Int. J. Contemp. Math. Sciences, Vol. 1, 2006, no. 15, 727 - 751

A Measure of Shape Dissimilarity for 3D Curves

Ernesto Bribiesca[1]

Department of Computer Science
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
Apdo. 20-726, México, D.F., 01000. Fax: (5255)5622-3620
ernesto@leibniz.iimas.unam.mx

Wendy Aguilar

Department of Computer Science
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
Apdo. 20-726, México, D.F., 01000. Fax: (5255)5622-3620
weam@turing.iimas.unam.mx

Abstract

We present a quantitative approach to the measurement of shape dissimilarity between two 3D (three-dimensional) curves. Any 3D continuous curve can be digitalized and represented as a 3D discrete curve. Thus, a 3D discrete curve is composed of constant orthogonal straight-line segments. In order to represent 3D discrete curves, we use the orthogonal direction change chain code. The chain elements represent the orthogonal direction changes of the contiguous straight-line segments of the discrete curve. This chain code only considers relative direction changes, which allows us to have a curve descriptor invariant under *translation* and *rotation*. Also, this curve descriptor may be *starting point normalized* and *mirroring curves* may be obtained with ease. Thus, using the above-mentioned chain code it is possible to have a unique 3D-curve descriptor.

To find out how close in shape two 3D curves are, a measure of shape-of-curve dissimilarity between them is introduced; analogous curves will have a low measure of shape dissimilarity, while different curves will have a high measure of shape dissimilarity. When this measure of shape dissimilarity is normalized, its values vary continuously from 0 to 1. If

[1] Author to whom correspondence should be addressed

two curves are identical, the value of the measure of shape dissimilarity is equal to 0. The computation of this measure for two curves is based on the analysis of their common and different subcurves represented by their chain elements. Finally, we present some results of the computation of the proposed measure for 15 curves.

Mathematics Subject Classification: 65D17

Keywords: Shape-of-curve dissimilarity, 3D discrete curves, chain coding, measure of shape dissimilarity, 3D curve representation

1 Introduction

The study of the representation and comparison of 3D curves is an important topic in different fields. This paper gives a procedure to measure the resemblance between any two 3D curves. With the help of procedures like this, a quantitative study of shape-of-curve dissimilarity may be possible. 3D curves are represented by means of chain coding. Chain-code methods are widely used because they preserve information and allow considerable data reduction. Chain codes are the standard input format for numerous shape analysis algorithms. Several shape features may be computed directly from the chain-code representation [8]. The first approach for representing 3D digital curves using chain code was introduced by Freeman in 1974 [7]. A canonical shape description for 3D stick bodies has been defined by Guzmán [10]. In the content of this work, we use the orthogonal direction change chain code [4] for representing 3D curves. The main characteristics of this chain code are the following: (1) it is invariant under *translation* and *rotation*; (2) optionally, it may be *starting point normalized* and *mirroring curves* may be obtained with ease; (3) there are only five possible orthogonal direction changes for representing any 3D curve, which produces a numerical string of finite length over a finite alphabet, which allows us to use grammatical techniques for 3D-curve analysis. Thus, using the above-mentioned chain code it is possible to have a unique 3D-curve descriptor.

The Hausdorff distance plays an important role in curve resemblance. Belogay *et al.* [3] propose a method for computing the Hausdorff distance between curves. As a measure for the resemblance of curves in arbitrary dimensions Alt and Godau [1] consider the so-called Frechet-distance, which is compatible with parametrizations of the curves. Arkin *et al.* [2] present an efficiently computable metric for comparing polygonal shapes based on turning angle representation.

Several authors have proposed different measures of shape dissimilarity for 3D curves. Li [13] proposes a system for matching and pose estimation of 3D curves under similarity transformation composed of translation, rotation and uniform scaling. Lo and Don [15] describe two invariant representations for 3D curves. One represents 3D curves by complex waveforms. The other illustrates 3D curves using 3D moment invariants of the data points on the curves.

Other methods are focused on 3D shape recognition, for instance: Jain and Hoffman [12] define a similarity measure between the set of observed features and the set of evidence conditions for a given object in a database. Dickinson *et al.* [6] present some techniques for recognition of 3D objects from a single 2D image. In such techniques, from an input image, a set of features or primitives may be extracted. Lohmann [16] considers a similarity measure based on the quotients of volumes of the studied 3D objects over well-known geometrical objects. Holden *et al.* [11] have evaluated eight different similarity measures applied to 3D serial magnetic resonance images. Recently, Rodriguez *et al.* [17] presented a new approach to measuring the similarity between 3D curves, this approach allows the possibility of using strings.

We present a measure of shape dissimilarity for 3D curves. In order to represent 3D curves, we use the orthogonal direction change chain code. Thus, using this chain code we obtain a unique curve descriptor represented by a chain. The measure of shape dissimilarity considers that two curves are more similar when they have in common more subcurves (one-to-one matching), and when these subcurves have the same orientation and position inside their curves. This paper is organized as follows. In Section 2 we summarize the orthogonal direction change chain code. In Section 3 we describe the proposed measure of shape dissimilarity. Section 4 gives some results. Finally, in Section 5 we present some conclusions.

2 The orthogonal direction change chain code

The purpose of this section is to summarize, in part, the orthogonal direction change chain code which was presented in [4]. The concepts of this chain code are important for describing the proposed measure of dissimilarity. Any 3D continuous curve can be digitalized and represented as a 3D discrete curve. Thus, a 3D discrete curve is composed of constant orthogonal straight-line segments. The chain elements represent the orthogonal direction changes of the constant straight-line segments of the 3D discrete curve. A chain A is an ordered sequence of elements, and is represented by $A = a_1 a_2 \ldots a_n = \{a_i : 1 \leq i \leq n\}$, where n indicates the number of chain elements. A chain element a_i indicates the orthogonal direction changes of the contiguous straight-line segments of the 3D discrete curve in that chain-element position. In the context of this work, the length l of each straight-line segment of any 3D curve is considered equal to one.

Each element of the chain labels a vertex of the 3D curve and indicates the orthogonal direction changes of the polygonal path in such a vertex. There are only five possible orthogonal direction changes for representing any 3D curve (figures 1(a)-(e) show these orthogonal direction changes):

1. The chain element "0" represents the direction change which *goes straight* through the contiguous straight-line segments following the direction of the last segment. This element is presented in Fig. 1(a).
2. The chain element "1" indicates a direction change to the *right* and is shown in Fig. 1(b)
3. The chain element "2" represents a direction change *upward* (stair-case fashion). This chain element is illustrated in Fig. 1(c).
4. The chain element "3" indicates a direction change to the *left* and is shown in Fig. 1(d).
5. The chain element "4" represents a direction change which is *going back* and is illustrated in Fig. 1(e).

Formally [5], if the consecutive sides of the reference angle have respective directions u and v (see Fig. 1(a)), and the side from the vertex to be labeled has direction w (from here on by direction we understand a vector of length 1), then the *chain element* is given by the following function,

$$\textit{chain element}(u,v,w) = \begin{cases} 0, & \text{if } w = v; \\ 1, & \text{if } w = u \times v; \\ 2, & \text{if } w = u; \\ 3, & \text{if } w = -(u \times v); \\ 4, & \text{if } w = -u; \end{cases}$$

where $\times$ denotes the cross product.

Fig. 1(f) illustrates an example of a continuous 3D curve. In order to observe curves in a three-dimensional way, they are represented as ropes. Fig. 1(g) shows the discrete version of the curve presented in (f) using only orthogonal directions. The origin of this curve is considered at the lower side. Fig. 1(h) shows the first element of the chain which corresponds to the element "3". Notice that the first direction change (which is composed of two contiguous straight-line segments) is used only for reference. Fig. 1(i) shows the next element obtained of the chain, which is based on the last direction change of the first element; the second element corresponds to the element "3", too. Fig. 1(j) show the next element obtained of the chain. Fig. 1(k) presents all the chain elements of the 3D curve. Finally, Fig. 1(l) is the chain of the above-mentioned curve, which is composed of seven chain elements.

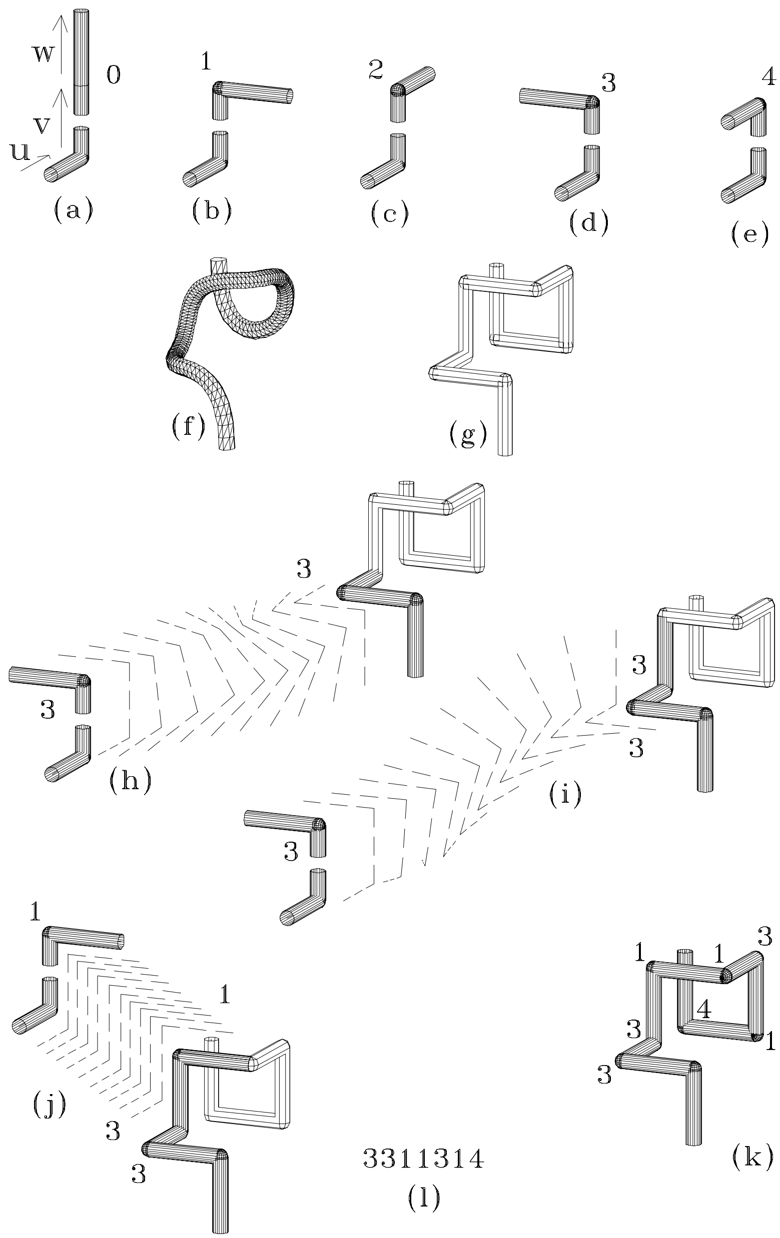


Figure 1: The orthogonal direction change chain code: (a) the chain element "0"; (b) the chain element "1"; (c) the chain element "2"; (d) the chain element "3"; (e) the chain element "4"; (f) an example of a 3D curve; (g) the discrete version of the curve shown in (f); (h) the first chain element of the chain; (i)-(k) next elements of the chain; (l) the chain of the above-mentioned curve.

Fig. 2 illustrates an example of a 3D closed curve. Fig. 2(a) presents the chain elements, again. Fig. 2(b) shows an example of a closed continuous curve. Fig. 2(c) illustrates the digitalized version of the curve shown in (b), its elements and its chain. The origin of this 3D discrete curve is represented by a sphere. Notice that when we are traveling a 3D curve in order to obtain its chain elements and find zero elements, we need to know what non-zero element was the last one in order to define the next element. In this manner orientation is not lost. Due to the fact that the orthogonal direction change chain code only considers relative direction changes, this chain code is invariant under rotation. Fig. 2(d) shows a rotation of the 3D discrete curve presented in (c) on the axis "X". Fig. 2(e) illustrates another example of a rotation of the above-mentioned curve on the axis "Z". Note that all chains are equal. Therefore, they are invariant under rotation. Using this notation its is possible to obtain mirror images of curves with ease. The chain of the mirror image of a curve is another chain (termed *mirroring chain*) whose elements "1" are replaced by elements "3" and vice versa. This replacement does not depend on the orthogonal mirroring plane used, it is valid for all possible orthogonal planes. The *inverse of a chain* of a curve is another chain formed of the elements of the first chain arranged in reverse order. Using the concept of the inverse of a chain, this notation may be starting point normalized by choosing the starting point so that the resulting sequence of elements forms an integer of minimum magnitude. For a complete review of the orthogonal direction change chain code and its properties, see ref. [4].

3 The measure of shape dissimilarity for curves

Since our goal is to give a measure of dissimilarity for 3D curves, we will start presenting an intuitive definition of what similarity is. In ref. [14], Lin states an intuitive definition of similarity as follows:

"**Intuition 1:** The similarity between A and B is related to their commonality. The more commonality they share, the more similar they are.

Intuition 2: The similarity between A and B is related to the differences between them. The more differences they have, the less similar they are.

Intuition 3: The maximum similarity between A and B is reached when A and B are identical, no matter how much commonality they share."

A similarity measure is a function that associates a numeric value with (a pair of) sequences (for example, two curves), with the idea that a higher value indicates greater similarity, while a dissimilarity measure is the opposite, a lower value indicates greater similarity.

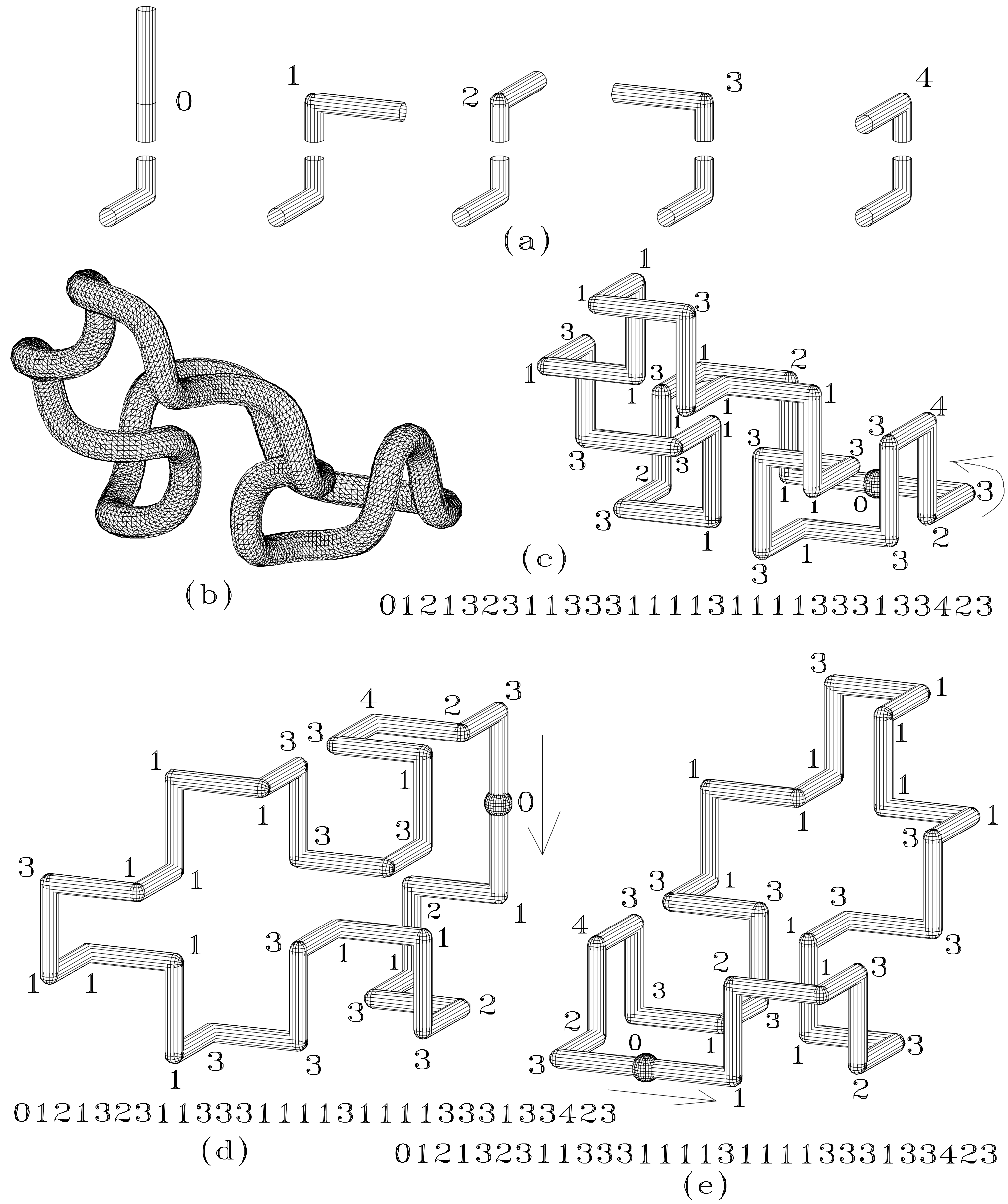


Figure 2: An example of a 3D closed curve: (a) the chain elements; (b) an example of a closed curve; (c) the 3D discrete curve of the curve shown in (b); (d) an example of a rotation on the axis “X”; (e) another example of a rotation on the axis “Z”.

In particular for 3D curves, we say that two curves are more similar when they have in common more subcurves (one-to-one matching), and when these subcurves have the same orientation and position inside their curves. In order to present our proposed measure of shape dissimilarity, we show ten examples of 3D curves which are illustrated in Fig. 3. Fig. 4 presents the digitalized versions of the 3D curves (and their chains) illustrated in Fig. 3. The origins of the 3D discrete curves are represented by spheres.

In the next subsections we present the proposed measure of shape dissimilarity which captures these aspects. First we show how to quantify their similarities, then their differences and finally we give a function which considers both aspects.

3.1 Quantifying their similarities

In order to measure the similarity between two given curves, we need first to quantify their similarities. Here we are interested in finding all subcurves of maximum length that belong to both curves. This leads us to the problem of common-subcurves detection.

3.1.1 Common-subcurves detection

Due to the fact that curves are represented by means of the orthogonal direction change chain code, we can ensure that every different curve has a unique representation (this is true only if their chains have been starting point normalized). Moreover, this representation is invariant under translation and rotation, which means that no matter what is the position and orientation of a curve, its representation will always be the same.

These properties are also true for subcurves. For example, consider the curve 41434 (the first stage of Hilbert curve [5]) which is shown in Fig. 5(a). If this curve is now part of another curve, its representation will be the same no matter its position and orientation, as can be seen in Fig. 5(b) (the second stage of Hilbert curve). Bold lines in Fig. 5(b) show multiple occurrences of curve presented in (a).

The position of the subcurve is given by its initial element index within the curve it belongs to. For example, the curve of Fig. 5(a) has initial element indices e_1, e_9, e_{17}, e_{25}, e_{33}, e_{41}, e_{49} and e_{57} in the curve of Fig. 5(b).

The first two non-zero preceding elements to the pattern 41434 (12, 41, 203, 41, 12, 41 and 31) represent the two orthogonal direction changes needed to define the first element of the common subcurve (element 4 in this example). These must represent an orthogonal direction change, so we have to consider all preceding elements until we find the first two elements different from zero.

Fig. 6 shows four examples of 3D discrete curves with common subcurve 2012031434 (which is drawn in bold lines and underlined in the chains) and

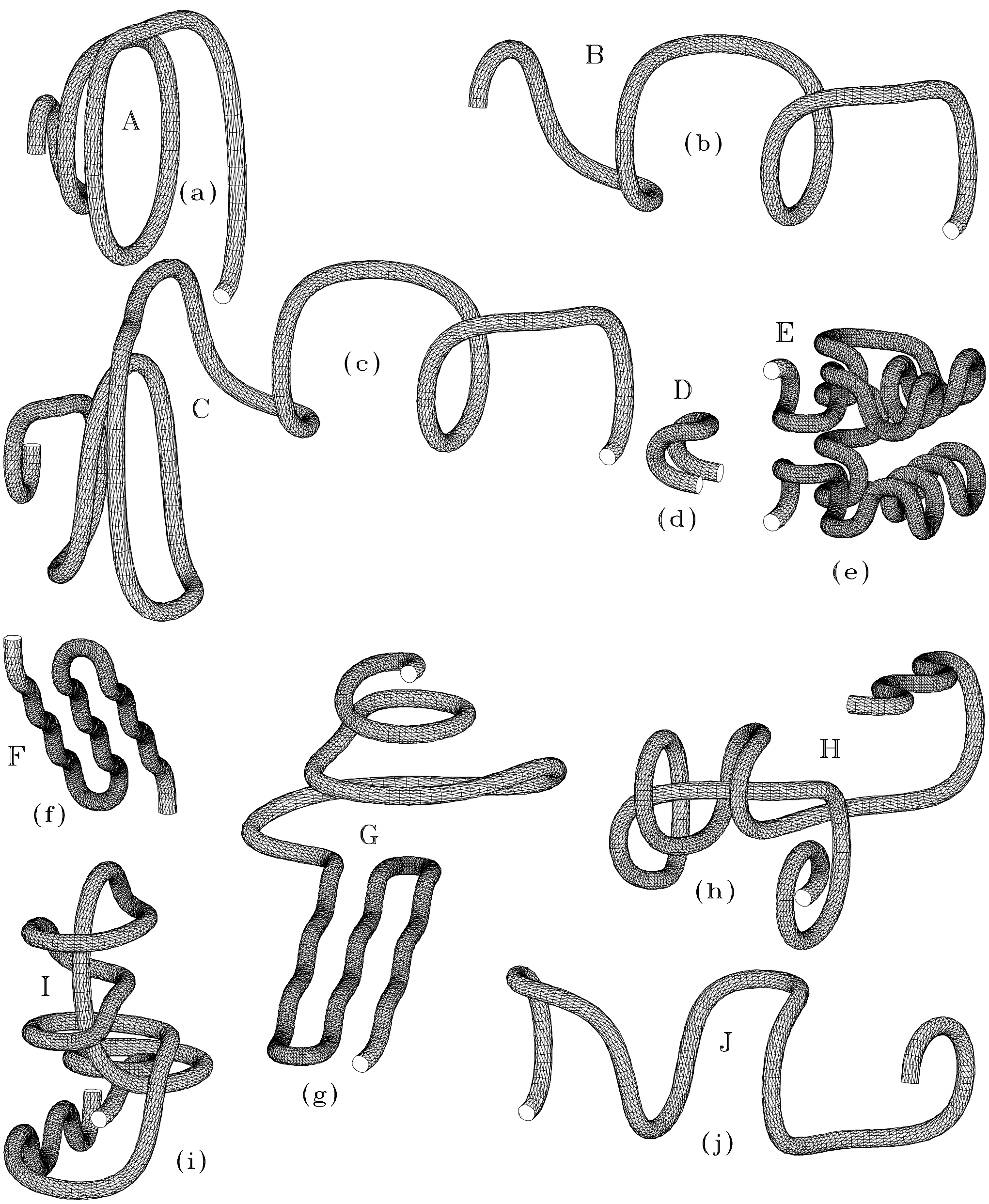


Figure 3: Ten examples of 3D curves: (a)-(j).

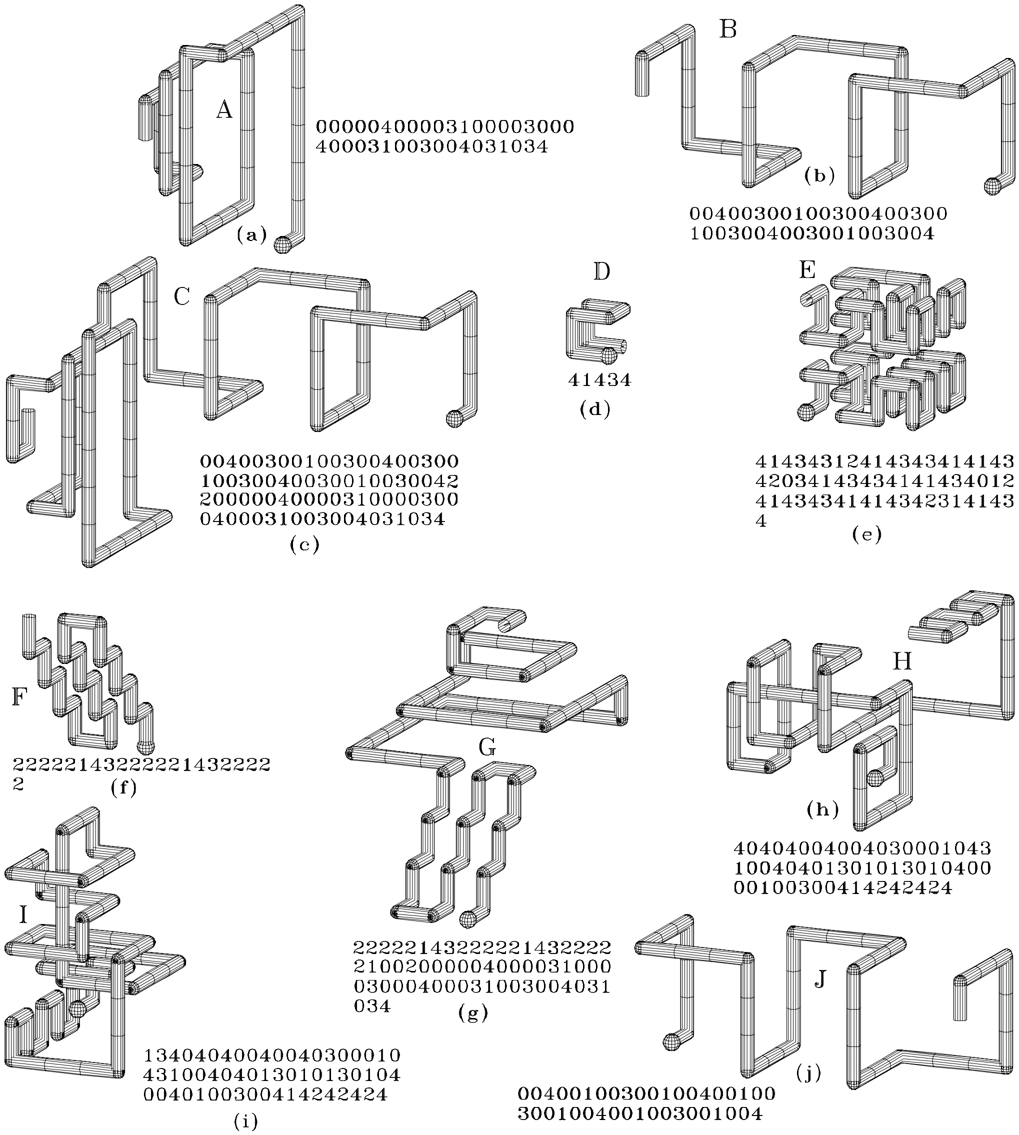
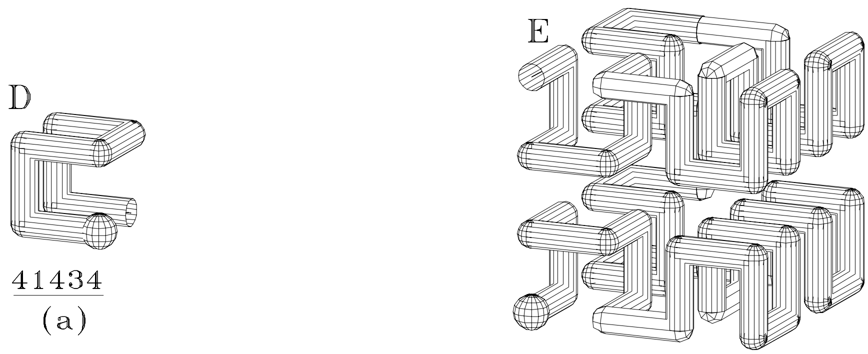


Figure 4: The digitalized versions of the ten 3D curves (and their chains) illustrated in Fig. 3, respectively: (a)-(j).



41434
(a)

4143431241434341
4143420341434341
4143401241434341
4143423141434
(b)

Figure 5: Common-subcurves detection: (a) the example of the 3D discrete curve D (the first stage of Hilbert curve); (b) the 3D discrete curve E (the second stage of Hilbert curve) is composed of multiple occurrences of the curve D.

whose chains have zero elements preceding the pattern and before the first two orthogonal direction changes (drawn in light lines and squared in their chains). Figures 6(a) and (c) are two examples where there are some zero elements just before the first element of the common-subcurve subchain. This causes an enlargement in the middle segment of the first element of the pattern (element 2 in this example). Figures 6(b) and (d) are examples of common subcurves where there are zero elements between the first two nonzero preceding elements. This causes an enlargement of the first segment of the first element. So now the whole description of the common subcurve is the chain which element by element is identical in both curves plus the previous elements needed to define two orthogonal direction changes. In the case where the common subcurve starts at index 1, the number of preceding elements to define two orthogonal direction changes is 2. This is because as you can see in figures 1(a)-(e) all chain elements are composed of three segments. The first two segments of any curve are represented by any two imaginary non-zero chain elements.

Notice that common subcurves are not now completely identical because their beginnings are a little bit different. This fact has to be remembered when computing their dissimilarity.

Summarizing, given two starting point normalized curves, a subcurve belonging to both curves can be found because it has the same elements as a chain plus the first n elements needed to define the first two orthogonal direc-

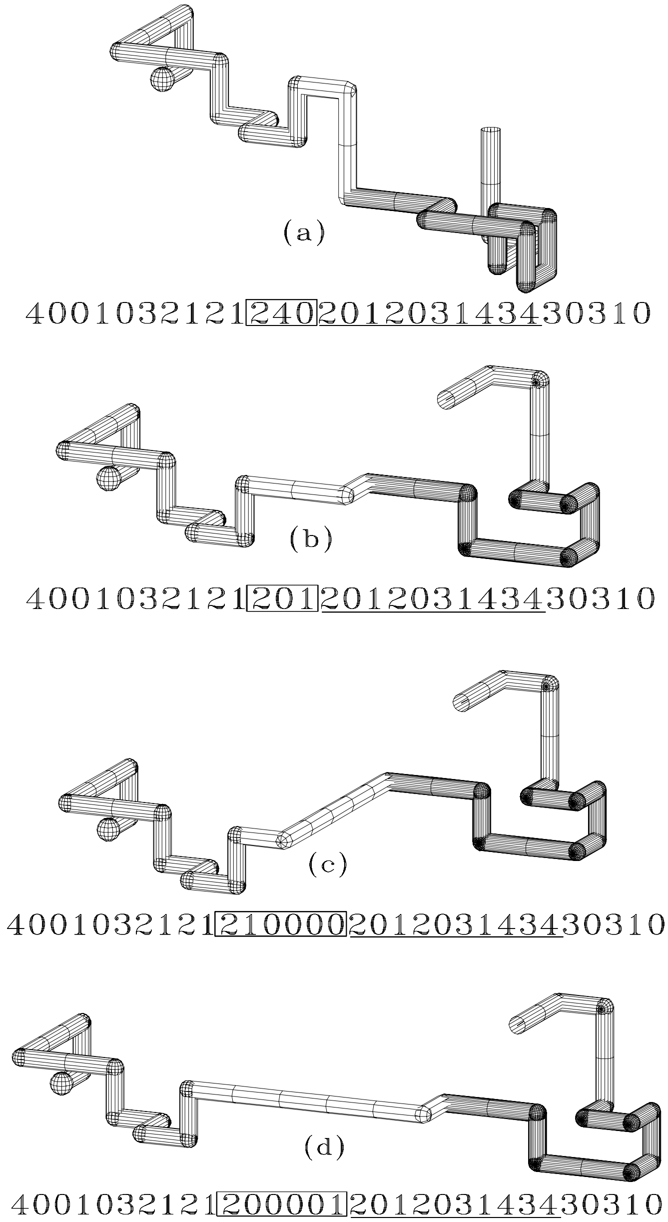


Figure 6: Four examples of 3D discrete curves ((a), (b), (c), and (d)) with the common subcurve 2012031434 that has the characteristic of having some zero elements preceding it and before the first two orthogonal direction changes.

tion changes. So, the problem of common-subcurves detection is reduced to the string-matching problem, in particular, to the problem of finding all the longest common substrings of two strings. Let A be the first string of length m, and B the second string of length n. A bruteforce matching algorithm can be used, which consist of starting at the first element of each string and then shifting A (with wrap-around in the case of closed curves) and keeping all common substrings of maximum length found during this process. Then, do the same for every element of B. This can be computed in $O(m \times n)$. However, better algorithms that use suffix trees can be found in [9].

3.1.2 Dissimilarity measure for common subcurves

Once all common subcurves of maximum length have been detected, it could be possible that two or more common subcurves have overlaps. For this reason we need to find a one-to-one matching between them. For this purpose we propose a dissimilarity measure for common subcurves. This measure gives us a parameter to choose the best matching in case there are overlaps. The characteristics used to evaluate their dissimilarity are: orientation, size and position within their curves. One extra characteristic is if their beginnings are identical as explained in the previous subsection. The size is given by the length of the subchain, the position by its starting element index and its beginning by the number of elements preceding the pattern needed to define two orthogonal direction changes for the first element of the common subcurve. The orientation of the curve can be calculated as follows.

3.1.2.1 Accumulated direction

The accumulated direction is the final orientation of a curve (or subcurve) after it has been affected by all its preceding chain elements. In other words, it is composed of two direction vectors which are the reference to define the next element of the curve.

For the first chain element, these two direction vectors are given arbitrarily, the only restriction is that they have to be orthogonal to each other. For example, u'$= (0, 1, 0)$ and d'$= (-1, 0, 0)$. We also need to append at the beginning of the chain two (non-zero) imaginary reference chain elements to be used as a reference for the first element of the chain.

The new direction vectors u and d, are calculated in terms of the current direction vectors u' and d' and depending on the current chain element, according to the next rule.

Element 0	$u = u$'	$d = d$'
Element 1	$u = d$'	$d = u$' $\times$ d'
Element 2	$u = d$'	$d = u$'

Element 3 $u = d'$ $d = -\,(u' \times d')$
Element 4 $u = d'$ $d = -u'$

Where $u' \times d'$ denotes the cross product for vectors. Thus, the accumulated direction for a curve is given by the final vectors u and d resulting from applying this rule to all the elements of the curve.

Definition 1. Pseudo-metric of accumulated direction.
Let A and B be two chains with accumulated directions (u,d) and (x,w) respectively. The pseudo-metric of accumulated direction is defined as follows.

$$\Delta(A, B) = \begin{cases} 0 & \textit{if } u = x \textit{ and } d = w; \\ 1 & \textit{otherwise.} \end{cases} \tag{1}$$

It can be proved that the pseudo-metric of accumulated direction is a pseudo-metric, since it satisfies the properties of non-negativity, symmetry, identity and the triangle inequality. Notice that it does not satisfy the property of uniqueness because there are different curves with the same accumulated direction. Curves of figures 4(a), (b), and (j) are examples of different curves with the same accumulated direction.

3.1.2.2 Dissimilarity measure between two subchains of a partial common couple

Definition 2. Subchain.
Let $A = a_1 a_2 ... a_m$ *be a chain. A subchain S of A with initial index i is defined as:*

$$S = a_i a_{i+1} ... a_n \quad \textit{such that} \quad 1 \leq i \leq m \quad \textit{and} \quad n \leq m$$

Definition 3. Maximum common subchain.
Let $A = a_1 a_2 ... a_m$ *and* $B = b_1 b_2 ... b_n$ *be two chains. A maximum common subchain S of A and B with initial indices q and r respectively, satisfy the following conditions:*

i) $1 \leq L(S) \leq \min(m, n)$
ii) S is a subchain of A and B, with initial indices q and r respectively.
iii) $a_{q-1} \neq b_{r-1}$ *and* $a_{q+L(S)} \neq b_{r+L(S)}$

where
$L(S)$ *is the length of the subchain.*
$1 \leq q \leq m$ *and* $1 \leq r \leq n$

Definition 4. Left maximization.
Let $A = a_1a_2...a_m$ *and* $B = b_1b_2...b_n$ *be two chains. Let* S *be a subchain of* A *and* B*, with initial index* i *in* A *and initial index* j *in* B*, such that:*

$A = a_1a_2...a_ia_{i+1}...a_{i+L(S)-1}...a_m$ *where,*
$a_i = s_1, a_{i+1} = s_2, ..., a_{i+L(S)-1} = s_{L(S)}$

$B = b_1b_2...b_jb_{j+1}...b_{j+L(S)-1}...b_n$ *where,*
$b_j = s_1, b_{j+1} = s_2, ..., b_{j+L(S)-1} = s_{L(S)}$

The function lmax is defined as:
$lmax(A, B, S) = a_ka_{k+1}...a_ia_{i+1}...a_{i+L(S)-1} = b_{k'}b_{k'+1}...b_jb_{j+1}...b_{j+L(S)-1} = u_{lmax}$

where u_{lmax} *is a maximum subchain of* $A' = a_1a_2...a_{i+L(S)-1}$ *and* $B' = b_1b_2...b_{j+L(S)-1}$.

Intuitively, l_{max} is a function that expands the subchain S to its left until S becomes a maximum subchain of A' and B'. Index k is named "left maximization index in A" and index k' is named "left maximization index in B".

Definition 5. Partial Common Couple.
Let $A = a_1a_2...a_m$ *and* $B = b_1b_2...b_n$ *be two chains. Let* P*' be a common subchain of* A *and* B *with initial indices* i'_1 *and* i'_2 *respectively, and* $L(P') = k'$*, for* $1 \leq k' \leq min(m, n)$.

$P' = p'_1p'_2...p'_{k'}$

Let $P = lmax(A, B, P')$*, with left maximization index in* A i_1 *and left maximization index in* B i_2.

$P = p_1p_2...p_k = a_{i_1}a_{i_{1+1}}...a_{i_1+L(P)-1} = b_{i_2}b_{i_{2+1}}...b_{i_2+L(P)-1}$

Let S *be a subchain of* A *with initial index 1 and* $L(S) = i_1 + L(P) - 1$.

$S = a_1a_2...a_{i_1}a_{i_1+1}...a_{i_1+L(P)-1} = a_1a_2...p_1p_2...p_k$

Let T *be a subchain of* B *with initial index 1 and* $L(T) = i_2 + L(P) - 1$.

$T = b_1b_2...b_{i_2}b_{i_2+1}...b_{i_2+L(P)-1} = b_1b_2...p_1p_2...p_k.$

For convenience the couple (S,T) will be named partial common couple.

Given two chains, A and B, the minimum number of partial common couples they could have is zero. This happens when A and B are completely different, in other words, the two chains do not have any chain element in common. While the maximum number of partial common couples they could have is given by the following expression:

$$(n0A \times n0B) + (n1A \times n1B) + (n2A \times n2B) + (n3A \times n3B) + (n4A \times n4B)$$

where the notation $n0A$ means "the number of zero elements in A".

As an example, consider the chains D and G shown in Fig. 4. The maximum number of partial common couples D and G could have is:

$$0 \times 27 + 1 \times 6 + 0 \times 16 + 1 \times 8 + 3 \times 6 = 32$$

All their partial common couples (S, T) are (P is shown in boldface):

1. (4**143**), 22222**143**)
2. (4**143** , 222221432222**143**)
3. (4143**4**, 222221**4**)
4. (4143**4**, 22222143222221**4**)
5. (414**34**, 2222214322222143222221002000004000031000030004000310030040310**34**)
6. (4**1** , 2222214322222143222222**1**)
7. (**4** , 222221**4**)
8. (**4** , 22222143222221**4**)
9. (41**4** , 222221432222214322222100200000400003100003000400031003004031034 → 22222143222221432222210020000040000310000300040003100300 403103**4**)
10. (4**1** , 2222214322222143222221002000004000031000030004000310030040 3**1**)
11. (**4** , 2222214322222143222221002000004000031000030004000310030040 3103**4**)
12. (414**3** , 22222143222221432222210020000040000310000300040003100 **3**)
13. (414**3** , 22222143222221432222210020000040000310000300040003100300 40**3**)
14. (4**1** , 22222143222221432222210020000040000310000300040000**31**)
15. (414**3** , 222221432222214322222100200000400003100000 **3**)
16. (4143**4**, 222221432222214322222100200000400003100003000400031003 00**4**)
17. (41**4** , 22222143222221432222210020000040000310000300040003100300 **4**)
18. (4**1** , 222221432222214322222100200000400003 **1**)
19. (**4** , 222221432222214322222100200000400003100003000400031003 00**4**)
20. (414**3** , 22222143222221432222210020000040000310000300040000 **3**)
21. (4143**4**, 222221432222214322222100200000 **4**)
22. (4143**4**, 22222143222221432222210020000040000310000300**4**)

23. (41**4** , 22222143222222143222221002000004**)**
24. (41**4** , 2222214322222214322222210020000040000310000300004)
25. (**4** , 2222221432222214322222210020000004)
26. (**4** , 2222214322222214322222210020000040000310000300004)
27. (414**3** , 222221432222214322222100200000400003)

Definition 6. Dissimilarity measure for partial common couples.
The dissimilarity for partial common couples is defined as:

$$d(S,T) = \begin{cases} 0, & when\ S = T \\ \frac{|L(S)-L(T)|}{max(L(S),L(T))-1} + (min(m,n) - L(P)) + \\ \Delta(S,T) + |ne_S - ne_T|\,, & when\ S \neq T \end{cases} \tag{2}$$

where,

$\frac{|L(S)-L(T)|}{max(L(S),L(T))-1}$ *measures the displacement of the two common subchains within their respective curves, in such a way that if the correspondence in position is exact this term becomes zero. This term is normalized to the range [0-1].*

$(min(m,n) - L(P))$ *measures how large are the common subchains with respect to the curve where they are contained.*

$\Delta(S,T)$ *is the pseudo-metric of accumulated direction of* S *and* T*. In other words, this term measures whether the final orientation of the subcurves* S *and* T *is the same. If it is the same, then this term becomes zero.*

$|ne_S - ne_T|$ *measures the number of preceding elements to* P *in* S *and* T *needed to define two orthogonal direction changes as explained in subsection 3.1.1.*

It can be proved that this dissimilarity measure for partial common couples is a metric, since it satisfies the properties of non-negativity, symmetry, identity, uniqueness and the triangle inequality.

This measure becomes more intuitive when it is bounded to the range of $[0,1]$, where 0 means that the two subchains are identical and 1 that they are completely different. For this purpose, the bounded dissimilarity measure for partial common couples, which is also a metric, is now defined.

Definition 7. Bounded dissimilarity measure for partial common couples.
The bounded dissimilarity measure for partial common couples is given by:

$$\overline{d}(s,t) = \frac{d(s,t)}{m+n}. \tag{3}$$

3.1.3 Finding one-to-one matching

After all maximum common subcurves have been detected, it is possible to have overlaps between them, in other words, it is posible that two or more common subcurves share some elements. Formally,

Definition 8. Partial common couples overlap.
Two partial common couples (S,T) and (S',T') of A and B, with maximum common subchains

$$P = p_1 p_2 ... p_k = a_{i_1} a_{i_1+1} ... a_{i_1+k-1} = b_{i_2} b_{i_2+1} ... b_{i_2+k-1}$$
$$Q = q_1 q_2 ... q_r = a_{i_3} a_{i_3+1} ... a_{i_3+r-1} = b_{i_4} b_{i_4+1} ... b_{i_4+r-1}$$

respectively, with initial indices i_1 in A and i_2 in B for P, and initial indices i_3 in A and i_4 in B for Q, do not have overlap if the subchains:

$$P' = p'_1 p'_2 ... p'_{k'} = a_{i'_1} a_{i'_1+1} ... a_{i'_1+k'-1} = b_{i'_2} b_{i'_2+1} ... b_{i'_2+k'-1}$$
$$Q' = q'_1 q'_2 ... q'_{r'} = a_{i'_3} a_{i'_3+1} ... a_{i'_3+r'-1} = b_{i'_4} b_{i'_4+1} ... b_{i'_4+r'-1}$$

before left maximization, $P = lmax(A, B, P')$ and $Q = lmax(A, B, Q')$, satisfy:

$$\{i'_1 - ne_S, ..., i'_1 - 1, i'_1, i'_1 + 1, ..., i'_1 + k' - 1\} \bigcap$$
$$\{i'_3 - ne_{S'}, ..., i'_3 - 1, i'_3, i'_3 + 1, ..., i'_3 + r' - 1\} = \emptyset$$

$$\{i'_2 - ne_T, ..., i'_2 - 1, i'_2, i'_2 + 1, ..., i'_2 + k' - 1\} \bigcap$$
$$\{i'_4 - ne_{T'}, ..., i'_4 - 1, i'_4, i'_4 + 1, ..., i'_4 + r' - 1\} = \emptyset$$

where the number of elements needed to define two orthogonal direction changes, ne_S, $ne_{S'}$,ne_T, $ne_{T'}$, are with respect to P' and Q'.

Let $\{(S_{j'}, T_{j'})\}$ be the set of all partial common couples of curves A and B. For example, the one shown in subsection 3.1.2.2. This set is first sorted according to their dissimilarity measure using Eq. (3). The computation of Eq. (3) is linear in the number of partial common couples of A and B. If two partial common couples in this set have overlap between them, then the one that has a lower dissimilarity d is chosen. The eliminated partial common couples are inspected to find out if there are still parts of them that could be matched without overlap, thus subchains P'_j must satisfy the following conditions:

- they are maximum common subchains of A and B, or

- they are subchains of maximum common subchains of A and B that were divided because they had overlap.

Finally, let $\{S_j, T_j\}$ be the set of all partial common couples of A and B, with no overlap with other partial common couple in this set. The similarities between A and B are given by:

$$\sum_{j=1}^{l} \overline{d}(S_j, T_j)$$

where

(S_j, T_j) is the j-th partial common couple found in A and B with no overlap with other partial common couples chosen.

l is the total number of partial common couples found in A and B without overlap.

3.2 Quantifying their differences

To quantify the differences between two given curves, the number of elements left without correspondence is computed. Formally, the differences between the curves A and B is given by

$$m + n - \left(\sum_{j=1}^{l} 2L(P'_j) + ne_{S_j} + ne_{T_j} \right) + 4$$

where

P'_j is the j-th subchain that corresponds to the j-th partial common couple, before the left maximization.

ne_{S_j}, ne_{T_j} are the number of elements in S_j and T_j needed to define and orthogonal direction change with respect to P'_j.

m is the number of elements of curve A.

n is the number of elements of curve B.

3.3 Dissimilarity measure for 3D curves

Definition 9. Dissimilarity measure for 3D curves.
Let $A = a_1 a_2 ... a_m$ and $B = b_1 b_2 ... b_n$ be two chains. The dissimilarity of A and B is defined as:

$$D(A,B) = \sum_{j=1}^{l} \overline{d}(s_j, t_j) + \left[m + n - \left(\sum_{j=1}^{l} 2L(P'_j) + ne_{S_j} + ne_{T_j} \right) + 4 \right] \quad (4)$$

Notice that the first term measures the similarities between A and B while the second term measures their differences.

It can be proved that this dissimilarity measure satisfy the properties of non-negativity, symmetry, identity and uniqueness. Notice that the property of triangle inequality is not satisfied, thus it is not a metric.

This measure of dissimilarity becomes more intuitive when it is bounded to the range of $[0, 1]$.

Definition 10. Bounded dissimilarity measure for 3D curves.
The bounded dissimilarity measure for 3D curves is defined as:

$$\overline{D}(A,B) = \frac{D(A,B)}{m+n+4}. \quad (5)$$

4 Results

To illustrate the capabilities of the measure of shape dissimilarity proposed here, we present the values of the measure for all 3D discrete curves shown in Fig. 4. Thus, Table 1 summarizes the computations of the different measures of shape dissimilarity for the ten analyzed curves. The values of these measures are normalized and vary continuously from 0 to 1. These values were computed using Eq. (5). When two curves are identical, the value of the measure of shape dissimilarity is equal to zero.

Analyzing the values of the measures of shape dissimilarity in Table 1, the two most similar curves of the ten studied above are the curves H and I (their value of measure of shape dissimilarity is equal to 0.1091). The curves B and J are also very similar their value is equal to 0.1131. An interesting example corresponds to the curves A and B, both have a value equal to 0.2264 which indicates that they are very similar. Note that both curves look like spirals. Thus, in this case the value of the measure was not affected too much by the intercalated zero elements. On the other hand, the most dissimilar curves are the following: curves D and G; D and H; C and D; and C and F, respectively. The values of the measure of shape dissimilarity are greater than 0.8.

	A	B	C	D	E	F	G	H	I	J
A	0.0	0.2264	0.3885	0.5823	0.6563	0.7741	0.2193	0.4092	0.4204	0.252
B		0.0	0.3278	0.6329	0.689	0.7314	0.4005	0.3445	0.3772	0.1131
C			0.0	0.8085	0.5249	0.8046	0.309	0.3988	0.3407	0.4003
D				0.0	0.8	0.6712	0.8618	0.8196	0.7965	0.6329
E					0.0	0.5612	0.5892	0.4748	0.54	0.5968
F						0.0	0.4772	0.7208	0.7951	0.7314
G							0.0	0.2722	0.3	0.4571
H								0.0	0.1091	0.3547
I									0.0	0.397
J										0.0

Table 1: Values of the measure of shape dissimilarity for the ten analyzed curves.

Finally, in order to show the properties of robustness of the proposed dissimilarity measure, we present three curves in Fig. 7 which have small deformations. Fig. 7(a) illustrates the curve A, its continuous and discrete representation and its chain. Fig. 7(b) shows the curve B which is very similar to curve A. Fig. 7(c) illustrates the curve C and the measures of shape dissimilarity between A and B, and between A and C, respectively. Note that in both measures curves are very similar. However, curves A and C are more similar than A and B which coincides with the shown curves. As a final example, we show in Fig. 8 two examples of curves to stand out how this representation and dissimilarity measure can be also applied to the case of 2D curves. Fig. 8(a) shows the 2D curve A which corresponds to a "fig leaf". Fig. 8(b) illustrates the curve B which is similar to A with a small deformation and the dissimilarity measure between both is presented too. Due to the fact that this proposed measure of shape dissimilarity was developed based on the orthogonal direction change chain code (in this code each element depends on the two previous non-zero elements and is invariant under rotation): it is not possible the use of this measure to other chain codes, such as Freeman chain coding [7] which uses absolute directions.

5 Conclusions

We have presented a measure of shape dissimilarity for 3D curves. This measure is based on a curve representation by means of the orthogonal direction change chain code; this curve representation is invariant under *translation* and *rotation*, may be *starting point normalized* and the *mirror image* of curves are obtained with ease. The above-mentioned code produces a robust notation for our proposed measure of shape dissimilarity. Thus, the computation of

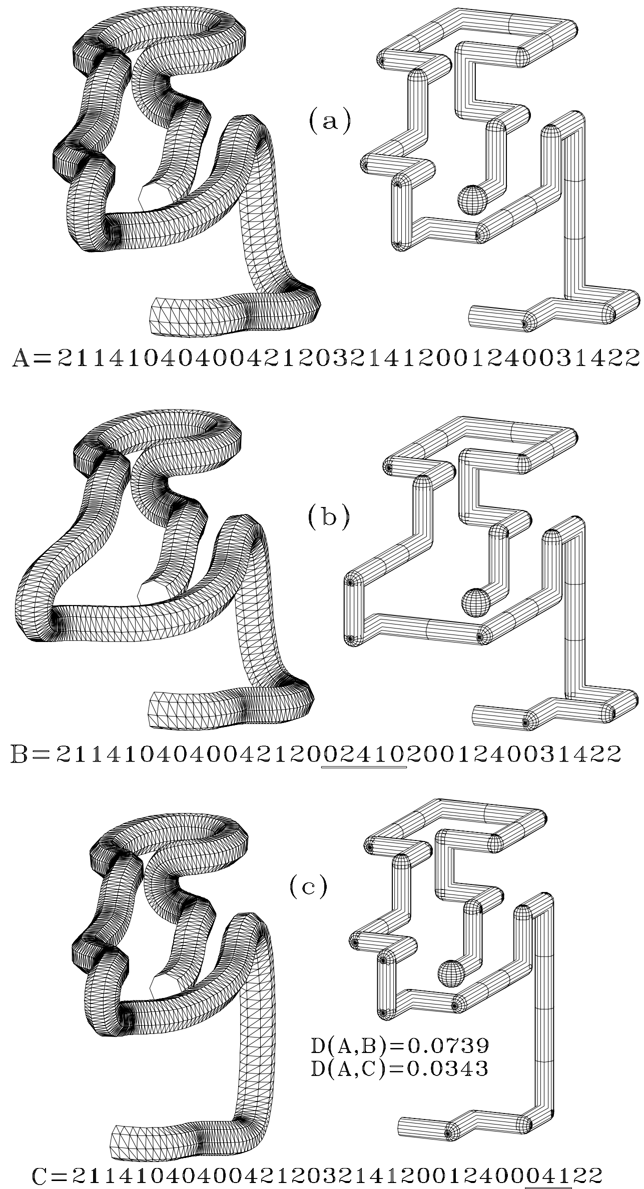


Figure 7: Three examples of 3D curves which have small deformations and their dissimilarity measures: (a) the continuous representation of the original curve A, its digitalized version, and its chain; (b) the curve B which has a small deformation at the middle part, in relation to A; (c) the curve C which has a small deformation at the beginning.

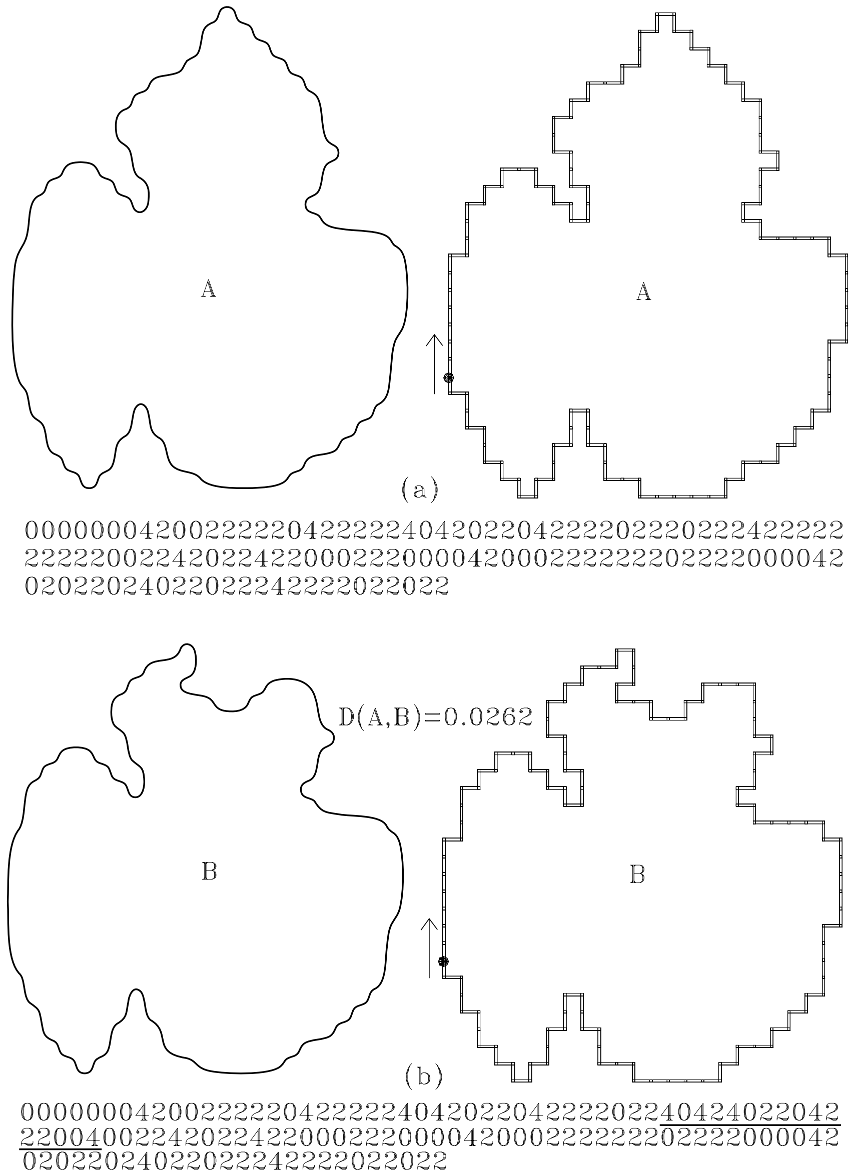


Figure 8: Two examples of 2D curves using the proposed dissimilarity measure: (a) the curve A; (b) the curve B and the dissimilarity measure between A and B.

this measure for two curves is based on the analysis of their common and different subcurves represented by their chain elements. This analysis allows us to detect dissimilarity of shape for curves. Generally speaking, we proposed that two curves are more similar when they have in common more subcurves, and when these subcurves have the same orientation and position inside their curves. Finally, we presented some results of the computation of the proposed measure for 15 curves.

ACKNOWLEDGEMENT. This work was supported by IIMAS-UNAM.

References

[1] H. Alt, M. Godau, Computing the Frechet distance between 2 polygonal curves, International Journal of Computational Geometry & Applications **5** (1-2) (1995), 75-91.

[2] E. M. Arkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, J. S. B. Mitchel, An efficiently computable metric for comparing polygonal shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence **13** (3) (1991), 209-216.

[3] E. Belogay, C. Cabrelli, U. Molter, R. Shonkwiler, Calculating the Hausdorff distance between curves, Information Processing Letters **64** (1997), 17-22.

[4] E. Bribiesca, A chain code for representing 3D curves, Pattern Recognition **33** (2000), 755-765.

[5] E. Bribiesca and C. Velarde, A formal language approach for a 3D curve representation, Computers & Mathematics with Applications **42** (2001), 1571-1584.

[6] S. J. Dickinson, A. P. Pentland, A. Rosenfeld, From volumes to views: an approach to 3-D object recognition, CVGIP: Image Understanding **55** (1992), 130-154.

[7] H. Freeman, Computer processing of line drawing images, ACM Computing Surveys **6** (1974), 57-97.

[8] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Second Edition, Prentice Hall, Upper Saddle River, New Jersey 07458, 2002.

[9] D. Gusfield, Algorithms on Strings, Trees, and Sequences, Cambridge University Press, Cambridge CB2 1RP, 1997.

[10] A. Guzmán, Canonical shape description for 3-d stick bodies, MCC Technical Report Number: ACA-254-87, Austin, TX. 78759 (1987).

[11] M. Holden, D. L. G. Hill, E. R. E. Denton, J. M. Jarosz, T. C. S. Cox, T Rohlfing, J. Goodey, D. J. Hawkes, Voxels similarity measures for 3-D serial MR brain image registration, IEEE Transactions on Medical Imaging **19** (2000), 94-102.

[12] A. K. Jain, R. Hoffman, Evidence-based recognition of 3-D objects, IEEE Transactions on Pattern Analysis and Machine Intelligence **10** (1988), 783-802.

[13] Stan Z. Li, Invariant representation, matching and pose estimation of 3D space curves under similarity transformations, Pattern Recognition **30** (3) (1997), 447-458.

[14] D. Lin, An Information-Theoretic Definition of Similarity. Proc. of the Fifteenth International Conference on Machine Learning, Madison, Wisconsin, pp. 296-304 (1998).

[15] Chong-Huah Lo and Hon-Son Don, Invariant representation and matching of space curves, Journal of Intelligent and Robotic Systems **28** (2000), 125-149.

[16] G. Lohmann, Volumetric Image Analysis, Wiley & Sons and B. G. Teubner Publishers, New York, NY, 1998.

[17] W. Rodriguez, M. Last, A. Kandel, H. Bunke, 3-Dimensional curve similarity using string matching, Robotics and Autonomous Systems **49** (3-4) (2004) 165-172.

Received: May 19, 2006