Assessing the efficacy of tomographic reconstruction methods through physical quantification techniques

OPEN ACCESS
IOP Publishing
Measurement Science and Technology
Meas. Sci. Technol. **32** (2021) 075404 (15pp)
https://doi.org/10.1088/1361-6501/abe337

Assessing the efficacy of tomographic reconstruction methods through physical quantification techniques

Sophia B Coban[1,2], **William R B Lionheart**[2] and **Philip J Withers**[3,4]

[1] Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, Netherlands
[2] Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom
[3] Henry Royce Institute, Henry Royce Institute Hub Building, Manchester M13 9PL, United Kingdom
[4] Department of Materials, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

E-mail: sophia.coban@cwi.nl

Received 4 September 2020, revised 25 January 2021
Accepted for publication 4 February 2021
Published 4 May 2021

Abstract

Conventionally the performance of computed tomography (CT) reconstruction algorithms is assessed by a voxel-to-voxel comparison between the object and the reconstructed volume, often using digital phantoms. However the real aim in the CT imaging community is not to develop a reconstruction algorithm to obtain the best-looking images, but one that allows us to extract the relevant information to a desired accuracy. Here, through various case studies, we quantify features of interest for the test object and use these as measures of the efficacy of the reconstructions. Where applicable, we compare the assessment technique against commonly used metrics to measure the quality of a reconstructed solution, and find that in most cases the popular metrics have no relation to the accuracy of the features we extract from a reconstruction. The assessment technique we demonstrate, which we refer to as physical quantification, is used to determine the shape, contacts and size of beads for a test dataset made available via the SophiaBeads Dataset Project. Using this image analysis approach a number of widely used reconstruction methods are evaluated. Our work shows that it is important to choose the optimal reconstruction strategy based on the features you want to quantify from the scan. For example, in our case we found that the shape of the beads could be measured using TV regularization with eight times fewer projections than the other methods, or that reconstructions obtained via many but noisy projections yield as accurate results as those obtained via less noisy but fewer projections.

Supplementary material for this article is available online

Keywords: computed tomography, reconstruction methods, image processing and segmentation, iterative reconstruction

(Some figures may appear in colour only in the online journal)

1. Introduction

As the name suggests, computed tomography (CT) is reliant on reconstruction algorithms to compute a faithful image of the object. Today the vast majority of applications where x-ray CT is employed (such as in medical, laboratory, security

or synchrotron settings) the images are reconstructed using filtered back-projection algorithms. In principle there are a large number of iterative reconstruction (IR) methods that could be used to achieve better reconstructions in specific circumstances, for example in a dynamic or limited angle experimental setup, or in electron tomography where the collected data is often noisy and has missing wedge of projection angles, one can obtain quantitative results with IR methods coupled with appropriate regularization. However it is difficult to know which method is best suited for a specific imaging problem. Indeed it is not even clear how to quantify the effectiveness of a particular method to reconstruct a circular scan cone beam CT data (typically in a laboratory, dental or industry CT system). For many, a high quality reconstruction means a noiseless image with clear contrast, a solution that satisfies a 'best trade-off' in convergence, and is close to the approximate, but robust and widely-used, Feldkamp–Davis–Kress (FDK) filtered back-projection algorithm [1]. Accuracy or quality metrics in image processing (signal-to-noise, contrast-to-noise) and in numerical analysis (l_2-norm error in the residual or consecutive iterations) are commonly used to assess the quality of reconstructed images and thus the suitability of the iterative methods. However, these are poor indicators of image quality since an iterative reconstructed solution would have spatially-dependent noise, and resolution dependent on image contrast [2].

In addition, FDK and IR methods produce different artefacts (artificial features that do not correspond to a physical feature of the object [3]) in a reconstructed image, and thus comparing these solutions with each other may not offer a fair comparison. As argued in [4], the assessment of image quality is challenged by these traditional techniques, and an established measure of merit for evaluating the quality of reconstruction; independent of the noise and the contrast, or its similarity to an approximate FDK solution, is necessary. An alternative, task-based method is introduced in [4] however this is designed for clinical studies, whereas our aim here is to establish a more general measure of quality for tomographic imaging.

The real aim in the CT imaging community is not to develop a reconstruction algorithm to obtain the best-looking images, but one that allows us to extract the relevant information from a reconstruction to a desired accuracy. The topic of physical accuracy in a tomographic reconstruction is often overlooked. It is demonstrated, to some degree, in studies [5–7], where the physical information about the object is known and is employed to assess an image reconstruction or segmentation method. However none of the cited work makes tomographic data available to test any other algorithms or validate their results.

In this paper, we examine the strategy of extracting physical information from a reconstructed volume in order to evaluate its quality, and thus the suitability of reconstruction methods. We suggest a series of appropriate physical assessment techniques, which we refer to as 'physical quantification'. We note that these assessment techniques are not universal, but specific to the object scanned. However the idea of extracting physical information is still applicable, and should be considered as a valid option to assess the efficacy of tomographic reconstruction methods. In this paper, we define physical quantification techniques for a specific set of CT data, where some of the characteristics of the scanned object are known. This data, collected for the purpose of developing and testing image reconstruction techniques, is summarised briefly here in section 2 having been first introduced in [8]. Similar strategies are often used in medical, geophysical and non-destructive material studies to build or verify a physical model. Numerous examples are given in [9–11], which experiment with the idea of using features in a CT reconstruction to extract information relevant to the study. In the same spirit, we consider four problems to examine the use of physical quantification techniques, and how the results could differ from traditional ways of assessing the suitability and performance of reconstruction methods. These are typical problems in the imaging community, and are included to offer perspectives relevant to both mathematicians and experimental scientists working in the field. The results of these case studies are presented in section 5, and their brief descriptions are as follows.

The first study compares three important types of reconstruction methods. These are the filtered back projection (FDK, an approximate direct method); the conjugate gradient least squares (CGLS, an iterative method with no explicit prior information) [12], and the total variation image reconstruction (henceforth TV Prior, an iterative method with strong prior information) [13]. The study focuses on making fair comparisons of these methods while physically quantifying the accuracy of the reconstructed results.

The second study looks at the convergence of CGLS in the absence of prior information, and highlights how the conclusions we can draw from the traditional and our technique of assessing convergence may differ.

The third study looks at the effect of varying the quality of the data acquired (this could be the exposure or dose) on the quality of reconstructed solutions. This involves studying the affect of increased noise as the amount of exposure (or dose) is reduced, posing the question as to whether there are any advantages to higher exposure scans.

The final study compares novel algorithms against more standard segmentation methods, in which we explore the idea of simultaneously reconstructing and segmenting images with the means of a 2-step discrete-CGLS algorithm (DCGLS, our modification to the discrete algebraic reconstruction technique (DART) algorithm) [14]. We compare a variation of the level set evolution formulation (the distance regularized level set evolution or DRLSE) [15], here adapted as a segmentation method, with simpler techniques such as basic, Otsu's [16] and locally adaptive [17] thresholding methods.

We finish this paper with more discussions and further analysis on the first case study along with some concluding remarks.

2. Establishing a reference CT dataset

2.1. Description of the test data

The SophiaBeads Dataset Project [18] is a collection of cone-beam x-ray CT data sets where the number of projections are varied while the total photon count (or the total exposure time) per acquired data set is kept constant [8]. These data sets enable a wide range of algorithm comparisons and information content optimizations to be examined. They are also suitable for testing segmentation methods since the sample comprises just two phases (glass and air). The project also contains scripts to read and pre-process the data prior to reconstruction, [19], which are essential in order to use the data set.

The data sets are available to download individually, with the largest one containing 2048 projections. The number of projections is reduced by a factor of 2 for the subsequent data sets, with the smallest one containing 64 projections. For the case studies in section 5, we made use of all the data sets available.

2.2. Scanner and sample details

The SophiaBeads data sets were collected using the 320/225 kV Nikon XTEK Bay [20] located at the Henry Moseley X-ray Imaging Facility, the University of Manchester.

The scanner consists of a cone-beam microfocus x-ray source that projects polychromatic x-rays onto a 2000×2000 pixel, 16-bit, flat detector panel. The optimal window size for the SophiaBeads reconstructions is 1564×1564, which was determined prior to the data acquisition, see [19]. The data sets were collected using the stop/start acquisition technique to reduce angular blurring. The source settings were 70 kV, 210 μA with a 0.25 mm copper filter attached to the front of the source to reduce any possible beam hardening artefacts. Both the source and the geometry settings were kept the same across all collected data. Annotated photographs of the setup and the mount are given in supplementary material appendix C (figure C3) (available online at stacks.iop.org/MST/32/075404/mmedia).

The beadpack sample comprised a 25 mm diameter (20 mm internal diameter) plastic tube, filled with homogeneous Soda-Lime Glass (SiO_2–Na_2O) beads[5] in the size range 2.3–2.6 mm, with a mean value of 2.5 mm. The beads in the plastic tube are packed tightly and do not move during data acquisition.

A pack of glass beads were chosen because their simple morphology lends itself to analysis yet is representative of many physical problems in non-destructive testing, such as a porous media [21–23] and particles undergoing granular flow [24, 25]. Glass was preferred to metal in order to reduce any possible beam hardening effects (glass provides good contrast to air while not as attenuating as metal).

3. Quantification strategies

3.1. Physical measures of the bead pack data

Our motivation for using the SophiaBeads data sets in particular is that the main features of the sample are well characterised: We know the sizes of the beads so we can physically quantify the quality of reconstructions. In section 5 where we solve typical questions in imaging, we quantify our results using physical measures, which we define below:

(a) **Shape of the beads**—Shape3D or Shape2D: this quantifies how close (in shape) a reconstructed bead is to a perfect sphere. This is defined by the ratio of the surface area of the sphere[6] to the surface area of the reconstructed bead. Similarly in 2D, this is given by the ratio of the area of the disk[6] to the area of the bead.
(b) **Aspect ratio of the beads**—AspectRatio3D or AspectRatio2D: The ratio of the smallest and the largest (orthogonal) width of the bead in 2D (circle) and 3D (sphere).
(c) **Sphericity of the beads**: The percentage of the reconstructed bead that fits to a perfect sphere, similar to Shape3D.

Such measures are often used to quantify particle shapes with the shape and aspect ratio functions available within the image-visualization and analysis software, Avizo Fire (versions 7–9) [26]. The sphericity measure is not native to Avizo and we implemented the measure independently, prior to the quantification stage. The sphericity percentage is defined by

$$\text{Sphericity} = 100\frac{\pi^{1/3}(6V)^{2/3}}{S}, \tag{1}$$

where V is the volume and S is the surface area of a bead [27, 28].

Unfortunately the actual beads imaged are neither monodisperse, nor perfectly spherical. This means that we have no unequivocal perfect physical measure to assess the reconstructed bead shapes. In order to establish a physical reference that is representative of the sample, we evaluated the mean and the standard deviation of the physical measures for the full 2048-projection FDK reconstruction (size: $1564 \times 1564 \times 2000$). These reference values, given in table 1, are based on 500 reconstructed and segmented beads, and are thus representative of the sample characteristics. For the planar (2D) quantification measures, we used the central horizontal slice (i.e. Slice 1000; size: 1564×1564) to obtain the reference values. A summary of this process is explained in supplementary material appendix C, and in greater detail in [29].

Although not employed here, one can also use the volume and the Gaussian or mean curvature of individual beads as well as the contact between two identical spheres to determine the accuracy of a reconstruction. These were not used as quantification measures since they were employed in the process of

[5] Lead-free Soda-Lime glass beads distributed by BioSpec Products, Inc. PO Box 788, Bartlesville, OK 74005, USA, www.biospec.com.

[6] The sphere (disk in 2D) is approximated based on the largest height or width of the reconstructed bead.

Table 1. Listing of the idealized ground truth of perfectly spherical beads and the reference values for the given physical measures obtained from the full 2048-projection FDK reconstruction.

		Reference physical measures			
Quantification measure	Idealized ground truth	Minimum	Maximum	Mean	Standard deviation
Shape3D	1	1.0832	1.9776	1.3274	0.00226
Shape2D	1	1.1248	2.3614	1.2143	0.00750
AspectRatio3D	1	0.7142	0.9719	0.9112	0.00172
AspectRatio2D	1	0.7948	0.9937	0.9319	0.00320
Sphericity	100	79.177	97.372	87.164	0.0132

separating and filtering the reconstructed beads in Avizo [29]. It should be noted that such physical measures could have been calculated just as easily using open-source image analysis software such as Blob3D [30], Fiji/ImageJ [31], Drishti [32] or VTK [33].

3.2. Traditional measures of reconstructed image quality

In addition to physical measures, we include traditional, pixel-based image quality assessment techniques (where applicable). These are standard measures used in mathematics to show the grey value differences between two images (the $\mathcal{L}_2$ error (in the image) relative to the ground truth image (the object), $\mathcal{E}_{2\text{norm}}$), or to observe changes in consecutive solutions or in estimated and measured data, at each iteration (iterative error relative to the current iterate, $\mathcal{E}_{\text{iter}}$ and the residual error relative to the measured data, $\mathcal{E}_{\text{res}}$, respectively). These are defined by

$$\mathcal{E}_{2\text{norm}} = \frac{\|x^* - x\|_2}{\|x^*\|_2}, \tag{2}$$

$$\mathcal{E}_{\text{iter}} = \frac{\|x_k - x_{k-1}\|_2}{\|x_k\|_2}, \tag{3}$$

$$\mathcal{E}_{\text{res}} = \frac{\|b - b_k\|_2}{\|b\|_2}, \tag{4}$$

where $\|\cdot\|_2$ denotes the (Euclidean) 2-norm, x^* the ground truth image; b the log of the measured data, and b_k the estimated data by the kth iterated solution x_k, defined by $b_k = Ax_k$. In cases where equation (2) is used, the 2048-projection reconstruction (of size $1564 \times 1564 \times 500$) is taken as the 'ground truth image'. We also consider the contrast-to-noise ratio (CNR) as an image quality assessment technique [34] (in dB), defined by

$$\text{CNR} = 10\log_{10}\frac{\mu_F - \mu_B}{\sqrt{\sigma_F^2 + \sigma_B^2}}, \tag{5}$$

where μ_F, μ_B are the mean grey values; σ_F and σ_B standard deviation in grey values in selected regions of foreground (F) and background (B). These regions are shown with blue rectangles in figure 1.

4. Reconstruction algorithms

4.1. Implementation notes

The IR methods considered in the case studies are implemented in MATLAB 2014b, with the forward- and back-projectors written in C [19]. Readers will note that the CGLS is used in all subsections. Other methods used in the case studies are FDK with a ramp filter (results obtained using the in-house software, CTPro); a formulation of total variation as a regularizer with a maximum *a posteriori* probability (MAP) estimate (TV Prior), and our modified version of the DART with segmentation methods such as the basic thresholding, Otsu's thresholding, locally adaptive thresholding and a variation of the level set evolution methods (DRLSE). All formulation and algorithm implementation details are summarised in supplementary material appendix A. All computations were run on machines with 3.10 GHz, 64-bit 16-core CPU (2x 8-core Intel Xeon E5-2687 W) using a single GPU with 12 GB RAM (Nvidia Quadro K6000).

4.2. Presentation of results

For the 3D results, we reconstruct the central subvolume of size $1564 \times 1564 \times 500$, and display the images of the central horizontal slice, (i.e. the Slice 1000). We focus on a central window of size 440 px $\times$ 440 px; smaller regions of size 100 px $\times$ 100 px of a bead in the centre, or 130 px $\times$ 130 px near the edge. These are shown in figure 1, henceforth referred to as ROI1, ROI2 and ROI3 throughout section 5, respectively.

The quantification plots are based on mean physical quantification values of 25 reconstructed beads. Details are given in the supplementary material appendix C.

5. Case studies

5.1. Comparing reconstruction methods using physical measures

Traditionally the efficacy of a reconstruction method is determined by voxel-to-voxel comparison. However in practice, the goal of a tomographic study is to extract physical information from a reconstructed volume. Thus we will use a physical quantification method to assess the suitability and the efficacy of three inherently different, yet commonly used and

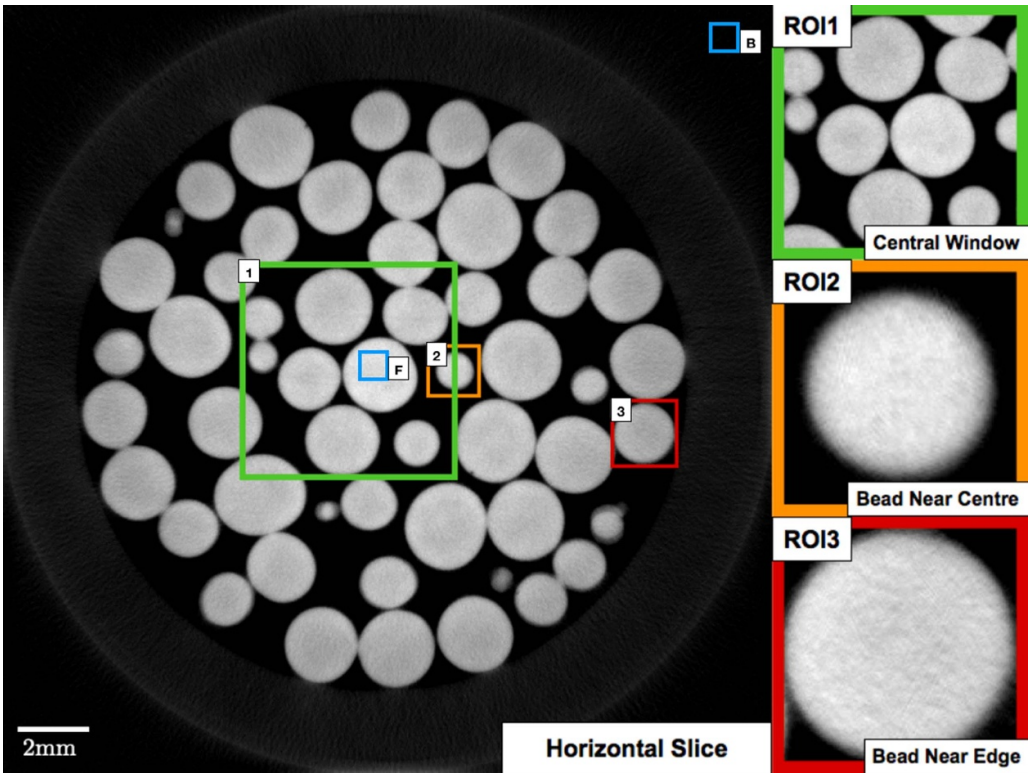


Figure 1. 2D CGLS central slice reconstruction of 512-projection data (four frames), along with three regions of interest, namely, a central window of the sample (ROI1; of size 440 px × 440 px or 704 μm × 704 μm), a bead near the centre (ROI2; size: 100 px × 100 px or 160 μm × 160 μm) and a bead near the edge (ROI3; size: 130 px × 130 px or 208 μm × 208 μm). The blue rectangular regions *F* and *B*, both size 50 px × 50 px, are the foreground (in this case, inside a bead) and the background (a region outside the sample holder) selected for the CNR estimations, respectively.

compared reconstruction methods. These methods are a popular approximate direct method with an implicit prior incorporated in a filter applied to the data (FDK); an iterative method with no explicit[7] prior information (CGLS), and an iterative method with a strong prior (TV Prior).

As the methods considered in this case study are of different types, the aim is to achieve comparable results as opposed to the best solution from either reconstruction method. In addition, the iterative methods require the consideration for an iteration cut-off point (which differs from data to data) and especially in the case of TV Prior, an additional set of parameters to fine-tune in order to achieve conceivable reconstructions. In order to obtain comparable results and avoid any bias towards a method, we performed parameter fine-tuning, as well as determining the iteration cut-off point, based on the 512-projection dataset. We note that the iteration cut-off point (i.e. the maximum number of iterations) is determined to be 12 iterations for CGLS, and 2000 iterations for TV Prior. Beyond these points, both methods started to either stall (no change above a certain threshold in the image domain) or diverge to a noisier image (assessed based on line profiles in the image domain). The difference in the maximum iteration numbers can be attributed to the influence of the strong prior. Of course running the TV Prior for 2000 iterations meant that the computation time increased drastically, with 2048-projection results taking approximately 4 weeks; 1024-projection results taking 2 weeks, and so on. We also note that we have not sought to accelerate the computations, but rather highlight the run times in order to provide relative indicators of the computational overhead associated with each method. The run time for each result is included in the quantification plot. Also note that there is no data fitting error for FDK as it is a direct inversion method [1, 35]. The reconstructed results of ROI1 for selected data sets are given in figure 2, followed by the quantification plot using Shape3D as the physical measure in figure 3.

It is not surprising that the TV Prior does better than the others for all reconstructions; the TV Prior takes into account the sparsity in the gradient reconstructions, enhancing the edges of the beads. Since the gradient image of SophiaBeads reconstructions is very sparse, the TV Prior works really well. The striking observation here is how long it takes to obtain a high quality result using this method. However, this is also expected since a more sophisticated formulation requires more computation—and in the case of TV Prior, this means multiple computations of the gradient and many iterations until a level of convergence is achieved. In our case, the numerical convergence for 1024- and 2048-projection data sets stalled, and the method was run until the limit of 2000 iterations was reached. The results did not show further improvements,

[7] An implicit prior information is provided for this method by the stopping criterion, which often is $\mathcal{E}_{\text{iter}}$ or $\mathcal{E}_{\text{res}}$. However in this paper, this is replaced by a user-defined number of iterations in order to eliminate the bias on the physical quantification.

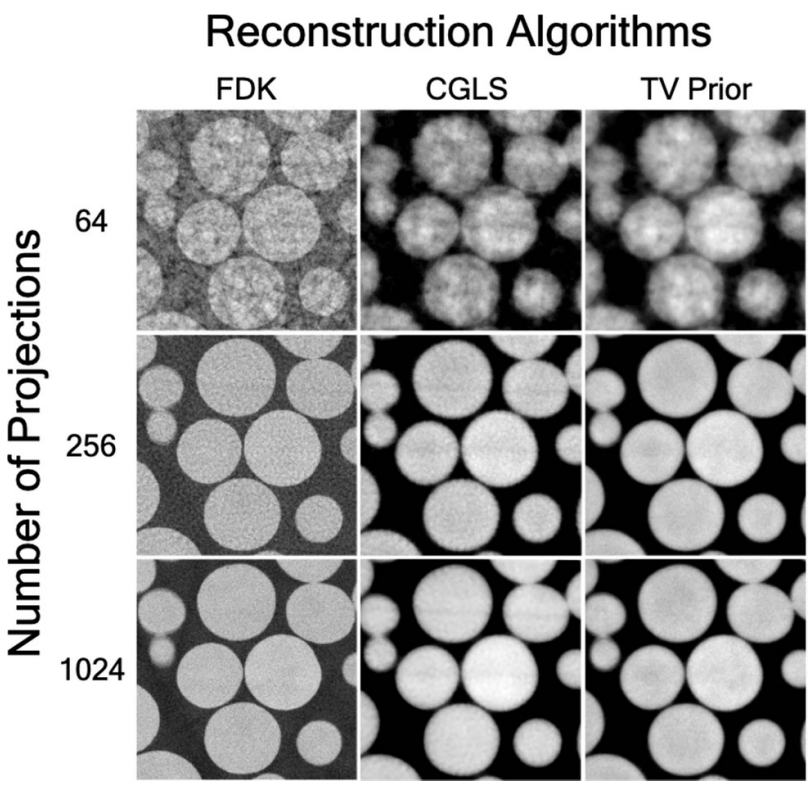


Figure 2. Central window reconstructions of Slice 1000 for (top to bottom) 64-, 256- and 1024-projection data sets, obtained using (left to right) FDK, CGLS and TV Prior.

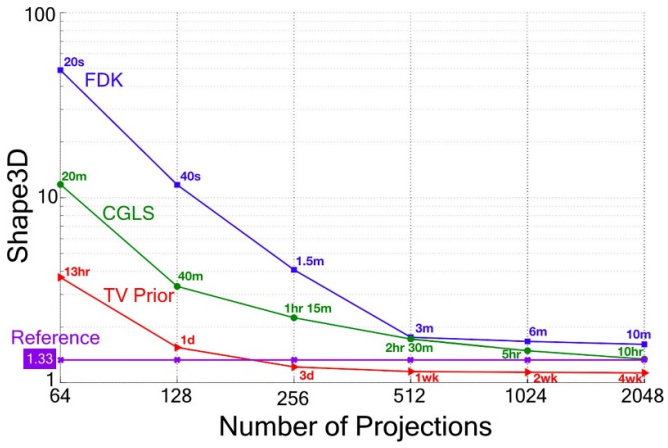


Figure 3. The results of the Shape3D analysis applied to the reconstructed volumes. The graph also includes the computational run times for each reconstruction, annotated above each result in matching colours.

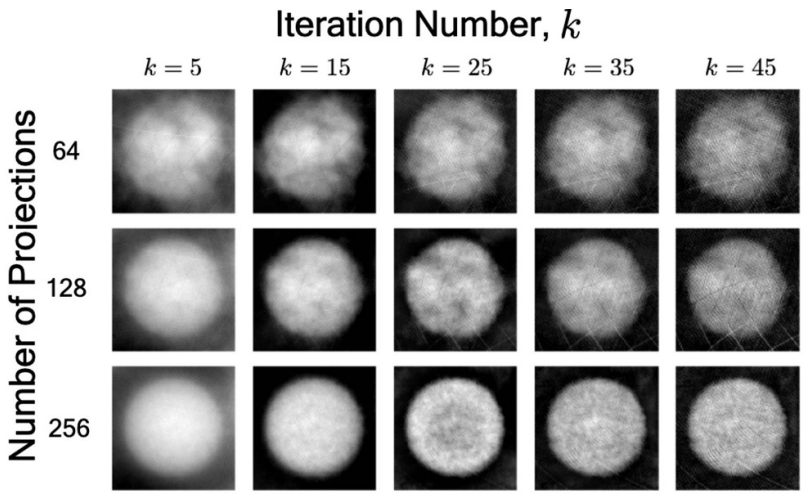
which is evident both in the Shape3D quantification plot and the reconstructed results.

TV Prior reaches the reference value threshold with 256-projection data set, although taking 3 days to get there. CGLS in comparison reaches the threshold in 10 h but with the highest number of projections. FDK never reaches the same level of accuracy as the other iterative methods (at 2048-projection data set, FDK solution returns Shape3D value of 1.61, while CGLS gives 1.33 and TV Prior 1.13).

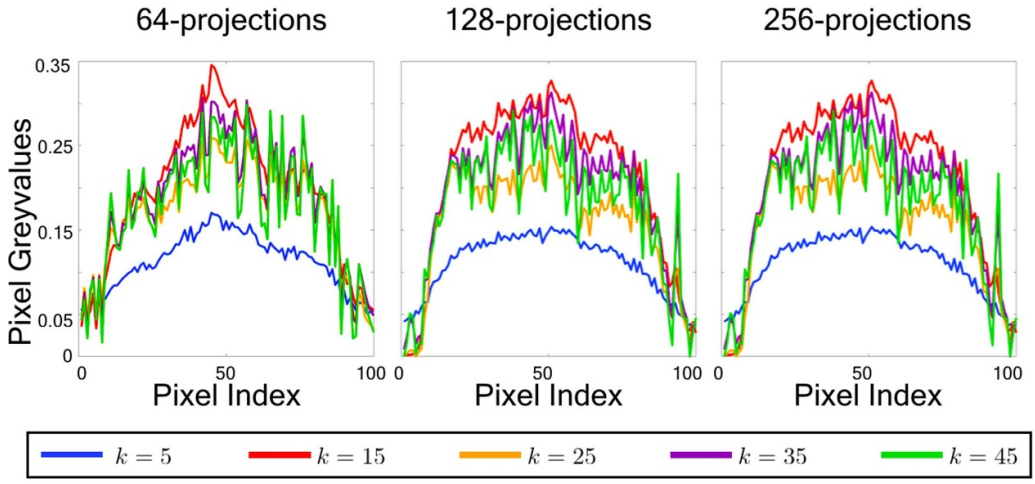
An interesting conclusion we can draw from the Shape3D quantification is that one can achieve the quality of the 2048-projection FDK result with a 512-projection, or even a 256-projection CGLS result. This could be particularly important in medical imaging where the dose required to record certain features is critical, or when imaging dynamic events where the acquisition speed is important. Using a simple prior, the CGLS result can be further improved upon with a small increase to the run time, thus the method can be seen as a good replacement for FDK in the case of a low-dose or few-projections problem. Conversely, when we collect data with many projections (e.g. 2048-projection), we obtain reconstructions in 10 min using FDK, which would otherwise take hours with CGLS and weeks with TV Prior.

5.2. Using physical measures to study the convergence of methods

In numerical analysis, typically a small test matrix (often with special properties or obtained via simulations) is used and the iterated solutions are compared to a true solution (a predetermined phantom) to study the behaviour of a method. The results however accurate and experiments representative of the physical case, do not always reflect what happens when applied to real world problems. Here, we have real CT data sets of varying projection numbers and a number of iterative methods, meaning we have the opportunity to study the behaviour of an iterative method with data obtained in a practical setting, using physical information we have of the beads.



(a)



(b)

Figure 4. Reconstructions of the bead near the centre (ROI2) at CGLS iterations 5 to 45 (a), and a line profile through the horizontal centre plotted at each iteration for 64-, 128- and 256-projection data sets (b).

We choose CGLS for this study as it is a fast iterative method where one 3D iteration ($x \in \mathbb{R}^{1564\times1564\times500}$) of 64-projection data set takes (on average) 90 s; 128-projection 180 s; 256-projection 360 s, and so on. Note that here we exclude the 2048-projection data set due to the excessive run time and limited contribution to the discussion (the results are similar to those of 1024-projection data set).

Two questions are considered in this study: How does CGLS converge in the case of real CT data, and how would this behaviour change as the amount of data is reduced and thus becomes insufficient?

We performed 50 iterations of CGLS, saving the reconstructed volume every 5th iteration to be analysed using a physical measure, here chosen to be AspectRatio3D. Also evaluated at each iteration, k, are $\mathcal{E}_{\text{iter}}$ and $\mathcal{E}_{\text{res}}$ (equations (3) and (4)), where the starting point, $x_{k=0}$ is a vector of zeros (an all-black image). We also evaluate CNR (equation (5)) for every 5th iteration. The following figures show the bead near the centre (ROI2) reconstructions for 64-, 128- and 256-projection data sets at iterations 5 to 45 (figure 4(a)); a line profile of through the (horizontal) centre of ROI2 every 5th iteration for each dataset (figure 4(b)), the linear plots of the pixel-based error measures (figure 5), CNR plot (figure 6) and the linear plot of AspectRatio3D values (figure 7).

It is difficult to draw any conclusions from the traditional error measures in figure 5. The relative iterative error plot in figure 5(a) indicates that as the iteration number, k, increases, the changes between reconstructions at consecutive iterations decrease. However, this does not mean that the method converges to a 'better' solution[8]. This simply means that the difference between the current and previous solution

[8] Better in the sense that the reconstructed image has less noise or the beads are accurately segmentable.

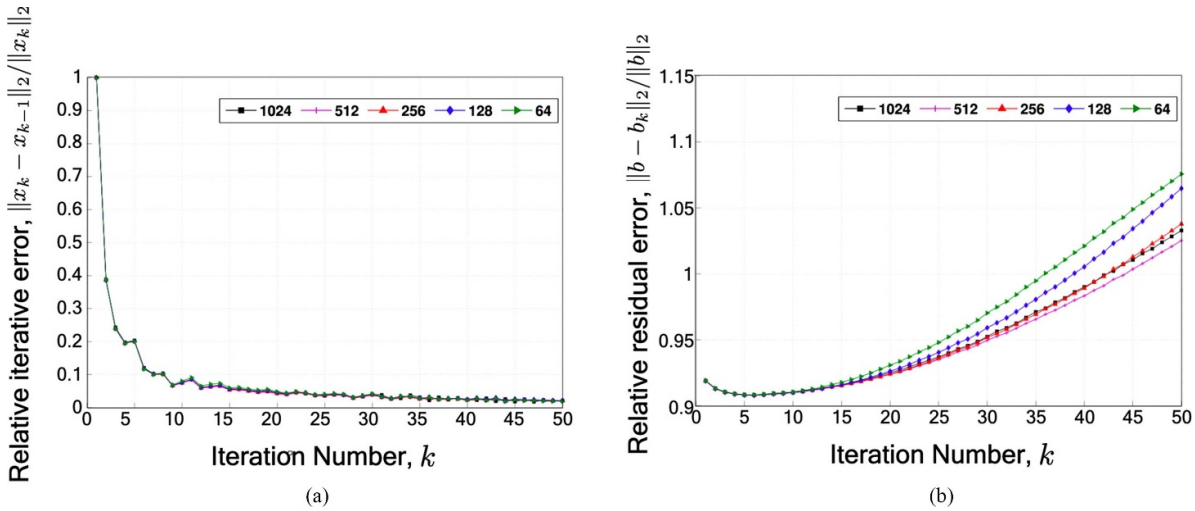


Figure 5. The relative iterative error, $\mathcal{E}_{\text{iter}}$ (a), and the relative residual error, $\mathcal{E}_{\text{res}}$ (b), for the CGLS reconstructions based on 64- to 1024-projection data sets.

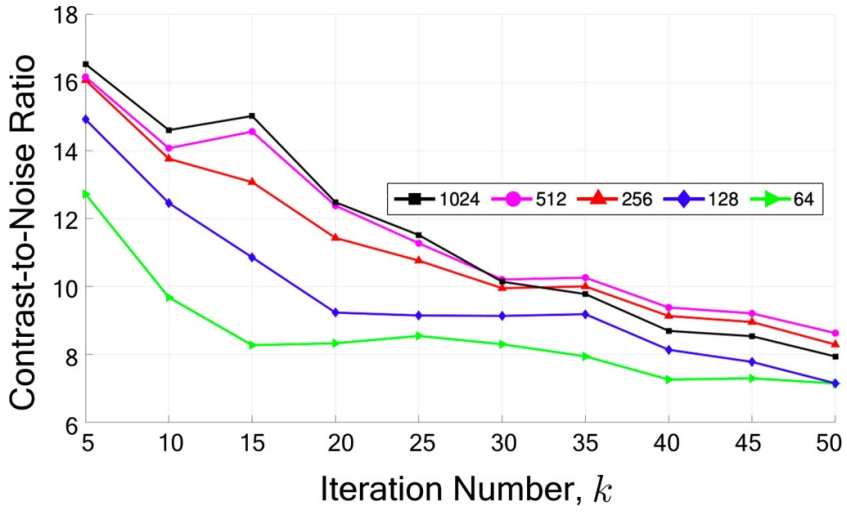


Figure 6. The CNR plot (in dB) for every 5th CGLS reconstruction based on 64- to 1024-projection data sets.

is decreasing, whereas at this point we would expect to have a noisier image at iteration 50. The relative residual error plot in figure 5(b), shows that as the iterations continue, the difference between the estimated (b) and predicted (b_k) data decreases initially, but later increases rapidly, at different rates. The minimum value of this error measure for all data sets is at iteration 4, suggesting that the 'best' approximation to a sensible solution is achieved at this point. However, the visual inspection of the reconstructed solutions at iteration 5 suggests the beads are still too blurry for a computer to automatically segment, and in the few-projection case, challenging even for the advanced visual perception of a human being. This shows how inaccurate standard practices can be when working with real data sets to extract physical information.

The CNR plot (figure 6) shows the contrast change with respect to noise as iteration number increases. Observing the plot, we see a clear divide between 1024- and 512-projection and 128- and 64-projection results in the sense that the pairs follow similar trends: The former pair show a quick recovery at iteration 15 before steadily decreasing (though at differing rates). The latter trend shows a steady decline from the beginning until iteration 20; a slight or no change in the CNR results until iteration 35, before declining again until iteration 50. Only the 256-projection results show a steady, predictable decrease throughout with a small recovery at iteration 35. The difference in trends can be due to the combination of angular sampling frequency being more pronounced with 256- and fewer projection data sets, and more noise being introduced into the reconstruction by CGLS as we iterate further. There are instances where CNR recovers slightly. For example the CNR result for 64-projection at iteration 25 gets close to the same level as that of 128-projection. Conversely, the 1024-projection results lose contrast and stoop to the level of 512-projection at iteration 20 (and lower onwards). Even more surprising is that 1024-projection is overtaken by 256-projection at iteration 30. All CNR results agree that the initial results for

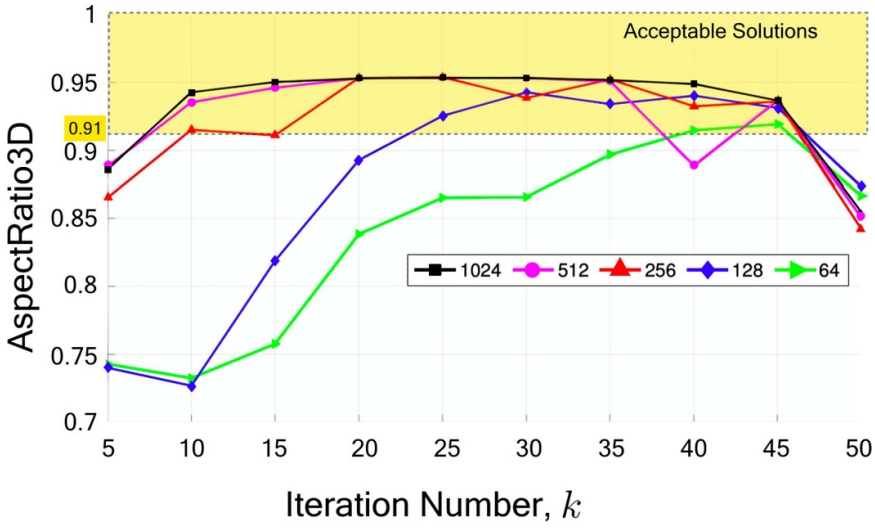


Figure 7. A linear plot of the physical quantification measure, AspectRatio3D. The quantification was performed for every 5th CGLS iteration for 64- to 1024-projection data sets. The shaded region indicates the acceptable solutions in the sense that all reconstructed objects within this space can be identified and segmented as beads, determined by the table 1 reference value for mean AspectRatio3D.

each data set (iteration 5) have the highest CNR, and as we iterate further, the values decline towards the background noise. However, a visual inspection of the first column in figure 4(a) shows blurry beads for all data sets.

We note that in our experiments employing CNR as an assessment technique, the sizes of regions chosen for the foreground and the background had a large effect on the resulting CNR plot, and at times, revealing completely different data trends. The difference was especially pronounced with low projection data sets (256-projection and fewer). In addition, none of the plots reliably pointed at a number of iterations where CGLS returned the 'best' results: they all agreed that the initial iterated solution (iteration 5) had the highest CNR, with the trend always decreasing for the remainder of the plot (although in unpredictable manner in comparison to one another). A future study could look into employing CNR for the initial sets of iteration (iterations 1–20) and what they reveal in challenging experimental setup (e.g. few-projection problems).

Figure 7 on the other hand shows the iterated values for which the reconstructed volumes resemble more defined beads. The yellow rectangle with dashed edges show the area for acceptable (i.e. accurately segmentable) AspectRatio3D solutions. This is determined using the table 1 value for mean AspectRatio3D (0.91 and above). The graph shows that the reconstructions quickly converge to an acceptable solution but later diverge. In the case of many-projections (512-projection and above), the convergence happens very quickly and results are stable for a while, whereas in the case of insufficient data (128-projection and below), the convergence is slower and results diverge shortly after, albeit at a slower speed than 256-, 512- and 1024-projection data. The results for the 256-projectiona are interesting in that their quality is as good as for the many-projection data sets, reaching the same level of quality as the 512- and 1024-projection results in just 20 iterations (3 h 20 min runtime).

It is worth noting that in contrast to the error and CNR plots, the AspectRatio3D analysis agrees with the visual inspection that the solutions in the middle are acceptable solutions. In addition the physical analysis agrees with the intuition that for data sets with many projections, the iterated solutions become more acceptable sooner. Meanwhile we observe a sudden dip in the physical assessment as we iterate further, which is a clear evidence of allowing for the noise to yield larger effects on the reconstructed results.

5.3. The effect of varying the acquired dose on the quality of reconstruction

Here, we attempt to answer an interesting question from an experimental point of view: The noise is undoubtedly reduced when we collect multiple instances (henceforth frames) of a projection and average over the total number. The intuition suggests reconstructions obtained via high-dose, fewer angle data would yield more accurate results as opposed to data with noisier projections taken more often. However, is it worth collecting multiple frames at conservative dose (exposure or photon count) in few-projection problems?

The acquisition plan of the SophiaBeads data sets as described in [8] is reproduced in figure 8, where line B represents data sets taken with fewer and fewer single-frame projections (i.e. the photon count is directly proportional to the number of projections, which is halved) whereas line A represents case where the number of projections decrease but the number of photons is compensated for by increasing the number of frames per projection (i.e. the same photon count in the data set as a whole). Line A data sets were released for the purpose of comparing algorithms, whereas line B data sets are not publicly available. However we can still obtain the line B data sets from line A by downsampling appropriately. For example, to obtain a 128-projection data set each projection with a single frame (line B), we can take every 16th projection

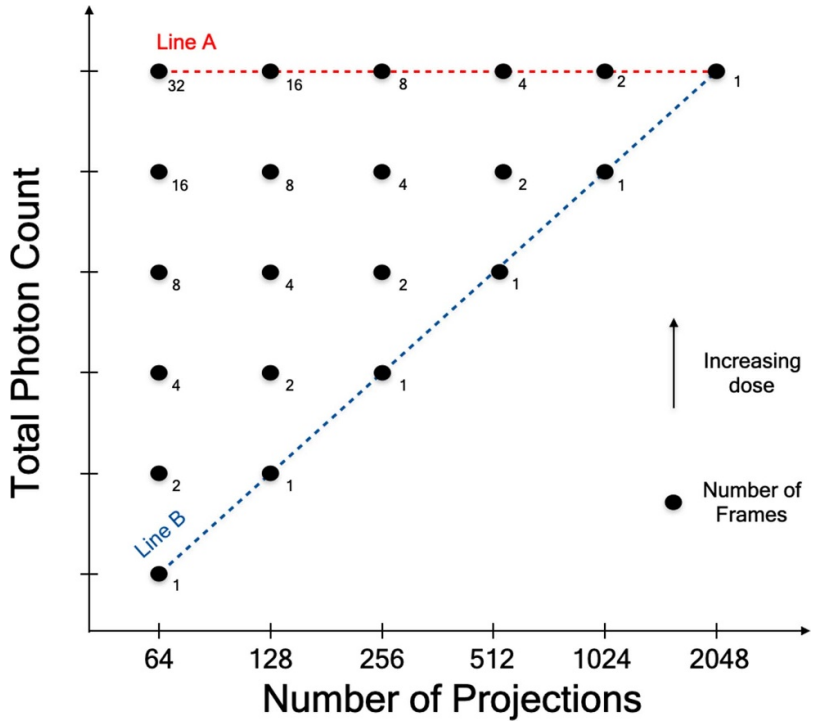


Figure 8. The tomographic acquisition plan for the beads pack data, reproduced following the description given in [8]. Line A here indicates the data sets with reduced number of projections while the photon count is kept constant; line B the data sets with halved photon count as the number of projections is decreased.

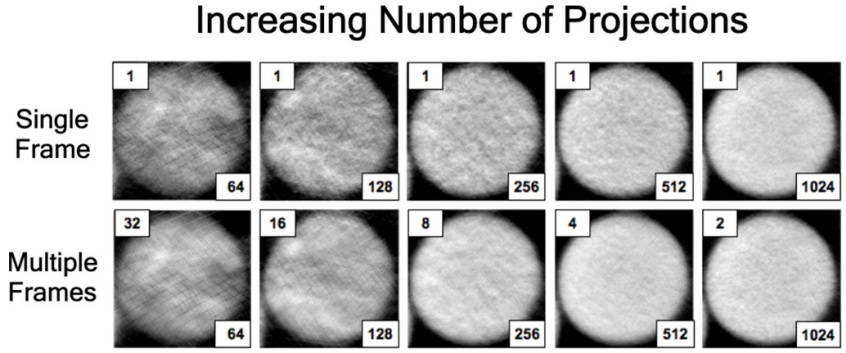


Figure 9. Reconstructions of the bead near the edge (ROI3) for all data set in line B (single frame, top row) and line A (multiple frames, bottom row).

from the 2048-projection data (1 frame). In fact, we can obtain a data set for any number of frames (i.e. dose) using the line A data sets: if we wanted to obtain 64-projection data set with 4 frames, we would downsample the 512-projection by 8 (32/4). In this respect it should be echoed here that this is possible because the SophiaBeads projections are not taken continuously but rather the rotation pauses to take one projection at a time (a stop-start scan).

The reconstructions were obtained using the same numerical set up as the previous case studies: by performing 12 CGLS iterations for each data set lying on lines A and B. Figure 9 shows ROI3 on Slice 1000 of each reconstructed volume.

The traditional image analysis technique is applicable here, and we use the $\mathcal{E}_{2\text{norm}}$ measure, taking the 2048 FDK result as the ground truth image (denoted by x^* in equation (2), see figure 10). For the physical quantification measure, we calculate the sphericity percentage of each solution in figure 11, including the mean reference value for sphericity, which is 87.16% (from table 1).

We observe that the traditional and the physical quantification measures agree, both confirming that collecting multiple frames generally give better results (as expected due to decrease in noise). However, when closely inspected, the trade off is not as big as intuition might suggest. In fact, looking at the trend of both lines A and B in the sphericity plot (figure 11), at its largest gap (64-projection data), both lines only exhibit 4.6% difference in the sphericity of the beads. This difference is reduced as the number of projections is increased. Studying either of the plots, we can further deduce that collecting a single-frame often gives results with highly comparable quality to multiple-frame data. This is observed in two locations.

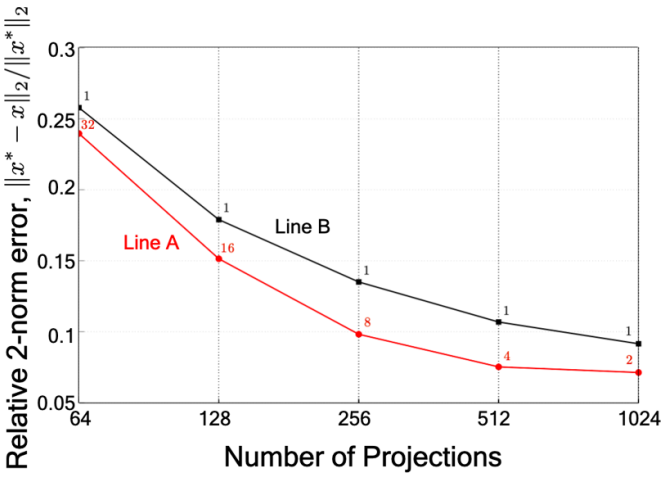


Figure 10. A log-linear plot of relative 2-norm error, $\mathcal{E}_{2\text{norm}}$, for all data sets. Each reconstruction is compared to the 2048-projection result, obtained using the same numerical setup.

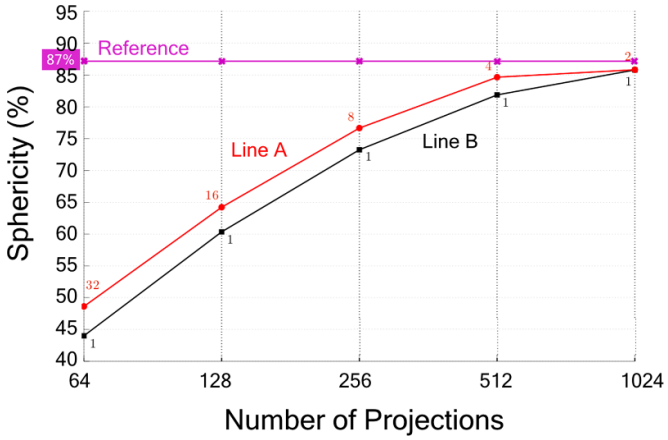


Figure 11. Sphericity of the reconstructed beads plotted as the number of projections increase. Mean reference physical value is denoted by the purple line with cross markers.

First, compare the line A 512-projection (4 frames) and 1024-projection (2 frames) results in either graphs: Here, the total exposure is the same but the exposure-per-projection is different. However, the level of reconstruction quality at these points is almost the same, meaning 2 frames of 1024 projections does not have any benefit over collecting 4 frames of 512. Secondly, observe the difference between line B 512-projection (1 frame) and line A 256-projection (8 frames) results in the sphericity plot. The line B solution with a single-frame gives a closer quantification value to the reference value (≈6% difference in sphericity) than the line A solution with 8 frames (≈11% difference). So despite having 4 times as much exposure-per-projection, the single-frame data set still gives a better solution. This also suggests that the SophiaBeads data sets were collected at too high an exposure level, reflecting a conservative tendency of CT practitioners to over-count per frame. A suggestion for future work could be looking at the limitations of iterative schemes where the exposure is conventionally deemed too short. Studying each line of sphericity individually, we note that the lines follow the same trend: the sphericity difference in line B data gets smaller as the number of projections is increased; line A following the same trend. We can attribute this to the effect of angular sampling: the trend of lines A and B is a good indicator of the effect of angular sampling in that, the gain in increasing the sampling frequency outweighs the gain that comes with higher exposure-per-projection. Another suggestion for a future study could be a comparison of the sphericity measure of individual beads across a reconstructed volume as a way to map the variations of image quality within the same volume. With the reduced exposure-per-projection coupled with fewer angles in the data, a reconstructed bead in the centre could exhibit different sphericity to one near the edge.

5.4. Integrated model-based reconstruction and segmentation—benchmarking implementations

The approach introduced in this paper can also be used to test out a novel imaging method, and to benchmark against more standard techniques to assess the accuracy and performance. The idea implemented here is to use segmentation as part of the reconstruction algorithm. Although fairly new in practice, the idea was considered in a more theoretical sense in [36, 37]. This recently gained attention and is being applied in more variations in the imaging community, namely to produce both continuous images using a Bayesian framework [38] or discrete images with a 2-step algorithm [14]. Also related is the idea of simultaneously obtaining both reconstructed and segmented image directly from the measured data [39], or using prior information about the phases in the sample [40].

The implementation in this study was inspired by the simple, 2-step discrete IR method detailed in [14], where a number of simultaneous algebraic reconstruction technique (SART) iterations are performed; the solution is segmented, boundary elements are determined, and SART iterations are continued for these boundary elements. Since the SophiaBeads sample only consists of homogeneous glass beads and air, the data sets are suitable for this study, in which we replace the SART method with CGLS (henceforth referred to as discrete-CGLS or DCGLS). This modification results in a large decrease in computational time. In [14], basic global thresholding is used, which is a simple method that classifies pixels in an image as 'black' (zero) or 'white' (above a user-defined threshold). To benchmark this aspect of the algorithm, we replace this step with other segmentation methods. Namely, the Otsu's thresholding algorithm, locally adaptive thresholding algorithm and the DRLSE method.

It is important to note that there were no parameter fine-tuning tests done for these segmentation algorithms, except for DRLSE. Of all the methods implemented, DRLSE was the most unstable as the quality depended heavily on a number of parameters. As it is outside the scope of this work, we do not go into the details of DRLSE but instead refer the reader to [15] for the derivation of the method (a summary is given in supplementary material appendix B along with the implementation details of this and other segmentation methods).

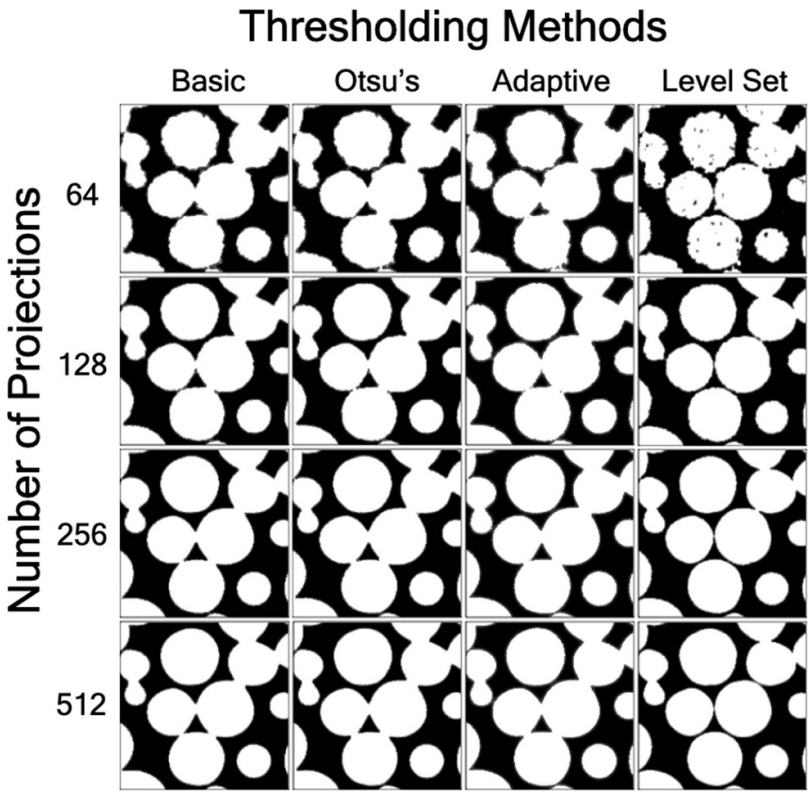


Figure 12. The central window of reconstructions for 64- to 512-projections data sets using (left to right) basic global thresholding (Basic), Otsu's thresholding (Otsu's), locally adaptive thresholding (Adaptive) and the distance regularised level set method (DRLSE).

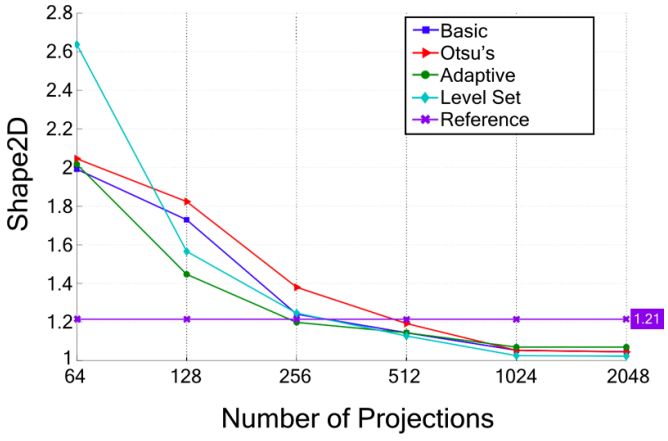


Figure 13. The log-linear plot of Shape2D quantification values for all segmentation methods for each data set, and of the physical reference value denoted by a purple dashed line with cross markers.

Initially five CGLS iterations were performed and the reconstruction at the 5th iteration was used as the starting point for DCGLS. We chose a small number of initial iterations to challenge the segmentation methods, which were shown to return acceptable solutions for 256-projection data set and above (see section 5.2). The DCGLS algorithm was applied to Slice 1000 to reduce the computation time but the implementations are easily expendable to 3D volumes. The ROI1 of selected reconstructions are shown in figure 12.

Traditional image analysis techniques are not feasible in this study since there is no segmented ground truth. We use Shape2D here as the physical quantification measure, applied to the entire slice, see figure 13. The reference value is the mean Shape2D value given in table 1.

As seen in both the reconstructions and the Shape2D quantification, DCGLS with the Level Set (DRLSE) at 64-projection returns relatively poor results. This may be due to the initial reconstructed volume (5 iterations of CGLS, recall the first column in figure 4) is not iterated long enough for the DRLSE to return a successful reconstruction. It could also be due to the DRLSE parameters being too strict for the few-projection case. Interestingly, for the same starting point, the most basic segmentation method returns better results, closely followed by the locally adaptive method. As the number of projections increase, the level set method takes over in quality, and appears to perform better than others.

Basic and level set methods reach a similar level of quality as the reference value at 256-projection with the locally adaptive method passing the reference value threshold (Basic: 1.24; locally adaptive: 1.20 and level set: 1.25), while Otsu's thresholding reaches this level with the 512-projection reconstruction. However, despite showing comparable Shape2D results to the basic and locally adaptive methods, the level set method produces more accurate reconstructions. This conclusion is drawn from the observation that the level set segmentations show the beads as almost touching or completely separate, which is more accurate than what the others produce (beads reconstructed with other methods show merged or

sintered beads for all cases, even at 2048-projection). We could of course have a stricter threshold value for the basic thresholding method, however this evidently means more unstable results for the data sets with fewer projections.

6. Discussion

The analysis approach presented here is applicable to various scenarios in imaging. As an example of this, as well as utilizing the SophiaBeads data sets, we have tried to answer typical questions that are of interest both to the mathematicians and the experimental scientists deploying CT. In this section, we consider the implications for each problem, as well as further analysis into comparison of algorithms.

The first case study showed that the iterative methods clearly gave better quality reconstructions. The prior information enhances the image quality and thus gives the best output measures. However, the results also highlighted an aspect often ignored, which is the computational time for reconstructions. As expected, more sophisticated algorithms take longer to reconstruct an image, and so there is a trade-off between computation time and quality. Our approach defines when the simpler methods are good enough.

The second problem involved studying the convergence behaviour through the use of the physical measures, and compared this to the standard practice of convergence analysis. Results obtained via the standard error norms did not yield conclusive results, demonstrating that the standard convergence analysis may be appropriate for simulations where a ground truth is predetermined but they are not applicable in practical cases. Studying the physical characteristics of a reconstructed volume to determine the convergence behaviour returns more trustworthy results, which can be verified by visual inspection.

The third study questioned the benefit of collecting data at constant exposure (total dose or photon count) as opposed to varying exposure (decreasing photon count). The results highlighted a conventional practice in CT, which is that even data sets with the least amount of information is collected with a 'comfortable' exposure level in order to collect the 'best projections'. For this reason collecting at high exposure did not have much benefit over collecting at lower exposure. An interesting study would be to repeat this assessment with data sets collected in more challenging (lower exposure) settings.

The final study looked at benchmarking an algoirthm integrating the reconstruction and segmentation steps, namely the 2-step algorithm, DCGLS. The results highlighted the importance of choosing parameters, which determined how unstable a method can be (especially in cases with few-projection data) and the accuracy of the reconstructed solutions.

Since the SophiaBeads data sets were collected specifically for comparing reconstruction algorithms, we include further analysis into the comparison presented in section 5.1.

6.1. Exploring the limits of reconstruction algorithms

The SophiaBeads data sets are collected with varying number of projections while keeping the total dose (or photon count) for each data constant. This was done to ensure fair comparison in the way that no other artefacts are introduced in the data as not to convolute the solution and therefore influence our conclusions. In section 5.1 where we compare three types of reconstruction algorithms, the resulting reconstructions were tested against a shape criterion, in this case Shape3D, to determine how close the reconstructed volumes were to a 'perfect solution'. Another aspect that could be tested here is the time it takes to obtain these reconstructions. Figure 3 has been annotated to include the reconstruction times, but these can also be presented in a graph in a similar manner. The presentation of results, given in figure 14, highlights the amount of time it takes to obtain a reconstruction, and raises important questions such as when does it make sense to implement more sophisticated methods and thus run longer reconstructions? Because it will take longer to reconstruct an image if there is more information, when does it become unnecessary to collect more information?

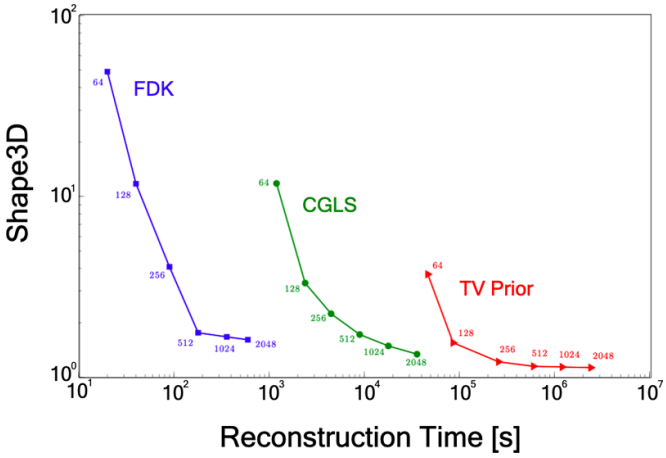


Figure 14. A log–log plot of the Shape3D results against the computational run time (in seconds) for each SophiaBeads data reconstruction.

It is clear from figure 14 that if we were to lower the dose, we can achieve improved results collecting 128 projections and reconstructing using TV Prior (in comparison to FDK). A similar degree of quality can be achieved with 512 projections, using CGLS, which would lead to a runtime that is approximately te times shorter than to obtain the TV Prior 128-projection reconstruction.

It is important to remember that the CGLS reconstruction results may be improved with simple or model-based regularization, and that the TV Prior reconstruction run time can be reduced by fine-tuning the parameters to avoid stalling in convergence. We observed stalling occurred more often with larger amounts of data, but in order to ensure that the TV Prior results were comparable with one another, we have used the same parameters for all TV Prior reconstructions. Similarly with CGLS, we chose not to include regularization in order to make a comparison between an iterative method with no prior information and one with model-based prior information (CGLS vs TV Prior).

7. Conclusions

A 3D reconstruction is used in applications where feature quantification plays a major role in e.g. quality control of a commercial product; testing and developing composite materials, or for tailoring a patient's treatments for a specific medical condition. Image analysis is crucial in assessing the suitability of a reconstruction method as well as the accuracy of information extracted from a reconstructed solution. Traditional techniques in image analysis (signal-to-noise ratio, L_2 norm, etc) were not adequate or able to give unbiased conclusions, and an established measure of merit was needed.

In this paper we introduced the idea of analysing a reconstructed solution using its characteristics such as shape, size and texture, which we referred to as physical quantifications. We defined a series of appropriate physical quantification measures to be applied to a CT data of a pack of glass beads, made available via the SophiaBeads Dataset Project, where the characteristics of the beads were known.

We exemplified the uses of physical quantification techniques through a number of case studies. These were chosen to be typical problems in CT in order to offer perspectives of mathematicians and experimentalists. Each study contained an example of physical characteristics being analysed, and a traditional analysis method is also included (where appropriate). This was done to highlight any possible discrepancies in conclusions drawn from both physical and traditional quantification techniques.

Our results demonstrate that physical quantification techniques can be applied to a wide range of problems, and offer clear indications as to the efficacy of a reconstruction strategy.

The results also showed that the introduced techniques allowed further analysis into the comparison of methods and analysis of various levels of dose in the acquisition. With these techniques we were able to conclude that the established beads pack data were collected at a conservative exposure level, which is the standard practice amongst the scientific experimentalists working in CT. A future task could be to collect a reference data set in more challenging settings, for example, when the projection view is severely limited or when the exposure-per-projection is decreased.

In addition, the results underlined the difference in computation times for a reconstruction, i.e. that the more sophisticated an algorithm gets, the longer the reconstruction takes. Another future task can be to limit the reconstruction time to assess the quality of results, with the goal of developing a sophisticated algorithm (correct forward model, appropriate stopping criterion, explicit prior and sensible optimization solver) that returns a reconstruction in relatively shorter amounts of time.

Acknowledgments

This work was funded by BP through the BP International Centre for Advanced Materials (BP-ICAM), the Department of Mathematics and EPSRC CCPi (EP/J010456/1) and the Netherlands Organisation for Scientific Research (NWO; Project No. 639.073.506). W R B L acknowledges the Royal Society for a Research Merit Award. P J W acknowledges European Research Council Grant CORREL-CT under Grant No. 695638. Funding for the Henry Moseley X-ray Imaging facility was provided by EPSRC (through Grants EP/M010619, EP/K004530, EP/F007906, EP/F028431). The authors would like to acknowledge the guidance of Dr Samuel A McDonald and Dr Julia Behnsen from the Manchester X-ray Imaging Facility (MXIF). An earlier version of this work was published as part of SBC PhD thesis [41, chapter 2].

ORCID iDs

Sophia B Coban https://orcid.org/0000-0002-9935-5906
William R B Lionheart https://orcid.org/0000-0003-0971-4678
Philip J Withers https://orcid.org/0000-0002-1946-5647

References

[1] Feldkamp L A, Davis L C and Kress J W 1984 Practical cone-beam algorithm *J. Opt. Soc. Am.* A **1** 612–19
[2] Solomon J and Samei E 2014 Quantum noise properties of CT images with anatomical textured backgrounds across reconstruction algorithms: FBP and SAFIRE *Med. Phys.* **41** 091908
[3] BS EN 16016-1 2011 *Non Destructive Testing. Radiation Methods. Computed Tomography. Terminology* (London: BSI)
[4] Vaishnav J Y, Jung W C, Popescu L M, Zeng R and Myers K J 2014 Objective assessment of image quality and dose reduction in CT iterative reconstruction *Med. Phys.* **41** 071904
[5] Kim S K, Kim C, Lee K Y, Cha J, Lim H J, Kang E Y and Oh Y W 2019 Accuracy of model-based iterative reconstruction for CT volumetry of part-solid nodules and solid nodules in comparison with filtered back projection and hybrid iterative reconstruction at various dose settings: an anthropomorphic chest phantom study *Korean J. Radiol.* **20** 1195–206
[6] Delingette H, Seguin O, Perrocheau R and Ménégazzi P 2008 Accuracy evaluation of 3D reconstruction from CT scan images for inspection of industrial parts (https://doi.org/10.1.1.156.5694)
[7] Kim M, Huh K H, Yi W J, Heo M S, Lee S S and Choi S C 2012 Evaluation of accuracy of 3D reconstruction images using multi-detector CT and cone-beam CT *Imaging Sci. Dentistry* **42** 25–33
[8] Coban S B, Withers P J, Lionheart W R B and McDonald S A 2015 When do iterative reconstruction methods become worth the effort? *Proc. 13th Int. Meeting of Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*
[9] Straumit I, Lomov S V and Wevers M 2014 Analysis and segmentation of a three-dimensional x-ray computed tomographyimage of a textile composite ed Y J Zhang and J M R S Tavares *CompIMAGE 2014* (*Pittsburgh, PA, USA, 3–5 September 2014*) (New York: Springer) pp 133–42
[10] Davour A, Svärd S J, Andersson P, Grape S, Holcombe S, Jansson P and Troeng M 2016 Applying image analysis techniques to tomographic images of irradiated nuclear fuel assemblies *Ann. Nucl. Energy* **96** 223–9
[11] Cortet B, Chappard D, Boutry N, Dubois P, Cotten A and Marchandise X 2004 Relationship between computed tomographic image analysis and histomorphometry for

microarchitectural characterization of human calcaneus *Calcified Tissue Int.* **75** 23–31
[12] Björck Å, Elfving T and Strakos Z 1998 Stability of conjugate gradient and Lanczos methods for linear least squares problems *SIAM J. Matrix Anal. Appl.* **19** 720–36
[13] Jensen T L, Jørgensen J H, Hansen P C and Jensen S H 2012 Implementation of an optimal first-order method for strongly convex total variation regularization *BIT Numer. Math.* **52** 329–56
[14] Batenburg K J and Sijbers J 2011 DART: a practical reconstruction algorithm for discrete tomography *IEEE Trans. Image Process.* **20** 2542–53
[15] Li C, Xu C, Gui C and Fox M D 2010 Distance regularized level set evolution and its application to image segmentation *IEEE Trans. Image Process.* **19** 3243–54
[16] Otsu N 1979 A threshold selection method from gray level histograms *IEEE Trans. Syst. Man Cybern.* **SMC-9** 62–6
[17] Bernsen J 1986 Dynamic thresholding of gray level images *ICPR'86: Int. Conf. on Pattern Recognition* vol SMC-9 pp 1251–5
[18] Coban S B and McDonald S A 2015 SophiaBeads dataset project *Zenodo* (available at: http://dx.doi.org/10.5281/zenodo.16474)
[19] Coban S B 2015 SophiaBeads dataset project codes *Zenodo* (available at: http://sophilyplum.github.io/sophiabeads-datasets/) (Accessed: 1 April 2020)
[20] Manchester X ray Imaging Facility 2015 The 320/225 kV Nikon XTEK bay (available at: www.mxif.manchester.ac.uk/resources/imaging-systems/nikon-custom-bay) (Accessed: 20 April 2020)
[21] Dobson K J, Coban S B, McDonald S A, Walsh J, Atwood R and Withers P J 2016 4D imaging of sub-second dynamics in pore-scale processes using real time synchrotron x-ray tomography *Solid Earth Discuss.* **1–27** 2016
[22] Murison J, Moosavi R, Schulz M, Schillinger B and Schröter M 2015 Neutron tomography as a tool to study immiscible fluids in porous media without chemical dopants *Energy Fuels* **29** 6271–6
[23] Davit Y, Iltis G, Debenest G, Veran-Tissoires S, Wildenschild D, Gerino M and Quintard M 2011 Imaging biofilm in porous media using x-ray computed microtomography *J. Microsc.* **242** 15–25
[24] Tsukahara M *et al* 2008 Coupled tomography and distinct-element-method approach to exploring the granular media microstructure in a jamming hourglass *Phys. Rev.* E **77** 061306
[25] Aste T, Saadatfar M and Senden T J 2005 Geometrical structure of disordered sphere packings *Phys. Rev.* E **71** 061302
[26] Avizo *Avizo User's Guide* (FEI Visualization Sciences Group)
[27] Wadell H 1935 Volume, shape and roundness of quartz particles *J. Geol.* **43** 250–80
[28] Friel J J 2000 *Practical Guide to Image Analysis* (Geauga County, OH: ASM International)
[29] Coban S B 2015 SophiaBeads datasets project documentation and tutorials *MIMS ePrints* **26** 1–22
[30] Ketcham R A 2005 Computational methods for quantitative analysis of three-dimensional features in geological specimens *Geosphere* **1** 32–41
[31] Rueden C T, Schindelin J, Hiner M C, DeZonia B E, Walter A E, Arena E T and Eliceiri K W 2017 Imagej2: Imagej for the next generation of scientific image data *BMC Bioinform.* **18** 529
[32] Limaye A 2012 Drishti: a volume exploration and presentation tool *Proc. SPIE* **8506** 85060X
[33] Schroeder W, Martin K and Lorensen B 2006 *The Visualization Toolkit* 4th edn (Clifton Park, NY: Kitware)
[34] Lifton J J, Malcolm A A and McBride J W 2015 An experimental study on the influence of scatter and beam hardening in x-ray CT for dimensional metrology *Meas. Sci. Technol.* **27** 015007
[35] Censor Y and Herman G T 1987 On some optimization techniques in image reconstruction from projections *Appl. Numer. Math.* **3** 365–91
[36] Mumford D and Shah J 1985 Boundary detection by minimizing functionals *Conf. on Computer Vision and Pattern Recognition*
[37] Mumford D and Shah J 1989 Optimal approximations by piecewise smooth functions and associated variational problems *Commun. Pure Appl. Math.* **42** 577–685
[38] Schneider M K, Fieguth P W, Karl W C and Willsky A S 2000 Multiscale methods for the segmentation and reconstruction of signals and images *IEEE Trans. Image Process.* **9** 456–68
[39] Ramlau R, Klann E and Ring W 2007 Simultaneous reconstruction and segmentation for tomography data *Proc. Appl. Math. Mech.* **7** 1050303–5
[40] Romanov M, Dahl A B, Dong Y and Hansen P C 2015 Simultaneous tomographic reconstruction and segmentation with class priors *Inverse Problems Sci. Eng.* 1–22
[41] Coban S B 2017 Practical approaches to reconstruction and analysis for 3D and dynamic 3D computed tomography PhD Thesis The University of Manchester (Accessed: 17 March 2019)