This ICCV2013 paper is the Open Access version, provided by the Computer Vision Foundation.
The authoritative version of this paper is available in IEEE Xplore.

Deblurring by Example using Dense Correspondence

Yoav HaCohen
Hebrew University
Jerusalem, Israel
yoav.hacohen@mail.huji.ac.il

Eli Shechtman
Adobe Research
Seatle, USA
elishe@adobe.com

Dani Lischinski
Hebrew University
Jerusalem, Israel
danix@cs.huji.ac.il

Abstract

This paper presents a new method for deblurring photos using a sharp reference example that contains some shared content with the blurry photo. Most previous deblurring methods that exploit information from other photos require an accurately registered photo of the same static scene. In contrast, our method aims to exploit reference images where the shared content may have undergone substantial photometric and non-rigid geometric transformations, as these are the kind of reference images most likely to be found in personal photo albums.

Our approach builds upon a recent method for example-based deblurring using non-rigid dense correspondence (NRDC) [11] and extends it in two ways. First, we suggest exploiting information from the reference image not only for blur kernel estimation, but also as a powerful local prior for the non-blind deconvolution step. Second, we introduce a simple yet robust technique for spatially varying blur estimation, rather than assuming spatially uniform blur. Unlike the above previous method, which has proven successful only with simple deblurring scenarios, we demonstrate that our method succeeds on a variety of real-world examples. We provide quantitative and qualitative evaluation of our method and show that it outperforms the state-of-the-art.

1. Introduction

Photographs often exhibit blur caused by camera defocus, camera motion, or motion in the scene. *Blind deblurring*, also known as *blind deconvolution*, refers to the problem of recovering a sharp image from a blurry one when the exact parameters of the blur operator are not known. Without any prior or additional information, this problem is inherently ill-posed, as there are many possible combinations of blur kernels and sharp images that can explain a given blurry image. In this paper we address a challenging variant of blind deblurring with unknown *spatially varying* camera motion blur, while assuming the availability of a sharp reference image containing *some* shared content under unknown geometric and photometric transformations.

While recent single-image approaches based on general priors or edge-based techniques [4, 6, 18, 21, 17] have shown significant progress, blind deblurring remains a very hard and ill-posed problem [20]. General assumptions may result in an inaccurate blur kernel estimation and/or an incorrect deblurred result. For example, the sparse derivatives prior suggested by Levin *et al.* [19] is often used in the (non-blind) deconvolution step to overcome artifacts due to an inaccurately estimated kernel. However, being based on generic natural image statistics, this prior biases the results towards piecewise smooth solutions and thus tends to reduce fine detail.

Several approaches perform deblurring by leveraging multiple photos of the scene [22, 27, 14]. While these methods demonstrate the benefit of having additional data, they require accurately registered photos that simultaneously capture exactly the same static scene. Although people often shoot several photos in a succession, unless the photos were taken intentionally to meet the requirements of these methods, suitable input image sets are unlikely to be found in personal photo albums.

Recently HaCohen *et al.* [11] presented the Non-Rigid Dense Correspondence (NRDC) method for simultaneously recovering a partial dense correspondence and a color transformation between pairs of images with shared content. NRDC was shown to be highly effective for finding large matching regions in typical personal photo collections. One of the applications demonstrated in [11] was deblurring by example: given a pair of images, where one image is sharp while another is blurry, NRDC was used to estimate a blur kernel simultaneously with the correspondence and the color transformation. However, this approach has been successfully demonstrated only on simple synthetic deblurring scenarios, involving simple blur kernels and no added noise.

Our method extends the above approach to more realistic scenarios, including real-world blur kernels, noise, and real-world sharp/blurry image pairs with spatially vary-

ing blur. Similarly to [11], we employ an iterative optimization scheme that alternates between finding correspondences, estimating the blur kernel, and recovering the sharp image with non-blind deconvolution. However, in order to increase the accuracy and robustness of our method, we introduce two key modifications: First, rather than exploiting the sharp reference only for kernel estimation, we also use this reference as a strong local prior for the non-blind deconvolution step. Second, we suggest a new fast and robust non-uniform blur kernel estimation method, which reduces the effect of outliers resulting from inaccuracies at the other steps.

After discussing relevant previous work (Section 2) and presenting our deblurring algorithm (Section 3), we evaluate our method and compare it to the state of the art deblurring methods (Section 4).

2. Background and Related Work

Generally, the problem of removing uniform camera motion blur is modeled as the problem of estimating a sharp *latent image* $\mathbf{x}$ given a blurry image $\mathbf{y}$, related by an unknown low-dimensional linear blur kernel $\mathbf{k}$ of size d with some additive noise $\mathbf{n}$ (all bold notations vectorized):

$$\mathbf{y} = \sum_{i=1}^{d} \mathbf{k}_i P_i \mathbf{x} + \mathbf{n} \tag{1}$$

where $P_i\mathbf{x}$ are geometric projections of the latent image. Each projection represents the image that the camera "saw" for the $\mathbf{k}_i$ portion of the shutter opening. If the classic uniform convolution model is assumed [9, 19, 27, 23, 16, 4, 21], $\{P_i\}$ represent translations in matrix form, and the model can be written as a uniform convolution of $\mathbf{k}$ with $\mathbf{x}$. Recent works [10, 26, 13, 18] use richer models where $\{P_i\}$ correspond to a subset of geometric transformations that better approximate the 6D camera motion (rotations and translations). In this paper, we also assume a spatially non-uniform blur model. Since both $\mathbf{x}$ and $\mathbf{k}$ are unknown, the problem is inherently ill-posed, and therefore additional prior or information must be used.

Recent single-image approaches for blind deblurring have shown significant progress by using general image priors [9, 21, 18] and edge-based techniques [16, 4, 6]. In order to recover both $\mathbf{x}$ and $\mathbf{k}$ most of these methods alternate between two main steps: First, updating the estimated sharp latent image $\mathbf{x}$ (non-blind deconvolution), and second, updating the blur kernel $\mathbf{k}$. These two steps typically involve solving equations of the form:

$$\mathbf{x} = \arg\min_{\mathbf{x}} \|A_{\mathbf{k}}\mathbf{x} - \mathbf{y}\| + \rho_L(\mathbf{x}) \tag{2}$$

$$\mathbf{k} = \arg\min_{\mathbf{k}} \|B_{\mathbf{x}}\mathbf{k} - \mathbf{y}\| + \rho_K(\mathbf{k}, \mathbf{x}) \tag{3}$$

where $A_{\mathbf{k}} = \sum_{i=1}^{d} \mathbf{k}_i P_i$ is the blur matrix, $B_{\mathbf{x}}$ is a matrix whose i-th column is $P_i\mathbf{x}$, and ρ_L and ρ_K are priors for the latent image and the kernel, respectively.

For ρ_K, Yuan *et al.* [27] and Cho and Lee [4] use Tikhonov regularization on the kernel ($\rho_K = \|\mathbf{k}\|_2$), Shan *et al.* [23] and Krishnan *et al.* [18] use $\rho_K = \|\mathbf{k}\|_1$, and Levin *et al.* [21] use the covariance around the estimated image (which depends both on $\mathbf{x}$ and $\mathbf{k}$) and a sparsity prior on the kernel ($\rho_K = \sqrt{\|\mathbf{k}\|_{0.5}}$).

For the latent image prior ρ_L, researchers commonly use a Gaussian prior [23, 4], or a sparsity prior [21, 18]. Cho *et al.* [7] suggested a deblurring method using region-specific priors. The method estimates a prior for each region based on its texture characteristics at a coarser scale, assuming scale invariance of such characteristics. Recently, priors based on patch-banks have been proven effective in reducing ringing artifacts [28, 24], but they might still lose high frequency texture details due to the limited number of example patches.

Despite the recent progress in single-image approaches, their general underlying assumptions are commonly violated resulting in strong visual artifacts. Methods like [23, 4, 21, 18, 7] that assume an objective gradient distribution (globally or locally), may fail when the gradient distribution of the true result is not as expected. Thus, their results are often not suitable for practical personal use, without further manual processing or image-specific tuning of many parameters.

Other approaches assume the existence of another accurately registered image of the same static scene, but blurred by a different kernel [22, 3, 1, 14] or containing noise [27], for estimating a blur kernel. Yuan *et al.* [27] use a noisy image taken from the same viewpoint as a prior for the non-blind deconvolution step, but only for recovering the low-frequencies of the latent image, leaving the high-frequencies prone to ringing and noise. Rav-Acha and Peleg [22] simultaneously deconvolve two registered images of the same static scene, each blurred by a different directional blur kernel, assuming the results should be equal (thus each image serves as a prior for the other). These approaches have shown the benefit of using additional images for image deblurring. However, they require accurately aligned input images which are not typically available in personal photo collections.

Several methods have attempted to abolish the requirement of registered images for deblurring. Ancuti *et al.* [2] use SIFT features to match between a blurry and a sharp image of a static scene. Joshi *et al.* [15] use photos of faces of the same person with a similar pose. HaCohen *et al.* [11] simultaneously deblurs and computes a partial dense correspondence between the blurry input image and a sharp reference, where the resulting correspondence is more dense and robust than previous methods. Cho *et al.* [5] removes blur

in video frames due to camera shake using patches sampled from nearby sharp frames. While our local prior is also based on reconstruction using sharp patches, our method interleaves finding correspondence, local prior and kernel estimation in an optimization framework and thus can handle more complex motions and blur kernels.

As already explained in Section 1, we introduce two important improvements to the approach described in [11], making it applicable to a much wider variety of deblurring scenarios and real-world image pairs. In particular, to the best of our knowledge, our method is the first to use an additional image as a local prior when estimating the latent sharp image, without requiring an accurate full registration between the images.

3. Method

Similarly to HaCohen *et al.* [11], we iteratively alternate between computing a dense correspondence, estimating the kernel, and estimating the latent image, while proceeding in a coarse-to-fine manner. However, there two crucial differences with respect to [11]: (i) we use the sharp reference image not only for blur kernel estimation, but also as a local prior for latent image recovery (Section 3.1); (ii) we robustly estimate a spatially varying blur kernel instead of a uniform one (Section 3.2).

Each iteration begins by applying the basic NRDC matching step [11], i.e., finding an approximate nearest neighbor field followed by aggregation of coherent regions, to obtain a dense correspondence between the sharp reference image $\mathbf{r}$ and the current latent image estimate $\mathbf{x}$. The resulting correspondence consists of: (i) a geometric pixel-wise mapping M from $\mathbf{r}$ to $\mathbf{x}$; (ii) an associated confidence map $\mathbf{w}$ that indicates the reliability of this mapping at each pixel using values between zero and one; and (iii) a parametric color transformation C between $\mathbf{x}$ and $\mathbf{r}_M$, where $\mathbf{r}_M$ is the reference image $\mathbf{r}$ geometrically transformed by M. Next, we estimate the blur kernel $\mathbf{k}$ from the pair of sharp and blurry images given by $\mathbf{r}_M$ and $C(\mathbf{y})$ (section 3.2). Finally, we use the estimated kernel $\mathbf{k}$, and a local prior given by the partial reconstruction of the latent image, $C^{-1}(\mathbf{r}_M)$, to update our latent image estimate $\mathbf{x}$ (section 3.1). Several such iterations take place at each scale. These steps are summarized in Algorithm 1.

To update M and $\mathbf{w}$ in each iteration (line 6 of Alg. 1) we use the previous estimates of M and $\mathbf{x}$ to initialize and to constrain the matching step. However, a known limitation of NRDC (shared by many other matching algorithms) is the lack of reliable matches inside large smooth regions. Note that in order to generate the kernel estimation equation (3) we need an entire neighborhood around each pixel of $\mathbf{r}_M$. Thus, while a small number of matches is not an issue for other applications of NRDC, in our case, we may be left with too few usable pixels for estimating the blur kernel. In order to avoid this problem, we fill smooth unreliably matched regions in M by interpolating the missing values from nearby reliably matched values. Specifically, we detect regions of adjacent pixels where the matching confidence $\mathbf{w}$ is low. We consider such a region as smooth if it contains only pixels with gradient magnitudes below some threshold. We then interpolate the mapping values inside such regions from nearby reliable matches by scattered data interpolation using the method of [8]. The effect of the hole filling step is demonstrated in Figure 1.

Algorithm 1 *DeblurByExample*$(\mathbf{r}, \mathbf{y})$

1: Initialize $\mathbf{x} = \mathbf{y}$
2: $[M, \mathbf{w}, C] = \mathit{Match}_{\mathrm{NRDC}}(\mathbf{r}, \mathbf{x})$ at the coarsest scale
3: **for** scale = coarse **to** fine **do**
4: **for** several iterations **do**
5: $\mathbf{k} = \mathit{EstimateKernel}(\mathbf{r}_M, C(\mathbf{y}))$ (Sec. 3.2)
6: $\mathbf{x} = \mathit{Deconvolve}(\mathbf{y}, \mathbf{k}, C^{-1}(\mathbf{r}_M), \mathbf{w})$ (Sec. 3.1)
7: $[M, \mathbf{w}] = \mathit{Match}_{\mathrm{NRDC}}(\mathbf{r}, \mathbf{x})$ at the current scale
8: Fill smooth holes in M
9: Fit a color transformation C between $\mathbf{x}$ and $\mathbf{r}_M$
10: **end for**
11: upscale M and $\mathbf{w}$
12: **end for**

Note also that unlike many other coarse-to-fine deblurring methods (e.g., [11, 4, 21, 18, 26]), we *do not* upscale the kernel $\mathbf{k}$ when switching from a coarser scale to a finer one. We found that any small interpolation/upsampling error in the kernel might result in large deconvolution artifacts. Instead, we upscale the mapping M and use it to recompute $\mathbf{k}$ and $\mathbf{x}$ at the next, finer, scale.

3.1. Latent image estimation

Given the blur kernel $\mathbf{k}$, a naïve way of obtaining the deblurred image would be to invert the kernel by solving Eq. 2. This often results in severe artifacts, such as ringing, because $A_{\mathbf{k}}$ is usually not well-conditioned; other sources for such artifacts might be inaccuracies in the estimate of $\mathbf{k}$, presence of noise, or other violations of the blurry image formation model.

To reduce these artifacts, the popular deconvolution approach of Levin *et al.* [19] introduces into the equation a *sparse prior* motivated by natural image statistics:

$$\mathbf{x} = \arg\min_{\mathbf{x}} \|A_{\mathbf{k}}\mathbf{x} - \mathbf{y}\|^2 + \lambda \sum_f \|T_f \mathbf{x}\|^{0.8}, \qquad (4)$$

where $\{T_f\}$ are Toeplitz (convolution) matrices corresponding to a set of derivative filters $\{f\}$. This prior penalizes small gradients, thereby encouraging piecewise smooth solutions.

However, while being commonly used, the sparse prior is still too generic and often overcomes ringing artifacts only

(a) blurry input (b) reference (c) with holes (d) after hole filling (e) final result

Figure 1. (c) and (d) compare the reconstructed image $\mathbf{r}_M$ with and without step 8 (hole filling) of Algorithm 1. The confidence map $\mathbf{w}$ (not shown) has high values where the colors in (d) are reconstructed and zero values otherwise; (e) is our final deblurred result.

at the price of overly (piecewise) smoothed results. Our approach is to leverage the existence of the sharp reference image $\mathbf{r}$, and the availability of a dense mapping M between $\mathbf{r}$ and the blurry image $\mathbf{y}$, to augment the generic sparse prior with a local non-parametric one, which we refer to as the *reconstruction prior*. This is only done in regions where we believe that we can reliably estimate $\mathbf{x}$ by the transformed reference image $\mathbf{r}_M$. Specifically, we use a linear combination of the two priors, where the relative weight of each prior is determined by the confidence associated with the reconstruction of each pixel.

Formally, we add a term that minimizes the difference between the result and the reconstructed image:

$$\begin{aligned} \mathbf{x} &= \arg\min_{\mathbf{x}} \|A_{\mathbf{k}}\mathbf{x}-\mathbf{y}\|^2 \\ &+ \lambda_1(I-D)\textstyle\sum_f \|T_f\mathbf{x}\|^{0.8} \\ &+ \lambda_2 D\textstyle\sum_{f'} \|T_{f'}\mathbf{x} - T_{f'}C^{-1}(\mathbf{r}_M)\|^2 \end{aligned} \quad (5)$$

where D is a diagonal matrix whose entries represent the reliability of the local prior at each pixel. Note, that beyond the obvious effect of this local prior on the pixels where it is applied, it also affects their neighborhood due to the mutual influence of adjacent pixels introduced by the first and the second term of Eq. 5.

The confidence map $\mathbf{w}$ associated with the mapping M is determined by the consistency of the matching, but it does not guarantee that the reconstruction $\mathbf{r}_M$ is consistent with the blurry input image (i.e., some elements of $\|A_{\mathbf{k}}C^{-1}(\mathbf{r}_M)-\mathbf{y}\|$ may be large). Therefore, we refine $\mathbf{w}$ by estimating the image regions where we can reliably estimate $\mathbf{x}$ using the transformed reference image $\mathbf{r}_M$. This is done by defining the i-th element of the diagonal of D based on $\mathbf{r}_M$, C, $\mathbf{y}$ and $\mathbf{k}$ as follows:

$$D_{i,i} = \mathbf{w}_i \exp\left(-\frac{\|A_{\mathbf{k}}C^{-1}(\mathbf{r}_M)_i - \mathbf{y}_i\|^2}{\sigma_{\mathbf{r}}}\right) \quad (6)$$

where $\sigma_{\mathbf{r}}$ is a constant.

In practice, we use the gradients of $\mathbf{y}$ and $C^{-1}(\mathbf{r}_M)$ rather than their intensity values in Eq. 6, having found it more robust to errors in C. A value of $\sigma_{\mathbf{r}} = 0.01$ was used to produce the results in the paper. To solve Eq. 5, we use iteratively reweighed least squares (IRLS), with Conjugate Gradient for solving the inner iterations. We use 15 IRLS iterations and a maximum of 25 CG iterations.

3.2. Non Uniform Kernel Estimation

3.2.1 Kernel estimation

If the estimate of the latent image $\mathbf{x}$ is correct up to some additive Gaussian noise, the optimal kernel $\mathbf{k}$ can be recovered by minimizing an objective function similar to Eq. 3 [4, 21, 18, 26, 10]. In single image deblurring methods, $\mathbf{x}$ is initially unknown so it is often approximated using edge prediction techniques ([4, 16, 6]) that rely on the existence of strong edges at multiple orientations. $\rho_K(\mathbf{k},\mathbf{x})$ is a prior on the kernel values and may be a function of the image $\mathbf{x}$.

When a sharp reference image $\mathbf{r}$ exists and a correspondence M is established, we can replace $\mathbf{x}$ with $\mathbf{r}_M$ at regions of high correspondence confidence as follows:

$$\mathbf{k} = \arg\min_{\mathbf{k}} \|W(B_{\mathbf{r}_M}\mathbf{k} - C(\mathbf{y}))\|^2 + \rho_K(\mathbf{k},\mathbf{x}) \quad (7)$$

where we also replace $\mathbf{y}$ by $C(\mathbf{y})$ (to compensate for the photometric differences between $\mathbf{r}$ and $\mathbf{y}$), and weight the differences with a diagonal matrix W whose main diagonal is the correspondence confidence $\mathbf{w}$.

We will now describe our blur model ($\{P_i\}$) and the prior $\rho_K(\cdot)$ that we use to regularize the kernel $\mathbf{k}$.

3.2.2 Non-uniform blur model

Many deblurring algorithms assume that the blur function (i.e. $A_{\mathbf{k}} = \sum_i \mathbf{k}_i P_i$) is a translation invariant (convolution) matrix. Recent work by Whyte *et al.* [26] and Gupta *et al.* [10] showed that general blur caused by 6D camera motion cannot be accurately represented by a translation invariant kernel and proposed spatially varying approaches. However, a recent review of deblurring algorithms [17] showed that the uniform translation-invariant blur model performs generally better than the spatially-varying models for large kernels, and not much worse when strong camera rotation was involved. Methods like [26, 10] are inherently less robust due to the large number of unknowns involved in the discretization of the multi-dimensional camera motion.

Motivated by this review and the approximation in [13, 12], we suggest a simpler model where each block has its own translation-invariant kernel but this kernel must not be too different from those of the adjacent blocks. Thus, we

Figure 2. An example for an improvement by using our non-uniform model applied on the top example in Fig. 5 (zoom on lower left corner). Left column: Our method with a single block (uniform model). Right column: Our method applied with the default 12 blocks.

model the blur as a set of convolution kernels, each estimated inside a different block (tile) of the image, while corresponding coefficients of kernels from adjacent blocks are regularized to be similar. Formally, we define the blur function as $A_{\mathbf{k}} = \sum_{b,i,j} \mathbf{k}_{b,i,j} P_{b,i,j}$ where $P_{b,i,j}$ denotes a matrix that offsets the image block b by $(i - c_x, j - c_y)$ pixels while zeroing the rest of the image (c_x and c_y are defined as half of the kernel dimensions).

Inspired by [12], the similarity between adjacent kernels is achieved by using the following prior in Eq. 7:

$$\rho_K(\mathbf{k}, \mathbf{x}) = \lambda_3 \sum_{b,i,j} \sum_{b' \in \mathcal{N}(b)} |\mathbf{k}_{b,i,j} - \mathbf{k}_{b',i,j}|^2 \qquad (8)$$

where $\mathcal{N}(b)$ is the set containing the indices of the blocks adjacent to b, and λ_3 is a weighting constant.

Our model's expressiveness depends on the number of blocks being used, and the weight of the connectivity prior between them. We found that using 3×4 blocks with $\lambda_3 = 5$ achieves a good balance between robustness and flexibility in all of the examples we have tried. Figure 2 demonstrates the contribution of our non-uniform model on one of our real world examples.

4. Evaluation and Results

We implemented our method using a combination of `matlab` and C++. The entire deblurring process (including the non-blind deconvolution step) takes about 1–2 minutes for 1024×768 images on a 2.3GHz Intel Core i7 (2820qm) MacBook Pro. Note that in addition to the deblurred result, the process also yields a partial dense correspondence between the two images, as well as a global color transformation between them.

We evaluate our method using the error metric proposed by Levin *et al.* [20] for quantitatively comparing deblurring methods. The metric assumes that the ground-truth blur kernel $\mathbf{k}_{gt}$ and sharp image $\mathbf{x}_{gt}$ are both known, and computes the error between the deblurring result $\mathbf{x}_{out}$ and $\mathbf{x}_{gt}$. This error is normalized by the error between the deconvolution with the ground-truth kernel $\mathbf{x}_{\mathbf{k}_{gt}}$ and $\mathbf{x}_{gt}$, resulting in the error ratio: $\|\mathbf{x}_{out} - \mathbf{x}_{gt}\| / \|\mathbf{x}_{\mathbf{k}_{gt}} - \mathbf{x}_{gt}\|$. An error ratio of 1 is ideal, while higher ratios indicate less accurate results. Furthermore, Levin *et al.* observed that error ratios above 2 are often visually implausible.

We use the 8 ground truth kernels extracted by Levin *et al.* [20] from real-world blurry photos. Although a dataset of 32 test images is also provided, we could not use this dataset since it does not include the reference images required by our approach. Hence, we used our own dataset, constructed from 5 color image pairs. These pairs were collected from personal photo albums and were not deliberately captured with our method or experiment in mind. One image from each pair was blurred with each of the 8 kernels, resulting in 40 test images, while the remaining 5 images serve as the sharp references. We added Gaussian noise with $\sigma = 1\%$ to each of the blurry images.

Figures 3 and 4 compare our method with three state-of-the-art single-image deblurring methods. The two-image deblurring methods of [22, 27] were not tested as they both require a pair of registered images of the same static scene, and [22] was designed for 1D motion blur only. The NRDC-based deblurring method of HaCohen *et al.* [11] was tested as well (using the authors' original implementation). However, it failed to find any correspondences on almost all of the 40 blurry test inputs, which were generated using complex real-world kernels, and therefore its results are not included in Figures 3 and 4.

Following [20], in Figure 3 we plot the cumulative error ratio (CER), where higher accuracy methods correspond to higher curves. As demonstrated by these plots, our method outperforms the other three methods that we compared to. Note that for most of the tested images the error ratio of our deblurring results is below 2. Figure 4 shows a visual comparison of our results to those of the other three methods on several different test cases from our dataset.

We also tested our method on several real-world image

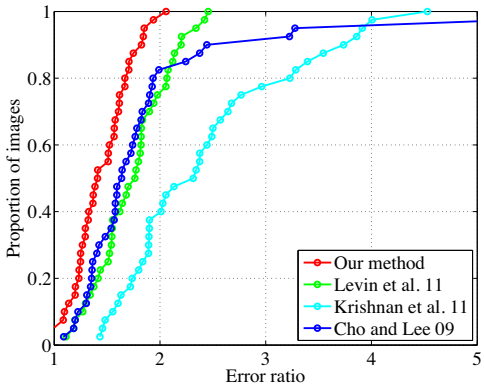

Figure 3. A quantitative comparison with several state-of-the-art single-image deblurring methods: Levin *et al.* [21], Krishnan *et al.* [18], and Cho and Lee [4]. Our method outperforms the other methods.

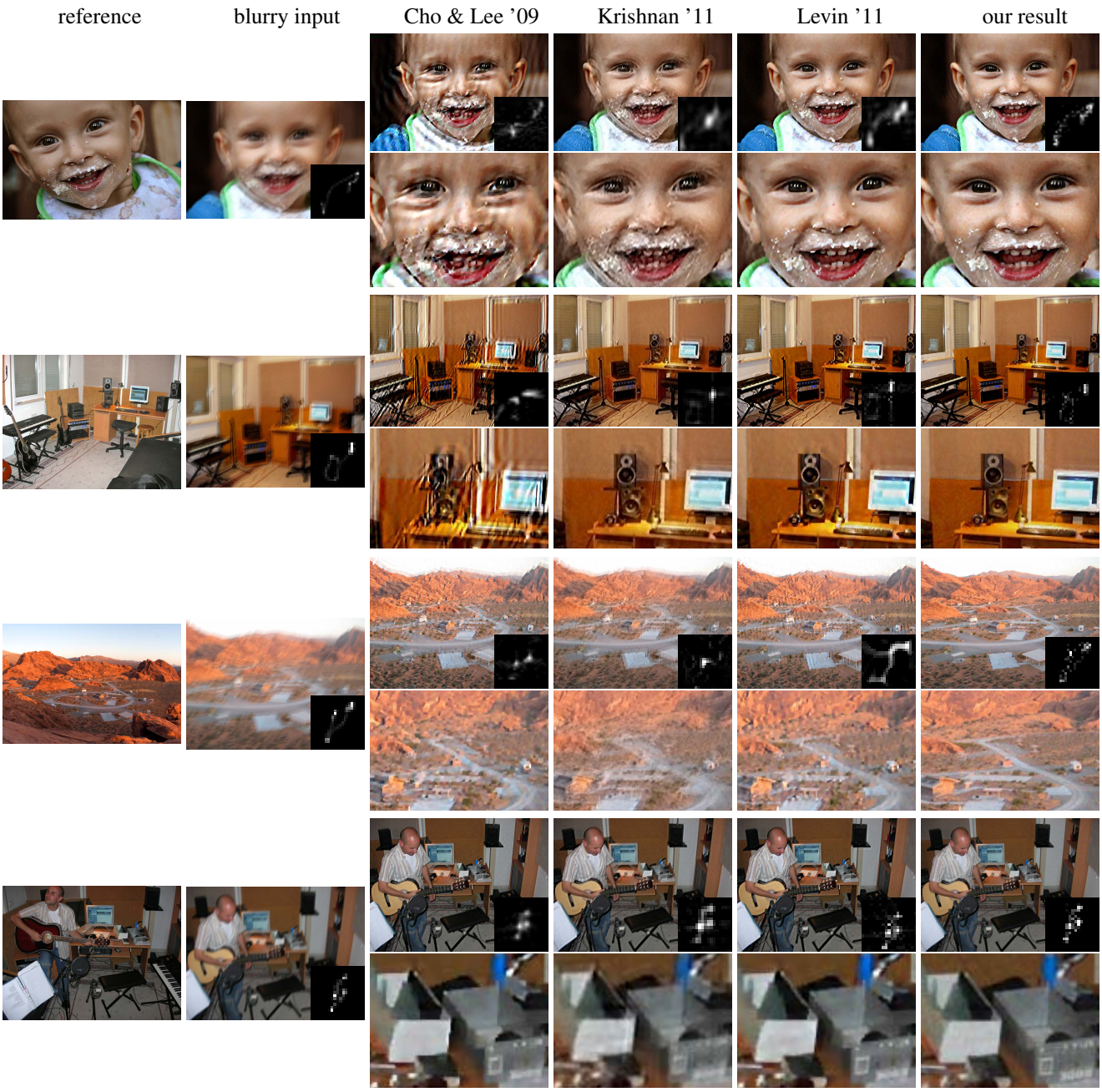

Figure 4. Example results from our test dataset. From left to right: sharp reference; blurry input image (with ground truth kernel superimposed); Cho and Lee [4]; Krishnan *et al.* [18]; Levin *et al.* [21]; our result. The odd rows show a deblurred result (with the recovered blur kernel) and the even rows show an enlarged portion of the deblurred results.

pairs, where one image exhibits significant blur, but a sharp reference is available that shares a significant portion of the visible content. Figure 5 shows our results on five such examples and compares them to three state-of-the-art single image deblurring methods. We also show the blur kernels recovered by each of the methods superimposed over the deblurred results. For our method, as well as that of Whyte *et al.* [25], which recover a non-uniform blur kernel, we superimpose the four kernels recovered near the corners of the image. Our method produces fewer visual artifacts and reconstructs fine detail better than the other methods. Similar qualitative advantages over additional methods is shown in the project webpage.

Limitations. The photometric and geometric variations that we can handle are similar to those of NRDC [11] and

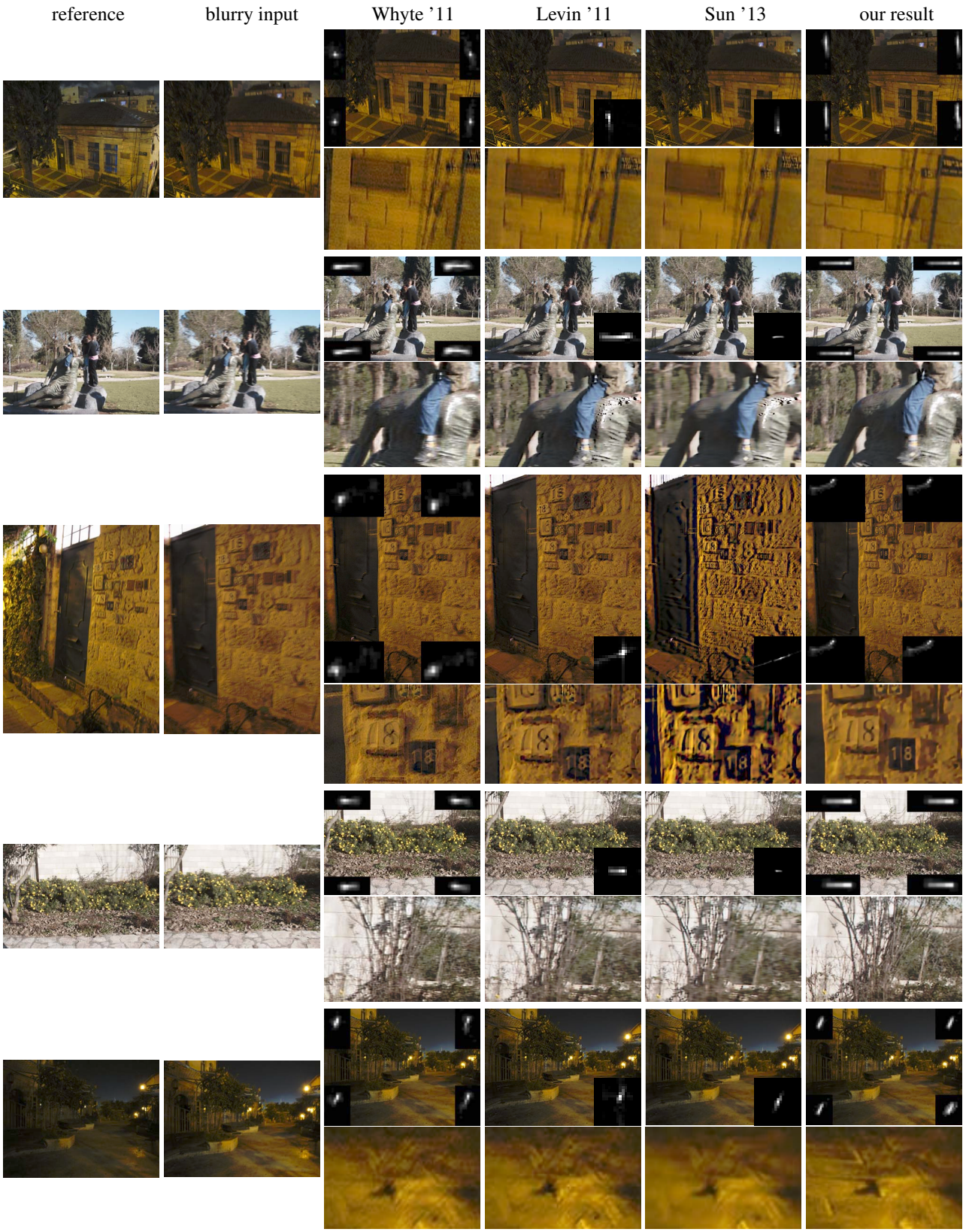

Figure 5. Five real-world sharp/blurry pairs. From left to right: sharp reference; blurry input image; Whyte *et al.* [25]; Levin *et al.* [21]; Sun *et al.* [24]; our result. The odd rows show a deblurred result with the recovered blur kernel(s) and the even rows show an enlarged portion of the deblurred results.

are discussed there in detail. As a result, content with extreme differences (e.g. day vs. night) will not be matched, as the search domain for similar patches is limited to a fixed range of color variations. Another limitation is that a good kernel estimation at one scale is required for correspondence at the next scale. Thus, the shared content must be large enough to be matched at the coarse scales. A possible workaround is to initialize the algorithm with a single image deblurring result at a finer scale. A third limitation is that the local prior can be used only for regions that are available in the reference image. Otherwise, our method will degrade gracefully by using the sparse prior. The sparse prior itself could be replaced with a more advanced image prior (e.g., [28, 24].

5. Conclusions and Future Work

We have presented a new method for deblurring photos using a sharp reference image that may often be found in personal photo collections. We have demonstrated that when a suitable reference exists, our method outperforms the state-of-the-art single image deblurring methods, while no other method can generally exploit such different examples for deblurring. Promising future research directions may be to further extend our method to more general blur models, (e.g. object motion blur), and using similar content from example images as a prior for non-shared regions.

Acknowledgments. This work was supported in part by the Israel Science Foundation and ICRI-CI (Intel Collaborative Research Institute for Computational Intelligence).

References

[1] A. Agrawal, Y. Xu, and R. Raskar. Invertible motion blur in video. *ACM Trans. Graph.*, 28:95:1–95:8, July 2009.

[2] C. Ancuti, C. O. Ancuti, and P. Bekaert. Deblurring by matching. *Comput. Graph. Forum*, 28(2):619–628, 2008.

[3] J. Chen, L. Yuan, C.-K. Tang, and L. Quan. Robust dual motion deblurring. In *Proc. CVPR*, 2008.

[4] S. Cho and S. Lee. Fast motion deblurring. *ACM Trans. Graph.*, pages 145:1–145:8, 2009.

[5] S. Cho, J. Wang, and S. Lee. Video deblurring for hand-held cameras using patch-based synthesis. *ACM Trans. Graph.*, 31(4):64:1–64:9, July 2012.

[6] T. S. Cho, S. Paris, B. K. P. Horn, and W. T. Freeman. Blur kernel estimation using the radon transform. In *CVPR*, pages 241–248. IEEE, 2011.

[7] T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R. Szeliski, and W. T. Freeman. Image restoration by matching gradient distributions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(4):683–694, 2012.

[8] Z. Farbman, R. Fattal, and D. Lischinski. Convolution pyramids. *ACM Trans. Graph.*, 30(6):175:1–8, Dec. 2011.

[9] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph.*, pages 787–794, 2006.

[10] A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless. Single image deblurring using motion density functions. In *Proc. ECCV: Part I*, pages 171–184, 2010.

[11] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski. Non-rigid dense correspondence with applications for image enhancement. *ACM Trans. Graph.*, 30(4):70:1–70:9, 2011.

[12] S. Harmeling, M. Hirsch, and B. Schölkopf. Space-variant single-image blind deconvolution for removing camera shake. In *NIPS*, pages 829–837, 2010.

[13] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Schölkopf. Fast removal of non-uniform camera shake. In *ICCV*, pages 463–470, 2011.

[14] M. Hirsch, S. Sra, B. Scholkopf, and S. Harmeling. Efficient filter flow for space-variant multiframe blind deconvolution. In *Proc. CVPR*, pages 607–614, 2010.

[15] N. Joshi, W. Matusik, E. H. Adelson, and D. J. Kriegman. Personal photo enhancement using example images. *ACM Trans. Graph.*, 29(2):12:1–12:15, April 2010.

[16] N. Joshi, R. Szeliski, and D. J. Kriegman. Psf estimation using sharp edge prediction. In *CVPR*, 2008.

[17] 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 *Proc. ECCV: Part VII*, pages 27–40, 2012.

[18] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, pages 233–240. IEEE, 2011.

[19] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26(3), July 2007.

[20] A. Levin, Y. Weiss, F. Durand, and W. Freeman. Understanding and evaluating blind deconvolution algorithms. In *Proc. CVPR*, pages 1964–1971, Los Alamitos, CA, USA, 2009. IEEE Computer Society.

[21] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *Proc. CVPR*. IEEE, June 2011.

[22] A. Rav-Acha and S. Peleg. Two motion-blurred images are better than one. *Pattern Recogn. Lett.*, 26:311–317, February 2005.

[23] Q. Shan, Z. Li, J. Jia, and C.-K. Tang. Fast image/video upsampling. *ACM Trans. Graph.*, 27(5):1–7, 2008.

[24] L. Sun, S. Cho, J. Wang, and J. Hays. Edge-based blur kernel estimation using patch priors. In *Proc. ICCP*, 2013.

[25] O. Whyte, J. Sivic, and A. Zisserman. Deblurring shaken and partially saturated images. In *Workshop on Color and Photometry in Computer Vision, with ICCV 2011*, 2011.

[26] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. In *CVPR*, 2010.

[27] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. *ACM Trans. Graph.*, 26(3), July 2007.

[28] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *Proc. ICCV, 2011*, pages 479–486, 2011.