# Tailor-Net

mixture of NN model
using joint, shape, style variation jointly
for 3D cloth deformation
trained on static pose

In this paper, we present TailorNet, a neural model which predicts clothing deformation in 3D as a function of three factors: pose, shape and style (garment geometry), while retaining wrinkle detail. This goes beyond prior models, which are either specific to one style and shape, or generalize to different shapes producing smooth results, despite being style specific. Our hypothesis is that (even non-linear) combinations of examples smooth out high frequency components such as fine-wrinkles, which makes learning the three factors jointly hard. At the heart of our technique is a decomposition of deformation into a high frequency and a low frequency component. While the low-frequency component is predicted from pose, shape and style parameters with an MLP, the high-frequency component is predicted with a mixture of shape-style specific pose models. The weights of the mixture are computed with a narrow bandwidth kernel to guarantee that only predictions with similar high-frequency patterns are combined. The style variation is obtained by computing, in a canonical pose, a subspace of deformation, which satisfies physical constraints such as inter-penetration, and draping on the body. TailorNet delivers 3D garments which retain the wrinkles from the physics based simulations (PBS) it is learned from, while running more than 1000 times faster. In contrast to

- PBS(물리 기반 시뮬)보다 1000배 빠르게 3D 의류 변형은 추론 신경망 모델
- 3가지 인자(pose, shape, style)
- 고주파(주름)-mixture of shape style specific pose model-와 저주파-MLP-로 나눠서 추론
- Style은 interpenetration(상호침투,collision)를 고려한 공간에서 canonical pose(고정된 표준자세)에서 계산된다.

## 1. Introduction

What is lacking is a unified model capable of generating different garment styles, and animating them on any body shape in any pose, while retaining wrinkle detail. To that end, we introduce *TailorNet*, a mixture of Neural Networks (NN) learned from physics based simulations, which decomposes clothing deformations into style, shape and pose – this effectively approximates the physical clothing deformation allowing intuitive control of synthesized animations. In the same spirit as SMPL [32] for bodies, TailorNet learns deformations as a displacements to a garment template in a canonical pose, while the articulated motion is driven by skinning. TailorNet can either take a real garment as input, or generate it from scratch, and drape it on top of the SMPL body for any shape and pose. In contrast to [17], our model predicts how the garment would fit in reality, *e.g.*, a medium size garment is predicted tight on a large body, and loose on a thin body.

- 물리 시뮬을 기반으로 학습한 Mixture of NN
- 실제 의류(이미지?)를 input으로 가질 수 있다

Learning TailorNet required addressing several technical challenges. To generate different garment styles in a static pose, we compute a PCA subspace using the publicly available digital wardrobe of *real static garments* [6]. To generate more style variation while satisfying garment-human physical constraints, we sample from the PCA subspace, run PBS for each sample, and recompute PCA again, to obtain a *static style subspace*. Samples from this subspace produce variation in sleeve length, size and fit in a static pose. To learn deformation as a function of pose and shape, we generated a semi-real dataset by animating garments (real or samples from the static style subspace) using PBS on top of SMPL [32] body for static SMPL poses, and for different shapes.

- PCA subspace를 계산함으로써 static(pose) style subspace를 얻는다.(앞에서 interpenetration 고려한다 했음)
- [6]이 핵심인듯
- SMPL에 PBS로 옷을 입혀서 데이터셋 만듬 -> blender로 데이터셋 만들자(tailor와 다르게 동적으로도?)

Our first observation is that, for a *fixed style* (garment instance) and body shape, predicting high frequency clothing deformations as a function of pose is possible – perhaps surprisingly, our experiments show that, for this task, a simple multi-layer perceptron model (MLP) performs as well as or better than Graph Neural Networks [31, 28] and Image-decoder on a UV-space [29]. In stark contrast, straightforward prediction of deformation as a function of style, shape and pose results in overly smooth un-realistic results. We hypothesize that any attempt to combine training examples smoothes out *high frequency* components, which explains why previous models [46, 17, 18], even for a single style, lack fine scale wrinkles and folds.

- 이전 연구(GNN, UV-Decoder)와 비교해서 간단한 MLP(TailorNet)의 성능이 더 뛰어났다.
- Deformation을 style shape pose의 함수로 모델링 하는 것은 너무 smooth한 결과를 줌
- Training example을 결합하는 것은 옷의 주름을 없앤다는 가설?

These key observations motivate the design of TailorNet: we predict the clothing low frequency geometry with a simple MLP. High frequency geometry is predicted with a mixture of high frequency style-shape specific models, where each specific model consists of a MLP which predicts deformation as a function of pose, and the weights of the mixture are obtained using a kernel which evaluates similarity in style and shape. A kernel with a very narrow bandwidth, prevents smoothing out fine scale wrinkles. Several experiments demonstrate that our model generalizes well to novel poses, predicts garment fit dependent on body shape, retains wrinkle detail, can produce style variations for a garment category (*e.g.*, for T-shirts it produces different sizes, sleeve lengths and fit type), and importantly is easy to implement and control. To summarize, the main contributions of our work are:

Several experiments show that our model generalizes to completely new poses of the AMASS dataset (even though we did not use AMASS to train our model), and produces variations due to pose, shape and style, while being more detailed than previous style specific models [46, 18]. Furthermore, despite being trained on static poses, TailorNet produces smooth continuous animations.

- Mixture of shape style specific model에서 각 model은 MLP로 구현되어 있다.
- mixture weight는 의상과 모델간 style shape 유사도로 구해진다.(SMPL의 GMM처럼?)
- 정적자세에서만 학습했음에도 animation에서 좋은 결과