MHEntropy: Entropy Meets Multiple Hypotheses for Pose & Shape Recovery

Glory Rongyu CHEN*, Linlin YANG*, & Angela YAO

Computer Vision & Machine Learning Group, School of Computing, National University of Singapore
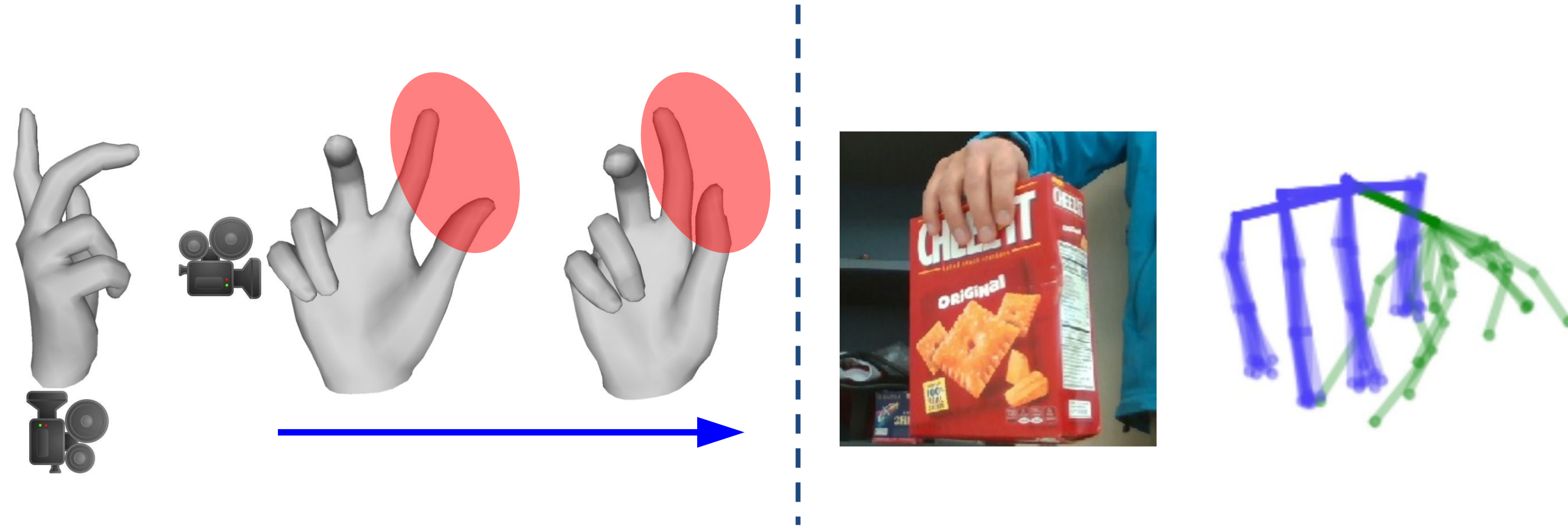
Multi-Hypothesis Pose & Shape Recovery

Problem Setting

- Input single-view RGB → output Multiple Hypotheses of human SMPL/hand MANO params.
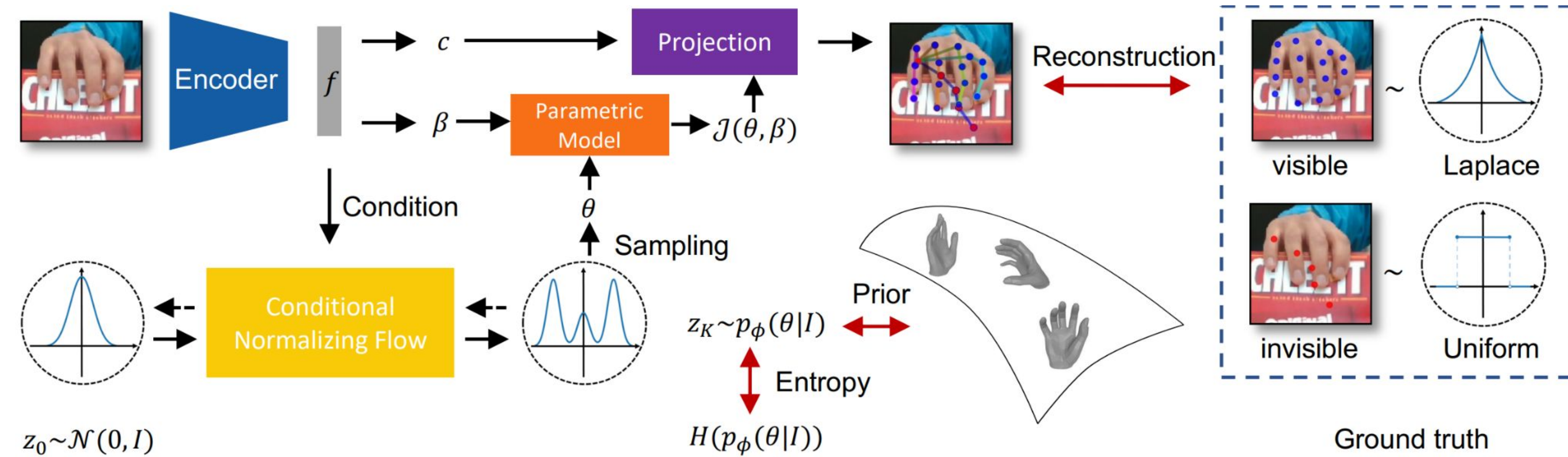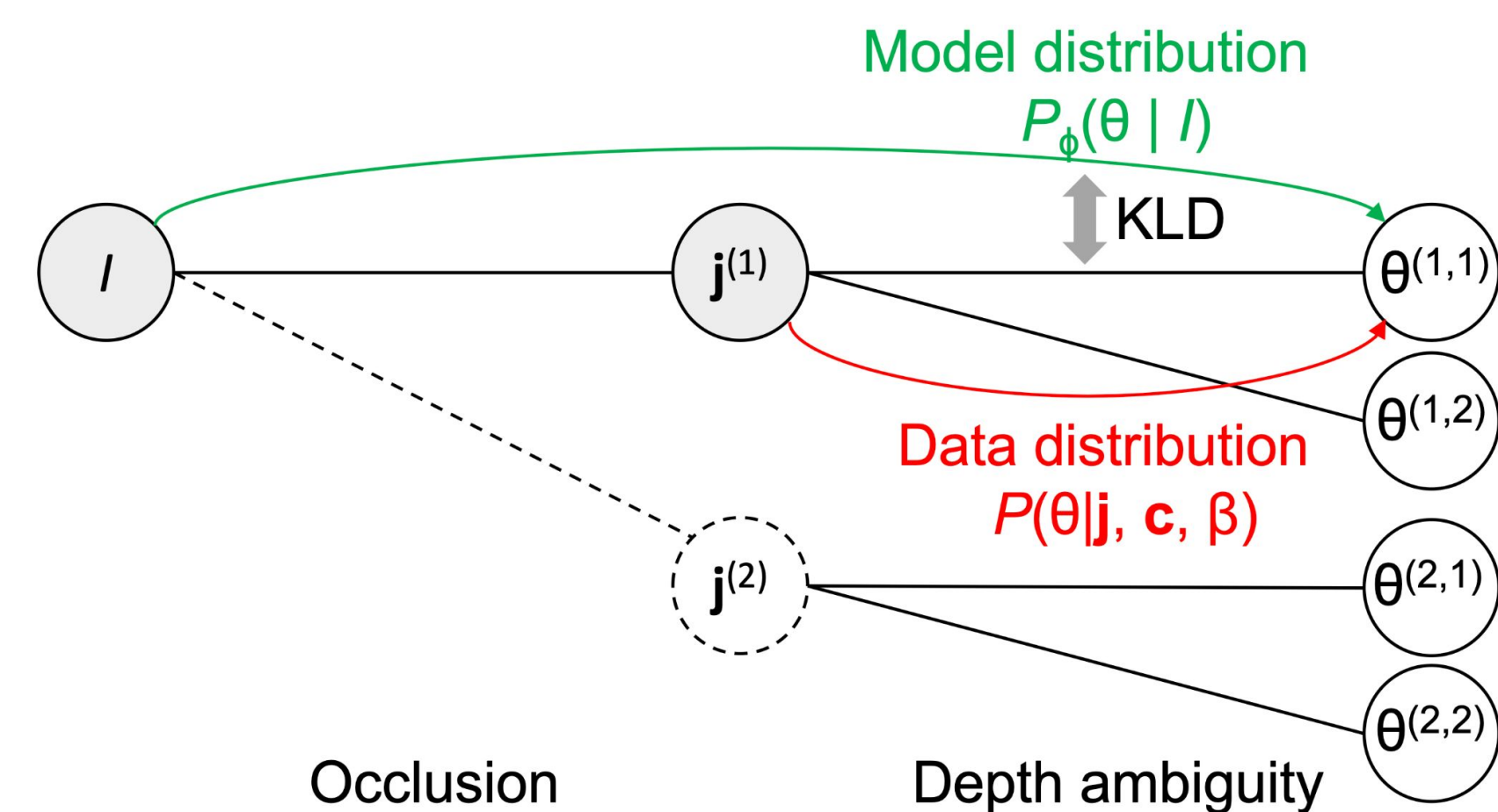
Challenges & Key Ideas

- [Annotations] Promote partial weak supervision (visible 2D).
- [Learning] Use knowledge to mitigate no 1-to-many data pair.
- [Evaluation] Gold criteria: *accurate on certain* (2D vis) & *diverse only on uncertain* (depth & invis).

Probabilistic Multi-Hypothesis Entropy

- No need Θ: sample the Model Distribution $P_{nf}(\Theta \mid I)$ & check…
- Knowledge: 1. feasible pose [Prior] $P(\Theta)$. 2. 2D [Reconstruction].
- Gold criteria $P(\mathbf{J} \mid I)$ = vis Laplace, invis Uniform.
- Use vis 2D & knowledge to define the Data Distribution
  $P_{data}(\Theta \mid I) \sim P(\mathbf{J} \mid \Theta) P(\Theta) P(\mathbf{J} \mid I)$.

Model distribution $P_\phi(\theta \mid I)$

KLD

Data distribution $P(\theta \mid \mathbf{j}, \mathbf{c}, \beta)$

Occlusion   Depth ambiguity

Reconstruction

visible ~ Laplace

invisible ~ Uniform

Ground truth

Project

CVML@NUS

$z_0 \sim \mathcal{N}(0, I)$

$z_K \sim p_\phi(\theta \mid I)$

Prior

Entropy

$H(p_\phi(\theta \mid I))$

a missing term

$$KL(p_\phi(\theta|I,\mathbf{j},\mathbf{c},\beta)\|p(\theta|I,\mathbf{j},\mathbf{c},\beta)) = -\left(\underbrace{\mathop{E}_{p_\phi(\theta|I)}[\log p(\mathbf{j}|\mathbf{c},\beta,\theta)]}_{\text{reconstruction}} + \underbrace{\mathop{E}_{p_\phi(\theta|I)}[\log p(\theta)]}_{\text{prior}} + \underbrace{H(p_\phi(\theta|I))}_{\text{entropy}} \right)$$

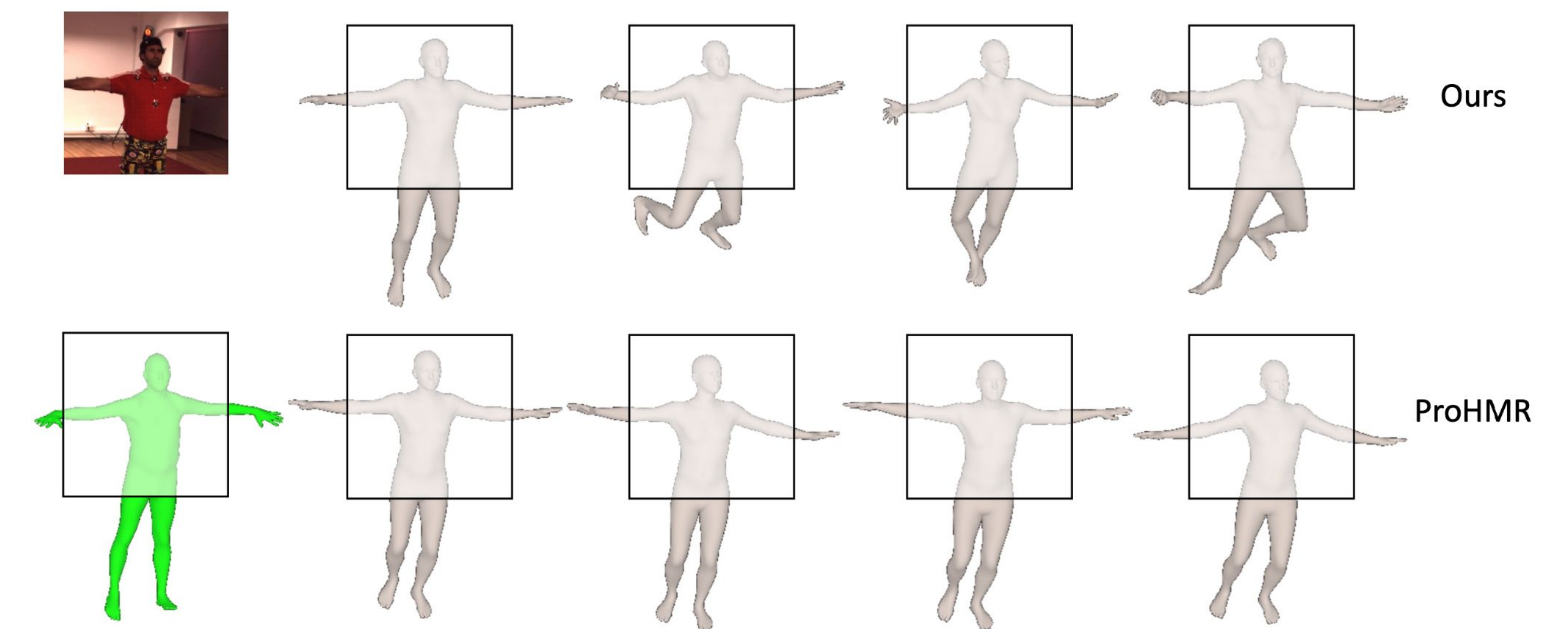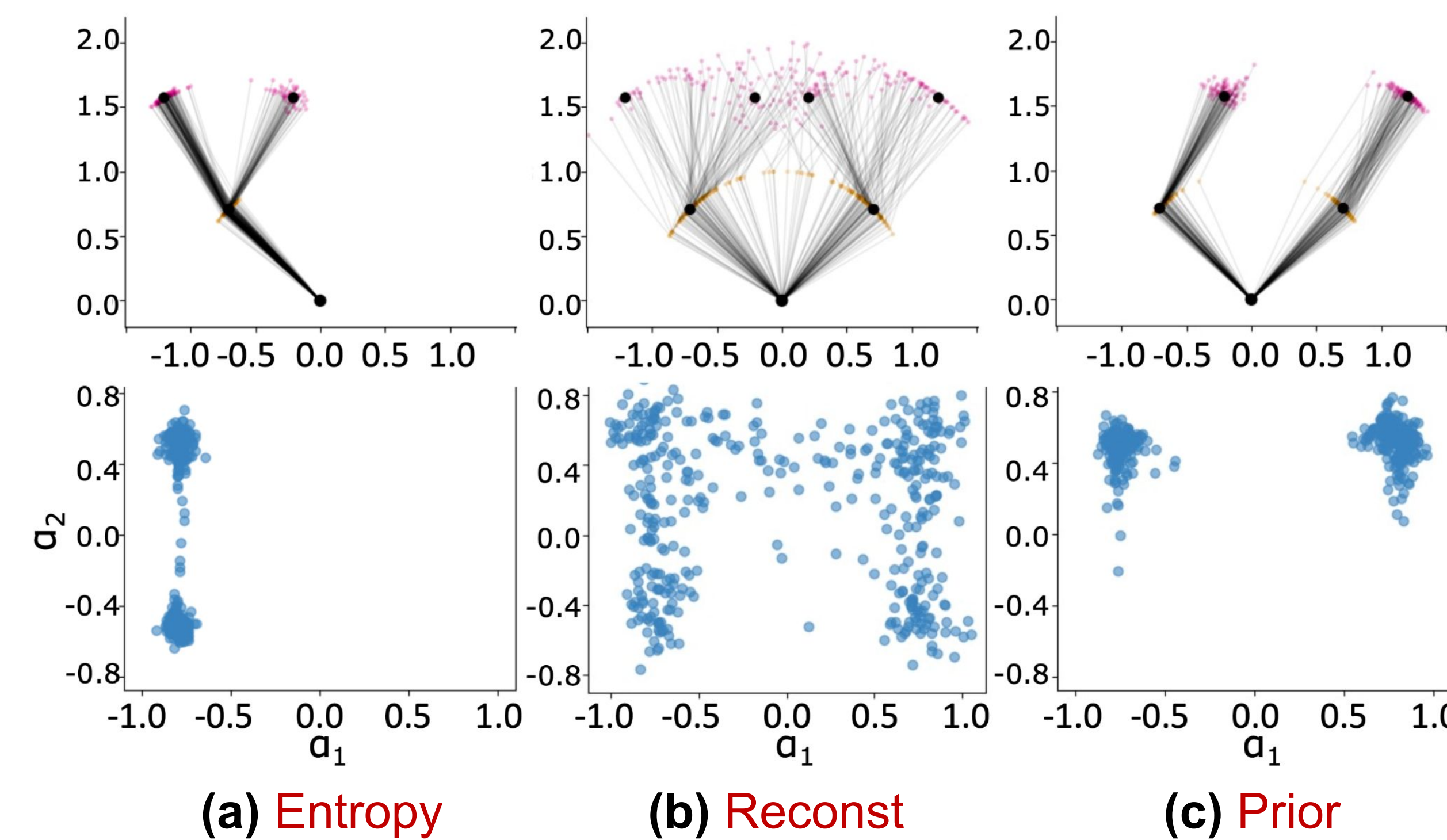## Accurate & Diverse Results & Discussions

(a) Entropy   (b) Reconst   (c) Prior

Supervision		MH	H36M	AH36M
2D Vis	HMR		67.4	85.2
	ProHMR	✓	64.3	82.6
	Ours	✓	**51.3**	**66.4**
3D	HMR		56.8	-
	SPIN		41.1	-
	MDN	✓	42.7	69.5
	CVAE	✓	46.2	75.1
	Multi-bodies	✓	42.2	64.2
	ProHMR	✓	**36.8**	60.1
	Ours	✓	**36.8**	50.6

	Acc	Div		Div|Acc
	AH (pix)↓	PJD		RD↓
		2D Vis	3D Occ	
ProHMR	10.92	0.06	0.26	0.23
Ours	**9.75**	4.56	64.05	**0.07**
ProHMR	13.38	3.98	24.27	0.16
Ours	**10.73**	4.23	47.95	**0.09**

We comprehensively evaluate Diversity & Accuracy of MH.

Ours

ProHMR

Downstream Tasks: Hand-Object Interaction post-selection & multi-view fitting...

Extensions: incorporate more consistency (*e.g.*, masks) & ambiguities (*e.g.*, blur) for meaningful diversity.