Internation Journal of Computer Vision, 28(3), 219–231 (1998)

Relaxing the Triangle Inequality in Pattern Matching

RONALD FAGIN AND LARRY STOCKMEYER
IBM Almaden Research Center, 650 Harry Road, San Jose, CA 95120-6099
fagin@almaden.ibm.com
stock@almaden.ibm.com
Received October 29, 1997; Accepted August 10, 1998

Abstract. Any notion of "closeness" in pattern matching should have the property that if A is close to B, and B is close to C, then A is close to C. Traditionally, this property is attained because of the triangle inequality ($d(A,C) \leq d(A,B) + d(B,C)$, where d represents a notion of distance). However, the full power of the triangle inequality is not needed for this property to hold. Instead, a "relaxed triangle inequality" suffices, of the form $d(A,C) \leq c(d(A,B) + d(B,C))$, where c is a constant that is not too large. In this paper, we show that one of the measures used for distances between shapes in (an experimental version of) IBM's QBIC[1] ("Query by Image Content") system (Niblack et al., 1993) satisfies a relaxed triangle inequality, although it does not satisfy the triangle inequality.

Keywords: pattern matching, shape matching, triangle inequality, distance measure, image database

1. Introduction

Traditionally, databases have been used to store and retrieve textual and numerical information. More recently, applications such as multimedia have led to the development of database systems that can handle images. One such system is the QBIC ("Query by Image Content") system (Niblack et al., 1993), developed at the IBM Almaden Research Center. An experimental version of the QBIC system (henceforth in this paper called simply "QBIC") can search for images by various visual characteristics such as color, shape, and texture. While the result of a query to a traditional database is usually some specific set of items (e.g., the names of all employees in the computer science department), the result of a query to a database of images might not be so well-defined. Consider, for example, a query that should return all items that look like a tree; such a query could be entered by having the user draw the desired tree-like shape on a screen, or by extracting the shape from a visual scene. Questions of the form "Does the shape D in the database look like the query tree shape Q?" do not have definite yes/no answers (unlike questions of the form "Is employee E in the computer science department?"). Rather, the answer to such a question is more reasonably given as a numerical "distance" that measures how well the shape D matches the shape Q. The answer to the query could then be an ordered list of shapes from the database, ordered by how closely they match the query shape Q. This raises the issue of how to define a measure of "distance" between shapes.

There is an extensive literature about various ways to define distances between shapes. These include methods based on turning angles (Arkin et al., 1990; McConnell et al., 1991), on the Hausdorff distance (Huttenlocher et al., 1992), on various forms of moments (Kim and Kim, 1997; Taubin and Cooper, 1991), and on Fourier descriptors (Jain, 1989).

Mehtre et al. (1997) and Mumford (1991) discuss and compare various approaches. Scassellati et al. (1994) compare methods on the basis of how well they correspond to human perceptual distinctions. In Section 2, we discuss a particular distance measure between shapes, that is one of the measures used in the QBIC system. Intuitively, it measures how well the boundary of one shape matches the boundary of the other, allowing either boundary to stretch when doing the matching. A variation of this method provided the best overall results in the Scassellati et al. study.

Let us reconsider the problem we mentioned earlier, where Q is a shape, and where we wish to obtain an ordered list of shapes from the database, ordered by how closely they match Q. Let us say that as in the QBIC system, we wish to see the best 10 matches, and then upon request the next best 10 matches, and so on. This is a computationally expensive process, for several reasons. For a given shape D in the database, computing the distance between Q and D may well be expensive in itself: for example, for the distance measure used in QBIC that is discussed in Section 2, a dynamic programming algorithm is used that has quadratic complexity. Furthermore, even if we wish to see only the best 10 matches, we may have to compute the distance between Q and *every* shape D in the database: this is because there is no obvious indexing mechanism that can be used.

A potential avenue for speeding up the search is to preprocess the database, clustering shapes according to their distance amongst themselves. Then, for example, if we have found that Q is far from the database shape D_1, and if the preprocessing tells us that D_1 is close to another database shape D_2, we might be able to infer that Q is sufficiently far from D_2 that we do not need to actually compute the distance between Q and D_2. Similarly, if Q is close to D_1, and if the preprocessing tells us that D_1 is far from D_2, we might be able to infer that Q is sufficiently far from D_2. For this to work, we must be able to relate the distance between Q and D_2 to the distance between Q and D_1 and the distance between D_1 and D_2, for example, by the triangle inequality. The triangle inequality for a distance measure d states that, for all A, B, and C,

$$d(A,C) \leq d(A,B) + d(B,C).$$

In considering similarity measures between shapes, Arkin et al. (1990) say that such a measure should be a metric. In particular, they say:

> *The triangle inequality is necessary since without it we can have a case in which $d(A,B)$ and $d(B,C)$ are both very small, but $d(A,C)$ is very large. This is undesirable for pattern matching and visual recognition applications.*

The theme of this paper is that we agree completely that a distance measure d where $d(A,B)$ and $d(B,C)$ are both very small, but where $d(A,C)$ is very large, is certainly undesirable. Instead, we want a distance measure d to have the property that if A is close to B, and B is close to C, then A is close to C. But to obtain this property, it is not necessary that d satisfy the triangle inequality. Instead, it is sufficient for d to satisfy a "relaxed triangle inequality" of the form

$$d(A,C) \leq c(d(A,B) + d(B,C)), \tag{1}$$

where c is a constant that is not too large. We show that a measure used for distances between shapes in the QBIC system satisfies a relaxed triangle inequality, although it does not satisfy the triangle inequality.

What if we are in a scenario where a relaxed triangle inequality holds? Recalling the situation described above, where we know distances $d(Q,D_1)$ and $d(D_1,D_2)$ and we want to conclude something about $d(Q,D_2)$, if d satisfies (1) and is symmetric we can infer the bounds

$$\begin{aligned} d(Q,D_2) &\geq (1/c)\cdot d(Q,D_1) - d(D_1,D_2) \\ d(Q,D_2) &\geq (1/c)\cdot d(D_1,D_2) - d(Q,D_1) \\ d(Q,D_2) &\leq c(d(Q,D_1) + d(D_1,D_2)). \end{aligned}$$

The first two inequalities correspond to the situations described earlier, where we conclude that Q is sufficiently far from D_2, without actually computing this distance. The third inequality corresponds to a situation where we conclude that Q is sufficiently close to D_2, by knowing that Q is close to D_1, and that D_1 is close to D_2. We note that this last case might not provide useful information in a system such as QBIC, where we want to know, in the case of close matches, just how close the match is (because the results are presented in sorted order based on closeness of match).

The remainder of the paper has three sections and an appendix. In Section 2, we formally define the distance

NEM$_r$, one of the measures used in the QBIC system. In Section 3, the definition is illustrated by an example. In Section 4, we sketch the proof of the relaxed triangle inequality; the full proof is given in the appendix. We give the definitions and results in greater generality than for the specific application to distances between shapes. The relaxed triangle inequality for shape distance follows immediately from the more general results. We also show in Section 4 that the value of the constant c we give in the relaxed triangle inequality is essentially the best possible within the more general framework. However, for the specific application to shape distance, some smaller constant might be possible, particularly when restricted to shapes meeting some naturalness property. In Section 4 we remark on ways that the relaxed triangle inequality might be improved, by using extra information contained in the boundary matching between two shapes (that is, in addition to the *NEM*$_r$-distance obtained from the boundary matching). An example of extra information that could be helpful is the amount of stretching done. Such improvements may be necessary for the relaxed triangle inequality to be useful in practice.

Even though the technical results in this paper apply to a specific distance measure, the results carry a more general message: A distance measure should not be judged unsuitable simply because it does not satisfy the triangle inequality; it might be possible to prove that the measure satisfies a relaxed triangle inequality. Our specific results give a concrete example of this, by proving that a natural measure of distance between shapes satisfies a relaxed triangle inequality, although it does not satisfy the triangle inequality.

2. The Distance Measure *NEM*$_r$

One intuitively appealing way to measure the distance between shapes is to measure how well the boundary of one shape matches the boundary of the other, allowing either boundary to stretch when doing the matching. This measure has been used, for example, in (Cortelazzo et al., 1994) for trademark shapes and in (McConnell et al., 1991) for ice floes. As in (Cortelazzo et al., 1994), we call this distance measure *nonlinear elastic matching* (*NEM*). After we define this measure formally, we shall show that *NEM* does not satisfy the niceness property we discussed in the introduction: it is possible for the *NEM*-distance between A and B to be small, and the *NEM*-distance between B and C to be small, with the *NEM*-distance between A and C being large. That *NEM* does not satisfy the triangle inequality was known previously (cf. (Cortelazzo et al., 1994)); we show further that it does not even satisfy a relaxed triangle inequality.

Niblack and Yin (1995) defined a modified version of *NEM*, which is essentially one of the methods implemented in the QBIC system. It is related to a distance notion described in (McConnell et al., 1991). Niblack and Yin's definition depends on a parameter r, a positive number, which we call the *stretching penalty*. The idea, informally, is that we add to the distance an amount equal to r times the amount of stretching that was done to make the two boundaries match. Thus, we pay a penalty for excessive stretching. Letting *NEM*$_r$ denote the modified measure, we show that *NEM*$_r$ satisfies a relaxed triangle inequality (1) with constant $c = 1 + O(1/r)$. Thus, c approaches 1 as r increases. As we shall show in Section 3, the version of the *NEM*-distance involving a stretching penalty as described in (McConnell et al., 1991) does not satisfy a relaxed triangle inequality.

We now consider the definition of *NEM*$_r$. Fix some stretching penalty $r \geq 0$. (Although we are primarily interested in the case $r > 0$, we allow $r = 0$ since *NEM*$_0$ is equivalent to *NEM*, so we get the definition of *NEM* as a special case.) Shortly, we shall define the distance *NEM*$_r(X, Y)$ between two sequences

$$\begin{aligned} X &= x_1, x_2, \ldots, x_m \\ Y &= y_1, y_2, \ldots, y_n. \end{aligned}$$

In general, we allow $m \neq n$ and we allow the elements x_i and y_j of the sequences to belong to some metric space S with distance metric b. We refer to (S, b) as the *base*. In particular, we assume that b is symmetric and satisfies the triangle inequality for all points in S, and that $b(x, x) = 0$ for all $x \in S$. We show that the *NEM*$_r$-distance satisfies a relaxed triangle inequality for any $r > 0$ and any S that is bounded, i.e., such that b_{sup} is finite, where

$$b_{sup} = \sup\{\, b(x, y) \mid x, y \in S \,\}.$$

The constant c in the relaxed triangle inequality depends on r and b_{sup}. In the application to shape matching, as we shall now discuss, the elements x_i and y_j represent tangent angles, and b measures the difference between two angles. Hence, in this case, $S = [0, 2\pi)$ and

$$b(x, y) = \min\{\, |x - y|, 2\pi - |x - y| \,\},$$

so $b_{sup} = \pi$.

We now discuss Niblack and Yin's approach to shape matching. We assume that each shape is given by a simple (non-self-intersecting) closed curve in the plane. We measure how well a particular point a on the boundary of one shape matches a particular point b on the boundary of another shape as the difference between the tangent angle to the boundary at point a and the tangent angle to the boundary at point b. Thus, we begin by replacing each shape by a sequence of tangent angles taken at some number n of points spaced equally in distance around the boundary of the shape. If $X = x_1, x_2, \ldots, x_n$ is the sequence of tangent angles for the first shape, and $Y = y_1, y_2, \ldots, y_n$ is the sequence of tangent angles for the second shape, then the NEM_r-distance between the shapes is taken to be the NEM_r-distance (which we shall define shortly) between the sequences X and Y. The NEM_r-distance between two shapes depends on the "starting points" on the boundaries of the two shapes (that is, where the comparisons begin). Ideally, the distance between two shapes should be taken as the min of the distance over all possible starting points. In fact, Niblack and Yin (1995) focus on this issue of starting points, based, for example, on the shape's moments. In this paper, we shall not consider this issue: we will assume that the starting points are given. It is easy to see that our results on the existence of a relaxed triangle inequality would continue to hold even if we were to define the distance by taking the min of the distance over all possible starting points. In the QBIC system, there is a fixed number of points, equally spaced around the boundary of the shape, and so the starting point determines the sequence. Other papers consider notions of distance that depend only on the shapes. For example, in (Arkin et al., 1990), where a distance function is given for polygonal shapes, this distance function does not depend on any other parameters such as auxiliary points taken along the boundary.

When we say that NEM_r satisfies a relaxed triangle inequality $NEM_r(A, C) \leq c(NEM_r(A, B) + NEM_r(B, C))$, we mean that the constant c does not depend on the length of the sequences A, B, C. In the application to shape matching, this means that c does not depend on the number of sample points. Specifically, we show that $c = (1 + \pi/2r)$ works if the same number of sample points is used for all shapes. If the number of sample points varies from shape to shape, we still obtain a relaxed triangle inequality, but with the larger constant $c = (1 + \pi/r)$. (One can imagine weaker versions of the concept of a relaxed triangle inequality where the "constant" c might depend on the dimensionality of the space from which the points A, B, C are drawn. However, for NEM_r there is no need to weaken it in this way.)

We return to the definition of $NEM_r(X, Y)$. An (m, n)-*mapping* is a set

$$M \subseteq \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\},$$

where we call each pair $\langle i, j\rangle \in M$ an *edge*, satisfying the following conditions:

1. Every number in $\{1, 2, \ldots, m\}$ is the first component i of some edge $\langle i, j\rangle \in M$;
2. Every number in $\{1, 2, \ldots, n\}$ is the second component j of some edge $\langle i, j\rangle \in M$; and
3. No two edges "cross", that is, there do not exist i, i', j, j' with $i < i'$, $j < j'$, and $\langle i, j'\rangle, \langle i', j\rangle \in M$.

An (m, n)-mapping M is *minimal* if no proper subset of M is an (m, n)-mapping. Note that in any minimal mapping, there cannot be three edges $\langle i, j\rangle, \langle i', j\rangle, \langle i', j'\rangle$, since the subset obtained by removing the edge $\langle i', j\rangle$ is a mapping. For example, Figure 1 shows a minimal $(9, 9)$-mapping. We sometimes refer to an (m, n)-mapping simply as a *mapping* when m and n are clear from context or unimportant.

An edge $\langle i, j\rangle \in M$ is a *stretch-edge (of M)* if either $\langle i-1, j\rangle \in M$ or $\langle i, j-1\rangle \in M$. For an edge $\langle i, j\rangle$ in the mapping M, define *s-cost*$(\langle i, j\rangle, M)$, the *stretch-cost of* $\langle i, j\rangle$ *with respect to* M, as

$$\textit{s-cost}(\langle i, j\rangle, M) = \begin{cases} r & \text{if } \langle i, j\rangle \text{ is a stretch-edge} \\ 0 & \text{otherwise.} \end{cases}$$

For example, in the mapping shown in Figure 1, the edges $\langle 2, 3\rangle, \langle 3, 5\rangle, \langle 5, 6\rangle, \langle 6, 6\rangle, \langle 7, 6\rangle$ and $\langle 9, 9\rangle$ are stretch-edges and each has stretch-cost r, while the other edges have stretch-cost 0.

Define *d-cost*$(\langle i, j\rangle, X, Y)$, the *distance-cost of* $\langle i, j\rangle$ *with respect to the sequences* X, Y, as

$$\textit{d-cost}(\langle i, j\rangle, X, Y) = b(x_i, y_j).$$

The stretch-cost and the distance-cost of the mapping M, the latter with respect to the sequences X, Y, are defined by summing the respective costs of all edges in M; that is

$$\textit{s-cost}(M) = \sum_{e \in M} \textit{s-cost}(e, M)$$

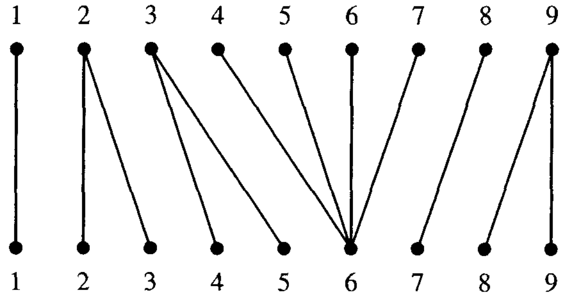

Fig. 1. A minimal $(9, 9)$-mapping. The stretch-cost of this mapping is $6r$.

$$\textit{d-cost}(M, X, Y) = \sum_{e \in M} \textit{d-cost}(e, X, Y).$$

The (total) cost of M is given by

$$\textit{cost}(M, X, Y) = \textit{s-cost}(M) + \textit{d-cost}(M, X, Y).$$

Finally,

$$\textit{NEM}_r(X, Y) = \min\{\, \textit{cost}(M, X, Y) \mid M \text{ is an } (m, n)\text{-mapping}\,\}.$$

In the sequel, we abbreviate $\textit{d-cost}(M, X, Y)$ by $\textit{d-cost}(M)$ and $\textit{cost}(M, X, Y)$ by $\textit{cost}(M)$ whenever the sequences X and Y are clear from context. Similarly, for an edge e in a mapping M, we may abbreviate $\textit{s-cost}(e, M)$ by $\textit{s-cost}(e)$ when M is clear.

Clearly the value of $\textit{NEM}_r(X, Y)$ does not change if we minimize over only the minimal (m, n)-mappings. It is also easy to see that $\textit{NEM}_r(X, Y) = \textit{NEM}_r(Y, X)$ for all X and Y, because for any (m, n)-mapping M, the set of edges obtained by reversing the first and second components of each edge in M gives an (n, m)-mapping M' having the same stretch-cost and the same distance-cost as M.

Although this definition of $\textit{NEM}_r(X, Y)$ involves a search over a number of mappings that grows exponentially in the minimum of m and n, it is well known that functions such as $\textit{NEM}_r(X, Y)$ can be computed in time $O(mn)$ by dynamic programming (see, for example, (McConnell et al., 1991; Cortelazzo et al., 1994; Niblack and Yin, 1995)). The algorithm iteratively computes the quantities $D(i, j)$, where $D(i, j)$ is the $\textit{NEM}_r$-distance between the length-i prefix of X and the length-j prefix of Y. The values of $D(i, j)$ can be computed by $D(1, 1) = b(x_1, y_1)$ and, for $i, j > 1$,

$$D(i, 1) = b(x_i, y_1) + D(i-1, 1) + r$$

$$D(1, j) = b(x_1, y_j) + D(1, j-1) + r$$

$$D(i, j) = b(x_i, y_j) + \min \begin{cases} D(i-1, j) + r \\ D(i-1, j-1) \\ D(i, j-1) + r \end{cases}$$

Then $\textit{NEM}_r(X, Y) = D(m, n)$.

3. An Example

We now illustrate the definitions with a simple example. Another purpose of the example is to show that the *NEM*-distance, where the stretching penalty r is 0, does not satisfy a relaxed triangle inequality, and to show that the $\textit{NEM}_r$-distance does not satisfy the triangle inequality for a small enough positive r. (In Section 4, we give a lower bound on the constant c in the relaxed triangle inequality for $\textit{NEM}_r$; since in particular this lower bound is bigger than 1 for every r, this shows that for every r, the $\textit{NEM}_r$-distance fails to satisfy the triangle inequality.) The example in this section also shows that the version of the *NEM*-distance involving a stretching penalty as described in (McConnell et al., 1991) does not satisfy even a relaxed triangle inequality. Thus, it is important how the stretching penalty r enters into the distance calculation: the method of (Niblack and Yin, 1995), where r is additive, gives a relaxed triangle inequality, whereas that of (McConnell et al., 1991), where r is multiplicative, does not.

Consider the three shapes shown in Figure 2. Note that each shape consists of five "short" line segments and three "long" line segments. (Although the shapes in Figure 2 were chosen to be polygons for simplicity, the $\textit{NEM}_r$-distance can be applied to more general shapes whose boundaries are curved.) The first step is to convert each shape into a sequence of tangent angles by placing sample points around the boundaries. To simplify the example suppose that, for each shape, one sample point is placed on each short line segment, k sample points are placed along each of the two long line segments that are part of the top of the shape, and m sample points are placed along the long line segment forming the bottom of the shape. The total number of sample points is therefore $n = 2k+m+5$. In each case we mark the starting point with an arrow, and we move clockwise around the shape. These sample points give

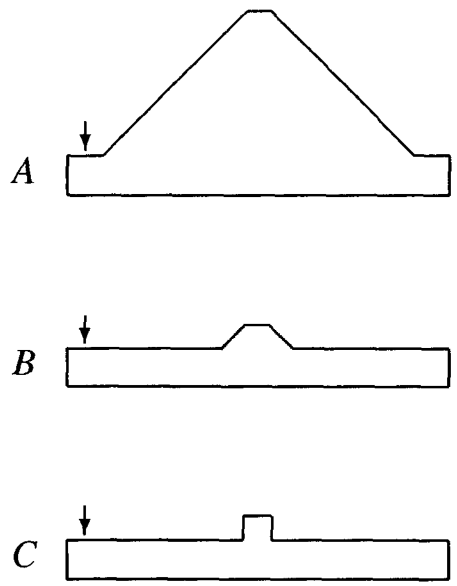

Fig. 2. Three shapes used to illustrate distances between shapes.

the following sequences of tangent angles:

$$
\begin{aligned}
A &: 0, \overbrace{\frac{\pi}{4},\ldots,\frac{\pi}{4}}^{k}, 0, \overbrace{\frac{7\pi}{4},\ldots,\frac{7\pi}{4}}^{k}, 0, \frac{3\pi}{2}, \overbrace{\pi,\ldots,\pi}^{m}, \frac{\pi}{2} \\
B &: \overbrace{0,\ldots,0}^{k}, \frac{\pi}{4}, 0, \frac{7\pi}{4}, \overbrace{0,\ldots,0}^{k}, \frac{3\pi}{2}, \overbrace{\pi,\ldots,\pi}^{m}, \frac{\pi}{2} \\
C &: \overbrace{0,\ldots,0}^{k}, \frac{\pi}{2}, 0, \frac{3\pi}{2}, \overbrace{0,\ldots,0}^{k}, \frac{3\pi}{2}, \overbrace{\pi,\ldots,\pi}^{m}, \frac{\pi}{2}.
\end{aligned}
$$

Let *angles*(S) denote the sequence of tangent angles associated with the shape S, where S is A, B, or C.

Consider first *NEM*, where the stretching penalty r is 0. In Figure 2, *NEM*$(A, B) = 0$: the small triangular protrusion in shape B is stretched to perfectly match the large triangular protrusion in shape A, and the short horizontal segments to the left and right of the large triangular protrusion in shape A are stretched to exactly match the long horizontal segments to the left and right of the small triangular protrusion in shape B; the rest of the boundaries of shapes A and B match exactly without any stretching. For future reference, call this mapping the *stretch mapping*. For example, the stretch mapping begins

$$
\langle 1,1\rangle,\ldots,\langle 1,k\rangle,\langle 2,k+1\rangle,\ldots,\\
\langle k+1,k+1\rangle,\langle k+2,k+2\rangle,\ldots .
$$

Since a total of four short line segments of length 1 are stretched to match four long line segments of length k, this mapping contains $4(k-1)$ stretch-edges. But since $r = 0$, the stretch-cost is 0. The distance-cost is 0 because each angle in *angles*(A) is mapped to the same angle in *angles*(B). The *NEM*-distance between shapes B and C is small (although not zero): in this case, the small triangular protrusion in shape B does not match the small square protrusion in shape C, although this mismatch occurs only in a small part of the boundary, so the distance is small. Specifically, *NEM*$(B, C) = 2(\pi/4) = \pi/2$. The upper bound *NEM*$(B, C) \le \pi/2$ is shown by the *no-stretch mapping* containing edges $\langle i, i\rangle$ for $1 \le i \le n$. However, *NEM*$(A, C) = k\pi/2$. The lower bound, *NEM*$(A, C) \ge k\pi/2$, holds because the angles $\pi/4$ and $7\pi/4$, occurring a total of $2k$ times in *angles*(A), differ by at least $\pi/4$ from every angle occurring in *angles*(C). The upper bound, *NEM*$(A, C) \le k\pi/2$, is shown by the no-stretch mapping. Since *NEM*(A, C) increases as k increases, whereas *NEM*(A, B) and *NEM*(B, C) are constant independent of k, the *NEM*-distance does not satisfy a relaxed triangle inequality (where the constant c is independent of the number of sample points).

It is instructive to see why the example of Figure 2 does not cause the relaxed triangle inequality to fail for *NEM*$_r$, like it does for *NEM*. For *NEM*$_r$, it is no longer true that the distance between A and B is zero; it is not even "small". If we do much stretching to make the triangular protrusions match at many points, then the distance includes a large term due to a large multiple of the stretching penalty. If, on the other hand, we do little stretching, then the distance includes a large term due to mismatch of tangent angles at many points. If we believe for aesthetic reasons that shapes A and B are not "close", then another advantage of *NEM*$_r$ over *NEM* (in addition to the advantage that *NEM*$_r$ satisfies a relaxed triangle inequality whereas *NEM* does not) is that *NEM*$_r$ better fits our aesthetic idea of "closeness" of shapes. Although *NEM*$_r$ satisfies a relaxed triangle inequality (as sketched in Section 4 and shown in the appendix), the shapes in Figure 2 show that it does not satisfy the triangle inequality if $r < \pi/8$. First, *NEM*$_r(A, B) \le 4(k-1)r$ is shown by the stretch mapping; the distance-cost of this mapping is still 0 as above, but its stretch-cost is now $4(k-1)r$. As above, *NEM*$_r(B, C) \le \pi/2$ is shown by the no-stretch mapping. But *NEM*$_r(A, C) \ge k\pi/2$, by the same argument given above for *NEM*. Using these bounds, it is easy to check that $NEM_r(A,C) > NEM_r(A,B) + NEM_r(B,C)$ if $r < \pi/8$.

Finally, we note that the version of the *NEM*-distance involving a stretching penalty as described in (McConnell et al., 1991) does not satisfy a relaxed triangle inequality. In this version, the stretching penalty r multiplies the distance-cost of a stretch-edge, instead of being added to it. So we need $r > 1$ in order that r impose a penalty. More formally, for sequences X and Y, a mapping M between them, and an edge $\langle i, j\rangle$ in M, define $cost'(\langle i,j\rangle, M, X, Y) = r \cdot b(x_i, y_j)$ if $\langle i, j\rangle$ is a stretch-edge of M, or $b(x_i, y_j)$ otherwise. Let $cost'(M, X, Y) = \sum_{e \in M} cost'(e, M, X, Y)$. Let $NEM'_r(X,Y)$ be the minimum $cost'$ of a mapping between X and Y. The shapes in Figure 2 show that if $r > 1$, then NEM'_r does not satisfy a relaxed triangle inequality. The reason is that $NEM'_r(A,B) = 0$, as shown by the stretch mapping. The distance-cost of all edges is 0 in the stretch mapping between A and B, so multiplying by r does not increase the cost. It is still true, as described above for *NEM*, that $NEM'_r(B,C) = \pi/2$ and $NEM'_r(A,C) \geq k\pi/2$. So a relaxed triangle inequality does not hold for NEM'_r.

4. The Relaxed Triangle Inequality

In this section we show that NEM_r satisfies a relaxed triangle inequality if $r > 0$ and if b_{sup} is finite. We consider first the case of equal-length sequences.

Theorem 1. *For any base (S, b), any real $r > 0$, any integer $n > 0$, and any three sequences X, Y, Z of length n,*

$$NEM_r(X,Z) \leq \left(1 + \frac{b_{sup}}{2r}\right)(NEM_r(X,Y) + NEM_r(Y,Z)).$$

Proof sketch: We outline the main steps of the proof. A full proof is given in the appendix.

The basic strategy is to take a mapping M_{XY} between X and Y having cost $NEM_r(X,Y)$, and a mapping M_{YZ} between Y and Z having cost $NEM_r(Y,Z)$, and paste them together in a certain way to obtain a mapping M_{XZ} between X and Z. The method of pasting together allows us to place an upper bound on the cost of M_{XZ} in terms of the cost of M_{XY} and M_{YZ}, that is, in terms of $NEM_r(X,Y)$ and $NEM_r(Y,Z)$. And once we have an upper bound on the cost of some mapping M_{XZ} between X and Z, we have an upper bound on $NEM_r(X,Z)$. As a simple example, suppose that the mappings M_{XY} and M_{YZ} have no stretch-edges; i.e., these mappings both consist of the edges $\langle i, i\rangle$ for $1 \leq i \leq n$. Then we take M_{XZ} to also consist of edges $\langle i, i\rangle$ for $1 \leq i \leq n$. Since the base distance b satisfies the triangle inequality (by assumption), it is easy to see that the distance-cost of M_{XZ} is at most the sum of the distance-cost of M_{XY} and the distance-cost of M_{YZ}. Since the stretch-cost of all three mappings is zero, we actually get the triangle inequality, $NEM_r(X,Z) \leq NEM_r(X,Y) + NEM_r(Y,Z)$, in this case. In general, however, the mappings M_{XY} and M_{YZ} can have stretch-edges, and this makes the construction of M_{XZ} and the bounding of its cost more complicated, and it also means that we do not get the triangle inequality in general.

Let M_{XY} and M_{YZ} be minimal (n,n)-mappings such that

$$cost(M_{XY}) = NEM_r(X,Y) \quad (2)$$

$$cost(M_{YZ}) = NEM_r(Y,Z). \quad (3)$$

Since we will be referring to edges in different mappings, for clarity we name the points of X, Y, Z using the notation $x[i], y[j], z[k]$, respectively, for $1 \leq i, j, k \leq n$. For example, an edge of M_{XY} has the form $\langle x[i], y[j]\rangle$ for some i and j.

To prove the relaxed triangle inequality, we construct a minimal (n,n)-mapping M_{XZ} and place an upper bound on $cost(M_{XZ})$. Since we want to use the fact that b satisfies the triangle inequality to help us bound the distance-cost of M_{XZ}, we want M_{XZ} to be a minimal (n,n)-mapping with the following "midpoint property": For every edge $\langle x[i], z[k]\rangle \in M_{XZ}$, there is a "midpoint" $y[j]$ such that $\langle x[i], y[j]\rangle \in M_{XY}$ and $\langle y[j], z[k]\rangle \in M_{YZ}$. Then, the distance-cost of the edge $\langle x[i], z[k]\rangle$ is at most the sum of the distance-costs of $\langle x[i], y[j]\rangle$ and $\langle y[j], z[k]\rangle$.

The first step is to show that some M_{XZ} with the midpoint property exists. This is done in the appendix by describing a construction of one such mapping by adding edges one at a time, such that each added edge has a midpoint.

To bound the cost of M_{XZ}, it is useful to divide the stretch-edges of a mapping into two classes, depending on which sequence receives the stretching. For M_{XY}, the stretch-edge $\langle x[i], y[j]\rangle$ is an *X-stretch-edge* if $\langle x[i-1], y[j]\rangle \in M_{XY}$, or a *$Y$-stretch-edge* if $\langle x[i], y[j-1]\rangle \in M_{XY}$ (since edges cannot cross, exactly one of these holds). For M_{YZ}, the stretch-edges are divided similarly into Y-stretch-edges and Z-stretch-edges. It is also useful to divide the stretch-cost of a mapping into two parts, based on this division of the stretch-edges, as follows. Define X-*s-cost*(M_{XY}) (resp., Y-*s-cost*(M_{XY})) to be r times the number of X-stretch-edges (resp., Y-stretch-edges) of M_{XY}. Similarly define Y-*s-cost*(M_{YZ}) and Z-*s-cost*(M_{YZ}). Since X and Y have the same length, the number of X-stretch-edges of M_{XY} equals the number of Y-stretch-edges of M_{XY}. Therefore, we have the following equalities involving the stretch-cost *s-cost*:

$$\begin{aligned} X\text{-}s\text{-}cost(M_{XY}) &= Y\text{-}s\text{-}cost(M_{XY}) \\ &= s\text{-}cost(M_{XY})/2. \end{aligned} \quad (4)$$

Similarly, since Y and Z have the same length,

$$\begin{aligned} Y\text{-}s\text{-}cost(M_{YZ}) &= Z\text{-}s\text{-}cost(M_{YZ}) \\ &= s\text{-}cost(M_{YZ})/2. \end{aligned} \quad (5)$$

To prove the relaxed triangle inequality, it suffices to prove the following two bounds on the stretch-cost *s-cost* and the distance-cost *d-cost* of M_{XZ}.

Claim 1.

$$s\text{-}cost(M_{XZ}) \leq s\text{-}cost(M_{XY}) + s\text{-}cost(M_{YZ}).$$

Claim 2.

$$\begin{aligned} d\text{-}cost(M_{XZ}) &\leq d\text{-}cost(M_{XY}) + d\text{-}cost(M_{YZ}) \\ &\quad + \frac{b_{sup}}{r}(X\text{-}s\text{-}cost(M_{XY}) + Z\text{-}s\text{-}cost(M_{YZ})). \end{aligned}$$

The relaxed triangle inequality stated in the theorem follows by algebraic manipulation from these two claims and (2), (3), (4), and (5).

To justify Claim 1, with each stretch-edge in M_{XZ} we associate a distinct stretch-edge in either M_{XY} or M_{YZ}. Clearly such an association (which is given in the appendix) suffices to prove Claim 1.

The final step is to justify Claim 2. Since we know that M_{XZ} has the midpoint property, we would like to use the fact that b satisfies the triangle inequality. A complication is shown by the situation in Figure 3 where the distance-cost of $\langle x[i], y[j] \rangle$ contributes t times to the distance-cost of M_{XZ}. The key observation in handling this complication is that each of the $t-1$ contributions of *d-cost*$(\langle x[i], y[j] \rangle)$ after the first contribution can be "balanced" by a Z-stretch-edge of M_{YZ} that contributes r to the stretch-cost of M_{YZ}. There is a symmetric case where an edge in

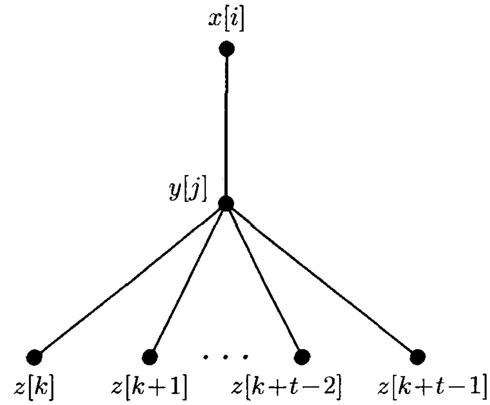

Fig. 3. A situation where the distance-cost of $\langle x[i], y[j] \rangle$ contributes t times to the distance-cost of M_{XZ}.

M_{YZ} contributes several times to the distance-cost of M_{XZ}, and the symmetric case is handled similarly, using X-stretch-edges of M_{XY} for the balancing. For more details, see the appendix. □

Remark. We suggest two ways that the relaxed triangle inequality might be improved. First, Claims 1 and 2 and (2), (3), (4), and (5) actually give the potentially tighter bound

$$\begin{aligned} NEM_r(X,Z) &\leq NEM_r(X,Y) + NEM_r(Y,Z) \\ &\quad + \frac{b_{sup}}{2r}(s\text{-}cost(M_{XY}) + s\text{-}cost(M_{YZ})) \end{aligned}$$

where M_{XY} and M_{YZ} are any mappings with $NEM_r(X,Y) = cost(M_{XY})$ and $NEM_r(Y,Z) = cost(M_{YZ})$. Therefore, in the application to image databases mentioned in the introduction, it might be advantageous in the clustering of database shapes to keep track of the stretch-cost of mappings as well as their total cost. It is easy to modify the dynamic programming algorithm to compute, together with the minimum total cost of a mapping, the minimum stretch-cost of a mapping among the mappings having minimum total cost.

Second, in the proof of Claim 2, we use b_{sup} as an upper bound on the distance-cost of any edge in M_{XY} and M_{YZ}. Therefore, another way to improve the relaxed triangle inequality in practice would be to replace the gross upper bound b_{sup} by the actual maximum distance-cost of edges in M_{XY} and M_{YZ}. This would require computing and storing these maximum distance-costs during the clustering preprocessing.

Although the remark above shows that we might get a better bound on $NEM_r(X, Z)$ in certain cases, the next result shows that the constant $(1 + b_{sup}/(2r))$ in the general relaxed triangle inequality is essentially the best possible. The proof is given in the appendix.

Theorem 2. *For any base (S, b) with $b_{sup} > 0$, any real $r > 0$, and any real $\varepsilon > 0$, there is an integer n and three sequences X, Y, Z of length n such that*

$$NEM_r(X, Z) \geq (1 + \frac{b_{sup}}{2r} - \varepsilon)(NEM_r(X, Y) + NEM_r(Y, Z)).$$

We now give analogues of Theorems 1 and 2 for the case of unequal-length sequences. The results are similar, except that the constant in the relaxed triangle inequality increases to $(1 + b_{sup}/r)$. The proofs in the unequal-length case are very similar to the proofs in the equal-length case. The differences are outlined in the appendix.

Theorem 3. *For any base (S, b), any real $r > 0$, and any three sequences X, Y, Z,*

$$NEM_r(X, Z) \leq (1 + \frac{b_{sup}}{r})(NEM_r(X, Y) + NEM_r(Y, Z)).$$

Theorem 4. *For any base (S, b) with $b_{sup} > 0$, any real $r > 0$, and any real $\varepsilon > 0$, there are three sequences X, Y, Z such that*

$$NEM_r(X, Z) \geq (1 + \frac{b_{sup}}{r} - \varepsilon)(NEM_r(X, Y) + NEM_r(Y, Z)).$$

Appendix

In this appendix, we prove Theorems 1, 2, 3, and 4.

Proof of Theorem 1: Let M_{XY} and M_{YZ} be minimal (n, n)-mappings such that

$$cost(M_{XY}) = NEM_r(X, Y) \quad \text{(A1)}$$
$$cost(M_{YZ}) = NEM_r(Y, Z). \quad \text{(A2)}$$

(Here we have abbreviated $cost(M_{XY}, X, Y)$ by $cost(M_{XY})$ and $cost(M_{YZ}, Y, Z)$ by $cost(M_{YZ})$, since the relevant sequences are clear from the name of the mapping. Similar abbreviations are made throughout the proof.) Since we will be referring to edges in different mappings, for clarity we name the points of X, Y, Z using the notation $x[i], y[j], z[k]$, respectively, for $1 \leq i, j, k \leq n$. For example, an edge of M_{XY} has the form $\langle x[i], y[j]\rangle$ for some i and j.

To prove the relaxed triangle inequality, we construct a minimal (n, n)-mapping M_{XZ} and place an upper bound on $cost(M_{XZ})$. The mapping M_{XZ} can be any minimal (n, n)-mapping with the following "midpoint property": If $\langle x[i], z[k]\rangle$ is an edge of M_{XZ}, say that $y[j]$ is a *midpoint* of $\langle x[i], z[k]\rangle$ if $\langle x[i], y[j]\rangle \in M_{XY}$ and $\langle y[j], z[k]\rangle \in M_{YZ}$. A mapping M_{XZ} has the *midpoint property* if every edge of M_{XZ} has at least one midpoint.

The first step is to show that some M_{XZ} with this property exists. We show how to construct one such mapping by adding edges one at a time, such that each added edge has a midpoint. Begin by adding $\langle x[1], z[1]\rangle$ to M_{XZ}. By the definition of a mapping, it must be that $\langle x[1], y[1]\rangle \in M_{XY}$ and $\langle y[1], z[1]\rangle \in M_{YZ}$, so $y[1]$ is a midpoint of $\langle x[1], z[1]\rangle$. To describe how to continue the edge-adding procedure, let $\langle x[i], z[k]\rangle$ be the edge last added to M_{XZ}, and let $y[j]$ be a midpoint of $\langle x[i], z[k]\rangle$, that is, $\langle x[i], y[j]\rangle \in M_{XY}$ and $\langle y[j], z[k]\rangle \in M_{YZ}$. Consider first the case that $i < n$ and $k < n$. We show that at least one of the edges $\langle x[i+1], z[k]\rangle$, $\langle x[i], z[k+1]\rangle$, or $\langle x[i+1], z[k+1]\rangle$ has a midpoint, so it can be added to M_{XZ}.

Case 1. $\langle x[i+1], y[j]\rangle \in M_{XY}$.

If Case 1 holds, then $y[j]$ is a midpoint of $\langle x[i+1], z[k]\rangle$, so we can add $\langle x[i+1], z[k]\rangle$ to M_{XZ}.

Case 2. $\langle y[j], z[k+1]\rangle \in M_{YZ}$.

If Case 2 holds, then $y[j]$ is a midpoint of $\langle x[i], z[k+1]\rangle$, so we can add $\langle x[i], z[k+1]\rangle$ to M_{XZ}.

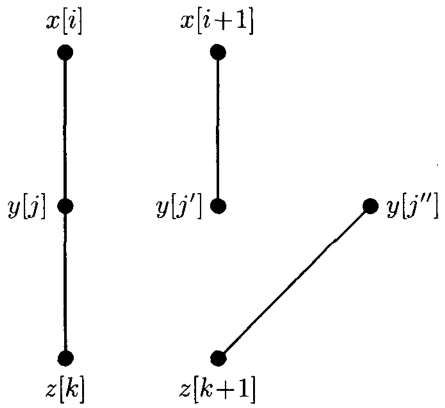

Fig. A.1. A case in the edge-adding procedure.

So suppose that neither Case 1 nor Case 2 holds. Let j' be the smallest integer such that $\langle x[i+1], y[j']\rangle \in M_{XY}$. Since Case 1 does not hold, since $\langle x[i], y[j]\rangle \in M_{XY}$, and since edges of M_{XY} do not cross, it follows that $j' > j$. See Figure A.1. Similarly, using the fact that Case 2 does not hold, if j'' is the smallest integer such that $\langle y[j''], z[k+1]\rangle \in M_{YZ}$, then $j'' > j$. If $j' = j''$, then we can add $\langle x[i+1], z[k+1]\rangle$ to M_{XZ} since it has a midpoint $y[j']$. So say that $j' \neq j''$, and assume without loss of generality that $j' < j''$; again see Figure A.1. By minimality of j'', it follows that $\langle y[j'], z[k+1]\rangle \notin M_{YZ}$. Since $y[j']$ must belong to some edge of M_{YZ}, since $\langle y[j'], z[k+1]\rangle \notin M_{YZ}$, and since edges cannot cross, it must be that $\langle y[j'], z[k]\rangle \in M_{YZ}$. So $y[j']$ is a midpoint of $\langle x[i+1], z[k]\rangle$, and we can add $\langle x[i+1], z[k]\rangle$ to M_{XZ}. This completes the case that $i < n$ and $k < n$. The cases where one of i or k is equal to n and the other is less than n are similar to the above (and simpler), and these cases are left to the reader. By continuing the edge-adding procedure we eventually reach an (n, n)-mapping M_{XZ}. If the mapping M_{XZ} constructed in this way is not minimal, then remove edges until a minimal mapping is reached.

To bound the cost of M_{XZ}, it is useful to divide the stretch-edges of a mapping into two classes, depending on which sequence receives the stretching. For M_{XY}, the stretch-edge $\langle x[i], y[j]\rangle$ is an *X-stretch-edge* if $\langle x[i-1], y[j]\rangle \in M_{XY}$, or a *$Y$-stretch-edge* if $\langle x[i], y[j-1]\rangle \in M_{XY}$ (since edges cannot cross, exactly one of these holds). For M_{YZ}, the stretch-edges are divided similarly into Y-stretch-edges and Z-stretch-edges. Define X-*s-cost*(M_{XY}) (resp., Y-*s-cost*(M_{XY})) to be r times the number of X-stretch-edges (resp., Y-stretch-edges) of M_{XY}. Similarly define Y-*s-cost*(M_{YZ}) and Z-*s-cost*(M_{YZ}). Clearly,

$$\textit{s-cost}(M_{XY}) = X\text{-}\textit{s-cost}(M_{XY}) + Y\text{-}\textit{s-cost}(M_{XY}).$$

Since X and Y have the same length, the number of X-stretch-edges of M_{XY} equals the number of Y-stretch-edges of M_{XY}. Therefore,

$$\begin{aligned} X\text{-}\textit{s-cost}(M_{XY}) &= Y\text{-}\textit{s-cost}(M_{XY}) \\ &= \textit{s-cost}(M_{XY})/2. \end{aligned} \tag{A3}$$

Similarly,

$$\begin{aligned} Y\text{-}\textit{s-cost}(M_{YZ}) &= Z\text{-}\textit{s-cost}(M_{YZ}) \\ &= \textit{s-cost}(M_{YZ})/2. \end{aligned} \tag{A4}$$

As we shall show, to prove the relaxed triangle inequality, it suffices to prove the following two bounds on the *s-cost* and the *d-cost* of M_{XZ}.

Claim 1.

$$\textit{s-cost}(M_{XZ}) \le \textit{s-cost}(M_{XY}) + \textit{s-cost}(M_{YZ}).$$

Claim 2.

$$\begin{aligned} \textit{d-cost}(M_{XZ}) \le\ & \textit{d-cost}(M_{XY}) + \textit{d-cost}(M_{YZ}) \\ & + \frac{b_{sup}}{r}(X\text{-}\textit{s-cost}(M_{XY}) + Z\text{-}\textit{s-cost}(M_{YZ})). \end{aligned}$$

Before explaining why these two inequalities hold, we first show that they lead to the relaxed triangle inequality stated in the theorem. Adding the left and right sides of the two inequalities, using that $\textit{cost}(M) = \textit{s-cost}(M) + \textit{d-cost}(M)$ for any mapping M, and using (A1), (A2), (A3), and (A4), gives

$$\begin{aligned} \textit{cost}(M_{XZ}) &\le \textit{cost}(M_{XY}) + \textit{cost}(M_{YZ}) \\ &\quad + \frac{b_{sup}}{2r}(\textit{s-cost}(M_{XY}) + \textit{s-cost}(M_{YZ})) \\ &\le (1 + \frac{b_{sup}}{2r})(\textit{cost}(M_{XY}) + \textit{cost}(M_{YZ})) \\ &= (1 + \frac{b_{sup}}{2r})(\textit{NEM}_r(X,Y) + \textit{NEM}_r(Y,Z)). \end{aligned}$$

Therefore,

$$\textit{NEM}_r(X,Z) \le (1 + \frac{b_{sup}}{2r})(\textit{NEM}_r(X,Y) + \textit{NEM}_r(Y,Z)),$$

as desired.

Some additional terminology will be useful in justifying the two claims. For each edge $e \in M_{XZ}$, let *mid*(e) be some midpoint of e; if e has more than one midpoint, the choice can be arbitrary. If $e = \langle x[i], z[k]\rangle$ and $y[j] = \textit{mid}(e)$, then let *first*$(e) = \langle x[i], y[j]\rangle$ and *second*$(e) = \langle y[j], z[k]\rangle$.

The next step is to justify Claim 1. With each stretch-edge in M_{XZ} we shall associate some stretch-edge *assoc*(e) in either M_{XY} or M_{YZ}, so that *assoc* is injective; that is, for every two distinct stretch-edges e and e' of M_{XZ} we have $\textit{assoc}(e) \neq \textit{assoc}(e')$. Clearly such an association suffices to prove Claim 1. It turns out that *assoc*(e) is either *first*(e) or *second*(e). Let $e = \langle x[i], z[k]\rangle$ be a Z-stretch-edge of M_{XZ}, so that $e' = \langle x[i], z[k-1]\rangle \in M_{XZ}$. Let $y[j] = \textit{mid}(e)$ and $y[j'] = \textit{mid}(e')$. If $j = j'$, then *assoc*(e) is the Z-stretch-edge $\textit{second}(e) = \langle y[j], z[k]\rangle$ in M_{YZ}. If $j \neq j'$, then we must have $j' < j$ and $\langle x[i], y[j]\rangle$ must be a Y-stretch-edge of M_{XY}. In this case, *assoc*(e) is the Y-stretch-edge $\textit{first}(e) = \langle x[i], y[j]\rangle$ in M_{XY}. If e is an X-stretch-edge of M_{XZ}, then in a completely symmetric way, either $\textit{assoc}(e) = \textit{first}(e)$ and *first*(e) is an X-stretch-edge in M_{XY}, or $\textit{assoc}(e) = \textit{second}(e)$ and *second*(e) is a Y-stretch-edge in M_{YZ}. It is easy to see that *assoc* is injective.

For readers who would like a formal argument, one follows. First, if e is an X-stretch-edge of M_{XZ} and e' is a Z-stretch-edge of M_{XZ}, then we cannot have $\textit{assoc}(e) = \textit{assoc}(e')$, since *assoc*$(e)$ is either an X-stretch-edge of M_{XY} or a Y-stretch-edge of M_{YZ}, whereas *assoc*(e') is either a Y-stretch-edge of M_{XY} or a Z-stretch-edge of M_{YZ}. So let e and e' be distinct Z-stretch-edges of M_{XZ} and suppose for contradiction that $\textit{assoc}(e) = \textit{assoc}(e')$. (The case where e and e' are both X-stretch-edges is symmetric.) We must have either $\textit{first}(e) = \textit{first}(e')$ or $\textit{second}(e) = \textit{second}(e')$, and in either case $\textit{mid}(e) = \textit{mid}(e')$. Since e and e' are Z-stretch-edges, it must be that $\textit{first}(e) = \textit{first}(e')$ and $\textit{second}(e) \neq \textit{second}(e')$. Let $e = \langle x[i], z[k]\rangle$ and $e' = \langle x[i], z[k']\rangle$. Since $e \neq e'$, it follows that $k \neq k'$; without loss of generality, assume that $k > k'$. Then by definition of *assoc* we would take $\textit{assoc}(e) = \textit{second}(e)$. Therefore, we cannot have $\textit{assoc}(e) = \textit{assoc}(e')$.

The final step is to justify Claim 2. Using the fact that b satisfies the triangle inequality,

$$
\begin{aligned}
\textit{d-cost}(M_{XZ}) &= \sum_{e \in M_{XZ}} \textit{d-cost}(e) \\
&\leq \sum_{e \in M_{XZ}} (\textit{d-cost}(\textit{first}(e)) + \textit{d-cost}(\textit{second}(e))).
\end{aligned} \tag{A5}
$$

If each edge in M_{XY} and M_{YZ} appeared at most once in the sum in (A5), as either *first*(e) or *second*(e) for at most one e, then $\textit{NEM}_r$ would satisfy the triangle inequality since we could conclude that $\textit{d-cost}(M_{XZ}) \leq \textit{d-cost}(M_{XY}) + \textit{d-cost}(M_{YZ})$. However, as shown in Figure A.2, the same edge of M_{XY} or M_{YZ} can appear several times in the sum.

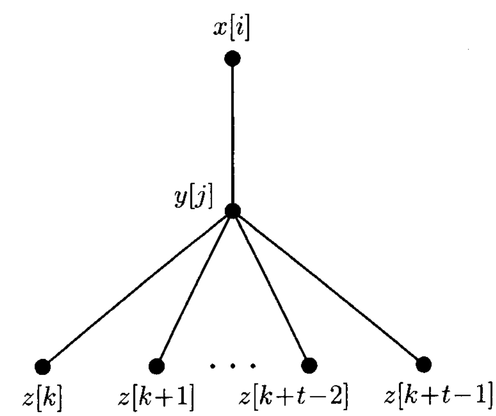

Fig. A.2. A situation where the distance-cost of $\langle x[i], y[j]\rangle$ contributes t times to the distance-cost of M_{XZ}.

This figure shows a situation where an edge $\langle x[i], y[j]\rangle$ in M_{XY} appears t times as *first*(e) for t edges $e = \langle x[i], z[k+l]\rangle$ for $0 \leq l \leq t-1$. There is a symmetric case where an edge in M_{YZ} appears several times as *second*(e) for several e's. We focus on the case shown in Figure A.2, the symmetric case being handled similarly. The key observation is that each of the $t-1$ occurrences of $\langle x[i], y[j]\rangle$ after the first occurrence can be "balanced" by a Z-stretch-edge of M_{YZ}. Break the sum in (A5) into pieces, each piece corresponding to a situation like the one shown in Figure A.2, or a symmetric situation. Focusing on the piece of the sum corresponding to Figure A.2,

$$\begin{aligned}
&\sum_{l=0}^{t-1}(d\text{-}cost(\langle x[i], y[j]\rangle) + d\text{-}cost(\langle y[j], z[k+l]\rangle)) \\
&\quad \leq\ d\text{-}cost(\langle x[i], y[j]\rangle) + (t-1)b_{sup} \\
&\qquad + \sum_{l=0}^{t-1} d\text{-}cost(\langle y[j], z[k+l]\rangle) \\
&\quad =\ d\text{-}cost(\langle x[i], y[j]\rangle) + \sum_{l=0}^{t-1} d\text{-}cost(\langle y[j], z[k+l]\rangle) \\
&\qquad + \frac{b_{sup}}{r}\sum_{l=1}^{t-1} s\text{-}cost(\langle y[j], z[k+l]\rangle). \qquad \text{(A6)}
\end{aligned}$$

(We used the fact that $s\text{-}cost(\langle y[j], z[k+l]\rangle) = r$ for $1 \leq l \leq t-1$.) In the symmetric case where an edge in M_{YZ} appears several times as *second*(e), the calculation is similar, except that the edges appearing as arguments of *s-cost* in (A6) are X-stretch-edges of M_{XY}, rather than Z-stretch-edges of M_{YZ}.

After each piece of the sum in (A5) is replaced by an upper bound of the form in (A6), each edge e in M_{XY} or M_{YZ} appears at most once in a term of the form $d\text{-}cost(e)$, each X-stretch-edge of M_{XY} and each Z-stretch-edge of M_{YZ} appears at most once in a term of the form $s\text{-}cost(e)$, and each Y-stretch-edge of M_{XY} and each Y-stretch-edge of M_{YZ} does not appear at all. From this it is easy to see that Claim 2 holds. □

Proof of Theorem 2: For simplicity, suppose there are $x, y \in S$ with $b(x,y) = b_{sup}$. (If (S,b) is such that the supremum is not achieved, the proof is similar since $b(x,y)$ can be made arbitrarily close to b_{sup}.) For a sequence σ and a positive integer p, let σ^p denote $\sigma, \sigma, \ldots, \sigma$ where σ is repeated p times. For sufficiently large integers p and q, the three sequences are:

$$\begin{aligned}
X &= x^{p+1}, (y^p, x^p)^q, x \\
Y &= x, (y^p, x^p)^q, x^{p+1} \\
Z &= y, (y^p, x^p)^q, x^{p+1}.
\end{aligned}$$

First note that $NEM_r(X,Y) \leq 2pr$. This bound is shown by the mapping M_{XY} that maps the first $p+1$ occurrences of x in X to the first occurrence of x in Y, maps the subsequence $(y^p, x^p)^q$ of X to the same subsequence in Y, and maps the last occurrence of x in X to the last $p+1$ occurrences of x in Y. Therefore, $s\text{-}cost(M_{XY}) = 2pr$ and $d\text{-}cost(M_{XY}) = 0$. Second note that $NEM_r(Y,Z) \leq b_{sup}$. This is shown by the mapping M_{YZ} that does no stretching, so that $s\text{-}cost(M_{YZ}) = 0$ and $d\text{-}cost(M_{YZ}) = b(x,y) = b_{sup}$.

For each fixed p, we now show that $NEM_r(X,Z) = 2pr + (p+1)b_{sup}$ for all sufficiently large q. The upper bound $2pr + (p+1)b_{sup}$ is shown by the mapping identical to the mapping M_{XY} above (except that Y is replaced by Z). The stretch-cost of this mapping is again $2pr$. The distance-cost is now $(p+1)b_{sup}$ since the first $p+1$ occurrences of x in X are mapped to the first occurrence of y in Z. To show that $2pr + (p+1)b_{sup}$ is also a lower bound, let M_{XZ} be an (n,n)-mapping such that $cost(M_{XZ}) = NEM_r(X,Z)$. If $s\text{-}cost(M_{XZ}) < 2pr$, then the mapping does not do enough stretching to align the subsequence $(y^p, x^p)^q$ of X with the same subsequence appearing in Z. Therefore, if $s\text{-}cost(M_{XZ}) < 2pr$, then $d\text{-}cost(M_{XZ}) \geq qb_{sup}$, so M_{XZ} cannot have minimum cost for large enough q, as its cost would exceed the upper bound just shown. So we can assume that $s\text{-}cost(M_{XZ}) \geq 2pr$. Since X begins with x^{p+1} and Z begins with y^{p+1}, at least $p+1$ edges in M_{XZ} must have distance-cost b_{sup}. Therefore, $NEM_r(X,Z) \geq 2pr + (p+1)b_{sup}$.

Using the bounds just derived,

$$\frac{NEM_r(X,Z)}{NEM_r(X,Y) + NEM_r(Y,Z)} \geq \frac{2pr + (p+1)b_{sup}}{2pr + b_{sup}}.$$

This fraction approaches $(1 + b_{sup}/(2r))$ as p approaches infinity. □

Proof of Theorem 3: In the proof of Theorem 1, when constructing the mapping M_{XZ} and proving Claims 1 and 2, we did not use that the sequences are of equal length, so these parts of the proof are unchanged. The only place where we used that the sequences are of equal length was in equations (A3) and (A4). However, in the present case we have $X\text{-}s\text{-}cost(M_{XY}) \leq s\text{-}cost(M_{XY})$ and $Z\text{-}s\text{-}cost(M_{YZ}) \leq s\text{-}cost(M_{YZ})$. Using these inequalities in place of (A3) and (A4) in the calculation following the statement of Claims 1 and 2 gives the result. □

Proof of Theorem 4: The proof is similar to that of Theorem 2, but using the sequences

$$\begin{aligned}
X &= x^{p+1}, (y^p, x^p)^q \\
Y &= x, (y^p, x^p)^q \\
Z &= y, (y^p, x^p)^q.
\end{aligned}$$

Arguing as before, it can be shown that for q sufficiently large,

$$\begin{aligned} NEM_r(X, Y) &\leq pr, \\ NEM_r(Y, Z) &\leq b_{sup}, \\ NEM_r(X, Z) &= pr + (p+1)b_{sup}. \end{aligned}$$

The result follows as before. □

Acknowledgements

We are grateful to Byron Dom, Martin Farach, Myron Flickner, Wayne Niblack, Prabhakar Raghavan, and Baruch Schieber for helpful discussions and comments.

Notes

1. QBIC is a trademark of IBM Corporation.

References

Arkin, E.M., Chew, L.P., Huttenlocher, D.P., Kedem, K., and Mitchell, J.S.B. 1990. An efficiently computable metric for comparing polygonal shapes. In *Proc. First ACM-SIAM Symp. on Discrete Algorithms*, San Francisco, CA, pp. 129–137.

Cortelazzo, G., Mian, G.A., Vezzi, G., and Zamperoni, P. 1994. Trademark shapes description by string-matching techniques. *Pattern Recognition* 27(8):1005–1018.

Huttenlocher, D.P., Rucklidge, W.J., and Klanderman, G.A. 1992. Comparing images using the Hausdorff distance under translation. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Champaign, IL, pp. 654–656.

Jain, A.K. 1989. *Fundamentals of Digital Image Processing*. Prentice-Hall: Englewood Cliffs, NJ.

Kim Y-S. and Kim, W-Y. 1997. Content-based trademark retrieval system using visually salient feature. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, pp. 307–312.

McConnell, R., Kwok, R., Curlander, J.C., Kober, W., and Pang, S.S. 1991. Ψ–S correlation and dynamic time warping: two methods for tracking ice floes in SAR images. *IEEE Trans. Geoscience and Remote Sensing* 29(6):1004–1012.

Mehtre, B.M., Kankanhalli, M.S., and Lee, W.F. 1997. Shape measures for content based image retrieval: a comparison. *Information Processing and Management* 33(3):319–337.

Mumford, D. 1991. Mathematical theories of shape: do they model perception? In *Proc. Conf. on Geometric Methods in Computer Vision*, San Diego, CA, SPIE volume 1570, pp. 2–10.

Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., and Yanker, P. 1993. The QBIC project: querying images by content using color, texture and shape. In *Proc. Conf. on Storage and Retrieval for Image and Video Databases*, San Jose, CA, SPIE volume 1908, pp. 173–181. QBIC Web server is `http://wwwqbic.almaden.ibm.com/`.

Niblack, W. and Yin, J. 1995. A pseudo-distance measure for 2D shapes based on turning angle. In *Proc. IEEE Int. Conf. on Image Processing*, Washington, DC.

Scassellati, B., Alexopoulos, S., and Flickner, M. 1994. Retrieving images by 2D shape: a comparison of computation methods with human perceptual judgments. In *Proc. Conf. on Storage and Retrieval for Image and Video Databases II*, San Jose, CA, SPIE volume 2185, pp. 2–14.

Taubin, G. and Cooper, D.B. 1991. Recognition and positioning of rigid objects using algebraic moment invariants. In *Proc. Conf. on Geometric Methods in Computer Vision*, San Diego, CA, SPIE volume 1570, pp. 175–186.