

# Good Image Priors for Non-blind Deconvolution: Generic *vs.* Specific

Libin Sun<sup>1</sup>, Sunghyun Cho<sup>2,\*</sup>, Jue Wang<sup>2</sup>, and James Hays<sup>1</sup>

<sup>1</sup> Brown University, Providence, RI 02912, USA

<sup>2</sup> Adobe Research, Seattle, WA 98103, USA

{lbsun,hays}@cs.brown.edu, sodomau@postech.ac.kr, juewang@adobe.com

**Abstract.** Most image restoration techniques build “universal” image priors, trained on a variety of scenes, which can guide the restoration of any image. But what if we have more specific training examples, *e.g.* sharp images of similar scenes? Surprisingly, state-of-the-art image priors don’t seem to benefit from context-specific training examples. Re-training generic image priors using ideal sharp example images provides minimal improvement in non-blind deconvolution. To help understand this phenomenon we explore non-blind deblurring performance over a broad spectrum of training image scenarios. We discover two strategies that become beneficial as example images become more context-appropriate: (1) *locally* adapted priors trained from region level correspondence significantly outperform globally trained priors, and (2) a novel multi-scale *patch-pyramid* formulation is more successful at transferring mid and high frequency details from example scenes. Combining these two key strategies we can qualitatively and quantitatively outperform leading generic non-blind deconvolution methods when context-appropriate example images are available. We also compare to recent work which, like ours, tries to make use of context-specific examples.

**Keywords:** deblur, non-blind deconvolution, gaussian mixtures, image pyramid, image priors, camera shake.

## 1 Introduction

Deblurring is a long-standing challenge in the field of computer vision and computational photography because of its ill-posed nature. In non-blind deconvolution, even though the point spread function (PSF) is known, restoring coherent high frequency image details can still be very difficult. In this paper, we address the problem of non-blind deconvolution with the help of similar (but not identical) example images, and explore deblurring performance across a spectrum of example image scenarios. For each type of training data, we evaluate various strategies for learning image priors from these examples. In contrast to popular methods that apply a single universal image prior to all pixels in the image [12,15,10,22,16], we adapt the prior to local image content and introduce

---

\* Sunghyun Cho is now with Samsung Electronics.

a multi-scale patch modeling strategy to fully take advantage of the example images and show improved recovery of image details. Unlike the recent instance-level deblurring method of [7], we do not require accurate dense correspondence between image pairs and hence generalize better to a wide variety of example image scenarios.

In a typical deblurring framework, a blurry image  $y$  is often modeled as a convolution between a PSF  $k$  and a sharp image  $x$ , with additive noise  $n$ :

$$y = k * x + n. \quad (1)$$

In non-blind deconvolution, both  $y$  and  $k$  are given, and  $n$  is often assumed to be i.i.d Gaussian with known variance. A typical choice of image prior is to encode the heavy-tailed characteristics on image gradients [12,15,10], and regularize the deconvolution process via some form of sparsity constraints on image gradients:

$$x = \operatorname{argmin}_x \|y - k * x\|^2 + \lambda(\|D_x x\|^\alpha + \|D_y x\|^\alpha) \quad (2)$$

where  $\lambda$  is proportional to the noise variance. For Gaussian priors ( $\alpha = 2$ ), there exist fast closed-form solutions via Fourier transform [12,1]. However, Gaussian priors are not appropriate for capturing the heavy-tailedness of natural images, hence produce oversmoothed image gradients. Sparsity priors based on Laplace distribution ( $\alpha = 1$ ) [12] and hyper-Laplacian distributions ( $0.5 \leq \alpha \leq 0.8$ ) [10] have been shown to work well. Other forms of parameterization have also been introduced, such as the generalized Gaussian distribution [2] and mixture of Laplacians [14]. Constraints on image gradients alone are usually insufficient and methods that are able to reason about larger neighborhoods lead to state-of-the-art performance [17,25,23,18,19]. In particular, Zoran and Weiss [25] model image patches via a simple Gaussian mixture model (GMM). This prior turns out to be extremely powerful for removing blur and noise. More recently, discriminative methods trained on corrupted/sharp patch pairs [18,19] have shown impressive performance without specifically modeling the image prior. However, a common problem for these generic methods is that restoring coherent high frequency details remains a challenging task. Deblurred results often contain artifacts such as broken lines and painterly structure details (see Fig. 1).

One likely cause is that given only very local image evidence based on a few adjacent pixels [12,10,2] or image patches [17,21,25,13,23], there is insufficient contextual information to drive the solution away from the conservative smooth state. In addition, most existing methods apply a single image prior to the whole image, which will inevitably introduce a bias towards smooth solutions, since natural images are dominated by smooth gradients.

To combat the tendency to oversmooth details, several recent works consider a content-aware formulation of image priors to accommodate the spatially-varying statistical properties in natural images [2,3,26]. While such content-aware approaches are promising, it is difficult to choose the right prior in the presence of blur and noise. For example, [2,3] estimate content-aware parametric priors based on the downsampled input image. The power of such internal statistics can be rather limited when faced with limited resolution or large blur. However, constructing expressive, content-aware image priors becomes feasible if we have access to sharp example images that are similar to the input.

In the digital age, photographers are likely to take many photos of the same physical scene over time, and this is the type of context we exploit to restore an image and enable content-aware adaptation of image priors. As an experiment, we randomly picked 100 query photos on `Flickr` and found instance level scene matches right next to the query in their respective photostream 42% of the time. This is probably a conservative estimate because photographers are exercising editorial restraint and tend to only publish good and unique photos. For photos where the shutter count was visible, 29% of the time the photographer had taken additional (non-uploaded) photos between instance level matching scenes. It is frustrating for photographers that restoring a blurry photo, even when they can often provide *sharp* photos of the same scene, remains a problem seldom considered by the research community, with the exception of [7], which requires a dense correspondence between the input and the example. However, in the presence of blur and noise, such dense correspondence is unreliable and cannot handle occlusions (see Fig. 1).

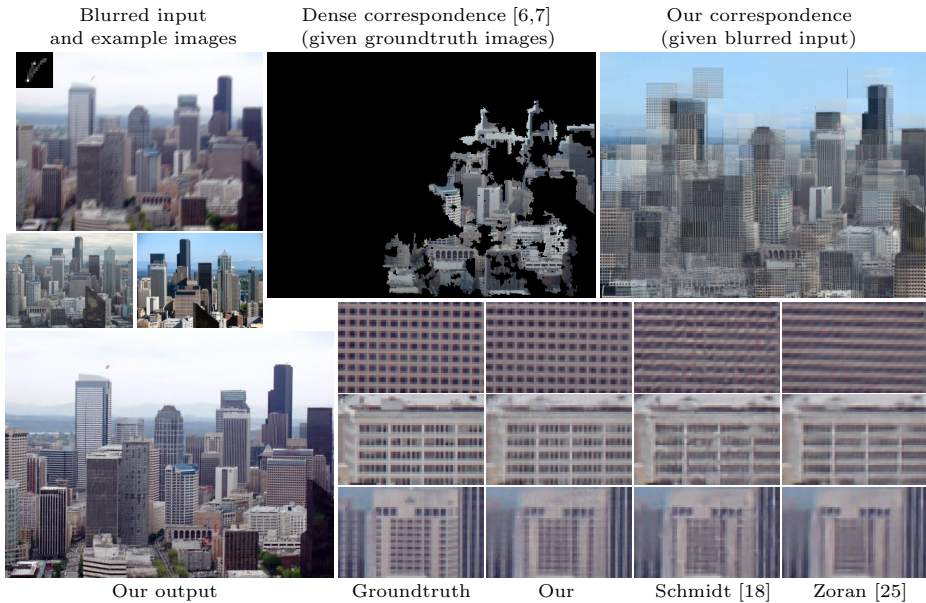
Given the recent advances in blur kernel estimation [5,15,1,22,20,9] and the fact that non-blind deconvolution can be regarded as separate step in the deblurring process, we consider the stand-alone problem of by-example *non-blind* deconvolution: given a blurry input image, a known PSF, and one or more *sharp* images with shared content, how can we reliably remove blur and restore coherent image details?

## 2 Overview

In order to explore non-blind deconvolution performance over a broad range of example image scenarios, we need to define a general deconvolution framework. We extend the EPLL/GMM framework from Zoran and Weiss [25] by augmenting the single-scale patch priors to a multi-scale formulation (Sec. 3). Once the form of image prior and deconvolution method is defined, we consider two training strategies: global training using data from example images, or local training using specific subsets of example data based on a region level correspondence (Sec. 4). Based on this setup, we can investigate various baseline methods that incorporate (1) different parameters in the prior configuration, and (2) different training strategies. We evaluate the performance of these baselines for each example image scenario (Sec. 5) and discover a set of key strategies that show significant benefit from having better example images. Finally, we compare experimental results (Sec. 6) using both synthetically blurred and real photos, against leading methods in generic non-deconvolution as well as by-example deblurring.

## 3 Patch-Pyramid Prior

Our work builds on Zoran and Weiss [25] in which a single-scale patch prior is trained from DC-removed patches. Natural images exhibit diverse yet structured content in different frequency bands that are tightly coupled. A single-scale



**Fig. 1.** The synthetically blurred input and sharp example images show different views of downtown Seattle. Even when given the groundtruth input image, the core correspondence algorithm in [6,7] returns partial (22%) correspondence from example 1 and zero matches from example 2. Our algorithm is able to establish meaningful region level correspondences, and locally adapt the prior to produce significantly more details than state-of-the-art non-blind deconvolution methods.

patch model lacks the ability to learn such statistical dependencies. We propose to jointly model multi-scale concentric patches extracted from an image pyramid, which we call *patch-pyramids*. This naturally extends the spatial scale of the patches without a geometric increase in dimensionality as would happen at a single scale. Furthermore, by capturing how mid and high frequency details covary, image details can be restored more coherently to remove common artifacts such as smudged-out structures, zigzag edges, and painterly appearance.

Consider an image  $x_1$  and its Gaussian pyramid layers  $\{x_1, \dots, x_m\}$ . Given a fixed patch width  $w$ , we denote a *patch-pyramid* by  $[x_1^m]^i$ , meaning a collection of  $m$  patches centered at the same relative coordinates  $i$  in each layer of the Gaussian pyramid. For conciseness, we use  $[x]^i$  to denote patch-pyramid at relative location  $i$  with some fixed size. We use bold fonts to indicate matrices.  $[\mathbf{x}]^i \in \mathcal{R}^{mw^2}$  is formed by concatenating patches in each layer of the pyramid.

We treat patch-pyramids with DC removed per layer as random variables and model the joint occurrence of these  $m$  layers via a Gaussian Mixture Model (GMM). For simplicity, a  $w \times w \times m$  GMM prior means that the model is trained using patch size  $w$  with  $m$  layers.

Let  $\mathbf{x}$  and  $\mathbf{y}$  be the latent and observed image. We follow the EPLL framework of [25] to minimize:

$$f_p(\mathbf{x}|\mathbf{y}) = \frac{\lambda}{2} \|\mathbf{Ax} - \mathbf{y}\|^2 - \sum_i \log p([\mathbf{x}]^i) \tag{3}$$

where  $\mathbf{A}$  represents the blur operator,  $\lambda = \frac{mw^2}{\sigma^2}$ ,  $\sigma^2$  is the noise variance in the image formation process, and  $p([\mathbf{x}]^i) \sim \sum_k \pi_k N([\mathbf{x}]^i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$  is the density function of the GMM prior for patch-pyramids.  $\{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$  are the mixture weight, mean, and covariance of the  $k^{th}$  Gaussian component, respectively. The single-scale patch model in [25] is a special case when  $m = 1$  and  $\boldsymbol{\mu} = \mathbf{0}$ .

### 3.1 Optimization

To optimize Eqn. (3) directly is challenging. A common strategy is to introduce auxiliary variables to assist the optimization process via half quadratic split [10,25]. To achieve this, we introduce auxiliary patch-pyramids  $[\mathbf{z}]^i$  to each location  $i$  and minimize the following global objective:

$$c_{p,\beta}(\mathbf{x}, \{[\mathbf{z}]^i\}|\mathbf{y}) = \frac{\lambda}{2} \|\mathbf{Ax} - \mathbf{y}\|^2 + \sum_i \frac{\beta}{2} (([\mathbf{x}]^i - [\mathbf{z}]^i)^T \boldsymbol{\Sigma}_{noise}^{-1} ([\mathbf{x}]^i - [\mathbf{z}]^i)) - \log p([\mathbf{z}]^i) \tag{4}$$

The diagonal matrix  $\boldsymbol{\Sigma}_{noise}$  reflects the varying relative noise level in each layer, with diagonal entries  $\sigma_j^2, j \in \{1, 2, \dots, m\}$ , each repeating  $w^2$  times. However, the noise across layers is correlated due to the effect of filtering and downsampling in the Gaussian pyramid. We empirically found the relationship  $\sigma_{j+1}^2 = \sigma_j^2/2$  to work well in our experiments. We set  $\sigma_1^2 = 1$ .

The optimization iterates between updating the auxiliary variables  $[\mathbf{z}]^i$  (Sec. 3.2) and solving the latent image  $\mathbf{x}$  (Sec. 3.3). Over iterations,  $\beta$  increases to tighten the coupling of  $[\mathbf{z}]^i$  and  $[\mathbf{x}]^i$  via the second term, which enables convergence. We empirically found the schedule  $\beta = 60 \cdot [1, 2, 4, \dots]$  to work well, typically converging within 8 iterations as shown in Fig. 4.

### 3.2 Z-Step

Given the current estimate for  $\mathbf{x}$ , finding  $[\mathbf{z}]^i$  amounts to solving for the MAP estimate, but computing the exact MAP solution is intractable. We follow the approximation procedure from [25] to obtain a Wiener filtering solution:

$$[\mathbf{z}]^i = (\boldsymbol{\Sigma}_{k_{max}} + \beta \boldsymbol{\Sigma}_{noise})^{-1} (\boldsymbol{\Sigma}_{k_{max}} [\mathbf{x}]^i + \beta \boldsymbol{\Sigma}_{noise} \boldsymbol{\mu}_{k_{max}}) \tag{5}$$

where  $k_{max}$  is the index of the Gaussian component with the highest responsibility.

### 3.3 X-Step

Keeping  $[\mathbf{z}]^i$  fixed, we solve for  $\mathbf{x}$  by the following update:

$$\hat{\mathbf{x}} = \left( \lambda \mathbf{A}^T \mathbf{A} + \sum_i \sum_j \beta_j (\mathbf{P}_{ij} \mathbf{H}_j)^T (\mathbf{P}_{ij} \mathbf{H}_j) \right)^{-1} \left( \lambda \mathbf{A}^T \mathbf{y} + \sum_i \sum_j \beta_j (\mathbf{P}_{ij} \mathbf{H}_j)^T [\mathbf{z}]_j^i \right) \quad (6)$$

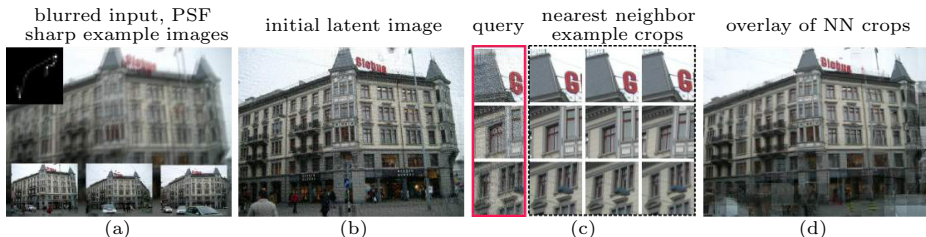
where  $j$  indexes over layers in the pyramid,  $\mathbf{H}_j$  is the Toeplitz matrix representation of the Gaussian filtering and downsampling operators associated with layer  $j$ ,  $\beta_j = \beta/\sigma_j^2$ ,  $[\mathbf{z}]_j^i$  is the  $j^{\text{th}}$  layer patch in  $[\mathbf{z}]^i$ , and  $\mathbf{P}_{ij}$  is the matrix operator extracting the patch at location  $i$  in layer  $j$ .

## 4 Locally Adapted Priors

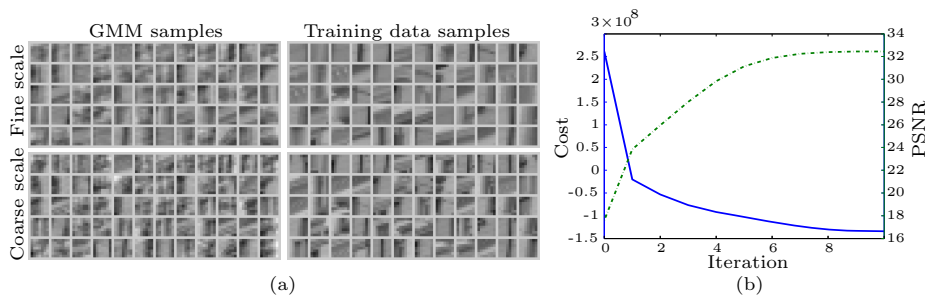
Clearly, the prior in Eqn. (3) plays a central role in the deblurring process. But how much can the prior benefit from example images? One way is to learn the GMM parameters globally using training data collected from the example images. Unfortunately, globally trained priors do not seem to benefit from having *better* example images, as we will show in Sec. 5. This may be because image statistics vary significantly across image locations and using a single global image prior for all image content inevitably compromises image details for smoothness. Instead, we show that priors can be adapted to local image content to provide significantly better recovery of image details.

To construct locally adapted priors, we operate on a half-overlapping grid of image crops and seek local correspondence as shown in Fig. 2. First, a fast  $L_2$ -based deconvolution is performed to provide a rough estimate of the latent image, which is then divided into half-overlapping  $64 \times 64$  crops. For each crop, a HOG descriptor [4] is computed and compared against a database of crops extracted from the sharp example images. We apply scale (factor of 1, 0.9, 0.8) and rotation ( $-3, 0, 3$  degrees) adjustments to each example image to better fit query image content. To reduce noise, we downsample the image by 0.5 in each dimension and apply Gaussian blur before computing the HOG features. A visualization of the nearest neighbor (NN) crops overlay is shown in Fig. 2, where salient image content is matched to reasonable example crops in the presence of noise. Additional visualizations across various example image scenarios can be found in Fig. 4.

Given the above crop-level correspondences, we train independent local GMM priors using patch data collected from 20 nearest neighbors for each query. For each query crop  $q_i$ , we adaptively choose the number of Gaussian components  $K_i \in [K_{min}, K_{max}]$  according to gradient complexity in the training data. Specifically, we first run canny edge detection on the sharp example images, and the total count of edge pixels  $N_i$  in the 20-NN crops for each  $q_i$  is recorded. We linearly scale  $K_i$ 's by  $K_i = K_{min} + (K_{max} - K_{min})(N_i - N_{min}) / (N_{max} - N_{min})$ , where  $N_{min}$  and  $N_{max}$  are the smallest and largest count among all queries. We



**Fig. 2.** (a) Input blurred image with known PSF and sharp example images, (b) initial latent image, (c) best matching example image crops for several query crops from the input, (d) visualization of the nearest neighbor crops overlaid on the input image. The initial latent image is very noisy, the nearest neighbor crops are misaligned and incoherent. Neither alone is a satisfactory image restoration, but we will use the information from both sources to restore blurry photos.



**Fig. 3.** (a) Using patch-pyramids from nearest neighbor crops for the bottom query crop in Fig. 2(c), we train a  $7 \times 7 \times 2$  local GMM and compare its random samples (left) against patches drawn directly from training data (right). The prior captures intricate coupling in different frequency bands. (b) The global objective function in Eqn. (3) converges over iterations with a fixed schedule for  $\beta$ , while the PSNR of the latent image increases. Locally trained  $7 \times 7 \times 2$  priors are used to restore the input image in Fig. 2.

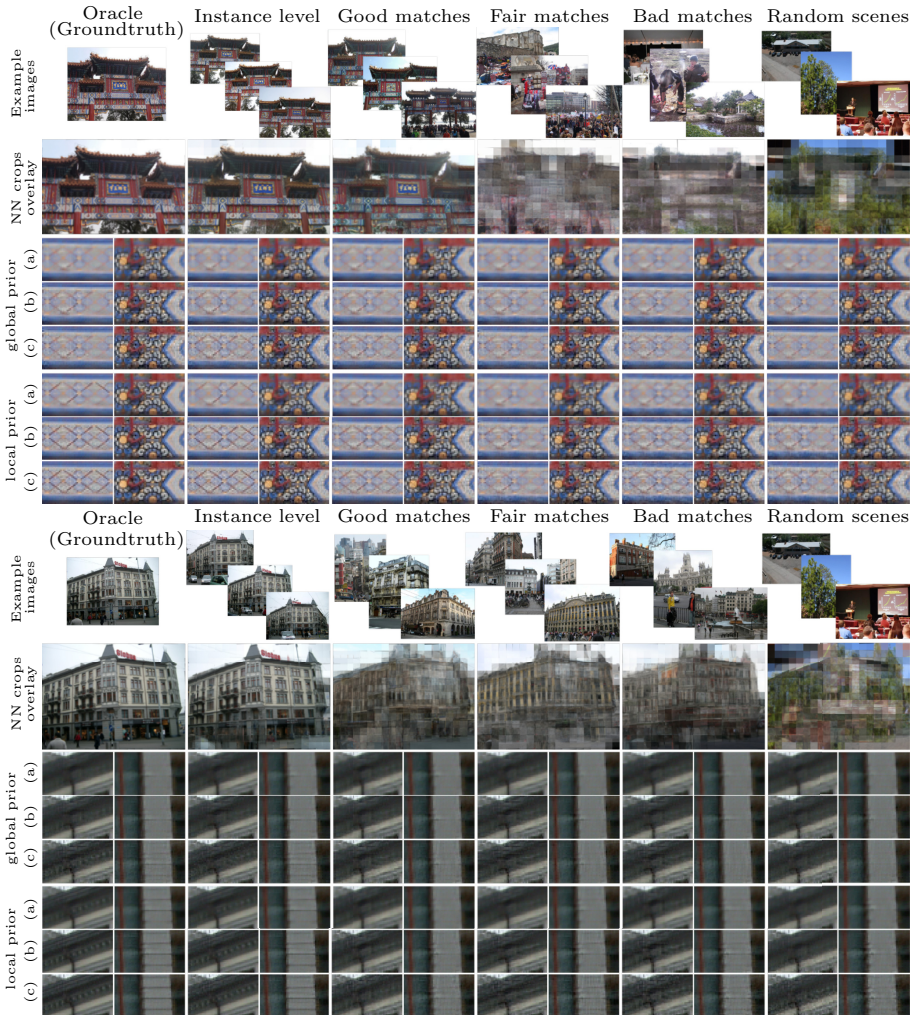
set  $K_{min} = 5, K_{max} = 50$  and learn the GMM via the Expectation-Maximization (EM) algorithm.

Due to the overlapping structure, each pixel is governed by at most four different local GMM priors. To be consistent with the overall objective in Eqn. (3), we choose the solution that gives the highest posterior log likelihood during the MAP approximation of  $[\mathbf{z}]^i$  (see Sec. 3.2).

### 5 How Do Example Images Help?

In order to answer this question, we consider how performance is affected by (1) various example image scenarios and (2) different parameters in our prior. Since the state-of-the-art by-example deblurring method of [7] requires instance-level examples, it is hard to evaluate its performance across a wide spectrum of examples. In this section only, we use the groundtruth image content for



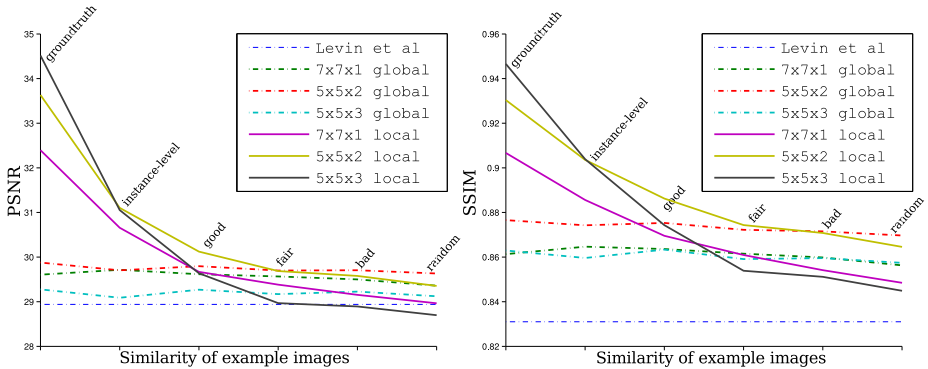


**Fig. 4.** Comparing various baselines across example scenarios and prior configurations. From top to bottom: various scenarios of example images, from the best possible (groundtruth) to similar scenes, to irrelevant images (random scenes); averaged overlay of 20 nearest neighbor crops; output using globally trained priors and locally adapted priors. Results obtained using  $7 \times 7 \times 1$ ,  $5 \times 5 \times 2$  and  $5 \times 5 \times 3$  GMM priors are shown in row (a), (b) and (c) respectively. Better image details can be recovered by (1) using better example images and (2) local training of patch-pyramid priors.

retrieving similar scenes as well as finding crop-level correspondences so that we can more accurately experimentally manipulate the quality of training data.

We consider a number of scenarios of example images: oracle, instance-level, scene matches, and random scenes. The test images are synthetically formed based on the landmark dataset of [24] (see Sec. 6.1 for details). The oracle





**Fig. 5.** Quantitative evaluation of different image priors across example images at various levels of similarity. The six groups of example images are the same visualized in Figure 4. Both PSNR and SSIM scores are reported. Each point is obtained by averaging scores from 20 test images.

scenario assumes that the groundtruth image is available for training the GMM priors. The instance-level examples come directly from the dataset of [24]. Scene matches are computed using the method and database described in [8]. The “good” scene matches (rank 1 to 3) are very similar scenes at similar scale under similar illumination, but typically not instance-level matches. The “fair” scene matches (rank 10-12) are usually less similar but still reasonable. The “bad” scene matches (rank 1998-2000) might only be of the same broad scene category. Finally, we select three random scenes from the database of [8] to act as the worst case scenario. See Figure 4 for examples of each set of training images.

For each example scenario, we consider six alternative prior configurations: (1) the prior can be either globally or locally trained, and (2) the patch-pyramid dimensions can be  $7 \times 7 \times 1$ ,  $5 \times 5 \times 2$ ,  $5 \times 5 \times 3$ . For the globally trained priors, we randomly sample  $2 \times 10^6$  patch-pyramids from the example images (with scale and rotation adjustments) and learn a 50-component GMM via mini-batch EM.

So how do example images help? Our experiments show that the answer is rather subtle: it depends on the priors. In Fig. 4, we show how the deblurring results change as the training examples become less similar to the blurry input. Using a test set of 20 images (see Sec. 6.1), we present quantitative evaluation of these baselines in Fig. 5.

We summarize several key observations below:

- 1 Better example images do help, but it also depends on the priors being used. Locally adapted priors appear to be very sensitive to example images, whereas global priors are not.
- 2 Given instance-level example images, local priors significantly outperform global priors. This is because local priors can provide fine-grained content-aware constraints whereas global priors apply a universal treatment to all image content, often introducing a bias towards smoothness.

- 3 Given sufficiently similar examples (not necessarily instance-level), multi-scale priors outperform single-scale priors. Quantitatively, the  $5 \times 5 \times 2$  prior consistently performs the best (both global and local). In Fig. 4, better connected edges and structured details become much more visible under multi-scale priors.

## 6 Comparison to Leading Methods

With the above analysis and observations, we combine local training and multi-scale patch-pyramid modeling, and report our results using  $7 \times 7 \times 2$  local priors for subsequent comparisons. For comprehensive evaluation, we consider a wide range of test images, containing both synthetic uniform blur and real unknown camera shake. We present quantitative and qualitative comparisons against leading methods in both generic and by-example deblurring methods.

### 6.1 Synthetically Blurred Images

For quantitative evaluation, we generate 20 synthetically blurred test images using four kernels (number 2, 4, 6, 8) from Levin *et al.* [11] and five color images with examples taken from [24]. 1% i.i.d Gaussian noise is added to the luminance channel. Evaluation is based on only the gray scale output images with the outer ring of 30 pixels removed. Color information is only used to assist the correspondence step in [7] and our pipeline (see Sec. 4). In Table 1, we show quantitative comparisons based on PSNR and SSIM scores. For comparisons against non-blind deconvolution methods [12,10,25,18], we assume the groundtruth PSF is known. In this case, our performance is better than the compared methods 100% of the time. A visual comparison of deblurred results can be found in Fig. 6.

When comparing to the recent by-example blind deblurring method of [7], we assume the groundtruth PSF is unknown, and run our system with the estimated blur kernels provided by the authors of [7] to ensure fair comparison. We report PSNR and SSIM performance in Table 1. In this case, we outperform the method of HaCohen *et al.* [7] 85% of the time. A qualitative comparison is shown in Fig. 8. Please note that a single example image is manually selected by the authors of [7] (out of all the examples we supplied) to generate their results since their system does not support multiple example images.

Our method clearly outperforms existing methods in terms of PSNR and SSIM scores, and is capable of restoring coherent mid to high level frequencies such as straight lines and structured details. The recent methods of [25,18] are very competitive without using context-specific example images, but can be quite limited in terms of recovering high frequency details, as shown in Fig. 1 and Fig. 6.

### 6.2 Real Photos with Unknown Blur

In Fig. 7, we show comparison on a test image from [7], where the input image exhibits unknown and spatially varying blur. Our latent image is produced with

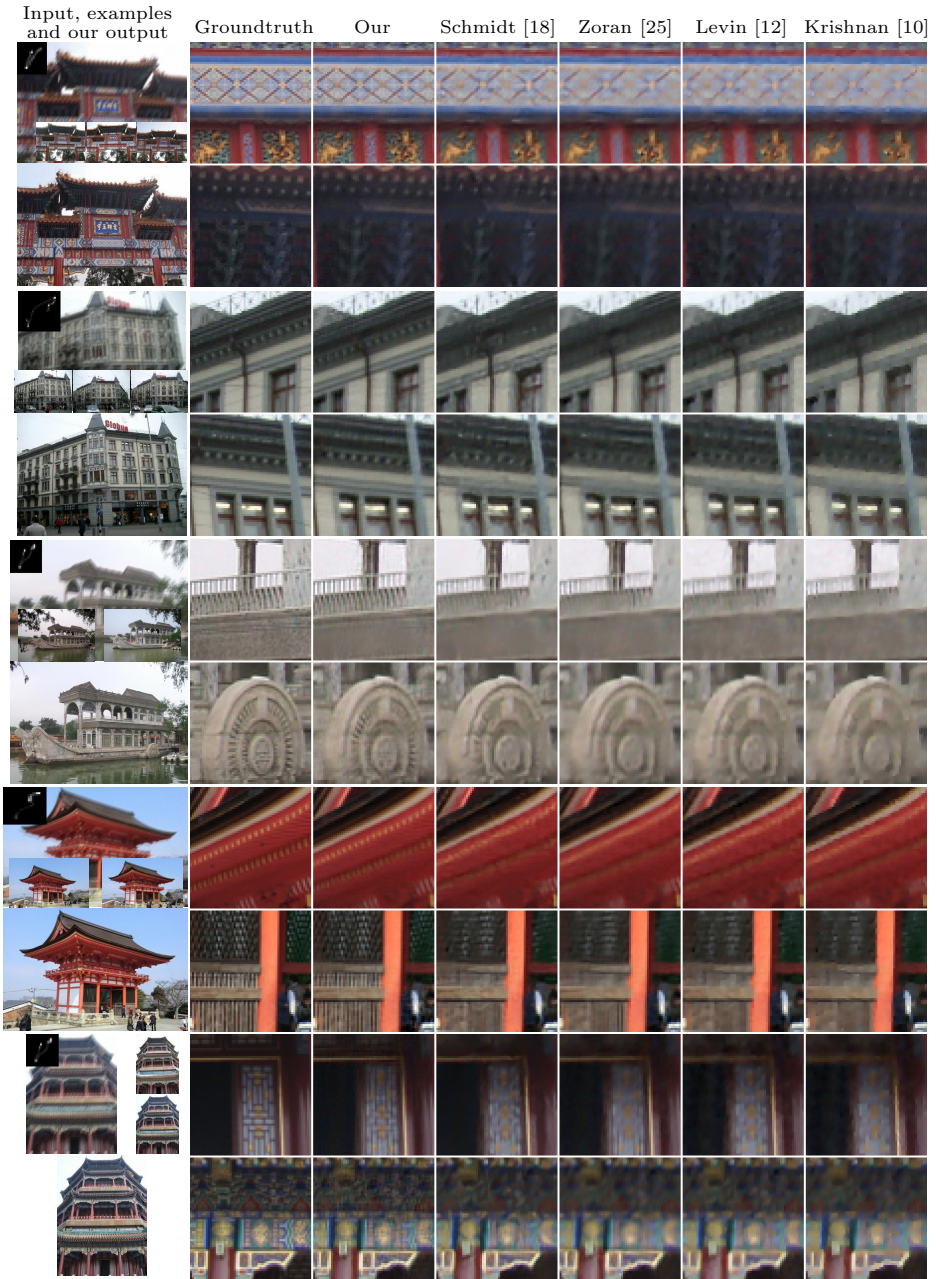
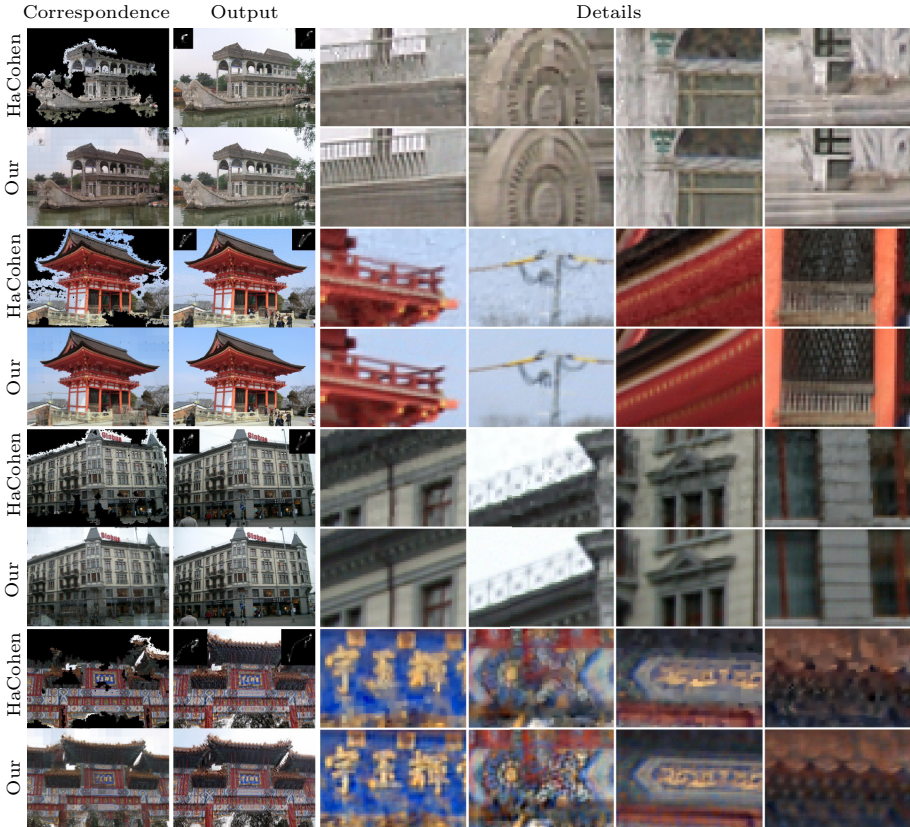


Fig. 6. Comparison on uniformly blurred synthetic test images. Groundtruth PSF's are assumed known and used by all competing methods.



**Fig. 7.** Test image from HaCohen *et al.* [7] with spatially varying PSF estimates. Our approach is highly competitive without requiring dense correspondence.



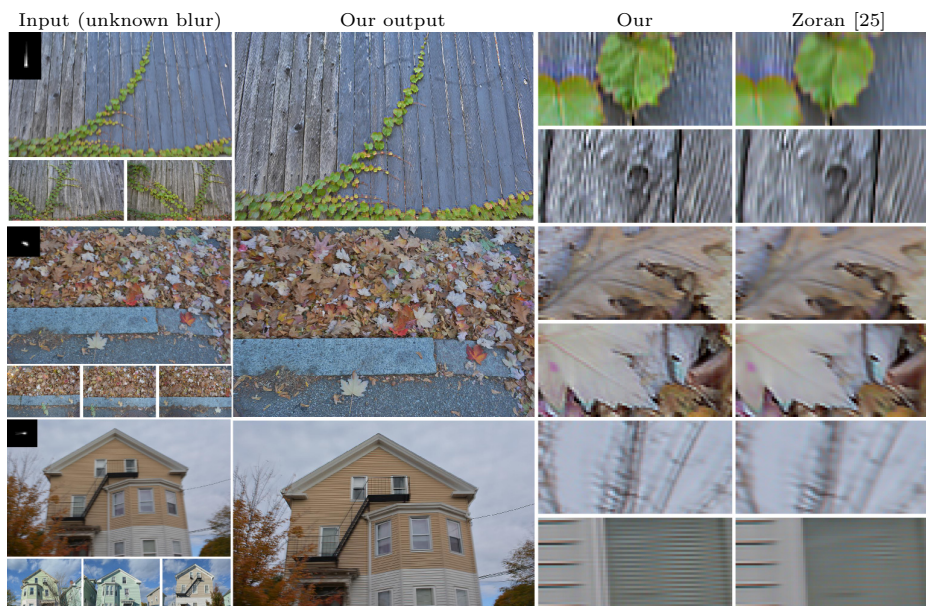
**Fig. 8.** Comparisons against the state-of-the-art by-example method of HaCohen *et al.* [7] on our uniformly blurred synthetic test images. Four examples are shown. Within each example, the first row shows (from left to right): dense correspondence found by [7], output of [7] with estimated PSF (top-left) and groundtruth PSF (top-right), close-up of [7]. The second row shows (from left to right): our nearest neighbor example crop overlay, our output, our close-up. The PSF estimates are supplied by the authors of [7]. All results are generated using the same input blurry images and PSF estimates, hence directly comparable. The last example shows a failure case due to inaccurate PSF estimate.



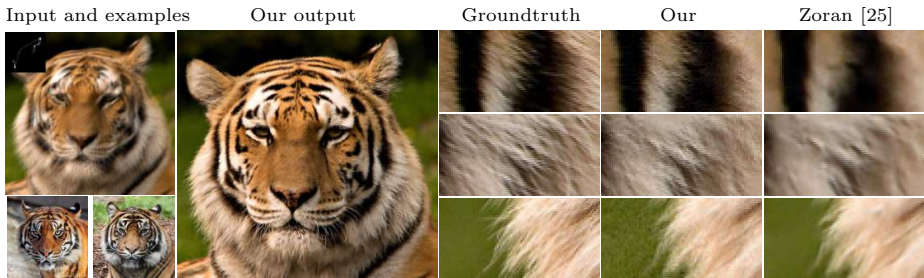
**Table 1.** Quantitative evaluation against existing methods. Methods [12,10,25,18] utilize universally learned image information for deconvolution, while [7] and our method focus on by-example deblurring. For fair comparison, our results in the last column are produced with the estimated PSF from [7]. Both methods make use of example images.

	given groundtruth PSF					given estimated PSF	
method	Levin[12]	Krishnan[10]	Zoran[25]	Schmidt[18]	Our	HaCohen[7]	Our
PSNR	28.94	28.43	29.85	29.90	<b>31.79</b>	27.00	<b>27.60</b>
SSIM	0.852	0.831	0.869	0.879	<b>0.915</b>	0.817	<b>0.843</b>

the PSF estimates from [7], and shows competitive restoration of details. In Fig. 9, we present additional results with unknown blur. All images are taken with the same camera. For most of the test cases, we were unable to obtain successful dense correspondences using the online code provided by the NRDC algorithm [6], which is at the heart of [7].



**Fig. 9.** Except for the third row, the core correspondence algorithm at the heart of [7] yields zero successful matches. For the third test image, it cannot explain more than 70% of the image. All input images are real photos with unknown blur. We estimate the blur kernel using [1].



**Fig. 10.** An example where our method produces convincing textures but also inappropriate high frequency content in background smooth regions (bottom crop)

### 6.3 Limitations

While our system achieves competitive restoration of details, it requires heavy computation especially in the training stage. Using our unoptimized MATLAB implementation, training a 50-component  $5 \times 5 \times 2$  GMM global prior takes roughly 5 hours on an Intel Xeon E5-2650 CPU, whereas training its local prior counterpart requires 12 minutes over a compute grid using 120 cores. However, we find that simply changing the stopping criteria for EM lets us speed up training by a factor of 100 at the expense of a 0.03 drop in PSNR on average. We speculate that further speedup can be obtained by reducing the number of parameters to learn via PCA and by optimizing our code. Finally, incorrect synthesis of details can occur near texture transitions, as shown in Fig. 10.

## 7 Conclusion

In this work, we have provided a novel analysis for by-example non-blind deconvolution by comparing performance against quality of example images for various scenarios using patch-based priors. In particular, we show that locally adapted priors with multi-scale patch-pyramid modeling leads to significant performance gains. We propose a method relying on mid-level correspondence of image crops that does not require dense correspondence at the pixel level. By modeling local image content using multi-scale patch-pyramids, our approach can efficiently take advantage of the sharp example images to restore coherent mid to high frequency image details. We conduct extensive evaluation based on images with both synthetic and real blur, comparing against leading methods in non-blind deconvolution as well as the state-of-the-art by-example deblurring method. By-example deblurring is a promising direction to alleviate the fundamental difficulty of existing algorithms to restore coherent high frequency details, and our method is one step closer to achieving high quality deblur results. For future work, we would like to investigate how our approach can be extended to utilize non-instance level (but still similar) example images and explore ways to improve blind deconvolution via examples.

## References

1. Cho, S., Lee, S.: Fast motion deblurring. *ACM Transactions on Graphics* (2009)
2. Cho, T.S., Joshi, N., Zitnick, C.L., Kang, S.B., Szeliski, R., Freeman, W.T.: A content-aware image prior. In: *CVPR* (2010)
3. Cho, T.S., Zitnick, C.L., Joshi, N., Kang, S.B., Szeliski, R., Freeman, W.T.: Image restoration by matching gradient distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2012)
4. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *CVPR* (2005)
5. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. *ACM Transactions on Graphics* (2006)
6. HaCohen, Y., Shechtman, E., Goldman, D., Lischinski, D.: Non-rigid dense correspondence with applications for image enhancement. *ACM Transactions on Graphics* (2011)
7. HaCohen, Y., Shechtman, E., Lischinski, D.: Deblurring by example using dense correspondence. In: *ICCV* (2013)
8. Hays, J., Efros, A.A.: Im2gps: estimating geographic information from a single image. In: *CVPR* (2008)
9. Köhler, R., Hirsch, M., Mohler, B., Schölkopf, B., Harmeling, S.: Recording and playback of camera shake: benchmarking blind deconvolution with a real-world database. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) *ECCV 2012, Part VII. LNCS, vol. 7578*, pp. 27–40. Springer, Heidelberg (2012)
10. Krishnan, D., Fergus, R.: Fast image deconvolution using hyper-laplacian priors. In: *NIPS* (2009)
11. Levi, E.: Using Natural Image Priors: Maximizing Or Sampling? Hebrew University of Jerusalem (2009), <http://leibniz.cs.huji.ac.il/tr/1207.pdf>
12. Levin, A., Fergus, R., Durand, F., Freeman, W.T.: Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics* (2007)
13. Levin, A., Nadler, B., Durand, F., Freeman, W.T.: Patch complexity, finite pixel correlations and optimal denoising. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) *ECCV 2012, Part V. LNCS, vol. 7576*, pp. 73–86. Springer, Heidelberg (2012)
14. Levin, A., Weiss, Y.: User assisted separation of reflections from a single image using a sparsity prior. *TPAMI* (2007)
15. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding and evaluating blind deconvolution algorithms. In: *CVPR* (2009)
16. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Efficient marginal likelihood optimization in blind deconvolution. In: *CVPR* (2011)
17. Roth, S., Black, M.J.: Fields of experts: A framework for learning image priors. In: *CVPR* (2005)
18. Schmidt, U., Rother, C., Nowozin, S., Jancsary, J., Roth, S.: Discriminative non-blind deblurring. In: *CVPR* (2013)
19. Schuler, C., Burger, H., Harmeling, S., Schölkopf, B.: A machine learning approach for non-blind image deconvolution. In: *CVPR* (2013)
20. Sun, L., Cho, S., Wang, J., Hays, J.: Edge-based blur kernel estimation using patch priors. In: *ICCP* (2013)
21. Weiss, Y., Freeman, W.T.: What makes a good model of natural images? In: *CVPR* (2007)



22. Xu, L., Jia, J.: Two-phase kernel estimation for robust motion deblurring. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 157–170. Springer, Heidelberg (2010)
23. Yu, G., Sapiro, G., Mallat, S.: Solving inverse problems with piecewise linear estimators: From gaussian mixture models to structured sparsity. *Transactions on Image Processing* (2012)
24. Yue, H., Sun, X., Yang, J., Wu, F.: Landmark image super-resolution by retrieving web images. *IEEE Transactions on Image Processing* (2013)
25. Zoran, D., Weiss, Y.: From learning models of natural image patches to whole image restoration. In: ICCV (2011)
26. Zuo, W., Zhang, L., Song, C., Zhang, D.: Texture enhanced image denoising via gradient histogram preservation. In: CVPR (2013)