Blind subpixel Point Spread Function estimation from scaled image pairs

Mauricio Delbracio, Andrés Almansa, Jean-Michel Morel, Pablo Musé

Blind subpixel Point Spread Function estimation from scaled image pairs

Mauricio Delbracio[†§] Andrés Almansa [‡]
Jean-Michel Morel [§] Pablo Musé [†]

September 19, 2011

Abstract

In most digital cameras, and even in high-end digital SLRs, the acquired images are sampled at rates far below the Nyquist critical rate, causing aliasing effects. This work introduces a blind algorithm for the subpixel estimation of the point spread function of a digital camera from aliased photographs. The numerical procedure simply uses two fronto-parallel photographs of any planar textured scene at different distances. The mathematical theory developed herein proves that the camera PSF can be derived from the inter-image kernel. Mathematical proofs supplemented by experimental evidence show the well-posedness of the problem and the convergence of the proposed algorithm to the camera in-focus PSF. An experimental comparison of the resulting PSF estimates shows that the proposed algorithm reaches the accuracy levels of the best non-blind state-of-the-art methods.

1 Introduction

Light diffraction, lens aberrations, sensor averaging and antialiasing filters are some of the inherent camera factors that unavoidably introduce blur in images. The blur that results from the combination of all these factors can be modeled locally as a convolution kernel known as *point spread function* (PSF), that corresponds to the space variant impulse response of the whole camera, including the sensor, before the final sampling.

The area enclosed by the first zero crossing of the PSF, usually called Airy pattern, is arguably the most reasonable characterization of the optical system resolution. Top camera/lens manufacturers use charts based on the PSF Fourier

[†]Instituto de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay. mdelbra@fing.edu.uy muse@fing.edu.uy

[‡]CNRS LTCI & Telecom ParisTech, 75634 PARIS Cedex 13, France. andres.almansa@telecom-paristech.fr

[§]CMLA ENS Cachan, 94235 Cachan Cedex, France. morel@cmla.ens-cachan.fr

spectrum modulus (the Modulated Transfer Function, MTF) to describe their products. But accurate knowledge of the PSF is not limited to quality assessment of optical devices, and proves to be extremely useful or even necessary for several image processing tasks such as deblurring [26], super-resolution [27] or shape from defocus [9].

PSF estimation procedures can be either blind or non-blind. Non-blind approaches assume perfect knowledge of a specially designed calibration pattern, and perform local kernel estimation by comparing one or several photographs of the calibration pattern with the ideal calibration pattern. Several patterns have been used for PSF estimation, ranging from pin-hole, slanted-edge [16, 29, 34, 10], or arc-step-edge patterns [18, 17] to random noise images [11, 19, 2, 3, 6].

Blind approaches estimate the PSF from a single or a set of photographs from one or several scenes, whose exact knowledge or exhaustive descriptions are not required. They do assume, however, that the scenes involved in the estimation follow some statistical model of sharp images, or include a significant amount of geometric cues such as sharp edges. Most blind PSF approaches attempt to detect edges, which are modeled as pure step-edge functions convolved with the PSF kernel [8, 22, 7, 30]. In this setting, blind estimation is very ill-posed; to solve the inverse problem, the solution space has to be constrained by considering kernels with a parametric model or with strong regularity assumptions. Such blind estimation techniques are therefore much less accurate than their non-blind counterparts.

The blind PSF estimation problem is closely related to depth-from-defocus (DFD) methods (see for example e.g [28]). However DFD methods typically assume a one-parameter-model for the out-of-focus PSF, e.g. a isotropic gaussian blur or a pillbox kernel in order to regularize the estimation.

In most typical digital cameras, and even in high-end DSLRs, the acquired images are sampled at rates far below the Nyquist critical rate, causing aliasing effects. Until recently, all non-blind sub-pixel PSF estimation methods from aliased images reported in the literature led to ill-posed inverse problems that, like in the blind case, required the imposition of simple parametric models, or equivalently of priors on the regularity or symmetry of the PSF. Nevertheless in a recent work [13] the authors of the present paper have shown that such *a priori* assumptions on the PSF are actually unnecessary and jeopardize the estimation accuracy. Using a realization of white Bernoulli noise as calibration pattern has been demonstrated to be nearly optimal in terms of well-conditioning of the matrix to be inverted. Thus, under reasonable lab conditions, the inverse problem is actually well posed and can deliver a very accurate regularization-free subpixel PSF estimation. Surprisingly, while other approaches using noise patterns had been previously reported, none of them had proposed a regularization-free subpixel PSF estimation scheme [11, 19, 2, 3, 6]. The main goal of the present paper is to extend this methodology to non-blind PSF estimation. This makes it far more practical. Indeed, non-blind PSF estimation approaches rely on a careful setup. The calibration grid has to be properly assembled; a good quality print is necessary. Also, although most methods can handle some variations in illumination and projective deformations, some care in the photograph acquisition

is also needed.

In contrast, the proposed acquisition procedure is simple and handy. It ensures regularization-free subpixel recovery of the PSF from a pair of two photographs of the same scene acquired at different distances with fixed camera configuration. Last but not least, experimental evidence will show that the resulting estimates do not exhibit any significant accuracy loss compared to their best non-blind competitors. The choice of the photographed scene is not critical. For a wide range of everyday scenes, the acquired image pairs lead to well posed inversions and highly accurate results.

The article is organized as follows: Section 2 presents a mathematical model of the image formation carried out by the camera, modeling accurately the aliasing. This model is used in Section 3, where it is shown that the camera PSF can be recovered from a pair of unknown scaled images, under reasonable conditions. The main algorithmic stages of the method derived mathematically in Section 3 are summarized in Section 4. Section 5 discusses a series of experiments on real and simulated images. Finally, Section 6 closes with a brief recapitulation and conclusions. The paper ends with appendices containing a detailed notation and complete mathematical proofs.

Indeed, this paper is written with a dual use in mind: Mathematicians and/or image processing specialists. We have tried to define accurately all mathematical objects necessary to deal rigorously with image formation. An accurate formalism is needed to justify the somewhat intricate interlacement of sampling and convolution operations. This forces one to check on the compatibility of all function or distribution spaces the objects belong to, and to verify that the formulas are mathematically consistent. Nevertheless the application-oriented reader can skip the proofs and the functional space details at a first reading, and simply focus on the standard image processing formalism and algorithms. Most proofs are anyway placed at the end of the paper. A glossary is appended to display all notation in a single place.

2 Image Formation Model

An accurate estimation of the PSF requires a proper modeling of the digital image formation process. The geometric component of this process is most often modeled in computer vision by a pin-hole camera. An *ideal pin-hole camera* with focal length f, shooting at a planar scene u from a distance d and at fronto-parallel pose, will produce an image $w(\mathbf{x}) = u(\lambda\mathbf{x})$ which is just an homothecy of zoom-out factor $\lambda = \frac{d}{f}$ of the original planar scene u. If the pose is not perfectly fronto-parallel, or the pin-hole camera is not ideal (non-standard internal calibration parameters), then the relationship between w and u will be a planar homography D, *i.e.* $w = u \circ D$. In a more accurate camera model (pin-hole + geometric distortion) the distortion D takes the form of a more general (but regular) diffeomorphism. This is required when the scene is a regular close-to-planar surface (as is assumed here), or when the geometric distortion due to the optical system is taken into account as suggested in [34, 18, 13].

For the purpose of PSF estimation this simple model needs to be augmented with an accurate radiometric component, comprising at least the following elements:

1. A model of continuous to digital conversion at the image plane, *i.e.* a *sampling* operator $\mathbf{S}_1$ and additive *noise* $\mathbf{n}$ due to measurement uncertainties;

2. The *blurring* kernel h due to intrinsic camera characteristics, such as diffraction when light goes through a finite aperture, light averaging within the sensor, lens aberration, etc.[1]

The whole image formation process can be summarized in a single equation

$$\tilde{\mathbf{v}} = g\left(\mathbf{S}_1\left((u \circ D) * h\right)\right) + \mathbf{n},$$

where $g(\cdot)$ is a monotone non-decreasing function that describes the non-linear sensor response. If the camera is working outside the saturation zone, the response should be linear, since the procedure deals directly with RAW images. Therefore, the function g will be assumed linear. It boils down to a rescaling of the dynamics of u and therefore disappears w.l.o.g. from the model. Hence, in the sequel, the image formation model will be

$$\text{(M)} \quad \tilde{\mathbf{v}} = \mathbf{S}_1\left((u \circ D) * h\right) + \mathbf{n}.$$

Since in practice our data consist exclusively of discrete sequences (or digital images), the image formation model will be rewritten in terms of discrete sequences. This requires the introduction of additional notation. It would be cumbersome to verify systematically all regularity requirements on all functions and distributions needed in the proofs. Thus, all necessary results are given in a precise form in the appendices. They will be invoked in the proofs and the reader is invited to check that their use was licit.

NOTATION In the sequel, *continuous images* are defined for $\mathbf{x} \in \mathbb{R}^2$, whereas *digital images* are sampled on a discrete grid $\mathbf{k} \in \mathbb{Z}^2$. $\mathbf{F}$ denotes the Fourier transform and $\hat{f}$ denotes the Fourier transform of a function f. The Shannon-Whittaker interpolator defined as $\mathbf{I}_1\mathbf{u}(\mathbf{x}) = \sum_{\mathbf{k}} \mathbf{u}(\mathbf{k})\text{sinc}(\mathbf{x} - \mathbf{k})$ is denoted by $\mathbf{I}_1$, $\mathbf{S}_1$ is the 1-sampling operator such that $\mathbf{u}(\mathbf{k}) = (\mathbf{S}_1 u)(\mathbf{k}) = u(\mathbf{k})$. The filter W_w is an ideal low-pass filter that cuts the sprectrum of a continuous signal to $[-w\pi, w\pi]^2$. Let us denote by $\mathcal{S}_s$ the s-to-1-resampling operator $\mathcal{S}_s = \mathbf{S}_1 H_s \mathbf{I}_1$ and the continuous homothecy $H_\lambda u(x, y) = \lambda^2 u(\lambda x, \lambda y)$. (*i.e.* $\lambda > 1$ is a zoom-out). The digital Nyquist homothecy operator of parameter α is defined by $\mathbf{H}_\alpha\mathbf{u} := \mathbf{S}_1 W_1 H_\alpha \mathbf{I}_1 \mathbf{u}$. We also denote the linear map associated to the convolution with a digital image $\mathbf{u}$ by $\mathbf{C}[\mathbf{u}]$. Let $\mathbf{L}$ be a linear operator, and denote

[1] Note that other blur effects, like motion or defocus blur, that may change from one snapshot to another will be carefully avoided by the experimental procedure.

by $\mathbf{L}^*$ its adjoint, and by $\mathbf{L}^+$ (if exits) its left pseudo-inverse $\mathbf{L}^+ := (\mathbf{L}^*\mathbf{L})^{-1}\mathbf{L}^*$.

A more precise definition of each term is presented in Appendix A.

Suppose that the PSF h is s-band-limited, that is $\text{supp}(\hat{h}) = [-s\pi, s\pi]^2$. Then, if sampled at a rate s, the Nyquist sampling theorem guarantees a perfect reconstruction of h from its samples $\mathbf{h} = \mathbf{S}_1 H_{\frac{1}{s}} h$. We are actually interested in the case $s > 1$, usual for digital cameras. This means that the images obtained from (M) are aliased, being obtained after convolution with h and 1-sampling.

The image formation model (M) can be written in terms of discrete sequences, which are the real numerical data that are dealt with:

$$\tilde{\mathbf{v}} = \mathcal{S}_s(\bar{\mathbf{u}}_D * \mathbf{h}) + \mathbf{n}, \tag{1}$$

where the digital image $\bar{\mathbf{u}}_D = \mathbf{S}_1 W_1 H_{\frac{1}{s}} u(D(\cdot))$ is a well-sampled version of the distorted image $u \circ D$. The value s is the resampling rate from the high resolution lattice $s\times$, where the PSF estimation will take place, to the $1\times$ sensor grid. The derivation of Eq. (1) follows from Lemma 6, given in the appendix.

Equation (1) can be rewritten as

$$\tilde{\mathbf{v}} = \mathcal{S}_s \mathbf{C}[\bar{\mathbf{u}}_D]\mathbf{h} + \mathbf{n}. \tag{2}$$

The numerical method will recover only a finite number of samples $\mathbf{h}$ of h. Strictly speaking h being band limited cannot be compactly supported. Nonetheless, the error introduced by assuming that the support of h is bounded will prove negligible in comparison to the other sources of error: image noise, quantization, slight estimation errors of D, etc. Indeed, the retrieved solution $\mathbf{h}$ will prove to be experimentally independent from variations of its assumed support as long, as it is large enough for errors to be negligible, and small enough for the operator to be still well conditioned.

When $\mathbf{n}$ is a zero-mean white discrete Gaussian noise, it follows from the previous formula that $\mathbf{h}_e = (\mathcal{S}_s \mathbf{C}[\bar{\mathbf{u}}_D])^+ \tilde{\mathbf{v}}$ is an unbiased estimator of $\mathbf{h}$, as long as the linear operator $\mathcal{S}_s \mathbf{C}[\bar{\mathbf{u}}_D]$ is injective. It can be shown that the estimator variance is proportional to the Hilbert-Schmidt norm of $(\mathcal{S}_s \mathbf{C}[\bar{\mathbf{u}}_D])$ (for matrices, the Frobenius norm[2]), and that it is nearly minimal when $\mathbf{u}_D$ is a white noise realization (see [13]).

3 PSF Estimation from two unknown scaled images

Assume that we perfectly know the latent sharp image u that produced the blurry aliased observation $\mathbf{v}$. Under this non-blind assumption solving for the

[2]Recall that the Hilbert-Schmidt norm is $\sum_i \|\mathbf{L}e_i\|^2$, where $\{e_i\}$ is any Hilbert basis of the domain of $\mathbf{L}$. If the linear operator is a matrix then the Hilbert-Schmidt norm is the Frobenius norm of the matrix.

PSF amounts to solve an inverse problem governed by the image formation model (M). The common approach that relies on regularization techniques allows to correctly recover the geometric and radiometric distortions (D and g), but severely distorts the high-frequency components of h. Notwithstanding, it is shown in [13] that this non-blind inverse problem is well posed as long as a white noise image u is chosen as calibration pattern, and that it yields an unbiased estimation of h, no regularization being used. In contrast the widespread calibration patterns based on step-edges (see for instance [18]) lead to ill-posed inverse problems, thus precluding an accurate subpixel estimation of h, unless the family of admissible kernels h is drastically reduced by regularization techniques.

Between these two extremes (perfect step edges and noise patterns), many highly textured natural scenes exist which, while not being optimal, potentially lead to a well-posed inverse problem. They will be exploited in this article to circumvent the non-blind hypothesis, by taking two snapshots of the same scene.

More precisely, this section will prove that the complete recovery of the camera PSF is theoretically possible, from the estimate of the kernel blur between two images of the same scene from different distances. The presentation is divided in two parts. First this kernel will be characterized for a pair of fronto-parallel views of a planar scene, and precise conditions will be given under which this inter-image kernel can be estimated. Then the camera PSF will be derived from the inter-image kernel, under very weak and reasonable conditions.

3.1 Estimating the relative blur between two images

Consider two digital images $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2$ of the same planar scene u, captured from different distances in a fronto-parallel position with negligible rotation around the optical axis. Let λ_1 and λ_2 denote the corresponding zooms between the scene and each of the images. This can be written

$$\begin{aligned} \tilde{\mathbf{v}}_i &= \mathbf{S}_1 H_{\lambda_i} u * h + \mathbf{n}_i \quad \text{for } i = 1, 2 \\ &= \mathbf{S}_1 v_i + \mathbf{n}_i \\ &= \mathbf{v}_i + \mathbf{n}_i, \end{aligned} \tag{3}$$

where $v_i := H_{\lambda_i} u * h$ and $\mathbf{v}_i := \mathbf{S}_1 v_i$. We will realistically assume that $h \in L^1 \cap \mathcal{BL}^2_0(\mathbb{R}^2)$ is non-negative with $\|h\|_{L^1} = 1$, and $u \in L^1(\mathbb{R}^2)$. (This classic notation is recalled at the end of the paper. The reader may skip the functional considerations at first reading). Also, it will be assumed that the acquisition distances are such that $s\lambda_1 < \lambda_2$; the importance of this assumption will soon become clear.

Definition 1. *Let $v_1, v_2 \in \mathcal{BL}^2_0(\mathbb{R}^2)$ be two fronto-parallel continuous views of the same scene, acquired from different distances λ_1 and λ_2 respectively. We call* ***inter-image kernel*** *between v_1 and v_2, any kernel $k \in \mathcal{BL}^2_0(\mathbb{R}^2)$ satisfying*

$$v_2 = H_{\lambda_2/\lambda_1} v_1 * k.$$

The following lemma provides a characterization of the inter-image kernel.

Lemma 1. *Let $h \in L^1 \cap \mathcal{BL}_0^2(\mathbb{R}^2)$ be non-negative, band-limited with supp$(\hat{h}) \subset [-s\pi, s\pi]^2$ and $\hat{h}(0) = 1$. Let ρ be the largest positive number such that $\hat{h}(\zeta) > 0$ for every $\|\zeta\|_\infty < \rho\pi$ and assume that $\lambda_2\rho \geq s\lambda_1$. Then there is an inter-image kernel $k \in \mathcal{BL}_0^2(\mathbb{R}^2)$ with support in $[-s\pi, s\pi]^2$ between (fronto-parallel views) v_1 and v_2 that satisfies*

$$H_\lambda h * k = h, \qquad \text{where } \lambda = \frac{\lambda_2}{\lambda_1}. \tag{4}$$

If $\hat{u}$ does not vanish inside $[-s\frac{\pi}{\lambda_2}, s\frac{\pi}{\lambda_2}]$ then the inter image-kernel is unique and only depends on h and λ.

Proof. In the Fourier domain

$$\mathbf{F}(H_\lambda v_1)(\zeta) = \hat{v}_1\left(\zeta/\lambda\right) = \hat{u}\left(\zeta/\lambda_2\right)\hat{h}\left(\zeta/\lambda\right).$$

On the other hand

$$\hat{v}_2\left(\zeta\right) = \hat{u}\left(\zeta/\lambda_2\right)\hat{h}\left(\zeta\right).$$

If k is an inter image kernel between v_1 and v_2 we have

$$\mathbf{F}(H_\lambda v_1)(\zeta)\hat{k}(\zeta) = \hat{v}_2(\zeta).$$

Thus by Definition 1 k is an inter-image kernel if and only if

$$\hat{u}\left(\zeta/\lambda_2\right)\hat{h}\left(\zeta/\lambda\right)\hat{k}(\zeta) = \hat{u}\left(\zeta/\lambda_2\right)\hat{h}\left(\zeta\right).$$

Thus any kernel k satisfying $\hat{h}\left(\zeta/\lambda\right)\hat{k}(\zeta) = \hat{h}\left(\zeta\right)$ where $\hat{u}\left(\zeta/\lambda_2\right)$ is not zero will be an inter-image kernel.

Since $\hat{h}$ is the Fourier transform of an $L^1(\mathbb{R}^2)$ image, it is continuous and, since $\hat{h}(0) = 1 > 0$, then ρ is necessarily positive. In addition, as $\lambda > \frac{s}{\rho}$ by hypothesis, $\mathbf{F}(H_\lambda h)(\zeta) = \hat{h}(\zeta/\lambda)$ does not vanish inside $[-s\pi, s\pi]^2$. Thus

$$\hat{k}(\zeta) = \frac{\hat{h}(\zeta)}{\hat{h}(\zeta/\lambda)} \tag{5}$$

is well defined all over its support supp$(\hat{k}) \subset [-s\pi, s\pi]^2$ and therefore so is k. Moreover, $k \in L^2(\mathbb{R}^2)$, being bandlimited with $\hat{k}$ is continuous. Finally, if $\hat{u}\left(\zeta/\lambda_2\right)$ is not zero in the support of $\hat{h}$, k is unique. □

Remark 1. *In the previous Lemma 1 it is assumed that the* PSF *h is the same for the two images. This has at least two practical implications. First, extreme precautions should be taken to ensure that both images are strictly in focus to avoid inconsistence with the mathematical formulation. Also, we are assuming that as long as both images are taken in focus the* PSF *remains constant (of course with all the other camera parameters fixed). Secondly, since the common area between $\mathbf{v}_1$ and $\mathbf{v}_2$ is an important part of the closest image, i.e. $\mathbf{v}_1$* PSF *space variance all over the image $\mathbf{v}_1$ may degrade the estimation. This can be avoided taking images that cover only the central part of the image, where the assumption of having a uniform kernel is reasonable.*

Remark 2. *The constraint on $\lambda > s/\rho$ ensures that the continuous zoomed-in image is not destroying any necessary information to generate the continuous zoomed-out version.*

The next goal is to estimate k. Some intuition will be built up first on how to derive the proposed estimator. Conditions will also be given under which the estimator is unbiased. Being s-bandlimited k can be exactly reconstructed from its $s\times$ samples denoted by $\mathbf{k}$. Let us denote by $\mathring{\mathbf{v}}_1 = \mathbf{S}_1 W_1 H_{\frac{\lambda}{s}} v_1$ a well sampled homothety of parameter $\frac{\lambda}{s}$ of v_1.

Proposition 1. *Under the assumptions of Lemma 1 if the operator $\mathcal{S}_s\mathbf{C}[\mathring{\mathbf{v}}_1]$ is injective, then*

$$\mathbf{k} = (\mathcal{S}_s\mathbf{C}[\mathring{\mathbf{v}}_1])^+\mathbf{v}_2.$$

Proof. Since k satisfies Eq. (4), it is an inter-image kernel between v_1 and v_2. Then,

$$\begin{aligned}\mathbf{v}_2 &= \mathbf{S}_1(v_2)\\ &= \mathbf{S}_1(H_\lambda v_1 * k).\end{aligned}$$

Using that k is s-band-limited, it follows that

$$\mathbf{v}_2 \overset{(14)}{=} \mathbf{S}_1(W_s H_\lambda v_1 * W_s k). \tag{6}$$

By applying the Nyquist-Shannon theorem for a band-limited signal, and other properties that are detailed in the Appendix we have

$$\begin{aligned}\mathbf{v}_2 &\overset{s\geq 0}{=} \mathbf{S}_1 H_s H_{\frac{1}{s}}(W_s H_\lambda v_1 * W_s k)\\ &\overset{(16)}{=} \mathbf{S}_1 H_s(H_{\frac{1}{s}} W_s H_\lambda v_1 * H_{\frac{1}{s}} W_s k)\\ &\overset{(15)}{=} \mathbf{S}_1 H_s(W_1 H_{\frac{\lambda}{s}} v_1 * W_1 H_{\frac{1}{s}} k)\\ &\overset{(12)}{=} \mathbf{S}_1 H_s \mathbf{I}_1 \mathbf{S}_1(W_1 H_{\frac{\lambda}{s}} v_1 * W_1 H_{\frac{1}{s}} k)\\ &\overset{(17)}{=} \mathbf{S}_1 H_s \mathbf{I}_1(\mathbf{S}_1 W_1 H_{\frac{\lambda}{s}} v_1 * \mathbf{S}_1 W_1 H_{\frac{1}{s}} k)\\ &\overset{\text{def}}{=} \mathbf{S}_1 H_s \mathbf{I}_1(\mathring{\mathbf{v}}_1 * \mathbf{k})\\ &\overset{\text{def}}{=} \mathcal{S}_s(\mathring{\mathbf{v}}_1 * \mathbf{k})\\ &\overset{\text{def}}{=} \mathcal{S}_s\mathbf{C}[\mathring{\mathbf{v}}_1]\mathbf{k}.\end{aligned} \tag{7}$$

The result follows from the fact that $\mathcal{S}_s\mathbf{C}[\mathring{\mathbf{v}}_1]$ is injective, thus it can be left inverted.

□

Remark 3. *If $\lambda < s$ then the convolution of $\mathbf{k}$ with $\mathring{\mathbf{v}}_1$ is not invertible so the operator $\mathcal{S}_s\mathbf{C}[\mathring{\mathbf{v}}_1]$ will not be injective. This constraint on λ ensures that the zoomed-in image presents enough information to generate the digital zoomed-out version.*

All digital images are in practice defined in a compact domain. It can be assumed realistically that the kernel support $\mathbf{k}$ is much smaller than the support of $\tilde{\mathbf{v}}_2$. It is shown in [13] that for a typical $s = 4$ value the operator $\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\lambda}}\tilde{\mathbf{v}}_1]$ is injective and the pseudo inverse is well posed as long as $\tilde{\mathbf{v}}_1$ is close to white noise. Therefore the chosen scene image u should exhibit the characteristics of a white noise.

Of course, in practice we do not have access to $\mathbf{v}_1$ (or $\mathring{\mathbf{v}}_1$) nor to $\mathbf{v}_2$, but to their noisy, aliased versions $\tilde{\mathbf{v}}_1$ and $\tilde{\mathbf{v}}_2$. Thus k must be estimated in the same way, with the sequences that are available.

$$\begin{aligned}\mathbf{H}_{\underset{s}{\lambda}}\tilde{\mathbf{v}}_1 &\overset{\text{def}}{=} \mathbf{S}_1 W_1 H_{\underset{s}{\lambda}} \mathbf{I}_1 \mathbf{v}_1 + \mathbf{n} \\ &= \mathbf{S}_1 W_1 H_{\underset{s}{\lambda}} v_1 + \mathbf{S}_1 W_1 H_{\underset{s}{\lambda}} (\mathbf{I}_1 \mathbf{v}_1 - v_1) + \mathbf{n} \\ &\overset{\text{def}}{=} \mathring{\mathbf{v}}_1 + \mathbf{r} + \mathbf{n},\end{aligned}$$

where $\mathbf{n} = \mathbf{H}_{\underset{s}{\lambda}}\mathbf{n}_1$, and $\mathbf{r} = \mathbf{S}_1 W_1 H_{\underset{s}{\lambda}}(\mathbf{I}_1\mathbf{v}_1 - v_1)$ is the aliasing term introduced when sampling v_1.

Neglecting the aliasing It was assumed that the aliasing term $\mathbf{r}$ could be neglected. Let us explore under which conditions this is a reasonable assumption. Since

$$\begin{aligned}\mathbf{r} &= \mathbf{S}_1 W_1 H_{\underset{s}{\lambda}} (\mathbf{I}_1 \mathbf{v}_1 - v_1) \\ &\overset{(15)}{=} \mathbf{S}_1 H_{\underset{s}{\lambda}} W_{\frac{s}{\lambda}} (\mathbf{I}_1 \mathbf{v}_1 - v_1),\end{aligned}$$

the estimation will only be affected by aliasing if there are aliasing components in the frequency interval $\left[-\frac{s}{\lambda}\pi, \frac{s}{\lambda}\pi\right]^2$. This allows us to choose $v_1 = H_{\lambda_1} u$ such that $\mathrm{supp}(\hat{v}_1) \subset \left[-2\pi + \frac{s}{\lambda}\pi, 2\pi - \frac{s}{\lambda}\pi\right]^2$ (see Figure 1). Last, to minimize the impact of the aliasing term the images should be acquired from a pair of fronto-parallel locations as far as possible one from the other, since that amounts to increase the value of λ.

Finally, one simple way of estimating k is by considering the least squares estimator

$$\mathbf{k}_e := \left(\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\lambda}}\tilde{\mathbf{v}}_1]\right)^+ \tilde{\mathbf{v}}_2.$$

This estimator will be accurate as long as the noise is much smaller than the signal power and the aliasing term $\mathbf{r} = \mathbf{S}_1 W_1 H_{\underset{s}{\lambda}}(\mathbf{I}_1\mathbf{v}_1 - v_1)$. More details of the estimation stability are presented in Section 4.

3.2 From relative to absolute blur

Notice that h appears on both sides of Equation (4). However, it can be obtained from k by calculating the following limit

$$h = \lim_{n\to\infty} H_{\lambda^{n-1}} k * \ldots * H_\lambda k * k.$$

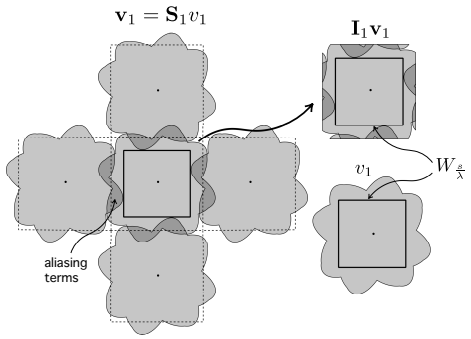


Figure 1: *Neglecting the aliasing* The estimation will only be affected by aliasing if there are aliasing components in the interval $\left[-\frac{s}{\lambda}\pi, \frac{s}{\lambda}\pi\right]^2$. Hence, to avoid aliasing one can choose $v_1 = H_{\lambda_1} u$ such that $\text{supp}(\hat{v_1}) \subset \left[-2\pi + \frac{s}{\lambda}\pi, 2\pi - \frac{s}{\lambda}\pi\right]^2$.

The proof is given in Appendix (Proposition 5). This proves that one can recover the camera PSF at $s\times$ resolution $\mathbf{h}$ from the inter-image kernel $k = H_s \mathbf{I}_1 \mathbf{k}$.

4 The complete PSF Estimation Procedure

This section describes the algorithmic steps that lead to local subpixel PSF estimates. The complete chain is summarized in the block diagram of Fig. 2. The next paragraphs present brief summaries for each block.

Image Alignment In order to estimate the geometric transformation between both images, they need to be precisely aligned. For that purpose SIFT descriptors [21] which are scale and contrast-change invariant are obviously the right choice. The IPOL implementation of ASIFT [24, 25] was chosen because of the efficiency of the Optimized Random Sampling Algorithm (ORSA) rejection of false matches.

Geometric Transform Estimation The complete geometric transformation from one image to the other was approximated with thin-plate splines [5, 31] from the matched SFT pairs. Modeling the geometric transformation in this way, permits to correct small non-affine distortions, especially in the zoomed-in view and little deviations from the parallel assumption when interpolating $\tilde{\mathbf{v}}_2$ to generate $\mathbf{H}_{\frac{\lambda}{s}} \tilde{\mathbf{v}}_2$. Of course, if the distortion is significant the assumed inter-image kernel Eq. (4) will not be accurate. The thin-plate representation as affine + non-affine parts of the geometric distortion is specially handful to estimate the relative scale $\lambda = (\lambda_x, \lambda_y)$ between both views, since this can be

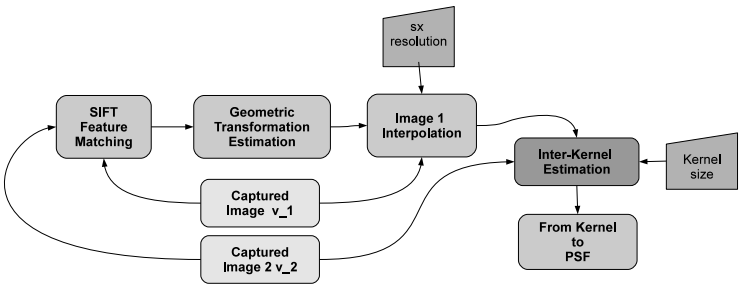


Figure 2: *Algorithm Description.* Both captured images are aligned via SIFT feature matching followed by the estimation of a smooth geometric distortion through thin-plate approximation of matched features. The relative geometric distortion and gray-level corrections are applied to a low-pass (unalised) version of the finest scale image 1. Then the interpolated image 1 and image 2 are compared to obtain a relative blurring kernel h_0 which is later iteratively updated to obtain the absolute camera PSF.

directly estimated from the affine part.

Gray level adjustment Both snapshots should be taken with exactly the same camera configuration. This implies that (provided that the scene illumination remained unchanged) it should not be necessary to compensate the gray levels between the two scenes.

Resampling and Distortion Correction of $\mathbf{v}_1$ To generate the rescaled samples $\mathbf{H}_{\frac{\lambda}{s}}\tilde{\mathbf{v}}_1$ from the digital image $\tilde{\mathbf{v}}_1$ requires its interpolation at the desired scale $\frac{\lambda}{s}$. This is done by using the estimated geometric transformation. Notice that since $\tilde{\mathbf{v}}_1$ is not very aliased, one can correctly interpolate it without introducing artifacts and thus correct small non-homothetic distortions between both views without affecting the kernel estimation.

Numerical Methods for Inter-Image Kernel Estimation Suppose that the image $\tilde{\mathbf{v}}_2$ has size $m \times n$. The goal is to estimate k at $s\times$ the resolution of $\tilde{\mathbf{v}}_2$ (camera sensor resolution). Also suppose that the estimated support of the inter-image kernel k is inside a $r \times r$ image. The matrix $\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\frac{\lambda}{s}}\tilde{\mathbf{v}}_1]$ corresponding to the s-downsampling operator and the convolution with the interpolated $\mathbf{H}_{\frac{\lambda}{s}}\tilde{\mathbf{v}}_1$ image, has size $mn \times r^2$. Then, the problem to be solved is

formally written as

$$(P) \quad \arg\min_{\mathbf{k}} \left\| \mathcal{S}_s \mathbf{C}[\mathbf{H}_{\frac{\lambda}{s}} \tilde{\mathbf{v}}_1]\mathbf{k} - \tilde{\mathbf{v}}_2 \right\|^2 .$$

Problem (P) can be solved by a simple least-squares algorithm. If the noise and the aliasing term are small then the solution to (P) will be close to the solution of $\mathcal{S}_s \mathbf{C}[\mathring{\mathbf{v}}_1]^+ \mathbf{v}_2$. This is a direct consequence of applying Lemma 6 with matrix $\mathbf{A} = \mathcal{S}_s \mathbf{C}[\mathring{\mathbf{v}}_1]$ and perturbated matrix $\mathbf{A} + \Delta\mathbf{A} = \mathcal{S}_s \mathbf{C}[\mathbf{H}_{\frac{\lambda}{s}} \tilde{\mathbf{v}}_1]$, and the respective vector $\mathbf{b} = \mathbf{v}_2$ and $\mathbf{b} + \delta\mathbf{b} = \tilde{\mathbf{v}}_2$.

Transforming the kernel: from k to h Recovering the camera PSF h requires to obtain the limit

$$h = \lim_{n \to \infty} H_\lambda^{n-1} k * \ldots * H_\lambda k * k.$$

Since $\mathbf{k}$ samples of k have been estimated at a small enough rate to be well sampled according to Shannon theory, one can interpolate them to generate k. Ideally this would require using a Shannon-Nyquist interpolator but a smoother one, the cubic spline, was adopted, which is reasonable, k proving *a posteriori* to be very smooth.

Directly working with the digital sequences requires some care about how the successive convolutions are computed. For example, since k is bandlimited $H_\lambda k$ should be lowpass filtered before applying the homothecy operator in order to avoid aliasing artifacts when convolving with k (consequence of Lemma 17 and the fact that $\lambda > 1$). To avoid this inconvenient one can re-state the limit convolution as follows:

$$h = \lim_{n \to \infty} H_{\lambda^n} (k * H_{\frac{1}{\lambda}} k \ldots * H_{\frac{1}{\lambda^n}} k).$$

If implemented in this way, the successive convolutions can be computed without any special care. Thus one can proceed as follows:

1. Initialize $\mathbf{u}_0 = \mathbf{k}$, $n = 1$
2. Compute $\mathbf{H}_{1/\lambda^n}\mathbf{k}$ by using $\lambda = (\lambda_x, \lambda_y)$ (from thin-plates affine part).
3. Calculate $\mathbf{u}_n = \mathbf{H}_{1/\lambda^n}\mathbf{k} * \mathbf{u}_{n-1}$.
4. If $\min\{\lambda_x^n, \lambda_y^n\} > \lambda_{\max}$ go to 5. Else update $n := n+1$ and repeat from 2.
5. Calculate $\mathbf{h} = \mathbf{H}_\lambda^n \mathbf{u}_n$.

The algorithm converges after a few iterations since λ^n grows very fast. Setting $\lambda_{\max} = 50$, since the convolution with the inter-image kernel zoomed-out $50\times$ or greater produces a negligible change in the final result.

Finally, since negative light does not exist the estimated PSF should be positive. We can therefore constrain the solution to be non-negative projecting the result to the non-negative hyperplane.

5 Experimental Results

Since there is no PSF ground-truth available, the validation of the proposed method was carried out by simulations and also by comparing the results with state-of-the-art non-blind methods [18, 17, 13, 20]. We opted to compare only to non-blind algorithms, the accuracy of blind methods being significantly lower. The behavior of the proposed approach was tested for several different image pairs and for super-resolution rates ranging from $1\times$ to $4\times$. The experiments were performed using a a Canon EOS 400D camera equipped with a Tamron AF 17-50mm F/2.8 XR Di-II lens. The focal length was fixed to 50.0 mm. A complete algorithmic description, an online demo facility and a source code can be found at the IPOL workshop by [14].

Simulation as a Sanity Check A synthetic random image u was generated and re-interpolated $4\times$ in order to get the "continuous" sharp homothecy of the image u. Next both images were convolved with a PSF like kernel (in this case a Gaussian isotropic kernel), and down-sampled to get the respective observed digital images at the camera resolution (i.e. $1\times$). The kernel was chosen so that the low resolution image presents aliasing artifacts. By generating the views of u in this way, there are no aliasing artifacts in the camera resolution zoomed-in image. This experiment was done as a sanity check of the proposed method. A $4\times$ kernel was estimated from the observed image pairs. The results are shown in Figure 3 and 4. The camera kernels estimation used the automatic registrations from the detected SIFT points / thin-plates and the ideal known alignment. Both estimations are accurate. Nevertheless, as shown in the difference images the automatic alignment introduces a small mis-alignment. Since the two images were simulated at distances in a $4\times$ ratio, the difference between the inter-image kernel and the PSF are very small due to the large λ value (shown in figure 3 bottom row, on the right).

Blind vs. Non-blind Figure 6 shows the $4\times$ PSF estimated by the proposed blind method from a pair of views of a wall shown in Figure 5. The estimation was conducted for one of the green channels (half of the green pixels of the Bayer matrix), with the camera aperture set to f/5.7. The estimated PSF is quite close to the one obtained with by the non-blind algorithm described in [13, 12]. In particular their sizes are similar and their corresponding MTFs present zeros at the same locations.

Bayer channels estimations Two pictures of a another textured wall shown in Figure 7 were used to estimate the PSF of the four color Bayer channels (RAW camera output). This wall texture presents characteristics similar to white noise. The results for the $4\times$ PSF estimated at the image center are shown in Fig. 7. Notice that the red channel PSF is wider than the green and the blue one, as

expected from the physics of diffraction-limited optical systems, since the wavelengths associated to red light are larger than the rest. The differences between the dominant orientations of the red/blue and green PSF spectra can be explained by the sensor shape. In fact, the sensor active zone is usually L-shaped (see for example [33]). Thus the red/blue sensors in the Bayer pattern should have, as observed, the same sensor term MTF and will be rotated 90° with respect to the green channels.

Different kinds of scenes The images of a wall shown in the previous experiments present good characteristics to use our blind PSF estimation algorithm. Despite having white noise appearance, they present enough structure to give sufficiently accurate SIFT points. The counterpart is that the spectrum will not be flat, thus the estimation at high frequencies may be affected. The SIFT descriptors are not robust in high aliasing situations. Hence, there is a trade-off between having accurate SIFT matches and textures with high frequency information. The texture shown in Figure 7 is an example of a good compromise for this trade-off.

Figure 8 shows the 1× to 4× PSF estimations for the first green channel from an image pair of a photograph from a magazine. The estimation was done at the image center for the camera working at f/5.7 aperture. All the subpixel estimations are consistent. As can be seen in this figure, they have overlapping MTFs in the common regions. These newspaper images produce accurate SIFT points. However the 4× PSF estimation is noticeably noisier than the one produced from the wall images. The main reason is that the spectrum of this magazine image, made of edges, has a faster decay. Despite of this, the quality of the 4× PSF estimation is still reasonable.

What kind of textures should be used? It follows from the previous analysis that textures composed by elements with different sizes, are preferable, to produce simultaneously good SIFT points and a sufficiently slow frequency decay. 3D textures like those shown in Figure 9 can be problematic for this approach. Even though they respect the two previous conditions, their 3D nature produces disparities and occlusions which change the image beyond a simple zoom. Likewise, non-Lambertian surfaces and dynamical scenes are not appropriated.

Comparison to non-blind methods Figure 10 shows a comparison of several non-blind methods with the blind method proposed here, obtained from the pair of wall images shown in Fig. 7. Since the non-blind methods achieve a much higher accuracy than blind methods, the choice was to compare the blind results obtained by the proposed method against those of state of the art non-blind algorithms. Among the non-blind methods is the recent one introduced in [13], which outperforms all methods shown in Figure 10. For this reason,

the MTF estimated from this later non blind method will serve as ground-truth. The Imatest commercial software for MTF estimation [20] is one of the evaluated methods. The Imatest estimation is performed from a slanted-edge image and only gives an estimate of the MTF at the direction orthogonal to the slanted-edge.

The Figure clearly shows that the proposed blind method performs at least as well as the other two non-blind methods under comparison. The Joshi et. al. non-blind algorithm [18, 17] shows similar performance for a carefully, manually chosen regularization parameter. Indeed, this slanted-edge algorithm depends on a regularization parameter which must be carefully chosen.

6 Discussion

This work has presented a blind subpixel accurate algorithm for the estimation of the point spread function of a digital camera from aliased photographs. The procedure is based on taking two fronto-parallel photographs of the same flat textured scene, from different distances leading to different geometric zooms, and then estimating the kernel blur between them.

The estimation method is regularization-free. In that sense, the technique is closely related to the recent non-blind estimation method [13] using a random noise pattern. This later paper shows that with such patterns the estimation problem is well posed and leads to accurate regularization-free estimates. The main difference is that non blind methods can estimate directly the PSF using the perfect knowledge of the pattern. In the blind methods the question is far more intricate because only an inter-image kernel can be recovered. Thus a mathematical analysis and new algorithms have been introduced proving how the PSF can be recovered from the inter-image kernel.

To reach a high accuracy, images of textured scenes with flat enough spectrum are preferable. It was experimentally verified that many textures found in nature are well adapted to these requirements. A comparison of the resulting PSF estimates with other subpixel PSF estimation methods shows that the proposed algorithm reaches similar accuracy levels to non-blind state of the art methods, with the advantage of not requiring any special acquisition setup, thus being much more practical.

Acknowledgements

The authors would like to thank Saïd Ladjal for fruitful comments and discussions. Research partially funded by the Centre National d'Etudes Spatiales (R&T), the European Research Council advanced grant Twelve Labours, the Office of Naval Research (grant N00014-97-1-0839), STIC-AmSud (11STIC-01 - MMVPSCV), and the Uruguayan Agency for Research and Innovation (ANII) under grant PR-POS-2008-003.

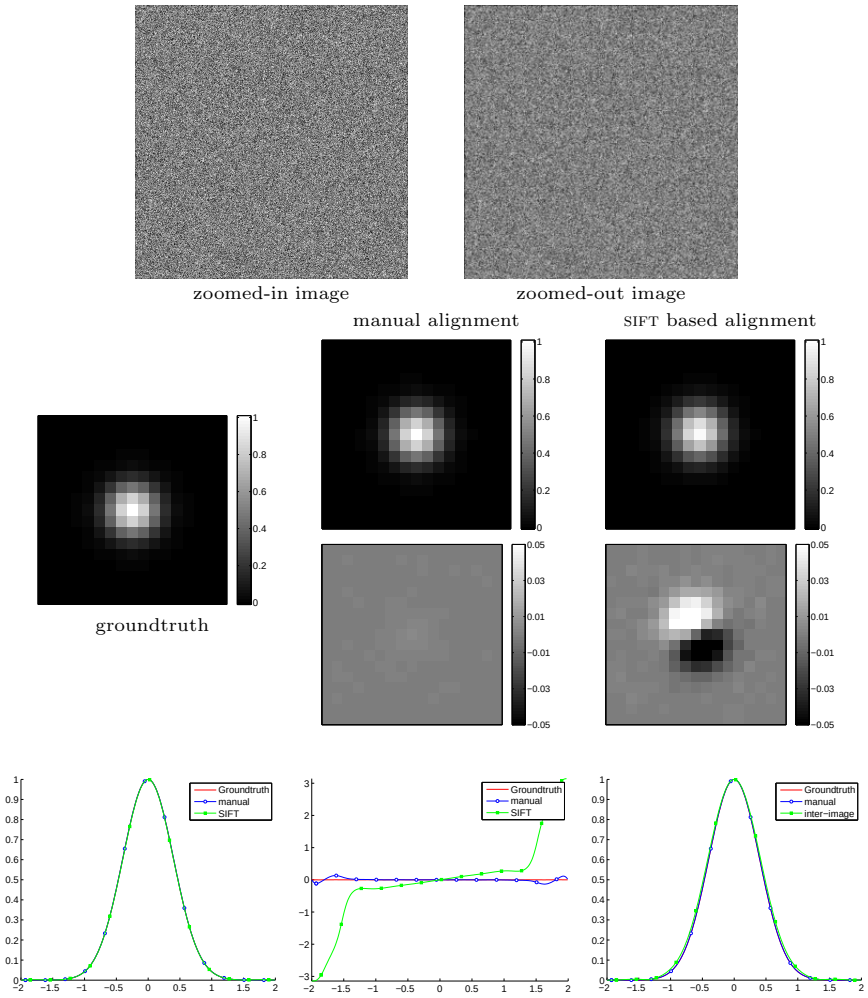
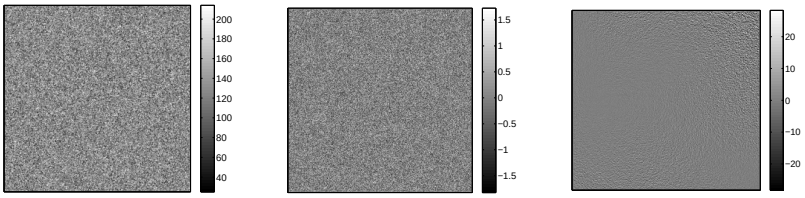


Figure 3: Synthetic example. 4× PSF estimation for simulated data. Top row: the two views zoomed-in / zoomed-out images. Middle row: the simulated PSF (ground truth) and the respective PSF estimations using the automatic SIFT points / thin-plates alignment and the ideal alignment. Both estimations are accurate. However, as shown in the difference images the automatic alignment introduces a small mis-alignment. This can also be seen in the phase and modulus of the PSF Fourier Transform vertical profile which is shown in the bottom row. Since the two images are simulated at distances ranging from 1 to 4, the difference between the inter-image kernel and the PSF are very small due to the large λ value (bottom row, on the right).



zoomed-out image difference manual difference SIFT

Figure 4: Synthetic example. 4× PSF estimation for simulated data. From left to right: Top row: the zoomed-out image, and the residual image $\mathcal{S}_s(\mathbf{H}_{\frac{s}{\lambda}}\tilde{\mathbf{v}}_1 * \mathbf{k}) - \tilde{\mathbf{v}}_2$ with the estimated kernel from the ideal alignment and the automatic alignment through the SIFT points. The residual in the automatic alignment case is significantly larger than in the ideal alignment case. However, the difference in the PSF seems to be negligible up to a subpixel translation as is shown in Figure 3.

zoomed-out photograph zoomed-in photograph

Figure 5: Wall image: blind vs. non-blind estimation. Two distant, parallel views of a textured wall.

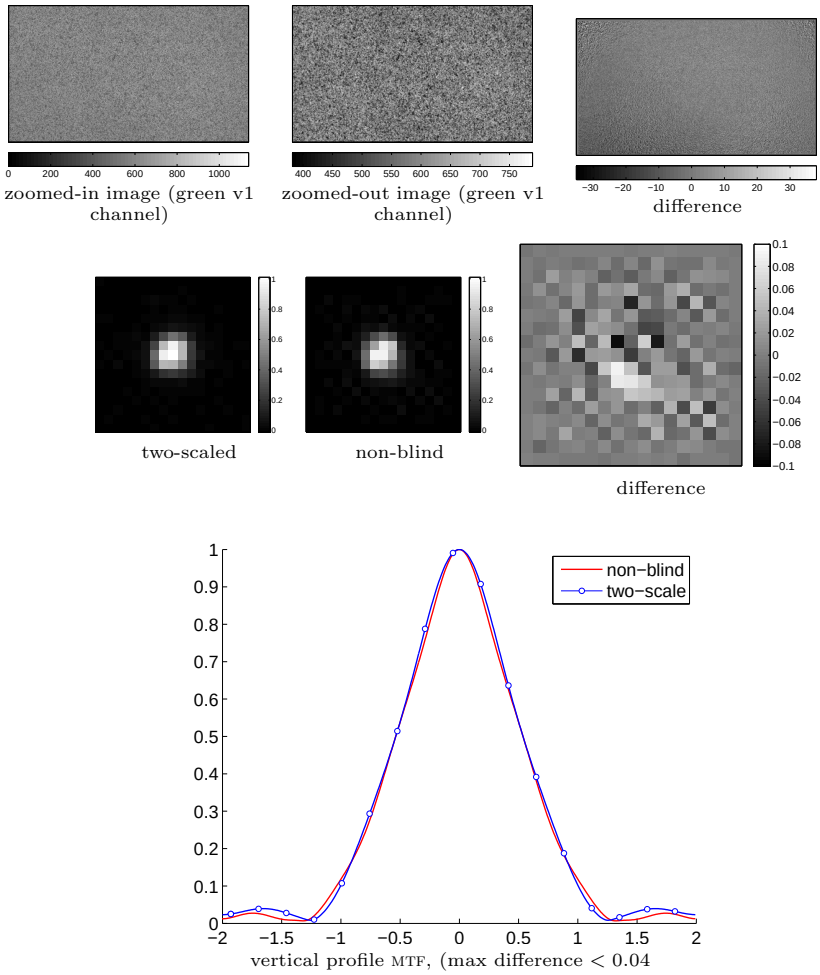


Figure 6: Wall image: blind vs. non-blind estimation. 4× PSF estimation for one of the green channels from a camera RAW output. Top row: two distant, parallel views of a textured wall. Middle row: the PSF estimated with the proposed algorithm and the one estimated using the previous non-blind algorithm. Bottom row: MTF vertical profile. Both estimations seem to be quite close. In particular the associated airy disks have similar sizes and the MTFs vanish in the same locations.

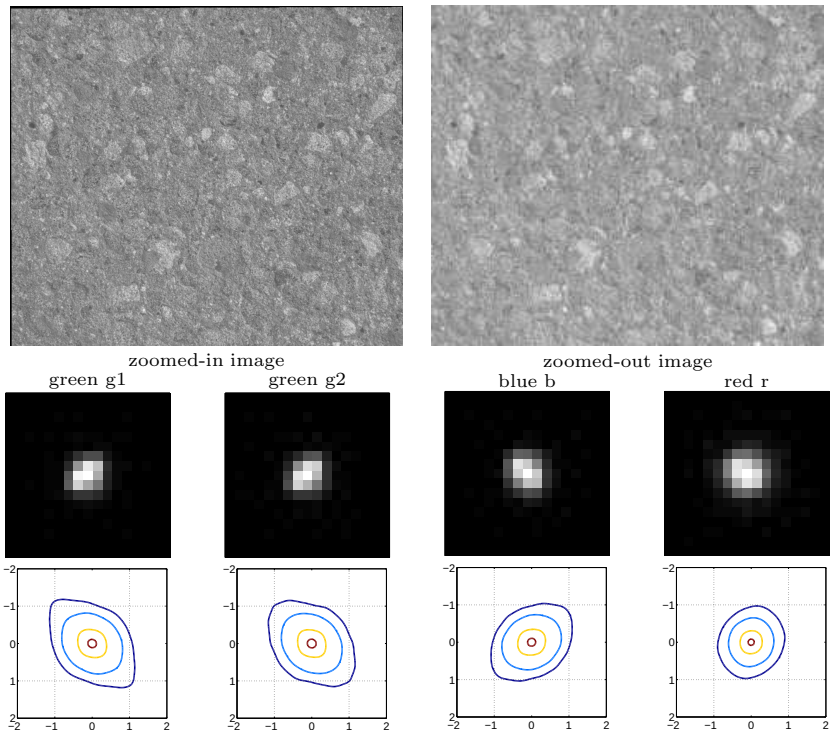


Figure 7: *Different color channels.* 4× PSF estimation for the four Bayer pattern channels (two greens, red and blue) from a camera RAW output. Top row: two distant, parallel views of a concrete wall. Middle row: the 4× PSF estimated for the four channels and their corresponding Fourier spectrum modulus. Bottom row: MTF horizontal profile. The estimation was performed with images captured at aperture f/5.7. The red PSF is larger than the green and the blue ones. Since the wavelengths associated to red are larger than the rest, the diffraction components for the red channel will be larger than those for green and blue for the same camera configuration. Also notice the differences between the shape of the red/blue and green PSF spectra (bottom row). Red and blue MTF seem to be 90° rotated with respect to the green ones. This symmetrical behavior is plausible for an L-shaped active zone sensor array ([33]).

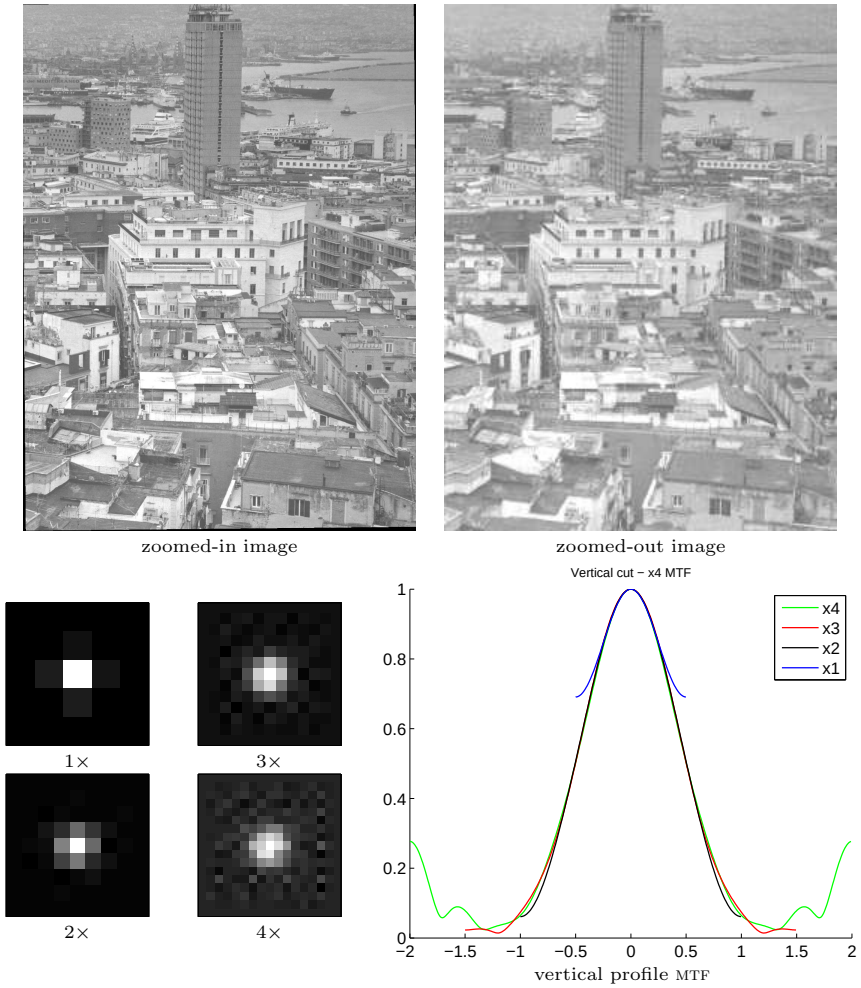


Figure 8: Magazine image : ×1, ×2, ×3 and ×4 estimations for the first green channel from an image pair of a newspaper. The estimation was done at the image center for the camera working at f/5.7 aperture. All the estimations are consistent. As we can be observed in this figure, they have overlapping MTFs in the common regions. The 4× PSF estimation is noisier than the one produced from the wall images. The main reason is that the faster spectral decay of this magazine image. However, the quality of the 3× PSF estimation is still acceptable.

Figure 9: Examples of problematic textures. Their 3D nature produces disparities and little changes in the angle-of-view would result in accuracy loss. Non-Lambertian surfaces and dynamical scenes are not appropriated either.

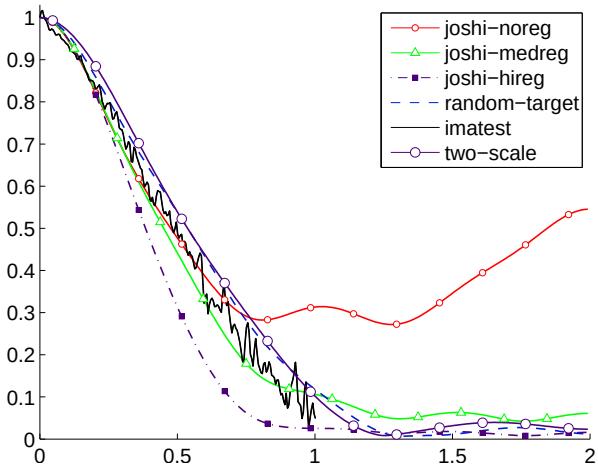


Figure 10: *Comparison of* PSF/MTF *estimation methods* Our implementation of Joshi et al. PSF estimation algorithm [18, 17], Imatest commercial software [20], the previously proposed non-blind random pattern based algorithm [13] and the blind two-scaled proposed in this work (applied to the images of the wall shown in Fig. 7). All estimations are done at the center of the image with a camera at aperture f/5.7 for one of the green channels. On the low frequencies all algorithms gave very similar estimations, while on the higher frequencies the Joshi et al. estimation depends strongly on the regularization level. Although much effort was made to get a noise free MTF estimation from the Imatest software, the final estimation is quite noisy. The Imatest estimation is done from a slanted-edge image and only gives an estimation for the MTF at the slanted-edge orthogonal direction. The proposed blind algorithm is the one presenting closest estimation to the non-blind estimation from [13], considered as ground-truth in virtue of its high accuracy.

A Mathematical framework and physical modeling

Functional spaces :

- $\mathbb{R}^2$ is the set of pairs of real numbers $\mathbf{x} = (x_1, x_2)$ and $\mathbb{Z}^2$ the set of pairs of integers $\mathbf{k} = (k_1, k_2)$. $L^1(\mathbb{R}^2)$ is the set of integrable functions on $\mathbb{R}^2$, $L^2(\mathbb{R}^2)$ the set of quadratically integrable functions, $C_b^0(\mathbb{R}^2)$ the set of continuous bounded functions, $C^\infty(\mathbb{R}^2)$ the set of infinitely differentiable functions, $\mathcal{S}(\mathbb{R}^2)$ the Schwartz class of C^∞ functions whose derivatives of all orders have fast decay, $\mathcal{S}'(\mathbb{R}^2)$ its dual, the space of tempered distributions, $\mathcal{E}'$ the subset of $\mathcal{S}'(\mathbb{R}^2)$ of compactly supported distributions. We shall use the properties of the convolution $L^1 * L^2 \subset L^2$, $L^1 * L^1 \subset L^1$, $L^2 * L^2 \subset C^0$, $\mathcal{E}' * \mathcal{S}' \subset \mathcal{S}'$.

- A non classical but convenient notation in Shannon sampling theory will be to denote by $\mathcal{BL}^2(\mathbb{R}^2)$ (or shortly $\mathcal{BL}^2$) the set of L^2 functions that are band limited inside $[-\pi, \pi]^2$.

- More generally, $\mathcal{BL}_0^2$ denotes the space of L^2 functions with compactly supported Fourier transform. This function space will play a particular role in what follows.

The following conventions and notations will be used in the sequel:

- $\mathbf{F}$ is the *Fourier Transform operator* defined on $\mathcal{S}'$; $\mathbf{F}(f)(\zeta) = \hat{f}(\zeta) = \int e^{-i\mathbf{x}\cdot\zeta} f(\mathbf{x}) d\mathbf{x}$ defines it for a function $f \in L^1(\mathbb{R}^2)$ in a point $\zeta = (\zeta_1, \zeta_2)$. This formula is still valid for functions belonging to $L^p(\mathbb{R}^2)$ with $1 < p \leq 2$ (see e.g. [32, 4]).

- *Continuous images* are defined for $\mathbf{x} \in \mathbb{R}^2$, whereas *digital images* are sampled on a discrete grid $\mathbf{k} \in \mathbb{Z}^2$.

- $\mathbf{S}_1 : C_b^0 \to \ell^\infty(\mathbb{Z}^2)$ is the 1-*sampling operator* such that $\mathbf{u}(\mathbf{k}) = (\mathbf{S}_1 u)(\mathbf{k})$. From the distribution viewpoint $\mathbf{S}_1$ is the product by a Dirac comb $\Pi_s := \sum_{\mathbf{k}} \delta_{s\mathbf{k}}$ with $s = 1$, namely $\mathbf{S}_1 u = \Pi_1 . u$ where u must be a continuous function. Both representations of the sampling operator will be identified, and it will be clear from the context, which one of both representations is intended. [3]

- A *digital image* $\mathbf{u}$ will be represented either as a sequence $(u(\mathbf{k}))_{\mathbf{k}}$ in $\ell^\infty(\mathbb{Z}^2)$ or as the corresponding Dirac comb $\mathbf{u} := \sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{u}(\mathbf{k}) \delta_{\mathbf{k}}$.

[3] Sampling at other rates $s \neq 1$ will be introduced later on, in terms of a scaling operator H_s which comes with the corresponding multiplicative constant. This helps to avoid cumbersome normalization factors.

- *Continuous images* will be assumed to be functions in $\mathcal{BL}^2_0(\mathbb{R}^2)$. This choice is consistent, since these functions are continuous (actually C^∞) and the sampling is well defined. Moreover, as suggested in [23] and later in [1, Appendix A] this choice is sufficiently general to model the continuous landscape observed by a camera just before sampling takes place at the sensors.
  In fact, even if the raw physical image before blur and sampling is, realistically, a positive Radon measure O (due to the photon-counting nature of sensitive digital systems) with compact support (imposed by the finite number of photons), it will still be blurred by a kernel h which will be regular enough for $h * O$ to be in $\mathcal{BL}^2_0$.
  How regular can it be realistically assumed to be? The kernel h originates in several physical phenomena from diffraction, through out-of-focus-blur to sensor integration. Each one of these phenomena, and their combination as well, lead to model h as a nonnegative function with finite mass $\int h = 1$ (normalized to 1). In addition the diffraction part ensures that $\hat{h}$ is compactly supported. From this one deduces that $h \in \mathcal{BL}^2_0 \cap L^1$.
  We now turn to the problem of simplifying O to a more manageable function u_0, which is indistinguishable from O after convolution with the PSF h.
  Let $B = \text{supp}\hat{h}$ be the (compact) spectral support of the PSF h. Hence h can be idempotently written as $h = h * h_0$, where $h_0 \in \mathcal{S}'$ has a compactly supported spectrum satisfying $\hat{h_0}(\eta) = 1$ for $\eta \in B$. The function $\hat{h}$ can easily constructed by an explicit formula as a C^∞ and compactly supported function satisfying $\hat{h}(\eta) = 1 on [-\pi, \pi]^2$. Then its inverse Fourier transform has all required properties.
  So we have
  $$u = h * O = h * u_0, \quad \text{where } u_0 = h_0 * O.$$
  In consequence, the observed landscape can be assumed w.l.o.g. to be $u_0 = h_0 * O$ instead of O. Being the convolution of a compactly supported positive Radon measure $O \in \mathcal{E}'$ with $h_0 \in \mathcal{BL}^2_0 \cap L^1$, u_0 also belongs to $\mathcal{BL}^2_0$, and its convolution with $h \in \mathcal{BL}^2_0 \cap L^1$ is the observed image $u \in \mathcal{BL}^2_0$.

- The operator $\mathbf{I}_1 : \ell^2(\mathbb{Z}^2) \to \mathcal{BL}^2(\mathbb{R}^2)$ denotes the *Shannon-Whittaker interpolator*, defined by $\mathbf{I}_1\mathbf{u}(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{u}(\mathbf{k})\text{sinc}(\mathbf{x} - \mathbf{k})$, where $\text{sinc}(\mathbf{x}) = \frac{\sin(\pi x)}{\pi x}\frac{\sin(\pi y)}{\pi y}$. We therefore have $\mathbf{I}_1\mathbf{u} = \mathbf{F}^{-1}(\sum_{\mathbf{k}} \mathbf{u}(\mathbf{k})e^{-i\mathbf{k}\cdot\xi}\mathbf{1}_{[-\pi,\pi]^2})$. When $\mathbf{u} \in \ell^2$, $\mathbf{F}(\mathbf{I}_1\mathbf{u})$ therefore belongs to L^2 and is compactly supported. Thus $\mathbf{I}_1\mathbf{u} \in \mathcal{BL}^2$ and we have $\mathbf{S}_1\mathbf{I}_1 = Id$.

- The filter $W_w u = \mathbf{F}^{-1}(\hat{u} \cdot \mathbf{1}_{[-w\pi,w\pi]^2})$ is an *ideal low-pass filter* that cuts the spectrum of u to $[-w\pi, w\pi]^2$. It is defined if $\hat{u}$ is a function.

- $H_\lambda u(\mathbf{x}) = \lambda^2 u(\lambda\mathbf{x})$ is the continuous homothety (*i.e.* $\lambda > 1$ is a zoom-out); the rationale for its normalization is to preserve the image mean

(its zero-frequency coefficient). In Fourier $\mathbf{F}(H_{\frac{\lambda}{u}})(\zeta) = \hat{u}(\frac{\zeta}{\lambda})$, so if u is α-band-limited then $H_{\frac{1}{\alpha}}u$ is band-limited.

- $\mathcal{S}_s : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$ denotes the s-to-1-resampling operator $\mathcal{S}_s = \mathbf{S}_1 H_s \mathbf{I}_1$ (*i.e.* $s > 1$ is a subsampling by s).

- $\mathbf{C}[\mathbf{u}] : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$ denotes the linear map associated to the convolution with a digital image $\mathbf{u}$. The convolved sequence belongs to $\ell^2(\mathbb{Z}^2)$ which in general is satisfied if $\mathbf{u} \in \ell^1(\mathbb{Z}^2)$.

- The digital Nyquist homothety operator $\mathbf{H}_\alpha : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$ is defined by $\mathbf{H}_\alpha \mathbf{u} := \mathbf{S}_1 W_1 H_\alpha \mathbf{I}_1 \mathbf{u}$. It is a digital zoom out if $\alpha > 1$.

- Let $\mathbf{L}$ be a bounded linear operator over a Hilbert space. $\mathbf{L}^*$ is its adjoint and $\mathbf{L}^+$ (if it exists) its pseudo-inverse, *i.e.* the minimum-norm solution of $(\mathbf{L}^*\mathbf{L})\mathbf{L}^+ := \mathbf{L}^*$.

B Standard results from Fourier analysis

The following two main results from standard Fourier analysis and distribution theory are stated without proof. The reader is referred to *e.g.* [32, 15] for the proofs in the particular setting chosen here.

Proposition 2 (Convolution through Fourier transform)**.** *The relation*

$$\mathbf{F}(f * g) = \mathbf{F}(f) \cdot \mathbf{F}(g) \tag{8}$$

is valid in any of these cases

1. *$g \in L^1(\mathbb{R}^2)$ and $f \in L^p(\mathbb{R}^2)$ for $1 \leq p \leq 2$. Then $f * g$ belongs to $L^p(\mathbb{R}^2)$ (see [32, Theorem 2.6]).*

2. *$g \in \mathcal{E}'$ and $f \in \mathcal{S}'$. Then $f * g$ belongs to $\mathcal{S}'$ (see [15, Theorem 7.1.15]).*

Applying the Fourier transform on both sides of Equation (8) and recalling that the squared Fourier Transform operator $\mathbf{F}^2(u) = (2\pi)^2[\mathbf{x} \mapsto u(-\mathbf{x})]$ is almost the identity (except for flipping and a constant factor), we obtain the following:

Corollary 1 (Product through Fourier transform)**.** *The relation*

$$\begin{aligned} \mathbf{F}(f \cdot g) &= \frac{1}{(2\pi)^2}\mathbf{F}(f) * \mathbf{F}(g) \\ \mathbf{F}^{-1}(f \cdot g) &= \mathbf{F}^{-1}(f) * \mathbf{F}^{-1}(g) \end{aligned} \tag{9}$$

holds when $\hat{g} \in \mathcal{E}'$ and $f \in \mathcal{S}'$. Then $f \cdot g$ belongs to $\mathcal{S}'$.

Proposition 3 (Poisson Formula in $\mathbb{R}^2$ for tempered distributions [15])**.**

$$\hat{\Pi}_1 = (2\pi)^2 \Pi_{2\pi}. \tag{10}$$

The Shannon-Whittaker sampling theorem is then a direct consequence of the two previous results:

Lemma 2. *If $\hat{u} \in \mathcal{E}'$, then*

$$\mathbf{F}(\Pi_1 \cdot u) = \Pi_{2\pi} * \hat{u}. \tag{11}$$

Proof. We can apply the first form of Corollary 1 where $f = \Pi_1 \in \mathcal{S}'$ and $\hat{g} = \hat{u} \in \mathcal{E}'$ to obtain

$$\mathbf{F}(\Pi_1 \cdot u) = (2\pi)^{-2} \hat{\Pi}_1 * \hat{u} = \Pi_{2\pi} * \hat{u}$$

where the last equality is deduced from the Poisson formula (10). □

Proposition 4 (Nyquist-Shannon Theorem)**.** *If $u \in \mathcal{BL}^2(\mathbb{R}^2)$, then*

$$u = \mathbf{I}_1 \mathbf{S}_1 u. \tag{12}$$

Proof. We can apply Lemma 2

Multiplying both sides of Equation (11) by $\mathbf{F}(\text{sinc}) = \mathbf{1}_{[B]}$ we obtain

$$\begin{aligned} \mathbf{F}(\text{sinc}) \cdot \mathbf{F}[\mathbf{S}_1 u] &= \mathbf{F}(\text{sinc}) \cdot [\Pi_{2\pi} * \hat{u}] \\ &= \sum_{\mathbf{k} \in \mathbb{Z}^2} \hat{u}(\cdot + 2\pi\mathbf{k}) \mathbf{1}_{[B]} \\ &= \hat{u} \end{aligned}$$

where in the right-hand side the only non-null term is $\mathbf{k} = 0$ because u is bandlimited in $B = [-\pi, \pi]^2$ and $\mathbf{F}(\text{sinc}) = \mathbf{1}_{[B]}$. Finally, using the second form of Corollary 1 we obtain

$$\text{sinc} * (\mathbf{S}_1 u) = u$$

and the left term is by definition $\mathbf{I}_1 \mathbf{S}_1 u$. □

Corollary 2. *If $u \in L^2$ is s-band-limited then*

$$u = H_s \mathbf{I}_1 \mathbf{S}_1 H_{\frac{1}{s}} u \tag{13}$$

C Proof of Main Results of Sections 2 and 3

Common hypotheses

According to the discussion in Appendix A, and in order to justify all the Lemmas and Propositions we will require that

- $h \in \mathcal{BL}^2_0 \cap L^1(\mathbb{R}^2)$, non-negative, $\hat{h}(0) = 1$.
- $u_0 \in \mathcal{BL}^2_0$

This ensures that the convolution $u_0 * h = u$ is well defined with $u \in \mathcal{BL}^2_0$.

For the uniqueness of the inter-image kernel we shall additionally assume that $\hat{h}$ does not vanish inside $[-\frac{s}{\lambda_2}\pi, \frac{s}{\lambda_2}\pi]$.

Main Results

We now prove several properties that are used throughout the article.

Lemma 3. *If $u \in \mathcal{BL}_0^2$, then $\mathbf{S}_1 u \in \ell^2(\mathbb{Z}^2)$.*

Proof. As u is in $\mathcal{BL}_0^2$ there exists $s > 0$ such that $\hat{u} \subset [-s\pi, s\pi]$. Furthermore, since $\hat{u} \in \mathcal{E}'$ applying (11) we have $\mathbf{F}(\mathbf{S}_1 u) = (2\pi)^2 \Pi_{2\pi} * \hat{u}$. Since u belongs to L^2 then $\hat{u}$ is again in L^2. Thus, $\Pi_{2\pi} * \hat{u}$ is the 2π-periodic version of a L^2 function in $[-s\pi, s\pi]$. Consequently the inverse Fourier transform of $\Pi_{2\pi} * \hat{u}$ is a Dirac comb whose coefficients are the Fourier series coefficients of $\hat{u}$. Thus the coefficients of $\mathbf{S}_1 u$ form a ℓ^2 sequence. □

Proposition 5. *Let $h \in L^1(\mathbb{R}^2)$ and $u, v \in L^1 \cup L^2(\mathbb{R}^2)$. The following equalities hold:*

$$W_1(h * v) = W_1 h * v = h * W_1 v \tag{14}$$

$$W_1 H_\lambda v = H_\lambda W_{\frac{1}{\lambda}} v \tag{15}$$

$$H_\alpha(u * v) = H_\alpha u * H_\alpha v \tag{16}$$

Proof. The proof of (14).
In Fourier,

$$\mathbf{F}(W_1(h * v)) \overset{\text{def}}{=} \mathbf{F}(h * v) \cdot \mathbb{I}_{[-\pi,\pi]^2} \overset{(8)}{=} \mathbf{F}(h) \cdot \mathbf{F}(v) \cdot \mathbb{I}_{[-\pi,\pi]^2}$$

Thus,

$$\mathbf{F}(h) \cdot \mathbf{F}(v) \cdot \mathbb{I}_{[-\pi,\pi]^2} = \mathbf{F}(h) \cdot \mathbb{I}_{[-\pi,\pi]^2} \cdot \mathbf{F}(v) \cdot \mathbb{I}_{[-\pi,\pi]^2}$$

and all results are dedudced from this last statement.

□

Proof. The proof of (15).

$$\mathbf{F}(H_\lambda v) = \lambda^2 \mathbf{F}(v(\lambda \cdot)) = \lambda^2 \frac{1}{\lambda^2} \hat{v}(\frac{\cdot}{\lambda}) = \lambda^2 H_{\frac{1}{\lambda}} \hat{v}.$$

Thus

$$\mathbf{F}(W_1 H_\lambda v) \overset{(8)}{=} \mathbf{F}(H_\lambda v) \cdot \mathbb{I}_{[-\pi,\pi]^2} = \lambda^2 H_{\frac{1}{\lambda}} \hat{v} \cdot \mathbb{I}_{[-\pi,\pi]^2};$$

On the other hand,

$$\begin{aligned} \mathbf{F}(H_\lambda W_{\frac{1}{\lambda}} u) &= \lambda^2 H_{\frac{1}{\lambda}} \mathbf{F}(W_{\frac{1}{\lambda}} v) \\ &\overset{(8)}{=} \lambda^2 H_{\frac{1}{\lambda}} (\hat{u} \cdot \mathbb{I}_{[-\frac{\pi}{\lambda}, \frac{\pi}{\lambda}]}) \\ &= \lambda^2 (H_{\frac{1}{\lambda}} \hat{v}) \cdot \mathbb{I}_{[-\pi,\pi]^2}. \end{aligned}$$

□

Proof. The proof of (16). The proof is a mere change of variables:

$$\begin{aligned} H_\alpha(u * v)(\mathbf{x}) &= \alpha^2 \int u(\mathbf{s}) v(\alpha\mathbf{x} - \mathbf{s}) \mathrm{d}\mathbf{s} \\ &= \alpha^4 \int u(\alpha\mathbf{s}) v(\alpha\mathbf{x} - \alpha\mathbf{s}) \mathrm{d}\mathbf{s} \\ &= (H_\alpha u * H_\alpha v)(\mathbf{x}). \end{aligned}$$

□

Lemma 4. *Let $u, v \in \mathcal{BL}^2_0(\mathbb{R}^2)$. If either u or v is band-limited then*

$$\mathbf{S}_1(u * v) = \mathbf{S}_1\bar{u} * \mathbf{S}_1\bar{v}, \tag{17}$$

where we have called $\bar{u} = W_1 u$ and $\bar{v} = W_1 v$.

Proof. We will prove this statement in the tempered distribution sense. We will consider $\mathbf{S}_1 u = \Pi_1 \cdot u = \sum_{\mathbf{k}} \delta_{\mathbf{k}} \cdot u$ is a Dirac comb. The application of $\mathbf{S}_1$ to $\bar{u}$, $\bar{v}$ and $u * v$ is well defined as all functions are in $\mathcal{BL}^2(\mathbb{R}^2)$ and by consequence they are in C^∞. Recall that if $u \in \mathcal{D}'$ and f is C^∞ then $f \cdot u \in \mathcal{D}'$ thus in this framework we need a function to be in C^∞ to be sampled.

From Lemma 3 we know that the coefficients sequence of $\mathbf{S}_1\bar{u}$ and $\mathbf{S}_1\bar{v}$ are in $\ell^2(\mathbb{Z}^2)$. Thus $(\mathbf{S}_1\bar{u}) * (\mathbf{S}_1\bar{v})$ is a bounded sequence and therefore every term is well defined.

Finally $\mathbf{F}(\mathbf{S}_1(u * v)) = \Pi_{2\pi} * (\hat{u}.\hat{v}) = (\Pi_{2\pi} * \hat{u}) \cdot (\Pi_{2\pi} * \hat{v})$ is true because all considered functions happen to be 2π-periodizations of compactly supported functions in $(-\pi, \pi)^2$, namely $\hat{u}$, $\hat{v}$ and their product. □

Proposition 6. *(Discrete Camera Model) Let $u \in L^2(\mathbb{R}^2)$ and $h \in L^1 \cap \mathcal{BL}^2_0$, band-limited in $[-s\pi, s\pi]^2$. Then*

$$\mathbf{S}_1(u * h) = \mathcal{S}_s(\bar{\mathbf{u}} * \mathbf{h}), \tag{18}$$

where we have called $\bar{\mathbf{u}} = \mathbf{S}_1 W_1 H_{\frac{1}{s}} u$ and $\mathbf{h} = \mathbf{S}_1 H_{\frac{1}{s}} h$.

Proof. We first derive the expression and then justify the application of each result.

$$\begin{aligned} \mathbf{S}_1(u * h) &= \mathbf{S}_1 H_s H_{\frac{1}{s}} (u * h) \\ &\overset{(14)}{=} \mathbf{S}_1 H_s H_{\frac{1}{s}} (W_s u * h) \\ &\overset{(16)}{=} \mathbf{S}_1 H_s (H_{\frac{1}{s}} W_s u * H_{\frac{1}{s}} h) \\ &\overset{(15)}{=} \mathbf{S}_1 H_s (W_1 H_{\frac{1}{s}} u * H_{\frac{1}{s}} h) \\ &\overset{(12)}{=} \mathbf{S}_1 H_s \mathbf{I}_1 \mathbf{S}_1 (W_1 H_{\frac{1}{s}} u * H_{\frac{1}{s}} h) \\ &\overset{(17)}{=} \mathbf{S}_1 H_s \mathbf{I}_1 (\mathbf{S}_1 W_1 H_{\frac{1}{s}} u * \mathbf{S}_1 H_{\frac{1}{s}} h) \\ &\overset{\text{def}}{=} \mathbf{S}_1 H_s \mathbf{I}_1 (\bar{\mathbf{u}} * \mathbf{h}) \\ &\overset{\text{def}}{=} \mathcal{S}_s (\bar{\mathbf{u}} * \mathbf{h}). \end{aligned}$$

First notice that as u and h are in L^1 we can apply (14) (16) directly. As $W_1 u$ is in L^2 we can apply (15). Nyquist theorem (12) is valid since $u, h \in L^1$ then $W_1 H_{\frac{1}{s}} u * H_{\frac{1}{s}} h$ belongs to $\mathcal{BL}^2$.

Both $W_1 H_{\frac{1}{s}} u$ and $H_{\frac{1}{s}} h$ are band-limited finite energy functions so we are free to apply (17). Since the sequence $(\bar{\mathbf{u}} * \mathbf{h})$ is the sampling of the bandlimited L^2 function $W_1 H_{\frac{1}{s}} u * H_{\frac{1}{s}} h$ it belongs to ℓ^2 (Lemma 3). Finally, the interpolation $\mathbf{I}_1(\bar{\mathbf{u}} * \mathbf{h})$ is well defined. □

Lemma 5. *Let $h \in L^1 \cap \mathcal{BL}^2_0(\mathbb{R}^2)$ and $k \in \mathcal{BL}^2_0$ such that $\hat{k}(\zeta) = \frac{h(\zeta)}{h(\frac{\zeta}{\lambda})}$. Assume λ large enough to ensure $\hat{h}(\zeta/\lambda)$ does not vanish in the support of $\hat{k}$. Then if $\lambda > 1$ we have*

$$\lim_{n\to\infty} H_{\lambda^{n-1}} k * \ldots * H_\lambda k * k = h$$

where the limit is in $L^2 \cap C^0$.

Proof. Let us call $u_n = H_{\lambda^{n-1}} k * \ldots * H_\lambda k * k$. Then in the Fourier domain we have

$$\begin{aligned} \lim_{n\to\infty} \hat{u}_n(\zeta) &= \lim_{n\to\infty} \prod_{i=0}^{n-1} \hat{k}\left(\frac{\zeta}{\lambda^i}\right) \\ &= \lim_{n\to\infty} \frac{\hat{h}(\zeta)}{\hat{h}(\zeta/\lambda^n)} \end{aligned}$$

Since $h \in L^1(\mathbb{R}^2)$ then $\hat{h} \in C^0(\mathbb{R}^2)$ and we have

$$\lim_{n\to\infty} \hat{h}(\zeta/\lambda^n) = \hat{h}(0) = 1$$

The convergence is uniform on a fixed compact set because $\hat{h}$ is continuous and compactly supported. This implies that the convergence holds in L^1 and L^2. Therefore

$$H_{\lambda^{n-1}} k * \ldots * H_\lambda k * k \overset{L^2 \cap C^0}{\longrightarrow} h$$

□

D Stability of the inter image kernel estimation

Lemma 6. *Let $\mathbf{A}$ be a full-rank $m \times n$ matrix, with $m > n$ and $\Delta\mathbf{A}$ a perturbation of the matrix $\mathbf{A}$ such that $\mathbf{A} + \Delta\mathbf{A}$ is full rank and $\|\mathbf{A}\|\|\Delta\mathbf{A}\| < 1$. Let $\mathbf{b}$ be a $m \times 1$ vector and $\Delta\mathbf{b}$ a perturbation of $\mathbf{b}$ then the solution of $\mathbf{x} = \mathbf{A}^+\mathbf{b}$ and $\mathbf{x}^* = (\mathbf{A} + \Delta\mathbf{A})^+(\mathbf{b} + \delta\mathbf{b})$ satisfy:*

$$\frac{\|\mathbf{x}^* - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\mathrm{cond}(\mathbf{A})}{1 - \|\mathbf{A}^+\Delta\mathbf{A}\|} \left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} + \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \right). \tag{19}$$

Proof. First notice as $\mathbf{A}$ is full rank the pseudo-inverse is the left-inverse of $\mathbf{A}$, namely $\mathbf{A}^+\mathbf{A} = \mathbf{I}$. Since $\|\mathbf{A}\|\|\Delta\mathbf{A}\| < 1$ we have that

$$(\mathbf{A} + \Delta\mathbf{A})^+ = (\mathbf{I} + \mathbf{A}^+\Delta\mathbf{A})^{-1}\mathbf{A}^+$$

and we also have

$$\|(\mathbf{I} + \mathbf{A}^+\Delta\mathbf{A})^{-1}\| = \left\|\sum(\mathbf{A}^+\Delta\mathbf{A})^k\right\| \leq \sum\|(\mathbf{A}^+\Delta\mathbf{A})\|^k = \frac{1}{1 - \|\mathbf{A}^+\Delta\mathbf{A}\|}.$$

Hence,

$$\begin{aligned}\mathbf{x}^* - \mathbf{x} &= (\mathbf{A} + \Delta\mathbf{A})^+(\mathbf{b} + \delta\mathbf{b}) - \mathbf{A}^+\mathbf{b} \\ &= (\mathbf{I} + A^+\Delta\mathbf{A})^{-1}\mathbf{A}^+(\mathbf{b} + \delta\mathbf{b}) - \mathbf{A}^+\mathbf{b}\end{aligned}$$

therefore

$$\begin{aligned}(\mathbf{I} + \mathbf{A}^+\Delta\mathbf{A})(\mathbf{x}^* - \mathbf{x}) &= \mathbf{A}^+(\mathbf{b} + \delta\mathbf{b}) - \mathbf{A}^+\mathbf{b} - \mathbf{A}^+\Delta\mathbf{A}\mathbf{A}^+\mathbf{b} \\ &= \mathbf{A}^+(\delta\mathbf{b} - \Delta\mathbf{A}\mathbf{x})\end{aligned}$$

and then

$$\begin{aligned}\frac{\|\mathbf{x}^* - \mathbf{x}\|}{\|\mathbf{x}\|} &\leq \frac{\|\mathbf{A}^+\|}{1 - \|\mathbf{A}^+\Delta\mathbf{A}\|}\frac{\|\delta\mathbf{b}\| + \|\Delta\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} \\ &= \frac{\mathrm{cond}(\mathbf{A})}{1 - \|\mathbf{A}^+\Delta\mathbf{A}\|}\frac{\|\delta\mathbf{b}\| + \|\Delta\mathbf{A}\mathbf{x}\|}{\|\mathbf{A}\|\|\mathbf{x}\|} \\ &\leq \frac{\mathrm{cond}(\mathbf{A})}{1 - \|\mathbf{A}^+\Delta\mathbf{A}\|}\left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{A}\mathbf{x}\|} + \frac{\|\Delta\mathbf{A}\|\|\mathbf{x}\|}{\|\mathbf{A}\|\|\mathbf{x}\|}\right) \\ &\leq \frac{\mathrm{cond}(\mathbf{A})}{1 - \|\mathbf{A}^+\Delta\mathbf{A}\|}\left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} + \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}\right)\end{aligned}$$

□

References

[1] A. Almansa, S. Durand, and B. Rougé, *Measuring and improving image resolution by adaptation of the reciprocal cell*, Journal of Mathematical Imaging and vision, 21 (2004), pp. 235–279.

[2] Sampo Backman, Anssi J. Maekynen, Timo T. Kolehmainen, and Kai M. Ojala, *Fast lens testing using random targets*, Opto-Ireland 2002: Optics and Photonics Technologies and Applications, 4876 (2003), pp. 1100–1109.

[3] S.M. Backman, A.J. Makynen, T.T. Kolehmainen, and K.M. Ojala, *Random target method for fast mtf inspection.*, Optics Express, 12 (2004).

[4] Jean M. Bony, *Cours d'analyse. Théorie des distributions et analyse de Fourier*, Les Éditions de l'École Polytechnique, 2001.

[5] F. L. Bookstein, *Principal warps: Thin-plate splines and the decomposition of deformations*, IEEE Trans. Pattern Anal. Mach. Intell., 11 (1989), pp. 567–585.

[6] Johannes Brauers and Til Aach, *Direct PSF estimation using a random noise target*, in IS&T/SPIE Electronic Imaging: Digital Photography VI, J. Allebach and S. Süsstrunk, eds., San Jose, USA, January 17–21 2010, SPIE-IST Vol. 7537, p. to appear.

[7] David Capel, *Image Mosaicing and Super-Resolution (Cphc/Bcs Distinguished Dissertations.)*, SpringerVerlag, 2004.

[8] Bernard Chalmond, *Psf estimation for image deblurring*, CVGIP: Graphical Models and Image Processing, 53 (1991), pp. 364 – 372.

[9] S. Chaudhuri and AN Rajagopalan, *Depth from defocus: a real aperture imaging approach*, Springer Verlag, 1999.

[10] Christopher D. Claxton and Richard C. Staunton, *Measurement of the point-spread function of a noisy imaging system*, J. Opt. Soc. Am. A, 25 (2008), pp. 159–170.

[11] A. Daniels, G.D. Boreman, A.D. Ducharme, and E. Sapir, *Random transparency targets for modulation transfer function measurement in the visible and infrared regions.*, Optical Engineering, 34 (1995), pp. 860 – 868.

[12] M. Delbracio, P. Musé, and A. Almansa, *Non-parametric sub-pixel local point spread function estimation*, Image Processing on Line, workshop, algorithmic description, online demo and source code accessible at `http://www.ipol.im/pub/algo/admm\_non\_blind\_psf\_estimation` (2011). 2011.

[13] M. Delbracio, P. Musé, A. Almansa, and J.-M. Morel, *The non-parametric sub-pixel local point spread function estimation is a well posed problem*, International Journal of Computer Vision, (2011), pp. 1–20. 10.1007/s11263-011-0460-0.

[14] A. Delbracio M., Almansa and P. Musé, *Blind subpixel point spread function estimation from scaled image pairs*, Image Processing on Line, workshop, algorithmic description, online demo and source code accessible at `http://www.ipol.im/pub/algo/damm_blind_psf_estimation_from_scaled_image_pairs/` (2011).

[15] Lars Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin,, 1983.

[16] ISO, *ISO 12233:2000: Photography – Electronic still-picture cameras – Resolution measurements*, (2000), p. 32.

[17] Neel Joshi, *Enhancing Photographs using Content-Specific Image Priors*, PhD thesis, Department of Computer Science and Engineering, University of California, San Diego, September 2008.

[18] Neel Joshi, Richard Szeliski, and David J. Kriegman, *Psf estimation using sharp edge prediction*, Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, 0 (2008), pp. 1–8.

[19] E. Levy, D. Peles, M. Opher-Lipson, and S.G. Lipson, *Modulation transfer function of a lens measured with a random target method.*, Applied Optics, 38 (1999), pp. 679 – 683.

[20] Imatest LLC, *Imatest 3.6*. `http://www.imatest.com/`, 2010.

[21] D. G. Lowe, *Object recognition from local scale-invariant features*, in Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, vol. 2, Los Alamitos, CA, USA, August 1999, IEEE Computer Society, pp. 1150–1157 vol.2.

[22] M. Luxen and W. Förstner, *Characterizing image quality: Blind estimation of the point spread function from a single image*, (2002), p. A: 205.

[23] JM Morel and S. Ladjal, *Notes sur l'analyse de Fourier et théorie de Shannon en traitement d'images*, in Analyse de Fourier et traitement dimages. Journées X-UPS, Ecole Polytechnique, F91128 Palaiseau cedex, France., 1998, http://www.math.polytechnique.fr/xups/, pp. 37–100.

[24] Jean-Michel Morel and Guoshen Yu, *Asift: A new framework for fully affine invariant image comparison*, SIAM J. Img. Sci., 2 (2009), pp. 438–469.

[25] ———, *ASIFT: A New Framework for Fully Affine Invariant Comparison*, Image Processing On Line, (2010).

[26] M.K. Ng, R.H. Chan, and W.C. Tang, *A fast algorithm for deblurring models with neumann boundary conditions*, SIAM Journal on Scientific Computing, 21 (2000), pp. 851–866.

[27] S.C. Park, M.K. Park, and M.G. Kang, *Super-resolution image reconstruction: a technical overview*, Signal Processing Magazine, IEEE, 20 (2003), pp. 21–36.

[28] D. Rajan and S. Chaudhuri, *Simultaneous estimation of super-resolved scene and depth map from low resolution defocused observations*, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 25 (2003), pp. 1102 – 1117.

[29] Stephen E. Reichenbach, Stephen K. Park, and Ramkumar Narayanswamy, *Characterizing digital image acquisition devices*, Optical Engineering, 30 (1991), pp. 170–177.

[30] Elisa H. Barney Smith, *Psf estimation by gradient descent fit to the esf*, Image Quality and System Performance III, 6059 (2006), p. 60590E.

[31] R. Sprengel, K. Rohr, and H.S. Stiehl, *Thin-plate spline approximation for image registration*, in Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE, vol. 3, Oct-3 Nov 1996, pp. 1190–1191 vol.3.

[32] Elias M. Stein and Guido Weiss, *Introduction to Fourier Analysis on Euclidean Spaces. (PMS-32)*, Princeton University Press, Nov. 1971.

[33] Orly Yadid-Pecht, *Geometrical modulation transfer function for different pixel active area shapes*, Optical Engineering, 39 (2000), pp. 859–865.

[34] J.A. Zandhuis, D. Pycock, S.F. Quigley, and P.W. Webb, *Sub-pixel non-parametric psf estimation for image enhancement*, IEE Proceedings - Vision, Image, and Signal Processing, 144 (1997), pp. 285–292.