

## Binary-image comparison with local-dissimilarity quantification

Étienne Baudrier<sup>a,\*</sup>, Frédéric Nicolier<sup>b</sup>, Gilles Millon<sup>b</sup>, Su Ruan<sup>b</sup>

<sup>a</sup>Laboratoire de Mathématiques et Applications, Université de La Rochelle, Avenue Crépeau, 17042 La Rochelle Cedex 1, France

<sup>b</sup>Centre de Recherche en STIC, IUT de Troyes, URCA, 9, rue de Québec, 10026 Troyes Cedex, France

Received 14 June 2006; received in revised form 9 January 2007; accepted 2 July 2007

### Abstract

In this paper, we present a method for binary image comparison. For binary images, intensity information is poor and shape extraction is often difficult. Therefore binary images have to be compared without using feature extraction. Due to the fact that different scene patterns can be present in the images, we propose a modified Hausdorff distance (HD) locally measured in an adaptive way. The resulting set of measures is richer than a single global measure. The local HD measures result in a local-dissimilarity map (LDMap) including the dissimilarity spatial layout. A classification of the images in function of their similarity is carried out on the LDMaps using a support vector machine. The proposed method is tested on a medieval illustration database and compared with other methods to show its efficiency.

© 2007 Elsevier Ltd. All rights reserved.

*Keywords:* Binary images; Hausdorff distance; Similarity measures; Spatial dissimilarity layout; Local analysis

### 0. Introduction

Image comparison is widely used in different domains: image retrieval [1], image classification [2], shape matching [3], image quality evaluation [4], registration [5]... Methods found in literature for image comparison can be classified into two approaches: (a) an image feature extraction (shape, curve, texture, histogram) followed by a feature comparison; (b) straight image comparison. For the first approach, conspicuous features must be captured in the signature of each image in order to be as discriminating as possible in some user defined way [6]. The choice of the signature attributes is not always easy and depends on the processed images [5,1]. In case that there are several patterns in the images, a segmentation is often necessary to compare locally the attributes. For binary images, object shapes cannot always be precisely identified, so it is difficult to find the features related to shapes. Moreover, the texture attribute is also difficult to extract (binary images are not always

textured), and the color attribute is poor (only black and white colors). Thus, the second approach, a straight image comparison, seems adapted in the case of binary images. In our work, the measure is windowed and the window size is adjusted so as to measure exclusively the local dissimilarity. The obtained result is composed of a set of local measures covering the image. The following state of the art gives a general presentation of dissimilarity measures and tends to choose the Hausdorff distance (HD). The computation of windowed measure all over the images is time consuming, nevertheless, for the HD, the algorithm comes down to a formula based on the distance transform (DT) of each image which accelerates the computation.

The paper is organized as follows: firstly we present an overview of the image retrieval in Section 1.1 and then focus on the HD in Section 1.2. Secondly the notion of local HD is introduced and its properties are exposed in Sections 2 and 3. This allows us to determine automatically the size of the local HD's window in function of the dissimilarity, as shown in Section 4. Then it is shown how the algorithm may be reduced to a formula based on the DT. The map of local HD defined without parameters is presented in Sections 4 and 5. Finally, qualitative results and an application to image classification are presented in Section 6 before concluding.

\* Corresponding author. Tel.: +33 5 46 45 83 04, fax: +33 5 46 45 82 40.  
E-mail addresses: [etienne.baudrier@univ-lr.fr](mailto:etienne.baudrier@univ-lr.fr) (É. Baudrier),  
[frederic.nicolier@univ-reims.fr](mailto:frederic.nicolier@univ-reims.fr) (F. Nicolier), [gilles.millon@univ-reims.fr](mailto:gilles.millon@univ-reims.fr)  
(G. Millon), [su.ruan@univ-reims.fr](mailto:su.ruan@univ-reims.fr) (S. Ruan).

## 1. State of the art

### 1.1. Image retrieval

Similarity measure is a difficult problem. The first approach using a feature extraction can be usually found in the field of image retrieval. Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. “Content-based” means that the search makes use of the contents of the images themselves, rather than relying on human-imputed metadata such as captions or keywords.

The ideal CBIR system from a user perspective would involve what is referred to as semantic retrieval. This type of open-ended task is very difficult for computers to perform. Real CBIR systems, being developed since the early 1980s [7] generally retrieve visual information by its content based on lower-level features. Features, that are generally color, shape and texture, are used to select good matches in response to the user’s query [1]. Different implementations of CBIR make use of different types of user queries.

- With query by example, the user searches with a query image (supplied by the user or chosen from a random set), and the software finds images similar to it based on various low-level criteria [8,9].
- With query by sketch, the user draws a rough approximation of the image they are looking for, for example with blobs of color, and the software locates images whose layout matches the sketch [10,11].
- Other methods include specifying the proportions of colors desired.

Once extracted, the features are compared with the query using a similarity measure, which is designed according to the nature of features. For shape feature, the similarity measure can be calculated on the curvatures, the moments or shape contours [3]. Feature histogram can also be used for the comparison [12]. Some improvements can be obtained when the spatial relationships between the features are considered [13]. Some authors represent the characterizations of the features in a multidimensional vector space. A distance measure is then used to estimate their similarity. The best known distances are the Euclidean distance, the Chamfer distance, the HD and the Mahalanobis distance [14]. Since the comparison is carried out between the extracted features, the distance measures are globally calculated. The decision upon similarity can then be taken from the measured distance.

In the case of a straight image comparison, the second approach, image intensity or color is generally used for the comparison. But for binary images, the intensity information is very poor, because there are only background pixels and foreground pixels (scene information). Therefore, binary image comparison can be considered as a comparison between a set of foreground points. In Ref. [15], several measures of

correspondence between binary images are described and compared. This study shows that distance-based measures perform better comparisons of binary images than measures based on set memberships. The most common measures are based on the Euclidean distance, on the 1–1 correspondence distances and on the HD. The measures based on a 1–1 correspondence comprise the bottleneck distance, minimum weight matching, uniform matching and minimum deviation matching [14]. They are used in graph theory and imply to find a correspondence between the points of the two images. In this case, it is necessary for images to have the same number of points or to determine a number of points to match, which is usually delicate. The HD on which our measure is based, is a *max–min* distance and it does not have this inconvenience. It is widely used, from face recognition [16] and binary pattern matching in images [17].

In case of binary images, shape extraction is often deemed difficult to carry out and the extraction of connected pixels does not allow efficient shapes extraction. Local information is therefore difficult to access to by this way. A better way to access to local representation is to measure a local distance. Recently Wang [4] has exploited the known characteristics of the Human Visual System to produce a Structural Similarity Index. This index of local measures could be seen as a measure of local dissimilarity but it is done through a fixed size window ( $8 \times 8$  pixels). Since the choice of the window size is fixed and independent of the image content, the obtained results cannot give precise comparison. In this paper, we will show the ability of giving an adaptive measure of the local distance in the case of straight binary image comparison.

### 1.2. Overview of the HD

#### 1.2.1. The HD as a dissimilarity measure over binary images

Among dissimilarity measures over binary images, the HD has often been used in the content-based retrieval domain and is known to have successful applications in object matching [18,19], in face recognition [20,21] or in learning [22]. It can be computed quickly using Voronoi diagrams [23]. Let us have a brief review of the definition and some properties related to the HD. Originally meant as a measure between two point collections  $F$  and  $G$  in a metric space  $E$  (whose underlying distance is  $d$ ), it can be viewed as a dissimilarity measure between two binary images  $\mathbf{F}$  and  $\mathbf{G}$ , considering  $F$  and  $G$  respectively the black pixels of  $\mathbf{F}$  and  $\mathbf{G}$ . For finite sets of points, the HD can be defined as [17]:

**Definition 1 (HD).** Given two non-empty finite sets of points  $F = \{f_1, \dots, f_{N_F}\}$  and  $G = \{g_1, \dots, g_{N_G}\}$  of  $\mathbb{R}^2$ , and an underlying distance  $d$ , the HD is given by

$$D_H(F, G) = \max(h(F, G), h(G, F)), \quad (1)$$

where

$$h(F, G) = \max_{f \in F} \left( \min_{g \in G} d(f, g) \right), \quad (2)$$

$h(F, G)$  is the so-called *directed HD*. For images, we use the same notation:  $D_H(\mathbf{F}, \mathbf{G}) = D_H(F, G)$ .

The interest of this measure comes firstly from its metric properties: non-negativity, identity, symmetry and triangle inequality. These properties correspond generally to our intuition for shape resemblance.

Moreover, the HD is a match methodology without point-to-point correspondence, so it is robust to local non-rigid distortions. Another source of interest is the following property:

**Proposition 2 (Translation).** *Let  $v$  be a vector of  $\mathbb{R}^2$ ,  $T_v$  translation of vector  $v$  and  $F$  a non-empty finite set of points, then*

$$D_H(F, T_v F) = \|v\|. \quad (3)$$

It implies that for a small translation, the value of the HD is small, which matches our expectation for a dissimilarity measure.

### 1.2.2. Some modified versions of the HD

The classical HD has good properties but it measures the most mismatched points between  $F$  and  $G$ . Indeed, considering two images containing the same pattern and one point added to the first image, far from the pattern, then the HD will measure the distance between the pattern and the point. As a consequence it is sensitive to noise [24].

Several modifications of the HD have been proposed to improve it. The next definitions are detailed in Ref. [16].

The directed distance of the partial HD (PHD) is defined in [17]:

$$h_K(F, G) = K_{f \in F}^{th} d(f, G), \quad (4)$$

where  $K_{f \in F}^{th}$  denotes the  $K$ th ranked value of  $d(f, G)$ . Thus, the PHD depends on a parameter  $p = K/N_F$  standing for the proportion of values taken into account. The PHD method yields good results for the case of impulse noise.

The directed distance of the modified HD (MHD) is defined in [25]:

$$h_{MHD}(F, G) = \frac{1}{N_F} \sum_{f \in F} d(f, G). \quad (5)$$

Unlike the PHD, the MHD measure does not require any parameters. The method adapts only to the case of Gaussian noise.

The directed distance of the weighted HD (WHD) is defined in [16]:

$$h_{WHD}(F, G) = \frac{1}{N_F} \sum_{f \in F} w(f) \cdot d(f, G), \quad (6)$$

where  $\sum_{f \in F} w(f) = N_F$ . An image can be divided into different parts, and the contribution of the different parts to the image matching may vary, therefore the HD should change. The WHD has been used in the Chinese character image matching [26,27] and in face recognition [28].

The directed distance of the censored HD (CHD) is defined by [24]:

$$h_{k,l}(F, G) = P_{f \in F}^{th} Q_{g \in G}^{th} d(f, g), \quad (7)$$

where  $P^{th}$  denotes the  $P$ th ranked value of  $Q_{g \in G}^{th} d(f, g)$ , with  $Q_{g \in G}^{th}$  representing the  $Q$ th ranked value of the underlying distance set. Since the CHD ranks the underlying distance, the effect of the impulse noise to the image is reduced.

The directed distance of the “doubly modified” HD (M2HD) is defined by [20]:

$$h_M = \frac{1}{N_F} \sum_{f \in F} d(f, G) \quad (8)$$

with  $d(f, G) = \max(\mathbf{I} \min_{g \in N_G^f} d(f, g), (1 - \mathbf{I})P)$ , where  $N_G^f$  is a neighborhood of the point  $f$  in the set  $G$  and  $\mathbf{I}$  indicates if there exists a point  $g \in N_G^f$ .

The directed distance of the least trimmed squared HD (LTS-HD) is defined in [29]:

$$h_{LTS}(F, G) = \frac{1}{K} \sum_{i=1}^K d(f, G)_{(i)}, \quad (9)$$

where  $K$  denotes  $p \times N_F$ , as in the PHD case, and  $d(f, G)_{(i)}$  represents the  $i$ th distance value in the sorted sequence  $d(f, G)_{(1)} \leq d(f, G)_{(2)} \leq \dots \leq d(f, G)_{(N_F)}$ . The measure  $h_{LTS}(F, G)$  is minimized by remaining distance values, after large distance values are eliminated. Even if the object is occluded or degraded by noise, this matching scheme yields good results.

### 1.2.3. Discussion

It is noticeable that except for the MHD, at least one arbitrary parameter has to be determined. The parameter must be chosen to make the measure as discriminating as possible and it depends upon the type of image to be studied, and sometimes on the characteristics of the compared images in the same application (e.g. more or less dark or noisy images). The MHD measure does not require any parameters. However its matching performance is not as efficient as those of the PHD and the CHD, due to the summation operator over all distances, some of which might be computed from outliers. Moreover, these measures are global and cannot account for local dissimilarities. Indeed, the HD is a “max–min” distance which implies that the value of the HD between two images is reached for at least one pair of points. But it does not indicate if the value is reached for a single pair of points or for several pairs of points. In this last case, if they are gathered in a part of the images or widespread all over the images, which corresponds to different degrees of dissimilarity. These remarks motivate us to design a local and parameter-free HD in the next section.

## 2. Local HD measure

In this section, the notion of local dissimilarity is first discussed, then a naive definition of a HD measure in a local window is presented. This naive measure is a simple adaptation of the global HD measure to the local window. Nevertheless, the HD measure is not defined if one of the two sets is empty, which can happen in the case of a local measure. Moreover,

during the local HD use, the window can be moved and resized. Consequently, a new definition of the local HD is given to make the measures coherent in this case. In all this section,  $F$  and  $G$  design two non-empty finite sets of points of  $\mathbb{R}^2$ , and  $W$  a convex closed subset of  $\mathbb{R}^2$ .

2.1. What is a local dissimilarity?

Producing locally a dissimilarity measure implies to compare the two images locally. It can be done thanks to a sliding window. The parts of both images viewed through this window are compared based on a dissimilarity measure. The sliding-window size plays an important rôle: it should fit the local dissimilarity so that the distance can give a local measure. As the dissimilarity size is a high-level notion related to semantics, it is not desirable to give a precise definition of it in this low-level study. Nevertheless, here is a rough idea: If the pixels located in the sliding window belong to coarse features, the window should be big enough to grasp feature’s dissimilarities. Similarly, for fine features, a window “bigger” than the features will include unwanted information on dissimilarities. Therefore, it is necessary to adapt the size of the window to obtain precise measures.

The meaning of “local” in “local dissimilarity” has also to be clarified. We make the assumption that a local dissimilarity must concern information involving the central pixels in the window, i.e. pixels whose distance from the central pixel is less than sliding step  $p$  (in our case, as  $p = 1$ , there is only the central pixel of the window).

2.2. Naive definition

It consists in modifying the definition of the global measure (def. 1) by introducing a subset standing for the window:

**Definition 3** (HD in a window (naive)).

$$HD_W(F, G) = \max(h_W(F, G), h_W(G, F)), \tag{10}$$

where

$$h_W(F, G) = \max_{f \in F \cap W} \left( \min_{g \in G \cap W} d(f, g) \right). \tag{11}$$

This definition is naive because it supposes that the elements of information extracted through a window remains an image. But it is possible that one of the images may not have any (black) point in the window. Yet, the HD is not defined on empty sets. In the next paragraph, the definition is modified to make it available for empty sets. Moreover, it is also designed to be coherent when the window is moved or resized.

2.3. Improvement of the naive definition

The issue is to attribute a value to a measure between a non-empty point set and an empty set. Indeed, this is an unforeseen case by the global HD because the maximum or the minimum that takes place in the definition cannot be

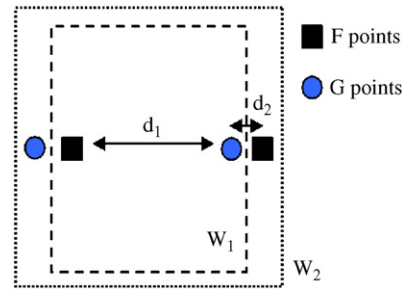


Fig. 1. For the naive definition,  $H_{W_1}(F, G) = d_1 > H_{W_2}(F, G) = d_2$ , which is not intuitive.

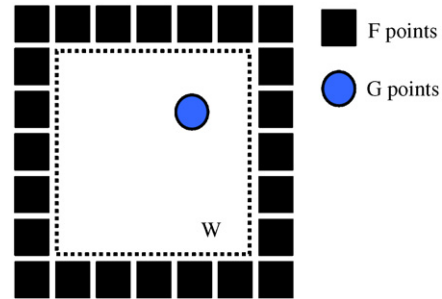


Fig. 2. Example for naive definition improvement. Which value can be attributed to the measure when one of the sets is empty in the window? This value depends on the point arrangement outside of the window. The arrangement which gives the lowest value is the one presented above.

calculated. The modification of the definition should respect the following principles:

- The measure value shall not decrease if the window size increases. More precisely, for a window  $V$  that contains the window  $W$ , the measure value in  $V$  shall not be strictly smaller than the measure in  $W$ .
- The different expressions obtained (in the case of one empty set, in the case of two non-empty sets, etc.) shall be consistent so as to have smooth transition when the window is modified.

The first item can be justified as follows: an increase of the window size brings a new piece of information on the dissimilarity, therefore the evaluation of the dissimilarity in the window shall not decrease (see Fig. 1). The second item gives the possibility to have a sliding window where the different expressions of the HD can be used successively.

Let us bring an intuitive justification of the following new definition for the HD in a window that takes into account the above items. The reason why the measure value decreases is the presence of new points (not in  $W$  but in  $V$ ) whose layout is not known in advance. A simple way to respect the non-decreasness principle is to envisage the case of the new points leading to the smallest measure value. This configuration is as following: there are  $G$  points all around  $W$  (see Fig. 2). Indeed, these points are the closest ones to the window border and thus give the smallest measure when the window increases. But considering the distance of  $F$  points to the  $G$  points all around the window

border comes down to consider the distance of the  $F$  points to the  $W$  border. That is what is improved in the new definition.

#### 2.4. Windowed HD

Definition 4 given in this section includes the distance to the window border so the mathematical notion of frontier and on its discrete version have to be detailed:

- The frontier  $Fr(W)$  comes from the topology defined by the metric  $d$ . In the application (Section 6), the distance is the one associated to the norm  $L_\infty$ .
- In the discrete case, we consider that the frontier is between the pixels. For example, the frontier of the ball  $B(x, n)$  is the line between  $B(x, n)$  and  $B(x, n + 1) \setminus B(x, n)$ . The distance of a point  $z \in B(x, n)$  to the frontier is equal to the distance to the pixels just after the frontier. Thus, for a point  $z$  in  $B(x, n)$  and on the side of  $B(x, n)$ , its distance to the frontier is equal to 1.

**Definition 4 (Windowed HD).** Let  $F, G$  be two bounded sets of  $\mathbb{R}^2$ .

$$HD_W(F, G) = \max(h_W(F, G), h_W(G, F)),$$

where there are three cases:

- (1) If  $F \cap W \neq \emptyset$  and  $G \cap W \neq \emptyset$ ,

$$h_W(F, G) = \max_{f \in F \cap W} [\min(\min_{g \in G \cap W} d(f, g), \min_{w \in Fr(W)} d(f, w))],$$

- (2) if  $F \cap W \neq \emptyset$  and  $G \cap W = \emptyset$ ,

$$h_W(F, G) = \max_{f \in F \cap W} [\min_{w \in Fr(W)} d(f, w)],$$

- (3) if  $F \cap W = \emptyset$ ,  $h_W(F, G) = 0$ .

**Remark 5.** (a) In case there is no point of  $F$  neither of  $G$  in  $W$ , both of the directed distances are equal to 0 and therefore the global distance too. Which is coherent with the fact that the two extracted parts are equal.

(b) In case there is exactly one set without point in  $W$ , one of the two directed distances is equal to 0 and the expression of the other one takes into account the edge distance.

The main difference with the classic HD definition is the introduction of the term  $\min_{w \in Fr(W)} d(f, w)$  in the first and in the second case. This term can be seen as the translation of the fact that the measure is done in a window and not on the whole images. Moreover in the second case, this term substitutes for  $\min_{g \in G \cap W} d(f, g)$  that is not available since  $G \cap W = \emptyset$ .

### 3. Properties of the windowed HD

In this section, useful properties are demonstrated. The criterion proposed in Section 4 is based on these properties:

- the  $HD_W$  is between 0 and  $HD(F, G)$ ,
- it is non-decreasing when embedded growing windows  $W$  are considered.

#### 3.1. General properties

$HD_W$  is non-negative and symmetric by definition.

**Proposition 6 (Identity).** Let  $F, G$  be two bounded sets of points of  $\mathbb{R}^2$ , and  $W$  a star-shaped closed subset of  $\mathbb{R}^2$ .

$$HD_W(F, G) = 0 \iff F \cap W = G \cap W. \quad (12)$$

**Proof.** ( $\Leftarrow$ ) Trivial. ( $\Rightarrow$ ) If  $HD_W(F, G) = 0$ .

Then both directed distances are equal to zero, and as the distance to the points of  $Fr(W)$  is never equal to zero, it remains to the property of the classical HD which is a metric.  $\square$

The following properties need the window  $W$  to be a ball.

**Proposition 7 (Boundary).** Let  $x \in \mathbb{R}^2$  and  $r > 0$ , and let define  $W = B(x, r)$  then  $HD_W(F, G) \leq HD(F, G)$ .

**Proof.** See Appendix A.1.

So when the window  $W$  slides all over the two images, the values in the produced dissimilarity map will remain between 0 and  $HD(F, G)$ . But what happens when the size of the window  $W$  is increased? It seems intuitive that the value increases as well. This is the object of the following proposition.

#### 3.2. Property depending on the window-size

This property ensures that the value measured in a window does not decrease when the window is enlarged. The information taken into account when the window is enlarged does not reduce the former dissimilarity-measure value.

**Proposition 8 (Growth).** Let  $V = B(x_v, r_v)$  and  $W = B(x_w, r_w)$  be two close discs such as  $V \subset W$  then  $HD_V(F, G) \leq HD_W(F, G)$ .

**Proof.** See Appendix A.2.

This property is important because it traduces the fact that the growth of the window will not reduce the measure of the local dissimilarity.

The properties of boundary and growth give a frame to a window-size-criterion definition. It remains now to find a criterion that finds the window growth in order to measure the local dissimilarity.

### 4. A parameter-free, adaptative, local HD

According to Section 2.1, the notion of local dissimilarity needs to be precised mathematically to define the criterion for the optimal window size. This will be treated in the next subsection.

In all the section,  $F$  and  $G$  design two non-empty finite sets of points of  $\mathbb{R}^2$ . The next property specifies the conditions on  $F$  and  $G$  to obtain the maximum value for  $HD_W(F, G)$ . Then,

the notion of local HD dissimilarity is defined. This finally enables the definition of an optimal size for a window  $B(x, r)$  in function of  $F$  and  $G$ .

**Lemma 9 (Maximum value).** *Let  $x \in \mathbb{R}^2$  and  $r > 0$ , and let a window  $B(x, r)$  then  $\sup_{F,G} HD_{B(x,r)}(F, G)$  is reached only when there is exactly one  $F$  point (resp.  $G$ ) at the center of  $W$  and no  $G$  point (resp.  $F$ ) except maybe on  $Fr(W)$ , then  $HD_{B(x,r)}(F, G) = r$ .*

**Proof.** Without loss of generality, we can focus on  $h_W(F, G)$ . For points of  $F$  in  $W$  that are not at the center of  $W$ ,  $\min_{w \in Fr(W)} d(f, w) < r$  then  $\min(\min_{g \in G \cap W}(d(f, g), \min_{w \in Fr(W)} d(f, w)) < r$ . For  $f_0$  at the center of  $W$ ,  $\min(\min_{g \in G \cap W}(d(f_0, g), \min_{w \in Fr(W)} d(f_0, w)) = r$  and it is reached for the points on  $Fr(W)$ . For the other points of  $F$  in  $W$  if there are any of them, this  $\min(\min_{g \in G \cap W}(d(f_0, g), \min_{w \in Fr(W)} d(f_0, w))$  is below  $r$ , and so the max over the points of  $F$  is  $r$ .  $\square$

The aim here is to quantify the local dissimilarity. We assume that the central pixel of the window  $W$  belongs to the features which compose the local dissimilarity. We also want the window to exclude other dissimilarities (i.e. related to others features). Thus, the measure must concern:

- a central point: if it is not involved, the window can be moved to put one of the included points at its center, and
- an edge point: if none of the edge points is involved, the window size can be reduced.

In this case, Lemma 9 implies that the measure in the window reaches its maximum value. Heuristically, the value measured in the window will be maximum if the window is smaller or equal to the ideal size (regarding the local dissimilarity size) and it will not be maximum if the window is too large. From which the local measure is defined as:

**Definition 10 (Local measure).** A window  $B(x, r)$  is said to give a local measure when the measure of the HD in the window  $B(x, r)$  is maximum:  $HD_{B(x,r)}(F, G) = r$ .

It is necessary to know if there is a maximum local measure. So, let  $x \in \mathbb{R}^2$  and  $r > 0$ , let us define:

**Definition 11 (Local-measure set).** The local-measure set  $\mathcal{R}$  is given by

$$\mathcal{R} = \{r > 0 / HD_{B(x,r)}(F, G) = r\}. \tag{13}$$

When  $\mathcal{R}$  is non-empty, it is bounded by  $HD(F, G)$  (Proposition 8), so it has an upper boundary  $r_{max}$ . This leads to the definition:

**Definition 12 (Maximum local measure).** For  $x \in \mathbb{R}^2$  fixed, if  $\mathcal{R}$  is not empty,  $r_{max} = \sup(\mathcal{R})$  is named the optimal radius and for this radius,  $HD_{B(x,r_{max})}$  is said to give the maximum local measure  $r_{max}$ .

It can be deduced from Lemma 9 that there are two possibilities:

- If  $x$  belongs neither to  $F$  nor to  $G$  or if  $x$  belongs to  $F$  and  $G$ , then by Lemma 9,  $HD_W$  will not be maximum for any window size: the measure will be equal to 0.
- If  $x$  belongs only to one of the two subsets  $F$  and  $G$ , for example  $F$ . Then, as  $G$  is finite,  $\exists g_{min} \in G \cap W$ , the closest point to  $x$  (for the distance  $d$ ). By Lemma 9, we know that for  $0 \leq r \leq d(f_0, g_{min}) = r_m$ ,  $W$  will give a local measure and for  $r > r_m$ , as  $\min_{b \in G \cap W}(f_0, g) = r_m < r$ ,  $HD_W$  cannot have its maximum value. So the measure will be

$$d(f_0, g_{min}) = d(f_0, G). \tag{14}$$

The following formula gives the explicit expression for the computation of the maximum local measure for both cases:

**Theorem 13.** *Let  $x \in \mathbb{R}^2$ , the maximum local measure centered on  $x$  is equal to*

$$|\mathbb{1}_F(x) - \mathbb{1}_G(x)| \max(d(x, F), d(x, G)),$$

where  $\mathbb{1}_F(x)$  is equal to 1 when  $x \in F$  and to 0 elsewhere.

**Proof.** The value of the maximum local measure and the value obtained from the formula are evaluated in different cases, so to show their equality:

- If  $x \in F \cup G$  or if  $x \notin F$  and  $x \notin G$ , the maximum local measure is equal to 0 and as  $|\mathbb{1}_F(x) - \mathbb{1}_G(x)| = 0$ , and thus the formula.
- If  $x$  belongs only to one of the two subsets  $F$  and  $G$ , for example  $F$ , from Definition 4, the local measures are equal to the directed distances from  $F$  to  $G$  and then with Eq. (14), the maximum local measure is equal to  $d(x, G)$ . As  $|\mathbb{1}_F(x) - \mathbb{1}_G(x)| = 1$ , the formula is also equal to  $d(x, G)$ .  $\square$

## 5. Local-distance map

This section deals with the formal and practical definition of the local-distance map (LDMap) which is based on the local distance measure for each point. Given two images, the LDMap is not an image dissimilarity measure. It is a map characterizing the local differences between the two input images. A subsequent classification step on the LDMap is necessary to access the similarity between the two images.

We first present a formal definition and some properties of the LDMap in the general case. Then a short study of the DT allows an evaluation of the complexity of the algorithm. Finally, some quantitative results are given.

### 5.1. Definition

The definition of the LD Map between two sets of points associates the value of the maximum local measure to each point

of  $\mathbb{R}^2$ . The formula obtained in Theorem 13 gives a reduced expression.

**Definition 14** (*Local-distance map (LDMaP)*). Let  $F$  and  $G$  be two non-empty finite sets of points of  $\mathbb{R}^2$ , the LDMaP is defined by

$$\forall x \in \mathbb{R}^2, \quad \text{LDMaP}(x) = |\mathbb{1}_F(x) - \mathbb{1}_G(x)| \times \max(d(x, F), d(x, G)). \quad (15)$$

The following corollary shows that the maximum value in the LDMaP is the HD between the two input sets.

**Corollary 15** (*Maximum value in the LDMaP*). The value  $v = \text{HD}(F, G)$  is reached at least once in the LDMaP:

$$\max(\text{LDMaP}(F, G)) = \text{HD}(F, G) = v. \quad (16)$$

**Proof.** As  $F$  and  $G$  are finite, Definition 1 implies that  $\text{HD}(F, G)$  is reached for two points, for example  $f_0$  and  $g_0$ :  $h(F, G) = d(f_0, g_0)$ . Theorem 13 implies that for  $f_0$ ,  $\text{LDMaP}(f_0) = d(f_0, G) = d(f_0, g_0)$ .  $\square$

## 5.2. The discrete case

The previous definition of the LDMaP is specialized here for digital images. In this case, the output is an image characterizing the local dissimilarities between the input images. Theorem 13 shows that the LDMaP's formula depends mainly on the DT  $x \mapsto d(x, F)$ . A short study of the DT is presented here so as to evaluate the computation complexity of the LDMaP.

### 5.2.1. Distance transform

The DT is well known in computer vision and in pattern recognition. It stands for the information of the distance to the point set of a binary image.

**Definition 16** (*DT*). Let  $A$  be a point set of  $\mathbb{R}^2$  the DT of  $A$  is given by

$$DT_A(x) = d(x, A) \quad \text{for } x \in \mathbb{R}^2. \quad (17)$$

The DT is a step in the HD computation. Most of the studies dedicated to it aim to improve the computation time. Here is a short presentation.

It was introduced by Rosenfeld [30,31], then in Ref. [32], the author gives a DT fast computation for the euclidean distance based on masks. More recently, several methods have been presented to give an exact and fast computation of the euclidean DT in a linear time [33,34]. The case of the chessboard distance has also been studied in Ref. [35], and Brown presents a fast algorithm in the case of the Manhattan distance [36]. Finally, in Ref. [37], the authors show a generalization of the DT to gray-level images. We will focus on the euclidean DT whose computation time is linear.

### 5.2.2. Computational complexity of the LDMaP

Let us consider two  $m \times m$  images. The computation of the LDMaP begins with the computation of the DTs of both of the images. Then three simple operations (a maximum, a difference and a product) are computed on the images. As the distance transform has a linear computation time  $O(m^2)$ , the whole computation complexity is a  $O(m^2)$ . The given complexity is thus only an upper bound. The real complexity depends on the input image contents.

### 5.3. Qualitative results

In order to have a better intuitive comprehension of the proposed LDMaP, some input images and their corresponding LDMaP are presented here. For all the following illustrations of LDMaPs, the darker the pixel, the higher the local distance. The first example is the comparison of simple images with lines (Fig. 3). The input images contain simple patterns (a vertical line, a horizontal one and a square). Let us give some comments about the comparison of the vertical and the horizontal lines: for the pixel where they cross, the value is equal to zero and for the other black pixels, the more distant from a line, the bigger the distance in the LDMaP. The value of the global HD is 11. The comparison between the vertical line and the square results in the same global HD (=11) but the spatial layout shows that this value is reached in numerous pixels (belonging to the square's vertical sides) while it is found only on four occasions in the comparison of the two lines (at their ends).

Fig. 4 illustrates the notion of local dissimilarity. Each image contains two letters. The dissimilarities are quantified: big dissimilarities are represented in *dark* and small ones in *light*. Moreover, they are spatially localized: from the LDMaP, one can see that the bigger dissimilarities are situated on the straight line of the "e" and on the top of the "t", so the corresponding dissimilarities are important. The ones between the left part of the "o" and the bottom of the "t" and between the loop of the "e" and the "c" are light, so they are small.

Fig. 5 offers an elaborate way to save time during the holidays. Indeed the images come from a ten-error game: the second image is a copy of the first one, but with ten differences (the ten errors). They have been photographed from slightly different points of view and straight comparison  $C = |B - A|$  failed to find the ten errors. The LDMaP (image D) highlights most of the errors (we have circled the 10 errors in black to highlight them), except the black stripe on the tracksuit (that has become two black stripes). The reason is the anisotropy of the "error": it is long and not wide. The window stops its growth as soon as it meets points of the stripes and so it is blind to the length of the "error".

### 5.4. Comparison

We present here some results regarding the advantages of the LDMaP with respect to the MHD existing in the literature. Fig. 6 presents two simple object's comparisons that result in the same HD value 11. As shown in Table 1, the comparisons

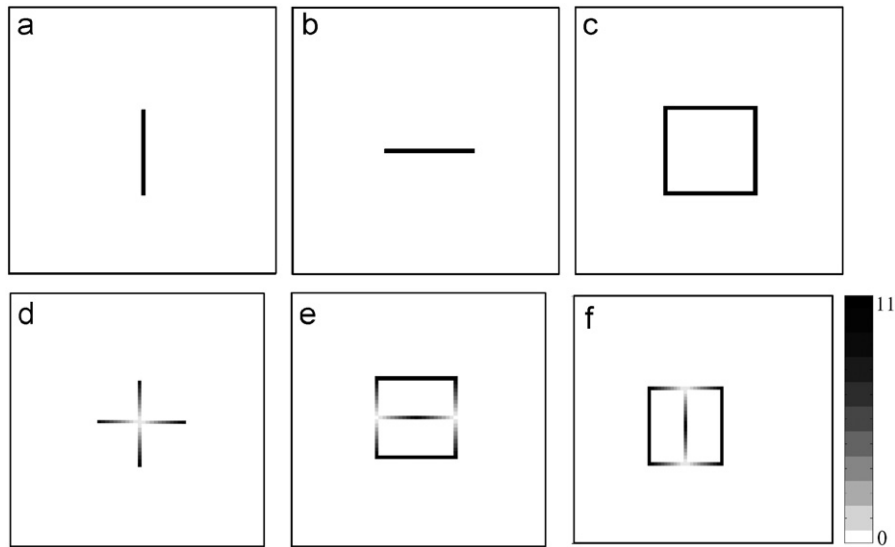


Fig. 3. Behavior of the LDMaP on simple patterns. A vertical line, a horizontal one, a square and their LDMaPs. The darker the pixel, the higher the distance measure.

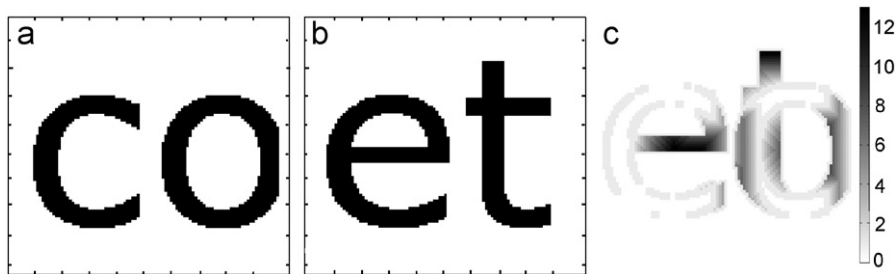


Fig. 4. Letters “co” et “et” and their LDMaP. The obtained LDMaP (c) shows clearly both locations and quantification of the dissimilarities.

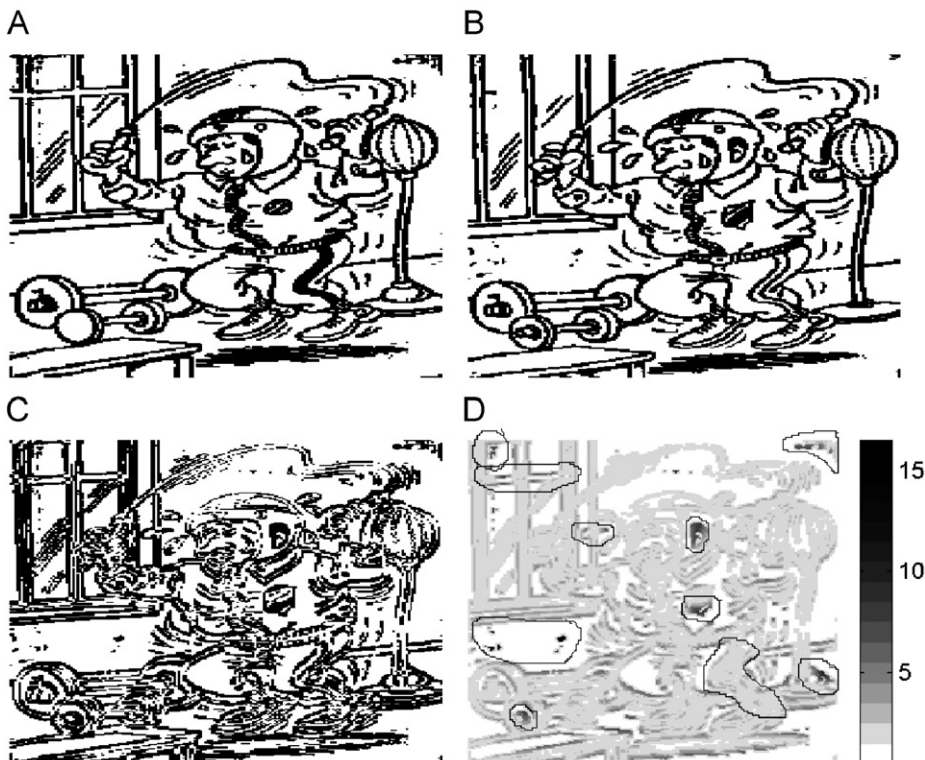


Fig. 5. The ten-error game. The two images to compare (A and B), the absolute difference  $C = |B - A|$  and their LDMaP (D) where we have circled the errors in black.

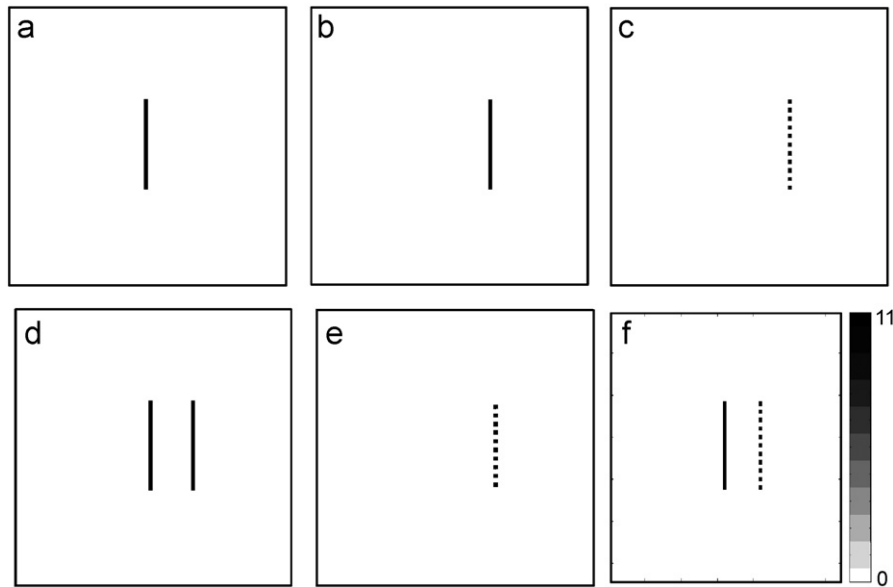


Fig. 6. LDMAP from the comparison of the image (a) (vertical line) with the image (b) (shifted line (a)) and the image (c) (dotted shifted line (a)). None of the presented global measures is precise enough to measure the different similarity degrees. On the contrary, this piece of information is contained in the LDMAP.

Table 1  
Results for several global HD measures of the pair A, D and A, E

Comparison method	HD	Partial HD	MHD	WHD (threshold)
A D	11	11	11	11
A E	11	11	11	11

The results show that the measures cannot account of the different degrees of similarity between the different image pairs.

of the image A with the images C and D give the same value for the PHD, the MHD, the WHD, although in the first case, two plain lines are compared and in the second one, a plain line is compared with a dashed one. The LDMAP accounts clearly for this distinction and enables a potential decision to distinguish these two cases. A more advanced comparison based on an application is given in Section 6.

### 5.5. Limitations

We present in this paragraph some kinds of images for which a LDMAP-based comparison will not give good information to evaluate their similarity.

- Firstly, it concerns images including a large ratio of pattern pixels, e.g. textured images. In this case, the measure values in the DT of the images are low. As the LDMAP is based on the DT, it is also composed of low values of distance measures, so it is difficult to evaluate the image similarity on this basis. This case is illustrated in Fig. 7. For the first comparison, the pattern pixel ratio is not large, and the obtained values (up to 17) reflect the similarity between the images (the image size is  $128 \times 128$ ). For the second comparison,

the binary images are textured and the pattern pixel ration is large. The obtained values (up to 2) do not reflect the similarity in this case.

- Secondly, it involves images with a small ratio of pixel images, e.g. segmented images. The reason is that the non-zero distance values are located on the pattern pixels of one of the images being compared. As a result, there are few non-zero values in the LDMAP, which can make the similarity evaluation difficult.

## 6. Application: CBIR

In order to show the interest of the proposed distance map with quantitative results, it has to be included in a global classification process. The aim of this section is to present the results of the comparison of the local-dissimilarity map (LDMAP) to other methods.

### 6.1. The database

The Troyes municipal library, which is our collaborator, has provided digitalized medieval illustrations for the test database [38]. These images, originally printed in books, have strong contrast which allows to binarize them with little loss. This database is composed of 68 images, some of them illustrating the same scene. The objective is to retrieve illustrations representing the same scene.

One of the difficulties comes from the numerous classes in the database: a class of images illustrating the same scene include between one and four images. So there are about 30 classes of similar images for the 68 images (some classes contain only one image). The number of classes makes a straight

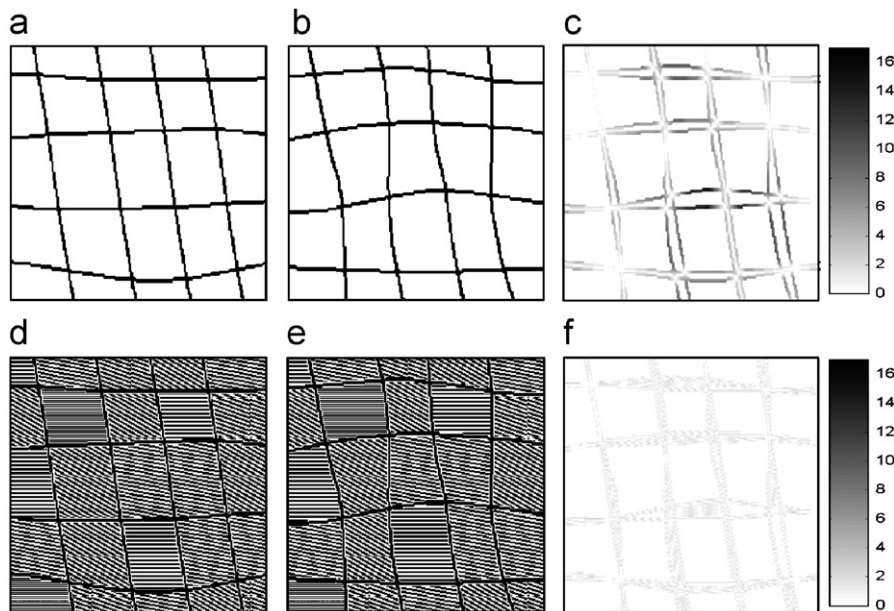


Fig. 7. Two images (a) and (b), their textured versions (d) and (e) and their LDMaPs. The values of the LDMaP (f) are lower than the LDMaP (c) because of the large pattern pixel ratio in the textured images (d) and (e).

comparison on the images complicated and in addition, the choice of the class number is difficult as far as new images can be introduced and need the creation of a new class. Unlike the images, the LDMaPs are classified into two classes: the ones obtained by the comparison of similar images  $C_{sim}$  and the other ones by the comparison of two dissimilar images  $C_{dissim}$ . The introduction of new images does not change the LDMaP class number.

The comparison of the 68 images results in 2278 LDMaPs, 125 of which are classified in  $C_{sim}$  and 2153 in  $C_{dissim}$  thanks to a manual expert comparison of the impressions. Examples of medieval impressions and their LDMaPs are given Fig. 8. Impressions 1 and 2 come from very similar wooden stamps and impression 3 illustrates the same scene with differences in the way of illustrating the grass and the helmets. Even when the global HD does not reflect the similarity degree ( $HD(imp1, imp2) = 25$ ,  $HD(imp1, imp3) = 15$ ),  $LDMaP(imp1, imp3)$  is locally darker (so with higher values) than  $LDMaP(imp1, imp2)$  where the illustration differs. The impression 4 illustrates a distinct scene and high values (comparing to those of the other LDMaPs) can be found all over  $LDMaP(imp1, imp2)$ .

## 6.2. The global classification process

It is composed of two stages: firstly the construction of a LDMaP between two images, secondly a classification using a Support Vector Machines (SVM) based on the obtained distance map. During the acquisition, the medieval impressions have been registered and binarized. The chosen size is  $64 \times 64$  pixels to save computation time. Instead of a comparison of two image signatures, the distance map allows a direct dissimilarity

comparison including spatial information on the dissimilarities. This information is exploited in the classification step.

### 6.2.1. Classification

LDMaPs are constructed by the comparison process. The set of the LDMaPs can be then classified in two classes:  $C_{sim}$  class including maps from similar images and  $C_{dissim}$  class including maps from dissimilar images.

A SVM method is used to classify the LDMaPs into these two classes. We will now briefly review of the SVM. SVM method is a classification and regression method introduced by Boser et al. [39], for a complete description see Ref. [40]. SVMs are particularly efficient for supervised classification because they can handle problems depending on numerous descriptors and have been successfully applied on large dimension real problems. For example, in pattern recognition, SVMs have been used for writing isolated figure recognition [41], object recognition [42], face detection in images [43] and text categorization [44].

In the frame of the LDMaP comparison, a study on the different SVM methods has been carried out [45] and shows that the most efficient choice to deal with our data was the classical SVM (C-SVM) with a linear kernel.

## 6.3. Experiments

Our objective is to test the method's efficiency in assessing local dissimilarities. The experiments were carried out by the following way: first a supervised machine learning is made on a set of 50 LDMaPs in  $C_{sim}$  and 50 in  $C_{dissim}$ . Then, the test is done on a distinct set of 75 LDMaPs of  $C_{sim}$  and 200 of  $C_{dissim}$ . The choice of the sets in each class is randomized. Secondly, Five comparison methods are carried out. Finally, the results obtained by the five comparison methods with the five

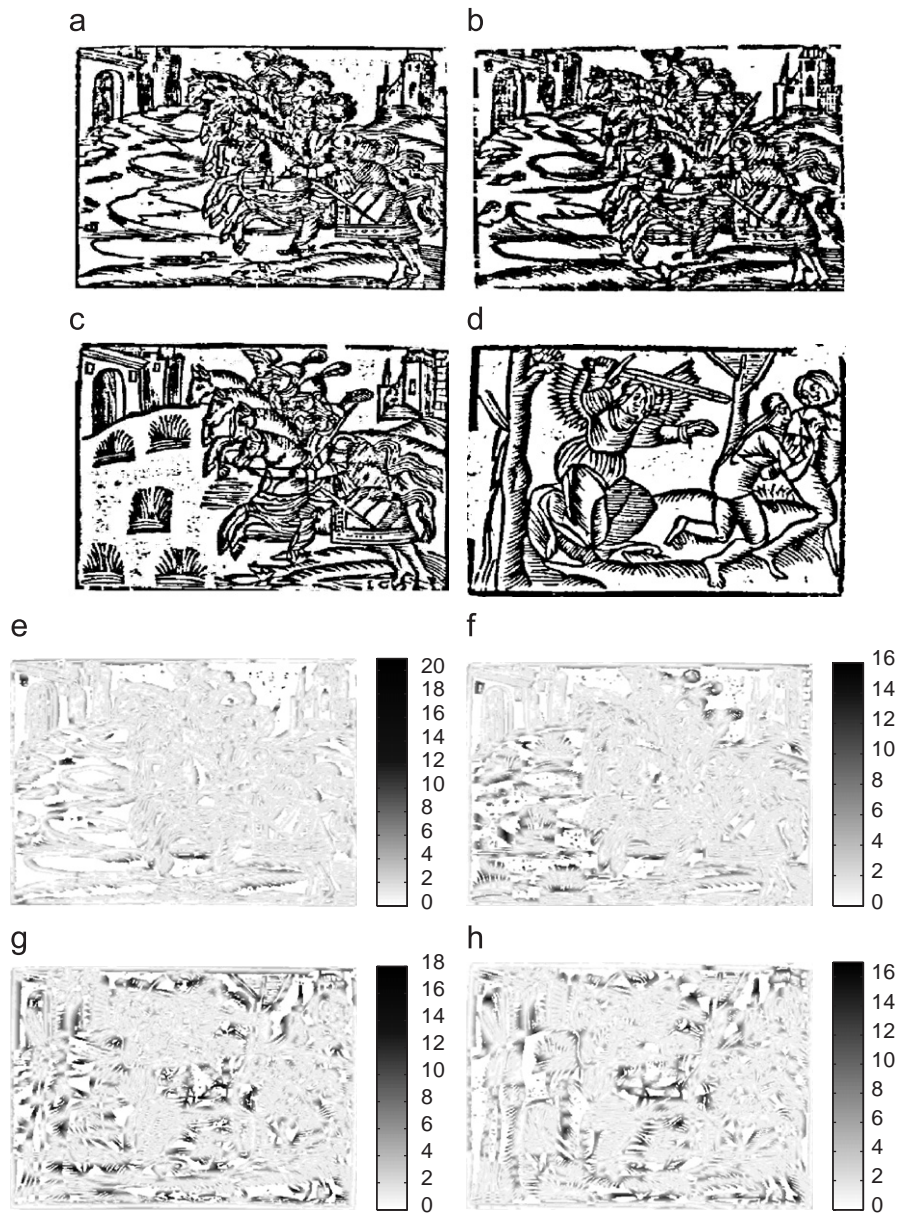


Fig. 8. Medieval impressions and their LDMs. Here are four medieval impressions. Imp. 1, Imp. 2 and Imp. 3 illustrate the same scene with a different kind of grass and helmets in Imp. 3. Imp. 4 illustrates a distinct scene.

comparison methods are compared with the ones obtained manually. The five classification methods are the following ones:

- our method based on the LDMap,
- the so-called Local Simple Difference Map (LSDMap) using the distance map, but with the simple difference locally instead of the HD:  $HSD_W(F, G) = |F \cap W - G \cap W|$ ,
- the global HD,
- the PHD,
- the MHD.

### 6.3.1. Test methods

A quantitative result is then obtained for each comparison method thanks to a decision step. The decision step is different

whether the measure result is an image (case of the LDMap and the LSDMap) or a real value (case of the HD and its variations). In the first case, the classification method is a SVM described in Section 6.2. In the second case, an empirical distribution for each class  $C_{sim}$  and  $C_{dissim}$  is computed from the learning set. As the modes of the empirical distributions are quite well defined, an easy and efficient classification method is the maximum likelihood method.

### 6.3.2. Results

Results are summarized in Table 2. They show the efficiency of the LDMap both concerning spatial information (comparison with the global HD and the PHD) and the ability of the local HD to catch the local dissimilarities (comparison with

Table 2

Results for  $D_{H,W}$ , the LDMMap used with the absolute difference, the global HD, the Partial HD (PHD) and for Modified HD (MHD)

Successful retrieval	LDMMap (%)	LSDMap (%)	HD (%)	PHD (%)	MHD (%)
Found in $C_{sim}$	98	90	60	83	77
Found in $C_{dissim}$	97	92	75	81	83

The PHD depends on a parameter  $p$  and the detailed results for the PHD are presented in Fig. 9.

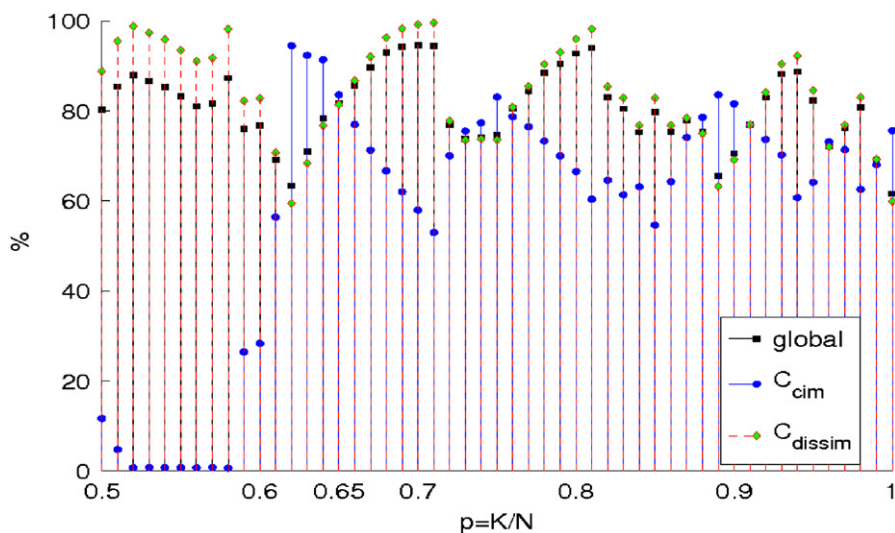


Fig. 9. Graphs of the successful rate for the partial HD (PHD) function of the parameter  $p = K/N$  (see formula Section 1.2.1). The graphs stand for the rates for the two classes  $C_{sim}$  and  $C_{dissim}$  and for all the LDMMaps (global rate).

the LSDMap). As the PHD depends on a parameter, only the results with the optimal parameter are presented in Table 2. The detailed results for the PHD are shown Fig. 9. This figure highlights the difficulty of the parameter choice, even with an a priori study. Indeed, this choice has to be precise because there is an efficiency difference up to 25% for two consecutive steps and it is difficult because there are several local maxima. The aim is to retrieve similar images, so  $C_{sim}$  successful rate is the most important and after the  $C_{dissim}$  rate. As  $C_{dissim}$  is really bigger than  $C_{sim}$ , the global rate is not very interesting. Thus, the best values are for  $p = 0.65$ : 83% for  $C_{sim}$  and 81% for  $C_{dissim}$ .

### 6.3.3. Comments

The results both for the HD and the PHD show that dissimilarity global information is less discriminant than a map of local information. The comparison between the LDMMap and the LSDMap shows that the local HD is better to measure local dissimilarity than the simple difference. This is illustrated on Fig. 10.

## 6.4. Robustness

### 6.4.1. Noise related to medieval impressions degradation

Robustness related to the application database is evaluated: The kind of noise related to medieval impressions can be

ink-stain noise and ink-erasing noise in order to simulate potential degradation that could damage illustrations. The robustness to these kinds of noise is tested in the following way:

- The learning stage is done on LDMMap from images without noise.
- The test stage is done on LDMMap from one unnoisy image and one noisy image.

This method enables to evaluate to which extent a stained or erased image is successfully compared to unnoisy images.

The successful retrieval rates for both classes  $C_{sim}$  and  $C_{dissim}$  have been measured for square ink stains and erasings, with an increasing square size. The results are presented Fig. 11. We can observe that the method is more robust to stains than to erasings. One reason is that the treated information comes from the black pixels. High values appear because of the erased square in the LDMMap and this leads the classifier to class the LDMMap as dissimilar.

### 6.4.2. Robustness relative to translation

In this paragraph, the robustness to small horizontal translation is evaluated. A small translation means a defect of registration between the images. This defect can occur during the digitalization of the impressions or after the pre-processing.

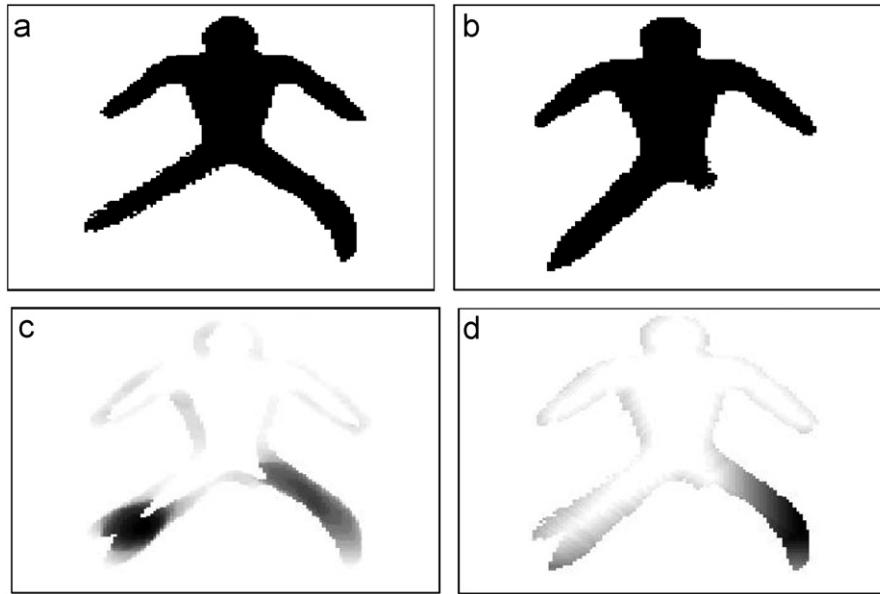


Fig. 10. Illustration of the LDMAP efficiency to measure the local dissimilarity. For the measured LSDMap between (a) and (b) (see (c)), it is not logical that the dissimilarities on the left leg are higher than those of the right leg. The LDMAP measured between (a) and (b) (see (d)) is close to the visual intuition with higher dissimilarities on the right leg than those of the left leg.

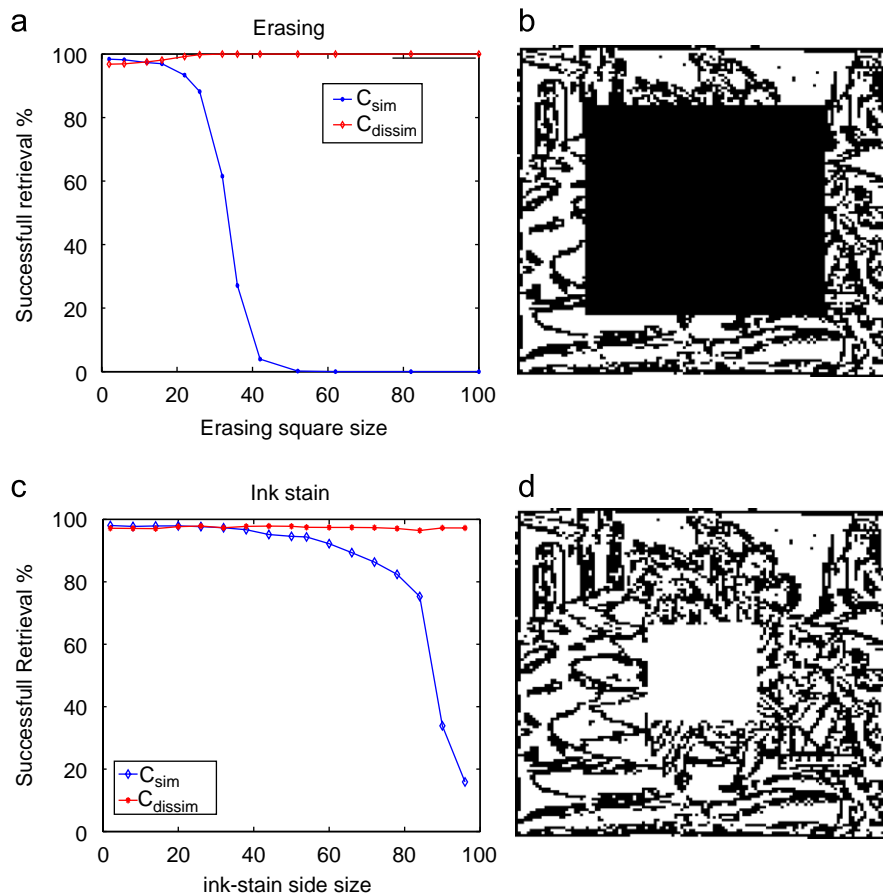


Fig. 11. Robustness to ink stains and erasings: the result is really different whether it is a stain or an erasing. One reason is that treated information is the one of black pixels.

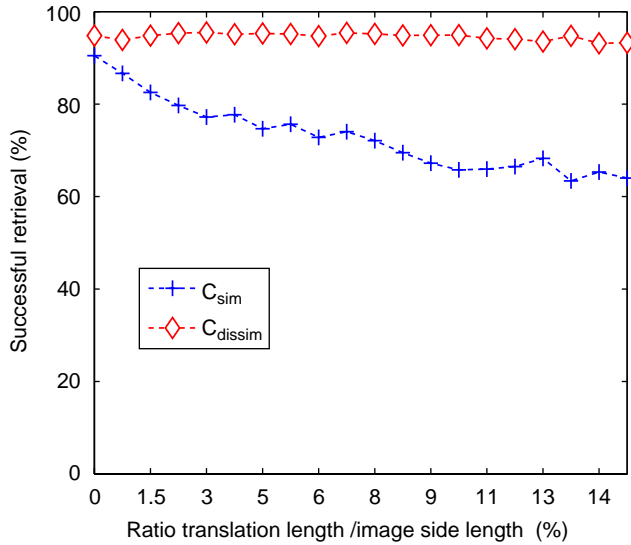


Fig. 12. Robustness to small translations. The successful retrieval rate is constant for  $C_{dissim}$  and decreases slowly for  $C_{sim}$ .

The robustness to translation is tested in the following way:

- The learning stage is done on LDMMap from images without translation.
- The test stage is done on LDMMap from one image without translation and one translated image.

This protocol allows to evaluate the extent to which a translated image is successfully compared to untranslated images. The successful retrieval rates for both classes  $C_{sim}$  and  $C_{dissim}$  have been measured with an increasing translation. The results are presented Fig. 12. The classification decreases slowly when the translation increases. It is normal since all the measures in the LDMMap increase. The property 2 shows indeed that the HD measure is proportional to the translation length. The LDMMap inherits of this property for small translations. For bigger translations, the measures in the LDMMap are no more local, and the successful retrieval rates are low. Nevertheless bigger translations do not correspond to registration errors, so it has not to be studied here.

## 7. Conclusion

This paper proposes a method to make an adaptive measure of the local dissimilarities between two binary images. With this end, a local HD is defined and its properties of boundary and growth are proven. This allows making the HD measure automatically fit the local image dissimilarity. The LDMMap contains the local distances and their spatial layout. These information may be exploited to compare images that are difficult to separate with a global similarity measures because of the presence of different contents. A supervised classification based on SVM is applied to the LDMMaps to decide the similarity of the compared images. As an application, the proposed method has been tested on an medieval wooden stamps base. The compar-

ison with global measures shows that the dissimilarity spatial distribution really improves the results of classification. The comparison with the simple-difference map shows that the HD is more efficient to catch local dissimilarities. Finally, the study of the global-process robustness carried out to show that the robustness to ink stains is good and the robustness to erasings is quite good. Further works will explore the gray-level case and study the use of the LDMMap for assessment of non-linear multiresolution properties.

## Appendix A.

### A.1. Proof of Proposition 7

**Proof.** Suppose  $HD_W(F, G) > HD(F, G)$ . Without loss of generality, suppose that  $HD_W(F, G) = h_W(F, G)$ . By hypothesis  $h_W(F, G) > 0$ , so from expression (4), there are two cases:

- (1) (If  $F \cap W \neq \emptyset, G \cap W \neq \emptyset$ ) then

$$h_W(F, G) = \max_{f \in F \cap W} \left[ \min \left( \min_{g \in G \cap W} d(f, g), \min_{w \in Fr(W)} d(f, w) \right) \right], \quad (\text{A.1})$$

so

$$\exists f_0 \in F \cap W / h_W(F, G) = \min \left( \min_{g \in G \cap W} d(f_0, g), \min_{w \in Fr(W)} d(f_0, w) \right). \quad (\text{A.2})$$

There are two cases:

- (a) The minimum is reached by the first member in parenthesis:

$$h_W(F, G) = \min_{g \in G \cap W} d(f_0, g),$$

then

$$\exists g_0 \in G \cap W / h_W(F, G) = d(f_0, g_0). \quad (\text{A.3})$$

Yet, for the point  $f_0$ , there are two pieces of information:

- the distance from this point to the edge of  $W$  is higher than  $d(f_0, g_0)$ , because the minimum is reached by the first member in parenthesis.
- As  $W$  is a ball, from the previous point, one has

$$B(f_0, d(f_0, g_0)) \subset W \quad (\text{A.4})$$

and by definition of  $g_0$ , there is no other point of  $G$  in  $B(f_0, d(f_0, g_0))$  than  $g_0$ , so

$$\min_{g \in G} d(f_0, g) = d(f_0, g_0),$$

i.e.  $d(f_0, g_0)$  is one of the minima, so

$$d(f_0, g_0) \leq \max_{f \in F} \left( \min_{g \in G} d(f, g) \right).$$

(a) Now one can deduce in this case (a) that

$$HD_W(F, G) = d(f_0, g_0),$$

so

$$HD_W(F, G) \leq HD(F, G),$$

which is in contradiction with the hypothesis.

(b) Let see the case the minimum in Eq. (A.1) is reached by the second member in parenthesis:

$$h_W(F, G) = \min_{w \in Fr(W)} d(f_0, w),$$

note  $r_0 = \min_{w \in Fr(W)} d(f_0, w)$ , then  $B(f_0, r_0) \subset W$ . Yet the minimum (of Eq. (A.1)) is not reached by a point of  $G$ , so  $G \cap B(f_0, r_0) = \emptyset$ . then

$$r_0 \leq \min_{g \in G} (f_0, g),$$

and so

$$r_0 \leq \max_{f \in F} \left( \min_{g \in G} d(f, g) \right), \quad (A.5)$$

yet, by definition

$$\begin{aligned} r_0 &= \min_{w \in Fr(W)} d(f_0, w) \\ &= h_W(F, G) \\ &= HD_W(F, G). \end{aligned}$$

Replace it in (A.5),

$$HD_W(F, G) \leq \max_{f \in F} \left( \min_{g \in G} d(f, g) \right), \quad (A.6)$$

What's more

$$\begin{aligned} \max_{f \in F} \left( \min_{g \in G} d(f, g) \right) &= h_W(F, G) \\ &\leq HD(F, G). \end{aligned}$$

Replace it in (A.6),

$$HD_W(F, G) \leq HD(F, G),$$

which is contradictory with the hypothesis.

(2) (If  $F \neq \emptyset, G = \emptyset$ ) then, from (4),

$$h_W(F, G) = \max_{f \in F \cap W} \left( \min_{w \in Fr(W)} d(f, w) \right),$$

So

$$\exists f_0 \in F \cap W / h_W(F, G) = \min_{w \in Fr(W)} d(f_0, w). \quad (A.7)$$

The situation is the same as in the case (b) treated above.  $\square$

## A.2. Proof of Proposition 8

**Proof.** The difference with the previous property is the distance to the edges in  $HD_W$ . First point:  $V \subset W$  implies that for all point  $v \in V$ ,

$$d(v, Fr(V)) \leq d(v, Fr(W)). \quad (A.8)$$

So

$$h_V(F, G) = \max_{f \in F \cap V} \min \left( \min_{g \in G \cap V} (d(a, g)), \min_{v \in Fr(V)} d(a, v) \right) \quad (A.9)$$

$$\leq \max_{f \in F \cap V} \min \left( \min_{g \in G \cap V} d(f, g), \min_{w \in Fr(W)} d(f, w) \right). \quad (A.10)$$

The demonstration follows the same way as the previous one: suppose  $HD_V(F, G) > HD_W(F, G)$  then without loss of generality,  $HD_V(F, G) = h_V(F, G)$ . By hypothesis  $h_W(F, G) > 0$ , so from expression (4),

$$\begin{aligned} \exists a_0 \in F \cap V / h_V(F, G) &= \begin{cases} \min \left( \min_{b \in G \cap V} d(a_0, b), \min_{v \in Fr(V)} d(a_0, v) \right) \\ \text{if } F \cap V \neq \emptyset, \quad G \cap V \neq \emptyset \\ \min_{v \in Fr(V)} d(a_0, v) \quad \text{if } F \cap V \neq \emptyset, \quad G \cap V = \emptyset. \end{cases} \end{aligned} \quad (A.11)$$

As in the previous proof, we treat the first case, the second one is then deduced from it.

(1) (If  $F \cap V \neq \emptyset, G \cap V \neq \emptyset$ ) two cases can be distinguished

(a) The minimum is reached by the first member in parenthesis:

$$h_V(F, G) = \min_{b \in G \cap V} d(a_0, b),$$

then

$$\exists b_0 \in G \cap V / h_V(F, G) = d(a_0, b_0). \quad (A.12)$$

Then two points regarding point  $a_0$ :

- As the minimum is reached by the first member in parenthesis, it is inferior to the second member, and so

$$\min_{b \in G \cap V} d(a_0, b) \leq \min_{v \in Fr(V)} d(a_0, v),$$

then, with (A.12),

$$d(a_0, b_0) \leq \min_{v \in Fr(V)} d(a_0, v). \quad (A.13)$$

- As a consequence of this first point:

$$B(a_0, d(a_0, b_0)) \subset V,$$

$$\forall b \in (B(a_0, d(a_0, b_0)) \cap G), \quad d(a_0, b) = d(a_0, b_0). \quad (A.14)$$

(a) So

$$\min_{b \in G \cap W} d(a_0, b) = d(a_0, b_0).$$

Let's prove now that

$$d(a_0, b_0) \leq \min_{w \in Fr(W)} d(a_0, w).$$

As  $a_0 \in V$  et  $V \subset W$ , one has

$$\min_{v \in Fr(V)} d(a_0, v) \leq \min_{w \in Fr(W)} d(a_0, w),$$

from Eq. (A.13), one has  $d(a_0, b_0) \leq \min_{v \in Fr(V)} d(a_0, v)$   
so

$$d(a_0, b_0) \leq \min_{w \in Fr(W)} d(a_0, w). \quad (\text{A.15})$$

What is more, from Eq. (A.14), one has

$$\forall b \in G \cap B(a_0, d(a_0, b_0)), \quad d(a_0, b) = d(a_0, b_0),$$

so

$$\forall b \in G \cap W, \quad d(a_0, b) \geq d(a_0, b_0),$$

i.e.,

$$d(a_0, b_0) \leq \min_{b \in B \cap W} d(a_0, b). \quad (\text{A.16})$$

From Eqs. (A.14) and (A.16), one has

$$d(a_0, b_0) \leq \min \left( \min_{b \in G \cap W} d(a_0, b), \min_{w \in Fr(W)} d(a_0, w) \right). \quad (\text{A.17})$$

By definition,

$$h_W(F, G) = \max_{a \in F \cap W} \left( \min \left( \min_{b \in G \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right),$$

and the maximum for all the points of  $F$  is greater than the value for the point  $a_0$ , so, with Eq. (A.17)

$$\max_{a \in F \cap W} \left( \min \left( \min_{b \in G \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right) \geq d(a_0, b_0),$$

so

$$h_W(F, G) \geq d(a_0, b_0).$$

What is in contradiction with hypothesis (A.8).

(b) The minimum of Eq. (A.11) is reached by the second term in parenthesis:

$$h_V(F, G) = \min_{v \in Fr(V)} d(a_0, v).$$

Note that  $r_0 = \min_{v \in Fr(V)} d(a_0, v)$ , as  $V$  is a ball by hypothesis, one has

$$B(a_0, r_0) \subset V.$$

What is more, as the minimum is reached by the second member of Eq. (A.11),

$$\forall b \in G \cap B(a_0, r_0), \quad d(a_0, b) \geq r_0,$$

so

$$\forall b \in G, \quad d(a_0, b) \geq r_0,$$

then

$$\forall b \in G \cap W, \quad d(a_0, b) \geq r_0,$$

and so

$$\min_{b \in G \cap W} d(a_0, b) \geq r_0. \quad (\text{A.18})$$

Yet, as  $a_0 \in V$  and  $V \subset W$ , then

$$\forall v \in Fr(V), \quad \forall w \in Fr(W), \quad d(a_0, v) \leq d(a_0, w)$$

so

$$r_0 = \min_{v \in Fr(V)} d(a_0, v) \leq \min_{w \in Fr(W)} d(a_0, w), \quad (\text{A.19})$$

with , one has:

$$\min \left( \min_{b \in G \cap W} d(a_0, b), \min_{w \in Fr(W)} d(a_0, w) \right) \geq r_0,$$

i.e.,

$$h_W(F, G) \geq r_0,$$

or  $r_0 = h_V(F, G)$  so

$$h_W(F, G) \geq h_V(F, G).$$

What is in contradiction with hypothesis (A.8).

(2) (if  $F \cap V \neq \emptyset$ ,  $G \cap V = \emptyset$ ) then one has

$$h_V(F, G) = \min_{v \in Fr(V)} d(a_0, v), \quad (\text{A.20})$$

and it comes down to the case (b).

## References

- [1] A.W.M. Smeulders, M. Worring, M. Santini, S. Gupta, R. Jain, Content based image retrieval at the end of the early years, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (2000) 1349–1380.
- [2] E. Baudrier, G. Millon, F. Nicolier, S. Ruan, A fast binary-image comparison method with local-dissimilarity quantification, in: *Proceedings of International Conference on Pattern Recognition (ICPR)*, vol. 3, IEEE, Hong-Kong, 2006, pp. 216–219.
- [3] R.C. Veltkamp, M. Hagedoorn, Shape similarity measures, properties and constructions, Technical Report UU-CS-2000-37, Utrecht University, October, 2000. (<http://ftp.cs.uu.nl/pub/RUU/CS/techreps/CS-2000/2000-37.pdf>).
- [4] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) (April 2004).
- [5] S. Antani, R. Kasturi, R. Jain, A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video, *Pattern Recognition* 35 (4) (2002) 945–965.

- [6] A.D. Lecce, V. Gerriero, An evaluation of the effectiveness of image features for image retrieval, *J. Visual Commun. Image Represent.* 10 (4) (1999) 351–362.
- [7] N.-S. Chan, K.-S. Fu, Query-by-pictorial-example, *IEEE Trans. Software Eng.* 6 (6) (1980) 519–524.
- [8] G. Sheikholeslami, W. Chang, A. Zhang, Semquery: semantic clustering and querying on heterogeneous features for visual data, *IEEE Trans. Knowl. Data Eng.* 14 (5) (2002) 988–1002.
- [9] A. Soffer, H. Samet, Pictorial queries by image similarity, in: *Proceedings of the 13th International Conference on Pattern Recognition*, vol. III, Vienna, Austria, 1996, pp. 114–119.
- [10] K. Matkovic, T. Psik, I. Wagner, W. Purgathofer, Tangible image query, in: *Proceedings of Smart Graphics*, Banff, Canada, 2004, pp. 31–42.
- [11] P. Agouris, J. Carswell, A. Stefanidis, A feature library approach to on-line image querying and retrieval for topographic applications, in: *Vision Interface '99*, Trois-Rivières, Canada, 1999.
- [12] R. Brunelli, O. Mich, Histograms analysis for image retrieval, *Pattern Recognition* 34 (8) (2001) 1625–1637.
- [13] J. Dombre, N. Richard, C. Fernandez-Maloigne, Use of spatial information for content-based image retrieval, in: *Color in Graphics, Image and Vision (CGIV)*, Poitiers, France, 2002, pp. 384–389.
- [14] R. Veltkamp, M. Hagedoorn, State-of-the-art in shape matching, Technical Report UU-CS-1999-27, Utrecht University, The Netherlands, 1999.
- [15] R. Klette, P. Zamperoni, Measures of correspondence between binary patterns, *Image Vision Comput.* 5 (4) (1987) 287–295.
- [16] C. Zhao, W. Shi, Y. Deng, A new Hausdorff distance for image matching, *Pattern Recognition Lett.* 26 (2004) 581–586.
- [17] D.P. Huttenlocher, D. Klanderman, W.J. Rucklidge, Comparing images using the Hausdorff distance, *Int. Conf. on Computer Vision and Pattern Recognition*, 1993, pp. 705–706.
- [18] D.P. Huttenlocher, W.J. Rucklidge, A multi-resolution technique for comparing images using the Hausdorff distance, *IEEE Trans. Pattern Anal. Mach. Intell.* 15 (9) (1993) 850–863.
- [19] O.-K. Kwon, D.-G. Sim, R.-H. Park, Robust hausdorff distance matching algorithms using pyramidal structures, *Pattern Recognition* 34 (10) (2001) 2005–2013.
- [20] B. Takács, Comparing faces using the modified Hausdorff distance, *Pattern Recognition* 31 (12) (1998) 1873–1881.
- [21] O. Jesorsky, K.J. Kirchberg, R.W. Frischholz, Robust face detection using the Hausdorff distance, in: J. Bigun, F. Smeraldi (Eds.), *Audio- and Video-Based Person Authentication—AVBPA 2001*, Lecture Notes in Computer Science, vol. 2091, Springer, Halmstad, Sweden, 2001, pp. 90–95.
- [22] A. Barla, F. Odone, A. Verri, Hausdorff kernel for 3d object acquisition and detection, in: R. Fu (Ed.), *ECCV '02: Proceedings of the Seventh European Conference on Computer Vision-Part IV*, Springer, London, UK, 2002, pp. 20–33.
- [23] G. Borgefors, Distance transformations in digital images, *Comput. Vision Graph. Image Process.* 34 (3) (1986) 344–371.
- [24] J. Paumard, Robust comparison of binary images, *Pattern Recognition Lett.* 18 (10) (1997) 1057–1063.
- [25] M.-P. Dubuisson, A.K. Jain, A modified Hausdorff distance for object matching, in: *Proceedings of the 12th IAPR International Conference on Pattern Recognition 1994*, pp. 566–568.
- [26] Y. Lu, C. Tan, W. Huang, L. Fan, An approach to word image matching based on weighted Hausdorff distance., in: *Proceedings of the Sixth International Conference on Document Analysis and Recognition*, 2001, pp. 921–925.
- [27] Y. Lu, C. Tan, Word spotting Chinese document images without layout analysis, in: *Proceeding of 16th International Conference on Pattern Recognition*, vol. 3, 2002, pp. 57–60.
- [28] B. Guo, K. Lam, W. Siu, S. Yang, Human face recognition using a spatially weighted Hausdorff distance, in: *The 2001 IEEE International Symposium on Circuits and Systems*, 2001, pp. 145–148.
- [29] D.-G. Sim, O.-K. Kwon, R.-H. Park, Object matching algorithms using robust Hausdorff distance measures, *IEEE Trans. Image Process.* 8 (3) (1999) 425–429.
- [30] A. Rosenfeld, J.L. Pfalz, Sequential operations in digital picture processing, *J. ACM* 13 (1966) 471–494.
- [31] A. Rosenfeld, A.C. Kak, *Digital picture processing*, Academic Press, New York, 1982.
- [32] G. Borgefors, A new distance transformation approximating the euclidean distance, in: *Proceedings of the International Joint Conference on Pattern Recognition*, 1986, pp. 336–338.
- [33] C. Huang, O. Mitchell, An euclidean distance transform using grayscale morphology decomposition, *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (4) (1994) 443–448.
- [34] H. Breu, J. Gil, D. Kirkpatrick, M. Werman, Linear time euclidean distance transform algorithms, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (5) (1995) 529–533.
- [35] Z.M. Kovács, R. Guerrieri, Computer recognition of handwritten characters using the distance transform, *Electron. Lett.* 28 (19) (1992) 1825–1827.
- [36] R. Brown, The fringe distance measure, *IEEE Trans. Systems Man and Cybernet.* 24 (1) (1994).
- [37] J. Arlandis, J. Perez-Cortes, The continuous distance transformation: a generalization of the distance transformation for continuous-valued images, in: A.M.I. Torres (Ed.), *Pattern Recognition and Applications*, *Frontiers in Artificial Intelligence and Applications*, vol. 56, IOS Press, 2000, pp. 89–98.
- [38] R. Seulin, O. Morel, G. Millon, F. Nicolier, Range image binarization: application to wooden stamps analysis, in: *International Conference on Quality Control by Artificial Vision (QCAV)—IEEE*, vol. 5132, SPIE, Gatlinburg, TN, USA, 2003, pp. 252–258.
- [39] B.E. Boser, I.M. Guyon, V.N. Vapnik, A training algorithm for optimal margin classifiers, in: *COLT '92: Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, ACM Press, New York, USA, 1992, pp. 144–152.
- [40] V. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.
- [41] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (3) (1995) 273–297.
- [42] V. Blanz, B. Schölkopf, H. H. Bülthoff, C. Burges, V. Vapnik, T. Vetter, Comparison of view-based object recognition algorithms using realistic 3d models, in: *ICANN 96: Proceedings of the 1996 International Conference on Artificial Neural Networks*, Springer, London, UK, 1996, pp. 251–256.
- [43] E. Osuna, R. Freund, F. Girosi, Training support vector machines: an application to face detection, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 1997, pp. 130–136.
- [44] T. Joachims, Text categorization with support vector machines: Learning with many relevant features, in: *ECML '98: Proceedings of the 10th European Conference on Machine Learning*, Springer, London, UK, 1998, pp. 137–142.
- [45] D. Ait Aouit, Classification d'images par la méthode des support vector machines: étude et applications, Master's Thesis, University of Reims-Champagne-Ardenne, Troyes, September 2004.

**About the Author**—ÉTIENNE BAUDRIER received a master degree from the University of Paris Sud in Pure Mathematics and the PhD degree in Image Processing from the University of Reims. He is now on a post-doctoral position in the University of La Rochelle and his research interests are the image comparison, the image classification and the application of geometrical algebra in image processing.

**About the Author**—GILLES MILLON received the PhD degree from the Reims Champagne Ardenne University (URCA, FRANCE) in 1999 working on the implementation of real time image processing algorithm on FPGA. He particularly studied the run time reconfiguration concept. He's actually an Assistant Professor in the URCA's CRESTIC laboratory. His current research interests are image segmentation, pattern recognition and classification by developing content-based image comparison algorithm.

**About the Author**—FRÉDÉRIC NICOLIER received a master degree in applied physics from the University of Burgundy in 1995 and the PhD degree in Image Processing from the same university in 2000. He is now Assistant Professor at the CReSTIC laboratory, IUT Troyes (URCA). His research interest includes wavelets transform, shape recognition and structural information processing.

**About the Author**—SU RUAN received the Ph.D. degree in Image Processing from the University of Rennes 1 in 1993. She was Assisted Professor at the University of Caen from 1993 to 2003. She is now Professor at the University of Reims, and works in the CReSTIC laboratory for the research. Her research is mainly in the fields of Image Processing and pattern recognition.