

# A unified approach to superresolution and multichannel blind deconvolution

Filip Šroubek, Gabriel Cristóbal and Jan Flusser

## Abstract

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

## Index Terms

Image restoration, multichannel blind deconvolution, resolution enhancement, superresolution, polyphase decomposition, regularization

## I. INTRODUCTION

Imaging devices have limited achievable resolution due to many theoretical and practical restrictions. An original scene with a continuous intensity function  $o(x, y)$  warps at the camera lens because of the scene motion and/or change of the camera position. In addition, several external effects blur images: atmospheric turbulence, camera lens, relative camera-scene motion, etc. We will call these effects *volatile blurs* to emphasize their unpredictable and transitory behavior, yet we will assume that we can model them as convolution with an unknown point spread function (PSF)  $v(x, y)$ . 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

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

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 an 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

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

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. The upper part of Fig. 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 super-resolved image can recover the original scene  $o(x, y)$  exactly. A reasonable goal of SR is a discrete version of  $o(x, y)$  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.

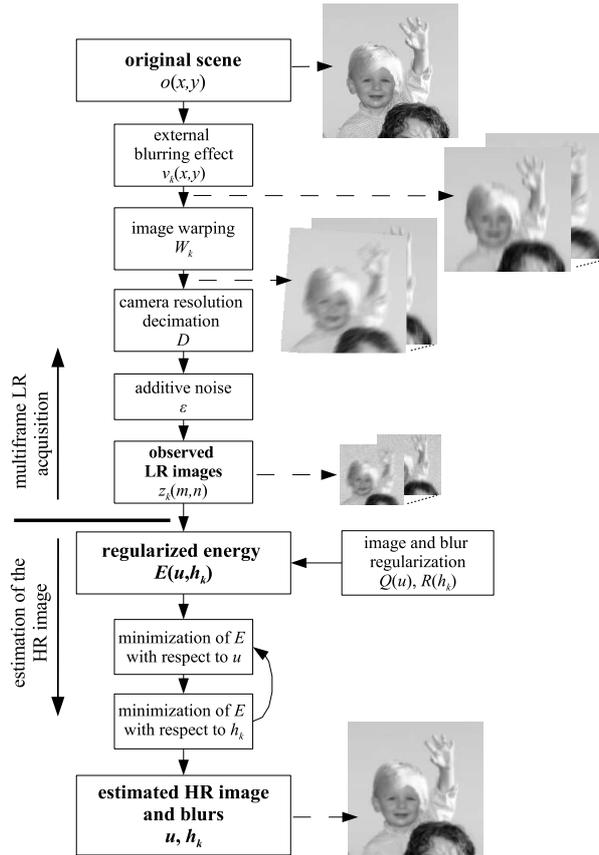


Figure 1: Low-resolution acquisition (top) and reconstruction flow (bottom).

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. [1] 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 [2]–[5]. Kundur *et al.* [6] provide a good overview. The problem 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 [7]–[12] 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 [13] 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 [14], [15]. 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 [16]–[18], include the shift estimation of the SR process. Other approaches focus on fast implementation [19], space-time SR [20] or SR of compressed video [17]. In general, most of the SR techniques assume *a priori* known blurs. However, few exceptions exist. Authors in [21], [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 [23], [24]. 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 [25], 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 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 prove that, under very mild conditions, the blurs can be recovered from the LR images up to small ambiguity. One can consider this as a generalization of the results proposed for blur estimation in the case of MBD problems. This fundamental observation enables us to

build a simple regularization term for the blurs even in the case of the SR problem. To tackle the minimization task, we use an alternating minimization approach (see Fig. 1), consisting of two simple linear equations.

Since the BSR problem requires complex matrix and vector manipulations, we dedicate the next section to notation used in the sequel. The rest of the paper is organized as follows. Section III states the acquisition model in the discrete case. Section IV reformulates the acquisition model using polyphase components. The fundamental theorems that provide solution to the BSR problem appear in Section V. The proposed regularized energy function together with the alternating minimization algorithm is described in Section VI. Performance of the proposed method under different conditions is demonstrated in the experimental Section VII and Section VIII concludes the paper.

## II. NOTATION

Our formulation of the BSR problem confines to the discrete domain with images of rectangular support and we use the following conventions throughout the sequel:

- $u(i, j)$  discrete image, lowercase letters
- $S^u := (S_1^u, S_2^u)$  size of the image  $u(i, j) : (i, j) \in \langle 0, S_1^u - 1 \rangle \times \langle 0, S_2^u - 1 \rangle$
- $U(\xi_1, \xi_2) := \sum_{0 \leq (i, j) \leq S^u} \xi_1^i \xi_2^j u(i, j)$ ,  $z$ -transform<sup>1</sup> of  $u$
- $\mathbf{u} := [u(0, 0), u(1, 0), \dots, u(S_1^u - 1, 1), u(0, 1), \dots, u(S_1^u - 1, S_2^u - 1)]^T$  image column vector, lowercase bold letters
- $\mathbf{C}$  matrix, uppercase bold letters
- $S^{\mathbf{C}} := (S_1^{\mathbf{C}}, S_2^{\mathbf{C}})$  size of the matrix  $\mathbf{C}$
- $[\cdot]_F$  operation in  $\mathbb{Z}_F$ , i.e., in the group of integers modulo  $F$
- $\lceil \cdot \rceil$  round up to the closest integer
- $(\cdot)^+$   $(x)^+ = x$  for  $x \geq 0$ , and  $(x)^+ = 0$  for  $x < 0$
- $\|\cdot\|$   $l^2$  norm

In order to shorten the notation, we define the following operators on the size pair  $(S^u)$ :

- “+”, “−” defined in a standard way
- $(i, j) + k := (i, j) + (k, k)$
- $\overline{(i, j)} := ij$
- $(i, j) < (k, l) := \{i < k \wedge j < l\}$  and similarly other binary relations “>”, “=”, etc.

Let us define convolution with a variable output support. We follow the definition in [7]. Let  $h$  and  $u$  be two images of size  $S^h := (S_1^h, S_2^h)$  and  $S^u := (S_1^u, S_2^u)$ , respectively, and  $A := (a_1 + 1, a_2 + 1)$ ,  $B := (b_1 + 1, b_2 + 1)$  define an arbitrary output rectangle, where  $A \leq B$ . Separate  $h$  column-wise and address individual columns as  $\mathbf{h}(0), \dots, \mathbf{h}(S_2^h - 1)$ . We denote a convolution matrix with a variable output support by  $\mathbf{C}_{S^u}^{A,B}\{h\}$ . It is a Toeplitz-block-Toeplitz (TBT) matrix of size  $(\overline{B - A + 1}, \overline{S^u})$  such that  $\mathbf{C}_{S^u}^{A,B}\{h\}\mathbf{u}$  gives the concatenated  $A \times B$  rectangle of  $h * u$ . More specifically

$$\mathbf{C}_{S^u}^{A,B}\{h\} := \underbrace{\begin{pmatrix} \mathbf{D}_{S_1^u}^{a_1, b_1}\{\mathbf{h}(a_2)\} & \dots & \mathbf{D}_{S_1^u}^{a_1, b_1}\{\mathbf{h}(a_2 - S_2^u + 1)\} \\ \mathbf{D}_{S_1^u}^{a_1, b_1}\{\mathbf{h}(a_2 + 1)\} & \dots & \mathbf{D}_{S_1^u}^{a_1, b_1}\{\mathbf{h}(a_2 - S_2^u + 2)\} \\ \vdots & \ddots & \vdots \\ \mathbf{D}_{S_1^u}^{a_1, b_1}\{\mathbf{h}(b_2)\} & \dots & \mathbf{D}_{S_1^u}^{a_1, b_1}\{\mathbf{h}(b_2 - S_2^u + 1)\} \end{pmatrix}}_{S_2^u \text{ blocks}} \quad (3)$$

and

$$\mathbf{D}_s^{a,b}\{\mathbf{h}(j)\} := \begin{pmatrix} h(a, j) & h(a - 1, j) & \dots & h(a - s + 1, j) \\ h(a + 1, j) & h(a, j) & \dots & h(a - s + 2, j) \\ \vdots & \vdots & \ddots & \vdots \\ h(b, j) & h(b - 1, j) & \dots & h(b - s + 1, j) \end{pmatrix}$$

where  $\mathbf{D}_s^{a,b}\{\mathbf{h}(j)\}$  is of size  $(b - a + 1, s)$  and  $h(i, j) = 0$  if  $(i, j)$  is outside the  $h$  support. Two important cases of different output size are: “full” convolution  $\mathbf{C}_{S^u}^{1, S^u + S^h - 1}\{h\}$  and “valid” convolution  $\mathbf{C}_{S^u}^{S^h, S^u}\{h\}$ . We have adopted the Matlab naming convention here. The abbreviated form  $\mathbf{C}_{S^u}\{h\}$  will always denote full convolution.

For further discussion, it is necessary to define notation for downsampled images and special submatrices. Let  $F$  denote a positive integer step (downsampling factor) and let  $\mathbf{S}_m^i$  be a 1-D sampling matrix of size  $(\lceil (m - i)/F \rceil, m)$ , where  $i = 0, \dots, F - 1$ . 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  $(m, 1)$ , the result is every  $F$ -th element of the vector starting from the  $(i + 1)$ -th element. In the 2-D case, the  $(\lceil (S^u - (i, j))/F \rceil, \overline{S^u})$  sampling matrix for the image size  $S^u$  is defined by

$$\mathbf{S}_{S^u}^{ij} := \mathbf{S}_{S_1^u}^i \otimes \mathbf{S}_{S_2^u}^j, \quad (4)$$

<sup>1</sup>We use the  $z$ -transform with positive powers to simplify the notation.

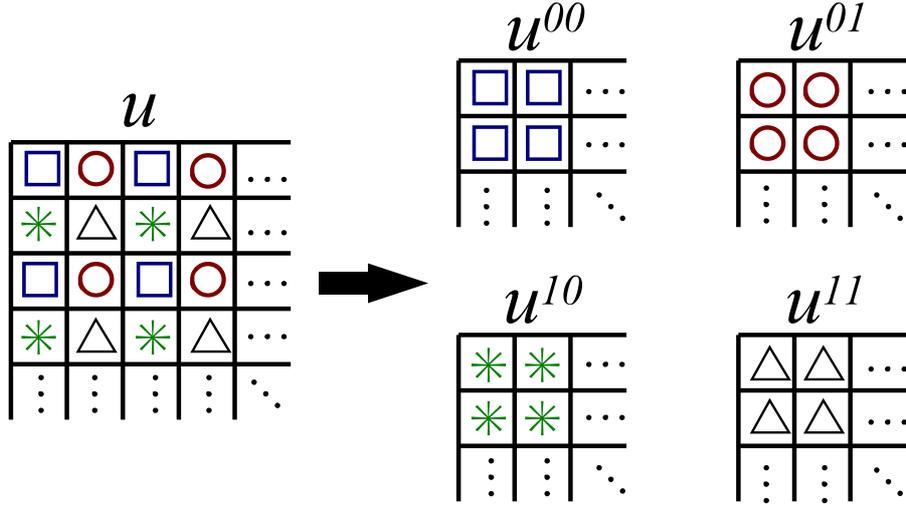


Figure 2: Polyphase decomposition for  $F = 2$ : Original image  $u$  decomposes into 4 downsampled images.

where  $\otimes$  denotes the matrix direct product (Kronecker product operator). Then on the image  $u$  we define the downsampled image in the vector form as

$$\mathbf{u}^{ij} := \mathbf{S}_{S_u}^{ij} \mathbf{u}, \quad (5)$$

which is equivalent to

$$\mathbf{u}^{ij} := [u(i, j), u(i + F), j), u(i + 2F, j), \dots, u(i, j + F), u(i + F, j + F), \dots]^T.$$

We have thus  $F^2$  distinct downsampled versions of one image; see Fig. 2. In the  $z$ -domain we refer to the downsampled images  $\mathbf{u}^{ij}$  as *polyphase components*  $U^{ij}(\xi_1, \xi_2)$  and one can readily see that

$$U^{ij}(\xi_1, \xi_2) := \sum_{0 \leq (x, y) \leq S_u - 1} \xi_1^x \xi_2^y u(Fx + i, Fy + j). \quad (6)$$

In the sequel we will often call  $\mathbf{u}^{ij}$  polyphase components as well and from the context it will be clear whether we refer to the components in the  $z$ -domain or in the image domain.

On the convolution matrix  $\mathbf{H} = \mathbf{C}_{S_u}^{A, B} \{h\}$  for  $A, B$  given above and  $C = B - A + 1$ , we define a submatrix  $\mathbf{H}^{ij, kl}$  by

$$\mathbf{H}^{ij, kl} := \mathbf{S}_C^{ij} \mathbf{H} (\mathbf{S}_{S_u}^{kl})^T \quad (7)$$

The matrices  $\mathbf{H}^{ij, kl}$ 's are still TBT as the original  $\mathbf{H}$  and one can verify that each performs convolution with the polyphase component  $\mathbf{h}^{[i-k+a_1]_F, [j-l+a_2]_F}$ .

### III. MATHEMATICAL MODEL

Let us assume we have  $K$  different LR frames  $\{z_k\}$  (each of equal size  $S^z$ ) 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  $u$ . 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

$$\mathbf{z}_k = \mathbf{D}\mathbf{V}_k\mathbf{W}_k\mathbf{u} + \mathbf{n}_k, \quad (8)$$

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 sensor PSF ( $\mathbf{G}$ ) 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 [26]. We assume that the downsampling factor (or SR factor, depending on the point of view), denoted by  $F$ , is the same in both directions. Note that  $F$  is a user-defined parameter. If  $F$  is an integer then  $\mathbf{S}$  is equivalent to (4) and is of size  $(\overline{S^z}, F^2\overline{S^z})$ . The variance of the sensor PSF together with the downsampling factor  $F$  characterize the digital sensor in use and fully define the decimation matrix. 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

$$\mathbf{z}_k = \mathbf{D}\mathbf{H}_k\mathbf{u} + \mathbf{n}_k = \mathbf{S}\mathbf{G}\mathbf{H}_k\mathbf{u} + \mathbf{n}_k. \quad (9)$$

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 annoying boundary effects, we assume that each observation  $z_k$  captures only part of  $u$ , which can be done by implementing  $\mathbf{H}_k$  and  $\mathbf{G}$  as “valid” convolution matrices  $\mathbf{C}_{S^u}^{S^h, S^u}\{h_k\}$  and  $\mathbf{C}_{S^u-S^h+1}^{S^g, S^u-S^h+1}\{g\}$ , respectively.

In the case of  $F = 1$ , the downsampling  $\mathbf{S}$  is not present and we face a standard MBD problem that has been solved elsewhere [7], [13]. Here we are interested in the case of

$F > 1$ , when the downsampling occurs. Can we estimate the blurs like in the  $F = 1$  case? The presence of  $\mathbf{S}$  prevents us to use results in [7], [13] directly. In the next section we use the polyphase formulation, transfer the problem from SIMO to MIMO and end up with a formulation of (9) without  $\mathbf{S}$ . We then show that similar conclusions obtained for MBD apply here as well.

#### IV. POLYPHASE FORMULATION

Polyphase components, as defined in (5) and illustrated in Fig. 2, consider only integer SR factors (2,3,...). This is a limiting fact since it has been shown in [27] that in real situations SR above 2 is problematic and that non-integer SR factors between 1 and 2 are required. We first derive polyphase formulation for integer SR factors and then by a simple trick extend it to any rational SR factor.

##### A. Integer downsampling factor $F$

Let us consider a simple convolution equation  $y = h * u$ , in the vector-matrix notation

$$\mathbf{y} = \mathbf{H}\mathbf{u}, \quad (10)$$

Let  $\mathbf{P}$  be a matrix defined as

$$\mathbf{P}_{S^y} := [(\mathbf{S}_{S^y}^{0,0})^T, \dots, (\mathbf{S}_{S^y}^{F-1,0})^T, (\mathbf{S}_{S^y}^{0,1})^T, \dots, (\mathbf{S}_{S^y}^{F-1,F-1})^T]^T, \quad (11)$$

where  $\mathbf{S}_{S^y}^{i,j}$  is the sampling matrix in (4). Note that  $\mathbf{P}_S$  is a permutation matrix and therefore  $\mathbf{P}_S^T \mathbf{P}_S = \mathbf{P}_S \mathbf{P}_S^T = \mathbf{I}_S$ . Multiplying (10) by  $\mathbf{P}_{S^y}$ , we get

$$\mathbf{P}_{S^y} \mathbf{y} = (\mathbf{P}_{S^y} \mathbf{H} \mathbf{P}_{S^u}^T) \mathbf{P}_{S^u} \mathbf{u} \quad (12)$$

and using the polyphase notation

$$\begin{bmatrix} \mathbf{y}^{00} \\ \mathbf{y}^{10} \\ \vdots \\ \mathbf{y}^{(F-1)(F-1)} \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{00,00} & \mathbf{H}^{00,10} & \dots & \mathbf{H}^{00,(F-1)(F-1)} \\ \mathbf{H}^{10,00} & \mathbf{H}^{10,10} & \dots & \mathbf{H}^{10,(F-1)(F-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}^{(F-1)(F-1),00} & \mathbf{H}^{(F-1)(F-1),10} & \dots & \mathbf{H}^{(F-1)(F-1),(F-1)(F-1)} \end{bmatrix} \begin{bmatrix} \mathbf{u}^{00} \\ \mathbf{u}^{10} \\ \vdots \\ \mathbf{u}^{(F-1)(F-1)} \end{bmatrix} \quad (13)$$

Similar decomposition was proposed, e.g., in [28]. To avoid long notation with the permutation matrix  $\mathbf{P}_S$ , we will use a more compact form

$$\mathbf{y}^{P(F)} = \mathbf{H}^{P(F)} \mathbf{u}^{P(F)}, \quad (14)$$

where  $\mathbf{y}^{P(F)} = \mathbf{P}_{S_y} \mathbf{y}$ ,  $\mathbf{u}^{P(F)} = \mathbf{P}_{S_u} \mathbf{u}$  and  $\mathbf{H}^{P(F)} = (\mathbf{P}_{S_y} \mathbf{H} \mathbf{P}_{S_u}^T)$ . Let us denote the row submatrices of  $\mathbf{H}^{P(F)}$  as  $\mathbf{H}^{(ij,:)} := [\mathbf{H}^{ij,00}, \dots, \mathbf{H}^{ij,(F-1)(F-1)}]$  and column submatrices as  $\mathbf{H}^{(:,ij)} := [(\mathbf{H}^{00,ij})^T, \dots, (\mathbf{H}^{(F-1)(F-1),ij})^T]^T$ . Every  $\mathbf{H}^{(ij,:)}$  (and likewise  $\mathbf{H}^{(:,ij)}$ ) comprises  $F^2$  convolution matrices  $\mathbf{H}^{ij,kl}$  and each convolution matrix performs convolution with one polyphase component of the PSF  $h$ . To get a better insight into the structure of  $\mathbf{H}^{P(F)}$ , we provide two statements easily verifiable:  $\mathbf{H}^{(ij,:)}$  first performs convolution with  $h$  shifted by  $(-i, -j)$  and then downsamples the result by  $F$ ;  $\mathbf{H}^{(:,ij)}$  first upsamples the input by  $F$ , performs convolution with  $h$  shifted by  $(-i, -j)$  and then shuffles the result to separate polyphase components.

We see that (14) is just a permutation of rows and columns of (10). The advantage of the polyphase formulation resides in the fact that downsampling is equivalent to considering only a section of (14), e.g.,  $\mathbf{y}^{00} = \mathbf{H}^{(00,:)} \mathbf{u}^{P(F)} = \mathbf{S}_{S_y}^{0,0} \mathbf{H} \mathbf{u}$ . To conclude this part, we reformulate the acquisition model (9) using polyphase components and obtain

$$\mathbf{z}_k = \mathbf{G}^{(00,:)} \mathbf{H}_k^{P(F)} \mathbf{u}^{P(F)} + \mathbf{n}_k, \quad (15)$$

where  $\mathbf{G}^{(00,:)}$  is analogous to  $\mathbf{H}^{(00,:)}$ . Any  $\mathbf{G}^{(ij,:)}$  can be used here but they are all equivalent from the reconstruction point of view, since they correspond to different translations of the original HR image  $u$ . The downsampling matrix  $\mathbf{S}$  is gone but at the cost of transforming the problem into the MIMO model. We have  $K$  output channels as before but  $F^2$  input channels (polyphase components of  $u$ ).

### B. Rational downsampling factor $F = p/q$

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 that can be 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  $F = p/q$  and the sampling frequency of the LR images  $z_k$  be  $q$ , then the sampling frequency of the HR image  $u$  will be  $p$ . From each LR image  $z_k$  we generate  $q^2$  polyphase components. We can consider these polyphase components as new output (low-LR) images with the sampling frequency 1 and we arrive to a SR problem with  $F = p$ . In other words, in order to get an integer SR factor we downsample the LR images and thus artificially increase the number of output channels. However, the number of unknown PSFs  $h_k$  remains the same.

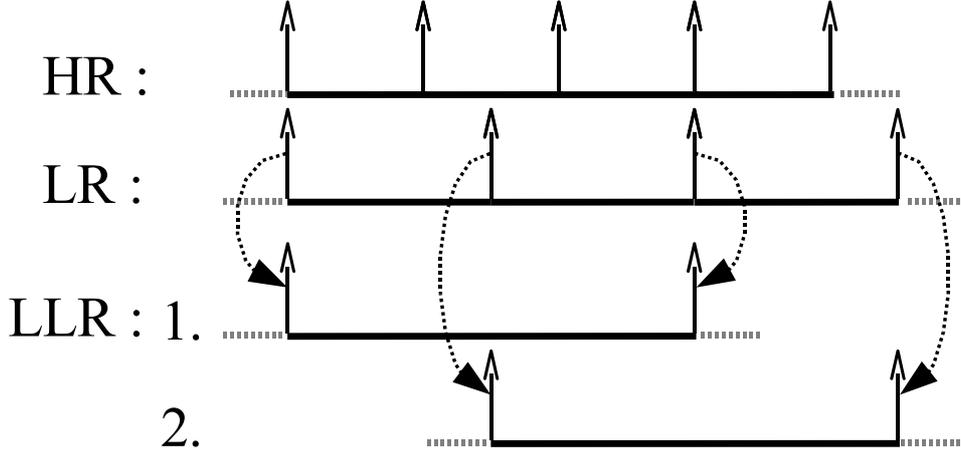


Figure 3: Rational downsampling  $F = 3/2$  in 1-D: We have LR signals (middle row) with the sampling frequency 2 and we want to obtain a HR signal (top row) with the sampling frequency 3. We convert this scenario to the one with the integer SR factor by considering every second sample of the LR signal and thus creating from each LR signal two signals (bottom row) of half size. These low-LR signals are then used in the SR problem with the integer factor 3.

We still have  $K$  PSFs since every pack of  $q^2$  low-LR images contains the same blur. An illustrative diagram of the process in 1-D for  $F = 3/2$  is given in Fig. 3.

Similarly to (15), we reformulate the acquisition model (9) using polyphase components and write

$$\mathbf{z}_k^{P(q)} = \begin{bmatrix} \mathbf{G}^{(00,:)} \\ \vdots \\ \mathbf{G}^{((q-1)(q-1,:)} \end{bmatrix} \mathbf{H}_k^{P(p)} \mathbf{u}^{P(p)} + \mathbf{n}_k, \quad (16)$$

where  $q^2$  submatrices  $\mathbf{G}^{(00,:)}, \dots, \mathbf{G}^{((q-1)(q-1,:)}$  are from  $\mathbf{G}^{P(p)}$ . We see that the rational SR factors can be expressed in a similar fashion as the integer factor. Only in this case, the resulting MIMO problem has  $Kq^2$  output channels and  $p^2$  input channels.

It remains to gloss the discretization of convolution with the sensor PSF  $g$  in the case of fractional SR factors. Since  $p$  is not divisible by  $q$ , the product  $\mathbf{S}\mathbf{G}$  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. The configuration for  $F = 3/2$  is illustrated in Fig. 4. In the above model (16), each submatrix  $\mathbf{G}^{(ij,:)}$  thus performs convolution with one of the

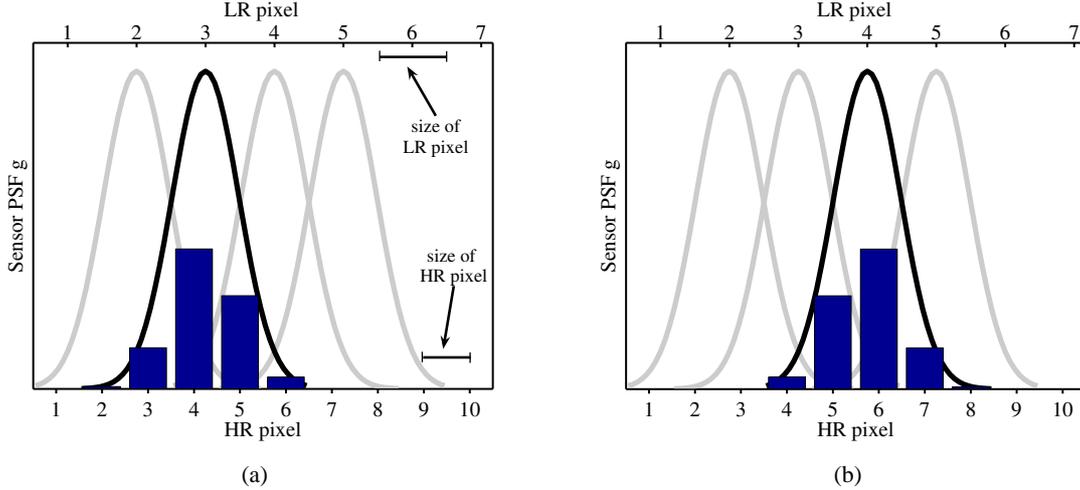


Figure 4: Sensor PSF discretization for the SR factor of  $3/2$ : Different discretizations of the PSF reside in a varying relative shift between LR and HR pixels. If the LR pixels is  $1.5\times$  the size of the HR pixel, then two distinct discretizations (a) and (b) show up in 1-D (in 2-D we have 4 discretizations). The plotted curves depict the sensor PSF in the continuous domain at different locations and the bar plots the discrete version of the PSF.

$q^2$  discretizations of  $g$  and downsamples the result by  $p$ .

## V. RECONSTRUCTION OF THE VOLATILE BLURS

We have the MIMO polyphase formulation for integer SR factors in (15) and its generalization for any rational factor in (16). The aim of this section is to use the polyphase formulation and derive fundamental theorems that will allow us to reconstruct the volatile PSFs  $h_k$ . We show that in the ideal case without any noise one can determine  $h_k$  solely from LR images  $z_k$ . Subsection V-A considers integer  $F$ , for which some ambiguity in the solution of  $h_k$  is inevitable. Subsection V-B discusses rational  $F$ , for which perfect solution is possible but only if the sensor blur  $g$  is known. Reducing the ambiguity of the solution is the topic of Subsection V-C. The last subsection addresses an issue of efficient calculation.

First, let us define notions of persistently exciting and co-primeness for polyphase components.

*Definition 1 (Polyphase persistently exciting):* The image  $u$  is called  $F$ -polyphase persistently exciting for size  $S^h$  iff all  $F^2$  polyphase components of  $\mathbf{C}_{S^u}^{S^h, S^u} \{h\} \mathbf{u}$  are nonzero for all  $h \neq 0$  of size  $S^h$ .

This assumption is very mild and it is satisfied for all common images as long as the image size is sufficiently larger than the PSF size.

*Definition 2 (Polyphase co-primeness):* Consider a set of PSFs  $\{h_k\}_1^K$ , their  $z$ -transforms  $\{H_k(\xi_1, \xi_2)\}_1^K$  and polyphase components  $\{H_k^{ij}(\xi_1, \xi_2)\}_1^K$ ,  $i, j \in \{0, \dots, F-1\}$ . They are  $F$ -polyphase co-prime iff for every  $(i, j)$  there does not exist a common factor  $C(\xi_1, \xi_2) \neq 1$  :  $H_k^{ij}(\xi_1, \xi_2) = C(\xi_1, \xi_2)\tilde{H}_k^{ij}(\xi_1, \xi_2)$ ,  $\forall k = 1, \dots, K$ .

The definition of co-primeness expresses precisely the disparity between different observation channels that is necessary in order to recover the volatile blurs.

In addition, we will assume in the following discussion that polyphase components of the PSFs are linearly independent across channels:

*Assumption A1:* For the given set of PSFs  $\{h_k\}_1^K$  and SR factor  $F$  it holds that

$$\sum_{i,j} \alpha_{ij} \begin{bmatrix} \mathbf{h}_1^{ij} \\ \vdots \\ \mathbf{h}_K^{ij} \end{bmatrix} = \mathbf{0} \quad \text{iff} \quad \alpha_{ij} = 0, \forall i, j = 0, \dots, (F-1),$$

where  $\mathbf{h}_k^{ij}$  are polyphase components of  $h_k$ .

#### A. Integer downsampling factor

The estimation of the volatile blurs  $h_k$  proceeds in two steps. First, we construct from the LR images  $z_k$  a convolution matrix  $\mathcal{Z}$  with a predetermined nullity. Second, we take the null space of this matrix and construct a new matrix  $\mathcal{N}$ , which then contains the true PSFs in its null space. The first step in its nature is similar to the procedure proposed in [9], [29] for solving the MBD problem.

We need to slightly rearrange the acquisition model in (15). Let  $S^\theta$  be the size of “nullifying” filters. The meaning of this name will be clear later. Let  $S^h$  be the maximum size of the PSFs  $\{h_k\}_1^K$ . Define “full” convolution matrices  $\mathbf{H}_k := \mathbf{C}_{F(S^\theta-1)+1}\{h_k\}$ ,  $\mathbf{G} := \mathbf{C}_{F(S^\theta-1)+S^h}\{g\}$  and “valid” convolution matrices  $\mathbf{U} := \mathbf{C}_{F(S^\theta-1)+S^g+S^h-1, S^u}\{u\}$ ,  $\mathbf{Z}_k := \mathbf{C}_{S^\theta, S^z}^{S^\theta, S^z}\{z_k\}$ . Further define

$$\mathcal{Z} := [\mathbf{Z}_1, \dots, \mathbf{Z}_K] \quad \text{and}$$

$$\mathcal{H} := [\mathbf{H}_1^{(:,00)}, \dots, \mathbf{H}_K^{(:,00)}]$$

of size  $(\overline{S^z - S^\theta + 1}, K\overline{S^\theta})$  and  $(\overline{F(S^\theta - 1) + S^h}, K\overline{S^\theta})$ , respectively. The degradation model (15) for the integer SR factor  $F$  becomes

$$\mathcal{Z} = \mathbf{U}^{(00,:)} \mathbf{G}^{P(F)} \mathcal{H} \tag{17}$$

*Lemma 1:* The null space of  $\mathcal{Z}$  is bounded from below with  $\text{nullity}(\mathcal{Z}) \geq (K\overline{S^\theta} - \overline{F(S^\theta - 1) + S^h})^+$ .

*Proof:* The dimensions of  $\mathcal{H}$  in (17) imply the lower bound. If it has more columns than rows  $\text{nullity}(\mathcal{Z}) \geq S_2^{\mathcal{H}} - S_1^{\mathcal{H}} = K\overline{S^\theta} - \overline{F(S^\theta - 1) + S^h}$  else  $\text{nullity}(\mathcal{Z}) \geq 0$ . ■

We know that since  $\mathbf{G}$  is a convolution matrix that has more rows than columns, it has full column rank (see proof in [7] for general convolution matrices). We assume that  $\mathbf{U}$  is  $F$ -polyphase persistently exciting for size  $F(S^\theta - 1) + S^g + S^h - 1$ , which implies that  $\mathbf{U}^{(00,:)}$  has full column rank. This is almost certainly true for real images if  $\mathbf{U}^{(00,:)}$  has at least as many rows as columns. If the assumption holds,  $\text{Null}(\mathcal{Z}) \equiv \text{Null}(\mathcal{H})$ . Set  $\mathbf{N} := \text{Null}(\mathcal{Z})$ . In accordance with the above lemma the size of  $\mathbf{N}$  is  $S^{\mathbf{N}} = (K\overline{S^\theta}, N)$ , where we assume  $N \geq K\overline{S^\theta} - \overline{F(S^\theta - 1) + S^h} > 0$ . We visualize the null space as

$$\mathbf{N} = \begin{bmatrix} \boldsymbol{\theta}_{1,1} & \dots & \boldsymbol{\theta}_{1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\theta}_{K,1} & \dots & \boldsymbol{\theta}_{K,N} \end{bmatrix}, \quad (18)$$

where  $\boldsymbol{\theta}_{kn}$  is the vector representation of the nullifying filter  $\theta_{kn}$  of size  $S^\theta$ ,  $k = 1, \dots, K$  and  $n = 1, \dots, N$ . Let  $\tilde{\theta}_{kn}$  of size  $F(S^\theta - 1) + 1$  denote upsampled  $\theta_{kn}$  by factor  $F$  such that  $\tilde{\theta}_{kn}^{00} = \theta_{kn}$  and  $\tilde{\theta}_{kn}^{ij} = 0, \forall (i, j) \neq (0, 0)$ , i.e.,  $\tilde{\theta}_{kn} = (\mathbf{S}_{F(S^\theta - 1) + 1}^{00})^T \theta_{kn}$ . Then, we define

$$\mathcal{N} := \begin{bmatrix} \mathbf{C}_{S^h} \{\tilde{\theta}_{1,1}\} & \dots & \mathbf{C}_{S^h} \{\tilde{\theta}_{K,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{S^h} \{\tilde{\theta}_{1,N}\} & \dots & \mathbf{C}_{S^h} \{\tilde{\theta}_{K,N}\} \end{bmatrix} \quad (19)$$

and conclude that

$$\mathcal{N}\mathbf{h} = \mathbf{0}, \quad (20)$$

where  $\mathbf{h} = [\mathbf{h}_1^T, \dots, \mathbf{h}_K^T]^T$ . This equation is a potential solution to the blur estimation problem. Unfortunately, since it was derived from (17), which is of the MIMO type, the ambiguity of the solution is high. It has been shown in [30] 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 it is summarized in the next theorem that determines the nullity of  $\mathcal{N}$ .

*Theorem 1:* Assume that  $S^\theta$  satisfies

$$N := K\overline{S^\theta} - \overline{F(S^\theta - 1) + S^h} > 0 \quad (21)$$

and

$$N\overline{F(S^\theta - 1) + S^h} \geq K\overline{S^h}. \quad (22)$$

Further assume that the HR image  $u$  is  $F$ -polyphase persistently exciting for size  $F(S^\theta - 1) + S^g + S^h - 1$ . For every polyphase index  $(i, j)$ , let there exist a common factor  $C_{ij}(\xi_1, \xi_2)$  of size  $S^{c_{ij}} : H_k^{ij} = C_{ij}(\xi_1, \xi_2) \tilde{H}_k^{ij}(\xi_1, \xi_2), \forall k = 1, \dots, K$ . It then holds that

$$\text{nullity}(\mathcal{N}) = \sum_{0 \leq (i,j) \leq (F-1)} F^2 \overline{S^{c_{ij}}} \quad (23)$$

For the proof see Appendix.

It has been shown in [7] that the assumption of co-primeness is mild, and that generically, any two common deterministic discrete filters are co-prime. The same holds true for polyphase co-primeness<sup>2</sup>. If this is the case, the following corollary applies.

*Corollary 1:* Suppose that  $S^\theta$  satisfies (21) and (22). If the HR image  $u$  is  $F$ -polyphase persistently exciting for size  $F(S^\theta - 1) + S^g + S^h - 1$  and the PSFs  $\{h_k\}_1^K$  are  $F$ -polyphase co-prime then  $\text{nullity}(\mathcal{N}) = F^4$ .

*Proof:* If  $\{h_k\}_1^K$  are  $F$ -polyphase co-prime then the only common factors  $C_{ij}(\xi_1, \xi_2)$  are scalars that are of size  $S^{c_{ij}} = 1$  and the result follows directly from Theorem 1. ■

The conclusion of the theorem and corollary may seem to be pessimistic (e.g., for  $F = 2$  the nullity is at least 16; however for  $F = 3$  the nullity is already 81). Nevertheless, we will show in Section VI that  $\mathcal{N}$  plays an important role in the regularized restoration algorithm and its ambiguity is not a serious drawback.

Up to this point we have assumed that the maximum size of the volatile PSFs ( $S^h$ ) is known or correctly estimated. In many real cases this is seldom true and one should analyze the dependence of  $\text{nullity}(\mathcal{N})$  on blur-size overestimation. Let  $S^{\hat{h}}$  denote the overestimated blur size such that  $S^{\hat{h}} - S^h = (L_1, L_2) \geq 0$ . It holds that the blur-size overestimation is equivalent under the  $z$ -transform to multiplying the original blurs by a spurious factor of degree at most  $(L_1, L_2)$ , i.e., the spurious factor in the image domain will be of size  $(L_1, L_2)$ . Therefore, one can apply Theorem 1 also to blur-size overestimation and derive the next corollary.

*Corollary 2:* Let  $(L_1, L_2) = (S^{\hat{h}} - S^h) \geq 0$ . Suppose that  $S^\theta$  satisfies (21) and (22), in which  $S^{\hat{h}}$  substitutes  $S^h$ . If the HR image  $u$  is  $F$ -polyphase persistently exciting for size  $F(S^\theta - 1) + S^g + S^{\hat{h}} - 1$  and the original PSFs  $\{h_k\}_1^K$  are  $F$ -polyphase co-prime then  $\text{nullity}(\mathcal{N}) = F^2 \overline{(L_1, L_2)} + 1$ .

*Proof:* If the blur size is overestimated by  $(L_1, L_2)$  pixels then for the size of the spurious factors  $C_{ij}(\xi_1, \xi_2)$  holds  $\sum_{0 \leq (i,j) \leq (F-1)} \overline{S^{c_{ij}}} = \overline{(L_1, L_2)} + 1$  and the result follows directly from Theorem 1. ■

<sup>2</sup>Note that co-primeness does not imply polyphase co-primeness and vice versa.

### B. Rational downsampling factor

Now we consider the rational downsampling factor  $F = p/q$  and start the analysis by rearranging the acquisition model in (16). Again, let  $S^\theta$  be the size of nullifying filters and  $S^h$  the maximum size of the PSFs  $h_k$ 's. In accordance with the discussion in the previous section, we have  $q^2$  distinct discretizations of the sensor PSF  $g$  that depend on the relative shift between HR and LR pixels. Let  $g_{ij}$  ( $i, j = 0, \dots, q-1$ ) denote such discretizations. Similarly, we define ‘‘full’’ convolution matrices  $\mathbf{H}_k := \mathbf{C}_{p(S^\theta-1)+1}\{h_k\}$ ,  $\mathbf{G}_{ij} := \mathbf{C}_{p(S^\theta-1)+S^h}\{g_{ij}\}$  and ‘‘valid’’ convolution matrices  $\mathbf{U} := \mathbf{C}_{p(S^\theta-1)+S^g+S^h-1}^{p(S^\theta-1)+S^g+S^h-1, S^u}\{u\}$ ,  $\mathbf{Z}_k := \mathbf{C}_{qS^\theta}^{qS^\theta, S^z}\{z_k\}$ . Then define

$$\begin{aligned}\mathcal{Z}' &:= [\mathbf{Z}_1^{(00,\cdot)}, \dots, \mathbf{Z}_K^{(00,\cdot)}], \\ \mathcal{G} &:= [\mathbf{G}_{00}^{P(p)}, \dots, \mathbf{G}_{(q-1)(q-1)}^{P(p)}] \quad \text{and} \\ \mathcal{H}' &:= [\mathbf{I}_{q^2} \otimes \mathbf{H}_1^{(:,00)}, \dots, \mathbf{I}_{q^2} \otimes \mathbf{H}_K^{(:,00)}]\end{aligned}$$

of size  $(\lceil (S^z - qS^\theta + 1)/q \rceil, Kq^2\overline{S^\theta})$ ,  $(p(S^\theta - 1) + S^g + S^h - 1, q^2p(S^\theta - 1) + S^h)$  and  $(q^2p(S^\theta - 1) + S^h, Kq^2\overline{S^\theta})$ , respectively. Note that  $\mathbf{Z}_k^{(00,\cdot)}$  is taken from  $\mathbf{Z}_k^{P(q)}$ , whilst  $\mathbf{H}_k^{(:,00)}$  from  $\mathbf{H}_k^{P(p)}$ . The degradation model for the rational SR factor  $F = p/q$  in (16) becomes

$$\mathcal{Z}' = \mathbf{U}^{(00,\cdot)} \mathcal{G} \mathcal{H}' \quad (24)$$

The integer SR factor is a special case of this equation. By setting  $q = 1$  we obtain (17).

In analogy with the derivation steps that led to (18), (19) and finally to (20), we proceed as follows. Set  $\mathbf{N}' := \text{Null}(\mathcal{Z}')$ . The size of  $\mathbf{N}'$  is  $(Kq\overline{S^\theta}, N')$ , where we assume  $N' \geq Kq\overline{S^\theta} - p(S^\theta - 1) + S^h + S^g - 1 > 0$ . We visualize the null space as

$$\mathbf{N}' = \begin{bmatrix} \boldsymbol{\theta}_{1,1} & \dots & \boldsymbol{\theta}_{1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\theta}_{q^2,1} & \dots & \boldsymbol{\theta}_{q^2,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\theta}_{Kq^2,1} & \dots & \boldsymbol{\theta}_{Kq^2,N} \end{bmatrix}, \quad (25)$$

where  $\boldsymbol{\theta}_{kn}$  is the vector representation of the nullifying filter  $\theta_{kn}$  of size  $S^\theta$ . Let  $\tilde{\theta}_{kn}$  denote upsampled  $\theta_{kn}$  by factor  $p$ . Then

$$\mathcal{N}' := \begin{bmatrix} \mathbf{C}_{S^g+S^h-1}\{\tilde{\theta}_{1,1}\} & \dots & \mathbf{C}_{S^g+S^h-1}\{\tilde{\theta}_{Kq^2,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{S^g+S^h-1}\{\tilde{\theta}_{1,N}\} & \dots & \mathbf{C}_{S^g+S^h-1}\{\tilde{\theta}_{Kq^2,N}\} \end{bmatrix} \times \mathbf{I}_K \otimes \begin{bmatrix} \mathbf{C}_{S^h}\{g_{00}\} \\ \vdots \\ \mathbf{C}_{S^h}\{g_{(q-1)(q-1)}\} \end{bmatrix} \quad (26)$$

and we conclude that

$$\mathcal{N}'\mathbf{h} = \mathbf{0}, \quad (27)$$

*Theorem 2:* Assume that  $S^\theta$  satisfies

$$N' := Kq^2\overline{S^\theta} - \overline{p(S^\theta - 1) + S^h + S^g - 1} > 0 \quad (28)$$

and

$$N'\overline{p(S^\theta - 1) + S^h + S^g - 1} \geq K\overline{S^h}. \quad (29)$$

If the HR image  $u$  is  $p$ -polyphase persistently exciting for size  $p(S^\theta - 1) + S^g + S^h - 1$  and the PSFs  $\{h_k\}_1^K$  are  $p$ -polyphase co-prime then  $\text{nullity}(\mathcal{N}') = 1$ .

For the proof see Appendix. The solution to (27) is unique apart from a scalar, but this magnitude ambiguity can be resolved by stipulating that the brightness of the image is preserved. While the conclusion of the above theorem is very optimistic, one should realize an important detail that distinguishes  $\mathcal{N}'$  from  $\mathcal{N}$ . The matrix  $\mathcal{N}$  does not depend on  $g$  and therefore the reconstruction of  $h_k$  for integer  $F$ , though ambiguous, can be carried out even without the knowledge of the sensor PSF. On the other hand,  $\mathcal{N}'$  contains  $q^2$  distinct discretizations of the sensor PSF and the reconstruction of  $h_k$  for rational  $F$  can fail if the sensor PSF is incorrectly estimated.

We can derive some interesting corollaries from conditions (28) and (29) in Theorem 2.<sup>3</sup>

*Corollary 3:* The minimum necessary  $S^\theta$  decreases with decreasing  $S^h$  and/or with increasing  $K$ .

*Proof:* The conclusion directly follows from conditions (28) and (29). ■

The dependence of  $S^\theta$  on  $F$ ,  $S^h$  and  $K$  is illustrated in Fig 5.

*Corollary 4:* If  $\overline{S^h + S^g - 1} \geq F^2$  then  $K > F^2$ .

*Proof:* If  $\overline{S^h + S^g - 1} \geq F^2 = (\frac{p}{q})^2$  it follows from (28) that  $Kq^2\overline{S^\theta} > \overline{p(S^\theta - 1) + S^h + S^g - 1} > \overline{p(S^\theta - 1) + S^h + S^g - 1} > p^2(\overline{S^\theta - 1} + \frac{1}{q^2}) \Rightarrow K > F^2$ . ■

The total support of the sensor and volatile blur should extend over the size of one LR pixel, otherwise the SR reconstruction is problematic. Therefore the necessary condition in the second corollary is reasonable and we see that the number of acquired images must be more than the SR factor squared. For example, for  $F = 3/2$ , 3 LR images are sufficient; for  $F = 2$ , we need at least 5 LR images to perform blur reconstruction.

<sup>3</sup>Likewise from conditions stated in Theorem 1 but they are equivalent to conditions in Theorem 2 for  $q = 1$ .

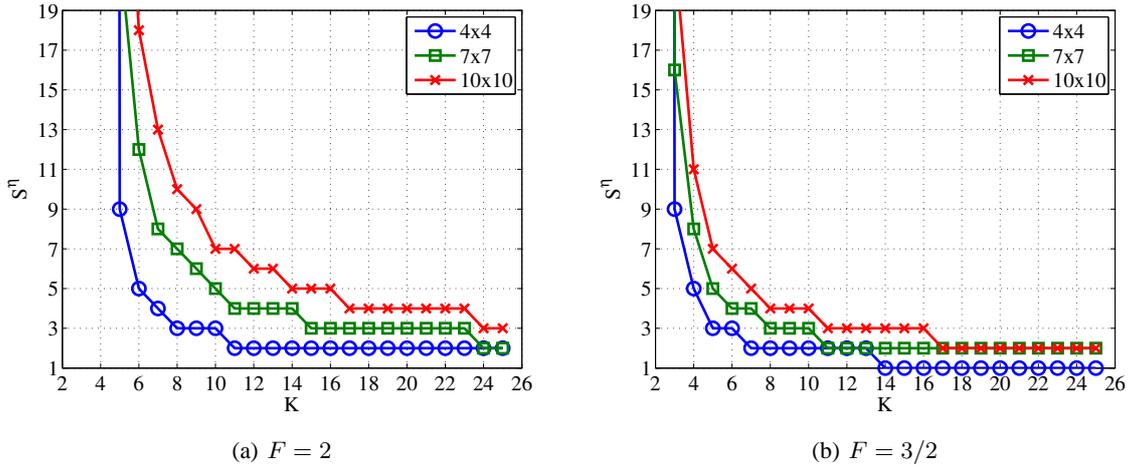


Figure 5: The size of nullifiers  $S^\theta$  depends on the SR factor ( (a)  $F = 2$ , (b)  $F = 3/2$ ), size of the blur (“o” -  $4 \times 4$ , “□” -  $7 \times 7$ , “x” -  $10 \times 10$ ) and number of output channels ( $K$ ). The plotted curves show the lower bound for the nullifier size.

### C. Reducing ambiguity

We would like to decrease the nullity of  $\mathcal{N}$  and thus avoid ambiguity in the solution under integer SR factors. A naive approach is to specify at least  $F^2$  values inside the unknown PSFs  $h_k$ . One can think of different scenarios. For example, if  $S^h \geq F$  then the knowledge of one  $h_k$  ( $\overline{S^h}$  values) suffices to fully determine all the other  $h_k$ 's. Unfortunately, in real cases one can not expect to poses such a strong knowledge. Theorem 2 suggests another possible path towards lesser ambiguity. Imagine that our acquisition device can produce in one shot two or more images that are shifted by predetermined vectors but contain the same volatile blur. Such a configuration inside acquisition devices is not rare. On the contrary, the majority of digital cameras contain CCD sensors arranged in the so-called “Bayer color pattern”, which has two times more green filters than red or blue ones; see [31]. One raw digital photography thus consists of 4 subimages: 1 red, 1 blue, and 2 greens shifted by  $[1, 1]$  pixels. Then the green channels have the desired property. We will demonstrate in the experimental section that if we have images with predetermined shifts, this information helps but with the increasing level of noise (common for real situations) its role becomes superfluous.

Let us assume we perform  $K$  measurements and in each  $R$  images are acquired that differ by predetermined displacements. We have thus  $RK$  LR images  $\{z_i\}_1^{RK}$  packed by  $R$ . In the  $k$ -th pack, we observe the same volatile PSF  $h_k$  ( $k = 1, \dots, K$ ) and the same set of  $R$

displacements. We model the  $r$ -th displacement as convolution with a filter  $t_r$  ( $r = 1, \dots, R$ ) that has one at the respective position and zeros elsewhere. Let  $S^t$  denote the maximum size of  $\{t_r\}_1^R$ . Following the same procedure that led to (24), we define “full” convolution matrices  $\mathbf{H}_k := \mathbf{C}_{F(S^\theta-1)+1}\{h_k\}$ ,  $\mathbf{T}_r := \mathbf{C}_{F(S^\theta-1)+S^h}\{t_r\}$ ,  $\mathbf{G} := \mathbf{C}_{F(S^\theta-1)+S^h+S^t-1}\{g\}$ , and “valid” convolution matrices  $\mathbf{U} := \mathbf{C}_{F(S^\theta-1)+S^g+S^h+S^t-2, S^u}\{u\}$ ,  $\mathbf{Z}_i := \mathbf{C}_{S^\theta, S^z}\{z_i\}$ . Further define

$$\begin{aligned}\mathcal{Z}^\dagger &:= [\mathbf{Z}_1, \dots, \mathbf{Z}_{RK}], \\ \mathcal{T} &:= [\mathbf{T}_1^{P(F)}, \dots, \mathbf{T}_R^{P(F)}] \quad \text{and} \\ \mathcal{H}^\dagger &:= [\mathbf{I}_R \otimes \mathbf{H}_1^{(:,00)}, \dots, \mathbf{I}_R \otimes \mathbf{H}_K^{(:,00)}],\end{aligned}$$

where  $\mathbf{H}_k^{(:,00)}$  is taken from  $\mathbf{H}_k^{P(F)}$ . The degradation model in (15) with  $R$  displacements is then given by

$$\mathcal{Z}^\dagger = \mathbf{U}^{(00,:)} \mathbf{G}^{P(F)} \mathcal{T} \mathcal{H}^\dagger. \quad (30)$$

Note that the product  $\mathbf{G}^{P(F)} \mathcal{T}$  plays a role of  $\mathcal{G}$  in (24) (convolution matrix of different discretizations of the sensor PSF). One can consider the model for the rational SR factor (24) as a special case of the above model (30).

Similarly to (25), we reconstruct nullifying filters  $\theta_{in}$  ( $i = 1, \dots, RK; n = 1, \dots, N$ ) from  $\text{Null}(\mathcal{Z}^\dagger)$ , build upsampled versions  $\tilde{\theta}_{in}$ , define

$$\mathcal{N}^\dagger := \begin{bmatrix} \mathbf{C}_{S^t+S^h-1}\{\tilde{\theta}_{1,1}\} & \dots & \mathbf{C}_{S^t+S^h-1}\{\tilde{\theta}_{RK,1}\} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{S^t+S^h-1}\{\tilde{\theta}_{1,N}\} & \dots & \mathbf{C}_{S^t+S^h-1}\{\tilde{\theta}_{RK,N}\} \end{bmatrix} \mathbf{I}_K \otimes \begin{bmatrix} \mathbf{C}_{S^h}\{t_1\} \\ \vdots \\ \mathbf{C}_{S^h}\{t_R\} \end{bmatrix} \quad (31)$$

and conclude that

$$\mathcal{N}^\dagger \mathbf{h} = \mathbf{0}, \quad (32)$$

The matrix  $\mathcal{N}^\dagger$  resembles in all aspects the matrix  $\mathcal{N}'$  from (26). Only the different discretizations of  $g$  are replaced with shift filters  $t_r$ . The following theorem formalizes the necessary conditions for a “unique” (nullity one) solution to (32).

*Theorem 3:* Assume that  $S^\theta$  satisfies

$$N^\dagger := KR\overline{S^\theta} - \overline{F(S^\theta - 1) + S^h + S^t - 1} > 0 \quad (33)$$

and

$$N^\dagger \overline{F(S^\theta - 1) + S^h + S^t - 1} \geq K\overline{S^h}. \quad (34)$$

Further assume that the HR image  $u$  is  $F$ -polyphase persistently exciting for size  $F(S^\theta - 1) + S^g + S^h + S^t - 2$  and the PSFs  $\{h_k\}_1^K$  are  $F$ -polyphase co-prime. Let  $p_r \in \langle 0, S_1^t - 1 \rangle \times \langle 0, S_2^t - 1 \rangle$  denote the position of one in  $t_r$ . If  $R \geq 3$  and vectors  $\{[p_1 - p_2]_F, [p_1 - p_3]_F, \dots, [p_1 - p_R]_F\}$  are linearly independent then  $\text{nullity}(\mathcal{N}^\dagger) = 1$ .

For the proof see Appendix. A unique solution (except for magnitude) exists for any SR factor if we have at least three different displacements that do not lie in one row, column or diagonal. The result of this theorem is particularly appealing for the SR factor of three and more. For  $F = 2$ , the benefits are not as evident, since in this case we have in total four possible displacements  $((0, 0), (1, 0), (0, 1), (1, 1))$  to fully define the HR image and we need three of them for the PSF reconstruction. One can reduce, though not eliminate, the ambiguity if only two displacements are available.

*Corollary 5:* Assume  $F = 2$ ,  $R = 2$  (displacements  $p_1, p_2$ ) and all the assumptions in Theorem 3. It then holds that if  $[p_1 - p_2]_F \in \{(1, 0), (0, 1)\}$  then  $\text{nullity}(\mathcal{N}^\dagger) = 4$  and if  $[p_1 - p_2]_F = (1, 1)$  then  $\text{nullity}(\mathcal{N}^\dagger) = 2$ .

The proof is in Appendix.

#### D. Efficient calculation

We consider here the integer downsampling case but the same holds for the rational (27) and predetermined shifts (32). The main difficulty resides in the computation of the null space  $\mathbf{N}$  of  $\mathcal{Z}$ . The “valid” convolution matrix  $\mathcal{Z}$  of size  $(\overline{S^z - S^\theta + 1}, K\overline{S^\theta})$  is too large in most of the cases to be handled as a whole. Note that though the size of the nullifying filters  $S^\theta$  is usually small (around  $10 \times 10$  pixels), the size of the LR images  $z_k$  is often over  $100 \times 100$ . Lemma 1 gives us the lower bound  $N$  of the nullity of  $\mathcal{Z}$ . Eigenvectors that correspond to the  $N$  smallest eigenvalues of  $\mathcal{Z}^T \mathcal{Z}$  span  $\text{Null}(\mathcal{H})$ . These eigenvectors then form the matrix  $\mathbf{N}$ . Therefore it suffices to construct the  $(K\overline{S^\theta}, K\overline{S^\theta})$  symmetric matrix  $\mathcal{Z}^T \mathcal{Z}$  and apply the eigenvalue decomposition thereon. This matrix product can be calculated in a very efficient way directly from  $z_k$  without building the full matrix  $\mathcal{Z}$ ; see [7], [12] for a detail discussion of calculating similar products.

## VI. 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 less ill-posed. The energy consists of three terms and takes the form

$$E(\mathbf{u}, \mathbf{h}) = \sum_{k=1}^K \|\mathbf{D}\mathbf{H}_k \mathbf{u} - \mathbf{z}_k\|^2 + \alpha Q(\mathbf{u}) + \beta R(\mathbf{h}), \quad (35)$$

where  $\mathbf{h} = [\mathbf{h}_1^T, \dots, \mathbf{h}_K^T]^T$ . The first term measures the fidelity to the data and emanates from our acquisition model (9). The remaining two are regularization terms with positive weighting constants  $\alpha$  and  $\beta$  that attract the minimum of  $E$  to an admissible set of solutions. The form of  $E$  very much resembles the energy we have proposed in [13] 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

$$Q(\mathbf{u}) = \mathbf{u}^T \mathbf{L} \mathbf{u}, \quad (36)$$

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  $\phi$  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 [32]. For the purpose of our discussion it suffices to state that after discretization we arrive again at (36), 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  $\mathcal{N}$  (or  $\mathcal{N}'$ ,  $\mathcal{N}^\dagger$ ) contains the correct PSFs  $h_k$  in its null space, we define the regularization term as a least squares fit

$$R(\mathbf{h}) = \|\mathcal{N}\mathbf{h}\|^2 = \mathbf{h}^T \mathcal{N}^T \mathcal{N} \mathbf{h}, \quad (37)$$

where for  $\mathcal{N}$  we can substitute  $\mathcal{N}'$  or  $\mathcal{N}^\dagger$ . The product  $\mathcal{N}^T \mathcal{N}$  is in general 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 1st term) 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

$$E(\mathbf{u}, \mathbf{h}) = \sum_{k=1}^K \|\mathbf{D} \mathbf{H}_k \mathbf{u} - \mathbf{z}_k\|^2 + \alpha \mathbf{u}^T \mathbf{L} \mathbf{u} + \beta \|\mathcal{N} \mathbf{h}\|^2. \quad (38)$$

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 (38) 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:

$$\begin{aligned} \text{step 1)} \quad \mathbf{u}^m &= \arg \min_{\mathbf{u}} E(\mathbf{u}, \mathbf{h}^m) \\ &\Leftrightarrow \left( \sum_{k=1}^K \mathbf{H}_k^T \mathbf{D}^T \mathbf{D} \mathbf{H}_k + \alpha \mathbf{L} \right) \mathbf{u} = \sum_{k=1}^K \mathbf{H}_k^T \mathbf{D}^T \mathbf{z}_k, \end{aligned} \quad (39)$$

$$\begin{aligned} \text{step 2)} \quad \mathbf{h}^{m+1} &= \arg \min_{\mathbf{h}} E(\mathbf{u}^m, \mathbf{h}) \\ &\Leftrightarrow \left( [\mathbf{I}_K \otimes \mathbf{U}^T \mathbf{D}^T \mathbf{D} \mathbf{U}] + \beta \mathcal{N}^T \mathcal{N} \right) \mathbf{h} = [\mathbf{I}_K \otimes \mathbf{U}^T \mathbf{D}^T] \mathbf{z}, \end{aligned} \quad (40)$$

where  $\mathbf{U} := \mathbf{C}_{S^h, S^u}^{S^h, S^u} \{u\}$ ,  $\mathbf{z} := [\mathbf{z}_1^T, \dots, \mathbf{z}_K^T]^T$  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 (38). 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  $\mathbf{h}_k \in \langle 0, 1 \rangle^{\overline{S^h}}$ . In order to enforce the bounds in step 2, we solve (40) 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  $\alpha$  and  $\beta$  depend on the level of noise. If noise increases,  $\alpha$  should increase and  $\beta$  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.

## VII. EXPERIMENTS

The experimental section consists of two parts. In the first part, a set of experiments on synthetic data evaluate performance of the BS algorithm with respect to noise and different regularization terms  $R(\mathbf{h})$ . The second part of this section demonstrates the applicability of the proposed method to real data and the performance gain with respect to the number of LR images. We compare the quality of SR reconstruction with three methods: two interpolation techniques and one state-of-the-art SR method.

In all the following experiments, we set the sensor blur to a Gaussian function of standard deviation  $\sigma = 0.35$  (relative to the scale of LR images). 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. For this reason, we have chosen  $\sigma$  that we believe is slightly smaller than the correct one.

The increasing SR factor  $F$  negatively influences the stability of the BSR algorithm. In addition, rational SR factors  $p/q$ , where  $p$  and  $q$  are incommensurable and large regardless of the effective value of  $F$ , make also the BSR algorithm unstable. It is the numerator  $p$  that determines the internal SR factor used in the algorithm. Hence we limit ourselves to  $F = 3/2, 5/3$  and 2.

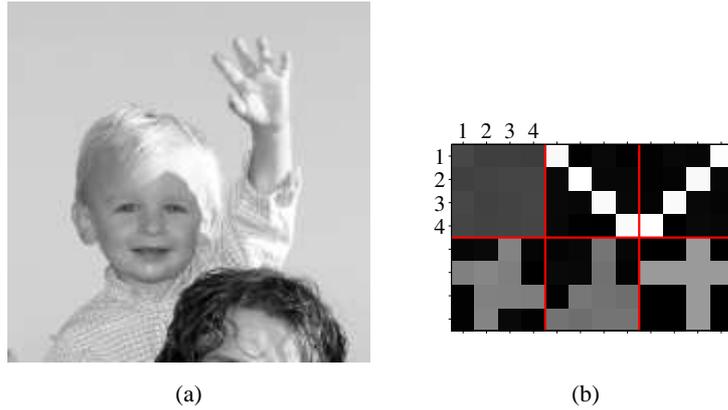


Figure 6: Simulated data: (a) original  $175 \times 175$  image; (b) six  $4 \times 4$  volatile PSFs used to blur the original image.

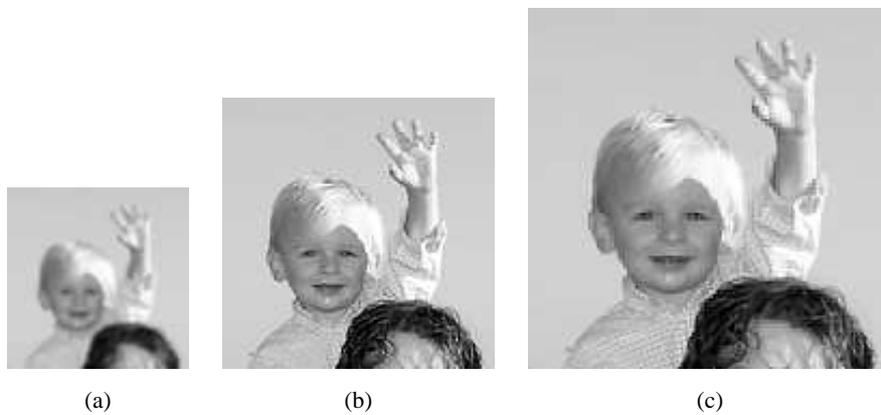


Figure 7: BSR of simulated data: (a) one of six LR images with the downsampling factor 2; (b) BSR for  $F = 3/2$ ; (c) BSR for  $F = 2$ . The shirt texture is not yet visible for the SR factor  $3/2$  but becomes well reconstructed for the SR factor 2.

#### A. Simulated data

First, let us demonstrate the BSR performance with a simple experiment. An  $175 \times 175$  image in Fig. 6(a) blurred with six masks in Fig. 6(b) and downsampled with factor 2 gives six LR images. Using the LR images as an input, we estimated the original HR image with the proposed BSR algorithm for  $F = 3/2$  and 2. Fig. 7 summarizes obtained results in their original size. One can see, that for  $F = 3/2$  (Fig. 7(b)), the reconstruction is good but some details, such as the shirt texture, are still fuzzy. For the SR factor 2, the reconstructed image in Fig. 7(c) is almost perfect.

In order to evaluate the noise robustness of the BSR method, we added noise to the above

LR images with SNR<sup>4</sup> between 50 dB and 10 dB and considered three different scenarios. In the first scenario, we downsampled the image with the integer factor 2 and performed BSR for  $F = 2$  using  $\mathcal{N}$  in the blur regularization  $R(\mathbf{h})$ . In the second scenario, we downsampled the image with the rational factor  $5/3$  and then performed BSR for  $F = 5/3$  using  $\mathcal{N}'$ . The last scenario simulated acquisitions with predetermined displacements as described in Section V-C. To each LR image (downsampling factor 2), we generated a second image shifted by  $(0.5, 0.5)$  pixels (relative the scale of the HR image) by translating the volatile blurs  $(1, 1)$  pixels. In the BSR algorithm, we then utilized  $\mathcal{N}^\dagger$  configured for two shifts:  $p_1 = (0, 0)$  and  $p_2 = (1, 1)$ . Fig. 8 summarizes the performance of the BSR method in terms of peak signal to noise ratio defined as  $\text{PSNR}(\hat{\mathbf{u}}) = 10 \log\left(\frac{\text{span}(\mathbf{u})^2}{\|\hat{\mathbf{u}} - \mathbf{u}\|^2 / S^{\bar{u}}}\right)$ , where  $\hat{\mathbf{u}}$  is the estimate of the original HR image  $\mathbf{u}$ , and  $\text{span}(\mathbf{u})$  denotes the span of gray-level values in the original image, typically 255. For the visual comparison of the obtained results refer to Fig. 9. The smaller nullity of  $\mathcal{N}'$  and  $\mathcal{N}^\dagger$  helps to improve estimates, yet as the noise level increases the performance boost diminishes. Reconstructed PSFs for the first scenario (Fig. 9(a)) exhibit a patch-like pattern due to relatively high nullity of  $\mathcal{N}$  ( $\text{nullity}(\mathcal{N}) = 16$ ). In the second and third scenario the matrix nullity is smaller and the PSF estimation is almost perfect for low noise levels; refer to Fig. 9(b) for the case of  $\mathcal{N}^\dagger$  ( $\text{nullity} = 2$ ). Results of all three scenarios become identical under more severe noise corruption with an example given in Fig. 9(c).

### B. Real data

To obtain real images, we used a standard 5 Mpixel color digital camera (Olympus C5050Z) equipped with an optical zoom  $3\times$ , which can store photos in a raw format. Since this work considers gray-level images, LR images correspond to green channels of color photos. To compare the quality of SR reconstruction, we provide results of three additional methods: two interpolation techniques and one state-of-the-art SR method. The first technique is simple bilinear interpolation (BI) of the LR image. The second technique combines the MBD method proposed in [13] and BI. The MBD method first removes volatile blurs and then BI of the deconvolved image achieves the desired spatial resolution. The third method, which we will call herein a ‘‘standard SR algorithm’’, is a MAP formulation of the SR problem proposed, e.g., in [16], [17]. 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

<sup>4</sup>SNR =  $10 \log(\sigma_u^2 / \sigma_n^2)$ , where  $\sigma_u$  and  $\sigma_n$  are the image and noise standard deviations, respectively.

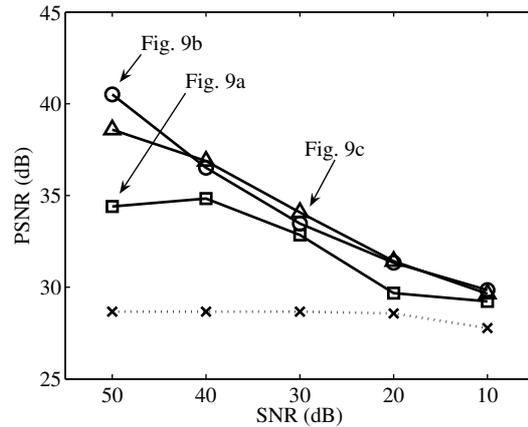


Figure 8: Performance of the BSR algorithm under different levels of noise: ( $\square$ ) BSR for  $F = 2$  with  $\mathcal{N}$  in the blur regularization  $R(\mathbf{h})$ ; ( $\triangle$ ) BSR for  $F = 5/3$  with  $\mathcal{N}'$ ; ( $\circ$ ) BSR for  $F = 2$  with  $\mathcal{N}^\dagger$ . The dotted line denotes the performance of bilinear interpolation. Note that as the noise level increases the advantage of the smaller nullity of  $\mathcal{N}'$  and  $\mathcal{N}^\dagger$  becomes less evident.

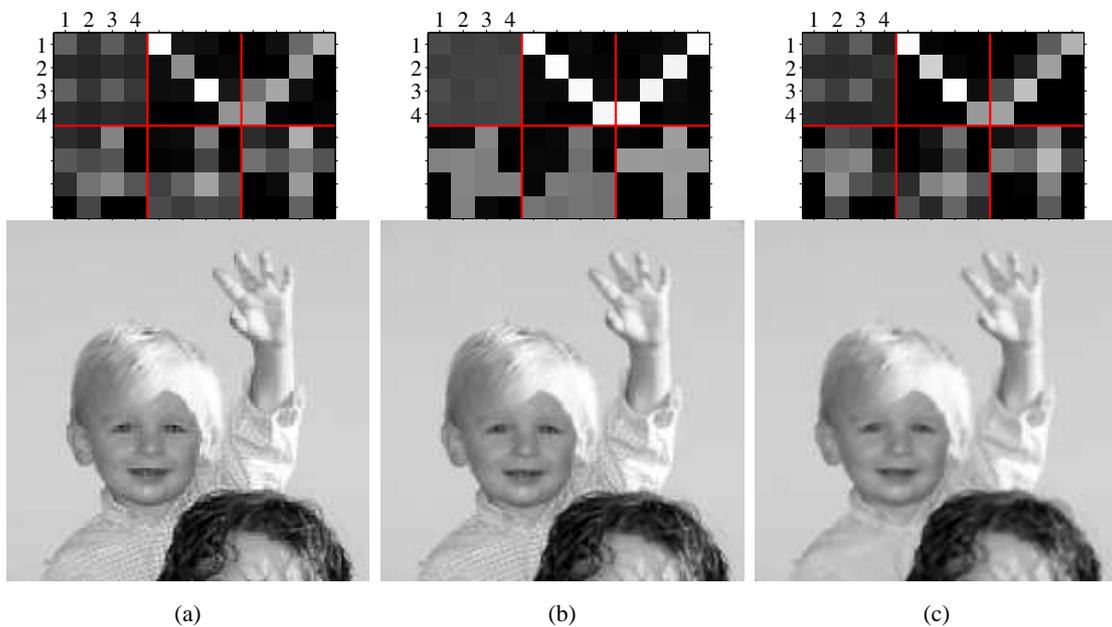


Figure 9: Results of the BSR simulation: The estimated HR image and PSFs in the case of (a) SNR = 50 dB with  $\mathcal{N}$ , (b) SNR = 50 dB with  $\mathcal{N}^\dagger$ , and (c) SNR = 30 dB with  $\mathcal{N}$ .

sensor blur ( $G$ ) and no volatile blurs. As an image prior, we use edge preserving Huber Markov Random Fields [26].

First, we show the performance of the proposed BSR method on data with negligible blur. We took eight images of a still object (see Fig. 10(a)) with a short shutter speed 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. 10(b). The estimated HR image using the BSR method is in Fig. 10(c). Close-ups in Fig. 10(d) demonstrate a very good performance comparable to the “ideal” one of the optical zoom.

The next two experiments demonstrate the true power of the BSR algorithm. In the first one, a long shutter speed introduced severe blurring in acquired images; see Fig. 11(a). We took four consecutive color images, and using both green channels, we generated altogether eight LR images. Owing to the Bayer pattern,  $(1, 1)$  shift exists between the green channels inside each color image. The standard SR approach in Fig. 11(e) gives unsatisfactory results. The MBD combined with BI in Fig. 11(c) provides very good results that lag a little behind BSR in Fig. 11(f) that does not consider the  $(1, 1)$  shift between the green channels. This is due to the character of the LR images that do not contain many details, and therefore, the SR part of the BSR algorithm cannot achieve much. The BSR approach with the predetermined shift (Fig. 11(g)) gives slightly better results that resemble the image (Fig. 11(d)) acquired with optical zoom and no blur. The second experiment compares different reconstruction techniques of a car front. In this case, the blur came from the car motion; see Fig. 12(a). The MBD with BI method (Fig. 12(b)) reconstructed well the banner, yet the license plate is not legible, since it contains tiny details. As in the previous experiment, the standard SR approach in Fig. 12(c) gives unsatisfactory results. The proposed BSR method outperforms all the other techniques and provides a sharp HR image. Figs. 12(d) and (e) illustrate BSR with the SR factor  $5/3$  and 2, respectively.

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 worsen. 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 results of one such experiment in Fig. 13 for

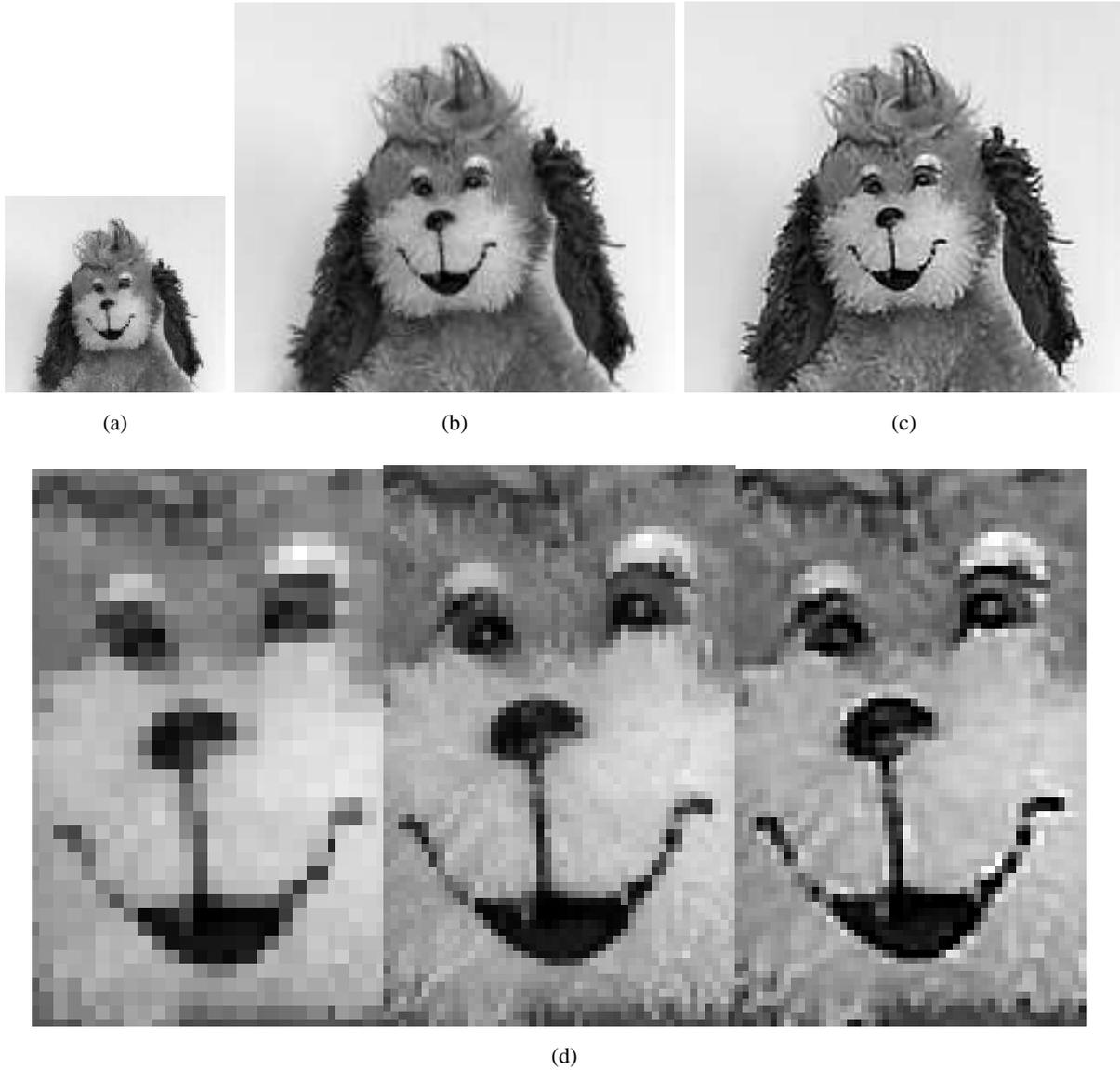


Figure 10: BSR of mildly blurred images ( $F = 2$ ): (a) One of eight LR images of size  $100 \times 90$ . (b) Image acquired with optical zoom  $2\times$ , which plays the role of “ground truth”. (c) Proposed BSR method. (d) Close-ups of images (a), (b) and (c) (from left to right).

$F = 3/2$  and the number of LR images<sup>5</sup> ranging from 3 to 8. A small improvement is apparent between using 3 and 4 LR images; compare Figs. 13(c) and (d). However, the result obtained with all 8 images in Fig. 13(e) shows a very little improvement. We deduce that for each SR factor exists an optimal number of LR images that is close to the minimum

<sup>5</sup>According to Lemma 4, the minimum number of LR images necessary to construct the blur regularization  $R(\mathbf{h})$  for  $F = 3/2$  is 3.

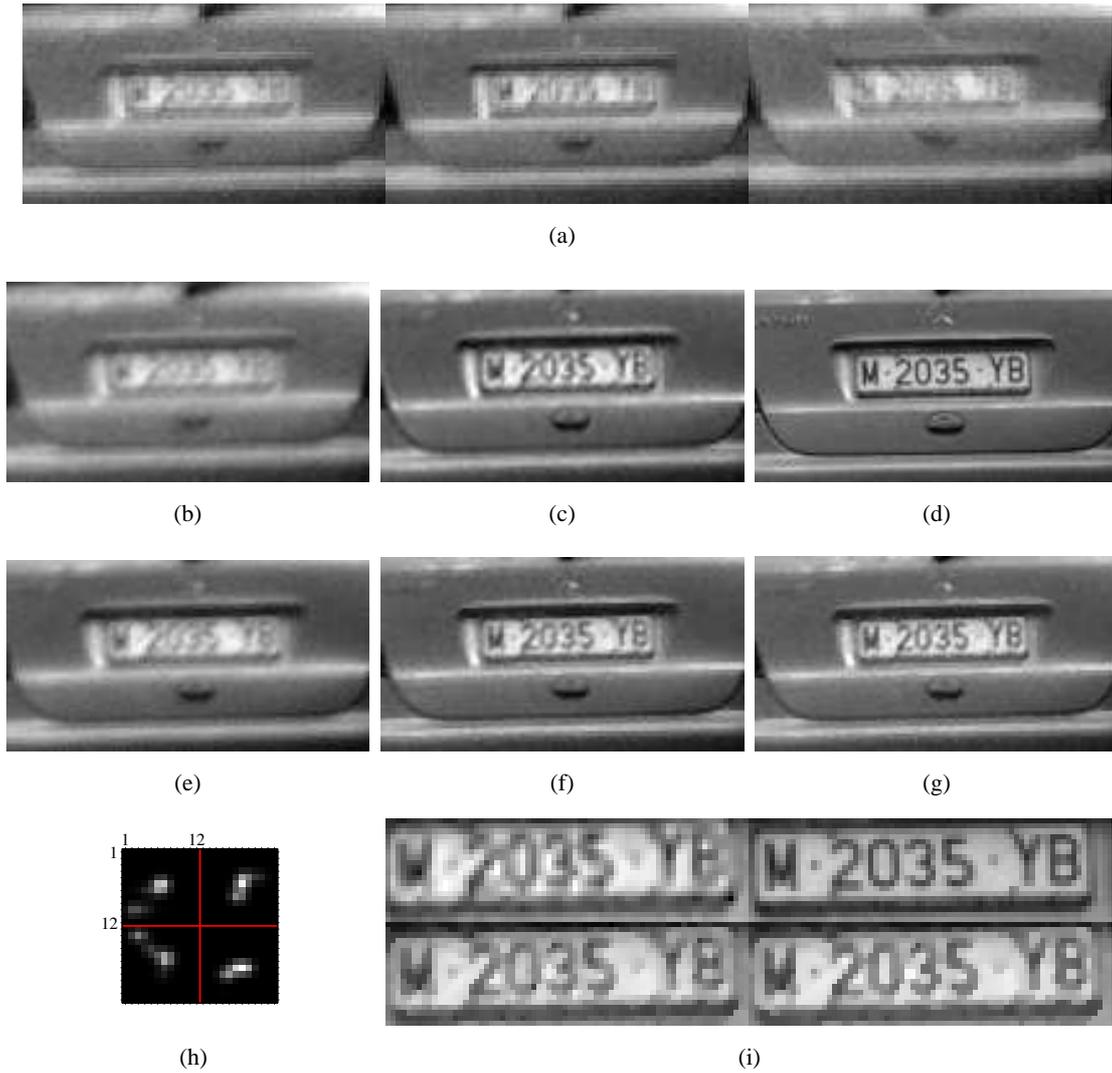


Figure 11: BSR of heavily blurred images ( $F = 5/3$ ): (a) Three out of eight LR images acquired with the digital camera; zero-order interpolation. (b) BI of the least blurred LR image. (c) MBD followed by BI. (d) Image acquired with the same camera mounted on a tripod and with optical zoom  $1.7\times$ . This image plays the role of “ground truth”. (e) Standard SR algorithm. (f) Proposed BSR without predetermined shifts ( $\mathcal{N}'$ ). (g) Proposed BSR with predetermined shifts ( $\mathcal{N}^\dagger$ ). (h) Four reconstructed  $12 \times 12$  PSFs for the result in (g). (i) Close-ups of the results (c), (d) on top and (f), (g) on bottom.

necessary number. Therefore in practice, we recommend to use the minimum or close to minimum number of LR images for the given SR factor.

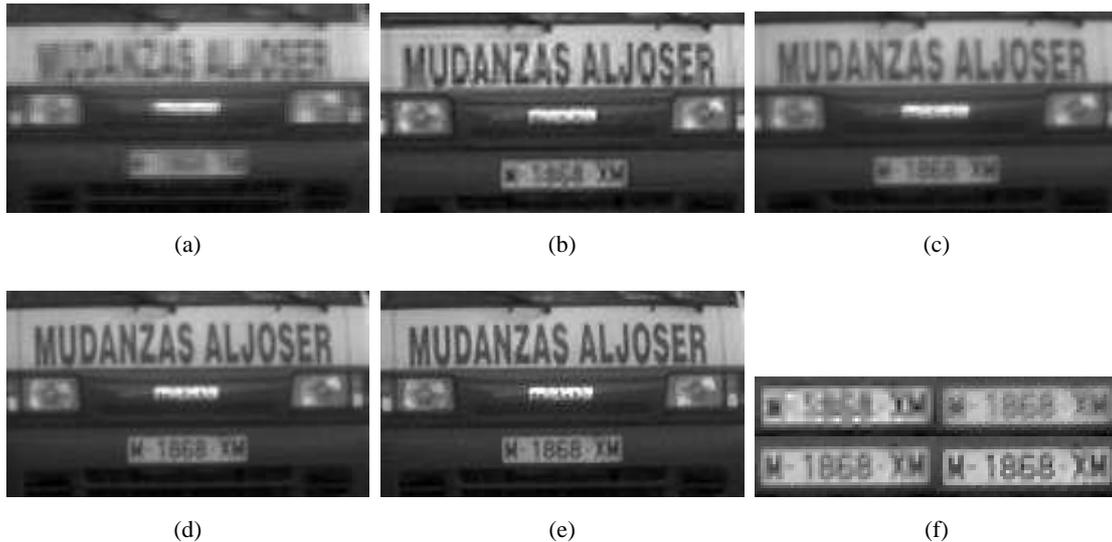


Figure 12: Car-in-motion experiment ( $F = 2$ ): (a) One of eight LR images acquired with the digital camera; zero-order interpolation. (b) MBD followed by BI. (c) Standard SR algorithm. (d) Proposed BSR algorithm for  $F = 5/3$  and bilinearly interpolated to have the same size as (e). (e) Proposed BSR algorithm for  $F = 2$ . (f) Close-ups of the results (b), (c) on top and (d), (e) on bottom.

## VIII. CONCLUSIONS

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

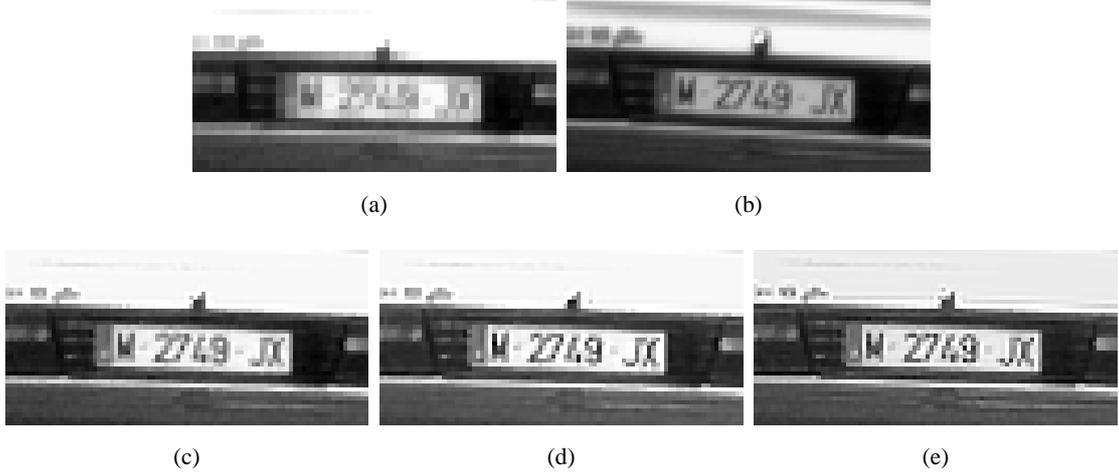


Figure 13: Performance of the BSR algorithm with respect to the number of LR images ( $F = 1.5$ ): (a) One of eight LR images of size  $40 \times 70$ , zero-order interpolation. (b) Image acquired with optical zoom  $1.5\times$ , which plays the role of “ground truth”. The proposed BSR algorithm using (c) 3, (d) 4, and (e) 8 LR images.

blur regularization and if the image regularization incorporates correlation of color channels, we will obtain a complex demosaicing methodology.

## APPENDIX

In order to prove the theorems we utilize the property of  $z$ -transform that  $y = h * u$  becomes a multiplication of polynomials  $Y(\xi_1, \xi_2) = H(\xi_1, \xi_2)U(\xi_1, \xi_2)$ . Using the polyphase components  $H^{ij}$  of  $H$  and  $U^{ij}$  of  $U$ , we can express the polyphase components of the resulting image  $Y$  as

$$Y^{ij} = \sum_{0 \leq (m,n) \leq F-1} \xi_1^{(\text{sgn}(m-i))^+} \xi_2^{(\text{sgn}(n-j))^+} H^{[i-m]_F [j-n]_F} U^{mn}. \quad (41)$$

For example, for  $F = 2$  we have

$$\begin{bmatrix} Y^{00} \\ Y^{10} \\ Y^{01} \\ Y^{11} \end{bmatrix} = \begin{bmatrix} H^{00} & \xi_1 H^{10} & \xi_2 H^{01} & \xi_1 \xi_2 H^{11} \\ H^{10} & H^{00} & \xi_2 H^{11} & \xi_2 H^{01} \\ H^{01} & \xi_1 H^{11} & H^{00} & \xi_1 H^{10} \\ H^{11} & H^{01} & H^{10} & H^{00} \end{bmatrix} \begin{bmatrix} U^{00} \\ U^{10} \\ U^{01} \\ U^{11} \end{bmatrix}. \quad (42)$$

Note that the matrices are polynomial matrices.

*Proof of Theorem 1:* In (17),  $\mathbf{G}$  has always full column rank. Since  $u$  is  $F$ -polyphase persistently exciting for size  $F(S^\theta - 1) + S^g + S^h - 1$ , then  $\mathbf{U}^{(00, \cdot)}$  has full column rank as well

and therefore  $\text{Null}(\mathcal{Z}) \equiv \text{Null}(\mathcal{H})$ . From Lemma 1 follows that the condition (21) guaranties nullity( $\mathcal{H}$ )  $> 0$  and thus the existence of the nullifying filters  $\theta_{kn}$ . The reconstruction equation  $\mathcal{N}\mathbf{h} = \mathbf{0}$  transforms in the  $z$ -domain to

$$\sum_{k=1}^K \tilde{\Theta}_{k,n} H_k = 0, \quad \forall n = 1, \dots, N, \quad (43)$$

where  $\tilde{\Theta}_{k,n}$  is the  $z$ -transform of  $\tilde{\theta}_{kn}$ . Using the polyphase formulation in (41), convolution with the upsampled filter  $\tilde{\theta}_{kn}$  is equivalent in the  $z$ -domain to multiplication by a diagonal matrix

$$\underbrace{\begin{bmatrix} \Theta_{k,n} & 0 & \dots & 0 \\ 0 & \Theta_{k,n} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \Theta_{k,n} \end{bmatrix}}_{F^2}, \quad (44)$$

where  $\Theta_{k,n}$  is the  $z$ -transform of  $\theta_{kn}$ . (Refer to an example in (42) for better understanding.) This splits every equation (43) into  $F^2$  equations  $\sum_{k=1}^K \Theta_{k,n} H_k^{ij} = 0, \forall i, j = 0, \dots, (F-1)$ . Then one can see, that the nullifying filters can not distinguish between different polyphase components. In addition,  $\sum_k \Theta_{k,n} H_k^{ij} = \sum_k \Theta_{k,n} C_{ij} \tilde{H}_k^{ij} \Rightarrow \sum_k \Theta_{k,n} \tilde{H}_k^{ij} = 0$ , where the degree of  $\tilde{H}_k^{ij}$  is by  $S^{c_{ij}}$  smaller than  $H_k^{ij}$ , and the nullifying filters can only recover the polyphase components up to the common factors  $C_{ij}$ . Hence for a polynomial matrix

$$\Theta = \begin{bmatrix} \Theta_{1,1} & \dots & \Theta_{K,1} \\ \Theta_{1,2} & \dots & \Theta_{K,2} \\ \vdots & \ddots & \vdots \\ \Theta_{1,N} & \dots & \Theta_{K,N} \end{bmatrix} \quad (45)$$

it holds that  $\text{nullity}(\Theta) = \sum_{i,j} \overline{S^{c_{ij}}}$ . If A1 is not true and the polyphase components are linearly dependent, the matrix nullity decreases. The matrix  $\mathcal{N}$  in (20) transforms to  $\Theta \otimes \mathbf{I}_{F^2}$  and if  $\mathcal{N}$  has more rows than columns (condition (22)) we conclude that  $\text{nullity}(\mathcal{N}) = \text{nullity}(\Theta \otimes \mathbf{I}_{F^2}) = F^2 \text{nullity}(\Theta) = F^2 \sum_{i,j} \overline{S^{c_{ij}}}$ .  $\square$

*Proof of Theorem 2:* The rational  $F = p/q$  factor can be considered as a special case of the problem of predetermined shifts. The matrix  $\mathcal{G}$  in (24) is similar to  $\mathbf{G}^{P(F)}\mathcal{T}$  in (30), where  $q^2$  distinct discretizations of the sensor blur  $g$  correspond to  $R = q^2$  shifts in  $\mathcal{T}$ . Hence, the result follows from the proof of Theorem 3 below.  $\square$

*Proof of Theorem 3:* We proceed as in the proof of Theorem 1. In (30),  $\mathbf{G}$  has full column rank. Since  $u$  is  $F$ -polyphase persistently exciting for size  $F(S^\theta - 1) + S^g + S^h + S^t - 2$ , then

$\mathbf{U}^{(00:\cdot)}$  has full column rank as well and therefore  $\text{Null}(\mathcal{Z}^\dagger) \equiv \text{Null}(\mathcal{T}\mathcal{H}^\dagger)$ . The condition (33) guaranties  $\text{nullity}(\mathcal{T}\mathcal{H}^\dagger) > 0$  and thus the existence of the nullifying filters  $\theta_{in}$ . In the  $z$ -domain,  $\mathcal{N}^\dagger \mathbf{h} = \mathbf{0}$  transforms to

$$\sum_{k=1}^K H_k \sum_{r=1}^R T_r \tilde{\Theta}_{R(k-1)+r,n} = 0, \quad \forall n = 1, \dots, N, \quad (46)$$

where  $T_r$  is the  $z$ -transform of  $t_r$ . Convolution with  $t_r$  shifts  $\tilde{\theta}_{in}$  by the integer vector  $p_r$  and in the above equation it is equivalent to reshuffling of polyphase components of  $\tilde{\Theta}_{in}$ . Note that all polyphase components of  $\tilde{\Theta}_{in}$  are zero except  $\tilde{\Theta}_{in}^{00} = \Theta_{in}$ . For the sake of simplicity, we will consider the case of  $F = 2$  and three linearly independent displacements  $p_1 = (0, 0)$ ,  $p_2 = (1, 0)$  and  $p_3 = (0, 1)$ . Using the polyphase formulation of convolution in (41), we can see, that for example for  $k = 1$ ,  $\sum_{r=1}^R T_r \tilde{\Theta}_{R(k-1)+r,n}$  becomes

$$\begin{bmatrix} \Theta_{1,n} & \xi_1 \Theta_{2,n} & \xi_2 \Theta_{3,n} & 0 \\ \Theta_{2,n} & \Theta_{1,n} & 0 & \xi_2 \Theta_{3,n} \\ \Theta_{3,n} & 0 & \Theta_{1,n} & \xi_1 \Theta_{2,n} \\ 0 & \Theta_{3,n} & \Theta_{2,n} & \Theta_{1,n} \end{bmatrix} \quad (47)$$

and in the same manner for all  $k$ . The shape of the above matrix forbids ambiguity in the order of polyphase components  $H_k^{ij}$  as was the case in (44). The same holds true for any integer SR factor  $F$  and any three shifts that do not lie on the same line. Since  $h_k$  are polyphase coprime, no factorization of  $H_k^{ij}$  is possible, and if  $\mathcal{N}^\dagger$  has more rows than columns (condition (34)) we conclude that  $\text{nullity}(\mathcal{N}^\dagger) = 1$ .  $\square$

*Proof of Corollary 5:* Consider the line of reasoning in the proof of Theorem 3 above. In the case of diagonal displacements (e.g.  $p_1 = (0, 0)$  and  $p_2 = (1, 1)$ ), the matrix in (47) looks as

$$\begin{bmatrix} \Theta_{1,n} & 0 & 0 & \xi_1 \xi_2 \Theta_{2,n} \\ 0 & \Theta_{1,n} & \xi_2 \Theta_{2,n} & 0 \\ 0 & \xi_1 \Theta_{2,n} & \Theta_{1,n} & 0 \\ \Theta_{2,n} & 0 & 0 & \Theta_{1,n} \end{bmatrix}. \quad (48)$$

Due to the presence of zeros, two polyphase components of  $H_k$  can remain zero without violating (46) and we have two possible configurations:  $H_k^{10} = H_k^{01} = 0$  and  $H_k^{00} = H_k^{11} = 0$ . We deduce that  $\text{nullity}(\mathcal{N}^\dagger) = 2$ .

In the case of vertical or horizontal displacements (e.g.  $p_1 = (0, 0)$  and  $p_2 = (1, 0)$ ), the

matrix in (47) takes the form

$$\begin{bmatrix} \Theta_{1,n} & \xi_1 \Theta_{2,n} & 0 & 0 \\ \Theta_{2,n} & \Theta_{1,n} & 0 & 0 \\ 0 & 0 & \Theta_{1,n} & \xi_1 \Theta_{2,n} \\ 0 & 0 & \Theta_{2,n} & \Theta_{1,n} \end{bmatrix} = \begin{bmatrix} \Theta_{1,n} & \xi_1 \Theta_{2,n} \\ \Theta_{2,n} & \Theta_{1,n} \end{bmatrix} \otimes \mathbf{I}_2. \quad (49)$$

As in the case of diagonal displacement, we have two configurations of zeros:  $H_k^{01} = H_k^{11} = 0$  and  $H_k^{00} = H_k^{10} = 0$ . However, since the Kronecker product of  $\mathbf{I}_2$  further reduces the matrix, the nullity increases twice and we conclude that  $\text{nullity}(\mathcal{N}^\dagger) = 4$ .  $\square$

## REFERENCES

- [1] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.
- [2] R. Lagendijk, J. Biemond, and D. Boeke, "Identification and restoration of noisy blurred images using the expectation-maximization algorithm," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 7, pp. 1180–1191, July 1990.
- [3] S. Reeves and R. Mersereau, "Blur identification by the method of generalized cross-validation," *IEEE Trans. Image Processing*, vol. 1, no. 3, pp. 301–311, July 1992.
- [4] T. Chan and C. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Processing*, vol. 7, no. 3, pp. 370–375, Mar. 1998.
- [5] M. Haindl, "Recursive model-based image restoration," in *Proceedings of the 15th International Conference on Pattern Recognition*, vol. III. IEEE Press, 2000, pp. 346–349.
- [6] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Magazine*, vol. 13, no. 3, pp. 43–64, May 1996.
- [7] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms," *IEEE Trans. Image Processing*, vol. 8, no. 2, pp. 202–219, Feb. 1999.
- [8] G. Giannakis and R. Heath, "Blind identification of multichannel FIR blurs and perfect image restoration," *IEEE Trans. Image Processing*, vol. 9, no. 11, pp. 1877–1896, Nov. 2000.
- [9] H.-T. Pai and A. Bovik, "On eigenstructure-based direct multichannel blind image restoration," *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1434–1446, Oct. 2001.
- [10] 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)*, 2000, pp. 22–27.
- [11] G. Panci, P. Campisi, S. Colonnese, and G. Scarano, "Multichannel blind image deconvolution using the bussgang algorithm: Spatial and multiresolution approaches," *IEEE Trans. Image Processing*, vol. 12, no. 11, pp. 1324–1337, Nov. 2003.
- [12] F. Šroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Trans. Image Processing*, vol. 12, no. 9, pp. 1094–1106, Sept. 2003.
- [13] —, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Processing*, vol. 14, no. 7, pp. 874–883, July 2005.
- [14] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Proc. Magazine*, vol. 20, no. 3, pp. 21–36, 2003.

- [15] S. Farsui, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *Int. J. Imag. Syst. Technol.*, vol. 14, no. 2, pp. 47–57, Aug. 2004.
- [16] R. Hardie, K. Barnard, and E. Armstrong, "Joint map registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1621–1633, Dec. 1997.
- [17] C. Segall, A. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Trans. Image Processing*, vol. 13, no. 7, pp. 898–911, July 2004.
- [18] N. Woods, N. Galatsanos, and A. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Trans. Image Processing*, vol. 15, no. 1, pp. 201–213, Jan. 2006.
- [19] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Processing*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
- [20] E. Shechtman, Y. Caspi, and M. Irani, "Space-time super-resolution," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 531–545, Apr. 2005.
- [21] N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement," *IEEE Trans. Image Processing*, vol. 10, no. 9, pp. 1299–1308, Sept. 2001.
- [22] N. Woods, N. Galatsanos, and A. Katsaggelos, "EM-based simultaneous registration, restoration, and interpolation of super-resolved images," in *Proc. IEEE ICIP*, vol. 2, 2003, pp. 303–306.
- [23] Wirawan, P. Duhamel, and H. Maitre, "Multi-channel high resolution blind image restoration," in *Proc. IEEE ICASSP*, 1999, pp. 3229–3232.
- [24] A. Yagle, "Blind superresolution from undersampled blurred measurements," in *Advanced Signal Processing Algorithms, Architectures, and Implementations XIII*, vol. 5205. Bellingham: SPIE, 2003, pp. 299–309.
- [25] D. Biggs, C.L.Wang, T. Holmes, and A. Khodjakov, "Subpixel deconvolution of 3D optical microscope imagery," in *Proc. SPIE*, vol. 5559, Oct. 2004, pp. 369–380.
- [26] D. Capel, *Image Mosaicing and Super-resolution*. New York: Springer, 2004.
- [27] Z. Lin and H.-Y. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 83–97, Jan. 2004.
- [28] R. Molina, M. Vega, and A. Katsaggelos, "Parameter estimation in Bayesian high-resolution image reconstruction with multisensors," *IEEE Trans. Image Processing*, vol. 12, no. 12, pp. 1655–1667, Dec. 2003.
- [29] H.-T. Pai and A. Bovik, "Exact multichannel blind image restoration," *IEEE Signal Processing Letters*, vol. 4, no. 8, pp. 217–220, Aug. 1997.
- [30] T. Moore, B. Sadler, and R. Kozick, "Regularity and strict identifiability in MIMO systems," *IEEE Trans. Signal Processing*, vol. 50, no. 8, pp. 1831–1842, Aug. 2002.
- [31] B. Gunturk, J. Glotzbach, Y. Altunbasak, R. Schafer, and R. Mersereau, "Demosaicking: Color filter array interpolation," *IEEE Signal Proc. Magazine*, vol. 22, no. 1, pp. 44–54, Jan. 2005.
- [32] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*. New York: Springer Verlag, 2002.