

# Inter and Intra-Modal Deformable Registration: Continuous Deformations Meet Efficient Optimal Linear Programming

Ben Glocker<sup>1,2</sup>, Nikos Komodakis<sup>1,3</sup>, Nikos Paragios<sup>1</sup>, Georgios Tziritas<sup>3</sup>, Nassir Navab<sup>2</sup>

4 July 2007

<sup>1</sup>GALEN group | laboratoire de mathématiques appliquées aux systèmes | école centrale paris

<sup>2</sup>computer aided medical procedures & augmented reality | technische universität münchen

<sup>3</sup>computer science department | university of crete

# Outline

- Introduction & Motivation
- Image Registration based on Discrete Labeling
- Optimization using Linear Programming
- Results & Conclusions

***In this presentation everything is intensity-based***

# Introduction

- Source and target image

$$f : \Omega \rightarrow \mathcal{R} \quad \Omega \subset \mathcal{R}^d \quad \text{with} \quad d \in \{2, 3\}$$

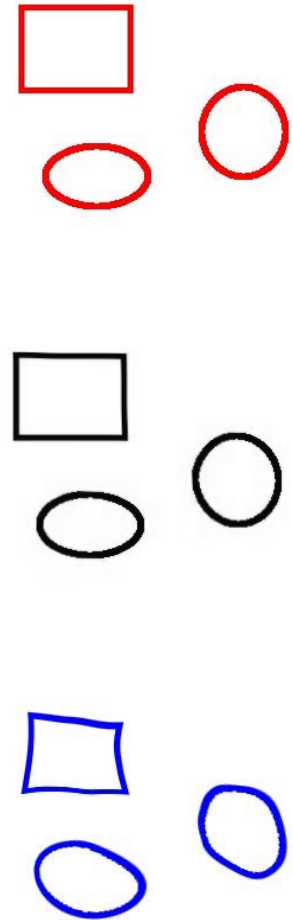
$$g : \Omega \rightarrow \mathcal{R}$$

- Image relation

$$g(\mathbf{x}) = h \circ f(\mathcal{T}(\mathbf{x}))$$

$$\mathcal{T} : \Omega \rightarrow \Omega \quad \text{non-linear transformation}$$

$$h : \mathcal{R} \rightarrow \mathcal{R} \quad \text{non-linear relation on intensities}$$



# Registration as an Optimization Problem

- Energy formulation

$$E(\mathcal{T}) = \int_{\Omega} \rho_h(g(\mathbf{x}), f(\mathcal{T}(\mathbf{x}))) d\mathbf{x} \rightarrow \min!$$

$\rho_h : \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R}$  distance measure

- The aim of registration is to recover the transformation which involves
  - the definition of a **transformation type**
  - the definition of a **distance/similarity measure**
  - the definition of an **optimization procedure**

# Review of Registration Methods

- Types of transformations
  - Rigid, affine, projective
  - Basis functions, Spline-based
  - Finite Element Models, ...
- Distance/Similarity measures
  - SAD, SSD, NCC, NMI, CR, ...
- Optimization methods
  - Variational
  - Gradient-based
  - Direct search (Simplex, Powell-Brent, Best Neighbor)

# Motivation

- What is expected from an optimal registration method?
  - Independent from the choice of the transformation type
  - Independent from the choice of the distance/similarity measure
  - Guarantee of a globally optimal solution
  - Reasonable computational complexity

# Our Contributions

- Novel deformable registration framework based on discrete labeling and linear programming
- Our framework bridges the gap between continuous deformations and discrete optimization
- Gradient-free and flexible in the choice of the distance measure
- Guaranteed optimality properties on the solution
- Computational efficient and tractable

# Image Registration based on Discrete Labeling

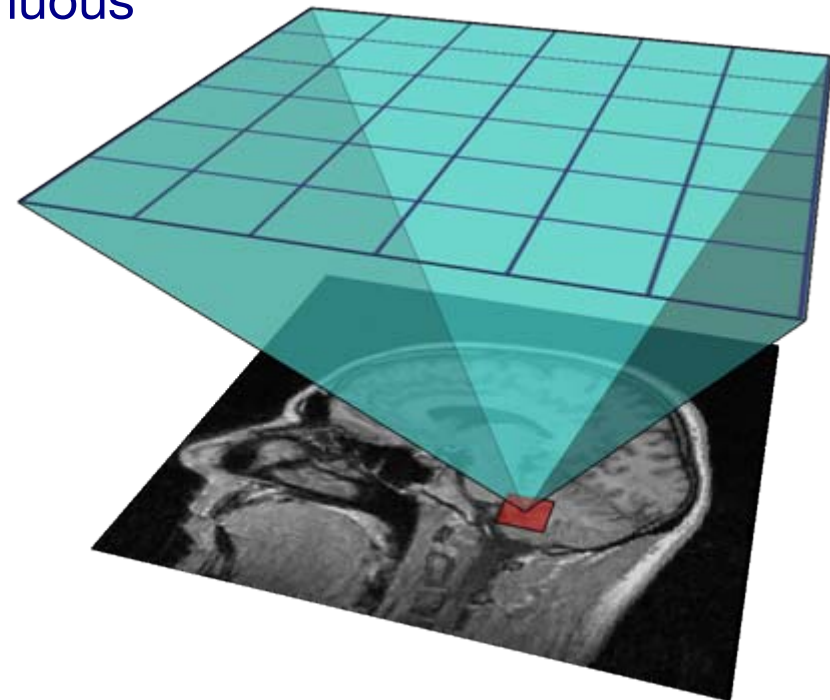


# Local Registration

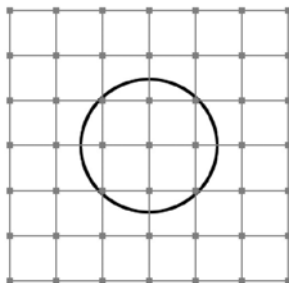
- Deformation grid providing a continuous and dense deformation field

$$\mathcal{T}(\mathbf{x}) = \mathbf{x} + \mathcal{D}(\mathbf{x})$$

with  $\mathcal{D}(\mathbf{x}) = \sum_{p \in \mathcal{G}} \eta(|\mathbf{x} - \mathbf{p}|) \mathbf{d}_p$



- In our implementation we use Free Form Deformation



$$\mathcal{D}(\mathbf{x}) = \sum_{l=0}^3 \sum_{m=0}^3 B_l(u) B_m(v) \mathbf{d}_{i+l, j+m}$$

[Rueckert99, Schnabel01, Rohlfing03, ...]

# Energy Formulation

- Reformulation of the optimization problem

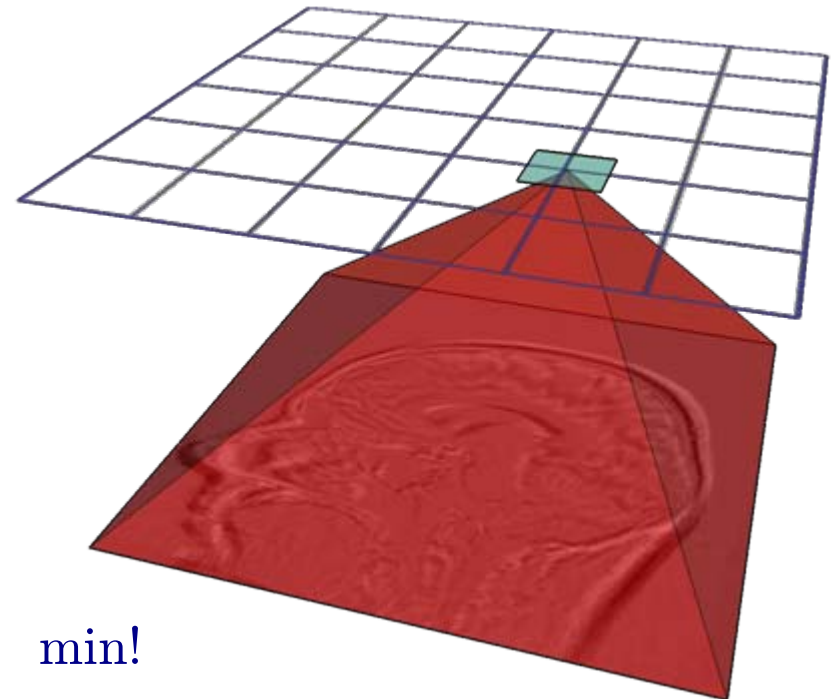
$$E_{\text{data}}(\mathcal{T}) = \sum_{p \in \mathcal{G}} \int_{\Omega} \eta^{-1}(|\mathbf{x} - \mathbf{p}|) \cdot \rho_h(g(\mathbf{x}), f(\mathcal{T}(\mathbf{x}))) d\mathbf{x}$$

- Smoothness term

$$E_{\text{smooth}}(\mathcal{T}) = \sum_{p \in \mathcal{G}} \phi(|\nabla_{\mathcal{G}} \mathbf{d}_p|)$$

- Registration task

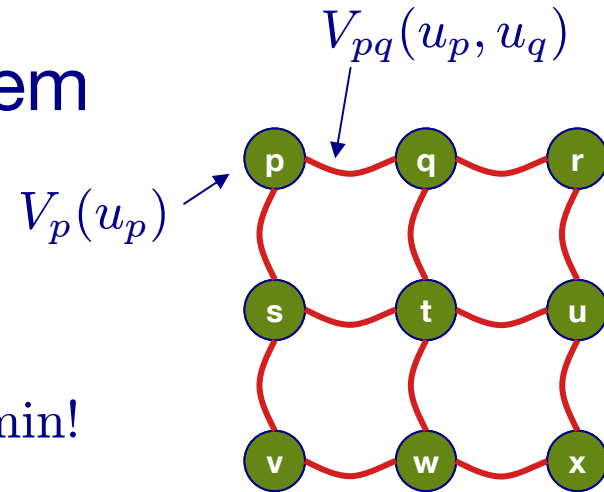
$$E_{\text{total}} = E_{\text{data}} + E_{\text{smooth}} \rightarrow \min!$$



# Discrete Optimization Problem

- General energy formulation

$$E_{\text{total}} = \underbrace{E_{\text{data}}}_{\text{green}} + \underbrace{E_{\text{smooth}}}_{\text{red}} \rightarrow \min!$$



- Markov Random Field (MRF) formulation for discrete labelings

$$E_{\text{total}}(u) = \sum_{p \in \mathcal{G}} \underbrace{V_p(u_p)}_{\text{green}} + \sum_{p, q \in \mathcal{E}} \underbrace{V_{pq}(u_p, u_q)}_{\text{red}}$$

Data term = singleton potentials

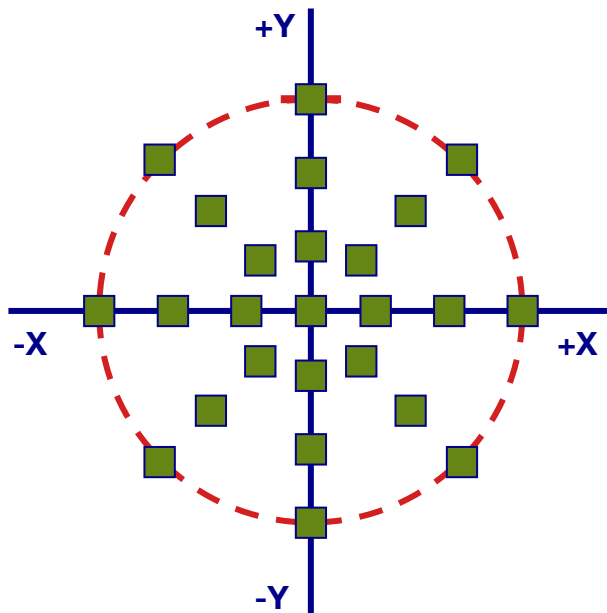
Smoothness term = pairwise potentials

# Discretization of Parameter Space

- Set of labels and a discretized deformation space

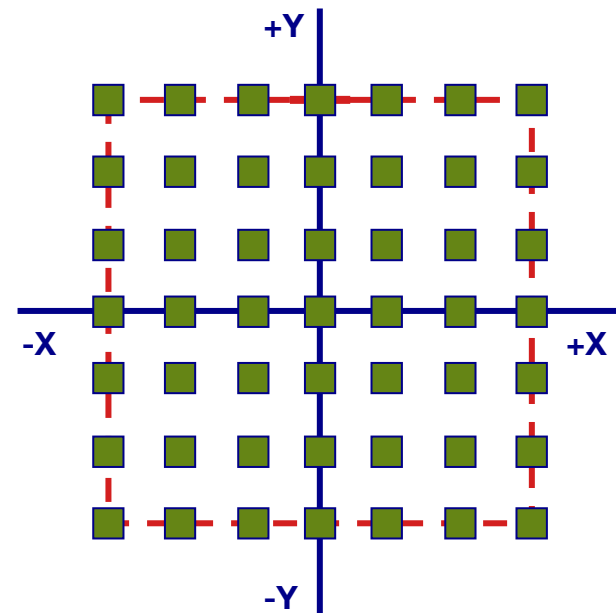
$$\mathcal{L} = \{u^1, \dots, u^i\}$$

$$\Theta = \{\underline{d^1}, \dots, \underline{d^i}\}$$



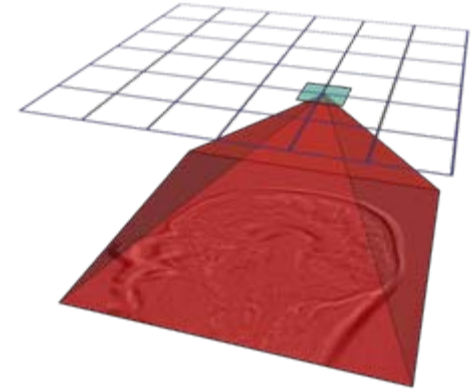
Sparse sampling

— Max  
■ Steps



Dense sampling

# Data Term



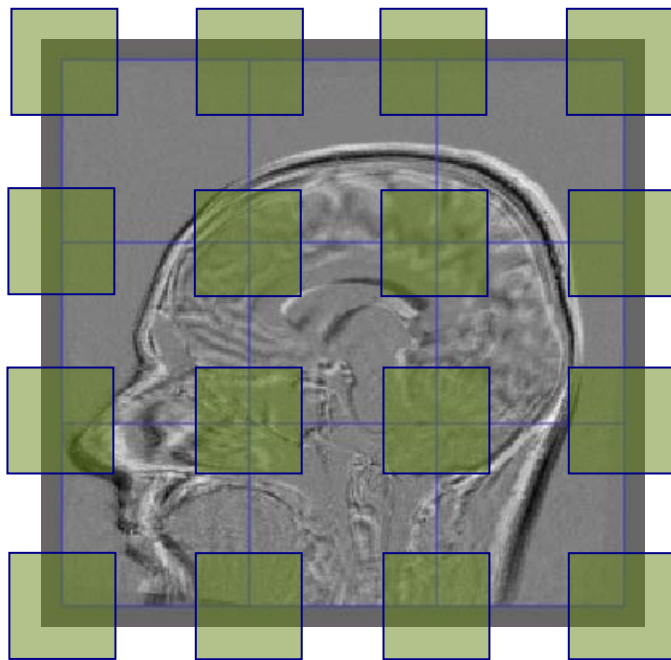
- MRF singleton potentials:  $|\mathcal{L}| \times |\mathcal{G}|$  cost matrix

$$\begin{aligned} E_{\text{data}}(u) &= \sum_{p \in \mathcal{G}} \int_{\Omega} \eta^{-1}(|\mathbf{x} - \mathbf{p}|) \rho_h(g(\mathbf{x}), f(\mathcal{T}(\mathbf{x}))) d\mathbf{x} \\ &\approx \sum_{p \in \mathcal{G}} V_p(u_p) \end{aligned}$$

- **Problem: singleton potentials are not independent!**

# Fast Approximation of Singleton Potentials

- Approximation of label costs simultaneously for all nodes



Current label  $x=0$   
 $y=0$

Nodes

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Single potential look-up table





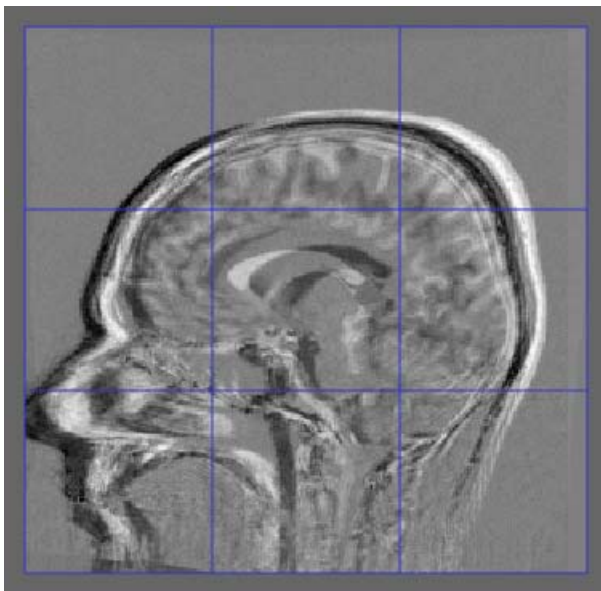




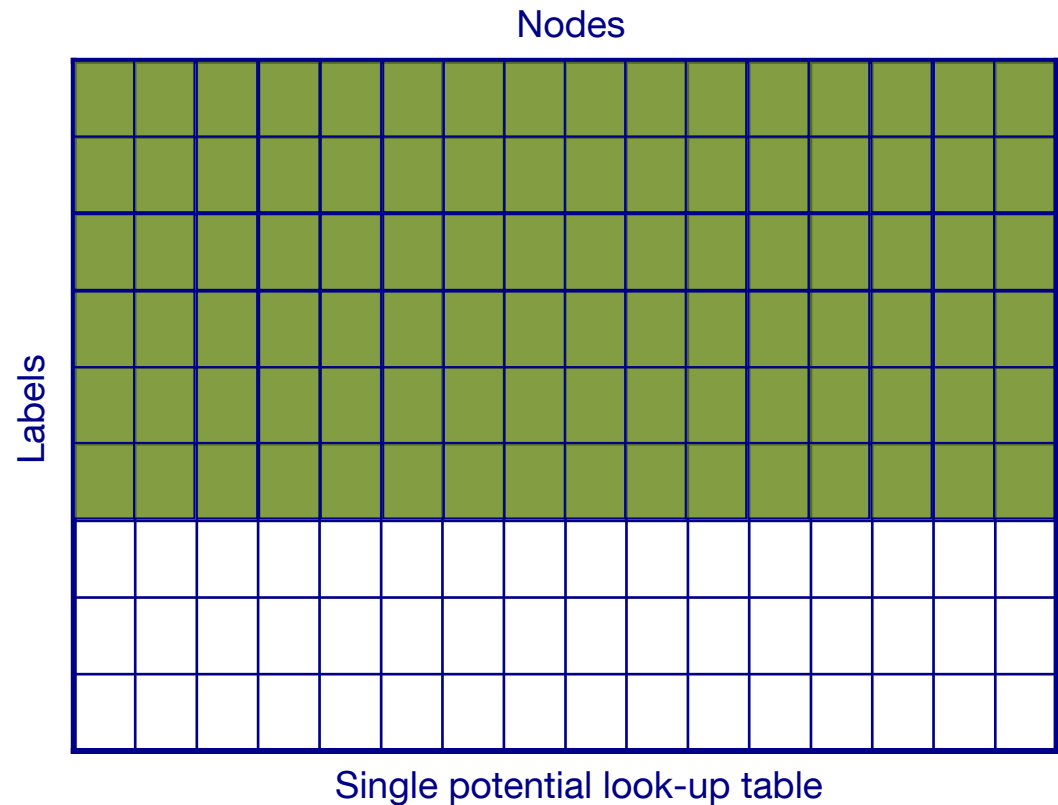


# Fast Approximation of Singleton Potentials

- Approximation of label costs simultaneously for all nodes

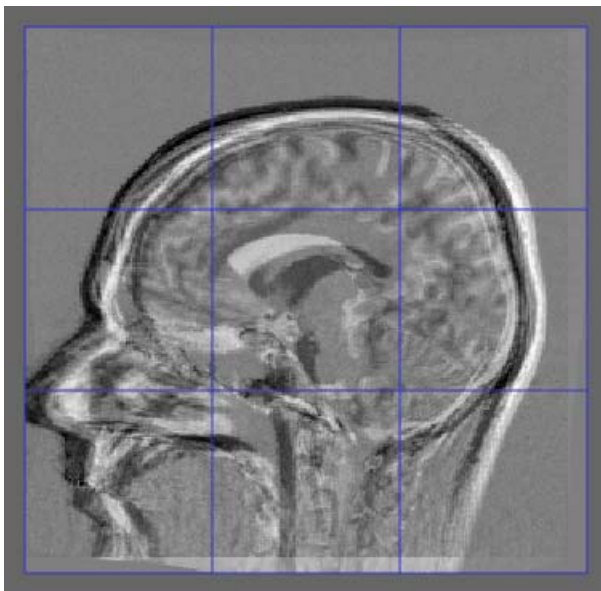


Current label  $x=-10$   
 $y=0$

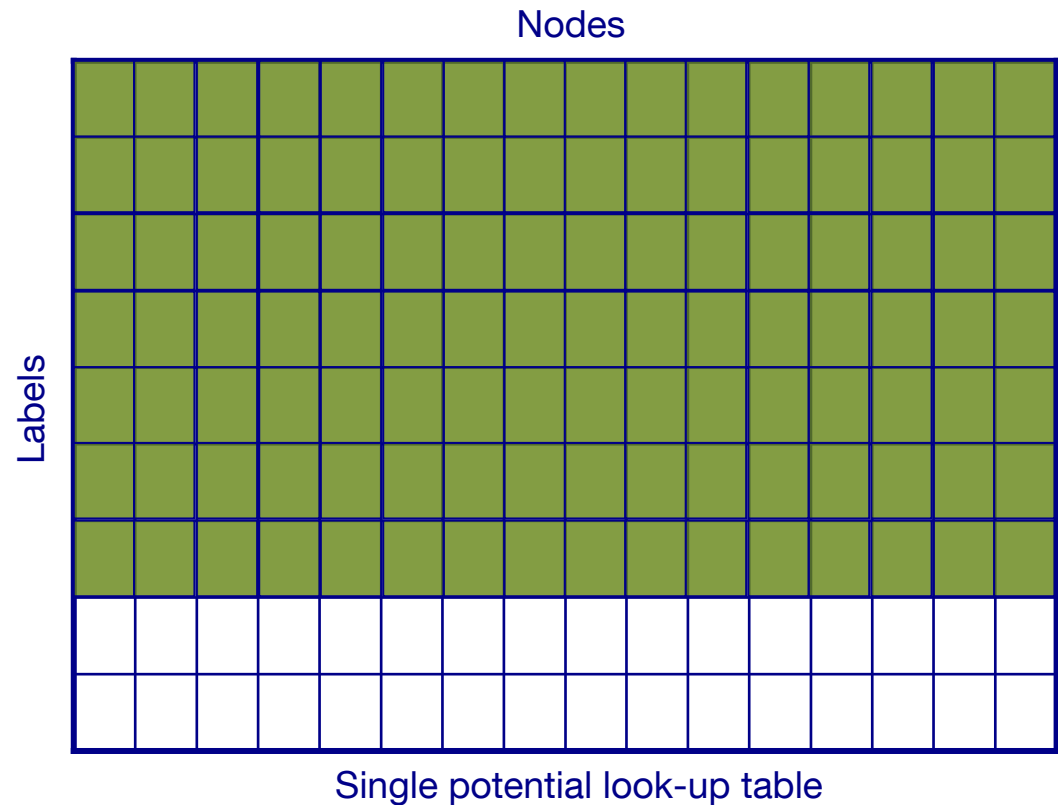


# Fast Approximation of Singleton Potentials

- Approximation of label costs simultaneously for all nodes

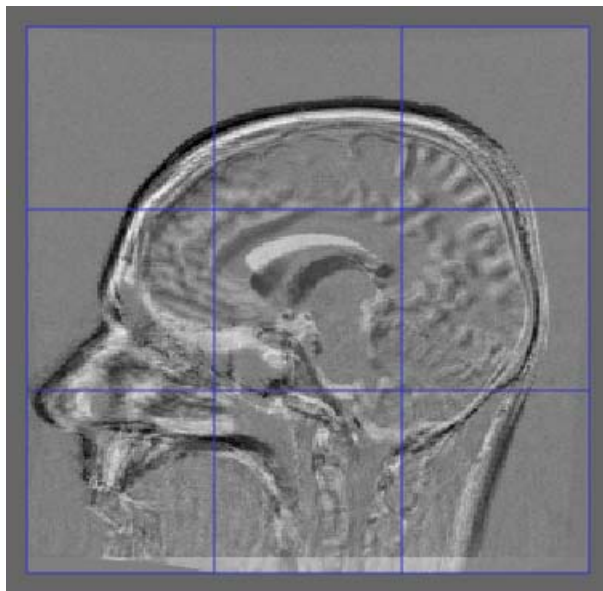


Current label  $x=-10$   
 $y=10$

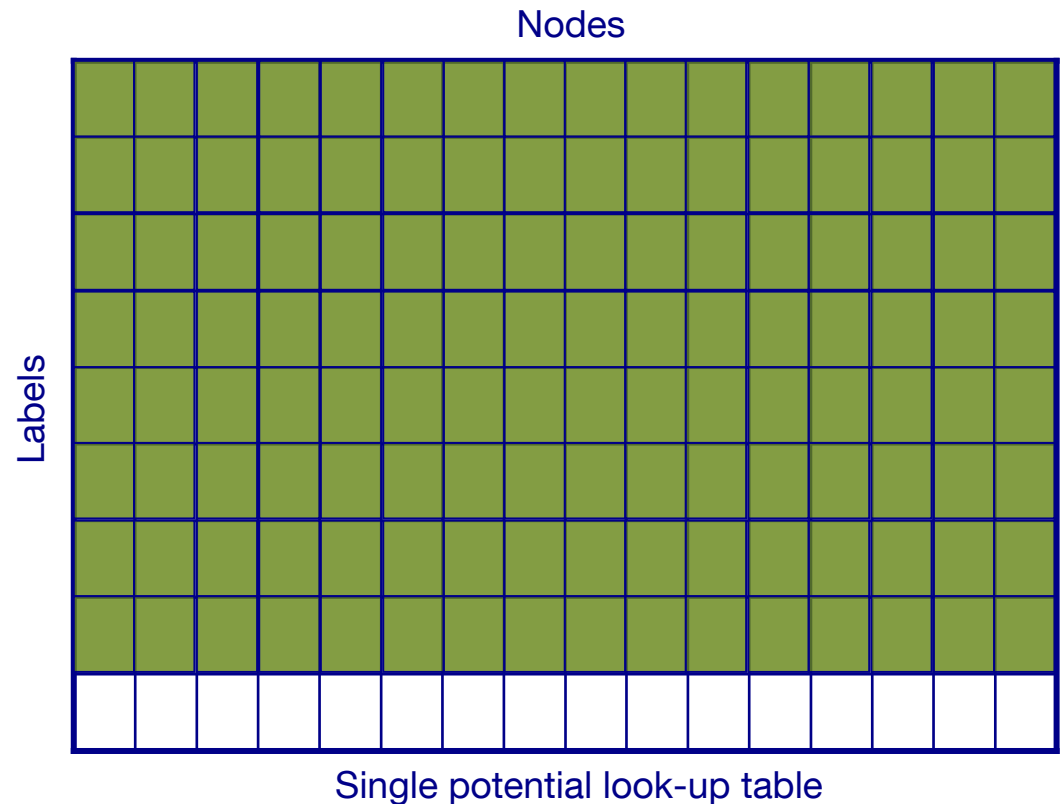


# Fast Approximation of Singleton Potentials

- Approximation of label costs simultaneously for all nodes

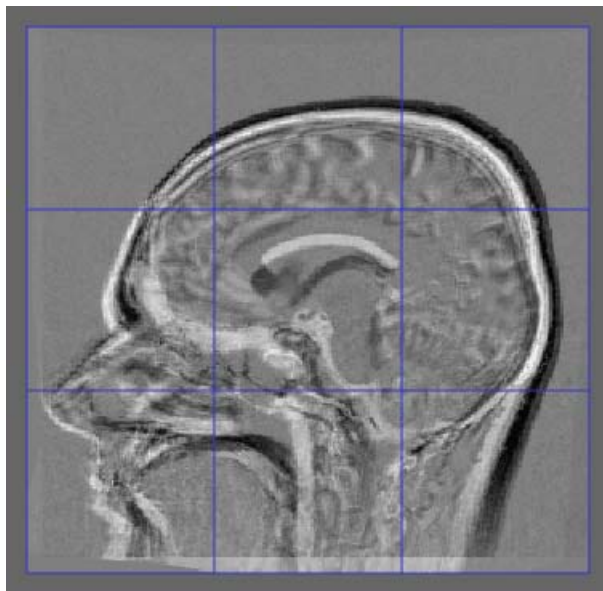


Current label  $x=0$   
 $y=10$

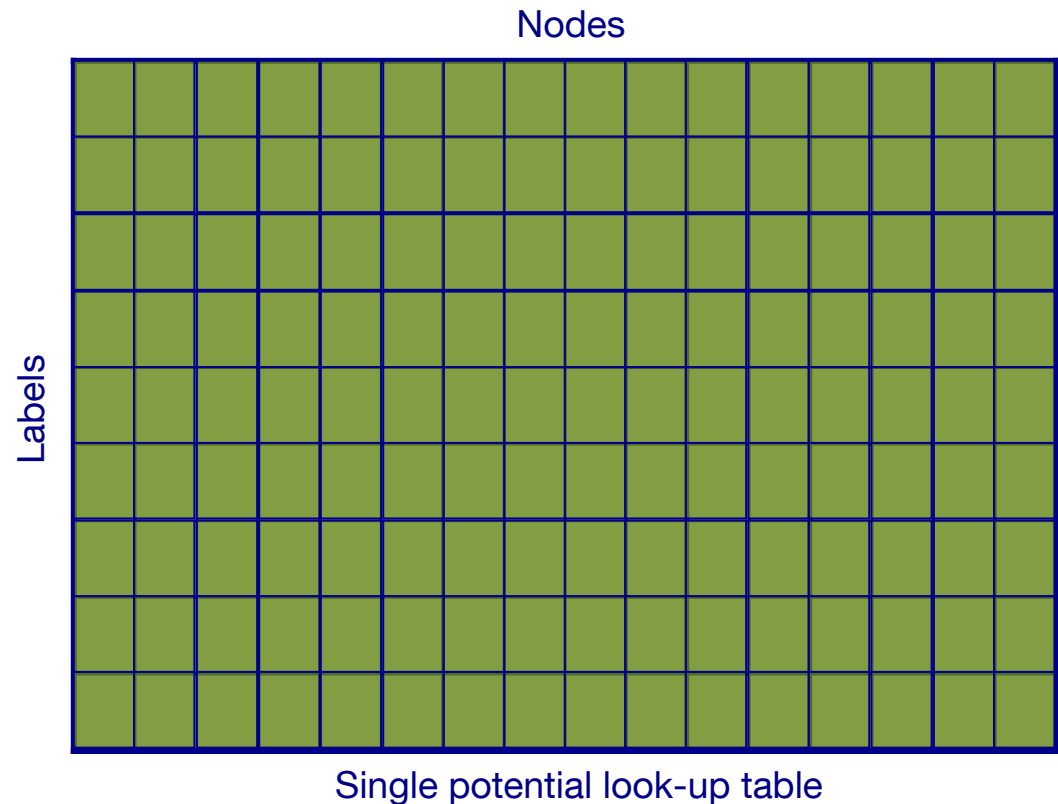


# Fast Approximation of Singleton Potentials

- Approximation of label costs simultaneously for all nodes



Current label  $x=10$   
 $y=10$



# Smoothness Term

- MRF pairwise potentials:  $|\mathcal{L}| \times |\mathcal{L}|$  cost matrix

$$E_{\text{smooth}}(u) = \sum_{p,q \in \mathcal{E}} V_{pq}(u_p, u_q)$$

- e.g. truncated absolute difference (piecewise smooth)

$$V_{pq}(u_p, u_q) = \lambda_{pq} \min(|\mathbf{d}^{u_p} - \mathbf{d}^{u_q}|, T)$$

- Note: smoothness function can vary locally

# MRF Formulation of Image Registration

$$E_{\text{total}}(u) = \sum_{p \in \mathcal{G}} V_p(u_p) + \sum_{p, q \in \mathcal{E}} V_{pq}(u_p, u_q)$$

Opens the door to MRF optimization techniques



# Optimization using Linear Programming

# Primal-Dual Schema

- Say we seek an optimal solution  $x^*$  to the following integer program (this is our **primal** problem):

$$\min \mathbf{c}^T \mathbf{x}$$

$$\text{s.t. } \mathbf{Ax} = \mathbf{b}, \mathbf{x} \in \mathbb{N}$$

(NP-hard problem)

- To find an approximate solution, we first **relax** the integrality constraints to get a primal & a dual linear program:

$$\text{primal LP: } \min \mathbf{c}^T \mathbf{x}$$

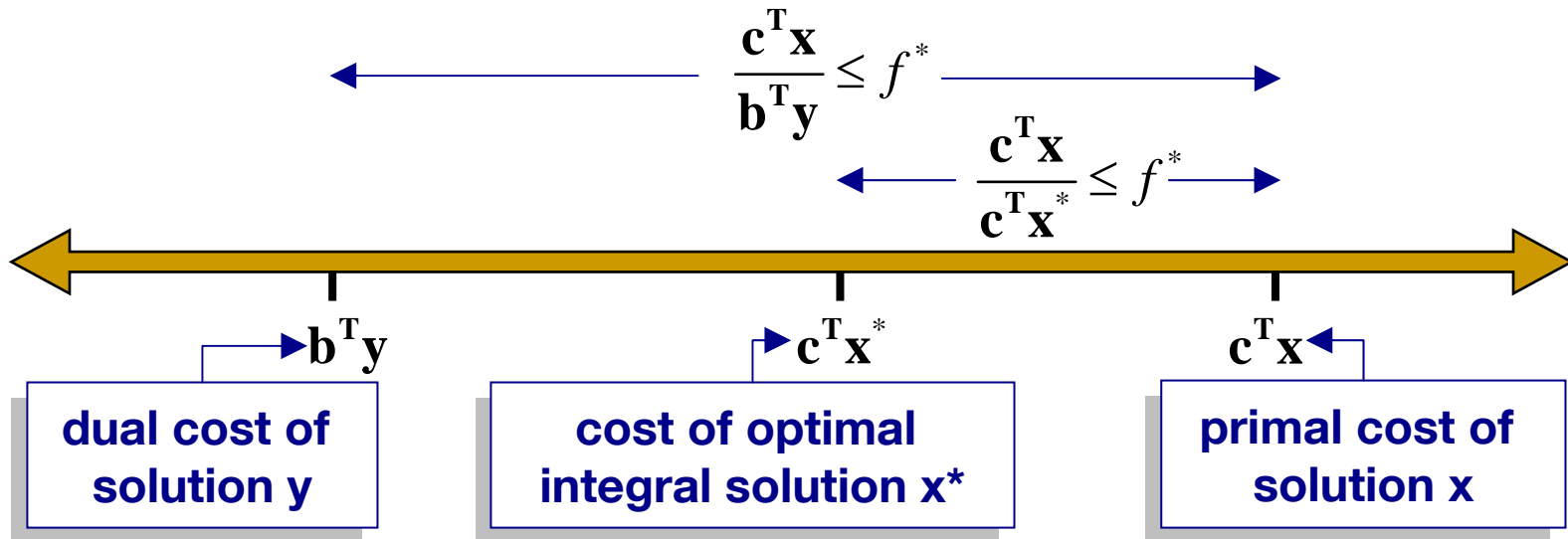
$$\text{s.t. } \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$$

$$\text{dual LP: } \max \mathbf{b}^T \mathbf{y}$$

$$\text{s.t. } \mathbf{A}^T \mathbf{y} \leq \mathbf{c}$$

# Primal-Dual Schema

- Goal: find integral-primal solution  $x$ , feasible dual solution  $y$  such that their primal-dual costs are “close enough”, e.g.,



Then  $x$  is an  $f^*$ -approximation to optimal solution  $x^*$



## Primal-Dual Schema for MRFs

$$\min \left[ \sum_{p \in G} \sum_{a \in L} V_p(a) x_{p,a} + \sum_{pq \in E} \sum_{a,b \in L} V_{pq}(a,b) x_{pq,ab} \right]$$

$$\text{s.t. } \sum_{a \in L} x_{p,a} = 1 \quad \leftarrow \text{(only one label assigned per vertex)}$$

$$\sum_{a \in L} x_{pq,ab} = x_{q,b}$$

$$\sum_{b \in L} x_{pq,ab} = x_{p,a}$$

← enforce consistency between variables  $x_{p,a}$ ,  $x_{q,b}$  and variable  $x_{pq,ab}$

$$x_{p,a} \geq 0, \quad x_{pq,ab} \geq 0$$

Binary variables

$x_{p,a}=1 \iff$  label  $a$  is assigned to node  $p$

$x_{pq,ab}=1 \iff$  labels  $a, b$  are assigned to nodes  $p, q$

# Primal-Dual Schema for MRFs

- During the PD schema for MRFs, it turns out that



- Resulting flows tell us how to update both:
    - the dual variables, as well as
    - the primal variables
- ← for each iteration of primal-dual schema

**Fast-PD**

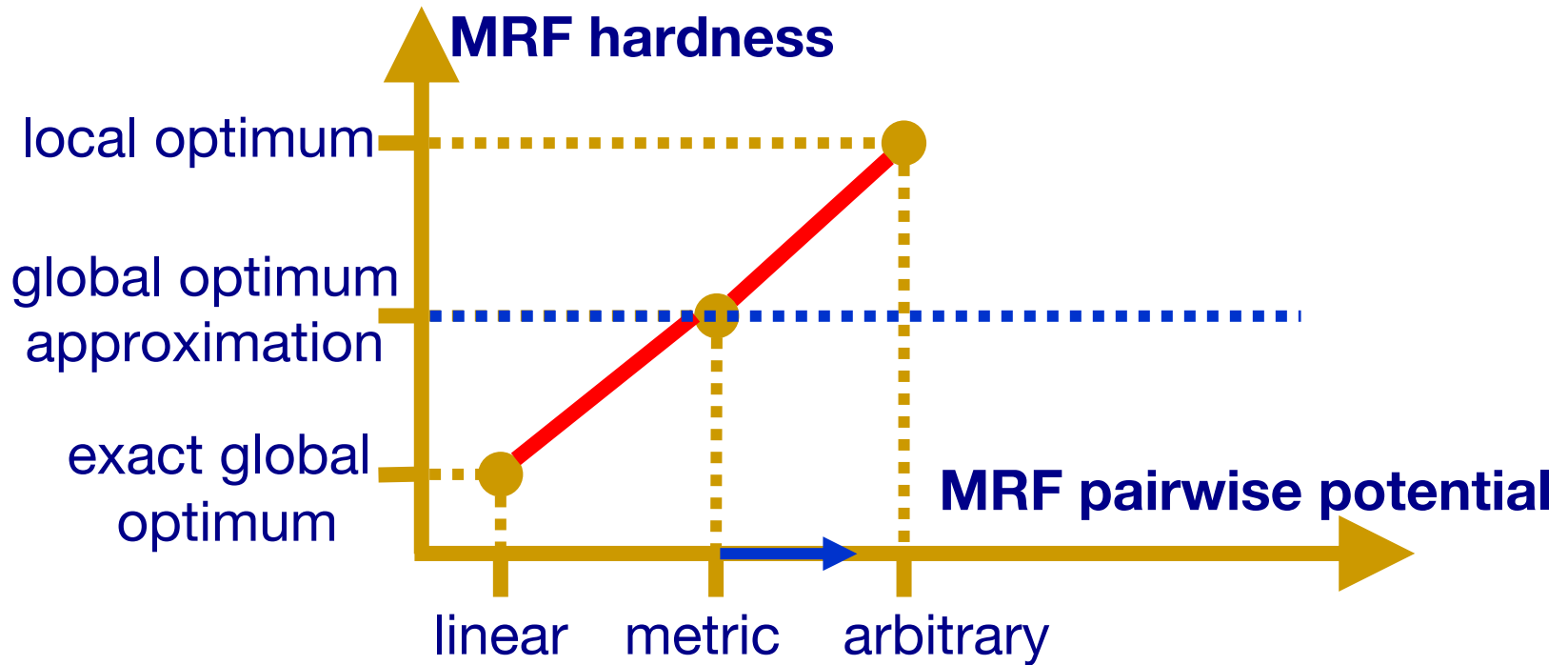
# Fast-PD

- MRF optimization method based on duality theory of Linear Programming (the Primal-Dual schema)
  - Can handle a very wide class of MRFs
  - Can guarantee approximately optimal solutions (worst-case theoretical guarantees)
  - Can provide tight certificates of optimality per-instance (per-instance guarantees)
  - Provides significant speed-up for static and dynamic MRFs

Komodakis, N., Tziritas, G., Paragios, N.

**Fast, Approximately Optimal Solutions for Single and Dynamic MRFs.**  
*Computer Vision and Pattern Recognition 2007*

# MRF Hardness



## Fast-PD



# Registration Algorithm

```
for i=1:no_iterations  
    1: setup_label_sets()  
    2: precompute_single_potential_matrix()  
    3: precompute_pairwise_potential_matrix()  
    4: compute_discrete_labeling()  
    5: update_deformation()  
end
```

Gaussian image pyramids & multi-level deformation grids are used for a hierarchical registration approach

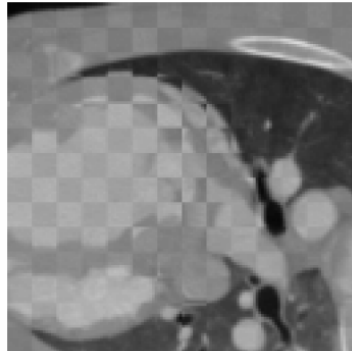
# Results & Conclusions

# Comparison

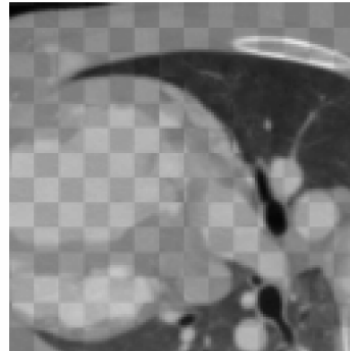
Schnabel01

Our Method

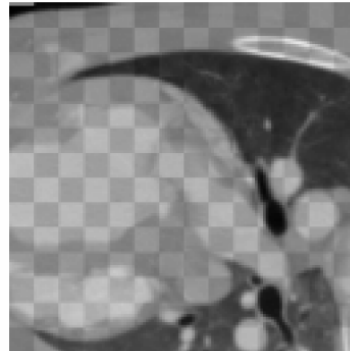
Our Method



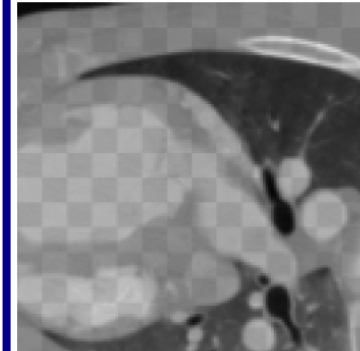
(a)



(b)



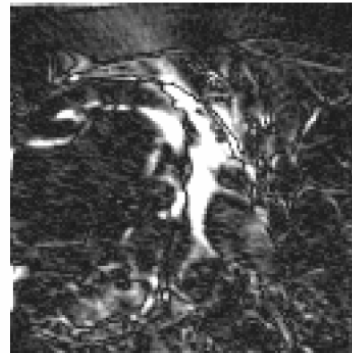
(c)



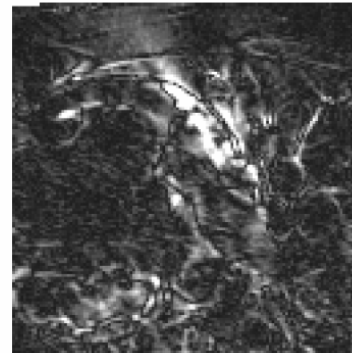
(d)



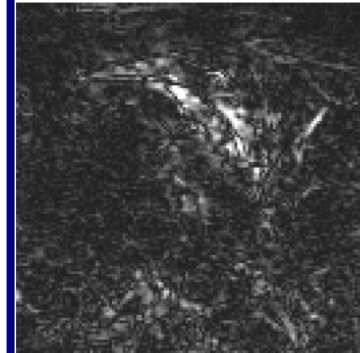
(e)



(f)



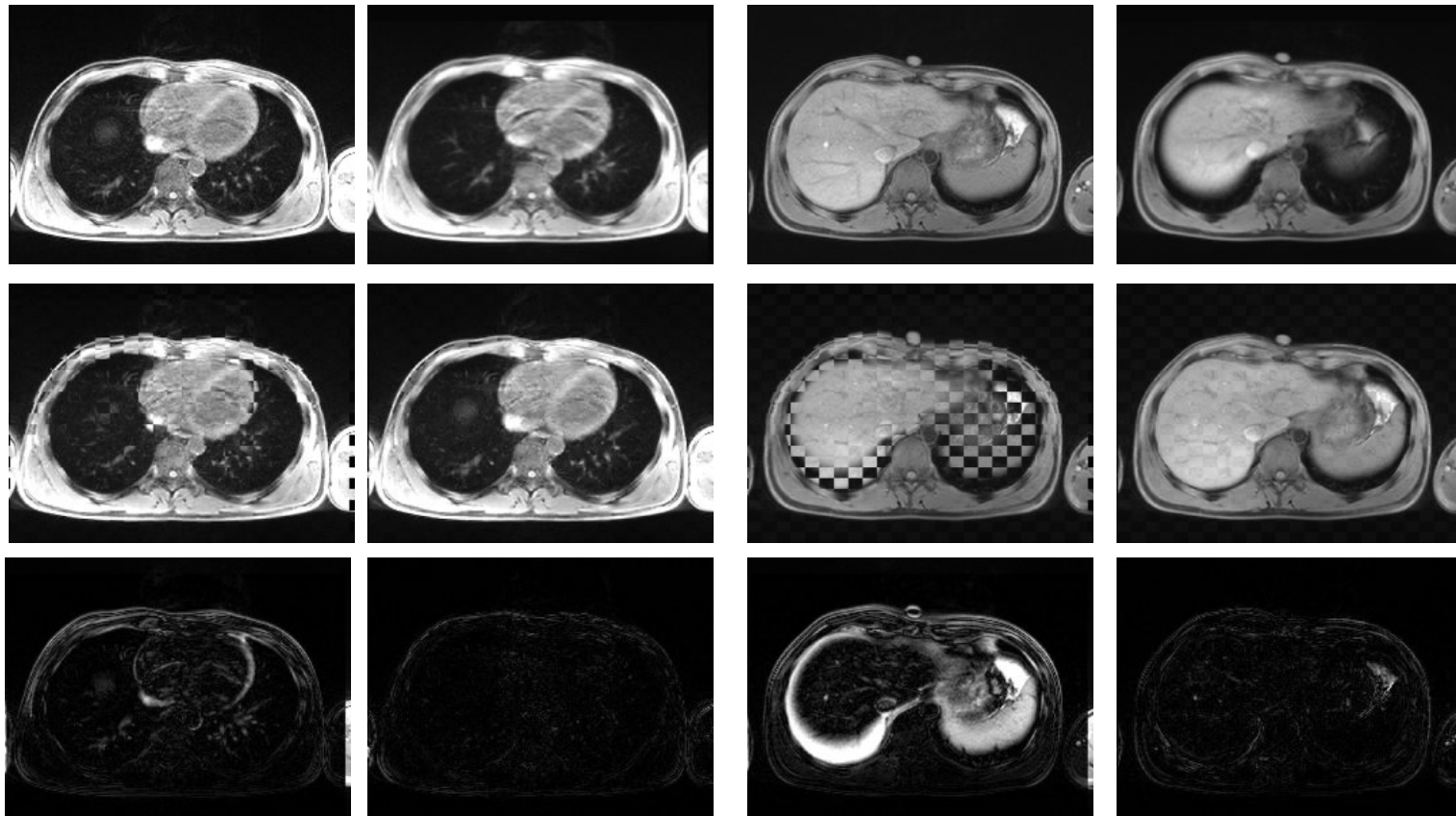
(g)



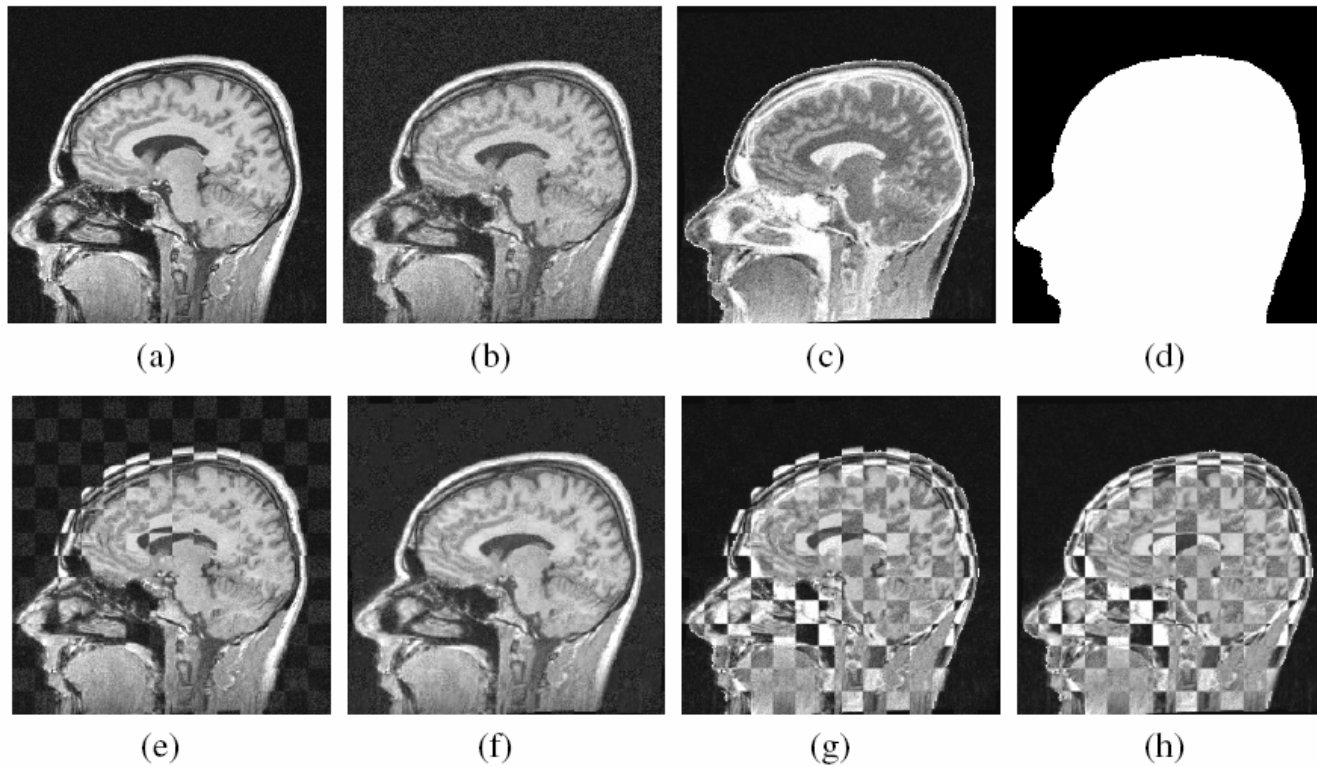
(h)

|              |                     |           |         |         |
|--------------|---------------------|-----------|---------|---------|
| SSD error    | 12278               | 3402      | 3180    | 1233    |
| Running time | Before registration | > 2 hours | < 2 min | < 2 min |

# Visual Results



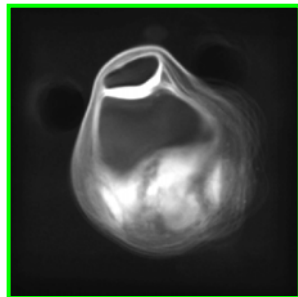
# Validation of Distance Measures



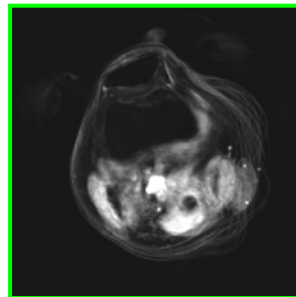
# Validation of Distance Measures

| Metric               | AE Mean | AE Median | AE Std | MOD Mean | MOD Median | MOD Std |
|----------------------|---------|-----------|--------|----------|------------|---------|
| SSD                  | 2.290   | 1.143     | 2.854  | 0.242    | 0.139      | 0.297   |
| SADG $\gamma = 1.0$  | 1.957   | 1.077     | 2.350  | 0.227    | 0.136      | 0.349   |
| SAD                  | 1.220   | 0.675     | 1.653  | 0.123    | 0.071      | 0.194   |
| SADG $\gamma = 0.75$ | 1.129   | 0.709     | 1.313  | 0.122    | 0.082      | 0.169   |
| SADG $\gamma = 0.25$ | 1.046   | 0.603     | 1.307  | 0.104    | 0.067      | 0.142   |
| SADG $\gamma = 0.5$  | 1.036   | 0.589     | 1.292  | 0.111    | 0.066      | 0.159   |
| NMI                  | 0.999   | 0.629     | 1.060  | 0.099    | 0.079      | 0.080   |
| CR                   | 0.927   | 0.536     | 1.116  | 0.089    | 0.068      | 0.092   |
| NCC                  | 0.765   | 0.402     | 1.082  | 0.070    | 0.047      | 0.076   |

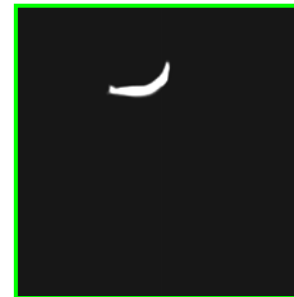
# Results (MICCAI 2007)



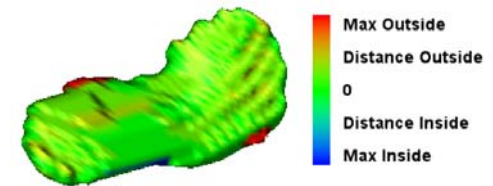
Mean Intensity Image



Variance Intensity Image



Segmentation



$$\rho_{\text{Atlas}}(I_{\mu}, I_{\sigma^2}, I_{\text{new}}) = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} \frac{(I_{\mu}(\mathbf{x}) - I_{\text{new}}(\mathbf{x}))^2}{2I_{\sigma^2}(\mathbf{x})}$$

| Method                  | DSC         | Sensitivity | Specificity | Interaction | Cartilage             |
|-------------------------|-------------|-------------|-------------|-------------|-----------------------|
| <i>Grau et al.</i>      | 0.90 (0.01) | 90.03 %     | 99.87 %     | 5-10 min    | Tibia, Femur, Patella |
| <i>Dam et al.</i>       | 0.92 (n/a)  | 93.00 %     | 99.99 %     | Max 10 min  | Tibia, Femur          |
| <i>Cheong et al.</i>    | 0.64 (0.15) | 74.00 %     | n/a         | 0           | Medial Tibia          |
| <i>Cheong et al.</i>    | 0.72 (0.09) | 79.00 %     | n/a         | 0           | Lateral Tibia         |
| <i>Folkesson et al.</i> | 0.80 (0.03) | 90.01 %     | 99.80 %     | 0           | Tibia, Femur          |
| <i>Our Approach</i>     | 0.83 (0.06) | 93.77 %     | 99.94 %     | 0           | Patella               |

# Conclusions

- What is provided by our registration framework?
  - Independent from the choice of the transformation type?  
Grid-based but independent from weighting function.
  - Independent from the choice of the distance/similarity measure?  
Yes. We are gradient-free.
  - Guarantee of a globally optimal solution?  
Quasi-yes (from an optimization point of view).
  - Reasonable computational complexity?  
Yes (256x256 in 2 seconds, 256x192x64 in 1 minute).



# Latest Work & Future Directions

- On-the-fly estimation of locally varying deformation spaces (done!)
- Incorporate global registration & different deformation models
- Introducing domain knowledge such as priors on the deformation space
- Belief propagation networks
- GPU implementation

