

# Table of Contents

Acknowledgements	ii
Table of Contents	iii
List of Tables	vii
List of Figures	viii
Chapter 1: Introduction	1
1.1 Learning-Based Human Body and Garment Estimation . . . . .	3
1.2 Differentiable Simulation for Material Optimization . . . . .	4
1.3 Simulation-Based Virtual Try-On . . . . .	5
1.4 Thesis Statement . . . . .	6
1.5 Main Results . . . . .	7
1.5.1 Shape-Aware Human Reconstruction Using Multi-View Images	7
1.5.2 Differentiable Simulation for Material Optimization . . . . .	8
1.5.3 Joint Estimation of Human and Garment from Video . . . . .	9
1.5.4 Time-Domain Parallelization for Accelerating Cloth Simulation	10
1.5.5 Dynamics-Inspired Garment Draping Prediction . . . . .	10
1.6 Outline of Dissertation . . . . .	11
Chapter 2: Shape-Aware Human Reconstruction Using Multi-View Images	13
2.1 Introduction . . . . .	13
2.2 Related Work . . . . .	16
2.2.1 Human Body Pose and Shape Recovering . . . . .	16
2.2.2 Learning-Based Pose/Shape Estimations . . . . .	19
2.2.3 Use of Synthetic Dataset . . . . .	19
2.3 Overview . . . . .	20
2.4 Model Architecture . . . . .	21
2.4.1 3D Body Representation . . . . .	23
2.4.2 Scalable Multi-View Framework . . . . .	24
2.4.3 Training and Inferring . . . . .	25
2.4.4 Implementation Details . . . . .	26
2.5 Data Preparation . . . . .	27
2.5.1 Parameter Space Sampling . . . . .	28
2.5.2 Human Body Motion Synthesis . . . . .	28

2.5.3	Cloth Registration and Simulation	29
2.5.4	Multi-View Rendering	29
2.6	Results	31
2.6.1	Ablation Study	31
2.6.2	Comparisons with Multi-View Methods	34
2.6.3	Real-World Evaluations	35
2.6.4	Multi-View Input in Daily Life	36
2.6.5	Extra Test Results	37
2.6.6	Additional Results on Real-World Images	38
2.6.7	Comparison on Human3.6M with Single-View Methods	40
2.6.8	Results Without Training on Synthetic Data	40
2.6.9	Detailed Errors on Real World Evaluation	41
2.6.10	Evaluation on <i>3D People in the Wild</i> .	42
2.6.11	Running Time	43
2.7	Conclusion and Future Work	43
Chapter 3: Differentiable Simulation for Material Optimization		45
3.1	Introduction	45
3.2	Related Work	47
3.3	Differentiable Cloth Simulation	49
3.3.1	Cloth Simulation Basics	50
3.3.2	Overview	53
3.3.3	Derivatives of the Physics Solve	54
3.3.4	Dynamic Collision Detection and Response	55
3.3.5	Derivatives of the Collision Response	56
3.3.6	Derivations of the Gradient Computation	62
3.4	Experiments	65
3.4.1	Ablation Study	66
3.4.2	Material Estimation	67
3.4.3	Motion Control	71
3.4.4	Collision-rich Motion Control	73
3.5	Conclusion	73
Chapter 4: Joint Estimation of Human and Garment from Video		75
4.1	Introduction	75
4.2	Related Work	77
4.3	Method Overview	80
4.4	Garment Auto-encoder	82
4.4.1	Two-Level Encoder-Decoder Structure	83
4.4.2	Representative Point Set Extraction	85
4.4.3	Training Losses	85
4.4.4	Recovery from Point Clouds to Garment Meshes	86
4.5	Material Estimation	87
4.5.1	Single Frame Closed-Loop Estimation	87
4.5.2	Temporal Estimation for Garment Material	90

4.6	Data Preparation and Training	91
4.6.1	Training Details	92
4.7	Experiments	93
4.7.1	Quantitative Analysis	93
4.7.2	Qualitative Results	95
4.7.3	Lab Experiments and User Study	98
4.7.4	Ablation Study	100
4.7.5	Latent Code Interpolation	104
4.7.6	Additional Qualitative Results	105
4.7.7	Application: Virtual Try-On	108
4.8	Conclusion	109
Chapter 5: Time-Domain Parallelization for Accelerating Cloth Simulation		112
5.1	Introduction	112
5.2	Related Work	115
5.2.1	Cloth Simulation	115
5.2.2	Time Parallel Time Integration Method	116
5.2.3	Parallel Cloth Simulation	117
5.2.4	Hierarchical Structures and Multi-level Methods	117
5.2.5	Mesh Upsampling	118
5.3	Overview	119
5.3.1	Two-Level Mesh Hierarchy Representation	120
5.4	Time Domain Parallelization	121
5.4.1	Static Temporal Partitioning	122
5.4.2	Adaptive Partitioning	124
5.4.3	Analysis on Performance Scalability	127
5.5	Smooth State Transitioning	128
5.5.1	Iterative Detail Recovery	129
5.5.2	Convergence and Continuity	131
5.5.3	Proof of Convergence of Algorithm 3	131
5.5.4	Iteration Number Estimation	134
5.5.5	Implementation Details	136
5.5.6	State Inconsistency	137
5.6	Results	138
5.6.1	Parameter and Scenario Setting	138
5.6.2	Performance	140
5.6.3	Smoothness	147
5.6.4	Memory and Render Latency	147
5.6.5	Limitations	148
5.7	Conclusion and Future Work	149
Chapter 6: Dynamics-Inspired Garment Draping Prediction		150
6.1	Introduction	150
6.2	Related Work	152
6.3	Method	155

6.3.1	Encoder . . . . .	156
6.3.2	GCN-Based Decoder . . . . .	157
6.3.3	Spectral Domain Decomposition . . . . .	158
6.3.4	Loss Functions . . . . .	160
6.4	Physics-Enforced Optimization . . . . .	162
6.5	Experiments . . . . .	165
6.5.1	Data Generation . . . . .	165
6.5.2	Ablation Study . . . . .	166
6.5.3	Optimization for Semi-Supervision . . . . .	168
6.5.4	Optimization for Graphic Print . . . . .	169
6.5.5	Quantitative Comparisons . . . . .	171
6.5.6	Qualitative Results . . . . .	172
6.5.7	Generalization to Different Garment Sizes . . . . .	173
6.6	Conclusion . . . . .	175
Chapter 7: Conclusion		176
7.1	Summary of Results . . . . .	176
7.2	Limitations . . . . .	179
7.3	Future Work . . . . .	181

## List of Tables

2.1	Comparison results on Human3.6M using MPJPE . . . . .	32
2.2	Comparison results on MPLINF_3DHP . . . . .	32
2.3	Comparison results on my synthetic dataset . . . . .	33
2.4	Comparison on Human3.6M . . . . .	35
2.5	Comparison results on tape-measured data . . . . .	35
2.6	Results on MPLINF_3DHP, validation set . . . . .	39
2.7	Results on MPLINF_3DHP, test set . . . . .	39
2.8	Results on Human3.6M . . . . .	41
2.9	Percentages of errors in common measurements . . . . .	42
2.10	Evaluation on an unseen dataset . . . . .	43
3.1	Statistics of the backward propagation . . . . .	67
3.2	Results on the material parameter estimation task . . . . .	70
3.3	Motion control results . . . . .	72
4.1	Comparison on material estimation . . . . .	94
4.2	Quantitative comparison . . . . .	95
4.3	Lab experiment results . . . . .	98
4.4	Ablation study for different parts . . . . .	100
4.5	Test errors on the Multi-Garment Net dataset . . . . .	104
4.6	Comparison with previous works . . . . .	108
5.1	Notations and definition of my method . . . . .	120
5.2	Results on a higher-resolution mesh . . . . .	139
5.3	Comparison between different partition schemes . . . . .	141
5.4	Results in the extreme case . . . . .	143
5.5	Comparison with GPU method . . . . .	144
6.1	Encoders ablation . . . . .	166
6.2	Decoders ablation . . . . .	167
6.3	Losses ablation . . . . .	168
6.4	Self-correcting pipeline ablation study . . . . .	169
6.5	Adaptation to new materials . . . . .	170
6.6	Comparison with TailorNet . . . . .	172
6.7	Comparison for models on different sizes . . . . .	174

## List of Figures

2.1	The network structure . . . . .	22
2.2	Detailed network structure of the regression block . . . . .	26
2.3	Examples of rendered synthetic images . . . . .	30
2.4	Prediction results compared to HMR . . . . .	37
2.5	Results on images with varying pose and shape . . . . .	38
2.6	Results on real-world multi-view images . . . . .	38
2.7	My model trained without synthetic data. . . . .	41
3.1	Impact of perturbation . . . . .	59
3.2	Example frame from the ablation study . . . . .	66
3.3	Example frame from the material estimation scene . . . . .	68
3.4	Example frame from the motion control experiment . . . . .	70
3.5	A motion control scene with more obstacles . . . . .	72
4.1	Overall network structure . . . . .	81
4.2	The network structure of the garment auto-encoder . . . . .	82
4.3	My estimation pipeline . . . . .	88
4.4	The network structure for body and garment estimation . . . . .	88
4.5	Qualitative comparison . . . . .	97
4.6	Material transfer between videos . . . . .	97
4.7	Qualitative results . . . . .	97
4.8	Qualitative comparison with a real-world video . . . . .	100
4.9	Sample test images for the comparisons . . . . .	101
4.10	Interpolation between different garments . . . . .	101
4.11	Material transfer examples . . . . .	102
4.12	Interpolation results . . . . .	106
4.13	Qualitative Results . . . . .	107
4.14	Virtual try-on example . . . . .	109
4.15	Training data examples . . . . .	109
4.16	User study examples . . . . .	110
5.1	Simulated ‘Karate’ animation using my method . . . . .	112
5.2	An overview of my method . . . . .	115
5.3	Adaptive partitioning Algorithm . . . . .	125
5.4	An example comparison of the meshes . . . . .	128
5.5	Performance scaling result (large time step) . . . . .	140

5.6	Results with increasing length of the simulation . . . . .	140
5.7	Performance scaling result (small time step) . . . . .	142
5.8	Small scale parallelization comparison . . . . .	145
5.9	Large scale parallelization comparison . . . . .	145
5.10	More simulation results . . . . .	146
5.11	Refining results . . . . .	146
6.1	My model learns how to drape garments . . . . .	150
6.2	Overall structure of my network . . . . .	155
6.3	Visualization of the eigen decomposition . . . . .	156
6.4	Reconstructions for different numbers of coefficients. . . . .	159
6.5	The semi-supervised self-correcting training pipeline. . . . .	164
6.6	Bodies at BMI percentiles 10, 30, 50, 70 and 90%. . . . .	165
6.7	Qualitative examples of self-correcting optimization . . . . .	171
6.8	Qualitative comparison with TailorNet . . . . .	173
6.9	Qualitative comparison with previous work . . . . .	174