

Shape Recognition by Combining Contour and Skeleton into a Mid-Level Representation

Wei Shen¹, Xinggang Wang², Cong Yao², and Xiang Bai²

¹ School of Communication and Information Engineering, Shanghai University,
149 Yanchang Road, Shanghai 200072, P.R. China

² Dept. of Electronics and Information Engineering, Huazhong University of Science
and Technology, 1037 Luoyu Road, Wuhan, Hubei Province 430074, P.R. China

Abstract. Contour and skeleton are two main stream representations for shape recognition in the literature. It has been shown that such two representations convey complementary information, however combining them in a nature way is nontrivial, as they are generally abstracted by different structures (closed string *vs* graph), respectively. This paper aims at addressing the shape recognition problem by combining contour and skeleton into a mid-level of shape representation. To form a mid-level representation for shape contours, a recent work named **Bag of Contour Fragments** (BCF) is adopted; While for skeleton, a new mid-level representation named **Bag of Skeleton Paths** (BSP) is proposed, which is formed by pooling the skeleton codes by encoding the skeleton paths connecting pairs of end points in the skeleton. Finally, a compact shape feature vector is formed by concatenating BCF with BSP and fed into a linear SVM classifier to recognize the shape. Although such a concatenation is simple, the SVM classifier can automatically learn the weights of contour and skeleton features to offer discriminative power. The encouraging experimental results demonstrate that the proposed new shape representation is effective for shape classification and achieves the state-of-the-art performances on several standard shape benchmarks.

Keywords: Shape Recognition, Mid-level Shape Representation, Bag of Contour Fragments, Bag of Skeleton Paths.

1 Introduction

Shape plays an important role in human perception for object recognition. The objects shown in Fig. 1 have lost their brightness, color and texture information and are only represented by their silhouettes, however it's not intractable for human to recognize their categories. This simple demonstration indicates that shape is stable to the variations in object color and texture and light conditions. Due to such advantages, recognizing objects by their shapes has been a long standing problem in the literature. Shape recognition is usually considered as a classification problem that is given a testing shape, to determine its category label based on a set of training shapes as well as their category label. The main challenges in shape recognition are the large intra-class variations induced

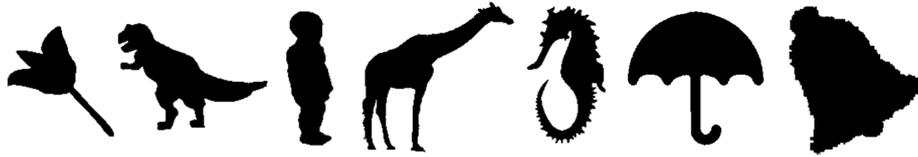


Fig. 1. Human biological vision system is able to recognize these object without any appearance information (brightness, color and texture)

by deformation, articulation and occlusion. Therefore, the main focus of the research efforts have been made in the last decade [5,12,17,18,3,4] is how to form a informative and discriminative shape representation.

Generally, the existing main stream shape representations can be classified into two classes: contour based [5,12,9] and skeleton based [1,2,16,14]. The former one delivers the information that how the spatial distribution of the boundary points varies along the object contour. Therefore, it captures more informative shape information and is stable to affine transformation. However, it is sensitive to non-ridge deformation and articulation; On the contrary, the latter one provides the information that how thickness of the object changes along the skeleton. Therefore, it is invariant to non-ridge deformation and articulation, although it only carries more rough geometric features of the object. Consequently, such two representations are complementary. Nevertheless, very few works have tried combining these two representations for shape recognition. The reason might be that combining the data of different structures is not trivial, as the contour is always abstracted by a closed string while the skeleton is abstracted either by a graph or a tree. So far as we know, ICS [3] is the only work to explicitly discuss how to combine contour and skeleton to improve the performance of shape recognition. However, the combination proposed in this work is just a weighted sum of the outputs of two generative models trained individually on contour features and skeleton features respectively. Therefore, how to combine contour and skeleton into a shape representation in a principled way is still an open problem.

In this paper, our goal is to address the above combination issue to explore the complementary between contour and skeleton to improve the performance of shape recognition. The main obstacle of the combination is how to transform the data of different structures into a common form. Recently, a contour based shape representation named **B**ag of **C**ontour **F**ragments (BCF) [20] was proposed, which is inspired by the well-known Bag of Features framework. In BCF, a contour is decomposed into a set of contour fragments, which will be then encoded and pooled to form a feature vector to represent the contour. BCF sheds lights onto the issue of the combination of contour and skeleton: Since a contour can be represented by a feature vector in BCF, a straightforward way to combine the contour and its skeleton is converting its skeleton to a feature vector as well followed by concatenating the two feature vectors. Toward this end, we propose a skeleton based representation named **B**ag of **S**keleton **P**aths (BSP) inspired by the framework of BCF.

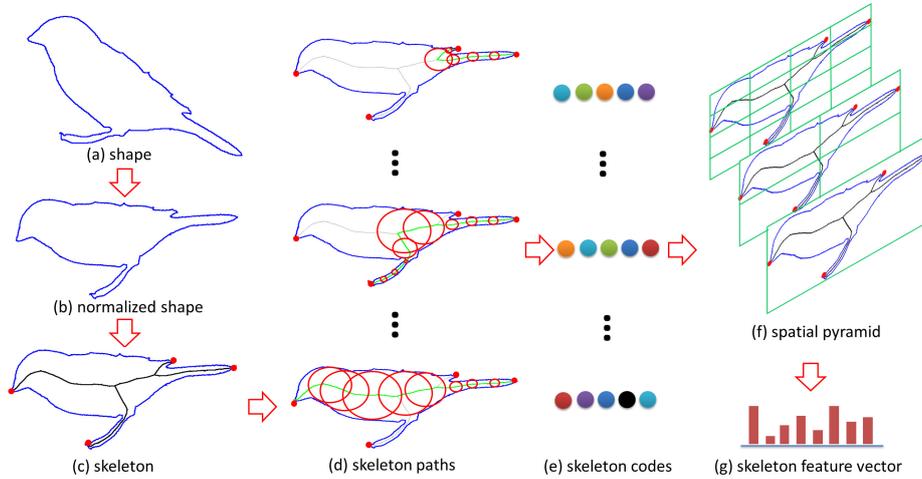


Fig. 2. The pipeline of building skeleton based shape representation by Bag of Skeleton Paths. (a) A shape. (b) The normalized shape obtained by aligned its major axis with the horizontal line. (c) Some examples of skeleton paths in green color. (d) The skeleton codes encoded from the skeleton paths. (e) $2^l \times 2^l$ ($l = 0, 1, 2$) subregions are used in SPM for max pooling. (f) The formed skeleton based shape representation.

Fig. 2 shows the pipeline of building skeleton based shape representation by Bag of Skeleton Paths. Given a shape, firstly a normalization step is performed to align the shape according to its major axis, as the spatial pyramid matching (SPM) [11] step shown in Fig. 2(f) is not rotation invariant. Then, the skeleton of the shape is extracted and decomposed into a set of skeleton paths. The skeleton paths, shown by the green curves in Fig. 2(d), are the shortest paths between pairs of end points of the skeleton. According to [2], a skeleton path is represented by a sequence of the radii of the maximal discs centered at skeleton points, as shown by the red circles in Fig. 2(d). Next, each skeleton path is encoded into a skeleton code. Finally, the skeleton codes are pooled into a compact skeleton feature vector by SPM. To encode skeleton paths, we adopt local-constrained linear coding (LLC) [19] scheme, as it has been proved to be efficient and effective for image classification. SPM provides additional spatial layout information and has been widely used in feature learning framework to boost the performance of image classification. It partitions the image into increasingly finer spatial subregions and computes histograms of local features from each sub-region. However, skeleton paths are different from the popular image features computed on rectangular image patches, such as SIFT [13] and HOG [7]. To perform SPM, we determine which sub-region a skeleton paths falls in by the location of the end point from which it emanate. The proposed BSP provides a effective way to convert a skeleton graph to a vector, a conventional form that can be dealt with by general classification models.

By concatenating the contour feature vector obtained by BCF with the skeleton feature vector obtained by BSP, a final shape feature vector is formed.

Any discriminative models, such as SVM and Random Forest, can be directly applied to the shape feature vector for shape classification. Using such discriminative models for shape recognition is more efficient than traditional shape classification methods, as the latter require time consuming matching and ranking steps. In addition, the weights of the contour features and skeleton features can be automatically learned by the discriminative models, such as linear SVM. This is an obvious advantage of the proposed combination method compared to ICS [3], which has to fine tune the weights between contour and skeleton models. Consequently, it's a natural way to combine the contour and the skeleton into a mid-level representation.

Our contributions can be summarized in two aspects. First, we propose a novel shape representation named Bag of Skeleton Paths, which can convert a skeleton graph to a compact single feature vector. Second, we provide a natural way to combine skeleton and contour for shape recognition, which achieves the state-of-the-arts on several shape benchmarks.

2 Related Work

There have been a rich body of works concerning shape recognition in recent years [5,12,17,18]. In the early age, the exemplar-based strategy has been widely used, such as [5,12]. Generally, there are two key steps in this strategy. The first one is extracting informative and robust shape descriptors. For example, Belongie *et al.* [5] introduce a shape descriptor named shape context (SC) which describes the relative spatial distribution (distance and orientation) of landmark points sampled on the object contour around feature points. Lin and Jacobs [12] use inner distance to extend shape context to capture articulation. As for skeleton based shape descriptors, the shock graph and its variants [16,14] are most popular, which are abstracted from skeletons by designed shape grammar. The second one is finding the correspondences between two sets of the shape descriptors by matching algorithms such as Hungarian, thin plate spline (TPS) and dynamic programming (DP). A testing shape is classified into the class of its nearest neighbor ranked by the matching costs. The exemplar-based strategy requires a large number of training data to capture the large intra-class variances of shapes. However, when the size of training set become quite large, it's intractable to search the nearest neighbor due to the high time cost caused by pairwise matching.

Generative models are also used for shape recognition. Sun and Super [17] propose a Bayesian model, which use the normalized contour fragments as the input features for shape classification. Wang *et al.* [18] model shapes of one class by a skeletal prototype tree learned by skeleton graph matching. Then a Bayesian inference is used to compute the similarity between a testing skeleton and each skeletal prototype tree. Bai *et al.* [3] propose to integrate contour and skeleton by a Gaussian mixture model, in which contour fragments and skeleton paths are used as the input features. Unlike their method, ours combine contour and skeleton into the mid-level of shape representation, and learn the weights between contour features and skeleton features automatically.

Recently, researchers begin to apply the powerful discriminative models to shape classification. Daliri and Torre [8] transform the contour into a string based representation according to a certain order of the corresponding contour points found during contour matching. Then they apply SVM to the kernel space built from the pairwise distances between strings to obtain classification results. Wang *et al.* [20] utilize LLC strategy to extract the mid-level representation BCF from contour fragments and they use linear SVM for classification. The proposed BSP is a extension of BCF for skeleton based mid-level representation.

3 Methodology

In this section, we will introduce our method for shape recognition, including the steps of shape normalization, BSP representation and shape classification by the combination of contour and skeleton.

3.1 Shape Normalization

As the SPM strategy assumes that the parts of shapes falls in the same sub-region are similar, it is not rotation invariant. To apply SPM to shape classification, a normalization step is required to align shapes roughly. One straight forward solution is to align each shape with its major axis. Here, we use principal component analysis (PCA) to compute the orientation of the major axis of each shape. Formally, given a shape $F \subset \mathbb{R}^2$, we apply PCA to the point set $\{p_i = (x_i, y_i) | p_i \in F\}_{i=1}^N$. First, the $N \times N$ covariance matrix Σ is computed by $\Sigma = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_i)(y_i - \bar{y}_i)$, where $\bar{x}_i = \sum_{i=1}^N x_i / N$ and $\bar{y}_i = \sum_{i=1}^N y_i / N$. Then, the two eigenvectors \mathbf{v}_1 and \mathbf{v}_2 of Σ form the columns of the $N \times N$ matrix V , and the two eigenvalues of Σ is $(\lambda_1, \lambda_2)^T = \mathbf{diag}(V^T \Sigma V)$. The orientation of the major axis of the shape F is the orientation of the eigenvector whose corresponding eigenvalue is bigger. All shapes are rotated to ensure their estimated major axes are aligned with the horizontal line, such as the example given in Fig. 2(b).

3.2 Bag of Skeleton Paths

In this section, we show how to build a BSP shape representation for a give shape F step by step.

Skeleton Paths Given a shape F , we obtain its skeleton $S(F)$ by the method introduced in [15], which does not require parameter tuning for skeleton computation. An end point in a skeleton is a skeleton point only have one adjacent skeleton point, such as the red points in Fig. 2(d). The shortest path between two end points, such as the green curves in Fig. 2(d), is called skeleton path, which is a informative skeleton descriptor and has been successfully used for skeleton matching [2]. Suppose there are m end points $\{e_i\}_{i=1}^m$ in the skeleton

$S(F)$. Let $h_{i,j}$ denote the skeleton path between e_i and e_j , then the set of the skeleton paths of $S(F)$ are

$$\mathcal{S}(S(F)) = \{h_{i,j} | i \neq j, i, j = 1 \dots m\}. \quad (1)$$

Note that, the skeleton paths $h_{i,j}$ and $h_{j,i}$ are two different skeleton paths. To represent a skeleton path $h_{i,j}$, a sequence of T skeleton points on it are sampled equally. Thus, the skeleton path $h_{i,j}$ is represented by $\mathbf{r}_{ij} = (\frac{R_t}{\bar{R}}; t = 1 \dots T)^T$, where R_t is the radius of the maximal disc centered at the t -th sampled skeleton points and \bar{R} is the mean value of the radii of all the skeleton points of $S(F)$. \bar{R} is a normalization factor to ensure \mathbf{r}_{ij} is scale invariant. The radius of the maximal disc centered at a skeleton point p is computed by the value of the distance transform of p w.r.t the object contour. In the following steps, we describe a skeleton path $h_{i,j}$ by its radii sequence descriptor $\mathbf{r}_{ij} \in \mathbb{R}^T$ for notation simplification.

Skeleton Paths Encoding. Encoding a skeleton path $\mathbf{r} \in \mathbb{R}^T$ is transforming it into a new space \mathcal{B} by a given codebook with K entries, $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_K) \in \mathbb{R}^{T \times K}$. In the new space, the skeleton path \mathbf{r} is represented by a skeleton code $\mathbf{c} \in \mathbb{R}^K$.

Codebook construction is usually achieved by unsupervised learning, such as k-means. Given a set of skeleton paths randomly sampled from all the skeletons in a dataset, we apply k-means algorithm to cluster them into K clusters and construct a codebook $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_K)$. Each cluster center forms a entry of the codebook \mathbf{b}_i .

To encode a skeleton path \mathbf{r} , we adopt LLC scheme [19], as it has been proved to be effective for image classification. Encoding is usually achieved by minimizing the reconstruction error. LLC additionally incorporates locality constraint, which solves the following constrained least square fitting problem:

$$\min_{\mathbf{c}_{\pi_k}} \|\mathbf{r} - \mathbf{B}_{\pi_k} \mathbf{c}_{\pi_k}\|, \quad \text{s.t.} \quad \mathbf{1}^T \mathbf{c}_{\pi_k} = 1, \quad (2)$$

where \mathbf{B}_{π_k} is the local bases formed by the k nearest neighbors of \mathbf{r} and $\mathbf{c}_{\pi_k} \in \mathbb{R}^k$ is the reconstruction coefficients. Such a locality constrain leads to several favorable properties such as local smooth sparsity and better reconstruction. The code of \mathbf{r} encoded by the codebook \mathbf{B} , i.e. $\mathbf{c} \in \mathbb{R}^K$, can be easily converted from \mathbf{c}_{π_k} by setting the corresponding entries of \mathbf{c} are equal to \mathbf{c}_{π_k} 's and others are zero.

Skeleton Code Pooling. Given a skeleton S , its skeleton paths are encoded into skeleton codes $\{c_i\}_{i=1}^n$, where n is the number of the skeleton paths in S . Now we describe how to obtain a compact skeleton based shape representation by pooling the skeleton codes. SPM is usually used to incorporate spatial layout information when pooling the image codes. It usually divide a image into $2^l \times 2^l$ ($l = 0, 1, 2$) subregions and then the features in each subregion are pooled respectively. While skeleton paths are quite different from the image features

computed on rectangular image patches. We find that skeleton paths emanate from one end point to others describe how the thickness of the object varies from the near to the distant (seeing the three skeleton paths emanate from one end point shown in Fig. 2(d)). For the aligned shapes belong to one category, the skeleton paths emanate from the end points falls in the same subregions should be similar. Therefore, we determine which subregion a skeleton path belong to by the location of the end point from which it emanates. More specifically, we divide a shape F into $2^l \times 2^l$ ($l = 0, 1, 2$) subregions, i.e. 21 subregions totally. Let $\mathbf{c}^e \in \mathbb{R}^K$ denote the skeleton code of a skeleton path emanate from an end point e , to obtain a skeleton based shape representation $\mathbf{g}(S(F))$, for each subregion $SR_i, i \in (1, 2, \dots, 21)$, we perform max pooling as follow:

$$\mathbf{g}_i(S(F)) = \max(\mathbf{c}^e | e \in SR_i), \quad (3)$$

where the ‘‘max’’ function is performed in an element-wise manner, i.e. for each codeword, we take the max value of all skeleton codes in a subregion. Max pooling is robust to noise and has been successfully applied to image classification. Thus, $\mathbf{g}(S(F))_i$ is a K dimensional feature vector of the subregion SR_i . The skeleton based shape representation $\mathbf{g}(S(F))$ is a concatenation of the feature vectors of all subregions:

$$\mathbf{g}(S(F)) = (\mathbf{g}_1^T(S(F)), \mathbf{g}_2^T(S(F)), \dots, \mathbf{g}_{21}^T(S(F)))^T. \quad (4)$$

Finally, $\mathbf{g}(S(F))$ is normalized by its ℓ_2 norm: $\mathbf{g}(S(F)) = \mathbf{g}(S(F)) / \|\mathbf{g}(S(F))\|_2$.

3.3 Shape Classification by Combining Contour and Skeleton

For a given shape F , to combine its contour $C(F)$ and skeleton $S(F)$ to classify it, we simply concatenate its BCF representation $\mathbf{f}(C(F))$ with its BSP representation $\mathbf{g}(S(F))$ to form a shape feature vector: $\mathbf{x}(F) = (\mathbf{f}^T(C(F)), \mathbf{g}^T(S(F)))^T$. Given a training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^M$ consists of M shapes from L classes, where \mathbf{x}_i and $y_i \in \{1, 2, \dots, L\}$ are the concatenated shape feature vector and the class label of i -th shapes respectively, we train a multi-class linear SVM [6] as the classifier:

$$\min_{\mathbf{w}_1, \dots, \mathbf{w}_L} \sum_{j=1}^M \|\mathbf{w}_j\|^2 + \alpha \sum_i \max(0, 1 + \mathbf{w}_{l_i}^T \mathbf{x}_i - \mathbf{w}_{y_i}^T \mathbf{x}_i), \quad (5)$$

where $l_i = \arg \max_{l \in \{1, 2, \dots, L\}, l \neq y_i} \mathbf{w}_l^T \mathbf{x}_i$ and α is a parameter to balance the weight between the regularization term (left part) and the multi-class hinge-loss term (right part). For a testing shape vector \mathbf{x} , its class label is given by

$$\hat{y} = \arg \max_{l \in \{1, 2, \dots, L\}} \mathbf{w}_l^T \mathbf{x}. \quad (6)$$

4 Experimental Results

In this section, we evaluate our method on two shape benchmarks and give the comparisons with the state-of-the-arts.

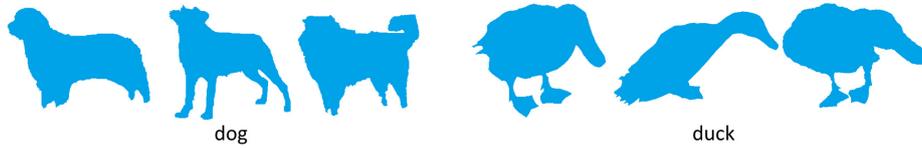


Fig. 3. Two classes with large intra-class variations in the Animal dataset [3]

4.1 Experimental Setup

The parameters introduced in our method are set as follows: the number of sampled points on a skeleton path $T = 50$, the codebook size $K = 1000$, the number of the nearest neighbors used for skeleton paths encoding $k = 5$ and the weight between the regularization term and the multi-class hinge-loss term in the multi-class linear SVM formulation $\alpha = 10$. We adopt the default parameter settings reported in [20] and [15] to extract BCF shape representation and skeletons, respectively.

For all the shape benchmarks, we randomly select half of the shapes in each class as the training samples and use the rest half for testing. To avoid the bias caused by randomness, such a procedure is repeated 10 times. Average classification accuracy and standard deviation are reported to evaluate the performance of different shape classification methods.

4.2 Animal Dataset

We first test our method on the Animal Dataset [3], which contains 2,000 animal shapes from 20 classes, such as bird, cat and dog. This dataset is the most challenging shape dataset, as each class has 100 shape images with large intra-class variations caused by view point change and significant deformation, such as the shapes shown in Fig. 3. The performances of the proposed method as well as other competing methods are depicted in Table 1. Our method achieves the best performance which outperforms BCF [20] by over 2%. This result proves that the skeleton based mid-level representation is complementary to contour based. Note that, our method is significantly superior to ICS [3], which proves that the proposed combination approach for contour and skeleton is more effective.

Table 1. Classification accuracy comparison on Animal dataset [3]

	Contour Segments [17]	IDSC [12]	ICS [3]	BCF [20]	Ours
Accuracy	69.7%	73.6%	78.4%	83.40 ± 1.30%	85.50 ± 0.88%

4.3 Mpeg7 Dataset

Mpeg7 dataset [10] is the most well-known shape dataset, which contains 1,400 shapes, including animals, artificial objects and symbols. It has 70 classes, in each

of which there are 20 different shapes. Table 2 demonstrates the classification accuracies obtained by competing methods. Our method also outperforms others on this dataset, which shows its generality.

The performance gain achieved by the proposed method compared to ICS [3] are mainly due to two reasons: (1) Skeleton is sensitive to contour noise, however max pooling provides the robustness to noise for BSP. Therefore, the combination of BCF and BSP does not induce additional noise; (2) The adopted multi-class linear SVM automatically learns the weights of contour and skeleton features and offers discriminative power, further increasing the accuracy.

Table 2. Classification accuracy comparison on Mpeg7 dataset [10]

	Contour Segments [17]	ICS [3]	BCF [20]	Ours
Accuracy	90.9%	96.6%	97.16 \pm 0.79%	98.35 \pm 0.63%

5 Conclusion

We have proposed a principled way to explore the complementary nature between contour and skeleton for shape recognition. A contour is represented by Bag of Contour Fragments; While for a skeleton, a novel skeleton based mid-level representation named Bag of Skeleton Paths has been proposed, for the purpose of capturing the geometric features along skeleton paths. Concatenating such two mid-level representations into one provides a compact and informative shape feature vector, which can be well handled by discriminative classifiers, such as multi-class linear SVM. The experimental results obtained on standard benchmarks verify the effectiveness of the proposed combination methods and demonstrate that it consistently outperforms the current state-of-the-art.

Acknowledgements. This work was supported in part by the National Natural Science Foundation of China under Grant 61303095 and Grant 61222308, in part by Innovation Program of Shanghai Municipal Education Commission under Grant 14YZ018, in part by Research Fund for the Doctoral Program of Higher Education of China under Grant 20133108120017 and in part by the Excellent Ph.D. Thesis Funding in Huazhong University of Science and Technology and Microsoft Research Asia Fellow 2012.

References

1. Aslan, C., Erdem, A., Erdem, E., Tari, S.: Disconnected skeleton: shape at its absolute scale. *IEEE Trans. Pattern Analysis and Machine Intelligence* 30(12), 2188–2203 (2008)
2. Bai, X., Latecki, L.: Path similarity skeleton graph matching. *IEEE Trans. Pattern Analysis and Machine Intelligence* 30(7), 1282–1292 (2008)

3. Bai, X., Liu, W., Tu, Z.: Integrating contour and skeleton for shape classification. In: ICCV Workshops, pp. 360–367 (2009)
4. Bai, X., Rao, C., Wang, X.: Shape vocabulary: A robust and efficient shape representation for shape matching. *IEEE Trans. Image Processing* 23(9) (2014)
5. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Analysis and Machine Intelligence* 24(4), 509–522 (2002)
6. Crammer, K., Singer, Y.: On the algorithmic implementation of multiclass kernel-based vector machines. *Journal of Machine Learning Research* 2, 265–292 (2001)
7. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR, pp. 886–893 (2005)
8. Daliri, M.R., Torre, V.: Robust symbolic representation for shape recognition and retrieval. *Pattern Recognition* 41(5), 1782–1798 (2008)
9. Felzenszwalb, P.F., Schwartz, J.: Hierarchical matching of deformable shapes. In: CVPR (2007)
10. Latecki, L.J., Lakämper, R., Eckhardt, U.: Shape descriptors for non-rigid shapes with a single closed contour. In: CVPR, pp. 1424–1429 (2000)
11. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: CVPR, pp. 2169–2178 (2006)
12. Lin, H., Jacobs, D.W.: Shape classification using the inner-distance. *IEEE Trans. Pattern Analysis and Machine Intelligence* 29(2), 286–299 (2007)
13. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2), 91–110 (2004)
14. Sebastian, T., Klein, P., Kimia, B.: Recognition of shapes by editing their shock graphs. *IEEE Trans. Pattern Analysis and Machine Intelligence* 26(5), 550–571 (2004)
15. Shen, W., Bai, X., Yang, X., Latecki, L.J.: Skeleton pruning as trade-off between skeleton simplicity and reconstruction error. *Science China Information Sciences* 56(4), 1–14 (2013)
16. Siddiqi, K., Shokoufandeh, A., Dickinson, S., Zucker, S.: Shock graphs and shape matching. *Int'l J. Computer Vision* 35(1), 13–32 (1999)
17. Sun, K.B., Super, B.J.: Classification of contour shapes using class segment sets. In: CVPR, pp. 727–733 (2005)
18. Wang, B., Shen, W., Liu, W., You, X., Bai, X.: Shape classification using tree-unions. In: ICPR, pp. 983–986 (2010)
19. Wang, J., Yang, J., Yu, K., Lv, F., Huang, T.S., Gong, Y.: Locality-constrained linear coding for image classification. In: CVPR, pp. 3360–3367 (2010)
20. Wang, X., Feng, B., Bai, X., Liu, W., Latecki, L.J.: Bag of contour fragments for robust shape classification. *Pattern Recognition* 47(6), 2116–2125 (2014)