

相互结构引导滤波 TV-L1 变分光流估计

葛利跃¹, 张聪炫^{1,2}, 陈震^{1,2}, 黎明², 陈昊²

(1. 南昌航空大学测试与光电工程学院, 江西南昌 330063; 2. 南昌航空大学无损检测技术教育部重点实验室, 江西南昌 330063)

摘要: 由于光流场既包含物体的运动信息, 又包含场景的三维结构信息, 因此光流计算技术是计算机视觉和机器视觉领域研究的重要任务之一. 针对现有光流计算方法在图像边缘保护方面存在过度平滑问题, 提出一种基于相互结构引导滤波的 TV-L1 (Total Variational with L1 norm, TV-L1) 变分光流估计方法. 通过提取置信度较高的图像相互结构区域, 构造基于相互结构引导滤波的全局目标函数, 并采用金字塔分层细化与交替迭代方案结合的策略进行优化, 该方法可以较好的保护图像边缘信息. 最后采用标准测试图像集对本文方法与现有代表性变分方法 LDOF (Large Displacement Optical Flow, LDOF), CLG-TV (Combined Local-Global Total Variation, CLG-TV), Classic + +, NNF (Nearest Neighbor Fields, NNF) 以及深度学习方法 FlowNet2.0 进行对比, 实验结果表明本文方法具有较高的光流估计精度与鲁棒性, 尤其对图像边缘保护具有显著的效果, 并且在运动目标检测, 机器人避障等方面具有一定应用前景.

关键词: 光流计算; 计算机视觉; 机器视觉; 图像边缘保护; 图像相互结构; 深度学习; 运动目标检测; 机器人避障

中图分类号: TP391 文献标识码: A 文章编号: 0372-2112 (2019)03-0707-07
电子学报 URL: <http://www.ejournal.org.cn> DOI: 10.3969/j.issn.0372-2112.2019.03.027

Mutual-Structure Guided Filtering Based TV-L1 Optical Flow Estimation

GE Li-yue¹, ZHANG Cong-xuan^{1,2}, CHEN Zhen^{1,2}, LI Ming², CHEN Hao²

(1. School of Measuring and Optical Engineering, Nanchang Hangkong University, Nanchang, Jiangxi 330063, China;

2. Key Laboratory of Nondestructive Testing, Ministry of Education, Nanchang Hangkong University, Nanchang, Jiangxi 330063, China)

Abstract: Because the optical flow field contains both the motion information of the object and the three-dimensional structure information of the scene, optical flow calculation technology is one of the important tasks in the field of computer vision and machine vision. But, for the existing optical flow methods, there is over-smoothing problem in image boundary preserving. This paper proposes a global TV-L1 (Total Variational with L1 norm, TV-L1) variational optical flow computation method based on the mutual-structure guided filtering. By extracting the mutual-structural regions of the image with higher confidence, we construct the global mutual-structure guided filtering objective function, and optimize the algorithm via combining the pyramid layering strategy with the alternating iteration scheme. This method can better preserve the image boundary information. Finally, we compare the proposed method with the existing representative variational methods LDOF (Large Displacement Optical Flow, LDOF), CLG-TV (Combined Local-Global Total Variation, CLG-TV), Classic + +, NNF (Nearest Neighbor Fields, NNF) and deep learning method FlowNet2.0 by using standard test image datasets. The experimental results demonstrate that the presented method has more accuracy and better robustness than the other evaluated methods, especially has the significant effect of boundary preserving in the areas, and it has application prospects in moving target detection, robot obstacle avoidance, and so on.

Key words: optical flow calculation; computer vision; machine vision; image boundary preserving; image mutual-structure; deep learning; moving target detection; robot obstacle avoidance

1 引言

光流场是图像像素点灰度变化的瞬时速度场, 其

包含了物体和场景的运动与结构信息, 因此光流计算技术是计算机视觉和机器视觉领域研究的重要任务之一. 近年来, 随着光流计算技术在估计精度、鲁棒性与时

收稿日期: 2018-03-12; 修回日期: 2018-07-09; 责任编辑: 覃怀银.

基金资助: 国家自然科学基金 (No. 61866026, No. 61772255, No. 61866025); 江西省优势科技创新团队计划 (No. 20152BCB24004, No. 20165BCB19007); 航空科学基金 (No. 2016ZC56005); 江西省青年科学基金 (No. 20171BAB212012); 江西省重点研发计划 (No. 20161BBE50080); 江西省研究生创新专项资金项目资助 (No. YC2017-S339)

效率等方面的大幅提升,已广泛应用于机器人视觉系统^[1]、医疗诊断^[2]等诸多领域.

根据研究内容的不同,光流计算技术研究可大致分为两类^[3]:一种是对光流能量泛函的研究,另一种是对光流计算策略的研究.其中,对能量泛函的研究主要是对数据项和平滑项进行设计^[4,5],以提高光流估计精度.对光流计算策略的研究,主要是研究使用特定方法策略处理某一类或几类问题^[6,7],进而提高特定场景下光流估计的精度.近年来,基于滤波优化的光流计算策略逐渐成为解决光流计算鲁棒性难题的重要手段^[8,9],但是该类方法易导致图像边缘模糊.针对该问题,本文提出一种基于相互结构引导滤波的 TV-L1 光流计算模型,通过在光流计算模型中引入置信度较高的相互结构区域,可以较好的保护图像边缘信息.实验结果表明本文方法具有较高的光流估计精度与鲁棒性,尤其对图像边缘保护具有显著的效果.

2 图像相互结构

2.1 图像区域定义

本文根据相互结构引导滤波的定义,将图像序列中的对应区域划分为以下三种不同结构类型^[10]:

表 1 图像不同区域类别划分

类别	特点
相互结构区域	具有相同或者相似的边缘轮廓结构,不受图像亮度变化影响.
不一致区域	图像序列中的运动与图像边缘不一致.
平滑区域	不包含重要结构信息,且易受噪声影响,通常可以归类为相互结构区域.

根据表 1 所示图像序列中不同对应区域的定义,为了克服由于中值滤波平滑作用所引起的边缘模糊问题,本文主要利用图像序列中置信度较高的相互结构区域对光流进行引导滤波,以保护图像与运动边缘轮廓信息.

2.2 图像区域提取

为了提取置信度较高的相互结构区域,本文通过定义图像序列相互结构区域提取因子来获取图像序列中的相互结构区域:

$$\rho(I_p^1, I_p^2) = \frac{\text{cov}(I_p^1, I_p^2)^2}{(\sigma(I_p^1) + \tau_1)(\sigma(I_p^2) + \tau_2)} \quad (1)$$

式(1)中, I^1, I^2 分别表示相邻两帧图像序列, p 是图像中任意局部区域的中心像素点, $\text{cov}(\cdot)$ 和 $\sigma(\cdot)$ 分别是图像序列对应区域的像素灰度协方差和方差, τ_1, τ_2 是正则化系数. 式(1)中, $\rho(I_p^1, I_p^2) \in [0, 1]$, 当 $\rho(I_p^1, I_p^2) \rightarrow 1$ 时, 表示图像序列对应区域的相关性越大, 该对应区域属于相互结构区域; 反之, 表示图像序列对应区域的相关性越小, 该对应区域属于不一致区域. 本文设置当

$\rho(I_p^1, I_p^2) \geq 0.8$ 时, 对应区域为相互结构区域.

3 相互结构引导滤波光流计算模型

3.1 相互结构引导滤波目标函数

令 $\mathbf{w} = (\mathbf{u}, \mathbf{v})^T$ 表示连续两帧图像间的估计光流场, 其中 \mathbf{u} 和 \mathbf{v} 分别表示光流水平和垂直方向分量. 假设 $\tilde{\mathbf{w}} = (\tilde{\mathbf{u}}, \tilde{\mathbf{v}})^T$ 表示两帧图像间的引导光流场, 则图像相互结构区域中任意局部区域的估计光流可用引导光流线性表示为:

$$f(\tilde{\mathbf{w}}, \mathbf{w}, a_p, a'_p) = \sum_{i \in N(p)} (a_p \tilde{\mathbf{w}}_i + a'_p - \mathbf{w}_i)^2 \quad (2)$$

式(2)中, $N(p)$ 表示图像中以像素点 p 为中心的任意局部区域, i 是该区域内任意邻域像素点, a_p 和 a'_p 表示局部区域内由引导光流到估计光流的线性系数.

同理, 该区域内的引导光流可用对应区域的估计光流线性表示为:

$$f(\mathbf{w}, \tilde{\mathbf{w}}, b_p, b'_p) = \sum_{i \in N(p)} (b_p \mathbf{w}_i + b'_p - \tilde{\mathbf{w}}_i)^2 \quad (3)$$

式(3)中, b_p, b'_p 与 a_p, a'_p 定义类似. 根据式(2)和式(3) 本文首先定义图像相互结构引导滤波的相似项 E_m :

$$E_m(\mathbf{w}, \tilde{\mathbf{w}}, a, a', b, b') = \sum_{p \in \Omega} (f(\tilde{\mathbf{w}}, \mathbf{w}, a_p, a'_p) + f(\mathbf{w}, \tilde{\mathbf{w}}, b_p, b'_p)) \quad (4)$$

式(4)中, Ω 表示图像区域, $\{a, a', b, b'\}$ 表示线性系数集合. 由于直接利用式(4)对图像估计光流进行优化会导致光流稀疏的问题. 因此, 为获得稠密的光流优化估计, 定义相互结构引导滤波平滑项 E_s :

$$E_s(a, b) = \sum_{p \in \Omega} (\tau_1 a_p^2 + \tau_2 b_p^2) \quad (5)$$

式(5)中, τ_1 和 τ_2 是权重系数, 决定了引导滤波的平滑程度. 同时为了抑制过度平滑, 本文引入平滑惩罚项 E_{prior} :

$$E_{\text{prior}}(\mathbf{w}, \tilde{\mathbf{w}}) = \sum_{p \in \Omega} \|\mathbf{w}_p - \mathbf{w}_p^0\| + \|\tilde{\mathbf{w}}_p - \tilde{\mathbf{w}}_p^0\| \quad (6)$$

式(6)中, \mathbf{w}_p^0 和 $\tilde{\mathbf{w}}_p^0$ 分别表示原始估计光流和引导光流中以像素点 p 为中心的任意局部区域光流. \mathbf{w}_p 和 $\tilde{\mathbf{w}}_p$ 分别表示相互结构引导滤波后的估计光流和引导光流. 惩罚项 E_{prior} 主要用于约束相互结构区域内原始光流与滤波优化后光流不发生严重偏离, 以克服平滑项可能引起的过度平滑问题.

结合式(4)~(6)可得相互结构引导滤波目标函数如下:

$$E_{\text{MS}}(\mathbf{w}, \tilde{\mathbf{w}}, a, a', b, b') = E_m(\mathbf{w}, \tilde{\mathbf{w}}, a, a', b, b') + E_s(a, b) + E_{\text{prior}}(\mathbf{w}, \tilde{\mathbf{w}}) \quad (7)$$

式(7)中, $\{a, a', b, b'\}$ 是线性系数集合, 本文将其计算引入到相互结构引导滤波 TV-L1 光流模型中, 在优化光流的同时优化相互结构引导滤波线性系数, 以获取最优光流估计.

3.2 基于相互结构引导滤波的 TV-L1 光流计算模型

基于 L1 范数的全变分 (Total Variational with L1 norm, TV-L1) 光流估计函数是现阶段应用最广泛的通用模型,主要由数据项 E_d 和正则化项 E_r 组成:

$$E(\mathbf{w}) = \int_{\Omega} \{ \varphi(|I(\mathbf{x} + \mathbf{w}) - I(\mathbf{x})|) + \varphi(|\nabla \mathbf{w}|) \} dx \quad (8)$$

式(8)中, $\mathbf{w} = (\mathbf{u}, \mathbf{v})^T$ 表示像素点集合 $\mathbf{x} = (x, y)^T$ 的估计光流, $\nabla = (\partial_x, \partial_y)^T$ 表示图像空间梯度算子, $\varphi(s) = (s^2 + \varepsilon^2)^{1/2}$ 表示惩罚函数,其中 $\varepsilon = 0.001$.

针对传统 TV-L1 光流估计模型易受图像噪声影响导致光流估计鲁棒性较差的问题,基于中值滤波的非局部 TV-L1 光流估计模型逐渐成为变分光流估计的主流方法,但中值滤波在抑制光流估计溢出点的同时会引起图像和运动边缘模糊.为了获取鲁棒光流估计的同时保护光流结果中的图像与运动边缘,本文在传统的 TV-L1 变分光流模型中引入基于相互结构引导滤波的边缘约束项,提出基于相互结构引导滤波的 TV-L1 光流估计模型:

$$\begin{aligned} E(\mathbf{w}, \tilde{\mathbf{w}}, a, a', b, b') = & \int_{\Omega} \varphi(|I(\mathbf{x} + \mathbf{w}) - I(\mathbf{x})|) dx \\ & + \int_{\Omega} \varphi(|\nabla \mathbf{w}|) dx \\ & + \sum_{p \in \Omega} (f(\tilde{\mathbf{w}}, \mathbf{w}, a_p, a'_p) \\ & + f(\mathbf{w}, \tilde{\mathbf{w}}, b_p, b'_p)) \\ & + \sum_{p \in \Omega} (\tau_1 a_p^2 + \tau_2 b_p^2) \\ & + \sum_{p \in \Omega} \|\mathbf{w}_p - \mathbf{w}_p^0\| + \|\tilde{\mathbf{w}}_p - \tilde{\mathbf{w}}_p^0\| \end{aligned} \quad (9)$$

式(9)中第一项为数据项,第二项为正则化项,其他求和项为相互结构引导滤波约束项.该模型利用图像中受信任的相互结构区域对光流进行引导滤波,在抑制光流溢出点的同时能够较好的保留图像和运动边缘信息.

3.3 线性化光流计算模型

为了优化方程(9),本文引入由粗到细估计策略,将式(9)的最小化线性求解转换为传统 TV-L1 光流模型金字塔分层估计和相互结构引导滤波交替迭代计算.首先由式(9)中传统 TV-L1 光流计算模型可推导出其对应的 Euler-Lagrange 方程如下:

$$\varphi'_s \cdot \text{div}(|\nabla \mathbf{w}|) = \varphi'_d \cdot (|I_x \cdot \mathbf{w} + I_t| \cdot I_x) \quad (10)$$

式(10)中, I_x 和 I_t 分别表示像素点亮度的空间和时间偏导数. $\varphi'_s = \varphi'_s(|I_x \cdot \mathbf{w} + I_t|)$ 和 $\varphi'_d = \varphi'_d(|\nabla \mathbf{w}|)$ 分别表示数据项和正则化项的一阶偏导数, div 表示光流散度.由于式(10)中包含 φ'_s 和 φ'_d 等半隐式,故采用金字塔分层

内外部迭代消除未知参数.假设金字塔采样层数为 n ,当前图像为金字塔第 k 层 ($1 \leq k \leq n$),则式(10)可写为:

$$\varphi'_s{}^{k'} \cdot \text{div}(|\mathbf{w}^k + d\mathbf{w}^k|) = \varphi'_d{}^{k'} \cdot (|I_x^k \cdot d\mathbf{w}^k + I_t^k| \cdot I_x^k) \quad (11)$$

式(11)中, \mathbf{w}^k 表示第 k 层图像光流初始值, $d\mathbf{w}^k$ 表示第 k 层图像光流计算增量. $\varphi'_s{}^{k'} = \varphi'_s{}^{k'}(|I_x^k \cdot d\mathbf{w}^k + I_t^k|)$, $\varphi'_d{}^{k'} = \varphi'_d{}^{k'}(|\mathbf{w}^k + d\mathbf{w}^k|)$. 根据图像金字塔第 k 层图像光流的初始值 \mathbf{w}^k 和增量 $d\mathbf{w}^k$ 可得第 k 层光流的输出值即第 $k+1$ 层图像光流估计初始值为:

$$\mathbf{w}^{k+1} = \mathbf{w}^k + d\mathbf{w}^k \quad (12)$$

在金字塔光流迭代过程中,为抑制各层光流估计结果中的溢出点,避免光流在图像和运动边缘过度平滑,本文将式(9)中相互结构引导滤波目标函数转换为图像金字塔层的相互结构引导滤波对当前层输出光流进行优化.

由于相互结构引导滤波目标函数中包含线性系数 $\{a, a', b, b'\}$ 等未知参数,本文利用图像金字塔层估计光流与线性系数集合进行交替迭代,在更新线性系数的同时实现图像金字塔层估计光流的引导滤波优化.令 $\{a^t, a'^t, b^t, b'^t\}$ 表示第 t 次迭代时的初始线性系数集合, \mathbf{w}_p^t 和 $\tilde{\mathbf{w}}_p^t$ 分别表示对应的估计光流和引导光流,则由式(1)、(2)和(3)可以推导出线性集合 $\{a^t, a'^t, b^t, b'^t\}$ 的迭代更新公式为:

$$\begin{cases} a_p^t = \frac{\text{cov}(\tilde{\mathbf{w}}_p^t, \mathbf{w}_p^t)}{\sigma(\tilde{\mathbf{w}}_p^t) + \tau_1}, & a'^t = \mu(\mathbf{w}_p^t) - a_p^t \cdot \mu(\tilde{\mathbf{w}}_p^t) \\ b_p^t = \frac{\text{cov}(\mathbf{w}_p^t, \tilde{\mathbf{w}}_p^t)}{\sigma(\mathbf{w}_p^t) + \tau_2}, & b'^t = \mu(\tilde{\mathbf{w}}_p^t) - b_p^t \cdot \mu(\mathbf{w}_p^t) \end{cases} \quad (13)$$

式(13)中, $\mu(\mathbf{w}_p^t)$ 和 $\mu(\tilde{\mathbf{w}}_p^t)$ 表示该区内估计光流和引导光流的平均值.更新相互结构引导滤波线性系数后,可得相互结构引导滤波的估计光流和引导光流交替优化迭代公式如下:

$$\begin{cases} \tilde{\mathbf{w}}_p^{t+1} = \frac{1}{M_w^t} (\Phi_w^t \mathbf{w}_p^{t+1} + K_w^t + \tilde{\mathbf{w}}_p^t) \\ \mathbf{w}_p^{t+1} = \frac{1}{M_w^t} (\Phi_w^t \tilde{\mathbf{w}}_p^{t+1} + K_w^t + \mathbf{w}_p^t) \end{cases} \quad (14)$$

式(14)中, Φ_w^t 、 Φ_w^t 、 K_w^t 、 K_w^t 、 M_w^t 和 M_w^t 是系数项,可由相互结构滤波系数集合 $\{a_p^t, a'^t, b_p^t, b'^t\}$ 计算得出.

根据式(13)、(14)中相互结构引导滤波线性系数和光流迭代公式,在图像金字塔第 k 层得到输出光流后,首先利用式(13)更新线性系数 $\{a_p^t, a'^t, b_p^t, b'^t\}$,然后利用更新后的线性系数通过式(14)计算更新引导光流 $\tilde{\mathbf{w}}_p^{t+1}$ 与估计光流 \mathbf{w}_p^{t+1} ,对以上步骤进行交替迭代,最终输出相互结构引导滤波后的估计光流 \mathbf{w}_p^{t+1} 作为当前

图像层的输出光流代入下一层图像进行计算,直至金字塔最底层原分辨率图像层输出最终光流估计结果。

4 实验与分析

4.1 光流评价标准

本文采用平均角误差 (AAE) 和平均端点误差 (AEE) 对光流估计结果进行量化评价,其中 AAE 反映估计光流整体偏离光流真实值的角度;AEE 反映估计光流整体偏离光流真实值的距离:

$$AAE = \frac{1}{N} \arccos \left(\frac{1 + \mathbf{u}_E \times \mathbf{u}_G + \mathbf{v}_E \times \mathbf{v}_G}{\sqrt{\mathbf{u}_E^2 + \mathbf{v}_E^2 + 1} \cdot \sqrt{\mathbf{u}_G^2 + \mathbf{v}_G^2 + 1}} \right) \quad (15)$$

$$AEE = \frac{1}{N} \sqrt{(\mathbf{u}_E - \mathbf{u}_G)^2 + (\mathbf{v}_E - \mathbf{v}_G)^2} \quad (16)$$

式中, $(\mathbf{u}_E, \mathbf{v}_E)^T$ 表示光流估计结果, $(\mathbf{u}_G, \mathbf{v}_G)^T$ 表示光流真实值, N 表示图像中像素点的总数。

4.2 对比方法

本文分别选取现有代表性的变分方法 LDOF^[6]、CLG-TV^[4]、Classic++^[9]、NNF^[11] 和深度学习方法 FlowNet2.0^[12] 进行量化对比与定性分析。其中 LDOF 采用的是基于特征匹配的光流计算模型;CLG-TV 采用的是全局与局部结合的光流计算模型,主要使用双边滤波提高光流估计的精度与鲁棒性;Classic++ 采用的是经典的鲁棒参数变分模型,优化过程中使用中值滤波去除光流结

果中的异常值;NNF 通过在优化光流过程中采用加权中值滤波提高了光流估计的鲁棒性;FlowNet2.0 是基于卷积神经网络的深度学习光流计算模型。

4.3 Middlebury 数据集实验

首先采用 Middlebury 数据库提供的标准测试序列 Grove2 和 Urban2 为例验证本文方法光流估计边缘保护效果。其中 Grove2 场景中存在复杂边缘结构和运动不连续,Urban2 场景存在大位移运动和小位移运动。图 1 分别展示了 Grove2 和 Urban2 的参考帧、光流真实值以及各对比方法的光流估计结果。从图 1 中可以看出,本文方法光流估计效果优于各对比方法,为了进一步说明本文方法的边缘保护特性,图 2 分别展示了图 1 中标签区域的局部放大图,从图中①、②区域可以看出,LD-OF 算法在树枝边缘处存在较大的误差;CLG-TV 算法结果存在模糊和大量噪声;Classic++ 算法在树叶边缘区域存在过度平滑;FlowNet2.0 算法没有估计出标签①、②中树枝部分的光流;NNF 算法在树枝边缘区域的细节上存在较多光流溢出点;本文方法在标签①、②区域的光流估计表现出较好效果,说明本文方法对细小边缘具有更好的保护效果。同样在标签③、④区域,可以观察到本文方法光流估计效果最接近于真实光流值,说明本文方法针对图像中相近亮度的弱边缘区域光流估计同样具有较好的估计效果和适用性。

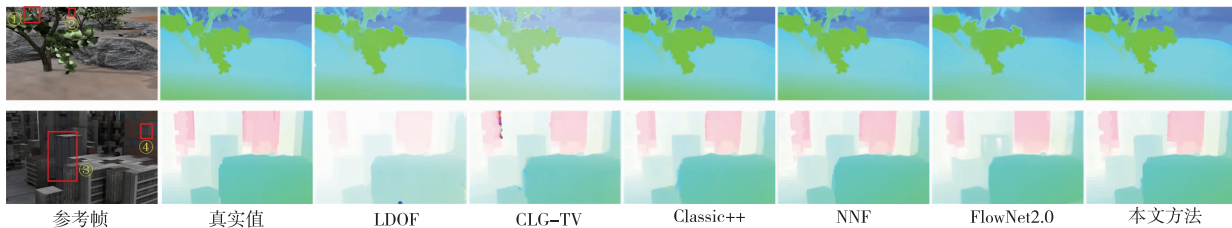


图1 Grove2和Urban2参考帧、光流真实值以及光流估计结果,第一行为Grove2,第二行为Urban2

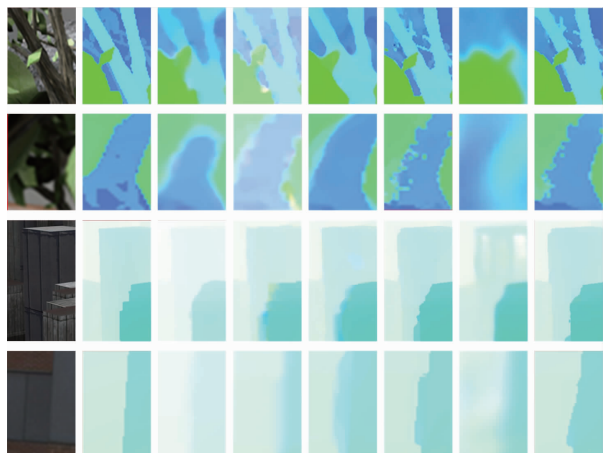


图2 标签区域局部放大图

在验证本文方法的边缘保护特性后,采用 Middlebury 数据库提供的八组测试图像序列对本文方法和各对比方法进行综合评价与量化对比,结果如图 3 所示。从图中可以看出,本文方法光流估计整体效果最优,尤其是针对复杂边缘结构、运动不连续、弱纹理区域等场景具有更好的适用性。

表 2 分别列出了本文方法与对比算法针对 Middlebury 测试图像序列的误差统计和时间消耗。可以看出,本文方法仅在 Dimetrodon 序列误差略大于 LDOF 和 Classic++ 算法,其他测试序列均取得最优表现,并且平均误差最小,说明本文方法具有更好的光流估计精度与边缘保护特性。在时间消耗方面,本文方法时间消耗相对较长,主要是因为金字塔每层图像上均需要完整遍历整幅图像以搜索相互结构区域,该过程导致时间消耗较大。

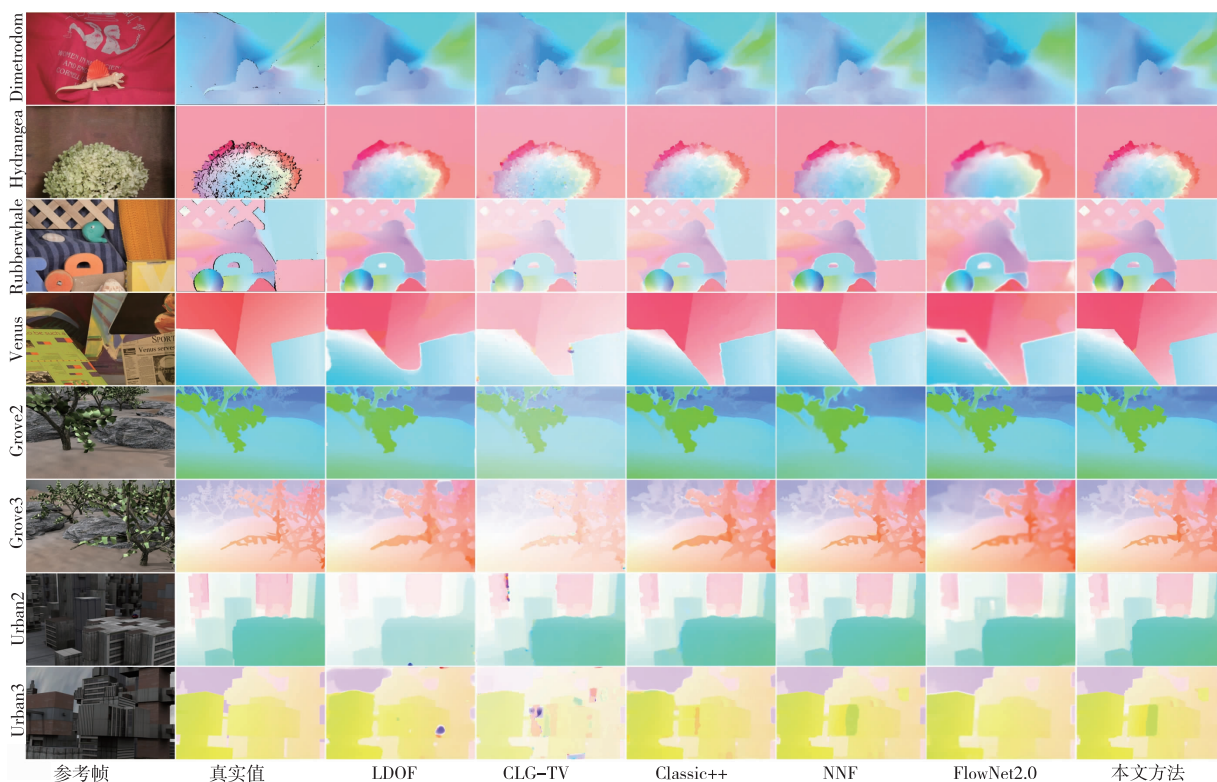


图3 Middlebury数据库图像序列光流估计结果

表 2 Middlebury 数据库图像序列光流估计误差与时间消耗对比 (时间消耗单位:s)

图像序列	LDOF	CLG-TV	Classic + +	NNF	FlowNet2.0	本文方法
	AAE/AEE/Time	AAE/AEE/Time	AAE/AEE/Time	AAE/AEE/Time	AAE/AEE/Time	AAE/AEE/Time
Dimetrodon	2.21/0.120/10.8	3.91/0.194/148.6	2.53/0.128/159.9	3.10/0.154/400.3	5.13/0.282/4.3	2.91/0.145/456.1
Hydrangea	2.07/0.176/13.1	2.62/0.216/154.9	1.79/0.152/156.0	1.87/0.153/335.9	3.14/0.286/4.2	1.79/0.150/395.1
RubberWhale	4.24/0.124/20.0	3.94/0.122/190.6	2.65/0.081/161.8	2.43/0.075/289.4	8.92/0.274/4.3	2.37/0.074/368.9
Venus	7.79/0.468/8.4	6.85/0.434/161.1	4.14/0.266/105.7	3.20/0.229/200.9	5.22/0.414/4.2	3.12/0.225/311.8
Grove2	2.27/0.178/25.5	2.62/0.185/255.5	2.07/0.141/241.3	1.40/0.102/495.8	4.57/0.310/4.4	1.35/0.097/594.1
Grove3	6.39/0.693/25.4	6.81/0.703/224.8	6.05/0.615/221.4	4.27/0.413/432.8	7.94/0.864/4.5	4.22/0.417/530.1
Urban2	2.90/0.403/31.8	3.13/0.402/241.7	2.54/0.362/222.8	2.14/0.260/383.0	4.81/0.517/4.4	1.88/0.235/619.4
Urban3	6.27/0.669/46.6	11.78/1.192/269.6	4.60/0.566/223.7	3.08/0.536/396.4	4.93/0.562/4.4	2.83/0.529/651.5
平均误差/时耗	4.27/0.354/22.7	5.21/0.431/205.9	3.30/0.289/186.6	2.69/0.240/366.8	5.58/0.439/4.3	2.43/0.234/490.8

4.4 MPI-Sintel 数据集实验

为了验证本文方法在困难场景下光流估计的效果,本实验选取 MPI-Sintel 数据库中的 Alley_2, Market_6, Temple_2 和 Bamboo_2 序列作为测试序列,它们对应分别包含复杂场景、非刚性大位移运动以及运动遮挡等困难场景,图 4 分别展示了测试图像序列的参考帧和光流真实值及各对比算法光流估计结果。

从图中可以看到,LDOF 算法在场景的边缘区域存在大范围过度平滑;CLG-TV 算法结果中存在大量噪声,场景边缘保护较差;Classic + + 算法由于受中值滤

波的影响,结果中存在严重的过度平滑;由于 MPI-Sintel 数据库提供较充分的训练样本,FlowNet2.0 算法的估计效果明显有所提升,但该方法在 Market_6 和 Bamboo_2 序列效果较差.NNF 算法在 Market_6 和 Temple_2 序列光流估计运动边缘区域出现明显的错误.本文方法整体效果较好,尤其边缘区域取得了较准确的估计结果,说明本文方法针对复杂场景、非刚性大位移运动以及运动遮挡等困难场景同样具有较高的光流估计精度与鲁棒性。

表 3 分别列出了本文方法与对比算法的误差统计

和平均时间消耗. 从表中可以看到, 本文方法平均 AEE 误差达到最小, 平均 AAE 误差仅略大于 LDOF 算法, 说明本文方法具有更好的光流估计精度与鲁棒性. 在时

间消耗方面, 本文方法由于测试图像序列尺寸增加, 在金字塔每层图像上完整遍历整幅图像搜索相互结构区域的时间消耗增加, 进而增加了运行时间消耗.

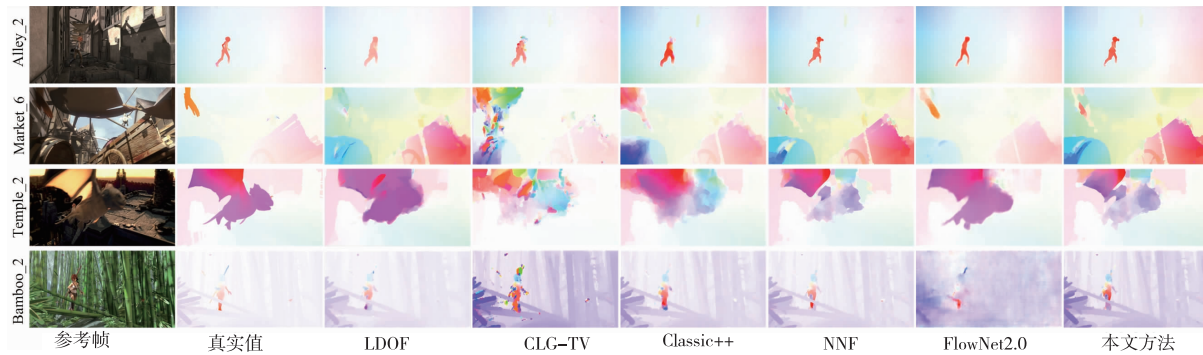


图4 MPI-Sintel数据库测试图像序列光流估计结果

表3 MPI-Sintel 数据库图像序列光流估计误差与时间消耗对比(时间单位:s)

对比方法	平均误差 AAE/AEE	平均耗时(s)	Alley_2	Market_6	Temple_2	Bamboo_2
			AAE/AEE	AAE/AEE	AAE/AEE	AAE/AEE
LDOF	6.99/4.545	40.2	2.23/0.122	9.10/4.586	11.21/13.180	5.43/0.293
CLG-TV	15.68/0.775	306.5	3.62/0.182	22.15/1.477	30.60/1.108	6.34/0.334
Classic + +	13.22/6.985	338.4	2.84/0.145	17.42/8.359	27.73/19.163	4.87/0.271
NNF	9.48/0.742	489.5	2.55/0.123	8.54/1.406	22.16/1.211	4.66/0.227
FlowNet2.0	11.61/1.456	5.0	3.28/0.166	11.14/2.204	15.24/2.773	16.79/0.682
本文方法	8.77/0.683	654.3	1.93/0.104	7.92/1.366	20.68/1.036	4.56/0.224

5 结论

本文提出一种基于相互结构引导滤波的 TV-L1 变分光流估计模型, 首先对图像中不同结构进行定义, 然后通过设计图像相互结构区域提取因子提取置信度较高的图像相互结构区域, 构造基于相互结构引导滤波的全局目标函数, 最后采用金字塔分层细化与交替迭代方案结合的策略进行优化. 实验结果表明本文方法具有较好的光流估计精度与鲁棒性, 尤其针对复杂场景、大位移运动以及运动遮挡等困难场景具有更好的边缘保护效果.

参考文献

- [1] Sanchez-Garcia A J, Rios-Figueroa H V, Marin-Hernandez A, et al. Decision making for obstacle avoidance in autonomous mobile robots by time to contact and optical flow [A]. Electronics, Communications and Computers [C]. Puebla; IEEE, 2015, 130 - 134.
- [2] 耿凤欢, 刘慧, 郭强, 等. 基于变分光流估计的肺部 4D-CT 图像超分辨率重建[J]. 计算机研究与发展, 2017, 54(8): 1703 - 1712.
Geng Fenghuan, Liu Hui, Guo Qiang, et al. Variational op-

tical flow estimation based super-resolution reconstruction for lung 4D-CT image [J]. Journal of Computer Research and Development. 2017, 54(8): 1703 - 1712. (in Chinese).

- [3] 张聪炫, 陈震, 黎明. 单目图像序列光流三维重建技术研究综述[J]. 电子学报, 2016, 44(12): 3044 - 3052.
ZHANG Cong-xuan, CHEN Zhen, LI Ming. Review of the 3D reconstruction technology based on optical flow of monocular image sequence [J]. Acta Electronica Sinica, 2016, 44(12): 3044 - 3052. (in Chinese).
- [4] Drulea M, Nedevschi S. Total variation regularization of local-global optical flow [A]. IEEE International Conference on Intelligent Transportation Systems [C]. Anchorage: IEEE, 2012. 318 - 323.
- [5] Zimmer H, Bruhn A, Weickert J. Optic flow in harmony [J]. International Journal of Computer Vision, 2011, 93(3): 368 - 388.
- [6] Brox T, Malik J. Large displacement optical flow; Descriptor matching in variational motion estimation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(3): 500 - 513.
- [7] 刘骏, 祖静, 张瑜, 等. 光照变化条件下的光流估计 [J]. 中国图象图形学报, 2014, 19(10): 1475 - 1480.

- Liu Jun, Zu Jing, Zhang Yu, et al. Optical flow estimation method under the condition of illumination change [J]. Journal of Image and Graphics, 2014, 19(10): 1475 – 1480. (in Chinese).
- [8] Xiao J, Cheng H, Sawhney H, et al. Bilateral filtering-based optical flow estimation with occlusion detection [A]. European Conference on Computer Vision [C]. Graz: Springer, 2006. 211 – 224.
- [9] Sun D, Roth S, Black M J. A quantitative analysis of current practices in optical flow estimation and the principles behind them [J]. International Journal of Computer Vision, 2014, 106(2): 115 – 137.
- [10] Shen X, Zhou C, Xu L, et al. Mutual-structure for joint filtering [J]. International Journal of Computer Vision, 2017, 125(1): 19 – 33.
- [11] Chen Z, Jin H, Lin Z, et al. Large displacement optical flow from nearest neighbor fields [A]. IEEE International Conference on Computer Vision and Pattern Recognition [C]. Portland: IEEE, 2013. 2443 – 2450.
- [12] Ilg E, Mayer N, Saikia T, et al. FlowNet 2.0: Evolution of optical flow estimation with deep networks [A]. IEEE International Conference on Computer Vision and Pattern Recognition [C]. Providence: IEEE, 2017. 2462 – 2470.

作者简介



葛利跃 男, 1992 年 10 月出生于安徽省蚌埠市. 现为南昌航空大学测试与光电工程学院硕士研究生. 主要研究方向为图像检测与智能识别.
E-mail: lygeah@163.com



张聪炫(通信作者) 男, 1984 年 7 月出生于河南省焦作市. 分别于 2007 年和 2014 年在南昌航空大学和南京航空航天大学获得学士和博士学位. 现为南昌航空大学副教授, 硕士生导师. 主要研究方向为图像检测与智能识别.
E-mail: zcxsg@163.com