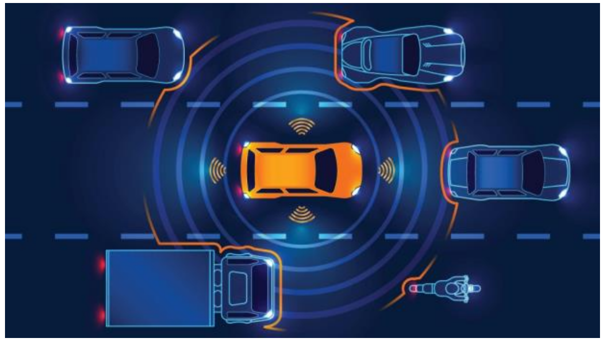
Utilizing Prior Knowledge for Non-Rigid Shape Completion

Omer Ben-Hayun

Supervisor: Ido Imanuel

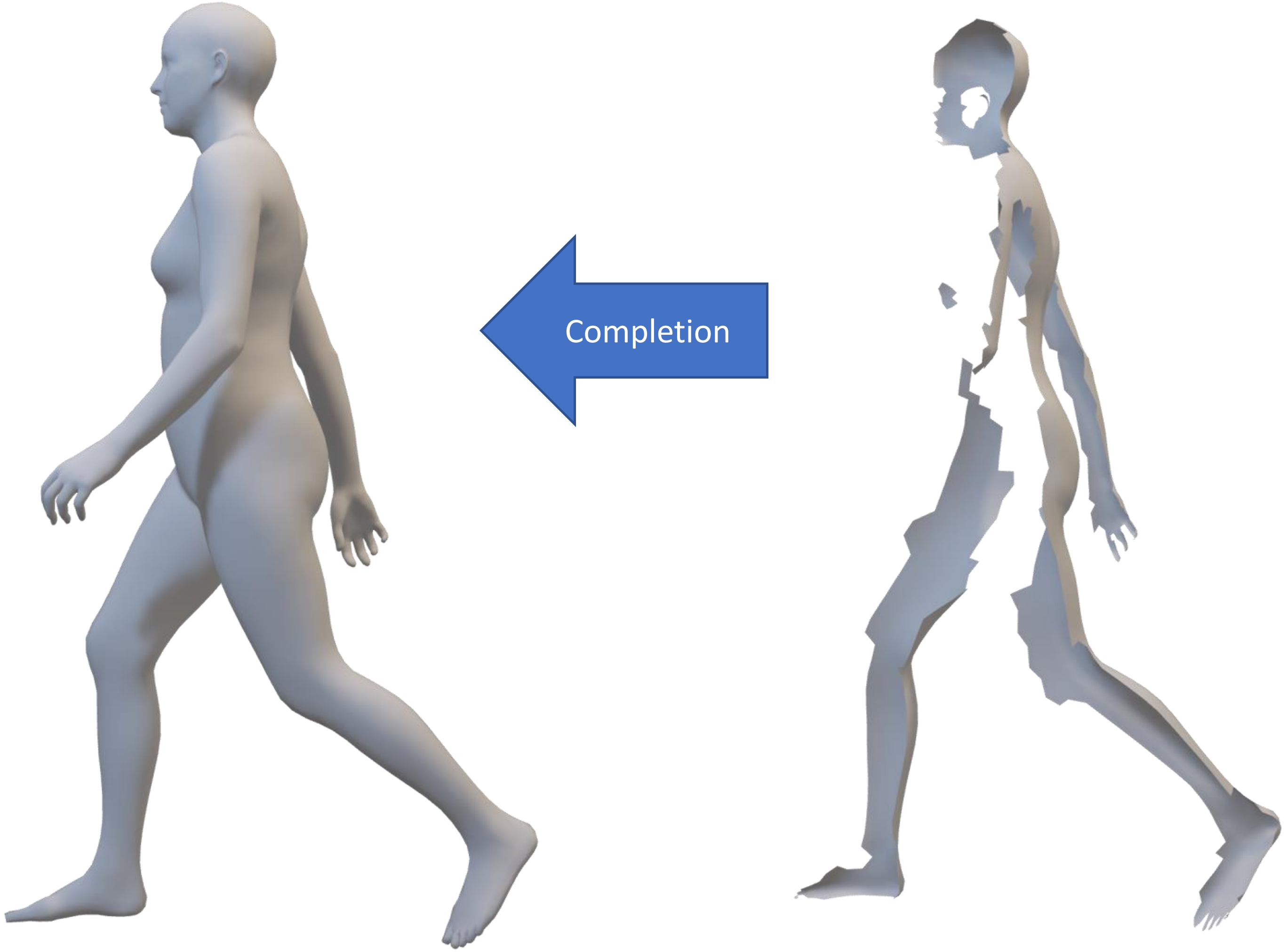
Shape completion applications



Images from left to right:
(1) https://topflightapps.com/ideas/computer-vision-in-medicine/,
(2) Andrey Suslov / Getty Images,
(3) **Gaming in the metaverse by meta https://www.youtube.com/watch?v=5FwztKGQmd8,**
(4) https://www.phonemore.com/news/how-to-3d-scan-an-object-with-your-phone/4291.

Scans source: Amass dataset. https://amass.is.tue.mpg.de

Goal:
shape completion of human shapes



Scans source: D-faust dataset. https://dfaust.is.tue.mpg.de/

Previous methods

Partial scan

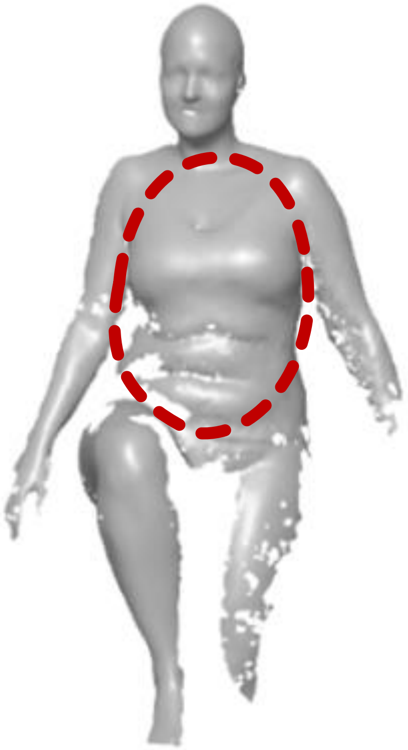
Reconstructions

SAL (2020) **AtlasNet (2018)** **SAL SignReg (2020)**

SAL: Sign Agnostic Learning of Shapes from Raw Data (2020), Atzmon et al
AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation (2018), Groueix et al

Previous methods

Partial scan



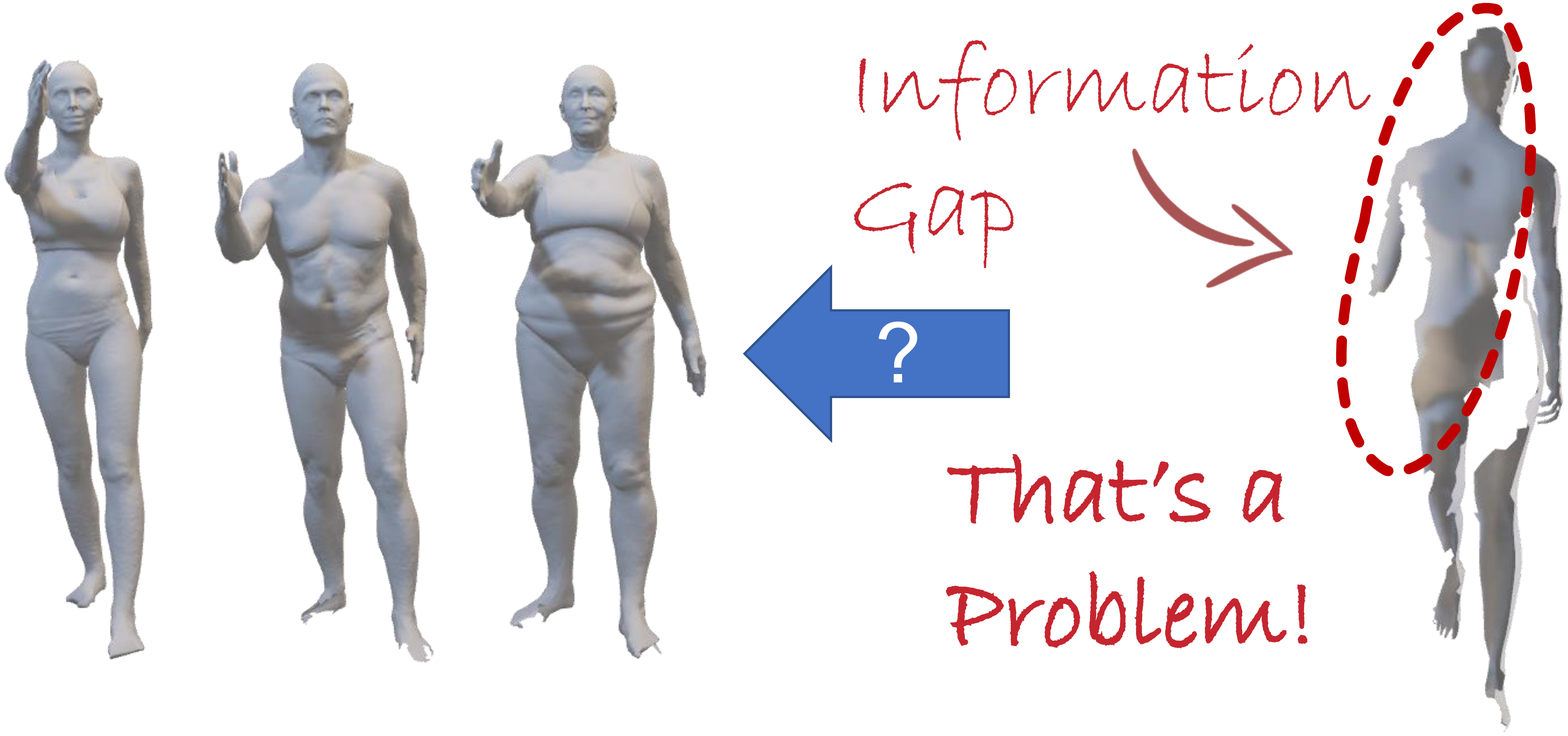
Reconstructions

Marin et al (2018) **Litany et al (2017)**

Deformable Shape Completion with Graph Convolutional Autoencoders (2017)
FARM: Functional Automatic Registration Method for 3D Human Bodies (2018)

Scans source: (Right) Amass dataset https://amass.is.tue.mpg.de
(Left) MPI-Faust dataset http://faust.is.tue.mpg.de

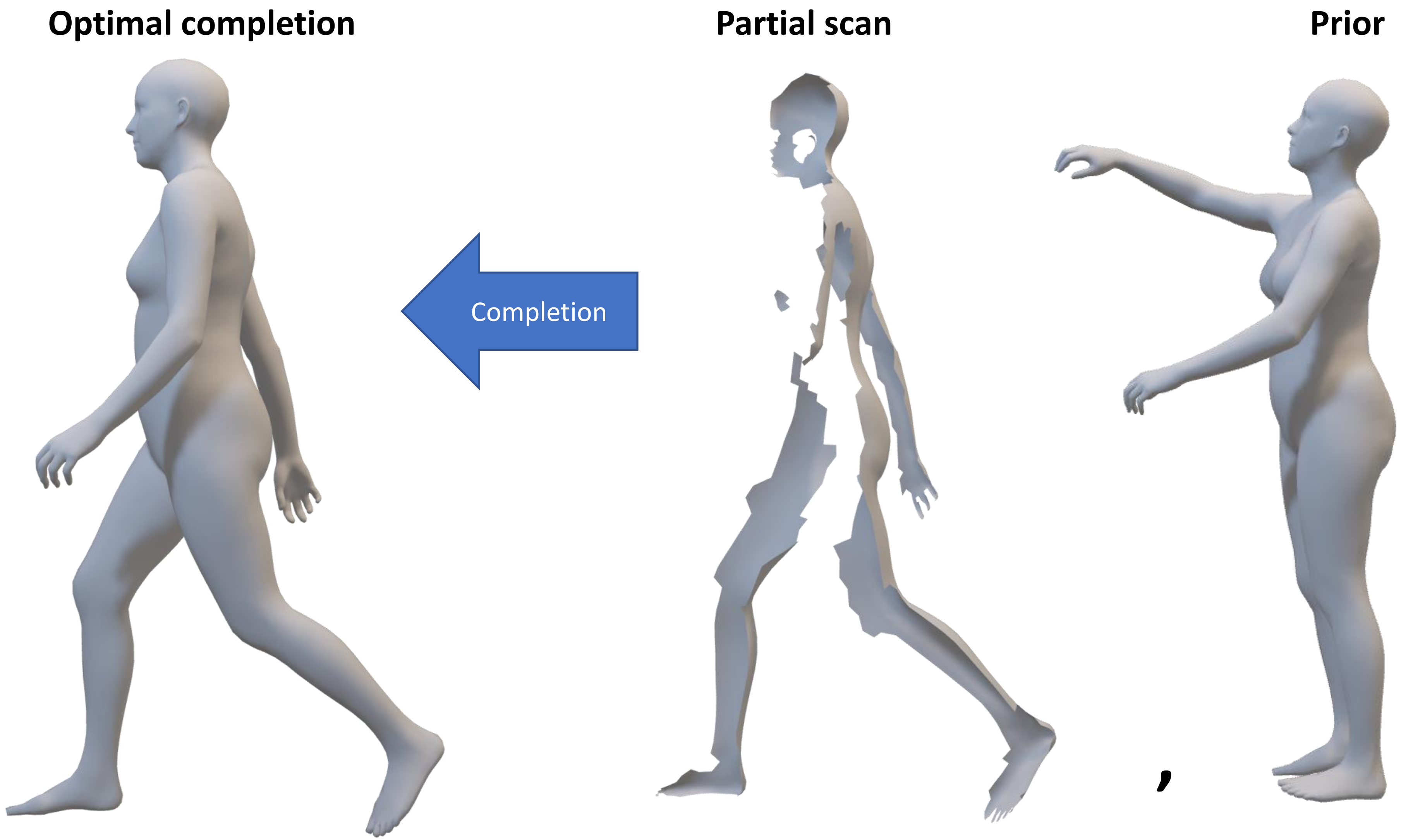
Challenge:



Scans source: Amass dataset. https://amass.is.tue.mpg.de

Previous method: **Towards Precise**

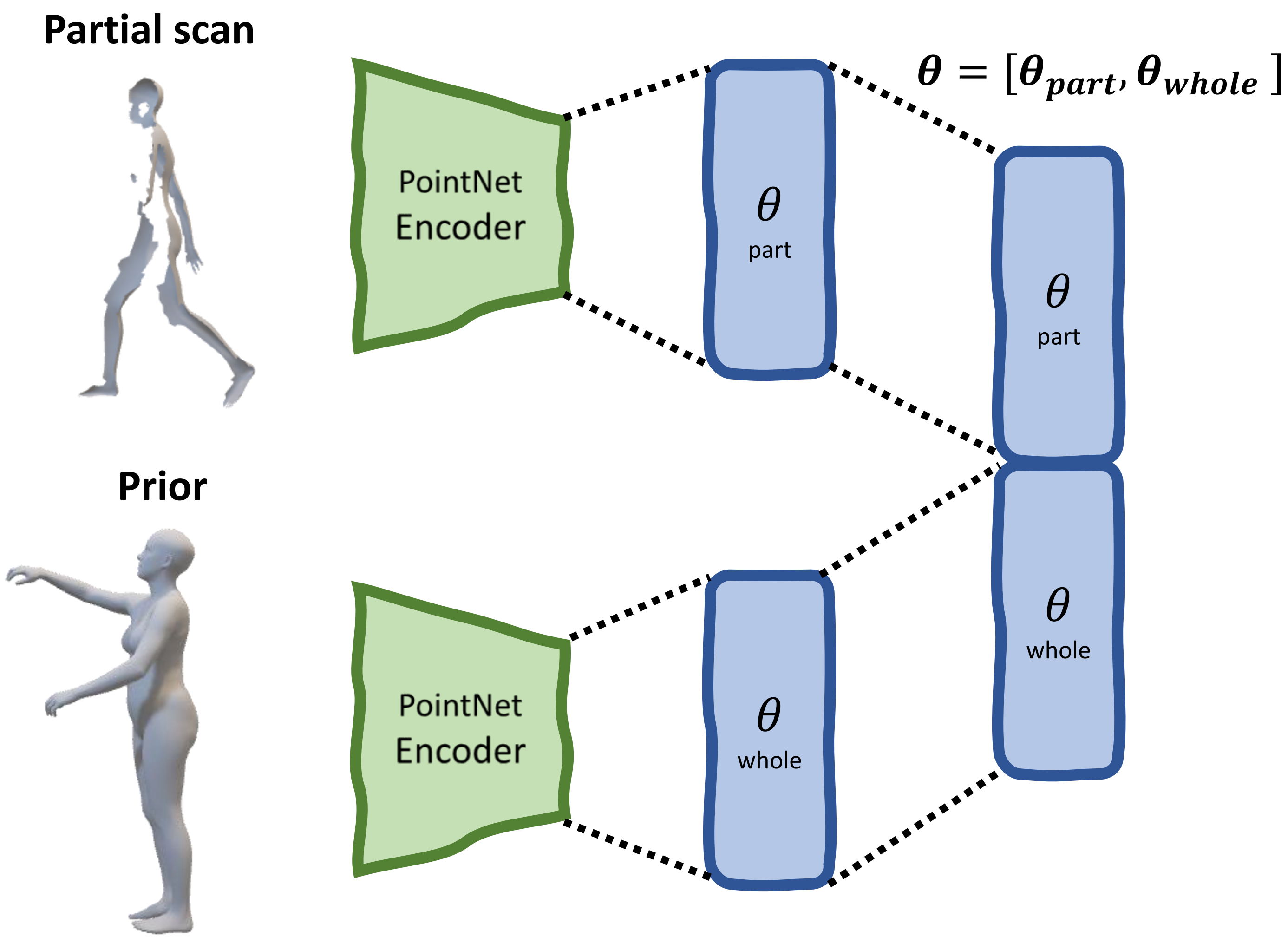
Bridging the information gap



Scans source: Amass dataset. https://amass.is.tue.mpg.de

Previous method: **Towards Precise**

Completion idea - encoder

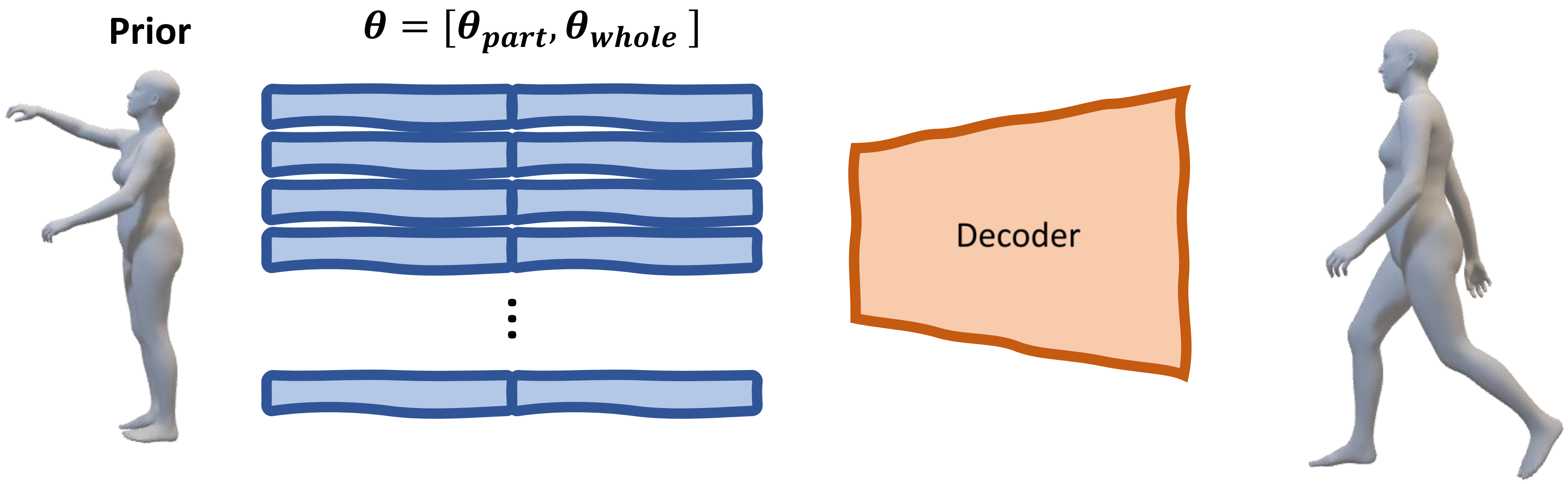


Towards precise completion of deformable shapes (2020)

Scans source: Amass dataset. https://amass.is.tue.mpg.de

Previous method: **Towards Precise**

Completion idea - decoder

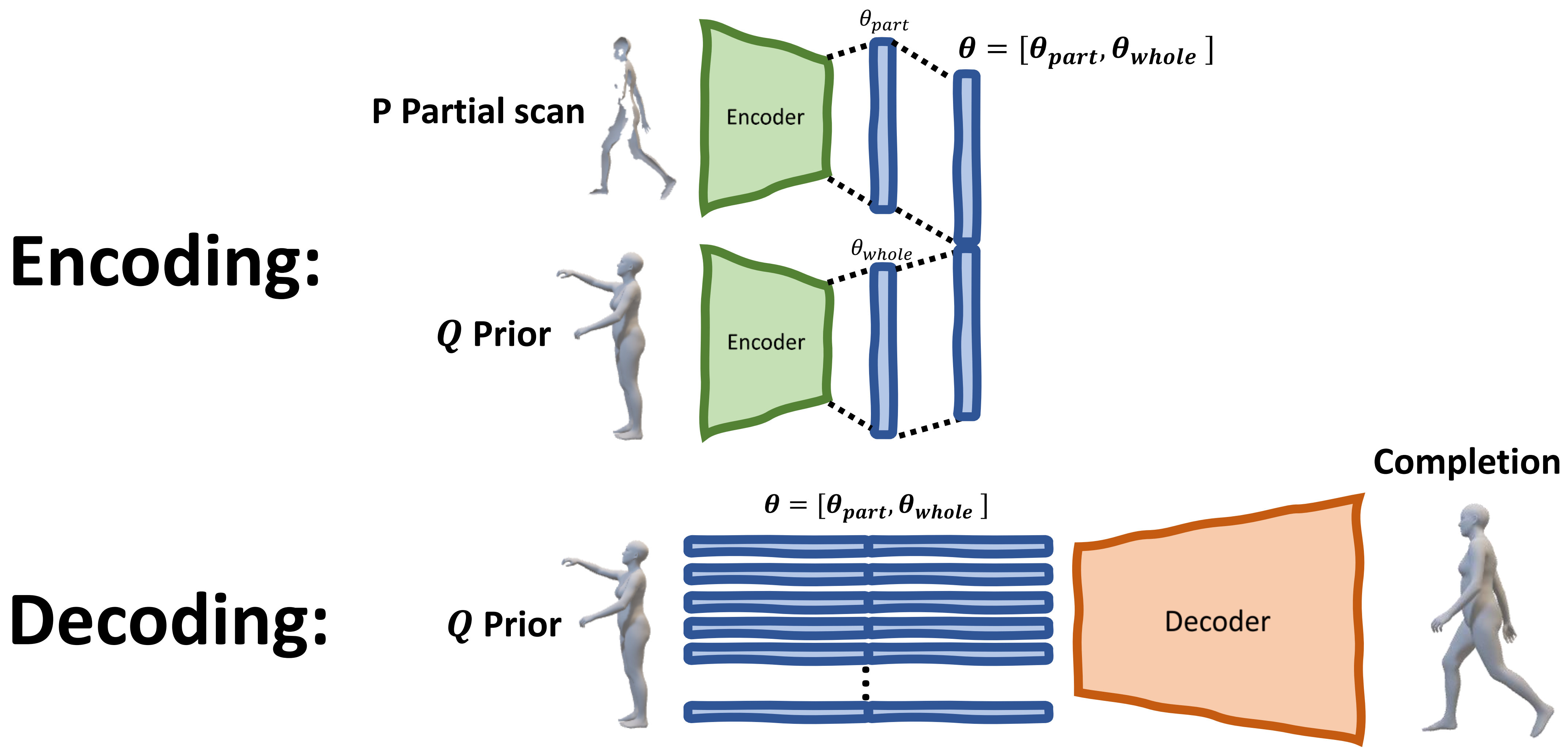


Towards precise completion of deformable shapes (2020)

Scans source: Amass dataset. https://amass.is.tue.mpg.de

Previous method: **Towards Precise**

Completion idea – full picture examples



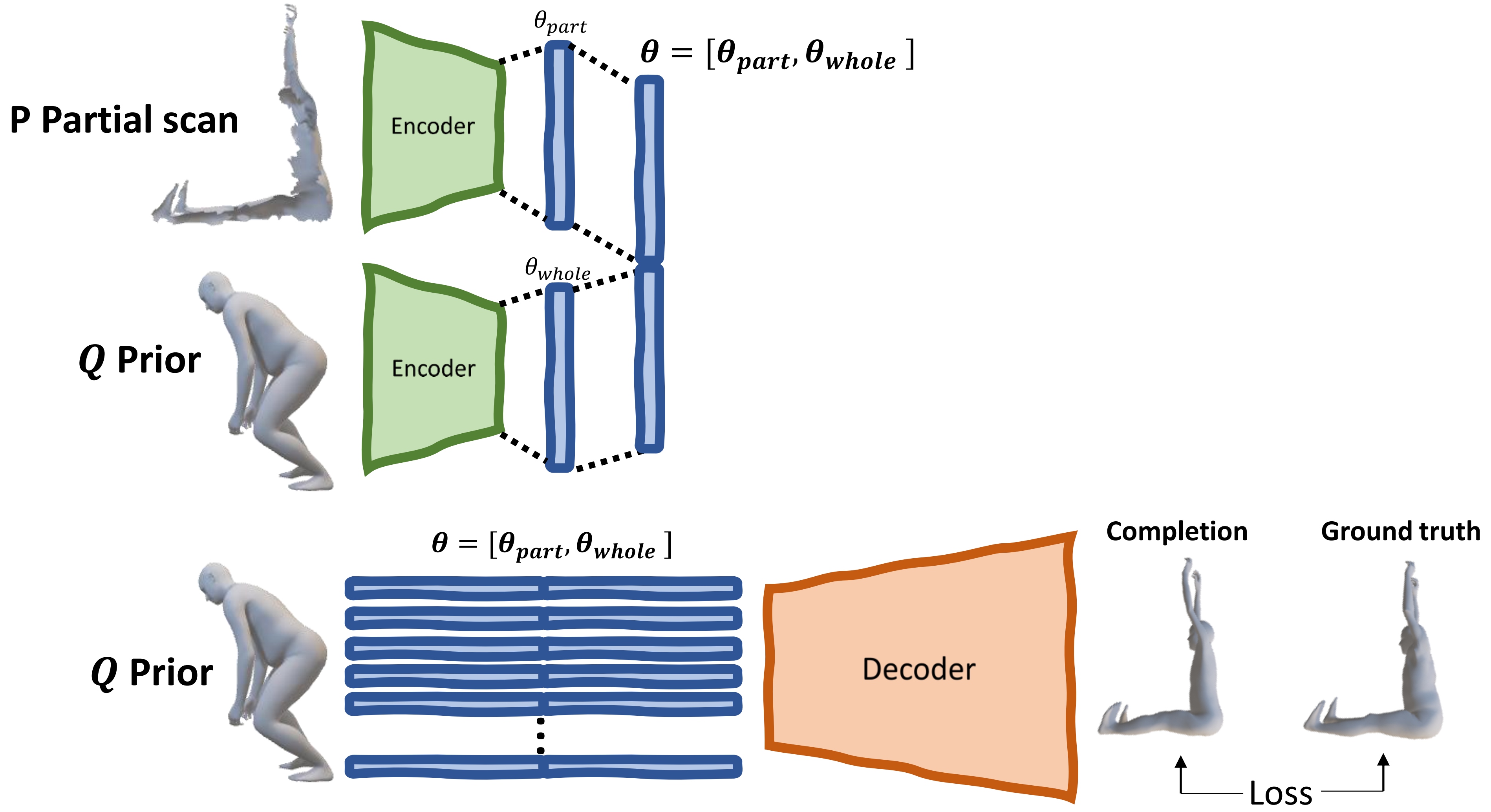
Scans source: Amass dataset. https://amass.is.tue.mpg.de

Previous method: **Towards Precise**

Completion idea – example 1

Encoding:

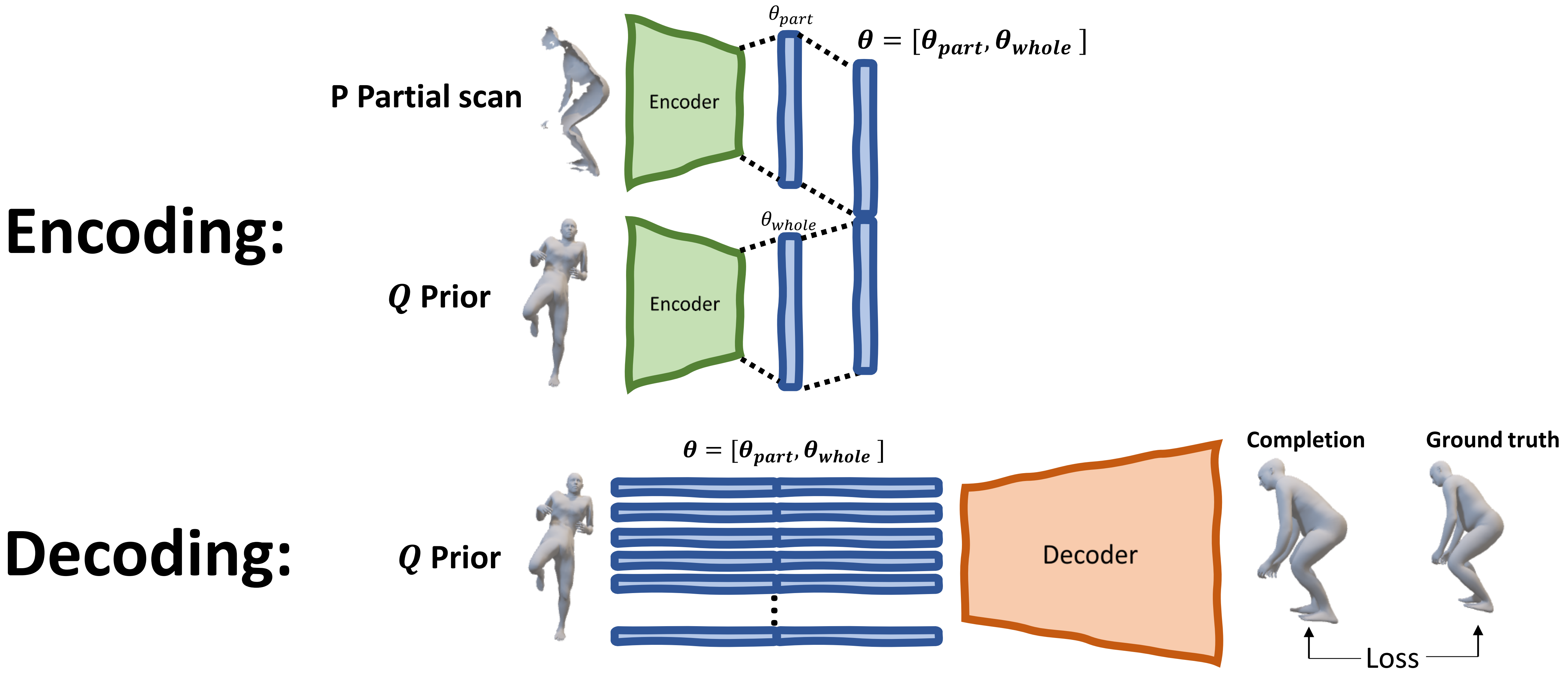
Decoding:



Scans source: Amass dataset. https://amass.is.tue.mpg.de

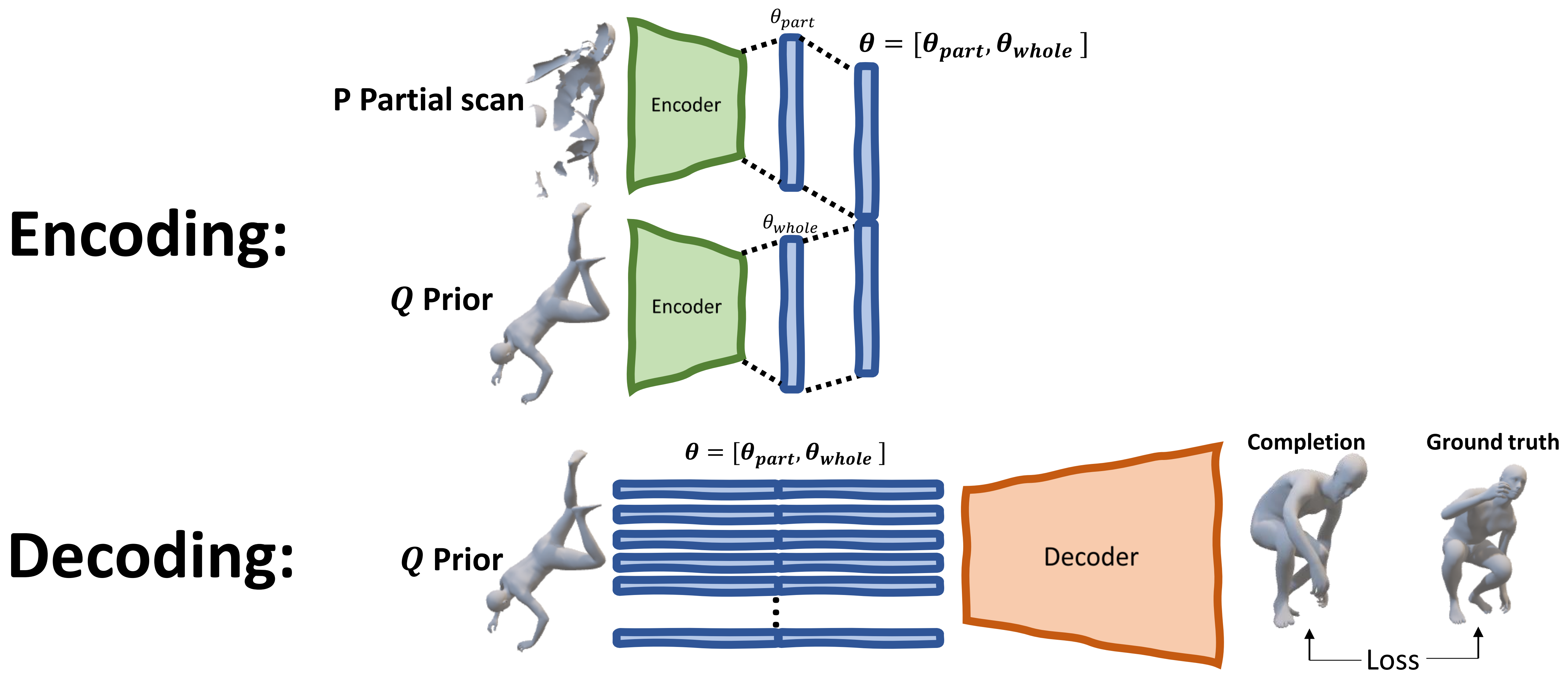
Previous method: **Towards Precise**

Completion idea – example 2



Previous method: **Towards Precise**

Completion idea – example 3

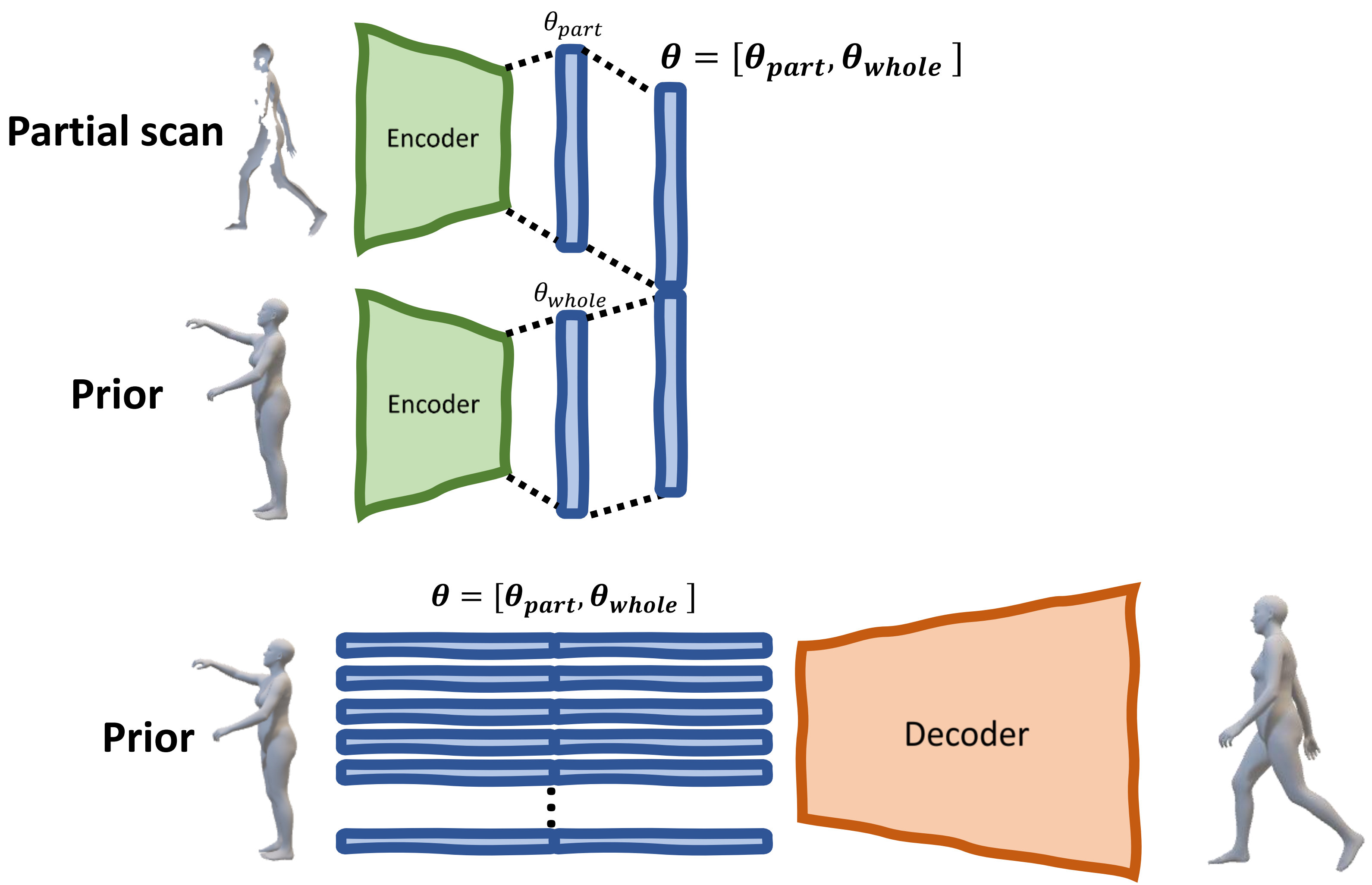


our method: FTS (fixed template shape)

Completion idea

Encoding:

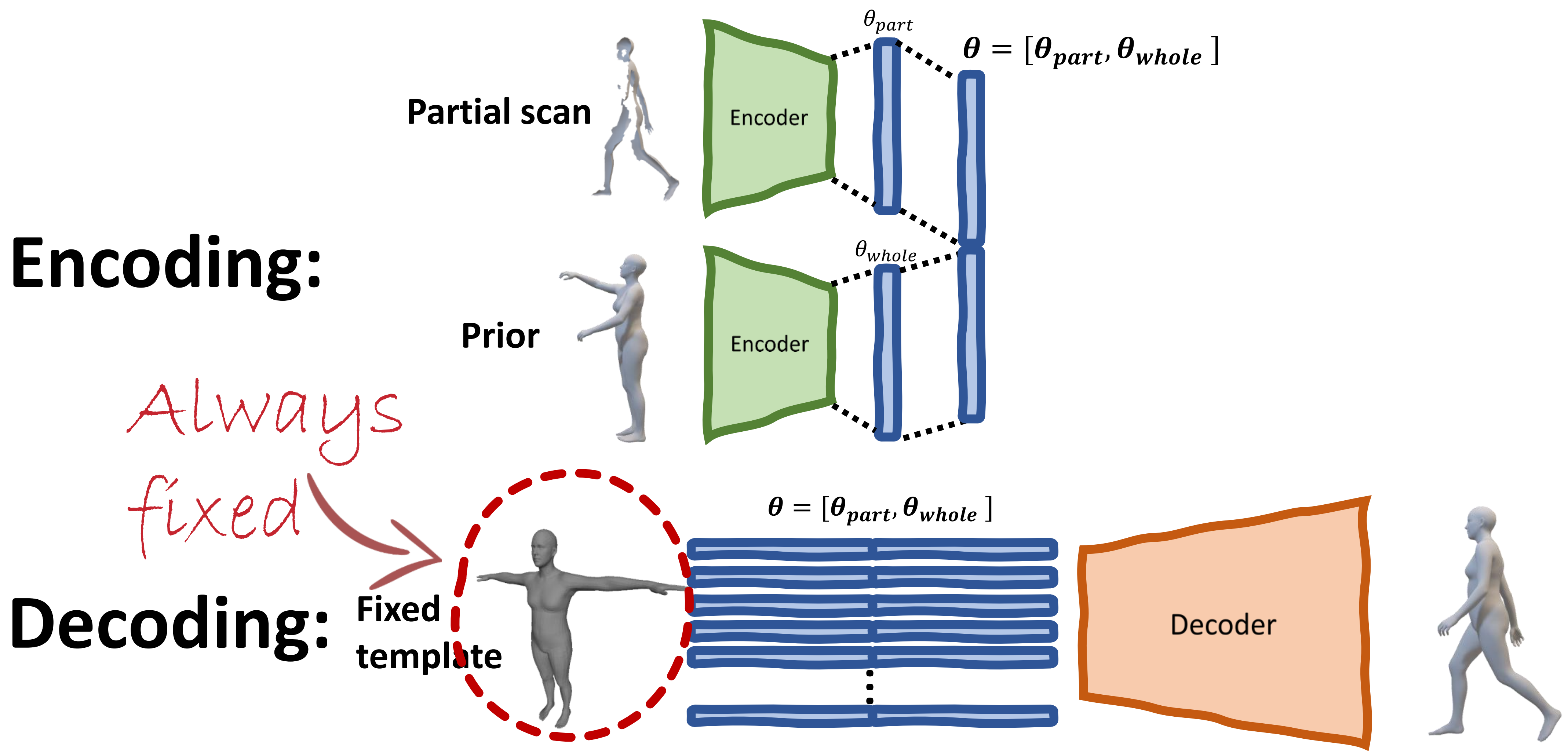
Decoding:



Scans source: Amass dataset. https://amass.is.tue.mpg.de

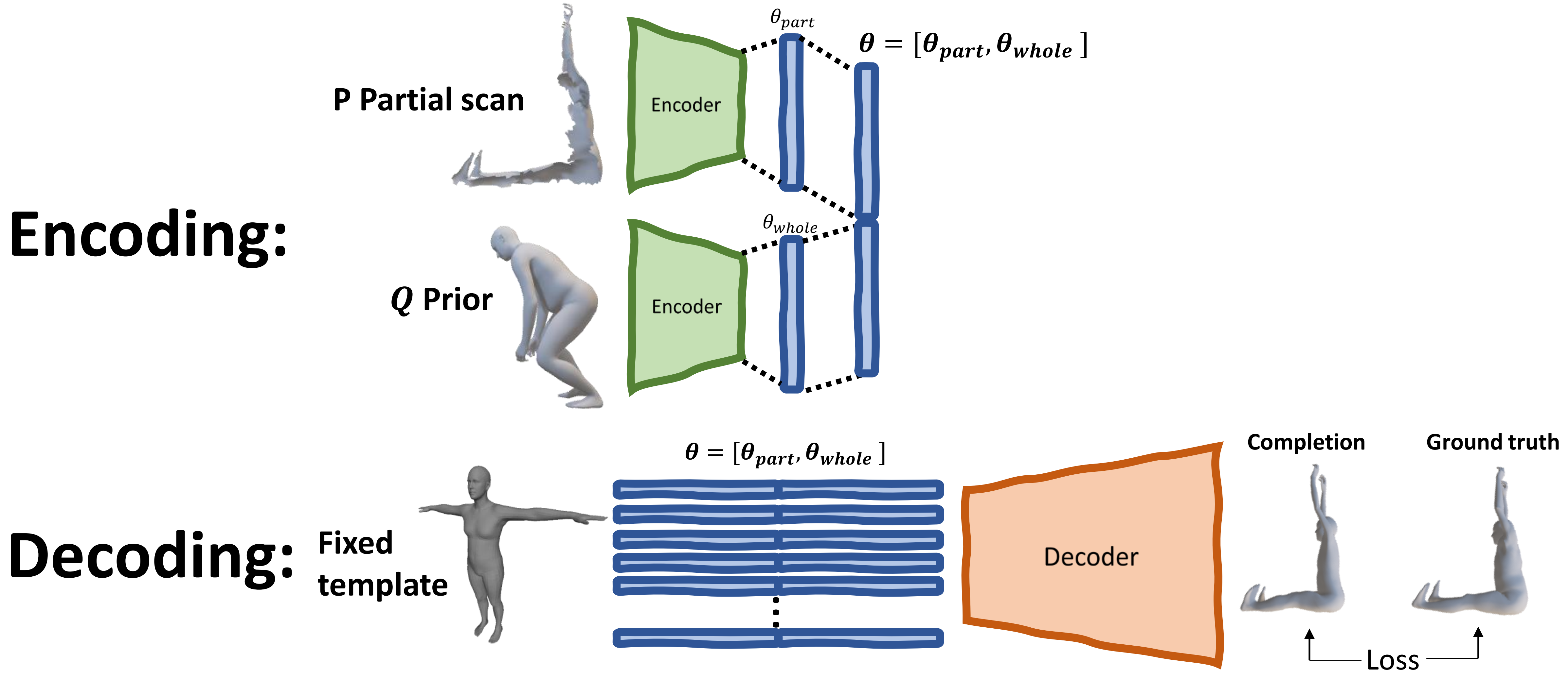
our method: FTS (fixed template shape)

Completion idea



our method: FTS (fixed template shape)

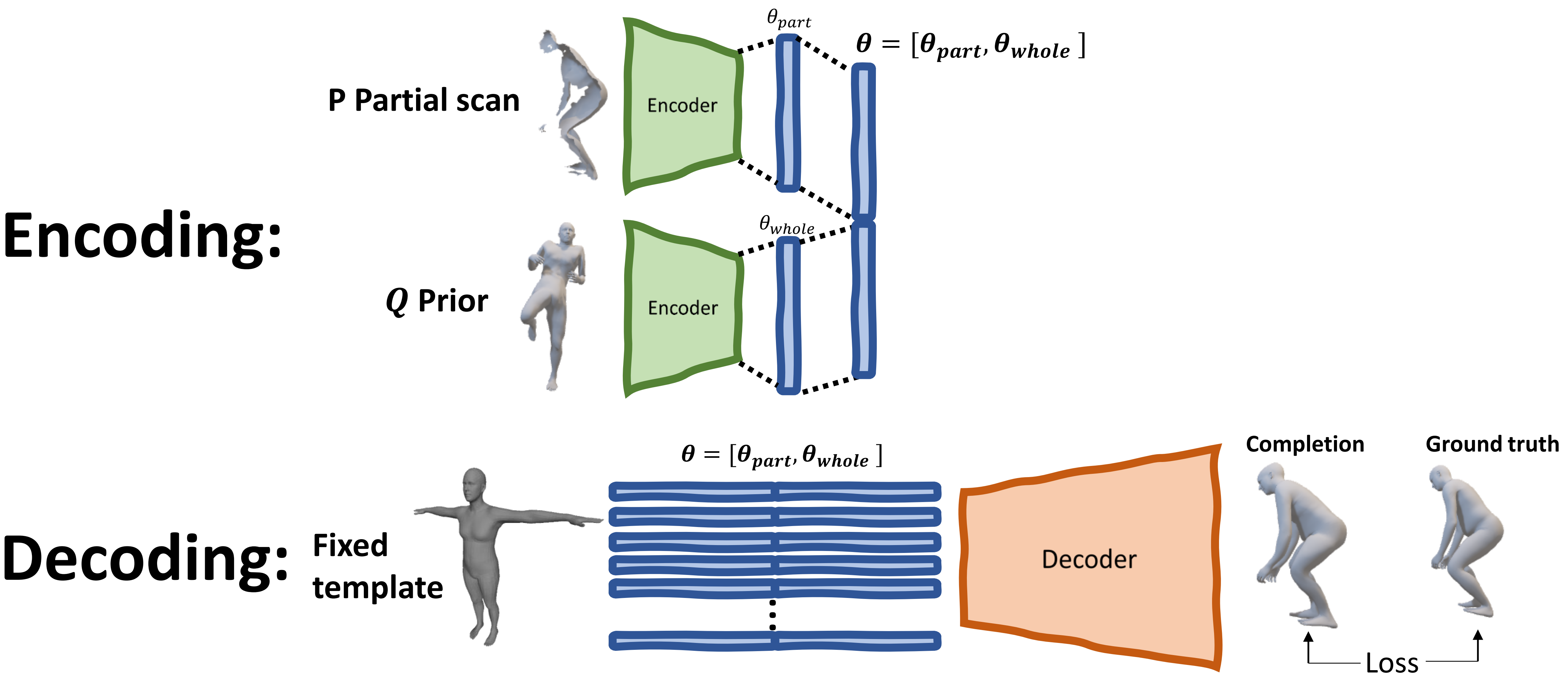
Completion idea – example 1



Scans source: Amass dataset. https://amass.is.tue.mpg.de

our method: FTS (fixed template shape)

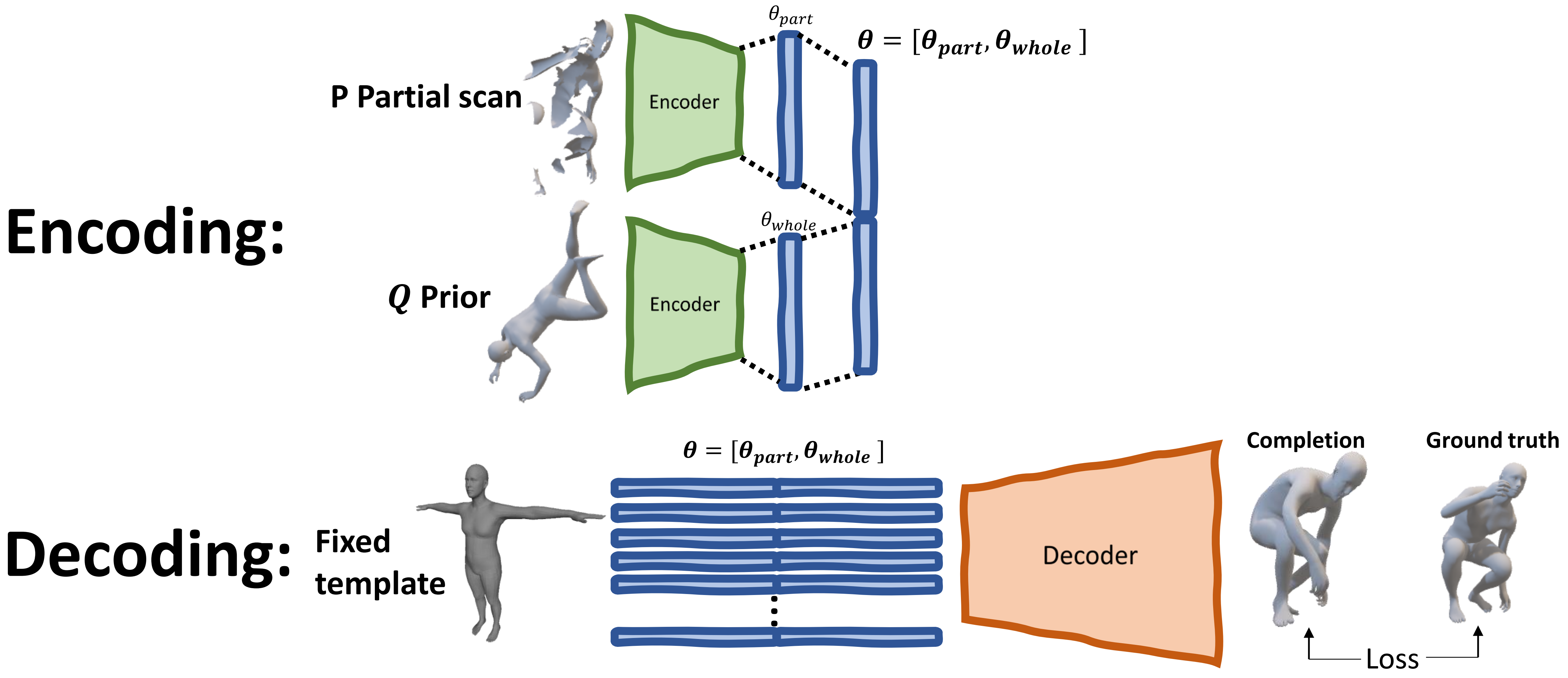
Completion idea – example 2



Scans source: Amass dataset. https://amass.is.tue.mpg.de

our method: FTS (fixed template shape)

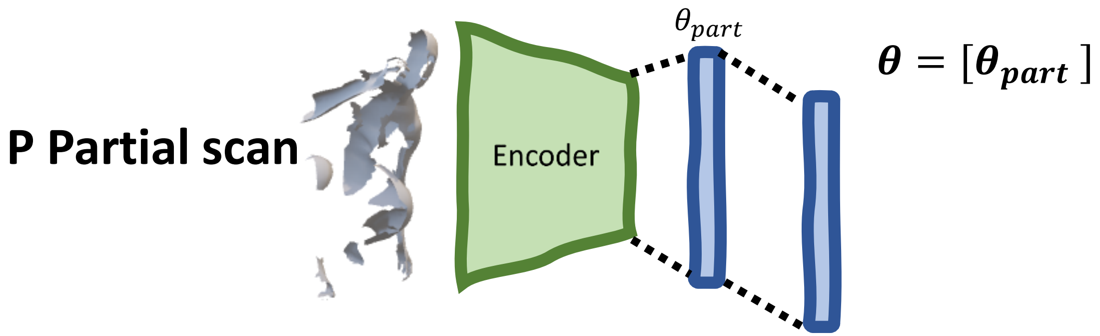
Completion idea – example 3



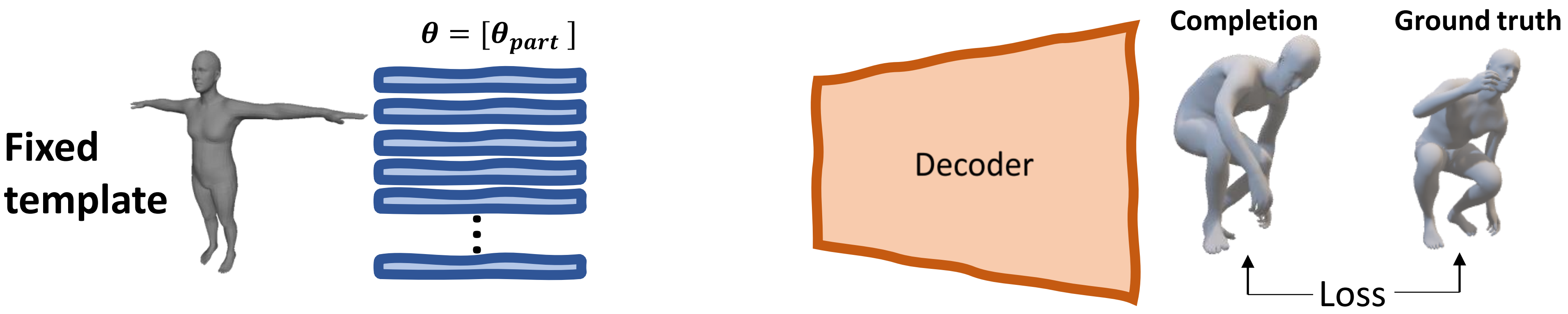
our method: **FTS-NP** (fixed template shape-no prior)

Completion idea

Encoding:



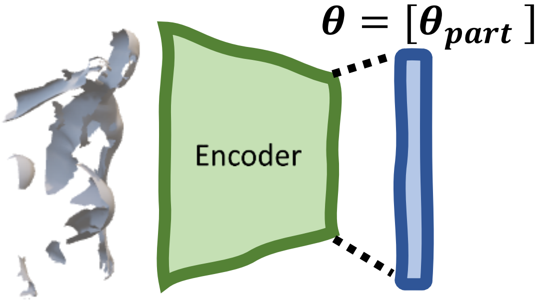
Decoding:



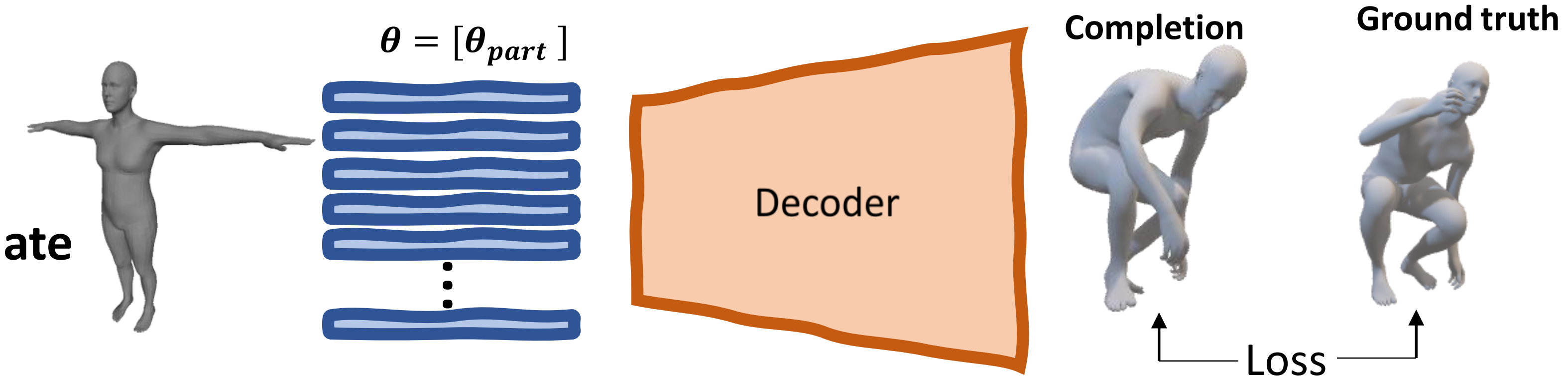
our method: FTS-NP (fixed template shape-no prior)

Completion idea

Encoding: **P Partial scan**



Decoding: **Fixed template**



Experiment

Methods in compression:

- **Towards Precise**
- **FTS**
- **FTS-NP**

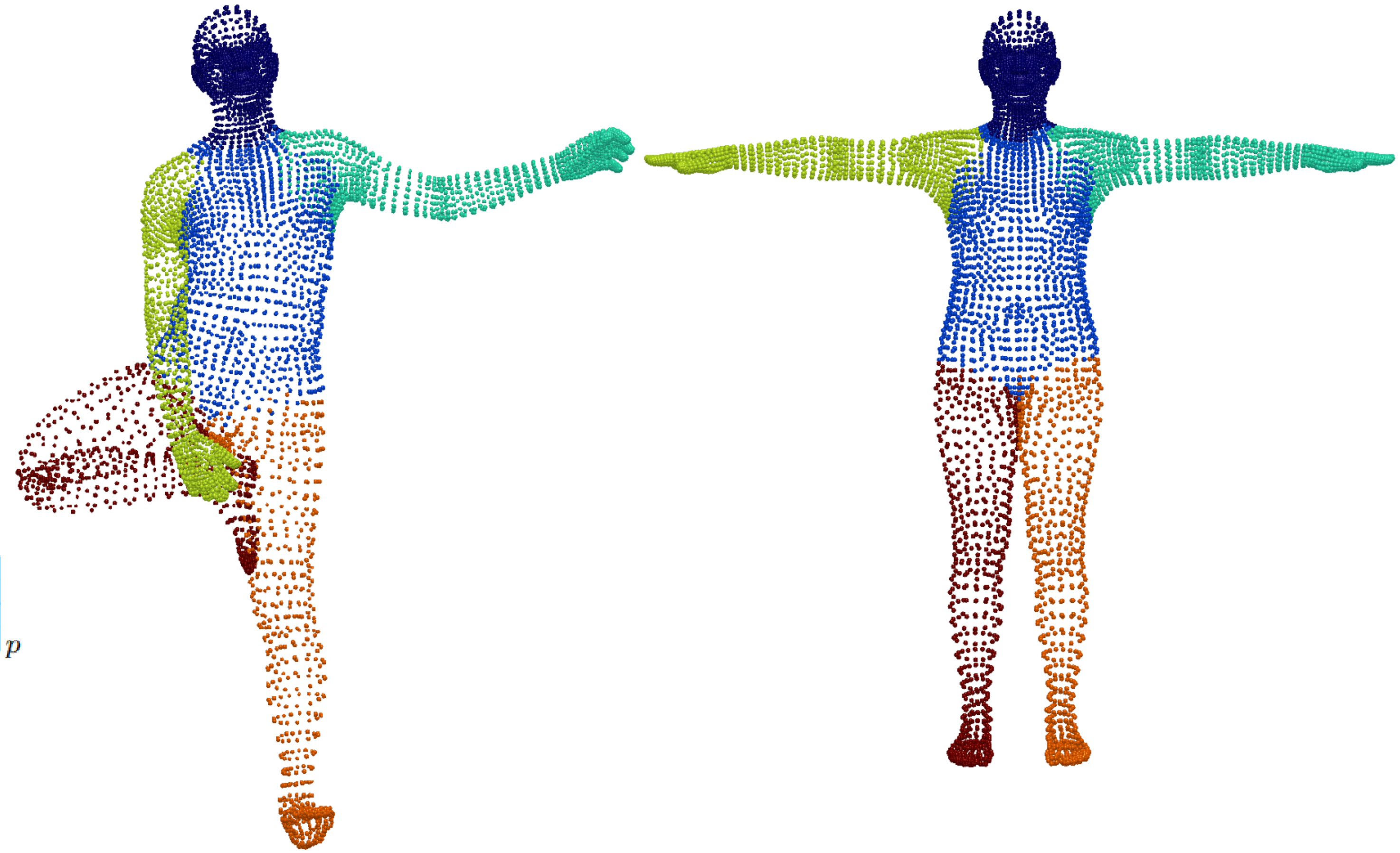
Tested on subsampled datasets from AMASS Dataset

(10 males, 50 samples for each actor)

The training set was sampled randomly.

Scans source: (Right) Amass dataset. https://amass.is.tue.mpg.de, (Left): SMPL+H rest pose

Evaluation metrics



- full body + **Segments**: head, torso , left arm , right arm , left leg, right leg.
- **errors:**

$$\forall p \in \{1, 2, \infty\} : \mathbb{E}_{l_n(G,R)} = \frac{1}{N} \sum_{i=1}^{n} \left|\left| g_i - r_i \right|\right|_p$$

$$\forall O \in \{Vol, Area\} : M_O\left(G, R\right) = \frac{|O(G) - O(R)|}{O(G)} \cdot 100$$

N = #Points in segment

Results

	Segment						
	full body	head	torso	left arm	right arm	left leg	right leg
$\mathbb{E}_{l_1}$ **Error [cm]**							
Towards Precise	0.240746	0.162363	0.166394	0.266586	0.270954	0.293877	0.303135
FTS-NP (ours)	0.172014	0.117620	0.121781	0.198461	0.200176	0.193018	0.199039
FTS (ours)	**0.147619**	**0.097046**	**0.103619**	**0.166146**	**0.170752**	**0.172403**	**0.182553**
$\mathbb{E}_{l_2}$ **Error [cm]**							
Towards Precise	0.162575	0.109220	0.111560	0.179142	0.183097	0.200223	0.206027
FTS-NP (ours)	0.116217	0.078855	0.081467	0.133859	0.135026	0.132671	0.135104
FTS (ours)	**0.099525**	**0.064675**	**0.069361**	**0.111625**	**0.115255**	**0.117476**	**0.124246**
$\mathbb{E}_{l_\infty}$ **Error [cm]**							
Towards Precise	0.137028	0.091874	0.093337	0.150302	0.154346	0.170173	0.174576
FTS-NP (ours)	0.097944	0.065986	0.068019	0.112450	0.113825	0.113498	0.114432
FTS (ours)	**0.083730**	**0.053837**	**0.057882**	**0.093691**	**0.097165**	**0.099761**	**0.105167**
Volumetric error [%]							
Towards Precise	46.854080	45.355503	44.622883	55.759094	57.923038	51.874989	52.759754
FTS-NP (ours)	31.126259	32.397682	28.613161	39.795094	42.709961	32.265114	32.649700
FTS (ours)	**26.938036**	**24.644217**	**24.285990**	**32.891121**	**35.859035**	**29.393967**	**29.738489**
Surface area error [%]							
Towards Precise	30.000505	32.000538	27.276175	36.425770	39.768318	33.572063	34.273651
FTS-NP (ours)	22.995281	22.502445	19.693295	29.450409	31.579748	22.296270	23.120274
FTS (ours)	**19.846867**	**16.847292**	**17.354050**	**24.950127**	**26.357685**	**19.766914**	**20.035872**

Table 1: Comparison of Towards-precise , FTS-NP (ours) and FTS (ours) shape completion methods with respect to the described evaluation metrics on each segment. All the methods was trained on the high-variance males random (MR) dataset. The minimum value on each column appear in bold.

Results

- Our completion methods outperform the existing state-of-the-art Towards Precise method across **all segmentations** and **all errors**.
- **FTS** method **surpass FTS-NP** method in all the same aspects as well, which imply the **prior importance** in shape completion.

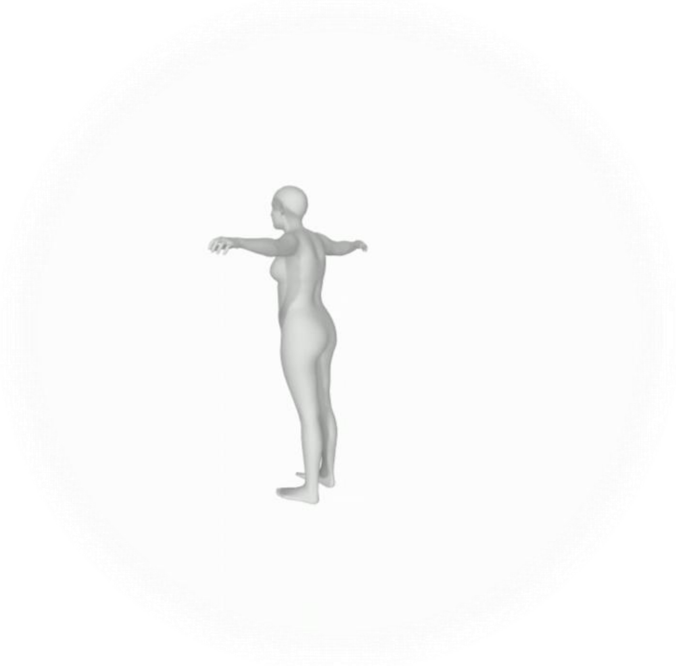
How can we improve results?

- Hyper parameters tunning
- Dataset sampling
- Achieving similar results using weaker prior

Our focus

AMASS Dataset

- 14 inner datasets
- 344 actors
- 11,265 animations, 17 million frames



Scans source: Amass dataset. https://amass.is.tue.mpg.de

AMASS – deep dive

- AMASS collected motion capture scans from existing datasets.
- Each scan was transformed into SMPL+H **parametric body model**.
- Finally, the feature vector is decoded into a synthetic body model.

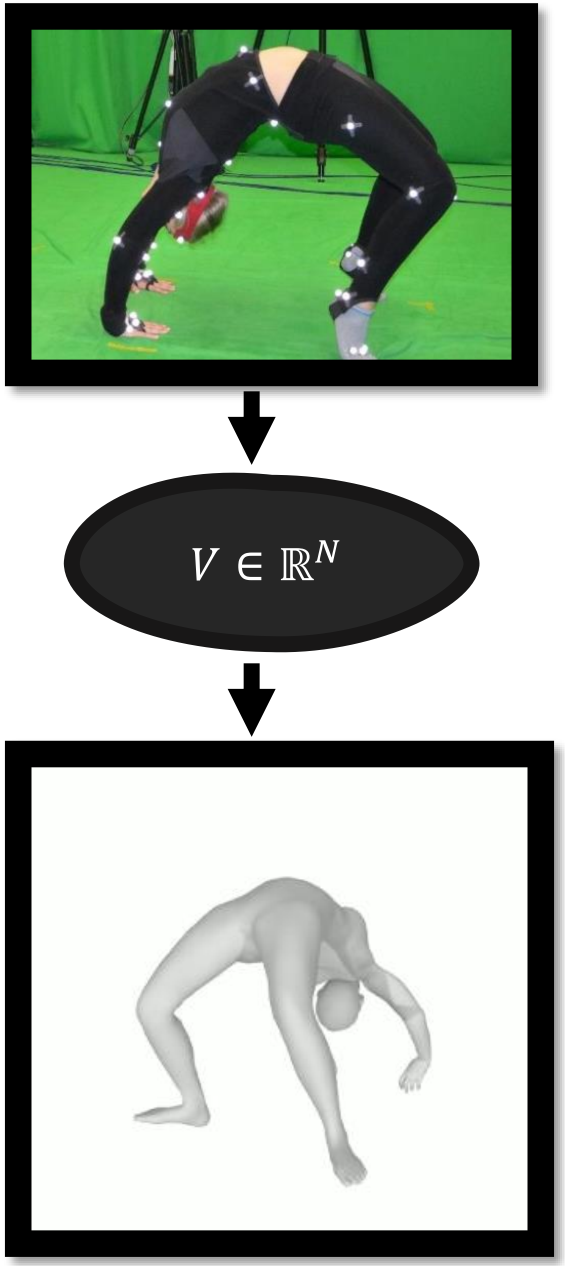
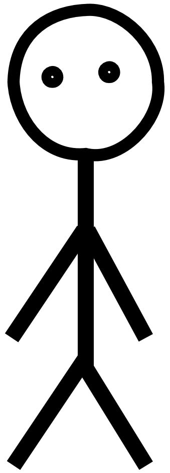


Image :
pose-Conditioned Joint Angle Limits for 3D Human Pose Reconstruction
Scan :
Amass dataset. https://amass.is.tue.mpg.de

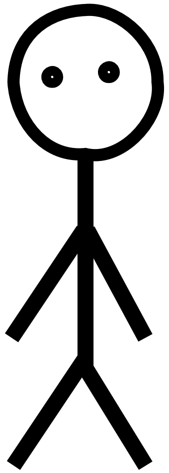
Parametric body models 2D example



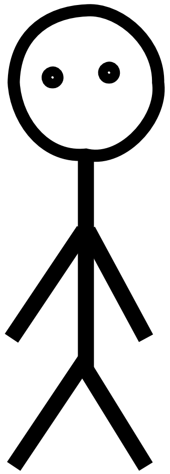
Parametric body models 2D example

Hands legs

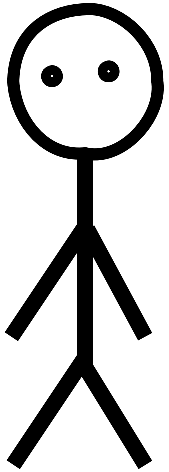
$$\Theta_{pose} = (\theta_1, \theta_2, \theta_3, \theta_4)$$

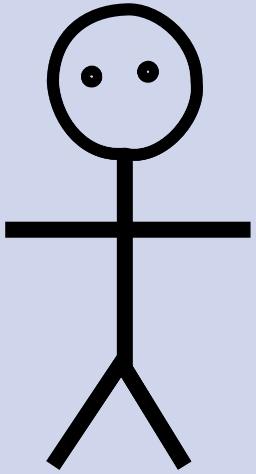
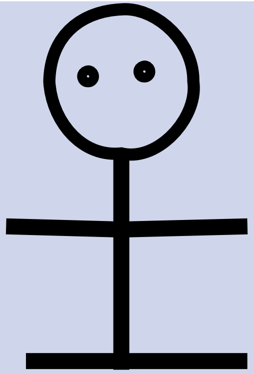
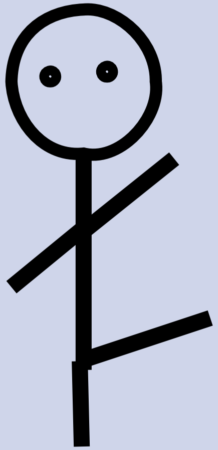


Parametric body models 2D example



$$\Theta_{pose} = (\overset{\text{Hands}}{\theta_1, \theta_2,} \overset{\text{legs}}{\theta_3, \theta_4})$$

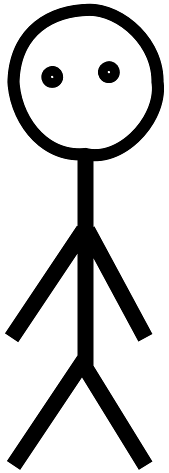


$\Theta_{pose} = (\theta_1, \theta_2, \theta_3, \theta_4)$	(30,30,90,90)	(90,90,90,90)	(110,0,45,135)
Model pose			

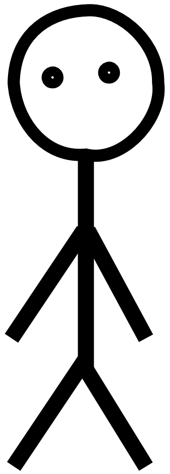
Parametric body models 2D example

Head size, Hight

$$\mathrm{B}_{shape} = (\beta_1, \beta_2)$$

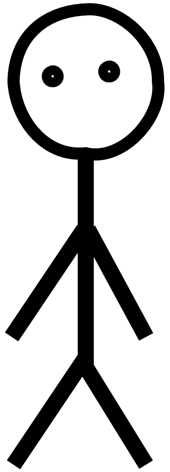


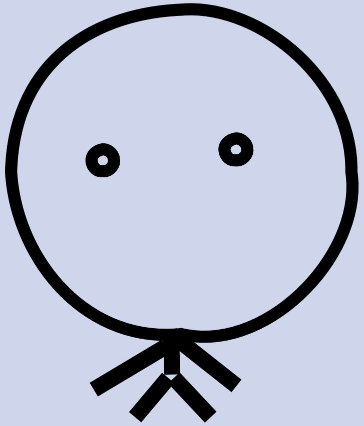
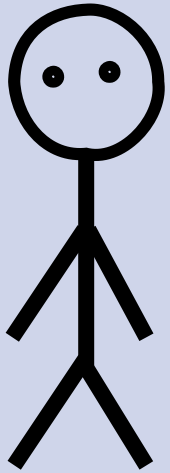
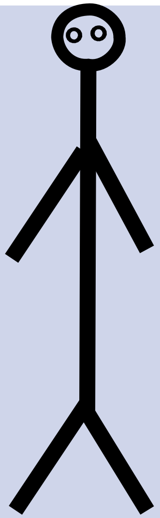
Parametric body models 2D example



Head size, Hight

$$B_{shape} = (\beta_1, \beta_2)$$

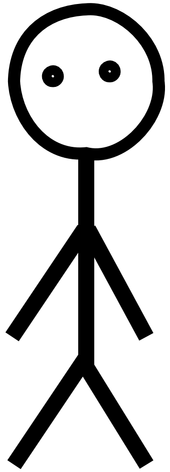


$B_{shape} = (\beta_1, \beta_2)$	(100,1)	(10,10)	(1,100)
Model Shape			

Parametric body models 2D example

Hands legs

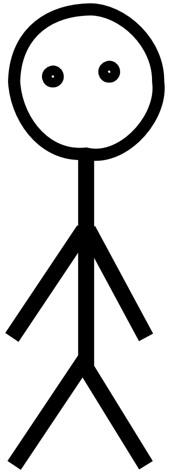
$$\Theta_{pose} = (\theta_1, \theta_2, \theta_3, \theta_4)$$



+

Head size, Hight

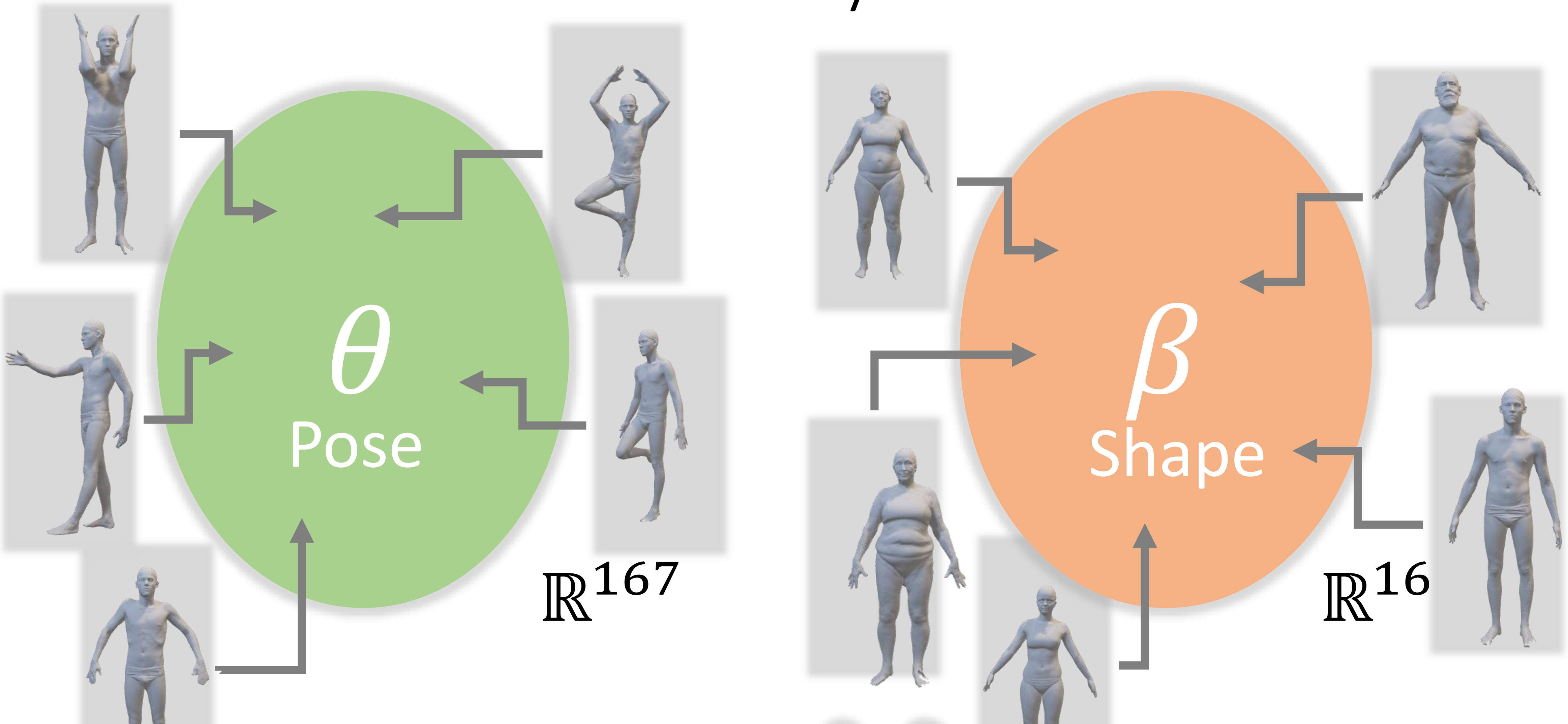
$$\mathrm{B}_{shape} = (\beta_1, \beta_2)$$



= Stickman model

Scans source: MPI-Faust dataset. http://faust.is.tue.mpg.de

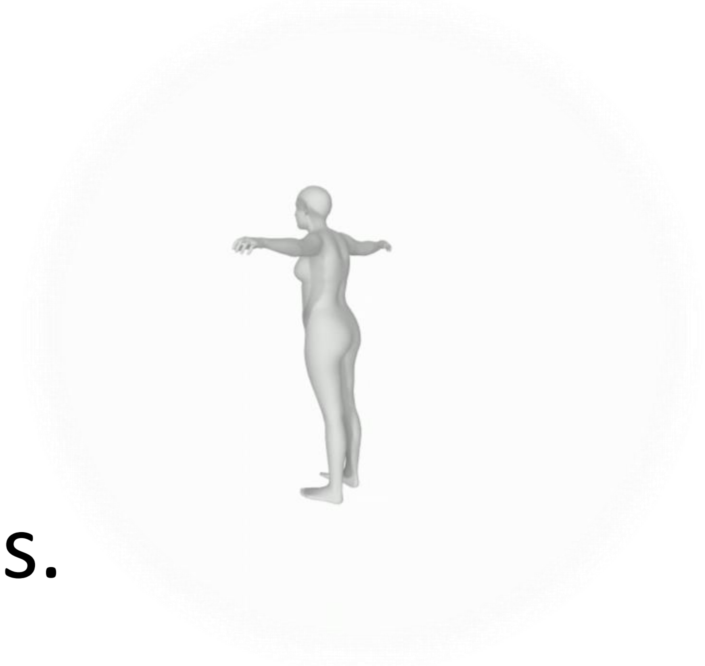
SMPL+H Parametric body model



Parametric human model $= (\theta, \beta)$

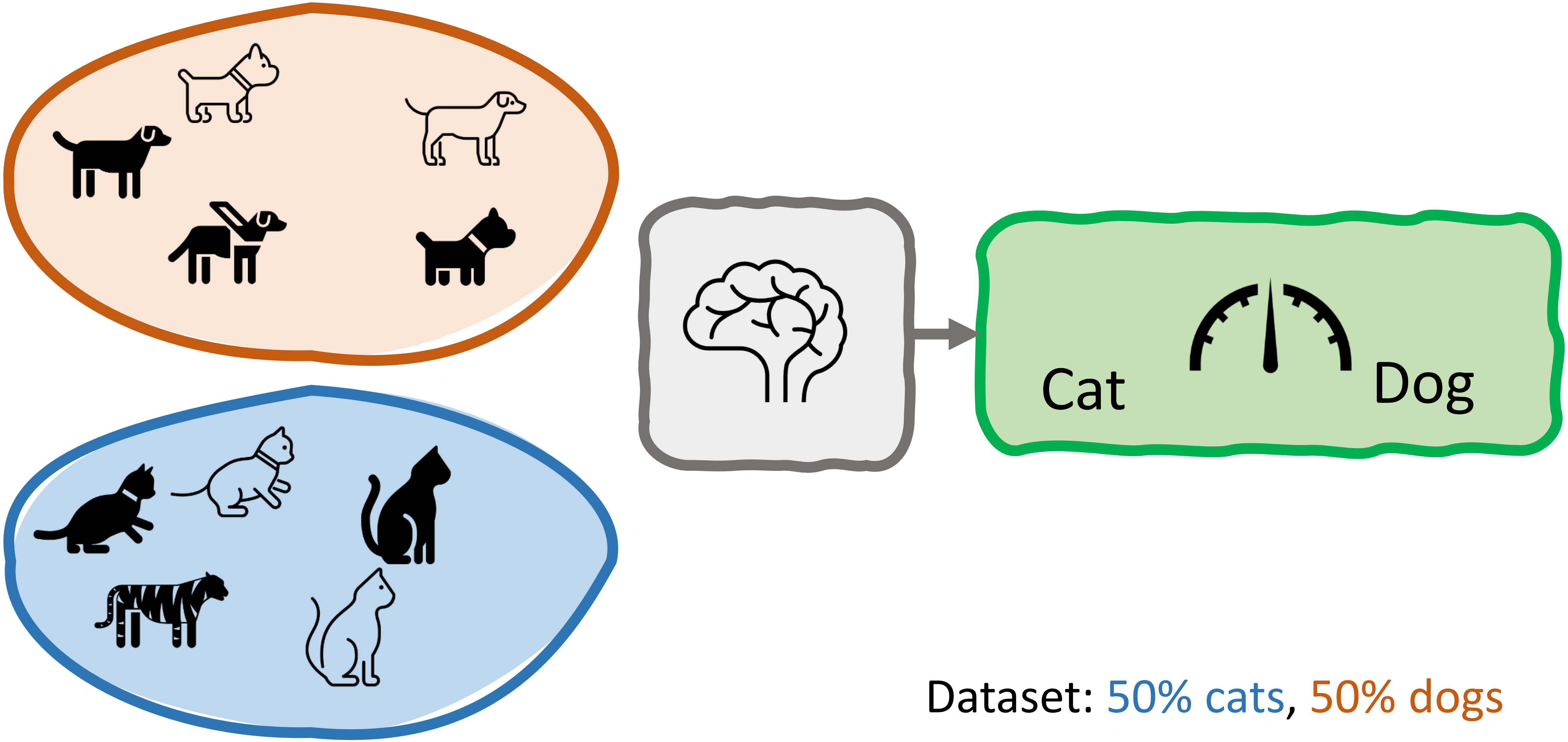
AMASS Dataset

- 14 inner datasets
- 344 actors
- 11,265 animations, 17 million frames
- **Problem:** Heavily biased towards rest poses.



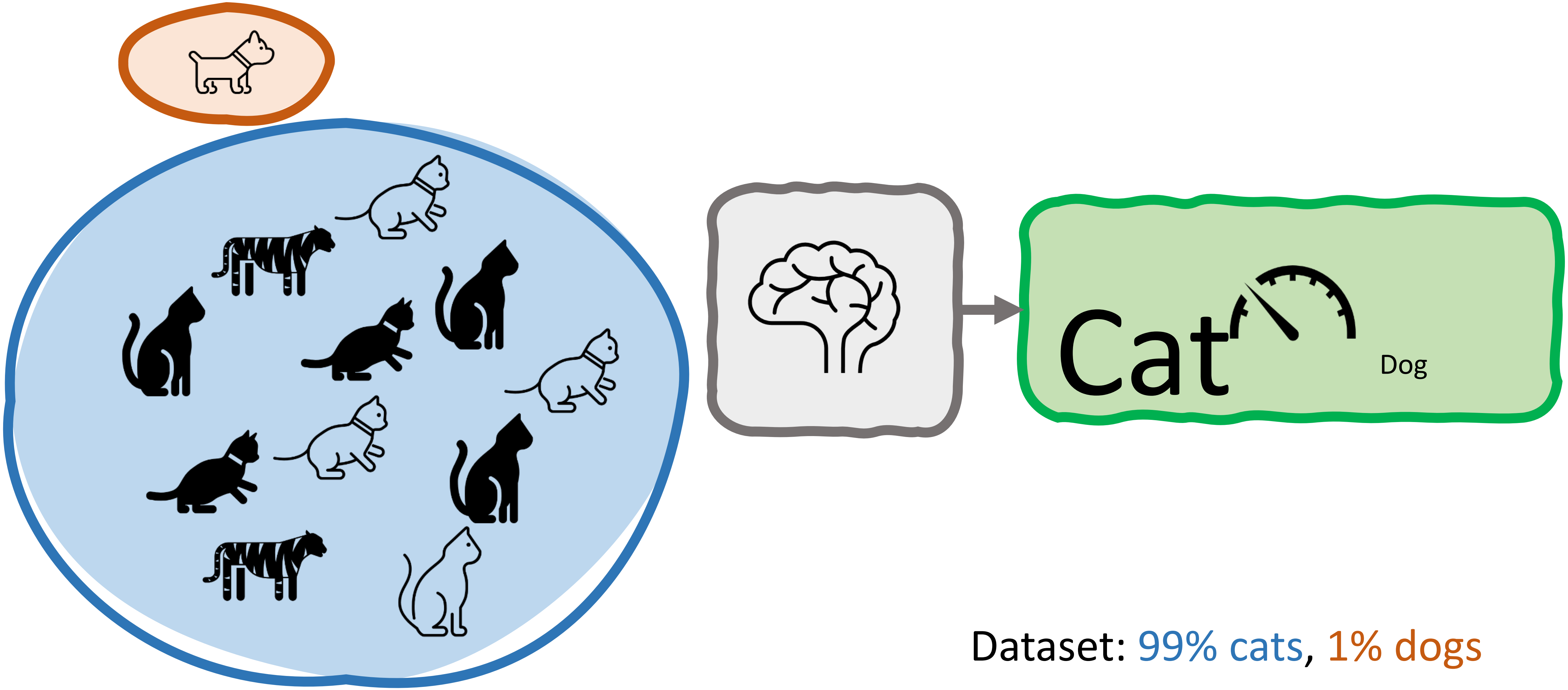
Scans source: Amass dataset. https://amass.is.tue.mpg.de

Biased dataset problem



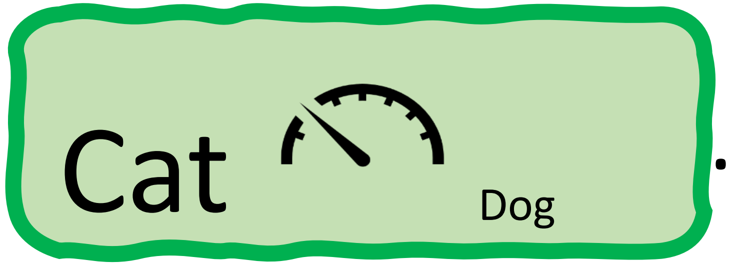
Dataset: 50% cats, 50% dogs

Biased dataset problem



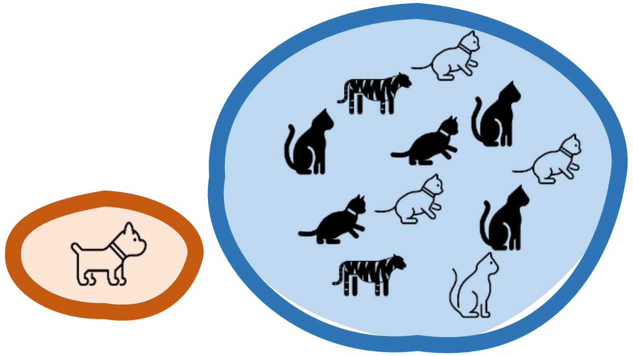
Dataset: 99% cats, 1% dogs

Biased dataset problem

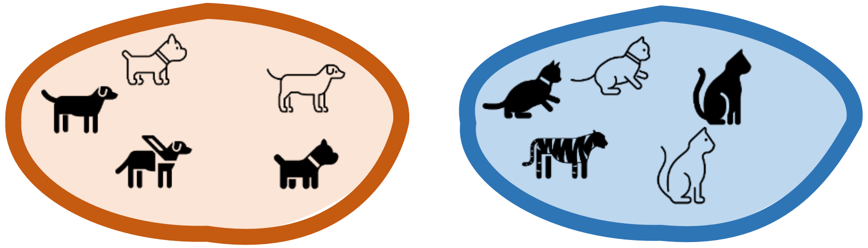
Testing the bad classifier 

.

- biased testing set of (99% cats, 1% dogs) leads to false evaluation of the model.

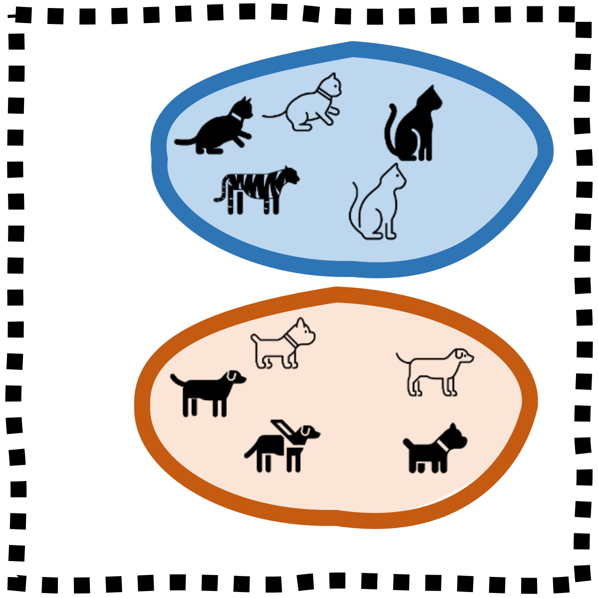


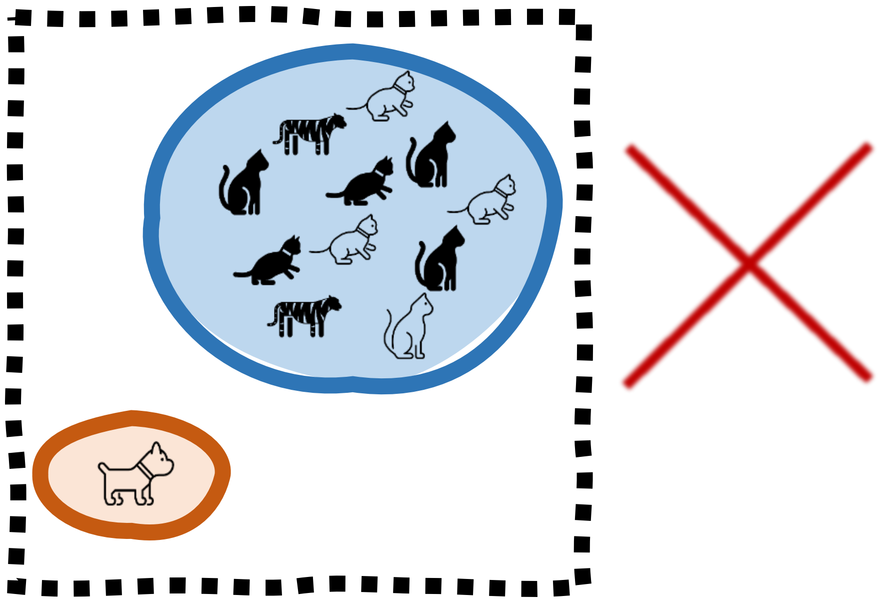
- However, testing with unbiased dataset of (50% cats, 50% dogs) will reveal the problem.



Biased dataset problem

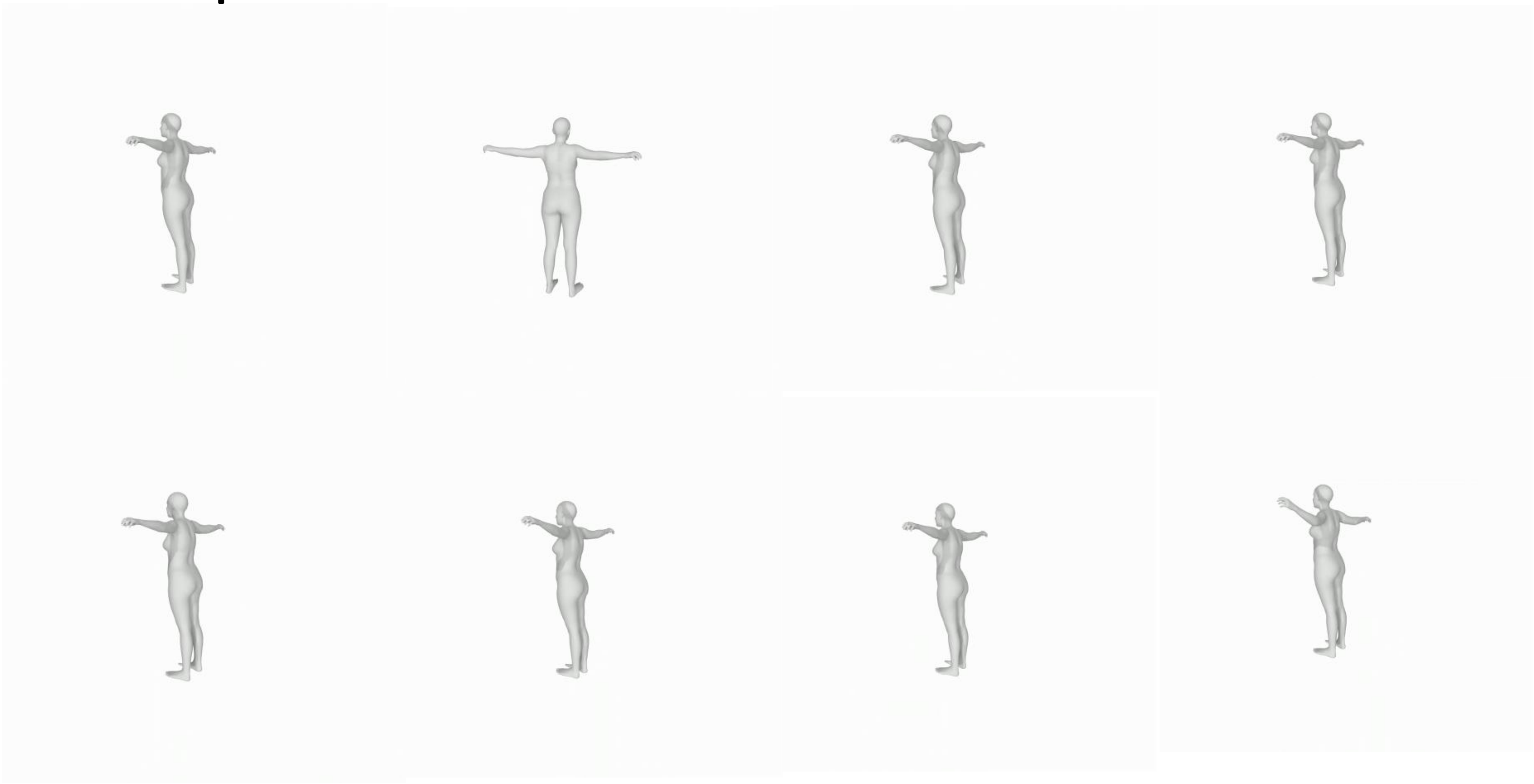
- Important Conclusion: the testing set should be as unbiased as possible.





Problem - AMASS is heavily biased towards rest poses

Aspect 1: Several animations for an actor:

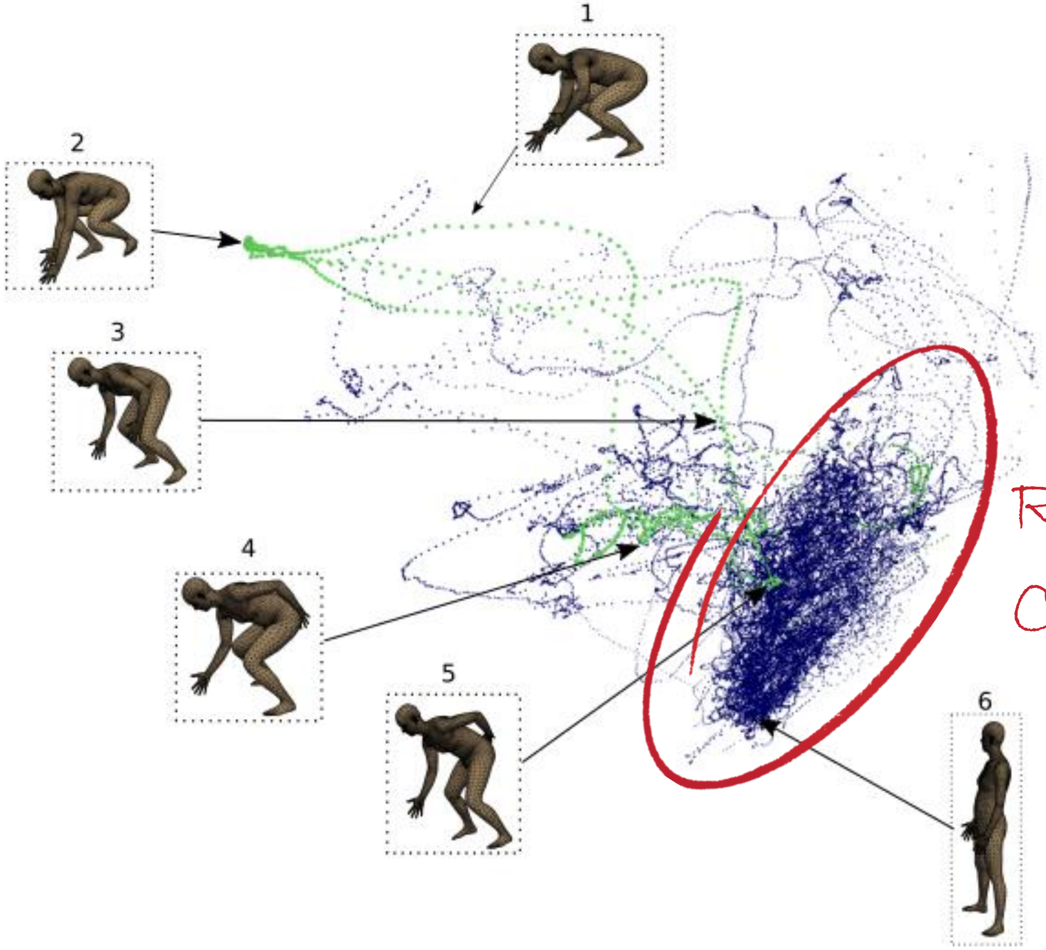


Scans source: Amass dataset. https://amass.is.tue.mpg.de

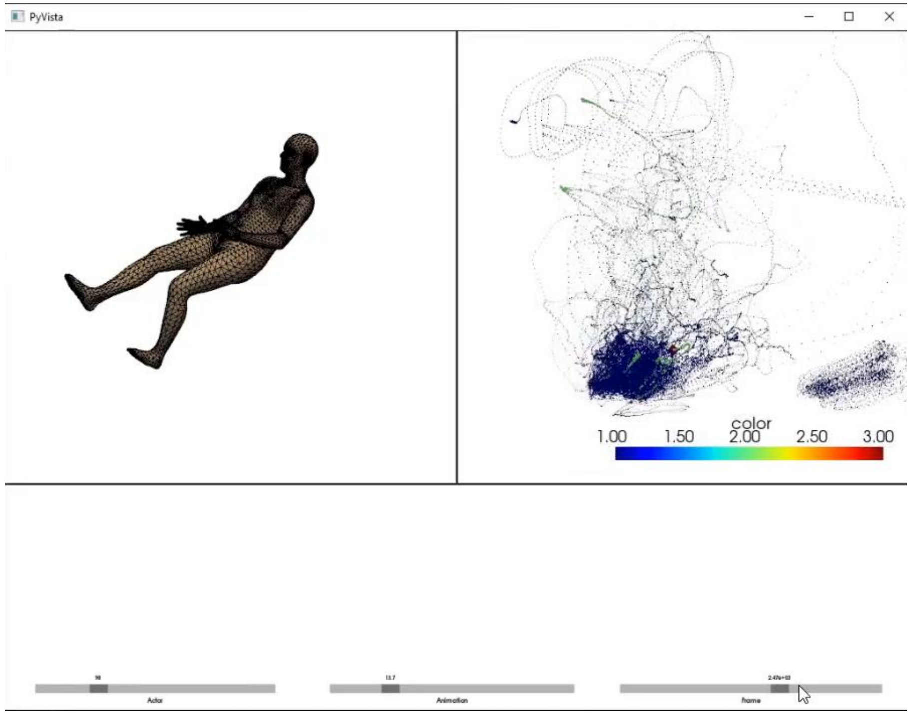
Scans source: Amass dataset. https://amass.is.tue.mpg.de

Problem - AMASS is heavily biased towards rest poses

Aspect 2:3D PCA projection of θ pose for an actor:

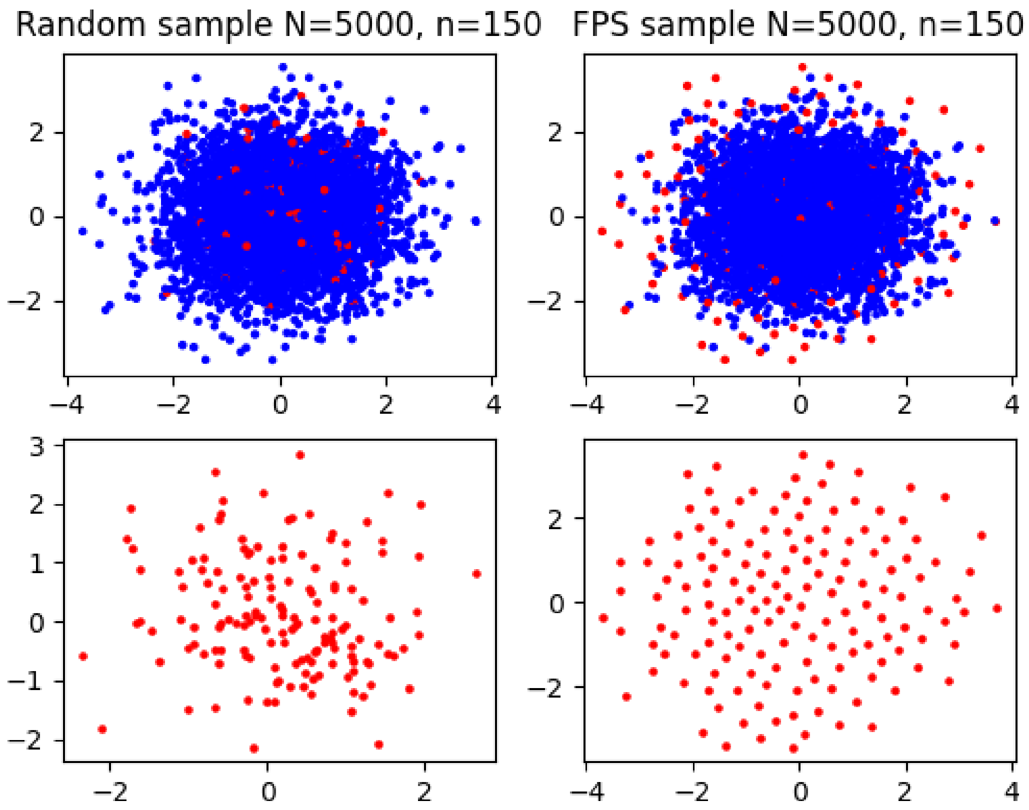


Rest pose Centroid!



Our solution: FPS sampling (farthest point sampling)

- We propose the FPS sampling methodology in order to sample subset of points that are farther away from each other.



Scans source: Amass dataset. https://amass.is.tue.mpg.de

Our solution: FPS sampling (farthest point sampling)

- We sample the frames using FPS sampling **over the shape vector** θ_{shape} on SMPL+H, resulting unbiased pose manifold with high variably.

FPS result for $n = 4$ on actor from BMLmovi dataset

Our solution: FPS sampling (farthest point sampling)

Algorithm 1: FARTHEST POINT SAMPLING (FPS)

Data:
$N \in \mathbb{N}^+$ the number of frames of a given actor.
$n \in \{1, 2, ..., N\}$ the number of frames to be sampled.
$\mathbf{\Theta}^{\text{body pose}} = \{\boldsymbol{\theta}^{\text{body pose}}\}_{i=1}^{N}$ the body pose vectors for each frame.
Result: selected frames $S \subseteq \{1, 2, ..., N\}$, $|S| = n$
Function *FPS*(N , n , $\mathbf{\Theta}^{body\ pose}$)
  sample randomly: $s \leftarrow \mathcal{U}\{1, 2, ..., N\}$
  $S \leftarrow \{s\}$
  $U \leftarrow \{1, 2, ..., N\} \setminus S$
  while $|S| < n$ **do**
    $\forall i \in U : d_i^{min} = \min_{j \in U; j \neq i} \left|\left| \boldsymbol{\theta}_j^{\text{body pose}} - \boldsymbol{\theta}_i^{\text{body pose}} \right|\right|_2$
    $s \leftarrow \arg\max_{i \in U} d_i^{min}$
    $S \leftarrow S \cup \{s\}$
    $U \leftarrow U \setminus \{s\}$
  return S

Sampling Experiment

Names:

	Males	Females
Random	high-variance males random (MR)	high-variance females random (FR)
FPS	high-variance males fps (MF)	high-variance females fps (FF)

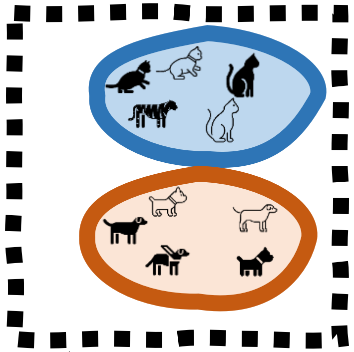
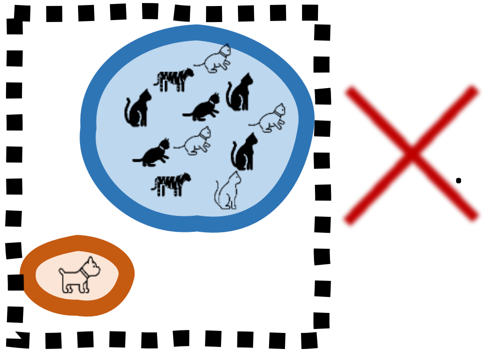
[M/F] [R/F]

Gender: Males or Females

Sample: Random or FPS

Splits:

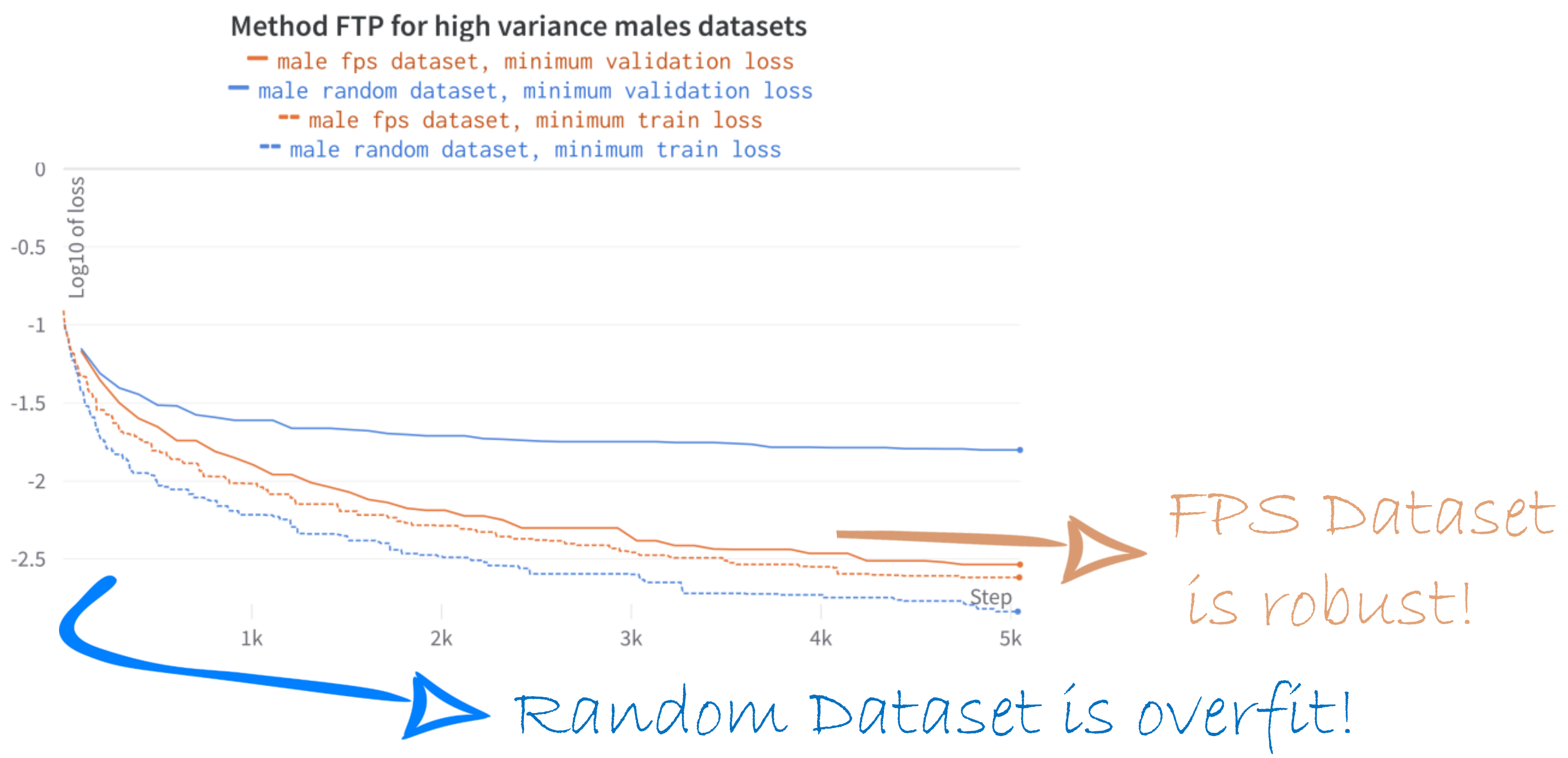
	Train	Validation	Test
males random	MR	MF	MF
males fps	MF	MF	MF
females random	FR	FF	FF
females fps	FF	FF	FF

Remember, we want to test with  and **not** with .

unbiased dataset

biased dataset

Results

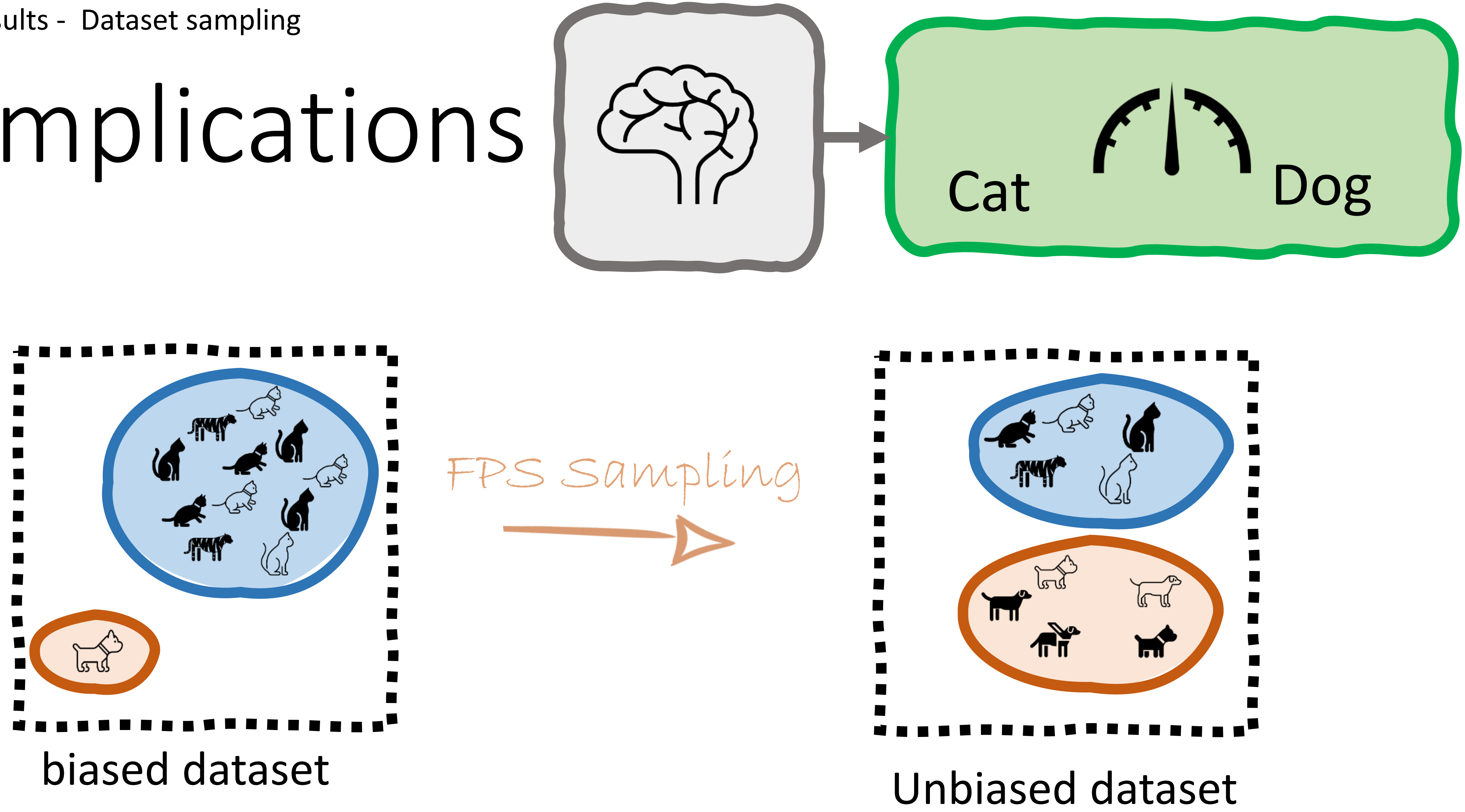

Method FTP for high variance males datasets
male fps dataset, minimum validation loss
male random dataset, minimum validation loss
male fps dataset, minimum train loss
male random dataset, minimum train loss
Log10 of loss
0
-0.5
-1
-1.5
-2
-2.5
Step
1k
2k
3k
4k
5k
FPS Dataset is robust!
Random Dataset is overfit!

Results

$\mathbb{E}_{l_2}$ Error [cm]	FTS (ours)	FTS-NP (ours)	Towards precise
males random	0.0977	0.1142	0.1584
males fps (ours)	0.0390	0.0474	0.0908
imrprovement factor	**2.505**	**2.409**	**1.744**
females random	0.0786	0.0715	0.1399
females fps (ours)	0.0348	0.0363	0.08244
imrprovement factor	**2.258**	**1.969**	**1.696**

significant performance improvement factor: **at least 1.5**, **across all cases**.

Implications



How can we improve results?

- Hyper parameters tunning
- Dataset sampling
- Achieving similar results using weaker prior

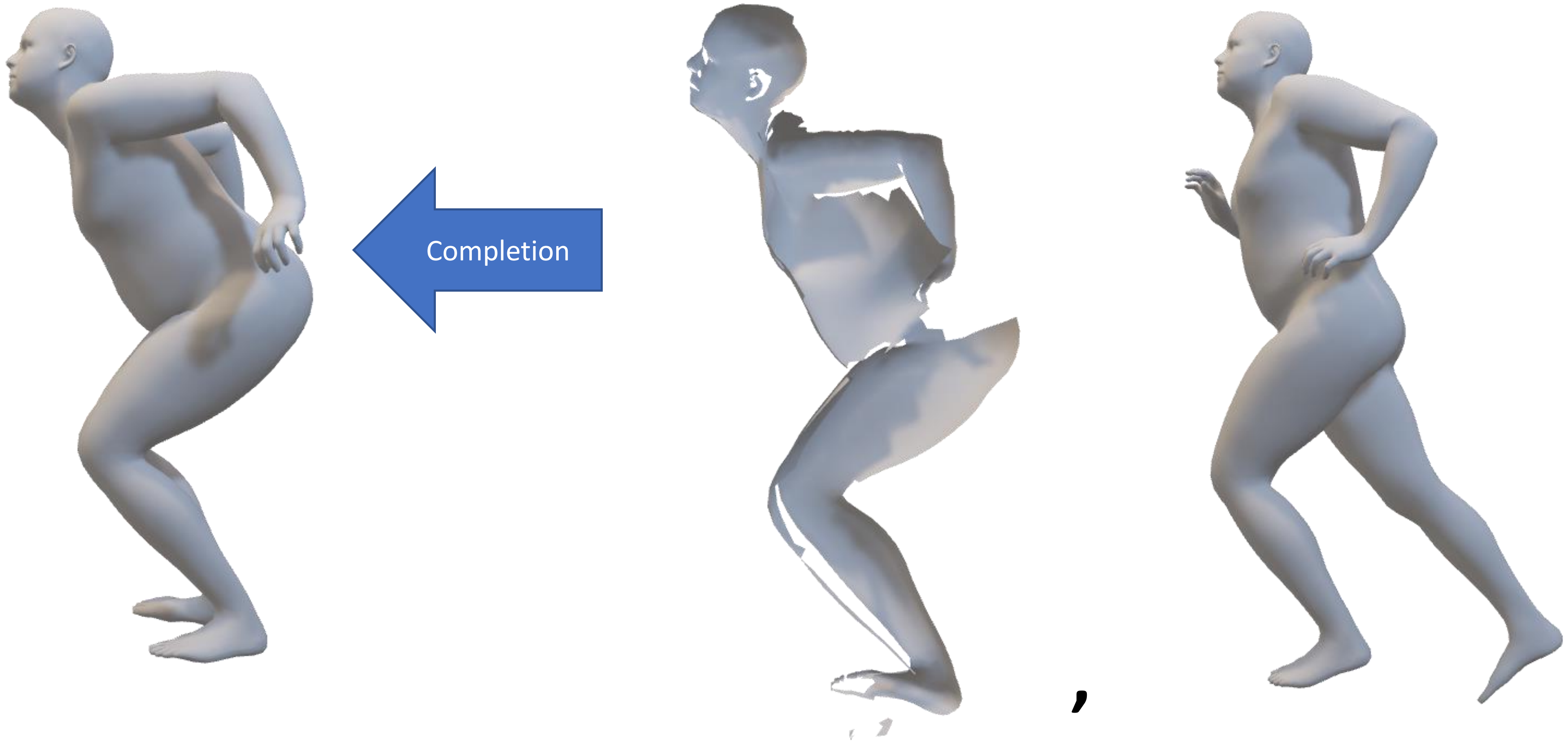
Our focus

Scans source: Amass dataset. https://amass.is.tue.mpg.de

What if we don’t have the full prior?

Bridging the information gap

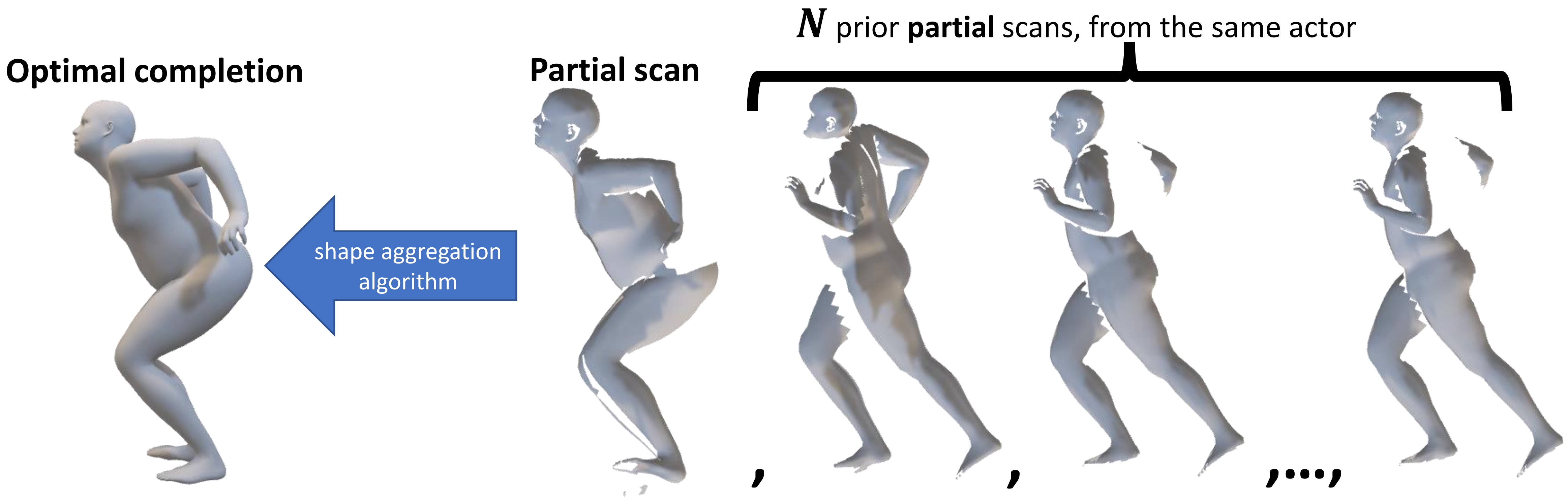
Optimal completion **Partial scan** **Prior**



Scans source: Amass dataset. https://amass.is.tue.mpg.de

What if we don’t have the full prior?

Bridging the information gap



Improve results – utilize weaker prior

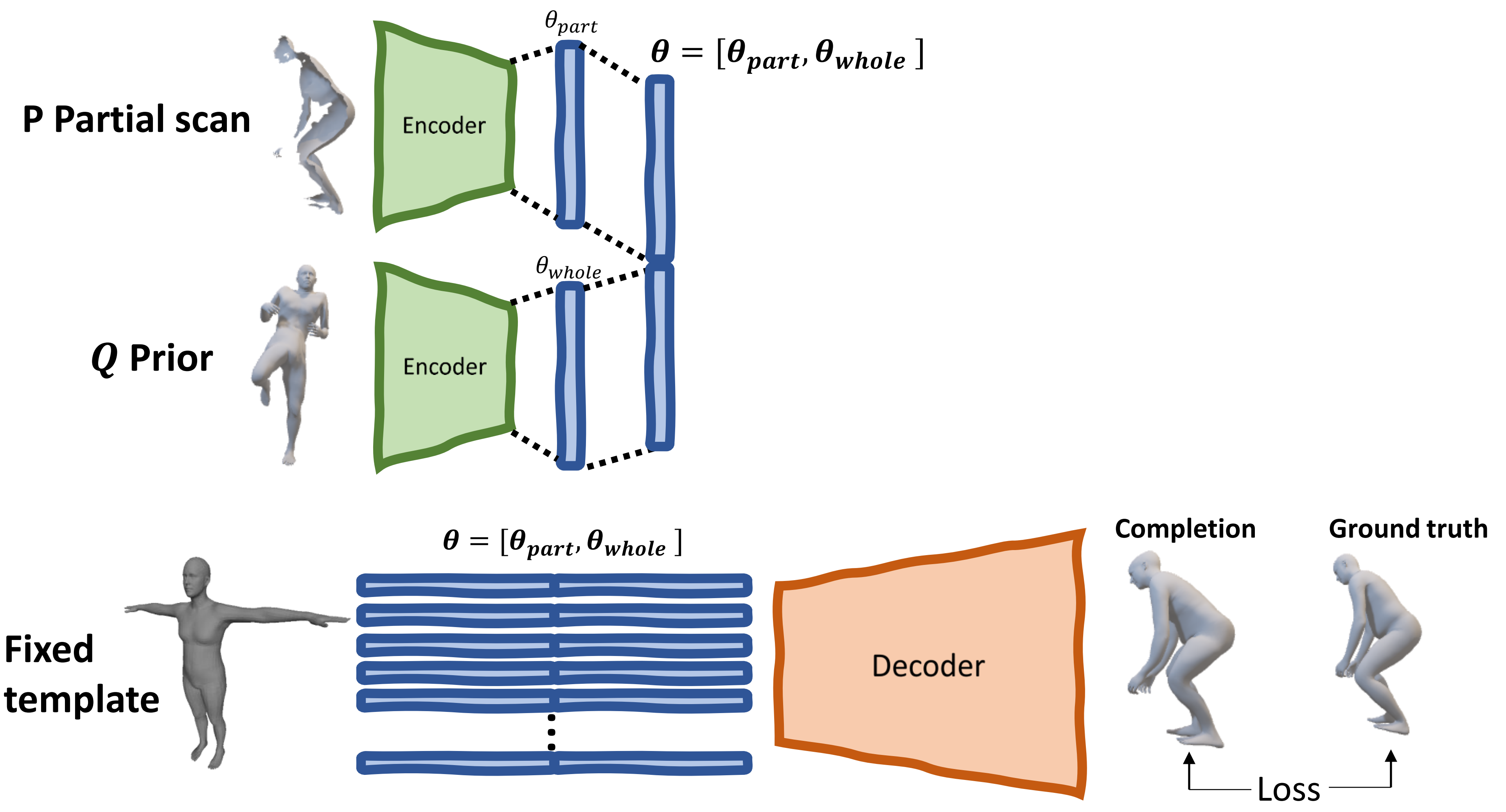
Scans source: Amass dataset. https://amass.is.tue.mpg.de

our method: FTS (fixed template shape)

Completion idea

Encoding:

Decoding:

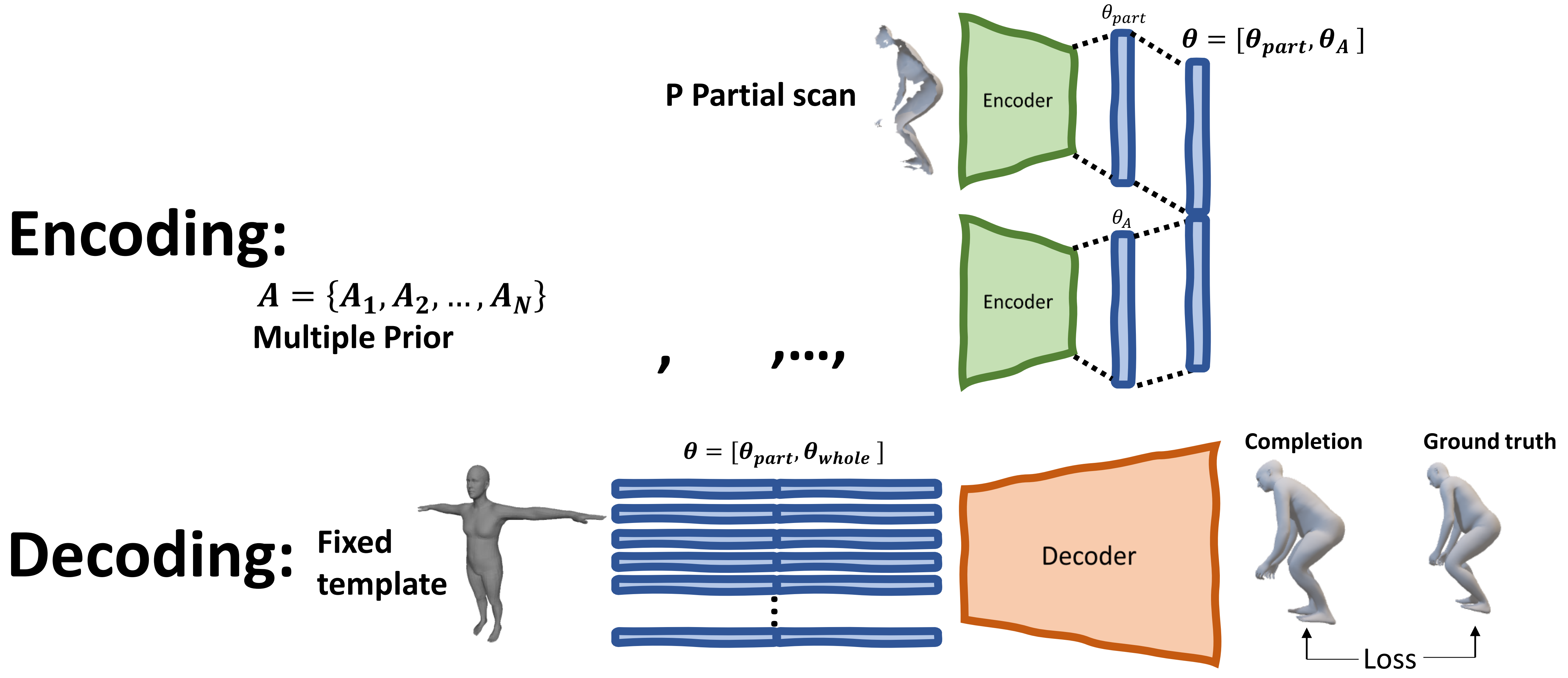


Improve results – utilize weaker prior

Scans source: Amass dataset. https://amass.is.tue.mpg.de

our method: FTMP (fixed template multiple prior)

Completion idea



Scans source: Amass dataset. https://amass.is.tue.mpg.de

Experiment

FTMP with different multiple prior scenarios:

"Same pose"

- N Partial point clouds acquired from the **same** actor (as the partial scan) and the **same** pose (Multi-view stereo)

P Partial scan

$\boldsymbol{A = \{A_1, A_2, \dots, A_N\}}$

Multiple Prior

, ,…,

"Other pose"

- N Partial point clouds acquired from the **same** actor (as the partial scan), in **another pose**

P Partial scan

$\boldsymbol{A = \{A_1, A_2, \dots, A_N\}}$

Multiple Prior

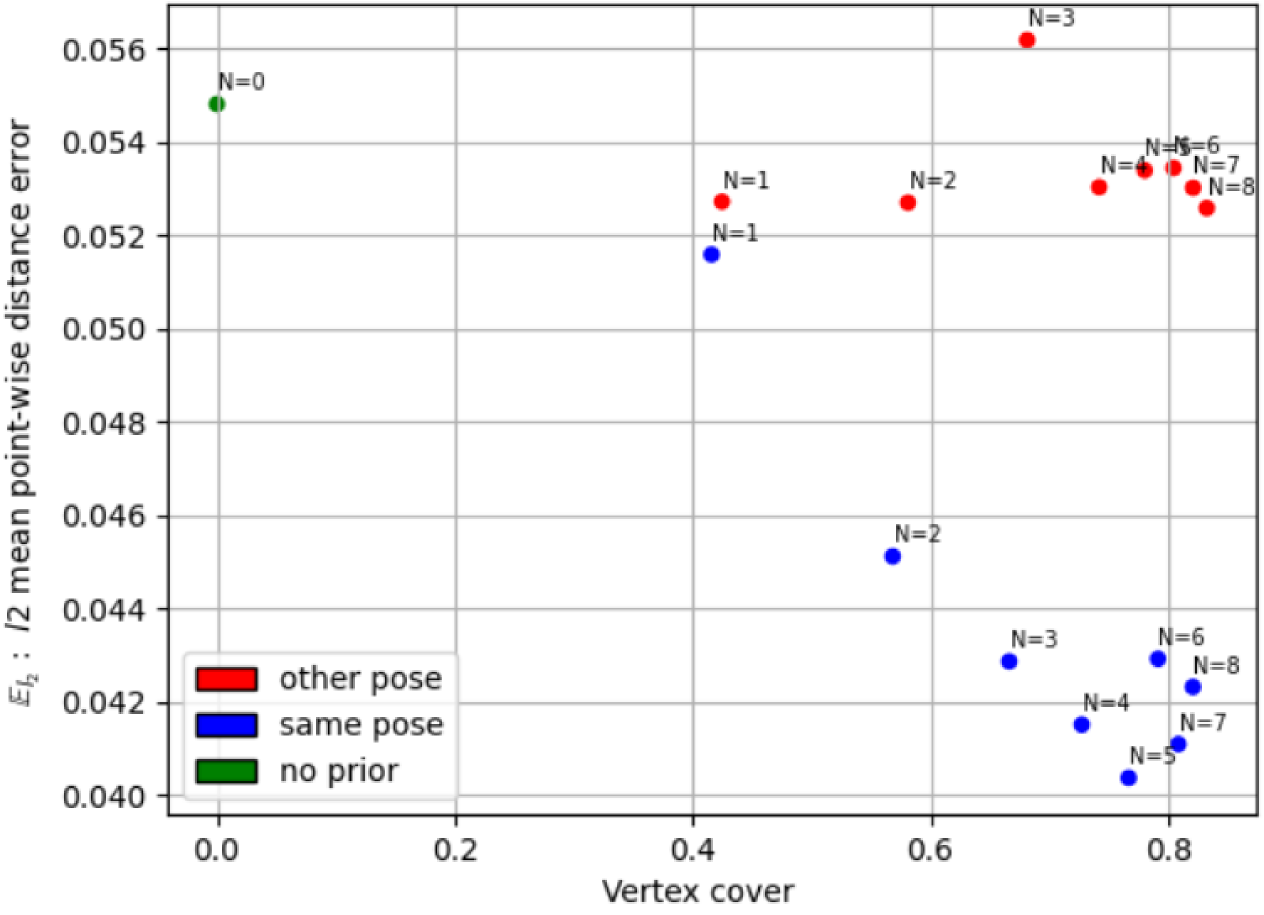
, ,…,

The scenarios compared on extensive version of AMASS, over 250 actors with 1K frames each, across $N \in \{0,1 \dots, 8\}$ partial point clouds.

Results

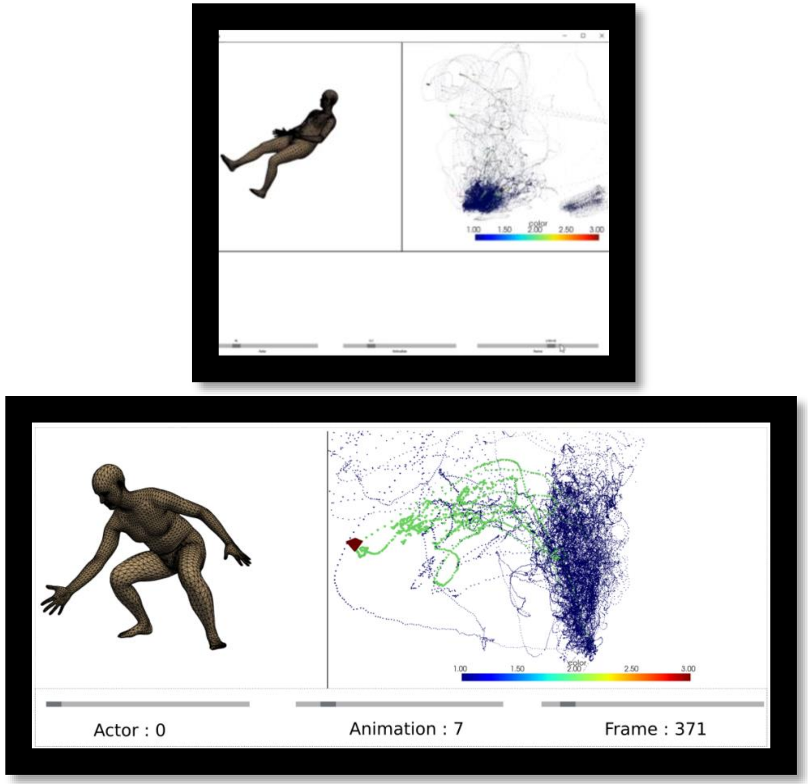
- Almost across all the cases, results with $N > 0$ were better than the no prior case $N = 0$.
- On the Same pose scenario:
  - Steady error decrease for $N \in \{1, \dots, 5\}$
  - On larger N's, the error stop decrease. (curse of dimensionality)
- On the Other pose scenario:
  - There was no significant difference between the experiments with respect to the error rate.
  - Might be related to the usage of the size of the shape descriptor components sizes relation $\theta = [\theta_{part}, \theta_A]$. In our experiment, θ comprised of two vectors with same lengths $|\theta_{part}| = |\theta_A| = 512$.

FTMP $\mathbb{E}_{l_2}$ error for $N \in \{0, \dots, 8\}$ **Partial** point clouds

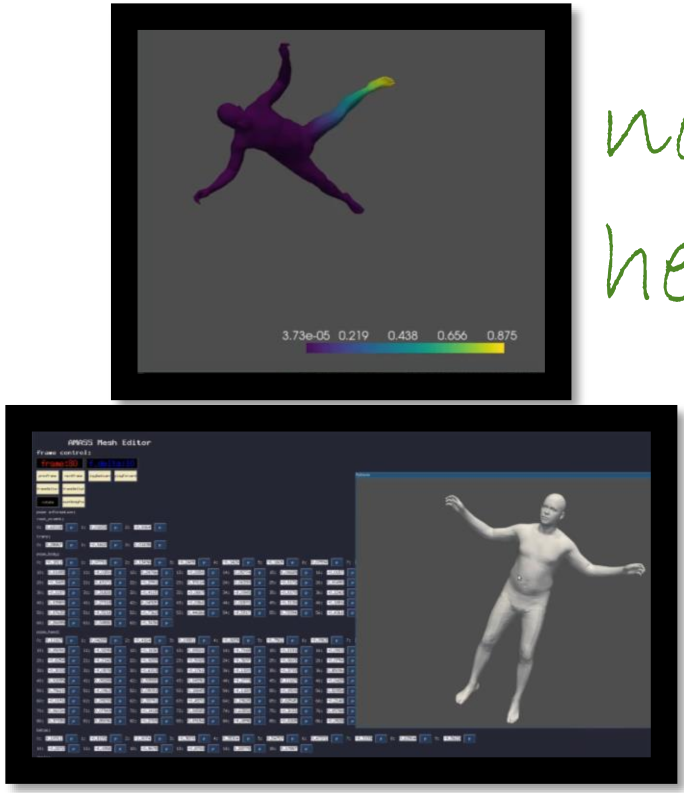


Contributions
visualization tools

1. We developed **strong visualization tools** to explore the shape manifold of parametric body models.



AMSS Pose manifold explorer



not covered here

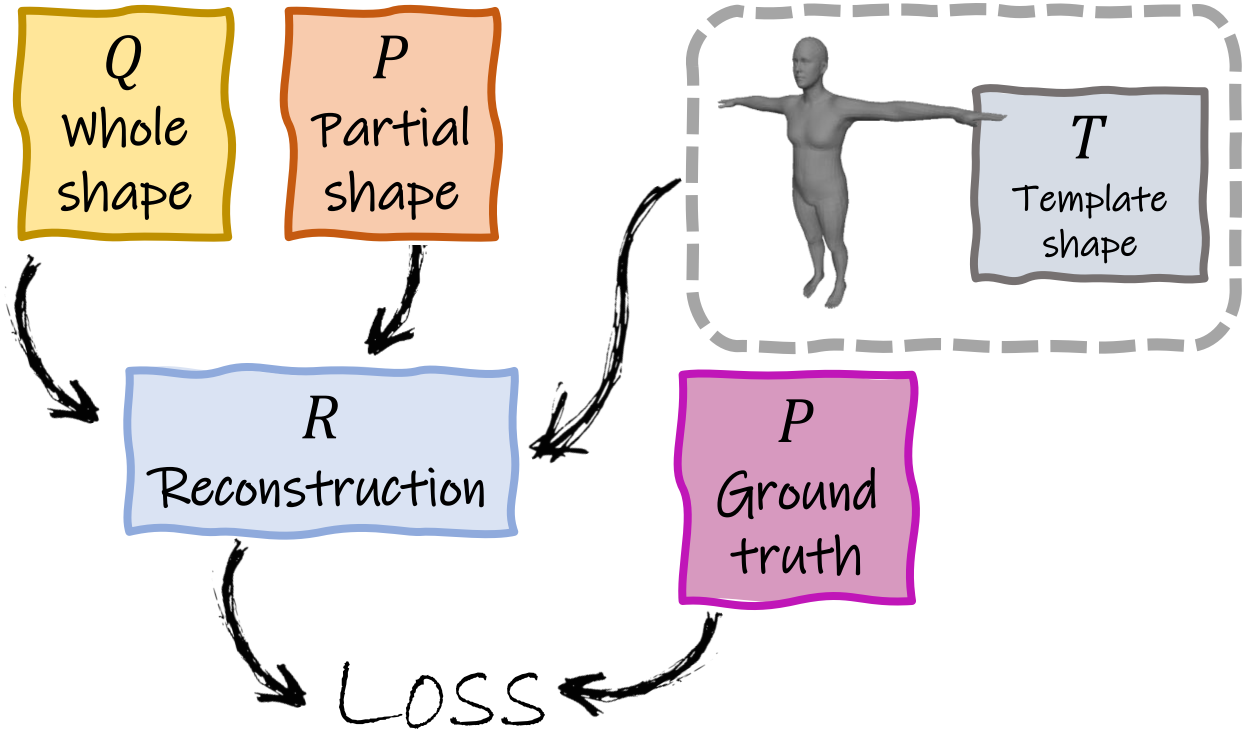
SMPL+H explorer

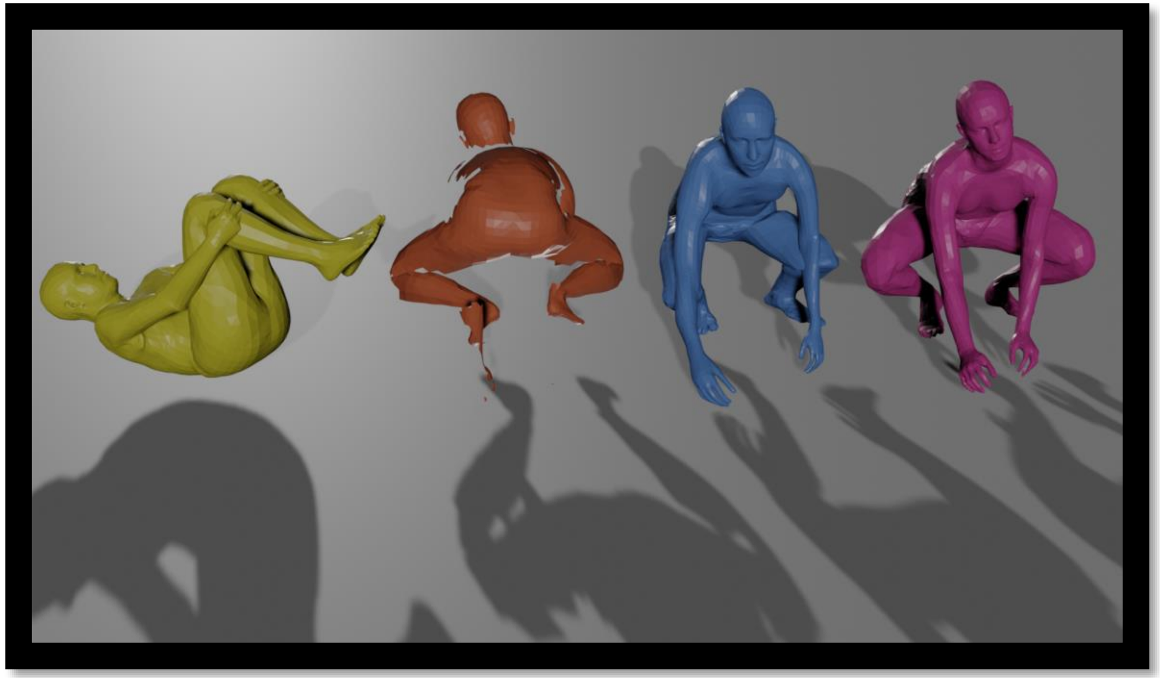
Scans source: (Right) Amass dataset. https://amass.is.tue.mpg.de, (Left): SMPL+H rest pose

Contributions

3D shape completion algorithm

2. We developed **state-of-the-art** 3D shape completion algorithm for shape completion from single partial view and another complete view in another pose. The algorithm significantly improve existing methods.





Scans source: Amass dataset. https://amass.is.tue.mpg.de

Contributions

FPS sampling methodology

3. We show a **new methodology** for choosing samples from large datasets that decrease the dataset bias and leads to **significant performance improvement** factor **of at least 1.5**, **across all cases**. It is the **first time** FPS sampling is implemented on the pose domain of human body parametric models. This method can be used in another machine-learning scenarios.

FPS result for $n = 4$ on actor from BMLmovi dataset

biased dataset

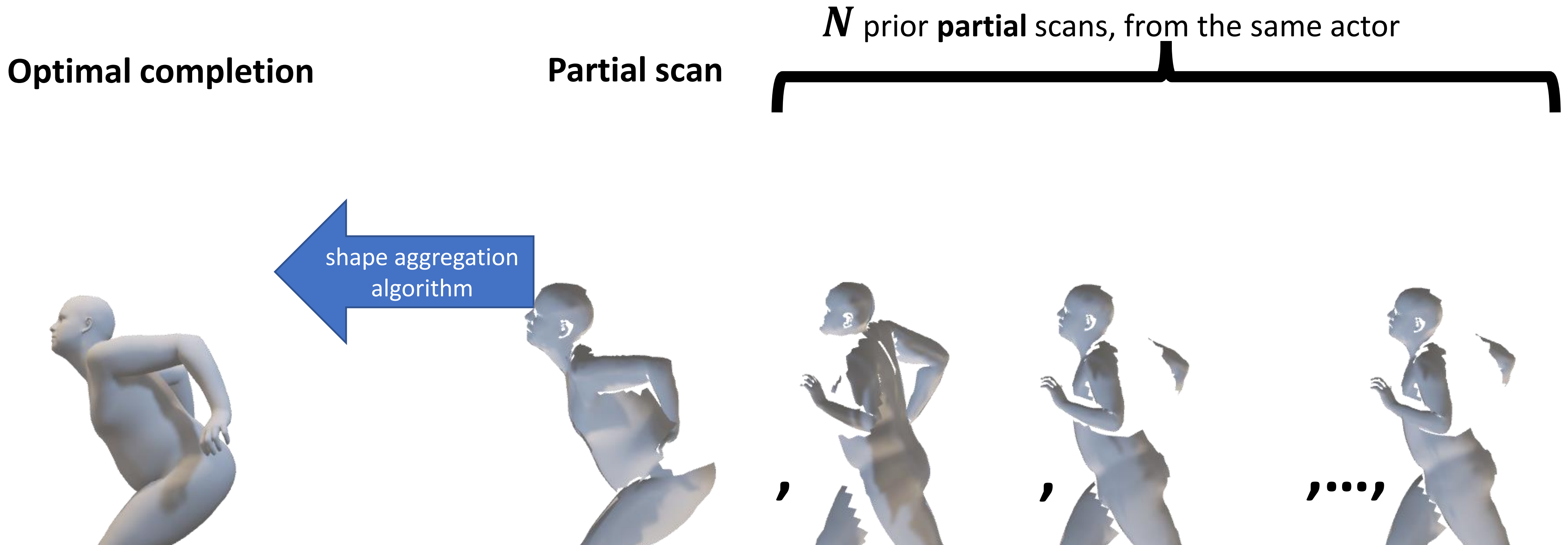
Unbiased dataset

Scans source: Amass dataset. https://amass.is.tue.mpg.de

Contributions

Non – rigid shape aggregation algorithm

4. We propose **new** architecture for shape completion from a single complete view and another set of multiple partial views.



Utilizing Prior Knowledge for Non-Rigid Shape Completion

Omer Ben-Hayun
Supervised by Ido Imanuel
Electrical Engineering Department
The Technion - Israel Institute of Technology
Haifa 32000, Israel

March 2, 2022

Abstract

In recent years, researchers have shown an increased interest in 3D human pose and shape estimation. Most studies in the field relies solely on completion from partial shape without additional information, resulting a limited models that cannot always reconstruct the partial shape precisely. The study utilized prior based approach for shape reconstruction of human partial scans that significantly improved the performance of existing methods. Additionally, in this study we developed and applied new technique for sampling from large datasets resulting solid increase of the performance across all tested learning models. The sampling methodology presented here has profound implications for future studies of machine-learning models that relies on learning from large datasets. Finally, we designed new visualization tools to explore the shape and the pose manifold of parametric body models and datasets.

Keywords: 3D shape completion, Non-rigid geometry, FPS sampling, Single View Reconstruction.

1 Introduction

In recent years, major advances in computational capabilities have arise a growing demand for creating and consuming 3D content. However, professional scanning devices are too expensive to be used for the typical user. As a result, acquisition

geometry of the original partial shape. Therefore, they often cannot be based solely on symmetry properties due to their non-rigid nature, and should be established upon more data in addition to the original scan.

[11] lists two basic approaches currently being adopted in research into shape completion of non-rigid shapes. One is generative based method and the second is alignment based method. Generative based approaches learn to approximate the class distribution and achieved impressive results in shape completion tasks. Yet, they suffer from notable methodological weaknesses, i.e. they are limited in that they only considers the partial shape during the completion time and does not take into account additional information that derived from the object. Hence, they failed to demonstrate generalization capabilities and cannot provide a accurate completion for unseen partial shapes. On the other hand, alignment based methods aiming to fit a complete shape to a partial shape. Since they exploit additional data during the inference time, they have potential advantage in terms of generalization and precise completions. However, current alignment based methods can carry only moderate partiality and considered to be slow.

This study set out to shine new light on shape completion tasks from several angles:

1. We introduce the design of visualization tools to explore the shape manifold of parametric body models.

2. We show a new methodology for choosing samples from large datasets that increase the performance of the learning process.

Future Work

- **FPS Sampling:** this sampling methodology could increase the performance of existing models in different scenarios: noise-reduction ,image classification, etc..
- **3D shape completion algorithm:** hyper-parameters tunning, improving the encoder-decoder architectures with newer point cloud learning shape encoders like PointNet++ .
- **Non – rigid shape aggregation algorithm:** improve the hyper-parameters , mainly $\left|\theta_{part}\right|$, $|\theta_A|$ in order to achieve better results.

MPI-Faust dataset http://faust.is.tue.mpg.de

Thanks