Journal of X-Ray Science and Technology 10 (2001) 59–66
IOS Press

Linear diophantine equations for discrete tomography

Yangbo Ye[a], Ge Wang[b] and Jiehua Zhu[a]
[a]*Department of Mathematics, The University of Iowa, Iowa City, IA 52242, USA*
[b]*Department of Radiology, The University of Iowa, Iowa City, IA 52242, USA*

Abstract. In this report, we present a number-theory-based approach for discrete tomography (DT),which is based on parallel projections of rational slopes. Using a well-controlled geometry of X-ray beams, we obtain a system of linear equations with integer coefficients. Assuming that the range of pixel values is $a(i, j) = 0, 1, \ldots, M - 1$, with M being a prime number, we reduce the equations modulo M. To invert the linear system, each algorithmic step only needs $\log_2^2 M$ bit operations. In the case of a small M, we have a greatly reduced computational complexity, relative to the conventional DT algorithms, which require $\log_2^2 N$ bit operations for a real number solution with a precision of $1/N$. We also report computer simulation results to support our analytic conclusions.

Keywords: Computed tomography (CT), discrete tomography (DT), number theory, Diophantine equations, computational complexity

1. Introduction

Computed tomography (CT) refers to computerized synthesis of an image from external measurements of a spatially varying function. Line integrals are the most common external measures, which are also known as projections. This underlying function is traditionally real-valued. However, the range of the underlying function is practically often discrete in industrial imaging and other applications. The problems of discrete tomography (DT) are to determine the underlying function from weighted sums over subsets of its domain. In the most essential sense, DT allows that given a discrete range of the underlying function we find its values that could not be found without the knowledge on the discrete range. Recently, DT attracts increasing interests, and becomes challenging and exciting.

Current studies (e.g. [3,6]) of DT are focused on algorithms derived from horizontal and vertical projections, plus sometimes diagonal projections. If the number of unknowns is more than the number of equations, these projections cannot determine a unique solution in the general case. Furthermore, existing algorithms do not directly make use of the discreteness of pixel values to reduce the computational complexity (see [1,7]). In this paper, we formulate a system of linear equations with integer coefficients, and develop a number-theory-based algorithm for DT. In the next section, we describe a pre-specified geometry for the formation of appropriate projections. In the third section, we propose a number-theory-based algorithm that can be applied to projection data collected according to the pre-specified geometry. Then, we present representative results obtained via computer simulation. In the last section, relevant issues are discussed for futher research, and conclusions are made.

2. Projection formation

Suppose an image $a(x,y)$ is divided into a square of grid $n \times n$. The total number of cells is n^2. Assume each cell has size of 1×1 and in each cell $a(x,y)$ takes a constant value. For $1 \leqslant i,j \leqslant n$, let us denote by $a(i,j)$ the constant function value in the (i,j)th cell, which is the cell enclosed by the lines: $x = i-1$, $x = i$, $y = j-1$ and $y = j$.

It is easy to have horizontal and vertical projections. They are rays of width 1. The horizontal projections give n equations:

$$b_{0/1}(j) = \sum_{i=1}^{n} a(i,j), \qquad j = 1,\dots,n, \tag{1}$$

where $b_{0/1}(j)$ is the density detected by the ith ray. Here the subscript $0/1$ denotes the slope of the rays. Similarly from the vertical projections we get

$$b_{1/0}(i) = \sum_{j=1}^{n} a(i,j), \qquad i = 1,\dots,n, \tag{2}$$

where $1/0$ represents the vertical slope.

There are two diagonal directions, one with slope $1/1$ and the other with slope $-1/1$. We take the width of the rays equal to $1/\sqrt{2}$ and divide neighboring rays by lines $y = x+k$ for slope$= 1/1$, and by $y = -x+k$ for slope$= -1/1$, for $k \in \mathbb{Z}$. Note that each ray passes through $1/2$ area of a cell. Therefore the diagonal projections with slope $1/1$ are

$$b_{1/1}(k) = \sum_{j=0}^{n} \left(\frac{1}{2}a(j,j+k) + \frac{1}{2}a(j+1,j+k)\right), \qquad -n < k \leqslant n, \tag{3}$$

where we set $a(i,j) = 0$ if (i,j) is outside the range $1 \leqslant i,j \leqslant n$. Similarly from the other diagonal direction we get equations

$$b_{-1/1}(k) = \sum_{j=0}^{n} \left(\frac{1}{2}a(j,n-j-k+1) + \frac{1}{2}a(j+1,n-j-k+1)\right), \qquad -n < k \leqslant n. \tag{4}$$

The key idea of our new algorithm is to have projections in other directions, and to do so in such a way that the resulting equations have integral coefficients. In this sense, our projections are different from those in [1,8]. Let p/q be a positive rational number in its reduced form. Then the greatest common divisor $(p,q) = 1$. Assume that $p > q > 0$. We want to construct parallel rays of slope p/q, by dividing our square grid into stripes using parallel lines $y = (p/q)(x-j)$, $-n < j \leqslant n$. Then each stripe and hence each ray have width $p/\sqrt{p^2+q^2}$ which is less than 1 but greater than $1/2$, because $p > q > 0$.

Such a ray passes through several cells. By computing the area of each cell (i,j) it passes through, we can determine the coefficient of $a(i,j)$ in the equation for this ray. The following theorem summarizes our results. Here we denote by $[x]$ the integral part of a real number x, that is, the largest integer $\leqslant x$. Denote by $\{x\}$ the fractional part of x, that is, $\{x\} = x - [x]$.

Theorem 1. *Let p and q be positive integers with $(p,q)=1$ and $p>q$. The equations from projections of slope p/q are*

$$\begin{aligned}
b_{p/q}(j) = &\sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}=0}} \left(\frac{q}{2p}a(j+ql/p,l) + \left(1-\frac{q}{2p}\right)a(j+ql/p+1,l)\right)\\
&+ \sum_{\substack{1\leqslant l\leqslant n,\\ 0<\{ql/p\}<q/p}} \left(\left(\frac{q}{2p} - \left\{\frac{ql}{p}\right\} + \frac{p}{2q}\left\{\frac{ql}{p}\right\}^2\right)a(j+[ql/p],l)\right.\\
&+ \left(1-\frac{q}{2p}+\left\{\frac{ql}{p}\right\} - \frac{p}{q}\left\{\frac{ql}{p}\right\}^2\right)a(j+[ql/p]+1,l)\\
&\left.+\frac{p}{2q}\left\{\frac{ql}{p}\right\}^2 a(j+[ql/p]+2,l)\right)\\
&+ \sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}\geqslant q/p}} \left(\left(1+\frac{q}{2p}-\left\{\frac{ql}{p}\right\}\right)a(j+[ql/p]+1,l)\right.\\
&\left.+\left(\left\{\frac{ql}{p}\right\}-\frac{q}{2p}\right)a(j+[ql/p]+2,l)\right),
\end{aligned} \tag{5}$$

for $-n<j\leqslant n$.

We will prove this Theorem in the next section. Here we point out that the coefficients are all rational numbers, and indeed, if multiplied by $2pq$, they become integers. This means that all our Eqs (1) through (5), together with Eqs (6) to (8) below, can be rewritten as linear equations of integral coefficients. If we seek integer solutions $a(i,j)$, we are dealing with a system of linear Diophantine equations. If the cells are supposed to take values $a(i,j)=0,1,\ldots,M-1$, we can further reduce these linear Diophantine equations modulo M.

By simple reflections or change of coordinates, we can get equations for projections of slopes q/p, $-p/q$, and $-q/p$.

Corallary 1. *Let p and q be positive integers with $(p,q)=1$ and $p>q$. The equations from projections of slope q/p are*

$$\begin{aligned}
b_{q/p}(j) = &\sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}=0}} \left(\frac{q}{2p}a(l,j+ql/p) + \left(1-\frac{q}{2p}\right)a(l,j+ql/p+1)\right)\\
&+ \sum_{\substack{1\leqslant l\leqslant n,\\ 0<\{ql/p\}<q/p}} \left(\left(\frac{q}{2p} - \left\{\frac{ql}{p}\right\} + \frac{p}{2q}\left\{\frac{ql}{p}\right\}^2\right)a(l,j+[ql/p])\right.\\
&+ \left(1-\frac{q}{2p}+\left\{\frac{ql}{p}\right\} - \frac{p}{q}\left\{\frac{ql}{p}\right\}^2\right)a(l,j+[ql/p]+1)
\end{aligned}$$

$$+\frac{p}{2q}\left\{\frac{ql}{p}\right\}^2 a(l, j+[ql/p]+2)\Bigg) \tag{6}$$

$$+\sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}\geqslant q/p}} \left(\left(1+\frac{q}{2p}-\left\{\frac{ql}{p}\right\}\right) a(l, j+[ql/p]+1) + \left(\left\{\frac{ql}{p}\right\}-\frac{q}{2p}\right) a(l, j+[ql/p]+2)\right),$$

for $-n < j \leqslant n$.

Corallary 2. *Let* p *and* q *be positive integers with* $(p, q) = 1$ *and* $p > q$*. The equations from projections of slope* $-p/q$ *are*

$$\begin{aligned}
b_{-p/q}(j) = &\sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}=0}} \left(\frac{q}{2p}a(n+1-j-ql/p, l) + \left(1-\frac{q}{2p}\right) a(n-j-ql/p, l)\right)\\
&+\sum_{\substack{1\leqslant l\leqslant n,\\ 0<\{ql/p\}<q/p}} \left(\left(\frac{q}{2p}-\left\{\frac{ql}{p}\right\}+\frac{p}{2q}\left\{\frac{ql}{p}\right\}^2\right) a(n+1-j-[ql/p], l)\right.\\
&+\left(1-\frac{q}{2p}+\left\{\frac{ql}{p}\right\}-\frac{p}{q}\left\{\frac{ql}{p}\right\}^2\right) a(n-j-[ql/p], l)\\
&\left.+\frac{p}{2q}\left\{\frac{ql}{p}\right\}^2 a(n-j-[ql/p]-1, l)\right)\\
&+\sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}\geqslant q/p}} \left(\left(1+\frac{q}{2p}-\left\{\frac{ql}{p}\right\}\right) a(n-j-[ql/p], l)\right.\\
&\left.+\left(\left\{\frac{ql}{p}\right\}-\frac{q}{2p}\right) a(n-j-[ql/p]-1, l)\right),
\end{aligned} \tag{7}$$

for $-n < j \leqslant n$.

Corallary 3. *Let* p *and* q *be positive integers with* $(p, q) = 1$ *and* $p > q$*. The equations from projections of slope* $-q/p$ *are*

$$\begin{aligned}
b_{-q/p}(j) = &\sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}=0}} \left(\frac{q}{2p}a(l, n+1-j-ql/p) + \left(1-\frac{q}{2p}\right) a(l, n-j-ql/p)\right)\\
&+\sum_{\substack{1\leqslant l\leqslant n,\\ 0<\{ql/p\}<q/p}} \left(\left(\frac{q}{2p}-\left\{\frac{ql}{p}\right\}+\frac{p}{2q}\left\{\frac{ql}{p}\right\}^2\right) a(l, n+1-j-[ql/p])\right.
\end{aligned}$$

$$
\begin{aligned}
&+\left(1-\frac{q}{2p}+\left\{\frac{ql}{p}\right\}-\frac{p}{q}\left\{\frac{ql}{p}\right\}^2\right)a(l,n-j-[ql/p])\\
&+\frac{p}{2q}\left\{\frac{ql}{p}\right\}^2 a(l,n-j-[ql/p]-1)\Bigg)\\
&+\sum_{\substack{1\leqslant l\leqslant n,\\ \{ql/p\}\geqslant q/p}}\left(\left(1+\frac{q}{2p}-\left\{\frac{ql}{p}\right\}\right)a(l,n-j-[ql/p])\right.\\
&+\left.\left(\left\{\frac{ql}{p}\right\}-\frac{q}{2p}\right)a(l,n-j-[ql/p]-1)\right),
\end{aligned}
\tag{8}
$$

for $-n < j \leqslant n$.

3. Algorithm and simulation

We select parallel rays whose slopes are rational numbers of the form $\pm p/q$, and control the width and location of each ray in a specific way. The resulting linear equations then have all coefficients being integers. Because of these integral coefficients, we may reduce this system of linear Diophantine equations modulo M, if the cells are supposed to take values $a(i,j) = 0, 1, \ldots, M-1$. In other words, we seek solutions in the field $\mathbb{Z}/M\mathbb{Z}$, when M is a prime number.

This linear system can be solved over $\mathbb{Z}/M\mathbb{Z}$ in the same way as over $\mathbb{R}$, using the standard Gaussian elimination method, for example. The main advantage of our approach is that, over $\mathbb{Z}/M\mathbb{Z}$, each operation uses only $\log_2^2 M$ bit operations (cf. Cohen [2]). On the other hand, for solution procedures of a linear system over $\mathbb{R}$, each operation needs $\log_2^2 N$ bit operations for a real number solution of precision $1/N$. For small M, the time $\log_2^2 M$ represents a big saving of computation time over the usual $\log_2^2 N$ time.

As an example, let us look at the case of $M = 2$, i.e., the case when every cell can only take value 0 or 1. Now $\mathbb{Z}/2\mathbb{Z}$ is the finite field of two elements 0 and 1. Operations in $\mathbb{Z}/2\mathbb{Z}$ are fast: $0 \times 0 = 0$, $0 \times 1 = 0$, $1 \times 1 = 1$, $0 + 0 = 0$, $0 + 1 = 1$, $1 + 1 = 0$, etc. Each operation only need $\log_2^2 2 = 1$ bit operation.

On the other hand, the solution procedure for the same linear system is much slower if we seek a real number solution to the precision of, let us say, $0.01 = 1/100$. With $N = 100$, one operation here need $\log_2^2 100 \simeq 44$ bit operations. In other words, our new algorithm uses only $1/44$ of the computation time of a standard CT algorithm. This huge saving of computation time was observed in a computation of a 30×30 DT procedure using Matlab. For our new modulo 2 algorithm, the CPU time is significantly faster than using a standard algorithm seeking real number solutions.

The linear equations are given in Section 2 from Eqs (1) to (8). In these equations, Eqs (1) and (2) are derived from horizontal and vertical projections and are of integral coefficients. Equations (3) and (4) are from diagonal projections; we need to multiply them by a factor 2 to get the equations of integral coefficients we want. Note that if the cells are supposed to take integral values $a(i,j) = 0, 1, \ldots, M-1$, $b_{1/1}(k)$ and $b_{-1/1}(k)$ will be integers or half integers. Consequently after multiplying them by 2, Eqs (3) and (4) become linear equation of integral coefficients. Similarly, we can multiply Eqs (5) to (8) by $2pq$ and change them to linear equations of integral coefficients.

In order to get enough equations from Eqs (5) through (8), we may fix a positive integer L and select all p and q with $1 \leqslant q < p \leqslant L$ and $(p,q) = 1$. These fractions q/p, together with 0 and 1, form a Ferry sequence of level L. There are about $(3/\pi^2)L^2$ such pairs of p and q, which produce about $(12/\pi^2)L^2$ directions: $\pm p/q$ and $\pm q/p$. In each of these directions we can get more than n but less than $2n$ rays. Since we need n^2 linearly independent equations, we may set $(12/\pi^2)L^2 n = n^2$. This means that L may be set to equal to $(\pi/\sqrt{12})\sqrt{n}$. For example, if $n = 256$, we may set $L = 14$.

4. Discussions and conclusion

Much efforts are needed to optimize imaging protocols and reconstruction algorithms. Clearly, image quality would depend on domain knowledge, imaging geometry, data completeness, noise level, and algorithmic details. Extensive simulation and physical experiments must be performed in selected cases to gain some insight into interplays among all the major factors.

It is ackowledged that the most popular imaging geometry is fan-beam-oriented, instead of parallel-beam-oriented. However, our findings are still of significance, because we can always rebin fan-beam projection data to parallel-beam projection data in the format we prefer. This rebinning process is always much faster than the reconstruction process. Furthermore, the rebinning is even a necessary intermediate step in spiral/helical CT, in which an X-ray source rotation and an object translation are performed simultaneously. We believe that fan-beam DT should be a most practical problem to solve in near future.

Although we have not addressed data inconsistence in this proof-of-concept report, we point out that noise in data can be suppressed in the same spirit of the conventional linear system solution algorithms. On the other hand, given a specific knowledge on the discrete range of the underlying function, more efficient algorithms may be designed for various DT applications. We are particularly interested in pursuing research along a number-theory direction. Relevant results will reported in future publications.

In conclusion, we have proposed a number-theory-based approach for discrete tomography (DT), and demonstrated that our method can greatly reduce the computational complexity. Further work is underway to generalize the findings, refine the algorithm and apply it to various applications.

Appendix. Proof of Theorem 1.

Suppose the slope of a ray is p/q, where $(p,q) = 1$, $p > q > 0$, and the ray is bounded by $y = (p/q)x$ and $y = (p/q)(x-1)$. Then this ray passes through cells in three different cases.

The first case is that there exists an integer $m \in \mathbb{Z}$ such that $(l-1)q/p < m < ql/p$. This condition is equivalent to $0 < \{ql/p\} < q/p$. In this case, the ray runs through three cells on level l with unknowns $a([ql/p], l)$, $a([ql/p]+1, l)$, and $a([ql/p]+2, l)$. The areas passed by the ray for the three cells are

$$w_{[ql/p],l} = \frac{1}{2}\left(\left[\frac{ql}{p}\right] - \frac{q}{p}(l-1)\right)\left(\frac{p}{q}\left[\frac{ql}{p}\right] - (l-1)\right)$$
$$= \frac{q}{2p} - \left\{\frac{ql}{p}\right\} + \frac{p}{2q}\left\{\frac{ql}{p}\right\}^2,$$

$$w_{[ql/p]+1,l} = 1 - \frac{p}{2q}\left(\frac{q}{p} - \left\{\frac{ql}{p}\right\}\right)^2 - \frac{1}{2}\left(\frac{q}{p}l - \left[\frac{q}{p}l\right]\right)\left(l - \frac{p}{q}\left[\frac{q}{p}l\right]\right)$$

$$= 1 - \frac{q}{2p} + \left\{\frac{ql}{p}\right\} - \frac{p}{q}\left\{\frac{ql}{p}\right\}^2,$$

$$w_{[ql/p]+2,l} = \frac{1}{2}\left(\frac{ql}{p} + 1 - \left[\frac{q}{p}l\right] - 1\right)\left(l - \frac{p}{q}\left[\frac{ql}{p}\right]\right)$$
$$= \frac{p}{2q}\left\{\frac{ql}{p}\right\}^2.$$

Thus the projection over these three cells produces three terms for the linear equation

$$\begin{aligned}&\left(\frac{q}{2p} - \left\{\frac{ql}{p}\right\} + \frac{p}{2q}\left\{\frac{ql}{p}\right\}^2\right) a([ql/p], l)\\ &+ \left(1 - \frac{q}{2p} + \left\{\frac{ql}{p}\right\} - \frac{p}{q}\left\{\frac{ql}{p}\right\}^2\right) a([ql/p]+1, l) + \frac{p}{2q}\left\{\frac{ql}{p}\right\}^2 a([ql/p]+2, l).\end{aligned} \tag{A1}$$

The second case is that $ql/p \in \mathbb{Z}$, i.e., $\{ql/p\} = 0$. In this case the ray passes through two cells on level l. The coefficients of the unknowns $a(ql/p, l)$ and $a(ql/p + 1, l)$ are $w_{ql/p,l} = q/2p$ and $w_{ql/p+1,l} = 1 - q/2p$, respectively. This case therefore contributes two terms to the linear equation

$$\frac{q}{2p} a(ql/p, l) + \left(1 - \frac{q}{2p}\right) a(ql/p + 1, l). \tag{A2}$$

The third case is when $\{ql/p\} \geqslant q/p$. In this case, the ray runs through two cells on level l and the corresponding coefficients are:

$$w_{l,[ql/p]+1} = \frac{1}{2}\left(\left[\frac{ql}{p}\right] + 1 - \frac{ql}{p} + \left[\frac{ql}{p}\right] + 1 - \frac{q}{p}(l-1)\right)$$
$$= \frac{q}{2p} + 1 - \left\{\frac{ql}{p}\right\}$$

$$w_{l,[ql/p]+2} = 1 - \left(\frac{q}{2p} + 1 - \left\{\frac{ql}{p}\right\}\right) = \left\{\frac{ql}{p}\right\} - \frac{q}{2p}.$$

Hence the projection over these two cells has the following two terms

$$\left(\frac{q}{2p} + 1 - \left\{\frac{ql}{p}\right\}\right) a([ql/p]+1, l) + \left(\left\{\frac{ql}{p}\right\} - \frac{q}{2p}\right) a([ql/p]+2, l). \tag{A3}$$

Collecting Eqs (A1), (A2), and (A3), we prove Theorem 1.

References

[1] M.T. Chan, G.T. Herman and E. Levitan, Bayesian image reconstruction using image-modeling Gibbs priors, *Int. J. Imaging Syst. Tech.* **9** (1998), 85–98.

[2] H. Cohen, *A course in computational algebraic number theory*, Springer, Berlin, 1996.

[3] A. Fazekas, G.T. Herman and S. Matej, On processing binary pictures via their projections, *Int. J. Imaging Syst. Tech.* **9** (1998), 99–100.

[4] G. Gritzmann, D. Prangenberg, S.S. de Vries and M. Wiegelmann, Success and failure of certain reconstruction and uniqueness algorithms in discrete tomography, *Int. J. Imaging Syst. Tech.* **9** (1998), 101–109.

[5] J.H.B. Kamperman and A. Kuba, Reconstruction of two-valued matrices from their two projections, *Int. J. Imaging Syst. Tech.* **9** (1998), 110–117.

[6] A. Del Lungo and M. Nivat, Reconstruction of connected sets from two projections, in: *Discrete Tomography Foundations, Algorithms, and Applications,* G.T. Herman and A. Kuba, eds, Birkhäuser, Boston, 1999, pp. 163–188.

[7] Y. Vardi and C.-H. Zhang, Reconstruction of binary images vie the EM algorithm, in: *Discrete Tomography Foundations, Algorithms, and Applications,* G.T. Herman and A. Kuba, Birkhäuser, Boston, 1999, pp. 297–316.

[8] A.E. Yagle, An explicit closed-form solution to the limited-angle discrete tomography problem for finite-support objects, *Int. J. Imaging Syst. Tech.* **9** (1998), 174–180.