Binary Symbol Recognition from Local Dissimilarity Map

Frédéric Morain-Nicolier, Jérôme Landré, Su Ruan

CReSTIC - URCA, 9 rue de Québec, 10026 Troyes Cedex, France
E-mail: frederic.nicolier@univ-reims.fr

Abstract

In previous works, we have proposed a local dissimilarity map (LDM) in order to compare images. In this research, we show how the LDM can be applied in the field of symbol recognition. A global dissimilarity measure (GDM) is obtained from the LDM. This versatile allow to measure symetric as well as asymetric similarities. A matcher is derived by summing the values of the LDM. The obtained matcher is compared to the chamfer matching. Its properties are related to the human similarity judgement from Tversky results. It is tested on the grec2005 symbol recognition database. Good to excellent results are obtained without any knowledge on images, and no pre-processing nor segmentation involved.

Keywords: binary images, symbol recognition, similarity, asymetry, logpolar transform.

1 Introduction

The field of symbol recognition is a key step in a graphic recognition system. A large set of application domains takes benefit of its advances: electronics, architecture, musical score and so on. Mainy solutions are specifics responses to specifics problems, leading to *ad hoc* denoising or segmentation [7],

In previous works [2], we have designed a local dissimilarity measure leading to a local dissimilarity map (LDM). This map is computed using sliding window with an auto-adaptative size. Here is the general idea: if the pixels located in the sliding window belong to coarse features, the window is growed to be big enough to grasp the feature's distances. The LDM only allows to compare binary images, but without any signature extraction. This allows to avoid preprocessing. We propose in section 1 a reformulation of the LDM computation, in order to obtain only linear operations.

Binary images can be matched using the chamfer matching algorithm proposed by Borgefors [3]. This algorithm is popular due to its good results. In section 2, we show how a global dissimilarity measure (GDM) between two binary images can be computed from the LDM. Moreover, we links the chamfer matching with the GDM. Section 4 is concerned with making rotation and scale invariant the matcher, thanks to a log-polar transformation. Section 5 presents the final algorithm and the results.

2 Local Dissimilarity Map reformulation

Among distance measures over binary images, the Hausdorff distance (HD) has often been used in the content-based retrieval domain and is known to have successful applications in object matching [6]. The HD is defined by $HD(A, B) = max(h(A, B), h(B, A))$, with $h(A, B) = \max_{a\in A}(\min_{b\in B} d(a, b))$ where d is a underlying distance. This a global distance. We have proposed a local measure and a local dissimilarity map (LDM) computed with an iterative algorithm. We have also proposed a LDM formula allowing fast computations. For binary images it is :

$$\mathrm{LDM}_{A,B}(p) = |A(p) - B(p)| \max(\mathrm{dt}_A(p), \mathrm{dt}_B(p)), \tag{1}$$

where dt_X is the distance transform of X. Still for binary images, this equation can be simplified in :

$$\mathrm{LDM}_{A,B} = B\mathrm{dt}_A + A\mathrm{dt}_B. \tag{2}$$

Proof. The proof is quite forward. LDM_1 (resp. LDM_2) is LDM from eq. (1) (resp. eq. (2)). For a given pixel p, there is few alternatives :

1. $A(p) = B(p)$

$$\begin{aligned} A(p) = B(p) &\Rightarrow A(p) - B(p) = 0 \\ &\Rightarrow \mathrm{LDM}_1(p) = \mathrm{LDM}_2(p) = 0 \end{aligned}$$

2. $A(p) = 1 \; and \; B(p) = 0$

$$\begin{aligned} A(p) = 1 \; and \; B(p) = 0 &\Rightarrow A(p) - B(p) = 1 \\ A(p) = 1 &\Rightarrow \mathrm{dt}_A(p) = 0 \; (definition \; of \; \mathrm{dt}) \\ &\Rightarrow \mathrm{LDM}_1(p) = \max(\mathrm{dt}_A(p), \mathrm{dt}_B(p)) = \mathrm{dt}_B(p) \\ \mathrm{LDM}_2(p) &= B(p)\mathrm{dt}_A(p) + A(p)\mathrm{dt}_B(p) \\ &= \mathrm{dt}_B(p) \\ \Rightarrow \mathrm{LDM}_1(p) &= \mathrm{LDM}_2(p) \end{aligned}$$

3. $A(p) = 1 \; and \; B(p) = 0$: same as 2).

□

This new equation is of big interest due to $\max$ and abs operators removal. Only linear operations remain.

3 Symbol recognition

Let's assume I is an image containing a symbol to be picked in a range of models. The key problem is to measure a similarity between I and M (the model image). The best model selection can then be done via the choice of the highest similarity measure. The LDM is representative of the dissimilarities between I and a model M, examples can be find in [1] or [2]. A global measure is computed from the sum of the LDM values :

$$\begin{aligned} \sum_p LDM_{I,M}(k,l) &= \sum_p [M(p)\mathrm{dt}_I(p) + I(p)\mathrm{dt}_M(p)] && (3) \\ &= \sum_p M(p)\mathrm{dt}_I(p) + \sum_p I(p)\mathrm{dt}_M(p). && (4) \end{aligned}$$

This equation can be interpreted as a "symetric chamfer-matching". The Chamfer-matching [3] minimizes a generalized distance between two sets of edge points. Each edge-model M is translated over $\mathrm{dt_I}$ the distance image of I. At each translation, the edge model M_t is superimposed on the distance image $\mathrm{dt_I}$. Then, the average of the distance values that edge model M_t hits gives Chamfer Score (CS). Thus for a given translation t:

$$\mathrm{CS}(I, M) = \frac{1}{N} \sum_p M(p)\mathrm{dt}_I(p). \quad (5)$$

The chamfer score $\mathrm{CS(I, M)}$ measures how the image is dissimilar to the model, or how the image looks like the model. By a comparison between eq. (4) and (5), it is clear that the global measure measures how M is similar to I, but also how M is similar to I. Tversky showed that similarity, from an human point of view, is asymetric [8]. For example, the judged similarity of North Korea to Red China is greather thant that of Red China to North Korea. The global measure of eq. (4) is symetric, but the introduction of weighting factors $\alpha, \beta \in [0, 1]$ leads to a versatile measure, the global dissimilary measure (GDM) :

$$\mathrm{GDM}(I,M) = \alpha\sum_p M(p)\mathrm{dt}_I(p) + \beta\sum_p I(p)\mathrm{dt}_M(p). \tag{6}$$

A symetric measure is obtained with $\alpha = \beta$. Asymetric ones are obtained with $\alpha \neq \beta$, such as $\alpha = 1, \beta = 0$ or $\alpha = 0, \beta = 1$. As suggested by [3] quadratic sums can produce more accurate results :

$$\mathrm{GDMq}(I,M) = \alpha\sum_p M(p)\mathrm{dt}_I^2(p) + \beta\sum_p I(p)\mathrm{dt}_M^2(p). \tag{7}$$

4 Rotation and scale invariant recognition

The proposed measure can only be applied on images with comparable sizes and orientations. In order to make rotation and scale invariant the recognition, we retained the scheme proposed in [4]. This method transform the rotation and scalings in translations thanks to a log-polar mapping. The scale and rotation parameters are estimated via a localisation of the transformed symbol in the transformed image. The images can then be geometrically corrected.

As in [4], assuming an image I is coded on a $N \times N$ grid, we use the following equations to resample the grid :

$$I_{\mathrm{p}}(\rho,\theta) = I(\rho\cos\theta + \frac{N}{2}, \rho\sin\theta + \frac{2}{2}) \tag{8}$$

$$I_{\mathrm{lp}}(m,k) = I_p\left(\frac{N/2-1}{2}(N-1)^{m/(N-1)}, \frac{2\pi k}{N} - \pi\right) \tag{9}$$

with $m \in [0, N-1]$ being the discrete radial coordinate and $k \in [0, N-1]$ being the discrete angular coordinate. I_{lp} is the log-polar transform of I. A nearest neighbour interpolation is used in order to obtain binary log-polar images.

Given two images I and M to be registrated, I_{lp} and M_{lp} are firstly computed. Thanks to the log-polar domain properties, it is only necessairy to localize I_{lp} in M_{lp} under translations. In other words, we need to estimate two integers (x, y) such that the translation of I_{lp} by (x, y) corresponds to M_{lp}. This is classicaly done using a cross-correlation product [4]. We show now how the GDM can be used for this purpose. The problem is to compute the similarity between M and $I_{(x,y)}$, that is the image with same size of M and centered at (x, y). This computation needs to be done for all positions (x, y) of I. If we keep the GDM as dissimilarity measure, the localisation is given by finding the minimum of :

$$L_{I,M}(x,y) = \mathrm{GDMq}(I_{(x,y)}, M). \tag{10}$$

This equation is rewritten with cross-correlation products ($\star$):

$$L_{I,M} = \mathrm{dt}_I^2 \star M + I \star \mathrm{dt}_M^2. \tag{11}$$

The cross-correlations lead to a very fast coding via the Fourier domain.

We compute $L_{I_{\mathrm{lp}},M_{\mathrm{lp}}}$ and localize its global minimum, providing two integers m and k. The rotation angle θ and the scaling factor σ are obtained with the following equations :

$$\theta = \frac{m - M/2}{M} + 1 \tag{12}$$

$$\sigma = (M-1)^{k/(M-1)} \text{ for } 0 \leq k < M/2 \text{ (enlargement)} \tag{13}$$

$$\sigma^{-1} = (M-1)^{(M-k)/M-1} \text{ for } M/2 < k < M \text{ (shrinkage)} \tag{14}$$

the model image M is then geometrically corrected from these parameters and compared to the image I.

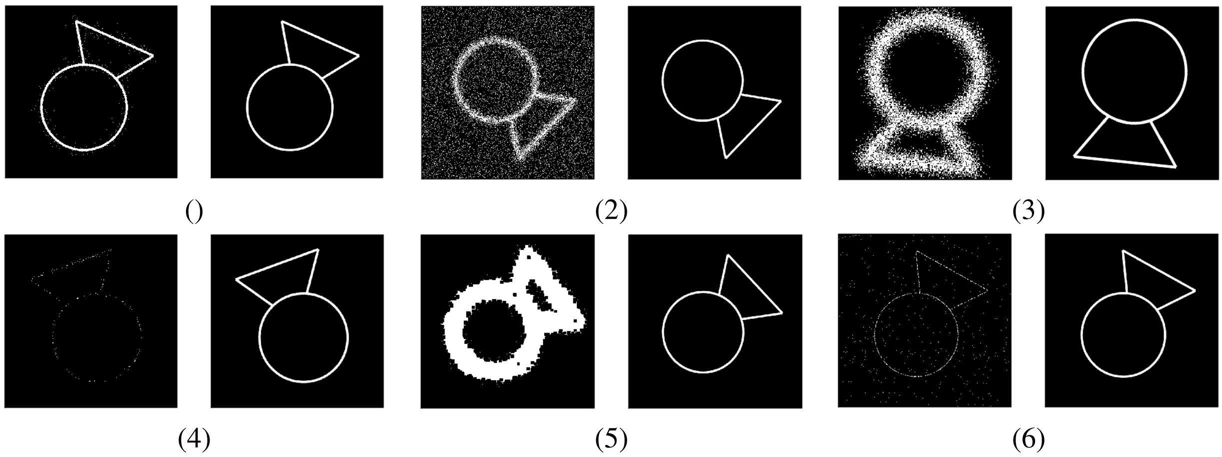

Figure 1: For each degradation model is shown the image (left) and its corresponding registrated model (right). The next step is the computation of the global dissimilarity between the two registrated images.

5 Final algorithm and results

5.1 Algorithm

The final algorithm is the following:

1. Compute I_{lp},log-polar representation of I, using eq. (8).
2. Foreach model M:
   (a) compute M_{lp}, log-polar representation of M
   (b) find the global minimum of $L_{I_{lp}}, M_{lp}$ (fast computation with eq. (11))
   (c) find (σ, θ) the scaling and rotation parameters of M (eq. (12)).
   (d) compute $M_{\text{corrected}}$from (ρ, θ), by applying reverse scale and rotation.
   (e) compute the dissimilarity $\text{GDMq}(I, M_{\text{corrected}})$ between $M_{\text{corrected}}$ and I (eq. (7)).
3. Keep the model with the lowest dissimilarity as winner.

5.2 Results

We have tested the proposed matcher on the grec2005 international symbol recognition contest tests. In this contest, six degradation models are proposed. Each test is available either with a rotation transformation or not, as well as either with a scaling transformation or not. In this research, we used only the binary representation of the symbols. An example of each degradation model is provided in fig. 1.

Complete tests have been conducted on the database with 25 symbols and 50 images. In order to check if a given model is corresponding to the input image, the global dissimilarity is computed with $\alpha = 1$, $\beta = 0$, leading to

$$\text{GDMq}(I, M) = \sum_p M(p)\text{dt}_I^2(p). \tag{15}$$

These parameters (α and β) are choosen in order to only take into account the information of I that belongs to the corresponding shape of M. Let's take an example from fig. 1.2. The image (on the left) is quite noisy. If we keep $\beta = 1$, all the noisy pixels impose to sum up a large number of the values of dt_M, even the ones localized far away from the model shape. The final value of the dissimilarity is thus always high, and will not allow to choose the right model. It is thus important to keep $\alpha = 1$ and $\beta = 0$.

	model 1	model 2	model 3	model 4	model 5	model 6
no rotation or scaling	100%	100%	100%	100%	100%	54%
with rotation and scaling	96%	40%	94%	70%	42%	12%

Table 1: Performances of the proposed matcher with respect to the 6 degradation models of the grec2005 tests. results obtained from 25 symbols and 50 images, when using 100 images and 150 symbols, the performances are lower by 5-10%, with an asymetric matcher ($\alpha = 1$, $\beta = 0$).

The final results are resumed in table 1. When no rotation nor scaling are present, the performance are very good. By example for model 1, it is reported that the best method obtains 100% without rotation/scaling and 88% with rotation/scaling (for the 25 symbols set) [5]. Our proposed matcher obtains 100% and 96%. The reader must remind that absolutely no hypothesis is made on the image or the symbol : no geometrical assumption, no denoising, no segmentation.

For the 6th degradation model, which is very destructive, the performances degrade severly. According to [5], the model 6 appears to be the more difficult to recognize in general. The report shows that the best recognition rate was only 59.81%. When rotation and scaling are involved, the performance depends of of the degradation (very good for models 1 and 3). A deep study of the behaviour of the matcher shows the main reason of bad performance is due to a bad registration.

6 Conclusion

We have proposed a symbol matcher based on the Local Dissimilarity Map between an image and a model. A fast computation is achieved by rewriting the matcher equation using intercorrelation products. The distinction between a symetric and an asymetric matching can be taken into account.

Tests were conducted from the grec2005 database with the raw matcher : no denoising, no segmentation. The performances are very good when no rotation nor scaling are involved, demonstrating the viability of the approach. We rotation and scaling are present, the performances are lower but still remain promising if not good.

We plan to work on :

- the removal the registration step by measuring the dissimilarity between the image and the model directly in the log-polar domain. This point should remove the critical step of parameter estimation. As the logpolar transformation modify the density of the grid, it will be necessary to take into account weighted factors depending on the localization of the grid.
- The application of gray-level LDM in order to allow comparison of gray-level images and log-polar correction via Fourier-Mellin transform.
- The application of the matcher to localize symbols in a bigger image.
- The application of some denoising and segmentation in order to boost the performances.

References

[1] E. Baudrier, N. Girard, J.-M. Ogier, ”A non-symmetrical method of image local-difference comparison for ancient impressions dating – Graphics Recognition”, *Recent Advances and New Opportunities: 7th International Workshop, GREC 2007*, Selected Papers, Curitiba, Brasil, 257–265, sept 2007.

[2] E. Baudrier, F. Nicolier, G. Millon, S. Ruan, ”Binary-image comparison with local-dissimilarity quantification”, *Pattern Recognition*, 41(5):1461–1478, 2008.

[3] G. Borgefors, ”Hierachical chamfer matching: a parametric edge matching algorithm”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):849–865, 1988.

[4] Q. Chen, M. Defrise, F. Deconinck, ”Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(12):1156-1168, 1996.

[5] P. Dosch, E. Valveny, *Report on the second symbol recognition contest*, in *Graphics Recognition, Ten Years Review and Future Perspectives*, Springer Berlin/Heidelberg, 381–397, 2006.

[6] D.P. Huttenlocher, W.J. Rucklidge, ”Comparing images using the hausdorff distance”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, 1993.

[7] J. Lladós, *Advances in Graphics Recognition*, in *Digital Document Processing*. Springer London, 281-303, 2007.

[8] A. Tversky, ”Features of similarity”, *Psychological Review*, 84(4):327–352, 1977.