New Shape Descriptors based on Tensor Scale with Global Features

Anderson M. Freitas, Paulo A. V. Miranda
Institute of Mathematics and Statistics,
University of São Paulo,
São Paulo, Brazil.
afreitas@ime.usp.br, pmiranda@vision.ime.usp.br

***Abstract*—In this work, two new shape descriptors are proposed for tasks in Content-Based Image Retrieval (CBIR) and Shape Analysis, which are built upon an extended tensor scale based on the Euclidean Distance Transform (EDT). First, the tensor scale algorithm is applied to extract shape attributes from its local structures as represented by the largest ellipse within a homogeneous region centered at each image pixel. In the new descriptors, the upper limit of the interval of local orientation of tensor scale ellipses is extended from π to 2π, to discriminate the description of local structures better. Then, the new descriptors are built based on different sampling approaches, aiming to summarize the most relevant features. Experimental results for different shape datasets (MPEG-7 and MNIST) are presented to illustrate and validate the methods. TSS can achieve high retrieval values comparable to state-of-the-art methods, which usually rely on time-consuming correspondence optimization algorithms, but uses a more straightforward and faster distance function, while the even faster linear complexity of TSB leads to a suitable solution for huge shape collections.**

I. Introduction

Content-based image retrieval (CBIR) concerns the problem of searching for digital images in large databases, which are similar to a query image. A particular type of CBIR system exploits the shape information as image descriptors. A shape descriptor should be simple, compact, insensitive to noise, affine-invariant, and at the same time contain all relevant information to distinguish different images [1]. Preferably, the matching algorithm used by a shape descriptor should also be fast to be suitable on large datasets, since the feature extraction can often be performed offline. Many proposed shape descriptors yield a high accuracy score but sacrifice performance with a high computational time, mainly relying on Dynamic Programming (DP) for the matching algorithm to establish correspondences.

In this work[1], two new shape descriptors are presented designed to achieve high accuracy with fast distance functions. Besides not using DP, the matching algorithms are also simpler than OCS (Optimal Correspondent Subsequence) used by BAS [2], OSB (Optimal Subsequence Bijection) and the Hungarian algorithm used in [3], and also more efficient than the one used in TSDIZ [4]. Comparing with especially proposed methods for dealing with large datasets, such as Hough Transform Statistics (HTS) and HTS-neighborhood (HTSn) [5], our matching algorithms have a lower computational complexity, for non-aligned shapes, and higher retrieval rates.

The methods presented here are based on the tensor scale concept (TS) [6] — a morphometric parameter yielding a simultaneous representation of local structure orientation, thickness, and anisotropy. The algorithm to compute tensor scale, as originally proposed [6], is computationally expensive. To address this problem, Andaló *et al*. proposed a simpler and yet effective implementation of the original method [4].

One contribution of this work was the revision of the algorithmic TS computation, as proposed by Andaló *et al*., extending the ellipse's orientation to 360°. Based on this richer TS model, two novel shape descriptors were proposed, with greater discrimination power in relation to previous TS-based works, for shape-based image retrieval: Tensor Scale Sector (TSS) and Tensor Scale Band (TSB) descriptors.

In relation to previous TS descriptors (TSD [7] and TSDIZ [4]), the proposed descriptors TSS and TSB are more accurate and have faster matching algorithms. TSS incorporates spatial information by the use of circular sectors and TSB by the use of concentric bands around a central point, which are much more discriminative than the simple normalized orientation histogram used by TSD. While TSDIZ is a contour-based method, TSS and TSB are region-based methods, opening new perspectives for novel applications. The features extracted from the ellipses are also more sophisticated, considering 360 degrees.

This paper is organized as follows: Section II presents the tensor scale previous relevant work, including its EDT-based implementation and how to extend it to ellipses with 360°, as used in this work. Then, our novel tensor scale descriptors, TSS and TSB, are shown in Sections III and IV. The combined descriptor approach is described in Section V, respectively, and the analysis of their computational complexity is presented in Section VI. The experimental evaluation is conducted in Section VII. Section VIII states the conclusions and derived work.

II. Background

Saha *et al.* have introduced a local scale method called tensor scale (TS) [6], which is the parametric representation of the largest ellipse (or ellipsoid in 3D), centered at a point p within the same homogeneous region under a predefined

[1]This work relates to a MSc dissertation, defended on October 24th, 2017.

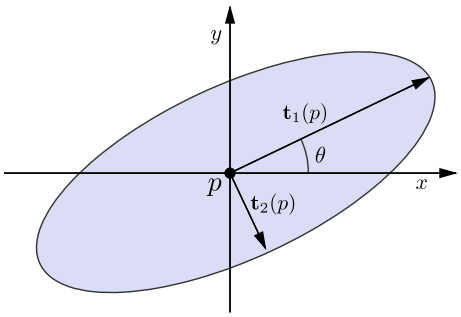


Fig. 1. Tensor scale representation at point p.

criterion (usually intensity). The tensor scale model (Fig. 1) provides three factors: orientation θ of the major semi-axis $\boldsymbol{t_1}$, anisotropy ($\sqrt{1 - \|\boldsymbol{t_2}(p)\|^2 / \|\boldsymbol{t_1}(p)\|^2}$) and thickness ($\|\boldsymbol{t_2}(p)\|$).

For a given pixel, the largest ellipse within the same homogeneous region is determined by tracing sample lines from 0 to 180 degrees around that pixel and computing the following steps: (i) Intensity computation along each sample line; (ii) Location of two optimum edge points on each sample line; (iii) Repositioning of the edge locations to points equidistant to the given pixel, following the axial symmetry of an ellipse; (iv) Computation of the best-fit ellipse to the repositioned edge locations using *Principal Component Analysis* (PCA). These steps are performed for each image pixel until all ellipses have been computed [6], [7].

In the case of binary images (or images with values taken from a fixed set of labels), Andaló *et al*. proposed a very fast TS implementation [4], which exploits the circles given by the *Euclidean Distance Transform* (EDT). The EDT is computed in linear time using the *Image Foresting Transform* (IFT) [8], a graph-based approach to the design of image processing operators based on connectivity.

A. EDT-based Tensor Scale

For each pixel, to locate the edge points on each sample line around it (Figure 2a), we first compute the *Euclidean Distance Transform* using the IFT framework. We insert in a priority queue all pixels that have a neighbour with different label as seeds with unitary costs. Then running the IFT with a proper path cost function leads to a cost map (Figure 2b) that has the Euclidean distance of every pixel to its nearest pixels with a distinct label [8].

To find the edge locations over the sample lines for a given central pixel, we exploit the EDT values to speed up the process using a sequence of jumps, by skipping all pixels in the radius given by the EDT, instead of performing a naïve pixel-by-pixel traversal (Figures 2c-e).

The computation of the best-fit ellipse is accomplished in two sub-steps: (i) Determination of ellipse orientation θ; and (ii) Computation of the lengths of the semi-axes.

In our approach we adopted the EDT value as being the thickness $b = \|\boldsymbol{t_2}(p)\|$. Therefore $a = \|\boldsymbol{t_1}(p)\|$ is the only variable left and this leads to simpler equations, with the final ellipse being more consistently contained in the same homogeneous region. In this sense:

$$a = \sqrt{\frac{C \cdot b^2}{b^2 \cdot A - B}} \tag{1}$$

where $A = \sum u_i^2$, $B = \sum u_i^2 v_i^2$, $C = \sum u_i^4$ and (u_i, v_i) are the relative coordinates of the repositioned edge points with respect to the central pixel and after rotation by angle $-\theta$, such that the ellipse's major semi-axis becomes aligned to the horizontal axis. To guarantee valid a values, points with $v_i^2 \geq b^2$ should not be considered during the computation of the above equations since they correspond to pixels out of the zone of possible boundary points of ellipses with the given thickness b. To make sure that at least one edge point lies on the valid zone, we trace an additional sample line on the direction given by the angle θ.

For the purposes of shape descriptor, θ can be assumed simply to be the orthogonal direction of the line that connects the central pixel to its nearest pixel with a distinct label given by the EDT. In our experiments, we adopted this simple, but yet effective variation, and considered a single sample line on the direction of θ to estimate a.

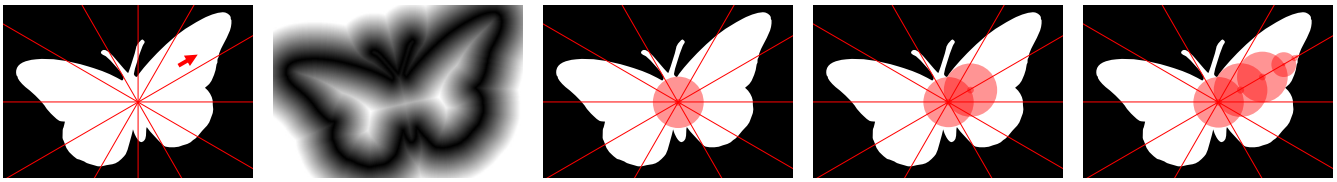
B. Extending TS to 360 degrees

In order to extend TS to 360°, we further distinguish the ellipses, by comparing the label value $L(p)$ at its central pixel p with the label $L(r)$ at its closest pixel r having a distinct label (Figure 3). Consider the vector $\vec{\boldsymbol{v}}_{rp}$. If $L(p) > L(r)$, then the ellipse orientation is taken as the angle formed by the vector $\vec{\boldsymbol{v}}_{rp}$ rotated 90° clockwise (Figure 3a), otherwise it is rotated 90° counterclockwise (Figure 3b). Figure 4 shows TS results coded in the HSV color space, where the hue corresponds to the orientation. TS by Andaló *et al*. [4] considered 180° only, such that opposite directions receive the same color/orientation (Figure 4a). By using the procedure depicted in Figure 3, we can extend the results to 360°, such that opposite directions are assigned to different colors (Figure 4b).

III. TSS: Tensor Scale Sector

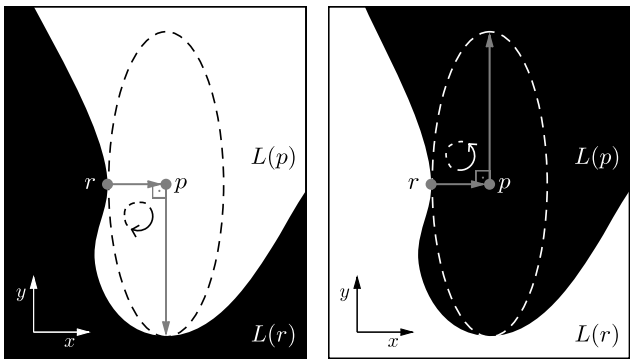
A single circular region around the object is considered, which is divided into sectors within concentric bands (Figure 5a). After that, tensor scale information is computed only inside this circular region, but the orientations of ellipses centered at pixels closer to the external circular border follow its round shape, and therefore do not present relevant information of the shape being analyzed. So, in this work these ellipses are disregarded (Figure 4). Figure 5a shows an example grid using four concentric bands and twenty sectors per band. The radius of the bands are uniformly sampled at regular intervals.

Ellipses attributes within sectors are used as features to compose a fixed-length feature vector, considering the orientation θ of each ellipse converted to an angle γ measured relatively to its sector, as shown on Figure 6. The usage of relative angles γ enables the rotational invariance of TSB and simplifies the



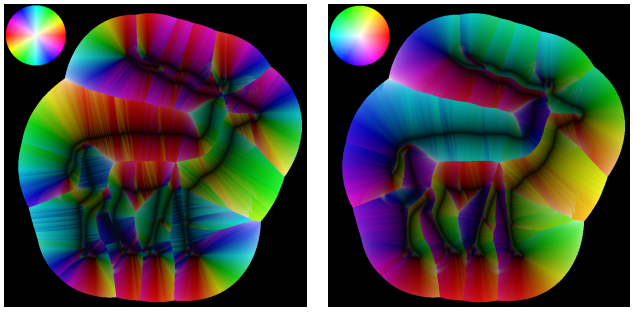
(a) (b) (c) (d) (e)

Fig. 2. Sample line traversal example in the direction pointed by the arrow in (a). (b) The auxiliary values of the EDT. (c) The first traversal step. (d) The second step. (e) The edge location is found after only six steps.



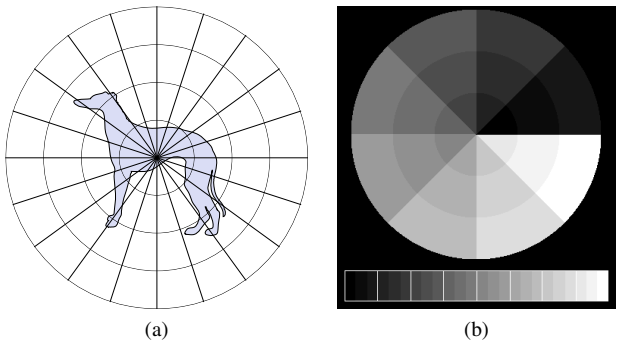
(a) $\theta = 270°$ $(L(p) > L(r))$ (b) $\theta = 90°$ $(L(p) \leq L(r))$

Fig. 3. TS with 360 degrees is obtained by comparing the label value $L(p)$ at its central pixel p with the label $L(r)$ of its closest pixel r having a distinct label.



(a) (b)

Fig. 4. (a) TS by Andaló *et al.* coded in the HSV colorspace, where the hue (H), saturation (S), and value (V) indicate the orientation (0°-180°), anisotropy, and thickness values, respectively. (b) TS extended to 360 degrees.



(a) (b)

Fig. 5. (a) Feature vector extraction of TSS. (b) The mapping in the simpler case of 3 bands ($B = 3$) with 8 sectors per band ($S = 8$). Each ellipse falls inside a radial band, which has an index $b \in [0, B-1]$, and has a sector index $s \in [0, S-1]$ in that band. Its position j in the feature vector is given by $j = b + s \times B$.

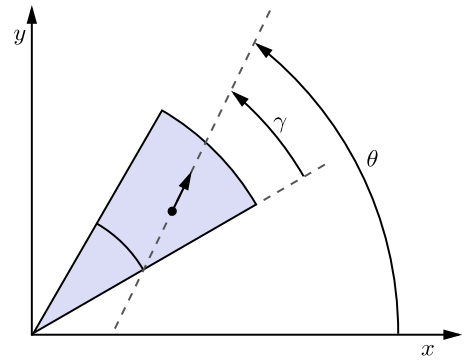


Fig. 6. For each ellipse, TSS and TSB consider angles γ measured relative to its sector, instead of working with θ.

distance function computation of TSS, as will be explained later.

To summarize the relative angles γ_i within each sector, each angle is treated as a vector $(v_x, v_y) = (w_i \cos(\gamma_i), w_i \sin(\gamma_i))$ such that the magnitude information is exploited as well. Then, a weight w_i is assigned to each ellipse which is given by its squared anisotropy value, to penalize ellipses with low anisotropy as they tend to be circles and therefore do not present a well-defined orientation. The local dominant relative orientation $(\bar{v}_x, \bar{v}_y)$ within these sectors is computed by Equations 2 and 3, where γ_i and w_i are the relative orientations and squared anisotropy values of all the n ellipses inside the sector being considered.

$$\bar{v}_x = \frac{1}{n} \sum_{i=1}^{n} w_i \cos(\gamma_i) \quad (2)$$

$$\bar{v}_y = \frac{1}{n} \sum_{i=1}^{n} w_i \sin(\gamma_i) \quad (3)$$

The usage of ellipses with 360° results in more complex orientation patterns (Figure 4), such as the discontinuities at pixels near the shape's skeleton. In order to bring in information about relative angles with opposite orientations, we also extract the sum of the absolute values of the responses

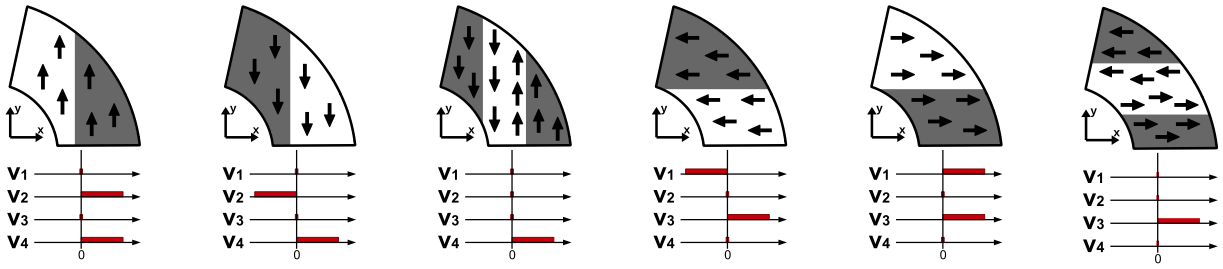


Fig. 7. (a-f) The features $\boldsymbol{v} = \{v_1, v_2, v_3, v_4\}$ of a sector effectively capture different underlying orientation patterns. (a-b) Vertical ellipses with the same orientation. (c) Vertical ellipses with opposite angles. (d-e) Horizontal ellipses with the same orientation. (f) Horizontal ellipses with opposite angles.

(Equations 4 and 5). Therefore, we will have for each sector a vector of four features $\boldsymbol{v} = \{v_1, v_2, v_3, v_4\}$, where $v_1 = \bar{v}_x$ and $v_2 = \bar{v}_y$ (Figure 7). This combination of ellipses with 360° and the features $\boldsymbol{v}$, that were inspired on a four-dimensional vector used in the SURF point detector and descriptor [9], leads to an improved retrieval performance.

$$v_3 = \frac{1}{n}\sum_{i=1}^{n} |w_i \cos(\gamma_i)| \quad (4)$$

$$v_4 = \frac{1}{n}\sum_{i=1}^{n} |w_i \sin(\gamma_i)| \quad (5)$$

Finally, the features of sectors on the same radial zones are consecutively grouped to compose the TSS feature vector (Figure 5b). To properly deal with flipped/mirrored images we should also consider a mirrored feature vector computed by considering inverted signals on the x components of the angles in Equations 2, 3, 4 and 5.

There are two parameters to be set, the central point coordinates and the radius of the circular region. These parameters can be tuned for different applications. In our experiments, the center of gravity (shape centroid) was used as the central point, and the radius was taken as three times the square root of the mean squared Euclidean distance between the shape pixels to the centroid.

The distance between two feature vectors is computed as a circular matching, required for the rotational invariance, given by the Algorithm 1, being the distance between corresponding sectors computed by the ℓ_2 distance (Lines 12–13), which empirically demonstrated superior results.

IV. TSB: Tensor Scale Band descriptor

The idea of TSB is to sacrifice the angular displacement of the sectors to obtain a faster distance function. In contrast to TSS, the spatial information in TSB is incorporated via the radial displacement only, by taking one normalized histogram with 60 bins per radial band, capturing the relative angular distribution of γ. Due to the high coverage area of the radial bands, the angular distributions in the form of a histogram is more appropriate than considering only the four-dimensional feature vector $\boldsymbol{v}$. The histograms are concatenated to compose a fixed-length feature vector, which is naturally rotational invariant because of the usage of relative angles γ.

To make the method less sensitive to histogram bins with very high values the square root of the normalized histogram values is used, so the difference between the largest and the smallest histogram bins becomes smaller than in the original scale. For the same reasons, we empirically observed that the ℓ_1 distance function resulted in superior performance than ℓ_2.

The distance function of TSB requires only the ℓ_1 distance computation between two feature vectors.

Algorithm 1. – TSS Distance Function

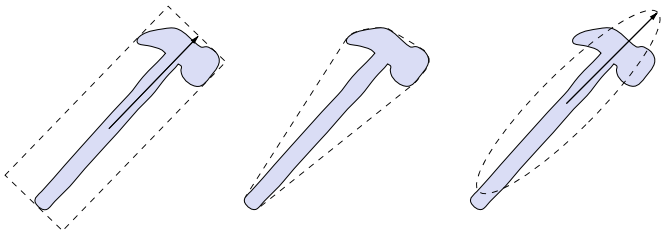
INPUT: Feature vectors fv_1 and fv_2
OUTPUT: Distance

1. $S \leftarrow$ *number of sectors per band*
2. $B \leftarrow$ *number of radial bands*
3. $N \leftarrow S \cdot B$
4. $d_{min} \leftarrow \infty$
5. ***for*** $k \leftarrow 0$, ***to*** $S-1$, ***do***
6. $\quad$ *offset* $\leftarrow B \cdot k$
7. $\quad d \leftarrow 0.0$
8. $\quad$ ***for*** $i \leftarrow 0$, ***to*** $N-1$, ***do***
9. $\quad\quad j \leftarrow (i + \textit{offset}) \ \% \ N$
10. $\quad\quad (a_1, a_2, a_3, a_4) \leftarrow fv_1[i]$
11. $\quad\quad (b_1, b_2, b_3, b_4) \leftarrow fv_2[j]$
12. $\quad\quad (d_1, d_2, d_3, d_4) \leftarrow (b_1 - a_1, b_2 - a_2, b_3 - a_3, b_4 - a_4)$
13. $\quad\quad d \leftarrow d + \sqrt{(d_1)^2 + (d_2)^2 + (d_3)^2 + (d_4)^2}$
14. $\quad\quad i \leftarrow i + 1$
15. $\quad$ ***if*** $d < d_{min}$ ***then*** $d_{min} \leftarrow d$
16. $\quad k \leftarrow k + 1$
17. ***return*** d_{min}/N

V. Combined Descriptor with Global Features

Considering that both TSS and TSB characterize a shape by summarizing only local TS features from delimited regions (sectors in TSS and bands in TSB), combining the descriptors with global geometric features can increase the original discrimination power. This strategy was also exploited in other works [10]–[12] to increase or decrease the distance between shapes based on their visually similarity.

In this work we used the following geometric global shape features to build the composite descriptor: aspect-ratio of the aligned bounding-box (Figure 8a), eccentricity, and solidity. The solidity is defined by the ratio between the area of the



(a) (b) (c)

Fig. 8. (a) Bounding-box aligned to the shape orientation. (b) Convex-hull. (c) Ellipse with the same second-order statistical moments of the shape.

shape and its convex-hull (Figure 8b), which in turn is defined by the smallest convex polygon that contains all the points belonging to the object. Eccentricity is defined by the ratio between the axes of the ellipse that has the same second-order statistical moments of the shape (Figure 8c), centered on the object's center of mass. It is worth mentioning that these global features are also invariant to rotation, scaling and mirroring operations.

The distance function δ_{GF} of the combined descriptor is obtained by the convex combination of the global geometric features with the original descriptor's distance (TSS or TSB), using the log-sum-exp [13] operation, given by:

$$\delta_{\text{GF}}(A,B) = \delta_D(A,B) + \log \sum_{k=1}^{m} \exp\left(\alpha_k \cdot \left(\varphi_k^A - \varphi_k^B\right)\right)$$

where δ_D is the distance function to be extended; $\varphi_k^S = (\varphi_1, \ldots, \varphi_m)$ a set of m geometric features from shape S; and α_k are the coefficients of the convex combination, with $\alpha_k > 0$ and $\sum_{k=1}^{m} \alpha_k = 1$.

VI. Computational Complexity

In the context of CBIR, the computational cost consists of two parts: (i) computing the feature vector; (ii) performing the shape dissimilarity by a distance function. On large collections, the latter is more important as the distance should be determined for every shape in the collection against the query shape, and the descriptor of all shapes are already stored and calculated beforehand. To perform a new search, only the descriptor of the query shape must be calculated, so for a fast retrieval time an efficient distance function is crucial.

The distance function of TSS presented in the Algorithm 1 requires the computation of the dissimilarity between each corresponding sector from both feature vectors. The feature vector size of TSS is proportional to the total number of sectors $N = S \cdot B$, where S is the number of sectors per band and B is the number of bands. The overall complexity for calculating the distance between two feature vectors is $O(S \cdot N)$. When the mirroring option is used, it must perform a second iteration over a reversed feature vector, and in this situation the matching is $O(2 \cdot S \cdot N) = O(S \cdot N)$.

The feature vector size of TSB is $N = 60 \cdot B$, where B is the number of bands, since we have one histogram with 60 bins per radial band. The distance function of TSB requires only the ℓ_1 distance computation between two feature vectors and its complexity is $O(N)$, which becomes $O(60 \cdot B) = O(B)$, if we consider the number of bins as a constant. It also allows the usage of fast indexing structures. Both distance functions are rotation-invariant and works with non-aligned shapes.

Table I compares the computational complexity of the evaluated descriptors used in Section VII.

VII. Experimental Results

The TSS and TSB descriptors were compared against commonly used shape descriptors. We evaluated the proposed descriptors using the MPEG-7 CE-Shape-1 (part B) dataset [1] and the MNIST dataset from [14]. The MPEG-7 dataset contains 1,400 shape images distributed along 70 classes, where each class contains 20 shapes with various rigid and non-rigid transformations, noise and change of viewpoint. The MNIST database consists of 10,000 images of handwritten digits and is commonly used in classification evaluation tasks.

In the experiments, we evaluated TSS with 4 bands and 20 sectors per band, and discarded the outermost band, since its ellipses essentially follow the round shape of the external circular region without adding relevant information. We also used a single mean feature for the inner band, instead of taking its sectors separately, since they are too small, resulting in a total of 40 sectors for the two remaining central bands plus one mean feature for the internal band (TSS-41). So, the number 41 of TSS-41 stands for $B + B + 1$, where B denotes the number of sectors per band ($B = 20$).

Since the shapes in the MNIST dataset are almost already aligned, we only considered the search from $-\pi/4$ to $\pi/4$.

In the case of TSB, we used 5 bands and 32 sectors per band, but we discarded the outermost band, since its orientation is mainly affected by the circular region around the object, resulting in a total of $60 \times 4 = 240$ bins for the 4 remaining concatenated histograms. Also, for the combined variations TSS+GF and TSB+GF, the parameter α were found by cross-validation.

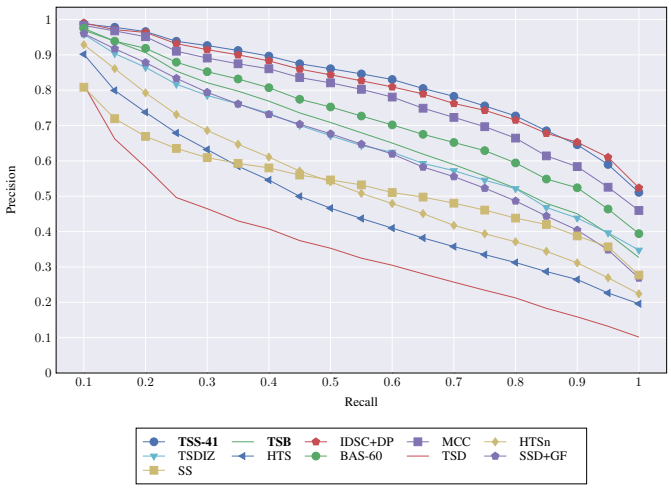


Fig. 9. Precision versus Recall curves (MPEG-7 CE-1 Part B dataset).

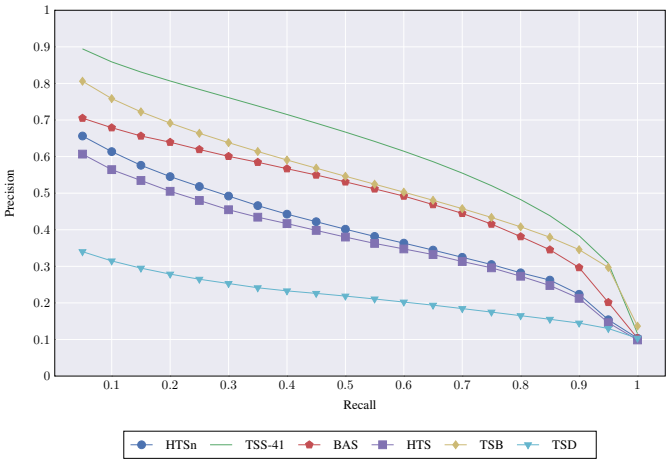


Fig. 10. Precision versus Recall curves (MNIST dataset).

TABLE I
BEP SCORES OF DIFFERENT DESCRIPTORS ON THE MPEG-7 DATASET AND ITS COMPUTATIONAL COMPLEXITY OF THE DISTANCE FUNCTION.

Descriptor	**BEP (%)**	**Complexity**	**Type**
Height Functions [15]	89.66	$O(N^3)$	Contour
CNSS [16]	89.47	$O(N)$	Contour
IDSC+DP [17]	88.11	$O(KN^2)$	Contour
TSS-41+GF	**87.22**	$\boldsymbol{O(SN)}$	Region
TSS-41	**86.20**	$\boldsymbol{O(SN)}$	Region
MCC [10]	84.93	$O(N^3)$	Contour
BAS-60 [2]	76.78	$O(N^3)$	Contour
SC+DP (Shape Contexts) [17]	76.51	$O(KN^2)$	Contour
TSB+GF	**75.20**	$\boldsymbol{O(N)}$	Region
Path Similarity [3]	75.16	$O(N^3)$	Skeleton
TSB	**74.01**	$\boldsymbol{O(N)}$	Region
SSD+GF [11]	71.00	$O(N^2)$	Contour
TSDIZ [4]	69.44	$O(CN^2)$	Contour
HTSn [5]	59.97	$O(N^2)$	Contour
HTS [5]	56.01	$O(N^2)$	Contour
TSD [7]	46.08	$O(N^2)$	Region

A. Evaluation measures

Precision vs. Recall ($P \times R$) curves are the commonest evaluation measure used in CBIR domain. Precision is defined as the fraction of retrieved images which are relevant to a query. In contrast, recall measures the fraction of the relevant images which has been retrieved. In general, the curve $P \times R$ closest to the top of the chart indicates the best performance. Figures 9 and 10 shows the $P \times R$ curves obtained for all evaluated descriptors on both datasets.

We also considered the bullseye precision (BEP) [1] for the evaluation of the MPEG-7 collection, which is defined as the count of true positives within the 40 top similar results for every shape as a query image. Then the score is normalized by the highest possible number of correct hits (which is 20×1400). The results are presented in Table I.

VIII. CONCLUSIONS

We presented two novel shape descriptors for CBIR, which are non-limited to a particular topology and that may be easily tailored for different applications. Their features extracted from circular sectors could be used to build a shape vocabulary according to the bag-of-words (BoW) paradigm to get a learning-based method, and also be combined with other features. The results demonstrate that TSS performs better than TSB, since it includes more spatial information, by exploiting each individual sector instead of only the bands as TSB.

This work resulted in the publication of the article *"TSS & TSB: Tensor Scale Descriptors within Circular Sectors for Fast Shape Retrieval"* [18], published on a special edition of *Pattern Recognition Letters* on *"Efficient Shape Representation, Matching, Ranking, and its Applications"*.

REFERENCES

[1] H. Kim and J. Kim, "Region-based shape descriptor invariant to rotation, scale and translation," *Signal Processing: Image Communication*, vol. 16, no. 1-2, pp. 87–93, 2000.

[2] N. Arica and F. T. Y. Vural, "BAS: A Perceptual Shape Descriptor based on the Beam Angle Statistics," *Pattern Recognition*, vol. 24, no. 9-10, pp. 1627–1639, 2003.

[3] X. Bai and L. J. Latecki, "Path Similarity Skeleton Graph Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 7, pp. 1282–1292, 2008.

[4] F. A. Andaló, P. A. V. Miranda, R. S. Torres, and A. X. Falcão, "Shape feature extraction and description based on tensor scale," *Pattern Recognition*, vol. 43, no. 1, pp. 26–36, Jan. 2010.

[5] G. B. Souza and A. N. Marana, "HTS and HTSn: New shape descriptors based on Hough Transform Statistics," *Computer Vision and Image Understanding*, vol. 127, pp. 43–56, 2014.

[6] P. K. Saha, "Tensor scale: A local morphometric parameter with applications to computer vision and image processing," *Computer Vision and Image Understanding*, vol. 99, no. 3, pp. 384–413, sep 2005.

[7] P. A. V. Miranda, R. S. Torres, and A. X. Falcão, "TSD: A Shape Descriptor Based on a Distribution of Tensor Scale Local Orientation," *XVIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)*, pp. 139–146, 2005.

[8] A. Falcão, J. Stolfi, and R. de Alencar Lotufo, "The Image Foresting Transform: Theory, Algorithms, and Applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 19–29, jan 2004.

[9] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, jun 2008.

[10] T. Adamek and N. E. O'Connor, "A multiscale representation method for nonrigid shapes with a single closed contour," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 5, pp. 742–753, May 2004.

[11] G. Pedrosa, C. Barcelos, and M. Batista, "Image feature descriptor based on shape salience points," *Neurocomputing*, vol. 120, pp. 156–163, 2013.

[12] G. V. Pedrosa, A. J. M. Traina, and C. A. Z. Barcelos, "Retrieving 2D shapes by similarity based on bag of salience points," *Multimedia Tools and Applications*, pp. 1–15, 2016.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[14] Y. LeCun and C. Cortes, "The MNIST database of handwritten digits," 1998. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[15] J. Wang, X. Bai, X. You, W. Liu, and L. J. Latecki, "Shape matching and classification using height functions," *Pattern Recognition Letters*, vol. 33, no. 2, pp. 134–143, 2012.

[16] A. B. Oliveira, P. R. Silva, and D. A. C. Barone, "A novel 2D shape signature method based on complex network spectrum," *Pattern Recognition Letters*, vol. 63, pp. 43–49, 2015.

[17] H. Ling and D. Jacobs, "Shape Classification Using the Inner-Distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 286–299, Feb 2007.

[18] A. M. Freitas, R. da S. Torres, and P. A. Miranda, "TSS & TSB: Tensor scale descriptors within circular sectors for fast shape retrieval," *Pattern Recognition Letters*, vol. 83, Part 3, pp. 303–311, 2016, special issue on "Efficient Shape Representation, Matching, Ranking, and its Applications".