

Title:

Stochastic Segmentation Models for Array-Based Comparative Genomic Hybridization Data Analysis

Author(s):

Tze Leung Lai, Haipeng Xing and Nancy Zhang

Technical Report number (Dept. of Statistics, Stanford Univ.):

2006-17

Date:

October 2006

Abstract:

Array-based comparative genomic hybridization (array-CGH) is a high-throughput, high-resolution technique for studying the genetics of cancer. Analysis of array-CGH data typically involves estimation of the underlying chromosome copy numbers from the log fluorescence ratios and segmenting the chromosome into regions with the same copy number at each location. We propose for the analysis of array-CGH data a new stochastic segmentation model and an associated estimation procedure that has attractive statistical and computational properties. An important benefit of this Bayesian segmentation model is that it yields explicit formulas for posterior means, which can be used to estimate the signal directly without performing segmentation, and for other posterior quantities that can be used to provide confidence assessments of any given segmentation. Simulation studies and applications to real array-CGH data illustrate the advantages of the proposed approach.