NEW AREA MATRIX-BASED AFFINE-INVARIANT SHAPE FEATURES AND SIMILARITY METRICS

Carlos R. P. Dionisio and Hae Yong Kim

Escola Politécnica, Universidade de São Paulo, Brazil
{carlos,hae}@lps.usp.br

ABSTRACT

A near-planar object seen from different viewpoints results in differently deformed images. Under some assumptions, viewpoint changes can be modeled by affine transformations. Shape features that are affine-invariant (af-in) must remain constant with the changes of the viewpoint. Similarly, shape similarity metrics that are af-in must rate the difference between two shapes, regardless of their viewpoints. Af-in shape features and similarity metrics can be used for the shape classification and retrieval. In this paper, we propose a new set of af-in shape features and similarity metrics. They are based on the area matrix, a structure that contains multiscale information about the shape. Experimental results indicate that the proposed techniques are robust to viewpoint changes and can rate correctly the dissimilarities between the shapes.

1. INTRODUCTION

An important problem in the object classification and retrieval is the fact that an object can be seen from different viewpoints, resulting in differently deformed images. Consequently, the invariance to viewpoints is a desirable property in many shape recognition systems. For near-planar objects, these deformations can be modeled approximately by affine transformations, if the viewpoints are sufficiently far away. In the literature, there are some papers on the shape recognition invariant under affine transformations. Some examples are affine-invariant (af-in) Fourier descriptors [1, 2], af-in moments [3, 4, 5] and affine curvature scale space [6, 7].

Area matrix is a structure that contains multiscale information about the shape. It was proposed by Shen et al. [8] and designed to be af-in. However, they did not correctly reparameterize the shape, and the resulting area matrix was not truly af-in. In a previous work [9], we reparameterized the shape using the affine-length before computing the area matrix, obtaining a truly af-in area matrix, insensible to the initial parameterization. Area matrix has too many information, and consequently it must be somehow summarized to become really useful. There are two ways doing it:

- The first is to extract af-in shape features, and construct a feature vector for each shape. Two feature vectors can be directly compared to rate the dissimilarity of the corresponding shapes. In our previous work [9], we extracted some simple features from the area matrix. In this paper, we propose a new set of af-in shape features: the central moments of the area matrix columns.

- The second is to construct an af-in shape similarity metric. Given the area matrices of two shapes, an af-in shape similarity metric performs some processing to rate their dissimilarity, regardless of their viewpoints. That is, the distance between two different views of the same object must be close to zero and the distance between two different objects must reflect their dissimilarity. In this paper, we propose a new set of af-in shape similarity metrics.

Experimental results indicate that the proposed techniques are robust to viewpoint changes, and make it possible to distinguish between two shapes that represent two different objects and two shapes that represent two different views of the same object.

2. FUNDAMENTAL CONCEPTS

2.1. Closed Curve

Given the image of a shape (for example, the images in figure 1), its edge pixels are extracted and traversed to yield a discrete closed curve that represents the shape. A discrete closed curve is a parametric equation $\mathbf{u}(t) = (x(t), y(t))$ where $t \in \{0, \ldots, N-1\}$ and $\mathbf{u}(N) = \mathbf{u}(0)$. A discrete curve is considered to be an approximate version of a continuous curve. The approximation of first and second derivatives of discrete curve $\mathbf{u}(t)$ can be computed by the central differences:

$$\dot{\mathbf{u}}(t) = \frac{\mathbf{u}(t+h) - \mathbf{u}(t-h)}{2h}$$

$$\ddot{\mathbf{u}}(t) = \frac{\mathbf{u}(t+h) - 2\mathbf{u}(t) + \mathbf{u}(t-h)}{h^2}$$

where h is a constant (usually $h = 1$).
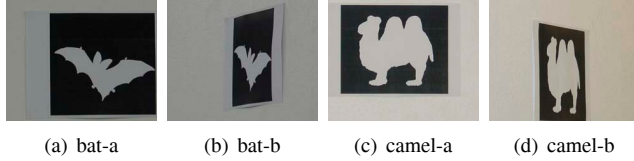
(a) bat-a   (b) bat-b   (c) camel-a   (d) camel-b

Fig. 1. Images of shapes viewed from different points.

2.2. Affine Transformation

An affine transformation $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ of a curve $\mathbf{u}$ is [10]:

$$\mathbf{F}(\mathbf{u}) = \mathbf{A}\mathbf{u} + \mathbf{b} \qquad (1)$$

where $\mathbf{A}$ is an invertible 2×2 matrix and $\mathbf{b} \in \mathbb{R}^2$ is the translation vector. A general affine transformation contains translation, scaling, rotation and shearing.

2.3. Affine Parameterization

The normalized affine-length of a continuous curve is [1, 3, 5, 11]:

$$\tau(p) = \frac{\int_0^p |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}} dt}{\int_0^T |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}} dt}, 0 \le p < T \quad (2)$$

and of a discrete curve is:

$$\tau(p) = \frac{\sum_0^{p-1} |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}}}{\sum_0^{N-1} |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}}}, \ p \in \{1, \dots, N\} \quad (3)$$

Given a discrete curve, we reparameterize it so that its vertices become equally affine-length spaced. Figure 2a shows an arc-length uniformly-spaced curve (in Euclidean distance sense) and figure 2b shows an affine-length uniformly-spaced curve. Note that equation 3 must be conveniently interpolated in order to perform the uniformly affine-spaced reparameterization. This interpolation can produce numerical errors. We found two ways to circumvent this error. The first is to perform repeated reparameterizations. It seems that reparameterizing an already almost affine-length uniformly-spaced curve leads to a more accurate parameterization. The second is to construct a (cubic) b-spline curve, using the original vertices as the control points. A b-spline curve is continuous, af-in [12] and its precise derivatives can be computed at any point of the curve (and not only at the vertices). Thus, it can be reparameterized precisely. We obtained the best numerical results (that is, the same shape viewed from different points yielded the most similar features and the smallest similarity distances) making first a discrete reparameterization, followed by another b-spline reparameterization.
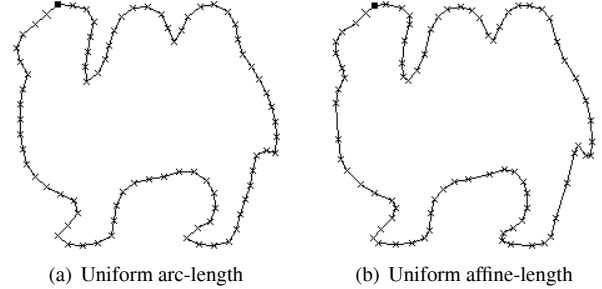

(a) Uniform arc-length   (b) Uniform affine-length

Fig. 2. Shape reparameterization with 80 vertices.

3. THE METHOD

3.1. Area and Affine Transformation

Let $\triangle$ be the triangle formed by vertices $u_1 = (x_1, y_1)$, $u_2 = (x_2, y_2)$ and $u_3 = (x_3, y_3)$. Then, the area of $\triangle$ is:

$$\text{Area}(\triangle) = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \qquad (4)$$

where $|.|$ is the determinant. The area is positive if u_1, u_2 and u_3 are counter-clockwise oriented, and negative if clockwise. Moreover, let $\mathbf{F}$ be an affine transformation and $\mathbf{A}$ its associated matrix. Then $\text{Area}(\mathbf{F}(\triangle)) = |\mathbf{A}|\text{Area}(\triangle)$.

3.2. Area Matrix

Let $\mathbf{u}(t)$, $t \in \{0, \dots, N-1\}$, be an affine-length uniformly-spaced closed curve and let $m = \lfloor (N-1)/2 \rfloor$. Let $\mathbf{A}'_{ti}$ be the area of the triangle formed by vertices $\mathbf{u}(t-i)$, $\mathbf{u}(t)$ and $\mathbf{u}(t+i)$. Then, the (non-normalized) area matrix of $\mathbf{u}$ is defined:

$$\mathbf{A}' = [\mathbf{A}'_{ti}]_{i=1,\dots,m}^{t=0,\dots,N-1} \qquad (5)$$

The central columns of $\mathbf{A}'$ contain coarse-scale or global information and the outer columns contain fine-scale or local information.

3.3. Normalized Area Matrix

The area matrix above is not scale-invariant, because the areas of the triangles increase and decrease with the changes of the scale. Moreover, different columns of the area matrix have different magnitudes (the central columns have large magnitudes, while the columns located at left or right regions have small magnitudes). We would like that all columns to be of similar magnitudes. To overcome these problems, we divide each matrix element $\mathbf{A}'_{t,i}$ by the absolute sum of the elements of column i, obtaining the normalized area matrix $\mathbf{A}$:

$$\mathbf{A}_{t,i} \leftarrow \frac{\mathbf{A}'_{t,i}}{\sum_{j=0}^{N-1} |\mathbf{A}'_{j,i}|} \qquad (6)$$

3.4. Central Moments

The normalized area matrix above is neither rotation-invariant nor invariant to the choice of the starting vertex, because a rotation (or a change of the starting vertex) will cause a circular shifting of the matrix rows. To obtain af-in shape features (inclusive rotation-invariant), in our previous paper [9] we have computed the sum, the sum of negative values and the sum of positive values of the weighted area matrix columns. In this work, we propose to use the central moments (of order n) of the columns of $\mathbf{A}$ as af-in shape features:

$$\Gamma_n(i) = \sum_{t=0}^{N-1} (\mathbf{A}_{t,i} - \bar{\mathbf{A}}_i)^n \quad i = 1, \ldots, m \qquad (7)$$

where $\bar{\mathbf{A}}_i$ is the mean value of column i of matrix $\mathbf{A}$.

3.5. Circular Matching Shape Similarity Metric

There may be different shapes with similar (positive/negative) sum of columns and similar central moments. In order to distinguish some shapes that are difficult to identify using only rotation-invariant features, we propose an af-in metric named circular matching shape similarity metric (CMSSM).

To compute CMSSM, the descriptor vector $\psi_{\mu,\sigma}$ is first extracted from the area matrix $\mathbf{A}$. The descriptor vector $\psi_{\mu,\sigma}$ depends on a gaussian curve $g_{\mu,\sigma}(.)$ with center μ and standard-deviation σ:

$$\psi_{\mu,\sigma}(t) = \sum_i g_{\mu,\sigma}(i)\mathbf{A}_{t,i} \quad t = 0, \ldots, N-1 \qquad (8)$$

Evidently, the index i should not wander outside the area matrix. We suggest not to place the center μ too close to the first or the last columns of $\mathbf{A}$, in order to minimize the influence of the noise in fine-scales. Given two descriptor vectors ψ_1 and ψ_2 (computed from the two shapes using the same gaussian curve $g_{\mu,\sigma}$), CMSSM distance can be computed by:

$$\text{CMSSM}(\psi_1, \psi_2) = \text{MIN}_{i=0}^{N-1}\{\text{RMSE}(\psi_1, \text{CS}_i(\psi_2))\} \quad (9)$$

where $\text{CS}_i(\psi_2)$ is the i-th circular shifting of vector ψ_2 and RMSE stands for "root mean square error". Obviously, other distance metrics between two vectors (such as the mean absolute error) can also be used.

4. EXPERIMENTAL RESULTS

Figure 1 depicts some animal figures viewed from different points. The edge pixels were extracted from these images and traversed to yield discrete closed curves. The closed curves were reparameterized using 200 vertices roughly equally affine-spaced, and then reparameterized again with cubic b-spline curves to yield 200 vertices precisely equally affine-spaced (figure 2 shows the reparameterization with 80 vertices).
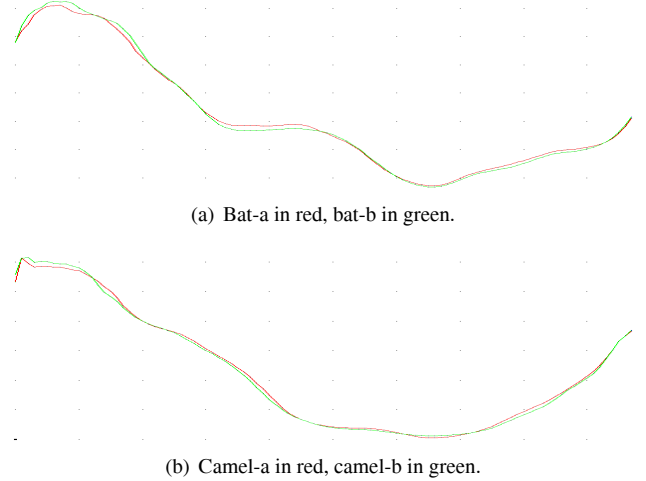

(a) Bat-a in red, bat-b in green.

(b) Camel-a in red, camel-b in green.
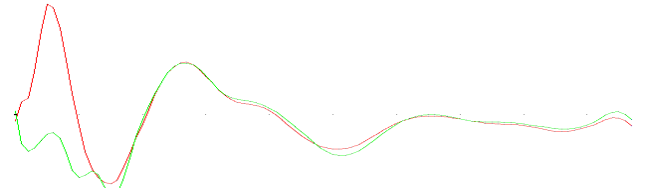
Fig. 3. Second central moments.

Fig. 4. Third central moments of the two bat figures.

Figure 3 depicts the second central moments of the four shapes. It is easy to see that the two bats yield almost the same features (figure 3a), as well as the two camels (figure 3b), while the curves between the bats and the camels are quite different. Indeed, table 1 shows that RMSEs between the second central moments of the same animal are approximately 10 times smaller than the RMSEs between different animals. The same happens with the fourth central moment.

However, using the third or fifth central moments, the bat can be hardly identified. The difference between the two bats is of similar magnitude of the differences between bats and camels. This problem seems to happen only using n-th central moments, with odd n. Figure 4 depicts the third central moments of the two bat figures and it is easy to see that the features are mismatching only at fine-scales. We suggest either not to use odd-order central moments, or to discard the fine-scale features.

To compute the CMSSMs between the four figures, we have used three Gaussians g_l, g_c, and g_r (left, central and right) with standard deviations $\sigma = 5$ columns and the centers located respectively at $\mu_l = 0.2m$, $\mu_c = 0.5m$ and $\mu_r = 0.9m$ (where m is the number of columns of the area matrix). Figure 5 depicts the aligned central descriptor vectors. It is easy to see that two figures of the same animal have yielded almost the same curves, while figures of different animals have yielded completely different curves. Table 2 shows

Table 1. Similarities of shapes according to the central moment descriptors.

	Γ_2	Γ_3	Γ_4	Γ_5
RMSE	$\times 10^{-6}$	$\times 10^{-9}$	$\times 10^{-10}$	$\times 10^{-12}$
(bat-a,bat-b)	11	664	32	180
(bat-a,cam-a)	108	871	132	158
(bat-a,cam-b)	109	978	120	159
(bat-b,cam-a)	109	959	127	131
(bat-b,cam-b)	109	1027	116	140
(cam-a,cam-b)	8	154	16	27

(a) Bat-a in red, bat-b in green.
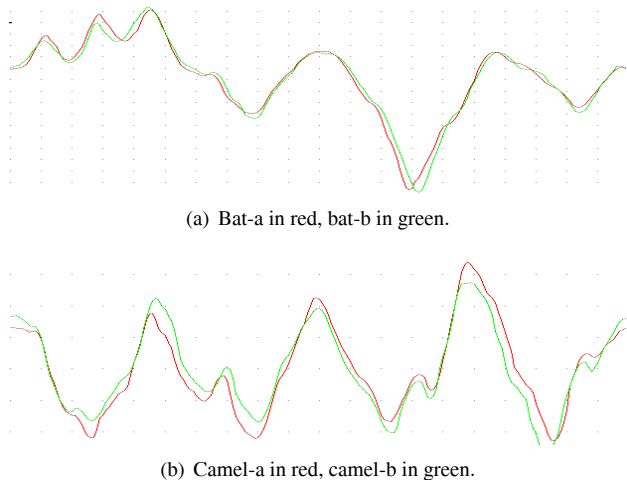

(b) Camel-a in red, camel-b in green.

Fig. 5. Descriptor vectors with the center $\mu_c = 0.5m$ and standard deviation $\sigma = 5$ columns.

the CMSSM distances between different figures. CMSSM distances between the same animals are more or less 5 times smaller than the distances between different animals.

Table 2. Similarities of shapes according to the CMSSM distances.

CMSSM $\times 10^{-6}$	left	central	right
(bat-a,bat-b)	76	32	32
(bat-a,cam-a)	305	207	185
(bat-a,cam-b)	317	207	187
(bat-b,cam-a)	310	204	183
(bat-b,cam-b)	327	204	186
(cam-a,cam-b)	59	31	32

5. CONCLUSIONS

This paper has introduced a new set of affine-invariant shape features and shape similarity metrics. They are based on the area matrix, properly reparameterized and normalized in order to become truly affine-invariant. The features are the cen-

tral moments of the area matrix columns. We have shown experimentally that even-order central moments are more stable under affine transformations than odd-order central moments. The similarity metric performs a circular matching of the area matrices in a chosen scale. The proposed features and metrics made it possible to distinguish clearly two different objects from two different views of the same object.

6. REFERENCES

[1] H. Burkhardt K. Arbter, W. E. Snyder and G. Hirzinger, "Applications of affine-invariant fourier descriptors to recognition of 3-d objects," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 640–647, 1990.

[2] F. Chaker, M. T. Bannour, and F. Ghorbel, "A complete and stable set of affine-invariant fourier descriptors," in *Proceedings of International Conference on Image Analysis and Processing*, Mantova, Italy, 17-19 Sept. 2003, pp. 578–581.

[3] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants," *Pattern Recognition*, vol. 26, no. 1, pp. 167–174, 1993.

[4] Z. Huang and F.S. Cohen, "Affine-invariant b-spline moment for curve matching," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA, 1994, pp. 490–495.

[5] A. Zhao and J. Chen, "Affine curve moment invariants for shape recognition," *Pattern Recognition*, vol. 30, no. 6, pp. 895–901, 1997.

[6] F. Mokhtarian and S. Abbasi, "Affine curvature scale space with affine length parametrisation," *Pattern Analysis & Applications*, vol. 4, pp. 1–8, 2001.

[7] F. Mokhtarian and S. Abbasi, "Shape similarity retrieval under affine transforms," *Pattern Recognition*, vol. 35, no. 1, pp. 31–41, 2002.

[8] D. Shen, W. Wong, and H.S.H. Ip, "Affine-invariant image retrieval by correspondence matching of shapes," *Image and Vision Computing*, vol. 17, pp. 489–499, 1999.

[9] C. R. P. Dionisio and H. Y. Kim, "New features for affine-invariant shape classification," in *Proceedings of IEEE Int. Conf. on Image Processing*, Singapore, 2004, pp. 2135–2138.

[10] M. Stanislaw, *Description invariante et locale des formes planes - Application à l'indexation d'une base d'images*, Doutorado, Universite de Valenciennes et du Hainaut Cambresis - ENIC/INT, 1999.

[11] S. Sapiro and A. Tannenbaum, "Affine invariant scale-space," *International Journal of Computer Vision*, vol. 11, no. 1, pp. 25–44, 1993.

[12] Z. H. Huang F. S. Cohen and Z. W. Yang, "Invariant matching and identification of curves using b-splines curve representation," *IEEE Transactions on Image Processing*, vol. 4, no. 1, pp. 1–10, 1995.