MFCM with the math done in continuum and with smoothness criterion on b

Awais Ashfaq[1] and Jonas Adler[2]

[1] KTH Royal Institute of Technology and Halmstad University
[2] KTH Royal Institute of Technology and Elekta Instrument AB

1 Introduction

This follows the paper "A modified fuzzy C means algorithm for shading correction in craniofacial CBCT images" but performs the deriviations in continuum and adds further regularity conditions on the bias field.

2 Bias-field estimation in continuum

The observed CBCT image is modelled as a sum of true signal and a spatially slow varying bias-field.

$$y(t) = x(t) + b(t) \quad \forall t \in \Omega \tag{1}$$

where
$\Omega \subset \mathbb{R}^n$ is the domain of the image
y is the observed intensity
x is the true intensity
b is the bias intensity

The FCM objective function is given by

$$J = \sum_{i=1}^{I} \int_{\Omega} \mu_i^m(x)(x(t) - c_i)^2 dt \tag{2}$$

where
I is the number of clusters
c_i is the center of the i^{th} cluster
$\mu_i(t)$ is the degree of membership of point t in the i^{th} cluster and $\mu_{in} \in [0\,,\,1]$.
m is the level of fuzziness. $m \in (1\,,\,\infty)$. $m = 1$ would imply a strict partitioning between clusters. Higher values of m imply smaller membership values and hence fuzzier clusters. For a given data point, sum of the membership values for every cluster is normalized to 1.

$$\sum_{i=1}^{I} \mu_i(t) = 1 \quad \forall\, t \in \Omega \tag{3}$$

The modified FCM objective function, following the idea as proposed by Ahmed et al is given by

$$J^* = \sum_{i=1}^{I} \int_{\Omega} \mu_i^m(t)[C_k(x - c_i)^2](t)dt \tag{4}$$

where $C_k x$ is the convolution of x with k, where k controls the neighbourhood influence.

$$C_k x = \int x(t - \tau)k(\tau)d\tau$$

In CBCT acquisition, intensity values correspond to attenuation coefficients of the material being scanned which in turn roughly correspond to the material's density. It is thus important to have a piece-wise smooth solution while keeping the true intensity values x close to the observed intensity values y (see Eq. 1). In order to preserve this information, we assume that the bias-field has zero mean.

$$\int_{\Omega} b(t)dt = \int_{\Omega} y(t) - x(t)dt = 0 \tag{5}$$

The task of minimizing J^* in Eq. 4 with constraints given by Eq. 3 and 5 is a nonlinear optimization problem. Mathematically,

$$\begin{aligned} \underset{\mu,c,b}{\text{minimize}} \quad & J^* \\ \text{subject to} & \sum_{i=1}^{I} \mu_i(t) = 1 \quad \forall\, t \in \Omega \\ & \int_{\Omega} (y(t) - x(t))dt = 0 \end{aligned}$$

2.1 Solution via coordinate descent

To minimize J^* we form the Lagrangian

$$\mathcal{L} = \sum_{i=1}^{I} \int_{\Omega} \mu_i^m(t)[C_k(x - c_i)^2](t)dt + \int_{\Omega} \gamma(t) \left(1 - \sum_{i=1}^{I} \mu_i(t)\right) dt + \lambda \int_{\Omega} (y(t) - x(t))dt \tag{6}$$

which we can find the explicit gradient of by straightforward computation

$$\begin{aligned}
\nabla_{\mu_i(t)}\mathcal{L} &= m\mu_i^{m-1}(t)[C_k(x-c_i)^2](t) - \gamma(t) \\
\nabla_{c_i}\mathcal{L} &= 2\int_{\Omega}[C_k^*\mu_i^m](t)(c_i - x(t))dt \\
\nabla_{x(t)}\mathcal{L} &= 2\sum_{i=1}^{I}[C_k^*\mu_i^m](t)(x(t)-c_i) - \lambda \\
\nabla_{\gamma(t)}\mathcal{L} &= \left(1 - \sum_{i=1}^{I}\mu_i(t)\right) \\
\nabla_{\lambda}\mathcal{L} &= \int_{\Omega}(y(t) - x(t))dt
\end{aligned}$$

where C_k^* is the adjoint of C_k. If k is symmetric, $C_k^* = C_k$. For an optimal solution, all of these derivatives would be simultaneously zero. However, finding such a point is non-trivial and we resort to coordinate descent to find an optimal solution. Solving for $\mu_i(t)$

$$\mu_i(t) = \left(\frac{\gamma(t)}{m[C_k(x-c_i)^2](t)}\right)^{\frac{1}{m-1}}$$

Using the constraint in Eq. 3 we calculate $\gamma(t)$ as

$$\gamma(t) = \frac{m}{\left(\sum_{j=1}^{I}\left(\frac{1}{[C_k(x-c_i)^2](t)}\right)^{\frac{1}{m-1}}\right)^{m-1}}$$

Update equation for membership matrix μ is given by

$$\mu_i(t) = \frac{1}{\sum_{j=1}^{I}\left(\frac{[C_k(x-c_i)^2](t)}{[C_k(x-c_j)^2](t)}\right)^{\frac{1}{m-1}}} \tag{7}$$

Update equation for cluster center c is given by

$$c_i = \frac{\int_{\Omega}[C_k^*\mu_i^m](t)x(t)dt}{\int_{\Omega}[C_k^*\mu_i^m](t)dt} \tag{8}$$

Update equation for x is given by

$$x(t) = \frac{\sum_{i=1}^{I} c_i[C_k^*\mu_i^m](t) + \lambda}{\sum_{i=1}^{I}[C_k^*\mu_i^m](t)} \tag{9}$$

and λ can be evaluated using constraint Eq. 5

$$\lambda = \frac{\int_{\Omega}\left(y(t) - \frac{\sum_{i=1}^{I} c_i[C_k^*\mu_i^m](t)}{\sum_{i=1}^{I}[C_k^*\mu_i^m](t)}\right)dt}{\int_{\Omega}\left(\sum_{i=1}^{I}[C_k^*\mu_i^m](t)\right)^{-1}dt} \tag{10}$$

3 Penalized square of bias

We here penalize on the square of the bias instead of using the zero mean condition 5, forcing it to be "low"

$$J^* = \sum_{i=1}^{I} \int_{\Omega} \mu_i^m(t)[C_k(x - c_i)^2](t)dt + \lambda \int_{\Omega} (y(t) - x(t))^2 dt \tag{11}$$

where λ is a regularization constant that can be selected. To solve this, we form the lagrangian

$$\begin{aligned}\mathcal{L} = \sum_{i=1}^{I} \int_{\Omega} \mu_i^m(t)[C_k(x - c_i)^2](t)dt + \\ \lambda \int_{\Omega} (y(t) - x(t))^2 dt + \int_{\Omega} \gamma(t) \left(1 - \sum_{i=1}^{I} \mu_i(t)\right) dt\end{aligned} \tag{12}$$

which we can find the explicit gradient of. All terms become the same as above, except the lambda derivative, which is no longer needed and x which is given by

$$\nabla_{x(t)}\mathcal{L} = 2\sum_{i=1}^{I} [C_k^* \mu_i^m](t)(x(t) - c_i) + 2\lambda(x(t) - y(t))$$

The updates are also the same as above, except for x

$$x(t) = \frac{\sum_{i=1}^{I} c_i [C_k^* \mu_i^m](t) + \lambda y(t)}{\sum_{i=1}^{I} [C_k^* \mu_i^m](t) + \lambda} \tag{13}$$

4 Penalized square of gradient of bias

We here additionally penalize on the square of the gradient of the bias, forcing it to be "low", so called Tichonov regularization

$$J^* = \sum_{i=1}^{I} \int_{\Omega} \mu_i^m(t)[C_k(x - c_i)^2](t)dt + \lambda_1 \int_{\Omega} (y(t) - x(t))^2 dt + \lambda_2 \int_{\Omega} ([\nabla(y - x)](t))^2 dt \tag{14}$$

To solve this we form the lagrangian

$$\begin{aligned}\mathcal{L} = \sum_{i=1}^{I} \int_{\Omega} \mu_i^m(t)[C_k(x - c_i)^2](t)dt + \\ \lambda_1 \int_{\Omega} (y(t) - x(t))^2 dt + \lambda_2 \int_{\Omega} ([\nabla(y - x)](t))^2 dt + \int_{\Omega} \gamma(t) \left(1 - \sum_{i=1}^{I} \mu_i(t)\right) dt\end{aligned} \tag{15}$$

which we can find the explicit gradient of. All terms become the same as above, except for x which is given by

$$\nabla_{x(t)}\mathcal{L} = 2\sum_{i=1}^{I}[C_k^* \mu_i^m](t)(x(t) - c_i) + \lambda_1(x(t) - y(t)) - 2\lambda_2[\varDelta(x - y)](t)$$

where $\varDelta = -\nabla^T\nabla$ is the Laplacian operator. The updates are also the same as above, except for x we get a system of equations

$$\sum_{i=1}^{I}[C_k^* \mu_i^m](t)x(t) + \lambda_1 x(t) - \lambda_2[\varDelta x](t) = \sum_{i=1}^{I} c_i[C_k^* \mu_i^m](t) + \lambda_1 y - \lambda_2[\varDelta y](t) \quad (16)$$

We can solve it using the CG method (`conjugate_gradient` in ODL).