Project Proposal: Temporally Consistent Video Colorization

Xuanyi Xie, Jiayi Hu

Problem Statement

We propose investigating the automatic colorization of grayscale videos with focus on temporal consistency. Unlike image colorization, video colorization faces unique challenges in maintaining color coherence between frames, which can cause flickering if not addressed. This problem is interesting because it enables restoration of historical footage while advancing spatiotemporal coherence techniques in computer vision.

Background Reading

Key papers to examine:
- Zhang et al. (2016) "Colorful Image Colorization"
- Lei et al. (2019) "Fully Automatic Video Colorization with Self-Regularization and Diversity"
- Lai et al. (2018) "Learning Blind Video Temporal Consistency"
- Recent advances from CVPR, ICCV, and ECCV (2020-2023)

Data

- Public video datasets: DAVIS dataset and Videvo stock footage
- Historical black and white footage from Internet Archive
- Self-created dataset: converting color videos to grayscale for ground-truth pairs

Proposed Method

1. **Frame-wise colorization**: U-Net architecture with VGG feature extractors. We'll start with pretrained models from Zhang et al. and add perceptual loss functions. We propose to build upon existing image colorization models (e.g., DeOldify)
2. **Temporal consistency**: Optical flow estimation (FlowNet2/RAFT) to propagate color between frames with a temporal consistency loss function.

Improvements over existing implementations:
- Specialized attention mechanism for spatial-temporal context
- Semantic understanding for consistent object colorization
- Lightweight architectures for longer video processing

Evaluation Plan

Qualitative: Visual comparisons with ground truth, consistency visualization across frames, attention heatmaps

Quantitative:
- Image quality: PSNR, SSIM, LPIPS
- Temporal consistency: warping error, temporal LPIPS
- User study (15-20 participants) rating naturalness and consistency
- Computational efficiency: processing time, memory usage

We expect to demonstrate improved temporal stability while maintaining colorization quality comparable to state-of-the-art methods.