A NEW SIMILARITY MEASURE USING HAUSDORFF DISTANCE MAP

Étienne Baudrier, Gilles Millon, Frédéric Nicolier, Su Ruan

CReSTIC (previously LAM)
IUT de Troyes, 9, rue de Québec, 10026 TROYES CEDEX, France
{e.baudrier, g.millon, f.nicolier, s.ruan}@iut-troyes.univ-reims.fr

ABSTRACT

Image dissimilarity measure is a hot topic. The measure process is generally composed of an information mining in each image which results in an image signature and then a signature comparison to take the decision about the image similarity. In the scope of binary images, we propose in this paper to replace the information mining by a new straight image comparison which does not require *a priori* knowledge. The second stage is then replaced by a decision process based on the image comparison. The new comparison process is structured as follows: a morphological multiresolution analysis is applied to the two images. Secondly a distance map is constructed at each scale by the computation of the Hausdorff distance restricted through a sliding-window. A signature is then extracted from the distance map and is used to take the decision. As an application, the algorithm has been successfully tested on an ancient illustration database.

Introduction

Image retrieval is an active domain. Retrieving images by their content, as opposed to meta-data, has become an important activity. It is classically composed of two stages: firstly, an information mining, which results in an image signature and secondly, a signature distance measure that is used to decide on the image similarity. In this process, the signature must capture conspicuous features in order to be as discriminating as possible in some user-defined sense. In general, the signature contains color, shape or texture information [1]. But the choice of the signature attributes is not easy and depends on the treated images. In the scope of binary images, we propose to replace this awkward information mining by a straight image comparison based on a modified Hausdorff distance (HD) [2, 3] producing a distance map. The second stage is then replaced by a decision process based on the distance map. While an information mining requires a priori knowledge on discriminating features before comparing the images, our process first expresses dissimilarities from the image comparison before taking a decision. This process developed for binary images is adaptable to pattern recognition. In this paper, we present the different stages of the measure process: firstly a morphological multiresolution analysis, secondly the construction of a HD map between two images. Then, a decision on the similarity of the images based on the distance map is presented at the end of section 2. Finally, we expose some results to show the efficiency of our method.

1. MULTIRESOLUTION

Human dissimilarity measure can be viewed as a coarse-to-fine process. As a consequence, in a first approximation, a dissimilarity measure can be carried on a low resolution, which allows in addition to save computation time. This can be done thanks to a Multiresolution Analysis (MRA). Nevertheless, the scale-space operator should satisfy conditions to preserve the binary image main features. Many classical scale-space operators use gaussian functions. They have good space-scale properties but they smooth transitions which results in a loss in the binary information and could produce errors in low resolutions. This drawback is a common property of linear filters. On the contrary, non-linear filters can avoid this problem. Among them, morphological operators are good candidates [4]. We have determined three criteria to choose the morphological MRA operator τ : τ should be edge-preserving, τ should be auto-dual (i.e. τ preserves the black-to-white pixels ratio) and τ should preserve the "main" features. Obviously, the last criterion is subjective and is satisfied *a posteriori*. This led us to choose the morphological so-called median operator which fulfills these conditions [5]. The morphological MRA is thus described by the following process :

1. non-linear median filtering (on a 2×2 window) of the approximation a_j,

2. down-sampling by a factor 2, giving a_{j-1} and details

3. repeat the process up to scale J.

Even if this study focuses on approximations, details may be exploited and are obtained by the following formula:

$$\mathbf{D_v}(i,j) = |\mathbf{I}(2i-1,2j-1) - \mathbf{I}(2i-1,2j)| \quad (1a)$$
$$\mathbf{D_d}(i,j) = |\mathbf{I}(2i-1,2j-1) - \mathbf{I}(2i,2j)| \quad (1b)$$
$$\mathbf{D_h}(i,j) = |\mathbf{I}(2i-1,2j-1) - \mathbf{I}(2i,2j-1)| \quad (1c)$$

where $I(2i-1,2j-1), I(2i-1,2j), I(2i,2j-1)$ and $I(2i,2j)$ stand for the four pixels in the 2×2 window of the filter.

2. DISTANCE MAP

Instead of processing a comparison of two image signatures, the distance map allows direct dissimilarity representation including topographic information on the dissimilarities. This piece of information can be exploited later in the decision process. The distance map rests on the two points explained below, the HD and a local computation.

2.1. Hausdorff distance

The HD has often been used in the content-based retrieval domain. Originally meant as a measure between two point collections A and B in a metric space $\mathcal{E}$ (whose distance is d), it can be viewed as a dissimilarity measure between two binary images $\mathbf{A}$ and $\mathbf{B}$ considering A and B, respectively the black pixels finite set of points of $\mathbf{A}$ and $\mathbf{B}$. For finite sets of points the HD can be defined as follows:

definition 1 $D_H(A,B) = \max\{\vec{d}(A,B), \vec{d}(B,A)\}$, $\vec{d}(A,B) = \max_{a \in A} \min_{b \in B} d(a,b)$ with $d(a,b)$ the underlying distance.

We use the same notation: $D_H(\mathbf{A}, \mathbf{B}) = D_H(A,B)$ for images. The interest of this measure comes partly from its metric properties (in our application, on the space of finite sets of points): non-negativity, identity, symmetry and triangle inequality. These properties correspond to our intuition for image resemblance. Another reason is that the HD measure between an image and its shifted copy is the norm of the translation vector. Thus, the HD matches our intuition in case of translation. Nevertheless, it measures the most mismatched points between A and B, which is not convenient. There is an extension that reduces this drawback, the so-called *partial Hausdorff distance* which measures the distance to the k nearest points. This is no longer a metric, but it does not take into account the most distant points which can be irrelevant for the measure. Nevertheless, it remains a global measure over images. As we want precise information about local differences, we introduce the local measure that is shown thereafter.

2.2. Local measurement

2.2.1. Definition of the distance map

The HD and the partial one give a global dissimilarity measure over images. In many cases, it is very interesting to measure local differences, that is why we have introduced a new distance measure based on the HD, designated by $D_{H,\mathrm{W}}$: it consists in making a local measure using a sliding-window W. At each sample point, the HD is computed on the portion of the images viewed through the sliding-window W. It results in a distance map $\mathcal{M}_{k,l}$ depending on the sliding-window size (w_x, w_y) and on the sliding-step **p**. The sliding-window size determines the difference size highlighted in the distance map: the size of measured dissimilarities increases with the sliding-window size.

The HD distance map has properties which cannot be exposed here in an exhaustive way because of the lack of space:

- a sliding-window reduced to one pixel (and a sliding-step of one pixel) produces the matrix of the simple difference between the images,

- a sliding-window with the same size as the images gives a matrix fulfilled with a single value: the global HD between the images,

- it reduces the HD inconvenience of "the most mismatched points measure".

2.2.2. Illustration of the distance map

We show four images (fig.1 left and top) and their distance maps (fig.1 bottom row). The more different the images, the darker the distance map, moreover, the topographical distribution is less regular when the images are dissimilar.

3. DECISION

The automatic decision cannot be made straight from the distance map which is too large: a distance-map signature has to be extracted. A simple one is the distance-map histogram. Indeed, the HD is a max-min distance and as the possible distance values number is finite in the sliding-window, the HD takes only $l+1$ values $v_0, v_1, \ldots, v_l$. They are discrete and remain between $v_0 = 0$ and a maximum value n. Intuitively, the average value in a distance-map produced from two similar images will be smaller than in one from dissimilar images. Moreover, in the first case, the distance measured in an area through the sliding-window is linked to the one measured in a close region thanks to the correlation between the images, whereas, in the second case, as the images are independent, there is no reason for them to be correlated. The distance-map set is divided into
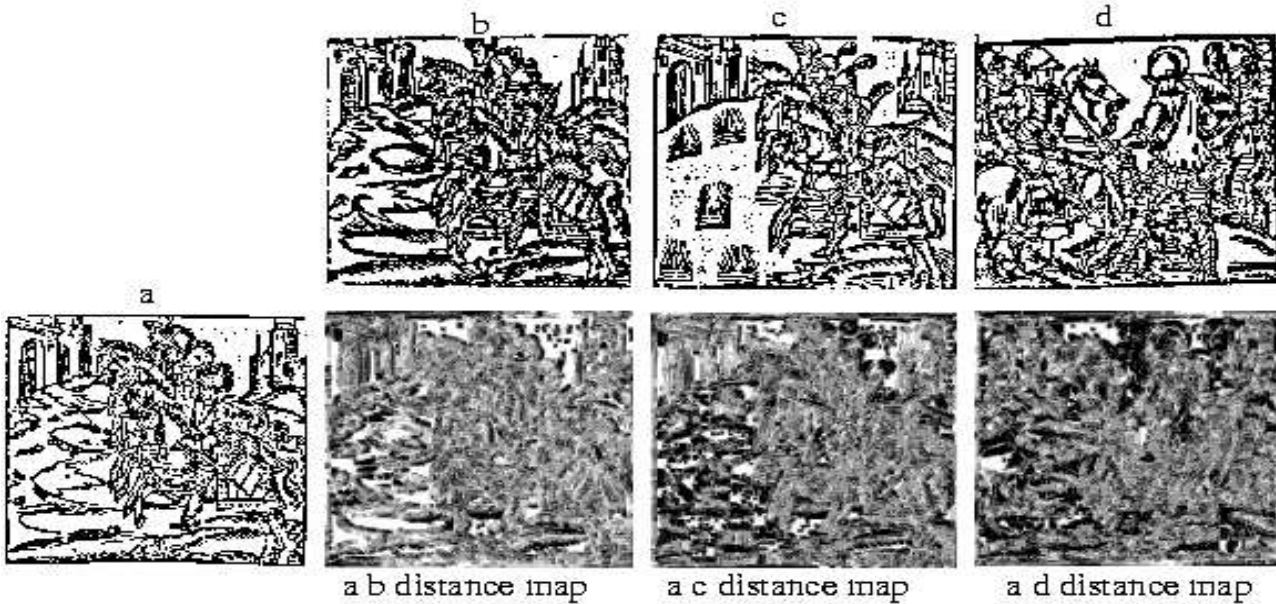
Fig. 1. Binary images (top and left) and their distance maps (bottom row) computed with a 5-pixel sliding-window side and a 1-pixel sliding-step. The distance values are from 0 (in white) to 5 (in black).

two classes -those that result from the comparison of *similar* images $\mathcal{C}_{sim}$, and those that result from the comparison of *different* images $\mathcal{C}_{dif}$. We represent the l first values of the histogram in a vector $X \in \mathbf{R}^l$. We make the assumption that the probability distribution of X is different in $\mathcal{C}_{sim}$ from the one in $\mathcal{C}_{dif}$. Then we model these two probability distributions by two gaussian distributions $G_{i \in \{sim, dif\}}$:

$$G_i(X) = \frac{1}{(2\pi)^{\frac{l}{2}} Det(\Gamma_i)} exp((X - m_i)^T \Gamma^{-1}(X - m_i))$$

Unlike the $l+1$ values of the histogram which are linearly linked, the l first values are linearly independent and so the variance-covariance matrix Γ_i is invertible. The values of Γ_i and of the averages m_i have to be computed on a learning set. Then we apply a bayesian decision to know if the compared images are similar or different.

4. RESULTS

The method is applied on a test database of digitalized ancient illustrations provided by Troyes' library within the framework of the project ANITA [6]. These images, originally printed on books with dark ink, have a strong contrast, which allows to binarize them with almost no loss. This database contains 70 images, some of which illustrate the same scene. Our objective is to test the method efficiency in retrieving similar images in the image database. The comparison of the images give 2307 distance maps. The distance maps have been computed from the third resolution

method	origine	found in $\mathcal{C}_{sim}$	found in $\mathcal{C}_{dif}$	successful retrieval
$D_{H,W}$	$\mathcal{C}_{sim}$	**145**	8	94%
	$\mathcal{C}_{dif}$	350	**1804**	92%
D_H	$\mathcal{C}_{sim}$	**70**	83	46%
	$\mathcal{C}_{dif}$	624	**1530**	71%
D_{pp}	$\mathcal{C}_{sim}$	**101**	52	66%
	$\mathcal{C}_{dif}$	474	**1680**	78%

Fig. 2. Table of results for $D_{H,W}$, the global HD D_H and the histogram of the pixel-to-pixel difference D_{pp}.

level of the AMR (see section 1) and are 128×128 pixels. First, we sort them manually in the two classes $\mathcal{C}_{sim}$ (153 items) and $\mathcal{C}_{dif}$ (2154 items) introduced in the former paragraph. Secondly, we apply our method so as to classify them automatically, which gives a decision ("the compared images are similar" or "are not similar"). Finally, we compare the results obtained manually with those obtained automatically. The results are summarized in fig. 2.

The results for $D_{H,W}$ have to be compared with those obtained with two other methods on the same data-base:

- the global HD,

- the pixel-to-pixel difference $|A - B|$ between the two images and with the same histogram model.

The efficiency is less satisfying. Thus, even if the information contained in the distance map is not completely ex-

PSNR	successful retrieval in $\mathcal{C}_{sim}$	successful retrieval in $\mathcal{C}_{dif}$
40 dB	100%	100%
35 dB	100%	70.5%
30 dB	100%	41.2%
25 dB	100%	47.0%
20 dB	92%	41.2%
15 dB	84%	41.2%

Fig. 3. Results for white noise robustness

ploited with this decision making, the use of a HD map improves the measure efficiency.

4.1. Robustness

4.1.1. Noise robustness

Method   The test is done on a 47 image database within similar and dissimilar images. For each image, six noisy images are produced with a *white noise* for a peak signal-to-noise ratio (PSNR) between 40 and 15 dB. At 40 dB, the noise is very small and at 15 dB, the original image is hardly visually detected. At each level, we have made 13 comparisons between similar images and 17 between dissimilar images.

Results   To measure robustness, we only consider image comparisons which have been well-classified by the method. Fig. 3 sums up the percentage of successful retrievals for both of the classes at each noise level.

Interpretation   The HD is sensitive to noise [7]. Our study shows that the use of a sliding-window reduces this sensitivity. Indeed if there is a faraway noise point, it will interfere with the distance measure only around the noise point, i.e. locally, in the distance map. Thus up to a 20dB PSNR, similar images have always been successfully retrieved. Concerning dissimilar comparisons, the successful retrievals represent only 41 % (the other 59 % are retrieved as similar images). The reason is that the images seem to become similar when the noise overflows the signal. This error corresponds to a false alarm. For a retrieval system, it means finding out more images than necessary, but it is less serious than missing correct images, which does not happen with this method.

5. CONCLUSION

In this paper, we have presented a dissimilarity measure that can be used on shapes or more complex images. This measure is based on a Hausdorff distance map which allows to specify the detail level to focus on. The use of a distance map makes the Hausdorff distance local and gives a good efficiency for the image comparison, thus better than the efficiency of the global Hausdorff distance. Furthermore the distance map allows the final user to catch at a glance the dissimilarity zones when comparing two images.

In future works, we aim at introducing the notion of shape to improve robustness, and at studying the choice of the sliding-window size; we then intend to exploit the coarse-to-fine aspect of the method in order to separate the "similar" class into two classes: "very similar" and "similar".

6. REFERENCES

[1] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, 2000.

[2] Daniel P. Huttenlocher and William J. Rucklidge, "A multi-resolution technique for comparing images using the hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 705–706, 1993.

[3] Remco C. Veltkamp and Michiel Hagedoorn, "Shape similarity measures, properties, and constructions," .

[4] Corinne Vachier, "Morphological scale-space analysis and feature extraction," in *Proc. of International Conference of Image Processing (ICIP)*, october 2001.

[5] H.J.A.M. Heijmans and J. Goutsias, "Multiresolution signal decomposition schemes. part 2: Linear and morphological pyramids," *IEEE Trans. Image Processing*, vol. 9, no. 11, pp. 1897–1913, november 2000.

[6] R. Seulin, Olivier Morel, G. Millon, and F. Nicolier, "Range image binarization: application to wooden stamps analysis," in *International Conference on Quality Control by Artificial Vision (QCAV) - IEEE*, Gatlinburg, Tenessee, USA, may 2003, vol. 5132, pp. 252–258, SPIE.

[7] R. Veltkamp and M. Hagedoorn, "State-of-the-art in shape matching," Tech. Rep. UU-CS-1999-27, Utrecht University, the Netherlands, 1999.