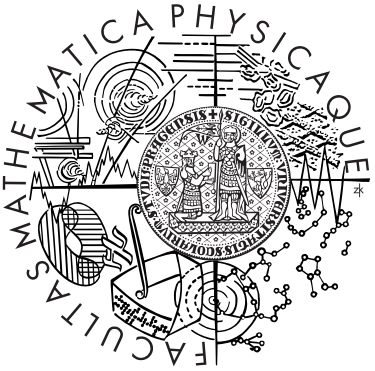
Univerzita Karlova v Praze
Matematicko-fyzikální fakulta



Habilitační práce

Multichannel Blind Image Restoration: from theory to practice

Filip Šroubek

Obor: Matematika – matematické modelování a numerická matematika

Contents

Resumé

This habilitation thesis tackles a complicated inverse problem of estimating a latent sharp image from observed blurred images. The blurring process is modeled by convolution and the inverse problem is referred to as "blind" deconvolution since we assume limited or zero knowledge of the convolution kernel. A wide range of degradation processes that occur during data acquisition can be modeled or at least well approximated by convolution. Camera/object motion, camera optics, turbulence of a measuring media such as the atmosphere are some of the examples. A frequent encounter of convolution in diverse application areas makes the deconvolution problem particularly appealing.

The thesis consists of eight scientific articles that survey author's contribution to the theory of image deconvolution. The common framework in the presented collection is a multichannel scenario, i.e., the same scene is captured more than once and each observed image contains a slightly different convolution kernel (blur). The presented work shows that under the assumption of multichannel acquisition we have tools to estimate blurs directly from the observed images without any prior knowledge of the kernel shape. Further it demonstrates that formulating blind deconvolution as an energy minimization problem provides the necessary robustness in the case of noisy acquisitions, which is essential for usability of blind deconvolution in practical applications.

Real data seldom follow the mathematical model precisely. This is either due to unknown perturbations or the acquisition model is more complicated than the assumed mathematical model. A common problem encountered in practice is misregistration of input images. It is hard to guarantee that during multiple acquisitions the observed images will be spatially aligned. The author shows that the proposed multichannel blind deconvolution method automatically estimates translation among images by shifting the estimated convolution kernels in the correct direction, which makes the method robust to slight misalignment of input images.

Another common problem is that input images have insufficient spatial resolution. Increasing the image resolution is called superresolution. The aliasing effect is important in this case as the high-resolution details are recovered from the overlapping image spectra. The author proposes to address both deconvolution and superresolution in one common framework resulting in a blind superresolution method, which simultaneously estimates convolution kernels and the sharp image in the high-resolution domain. In practice the maximum meaningful resolution factor we can achieve (often between $2\times$ and $3\times$) is limited by the number of input images and discrepancies from the mathematical model. The author proves that the theory of blind superresolution derived for integer resolution factors is easily extendable to rational factors using a polyphase decomposition.

Current images have many millions of pixels and fast, close to real time, deconvolution methods are preferred. Recent progress in the direction of fast numerical optimization methods is included in the thesis.

The final three articles in the collection illustrate applicability of blind deconvolution in ophthalmology and mobile phone photography. Images of eye retina, analyzed by ophthalmologists are often blurred due to eye movement and pupil imperfections. The first paper demonstrates that multichannel blind deconvolution could be a useful tool for obtaining sharp retina images and therefore improving retina defect diagnosis. The second paper extends blind deconvolution to a space-variant case and evaluates its performance on retina images. The final paper discusses implementation of blind deconvolution in embedded device such as smartphones.

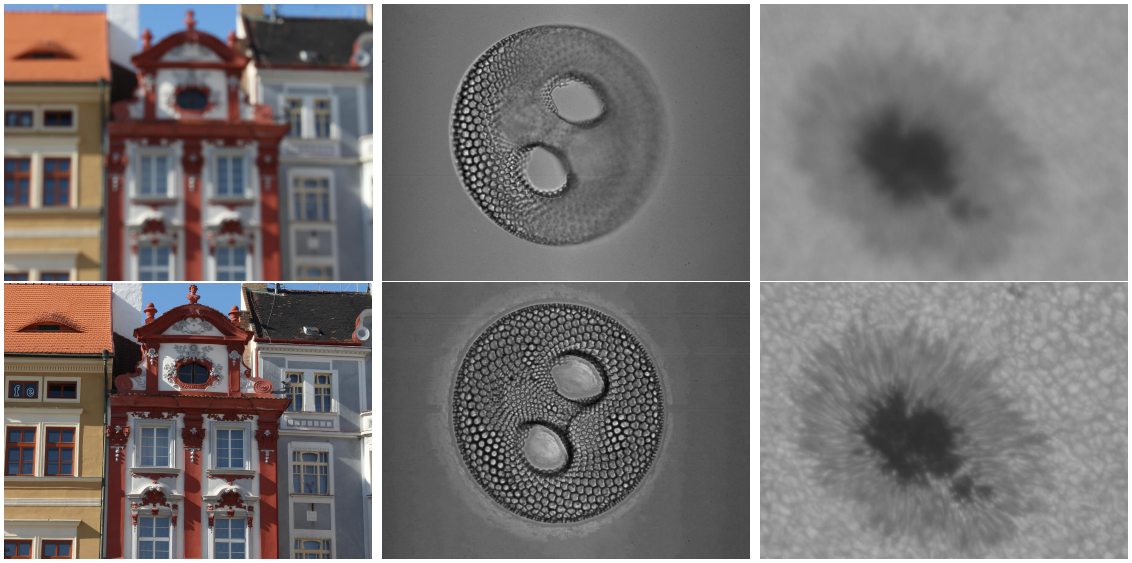


Figure 1: Examples of acquired images under degraded conditions (top row) and ideal nonviable conditions (bottom row) in three application areas: photography (house), microscopy (biological specimen), astronomy (sunspot).

1 Introduction

The habilitation thesis addresses one of the core problems of image processing, which is an estimation of sharp images from degraded observations (measurements). Processing images becomes an every-day practice in a wide range of applications in science and technology and we rely on images with ever growing emphasis. Our understanding of the world is however limited by measuring devices that we use to acquire images. Inadequate measuring conditions together with technological limitations of the measuring devices result in acquired images that represent a degraded version of the "true" image. Fig. 1 illustrates examples of acquired images under real conditions versus ideal conditions in three different application areas. It is important to underline that the ideal conditions may not be achievable in practice and the only solution to get the ideal image is to estimate it from the acquired ones.

The relation between the true latent image u and the degraded observed image g is given by a formula

$$g = Hu + n\,, \tag{1}$$

where H is the degradation operator and n is additive noise. By the word "degradation" we loosely mean an operator that diminishes or completely removes high frequency information (details) from images. The difficulty with H is that it is ill-conditioned, which means that during inversion noise n gets amplified and the solution is unstable. We face an *ill-posed inverse* problem that requires special handling. Our scenario is even more complicated as H is unknown, but we can assume that it belongs to a certain type of degradation.

The most common type of degradation, which is considered in the thesis, is *convolution*:

$$Hu(x) = \int h(x-t)u(t)\mathrm{d}t\,, \tag{2}$$

where $x, t \in \mathbb{R}^2$ for images. This definition extends to any number of dimensions and not just $\mathbb{R}^2$. For example in confocal microscopy, convolution is in $\mathbb{R}^3$. Function h is a convolu-

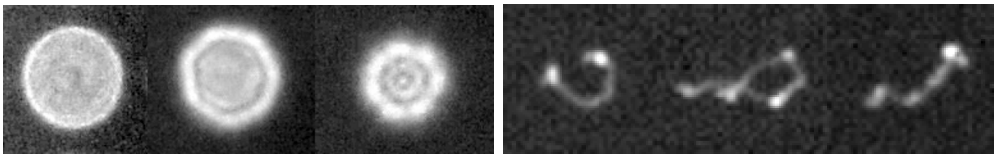


Figure 2: Examples of real camera blurs: (left) three blurs caused by out-of-focus lens with different lens parameters (focal length and aperture size), notice polygonal shape clearly visible in the central image, which corresponds to the aperture opening of 7-blade diaphragm; (right) three blurs caused by camera motion during exposure.

tion kernel (or simply blur) and defines the behavior of the convolution operator. It is also called a *point spread function* (PSF), because h is an image the device would acquire after measuring an ideal point source $\delta(x)$ (delta function). Image blur due to camera motion or improper camera focus setting can be modeled by convolution. The degree of blurring influences the PSF size and the physical nature of blurring determines the PSF shape. For example, out-of-focus camera lens causes convolution with a cylindrical PSF, and camera motion causes convolution with a curvy PSF, where the curve shape is related to the motion trajectory; see Fig. 2. There is a wide range of imaging devices, in which the acquisition process can be modeled by convolution. Apart from devices with classical optical systems, such as digital cameras, optical microscopes or telescopes, convolution degradation occurs also in atomic force microscopy (AFM) or scanning tunneling microscopy (STM), where the PSF is determined by the shape of a measuring tip. Media turbulence (e.g. atmosphere for terrestrial telescopes) can cause blurring that can be modeled by convolution, and there are many more examples. To make convolution more general, it is often necessary to allow the PSF to change over the image. Then in (2), h becomes also a function of x, i.e. $h(x,t)$. This is called *space-variant* convolution, though strictly speaking it is not mathematical convolution any more. Using space-variant convolution we can model a wider range of degradation types, such as blur induced by complex camera motion and rotation, out-of-focus blur in a wide-depth scene, or blur due to hot-air turbulence. Volumes acquired by a confocal microscope are in general degraded by 3D space-variant blur, which renders this particular problem even more challenging.

Deconvolution, as the name suggests, refers to the process of inverting the convolution operator H. *Blind deconvolution* denotes the case when the PSF is also unknown. If only one image g is observed then we call this problem *single-channel* blind deconvolution.

When we have more observations of the latent image u, we write

$$g_k = H_k u + n_k \,, \tag{3}$$

where g_k is the k-th degraded image. Notice the subscript k in H, which indicates that we assume different convolution kernels h_k across images. Indeed it is highly desirable that the PSFs differ, since then multiple observations may convey complementary information. The estimation of the latent image u from the multiple observations g_k's without any knowledge of h_k's is referred to as *multichannel blind deconvolution* and this is the unifying theme of the selected publications, which constitute the habilitation thesis.

An interesting extension of the above degradation operator, which is also discussed in the thesis, is to consider in addition to convolution H a decimation operator D and rewrite model (3) as

$$g_k = D H_k u + n_k \,. \tag{4}$$

The decimation operator D models sampling on a camera sensor, which is affected by diffraction, shape of light sensitive elements and void spaces between the elements. Including D in the model allows us to increase spatial resolution of images. The corresponding inverse problem is called *superresolution*.

The rest of the thesis is organized as follows. Section 2 overviews the evolution of state of the art in the last 12 years during which the author of the thesis has contributed to the field of image restoration. This section discusses the development in a slightly wider perspective than the thesis topic and includes also references to author's research articled that are not part of the thesis collection.

Section 3 lists eight published research articles that constitute the habilitation thesis and for each article gives a brief overview of the main ideas and contribution to the state of the art.

Section 4 concludes the habilitation thesis and the final Section 5 contains reprints of the research articles.

2 State of the art

Recovering u from g even in the nonblind case is not straightforward. A standard technique is to convert the deconvolution problem to energy minimization [1]. The core term in the energy function implied by the model (1) is called a data-fitting or fidelity term and takes the form

$$E(u) = \|g - Hu\|_2^2 , \tag{5}$$

where $\|\cdot\|_p$ denotes the L_p norm. In this case, finding the minimum of $\hat{u} = \arg\min_u E(u)$ is equivalent to a lease-square fit. The difficulty of finding the minimum of (5) resides in the degradation operator H. Since blurring diminishes high frequency information (image details), the spectrum of H contains zeros or values close to zero. Therefore, H is generally not invertible.

A classical way to solve such ill-posed minimization problems is to add regularization terms and convert the problem to a related one that admits a unique solution. Regularization conveys additional prior knowledge of the original image u to the energy function. Priors are application dependent and general rules for constructing the priors are hard to find. Nevertheless, studying image statistics shows that the majority of natural images contain smooth regions with abrupt changes of intensity values at object boundaries that correspond to edges. An image gradient is a useful feature, which can distinguish between edges and smooth regions. Therefore, regularization terms are often functions of $\nabla u = [u_{x_1}, u_{x_2}]$. Other features are also used, such as outputs of various differential filters. A typical example are wavelet coefficients. The L_p quasi-norms for $p \leq 1$ of the image gradient (a special case of $p = 1$ is called *total variation* [2]), is a popular choice for the image regularization term. Then the regularized energy becomes

$$E(u) = \|g - Hu\|_2^2 + \lambda \|\nabla u\|_p^p . \tag{6}$$

Parameter λ is a positive weighting constant. The first term forces the solution to be close to the observed data and the second one guarantees that the solution is sufficiently smooth in the L_p sense. Noise is removed in smooth regions, while edges are not excessively penalized, since we use L_p for $p \leq 1$ instead of L_2 norm.

In blind image restoration, the energy E becomes a functional of two unknowns, original image u and degradation H. In our case of convolution, H is parametrized by PSF

h. If both u and h are unknown, the problem is underdetermined and some additional information (e.g. regularization of h) or different minimization strategy is necessary. For example, so called "no-blur" solution, when $\hat{u} = g$ and $\hat{h} = \delta$, is one of the unwanted solutions, which gives the blurred input as the latent image and the delta function as the estimated PSF. Many commonly used energy functions get trapped in "no-blur" solution.

A more general formulation of the image restoration problem is stochastic [3]. Images and PSFs are regarded as random vector fields [4] with known prior probability distribution functions $p(u)$ and $p(h)$, respectively. The Bayesian paradigm dictates that the inference on the latent image and PSF should be based on the posterior probability

$$p(u, h|g) \propto p(g|u, h)p(u)p(h)\,, \tag{7}$$

where u and h are assumed to be independent. The conditional distribution $p(g|u, h)$ is given by our model (1). Suppose that n is white Gaussian noise then the logarithm of the conditional distribution is equivalent to (5). Different noise distributions result in different data-fitting terms. For example, the Laplace distribution implies $E(u) = \|z - Hu\|_1$. Estimating the pair $(\hat{u}, \hat{h})$ is equivalent to maximizing the posterior $p(u, h|g)$, which is commonly referred to as the maximum *a posteriori* (MAP) approach. Note that maximization of the posterior is equivalent to minimization of $-\log p(u, h|g)$, which is an energy minimization task (6), where the prior plays a role of the regularization term. The simplest but also the most common method maximizes the posterior in an alternative manner with respect to u and h. Unfortunately, since the posterior is not convex, it has an uneven shape with many local peaks and alternating maximization often returns incorrect solutions.

In the case of single-channel blind deconvolution, proposed approaches include stronger regularization both on the image and blur and above all must use more sophisticated estimation procedures as discussed in Sec. 2.1. The multichannel case discussed in Sec. 2.2 permits estimation of the blurs without any prior knowledge of their shape. The space-variant case with parametric approaches is covered in Sec. 2.3.

2.1 Single-channel blind deconvolution

One way to tackle the problem, when we have only one observation and no knowledge of the PSF, is to assume a parametric model of the PSF and search in the space of parameters and not in the full space of PSFs. Chang *et al.* in [5] investigated zero patterns of the Fourier transform or cepstrum, and assumed only parametric motion or out-of-focus blurs. More recently a similar idea was proposed in [6]. Low-level parametric methods for estimating general motion blurs were proposed in [7–10]. Parametric methods have two disadvantages. They are more restrictive than the fully blind ones and they can also be computationally more demanding. Even if minimization with respect to the unknown PSF is linear, minimization with respect to one of the parameters of the PSF does not have to be linear and thus effective methods for solving linear problems can not be applied. Real PSFs typically differ from their parametric models and this prevents the parametric methods to find an exact solution.

There has been a considerable effort in the image processing community in the last three decades to find a reliable algorithm for single-channel blind deconvolution. First algorithms appeared in telecommunication and signal processing in early 80's [11]. For a long time, the general belief was that blind deconvolution was not just impossible, but that it was

hopelessly impossible. Proposed algorithms usually worked only for special cases, such as astronomical images with uniform (black) background, and their performance depended on initial estimates of PSFs; see [12–14].

Over the last few years, single-channel blind deconvolution experiences a renaissance. The key idea of new algorithms is to address the ill-posedness of blind deconvolution by characterizing the prior $p(u)$ using natural image statistics and by a better choice of estimators. The idea of natural image statistics was also explored by the author of the thesis in [15]. A heated activity started with the work of Fergus *et al.* [16], who applied variational Bayes to approximate the posterior $p(u, h|g)$ by a simpler distribution $q(u, h) = q(u)q(h)$. Other authors [17–20] stick to the "good old" alternating MAP approach, but by using ad hoc steps, which often lack rigorous explanation, they converge to a correct solution. Levin *et al.* in [21, 22] proved that a proper estimator matters more than the shape of priors. They showed that marginalizing the posterior with respect to the latent image u leads to the correct solution of the PSF h. The marginalized probability $p(h|g)$ can be expressed in a closed form only for simple priors that are, e.g., Gaussian. Otherwise approximation methods such as variational Bayes [23] or the Laplace approximation [24] must be used.

More recently, even better results were achieved if the model of natural images statistics was abandoned and priors that force unnaturally sparse distributions were used instead, such as in [25,26] and similarly also by the author of the thesis in [27]. A formal justification of unnaturally sparse distributions was given in [28] together with a unifying framework for the MAP and variational Bayesian formulation.

2.2 Multichannel blind deconvolution

The framework of multiple observations as defined in (3) provides the necessary constraint to make the image restoration task well posed. One of the earliest intrinsic multichannel blind deconvolution methods [29] was designed particularly for images blurred by atmospheric turbulence. Harikumar *et al.* [30] proposed an indirect algorithm, which first estimates the blur functions and then recovers the original image by standard nonblind methods. The blur functions are equal to the minimum eigenvector of a special matrix constructed from the blurred images, which is the same idea published earlier for 1D signals in [31]. Necessary assumptions for perfect recovery of the blur functions are noise-free environment and channel coprimeness, i.e. a scalar constant is the only common factor of the blurs. Giannakis *et al.* [32] developed another indirect algorithm based on Bezout's identity of coprime polynomials which finds restoration filters and by convolving the filters with the observed images recovers the original image. Both algorithms are vulnerable to noise and even for a moderate noise level restoration may break down. In the latter case, noise amplification can be attenuated to a certain extent by increasing the restoration filter order, which comes at the expense of deblurring. Pai *et al.* [33] suggested two multichannel restoration algorithms that constract a special matrix from the input blurred images and estimate directly the original image from the null space or from the range of the matrix. Another direct method based on the greatest common divisor was proposed in [34]. Interesting approaches based on the ARMA (autoregressive moving average) model are given in [35]. Multichannel blind deconvolution based on the Bussgang algorithm was proposed in [36], which performs well on spatially uncorrelated data, such as binary text images and spiky images. Most of the algorithms lack the necessary robustness since they do not include any noise assumptions in their derivation and miss regularization terms. The author

of the thesis proposed an iterative multichannel algorithm [37] that performs well even on noisy images. It is based on least-squares deconvolution by anisotropic regularization of the image and between-channel regularization of the blurs.

An obstacle of multichannel methods, which is not present in the single-channel case, is that the observed images must be spatially aligned, which is seldom true in practice. A first attempt in this direction was done by the author in [38], where blind deconvolution of images that are mutually shifted by unknown vectors was proposed. The author extended this idea to superresolution in [39]. In superresolution, the physical resolution of the image is increased, which is equivalent to considering both convolution and decimation according to the model in (4).

The recent findings in the single-channel area, in particular the favorable property of unnaturally sparse priors, was incorporated in the multichannel case in [40].

An interesting specific branch of the multichannel methods, which was found appealing for its potential practical usage, works with a pair of acquired images: one correctly exposed but blurred and one underexposed (noisy) but sharp image. We can apply the general multichannel blind deconvolution methods, but in this particual case they are even better posed, as was demonstrated in [41–43] and also by the author of the thesis in [44].

Blind deconvolution in the multichannel framework is in general a well-posed inverse problem. However, in many practical situations we do not have multiple observation of the same scene, which would differ only by the convolution kernel, and we must revert to the single-channel case.

2.3 Space-variant blind deconvolution

Space-variant blind deconvolution is even more complicated as the PSF is also a function of the position vector. As a rule, the space-variant PSF cannot be expressed by an explicit formula but in many cases it has a special structure that can be exploited. For example, the blur caused by camera motion is limited by six degrees of freedom of rigid body rotation and translation. If we have an estimate of the camera motion from inertial sensors [45, 46] or other sources [47–49], we are able to reconstruct the PSF and recover the latent image. The author of the thesis contributed in this area by implementing blind deconvolution in mobile phones using gyroscopes [50, 51]. If the information about camera motion is not available, the PSF can be estimated directly from the input images but additional restrictions on camera motion are necessary for the methods to work. If for example only rotation is assumed, we can express the degradation operator as a linear combination of basis blurs (or images) and solve the blind problem in the space of the basis, which has much lower dimension than the original problem. Whyte *et al.* [52] considered rotations about three axes up to several degrees and described blurring using three basis vectors. For blind deconvolution, they used an algorithm analogous to [16] based on marginalization over the latent sharp image. Gupta et al. [53] and Hirsch et al. [54] adopted a similar approach, replacing rotations about x and y axes by translations. Recently, convolution neural networks were used to remove various types of blurs directly from the blurred image, as e.g. proposed by the author of the thesis in [55], or indirectly by first estimating a motion field and then performing space-variant nonblind deconvolution of the image [56].

Removing out-of-focus blur is a more complex problem, since the PSF depends on object distance and we need to estimate also the depth map as was proposed in [57, 58].

If the PSF changes smoothly over the image, the PSFs can be considered as constant

on a small neighborhood and estimated on a regularly spaced grid. This idea has been proposed by the author of the thesis in [44] and later extended to superresolution in [59]. For estimation, we can apply locally single-channel blind deconvolution methods [60,61], or if a pair of blurred and noisy/underexposed images is available, multichannel methods [41,62]. Recently, more accurate parametric interpolation of PSFs on the grid has been proposed by the author in [63].

An especially difficult situation is that of the blur caused by object motion, as objects usually move independently of each other and often in different directions. In order to achieve a good quality of deblurring, the object must be precisely segmented, taking into account partial occlusion close to object outline. Most of the methods [64–67] follow the pioneering paper of Levin [68] that assumed that objects move with a constant velocity and segmented objects based on a statistics of image derivatives. A completely different approach, without the need to segment objects, was proposed in [69] but the method works only with video sequences and not with single images.

3 Research articles in the habilitation thesis

The habilitation thesis represents a survey of the most important results in the area of multichannel blind deconvolution that the author has achieved in the period 2003–2014. The author has incorporated some of these results into the course "Variational Methods in Image Processing", which he opened at the Faculty of Mathematics and Physics, Charles University and which he has been teaching since 2011. The teaching provides him with a feedback which stimulates some of his recent and current research.

The habilitation thesis is a collection of eight research articles. The acquisition model is assumed to be of the form (3) and in all the cases except one (paper no. 2) the energy minimization approach is considered. The first five articles are chronologically sorted and summarize contribution of the author to the theory of blind deconvolution. The remaining three articles are examples of applied research articles that illustrate the use of blind deconvolution in practice.

1. F. Šroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Transactions on Image Processing*, vol. 12, no. 9, pp. 1094–1106, 2003.

   This research article was the first step in the direction of robust multichannel deconvolution methods. We use energy minimization approach with regularization. The PSF regularization is based on a simple but elegant idea presented originally by Harikumar *et al.* in [30]. They showed that by constructing a special matrix from blurred input images, we can determine PSFs as minimum eigenvalues of the matrix. However, stability of Harikumar's method deteriorates quickly with increasing noise. Instead of using the special matrix directly, we construct from the matrix a quadratic regularization term, which is intrinsically multichannel as it couples all the input images and approaches the minimum for correct PSFs. Then we minimize the regularized energy function with respect to the image and PSFs. To increase stability even further we include image regularization based on image gradients, such as Total Variation [70] or Mumford-Shah functional [71]. The final energy function is convex but the solution leads to nonlinear equations. This drawback is solved by a half-quadratic algorithm [72], which converts the problem to a set of linear equations.

A special attention is paid to discretization of image regularization terms using four-connectivity and eight-connectivity approximation. The performance of the proposed method is evaluated on synthetically blurred data and also on camera out-of-focus images and astronomical data.

The main contribution of this research article is in constructing a novel multichannel regularization term and proposing an iterative method for blind deconvolution, which is robust to noise and thus suitable for practical applications.

2. F. Šroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Transactions on Image Processing*, vol. 14, no. 7, pp. 874–883, 2005.

   In this research article, we adopt a stochastic approach to multichannel blind deconvolution and formulate the restoration problem as a MAP inference; see (7). Regularization terms are now replaced by prior distributions of images and blurs. This interpretation introduces covariance matrices that were omitted in the original formulation and allows us to better understand meaning of weighting parameters in front of these terms. We also prove that the proposed method can compensate for a misalignment of input blurred images. From the practical point of view, this is an important feature. We require multiple (minimum two) images of the same scene that are blurred in a slightly different way. Video sequences or continuous shooting in digital cameras often provide data where neighboring frames (images) depict the same scene with blurs slightly varying in time. However, such images are rarely spatially aligned (registered). We can use registration methods in [73] to geometrically align input blurred images, but registration of blurred images is imprecise. We show that by overestimating the blur support, the proposed method is able to automatically shift the estimated blurs and thus cancel spatial misalignment of the images.

   The main contribution of this research article is in built-in compensation for misalignment of input images, which further increases applicability of the proposed method.

3. F. Šroubek, G. Cristóbal, and J. Flusser, "A unified approach to superresolution and multichannel blind deconvolution," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2322–2332, 2007.

   In the previous two research articles, we have developed a theory of multichannel blind deconvolution. Multiple observations of the same scene give us one additional benefit. If the observations differ by subpixel shifts we can also increase spatial resolution of the latent image (superresolution), which is explored in this research article. We assume the multichannel acquisition model with decimation as defined in (4). We prove that even in the presence of the decimation operator D, which does not commute with the convolution operator H, we can construct a blur regularization term, which is similar to the regularization term in the classical multichannel blind deconvolution problem. The regularization term is not strictly convex and approaches the minimum on a subspace of dimensions proportional to the superresolution factor. The superresolution factor is a user parameter and determines the final high resolution of the latent image. With an increasing superresolution factor the minimum number of input images necessary to construct the regularization term increases proportionally. The regularized energy function is minimized with respect to the image and blurs as in the case of multichannel blind deconvolution. However in this case we

also recover the lost spatial resolution of the latent image. We named this problem blind superresolution.

The proposed blind superresolution method went way beyond standard superresolution techniques. While estimating the blurs in the high resolution grid of the final latent image, we calculate not only PSFs but also subpixel shifts. This made it one of the first methods that performs deconvolution and resolution enhancement simultaneously.

4. F. Šroubek, J. Flusser, and G. Cristobal, "Super-resolution and blind deconvolution for rational factors with an application to color images," *Computer Journal*, vol. 52, no. 1, pp. 142–152, 2009.

   The previous research article demonstrates that superresolution neatly combines with multichannel blind deconvolution. The minimum number of input images required for well-posed blind superresolution depends on the superresolution factor. For example estimating blurs and increasing resolution by a factor of 2, requires at least 5 images. If the factor is 3, we already need 10 images. In many practical applications it is difficult to guarantee this minimum number. In addition, acquired images do not follow precisely our mathematical model, which implies that superresolution factor of more than 2 provides negligible improvement in practice as was experimentally demonstrated. These facts show that non-integer superresolution factors below two are meaningful as they require less number of input images and recover high frequency information. This work uses the notion of polyphase decomposition to derive PSF regularization terms that work for any rational superresolution factor. We can thus extend conclusions derived in the previous research article about blind superresolution to factors such as $3/2 = 1.5$ (requires 3 input images) or $7/4 = 1.75$ (requires 4 input images). Other improvements discussed in the paper are image regularization terms for color images, and advantages of image registration performed in the decimation matrix versus registration done beforehand.

   The main goal of this paper has been to extend the theory of blind superresolution for integer factors to rational factors.

5. F. Šroubek and P. Milanfar, "Robust multichannel blind deconvolution via fast alternating minimization," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1687–1700, 2012.

   The quality of image deconvolution is very sensitive to accuracy with which the PSF is estimated. Disturbing artifacts appear in deconvolved images due to inaccurate PSF estimation. Using stronger image regularization we can avoid the artifacts but we inevitably loose details. In the first part of this research article, we analyze the multichannel blur regularization term and show that its dependence on noise may bias the estimation of PSFs in the noisy case. Using filtered images instead of the original intensity values to construct the blur regularization term diminishes the bias and improves the accuracy of PSF estimation. The second part of the paper is dedicated to a fast numerical optimization method, which would allow blind deconvolution of large images (several Mpixels) and large blurs (up to 100×100 pixels). Again we use alternating minimization between two steps: minimization with respect to the latent image and minimization with respect to the PSFs. However this time, we solve a nonlinear problem in each step by applying a variable splitting technique to convert

the problem to constrained optimization and then using an augmented Lagrangian method to solve the constrained optimization. The augmented Lagrangian method is a fast converging method, which can solve the blind deconvolution problem in an efficient way. Examples of blind deconvolution of high resolution photos captured with a DSLR camera conclude the paper.

The main contribution of this research article is in improving accuracy of PSF estimation and providing a fast and reliable multichannel blind deconvolution algorithm that copes with high-resolution image and large blurs.

6. A. Marrugo, M. Šorel, F. Šroubek, and M. Millan, "Retinal image restoration by means of blind deconvolution," *Journal of Biomedical Optics*, vol. 16, no. 11, pp. 116016-1-11, 2011.

   This applied research article is an example of direct application of multichannel blind deconvolution illustrating a step towards computer-assisted diagnosis and telemedicine in ophthalmology. Here we present a method for color retinal image restoration by means of multichannel blind deconvolution. The method is applied to a pair of retinal images acquired within a lapse of time, ranging from several minutes to months. It consists of a series of preprocessing steps to adjust the images so they comply with the considered degradation model (3), followed by the estimation of the PSF and, ultimately, image deconvolution. The preprocessing is composed of image registration, uneven illumination compensation, and segmentation of areas with structural changes. In addition, we have developed a procedure for the detection and visualization of structural changes. This enables the identification of subtle developments in the retina not caused by variation in illumination or blur. The method was tested on synthetic and real images.

   The main purpose of this paper has been to investigate a new approach for retinal image restoration based on multichannel blind deconvolution.

7. A. Marrugo, M. Millán, M. Šorel, and F. Šroubek, "Restoration of retinal images with space-variant blur," *Journal of Biomedical Optics*, vol. 19, no. 1, 2014.

   In this paper we extend results of the previous article by introducing a method for restoring retinal images affected by space-variant blur. To do so, we described a space-variant model of blur in terms of convolution with a PSF that changes depending on its position. We show that the PSFs need not be computed for all pixels, which is quite a demanding task, but for a small set of discrete positions. For any intermediate position bilinear interpolation suffices. In this way, we achieve a space-variant representation of the PSF. The estimation of accurate local PSFs proved difficult due to the very nature of the images; they usually contain texture-less or nearly homogeneous regions that lack retinal structures, such as blood vessels, to provide sufficient information. In this regard, we propose a strategy based on eye-domain knowledge to adequately identify and correct such non-valid PSFs. Without this, the restoration results are artifact-prone with an overall image quality that is worse than the original image. The proposal has been tested on artificially and naturally degraded retinal images coming from the clinical practice. The details from the restored retinal images show an important enhancement, which is also demonstrated with the improvement in the detection of the retinal vasculature.

The main contribution of this work is to demonstrate a relatively simple but effective remedy to smoothly and slowly varying space-variant blur by applying patch-wise blind multichannel deconvolution.

8. O. Šindelář and F. Šroubek, "Image deblurring in smartphone devices using built-in inertial measurement sensors," *Journal of Electronic Imaging*, vol. 22, no. 1, pp. 011003-1-8, 2013.

   This is another example of an applied research article. The target application is photography on embedded devices. Blur induced by camera motion is a frequent problem in photography mainly when the light conditions are poor. As the exposure time increases, involuntary camera motion has a growing effect on the acquired image. Image stabilization devices that help to reduce the motion blur by moving the camera sensor in the opposite direction are becoming more common. However, such hardware remedy has its limitations as it can compensate only for motion of a very small extent and speed. Deblurring the image offline using mathematical algorithms is usually the only choice we have in order to obtain a sharp image. Motion blur can be modeled by convolution and then the deblurring process is deconvolution. Many devices, such as modern smartphones, are now equipped with inertial sensors (gyroscopes and accelerometers) that can give us a very accurate information about camera motion. If we are able to reconstruct camera path then we can recover blur and perform nonblind image deblurring. This idea was originally described in [45] but the authors have designed an expensive measuring apparatus consisting of a DSLR camera and a set of inertial sensors, and perform image deblurring offline on a computer. Our work is based on the same idea but the aim is to show that image deblurring is feasible on modern smartphones without the requirement of other devices.

   The main contribution of this work is to illustrate that blur estimation with built-in inertial sensors is possible and to implement image deblurring on a smartphone, which works in practical situations and is relatively fast to be acceptable for a general user.

4 Conclusions

The presented habilitation thesis summarizes author's contribution to the theory of blind deconvolution in the last ten years. The underlying theme linking the collection of eight publications that comprise the thesis is multichannel blind deconvolution. The presented research articles summarize gradual improvements in the field of blind deconvolution that resulted in a robust algorithm, which works with misaligned high-resolution blurry images, can cope with large blurs, and is computationally efficient using advance numerical methods. Multichannel framework of blind deconvolution is extended to resolution enhancement (superresolution), which is covered by two research articles in the collection. The applicability of the approach has been demonstrated on many practical examples. A various versions of the restoration algorithm are available for free for research purposes on the institute web pages. To date, we file over 1000 downloads of the software, which indicates a high interest of the research community in this topic. This fact is also supported by a relatively high impact of the presented collection of articles, which is around 280 citations (according to SCOPUS) in total excluding self-citations.

References

[1] M. Banham and A. Katsaggelos, "Digital image restoration," *IEEE Signal Processing Magazine*, vol. 14, no. 2, pp. 24–41, Mar. 1997.

[2] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[3] P. Campisi and K. Egiazarian, Eds., *Blind Image Deconvolution, Theory and Application*. CRC Press, 2007.

[4] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, 2nd ed. Orlando, FL, USA: Academic Press, Inc., 1982.

[5] M. Chang, A. Tekalp, and A. Erdem, "Blur identification using the bispectrum," *IEEE Trans. Signal Processing*, vol. 39, no. 10, pp. 2323–2325, Oct. 1991.

[6] J. Oliveira, M. Figueiredo, and J. Bioucas-Dias, "Parametric blur estimation for blind restoration of natural images: Linear motion and out-of-focus," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 466–477, 2014. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6637096

[7] Q. Shan, W. Xiong, and J. Jia, "Rotational motion deblurring of a rigid object from a single image," in *Proc. IEEE 11th International Conference on Computer Vision ICCV 2007*, 14–21 Oct. 2007, p. 1–8.

[8] A. Stern and N. Kopeika, "Analytical method to calculate optical transfer functions for image motion and vibrations using moments," *J. Opt. Soc. Am. A*, vol. 14, no. 2, pp. 388–396, Feb. 1997.

[9] Y. Yitzhaky and N. Kopeika, "Identification of blur parameters from motion blurred images," *Graphical Models and Image Processing*, vol. 59, no. 5, pp. 310–320, Sep. 1997.

[10] J. Oliveira, M. Figueiredo, and J. Bioucas-Dias, "Blind estimation of motion blur parameters for image deconvolution," in *Pattern Recognition and Image Analysis*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007, vol. 4478, pp. 604–611. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-72849-8_76

[11] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, 1980.

[12] G. Ayers and J.C.Dainty, "Iterative blind deconvolution method and its application," *Optical Letters*, vol. 13, no. 7, pp. 547–549, Jul. 1988.

[13] T. Chan and C. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Processing*, vol. 7, no. 3, pp. 370–375, Mar. 1998.

[14] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Transactions on Image Processing*, vol. 15, no. 12, p. 3715–3727, Dec. 2006.

[15] J. Kotera, F. Sroubek, and P. Milanfar, "Blind deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors," in *Proceedings of the 15th International Conference on Computer Analysis of Images and Patterns (CAIP)*, 2013.

[16] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*. New York, NY, USA: ACM, 2006, pp. 787–794.

[17] J. Jia, "Single image motion deblurring using transparency," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR '07*, 17–22 June 2007, pp. 1–8.

[18] N. Joshi, R. Szeliski, and D. J. Kriegman, "PSF estimation using sharp edge prediction," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR 2008*, 23–28 June 2008, pp. 1–8.

[19] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," in *SIGGRAPH '08: ACM SIGGRAPH 2008 papers*. New York, NY, USA: ACM, 2008, p. 1–10.

[20] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proceedings of the 11th European conference on Computer vision: Part I*, ser. ECCV'10. Berlin, Heidelberg: Springer-Verlag, 2010, p. 157–170. [Online]. Available: http://portal.acm.org/citation.cfm?id=1886063.1886077

[21] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR '09*, 2009, pp. 1964–1971.

[22] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind deconvolution algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2354–2367, 2011.

[23] J. Miskin and D. J. MacKay, "Ensemble learning for blind image separation and deconvolution." in *Advances in Independent Component Analysis*, M. Girolani, Ed. Springer-Verlag, 2000, pp. 123–142.

[24] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, and A. K. Katsaggelos, "Hierarchical bayesian image restoration from partially known blurs," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1784–1797, 2000.

[25] S. D. Babacan, R. Molina, M. N. Do, and A. K. Katsaggelos, "Bayesian blind deconvolution with general sparse image priors," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 341–355.

[26] L. Xu, S. Zheng, and J. Jia, "Unnatural l0 Sparse representation for natural image deblurring," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 1107–1114. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6618991

[27] F. Šroubek, V. Šmídl, and J. Kotera, "Understanding image priors in blind deconvolution," in *Proc. IEEE International Conference on Image Processing*, 2014.

[28] D. Wipf and H. Zhang, "Revisiting bayesian blind deconvolution," *Journal of Machine Learning Research (JMLR)*, November 2014, an earlier version of this paper appeared in 2013 here: http://arxiv.org/abs/1305.2362. [Online]. Available: http://research.microsoft.com/apps/pubs/default.aspx?id=192590

[29] T. Schulz, "Multiframe blind deconvolution of astronomical images," *J. Opt. Soc. Am. A*, vol. 10, no. 5, pp. 1064–1073, May 1993.

[30] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms," *IEEE Trans. Image Processing*, vol. 8, no. 2, pp. 202–219, Feb. 1999.

[31] M. I. Gurelli and C. L. Nikias, "EVAM: An eigenvector based algorithm for multichannel blind deconvolution of input colored signals," *IEEE Trans. Acous., Speech, Signal Processing*, vol. 43, p. 134–149, 1995.

[32] G. Giannakis and R. Heath, "Blind identification of multichannel FIR blurs and perfect image restoration," *IEEE Trans. Image Processing*, vol. 9, no. 11, pp. 1877–1896, Nov. 2000.

[33] H.-T. Pai and A. Bovik, "On eigenstructure-based direct multichannel blind image restoration," *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1434–1446, Oct. 2001.

[34] S. Pillai and B. Liang, "Blind image deconvolution using a robust GCD approach," *IEEE Trans. Image Processing*, vol. 8, no. 2, pp. 295–301, Feb. 1999.

[35] M. Haindl and S. Šimberová, "Model-based restoration of short-exposure solar images," in *Frontiers in Artificial Intelligence and Applications*, ser. Knowledge-Based Intelligent Engeneering Systems, L. Jain and R. Howlett, Eds., vol. 87. Amsterdam: Publisher IOS Press, 2002, pp. 697–706.

[36] G. Panci, P. Campisi, S. Colonnese, and G. Scarano, "Multichannel blind image deconvolution using the bussgang algorithm: Spatial and multiresolution approaches," *IEEE Trans. Image Processing*, vol. 12, no. 11, pp. 1324–1337, Nov. 2003.

[37] F. Šroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Trans. Image Processing*, vol. 12, no. 9, pp. 1094–1106, Sep. 2003.

[38] ——, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Processing*, vol. 14, no. 7, pp. 874–883, Jul. 2005.

[39] F. Šroubek, G. Cristobal, and J. Flusser, "A unified approach to superresolution and multichannel blind deconvolution," *IEEE Trans. Image Processing*, vol. 16, no. 9, pp. 2322–2332, Sep. 2007.

[40] H. Zhang, D. Wipf, and Y. Zhang, "Multi-observation blind deconvolution with an adaptive sparse prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 8, pp. 1628–1643, 2014. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6848885

[41] M. Tico, M. Trimeche, and M. Vehvilainen, "Motion blur identification based on differently exposed images," in *Proc. IEEE Int. Conf. Image Processing*, 2006, pp. 2021–2024.

[42] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," in *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*. New York, NY, USA: ACM, 2007, p. 1.

[43] S. Babacan, J. Wang, R. Molina, and A. Katsaggelos, "Bayesian blind deconvolution from differently exposed image pairs," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2874–2888, 2010. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5482163

[44] M. Šorel and F. Šroubek, "Space-variant deblurring using one blurred and one underexposed image," in *Proceedings of the IEEE 16th International Conference on Image Processing ICIP 2009*. IEEE, 2009, pp. 157–160.

[45] N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski, "Image deblurring using inertial measurement sensors," *ACM Trans. Graph.*, vol. 29, pp. 30:1–30:9, July 2010. [Online]. Available: http://doi.acm.org/10.1145/1778765.1778767

[46] S. Park and M. Levoy, "Gyro-based multi-image deconvolution for removing handshake blur," in *Proceedings of IEEE CVPR*, 2014, pp. 3366–3373.

[47] M. Ben-Ezra and S. Nayar, "Motion deblurring using hybrid imaging," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 1, Jun. 2003, pp. I–657 – I–664 vol.1.

[48] M. Ben-Ezra and S. K. Nayar, "Motion-based motion deblurring," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 689–698, Jun. 2004.

[49] Y.-W. Tai, H. Du, M. S. Brown, and S. Lin, "Image/video deblurring using a hybrid camera," in *CVPR*, 2008.

[50] O. Šindelář and F. Šroubek, "Image deblurring in smartphone devices using built-in inertial measurement sensors," *Journal of Electronic Imaging*, vol. 22, no. 1, 2013.

[51] O. Šindelář, F. Šroubek, and P. Milanfar, "Space-variant image deblurring on smartphones using inertial sensors," in *CVPRW 2014 : Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops 2014*. IEEE Computer Society, 2014, pp. 191–192.

[52] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, Jun. 2010, pp. 491 –498.

[53] A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions," in *Proceedings of the 11th European conference on Computer vision: Part I*, ser. ECCV'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 171–184. [Online]. Available: http://portal.acm.org/citation.cfm?id=1886063.1886078

[54] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf, "Fast removal of non-uniform camera shake," in *Proc. IEEE Int Computer Vision (ICCV) Conf*, 2011, pp. 463–470. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6126276

[55] M. Hradis, J. Kotera, P. Zemcik, and F. Sroubek, "Convolutional neural networks for direct text deblurring," in *Proceedings of The British Machine Vision Association and Society for Pattern Recognition BMVC*, 2015.

[56] J. Sun, W. Cao, Z. Xu, and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, 2015, pp. 769–777. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7298677

[57] M. Šorel and J. Flusser, "Space-variant restoration of images degraded by camera motion blur," *IEEE Transactions on Image Processing*, vol. 17, no. 2, pp. 105–116, Feb. 2008.

[58] P. Favaro, M. Burger, and S. Soatto, "Scene and motion reconstruction from defocus and motion-blurred images via anisothropic diffusion," in *ECCV 2004, LNCS 3021, Springer Verlag, Berlin Heidelberg*, T. Pajdla and J. Matas, Eds., 2004, pp. 257–269.

[59] M. Šorel and F. Šroubek, "Restoration in the presence of unknown spatially varying blur," in *Image Restoration: Fundamentals and Advances*, G. Bahadir and L. Xin, Eds. CRC Press, 2012, pp. 63–88.

[60] T. L. Williams, *The optical transfer function of imaging systems.* Institute of Physics Publishing, London, 1999.

[61] N. Joshi, R. Szeliski, and D. J. Kriegman, "PSF estimation using sharp edge prediction," *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pp. 1–8, 2008.

[62] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," in *SIGGRAPH '07: ACM SIGGRAPH 2007 papers.* New York, NY, USA: ACM, 2007, p. 1.

[63] F. Šroubek, M. Šorel, I. Horackova, and J. Flusser, "Patch-based blind deconvolution with parametric interpolation of convolution kernels," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2013, pp. 1–4.

[64] A. Chakrabarti, T. Zickler, and W. T. Freeman, "Analyzing spatially-varying blur," in *CVPR'10*, San Francisco, CA, USA, Jun. 2010, pp. 2512–2519.

[65] S. Dai and Y. Wu, "Motion from blur," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, Jun. 2008, pp. 1 –8.

[66] R. Liu, Z. Li, and J. Jia, "Image partial blur detection and classification," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, Jun. 2008, pp. 1 –8.

[67] T. H. Kim and K. M. Lee, "Segmentation-free dynamic scene deblurring," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 2014, pp. 2766–2773. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6909750

[68] A. Levin, "Blind motion deblurring using image statistics," in *Advances in Neural Information Processing Systems (NIPS)*, 2006, pp. 841–848.

[69] H. Takeda and P. Milanfar, "Removing motion blur with space–time processing," *IEEE Transactions on Image Processing*, vol. 20, no. 10, pp. 2990–3000, 2011.

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

[71] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.

[72] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.

[73] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

5 Reprints

1094 IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 12, NO. 9, SEPTEMBER 2003

Multichannel Blind Iterative Image Restoration

Filip Šroubek and Jan Flusser, *Senior Member, IEEE*

***Abstract*—Blind image deconvolution is required in many applications of microscopy imaging, remote sensing, and astronomical imaging. Unfortunately in a single-channel framework, serious conceptual and numerical problems are often encountered. Very recently, an eigenvector-based method (EVAM) was proposed for a multichannel framework which determines perfectly convolution masks in a noise-free environment if channel disparity, called co-primeness, is satisfied. We propose a novel iterative algorithm based on recent anisotropic denoising techniques of total variation and a Mumford–Shah functional with the EVAM restoration condition included. A linearization scheme of half-quadratic regularization together with a cell-centered finite difference discretization scheme is used in the algorithm and provides a unified approach to the solution of total variation or Mumford–Shah. The algorithm performs well even on very noisy images and does not require an exact estimation of mask orders. We demonstrate capabilities of the algorithm on synthetic data. Finally, the algorithm is applied to defocused images taken with a digital camera and to data from astronomical ground-based observations of the Sun.**

***Index Terms*—Conjugate gradient, half-quadratic regularization, multichannel blind deconvolution, Mumford–Shah functional, subspace methods, total variation.**

I. Introduction

BLIND restoration of an image acquired in an erroneous measuring process is often encountered in image processing but a satisfying solution to this problem has not been yet discovered. The amount of *a priori* information about degradation, i.e., the size or shape of blurs, and the noise level, determines how mathematically ill-posed the problem is. Even nonblind restoration, when blurs are available, is in general an ill-posed problem because of zeros in the frequency domain of the blurs. The single-channel (SC) blind and nonblind deconvolution in two-dimensional (2-D) have been extensively studied and many techniques have been proposed for their solution [1], [2]. They usually involve some regularization which assures various statistical properties of the image or constrains the estimated image and/or restoration filter according to some assumptions. This regularization is required to guarantee a unique solution and stability against noise and some model discrepancies. SC restoration methods that have evolved from denoising applications form a very successful branch. Anisotropic denoising techniques play a prominent role due to their inherent ability to preserve edges in images. Total variation (TV) has proved to be a good candidate for edge-preserving denoising [3]. The TV solution is associated with highly nonlinear Euler-Lagrange equations but several linearization schemes were proposed to deal with this nonlinearity: the fixed point iteration scheme [4], [5], the primal-dual method [6] or a more general half-quadratic regularization scheme proposed in [7]. Recently, a more sophisticated approach, which minimizes the Mumford–Shah energy function [8], was successfully applied to image denoising and segmentation [9]. A trivial extension into the nonblind deconvolution problem exists for all these iterative denoising techniques.

A breakthrough in understanding of blind deconvolution was the method of zero sheets proposed by Lane and Bates [10]. They have shown that the SC blind deconvolution is possible in a noise-free case. Their arguments rest on the analytical properties of the z-transform in 2-D and on the fact that 2-D polynomials are not generally factorizable. Although conceptually the zero sheets are correct, they have little practical application since the algorithm is highly sensitive to noise and prone to numerical inaccuracy for large image sizes. A famous pioneering work in blind deconvolution has been done by Ayers and Dainty [11]. (Interesting also are enhancements proposed in [12]–[14].) Their iterative method based on Wiener-like filters with the possibility to include all sorts of constraints is robust to noise but lacks any reliability, since the problem of blind deconvolution is ill-posed with respect to both the image and the blur. If the images are smooth and homogeneous, an autoregressive model can be used to describe the measuring process. The autoregressive model simplifies the blind problem by reducing the number of unknowns and several techniques were proposed for finding its solution [15]–[17]. Very promising results have been achieved with a nonnegativity and support constraints recursive inverse filtering (NAS-RIF) algorithm proposed by Kundur and Hatzinakos [2] and extensions in [18], [19]. These methods, however, work on images that contain objects of finite support and have a uniform background. The area of the object support must be determined in advance. A bold attempt [20] has been made to use the TV-based reconstruction for the blind SC problems but with dubious results as the problem is ill-posed with respect to both the image and the blur. The alternating minimization algorithm has been proposed for this purpose and Chan *et al.* [21] have verified its convergence in case of the L_2 norm of the image gradient, but not in case of the TV functional.

The knowledge of the degradation process does not have to be the only source of useful *a priori* information. Multiple acquisition that generates several differently blurred versions of one scene may provided the information. Examples of such multichannel (MC) measuring processes are not rare and include remote sensing and astronomy, where the same scene is observed at different time instants through a time-varying

Manuscript received June 28, 2002; revised March 28, 2003. This work was supported by the Grant Agency of the Czech Republic under Project 102/00/1711. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Thierry Blu.

The authors are with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, 182 08 Prague 8, Czech Republic (e-mail: sroubekf@utia.cas.cz; flusser@utia.cas.cz).

Digital Object Identifier 10.1109/TIP.2003.815260

inhomogeneous medium such as the atmosphere; electron microscopy, where images of the same sample are acquired at different focusing lengths; or broadband imaging through a physically stable medium but which has a different transfer function at different frequencies. The MC acquisition refers in general to two input/output models that differ fundamentally, and from the mathematical point of view, should be distinguished: the single-input multiple-output (SIMO) model and the multiple-input multiple-output (MIMO) model. The SIMO model is typical for one-sensor imaging under varying environment conditions, where individual channels represent the conditions at time of acquisition. The MIMO model refers to multi-sensor or broadband imaging, where the channels represent, for example, different frequency bands or resolution levels. Color images are the special case of the MIMO model. An advantage of MIMO is the ability to model cross-channel degradations which occur in the form of channel crosstalks, leakages in detectors, and spectral blurs. Many techniques for solving the MIMO problem were proposed and could be found in [22]–[25]. In the sequel, we confine ourselves to the SIMO model exclusively and any reference to the term MC denotes the SIMO model.

Nonblind MC deconvolution is potentially free of the problems arising from the zeros of blurs. The lack of information from one blur in one frequency is supplemented by the information at the same frequency from others. It follows that the blind deconvolution problem is greatly simplified by the availability of several different channels. Moreover, it is possible to estimate the blur functions directly by a simple one-step procedure and reduce the blind problem to the nonblind one if certain conditions are met. Harikumar and Bresler proposed in [26], [27] a very elegant one-step subspace procedure (EVAM) which accomplishes perfect blind restoration in a noise-free environment by finding a minimum eigenvector of a MC condition matrix. One disadvantage of EVAM is its vulnerability to noise. Even for a moderate noise level the restoration may break down. Pillai *et al.* [28] have proposed another intrinsically MC method based on the greatest common divisor which is, unfortunately, even less numerically stable. A different, also intrinsically MC, approach proposed in [29] first constructs inverse FIR filters and then estimates the original image by passing the degraded images through the inverse filters. Noise amplification also occurs here but can be attenuated to a certain extent by increasing the inverse filter order, which comes at the expense of deblurring.

The above reasoning implies that the combination of the anisotropic denoising technique with the subspace procedure could provide both the numerical stability and the necessary robustness to noise. In the paper, we thus propose an MC alternating minimization algorithm (MC-AM) which incorporates the EVAM condition matrix into the anisotropic denoising technique as an extra regularization term. We derive the algorithm for two different denoising approaches: total variation and Mumford–Shah functional; and discuss in detail linearization and discretization schemes which lead in both cases to simple equations that differ only in the construction of one particular matrix.

The rest of this paper is organized as follows. Used notation and few numerical considerations are presented in Section II. Section III provides mathematical preliminaries for the development of the algorithm, which is then described in Section IV. Results of three experiments conducted on artificial and real data, and comparisons with the simple EVAM method are given in Section V.

II. Notation and Definitions

Throughout, Ω will denote a rectangle in $\mathbb{R}^2$ (although lower or higher dimensions may be also considered) which is the definition domain of image intensity functions. All the image intensity functions will be regarded as a bounded gray-level functions of the form $u : \Omega \to [0, 1]$. $\mathbf{x} = (x, y)$ denotes location in Ω, $|\mathbf{x}| = \sqrt{x^2 + y^2}$ denotes Euclidian norm, and $||\cdot||$ denotes the norm in $L_2(\Omega)$. $|E|$ stands for the Lebesgue measure of $E \subseteq \mathbb{R}^2$ which could be considered to be equal to the area of E.

To be able to implement the proposed algorithm a proper discretization is necessary. We will follow the CCFD (cell-centered finite difference) discretization scheme [5]. A square lattice is constructed on top of Ω with a constant step h. Let m and n denote the minimum number of cells in the y and x directions, respectively, that covers the total area of Ω. A cell $c_{ij} \subseteq \Omega$ is defined as

$$c_{ij} = \{(x, y) : (i - 1/2)h \le y \le (i + 1/2)h, \\ (j - 1/2)h \le x \le (j + 1/2)h\}$$

with area $|c_{ij}| = h^2$. The cell centers are given by (x_j, y_i) and indexed (i, j), where

$$x_j = (j - 1/2)h, \quad j = 1, \ldots, n \\ y_i = (i - 1/2)h, \quad i = 1, \ldots, m.$$

The cell middle edge points are given by $(x_{j\pm1/2}, y_{i\pm1/2})$ and indexed $(i \pm 1/2, j \pm 1/2)$, where

$$x_{j\pm1/2} = x_j \pm (h/2) \\ y_{i\pm1/2} = y_i \pm (h/2).$$

Function $u(\mathbf{x})$ is then approximated by a piecewise constant function $U(\mathbf{x})$ which has a constant value u_{ij} inside the cell c_{ij}. u_{ij} is often calculated as the mean of $u(\mathbf{x})$ over the cell c_{ij} or simply the value of u at the cell center (i, j). The set of u_{ij} values fully defines the piecewise constant function $U(\mathbf{x})$ which can be thus regarded as a discrete matrix $\mathbf{U} = \{u_{ij}\}$ of size (m, n). The 2-D discrete z-transform of $\mathbf{U}$ is defined as $\tilde{U}(z_1, z_2) = \sum_{i=1}^{m} \sum_{j=1}^{n} u_{ij} z_1^{-i} z_2^{-j}$, where $z_1, z_2 \in \mathbb{C}$. Finally, $\mathbf{u} \in \mathbb{R}^{mn}$ denotes the discrete vector representation of the image function $u(\mathbf{x})$ and is obtained by lexicographically ordering u_{ij} with respect to the index pair (i, j). Any linear operator $K(\cdot)$ and operation $K(u)(\mathbf{x})$ can be thus approximated by a discrete matrix $\mathbf{K}$ and matrix-vector multiplication $\mathbf{Ku}$, respectively.

In the sequel, the symbol $*$ will denote 2-D convolution. Using the vector-matrix notation, the convolution $h * u$ is approximated by $\mathbf{C_H u}$, where $\mathbf{C_H}$ is a block Toeplitz matrix with Toeplitz blocks. If spatial periodicity of functions is assumed, standard convolution could be replaced with circular

convolution, which is represented in the discrete space by a block circular matrix with circular blocks. The Fourier transform (FT) simplifies circular matrices to diagonal matrices, and clearly, this is a very useful property which justifies the periodic assumption.

Before we proceed on, it is crucial to investigate the discretization of flux variables. Let us consider the amount v of image gradient ∇u flowing in the direction $\mathbf{n}, v(\mathbf{x},\mathbf{n}) = \langle((\partial u)/(\partial x)(\mathbf{x}),(\partial u)/(\partial y)(\mathbf{x})),\mathbf{n}\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the scalar product. The discretization of $\nabla u(\mathbf{x})$ follows the CCFD scheme. However, the normal vector $\mathbf{n}$ has a finite number of directions in the discrete space. The most simplified approximation (four-connectivity) defines only two main directions (1, 0), (0, 1) and the corresponding discrete flux v is defined at the cell middle edge points as $v((x_j,y_i),(0,1)) \approx v_{i+1/2,j} = |u_{i+1,j} - u_{i,j}|/h, v((x_j,y_i),(1,0)) \approx v_{i,j+1/2} = |u_{i,j+1} - u_{i,j}|/h$. A more accurate approximation (eight-connectivity) would include, apart from the two main directions, additional two diagonal directions (1, 1), $(-1,1)$ that define flux values at the cell corners as $v_{i+1/2,i+1/2} = |u_{i+1,j+1} - u_{i,j}|/\sqrt{2h^2}$ and $v_{i+1/2,j-1/2} = |u_{i+1,j-1} - u_{i,j}|/\sqrt{2h^2}$.

III. Mathematical Preliminaries

Consider the MC (SIMO) model that consists of P measurements of an original image u. The relation between recorded images z_p and the original image u is described by

$$z_p(\mathbf{x}) = (h_p * u)(\mathbf{x}) + n_p(\mathbf{x}), \quad \mathbf{x}\in\Omega, \quad p = 1,\dots,P \tag{1}$$

where h_p is the point spread function (PSF) of the p-th channel blur, and n_p is signal independent noise. Note, that the only known variables are z_p. As the blind deconvolution problem is ill-posed with respect to both u and h_p, a constrained minimization technique is required to find the solution of (1). Constraints considered here are very common in real acquisition processes and thus widely accepted. Assuming white noise (with diagonal correlation matrix) of zero mean and constant variance σ^2, and PSF's preserving energy, the imposed constraints take the following form:

$$\int_\Omega (h_p * u - z_p)^2\,\mathrm{d}\mathbf{x} = |\Omega|\sigma^2, \quad p = 1,\dots,P \tag{2}$$

$$\int_\Omega (z_p - u)\,\mathrm{d}\mathbf{x} = 0, \quad p = 1,\dots,P. \tag{3}$$

Let $Q(u)$ and $R(h_p)$ denote some regularization functionals of the estimated original image u and PSFs h_p, respectively. The constrained minimization problem is formulated as $\min_{u,h_p} Q(u) + R(h_p)$ subject to (1)–(3). The unconstrained optimization problem, obtained by means of the Lagrange multipliers, is to find u and h_p which minimize the functional

$$E(u,h_1,\dots,h_P) = \frac{1}{2}\sum_{p=1}^{P}||h_p * u - z_p||^2 + \lambda Q(u) + \gamma R(h_1,\dots,h_P) \tag{4}$$

where λ and γ are positive parameters which penalize the regularity of the solutions u and h_p. Constraint (3) is automatically satisfied under certain conditions as it will be clear later. For now, the crucial question is how the functionals Q and R should look like. We proceed the discussion first with possible choices for $Q(u)$ and then for $R(h_p)$.

A. Regularization Term $Q(u)$

Regularization of (1) with respect to the image function can adopt various forms. The classical approach of Tichonov chooses $Q(u) = \int_\Omega |\nabla u|^2$. The corresponding nonblind minimization problem can be easily solved using FT and is equivalent to Wiener filtering. However, this advantage is only computational, because the obtained results are poor. The functional assumes u is smooth and any discontinuities in u create ringing artifacts. In the space of bounded variation functions where TV serves as seminorm, it is possible to define correctly image gradient together with discontinuities. Therefore, the TV convex functional was proposed by Rudin *et al.* [3] as the appropriate regularization functional

$$Q_{\mathrm{TV}}(u) \equiv \int_\Omega |\nabla u|. \tag{5}$$

The associated Euler-Lagrange equations of (4) with respect to u are

$$\frac{\partial E}{\partial u} = \sum_p \mathcal{C}^*_{h_p}\left(\mathcal{C}_{h_p}(u) - z_p\right) - \lambda\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right)$$
$$\frac{\partial u}{\partial n} = 0 \quad \text{on } \partial\Omega \tag{6}$$

where $\mathcal{C}_{h_p}(\cdot) \equiv (h_p * \cdot)$ and $\mathcal{C}^*_{h_p}(\cdot)$ denotes the adjoint operator, which is in our case $\mathcal{C}^*_{h_p}(\cdot) = (h_p(-\mathbf{x}) * \cdot)$. In the second equation, $(\partial u)/(\partial n)$ is the directional derivative in the direction of the vector normal to the domain boundary $\partial\Omega$. Let us assume that the PSFs h_p are known. It was mentioned in the introduction that this equation is highly nonlinear, and moreover, not defined for $|\nabla u| = 0$. Several techniques were proposed to solve (6). We follow the linearization scheme described in [30] which is similar to the half-quadratic regularization scheme of Geman [7] and which could be easily applied to more complex functionals of the Mumford–Shah kind. The scheme introduces "an auxiliary variable" which transfers the problem to a more feasible one. Note that for every $x \in \mathbb{R}, x \neq 0, |x| = \min_{v>0}((v/2)x^2 + (1/(2v)))$ and the minimum is reached for $v = 1/|x|$. For numerical reasons, it is necessary to restrict v on a closed set $K_\epsilon = \{v\colon \epsilon \le v \le 1/\epsilon\}$. Substituting the above relation into (5), we obtain a functional of two variables

$$Q_\epsilon(u,v) = \frac{1}{2}\int_\Omega \left(v|\nabla u|^2 + \frac{1}{v}\right) \tag{7}$$

and the algorithm consists of alternating minimizations of $F_\epsilon(u,v) = \lambda Q_\epsilon(u,v) + (1/2)\sum_p ||\mathcal{C}_{h_p}(u) - z_p||^2$ over u and v. For any starting values u^0 and v^0, the steps $n \ge 1$ are

$$u^n = \arg\min_u F_\epsilon(u, v^{n-1})$$

and

$$v^n = \arg\min_{v \in K_\epsilon} F_\epsilon(u^n, v) = \min(\max(\epsilon, 1/|\nabla u^n|), 1/\epsilon). \tag{8}$$

The minimization over v is trivial and the minimization over u is also simple, since $F_\epsilon(u,v)$ is convex and quadratic with respect to u. Convergence of the algorithm to the minimizer u_ϵ of F_ϵ is proved in [30]. Moreover, it is proved that F_ϵ converges to the original functional $F(u,v) = \lambda Q_{\mathrm{TV}}(u) + (1/2)\sum_p ||\mathcal{C}_{h_p}(u) - z_p||^2$ as $\epsilon \to 0$ but in a weak sense. This weaker notion of convergence, called Γ-*convergence*, was introduced for studying the limit of variational problems. It states that if the sequence (or a subsequence) of minimizers u_ϵ convereges to some u then u is a minimizer for F and $F_\epsilon(u_\epsilon) \to F(u)$. For each case, v is given by the second equation in (8).

In the late 80s, Mumford and Shah [8] have proposed a very complex energy function designed for image segmentation which depends on the image function u and the size of discontinuity set. In order to study the energy function, a weak formulation which depends solely on u was introduced. The regularization term of the weak Mumford–Shah energy is then

$$Q_{\mathrm{MS}}(u) \equiv \int_\Omega |\nabla u|^2 + \mu \mathcal{H}^1(S_u) \tag{9}$$

where $\mathcal{H}^1$ denotes the 1-D Hausdorff measure and $S_u \in \Omega$ is the 1-D set on which u is not continuous. The gradient ∇u is defined everywhere outside S_u. What follows is derived from Chambolle [9]. Let $U(\mathbf{x})$ denotes the piecewise constant approximation of $u(\mathbf{x})$ as described in Section II. Let the set of cell centers be $C_\Omega = \{(x_j, y_i)\colon i = 1,\ldots,m; j = 1,\ldots,n\} \subset \Omega$. Consider a functional

$$Q_h(U) = h^2 \sum_{\mathbf{x}\in C_\Omega} \sum_{\substack{\boldsymbol{\xi}\in\mathbb{Z}^2 \\ \mathbf{x}+h\boldsymbol{\xi}\in\Omega}} \frac{\mu}{h} f\left(\frac{(U(\mathbf{x}) - U(\mathbf{x}+h\xi))^2}{\mu h}\right)\phi(\xi) \tag{10}$$

where $\phi : \mathbb{Z}^2 \to \mathbb{R}^+$ is even, satisfies $\phi(0) = 0$, and $\phi(e_i) > 0$ for any $i = 1, 2$ where $\{e_1, e_2\}$ is the basis of $\mathbb{R}^2$; $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a nondecreasing bounded function that satisfies $f(0) = 0, f(+\infty) = 1$, and $f'(0) = 1$. A good candidate for f is, for example, $f(t) = (2/\pi)\arctan((\pi t)/2)$. According to [9], $Q_h\Gamma$-converges to a close approximation of the weak Mumford–Shah energy (9). The proximity is chiefly influenced by the course of function ϕ. Due to the high nonconvexity in (10), the numerical computation of an exact minimizer is not guaranteed. If, in addition to the previous assumptions about f, we assume that f is concave and differentiable, we may write

$$f(x) = \min_{0\le v\le 1} xv + \psi(v) \tag{11}$$

and the minimum is reached for $v = f'(x)$. We do not have to be concerned about the shape of $\psi(v)$, since ψ will vanish in the minimization procedure. We may therefore combine (11) with (10) and obtain a functional of two variables

$$Q_h(U,V) = h^2 \sum_{\mathbf{x}\in C_\Omega} \sum_{\substack{\boldsymbol{\xi}\in\mathbb{Z}^2 \\ \mathbf{x}+h\boldsymbol{\xi}\in\Omega}} \left[\mu\frac{\psi(V(\mathbf{x},h\xi))}{h} + V(\mathbf{x},h\xi)\left|\frac{U(\mathbf{x}) - U(\mathbf{x}+h\xi)}{h}\right|^2\right]\phi(\xi) \tag{12}$$

where $V : C_\Omega \times h\mathbb{Z}^2 \to [0,1]$. The minimization algorithm is similar to (8) and consists of alternating minimizations of $F_h(U,V) = \lambda Q_h(U,V) + (1/2)\sum_p ||\mathcal{C}_{H_p}(U) - Z_p||^2$ with respect to U and V. The iteration steps are as follows:

$$U^n = \arg\min_U F_h(U, V^{n-1})$$

and

$$V^n(\mathbf{x}, h\xi) = f'\left(\frac{(U(\mathbf{x}) - U(\mathbf{x}+h\xi))^2}{\mu h}\right). \tag{13}$$

The minimization over V is straightforward and the minimization over U is a simple problem, since $F_h(U,V)$ is convex and quadratic with respect to U.

B. Regularization Term $R(h_p)$

We show regularization of (1) with respect to the blurs h_p. The discrete noise-free representation of (1) that conforms to the discretization scheme in Section II is given as follows:

$$\mathbf{Z}_p = \mathbf{H}_p * \mathbf{U}, \quad p = 1,\ldots,P \tag{14}$$

where matrices $\mathbf{Z}_p, \mathbf{H}_p$, and $\mathbf{U}$ are of size $(m_z, n_z), (m_h, n_h)$, and (m_u, n_u), respectively, regardless of the channel index p. The assumption that sizes of $\mathbf{H}_p$ are equal, is not really restrictive, since any $\mathbf{H}_p$ with a smaller size can be padded with zeros up to the size of the largest one. Clearly, $m_z = m_h + m_u - 1$ and $n_z = n_h + n_u - 1$ if full convolution is considered.

It was mentioned earlier, that an exact solution exists for noise-free MC blind systems (using the subspace method) if certain disparity of channels is guaranteed. The following assumption clarifies the disparity notion and is fundamental to the MC blind deconvolution problem.

Assumption A1: Let $\tilde{H}_p$ be the discrete z-transform of $\mathbf{H}_p$. A set of 2-D polynomials $\{\tilde{H}_p(z_1,z_2), p = 1,\ldots,P\}$ is weakly co-prime.

The polynomials $\tilde{H}_p(z_1,z_2)$ are weakly (factor) co-prime if and only if the greatest common divisor is scalar, i.e., $\tilde{H}_p(z_1,z_2) = C(z_1,z_2)\tilde{H}'_p(z_1,z_2), \forall p = 1,\ldots,P$ hold true only for a scalar factor $C(z_1,z_2) = a$. A similar notion known as strong (zero) co-primeness is defined as follows. The polynomials are strongly co-prime if and only if they do not have common zeros, i.e., there does not exist $(\zeta_1,\zeta_2) : \tilde{H}_p(\zeta_1,\zeta_2) = 0, \forall p = 1,\ldots,P$. Clearly, both notions are equivalent for 1D polynomials. However, for 2-D polynomials weak co-primeness is much less restrictive than strong co-primeness. Strong co-primeness of two 2-D polynomials is an event of measure zero, since two zero lines on the (z_1,z_2) plane intersect with probability one, but weak co-primeness in practice holds for many common deterministic filters. Strong

co-primeness is almost surely satisfied for $P \geq 3$, since three or more zero lines pass through one common point on the (z_1, z_2) plane with probability zero.

The following proposition proved in [26] is regarded as the core stone of the subspace method.

Proposition 1: If $p \geq 2$, A1 holds and $\mathbf{U}$ has at least one nonzero element, then solutions $\{\mathbf{G}_i(m_g, n_g)\}$ to

$$\mathbf{Z}_i * \mathbf{G}_j - \mathbf{Z}_j * \mathbf{G}_i = \mathbf{O}, \quad 1 \leq i < j \leq P \tag{15}$$

have the form

$$\mathbf{G}_i = \begin{cases} \mathbf{H}_i * \mathbf{K}, & \text{for } m_g \geq m_h \wedge n_g \geq n_h \\ \alpha \mathbf{H}_i, & \text{for } m_g = m_h \wedge n_g = n_h \\ \emptyset, & \text{for } m_g < m_h \vee n_g < n_h \end{cases}$$

where $\mathbf{K}$ is some factor of size $(m_g - m_h + 1, n_g - n_h + 1)$ and α is a scalar.

In the presence of noise, the situation is different and for the correct support (m_h, n_h) system (15) is not equal to zero but rather to some measurement of noise. The strategy in this case is to find the least-squares solution of (15) for $\mathbf{G}_i$. In the framework of our proposed MC blind deconvolution algorithm, we can thus define the regularization of h_p as

$$R(h_1, \ldots, h_P) = \frac{1}{2} \sum_{1 \leq i < j \leq P} \|\mathcal{C}_{z_i}(h_j) - \mathcal{C}_{z_j}(h_i)\|^2 \tag{16}$$

where $\mathcal{C}_{z_i}(\cdot) \equiv (z_i * \cdot)$. It is clear that a correct estimation of the PSF support is crucial, since the support overestimation adds some spurious factor $\mathbf{K}$ to the true solution, and even worth, the support underestimation does not have any solution. It implies, that with respect to (15), the solutions $\mathbf{G}_i$ for different overestimated supports are indistinguishable, i.e., (16) is convex but far from strictly convex. It will be clear later, that the term $\sum_i \|h_i * u - z_i\|^2$ in (4) penalizes the overestimated solutions.

After substituting for R in (4), the Euler-Lagrange equations with respect to h_p are

$$\begin{aligned} \frac{\partial E}{\partial h_p} &= \mathcal{C}_u^*(\mathcal{C}_u(h_p) - z_p) \\ &\quad - \gamma \sum_{\substack{i=1 \\ i \neq p}}^{P} (\mathcal{C}_{z_i}^* \mathcal{C}_{z_i}(h_p) - \mathcal{C}_{z_i}^* \mathcal{C}_{z_p}(h_i)), \\ \frac{\partial h_p}{\partial n} &= 0 \quad \text{on } \partial\Omega, \quad p = 1, \ldots, P \end{aligned} \tag{17}$$

where $\mathcal{C}_u(\cdot) \equiv (u * \cdot)$ and the adjoint operator is $\mathcal{C}_u^*(\cdot) = (u(-\mathbf{x}) * \cdot)$. This is a simple set of linear equations and thus finding solutions h_p is a straightforward task. The Neumann boundary condition could be omitted since the support of h_p is assumed to be much smaller then the support of u.

It should be mentioned that Proposition 1 holds only in case that the acquired images $\mathbf{Z}_p$ are of full size, i.e., convolution in (14) is full and thus $\mathbf{Z}_p$ are not cropped. This is, however, seldom true in real applications. For the cropped scenario, a similar proposition holds which is also derived in [26]. We will not discussed this proposition in detail. For our purpose, it will suffice to note that the full convolution operator in (15) must be replaced with a cropped convolution operator. Cropped convolution differs from full convolution only in the size of the definition domain. It is not defined at image boundaries where one of the convolution arguments is not fully defined, i.e., the result of full convolution $\mathbf{Z} * \mathbf{H}$ is of size $(m_z + m_h - 1, n_z + n_h - 1)$, while the result of cropped convolution is of size $(m_z - m_h + 1, n_z - n_h + 1)$ if $m_z \geq m_h, n_z \geq n_h$. Cropped convolution is thus well defined even for cropped images and the results of Proposition 1 hold. By using cropped convolution, we get for free another advantage that the Neumann boundary condition in the Euler-Lagrange equation (6) will be automatically satisfied for the convolution term in this equation. A slight computational drawback is the fact that cropped convolution cannot be diagonalized with FT anymore. Nevertheless, we will assume cropped convolution in the following discussion for the reasons given above and show efficient computation of resulting matrices.

IV. MC-AM Algorithm

From the above discussion follows that the unveiled energy function E from (4) becomes

$$\begin{aligned} E(u, h_1, \ldots, h_P) &= \frac{1}{2} \sum_{p=1}^{P} \|h_p * u - z_p\|^2 \\ &\quad + \lambda \int_\Omega |\nabla u| + \gamma \frac{1}{2} \sum_{1 \leq i < j \leq P} \|\mathcal{C}_{z_i}(h_j) - \mathcal{C}_{z_j}(h_i)\|^2 \end{aligned} \tag{18}$$

for the TV regularization and we would obtain a similar equation for the Mumford–Shah regularization. Note that $E(u, h_p)$ as a functional of several variables is not convex everywhere and allows infinitely many solutions. If (u, h_p) is a solution, then so are $(\alpha u, (1/\alpha) h_p)$ (mean-value ambiguity), $(u(\mathbf{x} \pm \xi), h_p(\mathbf{x} \mp \xi))$ (shift ambiguity) for any $\alpha \in \mathbb{R}$ and $\xi \in \mathbb{R}^2$. On the other hand, for fixed u or h_p, $E(u, h_p)$ is a convex functional of h_p or u, respectively. The AM algorithm, for some initial value u^0, alternates between the following two steps:

$$\begin{aligned} h_p^n &= \arg\min_{h_p} E(u^{n-1}, h_p) \quad \text{by (17)} \\ u^n &= \arg\min_{u} E(u, h_p^n) \quad \text{by (8) or (13)} \end{aligned} \tag{19}$$

for $n \geq 1$. A minimizer of the first minimization equation can be determined by directly solving $(\partial E)/(\partial h_p) = 0$, i.e., (17). The second minimization equation can be solved via (8) if the TV functional is considered or via (13) if the Mumford–Shah functional is considered. The mean-value ambiguity is removed by constraint (3). It will be explained at the end of this section, that this constraint is automatically satisfied in the AM algorithm. A correct setting of the blur size (m_h, n_h) alleviates the shift ambiquity. In the noise-free case, the AM algorithm transforms into the EVAM method: the first step in (19) becomes perfect blur restoration and the second step calculates the least-squares solution of the image. When noise is present, any convergence analysis is difficult to carry out but results of our experiments are satisfying and illustrate a strong stability of the algorithm.

Consider the discretization scheme described in Section II. The P-channel acquisition model (1) becomes in the discrete space

$$\mathbf{z} = \mathcal{H}\mathbf{u} + \mathbf{n} = \mathcal{U}\mathbf{h} + \mathbf{n} \tag{20}$$

where $\mathbf{h} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_P^T]^T$ and $\mathbf{z} = [\mathbf{z}_1^T, \ldots, \mathbf{z}_P^T]^T$ denote vectors of size $P m_h n_h$ and $P m_z n_z$ representing discrete, con-

catenated and lexicographically ordered h_p and z_p, respectively. Matrices $\mathcal{U}$ and $\mathcal{H}$ are defined as

$$\mathcal{U} \equiv \underbrace{\begin{pmatrix} \mathbf{C_U} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{C_U} \end{pmatrix}}_{P}, \quad \mathcal{H} \equiv \begin{pmatrix} \mathbf{C_{H_1}} \\ \vdots \\ \mathbf{C_{H_P}} \end{pmatrix} \tag{21}$$

where $\mathbf{C_U}$ and $\mathbf{C_{H_p}}$ denote cropped convolution with $\mathbf{U}$ and $\mathbf{H}_p$, respectively. The size of $\mathcal{U}$ is (Pm_zn_z, Pm_hn_h) and of $\mathcal{H}$ is (Pm_zn_z, m_un_u). If the size of the recorded images is (m_z, n_z) then the minimum size of the original image is $m_u = m_z + m_h - 1, n_u = n_z + n_h - 1$.

Suppose that $\mathcal{Z}$ is a matrix defined by the iterative prescription

$$\mathbf{S}_{P-1} \equiv \begin{pmatrix} \mathbf{C_{Z_P}} & -\mathbf{C_{Z_{P-1}}} \end{pmatrix},$$

$$\mathbf{S}_t \equiv \left(\begin{array}{c|ccc} \mathbf{C_{Z_{t+1}}} & -\mathbf{C_{Z_t}} & & \\ \mathbf{C_{Z_{t+2}}} & & -\mathbf{C_{Z_t}} & \\ \vdots & & & \ddots \\ \mathbf{C_{Z_P}} & & & -\mathbf{C_{Z_t}} \\ \hline \mathbf{O} & \multicolumn{3}{c}{\mathbf{S}_{t+1}} \end{array}\right) \quad t = P-2, P-3, \ldots, 1,$$

$$\mathcal{Z} \equiv \mathbf{S}_1 \tag{22}$$

where $\mathbf{C_{Z_t}}$ denotes cropped convolution with the image $\mathbf{Z}_t$, then the right-hand side of (16) becomes $(1/2)\|\mathcal{Z}\mathbf{h}\|^2$ and the size of $\mathcal{Z}$ is $((P(P-1))/2)(m_z - m_h + 1)(n_z - n_h + 1), Pm_hn_h)$. We assume that $\mathrm{supp}(\mathbf{Z}_p) \gg \mathrm{supp}(\mathbf{H}_p)$ for $p = 1, \ldots, P$. From Proposition 1 follows, that for the noise-free case, $\mathcal{Z}$ has full column rank $(\mathrm{rank}(\mathcal{Z}) = Pm_hn_h)$ only if the blur size is underestimated, i.e., $m_h < m_h^* \vee n_h < n_h^*$, where (m_h^*, n_h^*) is the correct blur size. For the overestimated blur size $m_h \geq m_h^* \wedge n_h \geq n_h^*$, $\mathrm{rank}(\mathcal{Z}) = Pm_hn_h - (m_h - m_h^* + 1)(n_h - n_h^* + 1)$.

In case of the modified TV functional (7), we need to consider the discretization scheme of the flux variable v. For the simple four-connectivity approximation, one obtains (23) and for the more elaborated eight-connectivity approximation (24)

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(v_{i+\frac{1}{2},j}|u_{i+1,j} - u_{i,j}|^2 + v_{i,j+\frac{1}{2}}|u_{i,j+1} \right. \\ &\quad \left. - u_{i,j}|^2 + \frac{1}{v_{i+\frac{1}{2},j}} + \frac{1}{v_{i,j+\frac{1}{2}}} \right) \\ &\quad \equiv \frac{1}{2}\mathbf{u}^T\mathcal{L}_4(v)\mathbf{u} + c(v) \end{aligned} \tag{23}$$

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(v_{i+\frac{1}{2},j}|u_{i+1,j} - u_{i,j}|^2 + v_{i,j+\frac{1}{2}}|u_{i,j+1} \right. \\ &\quad - u_{i,j}|^2 + \frac{1}{\sqrt{2}}v_{i+\frac{1}{2},j+\frac{1}{2}}|u_{i+1,j+1} - u_{i,j}|^2 \\ &\quad + \frac{1}{\sqrt{2}}v_{i+\frac{1}{2},j-\frac{1}{2}}|u_{i+1,j-1} - u_{i,j}|^2 + \frac{1}{v_{i+\frac{1}{2},j}} \\ &\quad \left. + \frac{1}{v_{i,j+\frac{1}{2}}} + \frac{1}{v_{i+\frac{1}{2},j+\frac{1}{2}}} + \frac{1}{v_{i+\frac{1}{2},j-\frac{1}{2}}} \right) \\ &\quad \equiv \frac{1}{2}\mathbf{u}^T\mathcal{L}_8(v)\mathbf{u} + c(v) \end{aligned} \tag{24}$$

where both $\mathcal{L}_4$ and $\mathcal{L}_8$ are block tridiagonal matrices formed from $v_{i\pm1/2,j\pm1/2}$ and $c(v)$ is a sum of inverse values of v. More precisely, the diagonal blocks are tridiagonal in both $\mathcal{L}$'s, and the off-diagonal blocks in $\mathcal{L}_4$ are just diagonal matrices, while in $\mathcal{L}_8$ they are tridiagonal as well. Almost identical discrete equations can be obtained for the Mumford–Shah regularization by means of (12). For instance, if $\phi \equiv 0$ except for $\xi \in \{(0,1),(1,0),(0,-1)(-1,0)\}$ where $\phi(\xi) = 1/2$ then (12) takes the form of (23) and, if in addition, $\phi(\eta) = 1/(2\sqrt{2})$ for $\eta \in \{(1,1),(-1,-1),(1,-1),(-1,1)\}$ then (12) takes the form of (24). We should not forget, however, that the difference between TV and Mumford–Shah still resides in the calculation of the flux variable $v = \varphi(u)$, e.g., from (8) follows that for TV

$$v_{i\pm1/2,j\pm1/2} = \min(\max(\epsilon, 1/|u_{i\pm1,j\pm1} - u_{i,j}|), 1/\epsilon) \tag{25}$$

and from (13) for Mumford–Shah

$$v_{i\pm1/2,j\pm1/2} = \frac{1}{1 + \left(\frac{\pi(u_{i\pm1,j\pm1} - u_{i,j})^2}{2\mu}\right)^2}. \tag{26}$$

In the vector-matrix notation, the total energy function (18) for some overestimated blur size (m_h, n_h) is

$$E_{(m_h,n_h)}(\mathbf{u},\mathbf{h}) = \lambda\mathbf{u}^T\mathcal{L}\mathbf{u} + \gamma\|\mathcal{Z}\mathbf{h}\|^2 + \|\mathcal{H}\mathbf{u} - \mathbf{z}\|^2 \tag{27}$$

where $\mathcal{L}$ stands for $\mathcal{L}_4, \mathcal{L}_8$, or any other matrix of similar form resulting from a different approximation. The flux variable v is neglected to simplify notation. Using this equation, the minimization algorithm in (19) reduces to a sequence of solutions of simple linear equations. The discrete MC-AM algorithm thus consist of the following steps.

Require: initial value $\mathbf{u}^0$, blur size (m_h, n_h), where $m_h > m_h^*, n_h > n_h^*$, and regularization parameters $\gamma > 0$ and $\lambda > 0$

1: **for** $n \geq 1$ **do**
2: $\mathbf{h}^n \leftarrow \mathrm{solve}\,[(\mathcal{U}^{n-1})^T\mathcal{U}^{n-1} + \gamma\mathcal{Z}^T\mathcal{Z}]\mathbf{h}^n = (\mathcal{U}^{n-1})^T\mathbf{z}$, {$\mathcal{U}^{n-1}$ is constructed by $\mathbf{u}^{n-1}$}
3: set $\mathbf{g}^0 = \mathbf{u}^{n-1}$ and $v^0 = \varphi(\mathbf{u}^{n-1})$
4: **for** $k \geq 1$ **do**
5: $\mathbf{g}^k \leftarrow \mathrm{solve}\,[(\mathcal{H}^n)^T\mathcal{H}^n + \lambda\mathcal{L}(v^{k-1})]\mathbf{g}^k = (\mathcal{H}^n)^T\mathbf{z}$, {$\mathcal{H}^n$ is constructed by $\mathbf{h}^n$}
6: $v^k = \varphi(\mathbf{g}^k)$, {for φ use (25) or (26)}
7: **end for**
8: $\mathbf{u}^n \leftarrow \mathbf{g}^k$
9: **end for**

The linear equation at line 2 can be solved directly since the symmetric square matrix $[(\mathcal{U}^{n-1})^T\mathcal{U}^{n-1} + \gamma\mathcal{Z}^T\mathcal{Z}]$ is of relatively small size Pm_hn_h, and is almost surly regular due to full column rank of the convolution matrix $\mathcal{U}$. Any reasonable image u is "persistently exciting," i.e., $u * h \neq 0$ for every FIR filter h of size much smaller than u. It was already mentioned that for the noise-free case, the dimension of the null space of $\mathcal{Z}$ is proportional to the overestimated blur size (m_h, n_h), more precisely the dimension is equal to $(m_h - m_h^* + 1)(n_h - n_h^* + 1)$, and any $\mathbf{g} \in \mathrm{null}(\mathcal{Z})$ takes the form $\mathbf{g} = [\mathrm{vec}\{\mathbf{K} * \mathbf{H}_1\}^T, \ldots, \mathrm{vec}\{\mathbf{K} * \mathbf{H}_P\}^T]$, where $\mathbf{K}$ is some spurious factor and $\mathbf{H}_p$ are correct PSFs of size (m_h^*, n_h^*). The spurious factor spoils the correct solution but cannot be avoided if

the exact size of blurs is not known in advance and if only $\mathcal{Z}$ is considered. It is the fundamental constraint (2) included at line 2 which penalizes the spurious factor. To see this, consider $\min_n \sum_i \|\mathbf{U}^n * \mathbf{K} * \mathbf{H}_i - \mathbf{U} * \mathbf{H}_i\|^2$ which is strictly greater than zero, unless $\mathbf{K}$ is a factor of $\mathbf{U}$, which cannot happen almost surely. Hence, the minimum is reached only for $\mathbf{U}^n = \mathbf{U}$ and $\mathbf{K}$ reduced to the 2-D delta function.

Due to the large size of each matrix, it is not feasible to compute the products $\mathcal{U}^T\mathcal{U}$ and $\mathcal{Z}^T\mathcal{Z}$ by first constructing $\mathcal{U}$ and $\mathcal{Z}$ and then doing the matrix multiplication. Fortunately, there exists a very fast direct construction method for both products. Moreover, the latter product is constructed only once at the beginning. It is easy to observe that the products consist of P^2 square blocks $\mathbf{B}_{ij}$ of size $m_h n_h, i, j = 1, \ldots, P$. In case of $\mathcal{U}^T\mathcal{U}$, only the diagonal blocks are nonzero and defined as $\mathbf{B}_{ii} = \mathbf{C}_{\mathbf{U}}^T\mathbf{C}_{\mathbf{U}}$. In case of $\mathcal{Z}^T\mathcal{Z}$, the off-diagonal blocks are defined as $\mathbf{B}_{ij} = -\mathbf{C}_{\mathbf{Z}_j}^T\mathbf{C}_{\mathbf{Z}_i}$ and the diagonal blocks $\mathbf{B}_{ii} = \sum_{k \neq i} \mathbf{C}_{\mathbf{Z}_k}^T\mathbf{C}_{\mathbf{Z}_k}$. We assume that $\mathbf{C}$ denotes cropped convolution. After some consideration, one would derive that the elements of $\mathbf{B}_{ij}$ are calculated as $b_{kl}^{ij} = \sum_{m=1}^{m_z - m_h + 1} \sum_{n=1}^{n_z - n_h + 1} z_i(m + \mu(l), n + \nu(l)) z_j(m + \mu(k), n + \nu(k))$, where z_i and z_j are elements of $\mathbf{Z}_i$ and $\mathbf{Z}_j$, respectively, and index shifts are $\mu(k) = [(k-1) \bmod m_h], \nu(k) = \lfloor (k-1)/(m_h) \rfloor$. Likewise, if z_i, z_j are replaced with u we get the elements of the diagonal blocks in $\mathcal{U}^T\mathcal{U}$. This way, one block is computed in $O((m_h n_h) m_z n_z \log(m_z n_z))$ multiplies. On contrary, the full matrix multiplication requires $O((P m_h n_h)^2 m_z n_z)$ multiplies.

The second linear equation at line 5 contains the symmetric positive semidefinite matrix $[(\mathcal{H}^n)^T\mathcal{H}^n + \lambda\mathcal{L}(v^{k-1})]$ of size $m_u n_u$. Most of the common PSFs have zeros in the frequency domain and/or very small values at higher frequencies and the resulting convolution matrices $\mathcal{H}$ are strongly ill-conditioned. Hence, the problem at line 5 is ill-posed and contains too many unknowns to be solvable by direct methods. A common approach, which we have also adopted, is to use conjugate gradient (CG) or preconditioned CG methods, see [5], [31]. The flux variable v is calculated directly by means of (25) if TV is considered or by means of (26) if Mumford–Shah regularization is considered. The relaxation parameter ϵ in (25) influences both the converge speed of the algorithm and accuracy of solutions at line 5. Refer to [4] for a discussion about how ϵ alters the convergence rate and for comparison of different numerical methods. In our experiments, we have found values around 10^{-3} the most appropriate. The parameter μ in (26) acts as a weighting factor of the discontinuity term in the Mumford–Shah functional (9). There is no straightforward estimation of the parameter's correct value and an evaluation by trial and error is probably the only choice. In our implementation, we alternate between minimizations over $\mathbf{g}$ and v only five times before returning back to line 2.

A. Convergence Properties

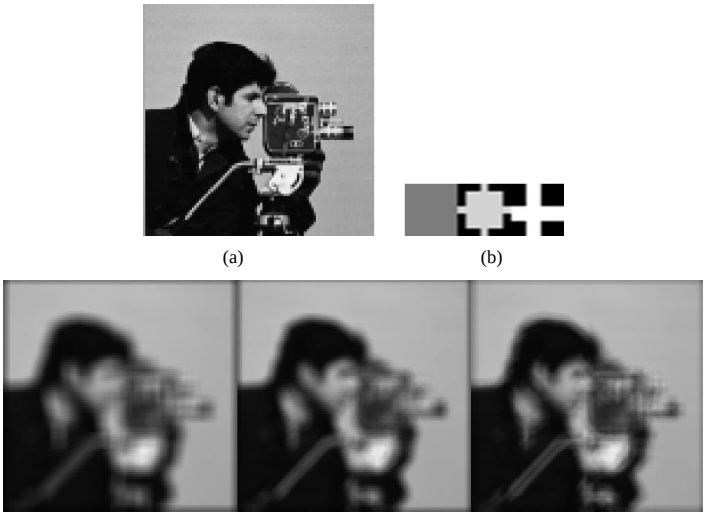
Convergence of the algorithm cannot be fully resolved on a purely theoretical basis. Nevertheless, we have made several interesting observations that rely on the fact that cropped convolution can be approximated by circular convolution and that eigenvalues of a circular convolution matrix are Fourier coefficients of the convolution mask.

Constraint (3), which was left aside at the beginning, is automatically satisfied in the algorithm if the mean values of the acquired images $\mathbf{z}_p$ and the initial estimate $\mathbf{u}^0$ are all equal, i.e., $\bar{\mathbf{u}}^0 = \bar{\mathbf{z}}_1 = \cdots = \bar{\mathbf{z}}_P$. To see this, we first approximate at line 2 cropped convolution with circular convolution and then apply FT to the equation. From the definition of $\mathcal{Z}$ in (22) and from the assumption of zero-mean noise follow, that the transformed $\mathcal{Z}^T\mathcal{Z}$ vanishes at the spatial frequency (0, 0). Since the (0, 0) frequencies refer to mean values, according to the the definition of FT, the solution $\mathbf{h}^n$ satisfies $\bar{\mathbf{h}}_p^n = 1$ if $\bar{\mathbf{u}}^{n-1} = \bar{\mathbf{z}}_p$. Likewise, if $\bar{h}_p^n = 1$, the solution $\mathbf{g}^k$ at line 5 satisfies $\bar{\mathbf{g}}^k = (1/P)\sum_{p=1}^P \bar{\mathbf{z}}_p = \bar{\mathbf{z}}_p$, since $\mathcal{L}(v^{k-1})$ has zero column-wise sums and hence vanishes at spatial frequencies $(0, \cdot)$ and $(\cdot, 0)$.

The AM algorithm is a variation on the steepest-descent algorithm. Our search space is a concatenation of the blur subspace and the image subspace. The algorithm first descends in the blur subspace and after reaching the minimum, i.e., $\nabla_{\mathbf{h}} E = 0$, it advances in the image subspace in the direction $\nabla_{\mathbf{u}} E$ orthogonal to the previous one, and this scheme repeats. To speedup the minimization, one may be tempted to implement direct set methods like Powell's that descend in arbitrary directions but this would require to solve nonlinear equations and the efficiency of such approach becomes problematic. Convergence is assured if the descent is restricted to a convex region of the functional which means that the Hessian matrix is positive semidefinite in the region. The Hessian of $E(\mathbf{u}, \mathbf{h})$ is a symmetric matrix

$$\begin{pmatrix} \nabla_{\mathbf{hh}} & \nabla_{\mathbf{hu}} \\ \nabla_{\mathbf{hu}}^T & \nabla_{\mathbf{uu}} \end{pmatrix}$$

where $\nabla_{\mathbf{hh}} = \mathcal{U}^T\mathcal{U} + \gamma\mathcal{Z}^T\mathcal{Z}, \nabla_{\mathbf{uu}} = \mathcal{H}^T\mathcal{H} + \lambda\mathcal{L}$ and the cross second derivative $\nabla_{\mathbf{hu}}$ is a combination of convolution and correlation matrices with $\mathbf{u}, \mathbf{h}, \mathbf{z}$. Let $\nabla_{\mathbf{hh}}$ and $\nabla_{\mathbf{uu}}$ be positive definite, which is true if $\mathbf{u}$ is persistently exciting and $\mathbf{h}_p$ are strongly coprime. The Hessian is then positive semidefinite if and only if $(\mathbf{x}^T\nabla_{\mathbf{hh}}\mathbf{x})(\mathbf{y}^T\nabla_{\mathbf{uu}}\mathbf{y}) \geq |\mathbf{x}^T\nabla_{\mathbf{hu}}\mathbf{y}|^2$ for all $\mathbf{x} \in \mathbb{R}^{P m_h n_h}$ and all $\mathbf{y} \in \mathbb{R}^{m_u n_u}$. If we assume that the convolution matrices can be block diagonalized with FT then the above semidefinite condition is satisfied if is satisfied for each spatial frequency alone. The multichannel term $\gamma\mathcal{Z}^T\mathcal{Z}$ is singular for each frequency and can be thus omitted. This leads us to a conclusion that this multichannel term does not directly enlarge the region of convexity. Instead, by defining mutual relations between the channel blurs, it penalizes any diversion of one blur from the rest. The necessary condition of convexity is thus expressed for each spatial frequency in each channel as $|\tilde{u}|^2(|\tilde{h}|^2 + \lambda|\tilde{l}|^2) \geq |\tilde{u}\tilde{h} + \tilde{u}\tilde{h}^* - \tilde{z}|^2$, where $\tilde{(\cdot)}$ denotes a Fourier coefficient of the corresponding signal, $\tilde{l}$ is a simplified expression that approximates eigenvalues of $\mathcal{L}$. Fundamental constraint (2) for a zero noise level takes the form $\tilde{u}\tilde{h} = \tilde{z}$ in the Fourier domain. After substituting the constraint into the above condition, we get $|\tilde{u}|^2(|\tilde{h}|^2 + \lambda|\tilde{l}|^2) \geq |\tilde{u}|^2|\tilde{h}|^2$ which is always true. In general, the condition is not satisfied only for the fundamental constraint but generates a periodic manifold that is difficult to visualize. It is important to note that the manifold size



(a) (b)

(c)

Fig. 1. (a) Original 100 × 100 cameraman image used for simulations; (b) three 7 × 7 convolution masks; and (c) blurred and noise-free images.

grows with λ, i.e., with increasing noise, convexity is guaranteed on a larger neighborhood of $\tilde{u}\tilde{h} = \tilde{z}$.

B. Estimations of Parameters γ and λ

To calculate precisely the regularization parameters is not only a tedious task but it also gives results that are of not much help in practical applications, since both parameters depend on a noise level which we usually do not know. Expressions derived here are very loose approximations that do not provide exact values but rather give a hint on the mutual relation of the parameters. Consider the equation at line 2 and let the values of $\mathbf{u}$ and $\mathbf{h}$ be equal to the original image and correct PSFs, respectively. Under the squared L^2 norm, we obtain $||\mathcal{U}^T(\mathcal{U}\mathbf{h}-\mathbf{z})||^2 = \gamma^2||\mathcal{Z}^T\mathcal{Z}\mathbf{h}||^2$, where $||\mathcal{U}\mathbf{h}-\mathbf{z}||^2 = ||\mathbf{n}||^2 \approx Pm_z n_z\sigma^2$. It is easy to verify that, if $\mathbf{n}$ is white Gaussian noise and $\mathcal{U}$ denotes convolution with $\mathbf{u}$, $||\mathcal{U}^T(\mathcal{U}\mathbf{h}-\mathbf{z})||^2 \approx Pm_h n_h||\mathbf{u}||^2\sigma^2$. Since $\mathbf{h}$ stands for the correct PSFs, it must be a linear combination of $\mathcal{Z}^T\mathcal{Z}$ eigenvectors that correspond to a cluster of minimum eigenvalues. Hence, $||\mathcal{Z}^T\mathcal{Z}\mathbf{h}||^2 = \lambda_1^2||\mathbf{h}||^2$, where λ_1 denotes the minimum eigenvalue of $\mathcal{Z}^T\mathcal{Z}$. From the definition of $\mathcal{Z}$ and Proposition 1 follow that $\lambda_1 \approx \sigma^2(P-1)m_u n_u$. Finally, we get the approximation

$$|\gamma| \approx \frac{1}{P-1}\frac{\sqrt{Pm_h n_h}}{||\mathbf{h}||}\frac{||\mathbf{u}||}{m_u n_u}\frac{1}{\sigma}. \tag{28}$$

The L^2 norms of $\mathbf{u}$ and $\mathbf{h}$ are of course not known in advance but $||\mathbf{u}||$ can be successfully approximated by $||\mathbf{z}_p||$ and if $\mathbf{h} > 0$ then $(P/(m_h n_h)) \leq ||\mathbf{h}||^2 \leq P$.

If we apply a similar procedure to the equation at line 5, we derive only the bottom limit of the regularization parameter λ. The uncertainty resides in the term $||\mathcal{L}\mathbf{u}||^2$, which cannot be simplified, since it totally depends on local behavior of the image function u. We may only formulate a generous upper limit which is $||\mathcal{L}\mathbf{u}||^2 \leq c^2||\mathbf{u}||^2$, where the constant c depends on the used approximation and the regularization term, i.e., for TV with $\mathcal{L}_4, c_4 = 4$ and for TV with $\mathcal{L}_8, c_8 = 4 + 4(1/\sqrt{2})$. The bottom limit is in general zero. Now, since $||\mathcal{H}^T(\mathcal{H}\mathbf{u}-\mathbf{z})||^2 \approx ||\mathbf{h}||^2\sigma^2 m_u n_u$, we obtain the approximated bottom limit of λ as

$$|\lambda| \gtrapprox \frac{||\mathbf{h}||\sigma}{c}\frac{\sqrt{m_u n_u}}{||\mathbf{u}||}. \tag{29}$$

The product of the parameters

$$|\gamma||\lambda| \geq \frac{1}{c(P-1)}\sqrt{\frac{Pm_h n_h}{m_u n_u}} \tag{30}$$

depends only on the dimensions of the problem and thus defines a fix relation between the parameters.

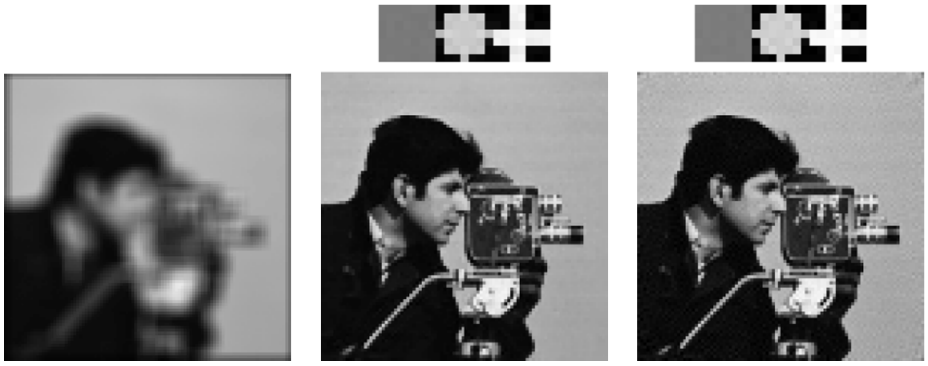
V. Experimental Results

In this section, we demonstrate the performance of our MC-AM approach on three different sets of data: simulated, real indoor and astronomical data. First, the simulated data for different SNR are used to compare results of MC-AM and EVAM. Second, the performance of MC-AM is evaluated on out-of-focus data acquired by a standard commercial digital camera. Last but not least, we demonstrate capabilities of the MC-AM approach on data from astronomical ground-based observations of the Sun.

For the evaluation of the simulated data, we use the percentage mean squared errors of the estimated PSFs $\hat{\mathbf{h}}$ and of the estimated original image $\hat{\mathbf{u}}$, respectively, defined as follows:

$$\begin{aligned}\mathrm{PMSE}(\mathbf{u}) &\equiv 100\frac{||\hat{\mathbf{u}}-\mathbf{u}||}{||\mathbf{u}||}\\ \mathrm{PMSE}(\mathbf{h}) &\equiv 100\frac{||\hat{\mathbf{h}}-\mathbf{h}||}{||\mathbf{h}||}.\end{aligned} \tag{31}$$

Both $\hat{\mathbf{u}}$ and $\hat{\mathbf{h}}$ are the outputs of MC-AM. In general, the mean squared errors do not correspond to our visual evaluation of image quality and visual comparison is often the only reliable evaluation technique. Nevertheless, the mean squared errors give us a hint how successful the restoration task was and therefore we present the calculated errors together with estimated images. In cases of the camera and astronomical



(a) (b) (c)

Fig. 2. Estimation of the cameraman image and blurs from three $\mathrm{SNR} = 50$ dB degraded images [(a) degradation with $\mathbf{h}_1$] using (b) the MC-AM algorithm and (c) the EVAM algorithm.

(a) (b) (c)

Fig. 3. Estimation of the cameraman image and blurs from three $\mathrm{SNR} = 40$ dB degraded images [(a) degradation with $\mathbf{h}_1$] using (b) the MC-AM algorithm and (c) the EVAM algorithm.

data, we use a wavelet-based focus measure [32] to compare results. It is necessary to remark that all the focus measures, which have been proposed in the literature, are easily deceived by possible artifacts which often occur in the reconstruction process. Artifacts are features (details) that were not present in original images and have been added to the images later due to erroneous image processing.

All the experiments were conducted for the TV regularization with the eight-connectivity discretization scheme. The Mumford–Shah regularization was found to produce similar results with one advantage of having a good edge detector in the flux variable v. Less advantages is the presence of the new parameter μ which influences the amount of edges. Since we were not interested in segmentation properties of the Mumford–Shah functional, the flexibility provided by μ was redundant.

A. Simulated Data

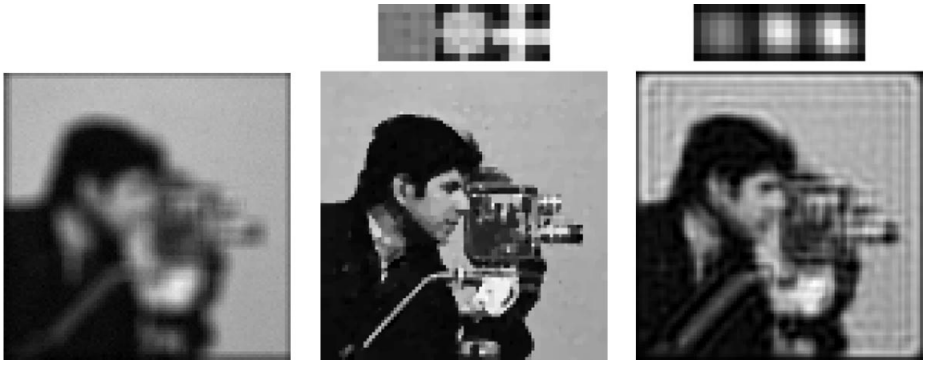
Cameraman image of size 100×100 in Fig. 1(a) was first convolved with three 7×7 masks in Fig. 1(b) and then white Gaussian noise at five different levels ($\mathrm{SNR} = 50, 40, 30, 20,$ and 10 dB) was added. This way we simulated three acquisition channels ($P = 3$) with a variable noise level that produced a series of degraded images $\mathbf{z}_1, \mathbf{z}_2$ and $\mathbf{z}_3$. The signal-to-noise ratio is calculated as usual

$$\mathrm{SNR} = 10 \log \left(\frac{\sum_{i=1}^{P} \|\mathbf{z}_i - \bar{\mathbf{z}}_i\|^2}{P m_z n_z \sigma^2} \right). \tag{32}$$

Both algorithms, our MC-AM and Harikumar's EVAM, were applied to the degraded data. The MC-AM algorithm was let to iterate over the main loop (lines 1 to 9) ten times, and within each iteration, the inner loop (lines 4 to 7) was iterated five times. The input parameters were initialized as follows: $\mathbf{u}^0 = \sum_{i=1}^{P} \mathbf{z}_i / P$; $(m_h, n_h) = (7, 7)$; λ was calculated from (29), since we know σ; and γ was estimated from the parameter product (30). Results for $\mathrm{SNR} = 50$ dB, $\mathrm{SNR} = 40$ dB, and $\mathrm{SNR} = 30$ dB are shown in Figs. 2, 3, and 4, respectively. Noise gets amplified in the EVAM reconstruction since it is not considered in the derivation of this method. The results for $\mathrm{SNR} = 30$ dB illustrate vividly this drawback. On contrary, the MC-AM algorithm is still stable even for lower SNRs (20 dB, 10 dB) as Fig. 5 demonstrates. The percentage mean squared errors of the results are summarized in Table I.

B. Real Indoor Data

Four images of a flat scene were acquired with a standard digital camera focused to 80 (objects in focus), 40, 39, and 38 cm



(a) (b) (c)

Fig. 4. Estimation of the cameraman image and blurs from three $\mathrm{SNR} = 30$ dB degraded images [(a) degradation with $\mathbf{h}_1$] using (b) the MC-AM algorithm and (c) the EVAM algorithm.

(a) (b) (c) (d)

Fig. 5. Estimation of the cameraman image and blurs from degraded images with low SNR using the MC-AM algorithm; (a)–(b) $\mathbf{h}_1$ degraded image with $\mathrm{SNR} = 20$ dB and restored image-blur pair; (c)–(d) $\mathbf{h}_1$ degraded image with $\mathrm{SNR} = 10$ dB and restored image-blur pair.

TABLE I
PERFORMANCE OF THE EVAM AND MC-AM ALGORITHMS ON SYNTHETIC DATA IN FIG. 1

SNR	EVAM		MC-AM	
	PMSE($\mathbf{h}$)	PMSE($\mathbf{u}$)	PMSE($\mathbf{h}$)	PMSE($\mathbf{u}$)
50dB	2.15	2.31	3.12	2.29
40dB	6.33	6.90	7.95	4.04
30dB	51.75	20.92	15.25	7.03
20dB	n/a	n/a	27.3	12.93
10dB	n/a	n/a	44.88	21.86

distance, respectively. The aperture was set at F2.8 and the exposure at $1/320$ s. The acquired data were stored as low resolution 480×640 24-bit color images and only the central rectangular part of the green channel of size 200×250 was considered for reconstruction. The central part of the first image, which captures the scene in focus, is shown in Fig. 6(a). Three remaining images, Fig. 6(c), were used as the input for the MC-AM algorithm. The parameter $\lambda = 1.6 \times 10^{-4}$ was estimated experimentally by running the algorithm with different λ's and selecting the most visually acceptable results. The parameter γ was calculated from (30). A defocused camera causes image degradation that is modeled by cylindrical blurs. A cepstrum analysis [33] was used to estimate diameters of these blurs, which were determined to be around 8 pixels. The size of blurs was then enlarged to 10×10 to assure inclusion of the whole cylinder. Obtained results after 10 iterations are shown in Fig. 6(b). Further iterations did not produce any visual enhancement. Simple visual comparison reveals that the letters printed on book covers are more readable in the restored image but still lack the clarity of the focused image, and that the reconstructed blurs resemble the cylindrical blurs as it was expected.

A quantitative evaluation of the amount of image blurring was done by wavelet-based focus measure [32]. The measured values, which rate the focus or the sharpness of images, are summarized in Table II. The three defocused images differ only slightly from each other and the difference is not visually detectable. However, the focus measure was able to distinguish different focus levels. It decreases as the difference from the correct focus distance increases. The focus measure of the restored image is significantly higher than the measures of the input images. It is remarkable how successful the restoration

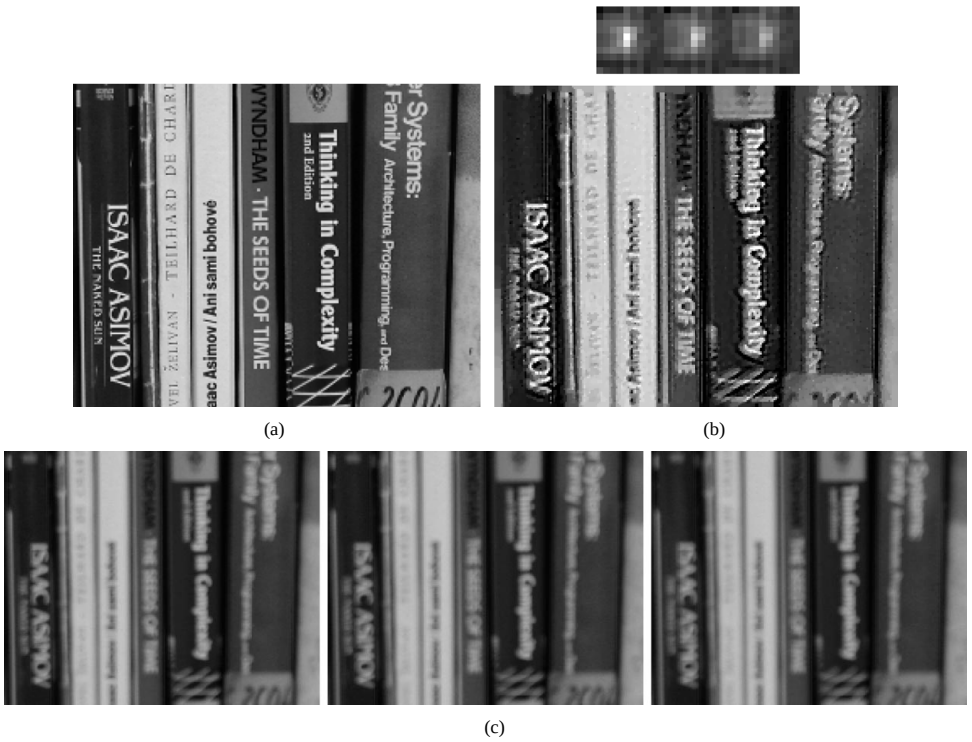


Fig. 6. Real indoor images: (a) 200×250 image acquired with the digital camera set to the correct focus distance of 80 cm; (b) MC-AM estimated image and 10×10 blurs obtained from three images (c) of false focus distances 40 cm, 39 cm, and 38 cm, after 10 iterations and $\lambda = 1.6 \times 10^{-4}$.

TABLE II
FOCUS MEASURES CALCULATED FOR THE REAL INDOOR IMAGES IN FIG. 6

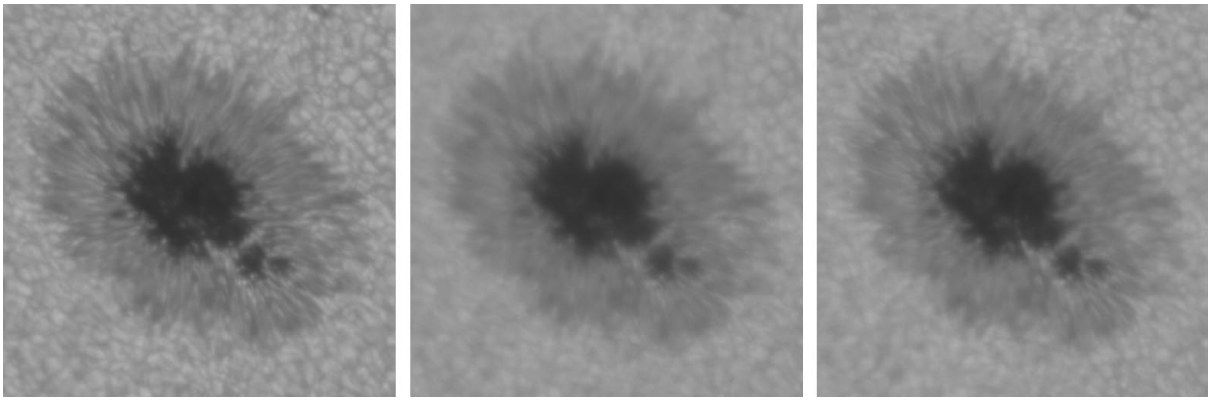
Image	focused	out of focus			restored
(focus distance)	(80cm)	(40cm)	(39cm)	(38cm)	
Focus Measure	0.3040	0.1064	0.0947	0.0859	0.2494

was, since one would expect that the similarity of blurs will violate the co-primeness assumption. It is believed that the algorithm would perform even better if a wider disparity between blurs was assured. Another interesting observation is the fact that the restored image gives a smaller response than the focused image. This is of course in agreement with our visual evaluation but it also supports a hypothesis that our restoration technique produces only few artifacts.

C. Astronomical Data

The last test which we have conducted was on real astronomical data obtained in the observation of the Sun. In the ground-based observations, the short-exposure images from the telescope are corrupted by "seeing." This degradation leads to image blurring, where the actual PSF is a composition of the intrinsic PSF of the telescope (which is constant over the observation period) and of a random component describing the perturbations of the wavefronts in the Earth's atmosphere. Different parts of the solar atmosphere are observed in different spectral bands. The lower part called photosphere is usually observed in visible light of $\lambda = 590$ nm while the medium part called chromosphere is best to observe in H_α ($\lambda = 656.3$ nm) wavelength. In visible light the effects of fluctuations in the refractive index of the air caused by temperature variations are more significant than in H_α. Since the atmospheric conditions may change very quickly, the acquired image sequence usually contains images of different quality from almost sharp to heavy blurred ones. Such sequence, which is a result of one observation session, may consist of several tens (or even hundreds) of images. Multichannel blind deconvolution is the way how to fuse the individual images of low quality to obtain one (or a few) "optimal" images which can be used for further investigation of astronomical phenomena.

In this experiment, we processed a sequence of images of a sunspot. Since the images were taken shortly one after another they are almost perfectly registered. The random nature of the atmospheric turbulence provides the necessary co-primeness of the individual PSFs. The least degraded image from the sequence, which is shown in Fig. 7(a), was selected as a reference image. Two other images of medium degradation, Fig. 7(b) and (c), were used as the input of the algorithm. The size of blurring masks was set to 12×12 which was believed to be large enough to contain the original blurring functions. The parameter λ was set to 10^{-4} which corresponds to $\mathrm{SNR} \approx 40$ dB and which is the expected noise level for this type of images. The restored image



(a) (b) (c)

Fig. 7. Astronomical data: (a) the least degraded 500×500 image of the sunspot from the sequence acquired with the terrestrial telescope (reference); (b)–(c) two blurred images from the sequence used for the reconstruction.

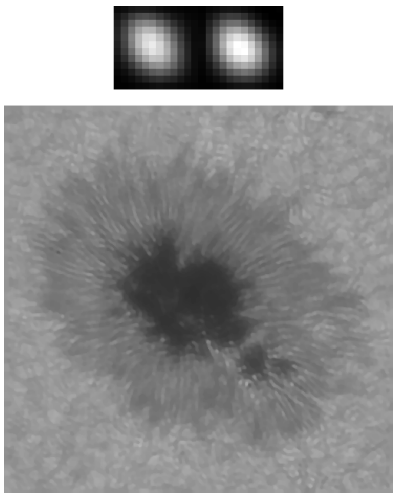


Fig. 8. Astronomical data: MC-AM reconstructed sunspot and 12×12 blurs with $\lambda = 10^{-4}$.

TABLE III
FOCUS MEASURES CALCULATED FOR THE SUNSPOT IMAGES IN FIGS. 7 AND 8

Image	reference	2 blurred (input)		restored (output)
Focus Measure	0.0149	0.0102	0.0112	0.0184

in Fig. 8 was obtained after three iterations of the MC-AM algorithm. It is worth noting that the used data are far from being "ideal" for the application of the MC-AM algorithm—there are only two channels, and their degradations are of similar nature. Nevertheless, the results are encouraging. By visual assessment, the restored image is clearly sharper than the two input images, contains no (or few) artifacts and its quality is comparable to the reference image. As in the previous experiment, we asses the quality also by quantitative focus measure (see Table III). The focus measure of the restored image is significantly higher than that of the input images and even slightly higher than the measure of the reference image. Along with the visual assessment, this illustrates a good performance of our method in this case.

VI. CONCLUSION

We have developed the algorithm for multichannel blind image restoration which combines the benefits of the edge preserving denoising techniques and the one-step subspace (EVAM) reconstruction method. This has been achieved by utilizing the multichannel EVAM constraint as a regularization term in the anisotropic denoising framework of total variation or the Mumford–Shah functional. The fundamental assumption is the weak co-primeness of blurs which guarantees the appropriate level of channel disparity and assures perfect restoration in a noise-free environment. The only input parameters, that are required, are the minimum order (size) of blurs and the noise level in the acquisition system. However, exact values of these parameters are not really needed and a rough estimate by trial and error is usually sufficient.

It was shown that the proposed algorithm gives satisfying results, compared to EVAM, even for low SNRs around 30 dB. This indicates that the denoising scheme significantly stabilizes the restoration process. The channel co-primeness is a mild condition especially in real applications, since the necessary channel disparity is probably always satisfied by random processes intrinsic to a given acquisition system. For example in case of the astronomical data, atmospheric turbulence is often modeled by Gaussian masks. In theory, any two Gaussian masks have a common nontrivial factor, but the algorithm was still able to recover the image, since small fluctuations in PSFs assured the co-prime condition.

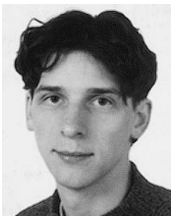
Although we have not addressed the question of computational complexity directly, we have demonstrated the ability of the algorithm to recover images of moderate size 500×500 with blurs up to 20×20.

We have not explored the influence of the blur order overestimation on image reconstruction and on convergence of the algorithm. A crucial issue for successful reconstruction, which to our knowledge has not been so far discussed in the literature,

is the spatial alignment of channels. In real applications, the channel misalignment occurs very frequently and therefore channel registration must precede the MC restoration task. Clearly by shifting the mask centers, we can compensate to a certain extent for small translation misalignments. It is expected that the overestimated blur orders provide the necessary freedom which nullifies such misalignments by automatically offsetting the blurs centers during the reconstruction process. The influence of the misregistration and the role of the order overestimation are matters for debate and will be considered in our future research.

REFERENCES

[1] M. Banham and A. Katsaggelos, "Digital image restoration," *IEEE Signal Processing Mag.*, vol. 14, pp. 24–41, Mar. 1997.
[2] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Mag.*, vol. 13, pp. 43–64, May 1996.
[3] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
[4] C. Vogel and M. Oman, "Iterative methods for total variation denoising," *SIAM J. Sci. Comput.*, vol. 17, no. 1, pp. 227–238, Jan. 1996.
[5] ——, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Image Processing*, vol. 7, pp. 813–824, June 1998.
[6] T. Chan, G. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 20, no. 6, pp. 1964–1977, July 1999.
[7] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 367–383, Mar. 1992.
[8] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.
[9] A. Chambolle, "Finite-differences discretizations of the Mumford-Shah functional," *RAIRO Math. Model Numer. Anal.*, vol. 33, no. 2, pp. 261–288, 1999.
[10] R. Lane and R. Bates, "Automatic multichannel deconvolution," *J. Opt. Soc. Amer. A*, vol. 4, no. 1, pp. 180–188, Jan. 1987.
[11] G. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its application," *Opt. Lett.*, vol. 13, no. 7, pp. 547–549, July 1988.
[12] N. Miura and N. Baba, "Segmentation-based multiframe blind deconvolution of solar images," *J. Opt. Soc. Amer. A*, vol. 12, no. 6, pp. 1858–1866, Sept. 1995.
[13] N. Miura, S. Kuwamura, N. Baba, S. Isobe, and M. Noguchi, "Parallel scheme of the iterative blind deconvolution method for stellar object reconstruction," *Appl. Opt.*, vol. 32, no. 32, pp. 6514–6520, Nov. 1993.
[14] R. Lane, "Blind deconvolution of speckle images," *J. Opt. Soc. Amer. A*, vol. 9, no. 9, pp. 1508–1514, Sept. 1992.
[15] R. Lagendijk, J. Biemond, and D. Boekee, "Identification and restoration of noisy blurred images using the expectation-maximization algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, July 1990.
[16] S. Reeves and R. Mersereau, "Blur identification by the method of generalized cross-validation," *IEEE Trans. Image Processing*, vol. 1, pp. 301–311, July 1992.
[17] A. Rajagopalan and S. Chaudhuri, "A recursive algorithm for maximum likelihood-based identification of blur from multiple observations," *IEEE Trans. Image Processing*, vol. 7, pp. 1075–1079, July 1998.
[18] C. Ong and J. Chambers, "An enhanced NAS-RIF algorithm for blind image deconvolution," *IEEE Trans. Image Processing*, vol. 8, pp. 988–992, July 1999.
[19] M. Ng, R. Plemmons, and S. Qiao, "Regularization of RIF blind image deconvolution," *IEEE Trans. Image Processing*, vol. 9, pp. 1130–1138, June 2000.
[20] T. Chan and C. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Processing*, vol. 7, pp. 370–375, Mar. 1998.
[21] ——, "Convergence of the alternating minimization algorithm for blind deconvolution," *Linear Algebra Applicat.*, vol. 316, no. 1–3, pp. 259–285, Sept. 2000.
[22] W. Zhu, N. Galatsanos, and A. Katsaggelos, "Regularized multichannel restoration using cross-validation," *Graph. Models Image Process.*, vol. 57, no. 1, pp. 38–54, Jan. 1995.
[23] B. Tom, K. Lay, and A. Katsaggelos, "Multichannel image identification and restoration using the expectation-maximization algorithm," *Opt. Eng.*, vol. 35, no. 1, pp. 241–254, Jan. 1996.
[24] A. Katsaggelos, K. Lay, and N. Galatsanos, "A general framework for frequency domain multi-channel signal processing," *IEEE Trans. Image Processing*, vol. 2, pp. 417–420, July 1993.
[25] M. Kang, "Generalized multichannel image deconvolution approach and its applications," *Optical Engineering*, vol. 37, no. 11, pp. 2953–2964, Nov. 1998.
[26] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms," *IEEE Trans. Image Processing*, vol. 8, no. 2, pp. 202–219, Feb. 1999.
[27] ——, "Efficient algorithms for the blind recovery of images blurred by multiple filters," in *Proceedings of ICIP 96*, vol. 3, Lausanne, Switzerland, 1996, pp. 97–100.
[28] S. Pillai and B. Liang, "Blind image deconvolution using a robust GCD approach," *IEEE Trans. Image Processing*, vol. 8, pp. 295–301, Feb. 1999.
[29] G. Giannakis and R. Heath, "Blind identification of multichannel FIR blurs and perfect image restoration," *IEEE Trans. Image Processing*, vol. 9, pp. 1877–1896, Nov. 2000.
[30] A. Chambolle and P. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.*, vol. 76, no. 2, pp. 167–188, Apr. 1997.
[31] R. Chan, T. Chan, and C.-K. Wong, "Cosine transform based preconditioners for total variation deblurring," *IEEE Trans. Image Processing*, vol. 8, pp. 1472–1478, Oct. 1999.
[32] J. Kautsky, J. Flusser, B. Zitová, and S. Šimberová, "A new wavelet-based measure of image focuse," *Pattern Recognit. Lett.*, vol. 23, pp. 1785–1794, 2002.
[33] M. Chang, A. Tekalp, and A. Erdem, "Blur identification using the bispectrum," *IEEE Trans. Signal Processing*, vol. 39, pp. 2323–2325, Oct. 1991.



Filip Šroubek received the B.Sc. and M.Sc. degrees in computer science from the Czech Technical University, Prague, Czech Republic, in 1996 and 1998, respectively, and is currently pursuing the Ph.D. degree in computer science at the Charles University, Prague.

Since 1999, he has been with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague. Since 2000, he has been with the Institute of Radiotechnique and Electronics, Academy of Sciences of the Czech Republic, Prague. His current research interests include all aspects of digital image processing and pattern recognition, particularly multichannel blind deconvolution, image denoising, image registration, and computer simulation and visualization of atomic collision processes.

Jan Flusser (M'94–SM'02) received the M.Sc. degree in mathematical engineering from the Czech Technical University, Prague, Czech Republic, in 1985, the Ph.D. degree in computer science from the Czechoslovak Academy of Sciences in 1990, and the D.Sc. degree in technical cybernetics in 2001.

Since 1985, he has been with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague. Since 1995, he has held the position of a head of Department of Image Processing. Since 1991, he has been also affiliated with the Charles University, Prague, and the Czech Technical University, Prague, where he teaches courses on digital image processing and pattern recognition. His current research interests include all aspects of digital image processing and pattern recognition, namely 2-D object recognition, moment invariants, blind deconvolution, image registration and image fusion. He has authored and coauthored more than 80 research publications in these areas.

874 IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 14, NO. 7, JULY 2005

Multichannel Blind Deconvolution of Spatially Misaligned Images

Filip Šroubek and Jan Flusser, *Senior Member, IEEE*

***Abstract*—Existing multichannel blind restoration techniques assume perfect spatial alignment of channels, correct estimation of blur size, and are prone to noise. We developed an alternating minimization scheme based on a maximum *a posteriori* estimation with *a priori* distribution of blurs derived from the multichannel framework and *a priori* distribution of original images defined by the variational integral. This stochastic approach enables us to recover the blurs and the original image from channels severely corrupted by noise. We observe that the exact knowledge of the blur size is not necessary, and we prove that translation misregistration up to a certain extent can be automatically removed in the restoration process.**

***Index Terms*—Image restoration, maximum *a posteriori* (MAP) estimator, multichannel blind deconvolution, subspace methods, variational integral.**

I. Introduction

In MANY applications, such as microscopy imaging, remote sensing, and astronomical imaging, observed images are often degraded by blurring. Examples of the most common sources of blur are atmospheric turbulence, relative motion between a camera, and an object or wrong focus. Restoration of the degraded images is a necessary step that precedes further image analysis.

First, a proper mathematical model that simulates the acquisition system is required. Images may be regarded as either deterministic or stochastic signals, blurred by linear or nonlinear processes and corrupted with additive or multiplicative noise. In the sequel, we assume a linear filter model with additive uncorrelated noise, i.e.

$$z(x) = (h * u)(x) + n(x) \tag{1}$$

where z, h, u, and n are the degraded image, system PSF (blur), original image, and noise, respectively, and $*$ denotes convolution. This model accurately describes many common degradations and that justifies its frequent use.

The amount of *a priori* information about the degradation, like the size or shape of blurring functions and the noise parameters, significantly influences the success of restoration. When the blur function is known, many conventional approaches have been developed to compensate for the distortion [1]. The problem is ill posed, and, to overcome this difficulty, it is common to use regularization. When the blur is unknown, we talk about *blind image restoration*. A basic survey of different blind restoration techniques is given in [2]. Most of the methods are iterative or recursive. They involve regularization terms based on available prior information which assure various statistical properties of the image and constrains the estimated image and/or restoration filter. As in the nonblind case, regularization is required to improve stability. For images with sharp changes of intensity, the appropriate regularization is based on variational integrals. A special case of the variational integral, total variation, was first proposed in [3]. Minimization of the variational integrals preserves edges and fine details in the image and it was applied to image denoising [4]–[6] and to blind restoration [7]–[9], as well. Since the blind case is strongly ill posed, all the methods suffer from convergence and stability problems. If the images are smooth and homogeneous, an autoregressive model can be used to describe the measuring process. The autoregressive model simplifies the blind problem by reducing the number of unknowns and several techniques were proposed for finding its solution [10]–[12].

There are many applications, where different blurred versions of the same original image are observed through multiple acquisition channels. We distinguish in general two multichannel (MC) models: the single-input multiple-output (SIMO) model and the multiple-input multiple-output (MIMO) model. The SIMO model (see Fig. 1) is typical for one-sensor imaging under varying environment conditions, where individual channels represent the conditions at time of acquisition. The MIMO model refers, for example, to multisensor imaging, where the channels represent different spectral bands or resolution levels. Color images are the special case of the MIMO model. An advantage of MIMO is the ability to model cross-channel degradations which occur in the form of channel crosstalks, leakages in detectors, and spectral blurs. Many techniques for solving the MIMO problem were proposed and could be found in [13]–[16]. In the sequel, we confine ourselves to the SIMO model exclusively and any reference to the term MC denotes the SIMO model. Sometimes the SIMO model is referred to in the literature as a *multiframe model*. Following the above notation, we define the SIMO model as:

$$z_k(x) = (h_k * u)(x) + n_k(x), \quad k = 1, \ldots, K \tag{2}$$

where K is the number of channels. Examples of such MC measuring processes are common, e.g., in remote sensing and astronomy, where the same scene is observed at different time instants through a time-varying inhomogeneous medium such as the atmosphere; in confocal microscopy, where images of

Manuscript received January 27, 2004; revised July 23, 2004. This work was supported by the Grant Agency of the Czech Republic under Grant 102/04/0155. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Robert D. (G. E.) Nowak.

The authors are with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, 182 08 Prague 8, Czech Republic (e-mail: sroubekf@utia.cas.cz; flusser@utia.cas.cz).

Digital Object Identifier 10.1109/TIP.2005.849322

the same sample are acquired at different focusing lengths; or in broadband imaging through a physically stable medium, but which has a different transfer function at different frequencies. Nonblind MC restoration is potentially free of the problems arising from the zeros of blurs. The lack of information from one blur in one frequency can be supplemented by the information at the same frequency from the others. Intuitively, one may expect that the blind restoration problem is also simplified by the availability of different channels. Two classes of MC blind image restoration algorithms exist. Extensions of single-channel blind restoration approaches form the first class, but since they suffer from similar drawbacks as their single-channel counterparts, they are of not much interest. The other class consists of intrinsic MC approaches and will be considered here.

One of the earliest intrinsic multichannel blind deconvolution (MBD) methods [17] was designed particularly for images blurred by atmospheric turbulence. Harikumar *et al.* [18] proposed an indirect algorithm, which first estimates the blur functions and then recovers the original image by standard nonblind methods. The blur functions are equal to the minimum eigenvector of a special matrix constructed by the blurred images. Necessary assumptions for perfect recovery of the blur functions are noise-free environment and channel coprimeness, i.e., a scalar constant is the only common factor of the blurs. Giannakis *et al.* [19] (and at the same time Harikumar *et al.* [20]) developed another indirect algorithm based on Bezout's identity of coprime polynomials which finds restoration filters and by convolving the filters with the observed images recovers the original image. Both algorithms are vulnerable to noise and even for a moderate noise level restoration may break down. In the latter case, noise amplification can be attenuated to a certain extent by increasing the restoration filter order, which comes at the expense of deblurring. Pai *et al.* [21], [22] suggested two MC restoration algorithms that, contrary to the previous two indirect algorithms, estimate directly the original image from the null space or from the range of a special matrix. Another direct method based on the greatest common divisor was proposed by Pillai *et al.* in [23]. In noisy cases, the direct algorithms are more stable than the indirect ones. Interesting approaches based on the ARMA model are given in [24], [25]. MC blind deconvolution based on the Bussgang algorithm was proposed in [26], which performs well on spatially uncorrelated data, such as binary text images and spiky images. Most of the algorithms lack the necessary robustness since they do not include any noise assumptions (except ARMA and Bussgang) in their derivation and miss regularization terms. Recently, we have proposed an iterative MC algorithm [27] that performs well even on noisy images. It is based on least-squares deconvolution by anisotropic regularization of the image and between-channel regularization of the blurs.

Unfortunately, all the above mentioned multichannel blind deconvolution methods contain two ultimate but unrealistic assumptions. They require exact knowledge of the PSFs support size and individual channels are supposed to be perfectly spatially aligned (*registered*). These strong assumptions are seldom true in practice and in fact they have prevented the usage of multichannel blind deconvolution methods in real applications.

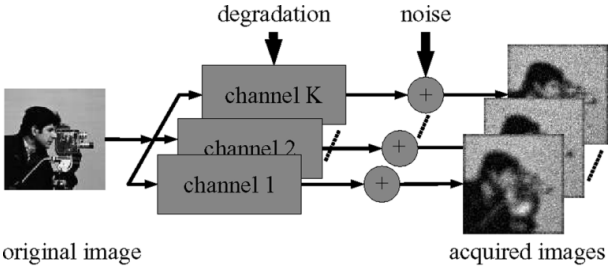


Fig. 1. Single-input multiple-output model: The original scene is captured by K different channels which are subject to various degradations.

A realistic image acquisition model contains spatial coordinate transformations τ_k that describe geometric differences between the original scene and the kth channel

$$z_k(\tau_k(x)) = (h_k * u)(x) + n_k(x), \quad k = 1, \ldots, K. \tag{3}$$

Image deformations τ_k originate from the fact that the channels are two-dimensional (2-D) projections of the three-dimensional world, generally acquired from different viewpoints and/or with different camera orientation. In simple cases, τ_k is limited to rotation and translation, but, in real applications, complex nonlinear deformations may be present, too. Image restoration then consists of two stages: image registration, which brings the channels into spatial alignment, followed by multichannel blind deconvolution.

There have been published hundreds of image registration methods (see [28] for the most recent survey) and even special algorithms for registering blurred channels were developed [29]–[34]. Despite this effort, perfect registration accuracy can rarely be achieved, namely in the case of blurred and noisy images. The registration error results in a slight between-channel shift of up to a few pixels, which cannot be further compensated in the registration stage. Thus, channel-to-channel registration simplifies (3) to the form

$$z_k(x + a_k) = (h_k * u)(x) + n_k(x), \quad k = 1, \ldots, K \tag{4}$$

where a_k is a small unknown alignment error.

Model (4) also applies to numerous practical tasks directly without the preceding registration process. Such situations typically occur when the camera is subject to vibrations or in multitemporal imaging when the camera slightly moves between consecutive channel acquisitions and/or the scene is not perfectly still.

The deconvolution methods mentioned above cannot restore images degraded according to model (4). If they were applied, the channel misregistrations would lead to strong artifacts in the restored image.

In this paper, we introduce the first MBD method which does not require perfect alignment of the individual channels and the knowledge of the blur size. Exploiting the stochastic model and the Bayes rule in Section IV, we express the *a posteriori* probability of the original image in terms of the conditional probability and two *a priori* probabilities, which are derived from properties of bounded variation functions and from the MC framework. An alternating minimization (AM) algorithm

as a solution to a maximum *a posteriori* probability (MAP) estimator is also given here. In Section V, we examine the minimization algorithm for its ability to alleviate the blur-oversized problem and demonstrate its convergence properties. We illustrate that the channel misalignment can be perfectly neutralized by properly oversizing the blur support in Section VI.

II. Notation

We use the following conventions throughout this paper:

$\mathbb{N}^2$	2-D space of integers;
$u:\mathbb{N}^2\to\mathbb{R}$	image function with a finite rectangular support;
$S_u\equiv(S_u^1,S_u^2)$	support size of the image u;
$x=(i,j)\in\mathbb{N}^2$	position at the ith row and the jth column in the image;
$u(x)\equiv u(i,j)$	image value at the position x;
$\mathbf{u}\equiv$	$[u(1,1),u(2,1),\ldots,u(S_u^1,1),u(1,2),\ldots,u(S_u^1,S_u^2)]^T$ image column vector, lowercase bold letters;
$\mathbf{C}$	matrix, uppercase bold letters;
$\|\cdot\|$	l^2 norm.

We endow the vector space $\mathbb{N}^2$ with the following operators:

"+" and "−"	defined in a standard way;
$(i,j)+k$	abbreviated form for $(i,j)+(k,k)$;
$\prod(i,j)$	$\equiv i\cdot j$;
$(i,j)<(k,l)$	$\equiv\{i<k\wedge j<l\}$ and similarly other binary relations ">," "=," etc.

For our next discussion, it is necessary to define convolution with a variable output support in matrix-vector notation. We follow the definition in [18]. Let $h(x)$ and $u(y)$ be two images with support $1\le x\le S_h=(S_h^1,S_h^2)$ and $1\le y\le S_u=(S_u^1,S_u^2)$, respectively, and $A=(a_1,a_2),B=(b_1,b_2)$ define an arbitrary output rectangle, where $A\le B$. We separate h column-wise and address individual columns as $h_1,\ldots,h_{S_h^2}$. We denote by $\mathbf{C}_{S_u}^{A,B}\{h\}$ a Toeplitz-block-Toeplitz matrix of size $\prod(B-A+1)\times\prod S_u$ such that the concatenated result of convolution $[\sum_x h(y-x)u(x)]_{(A\le y\le B)}$ is equal to $\mathbf{C}_{S_u}^{A,B}\{h\}\mathbf{u}$. This is given by

$$\mathbf{C}_{S_u}^{A,B}\{h\}=\underbrace{\begin{pmatrix}\mathbf{D}_{S_u^1}^{a_1,b_1}\{h_{a_2}\} & \cdots & \mathbf{D}_{S_u^1}^{a_1,b_1}\{h_{a_2-S_u^2+1}\}\\ \mathbf{D}_{S_u^1}^{a_1,b_1}\{h_{a_2+1}\} & \cdots & \mathbf{D}_{S_u^1}^{a_1,b_1}\{h_{a_2-S_u^2+2}\}\\ \vdots & \vdots & \vdots\\ \mathbf{D}_{S_u^1}^{a_1,b_1}\{h_{b_2}\} & \cdots & \mathbf{D}_{S_u^1}^{a_1,b_1}\{h_{b_2-S_u^2+1}\}\end{pmatrix}}_{S_u^2\ \text{blocks}}$$

and

$$\mathbf{D}_s^{a,b}\{h_j\}=\begin{pmatrix}h(a,j) & h(a-1,j) & \cdots & h(a-s+1,j)\\ h(a+1,j) & h(a,j) & \cdots & h(a-s+2,j)\\ \vdots & \vdots & \vdots & \vdots\\ h(b,j) & h(b-1,j) & \cdots & h(b-s+1,j)\end{pmatrix}\tag{5}$$

where $h(i,j)=0$ for $(i<1)\vee(i>S_h^1)\vee(j<1)\vee(j>S_h^2)$ and the size of $\mathbf{D}_s^{a,b}\{h_j\}$ is $(b-a+1)\times s$.

III. Problem Formulation

We first define the SIMO degradation model in the discrete domain $\mathbb{N}^2$ as follows. Suppose that an original (input) image $u(x)$ has support $1\le x\le S_u$. The input image propagates through K different channels that behave as linear filters each with a finite impulse response (blurs) $h_k,k\in\{1,\ldots,K\}$. Let the maximum support of the blurs be S_h. In each channel, the image is further degraded with additive white Gaussian noise (AWGN) n_k of zero mean and variance σ_k^2 and shifted by $t_k\in\mathbb{N}^2,t_k\ge 0$. Let S_t denote the maximum observed shift. On the output, we receive degraded and shifted images $z_k(x)$ with minimum support $1\le x\le S_z$, where $S_z=S_u-S_h-S_t+1$. The whole model can be expressed as

$$z_k(x)=\sum_{1\le y\le S_u-S_h+1}\delta_{t_k}(x-y+S_t+1)\times\sum_{1\le z\le S_u}h_k(y-z+S_h)u(z)+n_k(x)$$

where $x,y,z\in\mathbb{N}^2$, and δ_{t_k} is the delta function at S_t-t_k+1. By concatenating columns of the images, we can rewrite the previous equation in matrix-vector notation as

$$\mathbf{z}_k=\mathbf{T}_k\mathbf{H}_k\mathbf{u}+\mathbf{n}_k$$

where $\mathbf{z}_k,\mathbf{u}$, and $\mathbf{n}_k$ are corresponding column image vectors. $\mathbf{T}_k=\mathbf{C}_{S_u-S_h+1}^{S_t+1,S_u-S_h+1}\{\delta_{t_k}\}$ is of size $\prod(S_u-S_h-S_t+1)\times\prod(S_u-S_h+1)$ and $\mathbf{H}_k=\mathbf{C}_{S_u}^{S_h,S_u}\{h_k\}$ is of size $\prod(S_u-S_h+1)\times\prod S_u$. Both matrices are constructed according to (5) and perform discrete convolution with the reduced output size. We refer to this type of convolution as "valid," since the result is defined only on the area where both convolution arguments are properly defined. It is easy to verify that the matrix product $\mathbf{T}_k\mathbf{H}_k=\mathbf{G}_k$ denotes "valid" convolution with a mask $g_k(x)=h_k(x-t_k)$ of size $S_g=S_h+S_t$. This mask is a shifted version of the original blur h_k. By concatenating the output vectors $\mathbf{z}\equiv[\mathbf{z}_1^T,\ldots\mathbf{z}_K^T]^T$ and the shifted blur vectors $\mathbf{g}\equiv[\mathbf{g}_1^T,\ldots,\mathbf{g}_K^T]^T$, the MC model can be rewritten in two equivalent forms

$$\mathbf{z}=\mathbf{G}\mathbf{u}+\mathbf{n}=\mathbf{U}\mathbf{g}+\mathbf{n}\tag{6}$$

where $\mathbf{G}\equiv[\mathbf{G}_1^T,\ldots,\mathbf{G}_K^T]^T,\mathbf{n}\equiv[\mathbf{n}_1^T,\ldots,\mathbf{n}_K^T]^T$, and $\mathbf{U}$ is a block-diagonal matrix with K blocks each performing convolution with the image u, i.e.

$$\mathbf{U}\equiv\underbrace{\begin{pmatrix}\mathbf{C}_{S_g}^{S_g,S_u}\{u\} & \cdots & \mathbf{0}\\ \vdots & \ddots & \vdots\\ \mathbf{0} & \cdots & \mathbf{C}_{S_g}^{S_g,S_u}\{u\}\end{pmatrix}}_{K\ \text{blocks}}.$$

We have obtained a standard MC convolution model and all conclusions for blur restoration in [18] and [19] apply also to our shifted version. When noise is omitted, it follows from (6) that the "valid" convolution matrix $\mathbf{Z}_k\equiv\mathbf{C}_{S_{\bar g}}^{S_{\bar g},S_z}\{z_k\}$ for some arbitrary support $S_{\bar g}$ is given by

$$\mathbf{Z}_k=\mathcal{U}\mathcal{G}_k\tag{7}$$

where $\mathcal{U} \equiv \mathbf{C}_{S_g+S_{\bar{g}}-1}^{S_g+S_{\bar{g}}-1,S_u}\{u\}$ and $\mathcal{G}_k \equiv \mathbf{C}_{S_{\bar{g}}}^{1,S_g+S_{\bar{g}}-1}\{g_k\}$. The above equality determines the rank property of $\mathbf{Z}_k$ and is utilized in the following lemma.

First, we recall an important definition from [19]. Let $\tilde{h}_k(z_1, z_2)$ denote the 2-D z transform of the blur h_k. The polynomials $\{\tilde{h}_k(z_1, z_2)\}$ are called "weakly coprime" if their only common factor is a scalar constant.

Lemma 1: Suppose that $K \geq 2, \{\tilde{h}_k(z_1, z_2)\}_{k=1}^K$ are weakly coprime, $\mathcal{U}$ in (7) has full column rank and the noise term is not present in (6). Then, all solutions $\{\bar{\mathbf{g}}_k\}$ to

$$\mathbf{Z}_i\bar{\mathbf{g}}_j - \mathbf{Z}_j\bar{\mathbf{g}}_i = \mathbf{0}, \quad 1 \leq i < j \leq K \tag{8}$$

have the form

$$\bar{\mathbf{g}}_k = \begin{cases} \mathbf{C}_{S_g}^{1,S_{\bar{g}}}\{f\}\mathbf{g}_k, & \text{if } S_{\bar{g}} \geq S_h + S_t \\ \alpha\mathbf{g}_k, & \text{if } S_{\bar{g}} = S_h + S_t \\ \emptyset, & \text{otherwise} \end{cases}$$

where f is some spurious factor of size $S_{\bar{g}} - S_h - S_t + 1$ and α is some scalar.

The proof is similar in nature to the proof given in [18] except that S_t is included in size constraints as discussed below. The above lemma states that in the noiseless case, if the estimated blur size $S_{\bar{g}}$ is equal to the sum of the maximum size of the original blurs S_h and the maximum shift S_t, then the true shifted blurs can be recovered precisely except to some scalar factor. This magnitude ambiguity can be resolved by stipulating, e.g., $\sum_x h_k(x) = 1$, which is a standard energy preserving assumption. For oversized $S_{\bar{g}}$, the solutions lie in a subspace of dimension $\prod(S_{\bar{g}} - S_g + 1)$. The first assumption that the blurs are weakly coprime is satisfied for many practical cases, since the necessary channel disparity is mostly guaranteed by the nature of the acquisition scheme and random processes therein. Refer to [18] for a relevant discussion. The second assumption of full column rank is also a mild one. For persistently exciting[1] u, the matrix $\mathcal{U}$ has full column rank provided that it has more rows than columns. Let us assume that the blur size is correctly estimated, i.e., $S_{\bar{g}} = S_g$, then $\mathcal{U}$ is of size $\prod(S_u - 2(S_g+1)) \times \prod(2S_g - 1)$ from which follows a size constraint $\prod(S_u - 2(S_h+S_t+1)) \geq \prod(2(S_h + S_t) - 1)$. Generally, u is much larger than h and the size constraint is violated only if the channel shift is for example $S_t > (S_u + 3)/4 - S_h$. To rephrase and simplify the condition, $\mathcal{U}$ does not have full column rank if $S_t > S_u/4$.

There are $K(K-1)/2$ equations in (8), and after stacking them into one system, we get

$$\mathcal{Z}\bar{\mathbf{g}} = \mathbf{0} \tag{9}$$

where $\bar{\mathbf{g}} \equiv [\bar{\mathbf{g}}_1^T, \ldots, \bar{\mathbf{g}}_K^T]^T$

$$\mathcal{Z} \equiv \begin{pmatrix} \mathcal{Z}_1^T & \cdots & \mathcal{Z}_{K-1}^T \end{pmatrix}^T$$

$$\mathcal{Z}_i \equiv \underbrace{\begin{pmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{Z}_{i+1} & -\mathbf{Z}_i & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{Z}_K & \mathbf{0} & \cdots & -\mathbf{Z}_i \end{pmatrix}}_{\substack{i-1 \text{ blocks} \\ \qquad\qquad K-i+1 \text{ blocks}}} \tag{10}$$

for $i = 1, \ldots, K-1$.

[1]Function u is called persistently exciting for size S_g if $u * g$ is different from zero for any $g \neq 0$, which is almost certainly true for real images if $S_g \ll S_u$.

The motivation behind Lemma 1 is to reduce the problem of identifying the blurs to a null-space problem, where the dimension of the null-space of $\mathcal{Z}$ is $\prod(S_{\bar{g}} - S_g + 1)$.

IV. MAP Blind Deconvolution

Adopting a stochastic approach, the restoration problem can be formulated as a MAP estimation. We assume that the images $\mathbf{u}, \mathbf{g}$, and $\mathbf{z}$ are random vector fields with given probability density functions (pdf) $p(\mathbf{u}), p(\mathbf{g})$, and $p(\mathbf{z})$, respectively, and we look for such realizations of $\mathbf{u}$ and $\mathbf{g}$, which maximize the *a posteriori* probability $p(\mathbf{u}, \mathbf{g} \,|\, \mathbf{z})$. We assume that $\mathbf{u}$ and $\mathbf{h}$ are uncorrelated then, according to the Bayes rule, the relation between *a priori* densities $p(\mathbf{u}), p(\mathbf{g})$ and the *a posteriori* density is $p(\mathbf{u}, \mathbf{g} \,|\, \mathbf{z}) = p(\mathbf{z} \,|\, \mathbf{u}, \mathbf{g})p(\mathbf{u})p(\mathbf{g})/p(\mathbf{z})$. The pdf $p(\mathbf{z})$ is a constant and can, thus, be omitted. The conditional pdf $p(\mathbf{z} \,|\, \mathbf{u}, \mathbf{g})$ follows from our model (6) and from our assumption of AWGN, i.e.

$$p(\mathbf{z} \,|\, \mathbf{u}, \mathbf{g}) \propto \exp\left\{-\frac{1}{2}(\mathbf{z} - \mathbf{G}\mathbf{u})^T\Sigma^{-1}(\mathbf{z} - \mathbf{G}\mathbf{u})\right\} \tag{11}$$

where Σ is the noise diagonal covariance matrix with $\{\sigma_k^2\}_{k=1}^K$ on the corresponding positions on the main diagonal. If the same noise variance σ^2 is assumed in each channel, Σ^{-1} reduces to a scalar σ^{-2}.

A. A Priori Distribution of the Original Image

The necessity of meaningful *a priori* probabilities becomes often Achilles' heel of Bayesian approaches. Several different forms of the image *a priori* probabilities were proposed in the literature. Some are suitable only for a specific class of images and others are more general. The classical form chooses the Laplacian operator as the inverse of the covariance matrix of u, i.e., $p(\mathbf{u}) \propto \exp(-\mathbf{u}^T\mathbf{L}\mathbf{u})$, where $\mathbf{L}$ denotes the discrete Laplacian operator. The exponent $\mathbf{u}^T\mathbf{L}\mathbf{u}$ is the discretization of $\int |\nabla u(x)|^2\, dx$, where ∇u denotes the gradient of u. Apart from easy implementation, this pdf is not suitable for the prior model, since the L_2 norm of the image gradient penalizes too much the gradients corresponding to edges and an oversmoothing effect is observed. In real images, object edges create sharp steps that appear as discontinuities in the intensity function. It is the space of bounded variation (BV) functions that is widely accepted as a proper setting for real images. Rudin *et al.* [3] first demonstrated very good anisotropic denoising properties of the total variation $\int |\nabla u(x)|\, dx$. Existence and uniqueness of the minimum of total variation is possible only in the BV space, in which case ∇u denotes the gradient of u in the distributional sense. The same holds true for a more general case of convex functions of measures $\int \phi(|\nabla u(x)|)\, dx$, where ϕ is a strictly convex, nondecreasing function that grows at most linearly. Examples of $\phi(s)$ are s (total variation), $\sqrt{1+s^2}-1$ (hypersurface minimal function), or $\log(\cosh(s))$. For nonconvex functions, nothing can be said about the existence of the minimum. Nevertheless, nonconvex functions, such as $\log(1+s^2)$, $s^2/(1+s^2)$ or $\arctan(s^2)$ (Mumford-Shah functional [35]), are often used since they provide better results for segmentation problems.

Our *a priori* image distribution consists of the convex function of measures. The function is highly nonlinear and to overcome this difficulty we follow a half-quadratic scheme proposed

in [36] and [37], which introduces an auxiliary variable. Special attention must be paid to the discretization of the image gradient $\nabla u(x)$ and relaxation of ϕ. If a second-order centered approximation of the first derivative is used, the prior pdf takes the form

$$p(\mathbf{u},v) \propto \exp\left\{\frac{1}{2}\sum_{i,j} v\left(i+\frac{1}{2},j\right)(u(i+1,j)-u(i,j))^2 + v\left(i,j+\frac{1}{2}\right)(u(i,j+1)-u(i,j))^2\right\} \equiv \exp\left\{-\frac{1}{2}\mathbf{u}^T\mathbf{L}(v)\mathbf{u}\right\} \quad (12)$$

where v is given by

$$v(i+\zeta,j+\zeta) = \frac{\phi'(|u(i+2\zeta,j+2\zeta)-u(i,j)|)}{|u(i+2\zeta,j+2\zeta)-u(i,j)|} \qquad \zeta = \left\{0,\frac{1}{2}\right\}. \quad (13)$$

The auxiliary flux variable v is similar to Geman's line process [38]. It denotes the edge strength between point (i,j) and its neighbors $(i\pm 1,j),(i,j\pm 1)$. For example, in the case of the hypersurface minimal function $\phi(s)=\sqrt{1+s^2}-1$, the flux variable becomes $v(s)=\phi'(s)/s=1/\sqrt{1+s^2}$. Matrix $\mathbf{L}(v)$ is a positive semidefinite block tridiagonal matrix constructed by v that performs shift-variant convolution with v. In the above discretization, the norm of the image gradient is variant to rotation. A more precise discretization is possible if we take into account the diagonal values.

B. A Priori Distribution of the Blurs

We derive the *a priori* distribution $p(\mathbf{g})$ directly from the MC model. If the AWGN noise term $\mathbf{n}$ is present in model (6), then the left-hand side of (9) is not zero, but equal to a realization of a Gaussian process with zero mean and covariance $\mathcal{C}=\mathcal{G}\Sigma\mathcal{G}^T$. Matrix $\mathcal{G}$ takes the form of $\mathcal{Z}$ in (10) with $\mathbf{Z}_i$ replaced by $\mathbf{C}_{S_z}^{S_g,S_z}\{g_i\}$. Our first estimate of the *a priori* pdf $p'(\mathbf{g})$ is then given by

$$p'(\mathbf{g}) \propto \exp\left\{-\frac{1}{2}\mathbf{g}^T\mathcal{Z}^T\mathcal{C}^{-1}\mathcal{Z}\mathbf{g}\right\}.$$

From Lemma 1, it follows that $\mathcal{Z}^T\mathcal{C}^{-1}\mathcal{Z}$ is close to singular and the number of eigenvalues that cluster around the noise variance is proportional to the degree of overestimation $(\prod(S_{\overline{g}}-S_g+1))$. The expected blurs lie inside a subspace defined by eigenvectors that correspond to these eigenvalues. We propose to construct *a priori* $p(\mathbf{g})$ by constraining $p'(\mathbf{g})$ to a set of admissible solutions. The set of admissible solutions is defined by our assumption that the blurs are positive and preserve energy; $A\equiv\{\mathbf{g}\,|\,g_k(x)\geq 0\wedge\sum_x g_k(x)=1,k=1,\ldots,K\}$. We write

$$p(\mathbf{g}) = \begin{cases} p'(\mathbf{g}), & \text{if } \mathbf{g}\in A \\ 0, & \text{otherwise.}\end{cases} \quad (14)$$

This leads to a prior pdf that is data dependent. From a strictly theoretical point of view, one should use here a different set of input data but degraded by the same blurs as the data $\mathbf{z}$ in question, or use one part of the input data for $p(\mathbf{g})$ and perform restoration on the other part.

The main difficulties are connected with the matrix $\mathcal{C}$. The inverse of the matrix is not trivial and the matrix is constructed by the blurs $\mathbf{g}$ that are to be estimated. One way is to use an iterative algorithm and update $\mathcal{C}$ by $\mathbf{g}$ estimated in the previous iteration. This iterative maximization of $p'(\mathbf{g})$ w.r.t. $\mathbf{g}$ closely resembles the maximum-likelihood algorithm proposed by Harikumar [18]. However, this updating procedure is difficult to justify. We, therefore, propose to simplify $\mathcal{C}$ and approximated it by a diagonal matrix $\mathcal{D}$ such that $\mathrm{diag}(\mathcal{D})=\mathrm{diag}(\mathcal{C})$, where $\mathrm{diag}(\cdot)$ denotes the main diagonal of the matrix. The elements of $\mathcal{D}$ take the form $\sigma_i^2\|\mathbf{g}_j\|^2+\sigma_j^2\|\mathbf{g}_i\|^2$ for $1\leq i<j\leq K$. The value of $\|\mathbf{g}_i\|^2$ is not known in advance, but a good initial approximation can be given. If $\mathbf{g}\in A$, then $(1/\prod S_g)\leq\|\mathbf{g}_i\|^2\leq 1$, and we use the bottom limit for $\|\mathbf{g}_i\|^2$.

C. AM–MAP Algorithm

The *a posteriori* pdf $p(\mathbf{u},\mathbf{g}\,|\,\mathbf{z})$ is composed of (11), (12), and (14) and turns out to be

$$p(\mathbf{u},\mathbf{g}\,|\,\mathbf{z}) \propto \exp\left\{-\frac{1}{2}((\mathbf{z}-\mathbf{G}\mathbf{u})^T\Sigma^{-1}(\mathbf{z}-\mathbf{G}\mathbf{u}) + \mathbf{u}^T\mathbf{L}(v)\mathbf{u}+\mathbf{g}^T\mathcal{Z}^T\mathcal{D}^{-1}\mathcal{Z}\mathbf{g})\right\} \quad (15)$$

for $\mathbf{g}\in A$ and zero otherwise. The MAP estimation is then equivalent to minimizing $E(\mathbf{u},\mathbf{g})=-\log(p(\mathbf{u},\mathbf{g}\,|\,\mathbf{z}))$ subjected to $\mathbf{g}\in A$. To find a minimizer of the energy function E, we perform alternating minimizations of E over $\mathbf{u}$ and $\mathbf{g}$. The advantage of this scheme lies in its simplicity. Each term in (15) is convex (but not necessarily strictly convex, especially, when $\mathbf{g}$ is oversized) and the derivatives w.r.t. $\mathbf{u}$ and $\mathbf{g}$ can be easily calculated.

In summary, the AM–MAP algorithm alternates between two steps

1)
$$\mathbf{u}^{m+1} = \arg\min_{\mathbf{u}} E(\mathbf{u},\mathbf{g}^m) \Leftrightarrow (\mathbf{G}^T\Sigma^{-1}\mathbf{G}+\mathbf{L}(v))\mathbf{u}=\mathbf{G}^T\Sigma^{-1}\mathbf{z} \quad (16)$$

2)
$$\mathbf{g}^{m+1} = \arg\min_{\mathbf{g}\in A} E(\mathbf{u}^{m+1},\mathbf{g}) \Leftrightarrow (\mathbf{U}^T\Sigma^{-1}\mathbf{U}+\mathcal{Z}^T\mathcal{D}^{-1}\mathcal{Z})\mathbf{g} = \mathbf{U}^T\Sigma^{-1}\mathbf{z}\wedge\mathbf{g}\in A. \quad (17)$$

In step 1, the flux variable v is updated according to (13).

Our AM approach is a variation on the steepest-descent algorithm. The search space is a concatenation of the blur subspace and the image subspace. The algorithm first descends in the image subspace and after reaching the minimum, i.e., $\nabla_{\mathbf{u}}E=0$, it advances in the blur subspace in the direction $\nabla_{\mathbf{g}}E$ orthogonal to the previous one, and this scheme repeats. We use the preconditioned conjugate gradient method (function *pcg* in Matlab) to solve the unconstrained minimization problem (16) and *fmincon* (Optimization Toolbox) function

to solve the constrained minimization problem (17). E as a function of both variables $\mathbf{u}$ and $\mathbf{g}$ is not convex. We cannot, thus, guarantee that the global minimum is reached by the AM–MAP algorithm. Nevertheless, our thorough testing have shown good convergence properties of the algorithm for many real problems. In the following experiments, the hypersurface minimum function was used as $\phi(s)$, and the more precise discretization involving diagonal terms was implemented. We also assumed that the noise variance σ_k^2 is known and is the same in each channel. If this is not the case, the noise variance can be assessed by standard noise estimation methods or an approach of "trial and error" can be considered. The impact of wrong σ_k^2 can be easily observed. If the parameter is too small, i.e., we assume less noise, the restoration process begins to amplify noise in the image. If the parameter is too big, the restoration process starts to segment the image.

D. Initial Guess

Setting appropriately initial blurs can help our iterative algorithm to converge to the global minimum. This issue is important especially for the overestimated blur size. One can readily see that translated versions of the correct blurs give the same maximum probability $p(\mathbf{u}, \mathbf{g} \,|\, \mathbf{z})$ as long as they fit into our estimated blur size. We already know that the prior pdf of the blurs is unable to distinguish between the correct blurs and the correct blurs convolved with an arbitrary spurious factor. This makes a negative impact on the convergence mainly if the channel misalignment occurs, since new local minima appear for blurs that cope with the misalignment by convolving the correct blurs with an interpolating kernel. For example, to compensate a one-pixel shift between two channels, the suboptimal solution is to shift both blurs in the opposite direction by half a pixel and perform an interpolation, while the correct solution (global minimum) is to shift one of the estimated blurs by one pixel and leave the other. To get closer to the correct solution, we, thus, propose to set the initial blurs $\mathbf{g}^0$ to delta functions positioned at the centers of gravity of blurs $\tilde{\mathbf{g}} = \arg\max_{\mathbf{g}} p(\mathbf{g})$. More precisely, if the images were blurred with energy preserving PSFs $\mathbf{g} \in A$, the centers of gravity (cog) satisfy $\mathrm{cog}(g_i) - \mathrm{cog}(g_j) = \mathrm{cog}(g_i * f) - \mathrm{cog}(g_j * f)$ for any f that preserves energy. From Lemma 1, it follows that our estimate $\tilde{g}_i$ is a good approximation of $g_i * f$, and we can, thus, calculate the relative positions of the centers of gravity. This technique enables us to compensate for the channel shifts right from the start of the algorithm and get away from the incorrect interpolated solutions.

V. Oversized Blurs

It is difficult to analyze global convergence properties of the algorithm (16), (17) due to the nonlinear term $\mathbf{L}(v)$. Chan *et al.* in [9] transformed a similar alternating minimization problem into the Fourier domain and performed the analysis there. In the Fourier domain, it is difficult to apply the support constraint on the blurs, but in our case, the blur size plays a fundamental role in Lemma 1. The analysis should, thus, be performed in the z transform domain but this is difficult to carry out.

We have run a series of experiments on simulated data with an incorrectly estimated blur size. A standard 128×128 "Lena" image was degraded with three random blurs of size 3×3 and with additive Gaussian noise of SNR $= 20, 30, 40, 50$ dB (SNR $= 10 \log(\mathrm{var}(u)/\sigma^2)$, where $\mathrm{var}(u)$ is the variance of the original image). The original image was recovered from each image triplet using the alternating minimization algorithm with the blur size set to 3×3, 4×4, 6×6, and 8×8, respectively. The percentage mean squared error of the estimated image $\bar{\mathbf{u}}$ defined as $\mathrm{PMSE}(\bar{\mathbf{u}}) = 100\|\bar{\mathbf{u}} - \mathbf{u}\|/\|\mathbf{u}\|$ was used as the evaluation measure at each iteration. Calculated PMSEs are summarized in Fig. 2. For less noisy data (SNR $= 40, 50$ dB), the convergence rate is fast and not affected by the incorrect blur size estimation. In the case of SNR $= 20, 30$ dB, we observe a negative influence of the overestimated blur size on the convergence. However, not many practical applications provide data with such low SNRs.

VI. Shift-Invariant Restoration

In this section, we illustrate the ability of the method to handle channels which are not registered. This is the most important advantage of the new technique.

The first experiment demonstrates the capability of the AM–MAP algorithm to recover the original image from two degraded and shifted versions thereof, when the maximum shift between the two channels is known. The standard 128×128 "Lena" image was degraded with two 5×5 blurs. One blurred image was shifted by 10×20 pixels and then both images were cropped to the same size; see Fig. 3. The AM algorithm was initialized with the correctly estimated blur size 15×25. The restored image and blurs are shown in Fig. 4. The blurs are perfectly recovered and properly shifted. The restored image matches the original on the area where data from both channels were available. The same experiment was conducted again but Gaussian noise SNR $= 30$ dB was added to the blurred and shifted input images in Fig. 3. Obtained results are depicted in Fig. 5 and illustrate satisfying restoration.

In the second experiment, we overestimated the blur size and proceeded as follows. The 250×250 test image in Fig. 6(a) was degraded with two different 5×5 blurs and noise of SNR $= 50$ dB. One blurred image was shifted by 5×5 pixels and then both images were cropped to the same size; see Fig. 6(c). Contrary to the previous experiment, the shift was considered unknown and the AM–MAP algorithm was initialized with the overestimated blur size 12×12. The fused image and the estimated blur masks are shown in Fig. 7. Recovered blurs contain negligible spurious factors and are properly shifted to compensate for the misregistration. The fused image is by visual comparison much sharper than the input channels and very similar to the original, which demonstrates excellent performance.

We have also compared the performance of the AM–MAP algorithm with the performance of Pai's method [22] for different noise levels. The Pai approach directly recovers the original image by calculating the maximum singular vector of a special matrix. The *QR decomposition* is necessary for the construction of this matrix and the *power method* (or any other iterative method for eigenvector computation) is used to find the maximum singular vector, i.e., the original image.

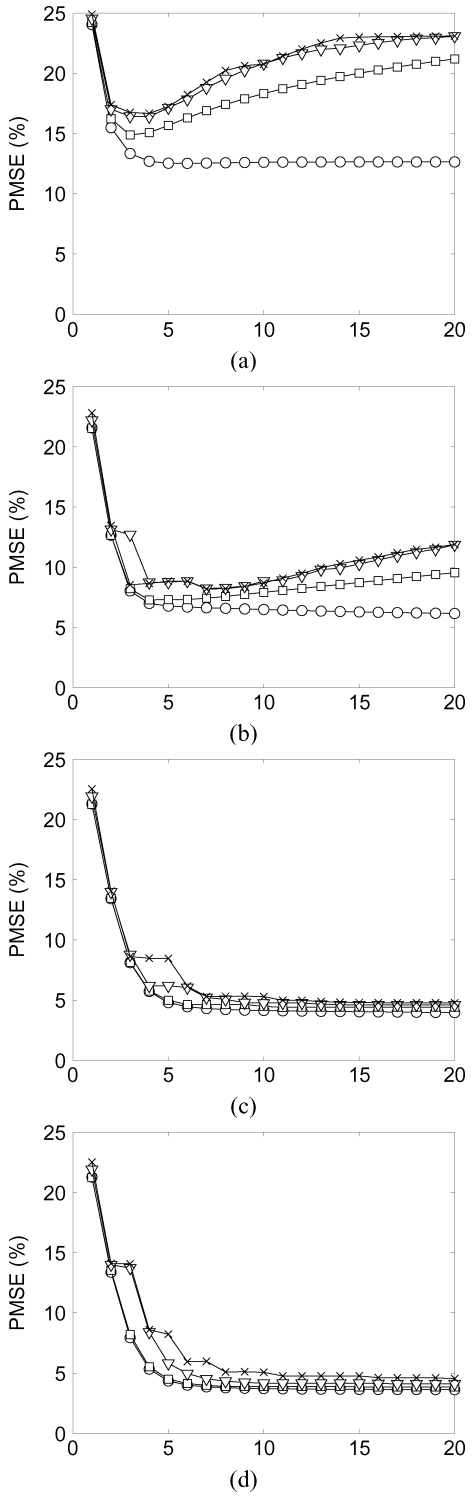


Fig. 2. PMSE of the estimated image as a function of iteration. The "Lena" image was degraded with three random blurs of size 3×3 and with AWGN of SNR (a) 20, (b) 30, (c) 40, and (d) 50 dB. The AM algorithm was executed with the estimated blur size ($\bigcirc$) 3×3 (correct size), ($\square$) 4×4, ($\triangledown$) 6×6, and ($\times$) 8×8.

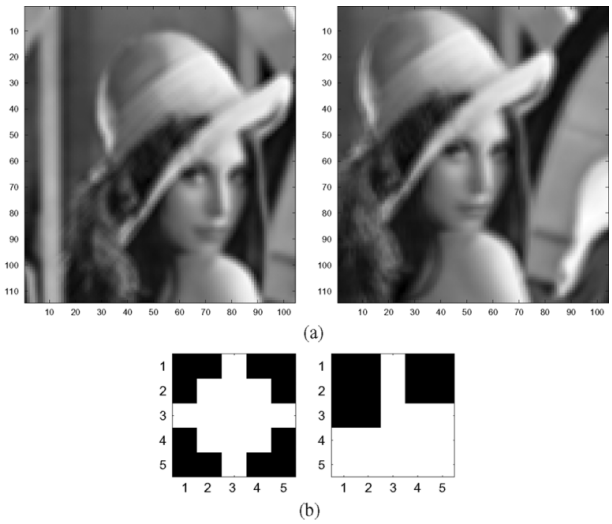


Fig. 3. (a) Input channels. "Lena" images degraded with two 5×5 blurs. Mutual translation between the images is 10 pixels vertically and 20 pixels horizontally. (b) Blurring masks.

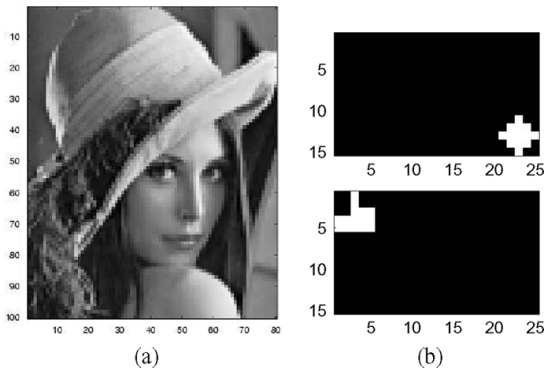


Fig. 4. Perfect noise-free AM restoration. (a) Recovered "Lena" image. (b) Recovered blurs and 10×20 shift between channels.

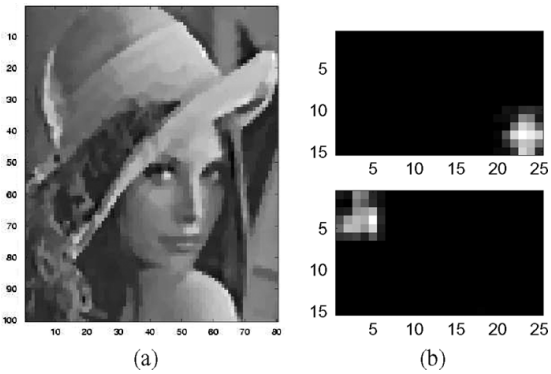


Fig. 5. Noisy AM restoration (30 dB). (a) Recovered "Lena" image. (b) Recovered blurs and 10×20 shift between channels.

Although the Pai method is not iterative in its definition, it requires numerical iterative methods and, thus, approaches the complexity of our inherently iterative algorithm. We used four randomly generated 3×3 blurs to obtain four blurred "Lena"

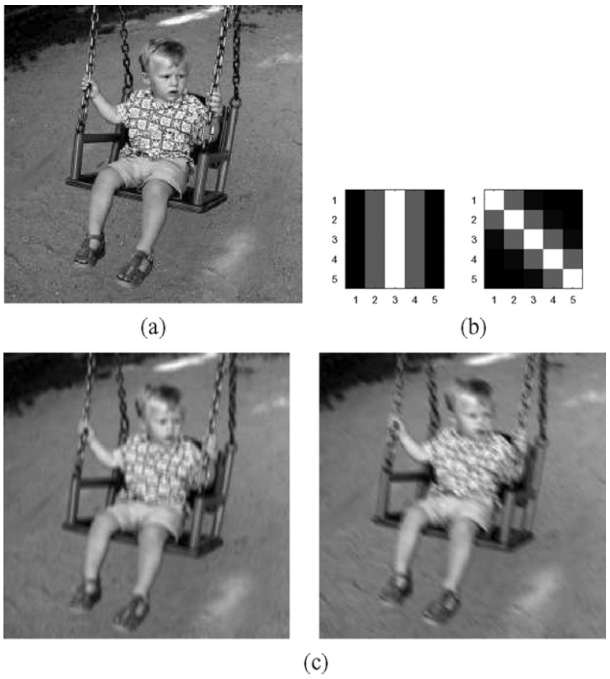


Fig. 6. (a) Original test image 250×250 pixels. (b) Two 5×5 PSFs. (c) Blurred and shifted images.

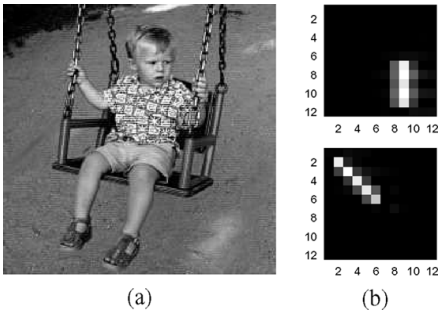


Fig. 7. Image restoration for the overestimated blur size. (a) Recovered image. (b) Estimated blur masks with the between-channel shift.

images. The images were then mutually translated so that centers of the images were in corners of a 5×5 square. Noise was added with SNR $= 10, 20, 30, 40$, and 50 dB, respectively. The maximum shift and the size of blurs were assumed to be known and, therefore, both methods were initialized with the correct blur size 8×8. For each SNR, the experiment was repeated with different blurs 10 times and stopped after 50 iterations in the AM–MAP case. The mean PMSE and standard deviation was calculated over these ten estimated images and plotted in Fig. 8. Clearly, the AM–MAP performs better then the Pai method for every SNR.

To evaluate the performance of the AM–MAP algorithm with respect to the knowledge of the channel misalignment, a different experiment was conducted. Degraded images were prepared in similar fashion as in the previous experiment but this time the translation between the channels varied from 0 to 5×5

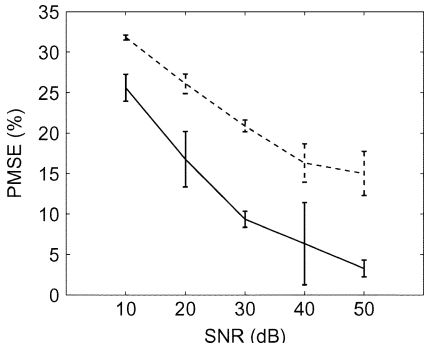


Fig. 8. Comparison of the (solid) AM–MAP algorithm and the (dashed) Pai method. Mean PMSE and (vertical abscissae) standard deviation of the restored images over ten different degradations and for different SNR.

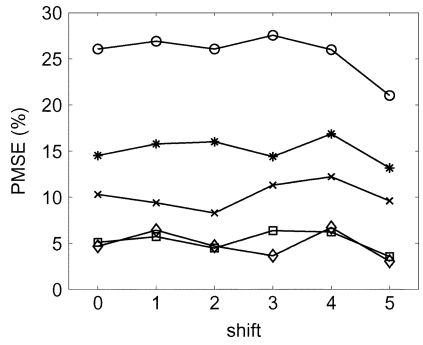


Fig. 9. AM–MAP algorithm performance on misaligned channels. Mean PMSE of restored images over ten different degradations for the channel misalignment 0 to 5 pixels and SNR ($\circ$) 10, ($*$) 20, ($\times$) 30, ($\square$) 40, and ($\diamond$) 50 dB.

pixels to simulate inaccurate registration. For each shift, the experiment was repeated ten times with different blurs and was every time initialized to 8×8 blur size. The calculated mean PMSE is plotted in Fig. 9. PMSEs are almost constant, which demonstrates very good stability of the algorithm against the mask overestimation.

Finally, to demonstrate the power of the AM–MAP algorithm, we performed an experiment with real data. This experiment was motivated by many practical situations where we have to handle images degraded by random motion and/or vibration blur. This problem appears frequently in industrial visual inspection when the camera is mounted on a vibrating machine or when a stationary camera monitors vibrating environment. A text label (a part of a standard newspaper page) was attached to a vibrating machine. The label was monitored under poor light conditions by a standard digital camera mounted on a tripod. The camera exposure time was set at $1/15$s which was comparable to the period of irregular vibrations of the machine. Three cropped images of the label acquired with the camera were used as the input channels of AM–MAP; see Fig. 10. The images were not binarized but only trasformed to grey-level images. Note strong motion blurs due to the machine movement and clear spatial misalignment of the channels. Since the shift and the size of the blurs were completely unknown and it was also difficult to estimate the upper bound, the AM–MAP algorithm was restarted with different parameters and the best results were

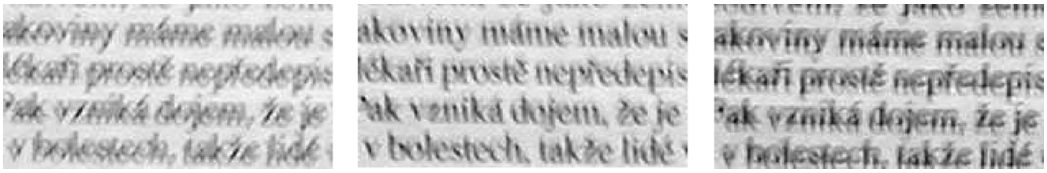


Fig. 10. Real data experiment. Three consecutive acquisitions of a text label attached to a vibrating machine. The grey-level images are cropped to 100×200 size. Shift blurs and spatial misalignment of the images are clearly visible.

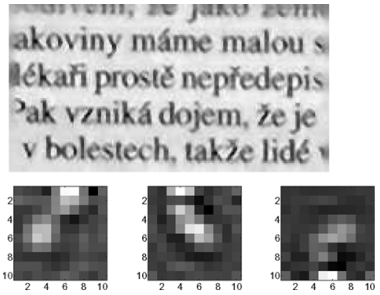


Fig. 11. Real data experiment. Reconstructed part of the label and the corresponding blurs (magnified) using the AM–MAP algorithm. The irregular vibration of the machine is well preserved in the blurs.

found for the blur size 10×10 and $\sigma^2 = 0.01$. We used the total variation in the *a priori* pdf to favor piecewise constant functions which is the case of the text label. The reconstructed label and the corresponding blur masks after 20 iterations are shown in Fig. 11. One can see that the restoration was successful (the text is clearly legible) and that the spatial misalignment inherent to this type of problems poses no threat to proper functionality of the algorithm. Observe that the restoration is slightly less successful near the image borders, especially close to the top edge, where only data from the third channel were available. Let us recall that no assumption about the shape of the blurring functions and no preprocessing of the input images were employed.

VII. Conclusion

We have developed the iterative algorithm for multichannels blind deconvolution that searches for the MAP estimator. The prior density functions were derived from the variational integral defined on bounded variation functions and from the mutual relation of weakly coprime channels. The restoration is regularized with an anisotropic term for edge preservation and performs well on heavily degraded images with high SNR and shows better performance then the most recent multichannel method. We have also shown that the inaccurate registration of channels can be alleviated by properly overestimating the size of blurs. All previously published MBD methods assumed perfectly registered channels or required, though not specifically considered in their formulation, an exact knowledge of the channel misalignment. To our knowledge, this is the only method dealing explicitly with misregistration of images in the multichannel framework and providing a successful solution to this problem.

References

[1] M. Banham and A. Katsaggelos, "Digital image restoration," *IEEE Signal Process. Mag.*, vol. 14, no. 2, pp. 24–41, Mar. 1997.

[2] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.

[3] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.

[4] C. Vogel and M. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Image Process.*, vol. 7, no. 6, pp. 813–824, Jun. 1998.

[5] T. Chan, G. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," in *SIAM J. Sci. Comput.*, vol. 20, Jul. 1999, pp. 1964–1977.

[6] A. B. Hamza, H. Krim, and G. Unal, "Unifying probabilistic and variational estimation," *IEEE Signal Process. Mag.*, vol. 19, no. 5, pp. 37–47, Sep. 2002.

[7] T. Chan and C. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.

[8] Y.-L. You and M. Kaveh, "Blind image restoration by anisotropic regularization," *IEEE Trans. Image Process.*, vol. 8, no. 3, pp. 396–407, Mar. 1999.

[9] T. Chan and C. Wong, "Convergence of the alternating minimization algorithm for blind deconvolution," *Linear Algebra Appl.*, vol. 316, no. 1–3, pp. 259–285, Sep. 2000.

[10] R. Lagendijk, J. Biemond, and D. Boekee, "Identification and restoration of noisy blurred images using the expectation-maximization algorithm," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 7, pp. 1180–1191, Jul. 1990.

[11] S. Reeves and R. Mersereau, "Blur identification by the method of generalized cross-validation," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 301–311, Mar. 1992.

[12] M. Haindl, "Recursive model-based image restoration," in *Proc. 15th Int. Conf. Pattern Recognition*, vol. III, 2000, pp. 346–349.

[13] W. Zhu, N. Galatsanos, and A. Katsaggelos, "Regularized multichannel restoration using cross-validation," *Graph. Models Image Process.*, vol. 57, no. 1, pp. 38–54, Jan. 1995.

[14] B. Tom, K. Lay, and A. Katsaggelos, "Multichannel image identification and restoration using the expectation-maximization algorithm," *Opt. Eng.*, vol. 35, no. 1, pp. 241–254, Jan. 1996.

[15] A. Katsaggelos, K. Lay, and N. Galatsanos, "A general framework for frequency domain multi-channel signal processing," *IEEE Trans. Image Process.*, vol. 2, no. 7, pp. 417–420, Jul. 1993.

[16] M. Kang, "Generalized multichannel image deconvolution approach and its applications," *Opt. Eng.*, vol. 37, no. 11, pp. 2953–2964, Nov. 1998.

[17] T. Schulz, "Multiframe blind deconvolution of astronomical images," *J. Opt. Soc. Amer. A*, vol. 10, no. 5, pp. 1064–1073, May 1993.

[18] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 202–219, Feb. 1999.

[19] G. Giannakis and R. Heath, "Blind identification of multichannel FIR blurs and perfect image restoration," *IEEE Trans. Image Process.*, vol. 9, no. 11, pp. 1877–1896, Nov. 2000.

[20] G. Harikumar and Y. Bresler, "Exact image deconvolution from multiple FIR blurs," *IEEE Trans. Image Processing*, vol. 8, no. 6, pp. 846–862, Jun. 1999.

[21] H.-T. Pai and A. Bovik, "Exact multichannel blind image restoration," *IEEE Signal Process. Lett.*, vol. 4, no. 8, pp. 217–220, Aug. 1997.

[22] ——, "On eigenstructure-based direct multichannel blind image restoration," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1434–1446, Oct. 2001.

[23] S. Pillai and B. Liang, "Blind image deconvolution using a robust GCD approach," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 295–301, Feb. 1999.

[24] A. Rajagopalan and S. Chaudhuri, "A recursive algorithm for maximum likelihood-based identification of blur from multiple observations," *IEEE Trans. Image Process.*, vol. 7, no. 7, pp. 1075–1079, Jul. 1998.
[25] M. Haindl and S. Šimberová, "Model-based restoration of short-exposure solar images," in *Frontiers in Artificial Intelligence and Applications*. ser. Knowledge-Based Intelligent Engeneering Systems, L. Jain and R. Howlett, Eds. Amsterdam, The Netherlands: IOS, 2002, vol. 87, pp. 697–706.
[26] G. Panci, P. Campisi, S. Colonnese, and G. Scarano, "Multichannel blind image deconvolution using the bussgang algorithm: Spatial and multiresolution approaches," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1324–1337, Nov. 2003.
[27] F. Šroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1094–1106, Sep. 2003.
[28] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, pp. 977–1000, 2003.
[29] Z. Myles and N. V. Lobo, "Recovering affine motion and defocus blur simultaneously," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 6, pp. 652–658, Jun. 1998.
[30] Y. Zhang, C. Wen, Y. Zhang, and Y. C. Soh, "Determination of blur and affine combined invariants by normalization," *Pattern Recognit.*, vol. 35, pp. 211–221, 2002.
[31] A. Kubota, K. Kodama, and K. Aizawa, "Registration and blur estimation methods for multiple differently focused images," in *Proc. Int. Conf. Image Processing*, vol. II, 1999, pp. 447–451.
[32] Z. Zhang and R. Blum, "A hybrid image registration technique for a digital camera image fusion application," *Inf. Fusion*, vol. 2, pp. 135–149, 2001.
[33] J. Flusser and T. Suk, "Degraded image analysis: An invariant approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 6, pp. 590–603, Jun. 1998.
[34] J. Flusser, J. Boldyš, and B. Zitová, "Moment forms invariant to rotation and blur in arbitrary number of dimensions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 234–246, Feb. 2003.
[35] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.
[36] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*. New York: Springer Verlag, 2002.
[37] A. Chambolle and P. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.*, vol. 76, no. 2, pp. 167–188, Apr. 1997.
[38] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.

Filip Šroubek received the B.Sc. and M.Sc. degrees in computer science from the Czech Technical University, Prague, Czech Republic, in 1996 and 1998, respectively, and the Ph.D. degree in computer science from the Charles University, Prague, in 2003.

Since 1999, he has been with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague. Since 2000, he has been also with the Institute of Radiotechnique and Electronics, Academy of Sciences of the Czech Republic, Prague. His current research interests include all aspects of digital image processing and pattern recognition, particularly multichannel blind deconvolution, image denoising, image registration, and computer simulation and visualization of atomic collision processes.

Jan Flusser (M'93–SM'03) received the M.Sc. degree in mathematical engineering from the Czech Technical University, Prague, Czech Republic, in 1985, and the Ph.D. degree in computer science and the D.Sc. degree in technical cybernetics from the Czechoslovak Academy of Sciences, Czech Republic, in 1990 and 2001, respectively.

Since 1985, he has been with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague, where, since 1995, he has held the position of Head of the Department of Image Processing. Since 1991, he has been also affiliated with the Charles University, Prague, and the Czech Technical University, Prague, where he teaches courses on digital image processing and pattern recognition. His current research interests include all aspects of digital image processing and pattern recognition, namely 2-D object recognition, moment invariants, blind deconvolution, image registration, and image fusion, and he has authored or coauthored more than 80 research publications in these areas.

2322 IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 16, NO. 9, SEPTEMBER 2007

A Unified Approach to Superresolution and Multichannel Blind Deconvolution

Filip Šroubek, Gabriel Cristóbal, *Senior Member, IEEE*, and Jan Flusser, *Senior Member, IEEE*

***Abstract*—This paper presents a new approach to the blind deconvolution and superresolution problem of multiple degraded low-resolution frames of the original scene. We do not assume any prior information about the shape of degradation blurs. The proposed approach consists of building a regularized energy function and minimizing it with respect to the original image and blurs, where regularization is carried out in both the image and blur domains. The image regularization based on variational principles maintains stable performance under severe noise corruption. The blur regularization guarantees consistency of the solution by exploiting differences among the acquired low-resolution images. Several experiments on synthetic and real data illustrate the robustness and utilization of the proposed technique in real applications.**

***Index Terms*—Image restoration, multichannel blind deconvolution, regularized energy minimization, resolution enhancement, superresolution.**

I. Introduction

IMAGING devices have limited achievable resolution due to many theoretical and practical restrictions. An original scene with a continuous intensity function $o[x, y]$ warps at the camera lens because of the scene motion and/or change of the camera position. In addition, several external effects blur images: atmospheric turbulence, camera lens, relative camera-scene motion, etc. We will call these effects *volatile blurs* to emphasize their unpredictable and transitory behavior, yet we will assume that we can model them as convolution with an unknown point spread function (PSF) $v[x, y]$. This is a reasonable assumption if the original scene is flat and perpendicular to the optical axis. Finally, the CCD discretizes the images and produces digitized noisy image $g[i, j]$ (frame). We refer to $g[i, j]$ as a *low-resolution (LR) image*, since the spatial resolution is too low to capture all the details of the original scene. In conclusion, the acquisition model becomes

$$g[i,j] = D((v * W(o))[x,y]) + n[i,j] \tag{1}$$

Manuscript received October 30, 2006; revised May 23, 2007. This work was supported in part by the Czech Ministry of Education under the project 1M0572 (Research Center DAR), in part by the Academy of Sciences of the Czech Republic under the project AV0Z10750506-I055, in part by the bilateral project 2004CZ0009, and in part by the projects TEC2004-00834, TEC2005-24739-E, and PI040765. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Tamas Sziranyi.

F. Šroubek and J. Flusser are with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 18208 Prague 8, Czech Republic (e-mail: sroubekf@utia.cas.cz; flusser@utia.cas.cz).

G. Cristóbal is with the Instituto de Óptica, CSIC, Serrano 121, 28006 Madrid, Spain (e-mail: gabriel@optica.csic.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2007.903256

where $n[i, j]$ is additive noise and W denotes the geometric deformation (warping). $D(\cdot) = S(g*\cdot)$ is the *decimation operator* that models the function of the CCD sensors. It consists of convolution with the *sensor PSF* $g[i, j]$ followed by the *sampling operator* S, which we define as multiplication by a sum of delta functions placed on a evenly spaced grid. The above model for one single observation $g[i, j]$ is extremely ill-posed. Instead of taking a single image we can take K $(K > 1)$ images of the original scene and this way partially overcome the equivocation of the problem. Hence, we write

$$g_k[i,j] = D((v_k * W_k(o))[x,y]) + n_k[i,j] \tag{2}$$

where $k = 1, \ldots, K$, and D remains the same in all the acquisitions. In the perspective of this multiframe model, the original scene $o[x, y]$ is a single input and the acquired LR images $g_k[i, j]$ are multiple outputs. The model is, therefore, called a single-input–multiple-output (SIMO) formation model. The upper part of Fig. 1 summarizes the multiframe LR acquisition process. To our knowledge, this is the most accurate, state-of-the-art model, as it takes all possible degradations into account. Several other authors, such as in [1]–[4], adopt this model, as well.

Superresolution (SR) is the process of combining a sequence of LR images in order to produce a higher resolution image or sequence. It is unrealistic to assume that the superresolved image can recover the original scene $o[x, y]$ exactly. A reasonable goal of SR is a discrete version of $o[x, y]$ that has a higher spatial resolution than the resolution of the LR images and that is free of the volatile blurs (deconvolved). In the paper, we will refer to this superresolved image as a *high resolution (HR) image* $f[i, j]$. The standard SR approach consists of subpixel registration, overlaying the LR images on an HR grid, and interpolating the missing values. The subpixel shift between images thus constitutes the essential assumption. We will demonstrate that assuming volatile blurs in the model explicitly brings about a more general and robust technique, with the subpixel shift being a special case thereof.

The acquisition model in (2) embraces three distinct cases frequently encountered in literature. First, if we want to resolve the geometric degradation W_k, we face a registration problem. Second, if the decimation operator D and the geometric transform W_k are not considered, we face a *multichannel* (or multiframe) *blind deconvolution* (MBD) problem. Third, if the volatile blur v_k is not considered or assumed known, and W_k is suppressed up to a subpixel translation, we obtain a classical SR formulation. In practice, it is crucial to consider all three cases at once. We are then confronted with a problem of *blind superresolution* (BSR), which is the subject of this investiga-

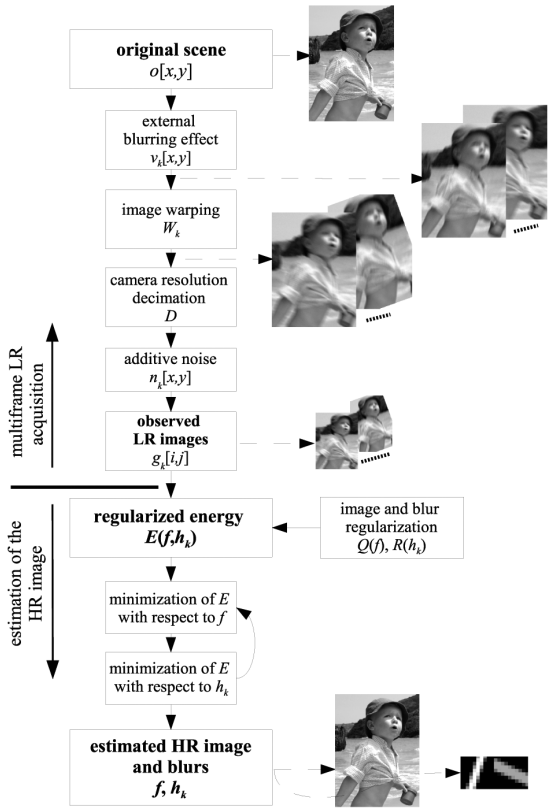


Fig. 1. (Top) Low-resolution acquisition and (bottom) reconstruction flow.

tion. The approach presented in this manuscript is one of the first attempts to solve BSR with only little prior knowledge.

Proper registration techniques can suppress large and complex geometric distortions (usually just up to a small between-image shift). There have been hundreds of methods proposed; see, e.g., [5] for a survey. In the rest of this paper, we will assume that the LR images are roughly registered and that W_ks reduce to small translations.

The MBD problem has recently attracted considerable attention. First blind deconvolution attempts were based on single-channel formulations, such as in [6]–[9]. Kundur *et al.* [10], [11] provide a good overview. The problem is extremely ill-posed in the single-channel framework and cannot be resolved in a fully blind form. These methods do not exploit the potential of the multichannel framework, because in the single-channel case missing information about the original image in one channel is not supplemented by information in the other channels. Research on intrinsically multichannel methods has begun fairly recently; refer to [12]–[16] for a survey and other references. Such MBD methods overpass the limitations of previous techniques and can recover the blurring functions from the degraded images alone. We further developed the MBD theory in [17] by proposing a blind deconvolution method for images, which might be mutually shifted by unknown vectors. To make this brief survey complete, we should not forget to mention a very challenging problem of shift-variant blind deconvolution, that was considered in [18] and [19].

A countless number of papers address the standard SR problem. A good survey can be found for example in [20] and [21]. Maximum likelihood, maximum *a posteriori* (MAP), the set theoretic approach using projection on convex sets, and fast Fourier techniques can all provide a solution to the SR problem. Earlier approaches assumed that subpixel shifts are estimated by other means. More advanced techniques, such as in [1], [2], and [4], include the shift estimation into the SR process. Other approaches focus on fast implementation [3], space-time SR [22] or SR of compressed video [2]. Most of the SR techniques assume *a priori* known blurs. However, in many cases, such as blurring due to camera motion, the blur can have a wild shape that is difficult to predict; see examples of real motion blurs in [23]. Authors in [24]–[26] proposed BSR that can handle parametric PSFs, i.e., PSFs modeled with one parameter. This restriction is unfortunately very limiting for most real applications. In [27], we extended our MBD method to BSR in an intuitive way but one can prove that this approach does not estimate PSFs accurately. The same intuitive approach was also proposed in [28]. To our knowledge, first attempts for theoretically correct BSR with an arbitrary PSF appeared in [29] and [30]. The interesting idea proposed therein is the conversion of the SR problem from SIMO to multiple input multiple output using so-called polyphase components. We will adopt the same idea here as well. Other preliminary results of the BSR problem with focus on fast calculation are given in [31], where the authors propose a modification of the Richardson–Lucy algorithm.

Current multiframe blind deconvolution techniques require no or very little prior information about the blurs, they are sufficiently robust to noise and provide satisfying results in most real applications. However, they can hardly cope with the decimation operator, which violates the standard convolution model. On the contrary, state-of-the-art SR techniques achieve remarkable results of resolution enhancement in the case of no blur. They accurately estimate the subpixel shift between images but lack any apparatus for calculating the blurs.

We propose a unifying method that simultaneously estimates the volatile blurs and HR image without any prior knowledge of the blurs and the original image. We accomplish this by formulating the problem as a minimization of a regularized energy function, where the regularization is carried out in both the image and blur domains. The image regularization is based on variational integrals, and a consequent anisotropic diffusion with good edge-preserving capabilities. A typical example of such regularization is total variation first proposed in [32]. However, the main contribution of this work lies in the development of the blur regularization term. We show that the blurs can be recovered from the LR images up to small ambiguity. One can consider this as a generalization of the results proposed for blur estimation in the case of MBD problems. This fundamental observation enables us to build a simple regularization term for the blurs even in the case of the SR problem. To tackle the minimization task, we use an alternating minimization approach (see Fig. 1), consisting of two simple linear equations.

The rest of the paper is organized as follows. Section II outlines the degradation model. In Section III, we present a procedure for volatile blur estimation. This effortlessly transforms into a regularization term of the BSR algorithm as described in Section IV. Finally, Section V illustrates applicability of the proposed method to real situations.

II. Mathematical Model

To simplify the notation, we will assume only images and PSFs with square supports. An extension to rectangular images is straightforward. Let $f[x, y]$ be an arbitrary discrete image of size $F \times F$, then $\mathbf{f}$ denotes an image column vector of size $F^2 \times 1$ and $\mathbf{C}_A\{f\}$ denotes a matrix that performs convolution of f with an image of size $A \times A$. The convolution matrix can have a different output size. Adopting the Matlab naming convention, we distinguish two cases: "full" convolution $\mathbf{C}_A\{f\}$ of size $(F+A-1)^2 \times A^2$ and "valid" convolution $\mathbf{C}_A^v\{f\}$ of size $(F-A+1)^2 \times A^2$. In both cases, the convolution matrix is a Toeplitz-block-Toeplitz matrix. We will not specify dimensions of convolution matrices if it is obvious from the size of the right argument.

Let us assume we have K different LR frames $\{g_k\}$ (each of size $G \times G$) that represent degraded (blurred and noisy) versions of the original scene. Our goal is to estimate the HR representation of the original scene, which we denoted as the HR image f of size $F \times F$. The LR frames are linked with the HR image through a series of degradations similar to those between $o[x, y]$ and g_k in (2). First f is geometrically warped $(\mathbf{W}_k)$, then it is convolved with a volatile PSF $(\mathbf{V}_k)$ and finally it is decimated $(\mathbf{D})$. The formation of the LR images in vector-matrix notation is then described as

$$\mathbf{g}_k = \mathbf{D}\mathbf{V}_k\mathbf{W}_k\mathbf{f} + \mathbf{n}_k \tag{3}$$

where $\mathbf{n}_k$ is additive noise present in every channel. The decimation matrix $\mathbf{D} = \mathbf{S}\mathbf{U}$ simulates the behavior of digital sensors by first performing convolution with the $U \times U$ sensor PSF $(\mathbf{U})$ and then downsampling $(\mathbf{S})$. The Gaussian function is widely accepted as an appropriate sensor PSF and it is also used here. Its justification is experimentally verified in [33]. A physical interpretation of the sensor blur is that the sensor is of finite size and it integrates impinging light over its surface. The sensitivity of the sensor is highest in the middle and decreases towards its borders with a Gaussian-like decay. Further, we assume that the subsampling factor (or SR factor, depending on the point of view), denoted by ε, is the same in both x and y directions. It is important to underline that ε is a user-defined parameter. In principle, $\mathbf{W}_k$ can be a very complex geometric transform that must be estimated by image registration or motion detection techniques. We have to keep in mind that subpixel accuracy in $\mathbf{g}_k$s is necessary for SR to work. Standard image registration techniques can hardly achieve this and they leave a small misalignment behind. Therefore, we will assume that complex geometric transforms are removed in the preprocessing step and $\mathbf{W}_k$ reduces to a small translation. Hence, $\mathbf{V}_k\mathbf{W}_k = \mathbf{H}_k$, where $\mathbf{H}_k$ performs convolution with the shifted version of the volatile PSF v_k, and the acquisition model becomes

$$\mathbf{g}_k = \mathbf{D}\mathbf{H}_k\mathbf{f} + \mathbf{n}_k = \mathbf{S}\mathbf{U}\mathbf{H}_k\mathbf{f} + \mathbf{n}_k. \tag{4}$$

The BSR problem then adopts the following form: We know the LR images $\{g_k\}$ and we want to estimate the HR image f for the given $\mathbf{S}$ and the sensor blur $\mathbf{U}$. To avoid boundary effects, we assume that each observation g_k captures only a part of f. Hence, $\mathbf{H}_k$ and $\mathbf{U}$ are "valid" convolution matrices $\mathbf{C}_F^v\{h_k\}$ and $\mathbf{C}_{F-H+1}^v\{u\}$, respectively. In general, the PSFs h_k are of different size. However, we postulate that they all fit into a $H \times H$ support.

In the case of $\varepsilon = 1$, the downsampling $\mathbf{S}$ is not present and we face a slightly modified MBD problem that has been solved elsewhere [12], [17]. Here, we are interested in the case of $\varepsilon > 1$, when the downsampling occurs. Can we estimate the blurs as in the case $\varepsilon = 1$? The presence of $\mathbf{S}$ prevents us from using the cited results directly. However, we will show that conclusions obtained for MBD apply here in a slightly modified form, as well.

III. Reconstruction of Volatile Blurs

Estimation of blurs in the MBD case (no downsampling) attracted considerable attention in the past. A wide variety of methods were proposed, such as in [12] and [13], that provide a satisfactory solution. For these methods to work correctly, certain channel disparity is necessary. The disparity is defined as weak co-primeness of the channel blurs, which states that the blurs have no common factor except a scalar constant. In other words, if the channel blurs can be expressed as a convolution of two subkernels, then there is no subkernel that is common to all blurs. An exact definition of weakly co-prime blurs can be found in [13]. Many practical cases satisfy the channel co-primeness, since the necessary channel disparity is mostly guaranteed by the nature of the acquisition scheme and random processes therein. We refer the reader to [12] for a relevant discussion. This channel disparity is also necessary for the BSR case.

Let us first recall how to estimate blurs in the MBD case and then we will generalize the results for integer downsampling factors. For the time being, we will omit noise n, until Section IV, where we will address it appropriately.

A. MBD Case

The decimation matrix $\mathbf{D}$ is not present in (4) and only convolution binds the input with the outputs. The acquisition model is of the SIMO type with one input channel f and K output channels g_k. Under the assumption of channel co-primeness, we can see that any two correct blurs h_i and h_j satisfy

$$g_i * h_j - g_j * h_i = 0. \tag{5}$$

There are $K(K-1)/2$ such relations and they can be arranged into one system. Let us define

$$\mathcal{N}_{\mathrm{MBD}} := \begin{pmatrix} \mathbf{Z}_1^T & \dots & \mathbf{Z}_{K-1}^T \end{pmatrix}^T$$

$$\mathbf{Z}_i := \underbrace{\begin{pmatrix} \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}}_{i-1 \text{ blocks}} \underbrace{\begin{pmatrix} \mathbf{G}_{i+1} & -\mathbf{G}_i & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_K & \mathbf{0} & \dots & -\mathbf{G}_i \end{pmatrix}}_{K-i+1 \text{ blocks}} \tag{6}$$

for $i = 1, \dots, K-1$, where $\mathbf{G}_i := \mathbf{C}_H^v\{g_i\}$. The complete system of relations (5) then takes the form

$$\mathcal{N}_{\mathrm{MBD}}\mathbf{h} = \mathbf{0} \tag{7}$$

where $\mathbf{h} = \left[\mathbf{h}_1^T, \dots, \mathbf{h}_K^T\right]^T$. In most real situations, the correct blur size (we have assumed square size $H \times H$) is not known

in advance, and, therefore, we can generate the above equation for different blur dimensions $\hat{H}_1 \times \hat{H}_2$. The nullity (null-space dimension) of $\mathcal{N}_{\mathrm{MBD}}$ is exactly 1 for the correctly estimated blur size. By applying SVD (singular value decomposition), we recover precisely the blurs except for a scalar factor. One can eliminate this magnitude ambiguity by stipulating that $\sum_{x,y} h_k[x,y] = 1$, which is a common brightness preserving assumption. For the underestimated blur size, the above equation has no solution. If the blur size is overestimated, then $\mathrm{nullity}(\mathcal{N}_{\mathrm{MBD}}) = (\hat{H}_1 - H + 1)(\hat{H}_2 - H + 1)$.

B. BSR Case

Before we proceed, it is necessary to define precisely the sampling matrix $\mathbf{S}$. Let $\mathbf{S}_1^\varepsilon$ denote a 1-D sampling matrix, where ε is the integer subsampling factor. Each row of the sampling matrix is a unit vector whose nonzero element is at such position that, if the matrix multiplies an arbitrary vector b, the result of the product is every εth element of b. If the vector length is M then the size of the sampling matrix is $(M/\varepsilon) \times M$. If M is not divisible by ε, we can pad the vector with an appropriate number of zeros to make it divisible. A 2-D sampling matrix is defined by

$$\mathbf{S}^\varepsilon := \mathbf{S}_1^\varepsilon \otimes \mathbf{S}_1^\varepsilon \tag{8}$$

where $\otimes$ denotes the matrix direct product (Kronecker product operator). Note that the transposed matrix $(\mathbf{S}^\varepsilon)^T$ behaves as an upsampling operator that interlaces the original samples with $(\varepsilon - 1)$ zeros.

A naive approach, as proposed in [27] and [28], is to modify (7) for the MBD case by applying downsampling, $\mathcal{N}_{\mathrm{BSR}} = \mathcal{N}_{\mathrm{MBD}}[\mathbf{I}_K \otimes \mathbf{S}^\varepsilon \mathbf{U}]$, and formulating the problem as

$$\min_{\mathbf{h}} \|\mathcal{N}_{\mathrm{BSR}}\mathbf{h}\|^2 \tag{9}$$

where $\mathbf{I}_K$ is the $K \times K$ identity matrix. One can easily verify that the condition in (5) is not satisfied for the BSR case as the presence of downsampling operators violates the commutative property of convolution. Even more disturbing is the fact that minimizers of (9) do not have to correspond to the correct blurs. We are going to show that if one uses a slightly different approach, reconstruction of the volatile PSFs h_k is possible even in the BSR case. However, we will see that some ambiguity in the solution of h_k is inevitable.

First, we need to rearrange the acquisition model (4) and construct from the LR images g_k a convolution matrix $\mathcal{G}$ with a predetermined nullity. Then, we take the null space of $\mathcal{G}$ and construct a matrix $\mathcal{N}$, which will contain the correct PSFs h_k in its null space.

Let $E \times E$ be the size of "nullifying" filters η_{kn}. (The meaning of this name will be clear later). Define $\mathcal{G} := [\mathbf{G}_1, \ldots, \mathbf{G}_K]$, where $\mathbf{G}_k := \mathbf{C}_E^v\{g_k\}$ are "valid" convolution matrices. Assuming no noise, we can express $\mathcal{G}$ in terms of f, u, and h_k as

$$\mathcal{G} = \mathbf{S}^\varepsilon \mathbf{F} \mathbf{U} \mathcal{H} \tag{10}$$

where

$$\mathcal{H} := \left[\mathbf{C}_{\varepsilon E}\{h_1\}(\mathbf{S}^\varepsilon)^T, \ldots, \mathbf{C}_{\varepsilon E}\{h_K\}(\mathbf{S}^\varepsilon)^T\right] \tag{11}$$

$\mathbf{U} := \mathbf{C}_{\varepsilon E+H-1}\{u\}$ and $\mathbf{F} := \mathbf{C}^v_{\varepsilon E+H+U-2}\{f\}$.

The convolution matrix $\mathcal{U}$ has more rows than columns, and, therefore, it is of full column rank (see proof in [12] for general convolution matrices). We assume that $\mathbf{S}^\varepsilon\mathbf{F}$ has full column rank as well. This is almost certainly true for real images if $\mathbf{F}$ has at least ε^2 times more rows than columns. Thus, $\mathrm{Null}(\mathcal{G}) \equiv \mathrm{Null}(\mathcal{H})$ and the difference between the number of columns and rows of $\mathcal{H}$ bounds from below the null space dimension, i.e.,

$$\mathrm{nullity}(\mathcal{G}) \geq KE^2 - (\varepsilon E + H - 1)^2. \tag{12}$$

Setting $N := KE^2 - (\varepsilon E + H - 1)^2$ and $\mathbf{N} := \mathrm{Null}(\mathcal{G})$, we visualize the null space as

$$\mathbf{N} = \begin{bmatrix} \mathbf{n}_{1,1} & \ldots & \mathbf{n}_{1,N} \\ \vdots & \ddots & \vdots \\ \mathbf{n}_{K,1} & \ldots & \mathbf{n}_{K,N} \end{bmatrix} \tag{13}$$

where $\mathbf{n}_{kn}$ is the vector representation of the nullifying filter η_{kn} of size $E \times E$, $k = 1, \ldots, K$ and $n = 1, \ldots, N$. The filters η_{kn} are made of values of $\mathcal{G}$'s null space and that is where their name comes from. Let $\tilde{\eta}_{kn}$ denote upsampled η_{kn} by factor ε, i.e., $\tilde{\eta}_{kn} := (\mathbf{S}^\varepsilon)^T \eta_{kn}$. Then, we define

$$\mathcal{N} := \begin{bmatrix} \mathbf{C}_H\{\tilde{\eta}_{1,1}\} & \ldots & \mathbf{C}_H\{\tilde{\eta}_{K,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_H\{\tilde{\eta}_{1,N}\} & \ldots & \mathbf{C}_H\{\tilde{\eta}_{K,N}\} \end{bmatrix} \tag{14}$$

and conclude that

$$\mathcal{N}\mathbf{h} = \mathbf{0} \tag{15}$$

where $\mathbf{h}^T = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$. We have arrived to an equation that is of the same form as (7) in the MBD case. Here, we have the solution to the blur estimation problem for the BSR case. However, since $\mathbf{S}^\varepsilon$ is involved, ambiguity of the solution is higher. Without proofs (for the sake of simplicity) we provide the following statements. For the correct blur size, $\mathrm{nullity}(\mathcal{N}) = \varepsilon^4$. For the underestimated blur size, (15) has no solution. For the overestimated blur size $\hat{H}_1 \times \hat{H}_2$, $\mathrm{nullity}(\mathcal{N}) = \varepsilon^2(\hat{H}_1 - H + \varepsilon)(\hat{H}_2 - H + \varepsilon)$. The conclusion may seem to be pessimistic. For example, for $\varepsilon = 2$ the nullity is at least 16, and for $\varepsilon = 3$ the nullity is already 81.

To shed more light on the above discussion about the nullity we have visualized the null space of $\mathcal{N}$ in Fig. 2. We convolved an image with six different 8×8 PSFs (the first PSF is in the top of Fig. 2), downsampled the blurred images with factor 2, and then constructed $\mathcal{N}$ from the images following the above derivation. We know that in this case 16 independent vectors span the null space of $\mathcal{N}$ and their arbitrary linear combination is a solution to (15). One such configuration of 16 independent vectors, where only the first PSF is extracted from each, is shown in the bottom of Fig. 2 arranged in a 4×4 table. One can see that the recovered PSFs contain parts of the original PSF and we have got four distinct parts each shifted to four different positions. Section IV will show that $\mathcal{N}$ plays an important role in the restoration algorithm as a consistency term and its ambiguity is not a serious drawback.

It is interesting to note that a similar derivation is possible for rational SR factors $\varepsilon = p/q$. We downsample the LR im-

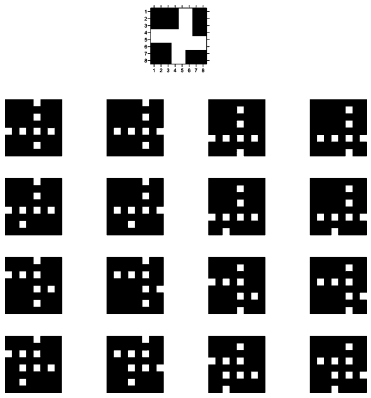


Fig. 2. Visualization of $\mathcal{N}$'s null space for $\varepsilon = 2$. (Top) Original 8×8 PSF and (bottom) one example of 16 PSFs that span the null space of $\mathcal{N}$. Proper linear combination of these 16 PSFs gives the original PSF.

ages with the factor q, thereby creating q^2K images and apply thereon the above procedure for the SR factor p.

Another consequence of the above derivation is the minimum necessary number of LR images for the blur reconstruction to work. The condition of the $\mathcal{G}$ nullity in (12) implies that the minimum number is $K > \varepsilon^2$. For example, for $\varepsilon = 3/2$, 3 LR images are sufficient; for $\varepsilon = 2$, we need at least 5 LR images to perform blur reconstruction. An intuitive explanation is that ε^2 input images are necessary for the SR problem to get a fully determined system of equations and additional input images are for the PSF estimation.

IV. Blind Superresolution

In order to solve the BSR problem, i.e., determine the HR image f and volatile PSFs h_k, we adopt an approach of minimizing a regularized energy function. This way, the method will be less vulnerable to noise and better posed. The energy consists of three terms and takes the form

$$E(\mathbf{f},\mathbf{h}) = \sum_{k=1}^{K} ||\mathbf{D}\mathbf{H}_k\mathbf{f} - \mathbf{g}_k||^2 + \alpha Q(\mathbf{f}) + \beta R(\mathbf{h}). \quad (16)$$

The first term measures the fidelity to the data and emanates from our acquisition model (4). The remaining two are regularization terms with positive weighting constants α and β that attract the minimum of E to an admissible set of solutions. The form of E very much resembles the energy proposed in [17] for MBD. Indeed, this should not come as a surprise since MBD and SR are related problems in our formulation.

Regularization $Q(\mathbf{f})$ is a smoothing term of the form

$$Q(\mathbf{f}) = \mathbf{f}^T\mathbf{L}\mathbf{f} \quad (17)$$

where $\mathbf{L}$ is a high-pass filter. A common strategy is to use convolution with the Laplacian for $\mathbf{L}$, which in the continuous case corresponds to $Q(f) = \int |\nabla f|^2$. Recently, variational integrals $Q(f) = \int \phi(|\nabla f|)$ were proposed, where ϕ is a strictly convex, nondecreasing function that grows at most linearly. Examples of $\phi(s)$ are s (total variation), $\sqrt{1+s^2} - 1$ (hypersurface minimal function), $\log(\cosh(s))$, or nonconvex functions, such as $\log(1+s^2)$, $s^2/(1+s^2)$ and $\arctan(s^2)$ (Mumford–Shah functional). The advantage of the variational approach is that it behaves as anisotropic diffusion. While in smooth areas it has the same isotropic behavior as the Laplacian, it also preserves edges in images. The disadvantage is that it is highly nonlinear. To overcome this difficulty one must use, e.g., the half-quadratic algorithm [34]. For the purpose of our discussion, it suffices to state that after discretization we arrive again at (17), where this time $\mathbf{L}$ is a positive semidefinite block tridiagonal matrix constructed of values depending on the gradient of f. The rationale behind the choice of $Q(f)$ is to constrain the local spatial behavior of images; it resembles a Markov random field. Some global constraints may be more desirable but are difficult (often impossible) to define, since we develop a general method that should work with any class of images.

The PSF regularization term $R(\mathbf{h})$ directly follows from the conclusions of the previous section. Since the matrix $\mathcal{N}$ in (14) contains the correct PSFs h_k in its null space, we define the regularization term as a least-squares fit

$$R(\mathbf{h}) = ||\mathcal{N}\mathbf{h}||^2 = \mathbf{h}^T\mathcal{N}^T\mathcal{N}\mathbf{h}. \quad (18)$$

If one replaces $\mathcal{N}$ with $\mathcal{N}_{\mathrm{BSR}}$, we have the naive approach. The product $\mathcal{N}^T\mathcal{N}$ is a positive semidefinite matrix. More precisely, R is a consistency term that binds volatile PSFs and prevents them from moving freely and, unlike the fidelity term [the first term in (16)], it is based solely on the observed LR images. A good practice is to include with a small weight a smoothing term $\mathbf{h}^T\mathbf{L}\mathbf{h}$ in $R(\mathbf{h})$. This is especially useful in the case of less noisy data in order to overcome the higher nullity of $\mathcal{N}$.

The complete energy then takes the form

$$E(\mathbf{f},\mathbf{h}) = \sum_{k=1}^{K} ||\mathbf{D}\mathbf{H}_k\mathbf{f} - \mathbf{g}_k||^2 + \alpha\mathbf{f}^T\mathbf{L}\mathbf{f} + \beta_1||\mathcal{N}\mathbf{h}||^2 + \beta_2\mathbf{h}^T\mathbf{L}\mathbf{h}. \quad (19)$$

Energy E as a function of both variables $\mathbf{f}$ and $\mathbf{h}$, is not convex due to convolution in the first term. On the other hand, the energy function is convex with respect to $\mathbf{f}$ if $\mathbf{h}$ is fixed and it is convex with respect to $\mathbf{h}$ if $\mathbf{f}$ is fixed. The minimization sequence $(\mathbf{f}^m, \mathbf{h}^m)$ can, thus, be built by alternating between two minimization subproblems. This procedure is called alternating minimizations (AM) and the advantage lies in its simplicity. For each subproblem a unique minimum exists that can be easily calculated. Derivatives w.r.t. $\mathbf{f}$ and $\mathbf{h}$ must be zero at the minima, which, in this case, leads to solving a set of simple linear equations. In conclusion, starting with some initial $\mathbf{h}^0$ the two iterative steps are

Step 1)
$$\mathbf{f}^m = \arg\min_{\mathbf{f}} E(\mathbf{f},\mathbf{h}^m) \Rightarrow \frac{\partial E}{\partial \mathbf{f}} = \mathbf{0} \Leftrightarrow \left(\sum_{k=1}^{K} \mathbf{H}_k^T\mathbf{D}^T\mathbf{D}\mathbf{H}_k + \alpha\mathbf{L}\right)\mathbf{f} = \sum_{k=1}^{K} \mathbf{H}_k^T\mathbf{D}^T\mathbf{g}_k \quad (20)$$

Step 2)
$$\mathbf{h}^{m+1} = \arg\min_{\mathbf{h}} E(\mathbf{f}^m,\mathbf{h}) \Rightarrow \frac{\partial E}{\partial \mathbf{h}} = \mathbf{0} \Leftrightarrow ([\mathbf{I}_K \otimes \mathbf{F}^T\mathbf{D}^T\mathbf{D}\mathbf{F}] + \beta_1\mathcal{N}^T\mathcal{N} + \beta_2\mathbf{L})\mathbf{h} = [\mathbf{I}_K \otimes \mathbf{F}^T\mathbf{D}^T]\mathbf{g} \quad (21)$$

where $\mathbf{F} := \mathbf{C}_H^v\{f\}$, $\mathbf{g} := [\mathbf{g}_1^T, \ldots, \mathbf{g}_K^T]^T$ and m is the iteration step. The AM approach is a variation on the steepest-descent algorithm. The search space is a concatenation of the blur subspace and the image subspace. The algorithm first descends in the image subspace and after reaching the minimum, i.e., $\partial E/\partial \mathbf{f} = \mathbf{0}$, it advances in the blur subspace in the direction $\partial E/\partial \mathbf{h}$ orthogonal to the previous one, and this scheme repeats. Due to the coupling of the variables by convolution, we cannot guarantee in theory that the global minimum is reached but thorough testing indicates good convergence properties of the algorithm for many real problems.

Convergence may further improve if we add feasible regions for the HR image and PSFs specified as lower and upper bounds constraints. To solve step 1, we use the method of conjugate gradients (function *cgs* in standard Matlab) and then adjust the solution $\mathbf{f}^m$ to contain values in the admissible range, typically, the range of values of $\mathbf{g}$. It is common to assume that PSF is positive and preserves image brightness, i.e., $h_k \geq 0$ and $\sum_{x,y} h_k[x,y] = 1$. We can, therefore, restrict the intensity values of PSFs between 0 and 1. In order to enforce the bounds in step 2, we solve (21) as a constrained minimization problem (function *fmincon* in Matlab Optimization Toolbox v.3) rather than using the projection as in step 1. Constrained minimization problems are more computationally demanding but we can afford it in this case since the size of $\mathbf{h}$ is much smaller than the size of $\mathbf{f}$.

The weighting constants α and β_i depend on the level of noise. If noise increases, α and β_2 should increase, and β_1 should decrease. One can use parameter estimation techniques, such as cross-validation [24] or expectation maximization [35], to determine the correct weights. However, in our experiments, we set the values manually according to a visual assessment. If the iterative algorithm begins to amplify noise, we have underestimated the noise level. On the contrary, if the algorithm begins to segment the image, we have overestimated the noise level.

V. Experiments

This section consists of two parts. In the first one, a set of experiments on synthetic data evaluate performance of the BSR algorithm with respect to the SR factor and compare the reconstruction quality with other methods mentioned below under different levels of noise. The second part demonstrates the applicability of the proposed method to real data.

In all the experiments the sensor blur is fixed and set to a Gaussian function of standard deviation $\sigma = 0.34$ (relative to the scale of LR images). One should underline that the proposed BSR method is fairly robust to the choice of the Gaussian variance, since it can compensate for insufficient variance by automatically including the missing factor of Gaussian functions in the volatile blurs.

Another potential pitfall that we have to take into consideration is a feasible range of SR factors. Theoretically there are no limitations on the upper bound of the SR factor. However, practical reasons impose limits. As the SR factor ε increases, we need more LR images ($K > \varepsilon^2$). The increasing number of LR images negatively affects the stability of BSR, since in real scenarios perturbations of the acquisition model occur, which disrupts the minimization scheme. SR factors beyond 2.5 are, thus, rare in real applications. A more elaborated discussion on fundamental limits of SR algorithms is given in [36]. In addition, rational SR factors p/q, where p and q are incommensurable and large regardless of the effective value of ε, also make the BSR algorithm unstable. It is the numerator p that determines the internal SR factor used in the algorithm. Hence, we limit ourselves to ε between 1 and 2.5, such as 3/2, 5/3, 2, etc., which is sufficient in most practical applications.

A. Simulated Data

First, let us demonstrate the BSR performance with a simple experiment. An 175×175 image in Fig. 3(a) blurred with six masks in Fig. 3(b) and downsampled with factor 2 gives six LR images. Using the LR images as an input, we estimated the original HR image with the proposed BSR algorithm for $\varepsilon = 1.5$ and 2. Fig. 4 summarizes obtained results in their original size. One can see, that for $\varepsilon = 1.5$ [Fig. 4(b)], the reconstruction is good but some details, such as the shirt texture, are still fuzzy. For the SR factor 2, the reconstructed image in Fig. 4(c) is almost perfect as most of the high-frequency information of the original image is correctly recovered.

Next, we evaluate noise robustness of the proposed BSR and compare it with other two methods: interpolation technique and state-of-the-art SR method. The former technique consists of the MBD method proposed in [17] followed by standard bilinear interpolation resampling. The MBD method first removes volatile blurs and then the interpolation of the deconvolved image achieves the desired spatial resolution. The latter method, which we will call herein a "standard SR algorithm," is a MAP formulation of the SR problem proposed, e.g., in [1] and [2]. This method uses a MAP framework for the joint estimation of image registration parameters (in our case only translation) and the HR image, assuming only the sensor blur ($\mathbf{U}$) and no volatile blurs. For an image prior, we use edge preserving Huber Markov random fields [33].

In the case of BSR, Section III shows that two distinct approaches exist for the blur estimation. Either we use the naive approach in (9) that directly utilizes the MBD formulation, or we apply the intrinsically SR approach given in (15). Depending on the approach, we use either $\mathcal{N}_{\mathrm{BSR}}$ or $\mathcal{N}$ in the blur consistency term $R(\mathbf{h})$ in the AM algorithm.

Altogether we have, thus, four distinct methods for comparison: standard SR approach, MBD with interpolation, BSR with naive blur regularization and BSR with intrinsic blur regularization. The experimental setup was the following. First, we generated six random motion blurs of size 4×4. Then we generated six LR images from the original HR image in Fig. 3(a) using the blurs and the downsampling factor of 2, and added white Gaussian noise with different SNR from 50 to 1 dB. The signal-to-noise ratio is defined as $\mathrm{SNR} = 10 \log\left(\sigma_f^2/\sigma_n^2\right)$, where σ_f and σ_n are the image and noise standard deviations, respectively. We repeated the whole procedure ten times for different realizations of noise. For each set of six LR images, the four methods were applied one by one. Parameters of each method were chosen to minimize the mean square error of the HR estimate. Fig. 5 summarizes the

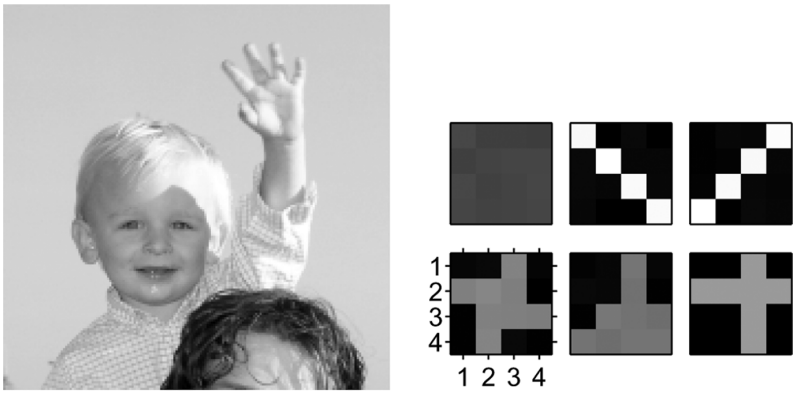


Fig. 3. Simulated data: (a) original 175 × 175 image; (b) six 4 × 4 volatile PSFs used to blur the original image.

Fig. 4. BSR of simulated data: (a) one of six LR images with the downsampling factor 2; (b) BSR for $\varepsilon = 1.5$; (c) BSR for $\varepsilon = 2$. The shirt texture is not yet visible for the SR factor 1.5 but becomes well reconstructed for the SR factor 2. On the other hand, face features probably lack very small details, and there is no visible improvement between 1.5 and 2.

obtained results in terms of peak signal to noise ratio defined as $\mathrm{PSNR}(\hat{\mathbf{f}}) = 10\log(\mathrm{span}(\mathbf{f})^2/\|\hat{\mathbf{f}} - \mathbf{f}\|^2/F^2)$, where $\hat{\mathbf{f}}$ is the estimate of the original HR image $\mathbf{f}$, and $\mathrm{span}(\mathbf{f})$ denotes the span of gray-level values in the original image, typically 255.

The standard SR method gives the poorest performance, since it lacks any apparatus for removing volatile blurs. MBD with interpolation removes blurs in the LR domain, which accounts for better performance. However, the best performance is apparent for the proposed BSR method, which accomplishes SR and blind deconvolution in the HR domain. The shape of the blur consistency term $R(\mathbf{h})$ plays its role, as well. In the case of the naive consistency term $(\mathcal{N}_{\mathrm{BSR}})$, estimated blurs are less accurate. This leads to tiny artifacts in the HR image and a small performance drop. On the other hand, the blur consistency term with $\mathcal{N}$ provides the most accurate estimations and outperforms all the other methods. For low SNR, all the tested methods tend to give similar results in the PSNR perspective and advantages of the proposed BSR method are less evident. Thus, for very noisy images (below 20 dB), it is sufficient to perform MBD with simple interpolation than to apply advanced SR methods, since MBD is definitely faster and the results look similar due to noise. The level of noise depends on the amount of light during acquisition and also on the quality of sensors. In our experience, most regular digital cameras have SNR around 50 dB, but

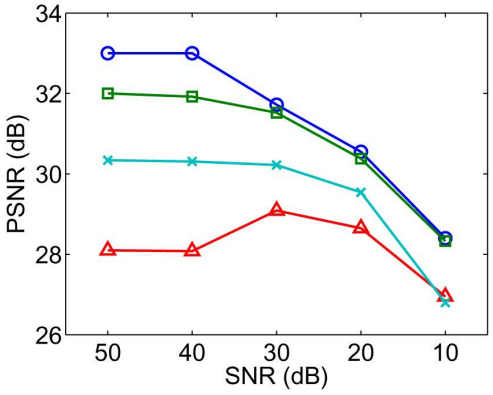
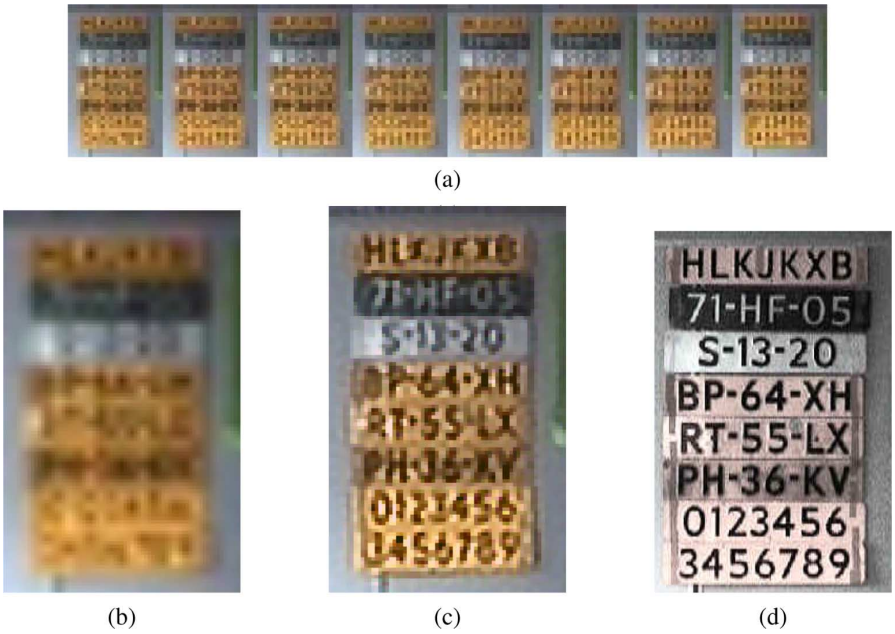
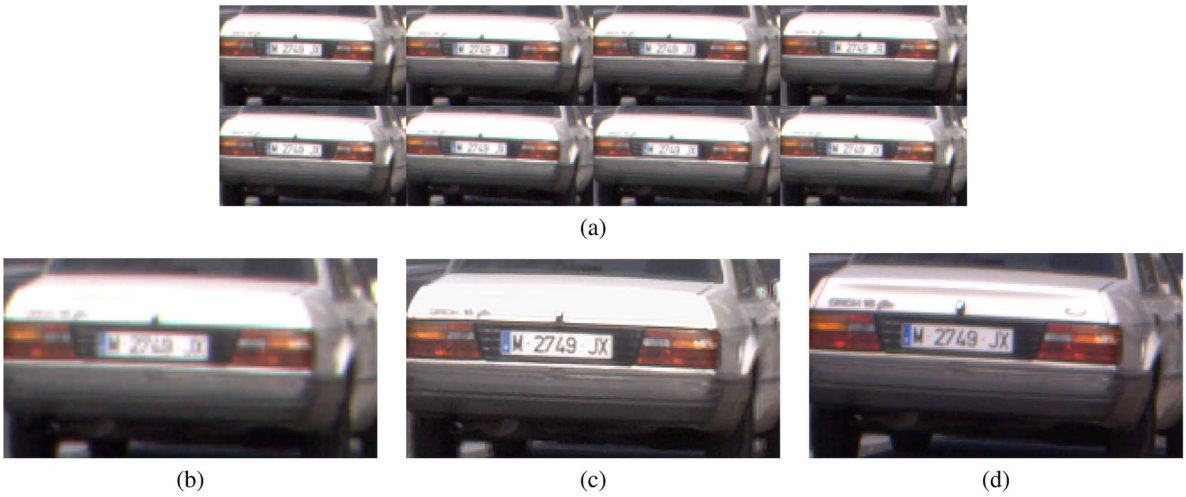


Fig. 5. Performance of the BSR algorithm and the other two methods under different levels of noise: (◯) BSR using $\mathcal{N}$ in the blur consistency term $R(\mathbf{h})$; (□) BSR using $\mathcal{N}_{\mathrm{BSR}}$; (×) MBD with bilinear interpolation; (△) standard SR method. Note that the proposed BSR outperforms any other method but as the noise level increases its supremacy becomes less evident.



(a) (b) (c) (d)

Fig. 6. Reconstruction of images acquired with a camcorder ($\varepsilon = 2.5$): (a) eight LR frames created from a short video sequence captured with the camcorder and displayed in their original size; (b) bilinear interpolation of one LR frame; (c) BSR estimate of the HR frame; (d) original HR frame.



(a) (b) (c) (d)

Fig. 7. Reconstruction of images acquired with a digital camera ($\varepsilon = 2$): (a) eight LR acquired shot with the digital camera and displayed in their original size; (b) bilinear interpolation of one LR image; (c) BSR estimate of the HR image; (d) image taken by the same camera but with optical zoom. The BSR algorithm achieves reconstruction comparable to the image with optical zoom.

with decreasing light, it can drop down to 30 dB. Webcameras have in general lower SNR around 30 dB, even in moderate light conditions.

B. Real Data

The next three experimental settings come from a license plate recognition task and they demonstrate the true power of the BSR algorithm. We used data from two different acquisition devices: camcorder and digital camera. The camcorder was Sony Digital Handycam and the digital camera was 5-Mpixel Olympus C5050Z equipped with 3× optical zoom. In order to work with color images, we extended the proposed BSR method by utilizing color TV [37] instead of standard TV in image regularization and by assuming the same blurring in all three color channels.

In the first scenario, we used a short video sequence provided by Dr. Z. Geradts from the Netherlands Forensic Institute (available at forensic.to/superresolution.htm). The video sequence was acquired with the camcorder and was artificially downsampled with factor 10. We extracted 16 frames from the downsampled video, of which eight are in Fig. 6(a), and applied the proposed BSR algorithm with the SR factor of 2.5. Fig. 6(b)

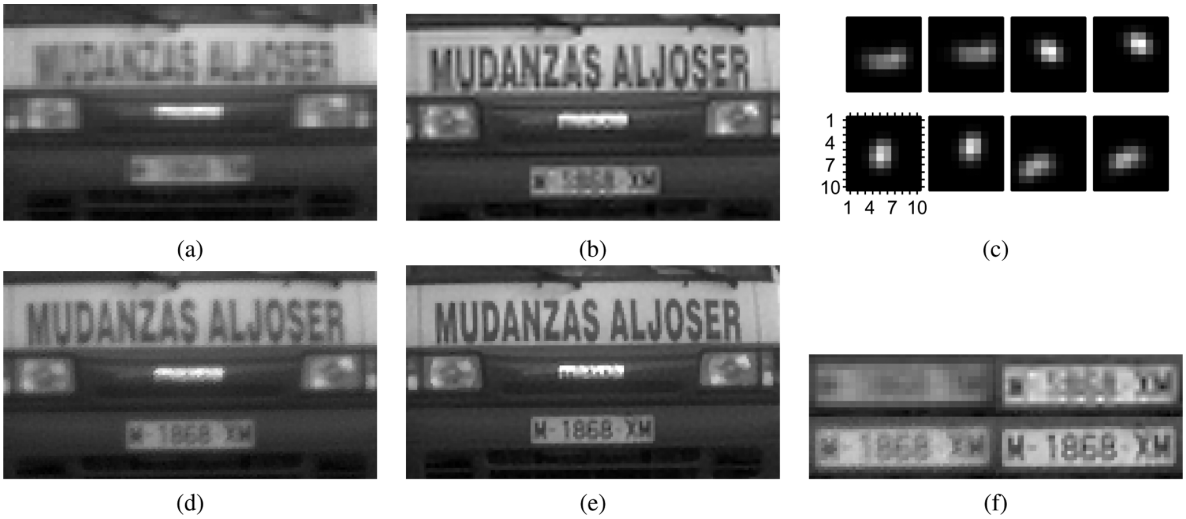


Fig. 8. License-plate recognition ($\varepsilon = 2$): (a) one of eight LR images acquired with a digital camera (zero-order interpolation); (b) MBD followed by bilinear interpolation; (c) PSFs estimated by the proposed BSR; (d) standard SR algorithm; (e) proposed BSR algorithm; (f) closeups of the images (a) and (b) on top and (d) and (e) on bottom. Note that only the BSR result (e) reconstructs the car brand name in such a way that we can deduce that it was a "Mazda" car.

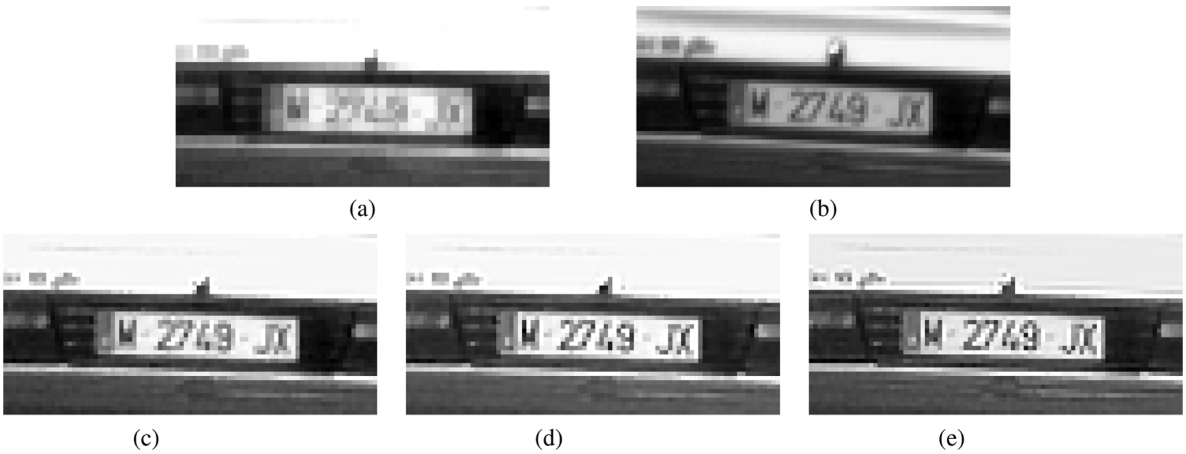


Fig. 9. Performance of the BSR algorithm with respect to the number of LR images ($\varepsilon = 1.5$). (a) One of eight LR images of size 40×70, zero-order interpolation. (b) Image acquired with optical zoom 1.5 ×, which plays the role of "ground truth." The proposed BSR algorithm using (c) 3, (d) 4, and (e) 8 LR images.

shows the first LR frame bilinearly interpolated to have the size of HR images. The HR frame estimated by BSR is in Fig. 6(c), and the original undecimated HR frame is in Fig. 6(d). The obtained result remarkably well recovers letters and numbers on the license plates.

In the second scenario, we used the digital camera and took eight photos of a stalled car, registered the photos with cross-correlation and cropped each to a 100×50 rectangle. All eight cuttings printed in their original size (no interpolation), including one image enlarged with bilinear interpolation, are in Fig. 7(a) and (b). We set the desired SR factor to 2 and applied BSR. In order to better assess the obtained results, we took one additional image with optical zoom set close to 2 ×. This image served as the ground truth; see Fig. 7(d). The proposed BSR method returned a well reconstructed HR image [Fig. 7(c)], which surpasses the image acquired with the optical zoom.

The third experimental setting consisted of a car moving towards a hand-held digital camera. We took four consecutive color images with the camera, and using both green channels (color image in digital cameras are made of two green channels and one red and one blue channel), we generated altogether eight LR images. The images were roughly registered with cross-correlation and cropped each to a 90×50 rectangle. One such image is in Fig. 8(a). We set the SR factor to 2 and applied different reconstruction techniques. The MBD with interpolation method [Fig. 8(b)] reconstructed the banner satisfactory, yet the license plate is not legible, since it contains tiny details that are beyond the resolution of LR images. The standard SR approach in Fig. 8(d) gives moderate results. The proposed BSR method in Fig. 8(e) outperforms all the other techniques and provides a sharp HR image. The PSFs estimated by BSR are in Fig. 8(c). Note that every second PSF is a shifted

version of the previous one, which was expected, since green channels in digital cameras are shifted diagonally by 1 pixel in each direction. For better visual comparison closeups of one of the input LR image and three reconstructed HR images appear in Fig. 8(f).

When dealing with real data, one cannot expect that the performance will increase without limits as the number of available LR images increases. At a certain point possible discrepancies between the measured data and our mathematical model take over, and the estimated HR image does not improve any more or it can even worsen. We conducted several experiments on real data (short shutter speed and motionless objects) with different SR factors and number of LR images K. See the results of one such experiment in Fig. 9 for $\varepsilon = 1.5$ and the number of LR images ranging from 3 to 8. A small improvement is apparent between using 3 and 4 LR images; compare Fig. 9(c) and (d). However, the result obtained with all eight images in Fig. 9(e) shows a very little improvement. We deduce that for each SR factor exists an optimal number of LR images that is close to the minimum necessary number. Therefore, in practice, we recommend to use the minimum or close to minimum number of LR images for the given SR factor.

VI. Conclusion

We have shown that the SR problem permits a stable solution, even in the case of unknown blurs. The fundamental idea is to split radiometric deformations into sensor and volatile parts and assume that only the sensor part is known. We can then construct a convex functional using the observed LR images and observe that the volatile part minimizes this functional. Due to the presence of resolution decimation, the functional is not strictly convex and reaches its minimum on a subspace that depends on the integer SR factor. We have also extended our conclusions to rational factors. To achieve robust solution, we have adopted the regularized energy minimization approach. The proposed BSR method goes far beyond the standard SR techniques. The introduction of volatile blurs makes the method particularly appealing to real situations. While reconstructing the blurs, we estimate not only subpixel shifts but also any possible blurs imposed by the acquisition process. To our knowledge, this is one of the first methods that performs deconvolution and resolution enhancement simultaneously.

References

[1] R. Hardie, K. Barnard, and E. Armstrong, "Joint map registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1621–1633, Dec. 1997.

[2] C. Segall, A. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Trans. Image Process.*, vol. 13, no. 7, pp. 898–911, Jul. 2004.

[3] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.

[4] N. Woods, N. Galatsanos, and A. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 201–213, Jan. 2006.

[5] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, pp. 977–1000, 2003.

[6] R. Lagendijk, J. Biemond, and D. Boekee, "Identification and restoration of noisy blurred images using the expectation-maximization algorithm," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 7, pp. 1180–1191, Jul. 1990.

[7] S. Reeves and R. Mersereau, "Blur identification by the method of generalized cross-validation," *IEEE Trans. Image Process.*, vol. 1, no. 7, pp. 301–311, Jul. 1992.

[8] T. Chan and C. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.

[9] M. Haindl, "Recursive model-based image restoration," in *Proc. 15th Int. Conf. Pattern Recognition*, 2000, vol. III, pp. 346–349.

[10] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.

[11] D. Kundur and D. Hatzinakos, "Blind image deconvolution revisited," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 61–63, Nov. 1996.

[12] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 202–219, Feb. 1999.

[13] G. Giannakis and R. Heath, "Blind identification of multichannel FIR blurs and perfect image restoration," *IEEE Trans. Image Process.*, vol. 9, no. 11, pp. 1877–1896, Nov. 2000.

[14] H.-T. Pai and A. Bovik, "On eigenstructure-based direct multichannel blind image restoration," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1434–1446, Oct. 2001.

[15] G. Panci, P. Campisi, S. Colonnese, and G. Scarano, "Multichannel blind image deconvolution using the bussgang algorithm: Spatial and multiresolution approaches," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1324–1337, Nov. 2003.

[16] F.Šroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 9, p. 1094, Sep. 2003.

[17] F.Šroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Process.*, vol. 14, no. 7, pp. 874–883, Jul. 2005.

[18] Y.-L. You and M. Kaveh, "Blind image restoration by anisotropic regularization," *IEEE Trans. Image Process.*, vol. 8, no. 3, pp. 396–407, Mar. 1999.

[19] A. Rajagopalan and S. Chaudhuri, "An MRF model-based approach to simultaneous recovery of depth and restoration from defocused images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 7, pp. 577–589, Jul. 1999.

[20] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, Mar. 2003.

[21] S. Farsui, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *Int. J. Imag. Syst. Technol.*, vol. 14, no. 2, pp. 47–57, Aug. 2004.

[22] E. Shechtman, Y. Caspi, and M. Irani, "Space-time super-resolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 531–545, Apr. 2005.

[23] M. Ben-Ezra and S. Nayer, "Motion-based motion deblurring," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 6, pp. 689–698, Jun. 2004.

[24] N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement," *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1299–1308, Sep. 2001.

[25] N. Woods, N. Galatsanos, and A. Katsaggelos, "EM-based simultaneous registration, restoration, and interpolation of super-resolved images," in *Proc. IEEE Int. Conf. Image Processing*, 2003, vol. 2, pp. 303–306.

[26] D. Rajan and S. Chaudhuri, "Simultaneous estimation of super-resolved scene and depth map from low resolution defocused observations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1102–1117, Sep. 2003.

[27] F.Šroubek and J. Flusser, "Resolution enhancement via probabilistic deconvolution of multiple degraded images," *Pattern Recognit. Lett.*, vol. 27, pp. 287–293, Mar. 2006.

[28] Y. Chen, Y. Luo, and D. Hu, "A general approach to blind image super-resolution using a PDE framework," *Proc. SPIE*, vol. 5960, pp. 1819–1830, 2005.

[29] Wirawan, P. Duhamel, and H. Maitre, "Multi-channel high resolution blind image restoration," in *Proc. IEEE ICASSP*, 1999, pp. 3229–3232.

[30] A. Yagle, "Blind superresolution from undersampled blurred measurements," in *Proc. Advanced Signal Processing Algorithms, Architectures, and Implementations XIII*, 2003, vol. 5205, pp. 299–309.

[31] D. Biggs, C. L. Wang, T. Holmes, and A. Khodjakov, "Subpixel deconvolution of 3D optical microscope imagery," *Proc. SPIE*, vol. 5559, pp. 369–380, Oct. 2004.

[32] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
[33] D. Capel, *Image Mosaicing and Super-Resolution*. New York: Springer, 2004.
[34] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*. New York: Springer Verlag, 2002.
[35] R. Molina, M. Vega, J. Abad, and A. Katsaggelos, "Parameter estimation in Bayesian high-resolution image reconstruction with multisensors," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1655–1667, Dec. 2003.
[36] Z. Lin and H.-Y. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 83–97, Jan. 2004.
[37] D. Tschumperlé and R. Deriche, "Diffusion PDE's on vector-valued images," *IEEE Signal Process. Mag.*, vol. 19, no. 5, pp. 16–25, Sep. 2002.

Filip Šroubek received the M.Sc. degree in computer science from the Czech Technical University, Prague, Czech Republic, in 1998, and the Ph.D. degree in computer science from the Charles University, Prague, in 2003.

From 2004 to 2006, he was in a postdoctoral position at the Instituto de Optica, CSIC, Madrid, Spain. He is currently with the Institute of Information Theory and Automation and also with the Institute of Photonics and Electronics, Academy of Sciences of the Czech Republic. He is an author of two book chapters and over 25 journal and conference papers on image fusion, blind deconvolution, superresolution, and related topics.

Gabriel Cristóbal (SM'96) received the M.Sc. and Ph.D. degrees in telecommunication engineering from the Universidad Politecnica of Madrid, Madrid, Spain, in 1979 and 1986, respectively.

He was Visiting Scholar at the International Computer Science Institute and an Associate Researcher at the University of California, Berkeley, from 1989 to 1992. He is currently a Research Scientist with the Instituto de Optica, Spanish Council for Scientific Research, Madrid. His current research interest are joint representations, vision modeling, resolution enhancement, and image compression.

Jan Flusser (SM'03) received the M.Sc. degree in mathematical engineering from the Czech Technical University, Prague, Czech Republic, in 1985, the Ph.D. degree in computer science from the Czechoslovak Academy of Sciences in 1990, and the D.Sc. degree in technical cybernetics in 2001.

Since 1985, he has been with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague. From 1995 to 2006, he held the position of Head of the Department of Image Processing. In 2007, he was appointed to Director of the Institute. Since 1991, he has also been affiliated with the Charles University, Prague, and the Czech Technical University, Prague, where he teaches courses on digital image processing and pattern recognition. He has been a Full Professor since 2004. His current research interests include all aspects of digital image processing and pattern recognition, namely 2-D object recognition, moment invariants, blind deconvolution, image registration, and image fusion. He has authored and coauthored more than 150 research publications in these areas.

Advance Access publication on December 10, 2007 doi:10.1093/comjnl/bxm098

Super-Resolution and Blind Deconvolution For Rational Factors With an Application to Color Images

FILIP ŠROUBEK[1,*], JAN FLUSSER[1] AND GABRIEL CRISTÓBAL[2]

[1]*Institute of Information Theory and Automation, AS CR, Pod vodárenskou věží 4, 182 08, Prague 8, Czech Republic*
[2]*Instituto de Óptica, CSIC, Serrano 121, 28006 Madrid, Spain*
**Corresponding author: sroubekf@utia.cas.cz*

In many real applications, traditional super-resolution (SR) methods fail to provide high-resolution images due to objectionable blur and inaccurate registration of input low-resolution images. Only integer resolution enhancement factors, such as 2 or 3, are often considered, but non-integer factors between 1 and 2 are also important in real cases. We introduce a method to SR and deconvolution, which assumes no prior information about the shape of degradation blurs, incorporates registration parameters, and is properly defined for any rational (fractional) resolution factor. The method minimizes a regularized energy function with respect to the high-resolution image and blurs, where regularization is carried out in both the image and blur domains. The blur regularization is based on a generalized multi-channel blind deconvolution constraint derived in the paper. An extension to color images is briefly discussed. Experiments on real data illustrate robustness to noise and other advantages of the method.

Keywords: super-resolution; blind deconvolution; alternating minimization; polyphase components

Received 19 December 2006; revised 24 October 2007

1. INTRODUCTION

Imaging devices have limited achievable resolution due to many theoretical and practical restrictions. An original scene with a continuous intensity function $o(x, y)$ warps at the camera lens because of the scene motion and/or the change of the camera position. In addition, several external effects blur images: atmospheric turbulence, camera lens, relative camera-scene motion, etc. We will call these effects *volatile blurs* to emphasize their unpredictable and transitory behavior, yet we will assume that we can model them as convolution with an unknown point spread function (PSF) $v(x, y)$. This is a reasonable assumption if the original scene is flat and perpendicular to the optical axis. Finally, charge-coupled devices (CCDs) discretize the image and produces digitized noisy image $g(i, j)$, which we refer to as a *low-resolution* (LR) *image*, since the spatial resolution is too low to capture all the details of the original scene. For one single observation $g(i, j)$, the problem is heavily underdetermined and lacks stable solution. To partially overcome the equivocation of the problem, we can take K ($K > 1$) images of the original scene and face the so-called multi-channel (multi-frame) problem. The acquisition model then becomes

$$g_k(i, j) = D([v_k * W_k(o)](x, y)) + n_k(i, j) \quad (1)$$

where $k = 1, \ldots, K$ is the acquisition index, $n_k(i, j)$ the additive noise and W_k the geometric deformation (warping), in general different for each acquisition. $D(\cdot)$ is the *decimation operator* that models the function of CCD sensors. It consists of convolution with a *sensor PSF* followed by a *sampling operator*, which we define as multiplication by a sum of delta functions placed on a grid. The above model is the state of the art as it takes all possible degradations into account.

Super-resolution (SR) is the process of combining a sequence of LR images in order to produce an image or sequence of higher resolution. It is unrealistic to assume that the super-resolved image can recover the original scene $o(x, y)$ exactly. A reasonable goal of SR is a discrete version of $o(x, y)$, which has higher spatial resolution than the resolution of the LR images and which is free of the volatile blurs (deconvolved). In the paper, we will refer to this super-resolved image as a *high-resolution* (HR) *image* $f(i, j)$ and the ratio between the

size of the sought HR image and input LR image will be called an *SR factor* and denoted by ε. The standard SR approach consists of subpixel registration, overlaying the LR images on an HR grid, and interpolating the missing values. The subpixel shift between images thus constitutes an essential feature. We will demonstrate that considering volatile blurs in the model explicitly brings about a more general and robust technique, with the subpixel shift being a special case thereof.

The acquisition model in equation (1) embraces three distinct cases frequently encountered in the literature. First, removal of the geometric degradation W_k is a registration problem. Second, if the decimation operator D and the geometric transform W_k are not considered, then we face a *multi-channel* (or multi-frame) *blind deconvolution* (MBD) problem. Third, if the volatile blur v_k is not considered or assumed known, and W_k is suppressed except to subpixel translations, we obtain a classical SR formulation. In practice, it is crucial to consider all three cases at once. We are then confronted with a problem of *blind SR* (BSR), the topic of this paper.

Proper registration techniques can suppress large and complex geometric distortions but usually a small between-image shift is still observable. There have been hundreds of methods proposed; see e.g. [1] for a survey. Here, we will assume that registration parameters can be calculated by one of the methods, and if applied, the LR images are registered except to small translations.

Research on intrinsically MBD methods has begun fairly recently; refer to [2–6] for a survey and other references. The MBD methods can directly recover the blurring functions from the degraded images alone. We further developed the MBD theory in [7] by proposing a blind deconvolution method for images, which might be mutually shifted by unknown vectors. To make this brief survey complete, we should not forget to mention a very challenging problem of shift-variant blind deconvolution, that was considered in [8, 9].

A countless number of papers address the standard SR problem. A good survey can be found in [10, 11]. Maximum likelihood, maximum *a posteriori* (MAP), the set theoretic approach using projection on convex sets and fast Fourier techniques can all provide a solution to the SR problem. Earlier approaches assumed that subpixel shifts are estimated by other means. More advanced techniques, such as in [12–14], include the shift estimation in the SR process. Other approaches focus on fast implementation [15], space–time SR [16], SR with complex image priors for joint image and segmentation estimation [17] or SR of compressed video [13]. Most of the SR techniques assume *a priori* known blurs. However, in many cases, such as camera motion, blurs can have wild shapes that are difficult to predict; see examples of real motion blurs in [18]. Nguyen *et al.* [19], Woods *et al.* [20] and Rajan and Chaudhuri [21] proposed BSR that can handle parametric PSFs, i.e. PSFs modeled with one parameter. This restriction is unfortunately very limiting for most real applications. In [22], we extended our MBD method to BSR in an intuitive way but one can prove that this approach does not estimate PSFs accurately. The same intuitive approach was also proposed in [23]. To our knowledge, first attempts for theoretically correct BSR with an arbitrary PSF appeared in [24,25]. The interesting idea proposed therein is to use the so-called polyphase components. We will adopt the same idea here as well. Other preliminary results of the BSR problem with focus on fast calculation are given in [26], where the authors propose a modification of the Richardson–Lucy algorithm.

Current MBD techniques require no or very little prior information about the blurs, they are sufficiently robust to noise and provide satisfying results in most real applications. However, they can hardly cope with the decimation operator, which violates the standard convolution model. On the contrary, state-of-the-art SR techniques achieve remarkable results of resolution enhancement in the case of no blur. They accurately estimate the subpixel shift between images but lack any apparatus to calculate the blurs.

Recently, in [27], we proposed a unifying method that simultaneously estimates the volatile blurs and HR image. The only prior knowledge required are estimates of the blur size and the level of noise in the LR images, which renders it a truly BSR method. The key idea was to determine subpixel shifts by calculating volatile blurs. As the volatile blurs are estimated in the HR scale, positions of their centroids correspond to sub pixel shifts. Therefore, by estimating blurs, we automatically estimate shifts with subpixel accuracy, which is essential for good performance of SR. This complex SR problem was solved by minimizing a regularized energy function, where the regularization was carried out in both the image and blur domains. The image regularization is based on variational integrals, and a consequent anisotropic diffusion with good edge-preserving capabilities. The blur regularization term is based on our generalized result of blur estimation in the SR case. To tackle the minimization task, we used an alternating minimization (AM) approach consisting of two simple linear equations.

In this work, we extend the BSR method by incorporating registration parameters and color images. To address correctly the correlation of color channels, we apply the regularizations proposed in [28]. Further, we show that by using the polyphase decomposition, we can formulate the SR problem not only for integer factors but also for any rational (fractional) factor, which is important in real applications where often only an SR factor between 1 and 2 is possible. If the noise removal and registration steps are not sufficiently reliable, a factor of 1.6 is the practical limit of SR [29]. The need for rational factors in SR was also reported in [30] with a focus on fast calculation using preconditioners.

Section 2 outlines the degradation model in the discrete domain for integer and non-integer SR factors. In Section 3, we present the regularized energy functional, derive the regularization terms and sketch an extension to color images.

The AM scheme and parameter estimation is given in Section 4. Finally, Section 5 illustrates the applicability of the proposed method to real situations.

2. DISCRETE MODEL

To simplify the notation, we will assume only images and PSFs with square supports. An extension to rectangular images is straightforward. Let $f(i, j)$ be an arbitrary discrete image of size $F \times F$ then $\mathbf{f}$ denotes an image column vector of size $F^2 \times 1$ and $\mathbf{C}_A\{f\}$ denotes a matrix that performs convolution of f with an image of size $A \times A$. The convolution matrix can have a different output size. Adopting the Matlab naming convention, we distinguish two cases: 'full' convolution $\mathbf{C}_A\{f\}$ of size $(F + A - 1)^2 \times A^2$ and 'valid' convolution $\mathbf{C}_A^v\{f\}$ of size $(F - A + 1)^2 \times A^2$. In both cases, the convolution matrix is of the Toeplitz-block-Toeplitz form. In the paper, we will not specify dimensions of convolution matrices if it is obvious from the size of the right argument.

Before we proceed, it is necessary to define precisely the *sampling matrix* $\mathbf{S}$. Let $\mathbf{S}^\varepsilon$ denote an 1-D sampling matrix, where ε is the integer subsampling factor and $i = 1,\ldots,\varepsilon$. Each row of the sampling matrix is a unit vector whose non-zero element is at such position that, if the matrix multiplies an arbitrary vector b, the result of the product is every εth element of b starting from b_i. If the vector length is M then the size of the sampling matrix is $\lceil (M - i + 1)/\varepsilon \rceil \times M$, where $\lceil \cdot \rceil$ rounds up to the closest integer. A 2-D sampling matrix is defined by

$$\mathbf{S}^\varepsilon_{i,j} := \mathbf{S}^\varepsilon_i \otimes \mathbf{S}^\varepsilon_j, \tag{2}$$

where $\otimes$ denotes the matrix direct product (Kronecker product operator). If the starting index (i, j) will be (1, 1) then we will omit the subscript and simply write $\mathbf{S}^\varepsilon$. Note that the transposed matrix $(\mathbf{S}^\varepsilon)^{\mathrm{T}}$ behaves as an upsampling operator that interlaces the original samples with $(\varepsilon - 1)$ zeros.

Polyphase components of an image $f(x, y)$ are

$$\mathbf{f}^{ij} := \mathbf{S}^\varepsilon_{i,j}\mathbf{f} \tag{3}$$

which is equivalent to

$$\mathbf{f}^{ij} := [f[i,j], f[i+\varepsilon, j], f[i+2\varepsilon, j], \ldots, f[i, j+\varepsilon], f[i+\varepsilon, j+\varepsilon], \ldots]^{\mathrm{T}}.$$

Therefore, each image breaks into ε^2 distinct polyphase components (downsampled versions of the image) (see Fig. 1).

Let us now define a discrete version of the acquisition model in equation (1). Assume that we have K different LR frames $\{g_k\}$ (each of size $G \times G$) that represent degraded (blurred and noisy) versions of the original scene. Our goal is to estimate the HR representation of the original scene, which we denoted as the HR image f of size $F \times F$. The LR frames are linked with the HR image through a series of degradations similar to those between $o(x, y)$ and g_k in equation (1). First f is geometrically warped ($\mathbf{W}_k$), then it is convolved with a volatile PSF ($\mathbf{V}_k$) and finally it is decimated ($\mathbf{D}$). The decimation operator $\mathbf{D}$ depends only on sensor characteristics and since we assume the same sensor in all acquisitions it appears without the index k. The formation of the LR images in vector–matrix notation is then described as

$$\mathbf{g}_k = \mathbf{D}\mathbf{V}_k\mathbf{W}_k\mathbf{f} + \mathbf{n}_k, \tag{4}$$

where $\mathbf{n}_k$ is additive noise present in every channel. In principle, $\mathbf{W}_k$ can be a very complex geometric transform that must be estimated by image registration or motion detection techniques. We have to keep in mind that the subpixel accuracy in $\mathbf{g}_k$'s is essential for SR to work properly. Standard image registration techniques can hardly achieve this and they leave a small misalignment behind. Then the warping operator splits into $\mathbf{W}_k = \mathbf{T}_k\mathbf{W}'_k$, where $\mathbf{W}'_k$ is estimated by some registration methods and $\mathbf{T}_k$ is the unknown translation. In order to change the order of geometrical warping and convolution, we consider only linear transformations. Unknown translation $\mathbf{T}_k$ combined with volatile PSF $\mathbf{V}_k$ gives us $\mathbf{V}_k\mathbf{T}_k = \mathbf{H}_k$, where $\mathbf{H}_k$ performs convolution with the shifted version of the volatile PSF v_k. The decimation matrix $\mathbf{D} = \mathbf{S}^\varepsilon\mathbf{U}$ simulates the behavior of digital sensors by first performing convolution with the $U \times U$ sensor PSF ($\mathbf{U}$) and then downsampling ($\mathbf{S}^\varepsilon$) by factor ε. Assuming linear transformations, $\mathbf{W}'_k$ may be grouped with the decimation operator resulting in $\mathbf{D}_k = \mathbf{S}^\varepsilon\ \mathbf{U}\mathbf{W}'_k$. From the numerical point of view, it is preferable to construct directly the whole $\mathbf{D}_k$ than the individual matrices $\mathbf{S}^\varepsilon$, $\mathbf{U}$ and $\mathbf{W}'_k$. Finally, the acquisition model reads

$$\mathbf{g}_k = \mathbf{D}_k\mathbf{H}_k\mathbf{f} + \mathbf{n}_k = \mathbf{S}^\varepsilon\mathbf{U}\mathbf{W}'_k\mathbf{H}_k\mathbf{f} + \mathbf{n}_k. \tag{5}$$

The BSR problem we are solving is the following: we know the LR images $\{g_k\}$ and $\mathbf{D}_k$'s, and we want to estimate the HR

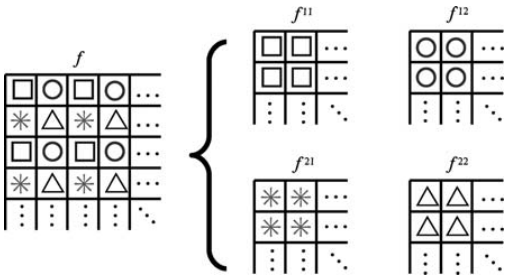


FIGURE 1. Polyphase decomposition for $\varepsilon = 2$: original image f decomposes into four downsampled images.

image f. In addition, the rough estimates of the size of volatile blur h_k and of the noise variance σ_n is necessary. PSFs h_k can be of different size. However, we postulate that they all fit into an $H \times H$ support. To avoid boundary effects, we assume that each observation g_k captures only a part of f. Hence, $\mathbf{H}_k$ and $\mathbf{U}$ are 'valid' convolution matrices $\mathbf{C}^v_F\{h_k\}$ and $\mathbf{C}^v_{F-H+1}\{u\}$, respectively.

To be able to deal correctly with non-integer SR factors, we need to express the above model using a sampling operator with integer ε. This will be necessary in the derivation of blur regularization in the next section. It can be done for factors expressed as a fraction p/q, where p and q are positive integers and $p > q$ (p and q are reduced so that they do not have any common factor).

Let $\varepsilon = p/q$ and the sampling frequency of the LR images g_k be q, then the sampling frequency (number of pixels per unit distance) of the HR image f is p. From each LR image g_k, we generate q^2 polyphase components. We consider these polyphase components as new input (downsampled-LR) images with the sampling frequency 1. Now, to obtain the HR image from the downsampled-LR images, we must solve an SR problem with the integer factor $\varepsilon = p$ and not with the rational one as before. In other words, in order to obtained an integer SR factor we downsample the LR images and thus artificially increase the number of channels. However, the number of unknown PSFs h_k remains the same. We still have K PSFs since every pack of q^2 downsampled-LR images contains the same blur.

An equivalent formulation of the model in equation (5) but for fractional SR factors p/q is

$$\begin{bmatrix} \mathbf{g}_k^{11} \\ \vdots \\ \mathbf{g}_k^{qq} \end{bmatrix} = \begin{bmatrix} \mathbf{S}^p\mathbf{U}_{1,1} \\ \vdots \\ \mathbf{S}^p\mathbf{U}_{q,q} \end{bmatrix} \mathbf{W}'_k\mathbf{H}_k\mathbf{f} + \mathbf{n}_k, \tag{6}$$

where each $\mathbf{U}_{i,j}$ performs convolution with one of the q^2 discretizations of the sensor PSF u and $\mathbf{g}_k^{ij} = \mathbf{S}^g_{i,j}\mathbf{g}_k$ are polyphase components of g_k for SR factor q. It is important to understand the discretization of the sensor PSF u in the case of fractional SR factors. Since p is not divisible by q, the product $\mathbf{S}^\varepsilon\mathbf{U}$ is shift-variant and it depends on a relative shift between the HR and LR pixels. One can readily see that the relative shift repeats every qth pixels (in both directions x and y) of the LR image and therefore we have q^2 distinct PSF discretizations. For further details see [31].

3. BLIND SUPERRESOLUTION

In order to solve the BSR problem, i.e. determine the HR image f and volatile PSFs h_k, we adopt an approach of minimizing a regularized energy function. This way the method will be less vulnerable to noise and better posed. The energy consists of three terms and takes the form

$$E(\mathbf{f}, \mathbf{h}) = \sum_{k=1}^{K} \|\mathbf{D}_k\mathbf{H}_k\mathbf{f} - \mathbf{g}_k\|^2 + Q(\mathbf{f}) + R(\mathbf{h}), \tag{7}$$

where $\mathbf{h} = [\mathbf{h}_1^{\mathrm{T}}, \ldots, \mathbf{h}_K^{\mathrm{T}}]^{\mathrm{T}}$. The first term measures the fidelity to the data and emanates from our acquisition model (5). The remaining two are regularization terms that attract the minimum of E to an admissible set of solutions. The form of E very much resembles the energy proposed in [7] for MBD. Indeed, this should not come as a surprise since MBD and SR are related problems in our formulation.

3.1. Image regularization

Regularization $Q(\mathbf{f})$ is a smoothing term of the form

$$Q(\mathbf{f}) = \alpha\mathbf{f}^{\mathrm{T}}\mathbf{L}\mathbf{f}, \tag{8}$$

where $\mathbf{L}$ is a high-pass filter and α is a positive regularization parameter. A common strategy is to use convolution with the Laplacian for $\mathbf{L}$, which in the continuous case, corresponds to $Q(f) = \int |\nabla f|^2$. Recently, variational integrals $Q(f) = \int \phi(|\nabla f|)$ were proposed, where ϕ is a strictly convex, non-decreasing function that grows at most linearly. Examples of $\phi(s)$ are s (total variation), $\sqrt{1 + s^2} - 1$ (hypersurface minimal function), $\log(\cosh(s))$ or non-convex functions, such as $\log(1 + s^2)$, $s^2/(1 + s^2)$ and $\arctan(s^2)$ (Mumford–Shah functional). The advantage of the variational approach is that while in smooth areas it has the same isotropic behavior as the Laplacian, it also preserves edges in images. The disadvantage is that it is highly nonlinear and to overcome this difficulty, one must use, e.g. half-quadratic algorithm [32]. For the purpose of our discussion, it suffices to state that after discretization we arrive again at equation (8), where this time $\mathbf{L}$ is a positive semi-definite block tridiagonal matrix constructed of values depending on the gradient of f. The rationale behind the choice of $Q(f)$ is to constrain the local spatial behavior of images; it resembles a Markov Random Field. Some global constraints may be more desirable but are difficult (often impossible) to define, since we develop a general method that should work with any class of images.

3.2. PSF regularization

Our PSF regularization term $R(\mathbf{h})$ consists of two terms. The first one is the same smoothing term as for images but applied to blurs, which is a typical prior that penalizes jagged blurs that are rare in real situations. The second term is a consistency term that binds the different volatile PSFs to prevent them from moving freely and unlike the fidelity term [the first term in equation (7)] it is based solely on the observed LR images. It takes the form of $\|\mathcal{N}\mathbf{h}\|^2$, where the

matrix $\mathcal{N}$ will be derived later. The complete PSF regularization is then given by

$$R(\mathbf{h}) = \beta \mathbf{h}^{\mathrm{T}} \mathbf{L} \mathbf{h} + \gamma \|\mathcal{N}\mathbf{h}\|^2, \tag{9}$$

where β and γ are positive regularization parameters that give different weights to the terms.

Consider the discrete model in equation (6) but without noise $\mathbf{n}_k$ and warping $\mathbf{W}_k'$ (this can be removed by registering the LR images). Let E be a positive integer and $\mathcal{G} := [\mathbf{G}_1^{11}, \ldots, \mathbf{G}_1^{qq}, \mathbf{G}_2^{11}, \ldots, \ldots, \mathbf{G}_K^{qq}]$, where $\mathbf{G}_k^{ij} := \mathbf{C}_E^v\{g_k^{ij}\}$. There are q^2 distinct discretizations of the sensor PSF u that depend on the relative shift between the HR and LR pixels. Let $u_{i,j}$ $(i, j = 1,\ldots,q)$ denote such discretizations.

Matrix $\mathcal{G}$ expressed in terms of f, u and h_k takes the form

$$\mathcal{G} = \mathbf{S}^p \mathbf{F}[\mathbf{U}_{1,1}, \ldots, \mathbf{U}_{q,q}]\mathcal{H}, \tag{10}$$

where $\mathbf{U}_{i,j} = \mathbf{C}_{pE-p+H}\{u_{i,\,j}\}$, $\mathbf{F} = \mathbf{C}^v_{pE-p+H+U-1}\{f\}$ and

$$\mathcal{H} := [\mathbf{I}_{q^2} \otimes (\mathbf{C}_{pE-p+1}\{h_1\}(\mathbf{S}^p)^{\mathrm{T}}), \ldots, \mathbf{I}_{q^2} \otimes (\mathbf{C}_{pE-p+1}\{h_K\} \times (\mathbf{S}^p)^{\mathrm{T}})].$$

The size of the upsampling matrix $(\mathbf{S}^p)^{\mathrm{T}}$ inside $\mathcal{H}$ is $(pE - p + 1)^2 \times E^2$.

If $\mathbf{S}^p\mathbf{F}$ is of full column rank, which is almost certainly true for real and sufficiently large images (see [31] for more details), then $\mathrm{Null}(\mathcal{G}) \equiv \mathrm{Null}([\mathbf{U}_{1,1}, \ldots, \mathbf{U}_{\mathrm{q,q}}]\mathcal{H})$. The difference between the number of columns and rows of $[\mathbf{U}_{1,1}, \ldots, \mathbf{U}_{\mathrm{q,q}}]\mathcal{H}$ bounds from below the dimension of $\mathcal{G}$'s null space, i.e.

$\mathrm{nullity}(\mathcal{G}) \geq$

$$N := K(qE)^2 - (pE - p + H + U - 1)^2. \tag{11}$$

Let $\mathbf{N}$ denote N null vectors of $\mathcal{G}$ stacked column-wise and η_{kn} are some $E \times E$ filters. We can visualize $\mathbf{N}$ as

$$\mathbf{N} = \begin{bmatrix} \boldsymbol{\eta}_{1,1} & \cdots & \boldsymbol{\eta}_{1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{q^2,1} & \cdots & \boldsymbol{\eta}_{q^2,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{Kq^2,1} & \cdots & \boldsymbol{\eta}_{Kq^2,N} \end{bmatrix}, \tag{12}$$

where $\boldsymbol{\eta}_{kn}$ is the vector representation of η_{kn}. Let $\tilde{\eta}_{kn}$ denote upsampled η_{kn} by factor p. Then

$$\mathcal{N} := \begin{bmatrix} \mathbf{C}_{U+H-1}\{\tilde{\eta}_{1,1}\} & \cdots & \mathbf{C}_{U+H-1}\{\tilde{\eta}_{Kq^2,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{U+H-1}\{\tilde{\eta}_{1,N}\} & \cdots & \mathbf{C}_{U+H-1}\{\tilde{\eta}_{Kq^2,N}\} \end{bmatrix} \times \mathbf{I}_K \otimes \begin{bmatrix} \mathbf{C}_H\{u_{1,1}\} \\ \vdots \\ \mathbf{C}_H\{u_{q,q}\} \end{bmatrix} \tag{13}$$

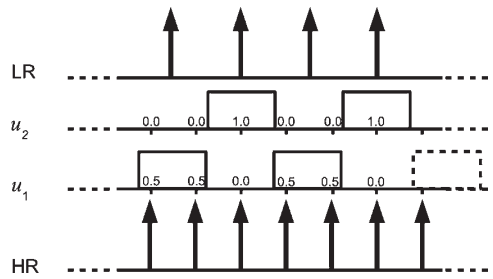
and we conclude (without proofs for the sake if simplicity) that

$$\mathcal{N}\mathbf{h} = \mathbf{0}. \tag{14}$$

An interesting observation follows from the nullity condition in equation (11): $K > (pE - p + H + U - 1)^2/(qE)^2$ and if $E \gg (p + H + U - 1)$ then $K > (p/q)^2$. It implies that the minimum number of input channels necessary for blur reconstruction to work is $K > \varepsilon^2$. For example, for $\varepsilon = 3/2$, three LR images are sufficient; for $\varepsilon = 2$, we need at least five LR images to perform blur reconstruction. Note that for no SR ($\varepsilon = 1$), the minimum number of input channels is 2, which is of course in accordance with the MBD theory.

To better understand the above derivation, the following example illustrates all the steps for an 1-D case.

EXAMPLE Let the HR signal be an 1-D periodic pulse f with the period [1, 1, 1, 1,0,0,0]. Such signal satisfies the necessary condition that $\mathbf{S}^p\mathbf{F}$ is of full-column rank. The HR signal is blurred by $h_1 = [0,1]$ and $h_2 = [1,\ 0]$ $(H = 2)$ and downsampled by factor $\varepsilon = p/q = 3/2$. To simplify the derivation, the sensor PSF u will be a pulse of length 1.5 HR pixels and its two $(q = 2)$ distinct discretizations are $u_1 = [0.5,\ 0.5,\ 0]$ and $u_2 = [0,\ 0,\ 1]$ $(U = 3)$ as depicted below:



The two $(K = 2)$ measured LR signals are thus

$$g_1 = [1, 1, 0.5, 0, 0.5, 1, 1, 0, 0, 1, 1, 1, 0, 0, \ldots],$$
$$g_2 = [1, 1, 0, 0, 1, 1, 0.5, 0, 0.5, 1, 1, 0, 0, 1, \ldots]$$

and their two ($q = 2$) polyphase components are

$$\begin{aligned} g_1^1 &= [1 \quad 0.5 \quad 0.5 \quad 1 \quad 0 \quad 1 \quad 0 \quad \ldots], \\ g_1^2 &= [1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad \ldots], \\ g_2^1 &= [1 \quad 0 \quad 1 \quad 0.5 \quad 0.5 \quad 1 \quad 0 \quad \ldots], \\ g_2^2 &= [1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad \ldots]. \end{aligned}$$

In the 1-D case, condition (11) reads $\text{nullity}(\mathcal{G}) \geq N := K(qE) - (pE - p + H + U - 1)$. Therefore, for the minimum admissible nullity $N = 1$, the size of filters η must be $E = 2$. From the LR signals, we thus construct

$$\mathcal{G} = \begin{bmatrix} 0.5 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0.5 & 0.5 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0.5 & 0 & 1 & 0.5 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0.5 & 0.5 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0.5 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

The null space of $\mathcal{G}$ is a single vector

$$\mathbf{N} = \begin{bmatrix} -0.6325 \\ 0 \\ -0.3162 \\ 0 \\ 0.6325 \\ 0 \\ 0 \\ 0.3162 \end{bmatrix}.$$

Extracting filters η's of length $E = 2$ from $\mathbf{N}$ and upsampling by $p = 3$, gives

$$\begin{aligned} \tilde{\eta}_{1,1} &= [\quad -0.6325 \quad 0 \quad 0 \quad 0 \quad], \\ \tilde{\eta}_{2,1} &= [\quad -0.3162 \quad 0 \quad 0 \quad 0 \quad], \\ \tilde{\eta}_{3,1} &= [\quad 0.6325 \quad 0 \quad 0 \quad 0 \quad], \\ \tilde{\eta}_{4,1} &= [\quad 0 \quad 0 \quad 0 \quad 0.3162 \quad]. \end{aligned}$$

Using equation (13) construct

$$\mathcal{N} = \begin{bmatrix} -0.3162 & 0 & 0 & 0 \\ -0.3162 & -0.3162 & 0.3162 & 0 \\ -0.3162 & -0.3162 & 0.3162 & 0.3162 \\ 0 & -0.3162 & 0.3162 & 0.3162 \\ 0 & 0 & 0 & 0.3162 \end{bmatrix}$$

and one can readily see that

$$\mathcal{N}[h_1, h_2]^{\mathrm{T}} = \mathbf{0}.$$

The matrix $\mathcal{N}$ contains the correct blurs in its null space. In real cases, when noise is present, we consider the l_2 norm as specified in equation (9).

In the course of $\mathcal{N}$'s derivation, we have to construct $\mathcal{G}$, which is huge even for images of moderate size, and then estimate its null space. However, efficient computation exists. The N smallest eigenvectors of $\mathcal{G}\mathcal{G}^{\mathrm{T}}$ give the null space matrix $\mathbf{N}$. The product $\mathcal{G}\mathcal{G}^{\mathrm{T}}$ is a square matrix of size proportional to E and can be calculated directly without first constructing $\mathcal{G}$; see e.g. [2] for details. E is the size of filters η and it is calculated from condition (11). It increases with the increasing SR factor ε and blur size H, but it decreases with the increasing number of LR images K. In general cases, the values of E are smaller or close to H. Therefore, the product $\mathcal{G}\mathcal{G}^{\mathrm{T}}$ is relatively small and the computation of $\mathcal{N}$ is fast.

3.3. Extension to color

There are three possible extensions of the acquisition model (1) to color images: assuming same blurs in color channels assuming different blurs in color channels and additionally assuming also intrachannel blurs between color channels ('crosstalks'). For the first two extensions, the PSF regularization term (9) can be used without any modifications. The third extension brings extra burden of crosstalks that prevents us from using the proposed PSF regularization and therefore we did not consider it here. Color channels are strongly correlated and it is highly desirable to introduce some coupling in the image regularization term (8). One can find a very good overview of different regularizations of color images in [28]. Here, we use the vector version of the variational approach, which is given by

$$Q(f) = \int \phi\left(\sqrt{\|\nabla f_r\|^2 + \|\nabla f_g\|^2 + \|\nabla f_b\|^2}\right) dx\, dy,$$

where $f_{\mathrm{r}}, f_{\mathrm{g}}, f_{\mathrm{b}}$ are red, greed, blue channels, respectively. Correlation of colors is appropriately addressed by this term and we can use the same numerical computation as in the case of gray-scale images (see Section 3.1).

4. ALTERNATING MINIMIZATION

The complete energy function reads

$$\begin{aligned} E(\mathbf{f}, \mathbf{h}) = &\sum_{k=1}^{K} \|\mathbf{D}_k \mathbf{H}_k \mathbf{f} - \mathbf{g}_k\|^2 \\ &+ \alpha \mathbf{f}^{\mathrm{T}} \mathbf{L} \mathbf{f} + \mathbf{h}^{\mathrm{T}} (\beta \mathbf{L} + \gamma \mathcal{N}^{\mathrm{T}} \mathcal{N}) \mathbf{h}. \end{aligned} \tag{15}$$

To find a minimizer, we perform AMs of E over $\mathbf{f}$ and $\mathbf{h}$. The advantage of this scheme lies in its simplicity. Each term of equation (15) is quadratic and therefore convex (but not

necessarily strictly convex) and the derivatives w.r.t. **f** and **h** are easy to calculate. This AM approach is a variation on the steepest-descent algorithm. The search space is a concatenation of the blur subspace and the image subspace. The algorithm first descends in the image subspace and after reaching the minimum, i.e. $\nabla_{\mathbf{f}} E = 0$, it advances in the blur subspace in the direction $\nabla_{\mathbf{h}}$E orthogonal to the previous one, and this scheme repeats. In conclusion, starting with some initial $\mathbf{h}^0$ the two iterative steps are:

Step 1.

$$\begin{aligned} \mathbf{f}^m &= \arg\min_{\mathbf{f}} E(\mathbf{f}, \mathbf{h}^m) \\ &\Leftrightarrow \text{solve for } \mathbf{f} \\ &\left(\sum_{k=1}^{K} \mathbf{H}_k^{\mathrm{T}} \mathbf{D}_k^{\mathrm{T}} \mathbf{D}_k \mathbf{H}_k + \alpha \mathbf{L}\right) \mathbf{f} = \sum_{k=1}^{K} \mathbf{H}_k^{\mathrm{T}} \mathbf{D}_k^{\mathrm{T}} \mathbf{g}_k \,. \end{aligned} \tag{16}$$

Step 2.

$$\begin{aligned} \mathbf{h}^{m+1} &= \arg\min_{\mathbf{h}} E(\mathbf{f}^m, \mathbf{h}) \\ &\Leftrightarrow \text{solve for } \mathbf{h} \\ &([\mathbf{I}_K \otimes \mathbf{F}^{\mathrm{T}} \mathbf{D}_k^{\mathrm{T}} \mathbf{D}_k \mathbf{F}] + \gamma \mathcal{N}^{\mathrm{T}} \mathcal{N} + \beta \mathbf{L}) \mathbf{h} \\ &= [\mathbf{I}_K \otimes \mathbf{F}^{\mathrm{T}} \mathbf{D}_k^{\mathrm{T}}] \mathbf{g}, \end{aligned} \tag{17}$$

where $\mathbf{F} := \mathbf{C}_H^v\{f\}$, $\mathbf{g} := [\mathbf{g}_1^{\mathrm{T}}, \ldots, \mathbf{g}_K^{T}]^{\mathrm{T}}$ and m is the iteration step. Note that both steps consist of simple linear equations.

Energy E as a function of both variables **f** and **h** is not convex due to coupling of variables via convolution in the first term of equation (15). Therefore, it is not guaranteed that the BSR algorithm reaches the global minimum, instead, one may get trapped in local minima. In our experience, convergence properties improve significantly if we add feasible regions for the HR image and PSFs specified as lower and upper bounds constraints. To solve Step 1, we use the method of conjugate gradients (function *cgs* in Matlab) and then adjust the solution $\mathbf{f}^m$ to contain values in the admissible range, typically, the range of values of **g**. It is common to assume that PSF is positive ($h_k \geq 0$) and that it preserves image brightness. We can therefore write the lower and upper bounds constraints for PSFs as $\mathbf{h}_k \in \langle 0,1 \rangle^{H^2}$. In order to enforce the bounds in Step 2, we solve equation (17) as a constrained minimization problem (function *fmincon* in Matlab) rather than using the projection as in Step 1. Constrained minimization problems are more computationally demanding but we can afford it in this case since the size of **h** is much smaller than the size of **f**.

Parameters α, β and γ depend on the level of noise. If noise increases, α and β should increase, and γ should decrease. One can prove that α and β are proportional to σ_n^2, which is the noise variance. Estimation techniques, such as cross-validation [19] or expectation maximization [33], can be used to determine the correct weights. However, we did not want to increase the complexity of the problem any further and thus we set the values in experiments manually according to our visual assessment. If the iterative algorithm begins to amplify noise, we have underestimated the noise level. On contrary, if the algorithm begins to segment the image, we have overestimated the noise level.

5. EXPERIMENTS

The following experiments with the proposed BSR method aim to first compare performance with other techniques and second demonstrate its applicability to real scenarios with misregistered input images and non-integer SR factors.

5.1. PSNR performance

We evaluated noise robustness of the proposed BSR and compared it with other two methods: interpolation technique and state-of-the-art SR method. The former technique consists of the MBD method proposed in [7] followed by the standard bilinear interpolation resampling. The MBD method first removes volatile blurs and then the interpolation of the deconvolved image achieves the desired spatial resolution. The latter method, which we will call herein a 'standard SR method', is a MAP formulation of the SR problem proposed, in [12, 13]. This method uses a MAP framework for the joint estimation of image registration parameters (in our case only translation) and the HR image, assuming only the sensor blur (**U**) and no volatile blurs. For an image prior, we used edge preserving

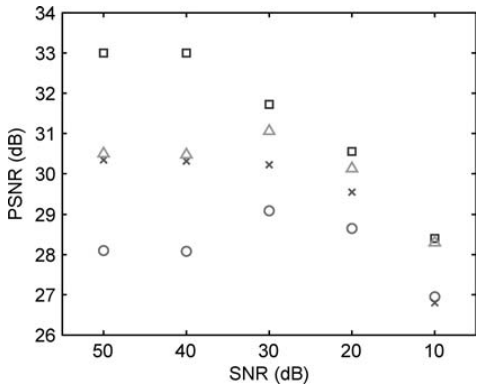


FIGURE 2. Performance of the BSR algorithm and the other two methods under different levels of noise: squares proposed BSR with $\beta = 0$; triangles proposed BSR with $\gamma = 0$; cross symbols represent MBD with bilinear interpolation; the circles represent the standard SR method. Note that the proposed BSR outperforms any other method but as the noise level increases its supremacy becomes less evident.

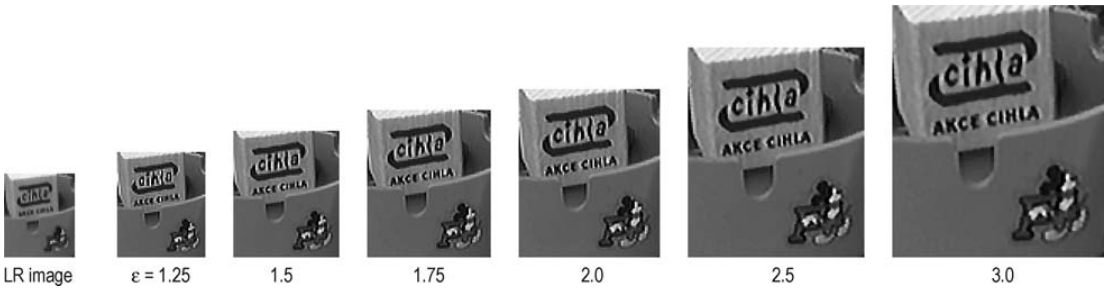


FIGURE 3. SR with non-integer factors of short-exposure images. The first left image is one of 10 LR frames acquired by a webcamera (exposure time 1/60 s) that were used to estimate HR images. The proposed BSR method was initialized with different SR factors from 1.25 to 3. The estimated HR images appear in their original size. An example of estimated PSFs for factor 2 is in Fig. 4.

Huber Markov random fields [34]. In order to evaluate the effect of the PSF regularization term (9), we ran the BSR method in two different modes. In the first mode, $\beta = 0$ and so the $\mathcal{N}$ term was considered. In the second mode, $\gamma = 0$ and so the standard smoothing term was considered.

The experimental setup was the following. First we generated six random motion blurs of size 4×4. Then we generated six LR images from the original 'Lena' image using the blurs and the downsampling factor of 2, and added white Gaussian noise with different signal-to-noise ratio (SNR) from 50 to 10 dB. SNR $= 10 \log(\sigma_f^2/\sigma_n^2)$, where σ_f and σ_n are the image and noise standard deviations, respectively. We repeated the whole procedure 10 times for different realizations of noise. For each set of six LR images, the three methods were applied one by one. Parameters of each method were chosen to minimize the mean square error of the HR estimate. Figure 2 summarizes the obtained results in terms of peak SNR defined as $\mathrm{PSNR}(\hat{\mathbf{f}}) = 10 \log\left(\frac{\mathrm{span(f)}^2}{\|\hat{\mathbf{f}} - \mathbf{f}\|^2/F^2}\right)$, where $\hat{\mathbf{f}}$ is the estimate of the original HR image $\mathbf{f}$, and span($\mathbf{f}$) denotes the span of gray-level values in the original image, typically 255.

The standard SR method gives the poorest performance, since it lacks any apparatus for removing volatile blurs. MBD with interpolation removes blurs in the LR domain, which accounts for better performance. However, the best performance is apparent for the proposed BSR method with $\mathcal{N}$ in the PSF regularization term. For low SNR, all the tested methods tend to give similar results in the PSNR perspective and advantages of the proposed BSR method are less evident. Thus, for very noisy images ($<$20 dB), it is sufficient to perform MBD with simple interpolation than to apply advance SR methods, since MBD is definitely faster and the results look similar due to noise.

5.2. Real data

We worked with a standard webcamera to record short video sequences of still scenes, extracted several consecutive frames and used the frames as input LR images. The input parameters of the BSR method were selected manually to give the best possible results. Common to all the experiments was the choice of the sensor blur, which was determined experimentally and was set to the Gaussian function of standard deviation $\sigma = 0.34$ (relative to the scale of LR images). One should underline that the proposed method is fairly robust to the underestimated size of the sensor blur, since it can compensate for insufficient variance by automatically including the missing factor of Gaussian functions in the volatile blurs. The quality of reconstructed HR images is not evaluated by any quantitative measure. Instead, we advocate the use of reader's subjective assessment.

The first experiment summarized in Figs. 3 and 4 compares results for different SR factors from 1.25 up to 3. In this case, the hand-held webcamera operated in good light conditions with the exposure time of 1/60 s and therefore no volatile blur was visible. Ten frames (see one such frame in the left side Fig. 3) extracted from the video served as input LR images for the BSR method with parameters $\alpha = 2 \times 10^{-3}$, $\gamma = 10$, $\beta = 0$, blur size 8×8 and six different SR factors $\varepsilon = 1.25, 1.5, 1.75, 2.0, 2.5, 3.0$. The HR images show improvement as the SR factor increases; however, the refinement becomes less visible after 2.0. The difference between 2.5 and 3.0 seems to be primarily only in size as no more details appear. In all the cases, estimated PSFs were more or less the same and an example for SR factor 2 is in Fig. 4. The PSFs are very localized and

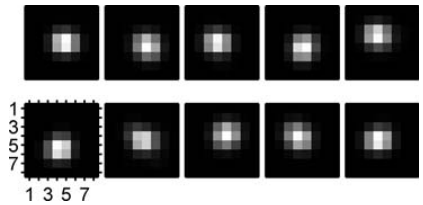


FIGURE 4. SR of short-exposure images. PSFs of LR images in Fig. 3 estimated by the proposed BSR method.

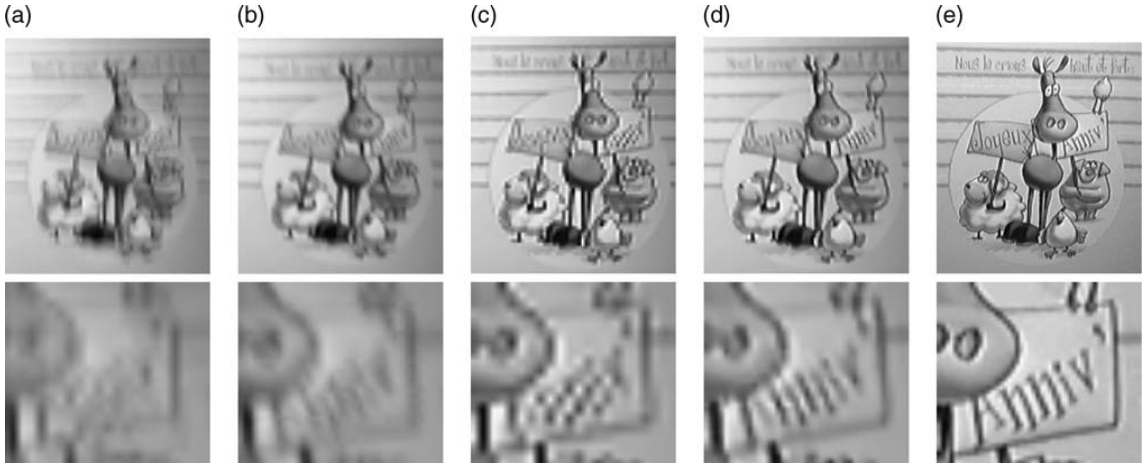


FIGURE 5. SR of long-exposure images. Five LR frames were extracted from a short video sequence captured by a webcamera (exposure time 1/10 s). An example of one frame and its close-up is in (a) top and bottom, respectively. The image is printed in the size of the output HR image for comparison reasons. Notice severe motion blur due to the long exposure time and motion of the hand-held camera. Estimated HR images for factor 5/3 were calculated by three different methods: standard SR in (b), MBD followed by bilinear interpolation in (c), and proposed BSR method in (d) (see PSFs estimated by BSR in Fig. 6). Compare obtained results with the image (e) acquired with the same webcamera but installed closer to the object.

they resemble delta functions (note that the displayed PSFs include the sensor blur), which was expected since the input images did not exhibit any volatile blur. Notice that the PSFs are shifted to compensate for subtle misalignments, which seamlessly performs subpixel registration and thus accurate SR.

In the second experiment, we recorded a greeting card with the hand-held webcamera in poor light conditions. The exposure time of the camera was 1/10 s and severe motion blur coming from the hand vibration is visible in images; see Fig. 5a. Then we move the camera toward the object, stabilize it, and grabbed one frame (Fig. 5e) to have an 'ideal' representation, which we show here only for evaluation reasons. Using five consecutive frames from the video, the HR image (SRF = 5/3) was estimated with three methods. The first result in Fig. 5b) was achieved by the standard SR method [13]. The result is poor as the method does not have means of removing blur. The second result in Fig. 5c shows MBD [7] with interpolation. The reconstructed image is sharper but many details are still missing, e.g. note the erroneous reconstruction of 'Anniv' on the flag held by the 'pig' right to the 'horse'. The third result is of the proposed BSR (Fig. 5d) run with parameters $\alpha = 2 \times 10^{-3}$ $\beta = 1$, $\gamma = 10$ and blur size 12×12. The obtained result after only three iterations of the AM algorithm is the sharpest with many details properly reconstructed. The key to successful reconstruction lies in the accurate estimation of PSFs within the iterative algorithm. As one could see in Fig. 6, the estimated PSFs model camera shake and since the whole procedure runs in the HR scale, the proposed method outperforms the former ones.

The third experiment demonstrates the advantage of using the decimation operators $\mathbf{D}_k$'s with registration instead of running BSR on registered LR images. A similar video sequence was recorded as in the previous case but this time we also rotated the camera during shooting to introduce geometric distortions that must be first eliminated by registration. An example of two such frames out of 10 is in Fig. 7a. Rotation is clearly visible and BSR cannot be applied without first registering the images. We estimated registration parameters and compared two approaches. First, we applied BSR on registered images; see the result in Fig. 7b. Second, we used the registration parameters to construct $\mathbf{D}_k$'s and applied BSR on the original unregistered images; see the result in Fig. 7c. In both cases, the parameters were set as in the previous experiment. Some small details are better reconstructed in the second approach, which indicates that using the registration parameters directly in BSR is preferable.

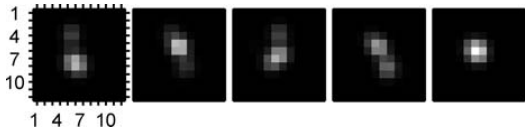


FIGURE 6. SR of long-exposure images. PSFs of images in Fig. 5(a) estimated by the proposed BSR method.

FIGURE 7. SR with registration. Ten frames were extracted from a video sequence recorded by a webcamera (exposure time 1/10 s). Two such frames are in (a) printed in the size of output HR images for comparison reasons. Due to slight rotation of the webcamera during shooting, registration parameters were first estimated by a registration method. Two approaches were compared: (b) Input frames were registered according to the parameters and BSR was applied on the registered images. (c) The registration parameters were used to modify decimation operators and BSR was applied to the original unregistered images. Better performance of the second approach is apparent from close-ups in the top row.

6. CONCLUSION

This paper presented an SR method for both integer and non-integer (rational) factors, which proved to be meaningful for cases when insufficient number of input LR images is available to perform SR with only integer factors, such as 2 or 3. To achieve truly robust methodology applicable in real situations, we adopted the regularized energy minimization approach, which we solve by alternating-minimization scheme. The fundamental improvement on previously proposed SR methods is the notion of estimating PSFs in the HR scale, which indirectly aligns LR images with subpixel accuracy. Using registration parameters inside, the algorithm instead of registering input images gives better results and paves the way for including methods of making registration parameters more accurate during reconstruction of the HR image [35].

FUNDING

Financial support of this research was provided by Czech Ministry of Education under the project 1M0572 (Research Center DAR), the Grant Agency of the Czech Republic under the project 202/05/0242, by the bilateral project 2004CZ0009 and by the projects TEC2004-00834, TEC2005-24739-E and PI040765.

REFERENCES

[1] Zitová, B. and Flusser, J. (2003) Image registration methods: a survey. *Image Vis. Comput.*, **21**, 977–1000.

[2] Harikumar, G. and Bresler, Y. (1999) Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms. *IEEE Trans. Image Process.*, **8**, 202–219.

[3] Giannakis, G. and Heath, R. (2000) Blind identification of multichannel FIR blurs and perfect image restoration. *IEEE Trans. Image Process.*, **9**, 1877–1896.
[4] Pai, H.-T. and Bovik, A. (2001) On eigenstructurebased direct multichannel blind image restoration. *IEEE Trans. Image Process.*, **10**, 1434–1446.
[5] Panci, G., Campisi, P., Colonnese, S. and Scarano, G. (2003) Multichannel blind image deconvolution using the bussgang algorithm: spatial and multiresolution approaches. *IEEE Trans. Image Process.*, **12**, 1324–1337.
[6] Šroubek, F. and Flusser, J. (2003) Multichannel blind iterative image restoration. *IEEE Trans. Image Process.*, **12**, 1094–1106.
[7] Šroubek, F. and Flusser, J. (2005) Multichannel blind deconvolution of spatially misaligned images. *IEEE Trans. Image Process.*, **14**, 874–883.
[8] You, Y.-L. and Kaveh, M. (1999) Blind image restoration by anisotropic regularization. *IEEE Trans. Image Process.*, **8**, 396–407.
[9] Rajagopalan, A. and Chaudhuri, S. (1999) An MRF model-based approach to simultaneous recovery of depth and restoration from defocused images. *IEEE Trans. Pattern Anal. Mach. Intell.*, **21**, 577–589.
[10] Park, S., Park, M. and Kang, M. (2003) Super-resolution image reconstruction: a technical overview. *IEEE Signal Process. Mag.*, **20**, 21–36.
[11] Farsui, S., Robinson, D., Elad, M. and Milanfar, P. (2004) Advances and challenges in super-resolution. *Int. J. Imag. Syst. Technol.*, **14**, 47–57.
[12] Hardie, R., Barnard, K. and Armstrong, E. (1997) Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Trans. Image Process.*, **6**, 1621–1633.
[13] Segall, C., Katsaggelos, A., Molina, R. and Mateos, J. (2004) Bayesian resolution enhancement of compressed video. *IEEE Trans. Image Process.*, **13**, 898–911.
[14] Woods, N., Galatsanos, N. and Katsaggelos, A. (2006) Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images. *IEEE Trans. Image Process.*, **15**, 201–213.
[15] Farsiu, S., Robinson, M., Elad, M. and Milanfar, P. (2004) Fast and robust multiframe super resolution. *IEEE Trans. Image Process.*, **13**, 1327–1344.
[16] Shechtman, E., Caspi, Y. and Irani, M. (2005) Space–time super-resolution. *IEEE Trans. Pattern Anal. Mach. Intell.*, **27**, 531–545.
[17] Humblot, F. and Muhammad-Djafari, A. (2006) Super-resolution using hidden markov model and bayesian detection estimation framework. *EURASIP J. Appl. Signal Process.*, **2006**, 36971.
[18] Ben-Ezra, M. and Nayer, S. (2004) Notion-based motion deblurring. *IEEE Trans. Pattern Anal. Mach. Intell.*, **26**, 689–698.
[19] Nguyen, N., Milanfar, P. and Golub, G. (2001) Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement. *IEEE Trans. Image Process.*, **10**, 1299–1308.
[20] Woods, N., Galatsanos, N. and Katsaggelos, A. (2003) EM-based simultaneous registration, restoration, and interpolation of super-resolved images. *Proc. ICIP*, Barcelona, September, pp. 303–306. IEEE Computer Society.
[21] Rajan, D. and Chaudhuri, S. (2003) Simultaneous estimation of super-resolved scene and depth map from low resolution defocused observations. *IEEE Trans. Pattern Anal. Mach. Intell.*, **25**, 1102–1117.
[22] Šroubek, F. and Flusser, J. (2006) Resolution enhancement via probabilistic deconvolution of multiple degraded images. *Pattern Recognit. Lett.*, **27**, 287–293.
[23] Chen, Y., Luo, Y. and Hu, D. (2005) A general approach to blind image super-resolution using a PDE framework. *Proc. SPIE*, Beijing, July, pp. 1819–1830. SPIE.
[24] Wirawan, D.P. and Maitre, H. (1999) Multi-channel high resolution blind image restoration. *Proc. ICASSP*, Phoenix, AR, March, pp. 3229–3232. IEEE Computer Society.
[25] Yagle, A. (2003) Blind superresolution from under-sampled blurred measurements. *Proc. SPIE*, Bellingham, December, pp. 299–309. SPIE.
[26] Biggs, D., Wang, C.L., Holmes, T. and Khodjakov, A. (2004) Subpixel deconvolution of 3D optical microscope imagery. *Proc. SPIE*, Denver, CO, October, pp. 369–380. SPIE.
[27] Šroubek, F., Cristóbal, G. and Flusser, J. (2007) A unified approach to superresolution and multichannel blind deconvolution. *IEEE Trans. Image Process.*, **16**, 2322–2332.
[28] Tschumperlé, D. (2002) PDE's-based regularization of multivalued images and applications. PhD Thesis, University of Nice-Sophia Antipolis.
[29] Lin, Z. and Shum, H.-Y. (2004) Fundamental limits of reconstruction-based superresolution algorithms under local translation. *IEEE Trans. Pattern Anal. Mach. Intell.*, **26**, 83–97.
[30] Pelletier, S. and Cooperstock, J. (2007) Fast super-resolution for rational magnification factors. *Proc. ICIP*, San Antonio, TX, September, pp. 65–68. IEEE Computer Society.
[31] Šroubek, F., Flusser, J. and Cristóbal, G. (2007) Multiframe Blind Deconvolution Coupled with Frame Registration and Resolution Enhancement. In Campisi, P. and Egiazarian, K. (eds.), *Blind Image Deconvolution: Theory and Applications*. CRC Press, FL.
[32] Aubert, G. and Kornprobst, P. (2002) *Mathematical Problems in Image Processing*. Springer, New York.
[33] Molina, R., Vega, M., Abad, J. and Katsaggelos, A. (2003) Parameter estimation in Bayesian highresolution image reconstruction with multisensors. *IEEE Trans. Image Process.*, **12**, 1655–1667.
[34] Capel, D. (2004) *Image Mosaicing and Super-Resolution*. Springer, New York.
[35] Chung, J., Haber, E. and Nagy, J. (2006) Numerical methods for coupled super-resolution. *Inverse Probl.*, **22**, 1261–1272.

IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 4, APRIL 2012 1687

Robust Multichannel Blind Deconvolution via Fast Alternating Minimization

Filip Šroubek, *Member, IEEE*, and Peyman Milanfar, *Fellow, IEEE*

***Abstract*—Blind deconvolution, which comprises simultaneous blur and image estimations, is a strongly ill-posed problem. It is by now well known that if multiple images of the same scene are acquired, this multichannel (MC) blind deconvolution problem is better posed and allows blur estimation directly from the degraded images. We improve the MC idea by adding robustness to noise and stability in the case of large blurs or if the blur size is vastly overestimated. We formulate blind deconvolution as an ℓ_1-regularized optimization problem and seek a solution by alternately optimizing with respect to the image and with respect to blurs. Each optimization step is converted to a constrained problem by variable splitting and then is addressed with an augmented Lagrangian method, which permits simple and fast implementation in the Fourier domain. The rapid convergence of the proposed method is illustrated on synthetically blurred data. Applicability is also demonstrated on the deconvolution of real photos taken by a digital camera.**

***Index Terms*—Alternating minimization, augmented Lagrangian, blind deconvolution.**

I. Introduction

IMAGE deconvolution is a classical inverse problem in image processing. Deconvolution appears in a wide range of application areas, such as photography, astronomy, medical imaging, and remote sensing, just to name few. Images deteriorate during acquisition as data pass through the sensing, transmission, and recording processes. In general, the observed degradation is a result of two physical phenomena. The first is of random nature and appears in images as noise. The second is deterministic and results in blurring, which is typically modeled by convolution with some blur kernel called the point spread function (PSF). Degradation caused by convolution can thus appear in any application where image acquisition takes place. The common sources of blurring are lens imperfections, air turbulence, or camera-scene motion. Solving the deconvolution problem in a reliable way has been of prime interest in the field of image processing for several decades and has produced an enormous number of publications.

Let us first consider problems with just one degraded image, i.e., single-channel deconvolution. The simplest case is if the blur kernel is known (i.e., a classical deconvolution problem). However, even here, estimating an unknown image is ill-posed due to the ill-conditioned nature of the convolution operators. This inverse problem can only be solved by adopting some sort of regularization (in stochastic terms, regularization corresponds to priors). Another option is to use techniques such as coded aperture [1], but this requires a modification of camera hardware, which we do not consider here. A popular recent approach is to let the unknown image be represented as a linear combination of few elements of some frame (usually an overcomplete dictionary) and to force this sparse representation by using the ℓ_p norm $(0 \leq p \leq 1)$. Either we can search for the solution in the transform domain (coefficients of the frame elements), which is referred to as the synthesis approach, or regularize directly the unknown image, which is called the analysis approach. Analysis versus synthesis approach has been studied earlier [2], [3]. If the frame is an orthonormal basis, both approaches are equivalent. More interesting however is the case of redundant representation (e.g., an undecimated wavelet transform), when the two approaches differ. Conclusions presented in [3] suggest that, for deconvolution problems, the analysis approach is preferable because sparsity should be enforced only on a part of the redundant representation (e.g., high-pass bands), and this can be easily implemented only in the analysis approach. Very recently, it has been shown that the analysis approach is solved efficiently using variable splitting and by applying a Bregman iterative method [4] or an augmented Lagrangian method (ALM) [5] (both methods lead to the same algorithm).

If the blur kernel is unknown, we face single-channel blind deconvolution, which is clearly even more complicated than the classical deconvolution problem. This inverse problem is underdetermined as we have more unknowns (image and blur) than equations. For a long time, the problem seemed too difficult to solve for general blur kernels. Past algorithms usually worked only for special cases, such as astronomical images with a uniform (black) background, and their performance depended on initial estimates of PSFs. To name a few papers from this category, consider [6]–[8] and survey [9]. Probably, the first attempt toward a more general blur estimation came from Fergus *et al.* [10], who proposed a variational Bayesian method [11] with natural image statistics. This triggered a furious activity in the computer vision community, and soon, several conference papers appeared on the same topic [12]–[17]. Levin *et al.* [15]

Manuscript received February 28, 2011; revised July 28, 2011 and October 24, 2011; accepted November 01, 2011. Date of publication November 09, 2011; date of current version March 21, 2012. This work was supported in part by the US Air Force under Grant FA9550-07-1-0365, by the National Science Foundation under Grant CCF-1016018, by the Czech Ministry of Education under Project 1M0572 (Research Center DAR), and by the Grant Agency of the Czech Republic under Project P103/11/1552. F. Sroubek performed the work while at the UCSC supported by the Fulbright Visit Scholar Fellowship. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ramin Samadani.

F. Šroubek is with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, 182 08 Prague, Czech Republic (e-mail: sroubekf@utia.cz).

P. Milanfar is with the Department of Electrical Engineering, University of California at Santa Cruz, Santa Cruz, CA 95064 USA (e-mail: milanfar@ee.ucsc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2011.2175740

pointed out that the joint posterior probability of the image–blur pair favors a trivial solution of the blur being a delta function and that marginalizing the posterior (integrating out the image variable) is more appropriate. However, a closed-form solution seldom exists, and a complicated approximation of the posterior is necessary, which leads to cumbersome methods that can hardly handle large blurs. In order to avoid these drawbacks, recent methods still try to minimize directly the joint posterior probability since it can be done in an efficient way but perform all sorts of tricks to avoid the trivial solution. Jia [12] uses an alpha matte to extract a transparency map and estimates the blur kernel on the map. Joshi *et al.* [13] predicts sharp edges using edge profiles and estimates the blur kernel from the predicted edges. Cho *et al.* [16] applies a shock filter and gradient thresholding to restore only strong edges and estimates the blur kernel from the truncated gradient image. A similar idea further improved by a kernel refinement step has been proposed recently by Xu *et al.* [17]. In general, the single-channel blind deconvolution methods get trapped in local minima and must estimate blurs using a multiscale approach. They have many parameters that influence the result considerably and are hard to tune. The common trick for the methods to work is to have means to predict strong edges. However, if the blurry image does not have salient edges or it is corrupted by noise, all the single-channel deconvolution methods usually fail.

The ill-posed nature of blind deconvolution can be remedied to a great extent by considering multiple images. In this case, the problem is referred to as multichannel (MC) blind deconvolution and will be the subject of our investigation. Acquired images must capture the same scene and differ only in the blur kernel. This may not seem to be easy to achieve in practice. However, the opposite is true. There are many situations where multiple images blurred in a slightly different way can be obtained. For example, if atmospheric turbulence causes blurring, we can capture several images (or video frames) in a row, and due to the random nature of turbulence, each image is almost surely blurred in a different way. If camera shake causes blurring, continuous shooting (or video capture) with the camera provides several images that are blurred in a different way since our hand moves randomly. MC deconvolution requires that the input images are properly registered, which is one drawback compared with the single-channel case. If the images are acquired as described above, misregistration is only minor, and even simple registration methods will provide accurate and stable results (see, e.g., [18]) for a survey of registration methods. We will thus assume that the input images are registered up to some global translation. A simple registration method for affine transforms is used in our experiments, as sketched in Section VI. More problematic is the occurrence of space-variant blur, which often arises in practice, such as rotating camera or profound depth of scene. We note that the method proposed here assumes a space-invariant case, but by applying the method locally, we can, in theory, deal with space-variant cases as well. We refer the interested reader to [19] and references therein for space-variant deconvolution.

One of the earliest intrinsic MC blind deconvolution methods [20] was designed particularly for images blurred by atmospheric turbulence. Harikumar *et al.* [21] proposed an indirect algorithm, which first estimates blur kernels and then recovers the original image by a standard nonblind method. The blur kernels are equal to the minimum eigenvector of a special matrix constructed from the blurred input images. Necessary assumptions for perfect recovery of the blurs are noise-free environment and channel coprimeness, i.e., a scalar constant is the only common factor of the blurs. Giannakis *et al.* [22] developed another indirect algorithm based on Bezout's identity of coprime polynomials, which finds restoration filters. In addition, by convolving the filters with the input images, it recovers the original image. Both algorithms are vulnerable to noise and, even for a moderate noise-level restoration, may break down. Pai *et al.* [23] suggested two MC restoration algorithms that, contrary to the previous two indirect algorithms, estimate directly the original image from the null space or from the range of a special matrix. Another direct method based on the greatest common divisor was proposed in [24]. In noisy cases, the direct algorithms are more stable than the indirect ones. Approaches based on the autoregressive moving average model are given in [25]. MC blind deconvolution using a Bussgang algorithm was proposed in [26], which performs well on spatially uncorrelated data, such as binary text images and spiky images. Sroubek *et al.* [27] proposed a method that reformulates Harikumar's idea in [21] as a MC regularization term and simultaneously minimizes an energy function with respect to the image and blur kernels. This allows us to handle inexact PSF sizes and to compensate for small misalignment in input images, which made MC deconvolution more practical. However, small PSFs (less than 15×15) and images of size couple of hundreds of pixels were only considered. It is mainly because of the inefficiency of the applied numerical algorithm that the method is not converging for larger blurs and images.

Here, we propose an MC blind deconvolution method that can handle very large blurs (e.g., 50×50) and images of several megapixels with even better accuracy and speed. The method is based on the same idea as in [27], and it is formulated as a constrained optimization problem. For image regularization, we use total variation (TV) [28], and for blur regularization, we use the MC constraint proposed in [21]. We show that the original MC constraint is not robust to noise and propose a simple remedy, which requires a negligible extra computation but achieves much better stability with respect to noise. Since the optimization problem mixes the ℓ_2 and ℓ_1 norms, we use the state-of-the-art numerical method of augmented Lagrangian [5] to solve the blind deconvolution problem and achieve very fast convergence. As it will be cleared later, positivity of blur kernels is an important constraint that must be included in the optimization problem. We show that positivity can be incorporated in augmented Lagrangian effortlessly without affecting the convergence properties.

This paper is organized as follows. Section II defines notation and presents the basic alternating minimization approach to blind deconvolution. Image regularization in the form of isotropic TV is given in Section III. Section IV discusses the problem of blur estimation in the MC scenario and influence of noise and blur size and proposes a novel blur kernel constraint with sparsity and positivity regularization. A description of the proposed algorithm is given in Section V, together with

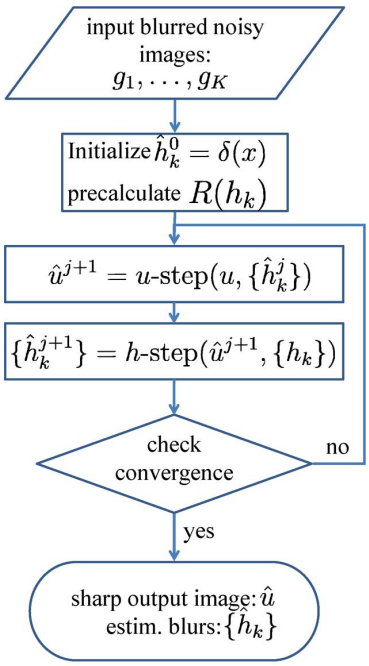


Fig. 1. Flowchart of the alternating minimization algorithm.

implementation details. The experimental section, Section VI, empirically validates the proposed method, and Section VII concludes this paper.

II. MC Blind Deconvolution Basics

We formulate the problem in the discrete domain and use frequently vector–matrix notation throughout the text. Images and PSFs are denoted by small italic letters and their corresponding vectorial representations (lexicographically ordered pixels) are denoted by small bold letters. The MC blind deconvolution problem assumes that we have $K > 1$ input images $\{g_1, \dots, g_K\}$ ($g_k : \mathbb{N}^2 \to \mathbb{R}$) that are related to an unknown image $u : \mathbb{N}^2 \to \mathbb{R}$ according to model

$$g_k(i) = (h_k * u)(i) + n_k(i), \qquad 1 \le k \le K \tag{1}$$

where h_k denotes an unknown blur (kernel or PSF = PSF) and n_k is the additive noise in the kth observation. Operator $*$ stands for convolution, and $i \in \mathbb{N}^2$. When no ambiguity arises, we drop multiindex i from the notation. In the vector–matrix notation, (1) becomes

$$\mathbf{g}_k = \mathbf{H}_k\mathbf{u} + \mathbf{n}_k = \mathbf{U}\mathbf{h}_k + \mathbf{n}_k \tag{2}$$

where matrices $\mathbf{H}_k$ and $\mathbf{U}$ perform convolution with h_k and u, respectively. To denote the ith element in the vector notation, we write $[\cdot]_i$, e.g., $u(i) = [\mathbf{u}]_i$. The size of images and blurs (matrices and vectors) will be discussed later when necessary.

In the case of multiple acquisitions, we cannot expect that input images are perfectly spatially aligned. One can model such misregistation by the geometric transformation $\mathbf{W}_k$ that precede blurring $\mathbf{H}_k$, i.e., $\mathbf{H}_k\mathbf{W}_k\mathbf{u}$. If $\mathbf{W}_k$ is invertible, then $\mathbf{H}_k\mathbf{W}_k = \mathbf{W}_k\mathbf{W}_k^{-1}\mathbf{H}_k\mathbf{W}_k = \mathbf{W}_k\widetilde{\mathbf{H}}_k$, where $\widetilde{\mathbf{H}}_k = \mathbf{W}_k^{-1}\mathbf{H}_k\mathbf{W}_k$. If $\mathbf{H}_k$ is a standard convolution with some PSF h_k and $\mathbf{W}_k$ is a linear geometric transformation, then the new blurring operator $\widetilde{\mathbf{H}}_k$ remains a standard convolution but with h_k warped according to $\mathbf{W}_k$. Therefore, for linear geometric transformations (such as affine), the order of geometric transformation and blurring can be interchanged. We thus assume that input images g_k can be accurately registered by linear transformations, and a registration step preceding blind deconvolution removes such geometric transformations.

It is well known that the problem of estimating u from g_k is ill-posed; thus, this inverse problem can only be solved satisfactorily by adopting some sort of regularization. Formally, this leads to the following optimization problem:

$$\min_{u,\{h_k\}} F(u, \{h_k\}) + Q(u) + R(\{h_k\}) \tag{3}$$

where F is the data fidelity term and Q and R are regularizers of the image and blurs, respectively. The formation model (1) determines the data term leading to a standard formulation $F(u, \{h_k\}) = (\gamma/2)\sum_{k=1}^{K} \|u * h_k - g_k\|^2$, where γ is inversely proportional to the variance of noise n_k and $\|\cdot\|$ denotes the ℓ_2 norm. For simplicity, we assume the same noise variance in all frames; therefore, single parameter γ suffices. The standard approach to solve (3) is called alternating minimization and will be adopted here as well. We split the problem into two subproblems, i.e.,

$$\text{“}u\text{-step"}: \quad \min_{u} F(u, \{h_k\}) + Q(u) \tag{4}$$

$$\text{“}h\text{-step"}: \quad \min_{\{h_k\}} F(u, \{h_k\}) + R(\{h_k\}) \tag{5}$$

and alternate between them (see the algorithm flowchart in Fig. 1). Convergence to the global minimum is theoretically not guaranteed since the unknown variables are coupled in the data term F. However, we show that each subproblem separately converges to its global minimum and that it can be solved efficiently by the ALM. This implies that, in general, the global minimum of (3) is attainable after few alternations between the subproblems. The next two sections describe in detail the image Q and blur R regularization terms.

III. Image Regularization

A popular recent approach to image regularization is to assume that the unknown image u is represented as a linear combination of few elements of some frame (usually an overcomplete dictionary) and to force this sparse representation by using the ℓ_1 norm (or ℓ_0). Arguably, the best known and most commonly used image regularizer, which belongs to the category of sparse priors, is the TV norm [28].

The isotropic TV model is the ℓ_1 norm of image-gradient magnitude values and takes the following form:

$$Q(u) = \sum_i \phi(\nabla u(i)) = \sum_i \sqrt{(\nabla_x u(i))^2 + (\nabla_y u(i))^2} \tag{6}$$

where $\phi(x) = \|x\|$. The TV regularizer thus forces the solution to have sparse image gradient. Depending on the type of data, one can have sparsity in different domains. This modification is however easy to achieve. All we have to do is to replace

derivatives with a transformation (e.g., a waveletlike multiscale transform), which gives sparse representation of our data.

Using the vector–matrix notation, the isotropic TV (6) can be written as

$$Q(\mathbf{u}) = \Phi(\mathbf{D}_x\mathbf{u}, \mathbf{D}_y\mathbf{u}) = \sum_i \sqrt{[\mathbf{D}_x\mathbf{u}]_i^2 + [\mathbf{D}_y\mathbf{u}]_i^2} \qquad (7)$$

where $\mathbf{D}_x$ and $\mathbf{D}_y$ are matrices performing derivatives with respect to x and y, respectively.

IV. Blur Estimation and Regularization

We first review an MC PSF estimation method proposed in [21], [22], which was later used in MC blind deconvolution as the PSF regularizer [27]. We demonstrate that the method is not robust to noise and show a novel improvement in this aspect. To keep the notation simple, let us assume 1-D data and the two-channel convolution model (1) ($K = 2$). The following discussion can be easily extended to 2-D data and any $K > 2$. The sizes of 1-D data g_k, u, and h_k is M, N, and L, respectively, with $N \gg L$. Noise n_k is of the same size as g_k. Kernels h_k can be of different sizes, but we can always pad the smaller ones with zeros to have the size of the largest one and therefore L refers to the size of the largest PSF. To deal correctly with convolution at image boundaries, we work with convolution that returns a "valid" part of the support and thus $M = N - L + 1$. The matrices $\mathbf{H}_k$ and $\mathbf{U}$ in the vector–matrix formation model (2) are thus of size $M \times N$ and $M \times L$, respectively.

Let $\widehat{h}_k$ be an estimate of h_k. In general, the original PSF size L is not known; therefore, $\widehat{h}_k$ can be of different size, which is denoted here as $\widehat{L}$. Let us study three cases that will be used in the following discussion: (a1) noiseless case ($n_k = 0$); (a2) PSF size is exactly known ($\widehat{L} = L$); and (a3) original PSFs are weakly coprime and images g_k are persistently exciting for size L. A set of kernels $\{h_k\}$ is called weakly coprime [22]; if there exists kernel s and set $\{\widetilde{h}_k\}$ so that, $\forall k$, $h_k = s * \widetilde{h}_k$, then s is a scalar. In other words, if the kernels are decomposable, they must not have a common kernel. An image z of size M is called persistently exciting [21] for size L if its "valid" convolution matrix $\mathbf{Z}$ of size $(M - L + 1) \times L$ has full column rank. Note that such an image will be also persistently exciting for any size smaller than L.

A. Noiseless Case

We first consider a situation, when all three assumptions (a1), a(2), and (a3) hold. If $\widehat{h}_k = h_k$, then

$$g_2 * \widehat{h}_1 - g_1 * \widehat{h}_2 = h_2 * u * \widehat{h}_1 - h_1 * u * \widehat{h}_2 = 0 \qquad (8)$$

where we used the commutative property of convolution. Rewriting the above relation in the vector–matrix notation, we get

$$[\mathbf{G}_2, -\mathbf{G}_1]\mathbf{h} = \mathbf{0} \qquad (9)$$

where $\mathbf{h} = [\mathbf{h}_1^T, \mathbf{h}_2^T]^T$. Matrices $\mathbf{G}_1$ and $\mathbf{G}_2$ denote "valid" convolution with g_1 and g_2, respectively, and they are of size $(M - L + 1) \times L$. Note that, in the case of $K > 2$, it is sufficient to consider all unordered pairs of images, which is equal to the combinatorial number $\binom{K}{2}$. Thus, for example, for $K = 3$, the number of image pairs is $\binom{3}{2} = 3$; (9) becomes

$$\begin{bmatrix} \mathbf{G}_2 & -\mathbf{G}_1 & \mathbf{0} \\ \mathbf{G}_3 & \mathbf{0} & -\mathbf{G}_1 \\ \mathbf{0} & \mathbf{G}_3 & -\mathbf{G}_2 \end{bmatrix} \mathbf{h} = \mathbf{0}.$$

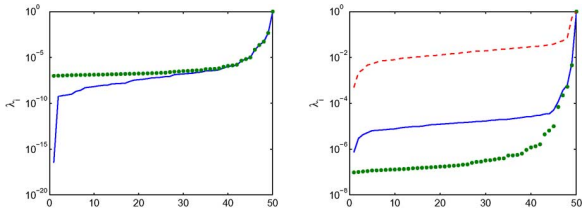


Fig. 2. Spectra of kernel regularization matrices $\mathbf{R}$ in (10), $\mathbf{R}_\Sigma$ in (15), and $\mathbf{R}_\Delta$ in (17). (a) $\mathbf{R}$ in the (solid line) noiseless and (dotted line) noisy case and (b) (solid line) $\mathbf{R}_\Sigma$ and (dashed line) $\mathbf{R}_\Delta$ in the noisy case.

Let us continue with $K = 2$ and define a symmetric positive semidefinite $2L \times 2L$ matrix, i.e.,

$$\mathbf{R} = [\mathbf{G}_2, -\mathbf{G}_1]^T[\mathbf{G}_2, -\mathbf{G}_1]. \qquad (10)$$

The computational complexity of constructing this matrix is discussed in Section V-C. It follows from (9) that the correct estimates of h_k lie in the null space of $\mathbf{R}$. We refer to eigenvalues of $\mathbf{R}$ as λ_i ($\lambda_1 < \lambda_2 < \ldots < \lambda_{2L}$) and the corresponding eigenvectors as $\mathbf{v}_i$. Since (a2) and (a3) hold, $\mathbf{R}$ has exactly one zero eigenvalue λ_1, and eigenvector $\mathbf{v}_1$ is equal to the correct PSFs h_k stacked in one vector multiplied by a scalar. Note that $\mathbf{R}$ is constructed solely from the input image $\mathbf{g}_k$ values, and it can be thus used for the PSF estimation. An example of the $\mathbf{R}$ spectrum (plot of λ_i values) is in Fig. 2(a) (solid line). Matrix $\mathbf{R}$ was constructed from images blurred by two 5×5 PSFs in Fig. 3(a). Notice the prominent kink at the first eigenvalue λ_1. The corresponding eigenvector $\mathbf{v}_1$ represents exactly the original PSFs. This fact is also illustrated in Fig. 4(a), which plots the representation $\{\alpha_i\}$ of $\mathbf{h}$ in basis $\{\mathbf{v}_i\}$, i.e., $\mathbf{h} = \sum_{i=1}^{2L} \alpha_i \mathbf{v}_i$. One can use $\mathbf{R}$ to build the following quadratic form:

$$R(\mathbf{h}) = \mathbf{h}^T\mathbf{R}\mathbf{h} \qquad (11)$$

and rewrite the $\mathbf{v}_1$ eigenvector estimation as a constrained optimization problem

$$\min_{\mathbf{h}} R(\mathbf{h}) \quad \text{s.t.} \quad \forall k \sum_i h_k(i) = 1. \qquad (12)$$

As proposed in [27], it is better to use the quadratic term R as a PSF regularization term in the blind MC deconvolution problem (3). Because of the favorable spectrum of $\mathbf{R}$, the convergence of such algorithms is very fast.

B. Noisy Case

Let us see what happens if we remove (a1) and allow noise to enter the formation model (1). We assume uncorrelated normally distributed noise $n_k \sim N(0, \sigma^2)$. It follows from (2) that the convolution matrices $\mathbf{G}_k$ in (9) take the form

$$\mathbf{G}_k = \mathbf{H}_k\mathbf{U} + \mathbf{N}_k \qquad (13)$$

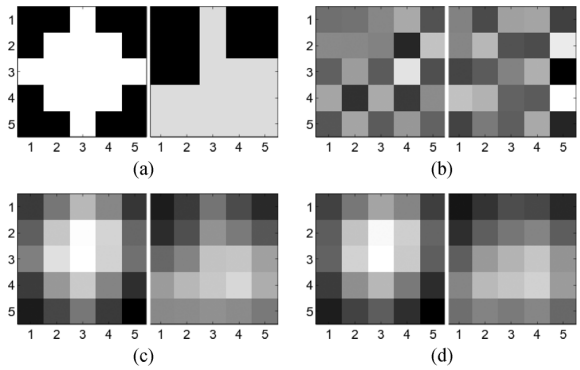


Fig. 3. PSFs and their estimates (first eigenvectors) in the noisy case. (a) two original PSFs of size 5×5. (b) Estimation using $\mathbf{R}$. (c) Estimation using $\mathbf{R}_\Sigma$. (d) Estimation using $\mathbf{R}_\Delta$.

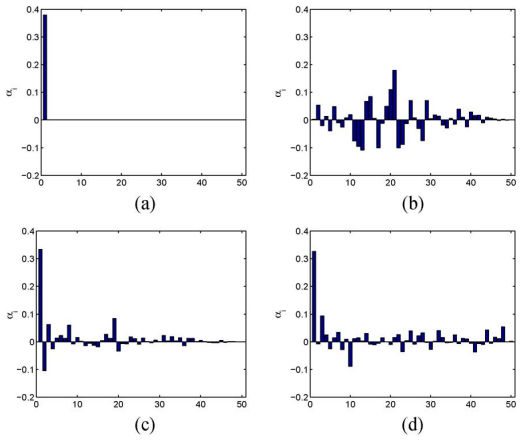


Fig. 4. Representation of PSFs in the eigenvector basis of regularization matrices. (a) $\mathbf{R}$ in the noiseless case. (b) $\mathbf{R}$ in the noisy case. (c) $\mathbf{R}_\Sigma$ in the noisy case. (d) $\mathbf{R}_\Delta$ in the noisy case.

where, this time, $\mathbf{H}_k$ is of size $(M-L)\times M$ and $\mathbf{N}_k$ is a noise convolution matrix constructed in the same way as $\mathbf{G}_k$ but using elements of n_k instead of g_k. Substituting for $\mathbf{G}_k$ in (9), we get

$$[\mathbf{G}_2, -\mathbf{G}_1]\mathbf{h} \sim N(\mathbf{0}, \Sigma) \tag{14}$$

where $\Sigma = \mathbf{cov}(\mathbf{H_2Uh}_1 + \mathbf{N}_2\mathbf{h}_1 - \mathbf{H_1Uh}_2 - \mathbf{N}_1\mathbf{h}_2) = \mathbf{cov}(\mathbf{N}_2\mathbf{h}_1 - \mathbf{N}_1\mathbf{h}_2) = \mathbf{cov}(\mathbf{H}_1\mathbf{n}_2 - \mathbf{H}_2\mathbf{n}_1) = \mathbf{H}_1\mathbf{cov}(\mathbf{n}_2)\mathbf{H}_1^T + \mathbf{H}_2\mathbf{cov}(\mathbf{n}_1)\mathbf{H}_2^T = \sigma^2(\mathbf{H}_1\mathbf{H}_1^T + \mathbf{H}_2\mathbf{H}_2^T)$, since $h_2 * u * h_1 = h_1 * u * h_2 \Longleftrightarrow \mathbf{H_2Uh}_1 = \mathbf{H_1Uh}_2$, which follows from (8). Because of noise, we cannot expect that the smallest eigenvalue of $\mathbf{R}$ will no longer be zero. Indeed, the kink visible in the noiseless case is completely leveled out in the noisy case. Fig. 2(a) (dotted line) shows the spectrum of $\mathbf{R}$ for the input data used before but corrupted by noise with SNR $= 40$ dB, which is a relatively small level of noise hardly detectable by human eyes. Eigenvector $\mathbf{v}_1$ is no longer informative and represents an erroneous solution, as shown in Fig. 3(b). The correct solution is a linear combination of all eigenvectors with the weights almost randomly distributed, as shown in Fig. 4(b).

The maximum-likelihood estimation of kernels must include the covariance matrix Σ in $\mathbf{R}$, i.e.,

$$\mathbf{R}_\Sigma = [\mathbf{G}_2, -\mathbf{G}_1]^T\Sigma^{-1}[\mathbf{G}_2, -\mathbf{G}_1]. \tag{15}$$

The spectrum of $\mathbf{R}_\Sigma$ retains the kink at the first smallest eigenvalue λ_1, as Fig. 2(b) (solid line) shows. For comparison, we show the original spectrum of $\mathbf{R}$ in (10), as a dotted line [also in Fig. 2(a)]. The eigenvector $\mathbf{v}_1$ of $\mathbf{R}_\Sigma$ captures the original PSFs, as shown in Fig. 3(c). Encoding of the true kernels $\mathbf{h}$ in the basis $\{\mathbf{v}_i\}$ is relatively sparse and cluster around the smallest eigenvalues [see Fig. 4(c)]. The same behavior persists even for much higher noise levels (around 10 dB). The construction of $\mathbf{R}_\Sigma$ has one severe drawback: We must know the correct kernels h_k *a priori* in order to build Σ. Since our aim is to estimate PSFs, this seem to be contradictory. One can apply an iterative procedure and update Σ with every new estimate of h_k, as proposed in [21]. Unfortunately, this framework is not guaranteed to converge. In addition, inversion of Σ can be very costly, which makes the whole calculation of $\mathbf{R}_\Sigma$ for large kernels (large L) impossible.

We propose to filter the blurred input images g_k in such a way so that $\mathbf{R}$ without Σ in (10) will be closed to $\mathbf{R}_\Sigma$ in (15). If we filter the input images with some kernel p, then

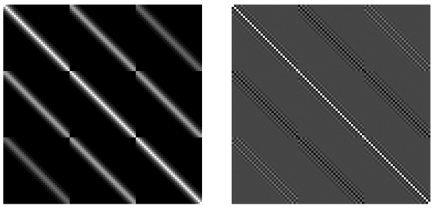
$$[\mathbf{PG}_2, -\mathbf{PG}_1]\mathbf{h} \sim N(\mathbf{0}, \Sigma_\mathbf{P}) \tag{16}$$

where $\mathbf{P}$ performs convolution with p and the covariance matrix is $\Sigma_\mathbf{P} = \sigma^2(\mathbf{H}_1\mathbf{PP}^T\mathbf{H}_1^T + \mathbf{H}_2\mathbf{PP}^T\mathbf{H}_2^T) = \sigma^2\mathbf{P}(\mathbf{H}_1\mathbf{H}_1^T + \mathbf{H}_2\mathbf{H}_2^T)\mathbf{P}^T$. The best choice of the filter p is such that $\Sigma_\mathbf{P} = \sigma^2\mathbf{I}$; since then, the covariance matrix can be neglected. However, this would again require *a priori* knowledge of unknown kernels h_k since p depends on h_k. Achieving a diagonal correlation matrix means that we want to spatially decorrelate the blur kernels. In the absence of any prior knowledge of the blurs, we wish to employ a decorrelation method that is sufficiently general. As such, given the well-accepted assumption of sparsity on high-frequency spatial structures, the natural choice is to apply a Laplacian operator. The justification is therefore empirical but quite reasonable. In Fig. 5(a), we show a small part of the covariance matrix Σ for our example with two blurs and, in Fig. 5(b), the covariance matrix $\Sigma_\mathbf{P}$ with $\mathbf{P}$ being the Laplacian. The covariance matrix of the filtered images is not diagonal but close to diagonal. The Laplacian produces images, which are relatively sparse and therefore spatially uncorrelated to a great extent. The same holds for PSFs that blur the images, which accounts for the close-to-diagonal covariance matrix.

Let Δ denote a matrix that performs convolution with the discrete Laplacian kernel l (in 1-D $l = [1, -2, 1]$). The proposed modification of the matrix $\mathbf{R}$ is

$$\mathbf{R}_\Delta = [\Delta\mathbf{G}_2, -\Delta\mathbf{G}_1]^T[\Delta\mathbf{G}_2, -\Delta\mathbf{G}_1]. \tag{17}$$

Matrix $\mathbf{R}_\Delta$ depends only on the input images g_k, and the construction is trivial. The spectrum of this matrix retains the kink [see dashed line in Fig. 2(b)] and relatively sparse representation of $\mathbf{h}$, as shown in Fig. 4(d). Eigenvector $\mathbf{v}_1$ estimates h_k in a similar way as ideal $\mathbf{R}_\Sigma$ [see Fig. 3(d)].



(a) (b)

Fig. 5. Covariance matrices. (a) Calculated from the original PSFs. (b) Calculated from the Laplacian of PSFs.

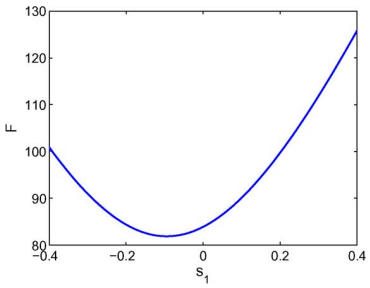


Fig. 6. Data term $F(\widehat{u}, \{s * h_k\})$ as a function of the first elements s_1 of the 2×1 spurious vector $s = [s_1, s_2]$, where $s_2 = 1 - s_1$. The minimum is not reached for $s = [0, 1]$ (delta function) but for s with a small negative value.

C. Overestimated Kernel Size

It is unrealistic to assume that the kernel size L is exactly known in practice. Let us thus consider the case when both (a1) and (a3) hold, but (a2) is violated with the kernel size being overestimated, i.e., $\widehat{L} > L$. We can readily see that if $\widehat{h}_k = s * h_k$, where s is an arbitrary spurious kernel of size $S = \widehat{L} - L + 1$, the MC constraint (8) still holds

$$g_2 * \widehat{h}_1 - g_1 * \widehat{h}_2 = h_2 * u * s * h_1 - h_1 * u * s * h_2 = 0. \quad (18)$$

In the language of matrix eigenvalues and eigenvectors, this fact translates as follows. Matrix $\mathbf{R}$ defined in (10) is of size $\widehat{L} \times \widehat{L}$. The correct kernels lie again in the null space of $\mathbf{R}$, but this time, the matrix nullity is of the size of the spurious kernel, i.e., nullity$(\mathbf{R}) = S$. The regularization term (11) built from $\mathbf{R}$ becomes less restrictive (more "flat") because of the increased nullity. Therefore, convergence of any minimization algorithm, which estimates PSFs using the proposed regularizer R, is seriously hindered in the case of overestimated kernel size. Note that if the kernel size $\widehat{L}$ is underestimated, (18) does not hold, and we cannot estimated the kernels at all. We will not consider the underestimated case and, instead, focus on improving the stability of the overestimated case.

One can be tempted to assume that the unconstrained optimization problem, as defined in (5), would eliminate the ambiguity inherent in $R(\mathbf{h})$. Using the vector–matrix notation, this problem rewrites as

$$\min_{\mathbf{h}} \frac{\gamma}{2} \sum_{k=1}^{2} \|\widehat{\mathbf{U}}\mathbf{h}_k - \mathbf{g}_k\|^2 + R(\mathbf{h}) \quad (19)$$

where $\widehat{\mathbf{U}}$ is the $M \times \widehat{L}$ convolution matrix with the estimate $\widehat{u}$ of the original image u. If estimate $\widehat{u} = u$, the above optimization problem is well posed, and in fact, we do not need regularizer R at all. However, this scenario is unrealistic since we do not know the original image. Alternating minimization often starts with $\widehat{u}$ equal to a so-called average image, i.e., $\widehat{u} = (1/K) \sum_k g_k$. To illustrate the behavior of the data term $F(\widehat{u}, \{\widehat{h}_k\})$ with respect to the spurious kernel s, we conducted the following experiment. We generated two blurry signals g_1 and g_2 using some random positive PSFs h_1 and h_2 of size L. We set $\widehat{L} = L + 1$; therefore, the spurious kernel s is of size 2, and $s = [s_1, s_2]$. Let us consider kernels of form $\widehat{h}_k = s * h_k$ that preserve energy $\sum_i \widehat{h}_k(i) = 1$, then $R(\widehat{\mathbf{h}}) = 0$ for any s and $s_2 = 1 - s_1$. The data term $F(\widehat{u}, \widehat{h}_k)$ with $\widehat{u}$ being the average image is a function of s_1, and we plot its values for different s_1 in Fig. 6. The minimum was reached for a negative value of s_1, and the same behavior was observed for any pair of blurs h_1 and h_2. The data term is thus biased toward kernels with small negative values, and the unconstrained optimization problem (19) is inappropriate if the kernel size is overestimated. An intuitive explanation is the following. Since we use the average image, the value of F would reach its minimum for some $\widehat{h}_k$ close to delta functions. Such a solution is however heavily penalized by R, which allows only PSFs of form $s * h_k$. In order to get closer to the delta-function solution, s must act as an inverse filter to all positive h_k, and this means that it must perform differentiation; hence, negative values in s are inevitable.

Forcing positivity on kernels is the remedy to the above problem. Clearly, this approach is possible only for positive kernels. We encounter positive-only kernels in many deconvolution problems, and making this assumption is thus not very restrictive. With the positivity constraint, the above problem can be solved by means of quadratic programming. Here, we show a different approach, which will allow us an elegant integration in the ALM and much faster implementation than quadratic programming. We have empirically observed that forcing sparsity on h_k further boosts convergence. In order to guarantee both positivity and sparsity, we propose to use a new kernel regularizer, i.e.,

$$R(\mathbf{h}) = \frac{\delta}{2}\mathbf{h}^T\mathbf{R}_{\triangle}\mathbf{h} + \Psi(\mathbf{h}) \quad (20)$$

where

$$\Psi(\mathbf{h}) = \sum_{k=1}^{K} \sum_{i=1}^{\widehat{L}} \psi\left(h_k(i)\right), \quad \psi(t) = \begin{cases} t, & \text{if } t \geq 0 \\ +\infty, & \text{otherwise} \end{cases} \quad (21)$$

and δ is the weight that controls the influence of the MC constraint $\mathbf{R}_{\triangle}$. The definition of ψ ensures positivity by absolutely penalizing negative values and forces sparsity by calculating the ℓ_1 norm of positive kernels.

Note that it is not necessary to explicitly include the constraint $\forall k \sum_i h_k(i) = 1$ as in (12), which preserves the average gray value in images. This constraint is automatically enforced by the fidelity term $\sum_{k=1}^{2} \|u * h_k - g_k\|$ in (19). If the mean value of the estimated image u is equal to the mean value of g_k, then by solving (19) (h-step), we always preserve $\sum_i h_k(i) = 1$. The u-step in (4) does not change the mean value of u either because the fidelity term is present there as well. Therefore, the

condition is not modified in alternating minimization, and we only have to guarantee that initial PSFs follow the constraint.

D. Kernel Coprimeness

Let us consider the assumption (a3) of persistently exciting images and weakly coprime kernels. The condition of persistently exciting image is a very mild one. Usually $M \gg L$, convolution matrices $\mathbf{G}_k$ have many more rows than columns, and the probability that the matrices will not have a full column rank is thus very small. We do not consider here degenerate cases, such as perfectly uniform or periodic images, that may not be persistently exciting.

The condition of weakly coprime kernels may seem to be more problematic. In the 1-D case (signals), any kernel of length L can be decomposed (factorized) into $L-1$ kernels (root factors) of size 2, which is the direct consequence of the fundamental theorem of algebra[1] (see, e.g., [29]). It is therefore likely that there might exist a factor common to all kernels h_k. In the 2-D case (images), no such factorization in general exists and, as also discussed in [21], coprimeness holds deterministically for most of the 2-D cases of practical interest.

If the common factor exists despite its low probability, kernel estimation still partially works. We are able to recover kernels without their common factor, and the common factor remains as a blur in the estimated image.

V. Optimization Algorithm

Alternating minimization, which solves the MC blind deconvolution problem (3), consists of two subproblems: minimization with respect to the image (u-step) and the minimization with respect to the blurs (h-step). Both subproblems share some similarities because both the image (7) and the blur regularizer (20) are not smooth and introduce nonlinearity in the problem. Direct minimization in each step would be thus a slow process. A simple procedure that solves such problems is called variable splitting, which decouples the ℓ_2 and ℓ_1 portions of the problem (3) by introducing auxiliary variable and converting each subproblem to two simpler minimization steps. We then apply the ALM, which is equivalent to the split Bregman iterative method [4], to solve the subproblems. Our derivation follows the work presented in [5] and partially in [4]. Unique aspects of our algorithm will be emphasized. From now on, we will exclusively use the vector–matrix notation and stack all observations into one system by using the compact notations $\mathbf{g} = [\mathbf{g}_1^T, \ldots, \mathbf{g}_K^T]^T$, $\mathbf{h} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_K^T]^T$, $\mathbf{H} = [\mathbf{H}_1^T, \ldots, \mathbf{H}_K^T]^T$, and the convolution matrix $\mathbf{U}$ will now denote a block diagonal matrix with K blocks, where each block is the original $\mathbf{U}$ from (2).

A. U-Step

Using the TV regularizer (7), minimization with respect to the image (4) writes as

$$\min_{\mathbf{u}} \frac{\gamma}{2}\|\mathbf{H}\mathbf{u}-\mathbf{g}\|^2 + \Phi(\mathbf{D}_x\mathbf{u}, \mathbf{D}_y\mathbf{u}). \tag{22}$$

[1]However, some of the factors may contain complex values.

Applying variable splitting, we replace $\mathbf{D}_x\mathbf{u}$ by $\mathbf{v}_x$ and $\mathbf{D}_y\mathbf{u}$ by $\mathbf{v}_y$. This yields a constrained problem

$$\min_{\mathbf{u},\mathbf{v}_x,\mathbf{v}_y} \frac{\gamma}{2}\|\mathbf{H}\mathbf{u}-\mathbf{g}\|^2 + \Phi(\mathbf{v}_x, \mathbf{v}_y) \quad \text{s.t.} \quad \mathbf{v}_x = \mathbf{D}_x\mathbf{u}, \mathbf{v}_y = \mathbf{D}_y\mathbf{u} \tag{23}$$

which is equivalent to (22). The ALM (or split-Bregman iteration) tackles the constrained problem (23) by considering the functional

$$\mathcal{L}_u(\mathbf{u}, \mathbf{v}_x, \mathbf{v}_y) = \frac{\gamma}{2}\|\mathbf{H}\mathbf{u}-\mathbf{g}\|^2 + \Phi(\mathbf{v}_x, \mathbf{v}_y) + \frac{\alpha}{2}\|\mathbf{D}_x\mathbf{u}-\mathbf{v}_x-\mathbf{a}_x\|^2 + \frac{\alpha}{2}\|\mathbf{D}_y\mathbf{u}-\mathbf{v}_y-\mathbf{a}_y\|^2 \tag{24}$$

and solving it with an iterative algorithm:

Algorithm: $\widehat{\mathbf{u}} = u\text{-step}(\mathbf{u}^0)$

1: Set $\mathbf{v}_x^0 = \mathbf{v}_y^0 = \mathbf{a}_x^0 = \mathbf{a}_y^0 = 0$ and $j = 0$

2: **repeat**

3: $\mathbf{u}^{j+1} = \arg\min_{\mathbf{u}} \mathcal{L}_u(\mathbf{u}, \mathbf{v}_x^j, \mathbf{v}_y^j) \iff [\mathbf{H}^T\mathbf{H} + (\alpha/\gamma)(\mathbf{D}_x^T\mathbf{D}_x + \mathbf{D}_y^T\mathbf{D}_y)]\mathbf{u}^{j+1} = \mathbf{H}^T\mathbf{g} + (\alpha/\gamma)[\mathbf{D}_x^T(\mathbf{v}_x^j + \mathbf{a}_x^j) + \mathbf{D}_y^T(\mathbf{v}_y^j + \mathbf{a}_y^j)]$

4: $\{\mathbf{v}_x^{j+1}, \mathbf{v}_y^{j+1}\} = \arg\min_{\mathbf{v}_x,\mathbf{v}_x} \mathcal{L}_u(\mathbf{u}^{j+1}, \mathbf{v}_x, \mathbf{v}_y) \iff$
$[\mathbf{v}_x^{j+1}]_i = [\mathbf{D}_x\mathbf{u}^{j+1} - \mathbf{a}_x^j]_i[\mathbf{s}]_i^{-1}\max([\mathbf{s}]_i - (1/\alpha), 0)$,
$[\mathbf{v}_y^{j+1}]_i = [\mathbf{D}_y\mathbf{u}^{j+1} - \mathbf{a}_y^j]_i[\mathbf{s}]_i^{-1}\max([\mathbf{s}]_i - (1/\alpha), 0)$,

where

$$[\mathbf{s}]_i = \sqrt{[\mathbf{D}_x\mathbf{u}^{j+1} - \mathbf{a}_x^j]_i^2 + [\mathbf{D}_y\mathbf{u}^{j+1} - \mathbf{a}_y^j]_i^2}$$

5: $\mathbf{a}_x^{j+1} = \mathbf{a}_x^j - \mathbf{D}_x\mathbf{u}^{j+1} + \mathbf{v}_x^{j+1}$

$\mathbf{a}_y^{j+1} = \mathbf{a}_y^j - \mathbf{D}_y\mathbf{u}^{j+1} + \mathbf{v}_y^{j+1}$

6: $j \leftarrow j + 1$

7: **until** stopping criterion is satisfied

8: **return** $\widehat{\mathbf{u}} \leftarrow \mathbf{u}^j$

This iterative algorithm consists of three update steps: lines 3, 4, and 5. Variables $\mathbf{a}_x$ and $\mathbf{a}_y$ are introduced by the ALM. Their update on line 5 is trivial. It is worth drawing a relation of the ALM to a penalty method. If we omit the updating step for $\mathbf{a}_x$ and $\mathbf{a}_y$, and keep $\mathbf{a}_x = \mathbf{a}_y = 0$, the above algorithm defaults to the penalty method. The penalty method converges to the solution of the constrained problem (23) only if we keep increasing α to infinity while iterating, as advocated in [30]. This is however not practical as the problem becomes gradually more ill-posed with increasing α. This drawback is avoided in the ALM. Since Φ is a lower semicontinuous proper convex function,[2] and $[\mathbf{D}_x^T, \mathbf{D}_y^T]^T$ has a full column rank, then, if (23) has a solution, the u-step algorithm converges to this solution even for α that is relatively small and fixed. This important theorem was proved in [31].

[2]In our case, Φ is continuous and thus lower semicontinuous

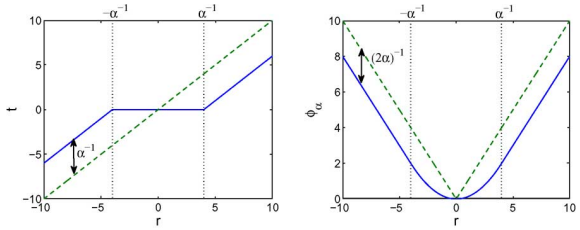


Fig. 7. Soft thresholding. (a) Shrinkage formula (26) for a nonzero threshold $1/\alpha$ (solid) and for $1/\alpha = 0$ (dashed). (b) Corresponding ϕ_α in (25) for a (solid) nonzero threshold $1/\alpha$ and (dashed) for $1/\alpha = 0$. Note that ϕ_α is a relaxed form of the ℓ_1 norm, which is the absolute value (dashed) in this simple case.

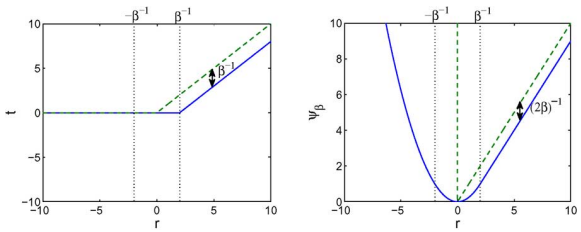


Fig. 8. Thresholding in the blur domain. (a) Shrinkage formula (32) for (solid) a nonzero threshold $1/\beta$ and for (dashed) $1/\beta = 0$. (b) Corresponding ψ_β in (31) for a nonzero threshold $1/\beta$ (solid) and for $1/\beta = 0$ (dashed).

Since $\mathcal{L}_u$ in (24) is quadratic with respect to $\mathbf{u}$, minimization on line 3 is a solution to a set of linear equations. We show later that this can be solved efficiently in the Fourier domain.

The beauty of variable splitting is that minimization with respect to $\mathbf{v}_x$ and $\mathbf{v}_y$ is, by definition, the Moreau proximal mapping [32] of Φ applied to $\mathbf{D}_x\mathbf{u}^{j+1} - \mathbf{a}_x^j$ and $\mathbf{D}_y\mathbf{u}^{j+1} - \mathbf{a}_y^j$. The problem can be solved for each ith element independently. Let $\mathbf{t} = [[\mathbf{v}_x]_i, [\mathbf{v}_y]_i]^T$ and $\mathbf{r} = [[\mathbf{D}_x\mathbf{u} - \mathbf{a}_x]_i, [\mathbf{D}_y\mathbf{u} - \mathbf{a}_y]_i]^T$ be vectors of size 2×1; the problem on line 4 is of the form

$$\phi_\alpha(\mathbf{r}) = \min_{\mathbf{t}} \left(\frac{\alpha}{2}\|\mathbf{r} - \mathbf{t}\|^2 + \|\mathbf{t}\| \right) \tag{25}$$

and, as proved in [30], the minimum is reached for

$$\mathbf{t} = \frac{\mathbf{r}}{\|\mathbf{r}\|} \max\left(\|\mathbf{r}\| - \frac{1}{\alpha}, 0 \right) \tag{26}$$

which is a generalized shrinkage formula for vectors. For the $\mathbf{r}$ scalar, (26) corresponds to a well-known soft-thresholding formula plotted as a solid line in Fig. 7(a). It is interesting to note that, after substituting for $\mathbf{t}$ in (25), $\phi_\alpha(\mathbf{r})$ [solid line in Fig. 7(b)] can be written in a closed form

$$\phi_\alpha(\mathbf{r}) = \begin{cases} \frac{\alpha}{2}\|\mathbf{r}\|^2, & \text{if } \|\mathbf{r}\| < \frac{1}{\alpha} \\ \|\mathbf{r}\| - \frac{1}{2\alpha}, & \text{otherwise} \end{cases} \tag{27}$$

which is a relaxed form of the original $\phi(\mathbf{r}) = \|\mathbf{r}\|$ in the isotropic TV definition (6). If $\alpha \to \infty$, then $\phi_\alpha \to \phi$, and the corresponding graphs are plotted as dashed lines in Fig. 7.

B. H-Step

The kernel estimation proceeds analogously to the u-step. Using the proposed regularizer (20), minimization with respect to the PSFs (5) writes as

$$\min_{\mathbf{h}} \frac{\gamma}{2}\|\mathbf{U}\mathbf{h} - \mathbf{g}\|^2 + \frac{\delta}{2}\mathbf{h}^T\mathbf{R}_\Delta\mathbf{h} + \Psi(\mathbf{h}) \tag{28}$$

Applying variable splitting $\mathbf{w} = \mathbf{h}$ yields the constrained problem

$$\min_{\mathbf{h},\mathbf{w}} \frac{\gamma}{2}\|\mathbf{U}\mathbf{h} - \mathbf{g}\|^2 + \frac{\delta}{2}\mathbf{h}^T\mathbf{R}_\Delta\mathbf{h} + \Psi(\mathbf{w}) \quad \text{s.t.} \quad \mathbf{w} = \mathbf{h} \tag{29}$$

Then, we consider the following functional:

$$\mathcal{L}_h(\mathbf{h}, \mathbf{w}) = \frac{\gamma}{2}\|\mathbf{U}\mathbf{h} - \mathbf{g}\|^2 + \frac{\delta}{2}\mathbf{h}^T\mathbf{R}_\Delta\mathbf{h} + \Psi(\mathbf{w}) + \frac{\beta}{2}\|\mathbf{h} - \mathbf{w} - \mathbf{b}\|^2 \tag{30}$$

and solve it with the following iterative algorithm:

Algorithm: $\widehat{\mathbf{h}} = h\text{-step}(\mathbf{h}^0)$

1: Set $\mathbf{w}^0 = \mathbf{b}^0 = 0$ and $j = 0$

2: **repeat**

3: $\mathbf{h}^{j+1} = \arg\min_{\mathbf{h}} \mathcal{L}_h(\mathbf{h}, \mathbf{w}^j) \iff [\mathbf{U}^T\mathbf{U} + (\delta/\gamma)\mathbf{R}_\Delta + (\beta/\gamma)\mathbf{I}]\mathbf{h}^{j+1} = \mathbf{U}^T\mathbf{g} + (\beta/\gamma)(\mathbf{w}^j + \mathbf{b}^j)$

4: $\mathbf{w}^{j+1} = \arg\min_{\mathbf{w}} \mathcal{L}_h(\mathbf{h}^{j+1}, \mathbf{w}) \iff [\mathbf{w}^{j+1}]_i = \max([\mathbf{h}^{j+1} - \mathbf{b}^j]_i - (1/\beta), 0)$

5: $\mathbf{b}^{j+1} = \mathbf{b}^j - \mathbf{h}^{j+1} + \mathbf{w}^{j+1}$

6: $j \leftarrow j + 1$

7: **until** stopping criterion is satisfied

8: **return** $\widehat{\mathbf{h}} \leftarrow \mathbf{h}^j$

Matrix $\mathbf{I}$ denotes identity of size $K\widehat{L} \times K\widehat{L}$. As in the u-step, the h-step iterative algorithm consists of three update steps: lines 3, 4, and 5. Since $\mathcal{L}_h$ in (30) is quadratic with respect to $\mathbf{h}$, minimization on line 3 is a solution to a set of linear equations. This time, the minimization with respect to $\mathbf{w}$ is again the Moreau proximal mapping of Ψ applied to $\mathbf{h}^{j+1} - \mathbf{b}^j$, and it is solved elementwise. Let $t = [\mathbf{w}]_i$ and $r = [\mathbf{h} - \mathbf{b}]_i$; the problem on line 4 is of the following form:

$$\psi_\beta(r) = \min_t \left(\frac{\beta}{2}(r - t)^2 + \psi(t) \right) \tag{31}$$

where ψ is our positivity–sparsity enforcing function defined in (21) and plotted as dashed line in Fig. 8(b). After some manipulation, one can see that the minimum is reached for

$$t = \max\left(r - \frac{1}{\beta}, 0 \right). \tag{32}$$

The plot of this "one-sided" thresholding function is the solid line in Fig. 8(a). Using the thresholding function, a closed form of ψ_β is

$$\psi_\beta(r) = \begin{cases} \frac{\beta}{2}r^2, & \text{if } r < \frac{1}{\beta} \\ r - \frac{1}{2\beta}, & \text{otherwise} \end{cases} \tag{33}$$

with a plot in Fig. 8(b), i.e., the solid line. The function linearly increases in the positive domain, whereas in the negative domain, it increases quadratically. If $\beta \to \infty$, then $\psi_\beta \to \psi$, and the thresholding function in (32) approaches the dashed line in Fig. 8(a). However, as in the u-step, we do not need to increase β to infinity for the h-step algorithm to converge to the solution of the constrained problem (29). The ALM approach with its extravariable $\mathbf{b}$ converges. Note, that ψ must be a lower semi-continuous proper convex function for the method to converge, which is the case. Interestingly, if we replaced in definition (21) infinity with some large but finite numbers, the resulting function would no longer be convex. Infinity in the definition might look dangerous, but it turns out to give an elegant solution in the form of the thresholding function (32).

C. Implementation

We have analyzed the main points (u-step and h-step) of the optimization algorithm. Now we proceed with the description of the main loop of the algorithm and the computational cost of individual steps. Let N denote the number of pixels in the output image u, and let $\widehat{L}$ denote the number of pixels in our overestimated PSF support. The main loop of the MC blind deconvolution alternating minimization algorithm looks as follows:

MC blind deconvolution

Require: $K \geq 2$ input images $\{\mathbf{g}_k\}$; blur size $\widehat{L}$; parameters α, β, δ, γ

1: Set $j = 0$, $\widehat{\mathbf{h}}_k^0$'s to delta functions, and $\widehat{\mathbf{u}}^0 = 0$

2: Calculate $\mathbf{R}_\Delta$

3: **repeat**

4: $\widehat{\mathbf{u}}^{j+1} = u\text{-step}(\widehat{\mathbf{u}}^j, \widehat{\mathbf{h}}^j)$

5: $\widehat{\mathbf{h}}^{j+1} = h\text{-step}(\widehat{\mathbf{u}}^{j+1}, \widehat{\mathbf{h}}^j)$

6: $j \leftarrow j + 1$

7: **until** stopping criterion is satisfied

8: **return** $\widehat{\mathbf{u}} \leftarrow \widehat{\mathbf{u}}^j$

The stopping criterion, which we typically use, is $\|\widehat{\mathbf{h}}^j - \widehat{\mathbf{h}}^{j-1}\|/\|\widehat{\mathbf{h}}^j\| < tol$. The same can be used in the h-step and, likewise, in the u-step using $\mathbf{u}$ instead of $\mathbf{h}$. The calculation of $\mathbf{R}_\Delta$ can be done using the fast Fourier transform (FFT) without explicitly constructing the convolution matrices $\mathbf{G}_k$. Since $\mathbf{G}_k$ values are "valid" convolutions, we can construct only one row of $\mathbf{R}_\Delta$ at a time, and the overall complexity is thus $O(K\widehat{L}N \log N)$.

In general, the most time-consuming is the u-step, which requires an inversion of the huge $N \times N$ matrix $[\mathbf{H}^T\mathbf{H} + (\alpha/\gamma)(\mathbf{D}_x^T\mathbf{D}_x + \mathbf{D}_y^T\mathbf{D}_y)]$. One can apply iterative solvers, such as conjugate gradient, to avoid direct inversion, but we can do even better and have a direct solver. In our formulation, $\mathbf{H}$, $\mathbf{D}_x$, and $\mathbf{D}_y$ are convolution matrices. To avoid any ringing artifacts close to image boundaries, they should perform "valid" convolution, i.e., the output image is smaller and covers a region where both the input image and the convolution kernel are fully defined. If we properly adjust the image borders, e.g., by using the function edgetaper in MATLAB, we can replace "valid" convolution with block-circulant one, and ringing artifacts will be almost undetectable. The TV regularizer also helps to reduce such artifacts. FFT diagonalizes block-circulant convolution matrices, and inversion is thus straightforward. The remaining update steps for $\mathbf{v}_x$ ($\mathbf{v}_y$) and $\mathbf{a}_x$ ($\mathbf{a}_y$) are simple and can be computed in $O(N)$ time. The u-step is thus carried out with an overall $O(N \log N)$ cost.

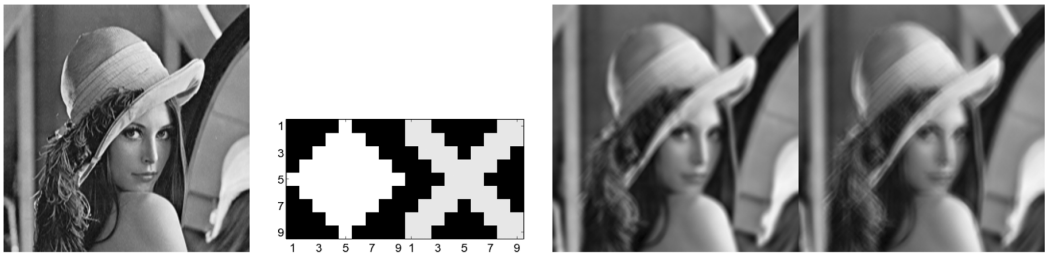
Unlike the u-step, which is calculated almost entirely in the Fourier domain, we perform the h-step in the image domain since we need the constrained kernel support $\widehat{L}$. Otherwise, $\mathbf{R}_\Delta$ becomes a very uninformative regularizer, as explained in Section IV-C. On line 3 of the h-step algorithm, we have to invert matrix $[\mathbf{U}^T\mathbf{U} + (\delta/\gamma)\mathbf{R}_\Delta + (\beta/\gamma)\mathbf{I}]$, which is of size $K\widehat{L} \times K\widehat{L}$ and thus much smaller than the matrix in the u-step. Typically, the size of blurs is not more than 40×40 pixels ($L = 1600$), and for two input images ($K = 2$), the matrix size is 3200×3200, which is still relatively small.[3] One can again apply an iterative solver such as a conjugate gradient, but we found it much more efficient to store the whole matrix and perform Cholesky decomposition to solve this problem. This can be computed in $O((K\widehat{L})^3)$ time. Again, update steps for $\mathbf{w}$ and $\mathbf{b}$ are very simple and require $O(K\widehat{L})$ operations.

Setting parameters is based solely on our empirical studies and cannot be considered as a rigorous procedure. The optimization method has four parameters. We have noticed that, in general, they can be fixed relative to one of them, i.e., γ, which depends on the noise level. This observation is not superficial. Afonso *et al.* [5] (as well as is [4] for the split Bregman method) also recommend to set parameters introduced by the ALM, i.e., in our case, α and β, with respect to the weight γ of the fidelity term. Parameter δ, which is the weight of the MC constraint term $\mathbf{h}^T\mathbf{R}_\Delta\mathbf{h}$, is proportional to the noise variance, as shown in (16), and therefore should be fixed to γ as well. The role of thumb is to set γ equal to a ratio of signal and noise variances, i.e., SNR $= 50$ dB $\Rightarrow \gamma = 10^5$ or SNR $= 20$ dB $\Rightarrow \gamma = 10^2$, etc.[4] Then, we have found that choosing $\alpha = 10^{-1}\gamma$, $\beta = 10^4\gamma$, and $\delta = 10^3\gamma$ usually results in good convergence. For higher noise levels (smaller γ), we observed that $\delta = 10^2\gamma$ is better.

In our experiments, the number of iteration in the main loop and in the u-step and h-step typically did not exceed ten. In order to further decrease computational time, we tried to modify the algorithm in several ways. For example, we found it very effective to divide the algorithm into two stages. In the first stage, we select a small (typically 256×256) central region from input images and run the algorithm on this selection. In the second stage, we take the estimated PSFs from the first stage and apply one u-step on the whole image in order to obtain the final reconstructed image. The usable output of the first stage are thus PSFs and not the reconstructed central region. We observed that fixing γ to ten (even for the SNR above 10 dB) in the first stage and

[3] A matrix of such size, if stored in double precision, occupies approximately 78 MB of memory, which current computers can easily handle.

[4] We use a standard definition of the signal to noise ratio, SNR $= 10\log(s^2/\sigma^2)$, where s^2 and σ^2 are the signal and noise variances, respectively.



(a) (b) (c)

Fig. 9. Test data set. (a) Original image 256 × 256. (b) Two blurs 9 × 9. (c) Example of an input blurry pair with SNR = 50 dB.

setting other parameters according to formulas as shown above produces accurate PSFs in a more reliable way. This convergence boost can be explained by noting that the reconstructed image for lower γ becomes more piecewise constant (patchy) with only strong edges preserved, which makes the h-step in the fidelity term focus only on areas around strong edges and neglect areas with details that are prone to noise.

Another modification, which proved to be a minor improvement, was to estimate PSFs in a multiscale fashion. Initializing with upsampled PSFs from the courser levels tend to decrease the number of iterations. However, we observed that more than two levels (half-sized and original scale) are not necessary and that the choice of the upsampling algorithm is important. Simple linear upsampling generates PSFs that are wider than the true PSFs on that scale, and we waste several iterations of the algorithm to shrink the PSFs back. In our tests, we were using a Lanczos interpolation method, which seems to give the best results.

To provide the cost of individual steps in terms of computer time, we performed blind deconvolution of two one-megapixel images with PSF size 40 × 40 on a 2.7-GHz Pentium Dual-Core CPU using our MATLAB implementation. The cost of one iteration inside the u-step and the h-step is around 0.8 and 4.5 s, respectively. Calculating matrix $\mathbf{R}$ using the whole images takes 11 min in this case, which is clearly the most time-consuming step. However, as pointed out earlier, we can calculate $\mathbf{R}$ on a small region. For example, for a 256 × 256 block, the calculation (same PSF size 40 × 40) then takes around 30 s.

VI. Experiments

In order to illustrate the favorable convergence properties of the proposed algorithm, we performed two sets of experiments. The first set works with synthetically blurred data and compares convergence and quality of PSF and image estimations for different SNRs and blur sizes. The second set of experiments compares the proposed algorithm with another MC blind deconvolution method of Katkovnik *et al.* [33] and demonstrates deconvolution of real photos taken with a standard digital camera.

The setup for the synthetic data experiment was the following. We took the Lena image in Fig. 9(a) and convolve it with two 9 × 9 blurs [see Fig. 9(b)] and add noise at three different levels SNR $= 50, 30,$ and10 dB. An example of blurry images for the least noisy case is in Fig. 9(c). To evaluate performance in every iteration j of the main loop, we use normalized root mean square error defined as NRMSE $= \|\widehat{\mathbf{h}}^j - \mathbf{h}^*\|/\|\mathbf{h}^*\|$, where $\widehat{\mathbf{h}}^j$ is the estimation of PSFs after j iterations and $\mathbf{h}^*$ are the true PSFs. NRMSE as a function of iterations and estimated PSFs for different situations are summarized in Fig. 10. NRMSE is plotted in logarithmic scale. Three graphs correspond to three levels of SNRs. In each case, we ran the algorithm with three different PSF supports: 9 × 9 (solid line), 15 × 15 (dotted line), and 21 × 21 (dashed line). The corresponding estimated sharp images for the PSF support 21 × 21 and are summarized in Fig. 11. One can see that the proposed method provides accurate results regardless of the degree of PSF size overestimation and shows robustness with respect to noise.

There are several interesting points we can draw from the obtained results. First of all, the MSE decreases very quickly. In most of the cases, after five iterations, MSE remains almost constant. For overestimated blur supports (dotted and dashed line) MSE reaches almost the same level as for the correct blur support (solid line), but the decrease is slightly less sharp (particularly visible for SNR = 50 dB). This is logical since, in the overestimated case, the dimensionality of the problem is higher, and the MC constraint $\mathbf{R}_\Delta$ is less effective, as discussed in Section IV-C. Clearly, as the noise level increases, the lowest attainable MSE increases as well. For SNR = 50 dB [see Fig. 10(a)], estimated PSFs are very accurate. The corresponding estimated image in Fig. 11(a) is almost perfect. For SNR = 30 dB, [see Fig. 10(b)], the estimated PSFs take the shape of the true PSFs but are slightly blurred. The estimated image in Fig. 11(b) still looks very sharp and artifact free. As the noise level increases further to SNR = 10 dB [see Fig. 10(c)], the quality of deconvolution starts to deteriorate, but the TV denoising feature of the method is evident, as shown in Fig. 11(c).

There are few data in the literature to which we can directly compare which uses multiple frames in the process. Most of the MC work presented in the introduction is mainly theoretical and presents no algorithms for large-scale problems. Comparison with single-channel results is possible, but we do not feel that this is fair to these other methods. To our knowledge, the only recent method, which is intrinsically MC and claims to work with large kernels, was proposed in [33]. This method performs alternating minimization by switching between minimization with respect to the image (corresponds to our u-step) and minimization with respect to the kernels (corresponds to our h-step). A variation of the steepest descent algorithm is used

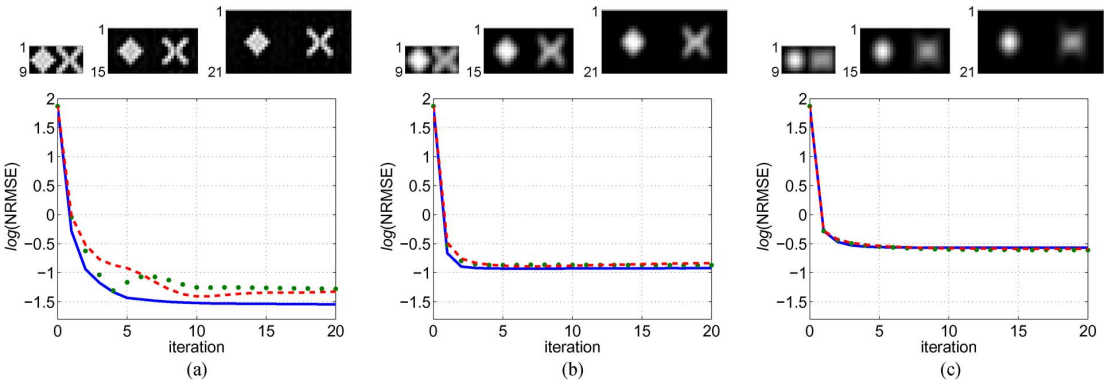


Fig. 10. Estimated PSFs and plots of NRMSE for different noise levels in input blurry images: (a) 50, (b) 30, and (c) 10 dB. Three different PSF supports were considered in each noisy case: (solid line) correct PSF size 9×9, (dotted line) two overestimated sizes 15×15, and (dashed line) 21×21. (a) 50 dB. (b) 30 dB. (c) 10 dB.

Fig. 11. Estimated sharp images for the PSF size set to 21×21 and three different noise levels: (a) 50, (b) 30, and (c) 10 dB. Results are arranged as in Fig. 10. The first row shows one of the input images and the second row shows the estimated image. (a) 50 dB. (b) 30 dB. (c) 10 dB.

for minimization. Everything is implemented in the Fourier domain, as in our case. For minimization, we use ALM in order to work with nonlinear regularization terms in an efficient manner. Katkovnik *et al.* use a variation of the steepest descent algorithm with only quadratic terms. Instead of using regularization, they project current estimation after every iteration into an admissible set of solutions (such as positive PSFs with limited support and image intensity values between 0 and 1) and perform spatially adaptive image denoising based on the intersection-of-confidence-interval rule. To compare the methods, we took a data set generated in [15], which contained four images blurred by eight PSFs providing 32 blurred images [see Fig. 12(a) and (b)]. The blurred images are real and captured by a digital camera. The ground-truth PSFs in Fig. 12(b) were estimated by a collection of point sources installed in the observed scene. We divided the blurred images into eight groups (each containing one image blurred by four blurs) and applied both methods. The NRMSE of the estimated images and blurs are plotted in Fig. 12(c) and (d). One can see that, in half of the cases, our method provides better PSFs (in the NRMSE sense) and outperforms the other method in the image NRMSE in all eight cases. In addition, our method requires only ten iterations of alternating minimization, whereas the other method requires roughly 100 iterations to achieve these results.[5]

In order to demonstrate that the algorithm works well in many practical applications, we took several pairs of images with a 3-megapixel digital camera Olympus C3020Z and applied the proposed algorithm. Light conditions were low, and the shutter speed of the camera was typically longer than 1/10 s. Such setting produces nice blurry images, when the camera is held in hands. It is of course necessary to first register the input photos before the algorithm can be applied. In our case, we do not have

[5]It is true that we perform at most ten iterations inside both u-step and h-step. Katkovnik's method cannot perform many iterations inside their u-step and h-step since they need to project into the admissible set frequently; therefore, they do ten steps of steepest descent in the h-step and one step in the u-step.

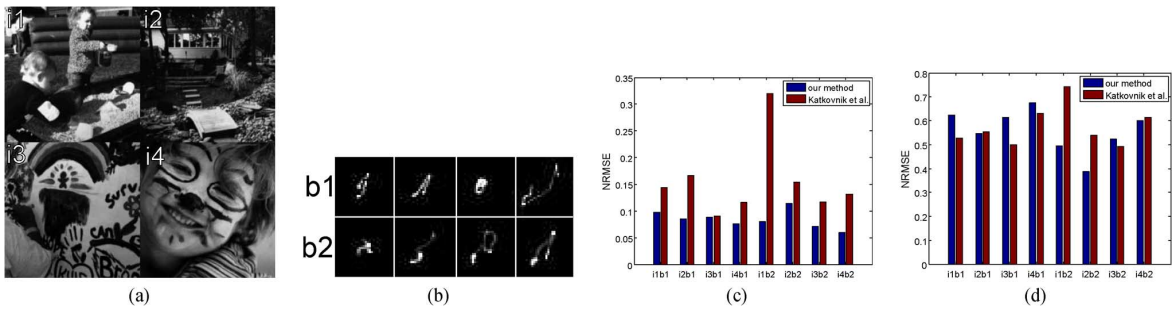


Fig. 12. Comparison with Katkovnik *et al* [33] Ground-truth data from Levin's data set [15]: (a) four images and (b) eight blur kernels, which generates 32 blurred images. We split the kernels into two groups (b1, b2) and got eight input sets each containing four blurred images. (c) NRMSE of estimated sharp images. (d) NRMSE of estimated kernels. Left bars are results of our method and right bars are results of [33].

Fig. 13. Real data set. (a) and (b) Ttwo input blurry images of size 2048 × 1536. (c) Estimated output sharp image using the proposed algorithm. (d) Closeups of the input images and the output, and estimated PSFs of size 50 × 30.

to deal with heavily misregistered data since the images have been taken one after another with minimum delay. A fast registration method, which proved to be adequate and was used in these experiments, works as follows. A reference image is selected from the input set $\{\mathbf{g}_k\}$, and the other images (called sensed images) are sequentially registered to the reference one. The reference and sensed image are first divided into several nonoverlapping blocks (typically 6×6). Phase correlation is applied in each block to determine the integer translation vector between the reference and the sensed block. The estimated shifts $(6 * 6 = 36)$ are used to calculate parameters of an affine transform. The sensed images are then interpolated using the estimated affine transforms.

Reconstruction results for two different data sets are in Figs. 13 and 14. Input image pairs exhibit relatively large blurring, but the reconstructed images are sharp and with negligible artifacts (see image closeups for better visual comparison). Estimated PSF pairs model very well motion blurs induced by camera shake. Some artifacts are visible in the second data set [see Fig. 14(c)] around the snow heap in the left bottom corner. It is very likely, that the blur is slightly different in this part due to a different distance from the camera or due to rotational movement during acquisition. Since our method assumes space-invariant blurs, such artifact are however inevitable.

VII. Conclusion

We have presented a new algorithm for solving MC blind deconvolution. The proposed approach starts by defining an optimization problem with image and blur regularization terms. To force sparse image gradients, the image regularizer is formu-

Fig. 14. Real data set. (a) and (b) Two input blurry images of size 2048 × 1536. (c) Estimated output sharp image using the proposed algorithm. (d) Closeups of the input images and the output and estimated PSFs of size 40 × 40.

lated using a standard isotropic TV. The PSF regularizer consists of two terms: MC constraint (matrix $\mathbf{R}_\Delta$) and sparsity–positivity. The MC constraint is improved by considering image Laplacian, which brings better noise robustness at little cost. Positivity helps the method to convergence to a correct solution, when the used PSF size is much larger than the true one. The proposed approach solves the optimization problem in an iterative way by alternating between minimization with respect to the image (u-step) and with respect to the PSFs (h-step). Sparsity and positivity imply nonlinearity, but by using the variable splitting and ALM (or split-Bregman method), we can solve each step efficiently, and moreover, convergence of each step is guaranteed. Experiments on synthetic data illustrate fast convergence of the algorithm, robustness to noise, and stability in the case of overestimated PSF sizes. Experiments on large real data underline practical aspects of the algorithm. Current and future work involves extending this approach to the space-variant blur and analyzing the convergence properties.

REFERENCES

[1] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Trans. Graph.*, vol. 26, no. 3, p. 70, Jul. 2007.

[2] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Probl.* vol. 23, no. 3, pp. 947–968, Jun. 2007 [Online]. Available: http://stacks.iop.org/0266-5611/23/i=3/a=007

[3] I. W. Selesnick and M. Figueiredo, "Signal restoration with overcomplete wavelet transforms: Comparison of analysis and synthesis priors," in *Proc. SPIE*, 2009, vol. 7446, p. 74 460D.

[4] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM J. Imag. Sci.* vol. 2, no. 2, pp. 323–343, Apr. 2009 [Online]. Available: http://portal.acm.org/citation.cfm?id=1658384.1658386

[5] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.

[6] G. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its application," *Opt. Lett.*, vol. 13, no. 7, pp. 547–549, Jul. 1988.

[7] T. Chan and C. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.

[8] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3715–3727, Dec. 2006.

[9] *Blind Image Deconvolution, Theory and Application*, P. Campisi and K. Egiazarian, Eds. Boca Raton, FL: CRC Press, 2007.

[10] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *Proc. SIGGRAPH: ACM SIGGRAPH Papers*, New York, 2006, pp. 787–794.

[11] J. Miskin and D. J. MacKay, "Ensemble learning for blind image separation and deconvolution," in *Advances in Independent Component Analysis*, M. Girolani, Ed. New York: Springer-Verlag, 2000.

[12] J. Jia, "Single image motion deblurring using transparency," in *Proc. IEEE Conf. CVPR*, Jun. 17–22, 2007, pp. 1–8.

[13] N. Joshi, R. Szeliski, and D. J. Kriegman, "PSF estimation using sharp edge prediction," in *Proc. IEEE CVPR*, Jun. 23–28, 2008, pp. 1–8.

[14] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," in *Proc. SIGGRAPH: ACM SIGGRAPH*, New York, 2008, pp. 1–10.

[15] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. CVPR*, 2009, pp. 1964–1971.

[16] S. Cho and S. Lee, "Fast motion deblurring," *ACM Trans. Graph. (SIGGRAPH ASIA)*, vol. 28, no. 5, p. 145, Dec. 2009.

[17] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proc. 11th ECCV*, Berlin, Germany, 2010, pp. 157–170 [Online]. Available: http://portal.acm.org/citation.cfm?id=1886063.1886077

[18] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[19] M. Sorel and J. Flusser, "Space-variant restoration of images degraded by camera motion blur," *IEEE Trans. Image Process.*, vol. 17, no. 2, pp. 105–116, Feb. 2008.

[20] T. Schulz, "Multiframe blind deconvolution of astronomical images," *J. Opt. Soc. Am. A*, vol. 10, no. 5, pp. 1064–1073, May 1993.

[21] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 202–219, Feb. 1999.

[22] G. Giannakis and R. Heath, "Blind identification of multichannel FIR blurs and perfect image restoration," *IEEE Trans. Image Process.*, vol. 9, no. 11, pp. 1877–1896, Nov. 2000.
[23] H.-T. Pai and A. Bovik, "On eigenstructure-based direct multichannel blind image restoration," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1434–1446, Oct. 2001.
[24] S. Pillai and B. Liang, "Blind image deconvolution using a robust GCD approach," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 295–301, Feb. 1999.
[25] M. Haindl and S. Šimberová, "Model-based restoration of short-exposure solar images," in *Frontiers in Artificial Intelligence and Applications*, L. Jain and R. Howlett, Eds. Amsterdam, The Netherlands: IOS Press, 2002, pp. 697–706.
[26] G. Panci, P. Campisi, S. Colonnese, and G. Scarano, "Multichannel blind image deconvolution using the bussgang algorithm: Spatial and multiresolution approaches," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1324–1337, Nov. 2003.
[27] F. Šroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Process.*, vol. 14, no. 7, pp. 874–883, Jul. 2005.
[28] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992.
[29] J. B. Reade, *Calculus With Complex Numbers*. Boca Raton, FL: CRC Press, 2003, ch. ch. Fundamental theorem of algebra, pp. 75–81.
[30] J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM J. Imag. Sci.* vol. 2, no. 2, pp. 569–592, Apr. 2009 [Online]. Available: http://portal.acm.org/citation.cfm?id=1658384.1658394
[31] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 3, pp. 293–318, Jun. 1992.
[32] J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.
[33] V. Katkovnik, D. Paliy, K. Egiazarian, and J. Astola, "Frequency domain blind deconvolution in multiframe imaging using anisotropic spatially-adaptive denoising," in *Proc. 14th EUSIPCO*, Sep. 2006, pp. 1–5.

Filip Šroubek (M'08) received the M.S. degree in computer science from the Czech Technical University, Prague, Czech Republic, in 1998 and the Ph.D. degree in computer science from Charles University, Prague, in 2003.

From 2004 to 2006, he was on a postdoctoral position with the Instituto de Optica, CSIC, Madrid, Spain. In 2010 and 2011, he was the Fulbright Visiting Scholar with the University of California, Santa Cruz. He is currently with the Institute of Information Theory and Automation and partially also with the Institute of Photonics and Electronics, Academy of Sciences of the Czech Republic, Prague. He is the author of seven book chapters and over 80 journal and conference papers on image fusion, blind deconvolution, super-resolution, and related topics.

Peyman Milanfar (F'10) received the B.S. degree in electrical engineering and mathematics from the University of California, Berkeley, in 1988 and the M.S., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1990, 1992, and 1993, respectively.

Until 1999, he was a Senior Research Engineer with SRI International, Menlo Park, CA. He is currently a Professor of electrical engineering with the University of California, Santa Cruz. From 1998 to 2000, he was a Consulting Assistant Professor of computer science with Stanford University, Stanford, CA. In 2002, he was also a Visiting Associate Professor with Standford University. His research interests include statistical signal, image processing, and inverse problems.

Dr. Milanfar is a member of the Signal Processing Society Image, Video, and Multidimensional Signal Processing Technical Committee. From 1998 to 2001, he was an Associate Editor for the IEEE Signal Processing Letters and was an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2005–2010. He is currently on the editorial board of the *SIAM Journal of Imaging Science* and *Image and Vision Computing*. He was the recipient of the US National Science Foundation CAREER award, and the best paper award from the IEEE Signal Processing Society in 2010.

Journal of Biomedical Optics 16(11), 116016 (November 2011)

Retinal image restoration by means of blind deconvolution

Andrés G. Marrugo,[a] Michal Šorel,[b] Filip Šroubek,[b] and María S. Millán[a]
[a]Universitat Politècnica de Catalunya, Department of Optics and Optometry, Group of Applied Optics and Image Processing, Violinista Vellsolà 37, Terrassa, Barcelona 08222 Spain
[b]Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, Prague 8, 18208 Czech Republic

Abstract. Retinal imaging plays a key role in the diagnosis and management of ophthalmologic disorders, such as diabetic retinopathy, glaucoma, and age-related macular degeneration. Because of the acquisition process, retinal images often suffer from blurring and uneven illumination. This problem may seriously affect disease diagnosis and progression assessment. Here we present a method for color retinal image restoration by means of multichannel blind deconvolution. The method is applied to a pair of retinal images acquired within a lapse of time, ranging from several minutes to months. It consists of a series of preprocessing steps to adjust the images so they comply with the considered degradation model, followed by the estimation of the point-spread function and, ultimately, image deconvolution. The preprocessing is mainly composed of image registration, uneven illumination compensation, and segmentation of areas with structural changes. In addition, we have developed a procedure for the detection and visualization of structural changes. This enables the identification of subtle developments in the retina not caused by variation in illumination or blur. The method was tested on synthetic and real images. Encouraging experimental results show that the method is capable of significant restoration of degraded retinal images. [DOI: 10.1117/1.3652709]

Keywords: blind deconvolution; image restoration; deblurring; retinal image.

Paper 11248RR received May 18, 2011; revised manuscript received Sep. 17, 2011; accepted for publication Sep. 22, 2011; published online Oct. 31, 2011.

1 Introduction

A fundus imaging device or retinal camera is a specialized low-power microscope with an attached camera designed to photograph the interior of the eye in association with the optical system of the eye. Retinal imaging is acknowledged to be an important tool for both detection and monitoring the progression of diseases affecting the eye, such as diabetic retinopathy, glaucoma, and age-related macular degeneration.[1] The digital format provides a permanent record of the appearance of the retina at any point in time.[2]

The imaging procedure is usually carried in two separate steps: Image acquisition and diagnostic interpretation. Image quality is subjectively evaluated by the person capturing the images, and they can sometimes mistakenly accept a low-quality image.[3] Low-quality image occurrence rate has been reported at 3.7–19.7% in clinical studies,[4–6] which is not a minor fact. A recent study by Abràmoff et al.[7] using an automated system for detection of diabetic retinopathy found that from 10,000 exams 23% had insufficient image quality. A major source of retinal image quality degradation are aberrations of the human eye, imperfections in the fundus camera optics, and improper camera adjustment, flash lighting, or focusing during the exam.[8] Moreover, regardless of how well controlled the aforementioned parameters are, in practice it may not always be possible to obtain good enough image quality as a result of additional factors such as lens opacities in the examined eye, scattering, insufficient pupil dilation or patient difficulty in steady fixating a target in the camera (such as in patients suffering from amblyopia).[3] Out of all possible retinal image degradations, some can be properly compensated via enhancement or restoration techniques (e.g., low-contrast, nonuniform illumination, noise, and blur).[2] However, this compensation is also dependent on the extent of the degradation. Regarding retinal image blurring, its main causes are relative camera-eye motion, inherent optical aberrations in the eye, and improper focusing.

In the past decade, many wavefront technologies—with its origins in astronomy—such as adaptive optics (AO)[9] and deconvolution from wavefront sensing (DWFS),[10] gave rise to the correction of monochromatic aberrations of the eye and also created new opportunities to image the retina at unprecedented spatial resolution. However, AO-corrected and DWFS-based fundus imagers usually aim at resolving details at the level of individual photoreceptors, thus have a field of view (FOV) of a couple degrees and a high resolution on the order of 1 or 2 μm.[11] Greater FOVs can be achieved (~5 deg)[12,13] with additional hardware constraints, beside the fact that diffraction limited imaging is not guaranteed due to an increase in aberrations.[14] Nevertheless, it is still a considerably narrow FOV and a major disadvantage with clinical subjects because of the need to examine larger areas of the retina. On the other hand, regular non-AO corrected fundus imagers used for routine checkups have a large FOV (typically, 30 deg) at the expense of lower spatial resolution, but still sufficient for practical detection and progression of observable

Address all correspondence to: Andrés G. Marrugo, Universitat Politècnica de Catalunya, Department of Optics and Optometry, Group of Applied Optics and Image Processing, Violinista Vellsolà 37, Terrassa, Barcelona 08222 Spain; Tel: 3493738678; E-mail: andres.marrugo@upc.edu.

1083-3668/2011/16(11)/116016/11/$25.00 © 2011 SPIE

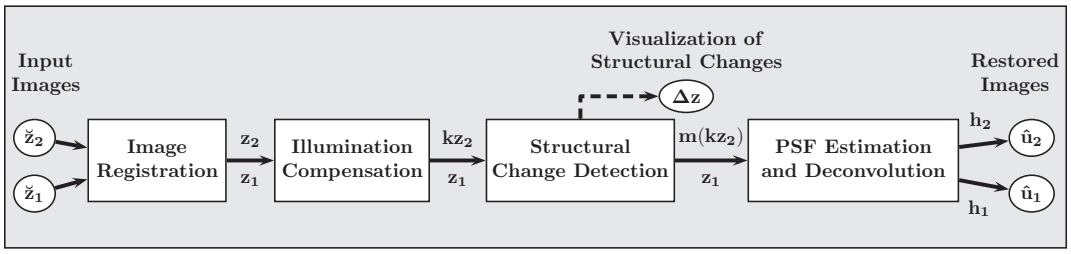


Fig. 1 Block diagram illustrating the proposed method. $\check{z}_i$ are the unregistered degraded input images, and $\hat{u}_i$ are their restored versions. The other variables are intermediate outputs of every stage; their meaning is given in the text.

clinical signs, such as microaneurysms, dot and blot hemorrhages, and exudates, among others. Consequently, large FOV fundus imagers are the major imaging modality available to patients visiting an eye-care clinic. The method proposed herein aims to restore images from conventional large FOV fundus imagers.

Among the normal retinal features, the blood vessel distribution exhibits a unique pattern in each individual and is highly stable in time. It is quite difficult to forge, and most common diseases do not change the pattern in a way that its topology is affected. For that reason, much effort has been put into the development of security systems based on the blood vessel distribution as a biometric signal for authentication purposes.[15] From this consideration, it is reasonable to assume the hypothesis that a pair of fundus images of the same retina, taken at different moments in time, contain enough common information to restore any of them by existing multichannel deconvolution techniques. We will demonstrate this fact later.

1.1 *Overview of Proposed Approach*

In this paper, we propose a new strategy for retinal image deblurring where we consider the most general image degradation case: blurred retinal images acquired in different moments in time, ranging from minutes to months; hence, disease progression is also considered. The main reason for this general image degradation case that considers long time lapses comes from the potential need to restore a degraded image acquired in the past being the only one available at that stage of the disease. This problem arises quite often in clinical practice. A correct assessment of a patient's state evolution requires sharp images from all moments in time; the method proposed here enables such opportunity. Disease progression characterization is embedded in the algorithm with the identification of areas of structural change (see Sec. 3.3).

Our restoration method is based on a technique called blind deconvolution (BD).[16,17] The goal of BD is to recover the original scene from a single image or a set of blurred images in the presence of a poorly determined or unknown point-spread function (PSF). The main assumption is that the blur can be described by a convolution of a sharp image with the unknown PSF. Restoration by deconvolution improves contrast and resolution of digital images (i.e., it is easier to resolve and distinguish features in the restored image). To avoid confusion with super-resolution, we briefly describe what we mean by resolution improvement. Digital deconvolution can be described as any scheme that sharpens up the PSF, while the spatial frequency bandwidth remains unchanged. This means that the spatial frequency response and the two-point resolution is improved, but the cutoff frequency is unchanged;[18] in the super-resolution context, the goal is to increase the cutoff frequency.

BD algorithms can be of single input [single-image blind deconvolution (SBD)] or of multiple images [multichannel blind deconvolution (MBD)]. Despite the fact that SBD is one of the most ill-posed problems, there are several reliable SBD algorithms,[19] although most of them require that the blurred image be governed by relatively strong edges, which is not case here. In Sec. 4.1 we compare our approach to a recent state-of-the-art SBD method.[20] The computational overhead from MBD (all of the preprocessing to adjust the time-sequence of images) in comparison to SBD is practically negligible, and the robustness of MBD is far superior and worth applying because SBD fails to produce a valid restoration. By the same token, the additional processing enables the identification of structural changes in the retina over time—a central task in medical practice. As a result, we have chosen a multichannel approach for the restoration of blurred retinal images.

An overview of the proposed approach is described in Fig. 1. We consider as input two-color retinal images acquired with a conventional fundus camera within a time lapse that can span from several minutes to months given by routine patient checkups. The images correspond to the same retina but can differ with respect to illumination distribution, blur, and local structural changes given by pathological developments. These differences cannot solely be accounted for by the convolutional model described in Sec. 2. For that reason, the images must be preprocessed before the blind deconvolution stage can take place. We register the images and compensate for inter-image illumination variation and structural changes. In fact, this preprocessing work becomes a great opportunity to meet one of the main concerns of ophthalmologists when they visually compare fundus images of the same retina over time: To identify true structural or morphological changes pertaining to possible pathological damage and, consequently, disregarding other changes merely caused by variation of illumination or blur. Ours is a two-stage blind deconvolution strategy. The first stage consists in the estimation of the PSFs following a multichannel scheme, and the second stage is the image deconvolution, where we restore every image with its corresponding PSF, independently. This has several advantages that will be explained in detail Sec. 3.5. The multichannel scheme is based on the method described in Ref. 21, which has proved to work well

in practice with sufficient experimental data. It is an alternating minimization scheme based on a maximum *a posteriori* (MAP) estimation, with *a priori* distribution of blurs derived from the multichannel framework and *a priori* distribution of the ideal sharp image defined by regularization with the total variation of the image.[22] MAP is formulated as an optimization problem, where regularization terms are directly related to priors. Regularization involves the introduction of additional information in order to solve an ill-posed problem in the form of a penalty or restriction in the minimization routine (see Sec. 3.4). This provides good quality of restoration—significantly better than, for example, Lucy–Richardson algorithm,[23] still widely used in biomedical applications. We have modified the algorithm in Ref. 21 to leave out regions where the eye fundus has structurally changed (it only takes into account one image in these regions) with the use of a masking operator, similarly to the solution proposed in Ref. 24 within the super-resolution context. This enabled us to restore both degraded input images.

In this work, our novel contributions to the retinal image processing task are twofold. First, we propose a degradation model for time-series retinal images, which captures the underlying distortions resulting from instrument limitations and changes between patient visits; we are also able to identify and highlight such changes. Second, we propose a restoration strategy based on blind deconvolution that is able to obtain image enhancement and resolution improvement using inexpensive digital methods applied to images acquired with a conventional fundus camera.

2 Mathematical Model of Image Degradation

The unregistered input images, as shown in Fig. 1, are $\breve{z}_1$ and $\breve{z}_2$. After registration, we obtain two degraded registered images z_1 and z_2, which we model as originating from an ideal sharp image u. Mathematically, the degradation model is stated as

$$\begin{aligned} z_1 &= u * h_1 + n_1, \\ z_2 &= (uk^{-1}) * h_2 + n_2 \ , \end{aligned} \tag{1}$$

where the asterisk is the standard convolution, h_i are called convolution kernels or PSFs, and k is a function accounting for relative local illumination change between images z_1 and z_2. For pixels where no illumination changes occur, $k \approx 1$. The noise n_i is assumed Gaussian additive with zero mean in both images. In our case, the PSFs and k comprise all radiometric degradations described above except structural changes in the eye, which is treated in Sec. 3.3. Despite the fact that we consider the PSFs to vary in time between the two image acquisitions, we assume them to be spatially invariant within each image. Because the FOV is of 30 deg or less, this assumption can be accepted in the first approach. This ideal sharp image u is actually unknown, and its estimation is the purpose of this paper. Thus to avoid confusion, the estimated (restored) image is denoted by $\hat{u}$. In Sec. 4.1, we test the performance of our method with synthetically degraded images, which means that we know u.

3 Description of the Method

In this section, we describe every stage of the proposed method. To illustrate each stage we use the images shown in Fig. 2. They were acquired using a nonmydriatic digital fundus camera system with conventional xenon flash lighting source (in the visible spectrum). The fundus images are from a patient that suffered from age-related macular degeneration and were captured within a seven-month time lapse. They are color RGB 24 bit-depth fundus images of size 1500×1200 digitized in TIFF format. This is a general example where both images do not correspond exactly to the same object field, the illumination distribution across both images is not exactly the same, and there are some structural differences between them given by the pathological development in the macula (centered yellowish region).

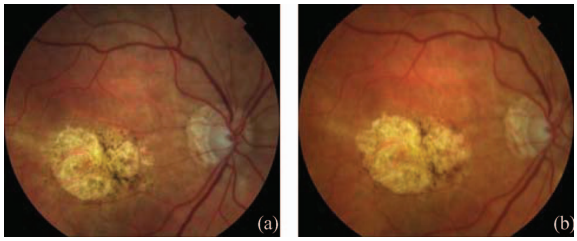


Fig. 2 Color fundus images of a human eye affected by age-related macular degeneration. Images (a) and (b) were captured within a seven-month time lapse, and (a) was captured before (b).

3.1 *Image Registration*

Image registration is a procedure that consists of spatial alignment of two or more images. General and application-specific image registration, such as in retinal imaging, has been investigated from the beginning of image-processing research. The interested reader is referred to the image registration review of Zitová and Flusser[25] and the recent work by Lee et al.[26] for objective validation of several retinal image registration algorithms. Image-registration techniques are usually divided into two groups: intensity-based and feature-based methods. Intensity-based methods have the drawback of poor performance under varying illumination conditions. Feature-based methods are robust to such effects but rely on accurate and repeatable extraction of the features. The retinal vasculature is known to provide a stable set of features for registration.

For registering the images, we use the robust dual-bootstrap iterative closest-point algorithm. We briefly describe it here; for a full description, of the method the reader is referred to Ref. 27. The vasculature from each image is automatically traced; starting from initial seed points extracted from a 1-D edge detection and, later, recursively tracking the vessels using directional templates. The vessel branching and crossover points are used as landmarks to register the images to subpixel accuracy. The registration algorithm starts from initial low-order estimates that are accurate only in small image regions called bootstrap regions. The transformation is then refined using constraints in the region, and the bootstrap region is expanded iteratively. The algorithm stops when the bootstrap region expands to cover the overlap between the images, and uses 12-dimensional quadratic mapping. This transformation model includes rotation, scale, translation, a shearing term, and a quadratic term that describes the spherical shape of the retina. We refer the interested reader to Ref. 28 for details on the model derivation. This registration algorithm is very robust to local changes and low overlap between

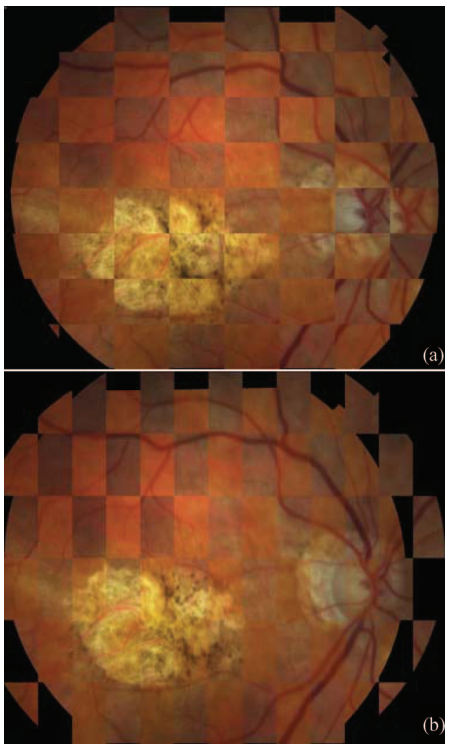


Fig. 3 Registration of images from Fig. 2 in checkerboard representation. (a) Before and (b) after registration.

images as demonstrated by its high success rate on test images with at least one common landmark point and overlaps even as low as 35%.[27] Even though the reported accuracy in Ref. 27 is of subpixel accuracy, in our case of degraded images this can be slightly worse without compromising the outcome. Minor local misregistration errors may occur when landmark points do not match precisely, but they will not be taken into account in the restoration because they will be masked out before the PSF estimation and image deconvolution stages (see Sec. 3.3).

To confirm the registration outcome, the pair of images before and after registration are shown in Fig. 3 in checkerboard representation, where the images are merged together in a chesslike pattern, where each square alternates information from one image to the other. Note how after registration the images have been correctly aligned, especially the blood vessel distribution.

3.2 *Compensation of Uneven Illumination*

Despite controlled conditions in retinal image acquisition, such as optical stops to prevent glares and provide a diffuse illumination, there are many patient-dependent aspects that are difficult to control and mainly affect the illumination component with gradual nonuniform spatial variations. Some of the contributing factors are (i) the curved surface of the retina (as a consequence, all regions cannot be illuminated uniformly); (ii) imaging requires either a naturally or an artificially dilated pupil (The degree of dilation is highly variable across patients); (iii) unexpected movements of the patient's eye; and (iv) presence of diseases. This nonuniform illumination across the image results in shading artifacts and vignetting. This effect hinders both quantitative image analysis and the reliable operation of subsequent global operators.

In our model, described by Eq. (1), the relative changes in intensity between the two fundus images cannot be described exclusively by convolution with different PSFs and must be compensated by k. A number of general-purpose techniques have been investigated to attenuate the variation of illumination. However, most techniques are oriented toward single-image compensation,[2] for instance, using the red channel to estimate background illumination.[29] Therefore, no consistency between two images is guaranteed. For our case, this uneven illumination can be compensated by properly adjusting the intensity values on one image to approximately match that of the other while satisfying a predetermined illumination model. This can be carried out if the blurring is not too large and the illumination changes smoothly, which is usually the case for fundus images. This assumption can be expressed mathematically as

$$(k^{-1} \cdot u) * h \approx k^{-1}(u * h).$$

The illumination of the fundus is formed by a slowly varying light field over a smooth surface, thus it can be modeled by a low-order parametric surface. In Ref. 30 they used a fourth-order polynomial to effectively model the light pattern formed by an illumination source passing through the attenuating ocular media. Here, we use a similar approach, but fitting the surface with respect to both images. The parametric surface fitting equation can then be formulated as

$$\arg\min_{k} \|z_1(x, y) - k(x, y)z_2(x, y)\|, \tag{2}$$

where $k(x, y) = \alpha_{15}y^4 + \alpha_{14}y^3x + \cdots + \alpha_2 y + \alpha_1$, and z_1, z_2 are the registered fundus images. We minimize Eq. (2) in the least-squares sense to estimate the 15 parameters. This procedure can be both carried out using the luminance channel or the green channel as usual in retinal image processing.[31] Here, we have used the green channel. Owing to the fact that the illumination can be compensated globally by the polynomial function k, it is important to realize that the structural changes remain unaffected. The interpretation of k from Eq. (2) is straightforward. If the registered images z_1 and z_2 had neither illumination changes nor structural changes, then $k \approx 1$ throughout the common object field. In Fig. 4, we show the resulting $k(x, y)$ for the images in Fig. 2. The different shades of gray indicate the average contrast and intensity difference between the two images. From the image, it can be seen that most areas have similar intensity values except for the upper left part (dark region).

3.3 *Segmentation of Areas with Structural Changes*

The pathological region is actually a structural change and cannot be taken as a variation of illumination. Image change analysis is of interest in various fields and many algorithms have been developed for change detection.[32,33] A survey of change detection methods can be found in Ref. 34. An initial step in order to identify these changes comes from computing the difference from the two registered images including the illumination compensation as

$$\Delta z(x, y) = z_1(x, y) - k(x, y)z_2(x, y) \ . \tag{3}$$

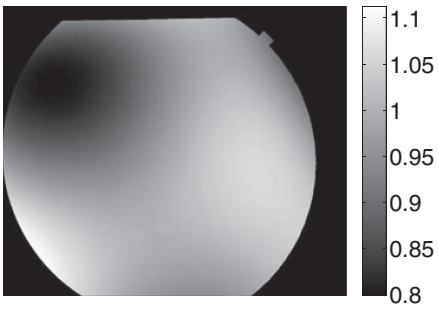


Fig. 4 Illumination compensation function $k(x, y)$.

The difference image is shown in absolute value Fig. 5(a). To better understand this result, in Fig. 5(b) we show one of the retinal images in gray scale, where the pixels related to structural changes are highlighted in pseudocolor. This image constitutes an important output of our algorithm. The structural changes can now be visualized and detected from the difference image $\Delta z(x, y)$ by taking a statistical significance test, in the same fashion as in Ref. 30. First, structural changes are often associated with a group of pixels; thus, the change decision at a given pixel j should be based on a small block of pixels in the neighborhood of j denoted as w_j. Second, in the absence of any change, the difference can be assumed to be due to noise alone. Therefore, the decision as to whether or not a change has occurred corresponds to choosing one of two competing hypothesis: the null hypothesis $\mathcal{H}_0$ or the alternative hypothesis $\mathcal{H}_1$, corresponding to no-change and change decisions, respectively. Assuming a Gaussian distribution for the difference values, the changes can be identified by comparing the normalized sum square of the differences within the neighborhood w_j to a predetermined threshold τ as described by Aach and Kaup.[32] The test is carried out as follows:

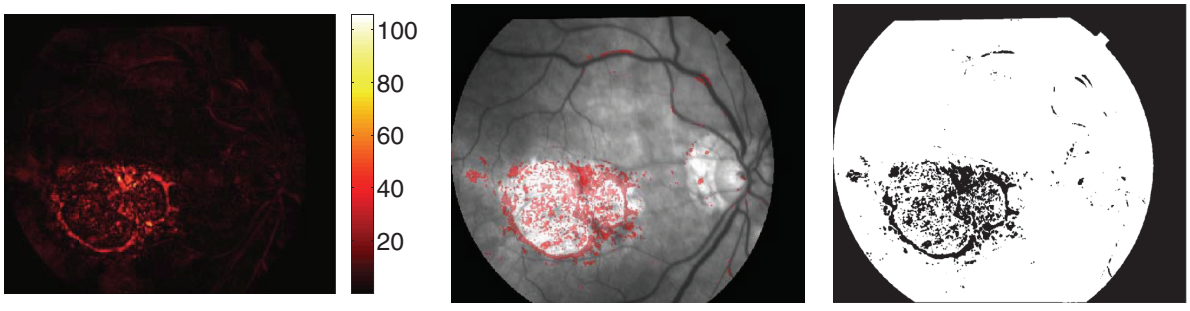
$$\Omega_j = \frac{1}{\sigma_n^2} \sum_{(x,y)\in w_j} \Delta z(x, y)^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau, \tag{4}$$

where σ_n is the noise standard deviation of the difference in the no-change regions. The threshold τ is derived from the fact that Ω_j follows a χ^2 distribution with N degrees of freedom, where N is the number of pixels in the window w_j. It can be obtained for a particular false-positive rate α from the χ^2 tables. The choice of an appropriate α is both guided by mathematical considerations (a 5% level for statistical significance is commonplace[35]) and the consequences that false alarms and misses might have. In this case, the effect of false alarms is unimportant because there would still be a large number of remaining pixels from where to compute the PSFs. On the other hand, misses do have a considerable impact in view of the fact that these pixels do not fulfill the convolutional model. As a result, α values of <0.05 might yield a more accurate change detection at the expense of possible undesirable misses. For all experiments, we use a 3×3 window ($N = 9$) and set $\alpha = 0.05$. The parameter σ_n was estimated by manually picking out no-change regions from a training set of images, computing Eq. (3) and the standard deviation inside these regions. Using Eq. (4) at each pixel, we can determine a change mask between the images or conversely a no-change mask. Given that, for the MBD procedure, we are interested in estimating the PSF from the no-change regions, the masking function m is obtained directly from the no-change mask of the significance test. The mask is shown in Fig 5(c). Note that the pathological region is the main cause of structural changes.

3.4 *Point-Spread Function Estimation*

In this section, we describe the basic principles of the blind deconvolution method used for the estimation of the PSFs. For this purpose, we have chosen one of the best working MBD methods.[21] MATLAB implementation of this method is available on the web of the authors.[36] The algorithm can be viewed as a Bayesian MAP estimation of the most probable sharp image and blur kernels. For our purposes, we used a modification of the original method that ignores regions affected by structural changes, which improves stability and precision of the computation. Without this modification, represented by the mask m in Eq. (5), the algorithm does not work reliably. The algorithm can be described as a minimization of the functional

$$\arg\min_{u,h_1,h_2} \left(\frac{1}{2}\|u * h_1 - z_1\|^2 + \frac{1}{2}\|m(u * h_2 - kz_2)\|^2 + \lambda_u \int |\nabla u| \, dx \, dy + \lambda_h \|m(z_1 * h_2 - kz_2 * h_1)\|^2 \right), \quad h_1, h_2 \geq 0, \tag{5}$$



(a) (b) (c)

Fig. 5 Intermediate outputs from the algorithm: (a) image difference $\Delta z(x, y)$ in absolute value, (b) image difference in pseudocolor on top of gray-scale fundus image, and (c) mask m for avoiding areas with structural changes.

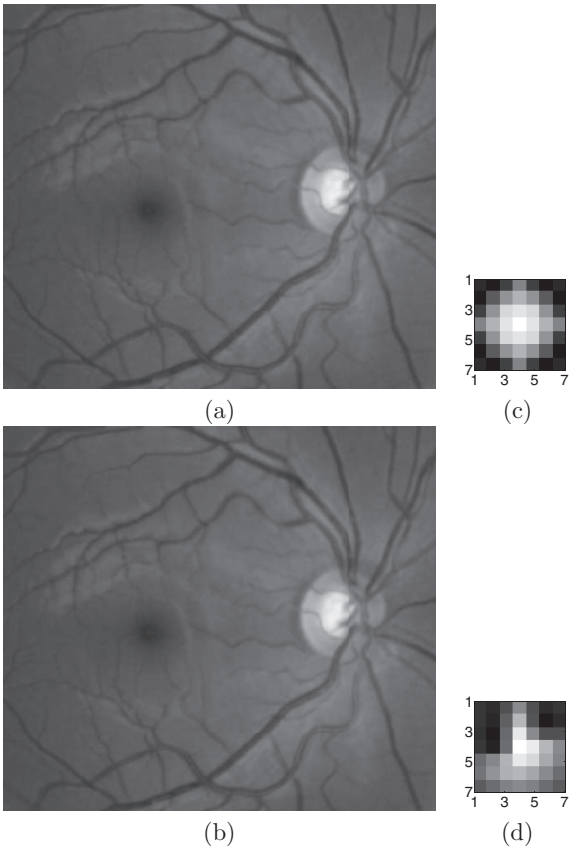


Fig. 6 (a, b) Degraded images (BSNR = 40 dB) and (c, d) PSFs.

with respect to the latent image u and blur kernels h_1 and h_2. The first and second terms measure the difference between the input blurred images and the searched image u blurred by kernels h_1 and h_2. The size of this difference is measured by L_2 norm $\|.\|$ and should be small for the correct solution; ideally, it should correspond to the noise variance in the given image. Function k compensates for uneven illumination as described in Sec. 3.2. The value of the masking function m is 1 in the valid points [white in Fig. 5(c)] and 0 in the pixels where the

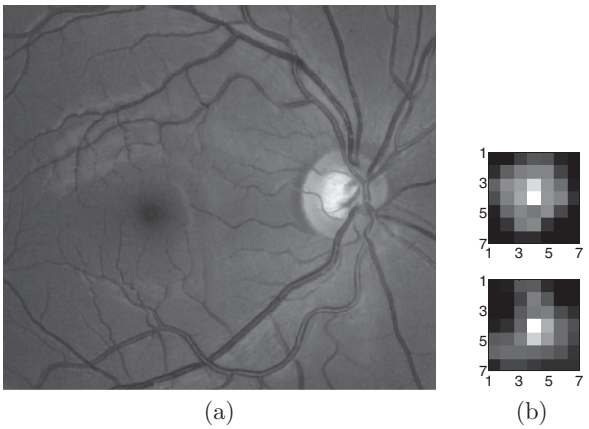


Fig. 7 (a) Restored image (ISNR = 4.45 dB) and (b) Estimated PSFs.

eye fundus has structurally changed. Any of the first two terms could be masked, but not both at the same time. This is because the latent image u cannot have pixels with no value at all; hence, these pixels must take values from any of the two images. In this case, z_2 is masked. As a result, these pixels take values from the first term. The two remaining terms are regularization terms with positive weighting constants λ_u and λ_h. The third term is nothing else than the total variation of image u. It improves stability of the minimization and from the statistical viewpoint incorporates prior knowledge about the solution. The last term is a condition linking the PSFs h_1 and h_2 of both images, which also improves the numerical stability of the minimization.

The functional is alternately minimized in the subspaces corresponding to the image and the PSFs. The advantage of this scheme lies in its simplicity, this alternating minimization approach is actually a variation of the steepest-descent algorithm. The minimization in the PSF subspace is equivalent to the solution of a system of linear equations in the least-squares sense with the non-negativity constraint, in our implementation solved by the MATLAB fmincon function. The nonblind deconvolution realized by the minimization in the image subspace, is solved by half-quadratic iterative scheme,[37] replacing the total variation by $\int \sqrt{|\nabla u|^2 + \epsilon^2}$, where ϵ is an auxiliary variable in the

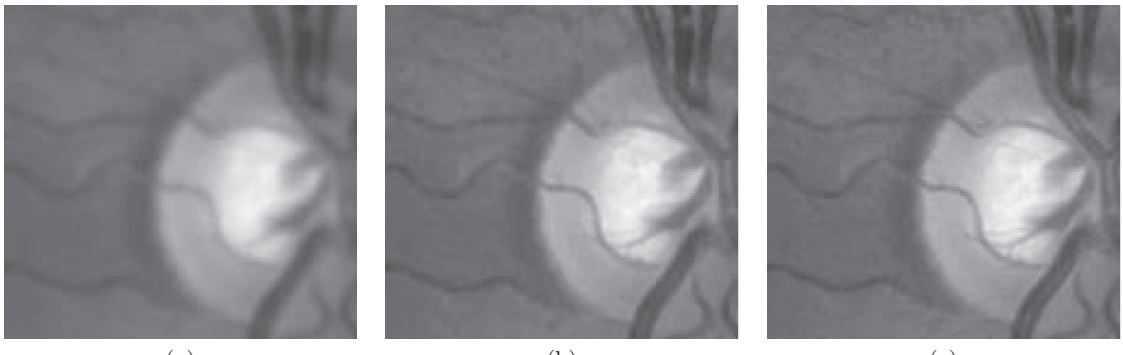
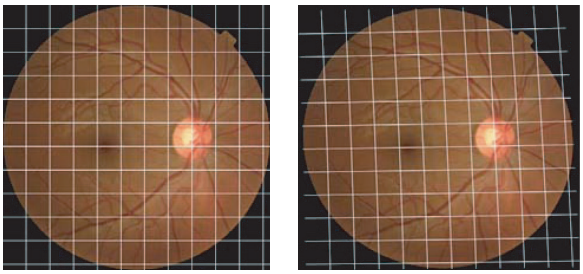


Fig. 8 Details from (a) degraded image, (b) restored image, and (c) original image.



(a) (b)

Fig. 9 (a) Original image and (b) geometrically distorted image.

range $0 < \epsilon \ll 1$. It is a small relaxation parameter that makes total variation differentiable around zero. A typical value for ϵ is 10^{-1}.

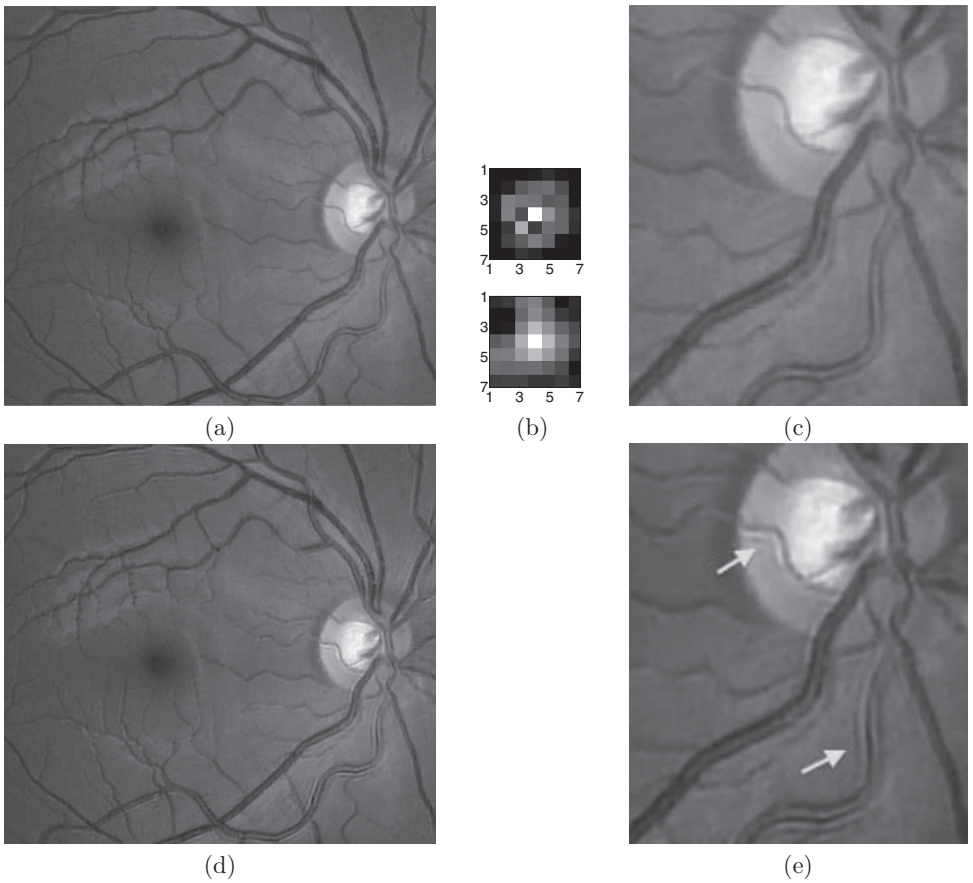
The main difference with respect to the original method[21] is the introduction of the masking function m, which is computed in the beginning of the algorithm as described in Sec. 3.3. During the minimization, the multiplication by m is included in the operator corresponding to the convolution with u (in the PSF minimization step) and in the operator corresponding to the convolution with h_2 (in the image minimization step). Because of the simplicity of this masking operation, the speed is practically the same as the speed of the original algorithm. In addition, even though we work with a complicated set of pixels, we can use the standard operation of convolution, which can eventually be speeded up using Fast Fourier transform (FFT).

3.5 *Image Restoration*

The aim of our algorithm is to restore both images as much as possible. Note that from Eq. (5) the restored version of z_1 ($\hat{u}_1$) is obtained because z_2 is masked; $\hat{u}_2$ could be obtained by minimizing Eq. (5) again with fixed PSFs and masking z_1. This procedure has the disadvantage that both images are restored only within the common object field. Therefore, an appropriate solution is to restore each image z_i via single-channel deconvolution with their corresponding PSF h_i (estimated from the previous step) by the minimization of the functional

$$\arg\min_{u_i} \left(\|u_i * h_i - z_i\|^2 + \lambda_u \int |\nabla u_i| \, \mathrm{d}x \, \mathrm{d}y \right). \tag{6}$$

This approach provides a further advantage in that the PSF estimation can be computed from a relatively small area of the common object field, provided that there are retinal structures



(a) (b) (c)

(d) (e)

Fig. 10 Image restoration from degraded and geometrically distorted images: (a) restored image by the proposed method (ISNR = 4.11 dB); (b) estimated PSFs; and (c) image detail, restored image by the method in Ref. 20 (ISNR = −0.72 dB); and (d) image detail.

within, thus greatly reducing the computational cost of the combined PSF estimation plus image deconvolution.

Finally, it should also be noted that the whole process of PSF estimation plus deconvolution can be computed for every channel of the RGB fundus image. However, in spite of the increase in computational burden, tests showed no real advantage to estimate the PSF for each channel. Moreover, the most suitable channel for PSF estimation is the green because it provides the best contrast. Whereas the blue channel encompasses the wavelengths most scattered and absorbed by the optical media of the eye; hence, the image has very low energy and a relatively high level of noise. As a result, the RGB deconvolved fundus image was computed by deconvolving every R, G, and B channel from the green channel PSF.

4 Experiments and Results

4.1 *Synthetic Images*

In this section, we use synthetically degraded retinal images to test the performance of the proposed method. We use blurred signal-to-noise ratio (BSNR) to measure the noise contained in the degraded image, and improvement in signal-to-noise ratio (ISNR) to measure the quality of restored images.[38] They are defined as follows:

$$\mathrm{BSNR} = 20\log_{10}\left(\frac{\|z\|}{\|n\|}\right),$$

$$\mathrm{ISNR} = 20\log_{10}\left(\frac{\|u-z\|}{\|u-\hat{u}\|}\right),$$

where u, z, $\hat{u}$, and n are the original image, degraded image, restored image, and noise vector, respectively. For ISNR, higher means better restoration; whereas for BSNR, lower means noisier degraded image. These metrics are mainly used to provide an objective standard for comparison to other techniques and they can only be used for simulated cases.

The first example is shown in Fig. 6, where the degraded images are synthesized from a sharp real image and the kernels shown in Fig. 6(c) and 6(d) plus Gaussian noise with zero mean and variance $\sigma^2 = 10^{-6}$ (BSNR=40 dB). The recovered image and PSFs are shown in Fig. 7. The restoration provides an ISNR=4.45 dB. In this case, for synthetically degraded images the masking operation of Sec. 3.3 was not applied. Visual inspection of the details shown in Fig. 8 clearly reveal the accuracy of the method. Under these circumstances, the algorithm is able to produce a significant restoration of fine details like small blood vessels around the optic disc.

To further test our approach under a more realistic degradation, we produced an initial geometrical distortion, via a quadratic model[26,28] as the one used for registration (Fig. 9). After the geometric distortion, the degradation (blur plus noise) is produced on both images (BSNR=40 dB). They are then registered, and the restored image is recovered via MBD. The restored image and the estimated PSFs are shown in Fig. 10. The ISNR is slightly less (4.11 dB) than in the previous case, but still sufficient to produce a significant restoration. To corroborate our assumption that MBD methods seem better suited for this type of images, we tried to restore the image with a recent SBD method proposed in Ref. 20. The result is shown in Fig. 10(e) and visually reveals that it does not follow the true nature of the blurring with artifacts around the blood vessels, thus being prone to produce a poor restoration evidenced by an ISNR$=-0.72$ dB.

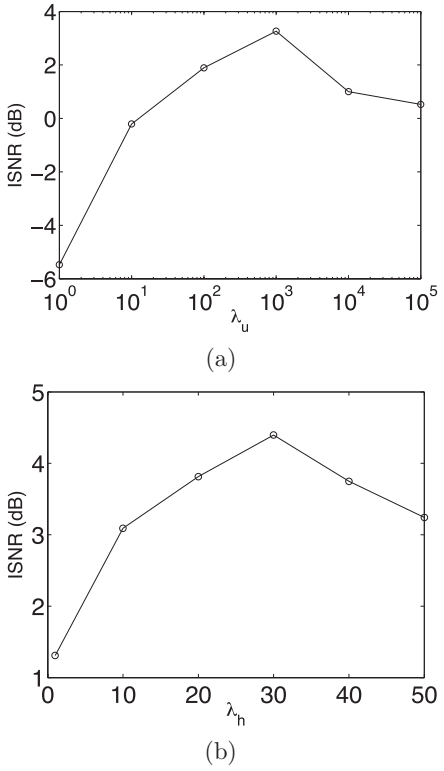


Fig. 11 Test on parameter setting (BSNR = 40 dB). Average ISNR with respect to different initial values of (a) λ_u and (b) λ_h.

Concerning parameter setting, in Fig. 11 we show the sensitivity of the two parameters λ_u and λ_h for the minimization of Eq. (5) in ISNR of the restored images. In Fig. 11(a), we fix the value of λ_h to 10 and check the ISNR of the restored images for different initial values of $\lambda_u = \{10^0, 10^1, 10^2, 10^3, 10^4, 10^5\}$. The best restoration is obtained with $\lambda_u = 10^3$; thus, in Fig. 11(b) we carried out the same procedure by fixing the value of λ_u to 10^3 and checking the ISNR of the restored image for different values of $\lambda_h = \{1, 10, 20, 30, 40, 50\}$. The best restoration was obtained with an initial value of $\lambda_h = 30$. For this type of image, when scaled to the interval $\langle 0, 1\rangle$, we find $20 < \lambda_h < 40$ to be a suitable range to produce an optimal restoration.

4.2 *Real Images*

The experiments shown in this section aim to demonstrate the applicability of the proposed method for retinal image deblurring in real scenarios. Three different cases are shown in Fig. 12, including the retinal images that were used to illustrate the method (Fig. 2). The estimated PSFs are shown at the bottom of the restored images. All images contain some pathological damage and have been acquired within considerable lapses of time (several months). In all examples, the resolution improvement can be visually assessed by the clear distinction of details, such as small blood vessels or the increase in sharpness

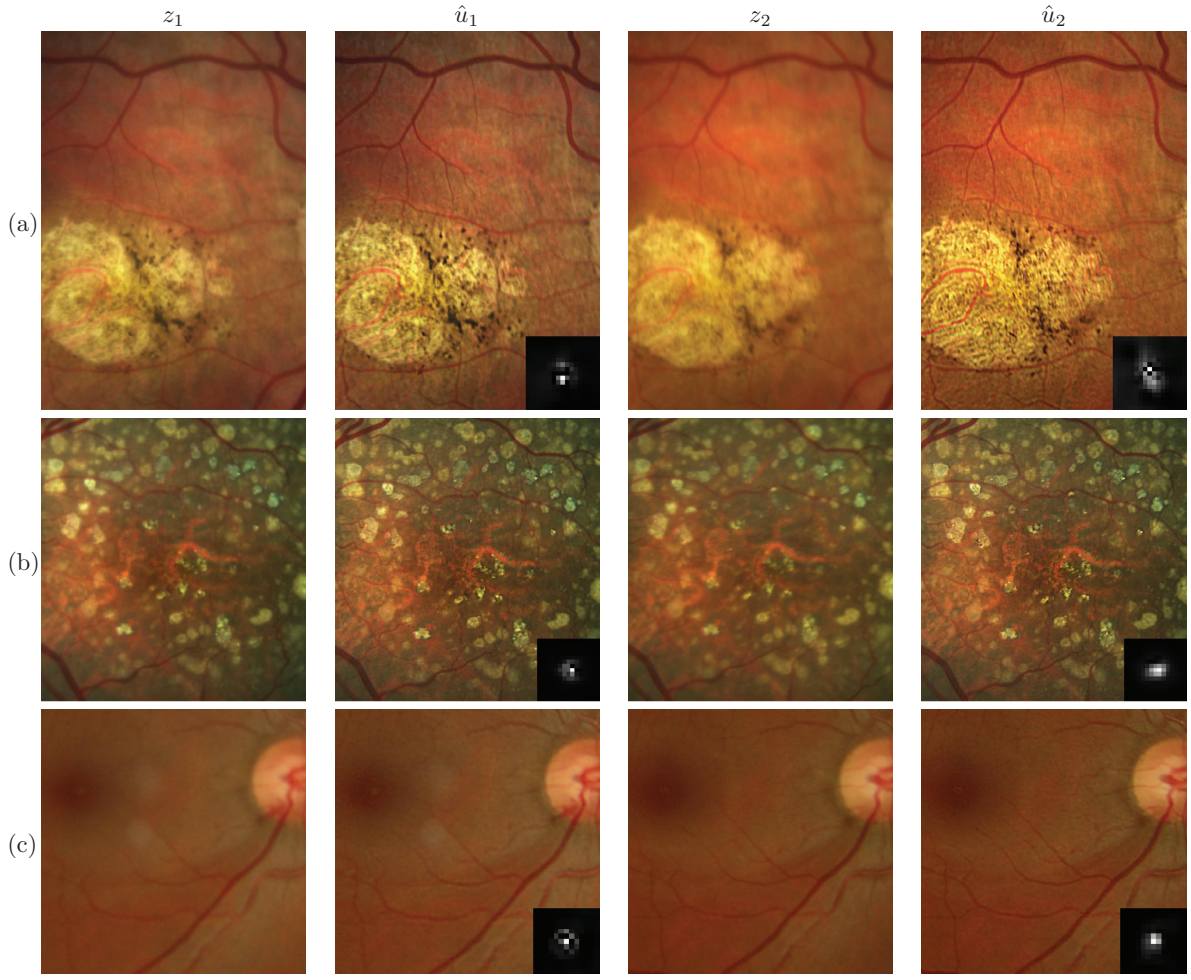


Fig. 12 Original and restored color retinal images; (a–c) indicate three separate cases arranged from left to right following our notation for degraded (z_i) and restored ($\hat{u}_i$) images. The images are cropped to represent the region of interest given by the pathological area. The estimated PSF is shown at the bottom of the restored image. Video files are also included for change detection in cases (a) and (b). (Video 1, Quicktime, 0.5 MB) [URL: http://dx.doi.org/10.1117/1.3652709.1]; (Video 2, Quicktime, 0.4 MB) [URL: http://dx.doi.org/10.1117/1.3652709.2]

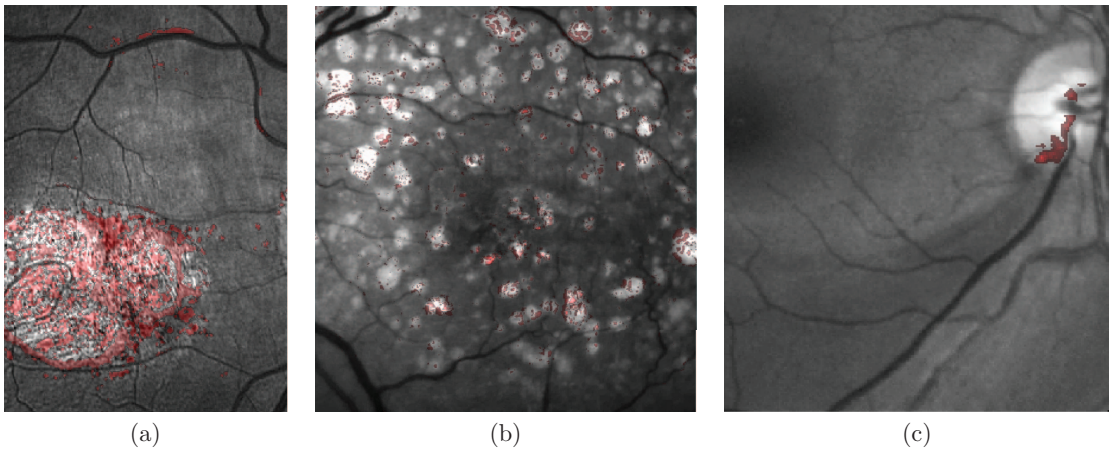


Fig. 13 Visualization of structural changes in pseudo-color for the images of Fig. 12.

of edges, especially in the pathological areas. We emphasize the fact that these images correspond to real routine patient follow-up and were not intentionally degraded. From a clinical viewpoint, the enhancement can be used for a more precise assessment of a patient's state. Likewise, the images are more suitable for subsequent processing such as for the detection of retinal pathology.[29,39]

In Fig. 13, the same images are shown but in gray scale to highlight the areas of structural change in pseudocolor. As mentioned earlier, this is an important result for its potential impact in medical practice. Subtle changes can be identified by this approach, such as the ones in Fig. 13(b) and the hemorrhage in the region of the optic disk in Fig. 13(c). Another technique to rapidly identify changes from the two images is by alternating both restored images in a video sequence. Videos 1 and 2 (Fig. 12) correspond to the first two real cases.

5 Conclusion

The main purpose of this paper has been to investigate a new approach for retinal image restoration based on multichannel blind deconvolution. In addition, we developed a strategy for identifying and highlighting areas of structural change with possible relation to pathological damage. We have verified that fundus images of the same retina over time contain enough common information to be restored with the proposed method. The method consists of a series of preprocessing steps to adjust the images so they comply with the convolutional model, followed by the final stages of PSF estimation and deconvolution. The synthetically degraded images enabled us to test the performance of the proposed approach and also to compare with a state-of-the-art single-channel blind deconvolution method. Results showed a remarkable enhancement evidenced by the increased visibility of details such as small blood vessels or pathological areas. The proposed method provides a novel practical approach for retinal image enhancement and, equally important the analysis of retinal changes over time. Central to the task of determining disease progression is the distinction of true change from variability.

The results of this study open several new avenues for research and applications. A possible application is found in the restoration of stereo retinal images for depth estimation. Most stereo images do not satisfy the brightness constancy assumption along with the expected blurring of some parts of the images because photographers find it difficult to focus two images simultaneously. Finally, research can also be conducted to compare to deconvolution from wavefront-sensing fundus imagers to determine if our method could be a suitable and inexpensive alternative.

Acknowledgments

This research has been partly funded by the Spanish Ministerio de Ciencia e Innovación y Fondos FEDER (Project No. DPI2009-08879). Financial support was also provided by the Czech Ministry of Education under the Project No. 1M0572 (Research Center DAR). The authors are also grateful to the ophthalmologist Jordi Monés, M. D., from the Institut de la Màcula i la Retina, Barcelona, for providing the images. The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

References

1. J. Kanski, *Diseases of the Ocular Fundus*, Elsevier/Mosby, New York (2005).
2. R. Winder, P. Morrow, I. McRitchie, J. Bailie, and P. Hart, "Algorithms for digital image processing in diabetic retinopathy," *Comput. Med. Imag. Graphics* **33**(8), 608–622 (2009).
3. H. Bartling, P. Wanger, and L. Martin, "Automated quality evaluation of digital fundus photographs," *Acta Ophthalmol.* **87**(6), 643–647 (2009).
4. A. Agrawal and M. A. McKibbin, "Technical failure in photographic screening for diabetic retinopathy," *Diabet. Med.* **20**(9), 777–780 (2003).
5. M. C. Boucher, J. A. Gresset, K. Angioi, and S. Olivier, "Effectiveness and safety of screening for diabetic retinopathy with two nonmydriatic digital images compared with the seven standard stereoscopic photographic fields," *Can. J. Ophthalmol.* **38**(7), 557–568 (2003).
6. H. M. Herbert, K. Jordan, and D. W. Flanagan, "Is screening with digital imaging using one retinal view adequate?" *Eye* **17**(4), 497–500 (2003).
7. M. D. Abràmoff, M. Niemeijer, M. S. A. Suttorp-Schulten, M. A. Viergever, S. R. Russell, and B. van Ginneken, "Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes," *Diabetes Care* **31**(2), 193–198 (2008).
8. A. V. Larichev, N. G. Irochnikov, P. Ivanov, and A. V. Kudryashov, "Deconvolution of color retinal images with wavefront sensing," *Proc. SPIE* **4251**, 102–116 (2001).
9. J. C. Christou, A. Roorda, and D. R. Williams, "Deconvolution of adaptive optics retinal images," *J. Opt. Soc. Am. A* **21**, 1393–1401 (2004).
10. J. Primot, G. Rousset, and J. C. Fontanella, "Deconvolution from wavefront sensing: a new technique for compensating turbulence-degraded images," *J. Opt. Soc. Am. A* **7**, 1598–1608 (1990).
11. D. Catlin and C. Dainty, "High-resolution imaging of the human retina with a Fourier deconvolution technique," *J. Opt. Soc. Am. A* **19**, 1515–23 (2002).
12. W. Wang, M. C.W. Campbell, M. L. Kisilak, and S. R. Boyd, "Optical design considerations when imaging the fundus with an adaptive optics correction," *Proc. SPIE* **7099**, 70990C (2008).
13. S. A. Burns, R. Tumbar, A. E. Elsner, D. Ferguson, and D. X. Hammer, "Large-field-of-view, modular, stabilized, adaptive-optics-based scanning laser ophthalmoscope," *J. Opt. Soc. Am. A* **24**, 1313–1326 (2007).
14. P. Bedggood, M. Daaboul, R. Ashman, G. Smith, and A. Metha, "Characteristics of the human isoplanatic patch and implications for adaptive optics retinal imaging," *J. Biomed. Opt.* **13**, 024008 (2008).
15. M. S. Millán, E. Pérez-Cabré, and B. Javidi, "Multifactor authentication reinforces optical security," *Opt. Lett.* **31**(6), 721–723 (2006).
16. D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.* **13**(3), 43–64 (1996).
17. E. P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*, CRC Press, Boca Raton (1996).
18. C. Sheppard, "Fundamentals of superresolution," *Micron* **38**(2), 165–169 (2007).
19. A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1964–1971 (2009).
20. L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," *Lect. Notes Comp. Sci.* **6311**, 157–170 (2010).
21. F. Šroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Process.* **14**(7), 874–883 (2005).
22. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, 259–268 (1992).
23. W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.* **62**(1), 55–59 (1972).
24. F. Šroubek, J. Flusser, and M. Šorel, "Superresolution and blind deconvolution of video," in *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1–4 (2008).
25. B. Zitová and J. Flusser, "Image registration methods: a survey," *Image Vis. Comput.* **11**(21), 977–1000 (2003).

26. S. Lee, J. Reinhardt, P. Cattin, and M. Abramoff, "Objective and expert-independent validation of retinal image registration algorithms by a projective imaging distortion model," *Med. Image Anal.* **14**(4), 539–549 (2010).
27. C. Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE Trans. Med. Imaging* **22**(11), 1379–1394 (2003).
28. A. Can, C. Stewart, B. Roysam, and H. Tanenbaum, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Trans. Pattern Anal.* **24**(3), 347–364 (2002).
29. C. Muramatsu, Y. Hayashi, A. Sawada, Y. Hatanaka, T. Hara, T. Yamamoto, and H. Fujita, "Detection of retinal nerve fiber layer defects on retinal fundus images for early diagnosis of glaucoma," *J. Biomed. Opt.* **15**, 016021 (2010).
30. H. Narasimha-Iyer, A. Can, B. Roysam, C. Stewart, H. Tanenbaum, A. Majerovics, and H. Singh, "Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy," *IEEE Trans. Biomed. Eng.* **53**(6), 1084–1098 (2006).
31. M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med. Image Anal.* **9**(3), 179–190 (2005).
32. T. Aach and A. Kaup, "Bayesian algorithms for change detection in image sequences using markov random fields," *Signal Process. Image Commun.* **7**, 147–160 (1995).
33. C.-C. Chang, T.-L. Chia, and C.-K. Yang, "Modified temporal difference method for change detection," *Opt. Eng.* **44**, 027001 (2005).
34. R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Trans. Image Process.* **14**(3), 294–307 (2005).
35. S. Stigler, "Fisher and the 5% level," *Chance* **21**, 12 (2008).
36. MATLAB application for multichannel blind deconvolution, http://zoi.utia.cas.cz/download (2008).
37. A. Chambolle and P. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.* **76**(2), 167–188 (1997).
38. Liyakathunisa and V. K. Ananthashayana, "Super resolution blind reconstruction of low resolution images using wavelets based fusion," *World Acad. Sci. Eng. Technol.* **40**, 177–181 (2008).
39. L. Xu and S. Luo, "Optimal algorithm for automatic detection of microaneurysms based on receiver operating characteristic curve," *J. Biomed. Opt.* **15**, 065004 (2010).

Journal of Biomedical Optics 19(1), 016023 (January 2014)

Restoration of retinal images with space-variant blur

Andrés G. Marrugo,[a,b,*] María S. Millán,[a] Michal Šorel,[c] and Filip Šroubek[c]
[a]Universitat Politècnica de Catalunya, Department of Optics and Optometry, Group of Applied Optics and Image Processing, Violinista Vellsolà 37, 08222 Terrassa, Spain
[b]Universidad Tecnológica de Bolívar, Facultad de Ciencias Básicas, Km 1 vía Turbaco, Cartagena, Colombia
[c]Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, 18208 Prague 8, Czech Republic

Abstract. Retinal images are essential clinical resources for the diagnosis of retinopathy and many other ocular diseases. Because of improper acquisition conditions or inherent optical aberrations in the eye, the images are often degraded with blur. In many common cases, the blur varies across the field of view. Most image deblurring algorithms assume a space-invariant blur, which fails in the presence of space-variant (SV) blur. In this work, we propose an innovative strategy for the restoration of retinal images in which we consider the blur to be both unknown and SV. We model the blur by a linear operation interpreted as a convolution with a point-spread function (PSF) that changes with the position in the image. To achieve an artifact-free restoration, we propose a framework for a robust estimation of the SV PSF based on an eye-domain knowledge strategy. The restoration method was tested on artificially and naturally degraded retinal images. The results show an important enhancement, significant enough to leverage the images' clinical use.
[DOI: 10.1117/1.JBO.19.1.016023]

Keywords: blind deconvolution; space-variant restoration; image restoration; deblurring; retinal image.

Paper 130016RR received Jan. 11, 2013; revised manuscript received Dec. 23, 2013; accepted for publication Dec. 30, 2013; published online Jan. 27, 2014.

1 Introduction

Blur is one of the main image quality degradations in eye fundus imaging, which along with other factors such as nonuniform illumination or scattering, hinder the clinical use of the images. Its main causes are inherent optical aberrations in the eye, relative camera-eye motion, and improper focusing. Eye motion is related to the patient inability to steady fixate a target in the fundus camera. Many patients have difficulty in fixating, like many elderly patients or those that suffer from amblyopia.[1] Because the optics of the eye is part of the optical imaging system, the aberrations of the eye are a common source of image quality degradation. To overcome this limitation, adaptive optics techniques have been successfully applied to correct the aberrations, thus producing high resolution images.[2] However, most commercial fundus cameras compensate for spherical refractive errors, but not for astigmatism[3]—let alone higher-order aberrations. In general, the aberrations of the eye have a stronger impact in image degradation than the aberrations introduced by the rest of the optical system, i.e., the retinal camera. Besides, even though it is possible to measure the optical quality of the camera, it would be exceptional to have readily available additional information related to the optical quality of the patient's eye. The described scenario is commonplace in the clinical setting, for which we assume the same conditions here. This brings about the need for a restoration procedure that accounts for the lack of information related to the origin of the image degradation.

The technique for recovering an original or unblurred image from a single or a set of blurred images in the presence of a poorly determined or unknown point-spread function (PSF) is called blind deconvolution. Removing blur from a single blurred image is an ill-posed problem as there are more unknowns (image and blur) than equations. Having more than one image of the same scene better poses the problem. In retinal imaging, it is not difficult to obtain a second image from the same eye, with the convenience that acquisition conditions remain quite similar. In fact, in Ref. 4, we took advantage of this condition and proposed a blind deconvolution method to restore blurred retinal images acquired with a lapse of time, even in the case where structural changes had occurred in the images. In that work, we detected structural changes, which in turn have clinical relevance, and applied a masking operator so that the images would comply with the considered degradation model. This enabled the successful restoration of many degraded retinal images coming from patient follow-up visits. However, the method is limited to images blurred uniformly; in other words, we assumed the blur to be space-invariant. In Sec. 4.2, we show an attempt at restoring an image degraded with spatially variant blur with this approach. The space-invariant assumption is commonplace in most of the restoration methods reported in the literature,[5] but in reality it is a known fact that blur changes throughout the image.[6] In this work, we consider the blur to be both unknown and space-variant (SV). This in itself is a novel approach in retinal imaging; relevant to such extent that many common eye related conditions, such as astigmatism, keratoconus, corneal refractive surgery, or even tear break-up, may contribute significantly to a decline in image quality[7,8] typically in the form of an SV degradation. An example of such a condition is shown in Fig. 1(a). The image corresponds to an eye from a patient with corneal abnormalities that lead to a loss in visual acuity and a quality degradation of the retinal image [Fig. 1(b)].

Restoration of images with SV blur from optical aberrations has been reported in the literature,[9] although the main limitation

*Address all correspondence to: Andrés G. Marrugo, E-mail: agmarrugo@utbvirtual.edu.co

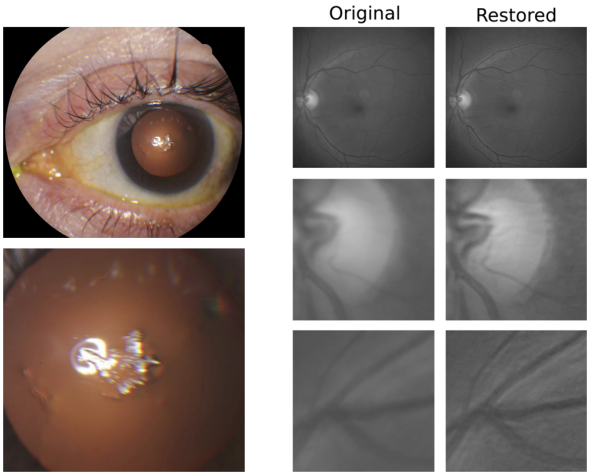


Fig. 1 (a) Top: Eye with corneal defects that induce retinal images with space-variant (SV) degradation. Bottom: zoomed region. (b) Left column: original image and details. Right column: restored image with proposed approach and details.

is that the blurred image is often restored in regions or patches, which are then stitched together. This inevitably leads to ringing artifacts associated with frequency-domain filtering like in Wiener filtering. Another clear disadvantage is a significant complexity for accurately estimating the SV PSF, for instance Bardsley et al.[10] use a phase-diversity based scheme to obtain the PSF associated with an image patch. This type of approach is common in atmospheric optics where the conditions and setup of the imaging apparatus (typically a telescope) are well known and calibrated. Unfortunately, this is not immediately applicable to retinal imaging, at least nonadaptive optics retinal imaging. Recently, there have been several works[11–13] that try to solve the SV blind deconvolution problem from a single image. The common ground in these works is that the authors assume that the blur is only due to camera motion. They do this in order to reduce the space in which to search for SV blurs. Despite their approach being more general, the strong assumption of camera motion is simply too restrictive to be applied in the retinal imaging scenario.

1.1 *Contribution*

In this work, we propose a method for removing blur from retinal images. We consider images degraded with SV blur, which may be due to factors like aberrations in the eye or relative camera-eye motion. Because restoring a single blurred image is an ill-posed problem, we make use of two blurred retinal images from the same eye fundus to accurately estimate the SV PSF. Before the PSF estimation and restoration stages take place, we preprocess the images to accurately register them and compensate for illumination variations not caused by blur, but by the lighting system of the fundus camera. This is depicted in the block diagram shown in Fig. 2. The individual stages of the method are explained in Sec. 3.

We assume that in small image patches, the SV blur can be approximated by a spatially invariant PSF. In other words, that in a small region, the wavefront aberrations remain relatively constant; the so-called isoplanatic patch.[6] An important aspect of our approach is that instead of deblurring each patch with its corresponding space-invariant PSF—and later stitching together the results—we sew the individual PSFs by interpolation and restore the image globally. This is intended to reduce some artifacts that otherwise would likely appear at the seams of the restored patches. The estimation of the local space-invariant PSFs may fail in patches with hardly any structural information (e.g., such as blood vessels). These poorly estimated or nonvalid PSFs introduce artifacts in the restored image. Detecting such artifacts and inferring the nonvalid PSFs is a difficult problem. Recently, Tallón et al.[14] developed a strategy for detecting these patches in an SV deconvolution and denoising algorithm from a pair of images acquired with different exposures: a sharp noisy image with a short exposure and a blurry image with a long exposure. Because they had two distinct input images that were able to: (i) Identify patches where the blur estimates were poor based on a comparison (via a thresholding operation) of the deconvolved patches with the sharp noisy patches. (ii) In those patches, instead of correcting the local PSFs and deconvolving the patches again, they performed denoising in the noisy sharp image patch. The end result is a patchwork-image of deconvolved patches stitched together with denoised patches. Their method is mainly oriented at motion blur, this is the reason for a dual exposure strategy. This is not readily implementable in the retinal imaging scenario where the SV blur is generally caused by factors like aberrations, including those belonging to the patient's eye optical system, the eye fundus shape, and the retinal camera. In this paper, we address the question "how to identify PSF estimation failure to improve the SV deconvolution of retinal images?" Retinal imaging provides a constrained imaging scenario from which we can formulate a restoration approach that incorporates prior knowledge of blur through the optics of the eye. The novelty in our approach is in the strategy based on eye-domain knowledge for identifying the nonvalid local PSFs and replacing them with appropriate ones. Even though methods for processing retinal images in a space-dependent way (like locally adaptive filtering techniques[15,16])

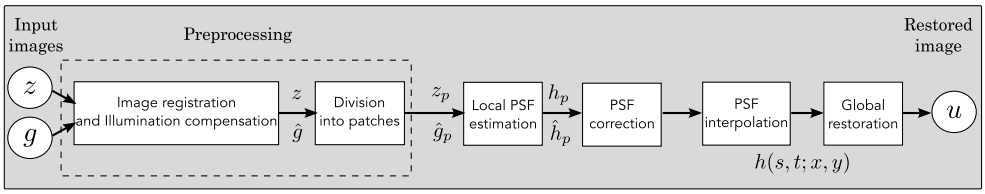


Fig. 2 Block diagram illustrating the proposed method. *z* is the degraded image, *g* is an auxiliary image of the same eye fundus used for the point-spread function (PSF) estimation, and *u* is the restored image. The other variables are intermediate outputs of every stage; their meaning is given in the text.

have been proposed in the literature; to the best of our knowledge, this is the first time a method for SV deblurring of retinal images is proposed.

2 SV Model of Blur

In our previous work,[4] we modeled the blurring of a retinal image by convolution with a unique global PSF. This approximation is valid as long as the PSF changes little throughout the field of view (FOV). In other words, that the blurring is homogenous. In reality, we know that the PSF is indeed spatially variant,[6] to such an extent that in some cases the space-invariant approach completely fails, bringing forth the need for an SV approach. To address this limitation, in this work, we model the blurred retinal image z by the linear operation

$$z(x,y) = [Hu](x,y) = \int u(s,t)h(s,t,x-s,y-t)\mathrm{d}s\mathrm{d}t, \tag{1}$$

where u is the unblurred retinal image and h is the SV PSF. The operator H is a generalization of standard convolution where h is now a function of four variables. We can think of this operation as a convolution with a PSF $h(s,t,x,y)$ that is now dependent on the position (x,y) in the image. Standard convolution is a special case of Eq. (1), where $h(s,t,x,y) = h(s,t)$ for an arbitrary position (x,y). Note that the PSF h is a general construct that can represent other complex image degradations which depend on spatial coordinates, such as motion blur, optical aberrations, lens distortions, and out-of-focus blur.

2.1 *Representation of SV PSF*

An obvious problem of spatially varying blur is that the PSF is now a function of four variables. Except trivial cases, it is hard to express it by an explicit formula. Even if the PSF is known, we must solve the problem of a computationally efficient representation.

In practice, we work with a discrete representation, where the same notation can be used but with the following differences: the PSF h is defined on a discrete set of coordinates, the integral sign in Eq. (1) becomes a sum, operator H corresponds to a sparse matrix and u to a vector obtained by stacking the columns of the image into one long vector. For example in the case of standard convolution, H is a block-Toeplitz matrix with Toeplitz blocks and each column of H corresponds to the same kernel $h(s,t)$.[17] In the SV case that we address here, as each column of H corresponds to a different position (x,y), it may contain a different kernel $h(s,t,x,y)$.

In retinal imaging, all typical causes of blur change in a continuous gradual way,[18] which is why we assume the blur to be locally constant. Therefore, we can make the approximation that locally the PSFs are space-invariant. By taking advantage of this property, we do not have to estimate local PSFs for every pixel. Instead, we divide the image into rectangular windows and estimate only a small set of local PSFs [see Fig. 3(a)] following the method described in Ref. 4 and outlined in Sec. 3. The estimated PSFs are assigned to the centers of the windows from where they were computed. In the rest of the image, the PSF h is approximated by bilinear interpolation from the four adjacent local PSFs. This procedure is explained in further detail in the following section.

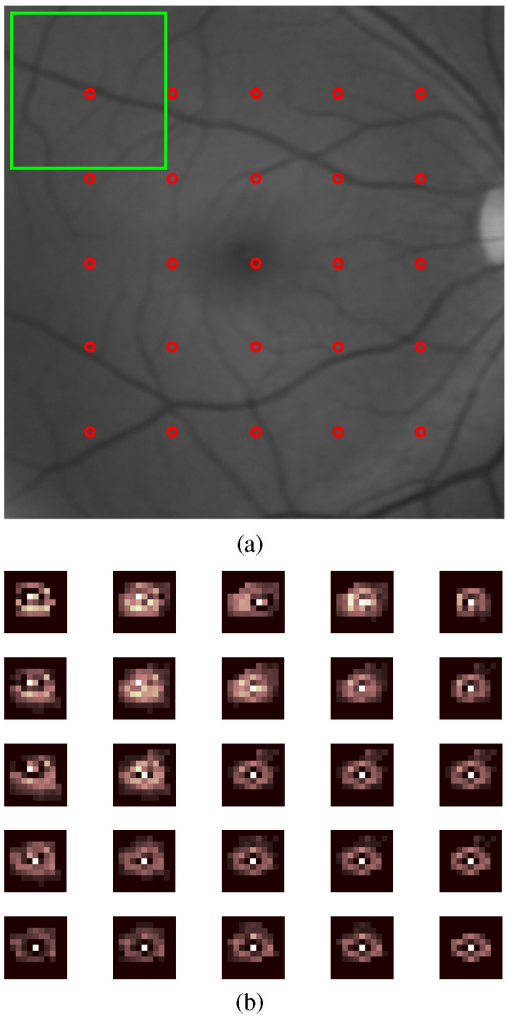


Fig. 3 (a) Retinal image degraded with artificial SV blur given by (b) grid of PSFs. The grid and the image patches shown in (a) are used for local PSF estimation.

3 Description of the Method

In this section, we describe the different stages of the proposed restoration method shown in Fig. 2. This paper follows Ref. 4 but addresses a more general problem: restoration of retinal images in the presence of an SV PSF. In Ref. 4, we showed that the single image blind deconvolution for blurred retinal images does not provide a suitable restoration. Moreover, in images with SV blur, the restoration is even worse. Alternatively, by taking two images of the same retina we can use a multichannel blind deconvolution strategy that is mathematically better-posed.[19]

In this paper, the estimation of the SV PSF is carried out via local multichannel deconvolution. To illustrate the method and to study its dependence on its tunable parameters, we use an original real image of the retina and obtain two artificially degraded versions from it, denoted by z and g. Figure 3(a) contains image z. The degraded images z and g have been obtained by blurring the original image with an SV PSF represented by the grid of local PSFs shown in Fig. 3(b) and adding Gaussian

zero-mean noise ($\sigma = 10^{-6}$). The PSF grid was built with realistic PSFs estimated from real blurred retinal images using the method of Ref. 4.

3.1 *Preprocessing*

Because we use a multichannel scheme for the estimation of the local PSFs, the images are preprocessed so that they meet the requirements imposed by the space-invariant convolutional model given by Eq. (2). This consists in registering the images and adjusting their illumination distribution following Ref. 4. By carrying out this procedure, the remaining radiometric differences between the images are assumed to be caused by blur and noise. Unlike the case considered in Ref. 4, where a relatively long lapse of time between the two image acquisitions may possibly involve a potential structural change, in this study, both the images z and g are originated from the same image or, in practice, they are acquired one shortly after the other. Therefore, no structural change is expected. Since image g is registered and its illumination matched to z, we denote this transformed auxiliary image as $\hat{g}$.

3.2 *Estimation of the Local PSFs*

In Sec. 2, we described the model for a spatially varying blur in which we assume the PSF h to vary gradually, which means that within small regions the blur can be locally approximated by convolution with a space-invariant PSF. For this reason, we approximate the global function h from Eq. (1) by interpolating local PSFs estimated on a set of discrete positions. The main advantage of this approach is that the global PSF needs not be computed on a perpixel basis which is inherently time-consuming.

The procedure for estimating the local PSFs is the following. We divide the images z and $\hat{g}$ with a grid of $m \times m$ patches [Fig. 3(a)]. In each patch p, we assume a convolutional blurring model where an ideal sharp patch u_p originates from two degraded patches z_p and $\hat{g}_p$ (for $p = 1, \ldots, m \times m$). The local blurring model is

$$\begin{aligned} z_p &= h_p * u_p + n, \\ \hat{g}_p &= \hat{h}_p * u_p + \hat{n}, \end{aligned} \tag{2}$$

where $*$ is the standard convolution and h_p and $\hat{h}_p$ are the convolution kernels or local PSFs. The noise (n and $\hat{n}$) is assumed to have a constant spectral density and a zero-mean Gaussian distribution of amplitude. Despite the fact that this may not be the most accurate representation of the noise, because the retinal images considered here are acquired by illuminating with a flash, the resulting signal-to-noise ratio is high enough that in the estimation of the PSFs, the impact of noise is not significant.

From this model, we can estimate the local PSFs with an alternating minimization procedure as described in Ref. 4 but applied locally. The general guideline is that the patch size should be large enough to include retinal structures and much larger than the size of the local PSF. In Sec. 4, we show further analysis on the robustness of the method to these parameters. Every local PSF is computed on each patch p by minimizing the functional

$$\begin{aligned} \arg \min_{u_p, h_p, \hat{h}_p} \Bigg(& \frac{1}{2}\|u_p * h_p - z_p\|^2 + \frac{1}{2}\|u_p * \hat{h}_p - \hat{g}_p\|^2 \\ & + \lambda_1 \int |\nabla u_p| \mathrm{d}x\mathrm{d}y + \lambda_2 \|z_p * \hat{h}_p - \hat{g}_p * h_p\|^2 \Bigg), h_p, \hat{h}_p(s,t) \geq 0, \end{aligned} \tag{3}$$

with respect to the ideal sharp patch u_p and the blur kernels h_p and $\hat{h}_p$. The blur kernel $h_p(s,t)$ is an estimate of $h(s,t,x_0,y_0)$, where (x_0, y_0) is the center of the current window z_p, and $\|.\|$ is the ℓ_2 norm. The first and second terms of Eq. (3) measure the difference between the input blurred patches (z_p and $\hat{g}_p$) and the sharp patch u_p blurred by kernels h_p and $\hat{h}_p$. This difference should be small for the correct solution. Ideally, it should correspond to the noise variance in the image. Although u_p is a restored patch, note that it is not used by our method, but discarded. This is because our method does not work by performing local deconvolutions and sewing restored patches together, which in practice would produce artifacts on the seams. Instead, we perform a global restoration method explained in Sec. 3.5. The two remaining terms of Eq. (3) are regularization terms with positive weighting constants λ_1 and λ_2, which we have set following the fine-tuning procedure described in Ref. 4. The tuning procedure consists of an optimization process where an artificially degraded retinal image is restored by varying λ and measuring a restoration error. This way an optimal λ is obtained. Typical values are $\lambda_1 = \sim 10^3$ and $\lambda_2 = \sim 10^1$. The third term is the total variation of u_p. It improves stability of the minimization and from a statistical perspective, it incorporates prior knowledge about the solution. The last term is a condition linking the convolution kernels which also improves the numerical stability of the minimization. The functional is alternately minimized in the subspaces corresponding to the images and the PSFs. The estimated PSFs for the artificially degraded [Fig. 3(a)] image are shown in Fig. 6(a).

3.3 *Identifying and Correcting Nonvalid PSFs*

3.3.1 *Strategy based on eye-domain knowledge*

The local PSF estimation procedure does not always succeeds. Consequently, such nonvalid PSFs must be identified, removed, and replaced. In our case, we replace them by an average of adjacent valid kernels. The main reason why the PSF estimation may fail is due to the existence of textureless or nearly homogenous regions bereft of structures with edges (e.g., blood vessels) to provide sufficient information.[14] To identify these nonvalid PSFs, we devised an eye-domain knowledge strategy. The incorporation of proper *a priori* assumptions and domain knowledge about the blur into the method provides an effective mechanism for a successful identification of poorly estimated PSFs.

The optics of the eye is part of the imaging system, therefore it is reasonable to assume that the PSF of the imaging system is determined by the PSF of the eye. The retinal camera can indeed be close to diffraction limited with a very narrow PSF, but the optics of the eye is governed by optical aberrations that change across the visual field[18] that lead to an SV PSF. The typical PSFs of the human eye, as reported in the literature,[18,20] display distinct shapes in many cases displaying long tails' evidence of the inherent optical aberrations. What is common to all PSFs of the human eye is that the energy is spread from a central lobe and

decreases outwardly. In this paper, we assume that PSFs that do not follow this general pattern are nonvalid PSFs which correspond to patches where the estimation failed.

In order to prove this, we designed an artificial experiment where we compare the estimated PSFs with the ground-truth PSFs by using a kernel similarity measure S proposed by Hu and Yang.[21] The measure is based on the peak of the normalized cross correlation between the two PSFs. The authors showed in the paper that this measure is more accurate than the root mean square error, especially because it is shift invariant. The measure is defined as the blur kernel similarity $S(h,\tilde{h})$ of two kernels, h and $\tilde{h}$,

$$S(h,\tilde{h}) = \max_{\gamma} \rho(h,\tilde{h},\gamma), \tag{4}$$

where $\rho(.)$ is the normalized cross-correlation function and γ is the possible shift between the two kernels. Let τ represent element coordinates, $\rho(.)$ is given by

$$\rho(h_p,\tilde{h}_p,\gamma) = \frac{\sum_{\tau} h_p(\tau)\cdot \tilde{h}_p(\tau+\gamma)}{\|h_p\|\cdot\|\tilde{h}_p\|}, \tag{5}$$

where $\|\cdot\|$ is the ℓ_2-norm. Larger similarity values reflect more accurate PSF estimation, thus better image restoration. The graphical representation of the similarity measure S is shown in Fig. 4(a). The darkest squares correspond to PSFs with the lowest similarity score.

The way we determine a nonvalid PSF is by characterizing the energy distribution along the local PSF space. We add the PSF values along concentric squares of radius r to build an energy distribution histogram $f(r)$, which is normalized to sum to 1. In Fig. 4(b), we show the histograms for the energy distribution characterization of the estimated PSFs from the artificially degraded retinal image [Fig. 3(a)]. To identify nonvalid PSFs, i.e., PSFs that do not follow the pattern we have previously described, we compute a shape descriptor for $f(r)$ defined as the probability $P(r \le r_m)$,

$$P(r \le r_{\mathrm{m}}) = \sum_{0}^{r_{\mathrm{m}}} f(r)\cdot r, \tag{6}$$

where r_{m} is the mode of $f(r)$. This descriptor gives an indication of how much a histogram is spread in relation to the peak (r_{m}) of the histogram. Particularly, it yields low values when the mode is located in the first few bins of the histogram and large values otherwise. This descriptor can be correlated with the PSF similarity measure to determine how it discriminates valid and nonvalid PSFs.

In Fig. 5, we plot the similarity measure S against $P(r \le r_{\mathrm{m}})$ for the estimated PSFs from two different retinal images which have been artificially degraded. The correlation coefficient for the two variables is -0.64. In addition to the correlation, from the plot we note that most PSFs with high $P(r \le r_{\mathrm{m}})$ have low kernel similarity (S) values, which is an indication that these are nonvalid or poorly estimated PSFs. A machine learning algorithm or clustering technique that automatically classifies nonvalid PSFs is out of the scope of this paper. Instead, our aim is to show that the energy characterization approach is sufficient to identify nonvalid PSFs, even at the expense of a few valid ones. As we show in Sec. 4.1, this is

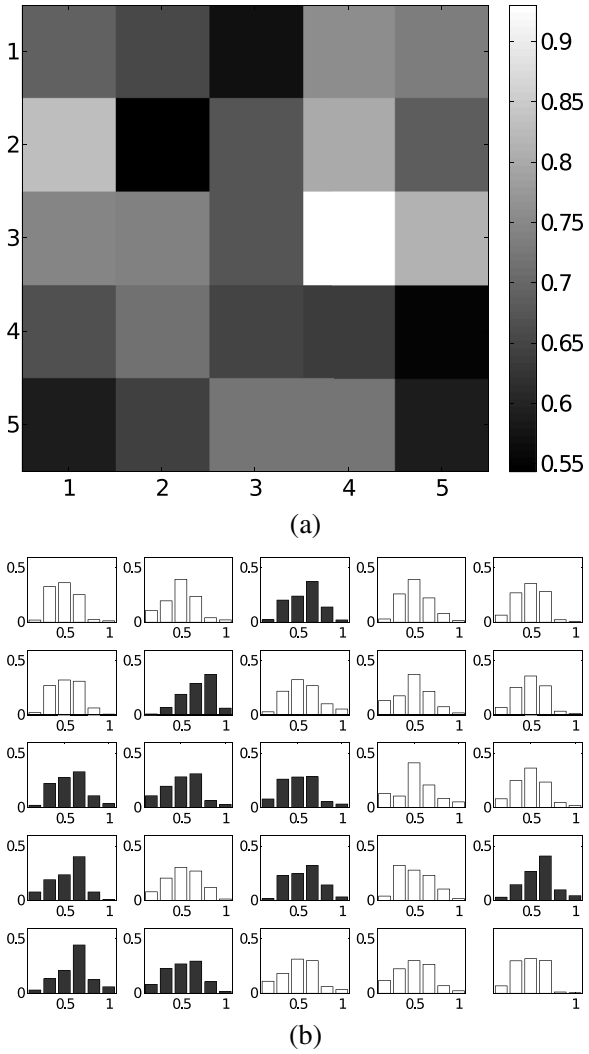


Fig. 4 (a) PSF similarity measure. (b) Characterization of estimated local PSFs by energy distribution. Histograms plotted in white bars in (b) have been labeled as valid PSFs based on the histogram probability descriptor (P) described in the text.

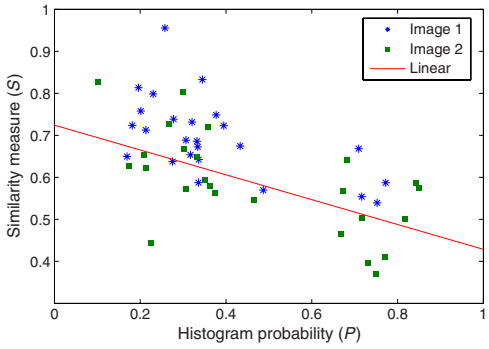


Fig. 5 PSF similarity measure (S) versus histogram probability (P). An S value of 1 represents an accurately estimated kernel. Low P values are associated with correctly estimated kernels.

not critical because the PSF changes smoothly throughout the FOV.

The histograms plotted in white bars in Fig. 4(b) correspond to PSFs with $P(r \leq r_m) < 0.3$, which we have labeled as valid PSFs. This means that we are favoring histograms skewed toward the left side. This is based on our assumption of the pattern that valid PSFs should follow. This correlates well with the similarity measure. Note that most PSFs with low similarity measure [darkest squares in Fig. 4(a)] have been correctly identified as nonvalid PSFs [black histograms in Fig. 4(b) and PSFs labeled with boxes in Fig. 6(a)].

The procedure for correcting the nonvalid local PSFs consists of replacing them with the average of adjacent valid kernels. Without this correction, the reconstruction develops ringing artifacts [see for example Fig. 11(b)]. The new set of valid local PSFs after replacing the nonvalid ones for the artificially degraded image is shown in Fig. 6(b).

3.4 *PSF Interpolation*

The computation of the SV PSF h is carried out by interpolating the local PSFs estimated on the regular grid of positions. The PSF values at intermediate positions are computed by bilinear interpolation of four adjacent known PSFs,[22] as shown in Fig. 7. Indexing any four adjacent grid points as $p = 1, \ldots, 4$ (starting from the top-left corner and continuing clockwise), the SV PSF in the position between them is defined as

$$h(s,t;x,y) = \sum_{p=1}^{4} \alpha_p(x,y) h_p(s,t), \tag{7}$$

where α_p is the coefficients of bilinear interpolation. Let us denote x_1 and x_2 as minimum and maximum x-coordinates of the subwindow, respectively. Analogously, y_1 and y_2 are the y-coordinates. Using auxiliary quantities

$$t_x = \frac{x - x_1}{x_2 - x_1}, \qquad t_y = \frac{y - y_1}{y_2 - y_1}, \tag{8}$$

the bilinear coefficients are

$$\begin{aligned} \alpha_1 &= (1 - t_y)(1 - t_x), \qquad \alpha_2 = (1 - t_y)t_x, \\ \alpha_3 &= t_y(1 - t_x), \qquad \alpha_4 = t_y t_x. \end{aligned} \tag{9}$$

In light of the definition for an SV PSF in Eq. (7), we can compute the convolution of Eq. (1) as a sum of four convolutions of the image weighted by coefficients $\alpha_p(x,y)$

$$[Hu](x,y) = \int u(s,t) h(s,t,x-s,y-t) \mathrm{d}s\mathrm{d}t, \tag{10}$$

$$= \int u(s,t) \sum_{p=1}^{4} \alpha_p(s,t) h_p(x-s,y-t) \mathrm{d}s\mathrm{d}t, \tag{11}$$

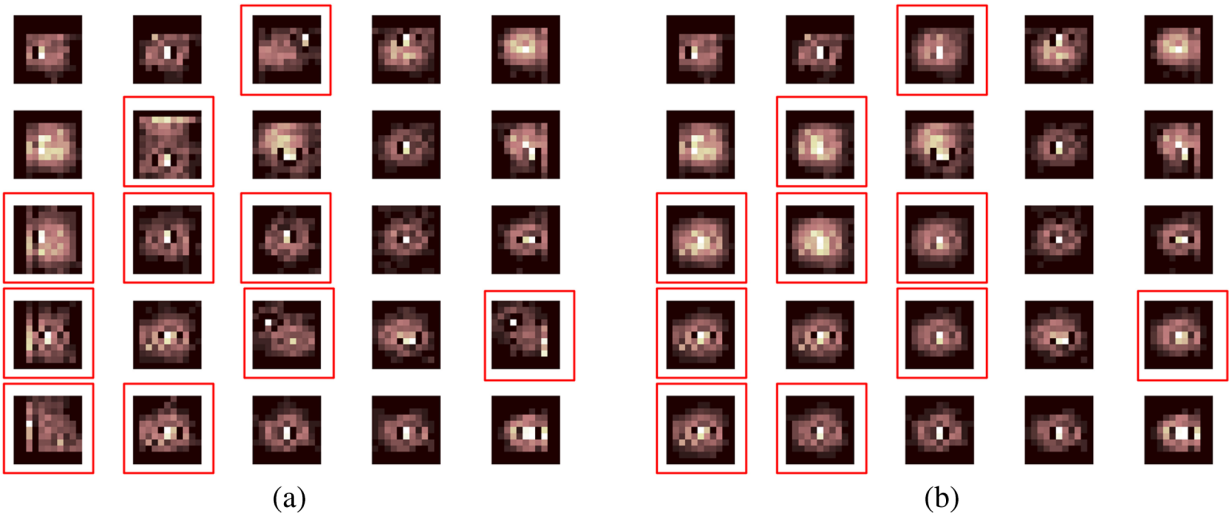
$$= \sum_{p=1}^{4} \int (\alpha_p(s,t) u(s,t)) h_p(x-s,y-t) \mathrm{d}s\mathrm{d}t, \tag{12}$$

$$= \left[\sum_{p=1}^{4} [\alpha_p u] * h_p \right](x,y). \tag{13}$$

In the same fashion, the operator adjoint to H (SV counterpart of correlation) denoted by H^* can also be defined in terms of the sums of four convolutions weighted by the α_p coefficients. These two operators are needed in all first-order minimization algorithms as the one used in the restoration stage (see Ref. 23 for further details).

3.5 *Restoration*

Having estimated a reliable SV PSF, we proceed to deblur the image. Image restoration is typically formulated within the Bayesian paradigm, in which the restored image is sought as the most probable solution to an optimization problem.



(a) (b)

Fig. 6 (a) Estimated set of 5×5 local PSFs. (b) New set of local PSFs with nonvalid PSFs replaced [compared with ground-truth set Fig. 3(b)]. Nonvalid PSFs have been labeled with a red square.

The restoration can be described as the minimization of the functional

$$\min_u \left[\frac{1}{2}\|z - Hu\|^2 + \lambda \int |\nabla u| \mathrm{d}x\mathrm{d}y\right], \tag{14}$$

where z is the blurred observed image, H is the blurring operator [Eq. (1)], u is the unknown sharp image, and λ is a positive regularization constant, which we have set according to a fine-tuning procedure.[4] The tuning procedure consists in artificially degrading a retinal image and restoring it with Eq. (14) by varying λ. Because the sharp original image is known we can compare it against the restored image using a metric like the peak-signal-to-noise ratio to determine an optimal value of λ. The first term penalizes the discrepancy between the model and the observed image. The second term is the regularization term which serves as a statistical prior. As regularization we use total variation, a technique that exploits the sparsity of image gradients in natural images. At present, solving the convex functional of Eq. (14) is considered a standard way to achieve close to state-of-the-art restoration quality without excessive time requirements.[24] We used an efficient method[25] to solve Eq. (14) iteratively as a sequence of quadratic functionals

$$u_{i+1} = \arg\min_u \left[\frac{1}{2}\|z - Hu\|^2 + \lambda \int \frac{|\nabla u|^2}{2|\nabla u_i|} + \frac{|\nabla u_i|}{2} \mathrm{d}x\mathrm{d}y\right]. \tag{15}$$

The functional of Eq. (15) bounds the original function in Eq. (14) and has the same value and gradient in the current u_i, which guarantees convergence to the global minimum. To solve Eq. (15), we used the conjugate gradient method.[17] The initial value of u_i for $i = 0$ is set to be equal to z. In order to avoid numerical instability for areas with small gradient ($|\nabla u_i|$ approaching zero), we use a relaxed ϵ-form of the minimized functional in Eq. (14), which implies that $|\nabla u_i| < \epsilon$ is equal to ϵ.

As regards the restoration of color RGB retinal images, we consider the following. The most suitable channel for PSF estimation is the green because it provides the best contrast.[26] This is mainly due to the spectral absorption of the blood in this band, which yields the dark and well contrasted blood vessels.[27] Conversely, the blue channel encompasses the wavelengths most scattered and absorbed by the optical media of the eye,[28] therefore the image in this band has very low energy and a relatively high level of noise. In the spectral zone of wavelengths larger than 590 nm, the light scattering on the red blood cells and the light reflection from the eye structures behind the vessel are dominant.[29] This produces the red band to be saturated and of poor contrast. As a result, we estimate the SV PSF from the green channel of the RGB color image, and later deconvolve every R, G, and B channels with the estimated SV PSF to obtain a restored RGB color image.

4 Experiments and Results

We performed several experiments on artificially and naturally degraded images to illustrate the appropriateness of the SV approach for restoring blurred retinal images. Moreover, to achieve an artifact-free restoration, we used our strategy for detecting and replacing the nonvalid local PSFs.

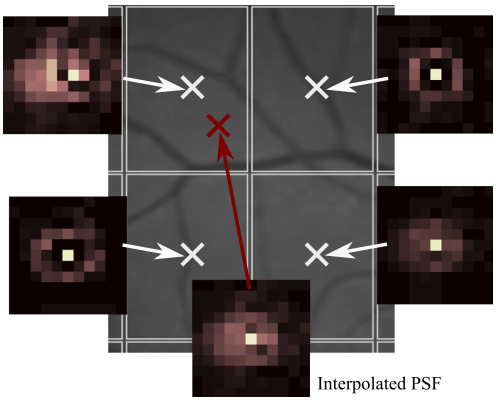


Fig. 7 Because the blur changes gradually, we can estimate convolution kernels on a grid of positions and approximate the PSF in the rest of the image (bottom kernel) by interpolation from four adjacent kernels.

4.1 *Artificially Degraded Images*

For the artificial experiment, we take a pair of images and degrade them with a 5×5 grid of realistic PSFs plus Gaussian noise ($\sigma = 10^{-6}$). The grid of PSFs was built upon realistic PSFs estimated from real degraded retinal images following the approach of Ref. 4. From the two input images, we restore one, and the other is used exclusively for the purpose of PSF estimation. We estimate the local PSFs by dividing the image into overlapping patches on a 5×5 grid [as shown in Fig. 3(a)]. The estimated PSFs are shown in Fig. 6(a). Because the PSF estimation may fail, we identify the nonvalid PSFs as described in Sec. 3.3. We replace them with with the average of adjacent valid kernels.

In Fig. 8(a), we show the restored artificial image with the directly estimated PSFs. The effect of nonvalid PSFs is evident in the poor quality of the restoration and the ringing artifacts. In Fig. 8(b), we show the restoration with the proposed method, where the nonvalid PSFs have been identified, removed, and replaced by the average of adjacent PSFs. To evaluate the restoration, we use the cumulative error histogram on a patch basis. The error[5] is the difference between a recovered image I_r with the estimated kernels and the known ground-truth sharp image I_g over the difference between the deblurred image I_{kg} with the ground-truth kernels. The error is given by $\|I_r - I_g\|/\|I_{kg} - I_g\|$. In Fig. 8(c), we show the cumulative error histogram for three restorations. H1 is the restoration with the directly estimated PSFs. It is important to note that shifted local PSFs warp the image which introduce additional artifacts and is the reason for such a low performance with approximately 40% of patches with an error lower than 2.5. After shifting the centroid of the PSFs to the geometrical center [restoration H2 in Fig. 8(c)], the reconstruction error is reduced significantly, about 60% of patches have an error lower than 1.5. Finally, the restoration (H3) with the removal of nonvalid PSFs increases significantly with all patches now displaying an error lower than 1.5. This means that after the nonvalid PSFs have been replaced the restoration quality is significantly increased.

To determine the limitations and robustness of the proposed method, we carried out several tests. First, we need to determine an optimal patch size for accurately estimating the local PSF. A patch that is too small compared to the kernel size may not have

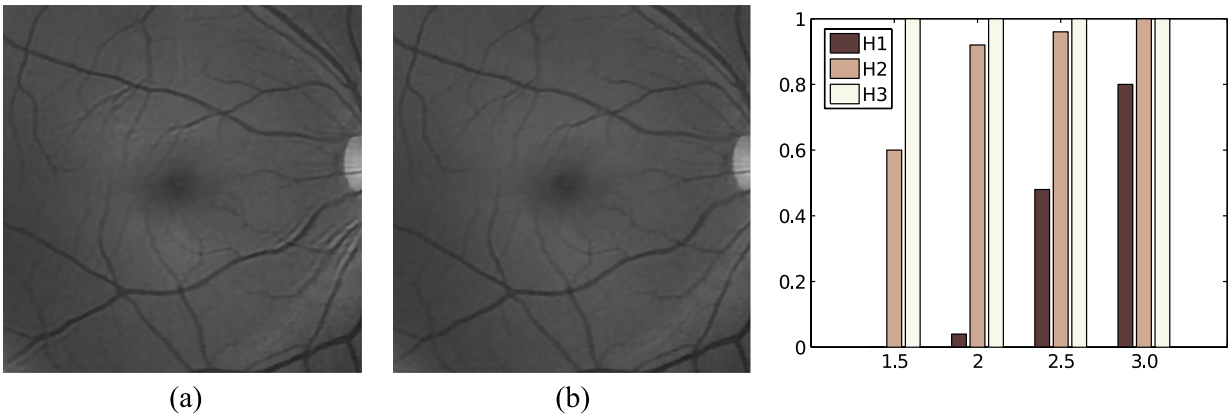


Fig. 8 (a) Restoration with the set of directly estimated PSFs shown in Fig. 6(a) (notice the artifacts due to nonvalid PSFs) and (b) restoration with the new set of PSFs with nonvalid PSFs replaced shown in Fig. 6. (c) Error histogram for evaluating the reconstruction using: H1—directly estimated PSFs, H2—PSFs shifted toward the geometrical center, and H3—the new set of valid PSFs.

enough information and is likely to favor the trivial solution (convolution with a delta). Conversely, a patch that is bigger than necessary may hinder the SV approach in addition to increasing the computational burden. In Fig. 9(a), we show the mean similarity measure versus the image patch size. It is important to note that a patch size of roughly four times the kernel size may be sufficient for an accurate PSF estimation. Increasing too much the patch size hinders the PSF estimation.

The second aspect to consider is under- and over-estimation of the local PSF size. In Fig. 9(b), we show the mean patch reconstruction error versus PSF size. As expected the error is minimum when the size is similar to the true size, yet under-estimating the PSF size is worse than over-estimating.

In relation to the SV characterization of the blur, we performed the PSF estimation with a different number of kernels. Initially with a coarse 2×2 grid increasing up to a 10×10 fine grid (the 5×5 grid is the ground-truth). A similar behavior is observed. Under-estimating the proper PSF grid size has a negative effect in that the SV nature of the blur is hardly identified which yields an error above 6 for the whole image. In this case, the error is not computed on a patch-basis because of the variable grid size.

4.2 *Naturally Degraded Images*

All of the naturally degraded images used in the experiments were acquired in pairs, typically with a time span between acquisitions of several minutes. Initially, to show the limits of the space-invariant approach we restored the blurred retinal image from Fig. 10(a) with a single global PSF with the space-invariant method we proposed in Ref. 4. This image corresponds to the eye fundus of a patient with strong astigmatism, which induces an SV blur as depicted by the image details shown in Figs. 10(b)–10(e). The restoration is shown in Fig. 11(a) and we can clearly observe various artifacts despite an increase in sharpness in a small number of areas. In view of this, it is evident that the space-invariant assumption does not hold in such cases. In the following, we move to the SV approach.

To carry out the SV restoration, we estimated the local PSFs on a 5×5 grid of image patches. From the estimated PSFs shown in Fig. 12(a), we notice a clear variation in shape mainly from the top-right corner where they are quite narrow, to the bottom left corner where they are more spread and wide. This variation is consistent with the spatial variation of the blur observed in the retinal image of Fig. 10(a). We restored the

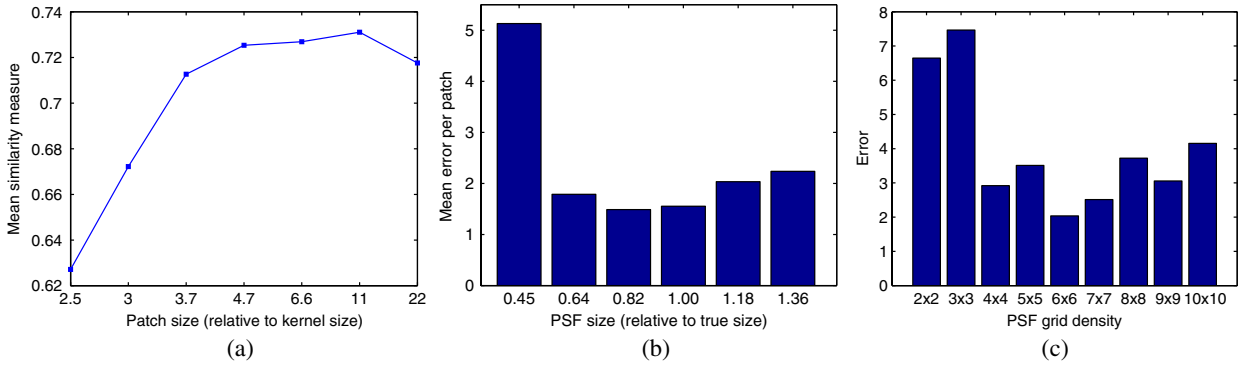


Fig. 9 (a) Mean similarity measure versus patch size. (b) Mean reconstruction error versus local PSF size. (c) Reconstruction error versus PSF grid size.

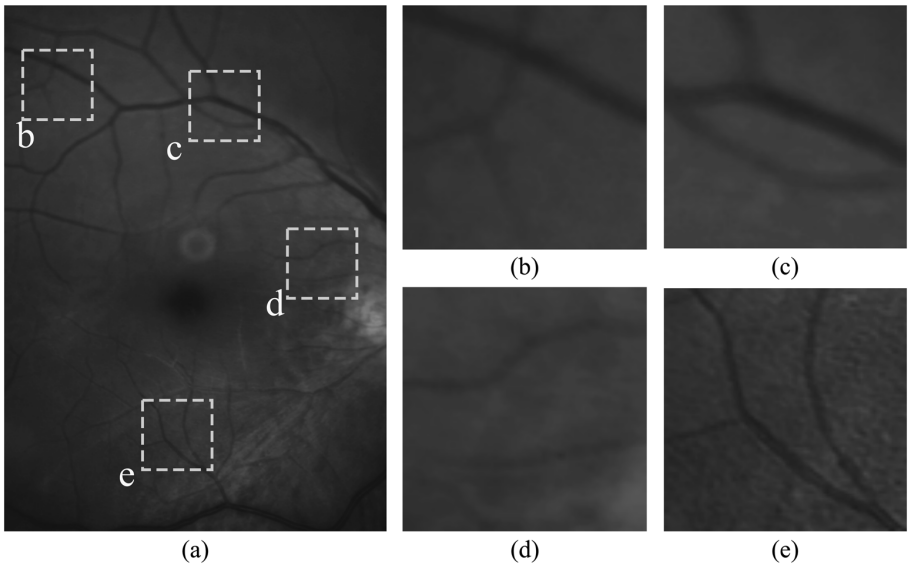


Fig. 10 (a) Retinal image degraded with real SV blur given by strong astigmatism. (b), (c), (d), and (e) zoomed regions to show the SV nature of the blur.

image with these local PSFs that were estimated directly without any adjustment. The restored image is shown in Fig. 11(b). One immediately obvious feature is that in several areas the restoration is rather poor, displaying ringing artifacts, whereas in others it is to some extent satisfactory. The local poor-restoration is linked to areas where the PSF estimation failed. By removing and correcting these nonvalid local PSFs, we obtained a noteworthy restoration shown in Fig. 11(c). Notice the overall improvement in sharpness and resolution with small blood vessels properly defined as shown by the image-details in the third column of Fig. 13. It could be said that without the replacement of the nonvalid PSFs the image quality after restoration is certainly worse than the original degraded image (see second column of Fig. 13).

To further demonstrate the capabilities of our method, additional restoration results on real cases from the clinical practice are shown in the following figures. As we mentioned in Sec. 1, a typical source of retinal image degradation comes from patients with corneal defects in which the cornea has an irregular structure [Fig. 1(a)]. This induces optical aberrations, which are mainly responsible for the SV blur observed in the retinal image. The image details shown in Fig. 1(b) reveal a significant improvement in which the retinal structures are much sharper and enhanced. In Fig. 14(a), a full color retinal image is shown, in which three small hemorrhages are more easily discernible in the restored image, along with small blood vessels. Another retinal image, shown in Fig. 14(b), reveals a clear improvement in resolution with a much finer definition of blood vessels.

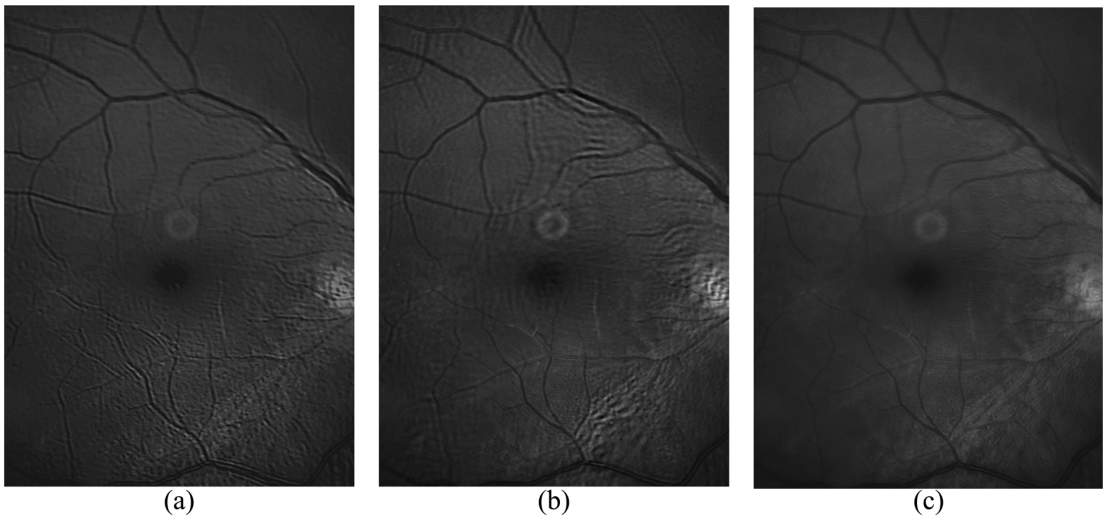


Fig. 11 (a) Space-invariant restoration, (b) SV restoration with directly estimated PSFs, and (c) SV restoration with the new set of PSFs. The reader is strongly encouraged to view these images in full resolution at http://www.goapi.upc.edu/usr/andre/sv-restoration/index.html.

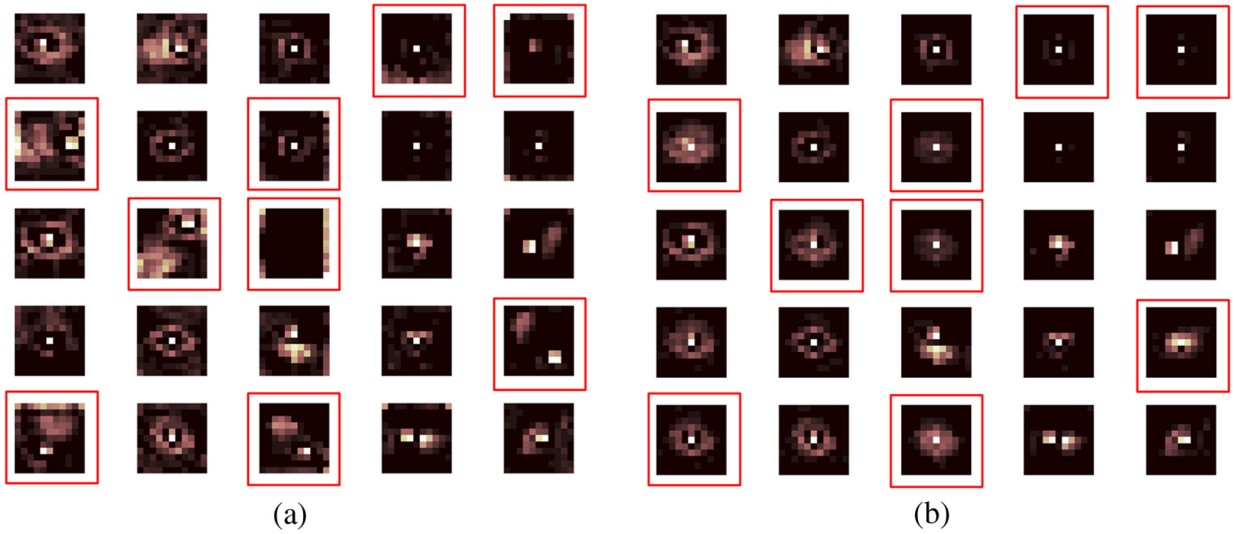


Fig. 12 (a) Estimated set of local PSFs from naturally degraded retinal image shown in Fig. 10(a). (b) Set of local PSFs after replacing nonvalid PSFs. Nonvalid PSFs have been labeled with a red square.

In addition, we processed retinal angiography images to test our method against a different imaging modality. Ocular angiography is a diagnostic test that documents, by means of photographs, the dynamic flow of dye in the blood vessels of the eye.[30] The ophthalmologists use these photographs both for diagnosis and as a guide to patient treatment. Ocular angiography differs from fundus photography in that it requires an exciter–barrier filter set (for further details see Ref. 30). The retinal angiography shown in Fig. 15 is degraded with a mild SV blur that hinders the resolution of small—yet important—details. The restoration serves to overcome this impediment; this can be observed from the zoomed-detail of the restored image.

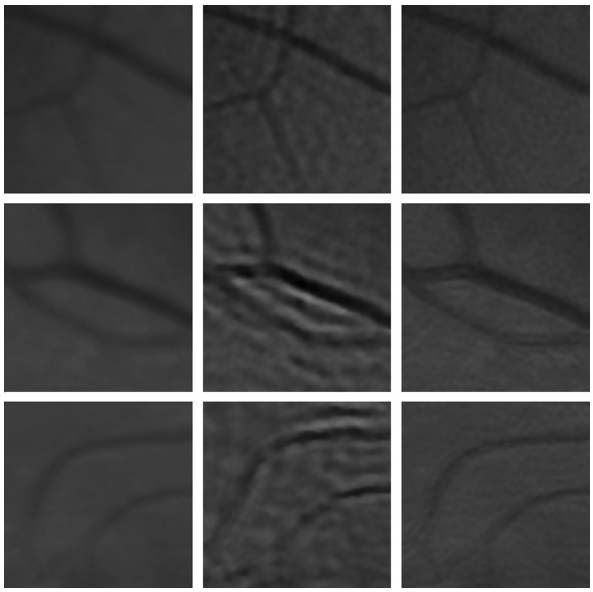


Fig. 13 Details of restoration. From left to right: the original degraded image, the SV restoration without correction of PSFs and the SV restoration with the correction.

The image enhancement may be useful for the improvement of recent analysis techniques for automated flow dynamics and identification of clinical relevant anatomy in angiographies.[31]

Finally, another way to demonstrate the added value of deblurring the retinal images is to extract important features, in this case detection of blood vessels. Such a procedure is commonly used in many automated disease detection algorithms. The improvement in resolution paves the way for a better segmentation of structures with edges. This is in great part due to the effect of the total variation regularization because it preserves the edge information in the image. To carry out the detection of the retinal vasculature, we used Kirsch's method.[32] It is a matched filter algorithm that computes the gradient by convolution with the image and eight templates to account for all possible directions. This algorithm has been widely used for

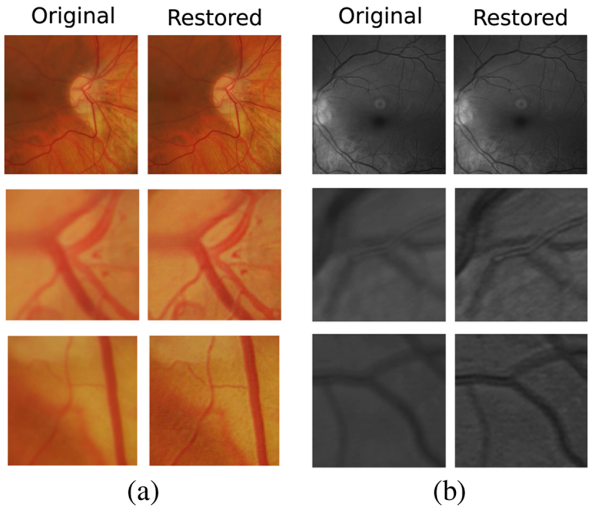


Fig. 14 Other retinal images restored with the proposed method. (a) First row: original and restored full-size retinal images. (b) Second and third rows: image details.

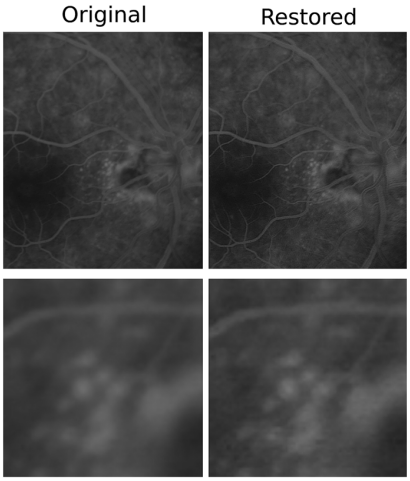


Fig. 15 Restoration of a retinal angiography. First row: original and restored full retinal images. Second row: image details.

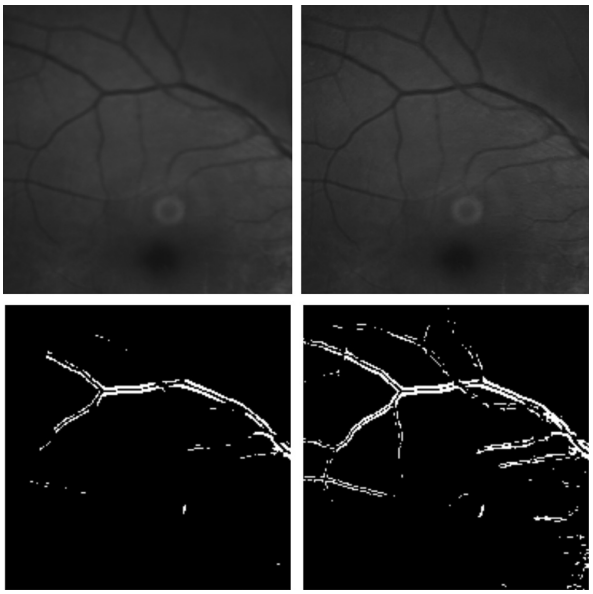


Fig. 16 First row (from left to right): original and restored retinal images. Second row: detection of blood vessels. Notice how the small blood vessels are better detected in the restored image.

detecting the blood vessels in retinal images.[33] In Fig. 16, we show the detection of the blood vessels from a real image of poor quality image and its restored version using our proposed method. A significant improvement in blood vessel detection is achieved. Smaller blood vessels are detected in the restored image, whereas the detection from the original image barely covers the main branch of the vasculature.

5 Conclusion

In this paper, we have introduced a method for restoring retinal images affected by SV blur by means of blind deconvolution. To do so, we described a spatially variant model of blur in terms of a convolution with a PSF that changes depending on its position. Since the SV degradation changes smoothly across the image, we showed that the PSF need not be computed for all pixels, which is quite a demanding task, but for a small set of discrete positions. For any intermediate position bilinear interpolation suffices. In this way, we achieve an SV representation of the PSF.

The estimation of accurate local PSFs proved difficult due to the very nature of the images; they usually contain textureless or nearly homogenous regions that lack retinal structures, such as blood vessels, to provide sufficient information. In this regard, we proposed a strategy based on eye-domain knowledge to adequately identify and correct such nonvalid PSFs. Without this, the restoration results are artifact-prone with an overall image quality that is worse than the original image. The proposal has been tested on artificially and naturally degraded retinal images coming from the clinical practice. The details from the restored retinal images show an important enhancement, which is also demonstrated with the improvement in the detection of the retinal vasculature.

In summary, it seems clear that the SV restoration of blurred retinal images is significant enough to leverage the images' clinical use. Improving the visibility of subtle details like small hemorrhages or small blood vessels may prove useful for disease screening purposes, follow-up monitoring, or early disease detection. With the new challenges faced by clinical services in the 21st century, automated medical image analysis tools are mandatory—this work is a step toward that direction.

Acknowledgments

This research has been partly funded by the Spanish Ministerio de Ciencia e Innovación y Fondos FEDER (project DPI2009-08879) and projects TEC2010-09834-E and TEC2010-20307. Financial support was also provided by the Grant Agency of the Czech Republic under project 13-29225S. Authors are grateful to Juan Luís Fuentes from the Miguel Servet University Hospital (Zaragoza, Spain) for providing images. The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

References

1. H. Bartling, P. Wanger, and L. Martin, "Automated quality evaluation of digital fundus photographs," *Acta Ophthalmol.* **87**(6), 643–647 (2009).
2. P. Godara et al., "Adaptive optics retinal imaging: emerging clinical applications," *Optom. Vision Sci.* **87**(12), 930–941 (2010).
3. J. Arines and E. Acosta, "Low-cost adaptive astigmatism compensator for improvement of eye fundus camera," *Opt. Lett.* **36**(21), 4164–4166 (2011).
4. A. G. Marrugo et al., "Retinal image restoration by means of blind deconvolution," *J. Biomed. Opt.* **16**(11), 116016 (2011).
5. A. Levin et al., "Understanding blind deconvolution algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2354–2367 (2011).
6. P. Bedggood et al., "Characteristics of the human isoplanatic patch and implications for adaptive optics retinal imaging," *J. Biomed. Opt.* **13**(2), 024008 (2008).
7. R. Tutt et al., "Optical and visual impact of tear break-up in human eyes," *Invest. Ophthalmol. Visual Sci.* **41**(13), 4117–4123 (2000).
8. J. Xu et al., "Dynamic changes in ocular zernike aberrations and tear menisci measured with a wavefront sensor and an anterior segment OCT," *Invest. Ophthalmol. Vis. Sci.* **52**(8), 6050–6056 (2011).
9. T. Costello and W. Mikhael, "Efficient restoration of space-variant blurs from physical optics by sectioning with modified Wiener filtering," *Digital Signal Process.* **13**(1), 1–22 (2003).
10. J. Bardsley et al., "A computational method for the restoration of images with an unknown, spatially-varying blur," *Opt. Express* **14**(5), 1767–1782 (2006).

11. S. Harmeling, M. Hirsch, and B. Scholkopf, "Space-variant single-image blind deconvolution for removing camera shake," *Adv. Neural Inf. Process. Syst.* **23**, 829–837 (2010).
12. O. Whyte et al., "Non-uniform deblurring for shaken images," *Int. J. Comput. Vis.* **98**(2), 168–186 (2012).
13. A. Gupta et al., "Single image deblurring using motion density functions," in *European Conf. on Computer Vision (ECCV 2010)*, K. Daniilidis, P. Maragos, and N. Paragios, Eds., pp. 171–184, Springer-Verlag, Crete, Greece (2010).
14. M. Tallón et al., "Space-variant blur deconvolution and denoising in the dual exposure problem," *Inf. Fusion* **14**(4), 396–409 (2012).
15. N. Salem and A. Nandi, "Novel and adaptive contribution of the red channel in pre-processing of colour fundus images," *J. Franklin Inst.* **344**(3–4), 243–256 (2007).
16. A. G. Marrugo and M. S. Millán, "Retinal image analysis: preprocessing and feature extraction," *J. Phys.: Conf. Ser.* **274**(1), 012039 (2011).
17. G. Golub and C. Van Loan, *Matrix Computations*, Vol. 3, Johns Hopkins University Press, Baltimore, Maryland (1996).
18. R. Navarro, E. Moreno, and C. Dorronsoro, "Monochromatic aberrations and point-spread functions of the human eye across the visual field," *J. Opt. Soc. Am. A* **15**(9), 2522–2529 (1998).
19. F. Sroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Process.* **14**(7), 874–883 (2005).
20. R. Navarro, "The optical design of the human eye: a critical review," *J. Optom.* **2**(1), 3–18 (2009).
21. Z. Hu and M.-H. Yang, "Good regions to deblur," in *European Conf. on Computer Vision (ECCV 2012)*, pp. 59–72, Springer-Verlag (2012).
22. J. G. Nagy and D. P. O'Leary, "Restoring images degraded by spatially variant blur," *SIAM J. Sci. Comput.* **19**(4), 1063–1082 (1998).
23. M. Sorel, F. Sroubek, and J. Flusser, "Towards super-resolution in the presence of spatially varying blur," in *Super-Resolution Imaging*, P. Milanfar, Ed., pp. 187–218, CRC Press, Boca Raton, Florida (2010).
24. P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*, CRC Press, Boca Raton, Florida (2007).
25. A. Chambolle and P. L. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.* **76**(2), 167–188 (1997).
26. A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE Trans. Med. Imaging* **22**(8), 951–958 (2003).
27. L. Gao, R. T. Smith, and T. S. Tkaczyk, "Snapshot hyperspectral retinal camera with the Image Mapping Spectrometer (IMS)," *Biomed. Opt. Express* **3**(1), 48–54 (2012).
28. M. Hammer and D. Schweitzer, "Quantitative reflection spectroscopy at the human ocular fundus," *Phys. Med. Biol.* **47**(2), 179–191 (2002).
29. V. Vucea et al., "Blood oxygenation measurements by multichannel reflectometry on the venous and arterial structures of the retina," *Appl. Opt.* **50**(26), 5185–5191 (2011).
30. P. Saine and M. Tyler, *Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging*, Butterworth-Heinemann, Woburn, Massachusetts (2002).
31. T. Holmes et al., "Dynamic indocyanine green angiography measurements," *J. Biomed. Opt.* **17**(11), 116028 (2012).
32. R. A. Kirsch, "Computer determination of the constituent structure of biological images," *Comput. Biomed. Res.* **4**(3), 315–328 (1971).
33. M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," *Comput. Biol. Med.* **37**(2), 262–267 (2007).

Andrés G. Marrugo received the BE degree (summa cum laude) in mechatronical engineering from the Technological University of Bolívar, Cartagena, Colombia, in 2008, the MSc in photonics, and the PhD in optical engineering from the Technical University of Catalonia, Barcelona, Spain, in 2009 and 2013, respectively. He is currently with the Technological University of Bolívar. He is a member of SPIE, the Spanish Optical Society, and the European Optical Society.

María S. Millán received the PhD in physics in 1990 and is a full professor of the College of Optics and Optometry in the Technical University of Catalonia, Barcelona, Spain. Her research work on image processing involves optical and digital technologies, algorithms, and development of new applications to industry and medicine. She has been a representative of the Spanish Territorial Committee in the International Commission for Optics. She is a member of the Optical Society of America, a fellow member of SPIE and a fellow member of the European Optical Society.

Michal Šorel received the MSc and PhD degrees in computer science from Charles University in Prague, Czech Republic, in 1999 and 2007, respectively. From 2012 to 2013 he was a postdoctoral researcher at Heriot-Watt University in Edinburgh, Scotland, and the University of Bern, Switzerland. Currently he is a research fellow in the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic.

Filip Šroubek received the MS degree in computer science from the Czech Technical University, Prague, Czech Republic, in 1998 and the PhD degree in computer science from Charles University, Prague, Czech Republic, in 2003. From 2004 to 2006, he was on a postdoctoral position in the Instituto de Optica, CSIC, Madrid, Spain. In 2010 and 2011, he was the Fulbright visiting scholar at the University of California, Santa Cruz. He is currently with the Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic.

Journal of Electronic Imaging 22(1), 011003 (Jan–Mar 2013)

Image deblurring in smartphone devices using built-in inertial measurement sensors

Ondřej Šindelář
Charles University in Prague
Faculty of Mathematics and Physics
Prague, Czech Republic
Email: ondrasindelar@gmail.com

Filip Šroubek
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 4, CZ-182 08, Praha 8, Czech Republic

Abstract. *Long-exposure handheld photography is degraded with blur, which is difficult to remove without prior information about the camera motion. In this work, we utilize inertial sensors (accelerometers and gyroscopes) in modern smartphones to detect exact motion trajectory of the smartphone camera during exposure and remove blur from the resulting photography based on the recorded motion data. The whole system is implemented on the Android platform and embedded in the smartphone device, resulting in a close-to-real-time deblurring algorithm. The performance of the proposed system is demonstrated in real-life scenarios.* © *2013 SPIE and IS&T [DOI: 10.1117/1.JEI.22.1.011003]*

1 Introduction

Blur induced by camera motion is a frequent problem in photography mainly when the light conditions are poor. As the exposure time increases, involuntary camera motion has a growing effect on the acquired image. Image stabilization (IS) devices that help to reduce the motion blur by moving the camera sensor in the opposite direction are becoming more common. However, such hardware remedy has its limitations, as it can compensate only for motion of a very small extent and speed. Deblurring the image offline using mathematical algorithms is usually the only choice we have in order to obtain a sharp image. Motion blur can be modeled by convolution, and the deblurring process is called deconvolution, which is a well-known ill-posed problem. In general, the situation is even more complicated, since we usually have no or limited information about the blur shape.

We can divide the deconvolution methods into two categories: methods that estimate the blur and the sharp image directly from the acquired image (blind deconvolution) and methods that use information from other sensors to estimate the blur (semi-blind deconvolution).

Paper 12326SSP received Aug. 22, 2012; revised manuscript received Dec. 20, 2012; accepted for publication Dec. 28, 2012; published online Feb. 4, 2013.

Over the last few years, blind deconvolution has experienced a renaissance. The key idea of new algorithms belonging to the first category is to address the ill-posedness of blind deconvolution by characterizing the image prior to using natural image statistics and by a better choice of estimators. A frantic activity started with the work of Fergus et al.,[1] who applied variational Bayes to approximate the posterior by a simpler distribution. Other authors[2,3,4,5] stick to the "good old" alternating maximum a posteriori estimation approach, but by using ad hoc steps, which often lack rigorous explanation, they converge to a correct solution. Levin et al. in Refs. 6 and 7 proved that a proper estimator matters more than the shape of priors. They showed that marginalizing the posterior with respect to the latent image leads to the correct solution of the blur. The marginalized probability can be expressed in a closed form only for simple priors that are, e.g., Gaussian. Otherwise approximation methods such as variational Bayes[8] or the Laplace approximation[9] must be used. Complex camera motion often results in blur that is space-variant, i.e., the blur is a function of a position vector. As a rule, the space-variant blur cannot be expressed by an explicit formula, but in many cases it has a special structure that can be exploited. If only one type of camera motion is considered (e.g., rotation), we can express the degradation operator as a linear combination of basis blurs (or images) and solve the blind problem in the space of the basis, which has much lower dimension than the original problem. Whyte et al.[10] considered rotations about three axes up to several degrees and described blurring using three basis vectors. For blind deconvolution, they used an algorithm analogous to Ref. 1 based on marginalization over the latent sharp image. Gupta et al.[11] adopted a similar approach, replacing rotations about x and y axes by translations. State-of-the-art blind-deconvolution algorithms achieve sometimes awesome results. However, their main limitation is that they work only in specific situations, they are prone to local extrema, and they are computationally very demanding.

The second category of deconvolution algorithms (semi-blind) tries to overcome these drawbacks by using information about the camera motion from other sources. One possibility is to acquire a pair of images: one correctly

exposed but blurred and one underexposed (noisy) but sharp image. Then we can apply multichannel blind deconvolution methods, which are better posed, as was proposed for example in Refs. 12, 13, and 14. Another possibility is to attach an auxiliary high-speed camera of lower resolution to estimate the point-spread function (PSF) using for example optical flow techniques.[15,16] Many devices, such as modern smartphones, are now equipped with inertial sensors (gyroscopes and accelerometers) that can give us accurate information about camera motion. If we are able to reconstruct camera path, then we can recover blur and perform nonblind image deblurring. This idea was recently described by Joshi et al., in Ref. 17, but they have designed an expensive measuring apparatus consisting of a digital single-lens reflex camera and a set of inertial sensors and perform image deblurring offline on a computer. This work is based on the same idea, but our aim is to show that image deblurring is feasible on modern smartphones and not requiring any other devices.

The main contribution of this work is to illustrate that blur estimation with built-in inertial sensors is possible and to implement image deblurring on a smartphone, which works in practical situations and is relatively fast to be acceptable for a general user. The next section shows the relation between the camera pose and the image blur, and discusses simplifications that we make. Section 3 briefly describes implementation on our test device (Samsung smartphone). Section 4 shows results of our experiments and addresses pitfalls that are common for cameras embedded in smartphones.

2 Camera Motion Blur Analysis

We start the discussion with a general camera motion. Since our primary goal is a handy implementation for mobile devices, we then introduce simplification of the problem that allows a fast and memory-conserving solution with promising results.

2.1 *Model*

The image degradation model is represented by relation

$$g = H(u) + n, \tag{1}$$

where H is a linear degradation operator and n is additive noise. Image coordinate indices are omitted here for simplicity. Our goal is to find an estimate of the original image u from the observed blurred image g.

To track the effect of camera motion on the output image, we first assume a standard perspective projection $\Pi:\mathbb{R}^3 \to \mathbb{R}^2$ that transforms a three-dimensional (3-D) point $[x, y, z]$ in the observed scene to a two-dimensional (2-D) location $[x', y']$ in the image plane:

$$\Pi([x, y, z]^T) = \left[\frac{xf}{z}, \frac{yf}{z}\right]^T. \tag{2}$$

For the sake of brevity, we assume here only the focal length f in the intrinsic camera matrix. The optical axis is identical with the z axis. During camera motion, projection of a point $p = [x, y, z]^T$ at time τ within the exposure period is given by

$$C(\tau) = \Pi\left(R(\tau)\begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} t_x(\tau) \\ t_y(\tau) \\ t_z(\tau) \end{bmatrix}\right) = \Pi[R(\tau)p + t(\tau)], \tag{3}$$

where R and t are the 3-D rotation matrix and translation vector, respectively, that define the camera pose at time τ. The rotation matrix $R(\tau)$ is given by three rotation angles $\phi_x(\tau)$, $\phi_y(\tau)$ and $\phi_z(\tau)$.

The resulting curve C makes up a trajectory of a trace that is left on the sensor by a point light source. Assuming a constant illuminance over the exposure period, the light energy emitted from the point is distributed evenly (with respect to time) over the curve C. This effectively gives us a time parametrization of a PSF for a given point, which forms the blur operator H. The operator H can be written in a form naturally generalizing standard convolution as

$$H(u)[x, y] = \int u(x - s, y - t)\tilde{h}(s, t, x - s, y - t)\mathrm{d}s\mathrm{d}t, \tag{4}$$

where $\tilde{h}$ depends on the position (third and fourth variable) and can be regarded as a space-variant PSF.

Now we can draw the relation between $\tilde{h}$ in Eq. (4) and the curve C. For any given 3-D point at position p rendered on the image plane to $[x', y'] = \Pi(p)$ the point-spread blur function $\tilde{h}(s, t, x', y')$ is a 2-D function of $[s, t]$, which can be interpreted as a blurred image of an ideal light point displayed at $[x', y']$. It can be thus obtained by rendering the curve C on a plane with the total integral of $\tilde{h}$ (which has to be equal to 1 to conserve distribution of energy) distributed along the path evenly in respect to the time parameter.

In the next section, we will show how to simplify this model and assume the space-invariant case, i.e., $\tilde{h}(s, t, x, y) = h(s, t)$.

2.2 *Space-Invariant Simplification*

We will consider a situation when the operator H is spatially invariant, so Eq. (1) becomes

$$g = h * u + n, \tag{5}$$

where "$*$" denotes convolution and h is a space-invariant PSF.

The PSF Eq. (3) is spatially variant in general, so it will be modified for our purposes. First of all, the translation t affects the projection differently depending on the object distance from the camera. The relation is inversely proportional, as shown in Fig. 1(a). In the case of our test device, if the camera shifts by 1 mm, objects at distance of 2 m or more move by less than 1 pixel in the image. We can thus effectively ignore translation as a cause of blur in many practical situations.

Rotation about the optical z axis (yaw) intuitively interferes with the space-invariant blur assumption. This type of rotation applied on a point light source placed in the center of the picture (on the optical axis) leaves the projection unchanged, but points outside the center form arc-shaped traces that grow toward the image borders. Provided that the camera is rotated with an equal amount around all three axes, which is a fair assumption under normal circumstances, a yaw has the least effect on the resulting blur, especially in the center of the sensor. Cellphone cameras typically

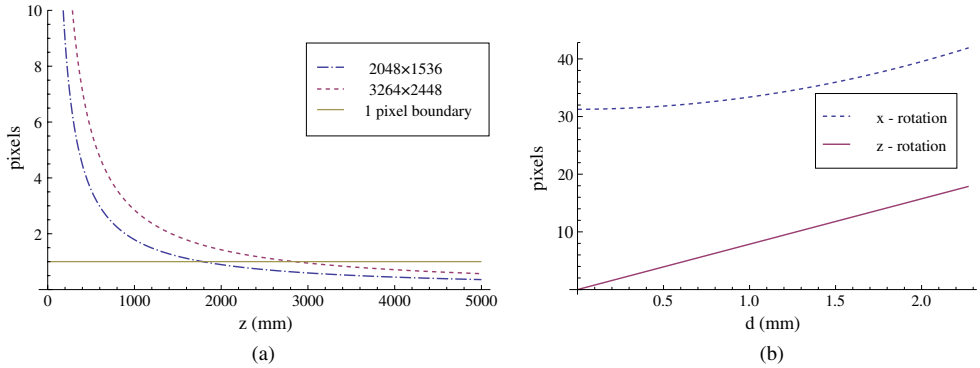


Fig. 1 Dependence of projection shift on translation and z rotation for a test device. (a) Influence of 1 mm x or y translation depending on object distance. Angle of view is 60 deg; two curves represent different image sensor resolution; (b) influence of 1 deg rotation about x and z axis depending on a distance d from the image sensor center. The full sensor extent corresponds to $d = 2.3$ mm; image resolution is 2048×1536.

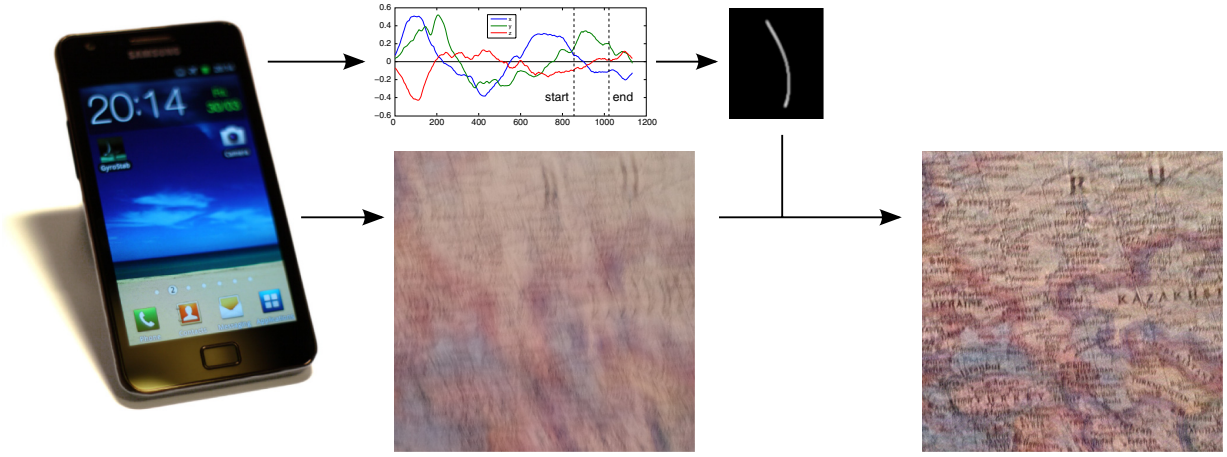


Fig. 2 Basic application workflow. Together with a taken photograph gyroscope data are recorded, which is a base for blur kernel estimation. A deconvolution is then performed to remove blur from the image.

have the focal length close to the sensor size, which means that only close to the image borders the blur size produced by yaw is approaching the blur size produced by rotation about x or y; see Fig. 1(b).

The last obstacle towards the space-invariant PSF is the perspective projection itself. Length of a trace caused by x and y rotations are projected slightly differently depending on the distance from the optical center, because the rectilinear projection in Eq. (2) casts a point at an angle α from the optical axis to a point at a distance of $f \cdot \tan(\alpha)$ from the image center. The tangent function is close to linear for small angels, so both x and y rotations by a small angle α shift a point in the sensor center approximately $f \cdot \alpha$ away in the direction of the given axis. Using the same rule for all points on the sensor gives us the space-invariant simplification of Eq. (3):

$$C(\tau) \approx \begin{bmatrix} x' \\ y' \end{bmatrix} + f \begin{bmatrix} \phi_x(\tau) \\ \phi_y(\tau) \end{bmatrix}, \tag{6}$$

where $[x', y']$ is the location of a point in the image. This approximation holds if z is large, and $x'\phi_x \ll f$ and $y'\phi_y \ll f$, which is true at least in the central part of the image.

3 Implementation

As a testing platform, we have chosen a *Samsung Galaxy S II* smartphone with Android OS. It is equipped with all the apparatus needed for our experiments; namely a relatively high-quality camera, motion sensors, a fast CPU, and enough RAM to perform computations.

3.1 *PSF Estimation*

During the photo acquisition, samples of angular velocity are recorded using the embedded gyroscopes, which are afterward trimmed to fit the exposure period. An estimation of the PSF is rendered by integrating the curve position from the recorded data using Eq. (6).

3.2 *Deconvolution*

State-of-the-art nonblind deconvolution methods use sparse image priors, and the solution is usually found by some iterative minimization algorithms, such as in Ref. 4. However, the limited computational power of the smartphone prevents us

from implementing these sophisticated deconvolution methods. We thus use a simple but fast Wiener filter in the form

$$\hat{U} = G\frac{H^*}{|H|^2 + \Phi}, \tag{7}$$

where Φ is an estimation of the inverse signal-to-noise ratio, and G, H, and $\hat{U}$ are discrete Fourier transforms of the observed image g, PSF h, and the estimated latent image $\hat{u}$, respectively.

Filtering in the frequency domain treats the image as a periodic function, which causes ringing artifacts around image borders. To overcome this problem, several less or more sophisticated techniques were proposed in the literature.[18,19] We have found sufficient to preprocess the input image g by blending the opposite image borders at the width of the PSF, which creates a smooth transition and eliminates the artifacts.

The intensity values of the output image $\hat{u}$ sometimes lie outside the 8-bit range (0 to 255), therefore we added optional normalization with clipping of outliers. The normalization is especially useful in the case of larger blurs and scene with high illumination.

For conversions of the images to frequency domain and back, we use fast Fourier transform (FFT) algorithm implemented in the fastest Fourier transform in the West (FFTW) library. Utilizing a fast ARM Cortex-A9 CPU with two cores and support for a single instruction, multiple data instruction set (NEON), FFTW proved to be remarkably fast on the tested smartphone (see Table 1).

The acquired images with native camera resolution of 3264×2448 is by default scaled down to 2048×1536 to take the advantage of better performance of FFTW when the image size is a factor of small primes. Image downsampling has a negligible effect on the image quality, because native camera resolution is unnecessarily high. The optical system of the camera has a very small aperture, which, because of diffraction and optical aberrations, limits the number of pixels that can be effectively captured by the image sensor.

To perform Wiener filtering, FFT must be applied several times: once for the PSF and twice (forward and backward-inverse) for each color channel. That yields a total of seven FFT operations. With some overhead of bitmap transfers, the deconvolution phase for the image resolution 2048×1536 takes about 2.6 s. The whole process starting from the camera shutter is done in a little over 6 s. This includes image resizing, PSF estimation, compressing, and saving the original and deblurred image files. The main application workflow is summarized on a schematic diagram in Fig. 2.

Table 1 Speed (in milliseconds) of FFT transform of grayscale images with different sizes and different CPU settings.

Resolution	No NEON, No hardware FPU	NEON, 1 core	NEON, 2 cores
1536×1152	2900	185	110
2048×1536	5300	330	195
2050×1538	—	1000	540
3264×2448	21200	1450	800

4 Results

In this section we display several of our results together with estimated PSFs; see Figs. 3, 4, and 5. All results were computed with the signal-to-noise parameter Φ set to 0.01. This value was determined experimentally to provide the best looking results. The original intention was to set Φ proportionally to the film speed (ISO value) extracted from the exchangeable image file format data of a photo, which should determine the amount of noise present in the image. However, we found the dependency of Φ on ISO very negligible. We explain this behavior by the denoising step that the mobile phone internally performs on the captured photos.

For comparison, we show an advanced nonblind iterative method (TV-L1) by Xu and Jia (Ref. 5)*, which minimizes image total variation and data term in the L_1-norm. We also tested blind deconvolution proposed in the same, which is probably currently the best blind deconvolution method. However, the result of the first test image shown in Fig. 3(e) illustrates a total failure of this method when applied to images taken by our test device. The PSF [Fig. 3(f)] estimated by the blind deconvolution method is close to a delta function and the estimated image [Fig. 3(e)] is thus a slightly sharpened image. We suspect that small PSF variations in space and/or the image post-processing done by the smartphone prevents a successful estimation of the correct motion blur. The same unsatisfactory behavior was observed in all our tests. However, our results in Figs 3(c), 4c, and 5c illustrate that in spite of a relatively simple approach, which incorporates the Wiener filter with the space-invariant PSF estimated by inertial sensors, the proposed method is capable of producing convincing images exposing many details that were hidden in the original. The nonblind algorithm of Xu and Jia, which is using the same PSF estimated by inertial sensors, tends to amplify the signal, which rather emphasizes noise and false edges than gains signal improvement. Conversely, high-frequency details are more suppressed, probably due to being treated as noise, despite of careful attempts to tune the parameters of the method. Within our testing environment, the simplified Wiener filter is more advantageous as it filters all frequencies evenly which apparently matches the spectrum characteristics of most of the tested images.

Our results seem to lack contrast, which is largely because of the normalization. On the other hand, it helps retaining the full dynamic range without saturation as clearly seen in the comparison Fig. 3.

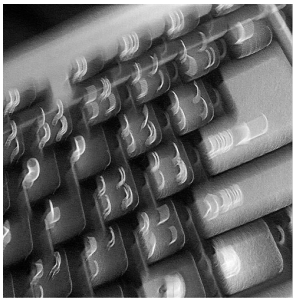
Our deconvolution process admittedly has downsides, as well. Focusing in a dark environment may be unsuccessful, and then the deconvolved result cannot be sharp even if the PSF estimation is correct, since we lack any means to estimate the out-of-focus blur.

The subjective quality of the deconvolution output is not entirely consistent. Images presented in this section are the best-looking results. Outputs of the similar quality are frequently achieved by our method, but sometimes the result is impaired by visual anomalies worsening its appearance. Most often it is manifested as ringing artifacts surrounding sharp edges in the picture, as demonstrated in Fig. 6.

*An executable is available for download at ~http://appsrv.cse.cuhk.edu.hk/~xuli/deconv.zip

(a) original (b) TV-L1 (c) our result

(d) PSF



(e) Xu, Jia — blind (f) PSF est. – Xu, Jia

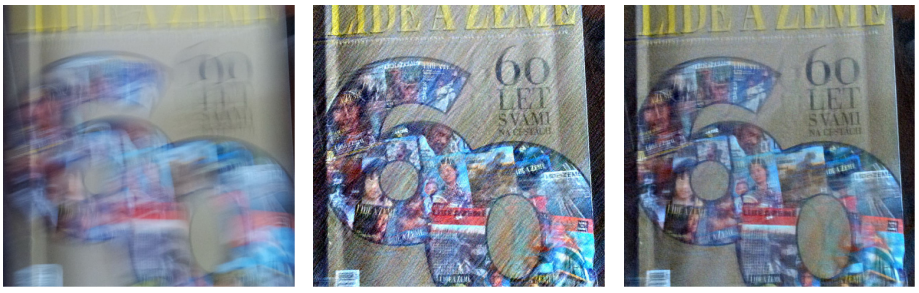
Fig. 3 Test 1: 1/7 s exposure, 16×59 estimated PSF.

The lack of control over camera hardware in the phone (no manual exposure settings, no access to raw data from the image sensor) and inaccurate timing of exposure events prevents us to systematically evaluate our method and find sources of malfunctioning.

The main problem is most likely the space-variant nature of the PSF as discussed in Sec. 2, which is particularly noticeable when a rotation about the z axis is significant or a translation movement is present and the scene depth is small. The example in Fig. 6 is influenced by a combination of both of these factors. The space-invariant approximation of camera projection is often apparent in parts close to image borders, because of a relatively wide camera field of view (60 deg).

However, another cause is the shutter mechanism. Contrary to systems with a mechanical shutter, values of illuminated pixels are here read successively line by line. The readout from the CMOS sensor takes several tens of milliseconds as shown in Fig. 7, which results in a picture not taken at a single moment, but with a slight time delay between the first and last pixel row. This process, called rolling shutter, is therefore another cause of the blur variance as the PSF depends on the vertical position in the image. The correct approach to PSF estimation is thus shifting inertial sensor data in time according to the vertical position in the image.

The application programming interface (API) of the tested device does not allow accurate synchronization between camera and gyroscope samples. Therefore we have



(a) original (b) TV-L1 (c) our result

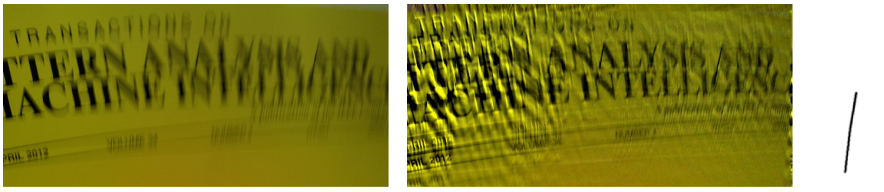
(d) PSF

Fig. 4 Test 2: 1/9 s exposure, 21×28 estimated PSF.

(a) original (b) TV-L1 (c) our result

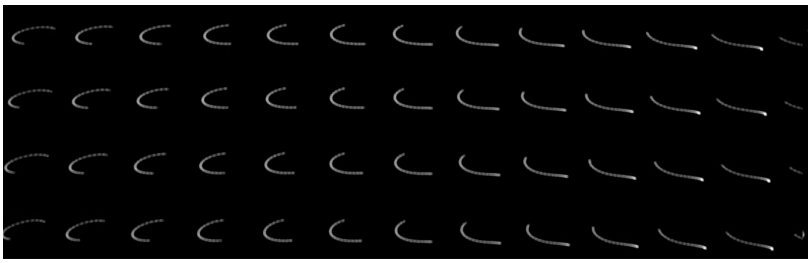
(d) PSF

Fig. 5 Test 3: 1/2 s exposure, 72×76 estimated PSF.



(a) original (b) result (c) PSF

Fig. 6 An example of an unsatisfactory result.



(a) traces of points on LCD

(b) 40 ms (c) 50 ms (d) 60 ms (e) 70 ms (f) 80 ms (g) 90 ms (h) 100 ms

Fig. 7 A snapshot of point grid displayed on a liquid crystal display screen showing the rolling shutter effect. The bottom row shows a series of blur kernels rendered using data from the gyroscope sensor shifted in time. Exposure 1/14 s, PSF images were created from sensor data starting 40 to 100 ms after a synchronization timestamp.

implemented a deconvolution preview, where the user picks the best option from a set of results created with time-shifted PSFs. The preview also partly solves the rolling shutter problem, since the selected time shift corresponds to a horizontal image band of a certain height that can be considered as acquired at one moment, thus eliminating the rolling shutter effect for that image part.

Image post-processing might also present a serious problem for the deconvolution. Since the original raw data from the image sensor are not available, we are forced to work with the JPEG-compressed image, which is most likely processed by a denoising, contrast-enhancement algorithm, or lens-distortion compensation. These adjustments are undesirable for our purposes, as they were not taken into account in our model.

Noise present in gyroscope measurement data can also be a problem, as displayed in Figs. 8 and 9. This has been examined in a following synthetic experiment. A test image was

Fig. 8 Noise in gyroscope data. Synthetically blurred Lena image using PSF from recorded gyroscope samples and afterward deblurred using PSF from measurements with variable amount of noise. Images are from left to right, top to bottom: original, blurred, and six deblurred images using original gyroscope data altered by random Gaussian noise with variance from 0 to 0.05 (gyroscope measurements are in rad/s).

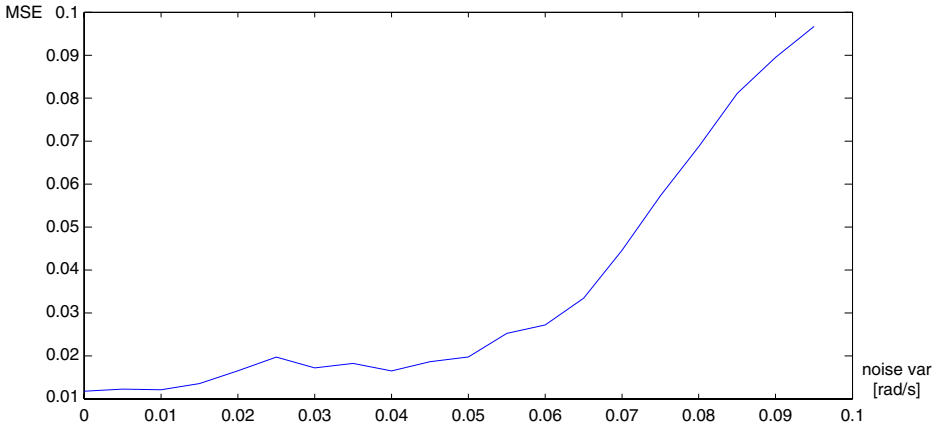


Fig. 9 Mean squared error (MSE) of difference between the original and deblurred image in relation to amount of added sensor noise. Gaussian noise of variance 0 to 0.1 was added to gyroscope measurements (angular velocity in rad/s). Deconvolution algorithm was then performed using computed blur kernels based on these altered measurements. MSE of difference to the original image is plotted in the graph (pixel value was normalized to $\langle 0, 1 \rangle$ range). The graph shows mean of 10 iterations for each of the variance values. Lena image was used for the test.

first blurred using convolution with a PSF counted from one set of gyroscope samples recorded in our mobile application. An additive noise was added to the image in accordance with the model 1 (40-dB Gaussian noise was used). Gaussian noise was also added to the gyroscope samples to simulate errors in sensor measurement. Corrupted image was then repaired using our deblurring algorithm from the altered motion data. Results for different amounts of noise in gyroscope samples are shown in Fig. 8. The mean square error of the result as a function of the gyroscope noise level (variance) is in Fig. 9. We can see that the performance starts to drop for noise levels above 0.05 rad/s. The gyroscope noise level typically encountered in the motion sensors inside mobile devices (in our case Samsung Galaxy S II) is 0.007 rad/s for our sampling rate, and it is therefore way below the critical level.

5 Conclusion

We have presented an image deblurring method that can effectively remove blur caused by camera motion using information from inertial sensors. The proposed method is fully implemented on a smartphone device, which is to our knowledge the first attempt in this direction and renders the method particularly appealing for end users. We have justified the space-invariant simplification for certain camera motions, but simultaneously we have uncovered intrinsic

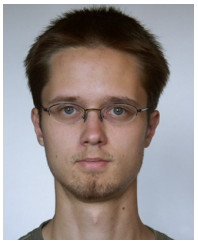
sources of space-variant blur, such as rolling shutter. The space-variant implementation of the deblurring algorithm, which would solve some of the current issues, is in theory possible, but the computational cost on the smartphone may be too high. It will be a topic of our future research to find out whether this is viable.

Acknowledgments

This work was supported in part by the Academy of Sciences of the Czech Republic under Grant M100751201 and by the Grant Agency of the Czech Republic under Project P103/11/1552.

References

1. R. Fergus et al., "Removing camera shake from a single photograph," in *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, pp. 787–794, ACM, New York, NY (2006).
2. J. Jia, "Single image motion deblurring using transparency," in *Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit., CVPR '07*, pp. 1–8, IEEE Computer Society Press, Silver Spring, MD (2007).
3. N. Joshi, R. Szeliski, and D. J. Kriegman, "PSF estimation using sharp edge prediction," in *Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit., CVPR 2008*, pp. 1–8, IEEE Computer Society Press, Silver Spring, MD (2008).
4. Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," in *SIGGRAPH '08: ACM SIGGRAPH 2008 papers*, pp. 1–10, ACM, New York, NY (2008).
5. L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proc. 11th European Conf. on Comput. Vis: Part I, ECCV'10*, pp. 157–170, Springer-Verlag, Berlin, Heidelberg (2010).
6. A. Levin et al., "Understanding blind deconvolution algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2354–2367 (2011).
7. A. Levin et al., "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit., CVPR '09*, IEEE Computer Society Press, Silver Spring, MD (2009).
8. J. Miskin and D. J. C. MacKay, "Ensemble learning for blind image separation and deconvolution," in *Advances in Independent Component Analysis*, M. Girolani, Ed., Springer-Verlag, Berlin (2000).
9. N. P. Galatsanos et al., "Hierarchical bayesian image restoration from partially known blurs," *IEEE Trans. Image Process.* **9**(10), 1784–1797 (2000).
10. O. Whyte et al., "Non-uniform deblurring for shaken images," in *Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit. CVPR '10* , pp. 491–498, IEEE Computer Society Press, Silver Spring, MD (2010).
11. A. Gupta et al., "Single image deblurring using motion density functions," in *Proc. 11th European Conf. on Comput. Vis.: Part I, ECCV'10*, pp. 171–184, Springer-Verlag, Berlin, Heidelberg (2010).
12. M. Tico, M. Trimeche, and M. Vehvilainen, "Motion blur identification based on differently exposed images," in *Proc. IEEE Int. Conf. Image Process.*, pp. 2021–2024, IEEE Computer Society Press, Los Alamitos, CA (2006).
13. L. Yuan et al., "Image deblurring with blurred/noisy image pairs," *in SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, p. 1, ACM, New York, NY (2007).
14. M. Šorel and F. Šroubek, "Space-variant deblurring using one blurred and one underexposed image," in *Proc. IEEE 16th Int. Conf. on Image Process., ICIP 2009*, IEEE Computer Society Press, Los Alamitos, CA (2009).
15. M. Ben-Ezra and S. K. Nayar, "Motion-based motion deblurring," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(6), 689–698 (2004).
16. Y.-W. Tai et al., "Correction of spatially varying image and video motion blur using a hybrid camera," *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(6), 1012–1028 (2010).
17. N. Joshi et al., "Image deblurring using inertial measurement sensors," *ACM Trans. Graph.* **29**(4), 30:1–30:9 (2010).
18. R. Liu and J. Jia, "Reducing boundary artifacts in image deconvolution," in *15th IEEE Int. Conf. on Image Process., ICIP 2008*, pp. 505–508 (2008).
19. M. Sorel, "Removing boundary artifacts for real-time iterated shrinkage deconvolution," *IEEE Trans. Image Process.* **21**(4), 2329–2334 (2012).



Ondřej Šindelář received an MS degree in computer science from Charles University in Prague, Czech Republic, in 2012. He is currently working on improvement of the project described in this article in collaboration with Dr. Šroubek and the Institute of Information Theory and Automation in Prague.

Filip Šroubek received an MS degree in computer science from the Czech Technical University, Prague, Czech Republic, in 1998 and PhD degree in computer science from Charles University, Prague, Czech Republic, in 2003. From 2004 to 2006, he was on a postdoctoral position in the Instituto de Optica, CSIC, Madrid, Spain. In 2010 and 2011, he was the Fulbright Visiting Scholar at the University of California, Santa Cruz. He is currently with the Institute of Information Theory and Automation and partially also with the Institute of Photonics and Electronics, where both institutes are part of the Academy of Sciences of the Czech Republic. He is an author of seven book chapters and over 80 journal and conference papers on image fusion, blind deconvolution, super-resolution, and related topics.