ELSEVIER

Pattern Recognition Letters 22 (2001) 1483–1502

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Application of Baddeley's distance to dissimilarity measurement between gray scale images

D. Coquin *, Ph. Bolon

Laboratoire d'Automatique et de Micro-Informatique Industrielle, LAMII/CESALP, Université de Savoie, CNRS-GdR-ISIS, B.P. 806 – 41, avenue de la Plaine, 74016 Annecy cedex, France

Received 6 April 2000; received in revised form 18 June 2001

Abstract

In this paper, we introduce a dissimilarity measure between two gray-scale images based on Baddeley's distance. To some extent, it can be regarded as a modification of that proposed by Wilson et al. [Internat. J. Comput. Vision 24 (1) (1997) 5–18]. Images are represented by surfaces in a 3D space, instead of their subgraphs. Distance calculations are performed by means of a 3D local distance operator adapted to parallelepipedic grids. No truncation effect is introduced. Properties of the new dissimilarity operator are compared to those of the Wilson–Baddeley–Owen operator in terms of sensitivity to gray level variations, spatial shifts and shape distortions. Compared with the Wilson–Baddeley–Owen operator, a more linear behavior is observed. A simplified phenomenological model is proposed in order to explain this behavior.

Keywords: Dissimilarity measure; Gray-scale image comparison; Objective metric

1. Introduction

A problem of both theoretical and practical importance in image processing is to compare two gray-scale images of equal sizes and gray value ranges, taking into account the eventual diversities due to translation, or gray level variations. A quantitative dissimilarity measure $\mathscr{D}(A,B)$ between two images A and B is desirable for a large number of practical applications (filtering, restoration and compression assessment, segmentation comparison). Thus, to estimate dissimilarity between images is a significant problem in image analysis.

Several distances have been proposed for measuring the dissimilarities between objects in binary images (e.g. Baddeley, 1992; Huttenlocher et al., 1993; Dubuisson and Jain, 1994). These dissimilarities are average measures and give a global value which corresponds to the average dissimilarity between two images.

Other objective criteria have been proposed for measuring dissimilarities between objects in gray-scale images (e.g. Coquin et al., 1995, 1997; Zamperoni and Starovoitov, 1996; Wilson et al., 1997; Di Gesù and Starovoitov, 1999), or for image retrieval purposes (Jacobs and Weinshall, 2000). Zamperoni and Starovoitov (1996) have proposed a multi-stage dissimilarity measure, in which each stage (pixel-to-pixel, pixel-to-window, window-to-window, and image-to-image) can be

* Corresponding author. Fax: +33-450-666-020.
E-mail address: didier.coquin@univ-savoie.fr (D. Coquin).

0167-8655/01/$ - see front matter
PII: S0167-8655(01)00104-0

based upon different distance measures, thus originating several variants of dissimilarities. They have shown the interest of choosing a distance computation based on a combined spatial and gray value match.

Wilson et al. (1997) have proposed an extension to gray scale images of Baddeley's error measure. Di Gesù and Starovoitov (1999) have presented three image distance-based functions for digital image comparison. Experimental results indicate better function sensitivity by combining both global intensity and local structural features with respect to conventional intensity-based measures. Unlike the RMS criterion which takes into account gray level variations only, those approaches include comparisons in both amplitude and spatial domains.

In this paper we introduce a new dissimilarity measure, based on a full 3D discrete distance operator, as well as a simplified phenomenological model describing its behavior. Unlike the Wilson–Baddeley–Owner (W–B–O) operator, no distance saturation effect is introduced. The main advantages of the new dissimilarity measure are:

- relative weighting of spatial and gray level distortion,
- invariance with respect to video inversion,
- quasi linear response to spatial and amplitude distortions,
- lower complexity (with respect to W–B–O operator).

Section 2 introduces the image representation. Section 3 briefly describes the W–B–O measure. In Section 4, the principle of the new operator is introduced. The experimental study of this criterion is developed in Section 5, and a comparison between the two methods is performed.

2. Image representation

Let $S \subset Z^2$ be the referential on which images are defined. Generally $S = \{0,1,\ldots,N-1\} \times \{0,1,\ldots,N-1\}$, with N the image dimension. Let $G = \{0,1,\ldots,255\}$ be the set of gray level values.

Let us consider two gray-scale images A and B, defined on referential S. Image A (resp. B) can be regarded as a function $f_A : S \to G$ (resp. f_B), or function $F_A : S \times G \to \{0,1\}$ (resp. F_B), with $s = (x,y) \in S$ a pixel.

$$F_A(s,g) = 1 \quad \text{iff } f_A(s) = g,$$
$$= 0 \quad \text{else.}$$

Hence, an image can be regarded as a binary set of points in a 3D volume. Thus we can apply Baddeley's distance of this set of points. It requires the computation of distances between voxels of volume $V = S \times G$. These distances are obtained by means of a 3D local distance operator (Borgefors, 1984; Verwer, 1991).

Voxels are regarded as parallelepipedic-shaped bricks of size $L \times H \times P$ (see Fig. 1). For the sake of simplicity, we take $L = H = 1$ (pixel size unit). Parameter P control the weight of gray level variations with respect to spatial distortions (see Section 4.4).

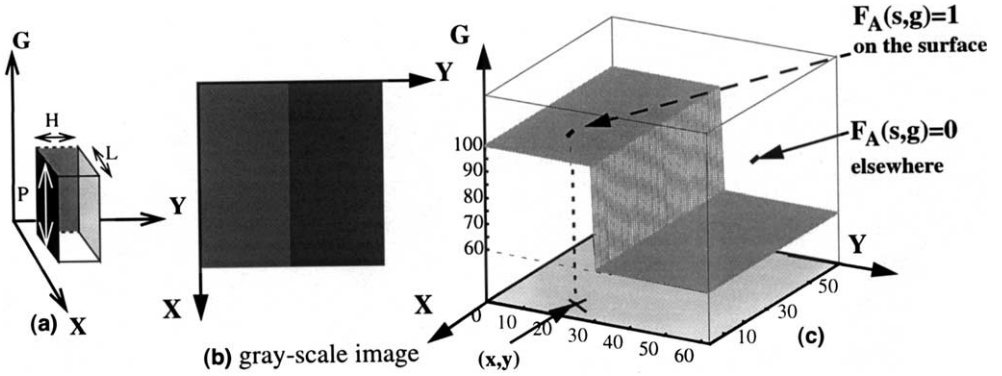

Fig. 1. Image representation: (a) voxel shape, (b) gray level representation of f and (c) binary function F.

Representing an image as a surface rather than a subgraph is preferable. In fact, the dissimilarity calculation takes all the voxels in set $S \times G$ into account. Using the subgraph representation, only voxels located above the image surfaces contribute to the dissimilarity measurements. This introduces an asymmetry in the processing. For instance, the dissimilarity between two images differs from that of their inverse video versions (see Appendix E).

3. Dissimilarity measures

In this section, the principle of Baddeley's distance between binary images is recalled. We then present its extension to gray level images proposed by Wilson et al. (1997).

3.1. Baddeley's distance (1992)

Baddeley's distance between two binary images A and B $(A, B \subseteq S)$, is defined as

$$\Delta_b(A,B) = \left[\frac{1}{\mathrm{card}(S)} \sum_{s \in S} |d_A(s) - d_B(s)|^E\right]^{1/E} \quad (1)$$

with $\mathrm{Card}(S)$ the number of elements in referential S, $d_A(s) = \min_{a \in A} d(s,a)$ is the distance between a point $s = (x,y)$ and a set A, and exponent E, such that $1 \leqslant E < \infty$. Baddeley's distance is an average measure. Hence, it is less sensitive than the Hausdorff distance to small and localized distortions. It can be shown that the Baddeley's distance tends to the Hausdorff distance as coefficient E tends to infinity (Baddeley, 1992).

3.2. Wilson–Baddeley–Owen dissimilarity measure (1997)

Wilson et al. introduced a new gray scale measure in 1997.

Let $f : S \to G$ be any picture function. According to Wilson et al. (1997), the subgraph of f is defined as: $\Gamma_f = \{(s,g); (s \in S, g \in G, g \leqslant f(s))\}$. This is the set of all points lying between the graph of f and the plane $g = 0$.

Let f_A and f_B be two picture functions having the same number of possible gray levels. Let Γ_{f_A} and Γ_{f_B} be the subgraphs of f_A and f_B, respectively. The gray level metric Δ_g is then defined, for $1 \leqslant E < \infty$, as

$$\Delta_g(\Gamma_{f_A}, \Gamma_{f_B}) = \left[\frac{1}{\mathrm{card}(S) \cdot \mathrm{card}(G)} \sum_{s \in S} \sum_{g \in G} |d^*[(s,g), \Gamma_{f_A}] - d^*[(s,g), \Gamma_{f_B}]|^E\right]^{1/E}, \quad (2)$$

where $d^*[(s,g), \Gamma_{f_A}] = \inf_{g';(|g-g'| \leqslant c)} \min(\max\{d[s, X_{g'}(A)], |g-g'|\}, c)$ is a distance function which gives the shortest distance between a voxel $(s,g) \in S \times G$ and the subgraph of $f_A \cdot X_{g'}(A) = \{s \in S; (f_A(s) \geqslant g')\}$ is the upper-level set at gray level g' (Fig. 2).

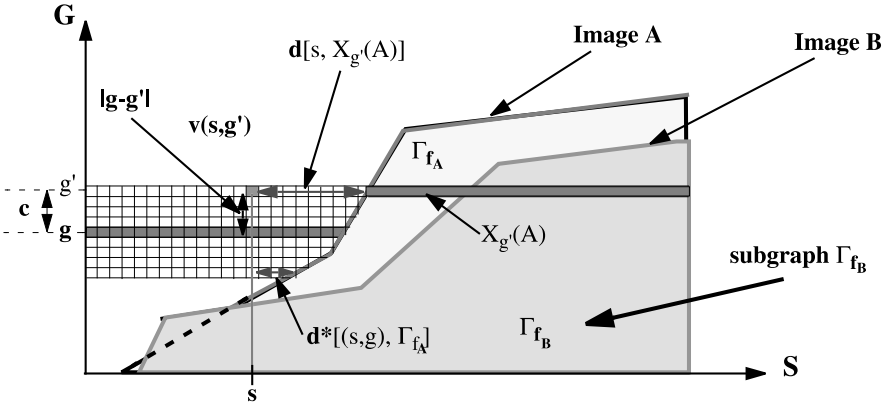

Fig. 2. Distance from a point s and image A.

The value of E determines the relative importance of large localization errors. For large E, the effect is similar to the Hausdorff metric. It should be noticed that distance d^* is bounded by value c. Wilson et al. suggests to truncate distances by constant c, such that only pixels s such that the amplitude is close to those of images are taken into account.

The principle of calculation of d^* is the following: for each g' layer one seeks the larger of two distances: the first one is the spatial distance between voxel (s,g) and the subgraph Γ_{f_A}; the second one is the gray level distance $|g-g'|$. d^* is then the smallest distance on all the gray levels g' on which search is made. In the implementation proposed by Wilson et al. S is a two-dimensional raster and the distance $d[s,X_{g'}(A)]$ between pixel s and the upper-level set $X_{g'}(A)$ in the domain S is calculated by means of the discrete 2D chamfer distance operator d_{5-7-11} (Borgefors, 1984). Calculations are carried out plane by plane.

The authors have shown that for certain types of distortions resulting from smoothing or compression, there is little difference between the errors detected by Δ_g and those found by the RMS criterion. However, as far as "feature" distortion is concerned, the response of Δ_g is a considerable improvement on the response of RMS (Wilson et al., 1997).

Setting the c value for the other image sizes can be done by considering noise magnitude. Let us consider a high resolution noisy image I and its low resolution version J. If J is obtained from I by a $N \times N$ size low pass filter and down sampling by factor N, the ratio of noise standard deviations will be $(\sigma_J/\sigma_I) \approx 1/N$. The scale of gray level distortions is divided by N. We use the same ratio for parameter c. In the following, we set constant c in proportion of the image size (e.g. $c = 8$ for 128×128 images). Anyway, it can be experienced that dissimilarity measures are not very sensitive to the value of parameter c. In their paper (Wilson et al., 1997), they suggest to take $c = 4$ for 64×64 images.

4. The new operator

In this section, we introduce a new dissimilarity measure based on Baddeley's distance between binary objects.

4.1. Principle

With the same notations as in Section 2, let A and B be the images to be compared. The new operator is defined, for $1 \leqslant E < \infty$, as

$$\mathscr{D}(A,B) = \left[\frac{1}{\mathrm{card}(V)} \sum_{v \in V} |d_A(v) - d_B(v)|^E \right]^{1/E}, \quad (3)$$

where $\mathrm{card}(V)$ is the number of voxels $v = (s,g)$ in the volume $V = S \times G$ on which this dissimilarity is computed, $d_A(v)$ (resp. $d_B(v)$) is the shortest distance between voxel v and the binary set characterizing image A (resp. B) (Fig. 3). Hence, we have $d_A(v) = \min_{p \in A} d(v,p)$.

The calculation of $\mathscr{D}$ is done on the whole space V, and not only voxels on the surfaces, so that the least difference between the two images is amplified, by accumulation.

4.2. Properties of $\mathscr{D}$

$\mathscr{D}$ is a metric, since it satisfies the following axioms:

(i) $\mathscr{D}(A,B) = 0$ if and only of $A = B$;

(ii) symmetry: $\mathscr{D}(A,B) = \mathscr{D}(B,A)$;

(iii) triangle inequality:

$$\mathscr{D}(A,B) \leqslant \mathscr{D}(A,C) + \mathscr{D}(C,B). \quad (4)$$

Proof.

(i) $A = B \Rightarrow \mathscr{D}(A,B) = 0$ is obvious. $\mathscr{D}(A,B) = 0 \Rightarrow A = B$? Let s be a pixel such that $f_A(s) \neq f_B(s)$. Let voxel $v = [s, f_A(s)]$, then $d_A(v) = 0$ and $d_B(v) \neq 0$. Hence $A \neq B \Rightarrow \mathscr{D}(A,B) \neq 0$.

(ii) Formula are symmetrical with respect to A and B.

(iii) We have

$$\left[\sum_{v \in V} |d_A(v) - d_B(v)|^E \right]^{1/E} = \left[\sum_{v \in V} |d_A(v) - d_C(v) + d_C(v) - d_B(v)|^E \right]^{1/E} \quad (5)$$

Let us consider functions $h(v) = d_A(v) - d_C(v)$ and $g(v) = d_C(v) - d_B(v)$.

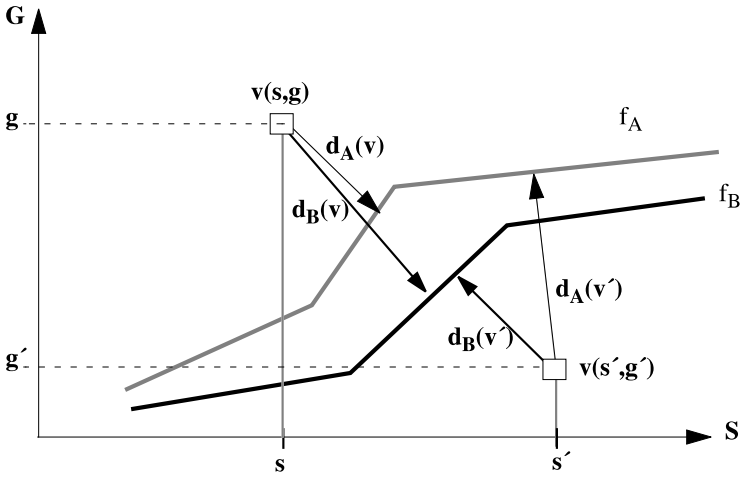

Fig. 3. Dissimilarity between two images.

The Lp-norm defined by $\|h\|_p = \left[\sum_{v\in V} |h(v)|^p\right]^{1/p}$.

We then have $\left[\sum_{v\in V} |d_A(v) - d_C(v)|^E\right]^{1/E} = \|h\|_E$ and $\left[\sum_{v\in V} |d_C(v) - d_B(v)|^E\right]^{1/E} = \|g\|_E$ and $\left[\sum_{v\in V} |d_A(v) - d_B(v)|^E\right]^{1/E} = \|h+g\|_E$.

Using Minkowski's inequality $\|h+g\|_E \leqslant \|h\|_E + \|g\|_E$, with $1 \leqslant E < \infty$, we have $\mathscr{D}(A,B) \leqslant \mathscr{D}(A,C) + \mathscr{D}(C,B)$ and triangle inequality is valid.

4.3. Implementation

In order to reduce the computational complexity, Euclidean distances $d_A(v)$ and $d_B(v)$ are approximated by means of a 3D local distance operator. A distance transformation image (Borgefors, 1984) is computed, with a $3\times3\times3$ local distance operator on a parallelepipedic grid. The distance transformation is calculated with integer coefficients, on the two sides of the reference surface. It should be noticed that since images are digitized, their representations may be nonconnected sets, as far as 26-connectivity is concerned. This does not introduce any singularity, since the distance transform is defined on the whole volume $V = S \times G$ and all the voxels of volume V are taken into account in the dissimilarity calculation. Sequential or parallel algorithms can be used (Rosenfeld and Pfaltz, 1966).

4.4. Parametrization

The voxel size is characterized by three parameters H, L and P (Fig. 1(a)). H and L values depend on the spatial sampling period of the data. Parameter P depends on the weight of gray level differences with respect to spatial distortions (Fig. 4).

Special case: $H, L \gg P$

The cost of spatial displacements is very large with respect to that of displacements along the gray level axis G. The minimal path between f_B and f_A is vertical.

Hence, we have

$$|d_A(v) - d_B(v)| = |f_A(s) - f_B(s)| = |g_A - g_B|. \qquad (6)$$

Only gray level differences are taken into account (see Appendix A).

5. Experimental results

In this section we compare the properties of the W–B–O dissimilarity and those of the new operator. Since it is a classical measurement, we compare them with the RMS criterion. We consider spatial displacements, gray level variations and shape distortions. If the dissimilarity between two

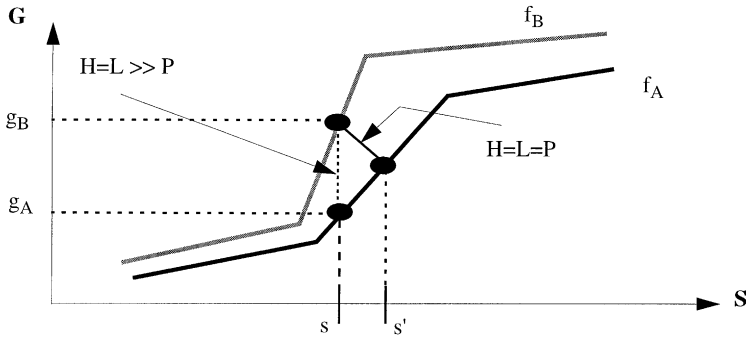

Fig. 4. Distances between images.

images has to be minimized with respect to a group of transformation, it may be desirable that the dissimilarity measure be proportional to the control parameter value.

In this section, we considered simple transformations such as spatial translations and gray level shifts. The results can be explained by means of a phenomenological model introduced in Appendix C. It is more difficult to analyze rotations because of the spatial discretization and interpolation processes. Further studies about transformations such as rotation, zooming, ..., have to be considered.

In order to make the comparisons easier, measurements are normalized on a 0–100% scale. 100% is the dissimilarity between a white image A_{white} (maximum gray level) and a black image A_{black} (minimum gray level).

Let $\mathscr{D}(A, A')$ be the dissimilarity between two images A and A', the *normalized dissimilarity* is then defined as

$$\mathscr{D}_N(A, A') = \frac{D(A, A')}{D(A_{\text{white}}, A_{\text{black}})}. \tag{7}$$

5.1. Effect of a spatial shift

The reference image is horizontally shifted to the left by increasing vectors (Fig. 5).

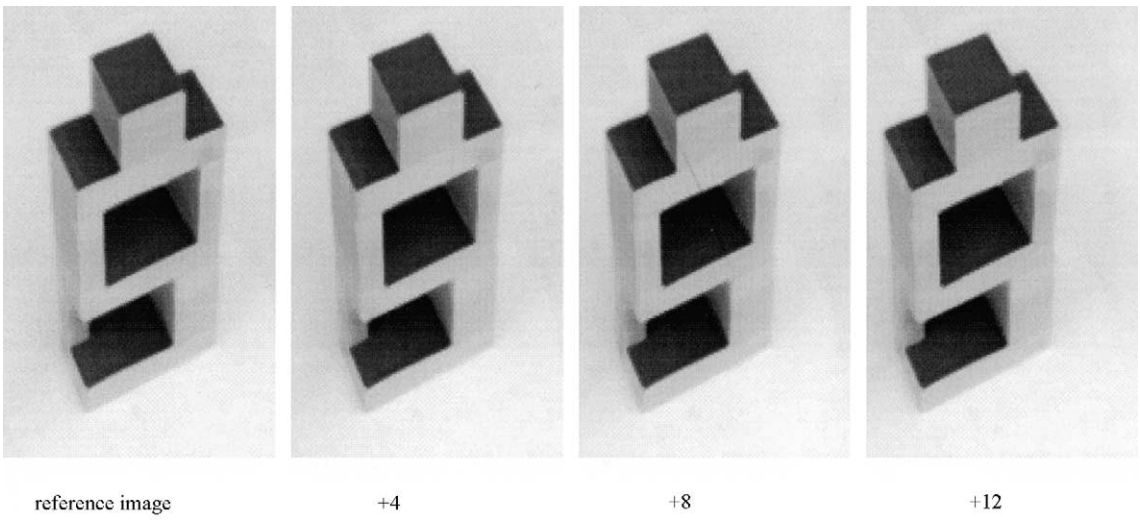

Fig. 5. Original image and shifted images.

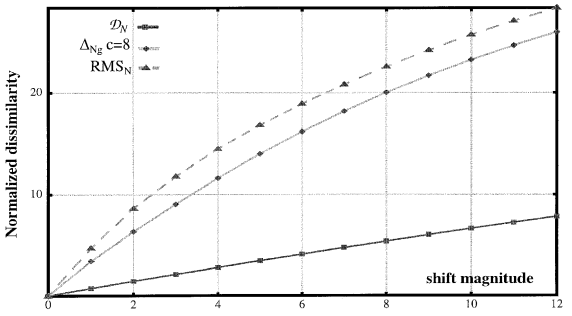

Fig. 6. Dissimilarity measurement versus shift magnitude.

Fig. 6 displays the normalized dissimilarity between the shifted image and the reference images as a function of the shift magnitude.

It can be noticed that the variation of the new dissimilarity $\mathscr{D}$ measure is linear. This is not the case for both RMS and W–B–O Δ_g criterion which tend towards a saturation value for large shift magnitude.

5.2. Constant gray level increase

The second experiment consists of an iterative increase of 10 gray levels on image. The results are obtained by comparing each modified image with the original undistorted image (Fig. 7). The original image Fig. 7(a) is available at http://www.inrialpes.fr/movi/pub/Images/index.html.

Fig. 7(b) shows synthetic images. The original image is a white noise, uniformly distributed with standard deviation $\sigma = 5$ and mean value $m = 128$.

It can be noticed of Fig. 8 that the RMS error variation is linear, as well as that of the new dissimilarity measure $\mathscr{D}$. This is not the case for the W–B–O measure Δ_g. Such a linear or quasi-linear behavior is obtained for lower gray level increases (see Appendix B). As mentioned in Section 3.2, parameter c plays a significant role. In (Wilson et al., 1997) the authors adapted the c value to the image size. For 64×64 images, they recommended using $c = 4$. In our experiment, the image size is 128×128. Parameter c is set to 8.

It should be noticed that the linear behavior is achieved if most of voxels are far from both image surfaces, e.g. if the gray level increase is low or the image contrast is low (see Appendix C).

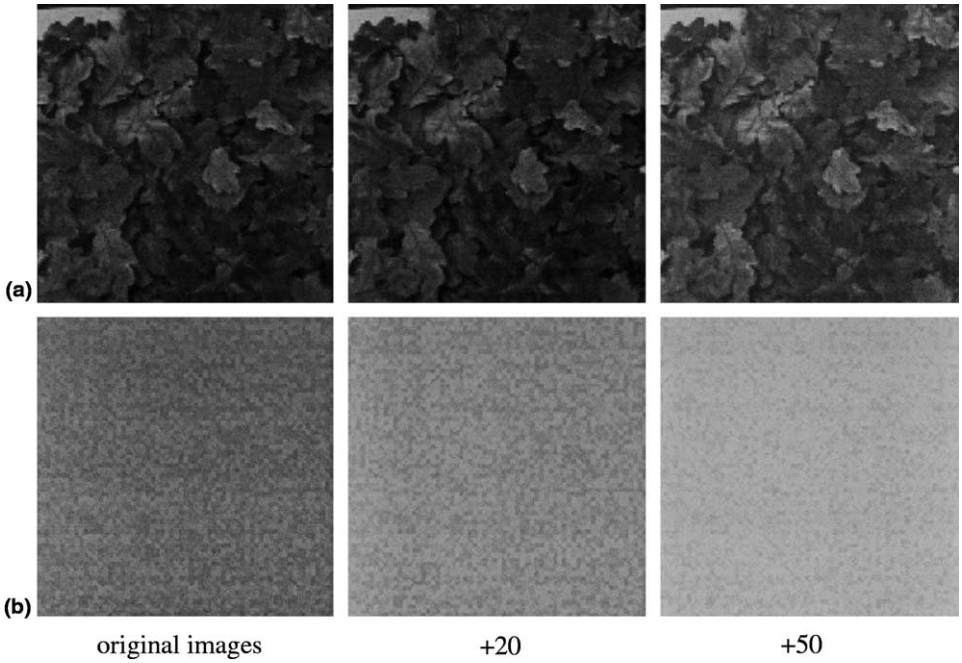

Fig. 7. Original images and gray level translated images.

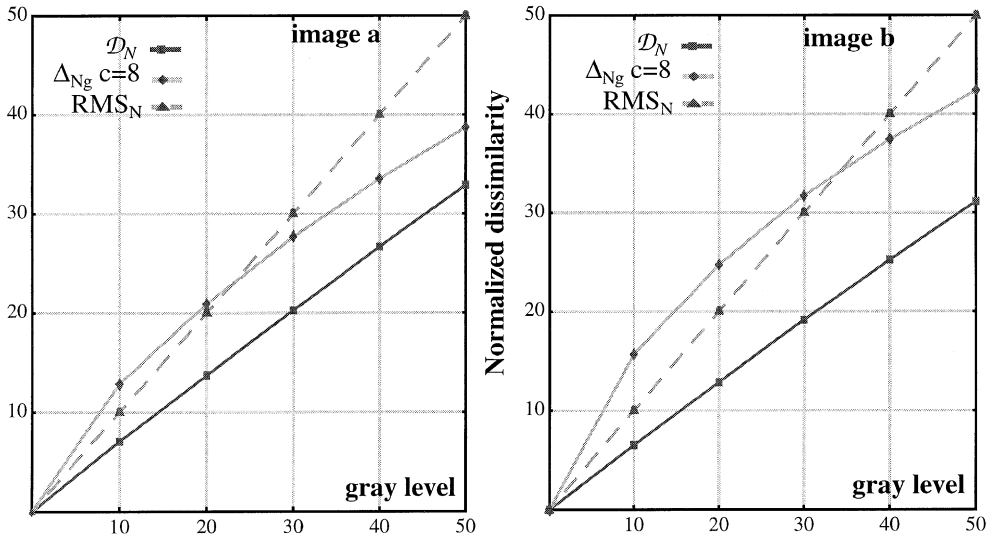

Fig. 8. Dissimilarity measurement versus gray level variation.

5.3. Combination of constant gray level increase and spatial shift

The linear behavior of the new dissimilarity operator can be observed by combining spatial and gray level distortions. Fig. 9 shows the reference image and three distorted images. The distortion consists of a spatial translation to the left by δ pixels and a gray level increase by δ gray level units. Fig. 10 displays the dissimilarity between the distorted images and the reference image as a function of the distortion parameter δ.

The linearity range is greater for operator $\mathcal{D}$ than for both the RMS and W–B–O Δ_g criteria. It should be noticed that the slope of the dissimilarity curve depends on the processed image.

5.4. Sensitivity to shape distortion

In this section, we study the effect of a shape distortion introduced by camera rotation. The image sequence is composed of 78 views of an object. The camera turns around the object. The whole rotation is slightly greater than 360°. This sequence was elaborated by Patrick Gros and is

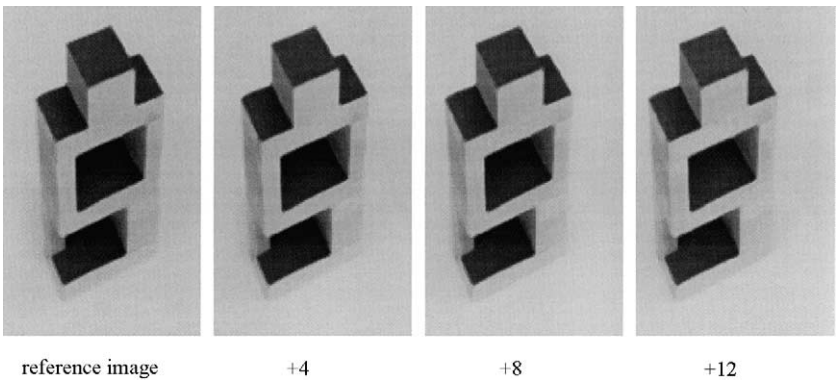

Fig. 9. Original image and distorted images.

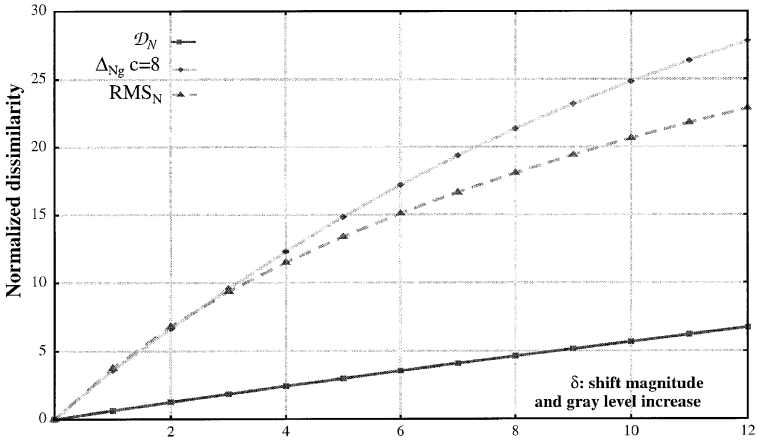

Fig. 10. Dissimilarity measurement versus shift magnitude, and gray level increase.

available at http://www.inrialpes.fr/movi/pub/Images/index.html. Fig. 11 shows some images from this sequence.

It should be noted that im1 is quite similar to im68. However image im33 displays the opposite side of the object. Image im12 and im47 are similar too, although they correspond to opposite sides.

The difference between im12 and im47 is located in the lower part of the object (Fig. 12). Image im12 presents a hole, which corresponds to dark pixels. This region is similar to the corresponding zone in image im1. Image im47 presents a planar facet composed of high value pixels. This is the reason for which the dissimilarity between im1 and im47 is greater (Fig. 13). Fig. 13 displays the dissimilarity measures against the image number. The reference image is im1. In our experiments exponent E was set to 2. For E ranging from 1 to 8, it was experienced that the results are fairly insensitive to this parameter.

5.5. Influence of ratio P/H

Fig. 14 shows the influence of ratio P/H on the dissimilarity measurement. It can be shown, as described in Section 4.4, that the dissimilarity measure tends towards the RMS measure as ratio P/H tends to zero. For the lowest value ($P/H = 0.1$), the curve obtained is similar to that of the RMS criterion.

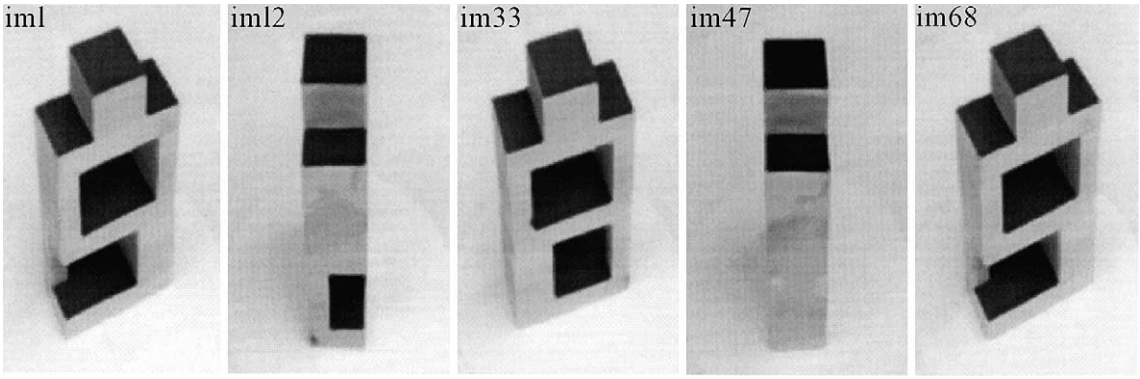

Fig. 11. Original image im1 and several images stemmed from the sequence.

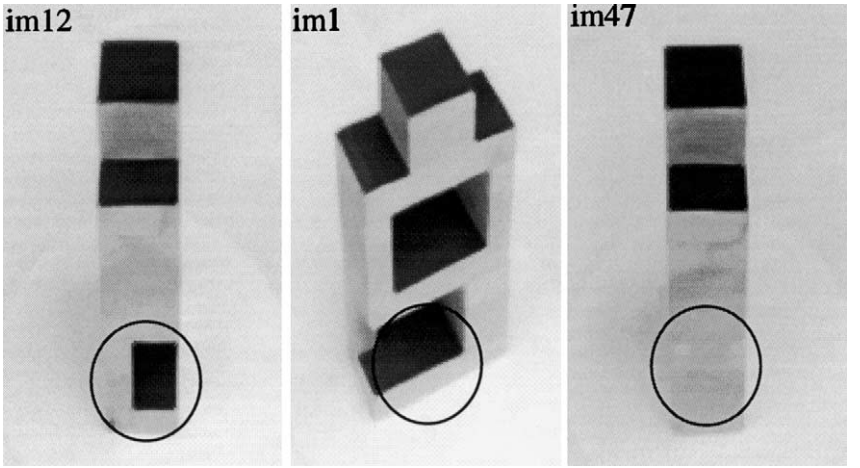

Fig. 12. Difference between original image im1, images im12 and im47.

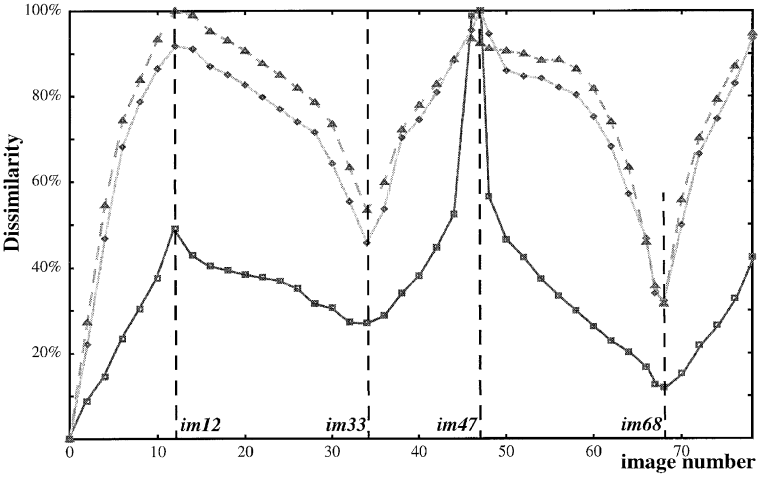

Fig. 13. Camera rotation. Dissimilarity measurement versus image number: (open square) new operator $\mathscr{D}$; (open triangle) RMS; (open diamond) W–B–O Δ_g $c=8$.

As P/H ratio increases, the geometrical deformations have more weight than the difference between gray levels. If we want that the dissimilarity operator is sensitive to both geometrical and gray level distortions, a value of P/H greater than 1 has to be used. Of course, the best choice is application dependent, according to whether gray level distortions or geometrical distortions are regarded as the most important criterion. There is no uniformly best P/H value. This raises the question of the dimension. In which units are distances measured? In fact, there is not a simple and definitive answer. If only gray level distortions are considered, the distance units might be Watts (power) or Joules (energy). If only translations are considered, it may be Meters. In our case, we consider arbitrary units, weighting both spatial and gray level distortions. In fact, the solution could be to consider not a distance, obtained for a given operator parametrization, but a signature, e.g. the set of distances, obtained for a given operator parametrization. Such an approach was used in (Coquin

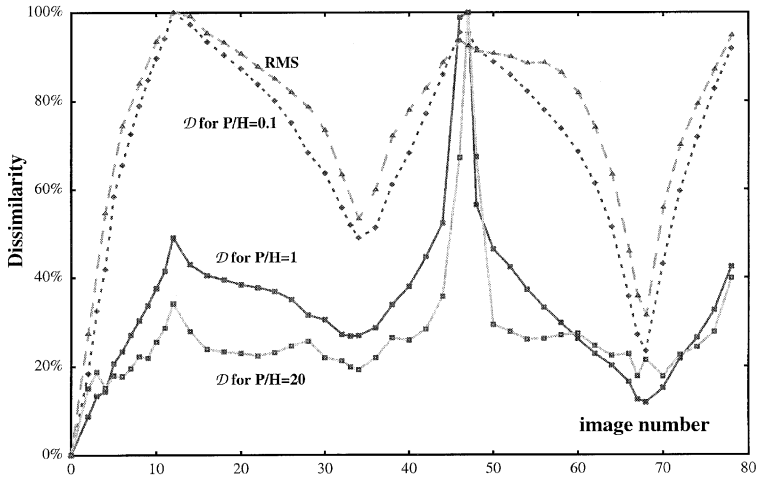

Fig. 14. New dissimilarity operator. Influence of ratio P/H. P/H values 0.1; 1; 20.

et al., 1995, 1997) but it makes comparisons more difficult.

5.6. Sensitivity to noise

Sensitivity to noise depends on the P/H ratio. If P/H tends to 0, the new operator behaves like the RMS (see Section 4 and Appendix A). Its sensitivity to noise is high. For larger P/H values, the new operator is sensitive to spatially coherent distortions. Hence, its sensitivity to noise is reduced provided that the noise correlation width is less than the size of the features of interest. This can be seen by considering the following experiment.

Let A be a reference image having a constant gray level. Let $B1$ and $B2$ be two distorted versions of A. Image $B2$ is obtained by adding a constant gray level 1 to image A. Image $B2$ is obtained by adding a random zero mean unit variance binary white noise to image A. The non-normalized dissimilarity measures are given in Table 1.

It can be seen that increasing the P/H ratio decreases the sensitivity to noise.

Table 1
Sensitivity to noise

	RMS	$\mathscr{D}$ with $P/H = 0.1$	$\mathscr{D}$ with $P/H = 1$
$d(A, B1)$	1	1	0.8
$d(A, B2)$	1	1	1

5.7. Application to database characterization

Image indexing and image retrieval systems hae to be tested by applying them to image databases. Results actually depend on both processing systems and database content. In fact, if the performance assessment depends on a classification rate, results are sensitive to within-class homogeneity and between-class diversity. Good performances obtained with a test database having a low within-class homogeneity and a low between-class diversity are more significant than those obtained with a database having a large within-class homogeneity and a large between-class diversity.

These characteristics can be evaluated by using the distances between the images of the test database. A simplified example of two databases is given below.

Fig. 15 displays 25 images gathered in two databases. Database 1 consists of 2 classes: 10 images resulting from Food and Agriculture bases. Database 2 consists of 3 classes: 15 images resulting from Couples, Children and Nature bases. A database is composed of N_c class $C_k,\ k = 1, 2, \ldots, N_c$. Each class C_k is composed of images $I_i^k,\ i = 1, 2, \ldots, \mathrm{card}(C_k)$.

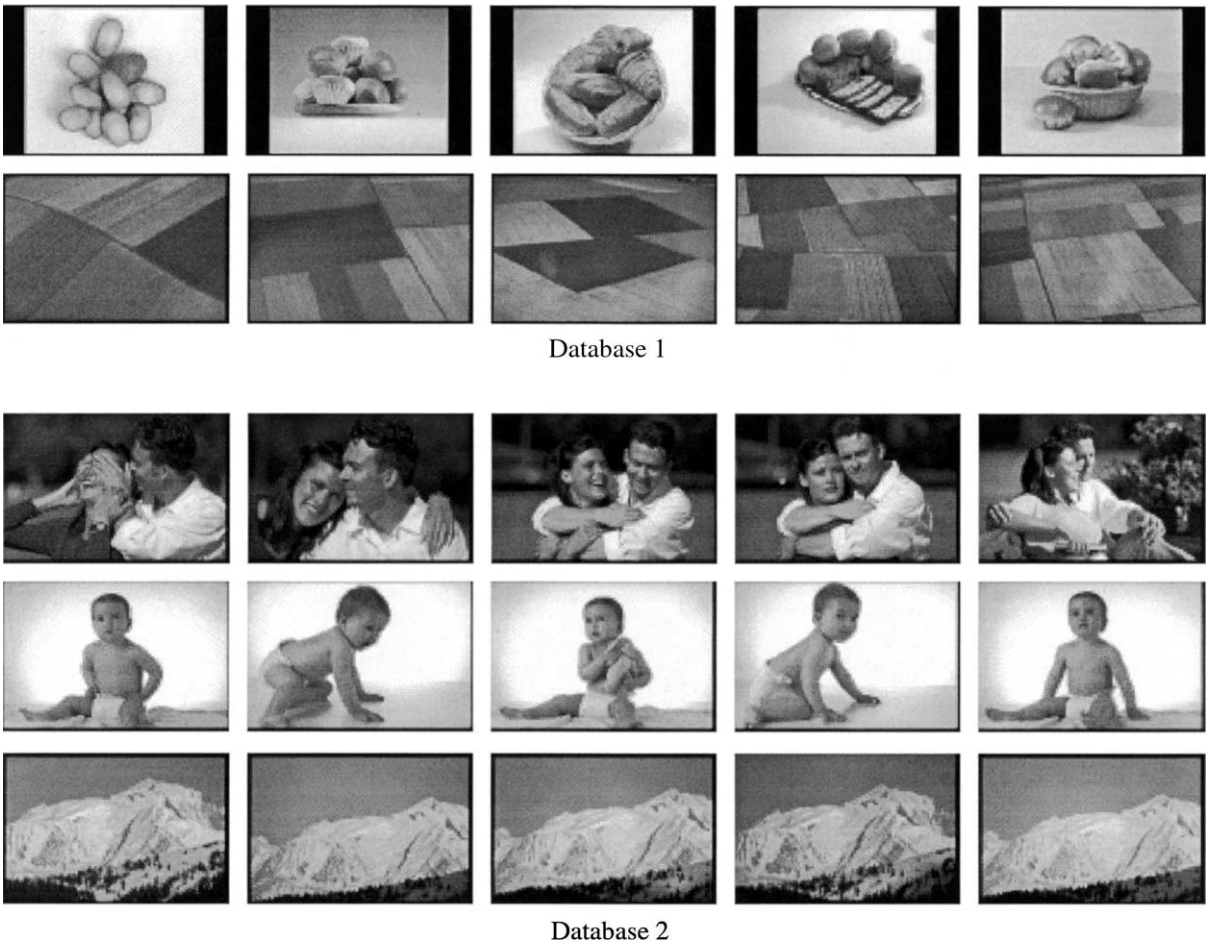

Fig. 15. Two databases.

Let $\mathscr{D}(I_i^k, I_j^k)$ be the distance between image I_i^k and image I_j^k of class C_k. The center of class C_k is the image I^k such that $G^k = \sum_{I_j^k \in C_k} \mathscr{D}(I^k, I_j^k)/\mathrm{card}(C_k)$ is minimal. The homogeneity of class C_k can be defined by $H^k = 1/G^k$.

The within-class homogeneity of the whole database is

$$H = \frac{\sum_k H^k}{N_c}.$$

The between-class diversity of the database can be defined by

$$Dv = \frac{\sum_{k=1}^{N_c} \sum_{j \neq k} \mathscr{D}(I^k, I^j)}{N_c(N_c - 1)}.$$

With the simplified database of Fig. 15, we have

$$H_1 = 0.2, \quad Dv_1 = 22.9 \quad \text{and} \quad H_2 = 0.31, \quad Dv_2 = 8.8.$$

Databases can be compared with respect to the dimensionless coefficient $r = Dv \cdot H$. This coefficient measures how easy the discrimination between classes is. With this example we have:

database 1: $r_1 = Dv_1 \cdot H_1 \approx 4.6$;

database 2: $r_2 = Dv_2 \cdot H_2 \approx 2.7$.

In this example, the classification results obtained with database 2 are more significant that those obtained with database 1.

Using W–B–O operator and RMS criterion gives the following results:

Table 2
Computation time (CPU in s)

Image size	W–B–O $\Delta_g(c=4)$	$\mathscr{D}$	Speedfactor
256×256	89.90	36.90	2.43
192×128	33.39	14.07	2.37
128×128	21.97	9.41	2.33
64×64	5.68	2.48	2.29

$r_{1(\mathrm{WBO})} \approx 1.96$, $r_{2(\mathrm{WBO})} \approx 2.17$ and
$r_{1(\mathrm{RMS})} \approx 2.25$, $r_{2(\mathrm{RMS})} \approx 2.43$.

With both criteria, the two databases are almost equally significant. As can be seen in Fig. 13, results obtained with RMS and W–B–O operator are similar.

As mentioned in Wilson et al. (1997), the RMS criterion is not the best one as soon as feature distortions are concerned. Hence, it should not be regarded as a reference. It should be noticed that, these results are given as illustrative examples. This study should be conducted with greater databases. This is beyond the scope of the paper.

5.8. Computation time

Compared with the W–B–O Δ_g operator, the time required to compute the new dissimilarity is decreased by a factor greater or equal to 2.3. The speed factor actually depends on both image content and image size. Typical values obtained with Sun Workstation Sparc Ultra 10.333 MHz, are shown in Table 2.

Distance maps are stored using 16 bit words, with an integer scaling factor. It should be noticed that the values are obtained without algorithmic optimization.

6. Conclusion

In this paper, an application to gray-scale images of the binary metric developed by Baddeley (1992) is introduced. It aims at combining both intensity variations and geometrical distortions. These two contributions can be weighted by means of a tuning parameter namely the ratio P/H. This new dissimilarity is compared with previously published ones and with the classical RMS operator. We have shown in various examples that the new operator has a linear behavior with respect to continuous distortions such as gray level variations and geometrical shifts. Because of its sensitivity to gray level variations and geometrical distortions, this operator can be helpful in different situations such as: image database characterization, image filtering operation assessment, digital elevation model and range image comparison, etc.

Acknowledgements

The authors are indebted to Dr. Piero Zamperoni (1939–1998) for valuable, helpful and friendly discussions about image comparison. They wish to thank Professor R. Mohr from INRIA Rhone-Alpes, Dr. P. Gros from IRISA and Mr. Ph. Bigard from GoodShoot for providing the data and helpful discussions, within the framework of Region Rhone-Alpes project ACTIV.

Appendix A. Special case: $H, L \gg P$

Let us consider a 3D local distance operator on parallelepipedic grid. Chamfer mask coefficients d_{ijk} (Fig. 16) are given by (Coquin et al., 1994; Coquin and Bolon, 1995):

$$d_{ijk} = T_{ijk} \cdot \frac{d_{100}}{L} \quad \text{with}$$
$$T_{ijk} = \sqrt{(iL)^2 + (jH)^2 + (kP)^2}, \tag{A.1}$$
$$d_{100} = \frac{-2H + 2H\sqrt{1+\lambda}}{\lambda} \quad \text{with}$$
$$\lambda = \frac{1}{L^2}(T_{110} - H)^2 + \frac{1}{p^2}(T_{111} - T_{110})^2. \tag{A.2}$$

These coefficients are function of displacement d_{100} and are given by

$$\begin{aligned} d_{010} &= H \cdot \frac{d_{100}}{L}, \quad d_{001} = P \cdot \frac{d_{100}}{L}, \\ d_{110} &= T_{110} \cdot \frac{d_{100}}{L}, \quad d_{101} = T_{101} \cdot \frac{d_{100}}{L}, \\ d_{011} &= T_{011} \cdot \frac{d_{100}}{L}, \quad d_{111} = T_{111} \cdot \frac{d_{100}}{L}. \end{aligned} \tag{A.3}$$

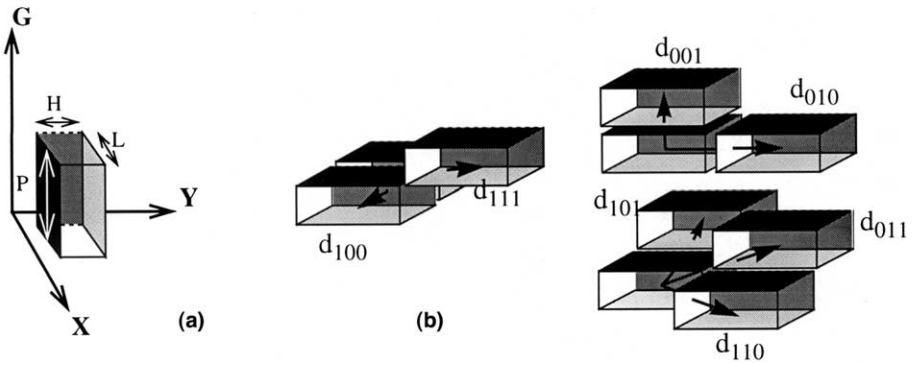

Fig. 16. Operator $3 \times 3 \times 3$ in a parallelepipedic gride: (a) voxel characteristics and (b) elementary displacements.

Hence, we have

$$\lambda = (\sqrt{2}-1)^2 + \frac{L^2}{P^2}\left(\sqrt{\frac{P^2}{L^2}+2}-\sqrt{2}\right)^2 \qquad (A.4)$$

since $L \gg P$ then $P/L \ll 1$ and $\lambda \approx (\sqrt{2}-1)^2 = 0.1715$, then

$$d_{100} = d_{010} \approx L\left(1-\frac{\lambda}{4}\right), \quad d_{110} \approx L\sqrt{2}\left(1-\frac{\lambda}{4}\right),$$
$$d_{001} \approx P \ll L, \quad d_{101} = d_{100}\frac{\sqrt{L^2+P^2}}{L} \approx L\left(1-\frac{\lambda}{4}\right)$$
$$d_{011} = d_{101} \approx L\left(1-\frac{\lambda}{4}\right), \quad d_{111} \approx L\sqrt{2}\left(1-\frac{\lambda}{4}\right). \qquad (A.5)$$

The distance between two voxels is the length of the shortest path between them. A path is composed of a finite number of elementary displacements. The length of an elementary displacement is d_{ijk}, according to its direction. Since $P \ll L$, any path containing a non-vertical elementary displacement is longer than a vertical path composed of a finite number of vertical elementary displacements (elementary length $= d_{001}$).

Let A (resp. B) be the subset of $S \times G$ representing the picture function f_A (resp. f_B) (Fig. 17). The distance between voxel $v = (s, g)$ and set A is

$$d_A(v) = |f_A(s) - g| \cdot P \qquad (A.6)$$

If we set $P = 1$, $L = H = \infty$ and if the total number of voxels is high with respect to the number of voxel lying between set A and set B, the distance between image A and image B is

$$\mathscr{D}(A,B) \approx \left[\frac{1}{\mathrm{card}(S \times G)}\sum_{v \in S \times G}|d_A(v) - d_B(v)|^E\right]^{1/E}, \qquad (A.7)$$

$$\mathscr{D}(A,B) \approx \left[\frac{1}{\mathrm{card}(S \times G)}\sum_{s \in S}\sum_{g \in G}|f_A(s) - f_B(s)|^E\right]^{1/E}, \qquad (A.8)$$

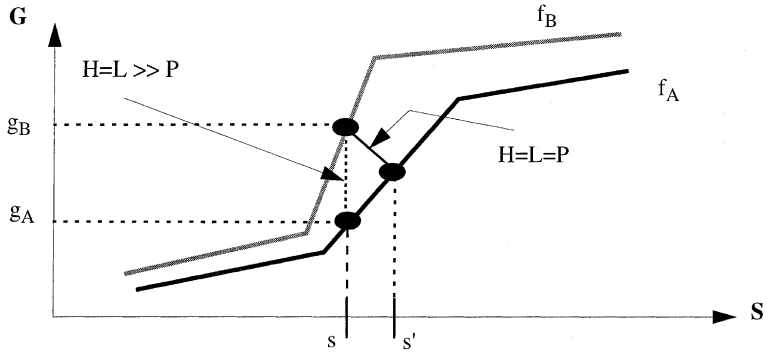

Fig. 17. Distance between two images.

$$\mathscr{D}(A,B) \approx \left[\frac{1}{\mathrm{card}(S)} \sum_{s \in S} |f_A(s) - f_B(s)|^E \right]^{1/E}. \quad \text{(A.9)}$$

If exponent $E = 2$, then $\mathscr{D}(A,B) \approx \mathrm{RMS}(A,B)$.

Appendix B. Example of constant gray level variation

With a simple example, we study the behavior of the W–B–O measure and the new $\mathscr{D}$ measure. This study points out the influence of voxels located outside the image surfaces on the linear behavior of the dissimilarity operator. In the following, the distorted image B is obtained by adding the constant gray level h to image A. Both images A and B have constant gray level.

B.1. Δ_g metric (Wilson et al. 1997)

Let f_A and f_B be two picture functions having the same number of possible gray levels. Let Γ_{f_A} and Γ_{f_B} be the subgraphs of f_A and f_B, respectively. The gray-level metric Δ_g is then defined, for $1 \leqslant E < \infty$, as

$$\Delta_g(\Gamma_{f_A}, \Gamma_{f_B}) = \left[\frac{1}{\mathrm{card}(S) \cdot \mathrm{card}(G)} \sum_{s \in S} \sum_{g \in G} |d^*[(s,g), \Gamma_{f_A}] - d^*[(s,g), \Gamma_{f_B}]|^E \right]^{1/E}, \quad \text{(B.1)}$$

where $d^*[(s,g),\Gamma_{f_A}] = \inf_{g';(|g-g'|)\leqslant c)} \min(\max\{d[s, X_{g'}(A)], |g-g'|\}, c)$ is a distance function which gives the shortest distance between a point $(s,g) \in S \times G$ and the subgraph of f_A. And $X_{g'}(A) = \{s \in S; (f_A(s) \geqslant g')\}$ is the upper-level set at gray-level g'.

Example. Let us consider two images A and B which have constant gray levels. For example a reference image A, (picture function f_A with constant gray level = 0) and a test image B, (picture function f_B with constant gray level $h = 10$), then

$$d^*[(s,g),\Gamma_{f_A}] = \inf_{g';(|g-g'|)\leqslant c)} \min(\max\{d[s, X_{g'}(A)], |g-g'|\}, c) \quad \text{(B.2)}$$

is equal to

$$d^*[(s,g),\Gamma_{f_A}] = \begin{cases} |g-g'| & \text{if } (|g-g'|) \leqslant c \\ c & \text{else.} \end{cases} \quad \text{(B.3)}$$

Therefore

$$\Delta_g(A,B) = \left[\frac{1}{N^2 \mathrm{card}(G)} \sum_{s \in S} \sum_{g \in G} |d_A - d_B|^2 \right]^{1/2} \quad \text{if exponent } E = 2 \quad \text{(B.4)}$$

$$\Delta_g(A,B) = \left\{ 2(1 + 2^2 + 3^2 + \ldots + (c-1)^2) + c^2(g+1-c)/\mathrm{card}\ (G) \right\}^{1/2} \quad \text{(B.5)}$$

here with $c = 8$ and $h = 10$

$$\Delta_g(A,B) = 1.36. \quad \text{(B.6)}$$

If image A_{black} is equal to 0 (*black image*) and image $B = A_{\mathrm{white}}$ is equal to 255 (*white image*) then with $c = 8$:

$$\Delta_g(A_{\mathrm{white}}, A_{\mathrm{black}}) = 7.94 \quad \text{(B.7)}$$

the *normalized dissimilarity* Δ_{Ng} is then defined as

$$\Delta_{Ng}(A,B) = \frac{\Delta_g(A,B)}{\Delta_g(A_{\mathrm{white}}, A_{\mathrm{black}})} = 17.1\%. \quad \text{(B.8)}$$

If $c = 4$ the *normalized dissimilarity* Δ_{Ng} is equal to

$$\Delta_{Ng}(A,B) = \frac{\Delta_g(A,B)}{\Delta_g(A_{\mathrm{white}}, A_{\mathrm{black}})} = 18.6\%. \quad \text{(B.9)}$$

Since the images have constant gray-levels, parameter c does not significantly modify the dissimilarity measure (see Fig. 18).

B.2. Application of Baddeley metric to gray level images

Let A and B be the images to be compared. The distance between A and B is defined by

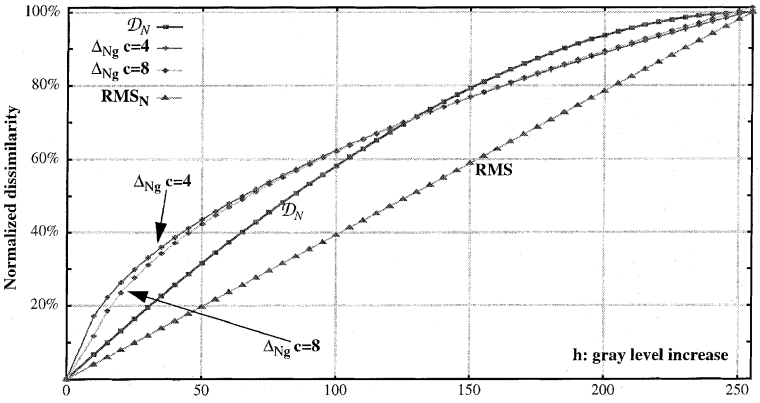

Fig. 18. Normalized dissimilarities versus constant gray level variations.

$$\mathscr{D}(A,B) = \left[\frac{1}{\mathrm{card}(V)} \sum_{v \in V} |d_A(v) - d_B(v)|^E \right]^{1/E} \quad \text{(B.10)}$$

with $1 \leqslant E < \infty$, voxel $v = (s,g)$, and $\mathrm{card}(V)$ is the number of pixel in the volume $V = S \times G$ on which we can compute this dissimilarity.

Image A is an image with constant gray level $= 0$ and image B is an image with constant gray level h.

Table 3 shows the difference between measures (W–B–O with $c = 8$ and $\mathscr{D}$) on calculation of the absolute difference $|d_A - d_B|$ between the two images A and B.

It should be noticed that the contribution of voxels located above the two image surfaces (e.g. $g \geqslant 10$) is equal to the gray level distortion h. The average contribution of voxels located between the two surfaces is proportional to h with a proportionality factor $\alpha < 1$.

Table 3
Some values of the difference between image A and image B

W–B–O operator				New operator		
Gray level: g	$d_A = d^*[(s,g),\Gamma_{f_A}]$	$d_B = d^*[(s,g),\Gamma_{f_B}]$	$\lvert d_A - d_B \rvert$	$d_A(v) - d_A(s,g)$	$d_B(v) - d_B(s,g)$	$\lvert d_A - d_B \rvert$
19	8	8	0	19	9	10
18	8	8	0	18	8	10
17	8	7	1	17	7	10
16	8	6	2	16	6	10
15	8	5	3	15	5	10
14	8	4	4	14	4	10
13	8	3	5	13	3	10
12	8	2	6	12	2	10
11	8	1	7	11	1	10
10	8	0	8	10	0	10
9	8	0	8	9	1	8
8	8	0	8	8	2	6
7	7	0	7	7	3	4
6	6	0	6	6	4	2
5	5	0	5	5	5	0
4	4	0	4	4	6	2
3	3	0	3	3	7	4
2	2	0	2	2	8	6
1	1	0	1	1	9	8
0	0	0	0	0	10	10

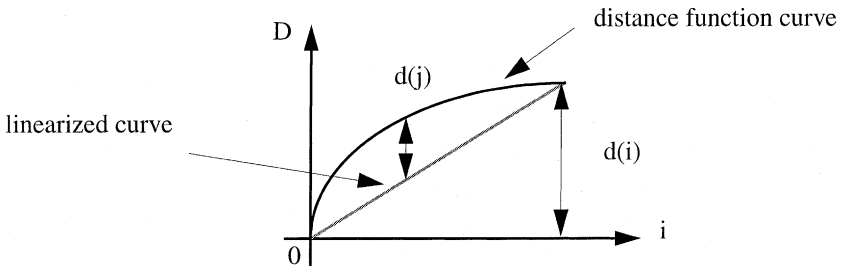

Fig. 19. Curvature parameter.

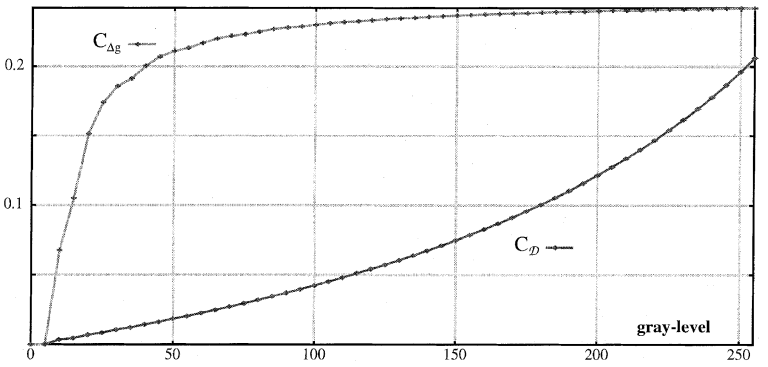

Fig. 20. Curvature parameter as function of the gray level variation.

If we take exponent $E = 2$, $\mathscr{D}(A,B)$ is given by

$$\mathscr{D}(A,B) = \sqrt{\frac{\sum_{g<h}(2g-h)^2 + \sum_{g \geqslant h} h^2}{\mathrm{card}(G)}}. \qquad \text{(B.11)}$$

If image A_{black} is equal to 0 (*black image*) and image $B = A_{\mathrm{white}}$ is equal to 255 (*white image*)

$$\mathscr{D}(A_{\mathrm{white}}, A_{\mathrm{black}}) = 147.8 \qquad \text{(B.12)}$$

the *normalized dissimilarity* $\mathscr{D}_N$ is then defined as

$$\mathscr{D}_N(A,B) = \frac{\mathscr{D}(A,B)}{\mathscr{D}(A_{\mathrm{white}}, A_{\mathrm{black}})} = 6.7\%. \qquad \text{(B.13)}$$

B.3. Comparison between the two operators

The reference image A has constant gray level 0. The distorted image B has constant gray level h. Fig. 18 shows the normalized dissimilarities obtained with the W–B–O operator ($c = 4$ and $c = 8$), the new dissimilarity $\mathscr{D}$ and the RMS, as a function of distortion parameter h.

To characterize the dissimilarity curve linearity, we calculate a curvature parameter defined as

$$C_d(i) = \max_{j \in \{0,1,2,\ldots i\}} \frac{\left| \frac{(j.d(i))}{g} - d(j) \right|}{d(i)}. \qquad \text{(B.14)}$$

This parameter is the normalized vertical distance between the distance function curve and the linearized curve (Fig. 19).

It can be noticed (Fig. 20) that the new dissimilarity measure has a more linear behavior than that of Wilson et al. For a gray-level variation equal to 100, the curvature parameter is less than 5%, whilst that of W–B–O operator is larger than 20%.

Appendix C. Constant gray level distortion, and spatial distortion

In this appendix we consider a simplified image model consisting of plateaus separated by sharp transitions. The typical transition height is T. The typical number of transitions is m. S is the image support, G is the set of gray levels.

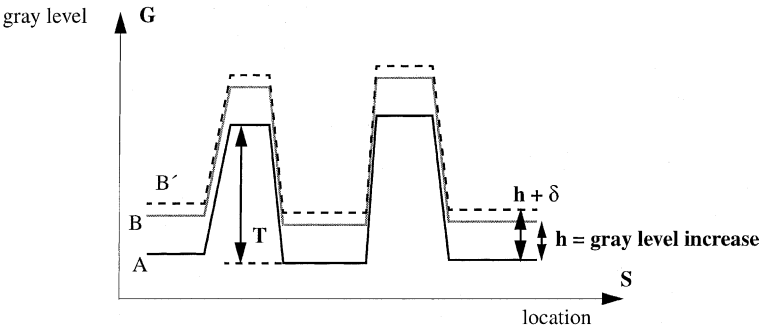

Fig. 21. Gray level distortion, A: reference image, B and B': distorted images.

C.1. Gray level distortion

Let A be the reference image. Let B be a distorted version of A obtained by adding a constant gray level h (Fig. 21). According to the results given in Appendix B, voxels v which are located far from surfaces A and B are such that $|d_A(v) - d_B(v)| = h$. Conversely, for voxels located near or between surfaces, the absolute value of the distance difference is less than h. Let N_1 be the number of those voxels. Let αh (with $0 \leqslant \alpha < 1$) be the average distance difference for those voxels. Coefficient α actually depends on the number of transitions, on their height and on their shape. Let N_2 be the number of far voxels, with $N_1 + N_2 = \mathrm{card}(G) \times \mathrm{card}(S)$.

The dissimilarity between image A and image B is given by

$$\mathscr{D}(A,B) = \left[\frac{1}{\mathrm{card}(V)}\sum_{v\in V}|d_A(v) - d_B(v)|^E\right]^{1/E} = \left[\frac{N_1(\alpha h)^E + N_2 h^E}{N_1 + N_2}\right]^{1/E}. \quad \text{(C.1)}$$

Let us consider a small increase δ in the gray level distortion. Let B' be the new distorted image. The number of voxels considered as near (resp. far) from the surfaces are then

$$N_1' = N_1 + k\delta \quad N_2' = N_2 - k\delta \quad \text{with} \quad k = \mathrm{card}(S). \quad \text{(C.2)}$$

Hence, we have

$$\mathscr{D}(A,B') = \left[\frac{N_1'[\alpha(h+\delta)]^E + N_2'(h+\delta)^E}{N_1' + N_2'}\right]^{1/E} = (h+\delta)\left(\frac{N_1\alpha^E + N_2}{N_1 + N_2}\right)^{1/E} \times \left[1 + \frac{k\delta(\alpha^E - 1)}{N_1\alpha^E + N_2}\right]^{1/E}. \quad \text{(C.3)}$$

By the first-order Taylor expansion, we can show that the dissimilarity increment is given by

$$\mathscr{D}(A,B') - \mathscr{D}(A,B) = \eta(\delta - Kh\delta) \quad \text{(C.4)}$$

with

$$\eta = \left[\frac{N_1\alpha^E + N_2}{N_1 + N_2}\right]^{\frac{1}{E}} \quad \text{and} \quad K = \frac{k(1-\alpha^E)}{E(N_1\alpha^E + N_2)}.$$

For lower values of gray level distortion h, $N_2 \approx \mathrm{card}(G) \times \mathrm{card}(S)$ then coefficient $K \approx 1/\mathrm{card}(G) \cdot E$ is small with respect to 1.

The dissimilarity increment is $\eta\delta$. Hence, the dissimilarity variation is linear with respect to the gray level distortion h. Coefficient η is less than 1. It is decreasing function of the average transition magnitude T. This means that a gray level increase by 1 is regarded as more important if T is low than if T is large. For large values of the gray level distortion h, the dissimilarity increment is decreased. This explains the curvature of the dissimilarity measurement as a function of the distortion parameter.

C.2. Spatial shift distortion

Let A be the reference image. Let B be the image obtained by shifting image A be Δ pixels. Let B' be the image obtained by an additional shift of δ pixels (Fig. 22).

Like in Section C.1, the dissimilarity between image A and image B can be given by

$$\mathscr{D}(A,B) = \left[\frac{1}{\mathrm{card}(V)}\sum_{v\in V}|d_A(v)-d_B(v)|^E\right]^{1/E} = \left[\frac{N_1(\beta\Delta)^E + N_2\Delta^E}{N_1+N_2}\right]^{1/E} \quad \text{(C.5)}$$

with $0 \leqslant \beta < 1$.

In fact, N_1, N_2, β depend on the transition height T and the number of voxels between two transitions. Parameter N_2 is the more larger than T is high. If the spatial shift magnitude is increased by δ, we have

$$N_1' = N_1 + mT\delta, \quad N_2' = N_2 - mT\delta. \quad \text{(C.6)}$$

Using a first-order Taylor expansion, the dissimilarity increment is given by

$$\mathscr{D}(A,B') - \mathscr{D}(A,B) = \mu(\delta - \Gamma\Delta\delta) \quad \text{(C.7)}$$

with

$$\mu = \left[\frac{N_1\beta^E + N_2}{N_1+N_2}\right]^{1/E} \quad \text{and} \quad \Gamma = \frac{mT(1-\beta^E)}{E(N_1\beta^E+N_2)}.$$

Like in the case of a gray level increase, the dissimilarity operator is linear with respect to the shift magnitude if this magnitude is small. The proportionality coefficient is less than 1.

Appendix D. Integer coefficients

The integer coefficients used in this study are given in Table 4.

The other coefficients are obtained taking symmetries into account.

Appendix E. Effect of video inversion

Let A and B be two images represented by their surface sets. Let A' and B' be their video inverse. With the new operator, we have $\mathscr{D}(A,B) = \mathscr{D}(A',B')$.

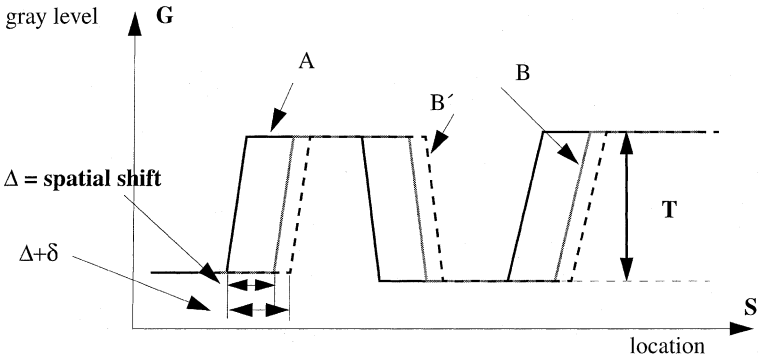

Fig. 22. Spatial shift distortion, A: reference image, B and B' distorted images.

Table 4
Integer coefficients

P/H	D_{100}	D_{010}	D_{001}	D_{110}	D_{101}	D_{011}	D_{111}
0.1	108	108	11	153	108	108	153
1	16	16	16	23	23	23	28
20	16	16	313	22	313	313	314

As defined in Section 4,

$$\mathscr{D}(A,B) = \left[\frac{1}{\mathrm{card}(V)}\sum_{v\in V}|d_A(v) - d_B(v)|^E\right]^{1/E}.$$

In fact, the Euclidean distance is invariant with respect to symmetry in the gray level domain. We have the same property if local distance operators are used, provided that their coefficients are symmetrical with respect to the horizontal plane. This is what is obtained following (Borgefors, 1984; Coquin et al., 1994; Verwer, 1991).

Let us consider the voxel $v = (s,g)$. Let voxel $a(v) \in A$ be such that $d_A(v) = d[v,a(v)]$. Let $v' = (s,g')$ be the symmetrical of voxel v with respect to a given horizontal plane (i.e. constant gray level). Let a' be the symmetrical of voxel a with respect to the horizontal plane.

We then have $d_A(v) = d[v,a(v)] = d[v',a'(v')] = d_{A'}(v')$. Since each voxel v can be associated with its symmetrical v' we obtain

$$\begin{aligned}\mathscr{D}(A,B) &= \left[\frac{1}{\mathrm{card}(V)}\sum_{v'\in V}|d_A(v') - d_B(v')|^E\right]^{1/E}\\ &= \left[\frac{1}{\mathrm{card}(V)}\sum_{v\in V}|d'_A(v) - d'_B(v)|^E\right]^{1/E}.\end{aligned}$$

Hence $\mathscr{D}(A,B) = \mathscr{D}(A',B')$.

If subgraph image representations are used, voxels v and their symmetrical v' do not play the same role. It can easily be shown that invariance with respect to video inversion cannot be guaranteed.

References

Baddeley, A.J., 1992., in: An Error Metric for Binary Images. Robust Computer Vision, Wichmann, Karlsruhe, pp. 59–78.

Borgefors, G., 1984. Distance transformations in arbitrary dimensions. CVGIP 27, 321–345.

Coquin, D., Chehadeh, Y., Bolon, Ph., 1994. 3D local operator on parallele pipedic grid. In: Proc. 4th Discrete Geometry for Computer Imagery, Grenoble, France, pp. 147–156.

Coquin, D., Bolon, Ph., 1995. Discrete distance operator on rectangular grids. Pattern Recognition Lett. 16, 911–923.

Coquin, D., Bolon, Ph., Chehadeh, Y., 1995. Opérateurs de distance 3D – Application à la comparison d'images. In: Proc. 15th colloque GRETSI, Juan-Les-Pins, France, pp. 761–764.

Coquin, D., Bolon, Ph., Chehadeh, Y., 1997. Evaluation quantitative d'images filtrées. In: Proc. 16th colloque GRETSI, vol. 2, Grenoble, France, pp. 1351–1354.

Di Gesù, V., Starovoitov, V., 1999. Distance-based functions for image comparison. Pattern Recoginition Lett. 20, 207–214.

Dubuisson, M.P., Jain, A.K., 1994. A modified Hausdorff distance for object matching. In: Proc. 12th Internat. Conf. on Pattern Recognition, Jerusalem, Israel, pp. 566–568.

Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J., 1993. Comparing images using the Hausdorff distance. IEEE Trans. Pattern Anal. Machine Intell. 15 (9), 850–863.

Jacobs, D.W., Weinshall, D., 2000. Classification with nonmetric distances: image retrieval and class representation. IEEE Trans. Pattern Anal. Machine Intell. 22 (6), 583–600.

Rosenfeld, A., Pfaltz, J., 1966. Sequential operations in digital picture processing. J. ACM 13, 471–494.

Verwer, B., 1991. Local distances for distance transformations in two and three dimensions. Pattern Recognition Lett. 12, 671–682.

Wilson, D.L., Baddeley, A.J., Owen, R.A., 1997. A new metric for grey-scale image comparison. Internat. J. Comput. Vision 24 (1), 5–18.

Zamperoni, P., Starovoitov, V., 1996. On measures of dissimilarity between arbitrary gray-scale images. Internat. J. Shape Modeling 2 (2&3), 189–213.