Improved Algorithms for Optimal Embeddings

NISHANTH CHANDRAN, RYAN MORIARTY, RAFAIL OSTROVSKY, AND OMKANT PANDEY

University of California, Los Angeles

MOHAMMAD ALI SAFARI

University of Alberta

AND

AMIT SAHAI

University of California, Los Angeles

Abstract. In the last decade, the notion of metric embeddings with small distortion has received wide attention in the literature, with applications in combinatorial optimization, discrete mathematics, and bio-informatics. The notion of embedding is, given two metric spaces on the same number of points, to find a bijection that minimizes maximum Lipschitz and bi-Lipschitz constants. One reason for the popularity of the notion is that algorithms designed for one metric space can be applied to a different one, given an embedding with small distortion. The better distortion, the better the effectiveness of the original algorithm applied to a new metric space.

The goal recently studied by Kenyon et al. [2004] is to consider all possible embeddings between two *finite* metric spaces and to find the best possible one; that is, consider a single objective function over the space of all possible embeddings that minimizes the distortion. In this article we continue this important direction. In particular, using a theorem of Albert and Atkinson [2005], we are able

45

This research of N. Chandran, R. Moriarty, O. Pandey, and A. Sahai was supported in part by NSF ITR and Cybertrust programs (including grants 0627781, 0456717, 0716389, and 0205594), a subgrant from SRI as part of the Army Cyber-TA program, an equipment grant from Intel, an Okawa Research Award, and an Alfred P. Sloan Foundation Research Fellowship. The research of R. Ostrovsky was supported in part by an IBM Faculty Award, a Xerox Innovation Group Award, NSF grants 0430254, 0716835, and 0716389 and by a U.C. MICRO grant.

Authors' addresses: N. Chandran, R. Moriarty, Department of Computer Science, University of California, Los Angeles, CA 90095, e-mail: {nishanth,ryan}@cs.ucla.edu; R. Ostrovsky, Department of Computer Science and Mathematics, University of California, Los Angeles, CA 90095, e-mail: {rafail}@cs.ucla.edu; O. Pandey, Department of Computer Science, University of California, Los Angeles, CA 90095, e-mail: omkant@cs.ucla.edu; M. Safari, Department of Computer Science, University of Alberta, Edmonton, Alberta, Canada T6G 2R3, Canada, e-mail: msafarig@cs.ualberta.ca; A. Sahai, Department of Computer Science, University of California, Los Angeles, CA 90095, e-mail: sahai@cs.ucla.edu.

to provide an algorithm to find the optimal bijection between two line metrics, provided that the optimal distortion is smaller than 13.602. This improves the previous bound of $3 + 2\sqrt{2}$, solving an open question posed by Kenyon et al. [2004]. Further, we show an inherent limitation of algorithms using the "forbidden pattern" based dynamic programming approach, in that they cannot find optimal mapping if the optimal distortion is more than $7 + 4\sqrt{3} (\simeq 13.928)$. Thus, our results are almost optimal for this method. We also show that previous techniques for general embeddings apply to a (slightly) more general class of metrics.

Categories and Subject Descriptors: F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*Computations on discrete structures, geometrical problems and computations, pattern matching*

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Optimal metric embeddings, line embeddings, forbidden patterns, metric spaces, shape matching, dynamic programming

ACM Reference Format:

Chandran, N., Moriarty, R., Ostrovsky, R., Pandey, O., Safari, M. A., and Sahai, A. 2008. Improved algorithms for optimal embeddings. ACM Trans. Algor. 4, 4, Article 45 (August 2008), 14 pages. DOI = 10.1145/1383369.1383376 http://doi.acm.org/10.1145/1383369.1383376

1. *Introduction*

For a bijection $\sigma : U \to V$ between two n-point metric spaces (U, d) and (V, d'), the expansion of σ is defined as

$$\text{expansion}(\sigma) = \max_{x,y \in U, x \neq y} \frac{d'(\sigma(x), \sigma(y))}{d(x, y)}.$$

The distortion σ is defined as follows: $\text{dist}(\sigma) = \text{expansion}(\sigma) \times \text{expansion}(\sigma^{-1})$. The *minimum distortion problem* is to find a bijection σ between two equal-sized finite metric spaces (U, d) and (V, d') such that $\text{dist}(\sigma)$ is minimum over all possible bijections.

The minimum distortion problem is interesting to study for both theoretical as well as practical reasons. From a complexity-theoretic point-of-view, it has interesting connections to graph isomorphism [Fortin 1996]. In particular, graph isomorphism on two input graphs G and H is trivially reduced to deciding if there exists an isometric (i.e., distortion 1) bijection between $\mathcal{M}_G$ and $\mathcal{M}_H$, where $\mathcal{M}_X$ denotes the shortest-path metrics of a graph X.

On the practical side, we note that applications dealing with shape matching and object recognition (e.g., signature matching, character recognition, matching facial features, pattern matching in complicated protein structures, etc.) require good measures of similarity. Distortion is an attractive measure of similarity between two point sets [Akutsu et al. 2003; Hoffmann et al. 1998; Belongie et al. 2002; Chazelle et al. 2003]. From the point-of-view of the aforementioned applications, good algorithms for finding minimum distortion bijection (or optimal bijection) are highly desirable.

Kenyon et al. [2004] show that the minimum distortion problem is NP-hard even to approximate (within a factor of 2), and provide two positive results, described next.

—A polynomial-time algorithm is given for *exactly* finding the minimum distortion bijection between two line metrics if the optimal bijection has distortion strictly less than $3 + 2\sqrt{2}$.

—A parameterized polynomial-time algorithm is provided for exactly finding the optimal bijection between a bounded-degree tree metric and an arbitrary unweighted graph metric.

In this article, we improve and generalize the results of Kenyon et al. [2004]. We detail these improvements as follows.

—In particular, we first provide a polynomial-time algorithm for exactly finding an optimal bijection between two line metrics if the optimal bijection has distortion strictly less than 13.602.

To achieve this improvement, we take a more general approach. In particular, Kenyon et al. [2004] look at a single pattern (partial bijection of size 4) and its inverse. They call this pattern a "forbidden pattern". The presence of such patterns guarantees high distortion $(3 + 2\sqrt{2})$. We generalize this approach and look for patterns of higher sizes whose presence will guarantee even higher distortion. We call these patterns *nonseparable permutations*. Absence of such permutations guarantees that the dynamic programming approach can be applied to find the optimal bijection/permutation.

We note that our nonseparable permutations are actually either the *simple* or the *exceptional* permutations as defined by Albert and Atkinson [2005]. This allows us to use a direct theorem of theirs to conclude that the minimum distortion of families of nonseparable permutations increases with size.

—Next, based on the idea of families of nonseparable permutations, we are able to design a dynamic programming algorithm which finds a minimum distortion bijection on more instances than given by Kenyon et al. [2004]. Thus our work answers a direct open question posed in Kenyon et al. [2004].

—We also show a limitation of the forbidden pattern approach, by showing that there exists arbitrarily large families of forbidden patterns with bounded distortion. This lower bound shows the extent to which this approach will be useful and indicates that a new approach must be taken to pass this bound.

We remark that, after the work of Kenyon et al. [2004], most research has focused on either approximating the distortion [Badoiu et al. 2005a, 2005b] or on proving the hardness of approximating it [Hall and Papadimitriou 2005; Papadimitriou and Safra 2005]. Hall and Papadimitriou [2005] show that line embeddings are hard to approximate, even within large factors, when the distortion is high. To the best of our knowledge, there have been no positive results for polynomial time algorithms that *exactly* find the minimum distortion bijection; our results are the first such improvement.

We have recently learned that, independently and concurrent to our work, Kenyon et al. also extended their result on the embedding between two line metrics, using an approach similar to ours. Their algorithm can now handle those cases where the distortion is less than $5 + 2\sqrt{6} \simeq 9.90$.

We also consider the case of embedding a bounded-degree unweighted tree metric into an arbitrary unweighted graph metric. We prove that the algorithm of Kenyon et al. [2004] actually works for a larger class of graphs: unweighted bounded-degree graphs with maximum cycle-length 3. In other words, we show that their algorithm finds an optimal bijection between a bounded-degree graph with maximum cycle-length 3 and an arbitrary unweighted graph metric.

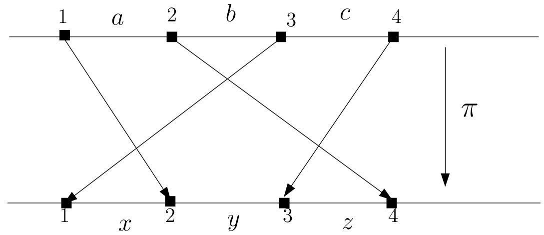


FIG. 1. The 4-separable permutation (2, 4, 1, 3).

1.1. RELATED WORK. The problem of embedding distance metrics into geometric spaces has been studied extensively [Kruskal 1964a, 1964b; Shepard 1962a, 1962b; Johnson and Lindenstrauss 2003; Linial 2002]. The minimum distortion problem is a natural variant of the bi-Lipschitz embeddings questions that were initially motivated by the study of Banach spaces.

A problem closely related to minimum distortion is that of minimum bandwidth. Minimum distortion can be viewed as a variation and generalization of the minimum bandwidth problem [Chinn et al. 1982; Diaz et al. 2002]. Good solutions for the latter, however, typically incur very large contraction and hence do not seem useful for solving minimum distortion.

After its introduction, the minimum distortion problem has received considerable attention in the research community. Most of the results, however, have been negative, showing the problem hard to even approximate. Among such results are those of Hall and Papadimitriou, who show the line embedding problem hard to approximate even within large factors if distortion is high [Hall and Papadimitriou 2005], and Papadimitriou and Safra [2005] show the general embedding problem hard to approximate within a factor of 3 in three dimensions.

Due to such results, some research has focused on *approximating* minimum distortion [Badoiu et al. 2005a, 2005b; Hall and Papadimitriou 2005] under certain circumstances (e.g., considering only injections, focusing on alternate definitions of distortion such as additive distortion, etc.). To the best of our knowledge, there have been no positive results on exactly solving the embedding problem with multiplicative distortion as the measure of similarity. After Kenyon et al. [2004], ours are the only positive results for finding an optimal embedding.

2. *Line Embeddings*

In this section, we focus on computing an optimal embedding between two fixed line metrics. A line metric is a set of points on a real one-dimensional line with the distance between any pair of points being their ℓ_1 distance (any ℓ_k distance is equivalent for one-dimensional points).

As mentioned earlier, Kenyon et al. [2004] consider the problem of optimally embedding one fixed line metric into another fixed one. They propose a polynomial-time, dynamic programming based algorithm that computes the optimal embedding if the distortion is less than $3 + 2\sqrt{2}$. They show that any bijection containing the bijection in Figure 1 as a partial bijection corresponds to an embedding with distortion at least $3+2\sqrt{2}$. These bijections have a nice structure that allows finding the optimal such permutation, using dynamic programming in polynomial time.

Notice that we can view any embedding as a mapping from source points to destination points, or simply as a permutation. In this section, we improve the result of Kenyon et al. [2004] by considering a less restricted class of permutations called k-separable permutations. We improve the threshold value on distortion, below which an optimal embedding can be found in polynomial time from $3+2\sqrt{2}$ to 13.602. Let us now introduce some basic definitions.

2.1. BASIC DEFINITIONS. Assume the optimal embedding between U and V is the permutation π. We specify a permutation π with the notation $(\pi(1), \pi(2), \ldots, \pi(n))$.

Permutation π_n of size n *contains* permutation π_k of size k if there exist indices $l_1 < l_2 < \cdots < l_k$ such that for all $1 \leq i < j \leq k, \pi_k(i) < \pi_k(j)$ iff $\pi_n(l_i) < \pi_n(l_j)$. In this case, we refer to π_k as a *subpermutation* of π_n. In particular, $\pi_n^{x,y}$ is the unique permutation of size $y - x + 1$ such that $\pi_n^{x,y}(i) < \pi_n^{x,y}(j)$ iff $\pi_n(i + 1 - x) < \pi_n(j + 1 - x)$.

By $[i, j]$, $i < j$, we mean the set of numbers from i to j. A *nice interval* I in π is either a singleton or a set of at least two consecutive numbers from 1 to n such that their mapping, via π, is still a set of consecutive numbers. For example, the permutation (4, 3, 1, 2) contains several nice intervals: [1, 2], [3, 4], [2, 4], and [1, 4].

If the interval $[1, n]$ can be decomposed into a constant number of subintervals such that each is mapped, via π, to a subinterval in V and if this property recursively holds for all subintervals, then we can use dynamic programming and find the optimal embedding. More formally, an interval I is *k-separable*, with respect to π, if either it has at most k points or it can be partitioned into nice subintervals $I_1, I_2, \ldots, I_m$ $(1 < m \leq k)$ such that each I_i is k-separable. Moreover, π is k-separable iff the interval $[1, n]$ is k-separable with respect to π. The *separability* of π is the minimum $k > 1$ such that π is k-separable.

For example, the permutation $\pi = (2, 4, 3, 6, 5, 1)$ is 3-separable. Specifically, $I_1 = [1, 3]$, $I_2 = [4, 5]$, $I_3 = [6]$, and it is clear that I_1, I_2, and I_3 are 3-separable as well.

Every 3-separable permutation is 2-separable, since for any three nice subintervals that partition a permutation, two may be merged to form a nice subinterval. Therefore, we don't have any permutation with separability 3. It's also easy to see that for $k \geq 4$, there exist permutations of size k with separability k. These permutations could be interpreted in a simpler way: They don't have any nice interval except the interval $[1, k]$. We refer to these special k-separable permutations as *nonseparable* permutations.

The distortion incurred by a permutation π, denoted by dist(π), is the minimum distortion incurred by embedding any two line metrics U and V via π. For example, dist(π) for the permutation in Figure 1 equals $3 + 2\sqrt{2}$ and happens when $[a, b, c, x, y, z] = [1, \sqrt{2}, 1, 1, \sqrt{2}, 1]$. As we see later, Theorem 2.3 states that dist(π) equals the largest eigenvalue of a 0-1 matrix corresponding to π.

Corresponding to every permutation π of size n, there exist three permutations π^0, π^1, and π^{-1} that are similar to π and incur the same distortion. For all i's, $\pi^0(i) = n + 1 - \pi(i)$, $\pi^1(\pi(i)) = i$, and $\pi^{-1}(i) = n + 1 - \pi^1(i)$. For example, if $\pi = (2, 4, 1, 3)$, $\pi^0 = \pi^1 = (3, 1, 4, 2)$ and $\pi^{-1} = \pi$. Throughout this section, we always assume that a permutation comes with all its four symmetric forms. For example, when we say 2-separable permutations avoid $\pi = (2, 4, 1, 3)$, we mean they avoid $(3, 1, 4, 2)$ as well.

Let Π_k be the set of all nonseparable permutations of size k. Let d_k be the minimum distortion over all permutations in Π_k. For example, $\Pi_4 = \{(2, 4, 1, 3)\}$, $\Pi_5 = \{(2, 4, 1, 5, 3), (2, 5, 3, 1, 4), (3, 5, 1, 4, 2)\}$, and it's not hard to see that $d_4 = d_5 = 3 + 2\sqrt{2}$. Note that by $\pi \in \Pi_k$ we implicitly mean $\pi^0, \pi^1, \pi^{-1} \in \Pi_k$ as well. So, $(3, 1, 5, 2, 4)$ is also in Π_5.

2.2. FORBIDDEN PERMUTATIONS. One commonly asked question regarding many permutation classes is whether they can be characterized by a finite forbidden set of permutations. For example, a permutation is 2-separable if and only if it contains neither $(2, 4, 1, 3)$ nor $(3, 1, 4, 2)$ [Bose et al. 1998].

Interestingly, we can generalize this statement for k-separable permutations.

THEOREM 2.1. *A permutation is k-separable if and only if the following holds.*

—*For odd k, it doesn't contain any permutation in* Π_{k+1}*; and*

—*for even k, it contains neither a permutation in* Π_{k+1} *nor* π^*_{k+2},

where π^*_{2m} *is the permutation of size* $2m$ *in which* $\pi^*_{2m}(2i) = i$ *and* $\pi^*_{2m}(2i-1) = i + m$.

PROOF. Assume π is not k-separable. Then, it must contain a nonseparable permutation π^0 of size $k_0 > k$. According to Albert and Atkinson [2005], every nonseparable permutation of size m either contains a nonseparable permutation of size $m-1$ or is identical to π^*_m. As π^*_m contains π^*_{m-2}, the statement of the theorem follows by repeatedly using the theorem of Albert and Atkinson [2005]. □

Note. Albert and Atkinson [2005, Theorem 4] use the notion *simple* for nonseparable and call π^*_{2m} an *exceptional* permutation. They obtain their result by using those from Schmerl and Trotter [1993] on partially ordered sets.

2.3. EMBEDDING BETWEEN TWO LINE METRICS. In this section we prove the following theorem, which is a generalization of Kenyon et al.'s [2004] result.

THEOREM 2.2. *For any two line metrics U and V and for any k, either the distortion of the optimal embedding between U and V is greater than* d_{k+1} *or there exists an* $O(k!n^{5k+2})$ *time algorithm (which is a polynomial in n when k is a constant) for computing the optimal embedding.*

Recall that d_{k+1} is the minimum distortion over all permutations in Π_{k+1}. Let π be the optimal embedding permutation. If π is not k-separable, then, according to Theorem 2.1, π contains either a permutation in Π_{k+1} or in π^*_{k+2} (in the case where k is even). From Sections 2.5 and 2.6, we can compute $\text{dist}(\pi^*_{k+2}) = 2k + 2\sqrt{k(k-1)} - 1$ and also conclude that $\text{dist}(\pi^*_{k+2}) \geq d_{k+1}$ for all k. Hence if π is not k-separable, then $\text{dist}(\pi) \geq d_{k+1}$. Otherwise, an algorithm for finding the optimal embedding follows.

2.4. THE ALGORITHM.

2.4.1. *Algorithm Intuition.* Our algorithm will guarantee that we solve all inputs whose optimal bijection π is k-separable, where k is the parameter. Note that if $\text{dist}(\pi) < d_{k+1}$, then π is k-separable. Setting $k = 46$, we get an algorithm that computes the bijection when the optimal bijection π has distortion $\text{dist}(\pi) < 13.602$.

At an intuitive level our algorithm will work as follows. It looks at every possible subinterval of the points in U against every possible subinterval of the points in V, starting with size 2 and working up to size n. It will break the subintervals into every possible k subsubinterval (including the empty sets). It will then try match these k subsubintervals by trying all $k!$ possible bijections of the subsubintervals. If a match is found with low enough distortion, the match will be saved for future reference. How the subintervals are mapped is no longer important; all we need know about the subinterval to continue the process is whether there was a bijection with distortion less than d_{k+1}, as well as the image and preimages, respectively, of the first and last point of U and the first and last point in V. The reason we need to keep the mappings of the first and last points in U and V is because when we try to combine two subintervals, we need to check the expansion and inverse expansion between them. We store this information in a table. When the subinterval is U and V, if we can map U to V by the same process with distortion less than d_{k+1}, then we output "yes".

Another way to think about the algorithm is to consider the algorithm to be looking for mappings that contain a pattern size k_1 for some $k_1 \leq k$. If it finds such a pattern it then thinks of this entire set as one mapping that could be part of another pattern of size $\leq k$, and it looks for such a pattern.

2.4.2. *Algorithm.* The algorithm gets as input two line metrics (U, d) and (V, d'). It also gets as parameters $\alpha = \sqrt{d_{k+1}}$, the maximum expansion and inverse expansion allowed, as well as a parameter k which is related to the bijections that the algorithm tries.

The algorithm proceeds by building a dynamic programming Boolean table T which is indexed by the following parameters:

—a subinterval $I = \{u_m, u_{m+1}, \ldots, u_{m+c-1}\}$ of U and a subinterval $J = \{v_{m'}, v_{m'+1}, \ldots, v_{m'+c-1}\}$ of V of the same size $c \geq 1$; and

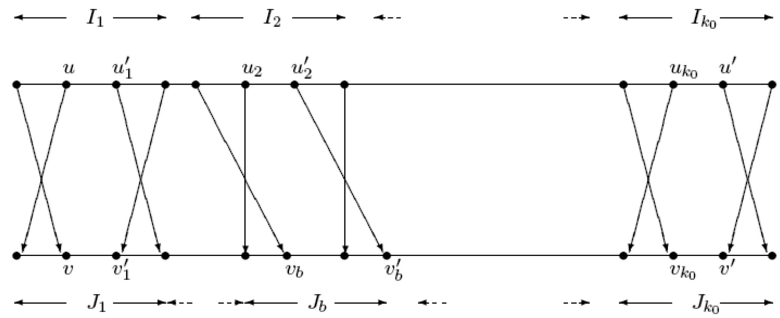
—four elements $v, v' \in J$ and $u, u' \in I$.[1] Specifically, v is the image of the first point in I, and v' is the image of the last point in I. Similarly, u is the preimage of the first point in J, and u' is the preimage of the last point in J.

We set the table entry $T[I, J, v, v', u, u']$ to true if there is a bijection $\sigma : I \to J$ such that $\sigma(u_m) = v$, $\sigma(u_{m+c-1}) = v'$, $\sigma^{-1}(v_{m'}) = u$, and $\sigma^{-1}(v_{m'+c-1}) = u'$, and with expansion and inverse expansion at most α.

The algorithm runs from $c = 1$ to n. The base case $c = 1$ is trivial, with all entries set to true. For $c > 1$, compute every entry $T[I, J, v, v', u, u']$ with $|I| = c$ and $|J| = c$ as follows: Consider all partitions of I and J into $2 \leq k_0 \leq k$ subintervals $I = \bigcup_{a=1}^{k_0} I_a$, $J = \bigcup_{b=1}^{k_0} J_b$. Try all possible combinations of pairs of I_a, J_b ($\sigma(I_a) = J_b$) over all a,b's and set $T[I, J, v, v', u, u']$ to true if and only if in at least one of the combinations, the following conditions hold.

—$\forall a, b \;\; T[I_a, J_b, v_b, v'_b, u_a, u'_a]$ is true, where $\sigma(I_a) = J_b$.

[1] Here, v' and u' do not denote images. They are just normal points. The same will hold throughout this subsection and we will specifically mention the images.



Filling in the table: a possible case.

—Let $J_{b_1} = \sigma(I_a)$, $J_{b_2} = \sigma(I_{a+1})$, $I_{a_1} = \sigma^{-1}(J_b)$, $I_{a_2} = \sigma^{-1}(J_{b+1})$. Then,

$$d(v_{b_2}, v'_{b_1}) \le \alpha \cdot d(\min(I_{a+1}), \max(I_a))$$

$$d(u_{a_2}, u'_{a_1}) \le \alpha \cdot d(\min(J_{b+1}), \max(J_b)).$$

These inequalities ensure that the edges connecting the subintervals have expansion and inversion expansion at most α.[2]

Once the table is prepared, the algorithm just checks whether $T[U, V, v, v', u, u']$ is true for some (v, v', u, u').

2.4.3. *Analysis.* The analysis proceeds as follows.

—*Correctness.* For the correctness of this algorithm we must show that we can solve any bijection whose optimal mapping is has distortion less than d_{k+1}. Since the distortion of the optimal mapping is less than d_{k+1}, the optimal mapping is k-separable. Hence, the permutation π contains only nice intervals of sizes at most k. Thus, the algorithm will try each of these partial mappings (on the nice intervals) and return a value of true for them.

—*Running Time.* The running time of the algorithm is easy to bound. Notice that the table size is just $O(n^7)$. Computing each entry $T[I, J, v, v', u, u']$ of the table is polynomial in n: The sets I and J can be split into $k_0 \le k$ sets in $O(n^{k-1})$ ways and for each such possible splitting we store $4(k_0 - 2) + 2 + 2 \le 4(k - 1)$ mappings, which can be done in $O(n^{4k-4})$; and, finally, there are $k!$ possible ways of mapping various I_a's to various J_b's. Thus computing each entry takes $O(n^{4k-4} \cdot n^{k-1} \cdot k!) = O(k!n^{5k-5})$ time. So, computing the whole table takes $O(k!n^{5k-5} \cdot n^7) = O(k!n^{5k+2})$.

This also completes the proof of Theorem 2.2.

2.5. LARGEST EIGENVALUE. In this subsection, we provide an interesting observation that the distortion of nonseparable patterns can be computed by computing the *largest eigenvalue* of the 0-1 matrix of their permutation. This observation suggests that we can find minimum distortions using a computer program.

Assume that the distortion corresponding to a permutation π of $[1, n]$ is λ. This means that for any two line metrics of n points each, the distortion using π is at least λ and there exists a pair of line metrics whose distortion, using π, is exactly

[2]Note that we need only consider the expansion and inverse expansion of edges [Kenyon et al. 2004].

λ. In fact, it is not hard to see that the maximum expansion and inverse expansion in embedding U to V happens for a pair of consecutive points, so we need to care only about them. Finding dist(π) corresponds to solving a set of linear equations. For example, for the permutation in Figure 1, the linear equations are

$$\begin{aligned} y+z &\leq \sqrt{\lambda}a \\ x+y+z &\leq \sqrt{\lambda}b \\ x+y &\leq \sqrt{\lambda}c \\ a+b &\leq \sqrt{\lambda}x \\ a+b+c &\leq \sqrt{\lambda}y \\ b+c &\leq \sqrt{\lambda}z, \end{aligned}$$

or equivalently $AX \leq \sqrt{\lambda}X$, where A is the adjacency matrix corresponding to π and X is $[a, b, c, x, y, z]^T$. In general, for a permutation π of size n that corresponds to embedding between two line metrics of size n, A has $2n-2$ rows and columns where, for all $0 \leq i, j < n$, $A[i, j] = A[n+i, n+j] = 0$ and $A[i, n+j] = A[n+i, j]$ is one iff the interval $[\pi(i), \pi(i+1)]$ (or $[\pi(i+1), \pi(i)]$ if $\pi(i) > \pi(i+1)$) contains the interval $[j, j+1]$, and is zero otherwise.

We can also assume that by scaling edge weights in U or V if necessary, the expansion and contraction both equal $\sqrt{\lambda}$. Thus, for any single edge in U and V we write an inequality to make sure that its corresponding expansion does not exceed $\sqrt{\lambda}$.

Since we are interested in minimizing λ, we had better make the equality $AX = \sqrt{\lambda}X$.[3] Therefore, $\sqrt{\lambda}$ is an eigenvalue of A. It is well known that when all entries of A are positive, the only eigenvalue whose corresponding eigenvector is positive is the largest eigenvalue [Horn and Johnson 1986, Chapter 8.2.]. Thus $\sqrt{\lambda}$ is the largest eigenvalue of A.

THEOREM 2.3. *Let A_π be the 0-1 matrix corresponding to π and let its largest eigenvalue be λ. Then, the distortion of π is exactly λ^2 and is obtained when the edge lengths are taken according to the eigenvector corresponding to λ.*

2.6. BOUNDING d_k. Although d_k is increasing in k, it remains bounded. This is somewhat disappointing, since if it were unbounded we could imagine an algorithm that finds an optimal embedding for any two line metrics, no matter how large the optimal distortion, whose running time is a function of the distortion.

THEOREM 2.4. *For any value k, there exists a nonseparable permutation π_k whose distortion is at most* $7 + 4\sqrt{3}$.

PROOF. Let $\hat{\pi}_{2n}$ be the permutation on $[1, 2n]$, where $\hat{\pi}_{2n}(1) = 2$, $\hat{\pi}_{2n}(2n) = 2n-1$, $\hat{\pi}_{2n}(2i) = 2i+2$, and $\hat{\pi}_{2n}(2i+1) = 2i-1$, for $i = 1, 2, \ldots, n-1$. Similarly, $\hat{\pi}_{2n+1}$ is defined as: $\hat{\pi}_{2n+1}(i) = \hat{\pi}_{2n}(i)$, for $i = 1, 2, \ldots, n-1$, $\hat{\pi}_{2n+1}(2n) = 2n+1$, and $\hat{\pi}_{2n+1}(2n+1) = 2n-1$(see Figure 2).

---

[3] To see this, let λ be the smallest distortion and assume $AX \leq \sqrt{\lambda}X$ for some positive vector X. Let X be the one with smallest sum of elements, that is, $X_1 + X_2 + ... + X_n$. If $AX < \sqrt{\lambda}X$ then $A_iX < \sqrt{\lambda}X_i$ (for some i), which means we can replace X_i by A_iX without violating the condition that $AX \leq \sqrt{\lambda}X$. This is a contradiction because now the new sum is the one with smallest sum of elements.

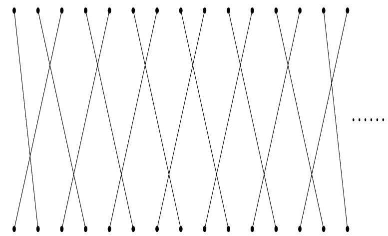


FIG. 2. Illustration of permutation $\hat{\pi}_{15}$.

TABLE I. DISTORTION OF $\hat{\pi}_k$ FOR SEVERAL VALUES OF k

k	5	7	9	11	13	15	17	19
distortion	8.352	10.896	12.045	12.651	13.007	13.233	13.385	13.492

TABLE II. d_k

k	4	6	9	12	15	24
d_k	5.828	8.352	9.899	10.896	11.571	12.850
k	30	34	38	42	46	
d_k	13.131	13.316	13.443	13.534	13.602	

Set $d_U(2i-1, 2i) = 1$, $d_U(2i, 2i+1) = \sqrt{3}$, $d_V(2i-1, 2i) = 2+\sqrt{3}$, and $d_V(2i, 2i+1) = 3+2\sqrt{3}$. The distortion corresponding to this pair of point sets is $7+4\sqrt{3}$, which means $d_k \leq 7+4\sqrt{3} \simeq 13.928$. □

Table I shows the exact distortion of such permutations for small values of k. Finding d_k for small k's (by computing the eigenvalue corresponding to all permutations in Π_k and taking the minimum) suggests that d_k converges to $7+4\sqrt{3}$. Table II shows the value of d_k for different k's.

Limitation of the approach. It is easy to see that if the pattern $\hat{\pi}_{15}$ keeps extending to infinity, then its distortion is $7+4\sqrt{3}$. Using the tightness property of edges in this pattern, we get the equations

$$\begin{array}{c} \alpha a = 2x+3y \ ; \ \alpha b = x+2y \\ \alpha x = 2a+3b \ ; \ \alpha y = a+2b \end{array} \Rightarrow$$
$$\alpha(2b-a) = y \ ; \ \alpha(2a-3b) = x$$
$$\alpha^2(2a-3b) = 2a+3b \ ; \ \alpha^2(2b-a) = a+2b,$$

from which we get $\alpha^2 = 7+4\sqrt{3} \approx 13.928$.

3. *Bounded-Degree Graphs with Short Cycles*

THEOREM 3.1. *Let (U, d) be the shortest-path metric of an unweighted graph G of maximum degree b ($b \neq 1$). Let (V, d') be the shortest-path metric of an arbitrary unweighted graph G'. Then, the problem of finding an optimal bijection between U and V is NP-hard.*

PROOF. This proof is based on the proof that it is NP-hard to approximate the minimum distortion problem within a factor better than 2, given in Kenyon et al. [2004]. Let G' be an unweighted, undirected graph on n vertices. Construct a metric

(V, d') by setting $d'(u, v) = 1$ if u, v is an edge of G', and $d'(u, v) = 2$ otherwise. Let the bounded-degree graph G be the unweighted cycle on n vertices, C. Clearly, C is of bounded degree $b = 2$ and construct the metric (U, d) in the same manner as (V, d'). It is easy to check that if G' contains a Hamilton cycle, then an optimal bijection between (U, d) and (V, d') has distortion exactly 2. If G' does not contain a Hamilton cycle, then any bijection must have distortion at least 4. Hence, the problem of finding an optimal bijection between (U, d) and (V, d') as described before is NP-hard. Since the given instance is a particular case of the metrics in the lemma, the lemma is true. □

In this section, we prove the following in a very similar manner to the algorithm presented in Kenyon et al. [2004].

THEOREM 3.2. *Let (U, d) be the shortest-path metric of an unweighted graph G of maximum degree b and largest cycle-length* 3. *Let (V, d') be the shortest-path metric of an arbitrary unweighted graph G'. Then, for any fixed constants b and α, there is an $O(n^2)$ algorithm that decides whether there exists a bijection between U and V with expansion and inverse expansion at most α.*

3.1. STRUCTURAL PROPERTIES. We begin with a few definitions. For a subset of vertices $A \subseteq G$, let $\Gamma(A)$ denote the set of neighbors of A that lie outside A. We also use $\Gamma(v)$ to denote the set of neighbors of a vertex $v \in G$.

Definition 3.3. We say that a graph G is *graph-rooted* at vertex r_0 by assigning every vertex $v \in G$ a value $l(v)$ that is equal to the length of the shortest path from v to r_0 in G (with $l(r_0) = 0$). By *level*(i), we denote the set of all vertices v in G such that $l(v) = i$.

Definition 3.4. G_r is the subgraph rooted at vertex r according to the following definition.

(1) r is in G_r.
(2) If there exists a path from r to a vertex v in G such that for all vertices v' along this path (including v), $l(v') > l(r)$, then $v \in G_r$.
(3) If (v_1, v_2) is an edge in G and both v_1 and v_2 are $\in G_r$, then the edge (v_1, v_2) is an edge in G_r.

We now prove the following lemma (based on the proof of Kenyon et al. [2004] in the case where (U, d) is the shortest-path metric of an unweighted tree T of maximum degree b). Let $B(u, l)$ (respectively, $B'(u, l)$) denote the closed ball of radius l around any vertex u in G (respectively, in G'). For a subset of vertices $A \subseteq G$ (respectively, in G'), let $\Gamma(A)$ (respectively, $\Gamma'(A)$) denote the set of neighbors of A that lie outside A. Assume that G is graph-rooted at an arbitrary vertex r_0. The subgraph rooted at any vertex r of G (as defined earlier) is denoted by G_r.

LEMMA 3.5. *Let $\sigma : U \to V$ be a bijection with expansion and inverse expansion at most α. Then, the following holds.*

(1) *G' has maximum vertex degree at most b^α.*
(2) *For any vertex $r \in G$, each connected component of $G' \setminus B'(\sigma(r), \alpha^2)$ lies either entirely in $\sigma(G_r)$ or entirely in $G' \setminus \sigma(G_r)$.*

(3) *For any $r \in G$, for any adjacent pair (u', v') in G' with $u' \in \sigma(G_r)$ and $v' \notin \sigma(G_r)$, both $\sigma^{-1}(u')$ and $\sigma^{-1}(v')$ are in $B(r, \alpha)$*

PROOF. For the first statement, for any $v \in G'$, the expansion of σ^{-1} implies that $\sigma^{-1}(B'(v, 1)) \subseteq B(\sigma^{-1}(v), \alpha)$, and the cardinality of this ball is at most b^{α} by the degree bound on G.

For the second statement, let G_r be the subgraph graph-rooted at $r \in G$. Let $v' = \sigma(v)$ be a vertex in $\Gamma'(\sigma(G_r))$. By the definition of Γ', v' is adjacent to some vertex $u' = \sigma(u)$ of $\sigma(G_r)$. From the inverse expansion bound, we have $d(u, v) \leq \alpha$. Now, assume that the shortest path from u to v goes through r. Then, clearly $d(r, v) \leq \alpha$. Thus we have $d'(\sigma(r), v') \leq \alpha^2$. From this we get

$$\Gamma'(\sigma(G_r)) \subseteq B'(\sigma(r), \alpha^2),$$

from which we get the second statement.

For the third statement, note that by the expansion of σ^{-1}, we get that $d(\sigma^{-1}(u'), \sigma^{-1}(v')) \leq \alpha$. Again assuming that the shortest path from u to v goes through r, we get that $d(r, \sigma^{-1}(u')) \leq \alpha$ and $d(r, \sigma^{-1}(v')) \leq \alpha$.

Now, the proof of Lemma 3.6 completes this proof. □

LEMMA 3.6. *Let $u \in G_r$ and $v \notin G_r$, then the shortest path from u to v goes through r.*

PROOF. We shall prove this by contradiction. Suppose the shortest path from u to v does not go through r. In this case, this path has to go through a node (r') such that $l(r') \leq l(r)$ (otherwise, v is a vertex of G_r). Note that there is a path from r to r' such that any vertex w on this path ($w \neq r, r'$) has $l(w) < l(r)$. Hence, there is a path from r to r' of length at least 2 that does not overlap with the paths from u to r and u to r'. Now, consider the nonoverlapping parts of the paths from u to r and u to r'. The lengths of these parts are at least 1 each, and hence we get a cycle of length at least 4 (by joining the path from r to r' completely at lower levels and the path from r to r' completely at higher levels). This is a contradiction to the maximum cycle-length restriction of 3 on G. Hence, the shortest path from u to v goes through r. □

We now present the algorithm, its analysis, and the proof of Theorem 3.2. This follows from the algorithm in the case of bounded-degree trees presented in Kenyon et al. [2004].

3.2. ALGORITHM AND PROOF OF THEOREM 3.2.

3.2.1. *Algorithm.* The algorithm is a dynamic programming algorithm in the same way as given in Kenyon et al. [2004]. The graph G is graph-rooted arbitrarily at a node r_0. The dynamic programming table T is indexed by the following parameters:

(1) $r \in \{u_1, \ldots, u_n\}$, the root of the subgraph G_r (with respect to the graph-rooting G);
(2) $r' \in \{v_1, \ldots, v_n\}$;
(3) an injection τ from $B(r, \alpha) \cap G_r$ into $B'(r', \alpha^2)$; and
(4) a subset S of the vertices of G' with the property that each connected component of $G' \backslash B'(r', \alpha^2)$ lies entirely within S or entirely outside of S.

An entry of the table is true if and only if there exists an injection $\sigma : G_r \to G'$ such that $\sigma(r) = r'$, σ coincides with r on $B(r, \alpha) \cap G_r$, $\sigma(G_r) = S$, and expansion of every edge of G_r and inverse expansion of every edge of $\sigma(G_r)$ are each at most α. To compute $T(r, r', \tau, S)$, we run through all combinations of entries $T(r_i, r'_i, \tau_i, S_i)_i$, all of which have value true. The r_i's are the children of a given root r. We set the result to be true if at least one of these combinations satisfies the conditions given next and to false otherwise.

(1) The map τ is consistent with all the maps τ_i's, the τ_i's are consistent among themselves, the S_i's do not include r', and S is the union of the S_i's plus the vertex r'.
(2) For each r'_i, we have $d'(r', r'_i) \leq \alpha d(r, r_i)$.
(3) For each adjacent pair v', w' in G', that belong to different sets S_i (or with $v' = r'$), both v' and w' are in the image of τ and satisfy $d(\tau^{-1}(v'), \tau^{-1}(v')) \leq \alpha$.

After all entries of the dynamic programming table are computed, the algorithm checks whether some table entry $T(r_0, ., ., .)$ is true.

3.2.2. *Running Time and Correctness.* The degree bound on G implies that $B'(v, \alpha^2)$ has size at most b^{α^3} for any v. We claim that the size of the table T is at most

$$n \times n \times (b^{\alpha^3})^{b^\alpha} \times 2^{2b^{\alpha^3}} = O(n^2).$$

The two n terms come from the r and r' in the table. The third factor bounds the number of maps from $B(r, \alpha)$ to $B'(r', \alpha^2)$. From the second part of the lemma, we get the number of possibilities for the set S as the fourth factor. Filling the table entries takes constant time, as, given r and r', we only have to consider r'_i such that $r'_i \in B'(r', \alpha)$; for further details, see Kenyon et al. [2004]. Thus the overall running time is $O(n^2)$.

The correctness of the algorithm follows in the same way as in Kenyon et al. [2004] by an induction (bottom-up the levels in G). This also completes the proof of Theorem 3.2.

ACKNOWLEDGMENTS. We wish to thank Kenyon et al. [2004] for pointing us to the Albert-Atkinson [2005] lemma. This replaced our previous independent proof of a similar lemma that we needed to establish our result. We also wish to thank M. Stanojević for helping with the computation of the eigenvalues.

REFERENCES

AKUTSU, T., KANAYA, K., OHYAMA, A., AND FUJIYAMA, A. 2003. Point matching under non-uniform distortions. *Discr. Appl. Math. 127*, 5–21.

ALBERT, M. H., AND ATKINSON, M. D. 2005. Simple permutations and pattern restricted permutations. *Discr. Math. 300*, 1-3, 1–15.

BADOIU, M., CHUZHOY, J., INDYK, P., AND SIDIROPOULOS, A. 2005. Low-Distortion embeddings of general metrics into the line. In *Proceedings of the ACM Symposium on the Theory of Computing (STOC)*, ACM.

BADOIU, M., DHAMDHERE, K., GUPTA, A., RABINOVICH, Y., RACKE, H., RAVI, R., AND SIDIROPOULOS, A. 2005. Approximation algorithms for low-distortion embeddings into low-dimensional spaces. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, ACM.

BELONGIE, S., MALIK, J., AND PUZICHA, J. 2002. Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell. 24*, 4, 509–522.

BOSE, P., BUSS, J. F., AND LUBIW, A. 1998. Pattern matching for permutations. *Inf. Process. Lett. 65*, 5, 277–283.

CHAZELLE, B., LIU, D., AND MAGEN, A. 2003. Sublinear geometric algorithms. In *Proceedings of the ACM Symposium on the Theory of Computing (STOC)*, ACM.

CHINN, P., CHVATALOVA, J., DEWDNEY, A., AND GIBBS, N. 1982. The bandwidth problem for graphs and matrices—A survey. *J. Graph Theory 6*, 223–254.

DIAZ, J., PETIT, J., AND SERNA, M. 2002. A survey on graph layout problems. *ACM Comput. Surv. 34*, 3, 313–356.

FORTIN, S. 1996. The graph isomorphism problem. Tech. Rep. 96-20, University of Alberta, Edomonton, Alberta, Canada.

HALL, A., AND PAPADIMITRIOU, C. 2005. Approximating the distoriton. In *Proceedings of the International Workshop on Randomization and Approximation Techniques (RANDOM APPROX)*. Lecture Notes in Computer Science, vol. 3624. Springer, 111–122.

HOFFMANN, F., KRIEGEL, K., AND WENK, C. 1998. Matching 2D patterns of protein spots. In *Proceedings of the ACM Symposium on Computational Geometry*. ACM, 231–239.

HORN, R. A., AND JOHNSON, C. R. 1986. *Matrix Analysis*. Cambridge University Press, New York.

JOHNSON, W., AND LINDENSTRAUSS, J. 2003. *Handbook of Geometry of Banach Spaces*. North-Holland.

KENYON, C., RABANI, Y., AND SINCLAIR, A. 2004. Low distortion maps between point sets. In *Proceedings of the ACM Symposium on the Theory of Computing (STOC)*, ACM.

KRUSKAL, J. 1964a. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Pyschometrika 29*, 1–27.

KRUSKAL, J. 1964b. Nonmetric multidimensional scaling: A numerical method. *Pyschometrika 29*, 115–129.

LINIAL, N. 2002. Finite metric spaces—Combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians III*, 573–586.

PAPADIMITRIOU, C., AND SAFRA, S. 2005. The complexity of low-distortion embeddings between point sets. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, ACM.

SCHMERL, J. H., AND TROTTER, W. T. 1993. Critically indecomposable partially ordered sets, graphs, tournaments and other binary relational structures. *Discr. Math. 113*, 1-3, 191–205.

SHEPARD, R. 1962a. The analysis of proximities: Multidimensional scaling with an unknown distance function 1. *Pyschometrika 27*, 125–140.

SHEPARD, R. 1962b. The analysis of proximities: Multidimensional scaling with an unknown distance function 2. *Pyschometrika 27*, 216–246.

RECEIVED FEBRUARY 2006; REVISED FEBRUARY 2008; ACCEPTED MARCH 2008