HodgeFormer: Transformers for Learnable Operators on Triangular Meshes through Data-Driven Hodge Matrices

Akis Nousias[1] Stavros Nousias[2]

[1]K3Y Labs [2]Technical University of Munich

WACV 2026 TUCSON, AZ 3/6 - 3/10

Background

- Traditional approaches for mesh learning typically rely on the computation of spectral components or spectral embeddings through eigenvalue decomposition of the Laplacian.
- HodgeFormer takes a fundamentally different approach. Instead of learning features on a fixed spectral basis, we learn the geometric operator itself — directly from data.

Landscape of Mesh Learning

Spatial methods are fast but lack global context. Spectral methods capture global geometry but require expensive eigendecomposition.

Method	Year	EVD?	Key Idea
Spatial Methods			
MeshCNN[1]	2019	No	Edge conv + collapse pooling
SubdivNet[2]	2021	No	Loop subdivision regularity
MeshWalker[3]	2020	No	Random walks + RNN
Spectral Methods			
HodgeNet[4]	2021	Yes	Learn diagonal Hodge stars
DiffusionNet[5]	2022	Yes	Heat diffusion + HKS
Laplacian2Mesh[6]	2022	Yes	Transformer on eigenvectors
MeT[7]	2023	Yes	Laplacian PE

The Hodge Star $\star_k$

$\star_k$ maps primal k-forms to dual $(2-k)$-forms via a ratio of measures — it is the *only* place shape information enters.

	Primal → Dual	Entry
$\star_0$	Vertex → Voronoi cell	A_v^* (area)
$\star_1$	Edge → dual edge	$\frac{\cot\alpha+\cot\beta}{2}$
$\star_2$	Face → dual vertex	$1/A_f$

Exterior Derivative d_k encodes connectivity:

$$d_0 \in \{-1,0,1\}^{n_e \times n_v}, \quad d_1 \in \{-1,0,1\}^{n_f \times n_e}$$

with $d_1 d_0 = 0$ (curl of gradient $= 0$).

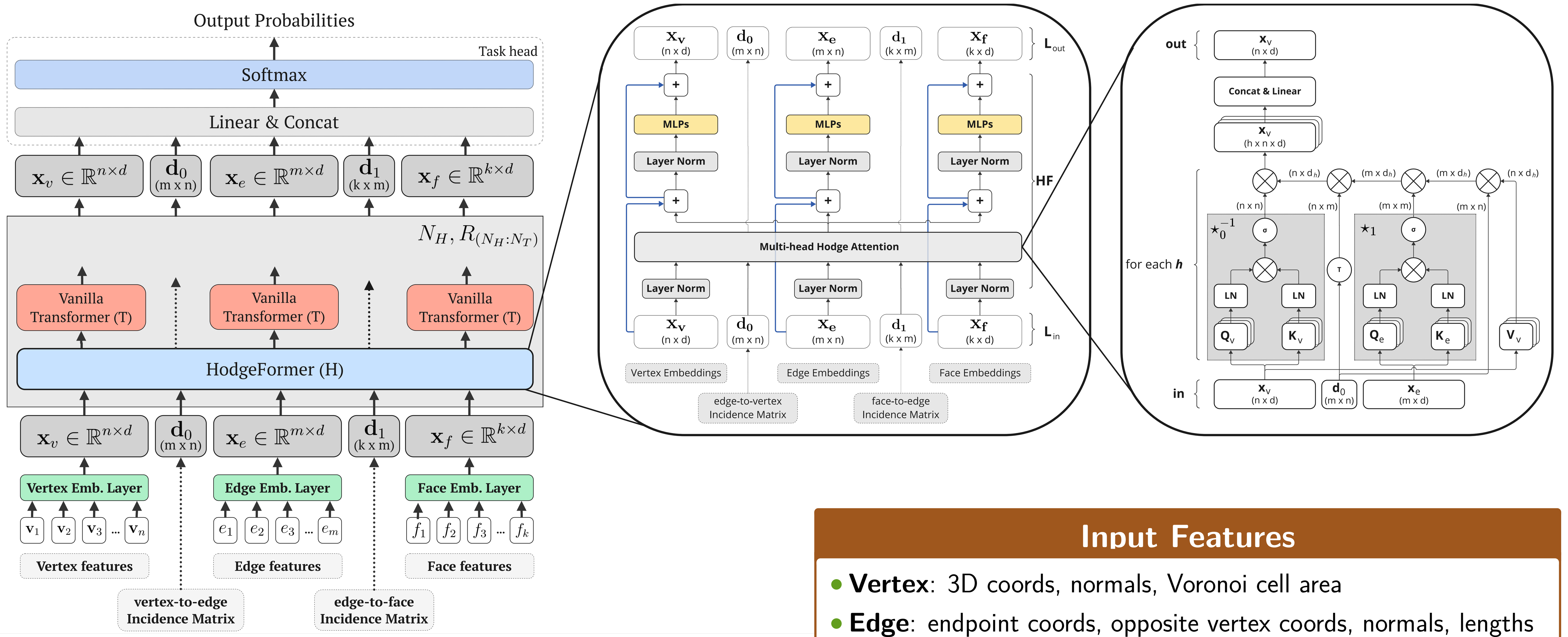
Galerkin Mass Matrices

Classical diagonal $\star_k$: one scalar per element, **no neighbor coupling**. Assumes orthogonal primal-dual. Breaks for irregular or obtuse meshes.

The **Galerkin** approach yields a non-diagonal mass matrix:

$$M_{ij} = \int_\Omega \phi_i(x) \cdot \phi_j(x)\, dx$$

If $M \approx \star_k^{\text{Gal}}$ and $A \approx M$, then **attention can learn Hodge stars**: $\star_k \approx \sigma(QK^\top/\sqrt{d})$



HodgeFormer Layer

Learned Hodge Laplacians:

- $L_v := \star_0^{-1}(x_v)\, d_0^T\, \star_1(x_e)\, d_0$
- $L_e := d_0\, \star_0^{-1}(x_v)\, d_0^T\, \star_1(x_e) + \star_1^{-1}(x_e)\, d_1^T\, \star_2(x_f)\, d_1$
- $L_f := d_1\, \star_1^{-1}(x_e)\, d_1^T\, \star_2(x_f)$

Parametrization:
$\star_k(x_k) = \sigma(Q_k K_k^T/\sqrt{d_h}),\ k \in \{v,e,f\}$.
Layer: $H_l(x_v, x_e, x_f) := G_l(A_{H_l}(\cdot) + (\cdot))$.
Multi-head:
$x_v = L_v V_v(x_v),\ x_e = L_e V_e(x_e),\ x_f = L_f V_f(x_f)$.

HodgeFormer Overview

1. **Fix** incidence matrices d_0, d_1 from mesh topology
2. **Learn** non-diagonal Hodge stars $\star_0, \star_1, \star_2$ via sparse attention
3. **Assemble** Hodge Laplacians: $L_v = \star_0^{-1} d_0^\top \star_1 d_0$ (and L_e, L_f)
4. **Apply** operators end-to-end

Input Features

- **Vertex**: 3D coords, normals, Voronoi cell area
- **Edge**: endpoint coords, opposite vertex coords, normals, lengths
- **Face**: vertex coords, face normal, face area

Embedding Layer $x_{k_{\text{in}}} \to x_k$

$x_k = \text{MLP}\left(x_{k_{\text{in}}} + A_k \cdot x_{k_{\text{in}}}\right)$

- One-hop aggregation via A_k
- MLP projects to latent dim d

Learned Hodge Star via Attention

- $Q_k = x_k W_{Q_k},\ K_k = x_k W_{K_k}$ $\quad W \in \mathbb{R}^{d\times d_h}$
- $\star_k(x_k) = \text{softmax}\left(\frac{Q_k K_k^\top}{\sqrt{d_h}}\right)$
- Non-diagonal

Results

Quantitative Results:

Method	EVD	SHREC	Cube	Human	Vases	Chairs	Aliens
Spectral Methods (require EVD)							
HodgeNet	Yes	94.7	—	85.0	90.3	95.7	96.0
DiffusionNet	Yes	99.5	—	90.8	—	—	—
Laplacian2Mesh	Yes	**100**	**91.5**	88.6	94.6	96.6	95.0
MeT	Yes	—	—	—	**99.8**	**98.9**	**99.3**
Spatial Methods (no EVD)							
MeshCNN	No	91.0	92.2	85.4	92.4	93.0	96.3
SubdivNet	No	99.5	98.9	**91.7**	96.7	96.7	97.3
MeshWalker	No	97.1	98.6	—	99.6	98.7	99.1
HodgeFormer	**No**	98.7	95.3	90.3	94.3	**98.8**	98.3

Runtime Comparison:

Mesh Size (n_v)	2^8	2^{10}	2^{12}	2^{14}
Compute Time (ms)				
HF Encoder (Train)	5.78	8.98	29.72	197.9
HF Encoder (Infer)	2.50	3.42	10.96	66.13
HF Layer (Infer)	1.08	1.91	9.42	55.25
Peak Memory Usage (GBs)				
HF Encoder (Train)	0.12	0.40	2.54	19.23
HF Encoder (Infer)	0.09	0.31	2.08	15.89
HF Layer (Infer)	0.09	0.31	2.08	15.89

	Train (s/ep)	GPU (ms)	Mem (GB)
MeshCNN	26.84	806.29	4.41
Laplacian2Mesh	641.34	423.57	4.80
HodgeFormer	**17.58**	**263.57**	14.21

Qualitative Results (Human Seg):

Robustness to perturbations (Human Seg):

Variant	Acc.	Drop
Original	90.3%	—
Gauss. Noise (λ=0.005)	88.3%	2.0%
Gauss. Noise (λ=0.020)	81.1%	9.2%
QEM Remesh (1000F)	87.2%	3.1%
Face Removal (p=0.10)	85.9%	4.4%
Patch Removal	86.8%	3.5%

Ablation Studies: Layer mixing :

Configuration	Accuracy(COSEG Vases)
N_H=4, ratio 4:0 (HF only)	92.02%
N_H=4, ratio 1:1	92.85%
N_H=4, ratio 4:2	**93.04%**

Conclusion: HodgeFormer achieves results comparable to the state-of-the-art without spectral features or eigenvalue decomposition.

Contact

Code:
https://github.com/hodgeformer/
Email:
anousias@k3y.bg
stavros.nousias@tum.de

References

[1] R. Hanocka, A. Hertz, N. Fish, R. Giryes, S. Fleishman, and D. Cohen-Or, "MeshCNN: a network with an edge," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, pp. 1–12, 2019.
[2] S.-M. Hu, Z.-N. Liu, M.-H. Guo, J.-X. Cai, J. Huang, T.-J. Mu, and R. R. Martin, "Subdivision-based mesh convolution networks," *ACM Transactions on Graphics (TOG)*, vol. 41, no. 3, pp. 1–16, 2022.
[3] A. Lahav and A. Tal, "MeshWalker: Deep Mesh Understanding by Random Walks," *arXiv preprint arXiv:2006.05353*, 2020.
[4] D. Smirnov and J. Solomon, "HodgeNet: learning spectral geometry on triangle meshes," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 1–11, 2021.
[5] N. Sharp, S. Attaiki, K. Crane, and M. Ovsjanikov, "DiffusionNet: Discretization Agnostic Learning on Surfaces," *ACM Transactions on Graphics*, vol. 41, no. 3, pp. 1–16, 2022.
[6] Q. Dong, Z. Wang, J. Gao, S. Chen, Z. Shu, and S. Xin, "Laplacian2Mesh: Laplacian-Based Mesh Understanding," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 1, pp. 492–502, 2023.
[7] G. Vecchio, L. Prezzavento, C. Pino, F. Rundo, S. Palazzo, and C. Spampinato, "MeT: A Graph Transformer for Semantic Segmentation of 3D Meshes," *arXiv preprint arXiv:2307.01115*, 2023.

Acknowledgements

This research was supported by TUM Innovation Network CoConstruct N2201 and TUM Georg Nemetschek Institute - Artificial Intelligence for the Built World.