Long-term dense motion estimation for video editing

Mini-project of UE Computer Vision

1 Objective

The objective of this project is to perform **video editing within an image sequence** by leveraging **dense optical flow fields** to propagate user-defined edits across time. Starting from a manual modification applied to a single frame, students will design and implement a method to transfer this editing consistently to the entire video sequence using optical flow integration.

This project aims to deepen the understanding of optical flow estimation, long-term motion propagation, and their practical use in video processing applications.

2 Background

Optical flow provides a dense pixel-wise motion field between consecutive frames in a video. When integrated over time, these motion fields can be used to track image content and propagate visual information across frames. Such techniques are fundamental in applications such as video editing, motion-aware filtering, and visual effects.

3 Topic selection

Students must rely on one of the following two references for optical flow computation:

Topic A – Classical Optical Flow

- D. Sun, S. Roth, M. J. Black, *Secrets of Optical Flow Estimation*, CVPR 2010. https://cs.brown.edu/people/dqsun/pubs/cvpr_2010_flow.pdf

Topic B – Learning-Based Optical Flow

- A. Dosovitskiy et al., *FlowNet: Learning Optical Flow with Convolutional Networks*, ICCV 2015. https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Dosovitskiy_FlowNet_Learning_Optical_ICCV_2015_paper.pdf

4 Data

Students may use **any image sequence or video of their choice**. The selected sequence must contain sufficient motion to justify the use of optical flow.

Recommended public sources include the **DAVIS dataset**, which provides high-quality, short video sequences well suited for optical flow estimation and video editing experiments: https://davischallenge.org/.

5 Methodology

The core of this project is the design and implementation **in Python** of a **dense and long-term motion estimation framework** and its application to a realistic video editing scenario.

5.1 Dense and long-term motion estimation

Students are required to develop a method to estimate dense pixel-wise motion across an image sequence. This method may rely on classical optical flow formulations (**Topic A**) or on learning-based approaches (**Topic B**).

Beyond pair-wise frame-to-frame estimation, the emphasis is placed on **long-term motion integration**, enabling the tracking of image content over multiple frames. Particular attention should be paid to temporal consistency, motion drift, and the handling of occlusions.

5.2 Video editing scenario

A video sequence of choice is selected, and a visible editing operation is applied to a single reference frame using an image-editing tool such as **GIMP**. Typical examples include texture replacement, logo insertion, object recoloring, or artistic modifications. This edited frame serves as the starting point for the video editing process.

5.3 Editing propagation via motion fields

Using the previously estimated dense and long-term motion fields, the goal is to propagate the editing consistently throughout the video sequence. The propagation strategy should preserve the spatial alignment of the edited content with the underlying motion in the scene.

Students are encouraged to explore creative solutions to limit error accumulation, handle occlusions, and maintain visual coherence over time.

5.4 Evaluation and analysis

The results are evaluated qualitatively by analyzing the temporal coherence and visual plausibility of the edited video. Failure cases, artifacts, and limitations of the proposed motion estimation and propagation strategy should be clearly identified and discussed.

6 Code

A baseline implementation is provided for reference, illustrating the propagation of a logo inserted in a single frame throughout a video sequence using optical flow: https://github.com/conze/UECompVis/tree/main/project. Students are expected to go beyond the provided baseline by demonstrating creativity in the video editing task. They are encouraged to experiment with different types of editing operations, to evaluate various optical flow estimation methods, and to explore multiple strategies for integrating optical flow over time in order to improve temporal consistency and reduce error accumulation.

7 Deliverables

1. **Code** with clear documentation
2. **Visual results**: original video, edited reference frame, final edited video sequence
3. **Oral presentation** of the work, supported by slides, summarizing the proposed method, results, and limitations

8 Expected learning outcomes

Students will gain a solid understanding of optical flow estimation, long-term motion integration, and motion-aware video editing.