LBP-based Segmentation of Defocus Blur

Xin Yi, Mark Eramian

Abstract—**Defocus blur is extremely common in images captured using optical imaging systems. It may be undesirable, but may also be an intentional artistic effect, thus, it can either enhance or inhibit our visual perception of the image scene. For tasks such as image restoration and object recognition, one might want to segment a partially blurred image into blurred and non-blurred regions. In this paper, we propose a sharpness metric based on LBP (local binary patterns) and a robust segmentation algorithm to separate in- and out-of-focus image regions. The proposed sharpness metric exploits the observation that most local image patches in blurry regions have significantly fewer of certain local binary patterns compared to those in sharp regions. Using this metric together with image matting and multi-scale inference, we obtained high quality sharpness maps. Tests on hundreds of partially blurred images were used to evaluate our blur segmentation algorithm and six comparator methods. The results show that our algorithm achieves comparative segmentation results with the state-of-the-art and have big speed advantage over the others.**

Index Terms—**defocus, blur, segmentation, LBP, local binary patterns, image restoration, object recognition, out-of-focus, blurred,**

I. Introduction

Defocus blur in an image is the result of an out-of-focus optical imaging system. In the image formation process, light radiating from points on the focus plane are mapped to a point in the sensor, but light from a point outside the focus plane illuminates a non-point region on the sensor known as a circle of confusion. Defocus blur occurs when this circle becomes large enough to be perceived by human eyes.

In digital photography, defocus blur is employed to blur background and "pop out" the main subject using large-aperture lenses. However, this inhibits computational image understanding since blurring of the background suppresses details beneficial to scene interpretation. In this case, separation of the blurred and sharp regions of an image may be necessary so that post-processing or restoration algorithms can be applied without affecting the sharp regions, or so that image features are only extracted from in-focus regions.

Most current image deblurring methods assume that the blur is spatially invariant [13], [41], [23], [27], [26], [22]. Typically, a global blur kernel is estimated and the original image is reconstructed by fitting it to different image priors with maximimum *a posteriori* estimation. Methods that explicitly model spatially variant blur typically restore small image patches within which blur can be treated as invariant, and restored patches are stitched together [48], [2], [6], [47], [16]. Efficient and accurate detection of blurred or non-blurred regions is useful in several contexts including: 1) in avoiding expensive post-processing of non-blurred regions

(e.g. deconvolution) [12]; 2) in computational photography to identify a blurred background and further blur it to achieve the artistic bokeh effect [4], [43], particularly for high-depth-of-field cellular phone cameras; and 3) for object recognition in domains where objects of interest are not all-in-focus (e.g. microscopy images) and portions of the object which are blurred must be identified to ensure proper extraction of image features, or, if only the background is blurred, to serve as an additional cue to locate the foreground object and perform object-centric spatial pooling [40].

The purpose of segmentation of defocus blur is to separate blurred and non-blurred regions so that any aforementioned post-processing can be facilitated. Herein, this problem is explicitly addressed without quantifying the extent of blurriness and a new sharpness metric based on Local Binary Patterns (LBP) is introduced.

II. Related works

The most commonly seen approach for defocus segmentation literature is via local sharpness measurement. There are many works in this area in the past two decades and most of them can be found in the image quality assessment field where images are rated by a single sharpness score that should conform to the human visual perception. These applications only require a single sharpness value to be reported for a single image, thus most of the measures only rely on sharpness around local edges [14], [31], [29] or some distinctive image structures determined in the complex wavelet transform domain [17]. Similarly, the line spread profile has been adopted for edge blurriness measurement in image recapture detection [46]. Since most of these metrics are measured around edges, they cannot readily characterize sharpness of any given local image content unless edge-sharpness is interpolated elsewhere as was done in [4], [57].

Measures such as higher order statistics [21], variance of wavelet coefficients [52], local variance image field [53] have been used directly in segmentation of objects of interest in low-depth-of-field images. These local sharpness metrics are based on local image energy which means that the measures will not only decrease if the energy of the PSF (point spread function) decreases (becomes more blurry), but also decreases if the energy of the image content drops. Thus, a blurry, high-contrast edge region could have a higher sharpness score than an in-focus, low-contrast one. These metrics are suitable for relative sharpness measures, e.g. in focal stacking, but do not behave very well for local sharpness measure across various image contents. This deficiency has already been pointed out in [55].

Recently, the authors of [43], [28] proposed a set of novel local sharpness features, e.g. gradient histogram span, kurtosis,

Xin Yi and Mark Eramian are with the Department of Computer Science, University of Saskatchewan, Saskatoon, SK, Canada

for training of a naïve Bayes classifier for blur classification of local image regions. The sharpness is interpreted as the likelihood of being classified as sharp patch. Su et al. used singular value decomposition (SVD) of image features to characterize blur and simple thresholding for blurred region detection [44]. Vu et al. used local power spectrum slope and local total variation for the measure in both the spectral and spatial domains. The final sharpness is the geometric mean of the two measures [51].

Instead of measuring sharpness only based on local information, Shi et al. proposed to learn a sparse dictionary based on a large external set of defocus images and then use it to build a sparse representation of the test image patch. The final measure was the number of non-zero elements of the corresponding words [20].

Depth map estimation is another approach that can also be used for defocus blur segmentation. Zhuo et al. used edge width as a reference for depth measurement under the assumption that edges in blurred regions are wider than those in sharp regions [57]. They obtained a continuous defocus map by propagating the sharpness measures at edges to the rest of the image using image matting [25]. Bae and Durand's work is similar, but they computed edge width differently by finding the distance of second derivative extrema of opposite sign in the gradient direction [4]. These methods tend to highlight edges in places where the blur measure is actually smooth.

Zhu et al. tried to explicitly estimate the space-variant PSF by analyzing the localized frequency spectrum of the gradient field [56]. The defocus blur kernel is parameterized as a function of a single variable (e.g. radius for a disc kernel or variance for Gaussian kernel) and is estimated via MAP_k estimation [26]. Similar work can be found in [9] but the blur kernel is restricted to a finite number of candidates. Khosro et al. estimate the blur kernel locally via blind image deconvolution by assuming the kernel is invariant inside the local block. But instead of fitting the estimated kernel to a parameterized model, they quantized the sharpness through reblurring [5]. Florent et al. treat the blur kernel estimation as a multi-label energy minimization problem by combining learned local blur evidence with global smoothness constraints [11]. These methods are inherently slow because of the iterative nature.

Unlike [4], [56], [57], we do not intend to construct a depth map. Our goal is only to separate in-focus regions from regions of defocus blur; Also unlike [20], we do not rely on external defocus images; in this respect our work is most similar to [44], [28], [43], [51].

We postulate that local-based defocus blur segmentation methods to date have been limited by the quality of the sharpness measures which they employ.

We now review local metrics of image sharpness that have been recently introduced for the segmentation of blurred regions. Generally, they fall into one of three categories: gradient domain metrics, intensity domain metrics, and frequency domain metrics.

A. Gradient Domain Metrics

1) Gradient Histogram Span [54], [44]: The gradient magnitude of sharp images exhibits a heavy-tailed distribution [13],

[41], [26], [22] and can be modelled with a two component Gaussian mixture model (GMM):

$$G = a_1 e^{-\frac{(g-\mu_1)^2}{\sigma_1}} + a_2 e^{-\frac{(g-\mu_2)^2}{\sigma_2}}, \qquad (1)$$

where means $\mu_1 = \mu_2 = 0$, variance $\sigma_1 > \sigma_2$, g is the gradient magnitude, and G is the gradient magnitude distribution in a local region. The component with larger variance is believed to be responsible for the heavy-tailed property. Thus the local sharpness metric is:

$$m_{GHS} = \sigma_1. \qquad (2)$$

2) Kurtosis [43]: Kurtosis, which captures the "peakedness" of a distribution, also characterizes the gradient magnitude distribution difference. It is defined as:

$$K = \frac{E[(g-\mu)^4]}{E^2[(g-\mu)^2]} - 3 \qquad (3)$$

where the first term is the fourth moment around the mean divided by the square of the second moment around the mean. The offset of 3 is to cause the peakedness measure of a normal distribution to be 0. The derived local sharpness metric is:

$$m_K = \min(\ln(K(g_x) + 3), \ln(K(g_y) + 3)), \qquad (4)$$

where g_x, g_y are gradient magnitudes along x and y axis respectively.

B. Intensity Domain Metrics

1) Singular Value Decomposition (SVD) [28]: An image patch $\mathbf{P}$ can be decomposed by SVD:

$$\mathbf{P} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{v}_i^T, \qquad (5)$$

where $\mathbf{U}, \mathbf{V}$ are orthogonal matrices, $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal entries are singular values arranged in descending order, $\mathbf{u}_i$ and $\mathbf{v}_i$ are the column vectors of $\mathbf{U}$ and $\mathbf{V}$ respectively, and λ_i are the singular values of $\mathbf{\Lambda}$. It is claimed that large singular values correspond to the rough shape of the patch whereas small singular values correspond to details. The sharpness metric is:

$$m_{SVD}(k) = 1 - \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i}, \qquad (6)$$

where the numerator is the sum of the k largest singular values.

2) Linear Discriminant Analysis (LDA) [43]: By sampling a set of blurred and non-blurred patches, this method finds a transform $\mathbf{W}$ that maximizes the ratio of the between-class variance S_b to the within-class variance S_w of the projected data with each variance:

$$S_b = \sum_{j=1}^{2} (\mu_j - \mu)^T (\mu_j - \mu),$$

$$S_w = \sum_{j=1}^{2} \sum_{i=1}^{N_j} (x_i^j - \mu_j)^T (x_i^j - \mu_j), \qquad (7)$$

where $j = 1$ represents the blurred class, $j = 2$ represents the sharp class, x_i is the intensity of the i-th pixel, and N_j is the

number of pixels in the corresponding region (see also Section 2.3 of [43]). This is solved by maximizing the ratio $\frac{det|S_b|}{det|S_w|}$ and the resulting column vectors of the projection matrix $\mathbf{W}$ are the eigenvectors of $S_w^{-1}S_b$. The final metric can be expressed as:

$$m_{LDA}(i) = \mathbf{w}_i^T \mathbf{p} \tag{8}$$

where $\mathbf{w}_i$ is the i-th column vector of matrix $\mathbf{W}$, and $\mathbf{p}$ is the vectorized patch intensity.

3) Sparsity [20]: This measure is based on sparse representation. Each patch is decomposed according to a learned over-complete dictionary which expressed as

$$\underset{\mathbf{u}}{\arg\min} ||\mathbf{p} - \mathbf{Du}||^2 + \lambda||\mathbf{u}||_1 \tag{9}$$

where $\mathbf{D}$ is the learned dictionary on a set of blur image patches. $\mathbf{p}$ is the vectorized patch intensity and $\mathbf{u}$ is the coeficients vector, each item of which is the weight used for the reconstruction. The reconstruction of a sharp patch requires more words than blurred patches. Thus the sharpness measure is defined as the number of non-zero elements in $\mathbf{u}$, i.e., the L_0 norm of $\mathbf{u}$.

$$m_S = ||\mathbf{u}||_0 \tag{10}$$

4) Total variation [51]: This metric is defined as

$$m_{TV} = \frac{1}{4} \max_{\xi \in P} TV(\xi)$$
$$\text{with } TV(\xi) = \frac{1}{255} \sum_{i,j} |x_i - x_j| \tag{11}$$

which is the maximum of the total variation of smaller blocks ξ (set as 2×2 in the original paper) inside the local patch P. The coefficient $\frac{1}{4}$ is a normalization factor since the largest TV of a 2×2 block is 4. The author argued that a non-probabilistic application of TV can be used as a measure of local sharpness due to its ability to take into account the degree of local contrast.

C. Frequency Domain Metrics

1) Power Spectrum [28], [43], [51]: The average of the power spectrum for frequency ω of an image patch is:

$$J(\omega) = \frac{1}{n}\sum_\theta J(\omega, \theta) \simeq \frac{A}{\omega^\alpha} \tag{12}$$

where $J(\omega, \theta)$ is the squared magnitude of the discrete Fourier transform of the image patch in the polar coordinate system, n is the number of quantizations of θ, and A is an amplitude scaling factor. It was shown that $\alpha = 2$ for sharp, natural images [49], [15], [8]. Since blurred images contain less energy in the high frequency components, the magnitude of their power spectra tend to fall off much faster with increasing ω, and the value of α is larger for such images. Rather than fitting a linear model to obtain α, the average of the power spectrum can be used instead as an indicator since the power spectra of blurred regions tend to have a steeper slope than for sharp regions, thus have a smaller average power. The metric is:

$$m_{APS} = \frac{1}{n}\sum_\omega \sum_\theta J(\omega, \theta) \tag{13}$$

In [51], [28], the authors directly use the fitted spectrum slope α as the measure. However, the author in [43] claimed that the average power spectrum is more robust to outliers and overfitting, thus we only evaluate m_{APS}.

D. Response Behavior of Sharpness Metrics

Given these sharpness metrics, we conducted a preliminary study to observe how they respond to different local image textures to see if they are limiting progress in blur detection as previously postulated. Since the proposed work is centred on local sharpness measure, this experiment excludes measures that rely on external informations, e.g. m_{LDA} and m_S.

Following the same methodology in [10], we assumed there are four common types of textures that appear in natural scenes, a random texture such as grass, a man-made texture, a smooth texture such as sky or fruit surface, and an almost smooth texture such as areas on the road sign (the characteristic of this texture is of low contrast and has more detail than pure smooth regions). Four such exemplar textures are shown in Figure 1. Gaussian blur of varying severity ($\sigma \in [0.1, 10.0]$) was applied to these image patches and each metric was computed for each texture and blur level. For the SVD-based metric, we tested with $k = 6$, that is, $m_{SVD}(6)$, but the response is similar for most values of k. The size of image patches were 21×21 pixels for all metrics.

Figure 2 shows the response of each metric to each of the four exemplar textures in Figure 1 over the tested range of σ. In addition, by evaluating 8000 sharp patches covering different scenes, an aggregate performance of these measures is also shown in Figure 2. The thick red curve shows the mean response over the 8000 patches and the dashed red curves show higher and lower quartiles (75th and 25th percentile). It can be seen from this figure that, in an aggregate manner, all measures decrease when blur increases (one exception is that m_K shows a slight increase after σ approaches 5). However, the aggregate data hides responses that are very different from the aggregate with m_{GHS} and m_K exhibiting minor to moderate non-monotonicity on some specific textures. Two patches are shown in Figure 3 with two levels of blur. The one with larger σ has larger m_K.

A smooth texture should elicit a constant, yet low response to the sharpness metrics since its appearance does not change with varying degrees of defocus blur, but the yellow curve shows big differences in responses for most of the sharpness metrics, with m_{GHS}, m_{TV} and m_{SVD} exhibiting the least variation. One would also expect that blurry regions would have smaller response than sharp regions, but that is not the case for all metrics. For example, at $\sigma = 1.5$ the range of values between the higher and lower quartile has a large overlap with range when $\sigma = 0$. In this respect, m_{APS} has the worse performance. Finally, none of the metrics are well-suited for measuring low contrast sharp regions, such as the almost smooth region in the example. This is because the low contrast region has very small intensity variance which leads to low gradient and low frequency response. The green and grey curve are almost inseparable for m_{GHS}, m_{SVD} and m_{TV}. This drawback is further shown in Figure 11. The low contrast
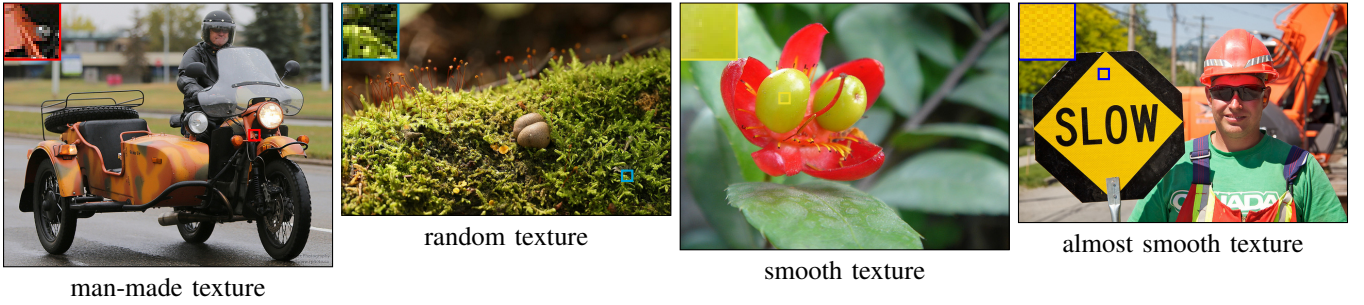
Fig. 1: Four commonly appeared textures in natural scenes.

mean —— higher and lower quartile - - - manmade —— random —— smooth —— almostsmooth ——
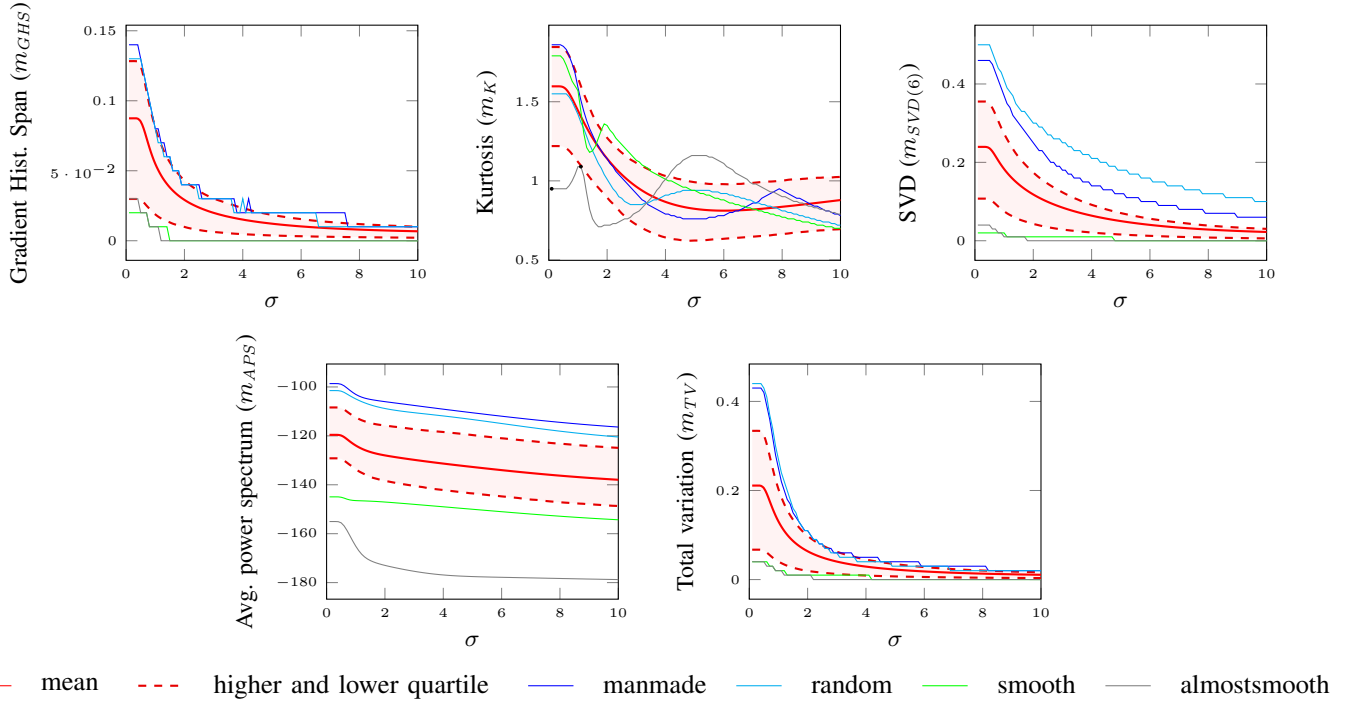
Fig. 2: Responses of different measures. The thick red curve shows the mean performance over 8000 patches and the dashed red line shows the higher and lower quartile. The responses to 4 exemplar patches are shown in blur, cyan, green, grey curves respectively.
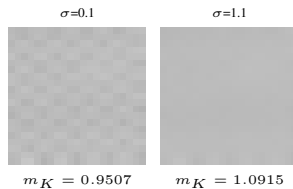

Fig. 3: An example of the non-monotonicity of the sharpness measure m_K. The patches showing here are the almost smooth patch under two levels of Gaussian blur as marked by black dots in m_K response in Figure 2.

yellow region of the road sign does not have a correct response for all measures even if it is in focus.

In the next section we present our new sharpness metric based on local binary patterns which is monotonic. The range of response values for blur patches has less intersection than that of sharp regions and it has a more appropriate response to low contrast region.

III. PROPOSED LBP BASED BLUR METRIC

Local Binary Patterns (LBP) [33] have been successful for computer vision problems such as texture segmentation [32], face recognition [3], background subtraction [19] and recognition of 3D textured surfaces [36]. The LBP code of a pixel (x_c, y_c) is defined as:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} S(n_p - n_c) \times 2^p$$

$$\text{with } S(x) = \begin{cases} 1 & |x| \geq T_{LBP} \\ 0 & |x| < T_{LBP} \end{cases} \quad (14)$$

where n_c is the intensity of the central pixel (x_c, y_c), n_p corresponds to the intensities of the P neighbouring pixels

located on a circle of radius R centered at n_c, and $T_{LBP} > 0$ is a small, positive threshold in order to achieve robustness for flat image regions as in [19]. Figure 4 shows the locations of the neighbouring pixels n_p for $P = 8$ and $R = 1$. In general, the points n_p do not fall in the center of image pixels, so the intensity of n_p is obtained with bilinear interpolation.
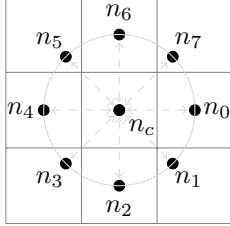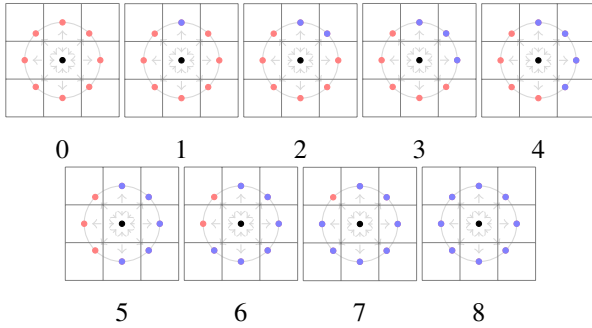

Fig. 4: 8-bit LBP with $P = 8, R = 1$.

Fig. 5: The uniform rotationally invariant LBP.

A rotation invariant version of LBP can be achieved by performing the circular bitwise right shift that minimizes the value of the LBP code when it is interpreted as a binary number [34]. In this way, number of unique patterns is reduced to 36. Ojala et al. found that not all rotation invariant patterns sustain rotation equally well [34], and so proposed using only uniform patterns which are a subset of the rotation invariant patterns. A pattern is uniform if the circular sequence of bits contains no more than two transitions from one to zero, or zero to one. The non-uniform patterns are then all treated as

one single pattern. This further reduces the number of unique patterns to 10 (for 8-bit LBP), that is, 9 uniform patterns, and the category of non-uniform patterns. The uniform patterns are shown in Figure 5. In this figure, neighbouring pixels are coloured blue if their intensity difference from centre pixel is larger than T_{LBP}, and we say that it has been "triggered", otherwise, the neighbours are coloured red.

Figure 6 shows the histogram of the nine uniform LBP patterns appearing in the blurred and non-blurred regions of 100 images randomly selected from a publicly available dataset of 704 partially blurred images [42], each of which is provided with a hand-segmented groundtruth image denoting the blurred and non-blurred regions. Bin 9 is the number of non-uniform patterns. The frequency of patterns 6, 7, 8, and 9 in blurred regions is noticeably less than that for sharp regions. The intuitive explanation for this is that in smoother areas, most neighbouring pixels will be similar in intensity to n_c, and the chance of a neighbour being triggered is lower, making the lower-numbered uniform patterns with fewer triggered neighbours more likely. Examples of the LBP histograms of specific sharp and blurred patches is given in Figure 7 which also exhibit this expected behaviour.

Our proposed sharpness metric exploits these observations:

$$m_{LBP} = \frac{1}{N} \sum_{i=6}^{9} n(LBP_{8,1}^{riu2} i) \tag{15}$$

where $n(LBP_{8,1}^{riu2} i)$ is the number of rotation invariant uniform 8-bit LBP pattern of type i, and N is the total number of pixels in the selected local region which serves to normalize the metric so that $m_{LBP} \in [0, 1]$. One of the advantages of measuring sharpness in the LBP domain is that LBP features are robust to monotonic illumination changes which occur frequently in natural images.

The threshold T_{LBP} in Equation 14 controls the proposed metric's sensitivity to sharpness. As shown in Figure 8, by increasing T_{LBP}, the metric becomes less sensitive to sharpness. However, there is a tradeoff between sharpness sensitivity and noise robustness, as shown in Figure 9. In situations where high sensitivity to sharpness is needed, a discontinuity-preserving noise reduction filter such as non-local means [7] should be employed.

Figure 10 shows our metric's response to various levels of blur ($T_{LBP} = 0.016$). There is a sharp fall-off between $\sigma = 0.2$ and $\sigma = 1.0$ which makes the intersection of response value range of sharp and blur much smaller than the other metrics. When σ approaches 2, responses for all patches shrinks to zero which facilitates segmentation of blurred and sharp regions by simple thresholding. Moreover, almost smooth region elicit a much higher response than smooth region compared with the other metrics. Finally, the metric response is nearly monotonic, decreasing with increasing blur, which should allow such regions to be distinguished with greater accuracy and consistency. Figure 11 shows maps of the local response of our metric and comparators for a sample image. Our metric has the most coherent response and responds the most consistently to low contras regions such as the road sign.
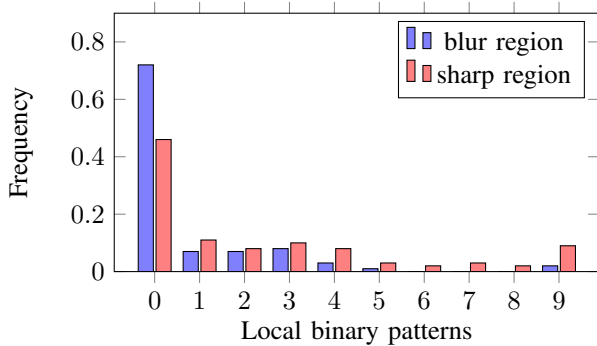

Fig. 6: LBP code distribution in blurred and sharp regions. Bins 0–8 are the counts of the uniform patterns; bin 9 is the count of non-uniform patterns. Data is sampled from 100 partial blurred images from [42].

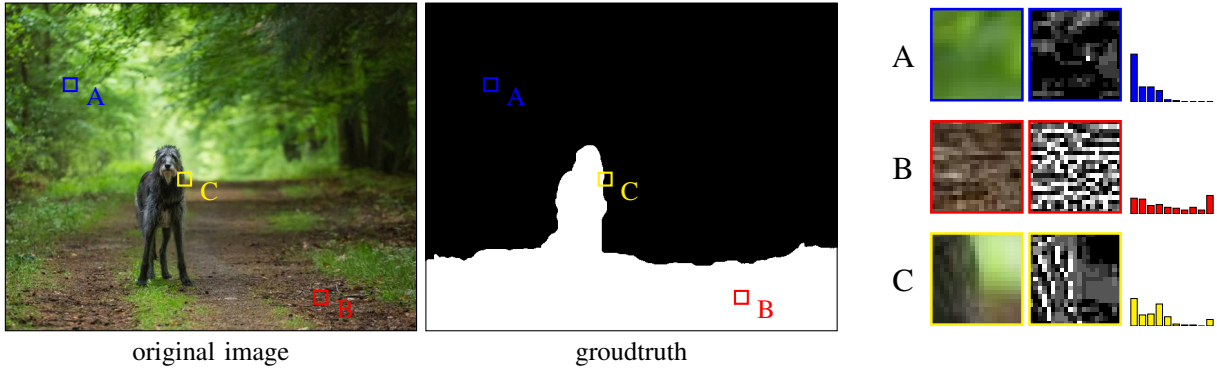original image   groudtruth

Fig. 7: Histogram of LBP patterns in three different patches which are sampled from blurred (A), sharp (B), and transitive (C) areas respectively. In the ground truth image, white denotes the sharp region and black the blurred region.
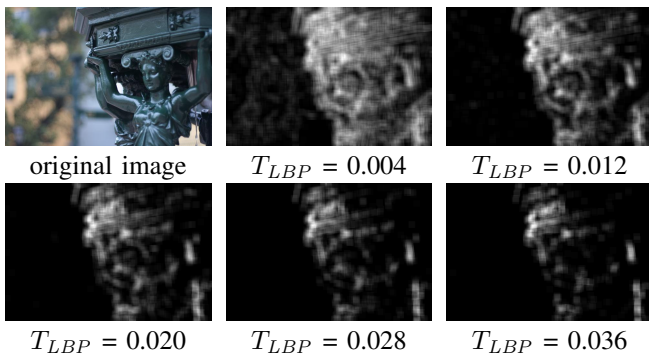

original image $T_{LBP} = 0.004$ $T_{LBP} = 0.012$

$T_{LBP} = 0.020$ $T_{LBP} = 0.028$ $T_{LBP} = 0.036$

Fig. 8: Response of m_{LBP} (Equation 15) for various values of threshold T_{LBP}. T_{LBP} determines the cutoff for the magnitude of intensity change that is considered an "edge", regardless of edge sharpness.

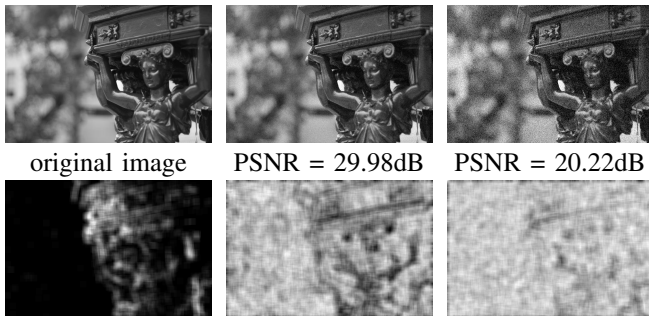original image   PSNR = 29.98dB   PSNR = 20.22dB

Fig. 9: Response of m_{LBP} in the presence of noise. Top: the original image and two copies corrupted by Gaussian noise; bottom: the corresponding sharpness maps. $T_{LBP} = 0.016$.

Fig. 10: Our metrics' response to the sample patches shown in Figure 1. As the same as in Figure 2, an aggregate response on 8000 sharp patches is also shown with the thick red curve showing the mean response and the dashed red curve showing the higher and lower quartile.

Sharpness Metric	Avg. Runtime
gradient histogram span (m_{GHS}) [43], [28]	273.19s
kurtosis (m_K) [43]	11.57s
singular value decomposition (m_{SVD}) [44]	38.66s
total variation (m_{TV}) [51]	50.00s
average power spectrum slope (m_{APS}) [43]	22.89s
our LBP-based metric (m_{LBP})	3.55s
our LBP-based metric (m_{LBP}, mex imp.)	26.5ms

TABLE I: Runtime comparison of various metrics. Note that the speed of our metric can be boosted by using integral image which makes the complexity independent of the size of local region.

Table I shows a comparison of the runtime of m_{LBP} and comparator metrics. Where available, author-supplied code for calculating the metrics was used, otherwise our own implementations were used. All implementations were in MATLAB and were unoptimized. 10 randomly selected images with approximate size of 640×480 pixels were tested on a Mac with 2.66 GHz intel core i5 and 8 GB memory. The average runtimes are reported in Table I.

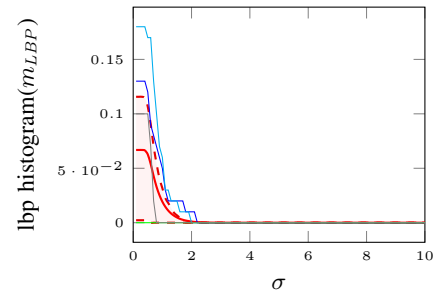The sharpness maps, response curves, and runtimes provide strong qualitative and quantitative evidence that our metric is superior. In the next section we present a blur segmentation method that achieves the state-of-the-art results by employing this metric.

IV. NEW BLUR SEGMENTATION ALGORITHM

This section presents our algorithm for segmenting blurred/sharp regions with our LBP-based sharpness metric; it is summarized in Figure 12. The algorithm has four main steps: multi-scale sharpness map generation, alpha matting initialization, alpha map computation, and multi-scale sharpness inference.
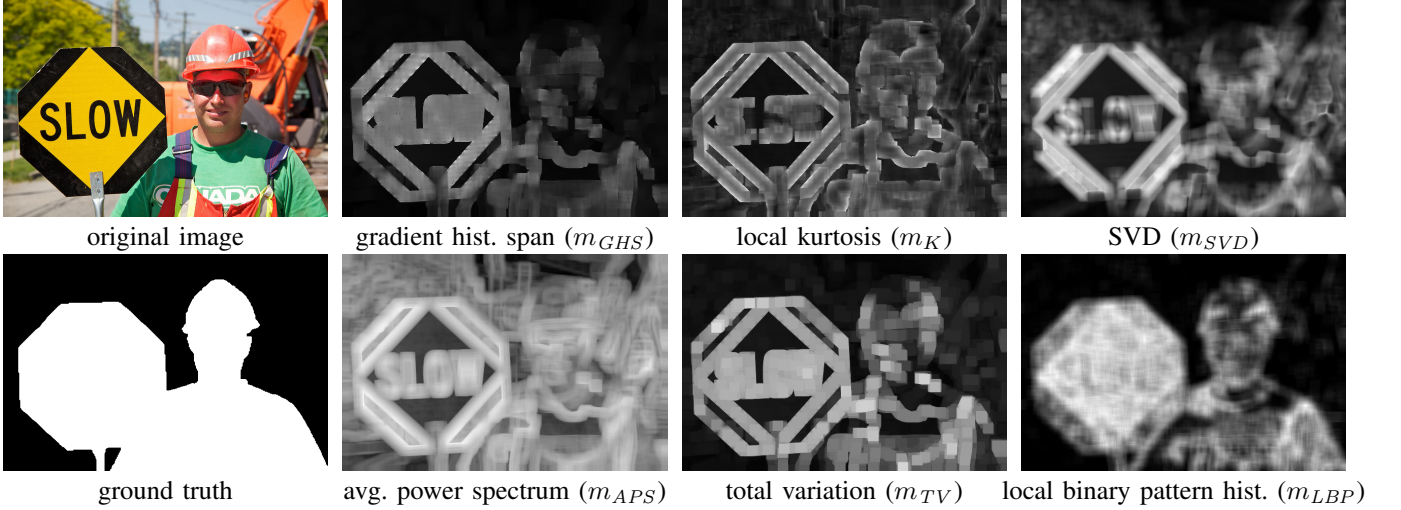
Fig. 11: Metric responses for a sample image for different sharpness metrics.

A. Multi-scale Sharpness Map Generation

In the first step, multi-scale sharpness maps are generated using m_{LBP}. The sharpness metric is computed for a local patch about each image pixel. Sharpness maps are constructed at three scales where scale refers to local patch size. By using an integral image [50], sharpness maps may be computed in constant time per pixel for a fixed P and R.

B. Alpha Matting Initialization

Alpha matting is the process of decomposing an image into foreground and background. The image formation model can be expressed as:

$$I(x,y) = \alpha_{x,y}F(x,y) + (1 - \alpha_{x,y})B(x,y) \qquad (16)$$

where the *alpha matte*, $\alpha_{x,y}$, is the opacity value on pixel position (x,y). It can be interpreted as the confidence that a pixel is in the foreground. Typically, alpha matting requires a user to interactively mark known foreground and background pixels, initializing those pixels with $\alpha = 1$ and $\alpha = 0$, repetively.

Interpreting "foreground" as "sharp" and background as "blurred", we initialized the alpha matting process automatically by applying a double threshold to the sharpness maps computed in the previous step to produce an initial value of α for each pixel:

$$mask^s(x,y) = \begin{cases} 1, & \text{if } m_{LBP}(x,y) > T_{m_1}. \\ 0, & \text{if } m_{LBP}(x,y) < T_{m_2}. \\ m_{LBP}(x,y), & \text{otherwise.} \end{cases}$$
$$(17)$$

where s indexes the scale, that is, $mask^s(x,y)$ is the initial α-map at the s-th scale.

C. Alpha Map Computation

The α-map was solved by minimizing the following cost function as proposed by Levin [25]:

$$E(\alpha) = \alpha^T \mathbf{L}\alpha + \lambda(\alpha - \hat{\alpha})^T(\alpha - \hat{\alpha}) \qquad (18)$$

where α is the vectorized α-map, $\hat{\alpha} = mask^i(x,y)$ is one of the vectorized initialization alpha maps from the previous step, and $\mathbf{L}$ is the matting Laplacian matrix. The first term is the regulation term that ensures smoothness, and the second term is the data fitting term that encourages similarity to $\hat{\alpha}$. For more details on Equation 18, readers are referred to [25].

The alpha matting was applied at each scale as shown in Figure 12. The final alpha map at each scale is denoted as $\alpha^s, s = 1, 2, 3$.

D. Multi-scale Inference

After determining the alpha map at three different scales, a multi-scale graphical model was adopted to make the final decision [43]. The total energy on the graphical model is expressed as:

$$E(h) = \sum_{s=1}^{3}\sum_{i} |h_i^s - \hat{h}_i^s| +$$
$$\beta\left(\sum_{s=1}^{3}\sum_{i}\sum_{j \in N_i^s} |h_i^s - h_j^s| + \sum_{s=1}^{2}\sum_{i} |h_i^s - h_i^{s+1}|\right)$$
$$(19)$$

where $\hat{h_i^s} = \alpha_i^s$ is the alpha map for scale s at pixel location i that was computed in the previous step, and h_i^s is the sharpness to be inferred. The first term on the right hand side is the unary term which is the cost of assigning sharpness value h_i^s to pixel i in scale s. The second is the pairwise term which enforces smoothness in the same scale and across different scales. The weight β regulates the relative importance of these two terms. Optimization of Equation 19 was performed using loopy belief propagation [30].

The output of the algorithm is h^3 which is the inferred sharpness map at the largest scale. This is a grayscale image, where higher intensity indicates greater sharpness.
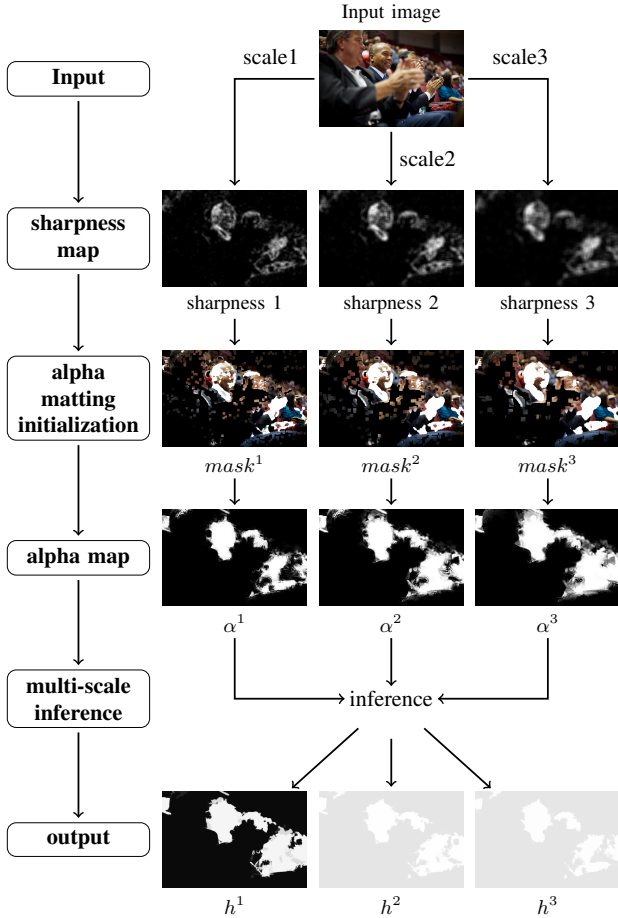
Fig. 12: Our blur segmentation algorithm. The main steps are shown on the left; the right shows each image generated and its role in the algorithm. The output of the algorithm is h^1.

V. BLUR SEGMENTATION ALGORITHM EVALUATION

Our blur segmentation algorithms were tested using a public blurred image dataset consisting of 704 partially blurred images and accompanying hand-segmented ground truth images [42] . Each image was segmented into sharp and blurred regions using the processes described in Section IV. Sharpness metric m_{LBP} was computed with $T_{LBP} = 0.016$. The sharpness map scales were square local regions of 11×11, 15×15, and 21×21 pixels. The thresholds used in the alpha matting step were $T_{m_1} = 0.3$ and $T_{m_2} = 0.01$. Weight $\beta = 0.5$ was used in the multi-scale inferencing step.

We compared our algorithm to six comparator methods briefly mentioned in Section II of which we now remind the reader. Su et al. simply calculated a sharpness map using m_{SVD} [44]; Vu combined both spectral and spatial sharpness (as called S_1 and S_2 in the original paper) using a geometric mean [51]. Shi14 used all of $m_{GHS}, m_K, m_{LDA}, m_{APS}$ together with a naïve Bayes classifier and multi-scale inference model [43]. Shi15 formed a sparse representation of image patches using a learned dictionary for the detection of slight perceivable blur [20]. Zhuo computed a depth map based on edge width [57]. Zhu estimated the space-variant PSF by statistical modelling of the localized frequency spectrum of
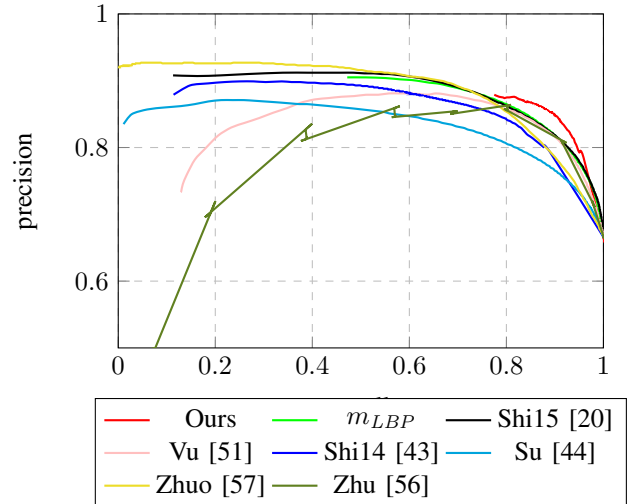

Fig. 13: Precision and recall curves for different methods on the blur dataset. The curves were obtained by thresholding the sharpness maps with threshold varying in the range of [0, 255]. Note that our method achieves the highest precision when recall is larger than 0.8. This comparison might be unfair for Zhu since their segmentation is based on graph cut rather than thresholding of the depth map. Therefore we compared their graph cut segmented binary map in the F-measure section.

the gradient field [56].

All the output of these methods are grayscale images where greater intensity indicates greater sharpness, and all (except for Zhu) use simple thresholding, T_{seg}, as a final step to produce a segmentation, as in our own algorithm. The parameters for the comparator algorithms were to the defaults as in their original code. Since we were unable to get the original code for Zhu's algorithm [56], which belongs to Adobe Systems Inc., the results shown here were produced by our own implementation of the algorithm as described in the published paper. The depth map was normalized by 8 (since the coherence labels are in the range of [0,8]) and inverted to get the sharpness map.

A. Precision and Recall

Precision and recall curves were generated for each algorithm by varying the threshold used to produce a segmentation of the final sharpness maps (i.e. similar to [43]).

$$precision = \frac{R \cap R_g}{R}, \quad recall = \frac{R \cap R_g}{R_g} \qquad (20)$$

where R is the set of pixels in the segmented blurred region and R_g is the set of pixels in the ground truth blurred region. Figure 13 shows the precision and recall curves for each method with the threshold T_{seg} sampled at every integer within the interval $[0, 255]$. Our algorithm achieves higher precision than the comparator algorithms when recall is above 0.8. Moreover, the proposed sharpness metric alone achieves results comparable to Shi15.

Figure 14 shows the sharpness maps (prior to final thresholding) for each algorithm for a few sample images. Our method is superior than the others under various background

Original images

Ground truth

Ours

m_{LBP}

Shi15 [20]

Vu [51]

Shi14 [43]

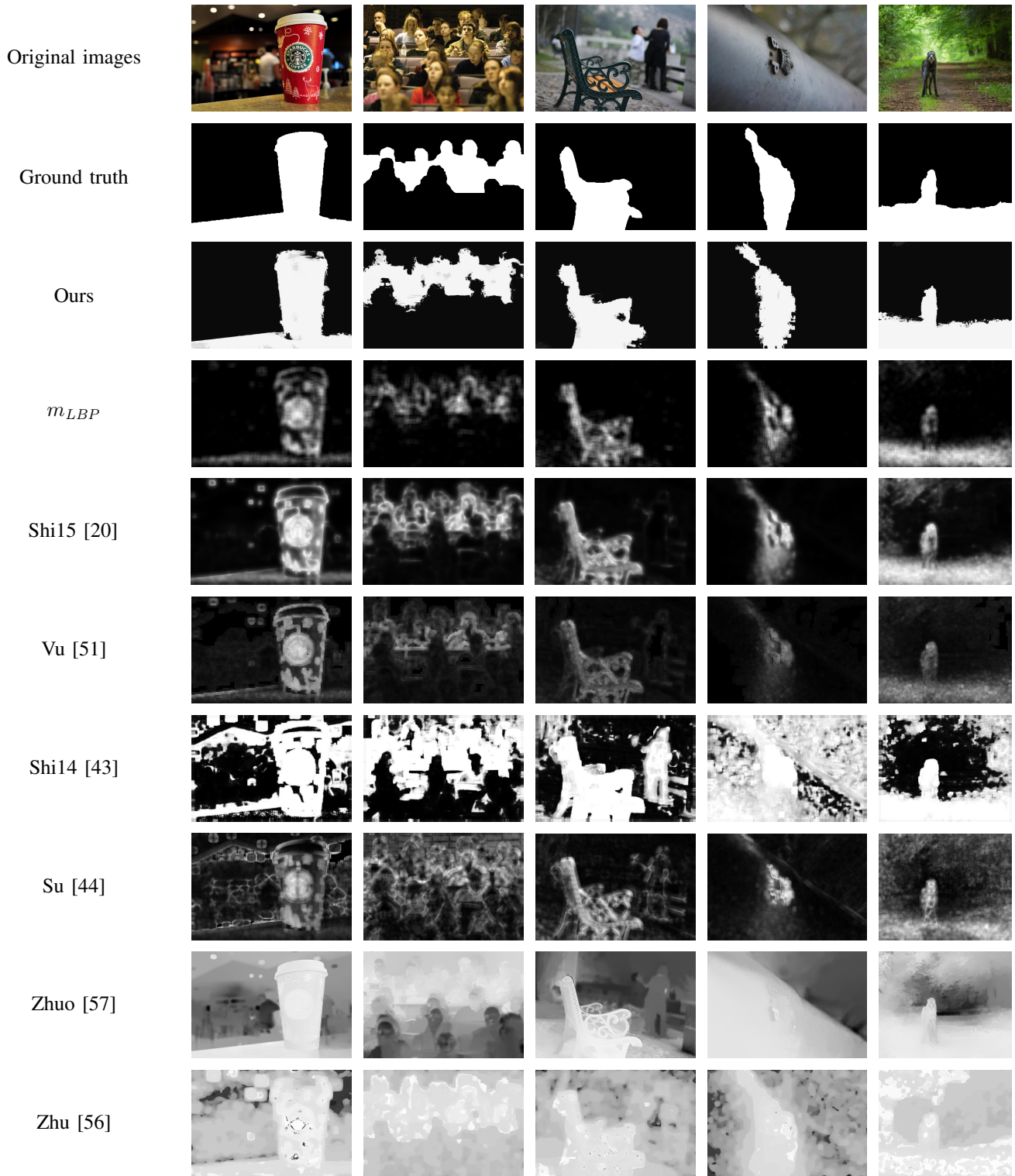Su [44]

Zhuo [57]

Zhu [56]

Fig. 14: Results achieved by different blur detection methods. Final sharpness maps, prior to thresholding for segmentation, are shown.
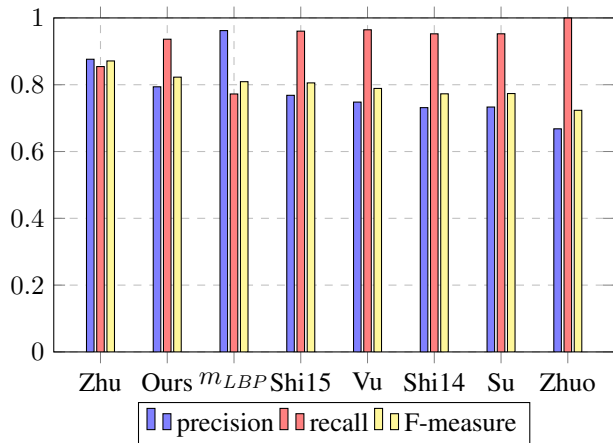
Fig. 15: Precision, Recall and F-measure for adaptive thresholds. The result of Zhu is achieved by using graph cut instead of simple thresholding as suggested in their paper. Note that by using a smaller threshold (i.e. $T_{seg} = 0.3$), our method can achieve comparative performance ($precision = 0.863$, $recall = 0.868$, F-measure $= 0.864$) with Zhu.

and blurs. We attribute errors mainly to the shortcomings of the sharpness metrics used by local based methods–Shi14, Vu, Su (Section II). Moreover, our detection maps contain mostly high- or low-confidence values which can be more correctly thresholded.

B. F-measure

In another experiment, we used an image dependent adaptive threshold proposed by [1] for the segmentation with the threshold defined as:

$$T_{seg} = \frac{2}{W \times H} \sum_{x=1}^{W} \sum_{y=1}^{H} I(x, y) \qquad (21)$$

where, W, H are the width and height of the final sharpness map I. Then, similar to [35], the weighted harmonic mean measure or F-measure of precision and recall was computed for comparison. The definition is as follows:

$$F_{\beta} = \frac{(1 + \beta^2) \times precision \times recall}{\beta^2 \times precision + recall} \qquad (22)$$

Here, β^2 was set to 0.3 as in [35], [1].

Note that, the segmentation map of Zhu was produced by graph cut instead of simple thresholding of the depth map. The parameters we used were the same as suggested in their paper which are $\lambda_0 = 1000$, $\sigma_\lambda = 0.04$, $\tau = 2$. Exemplar segmentation maps of images in Figure 14 is shown in Figure 16.

The reason our performance is worse than Zhu in Figure 15 is that the adaptive threshold is not the best threshold for our method. The best precision and recall we can achieve, as can be seen in Figure 13, is $precision = 0.863$, $recall = 0.868$ which is comparative with the one Zhu have achieved. However, even if only the adaptive-thresholded results were compared, our method ranked first among the comparators.

Blur segmentation	Avg. Runtime
Ours	27.75s
m_{LBP}	40ms
Shi15 [20]	38.36s
Vu [51]	19.18s
Shi14 [43]	705.27s
Su [44]	37s
Zhuo [57]	20.59s
Zhu [56]	12min

TABLE II: Run time comparison of different blur segmentation methods. The time for our method is based on a mex implementation of m_{LBP}.

A run time comparison of the complete segmentation algorithms is shown in Table II. The same setup was used for the measurement of runtime as in Table I. Given m_{LBP}'s performance in the precision and recall curve and F-measure, it has a significant speed advantage over the others. The time for our complete segmentation algorithm is mostly spent on the matting and multi-scale inference. It ranks the fourth among all these methods.

Finally, we give some examples of our algorithm applied to images other than those in our evaluation data set. Microscopy optics often have low depth of field and form an important class of images for blur detection. Figure 17 shows examples of our algorithm applied to such images. The first is a plant seed [39] whose roughly spherical shape results in a ring-shaped in-focus region. The other image is a microorganism [37] in fresh water. The threshold T_{LBP} for the sharpness metric was set to 0.012 and 0.04 respectively. Note how well our segmentation results conformed to the visual perception of the image sharpness. Additional results can be seen in the supplementary file.

C. Discussion

You may have noticed the jagged boundary of our segmentation map when there is a distinctive depth discontinuity between the foreground and background. This is because the sharpness is measured locally. It is inevitable to incorporate regions with various extents of sharpness by using a local window, especially around edges where the depth discontinuity occurs. Therefore, the sharp area is enlarged in the alpha matting initialization step (step B). Zhu solved this problem by taking smoothness and color edge information into consideration in the coherence labeling step but would also fail in cases where depth changes gradually.

There are certain situations that can cause our method to fail. Our method has difficulty differentiating an in-focus smooth region and a blurred smooth region since only a limited small size of local neighbour is considered, but this is a problem that will be inherently challenging for any algorithm. If noise level in the image is low, this problem can be overcome to some extent by reducing the T_{LBP} threshold. In addition, for object recognition purposes, this drawback would not weaken the feature representation too much since smooth regions contain little to no useful discriminating texture. An example of this type of failure case and proposed remedy can be seen in Figure 18(a).
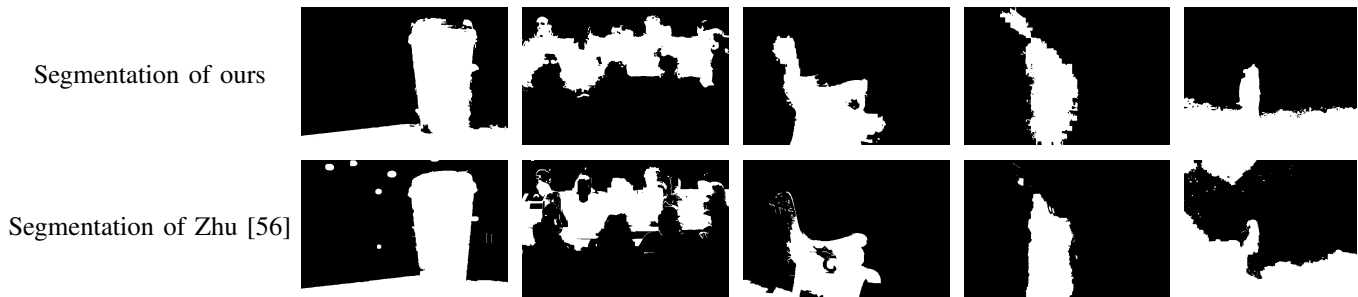
Fig. 16: Binary segmentation map comparison with Zhu et al.

Another failure case occurs due to image noise, but it can be mitigated by applying a noise reducing filter as mentioned in section III. An example of this type of failure and the proposed remedy is shown in Figure 18(b).

The selection of T_{LBP} is essential for obtaining a satisfactory segmentation. It controls how much sharp area would appear in the final segmentation result. For a image with little to no noise, T_{LBP} 0.016 should produce a reasonable result. Lowering the value would cause the inclusion of more low contrast sharp regions. For a image corrupted by noise, a noise reduction procedure should be employed.

The proposed metric was inspired by the statistical difference of local binary patterns of a set of partial blurred images. Since the source of blurriness is mainly defocus blur, our metric currently is only capable of detecting defocus blur. Given that there are other type of blurriness such as those introduced by low qualities of lens and materials in image systems and motion blur, it would be worth studying the blur model due to the properties of optical devices [24] and at the same time exploring properties of different patterns such as the non-uniform binary patterns and local ternary pattern (LTP) [45] on blur regions of different type. Moreover, the ideas used in noise-resistant LBP (NRLBP) [38], which treats pixels susceptible to noise as having uncertain state and then determines the corresponding bit value based on the other bits of the LBP code, might worth borrowing if explicit handling of noise in blur detection is desired.

VI. CONCLUSION

We have proposed a very simple yet effective sharpness metric for defocus blur segmentation. This metric is based on the distribution of uniform LBP patterns in blur and non-blur image regions. The direct use of the local raw sharpness measure can achieve comparative results to the stat-of-the-art defocus segmentation method that based on sparse representation, which shows the potential of local based sharpness measures. By integrating the metric into a multi-scale information propagation frame work, it can achieve comparative results with the state-of-the-art. We have shown that the algorithm's performance is maintained when using an automatically and adaptively selected threshold T_{seg}. Our sharpness metric measures the number of certain LBP patterns in the local neighbourhood thus can be efficiently implemented by integral images. If combined with real-time matting algorithms, such as GPU implementations of global matting [18],
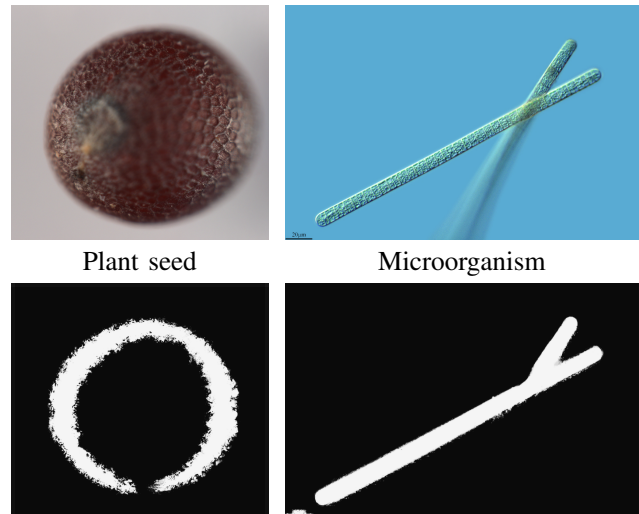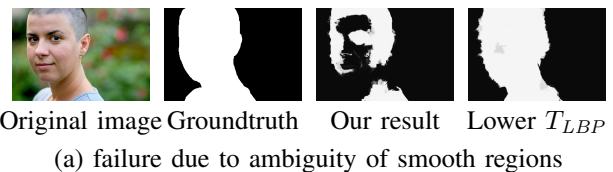

Plant seed   Microorganism

Fig. 17: Our algorithm applied to microscopy images. Top row: original images; bottom row: final sharpness maps.

Original image   Groundtruth   Our result   Lower T_{LBP}

(a) failure due to ambiguity of smooth regions

Original image   Groundtruth   Our result   Noise removal

(b) failure due to noise

Fig. 18: Blur segmentation algorithm failure cases and mitigation.

our method would have significant speed advantage over the other defocus segmentation algorithms.

also like to thank the authors of [57], [43], [51] for releasing their code.

REFERENCES

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on*, pages 1597–1604. IEEE, 2009.

[2] H.-M. Adorf. Towards hst restoration with a space-variant psf, cosmic rays and other missing data. In *The Restoration of HST Images and Spectra-II*, volume 1, page 72, 1994.

[3] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12):2037–2041, 2006.

[4] S. Bae and F. Durand. Defocus magnification. In *Computer Graphics Forum*, volume 26, pages 571–579. Wiley Online Library, 2007.

[5] K. Bahrami, A. C. Kot, and J. Fan. A novel approach for partial blur detection and segmentation. In *Multimedia and Expo (ICME), 2013 IEEE International Conference on*, pages 1–6. IEEE, 2013.

[6] J. Bardsley, S. Jefferies, J. Nagy, and R. Plemmons. A computational method for the restoration of images with an unknown, spatially-varying blur. *Optics express*, 14(5):1767–1782, 2006.

[7] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 60–65. IEEE, 2005.

[8] G. Burton and I. R. Moorhead. Color and spatial structure in natural scenes. *Applied Optics*, 26(1):157–170, 1987.

[9] A. Chakrabarti, T. Zickler, and W. T. Freeman. Analyzing spatially-varying blur. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2512–2519. IEEE, 2010.

[10] T. S. Cho. *Motion blur removal from photographs*. PhD thesis, Citeseer, 2010.

[11] F. Couzinie-Devy, J. Sun, K. Alahari, and J. Ponce. Learning to estimate and remove non-uniform image blur. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1075–1082. IEEE, 2013.

[12] S. Dai and Y. Wu. Removing partial blur in a single image. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2544–2551. IEEE, 2009.

[13] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 787–794. ACM, 2006.

[14] R. Ferzli and L. J. Karam. A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb). *Image Processing, IEEE Transactions on*, 18(4):717–728, 2009.

[15] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *JOSA A*, 4(12):2379–2394, 1987.

[16] D. Fish, J. Grochmalicki, and E. Pike. Scanning singular-value-decomposition method for restoration of images with space-variant blur. *JOSA A*, 13(3):464–469, 1996.

[17] R. Hassen, Z. Wang, M. M. Salama, et al. Image sharpness assessment based on local phase coherence. *Image Processing, IEEE Transactions on*, 22(7):2798–2810, 2013.

[18] K. He, C. Rhemann, C. Rother, X. Tang, and J. Sun. A global sampling method for alpha matting. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2049–2056. IEEE, 2011.

[19] M. Heikkila and M. Pietikainen. A texture-based method for modeling the background and detecting moving objects. *IEEE transactions on pattern analysis and machine intelligence*, 28(4):657–662, 2006.

[20] J. J. Jianping Shi, Li Xu. Just noticeable defocus blur detection and estimation. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, 2015.

[21] C. Kim. Segmenting a low-depth-of-field image using morphological filters and region merging. *Image Processing, IEEE Transactions on*, 14(10):1503–1511, 2005.

[22] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *Advances in Neural Information Processing Systems*, pages 1033–1041, 2009.

[23] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 233–240. IEEE, 2011.

[24] D. Kundur and D. Hatzinakos. Blind image deconvolution. *Signal Processing Magazine, IEEE*, 13(3):43–64, 1996.

[25] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2):228–242, 2008.

[26] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1964–1971. IEEE, 2009.

[27] 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.

[28] R. Liu, Z. Li, and J. Jia. Image partial blur detection and classification. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[29] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. A no-reference perceptual blur metric. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 3, pages III–57. IEEE, 2002.

[30] K. P. Murphy, Y. Weiss, and M. I. Jordan. Loopy belief propagation for approximate inference: An empirical study. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, pages 467–475. Morgan Kaufmann Publishers Inc., 1999.

[31] N. D. Narvekar and L. J. Karam. A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). *Image Processing, IEEE Transactions on*, 20(9):2678–2683, 2011.

[32] T. Ojala and M. Pietikäinen. Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, 32(3):477–486, 1999.

[33] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

[34] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, 2002.

[35] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung. Saliency filters: Contrast based filtering for salient region detection. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 733–740. IEEE, 2012.

[36] M. Pietikäinen, T. Nurmela, T. Mäenpää, and M. Turtinen. View-based recognition of real-world textures. *Pattern Recognition*, 37(2):313–323, 2004.

[37] C. C. P. A. PROJECT. Microagua. https://www.flickr.com/photos/microagua/.

[38] J. Ren, X. Jiang, and J. Yuan. Noise-resistant local binary pattern with an embedded error-correction mechanism. *Image Processing, IEEE Transactions on*, 22(10):4049–4060, 2013.

[39] J. J. Ruojing Wang, Jennifer Neudorf. Canadian food inspection agency seed dataset, 2015.

[40] O. Russakovsky, Y. Lin, K. Yu, and L. Fei-Fei. Object-centric spatial pooling for image classification. In *Computer Vision–ECCV 2012*, pages 1–15. Springer, 2012.

[41] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. In *ACM Transactions on Graphics (TOG)*, volume 27, page 73. ACM, 2008.

[42] J. Shi, L. Xu, and J. Jia. Blur detection dataset. http://www.cse.cuhk.edu.hk/~leojia/projects/dblurdetect/dataset.html, 2014.

[43] J. Shi, L. Xu, and J. Jia. Discriminative blur detection features. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 2965–2972. IEEE, 2014.

[44] B. Su, S. Lu, and C. L. Tan. Blurred image region detection and classification. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 1397–1400. ACM, 2011.

[45] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *Image Processing, IEEE Transactions on*, 19(6):1635–1650, 2010.

[46] T. Thongkamwitoon, H. Muammar, and P.-L. Dragotti. An image recapture detection algorithm based on learning dictionaries of edge profiles. *Information Forensics and Security, IEEE Transactions on*, 10(5):953–968, 2015.

[47] H. J. Trussell and S. Fogel. Identification and restoration of spatially variant motion blurs in sequential images. *Image Processing, IEEE Transactions on*, 1(1):123–126, 1992.

[48] H. J. Trussell and B. R. Hunt. Image restoration of space variant blurs by sectioned methods. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'78.*, volume 3, pages 196–198. IEEE, 1978.

[49] v. A. van der Schaaf and J. v. van Hateren. Modelling the power spectra of natural images: statistics and information. *Vision research*, 36(17):2759–2770, 1996.

[50] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.

[51] C. T. Vu, T. D. Phan, and D. M. Chandler. : A spectral and spatial measure of local perceived sharpness in natural images. *Image Processing, IEEE Transactions on*, 21(3):934–945, 2012.

[52] J. Z. Wang, J. Li, R. M. Gray, and G. Wiederhold. Unsupervised multiresolution segmentation for images with low depth of field. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(1):85–90, 2001.

[53] C. S. Won, K. Pyun, and R. M. Gray. Automatic object segmentation in images with low depth of field. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 3, pages 805–808. IEEE, 2002.

[54] J. Zhao, H. Feng, Z. Xu, Q. Li, and X. Tao. Automatic blur region segmentation approach using image matting. *Signal, Image and Video Processing*, 7(6):1173–1181, 2013.

[55] X. Zhu. *Measuring spatially varying blur and its application in digital image restoration*. PhD thesis, UNIVERSITY OF CALIFORNIA, SANTA CRUZ, 2013.

[56] X. Zhu, S. Cohen, S. Schiller, and P. Milanfar. Estimating spatially varying defocus blur from a single image. *Image Processing, IEEE Transactions on*, 22(12):4879–4891, 2013.

[57] S. Zhuo and T. Sim. Defocus map estimation from a single image. *Pattern Recognition*, 44(9):1852–1858, 2011.

PLACE
PHOTO
HERE

Xin YI Biography text here.

Mark Eramian Biography text here.