REAL-TIME ADAPTIVITY IN HEAD-AND-NECK AND LUNG CANCER RADIOTHERAPY IN A GPU ENVIRONMENT

Anand P Santhanam

Assistant Professor, Department of Radiation Oncology
OUTLINE

• Adaptive radiotherapy for head and neck and lung cancer
• Key tools used for adaptive radiotherapy
  • 3D Deformable Image Registration (DIR)
    • Real-time 3D DIR
  • Physics-based modeling
    • Quantification of systematic errors in DIR
  • 3D Dose Calculation
    • Real-time non-voxel based dose calculation
RADIOTHERAPY

- Treatment for unresectable tumors
- Procedure
  - Patient is already diagnosed with the type of cancer
  - A 3D/4D CT scan is acquired before the treatment
  - Clinical experts contour (or delineate) the tumor and surrounding critical organs
  - Appropriate radiation dose is planned
    - Max dose to the tumor
    - Min dose to the critical organs.
  - Patient is treated for several days
    - 5-35 days
Treatment Uncertainty

- Rigid Registration – neglects soft tissue changes
- Daily MVCT image quality - loss of detail and stratification
- Computational effort - accurate DIR is time consuming
Adaptive Therapy

- Calculate the dose delivered on deforming normal and diseased organs.
- Facilitate 3D structures for deforming anatomy.
- Effectively spare normal organs and tissues.
- Modify the dose delivered on subsequent fractions.
ADAPTIVE RADIOTHERAPY

• Accumulate Dose over Deformed Volumes
TOOLS FOR ADAPTIVE RADIOTHERAPY - 1

- 3D Image Registration
- 3D Biomechanical modeling
- 3D Dose Calculation
GPU BASED IMAGE REGISTRATION FOR ADAPTIVE RADIOTHERAPY
DEFORMABLE IMAGE REGISTRATION ACCURACY

- Registration error is typically quantified using manually placed landmarks
- Validation hampered by lack of ground truth data
SYSTEMATIC STUDY FOR DIR VALIDATION

• Registration parameters determined through exhaustive search.

• Validation:
  • Landmark based metric
    • Target Registration Error
  • Image based metrics
    • Mutual Information, Correlation Coefficient, Entropy
      Correlation Coefficient, DICE
11 Head and Neck Patients were used in the study.

6 levels of target volume reduction were examined:
- 0, 5, 10, 15, 20, and 30%

45 postures were created systematically at each volume reduction level:
- rotating the skull between 4 and -4 degrees along each axis.

Neylon J and Santhanam AP et al Medical Physics 2015
PATIENT SPECIFIC MODEL GENERATION

Initializing the mass-springs

- Load DICOM CT
- Load DICOM RTSTRUCT
- Volume Filling Algorithm
- Assign elements to structures
- Establish spring-damper connections
- Set material properties
Create a uniform cell grid, assign each element a hash value based on cell ID.

Sort by hash using a fast radix algorithm.

Search a 5x5x5 cell neighborhood and establish connections as a 3x3x3 cube, creating 26 ‘springs’ per element.

Record the rest lengths and orientations.
MODEL ACTUATION

• Control the skeletal anatomy
  • 1 degree rotations about each axis

• Soft Tissue deforms due to elastic forces

• The color map illustrates areas of compression (blue) and strain (red)
VOLUME CHANGES – WEIGHT LOSS

- Volume can be adjusted manually by increasing or decreasing the rest length of the internal connections of a structure.
- The update loop uses a two-pass system:
  - First - apply the internal structure forces
  - Second - propagate changes to surrounding tissues
SYNTHETIC DATA CREATION

• Find the voxelized coordinates of each element after deformation
• Rotation and regression causes hole and aliasing artifacts
• Holes are addressed by ray-casting along each spring connection to fill holes
• Aliasing is addressed using a GPU based texture smoothing on edges
• Record the vector displacement of each element and the structure to which they belong
• Randomly select 100 elements from each structure for landmark analysis
  • Compare to registration results to find TRE
PARAMETER SEARCH

- From a set of manually placed landmarks, calculated the target registration error (TRE) for a spectrum of registration parameters.

- Error for kV->MV registration with 5 Levels, 1 Warp

- Default parameters:
  - Smoothing: 500
  - Levels: 5
  - Warps: 2
  - Iterations: 150
PARAMETER SEARCH

- Similarly for kV->kV registrations
GROUND TRUTH REGISTRATION ACCURACY

2 Parameter optimization is convex

3 Parameter optimization is non-convex
GPU BASED COMPUTATIONS

GPU run time in dependence of the resolution levels and the solver iterations for a whole lung data (a) and separate lung data (b) on a NVIDIA GTX 680 GPU.
LANDMARK BASED DIR VALIDATION

Registration error by patient for head rotation of -4°, -2°, and -2° about the x, y, and z axes, respectively.

<table>
<thead>
<tr>
<th>Error (mm)</th>
<th>PTV 1</th>
<th>Parotids</th>
<th>Mandible</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>0.902</td>
<td>0.906</td>
<td>1.348</td>
<td>0.640</td>
</tr>
<tr>
<td>Patient 2</td>
<td>0.743</td>
<td>1.022</td>
<td>1.262</td>
<td>0.779</td>
</tr>
<tr>
<td>Patient 3</td>
<td>0.845</td>
<td>2.227</td>
<td>2.615</td>
<td>0.926</td>
</tr>
<tr>
<td>Patient 4</td>
<td>0.674</td>
<td>1.375</td>
<td>1.465</td>
<td>0.872</td>
</tr>
<tr>
<td>Patient 5</td>
<td>0.925</td>
<td>1.843</td>
<td>1.923</td>
<td>1.094</td>
</tr>
<tr>
<td>Patient 6</td>
<td>1.124</td>
<td>0.827</td>
<td>1.135</td>
<td>0.927</td>
</tr>
<tr>
<td>Patient 7</td>
<td>0.873</td>
<td>0.936</td>
<td>1.254</td>
<td>0.861</td>
</tr>
<tr>
<td>Patient 8</td>
<td>0.925</td>
<td>1.124</td>
<td>1.345</td>
<td>0.951</td>
</tr>
<tr>
<td>Patient 9</td>
<td>1.132</td>
<td>1.334</td>
<td>1.659</td>
<td>1.296</td>
</tr>
<tr>
<td>Patient 10</td>
<td>0.887</td>
<td>1.473</td>
<td>1.726</td>
<td>0.881</td>
</tr>
</tbody>
</table>

Santhanam AP and Neylon J, ASTRO 2014
GPU BASED DOSE CALCULATION

- Convolution/Superposition

\[ Dose(v) = \int T'(v') H(\rho_{v-v'}(v - v')) d^3v' \]

- Naïve Implementation
  - Port CPU algorithm directly
  - Calculate every voxel simultaneously

- Optimized Implementation
  - Coalesced Global memory - data size invariability
  - Texture memory caching - intrinsic linear interpolation
  - Shared memory utilization – 20x to 30x shorter latencies than Global memory

Neylon J and Santhanam A.P Medical Physics 2014
**PERFORMANCE - GPU PARALLELIZATION**

**Dependence on Field and Data Size**

Optimized GPU Convolution

<table>
<thead>
<tr>
<th>Angular Sampling Combination (Zenithal x Azimuthal)</th>
<th>64^3 Phantom (10x10 mm Field)</th>
<th>64^3 Phantom (50x50 mm Field)</th>
<th>64^3 Phantom (100x100 mm Field)</th>
<th>128^3 Phantom (10x10 mm Field)</th>
<th>128^3 Phantom (50x50 mm Field)</th>
<th>128^3 Phantom (100x100 mm Field)</th>
<th>256^3 Phantom (10x10 mm Field)</th>
<th>256^3 Phantom (50x50 mm Field)</th>
<th>256^3 Phantom (100x100 mm Field)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4x4</td>
<td>1E-1</td>
<td>1E-1</td>
<td>1E-1</td>
<td>1E-1</td>
<td>1E-1</td>
<td>1E-1</td>
<td>1E-1</td>
<td>1E-1</td>
<td>1E-1</td>
</tr>
<tr>
<td>6x6</td>
<td>1E-0</td>
<td>1E-0</td>
<td>1E-0</td>
<td>1E-0</td>
<td>1E-0</td>
<td>1E-0</td>
<td>1E-0</td>
<td>1E-0</td>
<td>1E-0</td>
</tr>
<tr>
<td>8x8</td>
<td>1E+1</td>
<td>1E+1</td>
<td>1E+1</td>
<td>1E+1</td>
<td>1E+1</td>
<td>1E+1</td>
<td>1E+1</td>
<td>1E+1</td>
<td>1E+1</td>
</tr>
<tr>
<td>10x10</td>
<td>1E+2</td>
<td>1E+2</td>
<td>1E+2</td>
<td>1E+2</td>
<td>1E+2</td>
<td>1E+2</td>
<td>1E+2</td>
<td>1E+2</td>
<td>1E+2</td>
</tr>
<tr>
<td>12x12</td>
<td>1E+3</td>
<td>1E+3</td>
<td>1E+3</td>
<td>1E+3</td>
<td>1E+3</td>
<td>1E+3</td>
<td>1E+3</td>
<td>1E+3</td>
<td>1E+3</td>
</tr>
<tr>
<td>16x16</td>
<td>1E+4</td>
<td>1E+4</td>
<td>1E+4</td>
<td>1E+4</td>
<td>1E+4</td>
<td>1E+4</td>
<td>1E+4</td>
<td>1E+4</td>
<td>1E+4</td>
</tr>
<tr>
<td>24x24</td>
<td>1E+5</td>
<td>1E+5</td>
<td>1E+5</td>
<td>1E+5</td>
<td>1E+5</td>
<td>1E+5</td>
<td>1E+5</td>
<td>1E+5</td>
<td>1E+5</td>
</tr>
<tr>
<td>48x48</td>
<td>1E+6</td>
<td>1E+6</td>
<td>1E+6</td>
<td>1E+6</td>
<td>1E+6</td>
<td>1E+6</td>
<td>1E+6</td>
<td>1E+6</td>
<td>1E+6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Field</th>
<th>100x100 mm</th>
<th>64^3 Phantom 4 mm voxels</th>
<th>128^3 Phantom 2 mm voxels</th>
<th>256^3 Phantom 1 mm voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU / Naïve</td>
<td>59.26 1.66</td>
<td>113.2 1.75</td>
<td>193.7 12.76</td>
<td></td>
</tr>
<tr>
<td>Naïve / Optimized</td>
<td>1.46 0.04</td>
<td>4.82 0.135</td>
<td>21.6 0.576</td>
<td></td>
</tr>
<tr>
<td>CPU / Optimized</td>
<td>86.63 3.49</td>
<td>546.4 20.3</td>
<td>4,175.5 354.96</td>
<td></td>
</tr>
</tbody>
</table>
### Performance – Sampling + GPU

<table>
<thead>
<tr>
<th>Angular Sampling Combination</th>
<th>100x100 mm Field</th>
<th>64^3 Phantom 4 mm voxels</th>
<th>128^3 Phantom 2 mm voxels</th>
<th>256^3 Phantom 1 mm voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU / Naïve</td>
<td>2,100</td>
<td>4,100</td>
<td>8,200</td>
<td></td>
</tr>
<tr>
<td>CPU / Optimized</td>
<td>3,100</td>
<td>19,500</td>
<td>176,000</td>
<td></td>
</tr>
</tbody>
</table>
CONCLUSION

- Adaptive radiotherapy is made possible by GPU based algorithms.
  - 3D Deformable image registration
    - Head and neck – X50 speed-up
    - Lungs - X200 speed-up
  - 3D Biomechanical modeling for motion tracking
    - Head and neck – No comparison
    - Lungs - X200 speed-up
  - 3D Dose calculation
    - X4200 speed-up
ACKNOWLEDGEMENTS

- National Science Foundation
- Varian Inc
- US Office of Naval Research
- UCLA Radiation Oncology