Joint Segmentation and Deformable Registration of Fractured Vertebrae Using a Synthesis of the Expectation Maximization and Belief Propagation Algorithms

MITACS-Fields Conference on the Mathematics of Medical Imaging

Special Session on Graph-Based Methods in Medical Imaging

June 24, 2011

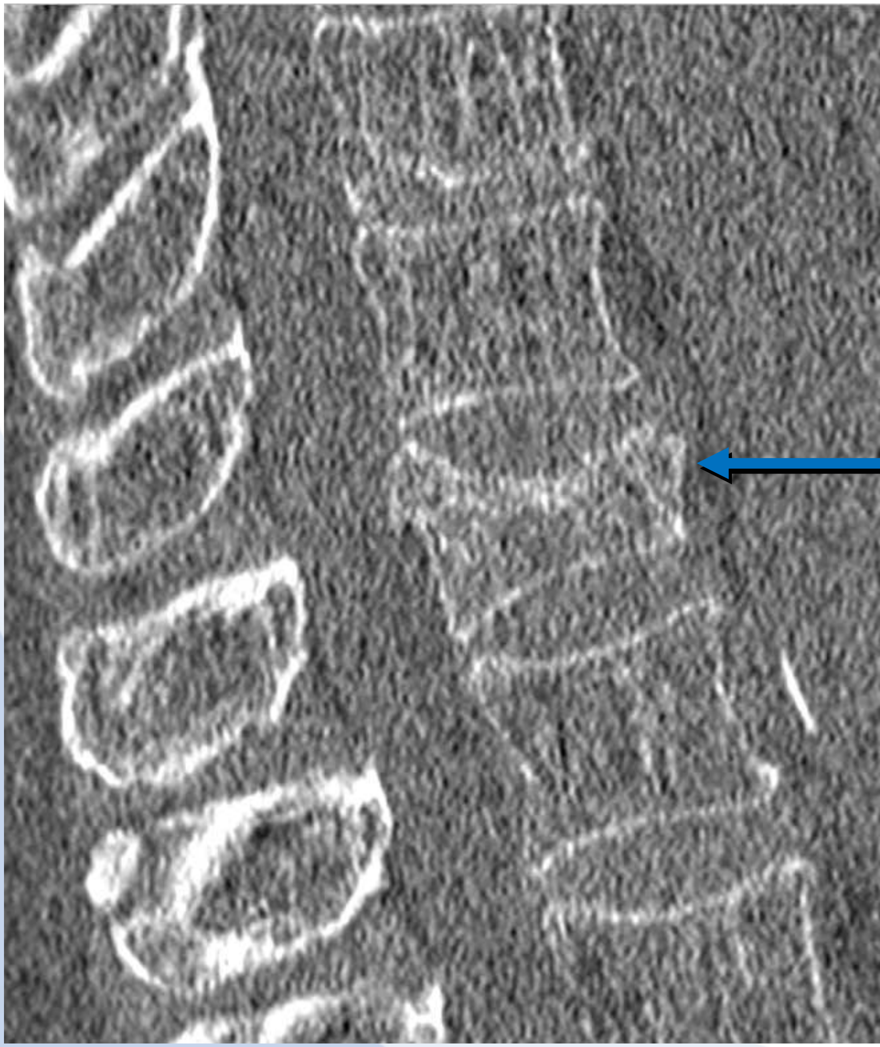
Paul Dufort, Ph.D.
Computational Imaging Scientist
University Health Network

The Joint Department of
MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

Overview

- Introduction of the clinical problem
- Overview of the proposed clinical solution
- Outline of the computational problem
- Description of the computational solution
- Demonstration of the program

The Clinical Problem

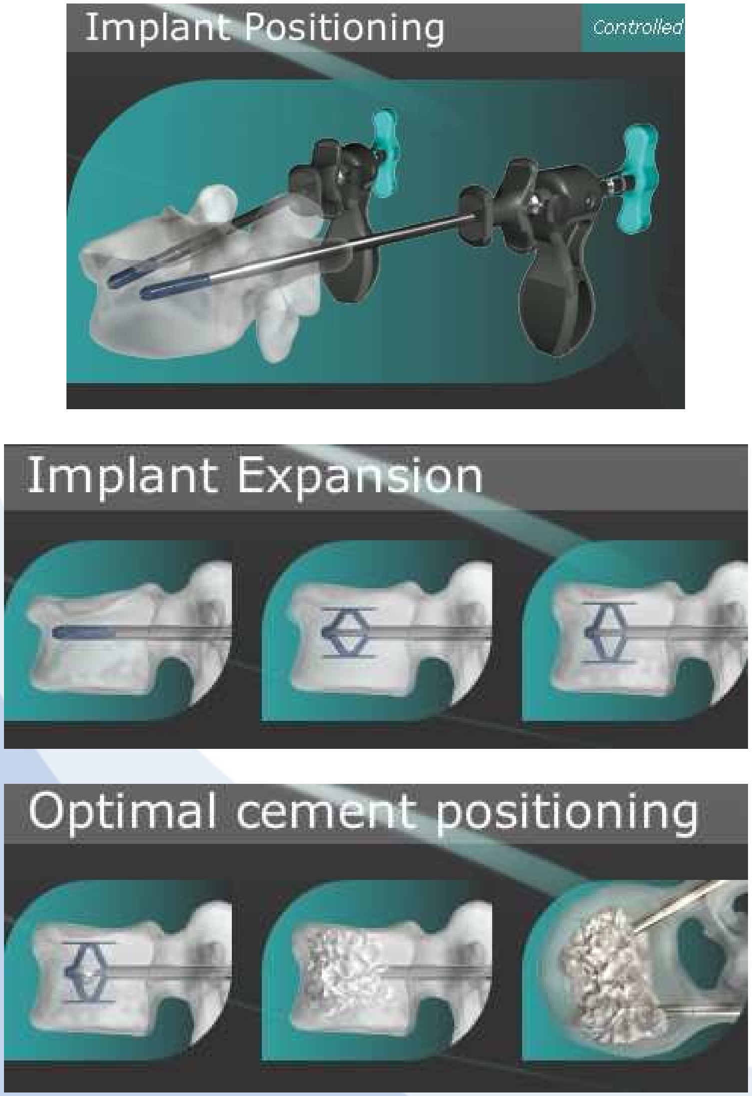


L1 vertebra with compression fracture

The Joint Department of MEDICAL IMAGING

The Proposed Clinical Solution

Instrument (SpineJack^{TM}) and surgical procedure



1) Insertion

2) Expansion

3) Cement

* From Vexim SAS web site

The Joint Department of MEDICAL IMAGING

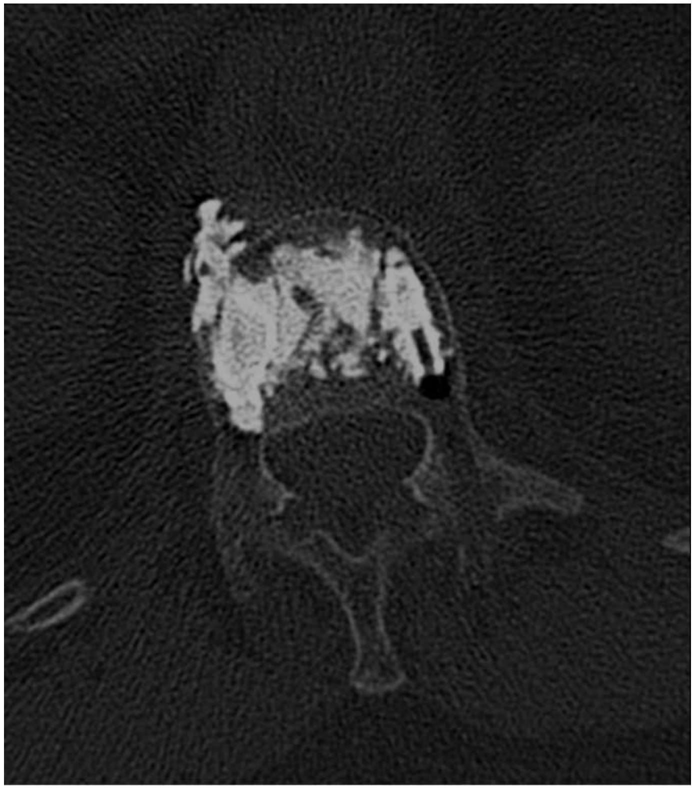
The Clinical Result

Sagittal Before

Sagittal After

The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

The Clinical Result



Axial After

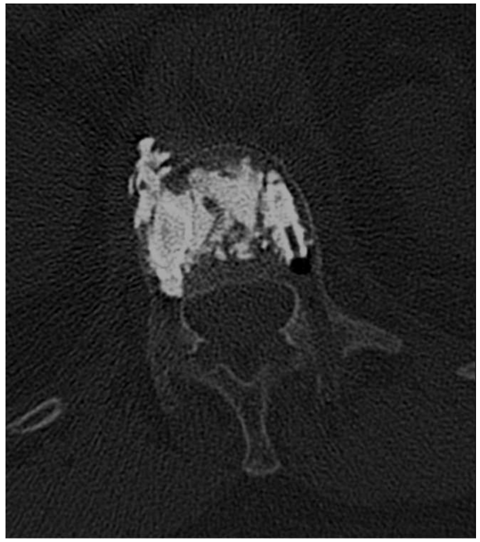
The Computational Problem

- Measure vertebral endplate height restoration given pre- and post-surgery CT scans
- Must work across a range of image qualities, vertebral deformations, and protocols
- Computational procedure must be automated to the greatest extent possible
- Procedure must not take more than two minutes

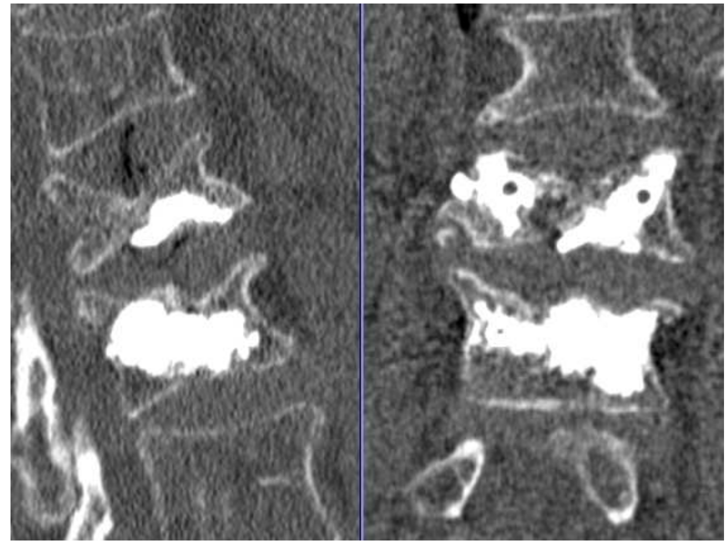
The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

The Computational Challenges

Extremely noisy images



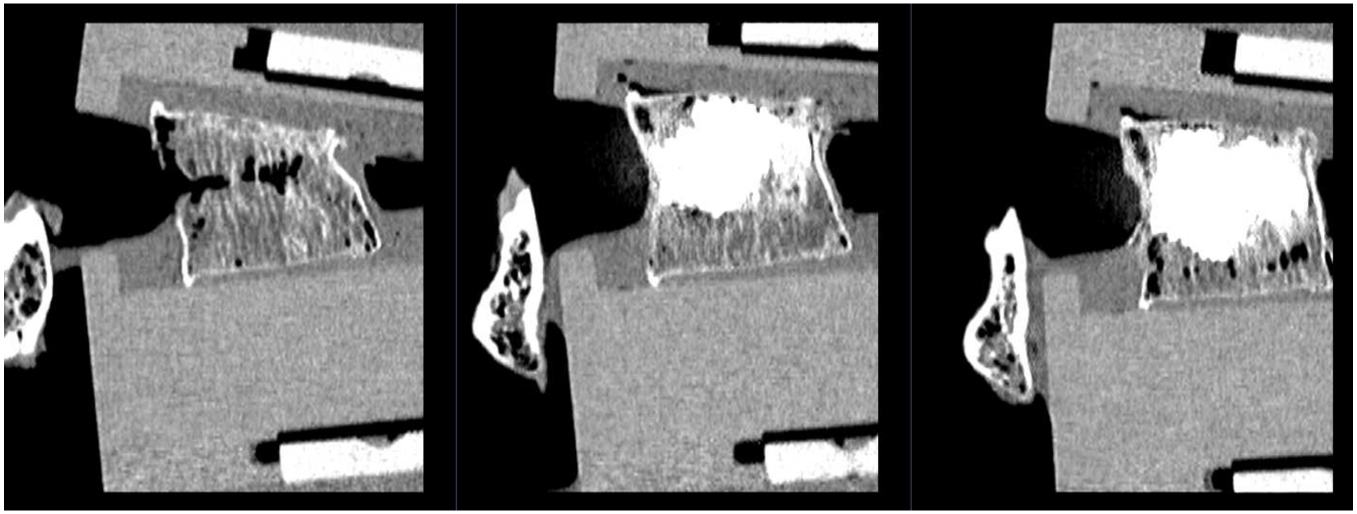
High intensity confounding materials



Extremely large deformations

The Joint Department of MEDICAL IMAGING

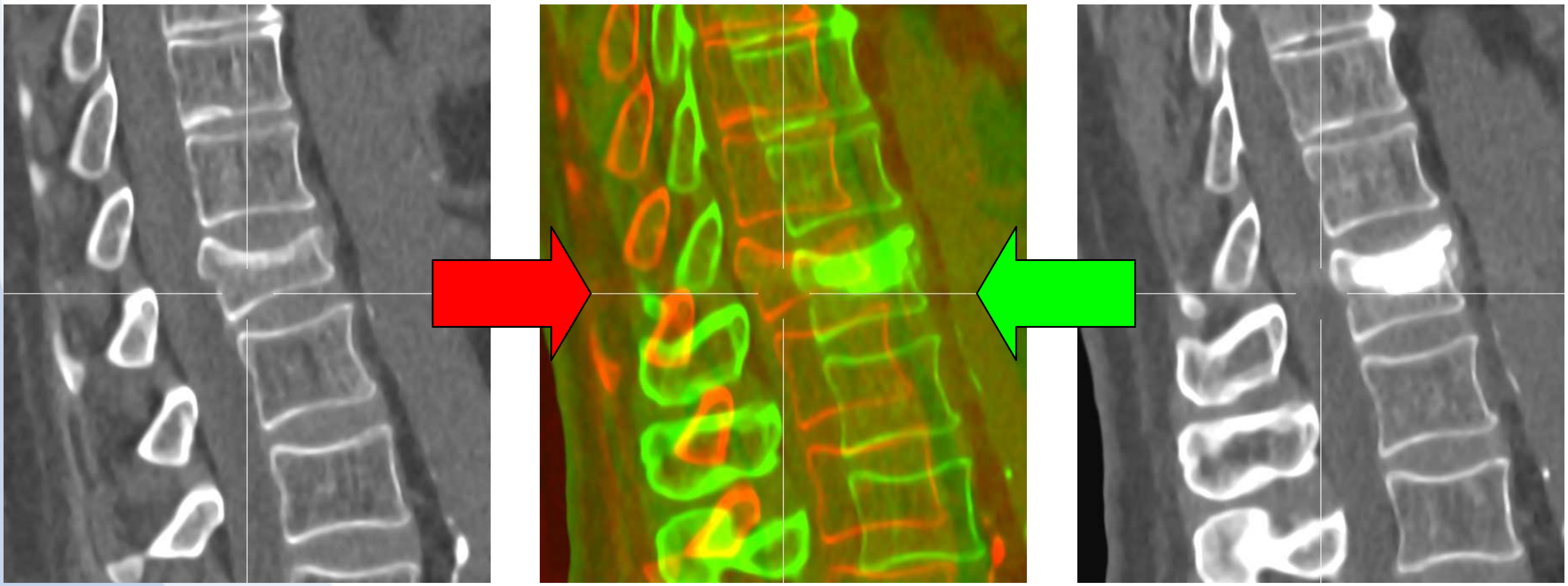
The Computational Challenges



Ex vivo protocol: completely different surrounding intensity context

The Computational Problem

Is this an image registration problem?



Seems like an image registration problem, but ...

The Joint Department of
MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

The Computational Problem

1) Global rigid registration won't work because:
   - (i) The spine is curved differently across scans
   - (ii) The vertebrae don't have the same shape
2) Global deformable registration won't work because:
   - (i) It will change the very shapes we are trying to measure
   - (ii) We could read the changes from the deformation field, but we won't know which vectors to examine

The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

The Computational Solution

Overall processing sequence:

1) Deformably register a segmented model vertebra to the pre- and post-surgery vertebrae in the CT scans

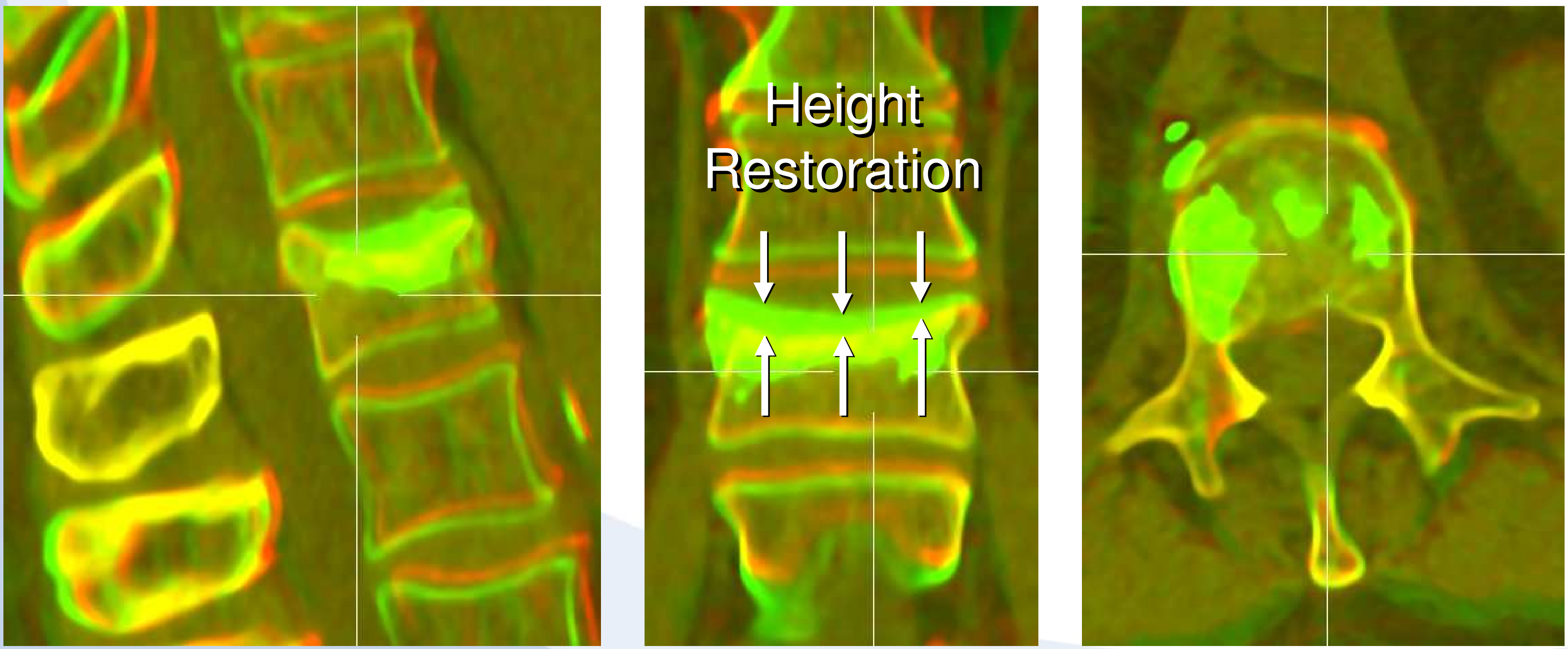
2) Use the correspondence to isolate the stable posterior regions based on segmentation of the model, and rigidly register these based on intensity

3) Use the correspondence to identify the vertebral endplates based on segmentation of the model, and measure the distance between them

The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

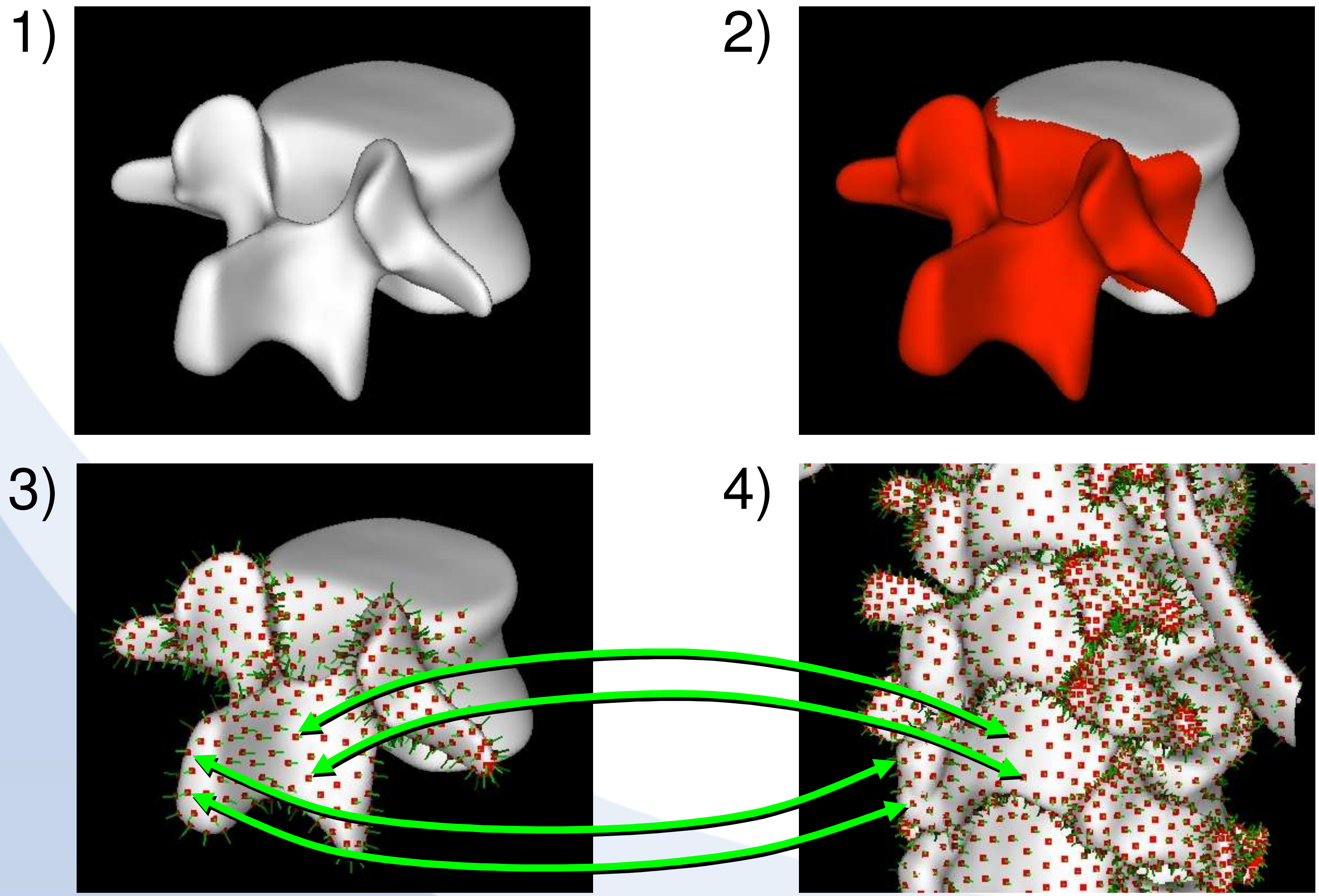
The Computational Solution

- Segment the posterior region of target vertebra in before and after scans using the correspondence with the segmented model vertebra
- Rigidly register posterior regions using a standard intensity-based approach
- Use the rigid transformation on the whole image so that the vertebral bodies are carried “along for the ride”



The Computational Solution

Use correspondence between features extracted from model and features extracted from data to deformably register model to data



The Expectation-Maximization Algorithm

- Formalized by Dempster, Laird and Rubin in 1977 to allow maximum likelihood estimation of parameters from data sets with missing variables

Height	Age	Weight
183 cm	34 yrs	95 kg
141 cm	?	67 kg
?	26 yrs	71 kg
165 cm	53 yrs	?

The Expectation-Maximization Algorithm

- Repeat until (guaranteed) convergence:

1) E-Step: Compute expectation of log-likelihood, using probability distribution over missing variables given current MLE estimates

$$Q\left(\Theta, \Theta^{(t-1)}\right) = \int \log\left[p\left(x, y \middle| \Theta\right)\right] p\left(y \middle| x, \Theta^{(t-1)}\right) dy$$

$$x = \text{Observed data}$$

$$y = \text{Missing data}$$

$$\Theta^{(t-1)} = \text{MLE parameters at step } t-1$$

2) M-Step: Maximize Q as a function of Θ

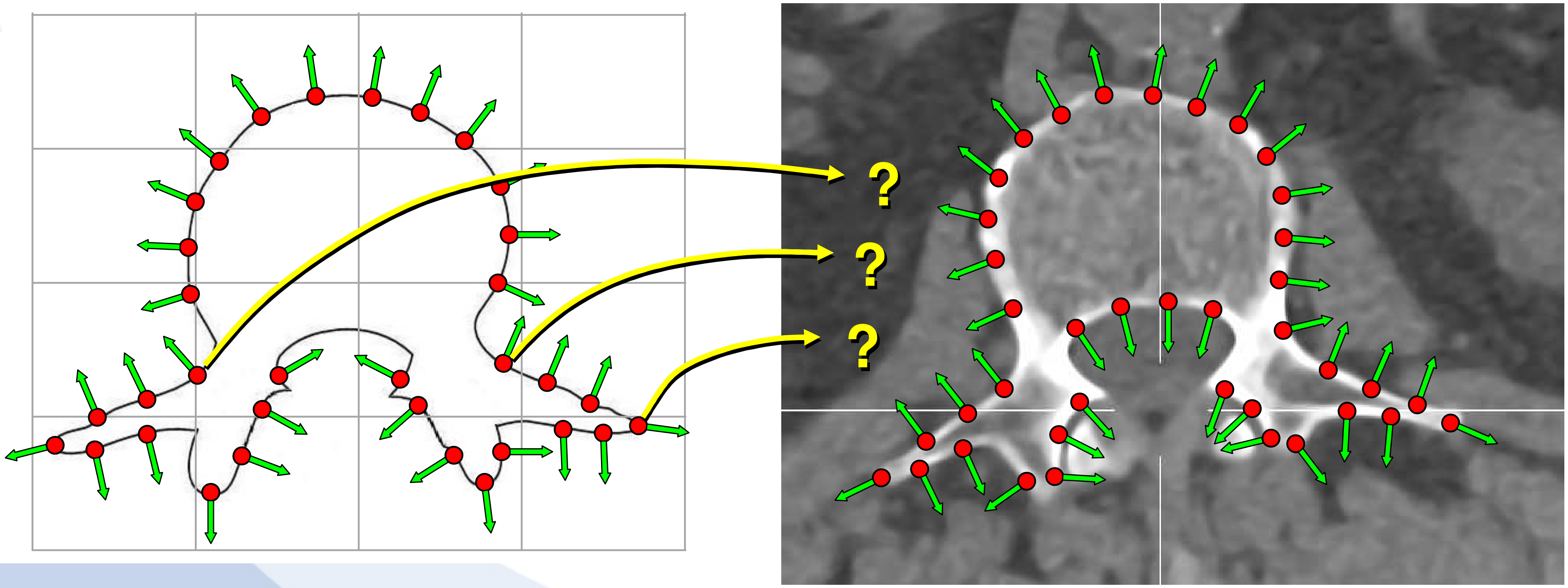
$$\Theta^{(t)} = \arg\min_{\Theta} Q\left(\Theta, \Theta^{(t-1)}\right)$$

The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

The EM Algorithm: Feature Correspondence

Joint deformable registration and feature correspondence with EM. Answer two questions at the same time:

1) Which model features should be mapped to which data features?

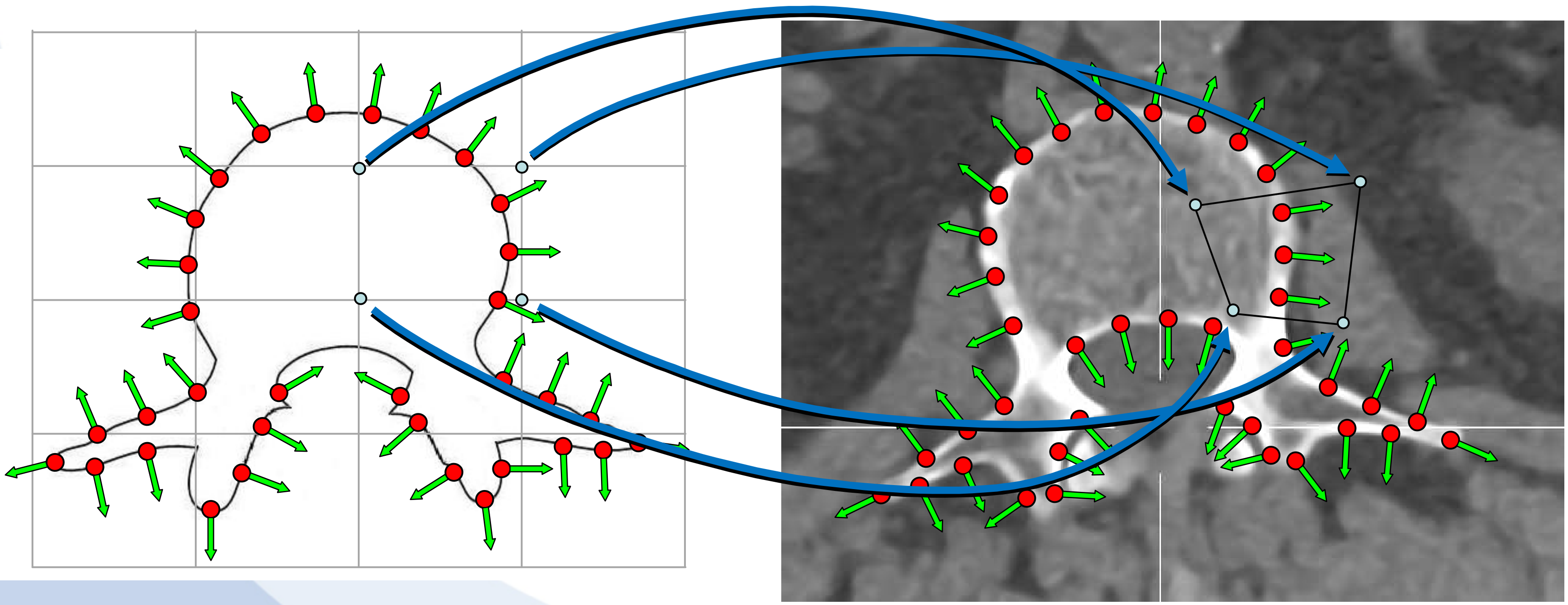


The Joint Department of MEDICAL IMAGING

The EM Algorithm: Feature Correspondence

Joint deformable registration and feature correspondence with EM. Answer two questions at the same time:

2)How should we deform the grid to make each model feature map to the correct data feature?



The EM Algorithm: Feature Correspondence

Joint deformable registration and feature correspondence with EM:

•Let the deformation field vectors be the maximum likelihood parameters to be estimated

•Let the correspondence variables mapping each model feature to one of many possible data features be the "missing data"

E-Step

$$p_{ij} = \exp\left[-\frac{\left(\vec{d}_{ij} - \vec{m}_i - \vec{F}^{(t-1)}(\vec{m}_i)\right)^2}{2\sigma_F^2}\right] \Bigg/ \left(u + \sum_{j \neq u} \exp\left[-\frac{\left(\vec{d}_{ij} - \vec{m}_i - \vec{F}^{(t-1)}(\vec{m}_i)\right)^2}{2\sigma_F^2}\right]\right)$$

p_{ij} = Probability that model feature i is mapped to data feature j

$\vec{d}_{ij}$ = Position of data feature j for model feature i

$\vec{m}_i$ = Position of model feature i

$\vec{F}^{(t-1)}$ = Deformation field at step t-1

σ_F = Field tolerance

u = Probability that model feature is unused

The Joint Department of MEDICAL IMAGING

Mount Sinai Hospital • University Health Network • Women's College Hospital

The EM Algorithm: Feature Correspondence

Joint deformable registration and feature correspondence with EM:

- Let the deformation field vectors be the maximum likelihood parameters to be estimated
- Let the correspondence variables mapping each model feature to one of many possible data features be the "missing data"

M-Step

$$\vec{F}^{(t)} = \arg\min_{\vec{F}} \sum_i \sum_{j \neq u} p_{ij} \left[\vec{d}_{ij} - \vec{m}_i - \vec{F}(\vec{m}_i) \right]^2 + \alpha \left[\nabla \vec{F} \right]^2 + \beta \left[\nabla^2 \vec{F} \right]^2$$

$\alpha = \text{Harmonic regularization weight}$

$\beta = \text{Biharmonic regularization weight}$

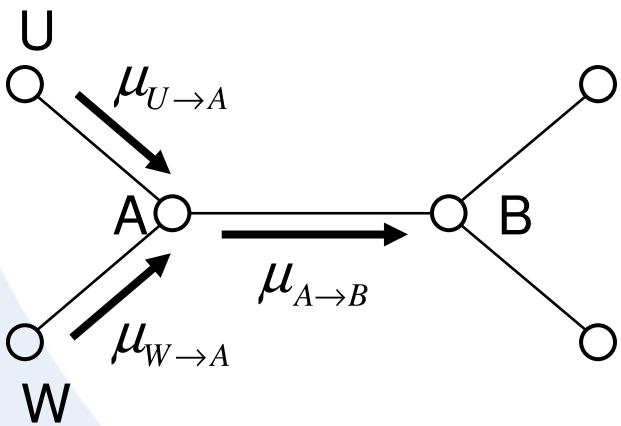
- After discretization, can solve as a sparse linear system in one step

The EM Algorithm: Limitations

- EM is a gradient ascent algorithm, so while guaranteed to converge, it will only converge to the maximum closest to its starting point in the parameter space
- Can accommodate a certain amount of clutter by keeping the regularization stiff until convergence to the vicinity of the target, then allowing greater deformation to occur
- With greater amounts of clutter, the algorithm will likely converge to a spurious maximum

Belief Propagation

- First described by Judea Pearl in 1988 for finding marginals and modes of probability distributions over tree-structured graphs

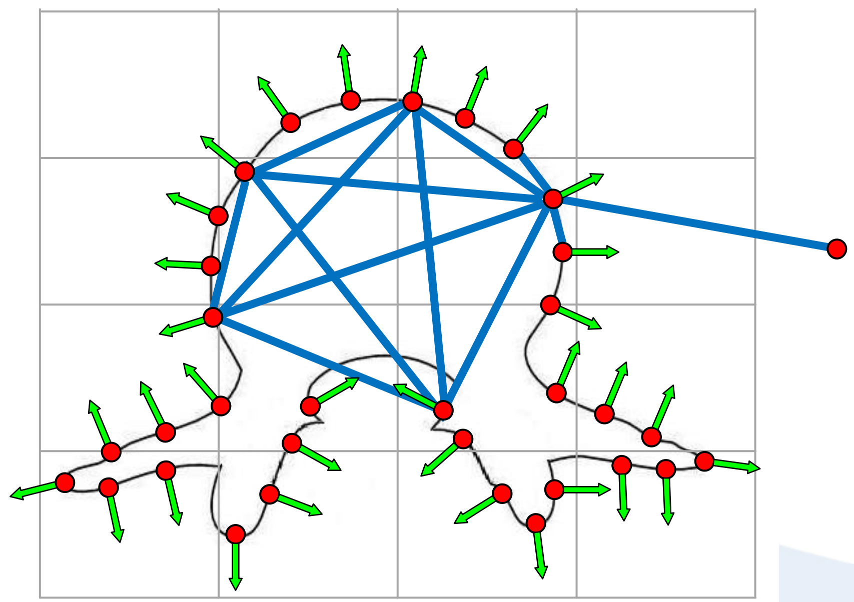


$$\mu_{A \to B}(x_B) = \min_{x_A} \left[E(x_A) + E(x_A, x_B) + \sum_{\substack{V \in N_A \\ V \neq B}} \mu_{V \to A}(x_A) \right]$$

- Converges to exact solution in finite number of steps for tree-structured graphs
- Convergence for “loopy” graphs not well characterized – links to Bethe free energy of statistical physics – but often works well
- Has revolutionized some fields – turbo-codes

EM + Belief Propagation

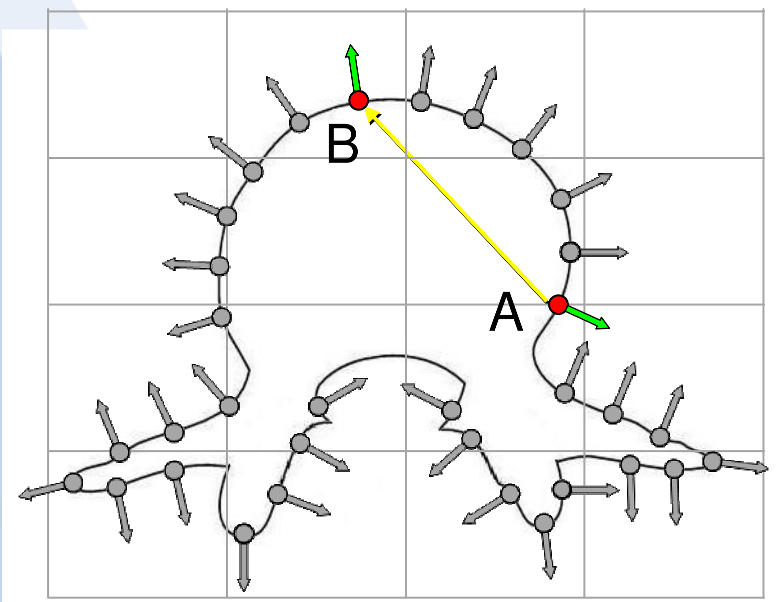
- Instead of using probabilities p_{ij} from E-step directly in the M-step, use them as priors in a Markov Random Field
- Node variables are “missing data” from EM – mappings from each model feature *i* to its possible data features *j*
- Edges enforce coherence between pairs of mappings from two model features to their respective data features

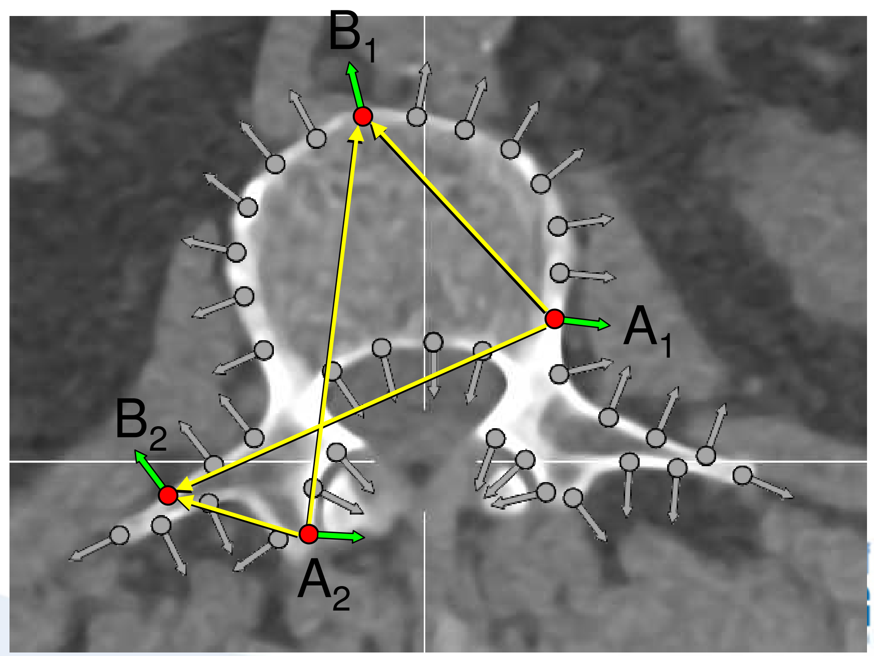


The Joint Department of MEDICAL IMAGING

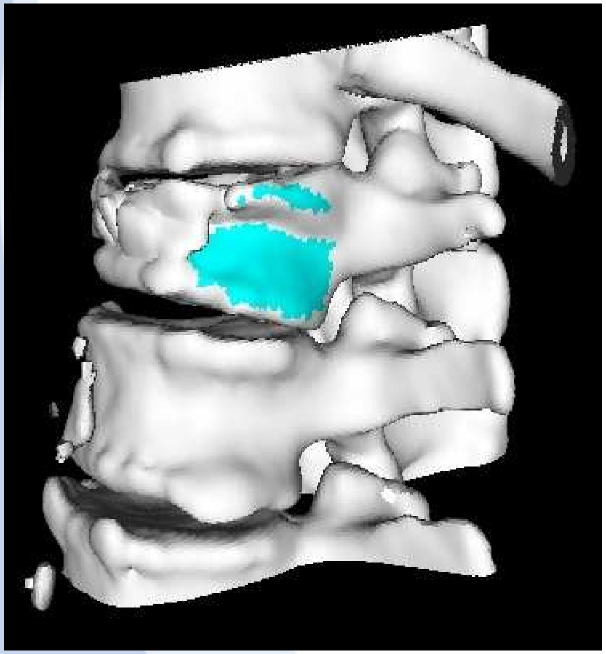
EM + Belief Propagation

- Match coherence is computed as the sum of squares of the differences between three pairs of normal and separation vector angles. The measure is translation, rotation, and scale invariant
- Scale-dependence can be computed (and weighted) separately based on the difference in the lengths of the separation vectors

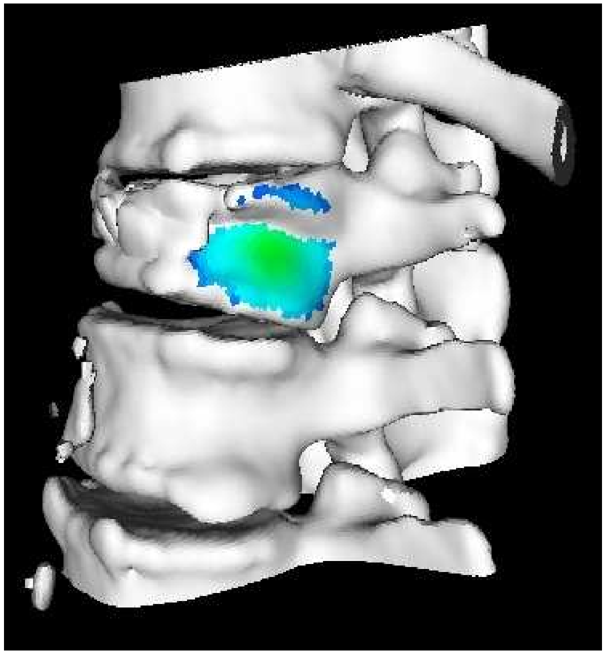




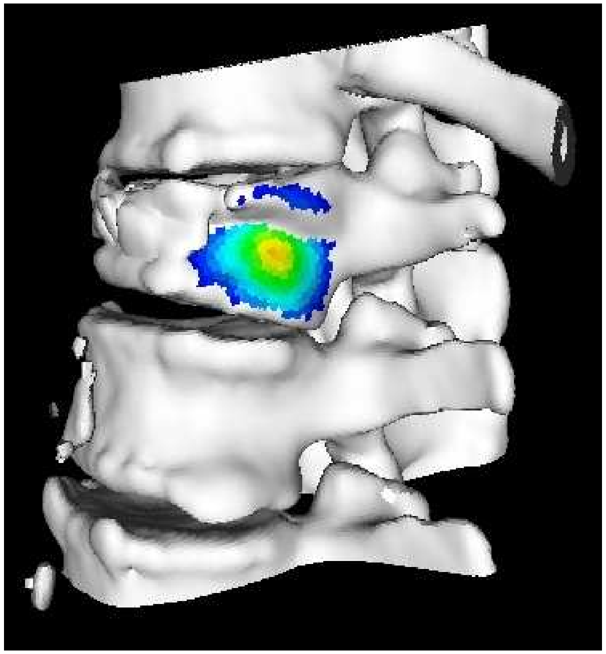
EM + BP: Iterations



Iteration 0



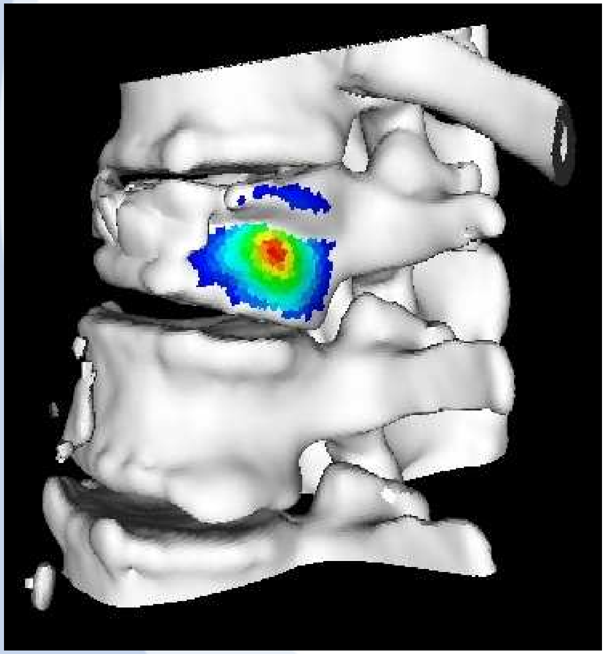
Iteration 1



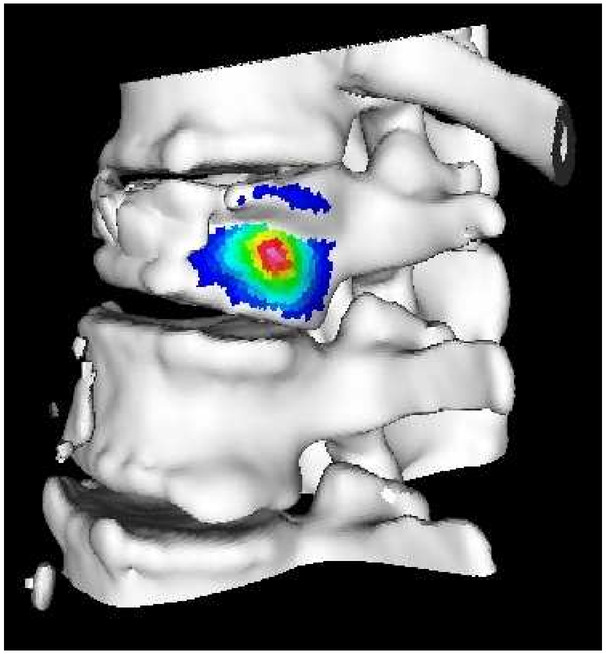
Iteration 2

The Joint Department of MEDICAL IMAGING

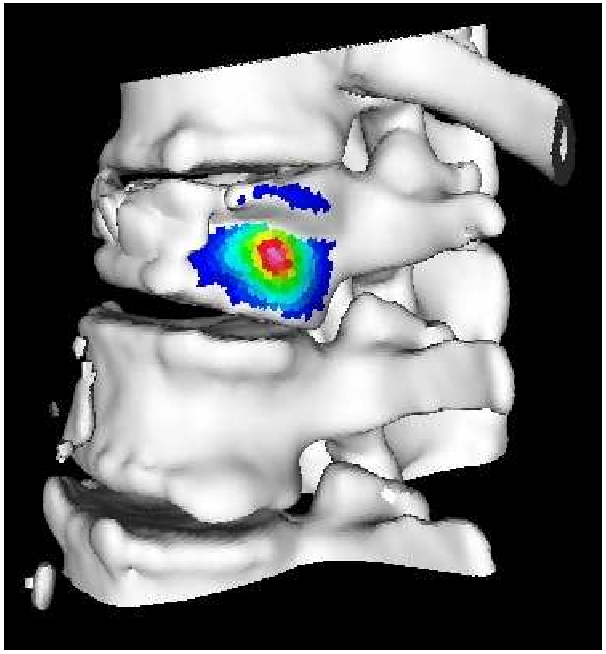
EM + BP: Iterations



Iteration 3



Iteration 5



Iteration 7

The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

EM + Belief Propagation: Experience

- Belief propagation on the min-sum ring is used to find the approximate mode of the probability distribution
- The BP max-marginals are fed back to the M-step of EM to update the deformable registration field
- Schedule: best results are obtained when one iteration of EM is interleaved with one iteration of BP
- Convergence is rapid – typically 10-20 iterations
- Much more robust to surrounding clutter than EM alone
- Major limitation is computational complexity: $O(MCD^2)$ where

  M = Number of features extracted from model
  C = Number of neighbours for each model feature in MRF
  D = Possible data feature matches for each model feature

EM + Belief Propagation: Demonstration

- Research conducted over the past year, software developed for Vexim SAS over the past six months
- Requires one mouse click at the approximate centre of the target vertebral body (to limit the volume of data that needs to be processed)
- Processing is completed in under two minutes for a typical 400 x 512 x 512 voxel CT scan

The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital

Conclusions

- Clinical problem: vertebral body height restoration following compression fracture
- Proposed clinical solution: SpineJack^{TM} by Vexim SAS
- Computational problem: measure the vertebral height restoration achieved in studies varying widely in image quality, degree of deformation, and protocol
- Computational solution: a hybrid application of the EM and BP algorithms to deformably register a segmented model vertebra to the CT data

The Joint Department of MEDICAL IMAGING
Mount Sinai Hospital • University Health Network • Women's College Hospital