SCV 2008: December 1-3, Taipei, Taiwan (Abstract)

Object Oriented Analysis for Image Segmentation

Yueh-Yun Chi
Department of Epidemiology and Health Policy Research, University of Florida, Gainesville, FL

Abstract

High throughput technologies in medical imaging, genetics, and chemical analysis generate an ever increasing number of variables for each independent sampling unit. Such high dimension, low sample size (HDLSS) data present a grand challenge to statisticians: how do we find good estimates and make credible inference? We demonstrate the process of identifying valid and principled data reductions in comparing human and computer accuracy in segmenting kidneys in CT scans. First, we reduced the three dimensional array of distances for each image comparison to a histogram (the object) to be modeled separately. Second, we used nonparametric kernel density estimation to explore distributional patterns and assess multi-modality. Third, systematic search for parametric distributions and truncated variations led to choosing a cube root transformation to an approximate Gaussian. Fourth, representing each histogram by an individually estimated distribution eliminated the HDLSS problem by reducing on average 26,000 distances per histogram to 2 parameter estimates. In the fifth and final step we used classical statistical methods to demonstrate that the two human observers disagreed significantly less with each other than with the computer segmentation. Nevertheless, the size of all disagreements was clinically unimportant relative to the size of a kidney. The hierarchal modeling approach to object oriented data created response variables deemed sufficient by both the scientists and statisticians. We believe the same strategy will succeed in many other arenas.