# Human Action Recognition Using Tensor Dynamical System Modeling

Chan-Su Lee

Yeungnam University, Rep. of Korea

# Contents

- Introduction
- Tensor Dynamical Models
- Statistical Analysis on Grassmann Manifold for Tensor Dynamical Models
- Experimental Result
- Conclusion

# Introduction

- Dynamics of human motion can be represented by spatiotemporal tensor time series
  - Motion capture data: 3D sequence in each joint(marker) time sequence

    → 3D vector x joint/marker # x t
  - Activity video: 2D image(x,y) x RGB x t

    → 2D vector x 3 (RGB) x t
- Useful to preserver original tensor structure of the data to analyze human motion
- Provide model to predict/generate future motion

# Tensor Normal Distribution

- Normal distribution

$$f_x(x) = (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

x: scalar

- Extension to vector: Multivariate normal distribution

x: vector

$$f_x(\boldsymbol{x}) = (2\pi)^{-\frac{1}{2}p} |\boldsymbol{\Sigma}|^{-1/2} e^{-\frac{1}{2}tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{u})(\boldsymbol{x}-\boldsymbol{u})'\}}$$

- Extension to matrix: Bilinear normal distribution

X: matrix

$$f_{\boldsymbol{X}}(X) = (2\pi)^{-\frac{1}{2}pn} |\boldsymbol{\Sigma}|^{-n/2} |\boldsymbol{\Psi}|^{-p/2} e^{-\frac{1}{2}tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{u})\Psi^{-1}(\boldsymbol{X}-\boldsymbol{u})'\}}$$

# Tensor Normal Distribution

- Bilinear normal distribution using unit vector

$$\sum_{ij} \quad X_{ij}\mathbf{e}_i^1(\mathbf{e}_j^2)'$$

$$= \quad \sum_{ij}\mu_{ij}\mathbf{e}_i^1(\mathbf{e}_j^2)' + \sum_{ij}\sum_{km}\tau_{ik}\gamma_{mj}U_{km}\mathbf{e}_i^1(\mathbf{e}_j^2)'$$

$$= \quad \sum_{ij}\mu_{ij}\mathbf{e}_j^2 \otimes \mathbf{e}_i^1 + \sum_{ij}\sum_{km}\tau_{ik}\gamma_{mj}U_{km}\mathbf{e}_j^2 \otimes \mathbf{e}_i^1$$

Unit basis: $\mathbf{e}_i^1 \, : \, p \times 1, \, \mathbf{e}_j^2 \, : \, n \times 1$

$U_{km} \sim N(0,1)$

$\mathbf{e}_i^1(\mathbf{e}_j^2)' \to \mathbf{e}_j^2 \otimes \mathbf{e}_i^1$

# Tensor Normal Distribution

- Tensor normal distribution

$$N_{p_1,p_2,\cdots,p_k}(\boldsymbol{\mu}, \sum_k, \sum_1, \sum_2, \cdots, \sum_{k-1})$$

$$\sum_{i_1,i_2,\cdots,i_k} X_{i_1,\cdots,i_k} \boldsymbol{e}^1_{i_1} \otimes \boldsymbol{e}^2_{i_2} \otimes \cdots \otimes \boldsymbol{e}^k_{i_k}$$

$$= \sum_{i_1,i_2,\cdots,i_k} u_{i_1,\cdots,i_k} \boldsymbol{e}^1_{i_1} \otimes \boldsymbol{e}^2_{i_2} \otimes \cdots \otimes \boldsymbol{e}^k_{i_k}$$

$$+ \sum_{i_1}^{p_1} \sum_{j_1}^{p_1} \sum_{j_2}^{p_2} \cdots \sum_{j_k}^{p_k} \tau^1_{i_1 j_1} \tau^2_{i_2 j_2}$$

$$\cdots \tau^k_{i_k j_k} U_{j_1 j_2 \cdots j_k} \boldsymbol{e}^1_{i_1} \otimes \boldsymbol{e}^2_{i_2} \otimes \cdots \boldsymbol{e}^k_{i_k},$$

$$\boldsymbol{\mathcal{A}} \circledast \boldsymbol{\mathcal{X}} = \sum_{j_1 \cdots j_M} \mathcal{A}_{i_1 \cdots i_M j_1 \cdots j_M} \mathcal{X}_{j_1 \cdots j_M}$$

$$vec(\boldsymbol{\mathcal{A}} \circledast \boldsymbol{\mathcal{X}}) = mat(\boldsymbol{\mathcal{A}}) vec(\boldsymbol{\mathcal{X}})$$

# Tensor Dynamical Models

- Linear dynamical system

$$\boldsymbol{x}_{n+1} = \boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{w}$$
$$\boldsymbol{y}_n = \boldsymbol{C}\boldsymbol{x}_n + \boldsymbol{v}$$

$$\begin{aligned} \boldsymbol{x}_n|\boldsymbol{x}_{n-1} &\sim N_k(\boldsymbol{x}_n|A\boldsymbol{x}_{n-1}, \boldsymbol{Q}) \\ \boldsymbol{y}_n|\boldsymbol{x}_n &\sim N_p(\boldsymbol{y}_n|C\boldsymbol{x}_n, \boldsymbol{R}) \end{aligned}$$

# Tensor Dynamical Models

- Tensor extension of linear dynamic models

$$\mathcal{Y}_{1,\cdots,N} = [\mathcal{Y}_1, \cdots, \mathcal{Y}_N] \quad \mathcal{Y}_n \in \mathbb{R}^{I_1 \times \cdots \times I_M}$$
$$\mathcal{X}_{1,\cdots,N} = [\mathcal{X}_1, \cdots, \mathcal{X}_N] \quad \mathcal{X}_n \in \mathbb{R}^{J_1 \times \cdots \times J_M}$$

$$\mathcal{X}_{n+1} | \mathcal{X}_n \sim \mathcal{N}(\mathcal{A} \circledast \mathcal{X}_n, \boldsymbol{Q})$$
$$\mathcal{Y}_n | \mathcal{X}_n \sim \mathcal{N}(\mathcal{C} \circledast \mathcal{X}_n, \mathcal{R})$$

# Tensor Dynamical Models

- Parameter estimation
  - Estimating parameters:

$$\theta = (\mathcal{U}_0, \mathcal{Q}_0, \mathcal{Q}, \mathcal{A}, \mathcal{R}, \mathcal{C}) \qquad \mathcal{X}_1 \sim \mathcal{N}(\mathcal{U}_0, \mathcal{Q}_0)$$

  - Metricized and vectorized representation

$$L(\theta|\mathcal{X}_{1,\cdots,N}, \mathcal{Y}_{1,\cdots,N}) = L(vec(\theta)|vec(\mathcal{X}_{1,\cdots,N}), vec(\mathcal{Y}_{1,\cdots,N})$$
$$vec(\theta) \ = vec(\mathcal{U}_0)), mat(\mathcal{Q}_0), mat(\mathcal{Q}), mat(\mathcal{A}), mat(\mathcal{R}), mat(\mathcal{C}))$$

  - Maximize vectorized projection matrix

$$v = [vec(C^{(1)})^T \cdots vec(C^{(M)})^T]^T$$

# Statistical Analysis for the Tensor Dynamical Models

- Characteristics of dynamic systems

$$\mathcal{A} = A^{(M)} \otimes \cdots \otimes A^{(1)}$$
$$\mathcal{C} = C^{(M)} \otimes \cdots \otimes C^{(1)}$$

- Observation of linear dynamical systems

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \end{bmatrix} y_0$$

$$O^T = [C^T (CA)^T (CA^2)^T \cdots (CA^{(m-1)})]$$

# Statistical Analysis for the Tensor Dynamical Models

- Observation of tensor dynamical systems

$$O^{(1)^T} = [C^T (CA^{(1)})^T (CA^{2(1)}))^T \cdots (CA^{(m-1)(1)})^T]$$

$$O^{(2)^T} = [C^T (CA^{(2)})^T (CA^{2(2)}))^T \cdots (CA^{(m-1)(2)})^T]$$

$$\vdots$$

$$O^{(n)^T} = [C^T (CA^{(n)})^T (CA^{2(n)}))^T \cdots (CA^{(m-1)(n)})^T]$$

  – Collection of N latent tensor dynamical models

→N Points on Grassmann manifold

→Geodesic distance measurement using Riemannian structure[2]

# Statistical Analysis for the Tensor Dynamical Models

- How to use the geodesic distance of Riemainnian manifold to Tensor dynamic model
  - Summation of Geodesic distance of linear dynamic model for each mode of tensor dynamic model→ order invariant distance
  - Pairwise distance of Geodesic distance of linear dynamic model for each mode of tensor dynamic model→ order preserving distance

# Experimental Results

- Dataset: INRIA human action database
  - 10 actors with 11 actions
  - Leave one-subject out evaluation
  - view invariant feature representation→ order invariant distance

| Method | Motion history+Mahl. [a] | Grassmann+Linear DS[b] | Proposed Tensor DS |
|---|---|---|---|
| Recognition rate | 93.3 | 93.9 | 95.46 |

[a] D. Weinland, R. Ronfard, and E. Boyer. Free viewpoint action recognition using motion history volumes. CVIU, 104(2):249–257, 2006
[b] P. K. Turaga, A. Veeraraghavan, A. Srivastava, and R. Chellappa. Statistical computations on grassmann and stiefel manifolds for image and video-based recognition. IEEE Trans. PAMI, 33(11):2273–2286, 2011

# Conclusion and future work

- Tensor dynamical system is useful for human action recognition
- Statistical analysis on the Grassmann manifold using geodesic distance enhances the classification performance of latent tensor dynamical systems

# Thank you