Decreasing Radius K-Nearest Neighbor Search using the Pyramid Technique

Qingxiu Shi and Bradford G. Nickerson

March 16, 2007

Abstract

A decreasing radius k-nearest neighbor search algorithm for the pyramid technique is presented. The Pyramid technique divides d-dimensional data space into $2d$ pyramids. Given a query point q, we initialize the radius of a range query to be the furthest distance of the k candidate nearest neighbors from q in the pyramid which q is in, then examine the rest of the pyramids one by one. After one pyramid is checked, the radius of the range query decreases or remains the same. For n uniformly and randomly distributed data points in $[0,1]^d$, we determine the search time $Q(n,d,k)$ for this algorithm to be bounded as $\Omega(d\log n+(2s)^d n) \leq Q(n,d,k) \leq O(d\log n+n(1-(1-4d^{1/d}s)^d))$, where s is the radius of the d-d ball enclosing a volume of k/n and $s \leq \frac{1}{4d^{1/d}}$. We experimentally compare the decreasing radius k-nearest neighbor search algorithm with the increasing radius k-nearest neighbor search algorithm using the Pyramid technique. The decreasing radius k-nearest neighbor search performance of the Pyramid technique is also compared to the BBD-tree, the SPY-TEC, the R*-tree and naive search. Experimental results show that our decreasing radius k-nearest neighbor search algorithm for the Pyramid technique is efficient when $d \leq \log_2 n$.