Edge sparsity criterion for robust holographic autofocusing

YIBO ZHANG,1,2,3 HONGDA WANG,1,2,3 YICHEN WU,1,2,3 MIU TAMAMITSU,1,2,3 AND AYDOGAN OZCAN1,2,3,*1

1Electrical Engineering Department, University of California, Los Angeles, California 90095, USA
2Bioengineering Department, University of California, Los Angeles, California 90095, USA
3California NanoSystems Institute (CNSI), University of California, Los Angeles, California 90095, USA
*Department of Surgery, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA
Corresponding author: ozcan@ucla.edu

Received 10 July 2017; revised 29 August 2017; accepted 30 August 2017; posted 31 August 2017 (Doc. ID 302019); published 22 September 2017

Auto-focusing is essential to digital holographic imaging. Previously used autofocusing criteria exhibit challenges when applied to, e.g., connected objects with different optical properties. Furthermore, in some of the earlier autofocusing criteria, the polarity, i.e., whether to search for the peak or the valley as a function of depth, changes for different types of samples, which creates another challenge. Here, we propose a robust and accurate autofocusing criterion that is based on the edge sparsity of the complex optical wavefront, which we termed the “sparsity of the gradient” (SoG). We demonstrated the success of SoG by imaging a wide range of objects, including resolution test targets, stained and unstained Papanicolaou smears, stained tissue sections, and blood smears. © 2017 Optical Society of America

Numerous autofocusing criteria have been demonstrated for different holographic imaging applications and were successfully used to image various samples [4,8–14]. For example, Gillespie and King proposed to use self-entropy (SEN) as an autofocusing criterion in digital holography and successfully applied it to test images [8]. As an alternative, Liebling and Unser used the sparsity of the Fresnelet coefficients (FRES) as an autofocusing criterion and applied it for imaging resolution test targets [9]. Dubois et al. used the integrated amplitude (AMP) as an autofocusing criterion for pure amplitude and pure phase objects, with opposite polarities [10]. Later a modified version of AMP was also proposed using integrated high-pass-filtered amplitude (HPA), which was demonstrated to lead to a minimum for both amplitude and phase objects [11]. Langehanenberg et al. compared four different autofocusing criteria for imaging of pure-phase objects, including the summed weighted power spectrum (SPEC), variance (VAR), summed gradient (GRA), and summed Laplacian (LAP) to validate the applicability of SPEC and GRA on pure-phase objects [4]. Memmolo et al. proposed to use the contrast texture measure (i.e., the Tamura coefficient, TC) as an autofocusing criterion and demonstrated successful results based on off-axis holography using macroscopic objects [12]. Later, Memmolo et al. also used a sparsity measure, i.e., the Gini index (GI), as an autofocusing criterion to image microscopic objects, e.g., cells [13]. Recently, Lyu et al. used the axial magnitude differential (DIF) as another autofocusing criterion and discussed its applicability to amplitude-contrast and phase-contrast samples [14].

Despite the wide variety of available autofocusing criteria, the choice of the specific criterion is, unfortunately, still highly subjective depending on the object type or application of interest. Development of a robust and accurate autofocusing criterion for digital holography that works universally well for different types of samples is a challenging task for several reasons. First, the coherent diffraction of light in holography results in complex interference patterns, in addition to speckle- and multiple-reflection-related interference artifacts, instead of providing homogeneous smoothening as in the case of incoherent microscopy and photography. Therefore, the sharpness or the high-frequency spatial content of a holographic image does not necessarily diminish with defocusing. Moreover, without

Digital holography [1] enables computational refocusing within a 3D sample volume after the image acquisition. This unique advantage has made it a useful technique in various applications such as bio-fluid analysis [2], dynamic 3D tracking of sperms [3], monitoring of living cells [4], among many others. Recently, it was also shown that digital holography in a lensless in-line format can achieve high-throughput imaging of tissue slides and cell smears on a chip [5,6].

Digital refocusing and reconstruction of holograms rely on the precise knowledge of the defocus distance (i.e., the “z distance”) of the object of interest. A widely accepted autofocusing method involves digital refocusing of a hologram using various z distances, where a certain function, i.e., an autofocusing criterion, is evaluated based on this set of refocused images. The z distance that corresponds to the maximum (or sometimes the minimum) of this criterion is used as the focus distance of the object for that given hologram. A good autofocusing criterion should in general be unimodal over a wide range of z distances and be accurate for various types of samples [7].

OCIS codes: (090.1995) Digital holography; (110.0180) Microscopy; (090.2880) Holographic interferometry; (100.2960) Image analysis.

https://doi.org/10.1364/OL.42.003824
As opposed to most of the other autofocus criteria that operate on the amplitude channel of a refocused image, SoG operates directly on the complex refocused image, which makes it much more general and robust for amplitude contrast, phase contrast, or mixed object types. The gradient operator in the definition of SoG can sense the sharp changes in an image regardless of the specific amplitude and/or phase-contrast mechanism of the object of interest. As a consequence, SoG is expected to always reach its maximum at the correct focus distance regardless of the type of the object, which is another advantage over some of the previously reported autofocus criteria, which usually exhibit opposite polarity for amplitude-versus phase-contrast objects and an ambiguous polarity for mixed amplitude/phase-contrast objects.

To compare the performance of GoG and ToG against the state-of-the-art, we selected eight previously proposed autofocus criteria, including SEN, FRES, HPA, SPEC, GRA, GI, TC, and DIF. Note that VAR and LAP [4] criteria are analogous to TC and GRA [4,12], respectively; therefore, they were not duplicated in our comparison. We used five different types of quasi-planar samples to validate our method, including a USAF resolution test target, an unstained Pap smear, a stained Pap smear (Pap stain, ThinPrep preparation), a 4 μm thick H&E-stained lung tissue section, and a Wright-stained whole blood smear. Among these objects, the USAF test target can be considered as an amplitude-contrast object, the unstained Pap smear can be considered as a phase-contrast object, while the others can be considered as mixed amplitude- and phase-contrast objects. We validated that both GoG and ToG outperform the existing autofocus criteria, providing superior accuracy and robustness for different types of amplitude-contrast, phase-contrast, and mixed amplitude- and phase-contrast objects.

Moreover, all the digital autofocus experiments reported here were performed without any phase-retrieval using in-line holograms of spatially connected samples that violate the “weakly scattering” condition; stated differently, the refocused images were severely contaminated by the twin image artifact, making autofocus much more challenging. This demonstrates that GoG and ToG are robust and accurate even under noisy conditions and also illustrates their practical usefulness for in-line holography, where the knowledge of the z distances is usually a prerequisite for the subsequent phase retrieval step; therefore, autofocus on the non-phase-retrieved hologram is usually an inevitable initial step for image reconstruction [5,6]. We also tested GoG and ToG on various phase-retrieved holograms, where they also performed very well [15]. Therefore, we believe GoG and ToG (or SoG in general) also can be applied to off-axis holographic imaging.

The raw holograms corresponding to the test objects were captured using a lens-free in-line holography setup, with pixel super-resolution implemented on a chip [5], achieving an effective pixel size of ~0.37 μm at both the hologram and object reconstruction planes. To create the “ground truth” focus distance for each sample, we took the following two steps: first, multithread phase retrieval [5] using eight hologram heights was performed using manually found relative z distances with respect to the first hologram height. The relative z distances were found such that their refocused images looked most alike. Second, we used the phase-retrieved complex-valued hologram at the plane closest to the sample to perform manual refocusing to create the ground truth absolute focus distance based on visual

phase retrieval, the twin image artifact in in-line holography makes autofocus much more challenging. These explain some of the failures of directly applying autofocus criteria developed for incoherent imaging to holography. Second, different mechanisms of light–matter interaction result in different complex-valued transmission functions corresponding to the samples. For example, a phase object, such as a transparent and unlabeled cell, interacts with the illumination light mainly through perturbations to the optical phase of the transmitted wavefront, leaving its amplitude mostly unchanged. On the other hand, an amplitude object, such as a metal-coated resolution test target, mainly modulates the amplitude of the transmitted wave, whereas a stained cell or tissue section that is semitransparent would both modulate the phase and amplitude. These variations among different types of objects make some of the autofocus criteria that are based on, e.g., the sharpness of the features, contrast, sparsity or other properties of the amplitude channel of the complex field, less robust. Furthermore, many of these criteria change polarity for amplitude-contrast versus phase-contrast samples and fail or perform suboptimal for mixed amplitude- and phase-contrast samples [10,14]. Some of the previous work also explored applying a criterion on the phase channel of the complex field for autofocus [8], but these approaches suffer from the same limitations discussed earlier, besides facing phase-wrapping-related challenges.

Here we propose a robust autofocus criterion for digital holography that performs accurate depth focusing across a wide range of samples, including amplitude-only, phase-only, and mixed-object types. We termed this criterion “sparsity of the gradient” (SoG), which is based on the edge sparsity of an object’s in-focus image. SoG of a complex optical wavefront is defined as

\[
\text{SoG}(U) = S(\nabla U),
\]

where \( \nabla \) is the gradient operator, \(| \cdot | \) is the modulus operator, and \( S(\cdot) \) is a sparsity measure. \( |\nabla U| \) can be calculated for a complex-valued image as

\[
|\nabla U|_{ij}^2 = |U_{ij} - U_{ij-1}|^2 + |U_{ij} - U_{i-1,j}|^2.
\]

Under the SoG framework, we used two different sparsity measures, GI and TC, which we term “Gini of the gradient” (GoG) and “Tamura of the gradient” (ToG), respectively. One should note that TC has recently been proven to be a sparsity measure [15]. These two sparsity measures (GI and TC) were chosen to quantify the edge sparsity of a complex image because they exhibit advantages, such as invariance under scaling, compared with other sparsity measures [15,16].

GI is defined for a real and non-negative image \( C \) as

\[
\text{GI}(C) = 1 - 2 \sum_{k=1}^{N} \frac{a_k}{\text{sum}(C)} \left( 1 - \frac{k}{N} + \frac{0.5}{N} \right),
\]

where \( a_k \) is the \( k \)-th sorted entry of the image matrix, \( C \), in ascending order, \( k = 1, \ldots, N \), and \( \text{sum}(C) \) is the sum of all the entries of the image. TC is defined as

\[
\text{TC}(C) = \sqrt{\sigma(C)/\langle C \rangle},
\]

where \( \sigma(\cdot) \) is the standard deviation, and \( \langle \cdot \rangle \) is the mean.

The design of the SoG criterion for autofocus assumes that the images of natural objects are mostly composed of flat regions and sharp edges, which is a widely accepted assumption utilized in image compression, image denoising, etc. When a refocused complex optical wavefront corresponding to a physical object is in focus, sharp edges should be sparse, i.e., the sharpest transitions in the image should only take up a small fraction of the image’s total area.
If we consider the smaller error of the two polarities (max versus object) and the unstained Pap smear (a phase-contrast object), performance on the USAF test target (an amplitude-contrast resulting from human error in manual focusing). Stained Pap smear, which are comparable to the 1 errors (1.66 μm and 1.52 μm) occurred for the stained Pap smear sample. To ensure the visibility of the curves, we see that HPA, GRA, GI, SPEC, DIF, TC, and FRES each failed for at least one sample. GRA had >2 μm but <4 μm errors for all three samples, but the polarity changed among the objects, which might be due to the relative extent of amplitude-contrast and phase-contrast of each sample.

To gain more insight into the behavior of different autofocusing criteria, we took the stained Pap smear sample as an example and plotted the values of each criterion against z, as shown in Fig. 1. The SPEC, SEN, and DIF exhibit significant oscillations, which make it difficult to search for a peak or valley. GI and TC do not have a peak/valley at the actual z, but they both have a large slope around it. GRA has a peak and a valley on the opposite sides of the correct z value, and the curve is also slightly fuzzy. FRES and HPA each have a peak/valley

Based on this evaluation process, the autofocusing accuracies of 10 different criteria (GoG, ToG, HPA, GRA, SEN, GI, SPEC, DIF, TC, and FRES) were compared with the gold standard focus distances for different objects (Table 1). To overcome the ambiguity of polarity based on the sample type, we automatically searched for both the peaks ("max") and valleys ("min") for all the criteria, except for GoG, ToG, and FRES (only max) and HPA (only min). To better visualize the results, we color-coded Table 1 with a color map, representing the errors made in focus distances, ranging between 0 and 10 μm. In the same table, >10 μm errors are colored as gray, considered as “failure.” “NaN” represents a case where max/min criterion is found at the boundary of the search range. Evidently, GoG and ToG performed very well for all the sample types. Their largest errors (1.66 μm and 1.52 μm, respectively) occurred for the stained Pap smear, which are comparable to the 1 μm uncertainty resulting from human error in manual focusing.

For the other autofocusing criteria, first let us focus on their performance on the USAF test target (an amplitude-contrast object) and the unstained Pap smear (a phase-contrast object). If we consider the smaller error of the two polarities (max versus min) wherever applicable, HPA, GRA, GI, and TC had errors less than or around 1 μm. We also observed polarity inversion for GRA, GI, and TC when switched from the USAF test target to an unstained Pap smear sample, which also has been reported in the literature [13,14]. The other criteria, SEN, SPEC, DIF, and FRES, failed for at least one of the two samples (Table 1). As for the performance of these autofocusing criteria for the other three samples, including the stained Pap smear, lung tissue and blood smear samples, we see that HPA, SEN, GI, SPEC, DIF, TC, and FRES each failed for at least one sample. GRA had >2 μm but <4 μm errors for all three samples, but the polarity changed among the objects, which might be due to the relative extent of amplitude-contrast and phase-contrast of each sample.

### Table 1. Absolute Errors (in μm) of Various Autofocusing Criteria on Different Samples

<table>
<thead>
<tr>
<th>Method</th>
<th>GoG Max</th>
<th>ToG Max</th>
<th>HPA Max</th>
<th>GRA Max</th>
<th>SEN Max</th>
<th>GI Max</th>
<th>SPEC Max</th>
<th>DIF Max</th>
<th>TC Max</th>
<th>FRES Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>USAF target</strong></td>
<td>0.96</td>
<td>1.02</td>
<td>1.11</td>
<td>494.1</td>
<td>0.07</td>
<td>0.84</td>
<td>505.2</td>
<td>0.34</td>
<td>151.8</td>
<td>12.52</td>
</tr>
<tr>
<td></td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td><strong>Unstained Pap smear</strong></td>
<td>0.10</td>
<td>0.04</td>
<td>0.98</td>
<td>0.06</td>
<td>10.43</td>
<td>10.54</td>
<td>436.0</td>
<td>0.70</td>
<td>0.00</td>
<td>148.5</td>
</tr>
<tr>
<td></td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td><strong>Stained Pap smear</strong></td>
<td>1.66</td>
<td>1.52</td>
<td>0.75</td>
<td>3.14</td>
<td>9.27</td>
<td>NaN</td>
<td>79.0</td>
<td>68.66</td>
<td>130.9</td>
<td>12.32</td>
</tr>
<tr>
<td></td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td><strong>Lung tissue</strong></td>
<td>0.12</td>
<td>0.54</td>
<td>NaN</td>
<td>2.35</td>
<td>6.97</td>
<td>16.46</td>
<td>296.1</td>
<td>17.59</td>
<td>NaN</td>
<td>6.93</td>
</tr>
<tr>
<td></td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td><strong>Blood smear</strong></td>
<td>1.08</td>
<td>1.10</td>
<td>19.78</td>
<td>4.38</td>
<td>2.90</td>
<td>5.21</td>
<td>81.77</td>
<td>41.02</td>
<td>7.14</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
</tbody>
</table>

For all criteria except GoG, ToG, HPA, and FRES, both minimum and maximum were searched.
near the correct $z$ but are not unimodal. GoG and ToG, on the other hand, are overall smooth, unimodal, and each has a strong global peak close to the correct $z$ value, as desired.

Figure 2 further provides a comparison of the refocusing results based on our criteria. To enable better visual judgment, the phase-retrieved hologram of a lung tissue sample using eight sample heights was refocused, where each focus distance was estimated from the non-phase-retrieved hologram. Figures 2(e)–2(m) clearly shows that GoG and ToG provide the best in-focus images, which contain sharp features of the cells and the tissue morphology that agree well with a $20 \times 0.5$ NA microscope objective image.

Summarizing this comparison, all the criteria other than GoG, ToG, and GRA failed for at least one sample. The average errors for GoG, ToG, and GRA for all the samples were 0.78, 0.84, and 1.70 $\mu m$, respectively, where the smaller errors between min and max were considered for quantification of the error in GRA (to its advantage). Polarity inversion was observed among different sample types as a disadvantage of GRA, which may cause confusion when autofocusing an unknown sample. An additional challenge with any autofocusing criterion that has polarity inversion is that there can be a particular specimen that has a certain level of amplitude and phase contrast such that it is exactly at (or close to) the point of polarity inversion, which can create a singularity point and the correct $z$ cannot be determined.

Although GoG and ToG performed similarly for dense objects, we also noticed that, for naturally sparse samples of small size, such as dilute Giardia lamblia cysts and sperm cells, the user sometimes needs to choose a region of interest that tightly bounds the object of interest when using GoG. This constraint is relaxed when using ToG, leading to more flexibility. Mathematical origins of this different behavior of GoG and ToG for naturally sparse samples are discussed in detail in [15].

In addition to the autofocus criteria discussed thus far, we also tested the recently proposed complex ratio [18], which failed for some of the samples. We also tried cascading existing criteria onto the gradient of the complex refocused image, creating various new criteria such as AMP of the gradient, SEN of the gradient, SPEC of the gradient, etc. These also failed for some of the samples and therefore were not included in our comparison.

Computational efficiency is another important feature of an autofocus criterion that needs to be considered. We performed a comparison of the run times of different autofocusing criteria using a laptop computer equipped with an Intel Core i7-7700HQ CPU at 2.80 GHz and 16 GB of memory; GPU acceleration was not used. The total time required to refocus a hologram of $1024 \times 1024$ pixels and evaluate each autofocus criterion is summarized in Table 2. For this image size, GoG and ToG are $\sim 1.8$ and $\sim 1.4$ times slower, respectively, compared with the most computationally efficient focusing criterion tested here (GRA).

In conclusion, we demonstrated a robust and accurate holographic autofocus criterion that is based on the edge sparsity of the complex wavefront. Analysis showed that it outperforms existing autofocus approaches for a wide range of objects, at the cost of a modest increase in computation time.

Funding. Howard Hughes Medical Institute (HHMI); Army Research Office (ARO); National Science Foundation (NSF); Mary Kay Foundation (TMKF).

### REFERENCES


### Table 2. Comparison of Computation Times

<table>
<thead>
<tr>
<th>Method</th>
<th>GoG</th>
<th>ToG</th>
<th>HPA</th>
<th>GRA</th>
<th>SEN</th>
<th>SPEC</th>
<th>DIF</th>
<th>TC</th>
<th>FRES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (ms)</td>
<td>177</td>
<td>134</td>
<td>106</td>
<td>97</td>
<td>170</td>
<td>142</td>
<td>109</td>
<td>190</td>
<td>114</td>
</tr>
</tbody>
</table>