

# High Performance Computing in Medical Physics

Roberto A. Isoardi, Pablo A. Cappagli

Comisión Nacional de Energía Atómica, Fundación Escuela de Medicina Nuclear Mendoza,  
Argentina

risoardi@ieee.org

**Abstract.** Within the myriad applications of Physics in Medicine, there are two major fields in terms of their relevance in clinical practice: Medical Imaging and Radiation Therapy. Both areas make extensive use of computational resources in order to provide a prompt response to physicians, if possible in real time. Although execution times were dramatically reduced in the last decade with faster than ever CPUs, it is still common to wait several minutes and on some occasions, several hours for certain processing tasks to yield clinically useful results. Some frequent examples include tomographic image reconstruction, internal dosimetry calculation and radiotherapy planning. Acceleration of such processes may be sometimes vital or extremely important, not only for the patient –whose quality of life improvement is the ultimate goal–, but also for optimizing professional work in a busy hospital environment. In recent years, Medical Physics benefited greatly from the implementation of new computing strategies for several applications, particularly making use of GPU [1]. This short paper describes some of the current fields of our expertise in Medical Physics, where High Performance Computing (HPC) plays indeed a key role.

## 1 Medical Image Processing

This field comprises a wide range of computing tasks, such as image reconstruction, processing and analysis, generally aimed at enhancing diagnostic efficiency and assisting on therapy planning. In this section we describe two typical lines of research in medical image processing, namely image segmentation and registration.

### 1.1 Brain tissue segmentation from Magnetic Resonance Imaging

Image segmentation is a major discipline in general image processing, consisting on the identification and further classification of different constituents or textures depicted in a given dataset [2]. In the case of biomedical images, automatic recognition and classification of different tissues becomes extremely important, for both diagnosis and therapy, in a qualitative or even also a quantitative fashion. (i.e. study of white matter lesions in the brain in patients with multiple sclerosis, volume-

try, etc.). A particularly interesting application is the automatic segmentation of Magnetic Resonance Imaging (MRI) of the brain. This imaging modality has a unique feature by producing high contrast in soft tissue (white matter, gray matter and cerebro-spinal fluid). The contrast can also be customized in MRI by selecting different radiofrequency excitation sequences (i.e. T1, T2, DP). An updated and extensive review on MRI segmentation methods can be found in [3].

One of the proposed statistical methods uses Bayesian Inference to segment the main tissues in the brain from MRI, using the Maximum Evidence criterion [4]. In a first approach, so called Discrete Model, it is assumed that every pixel can belong to only one tissue class. In a more complex approach, the Partial Volume Model, two classes are allowed to co-exist in every pixel with different probabilities, thus representing the situation at tissue border in a more realistic fashion. This models assume that every pixel takes an intensity value similar to its closest neighbors', by means of an interaction potential. As every neighborhood is computed independently, the algorithm calculates values pixel by pixel in sequential or nested loops for every iteration. Since in this case the classification outcome for every pixel does not influence the computation of others afterwards, this algorithm is a good candidate to be expedited with parallel programming, either on multi-core platforms or GP-GPU.

The image to be processed, usually consists of a 3D dataset of dimensions 256x256x256 voxels (2 bytes/voxel) and typical execution times using a gcc compiler are 5-10 minutes on an ordinary Core 2 Duo processor. We recently have implemented this segmentation algorithm on GPU with promising results [5], achieving 15x speedup in comparison to the optimized code on CPU.

## 1.2 Rigid and deformable registration of tomographic modalities

Another main field in medical image processing is the one bringing two or more datasets into spatiotemporal alignment. This process, known as registration, usually takes a *reference image* (2D or 3D of a given scan) to which two or more *floating images* of the same or different modalities (i.e. MRI- MRI, CT- MRI, PET- MRI, etc)<sup>1</sup> are aligned after applying a spatial transformation. This operation may be rigid or deformable, and a cost function is optimized in an iterative algorithm. The motivations for such a process are multiple, ranging from diagnostic power enhancement after comparing different modalities, disease follow-up, to assistance in radiotherapy planning [6].

A frequent problem is the combination of 3D tomographic scans of a given patient of the same or different modalities. The shape of certain body regions will not change noticeably between scans, such as the brain in most cases. In this case it is possible to apply rigid-body transformations which have only 6 degrees of freedom in 3D. Other regions such as thorax and abdomen, however, deform considerably between scans due to various reasons: different patient positioning and curvature of the scanner bed, respiratory motion, physiological and/or structural changes, either normal or pathological.

---

<sup>1</sup> CT: Computed Tomography ; PET: Positron Emission Tomography

In order to treat this challenging problem, we designed a deformable registration algorithm that applies a deformation field using Fourier basis functions on the floating image, choosing the *Mutual Information* as the similarity measure to be maximized [7]. A hierarchical subdivision scheme was applied to improve its efficiency and to avoid that the algorithm gets trapped in local maxima. In each subdivision stage, deformable registration is applied between corresponding subvolumes of the reference and floating image [8]. Finally, registered subvolumes are matched to neighbors using quaternion interpolation [9].

Normally, and under the same compiling conditions and platform as described in the previous section, this algorithm demands 30-40 minutes of CPU to register two 512x512x50 voxel datasets. Given the hierarchical subdivision scheme chosen in the algorithm design, this problem may be considerably accelerated using parallel computing as described in previous publications [1].

## 2 Radiotherapy

In Radiation Therapy, the calculation of the dose delivered by ionizing radiation and the use of optimization algorithms on advanced methods of treatment, are the main areas where GPU programming has its greatest impact. In the clinical practice, the performance and accuracy of dose calculation are the foundations for the quality of treatment plans. Currently, treatment plans are evaluated by creating several plans until, according to medical criteria, a final plan is approved, meeting a compromise between the dose delivered to the tumor and dose received by healthy organs located around it. In addition to this, advanced treatment modalities such as intensity-modulated radiation therapy (IMRT), require the execution of an optimization algorithm to find the proper intensity for each treatment beam to conform the dispensed dose. This points out that the reduction of the execution times of dose calculation and optimization algorithms play a key role in the selection of the treatment plan, allowing medical evaluation of a greater number of plans and the subsequent selection of the most suitable one. Fast elaboration of treatment plans also enables the implementation of adaptive treatments (Adaptive Radiation Therapy), a concept whereby the treatment plan created for a patient is adjusted periodically (along the weeks involving its execution) based on the patient response. In the extreme, the treatment plan could be adapted in each daily application.

The dose calculation algorithms that currently predominate on commercial treatment planning systems lie on resolving superposition integrals that compute the contribution to the dose produced by each element of the beam from the primary interaction site to neighboring deposition sites. When the radiation beam is decomposed into multiple diverging beams of differential cross-section (known as beamlets), the solution of the problem is 2D and is known as pencil beam algorithm. When calculating the contribution of the beam from specific sites of interaction, the problem becomes 3D and the algorithm is known as convolution/superposition. In both types of algorithms the dose is obtained by calculating the contribution of each beamlet by convoluting it with a kernel (appropriate for each problem) representing the dose dep-

osition produced by a unit intensity beamlet. The contribution of each interaction site is independent of the others, which makes these algorithms ideal candidates for parallel programming. Several research groups have reformulated pencil beam and convolution/superposition algorithms for GPU implementation [10, 11, 12], reporting speedups of 200-400x [13].

The first step in the implementation of this kind of algorithms is to calculate the distance from the source of radiation and from the interaction site to the dose deposition site. To do this, a line integral must be solved weighting the distance traveled with the electron density of the traversed CT voxels. These types of methods (known as raytracing) are computationally demanding and, for each beam into a treatment plan, the radiological distance of thousands of lines must be calculated. As each traced line is independent, raytracing algorithms could be adapted for parallel implementation [14, 15], reporting speedup of 6x. This type of increase, while modest, is essential to achieve real time dose calculation.

Monte Carlo methods offer more accurate dose calculation algorithms than those mentioned above, but are computationally more expensive and its implementation in clinical practice is not yet widely available. Monte Carlo techniques simulate, from first principles, how the history of each particle in the irradiation beam evolves. The addition of billions of particle histories gives statistical accuracy to the method. Since each particle of the primary beam is independent of the others, the Monte Carlo algorithms are ideal for parallel computation. However, due to the complexity of the method, only moderate acceleration has been reported [16, 17].

Currently, at Fundación Escuela de Medicina Nuclear de Mendoza, we are working to implement pencil beam type algorithm for electron beam dose calculation (along with a raytracing algorithm for radiological distance calculation) onto GPU. These algorithms will be used in modulated compensatory equivalent tissue design [18, 19, 20], which will calculate the necessary compensatory tissue for greater conformation of the delivered dose to the tumor in real time.

### 3 References

1. Pratz, G., Xing L.: GPU computing in medical physics: A review. *Med. Phys.* 38, 2685 (2011)
2. Gonzalez, R. C., Woods, R. E.: Digital Image Processing, 3rd edition, *Prentice Hall*, 2008.
3. Balafar, M. A., *et al.* Review of brain MRI image segmentation methods. *Artificial Intelligence Review* 33.3 261-274 (2010)
4. Isoardi, R. A., Oliva, D. E., Mato G.: Maximum Evidence Method for classification of brain tissues in MRI. *Pattern Recognition Letters* 32.1 12-18 (2011)
5. Fioretti F, Lemma D, Millan E, Isoardi R.A., Bringa E.M.: Implementation of a Bayesian Statistics Segmentation Algorithm in GPUs, High Performance Computing Latin America, HPC LatAm, Conference Proceedigs (2012)
6. Markelj, P., *et al.*: A review of 3D/2D registration methods for image-guided interventions. *Medical image analysis* 16.3 642-661 (2012)
7. Isoardi, R.A., Osorio, A.R., Mato, G: Medical Image Registration with Fourier basis Functions. In: Fourier Transforms - New Analytical Approaches and FTIR Strategies, *InTech Editors*, Rjeka. (2011) Available from: <http://www.intechopen.com/books/fourier>

transforms-new-analytical-approaches-and-ftir-strategies/medical-image-registration-with-fourier-basis-functions

8. Likar, B., Pernuš, F.: A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing*, 19(1), 33-44. (2001)
9. Walimbe, V., Shekhar, R.: Automatic elastic image registration by interpolation of 3D rotations and translations from discrete rigid-body transformations. *Medical Image Analysis*, 10(6), 899. (2006)
10. Gu, X., *et al*: A GPU-based finite-size pencil beam algorithm with 3D-density correction for radiotherapy dose calculation, *Phys. Med. Biol.* 56 (2011)
11. Hissoiny, S., Ozell, B., Després, P.: A convolution-superposition dose calculation engine for GPUs. *Med. Phys.* 37 (2010)
12. Lu, W.: A non-voxel-based broad-beam (NVBB) framework for IMRT treatment planning. *Phys. Med. Biol.* 55 (2010)
13. Chen, Q., Chen, M., Olivera, G., Lu, W.: Non-Voxel-Based-Broad-Beam (NVBB) Framework for EMRT Treatment Planning - Ultra-Fast Full Dose Calculation. *Med. Phys.* 37 (2010).
14. Greef, M., *et al*: Accelerated ray tracing for radiotherapy dose calculations on a GPU. *Med. Phys.* 36 (2009)
15. Xiao, K., *et al*: Efficient implementation of the 3D-DDA ray traversal algorithm on GPU and its application in radiation dose calculation. *Med. Phys.* 39 (2012)
16. Jia, X., *et al*: Development of a GPU-based Monte Carlo dose calculation code for coupled electron-photon transport. *Phys. Med. Biol.* 55 (2010)
17. Zhou, B., *et al*: GPU-accelerated Monte Carlo convolution/superposition implementation for dose calculation. *Med. Phys.* 37 (2010)
18. Hogstrom, K., Fields, R.: Electron beam dose calculations. *Phys. Med. Biol.* 26. (1981)
19. Low, D.A., *et al*: Electron bolus design for radiotherapy treatment planning: Bolus design algorithms. *Med. Phys.* 19, (1992)
20. Acevedo, Y.: Nueva técnica de fabricación de bolus para radioterapia con haces externos de electrones de alta energía. Tesis de Maestría en Física Médica de la UBA (2012)

Learn about Intel's high performance computing processors, architectures and frameworks, and why our technologies deliver better, faster results. High Performance Computing (HPC). This is HPC on Intel. Tackle complex workloads and tomorrow's challenges with one data-centric platform designed for HPC, AI, and analytics.<sup>1 2 3</sup> Many computing applications in medical physics can be formulated as data-parallel tasks that exploit the capabilities of the GPU for reducing processing times. The authors review the basic principles of GPU computing as well as the main performance optimization techniques, and survey existing applications in three areas of medical physics, namely image reconstruction, dose calculation and treatment plan optimization, and image processing. Author supplied keywords. dose calculation. graphics processing units. high-performance computing. image processing. image segmentation.