Appendix 1

In general, the registration problem is formulated as an optimization problem which maximizes a cost function $C(T; I_F, I_M)$ with respect to the transform T. The transform T is obtained by optimizing the control point location in a manner which maximizes the similarity metric between the fixed image (I_F) and the moving image (I_M),

$$\hat{T} = \arg\max_{T} C(T; I_F, I_M) \quad [1]$$

with

$$C(T; I_F, I_M) = -S(T; I_F, I_M) + \alpha P(T) \quad [2]$$

where $P(T)$ is a regularization term which constrains non-rigid deformation and α is a weighting factor which balances the similarity metric $S(T; I_F, I_M)$ and the regularization term $P(T)$. We model the non-rigid transform T with a B-Splines based deformation field, where parameter μ models the transformation T. Finding the optimal transformation $\hat{T}$ therefore is an optimization problem, determining the parameter μ that maximizes the cost function $C(\mu; I_F, I_M)$,

$$\widehat{T_\mu} = \arg\max_{T} C(T_\mu; I_F, I_M) \quad [3]$$

or

$$\mu = \arg\max_{T} C(\mu; I_F, I_M) \quad [4]$$

In every iteration k, the current parameter μ_k is updated by adding a small step in direction of the derivative of the cost function $\partial C / \partial \mu$,

$$\mu_{k+1} = \mu_k - a_k \frac{\partial C}{\partial \mu} \quad [5]$$

where $\alpha_k > 0$ is the size of the step which changes in every iteration. Klein *et al.* (37) proved that using a decay of α_k according to $a_k = a / (k+A)^\gamma$, where $\alpha>0$, A≥1, and $0 \le y \le 1$ are user-predefined constants, the convergence rate significantly reduces computation time without affecting the final result. Based on this result, we used the stochastic gradient descent in our study.

References

37. Klein S, Pluim JPW, Staring M, Viergever MA. Adaptive Stochastic Gradient Descent Optimisation for Image Registration. Int J Comput Vis 2008;81:227.

 http://dx.doi.org/10.21037/qims-20-705