IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 17, 2019, accepted December 4, 2019, date of publication December 9, 2019,
date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958416

Plant Species Recognition Using Triangle-Distance Representation

CHENGZHUAN YANG[1] AND HUI WEI[2]

[1]School of Information, Zhejiang University of Finance and Economics, Hangzhou 310018, China
[2]Laboratory of Cognitive Algorithm and Model, Shanghai Key Laboratory of Data Science, School of Computer Science, Fudan University, Shanghai 201203, China

Corresponding author: Chengzhuan Yang (chengzhuanyang@zufe.edu.cn)

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LQ19F020003, in part by the National Natural Science Foundation of China under Grant 61375122 and Grant 61902159, and in part by the Shanghai Science and Technology Development Funds under Grant 13dz2260200 and Grant 13511504300.

ABSTRACT Plant species recognition using leaf images is a highly important and challenging issue in botany and pattern recognition. A center problem of this task is how to accurately extract leaf image characteristics and quickly calculate the similarity between them. This article presents a new shape description approach called triangle-distance representation (TDR) for plant leaf recognition. The TDR descriptor is represented by two matrices: a sign matrix and a triangle center distance matrix. The sign matrix is used to characterize the convex/concave property of a shape contour, while the triangle center distance matrix is used to represent the bending degree and spatial information of a shape contour. This method can effectively capture the detailed and global characteristics of a leaf shape while keeping the similarity transformations (translation, rotation, and scaling) unchanged. In addition, this method is quite compact and has low computational complexity. We tested our method on four standard plant leaf datasets, including the famous Swedish, Smithsonian, Flavia, and ImageCLEF2012 datasets. The results confirm that our approach exceeds the prior state-of-the-art shape-based plant leaf recognition approaches. An extra experiment on the MPEG-7 shape dataset further shows that our method can be applied to general shape recognition.

INDEX TERMS Plant species recognition, shape matching, triangle-distance representation, shape descriptor.

I. INTRODUCTION

Plants provide us with oxygen and fuel, and they also play a significant role in maintaining the balance of the Earth's ecosystem. Thus, it is necessary to identify plant species to maintain biodiversity. There are a large number of plant species present on Earth [1], and accurate identification of these large numbers of plant species is a very challenging task because plant species identification requires specialized knowledge and in-depth training related to botany. Therefore, there is an urgent need to develop an automatic plant leaf recognition system. This is not only useful for general use, but also helpful to experienced botanists and plant ecologists.

Plant species identification usually involves observation of some morphological characteristics of plants, such as stems, leaves, fruits, flowers, and embryos. A review of several existing plant species recognition methods can be found in the literature [2], [3]. Among these features, leaves are considered to be very valuable clues in identifying plant species. This is because the leaves will remain on the plants for at least several months. Other organs, such as flowers and fruits, will change greatly in a short time. Thus, there has been a large amount of work carried out in identifying plant species using leaf images in recent years [4]–[14]. Leaves are usually described by shape, color, or texture, while the color of leaves changes by season or geographical location, and different plant species usually have the same leaf color, such as green or yellow. As a result, the color of a leaf is not a reliable factor in the identification of plant species. Thus, only shape and texture information are usually considered in similarity based plant leaf recognition schemes. Compared with the texture information of leaves, shape is a high-level visual feature and remains constant for illumination and object deformation. This article focuses on the shape of leaf images for plant

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

species identification, because the shape of leaf images contains a significant amount of valuable information that can be used to distinguish different types of plant species. Therefore, the shape of leaf images is often used as an important feature of plant species recognition. However, it is a highly challenging problem for retrieving similar leaf shapes from a leaf image database. There are two main difficulties: (1) Nowadays, image databases are becoming increasingly larger, and efficient retrieval from such large-scale leaf image databases has become very challenging (e.g., there are approximately 40 million plant species on Earth, and millions of plant leaf images are stored in such databases); (2) The shapes of leave images have large intra-class changes and small inter-class similarities, thus affecting the accuracy of retrieval. Fig. 1(a) shows that the shapes of different plant species have smaller inter-class differences, and Fig. 1(b) shows that the same species of leaves exhibit large intra-class changes.

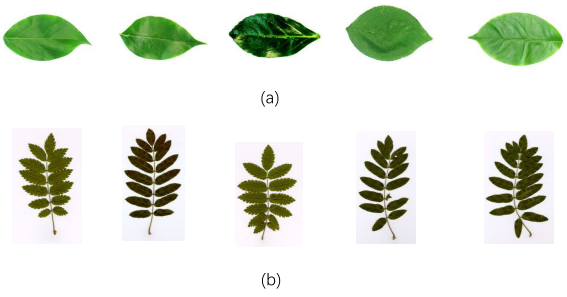

FIGURE 1. Shape of plant leaves having a large deformation: leaf sample images with (a) smaller inter-class distance and (b) larger intra-class differences.

In this article, a new shape description approach called triangle-distance representation (TDR) is presented for plant leaf recognition. This method can maintain retrieval efficiency while achieving better retrieval accuracy, so that the two above-mentioned difficulties can be well handled. The proposed TDR descriptor is described by two matrices: a sign matrix and a triangle center distance matrix. The sign matrix is used to characterize the concave and convex properties of the leaf shape, while the triangular center distance matrix is used to characterize the curvature and spatial information of the shape. The TDR descriptor can combine the global and detailed features of a leaf shape well, and it is invariant to similarity transformations (translation, rotation, and scaling). In addition, the TDR has a lower feature dimension and is a compact shape descriptor. In the stage of shape matching, we only use the L_1 distance to calculate the dissimilarity between shapes, and the shape matching efficiency is very high. Therefore, our method is very suitable for the tasks of large-scale image retrieval and real-time application. We performed experiments on four standard plant leaf datasets, including the well-known Swedish, Smithsonian, Flavia, and ImageCLEF2012 datasets. The results confirm that our approach exceeds the prior state-of-the-art shape-based plant leaf recognition approaches. An extra experiment on the MPEG-7 shape dataset further shows that our method can be applied to general shape recognition.

The rest of this article is structured as follows: In Section 2, we review several existing shape-based plant leaf recognition methods. In Section 3, we introduce the proposed TDR method in detail. The experimental results of our method on several datasets are given in Section 4. Section 5 concludes the article.

II. RELATED WORK

Plant species recognition using leaf images has recently attracted great attention from researchers in pattern recognition and botany [7], [10], [15]–[22]. This task can be regarded as a special case of the more general image classification problem and has been extensively studied in botany and pattern recognition. There are already some works that use only shape information for plant species identification. Below, we briefly review several important methods for shape-based plant leaf recognition.

Most shape-based plant species identification methods use geometric features and descriptors to describe the shape information of leaf images. The first class of methods extracts the morphological characteristics of plant leaves, such as aspect ratio, solidity, and convex area ratio. For example, Wu *et al.* [23] used aspect ratio, narrow factor, form factor, and rectangularity as features and adopted probabilistic neural network as a classifier for the identification of plant leaf. Belhumeur *et al.* [15] presented a plant leaf classification framework using complicated background. In their method, the Hu and Zernike moments serve as shape features, and then the moving center hypersphere (MCH) classifier is employed to identify plant species. Caballero and Aranda [24] combined geometric features with contour descriptors for efficient plant leaf identification. Cerutti *et al.* [25] presented a model-based approach for plant species recognition using high-level geometrical features and a mobile plant species identification application was implemented. Aptoula and Yanikoglu [26] proposed two morphological leaf descriptors: the first is a morphological covariance matrix based on leaf contours, and the second is a circular covariance histogram based on leaf vein structure. They combined these two descriptors and achieved very good recognition performance on the imageCLEF2012 dataset. Kalyoncu and Toygar [8] first proposed some new shape features to characterize the margins of plant leaves, and then combined this feature with geometric features and other shape descriptors for plant leaf classification. However, the drawback of this type of the method is that it is difficult to accurately extract the details of plant leaves, especially for plant leaves with large deformations. Thus, the accuracy of these methods in identifying plant leaves is relatively low.

The second class of methods makes use of shape descriptors for the identification of plant leaves. Soderkvist [27] built a tree-structured plant leaf classification system by combining curvature scale space, moment features and Fourier descriptors. They then tested this system on Swedish the leaf dataset and obtained a good recognition accuracy. Ling [28] first presented a new inner-distance shape context (IDSC)

descriptor, which takes the place of the Euclidean distance used in the shape context (SC) [29] descriptor with the inner distance. They then used the IDSC descriptor for plant leaf recognition and achieved good recognition accuracy. Since the dynamic programming method is used for shape matching, the computational complexity of the approach is extremely high. Mouine *et al.* [30], [31] proposed an extended SC approach for plant species recognition. They first considered two sources of information about the leaf: the leaf margin and leaf salient points. Then, two shape context descriptors are separately built for the leaf boundary and the leaf salient points. Finally, they combined these two shape description methods for plant leaf recognition. In [32], a shape-tree shape descriptor was proposed that represents the hierarchical geometric propensities of a shape. They then employed this descriptor for plant leaf classification, and obtained a higher recognition rate on Swedish leaf dataset. Similarity, this method also has a very high computational complexity. Hearn [33] presented an automated plant leaf recognition method by combining the Fourier descriptor and Procrustes analysis. Hu *et al.* [4] proposed a shape descriptor called the multiscale distance matrix (MDM) and applied it for the identification of plant leaves. This method is very efficient at characterizing the geometric properties of a shape and remains unchanged for translation, rotation, scale, and bidirectional symmetry of a shape. Kumar *et al.* [16] built a mobile app for automatic plant species recognition and obtained good performance on real-world leaf images. Zhao *et al.* [7] proposed a pattern counting approach for the identification of plant leaves. They first proposed a counting-based shape descriptor, and then used it for plant leaf recognition. However, the downside of this method is that it must use specialized training set learning to generate codebooks or dictionaries, and then uses them to generate shape descriptors. Zhang *et al.* [12] presented a plant species identification approach by combining sparse representation and singular value decomposition. This method requires neither the construction of specific classification features nor a training process. At the same time, this method has a higher efficiency. However, the recognition accuracy of this method is relatively low.

Recently, Yang *et al.* [10] proposed an efficient shape-based plant leaf recognition method using a multiscale triangular centroid distance (MTCD) descriptor. This method has obtained good recognition results on many standard leaf datasets and has very high retrieval efficiency at the same time. However, this method still has a high feature dimension and a low recognition accuracy. To further enhance the performance of plant leaf recognition and reduce the dimension of the shape features, the current work expands the MTCD method [10] from the following three perspectives. (1) Unlike the MTCD method, which uses many scale levels ($\lfloor (N-1)/2 \rfloor$), our method uses only the logarithmic distance as the scale level ($log_2(N/2)$), where N denotes the number of sample points of a leaf shape. The advantage of this improvement is to reduce the dimension of the feature and increase the speed of convergence, thereby improving the efficiency of retrieval. (2) Our method modifies the distance in the MTCD method to a signed distance. The sign of the distance is used to reflect the concavity and convexity of the shape contour, and the size of the distance can reflect the bending degree and spatial information of the shape contour. Based on these two kinds of information, we constructed two matrices to represent the shape, that is, a sign matrix and a triangle center distance matrix. The advantage of this improvement is that it can enhance the distinguishing ability of different types of shapes and improve the accuracy of plant leaf recognition. (3) In the shape matching stage, we use a simple L_1 distance instead of the cosine distance to measure the difference between the two shapes. These enhancements can be seen from our experimental results.

III. PROPOSED TDR METHOD

In this section, the proposed TDR approach for plant species identification is presented in detail. First, we give the definition of the proposed TDR. Second, we describe the shape similarity measure used for shape matching. Third, we analyze the computational complexity of TDR for the identification of plant leaves. Finally, the properties of the proposed method are discussed.

A. DEFINITION OF TDR

A shape boundary S can be represented by N sample points $S = \{P_i(x_i, y_i), i = 1, \ldots, N\}$, which is obtained by tracking and uniformly sampling the outer contour of a leaf shape in an anti-clockwise direction. P_1 is the starting point of the shape contour, and the x and y coordinates of point P_i are x_i and y_i, respectively. As the shape boundary is closed, we have $P_i = P_{i+N}$. For each point $P_i(x_i, y_i)$, we can find its two adjacent points $P_{i-l(k)}(x_{i-l(k)}, y_{i-l(k)})$ and $P_{i+l(k)}(x_{i+l(k)}, y_{i+l(k)})$, where $l(k)$ is the logarithmic distance between sample points of S, with $1 \le k \le T$, while l is an increasing function where $l(k) \le N/2$. T denotes the number of scales and N represents the number of sampling points. We set $T = \lfloor log_2(N/2) \rfloor$ and $h(k) = 2^{k-1}$ in the experiment. The above three consecutive points can form a triangle $\triangle P_{i-l(k)}P_iP_{i+l(k)}$. We can then compute the coordinates of the center point $g_{ik}(x_{g_{ik}}, y_{g_{ik}})$, which is given by

$$\begin{cases} x_{g_{ik}} = (x_{i-l(k)} + x_i + x_{i+l(k)})/3 \\ y_{g_{ik}} = (y_{i-l(k)} + y_i + y_{i+l(k)})/3 \end{cases} \tag{1}$$

For each sample point P_i of the boundary S, T triangles are obtained in the above way. Next, we can compute the distance between P_i and the center point $g_{ik}(i \in [1, N], k \in [1, T])$ of all triangles. Therefore, we can obtain the TDR of the sample point P_i, which is defined as follows:

$$TDR(P_i) = (|TCD(P_i, g_{i1})|, \ldots, |TCD(P_i, g_{iT})|), \tag{2}$$

where $|TCD(P_i, g_{ik})| = \sqrt{(x_i - x_{g_{ik}})^2 + (y_i - y_{g_{ik}})^2}$. The sign of $TCD(.)$ is determined by the following method: Consider a vector the starting point of which is $P_{i-l(k)}$ and the

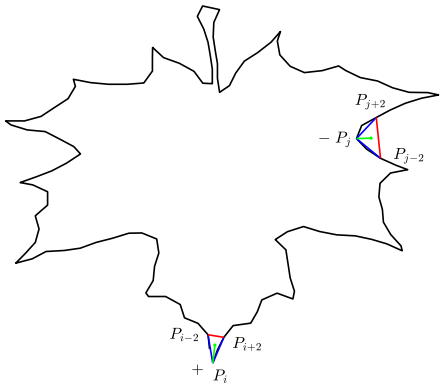

FIGURE 2. Two triangles of boundary points P_i and P_j, and their triangle center distances (shown in green color), where "+" and "−" are symbols of the triangle center distances.

ending point of which is $P_{i+l(k)}$. When the point P_i is located on the right-hand side of the vector, $TCD(.)$ has a positive value. When the point P_i is located on the left-hand side of the vector, $TCD(.)$ has a negative value. Fig. 2 graphically illustrates the definition of signed distances. The size of TCD value can reflect the bending degree and spatial information of the shape contour, while the sign of the TCD value can capture the concave and convex characteristics of the shape contour. By extracting the TDR of all the contour points, we acquire the TDR of the shape S, which is given by

$$\begin{aligned} TDR(S) &= (TDR(P_1), \ldots, TDR(P_N)) \\ &= \begin{pmatrix} |TCD(P_1, g_{11})| & \ldots & |TCD(P_N, g_{N1})| \\ \vdots & \ddots & \vdots \\ |TCD(P_1, g_{1T})| & \ldots & |TCD(P_N, g_{NT})| \end{pmatrix}. \end{aligned} \tag{3}$$

We can see that $MTCD(S)$ is a $T \times N$ matrix, where column i is the shape descriptor $TDR(P_i)$ of P_i of S. Next, we derive two matrices: a triangle center distance matrix $TCDM$ and a symbol matrix $SignM$, which are respectively given by

$$TCDM(S) = |TDR(S)|, \tag{4}$$

and

$$SignM(k, i) = \begin{cases} 1 & \text{if } TCD(P_i, g_{ik}) > 0 \\ 0 & \text{otherwise .} \end{cases} \tag{5}$$

According to the above definition, we observe that $TCDM$ represents the absolute value of the TDR, which is used to characterize the bending degree and spatial information of the shape contour, and $SignM$ is a binary matrix, which is used to characterize the concave and convex properties of the shape contour.

From the definition of the TDR, we can see that the TDR has an inherent translation invariance to the shape contour. Thus, according to Eqs. (4) and (5), $TCDM$ and $SignM$ have the same translation invariance to the shape contour. By analyzing the properties of the $TCDM$ and $SignM$ with regard to the scaling changes of a shape, we find that $SignM$ remains unchanged in shape scale, while $TCDM$ changes with scale. To make the $TCDM$ scale unchanged, we perform local normalization by dividing by the maximum value per row. In addition, when the shape contour is rotated, the starting point of the TDR descriptor changes, and a displacement h occurs, such as $TDR(P_i) = TDR(P_{i+h})$. Thus, we have $TCDM(P_i) = TCDM(P_{i+h})$ and $SignM(P_i) = SignM(P_{i+h})$. To keep the rotation of the shape contour constant, we use a Fourier transform for $TCDM$ and $SignM$ and only use the amplitude information of the Fourier coefficient to describe the shape. Meanwhile, the feature dimension of the proposed descriptor is further reduced after the Fourier transformation, thereby improving the efficiency and effectiveness of shape matching. To facilitate the interpretation of discrete Fourier transforms, we use c_k and d_k to represent the characteristics of each row of the $TCDM$ and $SignM$ descriptors, respectively. Thus, the amplitudes of their Fourier transform are computed by

$$\alpha_k(i) = |\frac{1}{N}\sum_{u=1}^{N} c_k(u) exp(\frac{-j2\pi(u-1)i}{N})|, \quad i = 1, 2, \ldots, N \tag{6}$$

and

$$\beta_k(i) = |\frac{1}{N}\sum_{u=1}^{N} d_k(u) exp(\frac{-j2\pi(u-1)i}{N})|, \quad i = 1, 2, \ldots, N, \tag{7}$$

where $j^2 = -1$. It can be easily proved that $\alpha_k(i)$ and $\beta_k(i)$ do not change with the rotation of a shape. Thus, we employ the amplitude information of the Fourier coefficients to characterize the shape of an object. To obtain a more compact and robust shape descriptor, the lowest M order coefficients are employed to describe the shape, where $M \ll N$. Therefore, we can obtain the final TDR for plant leaf recognition, which is defined as follows:

$$TDR = \{\alpha_k(v), \beta_k(v) | k = 1, 2, \ldots, T; v = 1, 2, \ldots M\}. \tag{8}$$

According to the definition of the final TDR, we observe that the size of the descriptor is becoming smaller. The dimension of the TDR changes from $T \times N$ to $T \times 2M$, where $M \ll N$. Hence, the TDR can further enhance the efficiency and reduce the storage space for the identification of plant leaves.

B. SHAPE DISSIMILARITY MEASURE

Given two shapes A and B, their TDRs are denoted $TDR^A = \{\alpha_k^A(v), \beta_k^A(v) | k = 1, \ldots, T; v = 1, \ldots M\}$ and $TDR^B = \{\alpha_k^B(v), \beta_k^B(v) | k = 1, \ldots, T; v = 1, \ldots M\}$, respectively. The dissimilarity measure between their TDRs is computed by

$$Dist(A, B) = \sum_{k=1}^{T}\sum_{v=1}^{M}(|\alpha_k^A(v) - \alpha_k^B(v)| + \lambda|\beta_k^A(v) - \beta_k^B(v)|), \tag{9}$$

where λ ($\lambda \in [0, 1]$) is a weight parameter. We apply the L_1 distance to calculate the degree of dissimilarity between the two shapes. A smaller distance value indicates that the two shapes are more similar. At the same time, we also note

that the TDR descriptors of the two shapes are compared at each scale. Therefore, the calculation of L_1 is quite simple and efficient for plant leaf recognition.

C. COMPUTATIONAL COMPLEXITY

In this subsection, we analyze the computational cost of our method for plant leaf recognition. The time consumption of the entire identification process is mainly composed of two parts. The first part is the time to calculate the TDR descriptor, and the second part is the time for shape matching. In the stage of calculating the TDR descriptor, the time cost of calculating the triangular center distance of all sampling points of the contour is $O(N)$ for each scale $k = 1, 2, \ldots, T$. Thus, the computational cost of extracting the TDR is $O(NT)$, where N and T represent the number of sampling points and scales, respectively. Subsequently, to maintain the rotational invariance of the shape boundary, we perform a discrete Fourier transform on each scale of the sign matrix $SignM$ and the triangle center distance matrix $TCDM$, respectively. The time cost of the Fourier transform of $SignM$ and $TCDM$ is $O(Nlog_2(N))$. As a result, the total time to calculate the TDR descriptor is $O(NT) + O(2TNlog_2N) = O(TNlog_2N) = O(\lfloor log_2(N/2) \rfloor Nlog_2N)$.

In the stage of shape matching, the time consumption is mainly employed to compute the distance between shapes. The time to calculate Eq. (9) is $O(2MT) = O(M\lfloor log_2(N/2) \rfloor)$, where $M \ll N$ represents the number of Fourier coefficients employed to describe the TDR. In the experiment, the values of T and M are small. Thus, the computation of L_1 distance is extremely fast. Our current implementation only needs about 0.4 ms to compute a matching on a 3.7 GHz computer. Thus, the proposed method is very efficient for the identification of plant leaves.

D. SUMMARY OF TDR PROPERTIES

The proposed TDR description method possesses the following characteristics which make it quite suitable for the tasks of real-time application and large-scale image retrieval. Below, we expound the merits of the TDR descriptor.

1) INVARIANCE TO SIMILARITY TRANSFORMATIONS

Similarity transformations include translation, rotation, and scaling, which is a basic requirement of the MPEG-7 standard [34]. Since we use relative distances, the proposed TDR has an inherent translational invariance. Subsequently, we employ a local normalization method to obtain the scaling invariance. Finally, we apply the discrete Fourier transforms for each scale of the TDR to maintain the rotational invariance of the shape.

2) COMPACTNESS OF THE TDR

Since the feature dimension at each scale $k = 1, \ldots, \lfloor log_2(N/2) \rfloor$ is $2M$, the total feature dimension of the proposed TDR descriptor is $2M\lfloor log_2(N/2) \rfloor$. We set $N = 512$ and $M = 6$ in the experiment, which makes the dimension of the TDR descriptor 96. Accordingly, only a small amount of memory is required to store the descriptor. We also compared the size of the TDR with the other three shape descriptors: IDSC [28], pattern counting [7], and MTCD [10]. The IDSC method is a very important and classic shape descriptor, and pattern counting is a recently proposed method for plant leaf recognition. The MTCD method is the most relevant to the proposed TDR method. Table 1 presents the comparison results of these shape descriptors. All experimental parameters are strictly in accordance with the requirements in their original paper. We can see from Table 1 that the proposed TDR descriptor has the lowest feature dimension and only requires less storage space. Therefore, the TDR is a quite compact shape description method.

TABLE 1. Feature dimensions of TDR, IDSC, pattern counting, and MTCD descriptors, where N and M denote the number of sampling points of the shape boundary and Fourier coefficients, respectively. N_d and N_θ are the number of inner-distance bins and inner-angle bins, respectively. In [7], the parameters of S and K represent the number of scales and number of dictionary atoms, respectively.

Methods	Feature dimension
IDSC [28]	$N \times N_d \times N_\theta = 128 \times 8 \times 12 = 12,288$
Pattern Counting [7]	$SK = 3 \times 512 = 1,536$
MTCD [10]	$M\lfloor (N-1)/2 \rfloor = 16 \times 63 = 1,008$
TDR	$2M\lfloor log_2(N/2) \rfloor = 2 \times 6 \times 8 = 96$

3) LOW COMPUTATIONAL COMPLEXITY

According to the analysis of the time complexity of our approach in Section III-C, the time complexity of our approach in the feature extraction phase is $O(\lfloor log_2(N/2) \rfloor Nlog_2N)$ and the time complexity of our approach in the shape matching phase is $O(M\lfloor log_2(N/2) \rfloor)$. For the well-known IDSC method, the time complexity of computing the IDSC is $O(N^3)$ and that of the shape matching phase is $O(KN^2)$, where K represents the number of possible starting points for shape alignment. For the MTCD method, the time complexity of the feature extraction and shape matching stages are $O(\lfloor (N-1)/2 \rfloor Nlog_2N)$ and $O(M\lfloor (N-1)/2 \rfloor)$, respectively. The proposed TDR method has lower computational complexity in the feature extraction and shape matching stages than the IDSC and MTCD methods. Therefore, the proposed TDR method is very suitable for the tasks of real-time application and large-scale image retrieval.

4) MULTISCALE REPRESENTATION STRUCTURE

The proposed TDR descriptor can well describe the local and global features of a leaf shape. We use the logarithmic distance as the scale to represent the shape. From the definition of the TDR, we find that larger scales can represent the global information of a shape, while smaller scales can represent the details of a shape. From the large scale to the small scale, a multi-scale TDR can be obtained. It can be seen from the experimental results that this multi-scale representation structure contains rich information on the shape of an object, thereby improving the ability to distinguish the shape.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETUP

We evaluated the performance of our approach on four popular leaf datasets: the Swedish [27], Smithsonian [28], Flavia [23], and ImageCLEF2012 [35] datasets. In addition, we also studied the potential of the proposed method for general shape recognition on the MPEG-7 shape dataset [36]. The same parameters ($N = 512, M = 6, T = \lfloor log_2(N/2) \rfloor = 8$, $\lambda = 0.3$) were set for the proposed method in all experiments.

We mainly used the evaluation standard defined by the dataset publishers when evaluating the performance of our method, and so it was also convenient to directly compare it with the previous methods. In addition, we also used the same performance evaluation benchmark, that is, the recognition rate versus top N images, to maintain the metric consistency across the test on all datasets, which can better reflect the performance of the algorithm.

It is worth pointing out that the main contribution of this article is to propose a new shape description method, so we used the existing technique for image pre-processing; that is, we first converted the color leaf into a grayscale leaf, and then used Otsu's method [37] to convert the grayscale leaf into a binary leaf. Third, we traced the leaf shape's boundary. Finally, we sampled the shape boundary into N points uniformly along the anti-clockwise direction and acquired a sequence of N points for subsequent feature extraction. Fig. 3 shows the pre-processing of a leaf image. In all experiments, we only used the outer contour of an object's shape to identify the plant leaves.

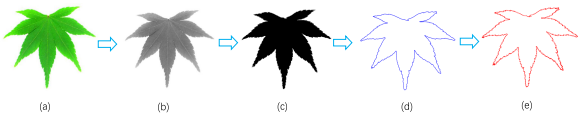

FIGURE 3. Preprocessing of leaf image: (a) original leaf, (b) grayscale leaf, (c) binary leaf, (d) leaf contour, and (e) uniformly sampled leaf contour with 512 points.

B. SWEDISH LEAF DATASET

The Swedish leaf dataset [27] is a famous dataset that is widely used to evaluate the classification performance of various shape descriptors. It contains 15 species of leaf images, each with 75 samples and a total of 1,125 leaf images. This dataset is very challenging for the identification of plant species because the shape of a leaf in the dataset has a small inter-class distance and a large intra-class difference. Fig. 4 shows several sample images of this dataset. We can see that the shapes of some species in the dataset are very similar, such as the first, third, and ninth species.

To facilitate a comparison with existing methods, we adopted the same performance evaluation criteria as in the literature [7], [10], [12], [13], [27], [28], [32]. For the 75 sample images in each category, we randomly selected 25 images as the training set, and the remaining 50 images served as the testing set. Then, we used one-nearest-neighbor (1-NN) to classify the testing set. The algorithm ran 10 times

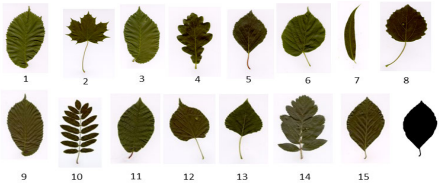

FIGURE 4. Samples for Swedish leaf dataset, one sample per category; Final leaf is an example of a mask image.

and took the average value as the final recognition rate of the algorithm. Table 2 lists the classification rates of our approach and the other methods on this dataset. The classification rates of other approaches are directly from the published results, except for the AlexNet [38] and VGG16 [39] methods. Since the AlexNet and VGG16 methods did not report the classification results on this dataset, we used the pre-trained AlexNet and VGG16 models and then fine-tuned them on the Swedish leaf dataset. Finally, we obtained the classification results of AlexNet and VGG16 methods on this dataset.

TABLE 2. Classification rates of different approaches on Swedish leaf dataset.

Method	Recognition rate
Soderkvist [27]	82.40%
SC+DP [28]	88.12%
Fourier Descriptor [28]	89.60%
SVDSR [12]	91.28%
IDSC+DP [28]	94.13%
SGLP [13]	95.24%
Shape-tree [32]	96.28%
Spatial PACT [40]	90.61%
MDM [4]	93.60%
Pattern Counting [7]	97.07%
MTCD [10]	96.31%
AlexNet+fine-tuning [38]	95.75%
VGG16+fine-tuning [39]	95.84%
Our Method	**97.27%**

We note that our approach achieves the best classification rate among all the competing approaches. Compared with the three complex dynamic programming matching methods of SC [28], IDSC [28], and shape-tree [32], the classification result of our approach is higher than 9.1%, 3.1%, and 1%, respectively. Compared with the two fast matching methods of MDM [4] and MTCD [10], the classification rate of our method is higher than 3.6% and nearly 1%, respectively. Compared with the recent generation model method SVDSP [12] and the manifold learning method SGLP [13], the classification accuracy of our method is higher than 6% and 2% of the two methods, respectively. Compared with the deep learning methods AlexNet [38] and VGG16 [39], the classification accuracy of our method is 1.5% and 1.4% higher than the two methods, respectively. The performance of our method is slightly superior to the pattern counting approach [7]. However, the pattern counting method must use a special training set to learn to generate codebooks or dictionaries, and then uses them to generate shape descriptors. If the training set does not exist, the shape descriptors cannot

be generated. The proposed TDR descriptor is generated independently and does not depend on any training samples.

To further evaluate the performance of our method, we also adopted the recognition rate versus top N leaves as the performance evaluation benchmark on the Swedish leaf dataset. In this measurement, each image of the dataset was used as a test sample, and then the distances between the test sample and all other images of the dataset were calculated. Finally, if there was a sample with the same class label as the test sample in the first N best matching (minimum shape difference), it was regarded as a successful recognition. The recognition rate of 1125 tests was counted as the recognition rate *Pre* corresponding to the parameter N. The recognition rate *Pre* versus the top N best matching curve was obtained when the N value was taken from 1 to 10. Obviously, the recognition rate *Pre* increases with N, and the corresponding value is the 1-NN recognition rate when $N = 1$. Fig. 5 shows the recognition results of our method and the other methods. As can be seen from the results, our method obtained the best recognition result in all the shape description methods compared. Compared with the deep learning methods AlexNet [38] and VGG16 [39], the recognition accuracy of our method was also comparable. This result further demonstrates the effectiveness of our approach in plant leaf recognition.

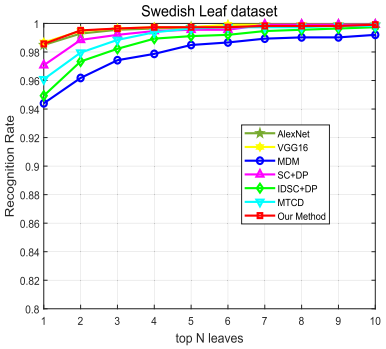

FIGURE 5. Recognition results of different methods on Swedish leaf dataset.

C. SMITHSONIAN LEAF DATASET

This dataset is a collection of isolated leaf images from the Smithsonian project [28]. It contains 343 leaf images from 93 plant species. The number of leaves from different plant species is different. Fig. 6 shows some sample images from this dataset. Leaf images in this dataset are susceptible to light changes and some leaf images are not very flat.

In the experiment, we used the same performance evaluation criteria used in the literature [10], [28]. We randomly selected 187 leaf images as the training samples, and the remaining 156 leaves were used for testing. The retrieval accuracy was assessed by a performance curve showing the recognition rate in the first N leaf images, where N ranges from 1 to 16. The recognition result versus the top N leaves curve is presented in Fig. 7.

The curves in Fig. 7 are the recognition rates, indicating how often the testing leaves can be correctly classified (or hit)

FIGURE 6. Several examples of Smithsonian leaf dataset with one example per species.

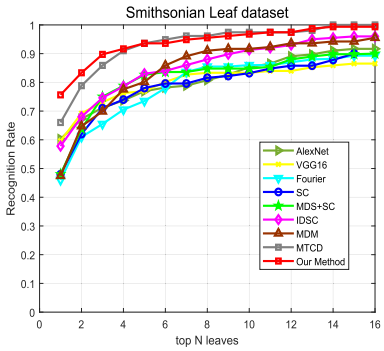

FIGURE 7. Recognition rates on Smithsonian leaf dataset.

by the top N candidates. It should be noted that this percentage increases monotonically with respect to N and decreases to the recognition rate of 1-NN when $N = 1$. Except for the results of the MDM method [4], AlexNet method [38], VGG16 method [39], and our method, the results of the other methods were derived from Refs. [10], [28]. Since the MDM method does not provide results on this dataset, we implemented the approach and reported the recognition result of the approach. At the same time, we also give the recognition results of the AlexNet and VGG16 methods on the Smithsonian dataset. We adopted the pre-trained VGG16 and AlexNet models based on the Caffe framework [41], and then used these two models to extract their FC7 features. Finally, we employed the L_2 distance to calculate the dissimilarity between FC7 features. We observe that our approach achieves the highest recognition accuracy among the nine approaches. It can be noted that the recognition rate of our method exceeds the IDSC [28] and MTCD [10] methods by 15% and 9%, respectively, when only one candidate result is returned. Compared with the two deep learning methods AlexNet [38] and VGG16 [39], the recognition accuracy of our method is higher than that of the two methods by approximately 15% and 16% when $N = 1$. This result indicates that our method has strong shape distinguishing ability.

D. FLAVIA LEAF DATASET

The Flavia leaf dataset is also widely employed to assess the performance of plant leaf recognition methods, which are primarily employed to evaluate the retrieval performance of algorithms. It has 32 plant species with a total of 1,907 leaf images. The number of leaf images for each species is

between 50 and 77. Fig. 8 presents some samples of this dataset.

FIGURE 8. Several example images of Flavia leaf dataset with one leaf per class.

Some methods are tested in [4], [10], [42] on the Flavia leaf dataset and their results are reported. To make a direct comparison with these approaches, we also used two performance evaluation criteria: the precision-recall (PR) curve and mean average precision (MAP). The precision P is the retrieved related image divided by the retrieved image, and the recall R is the retrieved related images divided by the total related images in the dataset. In the experiment, each image of the Flavia leaf dataset is sequentially served as the query, and the retrieval result is the average of all the queries.

The MAP value is computed by the following expression:

$$MAP = \frac{\sum_{q \in Q} AP(q)}{|Q|}. \tag{10}$$

Here, $|Q|$ denotes the number of queries. $AP(q)$ represents the average precision score of each query q, which is given by

$$AP(q) = \frac{\sum_{k=1}^{m} (P(k) \times f(k))}{n}, \tag{11}$$

where $P(k)$ denotes the precision at cutoff k in retrieved image sequences, and $f(k)$ takes 1 when the retrieved k-th image is related to the query q and 0 otherwise. m and n represent the number of retrieved images and retrieved images relating to the query q, respectively. The larger the MAP value, the better the performance. The MAP values of our method and other methods are shown in Table 3. Except for the AlexNet and VGG16 methods, the MAP scores of the other approaches are from the published results. For the AlexNet and VGG16 methods, we used the same process as the Smithsonian dataset in Section IV-C; that is, we used the pre-trained AlexNet and VGG16 models to extract their FC7 features, and then employed the L_2 distance to calculate the dissimilarity between FC7 features.

It can be seen from Table 3 that our approach achieved the highest MAP score among all nine approaches. Compared to other shape description methods, the MAP value of our method is higher than that by 3.3%-21.4%. Compared with the deep learning methods AlexNet [38] and VGG16 [39], the MAP score of our method is higher than the two methods by 1% and 3%, respectively. Although the shape of the dataset has very large inter-class similarities, our method still obtains a better recognition result.

TABLE 3. MAP values of different methods on Flavia leaf dataset.

Method	MAP
GEDT [42]	48.01%
AlexNet+FC7+L2 [38]	68.58%
VGG16+FC7+L2 [39]	66.42%
IDSC+DP [28]	59.32%
MDM [4]	47.91%
Riemannian metric [42]	57.21%
MTCD [10]	61.84%
SC+DP [29]	66.30%
Our Method	**69.61%**

Fig. 9 shows the PR curves of our approach and the other approaches on the Flavia leaf dataset. As can be seen, our approach achieves the best PR curve in all approaches compared. This result further demonstrates that our method is well suited for plant species identification.

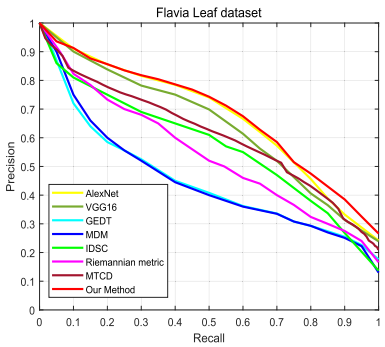

FIGURE 9. PR curves of proposed approach and other methods on Flavia leaf dataset.

To further evaluate the performance of our method, we also presented the recognition rate versus top N leaves for our method on the Flavia leaf dataset to maintain the metric consistency of all dataset tests. The test standard was the same as the Swedish leaf dataset in Section IV-B; that is, each image in the Flavia leaf dataset was used as a test sample, and then the distances between the test sample and all other images in the dataset were calculated. Finally, the top N best matching images (minimum shape difference) were returned. If there was an image with the same class label as the test sample in the returned image, it was regarded as a successful recognition. Fig. 10 shows the comparison results of our method and the other methods.

It can be seen from Fig. 10 that our method achieves the best recognition performance in all shape-based plant leaf recognition methods. This shows that our method has a strong ability to distinguish between shapes. At the same time, our method also obtains comparable results in comparison with the deep learning-based AlexNet and VGG16 methods. In addition, our method is very efficient in plant leaf classification and retrieval tasks. Therefore, our method is very effective for real-time application.

E. IMAGECLEF2012 LEAF DATASET

The fourth challenging dataset employed in our experiment was the famous ImageCLEF2012 plant leaf dataset [35].

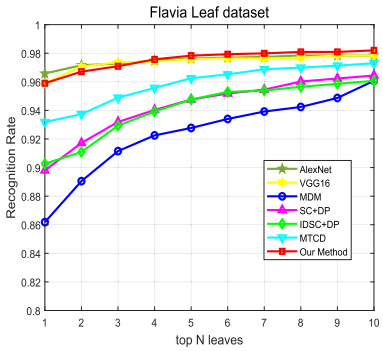

FIGURE 10. Comparison results of different methods on Flavia leaf dataset.

To the best of our knowledge, it is currently the largest dataset for the research of plant leaf identification. This dataset contains a total of 11,572 leaf images, which is divided into three categories of leaf images: ''scans'', ''pseudo-scan'', and ''photograph''. Leaves in the category ''scans'' contain clean backgrounds and minor shadowing, while those in category ''pseudo-scan'' are photographs with clean backgrounds, but have very heavy shadowing. Leaves in the category ''photograph'' contain unconstrained photos with complex natural backgrounds. We only employed the ''scans'' category images in the experiment, which accounts for 57% of the entire dataset. This is because we only focused on shape-based plant leaf recognition, and the shapes of the ''scans'' category images can be extracted accurately by using a pre-processing step. This subset consists of 6,630 leaves from 115 different species. The number of leaves per plant species is between 2 and 249. Fig. 11 shows some samples of the ''scanned'' class subset.

FIGURE 11. Some samples from the ''scans'' category subset of the ImageCLEF2012 leaf with one sample per species.

This dataset is very challenging because it contains a great number of different plant species, and the shape of different plant species has a large similarity, the shape of the same plant species varies greatly. This subset contains 4,870 leave images for training and 1,760 leaf images for testing. The performance evaluation standard employed was the same as for the Smithsonian leaf introduced in Section IV-C. Fig. 12 presents the recognition rates of our approach and other methods on this dataset.

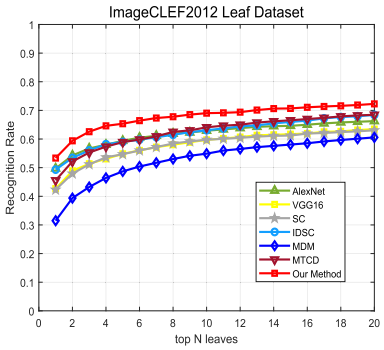

FIGURE 12. Recognition result on ''scans'' category of ImageCLEF2012 leaf dataset.

As can be seen from the results, our method achieves the best recognition performance among all the competing methods. The recognition rate of our method is much higher than that of the MDM [4] and SC [29] methods, and is higher than those two methods by 22% and 11%, respectively, when only one candidate result is returned. Compared to the IDSC [28] and MTCD [10] methods, the recognition accuracy of our method is higher than the two methods by 4% and 8%, respectively, when $N = 1$. Compared to the deep learning methods AlexNet [38] and VGG16 [39], the recognition accuracy of our method is higher than that of the two methods by 4% and 11% when $N = 1$. We can also see that the recognition performance of VGG16 on this dataset is almost identical to that of the SC method, and the result of AlexNet method on this dataset is slightly lower than the IDSC method. At the same time, we note that the recognition performance of the AlexNet method on this dataset is better than that of the VGG16 method. These results indicate that the proposed approach can distinguish different plant species well. In addition, the matching efficiency of our method is also very high. Therefore, our method is very suitable for real-time tasks and large-scale image retrieval tasks.

F. MPEG-7 SHAPE DATASET

To further test the potential of our approach for general shape recognition, we adopted the same experimental method as in [7], and used the MPEG-7 shape dataset [36] to evaluate the generality of the algorithm. This dataset has 1,400 images in 70 categories, with 20 images per category. Fig. 13 presents several samples from this dataset.

To facilitate comparison with other methods, we adopted the same performance evaluation criteria as in [7]. This performance metric is the same as that used for the Swedish leaf dataset in Section IV-B; that is, the recognition rate versus top N images evaluation standard. Fig. 14 shows the recognition rate curves of our method and the other six methods, that is, IDSC [28], MDM [4], pattern counting [7], MTCD [10], AlexNet [38], and VGG16 [39]. The result of the pattern counting method here is taken directly from the literature [7], and the result of the IDSC method is obtained by running the source code released by the author [28].

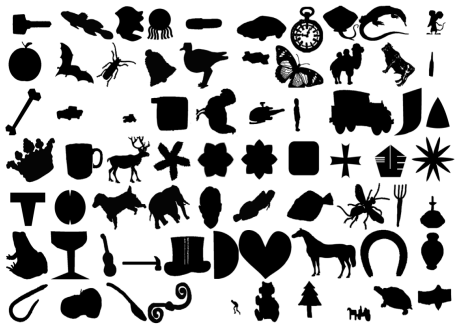

FIGURE 13. Several samples of MPEG-7 dataset, with one shape per class presented.

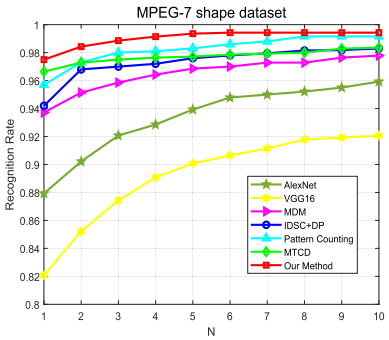

FIGURE 14. Recognition rates of different methods on MEPG-7 shape dataset.

For the MDM and MTCD methods, we directly implemented them and reported their experimental results on this dataset. For the AlexNet and VGG16 methods, we directly adopted the same processing as in Section IV-C; that is, we used the pre-trained AlexNet and VGG16 models to extract their FC7 features, and then employed the L_2 distance to calculate the dissimilarity between FC7 features. Finally, we obtained the recognition results of the AlexNet and VGG16 methods on the MPEG-7 shape dataset.

It can see from Fig. 14 that our method obtains the best recognition result in all methods. Compared with the classic IDSC [28] and MDM [4] methods, the recognition accuracy of our method is higher than that of the two methods by 3% and 4%, respectively, when only one candidate result is returned. Compared with the recent shape description methods MTCD [10] and pattern counting [7], the recognition accuracy of our method is higher than that of the two methods by approximately 1% and 2%, respectively, when $N = 1$. Compared to the deep learning methods AlexNet [38] and VGG16 [39], the recognition accuracy of our method is higher than that of the two methods by 9% and 15%, respectively, when $N = 1$. At the same time, it can be seen that the recognition accuracy of the AlexNet and VGG16 methods is lower than that of the shape-based description methods. The main reason is that AlexNet and VGG16 are pre-trained in natural images, while the images in the MEPG-7 dataset are binary images. This affects the AlexNet and VGG16 methods in extracting robust and effective image features. In addition, since the binary image on the MPEG-7 dataset is identified based on shape information, and AlexNet and VGG16 mainly extract high-level visual information of an object, their ability to describe the shape of an object is thus relatively weak, which further affects their recognition accuracy on the MPEG-7 shape dataset. The experimental results prove that our approach can be effectively applied to general shape recognition.

G. PARAMETER STUDY

We mainly analyzed the influence of selected parameters on plant leaf recognition performance. In our proposed method, there are three parameters: N, the number of sample points of the leaf boundary; λ, the weight parameter; and M, the number of Fourier coefficients. Here, N and M are mainly used for the shape description stage, and λ is used for the shape matching stage. We first analyzed the two parameters N and M of the shape description stage, and then the λ parameter. When analyzing the N and M parameters, λ remains unchanged; when analyzing the λ parameter, N and M remain unchanged. We used the Flavia leaf dataset for parameter analysis experiments. In the experiment, we set N to the power of two, so that the fast Fourier transformation could be used. The retrieval results were evaluated by MAP scores. Table 4 shows the MAP scores with $\lambda = 0.3$ at different values of N and M used in the shape feature extraction stage.

It can be seen from Table 4 that the best MAP score obtained by the proposed method on the Flavia leaf dataset is 70.58% when $N = 2048$ and $M = 6$. At the same time, it can also be seen that the MAP score increases with increasing N and decreases with increasing M. In addition, the increase in the MAP score becomes increasingly smaller and when the value of N is further increased, as can be seen from $N = 1024$ to $N = 2048$.

In general, a smaller M value can achieve better performance. This is because a smaller M value gives the main feature of a shape, while a larger M value provides more details of a shape, but also introduces too much noise. It can be seen from the experimental results that the MAP score will decrease when the M value is large. This is because the introduction of excessive noise affects the retrieval performance. Accordingly, $M = 6$ is the optimal value on the Flavia leaf dataset. Meanwhile, although more precise sampling can improve the accuracy of plant leaf recognition, the number of sampling points does not remarkably improve the recognition accuracy when N reaches a certain threshold. The satisfactory result of the Flavia leaf dataset is obtained when $N = 2048$. However, a larger value of N also increases the time consumption of plant leaf recognition and decreases the retrieval efficiency, which can be seen from the results in Fig. 15 in Section IV-H.

Therefore, the proposed TDR descriptor is a parameter-based method, which indicates that we can obtain the best recognition result by adjusting the parameters of different datasets. However, the parameters $N = 512$ and $M = 6$ can

TABLE 4. MAP scores of Flavia leaf dataset with $\lambda = 0.3$ at different values of N and M used in shape description stage.

	$N = 64$	$N = 128$	$N = 256$	$N = 512$	$N = 1024$	$N = 2048$
$M = 6$	0.6084	0.6513	0.6839	**0.6961**	**0.7042**	**0.7058**
$M = 8$	0.6018	0.6419	0.6733	0.6864	0.6963	0.7004
$M = 10$	0.5949	0.6317	0.6634	0.6769	0.6874	0.6926
$M = 12$	0.5884	0.6241	0.6551	0.6683	0.6795	0.6854
$M = 14$	0.5823	0.6166	0.6466	0.6603	0.6718	0.6787
$M = 16$	0.5777	0.6107	0.6396	0.6531	0.6649	0.6723
$M = 32$	0.5597	0.5899	0.6103	0.6196	0.6326	0.6423

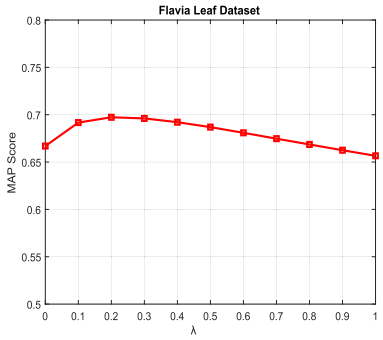

FIGURE 15. MAP scores of Flavia leaf dataset for 11 different values of $\lambda = 0$ weight when $N = 512$ and $M = 6$.

give better results for most applications, while maintaining the compactness and high efficiency of the descriptor. Thus, we employed the parameters $N = 512$ and $M = 6$ in all experiments.

We next analyzed the influence of λ parameter on the performance of plant leaf recognition. The λ parameter had 11 values: 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1. $\lambda = 0$ means that we only use the triangle center distance feature for the identification of plant leaves. Fig. 15 shows the MAP scores of the Flavia leaf dataset for 11 different λ weight values when $N = 512$ and $M = 6$. It can be seen that the MAP scores of the Flavia leaf dataset first increase and then decrease as λ increases, and the MAP score of the Flavia leaf dataset takes the optimal value when the $\lambda = 0.2$. At the same time, it is also noteworthy that better MAP scores can be obtained when λ is in the interval [0.2, 0.3]; that is to say, there is a reasonable range of λ values, within which the exact setting is not critical ($\lambda \in [0.2, 0.3]$). Thus, we set $\lambda = 0.3$ in all experiments.

In addition, it can be seen that we obtained a MAP score of approximately 5% over the MTCD method when $\lambda = 0$. This result shows that we can improve the accuracy of plant leaf recognition by using the logarithmic distance scale and L_1 distance. When $\lambda > 0$, the sign matrix is added to the shape feature, which indicates that the concave-convex property of the shape contour is taken into account. We can see that the MAP score is further improved, which is approximately 8% higher than that of MTCD method. This result shows that the concave-convex property of the shape contour can improve the ability to distinguish the shape, thereby improving the accuracy of plant leaf recognition.

Finally, to further explore the performance improvement of our method as being due to the three proposed improvements, we analyzed the performance of the MTCD method at different N and M values, so that we could further explore whether the performance is significantly improved when taking the same parameters as our method. Table 5 gives the MAP scores of the Flavia leaf dataset for the MTCD method at different N and M values. As can be observed, the MAP score of the MTCD method first increased and then decreased as N increased for all M values, and the MAP score of the MTCD method first increased and then decreased as M increased when N value is than or equal to 128. For $N = 64$, the MAP score of the MTCD method first increased, then decreased, and next increased, but the overall performance remained relatively stable. The best MAP score of the MTCD method on the Flavia leaf dataset is 62.74% when $N = 512$ and $M = 12$. The best performance achieved by the MTCD method is also approximately 8% lower than that of our proposed method. This experimental result further shows the ability to improve the performance of plant leaf recognition via the proposed three improvements. At the same time, it can be seen that the efficiency of our method is also significantly higher than that of the MTCD method from the subsequent efficiency experiment analysis. In [10], the parameter N is set to 128. This is because the efficiency of recognition will decrease when N is too large. In addition, the exact value of M is not critical (e.g., $M \in [8, 32]$) for the performance of plant leaf recognition after setting the value of N.

H. EFFICIENCY COMPARISON

To prove the efficiency and effectiveness of the proposed approach, we compared the time consumption of the proposed approach with other important methods. Table 6 shows the time consumption of various methods on the Flavia leaf dataset. The time consumption of all methods is obtained under the same conditions, which represents the average time cost in retrieving each query in the Flavia leaf dataset. In our experiment, all the algorithms were implemented by MatLab except for the DP step of the IDSC method, which is implemented in the C language. The experimental platform employed a PC with a Core(TM) i7-8700K 3.7-GHz CPU and MatLab 2017a. We chose the IDSC method [28] for comparison because it is the most classic and popular method for shape description. Meanwhile, the source code of this method is provided. Thus, we could directly compare it with our method under the same conditions. For the fast shape matching method, we chose the MDM [4] and MTCD [10] methods. These two methods are recently proposed and achieve better results for plant leaf recognition. We implemented both methods and compared them with our method under the same conditions.

TABLE 5. MAP scores of Flavia leaf dataset for the MTCD method at different *N* and *M* values.

	$N = 64$	$N = 128$	$N = 256$	$N = 512$	$N = 1024$	$N = 2048$
$M = 6$	0.5942	0.6096	0.6182	0.6188	0.6185	0.6183
$M = 8$	0.6008	0.6154	0.6237	0.6241	0.6239	0.6236
$M = 10$	0.6038	0.6180	0.6262	0.6266	0.6263	0.6261
$M = 12$	0.6047	0.6188	0.6271	**0.6274**	0.6272	0.6270
$N = 14$	0.6040	0.6185	0.6263	0.6267	0.6265	0.6262
$M = 16$	0.6038	0.6184	0.6263	0.6267	0.6265	0.6262
$N = 32$	0.6042	0.6182	0.6262	0.6264	0.6262	0.6261

TABLE 6. Time consumption of various approaches on Flavia leaf dataset.

Methods	Average retrieval time (ms)
IDSC+DP [28]	4.92×10^3
MDM [4]	1.78×10^2
MTCD (N=128, M=16) [10]	1.67×10^2
Our Method (N=512, M=8)	1.62×10^2
Our Method (N=128, M=16)	1.55×10^2

It can be seen from Table 6 that the retrieval time of our method is much lower than that of the IDSC method. The time consumption of our method is only approximately 3% of the IDSC method. At the same time, it can be noted that the IDSC method employs DP in the C language to calculate the dissimilarity between shapes. The C language is faster than the MatLab language for loop operations. Thus, the efficiency of the IDSC method will be further reduced when the dissimilarity is computed by MatLab. Even with the two fast matching methods of MDM and MTCD, the retrieval time of our method is lower than theirs, and the recognition accuracy of our method is higher than those of the two methods. In the last row of Table 6, the parameter settings of our method are the same as those of MTCD method. It can be seen that the retrieval time of our approach is less than that of the MTCD approach. This result shows that using logarithmic distance as a scale can improve the efficiency of retrieval. Thus, our method is very suitable for the tasks of real-time application and large-scale image retrieval.

To further explore the effect of the number of sampling points N on retrieval efficiency, we compared the time consumption of our approach and that of the MTCD method at different N values. Fig. 16 shows the time consumption of the two methods when taking different values of N, which represents the total time consumption in retrieving all images in the Flavia leaf dataset.

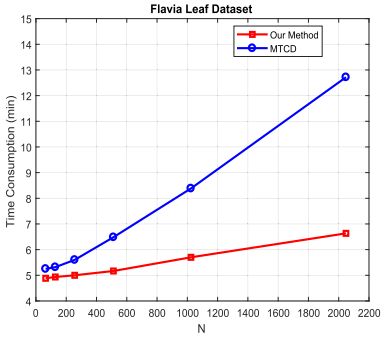

FIGURE 16. Time consumption of our method and the MTCD method for different *N* values.

It can be seen that the total retrieval time of our method is lower than that of the MTCD method at all N values. This result further demonstrates that the use of logarithmic distance as a scale can improve the retrieval efficiency. At the same time, it can be noted that the retrieval time will increase significantly when the value of N is greater than or equal to 1024. This is mainly due to the increased time of shape feature extraction. However, the time increase of our method is far less than that of the MTCD method. In addition, the total retrieval time of our method is relatively stable when N is less than or equal to 512. Therefore, this is also why we set the N to 512 in our experiment.

V. CONCLUSION

We have proposed a novel shape description approach named TDR for plant species recognition. The proposed TDR is described by two matrices: a sign matrix and a triangle center distance matrix. The sign matrix is used to characterize the concave and convex properties of a shape, while the triangle center distance matrix is used to reflect the bending degree and spatial distribution information of a shape. The TDR descriptor is a compact multi-scale representation method that can well characterize the local and global features of the shape while maintaining the invariance to similarity transforms. We conducted extensive experiments on four standard plant leaf datasets, and the results demonstrate that our approach exceeds the existing state-of-the-art shape-based plant species recognition approaches on the aspects of retrieval accuracy, efficiency, and storage space. In addition, our experimental results on the MPEG-7 shape dataset further confirm that our method has great potential in general shape recognition. However, we only used the shape information of plant leaves, and the texture and venation structure of plants are also very important features for identifying plant species. In future work, we plan to combine our method with textures or venation structures for plant leaf recognition. We believe that the performance of our method will be further enhanced when texture or venation information of plants is incorporated.

REFERENCES

[1] X. F. Wang, D. S. Huang, J. X. Du, H. Xu, and L. Heutte, "Classification of plant leaf images with complicated background," *Appl. Math. Comput.*, vol. 205, no. 2, pp. 916–926, 2008.

[2] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin, "Plant species identification using digital morphometrics: A review," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7562–7573, 2012.

[3] J. Wäldchen and P. Mäder, "Plant species identification using computer vision techniques: A systematic literature review," *Arch. Comput. Methods Eng.*, vol. 25, no. 2, pp. 507–543, 2018.
[4] R. Hu, W. Jia, H. Ling, and D. Huang, "Multiscale distance matrix for fast plant leaf recognition," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4667–4672, Nov. 2012.
[5] Y. K. Lei, J. W. Zou, T. Dong, Z. H. You, Y. Yuan, and Y. Hu, "Orthogonal locally discriminant spline embedding for plant leaf recognition," *Comput. Vis. Image Understand.*, vol. 119, no. 2, pp. 116–126, 2014.
[6] J. Fan, N. Zhou, J. Peng, and L. Gao, "Hierarchical learning of tree classifiers for large-scale plant species identification," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4172–4184, Nov. 2015.
[7] C. Zhao, S. S. F. Chan, W. K. Cham, and L. M. Chu, "Plant identification using leaf shapes—A pattern counting approach," *Pattern Recognit.*, vol. 48, no. 10, pp. 3203–3215, 2015.
[8] C. Kalyoncu and Ö. Toygar, "Geometric leaf classification," *Comput. Vis. Image Understand.*, vol. 133, pp. 102–109, Apr. 2015.
[9] Y. G. Naresh and H. S. Nagendraswamy, "Classification of medicinal plants: An approach using modified LBP with symbolic representation," *Neurocomputing*, vol. 173, pp. 1789–1797, Jan. 2016.
[10] C. Yang, H. Wei, and Q. Yu, "Multiscale triangular centroid distance for shape-based plant leaf recognition," in *Proc. Eur. Conf. Artif. Intell.*, 2016, pp. 269–276.
[11] K. Horaisová and J. Kukal, "Leaf classification from binary image via artificial intelligence," *Biosyst. Eng.*, vol. 142, pp. 83–100, Feb. 2016.
[12] S. Zhang, C. Zhang, Z. Wang, and W. Kong, "Combining sparse representation and singular value decomposition for plant recognition," *Appl. Soft Comput.*, vol. 67, pp. 164–171, Jun. 2018.
[13] Y. Shao, "Supervised global-locality preserving projection for plant leaf recognition," *Comput. Electron. Agricult.*, vol. 158, pp. 102–108, Mar. 2019.
[14] U. P. Singh, S. S. Chouhan, S. Jain, and S. Jain, "Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease," *IEEE Access*, vol. 7, pp. 43721–43729, 2019.
[15] P. N. Belhumeur, D. Chen, S. Feiner, D. W. Jacobs, W. J. Kress, H. Ling, I. Lopez, R. Ramamoorthi, S. Sheorey, and S. White, "Searching the world's herbaria: A system for visual identification of plant species," in *Proc. Eur. Conf. Comput. Vis. Comput. Vis. (ECCV)*, Marseille, France, Oct. 2008, pp. 116–129.
[16] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. B. Soares, "Leafsnap: A computer vision system for automatic plant species identification," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 502–516.
[17] P. Novotný and T. Suk, "Leaf recognition of woody species in central Europe," *Biosyst. Eng.*, vol. 115, no. 4, pp. 444–452, 2013.
[18] S. Zhang, Y. Lei, T. Dong, and X. P. Zhang, "Label propagation based supervised locality projection analysis for plant leaf classification," *Pattern Recognit.*, vol. 46, no. 7, pp. 1891–1897, 2013.
[19] J. X. Du, C. M. Zhai, and Q. P. Wang, "Recognition of plant leaf image based on fractal dimension features," *Neurocomputing*, vol. 116, no. 10, pp. 150–156, Sep. 2013.
[20] M. G. Larese, R. Namías, R. M. Craviotto, M. R. Arango, C. Gallo, and P. M. Granitto, "Automatic classification of legumes using leaf vein image features," *Pattern Recognit.*, vol. 47, no. 1, pp. 158–168, Jan. 2014.
[21] J. Hu, Z. Chen, M. Yang, R. Zhang, and Y. Cui, "A multiscale fusion convolutional neural network for plant leaf recognition," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 853–857, Jun. 2018.
[22] M. Turkoglu and D. Hanbay, "Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts," *Appl. Math. Comput.*, vol. 352, pp. 1–14, Jul. 2019.
[23] S. G. Wu, F. S. Bao, E. Y. Xu, Y. X. Wang, Y. F. Chang, and Q. L. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Dec. 2007, pp. 11–16.
[24] C. Caballero and M. C. Aranda, "Plant species identification using leaf image retrieval," in *Proc. ACM Int. Conf. Image Video Retr.*, 2010, pp. 327–334.
[25] G. Cerutti, L. Tougne, J. Mille, A. Vacavant, and D. Coquin, "Understanding leaves in natural images: A model-based approach for tree species identification," *Comput. Vis. Image Understand.*, vol. 117, no. 10, pp. 1482–1501, 2013.
[26] E. Aptoula and B. Yanikoglu, "Morphological features for leaf based plant recognition," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 1496–1499.
[27] O. J. O. Söderkvist, "Computer vision classification of leaves from swedish trees," M.S. thesis, Linköping Univ., Linköping, Sweden, Sep. 2001.
[28] H. Ling, "Shape classification using the inner-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.
[29] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.
[30] S. Mouine, I. Yahiaoui, and A. Verroust-Blondet, "Advanced shape context for plant species identification using leaf image retrieval," in *Proc. 2nd ACM Int. Conf. Multimedia Retr.*, 2012, pp. 1–8.
[31] S. Mouine, I. Yahiaoui, and A. Verroust-Blondet, "Plant species recognition using spatial correlation between the leaf margin and the leaf salient points," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 1466–1470.
[32] P. F. Felzenszwalb and J. D. Schwartz, "Hierarchical matching of deformable shapes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[33] D. J. Hearn, "Shape analysis for the automated identification of plants from images of leaves," *Taxon*, vol. 58, no. 3, pp. 934–954, 2009.
[34] J. M. Martínez, "MPEG-7 Overview (version 9)," Int. Org. Standardization-Org. Int. Normalization, Geneva, Switzerland, Tech. Rep. N5525, 2003.
[35] H. Goëau, P. Bonnet, A. Joly, I. Yahiaoui, D. Barthélémy, B. Nozha, and J.-F. Molino, "The ImageCLEF 2012 Plant identification Task," in *Proc. CLEF*, Rome, Italie, Sep. 2012, pp. 1–25.
[36] L. J. Latecki, R. Lakamper, and T. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2000, pp. 424–429.
[37] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[39] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1097–1105.
[40] J. Wu and J. M. Rehg, "Centrist: A visual descriptor for scene categorization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1489–1501, Aug. 2011.
[41] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 675–678.
[42] H. Laga, S. Kurtek, A. Srivastava, and M. Golzarian, "A Riemannian elastic metric for shape-based plant leaf classification," in *Proc. Int. Conf. Digit. Image Comput. Techn. Appl.*, 2012, pp. 1–7.

CHENGZHUAN YANG received the Ph.D. degree from the Department of Computer Science, Fudan University, China, in 2017. He is currently an Assistant Professor with the School of Information, Zhejiang University of Finance and Economics. His current research interests include computer vision, pattern recognition, and artificial intelligence.

HUI WEI received the Ph.D. degree from the Department of Computer Science, Beijing University of Aeronautics and Astronautics, in 1998. From 1998 to 2000, he was a Postdoctoral Fellow with the Department of Computer Science and the Institute of Artificial Intelligence, Zhejiang University. Since November 2000, he has been with the Department of Computer Science and Engineering, Fudan University. His research interests include artificial intelligence and cognitive science.