Optics & Laser Technology 62 (2014) 115–123

Contents lists available at ScienceDirect

Optics & Laser Technology

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

Pose recognition of articulated target based on ladar range image with elastic shape analysis

Zheng-Jun Liu*, Qi Li, Qi Wang

National Key Laboratory of Science and Technology on Tunable Laser, Harbin Institute of Technology, P.O. Box 3031, 2 YiKuang Street, Harbin 150080, China

ARTICLE INFO

Article history:
Received 15 July 2013
Received in revised form
10 February 2014
Accepted 14 February 2014
Available online 29 March 2014

Keywords:
Range image
Elastic shape analysis
Pose recognition

ABSTRACT

Elastic shape analysis is introduced for pose recognition of articulated target which is based on small samples of ladar range images. Shape deformations caused by poses changes represented as closed elastic curves given by the square-root velocity function geodesics are used to quantify shape differences and the Karcher mean is used to build a model library. Three kinds of moments – Hu moment invariants, affine moment invariants, and Zernike moment invariants based on support vector machines (SVMs) – are applied to evaluate this approach. The experiment results show that no matter what the azimuth angles of the testing samples are, this approach is capable of achieving a high recognition rate using only 3 model samples with different carrier to noise ratios (CNR); the performance of this approach is much better than that of three kinds of moments based on SVM, especially under high noise conditions.

1. Introduction

Ladar is capable of collecting intensity images and range images of targets. Because range images are able to represent high resolution details of three-dimensional (3D) shape information of complex targets, there is considerable interest in developing robust range image automatic target recognition (ATR) technologies as new ladar sensors become more widely available [1–3]. In ATR field, rigid targets are divided into two kinds: non-articulation and articulation. In the past decades, various approaches have aimed at non-articulated rigid targets, such as "spin-images" [4], "local surface patches" [5,6], moment invariants [7–9] and so on [10]. A few researchers discussed about recognizing articulated targets, such as Jones and Bhanu [11] and Weiss and Ray [12]. Grinnell et al. used Synthetic Aperture Radar (SAR) to scattering center locations and magnitudes as features that are invariant to articulation. Moreover, Weiss et al. applied the multi-scale region-based invariant transform to represent articulated targets of range images.

This paper does not discuss the invariance of articulation, but addresses the issue of pose recognition of articulated targets, e.g., tanks or excavators, from single range images. Pose recognition refers to the modeling and recognition of several poses performed by an articulated target in the form of a sequence of motions. These poses are typically used to communicate certain control commands and request a machine with vision capabilities, especially in applications such as human–computer interaction and robotics.

Shape is a very critical feature representation to describe transformations that are caused by different poses of articulated targets. Most approaches of shape representation focus on similarity measures of pairwise shapes. Shape representation includes two kinds of shape descriptors: contours and skeletons (or medial axis). Compared with a point set, a contour contains the relationship between orders of target boundary; compared with skeletons, contours, which have a good robustness for large deformation between similar objects, are insensitive to noise and boundary perturbation.

In this paper, we applied the elastic shape analysis to manifold geometry for boundary curves of articulated targets. Over the last few years, despite the proposed multitudes of metrics, there is an emerging consensus on the suitability of the elastic metric for curve shape analysis. This metric uses a combination of bending and stretching/compression to find optimal deformations from one shape to another [13]. These deformations are studied as the shortest paths, or geodesics, under this chosen metric in a certain shape space. This metric was first suggested by Younes et al. [14] and subsequently utilized by Mio et al. [15], who developed an algorithm to compute geodesic paths between arbitrary shapes. The choices of shape representation and Riemannian metric are critically important to improve understanding and to compute efficiently.

Recently, Joshi et al. [13] first introduced representation of square-root velocity function (SRVF) and Srivastava et al. [16]

* Corresponding author.
E-mail address: zhengjunliu360@aliyun.com (Z.-J. Liu).

http://dx.doi.org/10.1016/j.optlastec.2014.02.016
0030-3992/

subsequently extended the representation. In 2011, Abdelkader et al. applied the approach to the recognization of 2D human gestures from videos [17]. A similar idea was introduced by Younes et al. [14]. Compared with other presentations, elastic shape analysis of SRVF has the following advantages: 1) the elastic matching of curves allows nonlinear registration and improves the match of features across silhouettes [13]; 2) the method of representation is invariant to rotation, translation, and scaling, as well as reparameterization; 3) the reparameterizations of curves in this metric do not change Riemannian distance between different curves and thus help removing the parameterization variability from the analysis. Furthermore, the approach is useful because it allows us to perform intrinsic statistical analysis tasks, such as shape modeling and Riemannian spaces.

In this paper, we have applied the representation of SRVF for pose recognition of 3D articulated targets. However, the approach of elastic shape analysis cannot keep the out-of-plane rotation invariance. Moreover, in the real application, it is difficult and expensive to acquire enough samples for ladar. Thus, in small samples, we need a representation that is robust to deformations from translation, scaling, rotation (including in-plane and out-of-plane rotation) and is capable of distinguishing the difference among poses caused by articulation angles. Therefore, we analyze how to build an effective model library. Moreover, using 3 and 7 model samples for each pose, recognition rates of elastic shape analysis in arbitrary azimuth angles, namely out-of-plane rotation, are analyzed and compared with three kinds of moment invariants – Hu moment invariants (HMIs), affine moment invariants (AMIs), Zernike moment invariants (ZMIs) – with support vector machine (SVMs).

This paper is organized as follows. Section 2 reviews the approach of elastic shape analysis. In Section 3, three experiments are carried out to evaluate and verify the effectiveness of elastic shape analysis for the pose recognition of articulated vehicles with simulated range images. The conclusion is presented in Section 4.

2. Review of elastic shape analysis approach

In the entire shape analysis, a shape space is typically constructed in two steps: 1) a mathematical representation of curves with appropriate constraints leads to a preshape space and 2) one identifies elements of the preshape space that belongs to the same orbits of shape-preserving transformations (rotation, translation, and scaling, as well as re-parameterization) [17]. The resulting quotient space is the desired shape space. We describe such representations as below.

2.1. SRVF representation in preshape space and geodesics

Let β be a contour curve of shape, parameterized by an arbitrary SRVF: $q: S^1 \to R^2$ (S^1 is a unit circle in R^2 centered at the origin):

$$q(t) = \beta'(t)/\sqrt{\|\beta'(t)\|} \tag{1}$$

where $\beta'(t)$ is the velocity vector of β and $\|\cdot\|$ is the Euclidean norm in R^2. In fact, this curve can be obtained using the equation $\beta(t) = \int_0^t q(s)\|q(s)\|ds$ [16].

In order to compare curves quantitatively, suppose that they are made of an elastic material and adopt a metric that measures the difficulty in reshaping a curve into another taking elasticity into account. Infinitesimally, this can be done using a Riemannian structure on M (n-dimensional manifold). The Riemannian metric consists of inner products $\langle \cdot, \cdot \rangle_x$ in the tangent space $T_xM, x \in M$, which varies smoothly along the manifold [18]. If $a: [0, \ 1] \to M$ is a differentiable path in M, then its length is given by $L[a] = \int_0^1 \langle a'(t), a'(t) \rangle^{1/2} dt$. The Riemannian distance between two points $q_0, q_1 \in M$, denoted as $d(q_0, q_1)$, is defined as the minimum length over all the paths on the manifold between x and y. A geodesic path can locally minimize the length between points.

The geodesics are calculated using a path-straightening approach where the geodesic path between the two points is first initialized with a path $a(t)$. For any two closed curves, denoted by $q_0, q_1 \in M$, a geodesic path is $a: [0, \ 1] \to M$:

$$a(t) = \frac{1}{\sin(\theta)}(\sin(\theta(1-t))q_0 + \sin(\theta t)q_1) \tag{2}$$

where $\theta = \cos^{-1}(<q_0, q_1>)$ denotes the length of the geodesic [16]. Then, this path is iteratively straightened using a gradient approach and the limit point of this algorithm is a geodesic path. To be effective for the shape analysis, the representation and the geodesics between the points must be invariant to the shape-preserving transformations [19].

The representation of a parametric curve q is clearly an invariant of translation since it is based on the velocity field of β. The scale variance will be obtained by fixing the length to be 1, such that $\int_0^1 q(s)ds = 1$. Moreover, in order to study the shapes of the closed curves, an additional condition that the curve starts and ends at the same point is imposed, as $\int_0^1 q(s)\|q(s)\|ds = 0$. The space that satisfies these conditions is called preshape space, as $C \subset M$ [16].

2.2. Removing shape-preserving transformations

The preshape representation of a curve is invariant to translations and scales. However, the representation is sensitive to reparameterizations and rotations. The two are studied as groups acting on C: rotation by the action of $SO(2)$ and reparameterization by the action of Γ, where $\Gamma = \{\gamma: S^1 \to S^1\}$ is the space of all orientation-preserving diffeomorphisms and $SO(2)$ is the special orthogonal group of 2×2 matrixes. The parameterization γ is a nonlinear monotonic differentiable function that lends the elastic properties of curves. The shape space of closed curves S is defined as the quotient space of the poses of $SO(2)$ and Γ; that is $S = C/(SO(2) \times \Gamma)$. To compute the geodesics between the two shapes, solve the optimization equation first:

$$d([q_0], [q_1]) = \min_{O \in SO(2), \gamma \in \Gamma} d(q_0, O(q_1 \circ \gamma)\sqrt{\gamma'}) \tag{3}$$

To remove the two transformations, we need to solve the joint minimization problem on (γ,O) in formula (3), with the cost function $H: \Gamma \times SO(2) \to R$, $H(\gamma, O) = d(q_0, O(q_1 \circ \gamma)\sqrt{\dot{\gamma}})$. The solution is obtained by iteratively fixing one variable and optimizing over the other one. For a fixed γ, the optimization of $H(\gamma, \cdot)$ over $SO(2)$ is obtained using the singular value decomposition (SVD); while the optimization of $H(\cdot, O)$ over Γ is performed using the dynamic programming (DP) algorithm, for a fixed $O \in SO(2)$ [16].

2.3. Computation of mean shape

Let $\{q_1, q_2, \ldots, q_n\}$be a collection of closed shapes. This intrinsic mean $\bar{q}$ is given by

$$\bar{q} = \frac{1}{n} \arg\ \min_{q \in S} \sum_{i=1}^{n} d(q, q_i)^2 \tag{4}$$

where d denotes the geodesic distance in preshape space after a DP alignment. To search for a Karcher mean of the collection, a gradient-type strategy is adopted. Given a tangent vector $v \in T_qC$, there exists a locally unique geodesic, $a_v(t)$, starting at q with t as its initial velocity and traveling with constant speed. The Riemannian exponential map, exp_q: $T_qM \to M$, maps a tangent vector v to a point on the manifold that is reached in unit time by the geodesic $a_v(t)$. The inverse of exp_q is known as the logarithm map and is