I-013

Automatic Shape Classification through Descriptor Matching

Yingdi Xie† Jun Ohya†

`

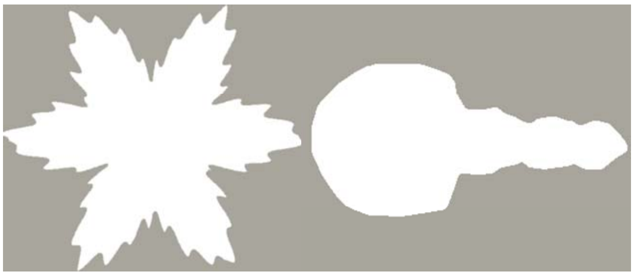
1. Introduction

Shape recognition and classification is of great importance in both human perception and visual computation. For its fundamentality toward object recognition, this topic has been drawing broad attention from specific shape, such as ellipse [1], to general shape classification [2]. The former category achieves recognition through parameterization which is subject to given formula of specific shapes. The nature of detection preciseness brings the specific shape fitting solvent wide application among accurate measurements and calibration, e.g. defect detection in industrial products. The general shape issue casts research emphasis on shape feature analysis with trained classifiers. For its functionality of handling exterior variation, general visual computation tasks (e.g. shape based object classification) frequently adopt the algorithms in this category, which is our focus in this paper.

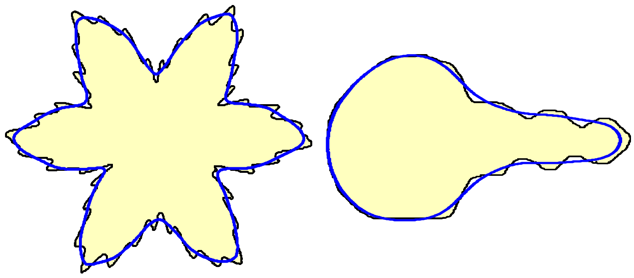
The overview of the proposed method and relevant assumption is described in section 2. Hierarchical edge orientation curve calculation with respect to the spline fitting to object contour by smoothed spline functions is elaborated in section 3. Section 4 introduces the Multi-layer k-nearest neighbors classifier, or ML-KNN classifier based shape recognition. Experiments and conclusion are given in section 5.

2. Proposal Overview

As for shape recognition, trivial parts of object shape often contain accurate but quantitative detail information. Facing the shape recognition task for human being, we tend to retrieve the result by selectively ignoring the trivial parts of object contour, until corresponding necessity is required. To select a suitable smoothing parameter, questionnaires are delivered to collect subject judgment for trivial parts ignorance. The questionnaire includes object silhouette that selected from the MPEG-7 database, and the spline fitting based contour fitting with the same smoothing parameter. The result manifests that for a general shaped object silhouette, e.g. left image in Fig.1, 95% shows satisfaction for the smoothing; while for the one with obvious corresponding natural interpretation, e.g. a key shaped silhouette as demonstrated in the right Fig.1, 80% of the experimenters indicate the smoothing is inappropriate and shall preserve the zigzag contour as a significant clue for recognition. Besides, experimenters also showed their preference for ignoring the small trivial parts comparing to the entire object. Since no consistent ignorance criterion can be concluded, the shape recognition and the smoothing parameter for contour fitting is obvious an egg and chicken problem, to solve which, we propose a ML-KNN based shape classification strategy to retrieve the shapes features with respective to different smoothing parameters. The processing steps include the contour extraction; shape feature computation based on spline fitting with different smoothing parameters, which is elaborated in section 3; recognition by the ML-KNN classifier to retrieve the final result, detail in section 4.



(a) Original images



(a) Smoothed Contour (blue), with same smoothing parameter

Figure.1 Images with different interpretations, left: general object; right: a key shaped object

3. Shape Feature Extraction

The contour of target object is first traced in counter-clockwise direction in order to facilitate the fitting computation. We then adopt cubic B-spline to perform parametric fitting to traced contour pixels' abscissa x and ordinate y respectively. Given the parameter t as a normalized common dependent variable of x and y, x(t) and y(t) functions are required to perform fitting computation by treating the traced contours' coordinates as control points. In order to achieve also the smoothing effect for trivia ignorance, the calculation is implemented by minimizing the following mixed energy function.

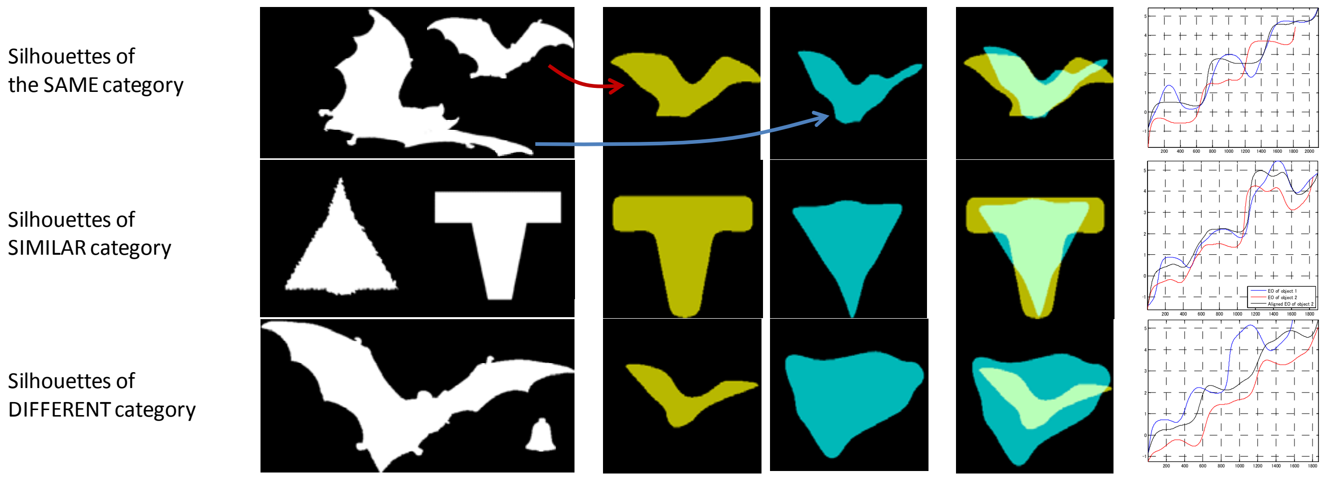
$$f = \underbrace{\operatorname{argmin}}_{f}\{\alpha E_{fit\,error}(f) + (1-\alpha)E_{bending}(f)\} \qquad (1)$$

where, α is the smoothing parameter, to indicate the balance between the preciseness of fitting to the control points (Eq.(2)), and the energy consumption over spline's bending (Eq.(3)) with respective to the second derivative of the fitted spline function.

$$E_{fit\,error}(f) = \sum_{i=1}^{n} w(i)|c(t_i) - f(t_i)|^2 \qquad (2)$$

$$E_{bending}(f) = \int \lambda(t)\left|\frac{d^2f(t)}{dt^2}\right|^2 dt \qquad (3)$$

where, f is the fitted spline function; c is either x or y function with respective to t; n is the total contour pixel number that involved; w(i) is the fit error weight function against the original object contour, and $\lambda(t)$ is the bending weight function, both weight function in this research are defined as uniform with value 1. Decreasing the smoothing parameter α within the range of [0 1] can generally construct the splines with less attention to the trivial parts, thus achieve smoothing effect.



a. Original silhouettes b. Resized and smoothed silhouettes with rotation alignment c. superposition d. EOCs matching

Figure 2, Rotation-Scale-Translation invariant realization though (1) regularizing object silhouette size, (2) minimal distance seeking. The original silhouettes demonstrate all the RST difference between two objects.

We select the edge orientation curve (eoc) as the shape descriptor of the target object contour. The eoc holds translation and scale-invariant feature which is desirable in shape recognition issue. However, from Eq.(1) and (3), one can identify that the fitted contour with certain smoothing parameter is sensitive to scale variation for the bending energy function. In order to solve this problem, the query images are cropped with minimum bounding circle, and resized to a fixed scale with object centroid locates at image center. T he important but not intrinsic feature that rotation invariant feature can be achieved by translational shifting one eoc against another to seek the minimal average error. The rotational difference can then be retrieved, specifically for an eoc with length n, the translational shift when reaching the minimal average error is computed by:

$$\Delta t = \underset{\Delta t}{\underline{\mathrm{argmin}}} \sum_{i=1}^{n} \left| g_1(t_i|\alpha) - P\big(g_2(t_i+\Delta t|\alpha)\big) \right| \qquad (4)$$

where, Δt is the translational shift, $P(\cdot)$ is extrapolation operation with continuity maintenance at both endpoints:

$$P\big(g(t)\big) = \begin{cases} g(t+n) - g(n) + g(1) & t < 0 \\ g(t) & 1 \le t \le n \\ g(t-n) + g(n) - g(1) & t > n \end{cases} \qquad (5)$$

Thus, the rotational difference between the two entry objects can be calculated as:

$$\theta = g_2(\Delta t|\alpha) - g_2(1|\alpha) \qquad (6)$$

4. Training and Classification

Traditional KNN classifies objects based on closest training examples in the feature space. We extend this algorithm by adopting the layer concept to achieve ML-KNN to solve the contour smoothing degree selection.

At the training stage, eoc of sample images as well as the manual labeling that contains classification and instructive smoothing parameter are input to the classifier. Then, the classifier constructs its layer structure by quantumizing the input smoothing parameters. At the automatic recognition stage, eoc features with different smoothing parameters that involved in the classifier are calculated. Find K nearest neighbors within all the layers respectively. The recognition result is then determined as the classification with major neighbors.

5. Experiments and Conclusion

We carry out the experiments against the image database of MPEG-7. As demonstrated in Fig.2, the above described eoc computation can measure shape's similarity and calibrate their rotation angle. Based on the proposed method, the classification result is given in Fig.3.

In this paper, we present an effective method for recognizing and classifying object shapes by taking advantage of smoothed contour's edge orientation curve. Recognition is achieved in a ML-KNN manner, for which the classifier is first trained by manually labeled sample images.

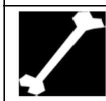
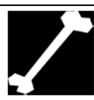
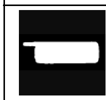
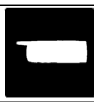
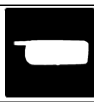
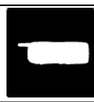
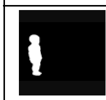
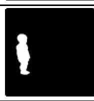
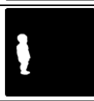
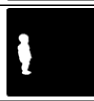
Query images	1st NN	2nd NN	3rd NN	Automatic classification result
				Bone
				Cellular Phone
				Children

Figure 3 Experimental results against the images from MPEG-7 database. "*NN*" stands for nearest neighbor. Right side of description of classification result is in accordance with human being's labeling in the training phase.

Reference

[1] Y. Xie, J. Ohya, "Elliptical Shaped Object Recognition via a Modified RANSAC with Edge Orientation Curve's Segmentation-Merge," *The Ninth IASTED International Conference on Visualization, Imaging and Image Processing (VIIP2009)*, Jul. 2009, to appear

[2] O.C. Hamsici, A.M. Martinez, "Rotation Invariant Kernels and Their Application to Shape Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16 Sept. 2008.