

Coastline Matching Process based on the discrete Fréchet distance

Ariane Mascret¹, Thomas Devogele¹, Iwan Le Berre², Alain Hénaff²

¹ Naval academy Research Institute (IRENav)
Lanvéoc, BP 600, F-29 240 Brest Naval, FRANCE
e-mail: {mascret, devogele}@ecole-navale.fr

² GEOMER Laboratory, LETG UMR 6554 CNRS,
Institut Universitaire Européen de la Mer (UBO),
Technopôle Brest-Iroise, 29280 Plouzané, FRANCE
e-mail: {iwan.leberre, alain.henaff}@univ-brest.fr

Abstract

Spatial distances are the main tools used for data matching and control quality. This paper describes new measures adapted to sinuous lines to compute the maximal and average discrepancy: Discrete Fréchet distance and Discrete Average Fréchet distance. Afterwards, a global process is defined to automatically handle two sets of lines. The usefulness of these distances is tested, with a comparison of coastlines. The validation is done with the computation of three sets of coastlines, obtained respectively from SPOT 5 orthophotographs and GPS points. Finally, an extension to Digital Elevation Model is presented.

Keywords

Data fusion, quality control, data matching, Fréchet distance, coastline monitoring.

1 Introduction

Computing the distance between two objects is a basic tool of geographic information systems. The most commonly used distance is the Euclidean distance (d_E) between two points. Others geometries (line and area in a two dimensional Cartesian system) need additional measures. In daily life, the notion of distance stand for the minimal effort required to reach one place from another. For example, the minimal distance between a pipeline and a river is the Euclidean distance between the two closest points from the river and the pipeline. Mathematically, a distance verifies three properties: non-negative, symmetry, triangle inequality. Thus, minimal distances that measure the distance between the closest points of geometries, can be completed by other distances like average distances or maximal distances. These lasts measure the average or maximal Euclidean distance between points of both geometries.

Maximal and average distances are useful to control or match data from different datasets. For example, in a quality control, they give the discrepancy between the encoded location and the location as defined in the specification (Veregin 1999). Likewise during the matching process, those measures permit to identify sets of data representing the same real world phenomenon in different data sets (Devogele et al. 1996).

Two different maximal distances are employed to calculate the maximal gap between lines: the Hausdorff distance and the Fréchet distance. The Hausdorff distance is the most popular maximal distance between two lines (L_1, L_2) (Deng et al. 2005) (Alt and Godau 1995). The Hausdorff distance (d_H) is defined as follows:

$$d_H(L_1, L_2) = \max \left(\sup_{p_1 \in L_1} \inf_{p_2 \in L_2} (d_E(p_1, p_2)), \sup_{p_2 \in L_2} \inf_{p_1 \in L_1} (d_E(p_1, p_2)) \right) \quad (1)$$

A line is an ordered set of points. Unfortunately, the Hausdorff distance does not take into account this property. Two lines can have a small d_H , without being similar each other at all. The inconvenient of the Hausdorff distance is the computation of Euclidean distance between closer points and not between homologous points (points, which can be visually matched). Hence, Hausdorff distance can not be used for sinuous lines. For this kind of lines, the Fréchet distance is more appropriated (Alt and Gadau 1995).

In the maritime context, the majority of the lines, like coastlines or isolines, used for making studies, are sinuous. The aim of this paper is to detail

how calculate the Fréchet distance, measure the average discrepancy for those kind of lines, and illustrate the result with a comparison between coastlines. This last part is realised thanks to the implementation of new methods based on Average Fréchet distance.

The remainder of this paper is as follows. Section 2 describes the discrete Fréchet distance which is a good approximation of the Fréchet distance. In section 3, this discrete distance is extended to introduce an average linear distance: the average Fréchet distance. Moreover it also explains how to compute this measure. A global process defined to match homologous objects from two datasets is proposed in section 4. Section 5 illustrates this matching process and these two distances by a real example of quality control on coastlines datasets. Related works on digital elevation model are described and discussed in section 6 and finding are summarized in section 7.

2 Discrete Fréchet distance

The Fréchet distance is the maximal distance between two oriented lines. Each oriented line is equivalent to a continuous function $f: [a, a'] \rightarrow V$ ($g: [b, b'] \rightarrow V$) where $a, a', b, b' \in \mathfrak{R}$, $a < a'$ ($b < b'$) and (V, d) is a metric space. d_F denotes their Fréchet distance defined as follows:

$$d_F(f, g) = \inf_{\substack{\alpha: [0,1] \rightarrow [a,a'] \\ \beta: [0,1] \rightarrow [b,b']}} \max_{t \in [0,1]} d(f(\alpha(t)), g(\beta(t))) \quad (2)$$

Let us give an illustration of the Fréchet distance: a man is walking with a dog on a leash. This man is walking on the one curve, the dog on the other one. Both may vary their speed, but backtracking is not allowed. Then the Fréchet distance of the curves is the minimal length of a leash that is necessary. The Fréchet method has the advantage of computing distances only on homologous points and not between closest points as for the Hausdorff distance.

Eiter and Mannila (1994) gave an approximation: the discrete Fréchet distance (d_{dF}) that computes in time $O(n \cdot m)$. L_1 and L_2 are interpreted as two oriented finite sets of points: $\langle L_{1,1} \dots L_{1,n} \rangle$ and $\langle L_{2,1} \dots L_{2,m} \rangle$. d_{dF} is the minimal length of leash such as a way from the pair of beginning points $(L_{1,1}, L_{2,1})$ to the pair of ending points $(L_{1,n}, L_{2,m})$ is possible. The path gives an ordered set of $(L_{1,i}, L_{2,j})$ such as the following pair of $(L_{1,i}, L_{2,j})$ is one of these three pairs:

- $(L_{1,i+1}, L_{2,j+1})$ man and dog are walking,
- $(L_{1,i+1}, L_{2,j})$ only the man is.
- $(L_{1,i}, L_{2,j+1})$ only the dog is.

These sets of points include end points of line segments (vertices). Some points on a line segment can also be integrated by resampling these sets.

d_{dF} is a good estimation of d_F because the approximation is limited by the maximal distance between two consecutive points (LengthMaxSeg) (Eiter and Mannila 1994):

$$d_F(L_1, L_2) \leq d_{dF}(L_1, L_2) \leq d_F(L_1, L_2) + \text{LengthMaxSeg} \quad (3)$$

In order to limit this approximation to ϵ , a resampling can be applied to both lines.

The discrete Fréchet between L_1 and L_2 can be computed recursively as follows:

$$d_{Fd}(L_1, L_2) = \max \left(\begin{array}{l} d_E(L_{1,n}, L_{2,m}) \\ \min \left(\begin{array}{l} d_{Fd}(\langle L_{1,1} \dots L_{1,n-1} \rangle, \langle L_{2,1} \dots L_{2,m} \rangle) \forall n \neq 1 \\ d_{Fd}(\langle L_{1,1} \dots L_{1,n} \rangle, \langle L_{2,1} \dots L_{2,m-1} \rangle) \forall m \neq 1 \\ d_{Fd}(\langle L_{1,1} \dots L_{1,n-1} \rangle, \langle L_{2,1} \dots L_{2,m-1} \rangle) \forall n \neq 1, m \neq 1 \end{array} \right) \end{array} \right) \quad (4)$$

$\langle L_{1,1} \dots L_{1,n-1} \rangle$ and $\langle L_{2,1} \dots L_{2,m-1} \rangle$ represent lines. Hence, it is possible to recursively apply this d_{dF} process with parameters: $\langle L_{1,1} \dots L_{1,n-1} \rangle$, $\langle L_{2,1} \dots L_{2,m-1} \rangle$. This process is terminated when the two lines are reduced to two single points ($\langle L_{1,1} \rangle$, $\langle L_{2,2} \rangle$).

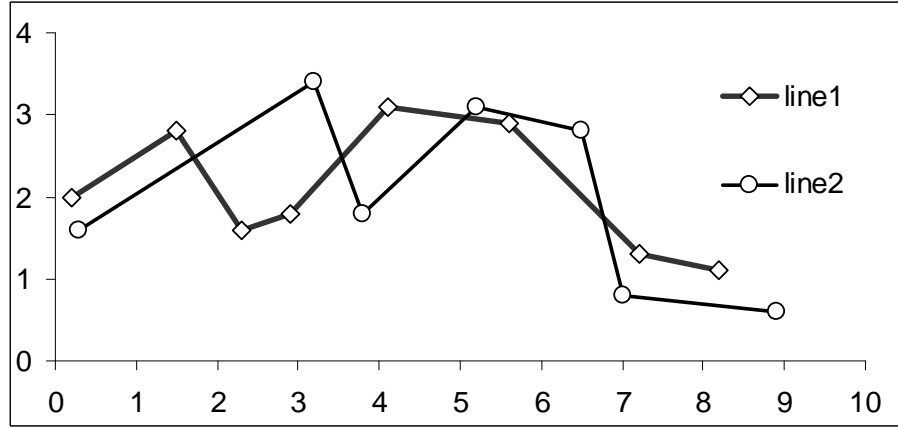


Fig. 1. Example of a couple of lines

The example presented in **Fig. 1**, illustrates the computing of the discrete Fréchet distance. The line 1 (L_1) is composed by 8 vertices ($L_{1,1}$ to $L_{1,8}$) and the line 2 (L_2) is composed by 7 vertices ($L_{2,1}$ to $L_{2,7}$).

The computing of the two matrices replaced the recursive process. The dimension of these two matrices is $n \times m$ (see table 1), where n and m are the number of vertices of L_1 and L_2 . These matrices are:

- The matrix of Euclidean Distance (MD). The value of the cell $MD_{i,j}$ is the distance between $L_{1,i}$, $L_{2,j}$.
- The Fréchet matrix (FM) which allows to calculate iteratively the Fréchet distance (Eiter et Mannila 94). The formula to compute $MF_{i,j}$ is :

$$MF_{i,j} = \max(d_E(L_{1,i}, L_{2,j}), \min(MF_{i-1,j}, MF_{i,j-1}, MF_{i-1,j-1})) \quad (5)$$

The discrete Fréchet distance is the value of $MF_{n,m}$. **Table 1** gives the d_{dF} between the two lines of figure 1: 1.8. d_{dF} is equal to the Euclidean distance between $L_{1,2}$ and $L_{2,2}$. These two points are homologous. For these two lines, the Hausdorff distance is smaller and is not significant in term of matching.

Table1 present a simple sample. The partial discrete Fréchet distance (d_{pdF}), with two smaller matrices (7×8) are computed in this table. If a re-sampling is processed with LengthMaxSeg equal to 0.1, two matrices (107×129) are computed and d_{pdF} is equal to 1.2.

Table 1. Matrix of Euclidean Distance and Fréchet Matrix for lines of figure 1.

L1.i.x	0.2	1.5	2.3	2.9	4.1	5.6	7.2	8.2	
L1.i.y	2	2.8	1.6	1.8	3.1	2.9	1.3	1.1	
Matrix of Euclidian distance between (L1.i, L2.j)									
L2.j.x	L2.j.y	1	2	3	4	5	6	7	8
0.3	1.6	0.41	1.70	2.00	2.61	4.09	5.46	6.91	7.92
3.2	3.4	3.31	1.80	2.01	1.63	0.95	2.45	4.52	5.50
3.8	1.8	3.61	2.51	1.51	0.90	1.33	2.11	3.44	4.46
5.2	3.1	5.12	3.71	3.26	2.64	1.10	0.45	2.69	3.61
6.5	2.8	6.35	5.00	4.37	3.74	2.42	0.91	1.66	2.40
7	0.8	6.91	5.85	4.77	4.22	3.70	2.52	0.54	1.24
8.9	0.6	8.81	7.72	6.68	6.12	5.41	4.02	1.84	0.86

Fréchet Matrix									
	1	2	3	4	5	6	7	8	
1	0.41	1.70	2.00	2.61	4.09	5.46	6.91	7.92	
2	3.31	1.80	2.01	2.00	2.00	2.45	4.52	5.50	
3	3.61	2.51	1.80	1.80	1.80	2.11	3.44	4.46	
4	5.12	3.71	3.26	2.64	1.80	1.80	2.69	3.61	
5	6.35	5.00	4.37	3.74	2.42	1.80	1.80	2.40	
6	6.91	5.85	4.77	4.22	3.70	2.52	1.80	1.80	
7	8.81	7.72	6.68	6.12	5.41	4.02	1.84	1.80	

New Measures derived from discrete Fréchet distance are introduced in (Devogele 2002):

- The partial discrete Fréchet distance (d_{pdF}). This measure is useful to match a line L_1 with a part of another line L_2 . d_{pdF} detects the partial homologous line $\langle L_{2.begin} \dots L_{2.end} \rangle$ and computes d_{pdF} . d_{pdF} is equal to $d_{dF}(L_1, \langle L_{2.begin} \dots L_{2.end} \rangle)$. Fig. 2a shows a case where the computing of d_{pdF} is necessary.
- The discrete Fréchet distance between 2 polygon borderlines. The process defines a function T to translate polygon borderlines P_1 and P_2 into lines L_1 and L_2 such as the d_{dF} between L_1 and L_2 is minimal. This process can also inverse the ordering of points. For example, d_{dF} between the two polygon borderlines of Figure 2b can be measured. No partial discrete Fréchet distance between 2 polygons borderlines are defined.
- The partial discrete Fréchet distance (d_{pdF}) between a line L_1 and a part of polygon borderlines P_2 . This measure is a mix of the two first measures. Figure 2c shows an example where the d_{pdF} can be computed between a line and a polygon borderline.

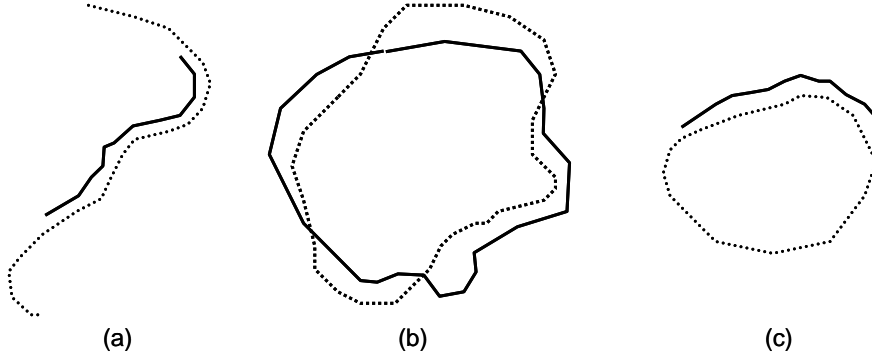


Fig. 2. Examples of interesting pairs of geometries for the computation of measure derived from the Fréchet distance

3 Average Fréchet distance

A new distance is defined from discrete Fréchet distance: the average Fréchet distance (d_{aF}). d_{aF} is the average Euclidean distance between points of pairs, which is based on the minimum path (MP).

As pre-processing, the path between pairs of points $(L_{1,1}, L_{2,1})$ and $(L_{1,n}, L_{2,m})$ is computed. This one is MP, compatible with discrete Fréchet distance. Several paths of the man and the dog with a length of leash equal to

Fréchet distance are possible. MP is the one where the man and the dog walk one their curves but choose to be as closer as possible. The two matrices MD and FM and an inferior or equal (\leq) operation between two pair of real numbers, are used to compute MP.

This operation \leq is defined as follow:

$$(a,b) \text{ and } (c,d) \in \mathbb{R}^2 \quad (6)$$

$$(a,b) \leq (c,d) \text{ if } a < c \text{ or if } a = c \text{ and } b \leq d$$

The minimal path is constructed by backtracking through the matrix. So the last pair $(L_{1,n}, L_{2,m})$ is added, while the previous one look for pair help to \leq operation. For $(L_{1,i}, L_{2,j})$, three previous candidate pairs are possible: $(L_{1,i-1}, L_{2,j-1})$, $(L_{1,i-1}, L_{2,j})$, $(L_{1,i}, L_{2,j-1})$. In order to chose the previous pairs of points, an associated pairs of real $C_{i,j}$ is defined, where $C_{i,j}$ is equal to $(FM(L_{1,i}, L_{2,j}), MD(L_{1,i}, L_{2,j}))$. The candidate pair, where the associated pair of real is inferior or equal to the two other real pair, is chosen. This construction is finished when i and j equal to 1. The algorithm is given by **Fig. 3**.

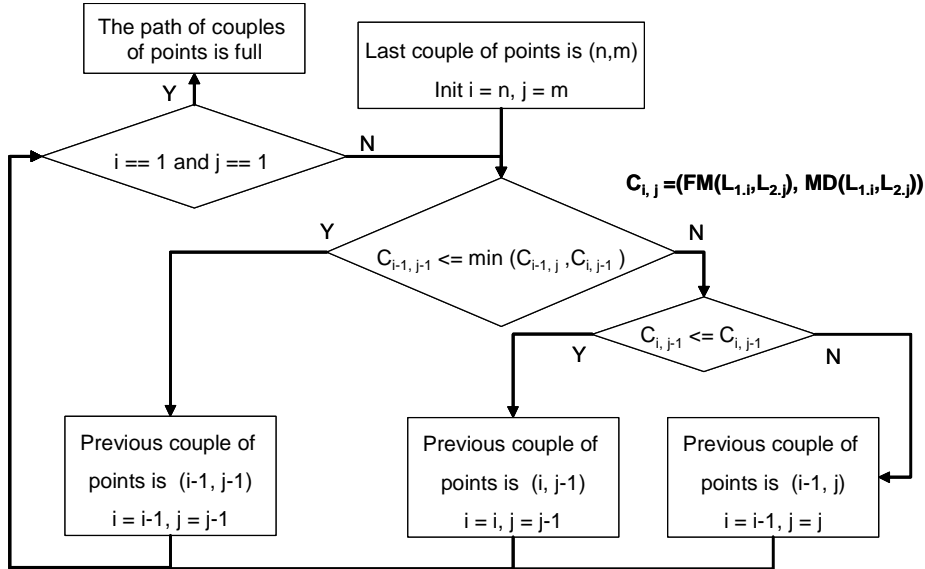


Fig. 3. Algorithm for computing the minimal path (MP)

In the example of **Fig. 1**, the minimal path is represented by the pairs of real in the grey cells of **Table 2**. After added the last couple $(L_{1,7}, L_{2,8})$, the previous couple $(L_{1,6}, L_{2,7})$ is chosen because $FM(L_{1,6}, L_{2,7})$ is inferior or equal to both $FM(L_{1,7}, L_{2,7})$ and $FM(L_{1,6}, L_{2,8})$. Indeed $(1.80, 0.54)$ is infe-

rior or equal to (1.84, 1.84) and (1.80, 1.24). This process is reused until i and j equal 1. So the minimal path is the order set of nine couples: $(L_{1,1}, L_{2,1})$, $(L_{1,2}, L_{2,2})$, $(L_{1,3}, L_{2,3})$, $(L_{1,3}, L_{2,4})$, $(L_{1,4}, L_{2,5})$, $(L_{1,4}, L_{2,6})$, $(L_{1,5}, L_{2,6})$, $(L_{1,6}, L_{2,7})$, $(L_{1,7}, L_{2,8})$. **Fig 4.** shows couples of $(L_{1,i}, L_{2,j})$ associated with minimal path. Points from pairs are homologous.

Table 2. Minimal path from $(L_{1,1}, L_{2,1})$ to $(L_{1,n}, L_{2,m})$ is defined by selecting the couple of real in grey cells. In cell i, j , the first number is $FM(L_{1,i}, L_{2,j})$ and the second one is $MD(L_{1,i}, L_{2,j})$

	1	2	3	4	5	6	7	8
1	0.41 0.41	1.70 1.70	2.00 2.00	2.61 2.61	4.09 4.09	5.46 5.46	6.91 6.91	7.92 7.92
2	3.31 3.31	1.80 1.80	2.01 2.01	2.00 1.63	2.00 0.95	2.45 2.45	4.52 4.52	5.50 5.50
3	3.61 3.61	2.51 2.51	1.80 1.51	1.80 0.90	1.80 1.33	2.11 2.11	3.44 3.44	4.46 4.46
4	5.12 5.12	3.71 3.71	3.26 3.26	2.64 2.64	1.80 1.10	1.80 0.45	2.69 2.69	3.61 3.61
5	6.35 6.35	5.00 5.00	4.37 4.37	3.74 3.74	2.42 2.42	1.80 0.91	1.80 1.66	2.40 2.40
6	6.91 6.91	5.85 5.85	4.77 4.77	4.22 4.22	3.70 3.70	2.52 2.52	1.80 0.54	1.80 1.24
7	8.81 8.81	7.72 7.72	6.68 6.68	6.12 6.12	5.41 5.41	4.02 4.02	1.84 1.84	1.80 0.86

Other processes have been described to define others minimal paths. In order to reconstruct three-dimensional solid from serial sections, Fuchs et al. (1977) propose to find the minimum cost cycles in a directed toroidal graph. To compute the minimum cost cycles, the matrix of Euclidean distance is transformed into a graph and a Dijkstra's Algorithm (Dijkstra 59) is employed to find the shortest path from the vertex $(L_{1,i}, L_{2,j})$ to the vertex $(L_{1,n}, L_{2,m})$.

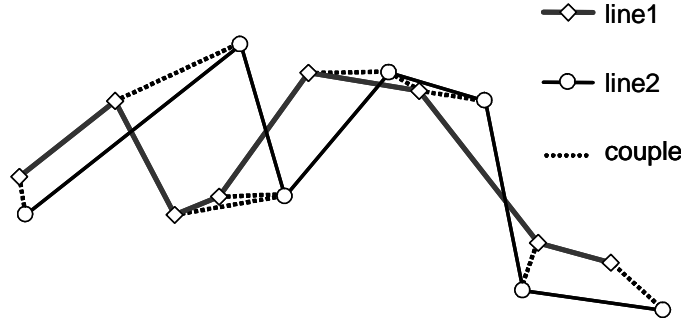


Fig. 4. Couple of $(L_{1,i}, L_{2,j})$ of the minimal path represented by dot lines

This method was translated in other domains such as 2D objects morphing (Sederberg and Greewood 1992). These methods minimize the sum of Euclidean distances between the points of pairs. Thus, a large distance can be chosen if the other distances are small. For computing average distance, MP has the advantage over these last paths to not select pair of points with large distance.

A few softwares (TCI 2005) (Geomod 2005) also employ list of pairs of homologous points as pre-processing for rubber-sheeting. Nevertheless, these pre-processing are semi-automatic and each point can only appear in one couple.

The average Fréchet distance (d_{aF}) used the MP. In the example of **Fig. 1**, the d_{aF} is equal to 0.95 and with a resampling pre-process with 0.1 as LengthMaxSeg, d_{aF} is equal to 0.51. **Table 3** complete the example, showing that resampling is necessary to obtain a value for d_{dF} closed to a value of d_F . Moreover the approximation is inferior to the one of LengthMaxSeg.

Table 3. d_{dF} and d_{aF} computed with different values for LengthMaxSeg

	without	1	0.1	0.01	0.001
d_{dF}	1.8028	1.2260	1.2015	1.2012	1.2012
d_{aF}	0.9524	0.5843	0.5116	0.5030	0.4997

The average Fréchet distance is an accurate measure to match data, to control quality and to merge data. For quality control, the d_{aF} with resampling, is an appropriate measure of the average discrepancy of lines whatever sinuosity.

4 Global process

The previous methods allow to compute measures (d_{dF} and d_{aF}) between homologous lines but a global process is required to work with sets of lines. Among all couple of lines, the program selects pair of lines: L_1 from the first set and L_2 from the second set as the distance between them could be inferior to a maximal distance (MaxDist). This process is divided in three steps (**Fig. 6**).

For the first step, the process makes a query on the line's "open" attribute to distinguish three different cases for matching:

- total or partial matching between open lines,
- partial matching between open line and a close one,
- total matching between close lines.

For the second step, the process determines if lines can be homologous. This step used two bounding-boxes for each line (shows **Fig 5.**):

- BB_i : The bounding-box of L_i : BB_i .
- BBE_i : The enlarged bounding-box of L_i . This rectangle is the BB_i enlarged to contain all points that distance is inferior to MaxDist to BB_i .

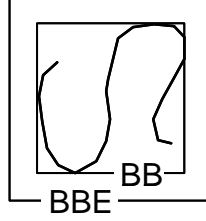


Fig. 5. Bounding-box (BB) and enlarged bounding box (BBE) of one line

If the bounding-box of L_1 is included in the Enlarge bounding-box of L_2 , consequently the Fréchet distance between those two lines can be inferior to the MaxDist previously defined. Hence there is a real probability for these lines to be homologous.

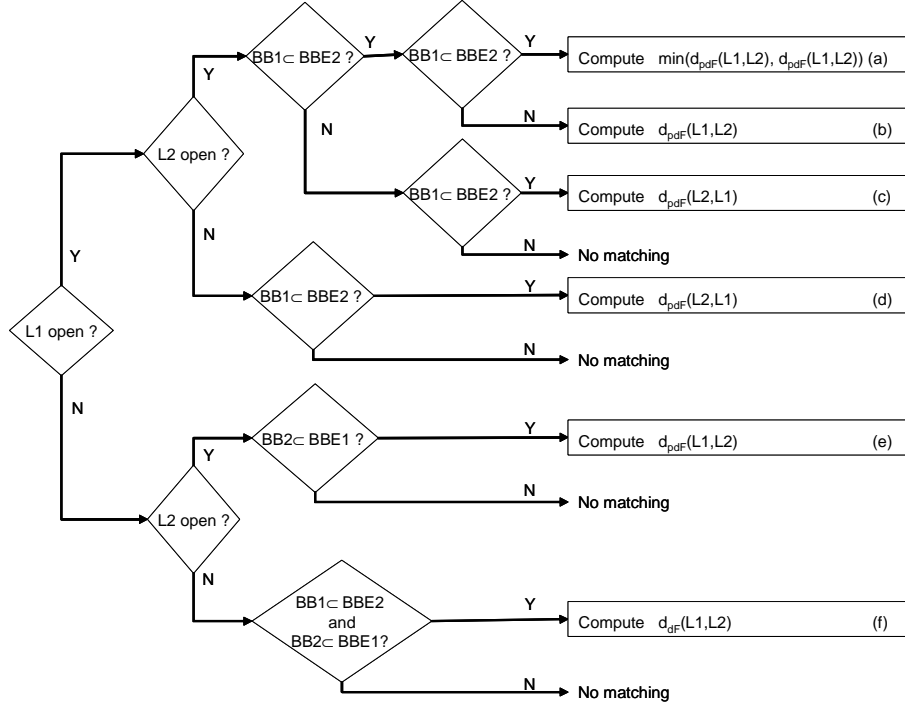


Fig. 6. Global process algorithm for each pair of (L_1, L_2)

The most common and tricky case is when a line is included in the enlarge bounding-box of another (**Fig. 6**, case (b) or (c)) but the reverse is false. This means that lines are homologous but one is shorter, a partial matching process between them is computed. If the reverse is true (**Fig. 6**, case (a)), the two partial Fréchet distances are calculated ($d_{pdf}(L_1, L_2)$,

$d_{\text{pdF}}(L_2, L_1)$ and the minimal distance is kept. The total matching is considered as a sub-set of a partial matching.

In the case of a partial matching between one open line and one close line (**Fig. 6**, case (d) or (e)), the close line is cut according to the open line, then the process calculates the Fréchet distance.

For two close lines (**Fig. 6**, case (f)), if one BB is include in other BBE, the discrete Fréchet distance between two polygons borderlines is computed.

Finally the global process selects only the homologous lines when their Fréchet distances are inferior to MaxDist. The result is a set of pairs of matching lines with their Fréchet distances (average and discrete) and a list of homologous points.

5 Example of Coastline Matching Process

Within the framework of CNES-IFEN (French space agency and French environmental Institute) littoral monitoring, Le Berre et al (2004) has tested the capabilities of SPOT 5 data as a relevant tool for coastal zone mapping and coastline updating. The aim of the project was to assess the potentialities of a high resolution sensor (satellite SPOT 5) to delineate a reference coastline used in many coastal applications like offshore dynamic monitoring, protection works against sea erosion or coastal land-cover mapping.

In addition to their visual interpretation, the measure derives from Fréchet distance complete the study by giving a quantitative evaluation of the distance between different digitized coastlines.

The shoreline is still an ambiguous concept despite a common use as a reference boundary between sea and land. Indeed, a gradual change on both sides and a permanent evolution during time don't allow to define an accurate and permanent boundary. The coastline is defined by IHO (International Hydrographic Organization, 2005) as the line where shore and water meet. Although the terminology of coasts and shores is rather confuse, shoreline and coastline are generally used as synonymous.

For marine application, the coastline is defined as the coast limit reached by the highest level of water (high seasonal tide). For the SHOM (French Marine Hydrographic and Oceanographic service) or Ifremer (French Institute for marine research) it is the conventional limit of the coastal domain at the neighbourhood of the High water line (Coastchart project, 2004). A theoretical definition could be either the highest astronomical tide or the extreme level high water limit on a period from 10 years.

From another point of view, one of geomorphological definitions is “morphological discontinuity area where the sea reaches the coast”. Thus using sedimentary, morphological and botanical features, it’s possible to avoid the problem of tide and therefore to use remote sensing data to map the coastline.

This last definition is chosen to digitize the shoreline from SPOT and orthophotographs data using together vegetation limits on cliffs, sand dunes and schorres, the foot sand dune, the erosion slope, the beach vegetation boundary, or the high water spring tide mark.

5.1 Datasets and methodology

The site of the experimentation is located in the North-West of Brittany (France). It has been chosen according to the availability of reference datasets along with its coastline diversity: cliffs of various heights and rocks (soft or hard), beaches and sand dunes, tidal flats, estuaries, and artificial coast.

The digitization of the shoreline is both based on SPOT 5 image and orthophotographs data from BD ORTHO® (IGN, 2003):

- The satellite image dated 2003/04/17 has 2.5 m resolution with a multispectral band (THR + XS) during low neap tide (tidal range: 114);
- The BD ORTHO® is produced by the French National Mapping Agency (IGN) with aerial photography shot at a 1:25 000 scale in June 2000. They are geometrically and orthogonally corrected with a Digital Elevation Model of the natural ground only (and not the superficial relief). The final product is a real colour picture, with a 50 cm spatial resolution and can be use at a scale of above 1:1 000.

In addition, several detailed topographical surveys of shoreline section were made in order to compare the digitized shoreline to reference lines.

These surveys were made at the same period than the Spot image acquisition with a laser tacheometer or with a differential Global Positioning System (GPS), with a precision close to centimetre, on various parts of the coast (namely sand dunes, low height soft cliff, artificial coast, cobble ridge and shore). The station position was determined with georeferenced positioning points and the topographical survey fit with the plot of the shoreline inflexion point layout in a plan.

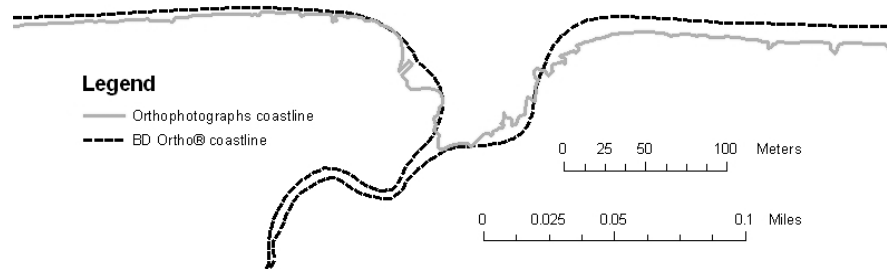


Fig. 7. Illustration of two different coastlines in estuary area.

As described previously, the Hausdorff distance is not accurate for sinuous lines like coastline (see **Fig. 7**), so Fréchet distance is used for this numerical comparison.

5.2 Results:

All the lines are resampled with a range equal to 0.5 meter. Empirically the distance accuracy is about 10 cm. The next tables summarize the results: in each cell, the first number is the discrete Fréchet distance and the second one between brackets is the average Fréchet distance, in meter.

For the digitalization two scales were employed: 1:1500 (SPOT 5 1500) and 1:6000 (SPOT 5 6000).

Table 4. Results for Artificial area

	GPS	Ortho
SPOT 5 1500	14.01 (3.58)	15.75 (3.70)
SPOT 5 6000	6.26 (2.37)	10.49 (2.67)
Ortho	7.37 (1.77)	×

Table 5. Results for Cobble ridge

	GPS	Ortho
SPOT 5 1500	3.39 (1.45)	5.13 (2.78)
SPOT 5 6000	4.18 (2.01)	6.10 (3.28)
Ortho	5.08 (1.82)	×

Table 6. Results for Vegetated cliffs

	GPS	Ortho
SPOT 5 1500	10.84(2.77)	12.20 (4.25)
SPOT 5 6000	11.98 (4.54)	10.32 (6.07)
Ortho	7.25 (2.8)	×

Table 7. Results for Shore area

	GPS	Ortho
SPOT 5 1500	5.20 (1.22)	4.93 (1.67)
SPOT 5 6000	3.90 (0.96)	4.97 (1.47)
Ortho	5.38 (1.41)	×

Table 8. Results for Cliffs top

	tacheometer	Ortho
SPOT 5 1500	3.78 (1.62)	3.84(1.10)
SPOT 5 6000	6.07 (3.34)	6.43 (2.62)
Ortho	3.10 (1.03)	×

Table 9. Results for Beach shoreline

	tacheometer	Ortho
SPOT 5 1500	3.79 (1.17)	10.22(2.37)
SPOT 5 6000	6.08 (2.25)	12.33 (4.14)
Ortho	8.64 (2.1)	×

First of all, for all tables two kinds of discrete Fréchet distance discrepancies are present: one upper to 10 meters (see **Table 4** and **Table 6**), and a second in average of 6 meters.

The discrepancy of ten meters could be explained either by a difference of interpretation between images and reality or changes in morphology (circle in **Fig. 8**). The second is mostly due to a problem of data resolution (see **Fig. 8**). Indeed, the digitization of points itself is an uncertain process. Even though we choose a precise scale to get the points, it's impossible to be sure that the coordinates are correct (Harvey and Vaughlin 1996).

In artificial area (see **Table 4**) the surprising best digitalization at 1:6000 than SPOT 5 1500 is only a consequence of a difference of interpreted features. Despite that, the best support remains orthophotographs (25% better than SPOT 5) due to its resolution which allow a better identification of build up areas.

The good results of average distance in “shore area” with satellite image (see **Table 7**) may be due to an evolution of the shore between 2000 and 2003, the dates the orthophotographs and the Spot image have been acquired. The same comment can be done for the cobble ridge (**Table 5**) and especially for the sand dune results (**Table 9**)

The measures obtained for cliffs are not homogeneous: **Table 6** gives a bigger Fréchet distance than **Table 8**. For “vegetated cliffs” the interpretation is disturbed, both with SPOT 5 or the orthophoto, by the vegetation that can mask the location of the *de facto* coastline. In comparison, “cliffs top” results (**Table 8**) show that orthophotographs are most convenient for the shoreline digitization.

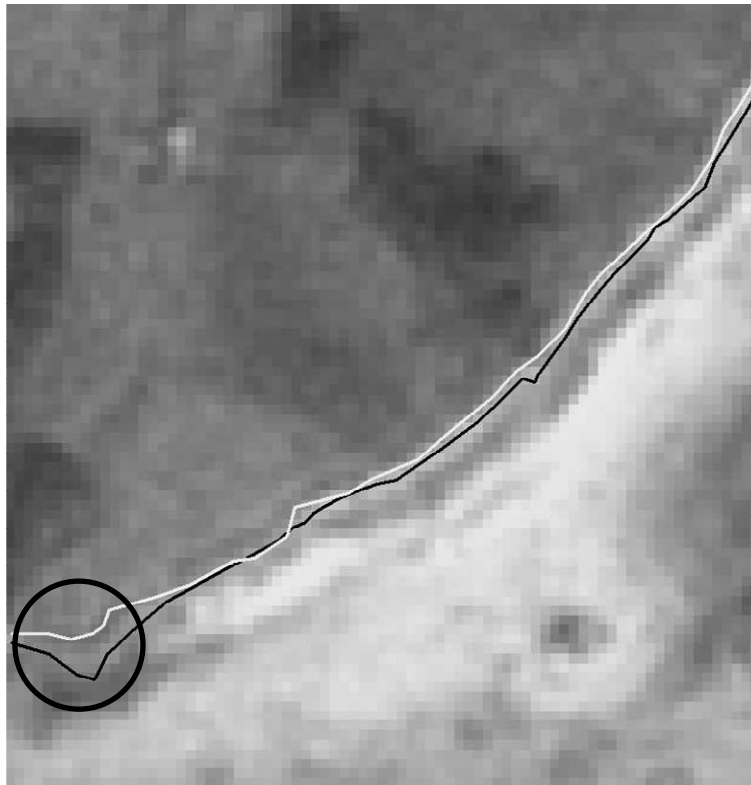


Fig. 8. Illustration of two kinds of errors: maximum Fréchet distance and average Fréchet distance. In the circle, the Fréchet distance between lines is bigger than between the reminder.

To summarize, the global process application had demonstrated that SPOT 5 and the orthophotographs may be the support of the digitizing of

coastlines with comparable planimetric accuracy. Actually, these quantified results were really surprising as we felt during the whole experimentation process that the identification of coastlines was more difficult on Spot 5 according to its lower spatial resolution. Indeed the results given by the program of discrete Fréchet distance show that SPOT 5 data are a relevant support for short term shoreline monitoring. This method can also be useful for over type of data matching. For example, the different processes concerning the matching of different networks (road, hydrological, electrical) present in (Mustière 2006), could be improved by an integration of this method.

6 Discussion

An important point of the research objectives is to extend the global process to the 3D. Thus future works tends to develop an integration matching method whatever the type of the Digital Elevation Model (DEM) is. The main issue is to solve the question of matching two DEMs in coastal area: by using matching surfaces between them, or forced lines obtained with a DEM enhancement.

In the context of seamless elevation model integration previous studies give a start point. In coastal domain, Gesch and Wilson (2001) propose an ad-hoc method which first converts each DEM in the same common vertical reference. Then, after remove false old bathymetric points, a raster surface model is produced from topographic and bathymetric points in the zero area elevation. Finally they merge the two first DEMs together using the third one to avoid interpolation edge effects.

In the continental area, Podobnikar (2005) proposes to average each DEM cell. As the obtained DEM is smoother than the input DEM, he enhances the result with geomorphological feature such as land-marks, hydrological network, or land registry points.

The problems with these methods are first a partial or total overlapping is required and second they didn't take care of an eventual planimetric shift between cells.

To solve that, a new method based on Fréchet matching must be defined. One of the most interesting contributions could be to process a geomorphological enhancement after landscape segmentation in order to obtain morphological forced lines. The landscape segmentation uses taxonomy of different land types: rock or reef formation, beach, estuaries, build up areas, and so on. For each type, main features of these relief ele-

ments (equivalent to “forced line”) are determined as ridge, thalweg, line of slope break, roughness.

Then the homologous force lines are identified and matched on each DEM. Finally, the distorted DEM are merged.

This method gives the advantage of merging different coordinate cells with the consideration of the type of landscape. So it's equivalent as making an integration with adapted enhancement and merging method with a planar control, taking care planimetric shifts and not only according the altimetric shift.

7 Conclusion

In the context of data matching and quality control, the Fréchet distance, where a good approximation is the discrete Fréchet distance (d_{Fd}), is a relevant tool for measure differences. A new distance is also defined from it: The Average Fréchet Distance (d_{aF}). While the d_{Fd} represents the maximal gap somewhere between two homologous points of lines, d_{aF} gives the average difference whatever sinuosity. These two measures are complementary to control lines quality. Furthermore, a global process is implemented in order to automatizes the matching process between two datasets.

The coastline matching process was done with the discrete Fréchet distance as it's the most appropriate tool to find sinuous homologous line, instead of Hausdorff distance.

In order to determine if SPOT 5 is a relevant tool for coastline mapping, we have implemented a program based on this method. The program gives a quantitative methodology to compare two chosen coastlines. The measures demonstrate that there is a good accuracy between SPOT data and reference data, despite a lower spatial resolution and temporal mismatch.

As these processes are appropriate to 2D data, further developments will extend the matching to 3D data, according to the type of relief in the coastal area.

References

- Alt H, Godau M (1995) Computing the Fréchet distance between two polygonal curves. In: International Journal of Computational Geometry & Applications, vol 5, nums 1-2, pp 75-91
- Coastchart project (2004) Demonstration of an EO based service for the update of marine charts, <http://www.logicacmg.com/COASTCHART/index.aspx.html>

- Deng M, Chen X, Li Z (2005) A Generalized Hausdorff Distance for Spatial Objects in GIS. In: Proceedings of the 4th ISPRS Workshop on Dynamic and Multi-dimensional GIS, University of Glamorgan, Archives of ISPRS, University of Glamorgan, pp 10-15
- Devogele T (2002) A new Merging process for data integration based on the discrete Fréchet distance. In: Proceedings of the 10th International Symposium on Spatial Data Handling (SDH), Richardson D, van Oosterom P (eds.), Springer, pp 167-181
- Devogele T, Trevisan J, Raynal L (1996) Building a multi-scale database with scale-transaction relationships. In : Proceedings of the 7th International Symposium on Spatial Data Handling (SDH), Kraak, Molenaar (eds.) Taylor & Francis, pp 337-351
- Dijkstra E (1959) A note on Two Problems in Connection with Graphs. *Numerische Mathematik*, Vol. 1, pages 269-271
- Eiter T, Mannila H (1994) Computing Discrete Fréchet Distance. Technical report of Christian Doppler Labor für Expertensysteme. Vienna University of technology, num. CD-TR 94/64
- Fuchs H, Kedem Z M, Uselton S P (1977) Optimal Surface reconstruction from planar contours, In: Graphics and image processing, Foley (ed), ACM, vol 20 n° 10, pp 693-702
- Gesch D, Wilson R (2001), http://nauticalcharts.noaa.gov/bathytopo/ESRI_UC2001/ESRI_UC_paper.html
- Geomod (2005) CadSIS, <http://www.geomod.fr/logiciels/dvpt/dvpt.htm>
- Harvey F et Vauglin F (1996) Geometric Match-processing : Applying Multiple Tolerances. In : Proceedings of the 7th International Symposium on Spatial Data Handling (SDH), Kraak, Molenaar (eds.) Taylor & Francis, pp 155-171.
- IGN (2003) BD ORTHO® Version 1, descriptif de contenu
http://www.ign.fr/telechargement/MPro/produit/BD_ORTHO/DC_BDORTH O.pdf
- International Hydrographic Organization (2005) HYDROGRAPHIC DICTIONARY 5th edition, Special Publications N° 32 of IHO
- Le Berre I, Henaff A, Wenzel F, Giraudet J (2004) Cartographie synthétique de l'environnement littoral du Finistère, exploitation de SPOT pour la cartographie de l'estran, du trait de côte et de l'occupation du littoral. In : Rapport final, Appel à proposition CNES/IFEN " Suivi du littoral par SPOT 5", GEOMER laboratory/Cetmef/DDE29
- Mustière S (2006) Results of experiments on automated matching of networks at different scales. In: Joint ISPRS Workshop on Multiple Representation and Interoperability of spatial Data, Hannover, Germany, February 2006
- Podobnikar T (2004) Production of integrated digital terrain model from multiple datasets of different quality. In: International Journal of Geographical Information Sciences, vol 19, num. 1, January 2005, pp. 69-89
- Sederberg T W, Greenwood E (1992) A Physically based Approach to 2-D shape blending, SIGGRAPH'92, vol 26, num. 2, pp. 25-34

- TCI software (2005) Adjust - True Rubbersheeting inside of AutoCAD, <http://tcicorp.com/html/adjust.html>
- Veregin H (1999) Data quality parameters. In : Geographical Information systems. Longley, Goodchild, Maguire, Rhind (eds), John Willey & Sons, inc, pp. 177-189
- Whitfield M, Pepper J (2003) Integrated coastal zone-Data research project (ICZMap®). http://www.thsoa.org/hy03/4b_2.pdf