Paper ID: #14

ASPIRE 2024, Hiroshima

Human Pose Refinement Using Line Graphs

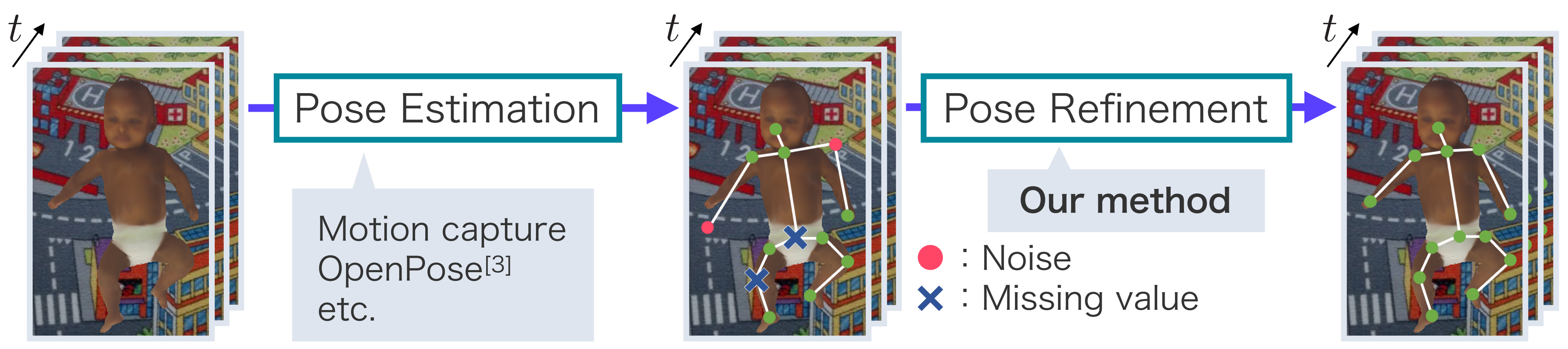
Hayate Kojima[1], Yuichi Tanaka[2]

[1]Tokyo University of Agriculture and Technology, [2]Osaka University

Introduction

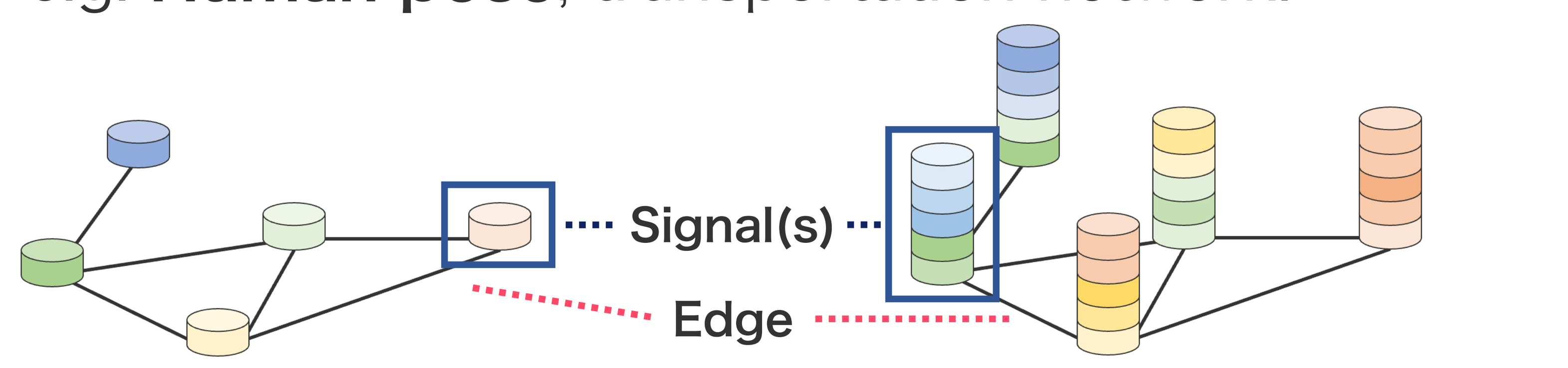
Purpose

- Recovering the original time-varying pose from the observed signals (with **missing values** and **noise**)



Graph Signal Processing (GSP)[1]

- Signals often have their underlying structures.
- GSP can consider the underlying structure of signals.
  e.g. **Human pose**, transportation network.



Proposed Method

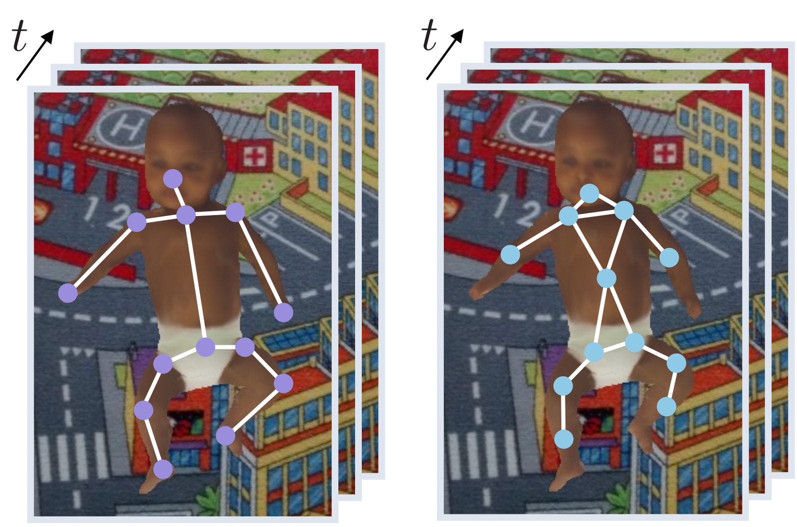
Main Idea

- Signals (=coordinates of joints) and weight of the edges (=joint length) change smoothly in a spatiotemporal.
- Coordinates are restored using refined joint length.
- Line graph transform is used for the restoration of the edges.

Line graph: A graph that represents the relationship between edges by making the edges of the original graph as nodes.

$$\mathbf{A}_L = \mathbf{B}^\top \mathbf{B} - 2\mathbf{I}$$

$\mathbf{A}_L$: Adjacency matrix of line graph
$\mathbf{B}$: Degree matrix of original graph
$\mathbf{X}$: Original time-varying graph
$\mathbf{Z}$: Time-varying line graph



Design of Optimization Problem

Edge Refinement Using Line Graphs

$$\min_{\tilde{\mathbf{Z}}} \|\mathbf{J}_L \circ \tilde{\mathbf{Z}} - \mathbf{Z}\|_2^2 + \alpha_{edge}\, \underline{\mathrm{tr}((\tilde{\mathbf{Z}}\mathbf{D})^\top \mathbf{L}_L (\tilde{\mathbf{X}}\mathbf{D}))}$$

Spatiotemporal smoothness of edge

Node Refinement Using Restored Edges

$$\min_{\tilde{\mathbf{X}}_{(d)}} \|\mathbf{J} \circ \tilde{\mathbf{X}}_{(d)} - \mathbf{Y}\|_2^2 + \alpha_{node}\, \underline{\mathrm{tr}((\tilde{\mathbf{X}}_{(d)}\mathbf{D})^\top \mathbf{L} (\tilde{\mathbf{X}}_{(d)}\mathbf{D}))}$$

Spatiotemporal smoothness of coordinates

$\alpha_{node}, \alpha_{edge}$: Hyperparameter
$\mathbf{Z} \in \mathbb{R}^{|\epsilon| \times T}$: Time-varying line graph signal
$\mathbf{J} \in \{0,1\}^{N \times T}$: Sampling operator
$\mathbf{J}_L \in \{0,1\}^{|\epsilon| \times T}$: Sampling operator
$\mathbf{L} \in \mathbb{R}^{N \times N}$: Graph Laplacian
$\mathbf{L}_L \in \mathbb{R}^{|\epsilon| \times |\epsilon|}$: Graph Laplacian of line graph

$\mathbf{X}_{(d)} \in \mathbb{R}^{N \times T}$: Observed time-varying graph signal of axis d

$$\mathbf{D} = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{bmatrix} \in \mathbb{R}^{T \times (T-1)}$$

: Temporal high pass filter

Refinement Algorithm

Algorithm 1 Human Pose Refinement Using Line Graphs

1) Calculate line graph signal
 $\mathbf{A}_L = \mathbf{B}^\top \mathbf{B} - 2\mathbf{I}$
 $\mathbf{L}_L, \mathbf{J}_L \leftarrow \mathbf{A}_L$
 $[\mathbf{Z}]_{e,t} \leftarrow$ set length of joint e on time t
2) Edge Refinement
 $\tilde{\mathbf{Z}} = f_{CD}(\mathbf{Z}, \mathbf{L}_L, \mathbf{J}_L)$
3) Calculate $\mathbf{L}$ from $\tilde{\mathbf{Z}}$
 for $t = 1$ to T
  $\mathbf{L} \leftarrow \tilde{\mathbf{Z}}$
 end for
4) Pose Refinement
 for $d = 1$ to D
  $\tilde{\mathbf{X}}_{(d)} = f_{CD}(\mathbf{X}_{(d)}, \mathbf{L}, \mathbf{J})$
 end for

Experiments

Dataset

- MINI-RGBD[2] (with pose estimated by OpenPose[3])

Compared Methods

- Graph temporal difference construction[4] (OGTR)
- Sobolev norm regularization[5] (TRSS)

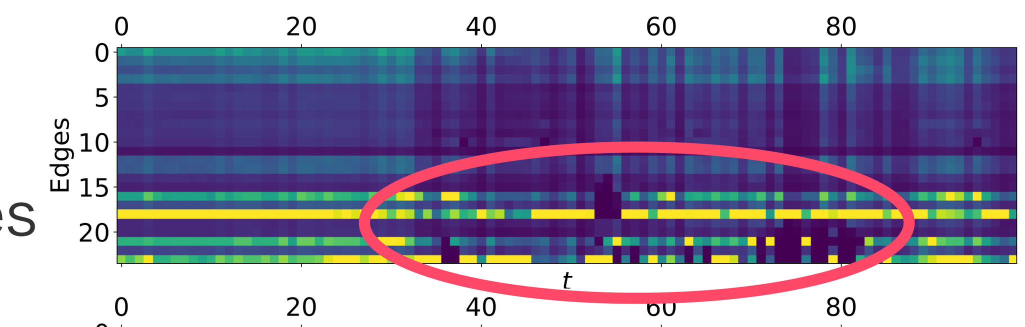
Results

① Restoration Performance (RMSE)

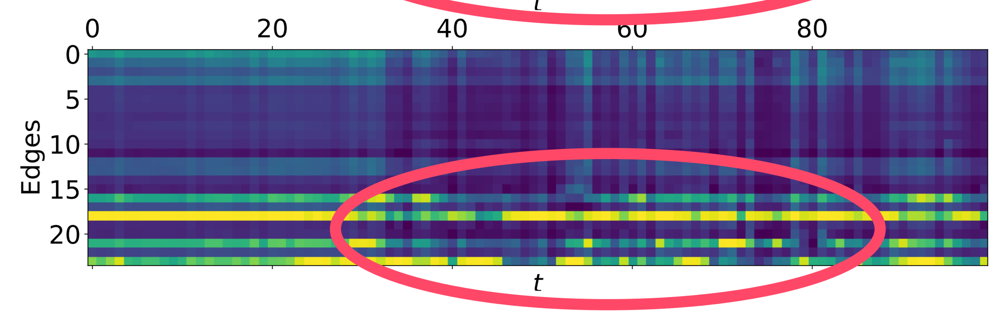
Method \# dataset	#1	#2	#3	#4	#5
OGTR	16.33	93.66	23.25	36.13	45.21
TRSS	16.33	95.97	24.60	38.24	47.29
Proposed	**9.71**	**92.90**	**22.52**	**32.82**	**44.47**

② Restored Edges

Before restoration:
 Contain missing values



After restoration:
 Smoothly interpolated



Pose refinement is performed using the information from the **restored edges**

Conclusion

- Our proposed method considers spatiotemporal smoothness of nodes and edges.
- It improves the accuracy of restoring human pose.

References

[1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processinon graphs: Extending high-dimensional data analysis to networks and other irregular domains," IEEE Signal Process. Mag., vol. 30, no. 3, pp. 83–98, May 2013.
[2] N. Hesse, C. Bodensteiner, M. Arens, U. G. Hofmann, R. Weinberger, and A. Sebastian Schroeder, "Computer Vision for Medical Infant Motion Analysis: State of the Art and RGB-D Data Set," in Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018, pp. 0–0.
[3] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," IEEE Trans Pattern Anal Mach Intell, vol. 43, no. 1, pp. 172–186, Jan. 2021.
[4] K. Qiu, X. Mao, X. Shen, X. Wang, T. Li, and Y. Gu, "Time-Varying Graph Signal Reconstruction," IEEE J. Sel. Top. Signal Process., vol. 11, no. 6, pp. 870–883, Sep. 2017.
[5] J. H. Giraldo, A. Mahmood, B. Garcia-Garcia, D. Thanou, and T. Bouwmans, "Reconstruction of Time-Varying Graph Signals via Sobolev Smoothness," IEEE Trans. on Signal and Inf. Process. over Networks, vol. 8, pp. 201–214, 2022.

Acknowledgement

This work was supported in part by JSPS KAK- ENHI under Grant 23K17461 and JST AdCORP under Grant JP- MJKB2307.