



HAL
open science

Deblurring Shaken and Partially Saturated Images

Oliver Whyte, Josef Sivic, Andrew Zisserman

► **To cite this version:**

Oliver Whyte, Josef Sivic, Andrew Zisserman. Deblurring Shaken and Partially Saturated Images. International Journal of Computer Vision, 2014, 10.1007/s11263-014-0727-3 . hal-01053888

HAL Id: hal-01053888

<https://inria.hal.science/hal-01053888>

Submitted on 17 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Deblurring Shaken and Partially Saturated Images

Oliver Whyte · Josef Sivic · Andrew Zisserman

Received: date / Accepted: date

Abstract We address the problem of deblurring images degraded by camera shake blur *and* saturated (over-exposed) pixels. Saturated pixels violate the common assumption that the image-formation process is linear, and often cause ringing in deblurred outputs. We provide an analysis of ringing in general, and show that in order to prevent ringing, it is insufficient to simply discard saturated pixels. We show that even when saturated pixels are removed, ringing is caused by attempting to estimate the values of latent pixels that are brighter than the sensor’s maximum output. Estimating these latent pixels is likely to cause large errors, and these errors propagate across the rest of the image in the form of ringing. We propose a new deblurring algorithm that locates these error-prone bright pixels in the latent sharp image, and by decoupling them from the remainder of the latent image, greatly reduces ringing. In addition, we propose an approximate forward model for saturated images, which allows us to estimate these error-prone pixels separately without causing artefacts. Results are shown for non-blind deblurring of

real photographs containing saturated regions, demonstrating improved deblurred image quality compared to previous work.

Keywords Non-Blind Deblurring · Saturation · Ringing · Outliers

1 Introduction

The task of deblurring “shaken” images has received considerable attention recently (Fergus et al. 2006; Cho and Lee 2009; Gupta et al. 2010; Joshi et al. 2010; Whyte et al. 2010; Shan et al. 2008). Significant progress has been made towards reliably estimating the point spread function (PSF) for a given blurred image, which describes how the image was blurred. Likewise, when the PSF for an image is known, many authors have proposed methods to invert the blur process and recover a high quality sharp image (referred to as “non-blind deblurring”).

One problematic feature of blurred images, and in particular “shaken” images, which has received relatively little attention is the presence of saturated (over-exposed) pixels. These arise when the radiance of the scene exceeds the range of the camera’s sensor, leaving bright highlights clipped at the maximum output value (e.g. 255 for an image with 8 bits per pixel). To anyone who has attempted to take handheld photographs at night, such pixels should be familiar as the conspicuous bright streaks left by electric lights, such as in Figure 1a. These bright pixels, with their clipped values, violate the assumption made by most deblurring algorithms that the image formation process is linear, and as a result can cause obtrusive artefacts in the deblurred images. This can be seen in the deblurred images in Figures 1b and 1c.

In this paper we address the problem of deblurring images containing saturated pixels, offering an analysis of the artefacts caused by existing algorithms, and a new algorithm

Parts of this work were previously published in the IEEE Workshop on Color and Photometry in Computer Vision, with ICCV 2011 (Whyte et al. 2011)

O. Whyte
Microsoft Corporation
Redmond, WA, USA
E-mail: oliverw@microsoft.com

J. Sivic
INRIA - Willow Project
Laboratoire d’Informatique de l’Ecole Normale Supérieure
(CNRS/ENS/INRIA UMR 8548), Paris, France
E-mail: josef.sivic@ens.fr

A. Zisserman
Visual Geometry Group
Department of Engineering Science
University of Oxford
Oxford, UK
E-mail: az@robots.ox.ac.uk

which avoids such artefacts by explicitly handling saturated pixels. Our method is applicable for all causes of blur, however in this work we focus on blur caused by camera shake (motion of the camera during the exposure).

The process of deblurring an image typically involves two steps. First, the PSF is estimated, either using a blind deblurring algorithm (Fergus et al. 2006; Yuan et al. 2007; Shan et al. 2008; Cho and Lee 2009) to estimate the PSF from the blurred image itself, or by using additional hardware attached to the camera (Joshi et al. 2010; Tai et al. 2008). Second, a non-blind deblurring algorithm is applied to estimate the sharp image, given the PSF. In this work we address the second of these two steps for the case of saturated images, and assume that the PSF is known or has been estimated already. Unless otherwise stated, all the results in this work use the algorithm of Whyte et al. (2011) to estimate a spatially-variant PSF. The algorithm is based on the method of Cho and Lee (2009), and estimates the PSF directly from the blurred image itself. Figure 1d shows the output of the proposed algorithm, which contains far fewer artefacts than the two existing algorithms shown for comparison.

1.1 Related Work

Saturation has not received wide attention in the literature, although several authors have cited it as the cause of artefacts in deblurred images (Fergus et al. 2006; Cho and Lee 2009; Tai et al. 2011). Harmeling et al. (2010b) address the issue in the setting of multi-frame blind deblurring by thresholding the blurred image to detect saturated pixels, and ignoring these in the deblurring process. When multiple blurred images of the same scene are available, these pixels can be safely discarded, since there will generally remain unsaturated pixels covering the same area in other images.

Recently, Cho et al. (2011) have also considered saturated pixels, in the more general context of non-blind deblurring with outliers, and propose an expectation-maximisation algorithm to iteratively identify and exclude outliers in the blurred image. Saturated pixels are detected by blurring the current estimate of the sharp image and finding places where the result exceeds the range of the camera. Those blurred pixels detected as saturated are ignored in the subsequent iterations of the deblurring algorithm. In Section 4 we discuss why simply ignoring saturated pixels is, in general, not sufficient to prevent artefacts from appearing in single-image deblurring.

In an alternative line of work, several authors have proposed algorithms for non-blind deblurring that are robust against various types of modeling errors, without directly addressing the sources of those errors. Yuan et al. (2008) propose a non-blind deblurring algorithm capable of suppressing “ringing” artefacts during deblurring, using multi-

scale regularisation. Yang et al. (2009) and Xu and Jia (2010) also consider non-blind deblurring with robust data-fidelity terms, to handle non-Gaussian impulse noise, however their formulations do not handle arbitrarily large deviations from the linear model, such as can be caused by saturation.

Many algorithms exist for non-blind deblurring in the linear (non-saturated) case, perhaps most famously the Wiener filter (Wiener 1949) and the Richardson-Lucy algorithm (Richardson 1972; Lucy 1974). Recently, many authors have focused on the use of regularisation, derived from natural image statistics, to suppress noise in the output while encouraging sharp edges to appear (Krishnan and Fergus 2009; Levin et al. 2007; Joshi et al. 2010; Afonso et al. 2010; Tai et al. 2011; Zoran and Weiss 2011).

For the problem of “blind” deblurring, where the PSF is unknown, single-image blind PSF estimation for camera shake has been widely studied using variational and maximum a posteriori (MAP) algorithms (Fergus et al. 2006; Shan et al. 2008; Cho and Lee 2009; Cai et al. 2009; Xu and Jia 2010; Levin et al. 2011; Krishnan et al. 2011). Levin et al. (2009) review several approaches and provide a ground-truth dataset for comparison on spatially-invariant blur. While most work has focused on spatially-invariant blur, several approaches have also been proposed for spatially-varying blur (Whyte et al. 2010; Gupta et al. 2010; Harmeling et al. 2010a; Joshi et al. 2010; Tai et al. 2011).

The remainder of this paper proceeds as follows: We begin in Section 2 by providing some background on non-blind deblurring and saturation in cameras. In Section 3 we analyse some of the properties and causes of “ringing” artefacts (which are common when deblurring saturated images), and discuss the implications of this analysis in Section 4. Based on this discussion, in Section 5 we describe our proposed approach. We present deblurring results and comparison to related work in Section 6.

2 Background

In most existing work on deblurring, the observed image produced by a camera is modelled as a linear blur operation applied to a sharp image, followed by a random noise process. Under this model, an observed blurred image \mathbf{g} (written as an $N \times 1$ vector, where N is the number of pixels in the image) can be written in terms of a (latent) sharp image \mathbf{f} (also an $N \times 1$ vector) as

$$\mathbf{g}^* = \mathbf{A}\mathbf{f} \quad (1)$$

$$\mathbf{g} = \mathbf{g}^* + \boldsymbol{\varepsilon}, \quad (2)$$

where \mathbf{A} is an $N \times N$ matrix representing the discrete PSF, \mathbf{g}^* represents the “noiseless” blurred image, and $\boldsymbol{\varepsilon}$ is some random noise affecting the image. Typically, the noise $\boldsymbol{\varepsilon}$ is

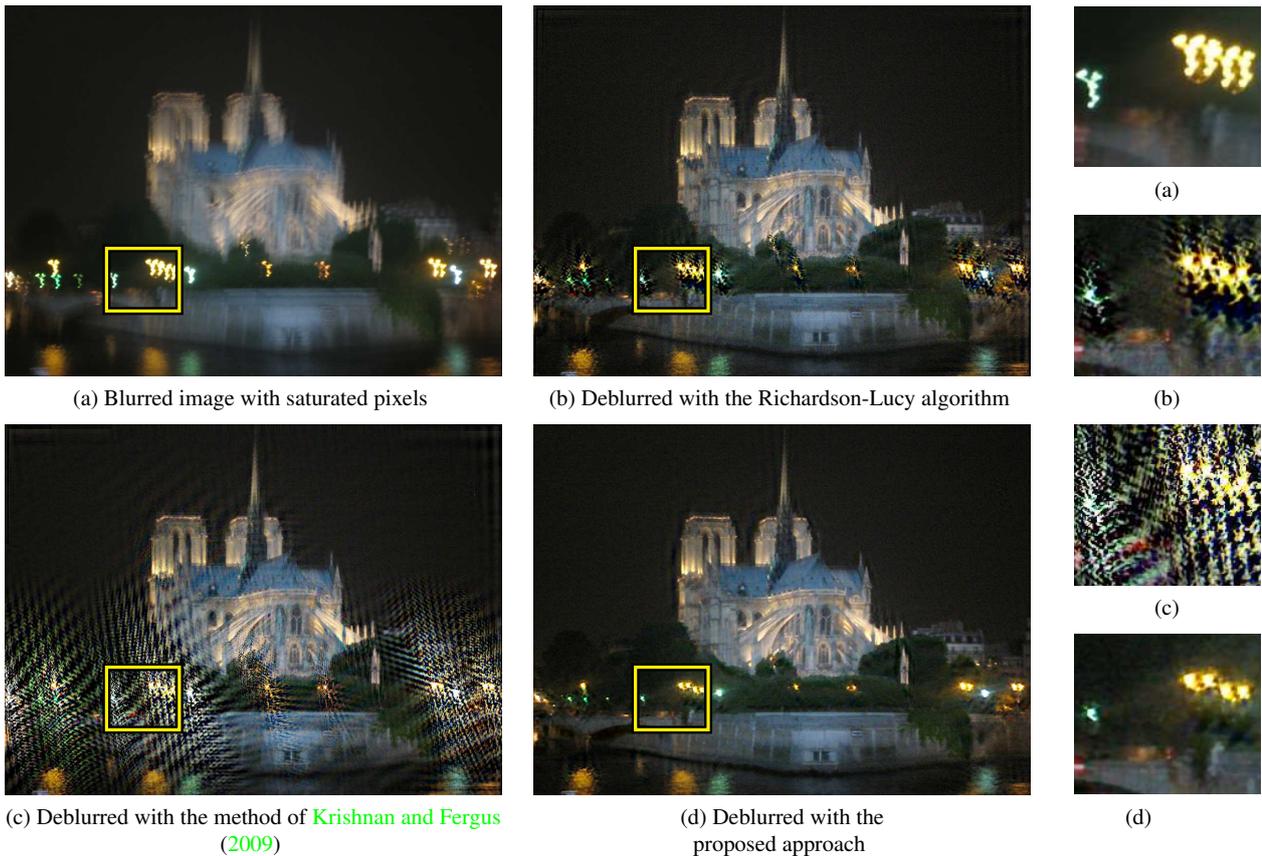


Fig. 1. **Deblurring in the presence of saturation.** Existing deblurring methods, such as those in (b) and (c), do not take account of saturated pixels. This leads to large and unsightly artefacts in the results, such as the “ringing” around the bright lights in the zoomed section. Using the proposed method (d), the ringing is greatly reduced and the quality of the deblurred image is improved. The PSF for this 1024×768 pixel image causes a blur of about 35 pixels in width, and was estimated directly from the blurred image using the algorithm described by Whyte et al. (2012).

modelled as following either a Poisson or Gaussian distribution, independent at each pixel.

For many causes of blur, the matrix \mathbf{A} can be parameterised by a small set of weights \mathbf{w} , often referred to as a *blur kernel*, such that

$$\mathbf{A} = \sum_k w_k \mathbf{T}_k, \quad (3)$$

where each $N \times N$ matrix \mathbf{T}_k applies some geometric transformation to the sharp image \mathbf{f} . Classically, \mathbf{T}_k have been chosen to model translations of the sharp image, allowing Equation (1) to be written as a 2D convolution of \mathbf{f} with \mathbf{w} . For blur caused by camera shake (motion of the camera during exposure), recent work (Gupta et al. 2010; Joshi et al. 2010; Whyte et al. 2010; Tai et al. 2011) has shown that using projective transformations for \mathbf{T}_k is more appropriate, and leads to more accurate modeling of the spatially-variant blur caused by camera shake. The remainder of this work is agnostic to the form of \mathbf{A} , and thus can be applied equally to spatially-variant and spatially-invariant blur.

Non-blind deblurring (where \mathbf{A} is known) is generally performed by attempting to solve a minimisation problem of the form

$$\min_{\mathbf{f}} \mathcal{L}(\mathbf{g}, \mathbf{A}\mathbf{f}) + \alpha\phi(\mathbf{f}), \quad (4)$$

where the *data-fidelity* term \mathcal{L} penalises sharp images that do not closely fit the observed data (i.e. \mathcal{L} is a measure of “distance” between \mathbf{g} and $\mathbf{A}\mathbf{f}$), and the *regularisation* term ϕ penalises sharp images that do not adhere to some prior model of sharp images. The scalar weight α balances the contributions of the two terms in the optimisation.

In a probabilistic setting, where the random noise ε is assumed to follow a known distribution, the data-fidelity term can be derived from the negative log-likelihood:

$$\mathcal{L}(\mathbf{g}, \mathbf{A}\mathbf{f}) = -\log p(\mathbf{g}|\mathbf{A}\mathbf{f}), \quad (5)$$

where $p(\mathbf{g}|\mathbf{A}\mathbf{f})$ denotes the probability of observing the blurry image \mathbf{g} , given a sharp image \mathbf{f} and PSF \mathbf{A} (often referred to as the likelihood). If the noise follows pixel-independent Gaussian distributions with uniform variance,

the appropriate data-fidelity term is

$$\mathcal{L}^G(\mathbf{g}, \mathbf{A}\mathbf{f}) = \sum_i (g_i - (\mathbf{A}\mathbf{f})_i)^2, \quad (6)$$

where $(\mathbf{A}\mathbf{f})_i$ indicates the i^{th} element of the vector $\mathbf{A}\mathbf{f}$. With Gaussian noise, Equation (4) is typically solved using standard linear least-squares algorithms, such as conjugate gradient descent (Levin et al. 2007). For the special case of spatially-invariant blur, and provided that the regularisation term ϕ can also be written as a quadratic function of \mathbf{f} , Equation (4) has a closed-form solution in the frequency domain, which can be computed efficiently using the fast Fourier transform (Wiener 1949; Gamelin 2001).

If the noise follows a Poisson distribution, the appropriate data-fidelity term is

$$\mathcal{L}^P(\mathbf{g}, \mathbf{A}\mathbf{f}) = - \sum_i (g_i \log(\mathbf{A}\mathbf{f})_i - (\mathbf{A}\mathbf{f})_i). \quad (7)$$

The classic algorithm for deblurring images with Poisson noise is the Richardson-Lucy algorithm (Richardson 1972; Lucy 1974), an iterative algorithm described by a simple multiplicative update equation. The incorporation of regularisation terms into this algorithm has been addressed by Tai et al. (2011) and Welk (2010). We discuss this algorithm further in Section 5.

A third data-fidelity term that is more robust to outliers than the two mentioned above, and which has been applied for image deblurring with impulse noise is the ℓ_1 norm (Bar et al. 2006; Yang et al. 2009) (corresponding to noise with a Laplacian distribution):

$$\mathcal{L}^L(\mathbf{g}, \mathbf{A}\mathbf{f}) = \sum_i |g_i - (\mathbf{A}\mathbf{f})_i|. \quad (8)$$

Although this data-fidelity term is more robust against noise values ε_i with large magnitudes, compared to the Gaussian or Poisson data-fidelity terms, it still produces artefacts in the presence of saturation (Cho et al. 2011).

For clarity, in the remainder of the paper we denote the data-fidelity term $\mathcal{L}(\mathbf{g}, \mathbf{A}\mathbf{f})$ simply as $\mathcal{L}(\mathbf{f})$, since we consider the blurred image \mathbf{g} and the PSF matrix \mathbf{A} to be fixed.

2.1 Sensor Saturation

Sensor saturation occurs when the radiance of the scene within a pixel exceeds the camera sensor’s range, at which point the sensor ceases to integrate the incident light, and produces an output that is clamped to the largest output value. This introduces a non-linearity into the image formation process that is not modelled by Equations (1) and (2). To correctly describe this effect, our model must include a non-linear function R , which reflects the sensor’s non-linear response to incident light. This function is applied to each

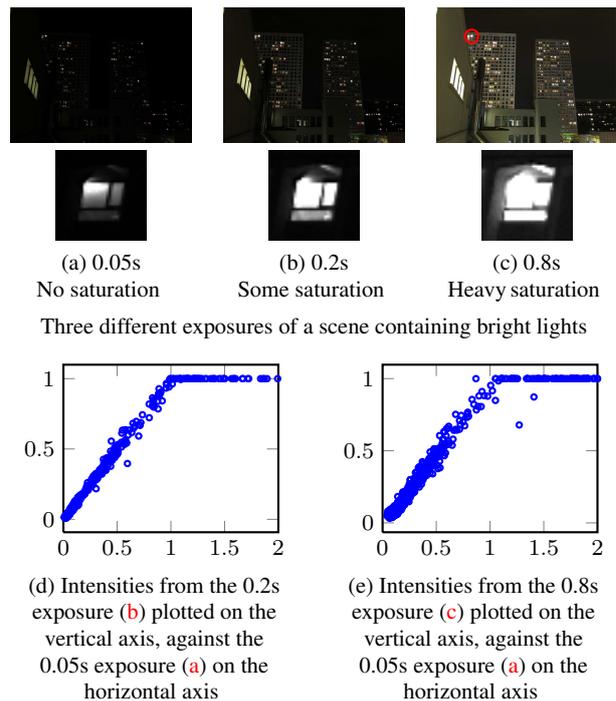


Fig. 2. **Saturated and unsaturated photos of the same scene.** (a)–(c) 3 different exposure times for the same scene, with bright regions that saturate in the longer exposures. A small window has been extracted, which is unsaturated at the shortest exposure, and increasingly saturated in the longer two. (d) Scatter plot of the intensities in the small window in (b) against those in the window in (a), normalised by exposure time. (e) Scatter plot of the intensities in the window in (c) against the window in (a), normalised by exposure time. The scatter plots in (d) and (e) clearly show the clipped linear relationship expected.

pixel of the image before it is output by the sensor, i.e.

$$g_i = R(g_i^* + \varepsilon_i), \quad (9)$$

where ε_i represents the random noise on pixel i .

For the purpose of describing saturation, we model the non-linear response function R as a truncated linear function, i.e. $R(x) = \min(x, 1)$, for intensities scaled to lie in the range $[0, 1]$. This model is supported by the data in Figure 2, which shows the relationship between pixel intensities in three different exposures of a bright light source. The pixel values in the short exposure (with no saturation) and the longer exposures (with saturation) clearly exhibit this clipped linear relationship. As the length of the exposure increases, more pixels saturate.

Due to the non-linearity in the image formation process, simply applying existing methods for non-blind deblurring (which assume a linear image formation model) to images affected by saturation, produces deblurred images which exhibit severe artefacts in the form of “ringing” – medium frequency ripples that spread across the image, e.g. in Figures 1b and 1c. Clearly, the saturation must be taken into

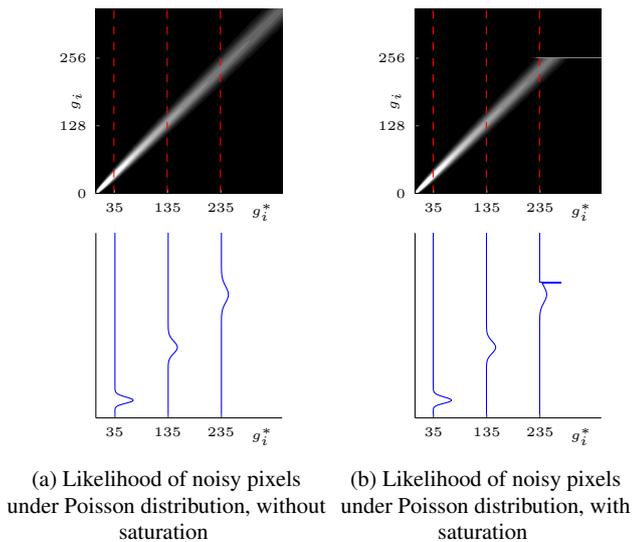


Fig. 3. **Noise distributions with and without sensor saturation.** The plots show the likelihood $p(g_i|g_i^*)$ of observed intensity g_i given the noiseless value g_i^* , under a Poisson noise model for the case with and without sensor saturation. The top row shows the 2D likelihood as an intensity-map, where black is zero and brighter is larger. The second row shows vertical slices for several values of g_i^* . Without saturation, the likelihood remains uni-modal around g_i^* even for bright pixels. With saturation (far-right plot), bright pixels have multi-modal likelihoods (one mode at g_i^* , and a second at 255). This makes the inversion of such a forward model particularly difficult.

account during non-blind deblurring in order to avoid such artefacts.

Given the non-linear forward model in Equation (9), it is tempting to modify the data-fidelity term \mathcal{L} to take into account saturation, and thus prevent artefacts from appearing in the deblurred image. The model in Equation (9), however, is not tractable to invert. Since the noise term ϵ_i lies inside R , the likelihood $p(g_i|g_i^*)$ (from which \mathcal{L} is derived) is distorted and, in general, can no longer be written in closed-form. Figure 3 shows the difference between the likelihoods with and without saturation for Poisson noise. In the saturated likelihood, pixels near the top of the camera’s range have distributions that are no longer smooth and uni-modal, but instead have a second sharp peak at 1. Furthermore, for some pixels this second mode at 1 is the maximum of the likelihood, i.e. $P(g_i = 1|g_i^*) > P(g_i = g_i^*|g_i^*)$, which clearly contradicts the normal behaviour of Poisson or Gaussian noise, where the likelihood is smooth and has a single mode at $g_i = g_i^*$.

Given the difficulty of inverting the true non-linear forward model, alternative approaches to handling saturated pixels are needed, in order to prevent ringing artefacts from appearing. Before discussing our approach to this, we provide some analysis of ringing in general.

3 Ringing

Ringing is common in deblurred images, and has been attributed to various causes, including the Gibbs phenomenon (Yuan et al. 2007) (arising from the inability of a finite Fourier basis to reconstruct perfect step edges), incorrectly-modelled image noise (i.e. using the incorrect likelihood for the data-fidelity term \mathcal{L} in Equation (4)), or errors in the estimated PSF (Shan et al. 2008). In the following we show that the root cause of ringing is the fact that, in general, blur annihilates certain spatial frequencies in the image. These spatial frequencies are very poorly constrained by the observed data (the blurred image) and can become artificially amplified in the deblurred image. Incorrectly-modelled noise and PSF errors are two causes of such amplification.

3.1 What does ringing look like?

For a discrete PSF given by a matrix \mathbf{A} , there is typically some set of vectors that are almost completely suppressed by \mathbf{A} . These are vectors that lie in, or close to, the nullspace of \mathbf{A} , i.e. $\{\mathbf{x} : \mathbf{A}\mathbf{x} \simeq \mathbf{0}\}$. When we attempt to invert the blur process by estimating the sharp image $\hat{\mathbf{f}}$ that minimises $\mathcal{L}(\hat{\mathbf{f}})$, these directions in the solution space are very poorly constrained by the data, since $\mathbf{A}(\hat{\mathbf{f}} + \lambda\mathbf{x}) \simeq \mathbf{A}\hat{\mathbf{f}}$, where λ is a scalar. As such, they can become artificially amplified in the deblurred image. These amplified components \mathbf{x} , lying close to the PSF’s nullspace, cause visible artefacts in the deblurred image.

The reason that poorly-constrained directions appear as ringing, and not as some other kind of visual artefact, is that the nullspace of a PSF tends to be spanned by vectors having small support in the frequency domain, but large spatial support. This can be seen from a spectral decomposition of the PSF. In the simplest case, where \mathbf{A} corresponds to a convolution with cyclic boundary conditions, its eigendecomposition can be obtained by the discrete Fourier transform. Thus, its eigenvectors are sinusoids, and its nullspace is spanned by a set of these sinusoids with eigenvalues close to zero. When the cyclic boundary conditions are removed, or the PSF is spatially-variant, e.g. due to camera shake, the exact correspondence with the Fourier transform and sinusoids no longer holds, however the nullspace is still spanned by vectors resembling spatial frequencies, as shown in Figure 4. These are vectors with a small support in the frequency domain but a large spatial support.

A concrete example is shown in Figure 5, which shows a blurred image with outliers added, deblurred using the Richardson-Lucy algorithm. Looking at the ringing artefacts in both the spatial domain and the frequency domain, it is clear that ringing occurs at frequencies that are poorly constrained by the blur kernel (frequencies at which the magni-

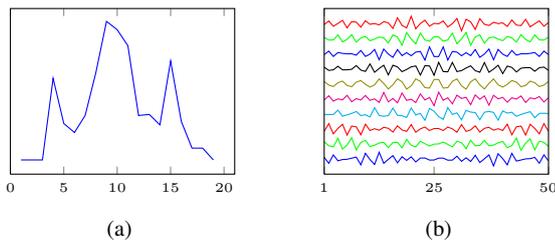


Fig. 4. **Poorly-constrained frequencies of a 1D blur kernel.** (a) A blur kernel. (b) We construct the 50×50 matrix \mathbf{A} that convolves a 50-sample signal with (a) (with zero-padding boundary conditions), and plot the 10 singular vectors with smallest singular values. The singular vectors have large spatial support, and correspond closely to spatial frequencies and their harmonics.

tude of the discrete Fourier transform (DFT) of the kernel is small).

3.2 How does ringing emerge?

Assuming that our estimate of the PSF contains no gross errors, the amplification of poorly-constrained directions in the solution may occur in two major ways. First, as mentioned above, there are spatial frequencies in the sharp image that are almost completely annihilated by \mathbf{A} . During non-blind deblurring, any such spatial frequency \mathbf{x} in the solution will be essentially unconstrained, since $\mathcal{L}(\mathbf{f} + \lambda\mathbf{x}) \simeq \mathcal{L}(\mathbf{f})$. If anything occurs during deblurring to cause λ to be non-zero, there is no force in the optimisation to reduce it again, and the deblurred result will be the superposition of the sharp image with some combination of sinusoids \mathbf{x} .

Figures 6g and 6h demonstrate this with a synthetic example, where the deblurred result depends highly on the initialisation. With some initialisations, the deblurred image contains ringing, while with other initialisations, no ringing appears. This indicates that the ringing components of the solution are indeed poorly constrained, but are not actively amplified by the data-fidelity cost. In Section 4 we will discuss the cause of the amplification in saturated images such as these.

The second, and harder to tackle, way that ringing can arise is if there are some unmodelled / outlier pixels in the blurred image \mathbf{g} , e.g. due to saturation or impulse noise. In this case, there may exist some \mathbf{x} such that $\mathcal{L}(\mathbf{f}^* + \lambda\mathbf{x}) < \mathcal{L}(\mathbf{f}^*)$, where \mathbf{f}^* denotes the true sharp image. This is possible because, by their very nature, the outliers cannot be well-explained by the true sharp image \mathbf{f}^* . In order for this to happen, the addition of $\lambda\mathbf{x}$ must (a) decrease the data-fidelity cost \mathcal{L} for the outlier pixels, while (b) not significantly increasing it for the remainder of the image. Clearly, to satisfy (a), \mathbf{x} cannot lie exactly in the nullspace of \mathbf{A} , otherwise $\mathbf{A}\mathbf{x} = \mathbf{0}$, however to satisfy (b), it also cannot lie too far from the nullspace, otherwise the data-fidelity cost

would grow quickly with λ . The result is that \mathbf{x} lies close to, but not in, the nullspace of \mathbf{A} , and the optimisation is more-or-less free to change λ in order to explain the outliers more accurately.

Figure 6f shows an example of this, where for all initialisations, some ringing appears, and the deblurred image $\hat{\mathbf{f}}$ has a lower data-fidelity cost than the true sharp image \mathbf{f}^* . Note that the ringing in Figure 6f is visually similar to that in Figures 6g and 6h, underlining the fact that in both cases, ringing appears in the poorly-constrained frequencies of the result.

One additional way that ringing may emerge in a deblurred image, which we do not address in this work, is if the PSF is incorrectly estimated. In this case, the PSF used to deblur is different from the one which caused the blur, and the deblurred image will be incorrect due to this discrepancy. We do not go into detail here, but for an estimated PSF $\hat{\mathbf{A}}$ that is related to the true PSF \mathbf{A} by $\hat{\mathbf{A}} = \mathbf{A} + \Delta\mathbf{A}$, the deblurred image will contain ringing in the spatial frequencies where $\hat{\mathbf{A}}$ has a small frequency response (i.e. close to the null-space of $\hat{\mathbf{A}}$) (Whyte 2012). This echoes the conclusions of the previous paragraphs, with the difference that the ringing frequencies are determined by the incorrect PSF $\hat{\mathbf{A}}$, instead of the true PSF \mathbf{A} . The problem of deblurring with an erroneous PSF has also been addressed recently by Ji and Wang (2012), who introduce additional latent variables in order to estimate \mathbf{f} in a manner that is invariant to a certain class of PSF errors.

3.3 Why doesn't regularisation suppress ringing?

Often, ringing appears in deblurred images despite the inclusion of some regularisation term ϕ in the optimisation problem in Equation (4). This seems counterintuitive, since the purpose of the regularisation is to provide some additional constraints on the solution, particularly in directions which are poorly constrained by the data. However, most popular forms of regularisation used in non-blind deblurring penalise only the magnitudes of the first or second spatial derivatives of the deblurred image, e.g.

$$\phi(\mathbf{f}) = \sum_i \rho(|(\mathbf{d}_x * \mathbf{f})_i|) + \rho(|(\mathbf{d}_y * \mathbf{f})_i|), \quad (10)$$

where the sum is taken over all pixels i in the image, the filters \mathbf{d}_x and \mathbf{d}_y compute derivatives, and ρ is a non-decreasing scalar function (Blake and Zisserman 1987; Bouman and Sauer 1993; Schultz and Stevenson 1994; Levin et al. 2007). These derivatives are computed from the differences between neighbouring pixels, and are essentially high-pass filters. Unsurprisingly, such regularisation works best at suppressing high spatial frequencies. At medium-to-low frequencies, the power of the regularisation decreases.

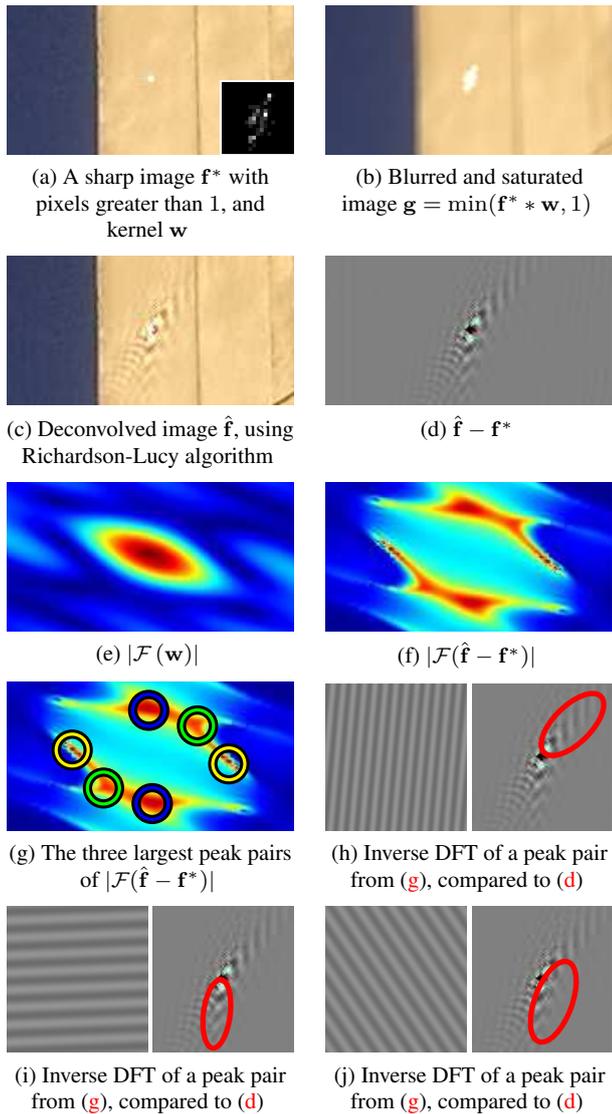


Fig. 5. **An example of ringing in the frequency domain.** A sharp image \mathbf{f}^* containing several pixels with values greater than 1 is convolved with a kernel \mathbf{w} (a), and clipped at 1 to produce a blurred, saturated image \mathbf{g} (b). When we deconvolve \mathbf{g} (c), ringing is produced (d). We can take the discrete Fourier transform (DFT) of the kernel $\mathcal{F}(\mathbf{w})$ (e), and the ringing artefacts $\mathcal{F}(\hat{\mathbf{f}} - \mathbf{f}^*)$ (f), and compare their magnitudes. In these plots, large values are red, while small values are blue. It is visible from these plots that the ringing frequencies with largest magnitude are those frequencies where the kernel has the smallest magnitude (i.e. the red regions in (f) correspond to dark blue regions in (e)). By isolating the largest peaks in the ringing spectrum $|\mathcal{F}(\hat{\mathbf{f}} - \mathbf{f}^*)|$ (g), and visualizing the spatial frequencies that they represent (h)–(j), we can see that they do indeed correspond to the ringing that is visible in the deblurred image.

Figure 7 demonstrates this, showing how the power of the regularisation is greatest at high frequencies, and falls to zero at low frequencies. The blurred data on the other hand constrains the lowest frequencies (i.e. at scales larger than the blur), but not the high frequencies. There is a region of medium frequencies that are poorly constrained by both. It is at these intermediate frequencies that some ringing \mathbf{x} can easily appear, since $\mathcal{L}(\mathbf{f}^* + \lambda\mathbf{x}) \simeq \mathcal{L}(\mathbf{f}^*)$, and $\phi(\mathbf{f}^* + \lambda\mathbf{f}) \simeq \phi(\mathbf{f}^*)$. Although the regularisation weight α in Equation (4) can be increased in order to reduce ringing, this will also begin to over-smooth the deblurred image.

Yuan et al. (2008) observe this fact and propose a multi-scale non-blind deblurring algorithm capable of preventing ringing caused by noise and numerical inaccuracies. However, although their method can suppress ringing at a wide range of frequencies, it is still generally unable to handle ringing caused by saturation, as shown in Figure 14.

4 Preventing Ringing when Deblurring Saturated Images

When applying existing non-blind deblurring algorithms to saturated images, which assume the linear model $\mathbf{g} = \mathbf{A}\mathbf{f} + \epsilon$, using Gaussian or Poisson data-fidelity terms \mathcal{L} , ringing is almost certain to appear. As discussed in the previous section, saturated pixels can cause the phenomenon where a deblurred image $\hat{\mathbf{f}}$ containing ringing actually has a lower data-fidelity cost than the true sharp image \mathbf{f}^* , due to the fact that the assumption of linearity is violated.

Although the saturation can be modelled with a non-linear response R (as discussed in Section 2.1) in order to prevent the case where $\mathcal{L}(\mathbf{f}^* + \lambda\mathbf{x}) < \mathcal{L}(\mathbf{f}^*)$, the model is not tractable to invert. For this reason, some recent authors have instead chosen to treat saturated pixels simply as outliers, and model the rest of the image as linear (Harmeling et al. 2010b; Cho et al. 2011). The saturated pixels are discarded and treated as missing data. Not only is this much more tractable to optimise, it arguably does not sacrifice much – saturated pixels are largely uninformative, due to being clamped at the maximum output value.

In order to perform non-blind deblurring with missing data, we can define a mask \mathbf{m} of binary weights, where $m_i = 0$ if pixel i is missing, and $m_i = 1$ otherwise. We then construct a weighted form of the data-fidelity term, using \mathbf{m} as the weights, and denote this $\mathcal{L}_{\mathbf{m}}$. For example, for Poisson noise, Equation (7) becomes:

$$\mathcal{L}_{\mathbf{m}}^{\text{P}}(\mathbf{f}) = - \sum_i m_i (g_i \log(\mathbf{A}\mathbf{f})_i - (\mathbf{A}\mathbf{f})_i). \quad (11)$$

By removing all the data that does not follow the linear model, it should no longer be possible for $\mathcal{L}_{\mathbf{m}}(\mathbf{f}^* + \lambda\mathbf{x}) < \mathcal{L}_{\mathbf{m}}(\mathbf{f}^*)$.

Harmeling et al. (2010b) estimate the mask \mathbf{m} directly from the blurred image, by finding all the pixels in \mathbf{g} that are above some threshold φ close to 1, i.e.

$$m_i = \begin{cases} 1 & \text{if } g_i < \varphi \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

On the other hand, Cho et al. (2011) estimate \mathbf{m} by blurring the current estimate of the sharp image $\hat{\mathbf{f}}$:

$$m_i = \begin{cases} 1 & \text{if } (\mathbf{A}\hat{\mathbf{f}})_i < 1 \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

Note that although some of the blurred pixels are discarded, the entire latent image \mathbf{f} is still estimated.

Although detecting and ignoring the (outlier) saturated pixels can reduce ringing in the deblurred images, it does not necessarily remove it entirely. By discarding saturated data, the data-fidelity term $\mathcal{L}_{\mathbf{m}}$ no longer actively *encourages* ringing. However, it may still be possible that $\mathcal{L}_{\mathbf{m}}(\mathbf{f}^* + \lambda\mathbf{x}) \simeq \mathcal{L}_{\mathbf{m}}(\mathbf{f}^*)$ for some ringing \mathbf{x} . As discussed in the previous section, if anything in the data, the initialisation, or in the deblurring algorithm serves to increase λ , there is no force to reduce it again, and the ringing will appear in the final result. Figure 6g demonstrates this with a synthetic example. Even when the outliers are known exactly, and are completely removed from the deblurring process, ringing can still appear. On the other hand, when we initialise the deblurring algorithm with the true sharp image, no ringing appears, indicating that there is nothing about $\mathcal{L}_{\mathbf{m}}$ that *causes* ringing, there is simply nothing to prevent or suppress it.

From this discussion, we conclude that in order to deblur saturated images without introducing ringing, we must do more than simply remove the saturated pixels from the observed image \mathbf{g} . We must *also* avoid amplifying the poorly-constrained spatial frequencies in the latent image \mathbf{f} . To achieve this, we could either introduce some form of regularisation to place additional constraints on \mathbf{f} , such as in Yuan et al. (2008), or we could directly modify the data-fidelity term. In this work we choose the latter approach, eschewing the use of regularisation to prevent ringing (although we revisit regularisation later in order to reduce other sources of noise in the deblurred results).

In this work, we posit that for the case of saturation, the main factor causing poorly-constrained spatial frequencies to become amplified is that there exist pixels in the *sharp* image that exceed the image’s range, i.e. $\exists j : f_j > 1$. During non-blind deblurring, we must estimate these “bright” pixels’ values, and it is the act of estimating these values that destabilises our estimates of other pixels. Note that removing the saturated pixels from \mathbf{g} does not, in general, remove the influence of these “bright” latent pixels from the blurred image; a latent pixel with intensity greater than 1 can contribute to a blurred pixel whose intensity is less than 1.

Indeed, a blurred image with no saturation at all may still correspond to a latent image containing pixels brighter than 1.

When we attempt to estimate a “bright” pixel’s value f_j , we are likely to make a relatively inaccurate estimate. The main reason for this is that the set of observations (blurred pixels) concerning f_j is likely to be incomplete. In the blurring process, each latent pixel in the sharp image is spread across multiple pixels in the blurred image \mathbf{g} . Given that $f_j > 1$, there is a good chance that some of these blurred pixels will be saturated, and thus uninformative as to the value of f_j . With fewer observations from which to estimate f_j , its value will be more susceptible to noise, and contain larger error than if we had a full set of unsaturated observations available. This is supported by Figure 6g, where the sharp image contains pixels with intensity greater than 1, and the appearance of ringing depends on the initialisation. When we initialise with the true sharp image, which already contains the bright (> 1) pixels, no ringing is caused. When we initialise with the blurred image, or random values in $[0, 1]$, the algorithm is forced to attempt to estimate the bright values, and in doing so causes ringing.

Since this amplification is caused when we attempt to estimate the “bright” pixels in \mathbf{f} , our approach to preventing ringing is based on the idea that we should avoid (or rather, delay) estimating such bright pixels. If we only estimate pixels that can be accurately estimated, we will not make gross errors, which in turn will avoid introducing ringing. For the synthetic example, Figure 6h shows the result of this approach. By refusing to estimate the bright (> 1) pixels in \mathbf{f} , the ringing is almost removed for all initialisations. In the next section we discuss how we do this in practice.

Note that this idea is similar to the notion of “Highest Confidence First” (Chou and Brown 1990). Chou and Brown point out that due to the coupling of pixels in a Markov Random Field image model, the estimate for a pixel with strong observations may be negatively impacted by a poor decision at a neighbouring pixel with weak observations. In such cases, the pixel with strong observations “can do better without the incorrect information of a neighbor.”

5 Proposed Method

In this section we describe our proposed algorithm for deblurring images containing saturated pixels. We begin by describing our approach to estimating the latent image without introducing ringing, motivated by the discussion in the previous sections. In addition, we propose a method for estimating the bright pixels separately, without introducing artefacts, and finally describe how these elements are combined into the complete proposed method.

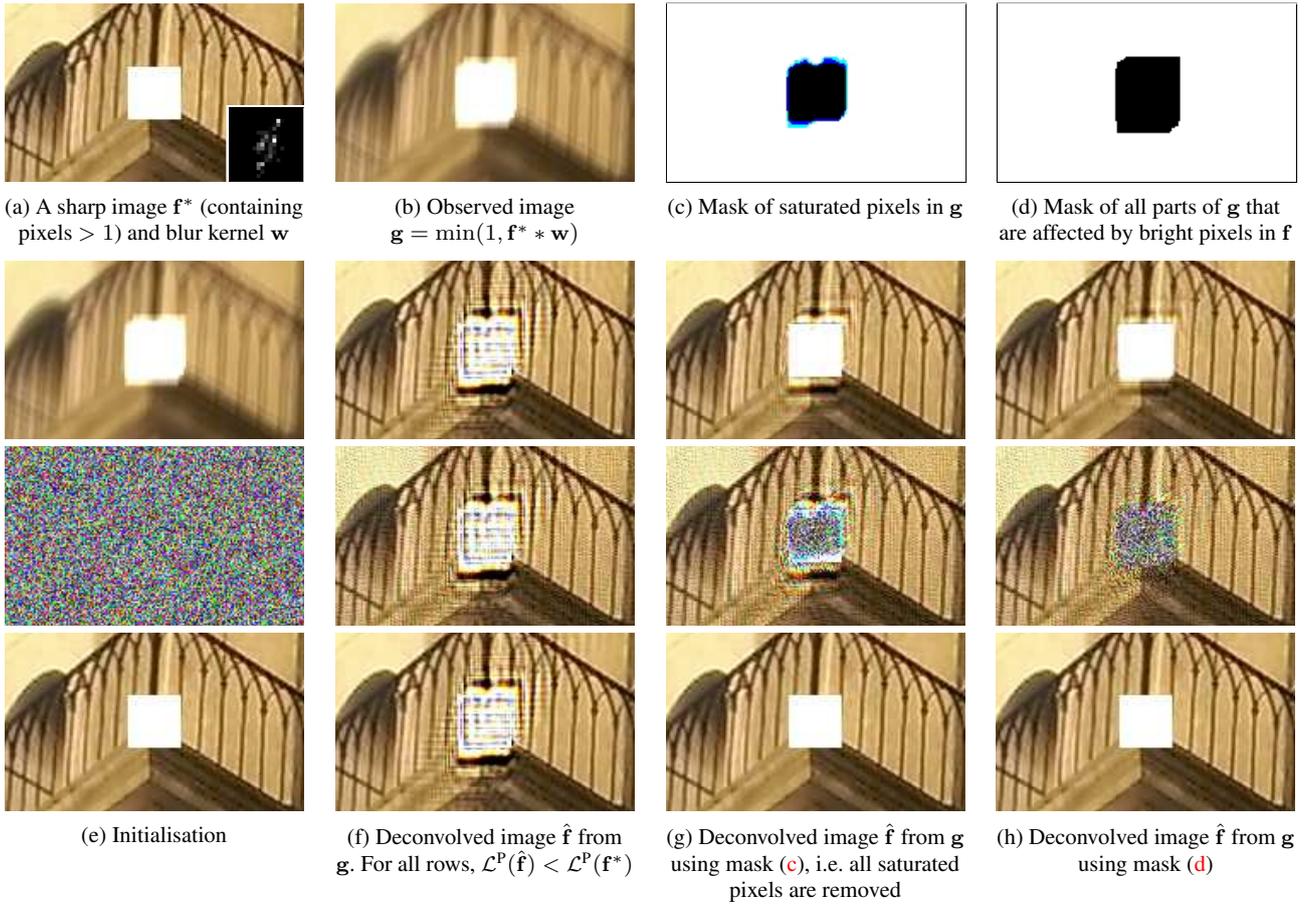


Fig. 6. **Synthetic example of deblurring with saturated pixels.** (a) A sharp image f^* containing some bright pixels with intensity twice the maximum image value. (b) The sharp image f is convolved with w and clipped at 1 to simulate saturation. No noise is added in this example. (c) The mask of pixels in g that are not saturated, i.e. $m_i = 1$ (white) if $(f^* * w)_i < 1$, and $m_i = 0$ (black) otherwise. (d) The mask of pixels in g that are not influenced by the bright pixels in f^* , i.e. $m_i = 1$ if $((f^* \geq 1) * w)_i = 0$. See Section 5.1 for further explanation. The following rows show the results of deblurring g with 1000 iterations of the Richardson-Lucy algorithm, starting from three different initialisations. (1st row) initialised with blurred image. (2nd row) initialised with random values in $[0, 1]$. (3rd row) initialised with true sharp image. As can be seen, in (f), ringing appears regardless of the initialisation, indicating that \mathcal{L} is encouraging this to occur. In (g), ringing appears with some initialisations but not others, indicating that \mathcal{L} does not encourage it, but does not suppress it either. In (h), ringing is almost gone, since we have removed the most destabilising data. Counterintuitively, although we have discarded more data, we end up with less ringing. Note that by removing all the blurred pixels shown in (d), we have no information about the bright pixels in f , and they simply retain their initial values. Note also that the deblurred images \hat{f} may contain pixel values greater than 1, however they are clamped to 1 before writing the image to file

5.1 Preventing ringing by refusing to make bad estimates

As we have seen, it is clear that when we attempt to perform non-blind deblurring on an image containing saturation, we will incorrectly estimate the “bright” latent pixels (with values close to or greater than 1) that caused the saturation. As discussed in the previous section, the errors we make at these bright pixels will cause ringing. In this section we propose a method that iteratively attempts to classify these latent pixels, and remove them from the estimation process, thereby mitigating their effect on the rest of the image.

Assume for now that we already know which pixels in the latent image are bright, and will thus be poorly estimated. We denote this set of pixels \mathcal{S} , and denote its set

complement (containing the rest of the image, which can be estimated accurately) by \mathcal{U} . We can write the latent image in terms of these two disjoint sets: $f = f_{\mathcal{U}} + f_{\mathcal{S}}$.

Given that we are unable to estimate $f_{\mathcal{S}}$ accurately, our aim is to prevent the propagation of errors from $f_{\mathcal{S}}$ to $f_{\mathcal{U}}$. To achieve this, we decouple our estimate of $f_{\mathcal{U}}$ from our estimate of $f_{\mathcal{S}}$. First, note that we can decompose the noiseless blurred image as:

$$g^* = \mathbf{A}f_{\mathcal{U}} + \mathbf{A}f_{\mathcal{S}}. \quad (14)$$

We denote by \mathcal{V} the set of blurred pixels that are not affected by $f_{\mathcal{S}}$, i.e. $\mathcal{V} = \{i | (\mathbf{A}f_{\mathcal{S}})_i = 0\}$, and construct the corresponding binary mask \mathbf{v} (where $v_i = 1$ if $i \in \mathcal{V}$, and $v_i = 0$ otherwise). Then, given that $\mathbf{v} \circ \mathbf{A}f_{\mathcal{S}} = \mathbf{0}$, we can

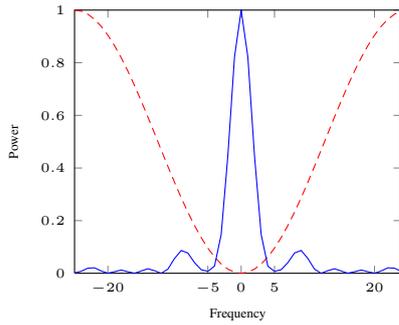


Fig. 7. **Power spectra of blur and gradient filters.** The power spectrum of the 1D blur kernel in Figure 4 (solid blue line), and the power spectrum of the derivative filter $[1, -1]$ (dashed red line), often used for regularisation in non-blind deblurring. The blur kernel has minima in its power spectrum, which correspond to poorly-constrained frequencies in the deblurred solution. At the low-frequency minima (e.g. near 5 cycles/image), the gradient filter has also lost most of its power, and so ringing at these frequencies is unlikely to be suppressed by gradient-based regularisation

write

$$\mathbf{v} \circ \mathbf{g}^* = \mathbf{v} \circ \mathbf{A}\mathbf{f}_{\mathcal{U}} \quad (15)$$

where $\cdot \circ \cdot$ represents the element-wise product between two vectors. From this, we can estimate $\mathbf{f}_{\mathcal{U}}$ independently of $\mathbf{f}_{\mathcal{S}}$. Note that we can obtain the mask \mathbf{v} simply by constructing the binary mask \mathbf{u} that corresponds to the set \mathcal{U} , and performing a binary erosion of \mathbf{u} with the non-zeros of the PSF.

Furthermore, given that the set \mathcal{U} does not contain any bright latent pixels, $\mathbf{A}\mathbf{f}_{\mathcal{U}}$ should not cause any saturation. Thus, we can estimate $\mathbf{f}_{\mathcal{U}}$ without modelling the non-linearity caused by saturation, and write the observed blurred image, including noise, as

$$\mathbf{v} \circ \mathbf{g} = \mathbf{v} \circ \mathbf{A}\mathbf{f}_{\mathcal{U}} + \mathbf{v} \circ \boldsymbol{\varepsilon}. \quad (16)$$

To estimate $\mathbf{f}_{\mathcal{U}}$ from $\mathbf{v} \circ \mathbf{g}$, we can adapt existing non-blind deblurring algorithms to handle missing data, as seen in Section 4. For Poisson noise, the data-fidelity term is simply weighted, as in Equation (11), using \mathbf{v} as the weights to form $\mathcal{L}_{\mathbf{v}}^{\text{P}}$. The classical algorithm for non-blind deblurring with Poisson noise is the Richardson-Lucy algorithm, which is described by the multiplicative update equation:

$$\hat{\mathbf{f}}^{t+1} = \hat{\mathbf{f}}^t \circ \mathbf{A}^{\top} \left(\frac{\mathbf{g}}{\mathbf{A}\hat{\mathbf{f}}^t} \right), \quad (17)$$

where the division is performed element-wise.

We can adapt the Richardson-Lucy update equation from Equation (17) to minimise the binary-weighted data-fidelity term $\mathcal{L}_{\mathbf{v}}^{\text{P}}$, instead of the homogeneous \mathcal{L}^{P} , leading to the fol-

lowing update equation for $\mathbf{f}_{\mathcal{U}}$:

$$\hat{\mathbf{f}}_{\mathcal{U}}^{t+1} = \hat{\mathbf{f}}_{\mathcal{U}}^t \circ \mathbf{A}^{\top} \left(\frac{\mathbf{g} \circ \mathbf{v}}{\mathbf{A}\hat{\mathbf{f}}_{\mathcal{U}}^t} + \mathbf{1} - \mathbf{v} \right). \quad (18)$$

Note that in order to avoid division by zero, we add a small positive constant to the denominator of the fraction.

This update equation is applicable to any kind of missing data in \mathbf{g} , the important distinction is in how we determine \mathbf{v} , the mask of missing pixels. In Figure 6h, we show the result of deblurring with \mathbf{v} determined as described above, by eroding \mathbf{u} . For comparison, Figure 6g shows the result when \mathbf{v} is determined simply by finding saturated pixels in \mathbf{g} , i.e. $v_i = 1$ if $g_i < 1$. Despite removing all the saturated pixels, the results in Figure 6g still contain ringing, while significantly less ringing appears in Figure 6h using the proposed approach.

5.2 Preventing artefacts inside bright regions

While the update method in the previous section is effective at preventing ringing from propagating outwards from bright regions into other parts of the image, we still wish to estimate values for the pixels in \mathcal{S} . Since these pixels are bright, and unlikely to be estimated accurately, we are not concerned with preventing propagation of information from \mathcal{U} to \mathcal{S} , and choose to update them using all the available data.

Using the standard Richardson-Lucy update for these bright regions however, can cause dark artefacts to appear, due to the fact that the linear model cannot explain the saturated data. Such artefacts are visible in Figure 10d. To prevent these artefacts, we propose a second modification of the Richardson-Lucy algorithm, to include a non-linear response function. Since the true forward model, discussed in Section 2.1 is not tractable to invert, we propose the following tractable but approximate alternative, which is sufficient for preventing artefacts in bright regions:

$$g_i = R(g_i^*) + \varepsilon_i, \quad (19)$$

where now the response R function comes before the noise. Re-deriving the Richardson-Lucy algorithm with this model gives the following update equation for the pixels in \mathcal{S} :

$$\hat{\mathbf{f}}_{\mathcal{S}}^{t+1} = \hat{\mathbf{f}}_{\mathcal{S}}^t \circ \mathbf{A}^{\top} \left(\frac{\mathbf{g} \circ R'(\mathbf{A}\hat{\mathbf{f}}^t)}{R(\mathbf{A}\hat{\mathbf{f}}^t)} + \mathbf{1} - R'(\mathbf{A}\hat{\mathbf{f}}^t) \right), \quad (20)$$

where $R'(\cdot)$ indicates the derivative of the response function R .

Since the ideal response function $R(x) = \min(x, 1)$ is not differentiable, we use a smooth, continuously differentiable approximation to this function (Chen and Mangasar-

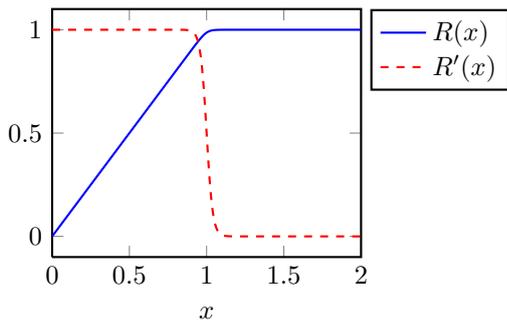


Fig. 8. **Modelling the saturated sensor response.** Smooth and differentiable approximation to the non-differentiable function $\min(x, 1)$ used to model sensor saturation, defined in Equation (21). The derivative is also smooth and defined everywhere.

ian 1996), where

$$R(x) = x - \frac{1}{a} \log\left(1 + \exp(a(x - 1))\right) \quad (21)$$

$$R'(x) = \frac{1}{1 + \exp(a(x - 1))}. \quad (22)$$

The parameter a controls the smoothness of the approximation, and we have found $a = 50$ to be a suitable compromise between smoothness and accuracy (we use this value in all results presented in this paper). Figure 8 shows the shape of these smooth versions of R and R' .

Equation (20) can be roughly interpreted as weighting the blurry pixels according to the value of R' : in the linear (unsaturated) portion where $x < 1$, $R(x) \simeq x$ and $R'(x) \simeq 1$, so that the term in parentheses is the same as for the standard RL algorithm. In the saturated portion where $x > 1$, $R(x) \simeq 1$ and $R'(x) \simeq 0$, so that the term in parentheses is equal to unity and has no influence on \mathbf{f} . We note that this behaviour is very similar to the method used by Cho et al. (2011) to handle saturation – given the current estimate $\hat{\mathbf{f}}$ of the sharp image, they compute the blurred estimate $\mathbf{A}\hat{\mathbf{f}}$, and ignore any blurry pixel i for which $(\mathbf{A}\hat{\mathbf{f}})_i > 1$.

5.3 Segmenting the latent image

So far in this section, we have assumed that \mathcal{S} and \mathcal{U} are known. Given a real blurred image however, we do not know *a priori* which parts of \mathbf{f} belong in \mathcal{U} and which in \mathcal{S} . We thus treat \mathcal{U} as another latent variable, in addition to the latent image \mathbf{f} , and perform the segmentation at each iteration t . Given the discussion in previous sections, we adopt a simple segmentation process, which consists of thresholding the *latent image* at some level φ close to 1:

$$\mathcal{U} = \{j | \hat{f}_j^t \leq \varphi\}. \quad (23)$$

We decompose $\hat{\mathbf{f}}^t$ according to

$$\hat{\mathbf{f}}_{\mathcal{U}}^t = \mathbf{u} \circ \hat{\mathbf{f}}^t \quad (24)$$

$$\hat{\mathbf{f}}_{\mathcal{S}}^t = \hat{\mathbf{f}}^t - \hat{\mathbf{f}}_{\mathcal{U}}^t, \quad (25)$$

where \mathbf{u} is the binary mask corresponding to \mathcal{U} . Since our aim is to ensure that no large errors are introduced in $\mathbf{f}_{\mathcal{U}}$, we set the threshold low enough that most potentially-bright pixels are assigned to \mathcal{S} . Empirically, we choose $\varphi = 0.9$ for the results in this paper, although we have not found the results to be particularly sensitive to the exact value of φ .

5.4 Adding regularisation

Although the focus of this work is mitigating ringing by modifying the data-fidelity term, rather than by designing new forms of regularisation, it may nonetheless still be useful to apply some form of regularisation to reduce other noise throughout the deblurred image. As discussed by Tai et al. (2011) and Welk (2010) it is possible to include a regularisation term in the Richardson-Lucy algorithm, and this remains true for our algorithm. In this work we include the ℓ_1 -norm of the gradients of the deblurred image as regularisation. Using the form of Equation (10), this is written

$$\phi(\mathbf{f}) = \sum_i |(\mathbf{d}_x * \mathbf{f})_i| + |(\mathbf{d}_y * \mathbf{f})_i|, \quad (26)$$

where the filters \mathbf{d}_x and \mathbf{d}_y compute derivatives.

We incorporate this in the manner described by Tai et al. (2011). Denoting the unregularised update of \mathbf{f} as $\hat{\mathbf{f}}_{\text{unreg}}^{t+1}$ (computed as in Equations (18) and (20) of the previous sections), we compute the update for the regularised problem as

$$\hat{\mathbf{f}}^{t+1} = \frac{\hat{\mathbf{f}}_{\text{unreg}}^{t+1}}{1 + \alpha \nabla \phi(\hat{\mathbf{f}}^t)}, \quad (27)$$

where $\nabla \phi(\hat{\mathbf{f}}^t)$ is the vector of partial derivatives of the regularisation¹ ϕ with respect to each pixel of the latent image \mathbf{f} , evaluated at $\hat{\mathbf{f}}^t$, i.e.

$$(\nabla \phi(\hat{\mathbf{f}}^t))_j = \left. \frac{\partial \phi(\mathbf{f})}{\partial f_j} \right|_{\hat{\mathbf{f}}^t}. \quad (28)$$

In Figure 9 we compare results obtained using our method with and without this regularisation. As can be seen, the ringing is largely suppressed by the steps described in the preceding sections. The addition of the regularisation term further improves the results by reducing noise elsewhere in the deblurred result.

We summarise our complete proposed algorithm in Algorithm 1. Figure 10 shows the contributions of the two pro-

¹ We note that the regulariser ϕ in Equation (26) is not continuously differentiable, however this can be avoided by using the standard approximation $|x| \simeq \sqrt{\epsilon + x^2}$, for some small number ϵ .

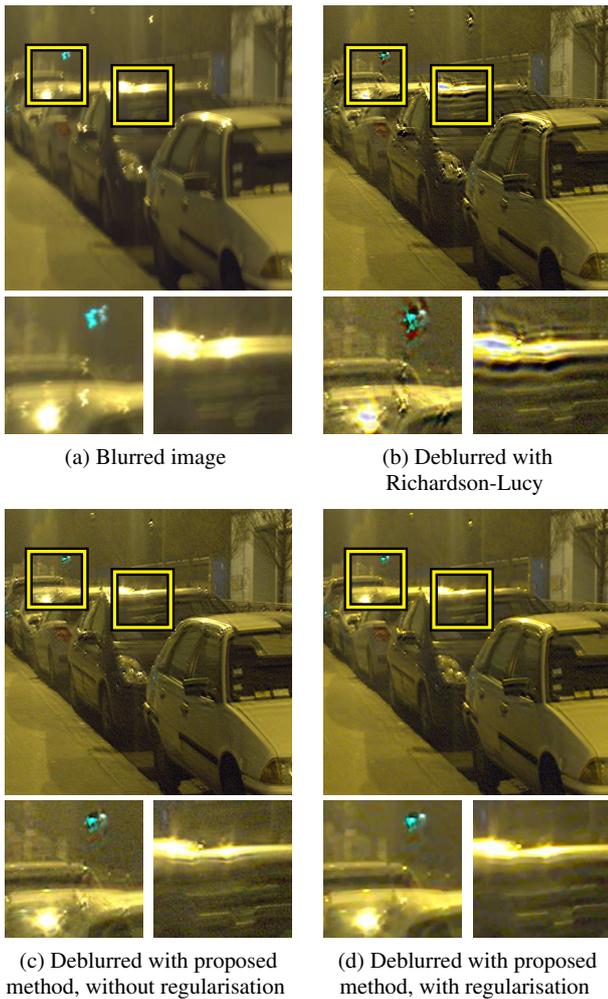


Fig. 9. **Effect of regularisation.** This figure shows a real shaken image, and the effect of our algorithm with and without regularisation. Note that the ringing around saturated regions, visible in (b) is removed by our method (c), even without regularisation. By adding regularisation (d), the remaining noise (visible in textureless regions) is ameliorated.

posed modifications for a synthetic 1D example. As is visible, the use of the decoupled update prevents ringing from spreading across the deblurred result, while the use of the non-linear forward model prevents dark artefacts from appearing inside the bright regions. By combining these two methods, we produce our best estimate of the pixels in \mathcal{S} and \mathcal{U} , while preventing the errors we make in \mathcal{S} from affecting the rest of the image.

5.5 Implementation

In this section we describe some of the implementation details of the proposed algorithm. When segmenting the current estimate of the latent image, we take additional steps to ensure that we make a conservative estimate of which pixels

Algorithm 1: Proposed algorithm

Input: Blurred image \mathbf{g} , blur descriptor \mathbf{w}
Output: Deblurred image $\hat{\mathbf{f}}$

- 1 $\hat{\mathbf{f}}^0 \leftarrow \mathbf{g}$
- 2 **for** $t = 0$ **to** num_iterations **do**
- 3 Decompose $\hat{\mathbf{f}}^t$ into $\hat{\mathbf{f}}_{\mathcal{U}}^t$ and $\hat{\mathbf{f}}_{\mathcal{S}}^t$ using (23) to (25)
- 4 Compute set \mathcal{V} in the blurred image by eroding \mathcal{U} with PSF
- 5 Compute $\hat{\mathbf{f}}_{\mathcal{U}}^{t+1}$ using only data from \mathcal{V} (18)
- 6 Compute $\hat{\mathbf{f}}_{\mathcal{S}}^{t+1}$ using all data, with non-linear modification (20)
- 7 $\hat{\mathbf{f}}_{\text{unreg}}^{t+1} = \hat{\mathbf{f}}_{\mathcal{U}}^{t+1} + \hat{\mathbf{f}}_{\mathcal{S}}^{t+1}$
- 8 Compute regularised update $\hat{\mathbf{f}}^{t+1}$ using (27)
- 9 **end**

can be estimated accurately. First, after thresholding the latent image in Equation (23), we perform a binary erosion on \mathcal{U} , such that

$$\mathcal{U} = \{j | f_j \leq \varphi\} \ominus \mathcal{M}, \quad (29)$$

where \ominus denotes binary erosion, and the structuring element \mathcal{M} used for erosion is a disk of radius 3 pixels. This ensures that all poorly-estimated pixels are correctly assigned to \mathcal{S} (perhaps at the expense of wrongly including some well-estimated pixels too). Performing this step improves the deblurred results, since it is not only the bright pixels whose value is likely to be inaccurate due to saturation, but their neighbours too. Fewer artefacts arise from wrongly assigning a well-estimated pixel into \mathcal{S} than the other way around. Second, in order to avoid introducing visible boundaries between the two regions, we blur the mask \mathbf{u} slightly using a Gaussian filter with standard deviation 3 pixels when decomposing the current latent image $\hat{\mathbf{f}}^t$ into $\hat{\mathbf{f}}_{\mathcal{U}}^t$ and $\hat{\mathbf{f}}_{\mathcal{S}}^t$ in Equations (24) and (25).

6 Results

Figures 1 and 12 show results of non-blind deblurring using the proposed algorithm described in Section 5 on real handheld photographs. For comparison, in Figure 1 we show results produced with the standard Richardson-Lucy algorithm, and the algorithm of Krishnan and Fergus (2009), neither of which is designed to take account of saturation. As such, both algorithms produce large amounts of ringing in the deblurred images. In Figure 12 we compare against a baseline approach, where saturated regions are simply masked out by detecting pixels in the blurred image that exceed a threshold value of $\varphi = 0.9$, as in Equation (12), and discarding those pixels. We found that dilating the masked regions using a

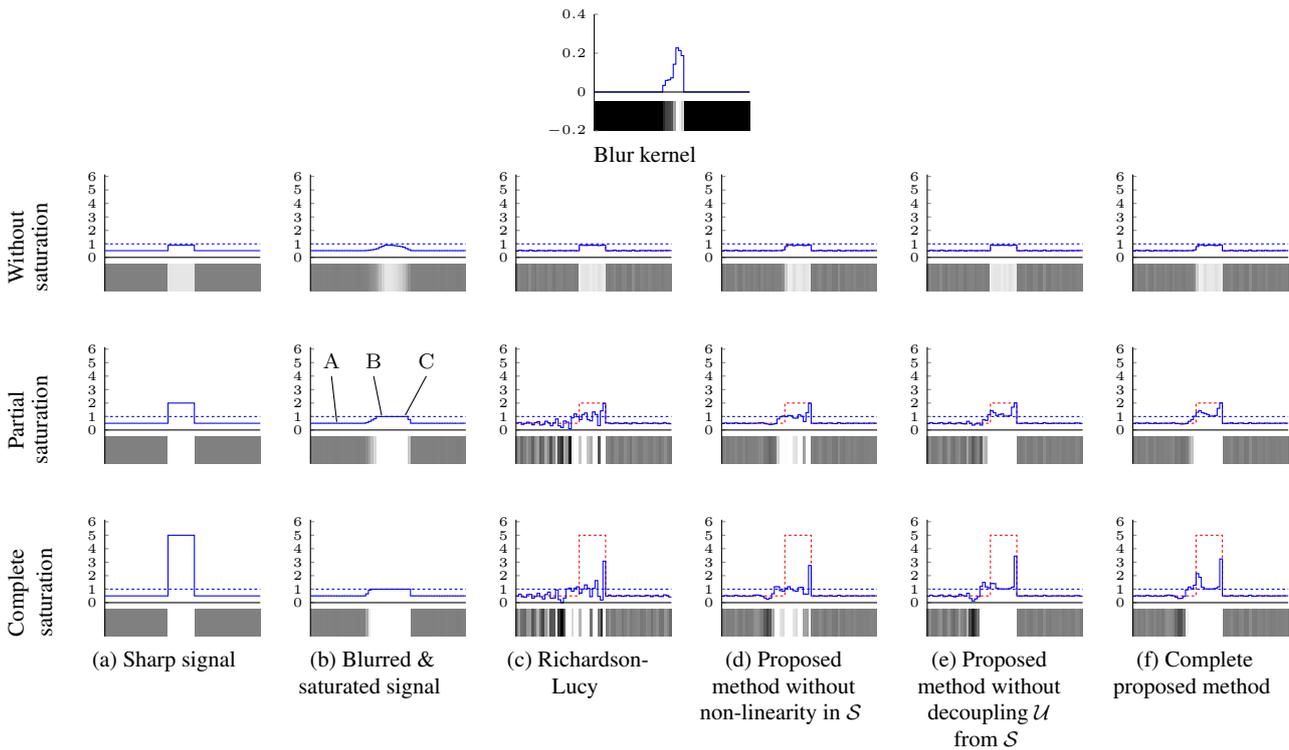


Fig. 10. **Synthetic 1D example of blur and saturation.** Each row shows a sharp “top-hat” signal, blurred using the filter shown at the top. Gaussian noise is added and the blurred signal is clipped. The last four columns show the deblurred outputs for the Richardson-Lucy algorithm, and for our two proposed modifications (the split update step in Section 5.1 and the non-linear forward model in Equation (19) in Section 5.2), separately and together. (Top row) With no saturation, all algorithms produce similar results. (Middle and bottom rows) When some of the blurred signal is saturated (region B), the Richardson-Lucy algorithm (c) produces an output with large ringing artefacts. Although region A is not itself saturated, the ringing propagates outwards from B & C across the whole signal. (d) Performing the split update, decoupling poorly-estimated bright pixels from the rest of the image, reduces ringing, but dark artefacts remain inside the bright region. (e) Deblurring the whole image with the non-linear model prevents artefacts inside the bright regions, but ringing is caused. (f) By combining the split update step with the non-linear forward model, the deblurred signal contains the least ringing, and has no dark artefacts inside the bright region

9×9 square window further reduced ringing, at the expense of leaving more blur around the saturated regions.

In both Figures 1 and 12, the (spatially-variant) PSFs for these images were estimated from the blurred images themselves using our MAP-type blind deblurring algorithm (Whyte et al. 2012), which is based on the algorithm of Cho and Lee (2009). The only modification required to estimate PSFs for saturated images using this blind algorithm is to discard potentially-saturated regions of the blurred image using a threshold. Since, in this case, the aim is only to estimate the PSF (and not a complete deblurred image), we can safely discard all of these pixels, since the number of saturated pixels in an image is typically small compared to the total number of pixels. There will typically remain sufficient unsaturated pixels from which to estimate the PSF.

Note in Figure 1 that the standard Richardson-Lucy algorithm and the algorithm of Levin et al. (2007) produce ringing around the saturated regions, while the proposed algorithm reduces this without sacrificing deblurring quality

elsewhere. In all results in this paper we performed 50 iterations of all Richardson-Lucy variants.

Figures 13 and 14 show the results of non-blind deblurring using our algorithm, alongside those of the recently-proposed algorithm of Cho et al. (2011). The blurred images and their spatially-invariant PSFs are provided by Cho et al., along with their deblurred results². In most cases our results exhibit less ringing than those of Cho et al. (2011), which is due to the fact that we explicitly decouple the estimates of bright pixels from other pixels, in addition to removing saturated blurred pixels.

In order to gauge the quantitative difference between deblurred images containing ringing and the results of our algorithm, we generated a set of synthetically blurred and saturated images, with varying degrees of saturation and blur. We then deblurred using the original Richardson-Lucy algorithm, which produces ringing, and our proposed algorithm. Figure 11 shows the results, indicating that our method pro-

² http://cg.postech.ac.kr/research/deconv_outliers/ (accessed November 12, 2011).

duces a measurable improvement in deblurred image quality, which increases as the amount of saturation increases.

7 Conclusion

In this paper we have developed an approach to the problem of non-blind deblurring of images blurred by camera shake and suffering from saturation. We have provided an analysis of the properties and causes of “ringing” artefacts in deblurred images, as they apply to saturated images. Based on this analysis, we have proposed a non-blind deblurring algorithm, derived from the Richardson-Lucy algorithm, which is able to deblur images containing saturated regions without introducing ringing, and without sacrificing detail in the result. The algorithm is underpinned by the principle that we should prevent gross errors in bright regions from propagating to other regions. We provide an implementation of our algorithm online at <http://www.di.ens.fr/willow/research/saturation/>.

As future work, our algorithm could potentially be extended to handle other sources of ringing, such as moving objects, impulse noise, or post-capture non-linearities (such as JPEG compression). Whenever it is possible to identify poorly-estimated latent pixels, our approach has the potential to reduce artefacts by decoupling these pixels from the rest of the image. In addition, this underlying principle of decoupling poorly-estimated latent pixels could also be applied within other non-blind deblurring algorithms, that are faster or more suitable for different noise models than the Richardson-Lucy algorithm, e.g. (Levin et al. 2007; Wang et al. 2008; Krishnan and Fergus 2009; Afonso et al. 2010; Almeida and Figueiredo 2013).

One alternative direction for future work is the investigation of new regularisers that are targeted at suppressing ringing. As discussed in Section 3.3, gradient-based regularisation is generally insufficient to suppress ringing, and regularisers with a larger bandwidth in the frequency domain are needed. The multi-scale method of Yuan et al. (2008) has this property, and patch-based methods such as that of Zoran and Weiss (2011) might also prove effective.

Acknowledgements This work was partly supported by the MSR-INRIA laboratory, the EIT ICT labs and ERC grant VisRec no. 228180.

References

- M. Afonso, J. Bioucas-Dias, and M. Figueiredo. Fast image recovery using variable splitting and constrained optimization. *IEEE Trans. Image Processing*, 19(9):2345–2356, 2010.
- M. S. C. Almeida and M. A. T. Figueiredo. Deconvolving images with unknown boundaries using the alternating direction method of multipliers. *IEEE Trans. Image Processing*, 22(8):3074–3086, 2013.
- L. Bar, N. Sochen, and N. Kiryati. Image deblurring in the presence of impulsive noise. *IJCV*, 70(3):279–298, 2006.
- A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, 1987.
- C. Bouman and K. Sauer. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Trans. Image Processing*, 2(3):296–310, 1993.
- J.-F. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion deblurring from a single image using sparse approximation. In *Proc. CVPR*, 2009.
- C. Chen and O. L. Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Computational Optimization and Applications*, 5(2):97–138, 1996.
- S. Cho and S. Lee. Fast motion deblurring. *ACM Trans. Graphics (Proc. SIGGRAPH Asia 2009)*, 28(5):145:1–145:8, 2009.
- S. Cho, J. Wang, and S. Lee. Handling outliers in non-blind image deconvolution. In *Proc. ICCV*, 2011.
- P. B. Chou and C. M. Brown. The theory and practice of Bayesian image labeling. *IJCV*, 4(3):185–210, 1990.
- R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graphics (Proc. SIGGRAPH 2006)*, 25(3):787–794, 2006.
- T. W. Gamelin. *Complex Analysis*. Springer-Verlag, New York, 2001.
- A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless. Single image deblurring using motion density functions. In *Proc. ECCV*, 2010.
- S. Harmeling, M. Hirsch, and B. Schölkopf. Space-variant single-image blind deconvolution for removing camera shake. In *NIPS*, 2010a.
- S. Harmeling, S. Sra, M. Hirsch, and B. Schölkopf. Multiframe blind deconvolution, super-resolution, and saturation correction via incremental EM. In *Proc. ICIP*, 2010b.
- H. Ji and K. Wang. Robust image deblurring with an inaccurate blur kernel. *IEEE Trans. Image Processing*, 21(4):1624–1634, 2012.
- N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski. Image deblurring using inertial measurement sensors. *ACM Trans. Graphics (Proc. SIGGRAPH 2010)*, 29(4):30:1–30:9, 2010.
- D. Krishnan and R. Fergus. Fast image deconvolution using hyper-Laplacian priors. In *NIPS*, 2009.
- D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *Proc. CVPR*, 2011.
- A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graphics (Proc. SIGGRAPH 2007)*, 2007.
- A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. Technical report, MIT, 2009.
- A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *Proc. CVPR*, 2011.
- L. B. Lucy. An iterative technique for the rectification of observed distributions. *Astronomical Journal*, 79(6):745–754, 1974.
- W. H. Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.*, 62(1):55–59, 1972.
- R. R. Schultz and R. L. Stevenson. A bayesian approach to image expansion for improved definition. *IEEE Trans. Image Processing*, 3(3):233–242, 1994.
- Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Trans. Graphics (Proc. SIGGRAPH 2008)*, 27(3), 2008.
- Y.-W. Tai, H. Du, M. S. Brown, and S. Lin. Image/video deblurring using a hybrid camera. In *Proc. CVPR*, 2008.
- Y.-W. Tai, P. Tan, and M. S. Brown. Richardson-Lucy deblurring for scenes under a projective motion path. *IEEE PAMI*, 33(8):1603–1618, 2011.
- Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sciences*, 1(3):248–272, 2008.



Blur size:	3 pixels					7 pixels					15 pixels				
Scale factor:	1.0	1.5	2.0	2.5	3.0	1.0	1.5	2.0	2.5	3.0	1.0	1.5	2.0	2.5	3.0
RL, SSIM:	0.990	0.980	0.954	0.924	0.894	0.956	0.939	0.898	0.859	0.827	0.905	0.891	0.858	0.818	0.782
Ours, SSIM:	0.990	0.981	0.959	0.933	0.906	0.955	0.946	0.923	0.896	0.868	0.903	0.893	0.873	0.847	0.820

Fig. 11. **Quantitative effect of our algorithm on reducing artefacts.** Starting from the sharp image shown on the left, we synthesize a set of blurry images with varying degrees of blur and saturation by scaling the intensities by increasing amounts, and blurring with several sizes of horizontal linear blur. As the scale factor increases, more pixels in the synthetic blurred image become saturated. In the table, we report the structural similarity index (SSIM) (Wang et al. 2004) between the true sharp image and the deblurred image, using the original Richardson-Lucy (RL) algorithm, which causes ringing, and our proposed algorithm. The SSIM of the deblurred images decreases both when the amount of blur increases, and when the amount of saturation increases. However, in almost all cases, our algorithm achieves the same or better results, and the improvement over the original Richardson-Lucy algorithm increases with the amount of saturation.

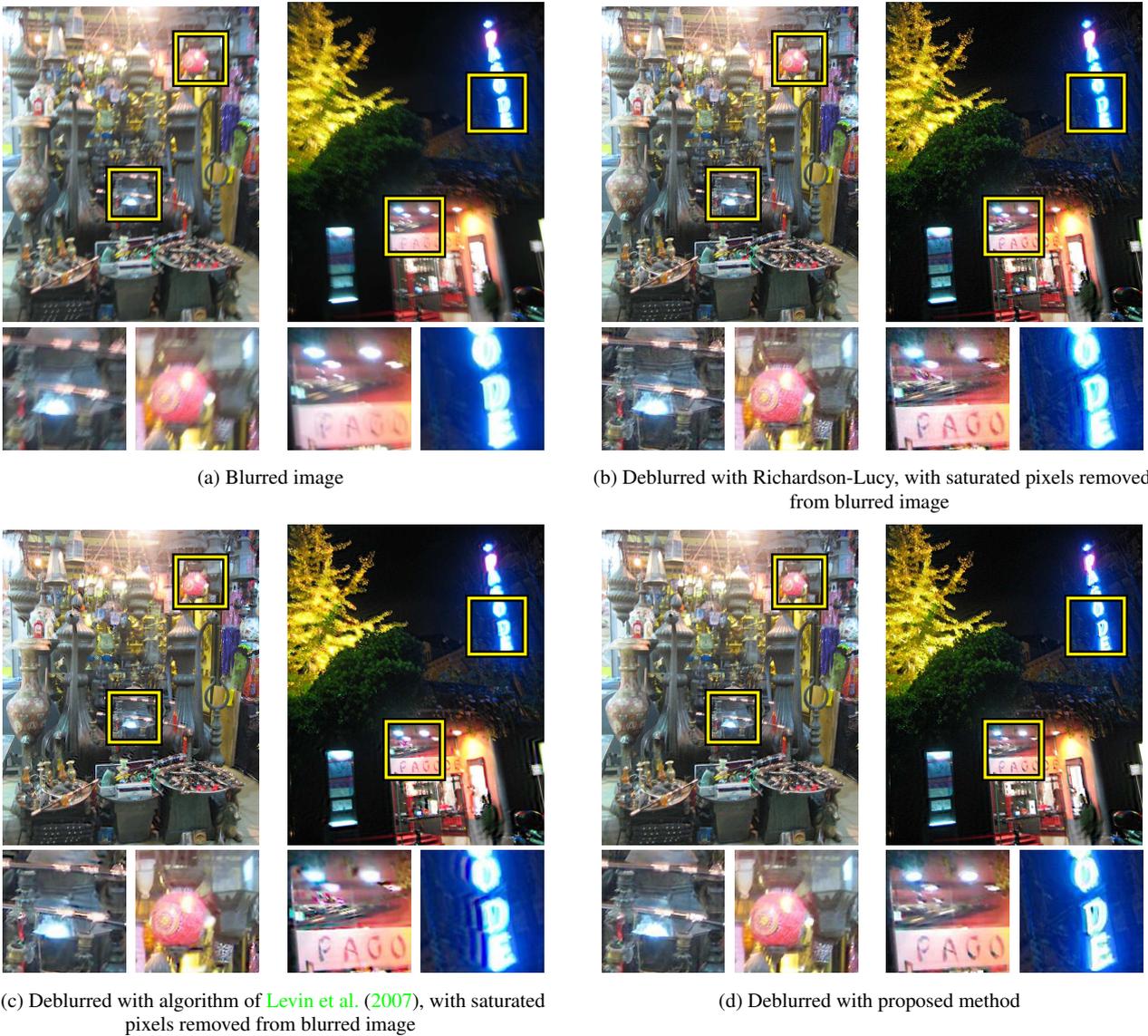


Fig. 12. **Deblurring saturated images.** Note that the ringing around saturated regions, visible in (b) and (c) is removed by our method (d), without causing any loss in visual quality elsewhere.

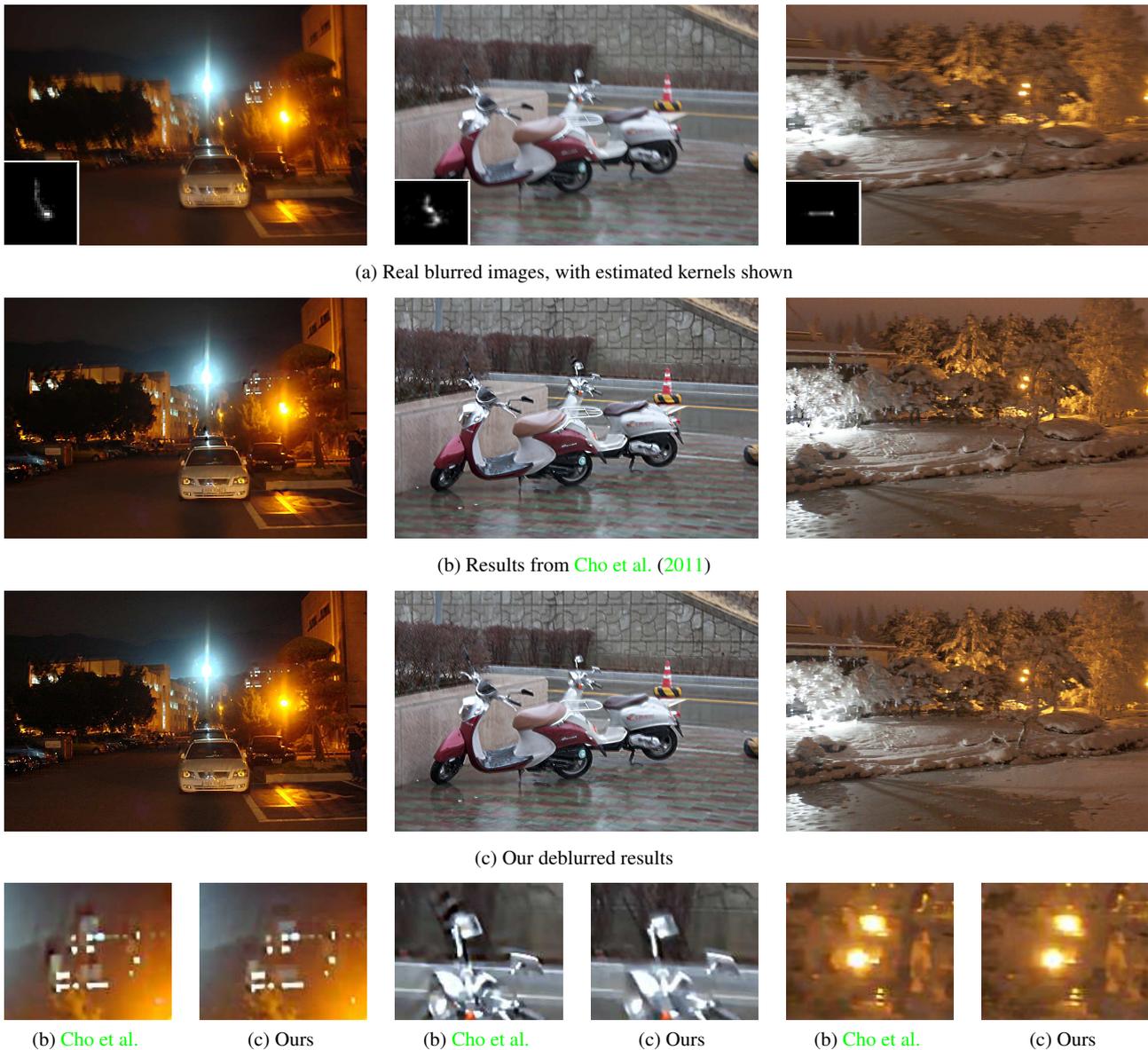


Fig. 13. **Comparison to the method of Cho et al. (2011).** The blurred images, the spatially-invariant PSFs and the results of their method were provided by the authors. Note in the close-ups and elsewhere in the images, our results generally contain less ringing than those of Cho et al. (2011), without sacrificing detail. For example, above the car in the left column, and along the top edge of the bikes in the middle column.

Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Processing*, 13(4):600–612, 2004.

M. Welk. Robust variational approaches to positivity-constrained image deconvolution. Technical Report 261, Saarland University, Saarbrücken, Germany, 2010.

O. Whyte. *Removing Camera Shake Blur and Unwanted Occluders from Photographs*. PhD thesis, ENS Cachan, 2012.

O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. In *Proc. CVPR*, 2010.

O. Whyte, J. Sivic, and A. Zisserman. Deblurring shaken and partially saturated images. In *Proc. CPCV, with ICCV*, 2011.

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

N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. MIT Press, 1949.

L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *Proc. ECCV*, 2010.

J. Yang, Y. Zhang, and W. Yin. An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise. *SIAM J. Sci. Comput.*, 31(4):2842–2865, 2009. doi: 10.1137/080732894.

L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. *ACM Trans. Graphics (Proc. SIGGRAPH 2007)*, 26(3), 2007.

L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM Trans. Graphics (Proc. SIGGRAPH 2008)*, 27(3), 2008.

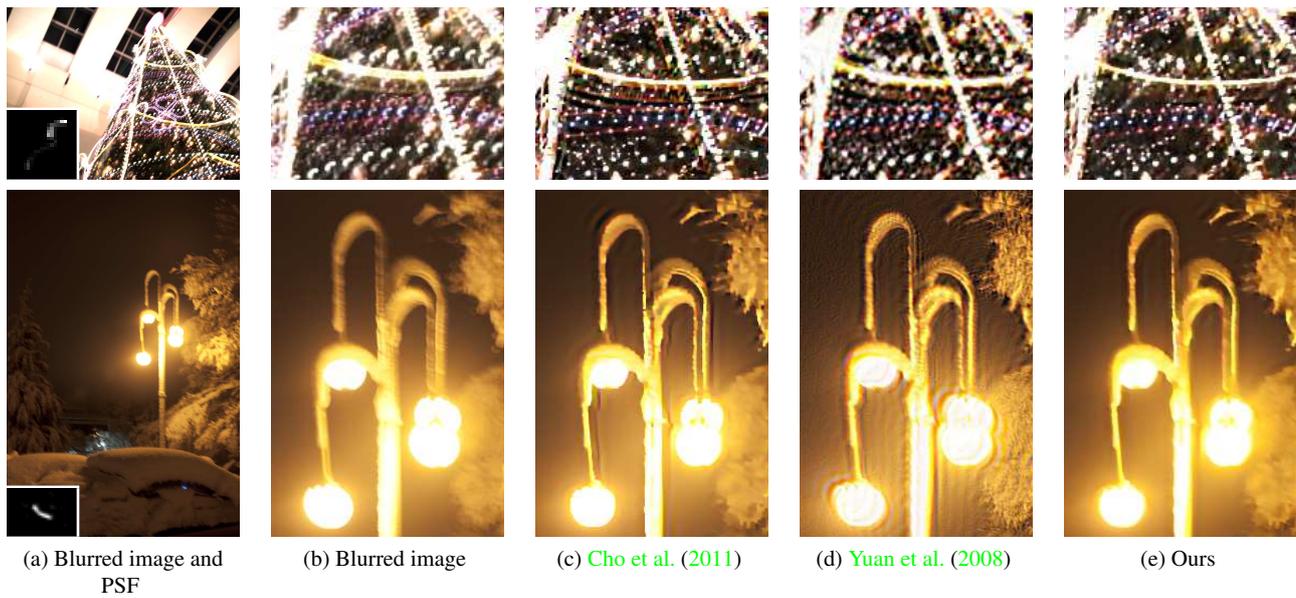


Fig. 14. **Comparison to Cho et al. (2011) and Yuan et al. (2008)**. This figure compares non-blind deblurring results, on saturated images, for (c) the algorithm of Cho et al. (2011), (d) the Richardson-Lucy-based algorithm of Yuan et al. (2008), and (e) our proposed algorithm. While both our algorithm and the algorithm of Yuan et al. are based on the Richardson-Lucy algorithm, our results contain much less ringing, due to handling the saturated pixels explicitly. Compared to the method of Cho et al., our results contain similar or less ringing. Results in (c) and (d) provided by Cho et al..

D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *Proc. ICCV*, 2011.