Volumetric Correspondence Networks for Optical Flow

Gengshan Yang¹, Deva Ramanan¹,²
Carnegie Mellon University¹, Argo AI²

Introduction

Optical flow / visual correspondences

Related work: cost volume filtering

Contributions

∙ Efficient higher-dimension (4D) cost volume processing
  ∙ Separable volumetric filters: reduce computation and parameters
  ∙ Multi-channel cost volumes: capture multiple dimensions of pixel similarity
  ∙ Adaptive cost volumes allow networks to generalize across tasks: train and test a single network for both flow+stereo

Approach

Key idea: 4D filters

Prior method
  1) reshape to $\mathcal{C}(x, y, z) \in \mathbb{R}^{T \times h \times w}$
  2) multi-channel 2D convs with $W \in \mathbb{R}^{n \times r \times c}$

Ours
  1) keep $\mathcal{C}(u, v, x, y) \in \mathbb{R}^{T \times c \times h \times w}$
  2) 4D convs with $W \in \mathbb{R}^{n \times r \times c \times 2 \times 3}$

Separable filters

$\mathcal{C}(u, v, x, y) = \mathcal{C}_x(u, v, x) * \mathcal{C}_y(u, v, y) + \mathcal{C}_y(u, v, x) * \mathcal{C}_x(u, v, y)$

Multi-channel 4D cost volumes

$\mathcal{C}(u, v, x, y) = \sum_{c=1}^{C} \mathcal{C}_c(u, v, x, y)$

Experiments

Benchmarks

Generalize: 1D to 2D disparity

Application: stereo matching with imperfect rectification

Ablation

Code available: github.com/gengshan-y/VCN