Pattern Recognition 60 (2016) 164–176

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Intrinsic sample mean in the space of planar shapes

M.J. Flores [a], X. Gual-Arnau [b], M.V. Ibáñez [c,*], A. Simó [c]

[a] Dep. de Psicologia Bàsica, Clínica i Psicobiologia, Universitat Jaume I. Spain
[b] Dep. de Matemàtiques. Institute of New Imaging Technologies, Universitat Jaume I, Spain
[c] Dep. de Matemàtiques. Institute of Mathematics and Applications, Universitat Jaume I, Spain

ARTICLE INFO

Article history:
Received 18 September 2015
Received in revised form
29 April 2016
Accepted 30 April 2016
Available online 22 May 2016

Keywords:
Intrinsic sample mean
Parameterized closed curve
Shape space
Gradient descent algorithm

ABSTRACT

In this paper we consider the shape space as the set of smooth simple closed curves in $\mathbb{R}^2$ (parameterized curves), modulo translations, rotations and scale changes. An algorithm to obtain the intrinsic average of a sample data (set of planar shape realizations), from the identification of the shape space with an infinite dimensional Grassmannian is proposed using a gradient descent type algorithm. A simulation study is carried out to check the performance of the algorithm.

1. Introduction

Most methods of scientific research use data which are collected through different techniques. The analysis of these data using statistical procedures becomes essential for many studies and, sometimes, to obtain averages of some observed variables becomes the main goal of the work.

In particular, shape is an important feature of objects and can be immensely useful in characterizing them for the purpose of detection, tracking, classification, and recognition [1–3].

In the last years, the definition and study of spaces of planar shapes has met a large amount of interest. Three major approaches can be identified in Shape Analysis, based on how the object's shape is treated in mathematical terms [4]. Shapes can be treated as sequences of labeled points in the Euclidean space (landmarks) [5–9], as compact sets on $\mathbb{R}^m$ [4,10,11], or they can be described by functions representing their contour.

In this paper, the shapes will be represented by smoothly immersed planar curves, which will constitute the boundaries of compact domains (boundaries of physical objects projected into the imaging plane). Different works [12,13] have identified, through isometry, the space of shapes with a given metric with the Grassmann manifold of 2-planes and infinite dimension. From this identification, all the characteristic geometric elements of a manifold (geodesics, distances, curvatures, etc.) have been obtained in the space of planar shapes.

Indeed, an important aspect of shape analysis is to obtain the empirical mean of several shapes. These empirical means can be regarded as prototypes in pattern recognition problems. However, this concept is not trivial. When the shape space M admits the structure of a Riemannian manifold, two kinds of means have been studied as Fréchet parameters associated with two types of distances on M [14,15]. If j is an embedding of M in a Euclidean space, the mean in the embedding space can be computed, and its projection to $j(M)$ yields the extrinsic mean shape [16,17]. On the other hand, a Riemannian distance on M yields the intrinsic mean set and intrinsic mean shape [15,18]. In this sense, the intrinsic mean shape can be defined as the set of points that minimize the mean squared distance as measured along geodesics.

In this paper we consider the intrinsic Fréchet mean in terms of a finite number of discrete observations rather than in terms of a probability measure of a distribution on the manifold. It is important to recognize that a mean on a Riemannian manifold is defined with respect to a particular Riemannian metric and that different metrics may generally give rise to different means.

The main contribution of this paper is to propose an algorithm to obtain the intrinsic average of a sample data (set of planar shape realizations), from the identification of the shape space with an infinite dimensional grassmannian, and using a gradient descent algorithm. That is, we consider a Riemannian metric in the shape space to obtain the intrinsic mean shape [19,18]. Similar algorithms to the obtained here using Grassmannians can be found in [20] applied on image and video-based recognition problems.

To reach this contribution, in Section 2 we define the shape space as a Riemannian manifold and we review some basic concepts about the shape space of geometric curves [21,22].

* Corresponding author.
E-mail address: mibanez@uji.es (M.V. Ibáñez).

http://dx.doi.org/10.1016/j.patcog.2016.04.025
0031-3203/

Subsequently, the definition of intrinsic sample mean is reviewed and an algorithm to compute it in the shape space is described in Section 3. The performance of the algorithm is checked in a simulation study in Section 4. Section 5 shows how the methodology is put into practice on several examples. In Section 6 a comparative study with other techniques is included and, finally, the conclusions are stated in Section 7. All the analysis have been carried out using the software MATLAB[1] and R [23].

2. Shape space

Let us define the space of smooth planar closed immersed curves as

$$M = \{\alpha \in C^\infty(\mathbb{S}^1, \mathbb{R}^2):\ |\alpha'(t)| \neq 0, \quad \forall\, t \in \mathbb{S}^1\}, \tag{1}$$

where $\mathbb{S}^1$ is the unit circle which is identified with $\mathbb{R}/(2\pi\mathbb{Z})$ and $\alpha'(t)$ is the usual parametric derivative of the parametric curve α.

M is the space of embedded curves, so curves which differ by a translation, scaling or reparameterization are different elements of M.

The pre-shape space of parameterized planar closed curves is defined as the set of these curves modulo translations, scalings and rotations; that is,

$$\mathcal{B} = \frac{M}{\text{Similarities}}, \tag{2}$$

with Similarities={translations, rotations, scalings} and M as in Eq. (1).

Finally, the shape space (the shape of geometric curves, i.e., curves considered up to similarities and reparameterizations) is defined as:

$$\mathcal{S} = \frac{\mathcal{B}}{Diff(\mathbb{S}^1)}. \tag{3}$$

where $Diff(\mathbb{S}^1)$ denotes the group of diffeomorphisms of $\mathbb{S}^1$.

2.1. The Riemannian manifold of the pre-shape space

Different Riemannian metrics can be defined on $\mathcal{B}$. Following [18,21], given a curve α: $\mathbb{S}^1 \longrightarrow \mathbb{R}^2$ in $\mathcal{B}$ (a representative element of an equivalence class), the tangent space $T_\alpha\mathcal{B}$ can be identified with the set of vector fields h: $\mathbb{S}^1 \longrightarrow \mathbb{R}^2$ along α modulo constant vector fields, and a Riemannian metric can be considered here, given by:

$$G_\alpha\left(h, k\right) = \frac{1}{l(\alpha)} \int_{\mathbb{S}^1} \left\langle \dot{h}, \dot{k} \right\rangle_{\mathbb{R}^2} ds, \tag{4}$$

where $\langle \cdot, \cdot \rangle_{\mathbb{R}^2}$ is the usual product in $\mathbb{R}^2$, $\dot{h}$ is the derivative with respect to arc length, s, and $l(\alpha)$ is the length of α.

Let V be the vector space of all C^∞ mappings f, g: $\mathbb{S}^1 \longrightarrow \mathbb{R}$, with the dot product,

$$\langle f, g \rangle_\infty = \int_0^{2\pi} f(x) g(x) dx$$

and let $Gr(2, V)$ be the Grassmannian of unoriented 2-dimensional subspaces of V defined by the orthonormal pairs $(e, f) \in V \times V$, i.e.,

$$Gr(2, V) = \{(e, f) \in V \times V\ :\ \| e \|_\infty = \| f \|_\infty = 1,\ \langle e, f \rangle = 0\}.$$

Let $Gr^0(2, V)$ be the subset of $Gr(2, V)$ such that $\{t / e(t) = f(t) = 0\} = \varnothing$.

Then, assuming that the planar curves are curves in the complex plane $\mathbb{C}$, it has been proved in [21] that the map

$$\Phi: Gr^0(2, V) \longrightarrow \mathcal{B}$$

$$(e, f) \longmapsto \Phi((e, f)) = \frac{1}{2} \int_0^t (e(s) + i f(s))^2\, ds = \alpha(t), \tag{5}$$

is an isometry, using the natural metric on $Gr(2, V)$ and the metric G_α on $\mathcal{B}$.

It is easy to verify that for $\alpha = \Phi((e, f))$ defined as in Eq. (5), $\alpha(0) = (0, 0)$ and it is a closed curve because $\| e \|_\infty = \| f \|_\infty$ and $\langle e, f \rangle_\infty = 0$. Moreover, since $\| e \|_\infty = \| f \|_\infty = 1$ it can be stated that $l(\alpha) = 1$.

Vice versa, given a curve $\alpha \in \mathcal{B}$, if θ_α denotes the tangent-angle function of α, it can be proved [22] that the inverse of Φ is given by:

$$e(t) = \sqrt{\frac{2|\alpha'(t)|}{l(\alpha)}} \cos\left(\frac{\theta_\alpha(t)}{2}\right),$$

$$f(t) = \sqrt{\frac{2|\alpha'(t)|}{l(\alpha)}} \sin\left(\frac{\theta_\alpha(t)}{2}\right). \tag{6}$$

2.2. Geodesics and distances in $\mathcal{B}$ and $\mathcal{S}$

To compute distances and geodesics in $\mathcal{B}$ we will use the isometry Φ (Eq. (5)) and the methods used in [21,24].

Given two closed curves $\alpha, \beta \in \mathcal{B}$ such that $\alpha = \Phi(e_1, f_1)$ and $\beta = \Phi(e_2, f_2)$, with $(e_1, f_1), (e_2, f_2) \in Gr^0(2, V)$, the distance between the curves is defined as the distance between the two dimensional subspaces W_1 and W_2, generated by $\{e_1, f_1\}$ and $\{e_2, f_2\}$ respectively.

The singular value decomposition of the orthogonal projection p of W_1 in W_2 gives orthonormal bases $\{\hat{e}_1, \hat{f}_1\}$ of W_1 and $\{\hat{e}_2, \hat{f}_2\}$ of W_2 such that $p(\hat{e}_1) = \lambda_1 \hat{e}_2$ and $p(\hat{f}_1) = \lambda_2 \hat{f}_2$, where $0 \le \lambda_1, \lambda_2 \le 1$, $\hat{e}_1 \perp \hat{f}_2$ and $\hat{f}_1 \perp \hat{e}_2$. In fact, λ_1 and λ_2 are the singular values of the (2×2)-matrix

$$A = \begin{pmatrix} \langle e_1, e_2 \rangle & \langle e_1, f_2 \rangle \\ \langle f_1, e_2 \langle & \langle f_1, f_2 \rangle \end{pmatrix}.$$

If we write $\lambda_1 = \cos\psi_1$, $\lambda_2 = \cos\psi_2$ then ψ_1, ψ_2 are the Jordan angles, $0 \le \psi_1, \psi_2 \le \pi/2$, and according to [24] the geodesic distance between $\alpha = \Phi(e_1, f_1)$ and $\beta = \Phi(e_2, f_2)$ is given by

$$d(\alpha, \beta) = d(W_1, W_2) = \sqrt{\psi_1^2 + \psi_2^2}. \tag{7}$$

An upper bound of this distance is given by $d(\alpha, \beta) \le \pi/\sqrt{2}$.

The geodesic joining the curves $\alpha(t)$ and $\beta(t)$ is defined by [21]

$$\gamma_{\alpha,\beta}(u) = \Phi(e(t, u), f(t, u)) = \frac{1}{2} \int_0^t ((e(s, u) + i f(s, u))^2) ds, \tag{8}$$

where

$$e(t, u) = \frac{\sin((1-u)\psi_1)\hat{e}_1(t) + \sin(u\,\psi_1)\hat{e}_2(t)}{\sin\psi_1},$$

$$f(t, u) = \frac{\sin((1-u)\psi_2)\hat{f}_1(t) + \sin(u\psi_2)\hat{f}_2(t)}{\sin\psi_2}.$$

Therefore, if $\alpha = \Phi(e_1, f_1)$ and $\beta = \Phi(e_2, f_2)$, in order to obtain the geodesic we have to diagonalize the matrix A by rotating the curve α by a constant angle ϕ_α, i.e., the basis (e_1, f_1) by the angle $\phi_\alpha/2$; and similarly the curve β by a constant angle ϕ_β. The angles ϕ_α and ϕ_β are given by the equations

$$\phi_\alpha + \phi_\beta = 2\ \arctan\left(\frac{\langle e_1, f_2 \rangle + \langle f_1, e_2 \rangle}{\langle e_1, e_2 \rangle - \langle f_1, f_2 \rangle}\right),$$

[1] MATLAB is a trademark of The MathWorks Inc.