BasicFlow: Lightweight SteadyFlow-Based Video Stabilization

Arnav Sharma (2023111033) Harshith Seera (2023101064)
CS7.404 Digital Image Processing

Introduction & Problem Statement

- Videos often suffer from jitter, wobble, parallax effects, and distortions due to handheld motion or scene depth changes.
- Video stabilization aims to correct these issues by estimating camera motion and producing a smoother visual trajectory between frames.
- SteadyFlow achieves high-quality stabilization by creating a spatially coherent optical flow, allowing us to smooth pixel profiles rather than feature trajectories.
- Our method, BasicFlow, provides a lightweight alternative by using RAFT flow, simple spatial masking, and efficient temporal smoothing.

Optical Flow

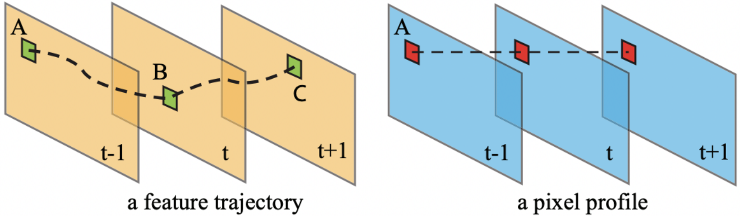
Optical Flow describes the apparent motion of pixels between two consecutive video frames. It provides a 2D motion vector (u, v) at each pixel, indicating how that pixel has moved.

Original Frame

Initial Optical Flow

BasicFlow

Feature Trajectories vs Pixel Profiles



- **Feature trajectories** track the motion of specific feature points across frames, but can break if features disappear or move out of view.
- **Pixel profiles** track the motion at the same pixel location over time, providing dense and stable information for every pixel.

Traditional Methods & Their Limitations

- Traditional stabilization uses global 2D models like affine or homography to approximate camera motion.
- These models assume the entire scene moves rigidly, which fails in the presence of depth variation or parallax.
- Feature-based tracking can be unreliable when features disappear, drift, or become unevenly distributed.
- Dynamic objects introduce inconsistent motion that global models cannot represent.
- As a result, traditional methods often produce wobble, distortion, or over-smoothing in complex scenes.

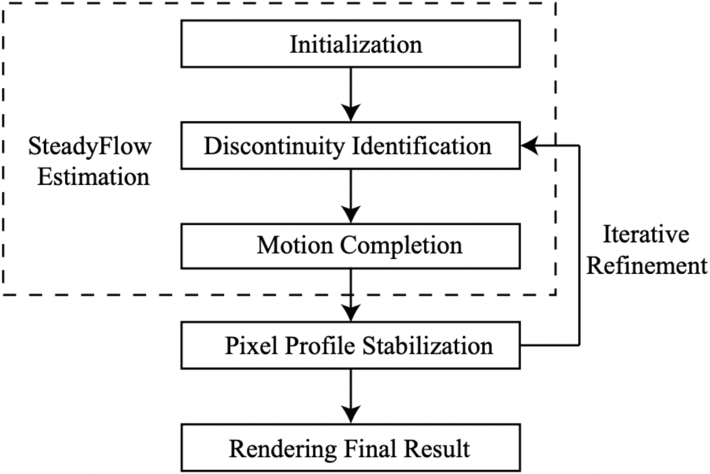
SteadyFlow

- SteadyFlow is a dense-flow video stabilization method designed to handle spatially-varying motion such as parallax and depth changes.
- Instead of relying on sparse feature trajectories, it analyzes pixel-wise motion across time using **pixel profiles**.
- The core idea is to build a spatially smooth optical flow field and then temporally smooth each pixel's motion to obtain a stable camera path.

SteadyFlow: The Process

- **1. Initialization**: Estimate dense optical flow and a global homography to capture coarse camera motion.
- **2. Discontinuity Identification**: Detect unreliable flow regions using spatial gradients and temporal inconsistency.
- **3. Motion Completion**: Fill in missing or inconsistent flow using mesh-based as-similar-as-possible warping to enforce spatial smoothness.
- **4. Pixel Profile Stabilization**: Build motion profiles for each pixel across time and smooth them using temporal optimization.
- **5. Iterative Refinement**: Repeat discontinuity detection and completion until the flow field becomes stable.

SteadyFlow Pipeline



- Iteratively refines discontinuous flow into a smooth field.
- Stabilizes the video by smoothing pixel motion profiles.

MeshFlow

- MeshFlow is a lightweight, online video stabilization method that replaces dense flow with a **sparse mesh-based motion model**.
- Instead of computing dense pixel motion, it tracks sparse features and propagates their motion to mesh vertices.
- Two median filters remove outliers and enforce spatial smoothness across the mesh.
- It avoids SteadyFlow's heavy flow refinement and ASAP motion completion.
- Predicts stabilization strength using PAPS (Predicted Adaptive Path Smoothing), removing the iterative search used by SteadyFlow.

Smoothing Formulas

SteadyFlow

$$E(P) = \sum_t \left(\|P(t) - C(t)\|^2 + \lambda_t \sum_r w_{t,r} \|P(t) - P(r)\|^2 \right)$$

MeshFlow (PAPS)

$$\lambda_t = \max\Big(0,\ \min(\lambda_t^{(1)}, \lambda_t^{(2)})\Big)$$

$$\lambda_t^{(1)} = -1.93\, T_v + 0.95, \qquad \lambda_t^{(2)} = 5.83\, F_a + 4.88$$

PAPS predicts the same smoothing weight λ_t used in the SteadyFlow energy, but estimates it directly from translation magnitude (T_v) and affine deformation (F_a), instead of performing SteadyFlow's slow iterative search.

What is BasicFlow?

- **BasicFlow is a lightweight stabilization pipeline we developed**, combining ideas from SteadyFlow and MeshFlow with modern deep optical flow.
- Uses **RAFT (Recurrent All-Pairs Field Transforms)** to obtain dense, accurate optical flow. RAFT performs iterative refinement over an all-pairs correlation volume, giving very reliable flow even in low-texture or complex regions.
- Instead of SteadyFlow's heavy discontinuity analysis, BasicFlow applies a **simple spatial gradient mask** to detect unreliable flow areas.
- Replaces SteadyFlow's ASAP motion completion with **iterative diffusion** (selective blurring) to smooth and fill discontinuities.
- Uses MeshFlow's **PAPS** (Predicted Adaptive Path Smoothing) to choose smoothing strength λ_t without iterative search.
- Produces the final stabilization by **smoothing per-pixel motion profiles**, similar in spirit to SteadyFlow but far more efficient.

BasicFlow: Pipeline

- **Dense Optical Flow (RAFT)** Compute high-quality dense optical flow between consecutive frames using RAFT with padding + cropping.
- **Discontinuity Detection** Compute Sobel gradients of the flow. Mark pixels with large gradient magnitude as unreliable (spatial mask).
- **Motion Completion via Diffusion** Run several iterations of selective blurring. Only update pixels inside the discontinuity mask $\rightarrow$ smooth, hole-free flow.
- **Pixel Profile Construction** Accumulate cleaned flows over time to build motion profiles $C(t)$ for every pixel.
- **Homography-Based Weight Estimation** Compute homographies between frames. Derive smoothing weights λ_t from translation and affine components.
- **Pixel Profile Smoothing (Jacobi Iteration)** Iteratively smooth each pixel's temporal motion profile to obtain stabilized motion trajectories $P(t)$.
- **Frame Warping and Output** Compute warp fields from $P(t) - C(t)$, warp each frame, crop borders, and write the stabilized video.

BasicFlow: Deviations from Literature

- **No temporal outlier detection** SteadyFlow uses spatial *and* temporal discontinuity checks; BasicFlow uses only a one-pass spatial gradient threshold.
- **No ASAP motion completion** SteadyFlow fills holes via a large sparse linear solve (ASAP warping); BasicFlow replaces this with simple **iterative diffusion** (selective blurring).
- **Lambda estimation from MeshFlow (PAPS)** SteadyFlow searches iteratively for the best λ_t; BasicFlow directly predicts it using MeshFlow's regression model.
- **Dense deep flow from RAFT** Removes reliance on sparse feature tracking (MeshFlow) or classical flow (SteadyFlow).
- **Simplified, linear pipeline** No iterative refinement, no heavy optimization $\rightarrow$ significantly faster in practice.

Dominating Foreground

Link:
`https://youtu.be/1FxWuKNQEiM`

Quick Zoom

Link:
`https://youtu.be/n95WHZevxD8`

Rolling Shutter

Link:
`https://youtu.be/QkzGF9ZRE0I`

All Comparisons

Link:
https://www.youtube.com/playlist?list=PL2gpCaN0OukK0CCQrF3CXxn65Ktbnit_8

Issues Addressed by BasicFlow

- **High computational cost in SteadyFlow** Replaces ASAP warping + iterative refinement with simple diffusion and predicted smoothing (PAPS), making the pipeline much faster.
- **Brittleness of feature tracking in MeshFlow** Uses dense RAFT optical flow instead of sparse features, avoiding failures in low-texture or uneven regions.
- **Latency from iterative parameter search** Removes SteadyFlow's iterative search for λ_t; BasicFlow predicts it instantly through MeshFlow's PAPS model.

Thank You!

Questions?