Model-Based Shape Recognition Using Recursive Mathematical Morphology

Michael Schauf, Selim Aksoy, Robert M. Haralick
Intelligent Systems Laboratory
Department of Electrical Engineering
University of Washington
Seattle, WA, 98195-2500
{mschauf,aksoy,haralick}@isl.ee.washington.edu

Abstract

This paper introduces a size invariant method to recognize two-dimensional binary shapes using the recursive erosion transform. Using recursive morphology with multiple structuring elements, the method takes constant time per pixel regardless of the scale of the shape model, and also works on noisy images without requiring noise removal. Results from experiments on 100 noisy images show the methodology is able to detect every shape model's scale and position with 13 false alarms and five misdetections out of 254 total translated and scaled models.

1. Introduction

The ultimate goal of binary shape recognition is to imitate the human ability to recognize a variety of different shapes in an image. A majority of the algorithms fall under the areas of boundary [2], feature [1], parts [6], and skeleton representations [7]. Many of the algorithms fail with the introduction of noise. Others are very complex, requiring large amounts of memory or processing time. The algorithm in this paper works with noisy images and requires only a small amount of memory and processing time.

The problem is defined such that a set of ideal, noise free, binary shape models is given. The input is a noisy image containing a handful of scaled shape models and the goal is to determine which shape models are placed in the noisy image and to determine their origin and scale.

2. Recursive Erosion Transform

The erosion of image A by a structuring element B is denoted by $A \ominus B$ and is defined

$$A \ominus B = \{x \in E^n | \cap_{b \in B} A_{-b}\} \tag{1}$$

In equation 1, each element a and b of A and B is the row and column coordinates of foreground pixels in the image and structuring element respectively. The operation A_{-b} is the operation of subtracting element b of B from every element a of A [4], [3].

The recursive morphology definition of the erosion transform breaks the structuring element into two parts, before the origin in top-down left-right scan order called Y, and after called Z. The function $f(x)$ equals the value at location x in image A and equals zero outside the image's width and column dimensions. The functions g and h act upon the coordinate elements in Y and Z respectively and are initialized at -1 for all elements. The erosion transform is then defined by a two pass forward, $(f \ominus_F b)(x)$, and backward, $(f \ominus_B b)(x)$, process.

$$(f \ominus_F b)(x) = \begin{cases} 0 & f(x) = 0 \\ \min\{\infty, [(f \ominus_F b) \ominus g](x)\} & \text{otherwise} \end{cases} \tag{2}$$

$$(f \ominus_B b)(x) = \min\{(f \ominus_F b)(x), [(f \ominus_B b) \ominus h](x)\} \tag{3}$$

$$(f \ominus g)(x) = \min_{y \in Y}\{f(x+y) - g(y)\} \tag{4}$$

$$(f \ominus h)(x) = \min_{z \in Z}\{f(x+z) - h(z)\} \tag{5}$$

3 Methodology

The basis of the methodology is the property that the erosion transform of any scaled shape model results in a scaled version of the original model's erosion transform. By this property, the values of the original shape model's erosion transform and the scaled model's erosion transform, as well as each values' coordinates, are related by a scaling factor. Therefore, corresponding original model values and coordinates and translated and scaled model values and coordinates are related by the following equations

$$q_m = sp_m \tag{6}$$

$$u_m = sr_m + t_r \tag{7}$$

$$v_m = sc_m + t_c \tag{8}$$

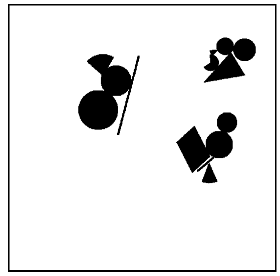
where r_m,c_m, and p_m are the row, column and erosion transform values of the original model, u_m,v_m, and q_m are the row, column, and erosion transform values of the translated and scaled model, s is the scale, t_r is the row translation, and t_c is column translation. (r_m, c_m, p_m) will be called model triples and (u_m, v_m, q_m) will be called image triples. Solving for s, t_r, and t_c results in

$$s = \frac{p_m}{q_m} \tag{9}$$

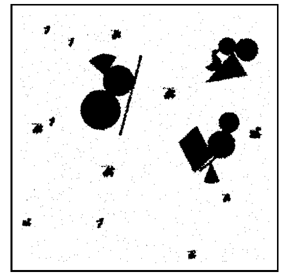
$$t_r = u_m - \frac{p_m r_m}{q_m} \tag{10}$$

$$t_c = v_m - \frac{p_m c_m}{q_m} \tag{11}$$

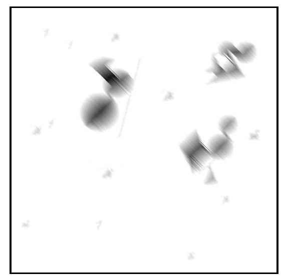
Using structuring elements spanning angles over 360 degrees results in different transforms which provide enough information, some not changed much by noise. Also, any noise around the boundaries of the foreground of an object has only slight effects on values within the boundary of the erosion transform.



(a) Image



(b) Noisy Image



(c) Erosion Transform of Noisy Image

Figure 1. Example Image

4 Algorithm

The algorithm contains both an offline and online part. The offline algorithm begins by calculating the sixteen erosion transforms for a single model. The sixteen structuring elements can be seen in Figure 2, Once these have been calculated, a list of triples is generated containing the relative maximum values in the shape model. This single relative maximum list is separated into smaller lists containing only those relative maxima that are closely connected. Once the smaller lists are generated, the center pixel for each list is determined and stored in another list. This pre-processing is done for all models.

The online part of the algorithm proceeds by first obtaining the image triple containing the global maximum erosion transform value in the image and the triple containing the global maximum erosion transform value in one of the shape models. For the moment it is assumed these are corresponding triples and the scale and translation are calculated by equations 9, 10, and 11. The model's existence at this location and scale is then verified by scaling and translating all the offline triples for that model using equations 6, 7, and 8 and searching in a three by three pixel square for an erosion transform value within one of the expected erosion transform value. If for a structuring element a given ratio of the triples match, the model matches for that structuring element. If a given ratio of the structuring elements match, the model exists at that translation and scale. Once a model is detected, the bounding box of the scaled model in the image is calculated by scaling the model row and column dimensions by the estimated scale, and the model is masked out of the image. If the model fails verification, the steps are repeated for all other models and the scale and translation is calculated using all erosion transforms. All these steps are repeated until the ratio of the number of foreground pixels over the background is less than five percent of the original ratio or the method does not detect another model using all models and scale and translation estimates.

5 Testing and Results

Testing the feasibility of the method required the generation of primitives, shape models, and noisy images containing translated shape models. A shape model is composed of a set of primitives which are circles, sectors, lines, triangles and quadrilaterals. For this feasibilty study the model size is 256 by 256. The shape model is constrained such

Table 1. Confusion matrix for the tests with noise

	Assigned Models										Mis.	Avg.Dist.	Avg.Scale
Groundtruthed Models	32	0	0	0	0	0	0	0	0	0	0	2.6939	0.0109
	0	28	0	0	0	0	0	0	0	0	0	2.2733	0.0081
	0	0	30	0	0	0	0	0	0	0	3	3.3781	0.0124
	0	0	0	30	0	0	0	0	0	0	0	3.1621	0.0120
	0	0	0	0	16	0	0	0	0	0	0	4.1056	0.0191
	0	0	0	0	0	22	0	0	0	0	0	2.7724	0.0111
	0	0	0	0	0	0	20	0	0	0	0	2.1297	0.0117
	0	0	0	0	0	0	0	21	0	0	0	1.4366	0.0054
	0	0	0	0	0	0	0	0	27	0	1	1.9564	0.0073
	0	0	0	0	0	0	0	0	0	28	1	2.5694	0.0119
False Alarm	0	0	0	0	0	11	2	0	0	0			

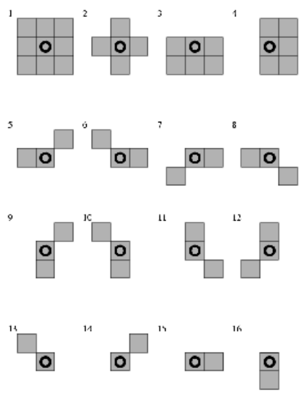


Figure 2. The sixteen structuring elements

that each primitive overlaps by less than 5%. Image generation requires scaling the shape model and randomly placing them in the image so they do not overlap with any other shape model already in the image. For this feasibility study, the image size is 512 by 512 and the scales are randomly generated between 0.3 and 1.0.

In testing, we constructed 10 different shape models, Each consisting of randomly generated primitives placed at random positions, From these 10 models, 100 different images were generated. The 100 images were also perturbed with noise by adding small irrelevant models, by adding noise to the boundaries of the shape models, and by adding pepper noise to the background [5]. The noise around the boundary of the object and the pepper noise was generated using 420 for the random seed $c_0 = 0.002$, $\alpha_0 = 0.7$, $\alpha = 0.7$, $\beta_0 = 0.7$, $\beta = 0.7$ and $stElSize = 0$. Irrelevant models were added by generating two additional models and scaling them between .08 and .10. This was done between two to four times. Figure 3 shows three of the 10 shape models generated an this experiment, and Figure 1, shows a generated image with and without noise. Tests were run on images with and without noise, and the results are summarized in Table 1.

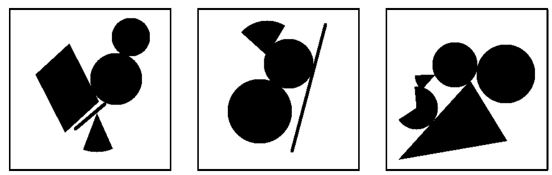


Figure 3. Three randomly generated models

6 Conclusion

This paper presented a successful algorithm to detect multiple scaled shape models in a noisy image. It was accomplished by using the recursive erosion transform since the erosion transform of any scaled shape model results in a scaled version of the original erosion transform. The feasibility tests show the method works well with the addition of noise. Future work includes the application of probabilistic methods to evaluate the errors caused by noise.

References

[1] F. Alt. Digital Pattern Recognition by Moments. *Journal of the ACM*, 11:240–258, 1962.

[2] H. Freeman. On the Encoding of Arbitrary Geometric Configurations. *IRE Transactions on Electric Computers*, EC-10:260–268, 1961.

[3] C. Giardina and E. Dougherty. *Morphological Methods in Image and Signal Processing*. Prentice-Hall, 1988.

[4] R. Haralick and L. Shapiro. *Computer and Robot Vision*, volume 1. Addison-Wesley, 1992.

[5] T. Kanungo. DDM Users Manual. Technical report, University of Washington, Seattle, WA, August 1993.

[6] L. Shapiro. A Structural Model of Shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2:111–126, 1980.

[7] P. Trahanias. Binary Shape Recognition Using The Morphological Skeleton Transform. *Pattern Recognition*, 25(11):1277–288, 1992.