Iterative Wiener Filtering for Deconvolution with Ringing Artifact Suppression

Filip Šroubek, Tomáš Kerepecký and Jan Kamenický
Institute of Information Theory and Automation
Czech Academy of Sciences
Prague, Czech Republic
sroubekf@utia.cas.cz

***Abstract*—Sensor and lens blur degrade images acquired by digital cameras. Simple and fast removal of blur using linear filtering, such as Wiener filter, produces results that are not acceptable in most of the cases due to ringing artifacts close to image borders and around edges in the image. More elaborate deconvolution methods with non-smooth regularization, such as total variation, provide superior performance with less artifacts, however at a price of increased computational cost. We consider the alternating directions method of multipliers, which is a popular choice to solve such non-smooth convex problems, and show that individual steps of the method can be decomposed to simple filtering and element-wise operations. Filtering is performed with two sets of filters, called restoration and update filters, which are learned for the given type of blur and noise level with two different learning methods. The proposed deconvolution algorithm is implemented in the spatial domain and can be easily extended to include other restoration tasks such as demosaicing and super-resolution. Experiments demonstrate performance of the algorithm with respect to the size of learned filters, number of iterations, noise level and type of blur.**

***Index Terms*—Wiener filter, LMMSE, deconvolution, total variation, ADMM, non-smooth optimization**

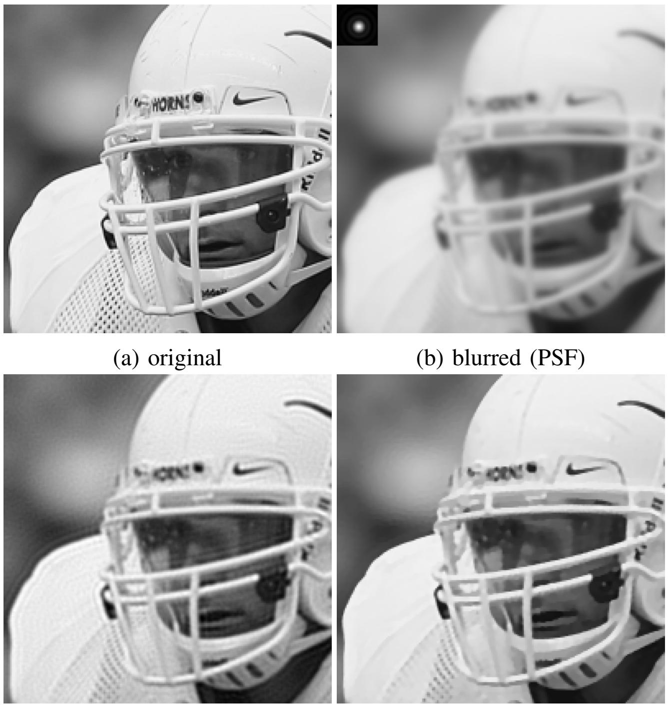
I. Introduction

Digital cameras are present in various measuring systems including microscopes, telescopes, and also small embedded systems like smartphones. Data acquired by camera sensors are subject to various types of signal degradation, for example lens and sensor blur, aberrations, color filter array (CFA) and noise. To obtain true images of the measured scene, it is necessary to correctly process the acquired data. The processing step is designed to run in the camera with limited computational capacity that allows only pixel-wise operations and some basic filtering.

Blur degradation often remains unattended in cameras as the cost to remove it is very high. Blur is modeled by convolution and even if the convolution kernel – called point spread function (PSF) – is known, the inverse problem of deconvolution is ill-posed due to values close to zero in spectra of common PSFs. For this reason, deconvolution methods based solely on linear operators, such as filtering, produce poor results.

From all linear filters, the optimal is the well-known Wiener filter, which is popular for having an explicit form in the frequency (Fourier) domain (FD) and estimates a sharp image in one step. However, the estimated image exhibits ringing artifacts around edges; see an example of Wiener output in Fig. 1(c) obtained by filtering the blurred image in Fig. 1(b). Another disadvantage is that implementation in the FD implicitly assumes circular convolution, which in real scenarios is violated and the so-called problem of boundary conditions results in disturbing artifacts close to image borders. Proposed remedies either solve the boundary pixels separately in the spatial (image) domain (SD) [1], [2], or modify the boundary pixels in the blurred image to better comply with circular convolution [3], [4].

Equivalently, the problem of boundary conditions can be



(a) original (b) blurred (PSF) (c) Wiener (d) proposed iterative Wiener

Fig. 1. Deconvolution without ringing artifacts: (a) the original image, (b) the blurred input image and PSF (inset), (c) restoration using the Wiener filter with power spectrum of the original image, (d) restoration using the proposed iterative Wiener filtering after 10 iterations. Notice ringing artifacts in the Wiener solution (c) and their absence in the proposed method (d).

This work was supported by Czech Science Foundation grant GA18-05360S and by the Praemium Academiae awarded by the Czech Academy of Sciences.

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

solved if only a 'valid' part of convolution is considered. Deconvolution formulated as a least squares optimization problem with Tikhonov quadratic regularization has a closed-form linear solution of type $Ax = b$; review classical restoration methods in [5]. However, when the 'valid' part of convolution is used, the inversion of A is typically not feasible and iterative numerical methods, such as Conjugate Gradient, must be used.

To achieve deconvolution results without artifacts, we have to leave the space of linear operators and allow non-linear ones. This is done by introducing non-smooth regularization terms, such as total variation [6], in the optimization problem; see for example [7]. Solving non-smooth convex problems requires specialized techniques, of which saddle-point methods [8] are probably the most popular. Generally, the solution is not in a closed form anymore and instead an iterative procedure is applied, which consists of multiple update equations and some of them are non-linear.

In this work, we propose to solve the deconvolution problem by combining the computational efficiency of Wiener filtering and the superior restoration quality of non-smooth optimization methods. We show that in the saddle-point methods linear update equations can be interpreted as Wiener-like filters and the non-linear update equations as soft thresholding. All steps are implemented in the SD using only filtering and element-wise operations, which naturally solves the problem of boundary artifacts. The filters are learned by solving a separate optimization problem on training data, which are specifically generated for the learning procedure. We foresee that the proposed algorithm easily extends to space-variant deconvolution, demosaicing or super-resolution.

The paper is organized as follows. In Section II, learning filters in the SD and the proposed algorithm is introduced. Section III experimentally validates the algorithm performance with respect to various conditions and Section IV concludes the paper with a short discussion of possible extensions of the algorithm.

II. METHODOLOGY

The discrete formation model considered in this work is a standard convolution process

$$g = Hu + n\,, \tag{1}$$

where g is the blurred and noisy image, u is the unknown sharp image, $H(\cdot) \equiv h * \cdot$ denotes a degradation operator (matrix) performing convolution with some known PSF h, and $n \approx \mathcal{N}(0, \sigma^2)$ is additive white Gaussian noise (AWGN) with zero mean and variance $\mathrm{Var}(n) = \sigma^2$. We consider scalar-valued digital images represented as column vectors $u \in \mathbb{R}^m$ and $g \in \mathbb{R}^p$. Pixels are indexed as $(u)_i$. In practice, H models 'valid' convolution and thus $m \geq p$. We define a discrete gradient operator $\mathsf{D} : \mathbb{R}^m \to \mathbb{R}^{m\times 2}$, which in its simplest form returns horizontal and vertical differences of pixels. It is a multidimensional array (tensor) consisting of two matrix components $[D_x, D_y]$ that perform convolution with $[1, -1]$ and $[1; -1]$ filters. The operator D can be more general and have more components, e.g. diagonal differences for isotropic behavior, or differences of pixels in a larger neighborhood to better capture correlation of pixels. On the vector-valued output of D, we define following norms:

$$\|\mathsf{D}u\|_{2,1}\colon \mathbb{R}^{m\times 2} \to \mathbb{R} \equiv \sum_i \left((\mathsf{D}u)_{i,1}^2 + (\mathsf{D}u)_{i,2}^2\right)^{1/2}\,,$$

$$\|\mathsf{D}u\|_{2,1}^2\colon \mathbb{R}^{m\times 2} \to \mathbb{R} \equiv \sum_i \left((\mathsf{D}u)_{i,1}^2 + (\mathsf{D}u)_{i,2}^2\right)\,.$$

If 'valid' convolution is replaced with circular convolution then (1) rewrites in the FD as

$$\mathcal{G} = \mathcal{H}\mathcal{U} + \mathcal{N}\,, \tag{2}$$

where capital calligraphic letters denote the Fourier transform $F(\cdot)$ of the corresponding function, e.g. $\mathcal{G} = Fg$, $\mathcal{H} = Fh$.

First let us formulate deconvolution as an optimization problem with Tikhonov regularization

$$\hat{u} = \arg\min_u \frac{\gamma}{2}\|Hu - g\|_2^2 + \|\mathsf{D}u\|_{2,1}^2\,, \tag{3}$$

where the norm of the first term is the classical ℓ_2-norm. A closed-form solution exists in this case and if circular convolution is assumed, the result in the FD has an explicit form of linear filtering

$$\hat{\mathcal{U}} = \hat{\mathcal{W}}\mathcal{G} = \frac{\mathcal{H}^*}{|\mathcal{H}|^2 + \frac{1}{\gamma}|\mathcal{D}|^2}\mathcal{G}\,, \tag{4}$$

where $\hat{\mathcal{W}}$ is a restoration filter in the FD. Note that since D is a tensor then $|\mathcal{D}|^2 = |\mathcal{D}_x|^2 + |\mathcal{D}_y|^2$, where $\mathcal{D}_{(\cdot)}$'s are Fourier transforms of gradient operator components. The power spectrum of D is identical to the spectrum of the Laplacian operator, which is f^2 with f being the spatial frequency.

The solution $\hat{\mathcal{W}}$ resembles a standard Wiener filter with a modified power spectrum of the original image. Recall that the Wiener filter is a linear minimum mean square error (LMMSE) estimator defined as

$$\hat{\mathcal{W}} = \arg\min_{\mathcal{W}} \mathbb{E}_{u,n}\{\|\mathcal{W}\mathcal{G} - \mathcal{U}\|_2^2\}\,, \tag{5}$$

where $\mathbb{E}_{u,n}\{\cdot\}$ denotes the expectation with respect to the distribution of images and noise. The solution is the well-known formula $\hat{\mathcal{W}} = \mathcal{H}^*/(|\mathcal{H}|^2 + \mathcal{S}_{nn}/\mathcal{S}_{uu})$ where $\mathcal{S}_{uu}$ and $\mathcal{S}_{nn}$ are power spectra of the original image and noise, respectively, and are assumed to be known. The filter $\hat{\mathcal{W}}$ in (4), which is the solution of Tikhonov regularization, is thus the Wiener filter for noise $n \approx \mathcal{N}(0, 1/\gamma)$ and images with the power spectrum $\mathcal{S}_{uu} = 1/|\mathcal{D}|^2$.

A. Learning restoration filters

To avoid the problem of boundary conditions in convolution, it is preferable to have restoration filters in the SD. We discuss two approaches. A straightforward one is to use the closed form solution in (4) and estimate the corresponding SD filter $\hat{w} \in \mathbb{R}^s$ for some given size s by solving

$$\hat{w} = \arg\min_w \|\hat{\mathcal{W}} - Fw\|^2 \quad \text{s.t.} \quad w \in \mathbb{R}^s,\ \sum_i (w)_i = (\hat{\mathcal{W}})_0\,, \tag{6}$$

where the second equality constraint guarantees that the filter mean is preserved. The above constrained optimization problem has a simple solution using the method of Lagrange multipliers: transform $\hat{\mathcal{W}}$ to the SD, crop it to size s, and add an appropriate constant to preserve the original mean value. However, there are two disadvantages of this approach. First, the filter is optimal in the sense of ℓ_2-norm calculated in the PSF domain, which has limited relation to the quality of the restoration. Second, it can be used only if the explicit form (4) in the FD exists. For example, if downsampling is present in the formation model, such as in super-resolution or demosaicing, the inversion must be done numerically and the FD explicit form is not viable.

A remedy to the above problems is the second approach that solves LMMSE (5) directly in the SD, see this approach applied to demosaicing in [9]. We take an arbitrary image $\tilde{u}$ and perform type of spectral whitening by modifying the image power spectrum to $\mathcal{S}_{\tilde{u}\tilde{u}} = 1/|\mathcal{D}|^2$. Then we generate a blurred image $\tilde{g}$ following the formation model (1) with $n \approx \mathcal{N}(0, 1/\gamma)$. The pair $(\tilde{u}, \tilde{g})$ is a training set, which we then use in optimization (5) by replacing the expectation with a sample mean. The complete algorithm is summarized in Alg. 1.

Algorithm 1 Learning restoration filters

Input: h – PSF, s – filter size, σ^2 – noise variance, $\mathcal{S}$ – power spectrum
Output: $\hat{w}$ – filter of size s

Generate a training pair $(\tilde{u}, \tilde{g})$ **:**
  Take some image $\tilde{u}$, modify its spectrum ($\mathcal{S}_{\tilde{u}\tilde{u}} = \mathcal{S}$), and generate a blurred image $\tilde{g} = h * \tilde{u} + n$ with $n \approx \mathcal{N}(0, \sigma^2)$.
Solve for $w \in \mathbb{R}^s$**:**
  $\hat{w} = \arg\min_{w} \|w * \tilde{g} - \tilde{u}\|_2^2$

B. Proposed iterative algorithm

Let us now reformulate deconvolution as an optimization problem with total variation regularization [6]

$$\hat{u} = \arg\min_{u} \frac{\gamma}{2}\|Hu - g\|_2^2 + \|\mathsf{D}u\|_{2,1} \,. \qquad (7)$$

Saddle-point methods are frequently used for solving such non-smooth convex problems. A popular choice is the 'alternating directions method of multipliers' (ADMM) [10], which is also considered here, however similar results are obtained also for 'primal-dual' methods of Chambolle and Pock [11]. The ADMM introduces an auxiliary variable $\mathsf{v} \in \mathbb{R}^{m\times 2}$ and an equality constraint $\mathsf{v} = \mathsf{D}u$, and rewrites (7) as a saddle-point problem for an 'augmented Lagrangian':

$$\min_{u,\mathsf{v}} \frac{\gamma}{2}\|Hu - g\|_2^2 + \|\mathsf{v}\|_{2,1} + \frac{\beta}{2}\|\mathsf{D}u - \mathsf{v} - \mathsf{a}\|_{2,1}^2 \,, \qquad (8)$$

where $\mathsf{a} \in \mathbb{R}^{m\times 2}$ is the Lagrange multiplier. Minimization with respect to the image u leads to a linear problem and if circular convolution is assumed, the result can be written in the FD as

$$\mathcal{U} = \frac{\mathcal{H}^*}{|\mathcal{H}|^2 + \frac{\beta}{\gamma}|\mathcal{D}|^2}\mathcal{G} + \frac{\mathcal{D}^*}{|\mathcal{D}|^2 + \frac{\gamma}{\beta}|\mathcal{H}|^2}(\mathcal{V} + \mathcal{A}) \qquad (9)$$

The first term is the solution of Tikhonov regularization (4) and is equivalent to Wiener filtering with PSF h, image power spectrum $1/|\mathcal{D}|^2$ and noise $n \approx \mathcal{N}(0, \beta/\gamma)$. Alg. 1 is applied with parameters $\sigma^2 = \beta/\gamma$ and $\mathcal{S} = 1/|\mathcal{D}|^2$ to learn the corresponding filter in the SD. We refer to this filter as '*restoration filter*' $w_1 \in \mathbb{R}^s$. The second term can be considered as another Wiener filtering of $(\mathsf{v} + \mathsf{a})$ with vertical and horizontal filters $[1, -1]$, image power spectrum $1/|\mathcal{H}|^2$ and noise $n \approx \mathcal{N}(0, \gamma/\beta)$. In this case, Alg. 1 is unstable since the PSF power spectrum $|\mathcal{H}|^2$ typically contains values close to zero in higher frequencies and setting the image power spectrum to $1/|\mathcal{H}|^2$ is not feasible. Instead, we apply the approach in (6) and refer to the estimated filters as '*update filters*' $\mathsf{w}_2 \in \mathbb{R}^{s\times 2}$. Note that w_2 is a vector-valued function and it consists of two filters one for each component of the gradient operator D (or more if D is more complex). Examples of restoration and update filters for two different PSFs are shown in Fig. 2. The remaining update equations for the auxiliary variable v and Lagrange multiplier a are in accordance with the ADMM and consist of simple element-wise operations. In the thresholding step, the norm on the vector-valued image is calculated per pixel as $\|\mathsf{D}u - \mathsf{a}\|_2 : \mathbb{R}^{m\times 2} \to \mathbb{R}^m \equiv \left((\mathsf{D}u - \mathsf{a})_{i,1}^2 + (\mathsf{D}u - \mathsf{a})_{i,2}^2\right)^{1/2}$ and the multiplication of the vector-valued image $(\mathsf{D}u - \mathsf{a})$ with the scalar-valued image $\max(\cdot)/\|\cdot\|_2$ is done element-wise by replicating the scalar-valued image.

Algorithm 2 Iterative Wiener filtering and thresholding (IWFT)

Input: g – blurred image, (w_1, w_2) – restoration and update filters, and N – number of iterations
Output: u – sharp image

1: **Initial estimation with restoration filter:**
2: $u_1 \leftarrow w_1 * g$
3: $k \leftarrow N$, $\mathsf{a} \leftarrow 0$, $\beta \leftarrow 10\max(g)$, $u \leftarrow u_1$
4: **repeat**
5: **Element-wise soft thresholding:**
6: $\mathsf{v} \leftarrow (\mathsf{D}u - \mathsf{a}) \cdot \dfrac{\max\left(\|\mathsf{D}u - \mathsf{a}\|_2 - \frac{1}{\beta}, 0\right)}{\|\mathsf{D}u - \mathsf{a}\|_2}$
7: $\mathsf{a} \leftarrow \mathsf{a} - \mathsf{D}u + \mathsf{v}$
8: **Improve the image with update filter:**
9: $u \leftarrow u_1 + \mathsf{w}_2 * (\mathsf{v} + \mathsf{a})$
10: $k \leftarrow k - 1$
11: **until** $k = 0$ **or** relative tolerance $< 10^{-4}$

The whole algorithm, which we call the iterative Wiener filtering and thresholding (IWFT) is summarized in Alg. 2. The filters w_1 and w_2 are inputs to the algorithm and they are precomputed for the given blur and noise level in the degraded image g. The algorithm consists of three main steps: initial filtering (line 2), element-wise computation (lines 6 and 7)

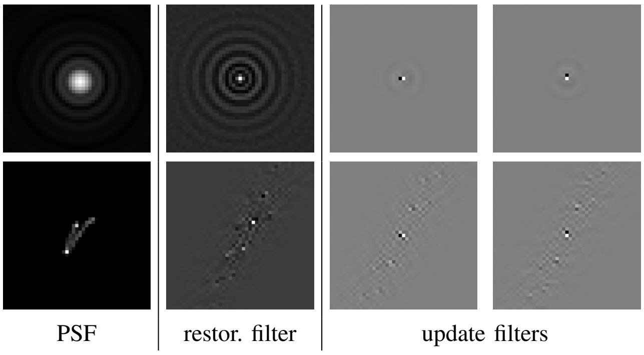


Fig. 2. Learned restoration and update filters for the Airy disk (top row) and motion blur (bottom row). From left to right: blur (h), initial restoration filter (w_1), two update filters (w_2) for horizontal and vertical differences. The filter size (s) is 45×45.

and update filtering (line 9). The first filtering step provides the initial estimator of u, which corresponds to the first term in (9). The remaining two steps are run iteratively. Update of v is element-wise soft thresholding with the threshold $1/\beta$. The parameter β is the same used in the construction of filters w_1 and w_2. Experimentally we have validated that the best results are achieved for β 10-times the range of intensity values, i.e. image gradients below $1/10$th of the intensity range are zeroed out in v. The update of the sharp image u is performed by filtering each component of the vector-valued image $(\mathsf{v} + \mathsf{a})$ with the corresponding update filter from w_2 and summing the results over the components. The algorithm finishes after satisfying one the convergence criteria: number of iterations or relative tolerance between the new estimation and the old one.

III. Experiments

The proposed IWFT algorithm solves the deconvolution problem (7) using the ADMM, which is guaranteed to converge. The appealing property of the proposed method is that all steps are implemented either by linear filters or simple element-wise operations, and thus the problem of boundary conditions in convolution is not present. The practical usage of the method is however determined by several other factors: by what margin the method outperforms the classical Wiener filter, how many iterations are generally required, and what is the minimum filter size to achieve these results. The following experiment addresses these issues in question.

The method performance was evaluated with respect to the filter size, number of iterations and noise level. The standard peak signal-to-noise (PSNR) ratio in dB was used as a performance measure. We also evaluated SSIM [12] and the results were equivalent. We took sharp images, blurred them with two types of blur – Airy disk modeling sensor blur and motion blur modeling camera shake – and added noise with SNR = 50, 30, 20dB. Figs. 1(a) and (b) illustrate an example of the original image and the corresponding one blurred by Airy disk, respectively. The restoration and update filters of different sizes were learned for each PSF and noise level with parameters $\gamma = 10^5$ (50dB), $\gamma = 10^3$ (30dB) and $\gamma = 10^2$ (20dB). Fig. 3 summarizes PSNR of the IWFT algorithm after $N = 0$, 1, 5 and 15 iterations for all generated images. $N = 0$ means that only the initial filtering with the restoration filter w_1 in step 2 is performed and the result is equivalent to the standard Wiener filter for the image power spectrum $1/|\mathcal{D}|^2$. In this case, strong ringing artifacts are present in the restored images as illustrated in the first column of Fig. 4. A noticeable improvement of the restored image both in the PSNR sense and visually is achieved after one application of the update filters w_2 (see the 1st iteration in the second column of Fig. 4). Additional iterations further improve the image, yet after 15 iterations (the last column) improvements are negligible.

The quality of restoration improves with the increasing filter size as expected. When the Airy disk is used, PSNR saturates for filter sizes of 45×45 in the case of 50dB. When noise increases in the image, the restoration and update filters perform more denoising than deconvolution and filters of smaller size become sufficient. So in the case of 30dB (20dB), PSNR saturates already for filter sizes of around 15×15 (10×10), however due to increased noise the achieved PSNR is lower. When the motion blur of similar effective size as the Airy disk is used, we notice that the maximum achievable PSNR is much higher. This is in accordance with the fact that Gaussian blurs (including Airy disk and out-of-focus) are more destructive than motion blurs. Examples of restored images for two noise levels and both blur types using filter size 45×45 are summarized in Fig. 5

IV. Conclusions

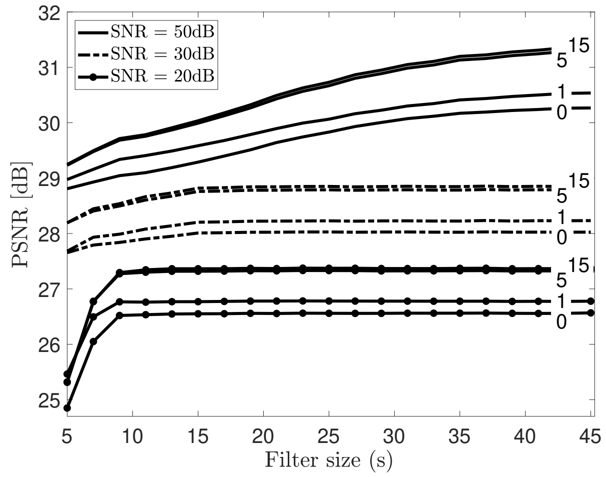
We have proposed a computationally efficient image restoration algorithm IWFT consisting of only filtering and element-wise operations, which makes it particularly suitable for implementation in digital cameras. The algorithm is based on the alternating directions method of multipliers and iteratively solves a non-smooth convex problem of deconvolution with total variation regularization using two linear filters. One filter is for initial restoration and another for updating the current estimate. Filters are implemented in the spatial domain and learned by two proposed learning methods. Experiments illustrate that the IWFT algorithm performs well for moderate filter sizes and removes ringing artifacts after only a few iterations in the case of realistic sensor and lens blurs.

A promising feature of the algorithm, which we plan to investigate in the near future, is the capability to seamlessly incorporate other restoration tasks. The algorithm for learning restoration filters is sufficiently general to estimate filters that in addition to deconvolution perform, e.g, demosaicing and super-resolution. In this case, we can replace the restoration filter for initial estimation with the newly learned filters and the rest of the algorithm remains the same.

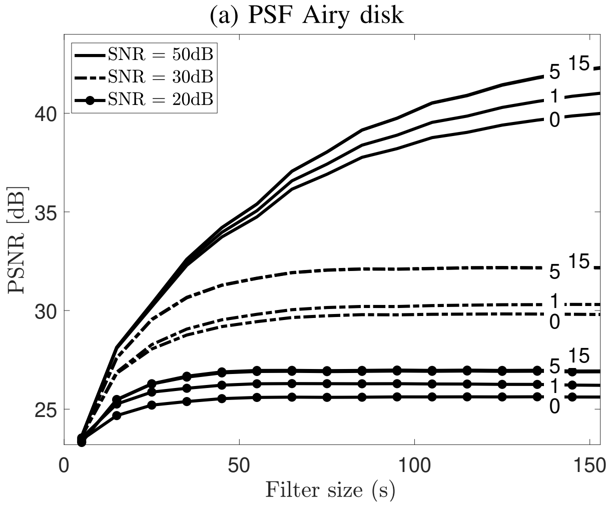
References

[1] S. J. Reeves, "Fast image restoration without boundary artifacts," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1448–1453, 2005.

[2] M. Šorel, "Removing boundary artifacts for real-time iterated shrinkage deconvolution," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2329–2334, Apr 2012.



(a) PSF Airy disk



(b) PSF motion

Fig. 3. PSNR performance of the proposed iterative Wiener filtering with respect to the size (s) of the restoration and update filters (w_1, w_2): (a) deconvolution of the Airy disk (top row in Fig. 2), (b) deconvolution of the motion blur (bottom row in Fig. 2). Performance curves are for 0 (only initial restoration filtering), 1, 5 and 15 iterations with noise levels 50dB (solid), 30dB (dashed) and 20dB (solid&marker). PSNR improves by a wide margin after the first 5 iterations and additional iterations are unnecessary. With increasing noise the minimum optimal filter size decreases. For the motion blur, achieved PSNRs are much higher yet the method is less practical as the performance plateau is reached for larger filter sizes.

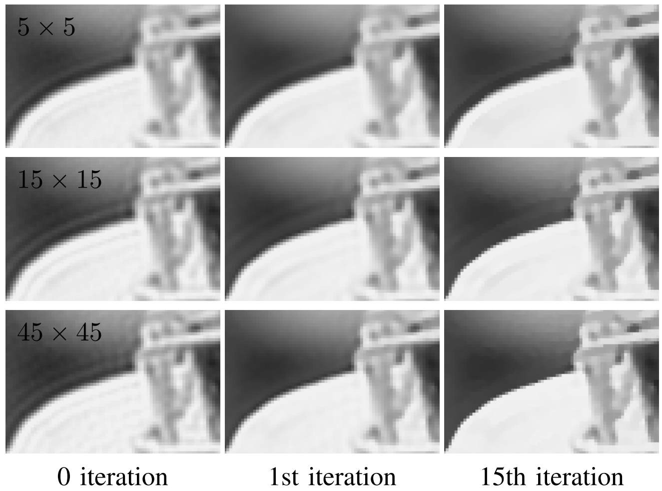


Fig. 4. Close-ups of restored images for different filter sizes and number of iterations: from top to bottom row – filter sizes of 5×5, 15×15 and 45×45; from left to right – 0 (only initial restoration filtering), 1 and 15 iterations. Images correspond to results of Airy disk deconvolution for SNR = 50dB with PSNR summarized in Fig. 3(a). Parts of the original and blurred image are in Fig. 1(a)-(b).

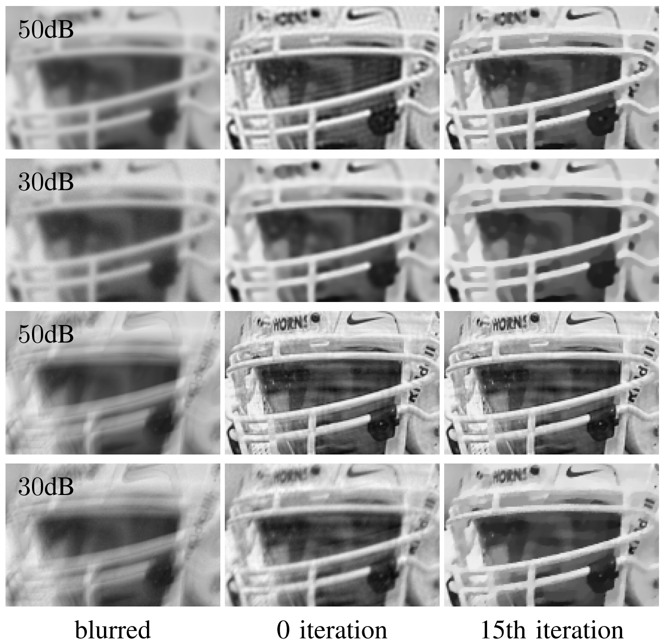


Fig. 5. Close-ups of restored images for different noise levels and number of iterations: from top to bottom row – noise levels of 50dB and 30dB; from left to right – input blurred image, 0 (only initial restoration filtering), 15 iterations. Images correspond to results of Airy disk (top two rows) and motion blur (bottom two rows) deconvolution for filter size 45×45.

[3] R. Liu and J. Jia, "Reducing boundary artifacts in image deconvolution," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*. IEEE, 2008, pp. 505–508.

[4] J. Portilla, "Maximum likelihood extension for non-circulant deconvolution," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Oct. 2014, pp. 4276–4279.

[5] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Process. Mag.*, vol. 14, no. 2, pp. 24–41, Mar. 1997.

[6] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[7] M. Almeida and M. Figueiredo, "Deconvolving images with unknown boundaries using the alternating direction method of multipliers," *Image Processing, IEEE Transactions on*, vol. 22, no. 8, pp. 3074–3086, Aug 2013.

[8] A. Chambolle and T. Pock, "An introduction to continuous optimization for imaging," *Acta Numerica*, vol. 25, pp. 161–319, 2016.

[9] J. Portilla, D. Otaduy, and C. Dorronsoro, "Low-complexity linear demosaicing using joint spatial-chromatic image statistics," in *Proc. IEEE Int. Conf. Image Processing 2005*, vol. 1, Sep. 2005, pp. I–61.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[11] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, no. 1, pp. 120–145, 2011.

[12] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.