

# 基于语义分割的双目场景流估计

陈震<sup>1</sup>, 马龙<sup>1</sup>, 张聪炫<sup>1,2</sup>, 黎明<sup>1</sup>, 吴俊劼<sup>1</sup>, 江少锋<sup>1</sup>

(1. 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; 2. 中国科学院自动化研究所, 北京 100190)

**摘要:** 针对现有场景流计算方法在复杂场景、大位移和运动遮挡等情况下易产生运动边缘模糊的问题, 提出一种基于语义分割的双目场景流估计方法. 首先, 根据图像中的语义信息类别, 通过深度学习的卷积神经网络模型将图像划分为带有语义标签的区域; 针对不同语义类别的图像区域分别进行运动建模, 利用语义知识计算光流信息并通过双目立体匹配的半全局匹配方法计算图像视差信息. 然后, 对输入图像进行超像素分割, 通过最小二乘法耦合光流和视差信息, 分别求解每个超像素块的运动参数. 最后, 在优化能量函数中添加语义分割边界的约束信息, 通过更新像素到超像素块的映射关系和超像素块到移动平面的映射关系得到最终的场景流估计结果. 采用 KITTI 2015 标准测试图像序列对本文方法和代表性的场景流计算方法进行对比分析. 实验结果表明, 本文方法具有较高的精度和鲁棒性, 尤其对于复杂场景、运动遮挡和运动边缘模糊的图像具有较好的边缘保护作用.

**关键词:** 语义分割; 场景流; 深度学习; 双目立体匹配; 最小二乘法; 超像素分割; 运动遮挡; 边缘保护

**中图分类号:** TP391      **文献标识码:** A      **文章编号:** 0372-2112 (2020)04-0631-06

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2020.04.002

## Binocular Scene Flow Estimation Based on Semantic Segmentation

CHEN Zhen<sup>1</sup>, MA Long<sup>1</sup>, ZHANG Cong-xuan<sup>1,2</sup>, LI Ming<sup>1</sup>, WU Jun-jie<sup>1</sup>, JIANG Shao-feng<sup>1</sup>

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

2. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** In order to address the issue of motion boundary blurring caused by the complex scenes, large displacement and motion occlusion, this paper proposes a binocular scene flow estimation method based on semantic segmentation. Firstly, by using the image semantic information, we classify the image regions into several categories with semantic labels through convolutional neural networks. Then we plan the motion models of various image regions according to the different semantic categories and compute the optical flow and disparity under the prior knowledge of semantic information. Secondly, we apply the superpixel segmentation to the input image and couple the optical flow and disparity information via least squares method to solve the motion parameters of each superpixel patch. Finally, we add the boundary information of semantic segmentation constraint to the optimization energy function, and estimate the scene flow by updating the mappings of pixels-to-superpixel and superpixel-to-plane. We evaluate the proposed approach and some state-of-the-art methods on the KITTI 2015 database to conduct a comparison experiment. The experimental results demonstrate that our method has high accuracy and good robustness, and especially has significant benefit of boundary preserving in the areas of complex scene, motion occlusion and motion boundary.

**Key words:** semantic segmentation; scene flow; deep learning; binocular stereo matching; least squares method; superpixel segmentation; motion occlusion; edge protection

收稿日期: 2018-12-07; 修回日期: 2019-05-13; 责任编辑: 李勇锋

基金项目: 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); 江西省青年科学基金 (No. 20171BAB212012); 中国博士后科学基金 (No. 2019M650894)

## 1 引言

场景流最初由 Vedula 等人<sup>[1]</sup>提出,根据光流将其定义为图像序列中像素点的三维运动矢量,从而建立了场景流计算基本模型.随后,针对场景流计算模型的优化方法不断涌现.针对复杂场景下物体运动的不连续性问题,Huguet 等人<sup>[2]</sup>首次提出基于变分理论的双目场景流模型,该模型通过多分辨率算法求解偏微分方程,能较好克服运动的不连续性问题.针对大位移运动场景流估计问题,Basha 等人<sup>[3]</sup>提出通过三维点云表示场景流,并采用金字塔分层策略,能够有效解决大位移运动下场景流估计问题.针对运动边缘模糊问题,Schuster 等人<sup>[4,5]</sup>提出联合光流与视差的场景流模型,通过稀疏匹配和边缘信息的稠密插值操作,能够有效保护运动边缘.Vogel 等人<sup>[6]</sup>提出分段刚性场景流模型,该模型将场景表示为有限数量的刚性移动平面,通过优化平面运动参数获取最终场景流结果.

现阶段,分段刚性场景流模型已经成为求解场景流的基本方法,但是该方法在复杂场景、大位移运动和运动遮挡等情况下易产生运动边缘模糊问题.本文提出一种基于语义分割的双目场景流计算方法,能够有效提高复杂场景、大位移运动和运动遮挡情况下场景流估计精度.

## 2 图像语义分割

语义分割是计算机视觉领域的研究热点之一.语义分割是根据语义对像素点进行分类,将图像中的场景分割为若干个目标区域,并最终得到语义标签图像,如图1所示.

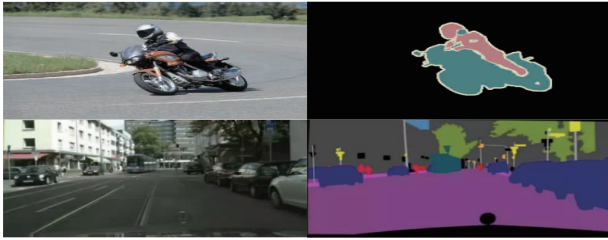


图1 语义分割示意图

卷积神经网络能够较好地实现图像分类,为增强语义分割结果中图像特征的表达性能,本文采用 Deeplab 语义分割模型<sup>[7]</sup>,结合深度卷积神经网络和条件随机场,采用多个卷积核实现多尺度图像语义分割. Deeplab 语义分割模型的输出为带有不同语义标签类别的图像区域,本文使用分层模型<sup>[8]</sup>将其分为物体、平面和其他三类.对三种不同的图像区域分别进行运动建模,通过融合算法<sup>[9]</sup>组合三种图像区域的运动得到最终的语义光流.

## 3 语义分割场景流估计模型

### 3.1 初始场景流运动模型

动态场景可表示为有限数量刚性移动的平面区域,通过超像素分割得到平面区域的运动边界,每一个超像素块表示一个刚性运动的3D平面  $\pi \equiv \pi(\mathbf{R}, \mathbf{t}, \bar{\mathbf{n}})$ ,旋转矩阵  $\mathbf{R}$  和平移向量  $\mathbf{t}$  表示平面的运动参数,法向量  $\bar{\mathbf{n}}$  表示平面的空间位置参数.

如图2所示,由于双目图像存在四个视图,为更高效地求解运动参数,本文将  $t_0$  时刻左目图像表示为参考图像,其他视图可通过参考图像转换得到.转换关系为:

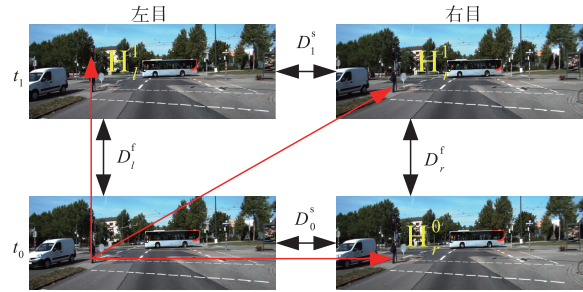


图2 双目视图转换关系

$$\begin{cases} \mathbf{H}_r^0 = (\mathbf{M} - \mathbf{m}\bar{\mathbf{n}}^T)\mathbf{K}^{-1} \\ \mathbf{H}_l^1 = \mathbf{K}(\mathbf{R} - \mathbf{t}\bar{\mathbf{n}}^T)\mathbf{K}^{-1} \\ \mathbf{H}_r^1 = (\mathbf{M}\mathbf{R} - (\mathbf{M}\mathbf{t} + \mathbf{m})\bar{\mathbf{n}}^T)\mathbf{K}^{-1} \end{cases} \quad (1)$$

其中  $\mathbf{H}_r^0$ 、 $\mathbf{H}_l^1$ 、 $\mathbf{H}_r^1$  分别表示参考图像到右目第一帧图像、左目第二帧图像、右目第二帧图像的单应性转换矩阵,符号  $\mathbf{M}$ 、 $\mathbf{m}$  和  $\mathbf{K}$  分别表示摄像机投影矩阵与标定参数矩阵.本文通过最小二乘法耦合语义光流和视差信息,对每个超像素块的深度和运动进行推理.首先通过最小二乘法求得初始解,然后通过洛伦兹惩罚函数  $\varphi(x) = \log(1 + x/2\sigma^2)$  对误差进行优化,公式如下:

$$\sum_{p \in \varepsilon} \varphi(\| \mathbf{T}(\mathbf{H}_r^0 p) - p' \|^2) \rightarrow \min_{\mathbf{n}} \quad (2)$$

$$\sum_{p \in \varepsilon} \varphi(\| \mathbf{T}(\mathbf{H}_l^1 p) - p' \|^2) \rightarrow \min_{\mathbf{R}, \mathbf{t}} \quad (3)$$

其中,超像素块  $\varepsilon$  中像素  $p$  与转换视图中像素  $p'$  相对应,  $\mathbf{T}$  代表投影算子.

### 3.2 优化能量函数模型

为优化初始场景流运动参数,首先定义两种映射关系  $S$  和  $P$ .  $S$  表示像素到超像素块的映射,  $P$  表示超像素块到移动平面的映射.定义场景流优化能量函数如下:

$$E(P, S) = E_D(P, S) + \lambda E_R(P, S) + \mu E_S(S) \quad (4)$$

其中,  $E_D(P, S)$  为数据项,表示为:

$$E_D(P, S) = D_0^s + D_1^s + D_l^f + D_r^f \quad (5)$$

式(5)中,  $D_0^s$ 、 $D_1^s$  分别表示当前帧和下一帧的立体匹配

项,  $D_l^f, D_r^f$  分别表示左右目的光流数据项:

$$\begin{cases} D_0^s = \sum_{p \in I_l^f} \rho(p, \mathbf{H}_r^0 p), D_1^s = \sum_{p \in I_l^f} \rho(\mathbf{H}_l^1 p, \mathbf{H}_r^1 p) \\ D_l^f = \sum_{p \in I_l^f} \rho(p, \mathbf{H}_l^1 p), D_r^f = \sum_{p \in I_l^f} \rho(\mathbf{H}_r^0 p, \mathbf{H}_r^1 p) \end{cases} \quad (6)$$

数据项通过汉明距离来评估不同视图的外观一致性, 汉明距离越小, 两个视图的外观相似程度则越高.

$E_R(P, S)$  为正则化项, 表示为:

$$E_R(P, S) = E_R^1(P, S) + E_R^2(P, S) \quad (7)$$

式(7)中,  $E_R^1(P, S)$  为几何正则化项, 表示为:

$$\begin{aligned} E_R^1(P, S) &= \sum_{(p, q) \in N} \omega_{p, q} \psi \left( 3 \int_0^1 \int_{-1}^1 f_\gamma(\alpha, \beta)^2 d\beta d\alpha \right) \\ &= \sum_{(p, q) \in N} \omega_{p, q} \psi \left( \| \mathbf{d}_1 \|^2 + \| \mathbf{d}_2 \|^2 \right. \\ &\quad \left. + \langle \mathbf{d}_1, \mathbf{d}_2 \rangle + \gamma^2 \| \mathbf{d}_n \|^2 \right) \end{aligned} \quad (8)$$

式(8)中,  $\omega_{p, q}$  为边界长度系数,  $\psi(\cdot)$  表示惩罚函数. 像素  $p$  和  $q$  共享相同边界, 边界端点  $c^1$  和  $c^2$  投影至三维平面得到  $c_p^1, c_p^2, c_q^1, c_q^2$  四个三维端点, 端点距离差分别为  $\mathbf{d}_1 = c_p^1 - c_q^1$  和  $\mathbf{d}_2 = c_p^2 - c_q^2$ , 通过  $\| \alpha \mathbf{d}_1 + (1 - \alpha) \mathbf{d}_2 \|^2$  表示边界在三维空间的距离. 由于空间表面存在曲率, 需考虑三维端点沿法向量方向移动后的距离. 两个三维平面的法向量分别为  $\mathbf{n}_p$  和  $\mathbf{n}_q$ , 法向量差为  $\mathbf{d}_n = \mathbf{n}_p - \mathbf{n}_q$ , 因此距离函数定义为:

$$f_\gamma(\alpha, \beta) = \| \alpha(\mathbf{d}_1 + \gamma \beta \mathbf{d}_n) + (1 - \alpha)(\mathbf{d}_2 + \gamma \beta \mathbf{d}_n) \|^2 \quad (9)$$

式(9)中,  $\alpha$  为边界距离系数,  $\beta$  为法向量系数,  $\gamma$  为调节曲率和边界距离的权重系数.

$E_R^2(P, S)$  为运动正则化项. 其中,  $c_p^1, c_p^2, c_q^1, c_q^2$  分别更新为  $\mathbf{R}_p c_p^1 + \mathbf{t}_p, \mathbf{R}_p c_p^2 + \mathbf{t}_p, \mathbf{R}_q c_q^1 + \mathbf{t}_q, \mathbf{R}_q c_q^2 + \mathbf{t}_q, \mathbf{n}_p$  和  $\mathbf{n}_q$  分别更新为  $\mathbf{R}_p \mathbf{n}_p - \mathbf{n}_p$  和  $\mathbf{R}_q \mathbf{n}_q - \mathbf{n}_q$ . 三维端点的距离差为  $\mathbf{d}_1^m = \mathbf{R}_p c_p^1 + \mathbf{t}_p - \mathbf{R}_q c_q^1 - \mathbf{t}_q$  和  $\mathbf{d}_2^m = \mathbf{R}_p c_p^2 + \mathbf{t}_p - \mathbf{R}_q c_q^2 - \mathbf{t}_q$ , 法向量之差为  $\mathbf{d}_n^m = (\mathbf{R}_p \mathbf{n}_p - \mathbf{n}_p) - (\mathbf{R}_q \mathbf{n}_q - \mathbf{n}_q)$ , 则运动正则化项表示为:

$$E_R^2(P, S) = \sum_{(p, q) \in N} \omega_{p, q} \psi \left( \| \mathbf{d}_1^m \|^2 + \| \mathbf{d}_2^m \|^2 + \langle \mathbf{d}_1^m, \mathbf{d}_2^m \rangle \right. \\ \left. + \gamma^2 \| \mathbf{d}_n^m \|^2 \right) \quad (10)$$

$E_S(S)$  为语义分割项, 表示为:

$$E_S(S) = E_S^1(S) + E_S^2(S) \quad (11)$$

$E_S^1(S)$  为语义项, 像素  $p$  和像素  $q$  为语义标签内两个相邻的像素. 为了避免过度分割并在空间上保持运动一致性, 对语义标签内的像素进行优化, 因此语义项表示为:

$$E_S^1(S) = \sum_{\substack{(p, q) \in S_{sc}, S_{sc} \in I_l^f \\ S(p) \neq S(q)}} \exp \left( \frac{-\delta |I_l^0(p) - I_l^0(q)|}{\sigma_I(p, q) + \varepsilon} \right) \quad (12)$$

式(12)中  $S_{sc}$  表示图像中的语义标签区域,  $I_l^0(p)$  和  $I_l^0(q)$  分别表示像素  $p$  和像素  $q$  的灰度值,  $\delta$  表示灰度梯度系数,  $\varepsilon$  表示标准差调节范围.

$E_S^2(S)$  为限制项, 避免在映射关系  $S$  更新过程中超像素块尺寸过大导致场景过于简化. 限制项表示为:

$$E_S^2(S) = \sum_{p \in I_l^f} \begin{cases} 0, & \exists e \in \varepsilon: \| e - p \|_\infty < N_s \\ \infty, & \text{else} \end{cases} \quad (13)$$

式(13)中,  $e$  表示超像素块的中心点像素,  $\varepsilon$  表示超像素块区域,  $N_s$  表示限制超像素块大小的尺寸.

### 3.3 计算步骤

Step1: 通过将图像划分为带有语义标签的区域.

Step2: 以语义标签作为先验知识计算光流信息.

Step3: 采用半全局匹配算法计算视差信息.

Step4: 通过最小二乘法耦合语义光流和视差信息求得初始运动参数, 可得到映射关系  $P$ .

Step5: 固定映射  $P$ , 更新映射  $S$ .

Step6: 固定映射  $S$ , 更新映射  $P$ .

经过能量函数优化后, 两种映射关系  $S$  和  $P$  为最优, 即求得更新后每个超像素块的运动参数.

## 4 实验与分析

### 4.1 评估指标

本文采用异常值百分比评估场景流:

$$SF - all = \frac{P_1 \& P_2}{ALL} \times 100\% \quad (14)$$

$$D1 - all = \frac{P_2}{ALL} \times 100\% \quad (15)$$

其中,  $SF - all$  和  $D1 - all$  分别表示场景流和视差异常值百分比, 式中  $P_1$  表示光流端点误差大于 3 个像素的像素个数,  $P_2$  表示视差端点误差大于 3 个像素的像素个数,  $ALL$  表示像素总个数.

### 4.2 对比方法

为验证本文方法的精度以及对运动边缘的保护作用, 选取 VSF<sup>[2]</sup>、SGM&FlowFie<sup>+</sup><sup>[5]</sup>、PCOF-LDOF<sup>[10]</sup>、CSF<sup>[11]</sup>、FSF + MS<sup>[12]</sup> 等方法与本文方法进行定性分析和量化对比. 其中 VSF 算法是变分框架场景流估计的经典方法; SGM&FlowFie<sup>+</sup> 算法使用双线性插值的策略进行估计; PCOF-LDOF 算法首先预测自身运动, 然后通过校正步骤来改进结果; CSF 算法通过最小二乘法解决非线性问题实现更好的优化; FSF + MS 算法通过判断与相机运动不一致的区域估计场景流.

### 4.3 参数设置

本文设置参数主要为语义分割项系数  $\mu$ . 本文以 000042 序列和 000169 序列为例, 设定语义分割项系数为 0.01, 0.05, 0.2, 0.5, 0.7, 0.9, 对测试图像误差结果进行定性分析. 由图 3 看出当参数取 0.7 时误差最小, 因此本文选取  $\mu$  为 0.7.

本文根据文献介绍设置各对比方法的关键参数如

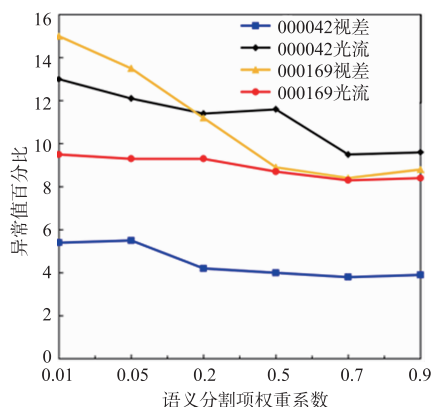


图3 语义分割项权重系数对场景流估计的影响

下:VSF 算法中平滑项权重  $\alpha_s = 0.5$ , 视差范围  $\lambda_d = 0.9$ ; SGM&FlowFie+ 算法中数据项权重  $\gamma_r = 0.77$ , 平滑项权重  $\lambda_s = 10$ ; PCOF-LDOF 算法中分辨率尺度系数  $J = 5$ , 迭代次数  $K = 4$ . CSF 算法中标准偏差  $\sigma_s = 1$ , 视差最大距离  $\lambda_{\text{census}} = 3$ ; FSF + MS 算法中掩膜参数  $\lambda_{\text{mask}} = 2$ , 截断阈值  $T_{\text{ncc}} = 0.25$ .

#### 4.4 KITTI 数据集实验

本文采用 KITTI 数据库提供的双目可见光图像测试序列验证本文方法场景流估计效果与鲁棒性. 图4分别

展示了 KITTI 数据集 000005、000010、000012 和 000013 四组双目图像序列, 其中 000005 序列存在复杂边缘结构的车辆信息, 000010 序列场景中存在大位移运动, 000012 序列和 000013 序列存在遮挡问题和运动不连续.

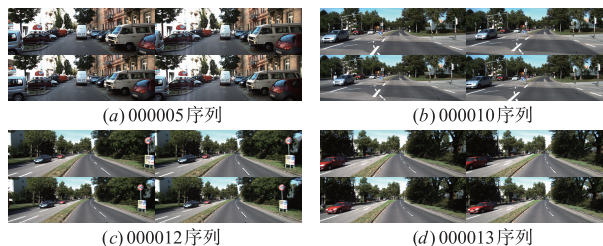


图4 KITTI测试集图像序列

图5展示了本文方法与 VSF、SGM&FlowFie+、CSF、FSF + MS、PCOF-LDOF 等对比算法的场景流误差图, 红色表示误差较大区域. 由图4可以看出本文方法红色误差区域最少, 场景流估计效果最好. 为验证本文方法综合性能, 表1列出了本文方法和各对比算法在 200 组 KITTI 测试图像的误差结果和时间消耗情况, 其中 bg 和 fg 分别对应背景误差值和前景误差值. 由表1可以看出, 本文方法各项误差均最小, 说明本文方法具有较好的估计精度.

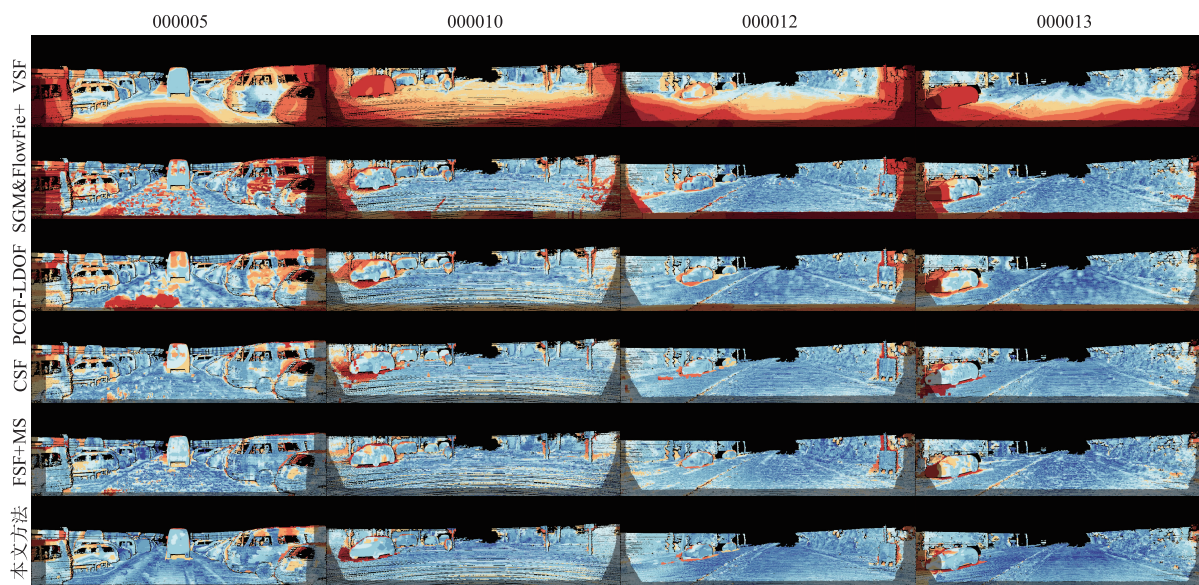


图5 不同对比方法场景流计算误差图

表1 场景流估计误差和时间消耗对比

对比方法	SF-bg	SF-fg	SF-all	时间(s)
VSF	67.69	62.93	66.90	7500
SGM&FlowFie+	32.26	40.12	33.57	29
PCOF-LDOF	25.26	49.39	29.27	50
CSF	12.21	33.21	15.71	80
FSF + MS	11.17	33.91	14.96	2.7
本文方法	9.64	32.88	13.51	180

在时间消耗方面, FSF + MS 算法时间消耗最少, 源于该算法采用物体与相机运动相融合的统一框架计算策略. VSF 算法时间消耗最大, 主要因为该方法使用的变分计算框架优化需要进行大量迭代操作. 本文方法时间消耗主要源于语义标签生成、语义光流计算、初始运动参数计算以及优化迭代等过程.

为更好展示本文方法在图像边缘的保护作用, 图6分别展示了本文方法与各对比算法的视差估计结果.

图中黑框区域为包含复杂结构信息和遮挡的区域,如车辆、信号灯等处.从图中可以看出,在车辆边缘、信号

灯边缘和标志牌轮廓细节上,本文方法通过语义分割能提供更准确的分割信息.

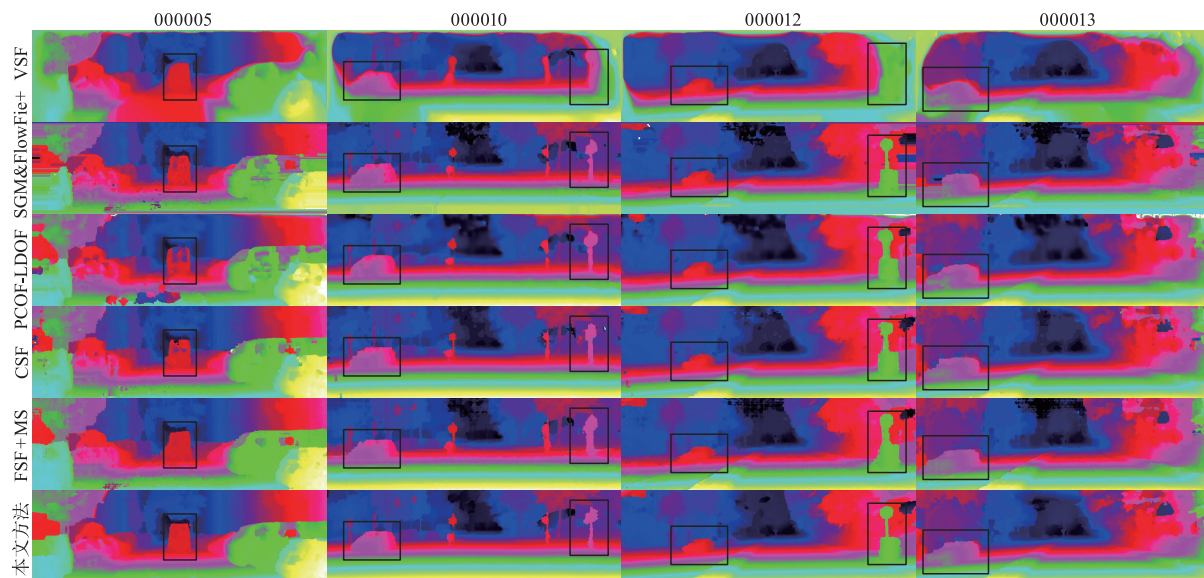


图6 不同对比方法视差结果图

为了进一步验证本文方法的估计精度,表 2 分别列出了本文方法与各对比算法的视差误差结果.从表中可以看出,本文方法各项误差均最小,说明本文方法在复杂场景、大位移、遮挡等情况下,针对立体匹配估计同样具有较高的精度和鲁棒性.

表 2 视差估计误差

方法	D1 - bg	D1 - fg	D1 - all
VSF	27.31	21.72	26.38
SGM&FlowFie +	11.93	20.57	13.37
PCOF-LDOF	6.31	19.24	8.46
CSF	4.57	13.04	5.98
FSF + MS	5.72	11.84	6.74
本文方法	3.59	13.11	5.18

## 5 结论

本文提出一种基于语义分割的双目场景流估计模型.首先对场景进行语义分割,根据语义标签进行运动建模,然后通过最小二乘法耦合光流和视差信息获取初始运动参数,最后添加语义分割约束条件优化运动参数得场景流结果.本文通过获取场景语义信息优化图像分割,针对复杂场景和运动遮挡具有很好的边缘保护作用.

### 参考文献

[1] Vedula S, Baker S, Collins R, et al. Three-dimensional scene flow[A]. International Conference on Computer Vi-

sion[C]. Kerkyra: IEEE, 1999. 722 - 729.

- [2] Huguet F, Devernay F. A variational method for scene flow estimation from stereo sequences[A]. IEEE International Conference on Computer Vision[C]. Rio de Janeiro: IEEE, 2007. 1 - 7.
- [3] Basha T, Moses Y, Kiryati N. Multi-view scene flow estimation: a view centered variational approach[J]. International Journal of Computer Vision, 2013, 101(1): 6 - 21.
- [4] Schuster R, Wasenmuller O, Kusch G, et al. Scene flow fields: dense interpolation of sparse scene flow correspondences[A]. IEEE Winter Conference on Applications of Computer Vision[C]. Nevada: IEEE, 2018. 1 - 10.
- [5] Schuster R, Bailer C, Wasenmuller O, et al. Combining stereo disparity and optical flow for basic scene flow[A]. International Commercial Vehicle Technology Symposium[C]. Kaiserslautern: Springer, 2018. 90 - 101.
- [6] Vogel C, Schindler K, Roth S. Piecewise rigid scene flow[A]. IEEE International Conference on Computer Vision[C]. Sydney: IEEE, 2013. 1377 - 1384.
- [7] Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 40(4): 834 - 848.
- [8] Sevilla-Lara L, Sun D, Jampani V, et al. Optical flow with semantic segmentation and localized layers[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Las Vegas: IEEE, 2016. 3889 - 3898.
- [9] Lempitsky V, Roth S, Rother C. Fusionflow: discrete-con-

- tinuous optimization for optical flow estimation[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Anchorage:IEEE,2008. 1 – 8.
- [10] Derome M, Plyer A, Sanfourche M, et al. A prediction-correction approach for real-time optical flow computation using stereo[A]. German Conference on Pattern Recognition[C]. Hannover:Springer,2016. 1 – 7.
- [11] Lü Z, Beall C, Alcantarilla P F, et al. A continuous optimization approach for efficient and accurate scene flow [A]. European Conference on Computer Vision[C]. Amsterdam:Springer,2016. 757 – 773.
- [12] Taniai T, Sinha S N, Sato Y. Fast multi-frame stereo scene flow with motion segmentation[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Hawaii: IEEE,2017. 729 – 739.

### 作者简介



**陈震** 男,1969年11月生,江西九江人. 分别于1993、2000和2003年在西北工业大学获得学士、硕士和博士学位. 现为南昌航空大学教授,博士生导师,主要研究方向为计算机视觉、图像处理与模式识别.

E-mail: dr\_chenzhen@163. com



**马龙** 男,1993年9月出生于河南省鹤壁市. 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:1007516637@qq. com



**张聪炫(通讯作者)** 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,中国科学院自动化研究所博士后. 主要研究方向为图像检测与智能识别.

E-mail: zcxdsg@163. com