ROBUST MULTISCALE TRIANGLE-AREA REPRESENTATION FOR 2D SHAPES

Ibrahim El Rube[a], Naif Alajlan [b], Mohamed Kamel[b], Maher Ahmed[c], and George Freeman[b]

[a] Systems Design Engineering, University of Waterloo, Waterloo, N2L 3G1, Ontario, Canada.
[b] Electrical and Computer Engineering, University of Waterloo, Waterloo, N2L 3G1, Ontario, Canada.
[c] Physics and Computer Science Department, Wilfrid Laurier University, Waterloo, N2L 3C5, Ontario, Canada.

ABSTRACT

In this paper, a new 2D shape Multiscale Triangle-Area Representation (MTAR) is proposed. This representation utilizes a simple geometric principle, the area of a triangle, in obtaining a robust and efficient shape representation. The use of the wavelet transform for decomposing the boundary of the shapes improves the efficiency and robustness of the representation. The MTAR is more robust to the affine transformation, less affected by noise, and more selective than similar methods, e.g., the curvature scale-space CSS. Two tests, using MPEG-7 CE-shape-1 database, show that MTAR achieves better performance than the CSS under affine transformation and in the general shape retrieval.

1. INTRODUCTION

Shape representation is a crucial step in shape analysis and matching systems. The complexity and the performance of the subsequent steps in shape analysis systems are largely dependent on the invariance, robustness, stability, and uniqueness of the applied shape representation method.

In the past decade, different techniques were adopted for shape representation and matching. Curvature scale space (CSS) [10],[1] and [8], fuzzy-based matching [6], dynamic programming [12], shape contexts [3], shock graphs [14], geodesic paths [7], Fourier descriptors [4], and wavelet descriptors [5] are examples of these techniques.

The multi-scale approach for shape representation and matching is considered the most promising. It can be argued that human perception of shapes is a multi-scale by nature. In addition, many interesting shape properties are revealed at different scale levels. Another advantage includes its invariance to moderate amounts of deformations and noise.

In this paper, MTAR is proposed as a new shape representation for closed boundaries. The MTAR enjoys many advantages over the curvature scale space (CSS) representation [10], which has been selected for MPEG-7 standardization after comprehensive comparative experiments with other methods [11]. Unlike the CSS representation where just the zero crossing points are invariant to affine transformations, the MTAR is totally invariant to affine transformations. In addition, the MTAR is more robust to noise and it provides selectivity in the matching process, that is, coarse-to-fine matching.

2. RELATED WORK

One of the most well-researched closed-contour shape representations is the curvature scale space (CSS) method proposed by Mokhtarian and Mackworth [9], [10]. A Gaussian kernel with increasing standard deviation σ is used to gradually smooth the contour at different scale levels. At each scale, the curvature of each contour point is measured by:

$$c(u,\sigma) = \frac{\dot{x}(u,\sigma)\ddot{y}(u,\sigma) - \ddot{x}(u,\sigma)\dot{y}(u,\sigma)}{(\dot{x}(u,\sigma)^2 + \dot{y}(u,\sigma)^2)^{\frac{3}{2}}} \quad (1)$$

Where c is the curvature at location u and scale σ, $\dot{x}$ and $\ddot{x}$ are the first and second derivatives of x, respectively. By setting (1) to zero, the inflection points (curvature zero crossings) are located at each scale which results in an image, called CSS image (see Fig. 1). This image shows the end-points of the concave segments along the contour (the horizontal axis) at each scaling iteration (the vertical axis). As the scale level increases, the smoothing effect increases and the number of inflection points decreases until the contour becomes totally convex.

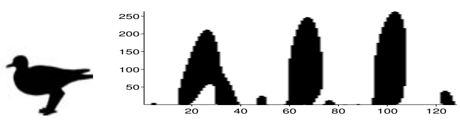

Fig. 1. CSS image for a bird shape.

Only the maxima of the CSS images' contours are used for matching two shapes [2]. Many heuristics are employed to find the best correspondence efficiently. This method exhibits certain degree of robustness to affine transformation [1]. The main limitations of the CSS method include its limited representation to only the concave segments, its

0-7803-9134-9/05/$20.00 ©2005 IEEE

failure to discriminate a shallow concavity from a deep one, and that it requires a large number of smoothing iterations (scales) to obtain the CSS image (may exceed 100).

An efficient computation of the curvature scale space representation using B-spline wavelets was proposed by Wang *et al.* [16], [15]. Their method provides an alternative to the classical Gaussian-based scale space representation while being much more efficient and relying on the well-established wavelet theory.

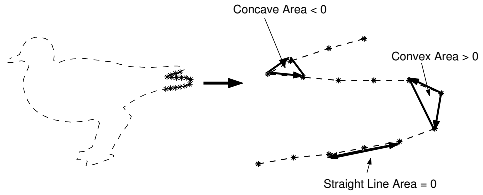

Fig. 2. Illustration of TAR.

3. MULTI-SCALE TRIANGLE-AREA REPRESENTATION (MTAR)

In this section, the MTAR is introduced as a closed-boundary representation. The MTAR has many desirable properties over other representations such as compactness, robustness to moderate noise and deformations, robustness to the affine transformations, provision for efficient matching, and handling partial occlusions. In addition, MTAR provides flexible coarse-to-fine matching. In the following, a brief explanation of how to obtain the MTAR images for an arbitrary closed contour is introduced.

3.1. MTAR Images

The MTAR starts by extracting the external contour of the shape of interest from the 2D image. Each contour point is represented by its x and y coordinates. Then, separated parameterized contour sequences $x(n)$ and $y(n)$ are obtained into order to facilitate applying 1D techniques to each sequence alone.

The contour is re-sampled to N points and the curvature of each point is measured using the triangle area representation (TAR). For each three equidistance points P_1, P_2 and P_3, the area A of the resultant triangle is given by:

$$A = \frac{1}{2} \begin{vmatrix} P_{x_1} & P_{y_1} & 1 \\ P_{x_2} & P_{y_2} & 1 \\ P_{x_3} & P_{y_3} & 1 \end{vmatrix} \quad (2)$$

This process is demonstrated in Fig. 2. When the contour is traversed in CCW direction, positive, negative and zero values of A mean convex, concave and straight-line points, respectively. By Increasing the length of the triangle sides, i.e., considering farther points, A will represent longer variations along the contour. A TAR image is obtained by thresholding A at zero and taking the locations of the negative values. So, it shows the concavities along the contour at increasing values of triangle sides, as shown in Fig. 3 (top). Thus, the horizontal axis in a TAR image shows the locations of the contour points and the vertical axis represents the triangle side length.

The TAR image, like the CSS image, is affected by noise. To reduce this effect, the dyadic wavelet transform is applied to each contour sequence at various scale levels, as shown in Fig. 3 (left column). The wavelet approximation coefficients are adopted in this paper to reduce the noise effect on the shape boundary. At each wavelet scale level a TAR image is computed. This results in $L + 1$ TAR images, where $L = log_2(N)$ and represent the number of the wavelet scale levels. Fig. 3 shows the first six MTAR images of the bird shape.

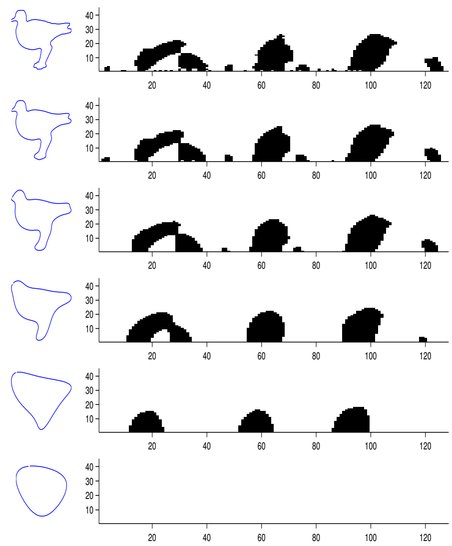

Fig. 3. MTAR of a bird shape.

A highly desirable property of the MTAR is the selectivity, which facilitates progressive or coarse-to-fine matching. Also, High-scale MTAR images can be used to eliminate dissimilar shapes and low-scale images can discriminate between similar shapes.

3.2. Affine Transformations and MTAR

The advantage of using (2) is that it is invariant to the affine transformations, which makes it more suitable for real-world imaging applications. A general affine transformation applied to a shape is given by:

$$\begin{bmatrix} P_x^t \\ P_y^t \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} P_x \\ P_y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} \tag{3}$$

where P_x^t and P_y^t are the transformed coordinates of P_x and P_y, respectively, e and f represent translation and a, b, c and d reflect scale, rotation and shear. By substituting (3) into (2), we obtain:

$$A^t = (ad - bc)\, A \tag{4}$$

where A^t is the affine transformed version of A. It is clear that A^t, and hence MTAR, is relatively invariant to affine transformations. Absolute invariance can be achieved by dividing A^t by its maximum value.

4. MATCHING

There are some resemblances between an MTAR image and CSS image produced for the same shape. They both describe concavities along the contour. However, the CSS method measures the curvature as the contour is smoothed using convolution with Gaussian kernels with different widths. On the contrary, each MTAR image represents the locations of the concavities using different triangle sides at a specific wavelet-smoothed scale level. Therefore, we followed a similar approach to that of CSS [13] in order to match two MTAR image sets of two shapes. In their method, matching is performed in two stages as follows.

In the first stage, a set of global features are used to eliminate very dissimilar shapes and exclude them from further processing. In our work, these features include aspect ration AR, circularity C, eccentricity E and solidity S. In the second stage, a dissimilarity measure (D_s) between each two MTAR images of the two shapes at certain scale is computed as described in [2]. D_s is based on finding a number of initial nodes or correspondences between two sets of maxima in the MTAR images using only two maxima in each image. Then, the lowest cost node is extended to include all other maxima and its cost is considered as D_s. However, we expand this approach to extend the two initial nodes with lowest costs and consider their minimum as D_s. This modification enhances the accuracy with a slight increase in computations.

The dissimilarity measure between each two MTAR images at a specific scale is a weighted sum of the global parameters and D_s. Consequently, the final dissimilarity distance between two shapes is a weighted sum of the dissimilarity measures between their MTAR images. Practically, the MTAR images at scale zero (noise sensitive) and the last four scales (less informative) are avoided in calculations. Therefore, only four MTAR images are computed which reduced the computational complexity by more than a half. Note that MTAR images require less iterations to compute than CSS images.

5. EXPERIMENTAL RESULTS

In order to test the proposed MTAR, the MPEG-7 CE-shape-1 dataset is chosen for the experiment. This dataset consists of 1400 shapes grouped in 70 classes. These shapes were taken from: natural objects, man-made objects, objects extracted from cartoons, and manually drawn objects.

Two experiments are conducted in this section to investigate the performance of the MTAR under the affine transformation distortions and to evaluate its retrieval performance as compared to the CSS method. In both experiments, the same matching algorithm is used for both methods to have a fair comparison of the two representations.

5.1. Affine Invariance test

In order to test the MTAR and CSS against the affine transformation distortions, the original 2D images of 70 shapes representing the MPEG7 dataset groups are transformed using (3). The parameters used to obtain these shapes are $b = [0, 0.4, 0.8, 1.5, 2, 3, 4, 6]$. Fig. 4 shows a sample of these distorted shapes. The performances of the MTAR and CSS are evaluated using the precision-recall curves as shown in Fig. 5. These curves are plotted for the CSS, the MTAR at the second scale level, the MTAR at the third scale level, and the combined MTAR. It can be easily noticed that the second and third scale levels are comparable to the CSS, whereas the MTAR that combines all the four levels outperforms the CSS.

Fig. 4. Affine distorted shapes

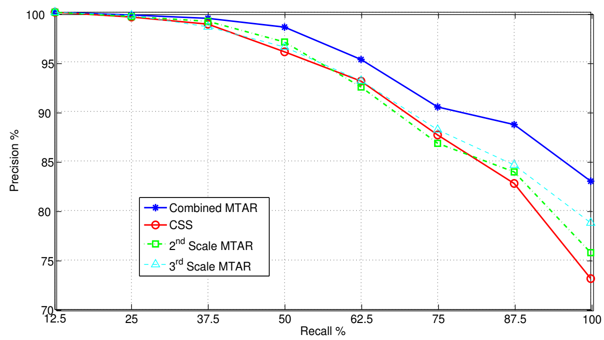

Fig. 5. Affine test precision-recall curves.

5.2. Shape Retrieval test

In this test, the retrieval performances of MTAR- and CSS-based systems are compared. In addition, the performance of MTAR using only the second scale level is also evaluated. For this purpose, the precision-recall pairs are computed for the three systems using all 1400 shapes of the MPEG-7 database. Fig. 6 shows that MTAR achieves higher accuracy than CSS at all recall values. It also shows that the performance of the second scale of MTAR is comparable to the CSS. Thus, MTAR outperforms CSS in retrieval performance.

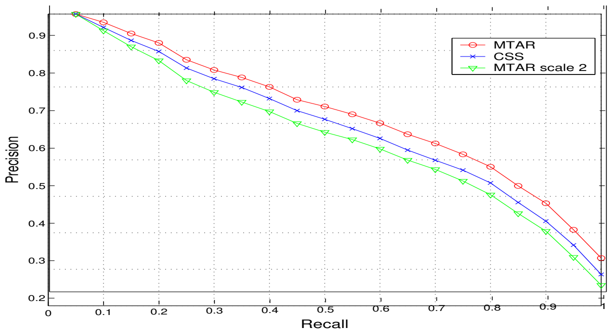

Fig. 6. Retrieval test precision-recall curves.

6. CONCLUSIONS

In this paper, a new multiscale shape representation is introduced. The representation utilizes the area of the triangles formed by each three equally apart points on the shape boundary. The MTAR images obtained from this representation are used in shape matching and retrieval. The results of two tests show that the proposed MTAR outperforms the CSS method in terms of both the robustness to the affine transformations and the retrieval accuracy.

7. REFERENCES

[1] S. Abbasi and F. Mokhtarian, *Affine-similar shape retrieval: application to multiview 3-d object recognition*, IEEE Trans. on Image Processing **10** (2001), no. 1, 131–139.

[2] S. Abbasi, F. Mokhtarian, and J. Kittle, *Curvature scale space image in shape similarity retrieval*, Multi-Media Systems **7** (1999), no. 6, 467–476.

[3] S. Belongie, J. Malik, and J. Puzicha, *Shape matching and object recognition using shape contexts.*, IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002), no. 24, 509–522.

[4] T. Bui and G. Chen, *Multiresolution moment-wavelet-fourier descriptor for 2-*D *pattern recognition*, Proceedings of SPIE-The International Society for Optical Engineering **3078** (1997), 552–557.

[5] G. Chauang and C. Kuo, *Wavelet descriptor of planar curves: Theory and applications*, IEEE Transaction on Image Processing **5** (1996), no. 1, 56–70.

[6] Y. Chen and J. Wang, *A region-based fuzzy feature matching approach to content-based image retrieval.*, IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002), no. 9, 1252–1267.

[7] E. Klassen, A. Srivastava, W. Mio, and S. Joshi, *Analysis of planar shapes using geodesic paths on shape spaces.*, IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004), no. 3, 372–383.

[8] F. Mokhtarian and M. Bober, *Curvature scale space representation: Theory, applications, and mpeg-7 standardization*, Kluwer Academic Publishers, 2003.

[9] F. Mokhtarian and A. Mackworth, *Scale-based description and recognition of planar curves and two-dimensional shapes*, IEEE Trans. on PAMI **8** (1986), no. 1, 34–43.

[10] ______, *A theory of multi-scale, curvature-based shape representation for planar curves*, IEEE Trans. on PAMI **14** (1992), no. 8, 789–805.

[11] The MPEG Home Page, *http://www.chiariglione.org/mpeg/index.htm*.

[12] E.G.M. Petrakis, A. Diplaros, and E. Milios, *Matching and retrieval of distorted and occluded shapes using dynamic programming*, IEEE Trans. on PAMI **24** (2002), no. 11, 1501–1516.

[13] W. K. Pratt, *Digital image processing*, second ed., John Wiley and sons Inc, 1991.

[14] T. Sebastian, P. Klein, and B. Kimia, *Recognition of shapes by editing their shock graphs.*, IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004), no. 5, 550–571.

[15] Yu-Ping Wang and S. L. Lee, *Scale-space derived from b-splines*, IEEE Trans. on PAMI **20** (1998), no. 10, 1040–1055.

[16] Yu-Ping Wang, S. L. Lee, and K. Toraichi, *Multiscale curvature-based shape representation using b-spline wavelets*, IEEE Trans. on Image Processing **8** (1999), no. 11, 1586–1592.