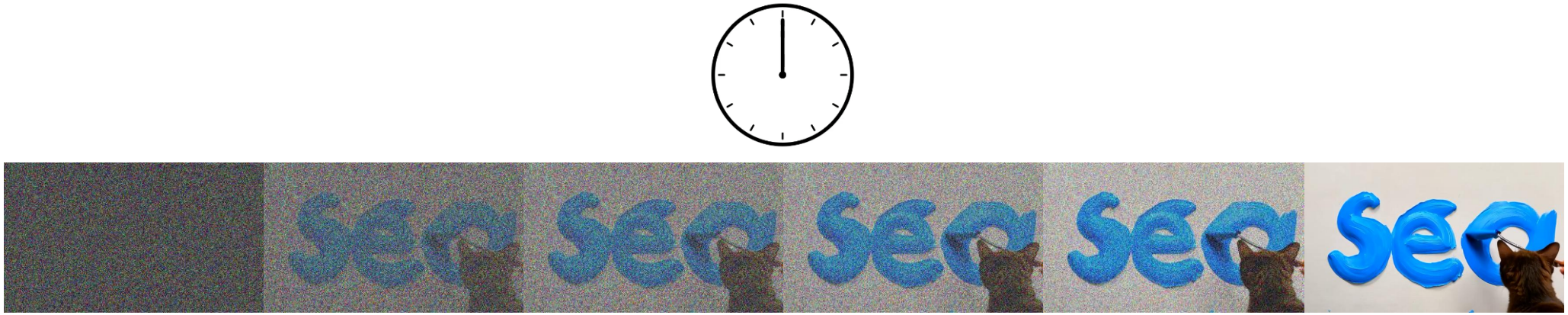
SeaCache: Spectral-Evolution-Aware Cache for Accelerating Diffusion Models

Jiwoo Chung[1], Sangeek Hyun[1], Minkyu Lee[1], Byeongju Han[2],
Geonho Cha[2], Dongyoon Wee[2], Yongjun Hong[2*], Jae-Pil Heo[1*]

[1]Sungkyunkwan University, [2]NAVER Cloud

Slow Diffusion

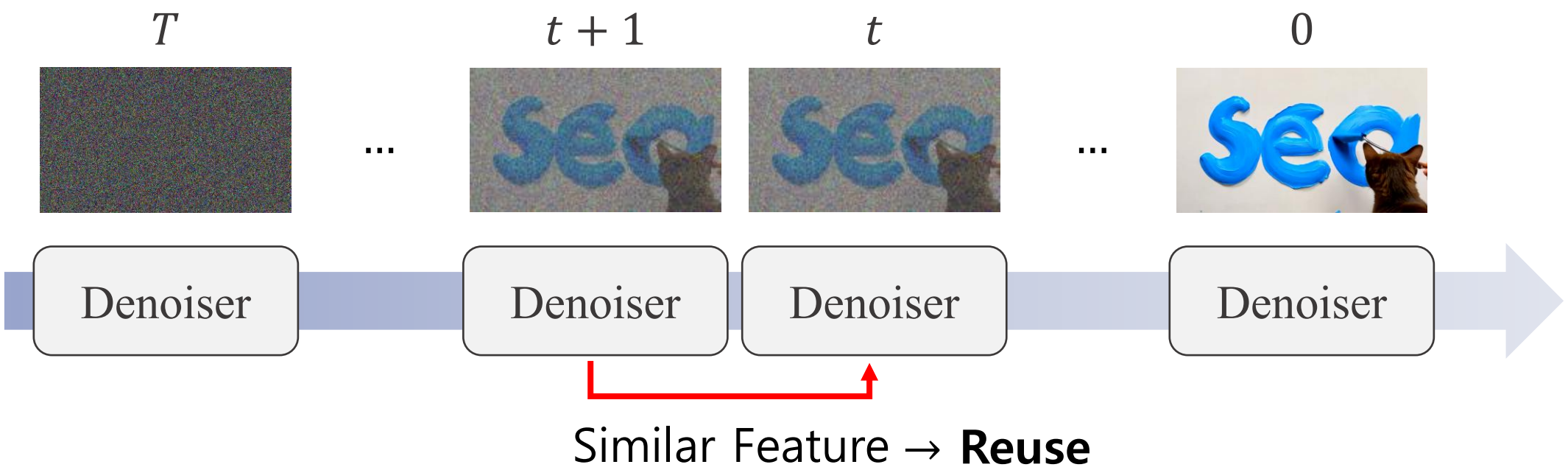
Diffusion is slow due to its iterative denoising process.



Diffusion Denoising Process

Motivation of Diffusion Caching

Adjacent timesteps often produce similar features.



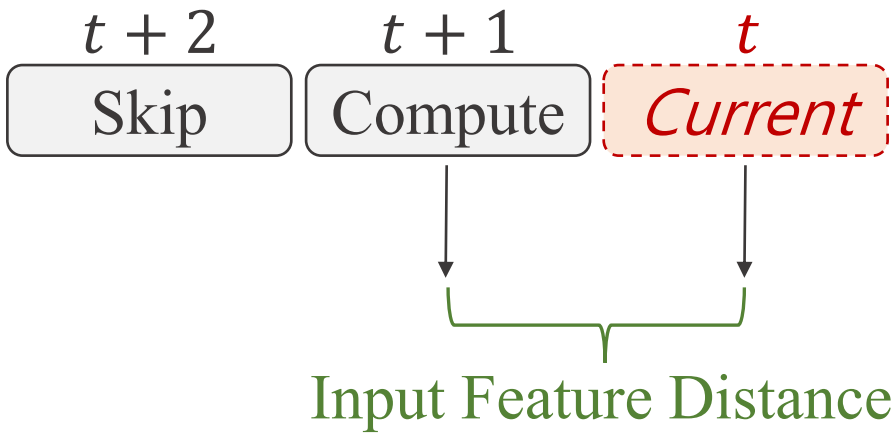
Goal of Diffusion Caching

The goal is to predict output similarity to decide when to skip computation.

Use **input similarity as a proxy** for output similarity

Diffusion Caching

Diffusion caching accelerates inference by reusing output feature maps.

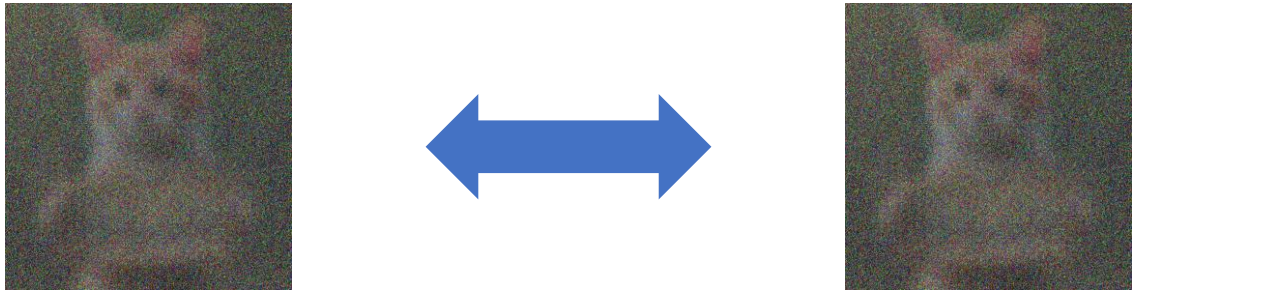


Large distance → **Compute**

Small distance → **Reuse**

Straightforward Approach: Raw Input Distance

The straightforward approach is to compare **raw** input features.



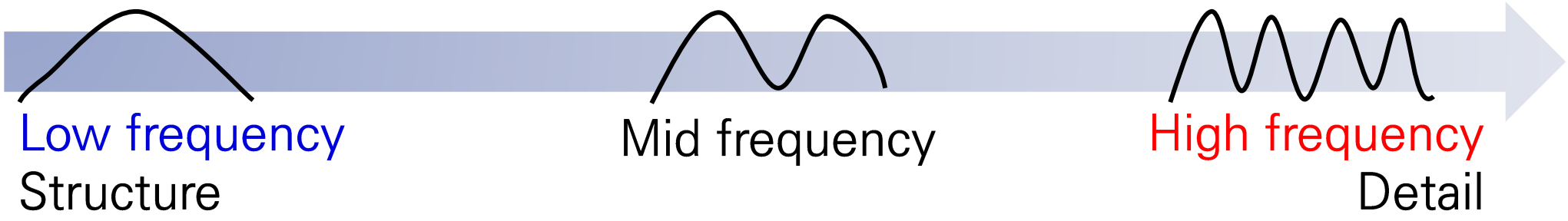
Raw input feature x_t Raw input feature x_{t+1}

But raw distance applies the same criterion at every timestep.

Motivation: Spectral Evolution

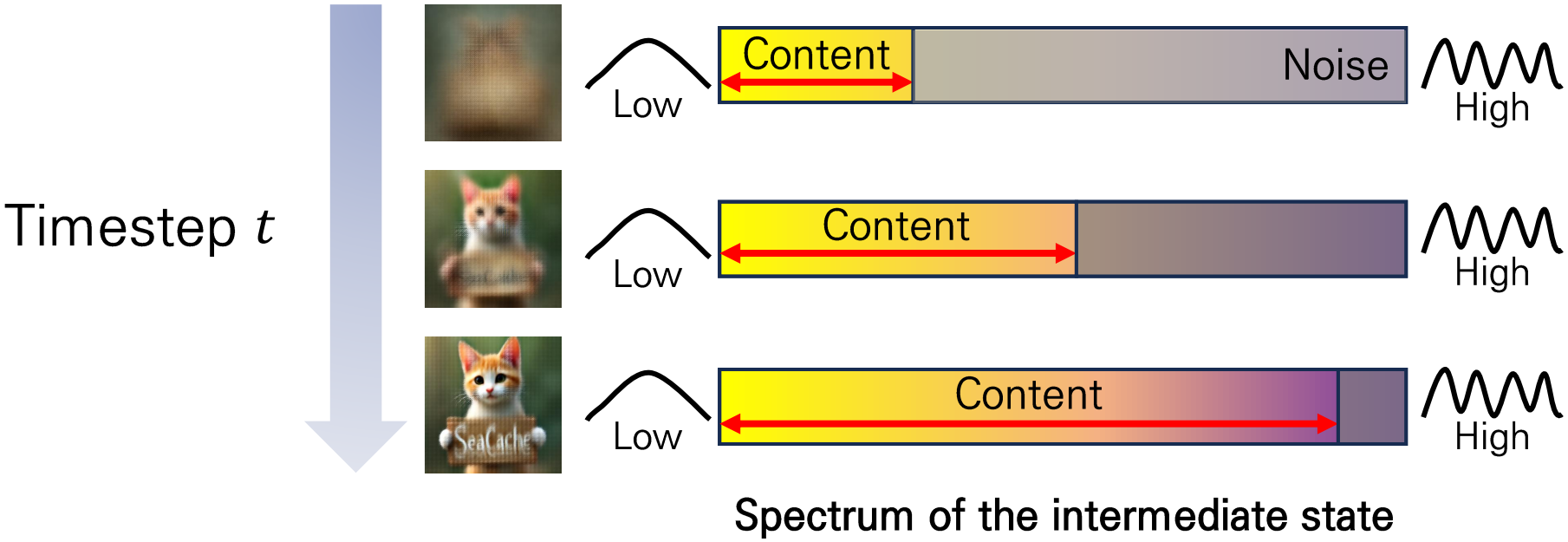
Content evolves from low freq. (coarse structure) to high freq. (fine details).

Therefore, the important frequency component changes over timesteps.



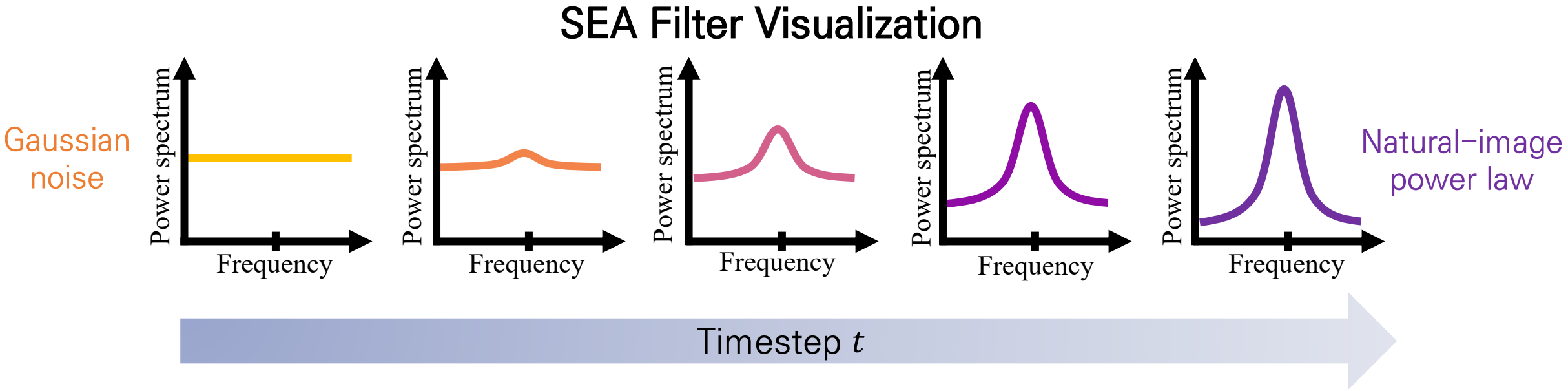
Motivation for SEA Filter

Caching distance should focus on the informative frequency range.



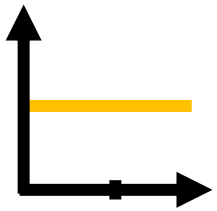
SEA Filter Visualization

The SEA filter identifies the informative frequency range at each timestep.



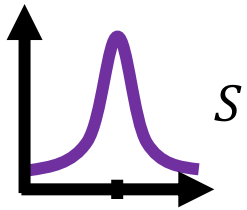
SEA Filter Derivation

We derive the SEA filter from linear denoising filter.



$\epsilon \sim \mathcal{N}(0, I)$

Gaussian noise

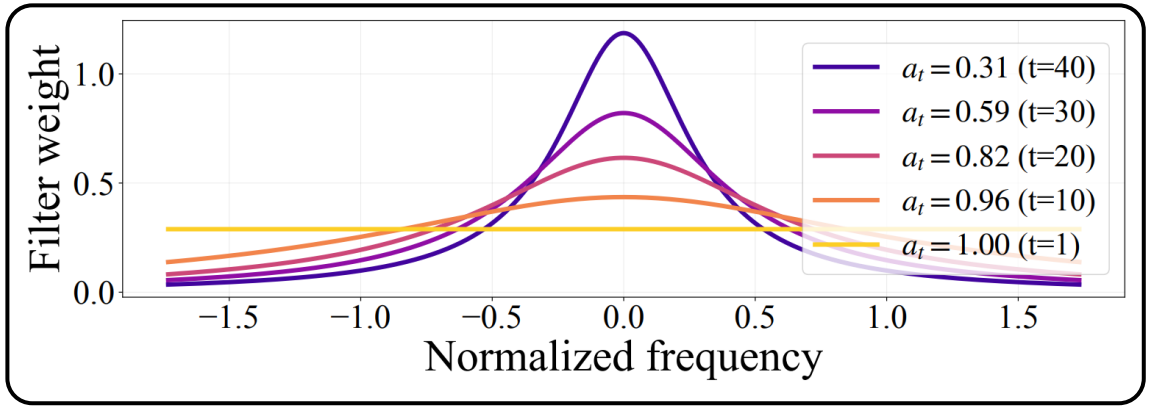


$S(f) \simeq A|f|^{-\beta}$

Natural-image power law

$x_t = a_t x_0 + b_t \epsilon$

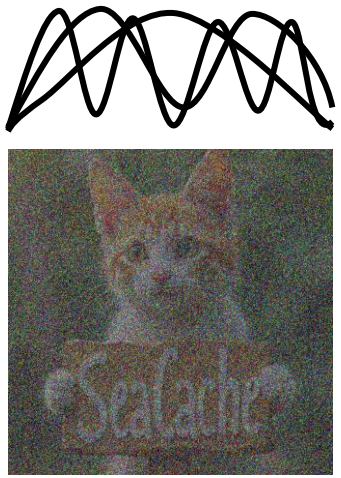
Mixture at timestep t



SEA filter : $\mathcal{H}_t(f) = \dfrac{a_t S(f)}{a_t^2 S(f) + b_t^2}$

Cache What Matters

Measure change in a spectrally-aligned space.



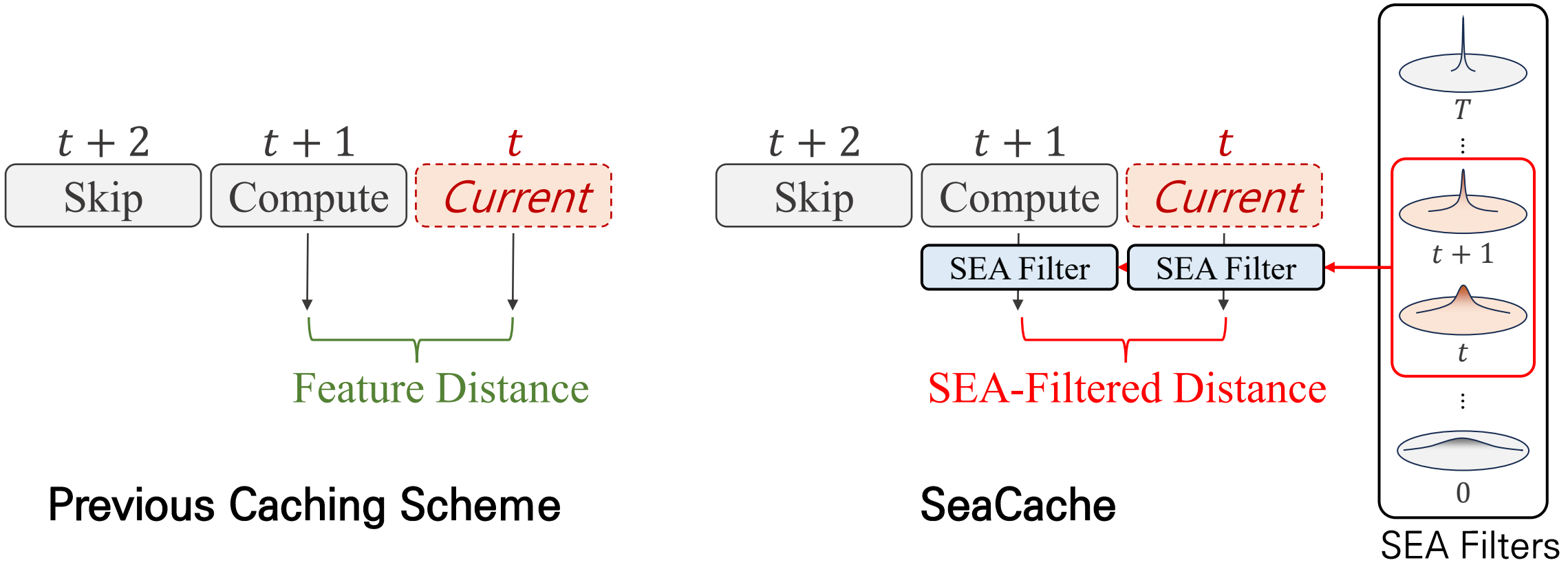
Raw Feature x_t

SEA Filter

Filtered Content $\hat{x}_0$

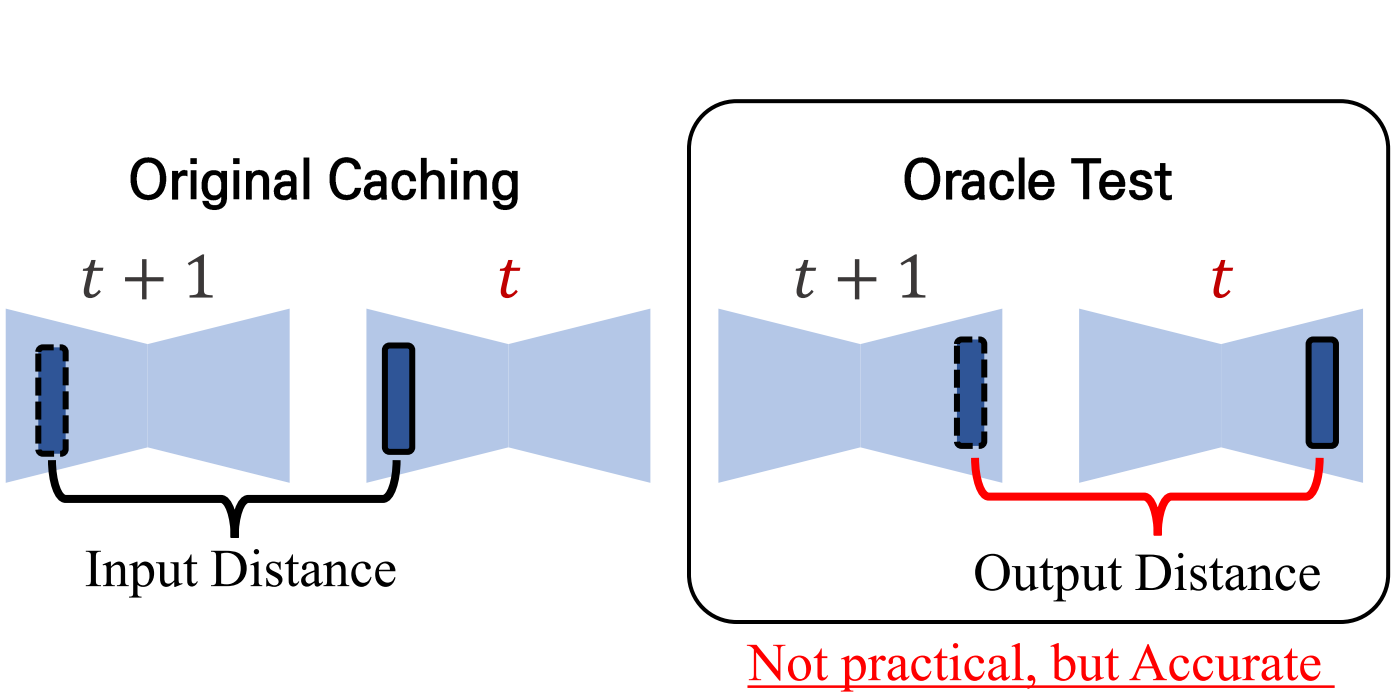
SeaCache Framework

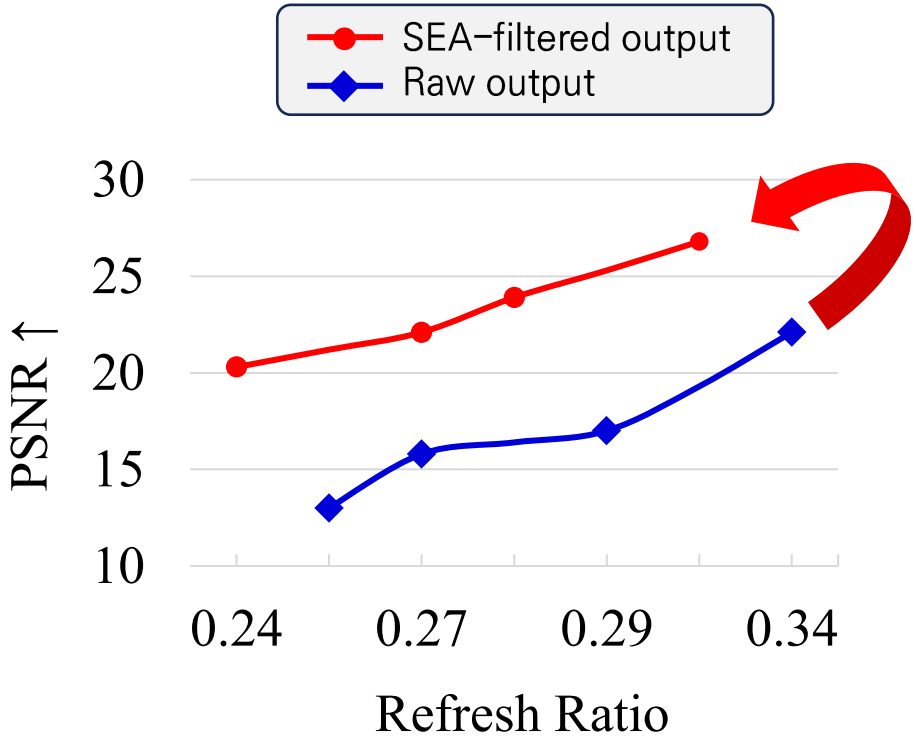
Measure distance between filtered input features.



Analysis: Oracle Test

We perform an oracle test to identify isolating the effect of spectral filtering.

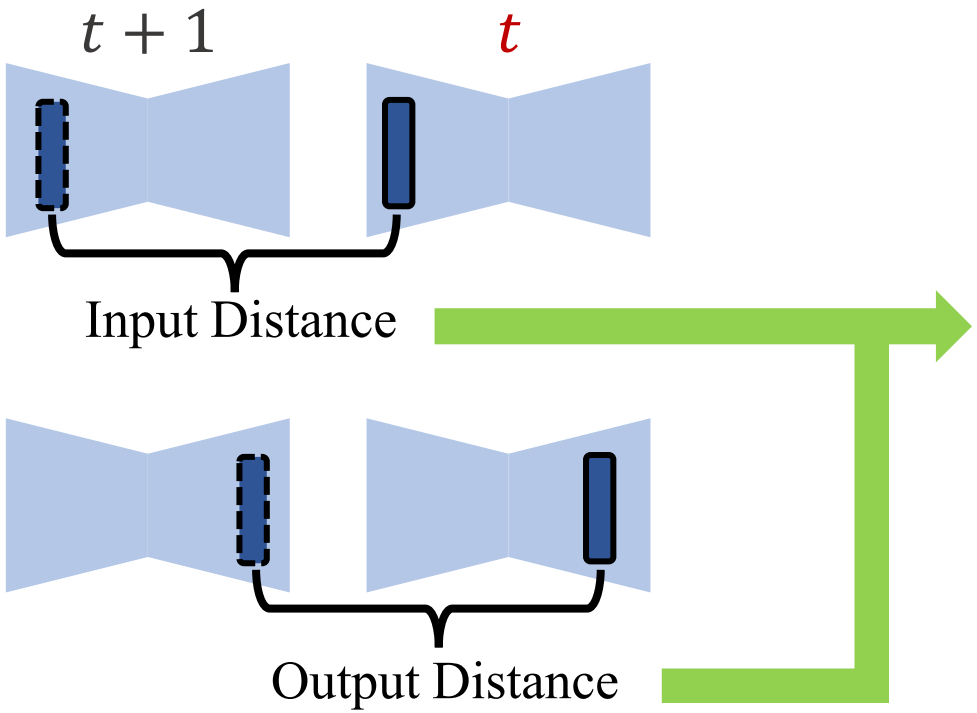


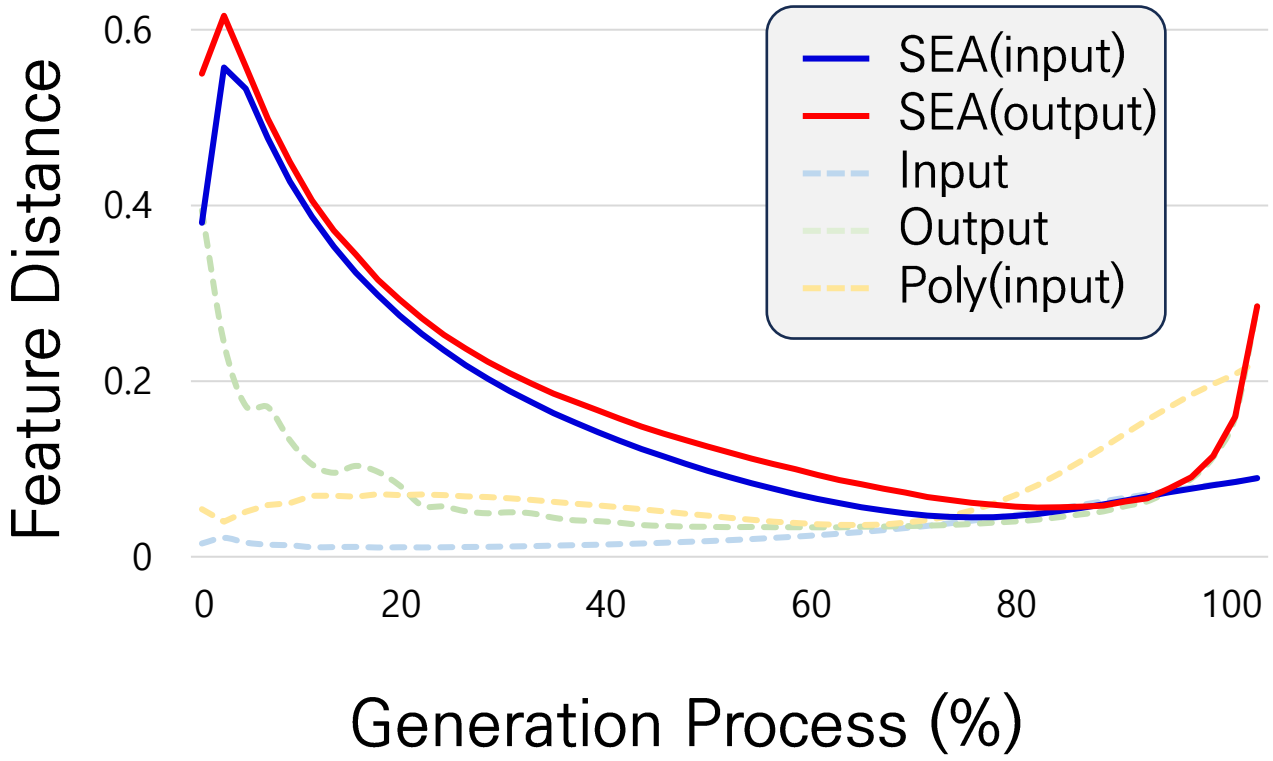


SEA-filtered outputs better track full-compute trajectories

Analysis: Input as Proxy

In practice, SEA(input) distance approximates SEA(output) distance.





Video Qualitative Results

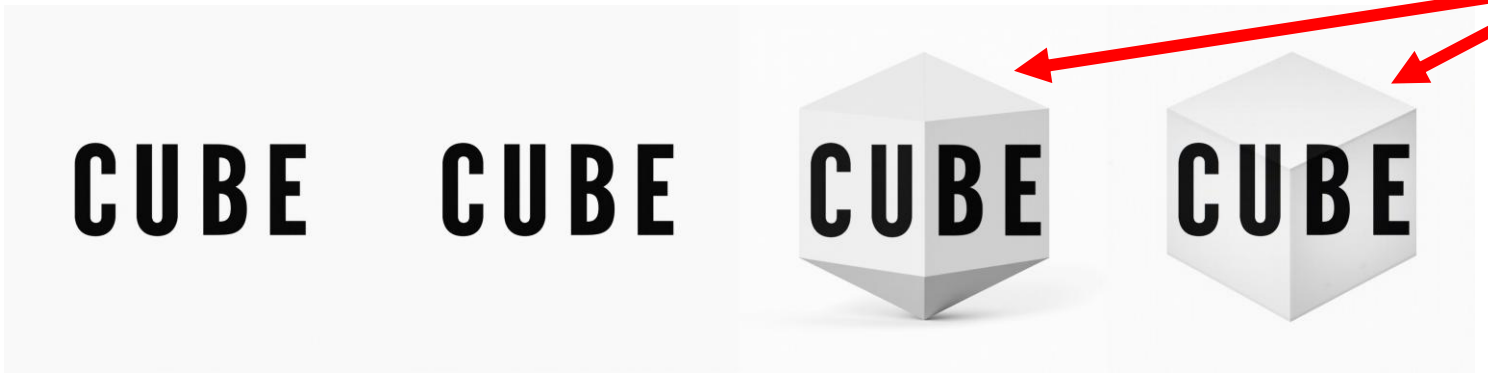
Original | SeaCache | TeaCache | TaylorSeer

“A person is tai chi.”

“A person is running on a treadmill.”

Image Qualitative Results

Original SeaCache TeaCache TaylorSeer



Not text only

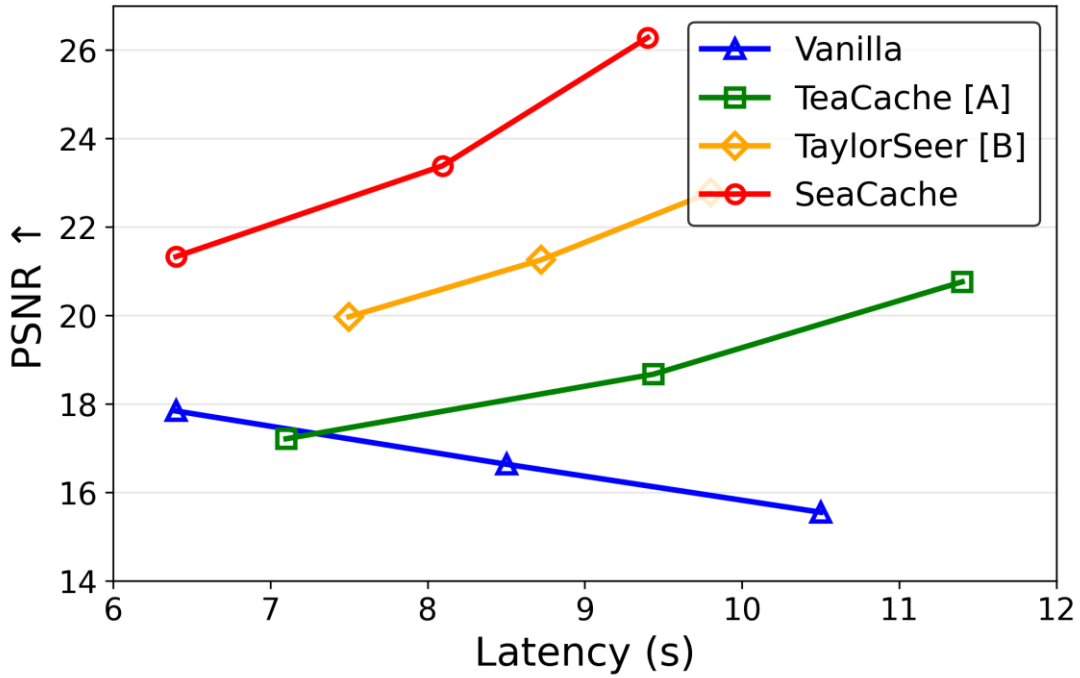
"Exactly "CUBE" text only, centered, white background"

Wrong count

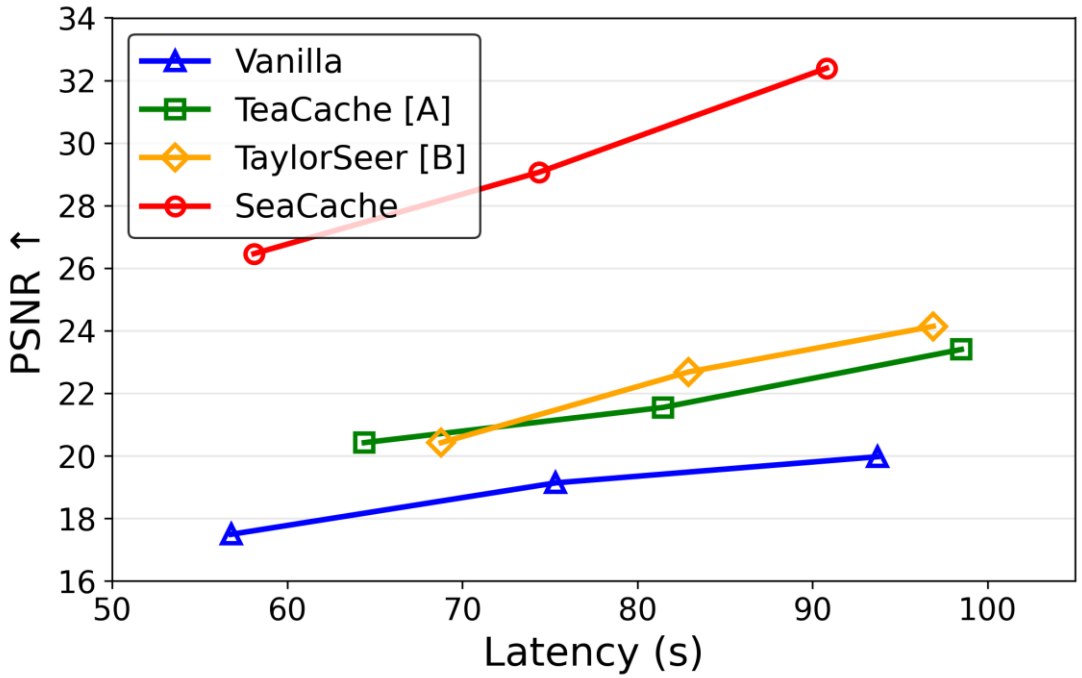
"7 white glue stars, no distortion, accurate shadows"

Quantitative Results

SeaCache achieves a better quality–latency trade-off.



FLUX (Text-to-Image)

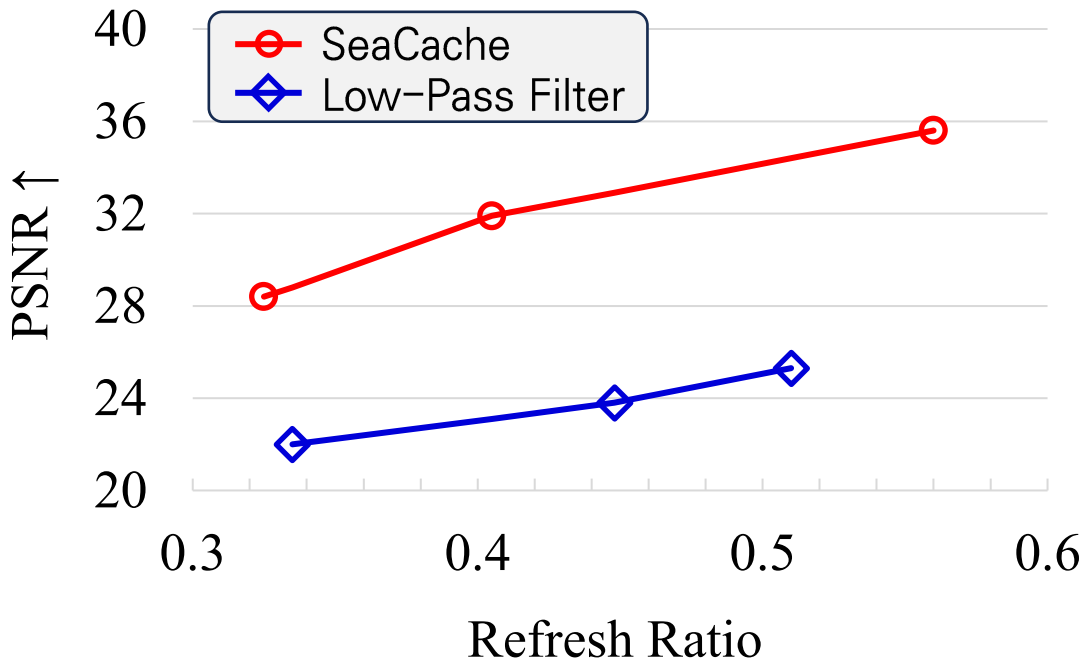


Wan2.1 1.3B (Text-to-Video)

[A] Liu, Feng, et al. "Timestep Embedding Tells: It's Time to Cache for Video Diffusion Model." CVPR 2025
[B] Liu, Jiacheng, et al. "From reusing to forecasting: Accelerating diffusion models with taylorseers." ICCV 2025.

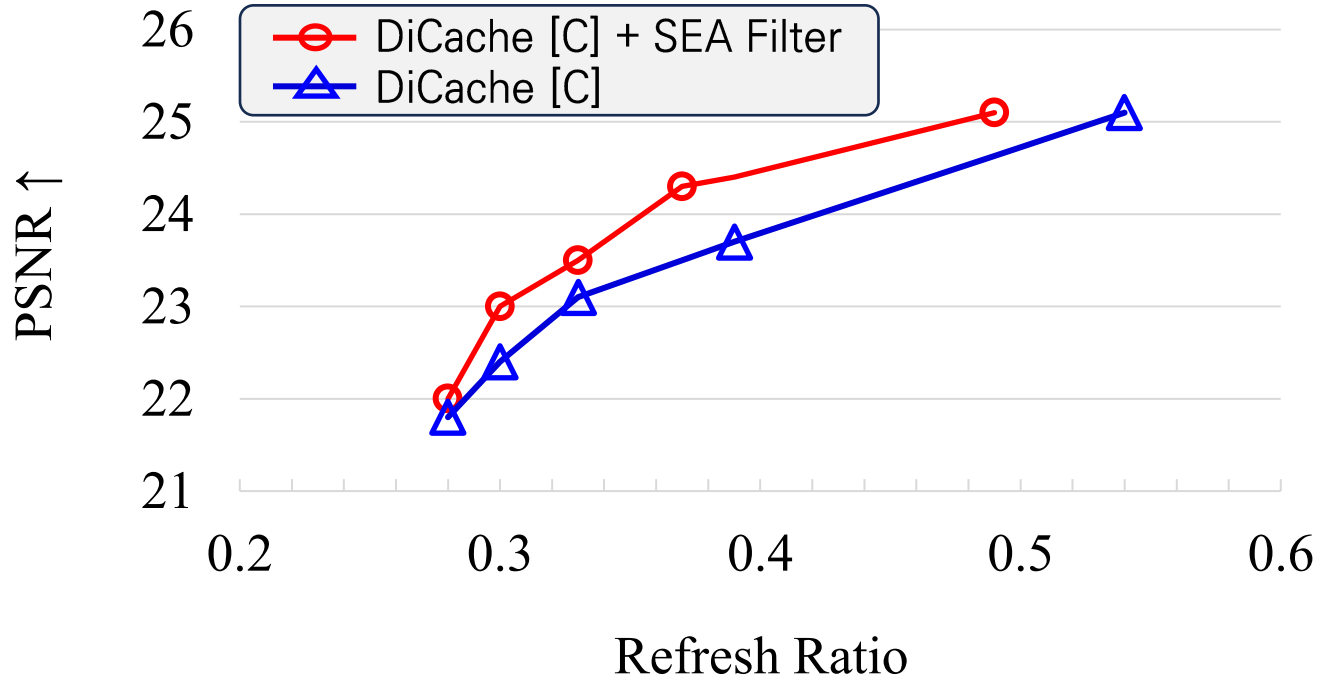
SeaCache Is Not Just Low-Pass

SeaCache outperforms a low-pass filter.



Compatibility with Existing Cache Methods

SEA filtering further improves DiCache across refresh ratios.



[C] Bu, Jiazi, et al. "Dicache: Let diffusion model determine its own cache." ICLR 2026

Conclusion

01

Plug-and-play

- Architecture-Agnostic
- Sampler-Agnostic

02

Compatible

- Compatible with other caching methods
- Compatible with other acceleration works

03

Practical

- Negligible overhead
- Single hyperparameter

Thank you!

Original
Video Gen.
(6 minutes)

SeaCache
(2 minutes)

Original SeaCache

Project Page

Original SeaCache Original SeaCache