

Indexing Structure For Handling Uncertain Spatial Data

Bir Bhanu, Rui Li, Chinya Ravishankar, Michael Kurth, Jinfeng Ni
College of Engineering,
University of California, Riverside
{bhanu, rli}@vislab.ucr.edu ; {ravi, kurthm, jni,}@cs.ucr.edu

Key words: *Uncertainty, Indexing, Mixture of Gaussian, Expectation-Maximization, Mojave Desert*

Abstract:

Consideration of uncertainty in manipulation and management of spatial data is important. Unlike traditional fuzzy approaches, in this paper we use a probability-based method to model and index uncertain data in the application of Mojave Desert endangered species protection. The query is a feature vector describing the habitat for certain species, and we are interested in finding geographic locations suitable for that species. We select appropriate layers of the geo-spatial data affecting species life, called habitat features, and model the uncertainty for each feature as a mixture of Gaussian. We partition the geographic area into grids, assign an uncertain feature vector to each cell, and develop a *filter-and-refine* indexing method. The filter part is a bottom-up binary tree based on the automated clustering result obtained using the EM algorithm. The refine part processes the filtered results based on the “similarity” between the query and properties of each cell. We compare the performance of our proposed indexing structure with R-tree on the TIGER-Line dataset with synthetic uncertainty and real Mojave Desert data, and show that our scheme outperforms R-tree.