Rolling Rotations for Recognizing Human Actions from 3D Skeletal Data

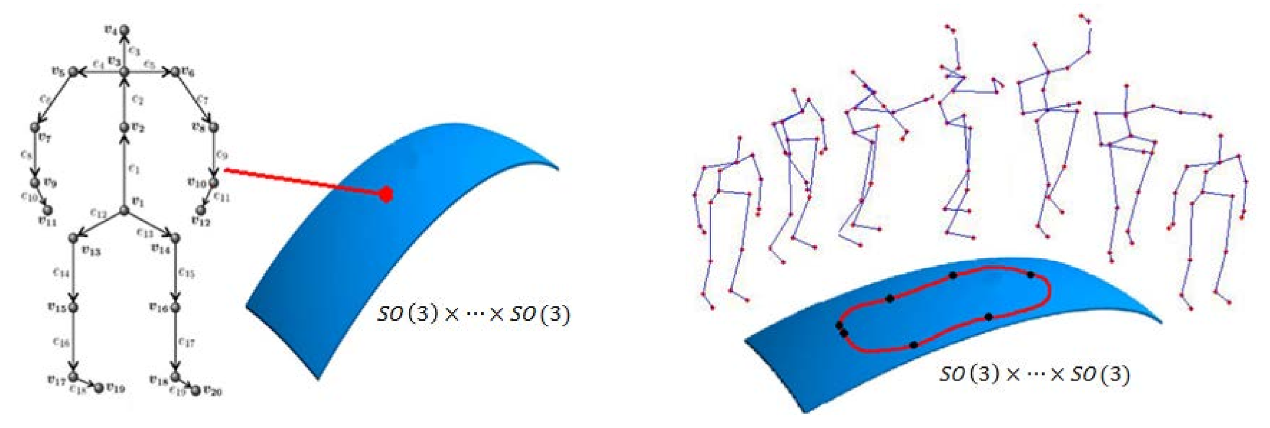
Raviteja Vemulapalli and Rama Chellappa

Department of Electrical and Computer Engineering, University of Maryland, College Park.

Main idea: Using rolling and logarithm maps to map temporal sequences from a manifold to its tangent space.

Skeletal Representation

- We represent a human skeleton using the relative 3D rotations between all pairs of body parts.
- 3D rotations are members of the special orthogonal group $SO(3)$.
- This rotation-based representation is scale-invariant.

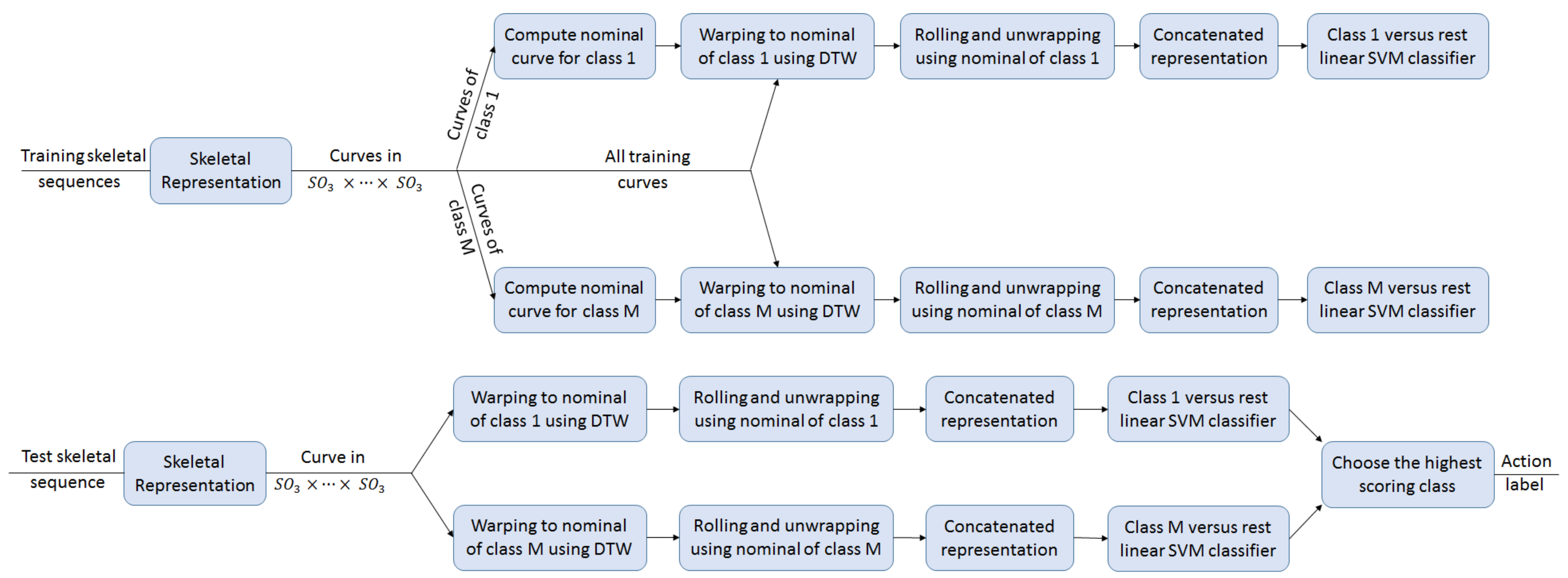


- $SO(3)\times\cdots\times SO(3)$ is a curved, smooth manifold.

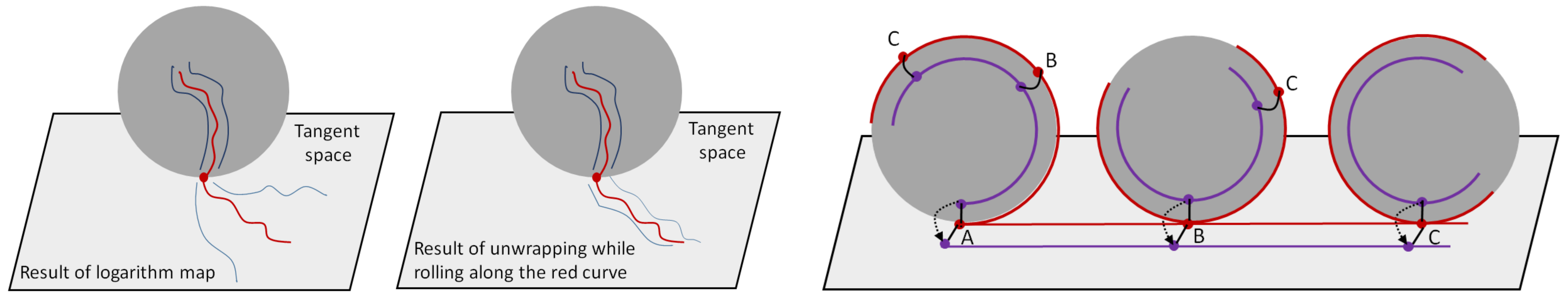
Action Recognition

- Dynamic Time Warping (DTW) is used to handle rate variations.
- Rolling and logarithm maps are used to map the curves from the Lie group $SO(3)\times\cdots\times SO(3)$ to its Lie algebra $\mathfrak{so}(3)\times\cdots\times\mathfrak{so}(3)$, which is the tangent space at the identity element.
- Linear SVM is used for classification in the tangent space.

Proposed Action Recognition System



Unwrapping while rolling



Theorem: Let $\{R_0, R_1, \ldots, R_T\}$ be a (discrete) curve in $SO(3)$. Let $\Omega_0, \Omega_1, \ldots, \Omega_T$ be T skew-symmetric matrices defined recursively using

$$\Omega_n = \log\left(e^{-\frac{\Omega_{n-1}}{2}} \ldots e^{-\frac{\Omega_1}{2}} R_n R_1^T e^{-\frac{\Omega_1}{2}} \ldots e^{-\frac{\Omega_{n-1}}{2}}\right).$$

Let $C(t) = \ (U(t), V(t), X(t)) \in SO(3)^2 R^9$ be a curve defined as

$$U(t) = e^{-\frac{(t-n+1)\Omega_n}{2}} e^{-\frac{\Omega_{n-1}}{2}} \ldots e^{-\frac{\Omega_1}{2}},\ \ V(t) = R_0^T e^{\frac{(t-n+1)\Omega_n}{2}} e^{\frac{\Omega_{n-1}}{2}} \ldots e^{\frac{\Omega_1}{2}} R_0,$$

$$X(t) = R_0 \textstyle\sum_{i=1}^{n-1} \Omega_i + (t-n+1)\Omega_n R_0,\ \ t \in [n-1, n],\ \ n = 1,2,\ldots,T.$$

Then, the action of $C(t) = \ (U(t), V(t), X(t))$ on $SO(3)$ defined as $(U, V, X) \circ Z = UZV^T + X;\ Z \in SO(3)$ results in rolling of $SO(3)$ over the tangent plane at R_0 along the curve $R(t)$.

Unwrapping a curve $\beta(t) \in SO(3)$ while rolling along $R(t)$ gives the following curve $\bar{\beta}(t) \in T_{R_0}SO(3)$

$$\bar{\beta}(t) = log_{SO(3)}(R_0, U(t)\beta(t)V(t)^T) + R_0 + X(t).$$

Theorem: Unwrapping while rolling preserves the distances between the (blue) curves that are being unwrapped and the (red) rolling curve.

Experimental Results

Dataset	Logarithm map at a point	Unwrapping while rolling
Florence3D	86.83	**89.82**
MSR Pairs	92.96	**94.09**
G3D-Gaming	87.82	**87.95**

- Comparison with state-of-the-art

Florence3D dataset

Approach	Accuracy
Elastic functional coding	89.67
Relative 3D geometry	90.71
Proposed (no FTP)	89.82
Proposed (FTP)	**91.40**

MSR Pairs dataset

Approach	Accuracy
Relative 3D geometry	93.65
Proposed (no FTP)	94.09
Proposed (FTP)	**94.67**

G3D Gaming dataset

Approach	Accuracy
RBM + HMM	86.40
Relative 3D geometry	**91.09**
Proposed (no FTP)	87.95
Proposed (FTP)	90.94

References

- K. Hupper and F. S. Leite, "On the Geometry of Rolling and Interpolation Curves on Sn, SOn, Grassmann Manifolds", Journal of Dynamical and Control Systems, 13(4):467–502, 2007.
- R. Vemulapalli, F. Arrate, and R. Chellappa, "Human Action Recognition by Representing 3D Skeletons as Points in a Lie Group", In CVPR, 2014.
- R. Caseiro, P. Martins, J. F. Henriques, F. S. Leite, and J. Batista, "Rolling Riemannian Manifolds to Solve the Multiclass Classification Problem", In CVPR, 2013