

Efficient Face Image Deblurring via Robust Face Salient Landmark Detection

Yinghao Huang, Hongxun Yao^(✉), Sicheng Zhao, and Yanhao Zhang

School of Computer Science and Technology, Harbin Institute of Technology,
Harbin, China
h.yao@hit.edu.cn

Abstract. Recent years have witnessed great progress in image deblurring. However, as an important application case, the deblurring of face images has not been well studied. Most existing face deblurring methods rely on exemplar set construction and candidate matching, which not only cost much computation time but also are vulnerable to possible complex or exaggerated face variations. To address the aforementioned problems, we propose a novel face deblurring method by integrating classical L_0 deblurring approach with face landmark detection. A carefully tailored landmark detector is used to detect the main face contours. Then the detected contours are used as salient edges to guide the blind image deconvolution. Extensive experimental results demonstrate that the proposed method can better handle various complex face poses while greatly reducing computation time, as compared with state-of-the-art approaches.

Keywords: Efficient · Robust · Face · Image deblurring · Landmark detection

1 Introduction

With the wide popularity of various smart hand-held devices, more and more images and videos are captured and shared by people to record the daily life. Unfortunately, due to various reasons, such as the hand shake of photographer and the motion of target objects, quite a proportion of these images are undesirably blurred, which can degrade ideal sharp images heavily. This predicament catalyzes an important research in the past decade, known as image deblurring, which aims to restore the latent sharp images from the blurred ones.

Mathematically, the underlying process of image degrading due to camera motion can be formulated as:

$$Y = k * X + \varepsilon, \quad (1)$$

where Y refers to the blurred image, k denotes the blur kernel, X and ε correspond to the latent sharp image and the added noise, respectively. Because of the intrinsic illness of this problem, extra information is necessary to constrain

the solutions. One effective way is to figure out the common and dominating differences between the sharp images and the blurred ones. Then by requiring the restored images to have such properties of natural sharp images, the deblurring process is encouraged to move towards the desired direction. Along this way, various priors have been proposed, such as the sparsity constraints and variant normalized version [2, 11], spectral irregularities [7], patch priors [18] and heavy-tailed gradient distributions [6, 13]. In terms of solving strategies, most of the state-of-the-art methods are explicitly or implicitly dependent on the intermediate salient edge structures [4, 5, 10, 17, 19, 20]. The underlying observation is that the original strong edges tend to be salient even after blurring, which can provide useful guiding information for latent sharp image restoration.

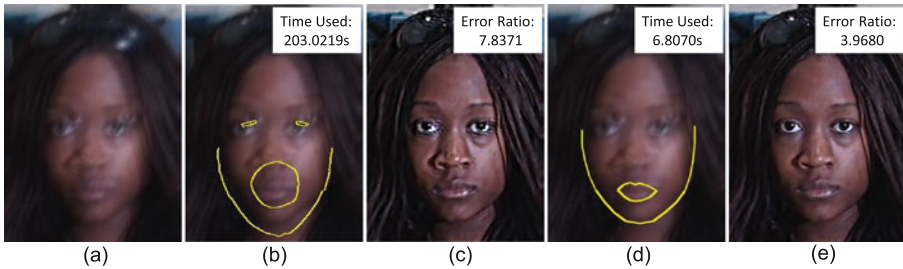


Fig. 1. Illustration of face image deblurring. (a): blurred image; (b) and (d): salient contour achieved by [16] and the proposed method; (c) and (e): corresponding results. Computation cost and error ratios are provided.

Unfortunately, these general-purpose methods fail in some specific cases, such as text images [3, 16], low-light images [9] and face images [1, 8]. In this paper, we mainly focus on face images. Though the method proposed in [8] can achieve decent results, it requires a sharp reference image sharing similar content with the blurred one, which is unrealistic in reality. Pan et al. [1] extended this approach by constructing a large exemplar set and selecting the best fit candidate as the guidance. Experiments showed that the method can achieve decent generalization ability and satisfying results. However, this method is heavily dependent on exemplar images, which makes it time-consuming and hard to handle complex or exaggerated face poses.

Another research area interesting us is face landmark detection. To accurately locate the key points on a face image, a lot of methods have been proposed [25–28]. Among them, the work of [25] achieved satisfying results in real time, making it feasible to fast detect various kinds of salient edges on face images. Our work is mainly motivated by this observation. As illustrated in Fig. 1, our method can achieve more promising visual quality when processing challenging images, while the computation time is much shorter. Especially when the shape of the testing face is not common or the face pose is exaggerated, the visual quality gain is much obvious.

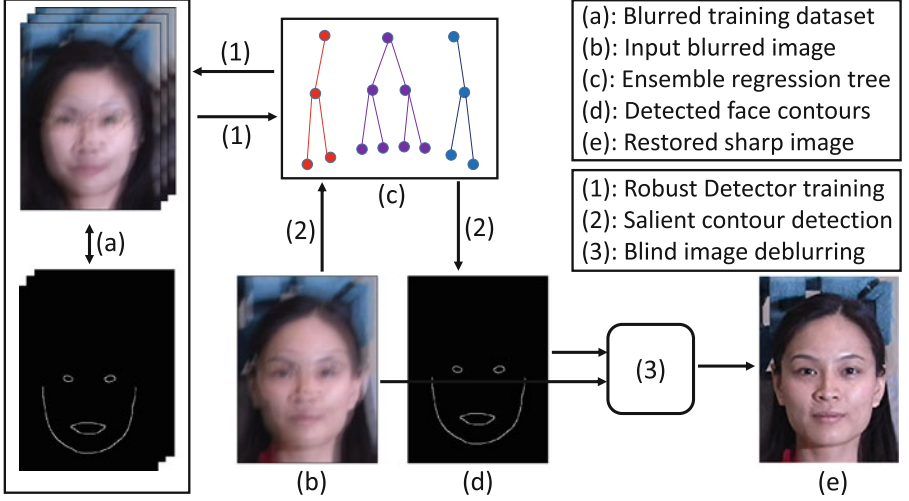


Fig. 2. Framework of the proposed method.

2 The Proposed Method

In this section, we firstly explain the motivation of this work, then give a description of our face deblurring method in detail. As shown in Fig. 2, our method mainly consists of three components: (1) robust face landmark detector training, (2) salient contour detection and (3) blind image deblurring.

2.1 Motivation

As stated previously, most existing face deblurring methods are heavily dependent on exemplar set construction and candidate matching. In consideration of the representation ability, a larger and more complete exemplar set is preferred. Inevitably, much time will be consumed in the candidate matching stage, where the best fit exemplar is chosen by comparing the test image with each candidate image. Taking the method in [1] for example, even in our C++ implementation using a moderate exemplar set with 2245 images, the entire candidate matching process still costs about 198 seconds. Compared with the 5s required by the remaining procedure, this process is too slow to endure. What's worse, due to the great flexibility of face shapes and poses, it is very hard to construct a complete candidate set. When the best matching exemplar image doesn't fit the test image well, the final restored results will be degraded sharply. This predicament is illustrated in Fig. 1 (b), where the mouse contour matches the real one poorly. To address these two problems, we propose to incorporate face landmark detection with image deblurring, which is elaborated in the following sections.

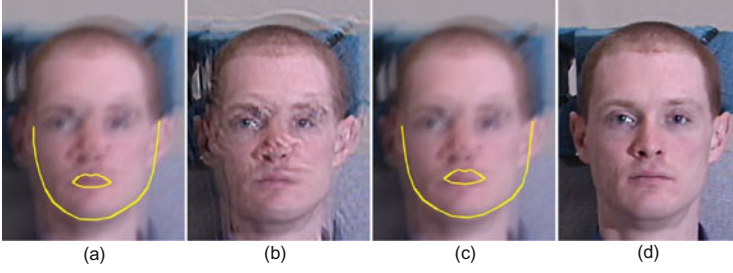


Fig. 3. Illustration of landmark detection. (a) and (c): detected contours using the landmark detector training on clear images and blurred images; (b) and (d): corresponding restored results.

2.2 Robust Face Landmark Detector Training

Accurately locating salient contours of testing face image within short time is a key step. Due to the difficulty of the problem and its intrinsic pairwise comparison of features, example-based methods cannot fulfill such goal, even with a limited training set. To address the above issues, an efficient and effective face landmark detection method similar to the one proposed in [25] is adopted, which is built upon the ensemble of regression forests.

Ensemble Regression Forests Landmark Detection. Given a training dataset $D = \{(I_1, S_1), \dots, (I_n, S_n)\}$ with I_i being a face image and S_i its corresponding shape vector, the goal of training ensemble of regression trees is to achieve a series of regression function $r_t(I, S^{(t)})$, which makes the closest landmark prediction to the ground-truth. In the t -th step, the target function can be formulated as

$$\Delta S_i^{(t)} = S_{\pi_i} - \hat{S}_i^{(t-1)}, \quad (2)$$

where $\hat{S}_i^{(t-1)}$ is the prediction result of the previous step. To this end, each regression function r_t is learned using the classical gradient boosting tree algorithm. The intensity difference of two pixels is used as the decision indicator, which is both light-weighted and relatively insensitive to global lighting changes. A piecewise constant function is used to approximate the underlying function for each regression tree, while leaf nodes are represented by constant vectors. To determine node split, we try to minimize the following loss function:

$$E(Q, \theta) = \sum_{s \in \{l, r\}} \sum_{i \in Q_{\theta, s}} \|r_i - \mu_{\theta, s}\|^2, \quad (3)$$

where Q is the set of indices of the training examples at a node, $Q_{\theta, l}$ are those sent to the left node induced by θ , r_i is the current residual vectors and

$$\mu_{\theta, s} = \frac{1}{|Q_{\theta, s}|}, \text{ for } s \in \{l, r\}, \quad (4)$$

The entire procedure is quite complicated and several other tricks introduced in [25] are also adopted, which speeds up the process greatly.

Training on Blurred Face Images. As demonstrated in Fig. 3, the landmark detector trained on standard dataset with sharp images often fails blurred ones, which can be a disaster to later processing. To endow the landmark detector with the ability of blur tolerating, we synthesised a blurred landmark detection dataset from Helen [29]. Specifically, 2000 images are randomly chosen from all the 2330 images of the Helen dataset, and then blurred with each of the eight kernels in [13] with their ground-truth labels unchanged. This results in a synthesised blurred face landmark detection dataset consisting of 160,000 images, which is then used to train the landmark detector. The left 330 images are similarly blurred to generate the test dataset. According to the same evaluation criteria [25], the average test error is 0.4. Though the test error seems large compared with the result reported in the original paper [25], the detection result is accurate enough to achieve good deblurring results in the next step.

2.3 Salient Contour Detection

In this stage, the main face contours are obtained by applying the trained detector on the test image. The experiments in [1] demonstrated that quite decent results can be achieved by using the contours of the lower face, the eyes and the mouth. However we found that in face landmark detection, as a quite difficult part, eyes usually cannot be accurately located to the extent of providing usable guiding information. So only the contours of the lower face and the mouth are singled out and used in the next stage.

2.4 Blind Image Deblurring

Now that the salient contours ∇S have been acquired, the remaining step is to recover the latent sharp image with the guide of the salient edges. A two-stage procedure similar to [1] is used in this paper:

Kernel Estimation. The two target function used to estimate the blur kernel is as follows:

$$\min_X \|X \star k - Y\|_2^2 + \lambda \|\nabla I\|_0, \quad (5)$$

$$\min_k \|\nabla S \star k - \nabla Y\|_2^2 + \gamma \|k\|_2^2, \quad (6)$$

where λ and γ are coefficients of the regularization terms. As shown in [20], some ringing artifacts in X can be effectively removed by using the L_0 -norm in Eq. 5. While in Eq. 6, the process of blur kernel estimation can be stabilized by applying the L_2 -norm based regularization. The subproblem 5 is solved by employing the half-quadratic splitting L_0 minimization method proposed in [20], while the other subproblem 6 can be solved via the conjugate gradient method.

Latent Image Recovery. After acquiring the blur kernel, quite a number of non-blind deconvolution methods can be used to restore the latent sharp image. Considering the fairness of experimental comparison, we also adopted the method with a hyper-laplacian prior $L_{0.8}$ [15] used in [1].

3 Experimental Results

In this section, extensive experiments are conducted to compare our method with the state-of-the-art in this area. All experiments are conducted on our desktop with an Intel Xeon CPU and 12 GB RAM. To further compare the computation cost with the baseline method [1], we reimplemented our method and the baline method in C++.

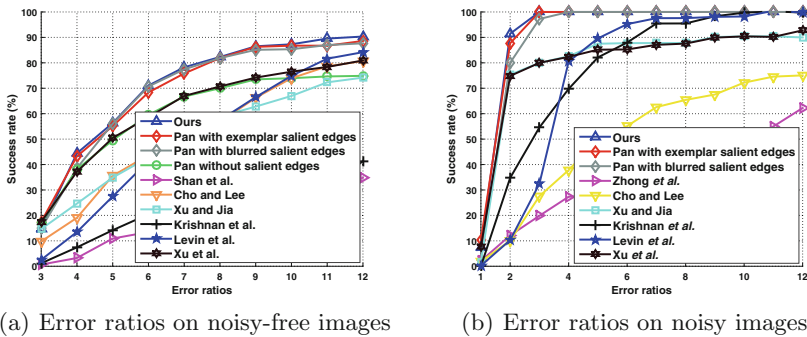


Fig. 4. Quantitative comparisons with some state-of-the-art single-image deblurring methods: Pan et al. [1], Shan et al. [17], Cho and Lee [4], Xu and Jia [19], Krishnan et al. [11], Levin et al. [14], Zhong et al. [23], and Xu et al. [20].

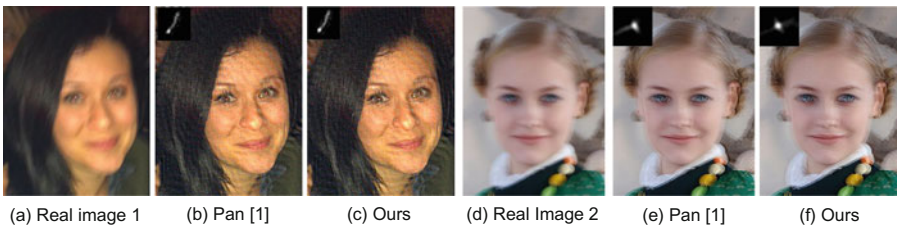


Fig. 5. Deblurring results of the real images used in [1]

3.1 Experiments on Synthesised Dataset and Real Images

To validate the effectiveness of our method, experiments are conducted on the synthesised dataset and real images used in [1] with the same settings. Qualitative and qualitative comparison results are provided in Figs. 4 and 5. Due to

space constraint, only restored images of our method and the baseline method are shown in Fig. 5. As we can see, our overall performance is comparable or even better than those achieved by the best baseline method.

3.2 Computation Cost Comparison

To further compare the computation cost, we divided the entire procedure of [1] and our method into two stages: salient edge acquisition (SEA) and blind deconvolution (BD). When carrying out the previous experiment on synthesised dataset, for each image the time used in the two stages is recorded. Then we computed the average time of the total 480 trails, as shown in Table 1.

Table 1. Running time comparison. The metric unit is second (s).

	SEA	BD	Total
[1]	198.0210	5.0009	203.0219
Ours	1.5070	5.3000	6.8070

It is easy to observe that, our method is much faster than the baseline method. Note that this result is on a quite small size image with $320 * 240$ pixels. When applying on a larger image, the time advantage of our method can be more notable.

3.3 Adaptation to Complex Face Poses

To further compare the ability to process complex face poses, we carefully cropped some face images having exaggerated expression or non-frontal orientation from the Helen dataset. Two experiments are conducted on this new dataset with the same experimental settings, one on the original dataset and another on its noisy version (one percent random noise is added). The result is shown in Fig. 6. Benefiting from the flexibility of the underlying face landmark detection, our method can better handle these difficult situations. Some experimental results are illustrated in Fig. 7. Besides, as shown in Fig. 8, our method is capable of processing images with more than one face, while the baseline method cannot handle this case naturally.

3.4 Rolling Guidance Face Deblurring

If speed is not the main concern, following the loopy style in [30], our method can be further boosted by simply adopting intermediate result image as the guidance and iterating the entire process for several times. Because in this paper algorithm efficiency is also one of the main focuses and the results achieved in a single operation is quite satisfying, this trait of our method is left for future work.

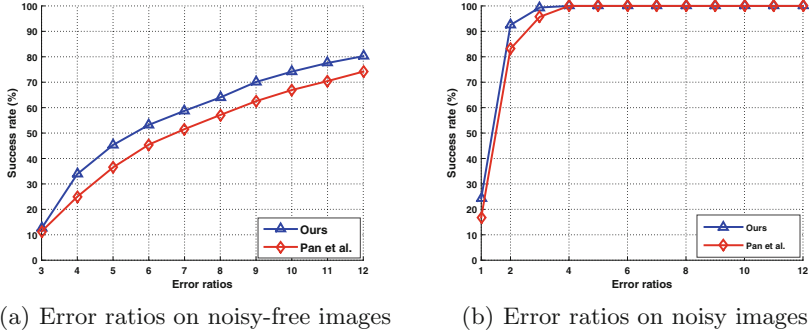


Fig. 6. Quantitative comparison on the new dataset: Pan et al. [1].

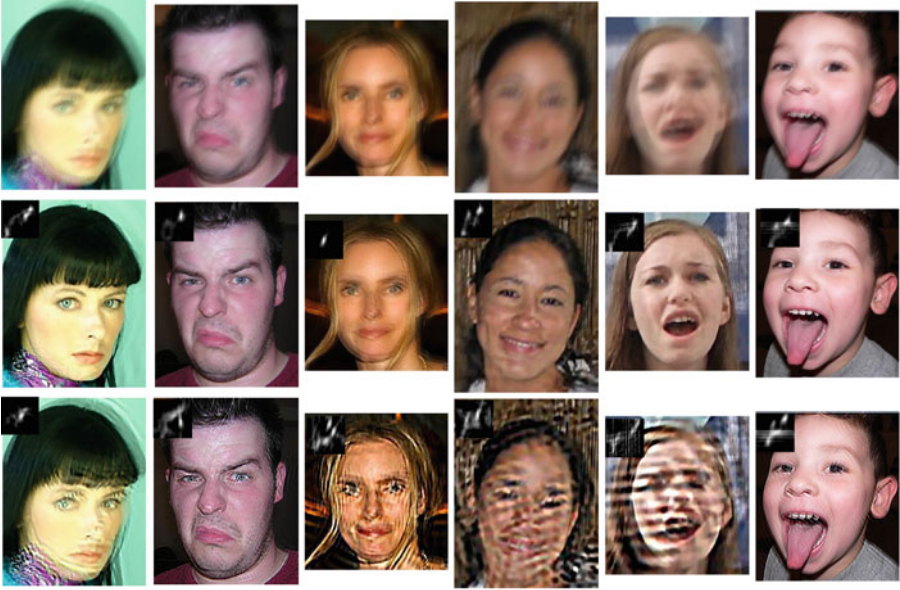


Fig. 7. Deblurring results on complex face poses. From top to bottom: the input blurred images; final restored results by our method; final restored results achieved by [1].

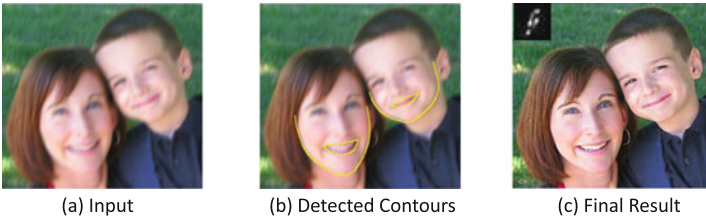


Fig. 8. Our deblurring result on the example image with more than one faces.

4 Conclusions

In this paper, we proposed a novel blind face image deblurring method by incorporating fast landmark detection with classical L_0 method. Benefiting from the accuracy and flexibility of the underlying face landmark detector, the proposed method can better handle challenging face images with various poses and expressions. At the same time, by avoiding time-consuming pair-wise comparison of face features, our method is much more efficient. Extensive experiments wiOur future work will focus on how to improve the deblurring quality further.

Acknowledgement. This work was supported in part by the National Science Foundation of China No. 61472103, and Key Program Grant of National Science Foundation of China No. 61133003.

References

1. Pan, J., Hu, Z., Su, Z., Yang, M.-H.: Deblurring face images with exemplars. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part VII. LNCS, vol. 8695, pp. 47–62. Springer, Heidelberg (2014)
2. 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)
3. Cho, H., Wang, J., Lee, S.: Text image deblurring using text-specific properties. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part V. LNCS, vol. 7576, pp. 524–537. Springer, Heidelberg (2012)
4. Cho, S., Lee, S.: Fast motion deblurring. *ACM Trans. Graph.* **28**(5), 145 (2009)
5. 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)
6. 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)
7. Goldstein, A., Fattal, R.: Blur-Kernel estimation from spectral irregularities. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part V. LNCS, vol. 7576, pp. 622–635. Springer, Heidelberg (2012)
8. HaCohen, Y., Shechtman, E., Lischinski, D.: Deblurring by example using dense correspondence. In: ICCV, pp. 2384–2391 (2013)
9. Hu, Z., Cho, S., Wang, J., Yang, M.H.: Deblurring low-light images with light streaks. In: CVPR, pp. 3382–3389 (2014)
10. Joshi, N., Szeliski, R., Kriegman, D.J.: PSF estimation using sharp edge prediction. In: CVPR, pp. 1–8 (2008)
11. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: CVPR, pp. 2657–2664 (2011)
12. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding and evaluating blind deconvolution algorithms. In: CVPR, pp. 1964–1971 (2009)
13. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Efficient marginal likelihood optimization in blind deconvolution. In: CVPR, pp. 2657–2664 (2011)
14. 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)
15. 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)

16. Pan, J., Hu, Z., Su, Z., Yang, M.H.: Deblurring text images via L0-regularized intensity and gradient prior. In: CVPR (2014)
17. Shan, Q., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. *ACM Trans. Graph.* **27**(3), 73 (2008)
18. Sun, L., Cho, S., Wang, J., Hays, J.: Edge-based blur kernel estimation using patch priors. In: ICCP, pp. 1–8 (2013)
19. 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)
20. Xu, L., Zheng, S., Jia, J.: Unnatural L0 sparse representation for natural image deblurring. In: CVPR, pp. 1107–1114 (2013)
21. 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)
22. 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)
23. 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)
24. He, K., Sun, J., Tang, X.: Guided image filtering. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 1–14. Springer, Heidelberg (2010)
25. Kazemi, V., Sullivan, J.: One millisecond face alignment with an ensemble of regression trees. In: CVPR, pp. 1867–1874 (2014)
26. Cao, X., Wei, Y., Wen, F., Sun, J.: Face alignment by explicit shape regression. *IJCV* **107**(2), 177–190 (2014)
27. Kazemi, V., Sullivan, J.: Face alignment with part-based modeling. In: BMVC, pp. 27–1 (2011)
28. Edwards, G.J., Cootes, T.F., Taylor, C.J.: Advances in active appearance models. In: ICCV, pp. 137–142 (1999)
29. Le, V., Brandt, J., Lin, Z., Bourdev, L., Huang, T.S.: Interactive facial feature localization. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part III. LNCS, vol. 7574, pp. 679–692. Springer, Heidelberg (2012)
30. Zhang, Q., Shen, X., Xu, L., Jia, J.: Interactive facial feature localization. In: ECCV, pp. 815–830 (2014)

Advances in Multimedia Information Processing -- PCM
2015

16th Pacific-Rim Conference on Multimedia, Gwangju,
South Korea, September 16-18, 2015, Proceedings,
Part I

Ho, Y.-S.; Sang, J.; Ro, Y.M.; Kim, J.; Wu, F. (Eds.)
2015, XXIV, 735 p. 358 illus. in color., Softcover
ISBN: 978-3-319-24074-9