

Supplementary Materials: Fuse Your Latents: Video Editing with Multi-source Latent Diffusion Models

Anonymous Author(s)
Submission Id: 1490

1 IMPLEMENTATION DETAILS

1.1 Dataset Details

We select 16 videos from the DAVIS dataset. The selected videos are listed in Table 1.

Table 1: Names of videos selected from DAVIS dataset.

bmx-rider	dog-competition	drift-turn	gold-fish
hurdles-race	ice-hockey	judo	juggling-selfie
lindy-hop	loading	mbike-trick	miami-surf
stunt	swimmer	trucks-race	varanus-cage

1.2 User Study Details

We recruit 20 evaluators to pair-wisely compare our method with baselines. For each comparison, we set three questions as following:

- Which one better matches the text description?

- Which one is more similar to the original video, considering modifications made to the original video?
- Which one is more continuous?

2 VIDEO DEMO

We provide video demos on our project page including

- Visualization of the denoising process for T2I and T2V LDMs’ latents.
- Video results of our method.
- Baseline and SOTA comparisons.

Our project page is available here.

3 LIMITATION AND FUTURE WORK

FLDM can serve as a versatile plugin that can be applied to off-the-shelf image and video LDMs to enhance the quality of video editing. This leads to that the capability of FLDM is influenced by T2I and T2V models. FLDM with stronger T2V and T2I models will bring about improved editing results.