Visual Information Processing

Menandro Roxas, Ph.D.
Post Doctoral Fellow, Oishi Laboratory
Institute of Industrial Science
The University of Tokyo

Today’s Class

• Optical Flow and It’s Applications
  • Background
    • What is optical flow
    • Techniques
    • State-of-the-art methods
  • Applications

Optical Flow Basic

• Brightness Constancy
  \[ I(x+u, y+v, t+1) - I(x, y, t) = 0 \]

• First-order approximation
  \[ I(x+u, y+v, t+1) = I(x, y, t) + \frac{dI}{dx} u + \frac{dI}{dy} v + \frac{dI}{dt} \]

Optical Flow (1D)

• Over-constrained
  − Neighboring pixels have equal motion?

• More equations
  \[
  \begin{bmatrix}
    I_x & I_y \\
    T_x & T_y
  \end{bmatrix}
  \begin{bmatrix}
    u \\
    v
  \end{bmatrix}
  =
  \begin{bmatrix}
    -I_t \\
    -T_t
  \end{bmatrix}
  \]

• Is this good enough?
Issues with Optical Flow Basic

- Aperture problem
- Occlusions
- Textureless Region
- Large Motion

Aperture Problem

![Barber's pole illusion](image)

- Actual Motion
- Projected Motion

Aperture Problem: Solution

- Regularization (Smoothness Constraint)
  - Neighboring pixels have similar motion (more or less)
- Isotropic information
  - Corners
  - High textures

\[
\begin{align*}
\text{argmin}_{u,v} & \left( I(x,y) + I(x+1,y) \right)^2 + \left( \frac{du}{dx} \right)^2 + \left( \frac{dv}{dy} \right)^2 \\
\text{subject to} & \frac{du}{dx} \approx 0 \\
& \frac{dv}{dy} \approx 0 \\
& u(x+1,y) - u(x,y) \approx 0
\end{align*}
\]

Aperture Problem: Solution

- Smoothness Constraint

Occlusion

- Undefined motion in occluded area
  - But still constrained. How?
Occlusion: Solution

- Step 1: Detection of occluded area
- Step 2: Deciding what “kind” of motion it is

Textureless Region

- Image gradients are zero
- $I_x = 0 \quad I_y = 0$

Textureless Region

- Regularization + Coarse-to-fine approach
  - Make the image small
  - Make neighboring pixels have similar motion
  - Increase resolution

Textureless Regions: Solution

Large Motion

- Coarse-to-fine approach
  - Bring the two corresponding points closer
- Sparse matching (global scale feature matching)
  - Search whole image for similar features

Large Motion: Solution
Different approaches

• Variational method
  – Naturally dense (complete) and continuous

• Feature-matching method
  – Patch matching
  – Naturally sparse and discrete

Variational method

• Sum of convex energies
  – Data (Brightness constancy)
  – Regularizer (Smoothness constraint)
  – Large Displacement Handling
  – Epipolar Geometry

Variational Method

• Brightness Constancy
  – Lukas-Kanade

• Regularizer
  – Least-Squares (Horn-Schunk) \( \frac{(du)^2}{dx^2} + \frac{(du)^2}{dy^2} \)
  – Total Variation \( \sqrt{\frac{(du)^2}{dx^2} + \frac{(du)^2}{dy^2}} \)
  – Non-Local (Median Filter + extension)

Variational Method

• Median filtering vs. Weighted Median filtering

NOISY SIGNAL
TV DENOISING

Variational Method

• Comparison of regularizers

Leastsquares
Total Variation
median filtering
weighted median filtering
Variational Method

- Large Displacement
  - Global feature matching
  - Constraint the pixel to the detected motion

\[ u(x, y) - u(x, y) = 0 \]

Variational Method

- Epipolar Geometry
  - Constrain the motion along a trajectory

Example of Minimization Function

\[
\arg \min_{u,v} \left( \| F_{12} \| + \| F_{21} \| + \lambda \| u \| + \lambda \| v \| \right)
\]

Solution to Variational Method

- Euler-Lagrange
- Gradient descent
- Primal-Dual decomposition
  - Used for non-differentiable functions such as l1-norm

Feature matching method

- Handles large displacement
- Mostly sparse
- SIFT, SURF, SIFT-Flow etc.
- DeepMatching
Evaluation Database

- Training + Testing
  - Middlebury
  - Kitti
  - MPI Sintel

Evaluation Database: Middlebury

- 12 training sets
- 12 testing sets
- Advantages
  - Real + synthetic images
  - Random object motion
- Disadvantages
  - Very small motion (~20 pixels)
  - Few images

Evaluation Database: KITTI2012

- 193 training
- 194 testing
- Advantages
  - Real world images
- Disadvantages
  - Partial ground truth
  - Outdoor only
  - Motions are due to camera movement (Static scene)

Evaluation Database: KITTI2015

- 200 training
- 200 testing
- Advantages
  - Real world images
  - Dynamic motion (moving objects)
  - More challenging
- Disadvantages
  - Challenging sets are very few (difficult to make conclusive remarks)

Evaluation Database: MPI Sintel

- 23 training
- 12 testing
- Advantages
  - Dynamic motion
  - Challenging large motion
  - Most difficult of the three database (Middlebury, Kitti and MPI)
- Disadvantages
  - Synthetic (CGI)

Best Performing Methods

- General Purpose
  - FlowFields
    - D. Bailer et al., "Flow Fields: Dense Correspondence Fields for Highly Accurate Large Displacement Optical Flow Estimation", ICCV 2015
  - EpicFlow
- Epipolar Geometry
  - PRSM
    - C. Vogel et al., "3D Scene Flow Estimation with a Piecewise Rigid Scene Model", IJCV 2015
  - SPS-FL
- Variational
  - Classic++, Classic++NL
  - Variational +
    - DeepFlow
FlowFields

For n=3

n=2

n=1

ED Free Initialization

Propagation

Range Search

SPS-Fl - Yamaguchi

MRF model
– Piecewise rigid moving planes (same as PRSM)

- Slanted-plane smoothing

\[ d(p, \theta_i) = A_i p_x + B_i p_y + C_i \]

Disparity at pixel \( p \)

A, B, C are plane parameters

(px, py) is the center of a segment

EpicFlow

\[ F_{W}(p) = \frac{\sum_{(p_n, p_m) \in M} k_D(p_n, p) |F(p_m)|}{\sum_{(p_n, p_m) \in M} k_D(p_n, p)} \]

Dense interpolation

Energy minimization

Results

(a) ANN [10]

(b) Our Flow Field

(c) Our outlier filtered Flow Field

(d) Ground truth

PRSM

- Piecewise rigid moving planes
**Results**

Classic++, Classic+NL
- Classic++, 5x5 median filter
- Classic+NL, Edge-adaptive median filter
  - Improved boundaries

**DEEPFLOW**
- Accurate, but DeepMatching is very slow

**Summary**
- Best methods
  - Handles large displacement
    - Typical case for databases
  - Handles occlusions
    - How to correctly guess the motion of missing pixels?

**Classic++, Classic+NL**

**Applications**
- Video inpainting
  - M. Roxas, “Video Completion via Spatio-Temporally Consistent Motion Inpainting”, CVA 2014
- Dense 3D reconstruction
- Video Segmentation
- Raindrop Detection and Removal
Video Inpainting

- Roxas

Iterative step:
- optical flow estimation
- improved color propagation
- update mask function

Dense 3D Reconstruction

Results

Methodology

Live Dense Reconstruction with a Single Moving Camera
Segmentation

(a) Frame 1
(b) Frame 2
(c) initial optical flow
(d) updated optical flow

Raindrops Detection

Fig. 1. An example of the results of our proposed detection and removal method. (a) Scene with raindrop. (b) Dense long trajectories. (c) Matching of trajectories obtained by matching. (d) Trajectory-based video completion.

Fig. 2. The timeline of our method.

Results

Summary

• What is optical flow? Specifically, variational optical flow

• Brief background of current state-of-the-art methods in optical flow

• Brief overview of applications using optical flow information