Abstract

Discovering the meaningful patterns and trends out of large datasets needs a very special attention now a days, and one of the most prevalent and widely studied problems in this area is the detection and formation of clusters accurately and correctly. Previous works on this field does not meet the problem of 3D spatial datasets with minimization of Input Parameters. The objective of this paper is to present a Tetrahedron-density based clustering technique for large 3D datasets which we have named as 3D-CATD (Three Dimensional-Clustering Algorithm using Tetrahedron Density), for efficient clustering of 3D spatial data. This algorithm is capable of identifying embedded clusters of arbitrary shapes as well as multi-density clusters over large 3D spatial datasets. The polyhedron approach is being incorporated to perform the clustering where the number of points inside a tetrahedron (tetrahedron density) of a polyhedron is calculated using barycentric formulae for tetrahedron. This is because of the fact that partitioning of the data set can be performed more efficiently in tetrahedron shape than in any other 3D shape due to its smaller space dimension. The ratio of number of points between two tetrahedrons can be found out which forms the basis of nested clustering of 3D data. Experimental results establish the superiority of the technique in terms of cluster quality and complexity.
References

- M. Ester, H. P. Kriegel, J. Sander and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in International Conference on Knowledge Discovery in Databases and Data Mining (KDD-96), Portland, Oregon, 1996, pp. 226-231.
- L. Ertoz, M. Steinbach and V. Kumar, "Finding Clusters of Different Sizes, Shapes, and Densities in Noisy, High Dimensional Data," in SIAM International Conference on Data Mining (SDM 03), 2003.
An Extended Density based Clustering Algorithm for Large Spatial 3D Data using Polyhedron Approach

- Lance Parsons, Ehtesham Haque, Huan Liu, &quot;Evaluating Subspace Clustering Algorithms&quot;; Supported in part by grants from Prop 301 (No. ECR A601) and CEINT 2004.

Index Terms

Computer Science Algorithms

Keywords

Clustering Density-based Density Confidence Polyhedron approach
Tetrahedron-density