Accurate Subpixel Point Spread Function Estimation from scaled image pairs

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

Accurate Subpixel Point Spread Function Estimation from scaled image pairs

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

Abstract

In most digital cameras, and even in high-end digital SLRs, the acquired images are sampled at rates below the Nyquist critical rate, causing aliasing effects. This work introduces an 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 these two images, under reasonable conditions. 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 photographs. 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 spectrum modulus (the Modulated Transfer Function, MTF) to describe their

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

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 [28], super-resolution [29, 31] or shape from defocus [10].

In most typical digital cameras, both compact and high-end DSLRs, images are sampled at frequencies below Nyquist critical rate. Consequently, only aliased versions of the camera PSF can be directly observed. Yet, to fully characterize the PSF, it is necessary to recover it at a subpixel resolution.

PSF estimation methods can be classified as non-blind or blind, depending on whether they use a snapshots of a specially designed calibration pattern. Blind approaches try to estimate the PSF from photographs from an unknown scene. They do assume, however, that the scene involved in the estimation follow some statistical model of sharp images, or include a significant amount of geometric cues such as sharp edges. Most of these PSF approaches attempt to detect edges, which are modeled as pure step-edge functions convolved with the PSF kernel [9, 24, 8, 4]. In this setting, the 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. Therefore, such blind estimation techniques do not lead to accurate PSF estimates, and are normally an ingredient in image restoration problems, where precision is not the main objective. For this reason, accurate PSF estimation procedures rely on the use of specially designed calibration patterns. A local kernel estimation is performed by comparing the ideal calibration pattern to its photographs.

Several patterns have been used for PSF estimation, ranging from pin-hole, slanted-edge [18, 30, 38, 11], or arc-step-edge patterns [20, 19] to random noise images [12, 21, 2, 3, 7]. Until recently, even non-blind sub-pixel PSF estimation methods reported in the literature led to ill-posed inverse problems. The inversion required the imposition of simple PSF parametric models, or other regularity or symmetry priors. In a recent work [14] we have shown that such *a priori* assumptions on the PSF are actually unnecessary and jeopardize the estimation accuracy. More precisely, by carefully modeling the image acquisition system, a calibration pattern made of a white noise realization is nearly optimal in terms of well-conditioning of the problem. This procedure leads to very accurate regularization-free subpixel PSF estimation.

The purpose of the present work is to explore the feasibility of obtaining accurate PSF estimates, while avoiding the explicit use of a calibration pattern. The motivation comes from the fact that, although very precise, the use of a calibration pattern can be sometimes tedious and impractical: these approaches rely on a careful setup, and the calibration grid has to be properly assembled where a good quality print is essential.

We show that, instead of using a photograph of a known calibration pattern, two photographs of the same scene acquired at different distances with fixed camera configuration are enough to recover a regularization-free subpixel PSF. The proposed acquisition procedure is simple and handy in comparison to a non-blind approach. 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 important but not critical. For a wide range of everyday textured scenes, the acquired image pairs lead to well posed inversions and highly accurate results.

This paper is written with a dual public 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.

The article is organized as follows: Section 2 presents a mathematical model of the digital image acquisition system. 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. We define the notion of blur between such pair of images, and we propose a method to perform its estimation. Then we prove that the camera PSF can be recovered from this inter-image blur. Section 4 presents an algorithm that implements the complete PSF estimation procedure described in Section 3. In Section 5 we discuss 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.

2 Image Formation Model

2.1 Generalized Digital Pin-hole Camera

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 scale factor $\lambda = \frac{d}{f}$ of the original planar scene u.

If the pose is not perfectly fronto-parallel, or the pin-hole camera presents non-canonical internal calibration parameters, w and u are related by a planar homography D, *i.e.* $w = u \circ D$. In a more accurate camera model 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 [38, 20, 14].

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

ments:

Blurring

The PSF kernel h models blur due to intrinsic camera characteristics, such as diffraction when light goes through a finite aperture, light averaging within the sensor, lens aberration, etc. Other blur sources like motion, atmospheric turbulence or defocus blur, that may change from one snapshot to another will be minimized by the experimental procedure and it is not the goal of the present work to estimate them.

The diffraction kernel is determined by the shape and size of the aperture, the focal length, and the wavelength of the considered monochromatic light. Under the Fraunhofer far-field approximation, for incoherent light this kernel is the squared Fourier transform modulus of the camera's aperture indicator function [15]. It follows that the PSF diffraction kernel is always non-negative and band-limited.

Besides the kernel due to diffraction, other sources of blur inherent to the optical system are present in real cameras. These are mainly optical aberrations, and anti-aliasing filters (that reduce aliasing but do not completely cancel it) introduced in the system prior to sampling [35, 39] . The sampling process also introduces blur. Indeed, each photo-sensor in the rectangular sampling grid integrates the light arriving at a particular exposure time. This corresponds to a convolution with the indicator function of the photo-sensor active area. To sum up, *the unknown* PSF results basically from the convolution of three non-negative kernels (diffraction, aberrations and anti-aliasing filters, and sensor averaging) one of them being band-limited. No parametrical model on the PSF will be adopted here. Nonetheless the physical modeling justifies our assumption that the PSF is *band-limited and non-negative.*

Sampling

A model of continuous to digital conversion at the image plane, *i.e.* an ideal *sampling* operator $\mathbf{S}_1$ and additive *noise* $\mathbf{n}$ due to measurement uncertainties. Physical models of digital camera sensors, both for CCD and CMOS sensors, suggest that the readout noise $\mathbf{n}$ is a mixture of luminance independent (Gaussian, thermal) noise, and luminance dependent (Poisson or photon counting) noise [16, 34, 25]. A usual simplification of this model, which we follow here, assumes the noise is image independent, white, Gaussian, with constant variance.

The whole image formation process can then 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, in RAW images this response can be reasonably assumed to be linear [14]. This boils

u, v	Images defined on continuous domain $\mathbf{x} \in \mathbb{R}^2$
$\mathbf{u}$, $\mathbf{v}$	Digital images are sampled on a discrete grid $\mathbf{k} \in \mathbb{Z}^2$
$\mathbf{F}$	Fourier transform
$\hat{f}$	Fourier transform of a function f
$\mathbf{I}_1$	Shannon-Whittaker interpolator: $\mathbf{I}_1\mathbf{u}(\mathbf{x}) = \sum_{\mathbf{k}} \mathbf{u}(\mathbf{k})\mathrm{sinc}(\mathbf{x}-\mathbf{k})$
$\mathbf{S}_1$	1-sampling operator: $\mathbf{u}(\mathbf{k}) = (\mathbf{S}_1 u)(\mathbf{k}) = u(\mathbf{k})$
W_w	Ideal low-pass filter that cuts the spectrum of continuous signals to $[-w\pi, w\pi]^2$
$\mathcal{S}_s$	The s-to-1-resampling operator $\mathcal{S}_s = \mathbf{S}_1 H_s \mathbf{I}_1$
H_λ	Continuous homothecy: $H_\lambda u(x,y) = \lambda^2 u(\lambda x, \lambda y)$. ($\lambda < 1$ dilation, $\lambda > 1$ contraction)
$\mathbf{H}_\alpha$	Digital Nyquist homothecy operator of parameter α: $\mathbf{H}_\alpha\mathbf{u} := \mathbf{S}_1 W_1 H_\alpha \mathbf{I}_1 \mathbf{u}$
$\mathbf{C}[\mathbf{u}]$	Linear map associated to the convolution with a digital image $\mathbf{u}$
$\mathbf{L}^*$	Adjoint of a linear operator $\mathbf{L}$
$\mathbf{L}^+$	Pseudo-inverse $\mathbf{L}^+ := (\mathbf{L}^*\mathbf{L})^{-1}\mathbf{L}^*$ of a linear operator $\mathbf{L}$
L^1	Integrable functions on $\mathbb{R}^2$ ($L^1(\mathbb{R}^2)$)
L^2	Square integrable functions ($L^2(\mathbb{R}^2)$).
$\mathcal{BL}^2$	L^2 functions, band limited in $[-\pi, \pi]^2$.
$\mathcal{BL}^2_0$	L^2 functions with compactly supported Fourier transform.

Table 1: Notation used in this article

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

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

2.2 Inverse problem statement in terms of digital sequences

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, summarized in Table 1 (a more precise definition of each term is presented in Appendix A). 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.

Suppose that the PSF h is s-band-limited, that is $\mathrm{supp}(\hat{h}) = [-s\pi, s\pi]^2$. Then, if sampled at a rate s, the Nyquist sampling theorem guarantees 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) may be subject to aliasing. Following Lemma 6, the image formation model (M) can be written in terms of discrete sequences:

$$\begin{aligned} \tilde{\mathbf{v}} &= \mathcal{S}_s(\bar{\mathbf{u}}_D * \mathbf{h}) + \mathbf{n}, \\ &= \mathcal{S}_s \mathbf{C}[\bar{\mathbf{u}}_D]\mathbf{h} + \mathbf{n}. \end{aligned} \tag{1}$$

The digital image $\bar{\mathbf{u}}_D = \mathbf{S}_1 W_1 H_{\frac{1}{s}} u_D$ is a well-sampled version of the distorted

image $u_D = u \circ D$. 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 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[1]), and that it is nearly minimal when $\bar{\mathbf{u}}_D$ is a white noise realization (see [14]).

3 PSF Estimation from an unknown pair of scaled images

Assume that we have perfect knowledge of the latent sharp image u that produced the blurry aliased observation $\tilde{\mathbf{v}}$. Under this non-blind assumption solving for the PSF amounts to solving an inverse problem governed by the image formation model (M). Of course, this would require the use of a specially designed calibration pattern. We are now interested in investigating to what extent the use of such pattern could be circumvented. We will propose a method that allows to accurately estimate the PSF from a pair of snapshots of the same scene, captured from different distances. In this method, the closest image will play a similar role to a calibration pattern in a classical non-blind approach.

In a previous work, we have shown that the highest accuracy in the PSF estimation is obtained by using a realization of a Bernoulli white noise as calibration target [14]. However, many highly textured scenes do exist in nature that, while not being optimal, may still lead to a well-posed inverse problem. In what follows, we prove that from two far apart snapshots of this kind of scene, complete recovery of the camera PSF is theoretically possible based on the estimation of the blur between this pair.

3.1 Relative blur between two images: the inter-image kernel

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

[1] 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.

the optical axis. Let λ_1 and λ_2 denote the corresponding scale factors between the scene and each of the images. Then,

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

where $v_i := H_{\lambda_i} u * h$ and $\mathbf{v}_i := \mathbf{S}_1 v_i$. We will realistically assume that $h \in L^1 \cap \mathcal{BL}_0^2$ is non-negative with $\|h\|_{L^1} = 1$, and $u \in \mathcal{BL}_0^2$ (details on the appropriateness of these assumptions are given in Appendix A.1). 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 $v_1, v_2 \in \mathcal{BL}_0^2$ 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 $k \in \mathcal{BL}_0^2$ 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 $h \in L^1 \cap \mathcal{BL}_0^2$ 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 $\lambda_2\rho > s\lambda_1$. Then there is an inter-image kernel $k \in \mathcal{BL}_0^2$ with support in $[-s\pi, s\pi]^2$ between (fronto-parallel views) v_1 and v_2 that satisfies*

$$H_\lambda h * k = h, \qquad \text{where } \lambda = \frac{\lambda_2}{\lambda_1}. \tag{3}$$

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. If k is an inter-image kernel between v_1 and v_2, according to Definition 1 it must satisfy

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

Since $v_i := H_{\lambda_i} u * h$, the right hand side of the previous equation is given by

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

In the same way, for the left-hand side,

$$H_\lambda v_1 = H_\lambda (H_{\lambda_2} u * h) \overset{(20)}{=} H_{\lambda_2} u * H_\lambda h,$$

i.e. $\mathbf{F}(H_\lambda v_1)(\zeta)\hat{k}(\zeta) = \hat{u}(\zeta/\lambda_2)\,\hat{h}(\zeta/\lambda)$. Hence,

$$\hat{u}(\zeta/\lambda_2)\,\hat{h}(\zeta/\lambda)\,\hat{k}(\zeta) = \hat{u}(\zeta/\lambda_2)\,\hat{h}(\zeta). \tag{4}$$

It follows that a sufficient condition for k to be an inter-image kernel is $\hat{h}\left(\zeta/\lambda\right)\hat{k}(\zeta) = \hat{h}\left(\zeta\right)$. Since $h \in L^1$, $\hat{h}$ is continuous. It follows that ρ is necessarily positive, since $\hat{h}(0) = 1 > 0$. In addition, as $\lambda > \frac{s}{\rho}$ by hypothesis, $\mathbf{F}(H_\lambda h)(\zeta) = \hat{h}(\zeta/\lambda)$ does not vanish inside $[-s\pi, s\pi]^2$ and

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

is well defined all over its support, $\mathrm{supp}(\hat{k}) \subset [-s\pi, s\pi]^2$. Finally, if $\hat{u}\left(\zeta/\lambda_2\right)$ does not vanish within the support of $\hat{h}$, from Eq. (4) 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. We are also assuming that for a fixed camera configuration, as long as both images are taken in focus, the* PSF *remains constant. Second, the common area between $\mathbf{v}_1$ and $\mathbf{v}_2$ covers an important part of the former one, and consequently its* PSF *may exhibit some space variance that may degrade the estimation. This can be avoided by taking snapshots with their common area covering only the central part of the frame, where the kernel does not change significantly.*

3.2 Estimation of the inter-image kernel

The next goal is to estimate the inter-image kernel k. Being k a s-band-limited function, we will work with its $s\times$ samples $\mathbf{k} = \mathbf{S}_1 H_{\frac{1}{s}} k$. We will show that under reasonable conditions, $\mathbf{k}$ can be recovered from the noisy aliased observations $\tilde{\mathbf{v}}_1$ and $\tilde{\mathbf{v}}_2$. Let us first build up some intuition on how to derive the proposed estimator. In what follows, $\mathring{\mathbf{v}}_1 = \mathbf{S}_1 W_1 H_{\frac{\lambda}{s}} v_1$ denotes a well sampled homothecy of parameter λ/s of v_1.

Proposition 1. *Under the assumptions of Lemma 1,*

$$\mathbf{v}_2 = (\mathcal{S}_s \mathbf{C}[\mathring{\mathbf{v}}_1])\mathbf{k}. \tag{6}$$

Proof. Being k an inter-image kernel between v_1 and v_2, it satisfies Eq. (3). Then,

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

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

$$\mathbf{v}_2 \overset{(18)}{=} \mathbf{S}_1(W_s H_\lambda v_1 * W_s k). \tag{7}$$

Using the Nyquist-Shannon theorem for a band-limited signal, and a set of properties detailed in the Appendix, yields

$$\begin{aligned}
\mathbf{v}_2 &\overset{s\geq 0}{=} \mathbf{S}_1 H_s H_{\frac{1}{s}}(W_s H_\lambda v_1 * W_s k) \\
&\overset{(20)}{=} \mathbf{S}_1 H_s (H_{\frac{1}{s}} W_s H_\lambda v_1 * H_{\frac{1}{s}} W_s k) \\
&\overset{(19)}{=} \mathbf{S}_1 H_s (W_1 H_{\frac{\lambda}{s}} v_1 * W_1 H_{\frac{1}{s}} k) \\
&\overset{(16)}{=} \mathbf{S}_1 H_s \mathbf{I}_1 \mathbf{S}_1 (W_1 H_{\frac{\lambda}{s}} v_1 * W_1 H_{\frac{1}{s}} k) \\
&\overset{(21)}{=} \mathbf{S}_1 H_s \mathbf{I}_1 (\mathbf{S}_1 W_1 H_{\frac{\lambda}{s}} v_1 * \mathbf{S}_1 W_1 H_{\frac{1}{s}} k) \\
&\overset{\text{def}}{=} \mathbf{S}_1 H_s \mathbf{I}_1 (\mathring{\mathbf{v}}_1 * \mathbf{k}) \\
&\overset{\text{def}}{=} \mathcal{S}_s (\mathring{\mathbf{v}}_1 * \mathbf{k}) \\
&\overset{\text{def}}{=} \mathcal{S}_s \mathbf{C}[\mathring{\mathbf{v}}_1]\mathbf{k}.
\end{aligned}$$

□

Of course, in practice we do not have access to $\mathring{\mathbf{v}}_1$ nor to $\mathbf{v}_2$, but to their noisy, aliased versions $\tilde{\mathbf{v}}_1$ and $\tilde{\mathbf{v}}_2$. Thus $\mathbf{k}$ cannot be directly estimated from Eq. (6). However, a relationship between $\mathring{\mathbf{v}}_1$ and $\tilde{\mathbf{v}}_1$ can be established as follows.

$$\begin{aligned}
\mathbf{H}_{\frac{\lambda}{s}} \tilde{\mathbf{v}}_1 &= \mathbf{H}_{\frac{\lambda}{s}} (\mathbf{v}_1 + \mathbf{n}_1) + \mathring{\mathbf{v}}_1 - \mathring{\mathbf{v}}_1 \\
&= \mathring{\mathbf{v}}_1 + \underbrace{\mathbf{S}_1 W_1 H_{\frac{\lambda}{s}} (\mathbf{I}_1 \mathbf{v}_1 - v_1)}_{\mathbf{r}} + \mathbf{H}_{\frac{\lambda}{s}} \mathbf{n}_1,
\end{aligned} \tag{8}$$

where the last equality results from the definition of the discrete homothecy operator. The term $\mathbf{r}$ is a consequence of aliasing when sampling v_1, and introduces an unknown bias in the estimation of $\mathbf{k}$. While this bias cannot be fully controlled, its impact can be mitigated. Indeed, since

$$\begin{aligned}
\mathbf{r} &= \mathbf{S}_1 W_1 H_{\frac{\lambda}{s}} (\mathbf{I}_1 \mathbf{v}_1 - v_1) \\
&\overset{(19)}{=} \mathbf{S}_1 H_{\frac{\lambda}{s}} W_{\frac{s}{\lambda}} (\mathbf{I}_1 \mathbf{v}_1 - v_1),
\end{aligned}$$

the aliasing term $\mathbf{r}$ will only be non-zero if there are aliasing components in the frequency interval $\left[-\frac{s}{\lambda}\pi, \frac{s}{\lambda}\pi\right]^2$. This allows 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). Thus, 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.

From now on, we assume that the snapshots are acquired following the previous considerations. Therefore, we can ignore the aliasing term in Eq. (8), what leads to

$$\mathbf{v}_2 = (\mathcal{S}_s \mathbf{C}[\mathring{\mathbf{v}}_1])\mathbf{k} = (\mathcal{S}_s \mathbf{C}[\mathbf{H}_{\frac{\lambda}{s}} \mathbf{v}_1 - \mathbf{H}_{\frac{\lambda}{s}} \mathbf{n}_1])\mathbf{k},$$

that is

$$(\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\Delta}}\tilde{\mathbf{v}}_1 - \mathbf{H}_{\underset{s}{\Delta}}\mathbf{n}_1])\mathbf{k} = \tilde{\mathbf{v}}_2 - \mathbf{n}_2.$$

One could be tempted to solve for $\mathbf{k}$ in the previous equation using a Total Least Squares based approach:

$$\text{(TLS)} \quad \underset{\mathbf{k},\boldsymbol{\delta},\boldsymbol{\epsilon}}{\arg\min} \ \|\boldsymbol{\delta}\| + \kappa\|\boldsymbol{\epsilon}\| \ \text{ subject to } \ \mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\Delta}}\tilde{\mathbf{v}}_1 + \boldsymbol{\delta}]\mathbf{k} = \tilde{\mathbf{v}}_2 + \boldsymbol{\epsilon}.$$

However, the particular structure of the operator $\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\Delta}}\tilde{\mathbf{v}}_1 + \boldsymbol{\delta}]$ makes this problem a difficult one. Instead we prefer to follow a simpler approach, that results from neglecting the noise term $\mathbf{H}_{\underset{s}{\Delta}}\mathbf{n}_1$. This yields to the least squares estimation problem

$$\text{(LS)} \quad \underset{\mathbf{k},\boldsymbol{\epsilon}}{\arg\min} \ \|\boldsymbol{\epsilon}\| \ \text{ subject to } \ \mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\Delta}}\tilde{\mathbf{v}}_1]\mathbf{k} = \tilde{\mathbf{v}}_2 + \boldsymbol{\epsilon},$$

whose solution is given by

$$\mathbf{k}_e = \left(\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\Delta}}\tilde{\mathbf{v}}_1]\right)^+\tilde{\mathbf{v}}_2. \tag{9}$$

If the noise $\mathbf{n}_1$ is small compared to $\mathbf{v}_1$, this solution would be very close to the one that would be obtained from (TLS). If in addition $\mathbf{n}_2$ is small compared to $\mathbf{v}_2$, both solutions would be close to the actual inter-image kernel $\mathbf{k} = \left(\mathcal{S}_s\mathbf{C}[\mathring{\mathbf{v}}_1]\right)^+\mathbf{v}_2$. This follows directly from the continuity and injectivity assumptions on $\mathcal{S}_s\mathbf{C}[\mathring{\mathbf{v}}_1]$, as a consequence of Lemma 6. This being said, we will consider the estimator of the inter-image kernel in Eq. (9).

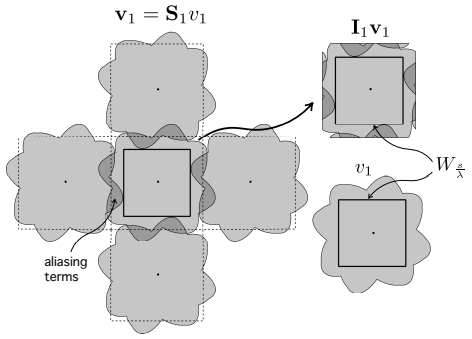


Figure 1: *Neglecting the aliasing* The estimation will only be affected by aliasing if there are aliasing components in the interval $\left[-\frac{s}{\lambda}\pi, \frac{s}{\lambda}\pi\right]^2$. Hence, to avoid aliasing one can choose $v_1 = H_{\lambda_1}u$ such that $\text{supp}(\hat{v_1}) \subset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

Remark 2. *If $\lambda < s$, the convolution between $\mathbf{k}$ and $\mathbf{H}_{\underset{s}{\lambda}}\tilde{\mathbf{v}}_1$ is not invertible so the operator $\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\lambda}}\tilde{\mathbf{v}}_1]$ will not be injective. This constraint on λ is necessary but not sufficient to make $\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\underset{s}{\lambda}}\tilde{\mathbf{v}}_1]$ invertible. In addition, it is required that the spectrum of the image $\mathbf{H}_{\underset{s}{\lambda}}\tilde{\mathbf{v}}_1$ exhibits slow-decay. Indeed, it is shown in [14] that the more flat the spectrum of the image scene, the better conditioned is the inverse problem. For that reason, in order to obtain accurate estimates of $\mathbf{k}$, it is desirable that the chosen scene u exhibits white noise characteristics.*

3.3 From relative to absolute blur

Now that we have a method to estimate the inter-image kernel $\mathbf{k}$, we will concentrate on how to recover the camera PSF. Notice that h is related to k by $H_\lambda h * k = h$, and therefore its derivation is not straightforward. However, as we prove in Appendix C, Proposition 5, it holds that

$$h = \lim_{n\to\infty} H_{\lambda^{n-1}}k * \ldots * H_\lambda k * k. \tag{10}$$

This equality shows that it is possible to recover the camera PSF h from the inter-image kernel k. Recall that in practice we only have access to discrete sequences, therefore it is convenient to derive a discrete equivalent of the previous limit. Since $\mathbf{k}$ is s-bandlimited,

$$\begin{aligned}
\mathbf{S}_1 H_{\underset{s}{1}}(H_\lambda k * k) &\overset{(18)}{=} \mathbf{S}_1 H_{\underset{s}{1}}(W_s H_\lambda k * k)\\
&\overset{(20)}{=} \mathbf{S}_1(H_{\underset{s}{1}} W_s H_\lambda k * H_{\underset{s}{1}} k)\\
&\overset{(19)}{=} \mathbf{S}_1(W_1 H_{\underset{s}{\lambda}} k * H_{\underset{s}{1}} k)\\
&\overset{(22)}{=} \mathbf{S}_1 W_1 H_{\underset{s}{\lambda}} k * \mathbf{S}_1 H_{\underset{s}{1}} k\\
&\overset{\text{def}}{=} \mathbf{H}_\lambda \mathbf{k} * \mathbf{k}.
\end{aligned}$$

Iteratively applying this result to Eq (10) yields

$$\mathbf{h} = \lim_{n\to\infty} \mathbf{H}_{\lambda^{n-1}}\mathbf{k} * \ldots * \mathbf{H}_\lambda \mathbf{k} * \mathbf{k}. \tag{11}$$

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. This alignment can be obtained by matching SIFT descriptors [23], which have the advantage of being

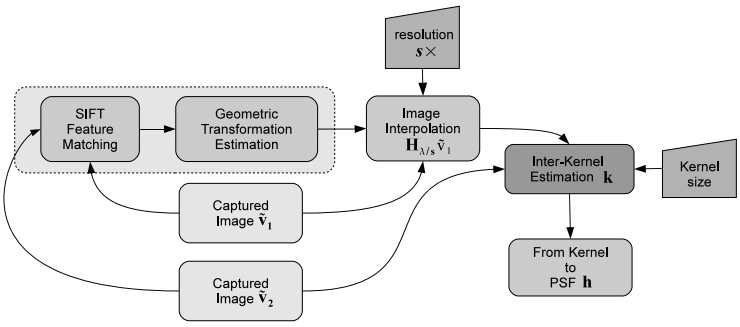


Figure 2: *Algorithm Description.* Both captured images are aligned via SIFT feature matching followed by the estimation of a smooth geometric distortion through thin-plate approximation of matched features. The relative geometric distortion and gray-level corrections are applied to a low-pass (unaliased) version of the finest scale image $\tilde{\mathbf{v}}_1$. Then the interpolated image $\mathbf{H}_{\frac{\lambda}{s}}\tilde{\mathbf{v}}_1$ and image $\tilde{\mathbf{v}}_2$ are compared to obtain the inter-image kernel k, which is later iteratively updated to obtain the absolute camera PSF h.

scale invariant. The IPOL implementation of ASIFT [27, 37] 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 [6, 32] from the matched SIFT pairs. This permits the correction of small non-affine distortions, and for deviations from the fronto-parallel assumption in the acquisition. Of course, if the distortion is significant the assumed inter-image kernel Eq. (3) will not be accurate. The thin-plate representation as affine + non-affine parts of the geometric distortion is especially helpful to estimate the relative scale $\lambda = (\lambda_x, \lambda_y)$ between both views, since this can be directly estimated from the affine part.

Gray level adjustment Both snapshots should be acquired with exactly the same camera configuration, and constant scene illumination. This ensures that there is no contrast change between them.

Resampling and Distortion Correction of $\tilde{\mathbf{v}}_1$ The generation of the rescaled samples $\mathbf{H}_{\frac{\lambda}{s}}\tilde{\mathbf{v}}_1$, requires the interpolation of $\tilde{\mathbf{v}}_1$ at the desired scale λ/s. This is done by using the estimated geometric transformation with bicubic interpolation. Notice that since $\tilde{\mathbf{v}}_1$ is not very aliased, one can correctly interpolate it

without introducing artifacts.

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 contained in a $r \times r$ patch. Then the matrix $\mathcal{S}_s\mathbf{C}[\mathbf{H}_{\frac{\lambda}{s}}\tilde{\mathbf{v}}_1]$ is of size $mn \times r^2$. A simple least-squares procedure yields the inter-image kernel estimator:

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

Transforming the kernel: from k to h Recovering the samples of the camera PSF $\mathbf{h}$ amounts to evaluate the limit in Eq (11). Directly working with the digital sequences requires some care about how the successive convolutions are computed. Since $\lambda > 1$, the application of $\mathbf{H}_\lambda$ would require a lowpass filter to avoid aliasing artifacts. To bypass this inconvenient one can re-state the limit convolution as follows:

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

If implemented in this way, the successive discrete convolutions can be computed without any special care. To apply the discrete homothecy operator to $\mathbf{k}$, we need to resample $\mathbf{k}$ using the Shannon-Whittaker interpolator. Because of its slow decay, in order to reduce ringing and other windowing effects, we opted to use bicubic interpolation. The whole derivation of $\mathbf{h}$ from $\mathbf{k}$ is summarized as follows:

Algorithm 1: From inter-image kernel to PSF

1. Initialize $\mathbf{w}_0 = \mathbf{k}$, $n = 1$
2. Compute $\mathbf{H}_{1/\lambda^n}\mathbf{k}$ by using $\lambda = (\lambda_x, \lambda_y)$ (from thin-plates affine part).
3. Calculate $\mathbf{u}_n = \mathbf{H}_{1/\lambda^n}\mathbf{k} * \mathbf{w}_{n-1}$.
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 $\mathbf{h} = \mathbf{H}_\lambda^n\mathbf{w}_n$.

This algorithm converges after a few iterations since λ^n grows very fast. In practice, we set $\lambda_{\max} = 50$, since the convolution with a $50\times$ scale-contracted inter-image kernel produces a negligible change in the final result (it amounts to convolve with a Kronecker delta).

In theory, as we already stated, the estimated PSF should be non-negative. In practice, small negative values may be observed, due to deviations form model

assumptions and numerical artifacts. To correct for these deviations, we simply set to zero all negative values.

5 Experimental Results

Since there is no PSF ground-truth available, the validation of the proposed method was carried out by simulations and by comparing the results with state-of-the-art methods [20, 19, 14, 22]. Comparison was made only to non-blind, target based methods, as the accuracy of blind methods is significantly lower. A complete algorithmic description, an online demo facility and a reference source code can be found at the IPOL workshop by [13].

5.1 Simulations 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 closest image. This experiment was done as a sanity check of the proposed method. A $4\times$ kernel was estimated from the observed image pair. The results are shown in Figure 3 and 4.

The procedure was tested for both automatic SIFT-based registration and the ideal (known) alignment. Although both estimates are significantly accurate, the automatic registration introduces a small mis-alignment, as shown in the difference images. See caption for details.

5.2 Real camera examples

The behavior of the proposed approach was tested for several different image pairs and for super-resolution estimations 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.

Two-scale vs. non-blind target-based method In [14] we proposed a non-blind method that uses a realization of white noise as calibration pattern. It was proved that, up to now, this method is the one estimates the PSF with highest accuracy. Therefore, the PSF resulting from this method will be used here as ground truth. Figure 6 shows the $4\times$ PSF estimated by the proposed two-scale 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

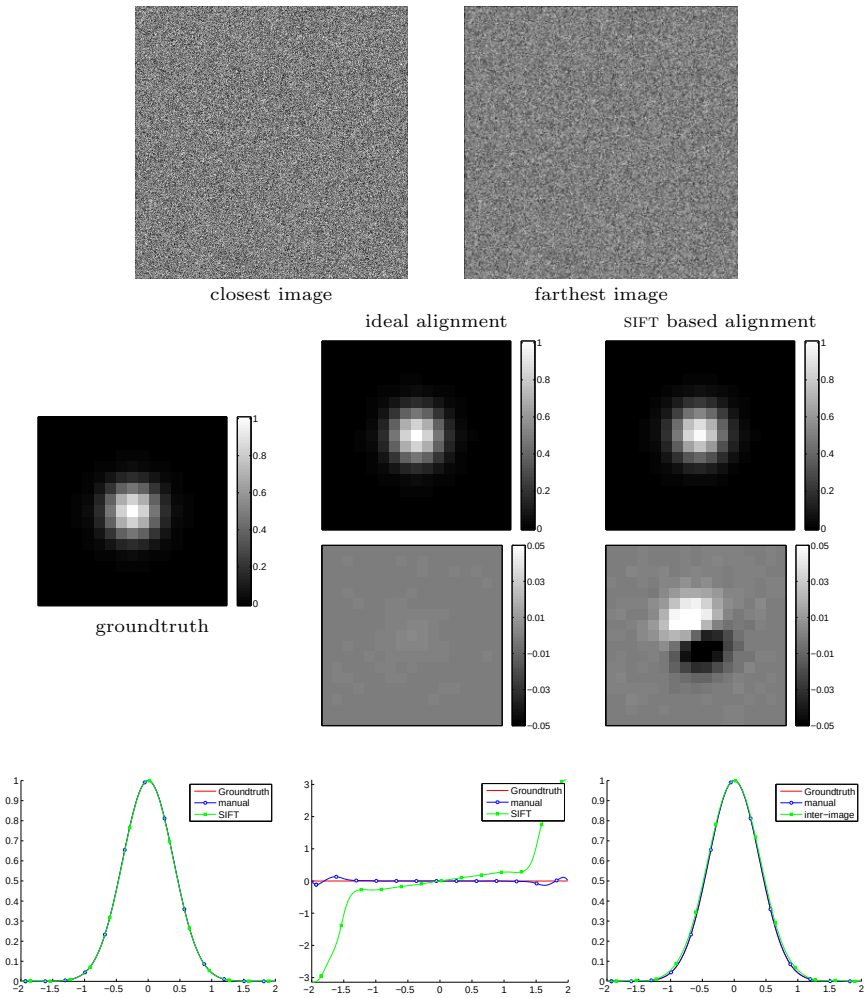
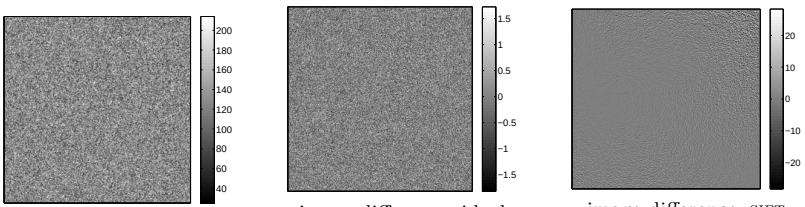


Figure 3: Synthetic example: 4× PSF estimation for simulated data. Top row: the two views closest / farthest images. Middle row: the simulated PSF (ground truth) and the respective PSF estimations using the automatic SIFT points / thin-plates alignment and the ideal alignment. Both estimations are accurate. However, as shown in the difference images the automatic registration introduces a small mis-alignment. This can also be seen in the phase and modulus of the PSF Fourier Transform vertical profile, shown in the bottom row. Bottom row, on the right: comparison of the inter-image and PSF kernels. Since both input images are simulated at distances in a ratio of $\lambda = 4\times$, h is very close to k.

the Bayer matrix), with the camera aperture set to f/5.7. The estimated PSF is quite close to the one obtained by using our random target based approach. In



farthest image | image difference, ideal alignment | image difference, SIFT alignment

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

particular their sizes are similar and their corresponding MTFs present zeros at the same locations.

farthest snapshot | closest snapshot

Figure 5: Wall image: two-scale vs. white noise target-based estimation. Two distant, parallel views of a textured wall.

Color filter array estimations Two pictures of 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× 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 and layout. In fact, each sensor active zone is usually L-shaped, and the red and blue sensors are rotated 90° w.r.t. the green ones (see for example [36]). These rotations are consistently observed in the PSFs and MTFs estimated with our two-scale method. This clearly illustrates

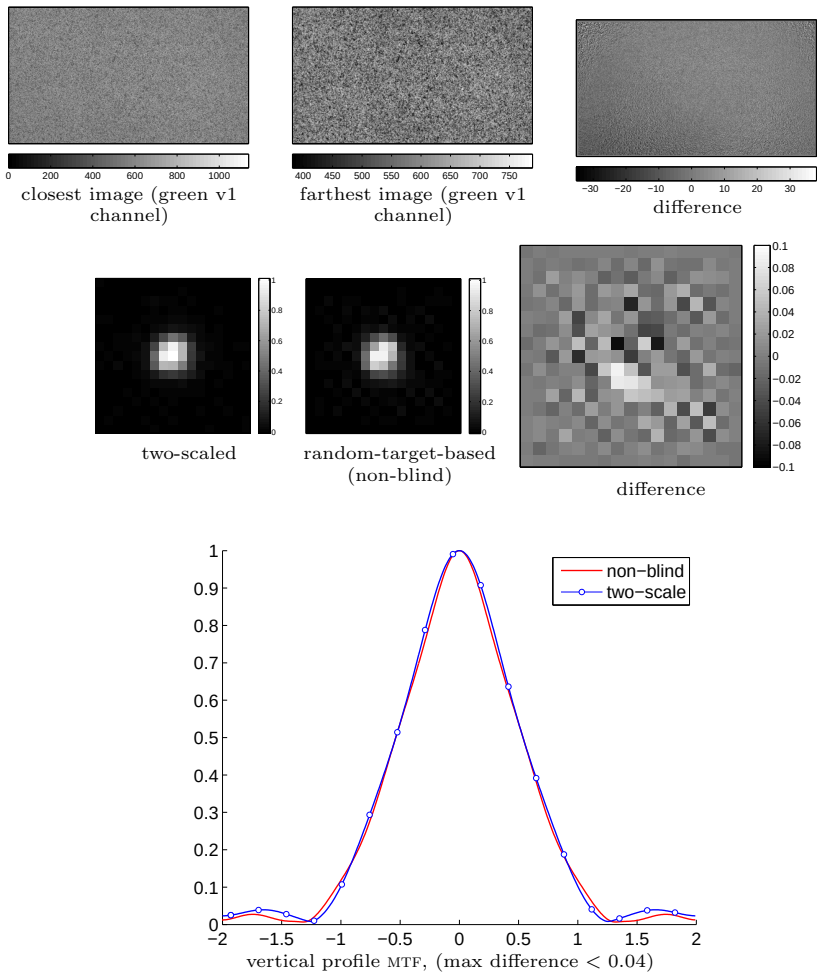


Figure 6: Wall image: two-scale vs. white noise target-based estimation. Estimation at 4× PSF resolution for one of the green channels from the camera RAW output. Top row: two distant, parallel views of a textured wall. Middle row: the PSF estimated with the proposed algorithm and the one estimated using the random target method. Bottom row: vertical profile of the MTF. Both estimations are close. In particular the associated airy disks have similar sizes and the MTFs vanish in the same locations.

the accuracy of the proposed approach.

Different kinds of scenes The wall images in the previous experiments are well adapted for our two-scale PSF estimation method, since their spectra show slow decay. A priori one would think that images from pure white noise would

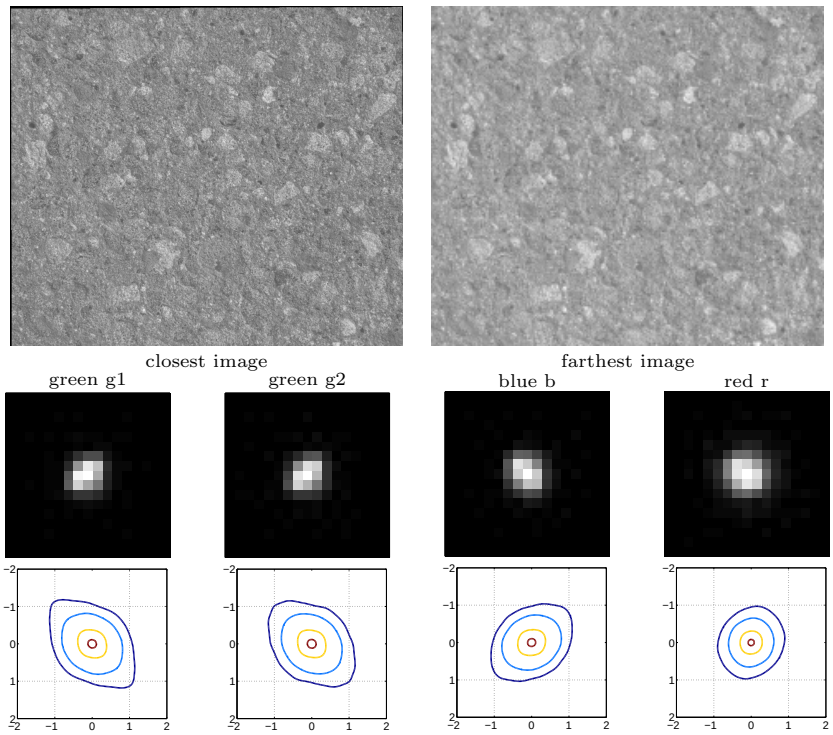


Figure 7: *Different color channels*: PSF estimation at 4× resolution for the four Bayer channels (two greens, red and blue). Top row: two distant, parallel views of a concrete wall. Middle row: the 4× PSF estimated for the four channels. Bottom row: their corresponding Fourier spectrum modulus. The estimation was performed with images captured at aperture f/5.7. The red PSF is larger than the blue and the green ones. This is consistent with the diffraction phenomenon: the red wavelengths red are larger than the rest, thus its diffraction kernel is wider. Also notice the differences between the shape of the red/blue and green PSF spectra (bottom row). Red and blue MTFs are rotated 90° with respect to the green ones. This symmetric behavior is consistent with the layout of L-shaped sensors [36].

yield to better estimates, since this is what happens in our previous target-based approach [14]. But for our two-scale approach, this would be true if both snapshots could be precisely aligned, which is not the case in practice. Indeed, SIFT descriptors are not stable in the presence of aliasing. 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 an appropriate trade-off.

Figure 8 shows two snapshots of a photograph in a magazine, with the corresponding 1× to 4× PSF estimations for the first green channel. The estimation

was performed at the image center for the camera working at f/5.7 aperture. All the subpixel estimations are consistent: their MTFs exhibit good overlap in common regions. While these newspaper images produce accurate SIFT points, their spectra decay faster than those of the wall images. Consequently, the high frequencies in the PSF estimate are noisier. This can be readily seen by comparing both estimates at $4\times$ resolution.

What kind of textures should be used? It follows from the previous analysis that, in order to simultaneously produce good SIFT points and a sufficiently slow frequency decay, textures composed by elements with different sizes are to be preferred. 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 either.

Comparison to other methods In this experiment we compare the performance of the two-scale method proposed here, with three state-of-the-art non-blind methods: Joshi et al.'s [20, 19], Imatest commercial software [22] and our previous random target method [14]. All the estimates were computed at the image center, with aperture f/5.7. For the two-scale approach, we used the wall image pair shown in Figure 7. Joshi et al.'s and Imatest use two different kinds of slanted-edge calibration patterns. The algorithm by Joshi requires to set a regularization parameter; we show the results obtained for three different levels of regularization.

Figure 10 shows the MTF profiles of the obtained PSF estimates. The proposed two-scale method performs at least as well as the non-blind methods under comparison. Joshi et. al. method shows similar performance for a carefully, manually chosen regularization parameter. See caption for details.

6 Conclusion

In this work we presented an algorithm for the subpixel 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 scales, and then estimating the kernel blur between them.

The estimation method is regularization-free. In that sense, the technique is closely related to our recent non-blind estimation method [14], which uses 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 proposed two-scale

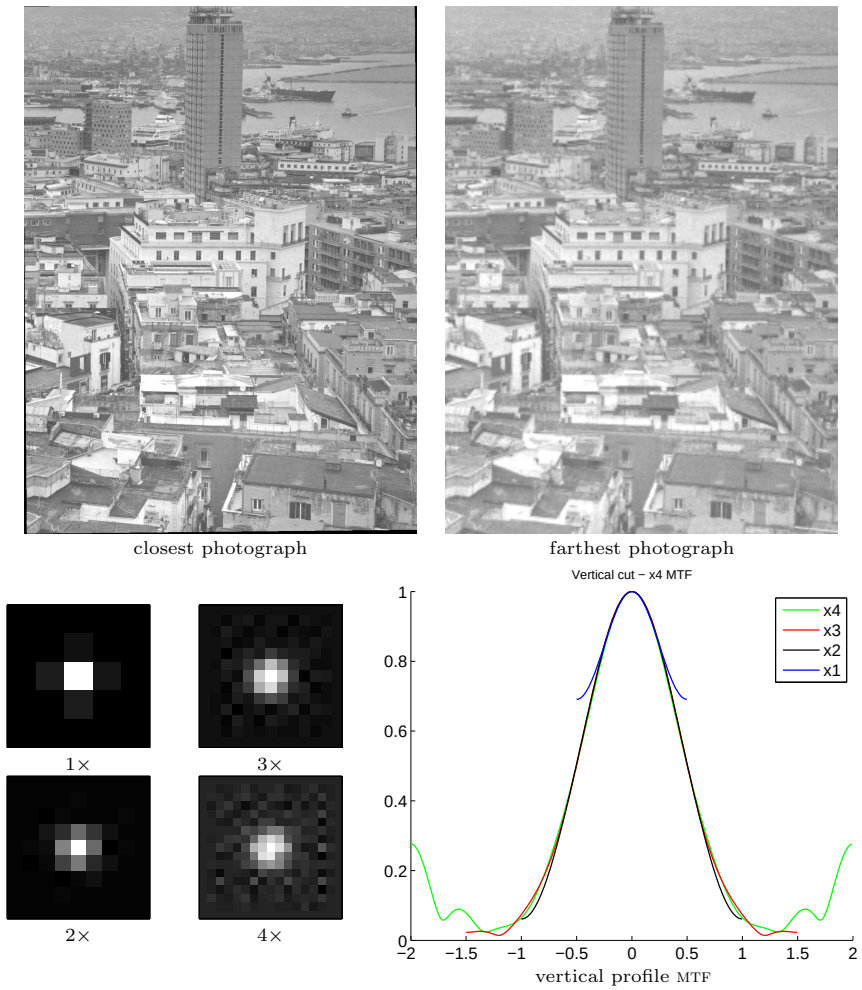


Figure 8: Magazine image: ×1, ×2, ×3 and ×4 estimations for the first green channel from a pair of photographs of a newspaper image. The estimation was done at the image center for the camera working at f/5.7 aperture. All the estimations are consistent: their MTFs show good overlap. The 4× PSF estimation is noisier than the one produced from the wall images. The main reason is that the spectrum of the magazine image decays faster.

method the question is far more intricate because only the blur between the acquisitions can be estimated. Thus a mathematical analysis and new algorithms have been introduced proving how the PSF can be recovered from the inter-image kernel.

To reach high accuracy, images of textured scenes with flat enough spectrum are preferred. It was experimentally verified that many textures found in nature

Figure 9: Examples of textures which are not adapted to the two-scale approach. 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.

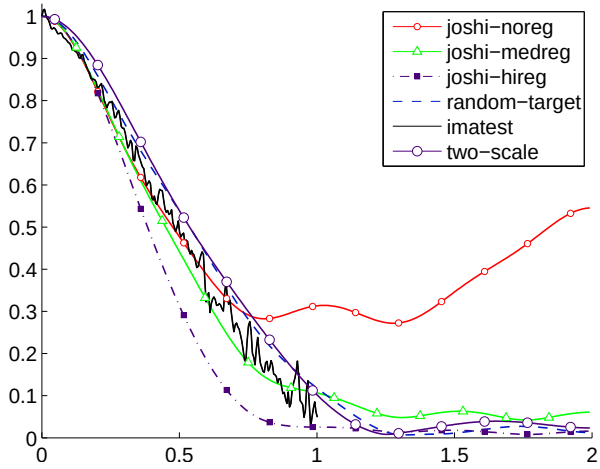


Figure 10: *Comparison of* PSF/MTF *estimation methods* Our implementation of Joshi et al. PSF estimation algorithm [20, 19], Imatest commercial software [22], our previous random target method [14], and the two-scale method proposed in this work (applied to the images of the wall shown in Fig. 7). On the low frequencies all algorithms produced very similar estimates, 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 two-scale algorithm is the one presenting closest estimation to the non-blind estimation from [14], considered as ground-truth in virtue of its high accuracy.

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 state of the art methods, with the advantage of not requiring any special acquisition setup or calibration pattern, 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.

A Mathematical framework and physical modeling

Functional spaces and other notation

- $\mathbb{R}^2$ 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 Square 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 $L^1 * L^2 \subset L^2$, $L^1 * L^1 \subset L^1$, $L^2 * L^2 \subset C^0$, $\mathcal{E}' * \mathcal{S}' \subset \mathcal{S}'$.

- We 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.

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. [33, 5]).

- *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.

- 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 $\mathbf{u} := \sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{u}(\mathbf{k})\delta_{\mathbf{k}}$.

- The operator $\mathbf{I}_1 : \ell^2(\mathbb{Z}^2) \to \mathcal{BL}^2(\mathbb{R}^2)$ denotes the *Shannon-Whittaker interpolator*, defined by $\mathbf{I}_1\mathbf{u}(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{u}(\mathbf{k})\mathrm{sinc}(\mathbf{x} - \mathbf{k})$, where $\mathrm{sinc}(\mathbf{x}) = \frac{\sin(\pi x)}{\pi x}\frac{\sin(\pi y)}{\pi y}$. 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 $\mathbf{u} \in \ell^2$, $\mathbf{F}(\mathbf{I}_1\mathbf{u})$ 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. Note that if $u \in L^1 \cup L^2$ then $W_1 u$ is in $\mathcal{BL}^2$.

- $H_\lambda u(\mathbf{x}) = \lambda^2 u(\lambda\mathbf{x})$ is the continuous homothecy (*i.e.* $\lambda > 1$ is a contraction); 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 homothecy 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 contraction 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 $(\mathbf{L}^*\mathbf{L})\mathbf{L}^+ := \mathbf{L}^*$.

A.1 Physical and mathematical modeling of continuous images

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 [26] 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 the camera PSF h which will be regular enough for $h * O$ to be in $\mathcal{BL}_0^2$.

How regular can it be realistically assumed to be? The kernel h originates in several physical phenomena from diffraction, through antialiasing filtering 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 $h \in \mathcal{BL}_0^2 \cap L^1$.

We now turn to the problem of simplifying O to a more manageable function u, which is indistinguishable from O after convolution with the PSF h. Let $B = \mathrm{supp}(\hat{h})$ 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 $\hat{h}_0$ can easily constructed by an explicit formula as a C^∞ and compactly supported function satisfying $\hat{h}_0(\eta) = 1$ on B. Then its inverse Fourier transform has all required properties.

So we have

$$v = h * O = h * u, \quad \text{where } u = h_0 * O.$$

In consequence, the observed landscape can be assumed without loss of generality to be $u = h_0 * O$ instead of O. Being the convolution of a compactly supported positive Radon measure $O \in \mathcal{E}'$ with $h_0 \in \mathcal{BL}_0^2 \cap L^1$, u also belongs to $\mathcal{BL}_0^2$, and its convolution with $h \in \mathcal{BL}_0^2 \cap L^1$ is the observed image $v \in \mathcal{BL}_0^2$.

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.* [33, 17] for the proofs in the particular setting chosen here.

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

$$\mathbf{F}(f * g) = \mathbf{F}(f) \cdot \mathbf{F}(g) \tag{12}$$

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 [33, Theorem 2.6]).*
2. *$g \in \mathcal{E}'$ and $f \in \mathcal{S}'$. Then $f * g$ belongs to $\mathcal{S}'$ (see [17, Theorem 7.1.15]).*

Applying the Fourier transform on both sides of Equation (12) 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*

$$\begin{aligned} \mathbf{F}(f \cdot g) &= \frac{1}{(2\pi)^2}\mathbf{F}(f) * \mathbf{F}(g) \\ \mathbf{F}^{-1}(f \cdot g) &= \mathbf{F}^{-1}(f) * \mathbf{F}^{-1}(g) \end{aligned} \tag{13}$$

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 [17])**.**

$$\hat{\Pi}_1 = (2\pi)^2 \Pi_{2\pi}. \tag{14}$$

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

$$\mathbf{F}(\Pi_1 \cdot u) = \Pi_{2\pi} * \hat{u}. \tag{15}$$

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 (14).

□

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

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

$$u = \mathbf{I}_1\mathbf{S}_1 u. \tag{16}$$

Proof. We can apply Lemma 2

Multiplying both sides of Equation (15) 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 for $\mathbf{k} = 0$ 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*

$$u = H_s \mathbf{I}_1 \mathbf{S}_1 H_{\frac{1}{s}} u \tag{17}$$

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

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

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

For the uniqueness of the inter-image kernel we shall additionally assume that $\hat{u}$ 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 (15) 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:*

$$W_1(h * v) = W_1 h * v = h * W_1 v \tag{18}$$

$$W_1 H_\lambda v = H_\lambda W_{\frac{1}{\lambda}} v \tag{19}$$

$$H_\alpha(u * v) = H_\alpha u * H_\alpha v \tag{20}$$

Proof. The proof of (18).
In Fourier,

$$\mathbf{F}(W_1(h * v)) \overset{\text{def}}{=} \mathbf{F}(h * v) \cdot \mathbf{1}_{[-\pi,\pi]^2} \overset{(12)}{=} \mathbf{F}(h) \cdot \mathbf{F}(v) \cdot \mathbf{1}_{[-\pi,\pi]^2}$$

Thus,

$$\mathbf{F}(h) \cdot \mathbf{F}(v) \cdot \mathbf{1}_{[-\pi,\pi]^2} = \mathbf{F}(h) \cdot \mathbf{1}_{[-\pi,\pi]^2} \cdot \mathbf{F}(v) \cdot \mathbf{1}_{[-\pi,\pi]^2}$$

and all results are deduced from this last statement.

□

Proof. The proof of (19).

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

Thus

$$\mathbf{F}(W_1 H_\lambda v) \overset{(12)}{=} \mathbf{F}(H_\lambda v)\cdot \mathbf{1}_{[-\pi,\pi]^2} = \lambda^2 H_{\frac{1}{\lambda}}\hat{v}\cdot \mathbf{1}_{[-\pi,\pi]^2};$$

On the other hand,

$$\begin{aligned}\mathbf{F}(H_\lambda W_{\frac{1}{\lambda}} u) &= \lambda^2 H_{\frac{1}{\lambda}}\mathbf{F}(W_{\frac{1}{\lambda}} v)\\ &\overset{(12)}{=} \lambda^2 H_{\frac{1}{\lambda}}(\hat{u}\cdot \mathbf{1}_{[-\frac{\pi}{\lambda},\frac{\pi}{\lambda}]})\\ &= \lambda^2 (H_{\frac{1}{\lambda}}\hat{v})\cdot \mathbf{1}_{[-\pi,\pi]^2}.\end{aligned}$$

□

Proof. The proof of (20). 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 $u, v \in \mathcal{BL}_0^2(\mathbb{R}^2)$. If either u or v is band-limited then*

$$\mathbf{S}_1(u * v) = \mathbf{S}_1\bar{u} * \mathbf{S}_1\bar{v}, \tag{21}$$

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 $u \in \mathcal{BL}_0^2$ and $h \in L^1 \cap \mathcal{BL}_0^2$, band-limited in $[-s\pi, s\pi]^2$. Then*

$$\mathbf{S}_1(u * h) = \mathcal{S}_s(\bar{\mathbf{u}} * \mathbf{h}), \tag{22}$$

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.

$$\begin{aligned}
\mathbf{S}_1(u*h) &= \mathbf{S}_1 H_s H_{\frac{1}{s}}(u*h) \\
&\overset{(18)}{=} \mathbf{S}_1 H_s H_{\frac{1}{s}}(W_s u * h) \\
&\overset{(20)}{=} \mathbf{S}_1 H_s (H_{\frac{1}{s}} W_s u * H_{\frac{1}{s}} h) \\
&\overset{(19)}{=} \mathbf{S}_1 H_s (W_1 H_{\frac{1}{s}} u * H_{\frac{1}{s}} h) \\
&\overset{(16)}{=} \mathbf{S}_1 H_s \mathbf{I}_1 \mathbf{S}_1 (W_1 H_{\frac{1}{s}} u * H_{\frac{1}{s}} h) \\
&\overset{(21)}{=} \mathbf{S}_1 H_s \mathbf{I}_1 (\mathbf{S}_1 W_1 H_{\frac{1}{s}} u * \mathbf{S}_1 H_{\frac{1}{s}} h) \\
&\overset{\text{def}}{=} \mathbf{S}_1 H_s \mathbf{I}_1 (\bar{\mathbf{u}} * \mathbf{h}) \\
&\overset{\text{def}}{=} \mathcal{S}_s (\bar{\mathbf{u}} * \mathbf{h}).
\end{aligned}$$

First notice that as $u \in \mathcal{BL}^2_0$ and $h \in L^1$ are in we can apply (18) (20) directly. As $W_1 u$ is in $\mathcal{BL}^2$ we can apply (19). Nyquist theorem (16) is valid since $u \in L^2$ and $h \in L^1$ 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 (21). 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 $h \in L^1 \cap \mathcal{BL}^2_0$ 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 $h \in L^1$ then $\hat{h} \in C^0$ 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 an injective bounded linear operator (*IBLO*) defined on a Banach space* X *and* $\Delta\mathbf{A}$ *a perturbation of* $\mathbf{A}$ *such that* $\mathbf{A}+\Delta\mathbf{A}$ *is also* IBLO *and* $\|\mathbf{A}\|\|\Delta\mathbf{A}\| < 1$*. Let* $\mathbf{b} \in X$ *and* $\Delta\mathbf{b}$ *be 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:*

$$\frac{\|\mathbf{x}^*-\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\mathrm{cond}(\mathbf{A})}{1-\|\mathbf{A}^+\Delta\mathbf{A}\|}\left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}+\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}\right). \tag{23}$$

where $\mathrm{cond}(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^+\|$.

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] S.M. Backman, A.J. Makynen, T.T. Kolehmainen, and K.M. Ojala, *Fast lens testing using random targets*, Opto-Ireland 2002: Optics and Photonics Technologies and Applications, 4876 (2003), pp. 1100–1109.

[3] SM. Backman, AJ. Makynen, TT. Kolehmainen, and KM. Ojala, *Random target method for fast MTF inspection.*, Optics Express, 12 (2004).

[4] EH. Barney Smith, *PSF estimation by gradient descent fit to the ESF*, Image Quality and System Performance III, 6059 (2006), p. 60590E.

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

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

[7] J. Brauers and T. 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.

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

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

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

[11] 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.

[12] 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.

[13] M. Delbracio, A. Almansa, and P. Musé, *Blind subpixel point spread function estimation from scaled image pairs.* `http://www.ipol.im/pub/algo/damm_blind_psf_estimation_from_scaled_image_pairs/`, accessed February 9, 2012.

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

[15] JW. Goodman, *Introduction To Fourier Optics*, McGraw-Hill Science/Engineering/Math, 1996.

[16] G. Healey and R. Kondepudy, *Radiometric CCD camera calibration and noise estimation*, IEEE Trans. Pattern Anal. Mach. Intell., 16 (1994), pp. 267–276.

[17] L. 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.

[18] ISO, *12233:2000 photography - electronic still-picture cameras - resolution measurements*, (2000), p. 32.

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

[20] N. Joshi, R. Szeliski, and DJ. Kriegman, *PSF estimation using sharp edge prediction*, Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, 0 (2008), pp. 1–8.

[21] 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.

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

[23] 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.

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

[25] A. Marion, *Acquisition et Visualisation des Images*, Eyrolles, 1997.

[26] 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.

[27] JM. Morel and G. Yu, *ASIFT: A new framework for fully affine invariant image comparison*, SIAM J. Img. Sci., 2 (2009), pp. 438–469.

[28] 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.

[29] 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.

[30] SE. Reichenbach and R. Park, SK. Narayanswamy, *Characterizing digital image acquisition devices*, Optical Engineering, 30 (1991), pp. 170–177.

[31] D. Robinson and P. Milanfar, *Statistical performance analysis of super-resolution*, Image Processing, IEEE Transactions on, 15 (2006), pp. 1413–1428.

[32] 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.

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

[34] H. Tian, B. Fowler, and AE. Gamal, *Analysis of temporal noise in CMOS photodiode active pixel sensor*, IEEE Journal of Solid-state Circuits, 36 (2001), pp. 92–101.

[35] CS. Williams and OA. Becklund, *Introduction to the Optical Transfer Function (SPIE Press Monograph Vol. PM112)*, SPIE Publications, 2002.

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

[37] G. Yu and JM. Morel, *ASIFT: An Algorithm for Fully Affine Invariant Comparison*, Image Processing On Line, (2011). DOI:`http://dx.doi.org/10.5201/ipol.2011.my-asift` 10.5201/ipol.2011.my-asift.

[38] 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.

[39] T. Zhao, R. Wang, Y. Liu, and F. Yu, *Characteristic-analysis of optical low pass filter used in digital camera*, vol. 6034, SPIE, 2006, p. 60340N.