PSF ACCURACY MEASURE FOR EVALUATION OF BLUR ESTIMATION ALGORITHMS

Jan Kotera[1,2], *Barbara Zitová*[1], *Filip Šroubek*[1]

[1]UTIA, Czech Academy of Sciences, Prague, Czech Republic
[2]Charles University in Prague, Faculty of Mathematics and Physics, Czech Republic

ABSTRACT

Given the large amount of blur estimation and blind deconvolution methods just in the last decade, there is an increasing need to compare the performance of a particular method with others. Unlike in other fields in image processing, there are very few well-established benchmark databases of test data and, more importantly, no standard way of performance evaluation. In this paper, we focus on the latter. We propose a new error measure for the blur kernel – a method for comparison of the blur estimate with the ground truth – which correctly reflects how inaccuracies in the blur estimation affect the subsequent image restoration, without the necessity to perform the actual deconvolution.

***Index Terms*—** PSF comparison, PSF error, blur estimation, image restoration, image deconvolution

1. INTRODUCTION

Single image blind deblurring is the task of estimating the latent sharp image from a single blurred image, without any prior knowledge of the blur itself. In most cases, the blurring process is modeled as a convolution of the unknown sharp image u with an unknown blur trace h (called Point-Spread Function, PSF) followed by a corruption with additive noise n, resulting in the observed image g

$$g = h * u + n. \tag{1}$$

Virtually every blind deconvolution method can be separated into two steps: blur estimation, in which the PSF h is computed, and image restoration, in which the desired image u is estimated using the blur estimate from the previous step. The latter part is called non-blind deconvolution (because the blur kernel is now considered known) and while it is still a non-trivial process in itself, it is a relatively straightforward and well-posed problem. Thus the success of the image restoration largely lies in the accuracy of blur estimation and this is where new blind deblurring methods usually center their contribution.

This work was supported by the GACR agency project GA13-29225S, GA UK project 938213/2013, the ARTEMIS joint undertaking under grant agreement no 641439 (ALMARVI), and SVV-2015-260223.

As with any task in image processing, there is a need to quantitatively compare the results of one method with others. Arguably the most common way is to measure the mean-square-error (MSE) of the restored image and the ground truth (or other derived measure of similarity, like ISNR), [1, 2, 3, 4]. The downside of this approach is that it does not give enough insight into how successful the key blur estimation alone actually was, because we evaluate the overall result only after the image restoration. Furthermore, such result is highly dependent on the used non-blind deconvolution method. Some non-blind methods are more forgiving to PSF error than others, some explicitly deal with border effects while other do not and the introduced error may overshadow the PSF error; all methods require tuning of usually several parameters. As a result, we can get completely different orderings (best to worst) of different PSF estimates just by using different non-blind methods or one method with different parameter settings, meaning that such process is useless as a performance evaluation tool across different author teams, where test conditions differ.

It is thus advantageous to have means of evaluating the accuracy of the blur estimation alone, rather then have the assessment of the first crucial step blurred by another non-trivial process. Surprisingly, there are no well-established standard methods for evaluation of blur estimation. One possible approach is to measure directly the MSE, or equivalently cross correlation, of the estimated PSF ([5, 4, 6]). The problem with PSF MSE is that it tells us how similar the PSFs look, it doesn't tell us how similar the restored images would look. While PSF estimation is a necessary intermediate step, the ultimate goal is the sharp restored image. Therefore, PSF error should be measured as it would propagate to image restoration. Two PSFs with the same MSE to the common ground-truth can produce vastly different results (both in terms of image MSE and human perception) when used for non-blind deconvolution.

To our best knowledge, the only attempt to better evaluate PSF estimation was made as a minor side product in [7, 8], where Levin et al. used the ratio between MSE of the image restored with the estimated kernel and MSE of the image restored with the ground truth kernel as the PSF error measure. This measure was adopted in later papers (e.g. [9, 10, 4, 11]), although it has the main drawback explained above – it heavily depends on the used non-blind deconvolution and measurement method. Many details, such as deconvolution method and its parameters, handling image borders, whether or not include extrapolated borders in the MSE calculation etc., were not specified by the authors of [7] and we know from experience that different choices in these aspects can change the error measure from small (close to 1) to several times higher on the same result set. For quantitative evaluation, such measure is then useless, but still used for lack of anything better.

We address this problem and, as a solution, propose a method which can accurately calculate the MSE introduced by the PSF to the image restoration just by one direct formula, without the need to perform the actual deconvolution. Such measure can very well serve as a standard because test results are easily reproducible, therefore meaningful. The next section contains mathematical description and derivation of the proposed measure, and Section 3 contains experimental verification of the efficacy of the proposed measure.

2. MATHEMATICAL DERIVATION

The error measure between the true PSF h and its estimate $\hat{h}$ we propose has the form

$$\rho[h,\,\hat{h}] = \mathbb{E}\left[\mathrm{MSE}(u_h, u_{\hat{h}})\right], \tag{2}$$

where u_h and $u_{\hat{h}}$ are sharp images restored with the correct PSF h and the estimated PSF $\hat{h}$, respectively, and $\mathbb{E}$ is probabilistic expectation over images and noise, which are treated as random variables. In the following subsection we provide a formula for evaluating the expression in (2) without having to actually compute u_h or $u_{\hat{h}}$, which of course we want to avoid. To summarize, we calculate the mean error of the image restoration caused by the PSF inaccuracy. For correctly estimated PSF, the error is zero and increases with the PSF error, but the increase reflects how the PSF error affects the image restoration.

Commonly used measure of improvement in image restoration is ISNR, defined as $10\log_{10}(\|u-g\|^2/\|u-u_{\hat{h}}\|^2)$. Based on PSF error measure (2) we propose an analogy, PSF-ISNR, as a measure of improvement in image restoration solely due to PSF estimation (with the effect of non-blind deconvolution eliminated)

$$\text{PSF-ISNR}[h,\,\hat{h}] = 10\log_{10}\left(\frac{\rho[h,\,\delta]}{\rho[h,\,\hat{h}]}\right), \tag{3}$$

where δ is the Dirac delta function (identity for convolution). We now proceed with the derivation of the formula for the error measure (2).

2.1. Derivation

Let us assume the blurring model in (1) and let us carry out the analysis in the Fourier domain, which is equivalent, since the error (2) is quadratic and L^2 norm is preserved. Fourier transforms of respective quantities will be denoted by capital letters, e.g. $G = HU + N$, omitting the particular frequency $G(\omega)$ etc. for brevity. If the correct PSF H is known, the optimal (in the sense of statistically mean squared error) estimate U_H of U obtainable by linear space-invariant filtering is the Wiener deconvolution

$$U_H = WG, \quad W = \frac{H^*}{|H|^2 + R}, \quad R = \frac{S_N}{S_U}, \tag{4}$$

where W is the Wiener filter, R is the noise-to-signal ratio, $S_U = \mathbb{E}[U^*U]$ is the image power spectrum, and $S_N = \mathbb{E}[N^*N]$ is the noise power spectrum. Similarly, if the PSF is estimated as $\hat{H}$, $\hat{H} \neq H$, the corresponding restored image $U_{\hat{H}}$ is

$$U_{\hat{H}} = \hat{W}G, \quad \hat{W} = \frac{\hat{H}^*}{|\hat{H}|^2 + R}. \tag{5}$$

The measure of PSF estimation error ρ (2) is the squared difference of the restored images. In Wiener deconvolution, the image U and noise N are regarded as random variables, therefore the difference must be considered in the sense of statistical expectation over all realizations of U and N

$$\rho[h,\hat{h}] = \mathbb{E}\left[\frac{1}{M^2}\sum_{\omega}\left|U_H - U_{\hat{H}}\right|^2\right], \tag{6}$$

where M is the number of image pixels. Plugging from (4) and (5) into the difference $U_H - U_{\hat{H}}$ and considering that $G = HU + N$ we have

$$\begin{aligned} U_H - U_{\hat{H}} &= (W - \hat{W})G \\ &= \left(\frac{H^*}{|H|^2 + R} - \frac{\hat{H}^*}{|\hat{H}|^2 + R}\right)(HU + N) \\ &= \frac{\left(H^*(|\hat{H}|^2 + R) - \hat{H}^*(|H|^2 + R)\right)(HU + N)}{(|H|^2 + R)(|\hat{H}|^2 + R)}. \end{aligned} \tag{7}$$

The expectation operator $\mathbb{E}$ acts non-trivially only on the last term in the numerator, $HU + N$. In the squared magnitude we get

$$\begin{aligned} \mathbb{E}\left[|HU + N|^2\right] &= |H|^2\mathbb{E}\left[|U|^2\right] + H\mathbb{E}\left[UN^*\right] \\ &+ H^*\mathbb{E}\left[U^*N\right] + \mathbb{E}\left[|N^2|\right] = |H|^2 S_U + S_N. \end{aligned} \tag{8}$$

U and N are assumed independent and N has zero mean, so the second and third term vanish.

Combining (7) and (8) and pluggin into (6), we finally get (expectation $\mathbb{E}$ and summation are interchangeable)

$$\begin{aligned} \rho[h,\hat{h}] &= \frac{1}{M^2}\sum_{\omega}\mathbb{E}\left[\left|U_H - U_{\hat{H}}\right|^2\right] \\ &= \frac{1}{M^2}\sum_{\omega}\frac{S_U\left|H^*(|\hat{H}|^2 + R) - \hat{H}^*(|H|^2 + R)\right|^2}{(|H|^2 + R)(|\hat{H}|^2 + R)^2}. \end{aligned} \tag{9}$$

2.2. Remarks

Equation (9) is our final PSF error measure. It is a single formula which can be readily evaluated, it does not require any iterative deconvolution algorithm, and it contains only the compared PSFs, the image power spectrum S_U and noise level (in R) as a single free parameter. The presence of S_U is understandable because different images respond differently to change in the (de)blurring PSF. For example, an image with horizontal translation symmetry will be unaffected by horizontal blur, therefore if h and $\hat{h}$ for such image differ only by a horizontal blur b (i.e., $h = \hat{h} * b$ or vice versa), then $\rho[h, \hat{h}] = 0$, which corresponds to the fact that the reconstructed images are identical. For calculation, we use the power spectrum of u as S_U, which presents no practical problem, because for quantitative testing, true h and therefore true u must be known.

The presence of R (eq. (4)) deserves some attention. Assuming that the input image u was corrupted with i.i.d Gaussian noise n, $n \sim N(0, \sigma^2)$, then $S_N \equiv M\sigma^2$. Therefore, for a fixed test image u, R is completely determined by single parameter σ, $R = M\sigma^2/S_U$. As a simplification with minor loss in fidelity, one can even use $R \equiv const$. In Wiener deconvolution, the parameter R can serve not only as an estimate of the input noise, but also as a regularizer compensating for the error in the PSF. If R is set too low for near-noiseless input image, even small errors in the PSF result in unpleasant visual artifacts in the restored image while with R set higher, the image restoration will be more "forgiving". The same situation occurs in the error measure ρ, if R is set too low, any deviation from the ground-truth PSF will result in high error values while high R will make the error measure less sensitive. Consequently, any published test results will be fully reproducible only if they include the value of R, the only free parameter, used for error measurement.

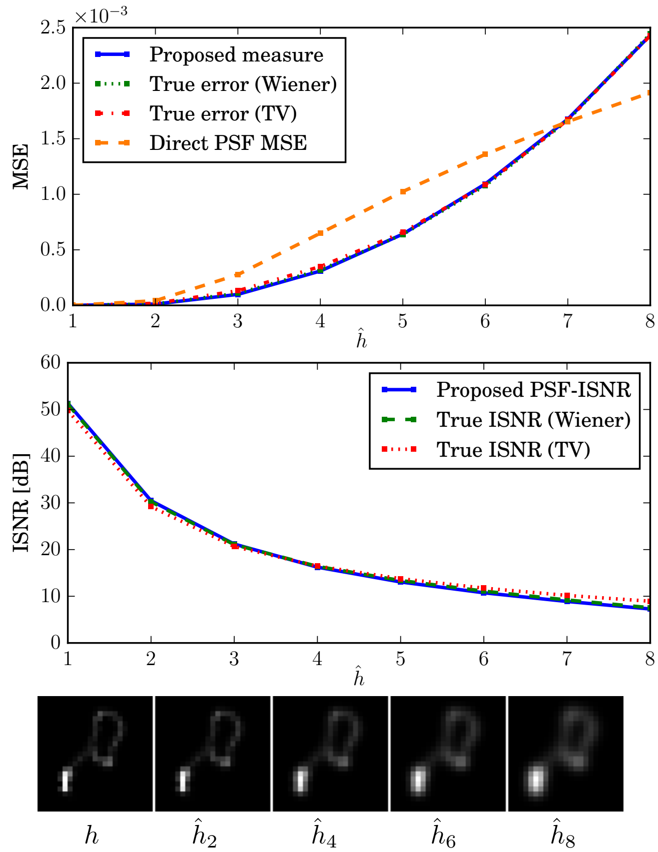

Fig. 1. Errors of PSFs with incorrect support. Top figure shows proposed PSF error compared with true image error obtained by deconvolution (Wiener filtering and total variation), and MSE calculated directly on the PSFs. Middle figure shows proposed PSF-ISNR measure compared with true ISNR obtained by deconvolution. Bottom strip contains the tested PSFs (selection): ground truth (first) and its mutations.

3. EXPERIMENTAL VERIFICATION

In this section, we present experimental verification, that the error measure ρ calculated for several types of error in PSFs corresponds to the true MSE measured after image restoration, and provide some comments. In Fig. 1-3 we show results of three experiments for three typical kinds of PSF error: incorrect support estimation (PSFs appear too thick, very common in blind PSF estimation), incorrect length of motion blur, and incorrect radius of out-of-focus blur. In each of these experiments, we constructed several PSFs with various degrees of aforementioned error and for each of these PSFs we calculated

- our proposed error measure ρ (solid blue line),
- true MSE of the images restored with Wiener filtering (dotted green),
- true MSE of the images restored by total variation (e.g. [12, 13]) deconvolution (dashdotted red), and
- direct MSE between PSFs themselves, scaled appropriately (dashed orange).

These error measurements are depicted in the top image of Fig. 1-3. All experiments were computed with the value $\sigma = 2^{-6}$ and the known `cameraman` image.

It is well apparent, that the calculated error measure ρ very accurately captures the true measured error introduced into the image by the PSF inaccuracy while retaining all the advantages of direct PSF error measure discussed above. Although our derivation is based in Wiener deconvolution, we can see that TV-based deconvolution behaves very similarly when it comes to image error, therefore conclusions form our analysis can be safely extended to more sophisticated non-blind deconvolution algorithms. On the other hand, direct PSF MSE provides absolutely no insight into how much particular PSF estimate affects the image restoration, this is well visible especially in Fig. 2 and 3 (see the orange dashed line).

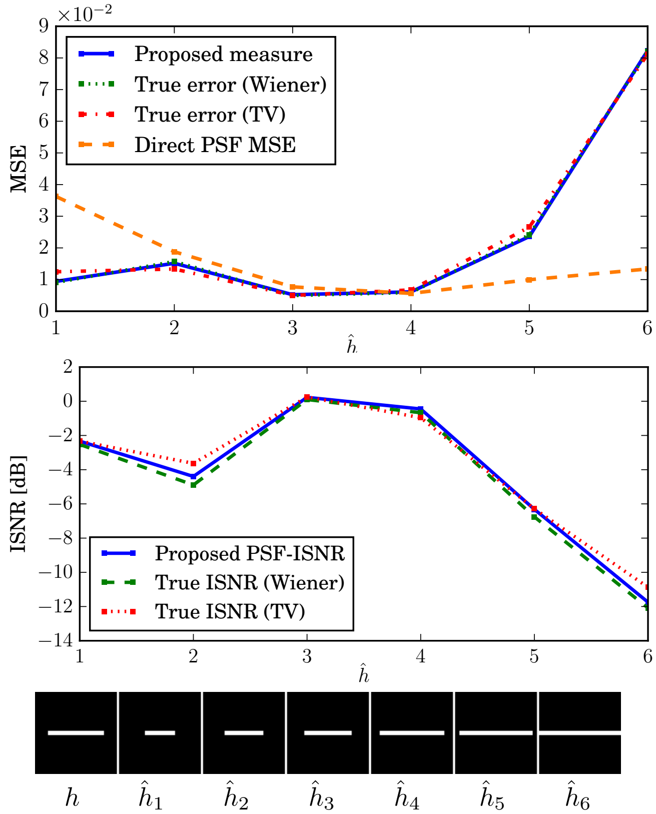

Fig. 2. Errors of PSFs with incorrect motion length (too short or long). Otherwise same setup as Fig. 1

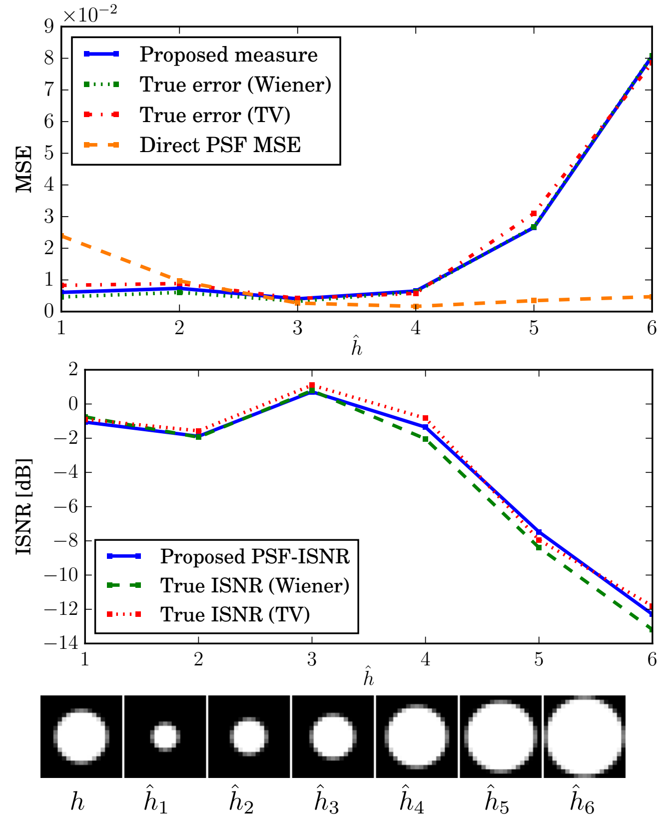

Fig. 3. Errors of PSFs with incorrect blur radius (too small or large). Otherwise same setup as Fig. 1

We can draw similar conclusions from the graphs depicting the comparison of the proposed PSF-ISNR measure and true image ISNR (middle image of Fig. 1-3), our measure has the fidelity of the ISNR computed after full deconvolution.

In the last presented experiment we changed the value of σ, Fig. 4 shows that while the absolute values of the error measure depend on σ, the relative comparison between PSFs and fidelity to true image error remain the same.

4. CONCLUSION

We proposed and presented mathematical derivation of a new error measure for direct blur estimation evaluation. We explained the necessity and advantages of evaluating blur estimation separately from image restoration and our proposed method fills the gap that is currently in this area. We also proposed PSF-ISNR as a counterpart for the well-accepted image ISNR but focused solely on blur estimation – a measure of improvement due to correct PSF estimation. MATLAB code for these measures can be readily downloaded from our website. We verified experimentally that the proposed error measure fulfills its promise and accurately captures the error introduced into image restoration by PSF error, regardless of the non-blind method used. Future work will include explicit dealing with border effects and space-invariant blur comparison.

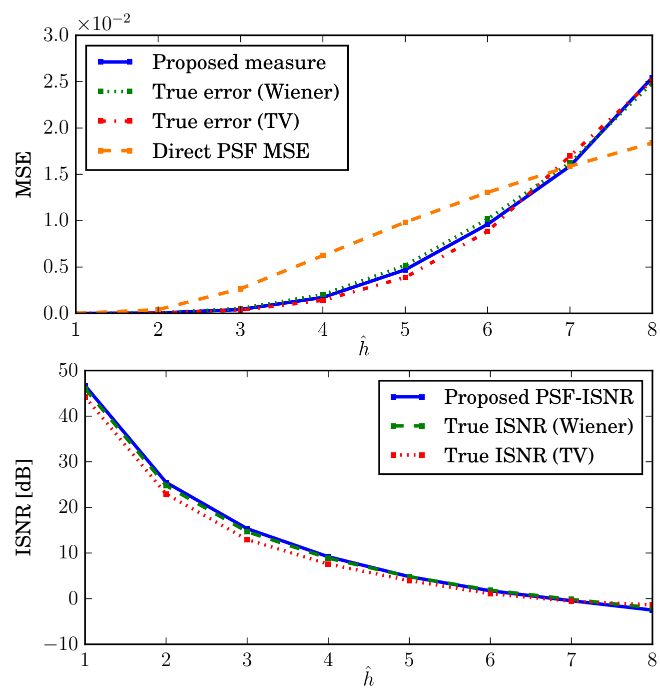

Fig. 4. Same setup as in Fig. 1 but for 8 times smaller value of σ.

5. REFERENCES

[1] M.S.C. Almeida and L.B. Almeida, "Blind and semi-blind deblurring of natural images," *Image Processing, IEEE Transactions on*, vol. 19, no. 1, pp. 36–52, Jan 2010.

[2] S Derin Babacan, Rafael Molina, Minh N Do, and Aggelos K Katsaggelos, "Bayesian blind deconvolution with general sparse image priors," in *Computer Vision–ECCV 2012*, pp. 341–355. Springer, 2012.

[3] Mariana SC Almeida and Mário AT Figueiredo, "Blind image deblurring with unknown boundaries using the alternating direction method of multipliers.," in *ICIP*. Citeseer, 2013, pp. 586–590.

[4] Li Xu, Shicheng Zheng, and Jiaya Jia, "Unnatural l0 sparse representation for natural image deblurring," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, June 2013, pp. 1107–1114.

[5] Sunghyun Cho and Seungyong Lee, "Fast motion deblurring," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 145:1–145:8, Dec. 2009.

[6] Jan Kotera, Filip Šroubek, and Peyman Milanfar, "Blind deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors," in *Computer Analysis of Images and Patterns*. Springer, 2013, pp. 59–66.

[7] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR '09*, 2009, pp. 1964–1971.

[8] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind deconvolution algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2354–2367, 2011.

[9] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, June 2011, pp. 233–240.

[10] Amit Goldstein and Raanan Fattal, "Blur-kernel estimation from spectral irregularities," in *Computer Vision–ECCV 2012*, pp. 622–635. Springer, 2012.

[11] D. Perrone and P. Favaro, "Total variation blind deconvolution: The devil is in the details," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, June 2014, pp. 2909–2916.

[12] Peter Blomgren, Tony F. Chan, Pep Mulet, and Chak-Kuen Wong, "Total variation image restoration: numerical methods and extensions," in *Image Processing, International Conference on*. 1997, vol. 3, pp. 384–384, IEEE Computer Society.

[13] S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, "An augmented lagrangian method for total variation video restoration," *IEEE Transactions on Image Processing*, vol. 20, no. 11, pp. 3097–3111, Nov. 2011.