33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Efficient and Effective Regularised ART for Computed Tomography

Alexey Buzmakov, Dmitry Nikolaev, Marina Chukalina and Gerald Schaefer

Abstract— **Algebraic Reconstruction Technique (ART) is a widely employed method in computed tomography since it has certain advantages, such as allowing reconstruction of data with missing projections in some angle ranges, over other techniques such as Filtered Back Projection (FBP). Recently, a regularisation technique for ART, RegART, was introduced which provides greatly reduced noise levels. However, a serious drawback of both ART and RegART is the computational complexity of the methods. In this paper, we present a fast version of RegART, which makes use of nVidia's CUDA technology, and show that this approach performs favourably compared to FBP.**

I. INTRODUCTION

Algebraic reconstruction technique (ART) [1] of tomographic data is an iterative way of reconstructing object spatial structure from a set of X-ray transmission projections collected at different angles of object rotation. ART approaches have several key advantages over transform-based methods. They can be used with irregular sampling geometries, incomplete noisy data sets and may incorporate curved ray paths [2,3]. Also, the application of algebraic techniques allows for better quality of tomography reconstruction data. However, one main disadvantage of ART-based techniques is that they are computationally rather expensive in comparison to integral reconstruction methods.

In [4], we introduced RegART, a regularised version of ART, which reduces the noise of reconstruction by introducing a spatial non-linear filtering stage between the iterations of the algorithm.

In this paper, we present a fast version of RegART which makes use of nVidia's CUDA technology for improved efficiency, and show that RegART performs favourably compared to Filtered Back Projection. Our system is developed and our experiments carried out at the A.V. Shubnikov Institute of Crystallography RAS [5]. Out of the several tomography data collection schemes [6], in our experiments the parallel scanning scheme (Fig.1) has been used; however RegART can be easily adapted to other scanning schemes.

This work was partly supported by RFBR under grants 09-07-00473-a and 09-02-12251-ofi_m, and by PICS 3470.

A Buzmakov is with the A.V. Shubnikov Institute of Crystallography RAS, Moscow, Russia.

D. Nikolaev is with the Institute for Information Transmission Problems RAS, Moscow, Russia.

M. Chukalia is with the Institute of Microelectronics Technology RAS, Chernogolovka, Russia

G. Schaefer is with the Department of Computer Science, Loughborough University, Loughborough, U.K.

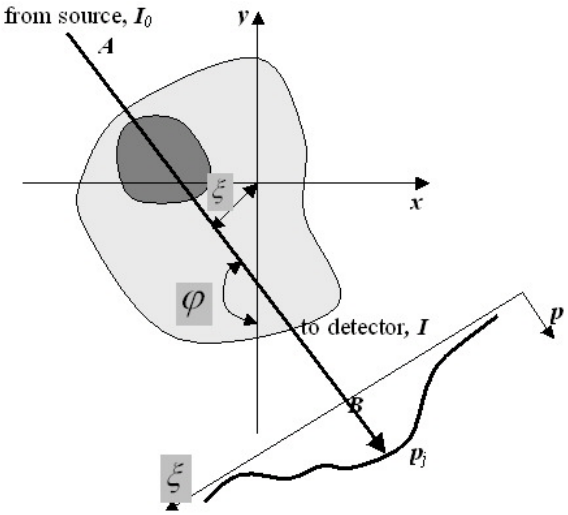

Fig. 1. Parallel scanning scheme.

II. ART AND REGART

Assuming a Cartesian co-ordinate system to describe projection formation, the line equation for *AB* in Fig.1 is

$$x\cos\varphi + y\sin\varphi = \xi \tag{1}$$

Let $f(x,y)$ describe the linear attenuation coefficient. Then, the transmission function of the fine beam *AB* is

$$I(\varphi,\xi) = I_0(\varphi,\xi)\exp(-\iint dxdyf(x,y)\delta(x\cos\varphi + y\sin\varphi - \xi)), \tag{2}$$

where I_0 is the intensity of the initial beam and δ is the Dirac delta-function.

Usually, a new function

$$p(\varphi,\xi) = \ln\left(\frac{I_0(\varphi,\xi)}{I(\varphi,\xi)}\right) \tag{3}$$

is introduced so that we get

$$p(\varphi,\xi) = \iint dxdyf(x,y)\delta(x\cos\varphi + y\sin\varphi - \xi). \tag{4}$$

which is known as the Radon transform of $f(x,y)$. The parallel projection is a collection of fine beam integrals for a constant φ.

For ART, a square grid is imposed on the image $f(x,y)$.

The function $f(x,y)$ is constant in each pixel, consequently we search the solution in the space of the piecewise constant functions. Let f_i denote a constant value in the *i*-th pixel and *N* be the total number of pixels. X-rays are lines running through the image plane (see Fig. 2), and we assume that the ray width is approximately equal to the pixel size.

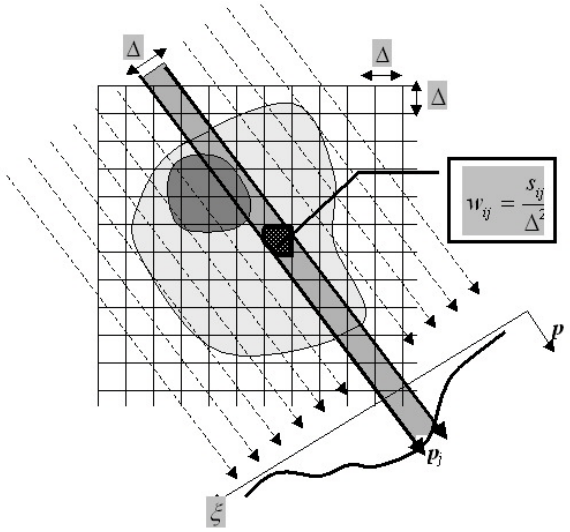

Fig. 2. Parallel scanning scheme (discrete representation).

The intensity p_j is called a ray sum. The relation between p_j and $\vec{f}$ is expressed as

$$p_j = \sum_{i=1}^{N} f_i w_{ij}, j = 1,...M \tag{5}$$

where *M* is the total number of rays (in all projections) and w_{ij} is the weighting factor that represents the contribution of the *i*-th pixel to the *j*-th ray sum.

For large *N* and *M* there exist iterative methods to solve the equation system (5). These are based on the "method of projections" first proposed by Kaczmarz [7]. An image, presented by $\vec{f}$, may be considered to be a single point in an *N*-dimensional space. Each of the linear equations of (1) defines a hyperplane. The unique solution to these equations is the intersection of all hyperplanes.

Let $\vec{f}^k$ be the estimated solution at the *k*-th iteration. The iteration scheme is represented by

$$\vec{f}^{k+1} = \vec{f}^k + \gamma \frac{p_j - \left(\vec{f}^k, \vec{w}_j\right)}{\left(\vec{w}_j, \vec{w}_j\right)} \vec{w}_j. \tag{6}$$

Here γ is the so-called relaxation parameter [8]. It can be shown [9] that the limits of cyclic sub-sequences generated by the method reduce to a weighted least squares solution of the system when the relaxation parameter approaches zero. This point minimises the sum of squares of Euclidean distances to the hyperplanes of the system.

The initial estimate denoted by $\vec{f}^0$ is assigned a value of zero, and it can be shown [10] that from any initial estimate the sequence generated by ART converges to a weighted least squares solution. The initial estimate is projected onto the hyperplane represented by the first equation in (6) to yield $\vec{f}_1^0$. The subscript here indicates how many hyperplanes are included in the $\vec{f}^0$ correction process. After each projection to a hyperplane, the estimated image $\vec{f}^0$ is updated. The first sub-iteration is finished once the correction over all hyperplanes has been performed.

One can note that $w_j \in \{0,1\}$ for $\varphi = 0$ and moreover that equation (5) may be rewritten as

$$p_j = \sum_{i=1}^{n} f_{ij}, \tag{7}$$

where f_{ij} denotes the *i*-th pixel in the *j*-th row and *n* is the size of the image. This allows us to use the following algorithm of projection calculation: for each rotation angle φ rotate image f_{ij} by this angle (e.g., using a bilinear approximation algorithm). After that apply equation (7) to obtain the projection.

Many projection access schemes have been discussed in the literature [11]. To minimise the influence of two neighbouring hyperplanes on each other we used the following scheme:

$$\begin{aligned} &p(\varphi_1,\xi_1), p\left(\varphi_1+\frac{\pi}{2},\xi_1\right), p(\varphi_1,\xi_2), p\left(\varphi_1+\frac{\pi}{2},\xi_2\right),... \\ &p\left(\varphi_1,\xi_{N/2}\right), p\left(\varphi_1+\frac{\pi}{2},\xi_{N/2}\right),...p\left(\varphi_M+\frac{\pi}{2},\xi_{N/2}\right). \end{aligned} \tag{8}$$

Because the projections are usually noisy, the intersection of the hyperplanes is not a point in the *N*-dimensional space but a polyhedron. Each iteration projects the estimated solution to a polyhedron hull area. On the other hand, the solution sought for belongs to the image class sub-space. The size, shape and position of the sub-space depend on the accuracy of the image description (accuracy of the image model). The image sub-space and the polygon can intersect or be close to each other. A regularisation operator brings the estimated solution from the polygon wall area to the image sub-space [12]. The space of piecewise constant functions is well suited for the description of tomography images. However, it is rather difficult to construct a projector which brings an estimated solution to this image sub-space. We have taken the space of piecewise smooth functions as the image space. That is, if the function belongs to this space it will belong to the same space after the median operator was implemented. Then the median filter operator [13, 4] can be used as the

projector from the polygon wall area to the image sub-space. We apply median filtering as the second sub-iteration. It is known that the median filter reduces speckle noise and salt and pepper noise, while its edge-preserving nature makes it useful in cases where edge blurring is undesirable. It should be noted that the type of projector depends on the chosen image sub-space and in the general case can be of any kind. The non-negativity constraint is reinforced when instead of $f < 0$ we set $f = 0$.

One iteration is completed once the full set of measurements has been processed. In the next iteration, $\vec{f}^{k}$ is projected onto the hyperplane represented by the first equation in (6), and successively onto the rest of the hyperplanes in (6). Then, filtering is applied and so on until the last iteration.

In the last iteration, all images $\vec{f}_1^{last},...\vec{f}_{N*M}^{last}$ are saved. The final step of the algorithm is the averaging over these images to exclude the influence of the last hyperplane projection.

III. RegART on CUDA

The main bottleneck in the implementation of the RegART algorithm is image rotation which is performed n_φ times during each iteration. We therefore employ nVidia's CUDA technology [14] to improve the performance of the method. CUDA (as well as ATI Stream and similar technologies) is a development kit that allows a large part of computationally intensive calculations to be propagated to a graphical processor (GPU). In our approach, we employ a rotation technique involving 2D texture fetching [15] using CUDA to reduce memory bandwidth. The bilinear rotation algorithm includes two almost independent stages for each pixel of the destination image: calculation of its exact co-ordinates on the source image, and bilinear interpolation of the value around this point. The used CUDA texture fetching mechanism allows implementation of the second stage on hardware level completely transparent for programmer. Moreover, texture memory is cached whereas general read-write memory is not.

IV. Performance Comparison

In Table 1 we present a performance comparison for image rotation with CUDA and a plain C implementation, run on a PC with 2 AMD Opteron 275, 8GB memory, nVidia GTX 285 with 240 cores, and Ubuntu Linux 9.10 64-bit as operating system. All computations were performed using double precision floating point numbers.

As the typical linear size of our tomography images is about 1000 pixels, we can expect a speed improvement of a factor of about 6. However, since the other steps of the algorithm (iterative modification and filtering of the image) still use the central processor, requiring memory transfers between CPU and GPU on each iteration, the overall speed-up is somewhat lower than those listed in the table (about 3).

TABLE I
SPEEDUP RESULTS ACHIEVED USING CUDA

Image size	CUDA / C speed-up
200x200	2.3
300x300	2.9
500x500	3.9
800x800	5.6
1000x1000	6.0
1500x1500	6.6

V. Reconstruction Quality Comparison

A reconstruction quality comparison between RegART and Filtered Back Projection (FBP) [16] was performed using the Shepp-Logan phantom which is the de-facto standard in computed tomography. Projections were modelled using normally distributed noise with a half-width about 1% of maximal signal (see Fig. 3). A 3x3 median filter was used as regularisation transform in RegART.

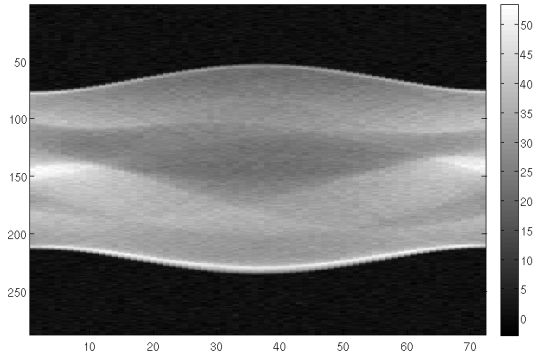

Fig.3. Simulated set of projections. 72 projection angles.

In Fig. 4, one scanline of the reconstructed phantome is presented. The dotted line is the result of the FBP method, the dashed line that of RegART, whereas the solid line represents the ground truth. It is easy to see that RegART is superior to FBP, showing far fewer fluctuations of the absorbtion coefficent.

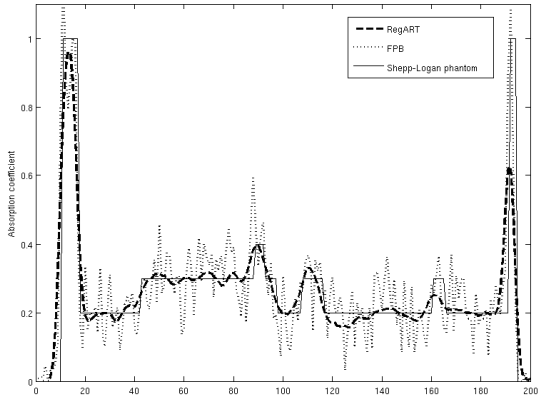

Fig.4. Reconstruction results.

VI. Conclusions

Speeding up of the RegART algorithm allows its use in real tomography hardware, which are currently often based on FBP, to provide less noisy images. However, to further compete with FBP in terms of computational efficiency, it ideally should be faster still. In order to achieve that, we are planning to optimise the speed of this procedure by incorporating a fast Hough transform calculation [17] which will reduce the complexity from $O(n^3)$ using rotations to $O(n^2 \log n)$.

References

[1] Gordon, R. A tutorial on ART (algebraic reconstruction technique). *IEEE Trans. Nuclear Science*, vol. 21, pp. 78-93, 1974.

[2] Wan, X., Gao, Y., Wang, Q., Le, S. and Yu, S. Limited-angle optical computed tomography algorithms. *Optical Engineering*, vol. 42, no.9, pp. 2659-2669, 2003.

[3] Schubert, F. Basic principles of acoustic emission tomography, *Journal of Acoustic Emission*, vol. 22, pp. 147-158, 2004.

[4] Chukalina, M., Nikolaev, D., and Simionovici, A. ART in X-ray tomography: Image noise reduction. *21th European Conference on Modelling and Simulation*, pp. 309-312, 2007.

[5] Asadchikov, V., Babak, V., Buzmakov, A. *et al.* X-ray difractometer with a source-detector moving system. *Engineering and Experimental Devices*, vol. 3, pp. 1-9, 2005 (in Russian).

[6] Natterer, F. Mathematical aspects of computerized Tomography. *Springer*, 1981.

[7] Kaczmarz, S. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres*, series A, vol. 35, pp. 335-357, 1937.

[8] Ros, D., Falcon, C., Juvells, I. and Javia, J. The influence of a relaxation parameter on SPECT iterative reconstruction algorithms. *Phys. Med. Biol*., vol. 41, no.5, pp. 925-937, 1996.

[9] Censor, Y., Eggermont, P.B.P. and Gordon, D. Strong underrelaxation in Kaczmarz's method for inconsistent systems. *Numerische Mathematik*, vol. 41, no. 1, pp. 83-92, 1983.

[10] Ming J. and Ge, W. Convergence of the simultaneous algebraic reconstruction technique (SART)." *IEEE Trans. Image Processing*, vol. 12, no. 8, pp. 957-961, 2004.

[11] Guan H. and Gordon, R. Computed tomography using algebraic reconstruction techniques (ARTs) with different projection access schemes: A comparison study under practical situations. *Phys. Med. Biol.*, vol. 41, pp. 1727-1743, 1996.

[12] Cheremuhin, E. and Chulichkov, A.I. Usage of the measurement computer systems in tomography. *Issue of Calc. Math. and Math. Phys.*, vol. 45, no.4, pp. 741-752, 2005.

[13] Davidsov, J.L., Garcia-Stewart, Ch. A., Ozanyan, K.B. Wright, P., Pegrum, S. and McCann, H. Image reconstruction for chemical species tomography with inregular and sparse beam array. *Proc. Photon 06*, 2006.

[14] Halfhill, T.R. Parallel processing with CUDA. *Microprocessor Report*, January 2008.

[15] nVidia. nVidia CUDA Programming Guide Version 2.3.1, pp. 31, 2009.

[16] Kak A.C. and M. Slaney. Principles of Computerized Tomographic Imaging. IEEE Press, NY, 1988.

[17] Nikolaev, D., Karpenko, S., Nikolaev, I. and Nikolayev, P. Hough Transform: Underestimated Tool in the Computer Vision Field. *22th European Conference on Modelling and Simulation*, pp. 238-243, 2008.