Multi-Organ Segmentation using deeds, Self-Similarity Context and Joint Fusion

Mattias P. Heinrich
Institute of Medical Informatics, Universität zu Lübeck

Multi-Atlas Labeling Beyond the Cranial Vault
Workshop and Challenge @ MICCAI 2015
Status Quo in Multi-Atlas Segmentation

Most publications based on MRI brain images
(many available public databases)

advances in recent years for
deformable registration and multi-atlas fusion
→ brain parcellation reaches performance ceiling

segmentation of abdomen and cervix is a very different challenge
approaches without registration play in important role
(often only evaluated on private datasets)

Are we using the right tools?
Multi-Atlas Segmentation of Abdominal Scans

calculate nonlinear transform between atlases & target space
transform manual segmentations towards target scan → label fusion

deformations can be extremely large - contrast and noise are challenges

assumption: optimising nonrigid registration is most important concern
Standard Registrations of NiftyReg

NiftyReg\(^1\) has excellent performance on brain MRI, and also on large lung motion EMPIRE10

→ yet performance for abdomen before label fusion is underwhelming

\[ D \approx 30\% \text{ (avg. single-atlas)} \]

potential issues: 1) local minima of gradient-based optimisation and 2) low performance of mutual information metric

**our approach:** 1) use discrete optimisation package **deeds** and 2) **self-similarity context** as similarity metric

3) employ widely used joint label fusion afterwards

\(^1\)Modat et al.: “Fast free-form deformation using graphics processing units” Comp Meth Prog Bio ‘10
**dense displacement sampling (deeds²)**

divide fixed image in non-overlapping blocks

→ calculate block-matching over ‘dense’ sampling region

*infer regularity* over joint displacement probabilities (of neighbouring blocks)
using MRF optimisation on minimum-spanning-tree

MRF avoids local minima - has high computational efficiency (<5 secs for inference)

²Heinrich et al.: “MRF-based deformable registration of lung CT” IEEE Trans Medical Imaging ‘13
dense displacement sampling (deeds\(^2\))

pre-processing: resampling to 2.2\(^3\) mm\(^3\) cropping 180x140x190

affine registration with block-matching and trimmed least squares

multi-scale approach with decreasing block-size and sampling region

initial level uses 2197 displacements and 66 mm capture range

local mutual information (over patch) with 8 bins, Parzen window
recalculate histogram for every search position

→ 90 secs. per 3D registration (compared to 5 sec. for optimisation)

\(^2\)Heinrich et al.: “MRF-based deformable registration of lung CT” IEEE Trans Medical Imaging ‘13
dense displacement sampling (deeds)

- spleen
- r kidney
- l kidney
- gallbladder
- esophagus
- liver
- stomach
- aorta
- vena cava
- portal vein
- pancreas
- r adrenal gland
- l adrenal gland
- average

NiftyReg: 0.239, 0.299, 0.394
Our affine: 0.239, 0.299, 0.394
Metric: Self-similarity Context

**self-similarity context** as similarity metric (extension from MIND)

self-similarity $D_p$ is measured by patch-wise SSD (sum of squared differences)

variance estimate $V$ makes descriptor insensitive to locally varying contrast

$$\text{MIND}(I, \mathbf{x}, \mathbf{r}) = \frac{1}{n} \exp \left( -\frac{D_p(I, \mathbf{x}, \mathbf{x} + \mathbf{r})}{V(I, \mathbf{x})} \right) \quad \mathbf{r} \in \mathcal{R}$$

**Pairwise** self-similarities in neighbourhood

**Context** → robust to noise, sensitive to curvature

**Quantise** descriptor into binary string

$(0.8,0.2,0.6,0.0) \rightarrow 0111,0001,0011,0000$

Distance between descriptors is **Hamming weight**

reduces computation time for similarity metric to **15 secs.**

---

3 Heinrich et al.: “Towards realtime multimodal fusion using self-similarity context” MICCAI ‘13
Self-similarity context

- NiftyReg
- our affine
- deeds MI
- deeds SSC

Values:
- spleen: 0.456
- r kidney: 0.394
- l kidney: 0.299
- gallbladder: 0.239
- esophagus: 0.8
- liver: 0.6
- stomach: 0.6
- aorta: 0.4
- vena cava: 0.4
- portal vein: 0.2
- pancreas: 0.2
- r adrenal gland: 0.2
- l adrenal gland: 0.2
average: 0.4

Organ labels:
- spleen
- r kidney
- l kidney
- gallbladder
- esophagus
- liver
- stomach
- aorta
- vena cava
- portal vein
- pancreas
- r adrenal gland
- l adrenal gland
Multi-Atlas Label Fusion

**majority vote:** same weight for every atlas

**locally weighted fusion:** with non-local search

**joint fusion**\(^4\): estimation of dependencies for local atlas weights

**corrective learning**\(^5\): fixes segmentation errors

*as comparison:* joint classification and regression forests\(^6\)

---

\(^4\)Wang et al.: “Multi-atlas segmentation with joint label fusion” IEEE PAMI ’13

\(^5\)Wang & Yushkevich: "Multi-Atlas Segmentation with Joint Label Fusion and Corrective Learning"
Frontiers in Neuroinformatics ’13

\(^6\)Glocker et al.: "Joint classification-regression forests for spatially structured multi-object .." ECCV ’12
Multi-Atlas Label Fusion

Joint CR Forest
- deedsSSC Single
- deedsSSC Majority
- NiftyReg Majority

- Multi-Atlas Label Fusion

Average: 0.79
Right kidney: 0.789
Liver: 0.95
Stomach: 0.85
Spleen: 0.92
Left kidney: 0.91
Gallbladder: 0.67
Esophagus: 0.66
Aorta: 0.80
Vena cava: 0.86
Portal vein: 0.83
Pancreas: 0.75
Right adrenal gland: 0.62
Left adrenal gland: 0.587

0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Conclusion and Outlook

Discrete optimisation helps overcome local minima

Self-similarity context is slightly more accurate and faster than MI

Very fast (45 secs. per atlas) and accurate

www.mpheinrich.de/software.html

Majority Voting alone is not sufficient

→ Joint label fusion (D=0.79, d=2.26mm) but very slow

Patch-MI workshop 2015 (faster joint label fusion)

Cervix segmentation even more challenging (D=0.63)

---

7Heinrich et al.: "Multi-Atlas Segmentation using Patch-Based Joint Label Fusion with Non-Negative Least Squares Regression" MICCAI Patch-MI ’15
Multi-Organ Segmentation using deeds, Self-Similarity Context and Joint Fusion
Non-negative Least Squares Regression

\[ \arg\min_c \| A c - b \|_2^2 \]

**additional non-negativity constraint**

\[ \text{subject to } c_i \geq 0 \ \forall \ i = \{1, 2, \ldots, n\} \]

implicitly this formulation promotes sparsity\(^3\) (c.f. LASSO \(L_1\) regularisation)

**advantages:** fast solvers are available, no weighting parameter required

visual example:

labels NNLS  
best atlas patches  
worst atlas patches  

\(^3\)Slawski and Hein: “Sparse recovery by thresholded non-negative least squares” NIPS ‘11
Comparison to state-of-the-art

Optimisation of regularisation parameter $\epsilon$ of least-squares
Weights are applied to the label patch (not only centre-voxel)

Significant improvements (~3% Dice) over NCC, LSQ, LASSO

Small improvements over Joint Fusion and 15-fold speed-up

Zhang et al.: “Sparse patch-based label fusion for multi-atlas segmentation” MBIA ’12