Medical Image Processing for Diagnostic Applications

Fuzzy C-means Clustering

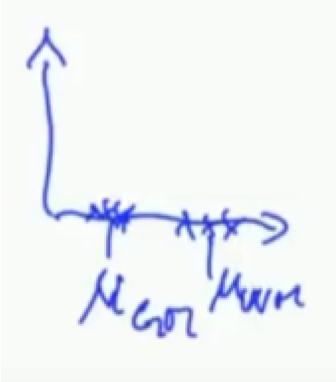
Online Course – Unit 25

Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics



find grey value mean values for different types of tissues

Fuzzy C-means Clustering

Regularization

The Regularized Optimization Problem

Summary

- Take Home Messages
- Further Readings

Fuzzy C-means Clustering

Definition

The ***fuzzy C-means objective function*** (FCM) for partitioning a set of observations into N_c classes which allows one data point to belong to more than one class:

$$J(x_1, x_2, \dots, x_n) = \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d ||x_k - c_i||^2.$$

- $c_1, c_2, \dots, c_{N_c}$ are the prototypes of the clusters. centers
- $x_1, x_2, \dots, x_n$ are the data points, in our particular case the logarithms of ideal intensities. -> grey values
- The ***probabilistic partition matrix*** $\boldsymbol{A} = [a_{i,k}]_{k=1,\dots,n}^{i=1,\dots,N_c}$ satisfies the probability constraint: -> prop for intensity in k belongs to cluster i

$$\sum_{i=1}^{N_c} a_{i,k} = 1 \quad \text{for all samples } k = 1, 2, \dots, n,$$

=1 so we get a propability

where $a_{i,k} \in [0, 1]$, $i = 1, \dots, N_c$.
- The ***fuzzifier*** $d \in [1, +\infty)$ is a weighting exponent.

Fuzzy C-means Clustering

Example

Let us assume we have three clusters $i = 1, 2, 3$, and five data points $k = 1, 2, 3, 4, 5$.

For k-means clustering the partition matrix is, for instance:

$$[a_{i,k}] = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

For fuzzy C-means clustering the partition matrix is, for instance:

$$[a_{i,k}] = \begin{pmatrix} 0.5 & 0.7 & 0 & 0.1 & 0.6 \\ 0.3 & 0.2 & 1 & 0.9 & 0.3 \\ 0.2 & 0.1 & 0 & 0 & 0.1 \end{pmatrix}.$$

prop for belonging to every class

Fuzzy C-means Clustering

There are a few drawbacks:

- The current objective function with the probabilistic assignment of data points to classes does not consider dependencies of neighboring data points. -> we only look at the grey value, but we also want to konsider the neighborhood
  Intuition: Neighboring data points most probably belong to the same class.
- The probabilistic approach requires mutually independent intensities.

The question now is, how can we incorporate dependencies of neighboring data points?

Topics

Fuzzy C-means Clustering: Regularization

Idea: Extend the FCM objective function by regularization.

- Regularization allows the
  - incorporation of *prior knowledge*, and/or -> same region = same class (propably)
  - introduction of *penalty terms*.
- This is achieved by adding a term that biases the solution of the optimization problem towards a piecewise homogeneous labeling.

Fuzzy C-means Clustering: Regularization

A possible **regularized** objective function is:

$$J_R(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d ||x_k - c_i||^2 + \lambda \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d \sum_{x_r \in \mathcal{N}_k} \frac{||x_r - c_i||^2}{\#\mathcal{N}_k},$$

produces low value if all pixels belon to the same class

where:

- $\mathcal{N}_k$ represents the particular set of neighbors of x_k,
- $\#\mathcal{N}_k$ is the cardinality of the considered neighborhood,
  size??
- $\lambda > 0$ is a weighting factor.

lambda big -> realy consider regularizer a lot

also consider aik -> how shure are we that the pixel belongs to the klass itself

Topics

The Regularized Optimization Problem

Let $[y_k]$ denote the biased logarithmized intensity values, and $[\beta_k]$ the logarithmized bias field. Now we replace the logarithm of the ideal intensity value x_k using $x_k = y_k - \beta_k$ and solve the optimization problem:

$$\left\{\hat{\boldsymbol{A}}, \hat{c}_i, \hat{\beta}_i\right\} = \underset{\boldsymbol{A}, c_i, \beta_k}{\arg\min} \left\{ \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d ||y_k - \beta_k - c_i||^2 + \lambda \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d \sum_{y_r - \beta_r \in \mathcal{N}_k} \frac{||y_r - \beta_r - c_i||^2}{\# \mathcal{N}_k} \right\},$$

subject to the probability constraint

A -> transition props.
C -> segmentation (centroids)
B -> Biast field

$$\sum_{i=1}^{N_c} a_{i,k} = 1 \quad \text{for all samples } k = 1, 2, \ldots, n.$$

The Regularized Optimization Problem

Membership evaluation:

- The optimization regarding $a_{i,k}$ has to satisfy the aforementioned probability constraint, i. e., the values must sum up to one for all k.
- For the incorporation of the probability constraint, we have to apply the Lagrange multiplier method.

Using Lagrange multipliers η_k, the extended objective function becomes:

because we have one constraint

$$J_R = \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^{d} \left(D_{i,k} + \frac{\lambda}{\#\mathscr{N}_k} E_{i,k} \right) + \sum_{k=1}^{n} \eta_k \left(1 - \sum_{j=1}^{N_c} a_{j,k} \right),$$

where

$$D_{i,k} = ||y_k - \beta_k - c_i||^2, \qquad \text{and} \qquad E_{i,k} = \sum_{(y_r - \beta_r) \in \mathscr{N}_k} ||y_r - \beta_r - c_i||^2.$$

Estimator for Partition Matrix

calculating the gradient gives us a ...

The computation of the zero crossings of the gradient w. r. t. the optimized variables results in the following estimator for the partition matrix:

$$\hat{a}_{i,k} = \frac{1}{\sum_{j=1}^{N_c} \left(\frac{\#\mathcal{N}_k D_{i,k} + \lambda E_{i,k}}{\#\mathcal{N}_k D_{j,k} + \lambda E_{j,k}} \right)^{\frac{1}{d-1}}},$$

and ...

Cluster Prototype Update

... the following update for the cluster prototypes:

$$\hat{c}_i = \frac{\sum_{k=1}^{n} a_{i,k}^d \left((y_k - \beta_k) + \frac{\lambda}{\#\mathcal{N}_k} \sum_{y_r \in \mathcal{N}_k} (y_r - \beta_r) \right)}{(1+\lambda) \sum_{k=1}^{n} a_{i,k}^d},$$

and ...

Bias Field Estimator

... the following estimator of the logarithmized bias field:

$$\hat{\beta}_k = y_k - \frac{\sum_{i=1}^{N_c} a_{i,k}^d c_i}{\sum_{i=1}^{N_c} a_{i,k}^d}.$$

since all the values (ABC) depend on each other we use a coordinated dencent methode -> find good A -> then good B -> ......

loop

Topics

Take Home Messages

- We have seen how to use (regularized) fuzzy c-means clustering to compute an estimator of the bias field.
- The regularization is needed to incorporate local dependencies of the image data points.

this algorithm does a lot of things at once -> can be computationaly expensive

Further Readings

The original paper on which the discussion in this unit is based on is:

Mohamed N. Ahmed et al. “A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data”. In: *IEEE Transactions on Medical Imaging* 21.3 (Mar. 2002), pp. 193–199. DOI: 10.1109/42.996338

If you want to know more about segmentation of MR images, e. g., consult

the Google Scholar record of ‘Sandy’ Wells’ publications.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. “A Review on MR Image Intensity Inhomogeneity Correction”. In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515