NEW FEATURES FOR AFFINE-INVARIANT SHAPE CLASSIFICATION

Carlos R. P. Dionisio and Hae Yong Kim

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

ABSTRACT

An object seen from different viewpoints results in differently deformed images. Affine-invariant shape classification must classify correctly the object, disregarding its viewpoint. In this paper, we propose new local and global features invariant under affine transformation. These features can be used for supervised or unsupervised shape classification, and for shape-based image indexing and retrieval. One of the proposed features is related to the convex deficiency and the others are extracted from the area matrix. Area matrix was used by Shen [1] for the similarity matching in image retrieval. However, differently from the Shen's work, we parameterize the shape contour using the affine-length parameter. This makes our features robust to affine parameterization, while Shen's's work does not have this property. Experimental results indicate that our method can classify correctly even highly deformed and noisy shapes using small training sets.

1. INTRODUCTION

An important problem in object classification is the fact that an object can be seen from different viewpoints, resulting in different images. Consequently, the invariance to viewpoints is a desirable property in many shape recognition systems. For near planar objects, these deformations can be modelled approximately by affine transformation, if the viewpoints is sufficiently far away. In the literature, there are many papers on the shape recognition invariant under affine transformation. Some examples are affine-invariant Fourier descriptors [2], affine-invariant moments [3, 4, 5] and affine curvature scale space [6]. In these works, the underlying idea is to use an affine parameterization, usually the affine-length parameter [2, 3, 5, 7].

Usually, a good classification rate can be achieved by using simultaneously local and global features. Convex deficiency is a global feature that contains information about concavities of the shape. This feature provides invariance to translations, rotations and can be easily made tolerant

to scaling. In this paper, we show that the convex deficiency can also become tolerant to affine transformation. The other proposed features are extracted from the area matrix. Area matrix was used by Shen [1] for the shape similarity matching in affine-invariant image retrieval. This matrix has two interesting properties. The first is that it can be made tolerant to affine transformation. The second is that its columns contain multiscale information: central columns contain coarse-scale or global information and outer columns contain fine-scale or local information. Differently from Shen's work, we parameterize shapes using the affine-length parameter. This procedure makes the area matrix to become invariant under affine parameterization. Computational complexity of Shen's work depends on the dimension of the area matrix. On the contrary, we extract and process only the relevant information, decreasing computational complexity. In the recognition phase, we use the k-nearest neighbors algorithm.

2. FUNDAMENTAL CONCEPTS

2.1. Closed Curve

Let us consider the discretized parametric equation of a closed curve:

$$\mathbf{u}(t) = (x(t), y(t)) \tag{1}$$

where $t \in \{0, \ldots, N-1\}$ and $\mathbf{u}(N) = \mathbf{u}(0)$. In general, curves of an application may have been parameterized with different number of vertices N. The first and second derivatives of $\mathbf{u}(t)$ are denoted, respectively, as $\dot{\mathbf{u}}(t) = (\dot{x}(t), \dot{y}(t))$ and $\ddot{\mathbf{u}}(t) = (\ddot{x}(t), \ddot{y}(t))$.

2.2. Affine Transformation

A general affine transformation T in $\mathbb{R}^2$ is defined as

$$T(X) = AX + b \tag{2}$$

where $X \in \mathbb{R}^2$, A is an affine matrix (a real matrix 2×2 with positive determinant) and $b \in \mathbb{R}^2$ is the translation vector. A general affine transformation contains translation, scaling, rotation and shearing. The following matrix represents a shear transformation and the parameter κ, called

shear ratio, controls the amount of deformation.

$$A_\kappa = \begin{bmatrix} 1 & \kappa \\ 0 & 1 \end{bmatrix} \qquad (3)$$

2.3. Area and Affine Transformation

If T is an affine transformation and A its matrix then:

$$\text{Area}(T(\triangle)) = |A|\text{Area}(\triangle) \qquad (4)$$

where $|A|$ is the determinant of A. Considering equation 4 and the fact that any simple 2-D polygon $\mathbf{u}$ can be decomposed as a set of triangles, we can state:

$$\text{Area}(T(\mathbf{u})) = |A|\text{Area}(\mathbf{u}) \qquad (5)$$

Let $\text{CH}(\mathbf{u})$ denote the convex hull of the curve $\mathbf{u}$. Clearly, a point belongs to $\text{CH}(\mathbf{u})$ if, and only if, it also belongs to $\text{CH}(T(\mathbf{u}))$. Consequently, we can also state:

$$\text{Area}(T(\text{CH}(\mathbf{u}))) = |A|\text{Area}(\text{CH}(\mathbf{u})) \qquad (6)$$

Based on equations 5 and 6, we define our first affine-invariant feature by dividing the convex deficiency of $\mathbf{u}$ by the area of convex hull of $\mathbf{u}$:

$$F_1 = \frac{\text{Area}(\text{CH}(\mathbf{u})) - \text{Area}(\mathbf{u})}{\text{Area}(\text{CH}(\mathbf{u}))} \qquad (7)$$

2.4. Affine Parameterization

We normalized the number of vertices of all shapes, because the proposed features depends on the choice of this number. We also re-parameterized all shapes using the affine-length parameter [2, 3, 5, 7]:

$$\tau(p) = \frac{\int_0^p |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}} dt}{\int_0^N |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}} dt} \qquad (8)$$

Hence, our features are also invariant under parameterization (besides being affine-invariant). To assure that τ is a strictly increasing function, we remove all vertices with the second derivative equal to zero. This is equivalent to removing vertices that belong to straight lines.

3. AFFINE-INVARIANT FEATURES

3.1. Area Matrix

Let $\mathbf{u}(t)$, $t \in \{0, \ldots, N-1\}$, be a closed curve and let $m = N/2$. 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 area matrix of $\mathbf{u}$ is defined:

$$\mathbf{A} = [\mathbf{A}_{ti}]_{i=1,\ldots,m}^{t=0,\ldots,N-1} \qquad (9)$$

The i-th column vector of $\mathbf{A}$ contains local (fine scale) information when i is small or close to m, and contains global (coarse scale) information when i is situated in the central region of $\mathbf{A}$. Indeed, each column vector of the area matrix analyzes the shape in a determined scale, and the whole area matrix contains multiscale information of the shape.

3.2. Normalized Area Matrix

We normalize an area matrix $\mathbf{A}$ in two steps: first, we make $\mathbf{A}$ to become affine invariant by dividing each element of $\mathbf{A}$ by the sum of all elements of $\mathbf{A}$. Second, we make all columns to be of similar magnitudes: We divide each column i of $\mathbf{A}$ by a weight $w(i)$. The weight vector w is computed from an affine-invariant area matrix $\mathbf{R}$ of a representative shape using the following equation:

$$w(i) = \max_{t=0,\ldots,N-1} \{|\mathbf{R}_{t,i}|\} \quad i = 1, \ldots, m \qquad (10)$$

3.3. Vectors with Local and Global Information

An area matrix contains too much information. To be useful, it must be reduced and organized somehow. We extract relevant information from a normalized area matrix $\mathbf{A}$ and store it in three vectors ψ_l, ψ_c and ψ_r that summarizes respectively information of columns situated at left, center and right of matrix $\mathbf{A}$. Vector ψ_l is computed by:

$$\psi_l(t) = \frac{\sum_i g_l(i)\mathbf{A}_{t,i}}{\sum_i g_l(i)} \quad t = 0, \ldots, N-1 \qquad (11)$$

where g_l is a Gaussian curve. The center of this Gaussian is situated at the left side of $\mathbf{A}$. It cannot be too close to the leftmost column, in order to reduce noise present in very fine scales. The standard deviation σ is chosen to extract all information present in the left side of $\mathbf{A}$, without mixing up with information of the central region of $\mathbf{A}$. Vectors ψ_c and ψ_r are similarly computed.

In this work, the three Gaussians were set according to Table 1:

Table 1. Configuration of the three Gaussians (m is the number of columns of $\mathbf{A}$).

Gaussian	σ	center
g_l	5	$0.20m$
g_c	5	$0.50m$
g_r	5	$0.85m$

3.4. The Proposed Features

We use five features to classify shapes:

- Feature F_1 was defined in equation 7.

- Feature F_2 is the sum of negative values of ψ_l, that is, $F_2 = \sum_{\psi_l(t)<0} \psi_l(t)$.

- Feature F_3 is the sum of positive values of ψ_l, that is, $F_3 = \sum_{\psi_l(t)>0} \psi_l(t)$.

- Feature F_4 is the sum of ψ_c, that is, $F_4 = \sum \psi_c(t)$.

- Feature F_5 is the sum of ψ_r, that is, $F_5 = \sum \psi_r(t)$.

All these features are invariants under affine transformation. They are also invariant to starting point. In order to use k-nearest neighbors classification, we normalized all features to have mean zero and standard deviation one.

4. EXPERIMENTAL RESULTS

We applied our method to the supervised recognition of static hand gestures. We used part of static hand gesture contours considered by Milios and Petrakis [8]. All contours were normalized with 140 vertices. We labelled manually 150 closed contours as examples for the supervised training. They are divided into six classes, each class with 25 examples, as depicted in Figure 1.
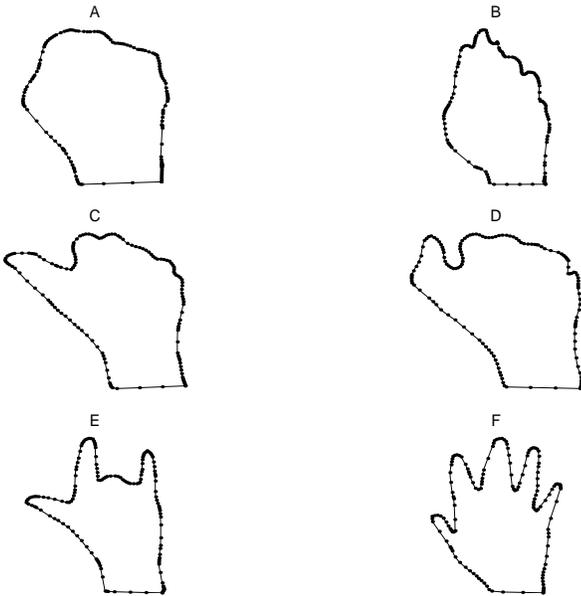

Fig. 1. Hand gestures of classes A, B, C, D, E and F.

We tested affine invariance of our features, specifically for the rotation and the shear transformation. We took one representative contour of each class (that does not belong to the training set) and generated 9 rotated versions, rotating the original contour by $20°$, $40°$, ..., $180°$. Then, we deformed the 6×9 rotated versions with shear ratios $\kappa = 1, 2$

and 3, obtaining $3 \times 6 \times 9$ deformed shapes. Figures 2 and 3 depicts the original and deformed shapes with shear ratios $\kappa = 2$ and 3, respectively.
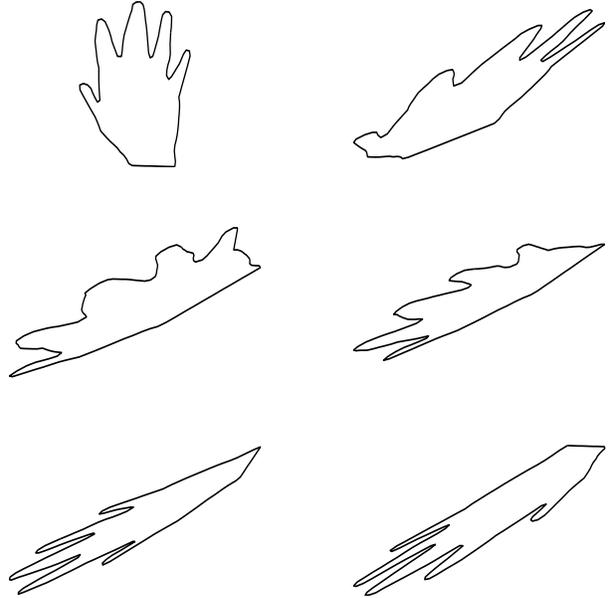

Fig. 2. The upper-left gesture belongs to class F. The others are deformed shapes with $\kappa = 2$.
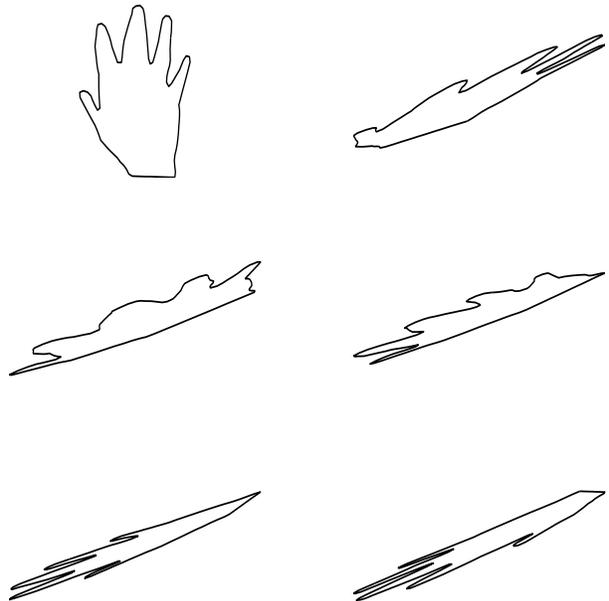

Fig. 3. The upper-left gesture belongs to class F. The others are deformed shapes with $\kappa = 3$.

For the recognition phase, we use k-nearest neighbors

Table 2. Classification of test data.

	A	B	C	D	E	F
A	25	0	0	0	0	0
B	1	24	0	0	0	0
C	0	0	25	0	0	0
D	0	1	0	24	0	0
E	0	0	0	0	25	0
F	0	0	0	1	0	25

algorithm. We tested this algorithm with k ranging from 1 to 25. All deformed shapes were correctly classified, using any $k \in \{1, \ldots, 25\}$. This shows the robustness of the proposes features under affine transformation.

We also labelled manually 150 non-deformed test contours, with 25 contours in each class. The set of test contours was disjoint from the set of training contours. Table 2 summarizes the classification of test hand gestures for non-deformed shapes, using the 5-nearest neighbors algorithm. In this table, the element (row i, column j) corresponds to the number of hand gestures manually labelled as i and classified as j by the algorithm. The error rate is very low, what shows the aptness of the proposed features to the shape classification. Figure 4 depicts some shapes that belong to classes E and F, and their classifications.
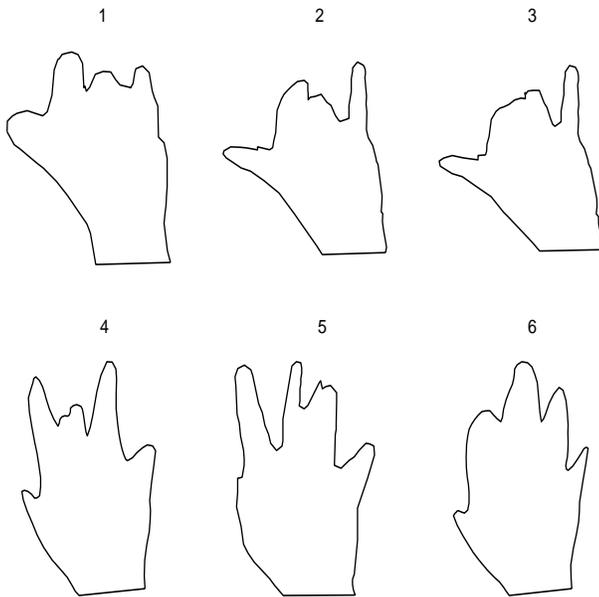

Fig. 4. Gestures 1, 2 and 3 belong to class E and they were correctly classified by our algorithm. Gestures 4, 5 and 6 belong to class F. Gestures 4 and 5 were correctly classified, but gesture 6 was classified as D.

5. CONCLUSIONS

In this paper, we have proposed a set of new features invariant under affine transformation. We have tested them in supervised shape classification, but they can also be used in unsupervised shape classification and in shape-based image indexing and retrieval. We have parameterized the shapes using an affine-invariant parameter, to make our features also invariant under parameterization. Experimental results have shown that our features are able to correctly discriminate shapes severely deformed by affine transformation.

6. REFERENCES

[1] 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.

[2] K. Arbter, W.E. Snyder, and G. Hirzinger H. Burkhardt, "Applications of affine-invariant fourier descriptors to recognition of 3-d objects," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 640–646, 1990.

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

[8] E. Milios and E.G.M. Petrakis, "Shape retrieval based on dynamic programming," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 141–146, 2000.