



**HAL**  
open science

# Superresolution Alignment with Innocence Assumption: Towards a Fair Quality Measurement for Blind Deconvolution

Martin Welk

► **To cite this version:**

Martin Welk. Superresolution Alignment with Innocence Assumption: Towards a Fair Quality Measurement for Blind Deconvolution. OAGM&ARW Joint Workshop 2017: Vision, Automation and Robotics. May 10-12, 2017, Vienna, Peter M. Roth; Markus Vincze; Wilfried Kubinger; Andreas Müller; Bernhard Blaschitz; Svorad Stolc, May 2017, Vienna, Austria. 10.3217/978-3-85125-524-9-29 . hal-01592311

**HAL Id: hal-01592311**

**<https://hal.science/hal-01592311>**

Submitted on 23 Sep 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Superresolution Alignment with Innocence Assumption: Towards a Fair Quality Measurement for Blind Deconvolution

Martin Welk<sup>1</sup>

**Abstract**—Quantitative measurements of restoration quality in blind deconvolution are complicated by the necessity to compensate for opposite shifts of reconstructed image and point-spread function. Alignment procedures mentioned for this purpose in the literature are sometimes not exactly enough specified; alignment-free approaches sometimes do not take into account the full variability of possible shifts. We investigate by experiments on a simple test case the errors induced by interpolation-based alignment procedures. We propose a new method for MSE/PSNR measurement of image pairs involving non-integer displacements that is based on a superresolution approach. We introduce an innocence assumption in order to keep deviations that can be explained by shifted sampling grids out of the error measurement. In our test case, the new measurement procedure reduces the variations in MSE/PSNR measurements substantially, creating the hope that it can be used for valid comparisons of blind deconvolution methods.

## I. INTRODUCTION

The removal of blur in images by blind image deconvolution has been studied for many years [2], [3], [4], [5], [6], [10], [16], [21], and received increasing interest during the last years [1], [7], [8], [9], [11], [12], [14]. A frequently used simplifying assumption is that the blur is spatially invariant, i.e. the redistribution of intensity is described by the same point-spread function (PSF)  $h$  at each image location. Blur is then described by a convolution between the unobserved sharp image  $g$  and the PSF  $h$ ; incorporating additive noise  $n$ , the observed image  $f$  is given by the blur model

$$f = g * h + n. \quad (1)$$

Whereas for non-blind deconvolution one assumes that  $f$  and  $h$  are known, and aims at an estimate  $u$  for the sharp image  $g$ , the knowledge of  $h$  is often not available in practice, thus necessitating blind deconvolution where the estimate  $u$  of the sharp image is to be obtained along with the PSF  $h$ , using only  $f$  as input image.

A variety of approaches to solve this task have been developed, creating the need for quality comparisons. Besides visual assessment, one is interested in quantitative measurements of reconstruction quality versus a known ground truth.

Frequently used standard measures for image reconstruction methods include the mean-square error (MSE) as well as the signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR) both of which are closely related to the MSE; furthermore, sometimes the average absolute error (AAE) is

advocated. Another measure that puts some more emphasis on important structural details of images such as contrast edges is the structural similarity index (SSIM), see [17]. Let us shortly recall the first three measures.

For a reference (ground-truth) image  $g$  and degraded (or reconstructed) image  $u$ , both of size  $n \times m$  pixels, their MSE is given by

$$\text{MSE}(u, g) = \frac{1}{nm} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (u_{i,j} - g_{i,j})^2. \quad (2)$$

Provided that  $u$  and  $g$  have equal mean intensity  $\mu$  (which we will assume in the following), this is the variance  $\text{var}(u - g)$  of  $u - g$ . Using the variance of  $g$  given by

$$\text{var}(g) = \frac{1}{nm} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (g_{i,j} - \mu)^2, \quad (3)$$

and the range  $R(g) := \max_{i,j} g_{i,j} - \min_{i,j} g_{i,j}$  (255 for saturated 8-bit images), one can compute the SNR

$$\text{SNR}(u, g) = 10 \log \frac{\text{var}(g)}{\text{var}(u - g)} \text{ dB} \quad (4)$$

and PSNR

$$\text{PSNR}(u, g) = 10 \log \frac{R(g)^2}{\text{var}(u - g)} \text{ dB}. \quad (5)$$

For non-blind deconvolution, both MSE/(P)SNR and SSIM are frequently used to assess reconstruction quality. Although these quantitative measures are not always in good agreement with visual assessments by humans, they are generally accepted as simple and objective measures. For a recent study on measures that approximate better the human perception of restoration quality see [13].

In blind deconvolution, however, their application meets a difficulty: If the reconstructed image  $u$  is translated by an arbitrary, often non-integer, displacement  $d$ , and the point-spread function  $h$  is translated by  $-d$ , these translations cancel in the convolution  $u * h$ . Blind deconvolution results that differ just by such opposite translations of  $u$  and  $h$  must therefore be considered equally valid reconstructions. An example of such shifts that indeed occur in blind deconvolution results is shown in Fig. 1. This precludes a straightforward (P)SNR or SSIM comparison of blind deconvolution results with ground truth. Obviously, some kind of alignment – rigid registration restricted to translations as transformations – has to be applied.

Nevertheless, blind deconvolution results are compared by PSNR and other quantitative measures in a number of works, e.g. [6], [7], [8], [9], [10], [14]. In many of these,

\*This work was not supported by any organization

<sup>1</sup>Martin Welk is with Department of Biomedical Informatics and Mechatronics, Private University for Health Sciences, Medical Informatics and Technology (UMIT), 6060 Hall/Tyrol, Austria  
martin.welk@umit.at

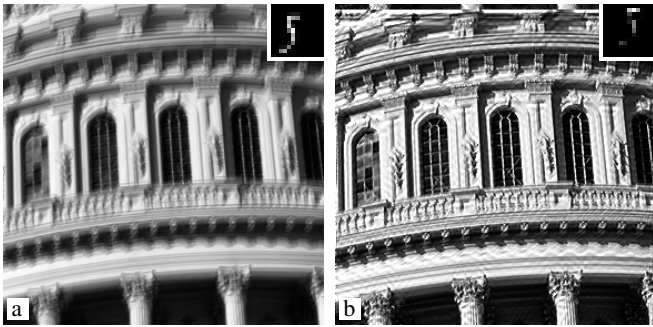


Fig. 1. (a) Synthetically blurred image with ground-truth PSF, from [11]. – (b) Blind deconvolution result with PSF, from [12]. Note the opposite shifts of image and PSF.

no alignment whatsoever is mentioned [6], [8], [9], [10]. Such an evaluation relies implicitly on the assumption that estimated PSFs are aligned with the ground truth PSF; probably this is approximately achieved by some test cases with small PSF support. Efforts to compensate shifts, either for images or for PSFs, are found in [7], [14], [20]. A benchmark established in [7] is based on simulating camera shake by generated trajectories. Multiple ground truth images are acquired directly along those trajectories, and the best match is used for error measurement. On one hand, a computational alignment step is avoided in this way. On the other hand, the procedure constrains shifts to the ground-truth trajectory which may be insufficient since blind deconvolution methods can well yield translations in which the coordinate origin of the PSF does not happen to be on the (unknown) trajectory that was used to generate the ground truth. The benchmark from [7] is also used in [20] and part of the evaluation in [14]. Further tests in [14] are based on data from [9]. Here, absolute errors of PSFs are measured, namely for “(aligned) blur estimates” with respect to ground truth PSFs. This allows indeed to handle unconstrained displacements. Details of the alignment procedure are not given, however.

In the following we discuss how to make precise such an alignment procedure. We focus on a scenario where a ground-truth image and PSF are available, and restoration quality is to be estimated by measuring the error between the ground truth and reconstructed images. In specifying the alignment procedure, some choices have to be made: first, should one register the reconstructed image to the ground truth image, or vice versa, or should perhaps both be transformed? Which interpolation procedure is to be used in the registration process? It is not a far-fetched guess that these details will influence the subsequent error measurements. In fact, we will demonstrate by a simple experiment in Section II that, dependent on details of the registration, the PSNR measures vary by 1.5 dB and more.

Given the fact that relative improvements of one blind deconvolution method over the other as reported in e.g. [7], [14] often amount to as little as 0.5 dB or even less, such a difference is significant.

This might be mitigated by using multiple test images and performing statistics on the errors measured for these. How-

ever, questions remain: Since errors introduced by interpolation can be expected to differ substantially between test cases where the displacement is approximately integer, and test cases where the displacement is near a half-pixel position, results may be strongly biased towards blind deconvolution methods that, for whatever reason, tend to reconstruct PSFs in similar pixel alignment as the ground-truth. Given the complexity of procedures both for constructing apparently realistic test cases, and of the blind deconvolution procedures themselves, it is such favourable alignments occur more often for some methods under investigation than for others. In such a case, the bias won’t necessarily average out for larger sample sizes.

For this reason, we pursue in this paper the goal to establish an alignment procedure for blind deconvolution results that avoids these pitfalls. We focus here on the MSE, from which (P)SNR can be derived via (4), (5).

**Structure of the paper.** In Section II we evaluate the errors introduced by interpolation-based alignment procedures using a simple test case. Section III establishes the fundamentals of an alignment procedure by superresolution in order to avoid these errors. The details of the procedure are discussed in Section IV, followed by experiments on the previously introduced test case in Section V. A short summary and outlook in Section VI concludes the paper.

## II. ALIGNMENT BY INTERPOLATION

To assess the errors introduced by alignment with interpolation, we set up a simple test case based on a ground truth grey-value image shown in Fig. 2 (a). We blur this image by 16 different PSFs shown in Fig. 2 (b); all these PSFs are downsampled versions of the same high-resolution PSF with horizontal and vertical shifts in  $1/4$  pixel steps. One blurred image is shown in Fig. 2 (c). Each of the blurred images is deconvolved with each of the 16 PSFs using the non-blind deconvolution method from [18] with the same parameters ( $\alpha = 0.01$ , 300 iterations). This yields 256 deblurred images with effective shifts w.r.t. the ground truth images from  $-0.75$  to  $+0.75$  pixels in  $x$  and  $y$  direction; one exemplary deblurred image is shown in Fig. 2 (d).

We can now measure the MSE (and resulting PSNR) for each deblurred image w.r.t. the ground truth image. In the following we report PSNR values as this is the most familiar measure in deconvolution literature. To reduce the impact of boundary artifacts, a 20 pixel wide margin is excluded from the measurement, thus using a  $88 \times 88$  central patch of the reference image.

We notice first that in the 16 translation-free cases (where the same PSF was used for blurring and deblurring) the PSNR varies between 29.74 and 30.41 dB, with an average of 30.07 dB and a standard deviation of 0.21 dB.

Next, we measure PSNR values for the entire set of 256 deblurred images. Here, the ground-truth and reconstructed images are aligned using either bilinear and bicubic interpolation with the ground-truth shift values. For the direction of alignment we consider three settings: (a) warping the

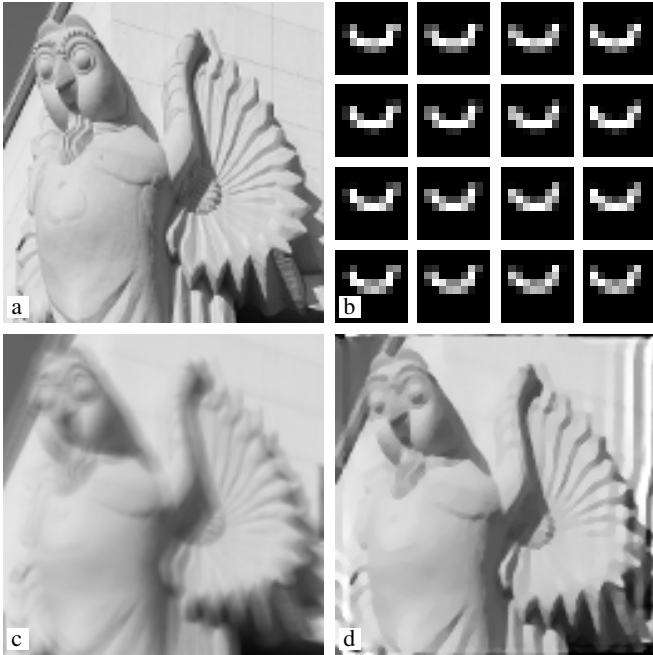


Fig. 2. (a) Ground truth image,  $128 \times 128$  pixels. (Clipped, downscaled and converted to greyscale from a photograph of the building of TU Vienna. Source of original image: [https://upload.wikimedia.org/wikipedia/commons/e/e9/TU\\_Bibl\\_01\\_DSC1099w.jpg](https://upload.wikimedia.org/wikipedia/commons/e/e9/TU_Bibl_01_DSC1099w.jpg), Author: Peter Haas. Available under licence CC BY-SA 3.0.) – (b) 16 PSFs,  $10 \times 10$  pixels each, subsampled from the same high-resolution input. The shift from row to row and from column to column is 0.25 pixels. – (c) Image (a) blurred by convolution with PSF from (b), first row, second column. – **Bottom right:** Image (c) deblurred with PSF from fourth row, third column, resulting in a shift relative to ground truth of  $(0.25, 0.75)$  pixels.

reconstructed image to match the ground-truth image; (b) warping the ground-truth image to match the reconstructed image; (c) applying half the shift to each of the ground-truth and reconstructed image. Statistics of the resulting PSNR values are presented in Table I.

To bring the previous procedure closer to a true blind deconvolution setting, we now switch to determining also the displacement from a minimisation of the MSE (or maximisation of the PSNR). To avoid analysing possible multiple optima, we employ here a brute-force optimisation varying the displacements in  $x$  and  $y$  direction in 0.01 steps from  $-1$  to  $1$ ; note that the exact displacements occur in the sequence

TABLE I

PSNR STATISTICS FOR 256 RECONSTRUCTED IMAGES WITH ALIGNMENT BY THE KNOWN (GROUND-TRUTH) SHIFT USING BILINEAR OR BICUBIC INTERPOLATION; (A) WARPING THE RECONSTRUCTED IMAGE, (B) WARPING THE GROUND TRUTH, (C) HALF-WAY WARPING GROUND TRUTH AND RECONSTRUCTED IMAGE.

Interpolation Alignment	bilinear			bicubic		
	(a)	(b)	(c)	(a)	(b)	(c)
min	27.47	29.74	29.74	28.35	29.74	29.39
max	30.41	33.54	33.55	30.41	31.84	31.55
(max–min)	2.94	3.80	3.81	2.06	2.10	2.16
mean	28.57	32.18	31.25	29.23	30.85	30.05
standard dev.	0.711	0.970	0.805	0.474	0.489	0.459

of displacements sampled thereby. Table II contains statistics of the misestimations  $\delta x$ ,  $\delta y$  of the  $x$  and  $y$  displacements, and the resulting PSNR. The latter values are slightly higher than in Table I but not seriously so.

As can be expected, warping the reconstructed image to match the ground truth (see columns marked (a) in Tables I and II) leads to lower PSNR values for image pairs with non-integer displacements. The variation is about 3dB with bilinear interpolation; bicubic interpolation reduces it to about 2dB which is still likely to warp comparisons substantially. When aligning instead the ground truth to the reconstructed images (columns (b) in Tables I and II) PSNR values are surprisingly higher for non-integer displacements, which means by comparison to the no-shift cases a clear overestimation of reconstruction quality. Apparently the warping of the ground truth image introduces some blur which matches well the remaining blur in the deconvolution results.

Inspection of the detail results corroborates that for the same image pair the choice which image is aligned to which one leads to discrepancies in PSNR of 4dB and more with bilinear, and still about 3dB with bicubic interpolation. Distributing the shift to both images (columns (c) in Table I) yields similar results as shifting the ground truth. As this proceeding does not offer an advantage, we do not pursue it further in the computationally more expensive scenario of Table II where also the displacements are optimised.

### III. ALIGNMENT BY SUPERRESOLUTION

We turn now to designing a procedure for image reconstruction error measurement with alignment. We give preference to the MSE as basis of our considerations because unlike the (P)SNR it treats the two images being compared in a completely symmetric way. We want to keep this symmetry also in the alignment procedure, thereby removing one of the arbitrariness of interpolation-based alignment procedures. For easier comparison to usual PSNR figures we will nevertheless report in the experiments later PSNR values computed from our MSE measurements.

An obvious requirement is that for perfectly aligned images the standard MSE measure has to be reproduced. Whereas the procedure will be described for prescribed

TABLE II

STATISTICS OF DISPLACEMENT MISESTIMATIONS  $\delta x$ ,  $\delta y$  AND PSNR FOR 256 RECONSTRUCTED IMAGES WITH ALIGNMENT ESTIMATED BY MSE MINIMISATION.

Interpolation Alignment	bilinear		bicubic	
	(a)	(b)	(a)	(b)
max $ \delta x $	0.18	0.17	0.07	0.16
std. dev. $\delta x$	0.079	0.081	0.028	0.080
max $ \delta y $	0.17	0.15	0.06	0.14
std. dev. $\delta y$	0.064	0.067	0.024	0.072
min PSNR	27.47	31.42	28.37	29.89
max PSNR	30.41	33.61	30.41	31.86
(max–min) PSNR	2.94	2.19	2.04	1.97
mean PSNR	28.67	32.67	29.25	31.10
std. dev. PSNR	0.697	0.527	0.471	0.465

displacement values, minimisation of the MSE measure is an obvious way to estimate also unknown displacements.

For the following, let us consider two images  $u$  and  $g$ , which are sampled representations of continuous-scale images. To specify the sampling process more precise, we assume that each pixel of  $g$  is the integral of the underlying continuous-scale image  $G$  over a rectangled region such that all pixels together tessellate (a rectangle of) the image plane:

$$g_{i,j} = \int_i^{i+1} \int_j^{j+1} G(x,y) dy dx, \quad (6)$$

and similarly for  $u$  whose grid is of equal resolution but shifted by  $d = (\alpha, \beta) \in \mathbb{R}^2$ ,

$$u_{i,j} = \int_{i+\alpha}^{i+1+\alpha} \int_{j+\beta}^{j+1+\beta} U(x,y) dy dx. \quad (7)$$

Without loss of generality, we assume  $0 \leq \alpha, \beta < 1$ .

Whereas in the special case of band-limited images sampled with at least their double limiting frequency, Shannon's sampling theorem guarantees that  $u$  and  $g$  contain full information on their continuous counterparts, this can usually not be expected to hold true for natural images; thus the continuous images  $U$  and  $G$  are in fact unknown.

A good measure for the discrepancy between  $u$  and  $g$  should essentially measure the discrepancy between their continuous versions  $U$  and  $G$ . In other words, we do not want to punish reconstructions for badly aligned grids, and formulate therefore an "innocence assumption" (*in dubio pro reo – in case of doubt for the defendant*): Whatever discrepancy between two images can plausibly be attributed to different sampling, shall not enter the discrepancy measure. In particular, if a sufficiently plausible continuous-scale image  $V \equiv U \equiv G$  exists from which both  $u$  and  $g$  can be obtained by sampling, their discrepancy should be measured as zero. Notice that the exact meaning of the word "plausible" remains to be specified later.

Our considerations can be boiled down to a discrete image  $v$  of size  $(2n+1) \times (2m+1)$  whose pixels are the intersections of pixels of  $u$  and  $g$ :

$$v_{i,j} = \int_{\xi_i}^{\xi_{i+1}} \int_{\eta_j}^{\eta_{j+1}} V(x,y) dy dx, \quad (8)$$

$i = 0, \dots, 2n$ ,  $j = 0, \dots, 2m$ , where  $\xi_i = i/2$  for even  $i$  and  $\xi_i = i/2 + \alpha$  for odd  $i$ ,  $\eta_j = j/2$  for even  $j$  and  $\eta_j = j/2 + \beta$  for odd  $j$ . Note that pixel  $(i, j)$  of  $g$  covers the four pixels  $(2i, 2j)$ ,  $(2i, 2j+1)$ ,  $(2i+1, 2j)$  and  $(2i+1, 2j+1)$  of  $v$  whereas pixel  $(i, j)$  of  $u$  covers the four pixels  $(2i+1, 2j+1)$ ,  $(2i+1, 2j+2)$ ,  $(2i+2, 2j+1)$  and  $(2i+2, 2j+2)$  of  $v$ . The image  $v$  is therefore a superresolution image [15] of  $g$  and  $u$ , albeit with pixels of different sizes. In  $x$  direction grid cells of size  $\alpha$  alternate with such of size  $1 - \alpha$ , whereas in  $y$  direction the same is true with  $\beta$  and  $1 - \beta$ .

In the general situation when  $U$  and  $G$  cannot be chosen as equal, we want to retain this idea and construct a super-resolution image  $v$  that tries to reconcile the information of  $u$  and  $g$  as good as possible. The discrepancy of  $u$  and  $g$

will then be measured by combining discrepancies between  $u$  and  $v$ , and between  $v$  and  $g$ .

In the perfectly aligned case,  $\alpha = \beta = 0$ , the MSE (2) of images  $g$  and  $u$  can be combined from the MSEs between each of  $g$  and  $u$  and their average  $v := \frac{1}{2}(g+u)$  via

$$\text{MSE}(u, g) = 2(\text{MSE}(u, v) + \text{MSE}(v, g)). \quad (9)$$

Moreover, using the parallelogram identity (or by an easy combination of Cauchy-Schwarz' inequality with the arithmetic-geometric mean inequality) we see that for any other image  $v$  the right-hand side of (9) will be greater than  $\text{MSE}(u, g)$ . This motivates the following definition.

**Definition.** Let images  $u$  and  $g$  of size  $n \times m$  sampled as in (6), (7) be given. Let a class  $X$  of  $(2n+1) \times (2m+1)$ -images  $v$  sampled as in (8) be given. For each image  $v \in X$ , define  $v_u, v_g$  as the downsamplings of  $v$  to the grids of  $u$  and  $g$ , respectively. The alignment-MSE  $\text{MSE}_X$  between  $u$  and  $g$  with respect to  $X$  is defined as

$$\text{MSE}_X(u, g) = \min_{v \in X} 2(\text{MSE}(u, v_u) + \text{MSE}(v_g, g)). \quad (10)$$

Application of this definition requires, first, the specification of the class  $X$  for given images  $u, g$ . The class  $X$  essentially defines what are plausible superresolution images. Second, a minimisation method will be needed to find the minimiser. We will turn to these issues in the next section.

#### IV. SPECIFYING CONSTRAINTS

Given  $u$  and  $g$ , a superresolution image  $v$  as specified in the previous section must satisfy the equations

$$\alpha\beta v_{2i,2j} + \alpha\bar{\beta} v_{2i,2j+1} + \bar{\alpha}\beta v_{2i+1,2j} + \bar{\alpha}\bar{\beta} v_{2i+1,2j+1} = g_{i,j}, \quad (11)$$

$$\bar{\alpha}\bar{\beta} v_{2i+1,2j+1} + \bar{\alpha}\beta v_{2i+1,2j+2} + \alpha\bar{\beta} v_{2i+2,2j+1} + \alpha\beta v_{2i+2,2j+2} = u_{i,j} \quad (12)$$

for  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ , where we have used the abbreviations  $\bar{\alpha} := 1 - \alpha$ ,  $\bar{\beta} := 1 - \beta$ .

On one hand, these are just  $2nm$  equations for  $4nm + 2n + 2m + 1$  pixels of  $v$  (from which two corner pixels could be eliminated as they are neither covered by  $g$  nor by  $u$ ); additional conditions will therefore be necessary to remove this underdetermination. On the other hand, for images  $u$  and  $g$  that do not match perfectly, we expect that the equations should be satisfied only approximately, which favours smoothness. Thus, we are led to reformulate our equation system into the minimisation of an energy function

$$E(v) = S_g(v) + S_u(v) \quad (13)$$

under suitable constraints, where  $S_g$  and  $S_u$  are quadratic

error terms for the equations above,

$$S_g(v) := \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (g_{i,j} - \alpha\beta v_{2i,2j} - \alpha\bar{\beta} v_{2i,2j+1} - \bar{\alpha}\beta v_{2i+1,2j} - \bar{\alpha}\bar{\beta} v_{2i+1,2j+1})^2, \quad (14)$$

$$S_u(v) := \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (u_{i,j} - \bar{\alpha}\bar{\beta} v_{2i+1,2j+1} - \bar{\alpha}\beta v_{2i+1,2j+2} - \alpha\bar{\beta} v_{2i+2,2j+1} - \alpha\beta v_{2i+2,2j+2})^2. \quad (15)$$

Up to constant factors,  $S_g$  and  $S_u$  are just the  $\text{MSE}(g, v_g)$  and  $\text{MSE}(u, v_u)$  from the alignment-MSE definition.

Let us therefore now discuss possible constraints for this minimisation problem. These constraints will constitute the class  $X$  of images to minimise over that appeared in the definition of the alignment-MSE.

Note first that in the equations (11), (12) for subsequent indices  $i$  or  $j$  the two input images  $u$  and  $g$  alternate. This suggests that for images  $u$  and  $g$  that do not perfectly match, solutions of (11), (12) are likely to develop oscillating patterns like stripes of alternating intensity or checkerboard structures, so the discrepancy between  $u$  and  $g$  can be translated to the image boundary where the first and last row and column of  $v$  are linked only to one of the input images and therefore provide degrees of freedom that can absorb the discrepancy. In extreme, this could mean that even for completely mismatching  $u$  and  $g$  highly oscillatory images  $v$  might exist that fulfil (11), (12) without any error. Such solutions should be rejected by a suitable class  $X$ .

In order to prevent  $v$  from developing strong high-frequency structures, a natural requirement could be that  $v$  should be essentially interpolating; thus each pixel intensity  $v_{i,j}$  should be in the interval bounded by the intensities  $g_{\lfloor i/2 \rfloor, \lfloor j/2 \rfloor}$ ,  $u_{\lfloor (i-1)/2 \rfloor, \lfloor (j-1)/2 \rfloor}$  of the two input pixels it is linked to by (11), (12). Whilst conceptually elegant and free of additional parameters, this constraint turns the minimisation of (13) into a quadratic minimisation problem on a highly nonconvex domain. We aim therefore at relaxing this constraint to a convex regularisation that warrants a unique solution as well as a practical minimisation procedure.

We extend therefore the energy function (13) to

$$E(v) = S_g(v) + S_u(v) + \gamma T(\nabla v) \quad (16)$$

where  $T$  is a regulariser that depends on the derivatives  $\nabla v = (v_x, v_y)$  of  $v$  approximated by finite differences, and  $\gamma > 0$  is a regularisation weight.

With regard to the quadratic nature of the mean square error to be measured, a Whittaker-Tikhonov regularisation

$$T(\nabla v) := \sum_{i,j} |\nabla v|^2 \quad (17)$$

lends itself as a natural candidate, which yields a convex quadratic minimisation problem, also removing completely the non-uniqueness of the original equations. Minimisers can efficiently be computed using standard iterative solution methods for the linear system of minimality conditions.

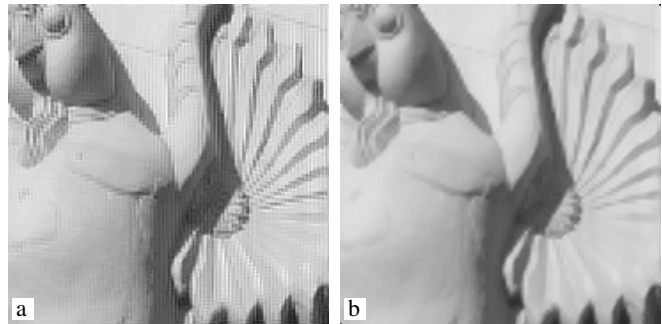


Fig. 3. (a) Superresolution image created in aligning the images from Fig. 2(a) and (d) with Whittaker-Tikhonov regularisation,  $\gamma = 0.003$ . Alignment-MSE measurement with this superresolution image yields a PSNR of 46.05 dB. – (b) Same with TV regularisation,  $\gamma = 0.03$ , yielding a PSNR of 29.92 dB.

A further candidate is total variation

$$T(\nabla v) := \sum_{i,j} |\nabla v|. \quad (18)$$

To find minimisers with this regularisation, one can use, e.g., a gradient descent approach where the regularisation is realised via a locally analytic scheme related to single-scale Haar wavelet shrinkage; we use here a variant of the scheme from [19] adapted to the unequal pixel sizes of  $v$ .

As a general rule, in order to just remove the underdeterminedness of the equation system (11), (12), it is desirable to keep the regularisation weight  $\gamma$  rather small.

## V. EXPERIMENTS

We evaluate the regularised superresolution alignment procedure from the preceding two sections by the test case from Section II. Starting with Whittaker-Tikhonov regularisation, we observe that for large regularisation weight such as  $\gamma = 0.3$  fairly precise estimates for the displacement can be obtained. However, the superresolution images in this case are severely blurred, leading to overestimation of alignment-MSE and low PSNR. For example, the resulting PSNR for the images from Fig. 2(a) and (d) is 28.61 dB. On the other hand, reducing the regularisation parameter to  $\gamma = 0.003$  yields extremely high PSNR estimates, e.g. 46.05 dB for the same two images. The reason is that the superresolution images are far away from interpolating between  $u$  and  $g$ , showing unnatural oscillations, see Fig. 3(a). In contrast, TV regularisation yields plausible results over a wide range of regularisation parameters, see the exemplary superresolution image in Fig. 3(b). For a more detailed evaluation we focus therefore on TV regularisation.

We measure first reconstruction errors for the known exact displacements, see column (a) of Table III. Next we estimate the displacements using the TV-regularised error measure itself, see column (b). Once more the minimisation is done by a grid search with  $x$  and  $y$  displacements varying from  $-1$  to  $+1$  in 0.01 steps. The TV regularisation weight  $\gamma$  is set to 0.03. As the application of the superresolution alignment in this brute-force minimisation is computationally expensive, we add a third scenario, column (c), in which a faster

variant of the superresolution alignment with Whittaker-Tikhonov regularisation and large regularisation parameter  $\gamma = 0.3$  is used for the displacement estimation, followed by the actual MSE/PSNR computation with TV regularisation and  $\gamma = 0.03$ . The latter method gives in a few cases a slightly lower PSNR than the ground-truth displacement, but otherwise approximates the previous scenario well.

It is evident that the variation of PSNR measures is reduced to about half with respect to the measurements with bicubic interpolation, both in terms of the amplitude between maximal and minimal PSNR and the standard deviation. With an amplitude of 1.2dB it is close to the variation of the shift-free subset of 0.7dB as reported in Section II.

## VI. SUMMARY AND OUTLOOK

In this paper we have studied the reliability of MSE/PSNR measurements for the quality assessment of blind deconvolution results, where the necessity arises to compare images that may be shifted relative to each other by non-integer displacements. An experimental study of simple alignment procedures with bilinear and bicubic interpolation showed that it introduces substantial deviations into the discrepancy measures in question. Comparisons of blind deconvolution methods should therefore not be based on such procedures. As an attempt to overcome this difficulty, we have designed a superresolution-based error measurement procedure that can substantially reduce the variations in MSE/PSNR estimates induced by the alignment step, leaving error margins that are closer to the uncertainty in shift-free cases.

In future work, these tests will have to be extended to more test cases. The applicability of the proposed procedure to other error measures such as MSSIM [17] or perceptual similarity measures [13] will be studied. Further analysis will be devoted to the observed variation of error measures among the shift-free reconstructed images. It will also be of interest to include the PSF into the displacement estimation.

Furthermore, the proposed approach will be used for comparisons between blind deconvolution methods. Taking into account the results from the present contribution and

the envisioned more extensive studies will help to assess the significance of method differences in such work.

## REFERENCES

- [1] M. S. C. Almeida and L. B. Almeida, "Blind and semi-blind deblurring of natural images," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 36–52, 2010.
- [2] L. Bar, N. Sochen, and N. Kiryati, "Variational pairing of image segmentation and blind restoration," in *Computer Vision – ECCV 2004, Part II*, ser. Lecture Notes in Computer Science, T. Pajdla and J. Matas, Eds. Berlin: Springer, 2004, vol. 3022, pp. 166–177.
- [3] T. F. Chan and C. K. Wong, "Total variation blind deconvolution," *IEEE Transactions on Image Processing*, vol. 7, pp. 370–375, 1998.
- [4] T. F. Chan, A. M. Yip, and F. E. Park, "Simultaneous total variation image inpainting and blind deconvolution," *International Journal of Imaging Systems and Technology*, vol. 15, no. 1, pp. 92–102, 2005.
- [5] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *Proc. SIGGRAPH 2006*, New York, NY, July 2006, pp. 787–794.
- [6] V. Katkovnik, D. Paliy, K. Egiazarian, and J. Astola, "Frequency domain blind deconvolution in multiframe imaging using anisotropic spatially-adaptive denoising," in *14th European Signal Processing Conference (EUSIPCO 2006)*. Florence, Italy: EURASIP, 2006.
- [7] R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, and S. Harmeling, "Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database," in *Computer Vision – ECCV 2012, Part VII*, ser. Lecture Notes in Computer Science, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Berlin: Springer, 2012, vol. 7578, pp. 27–40.
- [8] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1964–1971.
- [9] —, "Efficient marginal likelihood optimization in blind deconvolution," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 2657–2664.
- [10] D. Li, R. M. Mersereau, and S. Simske, "Blind image deconvolution through support vector regression," *IEEE Transactions on Neural Networks*, vol. 18, no. 3, pp. 931–935, 2007.
- [11] G. Liu, S. Chang, and Y. Ma, "Blind image deblurring using spectral properties of convolution operators," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5047–5056, 2014.
- [12] P. Moser and M. Welk, "Robust blind deconvolution using convolution spectra of images," in *1st OAGM-ARW Joint Workshop: Vision Meets Robotics*, K. Niel, P. M. Roth, and M. Vincze, Eds. Wels, Austria: Österreichische Computer-Gesellschaft, 2016, pp. 69–78.
- [13] R. Reisenhofer, S. Bosse, G. Kutyniok, and T. Wiegand, "A Haar wavelet-based perceptual similarity index," arXiv.org, Tech. Rep. cs:1607.06140, 2016.
- [14] K. Schelten, S. Nowozin, J. Jancsary, C. Rother, and S. Roth, "Interleaved regression tree field cascades for blind image deconvolution," in *IEEE Winter Conference on Applications of Computer Vision*, 2015, pp. 494–501.
- [15] J. Tian and K.-K. Ma, "A survey on super-resolution imaging," *Signal, Image and Video Processing*, vol. 5, pp. 329–342, 2011.
- [16] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Transactions on Image Processing*, vol. 7, pp. 813–824, 1998.
- [17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [18] M. Welk, "A robust variational model for positive image deconvolution," *Signal, Image and Video Processing*, vol. 10, no. 2, pp. 369–378, 2016.
- [19] M. Welk, J. Weickert, and G. Steidl, "A four-pixel scheme for singular differential equations," in *Scale Space and PDE Methods in Computer Vision*, ser. Lecture Notes in Computer Science, R. Kimmel, N. Sochen, and J. Weickert, Eds. Berlin: Springer, 2005, vol. 3459, pp. 585–597.
- [20] L. Xu, S. Zheng, and J. Jia, "Unnatural  $L_0$  sparse representation for natural image deblurring," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1107–1114.
- [21] Y.-L. You and M. Kaveh, "Anisotropic blind image restoration," in *Proc. 1996 IEEE International Conference on Image Processing*, vol. 2, Lausanne, Switzerland, Sept. 1996, pp. 461–464.

TABLE III

STATISTICS OF DISPLACEMENT MISESTIMATIONS  $\delta_x$ ,  $\delta_y$  AND PSNR FOR 256 RECONSTRUCTED IMAGES WITH SUPERRESOLUTION-BASED ALIGNMENT WITH TV REGULARISATION, (A) USING EXACT DISPLACEMENTS, (B) ESTIMATING DISPLACEMENTS BY MSE MINIMISATION WITH TV REGULARISATION, (C) ESTIMATING DISPLACEMENTS BY MSE MINIMISATION WITH WHITTAKER-TIKHONOV REGULARISATION.

Setting	(a)	(b)	(c)
$\max \delta_x $		0.09	0.09
std. dev. $\delta_x$		0.037	0.033
$\max \delta_y $		0.08	0.08
std. dev. $\delta_y$		0.031	0.036
min PSNR	29.38	29.40	29.40
max PSNR	30.47	30.63	30.46
(max–min) PSNR	1.09	1.23	1.06
mean PSNR	29.93	29.98	29.92
std. dev. PSNR	0.240	0.263	0.236