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 $\alpha$ 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 $\alpha$ to link both modalities into one reconstruction. The choice of $\alpha$ 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 $\alpha$ 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 $\alpha$ 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 $\alpha$ 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 $w$ 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 Research, 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 $e = 1, \ldots, E$ is associated with an unknown volumetric object $x^{(e)} \in \mathbb{R}^{N \times 1}$, where $N$ is the total number of equally-spaced voxels to be reconstructed. Let HAADF-STEM and EDS-STEM images be $p^h \in \mathbb{R}^{M \times 1}$ and $p^{(e)} \in \mathbb{R}^{M \times 1}$, 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

$$x^* = \arg \min_x \alpha^2 \left\| p^h - \sum_{e=1}^{E} Wx^{(e)} \right\|^2_2 + (1-\alpha)^2 \sum_{e=1}^{E} \left\| r^{(e)} p^{(e)} - Wx^{(e)} \right\|^2_2$$

(1)

in which $x = [x^{(1)}^T, \ldots, x^{(E)}^T]^T$, and each entry $w_{mn}$ in $W \in \mathbb{R}^{M \times EN}$ 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 $p^h_m = \sum_{e=1}^{E} r^{(e)} p^{(e)}_m$, $m = 1, \ldots, M$. Here, a weighting factor $\alpha \in (0, 1)$ is introduced to balance the residue terms of HAADF-STEM and EDS-STEM. In principle, $\alpha$ 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.

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 foreground 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 $\alpha$ for Au in the absence of its ground truth. The analysis of Ag follows the same principle.

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

In order to measure the cross-atomic contamination, we first generate a binary mask $B_{\text{Au}}$ 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 $G_{\nabla}^* = \sigma_{E}(L_2^2 + L_2^2)$ with $L = f * g_{\nabla}(:, \sigma_{E})$, 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 $g_{\nabla}(:, \sigma_{E})$ whose standard deviation is $\sigma_{E}$. Finally, we calculate the cross-atomic contamination metric $Q_{CC}$ by averaging the intensity of pixels outside the mask.

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

We evaluate the inhomogeneity of a non-ideal gray-scaled Au reconstruction by comparing it to its binary mask $B_{\text{Au}}$. In [11], similarity between two images $f_1$ and $f_2$ is measured by
the Pearson coefficient
\[
PC = \frac{\sum_i (f_{1,i} - \bar{f}_1)(f_{2,i} - \bar{f}_2)}{\sqrt{\sum_i (f_{1,i} - \bar{f}_1)^2 \sum_i (f_{2,i} - \bar{f}_2)^2}}
\]
(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 \(f_1\) and \(f_2\) are not subtracted, a new coefficient, the so-called overlap coefficient
\[
OC = \frac{\sum_i f_{1,i}f_{2,i}}{\sqrt{\sum_i f_{1,i}^2 \sum_i f_{2,i}^2}}
\]
(3)
is defined. We represent our two inhomogeneity metrics as \(Q_{\text{IH,1}} = 1 - PC\) and \(Q_{\text{IH,2}} = 1 - OC\), respectively.

3.3. Noise metrics \(Q_{\text{IH,1}}\) and \(Q_{\text{IH,2}}\)

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 \(\theta\), the corresponding Gabor response for an input image \(f(x, y)\) is \(G^\theta = f * g(\cdot; \varphi, \gamma, \sigma, F_g, \theta)\), in which \(\varphi\) is the phase offset, \(\gamma\) and \(\sigma\) the spatial aspect ratio and standard deviation of the Gaussian envelope, \(F_g\) and \(\theta\) 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 \(G^\theta(u, v)\); for each row \(i\) (or column \(j\)) in \(G^\theta\), 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 \(\theta\). 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 \(\theta\). The former is determined by the central frequency of the filter at the highest frequency \(F^\theta\), the ratio between two neighboring central frequencies \(F_{\pm}\) and the number of frequencies \(N_F\), and the latter by the number of orientations \(N_\theta\). We calculate our first noise metric \(Q_{\text{IH,1}}\) 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 \(\phi\) is defined as \(H^\phi = f * h(\cdot; N_h, F_h, B_h, \phi)\) 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 \(\phi\). \(N_h\), which relates to the orientation selectivity, is the number of filters along the \(\phi\)-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 \(\phi\) and denote it as our second noise metric \(Q_{\text{IH,2}}\).

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 \(\alpha\) for Au with 100 iterations adopted for HEBT at slice 150.](image)

As illustrated in Fig. 2, the background of Au reconstruction gets more contamination from Ag when \(\alpha\) 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 \(Q = Q_{\text{CC}} \times Q_{\text{IH,1}} \times Q_{\text{IH,2}} \times Q_{\text{H}}\). 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 \(\alpha\) and number of iterations \(N\), 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 \(\pm 0.03\), 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 \(\alpha\) around the minima of parabolas are almost the same.

4. RESULTS

Since HEBT only takes one \(\alpha\) value as the input but reconstructs both Au and Ag as the output, we consider them si-
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 \( \alpha \) 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**

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>[6]</td>
<td>Phase offset</td>
<td>( \varphi )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Spatial aspect ratio</td>
<td>( \gamma )</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Central frequency of filter at the highest frequency</td>
<td>( F_M )</td>
<td>( \sqrt{2}/4 )</td>
</tr>
<tr>
<td>[13]</td>
<td>Frequency ratio</td>
<td>( F_r )</td>
<td>( \sqrt{2} )</td>
</tr>
<tr>
<td></td>
<td>Number of orientations</td>
<td>( N_\theta )</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Number of frequencies</td>
<td>( N_F )</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Number of filters</td>
<td>( N_h )</td>
<td>33</td>
</tr>
<tr>
<td>[12]</td>
<td>Central frequency of Gaussian profile</td>
<td>( F_h )</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Bandwidth of Gaussian profile</td>
<td>( B_h )</td>
<td>0.5( F_h )</td>
</tr>
</tbody>
</table>

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

Fig. 5 illustrates the consistency of the optimal \( \alpha \) w.r.t.

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, \( \alpha \) calculated and/or predicted by our quality metric still achieves an accuracy of \( \pm 0.03 \).

**5. CONCLUSION**

In this paper, we propose a no-reference quality metric for HEBT to automatically determine its weighting factor \( \alpha \) 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 \( \alpha \) can be chosen in real-time without the need to perform reconstruction.
6. REFERENCES


