# First Project update

## Proposed topic

Reconstructing 3D Human Mesh Model from a Single Wild Image


## Student Name: Ting Liu ( UIN:730008149 )


## Summary of work

### Completed work

1. Study papers about SMPL[1], ResNet[2], HMR[3].

2. Set up the environment using Anaconda3:

Tensorflow 2.4, Python 3.6, Other libraries: keras, opencv-python, numpy, h5py, ipykernel, matplotlib, scipy, tqdm, pyglet, etc.

3. Test and analysis codes of Resnet50V2 and SMPL method.

3.1 Resnet50V2

From tensorflow.keras.applications.resnet_v2 import ResNet50V2 with pre-trained weights on ImageNet dataset.

3.2 SMPL generator

Construct a generator by regressors and test with the pre-trained regressor model on COCO dataset.

3.2.1 Regression module

Resnet50V2 extracts features [1*2048] to generate SMPL pose_and_shape parameters (24*3 pose and 10 shape parameters) by a regressor. It consists of 2 fully-connected layers with 1024 neurons each with dropout layer ( Dropout(0.5) ) in between, followed by a final fc layer of 85D neurons.

3.2.2 Blend shape parameters

SMPL generation module converts pose_and_shape parameters to 3D vertices, joints and rotations of the human model. SMPL is based on linear blend skinning (LBS) models. There are 6890 vertices and K=23 joints (23 joints and 1 root orientation) in a human SMPL model. Each vertex $t_i$ in the mesh is linearly controlled by all joints with lbs weights $\omega_{k,i}$.

$$\bar{\mathbf{t}}'_i = \sum_{k=1}^{K} w_{k,i} G'_k(\vec{\theta}, J(\vec{\beta}))(\bar{\mathbf{t}}_i + \mathbf{b}_{S,i}(\vec{\beta}) + \mathbf{b}_{P,i}(\vec{\theta}))$$

$G_k$ is the world transformation of joint k. $G_k$' is the following transformation after removing $G_k$ due to the rest pose. $\theta$ is a pose vector, $\beta$ is a shape vector, and J is the collection of all pose vectors.

a) For shape blend shapes, $B_{S,i}$ presents different body shapes of people. It is used to infer shape dependent SMPL joints locations by applying smpl_joint_regressor (J_regressor in pretrained paired base_model).

b) For pose blend shapes, $B_{P,i}$ is used in preparation for the split pose. It adds pose features to SMPL vertices, producing reposed parameters [6890*3] to reshape joints_controlled vertices.

c) Optimization objective for shape and pose parameters is to minimize the vertex reconstruction error. Shape parameter training: PCA. Pose parameter training: gradient-based, non-negative least squares.

d) Deforming vertices in a mesh model by dual quaternion skinning from weighted joints ( lbs weights are loaded from pretrained paired base_model).

**Achieved results**

- Test Image(from COCO dataset):

- Resnet50V2 extracted features: pose (1, 72)

```
[[ 2.99229789e+00  2.32006446e-01 -2.65995324e-01  7.28881508e-02
   1.16436847e-01  2.08306819e-01 -4.53822553e-01 -5.32653481e-02
  -2.65590429e-01  3.11724305e-01 -1.14719130e-01  3.53369489e-02
   1.25984192e-01 -3.03998053e-01 -3.64122652e-02  5.81212878e-01
   7.08573163e-02  8.50350857e-02  2.01636460e-02  6.76949248e-02
  -7.10003972e-02 -1.43410251e-01  2.14893341e-01 -8.87283385e-02
  -2.24928170e-01 -3.06966990e-01  2.19144985e-01  5.71423545e-02
  -2.15562377e-02  2.09669545e-02 -1.61535442e-01  1.01694748e-01
   2.27218032e-01  4.42382321e-02  1.79629862e-01 -2.83684999e-01
   1.02624893e-01 -5.31504676e-02  2.47309972e-02 -3.05911452e-02
   1.38713658e-01 -3.28828335e-01 -1.45027369e-01 -9.00704861e-02
  -1.97478533e-02  2.45430976e-01 -1.50686502e-03  3.01675275e-02
  -7.75373131e-02 -3.75955105e-02 -3.25453818e-01 -2.77324378e-01
   4.18180376e-02  6.98037744e-02 -1.90767422e-02 -6.05258346e-01
   1.14554919e-01 -4.92527783e-01  2.32917756e-01  1.33113265e-02
  -1.79201871e-01 -3.84014547e-02  5.85794225e-02  1.23685651e-01
  -1.20442621e-02 -3.81480977e-02 -1.30594328e-01 -4.52433154e-03
  -7.11707696e-02 -8.62428397e-02  4.39000130e-02  7.35650957e-02]]
```

- Resnet50V2 extracted features: shape (1, 10)

```
[[ 0.28703701  0.04517686 -0.8366228  -0.06514511  0.0571376   0.13152993
   0.2088913  -0.11868145 -0.34755307  0.29291362]]
```

**Unexpected complications**

1. Tensorflow 2.6 doesn't support Mac M1 chip, exports 'Illegal hardware instruction' error.

2. Tensorflow 2.6 for Windows can cause 'Value error / InvalidArgumentError' in excuting ResNet50V2 module from keras.

Solution: use tensorflow 2.4 on Windows.

## Plan for completion

1. Learning the Dual Quaternion Skinning method (SMPL model deformation approach): how shape and pose parameters affect the mesh surface vertices' coordinates (6890*3).

2. Add a discriminator to optimize shape and pose parameters based on adversarial learning. By 11/18, second update report.

3. Do training and testing on different datasets, parameter adjustment. By 12/7, final report and project presentation.

REFERENCES

[1] Loper, Matthew, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. "SMPL: A skinned multi-person linear model." ACM transactions on graphics (TOG) 34, no. 6 (2015): 1-16.

[2] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Identity mappings in deep residual networks." In European conference on computer vision, pp. 630-645. Springer, Cham, 2016.

[3] Kanazawa, Angjoo, Michael J. Black, David W. Jacobs, and Jitendra Malik. "End-to-end recovery of human shape and pose." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7122-7131. 2018.