

FAST DEFOCUS MAP ESTIMATION

Ding-Jie Chen, Hwann-Tzong Chen, and Long-Wen Chang

Department of Computer Science, National Tsing Hua University, Taiwan

ABSTRACT

This paper presents a fast algorithm for deriving the defocus map from a single image. Existing methods of defocus map estimation often include a pixel-level propagation step to spread the measured sparse defocus cues over the whole image. Since the pixel-level propagation step is time-consuming, we develop an effective method to obtain the whole-image defocus blur using oversegmentation and transductive inference. Oversegmentation produces the superpixels and hence greatly reduces the computation costs for subsequent procedures. Transductive inference provides a way to calculate the similarity between superpixels, and thus helps to infer the defocus blur of each superpixel from all other superpixels. The experimental results show that our method is efficient and able to estimate a plausible superpixel-level defocus map from a given single image.

Index Terms— Defocus estimation

1. INTRODUCTION

Common causes of image blur include camera shake [7, 17], object motion [4, 9, 17], and defocus [2, 6, 8, 10, 11, 15, 16, 17, 19, 20]. In this paper, we aim to study the problem of defocus blur, which provides important clues for estimating relative depths. Defocus maps are also useful for image manipulation. For example, refocus, segmentation, matting, background decolorization, and salient region detection are just some of the possible applications [8, 12, 13, 14].

A 3D point captured in an image will look sharp if it locates on the focal plane, because the rays from this point will converge to the same spot on the camera sensor. In contrast, a 3D point will look blur in the image if it deviates from the focal plane. This kind of blur is called the *defocus blur*. The blur pattern on the sensor is called the circle of confusion. The diameter c of the circle of confusion is proportional to the distance from the 3D point to the focal plane. This means the diameter of the circle of confusion, the distance from the 3D point to the focal plane, and the degree of blur are positively correlated.

The defocus blur is usually estimated around the edge pixels. The resulting estimate thus yields a sparse defocus map. Therefore, a subsequent propagation step is needed to spread out the sparse defocus values to the entire image. However,

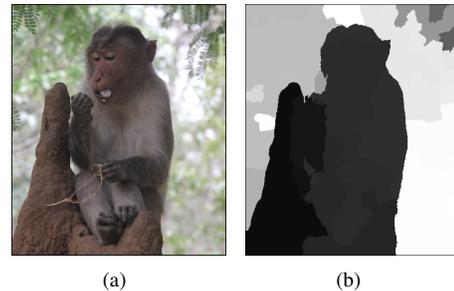


Fig. 1. An example of defocus map estimation. (a) Input image. This image is downloaded from flickr.com (<https://goo.gl/RZb9wl>). (b) The estimated superpixel-level defocus map. Note that brighter intensity values correspond to greater degrees of defocus.

we observe that the existing methods usually spend a lot of time on the propagation step, and this will limit the usefulness of defocus blur estimation. The excessively large computational cost on the propagation step motivates us to explore a faster strategy for propagating the sparse defocus map.

1.1. Related Work

The literature on defocus blur estimation can be roughly divided into two categories: gradient based methods and frequency based methods.

Gradient based methods [2, 6, 8, 10, 11, 16, 20] exploit the fact that the gradient magnitudes at the edge locations decline sharply after blurring. Elder and Zucker [6] propose to measure image blur using the first and second order gradients on edges, and hence generate a sparse defocus map. Bae and Durand [2] use a multi-scale edge detector to estimate the blur of edge pixels, and then propagate the blur cues over the image by an edge-aware colorization method. Namboodiri and Chaudhuri [10] model the defocus blur as a heat diffusion process. They estimate the defocus blur at edge locations according to the inhomogeneous inversion heat diffusion, and then propagate the defocus blur using graph-cuts. Tai and Brown [16] use local contrast prior on edge gradient magnitudes to define the amount of defocus blur. Then, they adopt MRF propagation to obtain the dense defocus map. The method of Zhuo and Sim [20] is based on the response of the

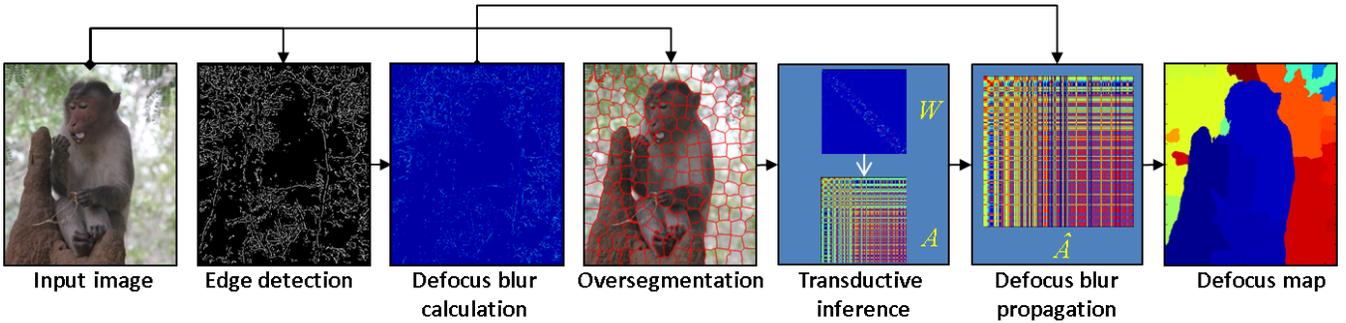


Fig. 2. The flowchart of the proposed method. *Edge detection:* The result of the structured edge prediction. *Defocus blur calculation:* The estimated sparse defocus map. *Oversegmentation:* The result of SLIC oversegmentation. *Transductive inference:* To transform a weight matrix W into an affinity matrix A . *Defocus blur propagation:* To propagate the sparse defocus map in superpixel-level based on the affinity matrix \hat{A} .

Gaussian blur on edge pixels for characterizing defocus blur. The matting Laplacian technique is used to interpolate the defocus blur of non-edge pixels. Jiang *et al.* [8] measure the pixel-level defocus blur with multiple Gaussian kernels. They directly define the region-level defocus blur of each region by the pixel-level blur values on the region boundary and edge pixels. Recently, Peng *et al.* [11] estimate the pixel blur using the difference between the before and after multi-scale Gaussian smoothing. They employ morphological reconstruction and the guided filter for blur map refinement.

Frequency based methods [15, 17, 19] draw upon the observation that blurring decreases high frequency components and increases low frequency components. For example, Zhang and Hirakawa [17] use double discrete wavelet transform to estimate defocus blur. Zhu *et al.* [19] measure the probability of blur scale via the localized Fourier spectrum. The blur scale of each pixel is then selected by a constrained pairwise energy function. Shi *et al.* [15] estimate the small defocus blur via the statistics of sparse just noticeable blur features. The result is further smoothed with an edge-preserving filter.

1.2. Our approach

In general, the aforementioned gradient based defocus estimation methods contain two phases. The first phase generates the sparse defocus blur estimation at edge locations. The second phase propagates the estimated sparse defocus blur to the entire image. The existing methods usually take too much time in the second phase. However, the pixel-level defocus information is not always required in practical applications, for example foreground/background segmentation and salient region detection. Therefore, instead of using graph-cuts, colonization, or matting Laplacian methods for pixel-level defocus propagation as in the previous work, we integrate oversegmentation and transductive inference to achieve highly efficient propagation at the superpixel level.

2. DEFOCUS BLUR ESTIMATION

Our method includes two phases, namely the *sparse defocus blur estimation* and the *defocus blur propagation*. The first phase aims to estimate the defocus blur on the edge pixels, and we adopt the gradient based approach in this phase. The second phase incorporates a new mechanism to propagate the sparse defocus blur to the whole image.

2.1. Sparse defocus blur estimation

2.1.1. Edge detection

Since the defocus blur is measured at the edge locations, we first apply edge detector to the image and extract the set E of edge pixels. We adopt Canny edge detector [3] and the structured edge prediction [5] for their efficiency and performance. The results from both methods are shown in the experiments.

2.1.2. Defocus blur calculation

Consider an edge function $e(x) = \alpha h(x) + \beta$, where α and β denote the amplitude and the offset of the edge respectively, $h(\cdot)$ denotes the step function, and x is a pixel location. The defocus blur can be modeled as a convolution of an edge pixel x with a Gaussian kernel $g(x, \sigma)$, where the standard deviation σ is proportional to the circle of confusion diameter c . The blurred edge is thus defined as $b(x) = e(x) \otimes g(x, \sigma)$. The unknown standard deviation σ means the blurriness on the edge pixel and can be used to represent the degree of defocus blur on that edge pixel.

If we re-blur the edge pixel using another Gaussian kernel, then the gradient of the re-blurred edge is represented as

$$\begin{aligned}
 \nabla(b(x) \otimes g(x, \sigma_r)) &= \nabla(e(x) \otimes g(x, \sigma) \otimes g(x, \sigma_r)) \\
 &= \frac{\alpha}{\sqrt{2\pi(\sigma^2 + \sigma_r^2)}} \exp\left(-\frac{x^2}{2(\sigma^2 + \sigma_r^2)}\right),
 \end{aligned} \tag{1}$$

where σ_r denotes the standard deviation of the re-blur Gaussian kernel. Zhuo and Sim [20] observe that the gradient magnitude ratio R between the original blurred edge and the re-blurred edge has maximum value at the *edge locations*. Furthermore, the maximum value is given by

$$R = \frac{|\nabla b(x)|}{|\nabla b(x) \otimes g(x, \sigma_r)|} = \sqrt{\frac{\sigma^2 + \sigma_r^2}{\sigma^2}}. \quad (2)$$

Hence, we can calculate the unknown blur σ using the gradient magnitude ratio R at the edge locations by

$$\sigma = \frac{\sigma_r}{\sqrt{R^2 - 1}}, \quad (3)$$

where σ_r is known and R can be derived from gradient magnitudes. Note that Eq. (3) is only applicable to the edge locations. Therefore, the intermediate outcome at the current step is just a sparse defocus map on edge pixels, as shown in the defocus blur calculation block in Fig. 2.

2.2. Defocus blur propagation

2.2.1. Oversegmentation

The goal of this step is to create the basic units (superpixels), and to define the similarity between the adjacent superpixels. Given an image, we first use the SLIC algorithm [1] to oversegment the image into a superpixel set $S = \{s_1, s_2, \dots, s_N\}$. According to the superpixel set S , we define a weighted connected graph $\mathcal{G} = (S, \mathcal{E}, \omega)$, where the vertex set is the superpixel set S and the edge set \mathcal{E} contains pairs of every two adjacent superpixels. That is, each vertex s_i denotes one single superpixel in S , and each edge $e_{ij} \in \mathcal{E}$ denotes the adjacency relationship between superpixels s_i and s_j . The weight function $\omega : \mathcal{E} \rightarrow [0, 1]$ defines the corresponding weight ω_{ij} to each edge e_{ij} , expressed in terms of feature similarities. We can thus define the weight matrix as $W = [\omega_{pq}]_{N \times N}$.

2.2.2. Transductive inference

The above weight matrix W describes the similarity between any two *adjacent* superpixels. According to the transductive inference method proposed by Zhou *et al.* [18], we can obtain an N -by- N affinity matrix A to describe the transductive similarity between any two superpixels, no matter they are adjacent or not. The affinity matrix A can be defined by

$$A = (D - \gamma W)^{-1} I, \quad (4)$$

where D is the diagonal matrix with each diagonal entry equal to the row sum of W , γ is a parameter in $(0, 1)$, and I is the N -by- N identity matrix. Since the affinity matrix encodes the transductive similarity between any two superpixels, it is possible to adjust the defocus blur of any superpixel pair using their affinity in A .

2.2.3. Defocus blur propagation

For each superpixel s_i , we define its initial defocus blur f_{s_i} as

$$f_{s_i} = \underset{x \in E_i}{\text{median}}\{f_x\}, \quad (5)$$

where E_i denotes the set of the interior edge pixels of the superpixel s_i . The defocus blur f_x on edge pixel x is computed using Eq. (3). Here we use the *median* to reduce the impact from the outliers. Now, take the affinity information into account, the propagated defocus blur \hat{f}_{s_i} of superpixel s_i is defined as

$$\hat{f}_{s_i} = \hat{A} \cdot [f_{s_1}, f_{s_2}, \dots, f_{s_N}]^T, \quad (6)$$

where $\hat{A} = [\hat{a}_{ij}]_{N \times N}$ denotes the modified affinity matrix with two conditions: *i*) column j is reset to zeros if E_j is empty; *ii*) each row is summed to one. The Eq. (6) means that the defocus blur of each superpixel is derived from not only its neighboring superpixels but also all other superpixels except the superpixels contain no edge pixels. An example of the final defocus map is shown in Fig. 2.

We summarize the steps of the proposed defocus blur estimation:

1. Edge detection: Detect edge pixels using Canny edge detector [3] or structured edge prediction [5].
2. Defocus blur calculation: Apply Eq. (2) and Eq. (3) to the edge pixels and generate the sparse defocus map.
3. Oversegmentation: Oversegment the image using SLIC [1] and construct the weight matrix W .
4. Transductive inference: Calculate the affinity matrix A by Eq. (4).
5. Defocus blur propagation: Use Eq. (5) and Eq. (6) to re-estimate the defocus blur of each superpixel.

3. EXPERIMENTAL RESULTS

We compare the proposed superpixel-level defocus blur estimation with some of the state-of-the-art methods. The evaluations are performed with respect to the execution time and visualization results. We use the blurry images from [20] and [8] for the experiments. There are five images in [20] and seven images in [8].

In this work, we set the number of superpixels $N = 200$ and the parameter $\gamma = 0.9999$ in Eq.(4). We calculate the weight matrix using the $YCbCr$ color features and Gaussian weighting function.

3.1. Execution Time Comparison

Fig. 3 shows that our method is about 5 to 13 times faster than previous methods. Fig. 3 also reveals the time-consuming

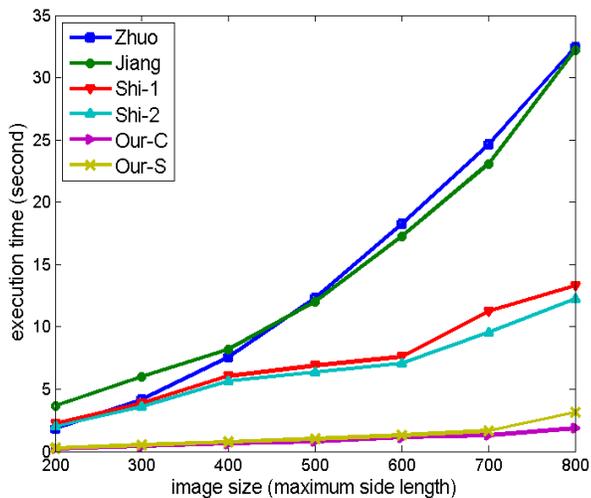


Fig. 3. Execution time comparison of our approach with other methods. The references of the diagram legends are Zhuo [20], Jiang [8], Shi [15], and our approach, respectively. Shi-1 and Shi-2 represent the case-1 and case-2 results of their released code. Our-C and Our-S represent the adopted edge detectors with Canny [3] or structured [5].

problem of pixel-level defocus estimation on large images. In general, the computation time of our method to process a 500×500 image is less than 1 second on an Intel Core i5-2500 CPU running at 3.30 GHz.

3.2. Visual Comparison

Fig. 4 shows the estimated defocus maps of different methods. For better visualization, we normalize the estimated defocus blur of all methods to the range of $[0, 1]$. The higher (brighter) intensity values represent the stronger defocus blur. With superpixel-level defocus blur propagation, superpixels that contain similar features could obtain similar defocus blur values. It can be seen that our defocus blur estimation is able to produce visually plausible defocus maps, and furthermore, the estimated defocus maps are readily usable for salient region detection or foreground/background segmentation.

4. CONCLUSION

We have shown that the proposed method can greatly speed up the propagation step of the gradient based defocus blur estimation. With the aids of the techniques from oversegmentation and transductive inference, the defocus blur propagation step becomes much more efficient. The experimental results show that our method for estimating superpixel-level defocus maps performs well in visualization results and computation time.

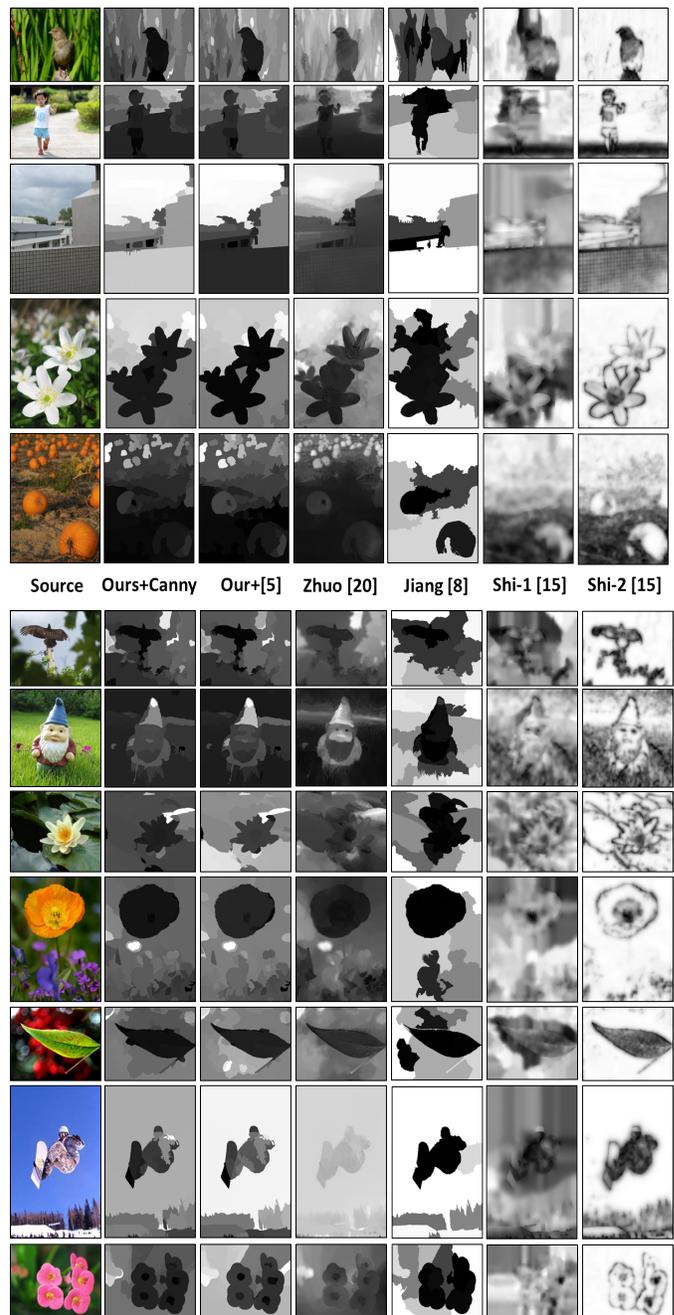


Fig. 4. Visual comparison of our approach with other methods. The top five test images are from [20] and the bottom seven images are from [8]. We normalize the estimated defocus blur of all methods to the range of $[0, 1]$. The higher (brighter) intensity values represent the stronger blur. The columns from left to right are source images, our results with Canny edge detector, our results with structured edge prediction, the results using Zhuo and Sim [20], the results of Jiang [8], and the results of case-1 and case-2 of Shi [15].

5. REFERENCES

- [1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(11):2274–2282, 2012.
- [2] S. Bae and F. Durand. Defocus magnification. *Comput. Graph. Forum*, 26(3):571–579, 2007.
- [3] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, 1986.
- [4] S. Cho and S. Lee. Fast motion deblurring. *ACM Trans. Graph.*, 28(5):145:1–145:8, 2009.
- [5] P. Dollár and C. L. Zitnick. Fast edge detection using structured forests. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(8):1558–1570, 2015.
- [6] J. H. Elder and S. W. Zucker. Local scale control for edge detection and blur estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(7):699–716, 1998.
- [7] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph.*, 25(3):787–794, 2006.
- [8] P. Jiang, H. Ling, J. Yu, and J. Peng. Salient region detection by UFO: uniqueness, focusness and objectness. In *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*, pages 1976–1983, 2013.
- [9] H. T. Lin, Y. Tai, and M. S. Brown. Motion regularization for matting motion blurred objects. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(11):2329–2336, 2011.
- [10] V. P. Nambodiri and S. Chaudhuri. Recovery of relative depth from a single observation using an uncalibrated (real-aperture) camera. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA, 2008*.
- [11] Y. Peng, X. Zhao, and P. C. Cosman. Single underwater image enhancement using depth estimation based on blurriness. In *2015 IEEE International Conference on Image Processing, ICIP 2015, Quebec City, QC, Canada, September 27-30, 2015*, pages 4952–4956, 2015.
- [12] C. Rhemann, C. Rother, P. Kohli, and M. Gelautz. A spatially varying psf-based prior for alpha matting. In *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*, pages 2149–2156, 2010.
- [13] J. Shi, X. Tao, L. Xu, and J. Jia. Break ames room illusion: depth from general single images. *ACM Trans. Graph.*, 34(6):225, 2015.
- [14] J. Shi, L. Xu, and J. Jia. Discriminative blur detection features. In *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23-28, 2014*, pages 2965–2972, 2014.
- [15] J. Shi, L. Xu, and J. Jia. Just noticeable defocus blur detection and estimation. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 657–665, 2015.
- [16] Y. Tai and M. S. Brown. Single image defocus map estimation using local contrast prior. In *Proceedings of the International Conference on Image Processing, ICIP 2009, 7-10 November 2009, Cairo, Egypt*, pages 1797–1800, 2009.
- [17] Y. Zhang and K. Hirakawa. Blur processing using double discrete wavelet transform. In *2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013*, pages 1091–1098, 2013.
- [18] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf. Learning with local and global consistency. In *Advances in Neural Information Processing Systems 16 [Neural Information Processing Systems, NIPS 2003, December 8-13, 2003, Vancouver and Whistler, British Columbia, Canada]*, pages 321–328, 2003.
- [19] X. Zhu, S. Cohen, S. Schiller, and P. Milanfar. Estimating spatially varying defocus blur from A single image. *IEEE Transactions on Image Processing*, 22(12):4879–4891, 2013.
- [20] S. Zhuo and T. Sim. Defocus map estimation from a single image. *Pattern Recognition*, 44(9):1852–1858, 2011.