



home
papers
study
life

About me

SVG

- Sparse VideoGen: Accelerating Video Diffusion Transformers with Spatial-Temporal Sparsity
- <https://arxiv.org/abs/2502.01776>
- ICML 2025

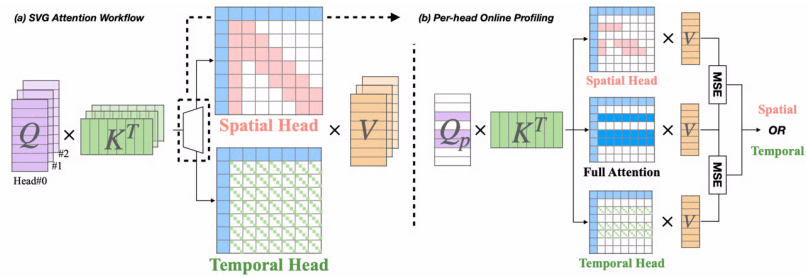


Figure 4. Overview of SVG framework. (a) During generation, SVG adaptively classifies each attention head as either a *spatial head* or a *temporal head* and applies a dedicated sparse attention computation accordingly. (b) This adaptive classification is driven by online profiling strategy, which extracts a small portion of Q , denoted as Q_p , to perform both spatial and temporal attention computations. SVG then selects the attention pattern that yields the minimal MSE compared to full attention, ensuring accurate classification.

3D full attention head Spatial head
Temporal head head CUDA
head Q + attention
head

- <https://github.com/svg-project/Sparse-VideoGen>

Newer

Older

2025-10-29

PAB

2025-10-29

TeaCache

leicheng © 2022-2025

Archive RSS feed GitHub Email QR Code

Made with Montaigne and by anton