UNIVERSAL SYSTEM FOR TOMOGRAPHIC RECONSTRUCTION ON GPUS

E.D. Kotina, V.N. Latypov, V.A. Ploskikh
Saint-Petersburg State University, Saint-Petersburg, Russian Federation
E-mail: ekotina123@mail.ru

MLEM methods are commonly used in PET and SPECT reconstruction. The paper presents a parallel implementation of MLEM using an OpenCL programming language. Evaluation of the algorithms is performed using simulated sinograms and real PET/SPECT data. Comparison with the CPU-based reference FBP and MLEM algorithms is performed.

PACS: 87.57.uk, 87.57.nf

1. INTRODUCTION

1.1. PREVIOUS WORK

During the last decade multiple attempts have been made to utilize the ever growing computational power of the graphical processing units in PET/SPECT image reconstruction. Standard MLEM and OSEM algorithms have been previously implemented [1, 2]. Modifications of these algorithms, such as Weighted Least-Squares [12], improve the image quality with some extra computations. Extensions to the 3D reconstructions are also proposed [8, 9]. Modifications also exist for the gated PET reconstruction [11]. Comparison of the MLEM algorithm with FBP on a large number of medical images is performed in [7].

MLEM/OSEM algorithms are based on a solution of linear algebraic equations, making them an ideal target for effective GPU-based parallelization. Previously linear algebraic equation solvers were implemented using graphics interfaces like OpenGL. With the invention of the GPGPU techniques and OpenCL [5] programming language it is possible to create the unified reconstruction system for PET and SPECT reconstruction which differs only in the system matrix calculation.

1.2. MLEM METHODS

The process of PET data acquisition is usually modelled as a system of linear algebraic equations.

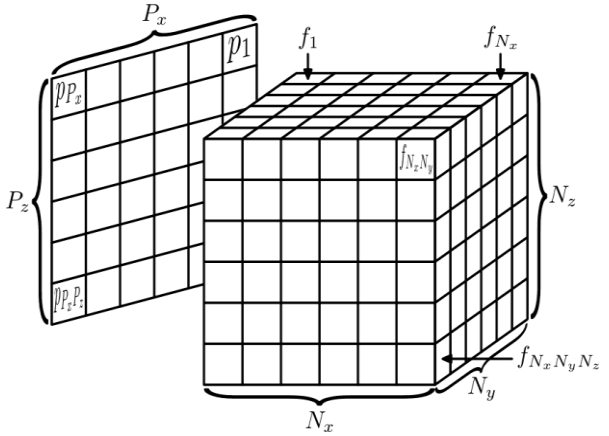

Fig. 1. The volumetric and projection data

If *f(x,y,z)* is the unknown density and *p(i)* is the set of recorded detection events, then after replacing *f* by the discrete set of *f(i,j,k)* we can write down the relation

$$p = Af, \qquad (1)$$

where A is the System Response Matrix (SRM) or the projection matrix, whose entries *a(i,j)* are proportional to the influence of an *i*-th volume element (voxel) to the *j*-th projection data element *p(j)*. Data indexing for SPECT is shown in Fig. 1. For the PET we use the *a(i,j)* equal to the probability of the detection of the photon emitted at the center of the *i*-th voxel on the *j*-th line of response (LOR). Typical LORs are shown in Fig. 2.

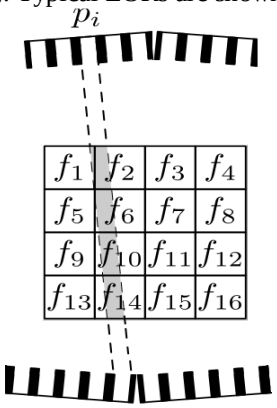

Fig. 2. PET lines of response

Given a detector configuration we calculate the cartesian coordinates for each detector pair and use the modification of Besenham's algorithm to trace the line from one detector to another. SPECT projection matrix has more non-zero elements, because the projection bin *p(i)* accumulates events from the conic volume (Fig. 3). The calculation of SRM for the SPECT is discussed in details in [15].

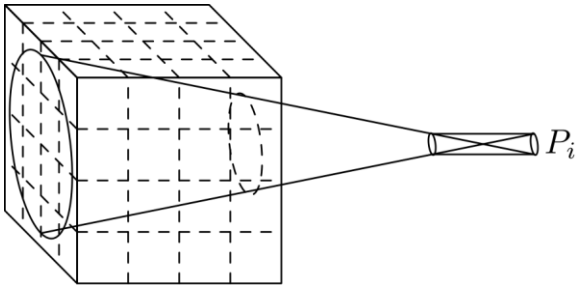

Fig. 3. SPECT geometry

To solve the equation (1) iterative methods are used. The most common algorithm is the Maximum Likelihood – Expectance Maximization (MLEM) by Shepp and Vardi [1] and its modification, Ordered Subsets EM (OSEM), by Hudson and Larkin [2]. Single iteration of the MLEM algorithm can be formulated as

$$f_{k+1} = f_k \frac{A\left(\frac{p}{Af}\right)}{A1}. \qquad (2)$$

In (2) the f is an $N \cdot N \cdot N$-dimensional vector and the division is performed element-wise. The 1 denotes a constant vector. The denominator is a vector with elements equal to the sum of A rows.

The iterations stop when two sequential approximations differ less than a specified threshold. The stopping criteria is based on the Normalized Root Mean-Square Deviation (NRMSD), given in [3]:

where *max* and *min* are the minimum and maximum values of the pixels found in f and g images.

At the end of each iteration we compare f_k and f_{k+1}, using the *NRMSD* f_k, f_{k+1}.

1.3. 2D RECONSTRUCTION

Restriction of the reconstruction to 2D slice allows to reduce the memory consumption for the SRM. Also the single reconstructed slice fits in the cache memory and minimizes the number of random memory accesses on each iteration of the MLEM algorithm.

2. IMPLEMENTATION DETAILS

2.1. ALGORITHM OUTLINE

The single iteration of the MLEM algorithm consists of three basic operations:

1) Projection or the multiplication by A.
2) Backprojection or the multiplication by A^T.
3) Correction – element-wise multiplication of the image by the projection and backprojection ratio.

Third step of the iteration is easily parallelizable and takes less time compared to step 1 and 2. First two steps are essentially the same and consist only of the matrix by vector multiplication.

In this paper we restrict the problem to 2D reconstruction. For the high-quality image and a symmetric PET detector configuration it is possible to upload the whole SRM to the GPU memory using the Yale compression scheme also known as compressed row storage [4].

2.2. MATRIX COMPRESSION

Both PET and SPECT projection matrices contain many zero entries. If the reconstructed slice is the $N \cdot N$ array and the number of lines of response is $M=D(D-1)/2$, where D is the number of detectors, then the complete SRM matrix consists of $(N \cdot N) \times M$ elements. Since each line of response may intersect with no more than $4*N$ voxels, each row may contain no more than $4*N$ non-zero elements, which is about 5% of the total memory space for a typical detector configuration and a high-resolution slice.

We store the list of non-zero matrix entries in three linear arrays, the IA, JA and A. IA[i] contains the index of the first non-zero element in i-th row, JA[i] contains the number of non-zero elements in the i-th row and A contains all the values of non-zero entries of the SRM.

For example, the matrix

[0.3 0.5 0.0 0.0]
[0.0 0.2 0.7 0.0]
[0.0 0.0 0.3 0.2]
[0.0 0.0 0.0 0.1]

has 7 non-zero elements and the corresponding A, IA and JA arrays are the following:

A = [0.3 0.5 0.2 0.7 0.3 0.2 0.1],
IA = [0 2 4 7],
JA = [0 1 1 2 2 3 3].

2.3. OPENCL PROGRAMMING

In this work the OpenCL [5] is chosen to implement the parallel version of the MLEM algorithm. This technology allows the program to run on any modern GPU from major manufacturers as nVidia, AMD and Intel.

The 2D variant of the MLEM algorithm reconstructs the volume data slice by slice. Each slice is initially filled with a single non-zero value. Then the projection data acquired from PET scanner is uploaded to the GPU memory and the iterations are performed. Each iteration uses a double-buffering scheme to store the old and new slice value (f_k and f_{k+1}). The steps of one iteration of the MLEM algorithm are based on the formula (2). The list of steps is the following:

1) Calculate backprojection vector X_{new}.
2) Calculate projection P_{new} from X_{old}.
3) Multiply X_{new} by the Correction factor.
4) Swap X_{new} and X_{old} vectors.
5) Calculate the $NRMSD(X_{new}, X_{old})$ and stop iteration if this value is smaller than the given threshold.

3. NUMERICAL EXPERIMENTS

3.1. GENERATING PROJECTION DATA

All numerical experiments in this article are performed in the same way. We start with a given volumetric image, add some Poisson noise [10] to each pixel, project the modified image using the projection operation of the MLEM algorithm and finally we use the obtained projection data as the source for our MLEM reconstructor. To validate the MLEM algorithm we use the filtered backprojections (FBP). The images are reconstructed using the FBP, MLEM and GPU-based MLEM implementation.

3.2. DERENZO PHANTOMS

The first numerical example is the reconstruction of a 256 by 256 pixels volume slice with the Derenzo phantom. Derenzo-type phantoms are widely used in the evaluation of PET reconstruction algorithms because they can be easily manufactured as thin tubes in a hollow plastic case [6].

Fig. 4 shows the same volume reconstructed with MLEM, FBP and GPU/MLEM algorithms. Left image, obtained with MLEM algorithm, uses the projection of a volume without noise. The central image is the slice reconstructed with filtered backprojections. The right image is the slice reconstructed with our GPU-based MLEM algorithm from the projection data with the same amount of Poisson noise as the central image.

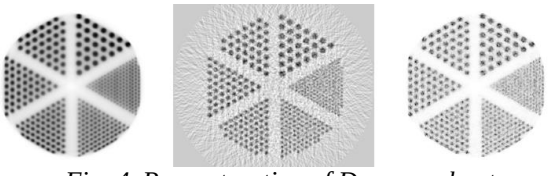

Fig. 4. Reconstruction of Derenzo phantom

3.3. HUMAN BRAIN SCANS

To demonstrate the applicability of the developed GPU reconstructor to SPECT, we use a slice of a human brain scan from EFATOM SPECT [13]. In this experiment we compare the FBP reconstruction from the standard EFATOM software suite and our new GPU-based MLEM reconstruction of low-statistics data (Fig. 5).

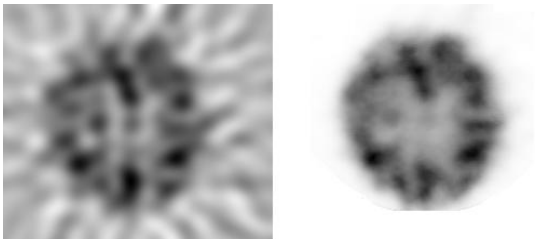

Fig. 5. Reconstruction of the EFATOM SPECT scans

The final experiment (Fig. 6) concerns the 128×128 PET human brain scan. We add the Poisson noise to the projection data and the reconstruct the slice using all reconstruction methods: FBP, MLEM and GPU-based MLEM.

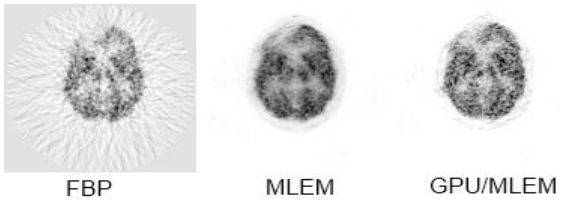

Fig. 6. Reconstruction of reprojected PET human brain scan

3.4. ALGORITHM PERFORMANCE

To make conclusions about the reconstruction speed of the algorithms we perform a series of experiments. Tables 1 and 2 present the reconstruction speed in seconds for the two tests: the high-resolution Derenzo phantom and a human brain scan. The first line of the tables contains the name of the method used and the second line contains the number of iterations for the MLEM methods (for filtered backprojections there is no iteration count).

Table 1

	FBP	MLEM			MLEM/GPU		
Iterations		5	10	20	10	20	40
Time	2,7	38	103	208	3.5	7	11.5

Derenzo, 256×256, 1 slice

Table 2

	FBP	MLEM			MLEM/GPU		
Iterations		5	10	20	10	20	40
Time	15	175	380	910	24.5	43	80

Brain scan, 128×128, 63 slices

CONCLUSIONS

A simple and effective parallel implementation of the MLEM algorithm is presented. The reconstruction speed is significantly faster than the CPU version of the same method and the algorithm is suitable for both SPECT and PET.

The work was supported by St. Petersburg State University (theme 9.39.1065.2012).

REFERENCES

1. L.A. Shepp, Y. Vardi. Maximum Likelihood Reconstruction for Emission Tomography // *IEEE Trans. Med. Imag*. 1982, v. 1, №.2, p. 113-121.
2. M. Hudson, R.S. Larkin. Accelerated Image Reconstruction using Ordered Subsets of Projection Data // *IEEE Trans. Med. Imag.* 1994, №13, p. 601-609.
3. A. Gaitanis, G. Kontaxakis, G. Spyrou, G. Panayiotakis, G. Tzanakos. PET image reconstruction: A stopping rule for the MLEM algorithm based on properties of the updating coefficients // *Comput Med Imaging Graph*. 2010, v. 34, №2, p. 131-141.
4. S. Pissanetzky. *Sparse Matrix Technology*. Academic Press, 1984.
5. A. Munshi ed. *The OpenCL Specification v. 2.0*. Khronos Group. http://www.khronos.org/registry/cl/. Last accessed 30 July 2013.
6. M.-A. Park, R.E. Zimmerman, A. Tabener, M.W. Kaye, S.C. Moore. Design and fabrication of phantoms using stereolithography for small-animal imaging systems // *Mol Imaging Biol.* 2008, v. 10(5), p. 231-236.
7. D.D. Duarte, M.S. Monteiro, El Hakmaoui F., J.O. Prior, L. Vieira, J.A. Pires-Jorge. Influence of Reconstruction Parameters During Filtered Backprojection and Ordered-Subset Expectation Maximization in the Measurement of the Left-Ventricular Volumes and Function During Gated SPECT // *J. Nucl. Med. Technol.* 2012, v. 40, p. 29-36.
8. P. Aguiara, M. Rafecas, J.E. Ortuno, G. Kontaxakis, A. Santos, J. Pavia, D. Ros. Geometrical and Monte Carlo projectors in 3D PET reconstruction // *Med. Phys.* 2010, v. 37(11), p. 5691-5702.
9. S. Moehrs, M. Defrise, N. Belcari, A. Del Guerra, A. Bartoli, S. Fabbri, G. Zanetti. Multi-ray-based system matrix generation for 3D PET reconstruction. // *Phys. Med. Biol.* 2008, v. 53, p. 6925-6945.
10. L. Devroye. *Non-Uniform Random Variate Generation.* New-York: Springer-Verlag, 1986.
11. J. Dey, M.A. King. Theoretical and Numerical Study of MLEM and OSEM Reconstruction Algorithms for Motion Correction in Emission Tomography // *IEEE Trans. Nucl. Sci*. 2009, v. 56(5), p. 2739-2749.
12. C.W. Stearns, J.A. Fessler. 3D PET Reconstruction with FORE and WLS-OS-EM // *IEEE Nuclear Science Symposium*. 2002, p. 912-915.
13. M.A. Arlychev, V.L. Novikov, A.V. Sidorov, A.M. Fialkovskii, E.D. Kotina, D.A. Ovsyannikov, V.A. Ploskikh. EFATOM Two-Detector One-Photon Emission Gamma Tomograph // *Technical Physics. The Russian Journal of Applied Physics.* 2009, v. 54, № 10, p. 1539-1547.

14. N. Mitsuhashi, K. Fujieda, T. Tamura, S. Kawamoto, T. Takagi, K. Okubo. Body Parts 3D: 3D structure database for anatomical concepts // *Nucleic Acids Research.* 2009, v. 37, p. 782-785.
15. Z. Liang, T.G. Turkington, D.R. Gilland, R.J. Jaszczak, R.E. Coleman. Simultaneous compensation for attenuation, scatter and detector response for SPECT reconstruction in three dimensions // *Phys. Med. Biol*. 1992, v. 37, p. 587-603.

Article received 04.10.2013

УНИВЕРСАЛЬНАЯ СИСТЕМА ДЛЯ ТОМОГРАФИЧЕСКОЙ РЕКОНСТРУКЦИИ С ИСПОЛЬЗОВАНИЕМ ГРАФИЧЕСКИХ ПРОЦЕССОРОВ

Е.Д. Котина, В.Н. Латыпов, В.А. Плоских

Реконструкция на основе метода максимального правдоподобия (MLEM) широко используется в позитронно-эмиссионной томографии. В статье представлена параллельная реализация метода MLEM на языке OpenCL. Оценка алгоритмов производилась на основе смоделированных синограмм, а также данных ПЭТ- и ОФЭКТ-исследований. Проведено сравнение базового метода фильтрованных обратных проекций, выполняемого на центральном процессоре, с алгоритмами MLEM, в том числе с реализуемыми с использованием современных графических процессоров.

УНІВЕРСАЛЬНА СИСТЕМА ДЛЯ ТОМОГРАФІЧНОЇ РЕКОНСТРУКЦІЇ З ВИКОРИСТАННЯМ ГРАФІЧНИХ ПРОЦЕСОРІВ

О.Д. Котіна, В.Н. Латипов, В.А. Плоскіх

Реконструкція на основі методу максимальної правдоподібності (MLEM) широко використовується в позитронно-емісійній томографії. У статті представлена паралельна реалізація методу MLEM на мові OpenCL. Оцінка алгоритмів проводилася на основі змодельованих сінограмм, а також даних ПЕТ- та ОФЕКТ-досліджень. Проведено порівняння базового методу фільтрованих зворотних проекцій, виконуваного на центральному процесорі, з алгоритмами MLEM, в тому числі з реалізованими з використанням сучасних графічних процесорів.