AbstractID: 11444 Title: A method to derive a probabilistic estimate of “truth” for the validation of tumor segmentation techniques

Purpose: To develop a robust technique that can combine biased segmentations from observers or multiple imaging modalities to estimate a “true” reference for evaluating tumor segmentation techniques.
Method and Materials: A probabilistic “truth” was estimated from a group of observers’ segmentations via the analysis of a maximum likelihood (ML) using simulated annealing (SA) and Expectation Maximum (EM) optimization algorithms. A loose condition, qualitative ranking of performance level of segmentations, was introduced to steer the SA based method to find an optimal solution when the relative qualities of input contours are known. The SA based method was compared with a previously published “truth” estimation method “STAPLE”, firstly, using digital phantoms, then, using clinical data - the gross tumor volumes (GTVs) of 12 head and neck cancer patients contoured by three experienced radiation oncologists using CT alone and subsequently, using PET /CT.
Results: The SA based method accurately determined the “truth” in digital phantom studies. In comparison, STAPLE was not correctly estimate the “truth”. In the study of the 12 head and neck patients, with a defined probabilistic “true” reference, the mean and range of sensitivity of the SA method were 0.90(0.70-0.98) compared to STAPLE 0.75 (0.45-0.94); the mean and range of volume ratio between the SA method and the reference were 0.964 (0.790-1.048), and between STAPLE and the reference were 0.762 (0.464-1.005). The reference volumes produced by the SA method were reliable and reasonable in all cases.
Conclusions: This work suggests that the SA method can accurately estimate a “truth” to validate segmentation methods which could potentially improve accuracy of radiation therapy or surgical target localization. The technique provides a reliable and convenient tool that could be used to optimally combine localization information from different sources to delineate clinical targets.