_2^2 + \lambda \|\nabla \mathbf{b}\|_1 + \mu \|\mathbf{1}^T \mathbf{b} - 1\|_2^2, \quad (18)$$

where the first term is a linear least-squares data fitting term, and the scalar $s = \sum_{k,l} I(k,l) / \sum_{k,l} J(k,l)$ accounts for the difference in exposure between the blurred and pinhole image. The second term represents a standard TV prior on the gradients of the recovered PSF, and the third term enforces an energy conservation constraint, i.e., $\sum_{k,l} B(k,l) = 1$.

Note that Eq. (18) is a convex optimization problem. We derive a first-order primal-dual algorithm adopting the framework described in Sec. 4. Specifically, Eq. (18) is expressed using the following operators adopting the notation from before:

$$\begin{aligned} \mathbf{K}(\mathbf{x}) &= \nabla \mathbf{x} \\ \mathbf{F}(\mathbf{y}) &= \|\mathbf{y}\|_1 \\ \mathbf{G}(\mathbf{x}) &= \frac{1}{\lambda} \|\mathbf{I}\mathbf{b} - s \cdot \mathbf{j}\|_2^2 + \frac{\mu}{\lambda} \|\mathbf{1}^T \mathbf{b} - 1\|_2^2 \end{aligned} \quad (19)$$

The following resolvent operators and convex conjugates necessary to apply Algorithm 1 are then provided in [Chambolle and Pock 2011]:

$$\mathbf{y} = \mathbf{prox}_{\sigma \mathbf{F}^*}(\tilde{\mathbf{y}}) \Leftrightarrow \mathbf{y}_i = \frac{\mathbf{y}_i}{\max(1, |\tilde{\mathbf{y}}_i|)} \quad (20a)$$

$$\mathbf{K}^* \mathbf{x} = \nabla^* \mathbf{x} = -\operatorname{div} \mathbf{x} \quad (20b)$$

However, the resolvent operator with respect to $\mathbf{G}(u)$ has to be derived for our problem. As we show in Appendix A, we can express it as a sequence of frequency domain operations:

$$\begin{aligned} \mathbf{u}_{\text{opt}} &= \mathbf{prox}_{\tau \mathbf{G}}(\tilde{\mathbf{u}}) \\ &= \mathcal{F}^{-1} \left(\frac{\tau s \mathcal{F}(\mathbf{I})^* \mathcal{F}(\mathbf{j}) + \frac{\lambda}{2} \mathcal{F}(\tilde{\mathbf{u}}) + \tau \mu \mathcal{F}(\mathbf{1})}{\tau |\mathcal{F}(\mathbf{I})|^2 + \frac{\lambda}{2} + \tau \mu \mathcal{F}(\mathbf{O})} \right), \end{aligned} \quad (21)$$

where $\mathbf{O}$ is a convolution matrix consisting only of ones.

Using these operators, we can again apply Algorithm 1. The computation of L has been described in Sec. 5.4 and the same τ and σ can be used.

6.2 Calibration Pattern

To allow for robust PSF estimation, the scene used for this purpose should have a broad spectrum. We therefore decide against both natural scenes and edge based patters (e.g., [Trimeche et al. 2005; Joshi et al. 2008]), and instead adopt a white noise pattern. See [Brauers et al. 2010] for a comparison of a white-noise pattern to different other calibration patterns. Our specific pattern, shown in Figure 9, is subdivided into several noise patches, separated by a white border. Each noise tile will produce a PSF for a single image tile, while the white frame helps suppress boundary effects in the PSF estimation process. This is because it enables a physically correct constant boundary condition around each patch if the frame width is $r_{\text{bnd}} \geq$ blur-radius (see again Fig. 9).

6.3 Full PSF Estimation Framework

To increase the efficiency of the PSF estimation, we apply the algorithm introduced above in scale space. We initialize our iterative minimization at each scale with the upsampled results from the next coarser scale, which yields a good starting point for the convex objective function from Eq. (18) and thus speeds up the minimization.

After the initial PSF estimation, we perform an additional step of smoothing by computing weighted averages of PSFs for a 3×3 set of neighboring tiles. Although the PSFs may contain high frequency features, these tend to change smoothly over the image plane. Combined with the relatively small tile size, we found that this spatial filtering does not cause feature loss, but reduces noise significantly.

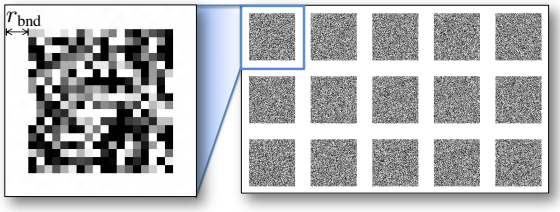

Fig. 9. Proposed PSF calibration target consisting of a grid of framed white noise patches. If the frame width is $r_{\text{bnd}} \geq$ blur-radius, a boundary condition propagating the boundary pixels outside (e.g., replicate the pixels) is a physically correct model for the blurred and sharp noise patch.

7. RESULTS

In this section we present a number of results captured with simple lens elements attached to a Canon EOS D40 camera, as well as detailed comparisons with state-of-the-art deconvolution methods. For figures where we choose to present cropped regions to highlight detail structure, the full native camera resolution image is provided in the supplemental material. We encourage the reader to zoom in on the individual images in the electronic version of this paper.

Efficiency.
First we compare the performance of our method to several other sparse optimization approaches. To this end, we implemented our objective function, including the cross-channel prior in the following solvers: Split-Bregman [Goldstein and Osher 2009], FISTA [Beck and Teboulle 2009], and Nesterov [Nesterov 2005], which are all recent, state-of-the-art sparse optimizers with guaranteed convergence to a global optimum. They therefore produce identical results to our method. To keep the comparison fair, we disabled the residual and multiscale parts of our algorithm. All experiments were performed on a single core of an Intel Core2 Quad CPU with 2.40GHz and 4GB RAM.

The test images are standard images used in imaging research, that were convolved with the PSFs shown in Figure 10. All algorithms were run until the per-iteration change in the result fell below a given threshold. Table I shows the timing results for the different algorithms and images. We note that our algorithm consistently outperforms the alternatives. For comparison we have added the last two rows in Table I which show the timings for the method from Levin et al. [2007] both in its original form, and a modified version with our cross-channel prior added. Our method is significantly faster than the version with the cross prior, but somewhat slower than Levin without our prior (as shown later in Figure 18, that method produces poorer results than our method). We note, however, that Levin's method only performs two iterations of IRLS, while we run our algorithm to convergence. Also, our use of multiscale deconvolution significantly speeds up the full algorithm.

Results from different lenses.
Figures 1 and 11 show several results from the most basic refractive optical system, a single plano-convex lens (focal length 130 mm, f/4.5) shown lying on the side in Figure 1. For all results in this paper, ISO100 and autoexposure settings were used, which we

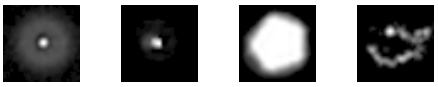

Fig. 10. Blur kernels used for Table I (all 45×45). From left to right: a) simple lens kernel estimated with our method, b) standard lens kernel estimated with our method, c) out-of-focus kernel from [Levin et al. 2007], d) motion blur from [Yuan et al. 2007].

Table I. Runtime comparison using different images and kernels. All times are in seconds.

Image:	Mandrill				Cameraman				Lena				Mean:
Kernel:	a)	b)	c)	d)	a)	b)	c)	d)	a)	b)	c)	d)	
Proposed method	14.5	11.7	22.8	18.2	19.8	12.9	23.6	19.0	20.7	10.2	25.5	21.6	18.4
Split Bregman	34.5	31.1	37.5	37.9	32.6	28.3	39.2	32.3	34.8	29.1	40.0	31.9	34.1
FISTA	320.7	304.8	385.4	341.6	443.9	413.5	422.5	404.4	425.5	409.9	407.5	376.3	388.0
Nesterov	344.8	320.5	378.7	361.9	462.4	440.5	453.2	431.7	451.7	440.8	424.3	405.5	409.7
Levin+cross term	53.5	27.6	72.3	71.6	40.6	26.4	69.4	71.9	40.2	24.3	76.4	74.8	54.1
Levin	11.1	11.4	11.3	11.6	11.8	11.8	11.4	11.0	11.3	11.8	11.2	11.4	11.4

consider as standard settings. The corresponding PSFs can be found on the right side of Figure 3. We note that all results show good detail down to the native sensor resolution, demonstrating that our method is indeed capable of producing high quality digital photographs with very simple lens designs. The parking meter image and the twigs show that deconvolving with the PSF for one particular scene depth preserves the deliberate defocus blur (bokeh) for objects at different depths without introducing artifacts. The manhole cover image at the top shows focus change over a tilted plane. Saturated regions as shown in the second inset of this image are handled properly by our method. The supplemental material contains even more plano-convex lens results than Fig. 11. It also includes a result under low-light conditions.

Figures 14 and 12 test the limits of our method using lenses with a larger aperture of $f/2$. Figure 14 shows the results for a 100 mm biconvex lens, with PSFs from Figure 3. The larger aperture increases the depth dependency of the PSF. Therefore our deconvolution method produces very sharp reconstructions in-plane (see color checker in Fig. 14), but suffers from some ringing out-out-of-plane (see text on the foreground object). A similar depth dependency is visible in Figure 12, where we show results for a 100 mm achromatic doublet (PSFs can be found in the supplemental material). Considering that the PSFs exhibit disk-like structures in all channels, our deconvolution method still produces detailed reconstructions. The depth-dependency of the PSFs (that is the ringing in the images) can be minimized through a stronger weight on the regularization terms, which, however, introduces cartoon-like artifacts at the highest resolution. Figure 13 shows a result of the same achromat stopped down to $f/4.5$. This significantly reduces the artifacts in results produced by our method. Note the strong chromatic artifacts in the naively deconvolved result, while our reconstruction contains significantly reduced ringing that is not chromatic. In the future one could imagine to improve the depth dependency by estimating depth-depending PSFs (see Section 8).

Results for a commercial camera lens, a Canon 28–105 mm zoom lens at 105 mm and f/4.5, are shown in Figures 15 and 16 (see supplemental material for full images). While this commercial lens shows much reduced aberrations compared to the uncorrected optics used above, there still are some residual blurs that can be removed with our method. In particular, the vertical edges of the calibration pattern in Figure 16 reveal a small amount of chromatic aberration that is removed by our method. The PSFs for this lens are around 11×11 pixels in diameter.

Fig. 11. Images captured with a simple plano-convex lens (left) and restored with our method (right). Note recovered spatial detail and the absence of color fringing. Also note the bokeh and graceful handling of out-of-focus blur in the top three images.

Fig. 12. Left: Cropped central region of the image from achromatic lens with $f/2.0$. Right: Deblurred version.

Fig. 13. Left: Captured input image from achromatic lens with $f/4.5$. Right: Deblurred version. The third inset on the right is the patch recovered with Levin et al. [2007] naively on independent channels.

Fig. 14. Left: Captured input image from biconvex lens at $f/2$. Right: Deblurred version. A very low cross-prior weight has been used here to illustrate the depth-dependency of the PSF.

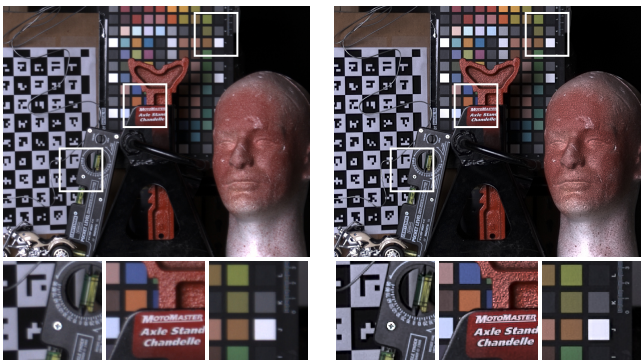

Fig. 15. Test scene captured with a commercial Canon lens (105 mm, f/4.5). Left: Captured input image. Right: Deblurred result.

Multispectral camera.
Our aberration correction method cannot only be used for regular RGB cameras but also for multispectral imagers. In this case, the cross-channel prior is applied for all pairs of frequency bands. Figure 17 shows results for a multispectral camera with custom multi-element optics. As with the conventional cameras, our method successfully removes chromatic aberration and restores lost image detail in the blurred channels. Considering the wavelength-dependent PSFs here, we want to point out that the assumption of fixed PSFs for each color-channel of an RGB-sensor is often violated. This assumption is made for all the RGB sensor results in this paper and is a classical assumption in deconvolution literature. However, one cannot tell from a tri-chromatic sensor the exact wavelength distribution. Metamers (different spectra that produce the same tri-

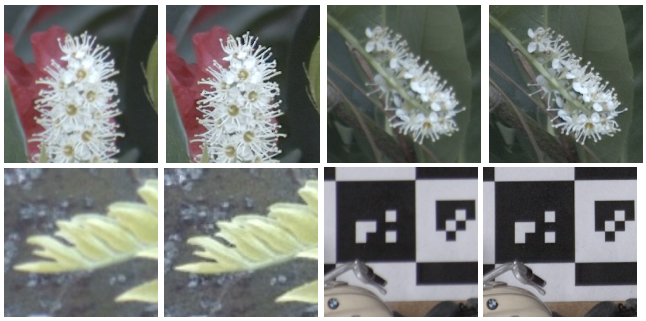

Fig. 16. Patches from an image captured with a commercial Canon lens (105 mm, f/4.5). Left: Captured input patches. Right: our results. The full image can be found in the supplemental material.

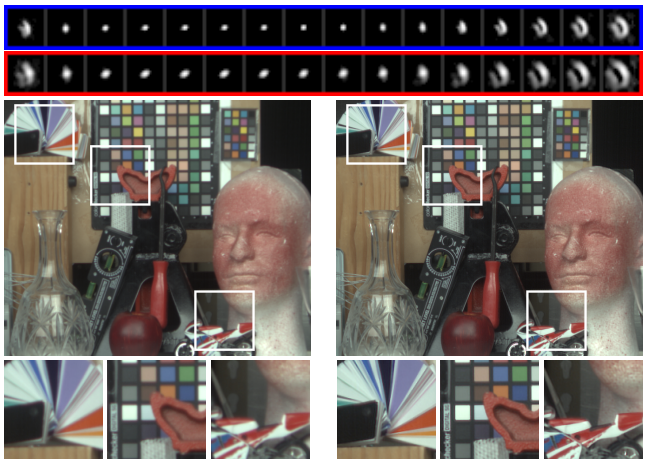

Fig. 17. Results for a multispectral camera with custom optics. The top row shows PSF estimation results for 20 nm wavelength increments, ranging from 420 to 720 nm (blue: image center, red: upper left corner). The bottom set of images and insets shows the captured image (left) and the deconvolved result (right) after mapping to sRGB space.

stimulus response) will have different blur kernels, so there can always be situations where the assumption of fixed per-channel PSFs will fail, such as for example light sources with multiple narrow spectra or albedos with very narrow color tuning. This will introduce errors in the data-fitting term of our objective function. Since we have a strong cross-channel and image prior (see comparisons to other priors below), we are still able to reconstruct images with high quality. For future research we imagine using a physical model of the full depth and wavelength-dependent PSF (see Section 8). Such a model may then allow to obtain a plausible spectral reconstruction of the image, similar to what Rump and Klein [2010] show.

We have compared our method against several state-of-the-art deconvolution algorithms, including the one from Levin et al. [2007]. Figure 13 shows an additional inset of per-channel deconvolution with their method. Compared to our result, it shows more prevalent ringing artifacts, which are particularly noticeable because they also have chromatic component. Our cross-channel prior successfully suppresses chromatic artifacts, but also helps to transfer information between channels for reduced ringing and increased detail overall. Figure 18 shows another comparison. Here we also implemented our cross-channel prior as an additional regularization term for Levin et al.'s method. While this improves the result, it does not match the quality of our method. The IRLS solver is

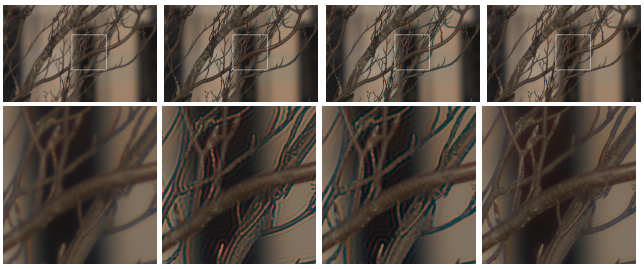

Fig. 18. Top row from left to right: Captured image, deconvolved result from [Levin et al. 2007] independently on each channel, IRLS method from [Levin et al. 2007] extended with our prior, our result. Bottom row: Patches in same order of the methods.

unstable, not guaranteed to converge, does not minimize the sparse norm close to its non-differentiable origin, and thus cannot achieve the results of our method. We especially found robustness to be a general issue once the cross-channel prior was added to IRLS, with the resulting method being very sensitive to parameter selection and often failing completely.

To further investigate the convergence properties of Hyper-Laplacian priors, we conducted more detailed experiments with the method by Krishnan and Fergus [2009] and compared it with a plain TV regularized solution (Appendix B) for which there are many solvers that guarantee convergence to a global optimum. We found that the TV solution often produces sparser gradient distributions (i.e., lower Hyper-Laplacian residual) than the solvers designed to optimize for these objectives. We can therefore conclude that Hyper-Laplacian regularization often fails at its objective to produce sparser gradient distributions, at least with the solvers available today.

We have also compared our method to Schuler et al. [2011]. First, in Fig. 19 we show results using their datasets (images and estimated PSFs). We show results with two different weights for the regularizer. Note that in *both* cases, our method recovers more fine-scale structure than their approach. Since Schuler et al. use a Hyper-Laplacian objective with the solver analyzed in Appendix B, we again achieve a more heavy-tailed gradient distribution in our result. However, the results for this dataset do not match the quality we can achieve with our own datasets – for varying regularization weight always a strong residual blur remains, especially around the leaves and window edges.

We believe these artifacts are due to issues with the estimated PSFs (which we take directly from Schuler et al.) rather than issues in the actual deconvolution method. We note that the PSFs lack a very low-frequency component, which we have observed with every single lens element we have tested in our setup. Instead, their blur kernels exhibit very thin structures akin to motion blur kernels and contain very little chromatic variation. A missing low-frequency component in the kernels could easily explain the poor quality of the deconvolved results. Unfortunately, Schuler et al. do not specify details about the lens, so we are unable to obtain similar images using our setup.

We note that Schuler et al.'s Yuv-cross-channel regularization *breaks down when the chromatic aberrations increase*. It completely fails for many of the large aberrations that we can handle. In Fig. 20 we compare our cross-channel prior against their Yuv-cross-channel regularization.

In this case we did not have access to their original implementation, so we reimplemented their method. To solve Eq.(5) from their paper, we have adapted the code from the Hyper-Laplacian solver [Krishnan and Fergus 2009] as suggested by the authors. All

Fig. 19. Top row from left to right: Captured image, our deconvolved result, result from [Schuler et al. 2011]. Patch matrix below from top left to bottom right: Captured image, result from [Schuler et al. 2011], our result with strong regularization (smoother reconstruction), our result with low regularization (more detailed reconstruction) .

parameters documented in their paper have been used. Note that our prior removes nearly all chromatic aberrations, while Schuler's Yuv regularization fails with the large aberrations that we can handle.

8. CONCLUSIONS AND FUTURE WORK

We have proposed a new method for high-quality imaging using simple lens elements rather than complex lens designs. This is made possible by a new, robust deconvolution with a novel cross-channel gradient prior that enforces sparsity of hue changes across the image. The prior allows detail transfer between different channels and suppresses reconstruction artifacts in individual channels. Both this deconvolution algorithm and a related PSF estimation method rely on a convex optimization framework [Chambolle and Pock 2011], which guarantees convergence to a globally optimal solution as well as optimal convergence rates. We demonstrate that this solver is also faster than alternative methods in practice. Further improvements in terms of both performance and quality are achieved by adding residual and multiscale deconvolution.

Overall our method produces image quality comparable to that of commercial point-and-shoot cameras for reasonable apertures (around f/4.5) even with the most basic refractive imaging optics imaginable: a single plano-convex lens. For much larger apertures of f/2 and more, the quality degrades due to the depth dependency of the PSF and significantly larger disk-shaped blur kernels, which destroy much of the frequency content. We point out that, in comparison to other deconvolution algorithms, our cross-channel prior still manages to transfer significantly more detail between channels to produce an acceptable result, although not at the quality level of a high-end lens in combination with a good SLR camera. We conclude that in order to achieve that level of image quality it may still be necessary to optimize the lens design to partially compensate for aberrations. However, we believe that even in this case, future lenses can be simpler than they are today, thus paving the road for lighter, cheaper, and more compact camera lenses.

At the moment, we are deconvolving the images with PSFs that are calibrated for a single scene depth and that are fixed per channel. One possible extension of our method is to perform deconvolution with PSFs for different image depths and wavelengths. How to handle depth-dependency has been demonstrated in, e.g., Levin et al. [2007], we have suggested a direction for wavelength-dependent deconvolution in Section 7. Doing so would require a large wavelength-dependent 3D table of PSFs, which may be difficult to calibrate and expensive to store. However, since we are

Fig. 20. Images from Fig. 11, 1 reconstructed using our cross-channel prior (left) and restored with the Yuv-regularization from [Schuler et al. 2011] (right).

working with very simple optics, an alternative similar to [Shih et al. 2012] may be to generate the PSFs procedurally from parameters of the optical system, such as lens curvature, glass type, as well as calibrated or measured focus and alignment parameters. We imagine to improve significantly on their results since our simple optical system leads to only a handful of optimization parameters.

REFERENCES

BECK, A. AND TEBOULLE, M. 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imag. Sci. 2*, 183–202.

BRAUERS, J., SEILER, C., AND AACH, T. 2010. Direct PSF estimation using a random noise target. In *Proceedings of SPIE, the International Society for Optical Engineering*. 75370B.

CHAMBOLLE, A. AND POCK, T. 2011. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis. 40*, 120–145.

CHAN, S., KHOSHABEH, R., GIBSON, K., GILL, P., AND NGUYEN, T. 2011. An augmented lagrangian method for total variation video restoration. *IEEE Trans. Image Proc. 20,* 11, 3097–3111.

CHO, T. S., JOSHI, N., ZITNICK, C. L., KANG, S. B., SZELISKI, R., AND FREEMAN, W. T. 2010. A content-aware image prior. In *Computer Vision and Pattern Recognition, IEEE Conference on*. 169–176.

CHO, T. S., ZITNICK, C. L., JOSHI, N., KANG, S. B., SZELISKI, R., AND FREEMAN, W. T. 2012. Image restoration by matching gradient distributions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 34,* 4, 683–694.

CHUNG, S.-W., KIM, B.-K., AND SONG, W.-J. 2009. Detecting and eliminating chromatic aberration in digital images. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*. 3905–3908.

COMBETTES, P. L. AND PESQUET, J.-C. 2011. Proximal splitting methods in signal processing. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, 185–212.

COSSAIRT, O. AND NAYAR, S. 2010. Spectral focal sweep: Extended depth of field from chromatic aberrations sweep. In *Computational Photography (ICCP), 2010 IEEE International Conference on*. 1–8.

FRIEDLANDER, M. P. 2011. Sparse optimization minitutorial. `http://www.cs.ubc.ca/~mpf/papers/mpf-tutorial-siopt11.pdf`.

GAUSS, C. 1841. *Dioptrische Untersuchungen*. Dieterich.

GIBSON, J. D. AND BOVIK, A., Eds. 2000. *Handbook of Image and Video Processing*. Academic Press, Inc.

GOLDLUECKE, B. AND CREMERS, D. 2011. Introducing total curvature for image processing. In *Proc. ICCV*. 1267–1274.

GOLDSTEIN, T. AND OSHER, S. 2009. The Split Bregman method for L1-regularized problems. *SIAM J. Img. Sci. 2*, 323–343.

GUICHARD, F., NGUYEN, H. P., TESSIRES, R., PYANET, M., TARCHOUNA, I., AND CAO, F. 2009. Extended depth-of-field using sharpness transport across color channels. In *Proceedings of SPIE, the International Society for Optical Engineering*. 72500N.

JOSHI, N., LAWRENCE, C., ZITNICK, R., SZELISKI, D., AND KRIEGMAN, J. 2009. Image deblurring and denoising using color priors. In *Computer Vision and Pattern Recognition, IEEE Conference on*. 1550–1557.

JOSHI, N., SZELISKI, R., AND KRIEGMAN, D. J. 2008. PSF estimation using sharp edge prediction. In *Computer Vision and Pattern Recognition, IEEE Conference on*. 1–8.

KANG, S. B. 2007. Automatic removal of chromatic aberration from a single image. In *Computer Vision and Pattern Recognition, IEEE Conference on*. 1–8.

KEE, E., PARIS, S., CHEN, S., AND WANG, J. 2011. Modeling and removing spatially-varying optical blur. In *Proc. IEEE ICCP*. 1–8.

KRISHNAN, D. AND FERGUS, R. 2009. Fast image deconvolution using Hyper-Laplacian priors. *Advances in Neural Information Processing Systems 22*, 1–9.

LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. T. 2007. Image and depth from a conventional camera with a coded aperture. In *ACM Trans. Graph. (Proc. SIGGRAPH)*.

LUCY, L. 1974. An iterative technique for the rectification of observed distributions. *The Astronomical Journal 79*, 745+.

MAHAJAN, V. N. 1991. *Aberration Theory Made Simple*. SPIE.

NESTEROV, Y. 2005. Smooth minimization of non-smooth functions. *Math. Program. 103*, 127–152.

RASHED, R. 1990. A pioneer in anaclastics: Ibn sahl on burning mirrors and lenses. *Isis 81,* 3, 464–491.

RICHARDSON, W. H. 1972. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am. 62*, 55–59.

ROWLAND, T. 2012. Operator norm. From MathWorld - A Wolfram Web Resource, created by Eric W. Weisstein. `http://mathworld.wolfram.com/OperatorNorm.html`.

RUMP, M. AND KLEIN, R. 2010. Spectralization: Reconstructing spectra from sparse data. In *SR '10 Rendering Techniques*, J. Lawrence and M. Stamminger, Eds. Eurographics Association, Saarbruecken, Germany, 1347–1354.

SCHULER, C. J., HIRSCH, M., HARMELING, S., AND SCHÖLKOPF, B. 2011. Non-stationary correction of optical aberrations. In *Computer Vision (ICCV), IEEE International Conference on*. IEEE, 659–666.

SEIDEL, L. 1857. *Über die Theorie der Fehler, mit welchen die durch optische Instrumente gesehenen Bilder behaftet sind, und über die mathematischen Bedingungen ihrer Aufhebung*. Abhandlungen der Naturwissenschaftlich-Technischen Commission bei der Königl. Bayerischen Akademie der Wissenschaften in München. Cotta.

SHIH, Y., GUENTER, B., AND JOSHI, N. 2012. Image enhancement using calibrated lens simulations. In *Computer Vision ECCV 2012*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Lecture Notes in Computer Science, vol. 7575. Springer Berlin Heidelberg, 42–56.

TRIMECHE, M., PALIY, D., VEHVILÄINEN, M., AND KATKOVNIK, V. 2005. Multichannel image deblurring of raw color components. In *Proc. Computational Imaging III*. 169–178.

WIENER, N. 1964. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. The MIT Press.

YUAN, L., SUN, J., QUAN, L., AND SHUM, H.-Y. 2007. Image deblurring with blurred/noisy image pairs. *ACM TOG (Proc. SIGGRAPH) 26,* 3, 1.

YUAN, L., SUN, J., QUAN, L., AND SHUM, H.-Y. 2008. Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM Trans. Graph. 27,* 3 (Aug.), 74:1–74:10.

APPENDIX

A. PROOFS

This section contains all the proofs for the mathematical results presented in the paper.

PROOF FOR (10).

$$
\begin{aligned}
&(\mathbf{i}_c)_{\text{opt}} \\
&= \underset{\mathbf{x}}{\operatorname{argmin}}\ \|\mathbf{B}_c\mathbf{x}-\mathbf{j}_c\|_2^2 + \lambda_c \sum_{a=1}^{5} \|\mathbf{H}_a\mathbf{x}\|_1 \\
&\qquad + \sum_{l\neq c} \beta_{cl} \sum_{a=1}^{2} \|\mathbf{H}_a\mathbf{x}\cdot\mathbf{i}_l - \mathbf{H}_a\mathbf{i}_l\cdot\mathbf{x}\|_1 \\
&= \underset{\mathbf{x}}{\operatorname{argmin}}\ \lambda_c \sum_{a=1}^{5} \|\mathbf{H}_a\mathbf{x}\|_1 + \sum_{l\neq c} \beta_{cl} \sum_{a=1}^{2} \left\|\left(\mathbf{D}_{\mathbf{i}_l}\mathbf{H}_a - \mathbf{D}_{\mathbf{H}_a\mathbf{i}_l}\right)\mathbf{x}\right\|_1 \\
&\qquad + \|\mathbf{B}_c\mathbf{x}-\mathbf{j}_c\|_2^2 \\
&= \left\| \begin{bmatrix} \lambda_c\mathbf{H}_1 \\ \vdots \\ \lambda_c\mathbf{H}_5 \\ \beta_{cl}\left(\mathbf{D}_{\mathbf{i}_l}\mathbf{H}_1 - \mathbf{D}_{\mathbf{H}_1\mathbf{i}_l}\right) \\ \vdots \\ \beta_{cl}\left(\mathbf{D}_{\mathbf{i}_l}\mathbf{H}_2 - \mathbf{D}_{\mathbf{H}_2\mathbf{i}_l}\right) \\ \vdots \end{bmatrix} \mathbf{x} \right\|_1 + \|\mathbf{B}_c\mathbf{x}-\mathbf{j}_c\|_2^2 \\
&= \|\mathbf{S}\mathbf{x}\|_1 + \|\mathbf{B}_c\mathbf{x}-\mathbf{j}_c\|_2^2
\end{aligned}
$$

□

PROOF FOR (17).

$$
\begin{aligned}
&\mathbf{u}_{\text{opt}} = \mathbf{prox}_{\tau\mathbf{G}}(\tilde{\mathbf{u}}) \\
&\quad = \underset{\mathbf{u}}{\operatorname{argmin}}\ \frac{1}{2\tau}\|\mathbf{u}-\tilde{\mathbf{u}}\|_2^2 \\
&\qquad + \|\mathbf{B}_c\mathbf{u}-\mathbf{j}_c\|_2^2 \\
&\qquad \underbrace{+\lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \|\mathbf{D}_{\mathbf{w}}\left(\mathbf{H}_a\mathbf{u}-\mathbf{H}_a\mathbf{i}_l\right)\|_2^2}_{\Phi(u)} \\
&\Leftrightarrow 0 = \frac{\partial\Phi(\mathbf{u}_{\text{opt}})}{\partial\mathbf{u}} = \frac{1}{\tau}(\mathbf{u}_{\text{opt}}-\tilde{\mathbf{u}}) + 2\left(\mathbf{B}_c^T\mathbf{B}_c\mathbf{u}_{\text{opt}} - \mathbf{B}_c^T\mathbf{j}\right) \\
&+ 2\lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \left(\left(\mathbf{D}_{\mathbf{w}}\mathbf{H}_a\right)^T\mathbf{D}_{\mathbf{w}}\mathbf{H}_a\mathbf{u}_{\text{opt}} - \left(\mathbf{D}_{\mathbf{w}}\mathbf{H}_a\right)^T\mathbf{D}_{\mathbf{w}}\mathbf{H}_a\mathbf{i}_l\right) \\
&\Leftrightarrow \frac{\partial\Phi(\mathbf{u}_{\text{opt}})}{\partial\mathbf{u}} = \frac{1}{\tau}(\mathbf{u}_{\text{opt}}-\tilde{\mathbf{u}}) + 2\left(\mathbf{B}_c^T\mathbf{B}_c\mathbf{u}_{\text{opt}} - \mathbf{B}_c^T\mathbf{j}\right) \\
&+ 2\lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \left(\mathbf{H}_a^T\mathbf{D}_{\mathbf{w}}^2\mathbf{H}_a\mathbf{u}_{\text{opt}} - \mathbf{H}_a^T\mathbf{D}_{\mathbf{w}}^2\mathbf{H}_a\mathbf{i}_l\right) = 0 \\
&\Leftrightarrow \left[2\tau\mathbf{B}_c^T\mathbf{B}_c + \mathbb{I} + 2\tau\lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \mathbf{H}_a^T\mathbf{D}_{\mathbf{w}}^2\mathbf{H}_a\right]\mathbf{u}_{\text{opt}} \\
&= 2\tau\mathbf{B}_c^T\mathbf{j}_c + \tilde{\mathbf{u}} + 2\tau\lambda_b \sum_{l\neq c}\sum_{a=1}^{2} \mathbf{H}_a^T\mathbf{D}_{\mathbf{w}}^2\mathbf{H}_a\mathbf{i}_l \\
&\Leftrightarrow \mathbf{A}\mathbf{u}_{\text{opt}} = \mathbf{b}
\end{aligned}
$$

□

PROOF FOR (21).

$$
\begin{aligned}
&\mathbf{u}_{\text{opt}} = \mathbf{prox}_{\tau\mathbf{G}}(\tilde{\mathbf{u}}) = \\
&\underset{\mathbf{u}}{\operatorname{argmin}}\ \underbrace{\frac{1}{2\tau}\|\mathbf{u}-\tilde{\mathbf{u}}\|_2^2 + \frac{1}{\lambda}\|\mathbf{I}\mathbf{u}-s\cdot\mathbf{j}\|_2^2 + \frac{\mu}{\lambda}\left\|\mathbf{1}^T\mathbf{u}-1\right\|_2^2}_{\Phi(u)} \\
&\Leftrightarrow \frac{\partial\Phi(\mathbf{u}_{\text{opt}})}{\partial\mathbf{u}} = \frac{2}{\lambda}\left(\mathbf{I}^T\mathbf{I}\mathbf{u}_{\text{opt}} - s\cdot\mathbf{j}\right) \\
&\qquad + \frac{1}{\tau}(\mathbf{u}_{\text{opt}}-\tilde{\mathbf{u}}) + \frac{2\mu}{\lambda}\left(\mathbf{1}^T\mathbf{u}_{\text{opt}}-1\right)\mathbf{1} = 0 \\
&\Leftrightarrow \frac{\lambda}{2}\mathbf{u}_{\text{opt}} - \frac{\lambda}{2}\tilde{\mathbf{u}} + \tau\left(\mathbf{I}^T\mathbf{I}\mathbf{u}_{\text{opt}} - s\cdot\mathbf{I}^T\mathbf{j}\right) + \tau\mu(\mathbf{O}\mathbf{u}_{\text{opt}}-\mathbf{1}) = 0 \\
&\Leftrightarrow \mathbf{u}_{\text{opt}} = \mathcal{F}^{-1}\left(\frac{\tau s\mathcal{F}(\mathbf{I})^*\mathcal{F}(\mathbf{j}) + \frac{\lambda}{2}\mathcal{F}(\tilde{\mathbf{u}}) + \tau\mu\mathcal{F}(\mathbf{1})}{\tau\left|\mathcal{F}(\mathbf{I})\right|^2 + \frac{\lambda}{2} + \tau\mu\mathcal{F}(\mathbf{O})}\right)
\end{aligned}
$$

□

B. COMPARISON OF HYPER-LAPLACIAN AND TV REGULARIZATION

Qualitative Comparison. Often IRLS-based solvers or solvers targeting non-convex objective functions fail to converge to a global minimum. Figure 21 shows an example for the non-convex function solved by Krishnan and Fergus [2009] using the kernels introduced in the same paper. A single channel grayscale deconvolution problem is solved using the unaltered code published from the authors' website. We compare these results against a single-channel

Table II. Comparison of the $\ell_{2/3}$-regularized objective $f(\mathbf{x}) = \frac{\lambda_d}{2}\|\mathbf{Bx} - \mathbf{j}\|_2^2 + \|\nabla\mathbf{x}\|_{2/3}$ and its regularizer component $r(\mathbf{x}) = \|\nabla\mathbf{x}\|_{2/3}$ using our solver with plain TV-regularization and the method from [Krishnan and Fergus 2009] for different images and kernels. All values in the table are in 10^3. Although we are actually minimizing the ℓ_1-regularized approximation of $f(\mathbf{x})$, our method finds in most cases a better objective with a more sparse solution than what the method from [Krishnan and Fergus 2009] produces.

Image:	Mandrill				Cameraman				Lake				Lena				House				
Kernel:	a)	b)	c)	d)	a)	b)	c)	d)	a)	b)	c)	d)	a)	b)	c)	d)	a)	b)	c)	d)	Mean:
$f(\mathbf{x})$ Hyper-Laplacian	230.86	155.80	254.67	324.29	313.79	173.04	330.45	394.65	271.05	150.83	296.63	**383.01**	248.62	94.68	253.23	351.58	**453.08**	**349.97**	488.47	525.24	302.19
$f(\mathbf{x})$ our TV-solver	**229.47**	**149.77**	**253.35**	**319.66**	**308.34**	**170.19**	**324.16**	**388.76**	**266.24**	**146.14**	**292.35**	384.79	**245.24**	**91.17**	**247.90**	**347.58**	445.21	350.53	**482.01**	**516.30**	**297.95**
$r(\mathbf{x})$ Hyper-Laplacian	**15.70**	31.24	13.75	22.31	12.86	17.37	12.28	15.67	18.05	24.39	17.96	22.41	15.87	20.02	15.84	18.52	14.17	16.68	14.95	17.02	17.85
$r(\mathbf{x})$ our TV-solver	16.72	**27.84**	**10.72**	**21.49**	**10.52**	**14.73**	**8.21**	**13.11**	**15.64**	**20.65**	**12.87**	**19.37**	**13.40**	**17.18**	**11.65**	**16.39**	**11.10**	**14.22**	**9.92**	**14.73**	**15.02**

TV-regularizer with the same kernels and without scale-space and residual deconvolution to make the comparison fair. The same regularization weights have been used, which means that the global optimum of the Hyper-Laplacian objective function should have a gradient distribution with a heavier tail than the one of the TV solution.

However, as seen in the figure the distribution obtained using Krishnan and Fergus' method [2009] exhibits a *lighter* tail than the solution from an ℓ_1-solver with guaranteed convergence, suggesting that the the global optimum for the Hyper-Laplacian objective has not been found.

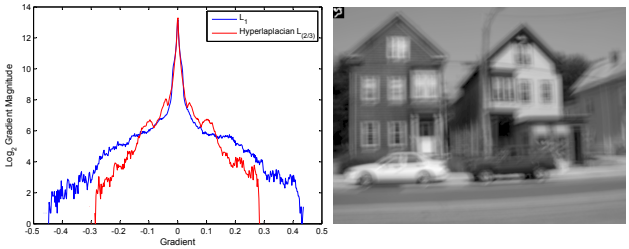

Fig. 21. Gradient distributions of the right synthetic blurred single-channel image deconvolved with the according PSF in its upper left. Gaussian noise $\sigma = 10^{-4}$ and quantization was applied. Note, that our ℓ_1-regularized solution features a heavier tail than the approximate solution for the $\ell_{\frac{2}{3}}$-regularized problem computed with the method from [Krishnan and Fergus 2009].

Quantitative Comparision. This suspicion can be further corroborated by systematic experiments with larger sets of images. Using the same images as for Table I, we compute both an $\ell_{2/3}$ regularized solution (again using the method by Krishnan and Fergus), as well as a TV solution, and then compute the numerical value of the $\ell_{2/3}$ objective function for both results (see Table II). We note that in most cases, the TV solver actually produces a more optimal solution to the $\ell_{2/3}$ problem than the dedicated Hyper-Laplacian solver.

These results also explain why Krishnan and Fergus [2009] in their work found only very minor differences in SNR values between the two methods. Our own experiments confirm these findings.

We conclude, that, although Hyper-Laplacian norms should in theory provide better regularization than simple TV terms, we in practice do not see these improvements due to the inability of current solvers to find the global optimum. These findings support our choice of ℓ_1 norms for the various regularization terms in our method, and stress the significance of convex methods with guaranteed convergence to a global optimum.