



16-18 March 2016 TIFPA e Centro di Protonterapia

Treatment Planning Simulations in Hadrontherapy with GPUs

Simona Giordanengo Istituto Nazionale di Fisica Nucleare Torino



My recent expertise





A GPU-based Planning and Delivery System to treat moving targets with therapeutic scanned ion beams

RIDOS: Real-Time Ion **Do**se Planning and Delivery **S**ystem

PI: Simona Giordanengo - INFN - Torino

First year activities CSN5 Padova 16 Febbraio 2015



INFN

ISTITUTO NAZIONALE DI FISICA NUCLEARE

Bandon. 15766

Concorso per il finanziamento di n. 1 progetto r giovani ricercatrici/ricercatori nell'ambito delle linee di ricerca della Commissione Scientifica Nazionale 5

IL PRESIDENTE

dell'Istituto Nazionale di Fisica Nucleare

zione della Giunta Esecutiva dell'Istituto in data 14 giugno 2013 n. 9875, da ratificare:

The modulated pencil beam scanning hadrontherapy



The modulated pencil beam scanning hadrontherapy



Scanning magnets in the beam line provide the beam movement in the transversal directions

The energy can be changed directly through the accelerator for **synchrotrons** or through energy selection system for **cyclotrons**

Thousands narrow beams over thousands voxels



20

The dose delivery system (DDS) to control the treatment



Transparent Ionization chambers to control the beams and drive the treatment

Scheme of the CNAO DDS

The dose delivery system (DDS) to control the treatment



Treatment planning simulations to provide the optimized beam characteristics

3D effects distributions

CT image : anatomical description



Dose prediction models

Radiation transport simulations and optimizations are used to predict the 3D distributions of:

- Physical Dose (Gy),
- Relative Biological Effectiveness (RBE)
- Biological Dose (Physical Dose * RBE),
- Linear Energy Transferred (LET),
- Survival

Different dose prediction models are available including:

- pencil beam,
- convolution-superposition
- Monte Carlo

Precision versus computation time being the relevant trade-off.

INFN-IBA TPS (DEK) for ion dose calculation



The INFN-IBA Beam Model allows computing the 3D effect of an ion field incident on a water-like material. This model is implemented in a research Dose Engine Kernel (DEK) to estimate the outcome of a therapeutic ion irradiation delivered through the spot-scanning technique.

A novel algorithm for the calculation of physical and biological irradiation quantities in scanned ion beam therapy: the beamlet superposition approach

> G Russo¹, A Attili¹, G Battistoni², D Bertrand⁵, F Bourhaleb⁶, F Cappucci², M Ciocca⁷, A Mairani⁷, F M Milian^{1,8}, S Molinelli⁷, M C Morone³, S Muraro², T Orts⁵, V Patera⁴, P Sala², E Schmitt¹, G Vivaldo¹ and F Marchetto¹

> > Q

Phys. Med. Biol. 61 (2016) 183-214



The RIDOS idea



Fast Dose Computation based on GPU



The RIDOS idea



The RIDOS goal





The RIDOS goal





Fast Forward Dose Computation (FDC) for *on-line* verification purposes



Prescription

Delivery

GPU-based Forward Planning: milestone of the RIDOS project



- M1: Fast Forward Planning implementation porting on GPU the INFN-IBA TPS libraries (SW)
- M2: Development of a simulator for preliminary out-of-beam tests (HW+SW)
- M3: HW design and development (HW)



DEK Dose Computation scheme GPU algorithms: (I) Raytracing



Raytracing (I) : Find intercepts with CT voxels

GPU code: check the intercept between the ray and each plan along the 3 voxels direction (stream compaction algorithm)



Max N intercepts: $(N_x+N_y+N_z) \times N_{ray}$ N = CT voxels ~ (512+512+256)

 N_{rav} (Min ÷ Max) = 1 ÷ 5000000



Raytracing(II): Sorting the intercepts along the beam propagation





GPU Code: Bitonic Sort

The intercepts are sorted in function of the distance from the CT entrance



Raytracing (III): Measure the ray Path Length (L)





 $L = \sum_{i} l_{i}$

 $\mathsf{Li} = \left[l(\vec{r}_{i+1}) - l(\vec{r}_i) \right]$



Xeon E5-1620v2-4C-16GB Nvidia TESLA-K20c.

Summary: Raytracing performances



CT size = 512 x 512 x 125 (~3x10⁷ voxels) COMP_GRID size = 170 x 170 x 125 (~4x10⁶ voxels)

number of rays = 124800 number of energies = 39 Radial cut-off = 10 mm Each ray is independent -- > raytracing for each ray in parallel

Different release and different NVIDIA GPUs

	QUADRO 2100M Laptop	TESLA K20c	TESLA K40c
Clock rate	666 MHz	705 MHz	745 MHz
Global Mem	2 GB	5 GB	12 GB
N SM	3	13	15
Performances			
Ray_Tr V.1	3.25 sec	1.59 sec	2.30 sec
Ray_Tr V.2	2.80 sec	0.71 sec	0.69 sec

Convert the path length of particles in Water Equivalent Path Length (WEPL)

The two paths are equivalent in the sense that a particle loses the same average energy (same dose) while traversing them.



Voxel selection: computing grid and cut-off

→ Select a Computing Grid (usually smaller than the CT grid)
 → Identify the voxels within a cut-off



Cut-off range: 20 – 100 mm

Large cut-off \rightarrow better accuracy

Large cut-off \rightarrow more voxels \rightarrow high computation time for standard C++ codes

1st Texture interpolation: find the WEPL value of each voxel within the cut-off

Hardware interpolation using the GPU Texture

Returns for each selected voxel the water-equivalent position in the beam reference system





2nd Texture Interpolation: dose evaluation from pre-computed LUTs

(1 LUT for each particle type)

The physical and biological dose contribution for each WEPL voxel **of the computing grid** is obtained through a 3D interpolation of pre-computed LUTs



Dose (Gy) Biological Dose (Gy[RBE]) LET (keV/μm) Survival probability

GPU-based DEK: PROTON dose performances



CPU: XeonE5-1620v2-3.7-4C-16GB GPU: Nvidia TESLA-K20c.

GPU-based DEK: CARBON ION dose performances



N voxels= 7292; radial cut-off=10 mm.

RIDOS GPU-based FP for inter-spill dose computation



Examples of Fast Dose Computation applications

- The INFN RIDOS project
- The FRED project (Fast GPU-based Monte Carlo)
- Recent sub-seconds pencil beam algorithm (by Joakim Da Silva)

FRED – GPU-based Monte Carlo for fast proton dose computation



- G. Battistoni INFN-MI
- V. Patera Univ. La Sapienza Roma / INFN
- S. Pioli INFN-LNF / Univ. La Sapienza Roma
- A. Schiavi Univ. La Sapienza Roma /INFN
- M. Senzacqua Univ. La Sapienza Roma

FRED <u>a fast-Monte Carlo tool for treatment planning and dose verification</u> <u>in proton therapy</u>

future use at CNAO for patient verification plans





- Text input file (flexible format)
- geometry and ROIs definition
- accelerator file (CNAO settings)
- pixel resampling: optimization parameters
- DICOM import
- RTSTRUCT (PTV,OAR,etc.)
- RTPLAN (raster file)





OUTPUT

- 3D maps using SILO library (VisIT/ParaView)
- binary output for post-processing (Matlab, Python)
- ROI:
 - Dose statistics
 - DVH



Gamma index pass rate and maps



FRED's hardware and performance



** DESKTOP: Mac Pro (AMD FirePro D300)

Recent Sub-second pencil beam algorithm by Joakim da Silva

Sub-second pencil beam dose calculation on GPU for adaptive proton therapy

Joakim da Silva^{1,2}, Richard Ansorge¹ and Rajesh Jena²



¹ Cavendish Laboratory, University of Cambridge, Cambridge, UK
 ² Department of Oncology, University of Cambridge, Cambridge, UK

Phys. Med. Biol. 60 (2015) 4777

Fast Pencil Beam Dose Calculation for Proton Therapy Using a Double-Gaussian Beam Model

Joakim da Silva^{1,2}*, Richard Ansorge¹ and Rajesh Jena²

¹Canandish Laboratory, Department of Physics, University of Cambridge, Cambridge, UK, ²Department of Oncology, University of Cambridge, Cambridge, UK



Front. Oncol. 5:281.

J. Da Silva PB Dose Calculation Results: skull base case



- $2\%/2 \text{ mm} \gamma$ -index compared to Fluka
- Passing rate: 96.7% (voxels receiving >10% of max dose)
- Corresponding value for Syngo: 96.8%

TIME PERFORMANCES

- Tesla K40 GPU (2880 cores @ 845 MHz)
- Skull base test case
 - 2 oblique beam directions
 - 38 + 45 energy layers, total 6776 spots
 - 2x2x2 mm³ dose resolution, 1x1x1 mm³ ray resolution
- Total calculation time 0.22 seconds
- Individual energy layers: 2.2-6.4 ms

Future and perspectives: Adaptive hadrontherapy (?)

Take advantage of better dose conformity -> need adaptive radiotherapy

For inter-fraction feedback Fast MC will be soon available for proton beams For intra-fraction a prompt feedback on real dose delivered will be possible BUT: *on-line patient imaging* is mandatory for significant clinical improvement

4D treatment (beam delivery synchronized with patient movements)

- Recalculate dose in 4D CT between each energy layer/spill
- Map dose back to reference (RIDOS)
- Compare with expected dose (RIDOS)
- VISION: act if different

Conclusions (I) Fast Dose Calculations allow

Developing new tools for the quality control of the effects related to the possible patient's movements during the delivery (RIDOS) before the irradiation (FRED)



Performing a step towards the control of the progressive motioncorrected dose distribution during treatment.

Conclusions (II)

GPUs are the present and the future for dose calculation

Thanks to RIDOS

A versatile FAST GPU-based Dose Engine Kernel for protons and ions is available at the INFN (To)

My research group is equipped with high performance GPUs and above all has learned the skill of CUDA programming

Additional References

Topical Review

GPU-based high-performance computing for radiation therapy

Xun Jia¹, Peter Ziegenhein² and Steve B Jiang¹

 ¹ Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
 ² Joint Department of Physics, The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust, SM2 5NG, SUTTON, UK

Phys. Med. Biol. 59 (2014) R151-R182

GPU technology is the hope for near real-time Monte Carlo dose calculations Xun Jia, X. George Xu, and Colin G. Orton

Citation: Medical Physics 42, 1474 (2015); doi: 10.1118/1.4903901



Vincenzo Monaco; Anna Vignati; Andrea Attili; Roberto Cirio; Mario Ciocca; Marco Donetti;

Flavio Marchetto; Felix Mas Milian; Germano Russo; Roberto Sacchi; Mohammad Varasteh Anvar.

Spare

Biological dose computation

$$\int_{ijk}^{bio}(\varphi) = \frac{-\alpha_x + \sqrt{\alpha_x^2 - 4\ln(S_{ijk})\beta_x}}{2\beta_x}$$
Cell Survival curve

$$L = Lethal events$$

$$S(D) = e^{-\langle L \rangle} = e^{-\alpha D - \beta D^2}$$

$$\int_{i}^{0} (\beta_{ijkl}) = \sum_{l} \alpha_{ijkl} \varphi_l D_{ijkl}^{0} + \left(\sum_{l} \sqrt{\beta_{ijkl}} \varphi_l D_{ijkl}^{0}\right)^2$$

$$\ell = \text{index of beams} \\
\psi_l = \text{fluence for each beam } \ell$$

$$\begin{bmatrix} D^0, \alpha_1 D^0, \sqrt{\beta_1} D^0, \alpha_2 D^0, \sqrt{\beta_2} D^0, \dots \\ \text{tissue 1} & \text{tissue 2} \end{bmatrix}$$
LUT values per single particle

$$D^0 = \text{Physical Dose}$$
Parameters for Biological Dose

CPU-DEK vs GPU-DEK time comparison



Computing Grid = 170x170x125; N of Rays = 1248; N of slice = 39;

State of art for Dose Distributions Comparison Low et al. Med. Phys. 25 (1998) ; , Wendling et al. Med. Phys 34 (2007); U-based Gamma-Index





Figure 2. CPU and GPU computation time as functions of the dose distribution resolution. Again, for the purpose of convenience, we scaled down the CPU computation time by a factor of 60.0 to illustrate it in the same axis of GPU computation time.