TE-NeRF: Triplane-Enhanced Neural Radiance Field for Artifact-Free Human Rendering

Sadia Mubashshira, Kevin Desai

The University of Texas at San Antonio

UTSA The University of Texas at San Antonio™

WACV 2025 TUCSON, ARIZONA • FEB 28 - MAR 4

Motivation

- The presence of artifacts significantly impacts the quality of rendered human avatars, undermining their potential in critical applications such as VR and telepresence.

- Existing NeRF-based methods often produce noticeable artifacts, particularly around complex regions like clothing edges and facial features, which detract from realism and immersion.

Introduction

- Introduce TE-NeRF, an enhanced NeRF framework that aims to achieve high-detail texture representation and seamless surface geometry in human avatar rendering.
- Integration of Triplane features with a density MLP to improve detail accuracy and reduce artifacts.
- Adaptive weight blending mechanism and silhouette loss to balance detail preservation and smooth rendering.

Background

🔹 Human-Specific Neural Rendering

- HyperNeRF → Captured topological deformations from motion.
- A-NeRF → Incorporated pose-dependent deformations.
- NeuralBody → Latent codes enabled better view synthesis.
- HumanNeRF → Separated skeletal and non-rigid motion for more realistic rendering.

⚠️ Limitation of Existing Methods

❌ These approaches primarily consider SMPL skeleton or pose, but not vertex-level information.

❌ This leads to artifacts around the body, particularly in areas with clothing and finer details.

✅ Our approach anchors Triplane features to SMPL vertices, capturing fine-grained details and eliminating artifacts around the body.

Model Architecture

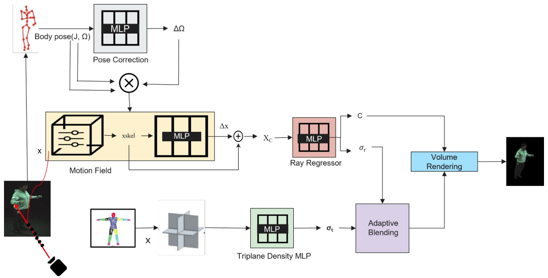

Figure 1: Overview of our proposed model architecture

- Pose correction module learns a **correction function** for pose parameters:

  $$P_{pose}(p) = (J, \Delta\Omega(p) \otimes \Omega)$$

- Motion Field T was decomposed into two parts.
  - Skeletal Motion Tskel
  - Non-rigid motion TNR
- **Ray Regressor** utilized an 8-layer MLP (width=256) with positional encoding γ for input position x, generating color c.
- Triplane Density MLP uses **Triplane Features** which is more **memory efficient** than volumetric cube.
- Features from the three planes f^x, f^y, f^z are concatenated and passed through MLP to predict density σ_t

  $$f = \text{concat}(f^x, f^y, f^z)$$

  $$\sigma_t = \text{MLP\_density}(f)$$

- Adaptive blending mechanism blends Triplane Density (σ_t) and Ray Regressor Density (σ_r) using:

  $$\sigma = w\_blend \cdot \sigma_r + (1 - w\_blend) \cdot \sigma_t$$

Result

Method	PSNR ↑	SSIM ↑	LPIPS ↓
Neural Body	29.08	0.9616	52.27
HumanNerf	30.24	0.9679	31.73
TE-NeRF	29.49	0.9526	42.25

Table 1. Quantitative results on the entire ZJU-MoCap dataset. Higher PSNR and SSIM values indicate better quality, while lower LPIPS values represent improved perceptual realism.

Ground Truth HumanNeRF TE-NeRF

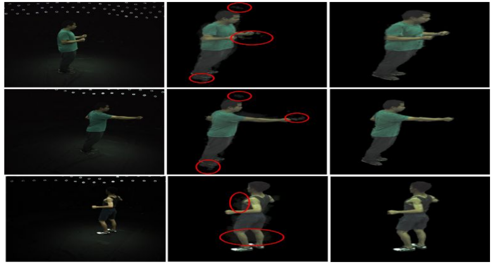

Figure 2: Qualitative comparison of rendering results on the ZJU-MoCap dataset.

Conclusion & Future Scope

- TE-NeRF removes external artifacts while maintaining geometric consistency.
- Triplane Features + Adaptive Blending → Improved spatially-aware density estimation.
- It outperforms Neural Body & HumanNeRF qualitatively.

Future Scope:

- Retain wrinkles, folds, and textures while eliminating external artifacts.
- Extend TE-NeRF to datasets with dynamic clothing.

VIR VISION & IMMERSIVE REALITIES LAB