

A geometric approach to quadratic optimization: an improved method for solving strongly underdetermined systems in CT[§]

DAN GORDON*[†] and RAWIA MANSOUR[‡]

[†]Department of Computer Science, University of Haifa, Haifa 31905, Israel

[‡]Ort-Achva High School, Naura, D.N. Gilboa 18950, Israel

(Received 17 February 2006; in final form 1 August 2006)

The problem of image reconstruction from projections in computerized tomography, when cast as a system of linear equations, leads to an inconsistent system. Hence, a common approach to this problem is to seek a solution that minimizes some optimization criterion, such as minimization of the residual (least-squares solution) by quadratic optimization. The QUAD algorithm is one well-studied method for quadratic optimization which applies the conjugate gradient to a certain so-called “normal equations” system derived from the original system. The concept of a *geometric algorithm* is introduced and defined as a method that is independent of any particular algebraic representation of the hyperplanes defined by the equations. An example of a geometric algorithm is the well known ART algorithm which, when implemented with a small relaxation parameter, achieves excellent results. Any non-geometric algorithm can be converted to a geometric one by normalizing the equations before applying the algorithm. This concept, when applied to QUAD, results in a geometric algorithm called NQUAD. ART, QUAD, and NQUAD were tested on various phantom reconstructions and evaluated with respect to image quality, convergence measures, and runtime efficiency. The results indicate that NQUAD is always preferable to QUAD, thus proving the usefulness of the geometric approach. Although ART is somewhat better than NQUAD in most standard cases, NQUAD is much better when the system is strongly underdetermined (a situation corresponding to low dose radiation), and this is even more pronounced with low-contrast images.

Keywords: Computerized tomography; Conjugate gradient; Image reconstruction; Iterative algorithms; Least squares; Normal equations; Quadratic optimization

2000 Mathematics Subject Classifications: 65F10; 65F50; 65K10; 65R32

*Corresponding author. Email: gordon@cs.haifa.ac.il

[§]Preliminary results presented at the 8th IASTED Conference on Computer Graphics & Imaging, August 2005.

1. Introduction

We consider the problem of image reconstruction from projections in transmission computerized tomography (CT) [11]. After a discretization of the image domain, the problem leads to a large and sparse system of linear equations, $Ax=b$, which is inevitably inconsistent, due to a variety of reasons (noise, inaccurate measurements, as well as the discretization process). Solution methods usually fall into *transform* methods or *iterative* methods, with the latter being generally slower but providing superior results under adverse conditions.

One of the earliest and most successful iterative approaches to image reconstruction from projections is algebraic reconstruction technique (ART) [11]. In this method, originally due to Kaczmarz [15], each iterate is projected onto a hyperplane (defined by an equation) in a cyclic manner. ART has been studied very extensively, both in theory and in practice. It is shown in [11] that the addition of a small relaxation parameter into ART produces excellent results. The convergence of ART with relaxation parameters, for consistent systems, has been shown by Herman *et al.* [13] and by Trummer [23]. Tanabe [22] has shown that when the system is inconsistent, ART converges *cyclically*, i.e., for each hyperplane, the sequence of projections on that hyperplane converges to a limit. Censor *et al.* [4] extended Tanabe's result to ART with a fixed relaxation parameter λ and also showed that for every hyperplane P , if the cyclic limit w.r.t. P is x_λ^P , then the limit $x^P = \lim_{\lambda \rightarrow 0} x_\lambda^P$ exists, and all the x^P 's coincide at a point which minimizes the (unweighted) sum of squares of the Euclidean distances to the hyperplanes. Note that such a minimization point is not necessarily unique; however, it was shown in [4] that if the iterations start with the zero vector, then the cyclic limits converge to a minimum norm point which achieves the above minimization.

A common approach to retrieving a "good" image is to seek a solution that minimizes some optimization criterion [7]. One such criterion is the minimization of $\|Ax - b\|$, achieved by quadratic optimization [1,12]. $\|Ax - b\|$ is the L_2 -norm of the residual, and a solution which minimizes this expression is also known as a least-squares solution. One way of obtaining a least-squares solution is to apply the conjugate gradient algorithm (CG) [14,21] to one of the symmetric "normal equations" systems $A^T Ax = A^T b$ or $AA^T y = b$ ($x = A^T y$); see [2]. A least-squares solution is not necessarily unique; however, there exists a unique *minimum norm* least-squares solution, and the iterates produced by the normal equations methods converge to this solution. Note that a point minimizing the L_2 -norm of the residual depends on the algebraic representation of the hyperplanes; multiplying any equation by a factor (different from zero or one) could lead to a different solution.

In contrast to the above normal equations methods, ART is a *geometric* algorithm in the following sense: The sequence of iterates generated by ART depends only on the hyperplanes defined by the equations and not on any particular algebraic representation of the hyperplanes. Any non-geometric iterative algorithm can be converted into a geometric one if, before applying the algorithm, the equations are *normalized* by dividing each equation by the L_2 -norm of the vector of its coefficients. This idea lies at the basis of our approach.

A simple example illustrating the difference between a geometric and a non-geometric method is the following system of two (inconsistent) equations in two variables x_1, x_2 : $x_1 + 0x_2 = 0$ and $10x_1 + 0x_2 = 10$. For any point $x = (x_1, x_2)$, the square of the L_2 -norm of the residual is $\|Ax - b\|^2 = x_1^2 + (10x_1 - 10)^2$. The minimum of this

quadratic polynomial is obtained at $x_1 \approx 0.99$. However, if the equations are normalized, then the first equation remains unchanged and the second equation becomes $x_1 = 1$. The L_2 -norm of the residual of this normalized system is minimized at $x_1 = 0.5$. This simple example illustrates how different algebraic representations of the same hyperplanes can affect the solution obtained by quadratic optimization. On the other hand, the iterates generated by ART are independent of any scaling of the equations.

In the application of CG, the literature abounds with a variety of preconditioning techniques, which can be regarded as multiplying A (on the left and/or the right) by some matrix. The particular method for quadratic optimization that we have used, which we call QUAD, is obtained by first performing *column normalization* on A , i.e., dividing each column of A by the L_2 -norm of the column, and then minimizing the residual of the resulting system [1]. Column normalization is equivalent to right-multiplying A by a certain diagonal matrix obtained from the column norms. Note that the resulting algorithm is not necessarily geometric in the above sense, because the resulting equations are not necessarily normalized.

In this study, we apply the geometric approach to QUAD by first normalizing the equations and then applying QUAD to the resulting system. The resulting algorithm, called NQUAD (Normalized QUAD), was implemented by Mansour [18]. Extensive experiments were carried out on many phantom images, comparing ART (with a small relaxation parameter), QUAD and NQUAD. Note [1, p. 245] that previous studies of QUAD did not scale the equations in any way. The comparisons employed two measures of the error (with respect to the phantom), the number of iterations required to achieve the optimal measures, and visual comparisons. Our results indicate that NQUAD is always better than QUAD on all counts, though the differences are not always visible. On an almost fully determined problem, ART produced better images than NQUAD, and in fewer iterations (though eventually NQUAD catches up).

Our most important finding is that with a strongly underdetermined system of equations (number of equations approximately equal to 25% of the number of variables) both QUAD and NQUAD were much better than ART, and NQUAD was better than QUAD. Furthermore, NQUAD achieved its best error measures in significantly fewer iterations than the other methods. Furthermore, the differences between the algorithms are much more pronounced on images of low contrast. The enhanced artifacts on low contrast images have been observed before in a somewhat different setting: “Even though these artifacts may have never been noticeable in high-contrast reconstructions, they become very visible in the low-contrast situation”— see [20, p. 520].

A strongly underdetermined system is compatible with a situation where low dose radiation is mandated for clinical reasons. Note that the dose is also reduced in such cases, but it is still preferable to use a technique that also requires fewer rays. Preliminary results of this research appear in [9].

The rest of this article is organized as follows: section 2 presents the image reconstruction problem in CT, and section 3 explains ART, QUAD, and NQUAD. Section 4 presents our results, and we conclude with some further research directions.

2. The reconstruction problem

In transmission CT, X-rays are passed through a cross section of an object and readings are taken by detectors at the opposite side. Different parts of the cross section attenuate

the X-rays differently, resulting in different readings at the detectors. The problem is to reconstruct the attenuation map—or image—of the cross section from the detector readings. Mathematically, the unknown attenuation is a nonnegative function of two variables defined on the cross section. The most elementary method for reconstructing the attenuation is to overlay the cross section with a Cartesian grid of pixels, and to consider the attenuation to be constant inside every pixel. Let n denote the total number of pixels.

In a typical setup, we assume that a line of parallel X-ray sources are aimed at a line of detectors, as shown in figure 1. X-rays are cast toward the detectors, which measure the attenuation along the rays' paths. The orientation of the sources and detectors is then rotated by some small angle, and the process repeats until the entire cross section has been "covered" by equiangular orientations. We denote by m the total number of rays obtained in this manner.

We assume that the pixels and rays are numbered consecutively in some manner from 1 to n and from 1 to m , respectively. Let x_j denotes the (assumed constant) value of the attenuation function in the j th pixel. It is assumed that the X-ray sources and detectors are points, and that the rays between them are lines. It is further assumed that the length of intersection of the i th ray with the j th pixel, denoted by a_j^i , represents the contribution of the j th pixel to the total attenuation along the i th ray.

The physical measurement of the total attenuation along the i th ray, denoted by b_i , is the line integral of the unknown attenuation function along the path of the ray. In the discretized setup, each line integral is approximated by a finite sum, so the (discretized) attenuation function is the solution of a system of linear equations

$$\sum_{j=1}^n a_j^i x_j = b_i \quad \text{for } 1 \leq i \leq m, \text{ or, in matrix form: } Ax = b. \quad (1)$$

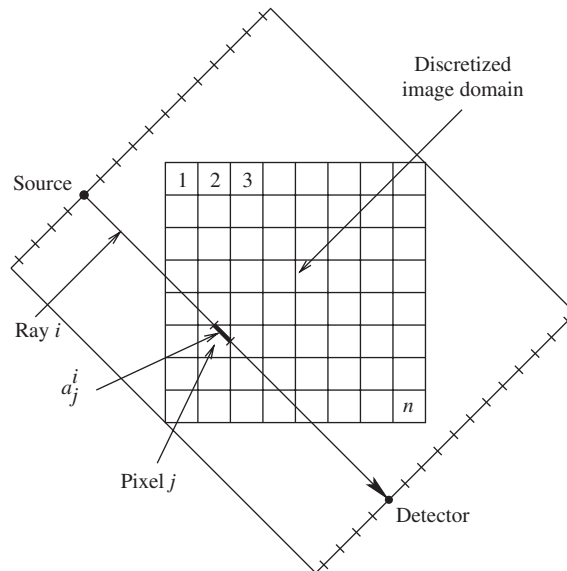


Figure 1. Tomographic image reconstruction in the discretized model.

3. ART, CG, and quadratic optimization

For $1 \leq i \leq m$, we denote by a^i the i th row vector of the matrix A in equation (1). $\langle u, v \rangle$ denotes the dot product of two vectors u, v and $\| \cdot \|$ denotes the L_2 -norm of a vector, i.e., $\|u\|^2 = \langle u, u \rangle$. ART can be described as follows: Starting from an arbitrary point $x^0 \in \mathbb{R}^n$, the k 'th iterate x^k is projected towards the next hyperplane, and the hyperplanes are chosen in cyclic order. We denote by $i(k)$ the k 'th index taken cyclically from 1 to m , i.e., $i(k) = k \bmod m + 1$. ART, with a fixed relaxation parameter λ , is the following algorithm:

Algorithm 1 (ART):

Initialization: $x^0 \in \mathbb{R}^n$ is arbitrary.

Iterative step: Given x^k compute

$$x^{k+1} = x^k + \lambda \frac{b_{i(k)} - \langle a^{i(k)}, x^k \rangle}{\|a^{i(k)}\|^2} a^{i(k)}.$$

For a symmetric $n \times n$ system matrix A , the conjugate gradient algorithm is the following (see [21]).

Algorithm 2 (CG):

(i) $x^0 \in \mathbb{R}^n$ is arbitrary, $p^0 := r^0 := b - Ax^0$

(ii) for $i = 0, 1, 2, \dots$ until convergence:

$$\begin{aligned} q^i &:= Ap^i \\ \alpha_i &:= \|r^i\|^2 / \langle p^i, q^i \rangle \\ x^{i+1} &:= x^i + \alpha_i p^i \\ r^{i+1} &:= r^i - \alpha_i q^i \\ \beta_i &:= \|r^{i+1}\|^2 / \|r^i\|^2 \\ p^{i+1} &:= r^i + \beta_i p^i \end{aligned}$$

(iii) output x^i .

Quadratic optimization is obtained by seeking a least-squares solution of the system (1), namely, a point $x^* \in \mathbb{R}^n$ minimizing $\|Ax^* - b\|$ (the L_2 -norm of the residual). It follows from [2, Theorem 1.2] that a solution of $A^T Ax = A^T b$ is a minimum-norm, least-squares solution of (1). As mentioned, QUAD applies CG to the system obtained from (1) by dividing each column of A by its L_2 -norm, as follows.

Algorithm 3 (QUAD):

(i) Normalize columns: For $1 \leq j \leq n$, let a_j be the j th column of A , $c_j = 1/\|a_j\|$ and $D = \text{diag}(c_1, \dots, c_n)$. Right-multiply the matrix A of the system (1) by D , and, denoting $E = AD$, obtain the system

$$Ey = b, \quad x = Dy.$$

(ii) Apply Algorithm 2 (CG) to the (symmetric) system $E^T E y = E^T b$, obtain a solution y^* , and output $x^* = Dy^*$.

Our geometric approach, when applied to QUAD, transforms it into the geometric algorithm NQUAD by first normalizing the equations of (1), as follows.

Algorithm 4 (NQUAD):

- (i) Normalize rows: For $1 \leq i \leq m$, let $r_i = 1/\|a^i\|$ and $D_1 = \text{diag}(r_1, \dots, r_m)$. Left-multiply both sides of equation (1) by D_1 , and, denoting $E = D_1A$, obtain the system

$$Ex = D_1b. \quad (2)$$

- (ii) Apply Algorithm 3 (QUAD) to the system (2).

4. Results

The aim of our experiments was to provide an initial evaluation of the notion of a geometric algorithm, and how the geometric and non-geometric versions of the same algorithm differ from each other. To this end, we compared ART, QUAD, and NQUAD. There are many other reconstruction algorithms which also deserve appropriate study, and they will be the subject of future research. In ART, we experimented with different values of the relaxation parameter λ in two different ways: In one set of experiments, λ was kept fixed throughout all the iterations, and different values of λ were tested. In another set of experiments, λ was allowed to change with the iteration number in various ways. None of these experiments produced better results than a fixed value of 0.1 for λ . The comparisons between the various algorithms were made both quantitatively using the convergence measures defined in the following subsection, and qualitatively, by visual comparisons. We also tested the two parallel algorithms CAV and BICAV [5,6], but these produced results similar to ART; they were less relevant to this study since their primary strength is in a parallel processing environment.

The three algorithms were tested on a large number of phantom test cases, with and without noise. The experiments were carried out within the framework of the SNARK93 package [3], which contains the means for creating various phantoms and includes a variety of standard reconstruction algorithms, including QUAD. NQUAD was implemented within the framework of SNARK93 by Mansour [18]. In all the tested algorithms, the initial estimate was taken as $x^0 = 0$.

Many test cases were examined, including cases with noise, and in most of these, the differences between the three methods were insignificant. Hence, we report only on cases that are relevant to the examined algorithms. These cases and the convergence measures used to compare the algorithms' behavior are detailed in the next subsection. Subsection 4.2 provides details of the results and a discussion. In subsection 4.3, we examine two more cases in order to isolate the different factors affecting the algorithms' behavior.

4.1. Test cases and measures

Two notable representative test cases are presented, both at a size of 255×255 . Case 1 is the Herman head phantom [11, section 4.3], which is specified by a set of ellipses, with a specific attenuation value (called *density*) attached to each elliptical region. This phantom can be seen in figure 2. The pixel size was set identical to the inter-ray

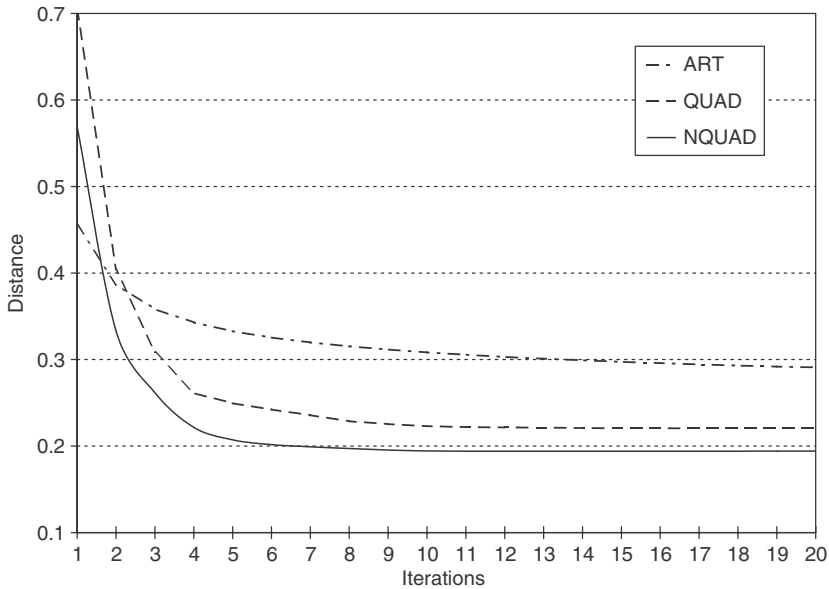


Figure 2. Distance measure for Case 2.

distance; this setup ensures the least amount of artifacts in the reconstructed images (as can be readily verified by experimentation). The reconstruction used 180 equidistant angles, resulting in a system of 65,025 variables and 64,980 rays—an almost fully determined system. Note that the concept of a fully determined system in CT also means that the projections are evenly distributed.

Case 2 is an example of a very adverse situation for any reconstruction method: a strongly underdetermined system of equations and a low contrast image. The phantom is based on example 7 of SNARK, modified to have very low contrast, with the low and high values differing by only 11% from the mean. This phantom can be seen in figure 3. The number of equations is only about 25% of the number of variables; this was obtained from the geometry of Case 1 by doubling both the inter-ray distances and the angle between successive projection angles. Table 1 summarizes the two cases.

The values of b_i , $1 \leq i \leq m$, are calculated by computing the line integrals through the elliptical regions (without reference to the discretization). Thus, the system (1) is basically inconsistent, because the left-hand-side is only an approximation to the actual integrals. This matches the real-life situation where the b_i 's are (path-lengths derived from) X-ray readings through an object but the region of interest is discretized as above. The algorithms were run on both cases for 40 iterations (an iteration means one cyclic sweep through all the equations).

Two measures were used to define how much a reconstructed image diverges from the phantom, the *distance* and the *relative error*. These are calculated by SNARK93 [3], and defined as follows. Let \hat{x} denote the phantom and x^k denote the reconstructed image after k iterations. Let S denote the set of indices j of pixels which are in the region of interest and let α be the number of elements in S . The *average* value of \hat{x} is defined by

$$\hat{\rho} = \frac{1}{\alpha} \sum_{j \in S} \hat{x}_j,$$

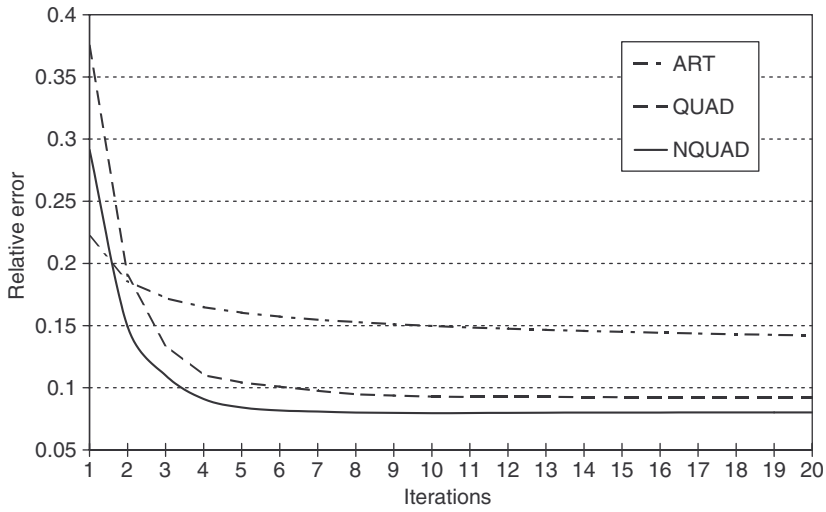


Figure 3. Relative error measure for Case 2.

Table 1. Details of the two test cases.

| Test cases | Case 1 | Case 2 |
|---------------|-------------|-----------|
| Image size | 255 × 255 | 255 × 255 |
| Variables | 65,025 | 65,025 |
| Equations | 64,980 | 16,290 |
| Projections | 180 | 90 |
| Rays | 361 | 181 |
| Density range | 0.113–0.726 | 0.4–0.5 |

and the Standard deviation of \hat{x} is defined as

$$\hat{\sigma} = \left(\frac{1}{\alpha} \sum_{j \in S} (\hat{x}_j - \hat{\rho})^2 \right)^{1/2}.$$

The *distance* between x^k and \hat{x} is defined as

$$\delta_k = \frac{1}{\hat{\sigma}} \left(\frac{1}{\alpha} \sum_{j \in S} (x_j^k - \hat{x}_j)^2 \right)^{1/2}.$$

Setting $\hat{\tau} = \sum_{j \in S} |\hat{x}_j|$, the *relative error* of x^k is defined by

$$\epsilon_k = \frac{1}{\hat{\tau}} \sum_{j \in S} |x_j^k - \hat{x}_j|.$$

4.2. Results and discussion

Tables 2 and 3 show the minimal distance and relative error measures for the three algorithms on the two test cases, together with the iteration numbers at which the minimum values were obtained. In Case 1, the measures are not significantly different from each other, though ART reaches the minimal relative error at fewer iterations. It can be seen that ART is somewhat better than NQUAD, and NQUAD is somewhat better than QUAD; this behavior is similar in a very wide range of examples that were tested. In Case 2, the differences between the three algorithms are much more pronounced, with NQUAD having a large lead over the other two methods. Even QUAD is much better than ART in this case. Furthermore, we can see that NQUAD reached its better values relatively early. In ART and QUAD, the two measures continued to decrease till the last iteration.

Figures 4 and 5 present plots of the two error measures as a function of the number of iterations for Case 2. One iteration of NQUAD took about twice the time of an ART iteration, but even so, NQUAD is clearly preferable.

Figures 2, 3, and 6 show the phantom and reconstructed images for Cases 1 and 2. The images were obtained by a program called SNARKVIEW [8] from the output file RECFIL which is produced by SNARK93. We set this program to assign grey level zero to the lowest-valued pixel, grey level 255 to the highest-valued pixel, and all other pixels are assigned grey levels at inbetween values by linear interpolation. This method of display differs from standard practice (which assigns grey levels by considering the phantom and images together); however, it is much better at showing up artifacts in the various images. For the sake of brevity, the reconstructed images are shown only at 10 iterations.

Figure 2 shows the phantom and images reconstructed by the three methods for Case 1. Note that there are no visible differences between the three methods, though at 5 iterations (image not shown), ART already achieved as good an image as NQUAD at 10 iterations. Figure 3 shows the phantom and reconstructions for Case 2. The ART reconstruction appears dark because some pixels in the reconstructed image have a high value. We conjecture that this problem is due to the fact that ART does not deal well with the low-contrast phantom when the number of equations is

Table 2. Case 1—minimal distance and relative error, and iteration number(s) at which they were obtained.

| Measure | ART | QUAD | NQUAD |
|-----------|-------------------|----------------|----------------|
| Distance | 0.0807, it. 12–13 | 0.0837, it. 15 | 0.0803, it. 13 |
| Rel. err. | 0.0497, it. 8–9 | 0.0558, it. 13 | 0.0521, it. 12 |

Table 3. Case 2—minimal distance and relative error, and iteration number(s) at which they were obtained.

| Measure | ART | QUAD | NQUAD |
|-----------|----------------|-------------------|-------------------|
| Distance | 0.2803, it. 40 | 0.2197, it. 40 | 0.1922, it. 11–14 |
| Rel. err. | 0.1373, it. 40 | 0.0918, it. 37–40 | 0.0771, it. 10 |

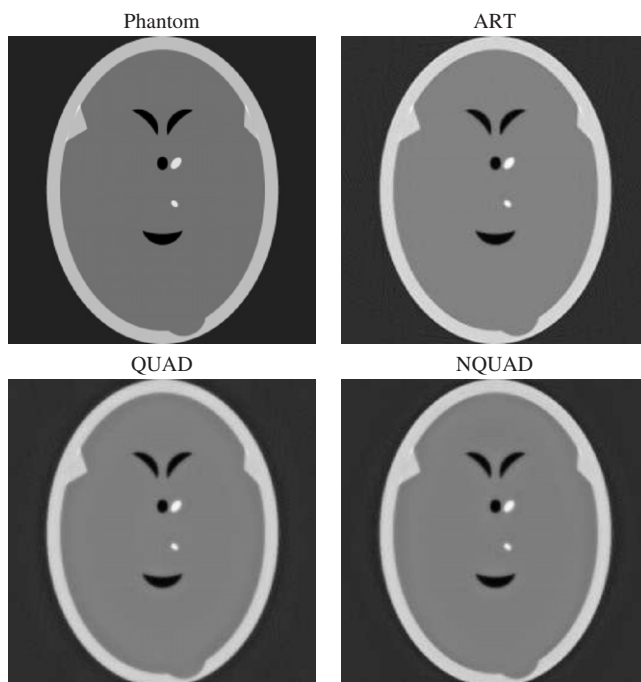


Figure 4. Case 1: phantom and reconstructed images after 10 iterations.

small. In order to compensate for the dark ART images, Figure 3 also shows the ART image after it was brightened to match the general level of the other images. This was done by mapping all grey levels of 174 and above to 255, with the others linearly mapped between 0 and 255. This figure shows that the ART image is inferior to the one produced by NQUAD. The differences between the images of QUAD and NQUAD can be discerned only with difficulty.

In order to better visualize the distinctions between the reconstructions in Case 2, we applied the standard image processing technique of histogram equalization to the phantom and the reconstructed images, as shown in figure 6. Note that now the phantom no longer appears as a low-contrast image. These images reveal more strongly that ART produces some artifacts, such as the concentric “rings”, and that the various elliptical regions are not as sharply defined as in NQUAD. To the best of our knowledge, there is no mathematical explanation as to why artifacts take one particular form or another. A careful examination of the QUAD and NQUAD images reveals that the QUAD image is slightly more noisy in some places—see for example the topmost dark rectangular patch.

4.3. *Intermediate cases*

Case 2 differs from Case 1 in two major aspects: the number of equations is only about 25% of the number of variables, and it is of very low contrast. In view of the results of

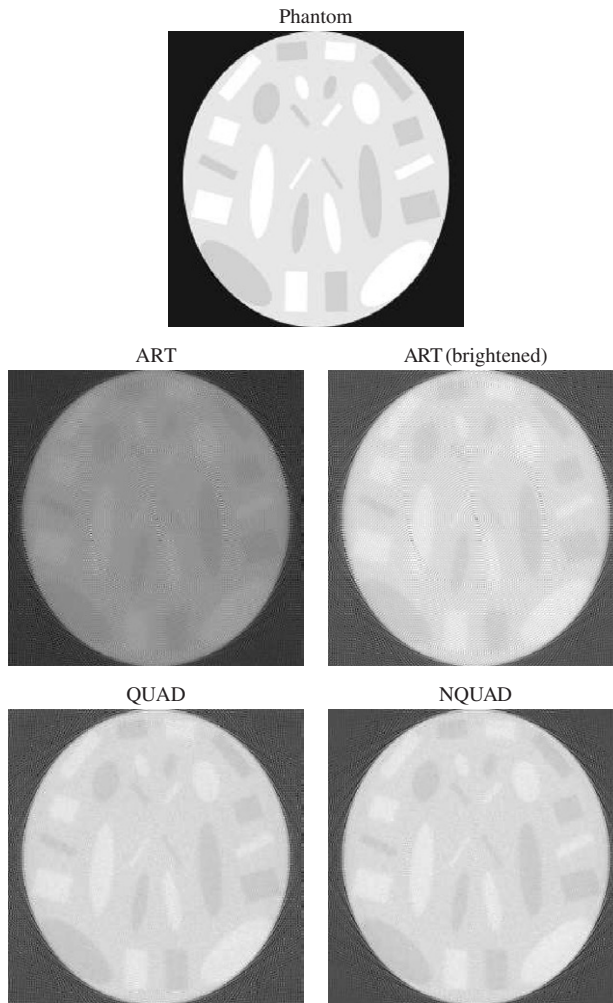


Figure 5. Case 2: phantom and reconstructed images at 10 iterations.

Cases 1 and 2, a natural question (raised by one of the reviewers) arises as to which of these two aspects is the primary cause of the different performance between ART and quadratic optimization. To this end, we examined two additional cases based on Cases 1 and 2:

- Case 3: An almost fully determined system of equations (as in Case 1), but with the same low contrast as Case 2. The phantom of this case can be seen in figure 7 below.
- Case 4: A strongly underdetermined system of equations (as in Case 2), but with an image of regular contrast. The phantom appears in figure 8 below.

Tables 4 and 5 show the minimal measures achieved by ART, QUAD, and NQUAD for Cases 3 and 4, respectively. From table 4 we can see that (similarly to Case 1) ART performs better on an almost fully-determined system, regardless of the fact that the phantom is of low contrast.

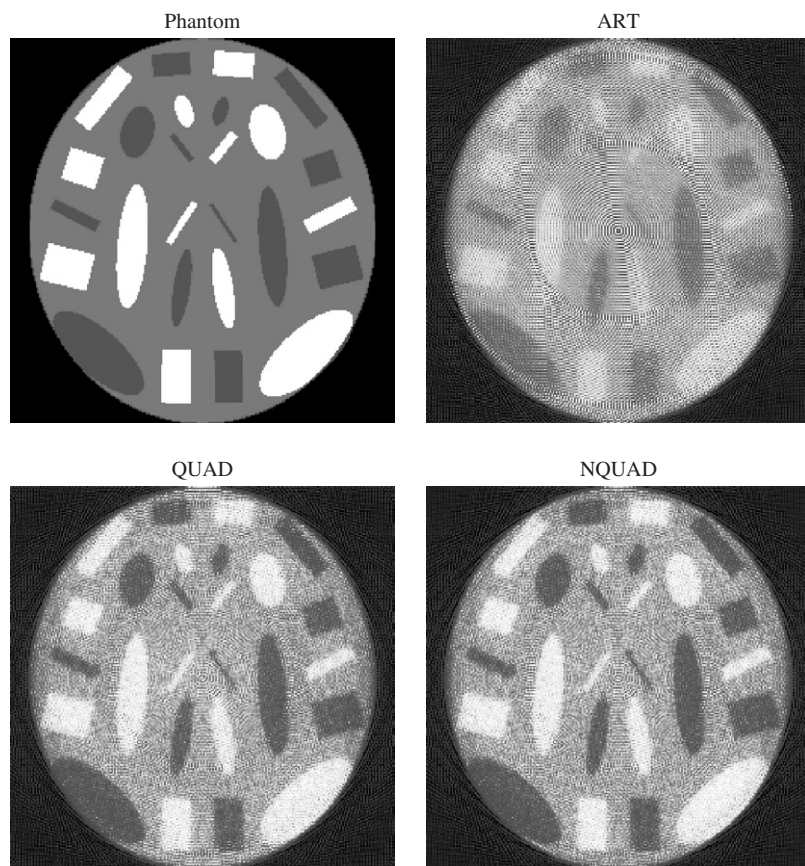


Figure 6. Case 2: phantom and reconstructed images (10 iterations) after histogram equalization.

Table 5 shows that in Case 4, with the strongly underdetermined system, both quadratic methods perform much better than ART, and that NQUAD is somewhat better than QUAD.

Figure 7 shows the phantom and reconstructed images for Case 3 after 10 iterations. The ART reconstruction appears “smoother” than the others, but differences between QUAD and NQUAD seem quite insignificant. Figure 8 shows the phantom and reconstructed images for Case 4 (strongly underdetermined system with high contrast). Both quadratic reconstructions are visibly better than that of ART, which shows some distinct artifact patterns. QUAD and NQUAD seem almost identical. However, all these differences are somewhat weaker than in Case 2 (which is of low contrast in addition to being strongly underdetermined).

Based on the convergence measures and the images, we can conclude that the major contributing factor to the superiority of the quadratic methods over ART in Cases 2 and 4 is the fact that the system of equations is strongly underdetermined. Also, under those conditions, NQUAD is somewhat better than QUAD. Furthermore, it is evident that the differences between the methods are more visibly pronounced when the data are of low contrast.

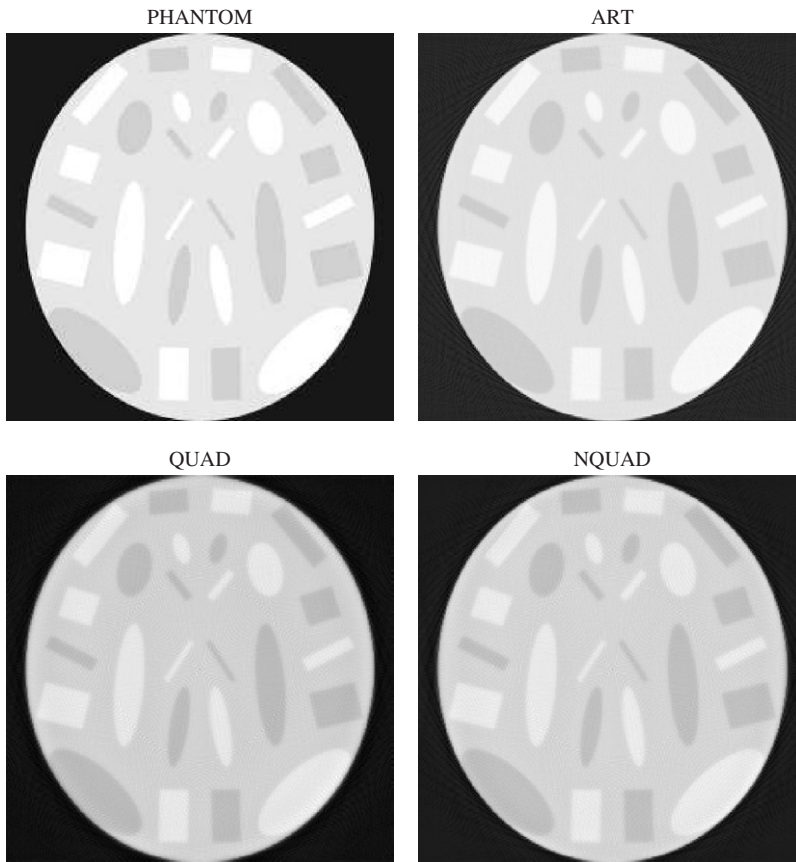


Figure 7. Case 3: phantom and reconstructed images after 10 iterations.

5. Conclusions and future work

This study introduced the concept of a geometric algorithm as a method (operating on a linear system) whose results are independent of any particular algebraic representation of the data. Any algorithm can be converted to a geometric one if the equations are normalized before the algorithm is applied. Quadratic optimization, in the context of image reconstruction from projections in CT, was studied by examining the QUAD algorithm. QUAD is the application of the conjugate gradient to the normal equations system derived from the column-normalization of the original system of equations. QUAD and its geometric/normalized form, NQUAD, were compared with ART (which is also a geometric algorithm) on various phantom test cases. The results can be summarized as follows:

- NQUAD was always better than QUAD, though the differences could not be visually detected when the equation system was almost fully determined.
- In an almost fully determined system, ART produced a better image than NQUAD during any given time period. However, given sufficient time, NQUAD eventually achieved the same quality.

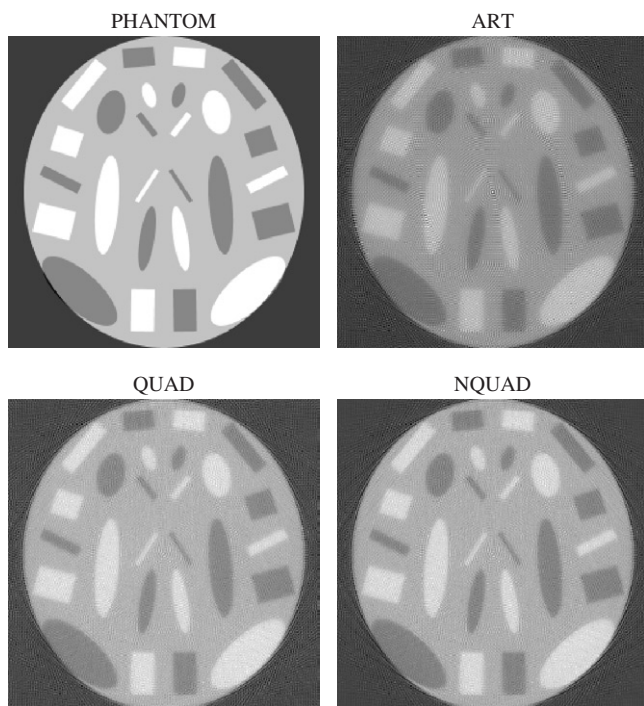


Figure 8. Case 4: phantom and reconstructed images after 10 iterations.

Table 4. Case 3—minimal distance and relative error, and iteration number(s) at which they were obtained.

| Measure | ART | QUAD | NQUAD |
|-----------|----------------|-------------------|-------------------|
| Distance | 0.0484, it. 10 | 0.0539, it. 16,17 | 0.0538, it. 14 |
| Rel. err. | 0.0176, it. 8 | 0.0236, it. 15 | 0.0230, it. 13,14 |

Table 5. Case 4—minimal distance and relative error, and iteration number(s) at which they were obtained.

| Measure | ART | QUAD | NQUAD |
|-----------|---------------|-------------------|-------------------|
| Distance | 0.3139 it. 20 | 0.2611 it. 19,20 | 0.2385, it. 18–20 |
| Rel. err. | 0.1712 it. 20 | 0.1387, it. 18–20 | 0.1256, it. 14–20 |

- In a strongly underdetermined system, both QUAD and NQUAD were much better than ART. NQUAD achieved its optimal error measures with significantly fewer iterations (and less time) than the other two and produced images that were somewhat better than those of QUAD.
- The advantages of the two quadratic methods over ART are even more pronounced when, in addition to the system of equations being strongly underdetermined, the underlying data is of low contrast.

The consequences are that NQUAD should always be preferred to QUAD, and that both ART and NQUAD should be readily available tools to use under different conditions.

These results suggest the following directions for future research, development, and experimentation:

- Examination of other methods with respect to their being geometric or not. If they are not geometric, they should be compared to their performance on the normalized equations.
- Our discretization model assumed that throughout a pixel, the attenuation value is constant. Lewitt [16,17] introduced a different mathematical model based on spherically symmetric volume elements, also known as “blobs”. The basis function of the pixel (or voxel in 3D) model has a value of 1 inside a unit square (cube in 3D) and 0 outside. The blob basis function is radially symmetric with a value of 1 at the center and smoothly and monotonically decreasing to zero with the distance from its center. See also [19] for further details. The normalization approach can also be applied to this model, and the results compared with the pixel model.
- Testing NQUAD in other settings that have proved to be problematic, such as cone-beam reconstructions of low-contrast objects—see [20].
- Incorporation of improved conjugate-gradient methods into image reconstruction software; see Gottlieb and Fischer [10].

Acknowledgments

The authors wish to thank Rachel Gordon for her help and advice. Thanks are also due to the anonymous reviewers for their helpful comments. Part of this research appeared in the second author’s MA thesis [18] carried out under the first author’s supervision at the University of Haifa.

References

- [1] Artzy, E., Elfving, T. and Herman, G.T., 1979, Quadratic optimization for image reconstruction, II. *Computer Graphics & Image Processing*, **11**, 242–261.
- [2] Björck, A. and Elfving, T., 1979, Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations. *BIT*, **19**, 145–163.
- [3] Browne, J.A., Herman, G.T. and Odhner, D., 1993, SNARK93: a programming system for image reconstruction from projections, Tech. Rep. MIPG198, The Medical Image Processing Group (MIPG), Dept. of Radiology, University of Pennsylvania.
- [4] Censor, Y., Eggermont, P.P.B. and Gordon, D., 1983, Strong underrelaxation in Kaczmarz’s method for inconsistent systems. *Numerische Mathematik*, **41**, 83–92.
- [5] Censor, Y., Gordon, D. and Gordon, R., 2001, BICAV: a block-iterative parallel algorithm for sparse systems with pixel-dependent weighting. *IEEE Transactions on Medical Imaging*, **20**(10), 1050–1060.
- [6] Censor, Y., Gordon, D. and Gordon, R., 2001, Component averaging: an efficient iterative parallel algorithm for large and sparse unstructured problems. *Parallel Computing*, **27**(6), 777–808.
- [7] Censor, Y. and Herman, G.T., 1987, On some optimization techniques in image reconstruction from projections. *Applied Numerical Mathematics*, **3**, 365–391.
- [8] Gordon, D., 1988, SNARKVIEW – user’s guide., Tech. Rep., University of Haifa.
- [9] Gordon, D. and Mansour, R., 2005, A geometric approach to quadratic optimization in computerized tomography. In: *8th IASTED International Conference on Computer Graphics & Imaging*, M.H. Hamza (Ed.), Acta Press, Calgary, Alberta, Canada, pp. 60–65.

- [10] Gottlieb, S. and Fischer, P.F., 1998, Modified conjugate gradient method for the solution of $A\underline{x} = \underline{b}$. *Journal of Scientific Computing*, **13**, 173–183.
- [11] Herman, G.T., 1980 *Image Reconstruction From Projections: The Fundamentals of Computerized Tomography* (New York: Academic Press).
- [12] Herman, G.T. and Lent, A., 1976, Quadratic optimization for image reconstruction, I. *Computer Graphics & Image Processing*, **5**, 319–332.
- [13] Herman, G.T., Lent, A. and Lutz, P.H., 1978, Relaxation methods for image reconstruction. *Communications of the ACM*, **21**, 152–158.
- [14] Hestenes, M.R. and Stiefel, E., 1952, Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, **49**, 409–436.
- [15] Kaczmarz, S., 1937, Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin de l'Académie Polonaise des Sciences et Lettres*, **A35**, 355–357.
- [16] Lewitt, R.M., 1990, Multidimensional digital image representations using generalized Kaiser–Bessel window functions. *Journal of the Optical Society of America A*, **7**(10), 1834–1846.
- [17] Lewitt, R.M., 1992, Alternatives to voxels for image representation in iterative reconstruction algorithms. *Physics in Medicine and Biology*, **37**, 705–716.
- [18] Mansour, R., 1999, Geometric least squares solutions in image reconstruction. Master's thesis, Dept. of Mathematics and Computer Science, University of Haifa, Israel.
- [19] Matej, S. and Lewitt, R.M., 1996, Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. *IEEE Transactions on Medical Imaging*, **15**, 68–78.
- [20] Mueller, K., Yagel, R. and Wheller, J.J., 1999, Anti-aliased three-dimensional cone-beam reconstruction of low-contrast objects with algebraic methods. *IEEE Transactions on Medical Imaging MI-18*, 519–537.
- [21] Saad, Y., 2003, *Iterative Methods for Sparse Linear Systems*, 2nd Edn (Philadelphia, PA: SIAM).
- [22] Tanabe, K., 1971, Projection method for solving a singular system of linear equations and its applications. *Numerische Mathematik*, **17**, 203–214.
- [23] Trummer, M.R., 1981, Reconstructing pictures from projections: on the convergence of the ART algorithm with relaxation. *Computing*, **26**, 189–195.