How many templates does it take for a good segmentation ?

Error analysis in multiatlas segmentation as a function of database size

Suyash P. Awate, Peihong Zhu, Ross T. Whitaker

Scientific Computing & Imaging (SCI) Institute, School of Computing, University of Utah

The authors gratefully acknowledge the support of this work through the National Alliance for Medical Image Computing (NAMIC) and the NIH/NCRR Center for Integrative Biomedical Computing (CIBC) grant P41-RR12553

1. Background : Challenges to Multiatlas Segmentation

- o Registration is imperfect
  - Low contrast-to-noise ratio for structure of interest
  - No homology with respect to surrounding structures
  - No global optimization
- o Stochastic relationship between structure's geometry & surrounding geometry
- o Quantify difficulty of segmentation
  - Segmentation error ("goodness measure") as a function of database size ("how many templates")

2. Contributions

- o Quantify difficulty, or characteristics, of a specific multiatlas segmentation problem (i.e. structure, modality, registration, label fusion, etc.)
  - Formulated multiatlas segmentation as nonparametric regression
  - Estimation's convergence behavior has a parametric form
  - Parameters characterize fundamental properties of the regression
  - Method for estimating these parameters for a given segmentation problem using a (small) multiatlas database
- o Predict segmentation error for large databases using small databases

3. Multiatlas Segmentation as Nonparametric Regression

- o For each target, transform database to factor out diffeomorphism (limit norm)
- o Independent variable = F = (deformed) biomedical image
- o Dependent variable = S = (deformed) segmentation image
- o Database = $a^M \equiv \{(f_m, s_m)\}_{m=1,\cdots,M}$
- o Target image = f_0
- o Regression function = $r(f) \equiv E_{P(S|f)}[S]$
  - Conditional expectation is optimal under mean squared error risk
- o Regression estimator $\hat{r}(F, A^M)$
- o Mean squared error (MSE) as a function of database size 'M'
  $MSE(M) \equiv E_{P(F,S,A^M)}[\| S - \hat{r}(F, A^M) \|^2] = E_{P(F)}[MSE(M,F)]$
- o At each voxel 'v', MSE is
  $MSE(M,f)[v] \equiv Var(S[v] \mid f) + Bias^2(\hat{r}(f, A^M)[v]) + Var(\hat{r}(f, A^M)[v])$
  = variance of conditional PDF + estimator bias² + estimator variance

4. Generalized-kNN Regression Estimator

$$\hat{r}(f, A^M)[v]) \equiv \left\{ \sum_{m=1}^{M} s_m[v] w\left(\frac{g(f_m, f)}{R_k} \right) \right\} \Big/ \left\{ \sum_{m=1}^{M} w\left(\frac{g(f_m, f)}{R_k} \right) \right\}$$

- o g(.) = some distance metric on images
- o R_k = distance to k-th nearest neighbor
- o w(.) = generalized weight function

$$Bias(\hat{r}(f, A^M)[v]) \approx \phi(r(\cdot)[v], P(F), f, D) \boxed{(k/M)^{2/D}}$$

$$Var(\hat{r}(f, A^M)[v]) \approx \psi(w(\cdot), D) \boxed{Var(S[v] \mid F)(1/k)}$$

[Local properties of kNN regression estimates. SIAM J. Alg. Disc. Meth. 1981]

5. Practical Interpretation

- o Given k (# nearest neighbors) and M (database size)
  - Perform Monte-Carlo sampling of targets and databases
  - Evaluate MSE for each voxel 'v' and database-size 'M'
  - Fit parametric curve

$$MSE(M)[v] = \alpha_v + \beta_v (k/M)^{4/D_v} + \gamma(1/k) = \delta_v + \beta_v (k/M)^{4/D_v}$$

$$\alpha_v = E_{P(F)}[Var(S[v] \mid F)], \beta_v = E_{P(F)}[\phi(r(\cdot)[v], P(F), D_v)]$$

$$\gamma_v = E_{P(F)}[Var(S[v] \mid F) \psi(w(\cdot), D_v)], \delta_v = \alpha_v + \gamma_v / k$$

Parameter alpha = lowest possible MSE
Parameter delta = lowest possible MSE for chosen k, w(.)
Parameter beta = complexity of regression
Parameter D = intrinsic dimension

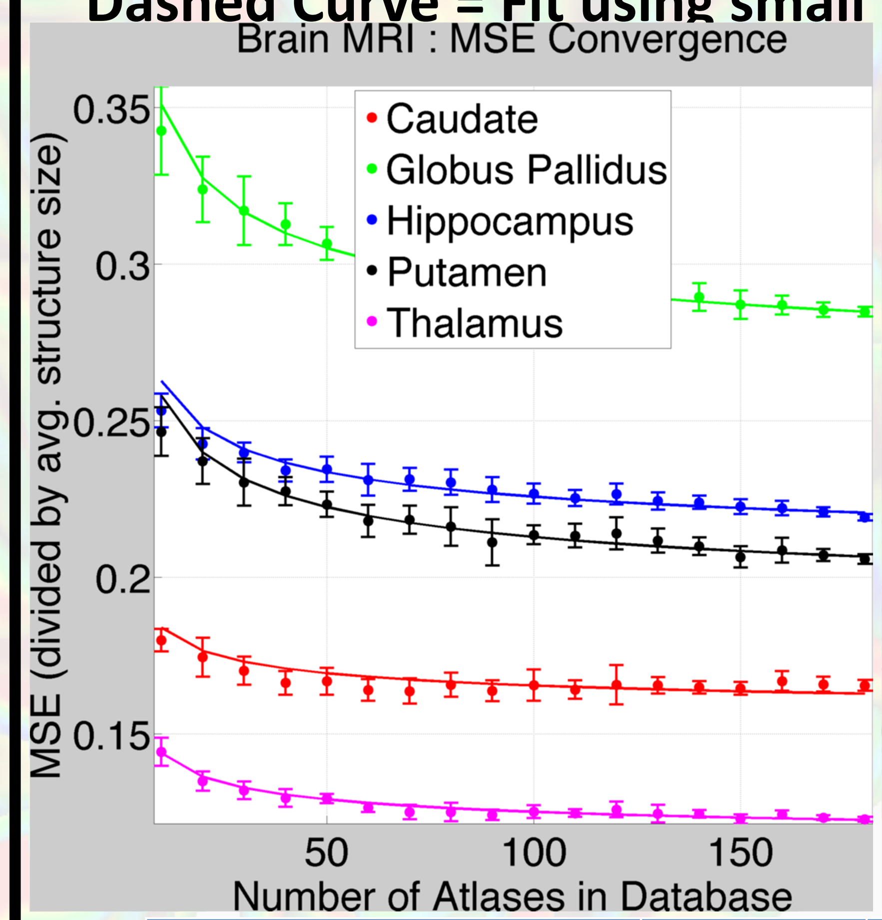
- o Extend model to entire structure by aggregating per-voxel analysis

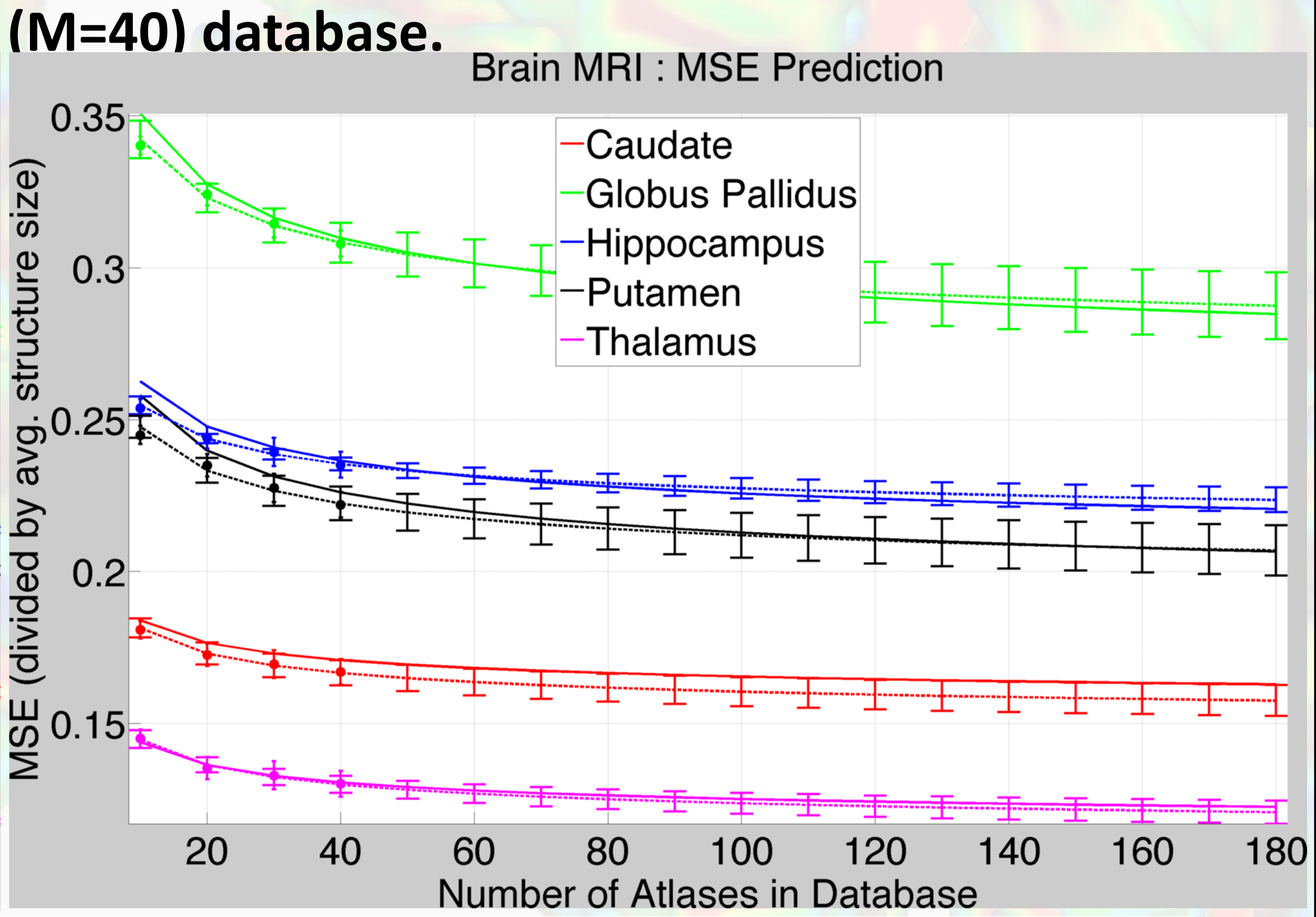
$$MSE(M) = \alpha + \beta (k/M)^{4/D} + \gamma(1/k) = \delta + \beta (k/M)^{4/D}$$

$$\alpha = \sum_{v=1}^{V} \alpha_v, \beta \approx \sum_{v=1}^{V} \beta_v, \gamma = \sum_{v=1}^{V} \gamma_v, \delta = \alpha + \gamma / k$$

6. Results

- o LEFT: Characteristics, or difficulty, of multiatlas segmentation of subcortical brain structures from T1-weighted MRI.
- o RIGHT: Prediction of error performance for large databases using small databases
- o Dot = Measured MSE.
  Solid Curve = Fit using large (M=186) database.
  Dashed Curve = Fit using small (M=40) database.





Parameters	Caudate	Globus Pallidus	Hippocampus	Putamen	Thalamus
delta = randomness	0.15	0.26	0.20	0.18	0.11
beta = complexity	0.03	0.10	0.06	0.08	0.03
D = dimension	10.1	10.0	10.0	10.0	10.0

- o Per-voxel parameters for hippocampus