How Many Points in Euclidean Space can have a Common Nearest Neighbor ?

Kenneth Zeger[1] Allen Gersho[2]

[1]Coordinated Science Laboratory, Dept. of Elect. and Comp. Engineering, University of Illinois, Urbana, IL 61801
[2]Dept. of Elect. and Comp. Engineering, University of California, Santa Barbara, CA 93106

Abstract — **A Euclidean code is a finite set of *codepoints* in n-dimensional Euclidean space, $\mathcal{R}^n$. The total number of nearest neighbors of a given codepoint in the code is called its touching number. We show that the maximum number of codepoints F_n that can share the same nearest neighbor codepoint is equal to the maximum kissing number τ_n in n dimensions, that is, the maximum number of unit spheres that can touch a given unit sphere without overlapping. We then apply a known upper bound on τ_n to obtain $F_n \leq 2^{n(0.401+o(1))}$, which improves upon the best known upper bound of $F_n \leq 2^{n(1+o(1))}$. We also show that the average touching number of all the points in a Euclidean code is upper bounded by τ_n.**

I. Introduction

A *Euclidean code* is a finite set Y of $M > 1$ points in n-dimensional Euclidean space $\mathcal{R}^n$. The *touching number* T_α of a codepoint α is the number of touching points of α. The *total touching number* of a code is the sum of the touching numbers of each point in the code. The *average touching number* is the total touching number divided by M, the size of the code. Also of interest for a given codepoint α is the total number of codepoints which have α as a touching point. Denote by F_n the maximum number of points in $\mathcal{R}^n$ that can share a common nearest neighbor, where the maximum is taken over all possible arrangements of points.

II. Main Result

In a sphere packing, the *kissing number* (or *contact number*) of any sphere is the number of spheres in the packing that it is tangent to. The *maximum kissing number* in $\mathcal{R}^n$, denoted by τ_n, is the largest kissing number that can be attained by any packing of n-dimensional spheres.

The main results of this paper (Theorem 1 and Theorem 2) show that τ_n equals the maximum number F_n of codepoints that can share a common nearest neighbor codepoint and that τ_n upper bounds the average touching number of a code.

The number of points in $\mathcal{R}^n$ that can share a common nearest neighor point plays an important role in a widely diversified set of fields of research. For example, some researchers in psychology have investigated the nearest neighbor problem from a statistical point of view [1], [3]. In [2], an extensive experimental study was done to compute a histogram for the values of F_n based on a statistical model. Other studies of the number of points having a given point as a nearest neighbor have been done in such fields as sociology, biology, cognitive psychology, and ecology (e.g. see the references listed in [1]).

The quantity F_n also plays an important role in the field of nonparametric (distribution free) estimation. In [4] it is shown that F_n is independent of the code size for metrics induced by arbitrary norms in $\mathcal{R}^n$. A bound of $F_n \leq 3^n - 1 \approx 2^{1.585n}$ for all $n \geq 1$ was cited in [5] as an upper bound which is independent of the code size. Rogers [7, Theorem 3] derived bounds on the number of unit spheres needed to cover a given sphere of arbitrary radius when $n \geq 9$. Fritz [8], citing the bounds of Rogers, noted that F_n can be approximately upper bounded by $F_n \leq 2^{n(1+o(1))}$. Stone [6, Proposition 12] has shown that F_n can be upper bounded by the minimum number of 60° cones emanating from a point that can cover space. Combining Stone's result with the earlier result of Rogers also gives $F_n \leq 2^{n(1+o(1))}$. The essential difference between our bound and the weaker one of Stone and Rogers is that theirs is based on a *minimal covering* while ours is based on a *maximal packing*.

Theorem 1 *The maximum number of points in $\mathcal{R}^n$ which can have a common nearest neighbor is equal to the maximum kissing number (i.e. $F_n = \tau_n$), and is thus bounded as* $2^{0.2075\ldots n(1+o(1))} \leq F_n \leq 2^{n(0.401+o(1))}$.

Theorem 2 *The average touching number of any Euclidean code in $\mathcal{R}^n$ is less than or equal to the maximum kissing number τ_n, and thus is upper bounded by* $2^{n(0.401+o(1))}$.

[1]Supported in part by NSF Grant No. NCR-92-96231.

References

[1] A. Tversky, Y. Rinott, C. Newman, "Nearest Neighbor Analysis of Point Processes: Applications to Multidimensional Scaling," *Journal of Mathematical Psychology*, vol. 27, pp. 235-250, 1983.

[2] L. T. Maloney, "Nearest Neighbor Analysis of Point Processes: Simulations and Evaluations," *Journal of Mathematical Psychology*, vol. 27, pp. 251-260, 1983.

[3] C. Newman, Y. Rinott, A. Tversky, "Nearest Neighbors and Voronoi Regions in Certain Point Processes," *Advances in Applied Probability*, vol. 15, pp. 726-751, 1983.

[4] P. J. Bickel and L. Breiman, "Sums of Functions of Nearest Neighbor Distances, Moment Bounds, Limit Theorems and A Goodness of Fit Test," *Annals of Probability*, vol. 11, no. 1, pp. 185-214, 1983.

[5] L. P. Devroye and T. J. Wagner, "Distribution-Free Inequalities for the Deleted and Holdout Error Estimates," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 5, pp. 202-207, March 1979.

[6] C. J. Stone, "Consistent Nonparametric Estimation," *Annals of Statistics*, vol. 5, pp. 595-645, 1977.

[7] C. A. Rogers, "Covering a Sphere with Spheres," *Mathematika*, vol. 10, pp. 157-164, 1963.

[8] J. Fritz, "Distribution-Free Exponential Error Bound for Nearest Neighbor Pattern Classification," *IEEE Trans. Info. Theory*, vol. IT-21, no. 5, pp. 552-557, Sept. 1975.