Computer Vision and Image Understanding
journal homepage: www.elsevier.com

Shape classification using bag of edge contexts

Nikita Lomov[a,b,**], Sergey Sidyakin[b]

[a]*Lomonosov Moscow State University, Moscow, Russia*
[b]*The Federal State Unitary Enterprise "State Research Institute of Aviation Systems", Moscow, Russia*

ABSTRACT

This paper addresses 2D shape representation and classification issues. Generally, the existing main shape representations are contour-based and skeleton-based. Pure contour-based representations are more successful than pure skeleton-based ones. However, the current progress in shape classification is bound to contour representations enriched with the information from skeleton either implicitly or explicitly. In this paper we propose a novel skeleton-based shape descriptor named Edge Context. Edge contexts retain information about the global structure of the entire shape from a local point of view. If non-rigid deformations do not break the topology and local thicknesses are not changed too much they are stable to flexible deformations of the shape parts. Edge contexts are encoded in the Bag of Features framework and Bag of Edge Contexts is produced. Such descriptor can explicitly improve a contour descriptor when both of them are combined into a single feature vector. We also provide quite efficient algorithms for construction of the mentioned descriptors using a continuous skeleton. Extensive experimentation have been conducted. Obtained results indicate that proposed shape representation achieves the state-of-the-art performances on several challenging shape datasets.

1. Introduction

The study of 2D shapes is hampered by the fact that color and texture information contained in conventional color images is not available. Therefore traditional approaches to the object classification in color images can not be directly applied.

According to the latest research (Bai et al., 2009; Shen et al., 2014; Wang et al., 2014), contour and skeleton representations are popular in the shape domain. Contour-based approaches (Belongie et al., 2002; Latecki and Lakämper, 2000; Ling and Jacobs, 2007) are suitable for representing external shape information. They are robust against occlusions to a certain extent. However contour approaches are sensitive to articulation and non-rigid shape deformations (Bronstein et al., 2008). Skeleton-based approaches (Aslan et al., 2008; Bai and Latecki, 2008; Mestetskiy, 2013, 2015; Sebastian et al., 2004; Siddiqi and Pizer, 2008) can cope well with non-rigid deformations and describe internal shape structure as the maximum inscribed circles centered on the shape skeleton carry information about the thickness of the figure parts.

Today many scientists begin to apply the powerful learning approaches and design new convenient shape feature representations for recognition and classification tasks. Modern contour descriptors are: a transformed contour segment ("signature") (Bai et al., 2009; Sun and Super, 2005), Shape Context (Belongie and Malik, 2000), a set of histograms of distances from each point of the contour segment to the remaining points on this segment (Contour Segment Context), Bag of Contour Fragments (BCF) (Wang et al., 2014). Modern skeleton descriptors are: an ordered normalized set of radii of the maximal inscribed circles (Bai et al., 2009), Bag of the Shortest Paths (BSP) (Shen et al., 2014). All of these descriptors can be represented in a vector form (e.g., 1D histogram). Such way of representation is convenient for shape comparison and for rapid search in the database. In this case shape comparison and search are reduced to a vector comparison, to which machine learning methods can be directly applied.

If we look at the shape classification results produced by modern methods we can see that contour approaches are more successful than skeleton approaches. We think that this might be due to occlusions and rigid deformations that appear more often in shape datasets then non-rigid deformations. Skeleton is also sensitive to the object holes. Therefore, the current

**Corresponding author: Tel.: +7-926-134-6498;
e-mail: nikita-lomov@mail.ru (Nikita Lomov)

progress in shape classification is bound to the contour descriptors (Bai et al., 2014; Wang et al., 2014) or the contour descriptors enriched with the skeleton information (Bai et al., 2009; Shen et al., 2014, 2016). Creation of pure skeleton descriptor that is more successful then a pure contour descriptor is still an open problem.

BCF (Wang et al., 2014) is the most effective pure contour descriptor today. BCF introduces the successful Bag of Contour Fragments framework to shape recognition that was used to build more powerful complex descriptors. BCF can be improved by information that is contained inside the shape. The first way to enrich BCF is introduced in paper (Shen et al., 2014). Authors propose to concatenate BCF and BSP for shape recognition. However, the result is not much satisfactory since accuracy is raised from 83,4% to 85,5% on the Animal dataset (Bai et al., 2009). Shen et al. do not report the accuracy of the BSP descriptor without BCF. So we have reasons to assume that pure BSP performs worse than BCF. We think that this happens because sequences of radii associated with skeleton paths are much less various and unique when compared to the shape contexts of contour fragments. Therefore they have a weak descriptive ability. BSP describe only local shape parts and do not analyze global shape characteristics.

The second approach to enrich BCF is proposed in paper (Shen et al., 2016), where each contour point is associated with a thickness value of the corresponding skeleton point. Skeleton-associated Shape Context (SSC) is build for three point features: thickness differences, distances and orientations. BCF framework is applied to SSC to build the Bag of Skeleton-associated Contour Parts (BSCP) descriptor. BSCP turned out to be more successful than BCF and BSP combination.

In this paper we have two goals. The first goal is to design a new pure skeleton descriptor that is superior to the other skeleton descriptors. The second goal is to address the contour skeleton combination issue.

To achieve the first goal we propose a novel skeleton-based shape descriptor similarly to the shape contexts (Belongie and Malik, 2000; Shen et al., 2016). In contrast to the previous works we do not use contour points at all. Our basis is a continuous skeleton and a radial function associated with it (Fig. 1d, Fig. 2) that could be obtained by means of analytical geometry (Mestetskiy, 1999, 2013). Since basic element of the continuous skeleton is an edge then we consider edges instead of points as basic elements. The continuous skeleton representation in the form of a geometric graph allows to represent a distance on the skeleton in the form of a continuous line along skeleton edges and easily determine its length. Therefore, we propose to build an Edge Context (EC) for each skeleton edge. Our edge context is also a 3D tensor, in which the three dimensions describe the distance differences, orientation differences and average thickness differences of the chosen skeleton edge relative to the other skeleton edges. So, the difference between the edge context and the skeleton-associated shape context is evident, i.e. SSC does not consider distance differences and orientation differences and it is computed for contour points. Another major difference from the usual shape context and SSC lies in the fact that each edge contributes to the edge context proportional to the area of the edge-associated region. We consider proper regions (Mestetskiy, 2015) as the edge-associated regions. Without diving too early in the details we define a proper region as a polygon, which is cut off from the figure by the line segments (Fig. 3). Proper regions define a partition of the figure into disjoint polygonal regions. Rarity and uniqueness of regions is higher than uniqueness of lines and points, so the shape region information is taken into account in the edge context. Therefore, our edge context has area nature that leads to more robust performance when compared to other skeleton descriptors. Proposed edge context retains information about the structure of the entire shape from a local point of view.

The idea of EC is closely related to the histograms of geodesic distances between shape points (Donoho et al., 2006; Hamza and Krim, 2005; Mahmoudi and Sapiro, 2008; Osada et al., 2001). Such histograms are not restricted to any specificity of the object domain, shape limitations or data dimension. Therefore they are capable of describing 2D, 3D and even multidimensional data. Classification of these histograms can be carried out on the basis of distances between points and on the basis of statistics that are calculated from distances. Euclidean distances, geodesic distances and diffusion distances (Donoho et al., 2006) between shape points are usually applied. Full matrix of pairwise distances, shape distributions (Osada et al., 2001), the average distance from each point to all others (Hamza and Krim, 2005) are often used as the descriptive statistics. The difference between edge contexts and histograms of geodesic distances lies in the fact that we do not compute distances between shape points, we compute distances between regions of shape points associated with skeleton edges and these distances are routed on the skeleton. The partition of shape points into regions happens automatically when we compute the continuous skeleton. So the problem of distribution of shape points between skeleton edges is implicitly solved for us. As a result, our descriptor can be quite efficiently computed using the conception of continuous skeleton. The computation time of the geodesic histogram is proportional to the square of shape points and the computation time of the edge context of the full shape is proportional to the square of skeleton edges. Fortunately, the number of skeleton edges is usually less than the number of shape points by one order of magnitude.

We are not restricted to the edge context. Following the framework of Bag of Contour Fragments we obtain a feature vector of the entire shape by encoding and then pooling edge contexts. We term our descriptor as Bag of Edge Contexts (BEC). Let us briefly describe it here. For a given shape firstly a normalization step is performed to align the shape according to its major axis as the Spatial Pyramid Matching (SPM) method (Lazebnik et al., 2006) is not rotation invariant. Then the continuous skeleton of the shape is extracted. For each continuous skeleton edge we build the corresponding edge context. After that edge contexts are encoded into edge codes. For this purpose we adopt local-constrained linear coding (LLC) (Wang et al., 2010). Finally, the edge codes are pooled into a compact shape feature vector by SPM. The obtained shape feature vector can be fed into any discriminative models (e.g. SVM) to perform shape classification.

To achieve the second goal we concatenate Bag of Edge Contexts with Bag of Contour Fragments similarly to the BSP method (Shen et al., 2014). It is possible to do since both descriptors are feature vectors. We propose to use the BEC-BCF combination for shape classification tasks.

Our contributions can be summarized in three aspects. First, we propose a new shape representation named Edge Context. It retains global skeleton information relative to the local skeleton edge. Second, we build a novel skeleton-based shape descriptor named Bag of Edge Contexts, which encodes edge contexts in the Bag of Features framework. BEC is a compact feature vector and it is suitable for learning discriminative classifiers. Third, we suggest to combine BEC and BCF for shape classification. Using this combination we achieve the state-of-the-art results on several most popular shape benchmarks surpassing the results of the other modern descriptors (Bai et al., 2014; Li et al., 2010; Lim and Galoogahi, 2010; Ling and Jacobs, 2007; Ozay et al., 2015; Shen et al., 2016; Sun and Super, 2005; Wang et al., 2014) and descriptor combinations (Bai et al., 2009; Shen et al., 2014).

The rest of the paper is organized as follows. We review the related works in Section 2. The details of the proposed shape descriptors are described in Section 3. Firstly, the focus is on the continuous skeleton representations, including the basic definitions and references. Secondly, we explain how to build the edge context and BEC, describing computation details in several subsections of section 3. Our descriptors are evaluated on several most used shape benchmarks in Section 4. Section 5 concludes the paper.

2. Proposed approach

In this section we propose our descriptors for shape recognition. Since our shape representations are constructed by means of continuous skeleton, the first subsection introduces a concept of continuous skeleton. It gives important definitions and describes fast algorithms for continuous skeleton computation.

2.1. Continuous skeletons of binary shapes

Originally, the concept of skeleton was introduced for the continuous objects by introduced in Blum (1967, 1973). According to Blum, skeleton of the object is the locus of the centres of the maximal disks that can be inscribed inside the object boundary. Then the skeleton concept was generalized for the discrete binary images (2D shapes) by discrete and continuous approaches. The discrete approach applies a morphological transformation based on distance maps, thinning and Voronoi diagrams of boundary points (Siddiqi and Pizer, 2008) to the binary shape (Fig. 1a). It constructs the discrete skeleton image (Fig. 1b). The main advantage of the discrete approach is the simplicity of the algorithms and a graphic visualization of the skeleton in the source raster format. The continuous approach approximates the binary shape by the polygonal figure (Fig. 1c) and constructs the skeleton by means of analytic geometry (Fig. 1d). This skeleton is considered as a continuous skeleton of binary shape. The main advantage of the continuous approach (Siddiqi and Pizer, 2008) is a continuous skeleton representation of binary shape as a geometrical graph with

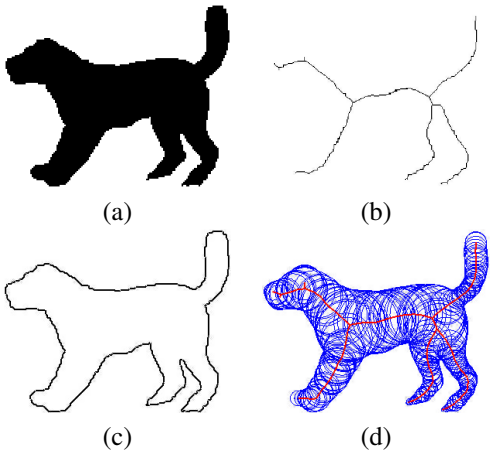


Fig. 1. The concept of skeleton: the binary shape (a); the discrete skeleton (b) obtained by thinning the binary shape; the polygonal shape boundary (c); the continuous skeleton (d) of the polygonal shape boundary. Inscribed circles (d) illustrate the values of radial function at skeleton points.

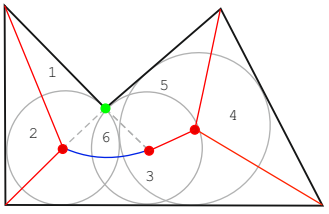


Fig. 2. The continuous skeleton and the Voronoi diagram. Voronoi sites: site-segments are shown in black, the site-point is green. Voronoi regions are represented with numbers. Skeleton linear edges, internal end-points of edges are shown in red. The parabolic edge is shown in blue. Inscribed circles for end-points are grey.

a radial function, which determines the width of shape parts. Continuous skeleton representations allow the usage of graph theory and computational geometry algorithms for image shape analysis and recognition. The second advantage of the continuous approach is its high computational efficiency. Experiments show that the continuous skeleton computation procedure is significantly faster than the discrete skeleton computation procedure. Even taking into account that the discrete approach is more suited for parallel computing (Mestetskiy, 2013). Other advantages are mathematical rigor and information content.

The continuous skeleton of the binary shape can be obtained from a Voronoi diagram of the polygonal shape boundary (Fig. 2). The polygonal shape boundary is a union of linear segments and vertices (points), which are considered as the Voronoi sites. The Voronoi diagram of these sites is generated and the skeleton is extracted as a subset of the diagram. The continuous skeleton has a fairly simple structure since it is a geometrical planar graph, which edges are described by straight line segments and quadratic parabola segments. Skeleton vertices are convex vertices of the polygonal figure (one degree vertices) and also points - centers of inscribed circles, tangent to the figure boundary in three or more points (three and more degree vertices). The radial function is defined in each skeleton point as the radius of inscribed circle centered in this point.

Note that the continuous skeleton of the polygonal figure always exists and it is unique.

There are known effective $O(n \cdot log(n))$ algorithms for the Voronoi diagram construction for the case of a simple polygon (Lee, 1982) and for the case of complex multiply-connected polygonal figures (Mestetskiy, 2013). In our implementation we stick to the more general Mestetskiy algorithm (Mestetskiy, 1999, 2013). According to it, the polygonal representation of the binary shape boundary is obtained by Minimal Perimeter Polygon (MPP) algorithm. It is guaranteed that the polygonal representation has no self-intersections. The idea of the skeleton computation (Mestetskiy, 2013) is to build a boundary adjacency tree from the polygonal description by a sweepline method and a Delaunay graph from the boundary adjacency tree. The Voronoi diagram is constructed from the Delaunay graph. Finally, the continuous skeleton is extracted as a subset of the Voronoi diagram and each skeleton edge together with its radial function gets a parametric description for convenience and compactness.

Thus, the continuous skeleton was discussed. We build our edge context descriptor based on the presented continuous skeleton representation in the next subsection.

2.2. Edge context

Suppose $B(X)$ is a polygonal shape boundary that was obtained by MPP algorithm (Mestetskiy, 2013) for the given shape X. We compute continuous skeleton $Sk(X)$ with edges $\{e_i\}$ from $B(X)$ by Mestetskiy algorithm (Mestetskiy, 1999, 2013). Then we assign a proper region P_i to each skeleton edge e_i (Fig. 3). A proper region is a polygon, which is cut off from the figure by the straight line segments. These line segments connect the edge ends and their projections to the site generators (site-segments and site-points). Proper regions of edges cover the entire polygonal figure completely and intersect each other only by their boundary parts. Therefore, the polygonal figure area is equal to the sum of the proper regions areas. Partition of the shape into proper regions provides a natural way to distribute shape points between skeleton edges. As a result, proper regions provide a way to assign edges with meaningful characteristics that can be computed from regions.

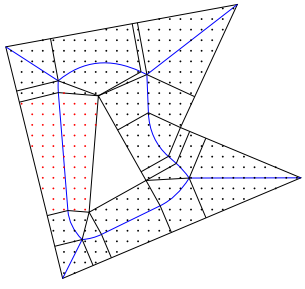


Fig. 3. The partitioning of the shape into proper regions. Each proper region corresponds to one continuous skeleton edge. Grid points of one proper region are shown in red.

We describe each edge in relation to the other skeleton edges using edge context ec_i. Suppose we have n edges $\{e_i\}$, $i \in [1, \dots, n]$. To find ec_i of e_i we compute the joint histogram of three edge features relative to other $n-1$ edges. Each edge e_j, $j \in [1, ..., n]$, $j \neq i$ contributes to the joint histogram proportional to the area s_j of its proper region P_j.

We consider the distance d_{ij} on the skeleton between centers (midpoints) of two edges e_i, e_j as the first relative edge feature. We use the difference $t_{ij} = t_i - t_j$ between average thicknesses t_i, t_j of two proper regions P_i, P_j that correspond to edges e_i, e_j as the second relative edge feature. We consider the rotation angle ϕ_{ij} between edges e_i and e_j as the third relative feature. The rotation angle ϕ_{ij} is calculated according to the following equation:

$$\phi_{ij} = \begin{cases} \alpha_i - \alpha_j, \text{ if } \alpha_i \geq \alpha_j; \\ \pi - (\alpha_i - \alpha_j), \text{ if } \alpha_i < \alpha_j, \end{cases}$$

where α is an inclination angle of the tangent to the axis of abscissas. The tangent is drawn in the edge center. Note that $d_{ij} = d_{ji}$, whereas for the other two relative features the order in which we consider the edges is important: $t_{ij} = -t_{ji}$; $\phi_{ij} = \pi - \phi_{ji}$, except for the case when $\phi_{ji} = 0$. In the presented form these tree relative edge features are invariant to shift and rotation. We additionally normalize edge features to ensure the scale invariance and to standardize the features:

$$d^*_{ij} = \left(\frac{d_{ij}}{\max_{i,j} d_{ij}} \right)^{\beta}$$

$$t^*_{ij} = \left(\frac{t_i}{r_{max}} \right)^{\gamma} - \left(\frac{t_j}{r_{max}} \right)^{\gamma},$$

$$\phi^*_{ij} = \frac{\phi_{ij}}{\pi}.$$

The parameters $\beta > 0$ and $\gamma > 0$ control the sensitivity to feature changes at different levels. For example, if $\beta < 1$ then edge context contains more information about changes in small distances, if $\beta > 1$ then edge context contains more information about changes in large distances. Features d^*_{ij} and ϕ^*_{ij} belong to the interval [0,1]; t^*_{ij} belongs to [-1,1]. The intervals are divided into equal parts: k parts for thickness differences, m parts for distances, n parts for rotation angle differences. Thus, we compute edge context ec_i of e_i as a histogram that contains $k \cdot m \cdot n$ bins:

$$\begin{aligned} & ec_i(a,b,c) = \sum_j \{ s_j : (d^*_{ij}, \phi^*_{ij}, t^*_{ij}) \in bin(a,b,c) \}, \\ & a = \lfloor (k-\epsilon)(t^*_{ij} + 1)/2 \rfloor, \\ & b = \lfloor (m-\epsilon) d^*_{ij} \rfloor, \\ & c = \lfloor (n-\epsilon) \phi^*_{ij} \rfloor, \end{aligned} \tag{1}$$

where ϵ denotes a very small positive value. We normalize edge context (1), so the sum of histogram values is equal to 1:

$$ec_i(a,b,c) = ec_i(a,b,c) / \sum ec_i(a,b,c). \tag{2}$$

Edge contexts retain global structure information of the entire shape, but they do it from a local point of view. This is a reason why they are quite informative and unique. For example, running a little forward, in the animal recognition task the proposed descriptor is able to distinguish the characteristic shape

parts: tail, ears, paws, head, body, etc. Therefore all shape parts have distinctive histograms (Fig. 4). We transform each edge context from a 3D tensor to a vector for further processing, so $ec_i \in R^{d\times 1}$, where $d = k \cdot m \cdot n$ is the dimension of the feature vector of e_i.

2.3. Bag of edge contexts

We adopt edge contexts as basic shape features for shape codebook learning and building our shape representation. Therefore, we apply Bag of Features approach to our edge contexts. Let us describe it.

Encoding edge contexts is to map feature vectors of skeleton edges into a new space spanned by a codebook V. In this new space edge contexts $\{ec_i\}$ are represented by edge codes $\{w_i\}$. A set of edge contexts for training is randomly selected from all available edge contexts. K-means algorithm (Duda et al., 2001) is applied on the selected feature vectors for clustering. The clustering centers are used as a codebook $V = [v_1, \ldots, v_K] \in R^{d\times K}$, where each column $v_i, i \in [1, \ldots, K]$ is a clustering center.

Local-constraint linear coding (LLC) (Wang et al., 2010) is applied to compute edge codes since it is both fast and effective. To represent ec_i in the space spanned by V, LLC uses m nearest neighbours in V as local bases for ec_i to form a local coordinate system. The m nearest neighbors of ec_i are denoted as $V_{\pi_i} \in R^{d\times m}$, where $\pi_i = \{\pi_i^1, \ldots, \pi_i^m\}$ is a set containing the indexes of the m nearest neighbors in V. V_{π_i} is a matrix consisting of the $\pi_i^1, \ldots, \pi_i^m$-th columns of V. We suppose that ec_i and its nearest neighbors lie close to a local linear patch of the manifold. The local geometry of ec_i is characterized by linear coefficients obtained through reconstruction of ec_i from V_{π_i}. The coefficients $w_{\pi_i} \in R^{m\times 1}$ are obtained by minimization of the following expression:

$$\min_{w_{\pi_i}} \|ec_i - V_{\pi_i} w_{\pi_i}\|^2 \text{ s.t. } \mathbf{1}^T w_{\pi_i} = 1, \qquad (3)$$

where weight vector w_i summarizes the contributions of local bases to edge contexts reconstruction, which is required to be summed to 1. The minimization problem in (3) is a small-scale least square problem and its time complexity is $O(m^2)$.

The edge code of ec_i is $w_i \in R^{K\times 1}$, where values of the $\pi_i^1, \ldots, \pi_i^m$ entries of w_i are equal to w_{π_i} and the rest w_i entries are set to zero.

We build a compact shape representation based on statistics of edge codes w_i utilizing spatial pyramid matching (SPM) (Lazebnik et al., 2006). For this purpose we divide shape into 1×1, 2×2 and 4×4 regions, as shown in Fig. 5. There are 21 regions in total.

We do max-pooling inside each region $r_l, l \in [1, \ldots, 21]$. Denote by w_z any encoded edge context at the position z inside X. The position of the edge code is defined as an edge center. Max-pooling operation takes the form:

$$h(X, r) = max(w_z | z \in r_l),\ h(X, r) \in R^{K\times 1}, \qquad (4)$$

where the max function works in a row-wise order, returning a feature vector of r_l. Accordingly, Bag of Edge Contexts (BEC) of X is a concatenation of $h(X, r)$ for all regions:

$$BEC(X) = [h(X, 1)^T, \ldots, h(X, 21)^T]^T. \qquad (5)$$

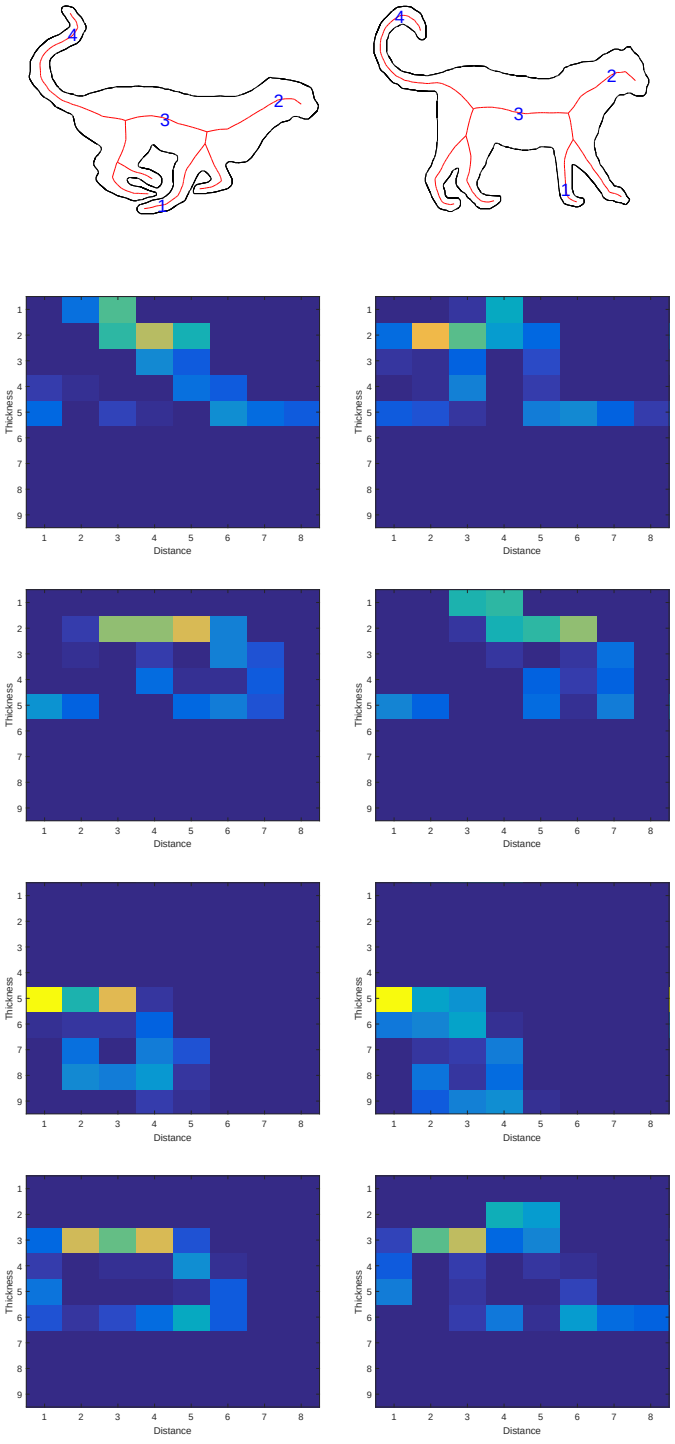


Fig. 4. Edge contexts of similar edges for different shapes. The first row: shapes, skeletons and edges of interest. Each next row shows contexts of corresponding edges (from 1 to 4) for two different shapes. Edge contexts are computed for distance and thickness features for the convenience of visualization. It can be seen that the feature distributions computed for similar parts are alike or quite close to each other. However, feature distributions computed for dissimilar parts are different.

The size of $BEC(X)$ is $21 \cdot K$.

Thus, the general view of edge contexts and BEC is proposed and we can proceed to computation of relative skeleton edge features that are mentioned above in subsection 3.2. Let us start

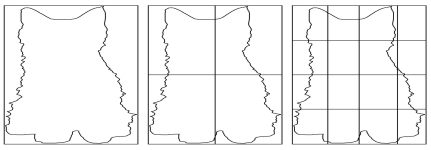


Fig. 5. The partitioning of the figure into 1×1, 2×2 and 4×4 regions.

with an average thickness of the proper region.

2.4. Proper region thickness

To compute an average thickness of each proper region, we need to find the thickness at each shape point q. We define the thickness at q as the maximum size r of inscribed circle that covers this point q.

The keystone idea of the effective thickness computation is derived from the fact that centers of all inscribed circles belong to the shape skeleton. This allows us to define the thickness function $T(q)$ for point q of the binary figure X in terms of the continuous skeleton $Sk(X)$ and the radial function $\rho(p), p \in Sk(X)$:

$$T(q) = \max_{p \in Sk(X)} \{\rho(p) : q \in D(p)\}, \tag{6}$$

where $D(p)$ is the maximal inscribed circle centered in p.

The skeleton can be represented as a union of subgraphs:

$$Sk(X) = \bigcup_{i=1}^{k} Sk_i$$

Consider Sk_i silhouettes:

$$G(Sk_i) = \bigcup_{p \in Sk_i} D(p), i = 1, \ldots, k.$$

Let the thickness function at the point q in relation to the subgraph Sk_i be given by:

$$T_i(q) = \begin{cases} \max_{p \in Sk_i}\{\rho(p) : q \in D(p)\} & \text{if } q \in G(Sk_i); \\ -1 & \text{otherwise.} \end{cases}$$

Consider equation (6) as

$$T(q) = \max_{i=1,\ldots,k} (T_i(q)).$$

A skeleton partitioning into subgraphs can simplify the procedure of calculating the thickness value at the point due to the following theorem.

Theorem 1. *Let $Sk_i = p(t) = (x(t), y(t)), t \in [0, 1]$ be a simple parametric curve such that $Sk_i \in Sk(X)$. The radial function $\rho(p)$ increases monotonically along Sk_i. The point q belongs to the silhouette $G(Sk_i)$. Then either the point q is covered by the maximal circle $D(p(1))$ of Sk_i, or there is a point $p', p' \in Sk_i$ such that $\rho(p')$ takes the value of $T_i(q)$ and q lies on the boundary of the inscribed circle centered at p'.*

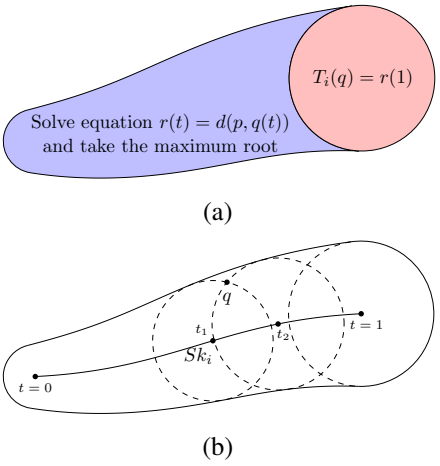


Fig. 6. The search for the covering circle of the maximum radius in the monotonic silhouette. The search scheme (a). The boundary circles for q (b). In this case $T_i(q) = r(t_2)$.

Proof. Consider the following function:

$$g(q, t) = r(t) - d(q, p(t)),$$

where $r(t) = \rho(p(t))$ and d is the Euclidean distance between the points. The function $g(q, t)$ is non-negative if the point q is covered by the inscribed circle centered at the point $p(t)$ and negative otherwise. Suppose that q does not belong to the maximal circle of the curve. Due to the continuity of the function $g(q, t)$ when q is fixed, the set $A = \{t : g(q, t) \geq 0\}$ is a compact set and $t', t' < 1$ is a maximum number of this set. It is obvious that $r(t') = T_i(q)$, but $g(q, t') = 0$ because of the continuity of the function $g(q, t)$ with reference to t and the inconstancy of its sign. □

Note. In the proof of the theorem we can replace the function $g(q, t)$ with $f(q, t) = r^2(t) - d^2(q, p(t))$. It is more convenient to operate with $f(q, t)$.

The usage of Theorem 1 is illustrated in Fig. 6.

Thus, we need only two steps for calculating the value of $T_i(q)$:

1. If $f(q, 1) \geq 0$, assign $T_i(q) = r(1)$.
2. Otherwise, solve the equation $f(q, t) = 0$ and assign $T_i(q) = r(t_{\max})$, where $t_{\max}$ is the maximum root.

We look for values of T_i at the points of a uniform grid N overlaid on the figure X with sufficiently small step (for example, points with integer coordinates). Computational procedure is organized as follows:

1. Represent skeleton of the figure X as a union of monotonic subgraphs $Sk_i, i = 1, \ldots, k$.
2. For all $q \in N$ take $T(q) = -1$
3. For all $i = 1, \ldots, k$
4. Change the value of $T_i(q)$ to $\max(T(q), r(1))$ for grid points that belong to the maximal circle of Sk_i
5. Find $T_i(q)$ by solving the equation $g(q, t) = 0$ for all other grid points in $G(Sk_i)$. To each $T_i(q)$ assign the value of $\max(T(q), T_i(q))$

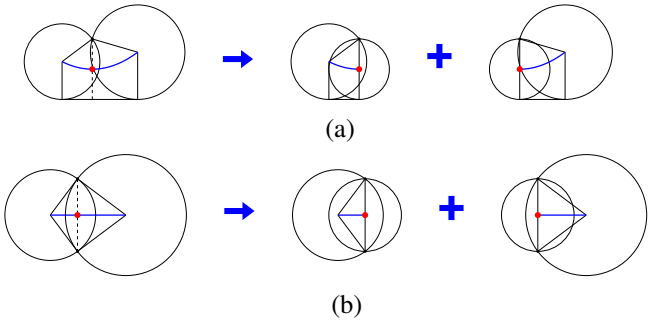


Fig. 7. Subdivision of non-monotonic parabolic (a) and hyperbolic (b) edges into monotonic parts. The cut (dashed line) passes through the point of radial function minimum (shown in red), which becomes the center of the small end circle for both new edges. Edges are shown in blue. End circles for all edges are displayed.

As been mentioned in subsection 3.1, if X is approximated by a polygon, the continuous skeleton can be efficiently obtained from the Voronoi diagram of its boundary elements (site-points and site-segments) cutting-off edges and vertices related to concave vertices of the figure. Therefore, all skeleton edges get explicit analytical description (Mestetskiy, 2013). The edges of three types are distinguished depending on the types of the site generators: linear edge (both sites are straight line segments), hyperbolic (both sites are points) and parabolic (one of the sites is a segment, the other is a site-point). A hyperbolic edge looks like a straight line segment.

A skeleton may contain edges that are not monotonic. In the non-monotonic edge there is always only one minimum of the radial function. So we divide such edges into two monotonic parts and consider them as separate skeleton subgraphs during the computation of thickness function at shape points (Fig. 7).

It remains to clarify the question of solving the equation $f(q,t) = 0$ for each of the three edge types. By (x_1, y_1), (x_2, y_2) denote coordinates of the edge ends. By r_1, r_2 denote radii of end circles such that $r_1 < r_2$. By (x_4, y_4) denote site-point coordinates of hyperbolic and parabolic edges.

1. The radial function $r(t)$ varies linearly during the uniform movement along the linear edge. So the linear edge is parameterized as follows:

$$x(t) = x_1 + t(x_2 - x_1);$$
$$y(t) = y_1 + t(y_2 - y_1);$$
$$r(t) = r_1 + t(r_2 - r_1).$$

The equation $f(q,t) = 0$ becomes quadratic as a result of parameterization:

$$(x_1+t(x_2-x_1)-x_0)^2+(y_1+t(y_2-y_1)-y_0)^2 = (r_1+t(r_2-r_1))^2.$$

2. The parabolic edge is a special case of the Bezier curve of order 2. Two of its control points coincide with the edge ends. The third is the intersection point of tangents to the parabola at the end points (Fig. 8). We have the following parameterization:

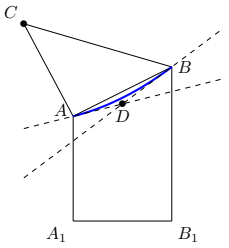


Fig. 8. The control point D of a parabolic edge; C is a site-point; A_1B_1 is a site-segment; AB is a parabolic edge, shown in blue.

$$x(t) = (1-t)^2 x_1 + 2tx_3 + t^2 x_2;$$
$$y(t) = (1-t)^2 y_1 + 2ty_3 + t^2 y_2;$$
$$r(t) = \sqrt{(x(t)-x_4)^2 + (y(t)-y_4)^2},$$

where (x_3, y_3) are the coordinates of the intermediate control point.

The equation $f(q,t) = 0$ takes on the form:

$$(x(t)-x_0)^2 + (y(t)-y_0)^2 = (x(t)-x_4)^2 + (y(t)-y_4)^2 \Rightarrow$$
$$-2x(t)x_0 + x_0^2 - 2y(t)y_0 + y_0^2 =$$
$$= 2x(t)x_4 + x_4^2 - 2y(t)y_4 + y_4^2,$$

that leads to a quadratic equation.

3. The hyperbolic edge is parameterized as follows:

$$x(t) = x_1 + t(x_2 - x_1);$$
$$y(t) = y_1 + t(y_2 - y_1);$$
$$r(t) = \sqrt{(x(t)-x_4)^2 + (y(t)-y_4)^2}.$$

The equation $f(q,t) = 0$ takes on the form

$$-2x(t)x_0 + x_0^2 - 2y(t)y_0 + y_0^2 = -2x(t)x_4 + x_4^2 - 2y(t)y_4 + y_4^2$$

That leads to a linear equation.

Thus, to calculate the thickness at the point in relation to the edge it is sufficient to solve the simplest equation.

The advantage of the proposed analytical approach to thickness calculation is calculating the exact real values of thickness function.

To calculate an average thickness of the edge proper region we select the grid points belonging to it (Fig. 3). It is easy to do since proper regions are unions of elementary geometric shapes taken from Voronoi diagram. Bypassing of the grid points is performed by means of the scan-line algorithm (Hearn et al., 2010).

Thus, an average thickness for each proper region P_i is computed as:

$$T(P_i) = \frac{\sum_{p \in N \cap P_i} T(p)}{\#\{p : p \in N \cap P_i\}}, \qquad (7)$$

where symbol # denotes the number of elements in the set.

Areas of the proper regions are quickly calculated analytically by means of continuous skeleton (Mestetskiy, 2013, 2015) in the next subsection.

2.5. Proper region area

The area of the proper region can be calculated using the radii r, R of end circles and the distance l between their centers.

The area S_{lin} of the proper region of a linear edge (Fig. 9) is defined as follows:

$$S_{lin} = (r + R)t, \quad t = \sqrt{l^2 - (R^2 - r^2)}.$$

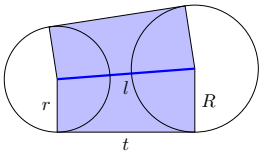


Fig. 9. The proper region of the monotonic linear edge.

The area S_{par} of the proper region of a parabolic edge (Fig. 10) is defined as follows:

$$S_{par} = \begin{cases} \Phi(R) + \Phi(r) & \text{if } t \geq 2\sqrt{r(R-r)} \\ \Phi(R) - \Phi(r) & \text{if } t < 2\sqrt{r(R-r)} \end{cases},$$

$$\Phi(z) = \frac{1}{2}(z + p)\sqrt{p(2z - p)},$$

$$p = \frac{t^2}{2l^2}(R + r + \sqrt{(R + r)^2 - l^2}).$$

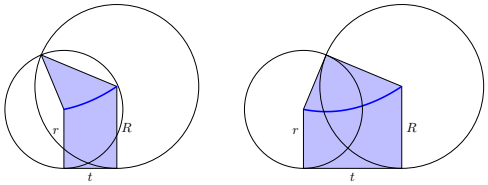


Fig. 10. The proper region of the monotonic (a) and non-monotonic (b) parabolic edge.

The area S_{hyp} of the proper region of a hyperbolic edge (Fig. 11) is defined as follows:

$$S_{hyp} = \begin{cases} \Psi(R) + \Psi(r) & \text{if } l^2 + r^2 \geq R^2 \\ \Psi(R) - \Psi(r) & \text{if } l^2 + r^2 < R^2 \end{cases},$$

$$\Psi(z) = \frac{q}{2}\sqrt{z^2 - \frac{q^2}{4}},$$

$$q = \frac{1}{l}(\sqrt{[(l + r)^2 - R^2] \cdot [R^2 - (l - r)^2]}).$$

It remains to clarify the question of the distance computation between two edges.

2.6. Distance between edges

Distance between any pair of edges is the shortest skeleton path between their centers (midpoints). It can be represented as a sequence of transitions from one middle point to another.

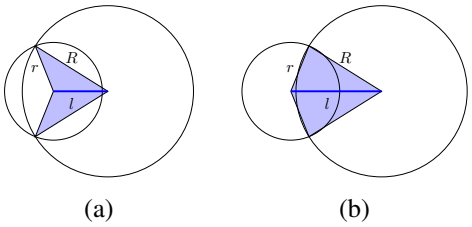


Fig. 11. The proper regions of the monotonic (a) and non-monotonic (b) hyperbolic edges.

Other midpoints are not visited under a single transition. Such transition is possible only for the adjacent edges within the skeleton, where the transition length is half of the sum of these edge lengths. In order to reduce the calculation of histogram to the classic problem of finding the shortest paths in a graph, we construct the graph Γ, where each edge e_i has the associated vertex v_i. If edges e_i and e_j are adjacent in the skeleton, then v_i and v_j are connected in both directions, otherwise the vertices are non-adjacent. We assign the weight to each edge of the new graph. The weight is equal to $(l_i + l_j)/2$. In other words, the weight equals distance on the skeleton between the centers of adjacent edges. An example of the skeleton and its dual graph Γ is shown in Fig. 12. The shortest path between the vertices of the dual graph corresponds to the shortest path between the edge centers in the skeleton. Therefore, Johnson algorithm (Johnson, 1977) is applied to the graph Γ to calculate the shortest paths between all pairs of edges.

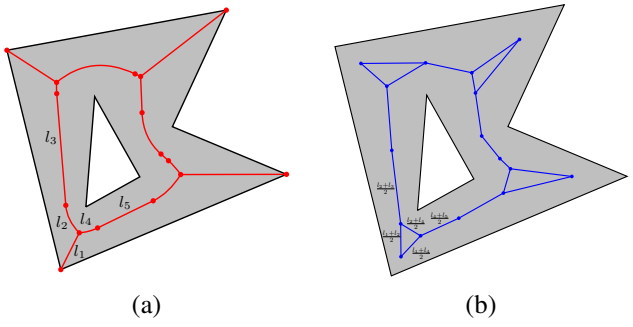


Fig. 12. Distances on the skeleton between pairs of skeleton edges: (a) the polygonal figure and its skeleton; (b) the dual graph, which is used to calculate distances; lengths (weights) of graph edges are specified locally.

Thereby, we explained how relative edge features are computed. In practice we are facing skeleton regularization problem. Our solution to this problem is given in the next section.

2.7. Regularization of continuous skeleton

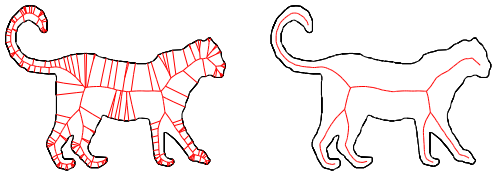
There are many uninformative noise edges in the continuous skeleton of a complex polygonal figure. Since for each skeleton edge we construct the edge context, noise edges produce redundant histograms and distort useful histograms of meaningful edges. Therefore, we provide a procedure for skeleton regularization. In this procedure the terminal edges are sequentially cut off from the skeleton until the given criteria is met. The terminal edge is an edge that have one terminal vertex among its

vertices. The terminal vertex is a vertex that belongs only to the one skeleton edge.

The entire shape is considered as a large region covered by inscribed circles centered at skeleton points. When the edge is clipped, the circles associated with this particular edge disappear. The union of the remaining circles is a silhouette of the skeleton subgraph. The skeleton subgraph silhouette is a subset of the entire shape and it is located inside the shape. If the Hausdorff distance between the shape and the skeleton subgraph silhouette is less than the specified threshold then the edge is cut off. Otherwise it is not cut off. Thus, the following regularization procedure is used:

1. Put all terminal edges in the candidate list L for cutting off.
2. While the list includes edges, extract one of them (e.g. e_i) and repeat the following steps:
   - 2..1 Consider the edge $e_i = AB$, where A is a terminal vertex. Suppose that AB is cut off. Compute the Hausdorff distance r between the shape and the obtained silhouette. The distance r is reached on a pair of points, one of which is the polygonal figure vertex achievable from the point B along previously cut edges, including AB. The other point lies on the boundary of the inscribed circle centered at B. Therefore, to determine r it is necessary to compute the values of $r = d(C_j, B) - \rho(B)$ for all achievable vertices C_j and take the minimum.
   - 2..2 If $r < \tau$, where τ is a maximal distance allowed, then
     - i. assign $t_j = \frac{s_i t_i + s_j t_j}{s_i + s_j}$ to all adjacent edges e_j;
     - ii. distribute the area s_i between adjacent edges equally;
     - iii. cut the edge e_i;
     - iv. if B becomes a terminal vertex, then add the edges connected to B to the end of L.
   - 2..3 Remove the edge from L.

Procedure of skeleton regularization is illustrated on Fig. 13.



(a) (b)

Fig. 13. Skeleton before and after regularization.

The distribution of the region area between the adjacent edges is required since the proper region of the cut edge must be considered. We also update the thickness since the average thickness of the adjacent edge changes after addition of the new region.

At this point we can proceed to the experimental section where classification performances are evaluated.

3. Experiments

We evaluate the proposed shape descriptors on the three 2D shape benchmarks with relatively larger number of shapes (Animal (Bai et al., 2009), MPEG7 (Latecki and Lakämper, 2000), Swedish leaf (Soderkvist, 2001) and give the comparisons with the state-of-the-arts.

3.1. Datasets preprocessing

Swedish leaf dataset consists of full-color images. So the segmentation procedure has to be carried out. We applied simple segmentation by thresholding the grayscale images. The threshold value was set automatically using the Otsu's method.

Holes represent another problem. The slightest change in the object topology critically affects the medial representation. Images in Animal contain only small redundant holes appearing not due to the self-occlusion. Filling the holes of less than 0.01% of the object area turned out to be sufficient. A similar option for the Swedish leaf was set to 0.6% in order to fill not only the redundant holes, but also the holes that really exist in the corrupted leaves. MPEG7 includes images with poor segmentation within the main contour. So all the holes in images of this dataset were filled. However, as a result some real holes were also filled. We adopt the default shape normalization technique reported in (Shen et al., 2014), (Shen et al., 2016) to extract BCF and BEC for all three datasets.

3.2. Implementation details

Parameters k, m and n are equal to 9, 8 and 6 respectively. Sensivity parameters β and γ are set to 0.5. Codebook size is set to 500 and extracted skeletons are regularized with maximum allowed Hausdorff distance $\tau = 8$. For BCF extraction we adopt the default parameter settings reported in (Wang et al., 2014).

Image classification is performed via SVM. For SVM learning we adopt the default parameters (Wang et al., 2014). The common strategy for evaluating shape classification performance (Bai et al., 2014; Wang et al., 2014) is half training. We randomly select half shapes in each class for training and use the rest shapes for testing in each round. This procedure is repeated for ten times and average classification accuracy, in percentage (%), is reported.

3.3. Animal

There are 20 animal classes in the Animal dataset. Each one comprises 100 examples. Objects are presented in various poses in each class, at the same time shapes significantly differ in size, location and articulation of parts. Some parts of objects are subjected to self-occlusions.

The classification accuracy obtained with BCE on the most challenging Animal dataset is 78.8%. This is quite high for a skeleton-based descriptor. BCE is the first skeleton descriptor designed for recognition that is better than ICS combination of two advanced contour (CS) and skeleton (SP) descriptors. At the same time it is better than the pure skeleton descriptor SP.

However, the lack of local contour information does not allow to achieve really advanced results. Therefore, we use our

descriptor in combination with BCF. BCE and BCF vectors are concatenated. BCE features are multiplied by 0.5 because we believe that contour information is more important.

The accuracy of the proposed descriptor with selected parameters is 90.44%. Thus, as can be seen from Table 1, the proposed approach is superior to the known modern methods (Bai et al., 2014; Li et al., 2010; Ling and Jacobs, 2007; Lim and Galoogahi, 2010; Ozay et al., 2015; Sun and Super, 2005; Wang et al., 2014), including those that combine contour and skeleton descriptions (Bai et al., 2009; Shen et al., 2014, 2016). BCF+BCE allows to improve the result for 7.04% compared to pure BCF. When Bag of Skeleton Paths allows to improve the result for 2.1% compared to pure BCF. As a result, BCF descriptor is successfully enriched by BCE.

To allow fair comparison with the algorithm based solely on contour features (Wang et al., 2014) we report classification accuracy for each of the 20 classes in Table 2. The proposed descriptor outperforms BCF for difficult classes that were previously recognized poorly (”Monkey”, ”Leopard”, ”Cat”, ”Cow”). So the benefits are demonstrated.

method	accuracy
CSS (Sun and Super, 2005)	69.7%
IDSC (Ling and Jacobs, 2007)	73.6%
CS (Bai et al., 2009)	71.7%
SP (Bai et al., 2009)	67.9%
ICS (Bai et al., 2009)	78.4%
BEC	78.8%
BCF (Wang et al., 2014)	83.4%
Shape Tree (Li et al., 2010)	80.0%
Lim (Lim and Galoogahi, 2010)	80.4%
Shape Vocabulary (Bai et al., 2014)	84.3%
BCF+BSP (Shen et al., 2014)	85.5%
MC-2 (Ozay et al., 2015)	85.9%
BSCP (Shen et al., 2016)	89.04%
BEC+BCF	90.44%

Table 1. Classification accuracy (half training) on the Animal dataset

3.4. MPEG7

MPEG7 dataset (Latecki and Lakämper, 2000) is a widely accepted shape dataset. It contains 1400 shapes of animals, artificial objects and symbols. It has 70 classes, including 20 different shapes for each class. Table 3 demonstrates the classification accuracies obtained by competing methods. For this dataset we have also performed leave-one-out cross-validation. Our method outperforms others on this dataset. So its generality is shown.

3.5. Swedish leaf dataset

The dataset contains leaves from 15 different swedish trees with 75 leaves per class. The difficulty with this dataset is that some leaves are indistinguishable to the untrained eye. We follow the experimental setting in (Ling and Jacobs, 2007), comparing classification accuracy obtained by our descriptor with other pure shape-based recognition methods in Table 4. One third of the dataset has been used for training. Classification accuracy on Swedish leaf has improved from 97.52% to 99.19%.

method	**bird**	**butterfly**	**cat**	**cow**
BCF	87.6%	92.2%	73.8%	77.4%
BEC+BCF	92.4%	97.6%	81.0%	86.2%
method	**crocodile**	**deer**	**dog**	**dolphin**
BCF	76.8%	90.4%	82.6%	89.0%
BEC+BCF	83.8%	93.0%	89.8%	93.4%
method	**duck**	**elephant**	**fish**	**flyingbird**
BCF	87.0%	95.2%	79.8%	72.0%
BEC+BCF	92.4%	96.4%	88.4%	83.6%
method	**hen**	**horse**	**leopard**	**monkey**
BCF	94.2%	95.4%	66.4%	58.4%
BEC+BCF	96.2%	96.8%	78.0%	78.4%
method	**rabbit**	**rat**	**spider**	**tortoise**
BCF	85.8%	70.6%	99.2%	93.6%
BEC+BCF	95.8%	87.8%	99.4%	98.0%

Table 2. Classification accuracy (half training) for each class on the Animal dataset

method	accuracy (half training)	accuracy (leave-one-out)
CSS (Sun and Super, 2005)	90.9%	97.93%
CS (Bai et al., 2009)	91.1%	-
SP (Bai et al., 2009)	86.7%	-
ICS (Bai et al., 2009)	96.6%	-
SoS (Daliri and Torre, 2006)	-	97.36%
RS (Daliri and Torre, 2008)	-	98.57%
KD (Daliri and Torre, 2010)	-	98.93%
BCF (Wang et al., 2014)	97.03%	98.86%
BCF+BSP (Shen et al., 2014)	98.35%	-
BEC+BCF	98.41%	99.00%

Table 3. Classification accuracy on MPEG7

method	accuracy
MAC (Soderkvist, 2001)	82.00%
Fourier (Ling and Jacobs, 2007)	89.60%
CS + DP (Ling and Jacobs, 2007)	88.12%
IDSC + DP (Ling and Jacobs, 2007)	94.13%
RS (Daliri and Torre, 2008)	95.47%
Shape Tree (Li et al., 2010)	96.28%
BCF (Wang et al., 2014)	97.52%
BEC+BCF	98.38%

Table 4. Classification accuracy on Swedish Leaf.

3.6. Performance

Dataset	BCF	EC	Encoding	BEC
Animal	3.518	0.682	0.141	0.823
MPEG7	2.390	0.709	0.103	0.812
Leaf	2.840	0.814	0.102	0.916

Table 5. Seconds per one shape on an average.

The experiments were carried out on a PC with Intel Core i5 750. The results shown in Table 5 confirm that edge contexts and BCE are quite efficiently computed.

4. Conclusions

In this paper we introduce two novel skeleton-based shape descriptors. The first one is a lower level descriptor named Edge Context. It is an extension of the shape context descriptor for skeleton case. The second one is a higher level shape descriptor named Bag of Edge Contexts. The edge context retains information about the global structure of the entire shape from a local point of view. Bag of Edge Contexts retains information about the global structure of the entire shape from all local points of view. So BEC is a truly global and rich descriptor. Both descriptors have not only a skeleton but an region (area) nature. Therefore they are stable to distortions of shape boundary. This makes them superior to the other known skeleton-based descriptors. For example, BSP and BCSP descriptors extract only thickness information from skeleton points and do not exploit structural internal information of the shape fully. Our descriptor extracts more robust information from the skeleton since it considers relations between shape regions. However this information is additionally enriched with shape region thickness features. And what is more, we consider orientation differences between skeleton edges, which make our descriptors more robust to rotations. Remembering the advantages of contour descriptors over skeleton descriptors, we propose to combine Bag of Edge Contexts and Bag of Contour Fragments for shape classification task. As a result, such combination is superior to any other known complex shape descriptor on several challenging shape benchmarks and achieves the state-of-the-art performance. We also describe effective analytical procedures to compute edge features, which are used in the edge context. The implementation (code) of our descriptors (including all experiments) is already available by request and soon will be available online.

Acknowledgement

This work was supported by Russian Science Foundation (RSF) under Grant 16-11-00082; by Russian Foundation For Basic Research (RFBR) under Grants 15-07-01323 A and 16-57-52042 MHT_a.

References

Aslan, C., Erdem, A., Erdem, E., Tari, S., 2008. Disconnected skeleton: Shape at its absolute scale. IEEE Trans. Pattern Anal. Mach. Intell. 30, 2188–2203.

Bai, X., Latecki, L.J., 2008. Path similarity skeleton graph matching. IEEE Trans. Pattern Anal. Mach. Intell. 30, 1282–1292.

Bai, X., Liu, W., Tu, Z., 2009. Integrating contour and skeleton for shape classification, in: Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, pp. 360–367.

Bai, X., Rao, C., Wang, X., 2014. Shape vocabulary: A robust and efficient shape representation for shape matching. IEEE Transactions on Image Processing 23, 3935–3949.

Belongie, S., Malik, J., 2000. Matching with shape contexts, in: Content-based Access of Image and Video Libraries, 2000. Proceedings. IEEE Workshop on, pp. 20–26.

Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. 24, 509–522.

Blum, H., 1967. A transformation for extracting new descriptors of shape, in: Wathen-Dunn, W. (Ed.), Models for the Perception of Speech and Visual Form. MIT Press, Cambridge, pp. 362–380.

Blum, H., 1973. Biological shape and visual science (part i). Journal of Theoretical Biology 38, 205–287.

Bronstein, A.M., Bronstein, M.M., Bruckstein, A.M., Kimmel, R., 2008. Analysis of two-dimensional non-rigid shapes. Int. J. Comput. Vision 78, 67–88.

Daliri, M.R., Torre, V., 2006. Shape recognition and retrieval using string of symbols, in: 5th International Conference on Machine Learning and Applications (ICMLA'06), pp. 101–108.

Daliri, M.R., Torre, V., 2008. Robust symbolic representation for shape recognition and retrieval. Pattern Recogn. 41, 1799–1815.

Daliri, M.R., Torre, V., 2010. Shape recognition based on kernel-edit distance. Comput. Vis. Image Underst. 114, 1097–1103.

Donoho, D., Chui, C., Coifman, R.R., Lafon, S., 2006. Diffusion maps. Applied and Computational Harmonic Analysis 21, 5 – 30.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern classification, 2nd ed.

Hamza, A.B., Krim, H., 2005. Probabilistic shape descriptor for triangulated surfaces., in: ICIP (1), IEEE. pp. 1041–1044.

Hearn, D.D., Baker, M.P., Carithers, W., 2010. Computer Graphics with Open GL. 4th ed., Prentice Hall Press.

Johnson, D.B., 1977. Efficient algorithms for shortest paths in sparse networks. J. ACM 24, 1–13. doi:10.1145/321992.321993.

Latecki, L.J., Lakämper, R., 2000. Shape similarity measure based on correspondence of visual parts. IEEE Trans. Pattern Anal. Mach. Intell. 22, 1185–1190.

Lazebnik, S., Schmid, C., Ponce, J., 2006. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), pp. 2169–2178. doi:10.1109/CVPR.2006.68.

Lee, D.T., 1982. Medial axis transformation of a planar shape. IEEE Trans. Pattern Anal. Mach. Intell. 4, 363–369.

Li, Y., Zhu, J., Li, F.L., 2010. A hierarchical shape tree for shape classification, in: 25th International Conference of Image and Vision Computing New Zealand, pp. 1–6.

Lim, K.L., Galoogahi, H.K., 2010. Shape classification using local and global features, in: Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on, pp. 115–120.

Ling, H., Jacobs, D.W., 2007. Shape classification using the inner-distance. IEEE Trans. Pattern Anal. Mach. Intell. 29, 286–299.

Mahmoudi, M., Sapiro, G., 2008. Three-dimensional point cloud recognition via distributions of geometric distances, in: Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on, pp. 1–8.

Mestetskiy, L., 1999. Skeletonization of polygonal figures based on the generalized delaunay triangulation. Programming and Computer Software 25, 131–142.

Mestetskiy, L., 2013. Binary Image Skeleton Representation by Compound Bezier Curves. iConcept Press.

Mestetskiy, L., 2015. Medial width of polygonal and circular figures - approach via line segment voronoi diagram, in: Proceedings of the 10th International Conference on Computer Vision Theory and Applications (VISIGRAPP 2015), pp. 379–386. doi:10.5220/0005261903790386.

Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D., 2001. Matching 3d models

with shape distributions, in: Proceedings of the International Conference on Shape Modeling & Applications, IEEE Computer Society. pp. 154–166.

Ozay, M., Aktas, U.R., Wyatt, J.L., Leonardis, A., 2015. Compositional hierarchical representation of shape manifolds for classification of non-manifold shapes, in: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), IEEE Computer Society. pp. 1662–1670.

Sebastian, T.B., Klein, P.N., Kimia, B.B., 2004. Recognition of shapes by editing their shock graphs. IEEE Trans. Pattern Anal. Mach. Intell. 26, 550–571.

Shen, W., Jiang, Y., Gao, W., Zeng, D., Wang, X., 2016. Shape recognition by bag of skeleton-associated contour parts. Pattern Recognition Letters 83, 321–329.

Shen, W., Wang, X., Yao, C., Bai, X., 2014. Shape recognition by combining contour and skeleton into a mid-level representation, in: Pattern Recognition - 6th Chinese Conference, CCPR 2014, Changsha, China, November 17-19, 2014. Proceedings, Part I, pp. 391–400.

Siddiqi, K., Pizer, S., 2008. Medial Representations: Mathematics, Algorithms and Applications. 1st ed., Springer Publishing Company, Incorporated.

Soderkvist, O.J.O., 2001. Computer vision classification of leaves from swedish trees. Master's thesis, Linkoping University, SE-581 83 Linkoping, Sweden.

Sun, K.B., Super, B.J., 2005. Classification of contour shapes using class segment sets, in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), IEEE Computer Society. pp. 727–733.

Wang, J., Yang, J., Yu, K., Lv, F., Huang, T., Gong, Y., 2010. Locality-constrained linear coding for image classification, in: In: IEEE Conference on Computer Vision and Pattern Recognition, p. 33603367.

Wang, X., Feng, B., Bai, X., Liu, W., Jan Latecki, L., 2014. Bag of contour fragments for robust shape classification. Pattern Recogn. 47, 2116–2125.