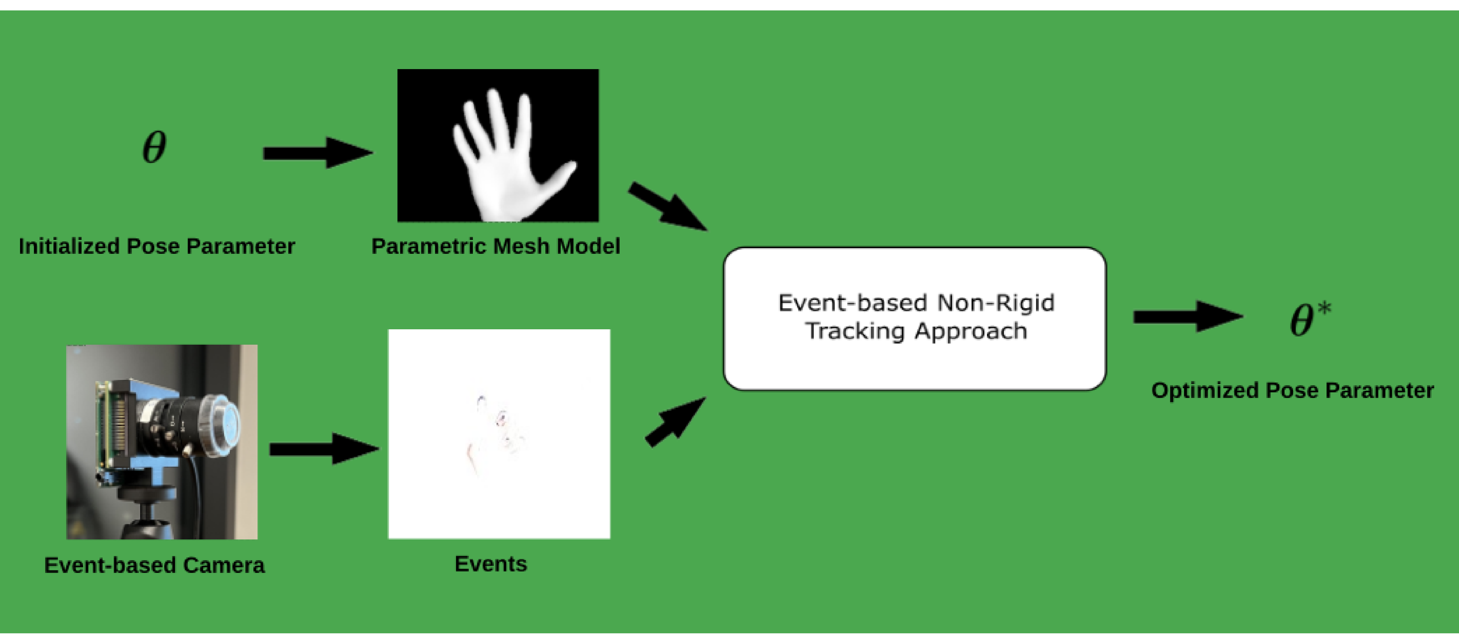
Event-based Non-Rigid Reconstruction from Contours

Yuxuan Xue[1,2], Haolong Li[1], Stefan Leutenegger[2], Joerg Stueckler[1]

[1]Embodied Vision Group, Max Planck Institute for Intelligent Systems, Tübingen, Germany
[2] Technical University of Munich, Munich, Germany

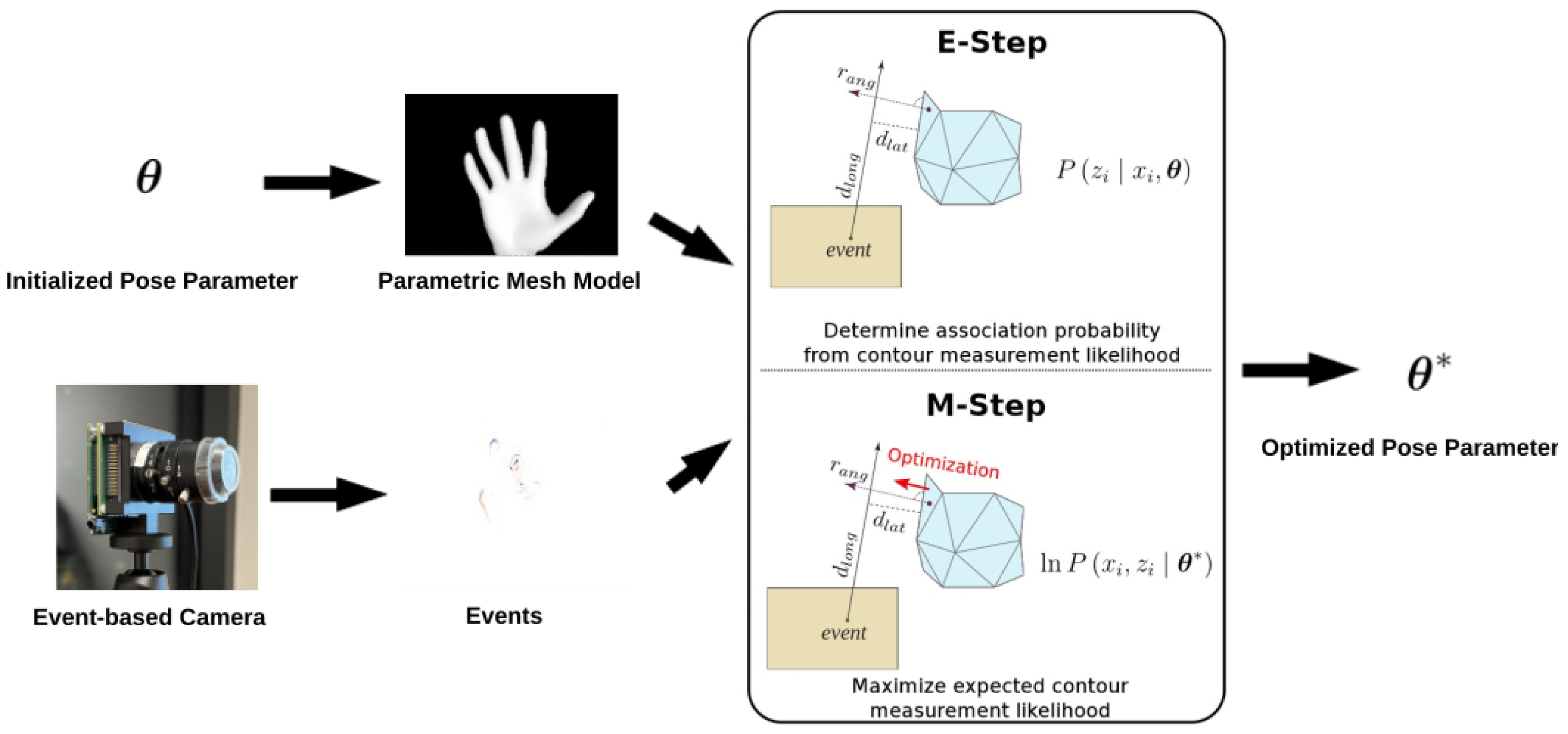


We propose a novel approach for **reconstructing fast non-rigid object deformations** using measurements from **event-based cameras**.

We observe that the majority of events of texture-less object motions are generated at the object **contour**. Our approach estimates the deformation of objects from **events generated at the object contour** in a **probabilistic optimization framework**.

Compared to baseline approaches, our method outperforms them in non-rigid reconstruction quantitatively and qualitatively.

Event-based Non-Rigid Reconstruction



E-step: Determine association probability based on latest MANO[6] pose parameter.

$$p(z_i = j \mid x_i, \boldsymbol{\theta}) = \frac{p(x_i \mid z_i = j, \boldsymbol{\theta})p(z_i = j \mid \boldsymbol{\theta})}{\sum_{j'} p(x_i \mid z_i = j', \boldsymbol{\theta})p(z_i = j' \mid \boldsymbol{\theta})} = \frac{p(x_i \mid z_i = j, \boldsymbol{\theta})}{\sum_{j'} p(x_i \mid z_i = j', \boldsymbol{\theta})}$$

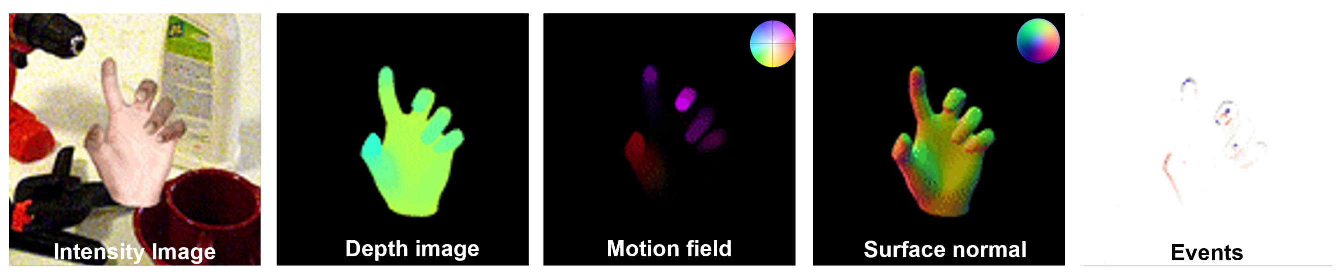
M-step: Maximize the expected contour measurement likelihood to update pose parameter.

$$p(x_i, z_i = j \mid \boldsymbol{\theta}) = p(x_i \mid z_i = j, \boldsymbol{\theta})p(z_i = j \mid \boldsymbol{\theta}) \propto p(x_i \mid z_i = j, \boldsymbol{\theta})$$

Iteration: optimize pose parameter by alternating E- and M-step.

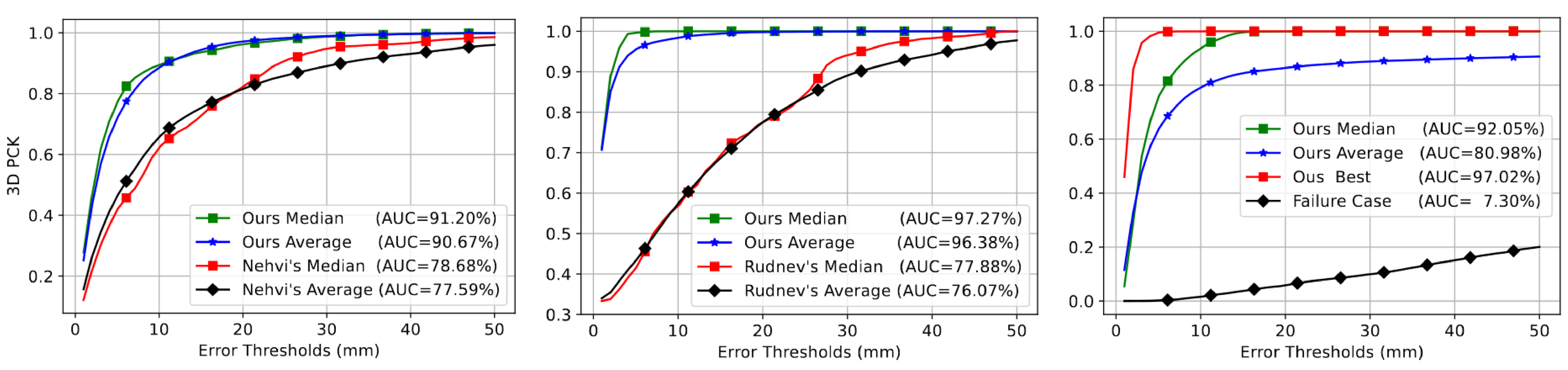
Event Stream Simulator

Inspired by state-of-the-art works, we propose our event stream simulator which supports **more data modalities** and **parametric body models**.



Quantitative Results on Synthetic Data

Scenario	Method	Mean MPJPE (mm)	Median MPJPE (mm)
MANO hand	Nehvi et al. [3]	11.61	10.85
	Ours	**4.52**	**4.27**
SMPL-X hand	Rudnev et al. [4]	11.88	10.73
	Ours	**1.11**	**0.76**
SMPL-X arm & hand	**Ours**	**15.39**	**3.93**



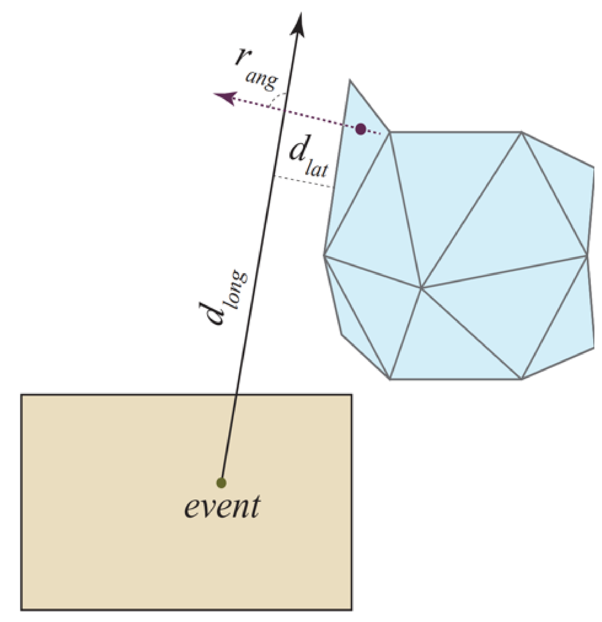
Results on MANO [6] (left), SMPL-X [2] hand (middle), and SMPL-X [2] arm & hand (right) scenario.

Conclusion & Outlook

Contribution: propose the definition of **contour events**; an **EM**-based **non-rigid reconstruction** approach from contour events; an efficient multi-modal **event stream simulator**
Limitation: ill-constrained settings, e.g. not enough contour events; not real-time capable yet.
Future Works: combine events and intensity frames to recover global rigid transformation; assign event to smaller range of mesh face and use parallel programming to increase efficiency

Measurement Likelihood

We use the measurement likelihood in E-step (to calculate the association likelihood) and M-step.



We formulate the measurement likelihood that an event x_i is caused by a mesh face f_j using

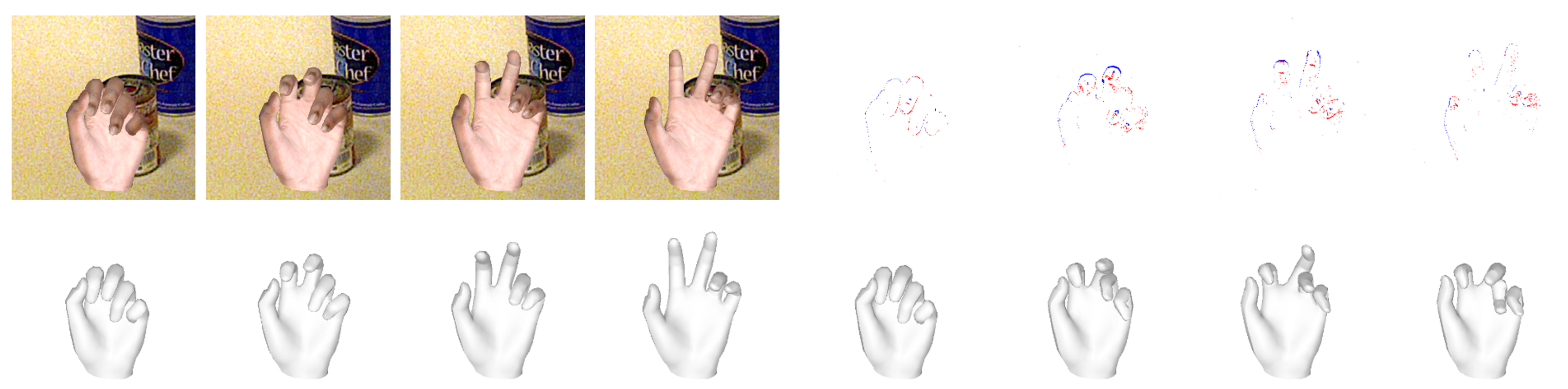
- lateral distance d_{lat}
- longitudinal distance d_{long} (not used in M-step)
- angular error r_{ang}

between the line of sight through event x_i and the mesh face f_j.

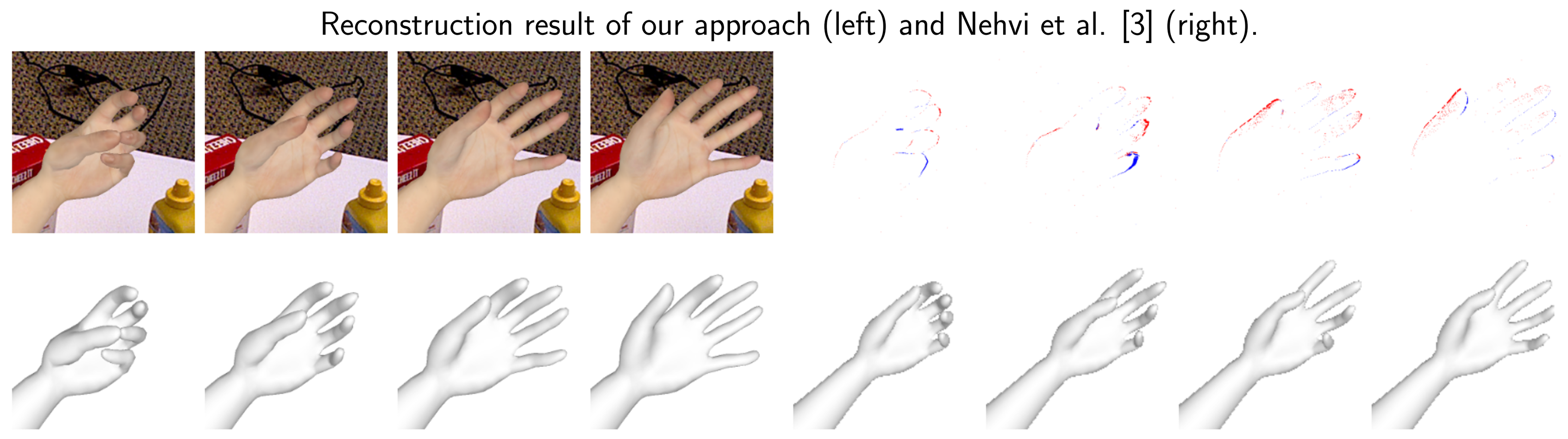
We formulate the measurement likelihood on the observed contour as

$$p(x_i \mid z_i = j, \boldsymbol{\theta}) \propto \underbrace{\sigma\left(\delta_j^i \frac{d_{\text{lat}}^2(i,j)}{\alpha}\right)}_{lateral} \underbrace{\exp\left(-\frac{d_{\text{long}}(i,j)}{\beta}\right)}_{longitudinal} \underbrace{\exp\left(-\frac{r_{\text{ang}}(i,j)}{\gamma}\right)}_{contour}$$

Results on Synthetic Data

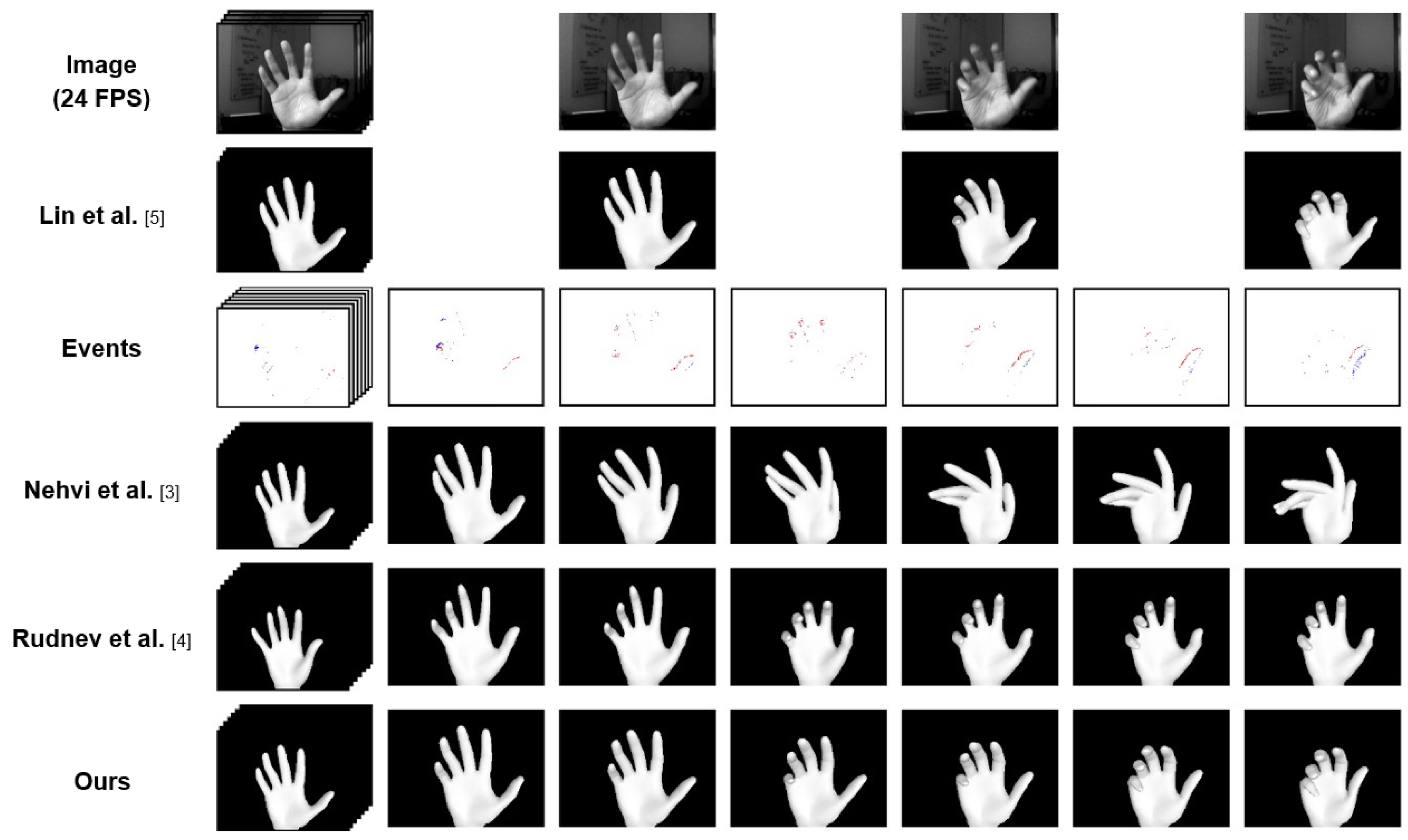


Reconstruction result of our approach (left) and Nehvi et al. [3] (right).



Reconstruction result of our approach (left) and Rudnev et al. [4] (right).

Results on Real Data



Reconstruction result on DAVIS 240C captured events.

Acknowledgements:

This work has been supported through the Max Planck Society and Cyber Valley. The authors thank the Empirical Inference Department at Max Planck Institute for Intelligent Systems in Tübingen for providing the DAVIS 240C event camera.

References:

[1] Xue, Y., Li, H., Leutenegger, S., and Stueckler, J. (2022). Event-based Non-Rigid Reconstruction from Contours. In *BMVC 2022*.

[2] G. Pavlakos et al. (2019). Expressive Body Capture: 3D Hands, Face, and Body from a Single Image. In *CVPR 2019*.

[3] J. Nehvi et al. (2021). Differentiable Event Stream Simulator for Non-Rigid 3D Tracking. In *CVPRW 2021*.

[4] V. Rudnev et al. (2021). EventHands: Real-Time Neural 3D Hand Pose Estimation from an Event Stream. In *ICCV 2021*.

[5] K. Lin et al. (2021). Mesh Graphormer. In *ICCV 2021*.

[6] J. Romero et al. (2017). Embodied Hands: Modeling and Capturing Hands and Bodies Together. In *SIGGRAPH Asia 2017*.