Focus Quality Metric Based on Visual Sensitivity

Mahdi S. Hosseini, Member, IEEE, Yueyang Zhang, Student member, IEEE, and Konstantinos N. Plataniotis, Fellow, IEEE

Abstract—In this paper, we propose a novel design of Human Visual System (HVS) response in a convolutional filter form to decompose meaningful features that are closely tied with image blur level. Non-reference (NR) based focus quality assessment (FQA) techniques have emerged as the standard of image quality assessment in diverse imaging applications. Despite their high correlation with subjective scoring, they are challenging for practical considerations due to high computational cost and lack of scalability across different image blurs. We bridge this gap by synthesizing the HVS response as a linear combination of Finite Impulse Response (FIR) derivative filters to boost the falloff of high band frequency magnitudes in natural imaging paradigm. The numerical implementation of the HVS filter is carried out with MaxPol filter library that can be arbitrarily set for any differential orders and cutoff frequencies to balance out the estimation of informative features and noise sensitivities. Utilized by HVS filter, we then design an innovative NR-FQA metric called “HVS-MaxPol” that (a) requires minimal computational cost, (b) produce high correlation accuracy with image sharpness level, and (c) scales to assess synthetic and natural image quality. Furthermore, we create a benchmark database in digital pathology for validation of image focus quality in whole slide imaging systems called “FocusPath” consisting of 864 blurred images. Thorough experiments are designed to test and validate the efficiency of HVS-MaxPol across different blur databases and state-of-the-art NR-FQA metrics. The experiment result indicates that our metric has the best overall performance with respect to speed, accuracy and scalability.

Index Terms—No-reference based image sharpness assessment, Visual sensitivity, MaxPol convolutional filters, Natural images, whole slide imaging (WSI) systems

I. INTRODUCTION

IMAGES of natural scenery in the context of image processing follow an inverse magnitude frequency response proportional to $1/\omega^2$, with $\omega$ being the spatial frequency. Such decay implies that the magnitude response of low frequencies is more powerful than that of its high band. Therefore, without processing the raw signals, humans may visually lose fine image edges and perceive only coarse information that can lead to blurry observations. However, in the human visual system (HVS), the distribution of frequency information is perceived equally across the whole spectrum. In fact, the HVS introduces a sensitivity response to the input visual contents to compensate the energy loss of high frequency information and make them as important as the low frequencies [1]–[3]. Biologically speaking, the neurons in the humans’ visual cortex are able to automatically tune the amplitude frequencies, balancing out the existing falloff of the high frequency spectrum. After such tuning, the average response remains constant (balanced) across a wide range of frequencies, which enables human beings to visualize any natural objects clearly, regardless of their size and distance.

The HVS could be modeled as a linear convolution process, with the natural scene modeled as an input power function that decays in the frequency domain, the processed signal in human’s visual cortex as another output power function that remains constant in the frequency domain, and HVS as a convolution filter. Under such assumptions, when a human visualizes a natural scene, the input power function of the scene is convolved with a pre-designed convolution filter inspired by this human’s HVS to generate the output power function of the processed image signal

$$I_O \approx I \ast h_{HVS},$$

where $I_O$ is the output image signal perceived by the human visual cortex, $I$ is the natural image signal and $h_{HVS}$ is the convolution filter simulating the HVS response. To understand how HVS works, this is no more than synthesizing the associated convolution filter, which boosts the power amplitude of the higher frequency to the same level as that of the lower frequency. Under such a model, the sharpness level of an image is obtained by “deconvolving” the visual content with HVS filtering process. In discrete computational modeling, many existing non-reference (NR) focus quality assessment (FQA) approaches make use of this fact to grade the sharpness level of images. Examples include but are not limited to the finite impulse response (FIR) filters [4]–[6], variational gradient map approaches which approximate HVS response by [7], [8], and the contrast map techniques [9]–[11]. However, these existing methods are suboptimal in the sense that they cannot fully manifest the reality of the HVS response.

In this paper, we aim to synthesize the HVS response by modeling the falloff of the natural scenery’s frequency using the generalized Gaussian (GG) distribution [12]. The shape and scale of the distribution conforms with the decay response of natural images in the frequency domain. To compensate for such falloff, the inverse frequency response of GG is considered to be the representative of the synthesized HVS. For numerical approximation, we fit this model with a linear combination of FIR derivative filters, where its numerical solution is provided by the MaxPol filter library introduced in [13]. This library is capable of generating various FIR kernels with different order of derivatives and cutoff frequencies. The multiple derivatives can be used to compensate the falloff of the blur images in the frequency domain and make a balanced average response for observation. The cutoff frequency of the filters also introduces a balanced way of keeping informative features in image decomposition and canceling the high band frequencies that are mostly related to noise artifacts. This newly synthesized HVS

Authors are with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada e-mail: mahdi.hosseini@mail.utoronto.ca
filter provides a unique framework for image decomposition where the blur features are well detected across different image frequencies. Employed by this HVS filter, we then propose our unique framework for image sharpness scoring for NR-FQA development.

A. Related Work in Blind Sharpness Assessment

NR-FQA has recently emerged as a special branch of image quality assessment (IQA) for blindly grading the blur or sharpness level of images. An overview of these methods is listed in Table I. Metrics based on gradient domain maps include maximum local variation [4], energy concentration in image patches [6], and local saliency map [8]. Contrast map is widely used to develop NR-FQA metric. Other methods combine brightness and color information [12], transform contrast values of pixel into the wavelet domain [13], apply the just noticeable blur metric (JNBM) to the pre-calculated contrast map [14], and merge the contrast map with structural change and luminance distortion [15]. Developments based on the wavelet domain have been also studied to decompose the image into several wavelet subbands to compute the log-energy of the decomposed features at pixel level for assisting sharpness evaluation [19]. The phase domain is also commonly used to develop NR-FQA. The theoretical development and relationship between phase domain and sharpness index is studied by Leclaire et al [20]. Furthermore, Hassen et al [21] compute the phase at each location of the image, called Local Phase Coherence (LPC), which is further defined as a sharpness score. Variational maps are also reported as a useful feature. The map has been used in several methods of NR metric development in conjunction with other feature methods [12], [22]. The transfer image domain into total variation and local variation for image analysis is another example to develop NR-FQA. The spectral and spatial domains is used in [11] based on the local maximum total variation domain to construct a third map for sharpness scoring. Metrics based on singular value decomposition to evaluate the image sharpness is also used in [23]. In recent years, various machine learning algorithms have been applied to the sharpness analysis of images. One common tool is the regression model in [12], [24] which performs regression on previously extracted features to approximate sharpness score. Another strong learning algorithm is support vector regression (SVR) in [7], [25] that passes the previously obtained features into convolution neural network (CNN) models for sharpness learning. The CNN model is usually trained to detect the strength of image edges. The CNN model is also used as a standalone pipeline in [26] to construct 50 feature maps, after which the maps are fed into a general regression neural network (GRNN) for scoring. The dictionary based method to learn sparse representation is also used in [5] to obtain sparse coefficients that are closely tied with image sharpness. Finally the k-mean algorithm is also used in [27] to extracted sharpness features.

B. Remaining Challenges and Contributions

The NR-FQA metric has high potential for different applications of the image acquisition pipeline, such as quality check control of high-throughput scanning solutions in digital pathology, astronomy, consumer imaging devices, etc. Owing to the limitations of computational speed, the majority of existing methods lag behind real-time analysis which is mandated by many imaging platforms. After testing several major recent NR-FQA metrics on various databases, we observed that some methods have relatively high accuracy but undesirable computational speed, while others have good performance on time complexity but are not accurate enough. In this context, there is a high demand for a fast and accurate NR-FQA metric. The following lists the contributions made in this paper

- Propose an effective approach to synthesize HVS response model by means of MaxPol library kernels.
- Implement a novel NR-FQA metric called “HVS-MaxPol” which uses the newly synthesized HVS filter to extract meaningful features that are closely tied with human visual blur perception.
- Construct a medical imaging database of digital pathology for sharpness assessment, called “FocusPath”.
- Perform comprehensive comparisons over diverse databases and state-of-the-art NR-FQA metrics in terms of accuracy, computational complexity and scalability.

Our early related work [30] proposes a NR-FQA metric called MaxPol. This metric directly employs MaxPol library for image sharpness assessment [17] that is highly capable of scoring synthetic blur images. In this paper, we employ the MaxPol library to synthesize the HVS by a newly designed

### Table I

| Author       | Year | Learning Algorithm | Learned Domain | Spatial Domain | Frequency Domain | Hybrid Domain | Spectral Domain |
|--------------|------|--------------------|----------------|----------------|-----------------|--------------|-----------------|
| Gu [12]      | 2018 | Regr.              | CN            |                |                 |              |                 |
| Yu [25]      | 2017 | CNN                |                |                |                 |              |                 |
| Yu [23]      | 2017 | CNN                |                |                |                 |              |                 |
| Li [7]       | 2017 | SVR                |                |                |                 |              |                 |
| Li [15]      | 2016 | Dict.              |                |                |                 |              |                 |
| Gu [24]      | 2015 | Regr.              |                |                |                 |              |                 |
| Kang [20]    | 2014 | CNN                |                |                |                 |              |                 |
| Mavridaki [29]| 2014 | SVM               |                |                |                 |              |                 |
| Ye [27]      | 2012 | KMean              |                |                |                 |              |                 |
| Liu [6]      | 2017 |                   | CN            |                |                 |              |                 |
| Li [9]       | 2016 |                   |                |                |                 |              |                 |
| Bahrami [4]  | 2014 |                   | CN            |                |                 |              |                 |
| Gvozden [13] | 2017 |                   |                |                |                 |              |                 |
| Guan [14]    | 2015 |                   |                |                |                 |              |                 |
| Liu [13]     | 2012 |                   | CN            |                |                 |              |                 |
| Vu [19]      | 2012 |                   |                |                |                 |              |                 |
| Leclaire [20]| 2015 |                 |                |                |                 |              |                 |
| Hassen [21]  | 2013 |                 |                |                |                 |              |                 |
| Lee [22]     | 2016 |                   |                |                |                 |              |                 |
| Bahrami [10] | 2016 |                   |                |                |                 |              |                 |
| Vu [11]      | 2012 |                   |                |                |                 |              |                 |
| Sang [23]    | 2014 |                   |                |                |                 |              |                 |
convolution filter for HVS-MaxPol metric development. This metric can efficiently process both synthetic and natural blur images, allowing a broader application to be addressed compared to sharpness metric in [30].

The remainder of the paper is as follows. We design HVS convolution filter in Section II and utilize the filter in Section III for NR-FQA metric development. Section IV introduces a new digital pathology database for NR-FQA benchmarking in natural medical imaging. The experiments are provided in Section V and the paper is concluded in Section VI.

II. VISUAL SENSITIVITY DESIGN MODEL

In this section, we design a new convolutional kernel to synthesize the HVS response. As mentioned earlier, the close approximation of the frequency spectrum of natural images follows a decay order of $\omega^{-1}$. Note that this is under the assumption that the image of interest is in focus. However, once the image is blurred, perhaps due to optical imperfections such as out-of-focus and lens aberrations, the falloff of the spectrum becomes steeper as a consequence. We model such falloff in the amplitude frequency domain with the generalized Gaussian distribution [16]

$$h_{GG}(x) = \frac{1}{2\Gamma(1+1/\beta)A(\beta, \alpha)} \exp\left(-\frac{x^2}{A(\beta, \alpha)}\right)^\beta$$  \hspace{1cm} (1)

where $\beta$ defines the shape of the distribution function, $A(\beta, \alpha) = (\alpha^{2\Gamma(1/\beta)/\Gamma(3/\beta)})^{1/2}$ is the scaling parameter, and $\Gamma(\cdot)$ is the Gamma function $\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}dt, \forall z > 0$. For instance, the standard Gaussian distribution, i.e. second order GG, is a variant of this model, where $\beta = 2$ and $A(2, \alpha)$ [16]. An example of the GG filter kernel is shown in Figure 4 for preset scale and shape parameters. The frequency spectrum of the filter (in the same figure) demonstrates the falloff of the frequency on the high frequency band, a property considered for natural images.

The objective here is to design a finite impulse response (FIR) kernel $h_{HVS}$, where its convolution with the GG kernel yields the delta response

$$h_{GG}(x) * h_{HVS}(x) = \delta(x).$$  \hspace{1cm} (2)

An equivalent representation of the above assumption is that the frequency response of HSV filter has an inverse relation to the natural image frequency model, i.e. $\tilde{h}_{HVS}(\omega) = h_{GG}^{-1}(\omega)$. We synthesize this inverse model by a symmetric FIR filter as a linear combination of multiple bases

$$h_{HVS}(x) \equiv c_1d_2(x) + c_2d_4(x) + \ldots + c_Nd_{2N}(x)$$  \hspace{1cm} (3)

where $d_{2n}(x)$ is the $2n$-th derivative operator. Therefore, the Fourier transform of HVS filter gives

$$\hat{h}_{HVS}(\omega) \equiv \sum_{n=1}^{N} c_n \hat{d}_{2n}(\omega) = \sum_{n=1}^{N} (-1)^nc_n\omega^{2n}. \hspace{1cm} (4)$$

Note that the Fourier transform of the even derivative operator is obtained by $\hat{d}_{2n}(\omega) = (i\omega)^{2n} = (-1)^n\omega^{2n}$. The unknown coefficients $c_n$ are approximated by fitting the model design in [4] to the inverse response $\hat{h}_{GG}^{-1}(\omega)$ using the non-linear least square in [32].

In particular, we consider the lowpass design

$$\hat{h}_{HVS}(\omega) = \begin{cases} \sum_{n=1}^{N} (-1)^nc_n\omega^{2n}, & 0 \leq \omega \leq \omega_c \\ 0, & \omega \geq \omega_c \end{cases} \hspace{1cm} (5)$$

where $\omega_c$ is the cutoff frequency. For numerical approximation of the lowpass derivative filters in (5), we used MaxPol [63294-maxpol-smoothing-and-differentiation-package] to solve numerical differentiation and find a package to solve numerical differentiation [17], [18]. Each derivative operator $d_{2n}(x)$ can be approximated using different orders of polynomials i.e. filter tap length and cutoff frequency to meet the lowpass criterion defined in (5). Figure (c)-(d) demonstrates an example of the impulse and frequency responses of the HVS filter design.

III. FOCUS QUALITY MEASUREMENT

In this section, we elaborate on the different steps used to build the focus quality metric from a digital image. Note that we avoid using color information by converting all color images to grayscale for processing. The proposed contains four main operations discussed in subsequent sections.

A. Background Check

We design a background check condition to exclude the image pixels with dark values. We set a simple hard threshold to exclude all pixels with normalized grayscale values smaller than 0.05. This is because the HVS can hardly extract useful information from an extremely dark environment for interpretation.

B. MaxPol variational decomposition

We decompose the grayscale image $I \in \mathbb{R}^{N_1 \times N_2}$ using the HVS filter $h_{HVS}(x)$ designed in Section II along the horizontal and vertical axes

$$\nabla_{HVS}I = [I * h_{HVS}, I * h_{HVS}^T]^T. \hspace{1cm} (6)$$

https://www.mathworks.com/matlabcentral/fileexchange/63294-maxpol-smoothing-and-differentiation-package
It is worth noting that the selection of the cutoff band $\omega_c$ in designing the HVS filter depends on the signal-to-noise-ratio (SNR) of the measurements. When the SNR measurement is high, then high cutoff should be selected to extract features from the wide frequency band and mitigate information loss. Figure 2(b)-(c) demonstrates the horizontal and vertical decomposition of an image shown in Figure 2a (we used only grayscale for decomposition). While the kernel is highly sensitive on sharp edges, it avoids noise amplification on smooth areas. The distribution of the decomposed features using the HVS filter is shown in Figure 3 for both horizontal and vertical features.

C. Rectified Linear Unit

The HVS filter is a symmetric FIR kernel and provides mostly redundant features on one side of the histogram distribution of decomposed features. To this end, we select features activating beyond zero similar to rectified linear unit (ReLU) function after the convolution operation

$$R(x) = \max(x, 0).$$

Figure 2(e)-(f) demonstrates the activated features beyond zero after decomposition. Figure 2(e)-(f) demonstrates the activated profiles using the decomposed features in Figure 2(b)-(c).

D. Feature map

We construct the feature map using the vector decomposition $\nabla_{HVS} I$ defined by (6) in the $\ell_2$-norm space as follows

$$M_{HVS} = (\|R(\nabla_{HVS} I(x))\|^2 + |R(\nabla_{HVS} I(y))|^2)^{\frac{1}{2}}. \quad (8)$$

The feature map $M_{HVS}$ encodes the edge significance in $\ell_2$-norm space. This promotes the sparsity of decomposed coefficients by weakening minor perturbations while preserving significant coefficients related to image edges. Figure 2d demonstrates the feature map using the activated profiles in Figure 2(e)-(f). The histogram distribution of the decomposed features along horizontal and vertical axes are also shown in Figure 2a.

E. Adaptive hard-thresholding

Similar to [11], we keep a subset of pixels $|\Omega|$ from feature map pixels $M_{HVS}$ to eliminate shallow coefficients that are less related to focus features

$$M_{HVS} = sort_d(M_{HVS}), \quad k \in \Omega \subset \{1, \ldots, N_1N_2\}, \quad (9)$$

where $sort_d$ is the sorting operator in descending form and the notation $|\Omega|$ refers to the cardinality of subset $\Omega$. In addition, we adaptively select the number of remaining pixels $|\Omega|$ from the histogram distribution of absolute decomposed coefficients $|\nabla_{HVS} I|$ in [6] and map the approximated variance of distribution using a nonlinear projection

$$p(\sigma) = \frac{1}{4}(1 - \tanh(60(\sigma - 0.095))) + 0.09. \quad (10)$$

Note that the approximated variance $\sigma$ is normalized in scale by the maximum amplitude of $|\nabla_{HVS} I|$. The image plot of this projection is shown in Figure 3b. A lower value of $\sigma$ is an indication of highly sparse image in the decomposed HVS domain $\nabla_{HVS} I$ where more pixel coefficients are kept to exploit pertinent focus information. High $\sigma$ is related to out-of-focus images where we keep less coefficients to avoid over fitting.

$$\mu_m = \mathbb{E} [(\Omega_{HVS} - \mu_0)^m] \quad (11)$$

F. Central moments information

The HVS map obtained from [6] contains high order moment features that are decomposed by the superposition of different order of derivatives defined in Section III. These features are related mostly to the image edges and contain high frequency components. To extract meaningful features from such a map, we measure the $m$th central moment of $\Omega_{HVS}$ for feature extraction
where $E[.]$ is the expectation and $\mu_0 = \frac{1}{N} \sum_{k \in \Omega} \mathbb{M}_{\text{HVS}}(k)$ is the average value of the remaining features. The $m$th moment $\mu_m$ encodes the $m$th power of deviation of variables $\mathbb{M}_{\text{HVS}}$ from its mean $\mu_0$. Here the negative logarithm of the moment is considered as the final score

$$C = - \log \mu_m$$

which represents the sharpness score related to an individual HVS filter.

G. Combination Weights Selection

Often it is meaningful to deploy more than one HVS filter for feature extraction due to the complexity of blur type in an image of interest. To simulate HVS using multiple kernels, we linearly combine multiple features extracted from different HVS filters. We defined the weights of the linear combination by regularizing the extracted features in IQA measure defined in Equation (13). The IQA measurement takes two sets of objective scores $X$ and subjective scores $Y$ as input and perform a nonlinear regression between these two sets. Next, the best projection of objective scores, $\hat{Y}$, is compared with the subjective scores $Y$ to show the correlation between $X$ and $Y$ in terms of the metrics such as PLCC, SRCC, KRCC, and RMSE [33], [34]. The nonlinear fitting is realized by tuning five parameters to minimize the RMSE between $\hat{Y}$ and $Y$, that is $\|Y - \hat{Y}\|_2^2$. The equation of $\hat{Y}$ is:

$$\hat{Y} = Q(X) = \beta_1\left(\frac{1}{2} + \frac{1}{1 + e^{\beta_2(X - \beta_3)}}\right) + \beta_4 X + \beta_5$$

(13)

where $\beta_i$ are the tuning parameters to be determined. Here, we revise this fitting problem for multiple objective scoring by defining a matrix-vector multiplication

$$X = MW,$$

(14)

where

$$M = \begin{bmatrix} C_{k1}^{(1)} & C_{k2}^{(1)} \\ C_{k1}^{(2)} & C_{k2}^{(2)} \\ \vdots & \vdots \\ C_{k1}^{(N)} & C_{k2}^{(N)} \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

(15)

where $C_{kj}^{(i)}$ refers to the objective scoring of the $i$th observation using the $j$th feature decomposed by the corresponding HVS filter. By substituting (15) into (13), the revised IQA metric for multiple feature measurement gives

$$\hat{Y} = Q(MW) = \beta_1\left(\frac{1}{2} + \frac{1}{1 + e^{\beta_2(MW - \beta_3)}}\right) + \beta_4 MW + \beta_5.$$ 

(16)

The parameters are identified by fitting the revised IQA $\hat{Y}$ in [16] to the subjective scores accordingly.

IV. FocusPath: A Digital Pathology Archive

The primary purpose of proposing this pathology database is to evaluate the performance of different non-reference sharpness assessment metrics on pathological images. In whole slide imaging (WSI) system, tissue slides are scanned automatically where out-of-focus scans are a common problem. The common practice in digital pathology (DP) is to install the tissue slide on a slide-holder and feed the slide holder into the scanner. The level of the tissue slide is automatically moved up/down in quarter micron (0.25$\mu$) resolution to adjust the focus level with respect to the focal length of the optical lens installed in the scanner. If the position of the stage is not adjusted properly, the tissue depth level will be mis-aligned with focal length, and hence the image will be perceived in blur. Such irregularities need to be detected and addressed accordingly by sending a feedback to the system to retake the scan or adding a quality check (QC) control on the scanning process.

All the digital pathological images included are captured as WSIs using the Huron TissueScope LE1.2 [35]. The scanner automatically digitizes, processes and stores the images of tissue slides in .tif format. To create a database with high diversity, the tissue slides used are cut from 9 distinct types of organs (i.e. 9 Slides). When scanning through a whole organ tissue slide, the scanner captures multiple image patches all in one, which are called Strips. We include 2 strips for each tissue slide to ensure we take different morphological structure across a wide tissue area. The WSI scanner generates several in-depth Z-stacks for each strip, where each z-stack is referred as a Slice. Different z-stacks represent different vertical positions of the WSI camera, which is highly related to the image blur. We include 16 z-stacks for every image strip of a certain tissue slide in the pathology database. In addition, we randomly choose 3 different positions and accordingly crop 3 image patches of 1024 $\times$ 1024 pixel size from the raw image. To summarize, the total number of images included is 864, cropped from 9 Slides $\times$ 2 Strips $\times$ 3 Positions $\times$ 16 Slices. The database is called FocusPath and is publicly published online. A sample of 16 different focus levels is shown in [16] Note that we have an extended version of FocusPath containing 8640 images extracted from 30 positions instead of 3 for FQA analysis. Since the database is too big, we avoid uploading it online, but it can be requested from the authors on demand.

[Image 1]

(a) Z1 (b) Z2 (c) Z3 (d) Z4 (e) Z5 (f) Z6 (g) Z7 (h) Z8

(i) Z9 (j) Z10 (k) Z11 (l) Z12 (m) Z13 (n) Z14 (o) Z15 (p) Z16

Fig. 4. The sample images of Slice 1 – 16. Sample images of different Slice index vary in blur levels, with the clearest one in the middle. All the samples come from Slide 1, Strip 0 and Position 1.

The subjective sharpness score assigned to each image in the database represents the blur level in terms of the human visual system. As mentioned above, different blur levels are caused by different vertical positions of the WSI camera. In this sense, the proposed subjective score is the difference between the best camera position $Z^*$ and the position of interest $Z_i$. In other words, the score represents how far the camera is

2FocusPath database: https://sites.google.com/view/focuspathuof
away from its best focus level when scanning a certain image slide. Specifically, every 16 images of the same content but different blur levels have a common $Z^*$ and every image has a corresponding $Z_i$. A sample set of 16 images is shown in Figure 4. The $Z^*$ value among every 16 images is determined by the proposed non-reference sharpness assessment metric in this paper. Since $Z^*$ is found among 16 different slices of the same tissue content, the scores will guarantee to obtain the best focus level. The plot of subjective scores across 16 images is shown in Figure 5. The mathematical representation of the subjective sharpness score is:

$$\text{Subjective Score at } Z_i = Z_i - Z^*, i = 1, 2, 3...16 \quad (17)$$

To make it clearer, we list the relevant information about the FocusPath in Table III. In addition, as mentioned before, we include 9 tissue slides in FocusPath. They are from different types of organs with different staining types, which guarantees diversity and effectiveness of the database. Further information about the tissue slides and sample images are available in Table II.

V. EXPERIMENTS

To evaluate the proposed HVS MaxPol model, we conduct experiments in terms of accuracy, time complexity and scalability. All of these three evaluation criteria are highly related to the practical application of a NR-IQA metric. For comparison, we select 9 state-of-the-art non-reference sharpness assessment techniques, including S$_3$ [11], MLV [4], Kang’s CNN [26], ARISMC [24], GPC [20], SPARISH [5], RISE [7], Yu’s CNN [25], and MaxPol [30].

A. Parameter Tuning

A single HVS-MaxPol kernel contains three parameters to tune, where two parameters are the scale $\alpha$ and shape $\beta$ of the associated GG blur kernel used for the construction of the HVS response, and the third parameter is the cutoff frequency $\omega_c$ set for the stop band design. For more information on these parameters, please refer to Section II. An additional parameter is also defined in Section III to extract different order of moment $\mu$ from decomposed image features. To determine the single kernel with proper parameter adjustment yielding the highest accuracy, we perform a grid search on $\alpha$, $\beta$, $\mu$ and $\omega_c$ with respect to the weighted overall accuracy across four synthetic and three natural databases. After determining the best single kernel, we further perform another grid search to find the second best kernel that achieves the best overall accuracy. The second kernel is combined with the previously-determined best kernel using the combination weight selection defined in Section III. Specifically, we combine 2-dimensional $W$ and 5-dimensional $\hat{\beta}$ into a single 7-dimensional vector $\hat{\beta}$ and pass it to nlinfit function in MATLAB for optimization. The optimal value for $\hat{\beta}$ will be returned. We denote the optimal value of $\hat{\beta}$ as $[\hat{\beta} \ W]$ and thus $M\hat{W}$ is the optimized objective scores. We select the optimal weight matrix $\hat{W}$ in terms of the PLCC performances tuned on FocusPath and CID2013 separately. In this case, HVS-MaxPol-1 uses the best single kernel and HVS-MaxPol-2 uses the combination of the best two kernels. All the parameters are independently tuned based for synthetic and natural databases and the final values for each parameter are listed in Table IV.

B. Performance Evaluation

The accuracy analysis are performed on seven distinct image blur databases. Specifically, we include four synthetic blur image databases, LIVE [34], CSIQ [36], TID2008 [37], and

---

**TABLE II**

| Slide # | Slide 1 | Slide 2 | Slide 3 | Slide 4 | Slide 5 | Slide 6 | Slide 7 | Slide 8 | Slide 9 |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Stain Type | Trichrome | H&E | Mucicarmine | IRON (PE) | ABF | Congo Red (CR) | PAS | Grocott | H&E |

**Example**

![Image 1](image1.png) ![Image 2](image2.png)

(a) quality assessment metric  
(b) absolute subjective score

---

**TABLE III**

| Features | Description |
|----------|-------------|
| # of Slides | 9           |
| # of Strips | 2           |
| # of Positions | 3           |
| # of Slices (Z-Stack) | 16         |
| Image Format | tiff         |
| Image Size | 1024 $\times$ 1024 |
| Pixel Resolution | 0.25$\mu$ |
| Optical Zoom | 40X         |
| Color Variation | Diverse Gamut |
| Focus Resolution | Diverse Gamut |
| Background Ratio | $<50\%$ |

Fig. 5. The objective scores generated using the proposed sharpness assessment metric and the absolute subjective score. The score peak indicates the clearest image among the set of sample images. In this sense, we pick the corresponding slice index of the peak as $Z^*$ for this set of images. The subjective scores are assigned in terms of the selected $Z^*$. All images are selected from Slice 1-16, Slide 1, Strip 0 and Position 1.
TABLE IV  
THE BEST SETTINGS FOR HVS MAXPOL-1 AND HVS MAXPOL-2 OVER NATURAL AND SYNTHETIC DATABASES FROM GRID SEARCH. THE SEARCH RANGE OF α IS 0.7–3, 0.8–2 FOR β, 3–23 FOR ωc, AND 2–14 FOR μ.

| Synthetic | HVS-1 | HVS-2 | Natural | HVS-1 | HVS-2 |
|-----------|-------|-------|---------|-------|-------|
| α         | 0.7   | 0.7   | 1.7     | 1.7   | 0.7   |
| β         | 0.8   | 0.8, 0.9| 1.4     | 1.4   | 0.8   |
| cutoff(ωc)| 19    | 19, 20| 13      | 13, 26|       |
| moment(μ) | 20    | 20, 12| 12      | 12, 4 |       |

TID2013 [38], including 145, 150, 100 and 125 Gaussian blurred images, respectively. We also include three natural blur image databases, BID [39], CID2013 [40], and our newly introduced FocusPath [7], with 586, 474 and 864 images respectively. The image content of BID and CID2013 varies from human beings to natural landscapes while the image content of FocusPath is related to pathological organs. Figure 6 demonstrates one sample image from all seven databases. The accuracy measurement of different NR focus quality metrics are reported using Pearson linear correlation coefficient (PLCC) and Spearman rank order correlation (SRCC) indicating the linear correlation and strength and direction of monotonicity between objective and subjective scoring, respectively [33, 34].

![Image](https://sites.google.com/view/focuspathuoft)

Fig. 6. One sample image from each natural (a)-(c) and synthetic (d)-(g) databases.

Table V demonstrates the performance scores for all methods using different databases. In natural imaging, HVS-MaxPol-2 achieves the highest overall (weighted) PLCC score of 0.7059, leading the formerly best metric RISE by 3.49%. HVS-MaxPol-1 earns the highest SRCC score of 0.6730, introducing a performance gain of 1.09% compared to RISE. Specifically, HVS-MaxPol-2 and HVS-MaxPol-1 outrank the other metrics over FocusPath, achieving at least 3% increase compared to the third best S3. Over BID and CID2013, the proposed HVS MaxPol metrics rank second while RISE ranks first on both databases. It worth noting that the RISE metric is trained with respect to each database where the scores are tested using only 20% of the whole databases, while our metrics are evaluated on the whole databases sharing a common parameter setting overall databases. This can also explain RISE’s poor performance over FocusPath (864 images) using the default mode. For synthetic image analysis, MaxPol metric outperforms all of the competing metrics across all databases. HVS-MaxPol metrics rank fourth and fifth in terms of overall performance where both exceed 0.92, and are only 2–3% lower than the best metric MaxPol and approximately 1% lower than the second best metric RISE. In particular, HVS-MaxPol-1 achieves the best performance on LIVE in terms of PLCC while HVS MaxPol-2 ranks second in terms of both PLCC and SRCC. HVS MaxPol-2 also earns the second-best PLCC score over CSIQ.

We conduct the statistical significance analysis between HVS-MaxPol-1, HVS-MaxPol-2, Maxpol and the rest of metrics over seven databases. In particular, the significance test is defined by the one-sided T-test with a 95% confidence level. As we can see from the result in Table VI, HVS MaxPol-1 outperforms seven and nine other methods over synthetic and natural databases, respectively. HVS MaxPol-2 outperforms 12 and nine other metrics over synthetic and natural databases, respectively. This indicates that the HVS-MaxPol metrics’ performances are above the average over synthetic databases and are significantly better than most other metrics over natural databases. In addition, Yu’s CNN, RISE and Kang’s CNN, the metrics that outperform HVS-MaxPol over some database, only use 20% of each database for testing (60% as the training set and 20% as the validation set for Yu’s CNN, 80% as the training set for RISE and Kang’s CNN). Besides, according to the analysis of MaxPol, it is statistically better than most of the metrics over synthetic databases, showing its outstanding performance.

To better visualize the performance of HVS-MaxPol-1 and HVS-MaxPol-2, we include the scatter plots of generated objective scores versus the corresponding subjective scores on every database in Figure 7. The plots also include the regression curves that are overlaid on the figures. The concentration of the scatter points around this curve implies the closeness of the metric to the subjective ratings. The deviation of points from the regression curve in natural images is higher compared to synthetic images, yielding inferior performance in general as expected. Furthermore, the objective scores are widely spread along the vertical axes with respect to their subjective scores, indicating that the measure has direct relationship (approximately linear) with subjective scores. Such relationship implies a better interpretation of the objective scores in the context the of HVS [41].

Overall, the proposed HVS-MaxPol-1 and HVS-MaxPol-2 perform well on scoring both synthetic and natural blur images. Our proposed metric does not require any training and can be reliably and effectively employed in practice regardless of the source and type of image blur. In addition, MaxPol is especially excellent in scoring synthetic blur images, and thus it can be another optimal option for synthetic blur assessment. It worth noting that we did not evaluate Yu’s CNN and Kang’s CNN on natural databases due to the lack of an existing pre-trained model. We also did not train the networks on these databases due to the complexity of their training procedural guidelines.

C. Scalability Analysis

In this section, we evaluate the scalability of all NR-FQA metrics. We define the scalability here as the consistency of
### TABLE V

Performance evaluation of 11 NR focus quality assessment metrics over seven image blur databases, including four synthetic and three natural blur databases. The performances are reported in terms of correlation accuracies PLCC and SRCC scores. The overall weighted average performances are also computed separately for synthetic blur and natural blur databases.

| Method          | Year | Measure | LIVE | CSIQ | TID2008 | TID2013 | BID | CID2013 | Overall-Synthetic | Overall-Natural |
|-----------------|------|---------|------|------|---------|---------|-----|---------|-----------------|----------------|
| S1 [11]         | 2012 | PLCC    | 0.9434 | 0.9175 | 0.8555 | 0.8816 | 0.4271 | 0.6863 | 0.7906 | 0.9042 | 0.6542 |
|                 |      | SRCC    | 0.9436 | 0.9058 | 0.8480 | 0.8609 | 0.4253 | 0.6460 | 0.7914 | 0.8944 | 0.6441 |
| MLV [2]         | 2014 | PLCC    | 0.9590 | 0.9069 | 0.8584 | 0.8830 | 0.3643 | 0.6890 | 0.3201 | 0.9063 | 0.4243 |
|                 |      | SRCC    | 0.9566 | 0.9246 | 0.8546 | 0.8785 | 0.3236 | 0.6206 | 0.3296 | 0.9090 | 0.3993 |
| Kang’s CNN [26] | 2014 | PLCC    | 0.9625 | 0.7743 | 0.8603 | 0.9308 | -     | -     | -     | 0.8848 | -     |
|                 |      | SRCC    | 0.9831 | 0.7806 | 0.8496 | 0.9215 | -     | -     | -     | 0.8842 | -     |
| ARISM [24]      | 2015 | PLCC    | 0.9590 | 0.9481 | 0.8544 | 0.8979 | 0.1841 | 0.5523 | 0.2263 | 0.9211 | 0.2936 |
|                 |      | SRCC    | 0.9561 | 0.9314 | 0.8681 | 0.9015 | 0.1742 | 0.4719 | 0.3043 | 0.9189 | 0.3059 |
| GPC [20]        | 2015 | PLCC    | 0.9242 | 0.9018 | 0.8684 | 0.8665 | 0.4409 | 0.6520 | 0.7499 | 0.8931 | 0.6317 |
|                 |      | SRCC    | 0.8369 | 0.8641 | 0.8729 | 0.8668 | 0.4361 | 0.6608 | 0.7811 | 0.8589 | 0.6334 |
| SPARISH [4]     | 2016 | PLCC    | 0.9595 | 0.9380 | 0.8900 | 0.9020 | 0.3460 | 0.6775 | 0.3439 | 0.9261 | 0.4273 |
|                 |      | SRCC    | 0.9593 | 0.9139 | 0.8836 | 0.8940 | 0.3413 | 0.6607 | 0.3566 | 0.9159 | 0.4267 |
| RISE [7]        | 2017 | PLCC    | 0.9620 | 0.9463 | 0.8289 | 0.9419 | 0.6017 | 0.7934 | 0.6509 | 0.9463 | 0.6710 |
|                 |      | SRCC    | 0.9493 | 0.9279 | 0.9218 | 0.9338 | 0.5839 | 0.7690 | 0.6566 | 0.9341 | 0.6621 |
| Yu’s CNN [25]   | 2017 | PLCC    | 0.9730 | 0.9416 | 0.9374 | 0.9221 | -     | -     | -     | 0.9449 | -     |
|                 |      | SRCC    | 0.9646 | 0.9253 | 0.9189 | 0.9135 | -     | -     | -     | 0.9322 | -     |
| MaxPol [30]     | 2018 | PLCC    | 0.9735 | 0.9657 | 0.9359 | 0.9412 | 0.3235 | 0.5674 | 0.7056 | 0.9563 | 0.5552 |
|                 |      | SRCC    | 0.9688 | 0.9581 | 0.9394 | 0.9448 | 0.2713 | 0.5310 | 0.7191 | 0.9514 | 0.5364 |
| HVS-MaxPol-1    | 2018 | PLCC    | 0.9877 | 0.9506 | 0.8811 | 0.8977 | 0.4112 | 0.7741 | 0.8212 | 0.9349 | 0.6847 |
|                 |      | SRCC    | 0.9722 | 0.9209 | 0.8813 | 0.8930 | 0.4363 | 0.7081 | 0.8144 | 0.9209 | 0.6730 |
| HVS-MaxPol-2    | 2018 | PLCC    | 0.9789 | 0.9507 | 0.8964 | 0.8980 | 0.4659 | 0.7329 | 0.8538 | 0.9355 | 0.7059 |
|                 |      | SRCC    | 0.9737 | 0.9216 | 0.8956 | 0.9014 | 0.4475 | 0.6102 | 0.8574 | 0.9263 | 0.6717 |

### TABLE VI

Statistical performance between HVS-MaxPol-1 and other competing methods (except HVS-MaxPol-2 and MaxPol). '+1' indicates that HVS-MaxPol-1 outperforms the competing method on a certain database. '-1' indicates the opposite. and '0' indicates that there is no significant difference between HVS-MaxPol-1 and the competing method.

| Metric          | Synthetic Blur Database | Natural Blur Database |
|-----------------|-------------------------|-----------------------|
|                | LIVE | CSIQ | TID2008 | TID2013 | BID | CID2013 | FocusPath |
| S1 [11]        | +1   | +1   | 0       | 0       | +1  | 0       | +1        |
| MLV [2]        | +1   | +1   | 0       | 0       | +1  | 0       | +1        |
| Kang’s CNN [26]| 0    | +1   | 0       | 0       | +1  | 0       | +1        |
| ARISM [24]     | 0    | 0    | +1      | +1      | 0   | +1      | +1        |
| GPC [20]       | 0    | 0    | +1      | +1      | 0   | +1      | +1        |
| SPARISH [4]    | 0    | 0    | 0       | 0       | +1  | 0       | +1        |
| RISE [7]       | 0    | 0    | −1      | −1      | −1  | 0       | +1        |
| Yu’s CNN [25]  | 0    | 0    | −1      | −1      | −1  | 0       | +1        |

### TABLE VII

Statistical performance between HVS-MaxPol-2 and other competing methods (except HVS-MaxPol-1 and MaxPol). '+1' indicates that HVS-MaxPol-2 outperforms the competing method on a certain database. '-1' indicates the opposite. and '0' indicates that there is no significant difference between HVS-MaxPol-2 and the competing method.

| Metric          | Synthetic Blur Database | Natural Blur Database |
|-----------------|-------------------------|-----------------------|
|                | LIVE | CSIQ | TID2008 | TID2013 | BID | CID2013 | FocusPath |
| S1 [11]        | +1   | +1   | 0       | 0       | +1  | 0       | +1        |
| MLV [2]        | +1   | +1   | 0       | 0       | +1  | 0       | +1        |
| Kang’s CNN [26]| 0    | +1   | 0       | 0       | +1  | 0       | +1        |
| ARISM [24]     | +1   | 0    | +1      | +1      | 0   | +1      | +1        |
| GPC [20]       | +1   | 0    | +1      | +1      | 0   | +1      | +1        |
| SPARISH [4]    | +1   | 0    | 0       | 0       | +1  | 0       | +1        |
| RISE [7]       | +1   | 0    | −1      | −1      | −1  | 0       | +1        |
| Yu’s CNN [25]  | 0    | 0    | −1      | −1      | −1  | 0       | +1        |
In particular, high-speed NR-FQA metric is a must in high-throughput digital image archiving platforms such as whole slide imaging in digital pathology and planetary observations in reconnaissance satellite orbiters such as MRO \[42\] and LRO \[43\]. In this section, we design two sets of experiments. The first one is the assessment of correlation accuracy of NR-FQA metrics versus CPU time over both natural and synthetic databases. We study the second experiment by analyzing the complexity of CPU time versus different image sizes. For the CPU time measure, all the experiments are done on a Windows station with an AMD FX-8370E 8-Core CPU 3.30 GHz. The relationship between PLCC performance and the real CPU time of each method is shown in Figure\[9\] where the overall weighted PLCC performances are used. A large y-axis value indicates a high accuracy and a small x-axis value indicates low time consumption. Thus, an excellent method should locate at the top-left corner in each plot. As we can see from the first plot in Figure\[9\] MaxPol, HVS-MaxPol-1 and HVS-MaxPol-2 are located at the top-left corner, indicating excellent accuracy performance for synthetic blur images. In particular, MaxPol has the highest accuracy and HVS MaxPol-1 consumes the least computational time. Additionally, although RISE has a higher accuracy compared to HVS MaxPol-1 and HVS MaxPol-2 over synthetic databases, its time consumption is approximately 100 times higher than the two HVS-MaxPol metrics. With respect to the natural image database shown in the second plot of Figure\[9\] the top three metrics located on the top-left corners are HVS-MaxPol-1, HVS-MaxPol-2, and

### D. Computational Complexity Analysis

Our final analysis is to study the computational complexity of all NR-FQA metrics. The speed of metric calculation in digital computing is highly critical for practical considerations.

**TABLE VIII**

| Metric          | Synthetic Blur Database | Natural Blur Database |
|-----------------|-------------------------|-----------------------|
|                 | LIVE        | CSIQ       | TID2008 | TID2013 | BID      | CID2013 | FocusPath |
| S3 [11]         | +1          | +1         | +1      | +1      | -1       | -1      | -1        |
| MLV [2]         | +1          | +1         | +1      | +1      | 0        | 0       | +1        |
| Kang's CNN [26] | 0           | +1         | +1      | +1      | +1       | 0       | +1        |
| ARISMA [24]     | +1          | 0          | +1      | +1      | +1       | 0       | +1        |
| GPC [20]        | 0           | +1         | +1      | +1      | -1       | 0       | -1        |
| SPARISH [5]     | +1          | 0          | +1      | +1      | 0        | -1      | +1        |
| Yü's CNN [25]   | 0           | 0          | 0       | +1      | -1       | -1      | +1        |

Fig. 7. Scatter plots of subjective MOS versus objective scoring using HVS MaxPol-1 (first row) and HVS MaxPol-2 (second row). The experiments are done for synthetic and natural blur databases as shown in the labels. The fitted regression curve is overlaid by a black solid line on all plots.

Performance over different image database volume. For the sake of experiment, we have selected the extended version of FocusPath which contains 8640 blur images in total for FQA analysis. We randomly select different database sizes of \{3\%, 6\%, 9\%, …, 30\%\} from the whole 8640 images. Monte-Carlo simulation is done over 50 different random selections from each database size. For each shuffle, we evaluate the correlation accuracy of different NR-FQA metrics and show them with a box-plot in the Figure\[8\]. The bottom line and top line of the box represent 25th and 75th percentile of all data and the central red line stands for the median value. A good metric should (a) yield high median value with minimal box size related to the standard deviation of the data, and (b) provide consistent performance over different database size. Overall, HVS-MaxPol-1 and HVS-MaxPol-2 have the best scalability with respect to the criteria defined in (a) and (b), where both methods generate the smallest box size and remain on the same level along different database size. We can also observe that S3 and MaxPol are slightly worse than HVS MaxPol-1 and HVS MaxPol-2 but much better than other metrics under the same analysis. Moreover, the performance of GPC drops dramatically as the size of database increases, showing its poor scalability.

**Objective Score**

**Subjective Score**

**Performance**

\( \text{Subjective Score} \)

\( \text{Objective Score} \)

**Regime**

- \( +1 \) indicates that MaxPol outperforms the competing method on a certain database.
- \( -1 \) indicates the opposite.
- \( 0 \) indicates that there is no significant difference between MaxPol and the competing method.

**Figure 8**

The fitted regression curve is overlaid by a black solid line on all plots.
We conduct further experiment to analyze the computational complexity (in the CPU time) over different image sizes. For this experiment, we select 20 samples of pathological images in square sizes \{64, 128, 256, 512, 1024, 2048\}. The CPU time is averaged over all 20 trials and the performances are shown in Figure 10. Each curve in the figure indicates the computational time complexity of the corresponding NR-FQA metric. According to the results, the rank observation of all metrics are consistent over different image size. HVS-MaxPol-1 is the most time efficient metric among all the evaluated metrics. It is almost 1000 times faster than ARISM, 100 times faster than RISE and S3, and almost 10 times faster than MLV. HVS-MaxPol-2 ranks second, two times slower than HVS-MaxPol-1, followed by GPC and MaxPol.

VI. CONCLUSION

In this paper, we have introduced a novel non-reference image sharpness assessment metric, which can process the input image efficiently and accurately. The foundation of this metric is based on the human vision system (HVS) response. We have simulated the HVS response using a symmetric FIR kernel as a superposition of multiple even-order derivative kernels which boosts high frequency domain magnitudes in a balanced way similar to the HVS response. The numerical approximation of the proposed HVS filter is accommodated through MaxPol filters image feature extraction that are closely related to focus blur. Thorough experiments are conducted on four synthetic blur image databases and three natural blur image databases for the comparison of eight state-of-the-art non-reference focus quality assessment metrics. The result shows that our metric significantly outperforms other metrics over both synthetic and real-world images.
and natural blur databases. We also conduct time complexity experiment and scalability experiments, both of which validate the superior performance of our metric. Besides, we construct and propose a natural blur digital pathology image database named “FocusPath”. Such a database can help the validation of non-reference image sharpness assessment metrics for medical imaging platforms. Unlike the previous metrics, our metric has the unique advantage of optimizing both speed and accuracy. The metric can be used for both synthetic and natural image assessment with a common parameter setting. This allows a wide range of practical applications to be addressed for focus quality assessment. Our future development will concern utilizing this metric in big image archiving database such as whole slide imaging systems and reconnaissance orbital imaging in satellites for quality check control application.

**ACKNOWLEDGMENT**

The authors would like to thank Huron Digital Pathology Inc. for providing digital pathology scan image database and their fruitful discussion during the development of the NR-FQA metric. The first and second authors research was partially supported by an NSERC Collaborative Research and Development Grant (contract CRDPJ-515553-17)

**REFERENCES**

[1] David J. Field, “Relations between the statistics of natural images and the response properties of cortical cells,” *Josa a*, vol. 4, no. 12, pp. 2379–2394, 1987.
[2] Nuala Brady and David J. Field, “What’s constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns,” *Vision Research*, vol. 35, no. 6, pp. 739–756, 1995.
[3] David J. Field and Nuala Brady, “What’s constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns,” *Vision Research*, vol. 35, no. 6, pp. 739–756, 1995.
[4] Khosro Bahrami and Alex C. Kot, “Image sharpness assessment based on discrete orthogonal moments,” *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 39–50, 2016.
[5] Leida Li, Dong Wu, Jinhian Wu, Haoliang Li, Weisi Lin, and Alex C. Kot, “Image sharpness assessment by sparse representation,” *IEEE Transactions on Multimedia*, vol. 19, no. 5, pp. 1030–1040, 2017.
[6] Leida Li, Weisi Lin, Xuesong Wang, Gaobo Yang, Khosro Bahrami, and Alex C. Kot, “No-reference image blur assessment based on discrete orthogonal moments,” *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 39–50, 2016.
[7] Leida Li, Wenhao Xia, Weisi Lin, Yuming Fang, and Shiqi Wang, “No-reference and robust image sharpness evaluation based on multiscale spatial and spectral features,” *IEEE Transactions on Multimedia*, vol. 19, no. 5, pp. 1030–1040, 2017.

**TABLE IX**

| Metric         | Speed | PLCC Synth. | SRCC Synth. | Scalability |
|----------------|-------|-------------|-------------|-------------|
| S3 [11]        | ✓     | x           | x           |             |
| MLV            | ✓     | ✓           | ✓           | x           |
| Kang’s CNN [26]| x     | ✓           | ✓           | x           |
| ARISM [24]     | ✓     | ✓           | ✓           | ✓           |
| GPC [20]       | ✓     | ✓           | ✓           | ✓           |
| RISE [7]       | ✓     | ✓           | x           | x           |
| Yu’s CNN [26]  | x     | ✓           | x           | x           |
| MaxPol         | ✓     | ✓           | ✓           | ✓           |
| HVS-MaxPol-1   | ✓     | ✓           | ✓           | ✓           |
| HVS-MaxPol-2   | ✓     | ✓           | ✓           | ✓           |

INDICATES THAT THE CORRESPONDING METRIC RANKS THE TOP FOUR UNDER A CERTAIN ANALYSIS AND X INDICATES THE OPPOSITE.

**Fig. 9.** PLCC values vs. CPU time among all methods on synthetic and natural databases. Markers located at the top-left corner are desirable, meaning the method has high accuracy and speed.

**Fig. 10.** The computational time vs the pixel number of an image using different methods. Lower curve shows computational efficiency.
[8] Yutao Liu, Ke Gu, Guangtao Zhai, Xianning Liu, Debin Zhao, and Wen Gao, “Quality assessment for real out-of-focus blurred images,” Journal of Visual Communication and Image Representation, vol. 46, pp. 70–80, 2017.

[9] Wufeng Xue, Lei Zhang, Xuanqin Mou, and Alan C. Bovik, “Gradient magnitude similarity deviation: A highly efficient perceptual image quality index,” IEEE Transactions on Image Processing, vol. 23, no. 2, pp. 684–695, 2014.

[10] Khosro Bahrami and Alex C. Kot, “Efficient image sharpness assessment based on content aware total variation,” IEEE Transactions on Multimedia, vol. 18, no. 8, pp. 1568–1578, 2016.

[11] Yu CT, Phan TD, and Chandler DM, “S3: A spectral and spatial measure of local perceived sharpness in natural images,” IEEE Transactions on Image Processing, vol. 21, no. 3, pp. 934–945, 2012.

[12] Ke Gu, Dacheng Tao, Jun-Fei Qiao, and Weisi Lin, “Learning a no-reference quality assessment model of enhanced images with big data,” IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 4, pp. 1301–1313, 2018.

[13] Goran Gvozden, Sonja Grgic, and Mislav Grgic, “Blind image sharpness assessment based on local contrast map statistics,” Journal of Visual Communication and Image Representation, vol. 50, pp. 145–158, 2018.

[14] Jingwei Guan, Wei Zhang, Jason Gu, and Hongliang Ren, “No-reference blur assessment based on edge modeling,” Journal of Visual Communication and Image Representation, vol. 29, pp. 1–7, 2015.

[15] Anmin Liu, Weisi Lin, and Manish Narwaria, “Image quality assessment based on gradient similarity,” IEEE Transactions on Image Processing, vol. 21, no. 4, pp. 1500–1512, 2012.

[16] M. T. Subbotin, “On the law of frequency of error,” vol. 31, no. 2, pp. 197–206, 1923.

[17] Mahdi S. Hosseini and Konstantinos N. Plataniotis, “Derivative kernels: Numerics and applications,” IEEE Transactions on Image Processing, vol. 26, no. 10, pp. 4596–4611, Oct 2017.

[18] Mahdi S. Hosseini and Konstantinos N. Plataniotis, “Finite differences in forward and inverse imaging problems: Maxpol design,” SIAM Journal on Imaging Sciences, vol. 10, no. 4, pp. 1963–1996, 2017.

[19] Phong V. Vu and Damon M. Chandler, “A fast wavelet-based algorithm for global and local image sharpness estimation,” IEEE Signal Processing Letters, vol. 19, no. 7, pp. 423–426, 2012.

[20] Arthur Leclaire and Lionel Moisan, “No-reference image quality assessment and blind deblurring with sharpness metrics exploiting fourier phase information,” Journal of Mathematical Imaging and Vision, vol. 52, no. 1, pp. 145–172, 2015.

[21] Rania Hasan, Zhou Wang, and Magdy M. A. Salama, “Image sharpness assessment based on local phase coherence,” IEEE Transactions on Image Processing, vol. 22, no. 7, pp. 2798–2810, 2013.

[22] Dohyoung Lee and Konstantinos N. Plataniotis, “Toward a no-reference image quality assessment using statistics of perceptual color descriptors,” IEEE Transactions on Image Processing, vol. 25, no. 8, pp. 3875–3889, 2016.

[23] Qingbing Sang, Huixin Qi, Xiaojun Wu, and Alan C. Bovik, “No-reference image blur index based on singular value curve,” Journal of Visual Communication and Image Representation, vol. 25, no. 7, pp. 1625–1630, 2014.

[24] Ke Gu, Guangtao Zhai, Weisi Lin, Xiaokang Yang, and Wenjun Zhang, “No-reference image sharpness assessment in autoregressive parameter space,” IEEE Transactions on Image Processing, vol. 24, no. 10, pp. 3202–3211, 2015.

[25] Shaode Yu, Shihbin Wu, Lei Wang, Fan Jiang, Yaoqin Xie, and Leida Li, “A shallow convolutional neural network for blind image sharpness assessment,” PloS One, vol. 12, no. 5, pp. e0176632, 2017.

[26] Le Kang, Peng Ye, Yi Li, and David Doermann, “Convolutional neural networks for no-reference image quality assessment,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1733–1740.

[27] Peng Ye, Jayant Kumar, Le Kang, and David Doermann, “Unsupervised feature learning framework for no-reference image quality assessment,” in 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2012, pp. 1098–1105.

[28] Shaode Yu, Fan Jiang, Leida Li, and Yaoqin Xi, “Cnn-grnn for image sharpness assessment,” in Asian Conference on Computer Vision. Springer, 2016, pp. 50–61.

[29] Efthymios Mavridaki and Vasileios Megaritis, “No-reference blur assessment in natural images using fourier transform and spatial pyramids,” in 2014 IEEE International Conference on Image Processing (ICIP). IEEE, 2014, pp. 566–570.