Establishing Concave Geometric Approximations to Better Capture Spatial Location, Extent and Shape for Geographic Information Retrieval

John Radke, Jianchun Xu , Patricia Frontiera

In the past decade digital collections that contain geographic data have gone from data poor to a data rich environments. Geographic data has grown at such a rate that the paradigm for geographic data discovery is changed forever, it is no longer practical to visually browse a library to determine if it contains relevant data for ones study. The adoption of metadata to catalog characteristics of each library entry provides an effective means by which queries can be performed and relevant data retrieved. One of the key metadata fields describing geographic content provides a description of a bounding extent or region over which the data is distributed.

The simplest, most conservative and most widely used structure to encode the spatial extent of a set of data is the bounding box, often referred to as the minimum bounding box (MBB). This convex shape scribes a box from the minimum x,y coordinates in the distribution to the maximum x,y coordinates. These bounding boxes can be compared to a query bounding box to determine the area of overlap and a probability that the data in the digital library satisfies the query. These statistics can be used to rank the relevance of each dataset with successful hits and offers a sound strategy for data retrieval.

Save for a few exceptions in a highly structured urban environment, it is rare that the spatial extent of a dataset would map onto a bounding box. This representation brings with it great imprecision. Although the convex hull offers some improvement, delineating concavity in polygonal representations promises higher precision in spatial matching and satisfying the query. We introduce and automate a spatial decomposition algorithm, an alpha shape, which delineates concave geometric approximations and better captures spatial location, extent and shape. We argue this method is a great improvement over both the MBB and the convex hull and will serve to better extract relevant data from digital geographic libraries or spatial databases.

Key words: Geographic data; digital library; minimum bounding box ; convex hull; spatial decomposition; alpha shape; geographic information retrieval; spatial databases.
.