Integration of Local and Global Shape Analysis for Logo Classification*

Jan Neumann, Hanan Samet and Aya Soffer**

Computer Science Department
Center for Automation Research
Institute for Advanced Computer Studies
University of Maryland
College Park, Maryland 20742
jneumann@cfar.umd.edu, hjs@cs.umd.edu and ayas@il.ibm.com

Abstract. A comparison is made of global and local methods for the shape analysis of logos in an image database. The qualities of the methods are judged by using the shape signatures to define a similarity metric on the logos. As representatives for the two classes of methods, we use the negative shape method which is based on local shape information and a wavelet-based method which makes use of global information. We apply both methods to images with different kinds of degradations and examine how a given degradation highlights the strengths and shortcomings of each method. Finally, we use these results to combine information from both methods and develop a new method which is based on the relative performances of the two methods.

Keywords: shape representation, shape recognition, image databases, symbol recognition, logos

1 Introduction

We examine three different approaches for classifying images with several components in an image database. One approach uses local methods to represent the image, the second uses global methods, while the third combines both using an adaptive weighting scheme based on relative performance. The local method uses so-called negative symbols, as described in [8], to compute a number of statistical and perceptual shape features for each connected component of an image and its background. The global method uses a wavelet decomposition of the horizontal and vertical projections of the global image as described in [5]. As a sample application of well-defined multi-component images, we use logos.

Several studies have reported results on some form of logo recognition. Each study used either global or local methods. These include local invariants [4, 7],

* The support of the National Science Foundation under Grants CDA-95-03994, IRI-97-12715, EIA-99-00268, and IIS-00-86162 is gratefully acknowledged.

** Currently at IBM Research Lab, Haifa 31905, Israel.

wavelet features [5], neural networks [3], and graphical distribution features [6]. The performance in case of certain degradations was examined.

In this paper we compare the local and global methods under the influence of several image degradations. The performance measure is the ranking of the original logo after inputing a degraded version of it into the classifier. The results exhibit the advantages and disadvantages of local methods, based on shape features, in contrast to global methods, rooted in signal processing. Finally, we present an algorithm that combines both methods into a single, robust framework by adaptively weighting the contributions of each method according to an estimate of their relative performance.

2 Preprocessing: Normalization of the images

The classification methods should be scale, translation, and rotation invariant. To achieve this, we apply some preprocessing steps to the input images before we start the computation of any features. The logos contained in the UMD-Logo-Database are gray-scale images that are scanned versions of black and white logos. Using an empirically determined preset threshold, we transform the input image into a binary image for which we compute its centroid.After shifting the image so that the centroid is located at the image center, which gives us translational invariance, we rotate the image around the centroid so that the major principal axis is aligned with the horizontal. This gives us rotational invariance. Finally, we resize the logo component so that its bounding box is a given percentage of the image size. This accounts for changes in scale of the input logos. These transformations make it possible to perform the following computations without reference to orientation, position, and scale.

3 The Wavelet Method

Given a normalized image we compute the horizontal and vertical projections of this binary image which are defined as $P(y) = \sum_{x=1}^{m} I(x,y)$ and $P(x) = \sum_{y=1}^{n} I(x,y)$. This means that we are counting the number of white pixels for each column and row. Next, we use a wavelet transform to apply a low-pass filter to the projections. In our experiments we used the Haar wavelet and the Daubechies wavelet s8 as implemented in the MATLAB wavelet toolbox and described in [9]. We do a 4-level Haar wavelet decomposition and for the 256x256 images that we used we get 16 low-pass coefficients per projection. In the case of the Haar wavelet this amounts to a repeated process of averaging and down-sampling. Finally, we end up with a 32-dimensional vector describing the logo as there are 16 coefficients for each of the two coordinate axes. This process is illustrated in Figure 1. These coefficient vectors, called *signatures*, are now used to compare different logos among each other. We use the L_1-Norm to compute the difference between their signatures, because the L_1-Norm is known to be robust against outliers and very fast to compute [9].

Proceedings of the 4th International Workshop on Visual Form (IWVF4) ,
Lecture Notes in Computer Science 2059, Springer, Berlin, 2001, pp. 769-778

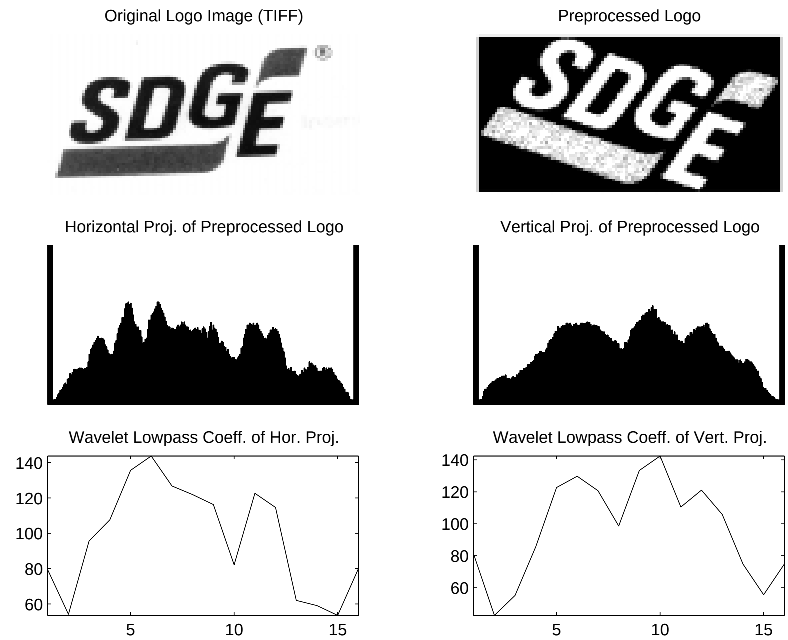


Fig. 1. The Wavelet signatures (from top-down, left-right): original image, normalized image, horizontal projection, vertical projection, low-pass wavelet coefficients of horizontal projection, low-pass wavelet coefficients of vertical projection (x-axis: index of coefficient, y-axis: coefficient magnitude).

4 The Negative Shape Method

The novel idea of the negative shape method as defined in [8] for the representation of symbol-like data such as found in logos is that we compute the shape features not just on the components of the foreground that constitute the symbol itself, but also on the components that make up the background of the image containing the symbol.

4.1 Choice of Shape Features

We start with the normalized images and do a connected component labeling of the image. For each component of the labeled image, we compute the following shape features:

1. **F1: Invariant moment:** The trace of the covariance matrix of the positions of the pixels that make up the logo, that is the sum of its diagonal entries.
2. **F2: Eccentricity:** The ratio between the length and width of the axis-aligned bounding box of the component after the normalization described in

Section 2. This gives us information about the extent of the elongation of a component.

3. **F3: Circularity:** The ratio between perimeter of the component and the perimeter of a circle of equivalent area: $CIRC = \frac{Perimeter^2}{4 \cdot \pi \cdot Area}$.
4. **F4: Rectangularity:** The ratio between the area of the component and the area of its bounding box.
5. **F5: Hole Area Ratio:** The ratio between the area of the holes inside the component and the area of the solid part of the component.
6. **F6,F7: Horizontal (Vertical) Gap Ratio:** The ratio of the square of the gap count to the area of the component where the gap count is defined as the number of pixels inside the component that have a right (bottom) neighbor that does not belong to the component.

4.2 The Classification Procedure

For the negative shape method we define the distance measure between two logos $Logo_1$ and $Logo_2$ as follows:

1. Normalize the value range for each element of the feature vector over all the logos of all the images in the dataset.
2. For each component of $Logo_1$ find the component of $Logo_2$ that has the smallest distance(L_2-norm) in feature space to it.
3. The average of these minimal distances over all the components of $Logo_1$ yields a measure for the distance between the two logos.

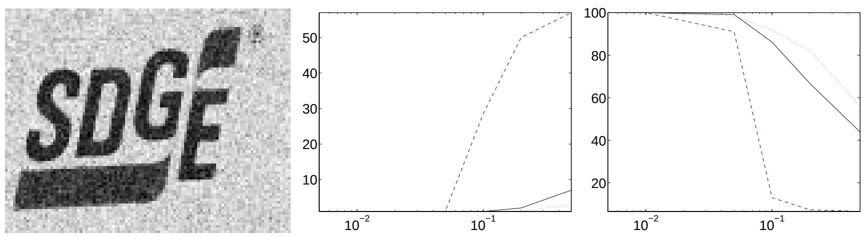
5 Comparison Between the Methods

All methods were implemented in Matlab^{TM} [1] and were applied to the logos contained in the UMD-Logo-Database (123 logos) [2]. The system was tested by providing it with an input logo and ranking the logos in the database based on their similarity to this logo. All methods always found the matching logo in the database. In particular, they ranked it first when the input logo is an uncorrupted version of one of the logos in the database. Below, we investigate the robustness of the methods when the logos are corrupted using four different image degradation methods as described in Figures 2a, 3a, 4a, and 5a. For each method, we degrade the images in the database to a varying degree, input them into the classifier, and then examine the rank (in terms of feature space distance) of the original, uncompromised logo. Here we examine the median of the rankings of the original logo over all the input logos (part b of all the figures) and how often in terms of the percent of all logos the original logo was ranked among the closest five of all logos (part c of all the figures). Each graph consists of three curves: the dashed curve corresponds to the negative shape method, the gray curve corresponds to the wavelet method, and the solid curve corresponds to the combined method which has not yet been described. The combined method was devised based on the results of these experiments and thus we defer its explanation and the analysis of the results using this method to the next section (i.e. Section 6) once we understand the pros and cons of the two methods.

Proceedings of the 4th International Workshop on Visual Form (IWVF4) ,
Lecture Notes in Computer Science 2059, Springer, Berlin, 2001, pp. 769-778

5.1 Additive Random Noise

To model the image degradation that is caused by processes such as fax transmissions or photo copying, we add Gaussian noise of zero mean and varying standard deviation (varying from 0.1 % to 50 % of the maximum possible pixel value of the image as indicated on the x-axis) to the gray-scale input images (e.g., Figure 2a).



(a) Example Image (b) Median Rank (c) Percentage of Top 5 Rankings

Fig. 2. Gaussian Noise: The x-axis denotes the standard deviation of the Gaussian noise with respect to the maximal pixel value of the original image. The dashed curves in (b) and (c) correspond to the negative shape method, the gray curves to the wavelet method, and the solid curves to the combined method.

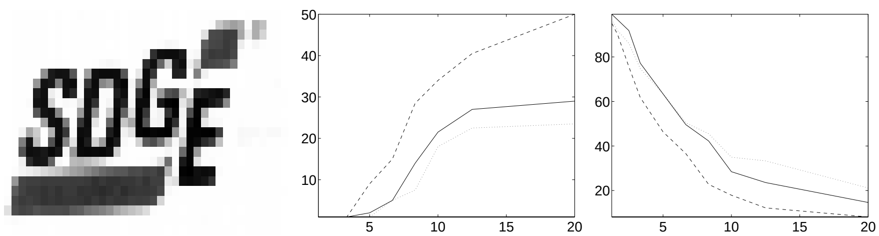
All the methods perform very well for small amounts of noise, but the wavelet method outperforms the negative shape noticeably (Figures 2b and 2c) for higher amounts of noise. Even when applying much noise (e.g., a standard deviation which is 20% of the possible pixel value), the average rank of the original logo is close to the top 10 (Figure 2b) and about 80% of the logos are ranked in the top 5 (Figure 2c). If we apply the negative shape method to such a heavily degraded image, the original logo is ranked in a nearly random manner (median rank 40th out of 123 logos as seen in Figure 2b) and the percentage of top 5 classifications is below 10% (Figure 2c).

It is to be expected that the wavelet method outperforms the negative shape method when adding random noise since the use of isotropic noise with an equal probability for adding or subtracting pixels should have only a small effect on the global histogram used in the wavelet method. We use noise of zero mean. Consequently, on the average, the distribution of white and black pixels in a row or column should not change much, and thus neither should the projection change much. On the other hand, in the negative shape method, we compute the feature vectors only on a small subset of pixels of each component. In this case, the noise will change the spatial distribution of the pixels more drastically because of the smaller number of pixels involved. Thus the negative shape method is less robust towards zero-mean Gaussian noise than the wavelet method.

Proceedings of the 4th International Workshop on Visual Form (IWVF4) ,
Lecture Notes in Computer Science 2059, Springer, Berlin, 2001, pp. 769-778

5.2 Reduced Resolution

To see how the methods handle differences in image resolution, which is obviously not offset by the scaling invariance since we work on digitized images, we reduce the size of the input images through sub-sampling using bilinear interpolation (e.g., Figure 3). The parameter value is the size ratio between the original and the sub-sampled image as indicated on the x-axis.



(a) Example Image (b) Median Rank (c) Percentage of Top 5 Rankings

Fig. 3. Reduced Resolution: The x-axis denotes the ratio between the size of the original and sub-sampled images.The dashed curves in (b) and (c) correspond to the negative shape method, the gray curves to the wavelet method, and the solid curves to the combined method.

As in Section 5.1, the wavelet method outperforms the negative shape method, although the negative shape method does not exhibit the same breakdown in performance as in the case of random noise. Since we use the low-pass wavelet coefficients for the classifier, the reduced resolution does not influence the performance of the wavelet method drastically. This is because sub-sampling an image by bilinear interpolation has a similar effect as low-pass filtering the image. The low-pass wavelet coefficients of a low-pass filtered image are in general very similar to the low-pass coefficients computed on the original image due to the fact that the low frequency components of the image are not affected noticeably by the sub-sampling operation. As before, the negative shape method is affected by this degradation because even when large scale changes are hardly visible, local shape features such as circularity, rectangularity and gap ratios are more susceptible to local changes due to a loss of detail.

5.3 Occlusion

To model the occlusion of parts of a logo, we add a component to the logo image which in this case is a black rectangle of varying size. The parameter here is the percentage of the image that is occluded by the rectangle (e.g., Figure 4a).

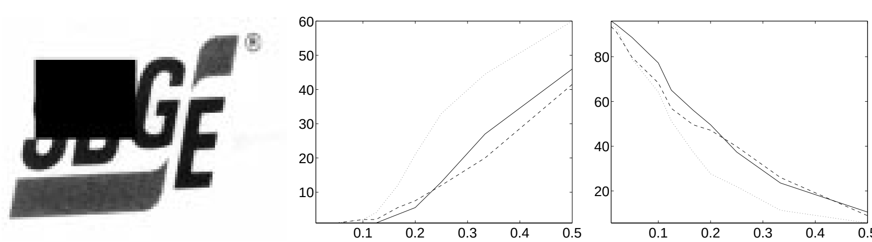
The performance graphs show that occlusion has a greater effect on the wavelet method than the negative shape method (Figures 4b and 4c) although both methods are able to handle small occlusions well.

Since the addition of an extra object or the omission of parts of the image causes global changes to the distribution of pixels in each row or column, the projections are strongly affected and thus so are the wavelet coefficients. Because of the local structure of the shape features, the components that are not occluded are not degraded at all and their feature values are unchanged. In the classifier we average the best feature vector matches for all the components in the input image. Since an occlusion is more likely to combine components into larger aggregates than to break them into many new ones, these few new components which do not have a corresponding component in the original image, are influencing the feature distance only to a small degree. Except for very degenerate configurations, the influence of the new components is averaged out by the continuing good matches of the feature vectors of the remaining uninfluenced components.

5.4 Swirling the Image

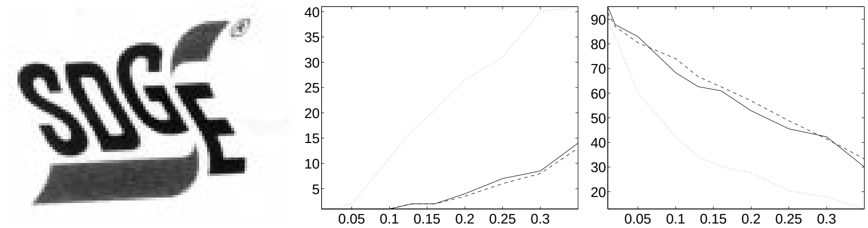
Swirling is a smooth deformation of an image which can be used to model a non-isotropic stretching of a logo. The relative position of each row is shifted to the left or right by an offset given by a smooth function, where the offset is limited to a certain percentage of the image width which is given as a parameter. This deforms the logo as if we would stretch a rubber sheet in different directions (e.g., Figure 5a).

This degradation has very different effects on the two methods. The performance of the wavelet method worsens rapidly with increasing swirl until we basically get a random ranking (our test size is 123, therefore, an average ranking of around 50 is nearly the expected median ranking for a logo that is not in the database). In contrast, the median rank of the original logo when using the negative shape method is lower than 10 (Figure 5b). It is possible to locally approximate this deformation as a combination of translations and rotations. The local features used by the negative shape method are rotation and translation



(a) Example Image (b) Median Rank (c) Percentage of Top 5 Rankings

Fig. 4. Occlusion of part of the image: The x-axis denotes the percentage of image area that is occluded. The dashed curves in (b) and (c) correspond to the negative shape method, the gray curves to the wavelet method, and the solid curves to the combined method.



(a) Example Image (b) Median Rank (c) Percentage of Top 5 Rankings

Fig. 5. Swirl of the image: The x-axis denotes the maximum horizontal displacement of an image row in percentage of image width. The dashed curves in (b) and (c) correspond to the negative shape method, the gray curves to the wavelet method, and the solid curves to the combined method.

invariant due to the component normalization. Therefore, it is much less affected by this degradation than the wavelet method. Recall that the wavelet method is only globally rotation and translation invariant due to the global preprocessing, but not locally.

6 Combination of both Methods

In Section 5 we saw that the wavelet and the negative shape methods perform very differently if the input logo is corrupted by either local or global degradations. To take advantage of the respective strengths of both methods we devised the following performance-dependent weighting scheme. First, for each undegraded logo l in the dataset we compute the average feature space distance of l to all other logos for both the wavelet and the negative shape methods. This is followed by calculating the average of these average distances for the two methods which we denote by A_w for the wavelet method and A_s for the negative shape method. We define the ratio between these two averages (i.e. $\frac{A_w}{A_s}$) to be the expected ratio E for the two methods. We determined how this ratio changed when we applied both methods to degraded inputs. The understanding of this relationship between the change in ratio and the relative performance of the two methods when applied to degraded images enabled us to adaptively weight the respective contributions of the two methods when combining them into a single distance measure. The relative weights are based on the change in the ratio because a a large increase of the feature space distance for one method compared to the other indicates a breakdown in its performance.

When classifying an input logo which has been degraded using one of the processes described in Section 5, we first compute the feature distances of this logo to all the other logos for the wavelet method which we denote by W and for the negative shape method which we denote by S. In addition, we define the averages of W and S over the whole dataset by D_w and D_s, respectively. Next, we compare the ratio between D_w and D_s (i.e., $\frac{D_w}{D_s}$) to the expected ratio between W and S which we assume to be similar to the precomputed value E. If the difference in the ratios indicates that one of the two methods is performing worse than expected, we decrease its weight in the final classification and increase the weight of the other method. The combined feature distance C for a single degraded input logo is a weighted sum of the wavelet method feature distance W and negative shape method feature distance S:

$$C = \frac{E \cdot D_s}{D_w} \cdot \frac{W}{E} + S \tag{1}$$

The factor E, that describes the average ratio between W and S, is only included in order to facilitate understanding the rationale behind the final weighting method. If we divide W by E, then we effectively normalize W, so that its magnitude is equal to the magnitude of S. Thus, if the ratio $\frac{D_w}{D_s}$ equals the expected ratio E, then we believe that both methods will perform well and we use an approximately equal weighting of the two feature distances W and S. If now the ratio $\frac{D_w}{D_s}$ either grows larger (smaller) than E because the degradation of the input logo causes the wavelet method to compute feature distances larger (smaller) than the negative shape method (up to the expected ratio E), then the contribution of W in equation 1 will be reduced (increased) because we have less (more) confidence in the wavelet method's ability to classify the input logo correctly.

This adaptive weighting scheme increases the robustness of the classification noticeably. When we examine the performance criteria in Section 5, we see that the combined method is able to capture the different behavior of the methods and adapts its weights accordingly. Comparing the performance of the combined method on images degraded as described in Section 5.1 and Section 5.4 where the wavelet and the negative shape method exhibit very different performances, we see that our weighting scheme is able to detect the change in relative performance and adjust the weights to mimic the classification of the better performing method. For the degradations described in Sections 5.2 and 5.3 where the performance difference between the two basic methods is not as pronounced, the combined method lags slightly behind the better performing method in the median rank criterion (part (b) of all the Figures), but equals or surpasses the performance of the better method in terms of the other criterion (part (c) of all the Figures). This shows that our combined scheme is effective in capturing global as well as local shape information and is thus able to deal well with the image degradations of the kind that we described.

7 Summary and Future Work

Both the wavelet as well as the negative shape method are well-suited for certain kinds of image degradations but are very sensitive to others. This discrepancy in performance can be explained by the difference between local shape feature-based and global, filter-based methods. On the one hand, we have the wavelet method that operates on the global image and computes features that are relatively invariant to degradations that are isotropic. On the other hand, we have the negative shape method which operates on local image regions. Thus its features are relatively invariant to changes that leave the image at other locations mostly intact such as occlusions or preserve the local image structure such as the swirl deformation. We take advantage of the fact that both basic methods perform very differently on images that exhibit degradations of either local or global nature by devising a performance-dependent weighting scheme that combines the results of both methods. Our combined algorithm shows a noticeable improvement in the robustness of the classification by combining the strengths and avoiding the weaknesses of the respective methods. This weighting scheme performs the better the more different the performances of the underlying methods are because this makes it easier to detect if one method is performing poorly with respect to the other method. Therefore, the wavelet and the negative shape methods are very well-suited to be combined by a performance-dependent weighting scheme.

For future work it is planned to improve the synergy between the two methods by using local image information to estimate how much an image region is degraded and then use this locality information to adaptively weigh the feature vectors on the component level.

References

1. Matlab - the language of technical computing. www.mathworks.com.
2. Umd logo database. http://documents.cfar.umd.edu/resources/database/UMDlogo.html.
3. F. Cesarini, E. Fracesconi, M. Gori, S. Marinai, J. Q. Sheng, and G. Soda. A neural-based architecture for spot-noisy logo recognition. In *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, pages 175–179, Ulm, Germany, August 1997.
4. D. Doermann, E. Rivlin, and I. Weiss. Applying algebraic and differential invariants for logo recognition. *Machine Vision and Applications*, 9(2):73–86, 1996.
5. M. Y. Jaismha. Wavelet features for similarity based retrieval of logo images. In *Proceedings of the SPIE, Document Recognition III*, volume 2660, pages 89–100, San Jose, CA, January 1996.
6. T. Kato. Database architecture for content-based image retrieval. In *Proceedings of the SPIE, Image Storage and Retrieval Systems*, volume 1662, pages 112–123, San Jose, CA, February 1992.
7. M. Kliot and E. Rivlin. Shape retrieval in pictorial databases via geometric features. Technical Report CIS9701, Technion - IIT, Computer Science Department, Haifa, Israel, 1997.
8. A. Soffer and H. Samet. Using negative shape features for logo similarity matching. In *Proceedings of the 14th International Conference on Pattern Recognition*, pages 571–573, Brisbane, Australia, August 1998.
9. G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley-Cambridge Press, Wellesley, MA, 1996.