Fast Adaptive Reparametrization (FAR) with Application to Human Action Recognition

Enjie Ghorbel, Girum Demisse, Djamila Aouada, Björn Ottersten

Abstract—In this paper, a fast approach for curve reparametrization, called Fast Adaptive Reparamterization (FAR), is introduced. Instead of computing an optimal matching between two curves such as Dynamic Time Warping (DTW) and elastic distance-based approaches, our method is applied to each curve independently, leading to linear computational complexity. It is based on a simple replacement of the curve parameter by a variable invariant under specific variations of reparametrization. The choice of this variable is heuristically made according to the application of interest. In addition to being fast, the proposed reparametrization can be applied not only to curves observed in Euclidean spaces but also to feature curves living in Riemannian spaces. To validate our approach, we apply it to the scenario of human action recognition using curves living in the Riemannian product Special Euclidean space $\mathbb{SE}(3)^n$. The obtained results on three benchmarks for human action recognition (MSRAction3D, Florence3D, and UTKinect) show that our approach competes with state-of-the-art methods in terms of accuracy and computational cost.

Index Terms—Reparametrization, action recognition Riemannian manifolds

I. INTRODUCTION

CURVES have been shown to be a very powerful representation in the fields of computer vision and pattern recognition. In fact, due to their simplicity, they are used for the description of various observations such as objects or shapes, through their contours in 2D [1][2][3], or level sets in 3D [4][5][6]. Curves are also used to represent temporal information such as speech [7][8], and motion field [9][10][11][12][13][14][15].

A curve may be defined as a mapping $\alpha : I \to \mathcal{M}$ where $I \subset \mathbb{R}$ is an interval of reals regulating the order of the data to be mapped into $\mathcal{M}$, the manifold of observations. The interval I governs what is called the *parametrization* of the curve. However, curve parametrization is not unique and the same curve in $\mathcal{M}$ can be expressed in an infinite number of different parameter spaces. Hence, this variability of parametrization can lead to the mismatch of similar curves. Thus, to compare and manipulate curves, a *reparametrization* step is usually undertaken where the domain I of a curve is to be redefined [16][17][3]. First curve matching techniques have been based on the detection of a set of finite points called landmarks [18][19] which are used to statistically describe curve
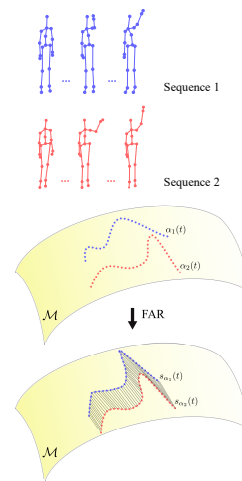
Fig. 1. Application of the proposed FAR method in the context of human action recognition: (Top) features are first extracted from skeleton sequences. (Middle) they are expressed as time dependant curves lying in a given manifold $\mathcal{M}$ denoted as α_1 and α_2. (Bottom) to overcome execution rate variability, the curves are reparametrized using the FAR approach by respectively computing s_{α_1} and s_{α_2}.

shapes. In later works, more sophisticated reparametrization techniques using the global information of curves [20][17] have been introduced. These curves are usually considered as varying in a Euclidean space. Nevertheless, some discriminative curves lie in nonlinear spaces. More particularly, curves living in Riemannian manifolds have shown great potential such as the space of Symmetric Positive Definite matrices [21] and the Special Euclidean group [22][3]. Thus, the non-linearity of these informative spaces makes classical reparametrization methods not directly applicable.

To overcome this issue, several attempts have been made towards the generalization of reparametrization to Riemannian manifolds. For instance, Dynamic Time Warping (DTW) [23], which has been initially designed for curves in $\mathbb{R}^n$, has been adapted to different Riemannian manifolds [24] by replacing the Euclidean distance used to compute the similarity measure by a more appropriate entity (geodesic distance, distance between curves projected to the Lie algebra). On the other hand, instead of using the similarity measure resulting from DTW, other approaches have proposed to define elastic distances [9][2][16]. This rate-invariant distance is usually computed by defining an optimization problem that finds the optimal parametrization. In [16], the Transported Square-Root Vector Field (TSRVF) which represents the extension of the Square-Root Vector Field (SRVF) method [3] to Riemannian manifolds has been introduced. Anirudh et al. [9] applied the TSRVF framework on two different Riemannian
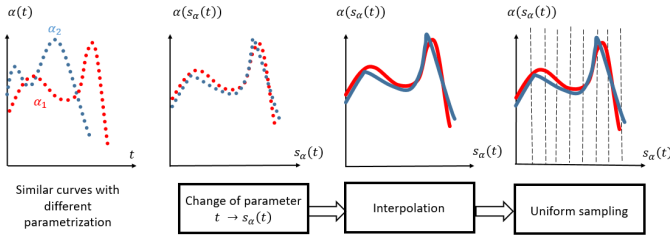
Fig. 2. Overview of the proposed reparametrization technique: First, a change of variable is realized by replacing the parameter t by the parametrization function $s_\alpha(t)$ described in (4). After that, an interpolation followed by a uniform sampling is performed to make the two curves comparable for the same level $s_\alpha(t)$.

manifolds ($\mathbb{SE}(3)$ and the Special Orthogonal group $\mathbb{SO}(3)$). Also, Demisse et al. [2] described curves as sequences of transformation matrices of $\mathbb{SE}(3)$ and formulated the problem of curve matching as an optimization problem using the geodesic distance of $\mathbb{SE}(3)$.

However, in spite of their effectiveness, these generalized methods designed for curves in Riemannian spaces present some drawbacks: (1) they are highly dependant on the quality of data since curves are compared in pairs; (2) they rely on optimization leading to an important computational complexity, and (3) they are performed using discrete data that are matched without necessarily corresponding to each other.

To address these limitations, a novel and fast curve reparametrization framework applicable to Euclidean and Riemannian spaces called *Fast Adaptive Reparametrization* (FAR) is introduced. Instead of comparing curves in pairs, this method reparametrizes any curve $\alpha : I \to \mathcal{M}$ by heuristically defining an adequate homeomorphism $s_\alpha : I_\alpha \to I$ such that $\alpha : I_\alpha \to \mathcal{M}$. To avoid relying on discrete data, the reparametrized numerical curves are interpolated and resampled uniformly. To evaluate our approach, FAR is adapted to the context of human action recognition in $\mathbb{SE}(3)^n$, as illustrated in Figure 1, and tested on three different benchmarks.

II. PROPOSED FAST ADAPTIVE REPARAMETRIZATION

The main goal of the proposed approach is to extract the properties of curves that are invariant under reparametrization in order to recognize equivalent curves. The question to be raised is, therefore, the following: *what is the meaning of equivalent curves?*

Let $\alpha_1 : I_1 \to \mathcal{M}$ and $\alpha_2 : I_2 \to \mathcal{M}$ be two curves with $\mathcal{M}$ as the observation manifold and $I_1 = [a_1, b_1], I_2 = [a_2, b_2] \in \mathbb{R}$ as their respective parameter domains. Follows the classical mathematical definition of equivalent curves.

Definition 1: The two curves α_1 and α_2 are considered to be equivalent ($\alpha_1 \sim \alpha_2$) if and only if there exists a one-to-one function $s(t) : I_2 \to I_1$ such that $\forall t, \dot{s}(t) \neq 0$ and $\alpha_2(s(t)) = \alpha_1(t)$.

The primary challenge is therefore to find the appropriate parametrization function $s(t)$.

Usually, reparametrization techniques as in [9][2][3] define $s(t)$ by minimizing a well-chosen similarity measure d between a pair of curves α_1 and α_2 such that

$$s^* = \min_s d\left(\alpha_1, \alpha_2 \circ s\right), \tag{1}$$

where s^* denotes the optimal parametrization of α_2 with respect to α_1. Unfortunately, this leads to polynomial complexity. Note that the similarity measure d_α between two curves is sometimes implicitly defined using an intermediate mapping such as TSRVF [16], SRVF [3], etc.

However, in pattern recognition, the definition of curve equivalence is slightly different from the original mathematical definition (*Definition* 1): curves are considered equivalent if they describe a similar pattern. Thus, the notion of equivalence can be confused with the notion of class. In fact, the goal of reparametrizing curves is to ensure that curves describing the same class look more similar. Therefore, it can be stated that the meaning of equivalence widely depends on the final application and on the properties that we aim to extract from curves. Thus, we propose to make use of our prior knowledge to define $s(t)$ in order to reduce the complexity of reparametrization algorithms. Instead of finding the optimal parametrization by comparing a given curve $\alpha : I_\alpha = [a_\alpha, b_\alpha] \to \mathcal{M}$ to a template, we define a homeomorphism as a parametrization function $s_\alpha : I_\alpha \to [0, 1]$ independently of any other reference curve. The function $s_\alpha(t)$ is therefore computed and applied to $\alpha(t)$ as described below.

$$\alpha(s_\alpha(t)) = \alpha \circ s_\alpha(t). \tag{2}$$

This heuristic approach is inspired by the TVR algorithm proposed in [10]. However, while TVR is restricted to the specific case of temporal normalization for human action recognition using curves lying in Euclidean spaces, our approach presents a generalized formulation applicable to various applications for curves lying not only in Euclidean spaces but also in Riemannian manifolds. Therefore, the main idea of our approach is to define s_α using a score function $f_\alpha : [a_\alpha, b_\alpha] \to \mathbb{R}$ which include a heuristic invariance under reparametrization of relevant curve properties. Instead of relying on the pure mathematical definition of curve equivalence, a physical component is included to orient the choice of the function f_α. In other words, our main goal is not to make the curves themselves invariant to reparametrization, but to restrict this invariance to the discriminative curve features in link with the application. The function f_α is computed based on the curve α itself. Therefore, we note that when designing f_α, it is important to take into account the topological structure of the space $\mathcal{M}$. Indeed, if curves are living in a non-linear space $\mathcal{M}$, classical measures such as Euclidean distance, derivative, etc. are not suitable.

Since open curves are numerically assimilated to ordered data, $s_\alpha(t)$ is constrained to be strictly increasing with respect to the variable t. To that aim, an increasing energy function E_α is designed using f_α as follows:

$$E_\alpha(t) = \int_a^t \|f_\alpha(t)\|_2^2 dt. \tag{3}$$

Furthermore, in order to ensure the variation in a single interval $I = [0, 1]$, the parametrization function $s_\alpha(t)$ is computed as follows:

$$s_\alpha(t) = \frac{E_\alpha(t)}{E_\alpha(b)}. \tag{4}$$

Note that arclength reparametrization can be considered as a specific case of this reparametrization, for which $f_\alpha(t) = \dot{\alpha}(t)$.

Numerical consideration: One of the strengths of FAR is that it takes into consideration the discrete nature of the data. Instead of trying to match non-corresponding curve points, we propose to follow a different strategy: we realize an interpolation of curves followed by uniform sampling. Figure 2 illustrates the different components of the proposed framework.

III. APPLICATION TO HUMAN ACTION RECOGNITION IN $\mathbb{SE}(3)$

In this section, we present the application of FAR in the context of action recognition using features expressed in $\mathbb{SE}(3)$. First, a brief mathematical background that overviews the properties of the Riemannian manifold $\mathbb{SE}(3)$ is recalled. Then, the used human motion representation, as well as the adaptation of FAR in this particular case, are described.

A. Special Euclidean Group $\mathbb{SE}(3)$

The Special Euclidean group $\mathbb{SE}(3)$ is defined as the set containing the 4×4 square matrices with the following form:

$$T(\mathbf{R}, \overrightarrow{\mathbf{t}}) = \begin{pmatrix} \mathbf{R} & \overrightarrow{\mathbf{t}} \\ 0 & 1 \end{pmatrix}, \tag{5}$$

where $\mathbf{R}$ is a three-dimensional rotation matrix and $\overrightarrow{\mathbf{t}}$ is a three-dimensional translation vector. As $\mathbb{SE}(3)$ is a Lie group and consequently is a non-linear structure, classical operations are not trivial. The more common practice is to map $\mathbb{SE}(3)$ elements to its Lie algebra at the identity element $\mathbf{I}_4$. The Lie algebra, denoted by $\mathbf{se}(3)$, represents the tangent space and is characterized by a vector space structure. To map elements of $\mathbb{SE}(3)$ to $\mathbf{se}(3)$, and vice-versa, the exponential and the logarithm maps respectively denoted by $\exp$ and $\log$ are used. For more details about differential geometry and the Lie group $\mathbb{SE}(3)$, we respectively refer the reader to [25] and [26].

Distance in $\mathbb{SE}(3)$: The geodesic distance d_G between two matrices $\mathbf{T}_A(\mathbf{R}_A, \overrightarrow{\mathbf{t}_A})$ and $\mathbf{T}_B(\mathbf{R}_B, \overrightarrow{\mathbf{t}_B}) \in \mathbb{SE}(3)$ can be computed as follows [2]:

$$d_G(\mathbf{T}_A, \mathbf{T}_B) = (\| \log(\mathbf{R}_A^T \mathbf{R}_B) \|_F^2 + \| \overrightarrow{\mathbf{t}_A} - \overrightarrow{\mathbf{t}_B} \|_2^2)^{1/2}. \tag{6}$$

Distance in $\mathbb{SE}(3)^n$: The cross product space $\mathbb{SE}(3)^n$ inherits the Lie group structure of $\mathbb{SE}(3)$. Therefore, the geodesic distance D between two elements $\alpha_1 = (\mathbf{T}_1, \mathbf{T}_2, ..., \mathbf{T}_n)$ and $\alpha_2 = (\mathbf{Q}_1, \mathbf{Q}_2, ..., \mathbf{Q}_n) \in \mathbb{SE}(3)^n$, with $\mathbf{T}_i$ and $\mathbf{Q}_i \in \mathbb{SE}(3)$ $\forall i$, is computed as follows:

$$D(\alpha_1, \alpha_2) = (\sum_{i=1}^{n} d_G(\mathbf{T}_i, \mathbf{Q}_i)^2)^{1/2}. \tag{7}$$

B. FAR Reparametrization in $\mathbb{SE}(3)^n$

To model human motion using skeletal data, we propose to use the representation introduced in [22] called Lie Algebra Relative Pairs (LARP) representing trajectories lying in $\mathbb{SE}(3)^n$.

In [22], authors suggest that a human skeleton S, composed of m edges, at an instant t can be represented by an ordered set of transformation matrices $\alpha(t) =$

$\{\mathbf{T}_{1,2}(t), \mathbf{T}_{2,1}(t), ..., \mathbf{T}_{m,m-1}(t), \mathbf{T}_{m-1,m}(t)\} \in \mathbb{SE}(3)^n$. These transformation matrices are estimated between using the skeleton edges. Thus, the full action can be seen as a trajectory $(\alpha(t))_{t \in [t_0, t_0+L]}$ varying in the product Special Euclidean Group $\mathbb{SE}(3)^n$ with L the duration of the action, t_0 the starting time and $n = 2C_m^2$, where C_m^2 is the number of combinations. More details about this representation can be found in the original paper [22].

To overcome execution rate variability, we propose to apply to the LARP representation, the FAR approach on the obtained curve α, as described in (4). To compute the parametrization function $s_\alpha(t)$, two different score functions $f_\alpha(t)$ are proposed, namely, arclength and motion quantity.

Arclength: To compute the arclength of the curve $\alpha(t)$ at an instant t on the manifold $\mathbb{SE}(3)^n$, we propose to compute the geodesic distance D between $\alpha(t)$ and $\alpha(t-1)$,

$$f_\alpha(t) = D(\alpha(t), \alpha(t-1)). \tag{8}$$

Therefore, the faster the motion is, the higher $f(t)$ is. Integrating this function in the parametrization function $s(t)$ makes α invariant to velocity variation. This can be seen as the extension of arclength reparametrization for the space $\mathbb{SE}(3)^n$.

Motion quantity: We define motion quantity as the distance between the current pose $\alpha(t)$ the rest pose denoted as α_{ref}. We propose to compute it using the geodesic distance D as follows:

$$f_\alpha(t) = D(\alpha(t), \alpha_{ref}). \tag{9}$$

The closer to the rest pose the current pose is, the less important it is considered. In the experiments, we assume that the rest pose is encoded in the first frame of the video since the used datasets are composed of segmented actions. Nevertheless, this approach can be further extended by learning a generic rest pose model.

Interpolation and uniform sampling: We propose to use the interpolation scheme for $\mathbb{SE}(3)$ as introduced in [27] and used in [22]. Each reparametrized temporal sequence $\mathbb{SE}(3)$ is interpolated individually. For each sequence α_j defined by $\mathbf{T}_j(t_0), \mathbf{T}_j(t_1), ..., \mathbf{T}_j(t_i), ..., \mathbf{T}_j(t_0 + L) \in \mathbb{SE}(3)$, the following interpolation formula is used. Considering $s(t) \in [s(t_i), s(t_{i+1})]$, we have:

$$\hat{\alpha}_j(s_\alpha(t)) = \mathbf{T}_j(s_\alpha(t_i)) \exp \left(\frac{s_\alpha(t) - s(t_i)}{s_\alpha(t_{i+1}) - s_\alpha(t_i)} \mathbf{B}_i \right), \tag{10}$$

with $\mathbf{B}_i = \log \left(\mathbf{T}_j(s_\alpha(t_i))^{-1} \mathbf{T}_j(s_\alpha(t_{i+1})) \right)$.
Finally, the interpolated curve $\hat{\alpha} = (\hat{\alpha}_j)_{j \in [1,n]}$ is uniformly sampled to recover the final descriptor.

IV. EXPERIMENTS

Our reparametrization approach is applied to LARP, as presented in Section III, and tested on three benchmarks collected for action recognition: Florence3D dataset [28], UTKinect dataset [29] and MSRAction3D dataset [30].

Florence3D dataset [28] is composed of 9 different actions. Each action is performed by 10 different subjects from two to three times. This dataset acquired using Kinect provides RGB images, depth maps as well as skeleton sequences. Each skeleton is constituted of 15 joints. UTKinect dataset proposed

in [29] contains 10 different classes of action. Each action class is repeated by 10 different subjects twice. Captured using Kinect, the three RGB-D modalities are also provided for this benchmark. However, in contrast to Florence3D, skeletons are formed by 20 joints. MSRAction3D introduced in [30] is probably one of the most well-known datasets. This dataset is composed of 20 actions performed by 10 different subjects from two to three times. It is also captured using Kinect. However, only skeletons and depth maps are provided. As in the UTKinect dataset, each skeleton is composed of 20 joints.

Our reparametrization applied to LARP is denoted as LARP+FAR-A (when using the arclength as a score function f_α) and LARP+FAR-MQ (when using the motion quantity as f_α). As discussed in [9], the Fourier Temporal Pyramid (FTP), which is a tool used to remove noise from skeletons, is not applied to non-Euclidean features. Indeed, analyzing and exploring the warping is not a trivial task when using FTP. In the experiments, we follow the same protocol proposed in [9], where five train-test protocols are realized and then averaged. For each protocol, half of the subjects are used for the training, while the rest is used for testing. For the classification, we also use the same one-vs-all SVM classifier proposed in [22].

To evaluate our approach, we combine the LARP representation with different reparametrization techniques. We compare our approach to LARP without carrying out any warping, using the modified DTW introduced in [22], the TSRVF method proposed in [9], the TSRVF with manifold functional variant of Principal Component Analysis (mfPCA) [9] and Principal Geodesic Analysis introduced in [31]. These approaches are respectively denoted as LARP, LARP+DTW, LARP+TSRVF, LARP+nfPCA, and LARP+PGA.

The results obtained on the Florence3D, UTKinect and MSRAction3D datasets, are reported in Table I.

Arclength (A) vs. Motion Quantity (MQ): The results show that LARP+FAR present better results on Florence3D and UTKinect datasets when associated to the MQ motion signal (with respectively 92.96% and 86.16% using LARP+FAR-MQ against 90.88% and 95.35% using LARP+FAR-A). This could be explained by the fact that MQ function highlights the frames containing key poses and is less sensitive to noise resulting from undesired motion. However, on MSRAction dataset, the obtained result using A as a score function is slightly higher than the one registered for LARP+FAR-MQ (only a difference of 0.15%).

Comparison with state-of-the-art: Compared to LARP+DTW and LARP+PGA, our method improves the performance on the three datasets by around 4% and 13%, respectively, on Florence3D, and 4% and 5% on UTKinect, and 4% and 9% on MSRAction3D. Also, compared to LARP+TSRVF and LARP+mfPCA , our approach enhances the results by respectively 3% and 3% on Florence3D and by 2% for both on UTKinect. Nevertheless, LARP+TSRVF and LARP+mfPCA present better performance than LARP+FAR on MSRAction3D by slightly exceeding it by respectively 1% and 2% . Despite that, it can be noted that globally the FAR approach achieves competitive results for action recognition.

Computational cost: In comparison to state-of-the-art approaches, our method has the advantage to present very

Method	Florence3D	UTKinect	MSRAction3D
LARP [22]	86.27*	93.57*	75.57*
LARP+DTW [22]	86.74*	92.17*	78.75*
LARP+PGA [31]	79.01	91.26	72.06
LARP+TSRVF [9]	89.50	94.47	84.62
LARP+nfPCA [9]	89.67	94.87	**85.16**
LARP+FAR-A (ours)	90.88	95.35	83.17
LARP+FAR-MQ (ours)	**92.96**	**96.16**	**83.03**

TABLE I
RECOGNITION ACCURACY ON FLORENCE3D, UTKINECT AND MSRACTION3D DATASETS. *THE VALUES CORRESPOND TO THE ONES REPORTED IN [9] AND NOT TO THE ONES REPORTED IN [22], SINCE FTP IS REMOVED.

low complexity. While LARP+TSRVF and LARP+mfPCA have a polynomial complexity, LARP+FAR requires only a complexity of $O(N)$, with N being the number of points forming the curve. This is explained by the fact that FAR is applied independently to each trajectory without the use of any reference, in contrast to other approaches which rely mainly on optimization between pairs of curves. As an illustration, DTW (having a complexity of $O(N^2)$) and FAR are compared in terms of execution time on the UTKinect dataset. Although DTW is implemented in C, FAR remains almost two times faster (LARP + DTW is executed in 272.10s, while LARP is executed in 164.72s).

Limitations and future works: It is important to note that the proposed approach in its current shape does not completely answer the equivalence formulation. For example, eliminating the variability by redefining the parameter space makes the calculation of a mean trajectory not straightforward in comparison to elastic approaches. Furthermore, the proposed functions might be affected by noisy data. Finally, an extension to closed curves could be important for some applications. All these points remain interesting to investigate.

V. CONCLUSION

In this paper, a fast open curve reparametrization technique in Riemannian spaces, called Fast Adaptive Reparametrization (FAR), is introduced. Instead of reparametrizing curves with respect to a reference, this approach reparametrizes each curve independently, ensuring a complexity of $O(N)$. This is done through the heuristic definition of a parametrization function depending on the application of interest. The FAR approach was applied to the LARP descriptor [22] in the domain of skeleton-based action recognition. This allowed evaluating our approach on curves in the Riemannian space $\mathbb{SE}(3)^n$.. The obtained results on three well-known datasets show the effectiveness of our approach as compared to other state-of-the-art techniques for action alignment in terms of accuracy and computational complexity.

AKNWOLEDGEMENTS

REFERENCES

[1] G. G. Demisse, D. Aouada, and B. Ottersten, "Similarity metric for curved shapes in euclidean space," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 5042–5050.

[2] ——, "Deformation based curved shape representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 6, pp. 1338–1351, 2018.

[3] A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn, "Shape analysis of elastic curves in euclidean spaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 7, pp. 1415–1428, 2011.

[4] D. Aouada and H. Krim, "Squigraphs for fine and compact modeling of 3-d shapes," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 306–321, 2010.

[5] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE transactions on pattern analysis and machine intelligence*, vol. 17, no. 2, pp. 158–175, 1995.

[6] M. Rousson and N. Paragios, "Shape priors for level set representations," in *European Conference on Computer Vision*. Springer, 2002, pp. 78–92.

[7] C. C. Tappert, "Cursive script recognition by elastic matching," *IBM Journal of Research and development*, vol. 26, no. 6, pp. 765–771, 1982.

[8] Y. Tohkura, "A weighted cepstral distance measure for speech recognition," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 35, no. 10, pp. 1414–1422, 1987.

[9] R. Anirudh, P. Turaga, J. Su, and A. Srivastava, "Elastic functional coding of riemannian trajectories," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 5, pp. 922–936, 2017.

[10] E. Ghorbel, R. Boutteau, J. Boonaert, X. Savatier, and S. Lecoeuche, "Kinematic spline curves: A temporal invariant descriptor for fast action recognition," *Image and Vision Computing*, vol. 77, pp. 60–71, 2018.

[11] R. Baptista, G. Demisse, D. Aouada, and B. Ottersten, "Deformation-based abnormal motion detection using 3d skeletons," in *IEEE International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2018.

[12] E. Ghorbel, R. Boutteau, J. Bonnaert, X. Savatier, and S. Lecoeuche, "A fast and accurate motion descriptor for human action recognition applications," in *Pattern Recognition (ICPR), 2016 23rd International Conference on*. IEEE, 2016, pp. 919–924.

[13] K. Papadopoulos, G. Demisse, E. Ghorbel, M. Antunes, D. Aouada, and B. Ottersten, "Localized trajectories for 2d and 3d action recognition," *Sensors*, 2019.

[14] E. Ghorbel, K. Papadopoulos, R. Baptista, H. Pathak, G. Demisse, D. Aouada, and B. Ottersten, "A view-invariant framework for fast skeleton-based action recognition using a single rgb camera," in *14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Prague, 25-27 February 2018*, 2019.

[15] R. Baptista, E. Ghorbel, A. E. R. Shabayek, F. Moissenet, D. Aouada, A. Douchet, M. André, J. Pager, and S. Bouilland, "Home self-training: Visual feedback for assisting physical activity for stroke survivors," *Computer methods and programs in biomedicine*, vol. 176, pp. 111–120, 2019.

[16] J. Su, S. Kurtek, E. Klassen, A. Srivastava *et al.*, "Statistical analysis of trajectories on riemannian manifolds: bird migration, hurricane tracking and video surveillance," *The Annals of Applied Statistics*, vol. 8, no. 1, pp. 530–552, 2014.

[17] L. Younes, "Computable elastic distances between shapes," *SIAM Journal on Applied Mathematics*, vol. 58, no. 2, pp. 565–586, 1998.

[18] D. G. Kendall, "A survey of the statistical theory of shape," *Statistical Science*, pp. 87–99, 1989.

[19] J. Wang, X. Bai, X. You, W. Liu, and L. J. Latecki, "Shape matching and classification using height functions," *Pattern Recognition Letters*, vol. 33, no. 2, pp. 134–143, 2012.

[20] M.-A. Charmi, S. Derrode, and F. Ghorbel, "Fourier-based geometric shape prior for snakes," *Pattern Recognition Letters*, vol. 29, no. 7, pp. 897–904, 2008.

[21] E. Ghorbel, J. Boonaert, R. Boutteau, S. Lecoeuche, and X. Savatier, "An extension of kernel learning methods using a modified log-euclidean distance for fast and accurate skeleton-based human action recognition," *Computer Vision and Image Understanding*, 2018.

[22] R. Vemulapalli, F. Arrate, and R. Chellappa, "Human action recognition by representing 3D skeletons as points in a Lie group," in *Conference on Computer Vision and Pattern Recognition, IEEE*, 2014, pp. 588–595.

[23] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series." in *KDD workshop*, vol. 10, no. 16. Seattle, WA, 1994, pp. 359–370.

[24] M. F. Abdelkader, W. Abd-Almageed, A. Srivastava, and R. Chellappa, "Silhouette-based gesture and action recognition via modeling trajectories on riemannian shape manifolds," *Computer Vision and Image Understanding*, vol. 115, no. 3, pp. 439–455, 2011.

[25] R. M. Murray, *A mathematical introduction to robotic manipulation*. CRC press, 2017.

[26] B. Hall, *Lie groups, Lie algebras, and representations: an elementary introduction*. Springer, 2015, vol. 222.

[27] M. Zefran and V. Kumar, "Two methods for interpolating rigid body motions," in *Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on*, vol. 4. IEEE, 1998, pp. 2922–2927.

[28] L. Seidenari, V. Varano, S. Berretti, A. Bimbo, and P. Pala, "Recognizing actions from depth cameras as weakly aligned multi-part bag-of-poses," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2013, pp. 479–485.

[29] L. Xia, C.-C. Chen, and J. Aggarwal, "View invariant human action recognition using histograms of 3D joints," in *Conference on Computer Vision and Pattern Recognition Workshops, IEEE*, 2012, pp. 20–27.

[30] W. Li, Z. Zhang, and Z. Liu, "Action recognition based on a bag of 3D points," in *Conference on Computer Vision and Pattern Recognition Workshops, IEEE*, 2010, pp. 9–14.

[31] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi, "Principal geodesic analysis for the study of nonlinear statistics of shape," *IEEE transactions on medical imaging*, vol. 23, no. 8, pp. 995–1005, 2004.

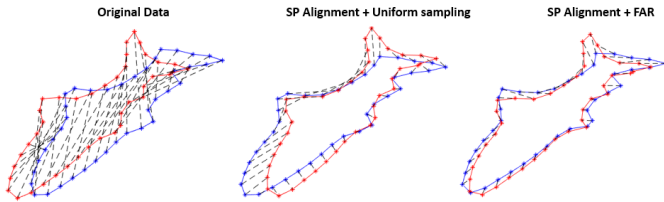# Supplementary Materials: Example of 2D Curve Representation with FAR

Fig. 1. Correspondence between two closed curves (red and blue): (Left) illustrates the correspondence between the two original curves before applying any reparametrization. It can be noted that the matching quality is very low. (Middle) shows the point matching of the two curves after applying the Starting Point alignment (SP) and the uniform reparametrization proposed in [1]. (Right) illustrates the point correspondence between the two curves after applying the SP alignment and the FAR reparametrization.

I. REPARAMETRIZATION OF 2D CURVES WITH FAR

To visualize the proposed concept, we apply the FAR approach to the simple case of correspondence between 2D curved shapes in $\mathbb{R}^2$ (Euclidean space), as shown in Figure 1.

A 2D curve can be represented by a function $\alpha : I = [a, b] \to \mathbb{R}^2$ with I regulating the ordering of the curve. To normalize a curve α, the following preprocessing is carried out:

$$\alpha^*(t) = \frac{\alpha(t) - \bar{\alpha}}{h}, \tag{1}$$

with $\bar{\alpha} = \frac{1}{b-a} \int_I \alpha(t) dt$ and $h = \int_I \|\alpha(t) - \bar{\alpha}\|_2^2$. Note that $\|\cdot\|_2$ denotes the L_2 norm. Since this kind of curve is usually closed and our algorithm is adapted to open curves, the starting point correspondence is found using the protocol proposed in [1].

In this case, to compute the parametrization function $\alpha(t)$, the score function $f_\alpha(t)$ is defined as the curvature where

$$f_\alpha(t) = \frac{1}{\ddot{\alpha}^*(t)}, \tag{2}$$

with $\ddot{\alpha}^*(t) = \frac{d^2 \alpha^*(t)}{dt^2}$. In this case, since the discrete points describing the curve are usually equally spaced, the velocity is not informative. In contrast, the curvature can be seen as an interesting feature that can orient the resampling. When the curvature increases, the sampling step should become smaller in order to take into account crucial information. Then, a linear interpolation is performed before uniformly resampling the data. The qualitative example depicted in Figure 1 shows that the matching quality using FAR is improved compared to the work of [1] while keeping the same computational complexity $O(N)$, with N the number of curve nodes.

Methods	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
uniform rep (200 points) [1]	91	89	73	75	63	57	51	44	35	30
ours (50 points)	99	96	86	80	70	66	62	49	44	37
ours (200 points)	99	96	92	86	83	77	70	57	47	48

TABLE I
RETRIEVAL RESULTS ON KIMIA99

II. 2D CURVED SHAPE RETRIEVAL APPROACH

We present additional experiments on KIMIA99 and KIMIA216 in the context of 2D curved shape retrieval, using the presented representation in previous section. Then, an L_2 similarity measure is used to compare the different shapes after reparametrizing the curves using (1) and (2). We follow the same experimental protocol used in [1] which is similar to the leave-one-out. The accuracy score is measured by respectively computing the overall results of the top 10 and top 11 in KIMIA-99 and KIMIA-216 retrievals from the same class, excluding the query shape [2].

Methods	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th
uniform rep (200 points) [1]	212	210	188	181	174	165	159	151	141	132	120
ours (50 points)	216	210	199	190	179	175	166	149	141	127	105

TABLE II
RETRIEVAL RESULTS ON KIMIA216

We compare our method to the uniform representation proposed in [1].

We respectively report the obtained results for KIMIA99 and KIMIA216 in Table I and Table II. The results show that FAR outperforms uniform representation Pamigirum on KIMIA99 while conserving the same complexity $O(N^2)$, with $N = 200$ the number of points. Even when decreasing the number of points to $N = 50$ to reduce the computational time, FAR remains more efficient than uniform representation on KIMIA99 and on KIMIA216.

REFERENCES

[1] G. G. Demisse, D. Aouada, and B. Ottersten, "Deformation based curved shape representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 6, pp. 1338–1351, 2018.

[2] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 26, no. 5, pp. 550–571, 2004.