Received April 2, 2019, accepted April 16, 2019, date of publication April 29, 2019, date of current version May 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2913688

Curvature Bag of Words Model for Shape Recognition

JIEXIAN ZENG[1,2], MIN LIU[2], XIANG FU[1,2], RUIYU GU[2], AND LU LENG[1,2], (Member, IEEE)
[1]Key Laboratory of Jiangxi Province for Image Processing and Pattern Recognition, Nanchang Hangkong University, Nanchang 330063, China
[2]School of Software, Nanchang Hangkong University, Nanchang 330063, China

Corresponding author: Xiang Fu (fxfb163@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61763033, Grant 61662049, Grant 61741312, Grant 61866028, Grant 61881340421, Grant 61663031, and Grant 61866025, in part by the Key Program Project of Research and Development (Jiangxi Provincial Department of Science and Technology) under Grant 20171ACE50024 and Grant 20161BBE50085, in part by the Construction Project of Advantageous Science and Technology Innovation Team in Jiangxi Province under Grant 20165BCB19007, in part by the Application Innovation Plan (Ministry of Public Security of P. R. China) under Grant 2017YYCXJXST048, and in part by the Open Foundation of the Key Laboratory of Jiangxi Province for Image Processing and Pattern Recognition under Grant ET201680245 and Grant TX201604002.

ABSTRACT The object shape recognition of nonrigid transformations and local deformations is a difficult problem. In this paper, a shape recognition algorithm based on the curvature bag of words (CBoW) model is proposed to solve that problem. First, an approximate polygon of the object contour is obtained by using the discrete contour evolution algorithm. Next, based on the polygon vertices, the shape contour is decomposed into contour fragments. Then, the CBoW model is used to represent the contour fragments. Finally, a linear support vector machine is applied to classify the shape feature descriptors. Our main innovations are as follows: 1) A multi-scale curvature integral descriptor is proposed to extend the representativeness of the local descriptor; 2) The curvature descriptor is encoded to break through the limitation of the correspondence relationship of the sampling points for shape matching, and accordingly it forms the feature of middle-level semantic description; 3) The equal-curvature integral ranking pooling is employed to enhance the feature discrimination, and also improves the performance of the middle-level descriptor. The experimental results show that the recognition rate of the proposed algorithm in the MPEG-7 database can reach 98.21%. The highest recognition rates of the Swedish Leaf and the Tools databases are 97.23% and 97.14%, respectively. The proposed algorithm achieves a high recognition rate and has good robustness, which can be applied to the target shape recognition field for nonrigid transformations and local deformations.

INDEX TERMS Curvature bag of words model, linear support vector machine, shape recognition.

I. INTRODUCTION

Shape is an inherent feature in understanding an image, that remains stable in spite of changes in an object's illumination, color and texture. Due to these advantages, shape features have been widely applied in object recognition tasks. With the advancement of contour detection proposed by Arbelaez *et al.* in [1], shape-based object recognition has become more practical and has attracted more attention in the field of computer vision.

Zhou Yu *et al.* [2] stated that the difficult problems in shape matching mainly include the follows: 1) Shape differences caused by various transformations (such as affine transformations, etc.). 2) The occlusion of objects and the movement of nonrigid objects (such as the shape changes caused by a horse running). 3) Changes between similar objects (such as the differences between people who are tall and thin). 4) Image noise and segmentation error. These problems make it more difficult to identify the shape of the object. The existing algorithms are still unable to cope with the problem of object recognition under severe occlusion. The image segmentation error is caused by the segmentation algorithm not obtaining the edge of the object from the real image. To address these problems, many invariant descriptors can be used to ensure the validity of the description of the object transform, such as the affine transform descriptor [3], the projective transform descriptor [4], and so on [5], [6]. To address the problems of occlusion and distortion of an object, a local descriptor is usually used to identify the shape.

Shape-based object recognition is generally considered to be a classification problem that classifies the tested object's shape through a given set of shapes. The general process

The associate editor coordinating the review of this manuscript and approving it for publication was Haiyong Zheng.

of the traditional object recognition algorithm is as follows: To assess the object's shape, first, the shape feature descriptors are extracted. Then, a matching algorithm is applied to obtain the distances between the descriptors. Finally, the shape recognition classification is completed by sorting and metric learning. During this process, the shape-distinguishing ability of the descriptor directly affects the recognition result; thus, most shape recognition work has concentrated on descriptors. The contour-based shape representation method uses a set of points extracted along the shape boundary curve to describe the shape's features, which can be more suitable for shape recognition and retrieval [2]. The classic contour method is shape context [7], which employs the spatial distribution information of statistical sampling points as a descriptor. Based on the shape context, Ling and Jacobs [8] replaced the Euclidean distance between the contour points used in the shape context algorithm with the inner-distance between the contour points and proposed the inner-distance shape context (IDSC) method. Sun *et al.* [9] and Wang *et al.* [10] applied the height function of the sampling point as a descriptor. Not only is the calculation required for this approach is simple but also it greatly improved the recognition rate. However, this type of method, which is based on the global spatial information of sample points, becomes unstable when objects are occluded or deformed. Nanni *et al.* [11] proposed a method based on the local phase quantization matrix description to improve the retrieval performance of traditional descriptors, such as the shape context, inner-distance shape context and height function, but its effect is not satisfactory. Shekar and Pilar [12] used a combined classifier models of the height function and the two-dimensional discrete cosine transform to classify object shapes and achieve better results. Cui *et al.* [13] employed the curvature of the curve and the unsigned curvature integral to provide a curve descriptor, which samples the contour curve using the same curvature integral so that the sampling points are more concentrated on the curve, and the descriptor has a larger entropy value. Although their algorithm was a significant improvement with respect to computational complexity, and can match occluded objects, its descriptor is susceptible to noise. Fu *et al.* [14] defined a new affine curvature descriptor based on [13] that is invariant under affine transformation, however, this descriptor is not defined at the inflection point and is difficult to apply in practice. Matsuda *et al.* [15] proposed a descriptor for retrieving shapes under unpredictable conditions and constructed an angle-length profile as a basis for the normalization of the curvature partition, but the result didn't turn out very well.

Effective shape descriptors play an important role in object recognition, and multi-scale descriptors can reflect more abundant shape information. Yang *et al.* [16] defined a descriptor as three types of invariants in multiple scales, in order to capture shape features from different aspects. Zhang *et al.* [17] proposed to segment a shape into multi-scale ellipse, then used the four invariant ellipse features to describe the relationship. Their experimental results were highly accurate. However, their multi-scale techniques used only different sizes of circles or ellipses on the shape to extract the multi-scale contours, then used some invariants to represent the features of these contours.

The traditional shape-matching algorithm achieves matching between features by aligning the sampling points, but in practical applications, it is difficult to obtain a strict correspondence. Wang *et al.* [18] employed a contour bag of words model to break through the limitations of the basic feature descriptors and promote the underlying features to semantic features of the middle layer, but it has difficulty describing the deformation objects by using the shape context descriptor. Jia *et al.* [19] applied a projective invariant descriptor combined with the curvature-graded shape-coded bag of words model to achieve shape recognition under the projective transformation.

To address the problem that the traditional algorithm cannot deal with the object recognition under nonrigid transformation and local deformation, a shape recognition algorithm based on the curvature bag of words (CBoW) model is proposed in this paper. The algorithm uses a multiscale curvature integral descriptor combine with the bag of words model to achieve shape recognition. Fig. 1 shows a flow chart of the method, which includes the following steps: a) Extract the object contour by using the contour extraction algorithm. b) Approximate the polygon of the object contour by applying the discrete contour evolution (DCE) algorithm [20]. c) Decompose the shape contour into the contour fragments by using the vertices of the polygon. d) Sample the contour fragment with the same curvature integral by applying different scales, calculating the curvature and obtaining the curvature description of the curve. e) Code each contour fragment with the locality-constrained linear coding (LLC) [21] algorithm, and use the max-pooling method to obtain the shape feature description. f) The shape feature descriptor obtained after e). g) Classify the shape feature descriptor by using the linear support vector machine (SVM) [22].

In summary, the proposed CBoW has several innovations: 1) It improves the existing curvature descriptors and propose multi-scale curvature integral descriptor, so that the description ability of the local descriptor is expanded. 2) It encodes the basic curvature descriptors, break through the correspondence of the sampling points, and forms a middle-level semantic description of the feature. 3) It Combines with equal-curvature integral grading pooling, so the feature distinguishing is stronger and the middle-level descriptor is further improved.

II. CONTOUR DECOMPOSITION

Contour fragments, which contain local and global information, are regarded as shape features with strong descriptive abilities [23]. The CBoW model uses contour fragments as shape features to learn the shape codebooks and create shape representations. The purpose of the contour decomposition is to decompose the contour of the object into

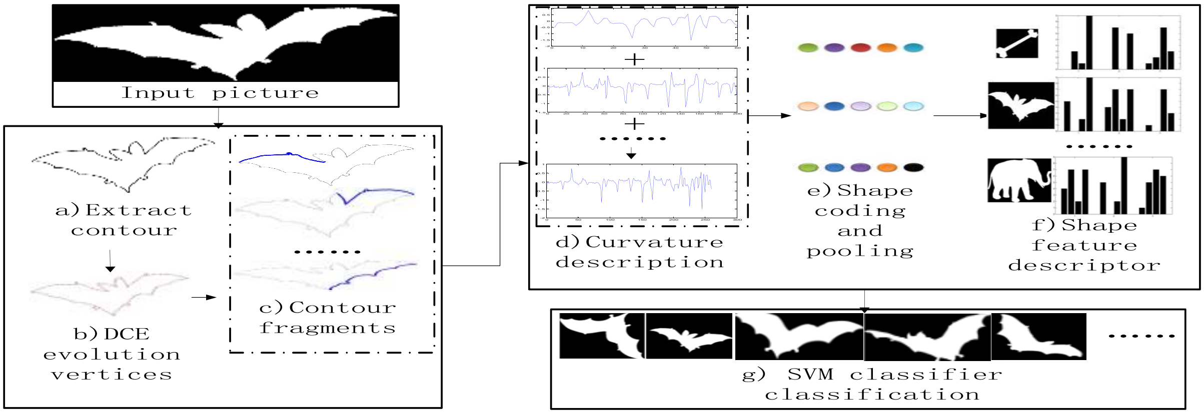

FIGURE 1. Algorithm flow chart.

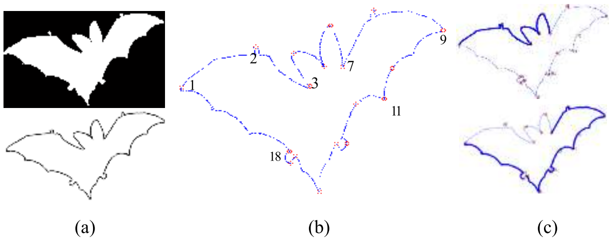

FIGURE 2. Target contour decomposition. (a) Target and contour; (b) DCE vertices; (c) c_{ij} and c_{ji}.

contour fragments. For binary images, the contour extraction algorithm is first applied to extract the object contour. Then, the DCE algorithm is used to obtain the approximate polygon of the object's contour. Finally, the vertices of the polygon are applied to decompose the shape contour into contour fragments.

A. CONTOUR EXTRACTION

Common contour extraction algorithms include edge detection methods, morphological methods, contour methods [24] and so on. The contour extracted by the edge detection method (such as the Canny operator) will retain some noise and cause some data information to be lost. The contours extracted by the morphological operator also retain some background noise and distort objects. The contour method can obtain more complete and ideal object contours and remove interference from noise in complex backgrounds. Therefore, this paper applies the contour method to perform contour extraction. Fig. 2(a) shows an object contour extracted by using the contour algorithm.

B. DCE EVOLUTION VERTEX

Assume that the object's shape contour is S, and we represent the shape contour S as $S(u) = (x(u), y(u))$. Then, we use the DCE algorithm to obtain a simplified polygon $\vec{p}$ with vertices on S. The representation of $\vec{p}$ is shown in (1):

$$\vec{p} = (p_1, \ldots., p_T). \tag{1}$$

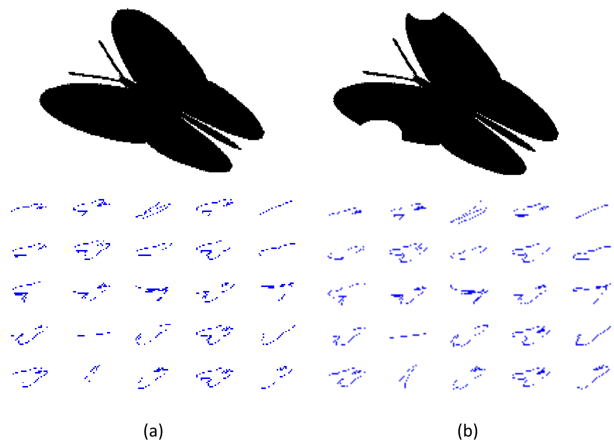

FIGURE 3. Part of the butterfly profile. (a) Complete butterfly picture and outline fragments; (b) Local deformed butterfly picture and contour fragment.

where T is the number of vertices, which is determined by parameter τ of the DCE algorithm, and $\vec{p}$ represents the polygon vertices on S.

Fig. 2(b) shows the vertices extracted from the input bat contour S by using the DCE algorithm.

C. CONTOUR FRAGMENT

For a given shape contour S and the vertices evolved by the DCE algorithm, the contour fragment set can be expressed as C(S). C(S) consists of the segments between each pair of vertices (p_i, p_j). Suppose that c_{ij} represents a contour segment between p_i and p_j. Then, we have the following:

$$C(S) = \{c_{ij} = (p_i, p_j); i \neq j; i, j \in [1, \ldots, T]\}. \tag{2}$$

Note that p_i and p_j are not necessarily neighboring vertices. Simultaneously, we have the following:

$$S = c_{ij} \cup c_{ji}. \tag{3}$$

Fig. 2(c) shows the contour segments of the vertices (1, 7) and vertices (7, 1).

Because the contour fragment between all the vertices are extracted, C(S) contains information on the very rich shape contours in S. As shown in Fig. 3, all the contour fragments extracted from one shape can effectively avoid differences caused by local deformations, and extract the same contour fragment that does not contain the deformation.

III. CURVATURE BAG OF WORDS MODEL

The bag-of-words (BOW) model is a commonly used document representation method in the field of information retrieval. In information retrieval, the BOW model assumes that a document, its word order, grammar, syntax and other elements are ignored, and it is only regarded as a collection of several words. The appearance of each word in the document is independent and does not depend on whether other words appear. In this paper, the contour segment is treated as a word.

In this section, we will study how to construct the CBoW model $f(S)$ for the shape contour segment set C(S) and classify the shape by using $f(S)$.

A. CURVATURE DESCRIPTION

Since the point coordinates of the contour are very unstable, they change with the transformation, rotation, scale scaling, etc. of the object, thus, the point coordinates of the contour cannot be used as a stable shape characteristic. The curvature is widely used to describe shapes as a similar invariant feature. This paper combines the research of Chai *et al.* [25] and Cui *et al.* [13] to propose a new curvature descriptor. We input the fragment c_{ij} and calculate its curvature feature descriptor x_{ij}.

The input contour fragment (curve) c_{ij} is parameterized by formula (4):

$$\Gamma(u) = (x(u), y(u)). \tag{4}$$

Then, the formula to calculate the curvature is shown in formula (5):

$$k(u) = (\dot{x}(u)\ddot{y}(u) - \ddot{x}(u)\dot{y}(u))/(\dot{x}^2(u) + \dot{y}^2(u))^{3/2}. \tag{5}$$

To eliminate the influence of noise, the curve c_{ij} is convoluted by using a Gaussian template $g(u, \sigma)$, and the curve after convolution is represented by (6):

$$\Gamma_\sigma = (x(u, \sigma), y(u, \sigma)). \tag{6}$$

Then, the curvature of the curve Γ_σ can be expressed as shown in (7):

$$k(u, \sigma) = \frac{X_u(u, \sigma)Y_{uu}(u, \sigma) - X_{uu}(u, \sigma)Y_u(u, \sigma)}{(X_u(u, \sigma)^2 + Y_u(u, \sigma)^2)^{3/2}}. \tag{7}$$

In the above formula, we have the following: $X_u(u, \sigma) = \partial x(u, \sigma)/\partial u^2$, $X_{uu}(u, \sigma) = \partial^2 x(u, \sigma)/\partial u^2$, $Y_u(u, \sigma) = \partial y(u, \sigma)/\partial u^2$, and $Y_{uu}(u, \sigma) = \partial^2 y(u, \sigma)/\partial u^2$.

As shown in Fig. 4, the curvature feature descriptor x_{ij} is calculated using the following steps:

(1) To standardize the arc length, this paper first resamples c_{ij}, and use p equidistant points, and then performs Gaussian smoothing to obtain the curve Γ_σ.

(2) Calculate the curvature k of Γ_σ according to (7) to obtain an S-k diagram.

(3) Integrate the curvature to obtain graphs of $\int |k| - s$ and $\int |k| - k$.

(4) For the graph of $\int |k| - k$, use a different step size $L = (\sum k)/N$ (where N is the number of sampling points, and $N<p$) to sample, obtain the point set Γ' corresponding to the sampling curvature, and calculate the curvature k_N of Γ' according to (7).

(5) Conduct multiscale sampling. In order to avoid attenuating the number of sampling points too fast, this article's methods have been tested many times. Suppose $N_{i+1} = 0.75N_i$. Repeat step (4) t times to obtain the curvature feature descriptor:

$$x_{ij} = [k_{N_1}, k_{N_2} \ldots k_{N_t}]. \tag{8}$$

Chai *et al.* [25] performed a Gaussian template on the curve before calculating the curvature. Cui *et al.* [13] proved that the curvature integral sampling points are stable to similar transformations. Furthermore, our algorithm makes use of curvature integral multiscale sampling on the basis of the previous two methods. And because Gaussian template could be used to eliminate Gaussian noise, we considered our descriptors to have the robustness against noise. Fig. 4 shows the steps involved in calculating the curvature descriptor.

B. SHAPE CODING AND POOLING

In the visual BoW model, the representation of the underlying features to the middle and high levels is achieved by coding [18]. In the CBoW model, the feature descriptor x_{ij} is encoded and mapped to a new space formed by the shape code book B. In the new space, the contour segment is represented as the shape code w_{ij}.

There are many kinds of codebook learning methods, which are mainly divided into supervised learning and unsupervised learning. Supervised learning requires training sets and testing samples, find rules in the training set and use these rules for the testing samples. Unsupervised learning requires no training set and only a set of data that looks for rules in the set of data. The unsupervised learning k-means [26] algorithm is stable, its spectral clustering effect is good, and its hierarchical clustering is fast. Therefore, the k-means algorithm is used as the method to learn codebook. The codebook is obtained by clustering the selected features using the k-means algorithm:

$$B = [b_1, b_2, \ldots b_M] \in R^{t \times M}. \tag{9}$$

Each column in the above formula is a cluster center, and M is the number of cluster centers.

This article uses the LLC algorithm for encoding. LLC introduces local constraints and operates at a higher speed. To express x_{ij} in the space formed by the codebook B, the LLC uses the l codewords(cluster centers) closest to x_{ij} in B to form a local coordinate system, which is denoted as $B_{\pi_{ij}} \in R^{t \times l}$. Here, π_{ij} is a set of recent codeword indices in B, denoted as $\pi_{ij} = \{\pi_{ij}^1 \ldots . \pi_{ij}^l\}$. $B_{\pi_{ij}}$ is the matrix space consisting of the $\pi_{ij}^1 \ldots . \pi_{ij}^l$ column in B. The feature descriptor x_{ij} can be

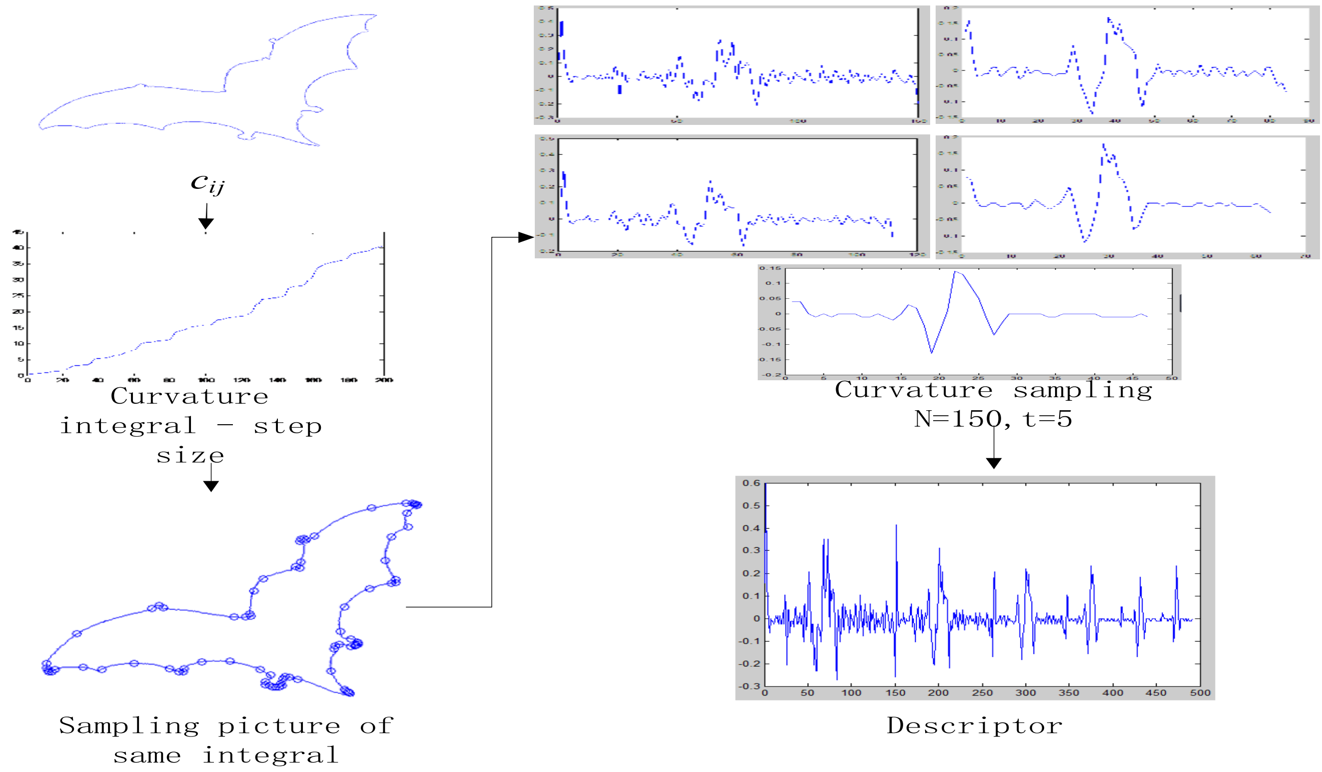

FIGURE 4. The calculation process of the curvature feature descriptor.

represented by the space of $B_{\pi_{ij}}$. The coordinate $w_{\pi_{ij}} \in R^{l\times 1}$ of x_{ij} in the space of $B_{\pi_{ij}}$ can be obtained by solving the objective function in (10):

$$\min_{\mathrm{w}_{\pi_{ij}}} \left\| x_{ij} - B_{\pi_{ij}} \mathrm{w}_{\pi_{ij}} \right\|^2 \quad s.t.\ 1^T \mathrm{w}_{\pi_{ij}} = 1. \tag{10}$$

where $w_{\pi_{ij}}$ is the spatial coordinate of x_{ij} and $B_{\pi_{ij}}$. In the experiment, l was set to 5. $w_{\pi_{ij}}$ has only l columns. Now, we can represent x_{ij} as $w_{ij}(t \times 1)$, where the value of column $\pi_{ij}^1 \ldots .\pi_{ij}^l$ corresponds to the corresponding value in $w_{\pi_{ij}}$, and the other codewords are set to zero.

There are many contour segments after shape coding. If the DCE algorithm evolves contours to obtain T vertices, then the number of divided contour segments is $T!/(T-2)!$ (where $T!$ is the factorial of T). Large number of contour segments result in high computational costs, and in order to obtain a compact representation of the contour, the shape contour needs to be pooled. Wang *et al.* [18] used the largest pooling method to obtain the contour representation of 21 regions. Jia *et al.* [19] used the curvature grading pool to obtain contour representations of 21 different curvature integral levels. Based on the method of Jia *et al.* [19], this paper performs equal-curvature integral grading pooling.

The size of each value in the contour shape variable w_{ij} represents the response from the corresponding codeword, and the maximum response on each codeword can be obtained by using max-pooling. Let w^z denote the codeword response of the curvature integral level at SS. Then max-pooling can be expressed as follows:

$$f(S, r) = \max(\mathrm{w}^z | z \in SS_r). \tag{11}$$

where r is the division level of the curvature integral.

By merging the maximum responses in all levels, the feature vector representing the CBoW model is as follows:

$$f(S) = \left[f(S,1)^T, \ldots, f(S,21)^T \right]^T. \tag{12}$$

IV. SVM SHAPE CLASSIFICATION

The outline shape representation obtained by the method in Section 2.2 is a vector. In this paper, the multiclass support vector machine is used for shape classification. Given the shape training set $\{f_i\}$ with the label $\{y_i \in [1\ldots.H]\}$, H is the number of training categories. The [27] summarized the multicategory SVM, which solves the following optimization problem:

$$\min_{\omega_1 \ldots . \omega_h} \sum_{h=1}^{H} \|\omega_h\|^2 + \lambda \sum_i \max(0, 1 + \omega_{r_i}^T f_i - \omega_{y_i}^T f_i). \tag{13}$$

In the above formula $r_i = \arg\max_{h\in[1\ldots.H], h\neq y_i} \omega_h^T f_i, \omega_h^T f_i$ obtains the variable h corresponding to the maximum value, f_i is the shape descriptor, the left side of the plus sign is a regular term, the right side is the loss function, and the parameter λ is the penalty factor. The off-shelf vector machine solver proposed in [28] can solve the optimization problem in (13). When testing, the shape classification formula is as follows:

$$\hat{y} = \arg\max_{h\in[1\ldots.H]} \omega_h^T f. \tag{14}$$

V. EXPERIMENTS AND ANALYSIS

To verify the effectiveness of the proposed algorithm, the performance of the algorithm was verified on the MPEG-7 database, the Swedish leaf database and the Tools database.

In this section, the parameters involved in the algorithm are as follows: the number of resampling points p of the contour, the number of points of the curvature integral sampling point N, the number of sampling times t, and the number of clustering centers of the codebook M. To select the optimal parameter settings, five categories of pictures are selected from the MPEG-7 image library to form a small image library. First, the test is performed on the database, and then the appropriate parameter settings are selected according to the experimental results. Fig. 5 shows the selected target categories.

FIGURE 5. Image categories selected in MPEG-7.

To make the sampling contour conform better to the true contour, the resampling point p of the contour is set to 400. Table 1 presents the retrieval accuracy rate in the case of $M = 1500$, the number of integral points N of different curvatures, and the number of sampling times t.

As shown in Table 1, more sampling points lead to a higher accuracy rate, but the required calculation time also increases. When $N = 160$ and $t = 5$, the accuracy rate is the highest. Furthermore, the dimension of the descriptor is lower than those of $N = 180$, $t = 6$ and $N = 200$, $t = 6$, and thus, the parameters N and t of the descriptor in this paper are set to $N = 160$ and $t = 5$, respectively.

TABLE 1. Recognition rate under different parameters.

$t \backslash N$	100	130	160	180	200
1	91.1%	97.2%	93.9%	96.1%	97.8%
2	96.7%	96.1%	96.1%	97.8%	98.9%
3	95.0%	97.2%	97.2%	97.8%	98.9%
4	96.7%	98.9%	98.9%	98.9%	97.8%
5	98.9%	98.3%	99.4%	98.9%	98.3%
6	96.7%	99.4%	98.9%	99.4%	99.4%

Table 2 shows the accuracy rate of different codebook parameters M when $N = 160$ and $t = 5$. As shown in Table 2, when the value of M is 1500, the experimental effect is the best. Therefore, the codebook clustering center M is set to 1500.

A. MPEG-7 DATASET

The MPEG-7 CE-Shape-1 Part B dataset has a wide range of applications in the field of shape analysis. It is primarily

TABLE 2. The effect of codebook size on recognition results.

M	600	1000	1400	1500	1800
Recognition rate	95.5%	97.7%	98.3%	99.4%	98.3%

used to measure the retrieval accuracy based on the similarity method. MPEG-7 CE-Shape-1 Part B has 70 directories, each directory contains a type of shape, and each shape has 20 different shapes. Therefore, MPEG-7 contains a total of 1400 shapes. Fig. 6 shows part of the shapes in the MPEG-7 dataset.

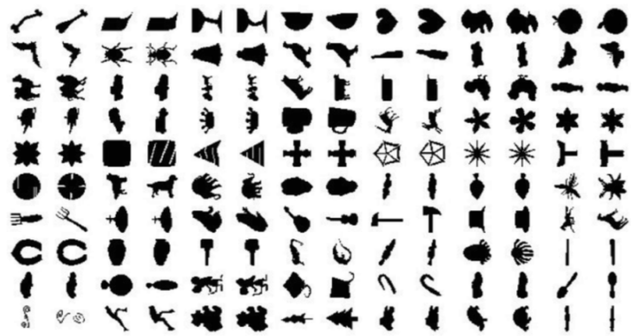

FIGURE 6. The part of the shapes in the MPEG-7 dataset.

Table 3 presents the evaluation results on the MPEG-7 CE-Shape-1 Part B dataset when using the leave-one-object-out evaluation method [2].

TABLE 3. Comparison of the recognition rates on the MPEG-7 dataset.

Algorithm	Recognition rate (%)
Projective transformation +BCF [19]	96.29
String of symbolic [29]	97.36
Polygonal multiresolution [30]	97.57
Class segment set [31]	97.93
BCF [18]	98.28
CBoW	98.21

It can be seen from Table 3 that the recognition rate of the proposed method is higher than other algorithms and slightly lower than the BCF algorithm. The reason why the recognition rate is slightly lower than the BCF algorithm is that the proposed algorithm uses the curvature as a descriptor and has rich local information, but it does not contain spatial position information, which has a slight influence on the recognition rate. However, we can see from the following experiments that the descriptor of CBoW model is more stable for shape deformation.

B. SWEDISH LEAF DATASET

The Swedish Leaf database is mainly used to evaluate the robustness of the characterization algorithm to small local deformation of the contour. As shown in Fig. 7, the database has 15 different categories, each containing 75 samples, for a total of 1125 Swedish leaf images.

FIGURE 7. Partial sample of the Swedish Leaf dataset.

It can be seen from Table 4 that the method in this paper achieves extremely high recognition accuracy even when only the contour features are used, regardless of the blade texture or color characteristics. This result is due to the good robustness of the curvature local descriptor to the local deformation of the contour.

TABLE 4. Comparison of the recognition rates on the swedish leaf dataset.

Algorithm	Recognition rate (%)
SC+DP [7]	88.12
IDSC+DP [8]	94.13
EHF+DP [9]	95.07
MARCH [10]	96.21
BCF [18]	96.56
CBoW	97.23

C. TOOLS IMAGE SET

Many categories of articulated shapes in the Tools image dataset vary widely, and the dataset is primarily to test the performance of the algorithm on nonrigid objects. As shown in Fig. 8, the Tools database contains 7 similar shapes. The statistical method of the recognition rate of the Tools database takes an arbitrary image as the shape to be retrieved, compares and sorts all the shapes, finds the top 5 most similar shapes, and finally counts the number of these shapes and the shapes to be identified belonging to the same category. Because the dataset is small, the effect of using the SVM classifier is often 100%. To compare the results of other algorithms, this paper uses the DP algorithm to match on this dataset. Table 5 shows the test results of the algorithm and other shape recognition algorithms on this the dataset.

The recognition rate of this method using the Tools dataset reaches 97.14%, which is better than other algorithms. It shows that the proposed algorithm is highly robustness to nonrigid transform object recognition.

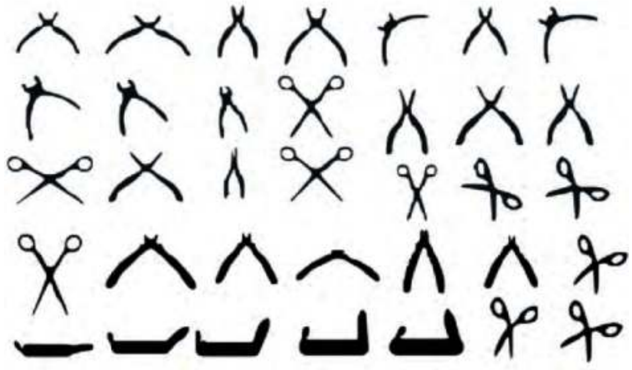

FIGURE 8. Tools dataset.

TABLE 5. Comparison of recognition rates on tools.

Algorithm	Recognition rate (%)
IDSC+DP [8]	56.57
LPQ+DP [11]	85.71
HF+DP [12]	89.71
EHF [9]	94.86
CBoW+DP	97.14

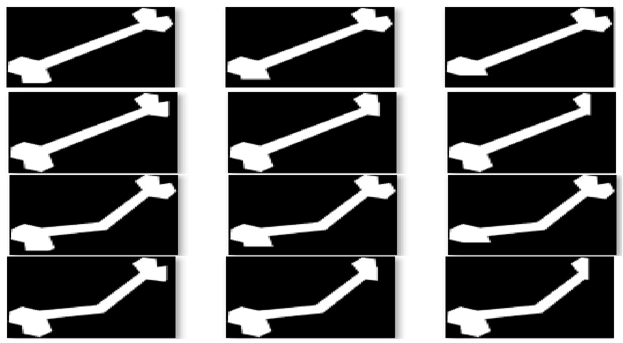

FIGURE 9. Occlusion image.

TABLE 6. Recognition rates on occluded images.

Occlusion	Bottom 5%	Bottom 10%	Bottom 15%	Right 5%	Right 10%	Right 15%
Recognition rate	92.56%	90.13%	87.25%	91.92%	90.34%	86.53%

D. OCCLUSION EXPERIMENT

Our descriptor of the CBoW model belongs to the local descriptor, which can handle occluded situations. In order to verify the robustness of the algorithm against occlusion, in this paper, the MPEG-7 CE-Shape-1 Part B image is occluded at the bottom of 5%, 10% and 15%, and the right of 5%, 10% and 15%, and the obtained images are shown in Fig. 9. Table 6 shows the test results of the algorithm under different occlusion situations.

The method of this paper maintains a good recognition rate even in the case of severe occlusion. This result shows that the proposed algorithm is highly robustness to occluded object recognition.

VI. CONCLUSION

To effectively improve the robustness of the visual BoW model in nonrigid transformation and local deformation targets, a curvature-based local feature descriptor is proposed to represent the contour segment features, and the BoW model is used for identification and classification. The CBoW model algorithm effectively improves the recognition performance and robustness for nonrigid transformation and local deformation targets. The main contributions and originality of this paper are as follows: 1) A multi-scale curvature integral descriptor is proposed to expand the description ability of local descriptors. 2) The curvature descriptor is encoded to break through the restriction of the corresponding relation of the shape-matching sampling points, thus forming the feature of intermediate semantic description. 3) Using the equal-curvature integral ranking pooling to enhance feature recognition and improve the performance of middle-level descriptor. Subsequent research, we will explore how to effectively apply the algorithm to contour-based natural images and 3D image retrieval.

REFERENCES

[1] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.

[2] Y. Zhou, J. T. Liu, and X. Bai, "Research and perspective on shape matching," *Acta Autom. Sinica*, vol. 38, no. 6, pp. 889–910, 2012.

[3] W. Wang, B. Xiong, X. Yan, Y. Jiang, and G. Kuang, "Affine invariant shape projection distribution for shape matching using relaxation labelling," *IET Comput. Vis.*, vol. 10, no. 2, pp. 124–133, 2016.

[4] Q. Jia, Y. Liu, and X. Fan, "Hierarchical shape feature extraction and matching algorithm based on cross ratio context," *J. Comput.-Aided Des. Comput. Graph.*, pp. 2247–2255, Dec. 2015.

[5] F. Mokhtarian, S. Abbasi, and J. Kittler, "Efficient and robust retrieval by shape content through curvature scale space," *Image Databases Multi-Media*, pp. 51–58, 1998.

[6] N. Alajlan, I. El Rube, M. S. Kamel, and G. Freeman, "Shape retrieval using triangle-area representation and dynamic space warping," *Pattern Recognit.*, vol. 40, no. 7, pp. 1911–1920, 2007.

[7] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," Dept. Comput. Sci. Eng., Univ. California San Diego, San Diego, CA, USA, Tech. Rep. 4, 2002.

[8] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.

[9] G.-D. Sun, Y. Zhang, P. Li, S.-Z. Mei, and D.-X. Zhao, "Feature description of exact height function used in fast shape retrieval," *Opt. Precis. Eng.*, vol. 25, no. 1, pp. 224–235, 2017.

[10] B. Wang, D. Brown, Y. Gao, and J. La Salle, "MARCH: Multiscale-arch-height description for mobile retrieval of leaf images," *Inf. Sci.*, vol. 302, pp. 132–148, May 2015.

[11] L. Nanni, S. Brahnam, and A. Lumini, "Local phase quantization descriptor for improving shape retrieval/classification," *Pattern Recognit. Lett.*, vol. 33, no. 16, pp. 2254–2260, 2012.

[12] B. H. Shekar and B. Pilar, "Discrete cosine transformation and height functions based shape representation and classification," *Procedia Comput. Sci.*, vol. 58, pp. 714–722, 2015.

[13] M. Cui, J. Femiani, J. Hu, P. Wonka, and A. Razdan, "Curve matching for open 2D curves," *Pattern Recognit. Lett.*, vol. 30, no. 1, pp. 1–10, 2009.

[14] H. Fu, Z. Tian, M. Ran, and M. Fan, "Novel affine-invariant curve descriptor for curve matching and occluded object recognition," *IET Comput. Vis.*, vol. 7, no. 4, pp. 279–292, Aug. 2013.

[15] Y. Matsuda, M. Ogawa, and M. Yano, "Shape retrieval with geometrically characterized contour partitions," *IEEE Access*, vol. 3, pp. 1161–1178, 2015.

[16] J. Yang, H. Wang, J. Yuan, Y. Li, and J. Liu, "Invariant multi-scale descriptor for shape representation, matching and retrieval," *Comput. Vis. Image Understand.*, vol. 145, pp. 43–58, Apr. 2016.

[17] X. Zhang, J. Xiang, and S. Xiong, "Shape retrieval using multiscale ellipse descriptor," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2017, pp. 1042–1046.

[18] X. Wang, B. Feng, X. Bai, W. Liu, and L. J. Latecki, "Bag of contour fragments for robust shape classification," *Pattern Recognit.*, vol. 47, no. 6, pp. 2116–2125, 2014.

[19] Q. Jia, M. Y. Yu, X. Fan, X. K. Gao, and H. Guo, "Shape coding and recognition method based on curvature classification," *Chin. J. Comput.*, vol. 40, no. 142, 2018.

[20] L. J. Latecki and R. Lakämper, "Convexity rule for shape decomposition based on discrete contour evolution," *Comput. Vis. Image Understand.*, vol. 73, no. 3, pp. 441–454, 1999.

[21] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3360–3367.

[22] D. Bryner, E. Klassen, H. Le, and A. Srivastava, "2D affine and projective shape analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 998–1011, May 2014.

[23] X. Bai, W. Liu, and Z. Tu, "Integrating contour and skeleton for shape classification," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops*, Sep./Oct. 2009, pp. 360–367.

[24] X. Liu, R. Shi, and D. Wang, "Research on contours of vehicles based on principle of contour," *Comput. Eng. Des.*, vol. 29, no. 13, pp. 3496–3498, 2008.

[25] X.-T. Chai, J.-Z. Liang, and L.-Y. Lang, "Full affine curve in curvature scale space image registration," *Comput. Sci.*, vol. 43, no. 1, pp. 53–56, 2016.

[26] M. Alhawarat and M. Hegazi, "Revisiting k-means and topic modeling, a comparison study to cluster arabic documents," *IEEE Access*, vol. 6, pp. 42740–42749, 2018.

[27] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," *J. Mach. Learn. Res.*, vol. 2, pp. 265–292, Dec. 2001.

[28] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *J. Mach. Learn. Res.*, vol. 9, pp. 1871–1874, Jun. 2008.

[29] M. R. Daliri and V. Torre, "Shape recognition and retrieval using string of symbols," in *Proc. 5th Int. Conf. Mach. Learn. Appl.*, Dec. 2006, pp. 101–108.

[30] E. Attalla and P. Siy, "Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching," *Pattern Recognit.*, vol. 38, no. 12, pp. 2229–2241, 2005.

[31] K. B. Sun and B. J. Super, "Classification of contour shapes using class segment sets," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2005, pp. 727–733.

JIEXIAN ZENG received the master's degree in engineering from Northwestern Polytechnical University, in 1997. He is currently a Professor with Nanchang Hangkong University. He has published over 100 papers. His research interests include image processing, pattern recognition, and computer vision research.

MIN LIU received the bachelor's degree from Nanchang Hangkong University, Jiangxi, China, where she is currently pursuing the master's degree. Her research interests include computer vision, image processing, and pattern recognition.

XIANG FU received the Ph.D. degree from Xidian University, in 2008. He is currently an Associate Professor with Nanchang Hangkong University. His research interests include computer vision, image processing, and pattern recognition.

RUIYU GU received the bachelor's degree from Hexi University, Gansu, China, and the master's degree from Nanchang Hangkong University, in 2018.

LU LENG received the Ph.D. degree from Southwest Jiaotong University, Chengdu, China, in 2012.

He did his Postdoctoral Research with Yonsei University, Seoul, South Korea, and Nanjing University of Aeronautics and Astronautics, Nanjing, China. He was a Visiting Scholar with West Virginia University, USA. He is currently an Associate Professor with Nanchang Hangkong University. He has published more than 60 international journal and conference papers, and been granted several scholarships and funding projects in his academic research. His research interests include image processing, biometric template protection, and biometric recognition.

Dr. Leng is a member of the Association for Computing Machinery (ACM), China Society of Image and Graphics (CSIG), and China Computer Federation (CCF). He is the Reviewer of several international journals and conferences.