NO-REFERENCE WEIGHTING FACTOR SELECTION FOR BIMODAL TOMOGRAPHY

Yan Guo and Bernd Rieger

Department of Imaging Physics, Delft University of Technology
E-mail: {y.guo-3, b.rieger}@tudelft.nl

ABSTRACT

Bimodal tomography introduces a weighting factor α to incorporate X-ray data into projection images acquired from scanning transmission electron microscope (STEM) for achieving an atom-specific three-dimensional (3D) reconstruction of an object on the nanoscale. Currently its value is chosen by computing reconstructions for a large range of $\alpha \in (0, 1)$ and comparing them to a hand-segmented ground truth with the mean square error (MSE). Since this is infeasible for an industrial application, in this paper we propose an image quality metric to quantify the quality of tomograms in terms of cross-atomic contamination and noise for selecting the weighting factor without a ground truth. Numerical results demonstrate that our framework can determine close-to-optimal weighting factor within an accuracy of ± 0.03. Moreover, approximating the shape of the minimum by a parabola effectively reduces the computational time by 90%.

***Index Terms*—** Image quality assessment, bimodal tomographic reconstruction

1. INTRODUCTION

Electron tomography (ET) is essential for studying specimens in materials science, as it reveals the 3D structure of an object from a series of its two-dimensional (2D) projections on the nanoscale [1]. In STEM, projections formed by a high-angle annular dark-field (HAADF) detector have high signal-to-noise ratio (SNR) but only contain aggregated information of all chemical elements along the projection direction [1]; projections obtained from energy dispersive X-ray spectroscopy (EDS) accomplish an atom-specific reconstruction but suffer from low SNR [2]. In order to simultaneously exploit these two complementary techniques, HAADF-EDS bimodal tomography (HEBT) proposed in [3] introduces a weighting factor α to link both modalities into one reconstruction. The choice of α depends on the noise level and influences the reconstruction result. However, there is no a priori way to determine the "best" value. In [3], the optimal α is found by computing reconstructions over the whole range of $\alpha \in (0, 1)$ and comparing them to a hand-segmented ground truth with the MSE. Since this is inapplicable for an industrial application, a quantitative quality control for reconstructions in the absence of a reference image is desired.

In recent years, no-reference image quality assessment has been widely investigated for different application scenarios [4] [5]. Proposed algorithms can be generally divided into two categories: (*i*) distortion-specific, that is, algorithms are designed specifically for one distortion. For instance, the framework presented in [6] uses Gabor filter to evaluate the streak (ringing) artifacts resulting from the iterative image restoration; (*ii*) non-distortion-specific, i.e. algorithms are generic and can respond to multiple degradations. Besides applications in computer vision, a lot of efforts have also been dedicated to developing assessment algorithms in the field of ET, such as evaluating the performance of tomographic reconstruction algorithms and/or the quality of tomograms. In [7], the length of phase boundary is treated as a quantitative morphological image characteristic to compare the commonly adopted filtered backprojection algorithm and the DIRECTT technique. In [8], Okariz *et al.* statistically analyze the intensity profiles at the edge of objects in the reconstructed volume to set the number of iterations used for the simultaneous iterative reconstruction technique.

In this paper, we propose an image quality metric to choose the close-to-optimal weighting factor α for HEBT by means of quantifying the reconstruction quality of a core-shell nanoparticle consisting of gold (Au) and silver (Ag). It can replace the MSE adopted in [3] if no ground truth is available. To begin with, Section 2 introduces the HEBT reconstruction technique and the methodology for deciding the optimal α with a hand-segmented ground truth. Related image quality assessment algorithms are briefly reviewed in Section 3 as prior work, followed by our proposed quality metric, and results presented in Section 4. Section 5 summarizes our work and discusses possible future extensions.

We use the following notations throughout this paper. Bold uppercase $\mathbf{W}$ and lowercase $\mathbf{w}$ represent matrices and column vectors, respectively, while non-bold letters W and w are scalars. Operators $(\cdot)^T$ and $*$ stand for transpose and convolution. $\mathbb{R}^{m \times n}$ denotes the space of all $m \times n$ matrices with real-valued elements.

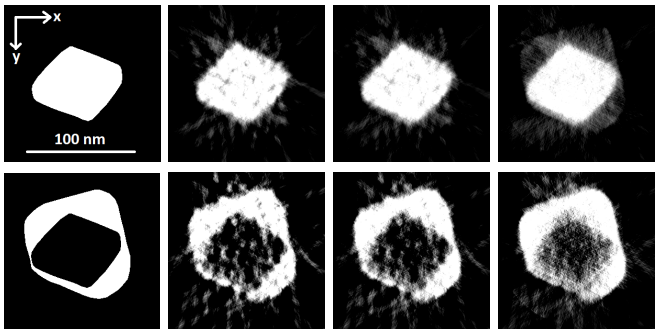
This work is partially supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organization for Scientific Reserarch, and partially by the Ministry of Economic Affairs, Agriculture and Innovation under project number 13314.

2. HAADF-EDS BIMODAL TOMOGRAPHY

Assume a specimen with E different chemical elements. Each element $e = 1, \cdots, E$ is associated with an unknown volumetric object $\mathbf{x}^{(e)} \in \mathbb{R}^{N\times 1}$, where N is the total number of equally-spaced voxels to be reconstructed. Let HAADF-STEM and EDS-STEM images be $\mathbf{p}^h \in \mathbb{R}^{M\times 1}$ and $\mathbf{p}^{(e)} \in \mathbb{R}^{M\times 1}$, respectively, where M is the total number of pixels in one projection image. In [3], HAADF-EDS bimodal tomographic reconstruction is defined as a least-square minimization problem

$$\begin{aligned}\mathbf{x}^* = \arg\min_{\mathbf{x}} \alpha^2 \left\| \mathbf{p}^h - \sum_{e=1}^{E} \mathbf{W}\mathbf{x}^{(e)} \right\|_2^2 \\ +(1-\alpha)^2 \sum_{e=1}^{E} \left\| r^{(e)}\mathbf{p}^{(e)} - \mathbf{W}\mathbf{x}^{(e)} \right\|_2^2\end{aligned} \tag{1}$$

in which $\mathbf{x} = [\mathbf{x}^{(1)T}, \cdots, \mathbf{x}^{(E)T}]^T$, and each entry w_{mn} in $\mathbf{W} \in \mathbb{R}^{M\times EN}$ is determined by the intersected area between the m-th ray integral and n-th voxel [9]. The response ratio factor $r^{(e)}$ for element e is calculated as $p_m^h = \sum_{e=1}^{E} r^{(e)} p_m^{(e)}$, $m = 1, \cdots, M$. Here, a weighting factor $\alpha \in (0, 1)$ is introduced to balance the residue terms of HAADF-STEM and EDS-STEM. In principle, α can be arbitrarily chosen between 0 and 1, whereas in practice it can neither be too small nor too large. The former makes the influence from HAADF-STEM hardly observable and the latter leads the minimization of EDS-STEM residue term to become inefficient.



(a) (b) (c) (d)

Fig. 1: Au (upper) and Ag (lower) images at slice 150. The size of the reconstruction volume is $300 \times 300 \times 300$. (a) ground truth; (b)-(d) HEBT reconstruction results with $\mathcal{N} = 100$ iterations and weighting factors $\alpha \in \{0.5, 0.7, 0.9\}$, respectively. For better visualization, we perform percentile contrast stretching from 0 to 87%.

We consider the same core-shell nanoparticle as in [3] that consists of Au (inner shell) and Ag (outer shell). Fig. 1 depicts HEBT reconstruction results for $\mathcal{N} = 100$ iterations and different weighting factors $\alpha \in \{0.5, 0.7, 0.9\}$ at slice 150 along the z-axis. The size of the reconstructed volume is $300 \times 300 \times 300$. Two binary images in the first column are the hand-segmented ground truth with homogeneous intensity. For $\alpha = 0.5$, $\mathcal{N} = 100$ introduces overfitting, that is, the least-square optimization fits to the noise rather than true patterns, and reconstructions in Fig. 1(b) are noisier with perceptible streaks showing up. In Fig. 1(d), $\alpha = 0.9$ is too large and hence Au leaks into the background of Ag reconstruction and vice versa. In order to find the "best" value of α beforehand, Zhong *et al.* compute the reconstruction $\mathbf{x}$ for a large range of $\alpha \in (0, 1)$ and compare it to the hand-segmented ground truth $\mathbf{x}_\mathtt{r}$ (Fig. 1(a)) via $\mathrm{MSE}(\mathbf{x}_\mathtt{r}, \mathbf{x}) = \min \|\mathbf{x}_\mathtt{r} - c\mathbf{x}\|_2^2$, where c is a scaling factor [3]. Since this is not feasible for an industrial application, an image quality metric to quantify the quality of reconstructions is desired, such that α can be determined without a hand-segmentation.

3. DETERMINE WEIGHTING FACTOR WITHOUT GROUND TRUTH

According to Fig. 1(a), ideal reconstructions of the core-shell nanoparticle should be binary with homogeneous foreground and zero-valued background. Inspired by the analysis of Fig. 1, we build our non-distortion-specific quality metric on assessing: (*i*) cross-atomic contamination, that is, how much Au is showing up in Ag regions and vice versa; (*ii*) inhomogeneity of the extracted fore- and background, and (*iii*) noise at the same time. In this section, we first present metrics that separately evaluate the aforementioned three, followed by our quality metric for choosing the close-to-optimal α for Au in the absence of its ground truth. The analysis of Ag follows the same principle.

3.1. Cross-atomic contamination metric $Q_{\mathtt{CC}}$

In order to measure the cross-atomic contamination, we first generate a binary mask $B_{\mathtt{Au}}$ for Au slice by slice based on the edge candidate points that are found in its volumetric reconstruction. In [10], edges are extracted by a scale-normalized differential entity $\mathcal{G}_\mathtt{g}^{\sigma_\mathtt{g}} = \sigma_\mathtt{g}(L_x^2 + L_y^2)$ with $L = f * g_\mathtt{g}(\cdot; \sigma_\mathtt{g})$, such that the scale at which an edge being detected can be automatically selected. Edge strength is defined as the gradient magnitude of a smoothed image L, which is obtained by convolving the input image $f(x, y)$ with a Gaussian kernel $g_\mathtt{g}(\cdot; \sigma_\mathtt{g})$ whose standard deviation is $\sigma_\mathtt{g}$. Finally, we calculate the cross-atomic contamination metric $Q_{\mathtt{CC}}$ by averaging the intensity of pixels outside the mask.

3.2. Inhomogeneity metrics $Q_{\mathtt{IH},1}$ and $Q_{\mathtt{IH},2}$

We evaluate the inhomogeneity of a non-ideal gray-scaled Au reconstruction by comparing it to its binary mask $B_{\mathtt{Au}}$. In [11], similarity between two images f_1 and f_2 is measured by

the Pearson coefficient

$$\mathrm{PC} = \frac{\sum_i (f_{1,i} - \bar{f}_1)(f_{2,i} - \bar{f}_2)}{\sqrt{\sum_i (f_{1,i} - \bar{f}_1)^2 \sum_i (f_{2,i} - \bar{f}_2)^2}} \tag{2}$$

where $f_{1,i}$ and $f_{2,i}$ are the intensity values of i-th pixel, $\bar{f}_1$ and $\bar{f}_2$ the average intensities over all pixels in f_1 and f_2, respectively. When $\bar{f}_1$ and $\bar{f}_2$ are not subtracted, a new coefficient, the so-called overlap coefficient

$$\mathrm{OC} = \frac{\sum_i f_{1,i} f_{2,i}}{\sqrt{\sum_i {f_{1,i}}^2 \sum_i {f_{2,i}}^2}} \tag{3}$$

is defined. We represent our two inhomogeneity metrics as $Q_{\mathtt{IH},1} = 1 - \mathrm{PC}$ and $Q_{\mathtt{IH},2} = 1 - \mathrm{OC}$, respectively.

3.3. Noise metrics $Q_{\mathtt{N},1}$ and $Q_{\mathtt{N},2}$

We investigate the noise level of Au reconstructions by computing the amount of streaks and oriented structures they contain. It is based on the previous work in [6] and [12].

In [6], streak artifacts are analyzed by a 2D Gabor filter, which can be regarded as modulating a Gaussian envelope by a sinusoidal wave with fixed frequency. Given a specific orientation θ, the corresponding Gabor response for an input image $f(x, y)$ is $\mathcal{G}^\theta = f * g(\cdot; \varphi, \gamma, \sigma, F_g, \theta)$, in which φ is the phase offset, γ and σ the spatial aspect ratio and standard deviation of the Gaussian envelope, F_g and θ the central frequency and orientation of the Gabor filter, respectively. Methodology proposed in [6] works as follows: decompose $f(x, y)$ using Gabor filter w.r.t. different orientations to obtain $\mathcal{G}^\theta(u, v)$; for each row i (or column j) in $\mathcal{G}^\theta$, calculate the maximum oscillation strength S_i (or S_j), which is defined as the response difference between the local maximum and its neighboring local minimum; compute the overall metric value for streak artifacts by finding the maximum oscillation strength S among all rows and columns and averaging over all orientations θ. Since we do not have a priori knowledge of the width of streaks, we further extend the original module to a filter bank based version. Its design involves two important parameters: F_g and θ. The former is determined by the central frequency of the filter at the highest frequency (F_M), the ratio between two neighboring central frequencies (F_r) and the number of frequencies (N_F), and the latter by the number of orientations (N_θ). We calculate our first noise metric $Q_{\mathtt{N},1}$ by modifying the algorithm proposed in [6] as follows: oscillation strength calculation and maximum value extraction are performed not only over all rows and columns but also over all frequency bands.

In [12], oriented structures are extracted by a Gaussian profile with orientation selectivity. A linear orientation space for a specific angle ϕ is defined as $\mathcal{H}^\phi = f * h(\cdot; N_h, F_h, B_h, \phi)$ where $h(\cdot; N_h, F_h, B_h, \phi)$ is obtained by rotating the orientation selective template filter $h(\cdot; N_h, F_h, B_h)$ over ϕ. N_h, which relates to the orientation selectivity, is the number of filters along the ϕ-axis, and F_h and B_h the central frequency and bandwidth of the Gaussian profile, respectively. After constructing the orientation space, we further find the maximum response over ϕ and denote it as our second noise metric $Q_{\mathtt{N},2}$.

3.4. Proposed method: A combination of individuals

For finding a proper combination for the aforementioned individual metrics, we first evaluate their own properties for different $\alpha \in [0.1, 0.9]$ with a step size of 0.01 at slice 150. Table 1 lists all important parameters, for choosing which we follow the guideline in [6] [12] [13] without fine tuning.

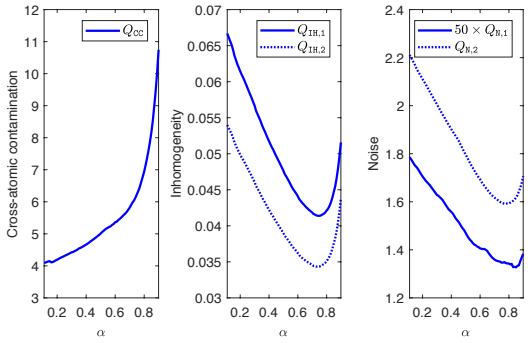
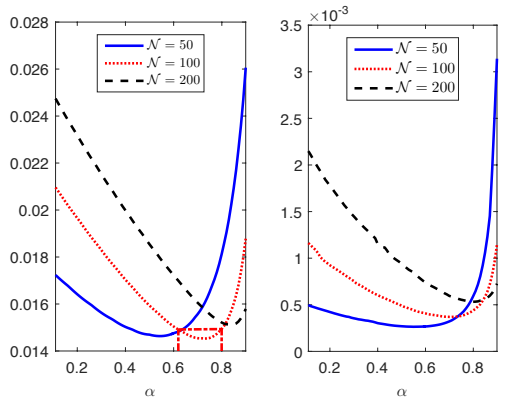


Fig. 2: Metric values of cross-atomic contamination, inhomogeneity and noise versus weighting factor α for Au with 100 iterations adopted for HEBT at slice 150.

As illustrated in Fig. 2, the background of Au reconstruction gets more contamination from Ag when α is increasing as it increases the ratio of HAADF-STEM term that contains aggregated information. Moreover, inhomogeneity and noise metrics have a clear unique minimum. We define our quality metric Q as the product of all individual metrics, namely $Q = Q_{\mathtt{CC}} \times Q_{\mathtt{IH},1} \times Q_{\mathtt{IH},2} \times Q_{\mathtt{N},1} \times Q_{\mathtt{N},2}$. Note that we do not normalize the individual metrics to $[0, 1]$, otherwise the minimum of each curve at zero would automatically dictate the minima of the multiplication. Fig. 3 depicts MSE and the combinational quality metric Q for Au w.r.t. different weighting factor α and number of iterations $\mathcal{N}$, in which Fig. 3(a) is the same as Fig. 7 in [3]. It can be observed that there is a relatively large range of $\alpha (\sim 0.18)$ within an uncertainty of ± 0.03, see red dash-dot lines in Fig. 3(a). Although parabolic curves in Fig. 3(b) are slightly different from the ones in Fig. 3(a), the optimal values of α around the minima of parabolas are almost the same.

4. RESULTS

Since HEBT only takes one α value as the input but reconstructs both Au and Ag as the output, we consider them si-



(a) MSE (b) Quality metric Q

Fig. 3: MSE and quality metric Q versus weighting factor α with different number of iterations $\mathcal{N}$ for Au at slice 150.

multaneously in this section by summing up their MSE (or Q) values and finding the minima. Note that it takes around two hours to generate one parabolic curve in Fig. 3(b), during which CPU time is mainly occupied for noise analysis. Therefore, we choose 8 equidistant samples for α from 80 points in total and perform polynomial fitting to predict parabolas and reduce the computational time by 90%.

Table 1: Parameters for reconstruction quality assessment

Ref.	Parameter	Symbol	Value
[6]	Phase offset	φ	0
	Spatial aspect ratio	γ	0.5
[13]	Central frequency of filter at the highest frequency	F_M	$\sqrt{2}/4$
	Frequency ratio	F_r	$\sqrt{2}$
	Number of orientations	N_θ	8
	Number of frequencies	N_F	3
[12]	Number of filters	N_h	33
	Central frequency of Gaussian profile	F_h	0.15
	Bandwidth of Gaussian profile	B_h	$0.5F_h$

Fig. 4 depicts the optimal values of α for different number of iterations $\mathcal{N}$ at slice 150, which are found by MSE, true and predicted quality metric Q, respectively. It demonstrates that besides α obtained from MSE, the other two also tend to increase with the increment of $\mathcal{N}$. This is because for large $\mathcal{N}$, large α guarantees that it converges to the true pattern rather than noise. Moreover, α from both true and predicted quality metric values achieve an uncertainty of ± 0.03 independent from the number of iterations $\mathcal{N}$ adopted for HEBT.

Fig. 5 illustrates the consistency of the optimal α w.r.t.

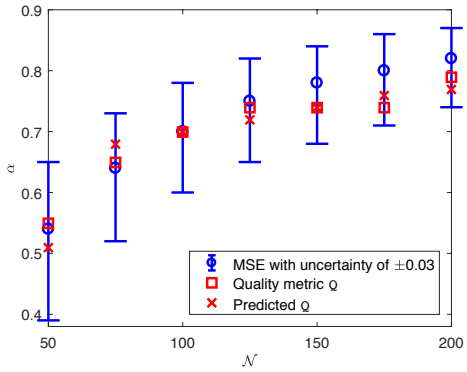


Fig. 4: Weighting factor α versus number of iterations $\mathcal{N}$ adopted for HEBT at slice 150.

different slices while Au and Ag are being considered simultaneously. Note that our quality metric Q is the closest to MSE at slice 150 because it is in the middle of the reconstruction stack ($300 \times 300 \times 300$) and thus suffers the least from boundary artifacts. However, even in the worst case where Q being the furthest to MSE, i.e. slices 80 and 170, α calculated and/or predicted by our quality metric still achieves an accuracy of ± 0.03.

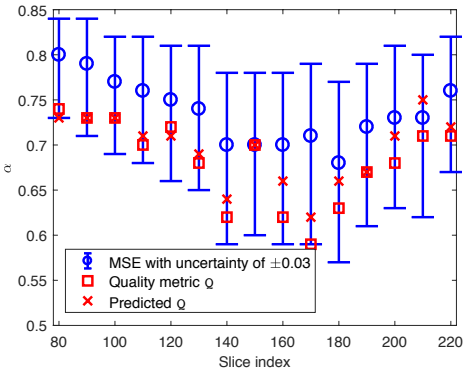


Fig. 5: Weighting factor α versus slice index with 100 iterations adopted for HEBT.

5. CONCLUSION

In this paper, we propose a no-reference quality metric for HEBT to automatically determine its weighting factor α by quantitatively evaluating the quality of tomograms. Furthermore, approximating the parabola by polynomial fitting reduces the computational time to 10%, which makes our quality metric more promising. As for the future work, we consider embedding the proposed assessment module into a learning system, such that α can be chosen in real-time without the need to perform reconstruction.

6. REFERENCES

[1] P.A. Midgley and M. Weyland, "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, pp. 413–431, 2003.

[2] T.J.A. Slater et al., "STEM-EDX tomography of bimetallic nanoparticles: a methodological investigation," *Ultramicroscopy*, vol. 162, pp. 61–73, 2016.

[3] Z. Zhong et al., "A bimodal tomographic reconstruction technique combining EDS-STEM and HAADF-STEM," *Ultramicroscopy*, vol. 174, pp. 35–45, 2017.

[4] V. Kamble and K.M. Bhurchandi, "No-reference image quality assessment algorithms: a survey," *Optik*, vol. 126, pp. 1090–1097, 2015.

[5] M. Shahid et al., "No-reference image and video quality assessment: a classification and review of recent approaches," *EURASIP Journal on Image and Video Processing*, vol. 40, 2014.

[6] B. Zuo, J. Tian, and D. Ming, "A no-reference ringing metrics for images deconvolution," in *Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition*, 2008, vol. 1, pp. 96–101.

[7] S. Lück et al., "Statistical analysis of tomographic reconstruction algorithms by morphological image characteristics," *Image Analysis and Stereology*, vol. 29, pp. 61–77, 2010.

[8] A. Okariz, T. Guraya, M. Iturrondobeitia, and J. Ibarretxe, "A methodology for finding the optimal iteration number of the SIRT algorithm for quantitative electron tomography," *Ultramicroscopy*, vol. 173, pp. 36–46, 2017.

[9] A.C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.

[10] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *Int. J. of Computer Vision*, vol. 30, no. 2, 1998.

[11] E.M.M. Manders, F.J. Verbeek, and J.A. Aten, "Measurement of co-localization of objects in dual-color confocal images," *Journal of Microscopy*, vol. 169, pp. 375–382, Mar. 1993.

[12] M. van Ginkel, *Image analysis using orientation space based on steerable filters*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2002.

[13] F. Bianconi and A. Fernandez, "Evaluation of the effects of Gabor filter parameters on texture classification," *Pattern Recognition*, vol. 40, pp. 3325–3335, 2007.