MoCoSys: Human Motion Correction based on Deep Learning Coupled with 3D+t Laplacian Motion Representation

Mansour Tchenegnon[1] Sylvie Gibet[1] Thibaut Le Naour[2]

[1] *IRISA (Expression), Université Bretagne Sud*

[2] *Motion-Up*

GTAS 2025, July 8th

Contents

Contents

Pose Estimation for Motion Reconstruction

- Estimation of joint positions from video
- Estimation with deep learning approaches

Pose Estimation for Motion Reconstruction

- Potential uses:
  - Motion analysis, anomaly detection, fall detection, etc.
  - **Data-driven animation**

Pose Estimation for Motion Reconstruction

Limitations for use in data-driven animation

- Temporal incoherence in the pose sequence
- Skeletal inconsistency (no preservation of the skeletal structure throughout the motion)

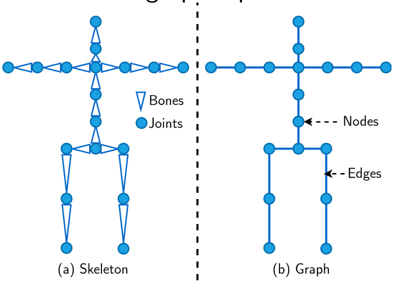
Objectives

- Estimate pose sequence from video with state-of-the-art solution
- Post-process pose sequence to produce motion usable for data-driven animation
  - Ensure skeletal consistency
  - Improve temporal coherence
  - Reduce bone length errors

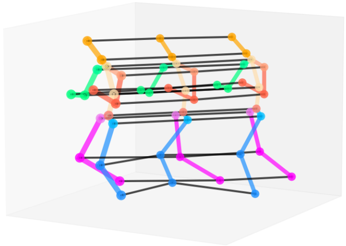
Contents

Motion Graph

Posture to graph representation

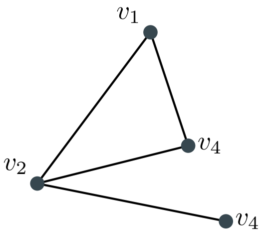


Motion graph representation



Discrete Laplacian Operator

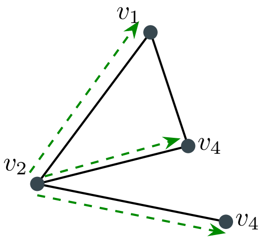
- Let a graph $G = (V, E)$ where $V = \boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n$ are Cartesian coordinates of the vertices in $\mathbb{R}^3$
- Laplacian operator : $\mathcal{L}(\boldsymbol{v}_i) = \boldsymbol{\delta}_i = \sum_{j \in \mathcal{N}(i)} w_{ij}(\boldsymbol{v}_i - \boldsymbol{v}_j)$



$$\begin{pmatrix} 2 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \boldsymbol{v}_3 \\ \boldsymbol{v}_4 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \boldsymbol{\delta}_3 \\ \boldsymbol{\delta}_4 \end{pmatrix}$$

Discrete Laplacian Operator

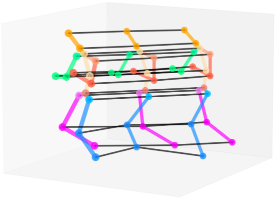
- Let a graph $G = (V, E)$ where $V = \boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n$ are Cartesian coordinates of the vertices in $\mathbb{R}^3$
- Laplacian operator : $\mathcal{L}(\boldsymbol{v}_i) = \boldsymbol{\delta}_i = \sum_{j \in \mathcal{N}(i)} w_{ij}(\boldsymbol{v}_i - \boldsymbol{v}_j)$



$$\begin{pmatrix} 2 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \boldsymbol{v}_3 \\ \boldsymbol{v}_4 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \boldsymbol{\delta}_3 \\ \boldsymbol{\delta}_4 \end{pmatrix}$$

Application to Motion Graph

- Applying Laplacian operator to the motion graph



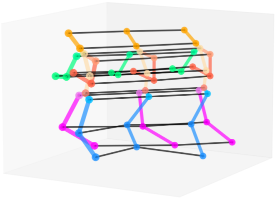
$$\mathcal{L}(\boldsymbol{v}_{i,t}) = \boldsymbol{\delta}_{i,t} = \sum_{j \in \mathcal{N}_t(i)} w_{ij,t}(\boldsymbol{v}_{i,t} - \boldsymbol{v}_{j,t})$$

$$+ w^{-}(\boldsymbol{v}_{i,t} - \boldsymbol{v}_{i,t-1}) + w^{+}(\boldsymbol{v}_{i,t+1} - \boldsymbol{v}_{i,t})$$

(Le Naour et al., 2013)

Application to Motion Graph

- Applying Laplacian operator to the motion graph



spatial links (same posture)

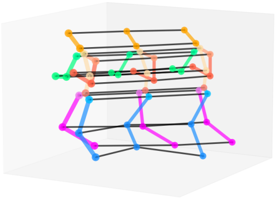
$$\mathcal{L}(\boldsymbol{v}_{i,t}) = \boldsymbol{\delta}_{i,t} = \boxed{\sum_{j \in \mathcal{N}_t(i)} w_{ij,t}(\boldsymbol{v}_{i,t} - \boldsymbol{v}_{j,t})}$$

$$+ w^{-}(\boldsymbol{v}_{i,t} - \boldsymbol{v}_{i,t-1}) + w^{+}(\boldsymbol{v}_{i,t+1} - \boldsymbol{v}_{i,t})$$

(Le Naour et al., 2013)

Application to Motion Graph

- Applying Laplacian operator to the motion graph



spatial links (same posture)

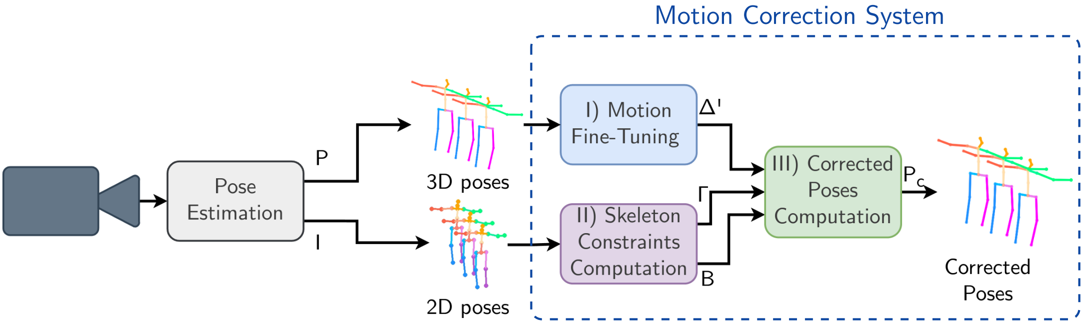
$$\mathcal{L}(\boldsymbol{v}_{i,t}) = \boldsymbol{\delta}_{i,t} = \boxed{\sum_{j \in \mathcal{N}_t(i)} w_{ij,t}(\boldsymbol{v}_{i,t} - \boldsymbol{v}_{j,t})}$$

$$+ \boxed{w^{-}(\boldsymbol{v}_{i,t} - \boldsymbol{v}_{i,t-1}) + w^{+}(\boldsymbol{v}_{i,t+1} - \boldsymbol{v}_{i,t})}$$

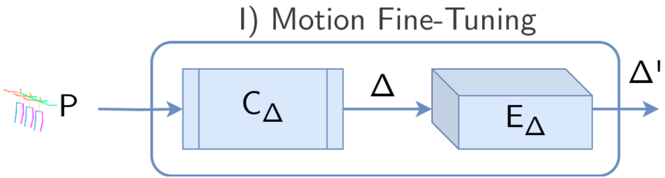
temporal links (adjacent postures)

(Le Naour et al., 2013)

Approach

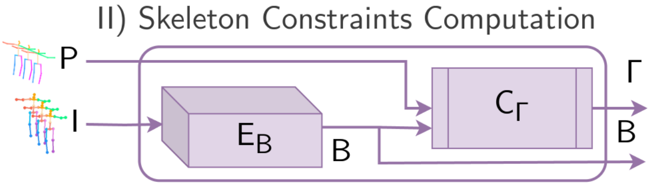

Motion Correction System
Pose
Estimation
P
I
3D poses
2D poses
I) Motion
Fine-Tuning
II) Skeleton
Constraints
Computation
Δ'
Γ
B
III) Corrected
Poses
Computation
P_c
Corrected
Poses

Motion Fine-Tuning



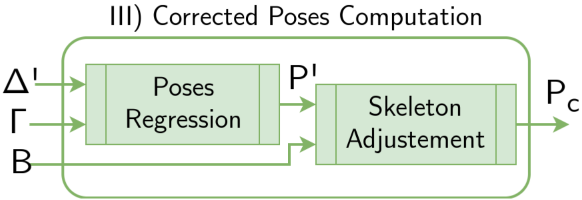
- **Algorithm** $C_\triangle$ computes the differential coordinates $\triangle$ for all joints
- **Neural networks** $E_\triangle$ fine tunes the differential coordinates

Skeleton Constraints Computation



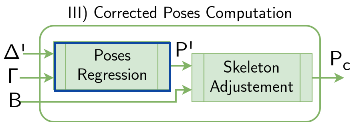
- **Neural networks** E_B estimate bone lengths
- **Algorithm** C_Γ computes skeleton constraints Γ as bone vectors with direction and length
  - **direction** from pose sequence P
  - **length** from bone lengths B

Corrected Poses Computation



- Two algorithms
  - **Algorithm "Poses Regression"**
  - **Algorithm "Skeleton Adjustment"**

Corrected Poses Computation

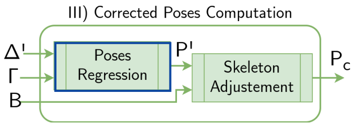


Algorithm "Poses Regression"

- Computes P' (fine-tuned pose sequence)
- By solving equation system

$$\begin{pmatrix} L \\ U \\ D \end{pmatrix} P' = \begin{pmatrix} \Delta' \\ C_p \\ \Gamma \end{pmatrix}$$

Corrected Poses Computation



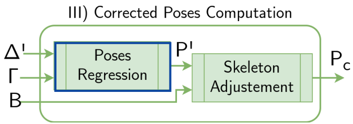
Algorithm "Poses Regression"

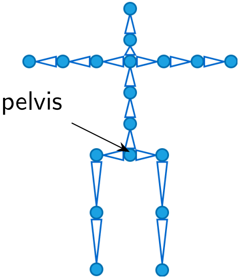
- Computes P' (fine-tuned pose sequence)
- By solving equation system
  - of fine tuned Laplacian coordinates Δ'

Laplacian matrix

$$\begin{pmatrix} \boxed{L} \\ U \\ D \end{pmatrix} P' = \begin{pmatrix} \boxed{\Delta'} \\ C_p \\ \Gamma \end{pmatrix}$$

Corrected Poses Computation





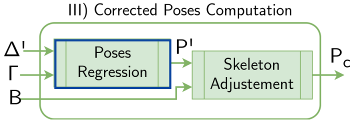
Algorithm "Poses Regression"

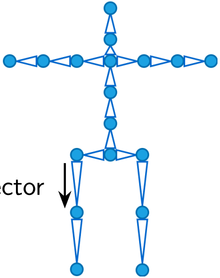
- Computes P' (fine-tuned pose sequence)
- By solving equation system
  - of fine tuned Laplacian coordinates Δ'
  - of positional constraints on the pelvis C_p

Pelvis constraints matrix

$$\begin{pmatrix} L \\ \boxed{U} \\ D \end{pmatrix} P' = \begin{pmatrix} \Delta' \\ \boxed{C_p} \\ \Gamma \end{pmatrix}$$

Corrected Poses Computation





constraint vector

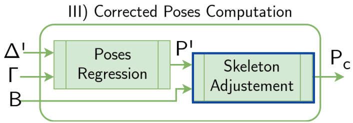
Algorithm "Poses Regression"

- Computes P' (fine-tuned pose sequence)
- By solving equation system
  - of fine tuned Laplacian coordinates Δ'
  - of positional constraints on the pelvis C_p
  - of skeleton constraints Γ

$$\begin{pmatrix} L \\ U \\ \boxed{D} \end{pmatrix} P' = \begin{pmatrix} \Delta' \\ C_p \\ \boxed{\Gamma} \end{pmatrix}$$

Skeleton constraints matrix

Corrected Poses Computation



Algorithm "Skeleton Adjustment"

- Ensure skeletal consistency in final pose sequence P_c
- By solving the equation system
  - of new skeleton constraints Γ' from P' and B
  - of positional constraints on the pelvis C_p

$$\begin{pmatrix} U \\ D \end{pmatrix} P_c = \begin{pmatrix} C_p \\ \Gamma' \end{pmatrix}$$

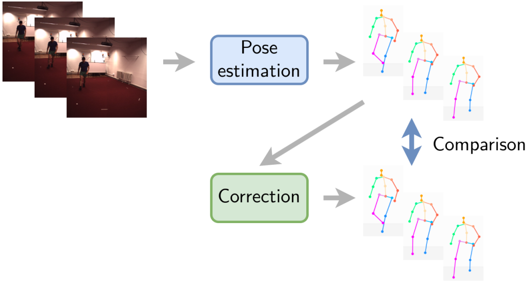
Contents

Experiment Setup

- **Dataset** : Human36m dataset
- **Training neural networks** E_Δ
  - Supervised learning (75% train, 25% test)
  - Input : estimated poses from a SOTA solution *(Chen et al., 2022)*
  - Loss function : $\mathcal{L}_\Delta = \frac{1}{N}\sum_{1}^{N} \|\Delta^{gt} - \Delta'\|$

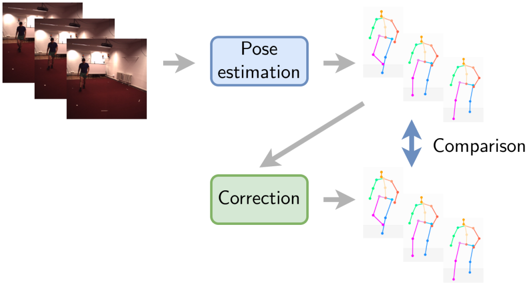
Evaluation Protocol

- **Evaluation** : comparison of error metrics between pose sequence after estimation and after correction
- **Evaluation metrics**
  - joint position, velocity, acceleration, bone length



Evaluation Protocol

- **Evaluation** : comparison of error metrics between pose sequence after estimation and after correction
- **Evaluation metrics**
  - joint position, velocity, acceleration, bone length
  - Metric for skeletal consistency ?



Evaluation Protocol

- Skeletal consistency metric
  - average standard deviation of bone length throughout the motion

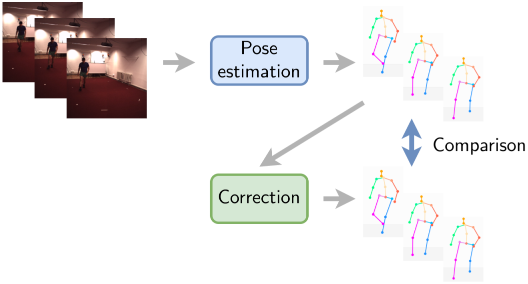
$$\sigma_b = \frac{1}{N} \sum_{t=1}^{T} \sqrt{(d_{b,t} - \mu_b)^2}$$

$$e_\sigma = \frac{1}{card(\mathcal{B})} \sum_{b \in \mathcal{B}} \sigma_b$$

$\mathcal{B}$ set of skeleton bones

T motion sequence length

$d_{b,t}$ length of bone b at frame t



Quantitative Results

	pos.(mm)	**vel.(mm/f)**	**acc.(mm/f²)**	**bone(mm)**	e_σ**(mm)**
Estimation	**44.63**	2.64	2.21	7.70	1.79
Correction	44.88	**2.27**	**1.00**	**3.76**	**0**

SOTA pose estimator : AANet *(Chen et al., 2022)*

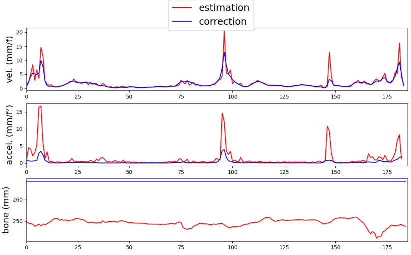
	pos.(mm)	**vel.(mm/f)**	**acc.(mm/f²)**	**bone(mm)**	e_σ**(mm)**
Estimation	47.61	2.69	1.55	10.28	7.56
Correction	**46.89**	**2.47**	**1.02**	**3.76**	**0**

SOTA pose estimator : PoseFormerV2 *(Zhao et al., 2023)*

	pos.(mm)	**vel.(mm/f)**	**acc.(mm/f²)**	**bone(mm)**	e_σ**(mm)**
Estimation	53.47	3.12	1.96	**3.08**	0
Correction	**52.85**	**2.73**	**1.22**	3.76	0

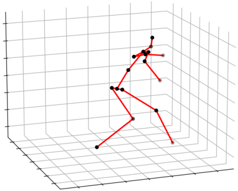
SOTA pose estimator : MotioNet *(Shi et al., 2020)*

Visual Results

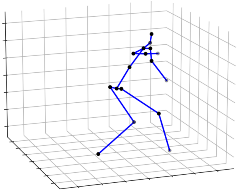


Curves

Estimation



Correction



Contents

Conclusion

- **Advantages**
  - Ensures skeletal consistency in corrected poses
  - Improves temporal quality of the motion
- **Limitations**
  - Spatial accuracy dependent on the pose estimator
- **Future work**
  - Build a complete pipeline body reconstruction (Mediapipe with correction)
  - Experiment correction on hand gesture reconstruction
  - Body and hand reconstruction (with Mediapipe)

Questions ?

Thank you !

References

(Chen, 2022) Chen, T., Fang, C., Shen, X., Zhu, Y., Chen, Z., Luo, J.: Anatomy-aware 3d human pose estimation with bone-based pose decomposition. IEEE Trans. Circuits Syst. Video Technol. 32(1), 198–209 (2022) https://doi.org/10.1109/TCSVT.2021.3057267

(Le Naour, 2013) Le Naour, T., Courty, N., Gibet, S.: Spatiotemporal coupling with the 3D+t motion Laplacian. Computer Animation and Virtual Worlds 24(3-4), 419–428 (2013) https://doi.org/10.1002/cav.1518 . Number: 3-4

(Peng, 2024) Peng, J., Zhou, Y., Mok, P.Y.: KTPFormer: Kinematics and Trajectory Prior Knowledge-Enhanced Transformer for 3D Human Pose Estimation. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1123–1132. IEEE, Seattle, WA, USA (2024). https://doi.org/10.1109/CVPR52733.2024.00113

(*(Shi, 2021)*) Shi, M., Aberman, K., Aristidou, A., Komura, T., Lischinski, D., Cohen-Or, D., Chen, B.: MotioNet: 3D Human Motion Reconstruction from Monocular Video with Skeleton Consistency. ACM Trans. Graph. 40(1), 1–15 (2021) https://doi.org/10.1145/3407659

(Zhao, 2023) Zhao, Q., Zheng, C., Liu, M., Wang, P., Chen, C.: PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8877–8886. IEEE, Vancouver, BC, Canada (2023). https://doi.org/10.1109/CVPR52729.2023.00857