



ELSEVIER

Contents lists available at ScienceDirect

Pattern Recognition

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

Integral invariants for space motion trajectory matching and recognition



Zhanpeng Shao, Youfu Li*

Department of Mechanical and Biomedical Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong

ARTICLE INFO

Article history:

Received 10 March 2014

Received in revised form

16 February 2015

Accepted 27 February 2015

Available online 11 March 2015

Keywords:

Motion trajectory

Integral invariants

Maximal blurred segment

Similarity measure

Trajectory matching

Sign recognition

ABSTRACT

Motion trajectories provide a key and informative clue in motion characterization of humans, robots and moving objects. In this paper, we propose some new integral invariants for space motion trajectories, which benefit effective motion trajectory matching and recognition. Integral invariants are defined as the line integrals of a class of kernel functions along a motion trajectory. A robust estimation of the integral invariants is formulated based on the blurred segment of noisy discrete curve. Then a non-linear distance of the integral invariants is defined to measure the similarity for trajectory matching and recognition. Such integral invariants, in addition to being invariant to transformation groups, have some desirable properties such as noise insensitivity, computational locality, and uniqueness of representation. Experimental results on trajectory matching and sign recognition show the effectiveness and robustness of the proposed integral invariants in motion trajectory matching and recognition.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

A motion trajectory, which records a sequence of moving positions of a tracked object, provides a compact and representative clue for motion characterization. It has been extensively studied for describing activities, behaviors, and motion patterns in different applications such as learning motion patterns [1,2], human action recognition [6,7], human–robot interaction [3], gesture recognition [4,8], and trajectory retrieval [49]. As these applications suggest, motion trajectories play an important role in determining the contents of video, perceiving similar motion patterns, retrieving actions, and so on. Min et al. [6] employed motion trajectories tracked from some body joints as features input to a discriminant model for classifying human activities. Yang et al. [8] modeled a gesture recognition system using a time-delay neural network, where motion patterns are learned from hand trajectories. Oikonomopoulos et al. [9] also tracked hand trajectories to understand human actions based on the RVM discriminative model [10]. Apart from modeling human actions, motion trajectories of objects of interest are often utilized to build

some activity models to understand and retrieve motion patterns in video surveillance [13,49] and for information visualization [5]. Nevertheless, in most related work motion trajectories were often directly used in the raw data form with naïve processing. The raw data rely on the absolute positions of motions in a coordinate system, and are, therefore, ineffective in computation and are sensitive to noise. Not surprisingly, they will change under different viewpoints. Therefore, most space motion trajectory features cannot be captured directly by the raw data.

Shape description has received considerable attention in computer vision for shape matching and classification. In this regard, shape descriptors for describing object contours are closely related to our research. In [21,18], Curvature scale space (CSS) was developed for shape matching. Curvatures of a shape contour at different scales are produced by convolving the shape contour with a series of Gaussian kernels in a coarse to fine manner, where the shape contour is deformed at varying scales, yielding undesirable distortions in the shape. By chain code [11], one can digitize a space curve in terms of relative direction changes of segmented lines. However the relative changes with respect to neighboring lines limited the use of chain code for complex space curves. Using algebraic curve, such as B-spline [24] and Bezier curve [12], a shape contour can be approximated through some key control points. These curve fitting methods show non-uniqueness when the sampling rate of motion trajectories varies or partial occlusions exist in a trajectory, because their approximation accuracies depend on those key control points. Shape context [16] was introduced to capture the histogram bins of neighboring points at each reference point of a curve using a log-polar weight kernel. As a local descriptor, shape context possesses

Abbreviations: MBS, Maximal 3D Blurred Segment; DTW, Dynamic Time Warping; 1-NN, 1-nearest neighbors; FD, Fourier descriptor; II, integral invariants; DI, differential invariants; DII, distance integral invariants; Z, integer set; Z^3 , 3-dimensional Euclidean space with only integer sets; I, non-empty interval of real numbers

* Corresponding author. Fax: +852 27888423

E-mail addresses: perry.shao@my.cityu.edu.hk (Z. Shao), meyfli@cityu.edu.hk (Y. Li).

<http://dx.doi.org/10.1016/j.patcog.2015.02.029>

0031-3203/© 2015 Elsevier Ltd. All rights reserved.

rich invariant properties, and is capable of handling occlusions, but it is not the best way to describe a space trajectory due to its coarse distributions captured for a shape. Therefore, even though these shape descriptors have shown good performance in specific applications, they are incapable of fully capturing motion trajectory features in 3D Euclidean space (3D), due to their limited representation capability for simple shape contours (Chain code and shape context), sensitivity to noise (CSS and chain code) and non-uniqueness (B-spline). Several moment invariants for 3D curves under similarity transformations were derived in [47], but they are global descriptors with two limitations: high-order moments are sensitive to noise [48] and they cannot admit invariant to occlusions. In addition, transform functions based on Fourier and Wavelet [22,23] extracted global features from a trajectory, but meanwhile the local features are lost. They are also not stable with respect to noise due to high-order Fourier coefficients involved. All of these shape or curve descriptors were initially constructed for simple planar shapes. They are thus insufficient and non-compact to semantically represent complex 3D space motion trajectories. Therefore, they cannot be extended straightforward to our research. As discussed above, a good descriptor for free-form motion trajectories is expected to be robust to noise and occlusions, and invariant to specific group transformations. Therefore, such a descriptor should necessarily satisfy a number of criteria, some of which are consistent with CSS [21]: uniqueness, invariance, noise resistance, and locality, and it needs to be applicable to both planar and space trajectories.

In this paper, we first review the related literature in invariants, and claim our contributions in invariant representation in Section 2 before proposing the definition and estimation of the new integral invariants in Section 3. In Section 4, we propose the similarity measure that allows warping motion trajectories with various temporal lengths onto each other. We conduct two experiments in Section 5 to show the properties, robustness of the proposed integral invariants through trajectory matching, and their effectiveness in sign recognition. Finally, we conclude the paper in Section 6.

2. Previous work on Invariants and our contributions

Invariants have played an important role for various applications in computer vision ranging from shape representation and matching [17] to object recognition [25] and gesture recognition [19]. Consequently plentiful features that are invariant to specific transformations (affine, similarity, Euclidean) have been investigated in [37,38,40–42]. Two invariant local descriptors related to our research are differential invariants and integral invariants, which have been investigated and put into applications in motion trajectory representation and recognition [32,33]. However, there is a major roadblock that is noise in motion trajectories. Whenever noise is present to affect the spatio-temporal primitives of motion trajectories, differential invariants would be dominated by even small-scale perturbations in that the computation of differential invariants involves high-order derivatives and hence are very sensitive to noise, even though they are approximated [25,27] in terms of joint invariants. Some approaches have tried to overcome this drawback for differential invariants by the introduction of the scale-space smoothing in [39], but a more effective and robust method has so far not been available in principle. There has been much work to attempt deriving integral invariants [14,15,41] based on integral operation. In [41], potentials were proposed to obtain integral invariants for planar shapes via integrating the potentials of the contour curves of shapes, but these integral invariants are global descriptors. Integral invariants for closed planar shapes [14,15] were derived by performing integration of a class of local kernels along the shape boundary represented by a planar curve, where the

locality is achieved by restricting integration to local neighborhoods at each point of the curve. Nevertheless, they cannot be extended to represent motion trajectories in 3D case in that the fact of open contours and varying orientations of space trajectories complicates the problems. Consequently, the idea of defining a class of kernel functions along a space trajectory, which admits invariant under group transformations, remains unresolved. Developing an effective, robust invariant representation for space trajectory matching and recognition is a promising topic. In our approach, we will extend the integral invariants for planar closed shapes [14] to define some new integral invariants for free-form space trajectories in 3D Euclidean space.

In this paper, we propose some new integral invariants for space motion trajectories using line integrals of a class of kernel functions along a motion trajectory. Depending on two designed kernel functions, we have two typical integral invariants of transformation groups, the distance and area integral invariants. In this paper, we favor the area integral invariants as our integral invariants in this paper, and define them as the line integrals over a dynamic domain of integration within instant Frenet–Serret frame along a motion trajectory. We then derive a novel estimation formula accordingly to approximate the proposed integral invariants for discrete trajectories based on the Maximal Blurred Segment of discrete curves. In order to match trajectories, we define a distance function of dynamical time warping as the similarity measure between a pair of motion trajectories able to deal with some nonlinear variations including different sampling rates, unequal lengths, and occlusions. Finally, we show the effectiveness, robustness of the integral invariants in motion trajectory matching and recognition.

3. Integral invariants

A space motion trajectory is a sequence of position vectors of a moving object in 3-dimensional Euclidean space, and we denote it with $\gamma: I \rightarrow \mathbf{R}^3$, parameterized by temporal sequence t :

$$\gamma(t) = \{x(t), y(t), z(t) \mid t \in [a, b]\}, \quad (1)$$

where $[a, b] \in I$ is the time interval and we assume the motion trajectory γ in this paper is a regular curve, i.e., $\|\dot{\gamma}(t)\| \neq 0$ at all t . Normally, a space motion trajectory can also be parameterized with respect to arc length s , $\gamma(s) = \{x(s), y(s), z(s)\}$. Note that in practical scenarios motion trajectories are often sampled discretely, and thus t is set to $[1, N] \in \mathbf{Z}$ in this case, where N is the trajectory length (frames). In numerical integration of integral invariants, this discrete temporal parameterization is suitable for modeling the approximation algorithm for numerical integration, whereas in the definition of integral invariants we assume a motion trajectory is a regular curve parameterized with respect to continuous arc length or time sequence.

In this section, we first propose a general definition of integral invariant for a space motion trajectory and then exemplify two specific integral invariants depending on two typical kernels, respectively. The estimation of the area integral invariants for a discrete trajectory is then derived based on the maximal blurred segment of discrete curves. In practical applications, motion trajectories commonly occur in both \mathbf{R}^2 and \mathbf{R}^3 Euclidean space. Therefore, the proposed integral invariants can be applied to both planar and space trajectories thanks to the property of \mathbf{R}^2 being the subgroup of \mathbf{R}^3 , where the integral invariants for planar trajectories are just the special instances without loss of generality.

3.1. Definition of integral invariant for space trajectory

As addressed in [14,25], it can be deduced that two motion trajectories are equivalent if and only if one can be mapped to another one by a group transformation. Furthermore, they are

equivalent as their invariants up to a group transformation are identical. Hence, a group-invariant for a space motion trajectory is defined in Definition 1.

Definition 1. Let G be a transformation group acting on \mathbf{R}^3 . The function $I: \mathbf{R}^3 \rightarrow \mathbf{R}$ is a G -invariant if it satisfies $I(\gamma) = I(g \bullet \gamma), \quad \forall g \in G$ (2)

The function $I(\bullet)$ at each point of a trajectory γ returns a real number. According to the definition, there are some familiar examples of invariants, such as curvature and torsion (differential invariants) in [25,33]. Curvature and torsion in 3D space are useful features for describing space curves at trajectory matching and recognition [32,33]. They are invariant to Euclidean group $E(3)$ and similarity group $S(3)$ which will be considered as the transformation group G for the following definition of integral invariants.

Integral invariants proposed by Manay et al. [14] are effective and powerful representations for planar shape matching. Extending them to represent space trajectories in \mathbf{R}^3 Euclidean space is quite promising. To this end, we have to first define a meaningful domain of integration over which a motion trajectory can be integrated in \mathbf{R}^3 space. Normally, the domain of integration can be an interval, area, or volume. Instead, we propose integral invariants for a motion trajectory based on the line integrals at a class of kernel functions along the motion trajectory, where the domain of integration is a trajectory (path) instead of an interval at a fixed axis. In addition, to preserve the locality of integral invariants, we employ a local function to restrict integration over the neighborhoods at each reference point. Taking into account these aspects, a general definition is made for a space trajectory with Definition 2. We denote transformation group acting on \mathbf{R}^3 with $G \in E(3) \cup S(3)$; we also use the formal notation μ to indicate either time instant t or arc length s along a trajectory.

Definition 2. A function $I(p)$ is an integral G -invariant for a space trajectory γ at point p if there exists a kernel function $k: \mathbf{R}^3 \times \mathbf{R}^3 \rightarrow \mathbf{R}$ such that the line integral of the kernel function along the trajectory γ :

$$I(p) = \int_{\gamma} k(p, \gamma(\mu)) d\mu \quad \forall p \in \gamma, \quad (3)$$

where $k(\ast, \ast)$ satisfies

$$\int_{\gamma} k(p, \gamma(\mu)) d\mu = \int_{g\gamma} k(gp, g \bullet \gamma(\mu)) d\mu \quad \forall g \in G, \quad (4)$$

In the definition, designing of the kernel function $k(p, \gamma)$ can be chosen freely depending on the final goal. There are two specific integral invariants for a space trajectory, the distance and area integral invariants depending on two designed kernel functions, respectively, exemplified in Sections 3.2 and 3.3.

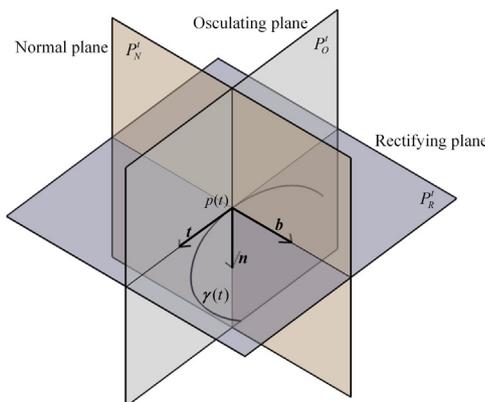


Fig. 1. Frenet-Serret frame at a reference point of a space trajectory.

3.2. Distance integral invariant

Consider $G = E(3)$ and the following integral invariant, given the kernel $k(\ast, \ast)$, computed at point $p \in \gamma$:

$$I(p) = \int_a^b k(p, \gamma(t)) dt = \int_a^b q(p, \gamma(t)) d(p, \gamma(t)) dt, \quad (5)$$

where $d(p, \gamma(t)) := \|p - \gamma(t)\|$ is the Euclidean distance in \mathbf{R}^3 , and the weighting function $q(\ast, \ast)$ can be designed freely to restrict the computation of $d(p, \gamma(t))$ in neighboring points to preserve the locality. In this paper, we let $q(\ast, \ast)$ be a ball kernel defined in Eq. (6). It is shown immediately that this is a Euclidean invariant since the distance is preserved under Euclidean transformations, but it needs to be further normalized to get invariance under similarity group transformations $S(3)$. In terms of the definition, the value of $I(p)$ is actually the average distance from the current point p to its neighboring points restricted by a kernel function. This distance integral invariant can be thought of as a continuous version of 3D shape context [16]. The difference is that the shape context signature at a reference point captures the histogram bins of the neighboring points using a log-polar weight kernel at the point, whereas in our case, we only retain the mean distance of the neighboring points relative to the reference point.

As suggested in [14], the distance invariants are not discriminative enough in that a unique distance invariant can correspond to different geometric features. This ambiguity for distance invariants certainly exists in 3D space, which is a motivation for us to introduce the next integral invariants.

3.3. Area integral invariants

3.3.1. The domain of integration over space trajectory

Considering some meaningful area integral invariants for a space trajectory, we have to first define an effective domain on the trajectory in which a kernel function can be integrated over in terms of Definition 2. Nevertheless, the fact of non-closed contours and varying motion orientations of space trajectories make it difficult to have an invariable domain of integration along motion trajectories. Therefore, we propose a dynamic domain of integration, where the domain at each point locates on a specific projection plane onto which a motion trajectory can be projected. In terms of differential geometry, the Frenet-Serret frame associated with a space curve is a typical descriptive framework to describe the dynamic evolution of the curve in 3D Euclidean space, where at each reference point there are three orthogonal unit vectors: the tangent vector \mathbf{t} , principle normal vector \mathbf{n} , and binormal vector \mathbf{b} . The osculating plane, $P'_O = span(\mathbf{t}, \mathbf{n})$, at the reference point $p(t) \in \gamma$ is spanned by the tangent vector \mathbf{t} and normal vector \mathbf{n} , while the other two planes, normal plane P'_N and rectifying plane P'_R , are spanned in the same way as shown in Fig. 1.

Based on the Frenet-Serret frame, the Frenet-Serret formulas are then configured using differential equations that describe the dynamic relations among the three unit vectors [26]. In the Frenet-Serret formulas, the most descriptive parameters for a space curve are curvature and torsion. Intuitively, we define the domain of integration at a reference point on the osculating and rectifying planes at the reference point in that the instant curvature and torsion are in principle evaluated on these two projection planes, through which we can easily connect the area integral invariants to traditional differential invariants at certain conditions discussed in Section 3.3.3 later. As a result, it comes to a series of osculating and rectifying planes, P'_O and P'_R , across all the points. These two planes at each point would serve as projection planes in which a space trajectory γ can be orthogonally projected onto, respectively, forming two projected trajectories, $\bar{\gamma}'_O$ and $\bar{\gamma}'_R$. The integration along these two projected trajectories can be computed independently as

shown in Fig. 3. And for the bounding of integration, it is the neighborhoods centered at the reference point bounded by a ball kernel in this paper shown in Fig. 2. More generally, the ball kernel can be substituted by a Gaussian kernel.

3.3.2. Definition of area integral invariants

A ball function, as shown in Fig. 2, is defined by $B_r: \mathbf{R}^3 \times \mathbf{R}^3 \rightarrow \{0, 1\}$ to indicate whether a point ε is located on the interior of a sphere with radius r centered at p defined by function $B_r(p, \varepsilon)$,

$$B_r(p, \varepsilon) = \begin{cases} 1 \dots |p - \varepsilon| \leq r \\ 0 \dots \text{otherwise} \end{cases} \cdot \forall p \in \gamma, \varepsilon \in \mathbf{R}^3, \quad (6)$$

Consider $G = E(3) \cup S(3)$ and for any given radius r , the area integral invariant at a reference point p on the osculating projection plane is defined as a line integral of a kernel function f_o along the projected trajectory $\bar{\gamma}_o^p$ in terms of Definition 2:

$$I_o^r(p) = \int_{\bar{\gamma}_o^p} f_o(p, \gamma(s)) ds = \int_a^b f_o(p, \gamma(t)) \|\bar{\gamma}_o^p(t)'\| dt, \quad (7)$$

where ds is arc length, indicating that we are moving along a trajectory $\gamma(s)$ parameterized by arc length. Kernel f_o can be an arbitrary continuous function, which we define here as a circle along the projected trajectory on the osculating plane, and where we set the radius of the circle as r equal to the ball kernel B_r acted on the trajectory, as indicated in Fig. 4. Since f_o is integrated along the projected trajectory $\bar{\gamma}_o^p$, it is actually the distance from the

projected trajectory to the circle

$$f_o(p, \gamma(t)) = B_r(p, \gamma(t)) \left(\sqrt{r^2 - \{(\gamma(t) - p) \cdot \mathbf{t}\}^2} - (\gamma(t) - p) \cdot \mathbf{n} \right) \Rightarrow$$

$$I_o^r(p) = \int_a^b B_r(p, \gamma(t)) \left(\sqrt{r^2 - \{(\gamma(t) - p) \cdot \mathbf{t}\}^2} - (\gamma(t) - p) \cdot \mathbf{n} \right) \|\bar{\gamma}_o^p(t)'\| dt, \quad (8)$$

As $(ds/dt) = \|\bar{\gamma}_o^p(t)'\|$, it here stands for the arc length derivative with respect to t in the projected trajectory $\bar{\gamma}_o^p$. We drive $\|\bar{\gamma}_o^p(t)'\|$ to be represented in terms of the expression $\gamma'(t)$ by transforming the original coordinate of the motion trajectory $\gamma(t)$ to the orthonormal basis of the Frenet–Serret frame at p :

$$\frac{ds}{dt} = \|\bar{\gamma}_o^p(t)'\| = \sqrt{(\gamma'(t) \cdot \mathbf{t})^2 + (\gamma'(t) \cdot \mathbf{n})^2} = \|\gamma'(t) \cdot [\mathbf{t} \ \mathbf{n} \ \mathbf{0}]\| \quad (9)$$

And, therefore, by Eqs. (8) and (9) we have

$$I_o^r(p) = \int_a^b B_r(p, \gamma(t)) \left(\sqrt{r^2 - \{(\gamma(t) - p) \cdot \mathbf{t}\}^2} - (\gamma(t) - p) \cdot \mathbf{n} \right) \|\gamma'(t) \cdot [\mathbf{t} \ \mathbf{n} \ \mathbf{0}]\| dt \quad (10)$$

In the same manner, we can obtain an area integral invariant on the rectifying plane accordingly

$$I_r^r(p) = \int_a^b B_r(p, \gamma(t)) \left(\sqrt{r^2 - \{(\gamma(t) - p) \cdot \mathbf{t}\}^2} - (\gamma(t) - p) \cdot \mathbf{b} \right) \|\gamma'(t) \cdot [\mathbf{t} \ \mathbf{0} \ \mathbf{b}]\| dt \quad (11)$$

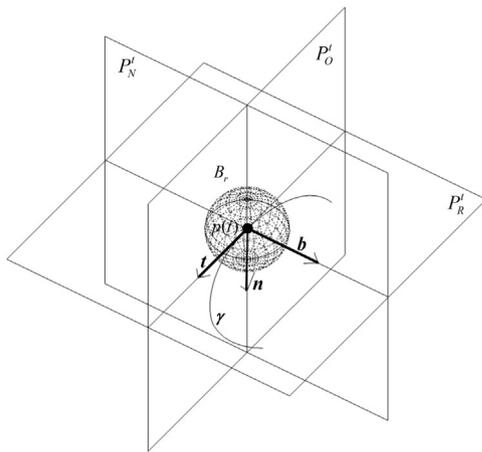


Fig. 2. Ball kernel acting on a reference point of a space trajectory.

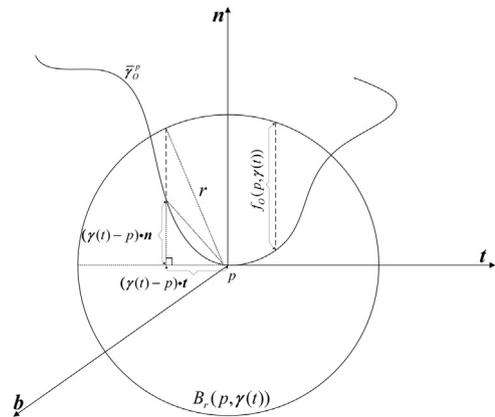


Fig. 4. Definition of the kernel function for the area integral invariant at a point of the projected trajectory on the osculating plane. In the same manner, the definition is the same as on the rectifying plane.

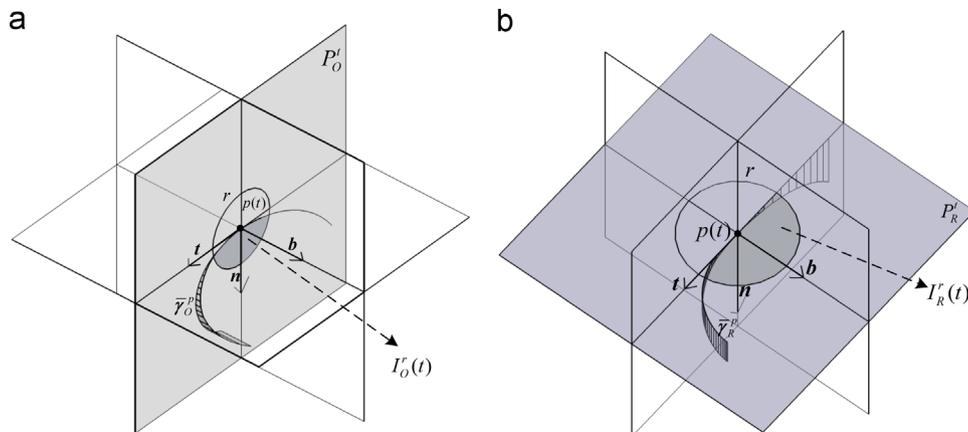


Fig. 3. Domain of integration on (a) the osculating plane and (b) the rectifying projection plane restricted by a ball kernel shown in Fig. 2. In the figure, the space trajectory is projected onto the osculating and rectifying planes of the Frenet–Serret frame at the reference point.

Finally, we can obtain the complete area integral invariants for a space trajectory at point p represented by the vector

$$I^r(p) = \{I^r_O(p), I^r_R(p)\}, \tag{12}$$

where all variables with the lower indicator R or O in equations denote their measure on the rectifying and osculating planes, respectively. As shown in Fig. 3, the area integral invariants $I^r(p)$ indicate the shaded areas, each of which lie in the interior of the circle with radius r centered at p on the corresponding projection plane as shown in Fig. 3, and their visualization along a space trajectory is shown in Fig. 5.

In 3D space, the area integral invariants here are in principle an extension of planar area integral invariants [14] in which all the points on a planar trajectory lie on a consistent osculating plane with

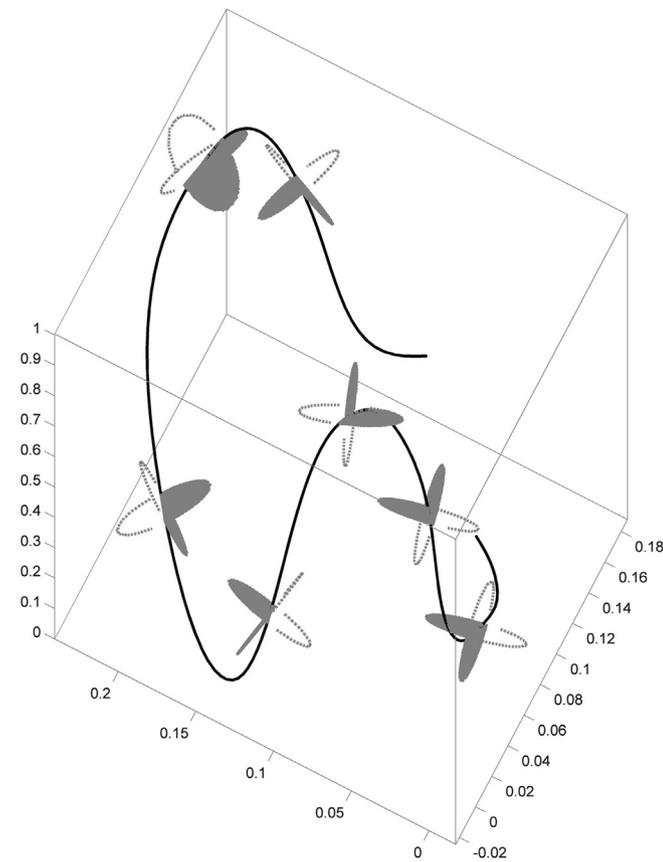


Fig. 5. Visualization of area integral invariants on both the osculating and rectifying planes of a space trajectory. The shaded areas represent the area integral invariants along the trajectory.

the torsion $\tau = 0$. We can observe from Eqs. (7)–(11) that the area integral invariants are derived based on local distances and areas within local Frenet–Serret frames, all of which are independent of specific coordinate systems. The dot and cross products of derivatives are purely local quantities of the trajectory features. Therefore, the area integral invariants are Euclidean invariants. They also can admit invariant to similarity group $S(3)$ by normalizing them

$$I^r(t)^* = I^r(p(t))^* = \frac{I^r(p(t))}{\pi * r^2} \tag{13}$$

where the normalized area integral invariants are bounded between 0 and 1, and would be the integral invariants we favor in the remainder of this paper.

3.3.3. Relation of local area integral invariants to curvature and torsion

It is noticed that there are some intrinsic connections between traditional differential invariants and the local area integral invariants under certain conditions. Torsion and curvature, traditional differential invariants for space trajectories [25], are considered as complete invariants since they allow the recovery of the original trajectory. Furthermore, all the differential invariants with any order can be in principle derived from functions of curvature and torsion in 3D space. Therefore it is an elegant way to tap into the results of differential invariants by linking our area integral invariants to curvature and its derivative, torsion, without confronting the drawbacks of high-order derivatives.

As the planar area integral invariants in [14] are a subset of our area integral invariants, the relation between curvature and the local area integral invariants can be thus derived straightforwardly in the same way as in [14]. Local curvature measures the deviance of a curve from being a straight line relative to the osculating plane at each point, at one of which the curvature can be computed locally based on the osculating circle as shown in Fig. 6(a). $I^r_O(p)$ denotes the interior region on the osculating plane, which is the intersection area between a circle with radius r and the projected curve $\bar{\gamma}_O$, and then can be approximated by the area of the shaded sector in Fig. 6 (b), i.e., $I^r_O(p) \approx r^2\theta$. Now, the angle θ can be computed from the cosine function $\cos \theta = r/2R$. Since the curvature is the reciprocal of the radius of the osculating circle, i.e., $\kappa = 1/R$, we have the same result as [14]

$$I^r_O(p) \approx r^2 \cos^{-1}(\frac{1}{2}r\kappa(p)) \tag{14}$$

Now, in terms of Eq. (14), the curvature at each point can be recovered from the area integral invariants whose approximation will reach the real values in limit when $r \rightarrow 0$.

Torsion, only existing for describing space curves, measures the deviance of a space curve from being the plane of the curvature and the turnaround of the binormal vector \mathbf{b} at each point, which

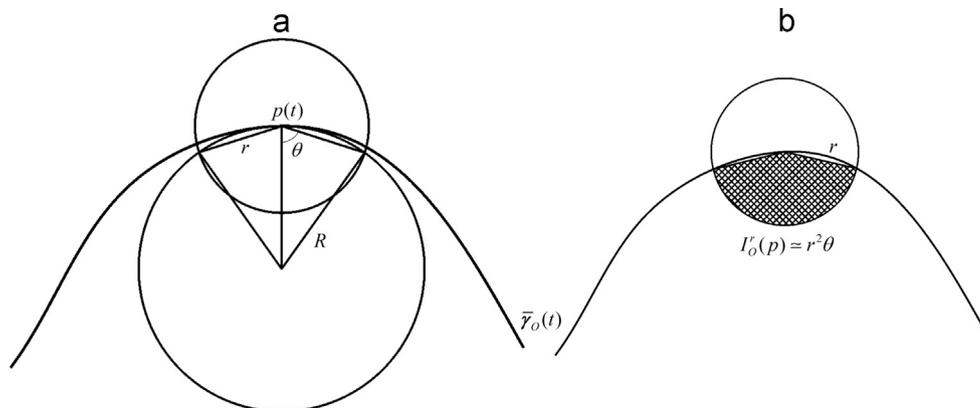


Fig. 6. Approximation of the area integral invariant on the osculating plane at a point. (a) The definition of θ . (b) Approximation of $I^r_O(p)$.

thus can be locally evaluated by projecting the space curve onto the rectifying plane. Then, the connection between the torsion at a point p and the area integral invariants $I_R^r(p)$ on the rectifying plane is exactly in the same formula as the approximation of $I_O^r(p)$ using the curvature. We thus have the result

$$I_R^r(p) \approx r^2 \cos^{-1} \left(\frac{1}{2} r \tau(p) \right) \tag{15}$$

3.4. Estimation of area integral invariants

As we claimed, integral invariants we propose do not involve high-order derivative operations. In practical applications, the distance integral invariant can be computed directly according to Eq. (5). However, the computation of the area integral invariants involves higher-order derivatives that cannot be computed directly for discrete trajectories in practical applications. In this section, the problem of how to estimate the area integral invariants is investigated. In practical applications, motion trajectories are some discrete temporal sequences sampled from the visual sensors in the presence of noise. For a discrete noisy motion trajectory, it is difficult to accurately estimate the projection planes along the motion trajectory, which immediately can compromise the accuracy and robustness of the area integral invariants even under small-scale perturbations. It is observed from Eqs. (10) and (11) that obtaining an instant Frenet–Serret frame at each point of a space trajectory is a key problem to calculate the area integral invariants. Thus, we formulate an estimation method of the Frenet–Serret frame based on Maximal 3D Blurred Segment of noise curves (MBS) [28] studied thoroughly in discrete geometry. We employ MBS to decompose a discrete noisy trajectory into some consecutive overlapped minimally thin blurred segments via eliminating and bypassing those noise points. Next, we further obtain the corresponding left and right key points of the blurred segments close to each reference point so that the triple norm vectors $(\mathbf{t}, \mathbf{n}, \mathbf{b})$ of the Frenet–Serret frame at each reference point can be estimated by these non-collinear triple points. The projection planes are then formed straightforwardly based on the estimated Frenet–Serret frame at each point.

3.4.1. Maximal 3D blurred segment of width v

The notion of blurred segment relies on the basic concept of 3D discrete line [30]. We need to review briefly its concept defined as follows.

Definition 3. A 3D discrete line [30], denoted by $D_{3D}(a, b, c, \mu, \mu', \omega, \omega')$, with a main vector $(a, b, c) \in \mathbf{Z}^3$ and $a \geq b \geq c$, is defined as a set of points (x, y, z) from \mathbf{Z}^3 verifying

$$D \begin{cases} \mu \leq cx - az < \mu + \omega \\ \mu' \leq bx - ay < \mu' + \omega' \end{cases} \tag{16}$$

with $\mu, \mu', \omega, \omega' \in \mathbf{Z}$, and ω and ω' are called the arithmetical width of D . According to Definition 3, the 3D discrete line is defined here to construct an envelope consisting of the upper and lower bounding lines on a segment of a discrete curve, also denoted as γ here, in limited width ω and ω' on the coordinate planes OXZ and OXY , respectively, as shown in Fig. 7 that illustrates a 2D projection version. Let $\gamma(i, j)$ be a set of successive points of a discrete curve γ indexing from i to j . The optimal bounding lines of $\gamma(i, j)$ with a minimal width are found by the thickness estimation of the convex hull of this set of points [29,31] on OXY and OXZ planes, respectively.

According to the definition of discrete line, a known 3D discrete curve can be segmented into a number of 3D discrete lines with a width v that is to control the segmentation level. Generally, a sequence of points $\gamma(i, j)$ from a 3D discrete curve γ is defined as a **3D blurred segment of width v** if there is an optimal discrete line enveloping $\gamma(i, j)$, named $D_{3D}(a, b, c, \mu, \mu', \omega, \omega')$, such that $(\omega' - 1/\max(|a|, |b|)) \leq v$ in the plane OXY and

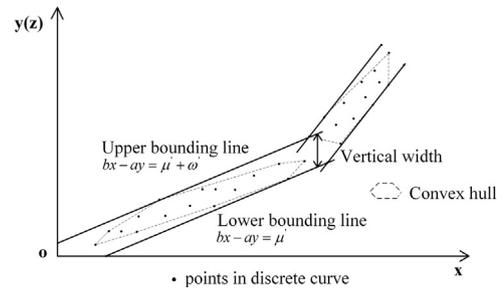


Fig. 7. Optimal bounding line D of two successive blurred segments of a discrete curve in the OXY plane.

$(\omega - 1/\max(|a|, |c|)) \leq v$ in the plane OXZ , which is denoted by $BS(i, j, v)$ and otherwise $\sim BS(i, j, v)$. The algorithm for recognizing the blurred segments of a discrete curve introduced in [28,30] allows us to work with a total complexity of $O(n \log^2 n)$ thanks to the dynamic thickness estimation of convex hulls in the discrete curve [31].

Furthermore, Nguyen et al. [28] and Faure et al. [29] proposed the notion of **Maximal Blurred Segment of width v (MBS)**. It is a 3D Maximal Blurred Segment of width v denoted as $MBS(i, j, v)$ if $BS(i, j, v)$, $\sim BS(i, j + 1, v)$ and $\sim BS(i - 1, j, v)$, i.e., an MBS of a discrete curve cannot be extended either at the right side or at the left side. Let us then consider decomposing a discrete space curve into a sequence of intercrossed Maximal 3D Blurred Segments (MBSSs) of width v with m length

$$MBS_v(\gamma) = \{MBS(B_1, E_1, v), MBS(B_2, E_2, v), \dots, MBS(B_m, E_m, v)\}, \tag{17}$$

with $B_1 < B_2 < \dots < B_m$ and $E_1 < E_2 < \dots < E_m$. $B_i, E_i | i \in [1, m]$ here denotes the beginning and ending positions of each Maximal Blurred Segment of the discrete curve.

3.4.2. Estimation of projection planes

Given a sequence of MBSs of a discrete trajectory, to estimate the Frenet–Serret frames along the discrete trajectory we first denote the estimated key points with $\gamma(B_i), \gamma(E_i) | i \in [1, m]$ as shown in Fig. 8. Obviously the MBSS, a sequence of estimated intercross discrete lines that connect the beginning and ending key points of each MBS, can approximate the real trajectory via bypassing some points that are not important to the trajectory shape. Inspired by this fact, let $\gamma(k)$ be the k th point of a 3D discrete trajectory γ , and $R(k), L(k) | k = 1 \dots N$ record a sequence of the positions of estimated right nearest key points and left nearest key points at the reference point $\gamma(k)$ such that $L(k) < k < R(k)$. Each triple point, $\{\gamma(R(k)), \gamma(k), \gamma(L(k)) | k = 1 \dots N\}$, is assumed to be not collinear as shown in Fig. 9. We then approximate the osculating circle at $\gamma(k)$ using the circumcircle of the triangle bounded by the triple points $\gamma(R(k)), \gamma(k), \gamma(L(k))$ as shown in Fig. 9. Therefore, the center of the circumcircle, denoted by $C(k)$, can be computed according to Heron’s formula.

Then, we define the unit norm vector at the k th point of the discrete trajectory by $\mathbf{n}_k = \gamma(k)C(k) / |\gamma(k)C(k)|$. The binormal vector \mathbf{b}_k is the unit vector that is perpendicular to the plane of the osculating circle. Then unit tangent vector \mathbf{t}_k is obtained straightforwardly: $\mathbf{t}_k = \mathbf{n}_k \wedge \mathbf{b}_k$. Now, Frenet–Serret frames associated with each point can be formed, and the triple projection planes are spanned by the triple norm vectors.

3.4.3. Approximation of area integral invariants

Given a sequence of estimated projection planes for a space discrete trajectory, we can approximate the area integral invariants at all points of the trajectory by numerical integration. At the k th reference point, we first project the space discrete trajectory onto the corresponding osculating and rectifying planes, forming the projected trajectory intersected by a circle with radius r at the k th point on either the osculating or the rectifying plane. Furthermore,

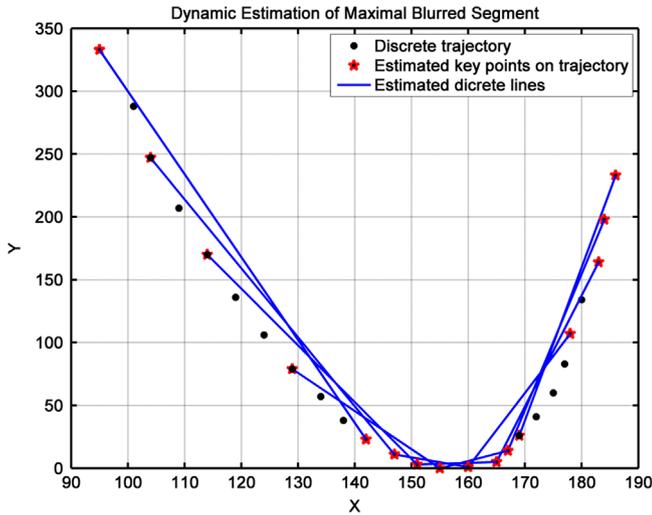


Fig. 8. Dynamic estimation of Maximal Blurred Segments for the 2D version (OXY plane) of a space discrete trajectory, where solid dots denote the points on the real discrete trajectory, stars denote the estimated key points on the trajectory and each blue line connects two key points in a maximal blurred segment.

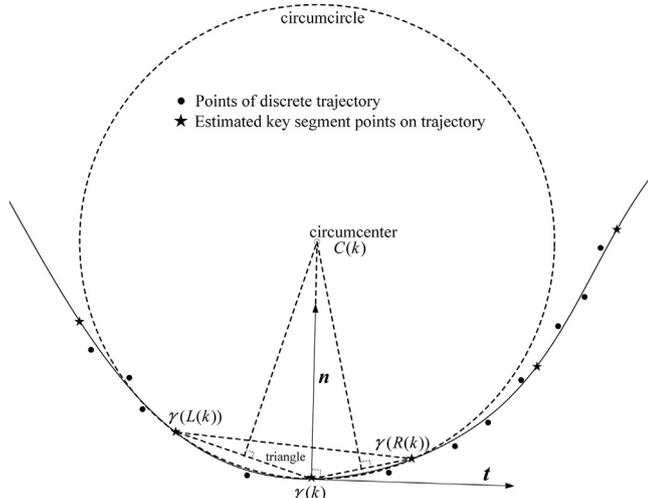


Fig. 9. Estimation of the normal and tangent vectors at a reference point of a trajectory.

we discretize the intersected trajectory by sampling it at small-scale grid resolution on the corresponding projection plane to approximate the intersected area using numerical integration.

The overall algorithms for approximating the integral invariants are described in Table 1 with Matlab pseudocodes in which the recognition of blurred segments refers to that in [28,30]. To test our numerical approximation, we consider an analytical space trajectory $\gamma(t) = (\cos t, \sin t, t)$ which is shown in Fig. 10(a). The analytical results of its differential invariants are given by theoretically computing

$$k(t) = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3} = \frac{1}{2}, \quad \tau(t) = \frac{(\gamma'(t) \times \gamma''(t)) \cdot \gamma'''(t)}{\|\gamma'(t) \times \gamma''(t)\|} = \frac{1}{2}.$$

The results show that this space trajectory moves along a fixed orientation with constant curvatures and torsions. Considering the relation of the area integral invariants to curvature and torsion, the integral invariants on two projection planes are supposed to be constants, consistent with the calculated differential invariants. We approximate the integral invariants via the proposed estimation formula, and compute the differential invariants by the approximation method in [27], respectively, with the discrete partition $\Delta t = 0.1$.

The results of the original trajectory are shown in Fig. 10(b) with an expected consistency with the theoretical results. Nevertheless, from Fig. 10(c) severe disturbances can be observed in the differential invariants when adding noise to the original trajectory.

It is worthy to note that for most of the relevant descriptors, the computational complexity tends inevitably to be high as the dimension of shapes rises to be higher [40]. Computing the descriptions for a trajectory in 3D space is far more complex than those in 2D space [25,27]. To describe a trajectory or shape in 3D space, there is a major bottleneck that is how to estimate the real kinematic properties of a moving particle along the trajectory. In differential geometry, they can be formulated by the Frenet–Serret frame. The descriptions for a 2D trajectory or shape lying on a consistent plane can be computed by performing discrete convolution [14,25,38] on an image or boundary without extra computations. In this paper, on the one hand we use the convex hull algorithm [31] to incrementally determine the width of blurred segments to reduce the computational cost, which makes the real-time computation of the integral invariants possible. On the other hand, in trajectory matching and recognition using the integral invariants, we have achieved superiority compared with other relevant descriptors in different situations including noise, occlusions, and group transformations, as demonstrated in the experiments (Section 5).

3.5. Occlusion invariants

Occlusion is a common phenomenon in motion trajectories due to some temporary failures of visual tracking and self-occlusions. A useful property of our integral invariants is that they can be applied to match trajectories despite occlusions, due to their computational locality. The partial integral invariants of non-occluded portion of a full trajectory can still be generated invariantly, which makes them well-suited for trajectory matching under occlusions. They are defined as occlusion invariants that enable a trajectory to be represented under reasonable occlusions. Note that heavy occlusions, however, may result in ambiguities in trajectory matching and recognition for similar trajectories.

An example of some occluded trajectories and their occlusion invariants are shown in Figs. 11 and 12, respectively. The invariants in Fig. 12(b and c) of the occluded instances can be completely localized in shape within the full invariants of the original trajectory in Fig. 12 (a), which enables the occluded trajectories to be matched and recognized from the partial matching of their occlusion invariants.

4. Similarity measure

In this section, we define a similarity measure between a pair of motion trajectories. A distance function between their integral invariants is defined to measure the similarity between them to match, cluster, and classify motion trajectories. Thanks to their invariance, the distance inherits invariant properties that are not affected by group transformations on the trajectories. Also, the robustness of the integral invariants makes the distance less sensitive to noise in practical scenarios. Moreover, we want the similarity insensitive to non-linear variations between them including different sampling rates, unequal lengths, and occlusions to trajectories. Therefore, we employ the DTW (Dynamic Time Warping) algorithm [44] to handle these variations.

Given a pair of motion trajectories, their integral invariants are represented by I_A and I_B with M and N length, respectively, defined in Eq. (12). To ultimately obtain the optimized distance, the DTW algorithm is performed based on dynamic programming to determine an optimal time warping path that could minimize the total distance $D(M, N)$ between two integral invariants, described as

Table 1

Algorithms 1–2 of approximation of area integral invariants.

Algorithm 1: Approximation of area integral invariants for a space discrete trajectory.**Input data:** a space discrete trajectory γ of n points, width v of the segmentation, radius r of the circle kernel for local area integral invariants;**Result:** I^r a sequence of area integral invariants with radius r ;

1. Employ the Maximal 3D Blurred Segment algorithm [28] to decompose the trajectory γ into a sequence of maximal blurred segments of width v , and obtain a sequence of the resulting segments: MBS_v ;
2. Build $MBS_v = \{MBS(B_i, E_i, v)\}$;
3. $I^r = \text{ones}(n, 2) * 0$;
4. $m = \text{length}(MBS_v)$; $E_{-1} = -1$; $B_m = n$;
5. for $i = 0$ to $m - 1$
 - for $k = E_{i-1} + 1$ to E_i do $L(k) = B_i$;
 - for $k = B_i$ to $B_{i+1} - 1$ do $R(k) = E_i$;
- end
6. for $k = 2$ to $n - 2$
 - call the **Algorithm 2** to get the $I^r(k)$ at k th point by inputting radius r , triple set $\{\gamma(L(k)), \gamma(k), \gamma(R(k))\}$; and the trajectory γ ;
7. end

Algorithm 2: Area Integral invariants approximation at k^{th} point**Input data:** a space discrete trajectory γ of n points: $n \times 3$ matrix, radius r of the circle kernel, the triple points $\{\gamma(L(k)), \gamma(k), \gamma(R(k))\}$;**Result:** $I^r(k)$, the area integral invariants with radius r at k^{th} point;

1. $I^r(k) = [0, 0]$; $grid = 2 * \pi / 1000$; $p = \gamma(k)$;
2. $C_k = \text{circumcenter}(\gamma(R(k)), \gamma(k), \gamma(L(k)))$; $\mathbf{n} = C_k - \gamma(k) / \text{norm}(C_k - \gamma(k))$; $\mathbf{b} = \text{cross}(\gamma(k) - \gamma(L(k)), \gamma(R(k)) - \gamma(k))$;
- $\mathbf{t} = \text{cross}(\mathbf{n}, \mathbf{b})$; % get the triple norm vectors by cross product of vectors;
3. $P_R(k) = \text{span}(\mathbf{b}, \mathbf{t})$; $P_O(k) = \text{span}(\mathbf{t}, \mathbf{n})$;
4. $\mathbf{D} = -(\mathbf{b} \cdot p)$; $t = \frac{\mathbf{b} \cdot \mathbf{D}}{\|\mathbf{b}\|}$; $\bar{\gamma}_O = \gamma - t \cdot \mathbf{b}$;
5. $\mathbf{D} = -(\mathbf{n} \cdot p)$; $t = \frac{\mathbf{n} \cdot \mathbf{D}}{\|\mathbf{n}\|}$; $\bar{\gamma}_R = \gamma - t \cdot \mathbf{n}$; % Projecting the trajectory onto the osculating and rectifying plane;
6. $\zeta_O = B_r(p, \bar{\gamma}_O)$; $\zeta_R = B_r(p, \bar{\gamma}_R)$; % Get the intersected trajectories ζ_O and ζ_R which are the intersection between the circle kernel centered at point $\gamma(k)$ on the projection planes $P_O(k)$ and $P_R(k)$, and the projected curve $\bar{\gamma}_O$ and $\bar{\gamma}_R$ respectively.
7. Discretize the integral intersected trajectories ζ_O and ζ_R with g grid resolution;
8. $N = \text{length}(\zeta_O)$; $M = \text{length}(\zeta_R)$;
9. $I^r_O(k) = \sum_{i=1}^N \left\{ \sqrt{r^2 - [\zeta_O(i) - p] \cdot \mathbf{t}}^2 - (\zeta_O(i) - p) \cdot \mathbf{n} \right\} * g$; $I^r_R(k) = \sum_{i=1}^M \left\{ \sqrt{r^2 - [\zeta_R(i) - p] \cdot \mathbf{t}}^2 - (\zeta_R(i) - p) \cdot \mathbf{b} \right\} * g$;
10. $I^r(k) = [I^r_O(k), I^r_R(k)]$;

follows:

$$D(m, n) = \min \{D(m-1, n), D(m, n-1), D(m-1, n-1)\} + d(m, n), \quad (18)$$

where $\{d(m, n) | m = 1 \dots M; n = 1 \dots N\} \in \mathbf{R}^{M \times N}$ denotes the local distance between samples m, n from two integral invariants, respectively, and it is defined as

$$d(m, n) = \Delta F = \frac{\Delta I^{m,n}}{S_I^{m,n}}, \quad (19)$$

where each specific term is defined to be the l_1 norm difference as follows:

$$\Delta I^{m,n} = \|I_A^r(m) - I_B^r(n)\|_1, \quad S_I^{m,n} = \| \{I_A^r(m), I_B^r(n)\} \|_1 \quad (20)$$

In Eq. (20), $S_I^{m,n}$ is to normalize the local distance to eliminate the effects of different kernel scales between integral invariants. Following the above definition, the total DTW distance $D(M, N)$ is computed recursively from $D(1, 1)$ to $D(M, N)$ by dynamically searching for a best alignment path using an optimized algorithm with slope constraints [44]. Then, the matching of a pair of motion trajectories is evaluated in terms of the total distance $D(M, N)$, which relatively reflects the similarity between two trajectories. An example of the results of dynamically finding an optimal path in two matching trajectories is shown in Fig. 13.

5. Experiments

This section presents experiments that show the locality, noise robustness, and invariance properties of the integral invariants, resulting in an effective and robust description for motion

trajectory representation, matching, and recognition. We conduct the following experiments on several datasets which contain a large variety of motion types and variations:

- **Trajectory matching:** we first evaluate the effectiveness and robustness of the integral invariants by matching a number of pairs of trajectories extracted from the HDM05 [34] and Berkeley MHAD [36] motion capture datasets. The motion capture datasets use a skeleton model of several joints of the body to represent the human motion dynamics, where a temporal sequence of 3D joint positions is recorded for each joint. We simulate a series of group transformations, noise, and occlusions on the extracted trajectories to match with each other to examine the noise robustness, invariance, and occlusions handling of the integral invariants, when compared to other descriptors.
- **Sign language recognition:** as a sign can be abstracted as hand motion trajectories, we show the effectiveness of the integral invariants applied to sign language recognition via some recognition benchmarks. ASL (Australian Sign Language) dataset [50] is employed here.

For comparison, we compare experimental results via the integral invariants (II) with other descriptor-based results including the distance integral invariant (DII), Fourier descriptor (FD) [23], and differential invariants (DI) [30,32] in terms of those mentioned evaluations, where the DTW (II, DII and DI) or Euclidean distance (FD) measures the trajectory similarity in trajectory matching and recognition based on a k -NN (k nearest neighbors algorithm).

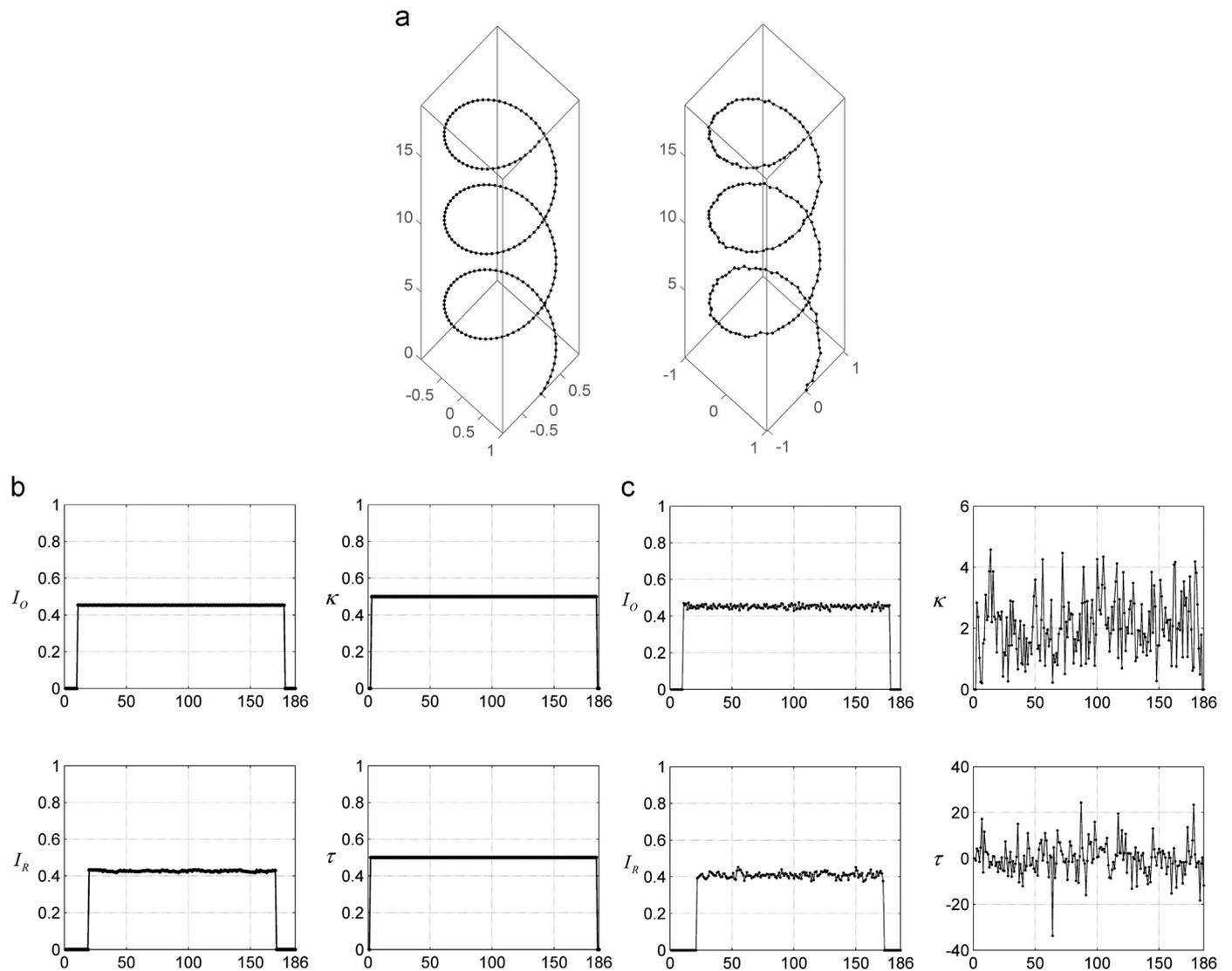


Fig. 10. Approximation results of integral invariants and differential invariants of an analytic trajectory and its noisy version. (a) Analytical curve $\gamma(t) = (\cos t, \sin t, t)$ for $0 \leq t \leq 6\pi$ and its noisy version (Gaussian white noise with standard deviation 0.01). (b) Results of the original version. (c) Results of the noisy version.

5.1. Parameters setting

For approximating the integral invariants in Algorithm 1, a trajectory is digitalized with 1000 grids. We then perform the Maximal Blurred Segment with width $v=8$ on all datasets. In addition, the radius of the ball kernel at $\gamma(k)$ is determined by the average of inter-distances between $\gamma(k)$ and its n th nearest points, $r = (|\gamma(k) - \gamma(k-n)| + |\gamma(k) - \gamma(k+n)|)/2$. In other words, the radius of the ball kernel at each point depends on the distributions of the neighboring points at each point instead of setting it empirically. In the experiments, we set $n=2$. In the same way, we obtain the average distance of the nearest 5 points at $\gamma(k)$ for the distance integral invariant.

For Fourier descriptor, the following Fourier transformation for 1D data is used to describe a motion trajectory by applying it to each dimensional data

$$X_l = \sum_{t=1}^N x(t) e^{-j2\pi t(l-1)/N}, \quad 1 \leq l \leq N, \quad (21)$$

As a result, we obtain a sequence of coefficients associated with each dimension: $F_l = \{X_l, Y_l, Z_l\}$. The Fourier descriptor is further

normalized [23] for keeping invariant to translation, start point, and scaling, provided that the first coefficient is ignored. The rest of the coefficients are scaled by the second coefficient. It should be noted that as suggested in [23] Fourier descriptor is not able to keep complete rotation invariance that is achieved by simply taking the magnitude of each Fourier coefficient.

5.2. Trajectory matching

For trajectory matching, the test is performed on two popular motion capture datasets, HDM05 dataset [34] and Berkeley MHAD dataset [36]. HDM05 dataset consists of around 100 motion classes performed by 5 different actors. Most of these classes contain 10–50 different realizations for each motion amounting to 1457 smaller motion clips. The duration of motion sequences ranges from 56 to 901 frames with the frequency of 120 Hz. Berkeley MHAD dataset contains 11 motion categories performed by 12 subjects with 5 repetitions per motion using active optical motion capture system at a 480 Hz sampling rate, yielding about 656 motion sequences (several erroneous motions are eliminated). The motion types contain: *jump, jumping jacks, bend, punch, wave one hand, wave two hands, clap, throw, sit down, stand up, and sit down/standup.*

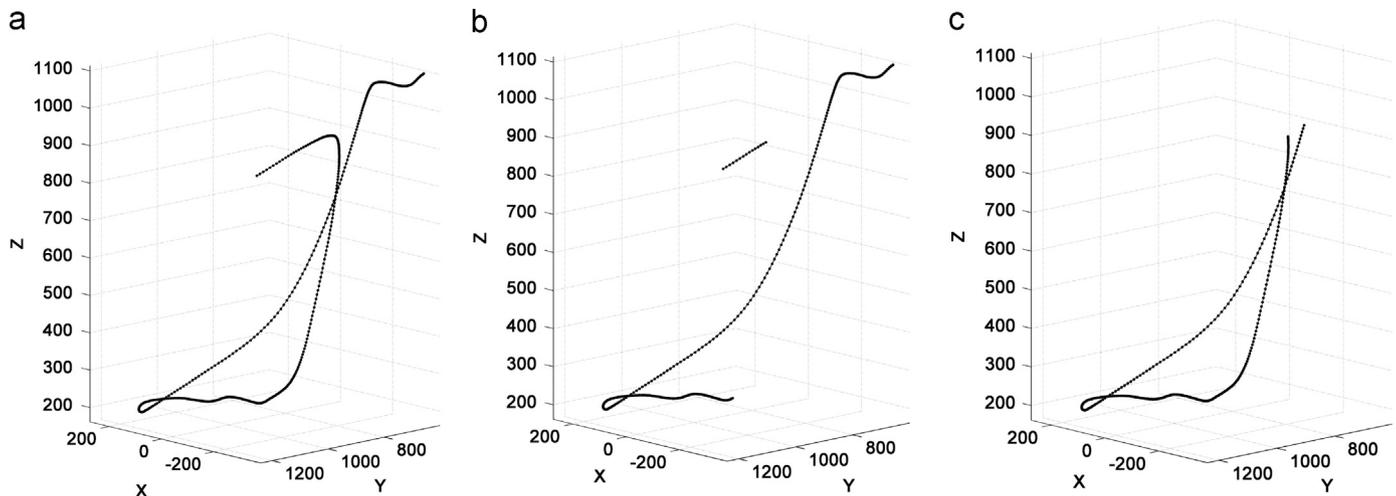


Fig. 11. Two occluded cases of (a) the original trajectory; (b) frames 14–148 occluded; and (c) frames 1–52 plus frames 360–432 occluded.

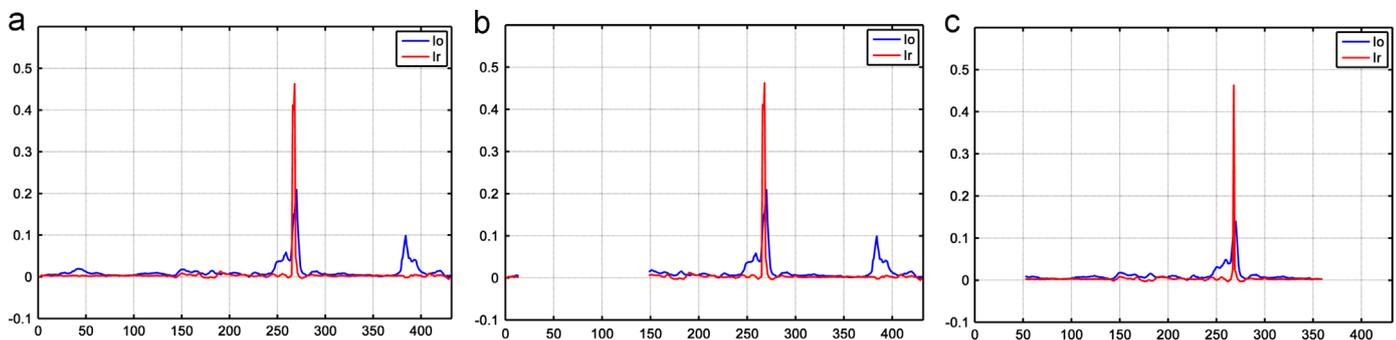


Fig. 12. The integral invariants I_O and I_R of the trajectories shown in Fig. 11 accordingly.

While each motion sequence in Berkeley MHAD contains multiple repetitions, we here collect the experimental dataset consisting of shorter motion clips by splitting each of them into individual sequences with one repetition using labels given by Ofli [36].

Firstly, for the test on the HDM05 dataset, we collect 16 motion classes including *deposit floor*, *elbow to knee*, *gab high*, *hop both leg*, *jog left*, *jump down*, *jumping jack*, *kick forward*, *lie down floor*, *rotate arm backward*, *sit down chair*, *sneak*, *squat*, *stand up*, *throw basketball*, and *throw*. Among them, 16 original motion trajectories of the right hand are randomly extracted each time from 16 motion classes, respectively. They are to match their transformation versions, each of which is obtained by a series of actions including first downsampling or upsampling randomly the original trajectory to a new one with 0–50% more or less length than the original frames, adding Gaussian noise (normalized standard deviation δ), then rotating 30° and 45° by x and z axes, respectively, then translating 200 mm and 500 mm along x and y directions, respectively, and finally scaling by 0.5 factor. Examples for a group of original and transformation versions of motion trajectories are shown in Fig. 14. This extracting and matching procedure is run 50 times getting the average matching accuracies as shown in Fig. 15 compared with other descriptors. Secondly, we also carry out the same matching test on the Berkeley MHAD dataset. In a similar manner, 11 pairs of motion trajectories of the right hand, corresponding to the respective 11 motion classes, are randomly extracted to match with their transformation versions, when compared to other descriptors.

The matching performance is evaluated by the matching accuracy that is the percentage of correct matching between pairs of trajectories. The average matching accuracies with the additive noise specified by δ are obtained by running the matching test 50 times

between original trajectories and transformed trajectories extracted from both datasets, when compared with other descriptors. Fig. 15 shows a plot of the matching accuracies on both the HDM05 and MHAD datasets as noise increases with $\delta = [0, 0.2]$. Not surprisingly, it can be observed that the matching accuracies obtained via all the descriptors decrease as the additive noise increases. The matching accuracies via the differential invariants (**DI**) and Fourier descriptor (**FD**) decrease more drastically than other descriptors as δ increases. The matching even fails via **FD** without noise added as shown in Fig. 15 due to its limited rotation invariance. Using integral invariants (**II** and **DII**), we can obtain some large improvements in matching performance, where the decreases of the matching accuracies are most insignificant, insensitive than other descriptors-based matching as noise increases. Especially for the trajectory matching on the MHAD dataset, the matching accuracies via **II** and **DII** decrease slightly. As shown in the first matching experiment, both the integral invariants (**II** and **DII**) are the most insensitive to noise and group transformations compared with other descriptors.

To quantitatively evaluate the similarity distance in the presence of additive noise, we compare the average distance matrixes between the original and noisy transformed trajectories, extracted from the HDM05 dataset, via the integral invariants (**II**) and differential invariants (**DI**) in Fig. 16. The distance matrix computed using **II** has lower distances on the diagonal as expected in the dataset, whereas the distance matrix computed using **DI** completely lacks the expected lower distances on the diagonal. Comparing this with the **DI**-based distance matrix, the robustness of the integral invariants to noise can be observed.

One of the advantages using the integral invariants is its potential to deal with occlusions. Occlusions are simulated here

by randomly breaking each transformed trajectory into two separated parts without additive noise when the trajectory matching is carried out. We can observe the matching results in Fig. 17 as

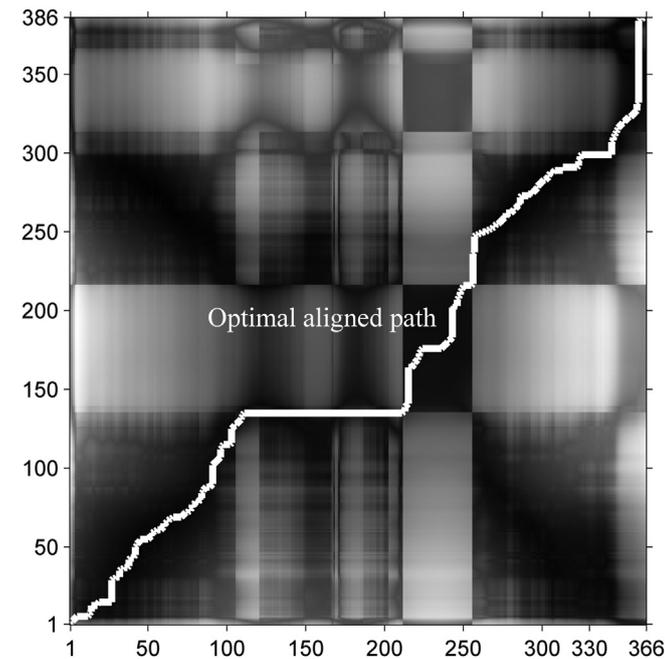


Fig. 13. Optimal path through a distance matrix. The optimal path is found by dynamically wrapping the time axis between integral invariants (vector features) of a pair of motion trajectories. The gray level indicates the similarity between two trajectories; darker shade indicates higher similarity.

the ratio of occlusion in trajectories, extracted from the HDM05 and Berkeley MHAD datasets, respectively, increases from 0% to 40% compared with other descriptors. The results show that the matching performance drops, based on both the integral invariants and differential invariants (**II** and **DI**), are relatively small thanks to their locality. The matching performance based on the distance integral invariant (**DII**) turns out to be worse. This is probably because the distance integral invariant is not sufficient to represent a trajectory especially under occlusions due to its non-uniqueness as suggested in Section 3.2. Not surprisingly, the matching accuracies via the Fourier descriptor (**FD**) decrease drastically as the ratio of simulated occlusions increases, especially on the HDM05, since its computation depends on the whole information of motion trajectories. It should be noted that to test how the matching performance is influenced only by occlusions using **FD**, the transformed trajectories here are only translated without rotation in this test. In summary, in occlusion handling, **II** is proved to be the best in matching occluded trajectories.

5.3. Sign language recognition

In motion recognition, we examine the effectiveness of the integral invariants in classifying motion labels on the Austral Sign Language (ASL) dataset. The ASL dataset consists of 2565 samples of Auslan signs, where 27 examples of each of 95 sign classes are captured from a native signer with high-quality data and each sample of Auslan signs is performed by moving the right hand and left hand simultaneously in 3D space. In this test, we only employ the root trajectory [20], the average of the right- and left-hand trajectories, to represent a sign. There is an instance of the sign word “make” as shown in Fig. 18. For test, we used the 1-NN classifier based on the

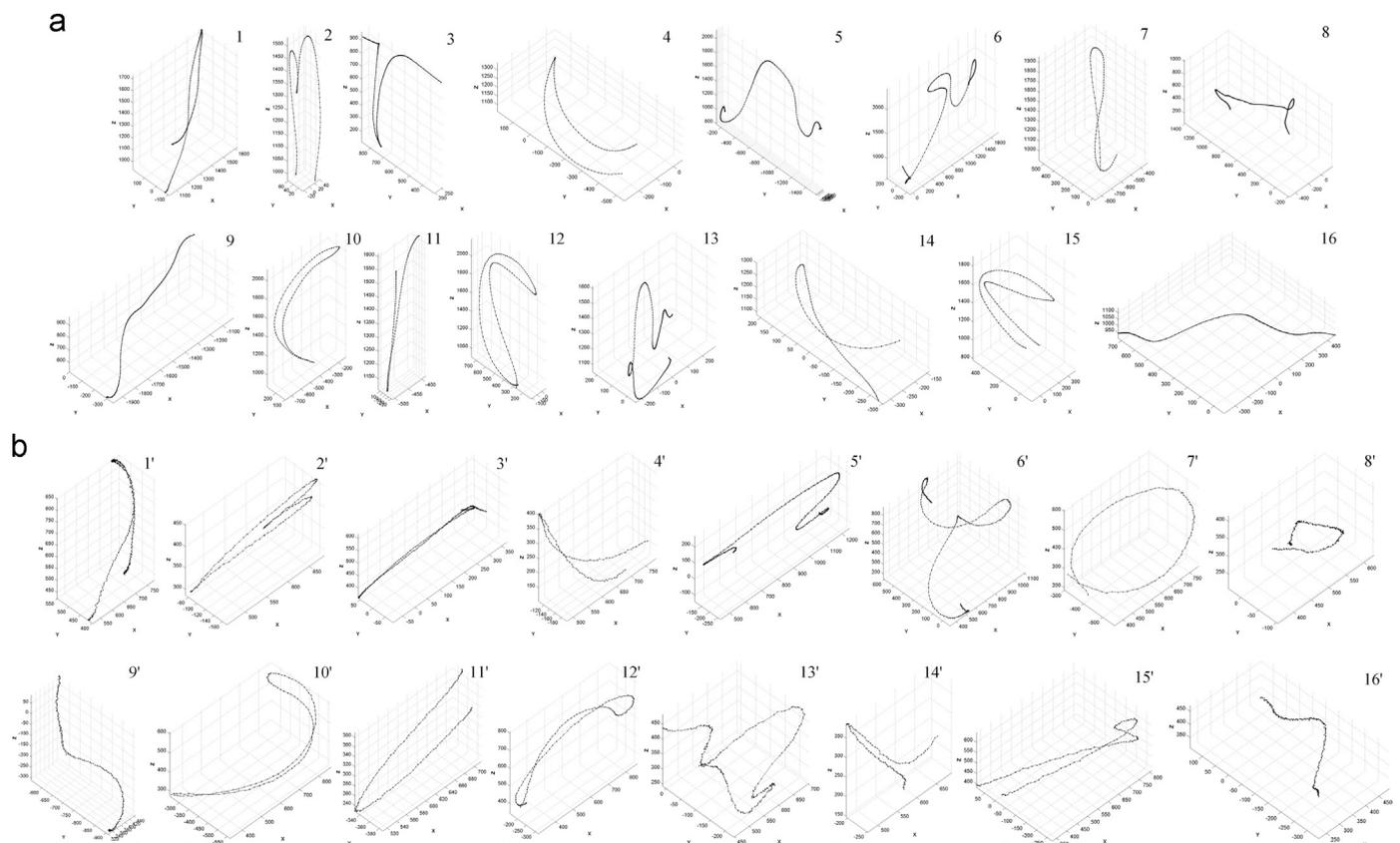


Fig. 14. Examples for a group of extracted 16 motion trajectories of the right hand from HDM05 motion dataset with two versions, (a) original versions, and (b) transformation versions without simulated occlusions.

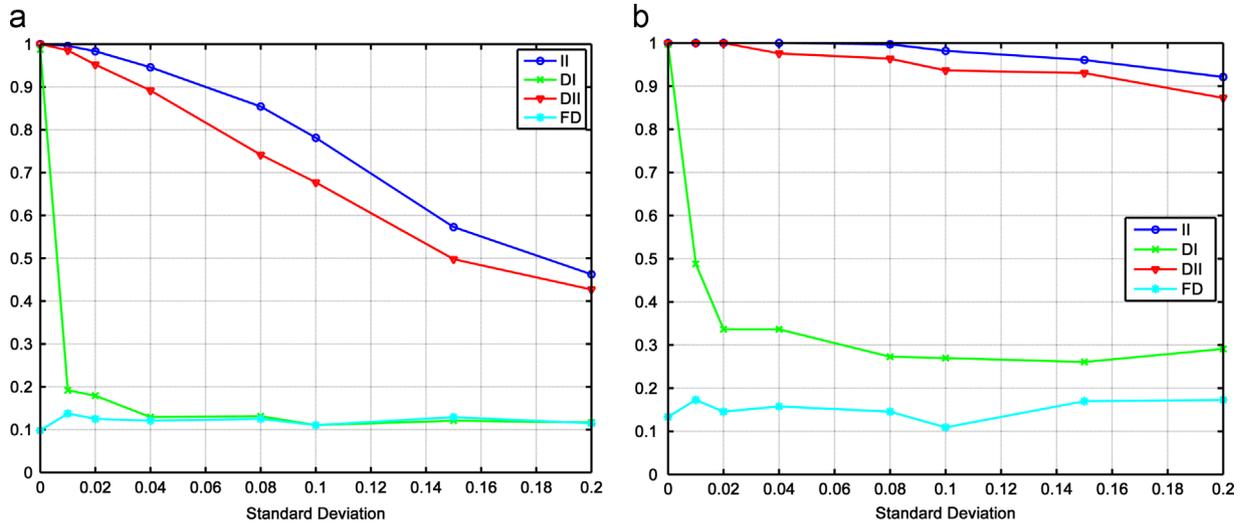


Fig. 15. Average matching accuracy as the additive noise in trajectories increases. (a) Average matching accuracy on the HDM05 dataset, and (b) on the Berkeley MHAD dataset. While the matching accuracies based on other descriptors strongly decrease as noise increases, the matching accuracies based on the integral invariants and the distance integral invariants are much more insensitive to noise.

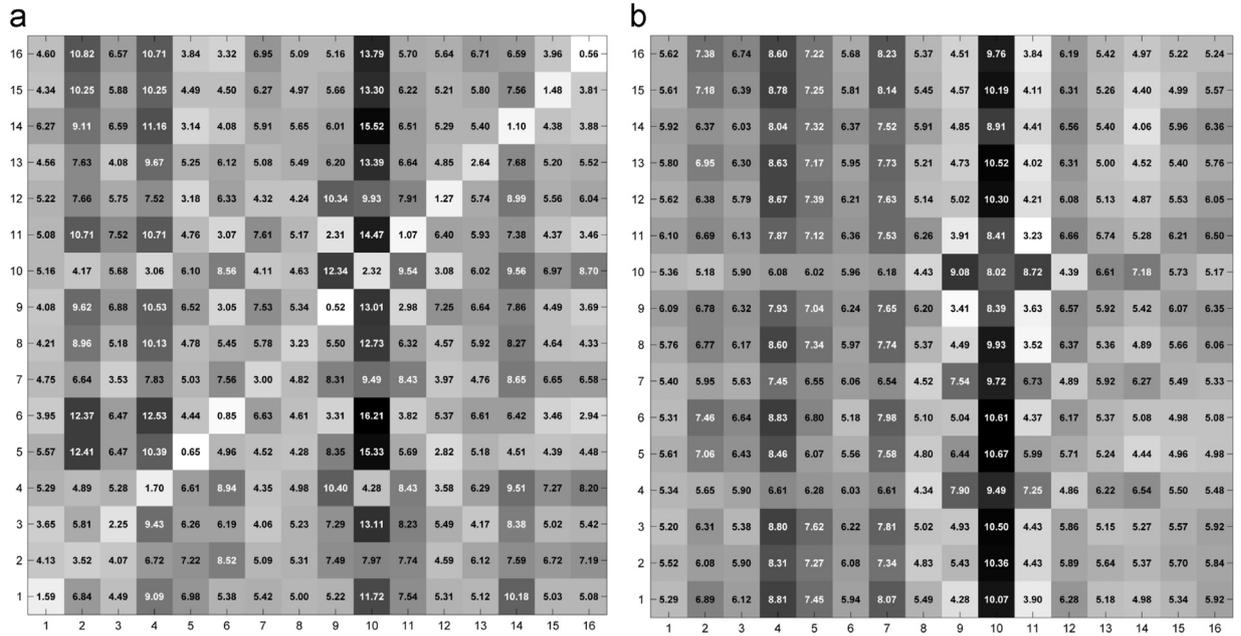


Fig. 16. Distance matrixes between noisy transformed trajectories with $\delta = 0.08$ (across bottom) and original trajectories (across left side) via (a) the integral invariants, and (b) differential invariants. Lighter shades indicate smaller distances.

DTW inter-descriptor distance as defined in Eq. (18). Each time 16 classes of samples are randomly selected in the ASL dataset to run the recognition test where half the samples of each class are for training and the other half are for testing. We repeated this test 50 times. In the same way, the other descriptor-based recognitions are also carried out to compare with our integral invariants.

The average recognition results are summarized in Table 2. The best recognition accuracy with 94.17% is obtained using the integral invariants (II) and the accuracy reduces to 93.85% when using the differential invariants (DI). Without additive noise, the recognition results show a consistency with the matching results in Section 5.2. The result based on the Fourier descriptor (FD) is much worse compared with DI and II. Compared with hand motions in matching experiment, sign languages are more complicated so that the lack of uniqueness of representation makes the

distance integral invariant (DII) ambiguous with the worst result in sign recognition.

6. Conclusion

In this paper, we propose a class of new integral invariants as an effective and robust description for motion trajectories to achieve effective and robust motion trajectory matching and recognition. Compared with differential invariants, the integral invariants involving integration along a motion trajectory have a smoothing effect and, therefore, are less sensitive to noise without preprocessing the motion trajectory. Also, regarding the estimation of the integral invariants, we can control both the scale of a kernel function and the width of blurred segments to achieve stability and robustness in

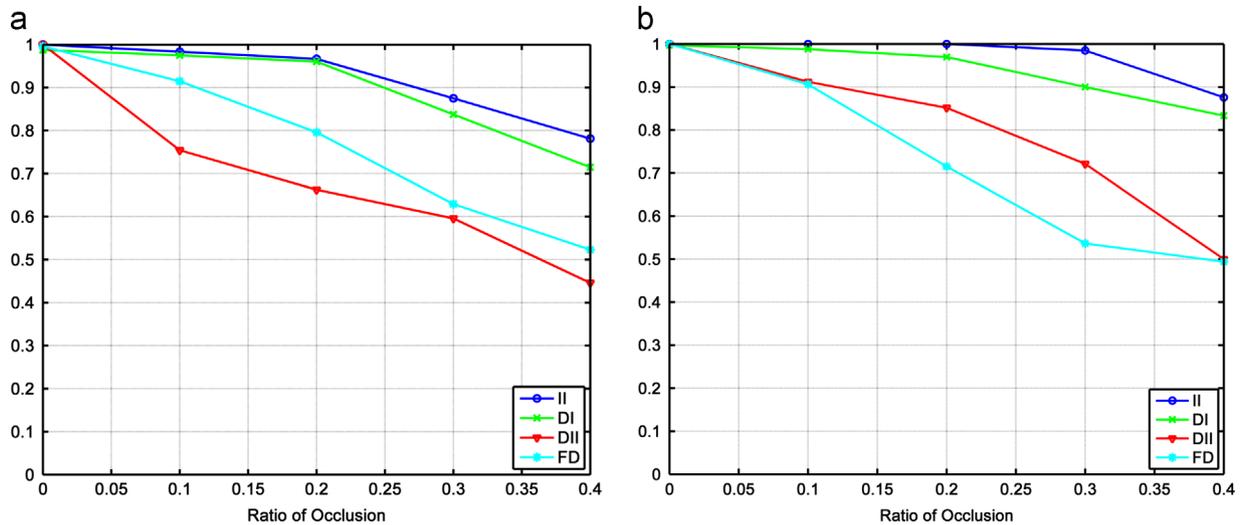


Fig. 17. Average matching accuracy as the ratio of occlusion increases. (a) Experimental results on the HDM05 dataset, and (b) experimental results on the Berkeley MHAD dataset. While the matching accuracies based on the Fourier descriptor and distance integral invariant strongly decrease as the ratio of occlusions increases, the matching accuracies based on the integral invariants and differential invariants are much more insensitive to occlusions.

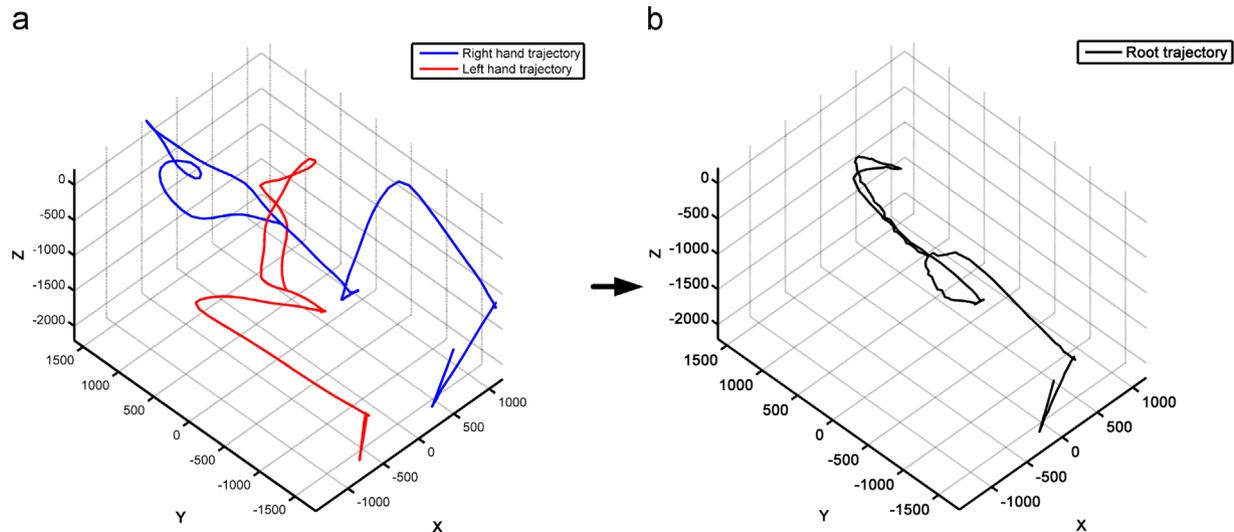


Fig. 18. Sign samples of the word "make" from the ASL dataset: (a) the right- and left-hand trajectories and (b) the root trajectory that is the average of the right- and left-hand trajectories.

the presence of noise. The computational locality of the integral invariants allows trajectory matching under occlusions, while some of the previous descriptors, such as the Fourier descriptor and moment invariants, are global descriptors where local changes can result in variations everywhere. Moreover, the invariance under Euclidean and similarity group transformations makes the integral invariants independent of observing viewpoints and viewing distances. On the experimental level, we show the claimed properties of the integral invariants including invariance, robustness to noise, and insensitivity to occlusions in trajectory matching and recognition compared with other descriptors.

Multiscale space methods [15,39] allow us to analyze motion features at multiple resolutions in a coarse-to-fine manner. Extending our integral invariants to multiscale representation is possible for tapping into different levels of motion features as we vary the scale of a kernel function in Definition 2, such as the radius of a ball kernel in this paper, which is our ongoing work now.

While the integral invariants are designed to describe space motion trajectories, they can also benefit invariant representation

in computer vision such as representing 3D contour shapes for object recognition and retrieval. Normally, both the Euclidean invariants and affine invariants are necessary for 2D shapes to address shape variations under Euclidean and affine transformations, for instance those shapes, captured from different viewpoints, are equivalent classes of each other. Nevertheless, in 3D space we argue that Euclidean invariants are sufficient to represent 3D shapes and trajectories in that the 3D information, tracked from stereo vision or other motion capture systems, would not generate affine transformations on them.

In practical applications, complex motions often involve multiple motion trajectories tracked from concurrent moving objects, such as articulated motions of the human body. As human motions can be represented in the form of 3D positions of joints in space over time, such joint trajectories can serve as effective features in motion retrieval [49] to accurately search similar motions from a large repository in games and animations. As addressed in our previous work [20], a group of multiple motion trajectories can be decomposed into a root trajectory and

Table 2

Recognition accuracy of 1-NN algorithm via integral invariants compared to other descriptors on ASL dataset.

Descriptors	Accuracy (%)
Fourier descriptor [23]	78.07
Differential invariants [32]	93.85
Distance integral invariant	24.43
Integral invariants	94.17

child trajectories. While the root trajectory represents the most discriminant one in multiple motion trajectories, the relative distributions of child trajectories with respect to the root trajectory are also a key clue to characterize motion contents. In this case, we can integrate both the integral invariants of the root trajectory and the relative distributions into one group to compose a hierarchical descriptor for multiple trajectories. The hierarchical descriptor is thus able to capture both the spatio-temporal distributions and key dynamic clues of multiple trajectories jointly thanks to the hierarchical structure [20]. Hence, the hierarchical descriptor for multiple trajectories can be used to retrieve human motion sequences. With our ongoing work, we are exploring the possibilities and potential of the hierarchical descriptor in motion retrieval.

A recognition system relies on two key factors, input features and recognition engines. While we have shown that the integral invariants are an effective representation for motion trajectories to be used in sign recognition, there is still room for improvement in the recognition efficiency for large-scale trajectories. Some learning-based algorithms, such as SVM and HMM, will be employed to boost the recognition efficiency by training these models using the integral invariants, which is our future work.

Conflict of interest

None declared.

Acknowledgment

This work was supported by the Research Grants Council of Hong Kong (Project no. CityU 118613), and the National Natural Science Foundation of China (No. 61273286) and the Center of Robotics and Automation at CityU.

References

- [1] M. Bennewitz, W. Burgard, G. Cielniak, S. Thrun, Learning motion patterns of people for compliant robot motion, *Int. J. Robot. Res.* 24 (1) (2005) 31–48.
- [2] S. Calinon, F. Guenter, A. Billard, On learning, representing, and generalizing a task in a humanoid robot, *IEEE Trans. Syst. Man Cybern. Part B Cybern.* 37 (2) (2007) 286–298.
- [3] J. Moldenhauer, I. Boesnach, T. Beth, V. Wank, K. Bos, Analysis of human motion for humanoid robots, in: *Proceedings of IEEE International Conference on Robotics and Automation*, Barcelona, 2005, pp. 311–316.
- [4] A. Psarrou, S. Gong, M. Walter, Recognition of human gestures and behavior based on motion trajectories, *Image Vis. Comput.* 20 (2002) 349–358.
- [5] Z. Wang, M. Lu, X. Yuan, J. Zhang, H. van de Wetering, Visual traffic jam analysis based on trajectory data, *IEEE Trans. Vis. Comput. Gr.* 19 (12) (2013) 2159–2168.
- [6] J. Min, R. Kasturi, Activity recognition based on multiple motion trajectories, in: *Proceedings of the 17th International Conference on Pattern Recognition*, Cambridge, UK, 2004, vol. 4, pp. 199–202.
- [7] S. Ali, A. Basharat, M. Shah, Chaotic invariants for human action recognition, in: *Proceedings of IEEE International Conference on Computer Vision*, Rio de Janeiro, Brazil, 2007, pp. 1–8.
- [8] M.H. Yang, N. Ahuja, M. Tabb, Extraction of 2D motion trajectories and its application to hand gesture recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (8) (2002) 1061–1074.
- [9] A. Oikonomopoulos, I. Patras, M. Pantic, N. Paragios, Trajectory-based representation of human actions, in: *Proceedings of International Conference on Artificial Intelligence for Human Computing*, Heidelberg, Germany, 2007, pp. 133–154.
- [10] C. Rao, A. Yilmaz, M. Shah, View-invariant representation and recognition of actions, *Int. J. Comput. Vis.* 50 (2) (2002) 203–226.
- [11] E. Bribiesca, A chain code for representing 3D curves, *Pattern Recognit.* 33 (5) (2000) 755–765.
- [12] M.C. Shin, L.V. Tsap, D.B. Goldgof, Gesture recognition using Bezier curves for visualization navigation from registered 3-D data, *Pattern Recognit.* 37 (5) (2004) 1011–1024.
- [13] B.T. Morris, M.M. Trivedi, Trajectory learning for activity understanding: unsupervised, multilevel, and long-term adaptive approach, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (11) (2011) 2287–2301.
- [14] S. Manay, D. Cremers, Byung-Woo Hong, A.J. Yezzi, S. Soatto, Integral invariants for shape matching, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (10) (2006) 1602–1617.
- [15] B. Hong, S. Soatto, Shape matching using multiscale integral invariants, *IEEE Trans. Pattern Anal. Mach. Intell.* (2015), <http://dx.doi.org/10.1109/TPAMI.2014.2342215> (in press).
- [16] G. Mori, S. Belongie, J. Malik, Efficient shape matching using shape contexts, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (11) (2005) 1832–1837.
- [17] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape context, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (4) (2002) 509–522.
- [18] F. Bashir, A. Khokhar, Curvature scale space based affine invariant trajectory retrieval, in: *Proceedings of IEEE International Multitopic Conference*, Lahore, Pakistan, 2004, pp. 20–25.
- [19] F.I. Bashir, A.A. Khokhar, D. Schonfeld, View-invariant motion trajectory-based activity classification and recognition, *Multimed. Syst.* 12 (1) (2006) 45–54.
- [20] Z. Shao, Y.F. Li, A new descriptor for multiple 3D motion trajectories recognition, in: *Proceedings of IEEE International Conference on Robotics and Automation*, Karlsruhe, 2013, pp. 4749–4754.
- [21] F. Mokhtarian, A.K. Mackworth, Scale-based description and recognition of planar curves and two-dimensional shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 8 (1) (1986) 34–43.
- [22] E. Bala, A.E. Cetin, Computationally efficient wavelet affine invariant functions for shape recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (8) (2004) 1095–1099.
- [23] P.R.G. Harding, T.J. Ellis, Recognizing hand gesture using Fourier descriptors, in: *Proceedings of International Conference on Pattern Recognition*, Cambridge, UK, vol. 3, 2004, pp. 286–289.
- [24] Y.u.e. Wang, E.K. Teoh, 2D Affine-invariant contour matching using B-spline model, *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (10) (2007) 1853–1858.
- [25] E. Calabi, P. Olver, C. Shakiban, A. Tannenbaum, S. Haker, Differential and numerically invariant signature curves applied to object recognition, *Int. J. Comput. Vis.* 26 (1998) 107–135.
- [26] Abraham Goetz, *Introduction to Differential Geometry*, Addison-Wesley, Massachusetts, 1970.
- [27] M. Boutin, Numerically invariant signature curves, *Int. J. Comput. Vis.* 40 (3) (2000) 235–248.
- [28] T.P. Nguyen, I. Debled-Rennesson, On the local properties of digital curves, *Int. J. Shape Model.* 14 (2) (2008) 105–125.
- [29] A. Faure, L. Buzer, F. Feschet, Tangential cover for thick digital curves, *Pattern Recognit.* 42 (10) (2009) 2279–2287.
- [30] I. Debled-Rennesson, F. Feschet, J. Rouyer-Degli, Optimal blurred segments decomposition of noisy shapes in linear time, *Comput. Gr.* 30 (1) (2006) 30–36.
- [31] L. Buzer, A simple algorithm for digital line recognition in the general case, *Pattern Recognit.* 40 (6) (2007) 1675–1684.
- [32] S.D. Wu, Y.F. Li, On signature invariants for effective motion trajectory recognition, *Int. J. Robot. Res.* 27 (8) (2008) 895–917.
- [33] S.D. Wu, Y.F. Li, Flexible signature descriptions for adaptive motion trajectory representation, perception and recognition, *Pattern Recognit.* 42 (1) (2009) 194–214.
- [34] M. Müller, T. Röder, M. Clausen, B. Eberhardt, B. Krüger, A. Weber, Documentation: Mocap database HDM05, Technical report CG-2007-2, Universität Bonn, 2007.
- [35] F. Ofli, R. Chaudhry, G. Kurillo, R. Vidal, R. Bajcsy, Berkeley MHAD: a comprehensive multimodal human action database, in: *Proceedings of the IEEE Workshop on Applications on Computer Vision*, Clearwater Beach, USA, 2013.
- [36] Ronald Alferetz, Yuan-Fang Wang, Geometric and illumination invariants for object, *IEEE Trans. Pattern Anal. Mach. Intell.* 21 (6) (1999) 505–536.
- [37] D. Mumford, J. Fogarty, F.C. Kirwan, *Geometric Invariance Theory*, 3rd edition, Springer-Verlag, New York, 1994.
- [38] F. Mokhtarian, A.K. Mackworth, A theory of multiscale, curvature-based shape representation for planar curves, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (8) (1992) 789–805.
- [39] H. Pottmann, J. Wallner, Q.X. Huang, Y.L. Yang, Integral invariants for robust geometry processing, *Comput. Aided Geom. Des.* 26 (1) (2009) 37–60.
- [40] C.E. Hann, M.S. Hickman, Projective curvature and integral invariants, *Acta Appl. Math.* 74 (2) (2002) 177–193.

- [42] J.L. Mundy, A. Zisserman, *Geometric Invariance in Computer Vision*, MIT, Cambridge, MA, 1992.
- [44] L. Rabiner, B. Juang, *Fundamentals of Speech Recognition*, Prentice-Hall, New Jersey, 1993.
- [47] H.L.i. Xu, 3-D curve moment invariants for curve recognition, *Lecture Notes in Control and Information Sciences*, Springer-Verlag, Berlin, Heidelberg (2006) 572–577.
- [48] C.-H. Teh, R.T. Chin, On image analysis by the methods of moments, *IEEE Trans. Pattern Anal. Mach. Intell.* 10 (4) (1988) 496–513.
- [49] W. Hu, G. Tian, X. Li, S. Maybank, An improved hierarchical Dirichlet process-Hidden Markov model and its application to trajectory modeling and retrieval, *Int. J. Comput. Vis.* 105 (3) (2013) 246–268.
- [50] U.C.I. KDD, ASL Archive, Available: <http://hdd.ics.uci.edu/databases/auslan2/auslan.html>.

Zhanpeng Shao received his B.S. and M.S. degrees in mechanical engineering from Xi'an University of Technology, Xi'an, China, in 2004 and 2007, respectively. From 2007 to 2011, he was an embedded system engineer in EVOIC Intelligent Technology Co., Ltd. Currently, he is a Ph.D. candidate in the Department of Mechanical and Biomedical Engineering at City University of Hong Kong. His research interests include pattern recognition, feature extraction, and robot vision.

Youfu Li received his B.S. and M.S. degrees in electrical engineering from Harbin Institute of Technology China. He obtained the Ph.D. degree in robotics from the Department of Engineering Science, University of Oxford, in 1993. From 1993 to 1995 he was a postdoctoral research staff in the Department of Computer Science at the University of Wales, Aberystwyth, UK. He joined City University of Hong Kong in 1995. His research interests include robot sensing, robot vision, 3D vision and visual tracking. He has served as an Associate Editor for the *IEEE Transactions on Automation Science and Engineering*, the *IEEE Robotics and Automation Magazine* and an Editor for *CEB, IEEE International Conference on Robotics and Automation*.