Blur Measurement for Partially Blurred Images with Saliency Constrained Global Refinement

Xianyong Fang[1], Qingqing Guo[1], Cheng Ding[1], Linbo Wang[1] *, and Zhigang Deng[2]

[1] Institute of Media Computing, Anhui University, Hefei 230601, China
{fangxianyong, wanglb}@ahu.edu.cn
[2] Department of Computer Science, University of Houston, TX 77204-3010 USA
{zdeng4@uh.edu}

Abstract. Blur measurement of partially blurred image is still far from being resolved. This calls for more distinctive blur features and, even more importantly, a global refinement strategy that has not been considered by existing studies. In this paper we propose a new spatial and frequencial coupled blur descriptor by composing the number of extreme points, the vector of all singular values and the entropy-weighted pooling of the high frequency DCT coefficients. We also introduce a global refinement scheme to explore the merits of saliency for further refining the initial measurements. Consequently, we propose a novel saliency constrained blur measurement method by integrating a neural network based blur metric and a superpixel-scale blur refinement together. Experimental results show the efficiency of our method qualitatively and quantitatively, especially for the images with flat textures.

Keywords: Partial blurred image, blur measurement, saliency

1 Introduction

Partially blurred images widely exist in our daily lives due to the intentionally or unintentionally motion or defocusing capture. Detecting the blur degree for each pixel in these images is important for many tasks, including image segmentation, depth estimation, image deblurring, and refocusing. There still lacks of an efficient method to this inverse problem although there have been many studies on it. In this work, we first introduce a new blur descriptor so that an initial measure for each pixel can be efficiently obtained. It is easy to see that the estimated metric usually appears inconsistent between neighborhoods if only local features are used and the global information is ignored. To address this problem, we further present a global refinement to improve the local metric by exploiting global information, which can effectively remove wrong measurements especially for regions with flat textures. To the best of our knowledge, the global refinement has not been reported in the previous literature.

---

* Corresponding author.

Our method does not assume particular blurring models [1, 2] and thus is general and applicable for real scenes. Our proposed blur descriptor includes both spatial and frequential features so that spatially visual appearance and anti-noisy frequency characteristics can be complementary to each other. It is effective from three aspects in comparison with existing multi-feature oriented methods [3–5].

First, spatially, we introduce the extreme point [6] as a blur feature, which was previously applied to texture analysis due to its embedded structure information. Extreme point is more sparse and thus distinguishable by counting its numbers than the widely used gradient [7, 8] among spatial features [9, 10].

Secondly, we take the complete set of singular values as a spatial complementary to the counting based extreme point because all singular values may vary significantly after blurred. A larger singular value often captures larger scale information, and the blurred image keeps the shape structures at large scale and discards the image details at small scale. The idea is also simple to apply without the burden of manually selecting the high-scale singular values presented in exiting methods [11, 12] which may obtain biased results for different images.

Thirdly, we adopt the state-of-art frequencial feature [13] which is the entropy-weighted pooling of a novel DCT based High-frequency multiscale Fusion and Sort Transform (HiFST) of gradient magnitudes. This metric is robust to noise distortion in comparison with existing frequenicial features [14, 15].

However, wrong measurements seem unavoidable because each blur metric is local and can be easily miscalculated when there are locally clear flat textures. Some existing refinement studies [9, 13, 12, 1] are still local because they only refer to the neighboring pixels. The low-frequency properties are shared by both the blurred region and the flat texture and thus makes the locally spatial or frequential features fail. Therefore, a global refinement strategy is highly desirable so that the overall scene distribution can be exploited to constrain the measurement, considering that flat texture appears very often in real scenes.

We can clearly see the object without difficulty if observing a blurred scene even with many flat textures. This insight inspires us to consider the importance of visual saliency [16] for a seemingly messy scene. Salient object [17] is the only one standing out in the scene considering the fact that humans have difficulty in paying attention to more than one simultaneously. The partially blurred image is also of no exception, where both the blurred and the unblurred regions in a nature scene can be either salient or unsalient. Therefore, we propose a saliency constrained measurement refinement method so that saliency is used as a global representation of an image and thus helps refining the blur measurement globally and eliminating wrong measures.

In the method, we adopt a neural network (NN) based learning approach to obtain the blur measure which is good at fuzzy inference by simulating the work of human brain and thus fits well with the visual perception of blur. We also take a superpixel based strategy to accelerate the saliency constrained blur refinement process. A bilateral filter is used to obtain a structure-preserved pixel-wise consistent blur measurement finally.

2   Our Blur Metric

2.1   Spatial and Frequencial Coupled Blur Descriptor

The proposed descriptor consists of three components: the number of extreme points, the vector of all singular values, and the entropy-weighted pooling of the high frequency DCT coefficients. Therefore, the blur descriptor, $\boldsymbol{B}_{i,j}$, for each pixel at (i,j), $p_{i,j}$, in the blurred $M \times N$ image $\boldsymbol{I}$ can be defined as

$$\boldsymbol{B}_{i,j} = (L_{i,j}, \boldsymbol{V}_{i,j}, D_{i,j}) \tag{1}$$

with $L_{i,j}$, $\boldsymbol{V}_{i,j}$ and $D_{i,j}$ denoting the number of extreme points, the vector composed of all singular values, and the entropy-weighted pooling of the high frequency DCT coefficients.

Extreme points. Local structures represent the local directions and intrinsic signal dimensions [6] and image extrema can be divided into five structure classes by intensity distribution: '|', '–', '/', '\' and '*'. The first four classes are one-dimensional except the last one being two-dimensional. The local minimum is weak for a blurred image and therefore the local maximum is considered in our work as extreme point. The extreme points show the high frequency information spatially and thus can depict the blur degree of each pixel.

Figure 1b shows the total number of each ideal extreme class in the four blurred or clear 40×40 regions specified in Figure 1a. The first four structure classes are few for all the regions while the blurred regions (Regions #1 and #2) are null and thus more sparse. However, for the fifth class, the clear regions (Regions #3 and #4) contain apparently more extreme points than the blurred regions (Regions #1 and #2). The map of the number of extreme points of the fifth structure class (Figure 1c) also shows this observation. Therefore, the fifth class can be more efficient and is adopted by us as a part of the blur descriptor.

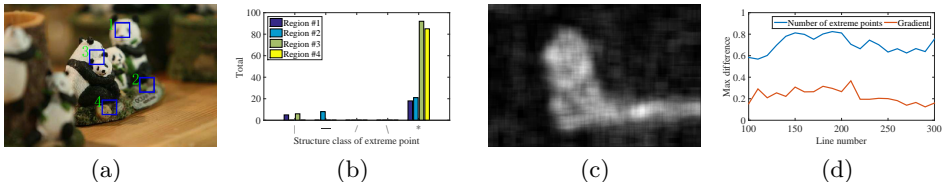(a)   (b)   (c)   (d)

Fig. 1: An example of the extreme point feature. (a): The example image with four specified clear or blurred regions; (b) the numbers of extreme points for the four regions in (a); (c): the map of the number of the extreme points belonging to the fifth structure class; and (d): the blur measurement abilities of extreme points and gradient.

Let us denote the intensity of $p_{i,j}$ being $E_{i,j}$. The noise-free definition of extreme point belonging to the fifth class, '*', for the center of a 3×3 area is

$$\bigcap_{\substack{i=-1 \\ |i|+|j| \neq 0}}^{+1} \bigcap_{j=-1}^{+1} E_{i,j} > E_{i+k,j+q} \tag{2}$$

where k and q are the numbers constraining the position of $E_{i+k,j+q}$ inside the area, and $|k| + |q| \neq 0$ so that the maximum condition is kept.

Figure 1d shows the performance of the extreme points in comparison with the widely adopted gradient. The number of extreme points belonging to the fifth structure class for a 8×8 region surrounding each pixel is counted. The max differences of either gradient or number of extreme points for each line between the top and the bottom lines of the clear object, No. 100 and 300, are computed, considering that the maximum and minimum of either gradient or number of extreme points in each line represent the clear pixel and the blurred one respectively. It can be seen that the extreme points are more discriminative than gradient and can be an effective feature for blur measurement.

Singular values. An image $\boldsymbol{I}$ can be decomposed into the weighted sum of n eigen-images by Singular Value Decomposition (SVD) as follows,

$$\boldsymbol{I} = \sum_{i=1}^{n} \lambda_i \boldsymbol{E}_i \tag{3}$$

where $\lambda_i (1 \leq i \leq n)$ are the eigen values in a decreasing order and $\boldsymbol{E}_i (1 \leq i \leq n)$ are rank-1 matrices called eigen-images. The eigen-images capture different detailed information: The more significant eigen-images, the larger scale information captured. A blurred image keeps the shape structure at large scale while discarding the image detail at small scales. Therefore, singular values can depict the subtle differences between blurred and non-blurred pixels.

We take the vector of all the singular values. Figure 2a shows this principle. Each region can have relatively high λ_1 for the largest scale, but its remaining eigen values for smaller scales vary differently: they are more significant in the clear region (Regions #4-#6) than the blurred one (Regions #1-#3) (most of them are nearly null). In addition, λ_1 for the blurred regions varies significantly with some regions (e.g., Regions #5 or #6) even close to the clear ones. Consequently, manually selecting the most significant singular values [11, 12] can be invalid. Therefore, the complete set of singular values is adopted by us as an efficient part of the blur descriptor, i.e., $\boldsymbol{V}_{i,j} = (\lambda_1, \ldots, \lambda_n)$.

DCT coefficients. DCT aims to transform a signal from spatial to frequencial space with cosine functions. High frequencies are sparse with mostly null in comparison with the other ones for a blurred image. Therefore the high frequency DCT coefficients can be used to measure the blur degree. We adopt the efficient
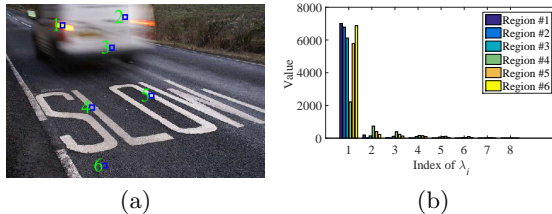
(a)   (b)

Fig. 2: Demonstration of the singular value feature. (a): The example image with six specified rectangular regions (three blurred regions and three clear ones); (b) comparison of the singular values among the six specified regions in (a). *Note*: The singular values are ordered in a decreasing order.

DCT feature, HiFST [13] which extracts and combines the high-frequency DCT coefficients of the local area in the gradient image multiresolutionally into a vector with their absolute values sorted in an increasing order. Assume the patch size as $W \times W$ and the vector consisting of the absolute high-frequency DCT coefficients of $p_{i,j}$ as $H_{i,j}^W$. The multi-scale HiFST decomposition is

$$F_{i,j} = sort(\bigcup_{r=1}^{m} H_{i,j}^{M_r}) \tag{4}$$

where $sort(\cdot)$ is a function for incremental ranking, m is the lay total, $M_r = 2^{2+r} - 1$ controls the patch size in different layers.

$F_{i,j}$ is normalized to effectively differentiate clear and blurred regions and the final blur measure is computed through a entropy-weighted max pooling,

$$D_{i,j} = T_{i,j}.\omega_{i,j} \tag{5}$$

where $T_{i,j}$ is obtained by the max pooling of $F_{i,j}$ among $\frac{W^2+W}{2}$ frequency layers and $\omega_{i,j}$ is the entropy of T computed by $k \times k$ neighborhood surrounding the corresponding pixel in T. Please refer to [13] for more details on HiFST.

2.2 Learning the Blur Measure with Neural Network

We take the NN approach to train a classifier as the metric to estimate the initial blur degree with the blur descriptor. Especially, backpropagation (BP) neural network [18] containing only three layers (input, output, and hidden layers) is used. It is simple to deploy, fast to train, and also experimentally robust.

3 Saliency Constrained Refinement of Blur Measurement

3.1 Relationship between Saliency and Blur Measure

The blur metric based on the above NN classifier may fail due to flat texture which constitutes a common difficulty for all existing blur measurement methods.

Interestingly, we can see clearly the whole blurred object(s) when observing a partially blurred image even with clear flattened texture. This observation intrigues us to look into the role of of visual saliency in separating the blurred and clear regions. A salient object is the only one viewed by a human due to the difficulty of viewing more than one simultaneously. Therefore, the corresponding saliency map can be exploited to refine the measurement globally.

Figure 3 illustrates the principle of saliency for blur refinement. The accurate blur measurement (Figure 3b) is obtained with the state-of-the-arts method [13] for fair demonstration. This measurement and the saliency map (Figure 3c) show the same intensity distributions which can be justified by comparing them directly (Figure 3d): The difference of blur measures increases when the difference of the saliencies increases. The bigger difference of the saliencies means the more possibility of the pixels belonging to different objects and having different blur degrees. Therefore, saliency can be used to refine the blur measures so that the pixels having the similar saliencies can be refined to have similar measures.

Fig. 3: Illustration of the relationship between blur measure and saliency. (a): The example image; (b) the accurate measurement of (a); (c): the saliency map; (d): the statistical relation between saliency and blur measure.
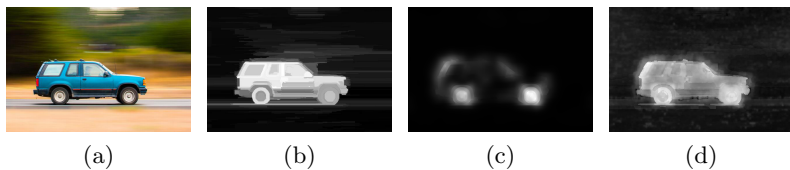

Fig. 4: An example of the saliency constrained blur measurement refinement. (a): The example image; (b): the saliency map; (c): the inaccurate measurement by Golestaneh and Karam [13]; (d): the refined measurement.

The observation from Figure 3 is important for images with flat texture which show low frequency distributions as the blurred regions. Saliency consistencies can then be adopted to refine the measures. Figure 4 shows the example car has apparent flat texture (Figure 4a). The state-of-the-art method (Figure 4c)

cannot correctly measure it with many false estimations in the car, while our saliency constrained method (Figure 4d) obtains more visually correct results. The details of the saliency constrained refinement is described as follows.

3.2   The Saliency Constrained Refinement Strategy

Assuming the blur measure and saliency of the current pixel p_c being B_c and S_c and its N neighbors being $\mathbb{B}_c = \{B_1, B_2, \ldots, B_N\}$ and $\mathbb{S}_c = \{S_1, S_2, \ldots, S_N\}$ respectively. The refined B_c, $\hat{B}_c$, relies on not only its own initial blur measurement (prior) $P(B_c) = B_c$, but also S_c and $\mathbb{S}_c$, considering that the neighbors sharing the similar saliency should be similarily blurred. Therefore, the refinement process can be formulated as a maximum a posteriori (MAP) estimation,

$$\begin{aligned} P(B_c|\mathbb{B}_c, S_c, \mathbb{S}_c) &\propto P(\mathbb{B}_c; S_c; \mathbb{S}_c|B_c)P(B_c) \\ &\propto P(\mathbb{B}_c|B_c)P(S_c, \mathbb{S}_c|B_c)P(B_c) \end{aligned} \tag{6}$$

by assuming the conditional independency between the blur measure and the saliency. It can be further rewritten as:

$$P(B_c|\mathbb{B}_c, S_c, \mathbb{S}_c) \propto \prod_i P(B_i|B_c)(\prod_i P(S_i|B_c))P(S_c|B_c)P(B_c). \tag{7}$$

The closer B_i to B_c is, the more similar they are. The same observation can also be applied to S_i and S_c. Therefore, this type of relationship can be described as a Gaussian distribution,

$$P(\tau_j|\tau_c) \propto e^{-(\tau_j - \tau_c)^2}, \tau \in \{B, S\} \tag{8}$$

Now taking the posterior distributions, $P(B_j|B_c)$ and $P(S_j|S_c)$ (Equation 8) into Equation 7 and deriving its right side w.r.t B_c, we can then deduce the solution of $\hat{B}_c$ by Equation 7 via MAP estimation as

$$\begin{aligned} \hat{B}_c &\propto \underset{B_c}{\arg\max} \ln \prod_i P(B_i|B_c) + \ln \prod_i P(S_i|B_c) + \ln P(S_c|B_c) + \ln P(B_c) \\ &\cong \frac{\sum_i B_i}{N} + g(S_c, \mathbb{S}_c, B_c) \end{aligned} \tag{9}$$

where $g(S_c, \mathbb{S}_c, B_c)$ is the derivative of $\ln \prod_i P(S_i|B_c) + \ln P(S_c|B_c)$ w.r.t. B_c and determines the contributions of S_c and $\mathbb{S}_c$ to B_c.

Directly solving $g(S_c, \mathbb{S}_c, B_c)$ in Equation 9 is difficult due to the lacking of proper formulations of the embedded distributions. Intuitively, the smaller the saliency difference between S_i and S_c is, the smaller the difference between B_i and B_c is and thus the more $\hat{B}_c$ relies on B_i. In addition, a neighboring pixel is invalid as the reference when its saliency is significantly different from S_c, i. e., they lie on different objects. Therefore, $g(\cdot)$ can be formulated as

$$g(S_c, \mathbb{S}_c, B_c) = \sum_i max((e^{-\alpha(\|S_i - S_c\|)} - \beta), 0)B_c \tag{10}$$

where α translates the weight to be in a more distinugishable range and β controls the threshold for a valid reference.

3.3   Our Measurement Algorithm

Directly applying the above pixel oriented idea can be slow and, therefore, we take a superpixel strategy to accelerate the measurement process, considering the regional constancy in the image. Accordingly, the average measure and saliency of each superpixel from the initial blur measurement and saliency map are adopted for a superpixel-scale refinement. Then a bilateral filtering is used so that pixel-consistent but also structure-preserving results can be obtained.

Figure 5 gives the pipeline. Both Figure 5b and Figure 5c are obtained in the pixel scale by the pre-trained NN classifier and the saliency detection algorithm, respectively. Then the superpixel-scale refinement (Figure 5e) is iteratively applied according to the superpixel segmentation (Figure 5d) until the saliency consistency is satisfied. Finally, bilateral filtering is applied to obtain the measurement output (Figure 5f).
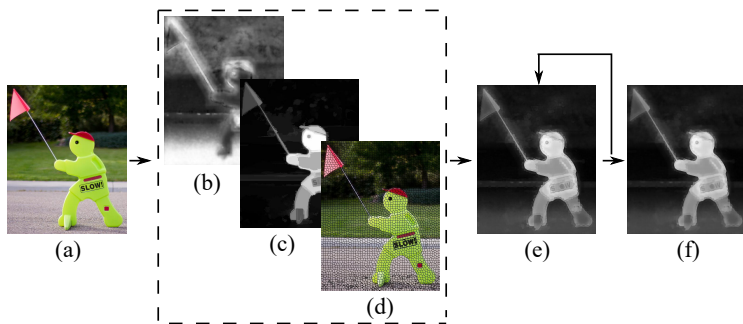

Fig. 5: The pipeline of the proposed measurement algorithm. (a): The input partially blurred image; (b): the blur measurement; (c): the saliency map; (d): the superpixel segmentation; (e): the superpixel-scale refinement; and (f): the final output by the bilateral filter.

The supperpixel segmentation is obtained by the simple linear iterative clustering (SLIC) [19] and the saliency is detected by the discriminative regional feature integration (DRFI) [20]. A sourrounding superpixel can be from a different object in a high probability if its saliency is larger than that of the center suprpixel. In this case, the surrounding superpixel is invalid as the reference for Equation 10. Therefore, only the saliency of the superpixel with the max weight among neighbors is considered in Equation 10.

4   Experimental Results

Our method is formatted into two versions to show the performance of the saliency constrained refinement: *Our method without saliency* and *our method*. The former does not include the saliency constrained refinement while the latter is

the complete implementation of the proposed method. Previous algorithms to be compared include Liu *et al.* [3], Chakrabarti *et al.* [21], Su *et al.* [11], Shi *et al.* [4], Shi *et al.* [9], Tang *et al.* [15], Yi and Eramian [10] and Golestaneh and Karam [13]. The experiments were performed on the dataset of Shi *et al.* [4] which includes 294 motion blurred images and 704 defocus ones. In the implementation of our method, there are 20 nodes in the middle layer of BP with the max iteration, error threshold and learning ratio are set to 5000, 0.00006 and 0.004, respectively; the approximate size of each supperpixel, and α and β in Equation 10 are set to 10, 10 and 0.7, respectively; and the max iterations of the refinement is set to 3.
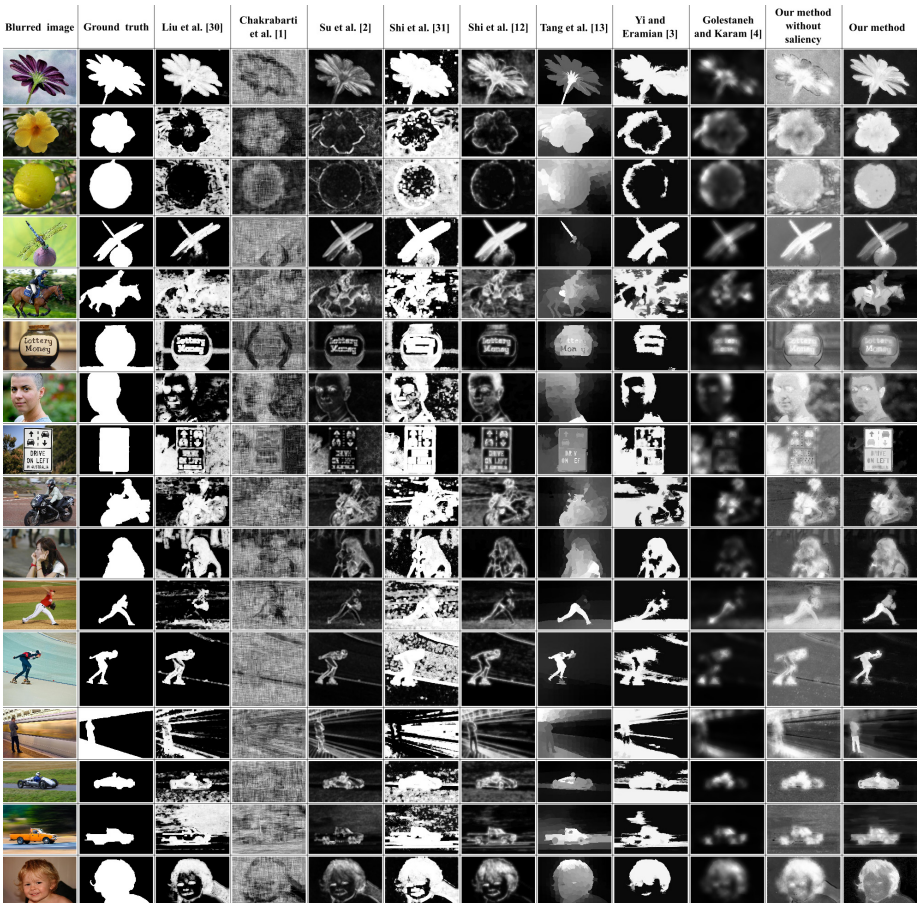

Fig. 6: Qualitative comparison of the blur measurement performance.

Figure 6 shows the qualitative results of different methods. The top-listed four images are richly textured while the remaining 12 images contain some flat

textures (Those nearer to the bottom generally contain larger flat textures). The results of *our method without saliency* are comparable with the best ones of the previous methods and *our method* produced consistently better results without significant wrong measure than all other methods, partially due to the benefits of the saliency constrained global refinement.

Statistical comparison of different methods was performed using the precision-recall curve. Two types of comparison were taken: One is the comparison with all images in the dataset of Shi *et al.* [4] and the other is with only the images containing apparently flat textures (Figure 7) in the dataset (There are 82 images in total). Figure 8 shows *our method* achieved better results than all other methods for both comparisons and even *our method without saliency* performed better for the flatten textures images than Golestaneh and Karam [13] (Figure 8b). The robust extreme points introduced in our approach are significantly sparse in the blurred area and thus are extraordinarily effective for the low textured images.

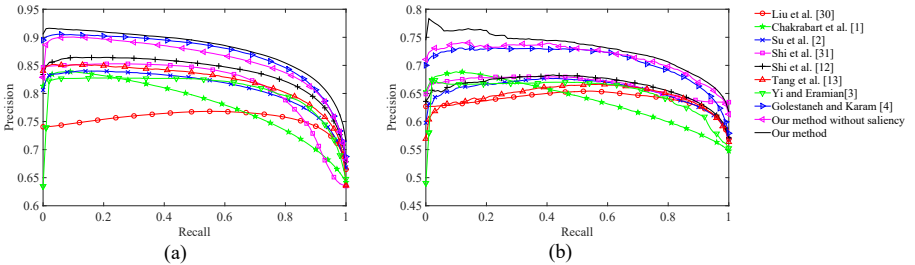Fig. 7: Examples of the image with flat textures for the statistical comparison.

Fig. 8: Statistical comparison of the blur measurement performance. (a): Comparison with all images in the dataset; and (b): comparison with only the images containing flat textures (Figure 7).

5   Conclusion

This paper proposes three contributions: 1) A novel blur descriptor consisting of the number of extreme points, the vector of all singular values, and the entropy-weighted pooling of high frequency DCT coefficients; 2) a global refinement

strategy using the saliency map to update the initial blur measurement; and 3) a novel blur measurement method which integrates the NN classifier as a blur metric for the initial blur measurement, the superpixel-scale refinement with the saliency constraint and the bilateral filtering for final structure-preserving oriented output. Experiments show the proposed method is effective, especially for the images with flat textures.

Our method can be very slow for a large image and thus we will study faster algorithm in the future. Current method may not work well when saliency detection fails. Therefore, more effective global refinement strategy is also one of our future directions for a better blur measurement.

Acknowledgement

References

1. Tai, Y.W., Brown, M.S.: Single image defocus map estimation using local contrast prior. In: ICIP. (2009) 1797–1800
2. Park, J., Tai, Y.W., Cho, D., Kweon, I.S.: A unified approach of multi-scale deep and hand-crafted features for defocus estimation. In: CVPR. (2017) 1063–6919.
3. Liu, R., Li, Z., Jia, J.: Image partial blur detection and classification. In: CVPR. (2008) 1–8
4. Shi, J., Xu, L., Jia, J.: Discriminative blur detection features. In: CVPR. (2014) 2965–2972
5. Vu, C.T., Phan, T.D., Chandler, D.M.: s_3: A spectral and spatial measure of local perceived sharpness in natural images. IEEE Trans. Image Processing **21**(3) (2012) 934–945
6. Xu, G., Wang, X., Xu, X.: Improved bi-dimensional emd and hilbert spectrum for the analysis of textures. Pattern Recognition **42**(5) (2009) 718–734
7. Elder, J.H., Zucker, S.W.: Local scale control for edge detection and blur estimation. IEEE Trans. Pattern Analysis and Machine Intelligence **20**(7) (1998) 699–716
8. Zhuo, S., Sim, T.: Defocus map estimation from a single image. Pattern Recognition **44**(9) (2011) 1852 – 1858
9. Shi, J., Xu, L., Jia, J.: Just noticeable defocus blur detection and estimation. In: CVPR. (2015) 657–665
10. Yi, X., Eramian, M.: LBP-based segmentation of defocus blur. IEEE Trans. Image Processing **25**(4) (2016) 1626–1638
11. Su, B., Lu, S., Tan, C.L.: Blurred image region detection and classification. In: ACM Multimedia. (2011) 1397–1400
12. Fang, X., Shen, F., Guo, Y., Jacquemin, C., Zhou, J., Huang, S.: A consistent pixel-wise blur measure for partially blurred images. In: ICIP, (2014) 496–500
13. Alireza Golestaneh, S., Karam, L.J.: Spatially-varying blur detection based on multiscale fused and sorted transform coefficients of gradient magnitudes. In: CVPR. (2017) 5800–5809

14. Javaran, T.A., Hassanpour, H., Abolghasemi, V.: Automatic estimation and segmentation of partial blur in natural images. The Visual Computer **33**(2) (2017) 151–161
15. Tang, C., Wu, J., Hou, Y., Wang, P., Li, W.: A spectral and spatial approach of coarse-to-fine blurred image region detection. IEEE Signal Processing Letters **23**(11) (2016) 1652–1656
16. Bhattacharya, S., Venkatesh, K., Gupta, S.: Visual saliency detection using spatiotemporal decomposition. IEEE Trans. Image Processing **27**(4) (2018) 1665–1675
17. Borji, A.: What is a salient object? a dataset and a baseline model for salient object detection. IEEE Trans. Image Processing **24**(2) (2015) 742–756
18. Hecht-Nielsen, R.: Theory of the backpropagation neural network. Neural Networks **1**(1) (1988) 445–445
19. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Ssstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Analysis and Machine Intelligence **34**(11) (2012) 2274–2282
20. Jiang, H., Wang, J., Yuan, Z., Wu, Y.: Salient object detection: A discriminative regional feature integration approach. In: CVPR. (2013) 2083–2090
21. Chakrabarti, A., Zickler, T., Freeman, W.T.: Analyzing spatially-varying blur. In: CVPR. (2010) 2512–2519