A Supervised Shape Classification Technique Invariant under Rotation and Scaling

Carlos R. P. Dionisio, Hae Yong Kim
Escola Politécnica, Universidade de São Paulo, Brazil

Abstract—**In this paper, we propose a new object classification technique based on polygonal approximation of the open profile of object. The vertices of polygonal approximation are formed by high curvature points of the profile and they are selected by Fourier transform of the object contour. A series of features are computed from the polygonal approximation and then the minimum distance classifier is used to recognize the object. The proposed technique is fast, simple and invariant under translation, rotation and scaling. Experimental results in recognition of static hand gestures show the performance of the proposed technique.**

I. INTRODUCTION

REPRESENTATION of the shape of an object can be based on its contour (i.e., its profile, outline or boundary) or on its region. The characterization of shape by its outline is a more natural choice because it resembles the way we, human beings, mentally represent an object. Our visual system itself focuses on boundaries and ignores uniform regions. This ability is "hard-wired" in our retina: two neuron layers that perform an operation similar to Laplacian are directly connected to rods and cones of retina. This operation is called lateral inhibition and helps us to extract contours and boundaries [5]. Consequently, in this paper we have chosen to represent shapes by outlines instead of by regions.

In this work, we simplify the complex shape of an object by a polygonal approximation. The polygonal approximation of the shape consists on finding significant vertices along the contour such that these vertices constitute a good approximation of the original contour. A classic approach to this problem is to take the high curvature points (i.e., points with high absolute value of curvature) as significant vertices.

The curvature is one of most important attributes that can be extracted from contours. For example, we can cite two psychophysical experiments carried out by Attneave [3]. In the first experiment, a series of 2D shapes has been presented to subjects, who were asked to represent each contour using a set of 10 points. The results showed clearly that the majority of subjects would rather use the high curvature points to represent each shape. In the second experiment, Attneave draw the portrait of a cat identifying high curvature points connected by straight lines, demonstrating that the majority of information in the original image is concentrated in the high curvature points. Similar experiments and results were reported also by Fischler and Wolf [4]. Consequently, the approximation of curves by straight lines connecting high curvature points retains the necessary information for a successful recognition of shape. This fact can be explained considering that our visual system focuses on singularities and ignores smooth curves using the lateral inhibition mechanism.

The shape representation by high curvature points considerably reduces number of points of contour, while keeping the necessary information for the shape recognition. Moreover, the high curvature points are robust descriptors of shape, in the sense that they are invariants under translation, rotation and scale changing [6].

In the literature, there are many methods to determine curvature of object boundary (for example, [7], [8], [9], [10]). In this paper, we will make use of a simple approach described by Costa and Cesar [1]. This approach utilizes the derivative of Fourier transform in order to determine the curvature. We have chosen this method due to its simplicity.

After approximating shapes of objects by polygons, a set of features are computed for each object. We have used eight features, all invariants under translation, rotation and scaling. In section 2 we describe these features in details. When all features are calculated, we compute the centroids (mean vectors) for each pattern class. Then, given a query pattern vector, the minimum distance classifier will be used to identify its class.

We have applied the method described above to the static hand gesture recognition problem. The experimental results demonstrate the performance of our technique. This paper is organized as follows. In Section 2, the method and the recognition algorithm are described. Experimental results are presented in Section 3. In Section 4, we present possible improvements of our work. Finally, in Section 5 we present our conclusions.

II. THE METHOD

It is not easy to find a set of high curvature points that constitute a good polygonal approximation of the outline of an object. Usually, a computationally-acquired outline is rough and there are many high curvature points in this rough outline that do not belong to a good polygonal approximation. There are some techniques that can reduce this problem (for example, interpolation or filtering), but they increase the total processing time. To prevent this problem, we delete too-near points of the contour. This simple pre-processing is enough to alleviate rough outline problem.

We present below the technique described by Costa and Cesar [1] for the curvature evaluation. Let $(x(t), y(t))$ be an object open contour and $u(t) = x(t) + jy(t)$ its representation in complex plane. The curvature $k(t)$ of $(x(t), y(t))$ is defined as

$$k(t) = \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}} \qquad (1)$$

The first and second derivative of $u(t)$, that is $\dot{u}(t)$ and $\ddot{u}(t)$, are defined as $\dot{u}(t) = \dot{x}(t) + j\dot{y}(t)$ and $\ddot{u}(t) = \ddot{x}(t) + j\ddot{y}(t)$,

Carlos R. P. Dionisio and Hae Yong Kim are with the Laboratório de Processamento de Sinais, Departamento de Engenharia de Sistemas Eletrônicos, Escola Politécnica, Universidade de São Paulo, São Paulo, SP, Brazil, Phone: (55-11) 3091-5605 Fax: (55-11) 3091-5718. E-mails: {carlos,hae}@lps.usp.br.

respectively. Taking into account the following relations

$$\begin{aligned}\dot{u}(t)\ddot{u}^*(t) &= \dot{x}(t)\dot{y}(t) + \ddot{x}(t)\ddot{y}(t) - j(\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)) \\ |\dot{u}(t)|^3 &= (\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}\end{aligned}$$

and Equation 1, we get the following representation of curvature $k(t)$ in terms of $u(t)$:

$$k(t) = \frac{-\mathrm{Im}\{\dot{u}(t)\ddot{u}(t)^*\}}{|\dot{u}(t)|^3}. \tag{2}$$

Let $U(f)$ be the Fourier transform of $u(t)$. This fact will be denoted as $u(t) \longleftrightarrow U(f)$. The derivatives of $u(t)$ and its Fourier transform present the following properties:

$$\begin{aligned}\dot{u}(t) &\longleftrightarrow (j2\pi f)\, U(f) \\ \ddot{u}(t) &\longleftrightarrow (j2\pi f)^2\, U(f)\end{aligned}$$

Thus, $U(f)$ can be used to estimate $\dot{u}(t)$, $\ddot{u}(t)$ and consequently the curvature $k(t)$.

To compute automatically high curvature points, the curvature is normalized by its perimeter. A point of contour $u(t)$ is classified as high curvature point if its curvature satisfies the following system:

$$\begin{cases} k(t) \geq \delta_p, & k(t) \geq 0 \\ k(t) \leq \delta_n, & k(t) < 0 \end{cases} \tag{3}$$

where δ_p and δ_n are the average values of positive and negative parts of the curvature, respectively. These high curvature points be the vertices of the polygonal approximation of contour. We also include the two extremity points of the open contour in the polygonal approximation.

Figure 1 depicts the contour of a hand gesture and its polygonal approximation. Note that this approximation may not be the best one. However, it is enough for our purposes. The features

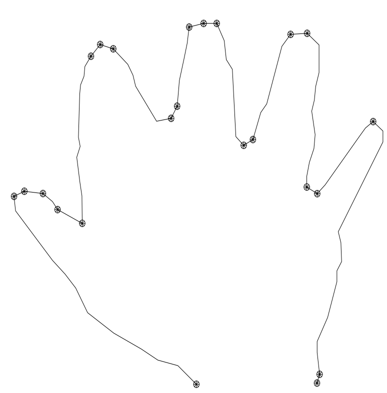

Fig. 1. Dots on the outline are the polygonal approximation determined by the proposed method.

to be used in classification are calculated using this polygonal approximation. We describe below the eight features to be used in the classification.

A. Features

Let u be the polygonal approximation of the outline of an object. Let us denote the perimeter and the area of u as P and A, respectively. The variable F will denote the area of the convex hull of u. Let D_{max} and D_{mean} be respectively the maximum and mean distances between the centroid and vertices of u. Let H be the perimeter of convex hull of u. A pattern vector will be composed by the following eight features:

C_1. *Circularity:* $\frac{P^2}{A}$.

C_2. $\frac{P-\sqrt{P^2-4\pi A}}{P+\sqrt{P^2-4\pi A}}$.

C_3. $(\log_2(\frac{2P}{P-H}))^{-1}$.

C_4. The average of curvature values of u divided by the perimeter of u.

C_5. The convex deficiency divided by the area of u, that is, $\frac{F-A}{A}$.

C_6. $\frac{A}{F}$.

C_7. The median of curvature values of u.

C_8. $\frac{D_{\text{max}}}{D_{\text{mean}}}$.

All these features are invariants under translation, rotation and scaling. The invariace is highly useful to obtain a robust classification.

B. Training Algorithm

The training algorithm can be summarized as follows. The input of this algorithm are the outlines of objects with respective labels. These labels certify the correct classification of each outline. The output of this algorithm is a table of centroids of each class.

- For each contour, delete points that have too close neighbors.
- Compute the curvature of each point of contour, using Equation 2.
- Determine the high curvature points of each contour using Equation 3. These points, together with the two extremity points of the contour, are the polygonal approximation of the contour.
- Compute the eight features for each polygonal approximation.
- Compute the centroid of each class of object.

C. Classifying Algorithm

The classifying algorithm receives as input a query contour with unknown classification and computes its class.

- Given a query contour with unknown classification, compute the eight features as described in the training algorithm. Let us denote the pattern vector so obtained as x.
- The pattern x will be classified as belonging to the class that has the shortest Euclidean distance between x and the its centroid.

III. Experimental Results

We have applied our method to supervised recognition of static hand gestures. We have used part of open contours of static hand gestures considered by Milios and Petrakis in [2]. These gestures are available at

"http://www.cs.yorku.ca/˜eem/gesturesDB/".

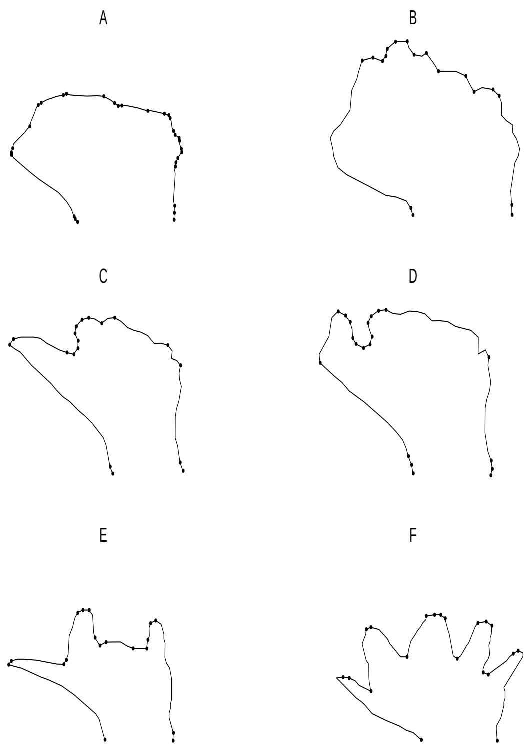

Fig. 2. Outlines of hand gesture classes A, B, C, D, E and F.

TABLE I
CLASSIFICATION OF TEST DATA.

	A	B	C	D	E	F
A	6	0	2	0	0	0
B	0	4	1	0	0	0
C	0	3	9	0	0	0
D	0	0	0	4	0	0
E	0	1	0	0	4	0
F	0	1	0	0	4	5

We have labelled manually 156 open contours for the training stage. They are divided into six classes, as depicted in Figure 2. We have also labelled 44 test contours. The test and training contours are disjoint.

Note that the polygonal approximation of gesture type B is very poor: there are too many high curvature points in the upper side of the contour. This is due to the roughness of that part of the contour. The remaining polygonal approximations have an acceptable quality.

Table I summarizes the classification results of the test hand gestures. In this table, the element (i, j) (row i, column j) corresponds to the number of hand gestures manually labelled as i classified as j by our algorithm. For example, the second row states that there are five test hand gestures type B. Four of them were correctly classified by our algorithm and one was classified as type C. We have also tested our method on training hand gestures and the results are listed in Table II.

Figures 3, 4, 5, 6, 7 and 8 depicts some correctly and uncorrectly classified test hand gestures.

TABLE II
CLASSIFICATION OF TRAINING DATA.

	A	B	C	D	E	F
A	22	0	0	0	0	0
B	0	13	3	0	0	0
C	0	2	24	1	0	0
D	0	0	0	22	0	0
E	0	0	1	0	15	0
F	0	0	0	0	2	51

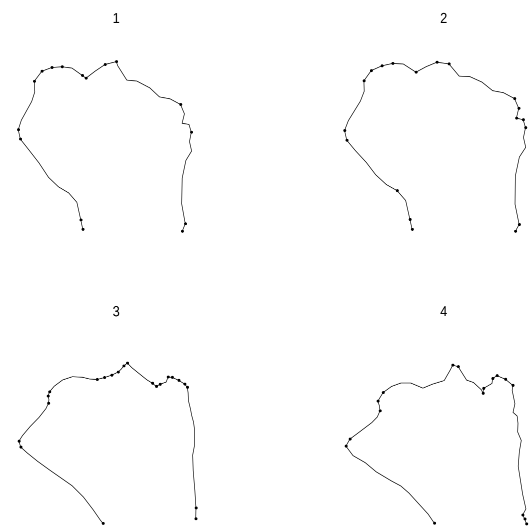

Fig. 3. Four test outlines manually labelled as class A. The gestures 1, 2 and 3 were classified by our algorithm as A. The gesture 4 was classified as C.

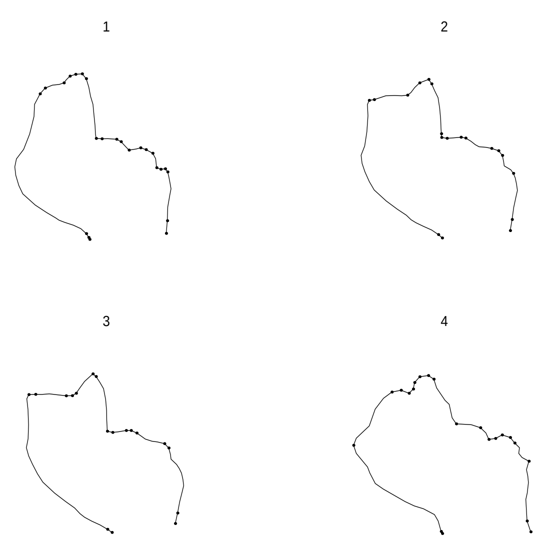

Fig. 4. Four test outlines manually labelled as class B. The gestures 1, 2 and 3 were classified by our algorithm as B. The gesture 4 was classified as C.

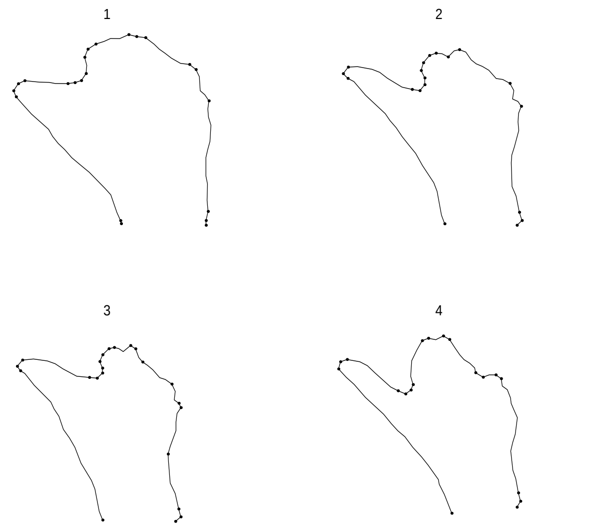

Fig. 5. Four test outlines manually labelled as class C. The gestures 1, 2 and 3 were classified by our algorithm as C. The gesture 4 was classified as B.

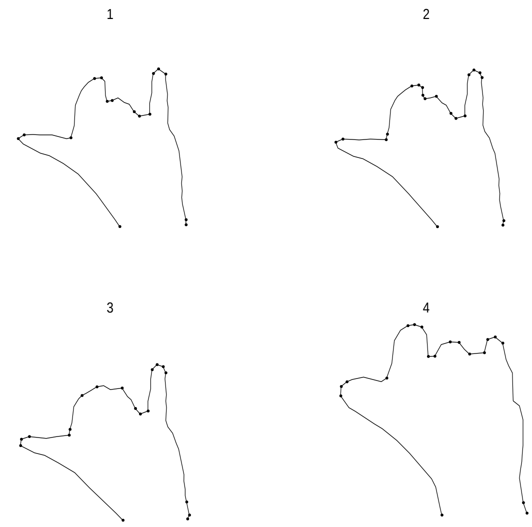

Fig. 7. Four test outlines manually labelled as class E. The gestures 1, 2 and 3 were classified by our algorithm as E. The gesture 4 was classified as B.

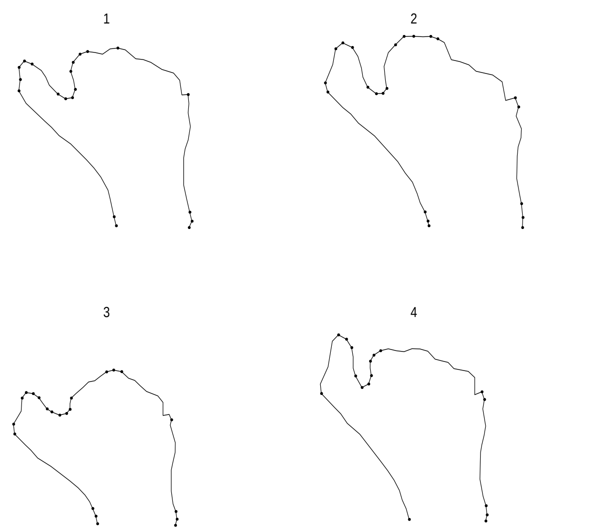

Fig. 6. Four test outlines manually labelled as class D. They were all classified by our algorithm as D.

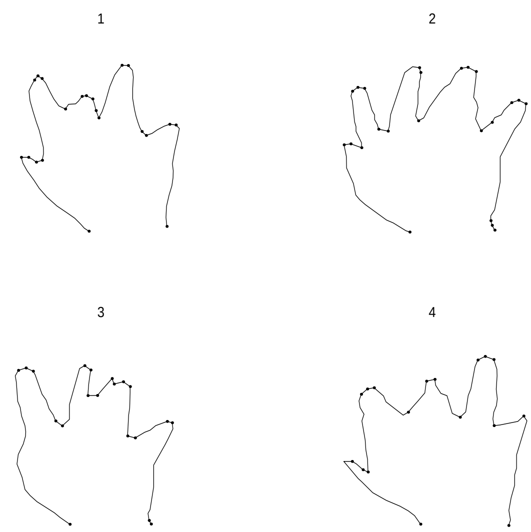

Fig. 8. Four test outlines manually labelled as class F. The gestures 1, 2 and 3 were classified by our algorithm as F. The gesture 4 was classified as E.

IV. Future Works

We deem that it is possible to get a better polygonal approximation of the contour using a multi-resolution approach like scale-space or wavelet. This would diminish the classification error rate. The classification process also can be improved using a more sophisticated machine learning scheme, like decision-tree or k-nearest neighbor classifier.

V. Conclusions

In this work, we have proposed a new technique for shape recognition invariant under translation, rotation and scaling. This technique makes use of Fourier transform to compute high curvature points. A set of features are extracted from these high curvature points and the minimum distance classifier is used as the recognition algorithm. We have applied the proposed technique to recognize static hand gestures and the high recognition rates obtained allow practical applications of our technique.

Acknowledgments

Carlos R. P. Dionisio is grateful to Capes for the financial support of this work under Ph.D. student grant. Hae Yong Kim is grateful to FAPESP and CNPq for the partial financial supports of this work under grants 1/02400-9 and 300689/98-5, respectively.

References

[1] L.F. Costa and R.M. Cesar Jr., *Shape Analysis and Classification: Theory and Practice*, Boca Raton, FL: CRC Press, 2001.

[2] E. Milios and E.G.M. Petrakis, *Shape Retrieval Based on Dynamic Programming*, IEEE Transactions on Image Processing, vol. 9, no. 1, 2000, 141-146.

[3] F. Attneave, *Some Informational Aspects of Visual Perception*, Psychological Review, 61:183-193, 1954.

[4] M.A. Fischler and H.C. Wolf, *Locating Perceptually Salient Points on Pla-*

nar Curves, IEEE Trannsactions on Pattern Analysis and Machine Intelligence, 16(2):113-129, February 1994.

[5] J.C. Russ, *The Image Processing Handbook*, 3rd edition, CRC, Springer and IEEE Press inc., 1995.

[6] C.H. Teh and R.T. Chin, *On the detection of dominant points on digital curves*, IEEE Trans. PAMI, vol. 11, pp. 859-872, 1989.

[7] H.L. Beus and S.S.H. Tiu, *An Improved Corner Detection Algorithm Based on Chain-Coded Plane Curves*, Pattern Recognition, 20:291-296, 1987.

[8] H.C. Liu and M.D. Srinath, *Corner detection from chain-code*, Pattern Recognition, 23(1-2):51-68, 1990.

[9] A. Rosenfeld and J.S. Weszka, *An improved method of angle detection on digital curves*, IEEE Tran. Comput., 24:940-941, 1975.

[10] J. Koplowitz and S. Plante, *Corner detection for chain coded curves*, Pattern Recognition, 28(6):843-852, 1995.