GFT-GCN: Privacy-Preserving 3D Face Mesh Recognition with Spectral Diffusion

Hichem Felouat[1,2] Hanrui Wang[2] Isao Echizen[1,2,3]
[1]The Graduate University for Advanced Studies, SOKENDAI, Kanagawa, Japan
[2]National Institute of Informatics, NII, Tokyo, Japan
[3]The University of Tokyo, Tokyo, Japan
{hichemfel, hanrui_wang, iechizen}@nii.ac.jp

Appendix

A. Diffusion Model

This section clarifies the distinction between classical diffusion models and the proposed spectral variant. We emphasize that our framework does not utilize diffusion for generative purposes, but rather for privacy-preserving feature transformation.

A.1. Classical Diffusion Models

Classical diffusion models (e.g., Denoising Diffusion Probabilistic Models) iteratively corrupt data by adding Gaussian noise and learn a reverse process to reconstruct the original input. Their primary goal is to generate and reconstruct high-fidelity samples that approximate the input distribution.

A.2. Spectral Diffusion for Privacy

In contrast, our spectral diffusion process is designed for biometric template protection, for which reversibility is undesirable. We operate in the graph spectral domain of 3D face meshes, and the forward noising process is controlled by a trained neural transformation ϕ. This process is guided by discriminability and unlinkability losses, ensuring that privacy protection does not compromise recognition accuracy.

A.3. Why Not Random Noise?

Our approach differs fundamentally from arbitrary stochastic perturbations. Simply adding random noise may severely reduce accuracy and irreversibly degrade biometric utility. Spectral diffusion, on the contrary, provides:

1. **Controlled irreversibility:** Noise is injected stepwise within a Markovian process, allowing a predictable scaling of entropy with the number of steps T.
2. **Learned transformations:** The network ϕ shapes the trajectory of noisy features, making inversion computationally intractable while still retaining discriminative information.
3. **Balanced trade-off:** Loss functions enforce a balance between privacy (irreversibility, unlinkability, renewability) and utility (high recognition accuracy).

The forward spectral diffusion process is not equivalent to an arbitrary perturbation. It yields representations that are systematically hardened against reconstruction attacks while remaining useful for recognition tasks.

B. Dataset Creation

This section describes the creation and splitting of datasets used for training, validation, and testing in the GFT-GCN framework, ensuring a robust evaluation of the proposed 3D face mesh recognition method.

B.1. Dataset Composition

For each dataset (e.g., BU-3DFE and FaceScape), let N_{subj} denote the number of subjects and N_{fs} the number of facial scans per subject. The total number of scans is given by

$$N_{\text{total}} = N_{\text{subj}} \times N_{\text{fs}}$$

B.2. Example Generation

For each subject, we construct training examples as follows:

- **Match pairs:** $N_{\text{fs}} \times (N_{\text{fs}} - 1)$ examples are generated by pairing scans from the same subject.
- **Mismatch pairs:** An equal number of examples are created by pairing scans from one subject with scans from other subjects.

This balanced strategy ensures that the dataset captures both intra-subject similarity and inter-subject differences, which is crucial for robust template-protection performance.

B.3. Data Splits

The dataset is divided into three subsets: 70% for training, 15% for validation, and 15% for testing. The split is applied randomly across subjects to preserve diversity while preventing subject overlap between sets. This configuration provides sufficient data for optimization, enables reliable hyperparameter tuning, and ensures an unbiased evaluation of generalization performance. Importantly, the design reflects the real-world variability in 3D facial data, supporting the framework's privacy-preserving objectives.

C. Theoretical Analysis of Diffusion Steps and Spectral Dimensions

This section provides a theoretical foundation for selecting the diffusion steps ($T \in \{25, 50, 75\}$) and spectral dimensions ($K \in \{10, 20, 25\}$) in the GFT-GCN framework. These parameters directly shape the **privacy-accuracy trade-off** in 3D face mesh recognition. The analysis draws on principles from diffusion models and Graph Fourier Transform (GFT) theory to justify the chosen ranges and quantify their impact on biometric template security and recognition performance.

C.1. Diffusion Steps (T) and Privacy

Each diffusion step adds Gaussian noise (Eq. 3), increasing entropy $H(Z_T)$ in proportion to T, consistent with the Markovian formulation of diffusion models [3, 4]. As entropy increases, the inverse problem becomes increasingly intractable, with the transition probabilities $P(Z_t|Z_{t-1})$ becoming increasingly difficult to compute, leading to reconstruction problems with NP-hard characteristics and improved protection against pre-image attacks (Section 4.3).

C.2. Diffusion Steps (T) and Accuracy

More steps improve irreversibility, but may also degrade recognition performance by over-smoothing discriminative features. The loss of spectral diffusion (Eq. 8) mitigates this by minimizing the loss of discriminability L_{disc} (Eq. 4) across T steps. Empirically, $T = 50$ achieves the best balance, yielding EER $= 0.01$–0.07% and F1 $= 0.93$–0.98 on BU-3DFE and FaceScape. The linear noise schedule β_t preserves sufficient features for reliable cosine similarity matching (Section 4.2).

C.3. Spectral Dimensions (K) and Privacy

Keeping only low-frequency coefficients reduces reconstructive detail. From graph signal processing theory [7], the first K eigenvalues capture global structure while discarding high-frequency identity cues, Figure 1 [2]. For 10k-vertex meshes, $K = 10$ yields an information loss factor of ~ 6.9, improving privacy relative to $K = 25$ (~ 6.0) [1] [5]. The low results of mutual information confirm this effect (Section 4.3).

C.4. Spectral Dimensions (K) and Accuracy

Smaller K increases privacy, but may discard useful identity information. The GCN refines these K-dimensional features ($F_{low} \in \mathbb{R}^{K\times n}$) under a Siamese contrastive loss (Eq. 2), maintaining class separability. $K = 10$–20 preserves F1 ≥ 0.96, while larger K increases computational cost with diminishing accuracy benefits, consistent with spectral dimensionality reduction [6].

C.5. Trade-off Model

We model how diffusion steps (T) and spectral dimensions (K) jointly affect privacy and accuracy.

Privacy: Each diffusion step adds Gaussian noise, so the entropy of the protected template can be approximated as:

$$\Delta H \;\approx\; \tfrac{K}{2}\log\bigl(1+cT\bigr) \;+\; C\log\bigl(\tfrac{N}{K}\bigr),$$

where the first term reflects the accumulated noise and the second term reflects the information loss from truncating to K coefficients ($N = |V|$ vertices). The larger T and the smaller K increase privacy.

Accuracy: Recognition depends on the signal-to-noise ratio of the retained features. A simple approximation is given below:

$$\Delta\text{EER} \;\approx\; A\exp\Bigl(-\alpha\,\tfrac{K}{1+cT}\Bigr),$$

Theoretically, the results indicate that accuracy improves with increasing K and decreases with increasing T.

D. Computational Costs and Deployment

We evaluated the efficiency of the proposed GFT-GCN framework in terms of computational cost and deployment feasibility. All experiments were conducted on NVIDIA Tesla A100 GPUs.

Pipeline Stages: The framework consists of three stages:

1. **Data Preparation:** Raw input loading, facial region cropping, and mesh normalization.
2. **Feature Extraction:** Computing $d = 10$ feature dimensions (3D coordinates, vertex normals, dihedral angles, and Gaussian curvature), and combining them with the low-frequency spectral form F_{low}.
3. **Inference:** Recognition using protected templates with diffusion-based template protection.

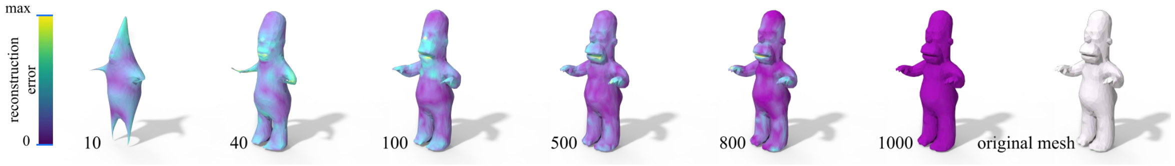

Figure 1. Effect of spectral dimensions K on mesh reconstruction. Small K captures only global shape, hiding identity details and improving privacy, while larger K recovers local geometry, reducing privacy but improving fidelity; the original mesh is shown for reference [2].

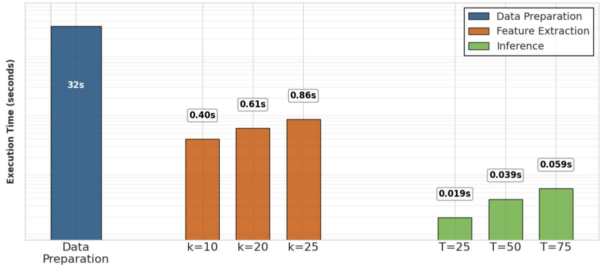

Figure 2. Execution time of GFT-GCN. Data preparation dominates (32 seconds), feature extraction grows with spectral dimension (k), and inference remains lightweight across diffusion steps (T).

Cost Distribution: The most expensive step is data preparation, which takes about 32 seconds per example, primarily due to I/O and preprocessing operations. The second contributor is feature extraction, dominated by the computation of eigenvalues and eigenvectors of the graph Laplacian (0.40–0.86 seconds depending on K). Inference is the lightest step, since diffusion adds only T iterative operations (0.019–0.059 seconds depending on T), Figure 2.

Deployment Feasibility: With typical settings ($K = 10$, $T = 50$), the framework achieves deployment-friendly performance, making it suitable for applications such as border control or secure device access. The client-server design further enhances scalability, as heavy computations (such as data preparation and feature extraction) are performed locally, thereby reducing server overhead while maintaining privacy.

Optimization Opportunities: Additional efficiency gains may be achieved by accelerating eigenvalue computation through parallel solvers, or by reducing diffusion steps using adaptive noise schedules.

References

[1] Yuzhou Chang, Jixin Liu, Yi Jiang, Anjun Ma, Yao Yu Yeo, Qi Guo, Megan McNutt, Jordan E Krull, Scott J Rodig, Dan H Barouch, et al. Graph fourier transform for spatial omics representation and analyses of complex organs. *Nature Communications*, 15(1):7467, 2024. 2

[2] Qiujie Dong, Zixiong Wang, Manyi Li, Junjie Gao, Shuangmin Chen, Zhenyu Shu, Shiqing Xin, Changhe Tu, and Wenping Wang. Laplacian2mesh: Laplacian-based mesh understanding. *IEEE Transactions on Visualization and Computer Graphics*, 30(7):4349–4361, 2023. 2, 3

[3] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020. 2

[4] Kotaro Ikeda, Tomoya Uda, Daisuke Okanohara, and Sosuke Ito. Speed-accuracy relations for diffusion models: Wisdom from nonequilibrium thermodynamics and optimal transport. *Physical Review X*, 15(3):031031, 2025. 2

[5] Tapan Ganatma Nakkina, Adithyaa Karthikeyan, Yuhao Zhong, Ceyhun Eksin, and Satish TS Bukkapatnam. Learnings graph-fourier spectra of textured surface images for defect localization. *Manufacturing Letters*, 41:1568–1578, 2024. 2

[6] Benjamin Ricaud, Pierre Borgnat, Nicolas Tremblay, Paulo Gonçalves, and Pierre Vandergheynst. Fourier could be a data scientist: From graph fourier transform to signal processing on graphs. *Comptes Rendus. Physique*, 20(5):474–488, 2019. 2

[7] Stefania Sardellitti, Sergio Barbarossa, and Paolo Di Lorenzo. On the graph fourier transform for directed graphs. *IEEE Journal of Selected Topics in Signal Processing*, 11(6):796–811, 2017. 2