Shape Descriptor Based on the Volume of Transformed Image Boundary

Xavier Descombes[1] and Sergey Komech[2]

[1] EPI Ariana, INRIA SAM, Sophia Antipolis, FR
`xavier.descombes@inria.fr`
[2] Dobrushin Lab., IITP, Moscow, RU
`komech@iitp.ru`

Abstract. In this paper, we derive new shape descriptors based on a directional characterization. The main idea is to study the behavior of the shape neighborhood under family of transformations. We obtain a description invariant with respect to rotation, reflection, translation and scaling. We consider family of volume-preserving transformations. Our descriptor is based on the volume of the neighbourhood of transformed image. A well-defined metric is then proposed on the associated feature space. We show the continuity of this metric. Some results on shape retrieval are provided on Kimia 216 and part of MPEG-7 CE-Shape-1 databases to show the accuracy of the proposed shape metric.

1 Introduction

Shape characterization is becoming a crucial challenge in image analysis. The increasing resolution of new sensors, satellite images or scanners provides information on the object geometry which can be interpreted by shape analysis. The size of data basis also requires some efficient tools for analyzing shapes, for example in applications such as image retrieval. Reviews of proposed representations can be found in [4,5]. One class of methods consists in defining shapes descriptors based on shape signatures histogram signatures, shape invariant moments, contrast, matrices or spectral features. A shape representation is evaluated with respect to its robustness, w.r.t. noise and/or intra-class variability, compaction of the description, its invariance properties and its efficiency in terms of computation time. According to Zhang and Lu [5], the different approaches can be classified into contour-based and region-based methods, and within each class between structural and global approaches. In this paper we consider shapes as binary silhouette of objects and concentrate on global approaches. Simple global shape descriptors embed area, orientation, convexity, bending energy [6,7]. Usually, these descriptors are not sufficiently sensitive to details to provide good scores in image retrieval. Distances between shapes or surfaces have been proposed, such as the Hausdorf distance or some modification to reduce sensitivity to outlier [8,9]. In this setting, the invariance properties can be obtained by taking the minimum distance over the corresponding group of transformation. A key issue is to consider a metric for which the minimum is computed with a low computational complexity.

In this paper, we derive a 2D signature of shapes and propose a metric on the associated feature space. The first idea consists of a description of the boundary regularity by comparing the volume of the boundary neighborhood with the shape volume. The second idea is to study the behavior of this descriptor under shape transformations. We thus define a family of diffeomorphisms consisting in expanding the shape in one direction and contracting it in the orthogonal direction. In that way, for a given detail, there exists at least such a transformation enlarging its contribution to the descriptor and another one reducing it. We then derive a well defined metric on the feature space, and show its performance for shape discrimination on databases of various size.

We describe the proposed shape space and define a metric on it in Section 2. A discretization of the metric is described and evaluated on two different databases in Section 3. Finally, conclusion and perspectives are drawn in Section 4.

2   A Topological Description of Shapes

2.1   Shape Space

We consider shapes as 2D silhouettes of bounded objects in the image plane.

Definition 1. *The pre-shape space S is the set of subsets of $\mathbb{R}^2$ satisfying the following conditions:*

1. *$\forall a \in S$, a is compact and connected, with a strictly positive area,*
2. *$\forall a \in S$, $\mathbb{R}^2 \setminus a$ is connected (a has no hole).*

Let us consider a shape $a \in S$. Define the closed ε-neighborhood of the set a in the sense of the Euclidean metric as $O^\varepsilon(a) = \{x \in \mathbb{R}^2 : e(x, a) \leq \varepsilon\}, \varepsilon \geq 0$, where $e(\cdot, \cdot)$ is the Euclidean distance.

On this pre-shape space, we consider the Hausdorff metric (which is well-defined, see, for example, [2]) for the sets in $\mathbb{R}^2$:

$$\rho(a, b) = \inf\{\delta > 0 : a \subset O^\delta(b), b \subset O^\delta(a)\},$$

where $a, b \subset S$.

A shape space should embed some invariance properties. Let G be the group of transformations of $\mathbb{R}^2$ generated by rotations, translations, reflections and scaling : $G = SL_2^{\pm}(\mathbb{R}) \times \mathbb{R}_+$. To define a shape space $\$$ isometry- and scale-invariant, we consider:

$$\$ = S/G \ . \tag{1}$$

For a given $A \in \$$, we note $r(A) = \{a \in S : vol(a) = 1, G(a) = A\}$, where $vol(\cdot)$ is the area of the set.

Therefore, on the shape space $\$$, the Hausdorff metric becomes:

$$d(A, B) = \inf\{\rho(a, b) \mid a \in r(A), b \in r(B)\}, \tag{2}$$

where $A, B \in \$$ (note that this metric can be compared with the Procrustes distance for sets consisting of finite number of points [3,1]). The proof of the following proposition is not very hard and we will omit it due to the space limit.

Proposition 1. $d(\cdot, \cdot)$ *is a well-defined metric on* $\$$.

2.2   Volume Descriptor and Family of Transformations

The main idea of the proposed description is to characterize the behavior of shapes under some transformations. These transformations aim at enlighting small characteristic details. We first consider the volume behavior under some dilation. Intuitively, this volume will increase more for sinuous shape boundaries than for smooth shapes. Let us consider a shape $a \in S$. Idea of our shape descriptor is based on analyzing the fraction

$$P^\varepsilon(a) = \frac{vol(O^\varepsilon(a) \setminus a)}{vol(a)} \quad , \qquad (3)$$

where $vol(\cdot)$ is the area of the set. This parameter is well-defined $(vol(a) \neq 0)$.

We consider a family $\{F_{(\theta,\beta)}\}$ of linear transformations of $\mathbb{R}^2$ in order to obtain more significant information about shapes. The goal of such transformations is to emphasize "features" of the shape in specific direction. These transformations are defined as follows:

$$F_{(\theta,\beta)} := \begin{pmatrix} \beta \cos^2 \theta + \frac{1}{\beta} \sin^2 \theta & (\beta - \frac{1}{\beta}) \sin \theta \cos \theta \\ (\beta - \frac{1}{\beta}) \sin \theta \cos \theta & \beta \sin^2 \theta + \frac{1}{\beta} \cos^2 \theta \end{pmatrix}, \qquad (4)$$

where $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}], \beta \geq 1$. Every $F_{(\theta,\beta)}$ is β-times expanding in one direction (with angle θ) and β-times contracting in orthogonal, so it is a volume-preserving transformation.

For every set $a \subset \mathbb{R}^2$, we obtain the map

$$a \mapsto P^\varepsilon_{(\theta,\beta)}(a) := \frac{vol(O^\varepsilon(F_{(\theta,\beta)}a) \setminus F_{(\theta,\beta)}a)}{vol(a)}, \quad \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}], \ \beta \geq 1, \ \varepsilon > 0 \ . \ (5)$$

It is clear that $P^\varepsilon_{(\theta,\beta)}(a)$ is a continuous function of ε, θ and β and $P^\varepsilon_{(-\frac{\pi}{2},\beta)}(a) = P^\varepsilon_{(\frac{\pi}{2},\beta)}(a)$. Note, that this function is a constant, if a is a ball $(\varepsilon, \beta$ are fixed).

Let R_γ be the rotation by an angle γ. We have the following property:

Property 1. $\forall \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}], P^\varepsilon_{(\theta,\beta)}(R_\gamma a) = P^\varepsilon_{((\theta+\frac{\pi}{2}+\gamma)mod(\pi)-\frac{\pi}{2},\beta)}(a)$.

We consider R_x, the reflection with respect to x axis (horizontal line). We have the following property:

Property 2. $\forall \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}], P^\varepsilon_{(\theta,\beta)}(R_x a) = P^\varepsilon_{(-\theta,\beta)}(a)$.

Let us denote R the group of transformations generated by Properties 1 and 2. The shape representation space $\mathcal{R}$ we consider for the fixed $\varepsilon > 0$ is then defined by the following mapping:

$$\Phi : \$ \to \mathcal{R} = C^0([-\frac{\pi}{2}, \frac{\pi}{2}] \times [1, \infty])/R,$$
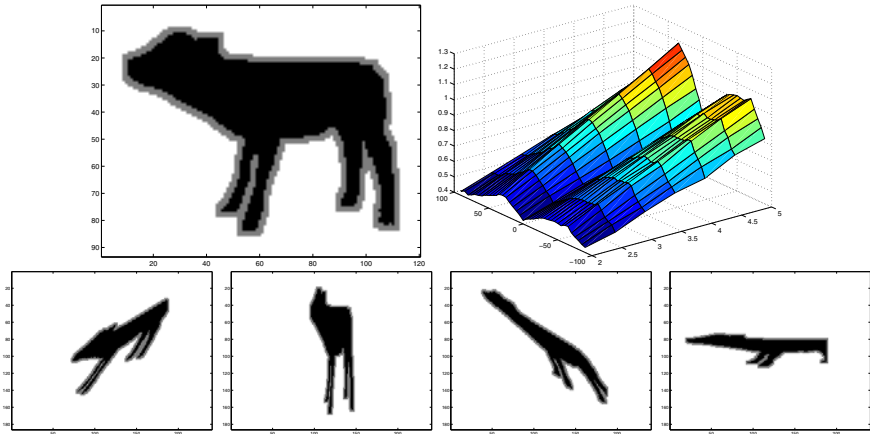$$A \mapsto P^\varepsilon_{(\theta,\beta)}(a)/R, \quad a \in r(A), A \in \$. \qquad (6)$$

Fig. 1. Calf (top left) and the associated representation $P_n(F_{(\gamma,\beta)})$. The bottom line represents $F_{(\gamma,\beta)}$ for $\beta = 2$ and $\gamma = 45, 90, -45, 0$ (the shape is in black and the neighborhood in grey).

On Figure 1, we can remark that the function $P^n_{(\gamma,\beta)}(a)$ increases more in two directions, corresponding to an extention of both the calf body and its legs.

We consider situation with no information about foreshortening and all rotations are equally likely to occur. Finally, we consider the following metric on $\mathcal{R}$:

$$l(\Phi(A), \Phi(B)) := \inf_{a,b} \left(\int_{[-\frac{\pi}{2},\frac{\pi}{2}]\times[1,\infty]} (P^\varepsilon_{(\theta,\beta)}(a) - P^\varepsilon_{(\theta,\beta)}(b))^2 \; e^{-\kappa\beta}d\beta \; d\theta \right)^{1/2} , \quad (7)$$

where $\kappa > 0$ is a parameter, $a \in r(A)$, $b \in r(B)$. The integral on the right-hand side converges due to $P^\varepsilon_{(\theta,\beta)}(a)$ is almost linear function of β for β big enough.

We thus have defined a map between the shape space and the feature space. Two similar shapes should be associated to close points in the feature space. This property can be established by the continuity of mapping Φ with respect to the metrics defined in both spaces. The proof of the following theorem we will omit.

Theorem 1. $\Phi : (\$, d(,)) \to (\mathcal{R}, l(,))$ *is a continuous map.*

3   Implementation and Results

3.1   Discretization

In practice, to compute the distance between two shapes, we have to discretize equation (7). When analysing the surfaces representing the function $P^n(F_{(\theta,\beta)}a)$

on Figure 1 it is clear that the embeded information is redundant. Indeed, the surfaces are very smooth, so that we can employ a drastic discretization scheme, without loosing information. We consider $n = 5$ pixels neighbourhood.

Before computing the proposed feature, we normalize the shapes to $V = 4000$ pixels (area) in order to satisfy the scale invariance.

3.2   MPEG-7 CE Shape-1 Part-B Data Set

We first evaluate the proposed approach for the 7 classes of well-known MPEG-7 CE Shape-1 Part-B data set (70 classes, each class containing 20 similar objects). Although the classes are quite distinct, this data set contains important within-class variations (see [11]).

We consider four directions, $\theta \in \{-\frac{\pi}{4}, 0, \frac{\pi}{4}, \pi\}$, and two expanding coefficients $\beta \in \{3, 5\}$. The coefficient defined the metric is $\kappa = \frac{1}{5}$.

We consider the proposed metric between each pair of shapes (except itself of course, cause distance is 0) and report in Table 1 the percentage of correct n^{th} neighbors for each class. The total correct answers correspond to 98% for the first neighbors. If we consider the tenth neighbors, we still obtain a total score of 85% of good retrieval. This shows the robustness of the proposed metric.

Table 1. Retrieval scores on the 7 classes of MPEG-7 database

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Bonefull	95	70	85	85	90	85	85	85	70	75
Heart	100	100	100	100	100	100	95	95	95	100
Glas	100	100	100	100	100	100	100	90	100	95
Fountain	100	100	100	100	100	100	100	100	100	100
Key	100	100	95	100	95	95	90	90	95	95
Fork	95	90	65	70	65	75	65	75	70	60
Hammerfull	95	95	80	30	30	35	30	40	35	15

3.3   Kimia Database

We considered the extended database defined by Kimia. It consists of 216 shapes divided into 18 classes (see [11]). We consider $\theta \in \{90, 65, 45, 30, 0, -30, -45, -65\}$, and two expanding coefficients $\beta \in \{3, 5\}$. The global retrieval score is 91.2% for the first neighbour, 80.1% for the second neighbour, 74.5% for the third neighbour, 72.7% for the fourth neighbour and 63% for the fifth neighbour. It shows the robustness of the proposed metric in case of a rather big database.

Our results for the first neighbour 91.2% (as well as 98% for the first database) are comparable with 98% for SC-method and 100% stated recently in [10] for Kimia 216 database. We besides have good result 80.6% for Bullseye score. It shows good localization of different classes in feature space. Our descriptor is very clear and transparent in implementation. It takes about 0.23 sec in Mat-Lab(Pentium 4, 2Gb) to find the first neighbour for one image among 215 others.

4   Conclusion

We have proposed a new metric on a shape space based on the shape properties after applying family of diffeomorphisms. The proposed metric is well-defined and continuous. Retrieval results on two databases, one of them consisting of 216 shapes, divided in 18 classes, have proven the relevance of this metric. Our image descriptor is invariant with respect to rotation, reflection, translation and scaling. We are currently studying the surjectivity of the associated mapping. We conjecture that surjectivity holds, at least for star shapes. Further studies (in particular 3-D shape retrieval) also include the definition of a shape classification algorithm, based on this description.

References

1. Dryden, I.L., Mardia, K.V.: Statistical shape analysis. Wiley, New York (1998)
2. Serra, J.: Image Analysis and Mathematical Morphology. Academic Press, New York (1982)
3. Kendall, D.G.: Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces. Bulletin of the London Mathematical Society 16, 81–121 (1984)
4. Loncaric, S.: A survey of shape analysis techniques, J. Pattern Recognition 31(8), 983–1001 (1998)
5. Zhang, D., Lu, G.: Reveiw of shape representation and description techniques. J. Pattern Recognition 37, 1–19 (2004)
6. Yong, I., Walker, J., Bowie, J.: An anlysis technique for biological shape. J. Comput. Graphics Image Process 25, 357–370 (1974)
7. Peura, M., Livirinen, J.: Efficiency of simple shape descriptors. In: Third International Workshop on Visual Form, Capri, Italy, May 1997, pp. 443–451 (1997)
8. Rucklidge, W.J.: Efficient locating objects using Hausdorff distance. Int. J. Comput. Vision 24(3), 251–270 (1997)
9. Belongie, S., Malik, J., Puzicha, J.: Matching shapes. In: IEEE Eighth Int. Conf. on Comput. Vision, Vancouver, Canada, July 2001, vol. I, pp. 454–461 (2001)
10. Bai, X., Latecki, L.J.: Path Similarity Skeleton Graph Matching. IEEE Trans Pattern Anal Mach Intell 30(7), 1282–1292 (2008)
11. Sebastian, T., Klein, P., Kimia, B.: Recognition of shapes by editing their shock graphs. IEEE Trans. Pattern Analysis and machine intelligence 25(5), 116–125 (2004)