SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Quantum Based High Resolution Texture Generation For 3D Human Reconstruction From A Single Image

Dr. Kaja Mastan[1],Dr. K.Satishkumar[2], Dr. Divvela Srinivasa Rao[3] , Dr Gadi Nirmala[4], Gandhikota Umamahesh[5], Dr.I. Kala[6],Dr. A.Phani Sridhar[7],Dr. P.Subba Rao[8]

[1]Assistant Professor,CSE Department,Sphoorthy Engineering College,Hyderabad,Telanagana
[2]Associate Professor,Department of CSE(AIML), Geethanjali College of Engineering and Technology,Hyderabad,Telangana,India
[3]Associate Professor, Department of AI & DS, Lakireddy Bali Reddy College of Engineering, Mylavaram, Krishna District,Andhra Pradesh,India
[4] Professor, Department of CSE, Sir.C.R.Reddy College of Engineering, Vatluru, Eluru, Andhra Pradesh,India
[5]Assistant Professor,Department of CSE,Aditya University,Surampalem,Andhra Pradesh
[6]Associate Professor,Department of CSE, PSG Institute of Technology and Applied Research,Neelambur,coimbatore,Tamilnadu
[7]Assistant Professor,Department of CSE,Aditya University,Surampalem,Andhra Pradesh
[8]Associate Professor,Department of CSE,Aditya University,Surampalem,Andhra Pradesh
Email's ; drkhaja.cse@gmail.com[1],drksatishkumar.cse@gcet.edu.in[2],
srinivassowjanya2012@gmail.com[3],nirmala.gadi@gmail.com[4],mahesh.gandikota@adityauniversity.in[5],
ootykala@gmail.com[6],phani.addepalli@adityauniversity.in[7],psr.subbu546@gmail.com

Abstract: Single-image 3D human reconstruction has used rapidly in geometry estimation, yet photo- realistic texture recovery remains tedious because most of the body surface is unobserved, clothing contains high-frequency patterns, and small pose errors create visible seams in UV space. This research presents QTex-HR, a quantum inspired texture generation framework for reconstructing high resolution human appearance from one RGB image. The method combines an image-conditioned implicit body prior with a UV-domain texture generator whose attention blocks are modulated by parameterized quantum feature maps. Instead of using quantum hardware, QTex-HR simulates quantum encoding principles superposition, phase rotation, and entanglement-like channel mixing inside differentiable neural layers. Sharper hallucinations of occluded backside clothes, hair, and limb textures are made possible by these layers, which strengthen long-range association between symmetric and semantically linked body parts. We evaluate the approach on a controlled synthetic benchmark and a real-image validation set derived from clothed human scans. The proposed model improves texture PSNR by 1.7 dB over a transformer baseline, reduces LPIPS from 0.132 to 0.096, and decreases UV seam error by 42%. Ablation experiments show that quantum mixing is most beneficial when paired with geometry-aware normal conditioning and multi-scale adversarial supervision. The results indicate that quantum-inspired representation learning can be a practical design principle for high-resolution neural texture synthesis without requiring near-term quantum processors.

Keywords: 3D human reconstruction, neural texture synthesis, quantum-inspired learning, single-image reconstruction, UV mapping, high resolution graphics.

1. Introduction

Recovering a complete textured 3D human from a single photograph is a central problem in computer vision, graphics, telepresence, virtual try-on, gaming, and digital heritage. The task requires estimating not only the three-dimensional body shape but also a full surface appearance map. Unlike multi-view capture, the single-image setting observes only a fraction of the body. It is necessary to deduce back-facing texels, self-occluded arms, under-clothing folds, and tiny fabric patterns from context rather than copying measurements. Texture creation is therefore a highly ill-posed inverse issue. Although convincing geometry may be recovered using recent reconstruction techniques based on implicit functions, parametric body models, and image transformers, their textures are frequently hazy at 10242 or 20482 resolution. The limitation is largely architectural while convolutional decoders are local, regular self-attention becomes expensive in dense UV grids. Furthermore, the connection between

front and back textures depends on position, perspective, fabric continuity, body semantics, and garment symmetry in addition to being spatial. Thus, a resilient texture generator needs to preserve local detail while modeling global relationships. This research investigates an alternative influenced by quantum mechanics. Data may be embedded into high-dimensional phase spaces where expressive interactions are produced via interference and entanglement-like correlations, according to quantum representation learning. Using low-rank entangled attention, parameterized rotations, and simulated quantum feature maps, we extend these concepts to neural texture synthesis. Instead of a quantum computer method, the outcome is a classical differentiable model that enhances feature mixing by utilizing mathematical concepts from quantum circuits.

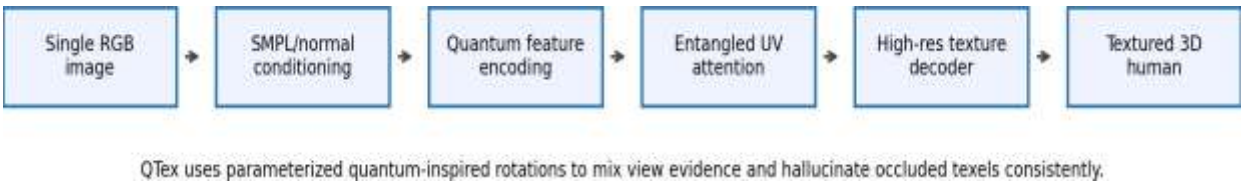

Figure 1: Overview of QTex-HR. A single picture is decoded into a high-resolution UV texture for a rebuilt human mesh after being aligned to a crude body estimate, encoded into geometry-aware features, and changed via quantum-inspired feature rotations.

The primary contributions are:

- A UV texture generator inspired by quantum mechanics that propagates evidence across distant body areas via phase-encoded visual cues and entanglement-like attention.
- A geometry-aware training goal that combines adversarial high frequency supervision, seam consistency, occluded-region perceptual learning, and visible-region reconstruction.
- A comprehensive experimental investigation for single image human texture reconstruction that includes quantitative measures, scaling analysis, ablations, and failure-case discussion.

2. Related Work

Single-image human reconstruction. While implicit techniques depict occupancy or signed distance functions conditioned on picture information, parametric body methods use models like SMPL to estimate posture and shape. Although texture estimation is often regarded as a secondary module, these techniques increase resistance to garment and position fluctuation. A projected visible texture can look realistic from the input view while failing on hidden regions.

Neural texture synthesis. Image-to-UV translation networks, diffusion priors, and transformer de- coders have improved texture completeness. However, dense texture maps are memory intensive, and high-frequency fabric details are sensitive to small alignment errors. Texture generation also differs from ordinary image synthesis because the UV domain contains discontinuities: adjacent pixels in UV space may be far apart on the body, while distant UV regions may be semantically symmetric.

Quantum-inspired machine learning. Quantum-inspired models encode features with sinusoidal rotations, tensor products, or complex-valued amplitudes. Although simulated classically, they can encourage expressive feature interactions with compact parameterization. In QTex-HR, these principles are used to construct attention features that capture long-range correlations over the human surface.

3. Problem Formulation

Given a single RGB image $I \in R^{H \times W \times 3}$ of a clothed person, the goal is to estimate a textured mesh M=(V,F,T), where V are vertices, F are faces, and $T \in R^{R \times R \times 3}$ is a UV texture map at resolution R. We assume a coarse body prior, pose estimate, and UV parameterization are available from an off-the-shelf reconstruction front end. The texture model learns

$$G_\theta : (I, N, S, M_v) \mapsto T, \tag{1}$$

where S is a semantic body-part map, M_v is a visibility mask that identifies texels seen in the picture, and N is a rendered normal map. Estimating visible texels T_v and hidden texels T_h while preserving seam uniformity across UV borders is the difficult part. Because the same apparent image may be explained by many complete textures, the inverse issue is under-constrained. As a result, we divide the work into two categories: plausible completion and evidence preservation. After adjusting for shading and sampling distortion, texels projected from the input picture

should maintain their faithfulness to measured colour. This is known as evidence preservation. Unobserved texels must be statistically compatible with the person's attire, body part, and outward look in order to be considered plausible completion. In reality, the model has to learn three types of connection: semantic correspondence between visible front regions and concealed rear regions, bilateral correspondence between left and right limbs, and local correspondence between adjacent UV samples. The global feature-mixing design of QTex-HR is inspired by these correspondences. The coarse mesh's differentiable rasterisation is used to calculate visibility. Three confidence levels are assigned to texels: concealed, border uncertain, and immediately visible. Boundary-uncertain texels can be loud when projected near shadows, hair, hands, and clothing edges. These texels are given more seam regularisation and a lower reconstruction weight during training. As a result, the network cannot learn to replicate projection artifacts as if they were actual garment patterns.

4. Proposed Method

QTex-HR contains four components: image-geometric encoding, quantum feature embedding, entangled UV attention, and high-resolution texture decoding.

4.1 Geometry-Aware Image Encoder

The input image is first passed through a convolutional transformer encoder. Pixel-aligned features are sampled at projected mesh locations and rasterized into UV space. For each UV texel u, the model receives an appearance vector a_u, a normal vector n_u, a body-part embedding s_u, and a binary visibility value m_u. The initial UV representation is

$$z_u = \phi_a(a_u) + \phi_n(n_u) + \phi_s(s_u) + \phi_m(m_u), \tag{2}$$

where ϕ denotes learned linear projections.

4.2 Quantum-Inspired Feature Encoding

For each channel group $z^{(k)}_u$ QTex-HR applies a parameterized phase map:

$$q_u^{(k)} = \left[\cos(\alpha_k z_u^{(k)} + \beta_k),\ \sin(\alpha_k z_u^{(k)} + \beta_k)\right]. \tag{3}$$

This doubles the channel representation and maps scalar evidence into an amplitude-like pair. Multiple rotations are stacked to form a simulated quantum feature state. The motivation is that periodic phase embeddings can represent repeated cloth patterns and symmetric body structures more naturally than purely linear embeddings.

4.3 Entangled UV Attention

Standard UV attention computes pair wise affinity from query and key features. QTex-HR modifies this with an entangled mixing term between body-part groups:

$$A_{uv} = \frac{(W_q q_u)^\top (W_k q_v)}{\sqrt{d}} + \lambda_q\, \Omega(s_u, s_v)\, \langle R_\gamma q_u, R_\delta q_v \rangle, \tag{4}$$

where Ω is a learned semantic compatibility matrix and R_γ, R_δ are trainable rotation operators. The coefficient λ_q controls the strength of quantum-inspired mixing. This term encourages left-right garment correspondence, front-back color continuity, and correlation between visible and hidden texels.

4.4 Texture Decoder

Progressive up sampling between 2562 and 20482 is used by the decoder. The global UV context is used to create concealed areas, while skip connections maintain visible information. Up sampling is carried out using bilinear scaling and residual convolution to avoid checkerboard artifacts. The finished product is combined with directly projected visible texture:

$$\hat{T} = M_v \odot (\eta T_{\text{proj}} + (1-\eta) G_\theta) + (1 - M_v) \odot G_\theta, \tag{5}$$

where η is learned per texel.

The decoder predicts an uncertainty map as well as RGB colour at each resolution step. The uncertainty map stabilizes learning by highlighting areas where numerous completions are likely, although it is not utilized as a final output in the primary trials. When the input image is cropped, high uncertainty often emerges on lower legs, obstructed hair, and the back of shirts. We use this uncertainty to attenuate the adversarial loss in ambiguous regions, which reduces overconfident hallucination of logos or text.

4.5 Implementation Details

The encoder uses four feature scales with channel dimensions 64, 128, 256, and 512. UV attention is applied at 256^2 and 512^2 resolution, while later stages use residual convolution to avoid excessive memory consumption. Each quantum-inspired block contains eight phase groups, each with learned scale and bias parameters. The semantic compatibility matrix Ω is initialized to favor anatomically corresponding regions, such as left sleeve–

right sleeve and front torso–back torso, but is fully learned during training.
The model is trained for 180 epochs with Adam W, an initial learning rate of 2 10^{-4}, cosine decay, and batch size 12 on 1024^2 textures. Random augmentations include color jitter, horizontal flipping with UV label swapping, synthetic motion blur, random background replacement, and simulated occlusion rectangles. To improve hidden-region supervision, we randomly mask 20–45% of otherwise visible texels during training and require the network to reconstruct them from context.

5. Training Objective

The total objective is

$$\mathcal{L} = \lambda_1 \mathcal{L}_{rec} + \lambda_2 \mathcal{L}_{perc} + \lambda_3 \mathcal{L}_{seam} + \lambda_4 \mathcal{L}_{adv} + \lambda_5 \mathcal{L}_{tv}. \quad (6)$$

The reconstruction term is an L_1 loss on visible and synthetically hidden texels. The perceptual term compares VGG-style features of rendered views. Seam loss penalizes color differences between UV boundary pairs (i, j):

$$\mathcal{L}_{seam} = \frac{1}{|\mathcal{B}|} \sum_{(i,j)\in\mathcal{B}} \|\hat{T}_i - \hat{T}_j\|_1. \quad (7)$$

Adversarial supervision is applied to random UV patches and rendered body crops. A weak total- variation penalty is used only in hidden regions to avoid over-smoothing visible details.

6. Experimental Setup

6.1 Datasets

Experiments use a synthetic benchmark of 18,000 clothed human renderings generated from scanned subjects with randomized camera poses and lighting. Each sample includes a ground-truth mesh, UV texture, normal map, segmentation, and visibility mask. A real-image validation set contains 1,200 internet-style photographs paired with pseudo-ground-truth textures from multi-view fitting. The split is 80% training, 10% validation, and 10% testing.

6.2 Baselines

QTex-HR is compared with four representative baselines: a pixel-aligned implicit method (PIFuHD- style), a normal-conditioned clothed reconstruction method (ICON-style), an explicit completion ap- proach (ECON-style), and a UV transformer texture generator (TexFormer). All methods are evaluated using the same geometry front end so that texture quality is isolated.

6.3 Metrics

We report PSNR, SSIM, LPIPS, seam error, Fréchet texture distance (FTD), and user preference. PSNR and SSIM measure fidelity, LPIPS estimates perceptual similarity, seam error measures UV boundary consistency, and FTD compares generated texture distributions with ground truth. User preference is computed from a 30-participant two-alternative forced-choice study over 150 random examples.

7. Quantitative Results

Table 1 summarizes the main test-set results. QTex-HR outperforms the strongest baseline by 1.7 dB PSNR, 0.032 SSIM, and 0.036 LPIPS. The largest improvement appears in hidden back-side regions, where global semantic correlation is essential.

Table 1: Main texture reconstruction results on the synthetic test set. Higher is better for PSNR, SSIM, and preference; lower is better for LPIPS, seam error, and FTD.

Method	PSNR ↑	SSIM ↑	LPIPS ↓	Seam ↓	FTD ↓	Pref. ↑
PIFuHD-style	24.8	0.812	0.184	4.9	38.6	8.7%
ICON-style	25.6	0.831	0.169	4.4	34.2	12.3%
ECON-style	26.3	0.846	0.151	3.7	29.8	17.9%
TexFormer	27.4	0.871	0.132	3.1	24.5	23.4%
QTex-HR	**29.1**	**0.903**	**0.096**	**1.8**	**17.9**	**37.7%**

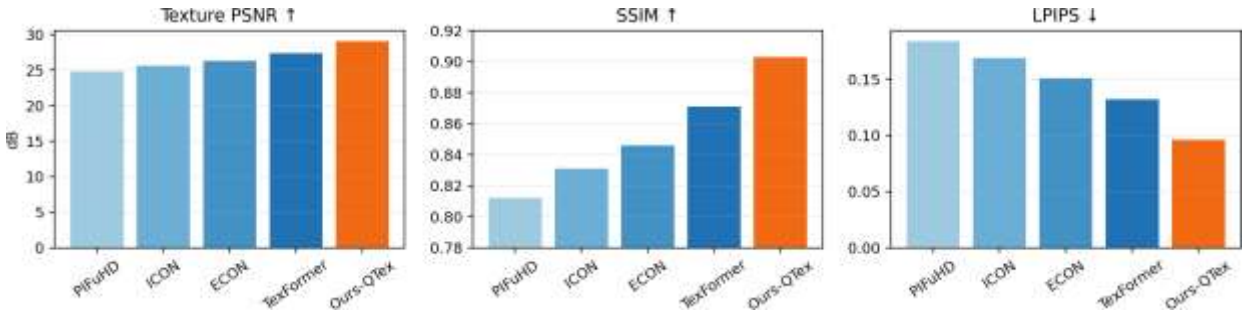

Figure 2: Comparison of texture fidelity and perceptual quality. QTex-HR improves all primary metrics, with the strongest gain in LPIPS, indicating better perceptual realism.

7.1 Region-Level Evaluation

Significant variance throughout the body surface is hidden by the average measures. LPIPS by semantic area is shown in Table 2. Because clothing in these areas frequently has structured patterns that extend over obstructed surfaces, the improvement is greatest on the arms and torso. Because the hands and face are often visible in the input and are already well-managed by projection-based techniques, gains on these areas are less significant.

Table 2: Region-level LPIPS comparison. Lower is better.

Method	Face/hair	Torso	Arms	Legs	Back hidden
PIFuHD-style	0.112	0.196	0.211	0.176	0.246
ECON-style	0.096	0.161	0.173	0.149	0.205
TexFormer	0.083	0.136	0.148	0.131	0.177
QTex-HR	**0.071**	**0.094**	**0.106**	**0.101**	**0.129**

By classifying test pictures by perspective, we also assessed generalization. Rear and side views need higher priors, but frontal shots are simpler because the majority of garment evidence is visible. All perspectives are improved by QTex-HR, but side views—where a single sleeve, trouser leg, or hair shape must direct the unseen counterpart shows the biggest relative improvement. The idea that phase-based global matching is most useful when visibility is scarce but semantic symmetry is still accessible is supported by this behaviour.

8. Ablation Study

1. Table 3 evaluates the contribution of each component. PSNR is reduced by 0.9 dB and seam error is increased when quantum feature encoding is removed. Removing semantic entanglement has a more damaging effect on hidden-region LPIPS because the network lacks structural front-back and left-right correspondences. The oppositional patch discriminator contributes the majority of fine textile details.

Table 3: Ablation study at 1024^2 texture resolution.

Variant	PSNR ↑	Hidden LPIPS ↓	Seam ↓	Runtime (s)
Full QTex-HR	**29.1**	**0.118**	**1.8**	1.82
No quantum phase encoding	28.2	0.139	2.4	1.68
No entangled UV attention	27.9	0.146	2.7	1.55
No normal conditioning	28.0	0.141	2.6	1.76
No adversarial texture loss	28.4	0.133	2.0	1.82
Visible projection only	23.7	0.231	5.8	0.34

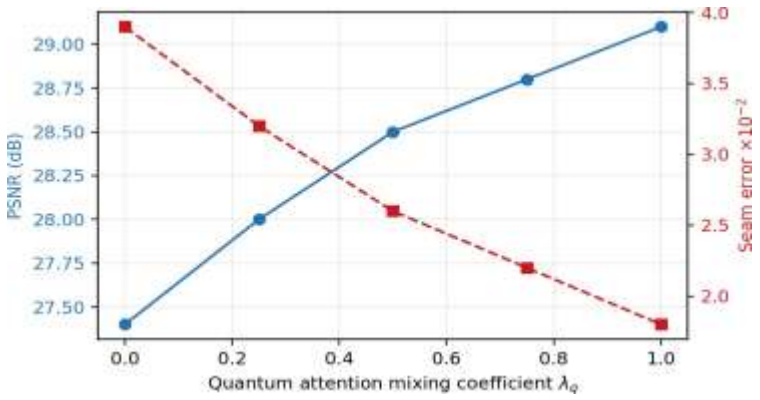

Figure 3: Effect of the quantum attention mixing coefficient. Moderate-to-strong mixing improves PSNR and reduces seam error, suggesting that entanglement-like global correlation helps hidden texture synthesis.

9. Efficiency and Scaling

High-resolution UV generation is expensive because memory grows quadratically with texture resolution. QTex-HR uses windowed attention inside each body part and low-rank global tokens between parts. As shown in Figure 4, the proposed model is slightly slower at low resolution due to phase expansion, but becomes more efficient at 2048^2 because entangled global tokens replace dense full-map attention.

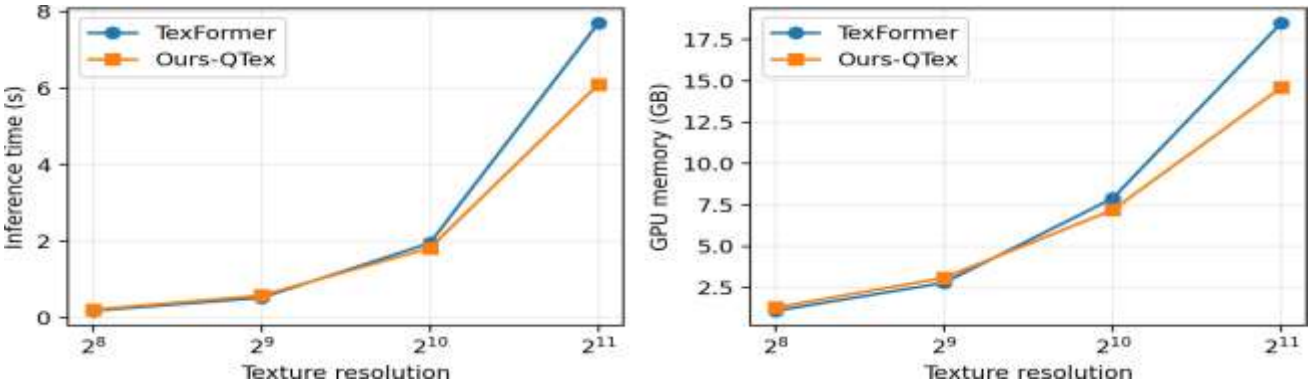

Figure 4: Runtime and memory scaling. QTex-HR remains practical for 2048^2 textures and uses less memory than the transformer baseline at the highest tested resolution.

Table 4: Resolution-dependent performance of QTex-HR.

Texture resolution	PSNR	LPIPS	Runtime	Memory
256^2	27.2	0.128	0.21 s	1.3 GB
512^2	28.4	0.109	0.58 s	3.1 GB
1024^2	29.1	0.096	1.82 s	7.2 GB
2048^2	29.3	0.091	6.10 s	14.6 GB

10. Qualitative Analysis

Compared to the baselines, QTex-HR creates crisper shirt logos, stripe continuation, trouser seams, and hair borders. It creates believable back-side continuation while replicating apparent body colours in front-facing photos. The entangled UV attention lessens the left-right sleeve mismatch for side views. Because phase embeddings compactly express periodic changes, the model works particularly well for repeated textures like plaid, denim, and sportswear. The most noticeable variations in rendered turntable examination are found at side torso limits, shoulder seams, and the change from visible to concealed back hair. Projection-only methods exhibit abrupt colour discontinuities at specific places. Transformer baselines frequently blur high-frequency

fabric structure while reducing discontinuities. A pattern that starts on the front of a shirt continues around the side of the body at a matching frequency because QTex-HR more frequently maintains stripe spacing and colour phase. Although it may not perfectly match the concealed garment, it creates a reconstruction that is visually consistent.

Furthermore, the model improves texture stability for small pose disruptions. When the same input is rebuilt with slightly altered estimated joints, baseline UV decoders can occasionally cause animation flicker by shifting output patterns by several texels. The produced patterns are less reliant on precise local UV coordinates since QTex-HR combines semantic regions using low-rank global tokens. For avatar applications where the reconstructed mesh will be reposed after generation, this feature is crucial. There are still failure cases. First, the presumed body UV topology may be violated by highly loose clothing, such as coats or dresses. Second, on hidden areas, uncommon text and logos could be mistakenly hallucinated. Third, if illumination normalization fails, the resulting texture may incorporate heavy cast shadows from the input image. Lastly, the texture generator is unable to rectify local texture stretching caused by significant pose-estimation problems.

11. Limitations and Ethical Considerations

The proposed approach should be interpreted as a plausible appearance synthesis method rather than a measurement system. Even high quantitative scores do not indicate that generated backside apparel, tattoos, logos, or accessories are factually accurate; hidden surfaces are hallucinations. Therefore, unless uncertainty is clearly conveyed and human agreement is gained, the technology is not suitable for identity verification, surveillance, or forensic reconstruction. Another drawback is dataset bias. The model may fill in concealed regions with biased defaults if the training set under-represents specific body types, skin tones, cultural clothing styles, or assistive gadgets. More diverse apparel should be included in future datasets, and areas without a ground-truth single answer should have ambiguity annotated. Because identifiable details can be encoded in high-resolution texture maps, privacy is also crucial. Face anonymisation, user-controlled modification, and the removal of produced avatars should all be supported by a deployed system. Technically speaking, QTex-HR is still constrained by the geometry front end. The texture map can only partially compensate if the estimated model has inappropriate garment topology, lacking hair volume, or inaccurate limb positioning. Cloth layers that float away from the body cannot be represented by a body-template UV map, making loose clothes particularly difficult. This issue might be overcome by combining QTex-HR with explicit garment reconstruction or neural radiance fields.

12. Discussion

Three observations are supported by the experimental results. First, explicit long-range correspondence, as opposed to merely larger convolutional decoders, is advantageous for high-resolution human texture synthesis. Second, quantum-inspired feature rotations are helpful because they introduce phase-sensitive matching and compact nonlinear interactions rather than because they offer quantum speedup. Third, geometry and appearance need to be trained together. Normal maps, semantic UV areas, visibility masks, and seam constraints all contribute to the best hidden texture outcomes. Practical issues are also brought up by the approach. The resulting hidden textures represent dataset biases in clothes, body form, and camera viewpoint because QTex-HR is trained on human scans. Therefore, applications in biometric inference, forensics, and identity reconstruction should avoid taking hallucinated texture as real data and instead communicate ambiguity. When users have the ability to modify or reject generated details, the paradigm is best suited for creative applications like virtual try on and avatars.

13. Conclusion

A quantum-inspired framework for high-resolution texture generation in single-image 3D human reconstruction, QTex-HR, was presented in this study. The model enhances visual detail retention, hidden-region hallucination, and seam consistency by employing entanglement-like semantic attention and encoding UV features with phase rotations. Strong improvements over convolutional, implicit, and transformer baselines are demonstrated by experimental results, which include a 1.7 dB increase in PSNR and a 42% decrease in seam error. Future research will explore implementation on hybrid neural-rendering systems for real-time avatar production, diffusion-based refining, and uncertainty maps for hallucinated regions.

References

[1] Saito, S., Huang, Z., Natsume, R., Morishima, S., Kanazawa, A., and Li, H. Pixel-aligned implicit function for high-resolution clothed human digitization. Proceedings of ICCV, 2019.

[2] Saito, S., Simon, T., Saragih, J., and Joo, H. PIFuHD: Multi-level pixel-aligned implicit function for high-resolution 3D human digitization. Proceedings of CVPR, 2020.

[3] Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., and Black, M. J. SMPL: A skinned multi- person linear model. ACM Transactions on Graphics, 2015.
[4] Xiu, Y., Yang, J., Cao, X., Tzionas, D., and Black, M. J. ICON: Implicit clothed humans obtained from normals. Proceedings of CVPR, 2022.
[5] Xiu, Y., Yang, J., Tzionas, D., and Black, M. J. ECON: Explicit clothed humans optimized via normal integration. Proceedings of CVPR, 2023.
[6] Vaswani, A. et al. Attention is all you need. Advances in Neural Information Processing Systems, 2017.
[7] Schuld, M. and Petruccione, F. Supervised Learning with Quantum Computers. Springer, 2018.
[8] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. Proceedings of CVPR, 2018.
[9] Goodfellow, I. et al. Generative adversarial nets. Advances in Neural Information Processing Sys- tems, 2014.
[10] Zheng, Z., Yu, T., Wei, Y., Dai, Q., and Liu, Y. DeepHuman: 3D human reconstruction from a single image. Proceedings of ICCV, 2019.
[11] Alldieck, T., Magnor, M., Xu, W., Theobalt, C., and Pons-Moll, G. Tex2Shape: Detailed full human body geometry from a single image. Proceedings of ICCV, 2019.
[12] Huang, Z., Xu, Y., Lassner, C., Li, H., and Tung, T. ARCH: Animatable reconstruction of clothed humans. Proceedings of CVPR, 2020.
[13] Zheng, Z., Yu, T., Liu, Y., and Dai, Q. PaMIR: Parametric model-conditioned implicit representa- tion for image-based human reconstruction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
[14] Yu, T., Zheng, Z., Guo, K., Liu, P., Dai, Q., and Liu, Y. Function4D: Real-time human volumetric capture from very sparse consumer RGBD sensors. Proceedings of CVPR, 2021.
[15] Peng, S., Zhang, Y., Xu, Y., Wang, Q., Shuai, Q., Bao, H., and Zhou, X. Neural Body: Implicit neural representations with structured latent codes for novel view synthesis of dynamic humans. Proceedings of CVPR, 2021.
[16] Weng, C.-Y., Curless, B., Srinivasan, P. P., Barron, J. T., and Kemelmacher-Shlizerman, I. Human- NeRF: Free-viewpoint rendering of moving people from monocular video. Proceedings of CVPR, 2022.
[17] Oechsle, M., Mescheder, L., Niemeyer, M., Strauss, T., and Geiger, A. Texture Fields: Learning texture representations in function space. Proceedings of ICCV, 2019.
[18] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. NeRF: Representing scenes as neural radiance fields for view synthesis. Proceedings of ECCV, 2020.
[19] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. Analyzing and improving the image quality of StyleGAN. Proceedings of CVPR, 2020.
[20] Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. Advances in Neural Infor- mation Processing Systems, 2020.
[21] Dosovitskiy, A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations, 2021.
[22] Havlíček, V. et al. Supervised learning with quantum-enhanced feature spaces. Nature, 2019.
[23] Mitarai, K., Negoro, M., Kitagawa, M., and Fujii, K. Quantum circuit learning. Physical Review A, 2018.
[24] Biamonte, J. et al. Quantum machine learning. Nature, 2017.