9

Multiframe blind deconvolution coupled with frame registration and resolution enhancement

Filip Šroubek, Jan Flusser

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 4, Praha 8, 182 08, Czech Republic
e-mail: (sroubekf,flusser)@utia.cas.cz

Gabriel Cristóbal

Instituto de Óptica
Consejo Superior de Investigaciones Científicas (CSIC)
Serrano 121, 28006 Madrid, Spain
email: gabriel@optica.csic.es

9.1   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) $v[x,y]$. 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

$$g[i,j] = D((v * o[W(n_1, n_2)])[x,y]) + n[i,j]\,, \tag{9.1}$$

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 an 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

$$g_k[i,j] = D((v_k * o[W_k(n_1, n_2)])[x,y]) + n_k[i,j]\,, \tag{9.2}$$

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. 9.1 illustrates 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.

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 $f[i,j]$*. 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 leads to a more general and robust technique, with the subpixel shift being a special case thereof.

The acquisition model in Eq. (9.2) 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
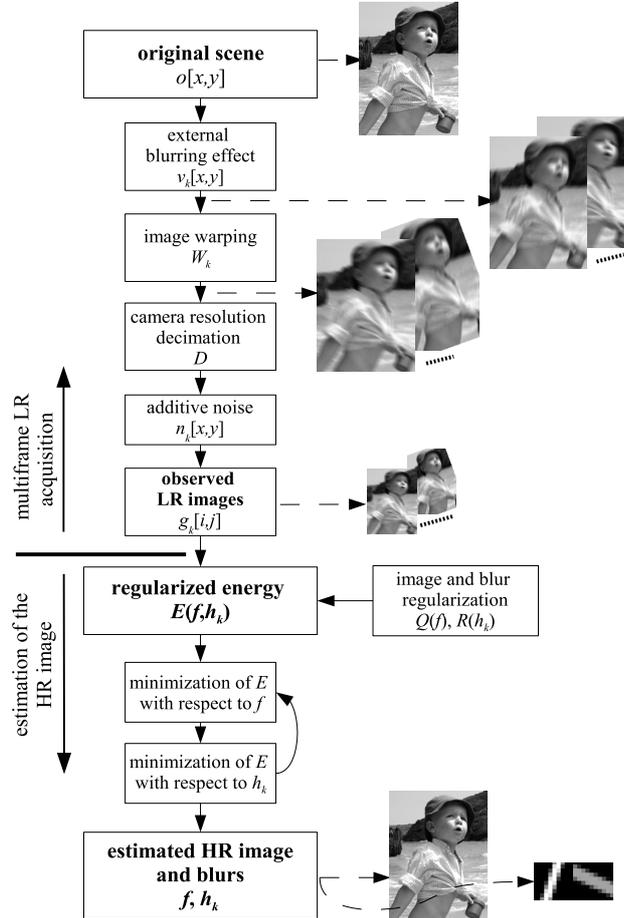
FIGURE 9.1: Low-resolution acquisition (top) and reconstruction flow (bottom).

consider all three cases at once. We are then confronted with a problem of *blind superresolution* (BSR), which is the subject of this investigation. The approach presented in this chapter 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. [31] for a survey. We will assume in the sequel that the LR images are roughly registered and that W_k's 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 [14, 21, 4, 9]. Kundur *et al.* [12, 13] provide a good overview. The problem is extremely ill-posed in the single-channel framework and cannot be resolved in the 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 [11, 8, 18, 19, 24] 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 [25] 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 can be found for example in [20, 7]. Maximum likelihood (ML), maximum a posteriori (MAP), the set theoretic approach using POCS (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 [10, 22, 29], include the shift estimation of the SR process. Other approaches focus on fast implementation [6], space-time SR [23] or SR of compressed video [22]. In general, most of the SR techniques assume *a priori* known blurs. However, few exceptions exist. Authors in [17, 28] 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 [27, 30]. 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. Other preliminary results of the BSR problem with focus on fast calculation are given in [2], 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 downsampling operator, which 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 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. How-

ever, 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. 9.1), consisting of two simple linear equations.

The rest of the chapter is organized as follows. Section 9.2 outlines the degradation model. Section 9.3 reformulates the degradation model using polyphase formalism, which we utilize in the next section and develop a procedure for the volatile blur estimation. These results effortlessly blend in a regularization term of the BSR algorithm as described in Section 9.5. Finally, Section 9.6 illustrates applicability of the proposed method to real situations.

9.2 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 a kernel of size $A \times A$, i.e., $\mathbf{C}_A\{f\}\mathbf{k}$ is the vector form of $f * k$, where k is 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 case the convolution matrix is a Toeplitz-block-Toeplitz (TBT) matrix. In the sequel 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 (9.2). First f 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

$$\mathbf{g}_k = \mathbf{D}\mathbf{V}_k\mathbf{W}_k\mathbf{f} + \mathbf{n}_k \,, \tag{9.3}$$

where $\mathbf{n}_k$ is additive noise present in every channel. The decimation matrix $\mathbf{D} = \mathbf{SU}$ simulates the behavior of digital sensors by performing first 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 [3]. A physical interpretation of the sensor blur is that the sensor is of finite size and it integrates impinging light over its surface. The sensitvity of the sensor is the highest in the middle and decreases towards its borders with 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 sub-pixel 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

$$\mathbf{g}_k = \mathbf{DH}_k\mathbf{f} + \mathbf{n}_k = \mathbf{SUH}_k\mathbf{f} + \mathbf{n}_k \, . \tag{9.4}$$

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 [11, 25]. Here we are interested in the case of $\varepsilon > 1$, when the downsampling occurs. Can we estimate the blurs like in the $\varepsilon = 1$ case? The presence of $\mathbf{S}$ prevents us to use the cited results directly. In the next section we use the polyphase formulation and transfer the problem from SIMO to MIMO. We then show that conclusions obtained for MBD apply here in a slightly modified form as well.

9.3 Polyphase Formulation

Polyphase formalism is an elegant way how to rewrite the acquisition model and thus get a better insight into BSR. First we will assume integer SR factors, for which the model is simple and easy to understand. Then we will generalize it for rational SR factors, for which we will take the full advantage of polyphase formalism. It will allow us to formulate the model for rational factors using a combination of integer factors.

Before we proceed, it is necessary to define precisely the sampling matrix $\mathbf{S}$. Let $\mathbf{S}_i^\varepsilon$ denote a 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 nonzero element is at such position that, if the matrix multiplies an arbitrary vector b, the result of the product is every ε-th element of b starting from b_i. 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

$$\mathbf{S}_{i,j}^\varepsilon := \mathbf{S}_i^\varepsilon \otimes \mathbf{S}_j^\varepsilon \,. \tag{9.5}$$

If the starting index (i, j) will be $(1, 1)$ we will omit the subscript and simply write $\mathbf{S}^\varepsilon$. Note that the transposed matrix $(\mathbf{S}^\varepsilon)^T$ behaves as an upsampling operator that interlaces the original samples with $(\varepsilon - 1)$ zeros. Now, we are ready to define *polyphase components* of an image $f[x, y]$ as

$$\mathbf{f}^{ij} := \mathbf{S}_{i,j}^\varepsilon \mathbf{f} \,, \tag{9.6}$$

which is equivalent to

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

Therefore, each image breaks into ε^2 distinct polyphase components (downsampled versions of the image); see Fig. 9.2. We will refer to this decomposition as a *polyphase decomposition*, and write $\mathbf{P}^\varepsilon \mathbf{f}$, where

$$\mathbf{P}^\varepsilon := [(\mathbf{S}_{1,1}^\varepsilon)^T, \ldots, (\mathbf{S}_{\varepsilon,1}^\varepsilon)^T, (\mathbf{S}_{1,2}^\varepsilon)^T, \ldots, (\mathbf{S}_{\varepsilon,\varepsilon}^\varepsilon)^T]^T \,. \tag{9.7}$$

Similar decomposition was proposed in [15]. Note that $\mathbf{P}$ is a permutation matrix and therefore $\mathbf{P}^T \mathbf{P} = \mathbf{P} \mathbf{P}^T = \mathbf{I}$. We first derive polyphase formulation for integer SR factors and then using a simple trick we extend it to rational ones.

9.3.1   Integer downsampling factor

Let us consider a simple convolution equation

$$\mathbf{g} = \mathbf{H} \mathbf{f} \,, \tag{9.8}$$

and explore the benefits of the polyphase decomposition. Multiplying by $\mathbf{P}^\varepsilon$, we get

$$[\mathbf{P}^\varepsilon \mathbf{g}] = [\mathbf{P}^\varepsilon \mathbf{H} (\mathbf{P}^\varepsilon)^T][\mathbf{P}^\varepsilon \mathbf{f}] \,. \tag{9.9}$$

The permutation matrix $\mathbf{P}^\varepsilon$ decomposes an image into ε^2 polyphase components, and in our case, $\mathbf{P}^\varepsilon \mathbf{g} = [(\mathbf{g}^{11})^T, \ldots, (\mathbf{g}^{\varepsilon\varepsilon})^T]^T$ and $\mathbf{P}^\varepsilon \mathbf{f} = [(\mathbf{f}^{11})^T, \ldots, (\mathbf{f}^{\varepsilon\varepsilon})^T]^T$. For the next discussion it is fundamental to make the observation that $[\mathbf{P}^\varepsilon \mathbf{H} (\mathbf{P}^\varepsilon)^T]$ consists of $\varepsilon^2 \times \varepsilon^2$ blocks. Each block retains the TBT shape of $\mathbf{H}$ but performs convolution with one polyphase component of h.
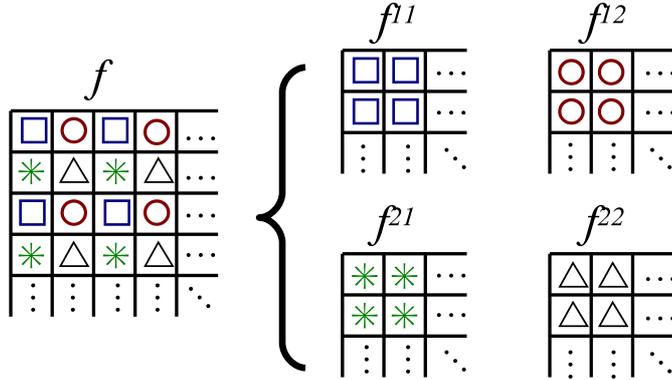
FIGURE 9.2: Polyphase decomposition for $\varepsilon = 2$: Original image f decomposes into 4 downsampled images.

We see that (9.9) is just a permutation of rows and columns of (9.8). The advantage of the polyphase formulation resides in the fact that downsampling is equivalent to a section of (9.9) that corresponds to one polyphase component. We conclude this part by reformulating the acquisition model (9.4) using polyphase components and obtain

$$\mathbf{g}_k = \mathbf{S}^\varepsilon \mathbf{U}\mathbf{H}_k \mathbf{f} + \mathbf{n}_k = [\mathbf{S}^\varepsilon \mathbf{U}\mathbf{H}_k (\mathbf{P}^\varepsilon)^T][\mathbf{P}^\varepsilon \mathbf{f}] + \mathbf{n}_k \,, \qquad (9.10)$$

for $k = 1, \dots, K$. Instead of $\mathbf{S}^\varepsilon = \mathbf{S}_{1,1}^\varepsilon$ one can use any $\mathbf{S}_{i,j}^\varepsilon$. However, they are all equivalent from the reconstruction point of view as they correspond to different translations of the HR image f. In the introduction we regarded the acquisition model as SIMO, with one input channel f and K output channels g_k. Under closer examination of the above polyphase formulation, one can see that $[\mathbf{S}^\varepsilon \mathbf{U}\mathbf{H}_k (\mathbf{P}^\varepsilon)^T]$ consists of $1 \times \varepsilon^2$ convolution matrices and that in reality the model is of the MIMO type with ε^2 input channels (polyphase components of f) and K output channels g_k.

9.3.2 Rational downsampling factor

Integer SR factors are too limiting. From the practical point of view, we would like to have non-integer SR factors as well. We can extend the above results to 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 output (downsampled-LR) images with the sampling frequency 1. Now, to obtain the HR image from the
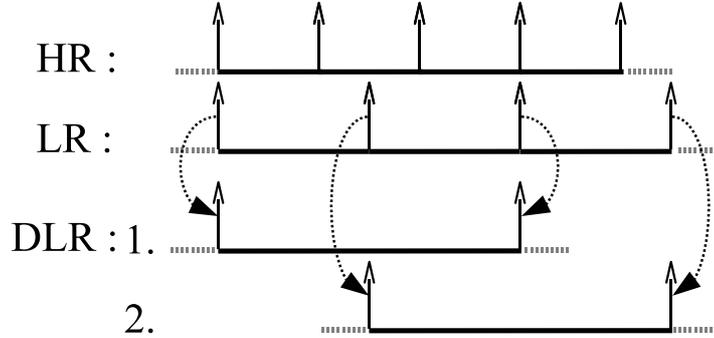
FIGURE 9.3: Rational downsampling $\varepsilon = 3/2$ in 1-D: We have LR signals (middle row) with the sampling frequency 2 and we want to obtain a HR signal (top row) with the sampling frequency 3. We convert this scenario to the one with the integer SR factor by considering every second sample of the LR signal and thus creating from each LR signal two signals (bottom row) of half size. These downsampled-LR signals are then used in the SR problem with the integer factor 3.

downsampled-LR images, we must solve a SR problem with the integer factor $\varepsilon = p$ and not with the rational one as before. In other words, in order to get 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 illustrative diagram of the process in 1-D for $\varepsilon = 3/2$ is given in Fig. 9.3.

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 $\mathbf{SU}$ 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 q-th pixels (in both directions x and y) of the LR image and therefore we have q^2 distinct PSF discretizations. To better understand this concept, see the configuration for $\varepsilon = 3/2$ in Fig. 9.4.

Similarly to (9.10), we reformulate the acquisition model (9.4) using polyphase components and write

$$[\mathbf{P}^q\mathbf{g}_k] = \left[\begin{bmatrix} \mathbf{S}^p\mathbf{U}_{1,1} \\ \vdots \\ \mathbf{S}^p\mathbf{U}_{q,q} \end{bmatrix} \mathbf{H}_k(\mathbf{P}^p)^T \right] [\mathbf{P}^p\mathbf{f}] + \mathbf{n}_k \,, \qquad (9.11)$$

where each $\mathbf{U}_{i,j}$ performs convolution with one of the q^2 discretizations of the sensor PSF u. We see that the rational and integer SR factors lead to similar expressions. Only in the rational case, the resulting MIMO problem has Kq^2
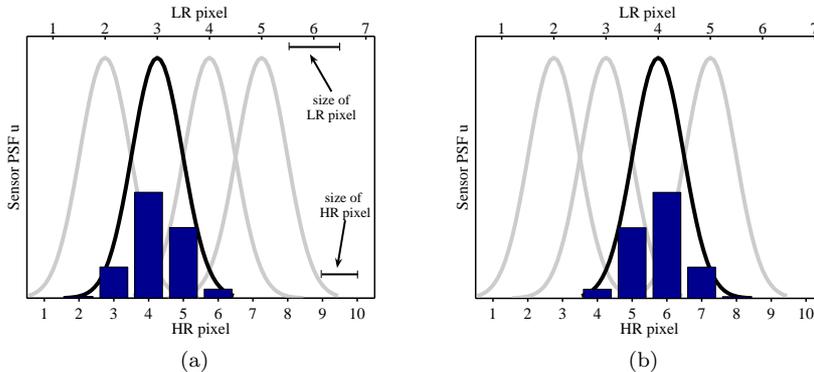
FIGURE 9.4: Sensor PSF discretization for the SR factor 3/2: Different discretizations of the PSF reside in a varying relative shift between LR and HR pixels. If the LR pixels is $1.5\times$ the size of the HR pixel, then two distinct discretizations (a) and (b) show up in 1-D (in 2-D we have 4 discretizations). The plotted curves depict the sensor PSF in the continuous domain at different locations and the bar plots its discrete version.

output channels and p^2 input channels.

9.4   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 [11, 8], 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 of 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 [8]. The channel co-primeness 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. We refer the reader to [11] 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 show how to generalize the results for integer and rational downsampling factors. In the following, we will assume that noise n is not present and wait till the next section 9.5, where we will address noise appropriately.

9.4.1 The MBD case

The downsampling matrix $\mathbf{S}$ is not present in (9.4) 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

$$\|g_i * h_j - g_j * h_i\|^2 = 0 \,. \tag{9.12}$$

Considering all possible pairs of blurs, we can arrange the above relation into one system

$$\mathcal{N}'\mathbf{h} = \mathbf{0} \,, \tag{9.13}$$

where $\mathbf{h} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_K^T]^T$ and $\mathcal{N}'$ is constructed solely by matrices that perform convolution with g_k. 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 $\mathcal{N}'$ is exactly 1 if the blur size is correctly estimated. By applying SVD (singular value decomposition), we recover precisely the blurs except to 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. If the blur size is underestimated, the above equation has no solution. If the blur size is overestimated, then nullity($\mathcal{N}'$) = $(\hat{H}_1 - H + 1)(\hat{H}_2 - H + 1)$.

9.4.2 The BSR case

A naive approach, e.g. proposed in [26, 5], is to modify (9.13) in the MBD case by applying downsampling and formulating the problem as

$$\min_{\mathbf{h}} \|\mathcal{N}'[\mathbf{I}_K \otimes \mathbf{S}^\varepsilon \mathbf{U}]\mathbf{h}\|^2 \,, \tag{9.14}$$

where $\mathbf{I}_K$ is the $K \times K$ identity matrix. One can easily verify that the condition in (9.12) 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 (9.14) do not have to correspond to the correct blurs. However, if we use the MIMO polyphase formulation in (9.10) or in (9.11), we will show that the reconstruction of the volatile PSFs h_k is possible even in the BSR case. We will see that for the integer SR factors ε, some ambiguity in the solution of h_k is inevitable, irrespective of the knowledge of the sensor blur u. For the rational ε, a solution is possible if and only if the sensor blur u is known and surprisingly the solution is without any ambiguity. Note that for correct reconstruction of the HR image, the sensor blur is necessary in any case.

First, we need to rearrange the acquisition model (9.4) 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. 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. Using (9.10) without noise, we can express $\mathcal{G}$ in terms of f, u and h_k as

$$\mathcal{G} = \mathbf{S}^\varepsilon \mathbf{F} \mathbf{U} \mathcal{H} \,, \tag{9.15}$$

where

$$\mathcal{H} := [\mathbf{C}_{\varepsilon E}\{h_1\}(\mathbf{S}^\varepsilon)^T, \ldots, \mathbf{C}_{\varepsilon E}\{h_K\}(\mathbf{S}^\varepsilon)^T] \,, \tag{9.16}$$

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

The convolution matrix $\mathcal{U}$ has more rows than columns and therefore it is of full column rank (see proof in [11] 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 $\text{Null}(\mathcal{G}) \equiv \text{Null}(\mathcal{H})$ and the difference between the number of columns and rows of $\mathcal{H}$ bounds from below the null space dimension, i.e.,

$$\text{nullity}(\mathcal{G}) \geq KE^2 - (\varepsilon E + H - 1)^2 \,. \tag{9.17}$$

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

$$\mathbf{N} = \begin{bmatrix} \boldsymbol{\eta}_{1,1} & \cdots & \boldsymbol{\eta}_{1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{K,1} & \cdots & \boldsymbol{\eta}_{K,N} \end{bmatrix} \,, \tag{9.18}$$

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., $\tilde{\eta}_{kn} := (\mathbf{S}^\varepsilon)^T \eta_{kn}$. Then, we define

$$\mathcal{N} := \begin{bmatrix} \mathbf{C}_H\{\tilde{\eta}_{1,1}\} & \cdots & \mathbf{C}_H\{\tilde{\eta}_{K,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_H\{\tilde{\eta}_{1,N}\} & \cdots & \mathbf{C}_H\{\tilde{\eta}_{K,N}\} \end{bmatrix} \tag{9.19}$$

and conclude that

$$\mathcal{N}\mathbf{h} = \mathbf{0} \,, \tag{9.20}$$

where $\mathbf{h}^T = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$. We have arrived to an equation that is of the same form as (9.13) in the MBD case. Here we have the solution to the blur estimation problem for the BSR case. However, since it was derived from (9.10), which is of the MIMO type, the ambiguity of the solution is higher. It has been shown in [16] 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 as well. Without proofs we provide the following statements. For the correct blur size, $\text{nullity}(\mathcal{N}) = \varepsilon^4$. For the underestimated blur size, (9.20) has no solution. For the overestimated blur size $\hat{H}_1 \times \hat{H}_2$, $\text{nullity}(\mathcal{N}) = \varepsilon^2(\hat{H}_1 - H + \varepsilon)(\hat{H}_2 - H + \varepsilon)$.

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. Nevertheless, Section 9.5 shows that $\mathcal{N}$ plays an important role in the regularized restoration algorithm and its ambiguity is not a serious drawback.

It remains to describe the procedure for the rational downsampling factors $\varepsilon = p/q$. The analysis starts by rearranging the acquisition model in (9.11). Again, let $E \times E$ be the size of nullifying filters. In the previous section, we have seen that there are q^2 distinct discretizations of the sensor PSF u that depend on the relative shift between HR and LR pixels. Let $u_{i,j}$ $(i, j = 1, \ldots, q)$ denote such discretizations. We define "full" convolution matrix $\mathbf{U}_{i,j} := \mathbf{C}_{pE+H-1}\{u_{i,j}\}$ and "valid" convolution matrices $\mathbf{F} := \mathbf{C}_{pE+U+H-2}^v\{f\}$, $\mathbf{G}_k := \mathbf{C}_{qE}^v\{g_k\}$. Then define

$$\mathcal{G} := [\mathbf{G}_1, \ldots, \mathbf{G}_K],$$
$$\mathcal{H}' := [\mathbf{I}_{q^2} \otimes \mathbf{C}_{pE}\{h_1\}, \ldots, \mathbf{I}_{q^2} \otimes \mathbf{C}_{pE}\{h_K\}][\mathbf{I}_{Kq^2} \otimes (\mathbf{S}^\varepsilon)^T]$$

The degradation model for the rational SR factor $\varepsilon = p/q$ becomes

$$\mathbf{S}^q \mathcal{G}[\mathbf{I}_K \otimes \mathbf{P}^q] = \mathbf{S}^p \mathbf{F}[\mathbf{U}_{1,1}, \ldots, \mathbf{U}_{q,q}]\mathcal{H}' \tag{9.21}$$

The integer SR factor is a special case of this equation. By setting $q = 1$ we obtain (9.15).

In analogy with the derivation steps for the integer case, we proceed as follows. Set $\mathbf{N} := \text{Null}(\mathbf{S}^q \mathcal{G})$. The size of $\mathbf{N}$ is $K(qE)^2 \times N$, where we assume $N \geq K(qE)^2 - (pE + H + U - 1)^2 > 0$. We visualize the null space as

$$\mathbf{N} = \begin{bmatrix} \boldsymbol{\eta}_{1,1} & \cdots & \boldsymbol{\eta}_{1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{q^2,1} & \cdots & \boldsymbol{\eta}_{q^2,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{Kq^2,1} & \cdots & \boldsymbol{\eta}_{Kq^2,N} \end{bmatrix}, \tag{9.22}$$

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

$$\mathcal{N} := \begin{bmatrix} \mathbf{C}_{U+H-1}\{\tilde{\eta}_{1,1}\} & \cdots & \mathbf{C}_{U+H-1}\{\tilde{\eta}_{Kq^2,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{U+H-1}\{\tilde{\eta}_{1,N}\} & \cdots & \mathbf{C}_{U+H-1}\{\tilde{\eta}_{Kq^2,N}\} \end{bmatrix} \times \mathbf{I}_K \otimes \begin{bmatrix} \mathbf{C}_H\{u_{1,1}\} \\ \vdots \\ \mathbf{C}_H\{u_{q,q}\} \end{bmatrix} \tag{9.23}$$

and we conclude that

$$\mathcal{N}\mathbf{h} = \mathbf{0}. \tag{9.24}$$

The presence of shifted versions of u eliminates any ambiguity of the solution and we can prove that for the correctly estimated blur size the nullity of $\mathcal{N}$ is 1.

While this conclusion may appear optimistic, one should realize an important detail that distinguishes $\mathcal{N}$ for the rational factors from $\mathcal{N}$ for the integer factors. The matrix $\mathcal{N}$ in the integer case does not depend on u and therefore the reconstruction of h_k, though ambiguous, can be carried out even without the knowledge of the sensor PSF. On the other hand, $\mathcal{N}$ in the rational case contains q^2 distinct discretizations of the sensor PSF and the reconstruction of h_k can fail if the sensor PSF is incorrectly estimated.

Another interesting 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 (9.17) 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.

9.5   Blind Superresolution

In order to solve the BSR problem, i.e, determine the HR image f 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

$$E(\mathbf{f}, \mathbf{h}) = \sum_{k=1}^{K} \|\mathbf{DH}_k\mathbf{f} - \mathbf{g}_k\|^2 + \alpha Q(\mathbf{f}) + \beta R(\mathbf{h})\,. \qquad (9.25)$$

The first term measures the fidelity to the data and emanates from our acquisition model (9.4). 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 [25] 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

$$Q(\mathbf{f}) = \mathbf{f}^T\mathbf{L}\mathbf{f}\,, \qquad (9.26)$$

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 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 (9.26), 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 (9.20) (integer factor) or in (9.24) (rational factor) contains the correct PSFs h_k in its null space, we define the regularization term as a least-squares fit

$$R(\mathbf{h}) = \|\mathcal{N}\mathbf{h}\|^2 = \mathbf{h}^T\mathcal{N}^T\mathcal{N}\mathbf{h}\,. \qquad (9.27)$$

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 (9.25)) 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 to overcome the higher nullity of integer-factor $\mathcal{N}$.

The complete energy then takes the form

$$E(\mathbf{f},\mathbf{h}) = \sum_{k=1}^{K} \|\mathbf{D}\mathbf{H}_k\mathbf{f} - \mathbf{g}_k\|^2 + \alpha\mathbf{f}^T\mathbf{L}\mathbf{f} + \beta_1\|\mathcal{N}\mathbf{h}\|^2 + \beta_2\mathbf{h}^T\mathbf{L}\mathbf{h}\,. \qquad (9.28)$$

To find a minimizer of the energy function, we perform alternating minimizations (AM) of E over $\mathbf{f}$ and $\mathbf{h}$. The advantage of this scheme lies in its simplicity. Each term of (9.28) is quadratic and therefore convex (but not necessarily strictly convex) and the derivatives w.r.t. $\mathbf{f}$ 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., $\nabla_{\mathbf{f}}E = 0$, 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) $\quad \mathbf{f}^m \quad = \arg\min_{\mathbf{f}} E(\mathbf{f}, \mathbf{h}^m)$

$$\Leftrightarrow (\sum_{k=1}^{K} \mathbf{H}_k^T\mathbf{D}^T\mathbf{D}\mathbf{H}_k + \alpha\mathbf{L})\mathbf{f} = \sum_{k=1}^{K} \mathbf{H}_k^T\mathbf{D}^T\mathbf{g}_k\,, \qquad (9.29)$$

step 2) $\quad \mathbf{h}^{m+1} = \arg\min_{\mathbf{h}} E(\mathbf{f}^m, \mathbf{h})$

$$\Leftrightarrow ([\mathbf{I}_K \otimes \mathbf{F}^T\mathbf{D}^T\mathbf{D}\mathbf{F}] + \beta_1\mathcal{N}^T\mathcal{N} + \beta_2\mathbf{L})\mathbf{h} = [\mathbf{I}_K \otimes \mathbf{F}^T\mathbf{D}^T]\mathbf{g}\,, \qquad (9.30)$$

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. Note that both steps consist of simple linear equations.

Energy E as a function of both variables $\mathbf{f}$ and $\mathbf{h}$ is not convex due to the coupling of the variables via convolution in the first term of (9.28). 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 $\mathbf{f}^m$ to contain values in the admissible range, typically, the range of values of $\mathbf{g}$. It is common to assume that PSF is positive ($h_k \geq 0$) 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 (9.30) 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 $\mathbf{h}$ 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 [17] or expectation maximization [15], 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.

9.6   Experiments

This section consists of two parts. In the first one, a set of experiments on synthetic data evaluate performance of the BSR algorithm[1] with respect to noise and different regularization terms $R(\mathbf{h})$. The second part demonstrates the applicability of the proposed method to real data and evaluates performance under different input scenarios. Moreover, we compare the reconstruction quality with other two methods: one interpolation technique and one state-of-the-art SR approach. A brief description of these methods follows later.

In all 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

---

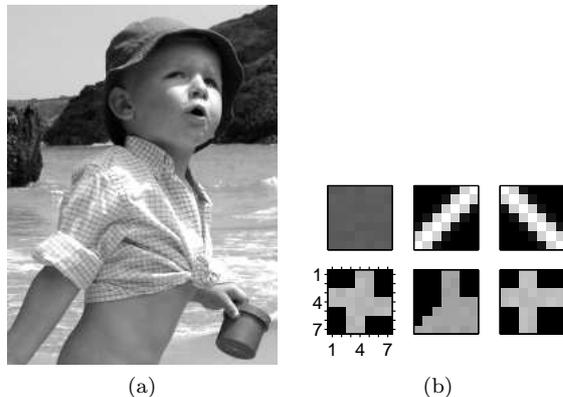[1]The BSR algorithm is implemented in Matlab v7.1 and is available on request.

(a)   (b)

FIGURE 9.5: Simulated data: (a) original 270×200 image; (b) six 7×7 volatile PSFs used to blur the original image.

the Gaussian variance, since it can compensate for the insufficient variance by automatically including the missing factor of Gaussian functions in the volatile blurs.

Another potential pitfall that we have take into consideration is a feasible range of SR factors. Clearly, as the SR factor ε increases we need more LR images and the stability of BSR decreases. 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; see Section 9.3.2. 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.

9.6.1   Simulated data

First, let us demonstrate the BSR performance with a simple experiment. A 270×200 image in Fig. 9.5(a) blurred with six masks in Fig. 9.5(b) and downsampled with factor 2 generated six LR images. Using the LR images as an input, we estimated the original HR image with the proposed BSR algorithm for $\varepsilon = 1.25$ and 1.75. In Fig. 9.6 one can compare the results printed in their original size. The HR image for $\varepsilon = 1.25$ (Fig. 9.6(b)) has improved significantly on the LR images due to deconvolution, however some details, such as the shirt texture, exhibit artifacts. For the SR factor 1.75, the reconstructed image in Fig. 9.6(c) is almost perfect.

Next, we would like to compare performance of different matrices $\mathcal{N}$ inside the blur regularization term $R(\mathbf{h})$ and robustness of the BSR algorithm to noise. Section 9.4 has shown that two distinct approaches exist for blur estimation. Either we use the naive approach in (9.14) that directly utilizes
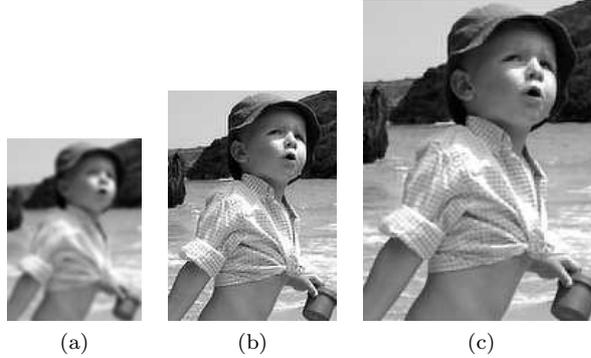
FIGURE 9.6: BSR of simulated data: (a) one of six LR images with the downsampling factor 2; (b) BSR for $\varepsilon = 1.25$; (c) BSR for $\varepsilon = 1.75$. The shirt texture shows interference for the SR factor 1.25 but becomes well reconstructed for the SR factor 1.75.

the MBD formulation, or we apply the intrinsically SR approach (polyphase formulation) proposed in (9.20) and (9.24) for integer and rational factors, respectively. Depending on the approach, the nullity of $\mathcal{N}$ varies, which influences the shape of the blur regularization functional. We considered three distinct scenarios with the original image and PSFs in Fig. 9.5. In the first one, we downsampled the blurred images with the integer factor 2 and performed BSR for $\varepsilon = 2$ using the naive approach inside blur regularization $R(\mathbf{h})$. In the second scenario, we again downsampled the images with the factor 2 and performed BSR for $\varepsilon = 2$ but utilized the polyphase approach for integer factors. The last scenario simulates a situation with rational SR factors. We downsampled the images with the rational factor 7/4 (= 1.75) and applied BSR with $\varepsilon = 7/4$ and with the polyphase approach for rational factors. In order to evaluate also the noise robustness, we added white Gaussian noise to the LR images with SNR 50dB and 30dB. Note that the signal-to-noise ratio is defined as SNR $= 10 \log(\sigma_f^2/\sigma_n^2)$, where σ_f and σ_n are the image and noise standard deviations, respectively. The BSR algorithm ran without the smoothing term on volatile blurs ($\beta_2 = 0$, refer to discussion in Section 9.5) to avoid artificial enhancement of blurs and to study solely the effect of different matrices $\mathcal{N}$.

Results are not evaluated with any measure of reconstruction quality, such as mean-square errors or peak signal to noise ratios. Instead we print the results and leave the comparison to a human eye as we believe that in this case the visual assessment is the only reasonable method. Estimated HR images and volatile blurs for three scenarios and two levels of noise are in Figs. 9.7 and 9.8. For 50dB (low noise), the performance strongly depend on the applied regularization. If we use the naive approach in the first scenario, the estimated PSFs are inaccurate and hence the reconstructed HR image contains many

artefacts as one can see in Fig. 9.7(a). In the second scenario (Fig. 9.7(b)), the reconstructed PSFs resemble the original ones but exhibit patch-like patterns due to higher nullity of $\mathcal{N}$, which is 16 for the integer SR factor 2. The patch-like pattern emanates from our inability to determine the mixing matrix for polyphase components. However, the reconstructed HR images display quality comparable to the original image. The third scenario (Fig. 9.7(c)) with the rational SR factor 1.75 provides the most accurate estimates of the PSFs (the nullity is 1) but for the HR image the improvement on the SR factor 2 is negligible. Clearly, more strict and accurate regularization terms improve HR images, yet as the noise level increases the performance boost diminishes. In the case of 30dB in Fig. 9.8, the reconstructed HR images are very similar for all three scenarios, though the PSFs still differ a lot and, e.g., the naive approach gives totally erroneous estimates. The reason for this higher tolerance to inaccurate PSFs resides in constants α and β (weights). To prevent amplification of noise, the weight α of the image smoothing term must be set higher and the other terms in the energy function become less important. Consequently, the estimated HR images tend to piecewise constant functions (general behavior of the TV seminorm) and discrepancies in the volatile blurs become less important. The BSR algorithm flattens interiors of objects to nearly homogeneous regions. Notice, e.g., that the shirt texture well reconstructed in 50dB is removed in 30dB reconstruction.

9.6.2 Real data

We tested the BSR method on real photos acquired with three different acquisition devices: mobile phone, webcamera and standard digital camera. The mobile phone of Nokia brand was equipped with a 1 Mpixel camera. The webcam was Logitech QuickCam for Notebooks Pro with the maximum video resolution 640×480 and the minimum shutter speed 1/10s. The last and most advanced device was a 5 Mpixel color digital camera (Olympus C5050Z) equipped with an optical zoom $3\times$. Since this work considers gray-level images, LR images correspond either to green channels or to gray-level images converted from color photos. To compare the quality of SR reconstruction, we provide results of two additional methods: interpolation technique and state-of-the-art SR method. The former technique consists of the MBD method proposed in [25] followed by standard bilinear interpolation (BI) resampling. The MBD method first removes volatile blurs and then BI 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 [10, 22]. 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 [3].

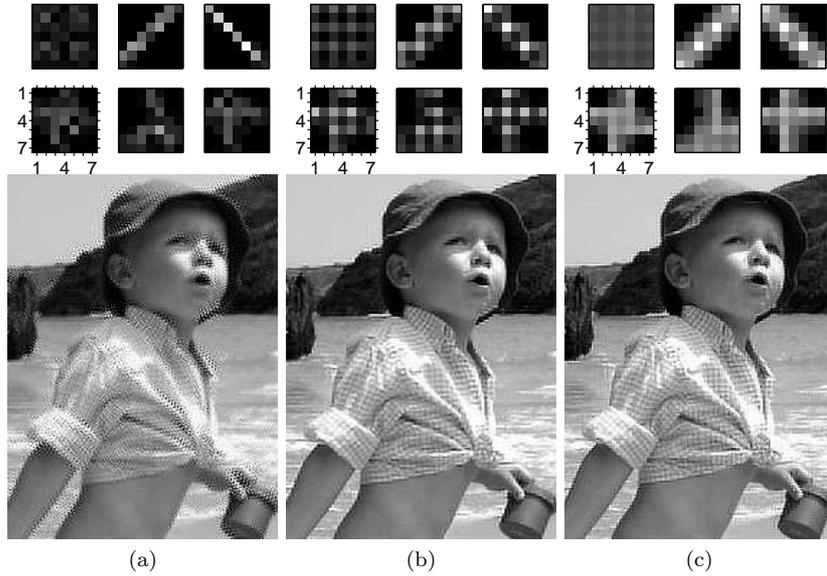First, the performance of the proposed BSR method was tested on data

FIGURE 9.7: BSR of simulated data with SNR = 50dB: Estimated HR images and PSFs for three different regularization matrices; (a) naive approach for the SR factor 2; (b) proposed polyphase approach for the integer SR factor 2; (c) proposed polyphase approach for the rational SR factor 7/4 (1.75). Note artifacts in (a) and the accurate PSF reconstruction in (c).

with negligible volatile blurs. Using the mobile phone, we took eight images of a sheet of paper with text. Fig. 9.9(a) shows one part (70×80) of the image (zero-order interpolation) considered in this experiment. Since the light conditions were good, the shutter speed of the mobile was short, which minimized any possible occurrence of volatile blurs. We set the desired SR factor to 2 and applied the standard SR method, MBD with BI, and BSR with outcomes in Figs. 9.9(b), (c) and (d), respectively. The standard SR technique gave results equivalent to those obtained by the BSR algorithm. In both cases the text is legible and the PSFs are almost identical, which indicates that the volatile blurs estimated by the BSR method were close to Dirac pulses. Consequently, the MBD method achieved only a little improvement as there was no blurring and bilinear interpolation does not create any new information.

The next two experiments demonstrate the true power of the BSR algorithm as we now consider LR images with substantial blurring. In the first one, we hold the webcam in hands and captured a short video sequence under poor light conditions. Then we extracted 10 consecutive frames and considered a small section of size 80×60; see one frame with zero-order interpolation

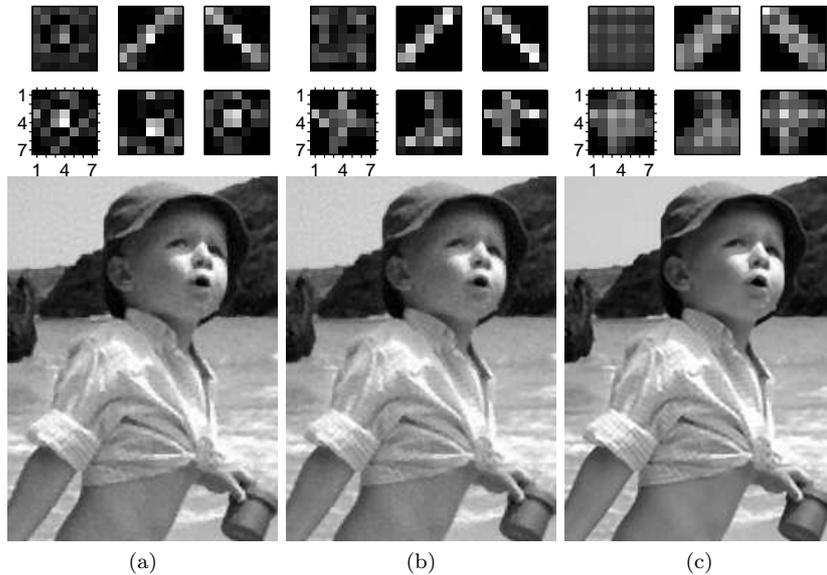**FIGURE 9.8:** BSR of simulated data with SNR = 30dB: Estimated HR images and PSFs for three different regularization matrices; (a) naive approach for the SR factor 2; (b) proposed polyphase approach for the integer SR factor 2; (c) proposed polyphase approach for the rational SR factor 7/4 (1.75). Due to noise, the image smoothing regularization term takes over and the HR images are smoother than for SNR = 50dB in Fig. 9.7.

in Fig. 9.10(a). The long shutter speed (1/10s) together with the inevitable motion of hands introduced blurring into the images. In this experiment, the SR factor was again 2. The standard SR algorithm could not handle this complicated situation with volatile blurs and the reconstructed HR image in Fig. 9.10(b) shows many disturbing artefacts. The MBD combined with BI removed blurring but subtle details in the image remained hidden; see Fig. 9.10(c). On the other hand, the proposed BSR algorithm removed blurring and performed SR correctly as one can compare in Fig. 9.10(d); note, e.g., the word "BŘEZEN". The PSFs estimated by the MBD and BSR look similar, as expected, but the BSR blurs contain more details.

In the second experiment, we compare the three reconstruction techniques on blurred photos of a car front. With the digital camera, we took four shots in a row and cropped a 120×125 rectangle from each. All four cuttings printed in their original size (no interpolation) are in Fig. 9.11(a). Similar to the previous experiment, the camera was held in hands, and due to the longer shutter speed, the LR images exhibit blurring. We set the SR factor to 5/3. Again we applied all three techniques as before. In order to better

asses the obtained results we took one additional image with optical zoom 1.7× (close to the desired SR factor 5/3) and with the camera mounted on a tripod to avoid any volatile blurs. This image served as the ground truth; see Fig. 9.11(e). Both MBD with BI in Fig. 9.11(b) and the standard SR approach in Fig. 9.11(c) failed to provide sharp HR images. The proposed BSR method outperformed both techniques and returned a well reconstructed HR image (Fig. 9.11(d)), which is comparable to the "ground-true" image acquired with the optical zoom and tripod. The PSFs estimated with BSR are in Fig. 9.11(f). To better evaluate the results, refer to four close-ups in Fig. 9.11(g).

9.6.3   Performance experiments

When dealing with real data, one cannot expect that the performance will increase indefinitely 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 get worse. We conducted several experiments on real data (short shutter speed and still shooting objects) with different SR factors and number of LR images K. See the results of one such experiment in Fig. 9.12 for $\varepsilon = 7/4$ and the number of LR images ranging from 4 to 10. Note, that at the end of Section 9.4.2 we concluded that the minimum number of LR images necessary to construct the blur regularization $R(\mathbf{h})$ for $\varepsilon = 7/4$ is 4. A certain improvement is apparent in using 6 instead of 4 LR images; see Fig. 9.12(c). However, results obtained with more images (8 and 10) show almost no 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.

The last experiment of this chapter demonstrates that the BSR method truly reconstructs high-frequency information, which is otherwise unavailable in the single LR image. With the digital camera, we took eight images of a ISO test pattern "Chart". The original chart is in Fig. 9.13(b) and one of the acquired LR images with zero-order interpolation is in Fig. 9.13(a). A riveting feature of this chart is that if BSR can recover high frequencies then we should be able to distinguish the rays closer to the center and thus eliminate the Moire effect apparent on the LR images. We applied the BSR algorithm with six different SR factors from $\varepsilon = 1$ to 2.5; see the estimated HR images in Fig. 9.13(c). From the obtained results it is clear that the amount of high-frequency information estimated by BSR depends on the SR factor. The Moire effect still visible for $\varepsilon = 1.25$, disappears for $\varepsilon = 1.5$. As the SR factor increases, the rays become better outlined closer to the center of the chart. However, this does not continue to infinity for real noisy data. In this case, we did not see any objective improvement beyond $\varepsilon = 2$.
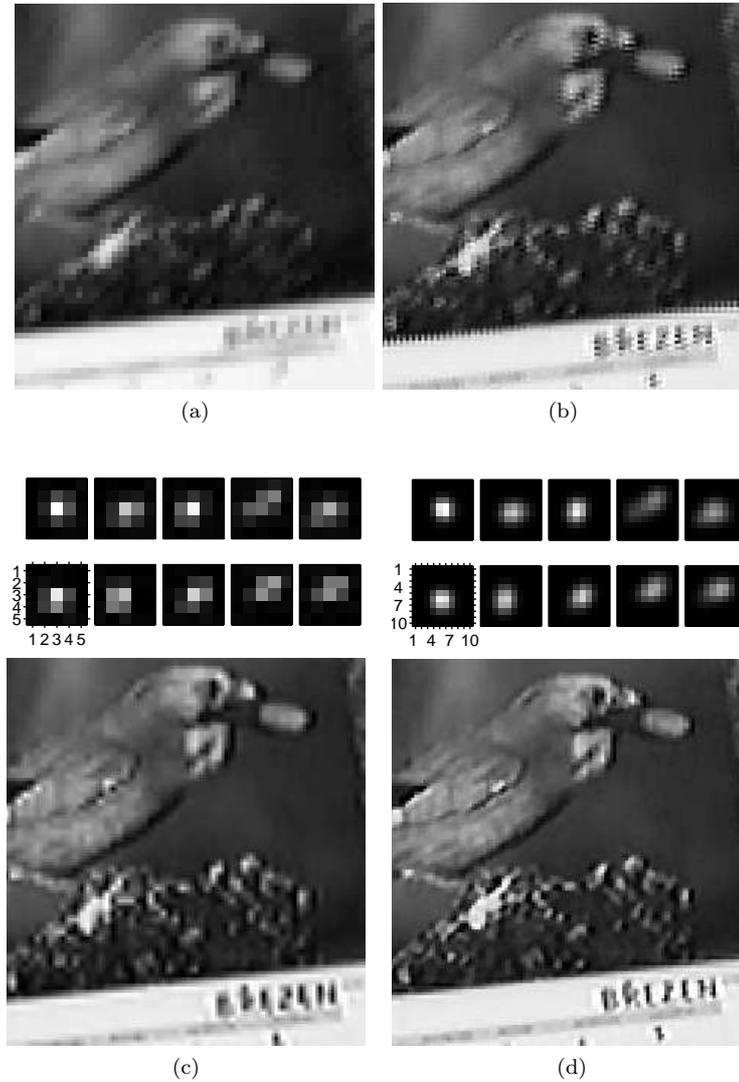
FIGURE 9.9: Reconstruction of images acquired with a mobile phone ($\varepsilon = 2$): (a) one of eight LR images shot with the mobile phone, zero-order interpolation; (b) HR image and blur shifts estimated by the standard SR algorithm; (c) HR image and blurs estimated by MBD with bilinear interpolation; (d) HR image and blurs estimated by the BSR algorithm. Volatile blurs were negligible in this case and the main source of degradation in (a) was the insufficient resolution of the device. Both, the standard SR (b) and BSR (d) methods, give similar results that improve legibility of the text significantly. MBD (c) can hardly achieve any improvement. The LR images in (a) have been courtesy of Janne Heikkila from the University of Oulu, Finland.

FIGURE 9.10: Reconstruction of images acquired with a webcam ($\varepsilon = 2$): (a) one of ten LR frames extracted from a short video sequence captured with the webcam, zero-order interpolation; (b) standard SR algorithm; (c) HR image and blurs estimated by MBD with bilinear interpolation; (d) HR image and blurs estimated by the BSR algorithm. Due to blurring in the LR frames, the standard SR method (b) gives unsatisfactory results. MBD (c) improves the image slightly but the true enhancement is achieved only with the BSR method (d).

(a)

(b) (c)

(d) (e)

(f) (g)
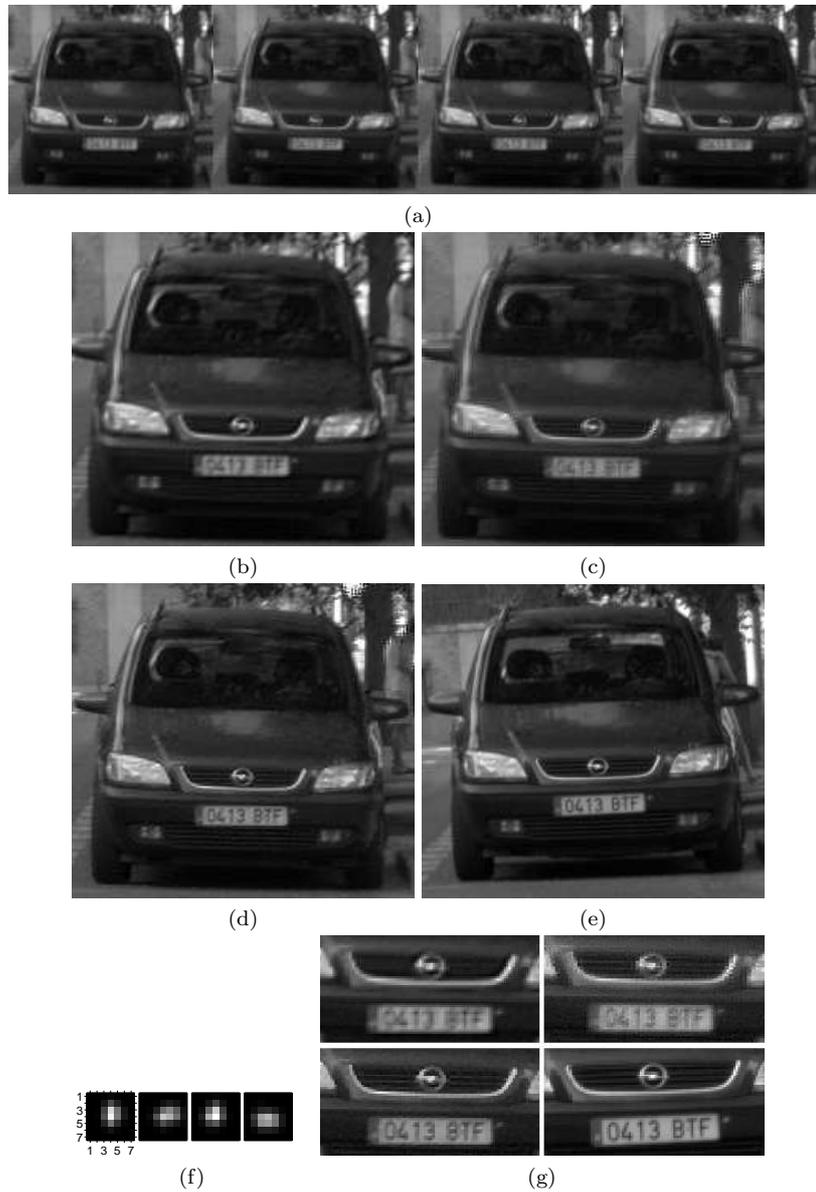
FIGURE 9.11: Reconstruction of images acquired with a digital camera ($\varepsilon = 5/3$): (a) four LR images used in the reconstruction; (b) MBD followed by bilinear interpolation; (c) standard SR algorithm; (d-f) BSR algorithm showing the HR image together with recovered blurs; (e) image acquired with the camera mounted on a tripod and with optical zoom 1.7×; (g) close-ups of the results (b), (c) on top and (d), (e) on bottom. Due to blurring and insufficient resolution of the LR images, both MBD (b) and the standard SR method (c) give unsatisfactory results. Only the BSR algorithm (d) achieves reconstruction comparable to the image with optical zoom (e).
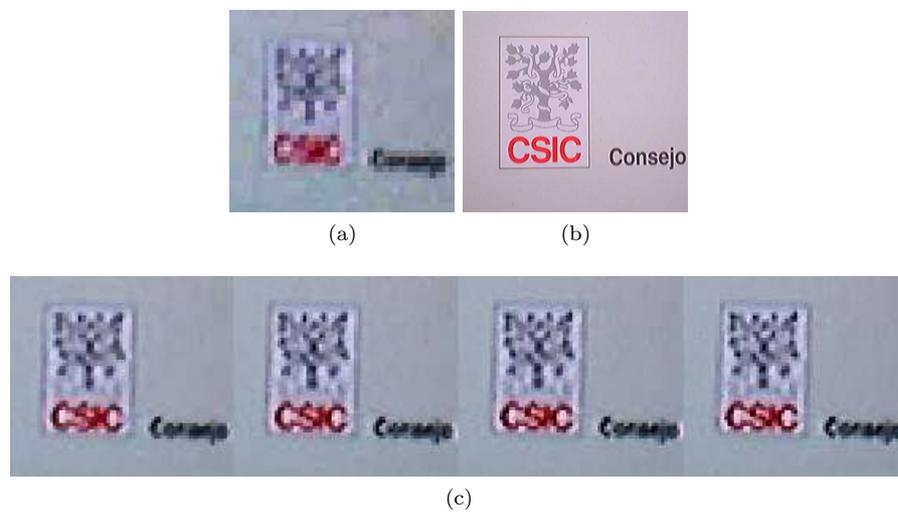
(a)   (b)

(c)

FIGURE 9.12:   Performance of the BSR algorithm with respect to the number of LR images ($\varepsilon = 7/4 = 1.75$): (a) one of ten LR images, zero-order interpolation; (b) original image; (c) HR images estimated by the BSR algorithm using 4, 6, 8 and 10 LR images (from left to right). A small improvement is visible between 4 and 6 images (compare letter "S" and details of the tree images). However, any further increase of the number of LR images proves fruitless.
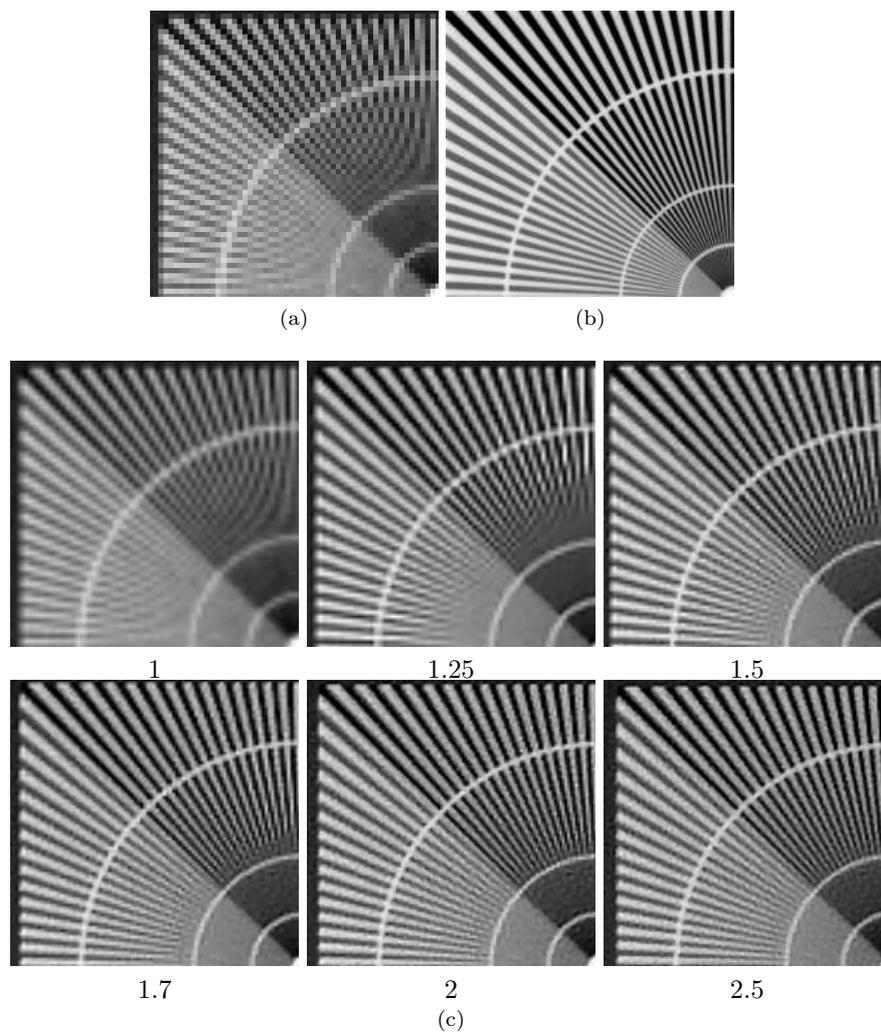
(a) (b)

1 1.25 1.5

1.7 2 2.5

(c)

FIGURE 9.13: Performance of the BSR algorithm with respect to the SR factor: (a) one of eight LR images acquired with a camera, zero-order interpolation; (b) original cropped "Chart" ISO image, courtesy of B. Brower (ITT Industries); (c) HR images estimated by the BSR algorithm with the SR factor 1, 1.25, 1.5, 1.7, 2, and 2.5. The HR images were bilinearly interpolated to have the same size. The BSR algorithm truly reconstructs high-frequency information. With the increasing SR factor, we can distinguish chart rays closer to the center. The Moire effect visible on the LR images (a) disappears completely for the SR factor 1.5 and more.

9.7 Conclusions

This chapter described a general method for blind deconvolution and resolution enhancement. 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 LR images and observe that the volatile part minimizes this functional. Due to 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 by means of polyphase decomposition. To achieve robust solution, the regularized energy minimization approach was adopted. 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 the only method that can perform deconvolution and resolution enhancement simultaneously. Several experiments with promising results give the reader a precise notion of the quality of the BSR methodology and wide applicability of the proposed algorithm to all sorts of real problems.

Acknowledgment

References

[1] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing.* Springer Verlag, New York, 2002.

[2] D.S. Biggs, C.L.Wang, T.J. Holmes, and A. Khodjakov. Subpixel de-convolution of 3D optical microscope imagery. In *Proc. SPIE*, volume 5559, pages 369–380, October 2004.

[3] D. Capel. *Image Mosaicing and Super-Resolution*. Springer, New York, 2004.

[4] T.F. Chan and C.K. Wong. Total variation blind deconvolution. *IEEE Trans. Image Processing*, 7(3):370–375, March 1998.

[5] Y. Chen, Y. Luo, and D. Hu. A general approach to blind image super-resolution using a PDE framework. In *Proc. SPIE*, volume 5960, pages 1819–1830, 2005.

[6] S. Farsiu, M.D. Robinson, M. Elad, and P. Milanfar. Fast and ro-bust multiframe super resolution. *IEEE Trans. Image Processing*, 13(10):1327–1344, October 2004.

[7] S. Farsui, D. Robinson, M. Elad, and P. Milanfar. Advances and chal-lenges in super-resolution. *Int. J. Imag. Syst. Technol.*, 14(2):47–57, August 2004.

[8] 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.

[9] 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.

[10] R.C. Hardie, K.J. Barnard, and E.E. Armstrong. Joint map registra-tion and high-resolution image estimation using a sequence of undersam-pled images. *IEEE Trans. Image Processing*, 6(12):1621–1633, December 1997.

[11] 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.

[12] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43–64, May 1996.

[13] D. Kundur and D. Hatzinakos. Blind image deconvolution revisited. *IEEE Signal Processing Magazine*, 13(6):61–63, November 1996.

[14] 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.

[15] R. Molina, M. Vega, J. Abad, and A.K. Katsaggelos. Parameter estimation in Bayesian high-resolution image reconstruction with multisensors. *IEEE Trans. Image Processing*, 12(12):1655–1667, December 2003.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] E. Shechtman, Y. Caspi, and M. Irani. Space-time super-resolution. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(4):531–545, April 2005.

[24] F. Šroubek and J. Flusser. Multichannel blind iterative image restoration. *IEEE Trans. Image Processing*, 12(9):1094–1106, September 2003.

[25] F. Šroubek and J. Flusser. Multichannel blind deconvolution of spatially misaligned images. *IEEE Trans. Image Processing*, 14(7):874–883, July 2005.

[26] F. Šroubek and J. Flusser. Resolution enhancement via probabilistic deconvolution of multiple degraded images. *Pattern Recognition Letters*, 27:287–293, March 2006.

[27] Wirawan, P. Duhamel, and H. Maitre. Multi-channel high resolution blind image restoration. In *Proc. IEEE ICASSP*, pages 3229–3232, 1999.

[28] 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.

[29] N.A. Woods, N.P. Galatsanos, and A.K. Katsaggelos. Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images. *IEEE Trans. Image Processing*, 15(1):201–213, January 2006.

[30] A.E. Yagle. Blind superresolution from undersampled blurred measurements. In *Advanced Signal Processing Algorithms, Architectures, and Implementations XIII*, volume 5205, pages 299–309, Bellingham, 2003. SPIE.

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