Visualizing the Segmentation Error of a Tomogram using the Residual Projection Error

Tom Roelandts, Kees Joost Batenburg and Jan Sijbers

***Abstract*—Tomographic reconstructions are often segmented to quantify structure parameters. However, to our knowledge, no efforts have yet been made to visualize the accuracy of the segmentation result. This paper introduces a way to visualize the segmentation error, based on the residual projection error, which is the difference between the recorded data and the forward projection of the segmented tomogram. From the residual projection error, a segmentation error tomogram is reconstructed. This error tomogram allows to detect errors in the gray levels of the segmented tomogram, or to discriminate between reconstruction artifacts and actual features of the scanned object. The proposed technique is independent of the algorithms that were used to create and segment the tomogram.**

***Index Terms*—Image segmentation, error reconstruction.**

I. Introduction

In many applications of tomography, the final tomographic reconstruction (the *tomogram*) must be segmented before the results can be analyzed. Segmentation amounts to the classification of image pixels into distinct classes, based on similarity with respect to some characteristic. Image segmentation is a well established field, and a range of methods has been developed, using diverse techniques such as global or local thresholding, region growing, and clustering [1], [2].

Most image segmentation methods are not specific towards the modality that was used to acquire the image. As a result, such methods do not exploit the raw data (X-ray radiographs in case of CT imaging) from which the image was reconstructed. Recently, global and local thresholding methods were proposed that do use the projection data to improve the selection of threshold parameters [3], [4]. In addition, reconstruction methods were recently developed in which the segmentation was directly incorporated into the reconstruction algorithm [5]–[8].

In this paper, the projection data is used to examine the *quality* of the segmentation. The proposed method assumes that the scanned object consists of homogeneous regions, and then determines how well a given segmentation represents that object. Moreover, the method can also be used to test for object homogeneity, since it generates large errors for non-homogeneous objects. The technique is independent of the reconstruction and segmentation algorithms.

This work was financially supported by the FWO (The Research Foundation - Flanders, Belgium) and by the NWO (the Netherlands Organisation for Scientific Research - The Netherlands, research programme 639.072.005).

Tom Roelandts, Kees Joost Batenburg and Jan Sijbers are with the IBBT-Vision Lab, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Belgium.

Kees Joost Batenburg is with the Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands.

Corresponding author: Tom Roelandts, E-mail: tom.roelandts@ua.ac.be.

The structure of this paper is as follows. In Section II, the method is introduced. Section III describes the simulation experiments that were performed to validate it. The results are discussed and a conclusion is reached in Section IV.

II. Method

The projection process in tomography can be modeled as a linear operator that is determined by the projection geometry. This leads to a system of linear equations,

$$\boldsymbol{W}\boldsymbol{x} = \boldsymbol{p}, \tag{1}$$

where $\boldsymbol{p} \in \mathbb{R}^m$ contains the projection data and $\boldsymbol{x} \in \mathbb{R}^n$ corresponds to the image. The linear operator is represented by the $m \times n$ matrix $\boldsymbol{W}$, the *projection matrix*. An approximate solution $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ of (1) can then be computed, in practice often by minimizing some norm $\|\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p}\|$. The image $\hat{\boldsymbol{x}}$ can then be segmented.

A segmentation method essentially partitions the pixels of an image into sets $Y_1, \ldots, Y_d$, where d is the number of classes (or gray levels) in the segmented image. Since we assume that the scanned object consists of homogeneous regions, the segmented image should have the same gray level for all pixels in a set Y_k. From $\hat{\boldsymbol{x}}$, we create a segmented image $\boldsymbol{s} \in \mathbb{R}^n$ by assigning a gray level $\rho_k \in \mathbb{R}$ to all pixels in a set Y_k, for each $k \in \{1, \ldots, d\}$. The values of $\rho_1, \ldots, \rho_d$ are not known a priori. Moreover, most of the segmentation algorithms that are commonly used, do not have these gray levels as an output. In such case, we use the mean of all pixel values in a class as an estimate,

$$\rho_k = \frac{1}{|Y_k|} \sum_{\hat{y} \in Y_k} \hat{y}, \text{ for each } k \in \{1, \ldots, d\}. \tag{2}$$

We now define the residual projection error. The segmented image $\boldsymbol{s}$ is forward projected to give $\boldsymbol{p_s} \in \mathbb{R}^m$, so $\boldsymbol{p_s} = \boldsymbol{W}\boldsymbol{s}$. The *residual projection error* $\boldsymbol{e} \in \mathbb{R}^m$ is then defined as

$$\boldsymbol{e} = \boldsymbol{p} - \boldsymbol{p_s}. \tag{3}$$

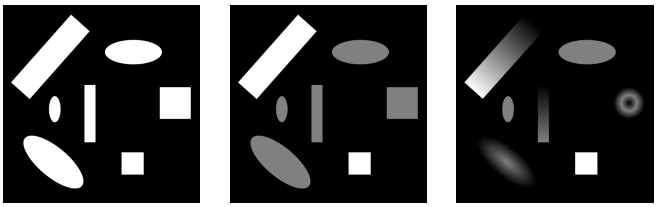
This projection error is then reconstructed by solving the system $\boldsymbol{W}\boldsymbol{y} = \boldsymbol{e}$, where $\boldsymbol{y} \in \mathbb{R}^n$ corresponds to the (unknown) error image. This results in an approximate solution $\hat{\boldsymbol{y}} \in \mathbb{R}^n$, the *reconstructed residual error*.

As mentioned in the introduction, this procedure does not depend on any particular reconstruction or segmentation algorithm. To create the segmented reconstructions, we use filtered backprojection (FBP), with a Ram-Lak filter, as an example of an analytical method. We use the simultaneous iterative reconstruction technique (SIRT) [9] as an example of an

iterative method. The reconstructions are globally thresholded using Otsu's method [10]. The residual projection error is reconstructed using SIRT.

III. Simulation Experiments

All simulation experiments were performed on a square reconstruction grid of 512×512 pixels. Three *phantom images* were created, as shown in Fig. 1. The size of each phantom is 2048×2048 pixels. The phantoms are much larger than the reconstruction grid, to reduce the effect of the pixelation on the reconstructions.



(a) Phantom 1 (b) Phantom 2 (c) Phantom 3

Fig. 1: Phantom images: (a) two gray levels, (b) three gray levels, (c) continuous grayscale.

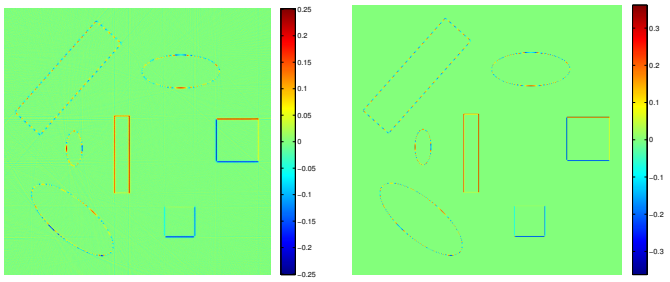
A. Optimally Segmented Reconstructions

In the first simulation experiment, the residual projection error was reconstructed for a dataset with no noise and a large number of projections. This demonstrates the result of the proposed technique with minimal hindrance from reconstruction artifacts. From Phantom 1, which is a binary phantom, a synthetic dataset was created using 360 parallel beam projections, evenly spaced at 0.5° intervals. A detector with 512 pixels was used, to simulate the practical situation where the detector pixel size equals the width of the reconstruction grid. The "ideal" segmented reconstruction was approximated by binning Phantom 1 to a 512×512 grid, and then segmenting it using a threshold of 0.5.

The "ideal" segmented reconstruction was then forward projected, and the difference with the original projections computed. This results in the residual projection error, which was then reconstructed using 300 iterations of SIRT. Fig. 2a shows the result. The true error (Fig. 2b) was computed as the difference between the original phantom (Fig. 1a) and an upscaled version of the ideal segmented reconstruction. For comparison, the true error is shown downscaled to the same resolution as Fig. 2a. The true error is only nonzero at the edges of the different structures, where errors are unavoidable, since the original phantom is not pixelated at the size of the reconstruction. The reconstructed residual error (Fig. 2a) is very close to the true error.

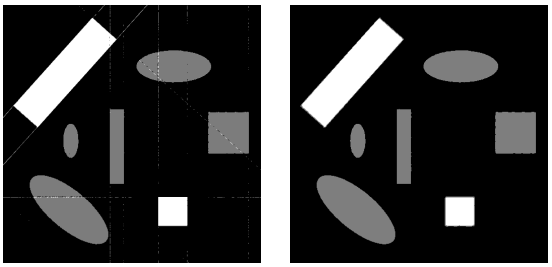
B. Reconstructions with Homogeneous Regions

The second simulation experiment is based on Phantom 2 (Fig. 1b), which contains three gray levels. Reconstructions of Phantom 2 should consist of homogeneous regions, since the ground truth object is divided into several such regions (compare with Phantom 3 (Fig. 1c), where this is clearly not the case). From Phantom 2, a synthetic dataset was created using 90 parallel beam projections, evenly spaced at 2° intervals. A detector with 512 pixels was again used. This synthetic dataset was then reconstructed twice, once using FBP and once using 300 iterations of SIRT. These reconstructions were then segmented using Otsu's method. The final segmented reconstructions are shown in Fig. 3. Due to the relatively low number of projection angles, a number of streaks are visible, especially in the FBP reconstruction of Fig. 3a.



(a) (b)

Fig. 2: (a) Reconstructed residual error for Phantom 1. (b) True error for Phantom 1, at the resolution of Fig. 2a.

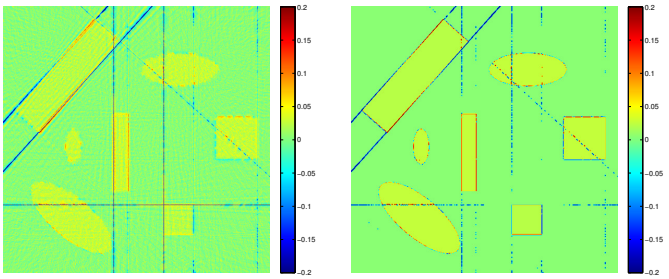


(a) FBP (b) SIRT

Fig. 3: Reconstructions of Phantom 2, segmented using Otsu's method.

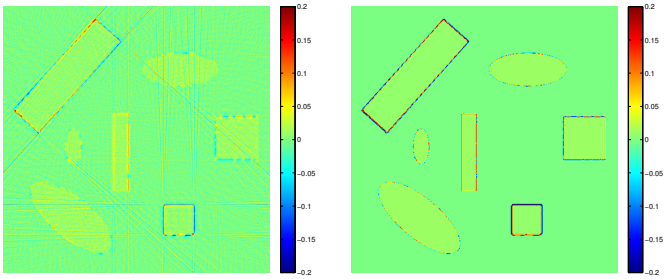
The segmented reconstructions were then forward projected, and the difference with the original projections computed, resulting in the residual projection error. This was then reconstructed using 300 iterations of SIRT. Figs. 4a (FBP) and 5a (SIRT) show the result. For comparison, the true error is also shown, in Figs. 4b and 5b. As before, there seems to be a close correspondence between the reconstructed error and the true error. The streaks from the FBP reconstruction are clearly visible in the reconstructed residual error (Fig. 4a). Hence, the error tomogram can be used to discriminate between this type of artifact and actual features of the scanned object. Apart from these streaks, the largest errors are again situated at the edges of the different structures. Figs. 4a and 5a further suggest that the gray levels of the different objects are underestimated, since the error is clearly positive and relatively uniform inside the objects. This is also confirmed by the true error.

Table I shows this quantitatively. It contains an overview of a number of parameters that were determined from the reconstructions, while the true gray levels were taken from Phantom 2. The computed gray levels were determined from the segmented reconstructions (Fig. 3), from which they were



(a) (b)

Fig. 4: (a) Reconstructed residual error for the FBP reconstruction of Phantom 2. (b) True error for that reconstruction, at the resolution of Fig. 4a. The scales have been limited to $[-0.2, 0.2]$ to keep the smaller errors from disappearing.



(a) (b)

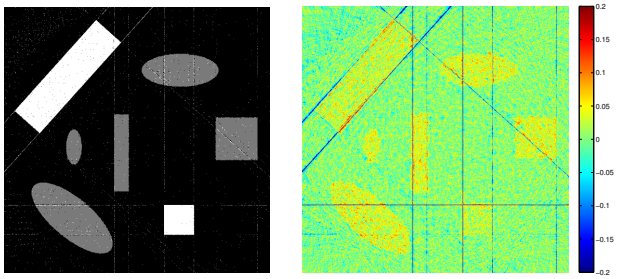
Fig. 5: (a) Reconstructed residual error for the SIRT reconstruction of Phantom 2. (b) True error for that reconstruction, at the resolution of Fig. 5a. The scales have been limited to $[-0.2, 0.2]$ to keep the smaller errors from disappearing.

computed using (2). The gray level errors were computed from Figs. 4a and 5a, using the same procedure. To verify the accuracy of the gray level errors, the last lines for each algorithm in Table I show the corrected gray levels, i.e., the sum of the computed gray levels and the gray level errors. The last lines show that the gray level error is a good estimate of the difference between the true gray level and the computed gray level, and that it can be used to correct the computed gray level. This result could potentially be improved even further by recreating the reconstructed residual error, starting with a segmented reconstruction that uses the corrected gray levels.

TABLE I: Estimated and true gray levels, noiseless dataset.

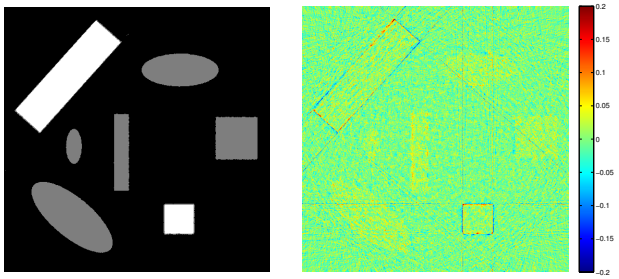
Algorithm	Parameter	ρ_1	ρ_2	ρ_3
FBP	True gray level	0.000	0.502	1.000
	Computed gray level	−0.004	0.476	0.979
	Gray level error	0.003	0.018	0.017
	Corrected gray level	−0.001	0.494	0.996
SIRT	True gray level	0.000	0.502	1.000
	Computed gray level	0.002	0.491	0.994
	Gray level error	−0.002	0.010	0.009
	Corrected gray level	−0.000	0.502	1.002

We also ran these experiments with Poisson noise applied to the synthetic dataset. The results are shown in Figs. 6 and 7, and in Table II. The quality of the FBP reconstruction (Fig. 6a) has suffered more from the noise than that of the SIRT reconstruction (Fig. 7a). In the reconstructed residual error (Figs. 6b and 7b), the noise from the original projections is apparent. The results in Table II, however, are still largely comparable with those from Table I. This means that the gray level errors in Figs. 6b and 7b, which seem drowned by the noise, are still quite accurate when averaged. This implies that we can decide, from the reconstructions of the residual error in Figs. 6b and 7b, that the segmentation from Fig. 7a is more accurate than that from Fig. 6a.



(a) (b)

Fig. 6: (a) FBP Reconstruction of Phantom 2, segmented using Otsu's method. Poisson noise was applied to the synthetic dataset. (b) Reconstructed residual error. The scale has been limited to $[-0.2, 0.2]$ to make it the same as in Fig. 4.



(a) (b)

Fig. 7: (a) SIRT Reconstruction of Phantom 2, segmented using Otsu's method. Poisson noise was applied to the synthetic dataset. (b) Reconstructed residual error. The scale has been limited to $[-0.2, 0.2]$ to make it the same as in Fig. 5.

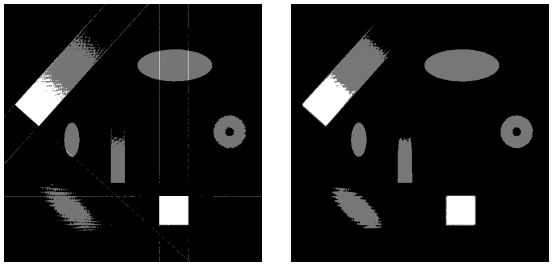
TABLE II: Estimated and true gray levels, dataset with noise.

Algorithm	Parameter	ρ_1	ρ_2	ρ_3
FBP	True gray level	0.000	0.502	1.000
	Computed gray level	−0.006	0.467	0.979
	Gray level error	0.002	0.022	0.020
	Corrected gray level	−0.004	0.489	0.999
SIRT	True gray level	0.000	0.502	1.000
	Computed gray level	0.002	0.491	0.994
	Gray level error	−0.002	0.010	0.009
	Corrected gray level	−0.000	0.502	1.003

C. *Reconstructions with Non-Homogeneous Regions*

The last simulation experiment is based on Phantom 3 (Fig. 1c), in which some of the homogeneous objects of

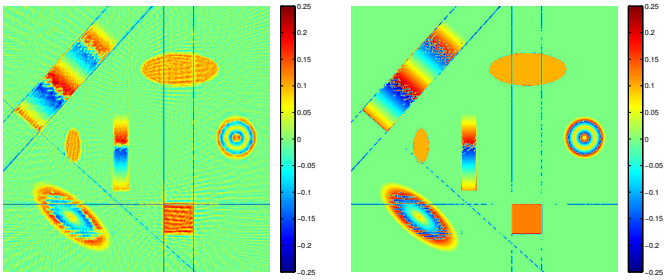
Phantom 2 are replaced by objects with a continuously varying intensity. Reconstructions of Phantom 3 are not expected to consist of homogeneous regions, since the ground truth object is not a collection of such regions. From Phantom 3, a synthetic dataset was again created using 90 parallel beam projections, evenly spaced at 2° intervals. This synthetic dataset was reconstructed using FBP and SIRT, and segmented using Otsu's method. The result is shown in Fig. 8.



(a) FBP (b) SIRT

Fig. 8: Reconstructions of Phantom 3, segmented using Otsu's method.

As before, the residual projection error was computed and reconstructed using 300 iterations of SIRT. The result is shown in Figs. 9a (FBP) and 10a (SIRT). Both figures show that the reconstructed residual error is useful to discriminate between objects that can be reconstructed as a homogeneous region (the homogeneous objects from Phantom 3), and those that cannot (the objects with continuously varying intensity, which have large structured errors in Fig. 9a and 10a). The reconstructed residual error is again quite close to the true error (Figs. 9b and 10b).



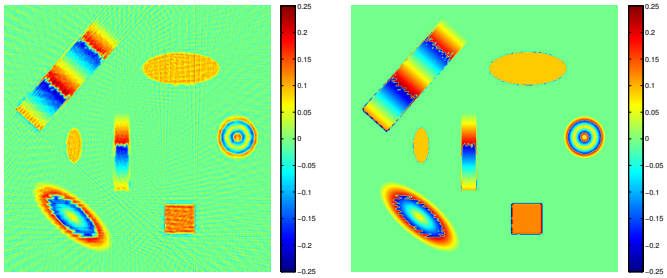
(a) (b)

Fig. 9: (a) Reconstructed residual error for the FBP reconstruction of Phantom 3. (b) True error for that reconstruction, at the resolution of Fig. 9a.

IV. Discussion and Conclusion

If tomograms are segmented without exploiting the projection data, which is often the case, the reconstructed residual error can still provide information on the quality of the segmentation.

The result for the ideal segmented reconstruction from Section III-A shows that the reconstructed residual error is a good approximation of the true error. For objects that consist of homogeneous regions, the experiments from Section III-B show that the proposed method can detect errors in the gray



(a) (b)

Fig. 10: (a) Reconstructed residual error for the SIRT reconstruction of Phantom 3. (b) True error for that reconstruction, at the resolution of Fig. 10a.

levels of the segmented reconstruction, even if those are quite small (Fig. 5), and in the presence of noise. The experiment from Section III-C shows that the technique can discriminate between homogeneous and non-homogeneous objects, without this knowledge being available a priori. Several experiments also show that artifacts that are visible in the segmented reconstruction, are also visible in the reconstructed residual error (e.g., Figs. 4a and 9a). Together, the simulation experiments demonstrate that the reconstructed residual error can be used as a visual map of the errors in the segmentation.

Reconstructing the residual projection error is a simple way to visualize the segmentation error of a tomogram. The technique is trivial to implement, since the only necessary tools are a forward projector and a reconstruction algorithm. The computational cost is modest, since only a single forward projection and a single reconstruction are needed.

References

[1] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recogn.*, no. 9, pp. 1277–1294, 1993.

[2] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annu. Rev. Biomed. Eng.*, pp. 315–337, 2000.

[3] K. J. Batenburg and J. Sijbers, "Optimal threshold selection for tomogram segmentation by projection distance minimization," *IEEE Trans. Med. Imag.*, vol. 28, no. 5, pp. 676–686, 2009.

[4] W. van Aarle, K. J. Batenburg, and J. Sijbers, "Optimal threshold selection for segmentation of dense homogeneous objects in tomographic reconstructions," *IEEE Trans. Med. Imag.*, vol. 30, no. 4, pp. 980–989, 2011.

[5] K. J. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. A. Midgley, J. C. Hernandez, U. Kaiser, E. R. Encina, E. A. Coronado, and G. Van Tendeloo, "3D imaging of nanomaterials by discrete tomography," *Ultramicroscopy*, vol. 109, no. 6, pp. 730–740, 2009.

[6] K. J. Batenburg and J. Sijbers, "DART: A practical reconstruction algorithm for discrete tomography," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2542–2553, 2011.

[7] T. Roelandts, K. J. Batenburg, and J. Sijbers, "PDART: A partially discrete algorithm for the reconstruction of dense particles," in *Proc. of the 11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Fully 3D)*, Potsdam, Germany, 2011.

[8] T. Roelandts, K. J. Batenburg, E. Biermans, C. Kübel, S. Bals, and J. Sijbers, "Accurate segmentation of dense nanoparticles by partially discrete electron tomography," *Ultramicroscopy*, In Press, doi:10.1016/j.ultramic.2011.12.003.

[9] J. Gregor and T. Benson, "Computational analysis and improvement of SIRT," *IEEE Trans. Med. Imag.*, vol. 27, no. 7, pp. 918–924, 2008.

[10] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.