

# TailorNet 模型

## 1. 简介

TailorNet 是首个同时利用模型身体姿势、模型身体形状和衣服风格（衣服本身的几何形状）预测衣服 3D 形变的 3D 衣物模型。核心技术是 Mesh 频率分解（mesh frequency decomposition），即把 mesh 形变分解为高频部分和低频部分：低频部分模拟大致形变，用一个以 pose，shape 和 style 参数的 MLP 来预测形变效果；高频部分突出衣服特性，用一个混合各种特定 shape-style（比如 t-shirt，裙子等）的 pose 模型来预测模拟效果，每个 style-shape 模型负责模拟一种特定形变，相似效果加权求和后得到最终预测的效果并生成新蒙皮覆盖到原模型上。Style 的变化是通过 PCA 计算标准姿势中的形变子空间得到，这个子空间满足某些物理约束，比如与身体的穿透或者覆盖在身体表面。

## 2. 数据、输入和输出

官方提供了多种衣物的原始数据集，男装女装均有，包括：t-shirt, 衬衫，长裤，短裙，并提供一个训练好的模型，可以直接使用，也可以按照文档指引生成自己的数据或者训练自己的模型。

模型的输入是一组以 SMPL 模型为基础的动作，且衣服节点已经集成到模型中，成为模型整体 mesh 的一部分。模型需要包含骨骼数据，蒙皮顶点数据等，以 obj 或者类似格式引入。

模型的输出是包含被修改的蒙皮的模型的一组 ply 文件，包含模型的面数、顶点、颜色等等。

## 3. 衣物模型

### 3.1 参数：身体姿势 $\theta$

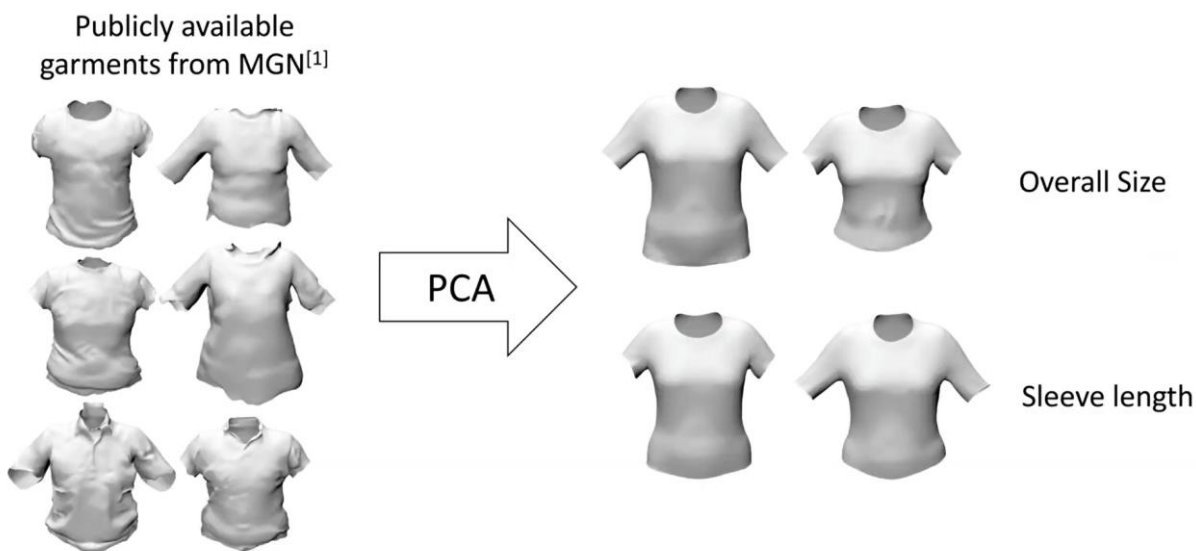
模型的身体姿势（Pose）是动画中的骨骼节点数据，提供了肢体的位置和走向等信息，影响衣服的整体布局和形变方向。

## 3.2 参数：身体形状 $\beta$

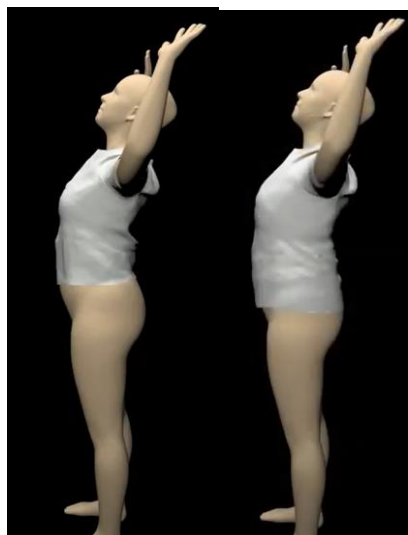
模型的身体形状 (Shape) 是动画中的蒙皮数据, 提供的肢体的半径, 人物胖瘦等信息, 影响衣服的形变程度, 比如是否贴身、褶皱程度等等。

## 3.3 参数：衣服风格 $\gamma$

衣服风格 (style) 是从衣物的真实数据中通过 PCA 提炼出来的衣服特性的特征矩阵。首先得到不同种类衣物在人物标准姿势和标准形状上的顶点数据, 执行 PCA 把数据集分解为由若干特征向量组成的特征矩阵, 每个特征向量可能对应一个或者多个衣物的特性, 比如紧身度、衣长、袖长等。



样例 1：相同 Pose、style 在不同 shape 的形变效果

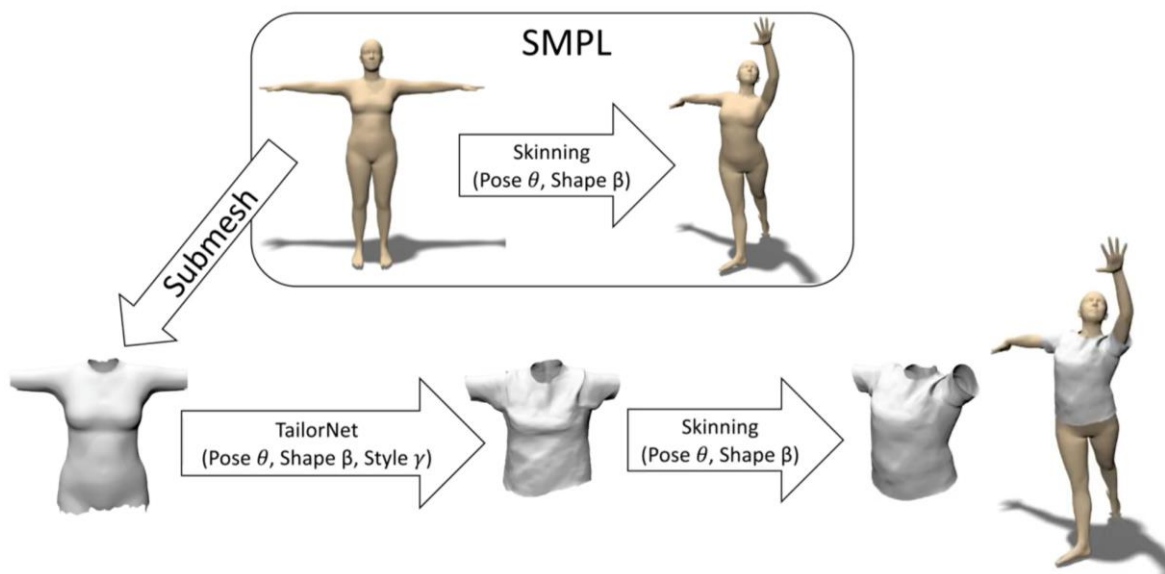


样例 2：相同 Pose、shape 在不同 style 的形变效果

### 3.4 工作流

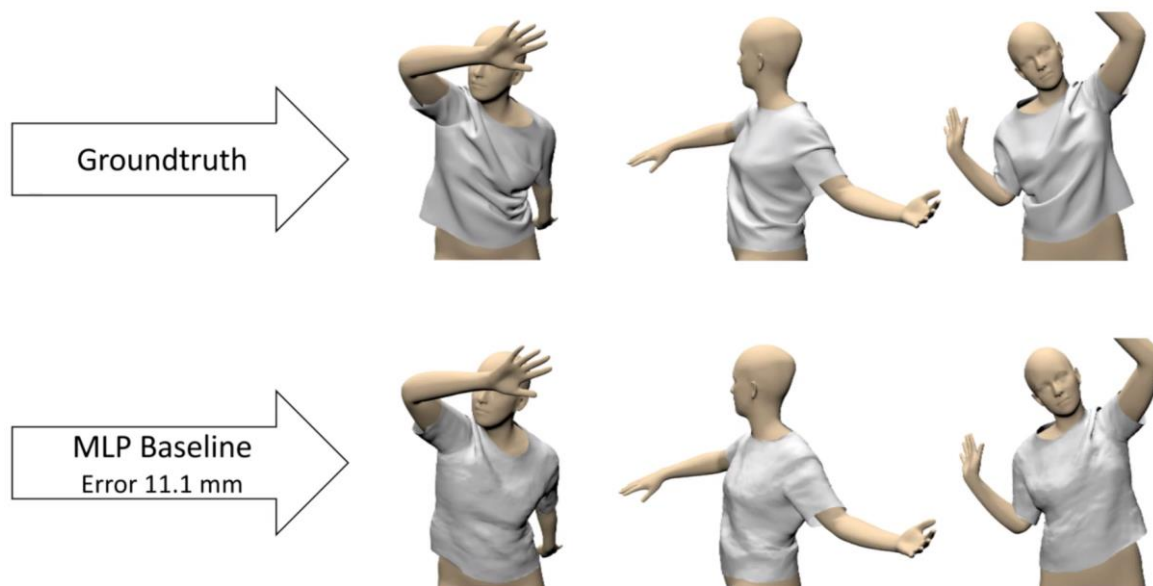
对于一段动画中的某个动作，衣服的模拟效果计算过程如下：

- (1) 提取 SMPL 身体模板的一个子 mesh 作为衣服模板，并获取骨骼结构  $\theta$  和蒙皮的权重  $\beta$  作为数据。
- (2) 以上述数据加上模拟的目标衣服的风格数据  $\gamma$  作为输入，调用训练好的 TailorNet 模型，重新生成蒙皮数据 ( $\theta'$ ,  $\beta'$ ) 作为输出。
- (3) 把新蒙皮数据重新绑在动画上。



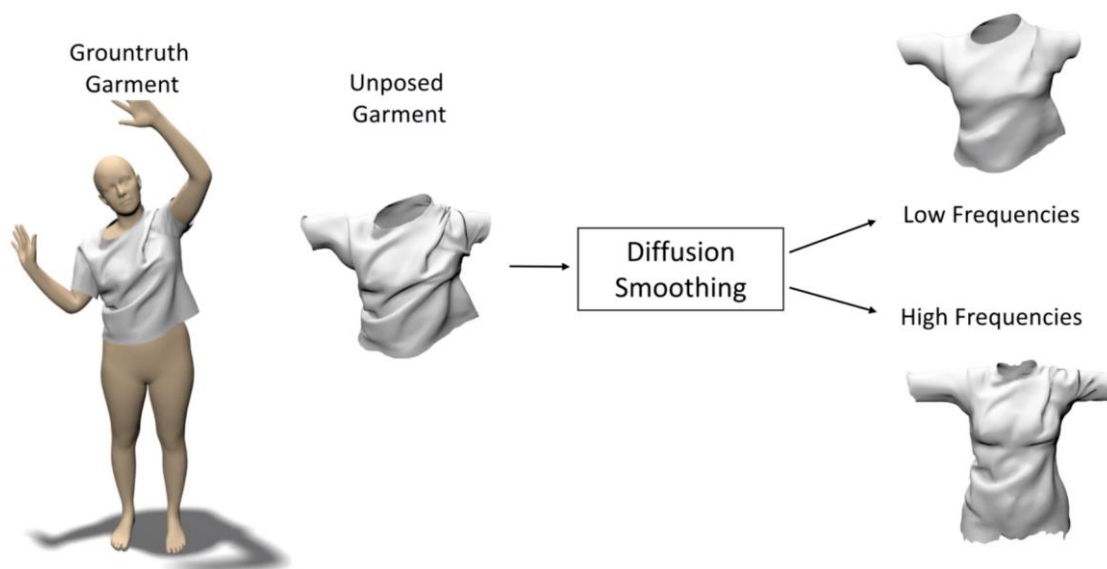
## 4. Mesh 频率分解法

论文作者发现，如果单单使用 Pose，Shape 和 Style 的联合多层感知器模型（MLP Joint Model）去模拟全部的衣服形变，得到的模拟结果是过度平滑的，特别是作者想要的褶皱效果，基本上被平滑了。



所以论文作者提出把平滑计算分散到两个频率的部分。低频部分决定了衣物的大致形状，只需要粗略的模拟就能提供良好的效果；高频部分决定了衣物的细节，比如褶皱效果，需要精细化的模拟。最终的模拟效果是两者的叠加。

### TailorNet – Frequency Decomposition

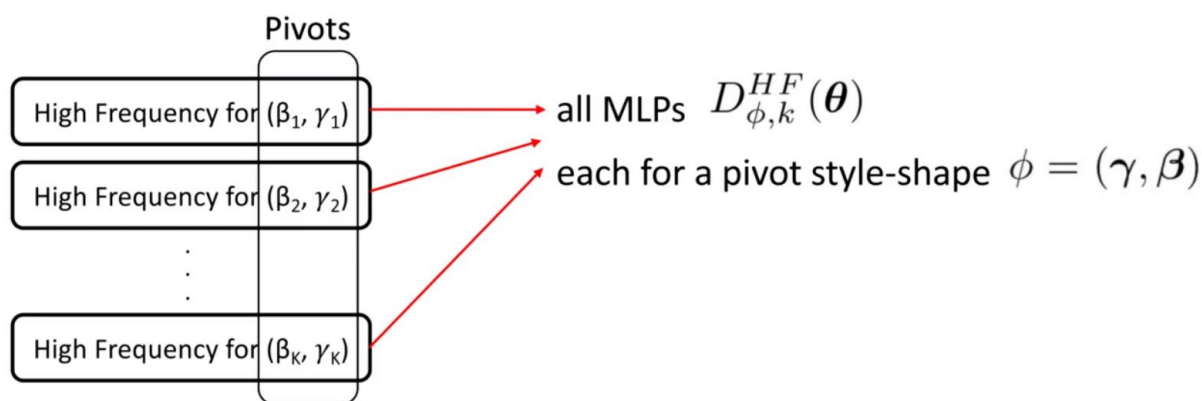


## 4.1 低频部分

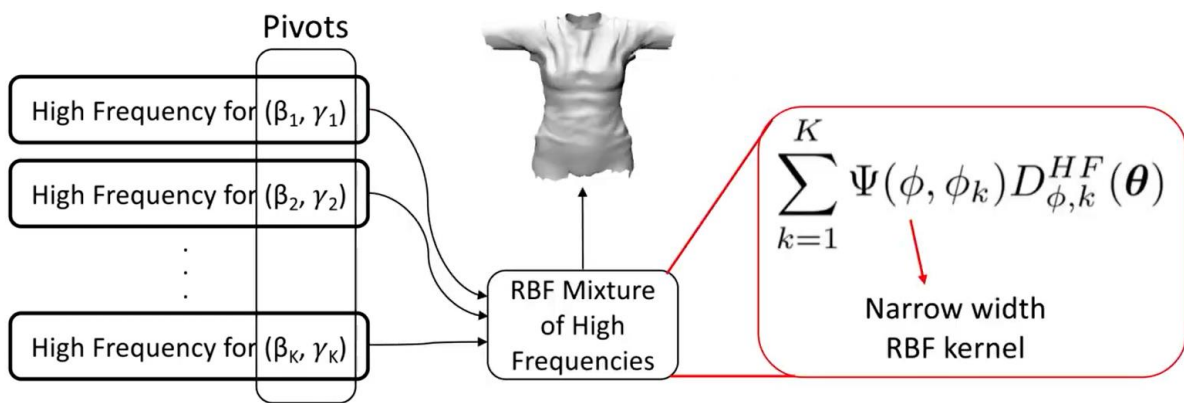
论文作者之前试验的联合 MLP 模型已经能较好的模拟衣服的大致形状，只是欠缺细节。所以低频部分直接沿用 MLP 模型，把 pose, shape, style 三组参数构建 MLP 模型来预测低频效果。

## 4.2 高频部分

高频部分对细节有很高的要求。发生形变时，不同衣物在细节上可能有完全不同的表现，而人物躯体形状（大小、长短等参数）也会直接影响衣服的形变（褶皱、修身程度、贴合度等），因此，可以把每组 shape 和 style 作为一组参数  $\phi = (\beta, \gamma)$ ，这一对参数再与姿势 pose 构建一系列 MLP 模型（每个 MLP 模型对应同一组  $\phi$  在不同  $\theta$  下的模拟结果），提高了神经网络模型的复杂度，从而实现细粒度模拟。

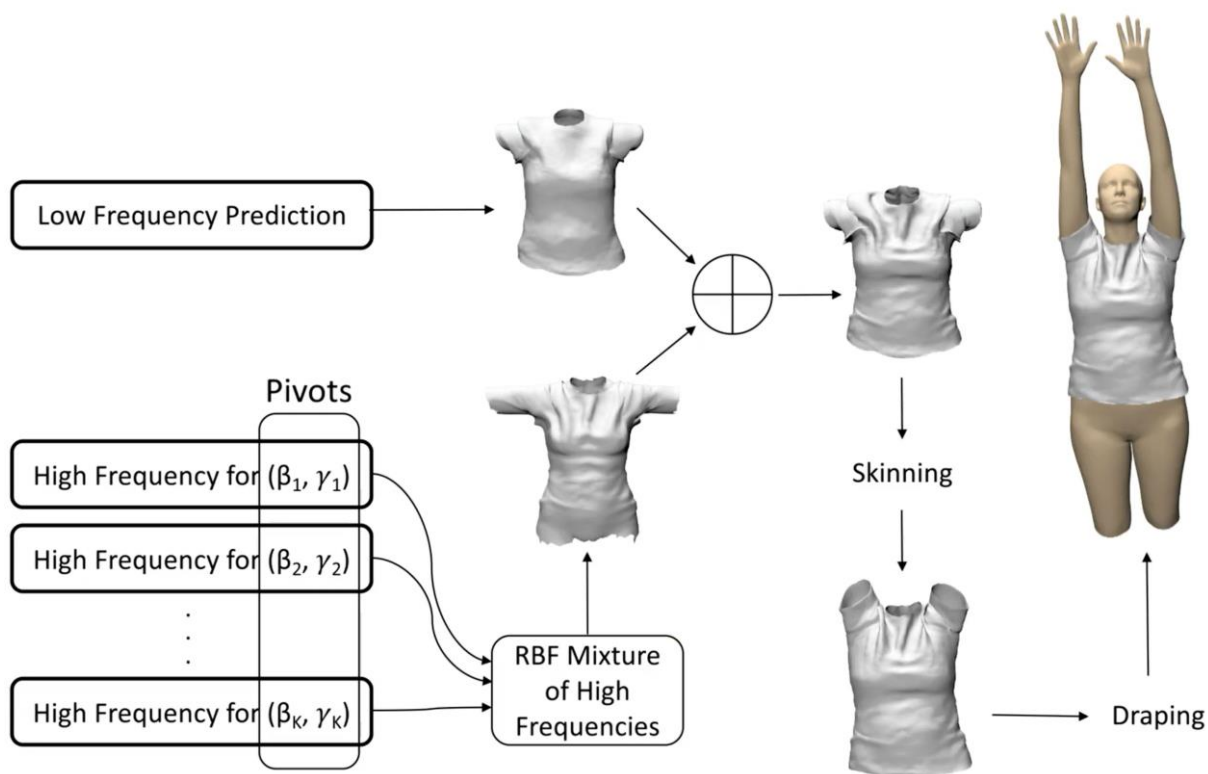


除了细粒度模拟，写实的效果要求突出衣服的特点。因此，对于每一组 shape-style 模型  $\phi$ ，加权求和得到它最后的输出。求和的权重是窄宽径向基函数核（Narrow Width RBF Kernel），它可以简单视为两个特征向量的距离的衡量函数，它的值介于 0 到 1 之间，核的值随距离的增大而减小，通常用来表示两个向量的相似度，而窄宽将使得核的半径减少，函数值的变化程度更加剧烈，能够显著地拉开两个向量的距离，让相似的向量对更加明显。使用窄宽 RBF 加权后，相似的 shape-style 输出得到保留，迥异的输出被弱化。

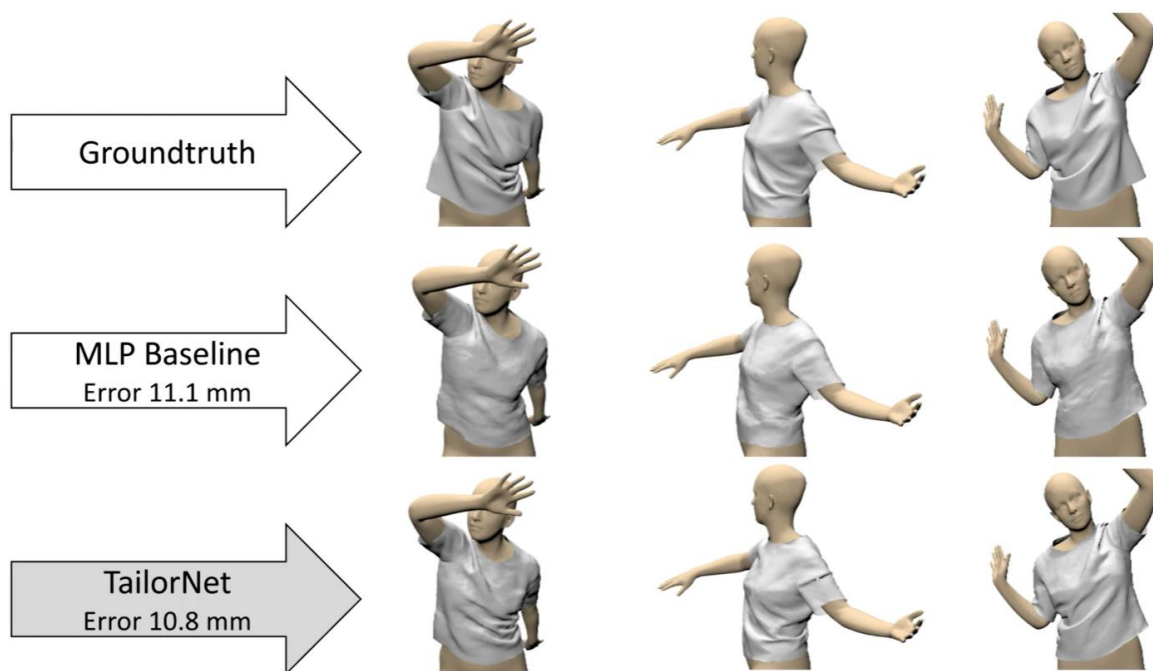


### 4.3 效果叠加

最后，低频部分与高频部分叠加在一起（如何叠加还有待细看），重新生成新的蒙皮，覆盖到裸体模型上。



最后的模拟效果，在褶皱等细节上，比简单的 MLP 模型好很多。



## 5. 对比

TailorNet 与基于物理的布料模拟 ( Physical Based Simulation, 简称 PBS ), 比如 Magica Cloth 等插件 :

- (1) PBS 有更好的通用性, TailorNet 在使用上有局限。除了衣物, PBS 的布料系统还能模拟头发等类布料; 而 TailorNet 只能模拟完整的衣物, 因为只有传统、完整的衣物才能提供大量数据训练模型, 类似天涯明月刀里的修仙衣服无法模拟。
  - (2) TailorNet 的模拟效果更写实 ( 比如褶皱 ), PBS 想要达到相同效果十分消耗性能, 速度也慢。TailorNet 根据动画离线计算了蒙皮上每个点的模拟数据, 细粒度控制 mesh 的运动变化, 运行时只需要加载蒙皮上的顶点坐标即可, 无需额外的物理解算。PBS 是实时计算的, 如果要达到写实的效果, 需要大幅提高模型的面数和顶点数, 并解算复杂的物理约束。
  - (3) PBS 有不确定因素, 难以完全控制每次的效果相同。PBS 是物理引擎的计算的结果, 除了布料自身的影响外, 还可能受场景内其他布料、外力的影响, 而且计算过程中有大量的随机数, 虽然大致的效果能够保证相近, 但细节的模拟效果难以复现。
- TailorNet 的模拟效果以离线坐标的方式传入, 模型输出后的结果是唯一的, 模拟过程中也不与外界交互, 因此每个动画里的衣服模拟效果是相同的。



- (4) TailorNet 把效果交给真实衣服的训练数据，PBS 的模拟效果则需要使用者有一定的物理知识。对 PBS 来说，底层的设计需要开发者精通物理知识，提炼出通用且关键的参数和约束，并实现解算的过程。TailorNet 要求用真实的衣物数据训练，通过 PCA 提取衣服的特征向量，重新组合后得到较为理想的模拟效果。

## 引用

TailorNet 训练模型 repo <https://github.com/chaitanya100100/TailorNet>

TailorNet 数据集 repo [https://github.com/zycliao/TailorNet\\_dataset](https://github.com/zycliao/TailorNet_dataset)

TailorNet 模型主页 <https://virtualhumans.mpi-inf.mpg.de/tailornet/>

SMPL 模型主页 <https://smpl.is.tue.mpg.de/>

TailorNet 论文 <https://virtualhumans.mpi-inf.mpg.de/tailornet/patel20tailornet.pdf>

TailorNet Oral Presentation <https://www.youtube.com/watch?v=vq7a52zObjsTailorNet>

效果演示 [https://www.youtube.com/watch?v=F0O21a\\_fsBQ](https://www.youtube.com/watch?v=F0O21a_fsBQ)