Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint Nr. 261

Robust Variational Approaches to Positivity-Constrained Image Deconvolution

Martin Welk

Saarbrücken 2010

Fachrichtung 6.1 – Mathematik
Universität des Saarlandes

Preprint No. 261
submitted: March 3, 2010

Robust Variational Approaches to Positivity-Constrained Image Deconvolution

Martin Welk

Saarland University
Mathematical Image Analysis Group
P.O. Box 15 11 50
66041 Saarbrücken
Germany
welk@mia.uni-saarland.de

Edited by
FR 6.1 – Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-Mail: preprint@math.uni-sb.de
WWW: http://www.math.uni-sb.de/

Robust Variational Approaches to Positivity-Constrained Image Deconvolution

Martin Welk

Abstract

Two approaches to the combination of robust variational deconvolution with positivity constraints are considered. The first approach modifies a standard robust variational deconvolution method by carrying out a gradient descent with respect to a multiplicative perturbation, which can also be considered as gradient descent in a hyperbolic metric. The second approach is based on the representation of the well-known Richardson-Lucy iterative deconvolution as fixed-point iteration for the minimisation of an information divergence functional, again under a multiplicative perturbation model. The asymmetric penaliser function contained in this functional is then varied into a robust penaliser, and complemented with a regulariser. The resulting functional gives rise to a fixed point iteration that we call robust and regularised Richardson-Lucy deconvolution. It achieves an image restoration quality comparable to state-of-the-art variational deconvolution with a computational efficiency similar to that of the original Richardson-Lucy method. Experiments on synthetic and real-world image data demonstrate the performance of the proposed methods.

1 Introduction

The sharpening of blurred images is a standard problem in many imaging applications. A large variety of different approaches to this severely ill-posed inverse problem have been developped over time which differ in the assumptions they make, and in their suitability for different application contexts.

Blur of an image is described by a *point-spread function (PSF)* which describes the redistribution of light energy in the image domain $\Omega \subset \mathbb{R}^2$. When blurring acts equally at all locations, one has a *space-invariant PSF* $h : \mathbb{R}^2 \to \mathbb{R}^+$ which acts by convolution. Accounting also for the impact of noise n, a typical blur model then reads

$$f = h * g + n \,, \tag{1}$$

where f is the observed image, and g the unknown sharp image.

In the more general case of a space-variant blur one needs a point-spread function with two arguments, $H : \Omega \times \Omega \to \mathbb{R}^+$, and $h * g$ is replaced with the integral operator

$$(H \circledast g)(\boldsymbol{x}) := \int_\Omega H(\boldsymbol{x}, \boldsymbol{y}) g(\boldsymbol{y}) \,\mathrm{d}\boldsymbol{y} \tag{2}$$

such that

$$f = H \circledast g + n \,. \tag{3}$$

Note that this includes the space-invariant case by setting $H(\boldsymbol{x}, \boldsymbol{y}) = h(\boldsymbol{x} - \boldsymbol{y})$.

Conservation of energy implies that $H \circledast 1 = 1$, however, this condition may be violated near image boundaries due to blurring across the boundary.

In deblurring, we want to obtain a restored image u that approximates g, from the degraded image f and the PSF H. This is the case of *non-blind* deconvolution (as opposed to blind deconvolution which would aim at inferring the sharp image and the PSF simultaneously from the degraded image).

Let us recall shortly some approaches to this problem. Linear filters offer efficient deblurring when noise is moderate and a sufficiently accurate blur model is available. The Wiener filter [33] is a prominent representative; it is optimal if the noise is Gaussian. As it makes use of Fourier representations, it is often used in conjunction with fast Fourier transform. Other linear methods are formulated using a matrix-algebraic framework [20]; they can also be adapted to spatially variant blurs [19]. The strength of linear filters lies in their efficiency and their equal suitability across a wide range of deconvolution problems. A broad variety of fast numerical algorithms has been developped for them.

Another class of approaches is based on statistical models [13, 35]. Their advantages include easy and precise adaptation to noise models, and the ability to make precise assertions about conservation of expectation values.

Richardson-Lucy (RL) deconvolution [17, 24] is an early representative of a nonlinear deconvolution procedure whose motivation comes from statistical considerations. It still enjoys popularity due to its simplicity and efficiency. Assuming that the degraded and sharp images, and the point-spread function are smooth functions over $\mathbb{R}^2$ with positive real values, one uses the iteration

$$u^{k+1} = \left(H^* \circledast \frac{f}{H \circledast u^k} \right) \cdot u^k \tag{4}$$

to generate a sequence of successively sharpened images $u^1, u^2, \ldots$ from the initial image $u^0 := f$. Here, H^* denotes the adjoint of the point-spread function H given by $H^*(\boldsymbol{x}, \boldsymbol{y}) = H(\boldsymbol{y}, \boldsymbol{x})$.

In the absence of noise the sharp image g is a fixed point of (4), as in this case the multiplier $H^* \circledast \frac{f}{H \circledast g}$ equals the constant function 1.

The single parameter of the procedure is the number of iterations. While with increasing number of iterations greater sharpness is achieved, the degree of regularisation is reduced at the same time, which in the long run leads to amplification of artifacts that after a while dominate the filtered image.

Variational methods [22, 34] tackle the deconvolution task by minimising a functional that consists of two parts: a *data term* that enforces the match between the sought image and the observed image via the blur model, and a *smoothness*

term or *regulariser* that brings in regularity assumptions about the unknown sharp image. A general model for variational deconvolution of a grey-value image u with known point-spread function H is based on minimising the energy functional

$$E[u] = \int_\Omega \left(\Phi\big((f - H \circledast u)^2\big) + \alpha\Psi\big(|\nabla u|^2\big)\right) \mathrm{d}\boldsymbol{x} \tag{5}$$

in which the data term $\Phi\big((f-H\circledast u)^2\big)$ penalises the reconstruction error or *residual* $f - H \circledast u$ to suppress deviations from the blur model, while the smoothness term $\Psi\big(|\nabla u|^2\big)$ penalises roughness of the reconstructed image. The *regularisation weight* α balances the influences of both terms. $\Phi, \Psi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ are increasing penalty functions. A popular representative is total variation deconvolution [8, 22, 28] in which Φ is the identity, and $\Psi(s^2) = |s|$.

Recent advances include the introduction of *robust* data terms with less-than-quadratic growth of Φ, e.g., the L^1 penaliser $\Phi(s) = |s|$, see [4], studies on various edge-preserving and edge-enhancing regularisers [3, 30], and spatially variant blur models [5, 31]. Variational deconvolution with nonlocal regularisation was studied in [15, 16].

In [29] an approach to perform variational deconvolution under inequality constraints was presented. This includes positivity of grey-values or an interval constraint for grey-value image deconvolution. The essential idea of this approach was to retain the same energy functional as for an unconstrained deconvolution but reparametrise the image in a way that enforces the desired constraint, e.g. $u = \exp(z)$ with a new image z to enforce positivity of u. Taking a differential geometric viewpoint, the resulting method could then be interpreted as gradient descent for u but with a new underlying metric in the space of image functions. In this form, the idea could immediately be adapted to design a variational deconvolution method for positive definite symmetric matrix-valued images that combines a linear blur model with guaranteed preservation of positive definiteness.

The strength of variational approaches lies in their great flexibility, and in the explicit way of expressing the assumptions made. They achieve often an excellent reconstruction quality, but their computational cost tends to be rather high.

Considering the relation between the before-mentioned approaches, it has been noted [26] that Richardson-Lucy deconvolution can be interpreted as a fixed point iteration for an optimisation problem. Here, Csiszár's *information divergence* [10] acts as an asymmetric penaliser function. By formulating the optimisation problem on a continuous spatial scale, Richardson-Lucy deconvolution is embedded into the context of variational deconvolution. This view permits to modify Richardson-Lucy deconvolution in the same flexible manner as standard variational approaches by introducing robust data terms and edge-preserving, or even edge-enhancing regularisers. While an edge-preserving regulariser in combination with Richardson-Lucy deconvolution has been used by Dey et al. [11, 12]

(in a space-invariant setting), the possibility of a robust data term has not been studied so far.

Our contribution. This paper focusses on deconvolution methods for grey-value and colour images subject to the positivity constraint that are derived or motivated from variational ideas. To this end, we first revisit the variational approach to inequality-constrained deconvolution from [29]. For the case of positivity-constrained deconvolution, we will rewrite this approach slightly using a multiplicative gradient descent.

Secondly, we modify the variational Richardson-Lucy approach by introducing robust data terms. To do this, we change the asymmetric penaliser function in the data term in such a way that larger residual errors are penalised less than with the standard Csiszár divergence term.

Using robust data terms together with a regulariser similar to [11, 12] we obtain a robust and regularised Richardson-Lucy variant that unites the high restoration quality of variational deconvolution methods with high efficiency that is not far from the original Richardson-Lucy iteration. This method has already been used for experiments on recovering information from diffuse reflections in [1]. We demonstrate that both robust data terms and regularisers contribute substantially to its performance.

Related work. The omnipresence of deblurring problems has made computer vision researchers address this problem since the beginnings of the discipline [33].

From the abundant literature on this topic, the most relevant work in our present context includes Richardson-Lucy deconvolution [17, 24], variational methods [8, 22, 25, 34], and their interplay [11, 12, 26], see also [7] for another approach to combine Richardson-Lucy deconvolution with regularisation.

Fundamental theoretical results on existence and uniqueness of solutions of deconvolution problems can be found in the work of Bertero et al. [6].

Robust data terms in deconvolution go back to Zervakis et al. [35] in statistical models, and have recently been used intensively in the variational context by Bar et al. [4, 5] and Welk et al. [29, 31]. Positivity constraints were studied in discrete iterative deconvolution by Nagy and Strakoš [21] and in a variational framework by Welk and Nagy [29]. The extension of variational approaches to multi-channel images has been studied in [14] and more specifically in deconvolution in [2, 29].

Structure of the paper. In Section 2 we revisit the constrained gradient descent from [29], focussing on the positivity case and reformulating it in terms of multiplicative gradient descent for this case. Section 3 is devoted to put Richardson-Lucy deconvolution into a variational context. Exploiting this connection, the RL algorithm can be modified in order to increase restoration quality and robustness with respect to noise and perturbations. An experimental comparison of the deconvolution techniques under consideration is provided in Section 4 based on both synthetic and real-world data. Conclusions in Section 5 end the paper.

2 Positivity-Constrained Gradient Descent, Revisited

In this section, we review the variational gradient descent approach to deconvolution under inequality constraints that has been introduced in [29], studying it in more detail and reformulating it in part.

2.1 Penalisers

We start with a short discussion of the penalisation functions in the energy functional (5). For the data term penaliser Φ, the most common choice is the least-squares penaliser $\Phi(s^2) = s^2$. With Φ growing slower than $\Phi(s^2) = s^2$, one has a *robust data term* [4] which reduces the influence of large residuals (outliers) on the energy compared to a least-squares term. The standard choice here is the regularised L^1-norm $\Phi(s^2) = \sqrt{s^2 + \beta^2}$ with some small $\beta > 0$. Robust data terms significantly improve the performance of variational deconvolution in the presence of noise [4] or data which fulfil model assumptions imperfectly, such as imprecise PSF estimates [31].

In the smoothness term, one could in principle use a simple Whittaker–Tikhonov regularisation [27, 32] given by $\Psi(s^2) = s^2$. Since this leads to unfavourable blurring effects, image processing literature mostly prefers non-quadratic terms like (regularised) total variation $\Psi(s^2) = \sqrt{s^2 + \varepsilon^2}$ (with small $\varepsilon > 0$) [4, 8, 18, 22, 25] or even non-convex ones like the Perona–Malik function $\Psi(s^2) = \lambda^2 \ln(1+s^2/\lambda^2)$ [23, 30]. The reason is the edge-preserving or even edge-enhancing behaviour of these regularisers.

2.2 Minimisation

One approach to compute a minimiser of (5) is based on gradient descent. To this end, one computes by the usual Euler-Lagrange formalism $\delta_v E := \frac{\mathrm{d}}{\mathrm{d}\varepsilon} E[u + \varepsilon v]\big|_{\varepsilon=0}$ for a small additive perturbation v. Requiring this to equal $\left\langle \frac{\delta E}{\delta u}, v \right\rangle$ for all test functions v, one derives the variational gradient (Gateaux derivative) $\delta E / \delta u$ and from this the integrodifferential equation

$$\begin{aligned} \partial_t u = -\frac{\delta E}{\delta u} = H^* \circledast \left(\Phi'\left((f - H \circledast u)^2\right) \cdot (f - H \circledast u)\right) \\ + \alpha \operatorname{div}\left(\Psi'\left(|\nabla u|^2\right) \nabla u\right) \end{aligned} \tag{6}$$

as the gradient descent for E. Complementing this equation with e.g. the blurred image f as initial condition and suitable boundary conditions, one has an initial-boundary value problem which is then approximated numerically until a steady state is reached. This can be done in the simplest case via an explicit Euler

scheme. More sophisticated approaches like semi-implicit schemes with conjugate gradients for (6) are possible; alternatively one can directly work with the elliptic problem. In terms of computational expense, the difference between these approaches is surprisingly small: The number of iterations up to convergence, be it explicit time steps or CG iterations, depends primarily on the extent and structure of the PSF. As each iteration contains two $\circledast$ operations, the minimisation of (5) remains expensive by any strategy.

With respect to gradient descent, one might also object to its use in minimising (5) in the case of non-convex regularisers because of the existence of multiple minima. One has to be aware, however, that non-uniqueness of minimisers is inherent to the deconvolution problem itself and can even occur with convex regularisers, depending on the PSF. Considering, for instance, a space-invariant PSF h whose Fourier transform possesses zeros, one sees that the corresponding frequencies vanish completely under convolution with h. As a consequence, the functional cannot distinguish between functions u and $\tilde{u}$ if $\tilde{u} - u$ is composed of these frequencies. In reality, the same holds true even for all frequencies whose coefficients in the Fourier transform of h are below a positive threshold that depends on the noise level. This insensitivity forms the root of the typical ringing artifacts in various deconvolution methods. Non-uniqueness of minimisers can therefore hardly be avoided, and is an issue for any other minimisation method, too. Practically, one is interested in a sufficiently good local minimum. Continuation strategies for the regularisation weight [30] are sometimes useful to steer the process into a favourable local minimum.

2.3 Constraints

From modelling considerations one can often infer additional information that helps to mitigate the ill-posedness of the deconvolution problem. An important type of such information are inequality constraints. For instance, in a scalar-valued image with grey-values proportional to radiance, negative values would correspond to unphysical negative radiance and can therefore be ruled out. Sometimes also an upper bound can be derived from image acquisition parameters.

To introduce an inequality constraint, like the positivity requirement for grey-values, into the variational framework, an obvious idea would be to add a penalty for negative values. In a straightforward realisation, however, this does not strictly enforce the desired inequality. Although more involved optimisation procedures like Bregman iterations might be considered to cure this, we follow a different approach here which has been introduced for variational deconvolution in [29], and which has been known before in the context of discrete deconvolution [21]. To this end, the grey-values are reparametrised via $u = \exp(z)$ with a new image function z whose values are unconstrained in $\mathbb{R}$.

As pointed out in [29], this idea can be put slightly more general using $u = \varphi(z)$ with a smooth invertible function $\varphi : \mathbb{R} \to \mathbb{R}$. By substituting this into (5),

the energy is expressed as a functional depending on z as

$$\tilde{E}[z] = \int_\Omega \left(\Phi\Big(\big(f - H \circledast \varphi(z)\big)^2 \Big) + \alpha \Psi\Big(\big(\varphi'(z)\,|\nabla z|\big)^2 \Big) \right) \mathrm{d}\boldsymbol{x} \; . \tag{7}$$

Calculating now the gradient descent for $\tilde{E}$, and eliminating z by the inverse function $z = \varphi^{-1}(u)$ leads to the new gradient descent

$$\partial_t u = -\varphi'\big(\varphi^{-1}(u)\big) \frac{\delta E}{\delta u} \; . \tag{8}$$

It is important to note firstly that this is still a gradient descent for the same energy as (5) but with the underlying varying function z instead of u. Secondly, it differs from the unconstrained gradient descent (6) only by the factor $\varphi'\big(\varphi^{-1}(u)\big)$. For the case of the positivity constraint $\varphi(z) = \exp(z)$ one has $\varphi'\big(\varphi^{-1}(u)\big) = u$. Similarly, to impose an interval constraint $a < u < b$ one can choose a sigmoid-shaped function like $\varphi(z) = \frac{a\exp(-z)+b}{\exp(-z)+1}$, which implies $\varphi'\big(\varphi^{-1}(u)\big) = (u-a)(b-u)/(b-a)$.

2.4 Reinterpretations

The reparametrisation approach can be rephrased in several ways, each of which is fruitful for further generalisations. Geometrically, performing the gradient descent based on the variation of the function z comes down to a function space transformation. Note that the derivation of (6) relied on the condition $\delta_v E = \big\langle \frac{\delta E}{\delta u}, v \big\rangle$ in which $\langle \cdot\,, \cdot \rangle$ denotes the standard inner product $\langle \delta u_1, \delta u_2 \rangle = \int_\Omega \delta u_1(\boldsymbol{x}) \delta u_2(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$ of perturbation functions which itself comes from the Euclidean metric $\mathrm{d}u$ on the range of u, i.e. the grey-values. Supposing positive values for u, one can easily replace this metric by a hyperbolic one, $\mathrm{d}_\mathrm{h} u := \mathrm{d}u/u$, which entails a new inner product $\langle \delta u_1, \delta u_2 \rangle_\mathrm{h}$. Requiring $\delta_v E = \langle g, v \rangle_\mathrm{h}$ immediately leads to the gradient $g = u \cdot \frac{\delta E}{\delta u}$ in accordance with (8) for $\varphi \equiv \exp$.

Thus, the positivity-constrained gradient descent turns out to be the gradient descent in a hyperbolic metric. From this viewpoint, zero and negative grey-values are avoided because they are put at infinite distance from any positive values. This reinterpretation can immediately be transferred to interval constraints with a suitable metric on (a, b).

Restricting to the positivity case, another reinterpretation is obtained by using a *multiplicative* gradient descent. To this end, one considers a multiplicative perturbation with a test function v and computes the variation $\delta_v^* E := \frac{\mathrm{d}}{\mathrm{d}\varepsilon} E[u(1+\varepsilon v)]\big|_{\varepsilon=0}$. The requirement $\delta_v^* E = \langle g, v \rangle$ again leads to $g = u \cdot \frac{\delta E}{\delta u}$.

2.5 Multi-Channel Images

Assume now that the blurred image is a multi-channel image $\boldsymbol{f} = (f_i)_{i \in \Gamma}$ with a channel index set Γ, e.g. an RGB colour image, whose channels are uniformly

blurred, i.e. the PSF H is equal for all channels. Denoting by Γ an index set for the channels, the energy functional (5) has to be slightly adapted by summing over Γ in the arguments of both penalisers,

$$E[\boldsymbol{u}] = \int_\Omega \left(\Phi\Big(\sum_{i\in\Gamma} (f_i - H \circledast u_i)^2\Big) + \alpha \Psi\Big(\sum_{i\in\Gamma} |\nabla u_i|^2\Big)\right) \mathrm{d}\boldsymbol{x} \,. \tag{9}$$

In the unconstrained gradient the sum over channels is found in the nonlinearities while linear operations are applied channelwise [14]:

$$\begin{aligned} -\frac{\delta E}{\delta u_i} = H^* \circledast \left(\Phi'\Big(\sum_{i\in\Gamma} (f_i - H \circledast u_i)^2\Big) \cdot (f_i - H \circledast u_i)\right) \\ + \alpha \operatorname{div}\left(\Psi'\Big(\sum_{i\in\Gamma} |\nabla u_i|^2\Big) \nabla u_i \right) . \end{aligned} \tag{10}$$

Positivity or interval constraints on the individual channels can be handled in a straightforward way by assuming channel-wise reparametrisation functions φ_i or channel-wise multiplicative perturbation.

Based on the geometric reinterpretation of the constrained gradient descent, even more complicated inequality constraints can be handled. In [29] this was demonstrated for images with positive definite symmetric matrices (diffusion tensors) as values.

3 Variational View on Richardson-Lucy Deconvolution

Let us now return to the Richardson–Lucy iteration (4). For given f, H, equation (4) can be understood as a fixed point iteration associated to the minimisation of the functional

$$E_{f,H}[u] := \int_\Omega \left(H \circledast u - f - f \ln \frac{H \circledast u}{f} \right) \mathrm{d}\boldsymbol{x} \tag{11}$$

compare [26]. This is the so-called *information divergence* introduced by Csiszár [10]. The asymmetric penaliser function $r_f(w) = w - f - f\ln(w/f)$ is strictly convex for all $w > 0$ and takes its minimum at $w = f$.

As a necessary condition for u to be a minimiser of (11), one can compute an Euler-Lagrange equation which in this case becomes particularly simple as no derivatives of u are present in the integrand. In view of the positivity requirement for u we base our derivation again on a multiplicative perturbation with a test

function v,

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} E_{f,H}[u(1+\varepsilon v)]\right|_{\varepsilon=0} = \int_\Omega \left(1 - \frac{f}{H \circledast u}\right) (H \circledast (uv))\, \mathrm{d}\boldsymbol{x} \tag{12}$$
$$= \int_\Omega \left(H^* \circledast \left(1 - \frac{f}{H \circledast u}\right)\right) uv\, \mathrm{d}\boldsymbol{x}\ .$$

Requiring that this expression vanishes for all test functions v yields the minimality condition

$$\left(H^* \circledast \left(1 - \frac{f}{H \circledast u}\right)\right) u = 0\ . \tag{13}$$

Because of the energy conservation property $H \circledast 1 = 1$ one sees that (4) is a fixed point iteration for (13).

3.1 Regularisation

In the presence of noise the functional (11) is not minimised by a smooth function u; in fact, the fixed-point iteration diverges for $k \to \infty$. From the variational viewpoint, the functional (11) needs to be regularised. In standard Richardson-Lucy deconvolution, this regularisation is provided implicitly by stopping the iteration after finitely many steps. The earlier the iteration is stopped, the higher is the degree of regularisation.

Although the functional (11) does not incorporate this sort of regularisation, the variational picture is advantageous because it allows to modify (11) in the same flexible way as standard variational approaches. The structure of the iterative minimisation procedure is preserved throughout these modifications, which leads to good computational efficiency.

Let us first note that by limiting the growth of high-frequency signal components, regularisation has a smoothing effect that in deconvolution problems acts contrary to the intended image sharpening. It is desirable to steer this effect in such a way that it interferes as little as possible with the enhancement of salient image structures, such as edges.

Implicit regularisation by stopping, however, is not easy to control. For this reason, it makes sense to introduce a variational regularisation term into the objective functional. This yields the functional

$$E_{f,H}[u] = \int_\Omega \left(H \circledast u - f - f \ln \frac{H \circledast u}{f} + \alpha \Psi(|\nabla u|^2)\right) \mathrm{d}\boldsymbol{x} \tag{14}$$

in which the Richardson-Lucy data term is complemented by a regulariser whose influence is weighted by the regularisation weight $\alpha > 0$. Concerning the penalisation function $\Psi(\,\cdot\,)$ in the regulariser, our discussion from Section 2 applies

analogously. With the total variation regulariser given by $\Psi(z^2) = |z|$, the energy functional (14) corresponds to the method proposed (in space-invariant formulation) by Dey et al. [11, 12].

The Euler-Lagrange equation for (14) is given by

$$0 = \left(H^* \circledast \left(1 - \frac{f}{H \circledast u} \right) - \alpha \operatorname{div} \left(\Psi'(|\nabla u|^2) \, \nabla u \right) \right) \cdot u \; . \tag{15}$$

In converting this into a fixed point iteration, we evaluate the divergence expression with $u = u^k$, yielding $D^k := \alpha \operatorname{div} \left(\Psi'(|\nabla u^k|^2) \, \nabla u^k \right)$. Dependent on whether the factor u with which the divergence term in (15) is multiplied is chosen as u^k or u^{k+1}, the right-hand side of (4) receives either the additional summand D^k, or is divided by $(1 - D^k)$. However, D^k can have either sign, and a negative value in the numerator or denominator will always lead to a violation of the positivity requirement.[1] For this reason, we choose the outer factor for D^k as u^k if $D^k > 0$, or u^{k+1} if $D^k < 0$. Using the abbreviations $[z]_\pm := \frac{1}{2}(z \pm |z|)$ we can therefore write our final fixed point iteration as

$$u^{k+1} = \frac{H^* \circledast \left(\frac{f}{H \circledast u^k} \right) + \alpha \left[\operatorname{div} \left(\Psi'(|\nabla u^k|^2) \, \nabla u^k \right) \right]_+}{1 - \alpha \left[\operatorname{div} \left(\Psi'(|\nabla u^k|^2) \, \nabla u^k \right) \right]_-} u^k \; . \tag{16}$$

We will refer to this method as *regularised RL*.

3.2 Robust Data Terms

Up to scaling and shifting, the asymmetric penaliser function $r_f(w)$ equals the logarithmic density of a Gamma distribution. Minimisation of the integral (11) therefore corresponds to a Bayesian estimation of the sharp image assuming a Poisson distribution for the image values whose conjugate prior is the Gamma distribution.

In the context of variational deconvolution, it has turned out useful to replace quadratic data terms that fit a Gaussian noise model by robust data terms that mirror noise distributions with "heavy tails". Not only can the resulting model handle extreme noise but it can also cope with imprecisions in the blur model. Following this idea, we will now replace the data term of (14) by one that is adapted to a broader distribution on $\mathbb{R}^+$. In order to preserve the structure of the fixed point iterations (4) and (16), we hold on to using $r_f(H \circledast u)$ in the data term, but apply a penaliser function Φ that grows less than linear.

Our modified functional therefore reads

$$E_{f,H}[u] = \int_\Omega \left(\Phi \left(H \circledast u - f - f \ln \frac{H \circledast u}{f} \right) + \alpha \Psi(|\nabla u|^2) \right) \mathrm{d}\boldsymbol{x} \; . \tag{17}$$

[1] In [12], D^k is placed entirely into the denominator, implying an upper bound on the regularisation weight.

Our particular choice for the data term penaliser will be $\Phi(z) = 2\sqrt{z}$. This choice resembles the function Φ that leads to L^1 robust data terms in conventional variational approaches. One difference, however, needs to be noted: Unlike its L^1 counterpart, the penaliser function $(\Phi \circ r_f)(w)$ is non-convex, which in principle can induce additional local minima of the functional E.

By an analogous derivation as before, one obtains for (17) the minimality condition

$$0 = \left(H^* \circledast \left(\Phi'(r_f(H \circledast u)) \left(1 - \frac{f}{H \circledast u} \right) \right) - \alpha \operatorname{div} \left(\Psi'(|\nabla u|^2) \nabla u \right) \right) \cdot u \tag{18}$$

that leads to the new fixed point iteration

$$u^{k+1} = \frac{H^* \circledast \left(\Phi'(r_f(H \circledast u)) \frac{f}{H \circledast u^k} \right) + \alpha \left[\operatorname{div} \left(\Psi'(|\nabla u^k|^2) \nabla u^k \right) \right]_+}{H^* \circledast \Phi'(r_f(H \circledast u)) - \alpha \left[\operatorname{div} \left(\Psi'(|\nabla u^k|^2) \nabla u^k \right) \right]_-} u^k \tag{19}$$

which we will call *robust and regularised RL deconvolution* (RRRL). Comparing to (16), the computational effort is increased by one more convolution and the evaluation of Φ'. With our particular choice $\Phi(z) = 2\sqrt{z}$ we have $\Phi'(z) = \frac{1}{\sqrt{z}}$. In the computation, this expression with its singularity at 0 is replaced for stabilisation with $(z^2 + \beta)^{-1/4}$ with a small $\beta > 0$.

Of course, (19) contains the regularised RL method (16) as special case ($\Phi' \equiv 1$). On the other hand, $\alpha = 0$ gives a non-regularised method which we will call *robust RL deconvolution*.

3.3 Multi-Channel Images

Replacing as in Section 2.5 the expressions $r_f(H \circledast u)$ and $|\nabla u|^2$ in the arguments of Φ and Ψ with their sums over the image channels, $R = \sum_{i \in \Gamma} r_{f_i}(H \circledast u_i)$ and $G = \sum_{i \in \Gamma} |\nabla u_i|^2$, we obtain as multi-channel analog of (19) the functional

$$E_{\boldsymbol{f},H}[\boldsymbol{u}] = \int_\Omega \left(\Phi(R) + \alpha \Psi(G) \right) \mathrm{d}\boldsymbol{x} \,. \tag{20}$$

This leads to the following multi-channel version of robust and regularised RL deconvolution

$$u_j^{k+1} = \frac{H^* \circledast \left(\Phi'(R) \frac{f_j}{H \circledast u_j^k} \right) + \alpha \left[\operatorname{div} \left(\Psi'(G) \nabla u_j^k \right) \right]_+}{H^* \circledast \Phi'(R) - \alpha \left[\operatorname{div} \left(\Psi'(G) \nabla u_j^k \right) \right]_-} u_j^k \,. \tag{21}$$

The same procedure works for the non-robust and/or non-regularised RL variants. Note that in the case of standard RL this boils down to channel-wise application.

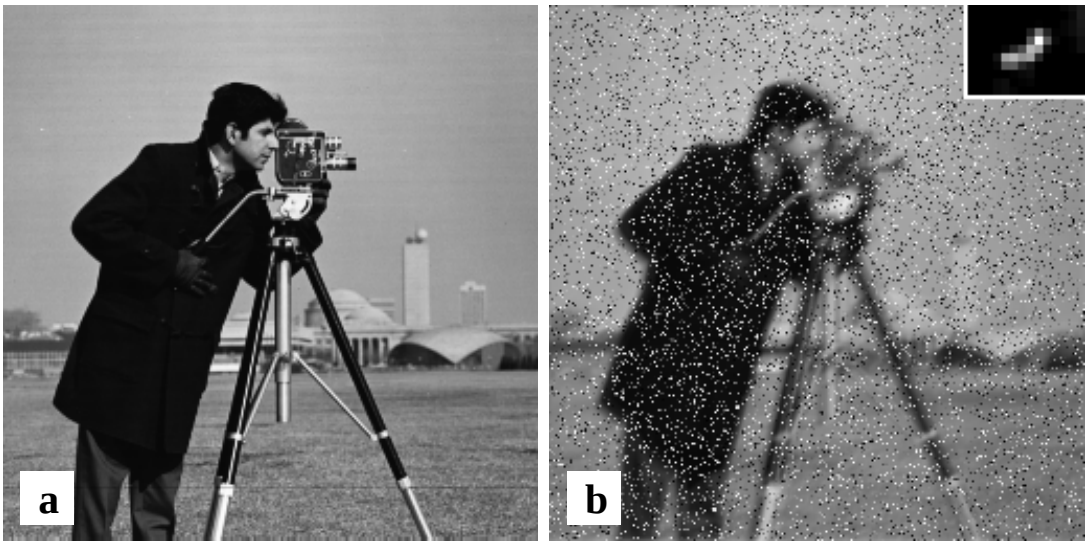

Figure 1: **(a)** *Cameraman* image, 256×256 pixels. **(b)** Moderately blurred, and 15 % of all pixels replaced by uniform noise, signal-to-noise ratio: 3.66 dB. Insert shows point-spread function (four times enlarged).

4 Experiments

Let us first evaluate the performance of our newly developped deconvolution methods on synthetic grey-value data. Since in this case the correct sharp image is known, we can also assess processed images by their signal-to-noise ratio

$$\mathrm{SNR}(u, g) = 10 \log_{10} \frac{\mathrm{var}(g)}{\mathrm{var}(g - u)} \,\mathrm{dB} \,. \tag{22}$$

Here, g and u are the original sharp image processed image, respectively. By $\mathrm{var}(v)$ we denote the variance of the image v. It should, however, be noted that SNR measurements do often not capture visual quality of deblurred images very well. The parameters in our experiments are optimised primarily for visual quality, not for SNR.

We remark also that in synthetically blurring images, we use convolution via the Fourier domain, which involves treating the image domain as periodic. In the deconvolution procedures, in contrast, the convolution operations are carried out in the spatial domain, extending the image where necessary constant along boundary normals. This discrepancy in the convolution procedure and boundary treatment is by purpose: it helps to prevent "inverse crimes" [9] that could unduly embellish results.

In our first series of experiments (Figures 1–4), we blur the *cameraman* image, Figure 1(a), by an irregularly shaped point-spread function of moderate size and apply impulsive (uniform) noise (b). The so degraded image can be deblurred to some extent by standard RL, Figure 2(a), but the noise is severely amplified, and

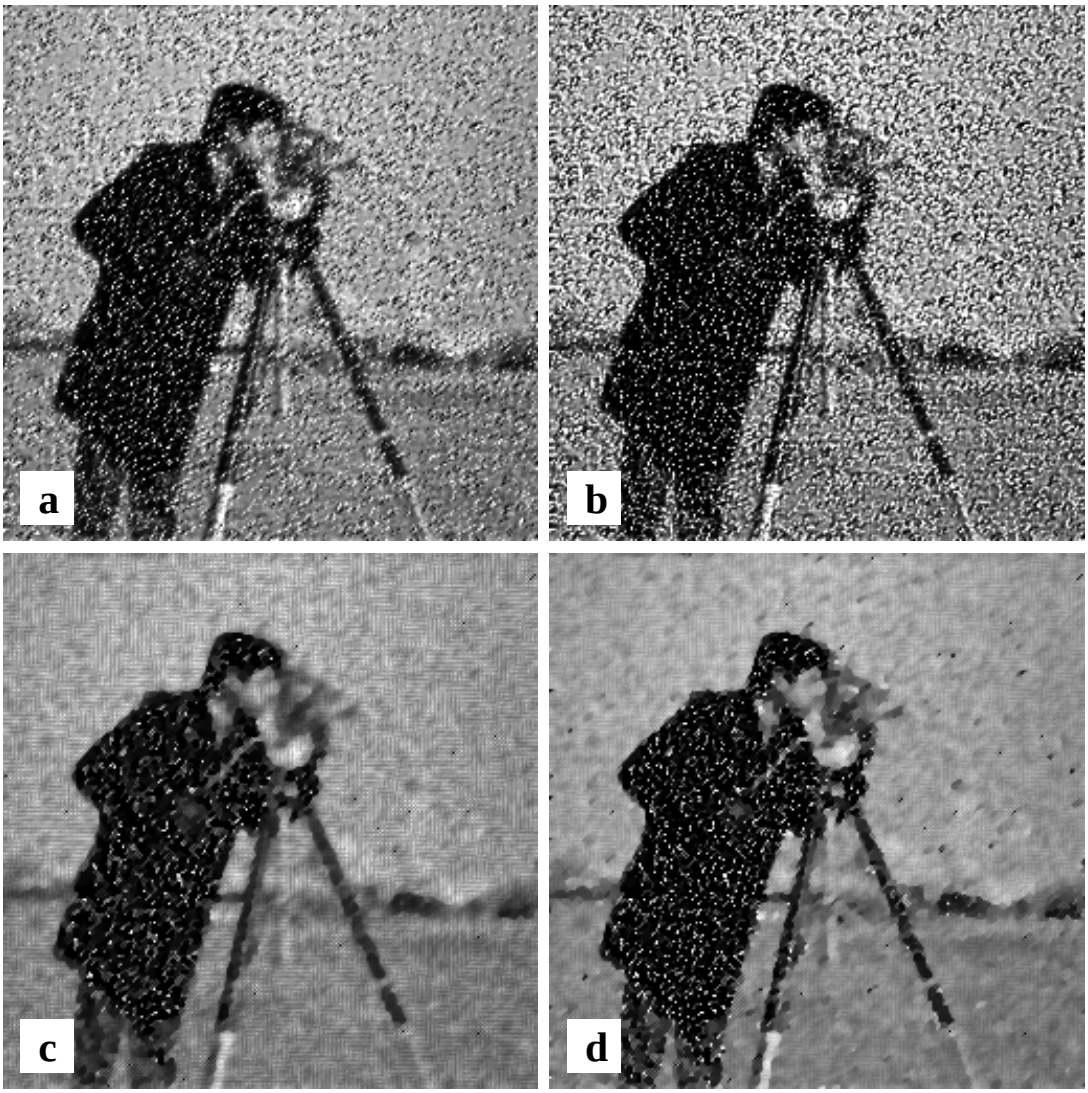

Figure 2: **(a)** Image from Figure 1(b) deblurred by 10 iterations of standard Richardson-Lucy deconvolution (4), SNR: 2.28 dB. **(b)** 30 iterations of RL deconvolution (4), SNR: −0.38 dB. **(c)** 100 iterations of regularised Richardson-Lucy deconvolution (16), $\alpha = 0.1$, SNR: 6.40 dB. **(d)** 100 iterations of regularised Richardson-Lucy deconvolution, $\alpha = 0.05$, SNR: 5.95 dB.

dominates the result when the iteration count is increased for further sharpening (b). Regularised RL (compare [11, 12]) visibly reduces the noise effect, see (c) and (d) for slightly different regularisation weights α.

Robust variational deconvolution achieves a significantly better restoration, see Figure 3(a) where the method from [31] without positivity constraint is used, and Figure 3(b) with positivity constraint as in [29] and Section 3. Visually, both variational approaches lead to a comparable restoration quality. It is, however,

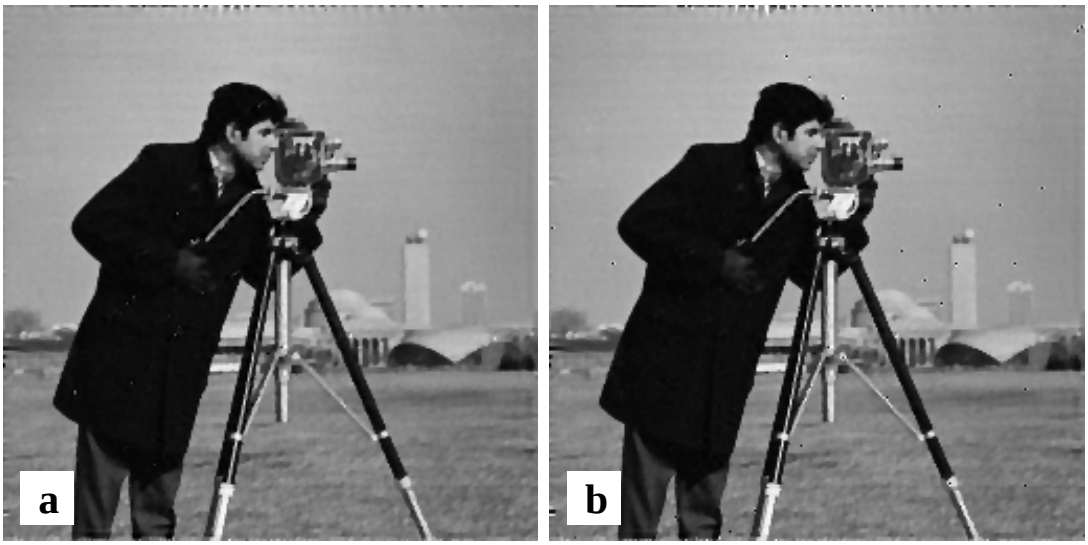

Figure 3: **(a)** Image from Figure 1(b) deblurred by robust variational deconvolution without constraints [31] using L^1 data term and Perona-Malik regulariser ($\lambda = 0.15$), regularisation weight $\alpha = 0.06$; SNR: 16.03 dB. **(b)** Same but with positivity constraint, see [29] and Section 3, regularisation weight $\alpha = 0.06$; SNR: 15.42 dB.

evident that in this example the positivity constraint bears no improvement. This is due to the fact that artifact suppression by the positivity constraint takes place mainly in dark image regions. In the cameraman image, however,

understandable given that dark regions play a minor role in the present test image, and the image degradation by blur and noise is only moderate.

Using robust RL without regularisation as in Figure 4 reduces the noise to an extent comparable to the regularised method in Figure 2(c, d). With more iterations, Figure 4(b), however, the noise still becomes dominant, thus limiting the possible deconvolution quality. Robust and regularised RL allows fairly good deblurring while only small rudiments of noise remain, see Figure 4(c). Even if the iteration count is drastically increased, such that the implicit regularisation by stopping takes only little effect, the explicit regularisation remains effective (d) and keeps the noise level low. Indeed, the reconstruction quality comes close, both visually and in terms of SNR, to that achieved by state-of-the-art robust variational deconvolution as shown before in Figure 3.

Encouraged by these findings, we increase blur and noise in our second series of experiments (Figures 5–6). Under the influence of a drastically larger simulated motion blur and doubled noise intensity, see Figure 5(b), classic RL gives no longer usable results (c). Regularised RL and robust RL can again cope better with the noise but their outcomes are far from satisfactory, see Figure 5(d) and 6(a). A significant improvement is obtained with robust and regularised RL, Figure 6(b), as well as with robust variational deconvolution without constraints

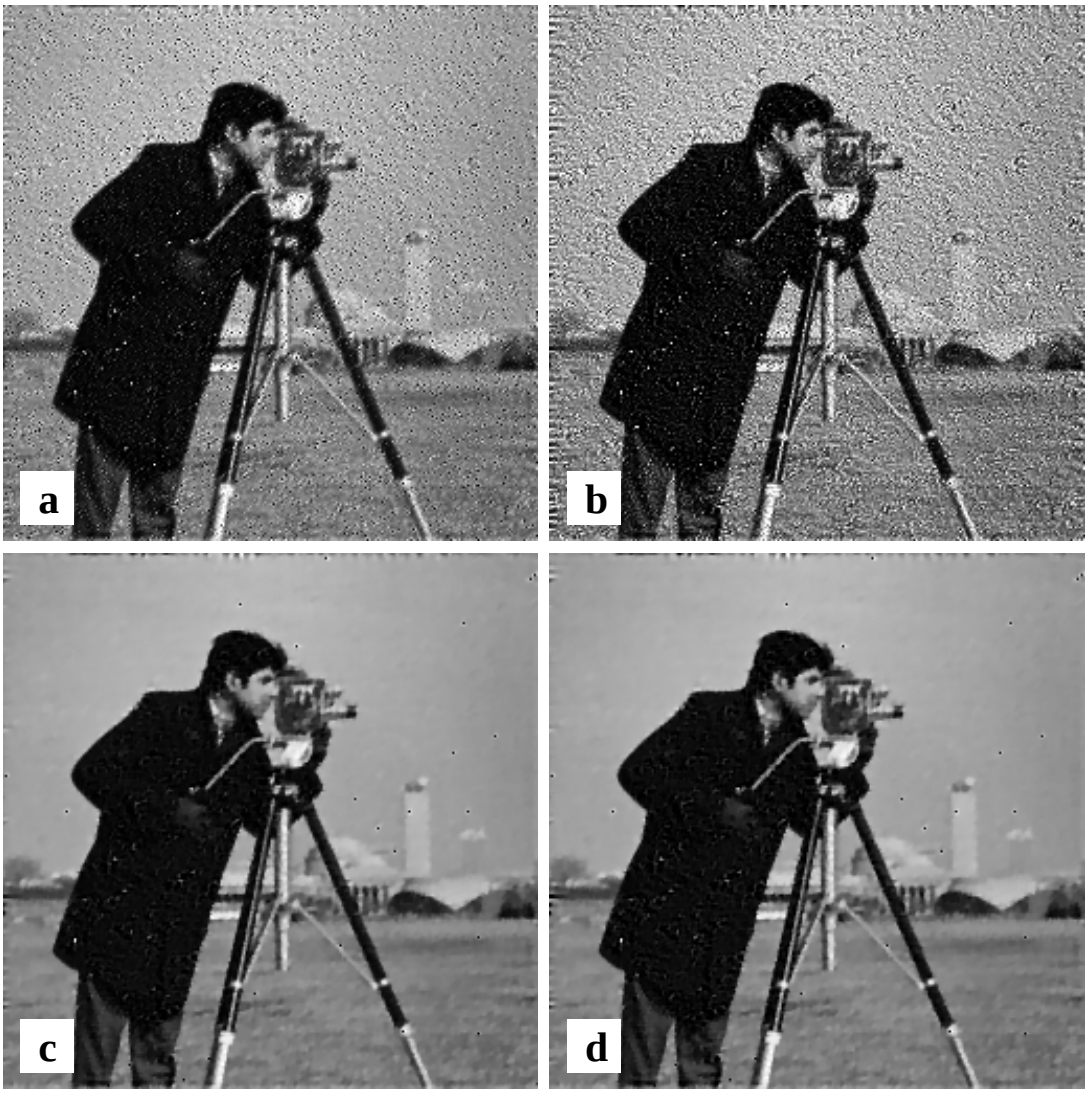

Figure 4: **(a)** Image from Figure 1(b) deblurred by 50 iterations of robust RL, SNR: 6.42 dB. **(b)** 200 iterations of robust RL, SNR: 4.00 dB. **(c)** 200 iterations of robust and regularised RL (19) with regularisation weight $\alpha = 0.005$; SNR: 14.41 dB. **(d)** Same but 2000 iterations; SNR: 14.21 dB.

(c) or with positivity constraint (d). Although the restoration quality is still imperfect, the three last mentioned methods are eye to eye in terms of visual quality and SNR.

In both synthetic experiments, the run times of the robust and regularised RL experiments had to exceed that of standard RL largely in order to achieve this high restoration quality, but remained by a factor $3 \ldots 5$ below those for the robust variational model from literature. The reason for this lies in the favourable structure of the minimality condition and the fixed point iteration obtained from

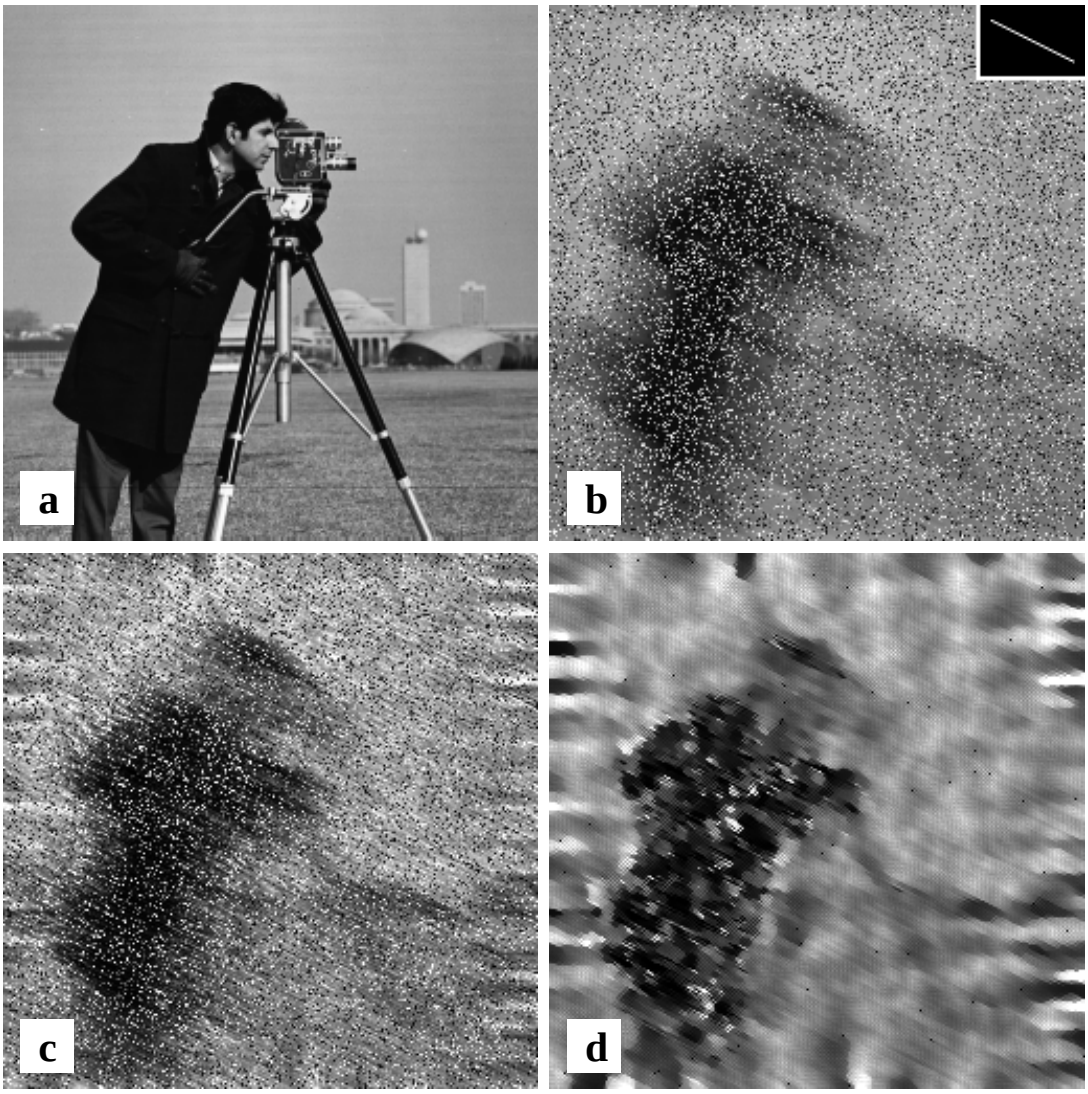

Figure 5: **(a)** *Cameraman* image, 256×256 pixels. **(b)** Severely blurred, and 30 % of all pixels replaced by uniform noise, signal-to-noise ratio: 0.26 dB. Insert shows point-spread function (true size). **(c)** Deblurred by 10 iterations of standard Richardson-Lucy deconvolution (4), SNR: 0.09 dB. **(d)** 100 iterations of regularised RL deconvolution (16) with regularisation weight $\alpha = 0.05$; SNR: 2.32 dB.

it. In contrast, the minimisation of the classical variational model becomes very slow when getting close to the optimum, thus requiring much more iterations.

Our last experiment (Figures 7–9) is based on real-world data. The colour photograph shown in Figure 7(a) was blurred during acquisition with an unknown point-spread function that can be inferred approximately from the shape of a point light source. For restoration, we use the multi-channel versions of all

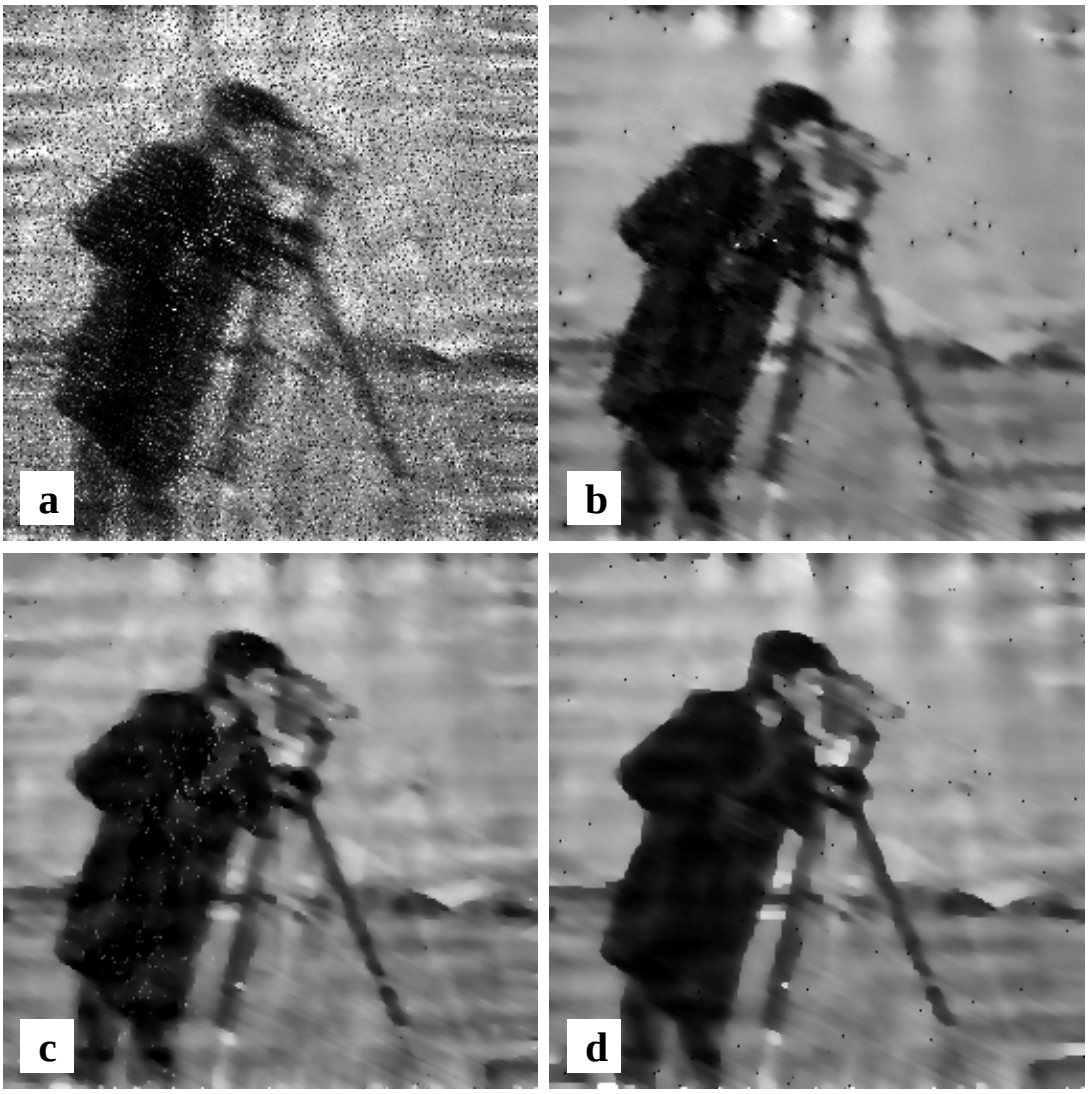

Figure 6: **(a)** Image from Figure 5(b) deblurred by 100 iterations of robust RL, SNR: 1.84 dB. **(b)** 400 iterations of robust and regularised RL (19) with regularisation weight $\alpha = 0.003$; SNR: 7.42 dB. **(c)** Robust variational deconvolution without constraint, $\alpha = 0.06$; SNR: 7.52 dB. **(d)** Robust variational deconvolution with positivity constraint, $\alpha = 0.09$; SNR: 7.60 dB.

methods.

Restoration by standard RL as in Figure 7(b) achieves a decent acuity at moderate computational cost, as also demonstrated in the detail view, Figure 9(a). Increasing the number of iterations quickly leads to ringing artifacts that are visible as shadows in the vicinity of all high-contrast image structures, see Figure 9(b).

Variational deconvolution with a robust L^1 data term and Perona-Malik reg-

Figure 7: **(a)** Colour photograph (Paris from Eiffel Tower) blurred during acquisition, 480×480 pixels. Insert shows approximate point-spread function used for deconvolution (same as in Figure 1, eight times enlarged). **(b)** Richardson-Lucy deconvolution, 20 iterations.

ulariser allows a visible improvement in acuity over RL deconvolution while reducing artifacts, see Figure 8(a) and the detail view in Figure 9(c). Using the positivity-constrained gradient descent (8) brings about a further significant improvement, see Figure 8(b) and 9(f). Due to the better suppression of ringing artifacts the regularisation weight α could be reduced by half here – in contrast, unconstrained variational deconvolution with the same reduced α creates much stronger artifacts, see Figure 9(d). Imposing the constraint but retaining the larger weight α, see Figure 9(e), already improves acuity but still smoothes out more fine details than in Figure 9(f).

The excellent restoration quality of variational deconvolution, however, comes at the cost of significantly increased computation time needed in order to approx-

Table 1: Approximate computational expense for deconvolution results shown in Figure 8. Absolute computation times (referring to single-threaded calculation on a Core2Duo at 1.86 GHz) are given for rough orientation.

Method	Iterations	Computation time (s)	Cost factor w.r.t. standard RL
Standard RL	20	20.9	1.0
Variational	1500	1735.5	83.0
RRRL (TV)	80	124.3	5.9
RRRL (PM)	300	448.4	21.5

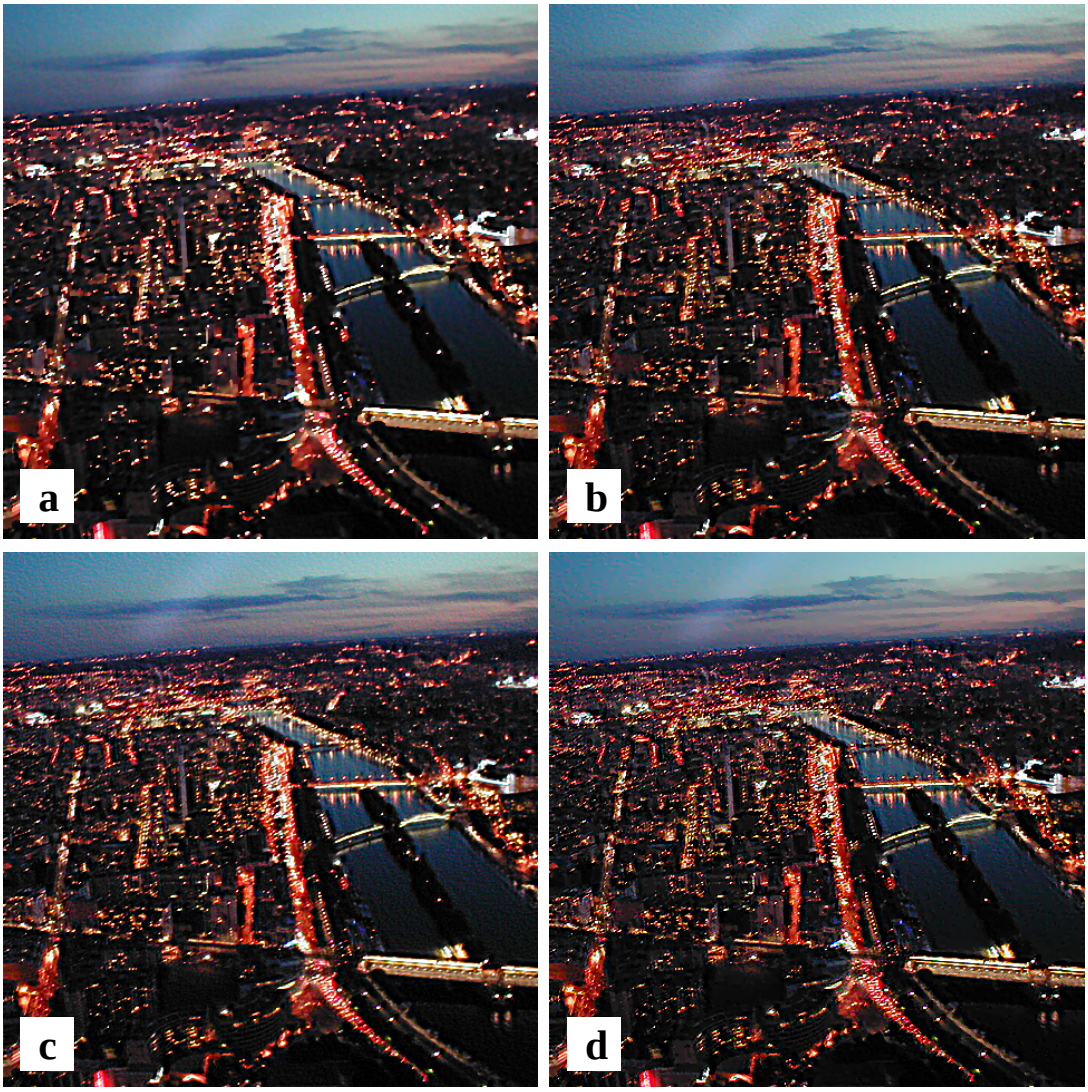

Figure 8: **(a)** Robust variational deconvolution of the image from Figure 7(a) using Perona-Malik regulariser ($\lambda = 26$) without constraints [31], regularisation weight $\alpha = 0.06$. **(b)** Same with positivity constraint [29], $\alpha = 0.03$. **(c)** Robust and regularised RL deconvolution, 80 iterations with TV regulariser, $\alpha = 0.002$. **(d)** RRRL, 300 iterations with Perona-Malik regulariser ($\lambda = 26$) and $\alpha = 0.18$.

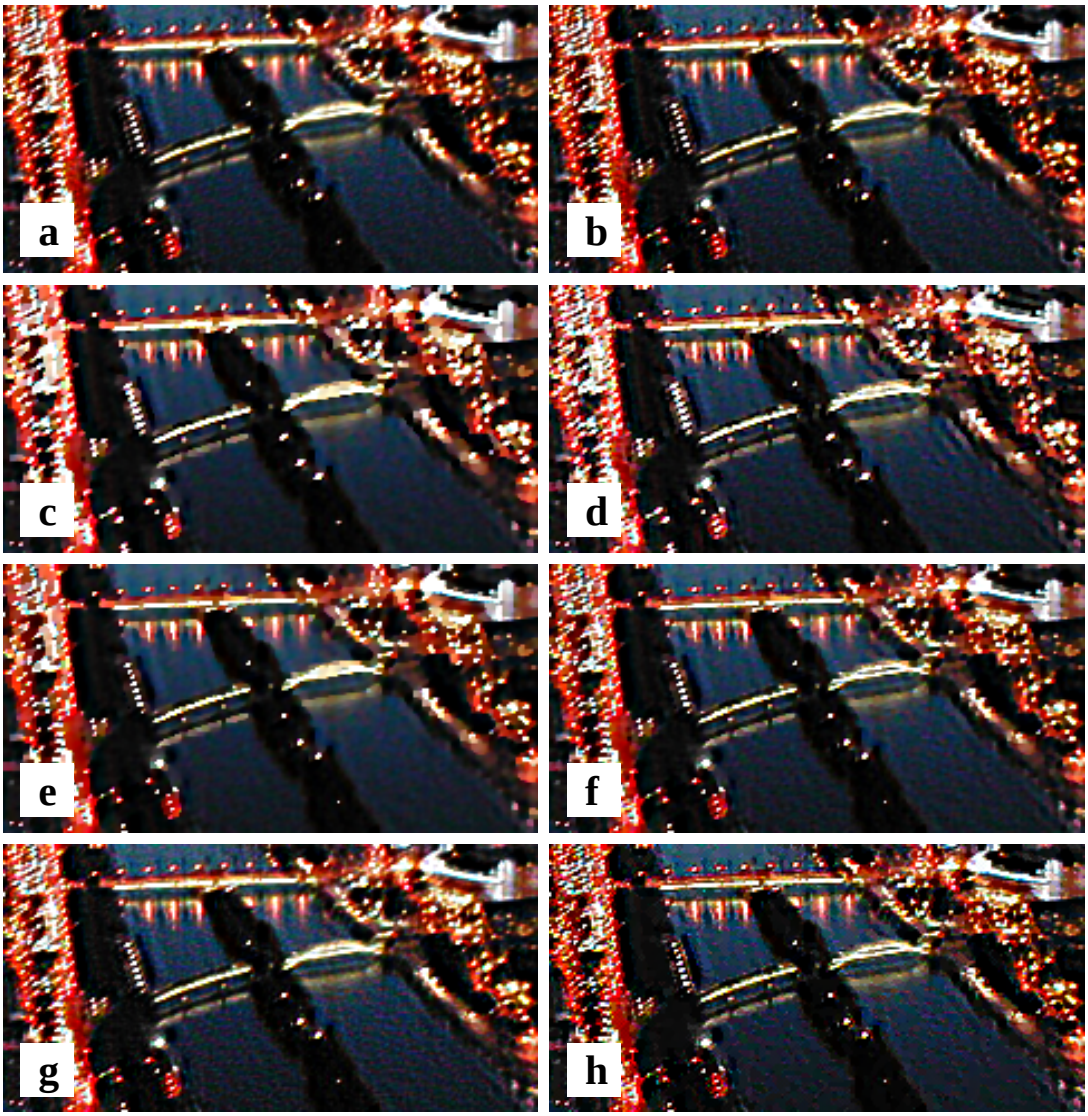

Figure 9: Detail from deconvolution results for the image from Figure 8(a). **(a)** Richardson-Lucy, 20 iterations, see Figure 8(b). **(b)** Same but 30 iterations. **(c)** Robust variational deconvolution using Perona-Malik regulariser ($\lambda = 26$) without constraints, regularisation weight $\alpha = 0.06$, see Figure 8(c). **(d)** Same but $\alpha = 0.03$. **(e)** Robust variational deconvolution using Perona-Malik regulariser ($\lambda = 26$) with positivity constraint, regularisation weight $\alpha = 0.06$. **(f)** Same but $\alpha = 0.03$, see Figure 8(d). **(g)** Robust and regularised RL deconvolution, 80 iterations with TV regulariser, $\alpha = 0.002$, see Figure 8(e). **(h)** RRRL, 300 iterations with Perona-Malik regulariser ($\lambda = 26$) and $\alpha = 0.18$, see Figure 8(f).

imate the steady state. Robust and regularised Richardson-Lucy deconvolution as shown in the last rows of Figures 8 and 9 provides an attractive compromise between standard RL and the variational gradient descent. Figures 8(c) and 9(g) show a RRRL result with TV regulariser which can be computed fairly fast. When the Perona-Malik regulariser is used instead, see Figures 8(d) and 9(h), a higher number of iterations is required in order for the edge-enhancing properties of the regulariser to pay off, but still the computation time is lower than with the gradient descent algorithm, compare also Table 1. In terms of restoration quality, both RRRL results range between the variational deconvolution without and with constraints. In assessing this finding, it has to be taken into account that the test image consisting of large dark regions with few highlights makes the positivity constraint particularly valuable.

5 Conclusions

In this paper, we have demonstrated firstly how a positivity constraint for grey- or colour-values can be accommodated in a robust variational deconvolution framework. Building on previous work [29] about a modified gradient descent for inequality constraints, we have presented this approach in a general formulation with a space-variant point-spread function and rewritten it in terms of a multiplicative gradient descent.

Secondly we have investigated Richardson-Lucy deconvolution from the variational viewpoint. Based on the observation [26] that the RL method can be understood as a fixed point iteration for the minimisation of the information divergence [10] in the role of an energy functional, it is embedded into the framework of variational methods. This allows in turn to apply to it the modifications that have made variational deconvolution the flexible and high-quality deconvolution tool that it is. As a result, we have obtained a novel robust and regularised Richardson-Lucy deconvolution method that competes in quality with state-of-the-art variational methods while preserving to a great extent the numerical efficiency of Richardson-Lucy deconvolution.

The asymmetric penaliser function $r_f(w)$ underlying the information divergence functional has played a little role in variational image analysis so far. We believe that this type of penaliser function can be applied advantageously in a variety of other variational approaches in image processing that handle positivity-constrained image data.

References

[1] M. Backes, T. Chen, M. Dürmuth, H. Lensch, and M. Welk. Tempest in a teapot: compromising reflections revisited. In *Proc. 30th IEEE Symposium*

on Security and Privacy, pages 315–327, Oakland, California, USA, 2009.

[2] L. Bar, A. Brook, N. Sochen, and N. Kiryati. Color image deblurring with impulsive noise. In N. Paragios, O. Faugeras, T. Chan, and C. Schnörr, editors, *Variational and Level Set Methods in Computer Vision*, volume 3752 of *Lecture Notes in Computer Science*, pages 49–60. Springer, Berlin, 2005.

[3] L. Bar, N. Sochen, and N. Kiryati. Variational pairing of image segmentation and blind restoration. In T. Pajdla and J. Matas, editors, *Computer Vision – ECCV 2004, Part II*, volume 3022 of *Lecture Notes in Computer Science*, pages 166–177. Springer, Berlin, 2004.

[4] L. Bar, N. Sochen, and N. Kiryati. Image deblurring in the presence of salt-and-pepper noise. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Scale Space and PDE Methods in Computer Vision*, volume 3459 of *Lecture Notes in Computer Science*, pages 107–118. Springer, Berlin, 2005.

[5] L. Bar, N. Sochen, and N. Kiryati. Restoration of images with piecewise space-variant blur. In F. Sgallari, F. Murli, and N. Paragios, editors, *Scale Space and Variational Methods in Computer Vision*, volume 4485 of *Lecture Notes in Computer Science*, pages 533–544. Springer, Berlin, 2007.

[6] M. Bertero, T. A. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76(8):869–889, Aug. 1988.

[7] E. Bratsolis and M. Sigelle. A spatial regularization method preserving local photometry for Richardson-Lucy restoration. *Astronomy and Astrophysics*, 375(3):1120–1128, 2001.

[8] T. F. Chan and C. K. Wong. Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7:370–375, 1998.

[9] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer, Berlin, 1992.

[10] I. Csiszár. Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems. *Annals of Statistics*, 19(4):2032–3066, 1991.

[11] N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. A deconvolution method for confocal microscopy with total variation regularization. In *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, April 2004.

[12] N. Dey, L. Blanc-Feraud, C. Zimmer, P. Roux, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microscopy Research and Technique*, 69:260–266, 2006.

[13] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[14] G. Gerig, O. Kübler, R. Kikinis, and F. A. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11:221–232, 1992.

[15] M. Jung, X. Bresson, T. F. Chan, and L. A. Vese. Color image restoration using nonlocal Mumford-Shah regularizers. In D. Cremers, Y. Boykov, A. Blake, and F. R. Schmidt, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 5681 of *Lecture Notes in Computer Science*, pages 373–387. Springer, Berlin, 2009.

[16] M. Jung and L. A. Vese. Nonlocal variational image deblurring models in the presence of Gaussian or impulse noise. In X.-C. Tai, K. Mørken, M. Lysaker, and K.-A. Lie, editors, *Scale-Space and Variational Methods in Computer Vision*, volume 5567 of *Lecture Notes in Computer Science*, pages 402–413. Springer, Berlin, 2009.

[17] L. B. Lucy. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79(6):745–754, June 1974.

[18] A. Marquina and S. Osher. A new time dependent model based on level set motion for nonlinear deblurring and noise removal. In M. Nielsen, P. Johansen, O. F. Olsen, and J. Weickert, editors, *Scale-Space Theories in Computer Vision*, volume 1682 of *Lecture Notes in Computer Science*, pages 429–434. Springer, Berlin, 1999.

[19] J. G. Nagy and D. P. O'Leary. Restoring images degraded by spatially-variant blur. *SIAM Journal on Scientific Computing*, 19:1063–1082, 1998.

[20] J. G. Nagy, R. J. Plemmons, and T. C. Torgersen. Iterative image restoration using approximate inverse preconditioning. *IEEE Transactions on Image Processing*, 5:1151–1162, 1996.

[21] J. G. Nagy and Z. Strakoš. Enforcing nonnegativity in image reconstruction algorithms. In D. C. Wilson, H. D. Tagare, F. L. Bookstein, F. J. Preteux, and E. R. Dougherty, editors, *Advanced Signal Processing Algorithms, Architectures, and Implementations*, volume 4121 of *Proceedings of SPIE*, pages 182–190. SPIE Press, Bellingham, 2000.

[22] S. Osher and L. Rudin. Total variation based image restoration with free local constraints. In *Proc. 1994 IEEE International Conference on Image Processing*, pages 31–35, Austin, Texas, 1994.

[23] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[24] W. H. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, 1972.

[25] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[26] D. Snyder, T. J. Schulz, and J. A. O'Sullivan. Deblurring subject to nonnegativity constraints. *IEEE Transactions on Image Processing*, 40(5):1143–1150, 1992.

[27] A. N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics Doklady*, 4:1035–1038, 1963.

[28] C. R. Vogel and M. E. Oman. Fast, robust total variation-based reconstruction of noisy, blurred images. *IEEE Transactions on Image Processing*, 7:813–824, 1998.

[29] M. Welk and J. G. Nagy. Variational deconvolution of multi-channel images with inequality constraints. In J. Martí, J. M. Benedí, A. M. Mendonça, and J. Serrat, editors, *Pattern Recognition and Image Analysis*, volume 4477 of *Lecture Notes in Computer Science*, pages 386–393. Springer, Berlin, 2007.

[30] M. Welk, D. Theis, T. Brox, and J. Weickert. PDE-based deconvolution with forward-backward diffusivities and diffusion tensors. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Scale Space and PDE Methods in Computer Vision*, volume 3459 of *Lecture Notes in Computer Science*, pages 585–597. Springer, Berlin, 2005.

[31] M. Welk, D. Theis, and J. Weickert. Variational deblurring of images with uncertain and spatially variant blurs. In W. Kropatsch, R. Sablatnig, and A. Hanbury, editors, *Pattern Recognition*, volume 3663 of *Lecture Notes in Computer Science*, pages 485–492. Springer, Berlin, 2005.

[32] E. T. Whittaker. On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, 41:63–75, 1923.

[33] N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. MIT Press, Cambridge, MA, 1949.

[34] Y.-L. You and M. Kaveh. Anisotropic blind image restoration. In *Proc. 1996 IEEE International Conference on Image Processing*, volume 2, pages 461–464, Lausanne, Switzerland, Sept. 1996.

[35] M. E. Zervakis, A. K. Katsaggelos, and T. M. Kwon. A class of robust entropic functionals for image restoration. *IEEE Transactions on Image Processing*, 4(6):752–773, June 1995.