JUNE 18-22, 2023
CVPR VANCOUVER, CANADA

tcs Research

IIT Delhi
Indian Institute of Technology Delhi

Transfer4D: A framework for frugal motion capture and deformation transfer



Shubh Maheshwari[1], Rahul Narain[2], Ramya Hebbalaguppe[1,2]
[1]TCS Research, India [2]Indian Institute of Technology Delhi, India


weta
DIGITAL

Low-Cost Motion Capturing System-HuCE cvprLab

Large Production Houses:

- ✗ Expensive
- ✗ Multiple cameras/sensors
- ✗ Markers/templates
- ✗ System Calibration (multi-camera synchronization)

Low-Cost Motion Capturing System-HuCE cvprLab

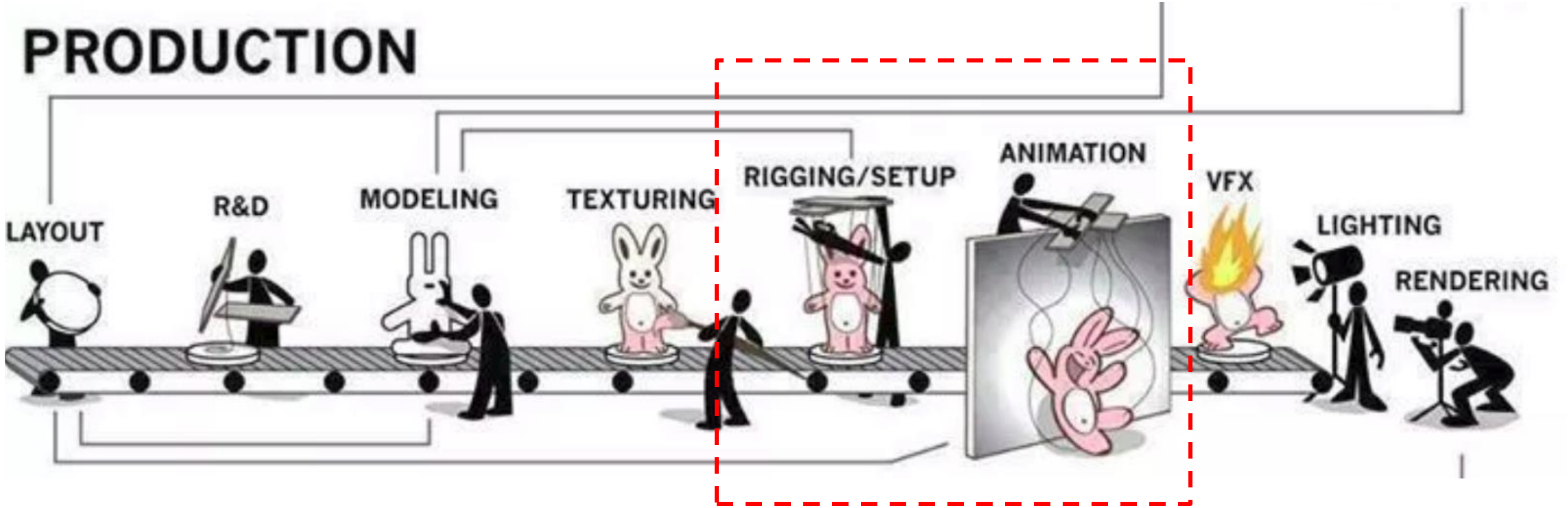
OcclusionFusion, CVPR 2020

Large Production Houses:

- ✗ Expensive
- ✗ Multiple cameras/sensors
- ✗ Markers/templates
- ✗ System Calibration

Ours:

- ✓ Frugal
- ✓ Single depth sensor
- ✓ No Markers/templates
- ✓ Minimal effort

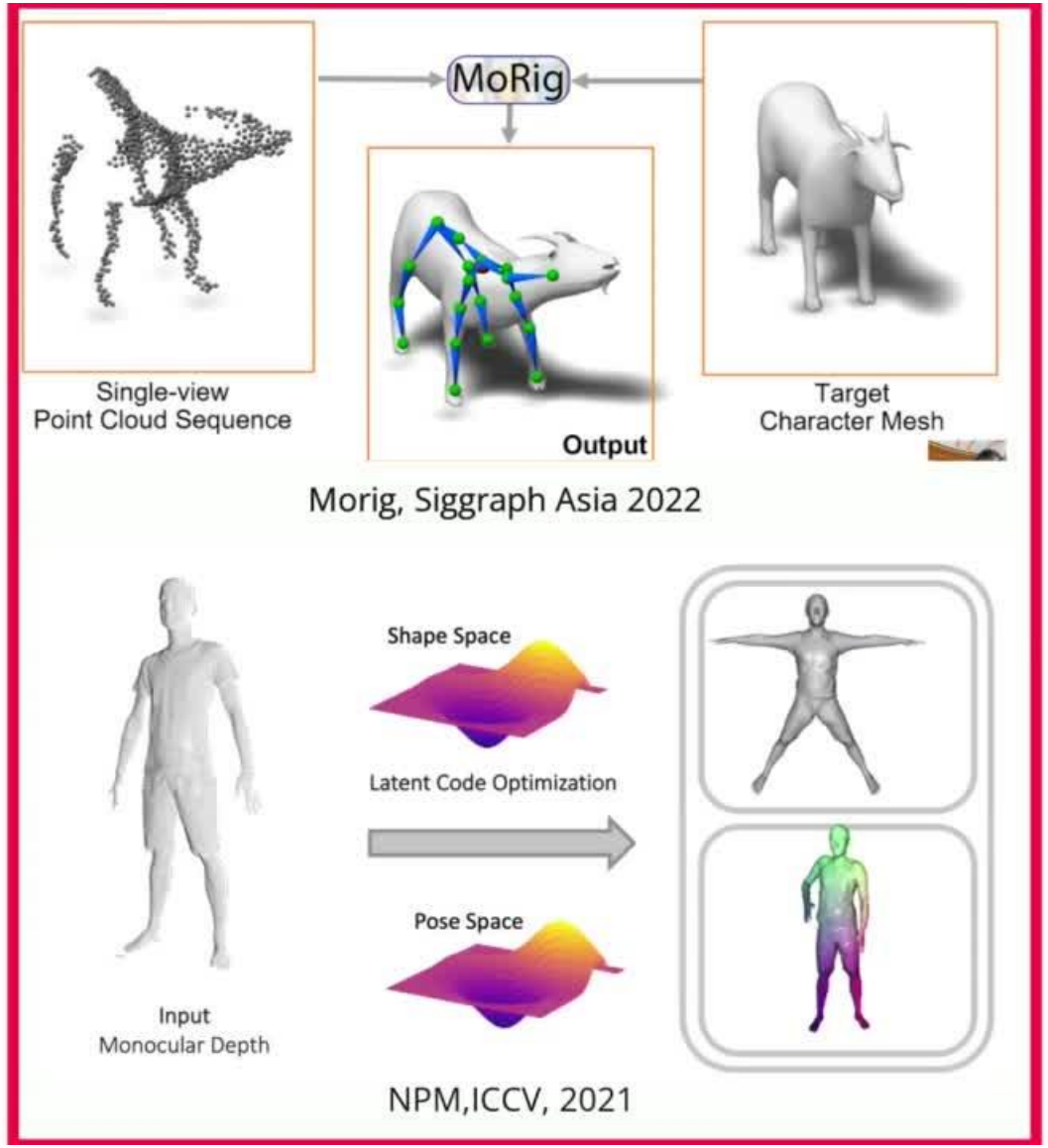


PROBLEM STATEMENT

Can we automate animators' effort in tedious and less creative components of animation?

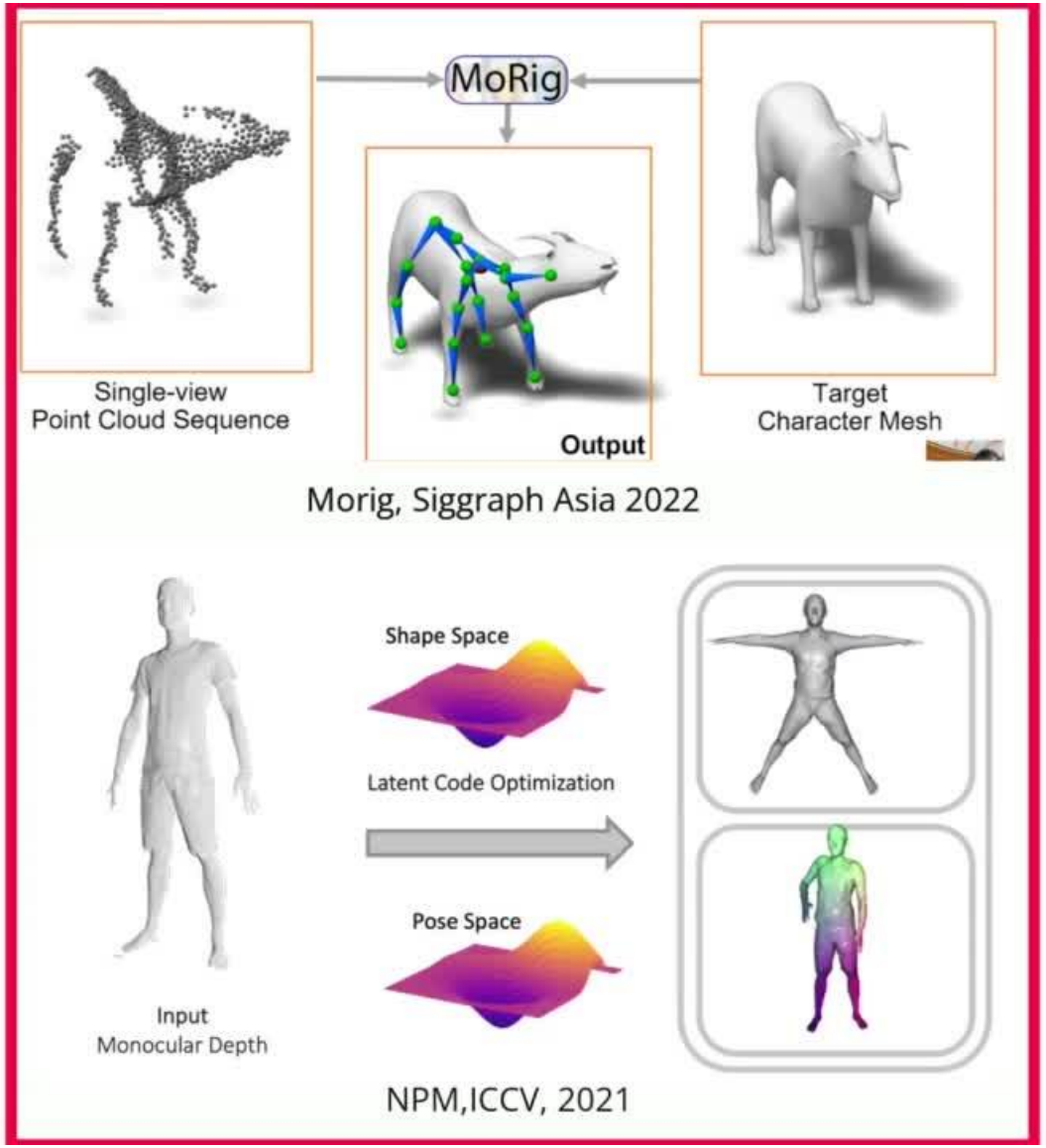
Image Source: Production Pipeline by Andrew Bean

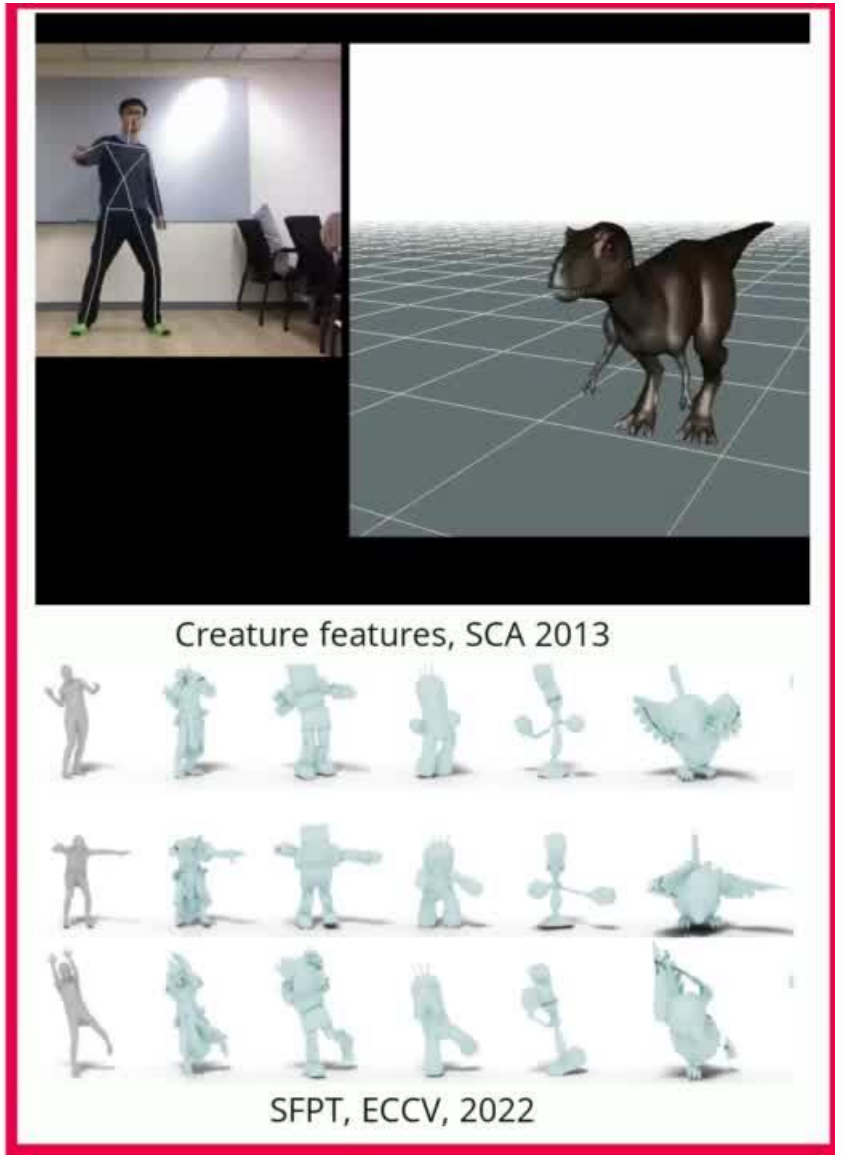
Related Works



✗ Require training supervision/custom dataset

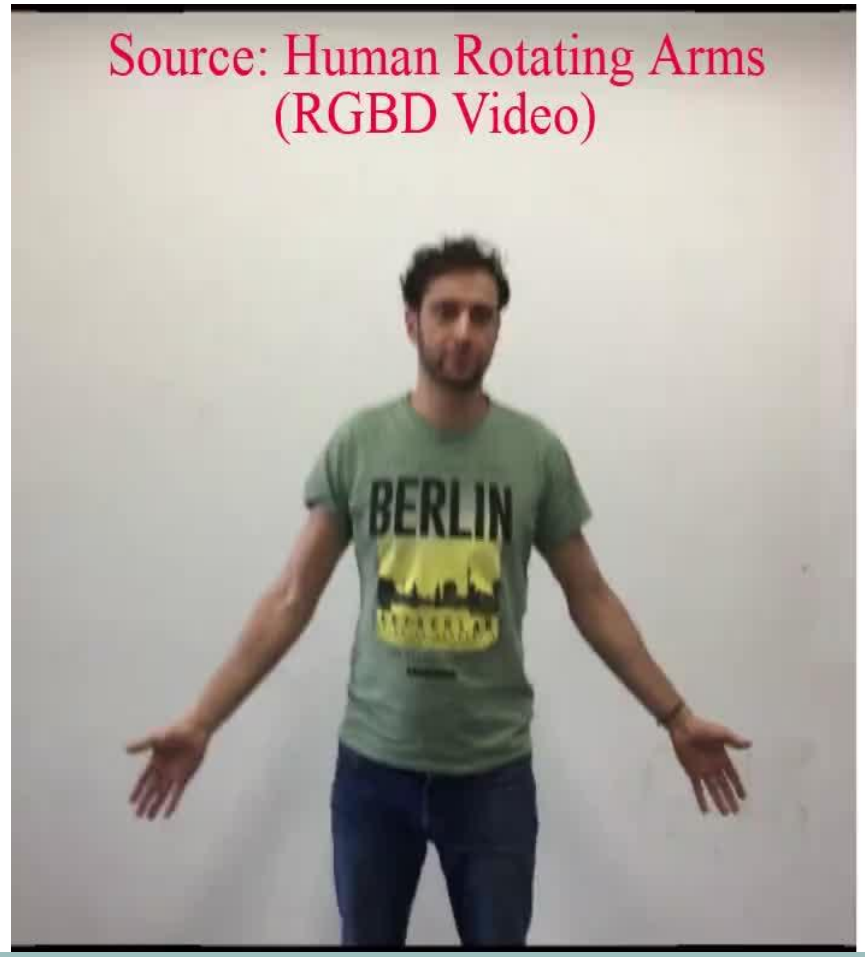
Related Works



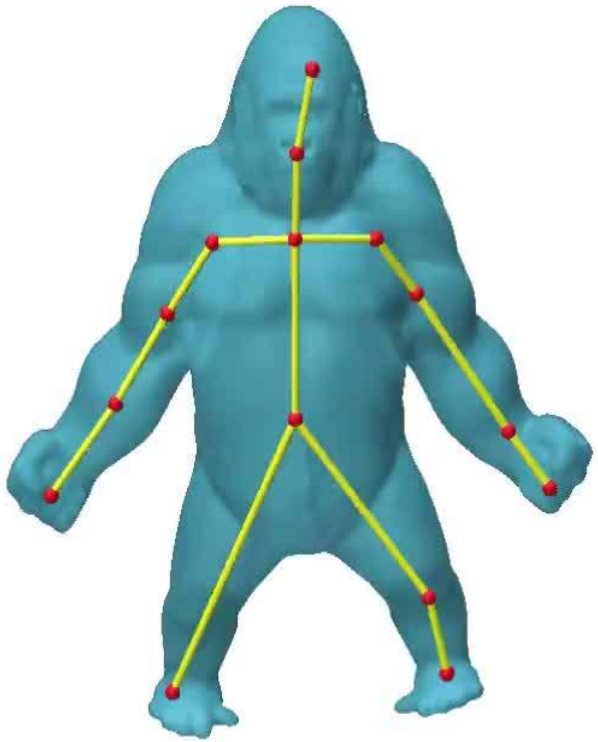


❌ Require training supervision/custom dataset

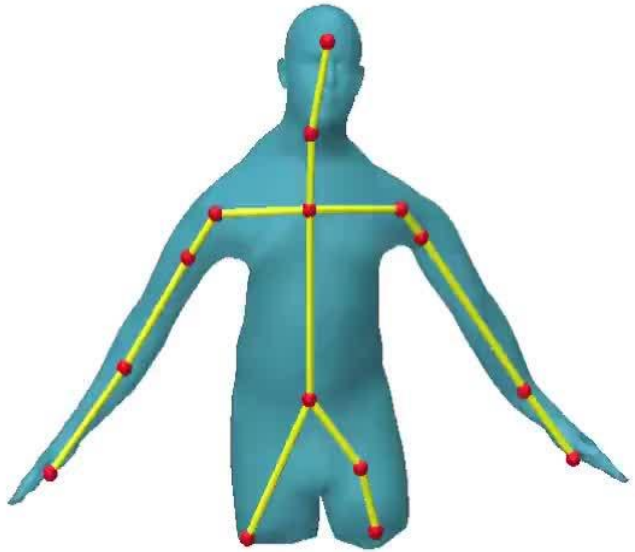
❌ Only transfer human motion



Target: Gorilla



Target: Human (Upper body)



Transfer4D

RESULT OF OUR UNSUPERVISED TECHNIQUE FOR INTRA-CATEGORY MOTION TRANSFER

Overview of Transfer4D

Transfer4D

CHALLENGES TO DEMOCRATIZATION VIA FRUGAL SENSORS

1. Input: A single-view video feed that implies partial information

- Could be noisy
- Sparse

Transfer4D

CHALLENGES TO DEMOCRATIZATION VIA FRUGAL SENSORS

1. Input: A single-view video feed that implies partial information

- Could be noisy
- Sparse

2. Output:

- Deformation of user-defined polygon mesh without a predefined skeletal rig

Transfer4D

CHALLENGES TO DEMOCRATIZATION VIA FRUGAL SENSORS

1. Input: A single-view video feed that implies partial information

- Could be noisy
- Sparse

2. Output:

- Deformation of user-defined polygon mesh without a predefined skeletal rig

Why this set-up?

The approach is frugal; it enables mass market reach and is more practical for deployment in future

Transfer4D

CHALLENGES TO DEMOCRATIZATION VIA FRUGAL SENSORS

- Shape matching between sparse source and complete target shape has not been explored extensively in literature for non-humans
- Prior works make a stricter assumption on the source to be noise-free and watertight.
- Our approach is category agnostic

Pipeline

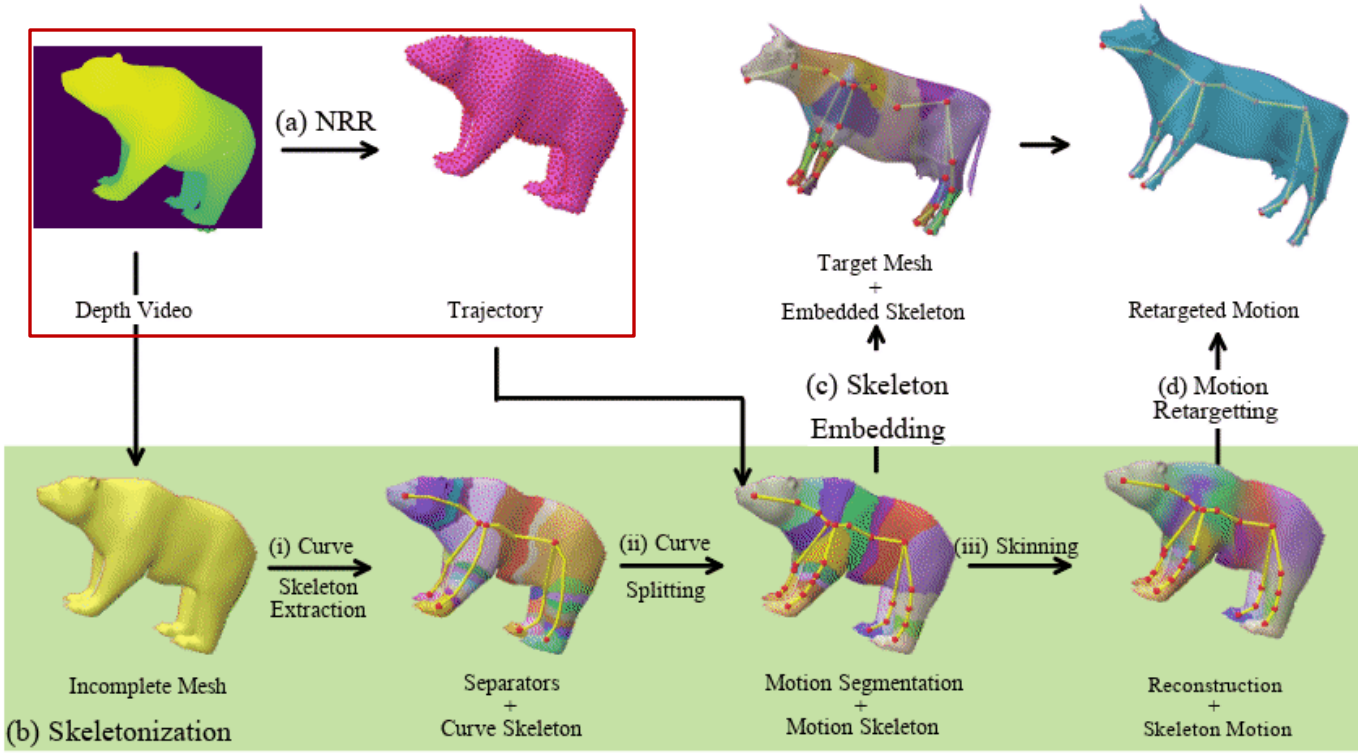
Transfer4D

OVERVIEW

Given the depth video,

(a) Non-rigid registration (NRR):

to align the source object at
a canonical frame to every frame,



Transfer4D

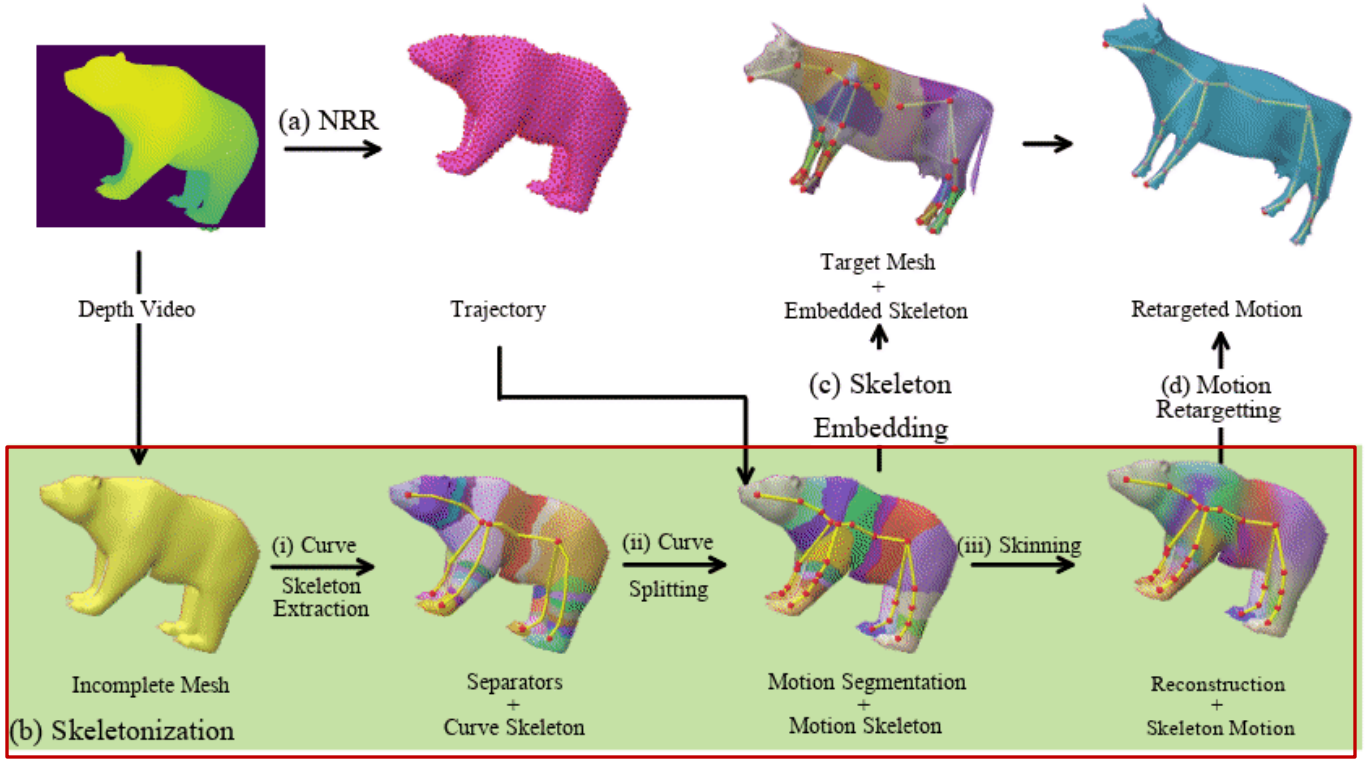
OVERVIEW

Given the depth video,

(a) Non-rigid registration (NRR):

to align the source object at
a canonical frame to every frame,

(b) Skeletonization: uses the geometry and estimated trajectory of the object to extract its underlying skeletal structure and the motion.



Transfer4D

OVERVIEW

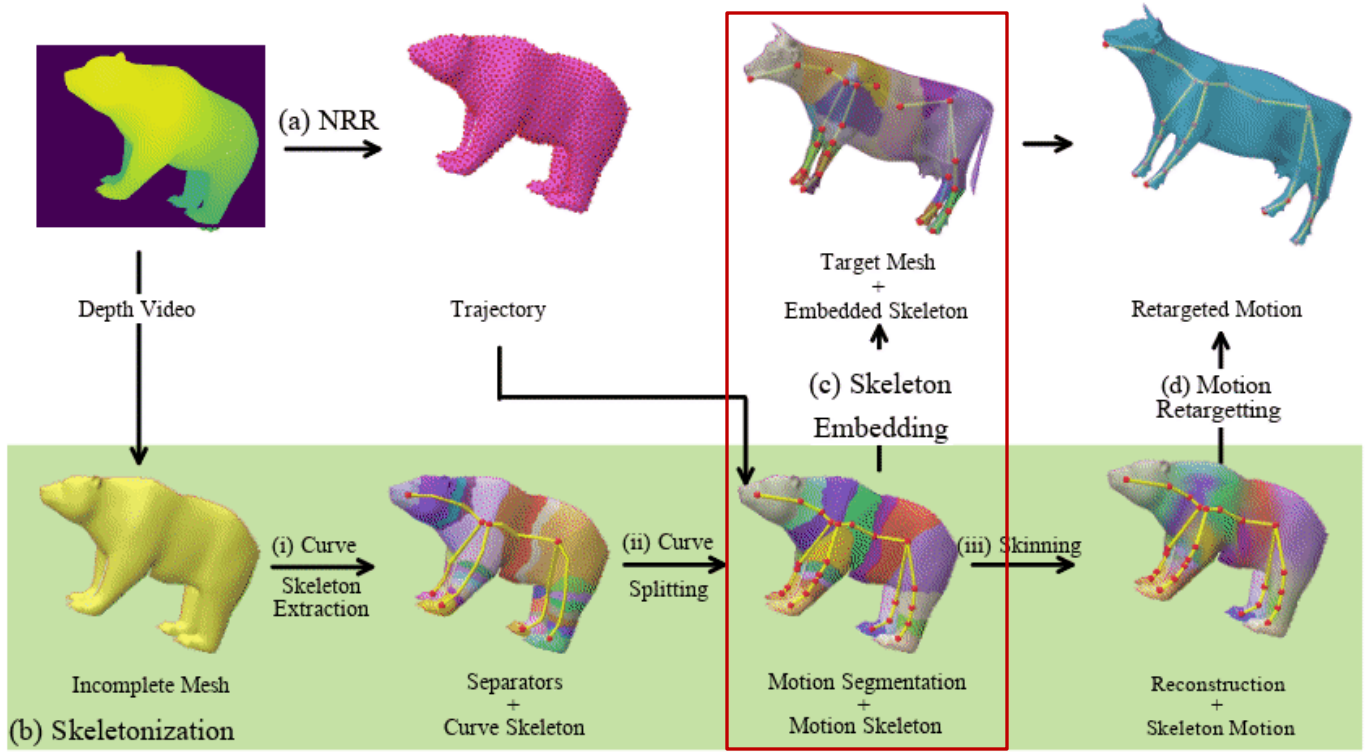
Given the depth video,

(a) Non-rigid registration (NRR):

to align the source object at
a canonical frame to every frame,

(b) Skeletonization: uses the geometry and estimated trajectory of the object to extract its underlying skeletal structure and the motion.

(c) Skeleton embedding: the obtained motion skeleton is embedded inside the target mesh and skinning weights are calculated.



Transfer4D

OVERVIEW

Given the depth video,

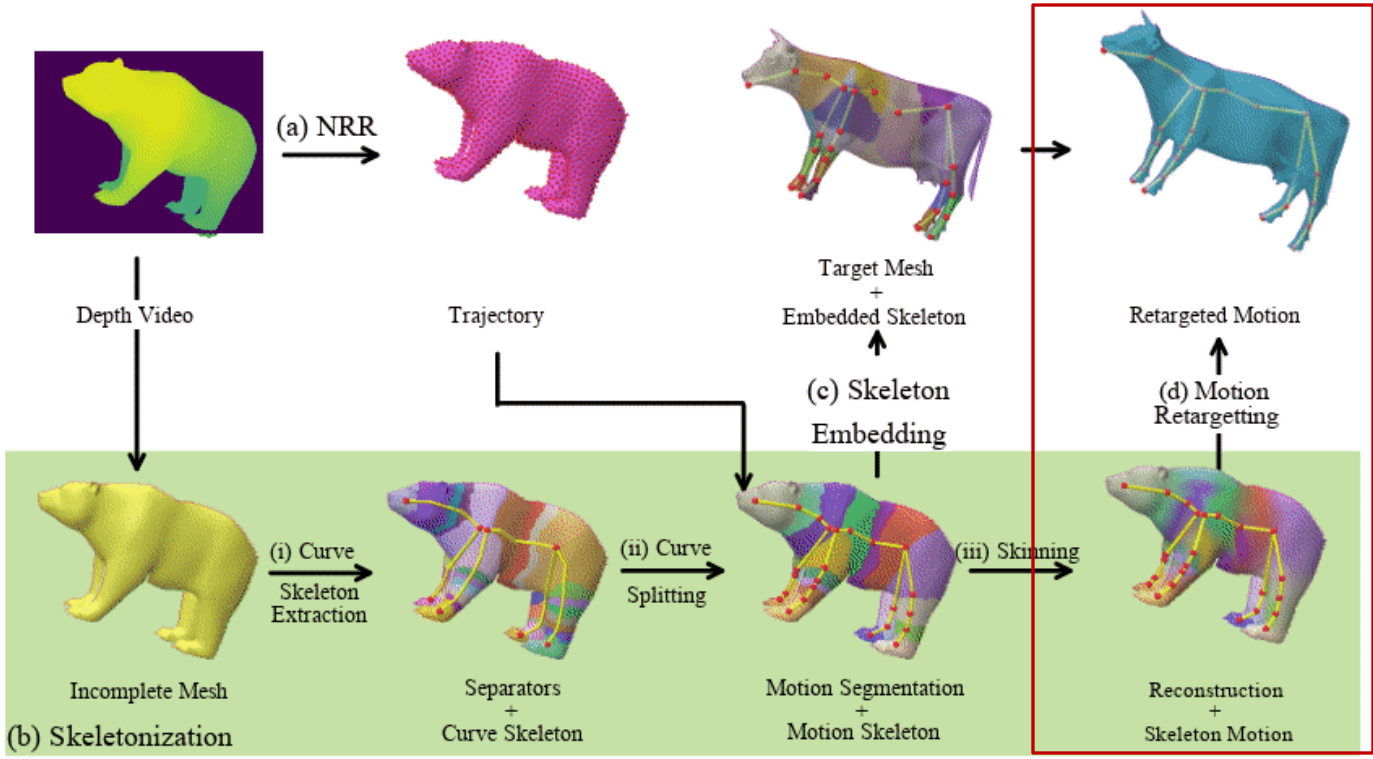
(a) Non-rigid registration (NRR):

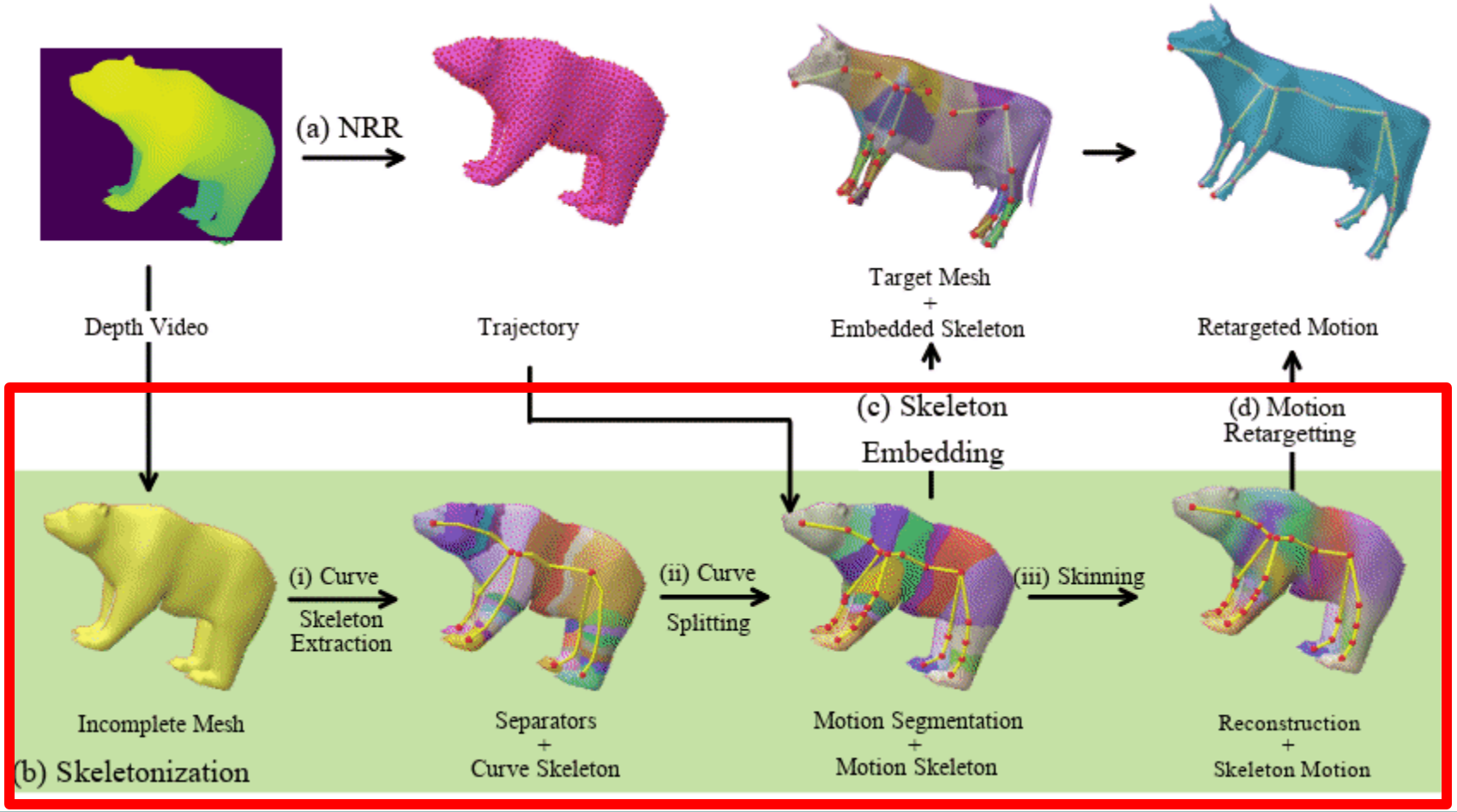
to align the source object at
a canonical frame to every frame,

(b) Skeletonization: uses the geometry and estimated trajectory of the object to extract its underlying skeletal structure and the motion.

(c) Skeleton embedding: the obtained motion skeleton is embedded inside the target mesh and skinning weights are calculated.

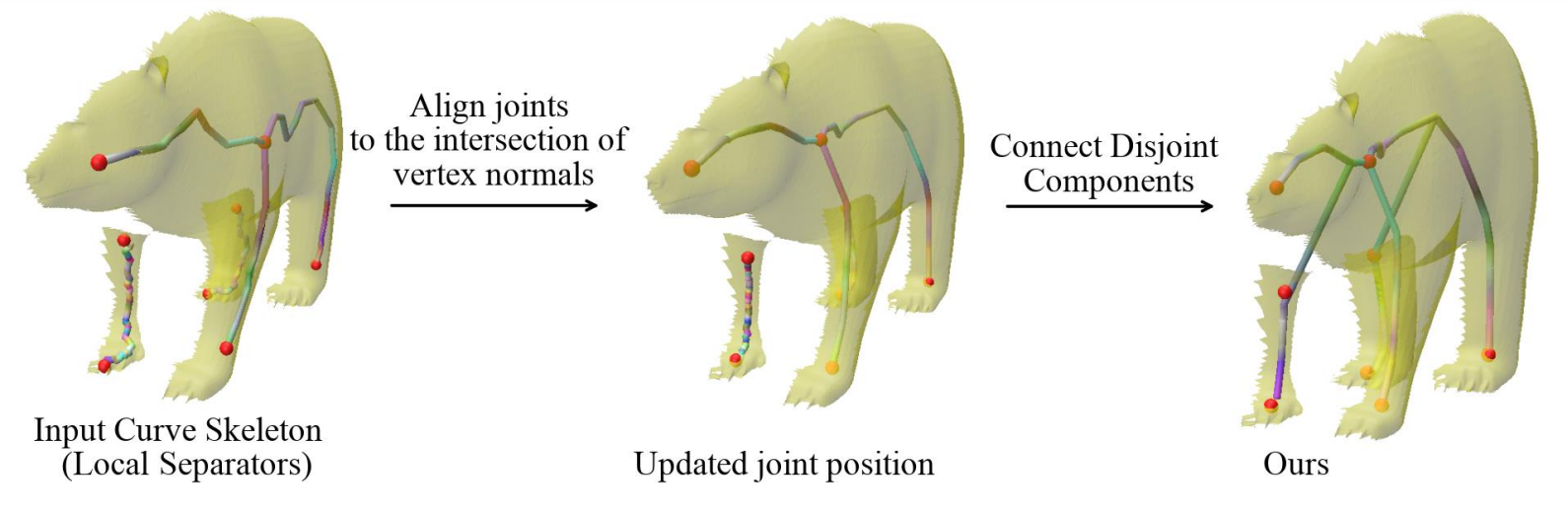
(d) Motion retargeting: the rotation of the bones is transferred from the motion skeleton to the target skeleton.





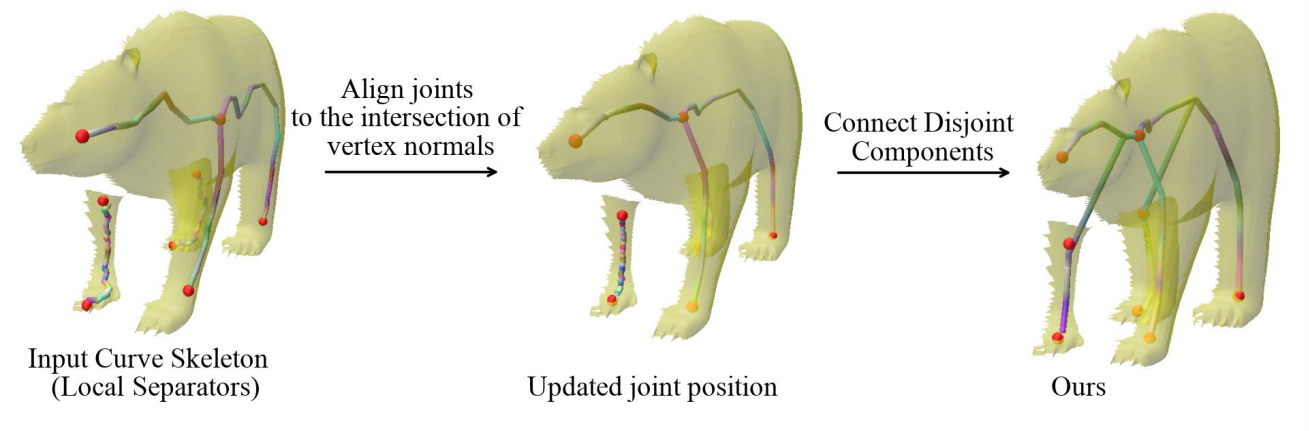
PROPOSED SKELETONIZATION PROCEDURE

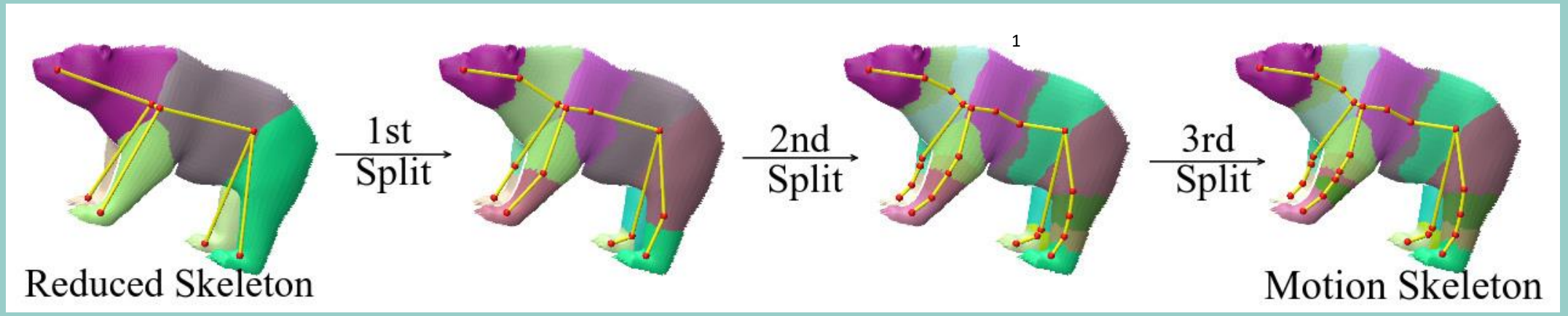
(i) Curve skeleton:
Notice that the extracted joint position from Local separators lies on the surface of the incomplete mesh. Our optimization aligns the joint position to the medial axis of the object.



[1] Skeletonization via Local Separators, ACM Transactions on Graphics

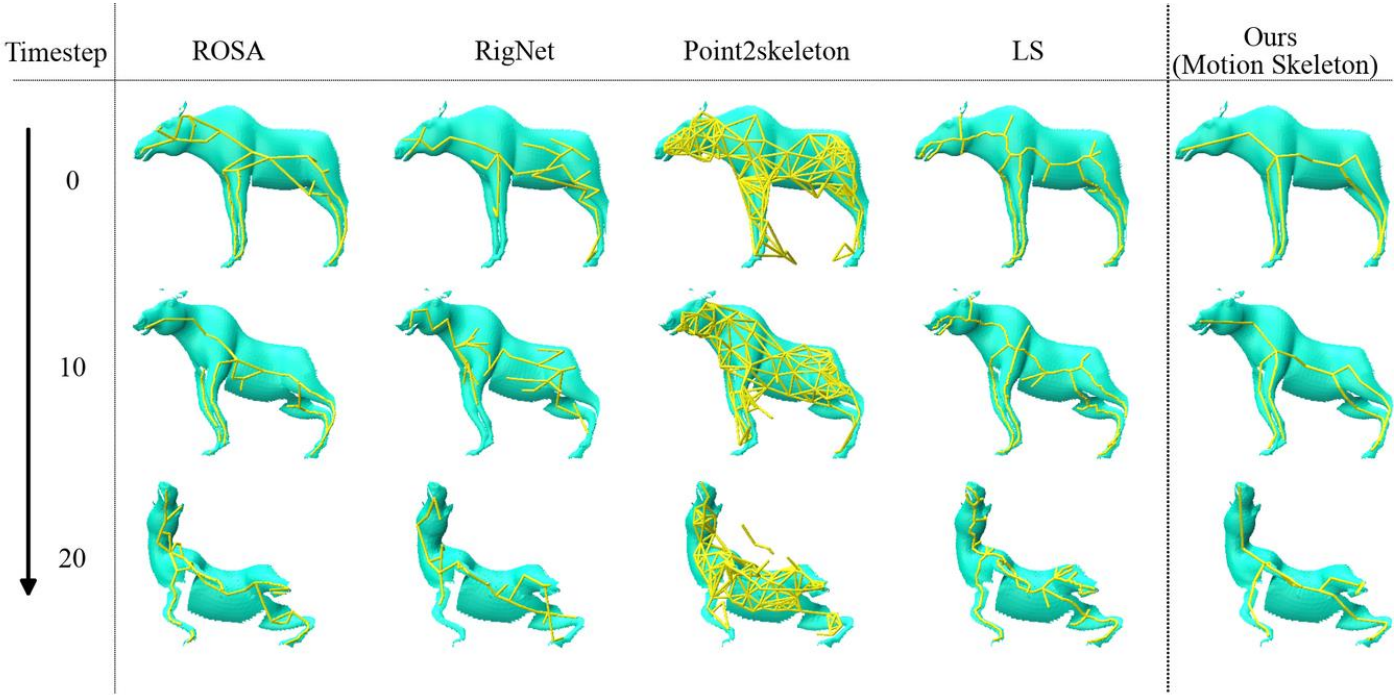
(i) Curve skeleton:
Notice that the extracted joint position from Local separators lies on the surface of the incomplete mesh. Our optimization aligns the joint position to the medial axis of the object.





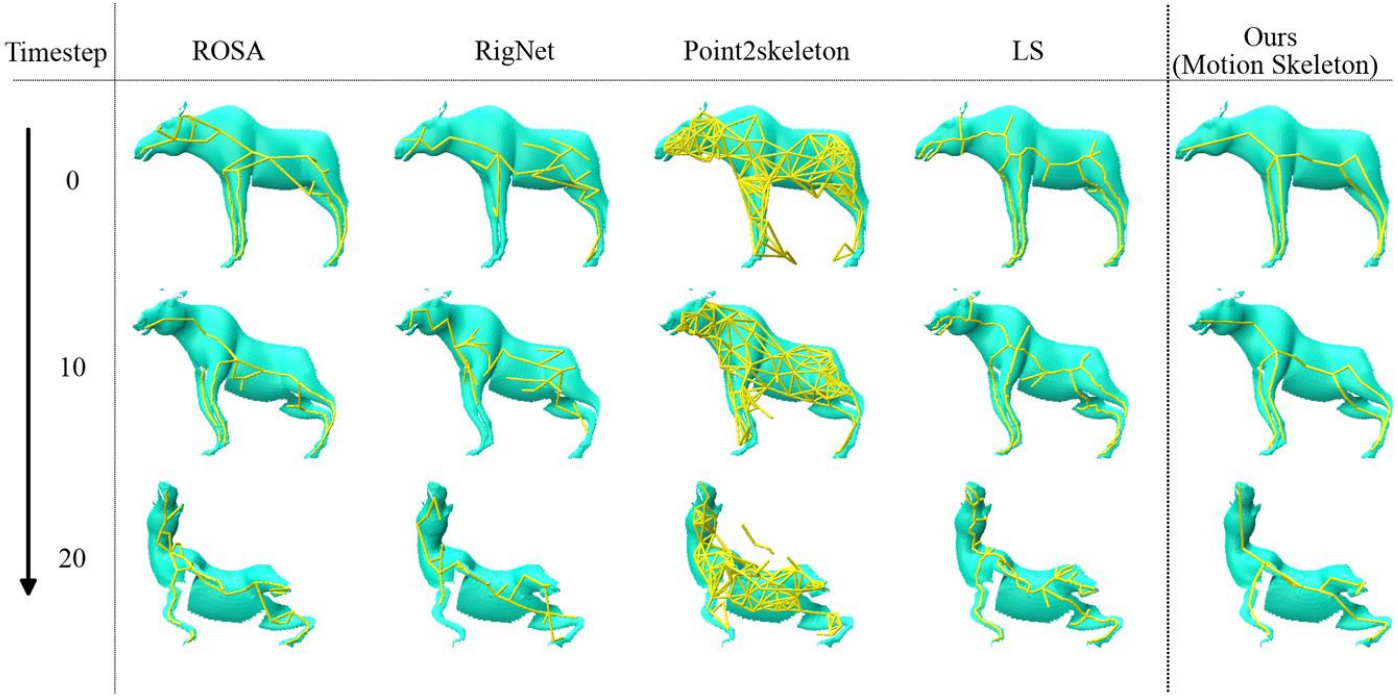
(ii) Motion Skeleton:
To incorporate the motion information, each curve of the extracted curve skeleton is split into multiple bones

Comparison

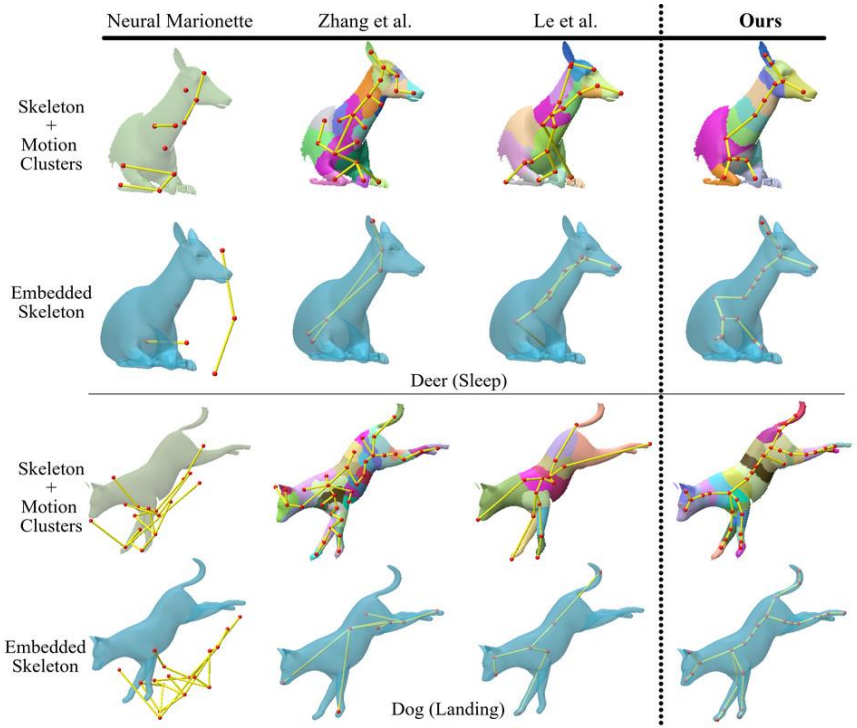


Compared to the skeleton extracted from static skeletonization methods, by utilizing motion information, ours method produces temporally coherent skeleton

Comparison



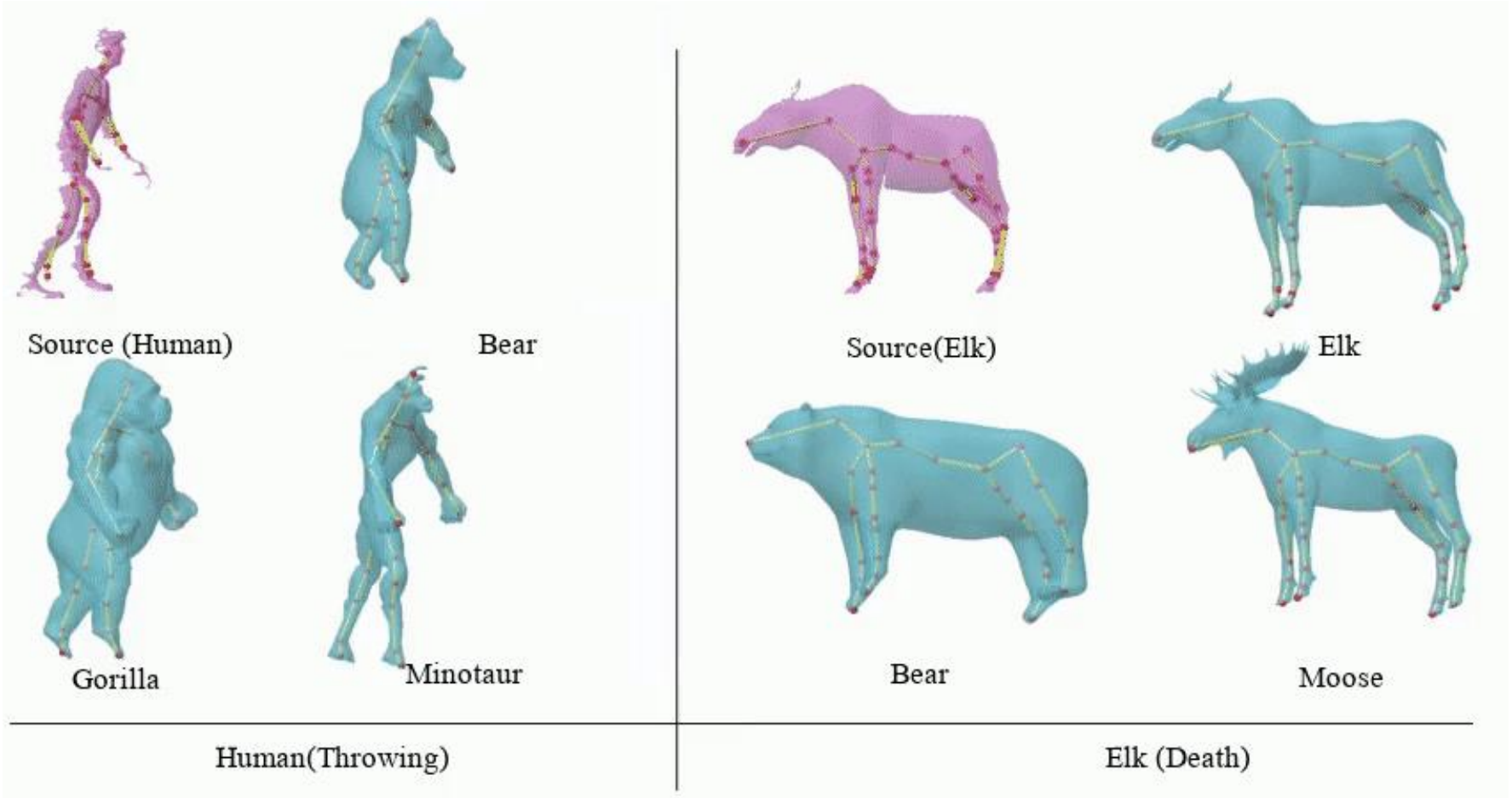
Compared to the skeleton extracted from static skeletonization methods, by utilizing motion information, ours method produces temporally coherent skeleton



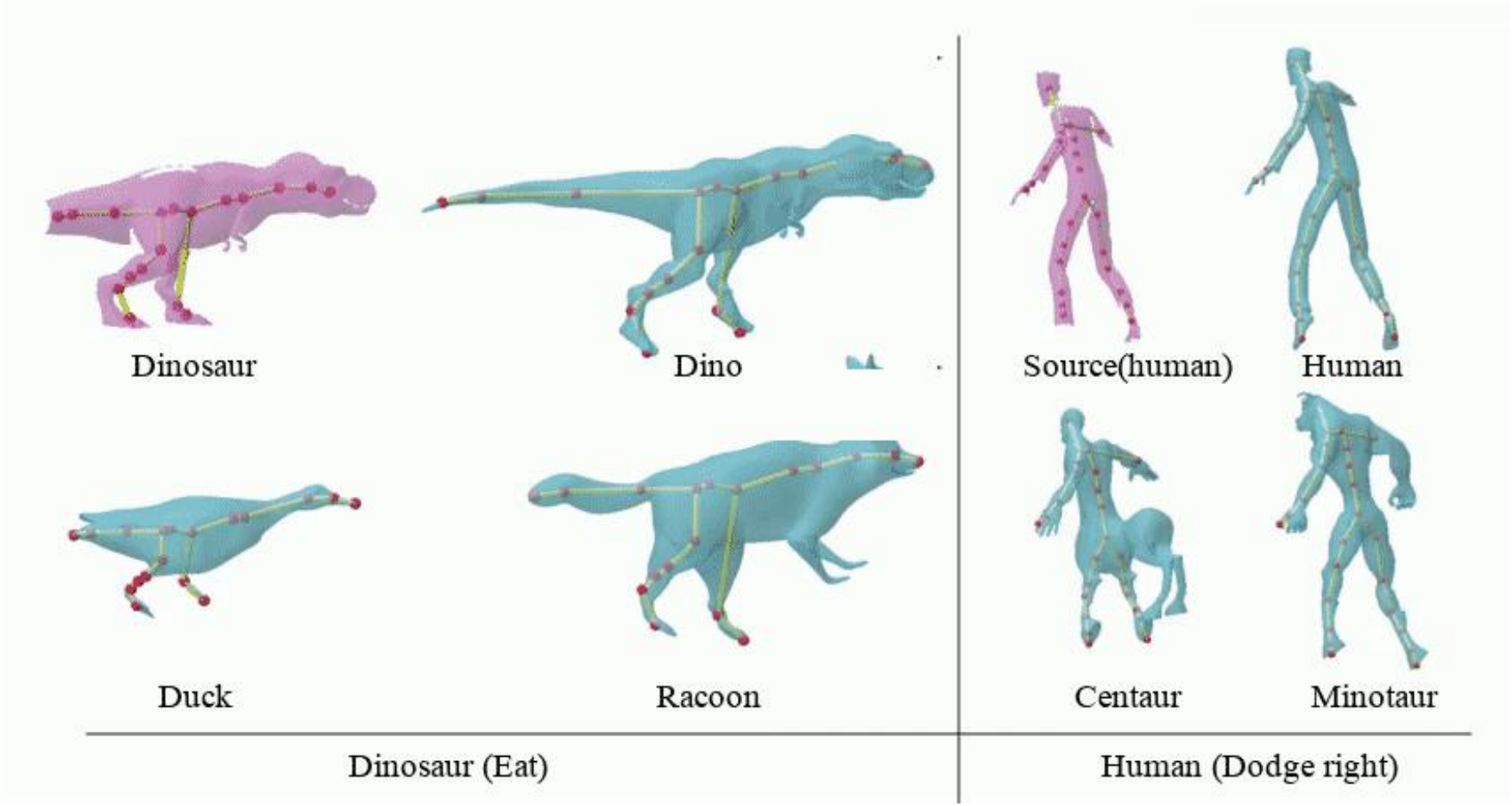
Compared to other motion skeletonization methods, by incorporating structural cues, ours is more effective at embedding the skeleton from incomplete mesh sequence.

Transfer4D

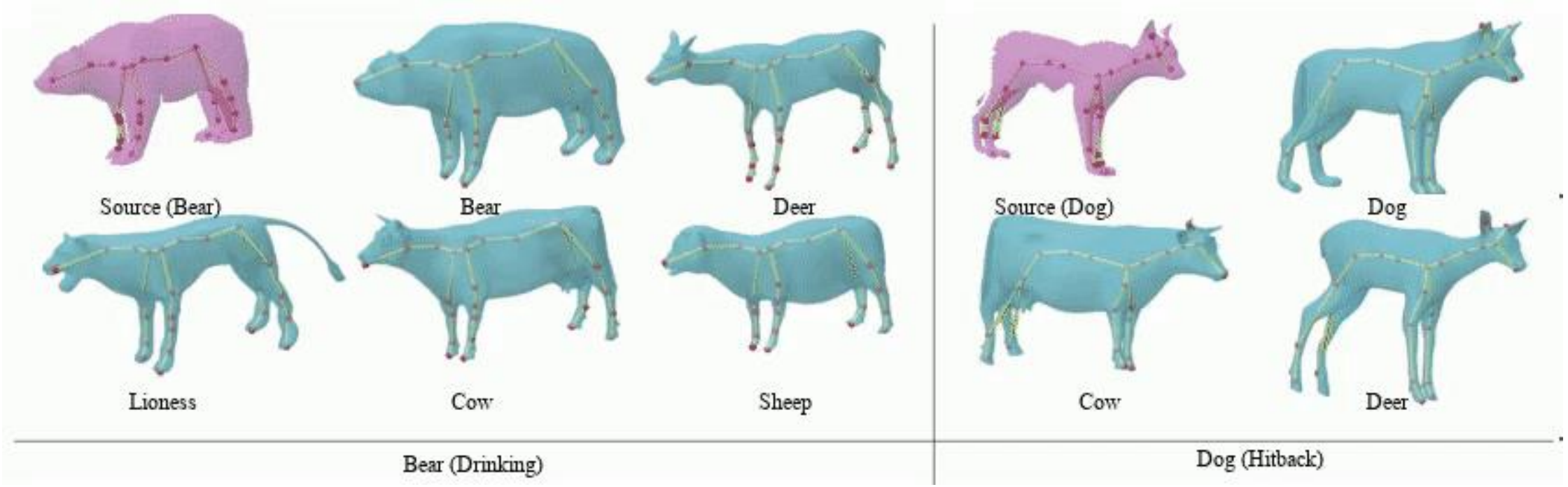
Results



ANIMATION TRANSFER



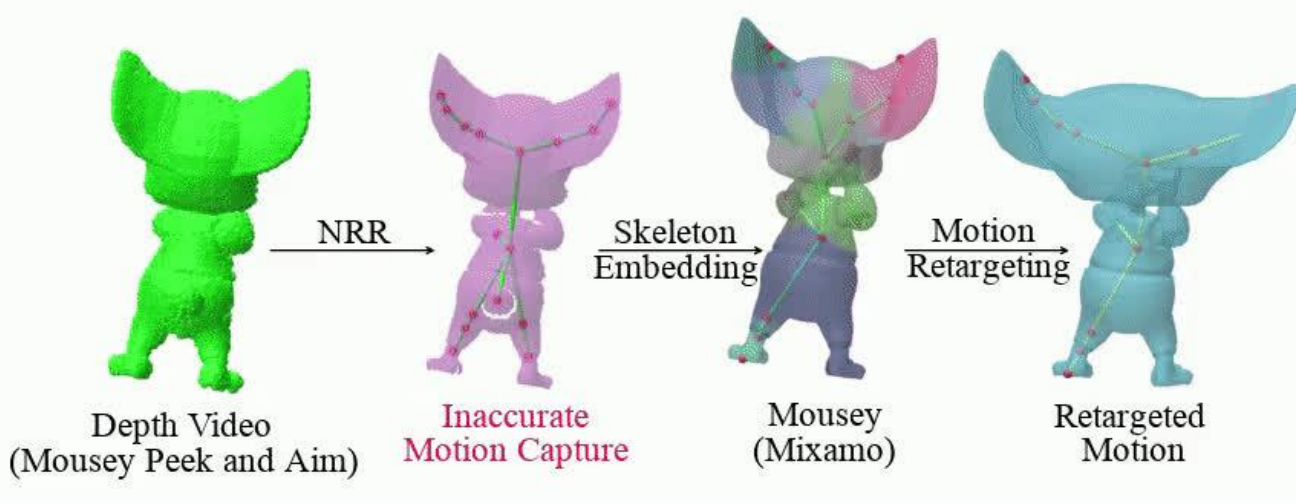
ANIMATION TRANSFER | BIPEDS

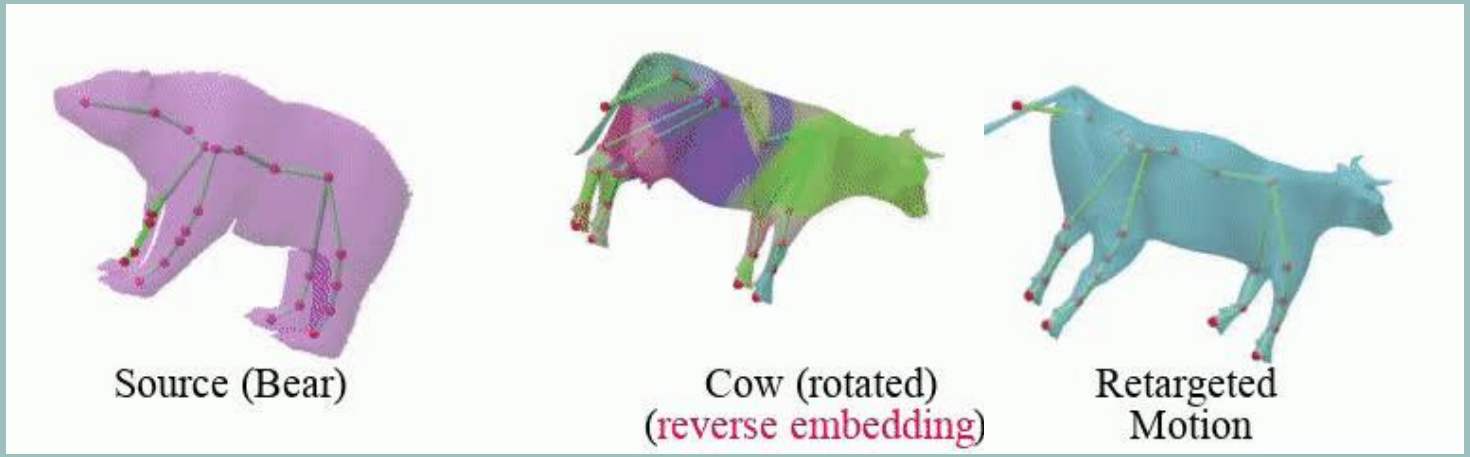


ANIMATION TRANSFER | QUADRUPEDS

Limitations

(1) Unable to capture large deformation between the source and target frame

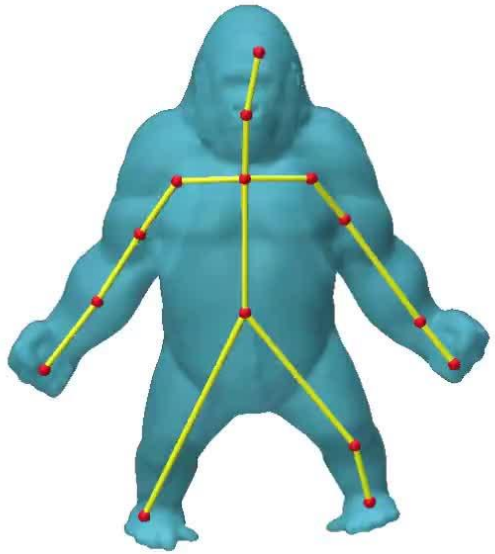




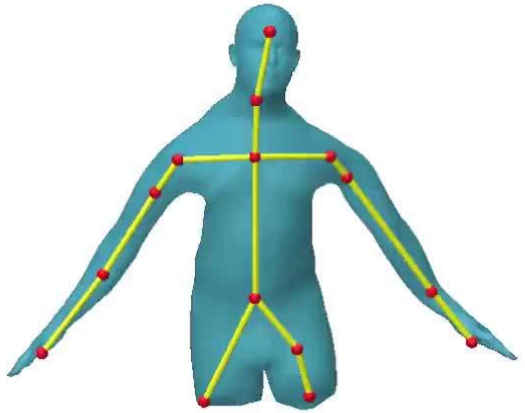
(2) source and target shapes should be in approximately the same pose

Conclusion

Target: Gorilla



Target: Human (Upper body)



Transfer4D is a frugal alternative that uses only commodity depth sensors and further reduces animators' effort by automating the rigging and animation transfer process