ELSEVIER

Discrete Applied Mathematics 117 (2002) 211–237

DISCRETE APPLIED MATHEMATICS

Towards optimal locality in mesh-indexings ☆

Rolf Niedermeier[a,*], Klaus Reinhardt[b,1], Peter Sanders[c]

[a] *Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13, D-72076 Tübingen, Germany*
[b] *Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13, D-72076 Tübingen, Germany*
[c] *Max-Planck-Institut für Informatik, D-66123 Saarbrücken, Germany*

Received 26 September 1997; revised 3 August 1999; accepted 26 July 2000

Abstract

The efficiency of many data structures and algorithms relies on "locality-preserving" indexing schemes for meshes. We concentrate on the case in which the maximal distance between two mesh nodes indexed i and j shall be a slow-growing function of $|i - j|$. We present a new two-dimensional (2-D) indexing scheme we call *H-indexing*, which has superior (possibly optimal) locality in comparison with the well-known Hilbert indexings. H-indexings form a Hamiltonian cycle and we prove that they are optimally locality-preserving among all cyclic indexings. We provide fairly tight lower bounds for indexings without any restriction. Finally, illustrated by investigations concerning 2-D and 3-D Hilbert indexings, we present a framework for mechanizing upper-bound proofs for locality. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Space-filling curve; Self-similar curve; Locality-preserving mesh-indexing; Locality-preserving grid-indexing; Hilbert-curve; Lower bound; Fractal

1. Introduction

For many fields in computer science, indexing schemes for meshes, that is, bijective mappings $\{0,\ldots,n-1\}^r \to \{0,\ldots,n^r-1\}$, plays a crucial role. For example, in computational geometry one often has to map an r-dimensional mesh onto a one-dimensional (1-D) traversal order or storage order. In this case, it is often advantageous if close-by raster points have close-by indices [3]. Analogous problems also arise in evaluating

☆ A preliminary version of parts of this work appear in the proceedings of the 11th International Symposium on Fundamentals of Computation Theory (Springer *LNCS* 1279, Springer, Berlin, pp. 364–375), held in Kraków, Poland, September 1–3, 1997.

* Corresponding author. Tel.: +49-7071-29-77568; fax: +49-7071-29-78962.

E-mail addresses: niedermr@informatik.uni-tuebingen.de (R. Niedermeier), reinhard@informatik.uni-tuebingen.de (K. Reinhardt), sanders@mpi-sb.mpg.de (P. Sanders).

[1] Work partially supported by DFG Project La 618/3-1 KOMET.

0166-218X/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0166-218X(00)00326-7

differential operators or even in a biological setting [20]. A conceptual problem with this notion of locality is that there are always raster points that are far apart from some other raster points. The converse notion of locality applies when a 1-D data structures is mapped to a multi-dimensional mesh. Here we are interested in indexing schemes which map close-by indices to close-by raster points. We will use the term $r \to 1$ locality for the first notion and the term $1 \to r$ locality for the latter. $1 \to r$ locality has the advantage that there are indexings for which locality can be achieved for all indices. Locality of type $1 \to r$ is also natural for applications in parallel processing on mesh-connected computers, where one often has to map 1-D data structures to the processor-mesh. If the communication requirements within this data structure are predominantly between close-by indices, it is advantageous to map them to close-by processors in order to decrease network contention and latency [6,7,21,25]. In this paper, we therefore concentrate on $1 \to r$ locality. We concentrate on worst case bounds – for example, this is the only way to exclude bottlenecks in parallel programs.

Several mesh-indexing schemes are well known. Most of these have been developed for the 2-D case, but they usually have generalizations for multiple dimensions, for example, row-major or snake-like row-major. However, taking a closer look at applications in parallel processing, one may observe that these kinds of indexings do not preserve locality of computation and communication very well. For example, for an r-dimensional mesh with side length n and generalized row-major indexing, processors 0 and $n-1$ are at distance $n-1$ from each other. Hence, a communication between these two processors ties up $n-1$ communication links and has a high latency. This is large compared to the distance of about $r\sqrt[r]{n}$ achievable if the first n processors could be arranged in a cube. A locality-preserving indexing should yield a distance $f(n) \in \mathrm{O}(\sqrt[r]{n})$. This should generalize to all pairs of processors within the mesh, that is, processors indexed i and j should be at most at distance $f(|i-j|)$ from each other. For example, a simple parallel variant of quicksort can be shown to run in average time $\Theta((n + \log m)m/n^r)$ for $m \geqslant n^r$ elements on n^r processors if a locality-preserving indexing scheme is used. This is asymptotically optimal and compared to other asymptotically optimal algorithms only $\Theta(\log n)$ rather than $\Theta(n)$ messages are sent on the critical execution path [25]. Quicksort, using row-major indexing and related schemes, needs time $\Theta((n \log n + \log m)m/n^r)$. Various other applications in parallel processing are discussed in [7,15,19]. Further applications of this kind of locality can be found in image processing and related fields (see [10] and the references cited there). See Section 3 for additional discussion.

In this paper, we consider $1 \to r$ locality in mesh-indexings using (discrete) space-filling curves. To analyze locality, we always make use of the three most important metrics in use: Manhattan, Euclidean, and maximum. One of the main contributions of this paper is the introduction of so-called *H-indexings* for 2-D meshes, which are based on a variant of the 2-D Sierpiński curve. H-indexings possess better locality than Hilbert indexings. In fact, we conjecture that they are optimally locality-preserving among all mesh-indexings. In other words, with respect to the Euclidean metric, we believe that for an $n \times n$-mesh, $n \geqslant 2$, in each indexing there must be indices i and j with $d_2(i,j) \geqslant \sqrt{4|i-j|} - c$, where c is some small constant.

We can show at least that this is true for the class of *cyclic indexings*. For example, we prove for H-indexings and the Euclidean metric $d_2(i,j) \leqslant \sqrt{4|i-j|-2}$ for arbitrary indices i and j. This is tight up to a small additive constant. This answers an open question from Gotsman and Lindenbaum [10] concerning the existence of a family of space-filling curves with locality properties better than those of Hilbert curves, where we have a constant factor of $\sqrt{6}$ instead of 2. Additionally, we have improved lower bounds for the locality attained through arbitrary indexings with respect to all three metrics mentioned above. Furthermore, we develop a technique for finding upper locality bounds by mechanically inspecting a finite number of cases. Consequently, this is applied to the 2-D Hilbert indexing and 3-D variants of the Hilbert indexing. This approach enables us to obtain simple and complete proofs of results that are new or previously relied on difficult to check proofs involving tedious manual case distinctions.

The paper is structured as follows. We introduce some notation in Section 2 and review related work in Section 3. In Section 4, we introduce H-indexings and show that they provide a better locality than 2-D Hilbert indexings. The general lower bounds indicating that the H-indexings may indeed be optimal are derived in Section 5. The technique for mechanically deriving upper bounds is developed in Section 6. This technique is shown by a simple yet complete proof for the locality properties of the 2-D Hilbert indexing with respect to the Manhattan metric. Then we adapt this method, so that it can be applied to 3-D variants of the Hilbert indexing and also include the Euclidean and maximum metrics. Section 7 summarizes the results of the paper and points out some areas requiring future research.

2. Preliminaries

In this paper, we work with 2-D and 3-D *meshes* (or, equivalently, *grids*). We concentrate on quadratic and cubic grids, where, for example, in the 2-D case we have n^2 points arranged in an $n \times n$-array. Meshes occur in various settings such as parallel computing, data structures, image processing, and many other fields of computer science. In the following, we restrict the description of some basic concepts to the 2-D case. Transferring this to a 3-D (and r-D) setting is straightforward.

We are interested in *indexing schemes* for meshes. An indexing scheme is simply a bijective mapping of $\{0,\ldots,n^2-1\}$ onto $\{0,\ldots,n-1\} \times \{0,\ldots,n-1\}$, thus providing a total ordering of the mesh points. We will study discrete space-filling curves and consider them to be special kinds of indexing schemes, which possess the desired property of locality preservation. To define locality, we first need a metric. We will use the *Manhattan* metric $d_1(a,b) = ||a-b||_1$, the *Euclidean* metric $d_2(a,b) = ||a-b||_2$, and the *maximum metric* $d_\infty = ||a-b||_\infty$ where $||(x,y)||_\alpha := \lim_{\beta\to\alpha}(|x|^\beta + |y|^\beta)^{1/\beta}$. By using the terms $x(i)$ and $y(i)$ we denote the position of a point i within the grid with respect to Cartesian coordinates.

A *discrete space-filling curve* $C: \{0,\dots,n^2-1\} \to \{0,\dots,n-1\} \times \{0,\dots,n-1\}$ fulfills $d_\infty(C(i), C(i+1)) = 1$. Thus one might say that space-filling curves provide *continuous* indexings. A space-filling curve traverses the grid-making unit steps and turning only at right angles. The meaning will always be clear from the context. Another feature of space-filling curves, besides being continuous, is usually their *self-similarity*. Self-similarity here simply means that the curve can be generated by putting together identical (basic construction) units, applying only rotation and reflection to these units. This becomes more obvious when considering the construction principles of Hilbert and H-curves in subsequent sections. To simplify presentation, in this paper the symbol i refers to its geometric location $(x(i), y(i))$ as well as to its index value. A *segment* $\overline{(i,j)}$ of a space-filling curve is the set $\{C(i),\dots,C(j)\}$ of mesh nodes. Our measure of locality is based on the requirement that for close-by indices i, j, with small $|i-j|$, the distance $d(i,j)$ defined by one of the above metrics should also be small. We call a continuous indexing *cyclic* if $d_2(0, n^2-1) = 1$. In this case we compute modulo n^2, that is, we use the additive group $(\{0,\dots,n^2-1\}, +)$ for adding and subtracting indices. Also, for cyclic indexings $|i|$ will denote the difference between i and 0 modulo n^2, thus $|i| \leqslant n^2/2$. Put simply, these assumptions express the following: For cyclic indexings it is unimportant at which point the numbering starts.

3. Related work

We cite some of the more recent papers from various fields dealing with locality questions for meshes and using space-filling curves as indexing schemes. We pay particular attention here to the field of parallel processing and give a short account of the development of locality-preserving indexings in this field.

Whereas we are studying $1 \to r$ locality, $r \to 1$ locality is for example studied by Mitchison and Durbin [20], who present some optimal results for this setting. Refer also to the paper of Gotsman and Lindenbaum [10], for a short discussion on various locality measures and related results. Locality of type $r \to 1$ is important when geometrical data is to be mapped onto a 1-D domain, e.g., in parallel gravitational particle simulation [26], for graph partitioning [14] and fast range queries for geometrical data stored on disks [3,4].

Whenever there is a requirement for some kind of locality in mesh-indexings, space-filling curves, and, in particular Hilbert indexings [2–4,6,7,9,10,12,13,23,25] seem to come into play.

Gotsman and Lindenbaum [10] study $1 \to r$ locality for the Euclidean metric that plays an important role in fields such as image processing and computer graphics. They primarily consider Hilbert's space-filling curve and provide upper and lower bounds. We improve their upper and lower bounds in the 2-D case.

The Manhattan metric is particularly important in the field of parallel processing on mesh-connected processor arrays. Here, good locality of an indexing scheme for the processors may lead to reduced communication costs [6,7,15,19,25]. (The same applies to the maximum metric, which is more suitable for grids with diagonal connections, cf. e.g. [16,17].) For the Manhattan metric and the field of parallel processing, we delve into more detail about the history of results and applications. Stout [27] seems to be the first who used the so-called *proximity orderings* in the context of 2-D mesh algorithms. We call them *Hilbert indexings* due to the direct relation to Hilbert's space-filling curve [11,24]. Subsequently, they have been used to speed up a wide variety of parallel algorithms: computational geometry [19], fast backtracking and branch-and-bound [15], mapping of pyramid networks [8], simulation of abstract parallel computation models [7,21], and parallel quicksort [25]. Quantitative analysis concerning the properties of locality-preserving indexing schemes have, so far, focused mainly on the 2-D Hilbert-indexing. According to Stout "there is a constant $c < 4$ such that processors numbered i and j are no more than $c\sqrt{|i-j|}$ communication links apart" [27, p. 27]. This was then proved by Kaklamanis and Persiano [15] for $c=4$. Recently, a bound of $3\sqrt{|i-j|}$ has been proved by Chochia et al. [7]. However, the proof is quite complicated. We present a fairly simple and complete proof of this result and show that H-curves, to be introduced in the next section, are better than Hilbert curves with respect to locality. Lately, Chochia and Cole [6] attained results for 3-D Hilbert indexings. These are also complemented by our results and more recent related work [2].

Buhrman et al. explain how *average* case lower bounds for the $1 \to r$ locality can be obtained using a simple counting argument and the concept of Kolmogorov complexity [5]. For the 2-D case and the Euclidean metric they show that $d_2(i,j) \geqslant \sqrt{0.636|i-j|}$ for any i and $\Omega(n^2)$ choices for j. Furthermore, $d_2(i,j) \geqslant \sqrt{2.5|i-j|}$ if i is mapped to a corner point.

4. The H-indexing

Gotsman and Lindenbaum [10, p. 797] posed the question as to "whether there exist families of space-filling curves with locality properties better than those of the Hilbert curves for all sizes". One of the main contributions of this paper is to answer this question affirmatively. Our result not only applies to the Euclidean metric as studied by Gotsman and Lindenbaum, but also to the Manhattan and the maximum metrics. In this section we introduce H-indexings and analyze their locality properties showing, the claimed improvement compared with Hilbert indexings. Section 5 argues that H-indexings are optimally locality-preserving among all discrete space-filling curves as they provide tight lower bounds.

4.1. Construction scheme

H-indexings are related to 2-D Sierpiński curves [24]. As the name indicates, H-indexings have an "H-shaped" form. In analogy to Hilbert indexings, we obtain

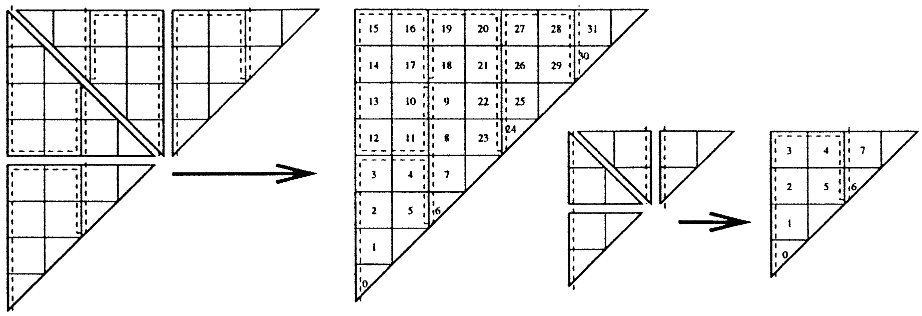


Fig. 1. H-indexings are built using triangles as building blocks.

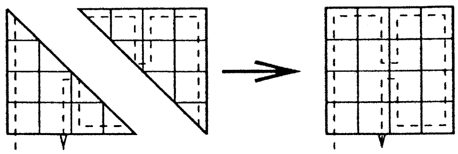


Fig. 2. Building an H-indexing for a square using two triangles.

indexings for $2^k \times 2^k$-meshes[2] by means of an inductive method. There is, however, a decisive difference. Whereas in the case of Hilbert indexings the building blocks are four smaller squares (cf. Section 6 and Fig. 7 there), the construction of H-indexings is easier to describe using right-angled triangles. For Hilbert indexings we only have one building block to which we apply rotation or reflection. To build the final mesh indexing, we put together two triangles. Fig. 1 shows the construction of a triangle from 4 smaller triangles. A triangle with 8 mesh nodes is constructed from triangles with only two nodes and a triangle with 32 nodes is constructed from those with 8 nodes. Observe that the triangles are constructed so that precisely every *other* mesh node along the diagonal belongs to the nodes of the triangle. Thus an indexing scheme for a square mesh can be obtained as shown in Fig. 2. Alternatively, Fig. 3 shows how for all $k > 1$ an H-indexing through a square of size 4^k is built from 4 H-indexings through squares of size 4^{k-1} each. For subsequent proofs, however, it is more convenient to make use of the construction principle based on triangles.

For computer-assisted construction, we can describe the H-indexing of a $2^k \times 2^k$ mesh by expressing the coordinates $x(i)$ and $y(i)$ of the ith point recursively in the following way. Fig. 4 best demonstrates the subsequently given recurrences for $x(i)$ and $y(i)$. The recurrences relate directly to the recursive construction principle of H-curves. Consider Fig. 4: The H-curve starts in the lower left corner with index 0. Let $h{:=}4^k/32$, where 4^k is the total number of mesh points. The H-curve first traverses the "triangle" (see Fig. 4) containing 0, then that containing h, then that containing $2h$, then that

[2] A Java program for the general case of non-cubic meshes with arbitrary side-lengths can be found at `http://www-fs.informatik.uni-tuebingen.de/~reinhard/hcurve.html`.

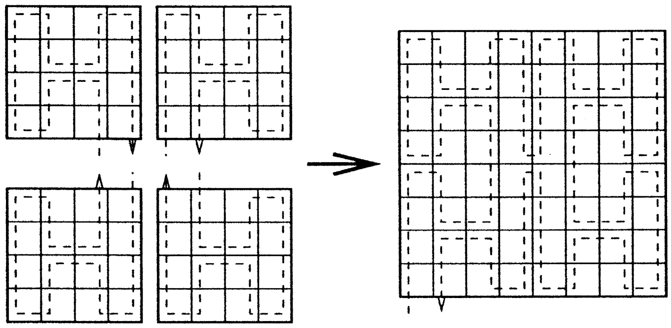


Fig. 3. Inductive construction principle of H-indexings.

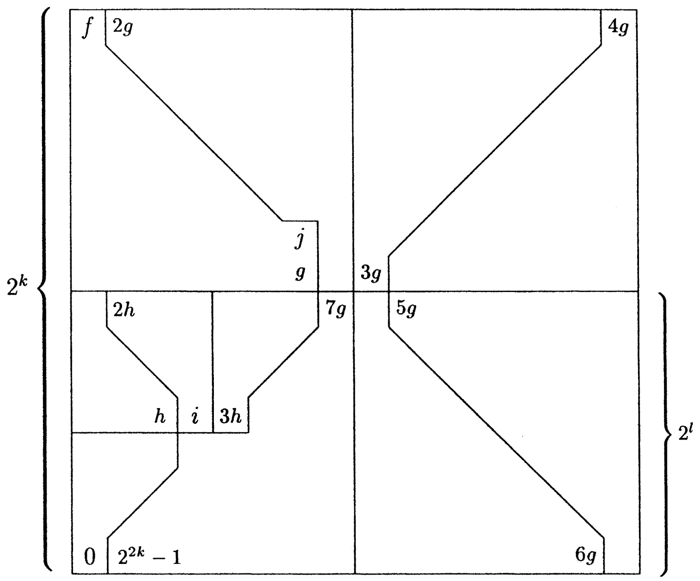


Fig. 4. The positions of the points i and j for the worst cases. The recursion is shown for $l = k - 1$. Let $g = 2^{2l-1}$ and $h = 2^{2l-3}$.

containing $3h$, until at $g = 4h$ it enters the upper left quadrant. From there it goes through f and then $2g$ and so on, always following some kind of triangle structure. Most importantly, this triangle structure acts recursively, thus leading to the somewhat complicated recurrence given below. Its correctness has been checked by computer. Note that in Fig. 4, i and j are located at some special points, which, as will later be shown, form a "worst case pair" of indices concerning the locality for the H-indexing.

Observe that the subsequent parameter l is uniquely determined in each recursive step by the if-conditions of the various cases; l ranges from $k-1$ to 1.

$$x(i)=\begin{cases} 2^k-1-x(i-2^{2k-1}) & \text{if } i\geqslant 2^{2k-1},\\ 2^l+x(i-3\cdot 2^{2l-1}) & \text{if } 4\cdot 2^{2l-1}>i\geqslant 3\cdot 2^{2l-1},\\ 2^l-1-x(3\cdot 2^{2l-1}-1-i) & \text{if } 3\cdot 2^{2l-1}>i\geqslant 2\cdot 2^{2l-1},\\ x(2^{2l}-1-i) & \text{if } 2\cdot 2^{2l-1}>i\geqslant 1\cdot 2^{2l-1},\\ 0 & \text{if } i\leqslant 1.\end{cases}$$

$$y(i)=\begin{cases} 2^k-1-y(i-2^{2k-1}) & \text{if } i\geqslant 2^{2k-1},\\ 2^l+y(i-3\cdot 2^{2l-1}) & \text{if } 4\cdot 2^{2l-1}>i\geqslant 3\cdot 2^{2l-1},\\ 2^l+y(3\cdot 2^{2l-1}-1-i) & \text{if } 3\cdot 2^{2l-1}>i\geqslant 2\cdot 2^{2l-1},\\ 2^{l+1}-1-y(2^{2l}-1-i) & \text{if } 2\cdot 2^{2l-1}>i\geqslant 1\cdot 2^{2l-1},\\ i & \text{if } i\leqslant 1.\end{cases}$$

The following results for "worst case distances" between points indexed by the H-curve are to be compared with the subsequent Theorem 1 presenting upper bounds for the locality of H-indexings. The Euclidean worst case (cf. Fig. 4) for each k are pairs of points $i=3\cdot 2^{2k-5}-1$ and $j=2^{2k-3}+1$ with $|i-j|=2^{2k-5}+2$ and

$$\begin{aligned} d_2(i,j) &= \sqrt{(x(i)-x(j))^2+(y(i)-y(j))^2}\\ &= \sqrt{(2^{k-2}-1-2^{k-1}+2)^2+(2^{k-2}-2^{k-1}-1)^2}\\ &= \sqrt{4(2^{2k-5}+2)-8+2}=\sqrt{4|i-j|-6}.\end{aligned}$$

The same pairs are also responsible for the worst case in the Manhattan metric

$$\begin{aligned} d_1(i,j) &= |x(i)-x(j)|+|y(i)-y(j)|\\ &= -2^{k-2}+1+2^{k-1}-2-2^{k-2}+2^{k-1}+1=2^{k-1}\\ &= \sqrt{8\cdot 2^{2k-5}}=\sqrt{8(|i-j|-2)}.\end{aligned}$$

Thus, in both cases we observe the worst cases on a diagonal direction (from i to j). In the maximum metric, however, the worst cases are from 0 to $f=2^{2k-2}-1$ (see Fig. 4) with $|0-f|=2^{2k-2}-1$ and

$$d_\infty(i,j)=2^k-1=2\sqrt{|0-f|+1}-1.$$

4.2. Upper bounds

In this subsection, we give results for locality properties of H-indexings with respect to the Euclidean, the Manhattan, and the maximum metric.

Theorem 1. *For two arbitrary indices i and j, $i\neq j$, on the H-indexing the following is true*:

1. $d_1(i,j) \leqslant \sqrt{8(|i-j|-2)}$ *for* $|i-j|>3$,
2. $d_2(i,j) \leqslant \sqrt{4|i-j|-2}$,
3. $d_\infty(i,j) \leqslant 2\sqrt{|i-j|+1}-1$.

Observe that upper and lower bounds match for the Manhattan metric and the maximum metric. For the Euclidean metric we had a lower bound of $\sqrt{4|i-j|-6}$ which is only $O(1/\sqrt{|i-j|})$ away from the upper bound – less than an additive constant.

Theorem 1 shows that H-indexings provide an improvement in locality compared to Hilbert-curves, answering an open question given by Gotsman and Lindenbaum [10]. Focusing their attention on the Euclidean metric, they proved that for Hilbert curves C with respect to their locality measure $L_1(C):=\max_{i,j\in\{1,\dots,n^2\},i<j} d_2(i,j)^2/|i-j|$ it holds $6\cdot(1-O(2^{-k}))\leqslant L_1(C)\leqslant 20/3$, where $n=2^{2k}$ with $k>1$. Our result implies that for H-indexings C we have $L_1(C)=4$. To present our result of Theorem 1, we preferred to make a more concrete and more precise statement (which even includes additive constants) than the "$L_1(C)$-notation" allows.

Both the maximum metric and the Manhattan metric are of specific relevance in parallel processing [7,21,25]. Another advantage of H-indexings over Hilbert indexings is that they do not just describe a Hamiltonian path, but a Hamiltonian cycle through the mesh as well. This is useful, e.g. for parallel algorithms which employ communication along a virtual ring network. Interestingly, H-indexings are optimally locality-preserving among all Hamiltonian cycles through a square mesh, as the next section shows.

As it turns out, proofs that give the above tight results *including* additive constants are fairly technical [22] and have been omitted here. As shown below, however, slightly weaker results regarding the additive constants can be proved in an elegant way.

Theorem 2. *For two arbitrary indices i and j on the H-indexing the following is true*:
1. $d_1(i,j) \leqslant \sqrt{8|i-j|}+4$,
2. $d_2(i,j) \leqslant 2\sqrt{|i-j|}+\sqrt{10}$,
3. $d_\infty(i,j) \leqslant 2\sqrt{|i-j|}+3$.

Proof. We concentrate on proving the result for the Euclidean metric $d_2(i,j)$. The statements for the Manhattan metric $d_1(i,j)$ and the maximum metric $d_\infty(i,j)$ then easily follow by the general relations

$$d_1(i,j) \leqslant \sqrt{2}d_2(i,j)$$

and

$$d_\infty(i,j) \leqslant d_2(i,j).$$

The proof for $d_2(i,j)$ works by induction on the size of the smallest triangle (according to the construction principle of H-curves) containing both i and j. Note that all these triangles are right-angled and contain 2^l mesh points for $l \geqslant 1$. Hence the induction operates on l. For $l=1$ and $l=2$ the claim can be trivially checked. Consider a triangle of size 8 (8-triangle for short), that is, $l=3$, as drawn in Fig. 5. For each of the nodes in an 8-triangle we assign a representative which is located on the corners of

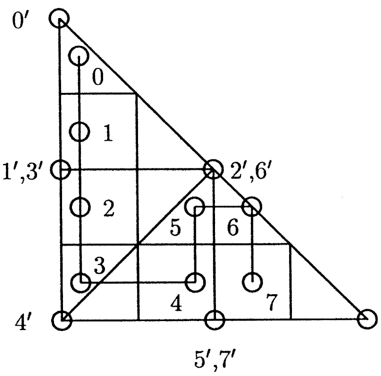


Fig. 5. Indexing nodes ($\{0,1,2,3,4,5,6,7\}$) in a triangle of size 8 and their representatives ($\{0',1',2',3',4',5',6',7'\}$). Note that $1'$ and $3'$, $2'$ and $6'$, and $5'$ and $7'$ each have the same location.

the 4 subtriangles as drawn in Fig. 5. The two representatives of a 2-triangle are determined as follows: If possible, rotate the 2-triangle in so that it has the same orientation (the vertical cathetus to the left, the horizontal cathetus to the bottom) as the original 8-triangle. The two representatives are then (in the case of Fig. 5) at the endpoints of the vertical cathetus. Observe that in Fig. 5, the 2-triangle containing nodes 4 and 5 cannot be rotated in so that it has the same orientation as the 8-triangle. In this case, we speak of the complementary[3] triangle and here the endpoints lie on the horizontal cathetus. Note that each right-angled triangle can be brought (by rotation) in one of the orientations “one cathetus as bottom line and one cathetus either to the left or to the right as vertical line”.

Let i and j be two arbitrary nodes and let $l > 2$. Let i' and j' be the representatives of i and j, respectively, which are obtained by applying the above rules to the 8-triangles containing i and j.

We show by induction on l that

$$d_2(i',j') \leqslant 2\sqrt{|i'-j'|}. \tag{1}$$

Observe that the numerical values of i and i', j and j', respectively, are the same, only their geometric positions differ a little. We introduce specifically the convention that a “2^l-triangle” may contain 2^l+1 representatives, where the 2^l+1st is also the first node of the subsequent triangle. This assumption is solely due to technical reasons. Our claim can be deduced from Eq. (1), because the Euclidean distance between an index i and its representative i' (for example, 2 and $2'$) may be at most $\sqrt{(1/2)^2+(3/2)^2}=\sqrt{10}/2$. Hence, $d_2(i,j) \leqslant d_2(i',j')+\sqrt{10}$, in the Manhattan case we have $d_1(i,j) \leqslant d_1(i',j')+4$, and in the maximum case we have $d_\infty(i,j) \leqslant d_\infty(i',j')+3$.

It remains to prove Inequality (1) by induction on l. The claim for $l=1$ and 2 can be easily checked (cf. Fig. 5). Now let i' and j' be in two different halves of their (smallest) “surrounding” triangle (otherwise the induction hypothesis applies). Due to

[3] The triangle mirrored at the vertical axis.

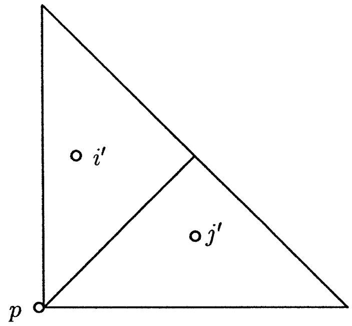


Fig. 6. Two representatives in the two halves of the smallest triangle containing both of them.

our definition of representatives we can assume (up to rotation) a situation as drawn in Fig. 6. In Fig. 6, the point p located at the right angle always represents a point in the indexing and the angle between i', p, and j' is at most 90°. Thus the Euclidean distance between i' and j' can be bounded from above using Pythagoras' theorem and the induction hypothesis:

$$\begin{aligned} d_2(i',j') &\leqslant \sqrt{d_2^2(i',p)+d_2^2(p,j')} \\ &\leqslant \sqrt{4|i'-p|+4|p-j'|} \\ &= 2\sqrt{|i'-j'|}. \end{aligned}$$

This verifies Inequality (1) and the proof is completed. □

In the next section, we show that H-indexings are quite close to optimal locality mesh-indexings.

5. Lower bounds

This section indicates that H-indexings might be optimal in locality-preservation among all indexings of 2-D meshes. We conjecture that they are optimal for the Euclidean, the maximum, and the Manhattan metric. Due to the fact that the difficulty for a general proof lies in “coming to grips with the loose ends”, we support this conjecture by showing the optimality among the cyclic indexings.

The idea at the core of the lower-bound proofs in this section is described in the following. As a rule, we pick a small number of points in the mesh. Every mesh indexing has to traverse these points in some specific order. Considering all possible orders and having picked out these mesh points carefully, we can focus on the argument that no matter what the indexing is, two of the indices picked, i and j, must have mesh distance $d(i,j)\leqslant c\sqrt{|i-j|}-d$ for constants c and d. In the subsequent proofs, we give values for c and d and prove their correctness by contradiction. The values for c and d were found by analyzing some concrete examples and deriving from these conjectures concerning c and d, which are proved here. Generally, these lower-bound proofs are based on case distinctions with respect to the order in which the selected mesh points are traversed by the indexing. The heart of all proofs is the well thought out selection of the appropriate mesh points. These points can be considered a "worst case configuration" valid for all mesh indexings, yielding our lower bounds.

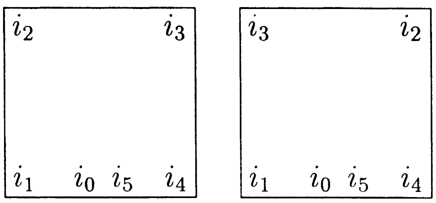
5.1. Euclidean and maximum metric

Theorem 1 of Gotsman and Lindenbaum [10] says that for any discrete 2-D space-filling curve on an $n \times n$-mesh, $d_2(i,j) > \sqrt{3(1-1/n)^2|i-j|}$. They also report that by a computerized exhaustive search they have improved the constant factor 3–3.25. We improve this to 3.5 by a direct proof. In addition, their result is only valid for continuous indexings, whereas poses no restrictions on the indexing. We conjecture that this can be raised to 4, implying the optimality of H-curves among all mesh-indexings (cf. Theorems 1 and 2).

In the following theorem we make use of the general relationship $d_\infty(i,j) \leqslant d_2(i,j)$ by proving only the result for the maximum metric.

Theorem 3. *For each indexing of an $n \times n$-mesh, $n \geqslant 2$, there must be indices i and j with $d_2(i,j), d_\infty(i,j) > n/4$ such that $d_2(i,j), d_\infty(i,j) \geqslant \sqrt{3.5|i-j|} - 1$.*

Proof. Due to $d_2(i,j) \geqslant d_\infty(i,j)$, it suffices to restrict our attention to the maximum metric. The proof is by contradiction. Assume on the contrary that for all i and j with $d_\infty(i,j) > n/4$ we have $d_\infty(i,j) < \sqrt{3.5|i-j|} - 1$, that means $|i-j| > (d_\infty(i,j)+1)^2/3.5$. In the following, we describe something like a "worst case configuration" of some index locations in the mesh. We consider the two cases represented by the two basic pictures below. All other cases are symmetric. Let $i_1 < i_2 < i_3$ and $i_2 < i_4$ be the indices of the 4 corner points of the $n \times n$-mesh. Since we leave the relation between i_3 and i_4 open, the following describes (except for symmetric cases) all possibilities (cf. [10]). Note that the right-hand picture is necessary for the case of non-continuous indexings.



Let i_0 be the rightmost point in the row between i_1 and i_4 with $i_0 < i_2$. Note that $i_0 = i_1$ is possible. The distance of i_0 from i_1 shall be $m-1$. Therefore, the neighboring point i_5 of i_0 with $i_2 < i_5$ has distance $n-m-1$ from i_4. Generally, we have two possible orders of i_0 and i_1 and six possible orders of i_3, i_4 and i_5. Thus, first assuming $n/4 < m < 3n/4$

in order to make subsequent use of our assumption $|i-j| > ((d_2(i,j)+1)^2)/3.5$, we derive the relationship shown below. Observe that the following is valid for both pictures above at the same time.

$$\begin{aligned}
n^2 &\geqslant \min\{|i_0-i_1|+|i_1-i_2|,\ |i_1-i_0|+|i_0-i_2|\}\\
&\quad+\min\{|i_2-i_3|+|i_3-i_4|+|i_4-i_5|,\ |i_2-i_3|+|i_3-i_5|+|i_5-i_4|,\\
&\qquad |i_2-i_5|+|i_5-i_4|+|i_4-i_3|,\ |i_2-i_5|+|i_5-i_3|+|i_3-i_4|,\\
&\qquad |i_2-i_4|+|i_4-i_3|+|i_3-i_5|,\ |i_2-i_4|+|i_4-i_5|+|i_3-i_3|\}\\
&> \frac{1}{3.5}\min\{(d_\infty(i_0,i_1)+1)^2+(d_\infty(i_1,i_2)+1)^2,\\
&\qquad (d_\infty(i_1,i_0)+1)^2+(d_\infty(i_0,i_2)+1)^2\}\\
&\quad+\frac{1}{3.5}\min\{(d_\infty(i_2,i_3)+1)^2+(d_\infty(i_3,i_4)+1)^2+(d_\infty(i_4,i_5)+1)^2,\\
&\qquad (d_\infty(i_2,i_3)+1)^2+(d_\infty(i_3,i_5)+1)^2+(d_\infty(i_5,i_4)+1)^2,\\
&\qquad (d_\infty(i_2,i_5)+1)^2+(d_\infty(i_5,i_4)+1)^2+(d_\infty(i_4,i_3)+1)^2,\\
&\qquad (d_\infty(i_2,i_5)+1)^2+(d_\infty(i_5,i_3)+1)^2+(d_\infty(i_3,i_4)+1)^2,\\
&\qquad (d_\infty(i_2,i_4)+1)^2+(d_\infty(i_4,i_3)+1)^2+(d_\infty(i_3,i_5)+1)^2,\\
&\qquad (d_\infty(i_2,i_4)+1)^2+(d_\infty(i_4,i_5)+1)^2+(d_\infty(i_5,i_3)+1)^2\}\\
&= \frac{1}{3.5}((m^2+n^2)+\min\{2n^2+(n-m)^2, 2n^2+(n-m)^2, n^2+(n-m)^2+n^2,\\
&\qquad 3n^2, 3n^2, n^2+(n-m)^2+n^2\})\\
&= \frac{m^2+3n^2+(n-m)^2}{3.5} = \frac{2m^2+4n^2-2nm}{3.5} = \frac{3.5n^2+2(n/2-m)^2}{3.5}.
\end{aligned}$$

This is a contradiction.

Now, turning to the case $m \leqslant n/4$, we do not use i_0 as a candidate point and a similar calculation as above yields

$$n^2 \geqslant \frac{3n^2+(n-m)^2}{3.5} \geqslant \frac{3n^2+(3n/4)^2}{3.5} = \frac{3.5625n^2}{3.5}$$

a contradiction. Analogously, if $m \geqslant 3n/4$, by eliminating i_5 we get

$$n^2 \geqslant \frac{m^2+3n^2}{3.5} \geqslant \frac{3n^2+(3n/4)^2}{3.5} = \frac{3.5625n^2}{3.5}. \qquad \square$$

Compared to Theorem 3, the lower bound for the special case of cyclic indexings can be obtained comparatively easily. Together with Theorem 1 it shows optimality of H-indexings among all cyclic indexings up to small additive constants.

Theorem 4. *For each cyclic indexing of an $n \times n$-mesh, $n \geqslant 2$, indices i and j must be present, so that $d_2(i,j), d_\infty(i,j) \geqslant 2\sqrt{|i-j|} - 1$. This lower bound specifically applies to the two corners i and j of the mesh.*

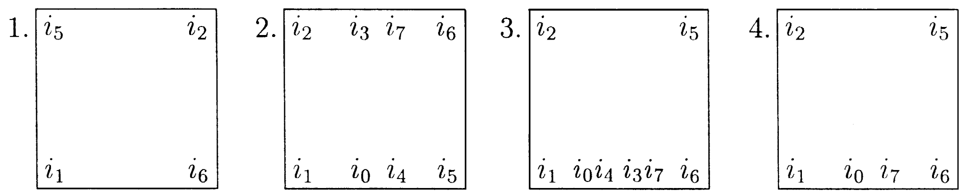
Proof. Let i_1, i_2, i_3, and i_4 be the 4 corner points of an $n \times n$-mesh. Because the indexing is cyclic (and thus also continuous, cf. Section 2) there must be two corner points i_j and i_k with $j,k \in \{1,2,3,4\}$ and $j \neq k$ such that $|i_j - i_k| \leqslant n^2/4$. On the other hand, $d_2(i_j,i_k) \geqslant d_\infty(i_j,i_k) \geqslant n-1 \geqslant 2\sqrt{|i_j - i_k|} - 1$. □

5.2. Manhattan metric

Whereas in the case of the Euclidean and the maximum metric we could give quite close bounds for the "general case", this seems to be more problematic when dealing with the Manhattan metric. In the general case, we obtain the following, comparatively weaker result, based on a more complicated case distinction concerning "worst case configurations" of some index locations (as shown by the subsequent pictures).

Theorem 5. *For each indexing of an $n \times n$-mesh, $n \geqslant 2$, indices i and j must be present with $d_1(i,j) > 2n/5$, so that $d_1(i,j) \geqslant \sqrt{6.5|i-j|} - 2$.*

Proof. Assume the contrary that for all i and j with $d_1(i,j) > 2n/5$ we have $d_1(i,j) < \sqrt{6.5|i-j|} - 2$, making $|i-j| > (d_1(i,j)+2)^2/6.5$. We describe the "worst case configurations" needed for proving our result by the following four pictures. Let $i_1 < i_2 < i_5 < i_6$ be the indices of the 4 corner points of the $n \times n$-mesh the indexing passes through in the given order. Then (except for symmetric cases) we have the following four possibilities. Observe that the first picture comes into play because we also allow non-continuous indexings.



In the second to fourth picture, i_0 is the rightmost point in the row containing i_1 with $i_0 < i_2$ and distance $m-1$ from i_1, and i_7 is the leftmost point in the row containing i_6 with $i_5 < i_7$ and distance $l-1$ from i_6. Moreover, i_3 and i_4 are immediate left-hand and right-hand neighbors of i_7 and i_0, respectively.

1. The case exhibited with the first picture is fairly easy to handle. Needing no further assumptions, we have

$$n^2 \geqslant |i_1 - i_6| = |i_1 - i_2| + |i_2 - i_5| + |i_5 - i_6|$$

$$> \frac{(d_1(i_1,i_2)+2)^2+(d_1(i_2,i_5)+2)^2+(d_1(i_5,i_6)+2)^2}{6.5}$$

$$\geqslant \frac{4n^2+n^2+4n^2}{6.5} = \frac{9n^2}{6.5},$$

a contradiction.

2. In the case referring to the second picture, if $i_4 < i_3$, then we have

$$n^2 \geqslant |i_0-i_7| = |i_0-i_2|+|i_2-i_4|+|i_4-i_3|+|i_3-i_5|+|i_5-i_7|$$

$$> \frac{(d_1(i_0,i_2)+2)^2+(d_1(i_2,i_4)+2)^2+(d_1(i_4,i_3)+2)^2}{6.5}$$

$$+ \frac{(d_1(i_3,i_5)+2)^2+(d_1(i_5,i_7)+2)^2}{6.5}$$

$$\geqslant \frac{(n+m)^2+(n+m)^2+(2n-m-l)^2+(n+l)^2+(n+l)^2}{6.5}$$

$$= \frac{8n^2+3m^2+2ml+3l^2}{6.5} \geqslant \frac{8n^2}{6.5}.$$

If $m+l \geqslant n/2$ then

$$n^2 \geqslant |i_0-i_7| = |i_0-i_2|+|i_2-i_5|+|i_5-i_7|$$

$$> \frac{(n+m)^2+4n^2+(n+l)^2}{6.5} = \frac{6n^2+2(m+l)n+m^2+l^2}{6.5} \geqslant \frac{7n^2}{6.5}$$

otherwise (i.e., $m+l<n/2$ and $i_3<i_4$) we have to distinguish between three sub-cases. First assume that $i_3 < i_1$. Then

$$n^2 \geqslant |i_3-i_5| = |i_3-i_1|+|i_1-i_2|+|i_2-i_5|$$

$$> \frac{(2n-l)^2+n^2+(2n)^2}{6.5} = \frac{9n^2-4ln+l^2}{6.5} \geqslant \frac{7n^2}{6.5}.$$

If $i_4 > i_6$, we get the same for reasons of symmetry.

Finally, if $i_1 < i_3$ and $i_4 < i_6$, then

$$n^2 \geqslant |i_1-i_6| = |i_1-i_3|+|i_3-i_4|+|i_4-i_6|$$

$$> \frac{(2n-l)^2+n^2+(2n-m)^2}{6.5} \geqslant \frac{9n^2-4(m+l)n}{6.5} \geqslant \frac{7n^2}{6.5}.$$

3. With respect to the third picture, we have

$$n^2 \geqslant |i_0-i_7| = |i_0-i_2|+|i_2-i_4|+|i_4-i_5|+|i_5-i_7|$$

$$> \frac{(n+m)^2+(n+m)^2+(2n-m)^2+(n+l)^2}{6.5}$$

$$\geqslant \frac{7n^2+3m^2+2nl+l^2}{6.5} \geqslant \frac{7n^2}{6.5}.$$

4. The last picture differs from the third case in that i_0 and i_7 are now immediate neighbors. In addition, for reasons of symmetry we assume without loss of generality that $m \leqslant n/2$ (otherwise, the roles of i_0 and i_7 will interchange). If $m \leqslant 0.418n$, then
$$n^2 \geqslant |i_1 - i_7| = |i_1 - i_5| + |i_5 - i_7|$$
$$> \frac{(2n)^2 + (1.582n)^2}{6.5} = \frac{(4 + 2.502)n^2}{6.5}.$$
If $i_0 < i_1$, then
$$n^2 \geqslant |i_0 - i_7| = |i_0 - i_1| + |i_1 - i_5| + |i_5 - i_7|$$
$$> \frac{m^2 + (2n)^2 + (2n - m)^2}{6.5} = \frac{8n^2 + 2m^2 - 4nm}{6.5}$$
$$= \frac{6.5n^2 + (n - 2m)^2/2 + (n - 2m)n}{6.5} \geqslant \frac{6.5n^2}{6.5}.$$
If $i_7 < i_6$ then
$$n^2 \geqslant |i_1 - i_6| = |i_1 - i_5| + |i_5 - i_7| + |i_7 - i_6|$$
$$> \frac{(2n)^2 + (1.5n)^2 + (0.5n)^2}{6.5} = \frac{(4 + 2.25 + 0.25)n^2}{6.5}.$$
Otherwise we have $0.418n < m \leqslant n/2$, $i_1 < i_0$, and $i_6 < i_7$. Then
$$n^2 \geqslant |i_1 - i_7| = |i_1 - i_0| + |i_0 - i_2| + |i_2 - i_6| + |i_6 - i_7|$$
$$> \frac{m^2 + (n + m)^2 + (2n)^2 + (n - m)^2}{6.5} = \frac{6n^2 + 3m^2}{6.5} > \frac{6.5n^2}{6.5},$$
again a contradiction.

This completes the proof. □

In the special cyclic case, however, we can again prove (asymptotic) optimality of H-curves due to the following theorem.

Theorem 6. *For each cyclic indexing of an $n \times n$-mesh, $n \geqslant 2$, indices i and j must be present, so that $d_1(i,j) \geqslant \sqrt{8|i-j|} - 2$. This lower bound specifically applies if i and j are in two diagonally opposite corners of the mesh.*

Proof. By definition of a cyclic indexing, $|i - j| \leqslant n^2/2$ for all i and j in an $n \times n$ square. Consequently, we have for two diagonally opposite corners i and j, $d_1(i,j) = 2n - 2 \geqslant 2\sqrt{2|i-j|} - 2 = \sqrt{8|i-j|} - 2$. □

6. Mechanizing proofs for upper bounds

The primary goal of this section is to introduce a technique, whereby it is possible to derive locality properties of self-similar indexings by mechanical inspection. In

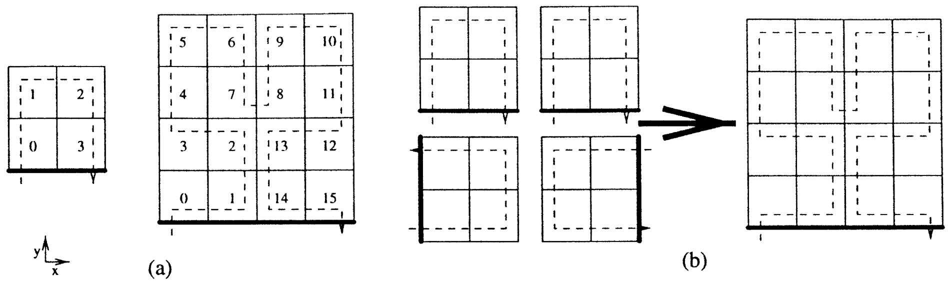


Fig. 7. Hilbert indexings of size 4 and 16 and the general construction principle.

Section 6.1, we start with the well-known 2-D Hilbert indexing and give a more complete proof of the tight bound for the Manhattan distance already found in [7], which does not need tedious manual case distinctions. Then, in Section 6.2, we develop a more widely applicable technique and apply it to other metrics and to 3-D Hilbert indexings.

6.1. The Hilbert indexing

Fig. 7(a) shows the two smallest Hilbert indexings for meshes of size 4 and 16. Fig. 7(b) shows the general construction principle. For any $k \geqslant 1$, four Hilbert indexings of size 4^k are combined into an indexing of size 4^{k+1} by rotating and reflecting them in such a way that concatenating the indexings yields a Hamiltonian path through the mesh. Note that the left-hand and the right-hand side of the curve are symmetrical to each other. Consequently, we need only keep track of the orientation of the edge which contains the start and end of the curve (drawn with bold lines here).[4] We start with a lower bound for the locality:

Theorem 7. *For every $k \geqslant 1$, indices i and j are present on the Hilbert indexing, so that $|i-j| = 4^{k-1}$ and the Manhattan-distance of i and j is exactly $3\sqrt{|i-j|} - 2 = 3 \cdot 2^{k-1} - 2$.*

Proof. Consider Fig. 8. It shows parts of the Hilbert indexing (rotated 90° to the right compared to Fig. 7). It suffices to show that the indices i and j in the lower left and upper right corner of the shaded area of Fig. 8 have Manhattan-distance $3\sqrt{|i-j|} - 2$. We must compute the size of the shaded area which denotes all nodes on the Hilbert indexing lying between i and j. We always draw the largest subsquare filled by the Hilbert indexing on the path from i to j. In this sense, the dotted line represents the path of the Hilbert indexing respective of the sizes of the largest subsquares it passes through. Except for the lower left corner and upper right corner, we have exactly

[4] We note without proof that the above rule uniquely defines the Hilbert indexing up to global rotation and reflection. In a sense, the Hilbert curve is the "simplest" self-similar, recursive, locality-preserving indexing scheme for square meshes of size $2^k \times 2^k$. More details can be found in [2].

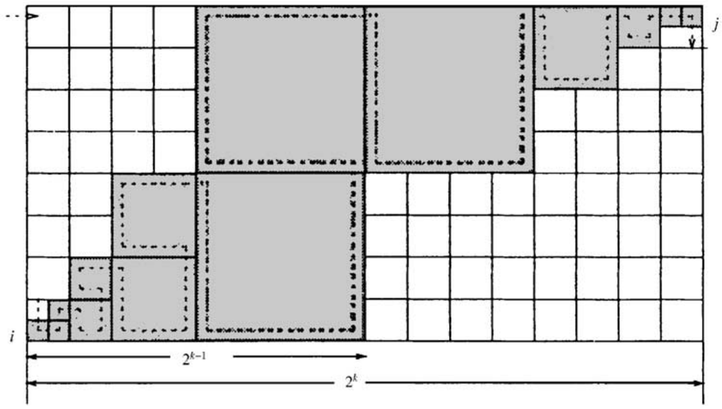


Fig. 8. Worst case for the Manhattan-distance between two indices i and j.

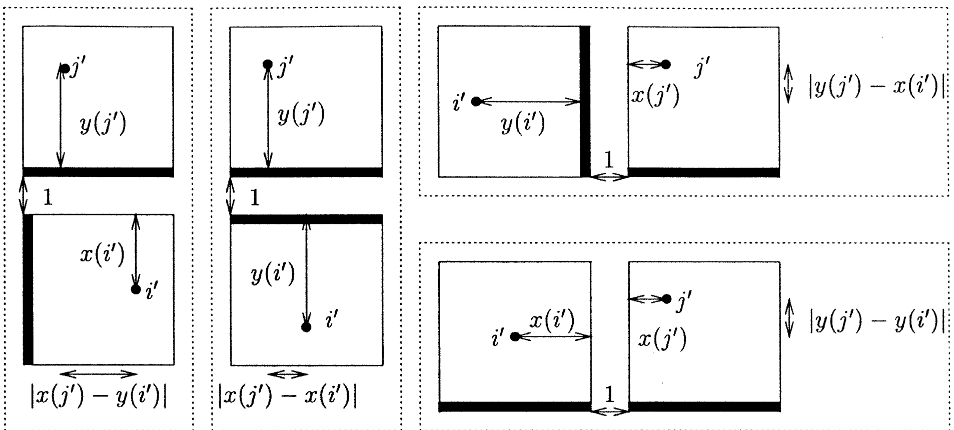


Fig. 9. Possible relative orientations of two Hilbert-squares, where i' corresponds to the term $l-i-1$ in the proof of Lemma 8 and j' corresponds to $j-l$.

three subsquares of size $2^l \times 2^l$ within the shaded area for each $0 \leqslant l < k-1$. As the shaded area of the left half can be mapped onto the unshaded area in the right half of Fig. 8 (except for one mesh node remaining), we get $|i-j| = 4^{k-1}$. Computing $3\sqrt{|i-j|} - 2 = 3 \cdot 2^{k-1} - 2$, we obtain the Manhattan-distance of i and j exactly, where the latter can easily be read from Fig. 8. □

Before we come to the matching upper bound, we need a technical lemma that shows how we can bound $\max_{|i-j|=m} d(i,j)$ for a fixed m by inspecting a finite number of segments. These are those segments of length m which either lie within a single indexing of size $4^{\lceil \log_4 m \rceil}$ or within two such-grids. For the latter case there are four subcases for the four different relative orientations of two subgrids shown in Fig. 9. This method works for an arbitrary norm $||\cdot||$.

Lemma 8. *Let $x(i)$ and $y(i)$ denote the x-coordinate and y-coordinate of the ith point in the Hilbert indexing. Let*

$$d_{\text{int}}(m) := \max\{d(i,j)\colon |i-j| = m \wedge 0 \leqslant i < j < 4^{\lceil \log_4 m \rceil}\} \quad \textit{and}$$

$$d_{\text{ext}}(m) := \max_{i'+j'=m-1} \max \begin{pmatrix} ||(|x(j') - y(i')|, 1 + y(j') + x(i'))|| \\ ||(|x(j') - x(i')|, 1 + y(j') + y(i'))|| \\ ||(1 + x(j') + y(i'), |y(j') - x(i')|)|| \\ ||(1 + x(j') + x(i'), |y(j') - y(i')|)|| \end{pmatrix}.$$

Then $\forall i,j\colon d(i,j) \leqslant \max(d_{\text{int}}(|i-j|), d_{\text{ext}}(|i-j|))$.

Proof. Consider any segment size m and any indices i and j with $|i-j| = m$. W.l.o.g. assume $j > i$ and let $k = \lceil \log_4 m \rceil$.

(1) *Case* $\forall l \in \{i+1,\dots,j\}$: $l \not\equiv 0 \bmod 4^k$: Due to the self-similarity of the Hilbert indexing, the segment $\overline{(i,j)}$ is isomorphic to the segment $\overline{(i \bmod 4^k, j \bmod 4^k)}$. This segment has already been checked by computing $d_{\text{int}}(m)$.

(2) *All other cases*: There is exactly one l with $i < l \leqslant j$ and $l \equiv 0 \bmod 4^k$. Due to the self-similarity and symmetry of the Hilbert indexing, the segments $\overline{(l,j)}$ and $\overline{(i,l-1)}$ are isomorphic to the segments $\overline{(0,i')}$ and $\overline{(0,j')}$, respectively, where $j' = j - l$ and $i' = l - i - 1$. There are only four different ways (disregarding rotation and reflection) the segments $\overline{(l,j)}$ and $\overline{(i,l-1)}$ can be oriented toward each other. Fig. 9 shows the ways in which this is possible. For each of these four cases, a formula describing the distance vector between the points i and j can be derived as follows: In one direction, the distance between the two points is one (the distance between the two subsquares) plus the sum of two coordinates from points i' and j' (using the standard orientation of the Hilbert indexing). In the other direction, the distance is the difference between the other two coordinates of i' and j'. For example, if the subsquares are arranged as in the leftmost part of Fig. 9, we have to add one, $y(j')$, and $x(i')$ in order to get the distance in the x-direction while the distance in the y-direction is $|x(j') - y(i')|$. The inner maximization for the definition of d_{ext} checks the norms of the four possible distance vectors. The outer maximization covers all possible values for l. □

This result will later be used in its full generality. It should be emphasized here that Lemma 8 can be verified mechanically by a simple computer program. For now, we concentrate on the Manhattan metric:

Theorem 9. *For the Manhattan-distance of two arbitrary indices i and j on the Hilbert indexing with $i \neq j$, we have* $d_1(i,j) \leqslant 3\sqrt{|i-j|} - 2$.

Proof. The fundamental goal here is to exploit the self-similarity of the Hilbert indexing for an inductive proof over $|i-j|$. In principle, the proof is quite simple. However, it proves to be the case that a special treatment is necessary for "small" meshes and for indices i and j which are close to the worst case described in Theorem 7.

(1) *Case* $|i-j|<16$: Apply Lemma 8 for $|i-j| \in \{1,\dots,15\}$.

(2) *Case* $|i-j| \geqslant 16$: By induction over $|i-j|$ we prove the following stronger statement: $d_1(i,j) \leqslant 3\sqrt{|i-j|}-2.5$ or i and j are arranged as in Theorem 7 (Fig. 8) and $d_1(i,j)=3\sqrt{|i-j|}-2$.

(2.1) *Basis of induction*, $16 \leqslant |i-j| \leqslant 80$: Apply Lemma 8 for $|i-j| \in \{16,\dots,80\}$. Note that this can be done mechanically by a simple computer program.

(2.2) *Inductive step for* $|i-j|>80$: We look at the "coarsened" indexing defined by considering each 2×2 subsquare starting at even coordinates as a single mesh node. Due to the self-similarity of the Hilbert indexing, the coarsened indexing is itself a Hilbert indexing.

Define $a \in \mathbb{N}$, $b \in \{0,1,2,3\}$, $c \in \mathbb{N}$ and $d \in \{0,1,2,3\}$, so that $i=4a+b$ and $j=4c+d$. In the coarsened indexing, the positions of i and j are a and c, respectively. Since $|a-c| \geqslant 16$, we can apply the induction hypothesis. Furthermore, $d_1(i,j) \leqslant 2d_1(a,c)+2$ because for each of the four mesh-positions in subsquare a there is a corresponding mesh-position in subsquare c which is $2d_1(a,c)$ steps away; at worst j can be another two steps away from the mesh-position corresponding to i. We now distinguish two cases regarding the relative positions of a and c.

(2.2.1) *a and c are* not *arranged as in Theorem* 7: By the induction hypothesis we have $d_1(a,c) \leqslant 3\sqrt{|a-c|}-2.5$ and therefore

$$d_1(i,j) \leqslant 2(3\sqrt{|a-c|}-2.5)+2 = 6\sqrt{|a-c|}-3.$$

Substituting $a=(i-b)/4$ and $c=(j-d)/4$ we get

$$|a-c| = \frac{|(i-b)-(j-d)|}{4} \leqslant \frac{|i-j|+|d-b|}{4} \leqslant \frac{|i-j|+3}{4}$$

and therefore $d_1(i,j) \leqslant 3\sqrt{|i-j|+3}-3$. A simple calculation shows that $3\sqrt{|i-j|+3} \leqslant 3\sqrt{|i-j|}+0.5$ for $|i-j| \geqslant 80$ and therefore $d_1(i,j) \leqslant 3\sqrt{|i-j|}-2.5$.

(2.2.2) *a and c* are *arranged as in Theorem* 7: With the exception of symmetrical cases the 2×2-subsquares for i and j are numbered $\binom{0\,3}{1\,2}$ and $\binom{0\,1}{3\,2}$ and the subsquare for j is above and to the right of the subsquare for i (refer to Fig. 8). There are two subcases:

(2.2.2.a) $b=d=1$: i and j are also arranged as in Theorem 7 and we get

$$d_1(i,j) = 2(3\sqrt{|a-c|}-2)+2 = 3\sqrt{4\left|\frac{i-1}{4}-\frac{j-1}{4}\right|}-2 = 3\sqrt{|i-j|}-2$$

as desired.

(2.2.2.b) *Else*: We can use the estimate $d_1(i,j) \leqslant 2d_1(a,c)+1$ because the worst case, in which $d_1(i,j)=2d_1(a,c)+2$, has already occurred in the case $b=d=1$. A calculation similar to the previous shows that

$$d_1(i,j) \leqslant 2(3\sqrt{|a-c|}-2)+1 = 6\sqrt{|a-c|}-3 \leqslant 3\sqrt{|i-j|}-2.5. \qquad \square$$

6.2. A generalized technique and its applications

There are few instances where the proof of Theorem 9 makes explicit use of the properties of the Hilbert indexing or the Manhattan metric. We now offer a generalized technique which can be applied to a wide spectrum of self-similar indexings in r-dimensional meshes made up of building blocks of size $q_1,\dots,q_r$ and a norm $||\cdot||$. However, for simplicity we restrict the presentation to cubic building blocks with side-length q and only show how slightly looser upper bounds than those of Theorem 9 can be proved. The latter relaxation allows us to avoid the special treatment of the worst case segments which is necessary in the proof of Theorem 9.

Theorem 10. *Given any indexing scheme for r-dimensional meshes with the property that combining each elementary cube of size q^r from a mesh of size q^{kr} into a single meta-node yields the indexing for a mesh of size $q^{(k-1)r}$:*
If $\forall q^{(k-1)r} \leqslant |i-j| \leqslant q^{kr} : d(i,j) \leqslant \alpha(\sqrt[r]{|i-j|}-\delta)-\beta$
where $\beta := ||(1,\dots,1)||$ and $\delta \geqslant \frac{\sqrt[r]{q^{kr}+q^r-1}-q^k}{q-1}$
then $\forall |i-j| \geqslant q^{(k-1)r} : d(i,j) \leqslant \alpha(\sqrt[r]{|i-j|}-\delta)-\beta$.

The proof of Theorem 10 is quite analogous to the Proof of Theorem 9.

Proof. By induction over $|i-j|$. Let $a=\lfloor i/q^r \rfloor$, $b=i \bmod q^r$, $c=\lfloor j/q^r \rfloor$, and $d=j \bmod q^r$. Due to the self-similarity of the indexing scheme, we can apply the induction hypothesis to a and c if $|i-j| \geqslant q^{kr}$. We find $d(i,j) \leqslant q \cdot d(a,c)+\beta(q-1)$ because for each of the q^r mesh-positions in subcube a there is a corresponding mesh-position in subcube c which is $q \cdot d(a,c)$ steps away; at worst j can be another $\beta(q-1)$ steps away from the mesh-position corresponding to i (the diameter of a cube of side length q). Using the induction hypothesis, we have $d(a,c) \leqslant \alpha(\sqrt[r]{|a-c|}-\delta)-\beta$ and therefore

$$d(i,j) \leqslant q(\alpha(\sqrt[r]{|a-c|}-\delta)-\beta)+\beta(q-1) = q \cdot \alpha(\sqrt[r]{|a-c|}-\delta)-\beta.$$

Substituting $a=(i-b)/q^r$ and $c=(j-d)/q^r$ we get

$$|a-c| = \frac{|(i-b)-(j-d)|}{q^r} \leqslant \frac{|i-j|+|d-b|}{q^r} \leqslant \frac{|i-j|+q^r-1}{q^r}$$

and therefore $d(i,j) \leqslant \alpha(\sqrt[r]{|i-j|+q^r-1}-q\delta)-\beta$. A simple calculation shows that $\sqrt[r]{|i-j|+q^r-1}-q\delta \leqslant \sqrt[r]{|i-j|}-\delta$ for $|i-j| \geqslant q^{kr}$ and $\delta \geqslant (\sqrt[r]{q^{kr}+q^r-1}-q^k)/(q-1)$. □

Theorem 10 can be applied so that it yields upper bounds for $d(i,j)$. However, the additive constant δ and – except for the Manhattan metric – the additive constant β are artifacts of the inductive proof. If we do not want to make case distinctions involving special properties of worst case segments as in the proof of Theorem 9, we have to accept a small increase in the multiplicative factor α which compensates for the additive constants if $|i-j|$ is large. The case of small $|i-j|$ can be resolved mechanically. Consider the following procedure for obtaining bounds of the form $d(i,j) \leqslant \alpha\sqrt[r]{|i-j|}+c$ where c is some constant to be determined.

- Determine q and r from the definition of the indexing.
- Fix a value k for the mesh size to be inspected.

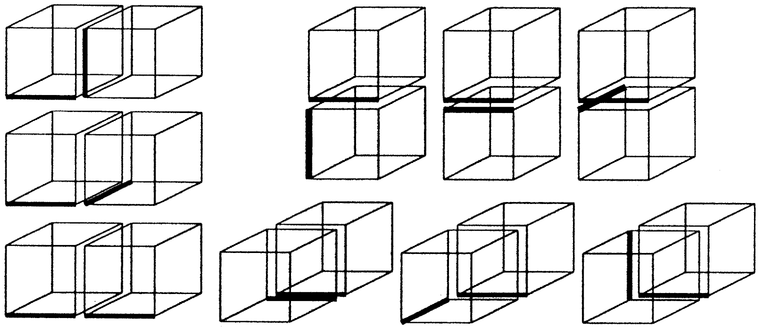


Fig. 10. Relative cube orientations to be checked for bounding maximum distances for a given segment size.

- Set $\delta = (\sqrt[r]{q^{kr} + q^r - 1} - q^k)/(q-1)$ and $\beta := ||(1,\dots,1)||$.
- Make use of the self-similarity of the indexing to find an analog to Lemma 8 which makes it possible to bound $d(i,j)$ for indices with $|i-j| = m$ using some mechanizable method. For example, Fig. 10 shows which relative cube orientations have to be checked for three dimensions.
- Find a constant α, so that $d(i,j) \leqslant \alpha(\sqrt[r]{|i-j|} - \delta) - \beta$ for $q^{(k-1)r} \leqslant |i-j| \leqslant q^{kr}$ where δ and β are defined as in Theorem 10. Applying Theorem 10 we can infer that the same is true for $|i-j| \geqslant q^{kr}$, i.e. $\forall |i-j| \geqslant q^{(k-1)r} : d(i,j) \leqslant \alpha(\sqrt[r]{|i-j|} - \delta) - \beta \leqslant \alpha\sqrt[r]{|i-j|} - \beta$.
- Find a constant $c \geqslant -\beta$ such that $d(i,j) \leqslant \alpha\sqrt[r]{|i-j|} + c$ for $|i-j| \leqslant q^{(k-1)r}$.
- We can now conclude from the two points above that *for all* i,j, $d(i,j) \leqslant \alpha\sqrt[r]{|i-j|} + c$.

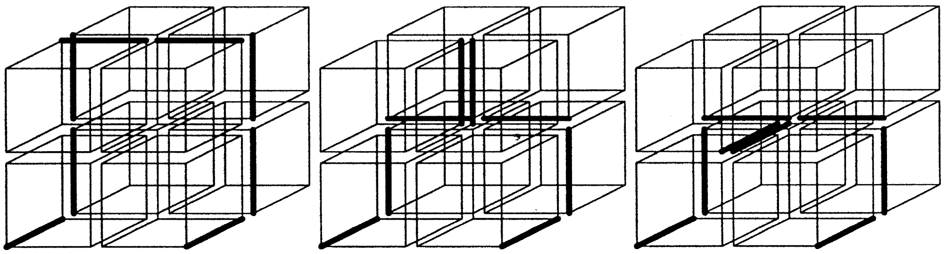
In the following, we will simply use $c = 0$ (which will always suffice) in order to indicate that the additive constants are not tight. Also, we will only cite the tightest constant factor for an upper bound as given by our method without repeating the point that the constructive nature of the method also yields a lower bound with a close-by constant factor.

6.2.1. 2-D Hilbert indexings

Using the above method and by applying a small computer program[5] to the case $k=8$, we can infer a bound for the Euclid metric of $d_2(i,j) \leqslant \sqrt{6+0.01}\sqrt{|i-j|}$, which is very close to the lower bound of $\sqrt{6|i-j|-2} - 1$ according to Gotsman and Lindenbaum [10]. A significant improvement of the upper bound $d_2(i,j) \leqslant \sqrt{6+\frac{2}{3}}\sqrt{|i-j|}$ is derived in the same paper.

Trivially, the same bound also applies to the maximum metric for which Gotsman and Lindenbaum reported the same constant factors of $\sqrt{6}$ and $\sqrt{6+\frac{2}{3}}$ for lower and upper bounds, respectively.

[5] Available under `http://www.mpi-sb.mpg.de/~sanders/programs/hilbert/euclid2.c`.



(a) (b) (c)

Fig. 11. Rule for building 3-D Hilbert indexings of order k from indexings of order $k-1$. The bottom front edge of the new cube is distinguished by the fact that the indexing starts and ends there. The corresponding edges of the component cubes are drawn with thick lines. The order $k-1$ cubes have to be rotated accordingly.

6.2.2. Symmetric 3-D Hilbert indexings

We have also applied the above technique to the three variants of a 3-D Hilbert indexing shown in Fig. 11. Up to rotation and reflections, these are the only variants which are symmetrical with respect to an axis. The maximum segment distances can be checked in a complete analogy of Lemma 8: Now nine relative orientations are to be checked. [6]

Applying the "method" for variants (b) and (c) with $k=5$ yields $d_1(i,j) \leqslant 4.820661 \sqrt[3]{|i-j|}$ and the systematic search discovers indices with $d_1(i,j) \geqslant 4.820248 \sqrt[3]{|i-j|}$. Variant (a) has a slightly better locality: $d_1(i,j) \leqslant 4.6161 \sqrt[3]{|i-j|} - 3$ for large $|i-j|$, which also applies for small $|i-j|$ using a slightly looser additive constant. In comparison, the best previous bound has the constant factor $8/\sqrt[3]{4} \approx 5.04$ [6].

Variant (a) is also slightly superior using the Euclidean metric, where we get $d_2(i,j) \leqslant 3.212991 \sqrt[3]{|i-j|}$ for variant (a) and $d_2(i,j) \leqslant 3.245222 \sqrt[3]{|i-j|}$ for variants (b) and (c) when we apply a simple program [7] for $k=4$. As opposed to the 2-D case, the maximum metric allows smaller bounds than the Euclidean metric in the 3-D case. We get $d_\infty(i,j) \leqslant 3.076598 \sqrt[3]{|i-j|}$ for variant (a) and $d_\infty(i,j) \leqslant 3.104403 \sqrt[3]{|i-j|}$ for variants (b) and (c). [8]

The method could also be applied to the asymmetrical variants of the Hilbert indexing described in [6]. We only have to change the procedure for checking maximum segment sizes in order to take segments starting at both ends of a cube indexing into account. Even generalizations to more complicated schemes, like the H^* indexing described in [6], seem possible. (This scheme appears to have a better locality than simple Hilbert indexings.) H^* uses two non-isomorphic building blocks to define larger indexings.

[6] A C-program doing the necessary checks is available under `http://www.mpi-sb.mpg.de/~sanders/programs/hilbert/check3d.c`.

[7] Available under `http://www.mpi-sb.mpg.de/~sanders/programs/hilbert/euclid3d.c`.

[8] The program is available under `http://www.mpi-sb.mpg.de/~sanders/programs/hilbert/max3d.c`.

Table 1

$d(i,j)$ (2-D)	Euclidean	Maximum	Manhattan
General lower bound	$\sqrt{3.5\|i-j\|}-1$	$\sqrt{3.5\|i-j\|}-1$	$\sqrt{6.5\|i-j\|}-2$
Cyclic lower bound	$\sqrt{4\|i-j\|}-1$	$\sqrt{4\|i-j\|}-1$	$\sqrt{8\|i-j\|}-2$
Upper bd. H-curve	$\sqrt{4\|i-j\|-2}$	$\sqrt{4\|i-j\|+4}-1$	$\sqrt{8(\|i-j\|-2)}$
Upper bd. 2-D Hilbert	$\sqrt{6.01\|i-j\|}$	$\sqrt{6.01\|i-j\|}$	$\sqrt{9\|i-j\|}-2$
Upper bd. Peano-curve	$\sqrt{8\|i-j\|}$	$\sqrt{8\|i-j\|}$	$\sqrt{(10.66\|i-j\|)}$
Upper bd. Peano-curve 2	$\sqrt{6.25\|i-j\|}$	$\sqrt{5.625\|i-j\|}$	$\sqrt{(10\|i-j\|)}$

But it still has the crucial property that the replacement of a $2 \times 2 \times 2$ cube with a unit cube yields an instance of the indexing.

7. Conclusion

Locality-preserving indexing schemes are increasingly becoming a standard technique by which to devise simple and efficient algorithms for mesh-connected computers, processing geometric data, image processing, data structures, and several other fields. The methods developed here help to use the term "locality-preserving" in an accurate quantitative sense. This makes it possible to show that for the most important 2-D case, the newly presented H-indexing is superior with respect to locality compared with the previously used Hilbert indexing. We conjecture that H-indexings are actually optimal among all possible indexing schemes, although we could only prove this for cyclic indexings thus far. This applies to the Euclidean as well as the maximum and the Manhattan metrics.

Our techniques for mechanically deriving upper bounds make it possible to quickly gain insight into the locality properties of indexing schemes. In particular, it was possible to give new, almost tight bounds for the 2-D Hilbert indexing with respect to the Euclidean metric and the maximum metric and also for the symmetric 3-D Hilbert indexings. In the following Table 1, we summarize our locality bounds for 2-D indexings and also include the results from [18] for Peano indexings, where it is remarkable that a variant of the Peano indexing yields better results than the Hilbert indexing in maximum metric:

With the advent of 3-D mesh-connected computers, such as the Cray T3E, the increasing interest in processing 3-D geometrical data and the growing importance of multidimensional data structures means that locality-preserving 3-D mesh indexings will become more important. [9] The following Table 2 summarizes locality bounds for 3-D indexings. The rather technical proofs of these results are contained in the technical report [22] corresponding to this paper. In particular, the table provides upper

[9] On modern parallel machines, good locality has mainly the indirect effect of increasing the usable bandwidth whereas the latency due to the distance in the network is negligible compared to other overheads. So is would also be interesting to study bandwidth directly.

Table 2

$d(i,j)$ (3-D)	Euclidean	Maximum	Manhattan
General lower bound	$\sqrt[3]{11.1\lvert i-j\rvert}-\sqrt{3}$	$\sqrt[3]{8.25\lvert i-j\rvert}-1$	$\sqrt[3]{42.625\lvert i-j\rvert}-3$
	$\approx 2.23\sqrt[3]{\lvert i-j\rvert}-\sqrt{3}$	$\approx 2.02\sqrt[3]{\lvert i-j\rvert}-1$	$\approx 3.49\sqrt[3]{\lvert i-j\rvert}-3$
Cyclic lower bound	$\sqrt[3]{12.39\lvert i-j\rvert}-\sqrt{3}$	$\sqrt[3]{9\lvert i-j\rvert}-1$	$\sqrt[3]{54\lvert i-j\rvert}-3$
	$\approx 2.31\sqrt[3]{\lvert i-j\rvert}-\sqrt{3}$	$\approx 2.08\sqrt[3]{\lvert i-j\rvert}-1$	$\approx 3.77\sqrt[3]{\lvert i-j\rvert}-3$
U. bd. 3-D Hil. (a)	$\sqrt[3]{33.2\lvert i-j\rvert}$	$\sqrt[3]{29.2\lvert i-j\rvert}$	$\sqrt[3]{98.4\lvert i-j\rvert}$
	$\approx 3.22\sqrt[3]{\lvert i-j\rvert}$	$\approx 3.08\sqrt[3]{\lvert i-j\rvert}$	$\approx 4.62\sqrt[3]{\lvert i-j\rvert}$
U. bd. 3-D Hil. (b,c)	$\sqrt[3]{34.2\lvert i-j\rvert}$	$\sqrt[3]{30.0\lvert i-j\rvert}$	$\sqrt[3]{112.1\lvert i-j\rvert}$
	$\approx 3.25\sqrt[3]{\lvert i-j\rvert}$	$\approx 3.11\sqrt[3]{\lvert i-j\rvert}$	$\approx 4.83\sqrt[3]{\lvert i-j\rvert}$

bounds for some symmetric 3-D variants of the Hilbert indexing. Note that here we still have a significant gap between upper and lower bounds.

Future work

There is a number of interesting open questions. One of these is to close the gap between the upper and lower bound for non-cyclic 2-D indexings and, in particular, for 3-D indexings.

Mechanical inspection methods will play an important role in investigating other indexings in particular for higher dimensions and for more complicated construction rules. The inspection methods themselves can be refined in various ways. They can be adapted to indexing schemes which are not based on combining cubic elements if we use a top-down decomposition rather than a bottom-up decomposition. For example, for some constant k', an H-indexing of size $2^k \times 2^k$ could be partitioned into $2 \cdot 4^{k'}$ triangles of area $2^{k-k'-1}$ without fixing k. The construction principle for the H-curve then defines a (cyclic) path traversing all the triangles. Thus, a computer can count the number of triangles on the (shortest) H-path between any two triangles. The algorithm can also be made faster by adaptively refining only those segments where computations for small k' could not rule out high diameter segments.

Initial work concerning the study of structural and combinatorial properties of Hilbert indexings in higher dimensions has recently begun [2]. In particular, it is clearly pointed out what characterizes an r-dimensional Hilbert curve for arbitrary $r \geqslant 2$.

Acknowledgements

We would like to thank Henning Fernau for valuable pointers regarding literature and Jochen Alber for helpful remarks concerning the notion of cyclic indexings and his clear presentation of some proofs using the maximum metric [1]. In addition, we are grateful to the anonymous referees for their insightful remarks, which have helped improve the presentation.

References

[1] J. Alber, Locality properties of discrete space-filling curves: results with relevance for computer science. Studienarbeit, Universität Tübingen, Wilhelm-Schickard-Institut für Informatik, July 1997 (in German). Available through http://www-fs.informatik.uni-tuebingen.de/ ∼alber/publications.html.

[2] J. Alber, R. Niedermeier, On Multi-Dimensional Hilbert property, Theory Comput. Systems 33 (2000) 295–312.

[3] T. Asano, D. Ranjan, T. Roos, E. Welzl, P. Widmayer, Space-filling curves and their use in the design of geometric data structures, Theoret. Comput. Sci. 181(1997) 3–15.

[4] E. Bugnion, T. Roos, R. Wattenhofer, P. Widmayer, Space filling curves versus random walks, In: M.v. Krefeld et al. (Eds.), Algorithmic Foundations of GIS, Lecture Notes in Computer Science, Vol. 1340, Springer, Berlin, 1997, pp. 199–211.

[5] H. Buhrmann, T. Jiang, M. Li, P. Vitányi, New applications of the incompressibility method (extended abstract), in: Proceedings of the 26th International Conference on Automata, Languages, and Programming, Lecture Notes in Computer Science, Vol. 1644, Springer, Berlin, July 1999, pp. 220–229.

[6] G. Chochia, M. Cole, Recursive 3D mesh indexing with improved locality, Technical Report, University of Edinburgh, 1997. Short version appeared in the Proceedings of HPCN'97, Lecture Notes in Computer Science, Vol. 1225, Springer, Berlin.

[7] G. Chochia, M. Cole, T. Heywood, Implementing the hierarchical PRAM on the 2D mesh: analyses and experiments, in: Symposium on Parallel and Distributed Processing, IEEE Computer Science Press, Silver Spring, MD, October 1995, pp. 587–595.

[8] A. Dingle, I.H. Sudborough, Efficient mappings of pyramid networks, IEEE Trans. Parallel Distributed Systems 5 (10) (1994) 1009–1017.

[9] C. Faloutsos, S. Roseman, Fractals for secondary key retrieval, Proceedings of the Eighth ACM Symposium on Principles of Database Systems, 1989, pp. 247–252.

[10] C. Gotsman, M. Lindenbaum, On the metric properties of discrete space-filling curves, IEEE Trans. Image Process. 5 (5) (1996) 794–797.

[11] D. Hilbert, Über die stetige Abbildung einer Linie auf ein Flächenstück, Math. Ann. 38 (1891) 459–460.

[12] H.V. Jagadish, Linear clustering of objects with multiple attributes, ACM SIGMOD Rec. 19 (2) (1990) 332–342.

[13] H.V. Jagadish, Analysis of the Hilbert curve for representing two-dimensional space, Inform. Process. Lett. 62 (1997) 17–22.

[14] M. Kaddoura, C.-W. Ou, S. Ranka, Partitioning unstructured computational graphs for nonuniform and adaptive environments, IEEE Parallel Distributed Technol.: Systems Appl. 3 (3) (1995) 63–69.

[15] C. Kaklamanis, G. Persiano, Branch-and-bound and backtrack search on mesh-connected arrays of processors, Math. Systems Theory 27 (1994) 471–489.

[16] M. Kunde, R. Niedermeier, K. Reinhardt, P. Rossmanith, Optimal deterministic sorting and routing on grids and tori with diagonals, Algorithmica 25 (1999) 438–458.

[17] M. Kunde, R. Niedermeier, P. Rossmanith, Faster sorting and routing on grids with diagonals, in: P. Enjalbert, E.W. Mayr, K.W. Wagner (Eds.), Proceedings of the 11th Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Computer Science, Vol. 775, Springer, Berlin, 1994, pp. 225–236.

[18] U.v. Luxburg, Locality measures of Peano curves, Studienarbeit, Universität Tübingen, Wilhelm-Schickard-Institut für Informatik, July 1998 (in German).

[19] R. Miller, Q.F. Stout, Mesh computer algorithms for computational geometry, IEEE Trans. Comput. 38 (3) (1989) 321–340.

[20] G. Mitchison, R. Durbin, Optimal numberings of an $N \times N$ array, SIAM J. Algebra Discrete Methods 7 (4) (1986) 571–582.

[21] R. Niedermeier, Towards realistic and simple models of parallel computation, Ph.D. Thesis, Universität Tübingen, July 1996.

[22] R. Niedermeier, K. Reinhardt, P. Sanders, Towards optimal locality in mesh-indexings, Technical Report IB12/97, Fakultät für Informatik, Universität Karlsruhe, Federal Report of Germany, September 1997.

[23] A. Pérez, S. Kamata, E. Kawaguchi, Peano scanning of arbitrary size images, International Conference on Pattern Recognition, 1992, pp. 565–568.

[24] H. Sagan, Space-Filling Curves, Universitext, Springer, Berlin, 1994.
[25] P. Sanders, T. Hansch, Efficient massively parallel quicksort, Proceedings of the Fourth International Symposium on Solving Irregularly Structured Problems in Parallel, Lecture Notes in Computer Science, Vol. 1253, Springer, Berlin, June 1997, pp. 13–24.
[26] J.P. Singh, Parallel hierarchical N-body methods and their implications for multiprocessors, Ph.D. Thesis, Stanford University, 1993.
[27] Q.F. Stout, Topological matching, Proceedings of the 15th ACM Symposium on Theory of Computing, 1983, pp. 24–31.