Fast Approximations of Shift-Variant Blur

Loïc Denis, Éric Thiébaut, Ferréol Soulez, Jean-Marie Becker, Rahul Mourya

Fast approximations of shift-variant blur

Loïc Denis · **Eric Thiébaut** · **Ferréol Soulez** · **Jean-Marie Becker** · **Rahul Mourya**

Abstract Image deblurring is essential in high resolution imaging, e.g., astronomy, microscopy or computational photography. Shift-invariant blur is fully characterized by a single point-spread-function (PSF). Blurring is then modeled by a convolution, leading to efficient algorithms for blur simulation and removal that rely on fast Fourier transforms. However, in many different contexts, blur cannot be considered constant throughout the field-of-view, and thus necessitates to model variations of the PSF with the location. These models must achieve a trade-off between the accuracy that can be reached with their flexibility, and their computational efficiency.

Several fast approximations of blur have been proposed in the literature. We give a unified presentation of these methods in the light of matrix decompositions of the blurring operator. We establish the connection between different computational tricks that can be found in the literature and the physical sense of corresponding approximations in terms of equivalent PSFs, physically-based approximations being preferable. We derive an improved approximation that preserves the same desirable low complexity as other fast algorithms while reaching a minimal approximation error.

Comparison of theoretical properties and empirical performances of each blur approximation suggests that the proposed general model is preferable for approximation and inversion of a known shift-variant blur.

Keywords Blur · Deconvolution · Inverse Problems · Image Restoration · PSF

L. Denis, J.M. Becker and R. Mourya
Université de Lyon, F-42023, Saint-Etienne, France; CNRS, UMR 5516, Laboratoire Hubert Curien, F-42000, Saint-Etienne, France; Université de Saint-Etienne, Jean Monnet, F-42000, Saint-Etienne, France
E-mail: {loic.denis,jean-marie.becker,r.mourya}@univ-st-etienne.fr

E. Thiébaut and F. Soulez
Université de Lyon, F-69000 Lyon, France; Université de Lyon 1, F-69622 Villeurbanne, France; Centre de Recherche Astrophysique de Lyon, Observatoire de Lyon, 9 Avenue Charles André, F-69561 Saint-Genis Laval CEDEX, France; CNRS, UMR 5574, Ecole Normale Supérieure de Lyon, F-69007 Lyon, France
E-mail: {eric.thiebaut,ferreol.soulez}@univ-lyon1.fr

1 Introduction

Image deconvolution is widely used to enhance the resolution, signal-to-noise ratio and contrast of blurred images. This well studied problem is classically stated as the minimization of a suitable cost function composed of a data fidelity and regularization terms Titterington (1985); Demoment (1989). The best restored image $f^\star$ reaches a compromise between fidelity to the data g and some priors:

$$f^\star = \arg \min_{f} \left\{ \Psi_{\mathsf{data}}(f,g) + \mu\, \Psi_{\mathsf{prior}}(f) \right\}, \qquad (1)$$

where Ψ_{data} is a likelihood term derived from the noise statistics and the image formation model, $\Psi_{\mathsf{prior}}(f)$ is a regularization term which has to be minimized to enforce the priors and $\mu > 0$ is a tuning parameter used to set the relative level of the priors.

As far as a small field-of-view is considered, the point spread function (PSF) can be considered as shift-invariant; as a result blurring amounts to a convolution. In many cases, however, blur is space-variant and thus cannot be modeled by a single convolution.

Blur variations across an image can be due to several causes: relative motion between the camera and the scene;

moving objects with respect to the background; variable defocusing of non planar scenes with some objects located in front of, or behind the in-focus plane; optical aberrations such as space-variant distortions, vignetting or phase aberrations. Occlusion phenomena occur, e.g., when some objects are in relative motion with respect to the scene's background, or due to parallax when a camera moves during acquisition, or with object masking in scenes displaying varying defocus blurring. There, however, exists a fundamental distinction between configurations with and without occlusion: occlusions arise from the 3-D geometry of the scene, thus a blurred image can not be related to a single planar crisp image. Occlusions are generally handled by compositing layers using their associated alpha matte (i.e., opacity image), see Porter and Duff (1984). The matting process, i.e., the extraction of foreground and alpha images is well studied for crisp images Chuang et al (2001); Wang and Cohen (2007). Deblurring images with space-variant blur involving occlusions requires a simultaneous recovering of the layers of crisp images and transparency channels, which is much more challenging, see Levin (2007); Almeida and Almeida (2009); Chakrabarti et al (2010).

The focus of this paper is shift-variant blur *without occlusion*. Several applications with practical interest indeed suffer from smooth blur variations, i.e., occlusion-free phenomena. They include 3-D microscopy, as described in Preza and Conchello (2004), astronomy with adaptive-optics correction Cresci et al (2005), or wide-field imaging, to name a few. Our focus on occlusion-free phenomena is justified by a methodological concern: PSF variations within the observed field are handled in a completely different way when multi-layer compositing is used compared to single-layer space-variant blurring. In the case of multiple layers, PSF are defined independently for each layer, and opacities (alpha values) locally govern which layer is visible, and hence, the local PSF. Occlusion-free blur requires modeling smooth PSF variations across the field. Any progress in this direction will be beneficial to matting-based occlusion modeling by handling blur variations within each layer.

While accurate modeling of PSF variations is essential for restoration, iterative deblurring techniques demand fast blurring computation. Blur models must address the two competing objectives of being *fast* and *accurate*.

State of the art: When PSFs have very small supports, computation of the blurring operator can be done in the spatial domain, see for example Sorel and Flusser (2008); Whyte et al (2010). PSF with wider supports require blur approximations. The most straightforward approach to shift-variant blur modeling is by decomposition of the observed field into small-enough regions so that the PSF may be considered invariant within each region. Thus these regions can be handled as usual with shift-invariant blur by computing

(FFT-based) convolutions, see Sec. 2.2. Such an approach however generates important artifacts at the region boundaries due to PSF discontinuities. In order to reduce these artifacts, Nagy and O'Leary (1998) propose to smooth out the transitions by interpolating between blurred images obtained by convolutions with different PSF. This modification is a strong improvement on piecewise constant PSF modeling for a reasonable additional computational cost (about 4 times more in 2D). This approach lacks physical basis in that it is not related to a natural approximation of PSFs; we show in Sec. 3 that a re-ordering of the operations (namely, weighting before convolving) improves the approximations without increasing the complexity. This formulation has been independently suggested in Gilad and Hardenberg (2006); Hirsch et al (2010). We have shown in Denis et al (2011) that it is a natural consequence of PSF interpolation. We recall this result in section Sec. 2.3. A different perspective is followed by Flicker and Rigaut (2005) and Miraut and Portilla (2012), see Sec. 2.4, where PSF are decomposed on a few modes computed by principal component analysis. While this paper was under review, we became aware of recent works that use wavelet transforms to efficiently encode the shift-variant blurring operator: Wei et al (2014); Escande and Weiss (2014). These approaches, not reviewed in details here, offer an appealing trade-off between accuracy and speed, at the cost of a rather intricate relationship between the PSFs and the approximated blurring operator. Finally, rather than modeling shift-variant blur, Maalouf et al (2011) computes several shift-invariant deconvolutions and interpolates the results. Although blur is incorrectly modeled, satisfactory results are shown with this crude approach on some microscopy images.

Our contributions: This paper extends our conference paper Denis et al (2011). We address the problem of defining fast and accurate models for space variant blur. In a discrete setting, we show that this problem can be recast as a matrix decomposition and approximation problem, thereby providing a unified framework for the description of existing models. We bridge the gap between approaches based on PSF interpolation and decomposition onto PSF modes by defining an optimal local PSF approximation. Both theoretical and practical performance of each model are compared. The formulations of Gilad and Hardenberg (2006); Hirsch et al (2010) appear better grounded and more effective in its approximation of PSF variations than the original method of Nagy and O'Leary (1998) that is still dominant in the literature. The proposed optimal local PSF approximation is shown to improve significantly (typically by an order of magnitude) the approximation error without increasing the computational cost.

Organization of the paper: Section 2 introduces the modeling of shift-variant blur. Existing models cited in the previ-

ous paragraphs are described in the light of matrix decompositions. A new approximation with established optimal trade-off between accuracy and computational efficiency is introduced in Sec. 2.5. All these models are compared in section 3 in terms of equivalent PSFs, preservation of physical properties, computational complexity, quality of PSF approximation and image restoration.

2 Approximations of shift-variant blur

Distortions caused by atmosphere turbulence, object / camera relative motion, or the instrument (limited aperture, optical aberrations) transform the original (crisp) image f into a blurry one, g. A fairly general modeling of this transform takes the form of a Fredholm integral equation of the first kind:

$$g(\boldsymbol{r}) = \int h(\boldsymbol{r}, \boldsymbol{s}) f(\boldsymbol{s}) \, \mathrm{d}\boldsymbol{s}, \tag{2}$$

where h denotes the PSF. The PSF h may be considered as the conditional density $\mathrm{p}(\boldsymbol{r}|\boldsymbol{s})$ describing the probability that a photon entering the system at location $\boldsymbol{s}$ leaves it at location $\boldsymbol{r}$, see Richardson (1972). In some cases, the PSF is shift-invariant ($\forall \boldsymbol{t}$, $h(\boldsymbol{r}, \boldsymbol{s}) = h(\boldsymbol{r} + \boldsymbol{t}, \boldsymbol{s} + \boldsymbol{t})$), i.e., depends only on the difference $\boldsymbol{r} - \boldsymbol{s}$. The right hand side of equation (2) is then a convolution and the system is referred to as isoplanatic.

More generally, the PSF may vary smoothly with the input location $\boldsymbol{s}$. In order to distinguish true PSF variations from simple shifts of the PSF $h(\boldsymbol{r}, \boldsymbol{s})$ due to changes in the input location $\boldsymbol{s}$, it will prove useful in the following to consider un-shifted PSF defined by: $k(\boldsymbol{r}, \boldsymbol{s}) = h(\boldsymbol{r} + \boldsymbol{s}, \boldsymbol{s})$. Equation (2) can then be rewritten under the form:

$$g(\boldsymbol{r}) = \int k(\boldsymbol{r} - \boldsymbol{s}, \boldsymbol{s}) f(\boldsymbol{s}) \, \mathrm{d}\boldsymbol{s}. \tag{3}$$

In the general case, the simulation of a blurred image by evaluation of the integral (2) is computationally expensive. As noted by Gilad and Hardenberg (2006), this evaluation is highly simplified if a *separable* linear approximation of the kernel is used:

$$k(\boldsymbol{r}, \boldsymbol{s}) \approx \sum_p m_p(\boldsymbol{r}) w_p(\boldsymbol{s}), \tag{4}$$

where m_p are components of the PSF model and w_p are weights depending on the location $\boldsymbol{s}$. With constant weights $w_p(\boldsymbol{s}) = w_p$, the corresponding kernel would be shift-invariant. By letting the weight $w_p(\boldsymbol{s})$ of each model m_p vary with the location $\boldsymbol{s}$, a shift-variant model is obtained. With this specific approximation, equation (2) simplifies to

a sum of convolutions:

$$g(\boldsymbol{r}) \approx \sum_p \int m_p(\boldsymbol{r} - \boldsymbol{s}) \left[w_p(\boldsymbol{s}) f(\boldsymbol{s}) \right] \mathrm{d}\boldsymbol{s}$$

$$\equiv \left[\sum_p m_p * (w_p \, f) \right](\boldsymbol{r}) \tag{5}$$

where $*$ is the classical notation for convolution. Equation (5) approximates the shift-variant operator as a sum of convolutions of weighted versions of the input image f. The existence of fast algorithms for discrete convolution makes this decomposition extremely useful. It will be central to all approximations that will be studied henceforth.

We now introduce a discrete form for blurring operations modeled by equation (2). Approximation of the discrete version of the blurring operator can then be considered from the point of view of matrix decomposition / approximation problems. A linear approximation of the integral in equation (2) leads to the following discrete linear model for blur:

$$\boldsymbol{g} = \boldsymbol{H} \cdot \boldsymbol{f}, \tag{6}$$

with $\boldsymbol{g} \in \mathbb{R}^N$ the N-pixels blurry image, $\boldsymbol{f} \in \mathbb{R}^M$ the M-pixels crisp image, $\boldsymbol{H} \in \mathbb{R}^{N \times M}$ the blurring operator, and $\cdot$ the matrix product. The matrix $\boldsymbol{H}$ defining the discrete operator is obtained by sampling the continuous operator h at locations $(\boldsymbol{r}_i)_{i=1,..,N}$ and $(\boldsymbol{s}_j)_{j=1,..,M}$:

$$\underset{\substack{1 \le i \le N \\ 1 \le j \le M}}{H_{i,j}} = h(\boldsymbol{r}_i, \boldsymbol{s}_j) \Delta_j, \tag{7}$$

with Δ_j the elementary volume measure ensuring normalization of $\boldsymbol{H}$ and possible non-uniform sampling of the input field $(\boldsymbol{s}_j)_{j=1,..,M}$. The jth column $\boldsymbol{H}_{\cdot,j}$ corresponds to the sampled PSF for a point-source located at $\boldsymbol{s}_j$.

Discretization (7) has some limitations. A more accurate discretization of integral (2) can be obtained using generalized sampling theory, as recently described by Chacko et al (2013). Let f^{int} be the continuous image defined by using the sequence of discrete coefficients $\boldsymbol{f}$ as weights of a set of basis functions:

$$f^{\text{int}}(\boldsymbol{s}) = \sum_j \vartheta_j^{\text{int}}(\boldsymbol{s}) \boldsymbol{f}_j, \tag{8}$$

with ϑ_j^{int} a shifted copy of a certain "mother" basis function ϑ^{int} (e.g., B-splines). Coefficients $\boldsymbol{f}$ are typically chosen so as to minimize the approximation error, i.e., the continuous image f^{int} corresponds to the orthogonal projection of f onto the subspace spanned by basis functions ϑ_j^{int}. Moreover, digitization of the blurred image by the sensor involves integration on the sensitive area of the pixel that will be modeled in the following way:

$$g_i = \int \vartheta_i^{\text{pix}}(\boldsymbol{r}) g(\boldsymbol{r}) \, \mathrm{d}\boldsymbol{r}, \tag{9}$$

with ϑ_i^{pix} a shifted copy of the pixel spatial sensitivity (e.g., indicator function of the sensitive area). Collecting together equations (2), (8) and (9), we get the following set of equations for the f_j coefficients:

$$g_i \approx \int \vartheta_i^{\text{pix}}(r) \int h(r,s) \sum_j \vartheta_j^{\text{int}}(s) f_j \, ds \, dr \qquad (10)$$

Thus the discrete operator H can be defined in a more general form than in equation (7):

$$H_{i,j}_{\substack{1 \le i \le N \\ 1 \le j \le M}} = \iint \vartheta_i^{\text{pix}}(r) h(r,s) \vartheta_j^{\text{int}}(s) \, ds \, dr \,. \qquad (11)$$

In order to separate changes that modify the PSF from shifts due to the displacement of the point source, we introduce a matrix K that collects all un-shifted PSFs[1], similarly to the un-shifted kernel k introduced in the continuous case:

$$K_{i,j}_{\substack{1 \le i \le N \\ 1 \le j \le M}} = \iint \vartheta_i^{\text{pix}}(r) k(r-s,s) \vartheta_j^{\text{int}}(s) \, ds \, dr \,. \qquad (12)$$

Using separable linear approximations (see equation (4))

$$K_{i,j} \approx \sum_p m_p(i) \, w_p(j) \qquad \leftrightarrow \qquad K \approx \sum_p m_p \cdot w_p^{\mathsf{T}} \quad (13)$$

provides decompositions of the shift-variant blurring operator as a sum of convolutions with prior weightings

$$H \approx \sum_p \text{conv}(m_p) \cdot \text{diag}(w_p) \,, \qquad (14)$$

where $\text{conv}(m_p)$ denotes the discrete convolution matrix with kernel m_p, $\text{diag}(w_p)$ is a diagonal matrix whose diagonal is given by vector w_p, and symbol T denotes matrix transpose.

Note that if the same kind of approximation were applied to the adjoint operator H^{T}, operator H would then be written as a sum of convolutions *followed* by spatial weighting (rather than *preceded* by spatial weighting). It would then resemble deformable filtering (a.k.a. steerable filtering), as developed in the seminal works of Freeman and Adelson (1991), Greenspan et al (1994) and Perona (1995). The key principle of steerable filtering is the ability to compute quickly the (shift-invariant) filtering result for any kernel orientation and size by simple linear combination of few pre-computed (shift-invariant) filtering outputs with well-chosen convolution kernels. The extension of such approaches to shift-variant filtering requires to approximate the rows of discrete operator H (i.e., the integration kernels) rather than its columns (i.e., the PSFs), which may be challenging since only PSF measurements are straightforward. We refer the interested reader to Miraut and Portilla (2012)

for an extended discussion on shift-variant filtering in the light of deformable filtering[2].

We give in Table 1 an illustration of some K and H matrices on very small-scale 1-D problems with Gaussian PSF. Such representations show how the matrices are structured and how their structure can be exploited to derive powerful decompositions. In contrast to Fish et al (1996), matrices K and H will neither be stored nor explicitly inverted in the following.

2.1 Shift-invariant PSF

For a shift-invariant PSF, K is a rank-one matrix with identical columns equal to a single PSF k:

$$K = k \cdot \mathbf{1}^{\mathsf{T}} \,, \qquad (15)$$

where $\mathbf{1}$ denotes a vector of convenient length composed of 1. Operator H then corresponds to a discrete convolution. While circular discrete convolutions are mapped to a simple product by the discrete Fourier transform, discrete convolutions require adequate zero-padding and cropping steps:

$$H \equiv \text{conv}(k) = R \cdot \underbrace{\mathscr{F}^{-1} \cdot \text{diag}(\hat{k}) \cdot \mathscr{F}}_{\text{circular convolution}} \cdot E \,, \qquad (16)$$

where E is an expansion operator that adds zeros to the edges of the input signal, R is a restriction operator that truncates the input signal to the size of the visible field of the output image, $\mathscr{F}$ and $\mathscr{F}^{-1}$ are the direct and inverse discrete Fourier transforms, and $\hat{k}$ is the discrete Fourier transform of the PSF:

$$\hat{k} = \mathscr{F} \cdot E \cdot k \,. \qquad (17)$$

Applications of operator H to an image f can then be very efficiently computed using fast Fourier transforms (FFT).

Correct implementation of discrete convolutions (i.e., without periodization artifacts) is essential in all following shift-variant blur approximations because they rely heavily on discrete convolutions. While circular convolutions applied to shift-invariant blurring generate artifacts located mainly at the image borders, they would yield strong artifacts all over the image when applied to shift-variant models of the form of equation (14).

In 1-D problems such as illustrated in the first row of Table 1, convolution matrix H has a Toeplitz structure. For 2-D problems, matrix H is Toeplitz-block Toeplitz and Eq. (16) and Eq. (17) involve 2-D discrete Fourier transforms. Matrix K is useful to exhibit PSF invariance compared to the more intricate structure of matrix H.

---

[1] this matrix can be considered as a dictionary of all PSFs.

[2] in particular, the adjoint operator H^{T} of operator H defined in equation (14) corresponds to a deformable filter

Table 1 Decomposition schemes of PSF K and blurring operator H in different scenarios of PSF variations. Columns of matrices K and H represent PSFs.

PSF variations	matrix	decomposition	properties

Note: Matrix entries are represented as dots whose area is proportional to their magnitude. Negative values are displayed as circles.

PSF variations	decomposition	properties
invariant (Sec. 2.1)	$K = k \cdot \mathbf{1}^\top$	rank one
	$H = \mathrm{conv}(k)$	discrete convolution
piecewise constant (Sec. 2.2)	$K = k_1 \cdot \iota_1^\top + k_2 \cdot \iota_2^\top$	low rank
	$H = \mathrm{conv}(k_1)\,\mathrm{diag}(\iota_1) + \mathrm{conv}(k_2)\,\mathrm{diag}(\iota_2)$	
smooth (Sec. 2.3)	$K \approx k_1 \cdot \varphi_1^\top + k_2 \cdot \varphi_2^\top + \cdots + k_4 \cdot \varphi_4^\top$	localized weights
	$H \approx \mathrm{conv}(k_1)\,\mathrm{diag}(\varphi_1) + \cdots + \mathrm{conv}(k_4)\,\mathrm{diag}(\varphi_4)$	
low rank (Sec. 2.4)	$K \approx u_1 \cdot \sigma_1 v_1^\top + u_2 \cdot \sigma_2 v_2^\top$	optimal global approximation
	$H \approx \mathrm{conv}(u_1)\,\mathrm{diag}(\sigma_1 v_1) + \mathrm{conv}(u_2)\,\mathrm{diag}(\sigma_2 v_2)$	
smooth (Sec. 2.5)	$K \approx c_1^\star \cdot w_1^{\star\top} + c_2^\star \cdot w_2^{\star\top} + \cdots + c_4^\star \cdot w_4^{\star\top}$	localized weights, optimal local approximation
	$H \approx \mathrm{conv}(c_1^\star)\,\mathrm{diag}(w_1^\star) + \cdots + \mathrm{conv}(c_4^\star)\,\mathrm{diag}(w_4^\star)$	

2.2 Piecewise constant PSFs

Let us now turn to the case of PSFs that are shift-invariant within regions defining a partition of the input field. Let P be the number of these regions. There are then P different PSF $\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_P\}$ and $\boldsymbol{K}$ is now a rank-P matrix:

$$\boldsymbol{K} = \sum_{p=1}^{P} \boldsymbol{k}_p \cdot \boldsymbol{\iota}_p^{\mathsf{T}}. \tag{18}$$

As illustrated in Table 1, the jth column of matrix $\boldsymbol{K}$ is equal to PSF $\boldsymbol{k}_p$ as long as $\boldsymbol{s}_j$ is located in the pth region $\mathscr{R}_p$ of the input field. The binary weights $\boldsymbol{\iota}_p$ define the indicator vector:

$$\iota_p(j) = \begin{cases} 1 & \text{if } \boldsymbol{s}_j \in \mathscr{R}_p, \\ 0 & \text{if } \boldsymbol{s}_j \notin \mathscr{R}_p. \end{cases} \tag{19}$$

Within each region, operator $\boldsymbol{H}$ acts like a discrete convolution. It is naturally expandable as a sum of discrete convolution operations with prior weighting in order to restrict each convolution to its specific region:

$$\boldsymbol{H} = \sum_{p=1}^{P} \text{conv}(\boldsymbol{k}_p) \cdot \text{diag}(\boldsymbol{\iota}_p). \tag{20}$$

A given column $\boldsymbol{h}_j$ of $\boldsymbol{H}$ (i.e., the system response to an impulse at $\boldsymbol{s}_j$) is thus equal to the corresponding column in the convolution matrix $\text{conv}(\boldsymbol{k}_p)$, with p the index of the region containing point $\boldsymbol{s}_j$. The order in which convolution and weighting are applied is essential: weighting first, then convolution. The converse order is used in Nagy and O'Leary (1998) and in many other subsequent works: Calvetti et al (2000); Nagy et al (2004); Preza and Conchello (2004); Bardsley et al (2006); Ng et al (2007); Rogers and Fiege (2011). The resulting operator is however not equivalent to $\boldsymbol{H}$, as illustrated in Table 2 and further discussed in section 3. We will denote these approximations under the term *image interpolation* to emphasize the difference with the *PSF interpolation* formulation described next.

2.3 Smoothly varying PSFs and their local approximation

In the applications discussed in the introduction, PSFs vary smoothly across the field. In such cases, a PSF (e.g., column $\boldsymbol{k}_j$ from $\boldsymbol{K}$) can be well approximated by other PSFs taken at nearby locations (columns $\boldsymbol{k}_p$ such that $\|\boldsymbol{s}_p - \boldsymbol{s}_j\|$ is small). Let us select P of these columns $\{\boldsymbol{k}_p \mid p \in \mathbb{G}_P\}$. Each column of $\boldsymbol{K}$ is then approximated by the weighted sum of P columns out of M (typically with $P \ll M$):

$$\boldsymbol{K} \approx \sum_{p \in \mathbb{G}_P} \boldsymbol{k}_p \cdot \boldsymbol{\varphi}_p^{\mathsf{T}}, \tag{21}$$

Interpolation weights $\varphi_p(j)$ are no longer constrained to take binary values. Weights are spatially localized: they are non-zero only on a spatial neighborhood surrounding location $\boldsymbol{s}_p$. The extent of that neighborhood depends on the interpolation order, it corresponds to a square twice the grid step along each dimension for first order (linear) interpolation.

Using this approximation, Eq. (14) becomes:

$$\boldsymbol{H} \approx \sum_{p \in \mathbb{G}_P} \text{conv}(\boldsymbol{k}_p) \cdot \text{diag}(\boldsymbol{\varphi}_p). \tag{22}$$

Weights localization makes decomposition of Eq. (22) very suitable for computations: in this way, full-field convolution computations are not necessary since the preceding weighting operation introduces zeros everywhere except on patches with size twice the grid step. This consideration led Hirsch et al (2010) to suggest to use Eq. (22) for shift-variant blur. The connection with PSF interpolation is however, to the best of our knowledge, not clearly stated in the literature. We further discuss the consequences of this connection in Sec. 3 where formulation (22) is compared with the non-equivalent *image interpolation* formulation introduced by Nagy and O'Leary:

$$\boldsymbol{H} \approx \sum_{p \in \mathbb{G}_P} \text{diag}(\boldsymbol{\varphi}_p) \cdot \text{conv}(\boldsymbol{k}_p), \tag{23}$$

in words, convolve then weight.

2.4 Low-rank approximation on PSF modes

It is often adequate to consider that PSF variations are well captured by a few number of modes; or in an equivalent way, that $\boldsymbol{K}$, being a concatenation of PSFs, has low-rank. A rank-P approximation of matrix $\boldsymbol{K}$ can be expanded as a sum of P rank-one matrices:

$$\boldsymbol{K} \approx \sum_{p=1}^{P} \boldsymbol{c}_p \cdot \boldsymbol{w}_p^{\mathsf{T}}. \tag{24}$$

The approximation error on $\boldsymbol{K}$ can be expressed as the sum of squared differences between corresponding elements in matrix $\boldsymbol{K}$ and its approximation:

$$\mathscr{E}^2(\boldsymbol{A}, \boldsymbol{B}) \equiv \|\boldsymbol{A} - \boldsymbol{B}\|_{\text{F}}^2 = \sum_{i,j} (A_{i,j} - B_{i,j})^2. \tag{25}$$

where $\|\ \|_F$ designates what is known as Frobenius norm. In terms of PSF, it is equivalent to the sum of squared Euclidean distances between corresponding PSF (i.e., columns) in $\boldsymbol{K}$ and its approximation.

By Eckart–Young theorem, the closest rank-P approximation (with minimum error $\mathscr{E}^2$) can be obtained from the singular value decomposition (SVD) of matrix $\boldsymbol{K}$ (Martin and Porter (2012)):

$$\boldsymbol{K} = \boldsymbol{U} \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{V}^{\mathsf{T}} = \sum_{p=1}^{N} \boldsymbol{u}_p \cdot \sigma_p \boldsymbol{v}_p^{\mathsf{T}}, \tag{26}$$

Table 2 Comparison of two non-equivalent formulations for shift-variant blur approximation: Approximation is far better using a decomposition that first applies a weighting, then convolves, rather than the converse as proposed originally in Nagy and O'Leary (1998). The curves represent respectively the exact and approximate PSFs with solid lines (resp. dashed lines), sampled at locations s_1 to s_{10}. Larger errors are observed in the left panel.

	Convolve, then weight	Weight, then convolve
References	Nagy and O'Leary (1998); Calvetti et al (2000); Nagy et al (2004); Preza and Conchello (2004); Bardsley et al (2006); Ng et al (2007); Rogers and Fiege (2011)	Flicker and Rigaut (2005); Hirsch et al (2010); Denis et al (2011); Ben Hadj and Blanc-Féraud (2012); Ben Hadj et al (2012) and Eq. (20), (22), (28) and (38)
Decomposition	$H \approx \sum\limits_{p=1}^{P} \left(\mathrm{diag}(\boldsymbol{w}_p) \right) \cdot \left(\mathrm{conv}(\boldsymbol{c}_p) \right)$	$H \approx \sum\limits_{p=1}^{P} \left(\mathrm{conv}(\boldsymbol{c}_p) \right) \cdot \left(\mathrm{diag}(\boldsymbol{w}_p) \right)$
Recentered PSF approximations k_i		
PSF error ($\times 4$)		
Properties	interpolates blurred images preserves positivity - breaks PSF symmetry - breaks PSF normalization	interpolates PSF preserves positivity + preserves PSF symmetry + preserves PSF normalization
	same computational complexity	

by retaining only the first P left and right singular vectors, weighted by the corresponding (largest) singular values:

$$K \approx \sum_{p=1}^{P} \boldsymbol{u}_p \cdot \sigma_p \boldsymbol{v}_p^{\mathsf{T}}, \qquad (27)$$

with $\boldsymbol{u}_p$ and $\boldsymbol{v}_p$ the p-th left and right singular vectors, and σ_p the corresponding singular value. In contrast to binary weights of piecewise constant PSF, or localized weights used when interpolating a PSF, components of vectors $\boldsymbol{v}_p$ take arbitrary values (positive or negative) and are defined over the whole input field. Vectors $\boldsymbol{u}_p$ can no longer be interpreted as PSF (no natural normalization[3] nor positivity), but rather as PSF modes as obtained by principal component analysis, see Table 1.

---

[3] PSF normalization is generally understood in the context of probability density functions: summation of the PSF must be equal to 1 for energy preservation, while it is the (Euclidean) norm of the left singular vectors $\boldsymbol{u}_p$ that is taken to be 1.

Decomposition of operator H follows directly from approximation K by reasoning as for Eq. (20) and Eq. (22):

$$H \approx \sum_{p=1}^{P} \mathrm{conv}(\boldsymbol{u}_p) \cdot \mathrm{diag}(\sigma_p \boldsymbol{v}_p). \qquad (28)$$

Since weights $\sigma_p \boldsymbol{v}_p$ are not localized, P full-field convolutions must be computed in this approximation, potentially a large computational load when $P \gg 1$.

The decomposition given by Eq. (28) has been proposed in Flicker and Rigaut (2005) and Miraut and Portilla (2012).

2.5 Proposed optimal local approximation

Low-rank decomposition of Eq. (27) is appealing because it is optimal with respect to error $\mathscr{E}^2$. The corresponding weights are however not localized, increasing the computational cost in proportion with the number of added PSF modes. PSF interpolation approach is preferable in this regard since weights localization prevents the computation
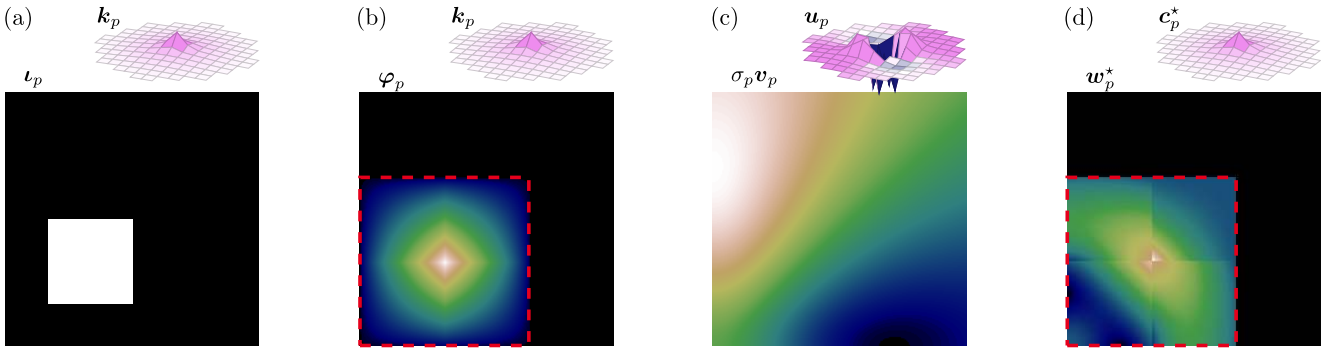
Fig. 1 2D illustration of shift-variant blur approximations: one of the convolution kernels and weights used with (a) piecewise constant PSF, Sec. 2.2; (b) interpolated PSF, Sec. 2.3; (c) decomposition on PSF modes, Sec. 2.4; (d) proposed optimal local approximation, Sec. 2.5. In (b) and (d), weights are non-zero only on a small region drawn with a dashed line (for visualization purposes, a 4×4 grid of PSF is considered here, weights are much more localized with denser PSF grids).

of full-field convolutions, saving computational effort especially for small PSF supports. Let us describe an intermediate solution that, while keeping localized weights, is optimal with respect to error $\mathscr{E}^2$. We define an optimal local approximation of matrix $\boldsymbol{K}$:

$$\boldsymbol{K} \approx \sum_{p=1}^{P} \boldsymbol{c}_p^{\star} \cdot \boldsymbol{w}_p^{\star\mathsf{T}}, \tag{29}$$

where PSF $\{\boldsymbol{c}_p^{\star}\}_{p=1}^{P}$ and weights $\{\boldsymbol{w}_p^{\star}\}_{p=1}^{P}$ are optimal solutions in the following minimization problem:

$$\{\boldsymbol{c}_p^{\star}, \boldsymbol{w}_p^{\star}\}_{p=1}^{P} = \underset{\{\boldsymbol{c}_p, \boldsymbol{w}_p\}_{p=1}^{P}}{\arg\min} \mathscr{E}^2\left(\boldsymbol{K}, \sum_{p=1}^{P} \boldsymbol{c}_p \cdot \boldsymbol{w}_p^{\mathsf{T}}\right). \tag{30}$$

Without constraint on the support of weights $\boldsymbol{w}_p$, the solution to Eq. (30) is given by the truncated SVD (Eq. (27)). We however restrict minimization to the support of interpolation weights $\text{supp}(\boldsymbol{\varphi}_p)$:

$$\forall p, \; \text{supp}(\boldsymbol{w}_p) \subset \text{supp}(\boldsymbol{\varphi}_p), \tag{31}$$

for a fixed PSF interpolation scheme $\{\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_P\}$, see Figure 1.

Eq. (30) is a biconvex minimization, as defined by Gorski et al (2007), for which a partial optimum can be found by alternate convex search. Let us describe how each minimization sub-problem is solved.

Computation of optimal PSF: We consider fixed weights $\boldsymbol{w}_p$, gathered into a single matrix $\boldsymbol{W}$:

$$\boldsymbol{W} = \begin{bmatrix} | & | & & | \\ \boldsymbol{w}_1 & \boldsymbol{w}_2 & \cdots & \boldsymbol{w}_P \\ | & | & & | \end{bmatrix}. \tag{32}$$

Let $\boldsymbol{C}$ be the matrix whose columns are vectors $\boldsymbol{c}_p$:

$$\boldsymbol{C} = \begin{bmatrix} | & | & & | \\ \boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_P \\ | & | & & | \end{bmatrix}. \tag{33}$$

Decompositions (29) and (30) can be rewritten as a single matrix product:

$$\sum_{p=1}^{P} \boldsymbol{c}_p \cdot \boldsymbol{w}_p^{\mathsf{T}} = \boldsymbol{C} \cdot \boldsymbol{W}^{\mathsf{T}}. \tag{34}$$

The matrix $\boldsymbol{C}^{\star}$ of optimal vectors $\boldsymbol{c}_p^{\star}$, given the weights, is:

$$\begin{aligned} \boldsymbol{C}^{\star} &= \underset{\boldsymbol{C}}{\arg\min} \|\boldsymbol{K} - \boldsymbol{C} \cdot \boldsymbol{W}^{\mathsf{T}}\|_{\mathrm{F}}^2 \\ &= \boldsymbol{K} \cdot \boldsymbol{W} \cdot (\boldsymbol{W}^{\mathsf{T}} \cdot \boldsymbol{W})^{-1}. \end{aligned} \tag{35}$$

Proof Frobenius norm can be rewritten using the trace function:

$$\|\boldsymbol{K} - \boldsymbol{C} \cdot \boldsymbol{W}^{\mathsf{T}}\|_{\mathrm{F}}^2 = \mathrm{tr}[(\boldsymbol{K} - \boldsymbol{C} \cdot \boldsymbol{W}^{\mathsf{T}}) \cdot (\boldsymbol{K} - \boldsymbol{C} \cdot \boldsymbol{W}^{\mathsf{T}})^{\mathsf{T}}]$$

whose expansion gives a constant term plus two terms depending upon $\boldsymbol{C}$, whose derivation with respect to matrix $\boldsymbol{C}$ is:

$$2\frac{\partial}{\partial \boldsymbol{C}} \mathrm{tr}(\boldsymbol{K} \cdot \boldsymbol{W} \cdot \boldsymbol{C}^{\mathsf{T}}) = 2\boldsymbol{K} \cdot \boldsymbol{W},$$

and

$$\frac{\partial}{\partial \boldsymbol{C}} \mathrm{tr}(\boldsymbol{C}^{\mathsf{T}} \cdot \boldsymbol{W} \cdot \boldsymbol{W} \cdot \boldsymbol{C}^{\mathsf{T}}) = 2\boldsymbol{C} \cdot \boldsymbol{W}^{\mathsf{T}} \cdot \boldsymbol{W},$$

giving the classical normal equations:

$$\boldsymbol{C} \cdot \boldsymbol{W}^{\mathsf{T}} \cdot \boldsymbol{W} = \boldsymbol{K} \cdot \boldsymbol{W}.$$

proving (35). ☐

In most cases a PSF has a support which is much smaller than the observed field. Let L be the number of pixels in the PSF support, N and M the number of pixels of the blurry and original images respectively. Matrix $\boldsymbol{K}$ can be reduced to a $L \times M$ matrix by discarding the $N - L$ lines of zeros. Matrix $\boldsymbol{C}$ is then $L \times P$ and matrix $\boldsymbol{W}$ is $M \times P$. Eq. (35) involves the inversion of a small $P \times P$ matrix and computing products involving matrix $\boldsymbol{W}$. These products are efficiently computed by skipping the many zeros in $\boldsymbol{W}$ due to the restricted supports of the weights (constraint of Eq. (31)).

(a) image interpolation: approach of Nagy and O'Leary (1998)
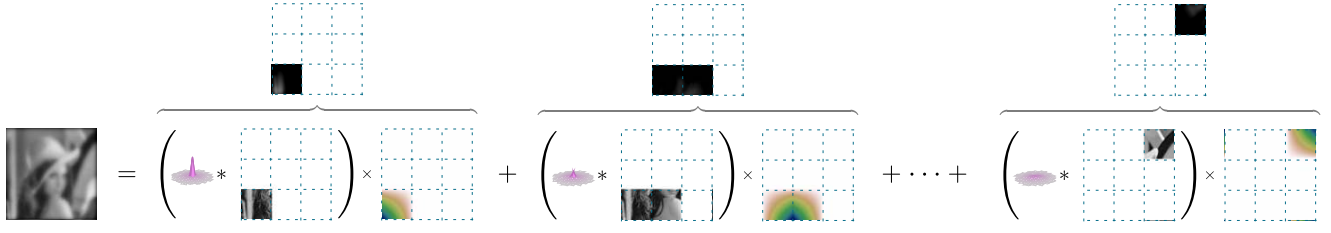
(b) decomposition on PSF modes: approach of Flicker and Rigaut (2005)
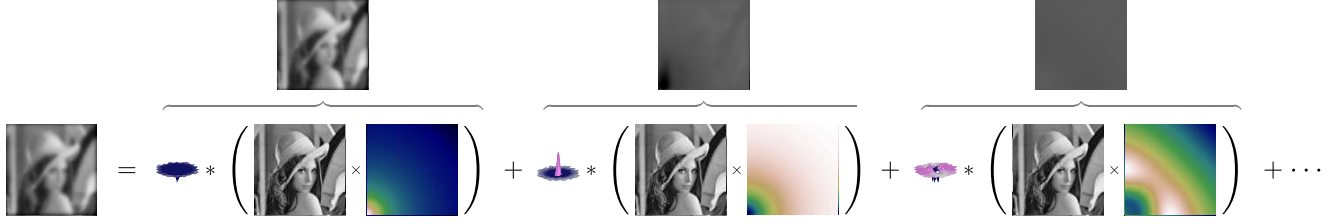
(c) PSF interpolation: shown to be equivalent to the approach of Hirsch *et al.* (2010)
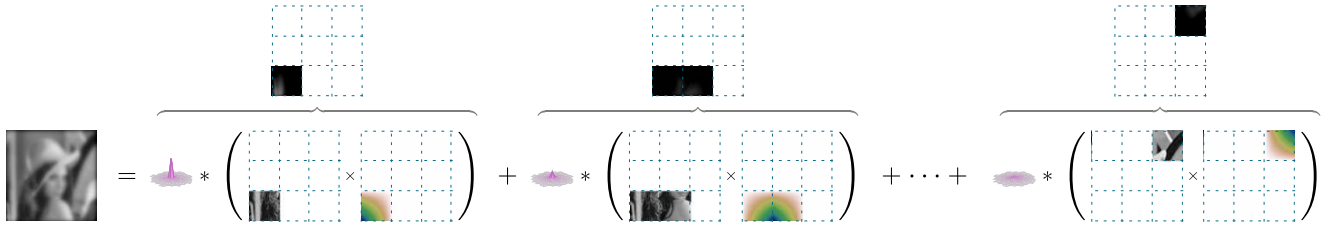
(d) optimal local approximation of PSF: this paper
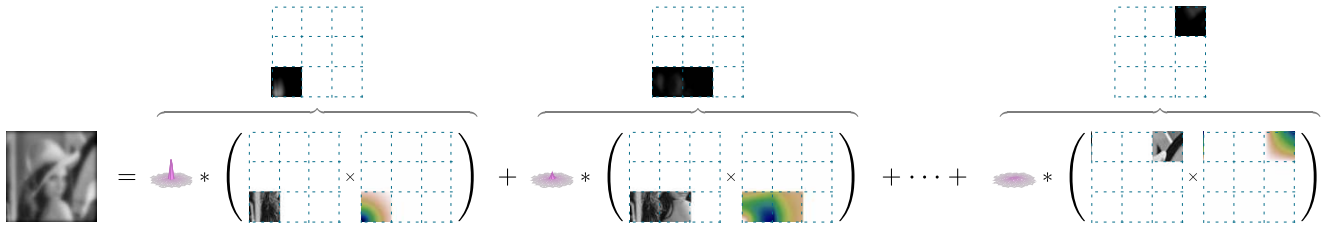
Fig. 2 Shift-variant blur applied to an image with 4 different models: (a) the model of Nagy and O'Leary (1998) first convolves image regions with different PSF and then interpolates the blurry results; (b) Flicker and Rigaut (2005) approximate local PSF on few PSF modes, the image is thus weighted according to the importance of each mode in the decomposition before convolving with PSF modes; (c) we show in this paper that interpolating PSF leads to the model proposed by Hirsch et al (2010), image blocks are first weighted according to the interpolation kernel, then convolved by the PSF; (d) the proposed optimal local approximation of PSF follows the same procedure, with weights and PSF that are chosen so as to minimize approximation error.

Computation of optimal weights: Weights updating requires a minimization of the error $\mathscr{E}^2$ subject to the support constraint of Eq. (31). Direct computation of the optimal weights matrix $\boldsymbol{W}$ as done for $\boldsymbol{C}$ in Eq. (35) would give non-zero values outside the supports. In fact, only few non-zero weights are used to form the approximation of a given PSF $\boldsymbol{K}_{\cdot,j}$. These weights are computed independently from weights at other locations. They are obtained by projection of PSF $\boldsymbol{K}_{\cdot,j}$ on the subspace spanned by active PSFs in $\boldsymbol{C}$, as detailed below.

With the local approximation of Eq. (29), element $K_{i,j}$ is approximated by:

$$K_{i,j} \approx \sum_{p=1}^{P} c_p^{\star}(i)\, w_p^{\star}(j)\,. \tag{36}$$

A given weight $w_p^{\star}(j)$ has a global influence on the jth column of $\boldsymbol{K}$. For any fixed j, only a few $w_p^{\star}(j)$ are non-zero due to weights localization, typically 2^d in d dimensions for supports corresponding to linear interpolation. Let $\boldsymbol{C}'_j$ be the sub-matrix obtained by discarding all columns $\boldsymbol{c}_p$ of $\boldsymbol{C}$ with

index p outside the support of the interpolation weights (i.e., such that $\varphi_p(j) = 0$). Let the weights inside the support be denoted $\boldsymbol{\omega}_j$ (i.e., weights allowed to be non-zero), then column $\boldsymbol{K}_{\cdot,j}$ is approximated by $\boldsymbol{C}'_j \cdot \boldsymbol{\omega}_j$. In contrast to Eq. (34), the product $\boldsymbol{C}'_j \cdot \boldsymbol{\omega}_j$ involves only weights inside the support. Optimal weights estimation boils down to a least squares formulation:

$$\boldsymbol{\omega}_j^{\star} = \arg\min_{\boldsymbol{\omega}_j} \|\boldsymbol{K}_{\cdot,j} - \boldsymbol{C}'_j \cdot \boldsymbol{\omega}_j\|_2^2$$

$$= (\boldsymbol{C}_j'^{\mathsf{T}} \cdot \boldsymbol{C}'_j)^{-1} \cdot \boldsymbol{C}_j'^{\mathsf{T}} \cdot \boldsymbol{K}_{\cdot,j}. \tag{37}$$

Since matrix $\boldsymbol{C}'_j$ is constant within a cell of the interpolation grid, Eq. (37) gives the weights of a cell (i.e., intersection of the supports of interpolation kernels; see an illustration of a support drawn with a dashed line in Fig. 1). $\boldsymbol{C}_j'^{\mathsf{T}} \cdot \boldsymbol{C}'_j$ is a very small $2^d \times 2^d$ matrix (dimension $d = 2$ for 2D images).

To summarize, minimization of the approximation error is performed by an alternate update of vectors $\boldsymbol{c}_p^{\star}$ with equation (35) and weights $\boldsymbol{w}_p^{\star}$ with equation (37). Starting from PSFs and weights given by the model of linear PSF interpolation described 2.3, each step of the alternate optimization further improves the accuracy of the model until a partial optimum is reached. In practice, only a few iterations are necessary to bring significant improvement (10 iterations where used to get the results illustrated in this paper).

The shift-variant blurring operator $\boldsymbol{H}$ is approximated following the decomposition of Eq. (28):

$$\boldsymbol{H} \approx \sum_{p=1}^{P} \mathrm{conv}(\boldsymbol{c}_p^{\star}) \cdot \mathrm{diag}(\boldsymbol{w}_p^{\star}). \tag{38}$$

Optimal vectors $\boldsymbol{c}_p^{\star}$ and optimal weights $\boldsymbol{w}_p^{\star}$ can be computed beforehand (i.e., for a given PSF model $\boldsymbol{H}$). The complexity of approximation (38) is the same as Eq. (23).

Figure 2 illustrates how the different models for a smoothly varying blur are applied on an image. Approaches (a), (c) and (d) process small blocks, while the method of Flicker and Rigaut (2005) that approximates PSFs on few modes operates on the whole image. When PSF variations can not be captured accurately using only the first PSF modes, this latter approach is time consuming; approaches based on local processing of image blocks are preferable. The method of *image interpolation* depicted in Fig. 2(a) is based on the fact that an image with varying blur can be obtained by interpolating between several blurred images computed by convolution with different PSFs. Approximating the blur with PSFs that are (linearly) interpolated from PSFs that are sampled on a grid, as described in section 2.3, leads to the method illustrated in Fig. 2(c). We showed that it is equivalent to the method of Hirsch et al (2010). The proposed optimal local approximation of blur is applied in the same way as the model of Hirsch, using only optimal weights and PSF instead of interpolation weights and sampled PSF.

3 Comparison of blur approximations

Let us now compare the different blur approximations that have been introduced on the standpoint of preservation of PSF properties (e.g., positivity or symmetry), computational efficiency, approximation error, or deblurring quality.

3.1 Equivalent PSF

In the literature devoted to shift-variant blur approximations, the main subject is computational efficiency; most often, the equivalent PSF is not considered. It is however essential to relate a given approximation method with the corresponding approximation in terms of PSF. Some approximations preserve important properties of the original PSF such as normalization, symmetry and positivity. We begin by giving the expression of an equivalent PSF for each model. We then review in section 3.2 the properties preserved by some of the models.

Equivalent PSF for each method: The (centered) PSF $\boldsymbol{k}_j$ for a point source located at $\boldsymbol{s}_j$ is approximated by an equivalent PSF $\tilde{\boldsymbol{k}}_j$ that depends on the model:

With a shift-invariant PSF model (Sec. 2.1):

$$\tilde{\boldsymbol{k}}_j^{(\mathrm{Cst})} = \boldsymbol{k}; \tag{39}$$

with a piecewise constant PSF model (Sec. 2.2):

$$\tilde{\boldsymbol{k}}_j^{(\mathrm{PCst})} = \boldsymbol{k}_p, \text{ with } p \text{ such that } \iota_p(j) = 1; \tag{40}$$

with a model based on PSF interpolation (Sec. 2.3):

$$\tilde{\boldsymbol{k}}_j^{(\mathrm{PSFInterp})} = \sum_{p \in \mathbb{G}_P} \varphi_p(j) \cdot \boldsymbol{k}_p, \tag{41}$$

where $\varphi_p(j)$ are interpolation weights;

with the model of *image interpolation*:

$$\tilde{\boldsymbol{k}}_j^{(\mathrm{ImageInterp})} = \sum_{p \in \mathbb{G}_P} \mathrm{diag}\left(\varphi_p^{\rightarrow j}\right) \cdot \boldsymbol{k}_p, \tag{42}$$

where $\varphi_p^{\rightarrow j}(i)$ is the interpolation weight at location $\boldsymbol{r}_i + \boldsymbol{s}_j$, see derivation in Appendix A;

with a model based on decompositions on modes (Sec. 2.4):

$$\tilde{\boldsymbol{k}}_j^{(\mathrm{Modes})} = \sum_p \sigma_p \boldsymbol{v}_p(j) \cdot \boldsymbol{u}_p; \tag{43}$$

and finally, with the optimal local approximation proposed in Sec. 2.5:

$$\tilde{\boldsymbol{k}}_j^{(\mathrm{OptLoc})} = \sum_p \boldsymbol{w}_p^{\star}(j) \cdot \boldsymbol{c}_p^{\star}. \tag{44}$$

Table 3 Summary of the main properties of shift-variant blur models (P is the number of terms in the approximation)

Method	Reference	Assumptions	Properties	Complexity (convolutions)
interpolate deconvolution results	[A]	slow PSF variations	− no shift-variant PSF model	$\approx P$
piecewise constant PSF	Sec. 2.2	large isoplanatic regions	− strong boundary artifacts	$\approx 1^{\star}$
convolve, then apply linear weighting	[B], Eq. (23)	smooth PSF variations	+ preserves PSF positivity	≈ 4 in 2D*
use linear weighting, then convolve	[C], Sec. 2.3	smooth PSF variations	+ interpolates PSF, + preserves PSF positivity, normalization and symmetry	≈ 4 in 2D*
decompose on PSF modes	[D], Sec. 2.4	PSF captured by few modes	+ optimal global approximation	P
use optimal weighting, then convolve	Sec. 2.5	smooth PSF variations	+ optimal local approximation	≈ 4 in 2D*

**if PSF support is small compared to the size of the regions; for approximations involving the 4 nearest PSFs*
References: [A] Maalouf et al (2011); [B] Nagy and O'Leary (1998); [C] Hirsch et al (2010); [D] Flicker and Rigaut (2005)

3.2 Properties

Depending on the approximation method, the equivalent PSF may fulfill some desirable properties that we list now.

Positivity: PSF of systems based on intensity measurements are necessarily positive-valued: $\forall (i,j), K_{i,j} \geq 0$.

Equivalent PSF with shift-invariant or piecewise constant models are naturally positive. Models based on PSF interpolation or image interpolation also preserve positivity as long as interpolation weights φ are positive. This is in particular the case with linear interpolation. PSF modes are positive- and negative-valued; thus approximations using PSF modes are not guaranteed to preserve positivity. Positivity is not enforced by the proposed local approximation; this can be done by restating criterion (30) under the form of constrained optimization.

Normalization: The incoming flux at location s_j is spread over the support of the PSF $h(\cdot, s_j)$ during the image formation process. If the system is flux-preserving, the sum of PSF coefficients should be equal to one: $\mathbf{1}^{\mathsf{T}} h_j = 1$. In practice, there are some losses in the system; more flux may be lost on the border of the field-of-view than in the center, a phenomenon called "vignetting". Correction of this space-variant attenuation requires an accurate modeling of the total collected flux for each input location, i.e., for all j the sums $\mathbf{1}^{\mathsf{T}} h_j$, or equivalently, the sums $\mathbf{1}^{\mathsf{T}} k_j$.

A shift-invariant PSF model can only account for a system without space-variant attenuation, as will be illustrated in section 3.5. A piecewise constant PSF model will account for piecewise constant changes of the attenuation; it requires a partitioning into very small regions to model smooth changes. Models based on the interpolation of PSF using positive interpolation weights lead to equivalent PSF with

sum $\mathbf{1}^{\mathsf{T}} \tilde{k}_j^{(\mathrm{PSFInterp})} = \sum_p \varphi_p(j) \, \mathbf{1}^{\mathsf{T}} k_p$. These models then account for variable attenuation by interpolating the attenuation of each sampled PSF k_p. In a system with flux conservation, the equivalent PSF conserves the norm of sampled PSF. By contrast, the model based on image interpolation does not preserve summation, even in the case of positive weights where $\mathbf{1}^{\mathsf{T}} \tilde{k}_j^{(\mathrm{ImageInterp})} = \sum_{p,i} \varphi_p^{\to j}(i) \cdot k_p(i)$. Table 2 summarizes the differences between the widespread model with image interpolation and approaches such as PSF interpolation. It can be observed in the simple case of a Gaussian PSF with standard deviation proportional to location s that the equivalent PSF over-estimates the actual PSF close to s_1, and under-estimates it near s_{10} (so that $\|\tilde{k}_1^{(\mathrm{ImageInterp})}\|_1 > \|k_1\|_1$ and $\|\tilde{k}_{10}^{(\mathrm{ImageInterp})}\|_1 < \|k_{10}\|_1$).

Invariance: In many models, the PSFs exhibit some invariance property. Let K be an un-shifted PSF invariant with respect to a linear transform T that preserves un-shifted PSFs: $\forall j, \, T \cdot k_j = k_j$. Operator T may represent for example a symmetry, or a rotation. Needless to say, shift-invariant and piecewise constant PSF models that approximate matrix K with columns sampled from K preserve the invariance with respect to transform T. The model based on PSF interpolation also preserves this invariance since column k_j is approximated by a linear combination of columns sampled from K. By linearity, $T \cdot \tilde{k}_j^{(\mathrm{PSFInterp})}$ is then equal to $\tilde{k}_j^{(\mathrm{PSFInterp})}$, i.e., the equivalent un-shifted PSFs are also invariant with respect to transform T. By contrast, the model based on image interpolation that applies varying weights to each sampled PSF does not preserve the invariance property. This can be again noticed on Table 2 where the reference shift-variant PSFs are Gaussian functions and the approximation arising from the image interpolation approach displays some asymmetry, unlike the interpolation model $\tilde{k}_j^{(\mathrm{PSFInterp})}$. The approximation based on a decomposition on

PSF modes does preserve invariance properties. Equivalent PSF $\tilde{\boldsymbol{k}}_j^{(\text{Modes})}$ are defined in a subspace of the column space of matrix $\boldsymbol{K}$. Each $\tilde{\boldsymbol{k}}_j^{(\text{Modes})}$ could then be re-written as a linear combination of columns of $\boldsymbol{K}$, and thus is unchanged after application of transform $\boldsymbol{T}$. Finally, the optimal local approximation method proposed in section 2.5 also preserves invariance properties. The equivalent un-shifted PSF at location $\boldsymbol{s}_j$ is defined as a linear combination of vectors $\boldsymbol{c}_p^\star$. These vectors necessarily lie in the column space of matrix $\boldsymbol{K}$. Otherwise, the rank of $\boldsymbol{K}$ would increase, in contradiction with the optimality of error $\mathscr{E}^2$.

3.3 Computational complexity

The inversion of shift-variant blur requires many evaluations of the approximate blur model. Each approximation discussed in section 2 requires a different computational effort which is analyzed in this section, regardless of the approximation quality that will be studied in the next section. Table 3 summarizes the properties of each model.

Approximating shift-variant blur by an average shift-invariant PSF is the crudest but fastest method and will be considered as a reference. Let t denote the time required to perform such a convolution for a given image with M pixels and a PSF with a rectangular support of L pixels. The processing time of a piecewise constant PSF model, i.e., constant inside rectangular regions, is also of order t if $L \ll M$ (so that the overhead required to compute values at the outer-border of the regions is negligible). Models based on PSF interpolation have a complexity that depends on the number of dimensions d along which PSFs vary (typically, $d = 1, 2$ or 3) and on the interpolation order o. With first-order interpolation and 2D shift-variant blur, PSFs are approximated by bi-linear interpolation; there are 2^2 non-zero terms in the sum of equation (22). More generally, there are $(o+1)^d$ non-zero terms and if outer-border computations are negligible (the support of weights φ_p being large compared to the support of the PSF), the total computation time is $\approx t \times (o+1)^d$.

A more accurate expression for complexity can be given. Let us consider a rectangular grid of P PSF sampled uniformly. The number of pixels of the crisp image is M, the number of pixels in the support of the PSF is L, and the number of pixels inside a grid cell is on average M/P. Due to the overlap of interpolation weights, each convolution in equation (22) is carried on $(o+1)^d$ cells. These convolutions can be computed using fast Fourier transforms (FFT) with an appropriate padding with $L^{1/d}$ zeros along each dimension. FFTs are then computed on areas of about $[(o+1) \times (M/P)^{1/d} + L^{1/d}]^d$ pixels. Since the whole operator involves P such computations, the total complexity is of order $P d [(o+1) \times (M/P)^{1/d} + L^{1/d}]^d \log[(o+1) \times (M/P)^{1/d} + L^{1/d}]$. If the PSF support is much smaller than

Table 4 Computation time relative to a simple convolution for an approximation of operator $\boldsymbol{H}$ based on linear interpolation of PSF (Sec. 2.3): (a) 512^2 pixels image, with 31×31 pixels PSFs; (b) 1000^2 pixels image, with 101×101 pixels PSFs. On our 3.3 GHz Intel Xeon Processor, a simple convolution takes 19 ms for a 512^2 pixels image (case a) and 43 ms 1000^2 pixels image (case b).

PSF grid:	5×5	10×10	20×20
case (a):	3.7	2.4	3.2
case (b):	4.2	15	39

each cell of the grid of PSF ($L \ll M/P$), the complexity is of order $(o+1)^d \times M \log M$, i.e., it corresponds to $(o+1)^d$ full-size convolutions (with first-order interpolation, it yields a multiplication by 2^d, where $d = 1, 2$ or 3 is the dimension). With very big PSF supports ($L \gg M/P$), the complexity rises to P full-size convolutions.

The method based on image interpolation and ours have the same complexity as the method described for PSF interpolation since convolutions are computed on areas that have similar sizes. By contrast, the decomposition on PSF modes described by Flicker and Rigaut (2005) does not enforce localization of weights. Thus, each of the P convolutions must be computed on the full image support, which is much more costly.

The size of the regions affects the efficiency of FFTs and other operations, which is not visible from the asymptotic computational complexity analysis. Smaller regions better fit into the different levels of memory cache; thus, it may be beneficial to use a finer grid of PSFs, as illustrated in table 4 case (a); larger PSFs involve much larger border effects and rapidly increasing computational times when the grid is refined, as illustrated in case (b).

Implementations of the approximations of shift-variant blur can benefit from parallelism at different levels: SIMD instructions for FFT and multiplications and additions, multi-threaded FFT supported by FFT libraries such as FFTW described in Frigo (1999), and coarse-grained parallelism where each term of the sum is computed in parallel in a different thread. In our implementation, we used both SIMD instructions and coarse-grained parallelism. Table 5 shows that our code scales well with the number of CPU cores: an average speedup of 5.3 is measured with 6 cores in our experiments. Using the same parallelization strategy for the image interpolation approach and the optimal method proposed in section 2.5, we reach similar timings.

3.4 Approximation error

The computational complexity should not be dissociated from the approximation error achievable by each model.

Let us first discuss about the ability of each model to exactly match the reference shift-variant operator when

Table 5 Average computation time for an approximation of operator H based on linear interpolation of PSF (Sec. 2.3): (a) 512^2 pixels image, with 31×31 pixels PSF; (b) 1000^2 pixels image, with 101×101 pixels PSF (Intel Xeon Processor with 6 cores). The speedup factor relative to a single-threaded computation is placed between brackets.

nb. of cores:		1	2	4	6
PSF grid		**time in ms (speedup gain w.r.t. one core)**			
	5×5	71	37 (1.9)	19 (3.7)	14 (5.1)
(a)	10×10	46	22 (2.1)	11 (4.2)	8 (5.8)
	20×20	60	30 (2.0)	15 (4.0)	11 (5.5)
	5×5	182	97 (1.9)	57 (3.2)	47 (3.9)
(b)	10×10	636	325 (2.0)	170 (3.7)	117 (5.4)
	20×20	1680	850 (2.0)	435 (3.9)	293 (5.7)

Table 6 Zernike coefficients used in the "two phase-screens" shift-variant simulation. Only non-zero coefficients are displayed.

aberration name	Zernike coefficients	
defocus	$a_4 = 0.3$	$a'_4 = 0.1$
third order astigmatism	$a_6 = 1.4$	$a'_6 = -1.4$
third order spherical	$a_{11} = 0.1$	$a'_{11} = -0.02$
fifth order coma	$a_{16} = 0.05$	$a'_{16} = 0$
fifth order coma	$a_{17} = 0.02$	$a'_{17} = 0$
fifth order spherical	$a_{22} = -0.5$	$a'_{22} = 0.5$

enough terms are used for the approximation. The piecewise constant PSF model matches the reference operator H when the number of terms P equals the number of input pixels M. Each region is then restricted to a single pixel and indicator vectors ι_j are elements e_j from the canonical basis (i.e., for all (i,j), $e_j(i) = 0$ when $i \neq j$ and $e_j(i) = 1$ when $i = j$). Similarly, using interpolation weights φ_j restricted to a single pixel with $\varphi_j = e_j$ gives a PSF interpolation model that exactly matches the operator H. By contrast, the model based on image interpolation with the same interpolation weights produces an approximation of operator H with an error that is bounded from below (with the consequence of a systematic irreducible error). Setting the interpolation weight φ_j equal to the values e_j in the expression of the equivalent PSF given in equation (42), one obtains:

$$\lim_{\varphi_j \to e_j} \tilde{k}_j^{(\text{ImageInterp})}(i) = \underbrace{k(r_i, \overbrace{r_i + s_j}^{s_q})\Delta_j}_{k_q(i)} \neq \underbrace{k(r_i, s_j)\Delta_j}_{k_j(i)}. \tag{45}$$

In the extreme case of a grid of PSF with the same density as the pixel grid, i.e., with as many terms as pixels in the image for the model based on image interpolation, the equivalent PSF does not correspond to the reference PSF. Only the central value ($r_i = 0$) of the equivalent PSF matches that of the reference PSF. Other values are sampled from PSFs located increasingly farther as the output location r_i in the PSF moves away from the PSF center. The model based on image interpolation offers a similar computational cost as PSF interpolation or optimal localized PSF/weights without reaching perfect approximation with regions as small as a single pixel. This is a serious reason for this model to be disregarded.

Finally, the model based on a decomposition on PSF modes provides an exact representation of the original operator H as long as the number P of terms is at least equal to the rank of H (at most $\min(M, N)$). In the case of optimal localized weights described in section 2.5, the approximation error is zero in the limit case already considered. The

choice $c_j^\star = k_j$ and $w_j^\star = e_j$ for all j corresponding to the PSF interpolation model indeed leads to a zero approximation error ($\mathscr{E}^2 = 0$) and is thus a solution to the optimization problem defined in equation (30).

In order to compare the different approximations of shift-variant blur on a realistic case, we modeled an optical system with wavefront aberrations located in two distinct pupil planes, see Figure 3. These aberrations may come from inhomogeneity of optical indices in the turbulent layers of atmosphere crossed by light in ground-based astronomical observations, or from optical aberrations of a multi-lens system. Depending on the incoming direction of the plane wave created by a point source at infinity, the optical aberrations do not sum up the same way, resulting in a shift-variant blur. Partial pupil masking for large incident angles creates an attenuation effect called vignetting. Variations of the incidence angle of the incoming wave lead to changing tip-tilt aberrations resulting in different shifts of the PSF centers. Using Fraunhofer diffraction theory, Goodman (2008), the PSF writes:

$$h(r, s) \propto \left| \int \Gamma(\rho) \exp\left\{ j\left[\mathscr{P}(\rho) + 2\pi \frac{r \cdot \rho}{\lambda} \right] \right\} d\rho \right|^2, \tag{46}$$

with $\Gamma(\rho)$ the modulus of the complex pupil function:

$$\Gamma(\rho) = \Pi(\rho)\Pi(\rho - s/s_{\max}), \tag{47}$$

and with $\mathscr{P}(\rho)$ the aberrated phase:

$$\mathscr{P}(\rho) = 2\pi \frac{\lambda_{\text{ref}}}{\lambda} \sum_k \left[a_k Z_k(\rho) + a'_k Z_k(\rho - s/s_{\max}) \right], \tag{48}$$

where Z_k is the k-th unnormalized Zernike polynomial as defined in Mahajan (1994), a_k and a'_k are the Zernike coefficients of the first and second phase screens, s and r are angular coordinates in radians, j stands for the imaginary unit, $\Pi(\rho)$ is the indicator function of a disk of unit radius centered at the origin: $\Pi(\rho)$ equals one if $\|\rho\|_2 \leq 1$, otherwise $\Pi(\rho)$ is zero, λ is the optical wavelength and $\lambda_{\text{ref}} = 546\,\text{nm}$ is a reference wavelength. Coefficients $\{a_k\}$ and $\{a'_k\}$ used in our numerical experiments are given in table 6.
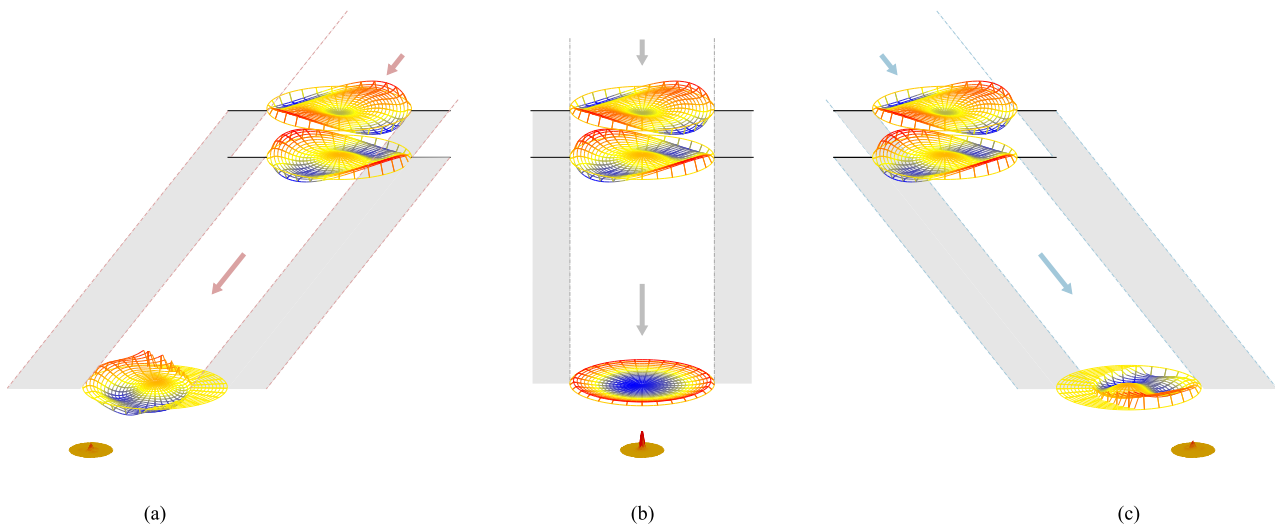
| (a) | (b) | (c) |

Fig. 3 Shift-variant blurring due to phase aberrations and vignetting: an incident plane wave is distorted according to phase aberrations of each lens. In the center of the field (b), most phase aberrations are compensated. Farther from the center of the field, part of the incoming wave front is masked (vignetting), and phase aberrations create a tilt and larger spread of the PSF.
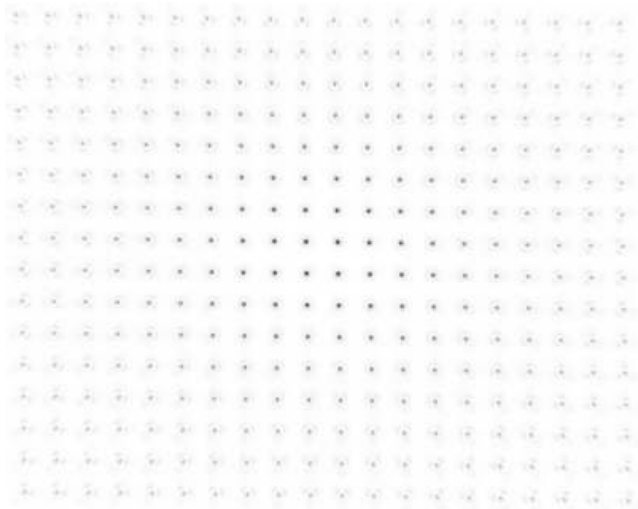
Fig. 4 Grid of PSFs generated from the shift-variant model depicted in Fig.3. Contrast is inverted in order to improve its visualization.

Figure 5 displays approximation error maps for the different shift-variant models in two cases: a coarse approximation using a 5×4 grid of PSFs, and a refined approximation with a 20×16 grid. For the model of Flicker and Rigaut (2005), we set the number of modes to 5 and 13 which represent a computational effort similar to the use of PSF grids. Unsurprisingly, using a shift-invariant PSF or piecewise constant PSF lead to the worst errors. PSF interpolation and image interpolation approaches are comparable when using a coarse grid, while the error is twice as much with image interpolation when the grid is refined (see subfigures (i) and (j)). Decomposition on PSF modes using the method of Flicker and Rigaut (2005) is penalized by the increase of the computational cost in proportion to the number

of modes (errors with 13 modes are over 4 times larger than with PSF interpolation on a 16×20 grid).Such an observation is however application-specific since the approximation error with the method of Flicker and Rigaut (2005) depends on the spectrum of singular values of matrix $\boldsymbol{K}$. Our simulation yields small shifts of the PSF (i.e., geometrical distortions) that reduce strongly the redundancy of the PSF (i.e., the decrease of singular values is slower than it would be with centered PSF). Finally, the highest accuracy is reached using the proposed optimization of localized weights and PSF. Approximation errors are reduced by an order of magnitude without increasing the complexity required to evaluate the operator.

The gain obtained by using weights and PSF refinement is further illustrated on Figure 6. This figure displays the evolution of the approximation error computed on the PSF or on a blurred image as a function of the density of the grid of PSF (solid lines). Approximation errors decrease when the grid is refined, except in some cases where a slight increase of the grid size gives a less favorable sampling of the PSF. From the normalized approximation errors given in log scale on the left-hand side y-axis one sees that PSF and weights optimization improve the approximation by about an order of magnitude. Figure 6 also displays the evolution of the computation times scaled as multiples of a full-size convolution. Increasing the grid size degrades the performance (except for some grid sizes that offer better FFT algorithms or better fit into memory caches). A trade-off must then be reached between the accuracy of the approximation and the time required to apply the operator. Thanks to the improved accuracy of the model with optimized PSF and weights, a given approximation error can be reached at a much lower computational cost.

(a) isoplanatic

(b) piecewise constant (4 × 5 grid)

(c) Nagy and O'Leary (1998) (4 × 5 grid)

(d) PSF interpolation (4 × 5 grid)

(e) Flicker and Rigaut (2005), (5 modes)

(f) proposed optimal local approximation (4 × 5 grid)
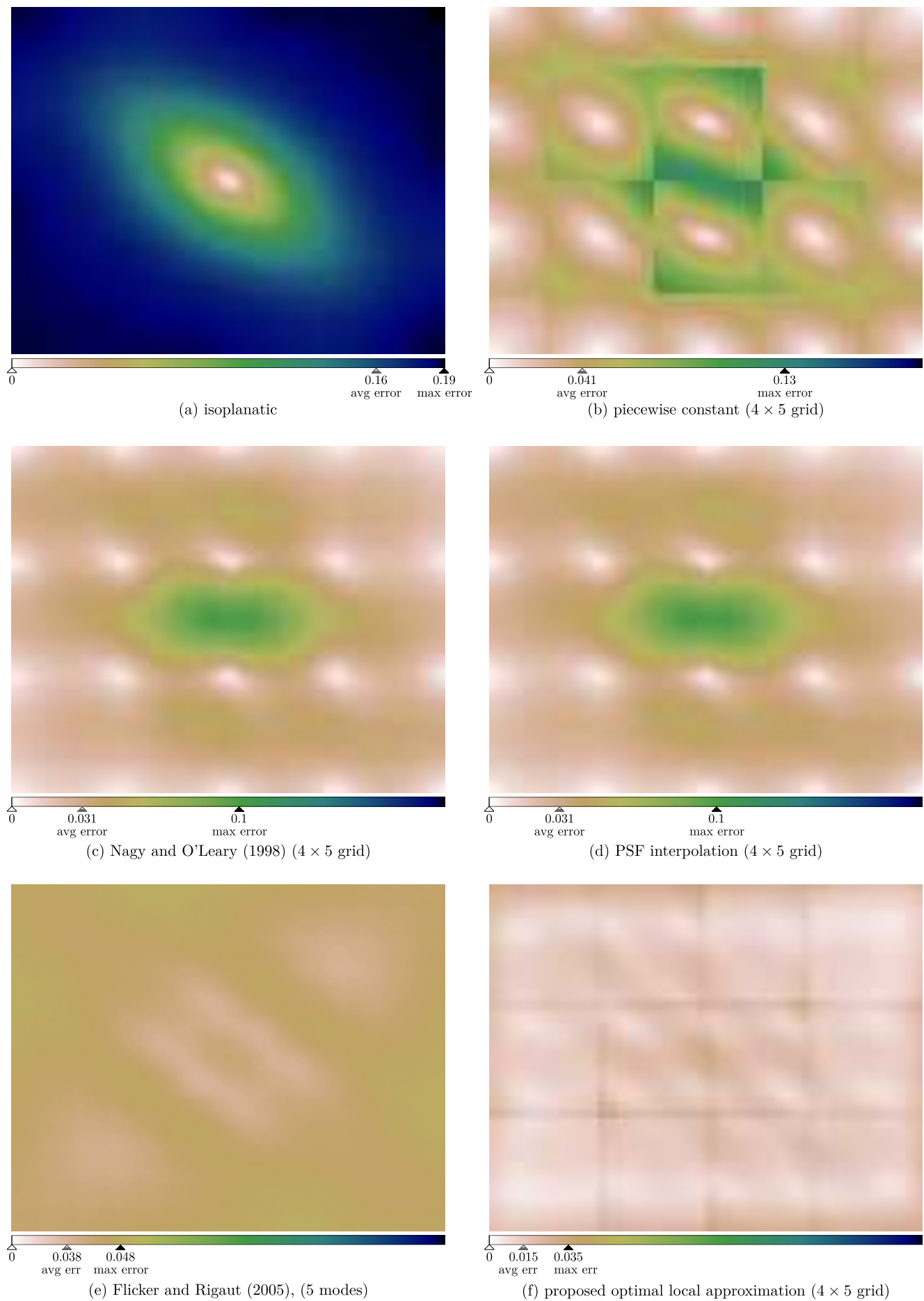
Fig. 5 Approximation errors of a shift-variant blur. Error maps shown give the RMS errors between the true PSF and its approximation depending on the location of the PSFs.

0   0.16   0.19
     avg error   max error

(g) isoplanatic

0.0094   0.04
avg error   max error

(h) piecewise constant (16 × 20 grid)

0.0066 0.024
avg   max
error error

(i) Nagy and O'Leary (1998) (16 × 20 grid)

0.0034 0.013
avg   max
error error

(j) PSF interpolation (16 × 20 grid)

0   0.016 0.031
     avg max
     error error

(k) Flicker and Rigaut (2005), (13 modes)

0.0017 0.0043
avg max
error error

(l) proposed optimal local approximation (16 × 20 grid)

Fig. 5 (cont.) Error maps for more accurate models.

Fig. 6 Trade-off between reduction of the approximation error and increase of the computational cost.

3.5 Deblurring performance

The blur models discussed in this paper are ultimately used within image restoration methods. In order to emphasize the impact of data modeling (i.e., the choice of a specific blur model), we considered numerical simulations in which degraded images were computed using a different PSF at each pixel of the field of view. In the first numerical experiments, no random noise was added, so the discrepancy between the simulated image and the output of a given blur model can only reflect modeling errors. The addition of noise tends to alleviate blur model differences. We used the same physically-grounded PSF model as described in Figure 3 and in the previous paragraph.

All our image restorations were carried out by solving the same (convex) MAP optimization problem (1), i.e., only the blur model $\boldsymbol{H}$ and the regularization weight μ change. Thus, differences in restoration quality can be directly ascribed to the blur model used. Assuming i.i.d. Gaussian noise, the data fidelity term becomes a simple least squares criterion:

$$\Psi_{\text{data}}(\boldsymbol{f}) = \|\boldsymbol{H} \cdot \boldsymbol{f} - \boldsymbol{g}\|_2^2. \tag{49}$$

In order to induce smoothness while preserving edges, we used an hyperbolic approximation of the *total variation* (Rudin et al (1992)) for the regularization of our images:

$$\Psi_{\text{prior}}(\boldsymbol{f}) = \sum_j \sqrt{\|\nabla_j \cdot \boldsymbol{f}\|_2^2 + \varepsilon^2}, \tag{50}$$

with ∇_j a finite difference operator which approximates the spatial gradient $\nabla f(\boldsymbol{s}_j)$ around j^{th} position $\boldsymbol{s}_j$ and $\varepsilon > 0$ a

relaxation parameter to have a criterion that is differentiable whatever $\boldsymbol{f}$. Thanks to this hyperbolic approximation, the image reconstruction problem (1) is a smooth convex minimization problem that can be solved using standard methods for large-scale optimization. We used a limited memory iterative quasi-Newton method[4] (Nocedal (1980)). Another advantage of introduci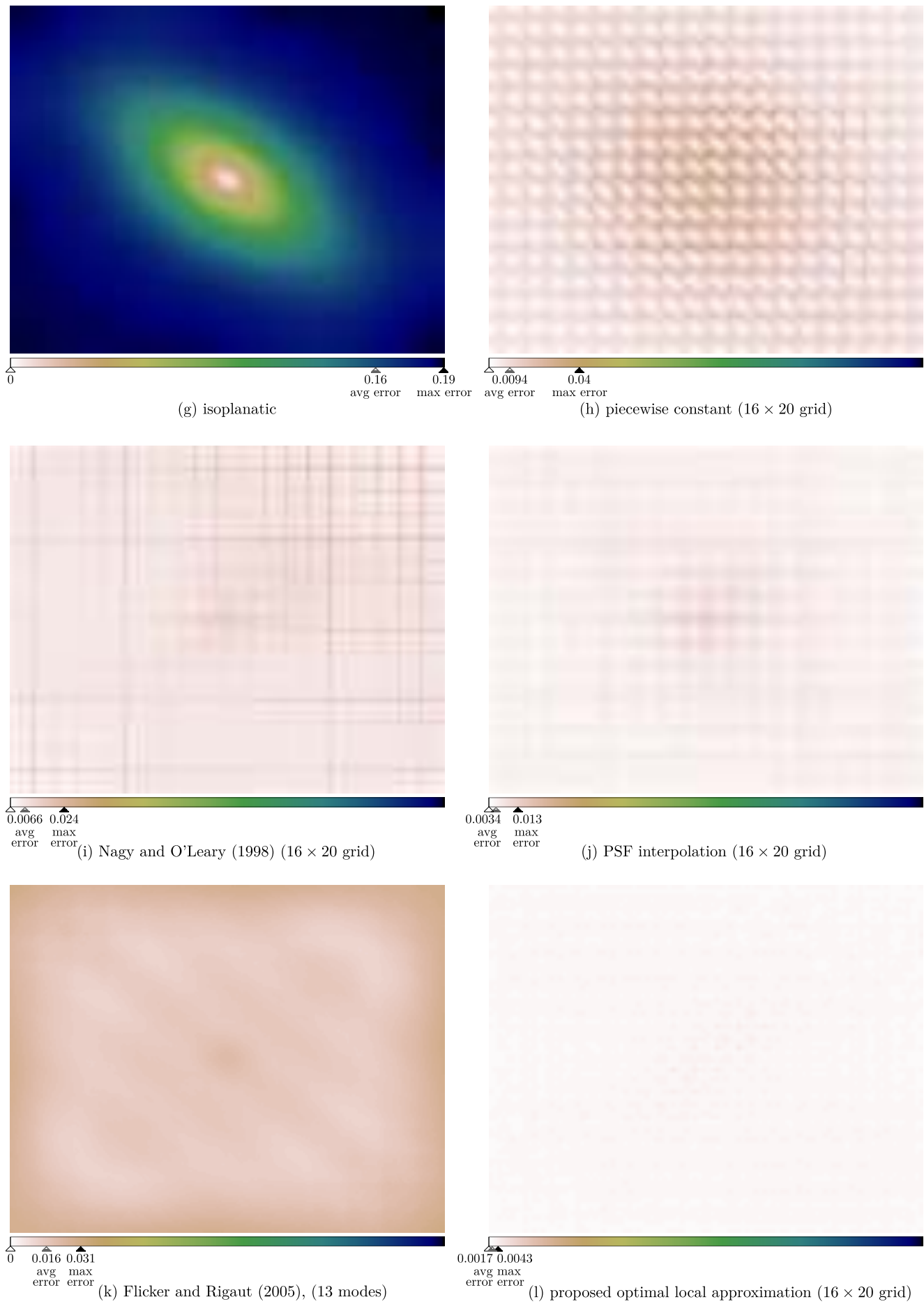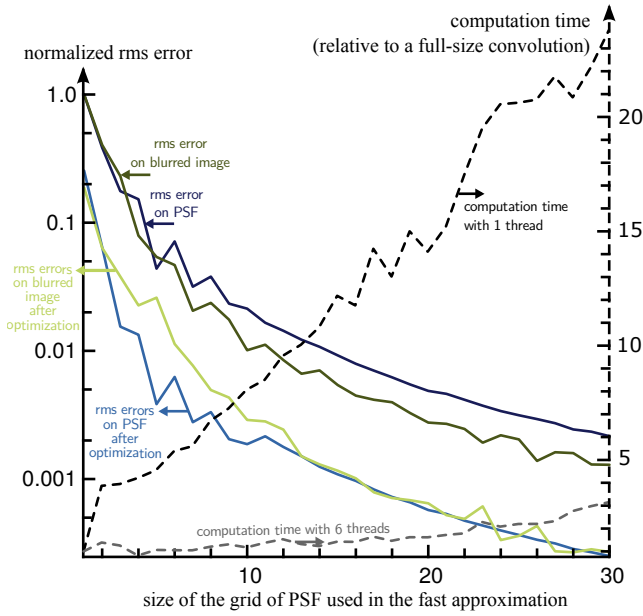ng parameter ε is to achieve an $\ell_2 - \ell_1$ behavior and avoid the classical staircase artifacts created by total variation minimization, Mugnier et al (2001). Choosing the best tuning parameter μ in equation (1) is the subject of intensive researches. For our purposes, since we knew the ground truth image, we simply tuned μ so as to maximize the PSNR. Note that many alternative regularization terms could be considered, such as recent models based on patch statistics: Zoran and Weiss (2011); Chen et al (2014).

To compare the performance of the shift-variant blur models described in this paper, we first consider the restoration of a simple test image made of white stripes on a black background. The simulated blurred version is shown in figure 7(a). Below each image a line profile is plotted corresponding to the red horizontal line. Blur is stronger in the periphery of the image compared to its center. Restoration of this test image using an invariant blur model and the central PSF produces the result shown in figure 7(b). The resolution in the center of the field is slightly improved but most of the blur remains in the periphery. Shift-variant models of Nagy and O'Leary (1998) (figure 7(c)) and bi-linear interpolation of PSFs (figure 7(d)) give comparable performance when a coarse 4×5 grid is used. The model of Flicker and Rigaut (2005) gives better performance when 5 modes are selected (figure 7(e)), probably because the error in PSF modeling is far more homogeneous in the field compared to the two previous models (see figure 5(e) compared to 5(c-d)). It can be noted that the location of the white stripes corresponds to areas where the errors in PSF modeling are the highest for shift-variant models with a 4×5 grid. The proposed blur model with optimal PSF and weights gives a far better restoration, as can be visually assessed on figure 7(f) and in terms of PSNR (9dB improvement compared to the shift-variant model with bi-linear interpolation of weights, with identical computational complexity once the optimal PSFs and weights are computed). Restoration results with 16×20 refined grids or more modes are shown on figure 7(g-j). The difference between image interpolation and PSF interpolation is clearly visible with this denser sampling of PSFs (a 3dB improvement is observed in favor of PSF interpolation). The model of Flicker and Rigaut (2005) offers a weaker improvement when the number of selected modes changes from 5 to 13 modes than other models with similar computational costs. The proposed model with optimized

---

[4] the source code can be downloaded at: `http://cral.univ-lyon1.fr/labo/perso/eric.thiebaut/?Software/OptimPack`

PSFs and modes reaches the best restoration performance. It can be observed that the value of the optimal weight $\mu^\star$ of the smoothness term is smaller when the modeling errors are smaller, i.e., the model of the data is more reliable. The values of $\mu^\star$ reported in figure 7 show a reduction by a factor 150 when replacing the shift-invariant model with the most accurate shift-variant model (i.e., with optimized weights and PSFs on a 16×20 grid).

We then consider the restoration of a natural image displaying different textured areas, figure 8(a). Since the image is in color, we considered a different PSF model for each channel. The phase aberration model described in figure 3 is adapted to different wavelengths by scaling. The resulting blurred image is shown in figure 8(b). We considered that the sensor covers only the area within the red box superimposed on all images. Let γ be the indicator of that area (i.e., $\gamma_k = 1$ iff pixel k is in the area covered by the sensor, otherwise $\gamma_k = 0$). We express our data-fitting term as:

$$\Psi_{\text{data}}(\boldsymbol{f}) = (\boldsymbol{H} \cdot \boldsymbol{f} - \boldsymbol{g})^\mathsf{T} \operatorname{diag}(\boldsymbol{\gamma}) (\boldsymbol{H} \cdot \boldsymbol{f} - \boldsymbol{g}). \tag{51}$$

In order to correctly model the image borders, we reconstruct an image that extends outside the actually observed area, see Soulez et al (2007); Reeves (2005); Matakos et al (2013). We chose to apply separately the relaxed total variation regularization to each color channel, instead of using the color total variation introduced in Blomgren and Chan (1998). Thus, artifacts introduced by each blur model are better identified (i.e., the regularization term does not compensate for inaccuracies of the blur model). The image is 400×320 pixels, PSFs are 51×51 pixels, ε is set to 10 (values in each color channel are in the range $[0, 255]$), and 500 iterations of the quasi-Newton method are performed (it has been checked in practice that 300 iterations were generally sufficient and that 2000 iterations did not bring noticeable improvement).

Restoration using a shift-invariant model gives the image shown in Figure 9(a). Several artifacts can be observed: the original flux is not restored; geometrical distortions are not corrected; and the original resolution is recovered only in the center of the field of view. Using a shift-variant blur model such as PSF interpolation leads to restored images with a more homogeneous resolution. When a coarse PSF grid is used, PSF interpolation and image interpolation produce very similar results, thus only the result for PSF interpolation is given in figure 9(b). The decomposition on few PSF modes fails to correctly model PSF tilt (i.e., decentering). The restored image shown in figure 9(c) suffers from strong color artifacts. The decomposition on PSF modes is not the best method for a simultaneous correction of shift-variant blur and geometrical distortion (i.e., variable PSF decentering). The model with optimized PSF and weights on a small 4×5 grid leads to the best PSNR among coarse models and a significantly improved resolution (e.g.,

region B in the zooms is much better recovered). However, abrupt changes of the PSF model at the boundary between two regions of the grid produce annoying artifacts. These artifacts can be reduced either by considering a refined grid, or by increasing the regularization weight at the cost of a resolution loss. They can be considered as a result of a misuse of the refined model, the grid being too coarse with respect to PSF variations. Figure 9(e-h) gives deblurring results with fine blur models of comparable computational complexity. With a denser grid of PSFs, the difference between deblurring with a model based on image interpolation (figure 9(e)) and PSF interpolation (figure 9(f)) is now noticeable in terms of PSNR (+0.25dB). When going from 5 to 13 modes, the results obtained with the model of Flicker and Rigaut (2005) improve notably. Within a comparable time budget, this model however gives less accurate restorations (PSNR is 2dB smaller). The proposed optimization of weights and PSF again leads to the best results in terms of PSNR. Artifacts at the boundaries of regions that could be observed with a coarse grid disappear with finer grids (compare figures 9(d) and 9(h)).

Table 7 provides further comparison of blur models based on the PSNR improvement brought by shift-variant blur models compared to a shift-invariant model based on the central PSF. Five natural images and three levels of noise are considered. Noise levels are given in terms of BSNR, i.e., the ratio between the noiseless blurred image range and the noise standard deviation, in log scale. In addition to Barbara image, we chose four images from Berkeley's segmentation dataset [5]. Some of the deblurring results of table 7, denoted by a star, are displayed in figure 10. Like in the previous experiments, we selected the number of modes to get comparable computational times with all blur model.

These deblurring results show a significant improvement for all models that model blur variations compared to the shift-invariant model. While the use of an accurate model brings small improvements for moderate SNRs, deblurring is strongly improved by accurate blur modeling in the high SNR regime. Interestingly, improving the SNR brings almost no improvement when using a shift-invariant blur model, for the blur under consideration. This means that in our setting modeling errors dominate stochastic errors from noise, so that only a model refinement can improve the deblurring results. Blur models based on a refined grid or a larger number of PSF modes consistently improve on coarser models. Among each group of models with similar computational complexity, the proposed locally optimal approximation provides the best deblurring performance, sometimes on par with the approach based on PSF interpolation. PSF interpolation and image interpolation be-

---

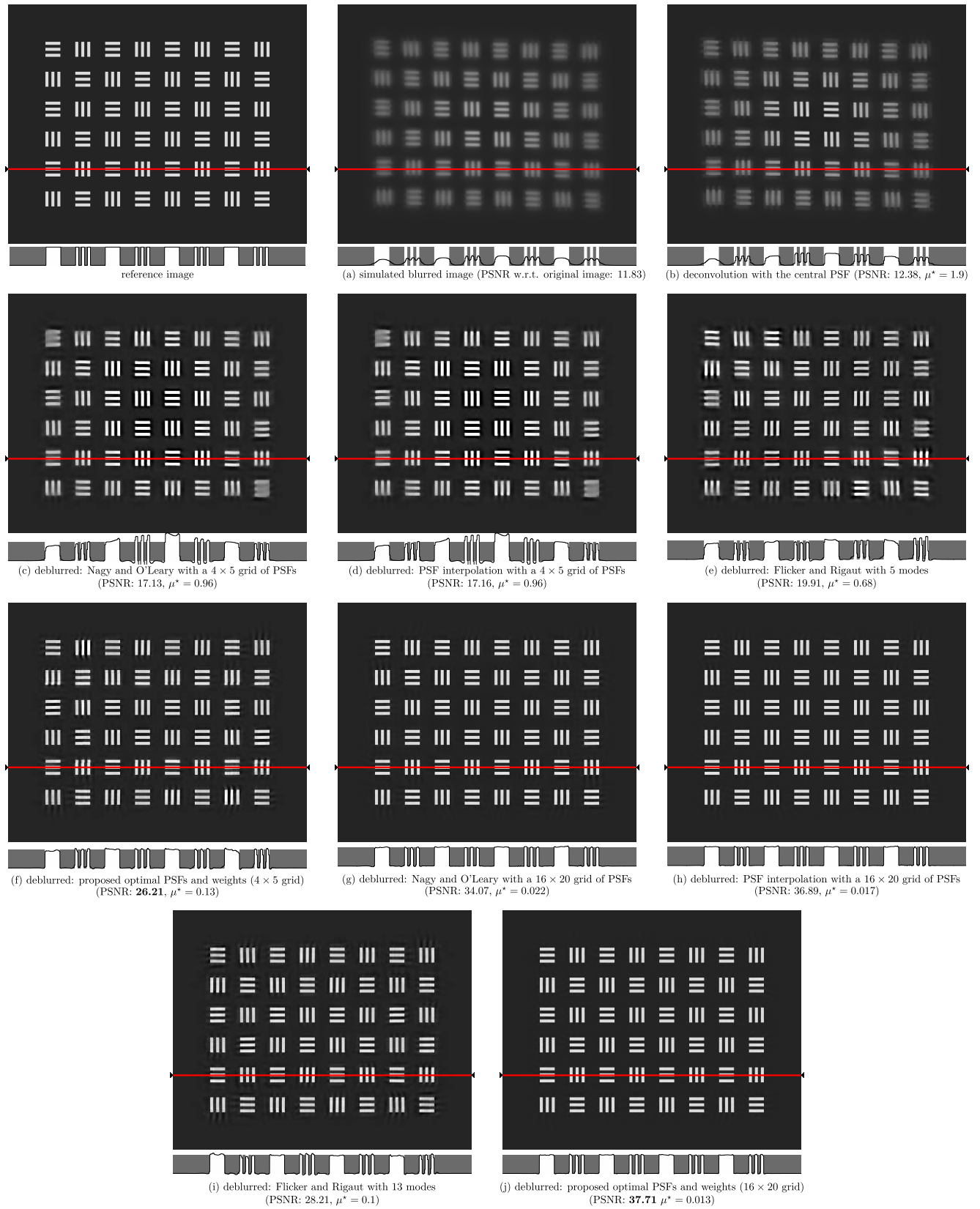**Fig. 7** Restoration of a resolution target degraded by shift-variant blur: (a) degraded image, (b) single-PSF deblurring, (c-f) deblurring with shift-variant PSF models of comparable computational complexity (coarse models), (g-j) deblurring with shift-variant PSF models of comparable computational complexity (fine models). A line profile along the red line indicated by the symbols ▶ and ◀ is drawn below each image.

(a) original image

sensor field of view

(b) simulated blurred image

Fig. 8 Image used to evaluate shift-variant deblurring of a color image: (a) original image; (b) simulated blurry image with a shift-variant chromatic PSF.

Table 7 Restoration performance in terms of PSNR gain in dB with respect to shift-invariant deblurring. The noise level used in our simulations is given in BSNR (see text). The images for the numbers followed by a star * are displayed in Fig. 10.

	image name / BSNR in dB														
	Barbara			**bird**			**man**			**tiger**			**flowers**		
	30	40	60	30*	40	60	30*	40	60	30	40*	60	30	40	60*
blur model															
shift-invariant	17.1	17.1	17.1	15.0*	15.0	15.1	16.1*	16.1	16.1	20.5	20.7*	20.8	20.3	20.5	20.5*
image interpolation (4 × 5 grid)	+5.5	+6.1	+6.4	+3.8	+4.8	+5.0	+8.0	+9.0	+9.3	+3.1	+3.6	+3.7	+5.9	+7.2	+7.7
PSF interpolation (4 × 5 grid)	+5.5	+6.1	+6.4	+3.8*	+4.8	+5.0	+8.0*	+9.1	+9.3	+3.1	+3.6*	+3.7	+5.9	+7.2	+7.7*
PSF modes (5 modes)	+4.7	+5.2	+5.4	+3.5*	+4.2	+4.4	+6.9*	+7.8	+8.0	+3.2	+4.3*	+4.7	+5.7	+7.0	+7.8*
proposed model (4 × 5 grid)	+5.2	+6.2	+7.3	+4.6*	+7.1	+9.0	+8.3*	+10.7	+12.7	+3.5	+5.4*	+7.5	+5.9	+7.9	+10.5*
image interpolation (16 × 20 grid)	+**5.9**	+7.2	+9.6	+4.8	+7.9	+11.8	+9.2	+12.6	+17.3	+**4.0**	+**6.4**	+10.7	+**6.6**	+9.2	+14.6
PSF interpolation (16 × 20 grid)	+**5.9**	+7.3	+9.8	+4.9*	+**8.0**	+12.2	+9.3*	+12.7	+17.9	+**4.0**	+6.4*	+11.0	+**6.6**	+**9.3**	+15.0*
PSF modes (13 modes)	+5.7	+6.8	+8.2	+4.6*	+7.1	+8.8	+8.8*	+11.5	+14.0	+3.9	+6.0*	+8.7	+6.4	+8.8	+12.5*
proposed model (16 × 20 grid)	+**5.9**	+**7.3**	+**9.8**	+4.9*	+**8.0**	+**12.3**	+9.3*	+**12.8**	+**18.1**	+**4.0**	+6.4*	+**11.1**	+**6.6**	+**9.3**	+**15.1***

have similarly when a coarse grid is considered, or for lower values of the SNR. The improved accuracy brought by PSF interpolation models is visible in the high SNR regime and with refined grids (size of the grid should be comparable or smaller than the PSF size). By close inspection of the images in figure 10, block artifacts are visible with the proposed locally optimal method only in the case of very high SNR and coarse grid (sixth row of images, last column).

4 Conclusions

We have provided a unified presentation of several shift-variant blur models that co-exist in the literature. Some of these models had earlier been introduced merely as computational tricks to produce a shift-variant blur effect, within a time lapse comparable to few convolutions. Detailed analysis of how those models relate to particular PSF approximations show that some are more sound, physically speak-

(zoom on 3 regions of the restored image)

A   B   C

(absolute errors in the red channel
dark indicates large errors)

(a) deconvolution with central PSF (PSNR: 17.14, $\mu^\star = 0.36$)

(zoom on 3 regions of the restored image)

A   B   C

(absolute errors in the red channel
dark indicates large errors)

(b) deblurred: PSF interpolation with a 4×5 grid of PSFs (PSNR: 23.54, $\mu^\star = 0.15$)

(zoom on 3 regions of the restored image)

A   B   C

(absolute errors in the red channel
dark indicates large errors)

(c) deblurred: Flicker and Rigaut with 5 modes (PSNR: 22.58, $\mu^\star = 0.13$)

(zoom on 3 regions of the restored image)

A   B   C

(absolute errors in the red channel
dark indicates large errors)

(d) deblurred: proposed optimal PSFs and weights (PSNR: 24.49, $\mu^\star = 0.015$)

Fig. 9 Restoration of a color image degraded by chromatic and shift-variant blur: (a) deconvolution with a shift-invariant model; (b-d) deblurring with shift-variant PSF models of comparable computational complexity (coarse models).

(zoom on 3 regions of the restored image)

A B C

(absolute errors in the red channel
dark indicates large errors)

(e) deblurred: Nagy and O'Leary with a 16×20 grid of PSFs (PSNR: 27.02, $\mu^\star = 0.001$)

(zoom on 3 regions of the restored image)

A B C

(absolute errors in the red channel
dark indicates large errors)

(f) deblurred: PSF interpolation with a 16×20 grid of PSFs (PSNR: 27.37, $\mu^\star = 0.0006$)

(zoom on 3 regions of the restored image)

A B C

(absolute errors in the red channel
dark indicates large errors)

(g) deblurred: Flicker and Rigaut with 13 modes (PSNR: 25.38, $\mu^\star = 0.0096$)

(zoom on 3 regions of the restored image)

A B C

(absolute errors in the red channel
dark indicates large errors)

(h) deblurred: proposed optimal PSFs and weights (PSNR: **27.43**, $\mu^\star = 0.0004$)

Fig. 9 (cont.) Restoration of a color image degraded by chromatic and shift-variant blur: (e-h) deblurring with shift-variant PSF models of comparable computational complexity (fine models).
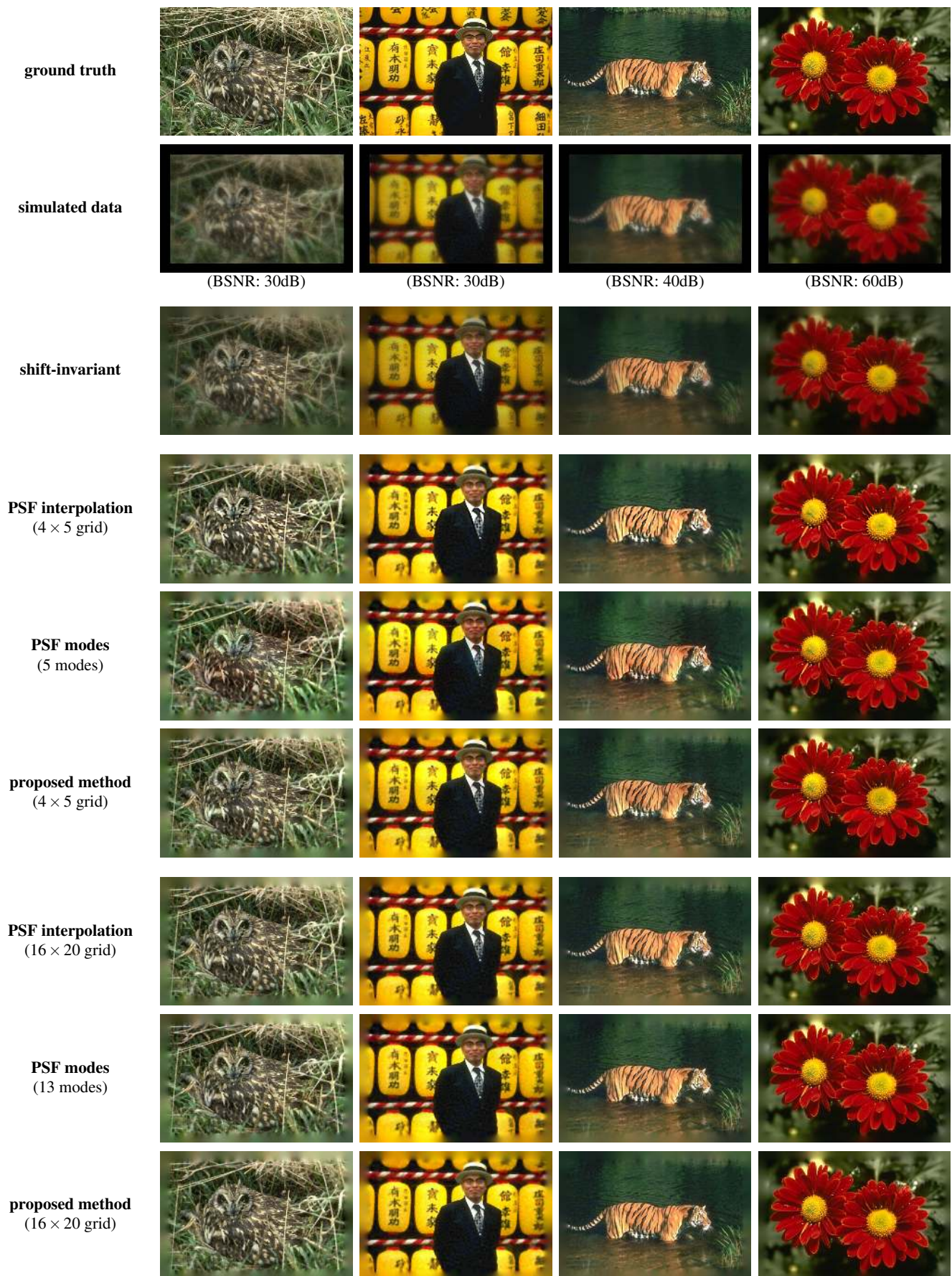
Fig. 10 Illustration of some of the deblurring results from the numerical experiment of Tab. 7.

ing, thus leading to improved performance for shift-variant deblurring. The most natural model follows from assuming that the PSF between two given locations can be interpolated, Denis et al (2011). It leads to a numerical implementation as efficient as the more widespread model based on image interpolation due to Nagy and O'Leary (1998), but with a better grounded approximation yielding systematic improvement of the deblurring quality. Incidentally, this PSF interpolation model has the same mathematical formulation as in Hirsch et al (2010).

These blur models can be refined in order to achieve high accuracy, at the cost of increased computational cost. We proposed to seek, within a given computational budget, the best possible model in terms of approximation quality. We modified the model based on PSF interpolation by considering optimal interpolation weights and PSF samples. This quite simple method improves the RMS modeling error of PSFs by as much as an order of magnitude. When applied to deblurring, this model consistently provides improved restorations, as quantified by PSNR values. Some block artifacts can appear if the grid of PSF is too coarse. These artifacts are due to some discontinuities of PSF approximations between neighboring blocks. They are visible only for high signal to noise ratios and disappear either by increasing the regularization or by refining the grid.

The cost of accurate models can be modest (an increase of computational time typically between a factor 4 to 10 in 2D) compared to convolution by a stationary PSF. The computational complexity should no longer be an obstacle to account for blur variations within the field of view.

Modeling a shift-variant PSF is necessary but not sufficient. The next important task is to properly calibrate this PSF. In practice, this is considerably more difficult than for a shift-invariant PSF because it requires the ability to put reference sources at any spot. The methodology proposed in Delbracio et al (2012) seems a promising approach to estimate shift-variant PSFs. A more general method would be to follow a blind deconvolution approach which is able to estimate the deblurred image and the blur from a single blurred image, see e.g., Ayers and Dainty (1988); Campisi and Egiazarian (2007); Soulez et al (2012) and Ben Hadj et al (2013). In some sense, the PSF and weight optimization strategy that we used in order to improve the quality of the model of the shift-variant blur is already a step in that direction.

A Derivation of the equivalent PSF for the model of Nagy and O'Leary (1998)

Nagy and O'Leary (1998) proposed to approximate the shift-variant blurring operator H by interpolating between the result of several

shift-invariant blur:

$$H \approx \sum_{p \in \mathbb{G}_P} \mathrm{diag}(\boldsymbol{\varphi}_p) \cdot \mathrm{conv}(\boldsymbol{k}_p). \tag{52}$$

Compared to expression (22) for PSF interpolation, the order of the weighting and blurring operations is exchanged. The right-hand side of equation (52) does not correspond to blurring with an equivalent PSF that is interpolated from a few sampled PSF. The equivalent un-shifted PSF $\boldsymbol{k}_j$ at location $\boldsymbol{s}_j$, as defined in section 2, is: $\boldsymbol{k}_j(i) = k(\boldsymbol{r}_i, \boldsymbol{s}_j)\Delta_j = h(\boldsymbol{r}_i + \boldsymbol{s}_j, \boldsymbol{s}_j)\Delta_j$. From equation (52), we get:

$$h(\boldsymbol{r}_i, \boldsymbol{s}_j) = \sum_{p \in \mathbb{G}_P} \boldsymbol{\varphi}_p(\boldsymbol{r}_i) \cdot k_p(\boldsymbol{r}_i - \boldsymbol{s}_j). \tag{53}$$

The equivalent PSF is then:

$$\begin{aligned}
\boldsymbol{k}_j(i) &= h(\boldsymbol{r}_i + \boldsymbol{s}_j, \boldsymbol{s}_j)\Delta_j \\
&= \sum_{p \in \mathbb{G}_P} \boldsymbol{\varphi}_p(\boldsymbol{r}_i + \boldsymbol{s}_j) \cdot k_p(\boldsymbol{r}_i + \boldsymbol{s}_j - \boldsymbol{s}_j)\Delta_j \\
&= \sum_{p \in \mathbb{G}_P} \boldsymbol{\varphi}_p^{\to j}(i) \cdot \boldsymbol{k}_p(i), \tag{54}
\end{aligned}$$

where $\boldsymbol{\varphi}_p^{\to j}$ designates the interpolation weights $\boldsymbol{\varphi}_p$ shifted by the vector $\boldsymbol{s}_j$. Using vector notation, the equivalent PSF can be written as in equation (42) that we recall here:

$$\tilde{\boldsymbol{k}}_j^{(\mathrm{ImageInterp})} = \sum_{p \in \mathbb{G}_P} \mathrm{diag}\left(\boldsymbol{\varphi}_p^{\to j}\right) \cdot \boldsymbol{k}_p.$$

References

Almeida M, Almeida L (2009) Blind deblurring of foreground-background images. In: Image Processing (ICIP), 2009 16th IEEE International Conference on, IEEE, pp 1301–1304

Ayers G, Dainty JC (1988) Iterative blind deconvolution method and its applications. Optics letters 13(7):547–549

Bardsley J, Jefferies S, Nagy J, Plemmons R (2006) A computational method for the restoration of images with an unknown, spatially-varying blur. Opt Express 14(5):1767–1782

Ben Hadj S, Blanc-Feraud L, Aubert G, Engler G (2013) Blind restoration of confocal microscopy images in presence of a depth-variant blur and poisson noise. In: Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, IEEE, pp 915–919

Blomgren P, Chan TF (1998) Color tv: total variation methods for restoration of vector-valued images. Image Processing, IEEE Transactions on 7(3):304–309

Calvetti D, Lewis B, Reichel L (2000) Restoration of images with spatially variant blur by the gmres method. In: Proceedings of SPIE, vol 4116, p 364

Campisi P, Egiazarian K (eds) (2007) Blind image deconvolution: theory and applications. CRC Press

Chacko N, Liebling M, Blu T (2013) Discretization of continuous convolution operators for accurate modeling of wave propagation in digital holography. JOSA A 30(10):2012–2020

Chakrabarti A, Zickler T, Freeman W (2010) Analyzing spatially-varying blur. In: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, pp 2512–2519

Chen Y, Ranftl R, Pock T (2014) Insights into analysis operator learning: A view from higher-order filter-based mrf model. Image Processing, IEEE Transactions on 23(3):1060–1072

Chuang Y, Curless B, Salesin D, Szeliski R (2001) A bayesian approach to digital matting. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, IEEE, vol 2, pp II–264

Cresci G, Davies R, Baker A, Lehnert M (2005) Accounting for the anisoplanatic point spread function in deep wide-field adaptive optics images. A& A 438:757–767, DOI 10.1051/0004-6361: 20052890, `arXiv:astro-ph/0504467`

Delbracio M, Musé P, Almansa A, Morel JM (2012) The non-parametric sub-pixel local point spread function estimation is a well posed problem. International Journal of Computer Vision 96:175–194, 10.1007/s11263-011-0460-0

Demoment G (1989) Image reconstruction and restoration: Overview of common estimation structures and problems. Acoustics, Speech and Signal Processing, IEEE Transactions on 37(12):2024–2036

Denis L, Thiébaut E, Soulez F (2011) Fast model of space-variant blurring and its application to deconvolution in astronomy. In: IEEE Int. Conf. on Image Process. (ICIP)

Escande P, Weiss P (2014) Numerical computation of spatially varying blur operators a review of existing approaches with a new one. arXiv preprint arXiv:14041023

Fish DA, Grochmalicki J, Pike ER (1996) Scanning singular-value-decomposition method for restoration of images with space-variant blur. J Opt Soc Am A 13(3):464–469, DOI 10.1364/JOSAA.13.000464

Flicker R, Rigaut FJ (2005) Anisoplanatic deconvolution of adaptive optics images. J Opt Soc Am A 22(3):504–513

Freeman WT, Adelson EH (1991) The design and use of steerable filters. IEEE Transactions on Pattern analysis and machine intelligence 13(9):891–906

Frigo M (1999) A Fast Fourier Transform compiler. In: ACM Sigplan Notices, ACM, vol 34, pp 169–180

Gilad E, Hardenberg J (2006) A fast algorithm for convolution integrals with space and time variant kernels. J Comp Phys 216(1):326–336

Goodman J (2008) Introduction to Fourier optics. McGraw-hill

Gorski J, Pfeuffer F, Klamroth K (2007) Biconvex sets and optimization with biconvex functions: a survey and extensions. Mathematical Methods of Operations Research 66:373–407, 10.1007/s00186-007-0161-1

Greenspan H, Belongie S, Goodman R, Perona P, Rakshit S, Anderson CH (1994) Overcomplete steerable pyramid filters and rotation invariance. In: Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on, IEEE, pp 222–228

Hirsch M, Sra S, Scholkopf B, Harmeling S (2010) Efficient filter flow for space-variant multiframe blind deconvolution. In: IEEE Comp. Vis. Pattern Recogn., pp 607–614

Levin A (2007) Blind motion deblurring using image statistics. Advances in Neural Information Processing Systems 19:841

Maalouf E, Colicchio B, Dieterlen A (2011) Fluorescence microscopy three-dimensional depth variant point spread function interpolation using Zernike moments. JOSA A 28(9):1864–1870

Mahajan VN (1994) Zernike circle polynomials and optical aberrations of systems with circular pupils. Applied optics 33(34):8121–8121

Martin CD, Porter MA (2012) The extraordinary svd. The American Mathematical Monthly 119(10):838–851

Matakos A, Ramani S, Fessler J (2013) Accelerated edge-preserving image restoration without boundary artifacts. Image Processing, IEEE Transactions on 22(5):2019–2029, DOI 10.1109/TIP.2013.2244218

Miraut D, Portilla J (2012) Efficient shift-variant image restoration using deformable filtering (part i). EURASIP Journal on Advances in Signal Processing 2012(1):1–20

Mugnier LM, Robert C, Conan JM, Michau V, Salem S (2001) Myopic deconvolution from wave-front sensing. JOSA A 18(4):862–872

Nagy J, O'Leary D (1998) Restoring Images Degraded by Spatially Variant Blur. SIAM J Sci Comp 19:1063

Nagy JG, Palmer K, Perrone L (2004) Iterative methods for image deblurring: A Matlab object-oriented approach. Numerical Algorithms 36:73–93, 10.1023/B:NUMA.0000027762.08431.64

Ng J, Prager R, Kingsbury N, Treece G, Gee A (2007) Wavelet restoration of medical pulse-echo ultrasound images in an em framework. Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 54(3):550–568

Nocedal J (1980) Updating quasi-Newton matrices with limited storage. Mathematics of Computation 35(151):773—782, DOI 10.1090/S0025-5718-1980-0572855-7

Perona P (1995) Deformable kernels for early vision. Pattern Analysis and Machine Intelligence, IEEE Transactions on 17(5):488–499

Porter T, Duff T (1984) Compositing digital images. ACM Siggraph Computer Graphics 18(3):253–259

Preza C, Conchello JA (2004) Depth-variant maximum-likelihood restoration for three-dimensional fluorescence microscopy. J Opt Soc Am A 21(9):1593–1601, DOI 10.1364/JOSAA.21.001593

Reeves SJ (2005) Fast image restoration without boundary artifacts. Image Processing, IEEE Transactions on 14(10):1448–1453

Richardson WH (1972) Bayesian-based iterative method of image restoration. JOSA 62(1):55–59

Rogers A, Fiege J (2011) Strong gravitational lens modeling with spatially variant point-spread functions. The Astrophysical Journal 743:68

Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Physica D 60:259—-268, DOI http://dx.doi.org/10.1016/0167-2789(92)90242-F

Sorel M, Flusser J (2008) Space-variant restoration of images degraded by camera motion blur. Image Processing, IEEE Transactions on 17(2):105–116

Soulez F, Denis L, Thiébaut É, Fournier C, Goepfert C (2007) Inverse problem approach in particle digital holography: out-of-field particle detection made possible. JOSA A 24(12):3708–3716

Soulez F, Denis L, Tourneur Y, Thiébaut E (2012) Blind deconvolution of 3d data in wide field fluorescence microscopy. In: Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, IEEE, pp 1735–1738

Titterington DM (1985) General structure of regularization procedures in image reconstruction. Astron & Astrophys 144:381–387

Wang J, Cohen M (2007) Image and video matting: a survey. Foundations and Trends in Computer Graphics and Vision 3(2):97–175

Wei J, Bouman C, Allebach J (2014) Fast space-varying convolution using matrix source coding with applications to camera stray light reduction. Image Processing, IEEE Transactions on 23(5):1965–1979, DOI 10.1109/TIP.2014.2311657

Whyte O, Sivic J, Zisserman A, Ponce J (2010) Non-uniform deblurring for shaken images. In: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, pp 491–498

Zoran D, Weiss Y (2011) From learning models of natural image patches to whole image restoration. In: Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE, pp 479–486