

Blur-Kernel Estimation from Spectral Irregularities

Amit Goldstein and Raanan Fattal

Hebrew University of Jerusalem

<http://www.cs.huji.ac.il/~raananf/projects/deblur>

Abstract. We describe a new method for recovering the blur kernel in motion-blurred images based on statistical irregularities their power spectrum exhibits. This is achieved by a power-law that refines the one traditionally used for describing natural images. The new model better accounts for biases arising from the presence of large and strong edges in the image. We use this model together with an accurate spectral whitening formula to estimate the power spectrum of the blur. The blur kernel is then recovered using a phase retrieval algorithm with improved convergence and disambiguation capabilities. Unlike many existing methods, the new approach does not perform a maximum a posteriori estimation, which involves repeated reconstructions of the latent image, and hence offers attractive running times.

We compare the new method with state-of-the-art methods and report various advantages, both in terms of efficiency and accuracy.

1 Introduction

In many practical scenarios, such as hand-held cameras or ones mounted on a moving vehicle, it is difficult to eliminate camera shake. Sensor movement during exposure leads to unwanted blur in the acquired image. Under the assumption of white noise and spatially-invariant blur across the sensor this process is modeled by

$$B(\mathbf{x}) = (I * k)(\mathbf{x}) + \eta(\mathbf{x}) \quad (1)$$

where $*$ denotes the convolution operation, B is the acquired blurry image, k is the unknown blur kernel and $\eta(\mathbf{x})$ is a zero-mean, identically- and independently-distributed noise term at every pixel $\mathbf{x} = (x, y)$. Blind image deconvolution, the task of removing the blur when the camera motion is unknown, is a mathematically ill-posed problem since the observed image B does not provide enough constraints for determining both I and k . Most deblurring techniques therefore introduce additional constraints over I and k . The most common framework for incorporating such prior knowledge is through maximum a posteriori (MAP) estimation. Norms favoring sparse derivatives are often used to describe I as a natural image [1–6]. While not failure-free, this approach was shown to recover very complex blur kernels and achieve impressive deblurred results. However, the maximization of these estimators is a time consuming task involving the computation of the latent image multiple times.

An alternative approach to the problem, which did not receive as much attention, extracts the blur kernel k *directly* from the blurry image B *without* computing I in the process. The basic idea is to recover k from the anomalies that B shows with respect to the canonical behavior of natural images. Yitzhaky et al. [7] compute the 1D autocorrelation of the derivative of B along the sensor movement direction. Normally, image derivatives are weakly correlated and hence this function should be close to a delta function. The deviation from this function provides an estimate for the power spectrum (PS) of the kernel, $|\hat{k}(\omega)|^2$. Hu et al. [8] also follow this strategy for recovering the PS of two-dimensional kernels and use the eight-point Laplacian for whitening the image spectrum. The blur kernel is then computed using a phase retrieval technique [9] that estimates the phase by imposing spatial non-negativity and compactness. This approach consists of evaluating basic statistics from the input B and, unlike methods that use MAP estimation, does not involve repeated reconstructions of I . While this makes it favorable in terms of computational-cost, the true potential of this approach in terms of accuracy was not fully explored.

In this paper we describe a new method for recovering the blur from irregularities in the statistics of motion-blurred images. We derive and use a power-law model to describe the PS of natural images, which refines the traditional $\|\omega\|^{-2}$ law. This new model better accounts for biases arising from the presence of large and strong edges in natural images. We use this model, as well as a more accurate spectral whitening formula, to recover the PS of the blur kernel, $|\hat{k}(\omega)|^2$, in a more robust and accurate manner compared to previous methods following this approach. We also describe a modified phase retrieval algorithm with improved convergence and ability to resolve ambiguities.

Unlike methods that rely on the presence and the identification of well-separated edges in the image, our purely statistical approach copes well with images containing under-resolved texture and foliage clutter, which are abundant in outdoor scenes. Similarly to [7, 8] we do not reconstruct the latent image repeatedly and access the input image only once to extract a small set of statistics. Thus, the core of our technique depends only on the blur kernel size and does not scale with the image dimensions.

The experiments we report show that our method is capable of achieving highly-accurate results, in various scenarios that challenge other approaches, and compares well with the state-of-the-art. Our CPU implementation achieves these results at running-times comparable to the fast method of Cho and Lee [10] and considerably faster than MAP-based methods.

2 Previous Work

Much effort was put into removal of image blur due to camera motion. Blind-deconvolution methods that recover the blur kernel k and the sharp image I rely on various regularities natural images exhibit. The most-dominant approach for tackling this problem, in the context of spatially-uniform blur kernel, is to formulate and solve a MAP problem. This requires the minimization of a log-likelihood

term that accounts for Eq. (1) plus additional prior terms over the image I and kernel k . Lagendijk et al. [11] use an autoregressive Gaussian prior for $I(\mathbf{x})$ and You and Kaveh [12] use a similar Gaussian prior over high-frequencies (derivatives) of I . Both priors are blind to the phase content of k and are not sufficient for recovering it. Hence, Lagendijk et al. further assume that the blur is symmetric (zero phase) and You and Kaveh incorporate adaptive spatial weighting which breaks this symmetry. Chan and Wong [1] replace the Gaussian image prior with a Laplace distribution defined by the L_1 norm over the image derivatives. This choice is more consistent with the heavy-tailed derivative distribution observed in natural images [13]. Levin et al. [4] show that this prior is not sufficient for uniqueness and may result in degenerate delta kernels. Indeed, methods that rely on sparse norms often introduce additional constraints such as smoothness of the blur-kernel [1] and two motion-blurred images [14] or use alternative image priors such as spatially-varying priors [3] and ones that marginalize over all possible sharp images $I(\mathbf{x})$ [2, 15]. To avoid the cost in computing marginalized distributions, Krishnan et al. [16] propose a normalized prior that favors sharp images over blurry ones. Levin et al. [6] reduce the marginalization cost by approximating the prior distribution via its mean and covariance. Gupta et al. [17] and Whyte et al. [18] use the MAP formulation for non-uniform blur model that accounts for more general camera motion.

Another line of works make a more direct use of the presence of sharp edges in images for recovering the blur. Joshi et al. [19] detect salient edges in the blurry image B and predict the underlying unblurred edges as ideal step edges. This correspondence allows recovering a one-dimensional slice or projection of k . In order to recover the complete kernel, this approach requires the availability and successful detection of strong edges at a variety of orientations. The explicit detection of edges limits this approach to unimodal blur kernels. Cho and Lee [10] reconstruct the sharp edges using inverse-diffusion shock filtering and recover the kernel efficiently in the gradient domain using fast Fourier transform. Xu and Jia [20] propose a two-phase scheme that computes a coarse estimation of the kernel from reconstructed sharp edges and then refine it iteratively. Cho et al. [21] also detect edges and use the Radon transform to reconstruct the blur kernel, given the projections of the blurry edges found. In [22] edges are modeled by a two-color model which is tailored using local clustering. The sharp image is then recovered by minimizing the color mixture. Jia [23] also relies on color mixtures and extracts the blur kernel of moving objects given their boundary alpha map.

Lastly, we would like to mention recent techniques that use dedicated hardware designed to acquire blur-free images. Ben-Ezra and Nayar [24] construct a hybrid camera that measures its own motion during light integration using another fast-shutter low-resolution camera. Raskar et al. [25] rapidly open and close the shutter during exposure and thereby shape the blur kernel to pass more high-frequency spatial details. Levin et al. [26] design a coded aperture which is easier to invert and allows a better discrimination between the blurs resulting from depths. Yuan et al. [27] use additional noisy yet unblurred short-exposure image in order to recover the blur kernel. Levin et al. [28] show that in images

captured by parabolic integration the blur of objects moving at different speeds can be eliminated using a single kernel. In [29] a combination of gyroscopes and accelerometers is used to estimate the camera acceleration and angular velocity during exposure.

3 Background

Field [30] and Burton and Moorhead [31] point out the following power-law describing the power spectra of images of natural scenes

$$|\hat{I}(\boldsymbol{\omega})|^2 \propto \|\boldsymbol{\omega}\|^{-\beta}, \tag{2}$$

for $\boldsymbol{\omega} = (\omega_x, \omega_y) \neq (0, 0)$, where I is a natural image, \hat{I} its Fourier transform and $\boldsymbol{\omega}$ denotes the frequency coordinates. Various studies suggest that $\beta \approx 2$, and explain this behavior by the presence of edges and other fractal-like features such as foliage and landscape. See [32] for a recent survey on the topic.

Compact second-order derivative filters, for example $l_x = [-1, 2, -1]$, provide a good approximation for $|\omega_x|^2$ especially at lower frequencies since

$$\hat{l}_x(\omega_x) = 2(\cos(\omega_x\pi) - 1) \approx \pi^2\omega_x^2 + \mathcal{O}(\omega_x^4), \tag{3}$$

where the last equality is based on Taylor approximation of cosine around $\omega_x = 0$. The four-point Laplacian filter¹ is given by the sum $l = \bar{d}_x * d_x + \bar{d}_y * d_y$, where $d_x = [1, -1]$ and $d_y = d_x^\top$. Therefore $\hat{l}(\boldsymbol{\omega}) \approx \pi^2(\omega_x^2 + \omega_y^2) \propto \|\boldsymbol{\omega}\|^2$. We use the bar sign $\bar{\cdot}$ to denote signal mirroring, i.e., $\bar{d}(\mathbf{x}) = d(-\mathbf{x})$.

Let us denote by d the ‘square-root’ filter of l which is defined as the symmetric filter giving $l = \bar{d} * d$. This factorization exists for every symmetric filter with non-negative frequency response, i.e., $\hat{l}(\boldsymbol{\omega}) \geq 0$. Filtering an image obeying Eq. (2) with d results in a signal with whitened spectrum, since

$$|\widehat{(I * d)}(\boldsymbol{\omega})|^2 = |\hat{I}(\boldsymbol{\omega})|^2 \cdot |\hat{d}(\boldsymbol{\omega})|^2 = |\hat{I}(\boldsymbol{\omega})|^2 \cdot \hat{l}(\boldsymbol{\omega}) \approx \pi^2 \|\boldsymbol{\omega}\|^{-2} \|\boldsymbol{\omega}\|^2 = c, \tag{4}$$

for some constant c , where the equality $|\hat{d}(\boldsymbol{\omega})|^2 = \hat{l}(\boldsymbol{\omega})$ is the Fourier analog of $l = \bar{d} * d$. Thus, in the case of a blurry image $B = I * k$, this procedure can be used to obtain the following estimate of the blur-kernel PS,

$$|\widehat{(B * d)}(\boldsymbol{\omega})|^2 = |\hat{I}(\boldsymbol{\omega})|^2 \cdot \hat{l}(\boldsymbol{\omega}) \cdot |\hat{k}(\boldsymbol{\omega})|^2 \approx c |\hat{k}(\boldsymbol{\omega})|^2. \tag{5}$$

The Wiener–Khinchin theorem relates the PS of any signal J to its autocorrelation by

$$\hat{R}_J(\boldsymbol{\omega}) = |\hat{J}(\boldsymbol{\omega})|^2, \tag{6}$$

¹ The four-point Laplacian filter is given by $l(0, 0) = 4$ and $l(x, y) = -1$ at x and y that $|x| + |y| = 1$ and zero otherwise. The eight-point Laplacian is given by $l(0, 0) = 8$ and $l(x, y) = -1$ when $|x|, |y| \leq 1$ ($|x| + |y| > 0$) and zero otherwise.

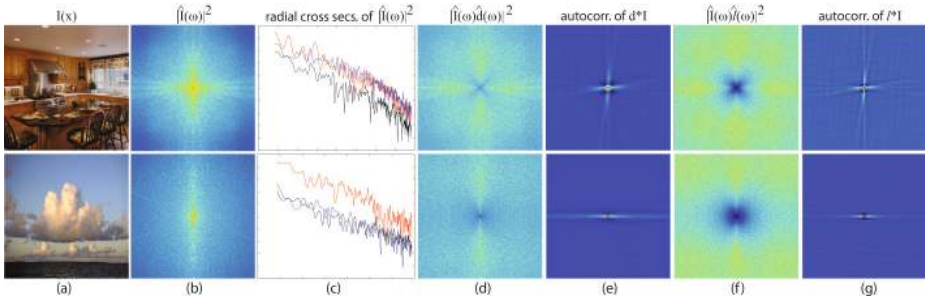


Fig. 1. Shows (a) non-blurred natural image, (b) its PS which has radial streaks of increased magnitudes, (c) 1D log-log plots of the PS along its (red) vertical, (blue) horizontal and (black) diagonal cross-sections. These plots differ by roughly an additive constant which corresponds to a multiplicative scalar in terms of the PS. (d) shows the PS of the image whitened with the Laplacian root filter d and (e) its autocorrelation in real-space, followed by (f,g) the same functions of the image whitened using the Laplacian filter itself (as used by [8]). The power spectra shown in this figure were mapped using logarithm for better display.

where the autocorrelation (autocovariance) is defined by $R_J(\mathbf{x}) = (\bar{J} * J)(\mathbf{x})$. This identity introduces real-space counterparts for the spectrum whitening in Eq. (4) and the blur approximation in Eq. (5), that are given by

$$R_{I*d}(\mathbf{x}) \approx c\delta(\mathbf{x}), \quad \text{and} \quad R_{B*d}(\mathbf{x}) \approx cR_k(\mathbf{x}), \tag{7}$$

respectively, where δ is the Dirac’s delta-function, i.e., $\delta(0,0) = 1$ and zero otherwise.

Yitzhaky et al. [7] exploit this regularity in natural images and recover the power spectra of one-dimensional blur kernels $|\hat{k}(\omega)|^2$ by differentiating the image along the blur direction (which they also estimate). In order to obtain the complete kernel k , they estimate its phase using the Hilbert transform under the assumption of minimal-phase blur. Hu et al. [8] use this regularity to recover general two-dimensional kernels by whitening the image spectrum using the eight-point Laplacian filter¹. Given the estimated $|\hat{k}|^2$, they compute k by recovering its phase using the *error-reduction* phase-retrieval algorithm [9].

While Eq. (2) models well certain images, the presence of long edges, both in natural and man-made scenes, undermines its accuracy. Fig. 1 shows two such cases where the power spectra show increased magnitudes along radial streaks which are orthogonal to the strong image edges. These streaks break the rotational symmetry predicted by Eq. (2). Analogously, the autocorrelation functions of these images, after being whitened with d , or l as used by Hu et al. [8], differ considerably from the delta function, predicted in Eq. (7), along the directions of the strong edges. Clearly, such deviations from the power-law in (2) undermine the accuracy of the recovered $|\hat{k}|^2$ from Eq. (5), or equivalently R_k recovered from Eq. (7).

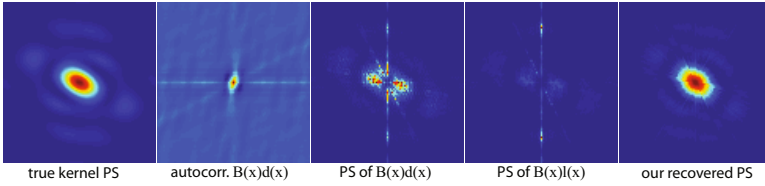


Fig. 2. Comparison between different spectral whitening procedures on an image containing strong horizontal edges

The question is, therefore, how does this discrepancy behave. To answer this, we plot the cross-sections of the PS along the most extreme orientations. The log-log plots of these 1D functions, shown in Fig. 1, can be approximately described by an additive offset which means that the PS of a natural image vary by multiplicative factors along different directions, i.e.,

$$|\hat{I}(\boldsymbol{\omega})|^2 \approx c_{\theta}(\boldsymbol{\omega}) \cdot \|\boldsymbol{\omega}\|^{-2}, \tag{8}$$

where $\theta(\boldsymbol{\omega}) = \tan^{-1}(\omega_x, \omega_y)$ is the angle of the vector $\boldsymbol{\omega}$.

In Fig. 2 we show the kernel PS estimated by whitening the image with d and l , and by using our method which relies on the refined power-law in Eq. (8). In the next section we describe an algorithm for recovering the blur kernel based on this model, i.e., solve for both $c_{\theta}(\boldsymbol{\omega})$ and the kernel phase, given autocorrelation functions computed from the input blurry image $B(\mathbf{x})$.

Finally, we would like to review the Fourier slice theorem which will become instrumental for us and is given by

$$(\widehat{P_{\theta}(J)})(\omega) = \hat{J}(\omega \mathbf{r}_{\theta}), \tag{9}$$

where $P_{\theta}(J)$ is a projection of a 2D signal into 1D by integrating it along the direction orthogonal to θ and \mathbf{r}_{θ} is a unit vector in 2D with the orientation of θ . Thus $\omega \mathbf{r}_{\theta}$ parameterizes with orientation of θ using the scalar ω . When applying this theorem in our context, where $J = |\hat{I}|^2$, we get

$$(\widehat{R_{P_{\theta}(I)}})(\omega) = |(\widehat{P_{\theta}(I)})(\omega)|^2 = |\hat{I}(\omega \mathbf{r}_{\theta})|^2 \approx c_{\theta} \cdot \|\omega\|^{-2}, \tag{10}$$

where the first equality follows from the Wiener–Khinchin theorem which applies between the 1D autocorrelation of the projected image $P_{\theta}(I)$ and its power spectrum in 1D Fourier space. The last equality follows from Eq. (8) where the restriction to a single slice leaves only a *single* unknown c_{θ} .

4 New Method

Given a blurry image $B = I * k$, the relation in Eq. (10) allows us to recover $|\hat{k}(\omega \mathbf{r}_{\theta})|^2$ up to a single scalar c_{θ} by

$$(\widehat{R_{d * P_{\theta}(B)}})(\omega) = \hat{l}(\omega) \cdot |\hat{B}(\omega \mathbf{r}_{\theta})|^2 = \hat{l}(\omega) \cdot |\hat{I}(\omega \mathbf{r}_{\theta})|^2 \cdot |\hat{k}(\omega \mathbf{r}_{\theta})|^2 \approx c_{\theta} \cdot |\hat{k}(\omega \mathbf{r}_{\theta})|^2. \tag{11}$$

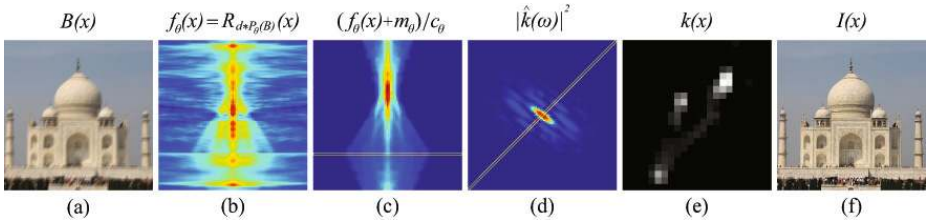


Fig. 3. Overview of the proposed method; (a) Input blurry image, (b) the 1D autocorrelation functions of the differentiated projections of the blurry image. Each row corresponds to a projection along a different θ . (c) shows the recovered correlations using the estimated c_θ and m_θ . The power spectrum recovered using the Fourier slice theorem is shown in (d), where the slice marked in yellow corresponds to the 1D Fourier transform of the marked row in (c). The kernel computed by the phase retrieval algorithm is shown in (e) and finally the deblurred image in (f). Note that our EM-like kernel recovery procedure iteratively computes (c-e).

We use $d = [3, -32, 168, -672, 0, 672, -168, 32, -3]/840$ which is a nine-point 1D differentiation filter (applied to the projected blurry image B in 1D). The resulting 1D Laplacian $l = \bar{d} * d$ has an accuracy of $\hat{l}(\omega) = \omega^2 + \mathcal{O}(\omega^8)$. This can be written in real-space, once again using the Wiener–Khinchin theorem, as

$$(R_{d*P_\theta}(B))(x) \approx c_\theta \cdot R_{P_\theta}(k)(x) \quad (12)$$

Thus, the first step of our algorithm is to compute $f_\theta(x) = R_{d*P_\theta}(B)(x)$ for every angle θ . Since we want to recover $|\hat{k}|^2$ on a grid of pixels, we pick angles θ that result in slices passing exactly through each pixel. Figure 3 shows an example of these functions. We implemented the projection operator P_θ using a nearest-neighbor sampling which achieved higher accuracy and runs in less time compared to other interpolation formulae we tested.

The resulting projected values correspond to averaging of a large number of pixels, proportional to the image dimensions n . The averaging of the independent noise terms $\eta(\mathbf{x})$ in Eq. (1) results in a noise reduction by a factor of $1/\sqrt{n}$. Therefore we ignore the noise present in $f_\theta(x)$.

Estimating $|\hat{k}(\omega)|^2$ slice-by-slice according to Eq. (12) introduces a set of unknowns c_θ , which are common to all values of $|\hat{k}|^2$ along the same angle θ . Additionally, the mean value of the projected slices $P_\theta(B)$ is lost due to the differentiation with d in Eq. (11), or equivalently $\hat{k}(0\mathbf{r}_\theta) = \hat{k}(0, 0)$ is missing in all the slices we compute. Let us denote these missing mean values by m_θ/c_θ and get altogether that $(f_\theta(x) + m_\theta)/c_\theta \approx R_{P_\theta}(k)(x)$. Therefore given $f_\theta(x)$ we need to estimate c_θ and m_θ in order to recover $R_{P_\theta}(k)(x)$.

We recover these values based on the following three kernel modeling assumptions:

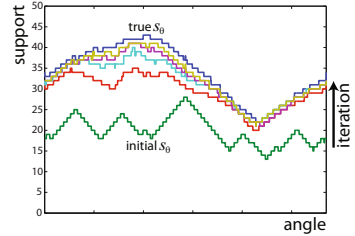
1. Camera blur kernels are proportional to the time period in which light is integrated at each camera offset as it moves during exposure. Thus, these numbers are all non-negative, i.e., $k \geq 0$, and so is its projection $P_\theta(k)$ and the 1D autocorrelation function $R_{P_\theta}(k)$.

2. Since the camera motion during exposure is finite, the blur kernel support must be compact. Similarly to the positivity case, the compact support is also inherited by $P_\theta(k)$ and $R_{P_\theta(k)}$. Thus, for each θ we have s_θ such that for every $|x| \geq s_\theta$ there is $R_{P_\theta(k)}(x) = 0$.
3. We can assume that the camera blur does not affect the total amount of light reaching the sensor, and get that $\int k(\mathbf{x})d\mathbf{x} = 1$. This remains to be the case for the projected kernel, i.e., $\int P_\theta(k)(x)dx = 1$, meaning that $\widehat{P_\theta(k)}(0) = 1$. Finally, since $|\widehat{P_\theta(k)}|^2(0) = 1$, we also get that $\int R_{P_\theta(k)}(x)dx = 1$.

Assuming we have the support extents s_θ we set $m_\theta = -f_\theta(s_\theta)$, based on the second assumption that $R_{P_\theta(k)}(s_\theta) = 0$. We then define $g_\theta(x) = f_\theta(x) + m_\theta$ for $|x| \leq s_\theta$ and zero otherwise. We make sure that $g_\theta \geq 0$ according to the first assumption above, by setting the negative values in g_θ within $[-s_\theta, s_\theta]$ to zero. Finally, we recover c_θ based on the third assumption by setting $c_\theta = \int g_\theta(x)dx$. Altogether, we expect that $g_\theta(x)/c_\theta \approx R_{P_\theta(k)}(x)$.

By repeating this procedure for all the θ we are considering, we obtain an approximation for the full 2D blur-kernel PS function $|\hat{k}(\boldsymbol{\omega})|^2$. We use this approximation to recover the blur kernel k using a phase-retrieval algorithm that we describe in Section 4.1. An actual example of this process is shown in Fig. 3. We conclude this section by explaining how the kernel support variables s_θ are determined and outlining the entire algorithm for kernel recovery.

Our algorithm consists of an iterative EM-like procedure where we switch between estimating the kernel $k(\mathbf{x})$ given the support variables s_θ and estimating s_θ from the recovered $k(\mathbf{x})$. We start this procedure with the initial guess $s_\theta = \operatorname{argmin}_x f_\theta(x)$. Then, given the retrieved kernel we update these values by setting $s_\theta = \operatorname{argmax}_x (R_{P_\theta(k)}(x) > 0.1 \cdot \max(R_{P_\theta(k)}))$.



The inset at the right shows the progression of s_θ along these iterations and Alg. 1 summarizes our method.

Algorithm 1: Iterative kernel recovery

Input: blurry image B ;
 calculate $f_\theta = R_{d^*P_\theta(B)}$;
 set $s_\theta = \operatorname{argmin}_x f_\theta(x)$;
for $i=1..N_{outer}$ **do**
 estimate $|\hat{k}|^2$ given s_θ ;
 estimate kernel using phase retrieval Alg. 2 given $|\hat{k}|^2$;
 update $s_\theta = \operatorname{argmax}_x (R_{P_\theta(k)}(x) > 0.1 \cdot \max(R_{P_\theta(k)}))$;
end
Output: recovered blur kernel k ;



Fig. 4. Comparison between the classic error-reduction phase retrieval versus our algorithm in Alg.2

4.1 Phase-Retrieval

Recovering the kernel k , given its power spectrum $|\hat{k}|^2$ requires estimating the phase component of $\hat{k}(\omega)$. There are quite a few algorithms that retrieve a signal or its phase given its power spectrum and additional constraints, such as the signal being positive and of limited spatial support. Largely speaking, these algorithms iteratively switch between Fourier and real-space domains to enforce the input PS and spatial constraints respectively.

Fienup [9] describes and compares several such algorithms, including the error-reduction algorithm used by Hu et al. [8]. According to Fienup and our own experience this algorithm tends to stagnate despite being far from a solution. Indeed, Fienup suggests an alternative strategy (that alternates between the error-reduction and another method called hybrid input-output) that avoids this problem. Luke [33] proposes a different variant, called Relaxed Averaged Alternating Reflections (RAAR), which we found to be more stable and converge more quickly.

Luke's algorithm treats the input PS as a hard-constraint and requires the recovered signal to possess this PS in full. This approach is less robust to noise or inaccuracies in the input PS and leads to oscillations between different configurations (as there might not be a real positive signal with limited support that have the exact input PS). In order to improve the performance on inaccurate input PS, we relax this requirement by treating the input PS as an attraction point instead of enforcing it (by averaging the current PS with the input). We find that this leads to a better exploration of this space and reduces the oscillations.

The input PS $|\hat{k}|^2$ and the spatial constraints may not guarantee a unique solution [34]. Moreover, and as Fienup discusses, such algorithms may converge to local minima configurations. Therefore, we repeat this procedure multiple times, starting from different random guesses for the phase component. Among the resulting solutions, we choose the one that maximizes the L_1/L_2 likelihood proposed by Krishnan et al. [16] when evaluated in a small window (100-by-100 pixels) of the image $B(\mathbf{x})$. In this step we test each kernel and its reflection to disambiguate this symmetry.

Algorithm 2: Phase retrieval algorithm

```

Input: estimated kernel magnitude spectrum,  $p(\omega) = |\hat{k}(\omega)|$ , kernel size  $s$ ;
for  $n=1 \dots N_{guesses}$  do
    // initiate the phase  $\phi(\omega)$  randomly;
    sample  $\phi(\omega)$  uniformly from  $[-\pi, \pi]$ ;
    // switch to real-space using inv. Fourier trans.;
     $g = \mathcal{F}^{-1}(p \cdot e^{i\phi})$ ;
    for  $m=1 \dots N_{inner}$  do
        // apply Fourier domain constraints;
         $g_2 = \mathcal{F}^{-1}((\alpha p + (1 - \alpha)|\hat{g}|) \cdot e^{i \cdot Phase(\hat{g})})$ ;
        // apply space domain constraints;
         $R(\mathbf{x}) = 2g_2(\mathbf{x}) - g(\mathbf{x})$ ;
         $\beta = \beta_0 + (1 - \beta_0)(1 - \exp(-m/7)^3)$ ;
         $\Omega = \{\mathbf{x} : R(\mathbf{x}) < 0\} \cup \{\mathbf{x} : \mathbf{x} \notin [0, s] \times [0, s]\}$ ;
         $g(\mathbf{x}) = \begin{cases} \beta g(\mathbf{x}) + (1 - 2\beta)g_2(\mathbf{x}) & \text{if } x \in \Omega \\ g_2(\mathbf{x}) & \text{if } x \notin \Omega \end{cases}$ ;
    end
     $\Omega = \{\mathbf{x} : g_2(\mathbf{x}) < 0\} \cup \{\mathbf{x} : \mathbf{x} \notin [0, s] \times [0, s]\}$ ;
     $k_n(\mathbf{x}) = \begin{cases} 0 & \text{if } x \in \Omega \\ g_2(\mathbf{x}) & \text{if } x \notin \Omega \end{cases}$ ;
end
Output:  $k_n$  with highest likelihood defined in [16] when deconvolving a small window;

```

We summarize our phase retrieval algorithm in Alg. 2. We used the values $\alpha = 0.95$ and $\beta_0 = 0.75$ in the code to produce the results reported in the paper. In Fig. 4 we compare between the error-reduction algorithm and our procedure.

5 Results

We implemented the phase retrieval and the projection operator P in C++ and the rest of our algorithm is implemented in Matlab. To produce the results presented here, we used three outer iterations (N_{outer}) in Alg. 1, started the phase retrieval with 30 different initialization ($N_{guesses}$) in Alg. 2 and run each 300 inner iterations N_{inner} . With this number of iterations and given $f_\theta(x)$, it takes our method about 13 seconds to recover the kernel independently of the image size. Calculating $f_\theta(x)$ involves projecting the image along various directions θ . Running this for all the angles required to reconstruct kernels of 30-by-30 pixels, takes about 2 seconds on a one-megapixel image. This number grows linearly with the image and kernel size. Running the phase retrieval from different initializations can be done in parallel. Similarly, projecting the image along each direction can be computed independently. Thus, we expect to achieve even faster running-times with a GPU implementation. We used the non-blind deconvolution of [5] to generate all the deblurred images shown here.

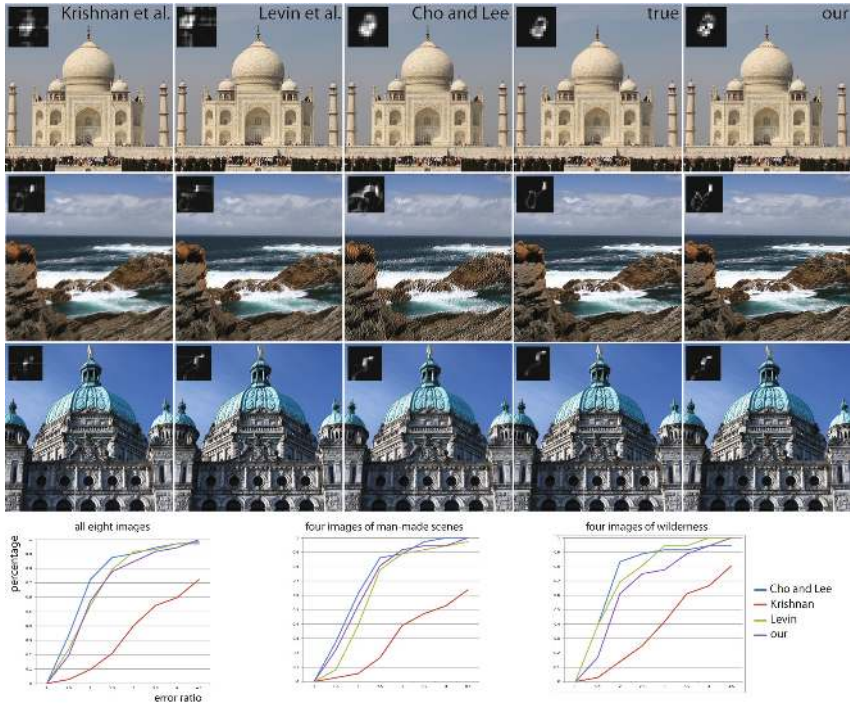


Fig. 5. Results obtained from a synthetic database of eight images and nine blur kernels. We show here three examples of recovered kernels and deblurred images. The graphs show the deconvolved-image metric introduced in [4] evaluated on these images as well as the scores obtained for the two image classes. Photos courtesy of Ryan Bushby, On The Go Tours, and Cooking in Tounques.

We compared between the methods of Cho and Lee [10], Levin et al. [6], Krishnan et al. [16] and ours over a database containing eight images, four of which contain man-made scenery and another four that do not. We blurred each of these images, according to Eq. (1) with Gaussian noise with $\sigma = 0.5\%$, using nine blur kernels that include the benchmark kernels used by [4]. We did not use the 255-by-255 gray-scale images used in [4] as this resolution does not provide reliable statistics and does not reflect nowadays real-world photos.

Fig. 5 shows a quantitative comparison between the different methods on this dataset using the deconvolved-image metric introduced in [4]. This metric consists of an error ratio given by $\|I_{out} - I_{gt}\|^2 / \|I_{k_{gt}} - I_{gt}\|^2$, where I_{out} is the image deblurred by the estimated kernel, $I_{k_{gt}}$ is the image obtained by deconvolving with the ground-truth kernel and I_{gt} is the ground-truth image. The metric indicates the percentage of images with ratios below a specified ratio.

The scores in Fig. 5 indicate that our method meets the accuracy of Levin et al. [6] with an advantage on man-made scenes. Our method achieves these results at significantly faster running times. The accuracy achieved by our method is comparable to the one Cho and Lee [10] achieve on man-made scenes. In this

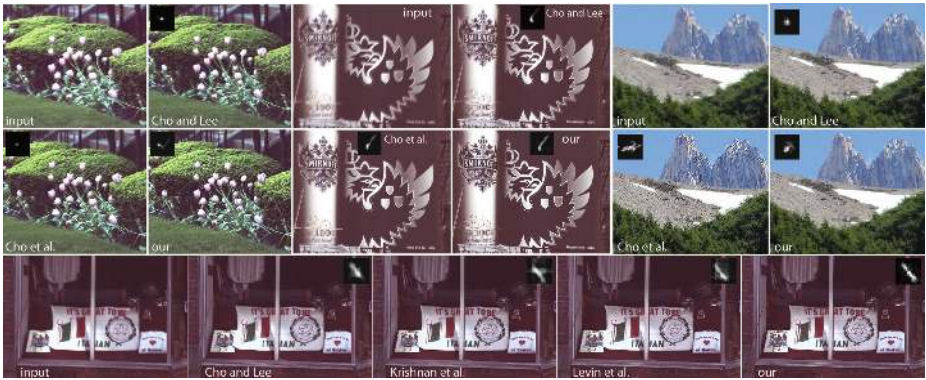


Fig. 6. Kernels estimated by different methods and the resulting deblurred images

figure we show three images (and kernels) that participated in this experiment and the estimated kernels and deblurred results.

We evaluated our method against [6, 10, 16] as well as Cho et al. [21] on images acquired with camera blur. We show the kernels estimated by the different methods and the resulting deblurred images in Fig. 6. This test shows our method’s robustness to foliage clutter and its ability compete with the state-of-the-art.

6 Conclusions

We presented a new method for recovering the blur kernel in motion-blurred natural images based on the statistical deviations they exhibit in their spectrum. Our method extracts a set of statistics from the input image, after properly whitening its spectrum, and uses them to recover the blur. Thus, it achieves favorable running-times compared to other methods that perform MAP estimation and recover the latent image repeatedly in the course of estimating the kernel.

The algorithm is based on a more refined model that we derived for modeling the power spectra of natural images. This model accounts for the presence of strong and long edges in the image and makes it possible to recover the blur in such scenarios. Our statistical approach does not rely on the presence nor the detection of well-defined step edges, needed at multiple orientations, as required by other methods [10, 21]. Finally, we described a modification that makes standard phase retrieval algorithms more robust to inaccuracies in the input.

As our algorithm relies on the canonical behavior of natural images, it is restricted to this class of images. Furthermore, since it estimates various statistics from the image, it cannot operate on small images where these statistics are unreliable. Our tests show successful operation on images of 512-by-512 pixels which fall well below the resolution of current digital cameras. Similarly to other methods that rely on the spatially-uniform blur, assumed in Eq. (1), the performance of our method deteriorates as the spatial variability of the blur increases.

This is the case when the camera motion contains rotation around its principal axis or in images with significant focal blur.

As future work we intend to follow this approach of non-MAP blur kernel estimation and consider other statistical regularities that can be exploited for this purpose.

Acknowledgements. The authors would like to thank Dilip Krishnan for his help. We also like to thank the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities as well as Microsoft New Faculty Fellowship Program for supporting this research.

References

1. Chan, T., Chiu-Kwong, W.: Total variation blind deconvolution. *IEEE Transactions on Image Processing* 7, 370–375 (1998)
2. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. In: *ACM SIGGRAPH Papers*, pp. 787–794. ACM, New York (2006)
3. Shan, Q., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. In: *ACM SIGGRAPH Papers*, pp. 73:1–73:10. ACM, New York (2008)
4. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding and evaluating blind deconvolution algorithms. In: *CVPR, IEEE*, pp. 1964–1971 (2009)
5. Krishnan, D., Fergus, R.: Fast image deconvolution using hyper-laplacian priors. In: *NIPS*, vol. 22, pp. 1033–1041 (2009)
6. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Efficient marginal likelihood optimization in blind deconvolution, pp. 2657–2664 (2011)
7. Yitzhaky, Y., Mor, I., Lantzman, A., Kopeika, N.S.: Direct method for restoration of motion-blurred images. *Journal of Opt. Soc. Am.* 15, 1512–1519 (1998)
8. Hu, W., Xue, J., Zheng, N.: Psf estimation via gradient domain correlation. *IEEE Transactions on Image Processing* 21, 386–392 (2012)
9. Fienup, J.R.: Phase retrieval algorithms: A comparison. *Appl. Opt.*, 2758–2769 (1982)
10. Cho, S., Lee, S.: Fast motion deblurring. In: *ACM SIGGRAPH Asia Papers*, pp. 145:1–145:8. ACM, New York (2009)
11. Legendijk, R.L., Biemond, J., Tekalp, A.M.: Maximum likelihood image and blur identification: a unifying approach. *Opt. Eng.* 29, 422–435 (1990)
12. You, Y.L., Kaveh, M.: A regularization approach to joint blur identification and image restoration. *IEEE Transactions on Image Processing* 5, 416–428 (1996)
13. Simoncelli, E.P.: statistical models for images: Compression, restoration and synthesis. In: *Proc. 31st Asilomar. Conf. on Signals, Systems and Computers*, vol. 1, pp. 673–678. *IEEE Computer Society* (1997)
14. Rav-Acha, A., Peleg, S.: Two motion-blurred images are better than one. *Pattern Recogn. Lett.* 26, 311–317 (2005)
15. Miskin, J., MacKay, D.: Ensemble learning for blind image separation and deconvolution. In: *Adv. in Independent Component Analysis*. Ed. Springer-Verlag (2000)
16. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. *CVPR* 1, 233–240 (2011)

17. Gupta, A., Joshi, N., Lawrence Zitnick, C., Cohen, M., Curless, B.: Single Image Deblurring Using Motion Density Functions. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 171–184. Springer, Heidelberg (2010)
18. Whyte, O., Sivic, J., Zisserman, A., Ponce, J.: Non-uniform deblurring for shaken images. In: CVPR, IEEE, pp. 491–498 (2010)
19. Joshi, N., Szeliski, R., Kriegman, D.: Psf estimation using sharp edge prediction. In: IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, Alaska (2008)
20. Xu, L., Jua, J.: Two-Phase Kernel Estimation for Robust Motion Deblurring. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 157–170. Springer, Heidelberg (2010)
21. Cho, T.S., Paris, S., Horn, B.K.P., Freeman, W.T.: Blur kernel estimation using the radon transform. In: CVPR, pp. 241–248 (2011)
22. Joshi, N., Zitnick, C.L., Szeliski, R., Kriegman, D.J.: Image deblurring and denoising using color priors. In: CVPR, pp. 1550–1557 (2009)
23. Jia, J.: Single image motion deblurring using transparency. In: Computer Vision and Pattern Recognition, pp. 1–8 (2007)
24. Ben-Ezra, M., Nayar, S.K.: Motion-based motion deblurring. *IEEE Trans. Pattern Anal. Mach. Intell.* 26, 689–698 (2004)
25. Raskar, R., Agrawal, A., Tumblin, J.: Coded exposure photography: motion deblurring using fluttered shutter. In: ACM SIGGRAPH Papers, pp. 795–804. ACM, New York (2006)
26. Levin, A., Fergus, R., Durand, F., Freeman, W.T.: Image and depth from a conventional camera with a coded aperture. In: ACM SIGGRAPH Papers. ACM, New York (2007)
27. Yuan, L., Sun, J., Quan, L., Shum, H.Y.: Image deblurring with blurred/noisy image pairs. In: ACM SIGGRAPH Papers. ACM, New York (2007)
28. Levin, A., Sand, P., Cho, T.S., Durand, F., Freeman, W.T.: Motion-invariant photography. In: ACM SIGGRAPH Papers, pp. 71:1–71:9. ACM, New York (2008)
29. Joshi, N., Kang, S.B., Zitnick, L.C., Szeliski, R.: Image deblurring using inertial measurement sensors. In: ACM SIGGRAPH Papers, pp. 30:1–30:9. ACM, New York (2010)
30. Field, D.J.: Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America. A, Optics and Image Science* 4, 2379–2394 (1987)
31. Burton, G.J., Moorhead, I.R.: Color and spatial structure in natural scenes. *Applied Optics* 26, 157–170 (1987)
32. Pouli, T., Cunningham, D., Reinhard, E.: Image statistics and their applications in computer graphics. In: Eurographics State of the Art Report, STAR (2010)
33. Luke, D.R.: Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems* 21, 37–50 (2005)
34. Bruck, Y., Sodin, L.: On the ambiguity of the image reconstruction problem. *Optics Communications* 30, 304–308 (1979)