On the Metric Properties of Discrete Space-Filling Curves *

C. Gotsman

Computer Science Department

Technion - Israel Institute of Technology

Haifa 32000, Israel

M. Lindenbaum

Computer Science Department

Technion - Israel Institute of Technology

Haifa 32000, Israel

Abstract

A space-filling curve is a linear traversal of a discrete finite multi-dimensional space. In order that this traversal be useful in many applications, the curve should preserve "locality".

We quantify "locality" and bound the locality of multi-dimensional space-filling curves. Classic Hilbert space-filling curves come close to achieving optimal locality.

EDICS: IP 3.1

Corresponding author:

Craig Gotsman

Dept. of Computer Science

Technion, Haifa 32000

Israel

Tel: +972-4-294336 Fax: +972-4-294353

Email: gotsman@cs.technion.ac.il

*A preliminary version of this work was presented at the IEEE International Conference on Pattern Recognition, Jerusalem, 1994.

1 Introduction

Denote $[N] = \{1, .., N\}$. A discrete m-dimensional space-filling curve of length N^m is a bijective mapping $C : [N^m] \longrightarrow [N]^m$ such that $d(C(i), C(i+1)) = 1$ for all $i \in [N^m - 1]$, where $d()$ is the Euclidean metric. In other words, the curve C of length N^m traverses all N^m points of the m-dimensional grid with side length N, making unit steps and turns only at right angles. For a historical account of classical space-filling curve constructions, see [8].

Space-filling curves are useful in applications where a traversal (scan) of a multidimensional grid is needed. Some algorithms perform local computations on neighborhoods, or exploit spatial correlation present in the data, so the preservation of "locality" during the traversal is desirable. By "locality", we mean that the traversal reflects proximity between the points of $[N]^m$, namely that points close in $[N]^m$ are also close in the traversal order, and vice versa. Sample applications are image halftoning ([11] and references therein), data organization [4], data compression [5] and color quantization [9].

The little work on this topic to date has addressed only one direction of this question, namely, how to design space-filling curves such that points close in the multidimensional space are also close along the curve. In general, as we shall show later (Theorem 2), this is impossible - for every space-filling curve C there will always be at least one pair of close points in $[N]^m$ which are very far apart along C. However, as these cases are rare, on the *average*, the situation will be much better. Perez, Kamata and Kawaguchi [7] quantify this using the average locality measure

$$L(C) = \sum_{i,j \in [N^m], i<j} \frac{|i-j|}{d(C(i), C(j))} \tag{1}$$

and describe a hierarchical construction for two-dimensional space-filling curves, which comes close to minimizing this measure.

Mitchison and Durbin [6] investigate similar measures of locality, taking into account only short (unit) Euclidean distances. This is because they regard the grid as an abstract lattice graph, ignoring its underlying geometry. They treat general two-dimensional mappings $C : [N^2] \longrightarrow [N]^2$ (not necessarily defining a curve). Their family of measures, parametrized by $q \in [0, 1]$, is:

$$\mathcal{L}_q(C) = \sum_{\{i,j \in [N^2] : i<j, d(C(i), C(j))=1\}} |i-j|^q \tag{2}$$

Interestingly enough, for the case $q = 1$, which may be compared to the measure (1), the optimal mapping turns out to be quite different from that in [7] (it is not even a curve).

For the more interesting case $q < 1$, which de-emphasizes longer distances along the curve, Mitchison and Durbin prove the lower bound

$$\mathcal{L}_q(C) \geq \frac{1}{1+2q} N^{1+2q} + O(N^{2q})$$

and provide an explicit construction C_N for any N with good, albeit suboptimal, locality. They conjecture that the optimal mapping must define a curve with a "fractal" character.

Voorhies [10] defines a more heuristic measure of locality, related to computer graphics applications, and experimentally compares the measures obtained for a variety of space-filling curves. He concludes that the Hilbert curve [3] is superior to other curves in this respect.

In this correspondence, we mainly address the converse question, i.e. to which extent can two points, which are close in the traversal order along the curve, be far apart in the multi-dimensional Euclidean metric. To quantify this, we use the following measures:

$$L_1(C) = \max_{i,j \in [N^m], i<j} \frac{d(C(i), C(j))^m}{|i-j|} \tag{3}$$

$$L_2(C) = \min_{i,j \in [N^m], i<j} \frac{d(C(i), C(j))^m}{|i-j|} \tag{4}$$

The use of the exponent m in (3) and (4) is justified by the fact that the maximal distance between points of $[N]^m$ is $O(N)$, and between two points of $[N^m]$ is $O(N^m)$. This correspondence presents bounds on $L_1(C)$ and $L_2(C)$.

Certain curve designs may be used to produce a *family* of curves for increasing values of N: $\mathcal{C} = \{C_N : N = 1, 2 \ldots\}$. In this case, it is interesting to investigate the limits (with a slight abuse of notation)

$$L_1(\mathcal{C}) = \lim_{N \to \infty} L_1(C_N)$$

$$L_2(\mathcal{C}) = \lim_{N \to \infty} L_2(C_N) \quad .$$

Essentially, we show that, for any m-dimensional curve family $\mathcal{C}$, if these limits exist, then $L_1(\mathcal{C}) \geq 2^m - 1$, and $L_2(\mathcal{C}) = 0$.

2 Locality of Space-Filling Curves

In this section, we provide a lower bound on L_1 and an upper bound on L_2 for multi-dimensional space-filling curves.

Theorem 1 *If C is a discrete m-dimensional $(m > 1)$ space-filling curve on $[N]^m$, then*

$$L_1(C) > (2^m - 1)(1 - 1/N)^m \quad .$$

Proof: Consider the 2^m corner points $\{1, N\}^m$ of the $[N]^m$ grid. Any space-filling curve must start at some arbitrary grid point, pass through these corner points in some order, and then end at another arbitrary grid point. Consider the increasing sequence of indices $\{P_i\}_{i=1}^{2^m}$ of these corner points along any such ordering. Fig. 1 shows the two possible distinct orderings of these 4 indices for the case $m = 2$. The Euclidean distance between any two consecutive points is $d(C(P_i), C(P_{i+1})) \geq N - 1$. Since $P_{2^m} - P_1 < N^m$, there exists an $1 \leq i < 2^m$ such that $P_{i+1} - P_i < N^m/(2^m - 1)$. For those two points we have $\frac{d(C(P_i),C(P_{i+1}))^m}{|P_i - P_{i+1}|} > (2^m - 1)(1 - 1/N)^m$. ▮

Remark: For the two-dimensional case, the limit constant given by Theorem 1 is 3. By a computerized exhaustive search, we have improved this to 3.25, implying that the bound of Theorem 1 is not tight. This was achieved by considering all possible paths through a specific configuration of 9 points in the plane, analogously to the 4 corner points in the theorem proof. ▮

Whether $|i - j|$ may be considered a good estimate for $d(C(i), C(j))^m$ depends on the existence of a positive constant lower bound on $L_2(C)$. The answer to this is negative, relying on the following discrete analog of the classic topological theorem ([2], Chap. 5, Theorem 2.3) that no mapping $f : [0, 1] \longrightarrow [0, 1]^m$ $(m > 1)$ can be continuous and also possess a continuous inverse.

Theorem 2 *If C is a discrete m-dimensional space-filling curve on $[N]^m$, then*

$$L_2(C) = O(N^{1-m}) \quad .$$

Proof: Choose a segment S of C of length at least $\frac{1}{4}N^m$ and at most $\frac{3}{4}N^m$. By the isoperimetric inequality on the multi-dimensional grid [1], the *boundary* of S, namely the set of grid points in S which have an immediate neighbor in $C - S$, denoted ∂S, satisfies $|\partial S| = \Omega(N^{m-1})$. [1] Form a list of the points of ∂S sorted by their

[1] $\Omega(f(N))$ denotes a quantity which is not less than $cf(N)$, for some constant c and sufficiently large N.

position along C, and choose the point p at the middle of this list. The distance along C from p to the nearest endpoint of S must now be at least $\lfloor \frac{1}{2}|\partial S| \rfloor$. This distance bounds from below the distance along C between p and any of its neighbors in $C - S$. The Euclidean distance between such a pair is 1, and the distance along the curve is $\Omega(N^{m-1})$, implying $L_2(C) = O(N^{1-m})$. ∎

3 Hilbert Curves

The standard m-dimensional raster space-filling curve on $[N]^m$, denoted R_N^m, does not have good locality properties, namely $L_1(R_N^m) = \Omega(N^{m-1})$ (see Fig. 2). On the other hand, the Hilbert curve [3] is an excellent example of a locality-preserving space-filling curve. The m-dimensional Hilbert curve of order k, denoted H_k^m, may be constructed recursively for any $N = 2^k$, as described in Fig. 3. The locality of the Hilbert curve is demonstrated by the following theorem, showing that it is close to optimal (compare with Theorem 1).

Theorem 3 *If H_k^m is a m-dimensional Hilbert curve on $[N]^m = [2^k]^m$, then*

$$L_1(H_k^m) \leq (m+3)^{m/2} 2^m \quad .$$

Proof: Consider any subpath of length n along H_k^m. There exists an integer r such that $(2^m)^{r-1} < n \leq (2^m)^r$. The fact that once H_k^m enters a grid quadrant of order r, it does not leave it until it has traversed all $(2^r)^m$ grid points in the quadrant, implies that the subpath must lie in the union of two adjacent quadrants containing $(2^r)^m$ grid points each (see Fig. 4(a)). If this were not true, the length of the subpath would be greater than $(2^r)^m$. The diameter d of the set of grid points traversed by the subpath satisfies $d^2 \leq (m-1+4)(2^r)^2$ (by Pythagoras' theorem), therefore

$$\frac{d^m}{n} \leq (m+3)^{m/2} 2^m \qquad ∎$$

The upper bound of Theorem 3 is far from tight in all cases. For the two-dimensional case, we can improve the constant 20 obtained from Theorem 3 almost to its optimal value:

Theorem 4 *If H_k^2 is a two-dimensional Hilbert curve on $[N]^2 = [2^k]^2$, then*

$$6(1 - O(2^{-k})) \leq L_1(H_k^2) \leq 6\frac{2}{3} \quad .$$

Proof: Using the terminology of the proof of Theorem 3, a more detailed analysis of subpath containment in quadrants of size 4^{r-1} (instead of quadrants of size 4^r) shows that one of the following six possibilities must hold:

1. $\frac{4}{16}4^r < n \leq \frac{5}{16}4^r$: $d^2 < \frac{5}{4}4^r$, hence $d^2/n \leq 5$.
2. $\frac{5}{16}4^r < n \leq \frac{6}{16}4^r$: $d^2 < \frac{29}{16}4^r$, hence $d^2/n \leq 5\frac{4}{5}$.
3. $\frac{6}{16}4^r < n \leq \frac{7}{16}4^r$: $d^2 < \frac{10}{4}4^r$, hence $d^2/n \leq 6\frac{2}{3}$.
4. $\frac{7}{16}4^r < n \leq \frac{8}{16}4^r$: $d^2 < \frac{10}{4}4^r$, hence $d^2/n \leq 5\frac{5}{7}$.
5. $\frac{8}{16}4^r < n \leq \frac{12}{16}4^r$: $d^2 < \frac{13}{4}4^r$, hence $d^2/n \leq 6\frac{1}{2}$.
6. $\frac{12}{16}4^r < n \leq 4^r$: $d^2 < 5 \cdot 4^r$, hence $d^2/n \leq 6\frac{2}{3}$.

For example, the subpath of Fig. 4(a) falls into category 5, as Fig. 4(b) illustrates. Taking the largest of the locality measures among these cases establishes the upper bound on $L_1(H_k^2)$.

For the lower bound, a subpath analogous to that illustrated in Fig. 5 (for $k = 6$, found by computer search) exists in H_k^2, for all $k > 1$, due to the recursive nature of the Hilbert curve. This subpath gives a locality measure of $6(1 - O(2^{-k}))$. Indeed, for H_k^2, it fills two adjacent quadrants of size $2^{k-3} \times 2^{k-3}$, two quadrants of size $2^{k-4} \times 2^{k-4}$ on either side of these two, aligned to a fixed direction, and so on until two quadrants of size 1. The Euclidean distance between the two endpoints is

$$d = 2\sum_{i=0}^{k-3} 2^i + 1 = 2^{k-1} - 1$$

and the distance along the curve is

$$n = 2\sum_{i=0}^{k-3} 4^i + 1 = \frac{2}{3}4^{k-2} + \frac{1}{3}$$

so $d^2/n = 6(1 - O(2^{-k}))$. ∎

Remark: For the three-dimensional case, Theorem 3 yields $L_1(H_k^3) \leq 117.56$. By computer simulation, we have found that $L_1(H_k^3) \leq 23$. ∎

4 Discussion

There remains a considerable gap between the lower bound on L_1 for general space-filling curves (Theorem 1), and the upper bound on L_1 for the Hilbert curve family (Theorem 3). This leaves open the question whether there exist families of space-filling curves with locality properties better than those of the Hilbert curves for all sizes.

It seems plausible that the Hilbert curves should also yield good results with respect to other measures of locality, such as that of Mitchison and Durbin [6]. These authors conjecture that the space-filling curve with optimal locality properties, measured by (2) with $q < 1$, must have a “fractal” character. Simulations performed by us show that, in agreement with this prediction, in some cases the (fractal) Hilbert curves indeed outperform the (non-fractal) curve constructed in [6].

In conclusion, we emphasize the practical implications of our results. Theorem 4 guarantees that if spatial correlation exists among the values of a discrete 2D data array, a 1D algorithm (such as that compressing a 1D data stream) may scan the the array along a Hilbert curve, and the loss in data correlation along the scan will be bounded.

Acknowledgements

Thanks to Gennady Agranov, Alfred Bruckstein and Lev Kaplan for helpful discussions on this subject.

References

[1] B. Bollobas and I. Leader. Edge-isoperimetric inequalities in the grid. *Combinatorica*, 11(4):299–314, 1991.

[2] J. Dugundji. *Topology*. Allyn and Bacon, Inc., Boston, 1966.

[3] D. Hilbert. Uber die stetige abbildung einer linie auf ein flachenstuck. *Mathematische Annalen*, 38:459–460, 1891.

[4] R. Laurini. Graphical databases built on Peano space-filling curves. In *Proceedings of Eurographics '85*, pages 327–338. North-Holland, 1985.

[5] A. Lempel and J. Ziv. Compression of two dimensional data. *IEEE Transactions on Information Theory*, 32(1):2–8, 1986.

[6] G. Mitchison and R. Durbin. Optimal numberings of an NxN array. *SIAM Journal on Discrete and Algebraic Methods*, 7(4):571–582, 1986.

[7] A. Perez, S. Kamata, and E. Kawaguchi. Peano scanning of arbitrary size images. In *Proceedings of the International Conference on Pattern Recognition*, pages 565–568. IEEE, 1992.

[8] H. Sagan. *Space-Filling Curves.* Springer Verlag, 1994.

[9] R.J. Stevens, A.F. Lehar, and F.H. Preston. Manipulation and presentation of multidimensional image data using the Peano scan. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(5):520–527, 1983.

[10] D. Voorhies. Space-filling curves and a measure of coherence. In J. Arvo, editor, *Graphics Gems II*, pages 26–30. Academic Press, 1991.

[11] Y. Zhang and R.E. Webber. Space diffusion: An improved parallel halftoning technique using space-filling curves. In *Computer Graphics Proceedings, Annual Conference Series*, pages 305–132. ACM, 1993.

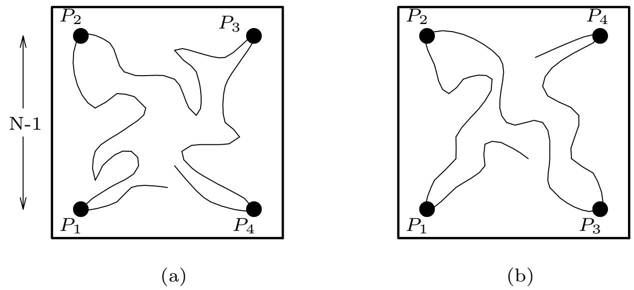


Figure 1: The two possible distinct traversals through the four corner points of the two-dimensional case considered in the proof of Theorem 1. All other traversals are symmetric to these.

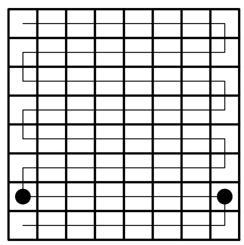


Figure 2: The two-dimensional raster space-filling curve R_N^2 on $[N]^2$. For the two grid points at the two ends of any scan line, both the Euclidean distance and that along the curve are $N-1$, so $L_1(R_N^2) = N-1$.

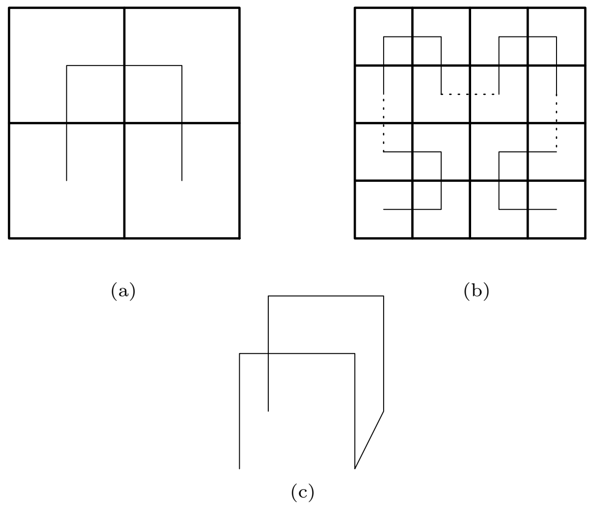
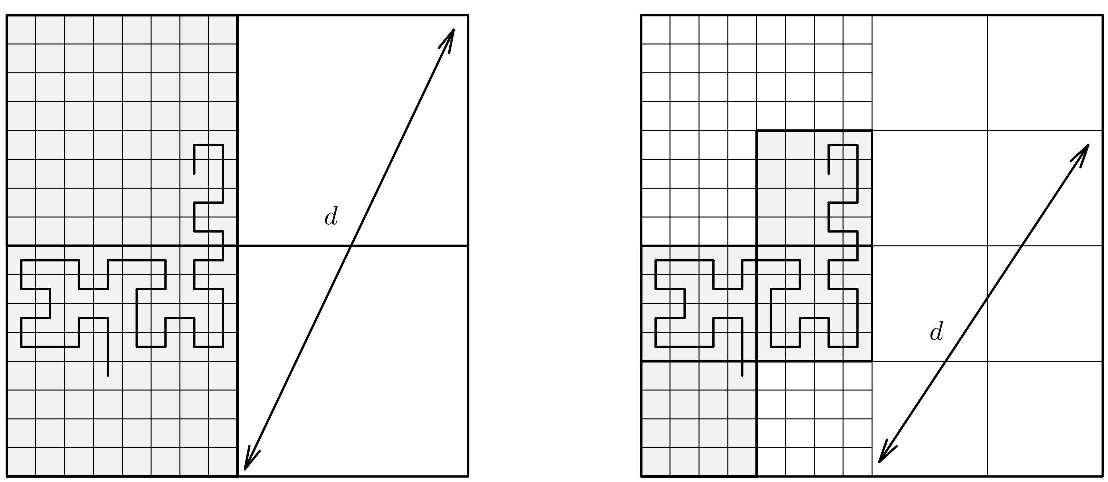


Figure 3: Recursive construction of the Hilbert curve: **(a)** The "seed" H_1^2. **(b)** H_2^2 constructed from 4 (rotated) versions of H_1^2. H_k^2 is constructed recursively in an analogous fashion. **(c)** The "seed" H_1^3. H_k^3 is constructed recursively from the seed curve.



$m = 2, n = 40, r = 3, d^2 = 8^2 + 16^2 = 320$

(a)

$m = 2, n = 40, r = 3, d^2 = 8^2 + 12^2 = 208$

(b)

Figure 4: Upper bounds on $L_1(H_4^2)$: **(a)** Quantities of Theorem 3. Any Hilbert subpath of length $16 < n \leq 64$ must lie within two adjacent quadrants of size 8×8. **(b)** Case 5 of Theorem 4. The subpath of (a), as any Hilbert subpath of length $32 < n \leq 48$, is actually contained in 4 adjacent quadrants of size 4×4, resulting in a tighter bound on d.

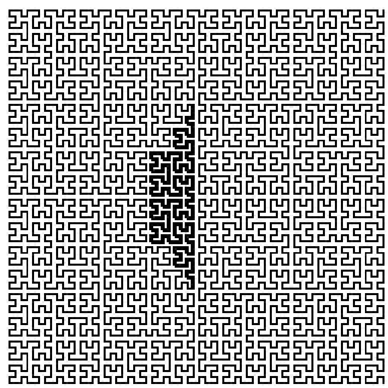


Figure 5: The "worst" subpath of H_6^2 determining the value of $L_1(H_6^2)$. This path was found by exhaustive search, using a computer. An analogous structure is present in H_k^2 for any value of k, due to the recursive nature of the Hilbert curve.