NeurIPS 2026 Poster

CycleSpectra: Cyclic Motion Spectra for Phase-Queryable 4D Cardiac Reconstruction

Xueming Fu[1,2] **Xin Li**[1,2] **Lixia Han**[6] **Ao Shen**[7] **Guangming Lu**[8] **Song Luo**[9] **Shaohua Kevin Zhou**[1,2,3,4,5]

[1]School of Biomedical Engineering, USTC [2]MIRACLE Lab, YRD-RIGHT, USTC Suzhou Institute for Advanced Research
[3]Jiangsu Key Lab of Multimodal Digital Twin [4]Biomedical Basic Research Center of Jiangsu [5]State Key Lab of Precision & Intelligent Chemistry
[6]Nanjing University of Aeronautics and Astronautics [7]Ruijin Hospital, SJTU [8]Jinling Hospital, NJU [9]Geriatric Hospital, NMU

University of Science and Technology of China

MIRACLE Lab

1 Motivation

Can a cyclic prior be built into the motion representation?

Cardiac motion is **largely cyclic** across a beat, yet existing methods typically fit discrete phases or pairwise transitions, without leveraging this cyclic prior in the representation.

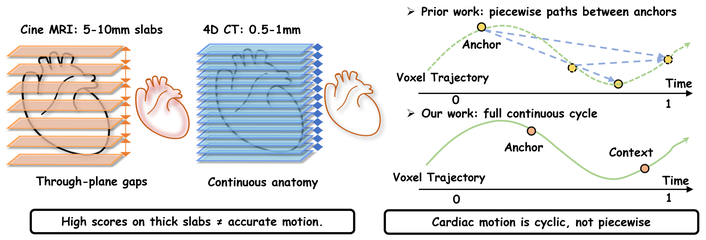


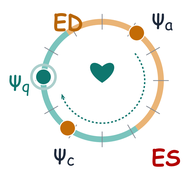
Fig. 1. **Left:** Thick cine slabs (5–10 mm) inflate image scores vs. thin-slice 4D CT (0.5–1 mm). **Right:** Real voxel trajectories (green) are non-monotonic and cyclic, whereas prior work learns piecewise paths between anchors (blue).

2 Key Idea: Cycle by Design

To encode the cyclic prior directly, we represent each site's motion as a **short Fourier series** in phase ψ:

$$\varphi(\mathbf{x}, \psi_a, \psi_q) = \sum_{k=1}^{K} \mathbf{A}_k(\mathbf{x})\left[\sin(2\pi k\psi_q) - \sin(2\pi k\psi_a)\right] + \mathbf{B}_k(\mathbf{x})\left[\cos(2\pi k\psi_q) - \cos(2\pi k\psi_a)\right]$$

- **Time never enters the network**; ψ_q is only the read-out coordinate.
- Periodic and continuous **by design**; displacement is zero at $\psi_q = \psi_a$.
- **Closed-form read-out** $\hat{\mathbf{X}}_{\psi_q} = \mathcal{W}(\mathbf{X}_a, \varphi)$ at any phase.



3 Contributions & Benchmark

1. **Spectral Motion Representation:** Per-site cyclic trajectories parameterized by K=4 latent Fourier bands $\{\mathbf{A}_k, \mathbf{B}_k\}$, queryable at arbitrary phases without per-case fitting.
2. **Phase-Queryable Framework:** DualMotion encoder + cross-frequency attention decoder + residual refinement; supports **test-time multi-frame (M=2→4) latent fusion** without retraining.
3. **Thin-Slice 4D-CT Benchmark:** 36 subjects (1.6 mm, 6–24 phases, 22/7/7 split) tested under canonical ED/ES and **85 non-canonical opposite pairs** (1,038 query frames), plus zero-shot OOD 4D-CT & 4D-cMR evaluation.

4 Method: Two Frames → Queryable Cycle

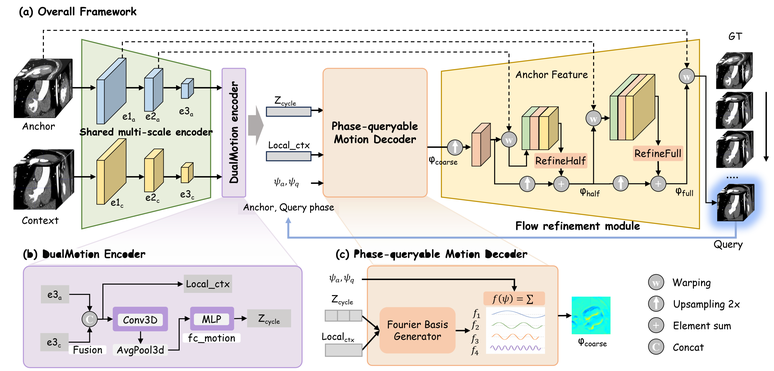


Fig. 2. Shared 3D encoder extracts multi-scale features from $(\mathbf{X}_a, \mathbf{X}_c)$; **DualMotion** outputs local context $\mathbf{Local}_{ctx}$ and K frequency tokens $\mathbf{Z}_{cycle}$; **Phase-Queryable Decoder** predicts Fourier maps $\{\mathbf{A}_k, \mathbf{B}_k\}$ and reads out flow analytically at any ψ_q, followed by two refinement blocks.

Training Objective: $\mathcal{L} = \mathcal{L}_{\ell_1} + \lambda_{ssim}\mathcal{L}_{SSIM} + \lambda_{fg}\mathcal{L}_{fg} + \lambda_{smooth}\mathcal{L}_{smooth} + \lambda_{fold}\mathcal{L}_{fold} + \lambda_{ep}\mathcal{L}_{ep}, \quad \mathcal{L}_{ep} = \left\|\mathcal{W}(\hat{\mathbf{X}}_{\psi_q}, \hat{\varphi}(\cdot, \psi_q, \psi_e)) - \mathbf{X}_e\right\|_1$

5 Results

- **Canonical ED/ES** (7 pairs): standard setting with inputs fixed at end-diastole (ED) and end-systole (ES).
- **Opposite-pair** (85 pairs): non-canonical systolic–diastolic pairs ($\Delta\psi \in [0.35, 0.65]$), testing generalization beyond ED/ES.

Method	PSNR↑	ROI PSNR↑	Dice avg↑	HD95↓	Curve corr↑	SV err. (%)↓	LVEF err. (%)↓
Canonical ED/ES							
Identity	20.31±0.91	18.56±0.87	0.751±0.066	11.04±6.07	-0.093±0.235	38.12±23.91	7.32±6.04
T0 linear-flow	21.70±0.91	20.03±0.86	0.776±0.059	10.64±6.26	0.158±0.282	39.58±23.45	7.65±6.18
UVI-Net (CVPR'24)	**24.18**±0.92	**22.29**±0.91	0.751±0.066	11.04±6.07	-0.093±0.235	38.12±23.91	7.32±6.04
TTT4D	21.06±1.26	19.39±1.21	0.719±0.062	12.14±6.09	-0.031±0.285	49.85±26.99	**5.52**±4.79
TMKD (MICCAI'25)	22.88±0.82	21.00±0.79	0.778±0.059	10.54±4.06	0.248±0.222	42.15±13.14	6.37±4.93
FusionNet (MICCAI'23)	21.24±0.93	19.71±1.00	0.778±0.061	13.32±8.26	**0.781**±0.181	41.44±39.08	9.01±13.49
Mo-Diff (MICCAI'25)	15.44±2.06	14.24±2.08	0.097±0.062	7.46±1.85*	0.143±0.233*	156.52±50.21	43.00±11.35†
Dyna3DGR (MICCAI'25)	19.95±0.40	19.66±0.80	0.713±0.061	11.96±5.97	-0.060±0.244	**37.57**±23.62	6.90±4.80
Ours	22.60±0.82	21.09±0.76	**0.800**±0.061	**9.83**±6.39	0.632±0.241	38.03±24.36	7.80±8.90
Opposite-pair							
Identity	21.30±0.77	19.54±0.69	0.785±0.047	8.87±2.65	0.102±0.188	57.39±6.93	24.41±5.90
T0 linear-flow	22.19±0.74	20.49±0.63	0.796±0.047	8.64±2.69	0.340±0.172	55.06±8.03	21.11±6.80
UVI-Net (CVPR'24)	**23.55**±0.75	21.62±0.64	0.785±0.047	8.87±2.65	0.102±0.188	57.39±6.93	24.41±5.90
TTT4D	20.61±1.23	18.91±1.17	0.740±0.048	10.29±2.71	0.185±0.153	46.83±7.58	11.40±2.54
TMKD (MICCAI'25)	23.01±0.75	21.09±0.64	0.778±0.058	10.62±3.83	0.187±0.166	64.53±7.41	24.47±5.96
FusionNet (MICCAI'23)	21.21±1.20	19.70±1.27	0.779±0.061	13.60±7.69	0.853±0.128	46.03±36.64	9.41±11.26
Mo-Diff (MICCAI'25)	15.76±0.86	14.59±0.92	0.061±0.014	13.13±9.39	0.196±0.224	97.56±34.50	43.00±11.35†
Dyna3DGR (MICCAI'25)	19.82±0.33	19.47±0.63	0.739±0.048	9.83±2.47	0.113±0.190	55.80±7.02	23.29±5.86
Ours	23.13±0.68	**21.67**±0.65	**0.821**±0.050	**7.94**±2.91	**0.862**±0.075	**44.12**±11.89	**8.00**±6.37

Tab. 1. Mean ± population SD (**bold**: best, underline: second best). Edges over FusionNet/TTT4D on Curve/LVEF are nominal (paired 95% CIs include 0). *n=5 valid subjects. †Mo-Diff LVEF saturates (empty LV masks).

6 Visualization & Ablations

▶ Reconstruction Frame

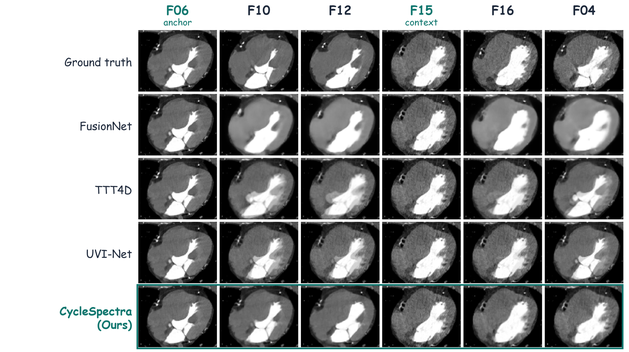


Fig. 3. Opposite-pair generation from **F06 (anchor)** and **F15 (context)** (columns **F10, F12, F16, F04** are unseen query phases). Baselines produce blurred or quasi-static intermediate frames away from canonical ED/ES, whereas ***CycleSpectra*** preserves chamber boundaries and recovers contraction–relaxation motion.

▶ Error Distribution

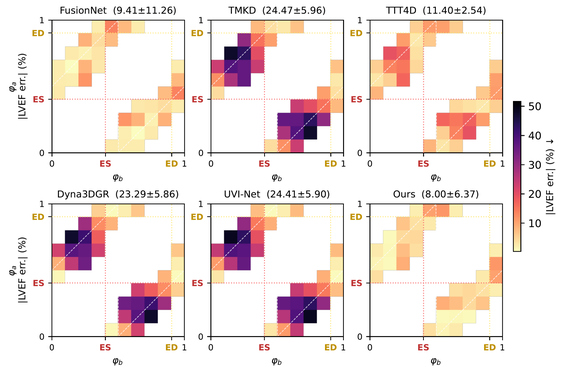


Fig. 4. Per-method |LVEF err.| over the (φ_a, φ_b) phase grid (85 opposite pairs; dotted lines: canonical $\bar{\varphi}_{ED} \approx 0.90$, $\bar{\varphi}_{ES} \approx 0.40$). Baselines spike away from ED/ES; **CycleSpectra stays low across the plane.**

Method / Configuration	PSNR↑	Dice avg↑	Curve corr↑	LVEF err. (%)↓
Decoder Parameterization (85 opposite pairs, 1,038 query frames)				
SinPE flow decoder (learned phase-cond.)	22.76±0.71	0.812±0.048	0.718±0.099	19.69±7.44
CycleSpectra (Ours, spectral K=4)	**23.13**±0.68	**0.821**±0.050	**0.862**±0.075	**8.00**±6.37
Test-Time Multi-Frame Latent Fusion (frozen model, common non-input frames)				
M=2 observed frames	23.58±0.80	0.836±0.027	0.846±0.151	11.09±10.91
M=3 observed frames	24.19±0.85	0.842±0.037	0.859±0.161	8.09±4.77
M=4 **observed frames**	**24.61**±1.08	**0.852**±0.044	**0.941**±0.034	**6.35**±4.19

Tab. 2. Learned phase-conditioned flow (SinPE) degrades sharply on cycle motion; latent fusion (M=2→4) lifts curve corr. to **0.941**.

Takeaways

Building the cardiac cycle into the representation itself—a compact periodic spectrum read out analytically at any phase—appears to follow the heart's own motion more closely, and may help recover motion trends over the whole cycle from only two sparse frames.

Limitations: 7 held-out test subjects (nested pairs); dense phase supervision is used in training; one low-EF failure case (GT LVEF 32.3%); stronger claims should await larger and more diverse cohorts.

Code