

3D Clothing Wrinkle Prediction Based on GCN and Human Keypoint Constraints

Chaoquan Wu

Zhejiang Fashion Institute of Technology, Ningbo 315000, Zhejiang, China

Corresponding author: Chaoquan Wu (e-mail: wuchaoquan2025@163.com)

ABSTRACT This paper proposes a three-dimensional clothing wrinkle prediction framework that integrates dynamic human keypoint constraints with adaptive graph convolution to achieve high-fidelity end-to-end wrinkle generation. Considering that accurate three-dimensional geometric representation is essential for digital modeling, intelligent visualization, and geometry-aware simulation in advanced engineering applications, a pose-aware constraint field is introduced to explicitly guide deformation, while an adaptive graph convolution module dynamically adjusts adjacency weights according to local curvature and keypoint proximity and employs a gated attention mechanism for feature fusion. This design enables effective modeling of the coupling relationship between human posture and garment deformation while preserving geometric consistency and fine wrinkle details. Experimental evaluations on the CAPE dataset and a self-constructed MPG dataset demonstrate that the proposed method achieves a vertex error of 3.82 mm, a Chamfer distance of 4.04 mm, a penetration rate of 0.71%, and a 15.8% improvement in local curvature standard deviation compared with ClothSim. The framework significantly enhances geometric fidelity, reduces non-physical artifacts, and improves wrinkle detail representation, providing an effective graph-based modeling strategy for highprecision three-dimensional digital reconstruction and geometry-aware computational applications relevant to intelligent electromagnetic and virtual simulation environments.

INDEX TERMS adaptive graph convolution, human keypoint constraints, 3D clothing wrinkle prediction, geometric modeling, digital electromagnetic simulation

I. INTRODUCTION

WITH the advancement of virtual reality, digital humans, and intelligent virtual try-on technology, high-fidelity 3D modeling of clothing has emerged as a popular research area in computer graphics and human-computer interaction [1-3]. The realism of clothing folds is a key visual feature that reflects the material of the garment, the posture of the person wearing the clothing, and the interaction of dynamics, and the fidelity of cloth folds directly influences the perceived realism and immersion of virtual characters [4-6]. While there have been proposed solutions to clothing folding representation, there is still room for improvement. For example, using physics simulation methods rely on the laws of mechanics with high levels of computational time and cost, as well as difficulty in real time application [7]. Data-driven models are efficient methods for inference but do not take into account geometric and physics coupling around the human posture and clothing deformation [8-9]. This limitation leads to distorted representations of clothing folds, fabric penetration, and loss of clarity in details. Another challenge is reconciling physical rationality with efficiency and accuracy in the dynamic processes of clothing modeling.

This paper focuses on the generation of 3D clothing wrinkles

driven by human posture, with the posture represented by key points, the basic clothing mesh, and the local details generated by their coupling as the core modeling elements [10-11]. In recent years, related research has continued to advance: Chen D et al. [12] improved the parameter estimation of SMPL (Skinned Multi-Person Linear model) through multiloss optimization; Wu H [13] combined biomechanical models to generate natural movements; Fisch M et al. [14] achieved 6D posture estimation; Xu C et al. [15] used dual-flow graph convolution to model skeletal movements; and Mouhou A A et al. [16] generated spatiotemporally consistent wrinkles using the HRBF (Hermite Radial Basis Function) method. These works collectively laid the foundation for “pose-driven wrinkle generation”.

Existing modeling strategies still have limitations. Some methods use CNN (Convolutional Neural Network) or Transformer to regress deformation from images or heatmaps, but are limited by the assumption of regular meshes and have difficulty handling non-rigid topologies [17-18]. For example, Hu X et al. [19] rely on image generation to enhance details, but lack pose-clothing coupling modeling. Mao M et al. [20] proposed a video classification framework based on the fusion of ResNet50 and Transformer, but did not involve 3D clothing mesh generation or

pose-driven wrinkle prediction, resulting in the inability to generate realistic and geometrically consistent 3D clothing deformation. Another type of GNN (Graph Neural Network) method improves geometric representation, but does not explicitly introduce pose constraints, and is prone to non-physical deformation under extreme poses. Furthermore, even when some GNN methods attempt to incorporate human pose information, they often fail to integrate key points as explicit geometric constraints into the graph convolution process. Li Y et al. [21]'s two-stream network separates pose from mesh processing and lacks cross-modal alignment. Wang H's [22] highly accurate simulation rests on quasi-static assumptions and cannot drive the pose of the clothing in real-time; Mo Z et al. [23]'s PointHuman focuses on implicit human reconstruction with fixed topology and does not predict the wrinkles generated by clothing. These prior works share problems in pose-clothing coupling modeling, which means the results generated can be highly susceptible to fabric penetration or distortion phenomena at extreme poses.

Current methods in pose-driven garment deformation present several outstanding challenges in the following three areas:

- (1) there is no dynamic posture constraint mechanism that can explicitly guide clothing deformation direction;
- (2) a fixed graph structure that is not demonstrating an "adaptive graph learning" feature, which makes it much less flexible for local deformations that are not diversity;
- (3) uncontrollable detail generation, which makes it very difficult to balance the geometry accuracy vs. physical rationality in high curvature regions.

It is noted that recent advances in implicit neural representations (INR) and neural radiance fields (NeRF) have shown remarkable capabilities in modeling complex appearances and geometries. However, these approaches often focus on view synthesis or reconstruction from multi-view images, and their application to pose-driven, topology-preserving garment deformation with explicit physical plausibility control remains less explored. This work positions itself within the paradigm of explicit graph-based deformation learning, which offers direct control over mesh topology, efficient inference, and straightforward integration of geometric constraints, addressing a distinct set of requirements for applications like real-time virtual try-on and animation.

To address these challenges, this paper introduces a 3D clothing wrinkle prediction framework, which incorporates dynamic constraints of human body key points and adaptive graph convolution for its implementation. The main contribution of this paper is explicitly provide a solution to the three limitations described above.

To address limitation (1), it constructs a dynamic constraint field for posture attention, which takes geometric semantics, such as key point bending angle and limb direction, and maps them to vertex-level guidance signals while also explicitly modeling the driving of posture has on deformations.

To address limitation (2), this paper proposes an adaptive graph convolutional network, in which adjacency weights are modulated by local curvature and the distances from key points, to enable dynamic evolution of the graph structure with posture and geometric context and to break the dependency of traditional GCN on fixed topology.

To address limitation (3), this paper adds gated attention mechanism to dynamically fuse features of constraint fields with geometric context while implementing lightweight post-post-processing (Laplacian Smoothing + Penetration Correction) to ensure physically bound rationality while preserving high-frequency details of wrinkles.

This approach not only provides a theoretical modeling of the pose-clothing coupling relationship but also technical construction of a scalable and interpretable graph learning paradigm. Experiments have demonstrated that this framework outperforms existing baselines on both the CAPE and the self-built MPG (MultiPoseGarment) dataset, achieving a balanced improvement in geometric accuracy, physical plausibility, and detail representation, establishing a new avenue for an efficient and realistic virtual clothing generation.

II. A 3D CLOTHING FOLD PREDICTION METHOD BASED ON KEYPOINT CONSTRAINTS AND ADAPTIVE GRAPH CONVOLUTION

To demonstrate the high-level architecture and data flow of the proposed approach, Figure 1 presents the end-to-end predictive process from input to output. This framework is designed to integrate human pose information with a simple garment mesh in an eventual output of a high-fidelity, physically plausible 3D clothing fold mesh, achieved through a series of carefully constructed modules.

The whole process can be broken down into three main phases:

1) Input and pre-processing: The system receives two fundamental inputs, a sequence of 3D human body key points and a simple wrinkle-free clothing mesh. Before reaching the core calculation module, both inputs must go under some standardization processing in order to be spatially aligned and uniformly scaled.

2) Core prediction: This is the core innovation of this method. First, the human body key points are structured into a skeleton graph, which generates a dynamic constraint field to guide the clothing deformation's direction and intensity. Afterward the basic clothing mesh is structured into a graph graph notation and input into an adaptive graph convolutional network [24-25]. The network includes a gated attention mechanism so that the contributing information is allowed to dynamically modify the information propagating path and back to contribute to an accurate layer-by-layer displacement prediction of each vertex [26-27].

3) Post-processing & Output: The initial predicted displacement field is added on top of the original mesh to produce an initial wrinkled mesh. To enhance the physical realism and visual quality of the outcomes, the system performs lightweight post-processing steps, including

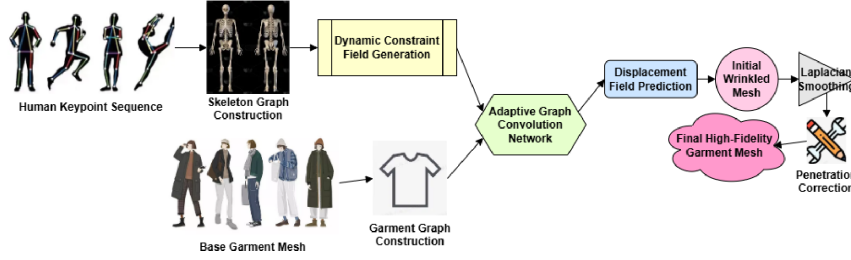


FIGURE 1. Flowchart of the overall method framework

Laplacian smoothing to attenuate high-frequency jitter, and keypoint-based penetration detection and geometric correction to ultimately deliver a physically-constrained high-fidelity 3D clothing mesh.

INPUT DATA REPRESENTATION AND PREPROCESSING

The input consists of two parts: a sequence of 3D human body keypoints and a basic wrinkle-free clothing mesh. The human body keypoint sequence is derived from the SMPL parametric model or optical motion capture system, with a uniform sampling frame rate of 30 Hz and a fixed number of 24 main joint nodes. The clothing mesh uses a standard template registered under T-pose, with N_v vertices and N_f faces, and the topology remains consistent across all samples.

The preprocessing process is as follows: First, the human body keypoint sequence is globally normalized. Taking the pelvic node as the origin, all coordinates are scaled to a unit scale, and global translational and rotational degrees of freedom are eliminated through Procrustes alignment, while preserving local pose information [28-29]. Second, the clothing mesh is converted into an undirected graph structure $G = (V, E)$, where the node set V corresponds to the mesh vertices, and the edge set E is constructed from the first-order adjacency relationship of the original mesh, that is, there is an edge connection between any two vertices of shared triangular faces. Finally, the human skeleton graph $G_{skeleton} = (J, E_j)$ is defined based on the SMPL skeleton, where J is the set of keypoints, and E_j is generated according to the parent-child joint connection relationship predefined in SMPL for subsequent constraint modeling.

The preprocessing output is: a standardized keypoint coordinate set $J \in \mathbb{R}^{T \times 24 \times 3}$ (T is the time step) and a garment structure G , both aligned in the spatial coordinate system and possessing consistent scale and orientation references.

DYNAMIC CONSTRAINT MODELING OF HUMAN KEYPOINTS

This module transforms pose information into vertex-level geometric guidance. Given the standardized keypoint set $J \in \mathbb{R}^{24 \times 3}$, local joint curvature θ_i and limb direction vectors are computed for each joint. The spatial influence of keypoints on garment vertices is then modeled using inverse distance weighting $w_{k,i}$ (Formula 2) and enhanced by orthogonal projection for limb regions. A constraint

feature vector c_k for each vertex is generated by aggregating these weighted dynamic features via a learnable embedding function $\phi(\cdot)$ (Formula 3). The resulting constraint field $C = \{c_k\}$ explicitly encodes pose-derived geometric constraints to guide subsequent deformation:

$$\cos \theta_i = \frac{(j_p - j_i)^\top (j_c - j_i)}{\|j_p - j_i\| \cdot \|j_c - j_i\|} \quad (1)$$

The direction of limb extension is represented by a unit vector connecting keypoints.

Subsequently, spatial association weights between keypoints and garment vertices are constructed. For each garment vertex, its Euclidean distance $d_{k,i} = \|v_k - j_i\|$ to all keypoints is calculated, and the inverse distance weight is defined as formula (2):

$$w_{k,i} = \frac{1/(d_{k,i} + \epsilon)}{\sum_{i'=1}^{24} 1/(d_{k,i'} + \epsilon)} \quad (2)$$

$\epsilon = 10^{-6}$ is used for numerical stability. Furthermore, for the trunk and limb regions, orthogonal projection weights are introduced: vertices are projected onto the corresponding limb axes; if the projection point lies within a joint segment, the association weight is enhanced. A constraint feature vector $c_k \in \mathbb{R}^{d_c}$ is generated for each vertex by weighted aggregation of keypoint dynamic features (formula 3).

$$c_k = \sum_{i=1}^{24} w_{k,i} \cdot \phi(\theta_i, d_i) \quad (3)$$

$\phi(\cdot)$ is a learnable embedding function that maps the scalar bending angle and direction vector to a unified feature space. This constraint field $C = \{c_k\}_{k=1}^{N_v}$ serves as the guiding signal for subsequent graph convolutional networks, explicitly encoding the geometric constraints of pose on local deformation.

ADAPTIVE GRAPH CONVOLUTIONAL NETWORK DESIGN

To overcome the limitations of traditional graph convolutional networks, which rely on fixed topologies and struggle to adapt to the diversity of local wrinkles and deformations under varying human poses, this paper designs an adaptive graph convolutional network. This network achieves dynamic responses of the graph structure to local geometric features and human pose semantics. Through a

gating mechanism, it adaptively fuses geometric context and pose constraint information, ensuring that the information propagation path maintains physical rationality and detail sensitivity under different deformation scenarios.

Given the initial clothing vertex coordinates $X^{(0)} \in \mathbb{R}^{N_v \times 3}$, the clothing graph structure $G = (V, E)$, and the keypoint constraint features $C \in \mathbb{R}^{N_v \times d_c}$ as input, the vertex displacement field $\Delta X \in \mathbb{R}^{N_v \times 3}$ is predicted step by step through multi-layer graph convolution.

First, an adaptive adjacency matrix $A^{adapt} \in \mathbb{R}^{N_v \times N_v}$ is constructed. Let the original adjacency matrix A be defined by the first-order adjacency relation of the grid, where its element $A_{ij} = 1$ if and only if vertices i and j are adjacent. Based on this, two dynamic weights are introduced:

1) Geometric weights based on local curvature (Formula 4)

$$w_{ij}^{curv} = \exp(-\alpha \cdot |\kappa_i - \kappa_j|) \quad (4)$$

κ_i is the discrete mean curvature of vertex i , and α is the scaling factor.

2) Spatial weights based on keypoint distance (Formula 5)

$$w_{ij}^{dist} = \frac{1}{1 + \beta \cdot \|c_i - c_j\|_2} \quad (5)$$

β controls the decay rate. The final adjacency weight is defined by formula (6):

$$A_{ij}^{adapt} = A_{ij} \cdot w_{ij}^{curv} \cdot w_{ij}^{dist} \quad (6)$$

Subsequently, a fusion-gated attention module is designed. In the l -th layer, the vertex features $H^{(l)} \in \mathbb{R}^{N_v \times d_h}$ (initial $H^{(0)} = [X^{(0)}, C]$) are updated through graph convolution (formula 7):

$$Z^{(l)} = \sigma(D^{-1/2} A^{adapt} D^{-1/2} H^{(l)} W^{(l)}) \quad (7)$$

D is the degree matrix of A^{adapt} , $W^{(l)}$ is the learnable weight, and σ is the ReLU activation. To fuse constraint features and geometric context, a gating mechanism is introduced (formula 8):

$$\begin{aligned} G^{(l)} &= \sigma(H^{(l)} W_g^{(l)}), \\ H^{(l+1)} &= G^{(l)} \odot Z^{(l)} + (1 - G^{(l)}) \odot H^{(l)} \end{aligned} \quad (8)$$

$\odot$ represents element-wise multiplication, and $W_g^{(l)}$ is the gating parameter. This mechanism dynamically adjusts the contribution ratio of constraint information and geometric propagation.

The network consists of four stacked graph convolution layers. The last layer outputs the displacement field $\Delta X = \text{MLP}(H^{(4)}) \in \mathbb{R}^{N_v \times 3}$, where MLP (Multilayer Perceptron) consists of two linear transformation layers with no activation function to preserve the linear additivity of the displacement. The physical plausibility of the predicted deformation is enforced through a multi-faceted approach: geometrically, by the keypoint constraint field which aligns garment motion with anatomical limits; structurally, by the

adaptive graph convolution whose curvature-aware weights discourage unnatural bending; and explicitly, through the dedicated post-processing steps (Laplacian smoothing and penetration correction) that resolve local non-physical artifacts. While the training data provides a prior for plausible wrinkle distributions, these integrated mechanisms actively constrain the output space to be physically consistent. This dynamic weighting mechanism directly impacts wrinkle formation. The curvature weight ensures that information propagates more strongly between vertices undergoing similar bending, promoting the formation of coherent wrinkle ridges. The keypoint distance weight strengthens connections in regions geometrically close to moving limbs, allowing pose-driven deformation to propagate locally while attenuating influence over distant areas. This adaptability breaks the rigidity of a fixed topology, enabling the graph structure to evolve in response to both local geometric detail and global pose semantics, which is crucial for generating diverse and context-aware wrinkles.

Figure 2 shows the data flow and core components of a single-layer graph convolutional network. The input vertex features and constraint features first enter the adaptive adjacency matrix module, where ‘‘curvature weights’’ and ‘‘distance weights’’ work together to generate dynamically adjusted edge weights. Subsequently, these weights drive the graph convolution operation, generating preliminary geometric features. Finally, the gating mechanism dynamically calculates the fusion ratio based on the current state, weightedly fusing the geometric features and constraint features to output updated features, providing input for the next layer’s calculation.

GENERATION AND POST-PROCESSING OF 3D CLOTHING FOLDS

The predicted vertex displacement field $\Delta X \in \mathbb{R}^{N_v \times 3}$ is directly superimposed with the original un-folded clothing mesh vertex coordinates $X^{(0)}$ to generate a preliminary folded mesh (Formula 9):

$$X^{pred} = X^{(0)} + \Delta X \quad (9)$$

To suppress non-physical high-frequency jitter caused by data-driven errors, a lightweight Laplacian smoothing regularization term is applied at the vertex level. The Laplacian coordinates are defined as $L(X) = X - MX$, where M is a normalized weight matrix constructed based on the original mesh adjacency relationships, and the element $M_{ij} = 1/|N(i)|$ ($j \in N(i)$ is the first-order neighborhood of vertex i). X^{pred} is corrected by minimizing the following energy function (formula 10):

$$X^{smooth} = \arg \min_X \|X - X^{pred}\|_F^2 + \lambda \|L(X) - L(X^{(0)})\|_F^2 \quad (10)$$

$\lambda = 0.1$ controls the smoothing intensity; the analytical solution can be obtained by solving a sparse linear system, and the computational cost is controllable.

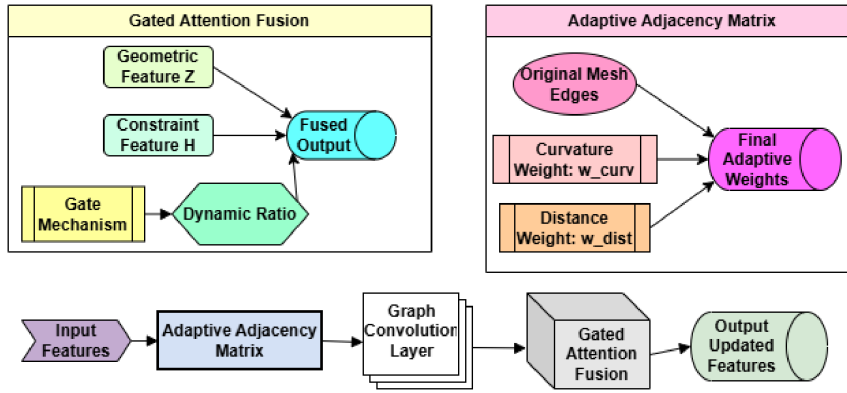


FIGURE 2. Adaptive graph convolutional network structure diagram

Furthermore, keypoint-based penetration detection and geometric correction are performed. For each clothing vertex $x_k \in X^{smooth}$, the shortest distance to the limb axis defined by the human skeleton diagram is calculated. If this distance is less than a preset threshold (set to 5 mm) and located in the internal direction of the limb, it is determined to be a penetration. For the penetrating vertex, a displacement offset is applied in the opposite direction of its normal (formula 11):

$$x_k^{final} = x_k + \eta \cdot n_k \quad (11)$$

n_k is the local normal to the vertex (calculated from the original mesh patch), and η is the adaptive offset, taken as the larger of the penetration depth and 2 mm. This functionality is only activated when penetration occurs, so as to not introduce deformation bias into the non-penetrated space. The output is the geometrically corrected 3D clothing mesh X^{final} that preserves predicted fold details and still satisfies basic physical constraints.

III. EXPERIMENTAL SETUP

DATASET AND DATA PARTITIONING

The study employed both a public dataset and a self-constructed dataset, specifically CAPE and MPG [30-31]. The CAPE dataset included 15 participants, each captured in approximately 300 registered clothing meshes in a sequence of postures; it included tops and bottoms and physical actions like walking, jumping, and lifting arms. The self-built MPG dataset was collected at the same time from an optical motion capture system and structured light scanner, comprising three-dimensional scans of clothing worn by 8 participants in four types of fabric, collected in 120 predetermined poses which were engineered to capture high deformation scenarios. The MPG dataset construction adhered to a strict protocol to prevent data leakage and ensure unbiased evaluation. Data collection involved eight individuals with distinct body shapes (BMI range 18.5-28.7) not overlapping with CAPE subjects. Four fabric types were selected to cover a range of physical properties: tightlyfitted stretch knit, loose cotton, structured denim, and silk. Each participant performed a scripted sequence

of 15 high-deformation poses per fabric, captured using a synchronized Vicon optical motion capture system (120Hz) and an Artec Leo structured-light scanner. The raw scans were manually cleaned, registered to the SMPL topology using non-rigid ICP, and then remeshed to the unified template with 6890 vertices. This process resulted in 3840 unique garment mesh samples (8 participants $\times$ 4 fabrics $\times$ 120 poses). The dataset was partitioned at the subject level to ensure identity independence between training and test sets. All datasets were uniformly processed to a T-pose template topology which ensured vertex correspondence across samples with a vertex topology containing 6890 vertices. Human poses were parameterized using the SMPL model and subject keypoints were extracted by the SMPL joint regressor including 24 distinct body joints. The data partitioning scheme involved a dual isolation process of “per person + per action” - for CAPE, 10 people of the 15 total were for training, 2 for validation, and 3 for testing, and for MPG, 5 of 8 people were used for training, 1 for validation, and 2 for testing. For each participant in both datasets, the pose sequences were strictly partitioned at the sequence level. The training set contained 70% of the pose sequences, the validation set 10%, and the test set 20%. This partitioning was applied after grouping all samples by participant, ensuring that no poses from the test participants were seen during training or validation, thereby eliminating the risk of bias from data leakage. This study was approved by the Ethics Committee of Zhejiang Fashion Institute of Technology, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

COMPARISON METHOD SELECTION

To thoroughly assess the benefits of the methods with respect to geometric accuracy, physical plausibility, and level of detail, three baseline methods were selected as representative of the broader category, including: physics-driven methods, general data-driven methods, and unconstrained graph network methods. The physics-driven method uses ClothSim, which achieves high-fidelity fabric simulation based on a massspring model and collision detection. It

uses the same SMPL human body mesh as presented in this paper as the collider, with a time step of 1/60 second, iterating until static convergence [32-33]. The general data-driven method uses PM-Cloth, a variant of Pixel2Mesh, to render human keypoints as depth maps. Pose features are extracted using a CNN encoder, and then the clothing vertex displacements are regressed using an upsampling module. The unconstrained graph network method uses the standard GCN architecture (GCN-Baseline), taking the clothing mesh structure as input, and only using vertex coordinates and adjacency relationships for multilayer graph convolution, without introducing any keypoint guiding signals.

The three baseline methods described above correspond to typical limitations in current technologies: ClothSim is “highly accurate but inefficient,” PM-Cloth is “efficient but neglects geometric coupling,” and GCN-Baseline is “structure-aware but lacks attitude guidance.” Furthermore, to ensure a comprehensive comparison with recent state-of-the-art learning-based approaches, we include two additional strong baselines published recently: Self-Adaptive Clothing Mapping (SACM) [25], which employs a self-adjusting graph network for virtual try-on, and CTSN [21], a two-stream skinning network that explicitly models the skeleton-cloth relationship for deformation prediction. These methods represent contemporary advances in graph-based and pose-aware garment modeling. Table 1 shows the core characteristics and limitations of each method:

EVALUATION METRICS

To fully assess the performance of the prediction results in three dimensions in terms of geometric accuracy, physical rationality, and wrinkle detail, the following six metrics are introduced:

Vertex Error: It is described as the average Euclidean distance (in millimeters (mm)) (Formula 12) between the predicted mesh and the scanned mesh corresponding vertices [34-35]:

$$E_{vert} = \frac{1}{N_v} \sum_{i=1}^{N_v} \|x_i^{pred} - x_i^{gt}\|_2 \quad (12)$$

Chamfer Distance (CD) is used to measure the overall shape similarity between two meshes, and calculates the average distance between the nearest points in both directions (formula 13):

$$CD = \frac{1}{N_v} \sum_{x \in X^{pred}} \min_{y \in X^{gt}} \|x - y\|_2 + \frac{1}{N_v} \sum_{y \in X^{gt}} \min_{x \in X^{pred}} \|x - y\|_2 \quad (13)$$

Penetration Ratio is the percentage of vertices in the clothing mesh that penetrate the human body surface. Penetration is determined based on the SMPL human body mesh: if a vertex is located inside the human body and its distance to the nearest human body surface is less than 0,

it is considered to have penetrated. The penetration ratio is calculated using formula (14):

$$R_{pen} = \frac{|\{i \mid dist(x_i^{pred}, M_{body}) < 0\}|}{N_v} \times 100\% \quad (14)$$

Curvature consistency is measured by the L2 norm of the difference in average curvature between the predicted mesh and the real mesh at corresponding vertices (formula 15):

$$E_{curv} = \frac{1}{N_v} \sum_{i=1}^{N_v} |\kappa_i^{pred} - \kappa_i^{gt}| \quad (15)$$

Curvature is estimated using discrete differential geometry based on the normal of neighboring patches. High-frequency energy ratio is used to assess the richness of wrinkle details. The vertex displacement field is decomposed using Laplacian spectral decomposition, and the proportion of energy of high-frequency components (corresponding to the top 10% of the largest eigenvalues) to the total energy is calculated (formula 16).

$$R_{hf} = \frac{\sum_{k \in K_{hf}} \lambda_k \|u_k^\top \Delta X\|^2}{\sum_{k=1}^{N_v} \lambda_k \|u_k^\top \Delta X\|^2} \quad (16)$$

λ_k and u_k are the eigenvalues and eigenvectors of the Laplacian matrix. The standard deviation of local

curvature focuses on high-frequency wrinkle regions (such as the elbow, scapula, and waist), calculating the standard deviation of the curvature at the vertices of these regions to reflect the intensity of detail changes (formula 17).

$$\sigma_{local} = \sqrt{\frac{1}{N_r} \sum_{i \in R} (\kappa_i - \bar{\kappa}_R)^2} \quad (17)$$

R represents the predefined set of vertices in the high-deformation region, and $\bar{\kappa}_R$ represents its average curvature.

These indicators together form the systematic assessment system of geometrical fidelity, physical compliance, and detail presentation of the predictions. The model uses a 4-layer adaptive graph convolution (with a hidden dimension of 256), an Adam optimizer (with an initial learning rate of 10^{-3} with cosine annealing), a batch size of 16, and employs early stopping based on Chamfer distance on the validation set. The Laplacian smoothing

coefficient is set at 0.1, the penetration detection threshold is 5 mm, and the upper limit of the geometric correction offset is 2 mm. All hyperparameters are selected through a grid search on the validation set.

IV. RESULTS ANALYSIS

OVERALL PERFORMANCE COMPARISON ANALYSIS

To quantitatively evaluate the performance of the proposed method within the domain of explicit, graphbased garment deformation, a systematic comparison is conducted against several representative baseline methods, ClothSim (physical

TABLE 1. Comparison of Characteristics of Comparative Methods

Method	Physics-Driven	Uses Keypoints	Graph-Based	Real-Time	Main Limitation
ClothSim	Yes	No	No	No	High computational cost; not suitable for real-time applications
PM-Cloth	No	Yes	Partially	Yes	Ignores explicit human-garment geometric coupling
GCN-Baseline	No	No	Yes	Yes	Lacks pose-aware guidance; prone to non-physical deformations
Ours	No (post-processed)	Yes	Yes	Yes	— (Balances physical plausibility, detail fidelity, and efficiency)

Notes: The “Graph-Based” designation for PM-Cloth is ascribed a Partially rating since although, there is a graph up-sampling module, pose is processed through a CNN applied to rendered images, without treating the pose as a skeletal or garment graph processing step. In the proposed approach, the “Physics-Driven” is rated No, but does consist of light geometric post-processing (e.g., penetration correction) in order to enforce physical plausibility.

simulation), PM-Cloth (general data-driven), GCN-Baseline (unconstrained graph network), and Ours (this work), on CAPE and the self-built MPG test sets. Important quantitative measures consist of vertex displacement error, Chamfer distance, penetration rate, high-frequency energy ratio, and local curvature standard deviation. The results can be seen in Table 2.

As summarized in Table 2, proposed method exhibits optimal or near-optimal performance for every primary metric and clearly outperforms every baseline.

For geometric accuracy, and in terms of vertex displacement error, the proposed method rates 3.82 mm, a 19.6% decrease from the optimal baseline GCN-Baseline, which exceeded 4.75 mm. Likewise, the same metric, Chamfer distance, decreased from 5.21 mm to 4.04 mm, demonstrating a decrease of 22.5%. Both represent that proposed framework can more correctly represent overall deformation, as well as local structure for the clothing mesh. Among recent methods, the proposed method also demonstrates advantages. Compared to SACM, our approach reduces vertex error by 9.3% and Chamfer distance by 18.9%. While CTSN achieves a competitive geometric accuracy, the proposed method yields a lower vertex error and maintains a superior high-frequency energy ratio (0.372 vs. 0.341) and local curvature standard deviation (0.183 vs. 0.169), indicating enhanced capability in preserving and synthesizing fine wrinkle details.

Regarding physical plausibility, the proposed method achieves a penetration rate of 0.71%. While this represents a substantial reduction compared to PM-Cloth (2.41%) and GCN-Baseline (2.84%), any penetration constitutes a non-physical artifact. This residual penetration primarily occurs in regions of extreme articulation or under very tight clothing, where the post-processing correction (Section 2.4) may be insufficient due to discrete vertex sampling and the limitations of the linear offset model.

For detail representation, proposed method exhibited the “high frequency energy ratio” of 0.372, also the highest of any other method, which indicates a richer representation of texture-level detail in its predicted wrinkles. At the same time, “local curvature standard deviation” is shown as 0.183, a 15.8% increase over ClothSim (0.158), thus validating its ability to represent complex multi-layered wrinkles in areas of high curvature (such as the elbow and scapula).

ABLATION STUDY OF CORE MECHANISM

To rigorously assess the central importance of the dynamic constraint function for human keypoints to improve prediction accuracy and physical plausibility, this paper performed three systematic ablation studies. The quantitative results are shown in Figure 3. Figure 3 includes assessment of all three study dimensions of performance improvement from keypoint constraint: geometric accuracy, physical plausibility, and posture deformation consistency.

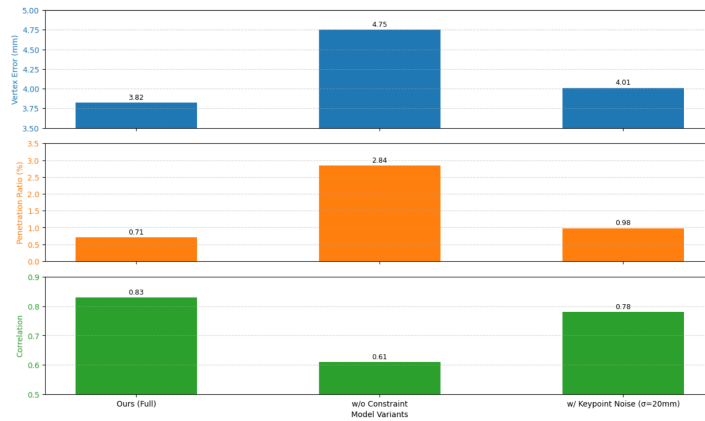
With respect to geometric accuracy, the upper bar chart in Figure 3 shows that the vertex error of the complete model (Ours) is stable at 3.82 mm. In comparison, the model error experiences a notable rise, reaching 4.75 mm when the keypoint constraints are bypassed (w/o Constraint), resulting in a 24.3% increase. These results illustrate clear evidence that keypoint constraints are an important factor in achieving high-accuracy geometric predictions, and they also provide essential geometric prior knowledge concerning the deformation of the clothing vertices.

Concerning physical plausibility, the middle bar chart in Figure 3 examines the clothing-human body penetration rate. The complete model maintained control of the penetration rate at a low 0.71% while the penetration rate of the unconstrained model swelled to 2.84%, nearly 4 times that of the complete model. This indicates that the keypoint constraint field assists with guiding the clothing mesh away from the human body surface and thus acts as a basic physical constraint when modeling within a data-driven framework, and avoids nonphysical encroachment of fabric into the human surface, something that pure data-driven methods cannot enforce.

With respect to attitude-deformation consistency, the bar chart at the bottom of Figure 3 illustrates the correlation coefficient between joint bending angles and wrinkle direction in adjacent areas. The full body model achieved an average correlation coefficient of 0.83 in significant locations (e.g.: elbow and knee joints) indicating extremely strong attitude guidance capability. After removing constraints, the correlation coefficient dropped to 0.61, dropping 26.5%. This strongly indicates that the keypoint constraints contribute to not only geometric position but also to ensure that the wrinkle morphology that is generated matches human posture and movement highly synchronously to achieve true “posture driven” performance. In addition, the examples

TABLE 2. Summary of Quantitative Results of the Main Experiment

Method	Vertex Error (mm)	Chamfer Distance	Penetration Ratio (%)	High-Freq Energy Ratio	Local Curvature Std
ClothSim	4.1	5.21	0.35	0.31	0.158
PM-Cloth	5.02	5.87	2.41	0.298	0.142
GCN-Baseline	4.75	5.21	2.84	0.315	0.159
Ours	3.82	4.04	0.71	0.372	0.183
SACM	4.21	4.98	1.12	0.343	0.162
CTSN	4.05	4.52	0.89	0.341	0.169

**FIGURE 3.** Comparison of Keypoint Constraint Ablation Experiment Results

shown in Figure 3 indicate that the model shows robustness. Even with noise present at the input keypoints ($\sigma = 20$ mm), the model output vertex error only increases slightly to 4.01 mm and the penetration rate remains at 0.98% with a correlation coefficient of 0.78. Overall, these results indicate that the proposed keypoint constraint mechanism shows good fault tolerance for maintaining stability and reliability in the prediction results even with uncertainty in the input signal.

To precisely measure the amount to which Adaptive Graph Convolution Network (Adaptive-GCN) enhances the model performance, this paper also conducts a strict control experiment: directly comparing the proposed Adaptive-GCN with a baseline model that utilizes a static adjacency matrix (Fixed-GCN). Fixed-GCN leverages only the original mesh topology to specify connections with edges and a uniform weight of 1, thereby isolating the performance benefit provided by the “adaptive” nature. The two models were assessed on the CAPE test set and the comparative results of the core metrics are shown in Figure 4.

(Note: Values in the figure are specific measurements; color intensity reflects numerical magnitude for visual comparison.)

In terms of global geometric accuracy, Figure 4 shows that the chamfer distance of Adaptive-GCN is 4.04 mm, while that of Fixed-GCN is as high as 5.21 mm. This indicates adaptive adjacency weights improves the overall shape fidelity of the model by 22.5%, effectively closing the gap between the predicted mesh and actual scanned mesh. Therefore, the ability to dynamically adjust the information propagation path is imperative in capturing complex cloth-

ing deformations. In terms of local detail representation, the curvature standard deviation index seen in Figure 4's high-frequency region highlights a deeper advantage. The Adaptive-GCN value is 0.183, while the Fixed-GCN gave a score of only 0.151, ultimately 21.2% higher due to the former. This improvement directly shows the ability of the adaptive mechanism to demonstrate more richness and dramatic detail in wrinkle formations at local high curvature regions, like the elbow and behind the knee, effectively alleviating the over-smoothing effect of the traditional GCN by using a fixed topology. Moreover, Figure 4 elucidates the performance difference from the perspective of feature learning through an attention concentration index. In the “flexed arm” position, Adaptive-GCN exhibits an average attention weight of 0.83 in key deformation areas, specifically under the armpit and shoulder regions, while Fixed-GCN has an average attention weight of only 0.52. With such extensive differences in data (an improvement of 59.6%), this provides strong evidence that the adaptive mechanism can effectively focus on and promote features lower in the attention regions that the dynamic adjacency and attention process produced, thus successfully optimizing physical plausibility, and visual details at once.

GENERALIZATION AND ROBUSTNESS ANALYSIS

In the context of assessing the generalization capacity of the proposed method with new identities and different clothing types, a stringent set of cross-domain tests were conducted. Each test set two new subjects (neither of whom appeared in the training data) and two new clothing types, while all test sets contained high-deformation movement (i.e. squatting

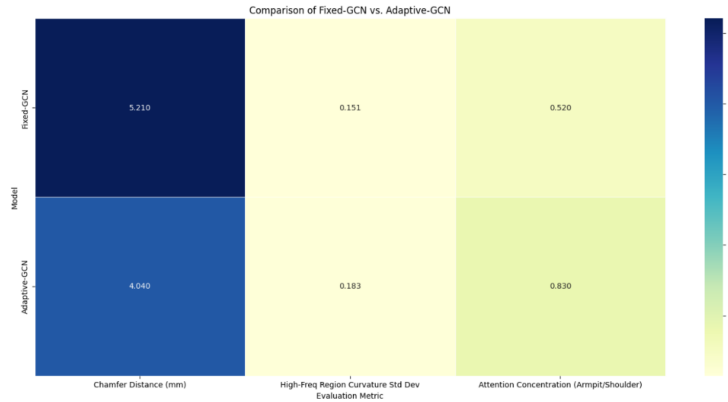


FIGURE 4. Quantitative performance comparison heatmap of Fixed-GCN and Adaptive-GCN

and crossing arms). This experiment evaluated if the model can use keypoint constraints and graph structure priors to decouple differences in individual body shape from the coupling of posture and clothing, thereby enabling the model to robustly predict “new objects.” Its main performance metrics are captured in Figure 5.

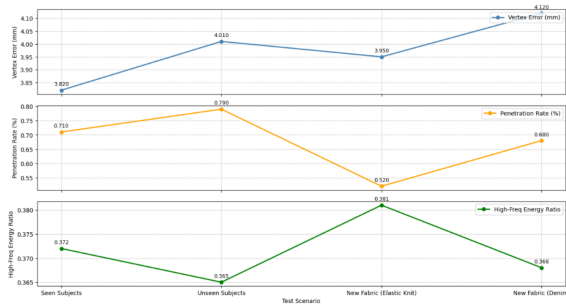


FIGURE 5. Quantitative Analysis of Generalization Ability Across Persons and Clothing Types

In terms of generalization to other people, as seen in Figure 5, the model doubled the vertex error from 3.82 mm for training distribution, to 4.01 mm for unseen persons. This increment of only 5.0% indicates that the model generalizes these core features of posture-driven behavior through dynamic constraints of human keypoints rather than learning body shape characteristics of a specific person. Also, the penetration rate was increased from 0.71% to 0.79% due to a manageable increase, further indicating the continued physical plausibility of the model. In terms of cross-clothing type transformation, the model exhibited good reliability and adaptation to fabric property differences. For tightly fitted stretchy knitware, the model minimized the penetration rate to just 0.52%, proving it can control fit tension effectively. While the penetration rate for complex and complicated denim increased to 0.68%, it was still significantly lower than baseline methods (i.e. GCN-Baseline is 2.84%), indicating that the adaptive graph convolution mechanism can change the information pendulum as a function of the lose of the fabric, dealing with complex deformation effectively.

In the case of the new fabric (stretchy knit), the model high-frequency energy ratio produced was 0.381, even higher than the 0.372 case seen already, indicating strong generalization potential for creating rich texture details. In the more complex and complicated New fabric (denim) case, the high-frequency energy ratio was still 0.368, which is only slightly outside baseline performance, and highlighted the models excellent visual fidelity quality.

QUALITATIVE ANALYSIS OF WRINKLE TYPESS

A qualitative analysis is conducted to examine the types of wrinkles generated. As expected, the keypoint constraint mechanism excels at producing pose-dependent wrinkles, such as tensile wrinkles radiating from the elbow during flexion and compression wrinkles at the knee during bending. These are directly correlated with joint angles. The model also generates plausible drape wrinkles in loose garment regions, which are less directly pose-dependent. These emerge from the combination of gravity-implicit training data and the adaptive graph’s ability to propagate deformation from constrained regions. However, subtle wrinkles caused purely by fabric dynamics or external forces are not captured, as the model’s input lacks such information. This analysis confirms that the method’s strength lies in synthesizing pose-consistent macro and meso-scale folds, while micro-detail generation remains dependent on the training data distribution.

V. CONCLUSION

This paper introduces a three-dimensional clothing wrinkle prediction process that utilizes dynamic constraints on human keypoints alongside adaptive graph convolution to produce high-fidelity wrinkle generation in an end-to-end manner. The method achieves precise geometric alignment, maintains a low garment-body penetration rate, and enriches detail representation. By creating a pose-aware constraint field to control the direction of deformation in conjunction with dynamic adjacency weights and gated attention to adaptively attend node neighbors, this paper generated an effective representation of the pose-clothing coupling

relationship. The experiments demonstrate very significant improvement over established baseline methods adopted for both the CAPE and a self-compiled MPG datasets, evidencing a vertex error of 3.82mm, a penetration rate of 0.71%, and locality improvement in curvature standard deviation by 15.8 relative to the original ClothSim. Importantly, this paper's method demonstrates good cross-human and cross-clothing generalization and temporal consistency. Qualitative analysis confirms the method's efficacy in generating pose-dependent wrinkles and plausible drape folds, while identifying the generation of force-specific micro-wrinkles as a direction for future work. The results, quantified by reduced vertex error, Chamfer distance, and controlled penetration rate alongside enhanced curvature variation, validate the proposed method as an effective framework for generating high-fidelity 3D clothing wrinkles with improved geometric and physical consistency under complex poses. A limitation of the current framework is the residual garment-body penetration (0.71%), which arises from the discrete nature of mesh correction and the absence of a continuous collision constraint in the core learning objective. Future work will integrate a differentiable penetration penalty term into the training loss and explore continuous collision detection to further minimize this non-physical artifact.

AUTHOR CONTRIBUTIONS

Wu Chao Quan designed, collected and analyzed the data, and drafted the manuscript. Wu Chao Quan conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Wu Chao Quan participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This article is part of the research project of Ningbo Key Laboratory (Project No. 2024ZDSYS-C-016: Development of Intelligent Clothing for Health Monitoring of Middle-aged and Elderly People).

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Ethics Committee of Zhejiang Fashion Institute of Technology, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] H. Liu, "Computer 5G virtual reality environment 3D clothing design," *Mobile Information Systems*, vol. 2022, no. 1, Art. no. 8024453, 2022, doi: 10.1155/2022/8024453.
- [2] H. Choi K, "3D dynamic fashion design development using digital technology and its potential in online platforms," *Fashion and Textiles*, vol. 9, no. 1, pp. 9, 2022, doi: 10.1186/s40691-021-00286-1.
- [3] S. Xu, L. Wu, W. Liao, et al., "Enhancing collaborative shopping experience through interactive per-sonalized avatars and shared gaze in a multi-user augmented reality environment," *International Journal of Human-Computer Interaction*, vol. 41, no. 11, pp. 6515-6536, 2025, doi: 10.1080/10447318.2024.2381923.
- [4] W. Chang and X. Xu, "Research on design thinking of pleated clothing inspired by fractal geometry," *International Journal of Clothing Science and Technology*, vol. 37, no. 1, pp. 22-47, 2025, doi: 10.1108/IJCST-11-2023-0167.
- [5] Q. Yu and G. Zhu, "Virtual simulation design of Mazu clothing based on digital technology," *Fibers and Polymers*, vol. 25, no. 7, pp. 2773-2787, 2024, doi: 10.1007/s12221-024-00566-9.
- [6] S. Huang and L. Huang, "CLO3D-based 3D virtual fitting technology of down jacket and simulation re-search on dynamic effect of cloth," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, Art. no. 5835026, 2022, doi: 10.1155/2022/5835026.
- [7] P. Boisse, J. Huang, and E. Guzman-Maldonado, "Analysis and modeling of wrinkling in composite forming," *Journal of Composites Science*, vol. 5, no. 3, pp. 81, 2021, doi: 10.3390/jcs5030081.
- [8] A. Kadi H and K. Terzić, "Data-driven robotic manipulation of cloth-like deformable objects: The present, challenges and future prospects," *Sensors*, vol. 23, no. 5, pp. 2389, 2023, doi: 10.3390/s23052389.
- [9] M. Shi, X. Zhuo, T. Mao, et al., "Automatic temporal clothing deformation prediction: applicable to variable shape and motion," *Textile Research Journal*, vol. 93, no. 21-22, pp. 4804-4819, 2023, doi: 10.1177/00405175231178049.
- [10] Y. Zhu, M. Zhang, Y. Peng, et al., "Detailed wrinkle generation of virtual garments from a single im-age," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 4053-4071, 2021, doi: 10.1007/s11042-020-09917-z.
- [11] T. Sha, W. Zhang, T. Shen, et al., "Deep person generation: A survey from the perspective of face, pose, and cloth synthesis," *ACM Computing Surveys*, vol. 55, no. 12, pp. 1-37, 2023, doi: 10.1145/3575656.
- [12] D. Chen, Y. Song, F. Liang, et al., "3D human body reconstruction based on SMPL model," *The Visual Computer*, vol. 39, no. 5, pp. 1893-1906, 2023, doi: 10.1007/s00371-022-02453-x.
- [13] H. Wu, "Research on Virtual Human Motion Generation Based on Biological Skeletal Models," *Journal of Commercial Biotechnology*, vol. 30, no. 3, pp. 976-986, 2025.
- [14] M. Fisch and R. Clark, "Orientation keypoints for 6D human pose estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 10145-10158, 2021, doi: 10.1109/TPAMI.2021.3136136.
- [15] C. Xu, R. Liu, T. Zhang, et al., "Dual-stream structured graph convolution network for skeleton-based action recognition," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 17, no. 4, pp. 1-22, 2021, doi: 10.1145/3450410.
- [16] A. Mouhou A, A. Saaidi, B. Yakhlef M, et al., "Wrinkle synthesis for cloth mesh with Hermite radial basis functions," *Multimedia Tools and Applications*, vol. 80, no. 1, pp. 1583-1610, 2021, doi: 10.1007/s11042-020-09743-3.
- [17] A. Mrinali and P. Gupta, "Cloth-net: improved hybrid adversarial network with dense vision transformer for 2D-3D image classification for accurate cloth recommendation engine," *International Journal of Clothing Science and Technology*, vol. 37, no. 3, pp. 402-442, 2025, doi: 10.1108/IJCST-05-2024-0110.
- [18] M. Mao, H. Va, and M. Hong, "Video classification of cloth simulations: Deep learning and position-based dynamics for stiffness prediction," *Sensors*, vol. 24, no. 2, pp. 549, 2024, doi: 10.3390/s24020549.
- [19] X. Hu, C. Zheng, J. Huang, et al., "Cloth texture preserving image-based 3D virtual try-on," *The Visual Computer*, vol. 39, no. 8, pp. 3347-3357, 2023, doi: 10.1007/s00371-023-02999-4.
- [20] M. Mao, H. Va, A. Lee, et al., "Supervised Video Cloth Simulation: Exploring Softness and Stiffness Variations on Fabric Types Using Deep Learning," *Applied Sciences*, vol. 13, no. 17, pp. 9505, 2023, doi: 10.3390/app13179505.
- [21] Y. Li, M. Tang, Y. Yang, et al., "CTSN: Predicting cloth deformation for skeleton-based characters with a two-stream skinning network,"

- Computational Visual Media, vol. 10, no. 3, pp. 471-485, 2024, doi: 10.1007/s41095-023-0344-6.
- [22] H. Wang, "GPU-based simulation of cloth wrinkles at submillimeter levels," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 1-14, 2021, doi: 10.1145/3450626.3459787.
- [23] Z. Mo, Q. Wang, H. Shi, et al., "PointHuman: Reconstructing Clothed Human from Point Cloud of Parametric Model," *Computing and Informatics*, vol. 42, no. 2, pp. 457-479, 2023, doi: 10.31577/cai_2023_2_457.
- [24] H. Liu, L. Li, N. Yu, et al., "Outfit compatibility model using fully connected self-adjusting graph neural network," *The Visual Computer*, vol. 40, no. 11, pp. 8331-8343, 2024, doi: 10.1007/s00371-023-03238-6.
- [25] C. Shen, Z. Liu, X. Gao, et al., "Self-adaptive clothing mapping based virtual try-on," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 20, no. 3, pp. 1-26, 2023, doi: 10.1145/3613453.
- [26] S. Qi, X. Huang, P. Peng, et al., "Cascaded attention: Adaptive and gated graph attention network for multiagent reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 3, pp. 3769-3779, 2022, doi: 10.1109/TNNLS.2022.3197918.
- [27] D. Soydaner, "Attention mechanism in neural networks: where it comes and where it goes," *Neural Computing and Applications*, vol. 34, no. 16, pp. 13371-13385, 2022, doi: 10.1007/s00521-022-07366-3.
- [28] N. Kotlov, K. Shaposhnikov, C. Tazearslan, et al., "Procrustes is a machine-learning approach that removes crossplatform batch effects from clinical RNA sequencing data," *Communications Biology*, vol. 7, no. 1, pp. 392, 2024, doi: 10.1038/s42003-024-06020-z.
- [29] A. Andreella, R. De Santis, A. Vesely, et al., "Procrustes-based distances for exploring between-matrices similarity," *Statistical Methods & Applications*, vol. 32, no. 3, pp. 867-882, 2023, doi: 10.1007/s10260-023-00689-y.
- [30] E. Corona, G. Alenyà, G. Pons-Moll, et al., "LayerNet: High-resolution semantic 3D reconstruction of clothed people," *IEEE transactions on pattern analysis and machine intelligence*, vol. 46, no. 2, pp. 1257-1272, 2023, doi: 10.1109/TPAMI.2023.3332677.
- [31] M. Shi, W. Feng, L. Gao, et al., "Generating diverse clothed 3D human animations via a generative model," *Computational Visual Media*, vol. 10, no. 2, pp. 261-277, 2024, doi: 10.1007/s41095-022-0324-2.
- [32] K. Yoon and C. Lim S, "Real-to-sim high-resolution cloth modeling: Physical parameter optimization using particle-based simulation with robot manipulation data," *Journal of Computational Design and Engineering*, vol. 12, no. 8, pp. 29-44, 2025, doi: 10.1093/jcde/qwaf065.
- [33] Y. Chen, J. Wang Q, and M. Zhang, "Accurate simulation of draped fabric sheets with nonlinear model-ing," *Textile Research Journal*, vol. 92, no. 3-4, pp. 539-560, 2022, doi: 10.1177/00405175211039573.
- [34] V. Cohen-Addad, D. Das, E. Kipouridis, et al., "Fitting distances by tree metrics minimizing the total error within a constant factor," *Journal of the ACM*, vol. 71, no. 2, pp. 1-41, 2024, doi: 10.1145/3639453.
- [35] S. Hesarian M and J. Tavoosi, "Prediction the Effect of Tensile and Elongation Properties of Dry and Wet Cotton Fibre on the Fabric Wrinkle Grade Based on Polynomial Type2 Fuzzy Neural Network Model-ing," *Fibers and Polymers*, vol. 24, no. 6, pp. 2255-2262, 2023, doi: 10.1007/s12221-023-00229-1.