Pattern Recognition (PR)

Prof. Dr.-Ing. Andreas Maier
Pattern Recognition Lab (CS 5), Friedrich-Alexander-Universität Erlangen-Nürnberg
Winter Term 2020/21

This is a printable version of the slides of the lecture

Pattern Recognition (PR)
Winter term 2020/21
Friedrich-Alexander University of Erlangen-Nuremberg.

Erlangen, January 8, 2021
Prof. Dr.-Ing. Andreas Maier

Adaptive Segmentation of MRI Data

Introduction

Magnetic Resonance Imaging (MRI) is an important acquisition technique.

It features:

- high spatial resolution
- good soft tissue contrast
- does not incorporate ionizing radiation (as computed tomography)

Several applications require the segmentation (classification) of the acquired images into tissue types.

Introduction (cont.)

Difficulties arise from:

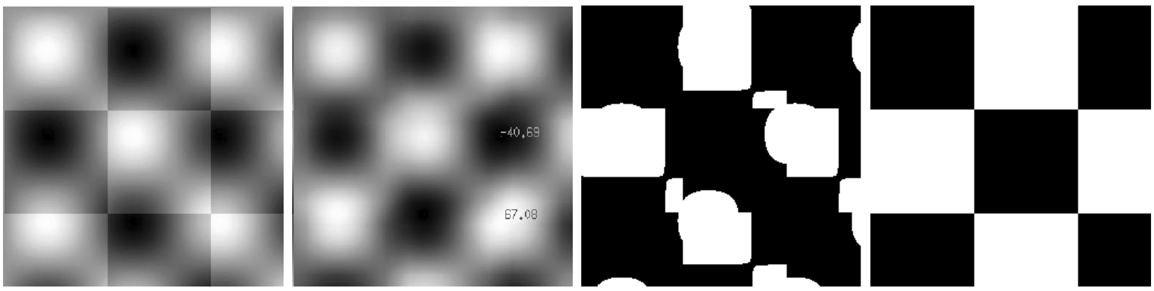
- missing intensity normalization (like Hounsfield units in CT)
- intensity inhomogeneities, also known as bias field
  (RF coils, acquisition sequences)

(a) with bias field (b) corrected

Fig.: MRI intensity inhomogeneity (Courtesy of F. Jäger)

Introduction (cont.)

Effect of the bias field on ML segmentation:



(a) (b) (c) (d)

Fig.: Synthetic image (a) overlaid with artificial bias field (b), result of ML segmentation (c), result after modeling bias field within segmentation (d) (Courtesy of W. Wells).

Introduction (cont.)

W. M. Wells et al. presented an approach to improve MR brain segmentation (1996):

- statistical approach to intensity-based segmentation of MRI
- statistical modeling of bias field (smoothness constraint)
- usage of EM algorithm for simultaneous computation of tissue classification and intensity inhomogeneity correction

Typical EM problem:

- The missing data is the tissue class assignment for each pixel.
- If the tissue was classified, the bias field could easily be computed.
- If the bias field was known, the tissue classification would be much easier.

Bias Field Model

- Let $\tilde{X}_i$ be the (unknown) intensity of the i-th voxel of the MRI data and B_i the corresponding bias field.
- The bias field is assumed to be multiplicative:

$$X_i = \tilde{X}_i \cdot B_i$$

- Using a log-transform on the intensities yields an additive bias field model:

$$Y_i = \log X_i = \log \tilde{X}_i + \beta_i \text{ , with } \beta_i = \log B_i$$

- The bias field is then:

$$\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{n-1})^T$$

with n being the number of voxels.

Bias Field Model (cont.)

- The bias field is assumed to change smoothly over the entire image domain.
- It is modeled by an n-dimensional zero mean Gaussian prior:

$$p(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}; 0, \boldsymbol{\Psi}_{\boldsymbol{\beta}})$$

Notes:

- $\boldsymbol{\Psi}_{\boldsymbol{\beta}}$ is a $n \times n$-dimensional covariance matrix
- $\boldsymbol{\Psi}_{\boldsymbol{\beta}}$ is too large to compute directly in practice
- Instead of the full covariance matrix, a banded estimate is chosen in practice.

Bayesian Approach

Variables:

Y_i	log-transformed observed intensity at i-th voxel
Γ_i	tissue class of the i-th voxel
μ_Γ	mean intensity for tissue class Γ
ψ_Γ	variance of tissue class Γ

The intensities are assumed to be scalar values, therefore: $\mu_\Gamma, \psi_\Gamma \in \mathbb{R}$

Bayesian Approach (cont.)

Assuming statistical independence of the intensities, the probability density for the entire image $\boldsymbol{Y} = (Y_0, Y_1, \dots, Y_{n-1})^T$ is:

$$p(\boldsymbol{Y}|\boldsymbol{\beta}) = \prod_i p(Y_i|\beta_i)$$

The probability of the observations is modeled as a Gaussian mixture over the tissue classes:

$$p(Y_i|\beta_i) = \sum_\Gamma p(Y_i, \Gamma|\beta_i) = \sum_\Gamma p(\Gamma)\, p(Y_i|\Gamma, \beta_i)$$

with

$$p(Y_i|\Gamma, \beta_i) = \mathcal{N}(Y_i; \mu_\Gamma + \beta_i, \psi_\Gamma)$$

Bayesian Approach (cont.)

Observations so far:

- Each tissue class is modeled with a normal distribution.
- The modeling of the observed intensity distribution yields a Gaussian mixture model.
- $p(\Gamma)$ is a stationary prior probability for the tissue class.
- The estimators for the GMM are non-linear!

Bayesian Approach (cont.)

Using Bayes rule to derive an objective function for the bias field:

$$p(\boldsymbol{\beta}|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\boldsymbol{\beta})\,p(\boldsymbol{\beta})}{p(\boldsymbol{Y})}$$

Applying the MAP principle yields an estimator for the bias field:

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \operatorname*{argmax}_{\boldsymbol{\beta}} p(\boldsymbol{\beta}|\boldsymbol{Y}) \\ &= \operatorname*{argmax}_{\boldsymbol{\beta}} \log p(\boldsymbol{\beta}|\boldsymbol{Y}) \\ &= \operatorname*{argmax}_{\boldsymbol{\beta}} \left(\log p(\boldsymbol{Y}|\boldsymbol{\beta}) + \log p(\boldsymbol{\beta})\right)\end{aligned}$$

Gradient Computation

At the optimum, the gradient w. r. t. $\boldsymbol{\beta}$ has to be zero:

$$\begin{aligned}\frac{\partial}{\partial \beta_i} \log p(\boldsymbol{\beta}|\boldsymbol{Y}) &= \frac{\partial}{\partial \beta_i}(\log p(\boldsymbol{Y}|\boldsymbol{\beta}) + \log p(\boldsymbol{\beta})) \\ &= \frac{\partial}{\partial \beta_i}\left(\sum_j \log p(Y_j|\beta_j) + \log p(\boldsymbol{\beta})\right) \\ &= \frac{\frac{\partial}{\partial \beta_i} p(Y_i|\beta_i)}{p(Y_i|\beta_i)} + \frac{\frac{\partial}{\partial \beta_i} p(\boldsymbol{\beta})}{p(\boldsymbol{\beta})} \\ &\stackrel{!}{=} 0.\end{aligned}$$

Gradient Computation (cont.)

$$\frac{\frac{\partial}{\partial \beta_i} p(Y_i|\beta_i)}{p(Y_i|\beta_i)} = \underbrace{\frac{\sum_\Gamma p(\Gamma)\, \frac{\partial}{\partial \beta_i} \mathscr{N}(Y_i; \mu_\Gamma + \beta_i, \psi_\Gamma)}{\sum_\Gamma p(\Gamma)\, \mathscr{N}(Y_i; \mu_\Gamma + \beta_i, \psi_\Gamma)}}_{\text{substitute GMM}}$$

$$= \frac{\sum_\Gamma p(\Gamma)\, \mathscr{N}(Y_i; \mu_\Gamma + \beta_i, \psi_\Gamma)\, \psi_\Gamma^{-1} (Y_i - \mu_\Gamma - \beta_i)}{\sum_\Gamma p(\Gamma)\, \mathscr{N}(Y_i; \mu_\Gamma + \beta_i, \psi_\Gamma)}$$

$$= \sum_\Gamma W_{i\Gamma} \left(\psi_\Gamma^{-1} (Y_i - \mu_\Gamma - \beta_i) \right)$$

Weight for the i-th voxel and tissue class Γ:

$$W_{i\Gamma} := \frac{p(\Gamma) \cdot \mathscr{N}(Y_i; \mu_\Gamma + \beta_i, \psi_\Gamma)}{\sum_\Gamma p(\Gamma) \cdot \mathscr{N}(Y_i; \mu_\Gamma + \beta_i, \psi_\Gamma)}$$

Gradient Computation (cont.)

Rewriting the equation:

$$\frac{\frac{\partial}{\partial \beta_i} p(Y_i|\beta_i)}{p(Y_i|\beta_i)} = \sum_\Gamma W_{i\Gamma} \left(\psi_\Gamma^{-1} (Y_i - \mu_\Gamma - \beta_i) \right)$$

$$= \sum_\Gamma W_{i\Gamma}\, \psi_j^{-1} (Y_i - \mu_\Gamma) - \sum_\Gamma W_{i\Gamma}\, \psi_\Gamma^{-1}\, \beta_i$$

$$= \overline{R_i} - \overline{\psi^{-1}}_i\, \beta_i$$

Mean residual:

$$\overline{R_i} := \sum_\Gamma W_{i\Gamma}\, \psi_\Gamma^{-1} (Y_i - \mu_\Gamma)$$

Mean inverse variance:

$$\overline{\psi^{-1}}_i := \sum_\Gamma W_{i\Gamma}\, \psi_\Gamma^{-1}$$

Gradient Computation (cont.)

Finishing gradient computation:

$$\begin{aligned}\nabla_\beta \log p(\beta|\boldsymbol{Y}) &= \overline{\boldsymbol{R}} - \overline{\Psi^{-1}}\beta + \frac{\nabla_\beta p(\beta)}{p(\beta)} \\ &= \overline{\boldsymbol{R}} - \overline{\Psi^{-1}}\beta - \Psi_\beta^{-1}\beta \\ &\stackrel{!}{=} 0\end{aligned}$$

It follows that:

$$\hat{\beta} = \boldsymbol{H}\overline{\boldsymbol{R}} \quad \text{with } \boldsymbol{H} \equiv \left[\overline{\Psi^{-1}} + \Psi_\beta^{-1}\right]^{-1}$$

$\boldsymbol{H}$ is a linear operator that is applied to the mean residual field.
In fact, $\hat{\beta}$ can be obtained by low pass filtering of $\overline{\boldsymbol{R}}$ and $\overline{\Psi^{-1}}$.

EM-Algorithm

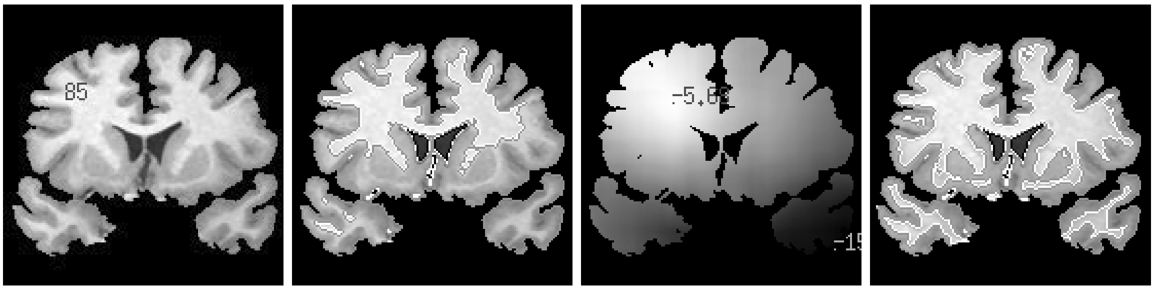
EM-Algorithm for the adaptive segmentation problem:

$$W_{i\Gamma} \leftarrow \frac{p(\Gamma)\cdot\mathcal{N}(Y_i|\mu_\Gamma+\beta_i,\psi_\Gamma)}{\sum_\Gamma p(\Gamma)\cdot\mathcal{N}(Y_i|\mu_\Gamma+\beta_i,\psi_\Gamma)} \tag{1}$$

$$\hat{\beta} \leftarrow \boldsymbol{H}\overline{\boldsymbol{R}} \tag{2}$$

- E-step: equation (1) yields the posterior tissue class probabilities for a known bias field
- M-step: equation (2) yields the new bias field for the current estimates for the tissue probabilities
- Result: iterating 5-10 times between the E- and the M-step is usually sufficient

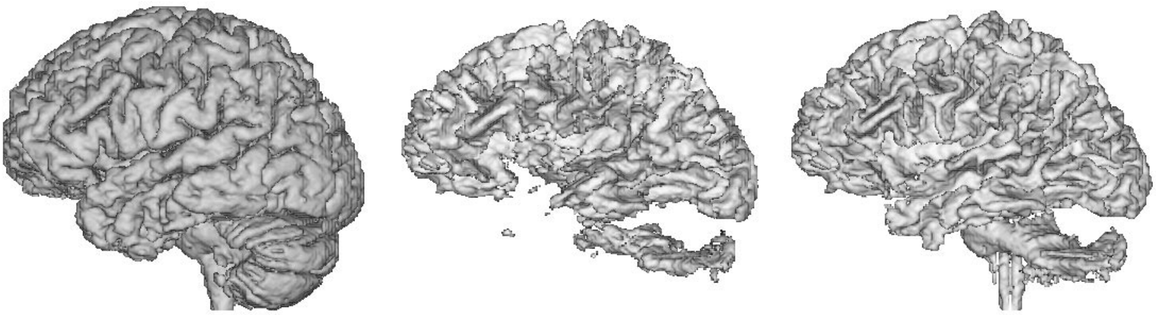
Results



(a) (b) (c) (d)

Fig.: Results of conventional segmentation (b) compared to adaptive segmentation (d) with computed bias field (c) on brain image (a) (Courtesy of W. Wells).

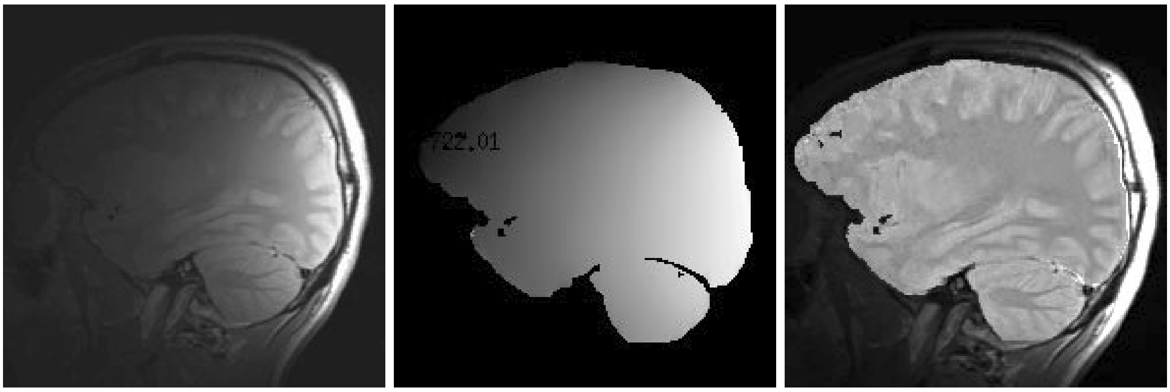
Results (cont.)



(a) (b) (c)

Fig.: Gray matter surface (a) for the previous image example, white matter surface of the conventional algorithm (b) and for the adaptive segmentation (c) (Courtesy of W. Wells).

Results (cont.)



(a) (b) (c)

Fig.: MRI image with bias field (a), computed bias field (b) and image corrected at the brain region (c) (Courtesy of W. Wells).

Model Extensions

Drawbacks of initial adaptive segmentation algorithm:

- brains should be extracted from entire data
- algorithm does not incorporate neighborhood of pixels
- purely intensity-based model

Extensions of the algorithm:

- incorporation of atlases for spatial probability maps of tissue classes
- definition of vector space for probabilistic atlases to get shape models
- voxel neighborhood relations modeled by Markov random fields
- incorporation into Bayesian model that is solved by EM approach

Model Extensions (cont.)

Result using an extended model:

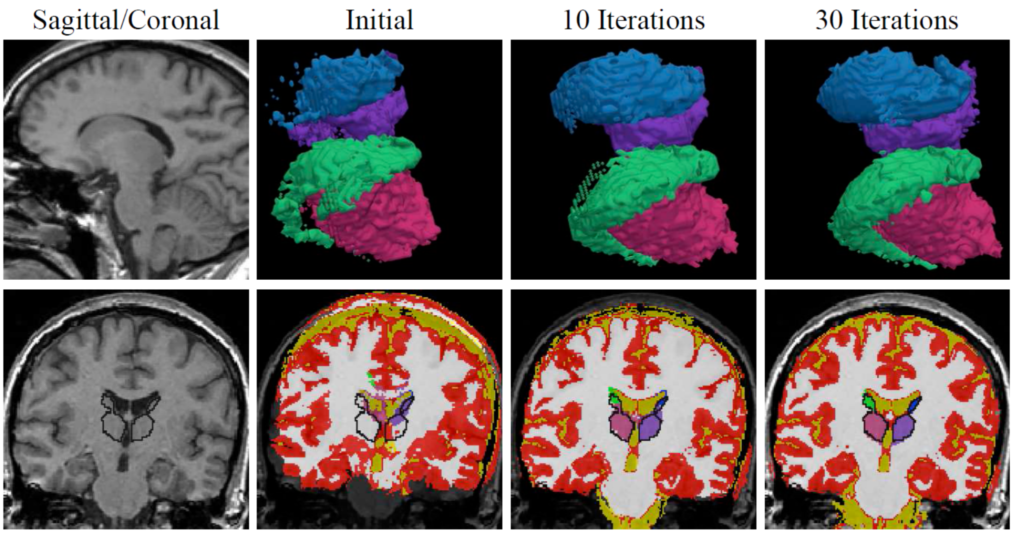


Fig.: MRI segmentation of the thalamus and caudate using an atlas-based EM segmentation algorithm (Courtesy of K. Pohl).

Lessons Learned

- Bayesian approach for MRI data segmentation
- incorporation of bias field estimation
- nonlinear problem is solved iteratively using EM algorithm
- improvement of results by incorporating atlases

Next Time in Pattern Recognition

Further Readings

- Original paper on adaptive MRI segmentation:

  W. M. Wells, R. Kikinis, W. E. L. Grimson, F. Jolesz:
  Adaptive segmentation of MRI data,
  IEEE Transactions on Medical Imaging, 15:429-442, 1996.

- F. Jäger, J. Hornegger:
  Nonrigid registration of joint histograms for intensity standardization in magnetic resonance imaging,
  IEEE Transactions on Medical Imaging, 28(1):137-150, 2009.

Further Readings (cont.)

Extensions of the model with shape models, atlas registration and MRFs:

- K. M. Pohl, J. Fisher, J. J. Levitt, M. E. Shenton, R. Kikinis, W. E. L. Grimson, W. M. Wells:
  A Unifying Approach to Registration, Segmentation, and Intensity Correction,
  Proc. MICCAI, pp. 310-318, 2005.
- K. M. Pohl, J. Fisher, S. Bouix, M. E. Shenton, R. W. McCarley, W. E. L. Grimson, R. Kikinis, W. M. Wells:
  Using the logarithm of odds to define a vector space on probabilistic atlases,
  Medical Image Analysis, 11(6), pp. 465-477, 2007.

Comprehensive Questions

- What is the idea of combined MR segmentation and bias field correction?
- What is the E-step in this context?
- What is the M-step?
- How can the update formulas be derived?