Probability of error as an image metric for the assessment of tomographic reconstruction of dense-layered binary-phase objects

Iksung Kang\textsuperscript{a} and George Barbastathis\textsuperscript{b,c}

\textsuperscript{a}Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA
\textsuperscript{b}Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA
\textsuperscript{c}Singapore-MIT Alliance for Research and Technology (SMART) Centre, Singapore 138602, Singapore

ABSTRACT

Each image metric represents different characteristics of images. For instance, similarity metrics, e.g., Structural Similarity Index Metric (SSIM) or Pearson Correlation Coefficient (PCC), utilize correlation between two images to calculate similarity between them; error metrics, e.g. Mean Absolute Error (MAE) or Root-Mean Squared Error (RMSE), compute the pixel-wise error between them according to different norms. As each of them highlights different aspects, a choice of the metric for an application may depend upon characteristics of the images. In this paper, we will show some tomographic reconstructions of dense-layered binary-phase objects, and as the objects are binary, we propose Probability of Error (PE) as an image metric for the assessment of the reconstructions in contrast to other metrics that are not constrained to the range of values. PE is equivalent to Bit-Error Rate (BER) in digital communications as both of the signals of interest are binary, and we are interested to a bit-wise deviation of the reconstructions of their corresponding ground truth images.

Keywords: Image metric, Probability of error, Binary object

1. INTRODUCTION

Inverse problems in optical imaging are often ill-posed due to a practical limitation of system aperture, detector noise, scattering, aberrations, etc. Many deep neural network (DNN) algorithms have dealt with the problems, e.g. phase retrieval\textsuperscript{1–3} under photon-limited conditions,\textsuperscript{4–6} limited-angle tomography,\textsuperscript{7,8} and ptychography,\textsuperscript{9,10} to get reconstructions with higher fidelity. Upon quantifying the quality of the reconstructions, certain image metrics have been adopted that effectively replace human perception, which can be only regarded as correct method to determine visual quality. Often, many have borrowed and used the concept of some image metrics to design their training loss functions for the optimization of deep learning architectures, e.g. negative Pearson Correlation Coefficient (NPCC)\textsuperscript{11} and cross-entropy.\textsuperscript{12}

Rather, the main purpose of the image metrics is to monitor and benchmark image processing applications,\textsuperscript{13} which also enables quantitative comparison between different applications. As each of the metrics highlights different aspects of the images, it should be carefully chosen upon the applications. For instance, Structural Similarity Index Metric (SSIM)\textsuperscript{13} and Pearson Correlation Coefficient (PCC)\textsuperscript{14} are representative correlative metrics. As they utilize the correlation for calculation, their values explain the statistical similarity between two images.

Further author information: (Send correspondence to Iksung Kang)
E-mail: iskang@mit.edu
Error metrics, e.g. Mean Absolute Error (MAE) and Root-Mean Squared Error (RMSE), are the simplest and most widely used quality metric. In image processing applications, they are computed as the pixel-wise error between two images, where the way to compute the error generally follows the convention of the mathematical norm. Thus, their outcomes indicate pixel-wise deviation of a reconstruction from its corresponding ground truth image.

There is a yet another widely used set of metrics, called probabilistic metrics, which include cross-entropy, Kullback-Liebler (KL) divergence, Jensen-Shannon (JS) divergence, mutual information (MI), and Wasserstein distance. These metrics involve intermediate computation of two different probability distributions of the images. Then, each of the metrics takes a different formulation to quantify the difference between distributions. Sometimes, a pre-trained neural network replaces conventional image metrics to better represent visual perception of natural images, i.e. perceptual loss metric.15 A large number of natural images in the ImageNet database is used for training a VGG network,17 and features of a certain layer of the network are extracted to compute the L2-norm as a feature loss.

Generally, image metrics listed above receive inputs with positive real numbers or often integers as a discretized form of digital images. However, here we propose an image metric that is specialized for binary objects, called Probability of Error (PE). The formulation of PE is equivalent to Bit-Error Rate (BER) in digital communications. We apply the image metric to binary-phase reconstructed objects from different inverse algorithms based on deep neural networks for limited-angle tomography. We also show a mathematical relationship of PE with Wasserstein distance, which is also known as Earth Mover’s Distance.

2. METHODS

We used three different deep neural networks for tomographic reconstruction of three-dimensional binary-phase objects in a limited angular view. Two of them are based on a static machine learning algorithm, and the other follows a dynamical machine learning approach. The first implementation (or Design 1) of the static algorithm is the same as the DNN used in Refs. 11,18, whose number of trainable parameters is 0.5 M, and hyperparameters are tuned for the second one (Design 2) to make it 21 M. The last neural network (Design 3) is different from the previous two as it is based on recurrent neural networks, whose implementation is the same as the one in Ref. 19, and this network also has 21 M of trainable parameters. These implementations are originally designed for a limited-angle tomographic reconstruction application.

All of the networks share the same experimental data from Ref. 18, where the optical apparatus used for experiments can also be found. Raw intensity measurements are pre-processed by Beam Propagation Method (BPM) and a gradient descent,18,20 but the loss function subject to be minimized is defined differently between the static and dynamic machine learning approaches. For Designs 1 and 2, the derivation of the gradient starts from the loss function18

\[
\mathcal{L} = \frac{1}{2} \sum_{n=1}^{N} \| H_n(f) - g_n \|_2^2,
\]

where \( H_n \) means a forward model defined with BPM for \( n \)-th angle of tomographic illumination, \( f \) a three-dimensional object to image, \( g_n \) an intensity measurement associated with \( n \)-th angular view, and \( N \) the number of the angles. Performing a gradient descent with the loss function, we get a three-dimensional object estimate (or Approximant) of the original object.

For Design 3, it is rather defined in a sequence to serve as input to the recurrent neural network as

\[
\mathcal{L}_n = \frac{1}{2} \| H_n(f) - g_n \|_2^2, \quad n = 1, 2, \cdots, N,
\]

where each trial using the loss functions generates a different Approximant of the original object, which is associated with each angular view, thus forming a length-\( N \) sequence.19

Proc. of SPIE Vol. 11653 116530T-2
Whereas the input to each design has different dimensions, the output is a volumetric object. In Refs. 18, 19, the objects consist of four layers, densely-placed.

3. RESULTS

In this section, we compare volumetric reconstructions of the object of interest from the different methods. We refer each of the methods to as Kang et al (21M) for the recurrent neural network approach as stated in Ref. 19, Goy et al (0.5 M) for the static neural network approach which was originally proposed in Ref. 18, and finally Goy et al (21 M) for the same approach but with more number of trainable parameters. Goy et al (21 M) was designed to set the computational complexity similar to our dynamical machine learning approach by setting the number of trainable parameters approximately the same.

Figure 1. Different volumetric reconstructions from three different reconstruction algorithms.
Table 1 lists values of image metrics of the volumetric reconstructions in Fig. 1. For the calculation of Pearson correlation coefficient and Wasserstein distance, we follow for PCC,

\[
\mathcal{L}_{\text{PCC}}(f, \hat{f}) = \frac{\sum_{x,y} (f(x,y) - \langle f \rangle)(\hat{f}(x,y) - \langle \hat{f} \rangle)}{\sqrt{\sum_{x,y} (f(x,y) - \langle f \rangle)^2} \sqrt{\sum_{x,y} (\hat{f}(x,y) - \langle \hat{f} \rangle)^2}},
\]

where \(f\) and \(\hat{f}\) are a ground truth object and its reconstruction, and for Wasserstein distance \((p = 1)\),

\[
W_{p=1} = \inf_{\gamma \sim \Pi} E_{(x,y) \sim \gamma} [|x - y|_1].
\]

Wasserstein distance means the minimum work required to transport the probability mass, where the cost is define as the 1-norm as above. For its application to 2D images, it can be re-formulated as

\[
W_{p=1} = \min_{(\gamma_{ij,kl})} \sum_{ij} \sum_{kl} \gamma_{ij,kl} |x_{ij} - x_{kl}|,
\]

where \(\sum_{kl} \gamma_{ij,kl} = f_{ij}, \sum_{ij} \gamma_{ij,kl} = \hat{f}_{kl}, \gamma_{ij,kl} \geq 0\).

As the object of interest is binary, Ref. 19 defines PE as follows.

\[
\text{PE} = \frac{\text{(# false negatives) + (# false positives)}}{\text{total # pixels}}.
\]

Here, the definition is the same as Bit-Error Rate or accuracy in digital communications. The use of PE for binary-phase objects is novel for a tomographic problem.\(^{19}\) We compute the metric with reconstructions, binarized by a thresholding method of which the threshold is adaptively determined according to the Maximum a Posteriori (MAP) decision rule. This way, even residual artifacts with intermediate values between two binary ends that occupy only small regions can be thresholded to the upper binary end, and thus they are dealt with more significance than as they would have been on their own without binarization. It has been also investigated that PE is closely related to the 2D Wasserstein distance, as an optimization problem on the optimal transport between two different distributions, but with a different cost tensor.\(^{19}\)

### Table 1. Reconstruction fidelity of each layer of the object according to Pearson correlation coefficient, Wasserstein distance, and probability of error.

|                  | Layer 1 | Layer 2 | Layer 3 | Layer 4 | Overall |
|------------------|---------|---------|---------|---------|---------|
| **Pearson**      |         |         |         |         |         |
| Kang et al (21 M)| 0.8580  | 0.6512  | 0.9539  | 0.6228  | 0.7715  |
| Goy et al (0.5 M)| 0.8453  | 0.5058  | 0.8751  | 0.5558  | 0.6955  |
| Goy et al (21 M)| 0.8487  | 0.5925  | 0.8917  | 0.6419  | 0.7437  |
| **Wasserstein**  |         |         |         |         |         |
| Kang et al (21 M)| 1.831   | 1.291   | 1.589   | 1.029   | 1.435   |
| Goy et al (0.5 M)| 2.994   | 2.177   | 2.142   | 1.486   | 2.200   |
| Goy et al (21 M)| 2.866   | 1.729   | 2.368   | 1.271   | 2.059   |
| **Probability**  |         |         |         |         |         |
| Kang et al (21 M)| 6.580   | 4.431   | 2.393   | 1.910   | 3.828   |
| Goy et al (0.5 M)| 7.233   | 5.969   | 6.677   | 2.399   | 5.569   |
| Goy et al (21 M)| 6.934   | 5.011   | 5.762   | 1.831   | 4.844   |

4. CONCLUSION

We have proposed and used a novel metric called Probability of Error (PE) for a tomographic problem of a pure-phase binary volumetric object. The computation involves binarization by an adaptive thresholding method,
because of which we believe even small residual artifacts are dealt with more significance than any other method without the binarization. This metric can be also interpreted as an optimal transport problem as it is closely linked to Wasserstein distance.

**ACKNOWLEDGMENTS**

Authors acknowledge support from IARPA (FA8650-17-C-9113). I. Kang acknowledges partial support from KFAS (Korea Foundation for Advanced Studies) scholarship.

**REFERENCES**

[1] Sinha, A., Lee, J., Li, S., and Barbastathis, G., “Lensless computational imaging through deep learning,” *Optica* 4(9), 1117–1125 (2017).

[2] Rivenson, Y., Zhang, Y., Günaydın, H., Teng, D., and Ozcan, A., “Phase recovery and holographic image reconstruction using deep learning in neural networks,” *Light: Science & Applications* 7(2), 17141–17141 (2018).

[3] Metzler, C., Schniter, P., Veeraraghavan, A., et al., “prDeep: robust phase retrieval with a flexible deep network,” in *International Conference on Machine Learning*, 3501–3510, PMLR (2018).

[4] Goy, A., Arthur, K., Li, S., and Barbastathis, G., “Low photon count phase retrieval using deep learning,” *Physical review letters* 121(24), 243902 (2018).

[5] Deng, M., Li, S., Goy, A., Kang, I., and Barbastathis, G., “Learning to synthesize: Robust phase retrieval at low photon counts,” *Light: Science & Applications* 9(1), 1–16 (2020).

[6] Kang, I., Zhang, F., and Barbastathis, G., “Phase extraction neural network (PhENN) with coherent modulation imaging (CMI) for phase retrieval at low photon counts,” *Optics Express* 28(15), 21578–21600 (2020).

[7] Zhang, H., Li, L., Qiao, K., Wang, L., Yan, B., Li, L., and Hu, G., “Image prediction for limited-angle tomography via deep learning with convolutional neural network,” *arXiv preprint arXiv:1607.08707* (2016).

[8] Würfl, T., Hoffmann, M., Christlein, V., Breininger, K., Huang, Y., Unberath, M., and Maier, A. K., “Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems,” *IEEE transactions on medical imaging* 37(6), 1454–1463 (2018).

[9] Nguyen, T., Xue, Y., Li, Y., Tian, L., and Nehmetallah, G., “Deep learning approach for Fourier ptychography microscopy,” *Optics express* 26(20), 26470–26484 (2018).

[10] Cherukara, M. J., Zhou, T., Nashed, Y., Enfedaque, P., Hexemer, A., Harder, R. J., and Holt, M. V., “AI-enabled high-resolution scanning coherent diffraction imaging,” *Applied Physics Letters* 117(4), 044103 (2020).

[11] Li, S., Deng, M., Lee, J., Sinha, A., and Barbastathis, G., “Imaging through glass diffusers using densely connected convolutional networks,” *Optica* 5(7), 803–813 (2018).

[12] Li, Y., Xue, Y., and Tian, L., “Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media,” *Optica* 5(10), 1181–1190 (2018).

[13] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P., “Image quality assessment: from error visibility to structural similarity,” *IEEE transactions on image processing* 13(4), 600–612 (2004).

[14] Pearson, K., “Notes on regression and inheritance in the case of two parents proceedings of the royal society of London, 58, 240–242,” (1895).

[15] Johnson, J., Alahi, A., and Fei-Fei, L., “Perceptual losses for real-time style transfer and super-resolution,” in *European conference on computer vision*, 694–711, Springer (2016).

[16] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L., “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, 248–255, Ieee (2009).

[17] Simonyan, K. and Zisserman, A., “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556* (2014).

[18] Goy, A., Rughoobur, G., Li, S., Arthur, K., Akinwande, A. I., and Barbastathis, G., “High-resolution limited-angle phase tomography of dense layered objects using deep neural networks,” *Proceedings of the National Academy of Sciences* 116(40), 19848–19856 (2019).
[19] Kang, I., Goy, A., and Barbastathis, G., “Limited-angle tomographic reconstruction of dense layered objects by dynamical machine learning,” arXiv preprint arXiv:2007.10734 (2020).

[20] Kamilov, U. S., Papadopoulos, I. N., Shoreh, M. H., Goy, A., Vonesch, C., Unser, M., and Psaltis, D., “Optical tomographic image reconstruction based on beam propagation and sparse regularization,” IEEE Transactions on Computational Imaging 2(1), 59–70 (2016).