2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

A Hierarchical Motion Trajectory Signature Descriptor

Shandong Wu, Y. F. Li and Jianwei Zhang

Abstract— **Motion trajectory is a compact clue for motion characterization. However, it is normally used directly in its raw data form in most work and effective trajectory description is lacking. In this paper, we propose a novel hierarchical motion trajectory signature descriptor, which can not only fully capture motion features for detailed perception, but also can be used for probabilistic fast recognition. The hierarchy enables the signature to exhibit high functional adaptability meeting different application requirements. At the first-level, differential invariants are employed to describe trajectory features and a nonlinear signature warping method is developed to perceive and recognize trajectories. The second-level signature is the condensation of the first-level signature by applying PCA based dimension optimization. It behaves more efficiently in recognition based on the Gaussian Mixture modeling and Bayesian classifier. The conducted experiments verified the signature's effectiveness.**

I. INTRODUCTION

Motion trajectory is a meaningful and informative clue in characterizing the motions of robot and human. No matter for simple actions, median behaviors or complex activities, the consequent motions can be characterized by identifying the involved subjects (human body, head, hands, or feet etc. [1]) and extracting the underlying motion trajectories. That is, motion can be analyzed spatiotemporally by the joint description of spatially parallel or (and) temporally sequential motion trajectory compositions. While the parallel trajectories entail integrated motion of multiple subjects moving at the same time, it is the sequential trajectory that captures the continuous motion. In this sense, describing, recognizing and perceiving motion trajectories are important for various motion analysis and applications. For example, in robot Programming by Demonstration (PbD), many kinds of human demonstrations can be well described by the extracted motion trajectories. Therefore, it is valuable to study motion trajectory in terms of effective description structure, fast recognition algorithm and semantic scene perception.

In the literature, motion trajectory has been explored intensively in diverse contexts [2], [3], [4]. However, in most work, trajectory was just used directly in its raw data form, which is quite inflexible and incapable of meeting the demands of effective description, recognition and perception. On the contrary, good motion trajectory descriptor can outperform the raw trajectory data.

In the existing work, some shape descriptors have been built [5], [6]. However, most of them do not perform as well as expected in descriptive capabilities especially in the functional adaptability. Simple contour functions such as chain code, centroid-contour distance (CCD) and R-S curve just admit ordinary performance. The descriptors based on Fourier Descriptor (FD) [7], wavelet coefficient [8] and Curvature Scale Space (CSS) images [9] can represent shape in a coarse-to-fine or multi-resolution manner, in which just partially salient features such as wavelet skeleton, the lower frequency information in FD and the curvature zero-crossing points in CSS are of concern for shape description. This explains why they actually are unable to represent shapes uniquely. Also, it may be undesirable to ignore much amount of less important information when the detailed features really matter. For example, in Fourier transform, it is difficult to perform local motion analysis in the frequency field because the time information is lost. The correspondence problem has to be regulated in CSS since the curve length shrinks in the Gaussian evolving. The algebraic curve and moment function [10] suffer from occlusion as they make use of global features. The mathematical curve NURBS [11] and B-spline [12] need a fitting process that causes inaccuracy. The B-spline method may result in recognition ambiguity as it is hard to compare B-spline parameters directly for recognition because a piece of trajectory is not uniquely described by a single set of control points [12].

The quality of a descriptor should be evaluated in terms of multiple functions. In essence, a descriptor's performance depends much on the kind of shape feature interested and the descriptive structure. In this paper, we propose a novel hierarchical trajectory signature descriptor. The signature is with a two-level structure hence admits higher functional adaptability. At the first level, motion trajectory is fully modeled using local differential invariants features, which is effective for detailed motion feature capturing and perception. At the second level, the full signature data is condensed to pursue faster recognition by introducing probabilistic methods.

II. THE SIGNATURE'S STRUCTURE

In this section, we brief the basic structure of the hierarchical signature descriptor. As illustrated in Fig. 1, the signature has two levels. While the first level is the full description to the entire raw trajectory data, the second level is the condensed description of the first-level signature data.

This work was supported by grants from the Research Grants Council of Hong Kong [Project No. CityU117106 and CityU117507]

Shandong Wu and Y. F. Li are with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Kowloon, Hong Kong s.d.wu@student.cityu.edu.hk, meyfli@cityu.edu.hk

Jianwei Zhang is with the Department of Informatics, University of Hamburg, Vogt-Kölln-Str. 30, D-22527, Hamburg, Germany zhang@informatik.uni-hamburg.de

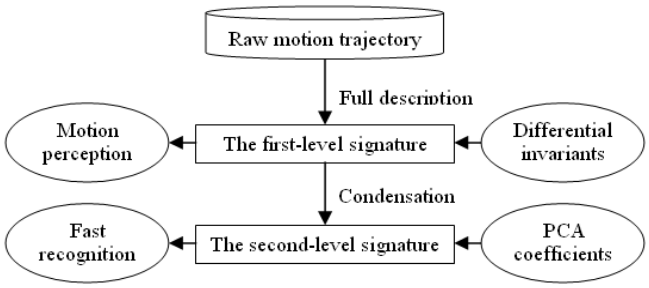

Fig. 1. The structure of the hierarchical motion trajectory signature

At the first level, Euclidean differential invariants features are employed to build a loose trajectory signature. As differential invariants are typical local features, they are not only capable of characterizing trajectory shapes, but also can offer richer invariants in trajectory representation with respect to rigid, metric and viewpoint change and insensitivity to occlusion. Focusing on the detailed motion features captured by the first-level signature, motion perception to different trajectory instances can be achieved readily via the visualization of the nonlinear inter-trajectory warping paths. The first-level signature is particularly appropriate for small scale application concerning both trajectory recognition and motion perception.

The second level of the signature is motivated by reducing the redundancy data in the first-level signature to obtain a compact signature description towards faster trajectory recognition. Applying PCA transform to the first-level signature, we get the second-level signature with less data but preserving most variance of the trajectory features. Further, based on the Gaussian Mixture modeling to the second-level signature, a Bayesian classifier is developed to replace the time-consuming matching algorithm. These resorts can speed much up the recognition especially for larger scale database.

III. THE FIRST-LEVEL SIGNATURE DESCRIPTION

A. Signature Definition and Robust Implementation

Motivated by the 2-D still curve representation [13], we have proposed a Euclidean 3-D trajectory signature definition based on typical differential invariants [14].

Definition 1. For a 3-D motion trajectory $\Gamma(t)$ parameterized by $\Gamma(t) = \{X(t), Y(t), Z(t)|t \in [1, N]\}$ (t is temporal index), its Euclidean signature S is defined in terms of curvature (κ), torsion (τ) and their first order derivatives with respect to Euclidean arc-length s ($\kappa_s = d\kappa/ds$ and $\tau_s = d\tau/ds$), in the following form,

$$S = \{[\kappa(t), \tau(t), \kappa_s(t), \tau_s(t)]|t \in [1, N]\} \quad (1)$$

To avoid the noise-sensitive high order derivatives in the accurate signature formula, an approximate signature S^* was implemented numerically by employing the joint Euclidean invariants (inter-point Euclidean distances). See our previous work [14] for the robust approximate signature calculation. We define the approximate signature as the first-level signature, which is diagrammatized by a curvature sub-signature (κ vs. κ_s) and a torsion sub-signature (τ vs. τ_s). Taking the trajectory Ω shown in Fig. 2 as an example, its first-level signature curves are illustrated in Fig. 3.

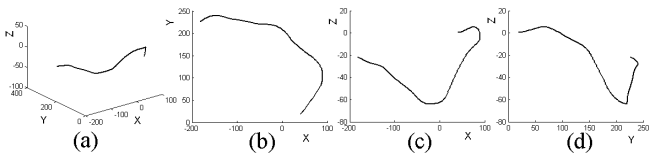

Fig. 2. A piece of trajectory Ω in 3-D form (a) and the views in plane $x-y$ (b), $x-z$ (c) and $y-z$ (d) respectively

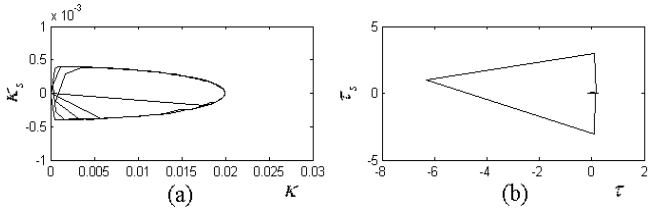

Fig. 3. The first-level signature of trajectory Ω. (a) Curvature sub-signature. (b) Torsion sub-signature

Besides the signature approximation, trajectory smoothing is also an effective method to enhance the signature's robustness by noise reduction. There are various smoothing methods like the anisotropic Perona-Malik diffusion algorithm [15]. However, while trajectory is smoothed, its shape may be affected too. Therefore, trajectory smoothing and shape-preserving have to be balanced. We design two trajectory smoothers, a moving average filter and a wavelet smoother. The smoothing parameters of the two smoothers can be interactively tuned by experience with a tradeoff between smoothing and shape-preserving. Fig. 4 demonstrates the trajectory smoothing to noisy trajectories. It can be observed that the smoothing effects are good and the shape deviations are also acceptable.

B. Motion Perception from the First-level Signature

As the first-level signature is based on the features extracted from all the sampled points, the signature data relies on the length and point distribution of a motion trajectory. However, as shown in the next, besides being used for trajectory recognition, the first-level signature is particularly useful for motion perception.

1) Motion Features Captured by the First-level Signature: The loose first-level signature provides an opportunity to

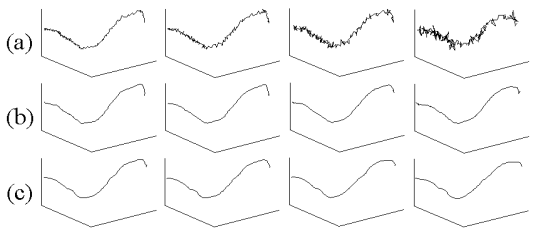

Fig. 4. Smoothing to noisy trajectories (a) using the moving filter (span = 11) (b) and the wavelet smoother (wavelet DB4 and the coefficients at level 2-5 respectively) (c)

perceive motion features, which is meaningful for human motion analysis for example in robot learning circumstance. On one hand, for various instances of the same one motion pattern, the perception to certain specific feature of an instance will make it intuitively distinctive and discriminative. In practice, motion inconsistency exists among multiple instances because robot or human cannot perform a motion exactly the same each time. On the other hand, the results of motion perception accordingly will be useful for robot to characterize the specific demonstrator (human), that is based on the fact that each different human may have specifically his (her) own motion characteristics embodied in the conducted motions. Motion perception can be mostly achieved in terms of the features attached in a motion trajectory, which can be extracted from the first-level signature description.

(1) Motion length. Measured in terms of the number of frames or sampling points, trajectory length can characterize the duration and spatial range of a motion. Attributed to the signature's computational locality, length perception to a signature is equivalent to the corresponding trajectory.

(2) Shape of motion. Basically, shape is one of the most important factors for motion pattern discrimination. From the curvature and torsion signature profiles, motion shape can be readily perceived. Intuitively, curvature measures how far a trajectory is from being on a straight line and the torsion measures how far it is from being in a plane.

(3) Motion speed. Speed is a key motion feature. User may perform the same motion with different speed profiles. Slower speed results in denser data sampled and faster speed with sparser data. Thus the speed feature can be perceived based on the points' distribution along trajectories.

(4) Sampling rate. The sampling rate of a tracking system may be unfixed (changing proportionally or randomly). Sometimes this is concerned and can be measured from the points' distribution.

(5) Occlusion. Occlusion often happens due to for example beyond viewing field or discontinuous tracking. According to the signature's computational locality, occlusion will only result in the signature shorter than the original in length. Thus it can be perceived from signature's length.

(6) Motion symmetry. Motion symmetry is an informative feature in certain motion patterns. We define the motion symmetry as two adjacent trajectory fragments symmetrical with respect to a central symmetrical point which connects these two fragments. Since the first-level signature can capture both shape and arc-length features, the central symmetry point can be easily detected by examining if $\kappa(t) = 0$ and $\tau(t) = 0$. Meanwhile, the symmetrical neighbor points around a central symmetry point satisfy the following conditions: $\kappa(t-1) = -\kappa(t+1)$, $\kappa_s(t-1) = -\kappa_s(t+1)$, $\tau(t-1) = -\tau(t+1)$ and $\tau_s(t-1) = -\tau_s(t+1)$.

2) DTW Based Nonlinear Inter-signature Matching: In the following, we present a method for the fully matching of two first-level signatures, through which the similarity of the two signatures can be measured, and the motion features along trajectory can also be highlighted for intuitive perception based on the inter-signature matching results.

According to the analysis in subsection B-1), we can infer that two signatures cannot be matched directly due to the possible difference in signature length, sampling point distribution and spatiotemporal shift of corresponding points, which may result from the variations in motion speed, sampling rate, occlusion or user's inconsistency in motion repeats. For example, Fig. 5 illustrates six instances (in solid lines) of the sampled point sequences for the same one motion. Therefore, it is crucial to customize a suitable inter-signature similarity metric to account for the consequences caused by above-mentioned factors for the signature matching.

The signatures' matching problem is analogous with the elastic time series/sequences comparison [16]. The key of an appropriate similarity measurement lies in finding the best matching of the element pairs along two signatures. Thus we adopted the Dynamic Time Warping (DTW) method to do nonlinear inter-signature matching [14]. For two first-level signatures S^{*i} and S^{*j} with respective length P and Q, let $S^{*i} = \{[\kappa^{*i}, \kappa_s^{*i}, \tau^{*i}, \tau_s^{*i}]_p | p \in [1, P]\}$ and $S^{*j} = \{[\kappa^{*j}, \kappa_s^{*j}, \tau^{*j}, \tau_s^{*j}]_q | q \in [1, Q]\}$ represent the sequences of signature quaternions. The cost function $d(p, q)$ reflecting the similarity between S^{*ip} and S^{*jq} is defined by

$$d(p,q) = \Delta S^{*ip,jq} = \frac{\Delta\kappa^{*ip,jq} \cdot \Delta\tau^{*ip,jq}}{\sqrt{(S^{*ip})^2} \cdot \sqrt{(S^{*jq})^2}} \tag{2}$$

where

$$\Delta\kappa^{*ip,jq} = \|(\kappa^{*ip}, \kappa_s^{*ip}) - (\kappa^{*jq}, \kappa_s^{*jq})\| \tag{3}$$

$$\Delta\tau^{*ip,jq} = \|(\tau^{*ip}, \tau_s^{*ip}) - (\tau^{*jq}, \tau_s^{*jq})\| \tag{4}$$

$$(S^{*ip})^2 = (\kappa^{*ip})^2 + (\kappa_s^{*ip})^2 + (\tau^{*ip})^2 + (\tau_s^{*ip})^2 \tag{5}$$

$$(S^{*jq})^2 = (\kappa^{*jq})^2 + (\kappa_s^{*jq})^2 + (\tau^{*jq})^2 + (\tau_s^{*jq})^2 \tag{6}$$

The accumulative minimum cost of aligning up to S^{*ip} and S^{*jq} is represented by $u(p, q)$, which is determined by

$$u(p,q) = min\{u(p-1,q-1), u(p,q-1), u(p-1,q)\} + d(p,q) \tag{7}$$

Following the above and working from $u(1, 1)$ to $u(P, Q)$, the best alignment of the two signatures is found giving rise to the minimum overall distance, say, the DTW distance, at $u(P, Q)$, which can serve as the similarity measurement of S^{*i} and S^{*j}.

3) Perception via the Visualization of Inter-trajectory Alignment: Applying the above formulated DTW algorithm to signature matching, Fig. 5 illustrates the nonlinear paths warping (in dotted lines) among the six trajectory instances. It is observed that the inter-trajectory alignments are really reasonable. Particularly, the one-to-many point correspondence makes the trajectories with different point distributions being matched well. That is, via visualizing the inter-trajectory alignments, motion can be perceived by looking into specific motion feature, which can help human (robot) observe how two trajectories differ in the features of interest.

Based on the DTW inter-signature matching, we can not only get a quantity as the overall similarity between two

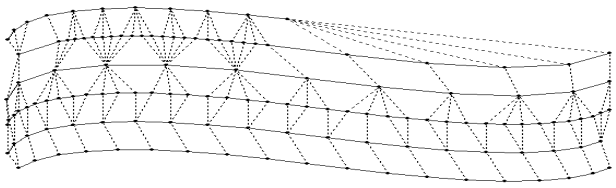

Fig. 5. Trajectory sampling instances (in solid lines) and the DTW based nonlinear paths warping (in dotted lines)

signatures, but also can visualize and perceive two trajectories intuitively referring to the inter-signature matching. In practice, the DTW based method provides a suitable metric for robot PbD learning to measure the quality of reproducing a motion by robot. This can be done by measuring the similarity between the reproduced trajectory's signature and the reference signature. Firstly, to examine the DTW distance that is a quantitative measurement to the error of a reproduction. The reproduction would be acceptable if this distance is within a predefined threshold. Secondly, human (robot) can in detail observe and perceive certain motion feature of the reproduction, for example, to check if the speed profile is consistent with that of the reference. In case certain portion is not reproduced as expected, it can be highlighted to remind robot to improve when that specific motion is repeated.

IV. THE COMPACT SECOND-LEVEL SIGNATURE

A. PCA Based Signature Dimension Reduction

Using the first-level signature, the DTW distance can be used for trajectory recognition. However, as observed, since the signature's length is equal to the length of motion trajectory, the full signature data will be large for complex and long trajectory. Also, DTW is a time-consuming matching algorithm. These two reasons may result in lower recognition efficiency. In fact, the first-level signature data is interrelated so that it can be optimized for the purpose of more compact description. This explains our motivation to build condensed second-level signature. To do that, the linear PCA (Principle Component Analysis) transform is used to reduce the dimensions of the first-level signature.

PCA transform is an effective method for dimension reduction. It projects the interrelated original data into another feature space where the projected data are small in dimension, uncorrelated, and capable of preserving most variance of the original data. PCA transform can be described by the corresponding PCA coefficients. A key problem in PCA is the selection of the number of principle components, which can be determined by an optimal threshold.

Assume the Singular Value Decomposition (SVD) is expressed by $X = U\Lambda V^T$, the PCA transform can be represented by

$$F = U^T X \tag{8}$$

where X is the original data and F the projected data. Note that only the first m columns of U (principle components) are picked to represent the PCA coefficients. If X is p-dimensional, F will be m-dimensional, and $m \ll p$. Therefore, the dimension is much reduced while the data variance is most preserved through the PCA transform.

The number of m depends on the selection of cumulative accuracy φ to account for the data variance, which can be analyzed based on the sorted eigenvalues λ_i of the covariance matrix of X, as formulated below,

$$\varphi = (\sum_{i=1}^{m} \lambda_i / \sum_{i=1}^{p} \lambda_i) \times 100\% \tag{9}$$

Hence m can be determined by a cut-off of φ like $\varphi = 95\%$ selected in our work. Then the first m columns of principle components are extracted as U_m, which is sufficient to guarantee the expected data variance (in terms of φ) being preserved in the PCA transform. In general, via the PCA transform (U_m), higher dimensional X_p is converted into lower dimensional F_m. All the first-level signatures are pre-normalized to equal length to be projected into the PCA space to obtain compact second-level signatures, before which each first-level signature X_p is re-arranged as univariate data in the form of $X_p = [\{\kappa(t)\}_{t=1}^N \{\tau(t)\}_{t=1}^N \{\kappa_s(t)\}_{t=1}^N \{\tau_s(t)\}_{t=1}^N]$.

B. Probabilistic Recognition from the Second-level Signature

The probabilistic method is used to do second-level signature recognition. According to the labeled signature classes, Gaussian Mixture model (GMM) is learned at first from the compact PCA coefficients of the training samples. Then an input signature is recognized based on Bayesian Theorem.

1) GMM Based Signature Modeling: Assume that $X_i = [X_{i,1} \cdots X_{i,2} \cdots X_{i,m} \cdots X_{i,M}]$ containing M signature samples, the underlying probability density function of X_i can be estimated by a mixture of Gaussian model Θ_i as follows,

$$P(X_i|\Theta_i) = \sum_{k=1}^{K} w_k N(X_i; \mu_k, \Sigma_k) \tag{10}$$

where K is the number of mixing Gaussian components, w_k is the mixing weights meeting $\sum_{k=1}^{K} w_k = 1$, and $N(X_i; \mu_k, \Sigma_k)$ demotes the Gaussian function,

$$f_{\mu,\Sigma}(X_i) = \frac{1}{\sqrt{2\pi}^d \sqrt{det\Sigma}} exp(-\frac{1}{2}(X_i - \mu)^T \Sigma^{-1}(X_i - \mu)) \tag{11}$$

where μ is the mean and Σ is the covariance. The dimension d is 1 as the second-level signature is univariate data.

The corresponding GMM model parameter $\Theta_i = \{w_k, \mu_k, \Sigma_k\}_{k=1}^K$ is estimated using the EM (Expectation-Maximization) algorithm [17] following the Maximum Likelihood Estimation (MLE) principle. Firstly, initialize the mixing number K and use the k-means method to estimate the initial GMM parameter $\Theta_i^{(0)} = \{w_k^{(0)}, \mu_k^{(0)}, \Sigma_k^{(0)}\}_{k=1}^K$. Then the EM algorithm iterates the E (Expectation) step and the M (Maximization) step defined as follows:

E-step:

$$P(k|X_{i,m}, \Theta_i^{(0)}) = \frac{w_k^{(0)} N(X_{i,m}; \mu_k^{(0)}, \Sigma_k^{(0)})}{\sum_{k=1}^{K} w_k^{(0)} N(X_{i,m}; \mu_k^{(0)}, \Sigma_k^{(0)})} \tag{12}$$

M-step:

$$\mu_k^{(1)} = \frac{\sum_{m=1}^{M} P(k|X_{i,m}, \Theta_i^{(0)}) X_{i,m}}{\sum_{m=1}^{M} P(k|X_{i,m}, \Theta_i^{(0)})} \quad (13)$$

$$\Sigma_k^{(1)} = \frac{\sum_{m=1}^{M} P(k|X_{i,m}, \Theta_i^{(0)})(X_{i,m} - \mu_k^{(1)})(X_{i,m} - \mu_k^{(1)})^T}{\sum_{m=1}^{M} P(k|X_{i,m}, \Theta_i^{(0)})} \quad (14)$$

$$w_k^{(1)} = \frac{1}{M} \sum_{m=1}^{M} P(k|X_{i,m}, \Theta_i^{(0)}) \quad (15)$$

The superscript (0) and (1) in above formula indicate the iteration indexes. The E step and M step are iterated until the predefined convergence condition arrived.

2) Bayesian Signature Recognition: Based on the GMM models $\{\Theta_i\}_{i=1}^{C}$ for C motion classes, the query signature X_q is recognized using the Bayes' decision rule. Considering the efficiency of likelihood calculation, the logarithm form of the Bayes' Theorem is adopted,

$$logP(\Theta_i|X_q) = logP(X_q|\Theta_i) + logP(\Theta_i) - logP(X_q) \quad (16)$$

Recognition is based on the posterior probability $logP(\Theta_i|X_q)$ following the MAP criterion. To do that, $P(X_q|\Theta_i)$ is calculated based on (10). The prior probability $P(\Theta_i)$ can be derived from the initial knowledge about the occurrence frequency of the samples. For example, it can be set to $1/C$ in case all the classes are equi-probable. The marginal probability $P(X_q)$ is calculated by $P(X_q) = \sum_{i=1}^{C} P(X_q|\Theta_i)P(\Theta_i)$.

V. EXPERIMENTS

A. Free Form Motion Trajectory Recognition and Perception

The first experiment is carried out to demonstrate free form trajectory recognition and perception from the first-level signature by applying the DTW signature matching. Three-dimensional trajectory is tracked and acquired using the method in [14]. Fig. 6 shows the stereo sensor setup and several snaps of the stereo trajectory tracking from human motion.

A small exemplary database is built containing ten (labeled by from '01' to '10') motion patterns, which are characterized in terms of the regular trajectories whose shapes are roughly close to the characters from '0' to '9', respectively. Note that we are not doing handwritten character recognition but illustrating the free forms of the trajectories. Each pattern is represented by a canonical instance, and all the canonical instances of the ten patterns are shown in Fig. 7. Next, as shown in Fig. 8, three series of input trajectories demonstrated by three different users with increasing shape variance (noise strength) are prepared to do recognition. They are smoothed by the moving average filter using span parameter 3, 5 and 9 respectively.

Fig. 7. The canonical instances of the ten motion patterns

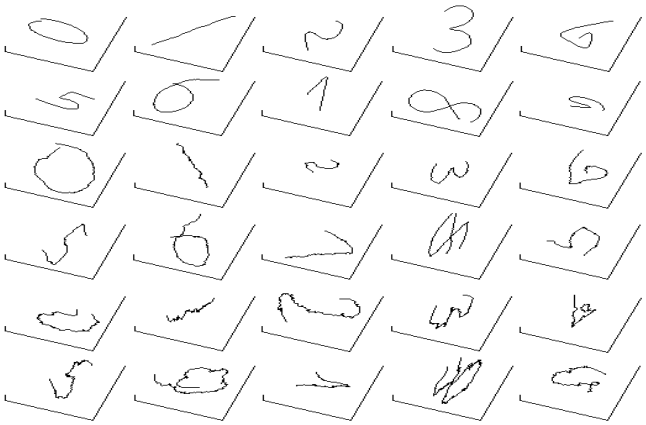

Fig. 8. The input series A (row 1-2), B (row 3-4) and C (row 5-6) with increasing variance in shape

A 1-NN classifier is developed by measuring the DTW distances of the first-level signatures between canonical instances (references) and query inputs. The classification results of the three series of inputs are illustrated in three corresponding tables in Fig. 9, where the DTW distances between the inputs (Y axis) and the references (X axis) are scaled to gray images. The lighter the cell is, the higher similarity is indicated. According to the distinctive lighter cells in the table diagonals, we get an intuitive confirmation that all the inputs are recognized correctly, in which the increasing variance in shape illustrates the signature's robustness against noise.

In addition, to illustrate the motion perception from the DTW inter-signature matching, Fig. 10 visualizes four cases of the inter-trajectory alignments. The warping paths reflect the difference on motion shape, motion speed and occlusion between two trajectories. This is helpful to analyze detailed

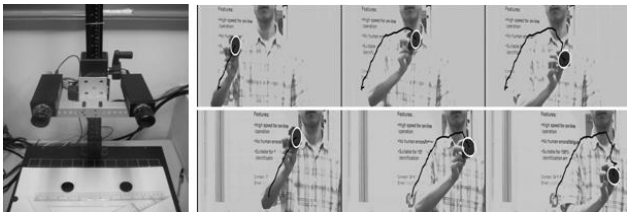

Fig. 6. The stereo setup and snaps of the stereo trajectory tracking

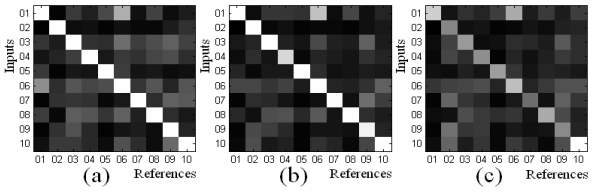

Fig. 9. The DTW based inter-signature distances comparison of the input series A (a), B (b) and C (c) with the canonical instances

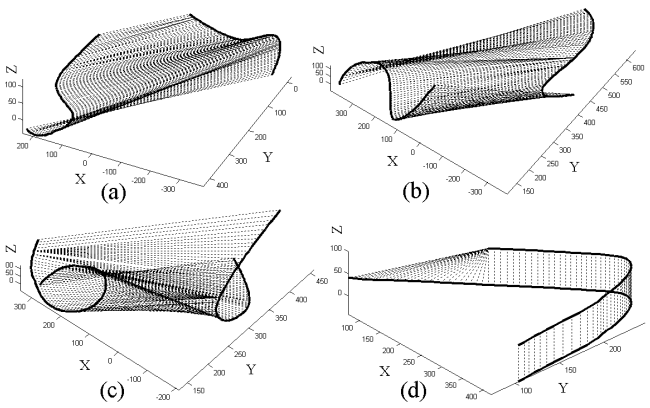

Fig. 10. Motion perception by visualizing the difference in trajectory shape (a)(b), motion speed (c) and occlusion (d)

motion features, indicate the aspects robot needs to improve in learning, and even identify the characterized users.

B. Fast Human Sign Language Recognition

The second experiment uses a larger trajectory dataset of UCI KDD high quality ASL [18] to test the signature's retrieval performance especially the improvement in efficiency benefitted from the second-level signature. The ASL trajectory dataset consists of 95 sign classes, and 27 samples were captured for each sign. All are re-sampled to 50 points specifically for the second-level signature calculation. To reduce noise and vibration, the wavelet smoother is applied using wavelet DB5 and the third level coefficients.

Half samples of a class are used to train a GMM model and the other half are input to do Bayesian signature recognition. We also compare the Fourier Descriptor based recognition by extracting the first 4 coefficients as trajectory representation and using the average Euclidean distance for classification. The retrieval experiment is repeated more than 50 times in a common PC (Pentium 4 CPU 3.00GHz, 512M RAM) by picking up a number of classes and samples randomly, which gives rise to an average recognition results. From Table I, we can find that the signature outperforms the FD representation in recognition accuracy. In addition, according to the recognition efficiency in Table II, the second-level signature is much faster than the first-level signature in trajectory retrieval, and as the classes number increases, the efficiency decreases much slower than that of the first-level signature. That is because the second-level signature is with optimized data and the Bayesian engine is used.

VI. CONCLUSIONS

Pursuing flexible human (robot) motion characterization, motion trajectory is particularly studied in this paper by building a hierarchical signature descriptor. The signature's novelty lies in its hierarchy and the resulted functional adaptability. While the first-level signature is good at full motion description and perception, the second-level signature is more efficient in recognition using the probabilistic methods. It can be concluded that the signature is effective.

TABLE I

ACCURACY COMPARISON OF THE SIGN TRAJECTORY RETRIEVAL

Classes No.	2	4	8
The second-level signature	92.38%	86.17%	78.66%
Fourier Descriptor	87.98%	75.74%	63.85%

TABLE II

RECOGNITION EFFICIENCY COMPARISON (UNIT: MS/QUERY)

Classes No.	2	4	8
The first-level signature (DTW matching)	753	1287	2206
The second-level signature	139	146	158

VII. ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their valuable comments.

REFERENCES

[1] M. H. Yang and N. Ahuja, "Recognizing hand gestures using motion trajectories," *in Proc. of IEEE Conference on CVPR*, Fort Collins, Colorado, Jun. 1999, pp. 468-472.

[2] Y. Sato, K. Bernardin, H. Kimura, and K. Ikeuchi, "Task analysis based on observing hands and objects by vision," *in Proc. IEEE/RSJ Int. Conf. on IROS*, Lausanne, Switzerland, 2002, pp 1208-1213.

[3] K. R. Dixon and P. K. Khosla, "Learning by observation with mobile robots: a computational approach," *in Proc. of IEEE Int. Conf. Robot. Automat.*, New Orleans, USA, Apr. 2004, vol. 1, pp. 102-107.

[4] C. Rao, A. Yilmaz, and M. Shah, View-invariant representation and recognition of actions, *International Journal of Computer Vision*, 50 (2), pp. 203-226, Nov. 2002.

[5] V. V. Kindratenko, On using functions to describe the shape, *Journal of Mathematical Imaging and Vision*, vol. 18, no. 33, pp. 225-245, May 2003.

[6] D. Zhang and G. Lu, Review of shape representation and description techniques, *Pattern Recognition*, vol. 37, no. 1, pp. 1-19, Jan. 2004.

[7] P. R. G. Harding and T. J. Ellis, "Recognizing hand gesture using Fourier descriptors," *in Proc. of the 17th International Conference on Pattern Recognition*, vol. 3, Cambridge, UK, 2004, pp. 286-289.

[8] G. Chuang and C. C. Kuo, Wavelet descriptor of planar curves: theory and applications, *IEEE Trans. on Image Processing*, vol. 5, no. 1, pp. 56-70, Jan. 1996.

[9] F. Mokhtarian, S. Abbasi and J. Kittler, "Robust and efficient shape indexing through curvature scale space," *in Proc. British Machine Vision Conference*, Edinburgh, UK, 1996, pp. 53-62.

[10] M. K. Hu, Visual pattern recognition by moment invariants, *IEEE Trans. Information Theory*, vol. 8, no. 2, pp. 179-187, Feb. 1962.

[11] J. Aleotti and S. Caselli, "Trajectory clustering and stochastic approximation for robot programming by demonstration," *Proc. IEEE/RSJ Int. Conf. on IROS*, Alberta, Canada, August 2-6, 2005, pp. 1029-1034.

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

[13] E. Calabi, P. J. Olver, C. Shakiban, A. Tannenbaum and S. Haker, Differential and numerically invariant signature curves applied to object recognition, *IJCV*, vol. 26, no. 2, 1998, pp. 107-135.

[14] S. D. Wu, Y. F. Li and J. W. Zhang, "Signature descriptor for free form trajectory modeling," *in Proc. IEEE International Conference on Integration Technology*, Shenzhen, China, 2007, pp. 167-172.

[15] P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, pp. 629-639, Jul. 1990.

[16] M. Vlachos, G. Kollios, and D. Gunopulos, Elastic translation invariant matching of trajectories, *Machine Learning*, vol. 58, no. 2-3, pp. 301-334, Feb. 2005.

[17] J. A. Bilmes, A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian Mixture and Hidden Markov Models, *International Computer Science Institute*, TR-97-021.

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