

3D Hand Pose Estimation from RGB Image

Eng Hock Lee
Chih-Tien Kuo

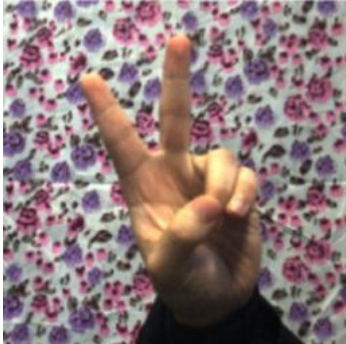


UNIVERSITY OF MINNESOTA
Driven to DiscoverSM

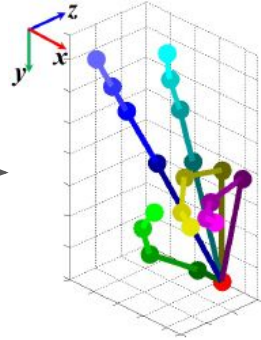
Problem Statement

Goal 1

Input RGB Image

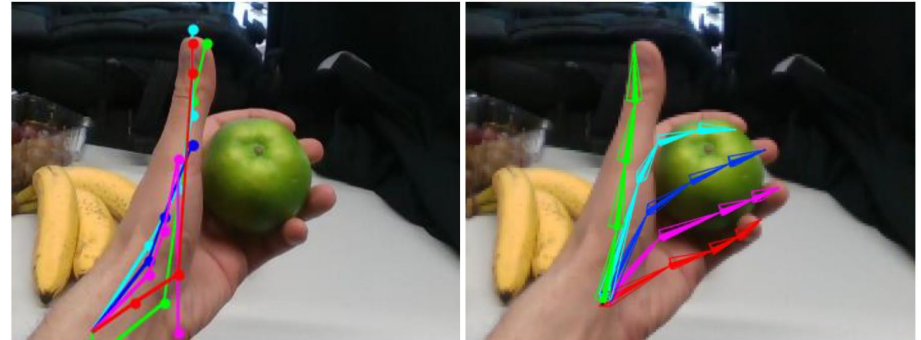


Output 3D Hand Pose



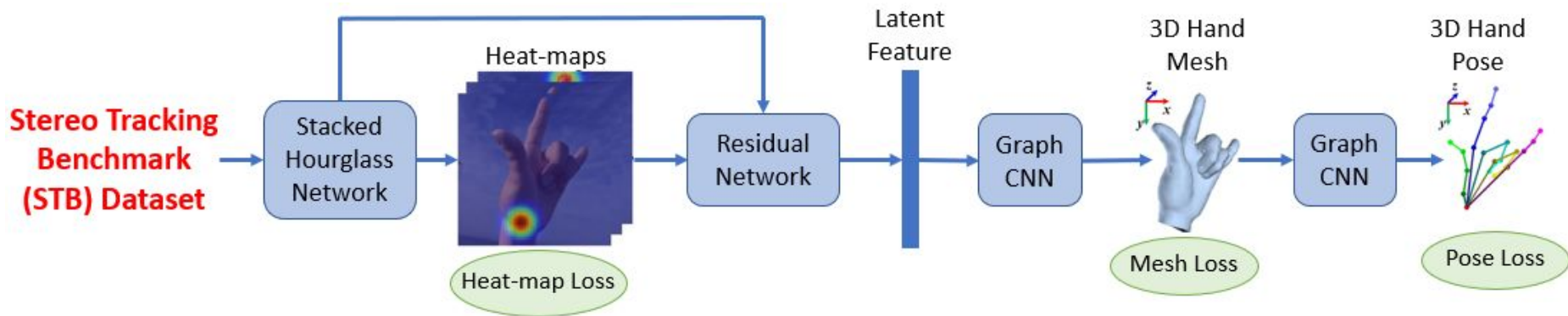
Goal 2

Overcoming Image Occlusion



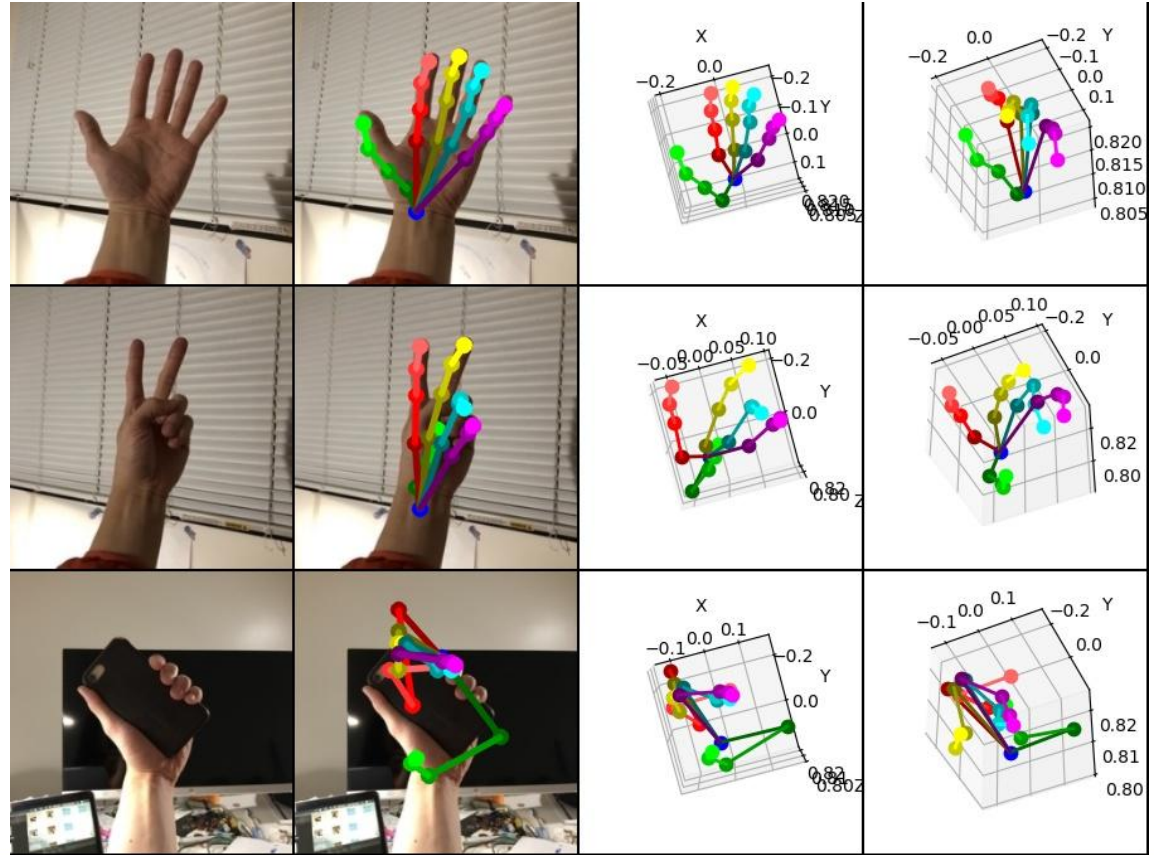
Baseline Method

- Following method from Ge, *et al.* [1]
- Using Stereo Hand Pose Benchmark (STB) Dataset for model training
- Issue:
 - STB dataset does not include images with occlusion



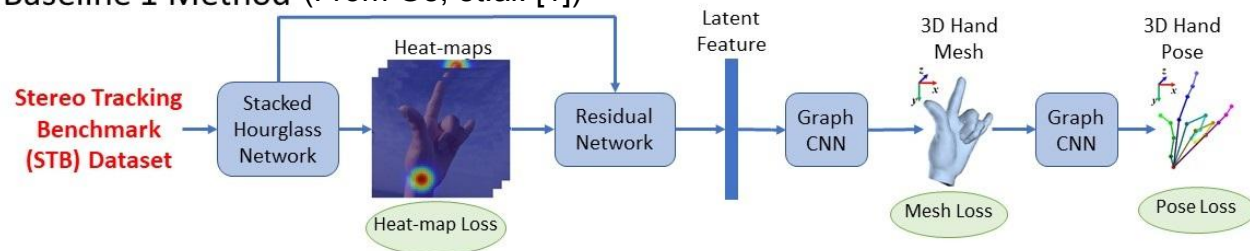
Baseline Method

Issue of occlusion

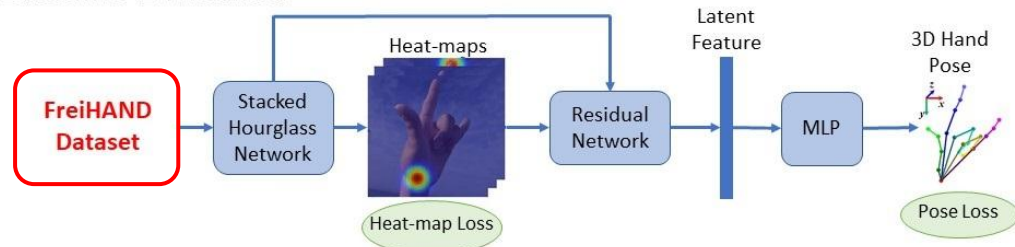


Baseline and Proposed Methods

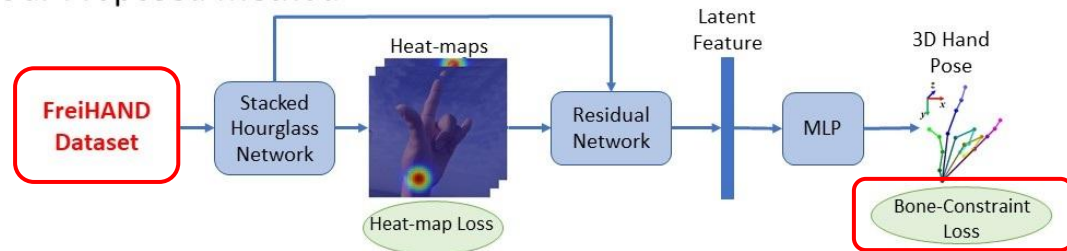
a) Baseline 1 Method (From Ge, et.al. [1])



b) Baseline 2 Method



c) Our Proposed Method



Proposed Method

- Original Loss

$$L_{pose} = \sum_{j=1}^{21} \left\| \phi_j - \hat{\phi}_j \right\|_2$$

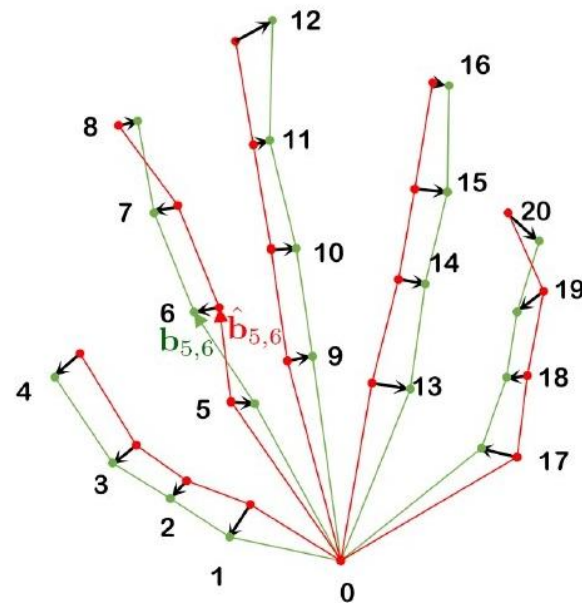
- Bone-Constraint Loss

$$L_{len} = \sum_{i,j} \left| \left\| \mathbf{b}_{i,j} \right\|_2 - \left\| \hat{\mathbf{b}}_{i,j} \right\|_2 \right|$$

$$L_{dir} = \sum_{i,j} \left\| \frac{\mathbf{b}_{i,j}}{\left\| \mathbf{b}_{i,j} \right\|_2} - \frac{\hat{\mathbf{b}}_{i,j}}{\left\| \hat{\mathbf{b}}_{i,j} \right\|_2} \right\|$$

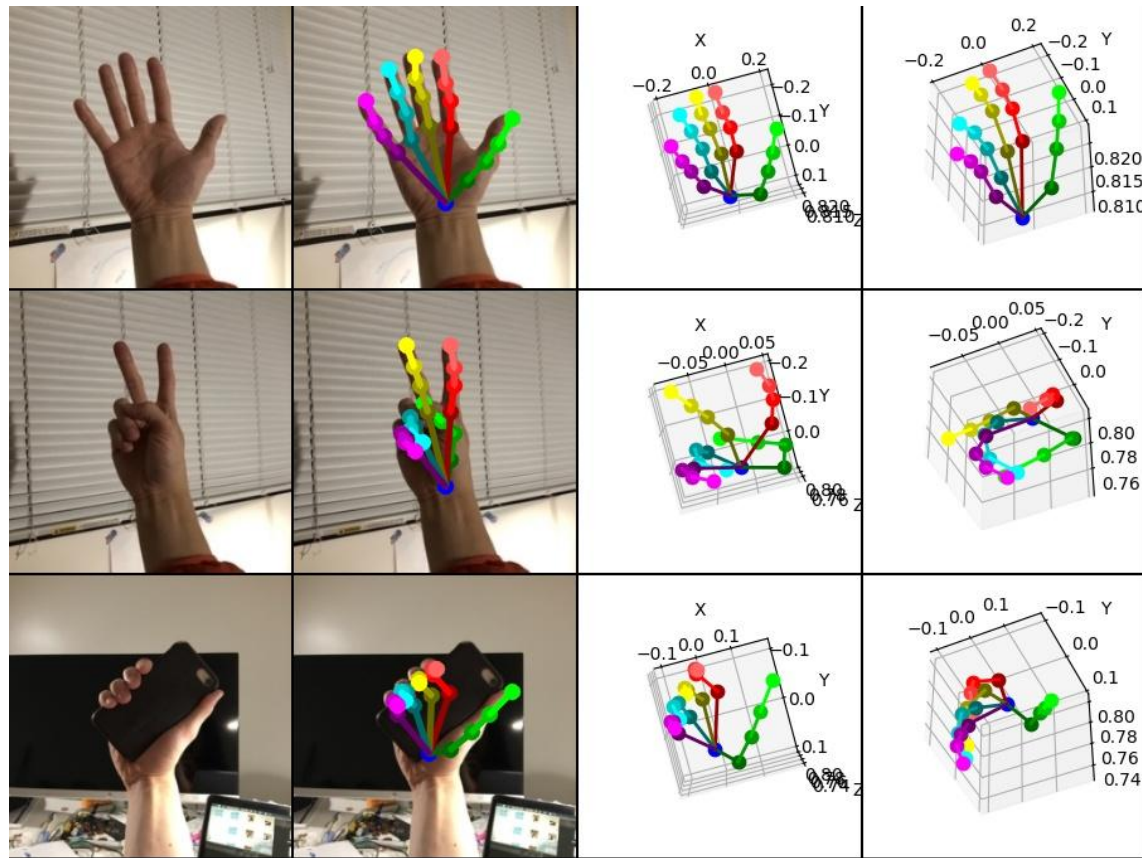
$$L = \lambda_{pose} L_{pose} + \lambda_{len} L_{len} + \lambda_{dir} L_{dir}$$

— Ground truth
— Estimates



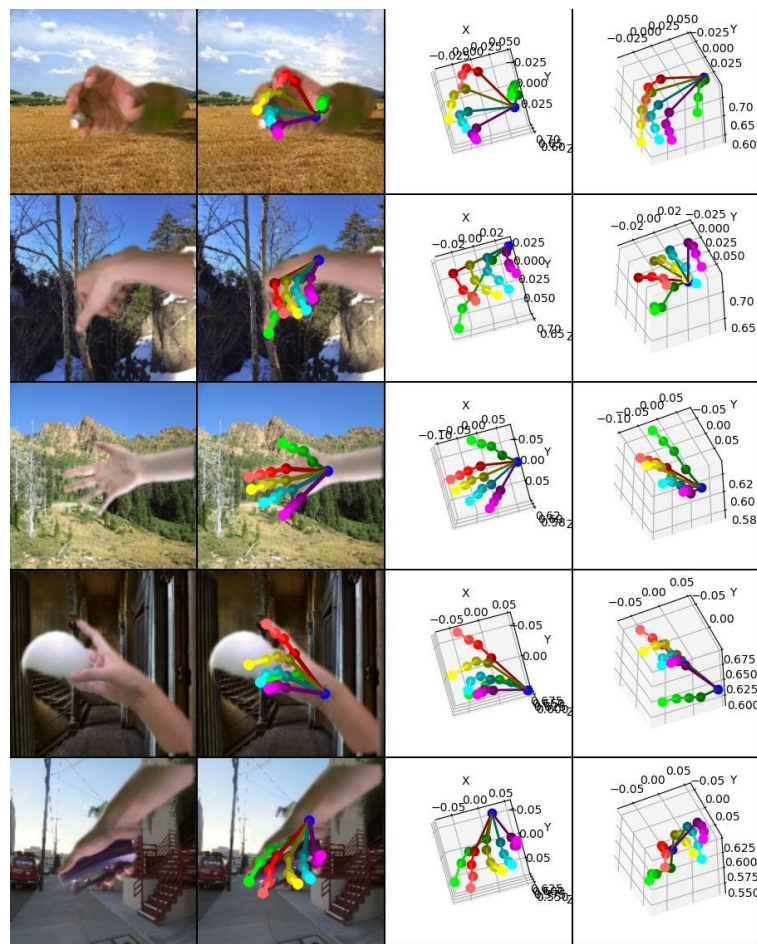
Qualitative Result

Pose reconstructed
with occlusion



UNIVERSITY OF MINNESOTA
Driven to DiscoverSM

Qualitative Result

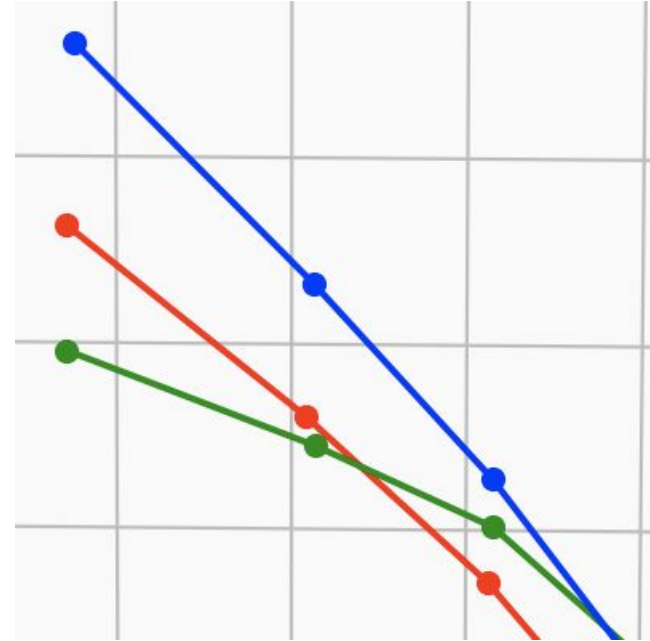
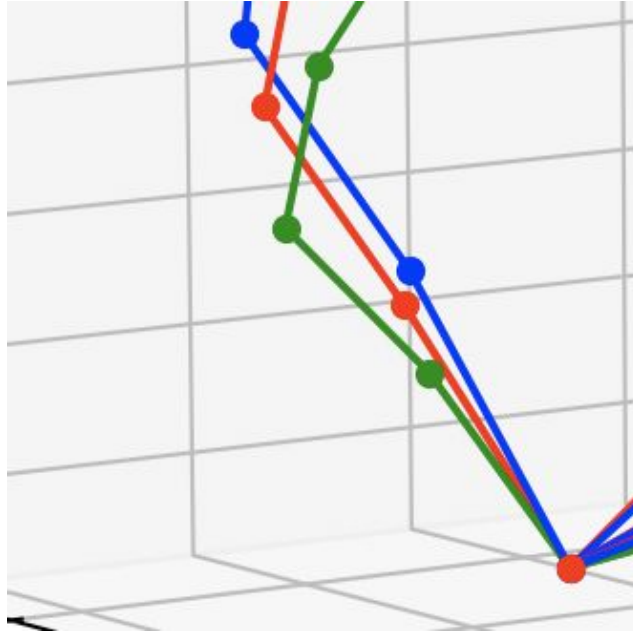


Examples with FreiHAND dataset



Qualitative Result

- Ground truth
- Baseline 2
- Proposed method



Quantitative Result

- Evaluated on 2000 testing images

3D Estimation Error		
Baseline 1	Baseline 2	Proposed Method
94.10 mm	22.84 mm	15.36 mm

- 3D PCK Plot

