Deblurring Face Images with Exemplars

Jinshan Pan[1⋆], Zhe Hu[2⋆], Zhixun Su[1], and Ming-Hsuan Yang[2]

[1]Dalian University of Technology [2]University of California at Merced

Abstract. The human face is one of the most interesting subjects involved in numerous applications. Significant progress has been made towards the image deblurring problem, however, existing generic deblurring methods are not able to achieve satisfying results on blurry face images. The success of the state-of-the-art image deblurring methods stems mainly from implicit or explicit restoration of salient edges for kernel estimation. When there is not much texture in the blurry image (e.g., face images), existing methods are less effective as only few edges can be used for kernel estimation. Moreover, recent methods are usually jeopardized by selecting ambiguous edges, which are imaged from the same edge of the object after blur, for kernel estimation due to local edge selection strategies. In this paper, we address these problems of deblurring face images by exploiting facial structures. We propose a maximum a posteriori (MAP) deblurring algorithm based on an exemplar dataset, without using the coarse-to-fine strategy or ad-hoc edge selections. Extensive evaluations against state-of-the-art methods demonstrate the effectiveness of the proposed algorithm for deblurring face images. We also show the extendability of our method to other specific deblurring tasks.

1 Introduction

The goal of image deblurring is to recover the sharp image and the corresponding blur kernel from one blurred input image. The process under a spatially-invariant model is usually formulated as

$$B = I * k + \varepsilon, \tag{1}$$

where I is the latent sharp image, k is the blur kernel, B is the blurred input image, $*$ is the convolution operator, and ε is the noise term. The single image deblurring problem has attracted much attention with significant advances in recent years [5, 15, 20, 3, 22, 12, 13, 6, 24]. As image deblurring is an ill-posed problem, additional information is required to constrain the solutions. One common approach is to utilize prior knowledge from the statistics of natural images, such as heavy-tailed gradient distributions [5, 15, 20, 14], L_1/L_2 prior [13], and sparsity constraints [1]. While these priors perform well for generic cases, they are not designed to capture image properties for specific object classes, e.g., text and face images. The methods that exploit specific object properties are likely

⋆ Both authors contributed equally to this work.

(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 1. A challenging example. (a) Blurred face image. (b)-(d) are the results of Cho and Lee [3], Krishnan et al. [13], and Xu et al. [24]. (e)-(f) are the intermediate results of Krishnan et al. [13] and Xu et al. [24]. (g) Our predicted salient edges visualized by Poisson reconstruction. (h) Our results (with the support size of 75×75 pixels).

to perform well, e.g., text images [2, 19] and low-light images [10]. As the human face is one of the most interesting objects that finds numerous applications, we focus on face image deblurring in this work.

The success of state-of-the-art image deblurring methods hinges on implicit or explicit extraction of salient edges for kernel estimation [3, 22, 12, 24]. Those algorithms employ sharp-edge prediction steps, mainly based on local structure, while not considering the structural information of an object class. This inevitably brings ambiguity to salient-edge selection if only considering local appearance, since multiple blurred edges from the same latent edge could be selected for kernel estimation. Moreover, for blurred images with less texture, the edge prediction step is less likely to provide robust results and usually requires parameter tuning, which would downgrade the performance of these methods. For example, face images have similar components and skin complexion with less texture than natural images, and existing deblurring methods do not perform well on face images. Fig. 1(a) shows a challenging face example which contains scarce texture due to large motion blur. For such images, it is difficult to restore a sufficient number of sharp edges for kernel estimation using the state-of-the-art methods. Fig. 1(b) and (c) show that the state-of-the-art methods based on sparsity prior [13] and explicit edge prediction [3] do not deblur this image well.

In this work, we propose an exemplar-based method for face image deblurring to address the above-mentioned issues. To express the structural information, we collect an exemplar dataset of face images and extract important structures from exemplars. For each test image, we compare it with the exemplars' structure and find the best matched one. The matched structure is used to reconstruct salient edges and guide the kernel estimation process. The proposed method is able to extract good facial structures (Fig. 1(g)) for kernel estimation, and better restore this heavily blurred image (Fig. 1(h)). We will also demonstrate its ability to extend to other objects.

2 Related Work

Image deblurring has been studied extensively and numerous algorithms have been proposed. In this section we discuss the most relevant algorithms and put this work in proper context.

Since blind deblurring is an ill-posed problem, it requires certain assumptions or prior knowledge to constrain the solution space. Early approaches, e.g., [25], usually use the assumptions of simple parametric blur kernels to deblur images, which cannot deal with complex motion blur. As image gradients of natural images can be described well by a heavy-tailed distribution, Fergus et al. [5] use a mixture of Gaussians to learn the prior for deblurring. Similarly, Shan et al. [20] use a parametric model to approximate the heavy-tailed prior of natural images. In [1], Cai et al. assume that the latent images and kernels can be sparsely represented by an over-complete dictionary based on wavelets. On the other hand, it has been shown that the most favorable solution for a MAP deblurring method with sparse prior is usually a blurred image rather than a sharp one [14]. Consequently, an efficient approximation of marginal likelihood deblurring method is proposed in [15]. In addition, different sparsity priors have been introduced for image deblurring. Krishnan et al. [13] present a normalized sparsity prior and Xu et al. [24] use L_0 constraint on image gradients for kernel estimation. Recently, non-parametric patch priors that model appearance of image edges and corners have also been proposed [21] for blur kernel estimation. We note that although the use of sparse priors facilitates kernel estimation, it is likely to fail when the blurred images do not contain rich texture.

In addition to statistical priors, numerous blind deblurring methods explicitly exploit edges for kernel estimation [3, 22, 12, 4]. Joshi et al. [12] and Cho et al. [4] directly use the restored sharp edges from a blurred image for kernel estimation. In [3], Cho and Lee utilize bilateral filter together with shock filter to predict sharp edges. The blur kernel is determined by alternating between restoring sharp edges and estimating the blur kernel in a coarse-to-fine manner. As strong edges extracted from a blurred image are not necessarily useful for kernel estimation, Xu and Jia [22] develop a method to select informative ones for deblurring. Despite demonstrated success, these methods rely largely on heuristic image filtering methods (e.g., shock and bilateral filters) for restoring sharp edges, which are less effective for objects with known geometric structures.

For face image deblurring, there are a few algorithms proposed to boost recognition performance. Nishiyama et al. [17] learn subspaces from blurred face images with known blur kernels for recognition. As the set of blur kernels is pre-defined, the application domain of this approach is limited. Zhang et al. [26] propose a joint image restoration and recognition method based on sparse representation prior. However, this method is most effective for well-cropped face images with limited alignment errors and simple motion blurs.

Recently, HaCohen et al. [8] propose a deblurring method which uses a sharp reference example for guidance. The method requires a reference image with the same content as the input and builds up dense correspondence for reconstruction. It has shown decent results on deblurring specific images, however, the usage of the same-content reference image restrains its applications. Different from this method, we do not require the exemplar to be similar to the input. The blurred face image can be of different identity and background compared to any exemplar images. Moreover, our method only needs the matched contours that encode the

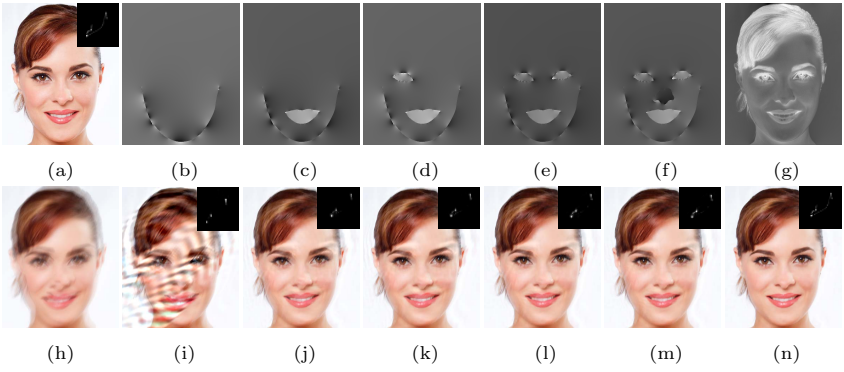


Fig. 2. The influence of salient edges in kernel estimation. (a) True image and kernel. (h) Blurred image. (b)-(f) are extracted salient edges from the clear images visualized by Poisson reconstruction. (g) shows the ground truth edges of (a). (i)-(n) are the results by using edges (b)-(g), respectively.

global structure of the exemplar for kernel estimation, instead of using dense corresponding pixels. In this sense, our method is more general on the object deblurring task with less constraints.

3 Proposed Algorithm

As the kernel estimation problem is non-convex [5, 15], most state-of-the-art deblurring methods use coarse-to-fine approaches to refine the estimated kernels. Furthermore, explicit or implicit edge selections are adopted to constrain and converge to feasible solutions. Notwithstanding demonstrated success in deblurring images, these methods are less effective for face images that contain fewer textures. To address these issues, we propose an exemplar-based algorithm to estimate blur kernels for face images. The proposed method extracts good structural information from exemplars to facilitate estimating accurate kernels.

3.1 Structure of Face Images

We first determine the types and number of salient edges for kernel estimation within the context of face deblurring. For face images, the salient edges that capture the object structure could come from the lower face contour, mouth, eyes, nose, eyebrows and hair. As human eyebrows and hair have small edges which could jeopardize the performance [22, 11], combined with their large variation, we do not take them into consideration as useful structures. Fig. 2 shows several components extracted from a clear face image as approximations of the latent image for kernel estimation (the extraction step will be described later). We test those edges by posing them as the predicted salient edges in the deblurring framework and estimate the blur kernels according to [15] by

$$k^* = \arg\min_k \|\nabla S * k - \nabla B\|_2^2 + \alpha\|k\|^{0.5}, \tag{2}$$

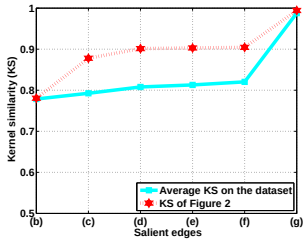
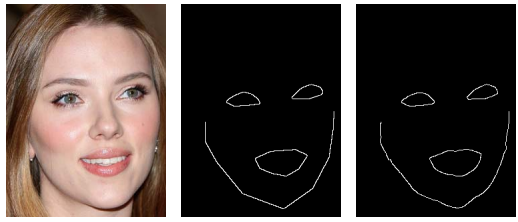


Fig. 3. The relationship between extracted salient edges and kernel estimation accuracy("KS" is the abbreviation of kernel similarity). The notation (b)-(g) for salient edges represent the 6 edge status as Fig. 2(b)-(g).

where ∇S is the gradients of the salient edges extracted from an exemplar image as shown in Fig. 2(b)-(g), ∇B is the gradient computed from the blurred input (Fig. 2(h)), k is the blur kernel, and α is a weight (e.g., 0.005 in this work) for the kernel constraint. The sparse deconvolution method [15] with a hyper-Laplacian prior $L_{0.8}$ is employed to recover the images (Fig. 2(i)-(n)). The results show that the deblurred result using the above-mentioned components (e.g., Fig. 2(l) and (m)), is comparable to that using the ground truth edges (Fig. 2(n)), which is the ideal case for salient edge prediction.

To validate the above-mentioned point, we collect 160 images generated from 20 images (10 images from CMU PIE dataset [7] and 10 images from the Internet) convolving with 8 blur kernels and extract their corresponding edges from different component combination (i.e., Fig. 2(b)-(g)). We conduct the same experiment as Fig. 2, and compute the average accuracy of the estimated kernels in terms of kernel similarity [11]. The curve in Fig. 3 depicts the relationship between the edges of facial components and the accuracy of the estimated kernel. As shown in the figure, the metric tends to converge as all the mentioned components (e.g., Fig. 2(e)) are included, and the set of those edges is sufficient (kernel similarity value of 0.9 in Fig. 3) for accurate kernel estimation.

For real-world applications, the ground-truth edges are not available. Recent methods adopt thresholding or similar techniques to select salient edges for kernel estimation and this inevitably introduces some incorrect edges from a blurred image. Furthermore, the edge selection strategies, either explicitly or implicitly, consider only local edges rather than structural information of a particular object class, e.g., facial components and contour. In contrast, we consider the geometric structures of a face image for kernel estimation. From the experiments with different facial components, we determine that the set of lower face contour, mouth and eyes is sufficient to achieve high-quality kernel estimation and deblurred results. More importantly, these components can also be robustly extracted [28] unlike the other parts (e.g., eyebrows or nose in Fig. 2(a)). Thus, we use these three components as the informative structures for face image deblurring.



(a) Input image (b) Initial contour (c) Refined contour

Fig. 4. Extracted salient edges (See Sec. 3.2 for details)

3.2 Exemplar Structures

We collect 2,435 face images from the CMU PIE dataset [7] as our exemplars. The selected face images are from different identities with variant facial expressions and poses. For each exemplar, we extract the informative structures (i.e., lower face contour, eyes and mouth) as discussed in Sec. 3.1. We manually locate the initial contours of the informative components (Fig. 4(b)), and use the guided filter [9] to refine the contours. The optimal threshold, computed by the Otsu method [18], is applied to each filtered image for the refined contour mask $\mathcal{M}$ of the facial components (Fig. 4(c)). Thus, a set of 2,435 exemplar structures are generated as the potential facial structure for kernel estimation.

Given a blurred image B, we search for its best matched exemplar structure. We use the maximum response of normalized cross-correlation as the measure to find the best candidate based on their gradients

$$v_i = \max_t \left\{ \frac{\sum_x \nabla B(x) \nabla T_i(x+t)}{\|\nabla B(x)\|_2 \|\nabla T_i(x+t)\|_2} \right\}, \tag{3}$$

where i is the index of the exemplar, $T_i(x)$ is the i-th exemplar, and t is the possible shift between image gradients $\nabla B(x)$ and $\nabla T_i(x)$. If $\nabla B(x)$ is similar to $\nabla T_i(x)$, v_i is large; otherwise, v_i is small. To deal with different scales, we resize each exemplar with sampled scaling factors in the range [1/2, 2] before performing (3). Similarly, we handle rotated faces by testing the rotation angle in [-10, 10] degree.

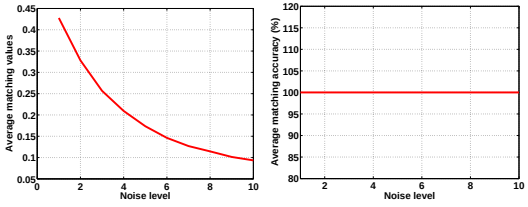
We denote the predicted salient edges used for kernel estimation as ∇S and it is defined as

$$\nabla S = \nabla S_{i^*}, \tag{4}$$

where $i^* = \arg\max_i v_i$, and $\nabla S_{i^*}(x)$ is computed as

$$\nabla S_{i^*}(x) = \begin{cases} \nabla T_{i^*}(x), & \text{if } x \in \{x | \mathcal{M}_{i^*}(x) = 1\}, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Here $\mathcal{M}_{i^*}$ is the contour mask for i^*-th exemplar. In the experiments, we find that using the edges of exemplars $\nabla T_{i^*}(x)$ as the predicted salient edges perform similarly as that of the input image $\nabla B(x)$, which can be found in Sec. 4. The reason is that $\nabla T_{i^*}(x)$ and $\nabla B(x)$ share similar structures due to the matching step, thus the results using either of them as the guidance are similar.



(a) (b)

Fig. 5. The influence of noise on the proposed matching criterion.

We conduct experiments with the quantitative accuracy to verify the effectiveness and robustness of our matching criterion. We collect 100 clear images on 50 identities, with 2 images for each. The images from the same person are different in terms of facial expression and background. In the test phase, we blur one image with random noise as the test image, and pose the others as exemplars. If the matched exemplar is the image from the same person, we mark the matching successful. We perform the test on each images with 8 blur kernels and 11 noise levels (0-10%) and show the matching accuracy in Fig. 5(b). We note that although noise will decrease the average matching values (see Fig. 5(a)), it does not affect the matching accuracy (Fig. 5(b)).

3.3 Kernel Estimation from Exemplar Structure

After obtaining salient edges ∇S, we estimate the blur kernel by alternately solving

$$\min_{I} \|I * k - B\|_2^2 + \lambda \|\nabla I\|_0 \tag{6}$$

and

$$\min_{k} \|\nabla S * k - \nabla B\|_2^2 + \gamma \|k\|_2^2, \tag{7}$$

where λ and γ are parameters for the regularization terms. Here the L_0-norm is employed to restore I and effectively remove some ringing artifacts in I as shown by [23]. In (7), the L_2-norm based regularization is employed to stabilize the blur kernel estimation with a fast solver.

For (6), we employ the half-quadratic splitting L_0 minimization method [23] to solve it. We introduce auxiliary variables $\mathbf{w} = (w_x, w_y)^\top$ corresponding to ∇I and rewrite (6) as

$$\min_{I,\mathbf{w}} \|I * k - B\|_2^2 + \beta \|\mathbf{w} - \nabla I\|_2^2 + \lambda \|\mathbf{w}\|_0, \tag{8}$$

where β is a scalar weight and increases by a factor of 2 over iterations. When β is close to ∞, the solution of (8) approaches that of (6).

We note that (8) can be efficiently solved through alternately minimizing I and $\mathbf{w}$ independently. At each iteration, the solution of I can be obtained by

$$\min_{I} \|I * k - B\|_2^2 + \beta \|\mathbf{w} - \nabla I\|_2^2, \tag{9}$$

Algorithm 1 Solving (6)

Input: Blur image B and estimated kernel k.
$I \leftarrow B$, $\beta \leftarrow 2\lambda$.
repeat
  solve for $\mathbf{w}$ using (11).
  solve for I using (10).
  $\beta \leftarrow 2\beta$.
until $\beta > 1e^5$
Output: Latent image I.

Algorithm 2 Blur kernel estimation algorithm

Input: Blur image B and predicted salient edges ∇S.
for $l = 1 \rightarrow n$ **do**
  solve for k using (7).
  solve for I using Algorithm 1.
  $\nabla S \leftarrow \nabla I$.
end for
Output: Blur kernel k and intermediate latent image I.

which has a closed-form solution computed in the frequency domain by

$$I = \mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}(k)}\mathcal{F}(B) + \beta(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(w_x) + \overline{\mathcal{F}(\partial_x)}\mathcal{F}(w_y))}{\overline{\mathcal{F}(k)}\mathcal{F}(k) + \beta(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(\partial_x) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(\partial_y))}\right). \tag{10}$$

Here $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ denote the Discrete Fourier Transform (DFT) and inverse DFT, respectively, ∂_x and ∂_y denote the vertical and horizontal derivative operators, and the $\bar{\cdot}$ is the complex conjugate operator.

Given I, the solution of $\mathbf{w}$ in (8) can be obtained by

$$\mathbf{w} = \begin{cases} \nabla I, & |\nabla I|^2 \geqslant \frac{\lambda}{\beta}, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

The main steps for solving (6) are shown in Algorithm 1.

Based on the above analysis, the main steps for the proposed kernel estimation algorithm are summarized in Algorithm 2. We use the conjugate gradient method to solve the least square problem (7).

3.4 Recovering Latent Image

Once the blur kernel is determined, the latent image can be estimated by a number of non-blind deconvolution methods. In this paper, we use the method with a hyper-Laplacian prior $L_{0.8}$ [16] to recover the latent image.

3.5 Analysis and Discussion

The initial predicted salient edges ∇S play a critical role in kernel estimation. We use an example to demonstrate the effectiveness of the proposed algorithm

(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 6. Results without and with predicted salient edges ∇S. (a)-(c) denote the 1st, 2nd, and 9th iteration intermediate results, respectively, with edge selection method [3] to predict salient edges ∇S in Algorithm 2. (d) Deblurred result with edge selection method [3] to predict salient edges ∇S in Algorithm 2. (e)-(g) denote the 1st, 2nd, and 9th iteration intermediate results, respectively, using our method to predict salient edges ∇S in Algorithm 2. (h) Our deblurred result. The blurred image in this figure is the same as that of Fig. 1.

for predicting initial salient edges ∇S. Fig. 6 shows that the deblurred result using the edge selection method [3] is unsatisfactory as it introduces artifacts by selecting ambiguous edges. However, the proposed method using the facial structure does not introduce ambiguous edges and thus avoids the misleading kernel estimation. Fig. 6(e)-(g) also demonstrate that the correct predicted salient edges ∇S lead to fast convergence.

We note that the proposed algorithm does not require coarse-to-fine kernel estimation strategies or ad-hoc edge selections. The coarse-to-fine strategy can be viewed as the initialization for the finer levels, which both constrains the solution and reduces the computational load. Recent results of several state-of-the-art methods [3, 13, 24] show that good salient edges at the initial stage are important for kernel estimation. If good initial edges can be obtained, it is not necessary to use coarse-to-fine strategies and specific edge selection, thereby simplifying the kernel estimation process significantly. Our method acts on the original scale only and exploits the exemplar-based structure information to regularize the solution. Benefiting from the facial structure, the proposed method performs well from the beginning without a coarse-to-fine strategy and achieves fast convergence. In the method [3], blur kernels are estimated in a coarse-to-fine manner based on an ad-hoc edge selection strategy. However, it is difficult to select salient edges from severely blurred images without exploiting any structural information (Fig. 6(a)). Comparing to the intermediate results using L_0 prior (Fig. 1(f)), our method maintains the facial components well (Fig. 6(g)), which boosts the performance of kernel estimation and the image restoration.

Robustness of exemplar structures: We use (3) to find the best matched exemplar in gradient space. If the face contour in the latent image is salient, it would present more saliently than other edges after blur. Thus the matched exemplar should share similar parts of the contours with the input, although not perfectly (e.g., Fig. 1(g)). Moreover, the shared contours encode global structures and do not contain many false salient edges caused by blur. We also note that most mismatched contours caused by facial expressions correspond to the small gradients in blurred images. In this situation, these components exert little effect on the kernel estimation according to the edge based methods [3, 22]. To

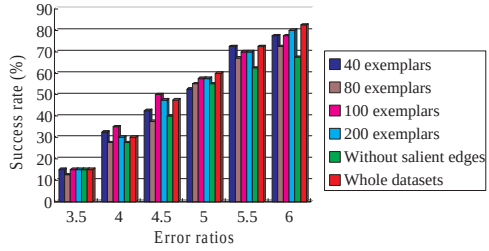


Fig. 7. Robustness to the size of dataset.

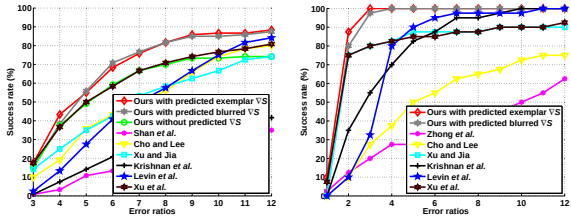
alleviate the problem, we update exemplar edges during the iteration to increase its reliability as shown in Fig. 6(e)-(g). For these reasons, along with the fact that a few correct contours would lead to high-quality kernel estimation, the matched exemplar guides kernel estimation well.

Robustness to dataset: Large dataset will provide reliable results in an exemplar-based method. However, since our method only requires partial matched contours as the initialization, it does not require a huge dataset for good results. To test the sensitivity, we evaluate our method with different numbers of exemplars. We use the k-means method on the exemplar dataset, and choose 40, 80, 100, and 200 clustering centers as the new exemplar datasets, respectively. Similar to [14], we generate 40 blurred images consisting of 5 images (of different identities as the exemplars) with 8 blur kernels for test. The cumulative error ratio [14] is used to evaluate the method. Fig. 7 shows that the proposed method can provide good results with very few exemplars (e.g., 40). With the increasing size of the exemplar dataset, the estimated results do not change significantly, which demonstrates the robustness of our method to the size of dataset.

Robustness to noise: If the blurred image contains severe noise, several edge selection methods [3, 22] and other state-of-the-art methods (e.g., [15, 13, 24]) may not provide reliable edge information for kernel estimation. However, our method will not be affected much due to the robustness of our matching criterion (See analysis in Sec. 3.2). We will show some examples in Sec. 4.

4 Experimental Results

In all the experiments, the parameters λ, γ and n are set to be 0.002, 1 and 50, respectively. We implement Algorithm 2 in MATLAB, and it takes about 27 seconds to process a blurred image of 320×240 pixels on an Intel Xeon CPU with 12 GB RAM. The MATLAB code and dataset are available at `http://eng.ucmerced.edu/people/zhu/eccv14_facedeblur`. As the method of [8] requires a reference image with same content as the blurred image which is not practical, we do not compare to [8] in this section, but we provide some comparisons in the supplementary material.



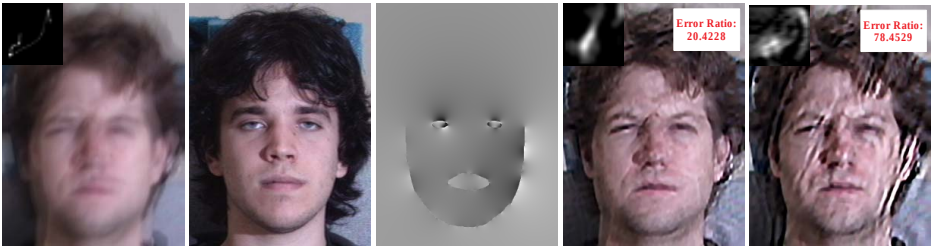
(a) Results on noise-free images (b) Results on noisy images

Fig. 8. Quantitative comparisons with several state-of-the-art single-image blind deblurring methods: Shan et al. [20], Cho and Lee [3], Xu and Jia [22], Krishnan et al. [13], Levin et al. [15], Zhong et al. [27], and Xu et al. [24].

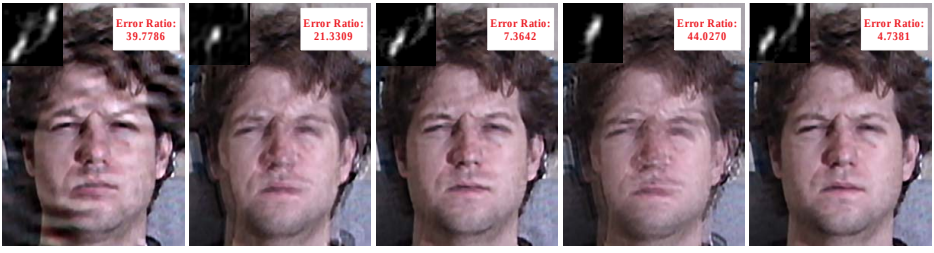
Synthetic Dataset: For quantitative evaluations, we collect a dataset of 60 clear face images and 8 ground truth kernels in a way similar to [14] to generate a test set of 480 blurred inputs. We evaluate the proposed algorithm against state-of-the-art methods based on edge selection [3, 22] and sparsity prior [20, 13, 15, 24] using the error metric proposed by Levin et al. [14]. Fig. 8 shows the cumulative error ratio where higher curves indicate more accurate results. The proposed algorithm generates better results than state-of-the-art methods for face image deblurring. The results show the advantages of using the global structure as the guidance comparing with those using local edge selection methods [3, 22, 24]. We also test different strategies for computing the predicted edges ∇S: 1) using the edges of exemplars $\nabla T_{i^*}(x)$ as ∇S (original); 2) using the edges of the input image $\nabla B(x)$ as ∇S; 3) not using ∇S at all. The first two approaches perform similarly as $\nabla B(x)$ and the matched $\nabla T_{i^*}(x)$ share partial structures, which also demonstrates the effectiveness of our matching step. Compared to the results without predicted edges ∇S, the ones using the predicted edges are significantly improved as shown in Fig. 8(a). It is noted that our method without predicted ∇S does not use coarse-to-fine strategy and generates similar results to [24], which indicates that the coarse-to-fine strategy does not help the kernel estimation much on face images with few textures.

To test the robustness to noise, we add 1% random noise to the test images and present the quantitative comparisons in Fig. 8(b). Compared to other state-of-the-arts methods, our method is robust to noise. We note that the results on noise images are of higher curve than that of noise-free images. The reason is that a noisy input increases the denominator value of the measure [14]. Thus the error ratios from noisy images are usually smaller than those from noise-free images, under the same blur kernel.

We show one example from the test dataset in Fig. 9 for discussion. The sparsity-prior-based methods [20, 13] generate deblurred images with significant artifacts as the generic priors are not effective for kernel estimation when blurred images do not contain rich texture. Edge based methods [3, 22] do not perform well for face deblurring as the assumption that there exists a sufficient number of sharp edges in the latent images does not hold. Compared to the method [24] based on an L_0-regularized method, the results by our method contain fewer



(a) Input & kernel (b) Exemplar image (c) Predicted ∇S (d) Shan [20] (e) Cho and Lee [3]



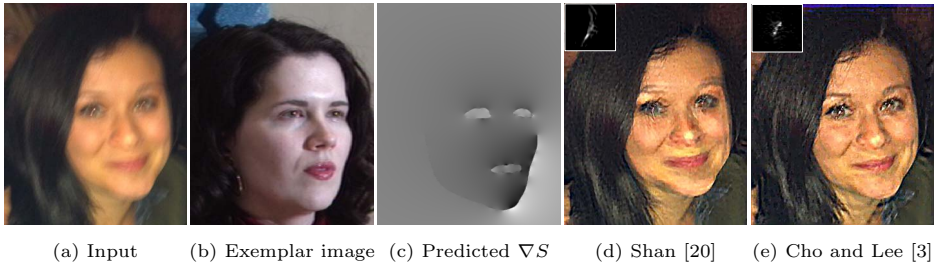
(f) Xu and Jia [22] (g) Krishnan [13] (h) Xu [24] (i) Ours without ∇S (j) Our results

Fig. 9. An example from the synthesized test dataset.

visual artifacts with lower error. Although the best matched exemplar is from a different person (the identities of exemplar and test sets are not overlapped) with different facial expressions, the main structures of Fig. 9(a) and (b) are similar, e.g., the lower face contours and upper eye contours. This also indicates that our approach via (3) is able to find the image with similar structure. The results shown in Fig. 9(i) and (j) demonstrate that the predicted salient edges significantly improve the accuracy of kernel estimation, while the kernel estimation result without predicted salient edges looks like a delta kernel. Although our method is also MAP-based, the predicted salient edges based on the matched exemplar provide good initialization for kernel estimation such that the delta kernel solution (e.g., Fig. 9(i)) is not preferred.

Real Images: We have evaluated the proposed algorithm on real blurred images and show some comparisons with the state-of-the-art deblurring methods. In this example, the input image (Fig. 10(a)) contains some noise and several saturated pixels. The results of [20, 3, 22, 13, 27] are not favorable with obvious noise and ringing artifacts. The proposed method generates a deblurred result with fewer visual artifacts and finer details compared with other methods despite the best matched exemplar visually bearing partial resemblance to the input image.

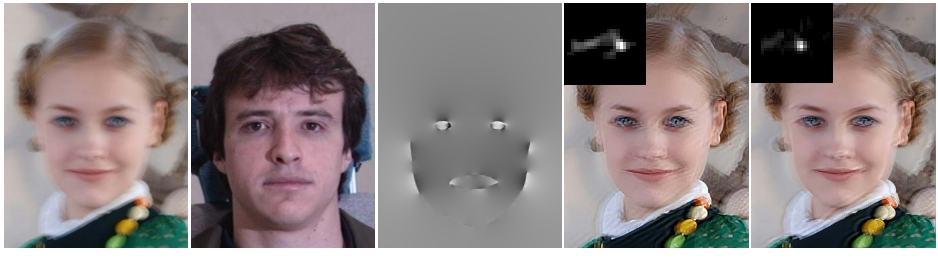
Fig. 11(a) shows another example of a real captured image. The edge selection methods [3, 22] do not perform well as ambiguous edges are selected for kernel estimation. Similarly, the sparsity prior based methods [20, 13, 24] do not perform well with unpleasant artifacts, while our method generates decent results.



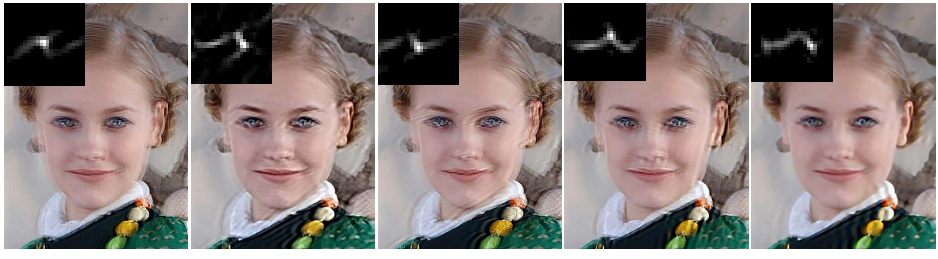
(a) Input (b) Exemplar image (c) Predicted ∇S (d) Shan [20] (e) Cho and Lee [3]

(f) Xu and Jia [22] (g) Krishnan [13] (h) Zhong [27] (i) Xu [24] (j) Our results

Fig. 10. Real captured example with some noise and saturated pixels. The support size is 35×35 pixels.



(a) Input (b) Exemplar image (c) Predicted ∇S (d) Shan [20] (e) Cho and Lee [3]

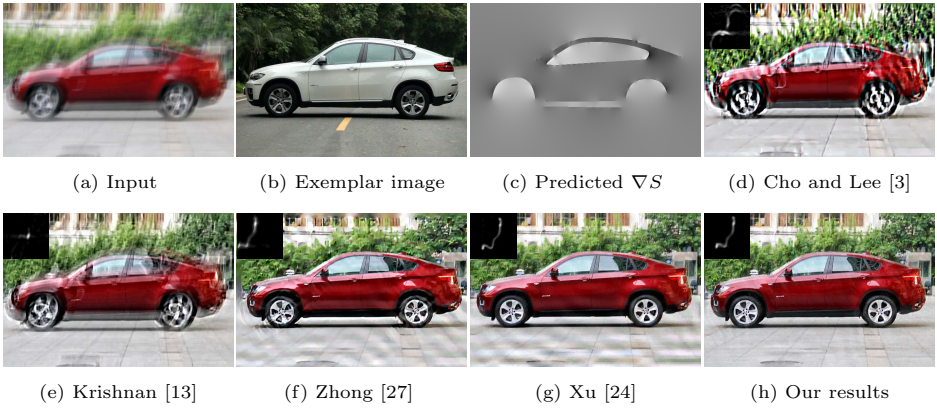


(f) Xu and Jia [22] (g) Krishnan [13] (h) Zhong [27] (i) Xu [24] (j) Our results

Fig. 11. Example of real captured image. The support size is 25×25 pixels.

4.1 Extension of the Proposed Method

In this work, we focus on face image deblurring, as it is of great interest with numerous applications. However, our exemplar-based method can be applied to other deblurring tasks by simply preparing exemplars with the extracted structure. We use an example on car images to demonstrate the extendability



(a) Input (b) Exemplar image (c) Predicted ∇S (d) Cho and Lee [3]

(e) Krishnan [13] (f) Zhong [27] (g) Xu [24] (h) Our results

Fig. 12. Our exemplar-based method on car image. Our method generates the deblurred result with fewer ringing artifacts.

of the proposed method in Fig. 12. Similar to the face images, we first prepare some exemplar images and extract the main structures (e.g., car body, windows and wheels) described in Sec. 3.2. For each test image, we use (3) to find the best exemplar image and compute salient edges according to (4). Finally, Algorithm 2 is used to generate the results. The results of [3, 13, 24, 27] still contain some blur and ringing artifacts. Compared to these methods, our method generates pleasant deblurred results with fewer noise and ringing artifacts.

5 Conclusion

We propose a novel exemplar-based deblurring algorithm for face images that exploits the structural information. The proposed method uses face structure and reliable edges from exemplars for kernel estimation without resorting to complex edge predictions. Our method generates good initialization without coarse-to-fine optimization strategies to enforce convergence, and performs well when the blurred images do not contain rich textures. Extensive evaluations with state-of-the-art deblurring methods show that the proposed algorithm is effective for deblurring face images. We also show the possible extension of our method on the other specific deblurring tasks.

Acknowledgment The work is partially supported by NSF CAREER Grant (No. 1149783), NSF IIS Grant (No. 1152576), NSFC (Nos. 61173103, 61300086, and 91230103), and National Science and Technology Major Project (2013ZX04005021).

References

1. Cai, J.F., Ji, H., Liu, C., Shen, Z.: Framelet based blind motion deblurring from a single image. IEEE Trans. Image Process. 21(2), 562–572 (2012)
2. Cho, H., Wang, J., Lee, S.: Text image deblurring using text-specific properties. In: ECCV. pp. 524–537 (2012)

3. Cho, S., Lee, S.: Fast motion deblurring. ACM Trans. Graph. 28(5), 145 (2009)
4. Cho, T.S., Paris, S., Horn, B.K.P., Freeman, W.T.: Blur kernel estimation using the radon transform. In: CVPR. pp. 241–248 (2011)
5. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. ACM Trans. Graph. 25(3), 787–794 (2006)
6. Goldstein, A., Fattal, R.: Blur-kernel estimation from spectral irregularities. In: ECCV. pp. 622–635 (2012)
7. Gross, R., Matthews, I., Cohn, J.F., Kanade, T., Baker, S.: Multi-pie. In: FG. pp. 1–8 (2008)
8. HaCohen, Y., Shechtman, E., Lischinski, D.: Deblurring by example using dense correspondence. In: ICCV (2013)
9. He, K., Sun, J., Tang, X.: Guided image filtering. In: ECCV. pp. 1–14 (2010)
10. Hu, Z., Cho, S., Wang, J., Yang, M.H.: Deblurring low-light images with light streaks (2014)
11. Hu, Z., Yang, M.H.: Good regions to deblur. In: ECCV. pp. 59–72 (2012)
12. Joshi, N., Szeliski, R., Kriegman, D.J.: PSF estimation using sharp edge prediction. In: CVPR. pp. 1–8 (2008)
13. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: CVPR. pp. 2657–2664 (2011)
14. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding and evaluating blind deconvolution algorithms. In: CVPR. pp. 1964–1971 (2009)
15. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Efficient marginal likelihood optimization in blind deconvolution. In: CVPR. pp. 2657–2664 (2011)
16. Levin, A., Fergus, R., Durand, F., Freeman, W.T.: Image and depth from a conventional camera with a coded aperture. ACM Trans. Graph. 26(3), 70 (2007)
17. Nishiyama, M., Hadid, A., Takeshima, H., Shotton, J., Kozakaya, T., Yamaguchi, O.: Facial deblur inference using subspace analysis for recognition of blurred faces. IEEE Trans. Pattern Anal. Mach. Intell. 33(4), 838–845 (2011)
18. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst., Man, and Cybern. 9(9), 62–66 (1979)
19. Pan, J., Hu, Z., Su, Z., Yang, M.H.: Deblurring text images via L_0-regularized intensity and gradient prior. In: CVPR (2014)
20. Shan, Q., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. ACM Trans. Graph. 27(3), 73 (2008)
21. Sun, L., Cho, S., Wang, J., Hays, J.: Edge-based blur kernel estimation using patch priors. In: ICCP (2013)
22. Xu, L., Jia, J.: Two-phase kernel estimation for robust motion deblurring. In: ECCV. pp. 157–170 (2010)
23. Xu, L., Lu, C., Xu, Y., Jia, J.: Image smoothing via L_0 gradient minimization. ACM Trans. Graph. 30(6), 174 (2011)
24. Xu, L., Zheng, S., Jia, J.: Unnatural L_0 sparse representation for natural image deblurring. In: CVPR. pp. 1107–1114 (2013)
25. Yitzhaky, Y., Mor, I., Lantzman, A., Kopeika, N.S.: Direct method for restoration of motion-blurred images. J. Opt. Soc. Am. A 15(6), 1512–1519 (1998)
26. Zhang, H., Yang, J., Zhang, Y., Huang, T.S.: Close the loop: joint blind image restoration and recognition with sparse representation prior. In: ICCV. pp. 770–777 (2011)
27. Zhong, L., Cho, S., Metaxas, D., Paris, S., Wang, J.: Handling noise in single image deblurring using directional filters. In: CVPR. pp. 612–619 (2013)
28. Zhu, X., Ramanan, D.: Face detection, pose estimation, and landmark localization in the wild. In: CVPR. pp. 2879–2886 (2012)