Harmonic Variance: A Novel Measure for In-focus Segmentation

Li, F.; Porikli, F.

Abstract

We introduce an efficient measure for estimating the degree of in-focus within an image region. This measure, harmonic mean of variances, is computed from the statistical properties of the image in its bandpass filtered versions. We incorporate the harmonic variance into a graph Laplacian spectrum based segmentation framework in order to accurately align the in-focus measure responses on the underlying image structure. Our results demonstrate the effectiveness of this novel measure for determining and segmenting in-focus regions in low depth-of-field images.

British Machine Vision Conference (BMVC)

Harmonic Variance: A Novel Measure for In-focus Segmentation

Feng Li
http://www.eecis.udel.edu/~feli/

Fatih Porikli
http://www.porikli.com/

Mitsubishi Electric Research Labs
Cambridge, MA 02139, USA

Abstract

We introduce an efficient measure for estimating the degree of in-focus within an image region. This measure, harmonic mean of variances, is computed from the statistical properties of the image in its bandpass filtered versions. We incorporate the harmonic variance into a graph Laplacian spectrum based segmentation framework in order to accurately align the in-focus measure responses on the underlying image structure. Our results demonstrate the effectiveness of this novel measure for determining and segmenting in-focus regions in low depth-of-field images.

1 Introduction

Detecting in-focus regions in a low depth-of-field (DoF) image has important applications in scene understanding, object-based coding, image quality assessment and depth estimation because such regions may indicate semantically meaningful objects.

Many approaches consider image sharpness as an in-focus indicator and attempt to detect such regions directly from the high-frequency texture components, e.g. edges [10]. However, in-focus edges with low intensity magnitudes and defocus edges with high intensity magnitudes may give similar responses. As a remedy, some early work [27, 29] normalize the edge strength with the edge width assuming the in-focus edges have steeper intensity gradients and smaller widths than the defocus edges.

In addition to edges, various statistical and geometric features, such as local variance [5, 1], variance of wavelet coefficients [30], higher order central moments [8], local contrast [23], multi-scale contrast [17], difference of histogram of oriented gradients [13], local power spectrum slope [16], gradient histogram span, maximum saturation, and local autocorrelation congruency are exploited to measure the degree of in-focus. These features often generate sparse in-focus boundaries, which are then propagated to the rest of the image. For instance, [4] determines a center line of edges and blur responses in a sparse focus map by using the first and second order derivatives from a set of steerable Gaussian kernel filters. To maintain smoothness, a Markov Random Field (MRF) and a cross-bilateral filter based interpolation scheme are described in [0]. Similarly, [28] imposes the ratio of maximum gradient to the local image contrast as a prior to the MRF label estimation for refinement while [20] applies an inhomogeneous inversion heat diffusion for the same purpose.

Alternatively, a number of methods analyze the difference between the filtered versions of the original image to estimate the degree of in-focus. For this purpose, [32] uses a Gaussian blur kernel to compute the ratio between the gradients of the original and blurred images. In a similar manner, [15] determines the blur extent by matching the histograms of the synthesized and defocus regions. [9] identifies in-focus regions by iterating over a localized blind deconvolution with frequency and spatial domain constraints starting with a given initial blur estimate. By sequentially updating the point spread function parameters, [14] applies bilateral and morphological filters to defocus map generated from a reblurring model, then uses an error control matting to refine the boundaries of in-focus regions.

In-focus detection can also be viewed as a learning task. [23] proposes a supervised learning approach by comparing the in-focus patches and their defocus versions on a training set to learn a discriminatively-trained Gaussian MRF that incorporates multi-scale features. [17] extracts multi-scale contrast, center surround histogram, and color spatial distribution for conditional random field learning. [16] proposes several features to classify the type of the blur.

There also exist approaches that use coded aperture imaging [11], active refocus [19] and multiple DoF images [13] for in-focus detection.

For accuracy, conventional approaches rely heavily on accurate detection of meaningful boundaries. However, object boundary detection has its own inherent difficulties. Besides, many methods require high blur confidence, i.e. considerably blurred backgrounds for a reliable segmentation of in-focus regions. In case the background features are not severely blurred, the existing statistic measures still give strong response. Dependence on blind morphology, initial segmentation, noise model heuristics, and training data bias are also among other shortcomings.

We note that natural images are not random patterns and exhibit consistent statistical properties. A property that has been used to evaluate the power spectrum of images is the rotationally averaged spatial power spectrum $\psi(\omega)$, which can be approximated as

$$\psi(\omega) = \sum_{\theta} \psi(\omega, \theta) \simeq \frac{k}{\omega^{\gamma}} \tag{1}$$

where γ is the slope of the power spectrum and k is a scaling factor. Eq.(1) demonstrates a linear dependency of $\log \psi$ on $\log \omega$. The value of γ varies from image to image, but lies within a fairly narrow range (around 2) for natural images, which could be used for blur identification as a defocus metric [16, 24].

2   Harmonic Variance

Here, we introduce an efficient in-focus measure to estimate the degree of in-focus within an image region. A harmonic mean of variances, which we call as *harmonic variance*, is computed from the statistical properties of the given image in its bandpass filtered versions.

To find the bandpass responses, we apply the 2D discrete cosine transform (DCT). For a given $N \times N$ image I, we construct the DCT response F as

$$F(u,v) = \alpha_x \alpha_y \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \cos\left(\frac{\pi u}{2N}(2x+1)\right) \cos\left(\frac{\pi v}{2N}(2y+1)\right) I(x,y) \quad ,$$

where the normalization factors α_x, α_y are

$$\alpha_x = \begin{cases} 1/N & u = 0 \\ \sqrt{2}/N & \text{else} \end{cases} \quad , \quad \alpha_y = \begin{cases} 1/N & v = 0 \\ \sqrt{2}/N & \text{else} \end{cases}$$

for $u, v = 1, \ldots, N$. A bandpass filter g_m retains frequencies in a desired range and attenuates frequencies below and above in the transform domain, i.e., $g_m(F) = F(u, v)$ if u, v is in the bandpass range $uv \in \mathcal{B}_m$ and $g_m(F) = 0$ otherwise. We obtain set of bandpass filtered images I_m $m = 1, .., M$ by applying the inverse DCT to g_m

$$I_m = F^{-1}\{g_m(F)\}. \tag{2}$$

Setting $F = \sum_{m=1}^{M} g_m(F)$, the summation of the bandpass filtered images becomes equal to the given image $I = \sum_{m=1}^{M} I_m$. Obviously, the number of the bands can at most be the number of DCT coefficients $M \ll N^2$. We apply the DCT bandpass filtering within separate 8×8 image patches with $M = 64$. In other words, the DCT transform of a patch around each pixel is individually bandpass filtered then the responses for all pixels are aggregated to find I_m.

Let us recall that the process of image blurring at a pixel $\mathbf{x} = (x, y)$ can be formulated as an image convolution, that is,

$$I_B(\mathbf{x}) = I(\mathbf{x}) \otimes K(\mathbf{x}) , \tag{3}$$

where I_B is the blurred version of the image I and K is the spatially varying blur kernel. We consider defocus blurring[1] and use a 2D isotropic Gaussian function with stand deviation σ to represent the blur kernel K at pixel position $\mathbf{x}$. From Fourier analysis, we know that this convolution is a low-pass filtering in the frequency domain with the cutoff frequency σ^{-1} for the patch centered at $\mathbf{x}$.

For defocus blurred patches, there at least exists a limited number of DCT filtered channels I_m below the cutoff frequency with large variances. Because the variance is the squared deviation, a few large variance values would greatly increase the estimated variance of the patch.

Note that, even if most of the variances σ_m^2 of a defocus region may have small values, one large variance estimate in a low-frequency channel, which is quite common for textured yet blurry regions, could still make the arithmetic mean of the variances larger than that of in-focus regions, causing the defocus region to be falsely considered as in-focus.

Our intuition is that, a competent in-focus measure should not be sensitive to such outliers of the standard deviation in the DCT filtered channels I_m. Thus, we use the harmonic mean instead of other possible first order statistics to combine the variances σ_m^2. For a given patch centered at $\mathbf{x}$, we define the in-focus measure as the harmonic mean of the variances $\sigma_m^2(\mathbf{x})$ omitting the variance of the zero-frequency channel σ_0^2

$$h(\mathbf{x}) = \left[\frac{1}{M-1} \sum_{m=1}^{M-1} \frac{1}{\sigma_m^2(\mathbf{x})} \right]^{-1}. \tag{4}$$

For positive data sets containing at least one pair of unequal values, the harmonic mean is always the smallest of the three means, while the arithmetic mean is the greatest, and the geometric mean is in between.

Harmonic mean has two advantages. First, an arithmetic mean estimate can be distorted significantly by the large variances σ_m^2 of the blurred patches, while the harmonic mean is robust. Second, the harmonic mean considers reciprocals, hence it favors the small σ_m^2 and increases their influence in the overall estimation. For the blurred patches, a large number of small variances would keep h at a low level and the effect of a few of large σ_m^2 is mitigated

---

[1]Motion blur kernels are also low-pass filters with multiple zero crossings.

by the use of harmonic mean. On the other hand, for the in-focus patches, h still gives larger estimates since most DCT channels have large variances.

Instead of the harmonic mean of variances, the median absolute deviation (MAD) can be substituted. Nevertheless, MAD disregards the contributions of the variances with small values by pulling the estimate towards the higher values especially when multiple large outliers exist.

2.1 Relation to Noise in Natural Images

It is intuitive to draw connections between the proposed in-focus measure and noise analysis of natural images reflecting the super-Gaussian marginal statistics in the bandpass filtered domains.

We assume the local image noise $\eta(\mathbf{x})$ is characterized as a Gaussian random variable with zero mean and $\sigma_\eta^2(\mathbf{x})$ variance[2]. The noise contaminated image is then $I_\eta(\mathbf{x}) = I(\mathbf{x}) + \eta(\mathbf{x})$. Based on the observation that noise tends to have more energy in the high frequency bands than natural images, estimation of the global noise variance in an image is usually based on the difference image between the noisy image and its low-pass filtered response.

The traditional methodology is to use the relationship between noise variance and other higher-order statistics of natural images to obviate the low-pass filtering step. In [26], a noise variance estimation function is learned by training samples and the Laplacian model for the marginal statistics of natural images in bandpass filtered domains.

Another family of methods uses the relationship between noise variance and image kurtosis in bandpass filtered domains. Kurtosis is a descriptor of the shape of a probability distribution and defined as $\kappa = \mu_4/(\sigma^2)^2 - 3$. It can also be viewed as a measure of nonnormality. Due to the independence of noise η to I, and the additive property of cumulants, for each bandpass filtered image I_m we have

$$\kappa_m(\mathbf{x}) = \tilde{\kappa}_m(\mathbf{x}) \left(\frac{\sigma_m^2(\mathbf{x}) - \sigma_\eta^2(\mathbf{x})}{\sigma_m^2(\mathbf{x})} \right)^2, \tag{5}$$

where $\tilde{\kappa}_m(\mathbf{x})$ and $\kappa_m(\mathbf{x})$ represent the kurtosis value of the latent and the noisy I_m at $\mathbf{x}$ for channel m.

For natural images, kurtosis has the concentration property in the bandpass filtered domains, thus, the kurtosis of an image across the bandpass filtered channels can be approximated by a constant $\kappa(\mathbf{x}) \approx \tilde{\kappa}_m(\mathbf{x})$ [21, 53]. In other words, we can estimate the optimal local noise variance $\sigma_\eta^2(\mathbf{x})$ and the kurtosis κ by designing a simple least squares minimization problem as

$$\min_{\sigma_\eta^2, \kappa} \sum_m \left[\sqrt{\kappa_m(\mathbf{x})} - \sqrt{\kappa(\mathbf{x})} \left(\frac{\sigma_m^2(\mathbf{x}) - \sigma_\eta^2(\mathbf{x})}{\sigma_m^2(\mathbf{x})} \right) \right]^2, \tag{6}$$

According to [21], a closed form solution exists

$$\sigma_\eta^2(\mathbf{x}) = \alpha(\mathbf{x}) \cdot h(\mathbf{x}), \tag{7}$$

where

$$\alpha(\mathbf{x}) = \frac{\sqrt{\kappa(\mathbf{x})} - \frac{1}{M-1} \sum_m \sqrt{\kappa_m(\mathbf{x})}}{\sqrt{\kappa(\mathbf{x})}}. \tag{8}$$

---

[2]We omit shot noise and other types of noise here. Shot noise follows a Poisson distribution, which can be approximated by a Gaussian distribution for adequately illuminated images.
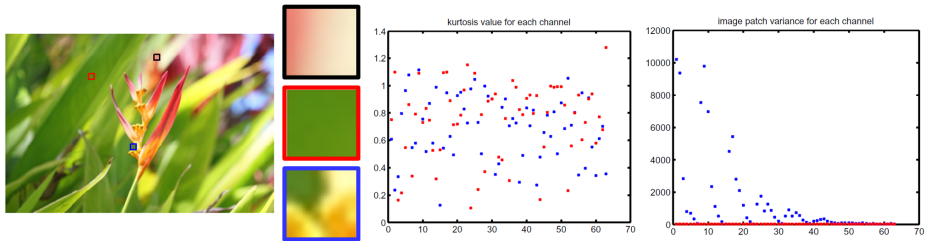
Figure 1: Three regions, defocused (black, red) and focused (blue) from a low DoF image. The corresponding kurtosis and variance distributions of blue and red regions in their power spectrum using 2D DCT (vectorized, horizontal axis from low-to-high frequencies).

Eqs.(7) and (8) indicate that the noise variance σ_η^2 is related to the in-focus measure h by the factor α.

As we show next, the noise variance estimation algorithms in [21, 53] are limited and do not apply to low DoF images.

Let us take a closer look at α in Eq.(8). Suppose we are given two different patches from a single image, one extracted from the in-focus region at $\mathbf{x}_1$ and the other from the defocused region at $\mathbf{x}_2$. Since they are extracted from the same image, their local noise variance should be the same, i.e., $\sigma_\eta^2(\mathbf{x}_1) = \sigma_\eta^2(\mathbf{x}_2)^3$. Moreover, α should be close to 0 for both in-focus and defocused patches since the arithmetic mean of κ_m is a close estimate of κ regardless of patch locations, that is, $\alpha(\mathbf{x}_1) \approx \alpha(\mathbf{x}_2) \approx 0$. Since $\mathbf{x}_1$ is from the in-focus region, it is very straightforward that $h(\mathbf{x}_1) \gg h(\mathbf{x}_2)$. Therefore, by Eq.(7), we have $\sigma_\eta^2(\mathbf{x}_1) \gg \sigma_\eta^2(\mathbf{x}_2)$, which is inconsistent with our previous assumption. This explains that why the kurtosis estimation in [21, 53] fails for low DoF images and their quadratic formulation shifts high frequency of the latent image into noise variance estimations.

This discrepancy can be seen in Fig. 1. As shown, we can easily compute the α and h for both in-focus region $\mathbf{x}_1$ (blue rectangle) and defocused region $\mathbf{x}_2$ (red rectangle), and we have $\alpha(\mathbf{x}_1) = 0.075$, $h(\mathbf{x}_1) = 97.12$, and $\alpha(\mathbf{x}_2) = 0.016$, $h(\mathbf{x}_2) = 0.17$, which is consistent with our discussion before. And by Eq.(7) we have $\sigma_\eta^2(\mathbf{x}_1) = 7.316$ and $\sigma_\eta^2(\mathbf{x}_2) = 0.003$, contrary to the assumption of the kurtosis-based noise estimation algorithm. To see our harmonic variance as a powerful in-focus metric, we mark another defocused region (black) that contains slightly more high frequency information than $\mathbf{x}_2$. Then we compute the ratio between the in-focus patch $\mathbf{x}_1$ and this region for our harmonic variance: $97.12/0.121 = 802$, and the arithmetic mean of the DCT channel variances: $1297.4/84.375 = 15$, which intuitively shows that our harmonic variance can competently differentiate the defocus background from the in-focus foreground (For clarity, we also show the kurtosis and the variance distributions for the blue and red regions in Fig.1).

3 In-Focus Region Segmentation

Contrary to the conventional seeded segmentation algorithms, e.g. graph-cuts [2, 3] and random walks [6], we use the estimated harmonic variance map $h(\mathbf{x})$ to automatically partition the in-focus regions without any user interaction. For more details please refer to [12].

---

[3]The assumption is that the input low DoF image is captured under good lighting condition, thus the image is mainly of amplifier noise with Gaussian distribution.

Algorithm 1: Segmentation Using Laplacian Spectrum

Input: Laplacian matrix L of I, harmonic mean h, penalty term β
Output: segmentation f
$W = I, t = 1$;
while $t < iter_{max}$ & $\|W(f-h)\|^2 < e$ **do**
$\quad$ $f^t = (\beta L^\top L + W)^{-1} W h$;
$\quad$ update W according to Eq.(11) ;
$\quad$ $t = t + 1$
return the optimal f

To incorporate the harmonic variance values, we compute a graph Laplacian matrix L from the given image I and use it as a constraint. The graph Laplacian spectrum constraint $Lf = 0$ enforces a given image structure on the prior information (in the data fidelity term $\|f-h\|^2$). With this constraint, the optimal f should lie in the null-space of L, that is, f should be constant within each connected component of the corresponding graph G. In most cases, the binary segmentation results consist of several disconnected components and segments. Since the estimated f can be represented by a linear combination of the 0 eigenvectors (the 'ideal' basis) of L we are still able to differentiate the foreground components from the background. In this way, we avoid computing L's nullity k and its basis, while still can use the image structure to regularize the data fidelity term.

The Laplacian matrix L can be used to regularize an optimization formulation by laying it on the structure inherent in I. This enables us to define the in-focus segmentation as a least-squares constrained optimization problem as

$$\min_f \|f-h\|^2, \quad \text{s.t. } Lf = 0. \tag{9}$$

From comparative analysis of the harmonic variance maps h and the optimal segmentation f, it is clear that the residual $\delta(\mathbf{x}) = |f-h|$ has many spatially continuous outliers. The least squares fidelity team, however, applies quadratic cost function with equal weighting that severely distorts the final estimate in case of outliers. One robust option is to weight large outliers less and use the structure information from the Laplacian spectrum constraint to recover the segmentation f. Therefore, we borrow existing principles from robust statistics [7] and adapt a robust functional to replace the least squares term as

$$\min_f \rho(f-h) + \beta \|Lf\|^2 , \tag{10}$$

where ρ is the robust function. We use the Huber function

$$\rho(\mathbf{x}) = \begin{cases} 1 & \text{if } \delta(\mathbf{x}) < \varepsilon \\ \varepsilon/\delta(\mathbf{x}) & \text{if } \delta(\mathbf{x}) \geq \varepsilon \end{cases} \tag{11}$$

for the reason that it is a parabola in the vicinity of 0 and increases linearly when δ is large. Thus, the effects of large outliers can be eliminated significantly. When written in matrix form, we use a diagonal weighting matrix W to represent the Huber weight function. Therefore, the data fidelity term can be simplified as $\|W(f-h)\|^2$. As a result, the problem Eq.(10) can be solved efficiently in an iterative least square approach. At each iteration, the optimal f is updated as

$$f = (\beta L^\top L + W)^{-1} W h . \tag{12}$$

Our algorithm is shown in Alg. 1.

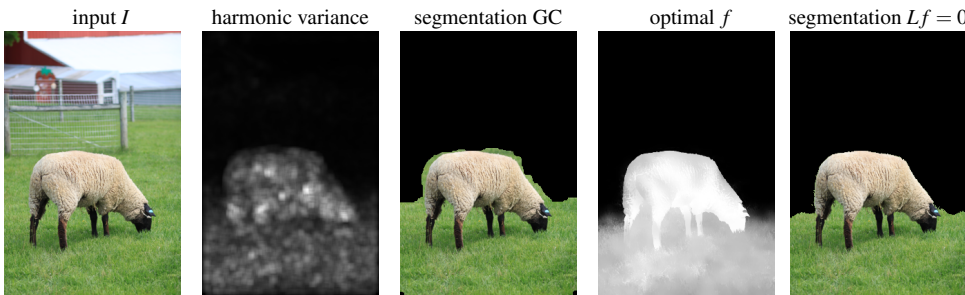input I harmonic variance   segmentation GC   optimal f segmentation $Lf = 0$

Figure 2: In-focus segmentation results using the proposed harmonic variance measure h.

4 Results

To test the performance of the proposed in-focus measure, we conducted experiments on benchmark images. Fig. 2 shows a typical low DoF scene captured with a large aperture lens. The input image I has gradually changing defocus blurs and the background (fences, farm houses, and etc.) contains strong edges. Our estimated harmonic variance measure accurately identifies the in-focus regions (sheep and near grass), and generate nearly consistent values for the foreground. When using our harmonic variance to initialize the data cost term of the graph-cut algorithm, it can well model the probability for each pixel belonged to the foreground. However, since our harmonic variance is a patch-based approach, it inevitably gives strong response across the in-focus boundaries, which can cause some artifacts for the graph-cuts algorithm, as shown in the center of Fig. 2. While the Laplacian based segmentation algorithm can robustly remove the effects of these outliers near the in-focus boundaries and generate an accurate segmentation for this challenging input.

We also compare our method with [32] as shown in Fig. 3. We use a traditional graph Laplacian based segmentation method using ℓ_2 norm for a fair comparison instead of the described robust approach in Eq.10. Our measure does not require an edge map (or weight matrix) yet generates a more accurate defocus map.

Next, we compare our harmonic variance measure with the high-order statistics (HOS) [8], the local power spectrum slope [16, 24], and a state-of-the-art visual saliency estimation algorithm [5] for low depth-of-field image feature extraction and in-focus object segmentation. We test the performance of these feature extraction algorithms using the same segmentation algorithm described in the previous section. In Fig. 4, we show five different natural/man-made scenes with varying defocus conditions and the corresponding defocus maps and segmentation results for HOS, saliency, slope and the harmonic variance measure.

In our analysis, we compute the local power spectrum slope $\gamma(\mathbf{x})$ for a patch size 17×17 around each pixel. The local power spectrum slope basically can detect all the in-focus regions, however, it also gives strong responses to complex background, which leads the segmentation algorithm to mislabel some parts of the background as the foreground.

Instead of computing the image statistics in the frequency domain, HOS works on the image spatial domain:

$$\text{HOS}(\mathbf{x}) = \min(C, \mu_4(\mathbf{x})/D) ,$$

where $\mu_4(\mathbf{x})$ is the 4-th order central moment at pixel $\mathbf{x}$ computed with local support, and C and D are empirical parameters for thresholding and scaling. The major drawback of these statistic measurements is that they still give strong response if the background features are not severely defocus blurred, which misleads the segmentation algorithm to label portions
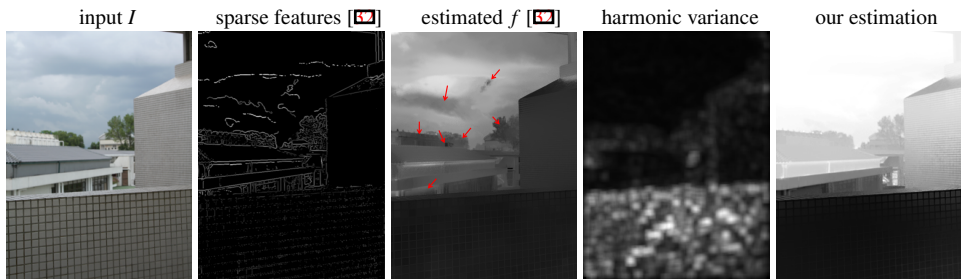
| input I | sparse features [ ] | estimated f [ ] | harmonic variance | our estimation |

Figure 3: Comparison of the estimated focus maps of a layered depth scene by [ ] and our algorithm. As visible, [ ] generates several incorrect depth regions such as the middle of clouds (depth inconsistency), tree (appears closer than the buildings), rooftop on the left (depth inconsistency), and etc. Harmonic variance measure and the proposed segmentation method accurately determine the defocus map for the input image.

of background as foreground. HOS performs well when the background of the input images is severely defocus blurred. As shown in the first and the fourth column, it can capture the majority of the in-focus objects, but still have some strong artifacts, such as disconnected in-focus regions. When the background still has strong high-frequency signals, it can completely fail because of simple threshold and scaling, as indicated by other scenes.

The visual saliency algorithm uses regional covariances to integrate several different low-level features for salient region detection. As demonstrated in Fig. 4, it is basically good for salient object indication, and can detect the centers of in-focus objects. However, the estimated saliency maps lacks the power to model the probability for segmentation, and the segmentation results can only capture parts of the in-focus regions.

As shown in Fig. 4, the harmonic variance measure can discriminate the in-focus and defocus regions efficiently. Its optimized version (second row) nicely aligns with the underlying image structure indicating the proposed measure is a reliable prior for the graph Laplacian based segmentation.

Table 1 presents the F-measure comparison scores of segmentation results for slope, saliency, HOS and the harmonic variance. All these are applied to the graph Laplacian based robust segmentation framework. F-measure is defined as $2\frac{Pr \cdot Re}{Pr+Re}$ where Pr is the precision $TP/(TP+FP)$ and Re is the recall $TP/(TP+FN)$ ratios. TP, FP, FN are true positives, false negatives, and false negatives. As shown, the segmentation results using the harmonic variance measure consistently have higher scores.

	slope γ [ ]	saliency [ ]	HOS [ ]	Harmonic h
paper	0.88	0.51	0.75	**0.93**
purple	0.37	0.79	0.34	**0.87**
can	0.60	0.61	0.40	**0.86**
flower	0.32	0.46	0.61	**0.88**
red	0.35	0.40	0.28	**0.76**

Table 1: F-measure scores for in-focus segmentation. All features are tested by the graph Laplacian based segmentation method.

5   Conclusions

We have introduced a new in-focus measure for estimating the degree of blurriness for low depth-of-field images. Compared with other low-level in-focus features, the harmonic vari-
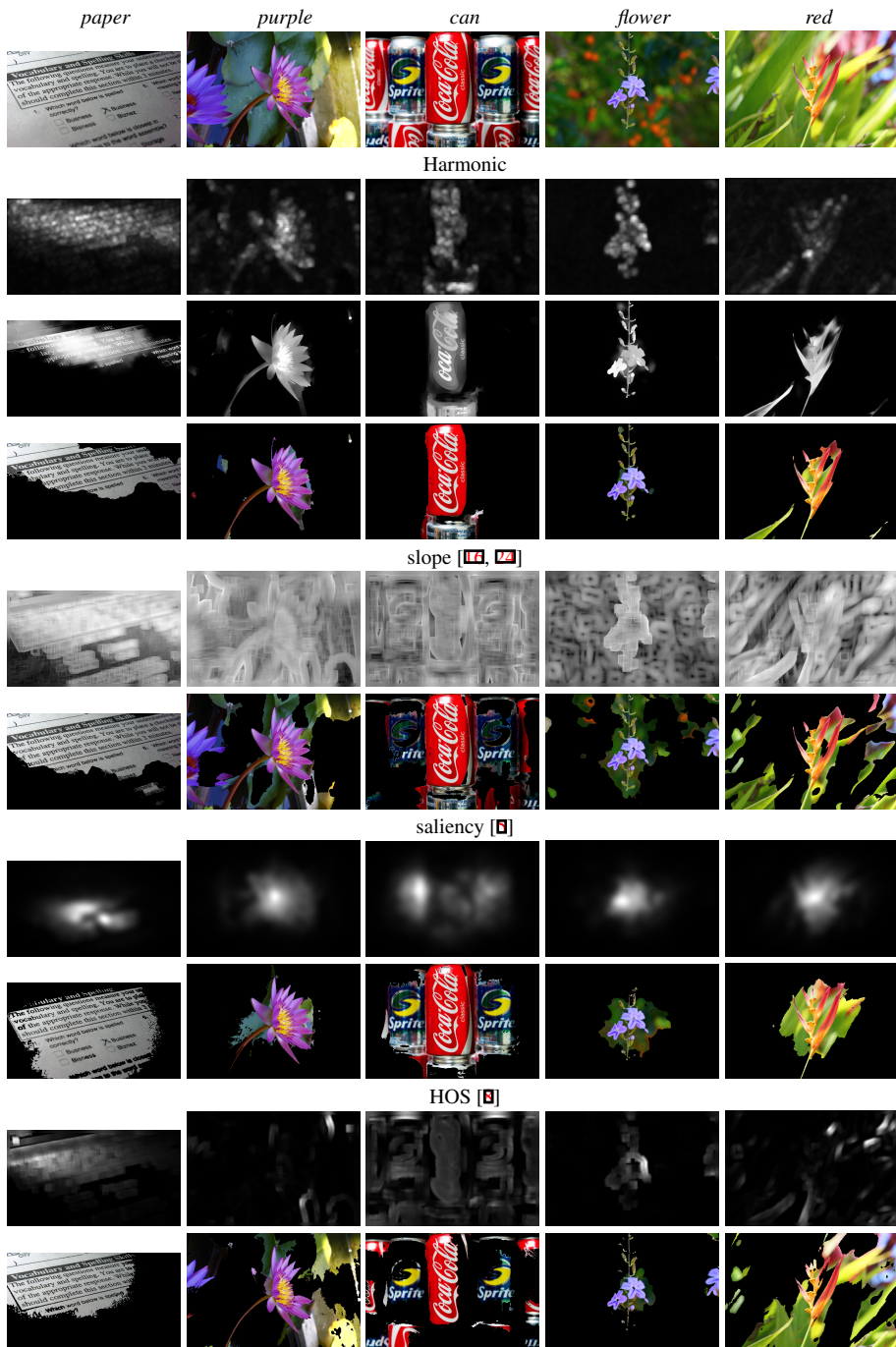
Figure 4: Top-to-bottom: input images, the harmonic variance measure *h* results, the optimized focus likelihood scores *f* of *h*, segmentation results. Other features: defocus maps and the corresponding segmentation results. Unlike the conventional features, the harmonic variance is an accurate indicator of in-focus regions. In addition, our method fits the underlying image structure.

ance can robustly differentiate the in-focus regions from the defocused background even when the background has strong high frequency responses. We also presented a segmentation algorithm derived from robust statistics and Laplacian spectrum analysis to accurately partition out the in-focus regions. In the future, we plan to extend our work for defocus matting and depth estimation.

References

[1] S. Bae and F. Durand. Defocus magnification. *Eurographics*, 2007.

[2] Y. Boykov and G. Funka-Lea. Graph cuts and efficient ND image segmentation. *International Journal on Computer Vision*, 70(2):109–131, 2006.

[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. PAMI*, 23(11):1222–1239, 2001.

[4] J. Elder and S. Zucker. Local scale control for edge detection and blur estimation. *IEEE Trans. PAMI*, 20(7):699–716, 1998.

[5] E. Erdem and A. Erdem. Visual saliency estimation by nonlinearly integrating features using region covariances. *Journal of Vision*, 13(4):1–20, 2013.

[6] L. Grady. Random walks for image segmentation. *IEEE Trans. PAMI*, 28(11):1768–1783, 2006.

[7] P. Huber. *Robust statistics. 1981*. Wiley, New York.

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

[9] L. Kovacs and S. Sziranyi. Focus area extraction by blind deconvolution for defining regions of interest. *IEEE Trans. PAMI*, 29(6):1080–1085, 2002.

[10] E. Krotkov. Focusing. *International Journal on Computer Vision*, 1:223–237, 1987.

[11] A. Levin, B. Fergus, F. Durand, and B. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. on Graphics*, 26(3):70, 2007.

[12] F. Li and F. Porikli. Laplacian spectrum: A unifying approach for enforcing point-wise priors on binary segmentation. *Submitted to ICCV*, 2013.

[13] F. Li, J. Sun, J. Wang, and J. Yu. Dual-focus stereo imaging. *Journal of Electronic Imaging*, 19 (4), 2010.

[14] H. Li and K. Ngan. Unsupervised video segmentation with low depth of field. *IEEE Trans. Circuits Syst. Video Technol.*, 17(12):1742–1751, 2007.

[15] H. Lin and X. Chou. Defocus blur parameters identification by histogram matching. *J. Opt. Soc. Am. A*, 29(8):1694–1706, 2012.

[16] R. Liu, Z. Li, and J. Jia. Image partial blur detection and classification. *CVPR*, 2008.

[17] T. Liu, J. Sun, N. Zheng, X. Tang, and H. Shum. Learning to detect a salient object. *CVPR*, 2007.

[18] Z. Liu, W. Li, L. Shen, Z. Han, and Z. Zhang. Automatic segmentation of focused objects from images with low depth of field. *Pattern Recognition Letters*, 31(7):572–581, 2010.

[19] F. Moreno-Noguer, P. Belhumeur, and S. Nayar. Active refocusing of images and videos. *ACM Trans. on Graphics*, 26(3):67–75, 2007.

[20] V. Namboodiri and S. Chaudhuri. Recovery of relative depth from a single observation using an uncalibrated camera. *CVPR*, 2008.

[21] X. Pan, X. Zhang, and S. Lyu. Exposing image splicing with inconsistent local noise variances. *ICCP*, 2012.

[22] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Trans. on Graphics*, 25(3):795–804, 2006.

[23] A. Saxena, S.H. Chung, and A.Y. Ng. Learning depth from single monocular images. *NIPS*, 2005.

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

[25] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Trans. on Graphics*, 27(3):73, 2008.

[26] A. Stefano, P. White, and W. Collis. Training methods for image noise level estimation on wavelet components. *EURASIP Journal on Applied Signal Processing*, 16:2400–2407, 2004.

[27] C. Swain and T. Chen. Defocus based image segmentation. *ICASSP*, 1995.

[28] Y. Tai and M. Brown. Single image defocus map estimation using local contrast prior. *ICIP*, 2009.

[29] D. Tsai and H. Wang. Segmenting focused objects in complex visual images. *Pattern Recognition Letters*, 19(10):929–949, 1998.

[30] J. Wang, J. Li, and R. Gray. Unsupervised multiresolution segmentation for images with low depth of field. *IEEE Trans. PAMI*, 23(1):85–90, 2001.

[31] C. Won, K. Pyun, and R. Gray. Automatic object segmentation in images with low depth of field. *ICIP*, 2002.

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

[33] D. Zoran and Y. Weiss. Scale invariance and noise in natural images. *ICCV*, 2009.