Dressing Avatars: Deep Photorealistic Appearance for Physically Simulated Clothing (Supplementary Document)

1 CLOTH SIMULATOR

Our real-time cloth simulator implements eXtended Position Based Dynamics [Fratarcangeli et al. 2016; Macklin et al. 2016]. XPBD is a constraint-based simulation model that often obtains much better performance compared to expensive non-linear solvers. It uses an iterative Gauss-Seidel solution for the linearized equations of motion. The method can be easily parallelized [Fratarcangeli et al. 2016] and implemented on hardware such as multi-core CPUs and GPUs, enabling interactive or real-time simulations on common modern hardware.

The method aims to solve Newton's equations of motion

$$\mathbf{M}\ddot{\mathbf{x}} = -\nabla U(\mathbf{x}), \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^{3n}$ encodes n vertex positions (of the cloth mesh in this case) and $\mathbf{M}$ is the mass matrix computed from element volumes and constant material density ρ. The energy potential $U(\mathbf{x})$ needs to be specified in terms of a vector of constraint functions $\mathbf{C}(\mathbf{x}) = [C_1(\mathbf{x}), C_2(\mathbf{x}), \cdots, C_m(\mathbf{x})]^\top$ as

$$U(\mathbf{x}) = \frac{1}{2}\mathbf{C}(\mathbf{x})^\top \boldsymbol{\alpha}^{-1}\mathbf{C}(\mathbf{x}), \tag{2}$$

where $\boldsymbol{\alpha}$ is a block diagonal compliance matrix. Any energy that can be written this way is suitable for XPBD. Using implicit Euler time integration, the XPBD algorithm reduces to solving for the constraint multiplier updates $\Delta\boldsymbol{\lambda}$ with

$$(\nabla \mathbf{C}(\mathbf{x}_i)^\top \mathbf{M}^{-1} \nabla \mathbf{C}(\mathbf{x}_i) + \tilde{\boldsymbol{\alpha}})\Delta\boldsymbol{\lambda} = -\mathbf{C}(\mathbf{x_i}) - \tilde{\boldsymbol{\alpha}}\boldsymbol{\lambda}_i, \tag{3}$$

where $\mathbf{x}_i$ and $\boldsymbol{\lambda}_i$ are the values of $\mathbf{x}$ and $\boldsymbol{\lambda}$ at iteration i, and $\tilde{\boldsymbol{\alpha}} = \frac{\boldsymbol{\alpha}}{\Delta t^2}$. Then the position is updated by

$$\Delta\mathbf{x} = \mathbf{M}^{-1}\nabla\mathbf{C}(\mathbf{x}_i)\Delta\boldsymbol{\lambda}. \tag{4}$$

The system in Eq. 3 is typically solved using Gauss-Seidel- or Jacobi-style updates. Stretching and shearing of the fabric is modeled as a mass-spring system whereas the bending is modeled as a zero angle constraint for dihedral elements. The underlying body is modeled as a triangle mesh and it is directly used for collision handling. Our solver is implemented using CUDA kernels and runs on the GPU.

1.1 Discussion: Simulation Parameters

In Fig. 1, we show example results generated by our system using different material parameters in the cloth simulation. Two critical parameters in the simulation are bending stiffness and stretching stiffness.

- (A) The bending stiffness controls the level of wrinkles. The larger the bending stiffness is, the fewer wrinkles remain in the output results.
- (B) The stretching stiffness influences the level of stretching of clothing at equilibrium. The larger the stretching stiffness is, the closer the output sticks to the rest length of the clothing template.

For the results in this paper, we experiment with different physical parameters and select the output that is visually most similar to the captured images. The process is well illustrated by Fig. 1. In addition, it can be observed in Fig. 1 that the performance of our appearance model is not sensitive to the parameters. Therefore, we only need to devote modest efforts into the selection of parameters. More sophisticated approaches can also be adopted, including using measured material data [Miguel et al. 2012; Wang et al. 2011], or building a perceptual control space for simulation [Sigal et al. 2015], which are beyond the scope of this paper.

1.2 Discussion: A Different Simulator

In this section, we demonstrate the possibility to use a different simulator from the our default XPBD simulator in our system. For this purpose, we adopt an open-source cloth simulator[1] based on Projective Dynamics with frictional contact modeling [Ly et al. 2020]. We compare the results generated by this simulator (named 'Projective Friction') with those from XPBD simulator in Fig. 2. The body configurations and the clothing appearance model are kept the same for the comparison.

As shown in Fig. 2, the XPBD simulator and Projective Friction simulator produce different clothing geometry due to the discrepancies in their formulations. However, our clothing appearance model can generate proper appearance with reasonable detail of wrinkles and shadow that agree with the corresponding geometry from both simulators. This suggests that the clothing appearance model is not tied to a specific simulator. The animation framework that we present in this paper has the potential to generalize to different implementation of physics-based cloth simulation.

1.3 Implementation Detail: Simulation Template

Our XPBD simulator supports using a 3D clothing mesh as rest shape, from which the the reference triangle sizes for the stretching and shearing energy terms are computed. For the bending energy, we still assume zero rest angles for dihedral elements. This allows us to directly use the same template mesh for clothing registration as the rest shape for the simulation.

The open-source simulator [Ly et al. 2020] requires a 2D template as the rest shape. We follow a previous work [Bang et al. 2021] to create a 2D template from the 3D registration template. The basic idea is to cut the 3D template mesh into several pieces, flatten them with minimal distortion, and enforce boundary smoothness requirements.

It should be also possible to adopt a very recent method [Pietroni et al. 2022] to create a 2D template. Another alternative is to modify the implementation [Ly et al. 2020] to support 3D template shapes. We leave a thorough discussion of these different possibilities to future work.

[1] https://gitlab.inria.fr/elan-public-code/projectivefriction

Fig. 1. We show results of different physical parameters used in the cloth simulation. We adjust the scale of bending stiffness on the left (A), and stretching stiffness on the right (B). Notice that our rendered results are reasonable despite the difference in simulation parameters.

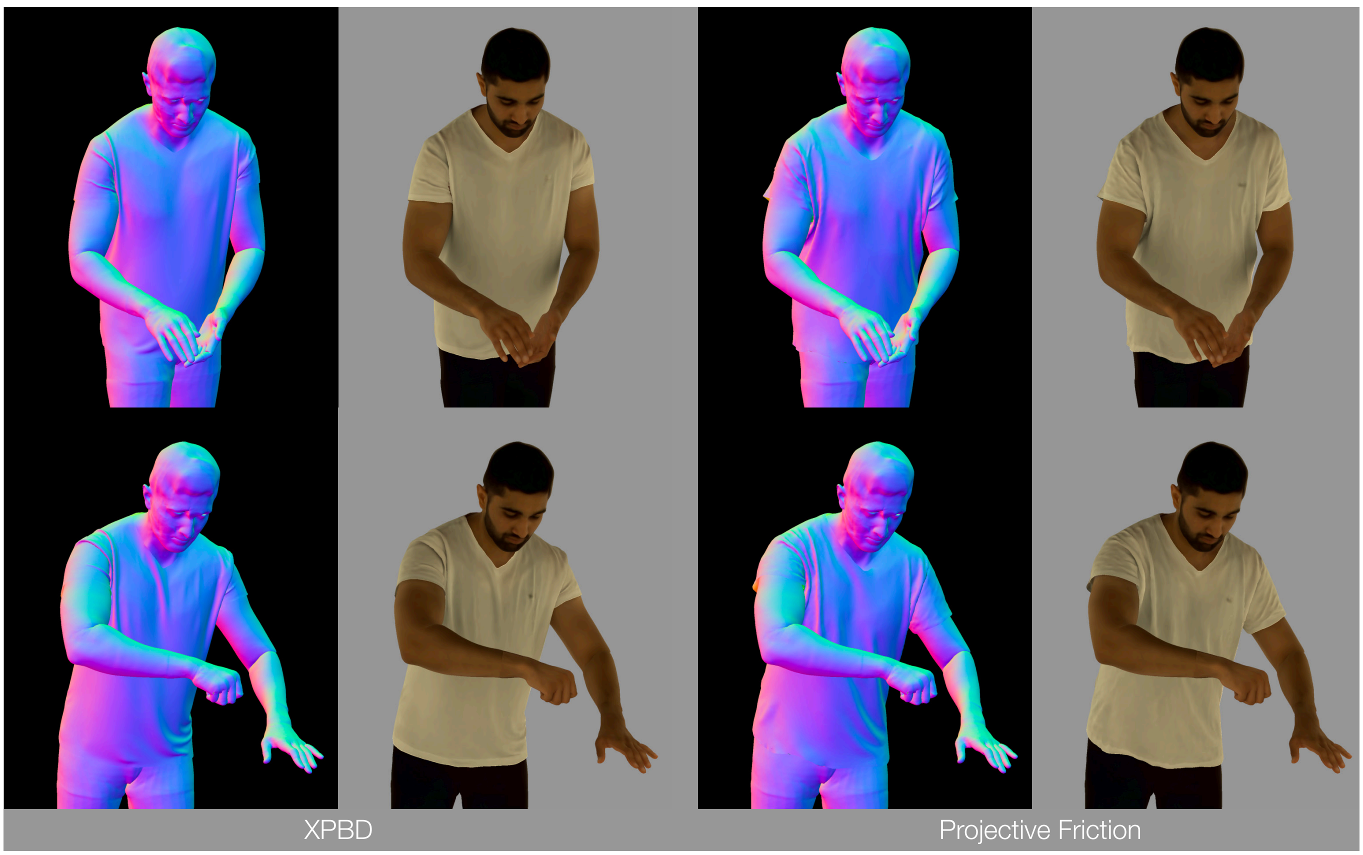

Fig. 2. We compare animation results using our default XPBD simulator with Projective Friction [Ly et al. 2020]. We show normal rendering of the simulated geometry together with the underlying body on the left, and the results of our photorealistic clothing appearance model on the right.

2 IMPLEMENTATION DETAIL

2.1 Base Body Avatars

In this work, we use two types of base body avatars, which we call minimally clothed avatars and underlying body avatars.

The minimally clothed avatars are built from captured sequences where the subjects wear only a green tight suit. To build these types of avatars, we use the same procedure as in previous work [Bagautdinov et al. 2021]. Since the capture suit tightly follow the body motion, the single-layer full-body avatar is able to model the full appearance. The avatars adopt a convolutional Variational Autoencoder (cVAE) architecture, conditioned on body pose, facial conditioning (if applicable), a latent code from the encoder, and ambient occlusion as input. At test (animation) time, we follow Bagautdinov et al. [2021] and use a fixed latent code (all zeros in our case) for all the body motion in the test sequence.

The underlying body avatars are built from normally clothed body capture as in previous work [Xiang et al. 2021]. To train these types of avatars, we register the clothing and body in two separate layers. The clothing registration method is described in Sec. 4 of our paper. In order to track body under loose clothing, we utilize the minimally clothed body data of the same subject as a prior. When tracking the skeleton poses using the clothed body reconstruction in [Xiang et al. 2021] (supplementary document, Sec. 1.2), we exclude the highly dynamic clothing region (bottom of the dress and the whole skirt) in the surface distance loss. To estimate the underlying body surface, we couple the invisible region of the body shape with the minimally clothed LBS model, and penalize collisions with the clothing surface similar to [Zhang et al. 2017]. With the tracked body data, we train the body-layer avatars with the same network architecture as described in the supplementary document of [Xiang et al. 2021], but without the clothing branch. The training process is similar to the description in Sec. 5.1 and 5.3 of [Xiang et al. 2021].

2.2 Clothing Appearance Model

In this section we provide more implementation detail on the clothing appearance model. We use the same architecture for both the view-independent and view-dependent networks, which is described in Fig. 3. For the shadow network, the architecture is kept to be the same as in previous work [Xiang et al. 2021]. The appearance model is trained end-to-end in PyTorch using the AdamW optimizer with an initial learning rate of 1×10^{-4}. The training goes on for 100k iterations with the batch size of 2.

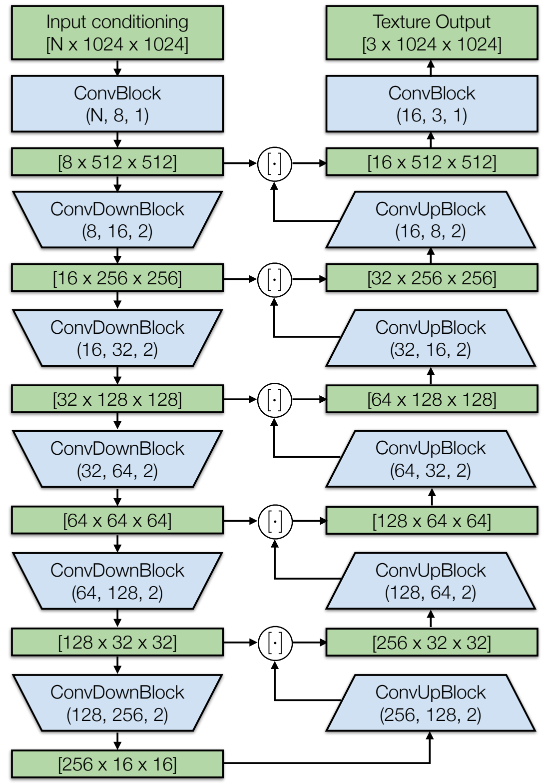

Fig. 3. The detailed architecture of view-independent and view-dependent networks used in the clothing appearance models. We adopt the UNet [Ronneberger et al. 2015] structure with skip connections. Data tensors are shown in green where the numbers represent '[channels, height, width]'. Network modules are shown in blue, where the numbers represent '(input channels, output channels, stride)'. For the exact detail of each block, please refer to the supplementary document of [Xiang et al. 2021].

3 ETHICAL CONSIDERATIONS

With the formulation described in this paper, building a high-quality avatar is not yet possible unless the subject agrees to be captured in a multi-camera system, a relatively high requirement on the capture setup. However, with the high-quality photorealistic avatars, there exist risks of others using the content to create fake imagery of the subjects. Engineering efforts must be made to promote security and limit access to intended users only.

REFERENCES

Timur Bagautdinov, Chenglei Wu, Tomas Simon, Fabian Prada, Takaaki Shiratori, Shih-En Wei, Weipeng Xu, Yaser Sheikh, and Jason Saragih. 2021. Driving-signal aware full-body avatars. *ACM Transactions on Graphics (TOG)* 40, 4 (2021).

Seungbae Bang, Maria Korosteleva, and Sung-Hee Lee. 2021. Estimating Garment Patterns from Static Scan Data. In *Computer Graphics Forum*, Vol. 40.

Marco Fratarcangeli, Valentina Tibaldo, and Fabio Pellacini. 2016. Vivace: A practical gauss-seidel method for stable soft body dynamics. *ACM Transactions on Graphics (TOG)* 35, 6 (2016).

Mickaël Ly, Jean Jouve, Laurence Boissieux, and Florence Bertails-Descoubes. 2020. Projective dynamics with dry frictional contact. *ACM Transactions on Graphics (TOG)* 39, 4 (2020).

Miles Macklin, Matthias Müller, and Nuttapong Chentanez. 2016. XPBD: position-based simulation of compliant constrained dynamics. In *Proceedings of the 9th International Conference on Motion in Games.*

Eder Miguel, Derek Bradley, Bernhard Thomaszewski, Bernd Bickel, Wojciech Matusik, Miguel A Otaduy, and Steve Marschner. 2012. Data-driven estimation of cloth simulation models. In *Computer Graphics Forum*, Vol. 31.

Nico Pietroni, Corentin Dumery, Raphael Falque, Mark Liu, Teresa Vidal-Calleja, and Olga Sorkine-Hornung. 2022. Computational pattern making from 3D garment models. *ACM Transactions on Graphics (TOG)* 41, 4 (2022).

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention.* Springer, 234–241.

Leonid Sigal, Moshe Mahler, Spencer Diaz, Kyna McIntosh, Elizabeth Carter, Timothy Richards, and Jessica Hodgins. 2015. A perceptual control space for garment simulation. *ACM Transactions on Graphics (TOG)* 34, 4 (2015).

Huamin Wang, James F O'Brien, and Ravi Ramamoorthi. 2011. Data-driven elastic models for cloth: modeling and measurement. *ACM Transactions on Graphics (TOG)*

30, 4 (2011).
Donglai Xiang, Fabian Prada, Timur Bagautdinov, Weipeng Xu, Yuan Dong, He Wen, Jessica Hodgins, and Chenglei Wu. 2021. Modeling clothing as a separate layer for an animatable human avatar. *ACM Transactions on Graphics (TOG)* 40, 6 (2021).
Chao Zhang, Sergi Pujades, Michael J Black, and Gerard Pons-Moll. 2017. Detailed, accurate, human shape estimation from clothed 3D scan sequences. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.