Medical Image Segmentation using MRF-based optimization

12th July, 2013
Il Dong Yun
Hankuk University of Foreign Studies
Contents

• Introduction
  • Medical Image segmentation
  • Mathematical Model for Image Segmentation

• Main Idea

• Overview of the algorithm

• Experimental Result

• Conclusion
Introduction: medical image segmentation

• Necessity of medical image segmentation
  • Organ segmentation is crucial element in monitoring and understanding of the progress of disease.
  • Segmentation has been mostly done totally by hand, or semi-automatic segmentation methods. But it needs experts effort & much time.
Introduction: medical image segmentation

• Difficulty of automatic segmentation
  • Inhomogeneous intensity according to modality
  • Low tissue contrast
  • Deformable and irregularly shape of organs
Introduction: image segmentation

coherent foreground
Introduction: image segmentation
Introduction: Pixel-wise segmentation

\[ \log P(z_i | x_i = 'b') \]

\[ \log P(z_i | x_i = 'f') \]

Maximum likelihood estimate:
More difficult image

coherent foreground
Need for spatial priors
Mathematical Model for Image Segmentation

- 1st order Markov Random Field (MRF)
  - Geman & Geman 1984; Besag 1974, 1986

Markov property:

\[
p(x_i \mid x \setminus \{x_i\}) = p(x_i \mid \mathcal{N}_i)
\]

where \( x = \{x_1, \ldots, x_N\} \)
Ising Model

- Binary variables: $x_i \in \{0, 1\}$

- Joint probability distribution:
  \[ p(x) = \prod_{(i,j) \in N} G_{i,j}(x_i, x_j) \]

  where
  \[ G_{i,j}(x_i, x_j) = 1 - K|x_i - x_j| \]

  and $0 < K < 1$
Simulation (MCMC sampling)
2D Hidden MRF model

\[ p(z, x) = \prod_{i} F_i(z_i, x_i) \prod_{(i,j) \in N} G_{i,j}(x_i, x_j) \]

- Probabilistic approach

\[
p(x \mid z) = \frac{p(x, z)}{p(z)} = \frac{p(z \mid x)p(x)}{p(z)} \propto p(z \mid x)p(x)
\]

- Maximum a Posteriori (MAP) solution: \( \hat{x} = \arg\max_x p(x \mid z) \)
Energy Minimization

- MAP inference can be formulated by equivalent energy minimization problem

MAP inference:
\[ \hat{x} = \arg\max_x p(x \mid z) = \arg\max_x \left[ \prod_i F_i(z_i, x_i) \prod_{(i,j) \in N} G_{i,j}(x_i, x_j) \right] \]

\[ p(x \mid z) \propto e^{-E(x)} \]

Energy minimization:
\[ \hat{x} = \arg\min_x E(x) = \arg\min_x \left[ -\sum_i \log F_i(z_i, x_i) - \sum_{(i,j) \in N} \log G_{i,j}(x_i, x_j) \right] \]

- **Unary term**
- **Pairwise term**
Discrete Energy Minimization

- Energy function:
  \[ E(x) = \sum \theta_i(x_i) + \sum_{(i,j) \in N} \theta_{ij}(x_i, x_j) \]

- (Finite) Labels: 1, 2, .., \( L \)

- Complexity of naïve approach: \( O(L^{|V|}) \)
Minimization Algorithms

- Iterated Conditional Modes (ICM)
- Simulated Annealing

- Dynamic Programming (DP)
- Belief Propagation (BP)
- Tree-Reweighted (TRW)

- Graph Cuts (GC)
- Branch & Bound

- Relaxation methods

Classical Move making algorithms

Message passing

Combinatorial Algorithms

Convex Optimization (Linear Programming, ...)

Hankuk Univ. of Foreign Studies
Graph Cut algorithm

Construct a graph such that:

1. Any st-cut corresponds to an assignment of $x$
2. The cost of the cut is equal to the energy of $x$: $E(x)$

[Hammer, 1965] [Kolmogorov and Zabih, 2002]
Main idea

• Motivation
  • Same organs have common global shape, but each subject has their own deformations locally.
  • We find the deformations in local view, and enforce the shape smoothness in global view.

• Main Idea
  • Based on prior knowledge of training set, adaptive local prior and global shape prior are considered.
  • Both the local prior and the global prior are incorporated into a single hierarchical structure MRF.
Overview of algorithm

Input Image
Overview of algorithm

Input Image

Divide the whole object into local interest regions
Overview of algorithm

Extract the similar patches in each regions from the training set (Reference patches)

Test patch

1  2  3  

. . . . .

N : # of training set

1  2  3  

. . . . .

N
Overview of algorithm

Local characteristics
(difference between FG/BG histogram)

\[ p_f(v) = w(v)p_s(v) + (1 - w(v))p_a(v). \]

Adaptive local prior
Shape prior (template)
Appearance prior (intensity histogram)

Make adaptive prior based on the shape and appearance information

Test patch

1  2  3  N : # of training set

Extract the similar patches in each regions from the training set (Reference patches)
Overview of algorithm

For each reference patch, MRF graph is constructed.

\[ \phi_l(x_v) = -\log(p_f(v)) \]

\[ E(X) = \sum_{v \in V} \phi(x_v) + \lambda \sum_{u,v \in \xi} \psi(x_u,x_v) \]

Adaptive prior  Smoothness  Graph cut Optimization

Construct local level MRF based on adaptive prior
Overview of algorithm

Construct local level MRF based on adaptive prior

Segmentation candidates in red patch

Segmentation candidates in orange patch

Obtain segmentation candidates in each local region
Overview of algorithm

How to select optimal candidates in each local region?

1 2 3 . . . . . N
Segmentation candidates in red patch

1 2 3 . . . . . N
Segmentation candidates in orange patch

Construct local level MRF based on adaptive prior
Obtain segmentation candidates in each local region
Overview of algorithm

Node = local region
Edge = relation of local regions

Construct global level MRF based on segmentation candidates

Node
Edges

Construct local level MRF based on adaptive prior

Segmentation candidates in red patch

Segmentation candidates in orange patch

Obtain segmentation candidates in each local region
Overview of algorithm

Construct global level MRF based on segmentation candidates

\[ E(X) = \sum_{v \in V} \phi(x_v) + \lambda \sum_{u,v \in \xi} \psi(x_u, x_v) \]

\[ \phi_g(x_i = j) = -\log \left( \frac{1}{|\xi(x_i^j)|} \sum_{v \in \xi(x_i^j)} \left| \frac{\partial I(v)}{I_{\text{max}}} \right| \right) \]

Boundary is likely to have high gradient

\[ \psi_g(x_i = j, x_{i'} = j') = -\log \left( \frac{2 \sum_{v \in \xi(x_i^j) \cap 

Overlapping regions are likely to be similar and smoothly connected,

Construct local level MRF based on adaptive prior

TRW-s Optimization
Overview of algorithm

Node = local region
Edge = relation of local regions

Construct global level MRF based on segmentation candidates

Final Result

Construct local level MRF based on adaptive prior

Segmentation candidates in red patch

Segmentation candidates in orange patch

Obtain segmentation candidates in each local region
Experiment Result

• Measurement

  • Dice similarity coefficient: \( \text{dsc}(S, R) = \frac{2 \cdot |S \cap R|}{|S| + |R|} \)

  \( S \) : Segmentation result
  \( R \) : Ground truth

• Quantitative Result

<table>
<thead>
<tr>
<th>Data</th>
<th>Organ</th>
<th>Initial</th>
<th>GP</th>
<th>LP</th>
<th>HMRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>knee MR</td>
<td>femur</td>
<td>0.906 (0.024)</td>
<td>0.946 (0.013)</td>
<td>0.966 (0.010)</td>
<td>0.970 (0.007)</td>
</tr>
<tr>
<td></td>
<td>tibia</td>
<td>0.886 (0.046)</td>
<td>0.926 (0.049)</td>
<td>0.968 (0.021)</td>
<td>0.973 (0.015)</td>
</tr>
<tr>
<td></td>
<td>cartilage (F)</td>
<td>-</td>
<td>-</td>
<td>0.879 (0.026)</td>
<td>0.887 (0.015)</td>
</tr>
<tr>
<td></td>
<td>cartilage (T)</td>
<td>-</td>
<td>-</td>
<td>0.858 (0.032)</td>
<td>0.878 (0.018)</td>
</tr>
<tr>
<td>brain MR</td>
<td>hippocampus (L)</td>
<td>0.642 (0.086)</td>
<td>0.832 (0.026)</td>
<td>0.854 (0.022)</td>
<td>0.868 (0.013)</td>
</tr>
<tr>
<td></td>
<td>hippocampus (R)</td>
<td>0.660 (0.074)</td>
<td>0.829 (0.027)</td>
<td>0.856 (0.018)</td>
<td>0.867 (0.015)</td>
</tr>
</tbody>
</table>

\( GP \) : Global prior based method
\( LP \) : Local prior based method
\( HMRF \) : Proposed method
Experiment Result

- Bone segmentation in knee MR image
Experiment Result

- Cartilage segmentation in knee MR image
Experiment Result

- Hippocampus segmentation in brain MR images
Experiment Result

- Pathological deformities (large shape variation)
Conclusion

• Automatic image segmentation algorithm for medical image
• Use a prior information efficiently
• Computational burden still high: Open MP, GPU implementation
• Interactive still needed for experts proof

Publications

• Sang Hyun Park, Soochahn Lee, Il Dong Yun, and Sang Uk Lee, Interactive 3D Segmentation Method based on Uncertain Local Region Updating in Hierarchical MRF Graph, SPIE MI 2013.
Application

• **Liver MR registration for tough cases**
  
  • Dense registration of Liver MR images with different phases (e.g., arterial / portal) containing tough cases
    
    • e.g., scattered small cysts, meta hiding in branching vessel
  
  • Up to **pixel-level accuracy** is suggested

![reference (arterial) and target (portal)](image)
Results

top left: reference (arterial)
top middle: target (portal)
top right: deformed
bottom left: initial difference
bottom right: final difference
Results

top left: reference (arterial)
top middle: target (portal)
top right: deformed
bottom left: initial difference
bottom right: final difference
Higher-Order MRF with Dense Local Descriptor

- Dense Local Descriptor
- Higher-Order MRF
- Optimization Strategy
Introduction

• Hybrid (Feature + Intensity-Based) Methods
  - [Liu’08, Ou’09, Sotiras’10, ...]

• Pros
  - Consistent relationship between intensity distributions is not needed
  - Can produce sophisticated displacements

• Cons
  - Limited accuracy
Previous Approaches

• Feature-based methods work badly on feature-less regions
• Intensity-based methods will fail on general multi-modal data
  - A hybrid approach of feature- and intensity-based methods is required

• Recently, dense local descriptors have been applied on MRF-based registration models
  - C. Liu, J. Yuen, A. Torralba, **SIFT Flow: Dense Correspondence across Scenes and Its Applications**, *TPAMI*, 2011.
    • SIFT descriptor (2D)
    • Gabor attribute (3D)
Previous Approaches

• These two approaches are using MRF models with first-order smoothness priors

  \[ E(x|\theta) = \sum_{s \in \mathcal{V}} \theta_s(x_s) + \sum_{(s,t) \in \mathcal{E}} \theta_{st}(x_s, x_t) \]

  - SIFT flow: shows blocky displacements (fronto-parallel bias in stereo); can not make sophisticated displacements required by medical applications
  - DRAMMS: first-order smoothness prior cannot represent thin plate spline energy (involving second order derivatives)

• In 3D, gradient histogram-based (SIFT descriptor) similarity measures are expected to be stronger than edge-based (Gabor attribute) similarity measures
  - There's no paper discovered this presumption
Proposed Method: Dense Local Descriptor

- Dense Local Descriptor
  - We use densely sampled standard SIFT descriptor
    - 4x4 grid bins, 4x4 windows for each bin, 8 gradient orientation histograms
    - $4 \times 4 \times 8 = 128$ dimensional vector
  - Visualizing dense SIFT descriptors
Proposed Method: Higher-Order MRF

- MRF Energy Model

\[ E(x|\theta) = \sum_{s \in V} \theta_s(x_s) + \sum_{(s,t) \in \mathcal{F}_P} \theta_{st}(x_s, x_t) + \sum_{(s,t,u) \in \mathcal{F}_H} \theta_{stu}(x_s, x_t, x_u) \]

- Mixed-order prior: first- and second-order smoothness prior

\[ \theta_{st}(x_s, x_t) = \lambda_{st} \min (\|d(x_s) - d(x_t)\|_1, T_{st}) \]

\[ \theta_{stu}(x_s, x_t, x_u) = \lambda_{stu} \min (\|d(x_s) - 2d(x_t) + d(x_u)\|_1, T_{stu}) \]
Mixed-Order Smoothness Prior

• Geman and Reynolds introduce a computational method based on simulated annealing in which the outcome of the first-order model is used as a starting point for the second-order model, etc.

• With the model parameters suitably chosen, planar and quadric structures may then be recovered with the higher-order models while existing jump discontinuities are maintained.

Experiments

- Single- and multi-modal brain MRIs
Experiments

• Face images with illumination change

Conclusion & Future Works

- The proposed method shows promising results on various registration experiments
  - Producing smoother displacements is enabled by higher-order smoothness priors
  - Appearance variations covered by dense local descriptors

- Efficient optimization is implemented
  - TRW message passing is parallel algorithm
  - Fast optimization using GPU (CUDA)
Conclusion & Future Works

• Performance comparison of dense local descriptor for registration will be performed in the future work

• Optimal methods for mixing-orders need to be analyzed

• For brain MRI registrations, we have several modality
  - T1, T2, PD, FL, FA, TR, ...
  - Multi-dimensional descriptors extracted from various modalities will be increase distinctiveness of descriptors
Segmentation based on prior knowledge

Optimization of local shape and appearance probabilities for segmentation of knee cartilage in 3-D MR images, CVIU 2011