COMPSTAT 2006, Rome
17th COMPSTAT Symposium of the IASC

Blind Superresolution

Filip Šroubek[1,2], Gabriel Cristóbal[2], and Jan Flusser[1]

[1] ÚTIA, Academy of Sciences of the Czech Republic, sroubekf@utia.cas.cz
[2] Instituto de Óptica, CSIC, gabriel@optica.csic.es

Summary. This paper presents a unifying 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. Experiments on real data illustrate the robustness and utilization of the proposed technique.

Key words: Blind deconvolution, superresolution, multiframe image restoration, MIMO

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 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) $v(x, y)$. Finally, the CCD discretizes the images and produces digitized noisy image $z(i, j)$ (frame). We refer to $z(i, j)$ 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

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

COMPSTAT 2006, Rome
17th COMPSTAT Symposium of the IASC

where n 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* $g(x, y)$ 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 $z(i, j)$ is extremely ill-posed. To partially overcome this difficulty, we assume that multiple LR observations of the original scene are available. Hence we write

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

where k is the acquisition index 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 $z_k(i, j)$ are multiple outputs. The model is therefore called a single input multiple output (SIMO) model. To our knowledge, this is the most accurate, state-of-the-art model, as it takes all possible degradations into account.

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 sequel, we will refer to this superresolved image as a *high resolution (HR) image* $u(i, j)$. 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 introduction of the volatile blurs brings about a more general and robust technique, with the subpixel shift being a special case thereof.

The acquisition model (2) embraces three distinct cases frequently encountered in literature. First, we face a registration problem, if we want to resolve the geometric degradation W_k. 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 investigation.

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. [23] for a survey. So we can assume in the sequel that the LR images are partially registered and that W_k reduces to a small translation.

The MBD problem has recently attracted considerable attention. First blind deconvolution attempts were based on single-channel formulations, such as in [3, 6, 10, 17]. Kundur *et al.* [9] provide a good overview. The problem

COMPSTAT 2006, Rome
17th COMPSTAT Symposium of the IASC

is extremely ill-posed in the single-channel framework and lacks any solution in the fully blind case. These methods do not exploit the potential of the multichannel framework, i.e., the missing information about the original image in one channel is supplemented by the information in other channels. Research on intrinsically multichannel methods has begun fairly recently; refer to [5, 8, 13, 14, 16, 19] for a survey and other references. Such MBD methods brake the limitations of previous techniques and can recover the blurring functions from the input channels alone. We further developed the MBD theory in [20] by proposing a blind deconvolution method for images, which might be mutually shifted by unknown vectors.

A countless number of papers address the standard SR problem. A good survey is for example in [15]. Maximum likelihood (ML), maximum a posteriori (MAP), the set theoretic approach using POCS, 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 [7, 18], include the shift estimation of the SR process. Other approaches focus on fast implementation [4]. In general, most of the SR techniques assume *a priori* known blurs. However, few exceptions exist. Authors in [12, 22] proposed BSR that can handle parametric PSFs, i.e., PSFs modeled with one parameter. This restriction is unfortunately very limiting for most real applications. To our knowledge, first attempts for BSR with an arbitrary PSF appeared in [21]. The interesting idea proposed therein is the conversion of the SR problem from SIMO to multiple input multiple output (MIMO) using so-called polyphase components. We will adopt the same idea here as well.

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 downsampling operator since this case violates the standard convolution model. On the contrary, state-of-the-art SR techniques achieve remarkable results in 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 or 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. 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 consisting of two simple linear equations.

The rest of the paper is organized as follows. Section 2 defines mathematical formalism in the sequel and the outlined degradation model. The section concludes with a procedure for estimating volatile blurs. A detail description of the BSR algorithm is given in Section 3. Final Section 4 illustrates applicability of the proposed method to real situations.

2 Mathematical Model

To simplify the notation, we will assume only images and PSFs with square support. An extension to rectangular images is straightforward. Let $u(i, j)$ be an arbitrary discrete image of size $U \times U$, then $\mathbf{u}$ denotes an image column vector of size $U^2 \times 1$ and $\mathbf{C}_A\{u\}$ denotes a matrix that performs convolution of u 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 $\mathbf{C}_A\{u\}$ of size $(U + A - 1)^2 \times A^2$ and "valid" convolution $\mathbf{C}^v_A\{u\}$ of size $(U - A + 1)^2 \times A^2$. For further discussion, it is necessary to define a sampling matrix. Let ε denote a positive integer step (downsampling factor) and let $\mathbf{S}^i_M$ be a 1-D sampling matrix of size $(M/\varepsilon) \times M$, where $i = 0, \dots, \varepsilon - 1$ and we assume that M is divisible by ε. Each row of the sampling matrix is a unit vector whose nonzero element is at the appropriate position so that, if the matrix is multiplied by a vector of size M, the result is every ε-th element of the vector starting with the $(i + 1)$-th element. In the 2-D case, the $(M/\varepsilon)^2 \times M^2$ sampling matrix for the image size $M \times M$ is defined by

$$\mathbf{S}^{ij}_M := \mathbf{S}^i_M \otimes \mathbf{S}^j_M \,, \tag{3}$$

where $\otimes$ denotes the matrix direct product (Kronecker product operator). If the size of the sampling matrix is evident from the context, we will omit the subscript M.

Let us assume we have K different LR frames $\{z_k\}$ (each of equal size $Z \times Z$) 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 u of size $U \times U$. The LR frames are linked with the HR image through a series of degradations similar to those between $o(x, y)$ and z_k in (2). First u is geometrically warped ($\mathbf{W}_k$), then it is convolved with an 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

$$\mathbf{z}_k = \mathbf{D}\mathbf{V}_k\mathbf{W}_k\mathbf{u} + \mathbf{n}_k \,, \tag{4}$$

where $\mathbf{n}_k$ is additive noise present in every channel. The decimation matrix $\mathbf{D} = \mathbf{S}\mathbf{G}$ simulates the behavior of digital sensors by performing first convolution with the $G \times G$ sensor PSF (g) 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 [2]. We assume that the subsampling factor (or SR factor, depending on the point of view), denoted by ε, is the same in both directions. Note that ε is a user-defined parameter. If ε is an integer then $\mathbf{S} := \mathbf{S}^{00}_{(\varepsilon Z)}$; see (3). 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 sub-pixel accuracy 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

$$\mathbf{z}_k = \mathbf{D}\mathbf{H}_k\mathbf{u} + \mathbf{n}_k = \mathbf{S}\mathbf{G}\mathbf{H}_k\mathbf{u} + \mathbf{n}_k\,. \tag{5}$$

In our formulation we know the LR images $\{z_k\}$ and we want to estimate the HR image u supposing that only $\mathbf{G}$ is known on the right hand side of the equation. To avoid boundary effects, we assume that each observation z_k captures only a part of u. Hence $\mathbf{H}_k$ and $\mathbf{G}$ are "valid" convolution matrices $\mathbf{C}^v_U\{h_k\}$ and $\mathbf{C}^v_{U-H+1}\{g\}$, respectively. The PSFs h_k can be of different size. However, we postulate that they all fit into the given $H \times H$ support.

In the case of $\varepsilon = 1$, the downsampling $\mathbf{S}$ is not present and we face a standard MBD problem that has been solved elsewhere [8, 20]. Here we are interested in the case of $\varepsilon > 1$, when the downsampling occurs. Can we estimate the blurs according to [8] and derive blur regularization as in [20]? The presence of $\mathbf{S}$ prevents us to use the cited results directly. First, we need to rearrange the acquisition model (5) and construct from the LR images z_k a convolution matrix $\mathcal{Z}$ with a predetermined nullity. Then we take the null space of $\mathcal{Z}$ and construct a matrix $\mathcal{N}$, which will contain the correct PSFs h_k in its null space. In the next section, we show how to utilize $\mathcal{N}$ in blur regularization.

Let $E \times E$ be the size of "nullifying" filters. The meaning of this name will be clear later. Define $\mathcal{Z} := [\mathbf{Z}_1, \ldots, \mathbf{Z}_K]$, where $\mathbf{Z}_k := \mathbf{C}^v_E\{z_k\}$ are "valid" convolution matrices. Using (5) without noise, we can express $\mathcal{Z}$ in terms of u, g and h_k as

$$\mathcal{Z} = \mathbf{S}^{00}\mathcal{U}\mathcal{G}\mathcal{H}\,, \tag{6}$$

where

$$\mathcal{H} := [\mathbf{C}_{\varepsilon E}\{h_1\}, \ldots, \mathbf{C}_{\varepsilon E}\{h_K\}] \times (\mathbf{I}_K \otimes (\mathbf{S}^{00}_{\varepsilon E})^T)\,, \tag{7}$$

$\mathcal{G} := \mathbf{C}_{\varepsilon E+H-1}\{g\}$ and $\mathcal{U} := \mathbf{C}^v_{\varepsilon E+H+G-2}\{u\}$. Matrix $\mathbf{I}_K$ denotes an identity matrix of size $K \times K$.

The convolution matrix $\mathcal{G}$ has more rows than columns and therefore it is of full column rank (see proof in [8] for general convolution matrices). We assume that $\mathbf{S}^{00}\mathcal{U}$ has full column rank as well. This is almost certainly true for real images if $\mathcal{U}$ has at least ε-times more rows than columns. Thus $\mathrm{Null}(\mathcal{Z}) \equiv \mathrm{Null}(\mathcal{H})$ and the difference between the number of columns and rows of $\mathcal{H}$ bounds from below the null space dimension, i.e., $\mathrm{nullity}(\mathcal{Z}) \geq KE^2 - (\varepsilon E + H - 1)^2 = N$. Setting $\mathbf{N} := \mathrm{Null}(\mathcal{Z})$, we visualize the null space as

$$\mathbf{N} = \begin{bmatrix} \boldsymbol{\eta}_{1,1} & \cdots & \boldsymbol{\eta}_{1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{K,1} & \cdots & \boldsymbol{\eta}_{K,N} \end{bmatrix}, \tag{8}$$

where $\boldsymbol{\eta}_{kn}$ is the vector representation of the nullifying filter η_{kn} of size $E \times E$, $k = 1, \ldots, K$ and $n = 1, \ldots, N$. Let $\tilde{\eta}_{kn}$ denote upsampled η_{kn} by factor ε, i.e., $\tilde{\eta}_{kn} := (\mathbf{S}^{00}_{\varepsilon E})^T \eta_{kn}$. Then, we define

$$\mathcal{N} := \begin{bmatrix} \mathbf{C}_H\{\tilde{\eta}_{1,1}\} & \ldots & \mathbf{C}_H\{\tilde{\eta}_{K,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_H\{\tilde{\eta}_{1,N}\} & \ldots & \mathbf{C}_H\{\tilde{\eta}_{K,N}\} \end{bmatrix} \tag{9}$$

and conclude that

$$\mathcal{N}\mathbf{h} = \mathbf{0}\,, \tag{10}$$

where $\mathbf{h} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_K^T]^T$. This equation is a potential solution to the blur estimation problem. Unfortunately, since it was derived from (6), which is of the MIMO type, the ambiguity of the solution is high. It has been shown in [11] that the solution of the blind 1-D MIMO case is unique apart from a mixing matrix of input signals. The same holds true here and without providing the proof we state that $\mathrm{nullity}(\mathcal{N}) = \varepsilon^4$.

It is interesting to note that similar derivation is possible for rational SR factors $\varepsilon = p/q$. We downsample the LR images with the factor q, create thus q^2K images and apply thereon the above procedure for the SR factor p.

3 Blind Superresolution

In order to solve the BSR problem, i.e, determine the HR image u and volatile PSFs h_k, we adopt a classical 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

$$E(\mathbf{u}, \mathbf{h}) = \sum_{k=1}^{K} \|\mathbf{D}\mathbf{H}_k\mathbf{u} - \mathbf{z}_k\|^2 + \alpha Q(\mathbf{u}) + \beta R(\mathbf{h})\,. \tag{11}$$

The first term measures the fidelity to the data and emanates from our acquisition model (5). The remaining two are regularization terms with positive weighting constants α and β that attract the minimum of E to an admissible

COMPSTAT 2006, Rome
17th COMPSTAT Symposium of the IASC

set of solutions. The form of E very much resembles the energy we have proposed in [20] for MBD. Indeed, this should not come as a surprise since MBD and SR are related problems in our formulation.

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

$$Q(\mathbf{u}) = \mathbf{u}^T \mathbf{L} \mathbf{u}\,, \tag{12}$$

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 $Q(u) = \int |\nabla u|^2$. Recently, variational integrals $Q(u) = \int \phi(|\nabla u|)$ were proposed, where ϕ is a strictly convex, nondecreasing function that grows at most linearly. Examples of $\phi(s)$ are s (total variation - used in our experiments), $\sqrt{1+s^2}-1$ (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 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 [1]. For the purpose of our discussion it suffices to state that after discretization we arrive again at (12), where this time $\mathbf{L}$ is a positive semidefinite block tridiagonal matrix constructed of values depending on the gradient of u. The rationale behind the choice of $Q(u)$ 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 input images.

The PSF regularization term $R(\mathbf{h})$ directly follows from the conclusions of the previous section. Since the matrix $\mathcal{N}$ in (10) contains the correct PSFs h_k in its null space, we define the regularization term as a least squares fit

$$R(\mathbf{h}) = \|\mathcal{N}\mathbf{h}\|^2 = \mathbf{h}^T \mathcal{N}^T \mathcal{N} \mathbf{h}\,. \tag{13}$$

The product $\mathcal{N}^T\mathcal{N}$ is a positive semidefinite matrix. More precisely, R 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 (11)) it is based solely on the observed LR images. A good practice is to include also a smoothing term $\mathbf{h}^T\mathbf{L}\mathbf{h}$ with a small weight in $R(\mathbf{h})$. This is especially useful in the case of very noisy data.

The complete energy then takes the form

$$E(\mathbf{u},\mathbf{h}) = \sum_{k=1}^{K} \|\mathbf{D}\mathbf{H}_k\mathbf{u} - \mathbf{z}_k\|^2 + \alpha \mathbf{u}^T\mathbf{L}\mathbf{u} + \beta\|\mathcal{N}\mathbf{h}\|^2\,. \tag{14}$$

To find a minimizer of the energy function, we perform alternating minimizations (AM) of E over $\mathbf{u}$ and $\mathbf{h}$. The advantage of this scheme lies in its simplicity. Each term of (14) is quadratic and therefore convex (but not necessarily strictly convex) and the derivatives w.r.t. $\mathbf{u}$ and $\mathbf{h}$ 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., $\nabla_{\mathbf{u}} E = 0$, 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:

$$
\begin{aligned}
\text{step 1)} \quad \mathbf{u}^m &= \arg\min_{\mathbf{u}} E(\mathbf{u}, \mathbf{h}^m) \\
&\Leftrightarrow \Big(\sum_{k=1}^{K} \mathbf{H}_k^T \mathbf{D}^T \mathbf{D} \mathbf{H}_k + \alpha \mathbf{L}\Big)\mathbf{u} = \sum_{k=1}^{K} \mathbf{H}_k^T \mathbf{D}^T \mathbf{z}_k \,, && (15) \\
\text{step 2)} \quad \mathbf{h}^{m+1} &= \arg\min_{\mathbf{h}} E(\mathbf{u}^m, \mathbf{h}) \\
&\Leftrightarrow \big([\mathbf{I}_K \otimes \mathbf{U}^T \mathbf{D}^T \mathbf{D} \mathbf{U}] + \beta \mathcal{N}^T \mathcal{N}\big)\mathbf{h} = [\mathbf{I}_K \otimes \mathbf{U}^T \mathbf{D}^T]\mathbf{z} \,, && (16)
\end{aligned}
$$

where $\mathbf{U} := \mathbf{C}_H^v\{u\}$, $\mathbf{z} := [\mathbf{z}_1^T, \ldots, \mathbf{z}_K^T]^T$ and m is the iteration step. Note that both steps are simple linear equations.

Energy E as a function of both variables $\mathbf{u}$ and $\mathbf{h}$ is not convex due to the coupling of the variables via convolution in the first term of (14). Therefore, it is not guaranteed that the BSR algorithm reaches the global minimum. 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 $\mathbf{u}^m$ to contain values in the admissible range, typically, the range of values of $\mathbf{z}$. It is common to assume that PSF is positive ($h_k \geq 0$) and preserves the image brightness ($\sum h_k = 1$). 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 (16) 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 them in this case since the size of $\mathbf{h}$ is much smaller than the size of $\mathbf{u}$.

The weighting constants α and β depend on the level of noise. If noise increases, α should increase and β should decrease. One can use parameter estimation techniques, such as cross-validation or expectation maximization, 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 contrary, if the algorithm begins to segment the image, we have overestimated the noise level.

4 Experiments

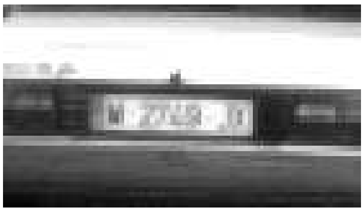
The experimental section demonstrates performance of the proposed method on two real-data sets. We compare the quality of SR reconstruction with two methods: interpolation technique and state-of-the-art SR method. The interpolation technique combines the MBD method proposed in [20] with bilinear interpolation (BI). MBD first removes volatile blurs and then BI of the deconvolved image achieves the desired spatial resolution. The second method, which we will call herein a "standard SR method", is a MAP formulation of the SR problem proposed, e.g., in [7, 18]. This method implements a MAP framework for the joint estimation of image registration parameters (in our case only translation) and the HR image, while assuming only the sensor blur ($\mathbf{G}$) and no volatile blurs. As an image prior we use edge preserving Huber Markov Random Fields.

In both the proposed BSR method and the standard SR method, we set the sensor blur to a Gaussian function of standard deviation $\sigma = 0.35$ (with respect to the scale of LR images). Contrary to the standard SR method, the proposed BSR method is fairly robust to the choice of the Gaussian variance, since it can compensate for the insufficient variance by automatically including the missing factor of Gaussian functions in the volatile blurs.

All images were captured with a standard 5 Mpixel color digital camera (Olympus C5050Z), which has an optical zoom up to $3\times$ and can capture 1.3 fps in a continuous mode. Since in this work we consider only gray-level images, we use the green channels of color photos as LR images.

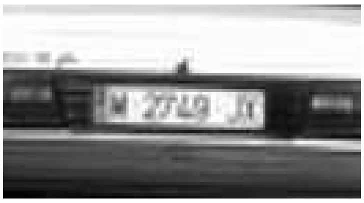
In the first experiment, see Fig. 1(a), we took eight images of a parked car. The shutter speed of the camera was short (1/320s) to minimize possible volatile blurs. We set the SR factor to 2. To compare the quality of reconstruction we acquired one additional image with optical zoom $2\times$ that plays the role of a "ground truth" image; see Fig. 1(b). Since the images contain mild blurs, MBD coupled with BI in Fig. 1(c) does not provide much improvement. On the other hand, the standard SR method in Fig. 1(d) gives results comparable to the ground truth. The proposed BSR algorithm in Fig. 1(e) shows slight improvement over the standard SR method. This is not a surprise, since the main source of degradation in this case is caused by the sensor blur, and both the standard SR method and BSR use the same sensor blur. Indeed, as illustrated in Fig. 1(f), blurs estimated by BSR consist primarily of the sensor blur (Gaussian).

The second experiment takes four images of a car in motion shot with the camera in the continuous mode and longer shutter speed (1/30s). We use the SR factor 5/3. The first frame with sever motion blur is in Fig. 2(a). In this case, the MBD+BI and the standard SR approach provide little improvement in contrast to the proposed BSR algorithm; compare Figs. 2(b), (c), (d) and close-ups in (f). Blurs estimated by BSR are in Fig. 2(e).



(a)

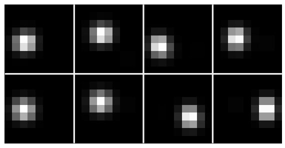
(b)



(c)

(d)

(e)



(f)

Fig. 1: SR of mildly blurred images ($\varepsilon = 2$): (a) One of eight LR image of size 40×70, zero-order interpolation; (b) Image acquired with optical zoom $2\times$, which plays the role of “ground truth”; (c) MBD with BI; (d) Standard SR method; (e) Results of the BSR algorithm with estimated PSFs in (f).

5 Conclusions

The proposed BSR method goes far beyond the standard SR techniques. The introduction of volatile blurs makes the method particularly appealing for real situations. While reconstructing the blurs, we estimate~~s~~ not only the subpixel shifts but also any possible blurs imposed by the acquisition process. To our knowledge, this is the only method that can perform deconvolution and resolution enhancement simultaneously. A possible future extension is into color imaging, which will lead to a powerful demosaicing methodology.

Acknowledgment

This work has been supported by Czech Ministry of Education under the project No. 1M6798555601 (Research Center DAR), the bilateral project

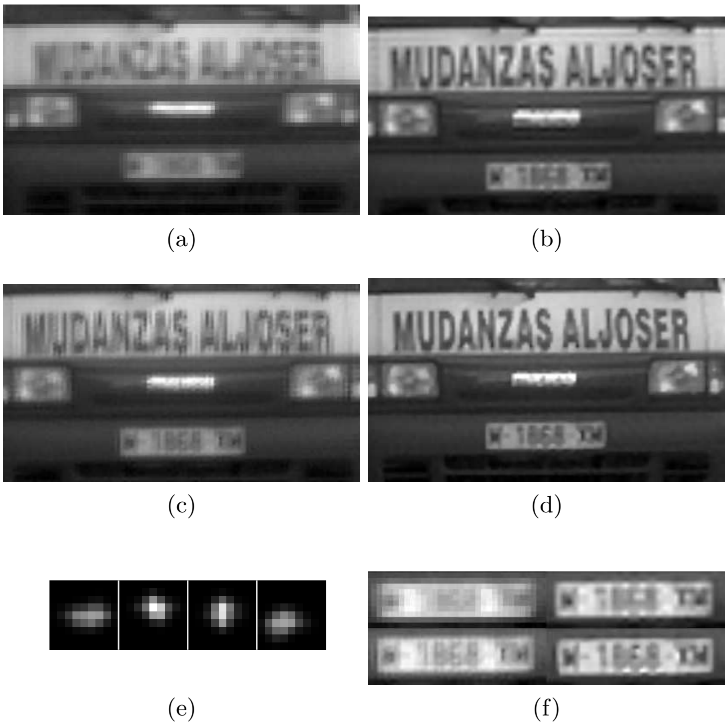


Fig. 2: SR of severely blurred images ($\varepsilon = 5/3$): (a) One of four LR images of size 50×100, zero-order interpolation; (b) MBD with BI; (c) Standard SR method; (d) Results of the proposed BSR algorithm with estimated PSFs in (e); Close-ups of the images in (a), (b) on top and (c), (d) on bottom.

2004CZ0009 of CSIC and Academy of Sciences of the Czech Republic, and the Grant Agency of the Czech Republic under the projects No. 102/04/0155 and No. 202/05/0242. F. Šroubek was also supported by the Spanish States Secretary of Education and Universities fellowship.

References

1. G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing.* Springer Verlag, New York, 2002.
2. D. Capel. *Image Mosaicing and Super-resolution.* Springer, 2004.
3. T.F. Chan and C.K. Wong. Total variation blind deconvolution. *IEEE Trans. Image Processing*, 7(3):370–375, March 1998.
4. S. Farsiu, M.D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Trans. Image Processing*, 13(10):1327–1344, October 2004.

5. G.B. Giannakis and R.W. Heath. Blind identification of multichannel FIR blurs and perfect image restoration. *IEEE Trans. Image Processing*, 9(11):1877–1896, November 2000.
6. M. Haindl. Recursive model-based image restoration. In *Proceedings of the 15th International Conference on Pattern Recognition*, volume III, pages 346–349. IEEE Press, 2000.
7. R.C. Hardie, K.J. Barnard, and E.E. Armstrong. Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Trans. Image Processing*, 6(12):1621–1633, December 1997.
8. G. Harikumar and Y. Bresler. Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms. *IEEE Trans. Image Processing*, 8(2):202–219, February 1999.
9. D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43–64, May 1996.
10. R.L. Lagendijk, J. Biemond, and D.E. Boekee. Identification and restoration of noisy blurred images using the expectation-maximization algorithm. *IEEE Trans. Acoust. Speech Signal Process.*, 38(7):1180–1191, July 1990.
11. T.J. Moore, B.M. Sadler, and R.J. Kozick. Regularity and strict identifiability in mimo systems. *IEEE Trans. Signal Processing*, 50(8):1831–1842, August 2002.
12. N. Nguyen, P. Milanfar, and G. Golub. Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement. *IEEE Trans. Image Processing*, 10(9):1299–1308, September 2001.
13. Hung-Ta Pai and A.C. Bovik. On eigenstructure-based direct multichannel blind image restoration. *IEEE Trans. Image Processing*, 10(10):1434–1446, October 2001.
14. 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*, 12(11):1324–1337, November 2003.
15. S.C. Park, M.K. Park, and M.G. Kang. Super-resolution image reconstruction: A technical overview. *IEEE Signal Proc. Magazine*, 20(3):21–36, 2003.
16. A. Rav-Acha and S. Peleg. Restoration of multiple images with motion blur in different directions. In *IEEE Workshop on Applications of Computer Vision (WACV)*, pages 22–27, 2000.
17. S.J. Reeves and R.M. Mersereau. Blur identification by the method of generalized cross-validation. *IEEE Trans. Image Processing*, 1(3):301–311, July 1992.
18. C.A. Segall, A.K. Katsaggelos, R. Molina, and J. Mateos. Bayesian resolution enhancement of compressed video. *IEEE Trans. Image Processing*, 13(7):898–911, July 2004.
19. F. Šroubek and J. Flusser. Multichannel blind iterative image restoration. *IEEE Trans. Image Processing*, 12(9):1094–1106, September 2003.
20. F. Šroubek and J. Flusser. Multichannel blind deconvolution of spatially misaligned images. *IEEE Trans. Image Processing*, 14(7):874–883, July 2005.
21. Wirawan, P. Duhamel, and H. Maitre. Multi-channel high resolution blind image restoration. In *Proc. IEEE ICASSP*, pages 3229–3232, 1999.
22. N.A. Woods, N.P. Galatsanos, and A.K. Katsaggelos. EM-based simultaneous registration, restoration, and interpolation of super-resolved images. In *Proc. IEEE ICIP*, volume 2, pages 303–306, 2003.
23. B. Zitová and J. Flusser. Image registration methods: A survey. *Image and Vision Computing*, 21:977–1000, 2003.