2013 IEEE International Conference on Robotics and Automation (ICRA)
Karlsruhe, Germany, May 6-10, 2013

A New Descriptor for Multiple 3D Motion Trajectories Recognition*

Zhanpeng Shao, *Student Member, IEEE*, Y. F. Li, *Senior Member, IEEE*

Abstract— **Motion trajectory gives a meaningful and informative clue in characterizing the motions of human, robots or moving objects. Hence, the descriptor for motion trajectory plays an importance role in motion recognition for many robotic tasks. However, an effective and compact descriptor for multiple 3D motion trajectories under complicated situation is lacking. In this paper, we propose a novel invariant descriptor for multiple motion trajectories based on the kinematic relation among multiple moving parts. There are two kinds of kinematic relation among multiple trajectories: articulated and independent trajectories. Spherical coordinate system is introduced to get a uniform and compact representation for both kinds of trajectories, where the relative trajectory concept are firstly defined based on orientation and distance changes in favor of acquiring relative movement features of each child trajectory with respect to the root trajectory. Then, by incorporating both the differential invariants of root trajectory and orientation, distance variations of each relative trajectory respectively, the new descriptor is constructed. Finally, effectiveness and robustness of our proposed new descriptor for multiple trajectories under complex circumstance are validated by the conducted two experiments for sign language and human action recognition.**

I. Introduction

In robotics, there is a need to characterize human or robot motion pattern. Motion trajectory, a sequence of successive moving positions of objects across time series, is a very compact and representative feature for motion representation that can perceive significant changes of motions easily. Therefore, motion trajectory due to its rich spatiotemporal and dynamic information always can be used to perform motion perception, recognition, and analysis in many special tasks including gesture recognition [1], human motion and behavior recognition [2], [7], RPD (Robot Programming by Demonstration) [4], [5] in which the system extracted a robot trajectory model from multiple human demonstrations. Examples also include rigid body motion recognition [6] and articulated motion analysis [22]. As these applications suggest, there always are multiple simultaneous motions, e.g., multiple trajectories extracted from body parts such as the head, hands or feet [2], for complex motion situation. Hence, an effective and compact descriptor for multiple trajectories is quite meaningful for motion analysis and recognition.

In the related works, motion trajectory and its corresponding descriptor have been extensively studied for years. Raw data of trajectories [2], [7], which relies on the absolute position of the data points and coordinate system, were mostly used directly. In certain situation when there are noise perturbations or under different viewpoints, the raw data might be not robust and flexible to represent motions. Shape descriptors were also employed to depict the motion trajectory including centroid-contour distance (CCD) [8], curvature scale space (CSS) [9], chain code [10] and etc. However, most of shape descriptors are just suitable for simple shape representation, where they admit limited invariance in different circumstances. In addition, B-Spline [11] and NURBS [12] descriptors based on mathematical model were used to model motion trajectory, shape contour and surface. Motion trajectories, there, are represented by a set of control point parameters of the spline curve or algebraic curve, which both admit limited invariance and suffer from occlusion case. Transform function, Fourier descriptor [13] and Wavelet transform [14], are also incapable for depicting the trajectory because the local features of motion trajectory are lost during the transform and condensation procedure of data. The differential invariant signatures for motion trajectory in previous descriptor proposed in [15], [16] and [17] are able to keep invariant under Euclidean and affine group action such as translation, rotation, scaling or occlusion. Both the dynamic instance and spatiotemporal information can be describe based on differential invariants allowing it perceive salient motion features. These advantages make it more applicable to depict complex and long-term motion trajectories with similar shape performed by different actors or under different viewpoints and occlusion situation.

However, most of related descriptors just focused on the description for single curve, 2D shape, or simple motion and thus are insufficient or not compact to represent complex, long-term 3D motions such as multiple interrelated and parallel objects motions. Although multiple trajectories have been extensively employed for human action representation and recognition [2], [18], all of them used the raw data of trajectories and didn't depict the kinematic features definitely among multiple trajectories. In the present study, the theory of chaotic systems [19] is introduced to model and analyze nonlinear dynamics of human action using trajectories of reference joints as representation. Later on [20] used a large number of particle trajectories and chaotic invariant features to recognize human action. In general, all these approaches suggest that multiple trajectories extracted from multiple objects are a complex and time varying nonlinear dynamical system. A complete description for multiple trajectories, nevertheless, needs to include all the independent variables and their dependence, certain defined constraints to be satisfied by the system [19]. However, in practical scenarios a complete analytic description is extremely hard to obtain. It is evident that existing literatures are lack of a complete, flexible, efficient and robust descriptor for multiple motion trajectories so that it is able to characterize the motion pattern explicitly for complex motions, which either are the regular motion with certain kinematic constraints or arbitrary, long-term motion.

*The work was supported by a grant from the Research Grants Council of Hong Kong [Project No. CityU118311] and NSFC [61273286]

Zhanpeng Shao and Y. F. Li are with the Department of Mechanical and Biomedical Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong (phone:852-34424724; fax: 852-34420172; e-mail: perry.shao@student.cityu.edu.hk, meyfli@cityu.edu.hk).

978-1-4673-5643-5/13/$31.00 ©2013 IEEE

Thus, the question addressed in this paper is how to construct a compact and effective descriptor for multiple 3D motion trajectories.

Our contribution, consequently, in this paper is to propose a new novel descriptor for multiple 3D motion trajectories based on the differential invariants proposed in [17]. Our scheme for constructing the descriptor consists of two primary components as follows:

1. We firstly pick up a representative trajectory among multiple trajectories as root trajectory that is represented by previous descriptor [17] based on differential invariant signatures, inheriting all the advantages addressed before. Other remaining trajectories are defined as child trajectories [21].

2. To incorporate both the spatial and temporal relation simultaneously between each child trajectory and root trajectory, spherical coordinate system is introduced to depict relative movement features of each child trajectory with respect to the root trajectory, where the relative orientation and distance changes across all the time instances are involved. Finally, two representations for root and each relative trajectory respectively are incorporating into one group.

This paper is organized as follows. Section II defines relative trajectory concept and its representation based on spherical coordinate system that involves the distance constraint in two cases. Section III introduces our constructing procedure and hierarchy of the new descriptor for multiple trajectories. Section IV discussed the recognition framework in this paper. Section V presents two recognition experiments based on ASL (Australian Sign Language) [24] and HDM05 (Human motion capture dataset) [25] dataset respectively using our proposed descriptor, gives a comparison between previous descriptor and our new compact descriptor for multiple trajectories. Finally, Section VI states the conclusion and discusses the future works.

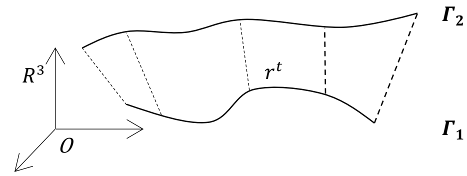
II. Representation for Relative Trajectory

A point trajectory in 3D space can be parameterized by $\boldsymbol{\Gamma}(\boldsymbol{t}) = \{\, x_t, y_t, z_t | t \in [1, N] \,\}$, where N is the number of frames (trajectory length) and $\{\, x_t, y_t, z_t \,\}$ is the Cartesian coordinate of a point at t^{th} time instant. Inspired by [21], the relative trajectory concept comes to be proposed firstly, that there are two trajectories, $\boldsymbol{\Gamma_1}$ and $\boldsymbol{\Gamma_2}$, generated by two simultaneous moving parts as shown in Fig. 1(a). If one of the trajectories, $\boldsymbol{\Gamma_1}$, is transformed to original, $\boldsymbol{O}$, another trajectory, $\boldsymbol{\Gamma_2}$, is mapped to the relative trajectory, $\Delta\boldsymbol{\Gamma}$, with respect to $\boldsymbol{\Gamma_1}$ as shown in Fig. 1(b). Relative trajectory is mathematically defined as:

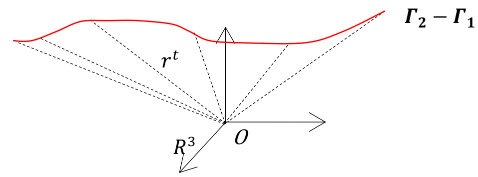
$$\Delta\boldsymbol{\Gamma} = \boldsymbol{\Gamma_2} - \boldsymbol{\Gamma_1} = \{\, \Delta x_t, \Delta y_t, \Delta z_t \mid t \in [1, N] \,\}. \quad (1)$$

Here, $\Delta\boldsymbol{\Gamma}$ actually represents the relative movement between any two simultaneously moving points. $\boldsymbol{\Gamma_1}$ and $\boldsymbol{\Gamma_2}$, are named root trajectory and child trajectory respectively [21] for convenience later.

There are two cases: Firstly, if two trajectories are moving independently without any constraint between them across all time instances they are shown in Fig. 1. Secondly, two trajectories are articulated that means child trajectory

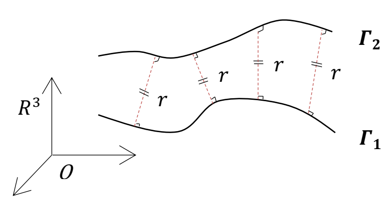


(a) Two independent trajectories without constraint.

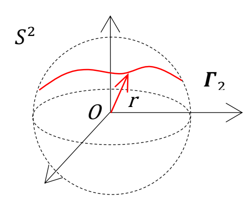


(b) Relative trajectory without constraint.

Figure 1. Relative trajectory without any constraint between two independent trajectories.



(a) Two trajectories with fixed distance.



(b) Relative trajectory with fixed distance (lie on a sphere).

Figure 2. Relative trajectory with fixed distance between two trajectories.

remains at fixed distance to root trajectory in whole occurrence as shown in Fig. 2, i.e. ,

$$\Delta x_t^2 + \Delta y_t^2 + \Delta z_t^2 = r^2, t = 1, \ldots N, \quad (2)$$

where the transformed relative trajectory $\Delta\boldsymbol{\Gamma}$ lies on a sphere with radius *r*. The representation of the relative trajectory for the articulated trajectories has to meet both the definition in (1) and additional quadratic equality constraints of (2). Thus, to represent the relative trajectory uniformly in both two cases regardless of constraint condition, instead of Cartesian coordinate representation, the spherical coordinate representation for the relative trajectory is introduced to depict the orientation and distance changes between child and root trajectory across all time instances definitely, i.e.,

$$\Delta\boldsymbol{\Gamma} = \boldsymbol{\Gamma_2} - \boldsymbol{\Gamma_1} \triangleq \{\, \theta_t, \emptyset_t, r_t \mid t \in [1, N]\}, \quad (3)$$

where θ is inclination from the z axis, $\emptyset$ is azimuth from the x axis in the xy plane, and *r is* the radius. By this representation, the space relative trajectory can be uniquely defined using two angles and one distance that also enables us to describe both the articulated trajectories and independent trajectories precisely. $\Delta\boldsymbol{\Gamma}$ is a relative trajectory lying on a sphere, as shown in Fig. 2 for articulated trajectories if r_t keep constant across all time instances, and an arbitrary relative trajectory without distance constraint as shown in Fig. 1 if r_t varies along all time series. Furthermore, this compact representation can satisfy the temporal and spatial constraints simultaneously. How to reach a smooth $(\theta_t, \emptyset_t, r_t)$ in spherical coordinate system here is straightforward that can be found in [3].It should be noted that the angles $\theta_t, \emptyset_t$are limited to $-\pi$ to π.

III. Compact Descriptor

This section presents the details of how to construct the new descriptor for multiple trajectories. In order to systematically construct a flexible and robust descriptor, we need to investigate the contexts difference between single and multiple parts motions. Simultaneous motions of multiple parts may be regarded as several individual movements independently, where modeling multiple motions is just by concatenating descriptors of all the individual trajectories piece to piece directly base on differential invariants not considering the kinematic relation among them [16]. As a result, the corresponding perception and recognition tasks hence have to only repeat one by one until all the trajectories are exhausted, giving rise to low efficient computation and bad performance of overall system. These stacked descriptors for multiple trajectories, however, are not practical and redundant in real world applications, where most of multiple objects move regularly with relative kinematic constraints or mutually keep certain spatiotemporal relations in a way that semantically represent certain motion intention [22]. Therefore, a compact, flexible, and efficient description for multiple trajectories is expected to be constructed systematically.

A. Descriptor Framework for Multiple Trajectories

There are multiple trajectories extracted from multiple concurrent moving objects such as multiple points from a rigid body, joints from articulated human body as shown in Fig. 3. Firstly, while one representative sample from multiple trajectories is picked up as root trajectory that is represented using the differential invariant signatures [17], the rest of trajectories are defined as child trajectories as shown in Fig. 4. There, root trajectory descriptor is defined as follows:

$$\overline{\boldsymbol{S}} = \{k(t), k_s(t),\ \tau(t), \tau_s(t),\ s(t),\ r(t) \mid t \in [1, N]\}, \quad (4)$$

where

$$k(t) = \left\| \Gamma'(t) \times \Gamma''(t) \right\| / \left\| \Gamma'(t) \right\|^3, \quad (5)$$

$$\tau(t) = (\Gamma'(t) \times \Gamma''(t)) \cdot \Gamma'''(t) / \left\| \Gamma'(t) \times \Gamma''(t) \right\|, \quad (6)$$

$$k_s(t) = \frac{dk(t)}{ds} = \frac{dk(t)}{dt} \cdot \frac{dt}{ds} = \frac{k'(t)}{\left\| \Gamma'(t) \right\|}, \quad (7)$$

$$\tau_s(t) = \frac{d\tau(t)}{ds} = \frac{d\tau(t)}{dt} \cdot \frac{dt}{ds} = \frac{\tau'(t)}{\left\| \Gamma'(t) \right\|}, \quad (8)$$

$$s(t) = \int_1^t \left\| \Gamma'(t) \right\| dt, \quad (9)$$

$$r(t) = \left\| \Gamma(t) - c \right\| = \left\| \Gamma(t) - \left(\int_1^N \Gamma(t) dt\right) / N \right\|. \quad (10)$$

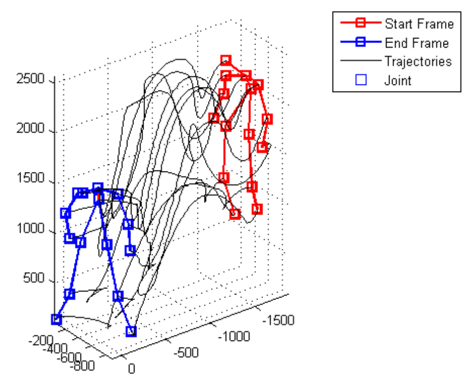


Figure 3. Sample sequences of "Jump down" classes from human motion capture dataset [25]. The red stick figure with joints depicts the start frame of the sequence, while blue stick figure with joints represents the end frame.

The parameters defined in (5-10) are the differential invariants: curvature k, torsion τ, and there derivatives k_s and τ_s, here, s is arc length from the beginning of the trajectory to the present point and r represents the distance from the present point to the geometric center c. The descriptor components depending on high order derivative and integral cannot be calculated directly due to its sensitivity to noise since the trajectory is a discrete sequence normally. To avoid this effect and make the descriptor robust, the approximation approaches in [16], [17] are employed to implement the descriptor computation, where its strong invariance has been verified under translation, scaling, rotation and occlusion. Secondly, based on relative trajectory concept in Section II, the relative trajectory representation of each child trajectory with respect to the root trajectory is obtained. Once we get all the relative trajectories and root trajectory representation, the compact descriptor is constructed by integrating them into one group for representing multiple trajectories that is actually a time sequence vector as follows:

$$\boldsymbol{F} = \{\overline{\boldsymbol{S}}, \Delta\boldsymbol{\Gamma}_1, \Delta\boldsymbol{\Gamma}_2, \dots, \Delta\boldsymbol{\Gamma}_i, \dots, \Delta\boldsymbol{\Gamma}_M\}, \quad (11)$$

where $\Delta\boldsymbol{\Gamma}_i$ is relative trajectory represented by (3), and there are $\boldsymbol{M}$ relative trajectories as show in Fig. 4. Here $\overline{\boldsymbol{S}}$ is the differential invariants descriptor of root trajectory.

Only orientation changes are preserved in $\Delta\boldsymbol{\Gamma}$ while the distance parameter r can be dropped in two situations, where relative trajectory lies on a sphere with fixed distance constraint or any two child trajectories keep articulated making one motion freedom of the trajectory lost. Also, the fixed distance r in multiple trajectories representation is not significant component for both perception and recognition tasks. As a result, here the final $\boldsymbol{F}$ descriptor is constructed in a compact way that only root trajectory is described uniquely and completely using differential invariants while the child trajectories are represented based on relative kinematic constraints between multiple trajectories. Thus, the new descriptor, $\boldsymbol{F}$, for multiple trajectories show more compact than the naive descriptor that may concatenated the trajectory piece by piece Moreover, $\boldsymbol{F}$ also inherits the invariance of previous differential invariants descriptor but needs to be normalized further in next section.

B. Preprocessing and Normalization

In order to reduce the level of noise in trajectory data different smoothing tools may be applied to filter the trajectory including B-spline smoothing, moving average and wavelet transform or other common smoothing approaches. While such tools are performed based on only the geometric aspects of trajectories data, it is often difficult to take into consideration time-related aspects. Thus, we employ the Kalman-based smoothing, a well-known time-based scheme for time sequence data in presence of noise and outlier. This allows us to take into consideration kinematic parameters.

Trajectories in 3D space may admit a similar motion pattern but differ in location, orientation, and scale due to different sampling rates and lengths, start position, affine transformation. To overcome and solve these ambiguities, the descriptor constructed before thus has to be normalized to keep invariance and uniform sale under different situations. The normalization of differential invariants for root trajectory, $\overline{\boldsymbol{S}}$, has been discussed to get $\overline{\boldsymbol{S}}^*$ in [17] and we only present the normalization of $\Delta\boldsymbol{\Gamma}$ in this paper.

We need to transform the relative trajectory discussed in Section II to canonical spherical coordinate representation by normalized the orientation $(\theta_t, \emptyset_t)$ and distance r_t respectively, where the orientation $(\theta_t, \emptyset_t)$ and distance r_t for relative trajectory are both normalized by averaging all the sampled points and taking into consideration start position defined as:

$$r_t^* = \frac{N \cdot r_t}{\sum_{t=1}^{N} r_t}, \tag{12}$$

$$\theta_t^* = \frac{N \cdot (\theta_t - \theta_1)}{\sum_{t=1}^{N} \|\theta_t - \theta_1\|}, \tag{13}$$

$$\phi_t^* = \frac{N \cdot (\phi_t - \phi_1)}{\sum_{t=1}^{N} \|\phi_t - \phi_1\|}. \tag{14}$$

Once normalization procedure is implemented, the descriptor, $\Delta\boldsymbol{\Gamma}^* = \{\theta_t^*, \emptyset_t^*, r_t^*\}$, shows strong invariance in different cases as shown in Fig. 5. As an example we provide a small snapshot of ASL (Australian Sign Language) dataset, where example trajectories from "wrong" class in the ASL dataset are performed by different actors. The corresponding orientation and distance changes are also shown in the second and third column of the figure respectively. As evident from this figure, the orientation and distance changes with different actors and length from one class perform the similar variation and shape. Finally, we get the normalized descriptor:

$$\boldsymbol{F}^* = \{\overline{\boldsymbol{S}}^*, \Delta\boldsymbol{\Gamma}_1^*, \Delta\boldsymbol{\Gamma}_2^*, \ldots, \Delta\boldsymbol{\Gamma}_i^*, \ldots, \Delta\boldsymbol{\Gamma}_M^*\}. \tag{15}$$

IV. DTW-Based Recognition

Dynamic time warping (DTW) [23] is an algorithm for measuring similarity between two elastic time sequences that may vary in time or speed. Here, it is effective for multiple trajectories group to overcome the diversity in motion speed and sampling rates, distributions. In this paper, we employ this nonlinear alignment method for motion recognition using new compact descriptor for multiple trajectories. 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:

$$\mathrm{D}(m,n) = \min\{D(m-1,n), D(m,n-1), D(m-1,n-1)\} + d(m,n), \tag{16}$$

where $d(m,n)$ denotes the distance of sample m and n from two nonlinear time sequences in multidimensional space, and $D(m,n)$ is the sum distance that can serve as similarity measurement between two time sequences.

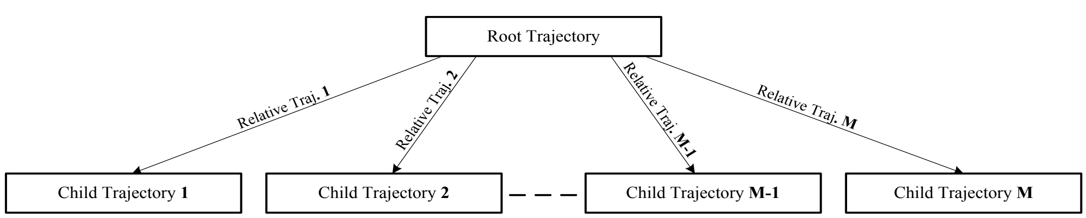
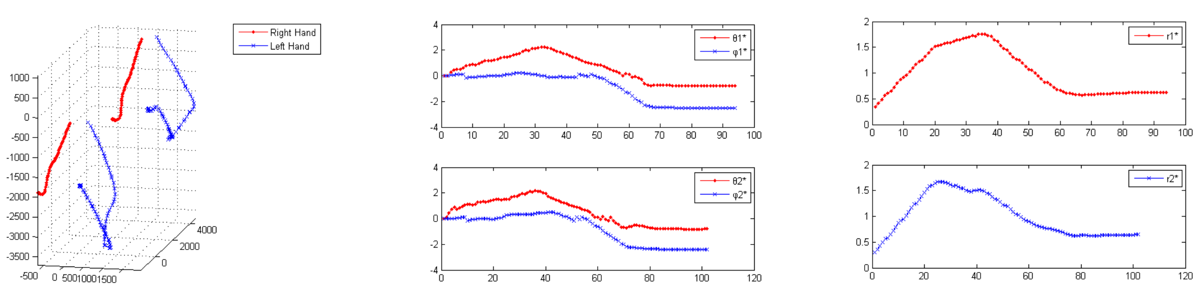


Figure 4. Multiple trajectories hierarchy



(a) One class with different 2 actors (b) Orientation changes (c) Distance changes

Figure 5. Two groups of example trajectories from "wrong" class in ASL dataset performed by different actors using two hands. The orientation and distance changes between left-hand and right-hand for two examples are shown in (b) and (c) respectively. It shows the similarity between two examples from one class.

For two groups of multiple trajectories, corresponding descriptors are $\boldsymbol{F}_A^*$ and $\boldsymbol{F}_B^*$ with respective lengths M and N. The distance (cost function) between samples $\boldsymbol{F}_A^{*m}$ and $\boldsymbol{F}_A^{*n}$ based on Euclidean distance is defined as follows:

$$d(m,n) = \Delta F^* =$$
$$\frac{\Delta k^{*m,n} \cdot \Delta \tau^{*m,n} \cdot (1+\lambda_1 \cdot \Delta h^{*m,n})}{\sqrt{(S^{*m})^2} \cdot \sqrt{(S^{*n})^2}} + \lambda_2 \cdot \Delta o^{*m,n} + \lambda_3 \cdot \Delta l^{*m,n}, \quad (17)$$

where

$$\Delta k^{*m,n} = \left\| (k^{*m}, k_s^{*m}) - (k^{*n}, k_s^{*n}) \right\|, \quad (18)$$

$$\Delta \tau^{*m,n} = \left\| (\tau^{*m}, \tau_s^{*m}) - (\tau^{*n}, \tau_s^{*n}) \right\|, \quad (19)$$

$$(S^{*m})^2 = (k^{*m})^2 + (k_s^{*m})^2 + (\tau^{*m})^2 + (\tau_s^{*m})^2, \quad (20)$$

$$(S^{*n})^2 = (k^{*n})^2 + (k_s^{*n})^2 + (\tau^{*n})^2 + (\tau_s^{*n})^2, \quad (21)$$

$$\Delta h^{*m,n} = \left\| (s_m^*, r_m^*) - (s_n^*, r_n^*) \right\|, \quad (22)$$

$$\Delta o^{*m,n} = \left\| (\theta_m^*, \phi_m^*) - (\theta_n^*, \phi_n^*) \right\|, \quad (23)$$

$$\Delta l^{*m,n} = \left\| r_m^* - r_n^* \right\|. \quad (24)$$

Equation (18-22) are the same as the definition in [15], [17] with the differential invariants for root trajectory of two groups, while distance between orientation and distance changes of two groups is defined in (23-24). Here the parameters λ_1, λ_2 and λ_3 are the weights of global shape, orientation and distance invariants respectively. The training method for weights λ_{1-3} is performed in same way with [17] to get the satisfied recognition performance for specific dataset.

V. Experiments

In the experiments, we verify the proposed compact descriptor by recognizing 3D trajectories of different classes in a large dataset [24], [25], which are high-quality Australian sign language and Human motion capture dataset respectively. Here, the DTW-based algorithm discussed in Section III is used to implement both the experiments. In addition, the previous descriptor [17] is also employed to implement the recognition experiments for comparison.

A. Sign Language Recognition of Different People

In first experiment, the ASL (Australian Sign Language) dataset are used to implement recognition task for classifying different sign classes. The dataset consists of 2565 samples of ASL signs, where 27 samples of each of 95 Australian sign classes were captured from a native signer with high-quality data and each sample is performed by moving the right-hand and left-hand simultaneously in 3-dimensional space. In this test, while the left-hand trajectory is picked up as root trajectory, the right-hand trajectory serves as child trajectory resulting in the new descriptor $\boldsymbol{F}^*$, 9-dimensional feature vector per sign. For testing, we used the cross validation approach with 1-NN classifier based on the DTW inter-descriptor distances as defined in (16-17), where both the signs classes and their samples were randomly picked up and recognized in our method. Half of the samples are further for training and the other half are for testing, where we repeated this procedure 50 times. In same way, the previous descriptor based on differential invariants is also tested by concatenating the left-hand and right-hand trajectory piece to piece directly. The average ratios of correct recognition for both descriptors are listed in Table I for different numbers of classes respectively. Fig. 6 shows the confusion degree among signs from eight classes. The lighter the cell is, the higher similarity is indicated. From the confusion table, it is evident that our new compact descriptor can obtain more discriminative result than previous descriptor, especially for classes 6-8 that are hardly distinguished from other class with classification ambiguity.

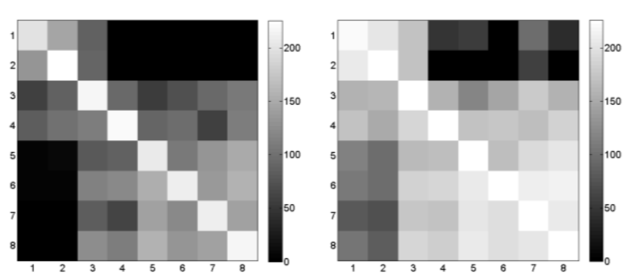
B. Human Action Recognition

The second set of experiment is performed on the Human motion capture dataset (HDM05) [25] containing 3D motion sequence, which are employed here to verify the effectiveness of the new compact descriptor in human action recognition. Human action consist of rich spatiotemporal patterns , where the 3D multiple trajectories are extracted from joints, e.g., elbow, knee as well as hand and feet that provide more accurate recognition of human action as shown in Fig. 3. The HDM05 dataset consists of around 100 different motion classes performed by five different actors. Most of these classes contain 10 and 50 different realizations for each action class amounting to 1457 smaller motion clips. These realizations may be slightly different from each other but semantically have the same motion pattern. Therefore, wide range of actions and variation for same action make human action recognition to be a very challenging task.

The initial input is in the form of 3-dimensional trajectories of several body joints of the stick figure shown in Fig. 3, but we use the 9 typical reference joints where the "Sternum" joint trajectory is selected as root sequences and other 8 reference joints serve as child trajectories per action. These raw data are then used to construct the new compact descriptor $\boldsymbol{F}^*$ according to our proposed procedure in Section III. It should be noted that there are fixed distance constraints among certain human joints including elbow and hand, knee and ankle. Thus, dropping the distance r parameters for corresponding relative trajectory here makes the new descriptor $\boldsymbol{F}^*$, 20-dimensional feature vector per action, more compact with eliminating the abundant information. For

TABLE I. AVERAGE RATIO OF CORRECT RECOGNITION(%)

Number of Classes	**2**	**4**	**8**
Previous Descriptor	91.23	87.89	87.71
New Descriptor	95.31	93.79	92.62

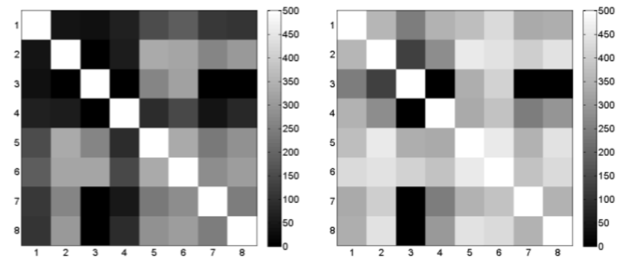


(a) (b)

Figure 6. Confusion degree among 8 classes of signs with different actors under our new descriptor in (a) and previous differential invariants descriptor in (b) for ASL dataset

TABLE II. AVERAGE RATIO OF CORRECT RECOGNITION(%)

Number of Classes	**2**	**4**	**8**
Previous Descriptor	95.01	83.39	73.11
New Descriptor	97.51	90.31	80.31



(a) (b)

Figure 7. Confusion degree among 8 action classes with different actors under our new descriptor in (a) and previous differential invariants descriptor in (b) for HDM05 dataset.

testing, we use the cross validation approach with 1-NN classifier in same way that ASL recognition experiment picks up training and test data randomly. Also, the previous descriptor based on differential invariants is tested by concatenating the each reference joint trajectory piece to piece directly. The classification results achieved by this approach are shown in Table II. In same way, the confusion degree is shown in Fig. 7. The discrimination of human motion recognition using new compact descriptor shows strong robustness that makes it more effective for recognition.

VI. CONCLUSION

In this paper, we proposed a new descriptor for multiple trajectories in motion recognition based on relative trajectory concept with kinematic constraint, which is invariant for group transformation. Using the new compact descriptor, the complicated motions are completely represented and the kinematic information is preserved. Experimental validation of the feasibility and potential advantages of new descriptor in performing motion recognition are demonstrated on two large and complex motion datasets.

In the future work, the high-level recognition engine will firstly be investigated based on the new descriptor due to an extra computational cost for DTW-based method. Moreover, the new descriptor for multiple trajectories contains rich kinematic information in 3-dimensional space, where we can further investigate the motion pattern perception for rigid body using the new descriptor and relative trajectory concept.

REFERENCES

[1] S. Waldherr, R. Romero, and S. Thrun, “A gesture based interface for human-robot interaction,” *Autonomous Robots*, vol. 9, no. 2, pp. 151-173, Sep. 2000.

[2] J. Min, R. Kasturi and O. Camps, "Extraction and temporal segmentation of multiple motion trajectories in human motion," *Image Vision Comput*., vol. 26 (12), pp. 1621-1635, 2008.

[3] Faria, D.R.; Dias, J., "3D hand trajectory segmentation by curvatures and hand orientation for classification through a probabilistic approach," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, St. Louis, MO, USA. 2009, pp. 1284-1289.

[4] Aleotti, J., and Caselli, S., “Trajectory Clustering and Stochastic Approximation for Robot Programming by Demonstration,” *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Edmonton, Canada, 2005, pp. 1029–1034.

[5] H. Friedrich, S. Munch, R. Dillmann, S. Bocionek and M. Sassin, "Robot programming by demonstration (RPD): Supporting the induction by human interaction," *Machine. Learning*, vol. 23, pp. 163-189, 1996.

[6] J. De Schutter, E. Di Lello, J. F. M. De Schutter, R. Matthysen, T. Benoit and T. De Laet, "Recognition of 6 DOF rigid body motion trajectories using a coordinate-free representation," *In Robotics and Automation (ICRA), 2011 IEEE International Conference*, Shanghai, 2011, pp. 2071-2078.

[7] A. Psarrou, S. Gong and M. Walter, "Recognition of human gestures and behaviour based on motion trajectories," *Image Vision Comput*., vol. 20, pp. 349-358, 2002.

[8] F. Bashir, A. Khokhar and D. Schonfeld, "View-invariant motion trajectory-based activity classification and recognition," *Multimedia Systems*, vol. 12, pp. 45-54, 2006.

[9] Bashir, F., Khokhar, A., “Curvature scale space based affine invariant trajectory retrieval,” *In: IEEE International Multi topic Conference, INMIC 2004*. Lahore, Pakistan, 2004, pp. 20-25.

[10] E. Bribiesca, “A chain code for representing 3D curves,” *Pattern Recognition*, vol. 33 no. 5, pp. 755-765, 2000.

[11] F. S. Cohen, Z. Huang, and Z. Yang, “Invariant matching and identification of curves using B-splines curve representation,” *IEEE Transactions on Image Processing*, vol. 4, no. 1, pp. 1–10, 1995.

[12] S. Y. Chen and Q. Guan, “Parametric shape representation by a deformable NURBS model for cardiac functional measurements,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 480–487, 2011.

[13] P. R. G. Harding and T. J. Ellis, “Recognizing hand gesture using Fourier descriptors,” *in Proc. Int. Conf. Pattern Recognition*, Cambridge, UK, , 2004, vol. 3, pp. 286-289.

[14] C. Cattani, "Shannon Wavelets for the Solution of Integrodifferential Equations," *Mathematical Problems in Engineering*, vol. 2010, pp. 1, 2010.

[15] S. D. Wu and Y. F. Li, “On signature invariants for effective motion trajectory recognition,” *International Journal of Robotics Research*, vol. 27, no. 8, pp. 895–917, 2008.

[16] S. D. Wu and Y. F. Li, “Flexible signature descriptions for adaptive motion trajectory representation, perception and recognition,” *Pattern Recognition*, vol. 42, no. 1, pp. 194–214, 2009.

[17] Jianyu Yang, Y. F. Li, Keyi Wang, Yuan Wu, Giuseppe Altieri, and Massimo Scalia, “Mixed Signature: An Invariant Descriptor for 3D Motion Trajectory Perception and Recognition,” *Mathematical Problems in Engineering*, vol. 2012, Article ID 613939, 29 pages, 2012. doi:10.1155/2012/613939.

[18] X. Ma, F. Bashir, A. Khokhar and D. Schonfeld, "Event Analysis Based on Multiple Interactive Motion Trajectories," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, pp. 397-406, 2009.

[19] S. Ali, A. Basharat, and M. Shah, “Chaotic invariants for human action recognition,” *In Proc. IEEE International Conf. on Computer Vision,* Rio de Janeiro, Brazil, 2007.

[20] S. D. Wu, O. Oreifej, and M. Shah, “Action recognition in videos acquired by a moving camera using motion decomposition of lagrangian particle trajectories,” *In Proc. IEEE International Conf. on Computer Vision*, Barcelona, 2011.

[21] Park, H.S., Sheikh, Y.A.: “3D reconstruction of a smooth articulated trajectory from a monocular image sequence.” *In: Proc. IEEE International Conf. on Computer Vision,* Barcelona, 2011, pp. 201-208.

[22] Kirk, A. G., O’Brien, J. F., & Forsyth, D. A., “Skeletal parameter estimation from optical motion capture data,” *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Los Alamitos, 2005.

[23] Sakoe, H. and Chiba, S., “Dynamic programming algorithm optimization for spoken word recognition,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), pp. 43-49, 1978.

[24] UCI KDD ASL Archive, [Online], Available: http://hdd.ics.uci.edu/databases/auslan2/auslan.html.

[25] M. Müller, T. Röder, M. Clausen, B. Eberhardt, B. Krüger, A. Weber. “Documentation: Mocap database HDM05,” *Technical report CG-2007-2*, Universität Bonn, 2007.