# Canine Mesh Recovery

Jake Deane

University of Bath

07-05-2019

# Table of Contents

# What is Animal Mesh Recovery? I

- The hypothesis of Canine Mesh Recovery (CMR) is based off the work by Kanazwa et al. [1] concerning Human Mesh Recovery (HMR).
- The goal of our end to end CMR framework is to construct a full 3D mesh of an canine body from a single RGB image.
- The overall framework can be seen in Figure 2 and the results can be seen in Figure 1.
- Though it is not perfect. For example, outlier physiques (e.g. very muscular) present problems for the framework.

# What is Animal Mesh Recovery? II

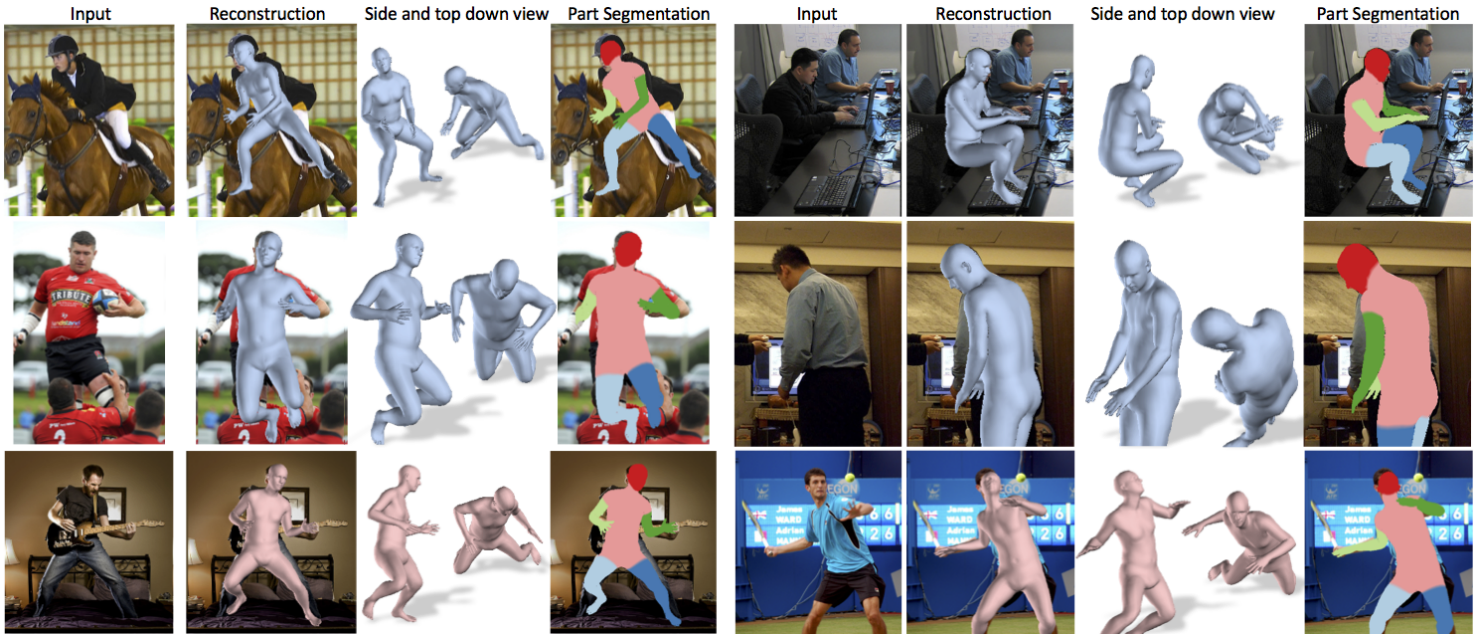


Figure 1: Results of the HMR framework

- The framework can be broken into three segments:
  - The Encoder - Transforms the high dimensional RGB image into smaller dimensional data that is fed into the Regressor.
  - The Iterative Regressor - Turns the encoded data into Pose (3D), Shape and Camera (scale, rotation and translation) parameters.
  - The Discriminator - Trained to tell whether the parameters from our own model correspond to a realistic canine body or not.

.

# Human Mesh Recovery I

- How can we adapt this framework to deal with quadrupeds (dogs, cats, mice e.t.c.)?
- In principal, this adaptation is quite simple.
- In practice however, this does not quite hold true.

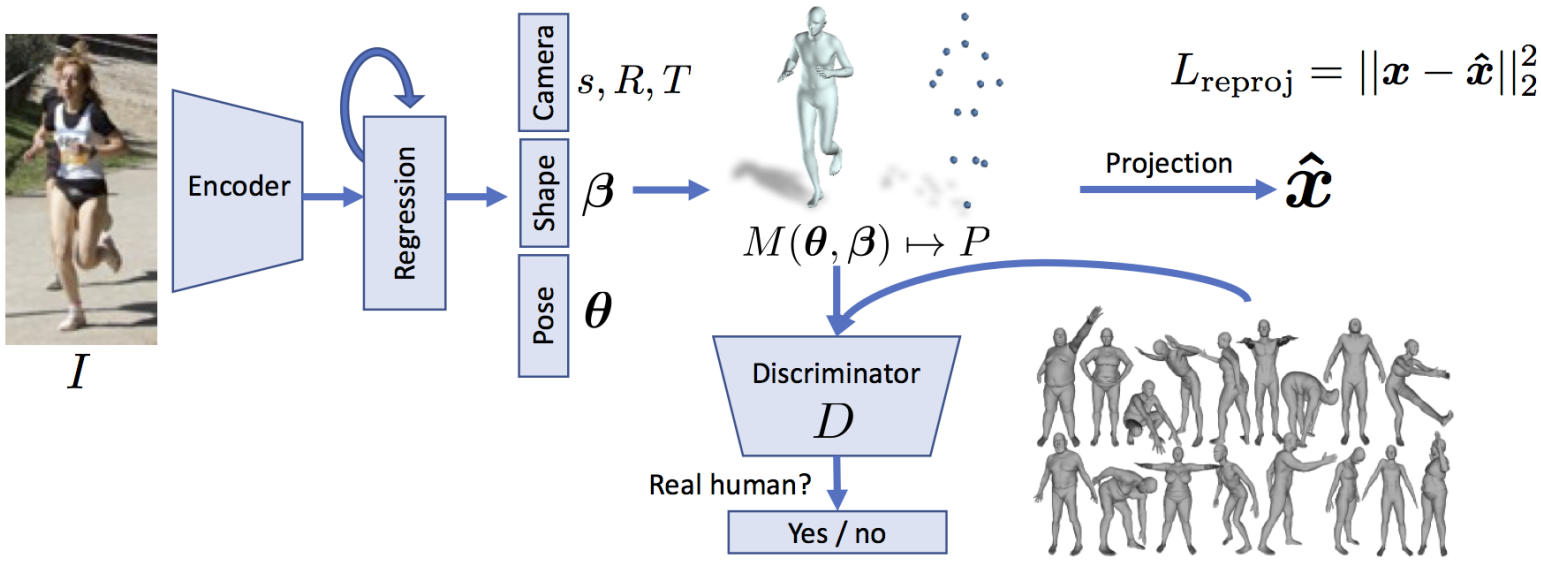


Figure 2: Overview of the HMR framework

# Problems We Face I

- There are three major problems with testing our hypothesis (along with minor ones).
- These two problems are:
  - Animal species have varying skeletal structures - see Figure 3.
  - The 2D joint coordinate datasets needed are either non-existent or unfit for purpose.
  - Similarly, 3D joint coordinate data is not readily available.
    - Both paired and unpaired 3D coordinates.


Figure 3: Very different body shapes

# How We Tackle These Problems - 2D

- We restrict ourselves to canines.
- This is done to avoid a too general framework that does not deliver accurate results.
- However, this does not reduce the problem entirely (see Figure 4).
- Still suffer from a lack of 2D/3D data.
- As a result, a dataset must be created.
- This has been done using a software called DeepLabCut[2][3].

Figure 4: Similar but still different body shapes

# How we tackle these problems - Other

- We hope that our lack of data can also be solved with synthetic/artificial data.
- This hypothesis is enforced by [4], where it has been shown that neural networks can be trained on synthetic datasets and produce equivalent results to real world data.

# What can this be used for? I

- Research
- Entertainment (e.g. video games)
- Animation
- Veterinary Science
- Zoological Science
- Security

# What's Next

- Complete the CMR framework.
- Increase the amount of 2D data to improve results.
- Obtaining paired 3D data to closely follow HMR[1].
- Extend the framework to RGB videos in order to produce moving graphical models in the vein of [5].
- Expand the framework to other quadrupeds.
- Create synthetic data to aid in training due to lack of real data available.

# Any Questions?

# References I

[1] Angjoo Kanazawa et al. “End-to-end Recovery of Human Shape and Pose”. In: *arXiv:1712.06584 [cs]* (18th Dec. 2017). arXiv: 1712.06584. URL: http://arxiv.org/abs/1712.06584 (visited on 17/05/2019).

[2] Alexander Mathis et al. “Markerless tracking of user-defined features with deep learning”. In: *arXiv:1804.03142 [cs, q-bio, stat]* (9th Apr. 2018). arXiv: 1804.03142. URL: http://arxiv.org/abs/1804.03142 (visited on 17/05/2019).

# References II

[3] Tanmay Nath et al. *Using DeepLabCut for 3D markerless pose estimation across species and behaviors*. preprint. Neuroscience, 24th Nov. 2018. DOI: 10.1101/476531. URL: http://biorxiv.org/lookup/doi/10.1101/476531 (visited on 16/04/2019).

[4] Gül Varol et al. "Learning from Synthetic Humans". In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (July 2017), pp. 4627–4635. DOI: 10.1109/CVPR.2017.492. arXiv: 1701.01370. URL: http://arxiv.org/abs/1701.01370 (visited on 16/05/2019).

# References III

[5] Angjoo Kanazawa et al. “Learning 3D Human Dynamics from Video”. In: *arXiv:1812.01601 [cs]* (4th Dec. 2018). arXiv: 1812.01601. URL: http://arxiv.org/abs/1812.01601 (visited on 17/05/2019).