On Integral Invariants for Effective 3-D Motion Trajectory Matching and Recognition

Zhanpeng Shao and Youfu Li, *Senior Member, IEEE*

***Abstract*—Motion trajectories tracked from the motions of human, robots, and moving objects can provide an important clue for motion analysis, classification, and recognition. This paper defines some new integral invariants for a 3-D motion trajectory. Based on two typical kernel functions, we design two integral invariants, the distance and area integral invariants. The area integral invariants are estimated based on the blurred segment of noisy discrete curve to avoid the computation of high-order derivatives. Such integral invariants for a motion trajectory enjoy some desirable properties, such as computational locality, uniqueness of representation, and noise insensitivity. Moreover, our formulation allows the analysis of motion trajectories at a range of scales by varying the scale of kernel function. The features of motion trajectories can thus be perceived at multiscale levels in a coarse-to-fine manner. Finally, we define a distance function to measure the trajectory similarity to find similar trajectories. Through the experiments, we examine the robustness and effectiveness of the proposed integral invariants and find that they can capture the motion cues in trajectory matching and sign recognition satisfactorily.**

***Index Terms*—Integral invariant, maximal blurred segment (MBS), motion trajectory, similarity measure, trajectory matching.**

I. Introduction

OBSERVING and interpreting motion trajectories, obtained by tracking moving objects of interest [3], [26], is an important topic in computer vision. In this paper, we propose to characterize motion patterns, contents by motion trajectories. Among the potential fields of applications are such areas as activity inference, motion retrieval, and learning motion pattern from demonstration. Despite the progress made in the past decade in trajectory-based analysis, many challenging problems still remain, such as 3-D trajectory matching, clustering, retrieval, and classification. All of these rely on an effective way to represent motion trajectories, where invariant to group transformations, tolerant to noise, and insensitive to occlusions in trajectories are demanded in most of the scenarios. In this paper, integral invariants are proposed to represent motion trajectories. The integral invariants are defined as line integrals of a class of kernel functions along a motion trajectory, and estimated by the proposed formula. The integral invariants admit invariant to translation, rotation, reflection, and scaling. They are also insensitive to the effects of noise and occlusions in motion trajectories, because the integral invariants are computed via integrating kernel functions over local neighboring points. In addition, the integral invariants can be extended to multiscale representation that allows us to capture varying detailed features of a motion trajectory at multiple scales. In order to match and cluster motion trajectories, dynamical time wrapping (DTW) distance of the integral invariants is defined to estimate the similarity between motion trajectories.

Motion trajectories can provide a key cue in motion analysis. In past decade, motion trajectories have been extensively studied in the domain of activity analysis [1]–[4], video surveillance [5], learning semantic scene [29], and motion retrieval [6]. In most of the related work, motion trajectories were often used directly in the raw data form or processed naively without invariance and robustness. Shape descriptor, a set of parameters, that are produced to describe the features of a given shape has been developed to attempt shape matching and classification. There are some task-specific shape descriptors [7] that are studied for shape or curve matching, such as curvature scale space [10], [11], B-spline [12], polynomial curve fitting [8], and shape context [13]. Especially, scale space methods [10], [11] have gained much attention in the contour description of 2-D shapes, where those descriptors are obtained by convolving the shape with a series of Gaussian kernels at multiple scales in a coarse-to-fine manner. However, in those scale space methods, the shape boundary is deformed at varying scales, thereby yielding undesirable distortions in the shape. B-spline and polynomial curve fitting show nonflexible when the sampling rates of motion trajectories varies or partial occlusions exist in a trajectory, because their approximation accuracies depend on some key control points. Shape context as a local descriptor owns some rich invariant properties, and is capable of handling occlusions in shape, but it is not the best way to describe a 3-D trajectory due to its coarse captured distributions of the shape. In addition, transformation functions such as Fourier and Wavelet [14], [15], extract global features from motion trajectories, but meanwhile the local features of motion trajectories are lost. Moment invariants for 3-D curves under similarity transformations were derived in [22] and [23], but they are global descriptors

Manuscript received November 10, 2014; revised February 7, 2015; accepted February 12, 2015. Date of publication March 3, 2015; date of current version January 13, 2016. This work was supported in part by the Research Grants Council of Hong Kong under Grant Project CityU 118613, in part by the Center for Robotics and Automation, and in part by the National Natural Science Foundation of China under Grant 61273286. This paper was recommended by Associate Editor J. Su.

The authors are with the Department of Mechanical and Biomedical Engineering, City University of Hong Kong, Hong Kong (e-mail: meyfli@cityu.edu.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2015.2404828

2168-2267 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

that are sensitive to the effects of occlusions, and high-order moments of them are sensitive to noise [23].

Most shape descriptors above mentioned were proposed initially for 2-D open or closed contour shapes. These shape descriptors cannot be directly moved to our research. We intend to represent a motion trajectory in both 2-D and 3-D Euclidean space with some essential properties including uniqueness of representation, invariance to specific transformation groups, noise resistance, and insensitivity to partial occlusions.

Invariants have played an important role for various applications in computer vision ranging from shape matching [13] to object recognition [24], gesture recognition [25], and motion perception [9]. Consequently, plentiful features that are invariant to some specific transformations have been studied in [24] and [27]. Two types of invariant descriptors that relate to our research are differential invariants and integral invariants, which have been thoroughly investigated and put into applications in [9], [30], and [32]. The locality is a useful property for overcoming occlusions in using differential invariants. However, differential invariants involves high-order derivatives and hence are very sensitive to noise, though approximated in terms of joint invariants [24]. To overcome the limitations of differential invariants, there have been a lot of work to attempt deriving integral invariants [21], [32], [37] based on integrals other than derivatives. In [32], potentials were proposed to get integral invariants for planar shapes via integrating the potentials of contour curves of shapes, but these integral invariants are global descriptors. Integral invariants for closed planar shapes [21], [37] were derived by performing integration of a local kernel along the shape boundary represented by a planar curve, where the locality is achieved by restricting integration on local neighborhoods at each point of the curve. Especially, multiscale integral invariants in [37] are obtained by integrating a series of isotropic kernels over a shape at multiple scales. However, they cannot be extended to represent 3-D temporal trajectories since that in 3-D space the fact of open contours and varying orientations of 3-D trajectories make the problems complicated. Consequently, the ideal of defining some functions on the domain of a motion trajectory, which can admit invariant under group transformations in 3-D space, still remains unresolved. In [35], the multiscale integral invariants for a 3-D motion trajectory have been explored primitively. Furthermore, in this paper, we will build a complete theory, and give some comprehensive discussions on them.

In this paper, we first propose a general definition of integral invariant for free form 3-D motion trajectories using the line integral of a kernel function along motion trajectories. Then, depending on the designed kernels, we develop two typical integral invariants of transformation groups, the distance and area integral invariants. Especially, for the area integral invariants, they are defined as the line integrals over a dynamic domain of integration within the instant Frenet–Serret frame at each point of the 3-D trajectory. To avoid the computation of high-order derivatives, a novel discrete geometric approach, maximal blurred segment (MBS) [18] of discrete curves, is then accordingly employed to estimate the Frenet–Serret frame along the 3-D trajectory, thereby getting the stable results of the area integral invariants. Moreover, to tap into detailed features of a motion trajectory at multiple scales, we evolve the integral invariants to their multiscale representation in a coarse-to-fine manner by varying the scale of the corresponding kernel function. In order to match trajectories, we define a DTW distance for measuring similarity between motion trajectories. Finally, to evaluate the robustness and effectiveness of the proposed integral invariants, we conduct three experiments in motion trajectory matching, sign recognition, where the integral invariants are used to measure the trajectory similarity under different situations including transformation, noise, and occlusion.

II. Integral Invariants for Motion Trajectory

A 3-D motion trajectory is a set of position vectors of a moving object in 3-D Euclidean space. Normally, it can be represented by a set of triple parametric functions with respect to the arc length s, $\boldsymbol{\gamma}(s) = \{x(s), y(s), z(s)\}$, being assumed to a regular curve in this paper. As motion trajectories related to our research are tracked in temporal sequence, it is preferable for a motion trajectory to be represented with respect to the temporal parameter t

$$\boldsymbol{\gamma}(t) = \{x(t), y(t), z(t) | t \in [a, b]\} \tag{1}$$

where $[a, b]$ is the time interval. Note that in practical applications motion trajectories are often sampled discretely. In this case, t is set to $[1, N]$ and N is the trajectory length (frames), whereas in the definition of integral invariant we assume a motion trajectory is a regular curve in terms of arc length parameter or continuous temporal parameter.

A. Definition of Integral Invariant

As addressed in [21] and [24], we can deduce 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 3-D motion trajectory is defined in Definition 1.

Definition 1: Let G be a transformation group acting on $\mathbb{R}^3$. The function $I : \mathbb{R}^3 \to \mathbb{R}$ is a G-invariant for the trajectory $\boldsymbol{\gamma}$ if it satisfies

$$I(\boldsymbol{\gamma}) = I(g \bullet \boldsymbol{\gamma}), \forall g \in G. \tag{2}$$

The function $I(\bullet)$ at each point of the trajectory $\boldsymbol{\gamma}$ returns a real number. The defined invariant includes some familiar examples, such as curvature and torsion (differential invariants) in [9] and [17]. In this paper, we define a new integral invariant for a 3-D motion trajectory based on the line integral of kernel function along the motion trajectory in Definition 2.

Definition 2: A function $I(p)$ is an integral G-invariant for the 3-D trajectory $\boldsymbol{\gamma}$ at point p if there exists a kernel $k : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ such that

$$I(p) = \int_{\boldsymbol{\gamma}} k(p, \boldsymbol{\gamma}(s)) ds, \forall p \in \boldsymbol{\gamma} \tag{3}$$

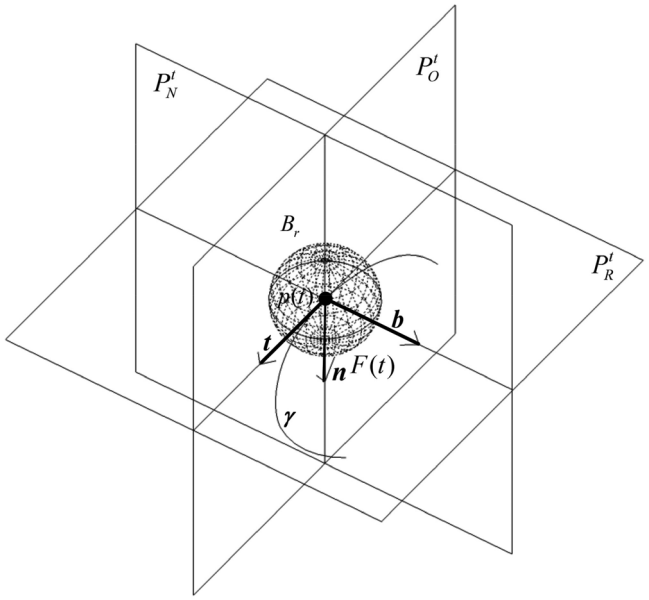


Fig. 1. Frenet–Serret frame and ball kernel at the reference point of a 3-D trajectory.

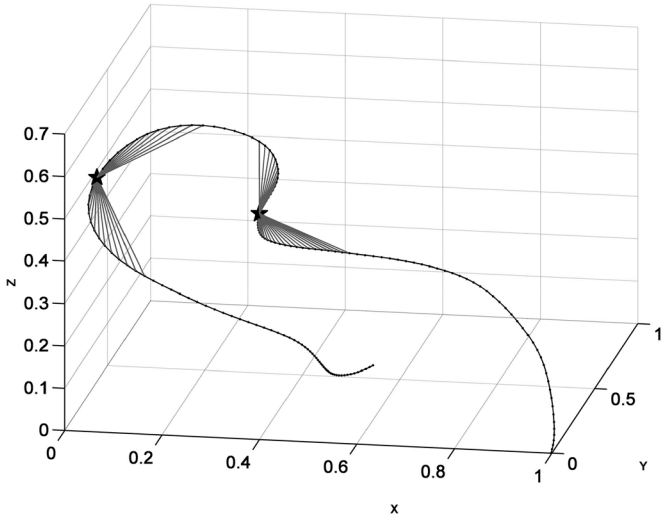


Fig. 2. Distance integral invariant defined in (6), computed by averaging the distances (represented by gray line) of the neighboring points with respect to the reference points denoted with the stars.

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

$$\int_{\boldsymbol{\gamma}} k(p, \boldsymbol{\gamma}(s))ds = \int_{g\boldsymbol{\gamma}} k(gp, g \bullet \boldsymbol{\gamma}(s))ds, \forall g \in G. \quad (4)$$

In most visual applications, the group G is typically Euclidean, similarity, affine or projective group. At this stage, this paper only focuses on the invariants of Euclidean and similarity group, since that in 3-D vision system the actions of projective and affine group in the reconstruction of stereo image pairs can be avoided as discussed in [9].

The kernel function in Definition 2 plays a key role to derive some meaningful integral invariants for motion trajectories, which can be designed depending on the final goal. It is expected to meet two criteria, where firstly it can preserve computational locality to handle the partial occlusions in motion trajectories, and secondly it can be integrated over the motion trajectories to get a group invariant. Considering these two aspects, we first define a ball kernel that can restrict the domain of integration to the local neighborhoods at each point. It is defined as follows.

A ball kernel function is defined with $B_r : \mathbb{R}^3 \times \mathbb{R}^3 \to \{0, 1\}$ to indicate whether the point ε is located on the interior of the sphere with radius r centered at point p of the motion trajectory $\boldsymbol{\gamma}$

$$B_r(p, \varepsilon) = \begin{cases} 1 \cdots \|p - \varepsilon\| \leq r \\ 0 \cdots \text{otherwise} \end{cases} \forall p \in \boldsymbol{\gamma}, \varepsilon \in \mathbb{R}^3. \quad (5)$$

The ball kernel is illustrated in Fig. 1, where we associate a ball with each point of the motion trajectory to restrict the domain of integration to the local neighborhoods with a specified radius. Given the defined ball kernel, we will derive two integral invariants for a 3-D motion trajectory by designing two typical kernel functions to be integrated as addressed in the next sections.

B. Distance Integral Invariant

Let $G \in E(3)$ be a Euclidean group, a symmetry group of 3-D Euclidean space. Given a distance kernel $d(p, \boldsymbol{\gamma}(t))$ and a ball B_r with radius r at point $p \in \boldsymbol{\gamma}$, we have

$$I^r(p) = \int_{\boldsymbol{\gamma}} k(p, \boldsymbol{\gamma}(s))ds = \int_a^b B_r(p, \boldsymbol{\gamma}(t))d(p, \boldsymbol{\gamma}(t))ds(t) \quad (6)$$

where $d(p, \boldsymbol{\gamma}(t)) = \|p - \boldsymbol{\gamma}(t)\|$ is the Euclidean distance in $\mathbb{R}^3$. This distance integral invariant is similar with shape context [13] which captures the local histogram bins of the neighboring points restricted by a log-polar weight kernel at the reference point, whereas in our case we store only the mean distance of the neighboring points restricted by the ball kernel. The distance integral invariant is a continuous local invariant provided by the distance kernel restricted on a ball neighborhood, as illustrated in Fig. 2. It is also insensitive to noise as shown in Fig. 8. Thus, the distance integral invariant shows insensitive to noise and occlusions in trajectory that can be demonstrated by experiments in Section VI. The distance integral invariant can be normalized to admit equivalent when scaling is performed on a trajectory. We normalize it by

$$I^r(p)^* = \frac{I^r(p)}{\max(B_r(p, \boldsymbol{\gamma}(t))d(p, \boldsymbol{\gamma}(t))) - min(B_r(p, \boldsymbol{\gamma}(t))d(p, \boldsymbol{\gamma}(t)))} \quad (7)$$

As suggested in [21], distance invariants are not discriminative enough in that a unique distance invariant can correspond to different geometric features. This ambiguity for distance invariant can be amplified in 3-D space especially for the complicated trajectories, which can be demonstrated in Section VI.

C. Area Integral Invariants

Regarding area integral invariants, we first introduce a special moving frame in 3-D Euclidean space, Frenet–Serret frame which the area integral invariants are defined based on.

In n-dimensional space, moving frame plays an important role in studying the extrinsic properties of smooth manifolds

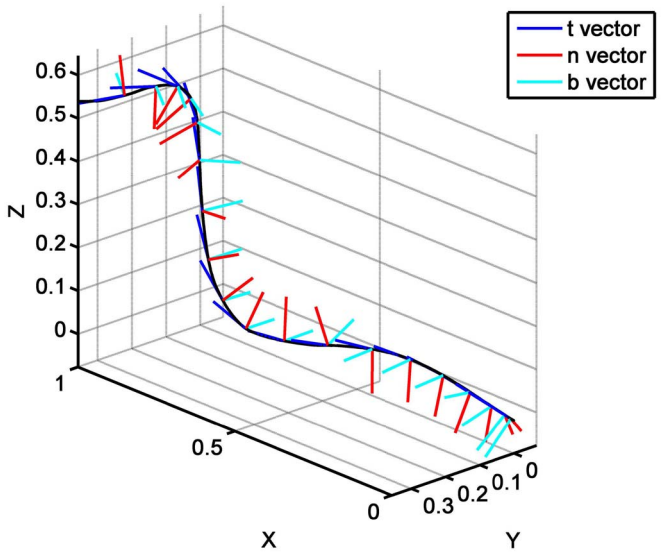


Fig. 3. Visualization of the $\boldsymbol{t}, \boldsymbol{n}$, and $\boldsymbol{b}$ vectors of Frenet–Serret frames of a 3-D motion trajectory.

in differential geometry. Joint invariants and differential invariants up to group transformations are developed in object recognitions in terms of moving frames [33], [34]. Frenet–Serret frame is a special moving frame which describes the kinematic properties of a particle along a regular curve in 3-D Euclidean space. We define our area integral invariants based on the Frenet–Serret frame, a real-valued frame $F(t)$ of reference that moves temporally with the observer along a trajectory $\boldsymbol{\gamma}(t)$ by the change of motion direction of the frame vectors, and exists locally. In Frenet–Serret frame, each point of a parametric trajectory is associated with a set of triple orthogonal unit vectors to describe its dynamic properties: 1) the tangent vector $\boldsymbol{t}$; 2) principle normal vector $\boldsymbol{n}$; and 3) binormal vector $\boldsymbol{b}$. These triple vectors $(\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b})$ at the reference point $p(t) \in \boldsymbol{\gamma}$ together form an orthonormal basis spanning $\mathbb{R}^3$ locally with three projection planes: 1) the osculating plane P_O^t; 2) normal plane P_N^t; 3) and rectifying plane P_R^t as shown in Fig. 1. The three vectors of the Frenet–Serret frame $F(t)$ at point $p(t)$ are an orthonormal basis constructed from the Gram–Schmidt process to the vectors and their derivatives, defined as follows:

$$F(t) = \{\boldsymbol{t}(t), \boldsymbol{n}(t), \boldsymbol{b}(t)\} \tag{8}$$

where

$$\begin{cases} \boldsymbol{t}(t) = \dfrac{\boldsymbol{\gamma}'(t)}{\|\boldsymbol{\gamma}'(t)\|} \\ \boldsymbol{n}(t) = \dfrac{\boldsymbol{t}'(t)}{\|\boldsymbol{t}'(t)\|} = \dfrac{\boldsymbol{\gamma}'(t) \times \left(\boldsymbol{\gamma}''(t) \times \boldsymbol{\gamma}'(t)\right)}{\|\boldsymbol{\gamma}'(t)\| \|\boldsymbol{\gamma}''(t) \times \boldsymbol{\gamma}'(t)\|} \\ \boldsymbol{b}(t) = \boldsymbol{t}(t) \times \boldsymbol{n}(t). \end{cases} \tag{9}$$

In this manner, all the points in a 3-D motion trajectory can be represented by their unit vectors, $(\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b})$. Fig. 3 shows a sequence of Frenet–Serret frames along a motion trajectory. Basing on the Frenet–Serret frame, we hence intuitively define the area integral invariants at point p as line integrals of a scalar field (kernel function) along two projected trajectories at point p, $\bar{\boldsymbol{\gamma}}_O^p$, and $\bar{\boldsymbol{\gamma}}_R^p$. They are orthogonally projected by the 3-D motion trajectory $\boldsymbol{\gamma}$ onto the osculating and rectifying planes of the Frenet–Serret frame at point p shown in Fig. 4. Geometrically,

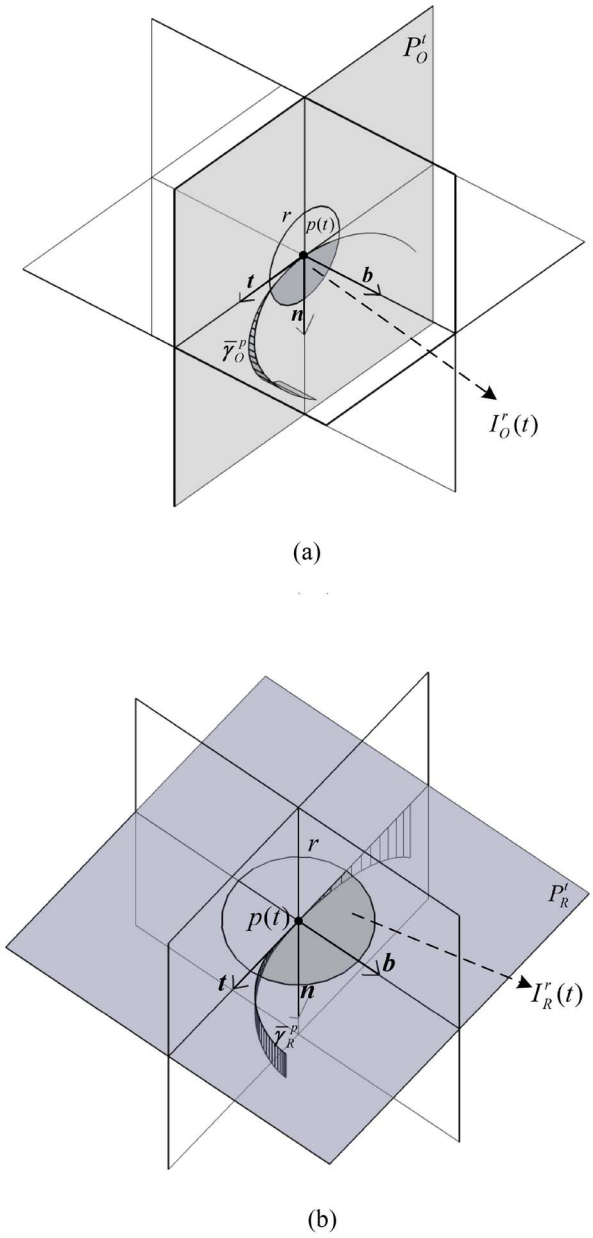


Fig. 4. Domain of integration on (a) osculating plane and (b) rectifying projection plane restricted by a ball kernel shown in Fig. 1. A 3-D trajectory is projected onto the osculating and rectifying plane of the Frenet–Serret frame at the reference point.

we want to define the area, fell into the interior of the circle intersected by the projected trajectory on either osculating or rectifying plane, as area integral invariants indicated as shaded area in Fig. 4. Following this fact, for any given ball kernel of radius r, the corresponding area integral invariants at point p can be defined as the line integrals of a scalar field f on both the osculating and rectifying plane, respectively:

$$\begin{cases} I_O^r(p) = \int_{\bar{\boldsymbol{\gamma}}_O^p} f_O(p, \boldsymbol{\gamma}(s))ds = \int_a^b f_O(p, \boldsymbol{\gamma}(t)) \left\| \bar{\boldsymbol{\gamma}}_O^p(t)' \right\| dt \\ I_R^r(p) = \int_{\bar{\boldsymbol{\gamma}}_R^p} f_R(p, \boldsymbol{\gamma}(s))ds = \int_a^b f_R(p, \boldsymbol{\gamma}(t)) \left\| \bar{\boldsymbol{\gamma}}_R^p(t)' \right\| dt \end{cases} \tag{10}$$

where the kernel function f_O or f_R, also called the integrand, is defined as the circle on either osculating or rectifying plane projected by the continuous sphere of the ball kernel B_r as shown in Fig. 5. Here, we first derive the area integral invariant

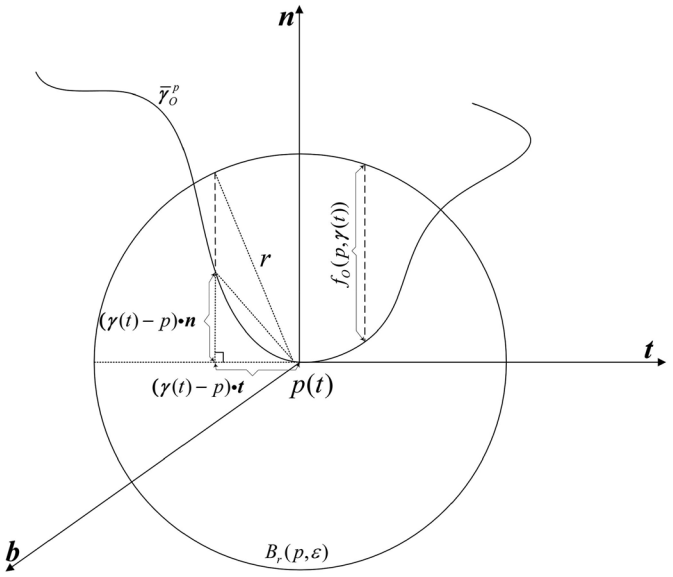


Fig. 5. Definition of the kernel functions for the area integral invariant on the osculating plane at the reference point. In the same way, the definition is same on rectifying plane.

on osculating plane at point p

$$f_O(p,\boldsymbol{\gamma}(t)) = B_r(p,\boldsymbol{\gamma}(t))\left(\sqrt{r^2-\{(\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{t}\}^2} - (\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{n}\right) \Rightarrow$$
$$I_O^r(p) = \int_a^b B_r(p,\boldsymbol{\gamma}(t))\left(\sqrt{r^2-\{(\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{t}\}^2} - (\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{n}\right)\left\|\bar{\gamma}_O^p(t)'\right\|dt. \tag{11}$$

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

$$\frac{ds}{dt} = \left\|\bar{\boldsymbol{\gamma}}_O^p(t)'\right\| = \sqrt{(\boldsymbol{\gamma}'(t)\bullet\boldsymbol{t})^2+(\boldsymbol{\gamma}'(t)\bullet\boldsymbol{n})^2} = \left\|\boldsymbol{\gamma}'(t)\bullet\left[\boldsymbol{t}\ \boldsymbol{n}\ \boldsymbol{0}\right]\right\|. \tag{12}$$

And therefore, by (11) and (12), we have

$$I_O^r(p) = \int_a^b B_r(p,\boldsymbol{\gamma}(t))\left(\sqrt{r^2-\{(\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{t}\}^2} - (\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{n}\right)\left\|\boldsymbol{\gamma}'(t)\bullet\left[\boldsymbol{t}\ \boldsymbol{n}\ \boldsymbol{0}\right]\right\|dt. \tag{13}$$

The area integral invariant on rectifying plane can be derived in the manner similar to the case of derivation of $I_O^r(p)$. Then, we have

$$I_R^r(p) = \int_a^b B_r(p,\boldsymbol{\gamma}(t))\left(\sqrt{r^2-\{(\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{t}\}^2} - (\boldsymbol{\gamma}(t)-p)\bullet\boldsymbol{b}\right)\left\|\boldsymbol{\gamma}'(t)\bullet\left[\boldsymbol{t}\ \boldsymbol{0}\ \boldsymbol{b}\right]\right\|dt. \tag{14}$$

The complete area integral invariants for a 3-D trajectory at point p then can be concatenated to the vector

$$I^r(p) = \left\{I_O^r(p), I_R^r(p)\right\} \tag{15}$$

where they can be further normalized by the area of the projected circle with radius r

$$I^r(t)^* = I^r(p(t))^* = \frac{I^r(p(t))}{\pi * r^2}. \tag{16}$$

The normalized area integral invariants are then bounded between 0 and 1.

We can observe from (10) to (14) that the area integral invariants are derived based on local distances and areas within the local Frenet–Serret frame which is independent of coordinate system. The dot and cross products of derivatives are purely local quantities of the trajectory features. Therefore, the area integral invariants are invariant up to Euclidean group transformations. Furthermore, by normalizing the area integral invariants, they can also admit invariant under similarity group.

III. Estimation of Area Integral Invariants

As we claimed, integral invariants do not involve high-order derivatives. The distance integral invariant can be computed directly according to (6), whereas the area integral invariants cannot be computed directly by the definition which requires the second-order derivatives at least. In practical applications, motion trajectories are discrete temporal sequences sampled from visual sensors in the presence of noise. This fact makes the computation of the area integral invariants sensitive to noise. It is observed from (8) to (14) that obtaining the instance Frenet–Serret frame $F(t)$ at each point is the key problem to compute the area integral invariants. Therefore, we here propose a formula to estimate $F(t)$ which the numerical integrations for the area integral invariants are based on.

3-D MBS of noise curve [18], [20] here is employed to decompose a discrete noise trajectory into some consecutive overlapped minimally thin blurred segments via eliminating, bypassing those noise points. Next, we further get the corresponding left and right key points of the blurred segments nearby each reference point so that $F(t)$ can be constructed by those noncollinear triple points. The detail of MBS is briefly recalled as follows.

We segment a known 3-D discrete curve into a number of 3-D discrete lines $D_{3\text{-D}}(a,b,c,\mu,\mu',\omega,\omega')$ [19] (blurred segment) of the width v that is to control the segmentation level of a sequence of points $\boldsymbol{\gamma}(i,j)$ in the 3-D discrete curve such that $\omega' - 1/\max(|a|,|b|) \le v$ on the plane OXY and $\omega - 1/\max(|a|,|c|) \le v$ on the plane OXZ. An example of segmentation of a discrete curve on the plane OXZ is shown in Fig. 6, where there are two consecutive blurred segments and the width is determined by the dynamical thickness estimation of convex hull [38] that consists of a set of successive discrete points as shown in Fig. 6.

Nguyen and Debled-Rennesson [18] and Faure *et al.* [20] further proposed the concept of MBS of width v, which means each segmented 3-D discrete line for a discrete curve cannot be extend neither at right side nor at left side for given width v. Basing on this concept, we can decompose a 3-D

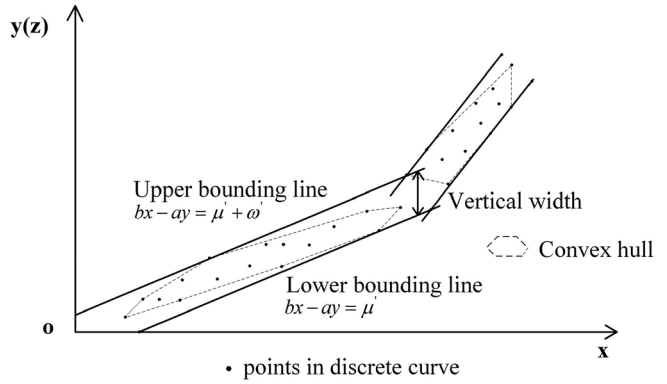


Fig. 6. Optimal bounding line D of two successive blurred segments of a discrete curve in OXZ plane.

discrete trajectory $\boldsymbol{\gamma}$ into a sequence of intercrossed MBSs of width v with m length

$$\mathrm{MBS}_v(\boldsymbol{\gamma}) = \{\mathrm{MBS}(B_1, E_1, v), \ldots, \mathrm{MBS}(B_m, E_m, v)\} \quad (17)$$

with $B_1 < B_2 < \cdots < B_m$ and $E_1 < E_2 < \cdots < E_m$. $\{B_i, E_i | i \in [1, m]$ here denote the beginning and ending positions of a MBS of the discrete trajectory.

Given a sequence of $\mathrm{MBS}_v(\boldsymbol{\gamma})$, to estimate the Frenet–Serret frame at a point, we first denote the estimated key points of the discrete trajectory with $\boldsymbol{\gamma}(B_i), \boldsymbol{\gamma}(E_i) | i \in [1, m]$ from the MBSs of the discrete trajectory. Then, let $R(k) \in \{B_i, E_i\}, L(k) \in \{B_i, E_i\} | k = 1 \cdots N$ record a sequence of positions of the estimated right nearest key points and left nearest key points at the reference point $\boldsymbol{\gamma}(k)$ such that: $L(k) < k < R(k)$, where we assume these triple points $\{\boldsymbol{\gamma}(R(k)), \boldsymbol{\gamma}(k), \boldsymbol{\gamma}(L(k)) | k = 1 \cdots N\}$ to be always not collinear. We, then, approximate the osculating circle at $\boldsymbol{\gamma}(k)$ using the circumcircle of the triangle bounded by these triple points. Let $C(k)$ be the center of the circumcircle.

Then, we define the norm vector at $\boldsymbol{\gamma}(k)$ as $\boldsymbol{n}(k) = \overline{\boldsymbol{\gamma}(k)C(k)} / |\overline{\boldsymbol{\gamma}(k)C(k)}|$. The unit tangent vector $\boldsymbol{t}(k)$ is the unit vector that is tangent with the osculating circle at point $\boldsymbol{\gamma}(k)$. Then, the binormal vector $\boldsymbol{b}(k)$ is obtained straightforwardly by cross product: $\boldsymbol{b}(k) = \boldsymbol{t}(k) \times \boldsymbol{n}(k)$. Now, the Frenet–Serret frame associated with each point can be formed and the triple projection planes are spanned by the triple norm vectors. We describe this estimation algorithm in Table I with MATLAB pseudocodes, in which the recognition of blurred segment refers to that in [18].

Given the estimated projection planes at the reference point, we first project the 3-D discrete trajectory onto the osculating and rectifying plane respectively, forming a projected trajectory intersected with the projected circle of sphere on either osculating or rectifying plane. Moreover, we further discretize the intersected trajectory by sampling it at small-scale grid resolution on the corresponding projection plane and thus approximate the intersected area using numerical integration. It should be noted that for our estimation, the trajectory is normalized first to keep the same scale with standard radius range (0–1) of ball kernel.

As addressed in our approach, the main burden of computing integral invariants falls on the recognitions of MBSs of a general discrete trajectory. In our formula, the recognition is sped up via the dynamical determination of the convex hull of a set of successive points as shown in Fig. 6. For blurred segment recognition [19], Melkman's algorithm is a classical method in discrete geometry to determine the thickness of convex hull that leads to linear complexity. However, to determine MBSs of a discrete trajectory, it is not efficient with the complexity $\mathcal{O}(n^2)$ (n is the number of points of the studied trajectory) in that convex hull has to be recalculated at each point when a new point is inserted or removed from a blurred segment. In order to achieve online determination of the set of MBS, we use the dynamical estimation of convex hull [38] to maintain a set of points online when inserting or removing a new point into the convex hull. This enhancement results in the computational cost in $\mathcal{O}(n \log^2 n)$ complexity, making the real-time computation of the integral invariants possible. Refer to [18] and [38] for details and complexity proofs.

TABLE I
ESTIMATION ALGORITHM FOR AREA INTEGRAL INVARIANTS

Algorithm 1: Estimation of area integral invariants for a 3D discrete trajectory.

Input data: a 3D discrete trajectory $\boldsymbol{\gamma}$ of n points, width v of the segmentation, radius r of the circle kernel for local area integral invariant;

Result: I^r a sequence of area integral invariants with radius r;

1. Employ the maximal 3D blurred segment algorithm [18] to decompose the trajectory $\boldsymbol{\gamma}$ into a sequence of maximal blurred segments of width v, and get a sequence of segments: MBS_v ;
2. Build $MBS_v = \{MBS(B_i, E_i, v)\}$;
3. $I^r = ones(\mathrm{n}, 2) * 0;$
4. $m = length(MBS_v)$; $E = -1; B_m = n;$;
5. for $i = 0$ to $m - 1$
6. for $k = E_{i-1} + 1$ to E_i do $L(k) = B_i$;
7. for $k = B_i$ to $B_{i+1} - 1$ do $R(k) = E_i$;
8. End
9. for $k = 2$ to $n - 2$
   compute the $I^r(k)$ at k^{th} point by inputting the radius r and the triple set $[\boldsymbol{\gamma}(L(k)), \boldsymbol{\gamma}(k), \boldsymbol{\gamma}(R(k))]$;
   % *project the trajectory* $\boldsymbol{\gamma}$ *onto the osculating and rectifying planes that are constructed via the circumcircle bounded by the triple set* $[\boldsymbol{\gamma}(L(k)), \boldsymbol{\gamma}(k), \boldsymbol{\gamma}(R(k))]$, *and then perform numerical integration of line integrals along the projected trajectories respectively.*
10. end

IV. MULTISCALE INTEGRAL INVARIANTS

In multiscale space theory [11], [37], a shape is analyzed at multiple scale spaces by convoluting the shape with a range of scales of kernels. Inspired by this theory, to more tap into the detailed features of a motion trajectory at different scales, we get the multiscale integral invariants by decreasing the radius of ball kernel from an initial value step-by-step. In this paper, we let the radius of ball kernel be decreased by half at each step, and then obtain the multiscale integral invariants at all the scales defined as

$$I(t)^* = \left\{ I^r(t)^* | r = \bar{r}/2^n, n = 0, \ldots, K \right\} \quad (18)$$

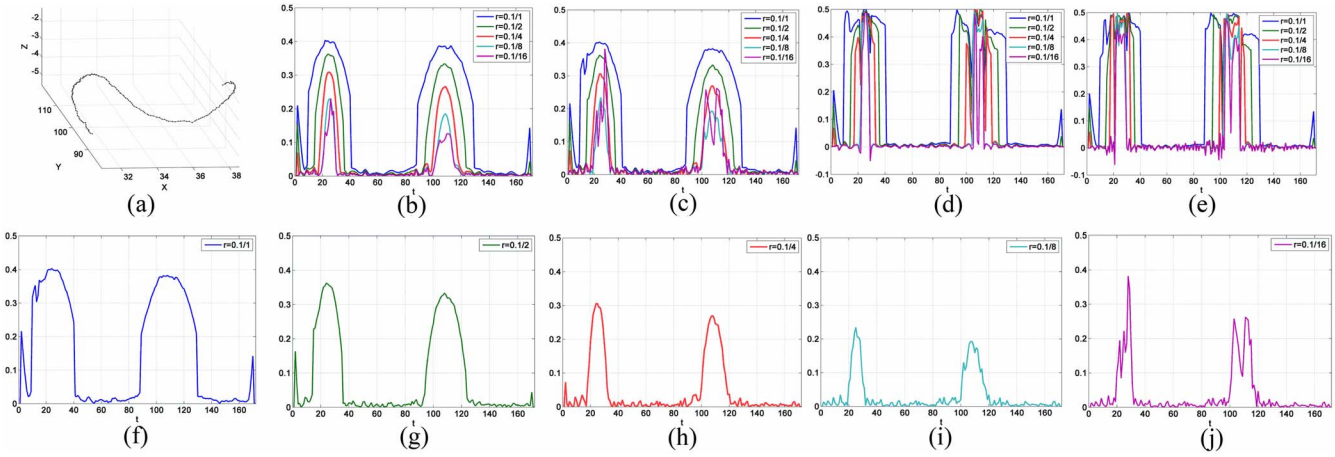


Fig. 7. Multiscale area integral invariants for (a) an example noisy motion trajectory. (b) and (d) plot the multiscale area integral invariants for the original trajectory on the osculating and rectifying planes respectively, whereas (c) and (e) plot their noise versions computed from the noisy trajectory (a). The multiscale area integral invariants are computed at 5 scales as the radius varied from initial value $\bar{r} = 0.1$, where we decompose the (c) into 5 individual one (f)–(j) at each scale.

where $\bar{r}$ denotes the initial radius, K stands for the number of scale that determines how deep a motion trajectory can be taped into and perceived with the multiscale integral invariants. As multiresolution analysis implied in signal processing, multiscale space theory can obtain abundant information about a contour at different scales. The features at more global scale can be interpreted from larger scale, while the features at more local scale can be interpreted from smaller scale. Benefiting from the redundancy of multiscale representation, trajectory alignment, and matching would show higher robustness to noise and a great coherence with human perception. We can generate a figure of the multiscale area and distance integral invariants for a motion trajectory as shown in Figs. 7 and 8 to show their multiscale features in a noisy example trajectory. As indicated in Fig. 7, the multiscale area integral invariants are able to capture not only the comparative global features but also the detailed features on the shape especially when the motion trajectory involves both the subtle changes and coarse changes. More importantly, the multiscale integral invariants show more competitive advantages in perception of noisy motion trajectory at its larger scale, especially for the distance integral invariant as demonstrated in Fig. 8.

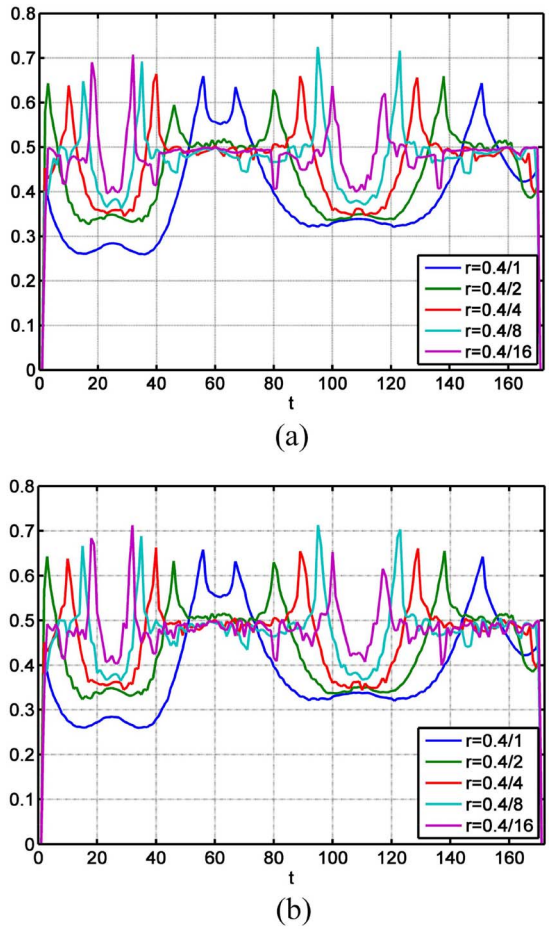


Fig. 8. Multiscale distance integral invariants for the trajectory are shown in Fig. 7(a). (a) Multiscale distance integral invariants for the original trajectory. (b) Corresponding noisy version computed from the noisy trajectory in Fig. 7(b), as the radius varies from initial value $\bar{r} = 0.4$.

V. Similarity Measure

In this section, we define a similarity measure between a pair of motion trajectories. By defining a distance function, we can measure how similar between them to match, cluster, and classify motion trajectories. We want the similarity to be insensitive to some variations such as unequal length, different sampling, and partial occlusions in trajectory. Therefore, we employ the DTW algorithm to deal with all of these variations, where the local distance in DTW alignment algorithm has to be defined firstly. In [35], we have defined a local distance function between groups of multiple motion trajectories, whereas in this paper, we only focus on the single trajectory matching. We can inherit the definition in [35] to define the local distance function between a pair of single trajectory by keeping only the distance of integral invariants of root trajectory as follows:

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

where $d(m,n) \in \mathbb{R}^{M\times N}$ denotes the local distance between samples m, n from the integral invariants of two motion

trajectories respectively, and the specific terms in (19) are defined to be the $\mathbb{L}^1$ norm difference as follows:

$$\Delta \boldsymbol{I}^{*m,n} = \left\| I_A^r(m)^* - I_B^r(n)^* \right\|_1 \quad S_I^{*m,n} = \left\| \left\{ I_A^r(m)^*, I_B^r(n)^* \right\} \right\|_1. \tag{20}$$

Consequently, 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 (3). We define the total DTW distance between their integral invariants to measure their similarity. To ultimately get the optimized distance, the DTW algorithm is performed to look for an optimal time warping path that minimizes the total distance $D(M, N)$ between two integral invariants. The alignment in every step of matching relies on the minimum sum distance in different routes up to previous step, which based on the dynamic programming, formulated as follows:

$$D(m, n) = \min\{D(m-1, n), D(m, n-1), D(m-1, n-1)\} + d(m, n). \tag{21}$$

Following the above definition, to improve computational efficiency and distance discrimination in (21), we can obtain the final total distance $D(M, N)$ computed recursively from $D(1, 1)$ to $D(M, N)$ by dynamically searching for the best alignment path using an optimized algorithm with slope constraint [31]. The total distance $D(M, N)$ relatively reflects how similar between a pair of trajectories.

VI. Experiments

In this section, we evaluate the effectiveness and robustness of the integral invariants via conducting three experiments on several datasets which contain a larger variety of motion types and variations.

1) *Motion Trajectory Matching:* We first evaluate the locality, invariance, and noise robustness of the integral invariants by matching a number of pairs of motion trajectories extracted from the HDM05 [28] and Berkeley Multimodal Human Action Dataset (MHAD) [16] motion capture datasets, when compared to other similar descriptors. Motion capture datasets use a skeleton model of several joints of body to represent human motion dynamics, where a temporal sequence of joint positions is recorded for each joint. We simulate a series of group transformations, noise, and occlusions on the trajectories to match with each other to examine the noise robustness, invariance to transformation groups, and occlusion handling.
2) *Sign Language Recognition:* As a sign can be abstracted as two hand trajectories, we demonstrate the effectiveness of the integral invariants applied to the sign recognitions via some large benchmarks. Australian Sign Language (ASL) dataset [36] is employed here.

For comparison, these experimental results via two classic descriptors, Fourier descriptor (FD) [14] and differential invariants (DI) [9], are added to compare with the results via our integral invariants that include the distance integral invariant (DII), area integral invariants (AII), and their multiscale representation (MDII and MAII). Note that DTW distance defined in (21) is used here to measure the trajectory similarity in matching, recognition based on the k-nearest neighbor algorithm for all the descriptors except the FD that uses the Euclidean distance. Here, let $k = 1$ in this paper.

To estimate the AII, we implement the MBS algorithm for motion trajectories with width $v = 8$ on all the datasets. The radius r of ball kernel at $\boldsymbol{\gamma}(k)$ is determined by the average of interdistances between $\boldsymbol{\gamma}(k)$ and its nth nearest points, $r = (\|\boldsymbol{\gamma}(k) - \boldsymbol{\gamma}(k-n)\| + \|\boldsymbol{\gamma}(k) - \boldsymbol{\gamma}(k+n)\|)/2$. In other words, the radius of ball kernel at each point depends on the distributions of the neighboring points instead of setting it empirically. In the experiments, we set $n = 3$. For the DII, we get the average distance of the nearest ten points at $\boldsymbol{\gamma}(k)$. For both the multiscale integral invariants, we start the multiscale process at the initial radius of ball kernel $r = 0.3$ and decrease it by half at each level with $K = 5$ steps.

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

$$X_l = \sum_{t=1}^{N} x(t)\ e^{-j2\pi(t-1)(l-1)/N}, 1 \le l \le N. \tag{22}$$

As a result, we get a sequence of coefficients associated with each dimension: $F_l = \{X_l, Y_l, Z_l\}$. The FD is further normalized [14] for keeping invariant to translation, starting point, and scaling, provided that the first coefficient is ignored, and the rest of the coefficients are scaled by the second coefficient. It should be noted that as suggested in [14] FD is not able to keep complete rotation invariance that is achieved by simply taking the magnitude of each Fourier coefficient.

A. Trajectory Matching

For trajectory matching, the test is performed on two popular motion capture datasets, the HDM05 [28] and Berkeley MHAD [16]. HDM05 dataset consists of around 100 action classes performed by five different actors. Most of these actions contain 10–50 different realizations for each action being amounted to 1457 smaller motion clips, and covering a broad spectrum of semantically meaningful variations. The duration of action sequences ranges from 56 to 901 frames. Berkeley MHAD dataset [16] contains 11 actions performed by seven male and five female subjects. Each action is performed five repetitions, yielding about 656 actions sequences (several erroneous actions are eliminated). In MHAD dataset, 11 action classes are recorded by camera, Kinect, motion capture, and accelerometer systems, which are: 1) jump; 2) jumping jacks; 3) bend; 4) punch; 5) wave one hand; 6) wave two hands; 7) clap; 8) throw; 9) sit down; 10) stand up; and 11) sit down/standup. We will use the optical data obtained by motion capture system in this paper, and segment the action sequences with multiple repetitions of the same action into individual sequences with one repetition using labels given by Ofli *et al.* [16].

Firstly, for the test on HDM05 dataset, 16 original motion trajectories of right hand are randomly extracted from 16 action classes respectively to implement trajectory matching including deposit floor, elbow to knee, gab high, hop both leg, jog left, jump down, jumping jack, kick forward, lie

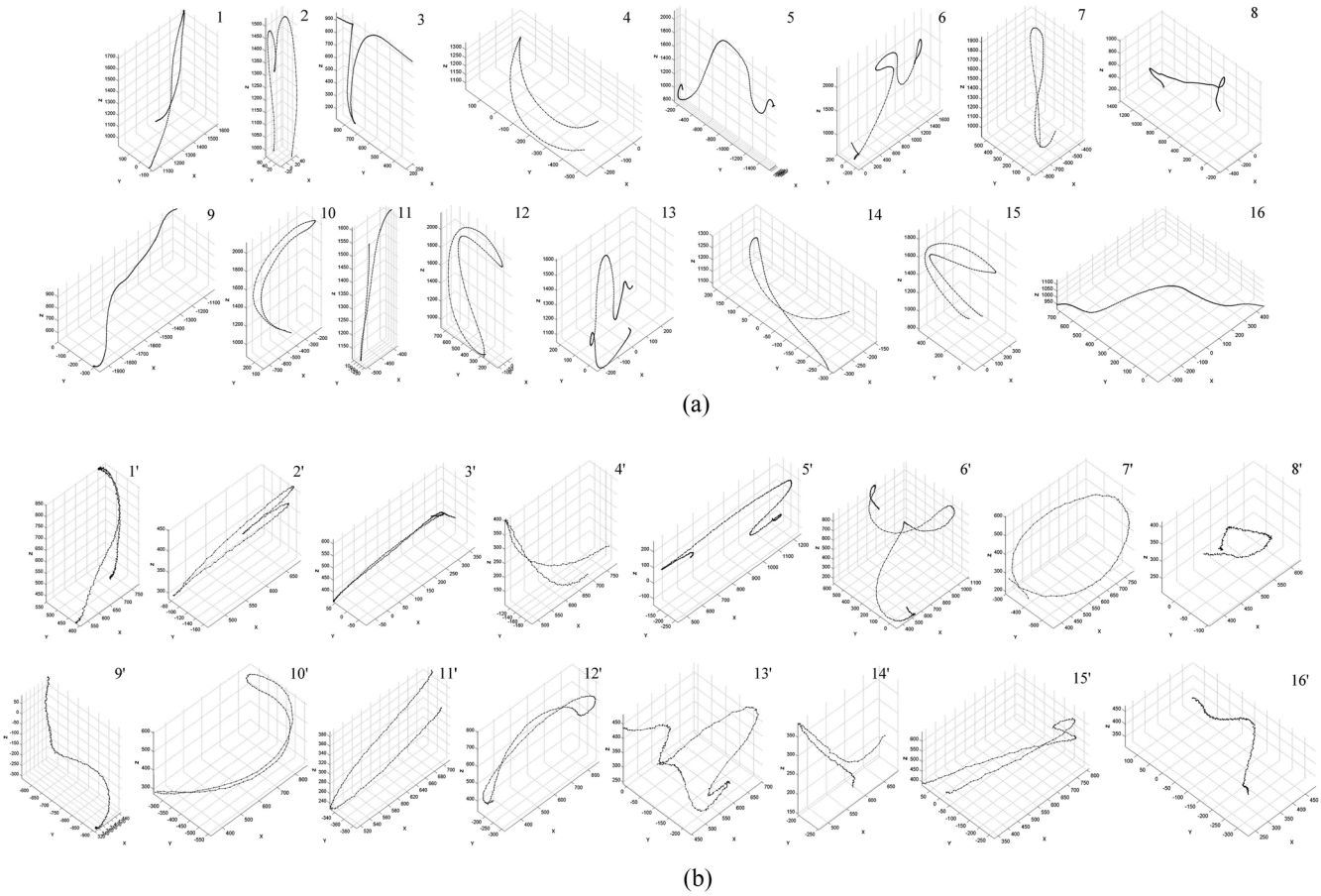
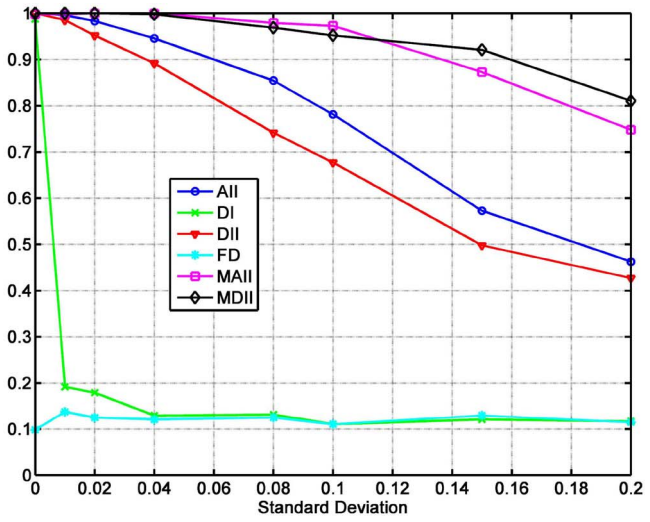


Fig. 9. Examples for a group of extracted 16 motion trajectories of the right hand from HDM05 motion dataset with their two versions, (a) original version and (b) transformation version, without simulated occlusions.

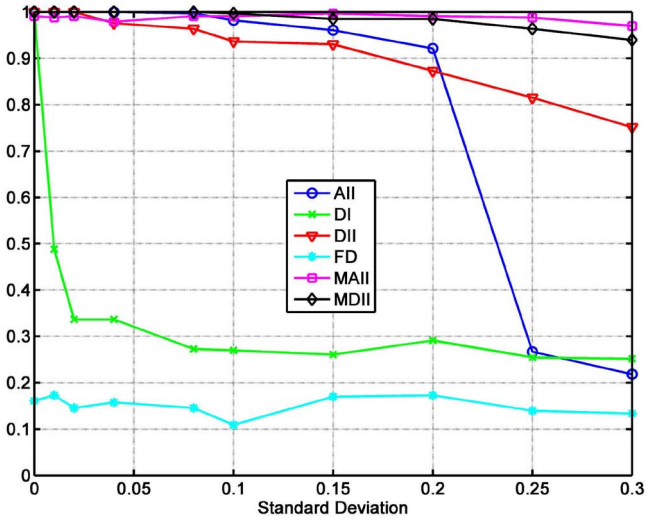
down floor, rotate arm backward, sit down chair, sneak, squat, stand up, throw basketball, and throw. The original motion trajectories are to match their transformation versions, each of which is obtained by a series of actions including first downsampling or upsampling randomly the trajectory to a new one with 0%–50% more or less length than original frames, adding Gaussian noise (normalized standard deviation δ), then rotating 30 and 45 degrees by x- and z-axis respectively, then translating 200 and 500 mm along x- and y-direction respectively, and finally scaling by 0.5 factor. Example for a group of original and transformation versions of motion trajectories are shown in Fig. 9. This extracting and matching procedure is run 50 times getting the average matching accuracies as shown in Fig. 10 compared with other descriptors. Secondly, we carry out the same matching test on MHAD dataset. In similar manner with the first matching test on HDM05 dataset, 11 pairs of motion trajectories of the right hand are randomly extracted to match each other, which correspond to the respective 11 action classes.

The matching performance is evaluated by matching accuracy that is the percentage of correct matching between pairs of trajectories. First, the average matching accuracies between original and transformed trajectories with the additive noise specified by δ are obtained by running the matching test 50 times between trajectories extracted from both the datasets compared to other descriptors. Fig. 10 shows a plot of the matching accuracies on HDM05 and MHAD datasets respectively as noise varies with $\delta = (0 - 0.2$ or $0.3)$. Not surprisingly, it can be shown that the matching accuracies via all of the descriptors decrease as added noise δ increases. The matching accuracies via the DI and FD decrease more drastically as δ increases. The matching even fails via the FD without noise added as shown in Fig. 10 due to its limited rotation invariance. Using the integral invariants (AII and DII), we can get some large improvements in matching performance, but on MHAD dataset the matching accuracy decreases suddenly via AII as δ is beyond 0.2. The decreases of matching accuracies via MAII and MDII are most insignificant, insensitive than other descriptor-based matching as noise increases, especially for the trajectory matching on MHAD dataset where it decreases only slightly. The matching results prove that both the MAII and MDII are most insensitive to noise and group transformations than other descriptors.

To quantitatively evaluate the similarity distance in the presence of the added noise, we compare the average distance matrixes between the original and noisy transformed trajectories in HDM05 dataset, computed via AII, DII, MAII, MDII, and DI as shown in Fig. 12. The distance matrixes computed using MAII and MDII own the most discriminative lower distances on the diagonal as expected in the

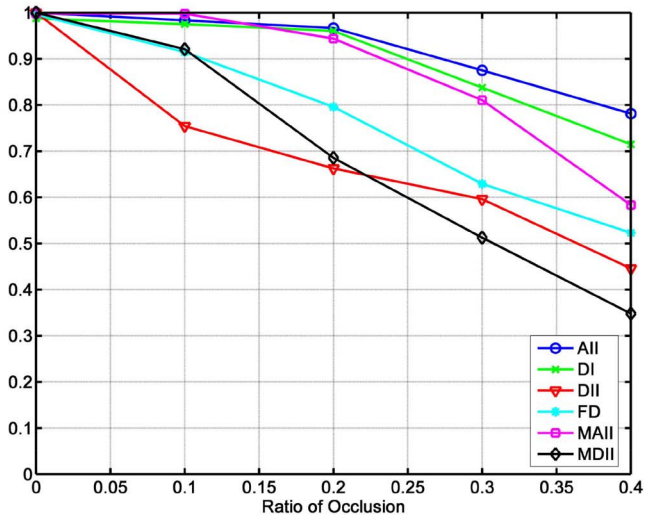


(a)

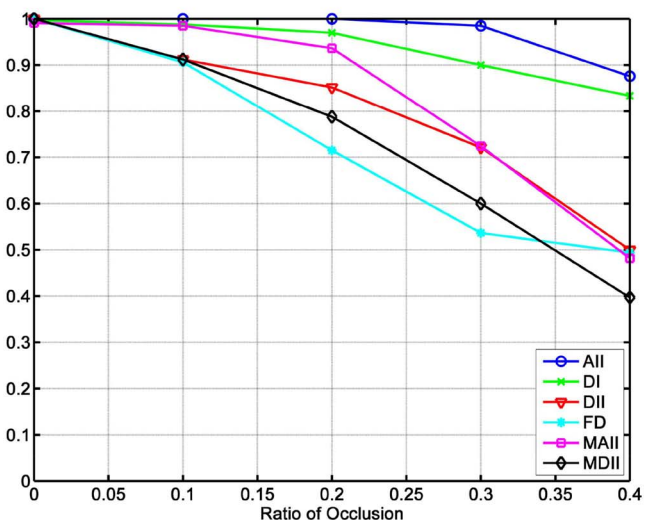


(b)

Fig. 10. Average matching accuracy as the additive noise in trajectories varies. Average matching accuracy on (a) HDM05 dataset and (b) Berkeley MHAD dataset.



(a)



(b)

Fig. 11. Average matching accuracy as the ratio of occlusions in trajectories varies. Experimental results on (a) HDM05 dataset and (b) Berkeley MHAD dataset.

dataset. The distance matrixes computed using AII and DII have relative more discriminative lower distances on the diagonal than the matrix using DI which completely lacks the expected lower distances on the diagonal. These matrix distances show the robustness of integral invariants to noise compared to DI.

One of the advantages using the integral invariants is its insensitive to the effects of partial occlusions. Occlusion is 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. 11 as the ratio of occlusion in motion trajectories increases from 0% to 40% compared to other descriptors. The results show that the matching performance drop, based on both the AII and DI, are relatively smaller thanks to their locality. The matching performance based on the DII and MDII turn out to be worse, even not better than FD-based matching results. This is probably because the DII is not sufficient to represent a trajectory especially under occlusions due to its nonuniqueness as suggested in Section II-B. As the computation of FD depends on the whole trajectory, not surprisingly the matching accuracies via FD decrease drastically as the simulated occlusion increases. It should be noted that to test how the matching performance is influenced only by occlusions using the FD, the transformed trajectories here are only translated without rotation. The matching via MAII is not expected to perform the best among all the descriptors, even worse than the FD-based matching on MHAD dataset. This is because multiscale representation involves multiple levels of descriptions computed from a smaller scale to a larger scale, where the computational locality of the multiscale representation is compromised at the larger scale. Therefore, in occlusion handling, the AII are proved to be the best in matching occluded trajectories.

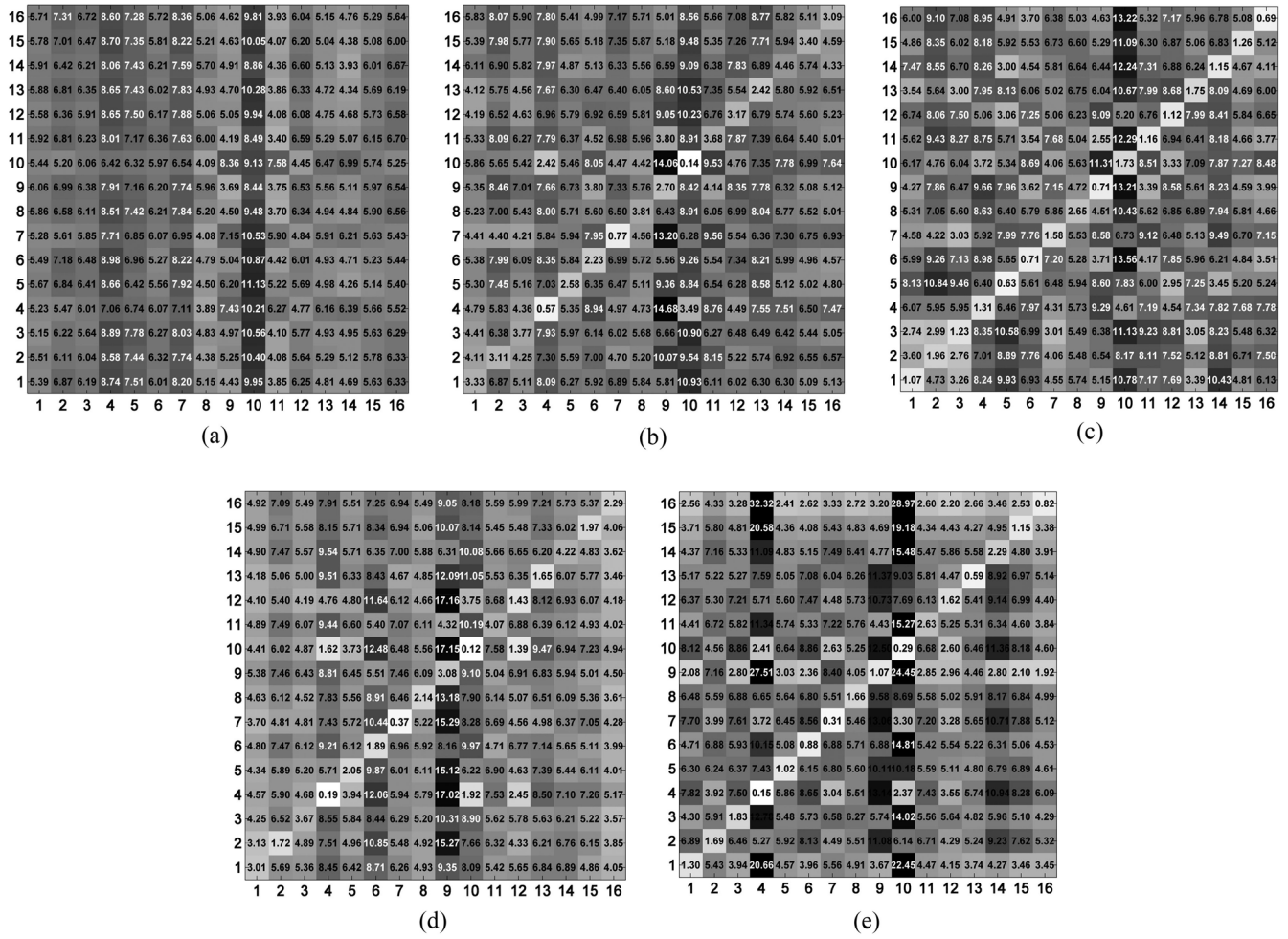


Fig. 12. Distance matrixes between noisy transformed trajectories with $\delta = 0.08$ (across bottom) and original trajectories (across left side) via (a) DI, (b) AII, (c) MAII, (d) DII, and (e) MDII. Lighter shades indicates smaller distances.

B. Sign Language Recognition

In sign language recognition, we examine the effectiveness of the integral invariants in classifying motion labels on ASL [36] dataset. 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 3-D space. In this test, as addressed in [17] and [35], we only employ the root trajectory, the average of right and left hand trajectories, to represent a sign. There is an instance of the sign word "make" as shown in Fig. 13. For test, we use the 1-NN classifier based on the DTW interdescriptor distance as defined in (21). Each time 16 classes of samples are randomly picked up in ASL dataset to run recognition test where half the samples of each class are for training and the other half are for testing. We repeat 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 II. The best recognition accuracy with 96.82% is obtained using the MAII and the accuracies reduce to 94.17% and 93.85% if using the AII and DI, respectively. While the recognition

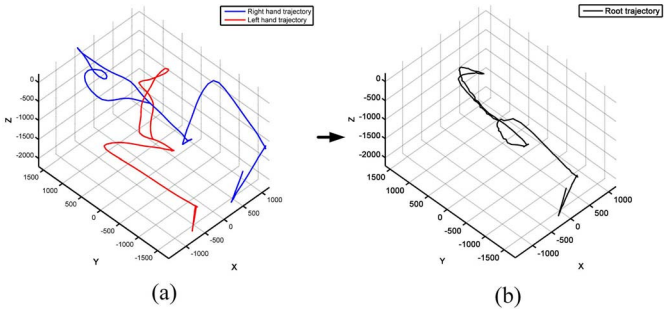


Fig. 13. Sign sample of the word make from ASL dataset. (a) Right and left hand trajectories and (b) root trajectory that is the average of the right and left hand trajectories.

performance using MAII is promising, the recognition accuracies via AII and DI are relatively lower. This is because multiscale representation can perceive more detailed information than those representations at one fixed scale (AII and DI). This fact can be demonstrated again by comparing the results based on both the DII and its multiscale representation (MDII). Nevertheless, for both the DII and MDII, the obtained accuracies are the worst. Compared

TABLE II
RECOGNITION ACCURACY VIA THE INTEGRAL INVARIANTS COMPARED WITH OTHER REPRESENTATIONS ON ASL DATASET

Representations	Accuracy
Fourier Descriptor [14] (**FD**).	78.07%
Differential Invariants [9] (**DI**).	93.85%
Distance Integral Invariant (**DII**).	58.23%
Area Integral Invariants (**AII**).	94.17%
Multiscale Distance Integral Invariants (**MDII**).	82.14%
Multiscale Area Integral Invariants (**MAII**).	**96.82%**

to hand motions in matching experiment, sign languages are more complicated so that the lack of uniqueness of representation makes DII and MDII ambiguous in recognizing sign languages.

VII. CONCLUSION

In this paper, we propose some integral invariants for motion trajectories to achieve effective and robust motion matching and recognition. Integral invariants are defined as line integrals of a kernel function along a motion trajectory, which have a smoothing effect and thus can achieve stability and robustness in the presence of noise without preprocessing the motion trajectory. We have two typical integral invariants that depend on the designed kernels to be integrated along the motion trajectory. The definition of integral invariants allows us to analyze the motion features at multiple resolutions by varying the scale of the kernel function, by which we extend the integral invariants to multiscale representation to more tap into detailed features at multiple scales in a coarse-to-fine manner. The integral invariants enjoy some useful properties including computational locality, invariance to transformation groups, and noise insensitivity. These properties make the integral invariants capable of dealing with viewpoint variations, occlusions and noise in motion trajectories.

On experimental level, we demonstrate the advantages including invariance, robustness to noise and effectiveness of the integral invariants in trajectory matching and recognition compared with other descriptors. As indicated in experiments, we have four typical integral invariants (DII, MDII, AII, and MAII) proposed in this paper, and their superiority are achieved in different situations. The insensitivity to the effects of occlusions requires the computation of invariants more locally, while the tolerance to noise requires the computation of invariants more globally. Therefore, multiscale integral invariants are not always the best representation in all situations. There are some trade-offs when choosing integral invariants or their multiscale versions to represent motion trajectories. Multiscale integral invariants are advisable for representing complex motion trajectories with minor occlusions. Integral invariants at a fixed scale are more preferable when there are relative severer occlusions in the motion trajectory. Nevertheless, in most cases, the performance via integral invariants is superior to previous invariant descriptors.

As the main computation burden for integral invariants results from the determination of MBSs, the computational cost is accordingly higher than those of the traditional descriptors such as Fourier and DI. This limits the applicability of our method in real-time scenarios. However, we can use the dynamical maintaining of convex hull to speed up the estimation of MBSs and reduce the complexity at $\mathcal{O}(n \log^2 n)$. Moreover, Faure *et al.* [20] claimed they can reduce the complexity in the determination of convex hull to linear time $\mathcal{O}(n \log n)$ that can be used to further improve the real-time performance using integral invariants, which is this paper. On the other hand, as motion tracking algorithms need to cope with noise and occlusions from the sensor, robustness is always a particular important issue in practical applications. The robustness to noise and occlusions using our method is much better than the previous methods although at a cost of higher computational complexity. Therefore, considering a tradeoff, the computation simplicity of Fourier and DI make them applicable when real-time performance is emphasized for specific applications or systems when only small noise and occlusions are present. Integral invariants will be a competitive choice when noise and occlusions are involved in motion tracking.

The integral invariants defined and investigated in this paper will also benefit the 3-D contour shape representation for object recognition, further providing rich experience and knowledge to invariant representation in computer vision areas.

REFERENCES

[1] J. Bandouch, O. C. Jenkins, and M. Beetz, "A self-training approach for visual tracking and recognition of complex human activity patterns," *Int. J. Comput. Vis.*, vol. 99, no. 2, pp. 166–189, Sep. 2012.

[2] D. Nyga, M. Tenorth, and M. Beetz, "How-models of human reaching movements in the context of everyday manipulation activities," in *Proc. IEEE Int. Conf. Robot. Autom.*, Shanghai, China, 2011, pp. 6221–6226.

[3] M. Devanne *et al.*, "3-D human action recognition by shape analysis of motion trajectories on Riemannian manifold," *IEEE Trans. Cybern.,* Doi: 10.1109/TCYB.2014.2350774.

[4] M. Bennewitz, W. Burgard, G. Cielniak, and S. Thrun, "Learning motion patterns of people for compliant robot motion," *Int. J. Robot. Res.*, vol. 24, no. 1, pp. 31–48, Jan. 2005.

[5] B. T. Morris and M. M. Trivedi, "A survey of vision-based trajectory learning and analysis for surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 8, pp. 1114–1127, Aug. 2008.

[6] W. Hu, X. Li, G. Tian, S. Maybank, and Z. Zhang, "An incremental DPMM-based method for trajectory clustering, modeling, and retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 5, pp. 1051–1065, May 2013.

[7] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," in *Pattern Recognition Techniques, Technology and Applications*, 1st ed., P.-Y. Yin, Ed. Rijek, Croatia: InTech, 2008, pp. 43–90.

[8] Y. K. Jung, K. W. Lee, and Y. S. Ho, "Content-based event retrieval using semantic scene interpretation for automated traffic surveillance," *IEEE Trans. Intell. Transp. Syst.*, vol. 2, no. 3, pp. 151–163, Sep. 2001.

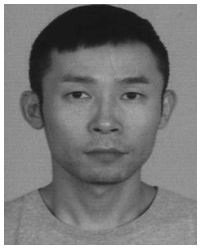
[9] S. Wu and Y. F. Li, "On signature invariants for effective motion trajectory recognition," *Int. J. Robot. Res.*, vol. 27, no. 8, pp. 895–917, 2008.

[10] F. Bashir and A. Khokhar, "Curvature scale space based affine-invariant trajectory retrieval," in *Proc. IEEE Int. Multitopic Conf.*, Lahore, Pakistan, 2004, pp. 20–25.

[11] F. Mokhtarian and A. K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 8, pp. 789–805, Aug. 1992.

[12] A. Oikonomopoulos, M. Pantic, and I. Patras, "Sparse B-spline polynomial descriptors for human activity recognition," *Image Vis. Comput.*, vol. 27, no. 12, pp. 1814–1825, 2009.

[13] G. Mori, S. Belongie, and J. Malik, "Efficient shape matching using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1832–1837, Nov. 2005.
[14] P. R. G. Harding and T. Ellis, "Recognizing hand gesture using Fourier descriptors," in *Proc. Int. Conf. Pattern Recognit.*, vol. 3. Cambridge, U.K., 2004, pp. 286–289.
[15] E. Bala and A. E. Cetin, "Computationally efficient wavelet affine invariant functions for shape recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1095–1099, Aug. 2004.
[16] F. Ofli, R. Chaudhry, G. Kurillo, R. Vidal, and R. Bajcsy, "Berkeley MHAD: A comprehensive multimodal human action database," in *Proc. IEEE Work. Appl. Comput. Vis.*, Tampa, FL, USA, 2013, pp. 53–60.
[17] Z. Shao and Y. F. Li, "A new descriptor for multiple 3D motion trajectories recognition," in *Proc. IEEE Int. Conf. Robot. Autom.,* Karlsruhe, Germany, 2013, pp. 4749–4754.
[18] T. P. Nguyen and I. Debled-Rennesson, "On the local properties of digital curves," *Int. J. Shape Model.*, vol. 14, no. 2, pp. 105–125, 2008.
[19] I. Debled-Rennesson, F. Feschet, and J. Rouyer-Degli, "Optimal blurred segments decomposition of noisy shapes in linear time," *Comput. Graph.,* vol. 30, no. 1, pp. 30–36, 2006.
[20] A. Faure, L. Buzer, and F. Feschet, "Tangential cover for thick digital curves," *Pattern Recognit.*, vol. 42, no. 10, pp. 2279–2287, 2009.
[21] S. Manay, D. Cremers, B. W. Hong, A. J. Yezzi, and S. Soatto, "Integral invariants for shape matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1602–1617, Oct. 2006.
[22] D. Xu and H. Li, "3-D curve moment invariants for curve recognition," in *Intelligent Computing in Signal Processing and Pattern Recognition*, vol. 345, D.-S. Huang, K. Li, and G. Irwin, Eds. Berlin, Germany: Springer, 2006, pp. 572–577.
[23] C.-H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 496–513, Jul. 1988.
[24] E. Calabi, P. Olver, C. Shakiban, A. Tannenbaum, and S. Haker, "Differential and numerically invariant signature curves applied to object recognition," *Int. J. Comput. Vis.*, vol. 26, no. 2, pp. 107–135, 1998.
[25] F. I. Bashir, A. A. Khokhar, and D. Schonfeld, "View-invariant motion trajectory-based activity classification and recognition," *Multimedia Syst.*, vol. 12, no. 1, pp. 45–54, 2006.
[26] T. Bai, Y. F. Li, and X. Zhou, "Learning local appearances with sparse representation for robust and fast visual tracking," *IEEE Trans. Cybern*., Doi: 10.1109/TCYB.2014.2350774.
[27] J. L. Mundy and A. Zisserman, "Numerical evaluation of differential and semi-differential invariants," in *Geometric Invariance in Computer Vision*, 1st ed. Cambridge, MA, USA: MIT Press, 1992, pp. 215–227.
[28] M. Meinard, M. Clausen, B. Eberhardt, and A. Weber, "Documentation: Mocap database HDM05," Dept. Comput. Sci., Univ. Bonn, Bonn, Germany, Tech. Rep. CG-2007-2, 2007.
[29] D. Makris and T. Ellis, "Learning semantic scene models from observing activity in visual surveillance," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 35, no. 3, pp. 397–408, Jun. 2005.
[30] M. Boutin, "Numerically invariant signature curves," *Int. J. Comput. Vis.*, vol. 40, no. 3, pp. 235–248, 2000.
[31] L. L. Rabiner and B.-H. Juang, "Pattern-comparison techniques," in *Fundamentals of Speech Recognition*, 1st ed. Englewood Cliffs, NJ, USA: Prentice Hall, 1993, pp. 141–238.
[32] C. E. Hann and M. S. Hickman, "Projective curvature and integral invariants," *Acta Appl. Math.*, vol. 74, no. 2, pp. 177–193, 2002.
[33] P. J. Olver, "Moving frames and differential invariants in centro-affine geometry," *Lobachevskii J. Math.*, vol. 31, no. 2, pp. 77–89, May 2010.
[34] P. J. Olver, "Joint invariant signatures," *Found. Comput. Math.*, vol. 1, no. 1, pp. 3–68, 2001.
[35] Z. Shao and Y. Li, "Multiscale integral invariant for motion trajectory matching and recognition," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Tianjin, China, 2014, pp. 126–131.
[36] (Nov. 4, 1998). *UCI KDD ASL Archive.* [Online]. Available: http://kdd.ics.uci.edu/databases/auslan2/auslan.html
[37] B. Hong and S. Soatto, "Shape matching using multiscale integral invariants," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 37, no. 1, pp. 151–160, Jan. 2015.
[38] L. Buzer, "A simple algorithm for digital line recognition in the general case," *Pattern Recognit.*, vol. 40, no. 6, pp. 1675–1684, 2007.



Zhanpeng Shao received the B.S. and M.S. degrees in mechanical engineering from the Xi'an University of Technology, Xi'an, China, in 2004 and 2007, respectively. He is currently pursuing the Ph.D. degree from the Department of Mechanical and Biomedical Engineering, City University of Hong Kong, Hong Kong.

From 2007 to 2011, he was an Embedded System Engineer at EVOC Intelligent Technology Company, Ltd., Hong Kong. His current research interests include pattern recognition, feature extraction, and robot vision.

Youfu Li (SM'01) received the B.S. and M.S. degrees in electrical engineering from the Harbin Institute of Technology, Harbin, China, and the Ph.D. degree in robotics from the Department of Engineering Science, University of Oxford, Oxford, U.K., in 1993.

From 1993 to 1995, he was a Research Staff at the Department of Computer Science, University of Wales, Aberystwyth, U.K. He joined the City University of Hong Kong, Hong Kong, in 1995, and is currently a Professor with the Department of Mechanical and Biomedical Engineering. His current research interests include robot sensing, robot vision, 3-D vision, and visual tracking.

Dr. Li has served as an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING and is currently serving as an Associate Editor of the IEEE ROBOTICS AND AUTOMATION MAGAZINE. He is an Editor of the IEEE Robotics and Automation Society's Conference Editorial Board, the IEEE Conference on Robotics and Automation.