Simultaneous Blurred Face Restoration and Recognition

Jun Li, Shasha Li, Jiani Hu and Weihong Deng
Beijing University of Posts and Telecommunications
Beijing, China
{buptlijun,lishasha6464}@163.com,{jnhu,whdeng}@bupt.edu.cn

Abstract

Blurred face restoration and recognition are closely related problems. Lots of existing works have proven that these two tasks contribute to each other, but only a few consider combining them together. In this paper, we propose Simultaneous Restoration and Recognition (SRR) method by iteratively dealing with these two tasks. Two new models are presented and then integrated into a two-staged method to handle the problems of the state-of-the-art exemplar based restoration method. The better-performing blurring method is chosen for recognition. Experimental results on FERET database demonstrate the efficiency of SRR in terms of restoration qualities and recognition accuracies for various blurs.

1. Introduction

Blur is one of the most common image degradations. For blurred faces, we usually perform two tasks, the high level blurred face recognition and the low level blurred face restoration.

Existing methods of blurred face recognition fall roughly into three categories. The deblurring method restores the sharp faces from blurred ones, and then feeds the restored image into classifier[6]. The blurring method blurs the gallery with the estimated kernel to construct a specific classifier for a query face[9]. The last method is to find blur invariant descriptors, such as Local Phase Quantization (LPQ) [3]. In practice, combining the first or second method with blur invariant descriptors is widely used [2].

As an ill-posed question, blurred face restoration, or face deblurring, means that we need to reconstruct the blur kernel and sharp image from a blurred one. Most existing deblurring methods are designed for generic images by exploiting natural image priors [10][5], which tend to fail in deblurring blurred faces [8]. Nishiyama *et al*. [6] build multiple PSFs and use the best matched one as the estimated PSF, which, however, should be powerless in handling cases where the blur kernels are complex. Recently, J Pan *et al*. [7] have proposed to use exemplar face to locate a favorable initial blur kernel, which can be considered as the state-of-the-art face deblurring method. However, as our later experiment reveals, it only works for deblurring faces containing lots of background, and fails to process face images produced by a face detector.

Only a few previous works consider the restoration and recognition tasks together, most of which simply fall into the deblurring methods of blurred face recognition. In fact, the restoration result could help blurred face recognition, and vice versa. It would make sense considering the restoration and recognition tasks together in a loop. H Zhang *et al*. [11] propose to combine blurred face restoration and recognition in an uniform sparse representation (SRC) framework, but the SRC method needs to build a complete dictionary with rich set of training faces, thus impractical for undersampled face recognition system [1]. In this paper, we introduce a better combination framework by exploiting intraclass variation and $L_{0.8}$ gradient prior to achieve simultaneous blurred face restoration and recognition (SRR).

The rest of the paper is organized as follows: Section 2 first describes our restoration and recognition methods separately, then presents our SRR method. Section 3 is the experiment part where some experiments on FERET database are presented. We finally conclude our methods and make some further discussions in Section 4.

2. Our method

2.1. Background

The blurring process could be modeled as:

$$\mathbf{y} = \mathbf{x} * \mathbf{k} + \mathbf{n} \tag{1}$$

where $\mathbf{x}$ is the original sharp face, $\mathbf{k}$ the blur kernel, or Point Spread Function (PSF), $\mathbf{n}$ the addictive noise, and $\mathbf{y}$ the final blurred face. This paper assumes that the gallery images are all sharp, which could be easily satisfied. For a blurred query face $\mathbf{y}$, we need to restore the sharp image $\mathbf{x}$, the blur kernel $\mathbf{k}$ and recognize the class label of $\mathbf{y}$ from the gallery set.

2.2. Recognition-based Blurred Face Restoration

For blind image restoration, or blind deblurring, the key is to correctly estimate the blur kernel $\mathbf{k}$. Two elements play an essential role in the state-of-the-art image deblurring methods, that is, proper image priors (either generic or specific) [8] and explicit edge prediction. Obviously, the recognized sharp face could provide strong prior for restoring the blurred query faces.

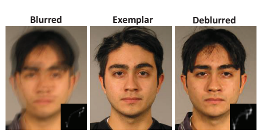
2.2.1 Model 0

The state-of-the-art face deblurring method is based on exemplars [7]. The objectives are rewrote as follows:

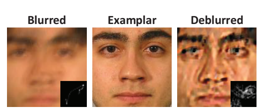
$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{x} * \mathbf{k} - \mathbf{y}\|_2^2 + \lambda \|\nabla \mathbf{x}\|_0 \\ &\min_{\mathbf{k}} \|\nabla \mathbf{e} * \mathbf{k} - \nabla \mathbf{y}\|_2^2 + \gamma \|\mathbf{k}\|_2^2 \end{aligned} \tag{2}$$

where $\mathbf{e}$ is the sharp exemplar $\mathbf{e_0}$ at the first step, and $\mathbf{e} = \mathbf{x}$ at later steps, λ and γ are parameters for the regularization terms. The optimization procedure could be found in [7].

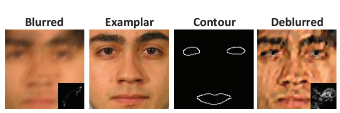
The authors recorded very good performance of the algorithm for a face image. A direct idea is to simply use the recognized sharp face as exemplar, which we call as model 0. Note that the face images in [7] contain lots of background, as presented in Fig. 1(a). A possible question would be: what if we only have a blurred face region outputted by a face detector? We give the answer in Fig. 1(b). We've also manually marked the salient edges of exemplar, but the deblurred result changes little, as shown in Fig. 1(c).



(a) Face image with lots of pure background



(b) Face image outputted by a face detector



(c) Face image outputted by a face detector. Mask is used to select salient edges.

Figure 1. Face deblurring using model 0. Each subfigure contains: blurred faces with ground truth kernel, exemplar sharp faces, (Contour mask for (c)), estimated kernels and deblurred faces.

Pitfalls of Model 0

Two problems are outstanding in model 0. First, $\mathbf{e}$ plays a decisive role, but the difference between $\mathbf{x}$ and $\mathbf{e}$ (recognized face) is large as faces from a same subject may vary with expression or illumination. Second, the L_0 prior on gradient is too strong to suit a simple face region. Therefore, we propose two new models in response.

2.2.2 Model 1

We assume the intraclass variations are similar in different subjects, and we could approximate $(\mathbf{x} - \mathbf{e_0})$ with a linear combination of our pretrained sufficient differences. Note that, W Deng *et al.* [1] employ a similar idea to successfully extend SRC. We change the first equation of the objective by adding model for the intraclass variation:

$$\begin{aligned} &\min_{\mathbf{x},\alpha} \|\mathbf{x} * \mathbf{k} - \mathbf{y}\|_2^2 + \lambda \|\nabla \mathbf{x}\|_0 + \nu \|\mathbf{x} - \mathbf{e_0} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \mu \|\boldsymbol{\alpha}\|_2^2 \\ &\min_{\mathbf{k}} \|\nabla \mathbf{e} * \mathbf{k} - \nabla \mathbf{y}\|_2^2 + \gamma \|\mathbf{k}\|_2^2 \end{aligned} \tag{3}$$

where λ and μ are the regularization parameters, $\mathbf{D} = (\mathbf{d_1}, \ldots, \mathbf{d_n})$ and $\mathbf{d_i} = (\mathbf{x_{i1}} - \mathbf{x_{i0}}, \ldots, \mathbf{x_{im}} - \mathbf{x_{i0}})$. Here $\mathbf{x_{ij}}$ denotes the jth face of the ith subject.

Optimizing for the second equation is same as model 0, and optimizing for the first one is through iteratively optimizing $\mathbf{x}$ and $\boldsymbol{\alpha}$. We describe it as follows:

The solving for x

When α is known, the first equation of Eq. 3 becomes:

$$\min_{\mathbf{x}} \|\mathbf{x} * \mathbf{k} - \mathbf{y}\|_2^2 + \lambda \|\nabla \mathbf{x}\|_0 + \nu \|\mathbf{x} - (\mathbf{e_0} + \mathbf{D}\boldsymbol{\alpha})\|_2^2 \tag{4}$$

To solve this L_0 minimization, we introduce auxiliary variable $\mathbf{w} = (\mathbf{w_x}, \mathbf{w_y})^{\mathbf{T}}$ corresponding to $\nabla \mathbf{x}$. Then Eq. 4 could be approximated by:

$$\min_{\mathbf{x},\mathbf{w}} \|\mathbf{x} * \mathbf{k} - \mathbf{y}\|_2^2 + \beta \|\mathbf{w} - \nabla \mathbf{x}\|_2^2 + \nu \|\mathbf{x} - (\mathbf{e_0} + \mathbf{D}\boldsymbol{\alpha})\|_2^2 + \lambda \|\mathbf{w}\|_0 \tag{5}$$

when β is $+\infty$, 5 equals 4. In our optimization framework, β increases by twofold to approximate the optimization value. We could then solve this problem by iteratively optimizing $\mathbf{x}$ and $\mathbf{w}$. We divide the objective into:

$$\min_{\mathbf{x}} \|\mathbf{x} * \mathbf{k} - \mathbf{y}\|_2^2 + \beta \|\mathbf{w} - \nabla \mathbf{x}\|_2^2 + \nu \|\mathbf{x} - (\mathbf{e_0} + \mathbf{D}\boldsymbol{\alpha})\|_2^2 \tag{6}$$

and

$$\min_{\mathbf{w}} \beta \|\mathbf{w} - \nabla \mathbf{x}\|_2^2 + \lambda \|\mathbf{w}\|_0 \tag{7}$$

Eq. 6 has a closed-form solution for $\mathbf{x}$ using Fourier Transformation, which is

$$\mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}(\mathbf{k})}\mathcal{F}(\mathbf{y}) + \beta(\overline{\mathcal{F}(\partial_\mathbf{x})}\mathcal{F}(\mathbf{w_x}) + \overline{\mathcal{F}(\partial_\mathbf{y})}\mathcal{F}(\mathbf{w_y})) + \nu \mathcal{F}(\mathbf{e_0} + \mathbf{D}\boldsymbol{\alpha})}{\overline{\mathcal{F}(\mathbf{k})}\mathcal{F}(\mathbf{k}) + \beta(\overline{\mathcal{F}(\partial_\mathbf{x})}\mathcal{F}(\partial_\mathbf{x}) + \overline{\mathcal{F}(\partial_\mathbf{y})}\mathcal{F}(\partial_\mathbf{y})) + \nu \mathbf{I}}\right) \tag{8}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the Fourier Transform and Inverse Fourier Transform, ∂_x and ∂_y derivative operators, and $\bar{\cdot}$ the complex conjugate operator.

The solution for $\mathbf{w}$ is:

$$\mathbf{w} = \begin{cases} \nabla \mathbf{x}, & \|\nabla \mathbf{x}\|^2 \geqslant \dfrac{\lambda}{\beta} \\ 0, & otherwise \end{cases} \tag{9}$$

The solving for $\boldsymbol{\alpha}$

The objective for solving $\boldsymbol{\alpha}$ could be simplified as:

$$\min_{\boldsymbol{\alpha}} \|(\mathbf{x}-\mathbf{e_0})-\mathbf{D}\boldsymbol{\alpha}\|_2^2+\rho\|\boldsymbol{\alpha}\|_2^2 \tag{10}$$

where $\rho=\mu/\nu$ serves as a regularization parameter. Eq. 10 has a closed-form solution:

$$\boldsymbol{\alpha}=(\mathbf{D^T D}+\rho\mathbf{I})^{-1}\mathbf{D^T}(\mathbf{x}-\mathbf{e_0}) \tag{11}$$

A fair comparison on model 0 and model 1 is presented in Fig. 2.

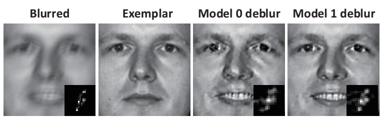


Figure 2. Comparison between model 0 and model 1. There is a intraclass variation between exemplar and ground truth sharp face.

2.2.3 model 2

L_0 is fit for a face image containing pure background because the gradient is very sparse. However, a simple face region needs a weaker sparse prior. J Pan *et al*. [7] use Hyper-Laplacian $L_{0.8}$ norm for the image prior in the last step after estimating the kernel. This prior delivers very good results though time-consuming. Some literature [5][4] has proved that $L_{0.8}$ works well for most natural images, we therefore replace L_0 with $L_{0.8}$. The objective becomes:

$$\begin{aligned}&\min_{\mathbf{x}} \|\mathbf{x}*\mathbf{k}-\mathbf{y}\|_2^2+\lambda\|\triangledown\mathbf{x}\|_{0.8}\\&\min_{\mathbf{k}} \|\triangledown\mathbf{e}*\mathbf{k}-\triangledown\mathbf{y}\|_2^2+\gamma\|\mathbf{k}\|_2^2\end{aligned} \tag{12}$$

The optimization procedure for $\mathbf{x}$ could be found in [4], and the optimization for $\mathbf{k}$ is the same to previous models.

The deblurring results are compared in Fig. 3.

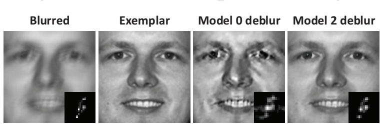


Figure 3. Ideal deblurring results using model 0 and model 1. The exemplars are same as ground truth images to remove the effects of intraclass variations.

2.2.4 Two-staged method

As we could see, the deblurring results after our modifications improve. Model 1 and model 2 each solve one and only one problem remaining in model 0. To handle the two problems, we propose a two-staged method combining the above two models. In the first stage, we use intraclass variation representation to roughly get a kernel; in the second stage, with the kernel as the initial case, we give it some fine tuning using $L_{0.8}$ to get the intended deblurring result. In brief, we locate a favorable initial point with model 1 and adjust it to the best result possible with model 2.

2.3. Restoration-based Blurred Face Recognition

After we have acquired the estimated kernel, we come to face-recognition. This can be accomplished with three tools, i.e. blurring the sharp gallery faces, deblurring the blurred query face, or blur invariant descriptor. The more common practice is to combine either one of the former two tools with the descriptor, as it doesn't conflict with neither of them.

In our case, we choose the state-of-the-art blur invariant descriptor LPQ. When it comes to blurring or deblurring methods, the deblurring methods are straight forward and could be directly acquired using non-blind deblurring method, while deblurring methods always aim to achieve high visual quality, resulting in some artificial effects to the deblurred face. Besides, the blurring methods undergo a sure process of blurring, but the drawback is the huge time costs when blurring the whole gallery set, especially when the gallery set is large.

We finally choose blurring method following LPQ for recognition in our SRR algorithm. Later we could see from the experiment part that the upper bound blurring method outperforms deblurring method.

2.4. Simultaneous Blurred Face Restoration and Recognition

From the discussions above, we could see that blurred face restoration and recognition contribute to each other. We propose to combine them into a joint framework, the procedure is described in Fig. 4.

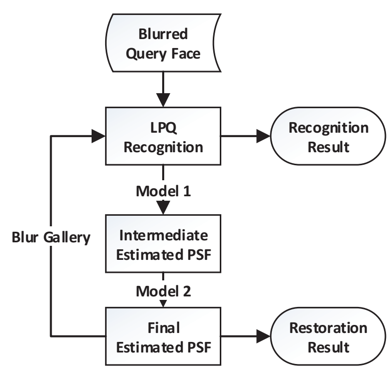


Figure 4. Overview of our SRR method.

First, we directly use LPQ to recognize the query set. Then, for each loop, in the restoration process, we get the intermediate estimated kernel, which is again used for "blurLPQ" recognition. After several iteration loops, we finally output the deblurred face and the recognition result.

3. Experiments

We conduct experiments on FERET database where all faces are sharp. After detecting and cropping face regions, we resize them to $128 * 128$ and perform face alignment. No further processing is added. We use "fa" subset as the gallery, totaling 1196 subjects and 1196 faces. "fb" subset contains 1195 subjects and 1195 faces, which are artificially blurred and added 30 dB white noise to form the query set. "ba", "bj" and "bk" each contains 200 subjects (same for 3 subsets) and 200 faces, and they are used for training intraclass variation dictionary $\mathbf{D}$. Nearest Neighbor (NN) is employed as classifier. The parameters γ, λ, ν and μ are set to be 1, 0.002, 0.08 and 0.0064. We use 50 iterations for model 0 and model 1, and 20 iterations for model 2.

3.1. Blurred Face Restoration

3.1.1 Evaluation Metric

The evaluation metric for blind restoration is always set to be ER (Error Ratio)

$$ER = \frac{\|\mathbf{x_{out}} - \mathbf{x_{gt}}\|}{\|\mathbf{x_{k_{gt}}} - \mathbf{x_{gt}}\|} \tag{13}$$

where $\mathbf{x_{out}}$ denotes the output of the blind deblurring system, $\mathbf{x_{k_{gt}}}$ the deblurring result knowing the ground truth kernel, and $\mathbf{x_{gt}}$ the ground truth sharp image. The value of ER is always bigger than 1 and negatively correlated to the quality of deblurring. This metric enables us to compare different deblurring models in a quantitive way.

3.1.2 Comparison of restoration models

We take a subject where the recognition results for the 8 complex Levin kernels [5] are the same. The deblurred results are shown in Fig. 5.

Though it's hard for us to decide whether model 1 or model 2 is better, a clear conclusion can be drawn that they are both significantly better than model 0.

3.2. Simultaneous blurred face recognition and restoration

Finally, we come to test our SRR method. Two types of blur are taken into consideration: Gaussian blur and motion blur. For Gaussian blur, the size is set to be $2 * ceil(2 * \delta) + 1$, where δ, the standard deviation, is given the value from 2 to 8 in increasement of 2. For motion blur, the blur length L is set from 5 to 20 in increasement of 5, and the blur angle 45^o. As for the choice of tools, we choose the two-staged tool for restoration, and "blurLPQ" for recognition purpose. After LPQ recognition, only two loops are taken (later we could see that this should suffice for our goal). The face recognition accuracies are presented in Fig. 6, and the mean error ratios in Table 1 and Table 2.

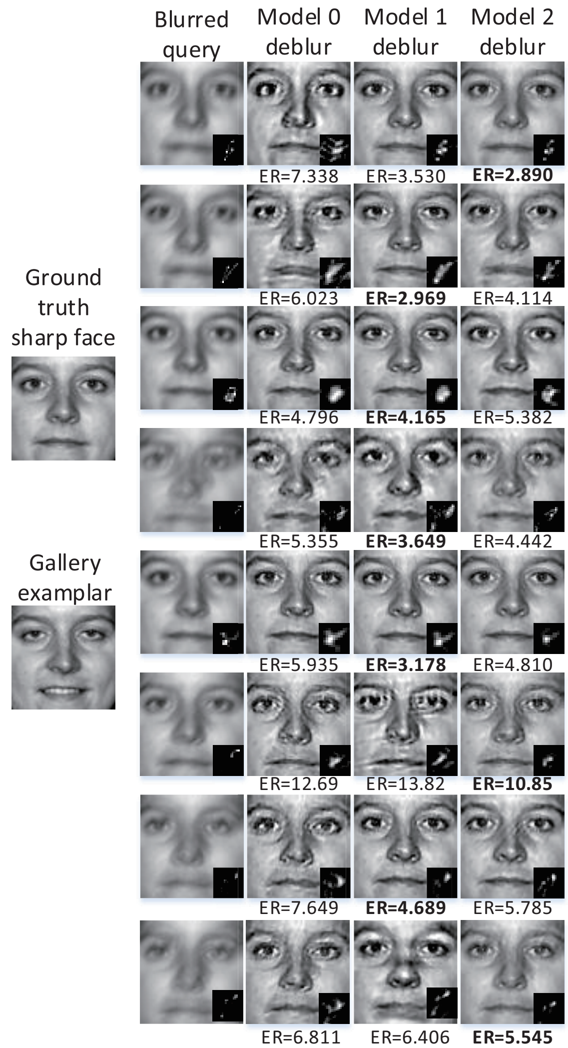
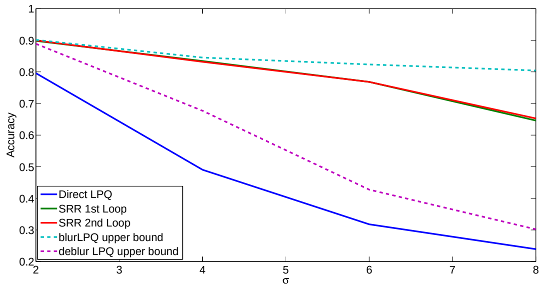


Figure 5. Face restoration results comparison. From left to right: ground truth sharp queries and recognized gallery exemplars, blurred query faces, deblurred results of model 0 to 2.

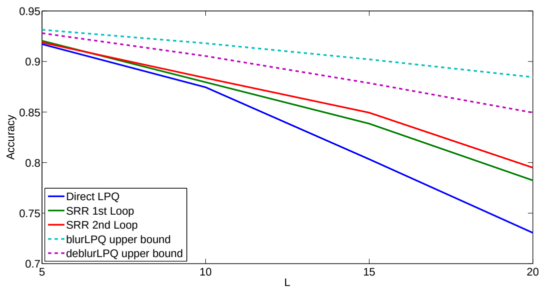
σ	2	4	6	8
model 0	1.602	1.464	1.357	1.269
model 1	1.558	1.513	1.422	1.337
model 2	1.378	**1.220**	**1.236**	**1.224**
SRR 1st loop	1.258	1.242	1.259	1.248
SRR 2nd loop	**1.255**	**1.240**	**1.258**	**1.245**

Table 1. Mean error ratios for Gaussian Blur

Analyzing the experiment results, we arrive at 4 interesting findings. First, after one loop, both the restoration and recognition delivered excellent results. Yet their performances can improve limitedly if new loops are added to. Second, with low ER and high face recognition accuracy (approximates the upper bounds), our method can handle Gaussian blur more effectively than motion blur. Third,



(a) Gaussian Blur



(b) Motion Blur

Figure 6. Face recognition performances

L	5	10	15	20
model 0	2.163	2.430	2.442	2.658
model 1	2.130	2.196	2.209	2.397
model 2	2.384	2.344	2.280	**2.178**
SRR 1st loop	1.816	1.914	2.100	2.225
SRR 2nd loop	**1.824**	**1.914**	**2.097**	**2.218**

Table 2. Mean error ratios for Motion Blur

though in some cases model 2 was recorded a better restoration effect than our two-staged method, the difference is so small that can be overlooked (usually less than 0.02). While in some other cases, our two-staged method is able to outperform model 2 by a large margin. Finally, blurring methods show better upper bound performance than deblurring methods, and our SRR is even better than the upper bound of deblurring methods for Gaussian blurs.

4. Conclusion

In this paper, we propose SRR method by iteratively solving restoration and recognition. We fixed the problems of the state-of-the-art exemplar based restoration method, model 0, by building up two models and proposed to combine these two models into a two-staged method. We employ the better-performing "blurLPQ" as our recognition method. SRR yields significant advantages both on restoration quality and recognition accuracy. For the next step, we plan to integrate model 1 and 2 in an uniform equation instead of by two stages, though the optimization of such an uniform equation may pose a great challenge.

5. Acknowledgements

This work was partially sponsored by supported by the NSFC (National Natural Science Foundation of China) under Grant No. 61375031, No. 61573068, No. 61471048, and No.61273217, the Fundamental Research Funds for the Central Universities under Grant No. 2014ZD03-01, This work was also supported by the Beijing Higher Education Young Elite Teacher Program, and the Program for New Century Excellent Talents in University.

References

[1] W. Deng, J. Hu, and J. Guo. Extended src: Undersampled face recognition via intraclass variant dictionary. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(9):1864–1870, 2012. 1, 2

[2] A. Hadid, M. Nishiyama, and Y. Sato. Recognition of blurred faces via facial deblurring combined with blur-tolerant descriptors. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1160–1163. IEEE, 2010. 1

[3] J. Heikkilä, E. Rahtu, and V. Ojansivu. Local phase quantization for blur insensitive texture description. In *Local Binary Patterns: New Variants and Applications*, pages 49–84. Springer, 2014. 1

[4] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Deconvolution using natural image priors. *Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory*, 2007. 3

[5] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2657–2664. IEEE, 2011. 1, 3, 4

[6] M. Nishiyama, A. Hadid, H. Takeshima, J. Shotton, T. Kozakaya, and O. Yamaguchi. Facial deblur inference using subspace analysis for recognition of blurred faces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(4):838–845, 2011. 1

[7] J. Pan, Z. Hu, Z. Su, and M.-H. Yang. Deblurring face images with exemplars. In *Computer Vision–ECCV 2014*, pages 47–62. Springer, 2014. 1, 2, 3

[8] L. Sun, S. Cho, J. Wang, and J. Hays. Good image priors for non-blind deconvolution. In *Computer Vision–ECCV 2014*, pages 231–246. Springer, 2014. 1, 2

[9] P. Vageeswaran, K. Mitra, and R. Chellappa. Blur and illumination robust face recognition via set-theoretic characterization. *Image Processing, IEEE Transactions on*, 22(4):1362–1372, 2013. 1

[10] R. Wang and D. Tao. Recent progress in image deblurring. *arXiv preprint arXiv:1409.6838*, 2014. 1

[11] H. Zhang, J. Yang, Y. Zhang, N. M. Nasrabadi, and T. S. Huang. Close the loop: Joint blind image restoration and recognition with sparse representation prior. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 770–777. IEEE, 2011. 1