Minimum-Description-Length Criterion for Image Interpretation and Data Analysis in Spatial Informatics

He-Ping Pan

Cooperative Research Centre for Sensor Signal and Information Processing
SPRI Building, Technology Park Adelaide, The Levels
SA 5095, Australia
Email: heping@cssip.edu.au

Abstract

Spatial Informatics typically involves interpretation of remotely sensed images and analysis of multi-sources of data. Random perturbation in the observations and diversity of hypothesized models give rise to the uncertainty and difficulty of image interpretation and data analysis. The Minimum-Description-Length (MDL) principle is a best established criterion, which selects the best model with the minimal length of jointly encoding the data and the model. Although in terms of probability, MDL criterion is equivalent to the Maximum Aposterior Probability (MAP) criterion, it is advantageous at combining different data types and different model structures in a uniform measure - the total number of bits. It is more realistic to computerized information processing, because everything is discrete with limited resolution. This paper clarifies the formulation of the MDL criterion, its relationships to information theory, stochastic complexity, and Bayesian decision strategy. To sufficiently demonstrate the applicability of this criterion, a number of cases where this criterion can be and has been applied are described, including global interpretation of remotely sensed images for landuse mapping, line generalization, digital terrain modelling, spatial indexing in GIS, and unsupervised clustering. The emphasis is on showing how each of these classic problems can be reformulated under this new criterion. The reformulations are likely to lead to breakthroughs or significant progresses in the fields.