

# 基于分层置信度传播的光流估计方法<sup>①</sup>

张子星, 文颖

(华东师范大学 计算机科学技术系, 上海 200062)

通讯作者: 文颖, E-mail: [ywen@cs.ecnu.edu.cn](mailto:ywen@cs.ecnu.edu.cn)

**摘要:** 置信度传播算法作为一种有效的寻找图像间对应点的方法, 近年来被广泛应用于光流估计. 但是在估计大位移高精度光流时, 将置信度传播直接应用于原图像会导致标签空间过大和处理时间过长的问題. 为了克服这个缺点, 我们提出了一种基于分层置信度传播的算法来估计高精度大位移光流. 本文方法将输入图像视作马尔科夫随机场, 为了提高效率, 在超像素和像素两个层面上执行置信度传播. 我们将超像素层得到的基础位移结果作为粗略的位移参考值, 可以有效地减小像素层置信度传播的标签空间, 并在有限的标签空间内得到高精度的光流估计结果. MPI Sintel 光流数据集上的实验结果显示本文提出的方法在精度和速度上都取得了较好的结果.

**关键词:** 光流; 大位移; 马尔科夫随机场; 置信度传播; 超像素

引用格式: 张子星, 文颖. 基于分层置信度传播的光流估计方法. 计算机系统应用, 2018, 27(9): 25-32. <http://www.c-s-a.org.cn/1003-3254/6518.html>

## Hierarchical Belief Propagation for Optical Flow Estimation

ZHANG Zi-Xing, WEN Ying

(Department of Computer Science and Technology, East China Normal University, Shanghai 200062, China)

**Abstract:** As an effective way to find correspondences between images, Belief Propagation (BP) is widely used for estimating optical flow in recent years. Nevertheless, its application to directly estimating high-accuracy large displacement optical flow needs huge label space and long time to process. In order to overcome the drawback of BP, we propose a Hierarchical Belief Propagation (HBP) algorithm to estimate high-accuracy large displacement optical flow. We treat input images as Markov Random Fields (MRFs). To accelerate computation, we perform BP on hierarchical MRFs, i.e., superpixel MRF and pixel MRF. The basic displacements obtained on the superpixel MRF are used as a coarse reference to constrain label space to a smaller size on the pixel MRF. Based on this constrained label space, we can estimate accurate optical flow efficiently. Experiments on MPI Sintel dataset show that the proposed method is competitive on speed and accuracy.

**Key words:** optical flow; large displacements; Markov Random Fields (MRFs); Belief Propagation (BP); superpixel

### 1 概述

光流估计是计算机视觉的一个基础研究领域, 其计算结果被广泛用于物体分割、跟踪、运动场景分类等许多领域<sup>[1]</sup>. 自从 Horn 和 Schunck 的工作<sup>[2]</sup>以来, 光流估计领域的研究已经持续了几十年. 传统的方法<sup>[3-5]</sup>

通过变分法框架最小化能量函数来求解光流场, 这在小位移场景下能获得较精确的结果, 但往往难以解决大位移光流问题. 尽管 Brox<sup>[3]</sup>采用由粗到细 (coarse-to-fine) 的策略将高分辨率下的大位移转化为低分辨率下的小位移来计算, 但对于位移比物体本身尺寸更大的

① 基金项目: 国家自然科学基金 (61773166); 上海市自然科学基金 (17ZR1408200)

Foundation item: National Natural Science Foundation of China (61773166); Natural Science Foundation of Shanghai Municipality (17ZR1408200)

收稿时间: 2018-01-15; 修改时间: 2018-02-09; 采用时间: 2018-02-26; csa 在线出版时间: 2018-07-26

情况仍然难以适应<sup>[6]</sup>。

为了克服变分法框架处理大位移时的缺陷, Brox 等人在变分法模型中集成了稀疏的描述子匹配的结果<sup>[6,7]</sup>, 用描述子匹配得到的运动场拉动变分法光流的计算. 通过将匹配结果与变分法相融合, 可以有效地提高变分法光流的精度<sup>[8]</sup>. Deepflow<sup>[9]</sup>通过非刚性匹配得到半稠密的运动场, 进一步提高了描述子对变分法的拉动能力. 文献<sup>[10]</sup>将图像进行分割并分别求取变换矩阵, 在单一运动场景和复杂运动场景下都取得了较好的结果, 近年来出现的基于 PatchMatch 的方法<sup>[11-13]</sup>通过随机搜索和邻域块之间的信息传播<sup>[14]</sup>加快了寻找匹配对的过程, 有效提高了匹配精度.

近年来, 马尔科夫随机场也被许多学者用来尝试作为寻找图像间匹配点的模型. 求解马尔科夫随机场的推断算法如置信度传播在实际应用中能得到较精确的匹配结果, 但是存在着计算量太大的问题. Felzenszwalb 等人提出了一种高效的循环置信度传播算法<sup>[15]</sup>, 在有限次的消息传递迭代以后就能获得较好的结果, 但该方法适用于小位移场景, 在大位移场景下仍然过于耗时. SPM-BP 方法<sup>[16]</sup>将消息传递算法与 PatchMatch 的随机搜索相结合, 在减小标签空间的同时获得了亚像素精度的位移场. FullFlow<sup>[17]</sup>提出了一种全局的消息传递优化方法并将消息传递的计算复杂度从二次降为线

性, 取得了较为精确的匹配结果. 将置信度传播应用于光流估计的关键在于在标签空间和位移场精度之间取得平衡. 高精度的大位移估计意味着较大的标签空间, 这会导致计算时间及内存上的巨大开销. FullFlow 虽然提出了改进方法, 但在大位移场景下的精确估计仍然比较耗时, SPM-BP 虽然利用随机搜索有效缩短了时间, 但是精度不稳定, 错误的标签空间选取可能导致严重的误差.

综上所述, 以往提出的方法存在的不足之处主要表现在光流计算精度有限和处理速度缓慢的问题, 为了解决上述问题, 我们提出了分层置信度传播算法用以计算精确光流. 图 1 展示了我们模型的主要流程. 我们构建了两层马尔科夫随机场: 超像素层与像素层. 我们首先在超像素层执行置信度传播得到一个大范围但低精度的位移场, 然后利用该位移场限定每个像素点可能的位移范围, 从而有效减小像素层置信度传播的标签空间, 然后在像素层执行置信度传播得到每个点的高精度位移. 通过分层的方式, 我们有效地提高了利用置信度传播计算光流的效率. 为了提高算法的鲁棒性, 我们还在置信度传播模型中加入了相邻多帧的图像信息作为参考, 对于存在严重遮挡的运动场景, 多帧信息能有效减少误匹配的发生.

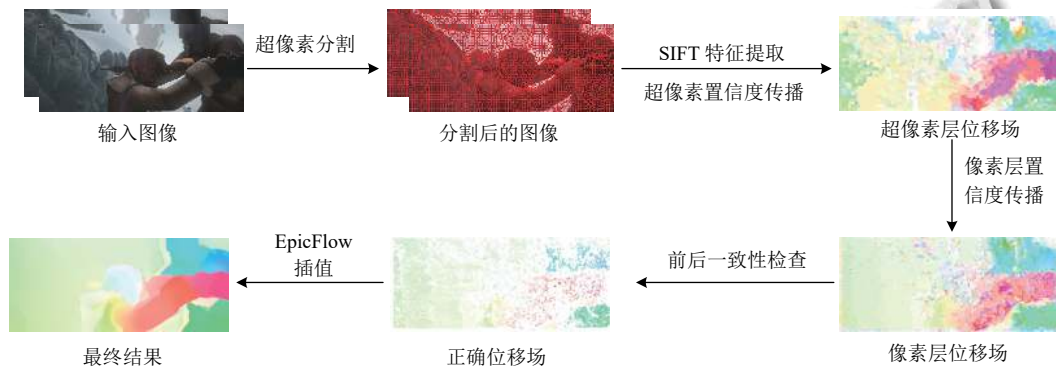


图 1 本文提出的分层置信度传播光流估计算法的流程示意图

## 2 相关工作

应用于计算机视觉的置信度传播算法的一般框架是将图像中的每一个像素点作为马尔科夫随机场中的一个节点, 并用边连接四邻接的像素点. 对每个节点而言, 每一种可能的取值都是一个标签. 对光流估计问题来说, 每一个标签都代表一种可能的位移. 目标函数一般定义为:

$$E(f) = \sum_{p \in \mathcal{P}} C_p(f_p) + \sum_{(p,q) \in \mathcal{N}} S(f_p, f_q) \quad (1)$$

其中,  $\mathcal{P}$  是节点构成的集合,  $f_p$  是赋给节点  $p$  的标签, 在光流估计问题中, 每个候选标签代表一种候选位移.  $\mathcal{N}$  代表四邻接节点间互相连接的边.  $C_p(f_p)$  代表数据项, 即节点  $p$  在标签  $f_p$  下的代价,  $S(f_p, f_q)$  代表平滑项, 用以衡量两个相邻节点  $p$  和  $q$  的标签差异.

根据 Felzenszwalb 等人的工作<sup>[15]</sup>, 在带环的马尔科夫随机场中最小化能量函数 (1) 可以通过循环置信度传播来实现. 循环置信度传播通过消息迭代传递来更新每个节点的状态, 消息迭代的公式如下:

$$m_{pq}^t(f_q) = \min_{f_p} (S(f_p, f_q) + C_p(f_p) + \sum_{s \in N_p \setminus q} m_{sp}^{t-1}(f_p)) \quad (2)$$

其中,  $m_{sp}^{t-1}(f_p)$  表示在  $t-1$  次迭代时节点  $s$  向节点  $p$  传递的消息.  $s$  是  $p$  的邻接节点中除了  $q$  以外的节点. 消息  $m_{pq}^t(f_q)$  是一个向量, 维度与节点的候选标签数量相同. 经过若干次迭代以后, 通过如下公式计算置信度向量:

$$b_p(f_p) = C_p(f_p) + \sum_{q \in N_p} m_{qp}(f_p) \quad (3)$$

并选择向量  $b_p$  中的最小值对应的标签作为节点  $p$  最后的标签.

当处理存在大位移场景的高分辨率图片时, 过大的节点数量和标签空间使得上述框架计算极为耗时. 上文提到的 FullFlow 虽然可以通过置信度传播取得较为精确的结果, 但是其计算极为消耗资源, 而 SPM-BP 利用随机选取来获取候选标签的方法并不稳定, 错误的标签可能造成较大的误差. 本文提出的方法不但能有效提升置信度传播的效率, 而且可以获得较高精度的结果, 在时间消耗和光流估计的精度上取得了较好的平衡.

### 3 本文方法

定义输入的两帧图像是  $I_1$ 、 $I_2$ , 将  $I_1$  视作马尔科夫随机场.  $\mathcal{L}$  代表候选标签的集合, 每个标签都是一个候选的二维位移向量. 马尔科夫随机场中的每个节点需要赋予一个合适的标签.

定义  $f_p \in \mathcal{L}$  为赋给节点  $p$  的标签,  $f_p = (u_p, v_p)$  代表通过位移  $(u_p, v_p)$  将  $I_1$  中的节点  $p(x, y)$  映射到  $I_2$  中的  $p'(x + u_p, y + v_p)$  处. 本文将能量函数定义为:

$$E(f) = \sum_{p \in \mathcal{P}} (\alpha C_{\text{color}}(f_p) + \beta C_{\text{desc}}(f_p)) + \sum_{(p, q) \in \mathcal{N}} S(f_p, f_q) \quad (4)$$

其中,  $p$  是  $I_1$  中的一个节点,  $(p, q)$  是两个相邻节点,  $C_{\text{color}}$  和  $C_{\text{desc}}$  分别代表移动前后的颜色差异、描述子差异. 平滑项  $S(f_p, f_q)$  用于惩罚相邻节点  $p$  与  $q$  运动的不一致性. 本文模型中的数据项由  $C_{\text{color}}$  和  $C_{\text{desc}}$  构成, 即:

$$C_p(f_p) = \alpha C_{\text{color}}(f_p) + \beta C_{\text{desc}}(f_p) \quad (5)$$

我们利用公式 (2) 和公式 (3) 来最小化能量函数 (4), 为了克服置信度传播在大位移场景下巨大计算量, 我们的模型引入了分层结构.

#### 3.1 分层结构

基于上述模型, 较直观的方法是将  $I_1$  中的每个像素点作为一个节点. 但是当进行大位移光流估计时, 直接应用该模型会导致  $\mathcal{L}$  包含成千上万的标签. 较大的节点数量和过大的标签空间使得计算变得过于耗时. 为了克服这个问题, 我们提出了分层结构来提高计算效率.

本文对于置信度传播过程的改进基于这样的假设: 一个物体表面局部区域内的像素点的运动往往是一致的, 因此可以将每一小块这样的区域作为一个节点构建马尔科夫随机场. 如果能得到每块区域的大致位移, 区域内的每个像素点在进行置信度传播时就能将候选位移控制在较小的范围内. 为了得到这些小块的区域, 我们将  $I_1$ 、 $I_2$  进行分割, 利用 SLIC 分割算法<sup>[18]</sup> 将原图分割成若干超像素块. 相较于规则分割, 超像素分割能更好地贴合物体的边缘, 保留边缘细节, 有利于保持运动场中物体边缘的清晰. 我们在超像素图像和原图像上分别构建马尔科夫随机场, 在超像素层和像素层两个层面执行循环置信度传播. 节点  $p$  在超像素层和像素层分别代表一个超像素块和一个像素点. 我们首先在超像素层执行置信度传播, 为了进一步加快计算, 本文方法在超像素层面获取的是一个低精度的位移场, 也就是说限定每个节点的候选位移步长为  $d$ , 即  $f_p \in R^2$ ,  $R = \{0, d, -d, 2d, -2d, \dots\}$ . 本文利用二维的 distance transform 算法<sup>[15]</sup> 来迭代地更新消息. 为了计算公式 (5), 需要确定超像素块  $p$  在经过位移  $f_p$  后在  $I_2$  上的对应块, 为此, 将超像素  $p$  的中心点  $x$  沿着  $f_p$  移动到  $I_2$  上的  $x'$  处, 将  $x'$  所属的超像素块  $p'$  作为  $p$  在  $f_p$  下的对应超像素块, 然后计算  $p$  与  $p'$  之间的颜色与描述子差异.

在完成超像素层的循环置信度传播后, 得到了一个基础的位移场. 利用这个基础结果来引导像素层的置信度传播, 能够有效地减少每个像素点的标签数量. 用  $(u_p, v_p)$  代表像素层节点  $p$  的一个候选位移, 注意如前文所述, 节点  $p$  在像素层代表一个像素点,  $(u_{sp}, v_{sp})$  代表节点  $p$  所在的超像素块在超像素层置信度传播中得到的位移, 则  $u_p = du + u_{sp}$ ,  $v_p = dv + v_{sp}$ , 其中  $du \in [-\delta, \delta]$ ,  $dv \in [-\delta, \delta]$ ,  $\delta$  决定了像素层每个节点的候选标签个数, 本文在像素层设置步长为 1, 则每个节点的标签数量为  $(2\delta + 1)^2$ . 像素层置信度传播是在超像素位移场基础上的进一步微调. 本文利用 min-convolution 算法<sup>[17]</sup> 来执

行像素层的消息迭代更新. 由于相邻的节点可能属于不同的超像素块, 因而它们可能拥有不同的基础位移, 需要对 min-convolution 的方法作修改以解决这个问题. min-convolution 的消息更新的过程被分解为两个求下包络的过程:

$$m_{pq}(u_q, v_q) = \min_{v_p} m_{pq|v_p}(u_q) + \rho(v_q - v_p) \quad (6)$$

$$m_{pq|v_p}(u_q) = \min_{u_p} \phi_{pq}(u_p, v_p) + \rho(u_q - u_p) \quad (7)$$

其中  $\rho(\cdot)$  是惩罚函数, 本文使用  $L_1$  范数.  $p$  和  $q$  是两个邻接节点,  $\phi_{pq}(u_p, v_p) = C_p(f_p) + \sum_{s \in N_p \setminus q} m_{sp}^{-1}(f_p)$ . 首先根据  $v_p$  的每一种可能的取值计算公式 (7), 然后利用 (7) 的结果根据  $v_q$  的每一种取值计算公式 (6). 若  $p$  与  $q$  拥有相同的基础位移, 则上述计算过程就是计算下包络的过程. 当  $p$  与  $q$  的基础位移不同时, 我们将上述求得的下包络函数图像向外延伸以覆盖  $p$  与  $q$  标签的取值范围. 图 2 显示了  $u_p$  与  $u_q$  的基础位移存在 1 个像素的相对偏移时的情况, 对于公式 (7), 当  $v_p$  值固定, 先假设  $u_p$  与  $u_q$  拥有相同的基础位移, 计算函数的下包络, 即图中加粗黑线部分, 然后再将下包络向外延伸 1 个单位以覆盖  $u_q$  的取值范围.

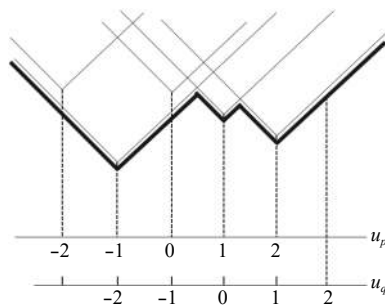


图 2 min-convolution 算法求取下包络

### 3.2 数据项定义

根据公式 (5), 数据项  $C_p(f_p)$  包含了图像的颜色信息和描述子信息. 本文在 CIELab 颜色空间中计算节点  $p$  在位移  $f_p$  下的颜色差异. 在超像素层, 节点  $p$  的颜色通过计算超像素块内的所有像素点的均值得到. 本文利用稠密 SIFT 描述子<sup>[19]</sup>来计算超像素块的  $C_{desc}$ . 首先在图像上提取稠密 SIFT 特征, 对于每个像素点, 我们将其  $8 \times 8$  的邻域空间分成  $4 \times 4$  的小块, 然后在每个小块内计算 8 个方向的梯度方向直方图, 最后所有小块的直方图合并成一个 128 维的向量. 超像素块的特征值由块内的所有像素点的特征值取均值得到, 我们利用欧

式距离计算移动前后对应的超像素块的特征值差异.

在像素层, 每个像素点都对应 MRF 中的一个节点, 由于节点数量多, 如果计算每个节点在每一种候选位移下对应点间的 SIFT 特征欧式距离将使得像素层的数据项计算变得极为耗时. 为了有效地减少计算时间, 本文方法在像素层上使用 Census 变换<sup>[20]</sup>提取每个点的局部特征, 然后利用汉明距离计算运动前后的  $C_{desc}$ . 相较于 SIFT 特征, Census 变换求取方便且特征维数低, 通过位运算实现的汉明距离也比欧式距离计算时间短. SPM-BP 方法<sup>[16]</sup>同样利用 Census 变换来计算数据项, 其工作验证了 Census 变换提取的特征作为数据项可以在较小的时间开销下获得稳定的性能.

### 3.3 多帧信息

为了解决由于大位移和物体运动到画面以外而导致的遮挡问题, 进一步提高结果精度, 本文提出的模型考虑了前后多帧图像的信息, 由于物体的运动具有惯性, 同一物体在相邻多帧图像上的运动往往具有一致性, 多帧信息在一定程度上可以弥补前后两帧图像中部分区域丢失的信息<sup>[21]</sup>. 我们在计算数据项  $C_p(f_p)$  时沿着运动方向向前和向后分别映射到相邻多帧图像上, 分别计算数据项然后取最小值作为最终的数据项结果, 以此提高模型的鲁棒性.

图 3 显示了使用多帧信息的示意图.  $I_{t-1}$ 、 $I_t$ 、 $I_{t+1}$  和  $I_{t+2}$  是连续的四幅输入图像, 我们要计算节点  $p_t$  在  $I_t$  与  $I_{t+1}$  之间的光流,  $(u, v)$  是一个候选标签, 则我们将  $p_t$  沿  $(-u, -v)$  移动到  $I_{t-1}$  上的  $p_{t-1}$  处, 沿  $(u, v)$  移动到  $I_{t+1}$  的  $p_{t+1}$  处, 沿  $(2u, 2v)$  移动到  $I_{t+2}$  的  $p_{t+2}$  处. 根据公式 (5) 分别计算  $p_t$  与  $p_{t-1}$ 、 $p_{t+1}$  和  $p_{t+2}$  之间的颜色与特征差异, 得到  $C_{t,t-1}(-f_p)$ 、 $C_{t,t+1}(f_p)$  和  $C_{t,t+2}(2f_p)$ , 取最小值作为  $p_t$  在  $(u, v)$  标签下的数据项, 即:

$$C_p(f_p) = \min\{C_{t,t-1}(-f_p), C_{t,t+1}(f_p), C_{t,t+2}(2f_p)\} \quad (8)$$

对于含有  $n$  个节点, 每个节点含  $k$  个候选位移的 MRF, 只利用单帧图像信息时计算  $C_p(f_p)$  的时间复杂度为  $O(nk)$ , 引入前后帧信息后, 计算  $C_p(f_p)$  的时间复杂度为  $O(3nk)$ , 其时间复杂度的在数量级上未有明显的上升.

### 3.4 后处理

在完成分层置信度传播以后, 通过前后一致性检查来去除错误的位移, 得到一个半稠密的光流场. 之后我们采用 EpicFlow<sup>[22]</sup>的插值方法, 通过现有的半稠密光流场插值填充空白区域的位移, 得到最终的亚像素精度的光流场结果.

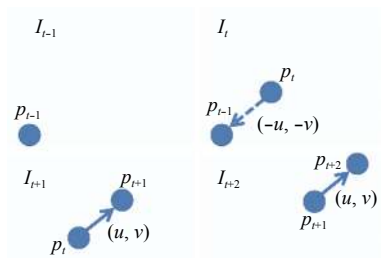


图3 多帧信息示意图

## 4 实验及结果分析

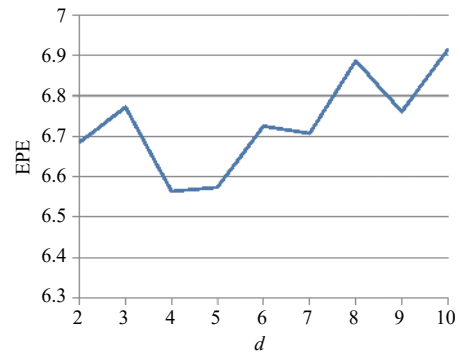
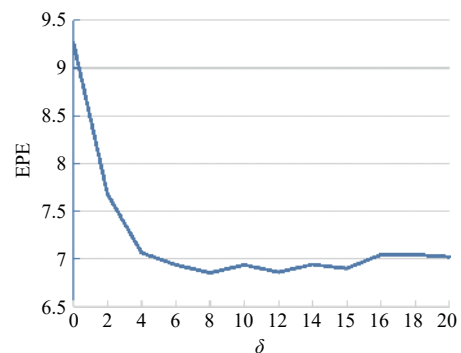
实验在常用的光流数据集 MPI Sintel 数据集<sup>[23]</sup>上对本文方法进行评估. MPI Sintel 数据集提供了包含大位移、运动模糊和大量遮挡等场景的动画视频序列. 数据集包含 23 个训练集图像序列, 12 个测试集图像序列, 每个图像序列都包含清晰版和加入了运动模糊、气流变化的最终版本. 由于最终版本图像更接近于真实场景图像, 实验在最终版本的图像上评估算法性能. 实验平台为 4 核 Intel i5-4590 3.30 GHz CPU 和 8 GB RAM 的个人台式计算机.

### 4.1 参数分析

我们随机选取 MPI Sintel 数据集中 10% 的训练集数据来进行模型参数的评估实验. 在公式 (4) 中, 系数  $\alpha$  和  $\beta$  分别控制着颜色和描述子的权重. 在 3.1 节中, 步长  $d$  影响超像素层置信度传播的标签空间大小,  $\delta$  决定了像素层标签空间的大小. 因为颜色信息和 SIFT 描述子在寻找点匹配中的作用已被许多研究者论证, 我们主要关注  $\delta$  和步长  $d$  对光流估计结果的影响. 实验采用终点误差 (EndPoint Error, EPE) 来衡量得到的光流场与真值之间的误差, 即求每个点的位移向量与真值的欧式距离作为该点的误差, 然后求所有点的误差的平均值作为最终的误差. 在超像素层, 参数设置为  $\alpha=1$ ,  $\beta=250$ , 在像素层, 参数设置为  $\alpha=0$ ,  $\beta=3$ .

图 4 显示了不同  $d$  取值对超像素层置信度传播结果的影响,  $d$  的取值越大, 则候选位移选取时的步长越大, 同样的位移范围下标签空间越小. 此处的误差是在超像素层完成置信度传播后计算得到的. 通过图中的曲线可以看出, 在超像素层的置信度传播运算并不需要高精度的候选位移,  $d=4$ 、 $d=5$  都取得了较好的结果. 因为  $d$  的取值越大, 节点的候选标签空间越小, 消息更新的速度更快, 在后续实验中设置参数  $d=5$ . 图 5 显示了像素点位移范围  $\delta$  的不同取值对误差的影响, 由于有超像素层的位移对像素点的位移范围进行了限制,

在像素层进行置信度传播时,  $\delta$  取较小的值就可以得到比较精确的结果, 扩大像素点标签空间并不会进一步提升光流估计的精度. 通过该实验可知分层结构可以有效地减小像素点的标签空间, 提高置信度传播的执行效率. 在后续试验中设置参数  $\delta=10$ .

图4  $d$  的不同取值对结果误差的影响图5  $\delta$  的不同取值对结果误差的影响

### 4.2 MPI Sintel 实验结果

实验模型参数设置为: 超像素层上  $\alpha=1$ ,  $\beta=250$ , 像素层上  $\alpha=0$ ,  $\beta=3$ ,  $\delta=10$ . 本文模型对原图像进行 1/2 降采样以后执行分层置信度传播, 在超像素层执行 2 轮消息更新迭代, 在像素层执行 3 轮消息更新迭代. 表 1 和表 2 展示了实验结果. “EPE All”表示整幅图像上所有点的位移的误差, “EPE Noc”表示在求取光流的两幅图像上都出现的区域的位移误差. “EPE Occ”表示在第一幅图像中可见而在第二副图像中被遮挡的区域的位移误差. 表 1 显示的是采用了置信度传播的相关方法的实验结果的比较. FullFlow 方法效果略优于我们的结果, 但是由于使用了全局优化的方法, 根据文献中报告的结果, 他们的时间开销将近 240 s, 且是在工作站上运行, 时间及内存开销远大于我们的方法. 而我们的方法在普通 8 GB 内存的台式计算机上的运行时

间为 45 s 左右,说明本文的模型能有效提高置信度传播的执行效率. 本文模型的时间消耗与 SPM-BP 接近,但是获得的光流估计的精度更高,说明本文的模型在效率和精度上取得了较好的平衡. HCOF<sup>[21]</sup>在模型中也加入的多帧信息,但是无论是在处理效率和结果精度上都低于本文方法.

表 1 本文方法与相关方法在 Sintel 上的结果

方法	EPE All	运行时间 (s)	平台
FullFlow <sup>[17]</sup>	5.895	240	Intel I7(3.6 GHz) 64 GB
本文方法	6.127	45	Intel I5(3.3 GHz) 8 GB
SPM-BP <sup>[16]</sup>	7.325	42	Intel I7(3.4 GHz) 8 GB
HCOF+multi <sup>[21]</sup>	8.799	1500	Intel Xeon(2 GHz) 4 GB

表 2 将本文模型的实验结果与目前前沿的一些方法作了比较. 从表中可以看到本文方法的结果优于大部分方法. 我们的模型使用了 EpicFlow<sup>[22]</sup>的插值方法,而其最终结果优于 EpicFlow,这说明本文提出的模型在精确匹配对应点上比 EpicFlow 表现更好. 从表中结果看出本文方法在精度上具有一定的竞争力.

图 6 显示了部分实验结果可视化图像的示例,图中将本文提出的方法与 EpicFlow 与 SPM-BP 的结果进行了对比. 每一列的第一行和第二行为输入图像,第三行为 EpicFlow 的结果,第四行为 SPM-BP 的结果,

第五行是本文方法的结果,最后一行为真值. 图中方框标出的区域显示了本文方法的优势. 从第一列我们可以看到 EpicFlow 的结果在图像右上角出现了明显的错误, SPM-BP 则在图像的右下角区域出现明显的错误,本文的方法与真值更接近. 第二列的 EpicFlow 的结果没有捕捉到人物手臂的运动, SPM-BP 的结果中手臂的运动方向与真值相比出现了明显的偏差,本文方法捕捉到了手臂及身后木棍的运动. 第三列中, EpicFlow 和 SPM-BP 都没有捕捉到左上角手部的运动,而本文提出的方法则成功地捕捉到了手部的运动. 第四列中, EpicFlow 和 SPM-BP 在图像左上方的运动方向都出现了明显的错误.

表 2 本文方法与部分常用方法在 Sintel 上的结果

方法	EPE All	EPE Noc	EPE Occ
FullFlow <sup>[17]</sup>	5.895	2.838	30.793
本文方法	6.127	2.856	32.783
EpicFlow <sup>[22]</sup>	6.285	3.060	32.564
FGI <sup>[24]</sup>	6.607	3.101	35.158
DeepFlow <sup>[9]</sup>	7.212	3.336	38.781
SPM-BP <sup>[16]</sup>	7.325	3.493	38.561
Classic+NLP <sup>[25]</sup>	8.291	4.287	40.925
HCOF+multi <sup>[21]</sup>	8.799	4.980	39.941
LDOF <sup>[6]</sup>	9.116	5.037	42.344
Classic+NL <sup>[4]</sup>	9.153	4.814	44.509

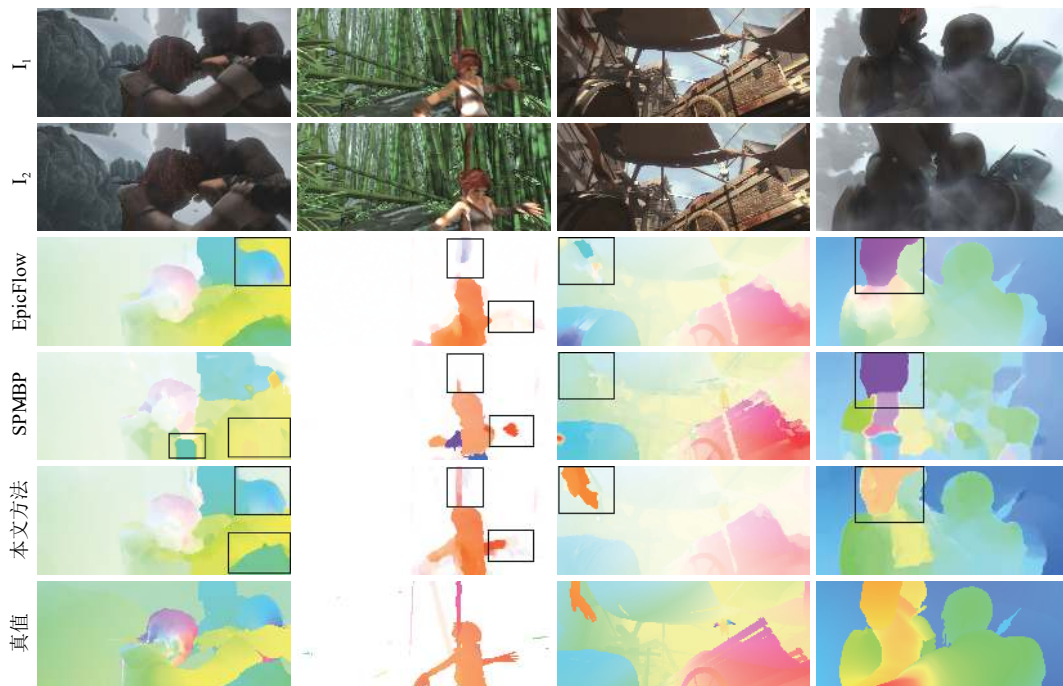


图 6 本文方法与相关方法的可视化结果比较

## 5 结束语

本文提出了一种分层执行置信度传播的方法来进行光流估计,能在较短的时间内获得高精度的大位移光流场.为了克服将置信度传播应用于大位移场景时巨大的内存与时间开销,本文方法首先在超像素图像上执行置信度传播,将每个超像素块作为节点并扩大候选位移的步长以提高运算效率,之后利用得到的超像素位移场来引导像素层的置信度传播,将每个像素节点的标签空间限制在一个较小的范围内,有效提高了像素层置信度传播的效率.本文模型仅使用了常见的SIFT描述子及Census变换就在大位移场景下取得了较为精确的光流估计结果,这说明本文提出的模型能有效地寻找精确的点对应,适用于高精度的大位移光流估计.

本文模型中使用了SLIC超像素分割算法和稠密SIFT描述子,在之后的工作中,设计更为优秀的分割算法和描述子,将进一步提升模型的性能.此外,利用GPU并行实现本文算法也将进一步提升分层置信度传播的执行效率.虽然本文提出的模型有效地提高了置信度传播的执行效率,但其执行速度距离实时计算仍有差距,且计算精度仍有提升空间.我们相信通过后续的工作改进,置信度传播算法将成为光流估计的一个有力工具.

### 参考文献

- 周智,张伟峰,赵斌,等.基于光流的ATM机异常行为实时检测.计算机系统应用,2017,26(9):232-237.[doi:10.15888/j.cnki.csa.005929]
- Horn BKP, Schunck BG. Determining optical flow. *Artificial Intelligence*, 1981, 17(1-3): 185-203. [doi: 10.1016/0004-3702(81)90024-2]
- Brox T, Bruhn A, Papenberger N, *et al.* High accuracy optical flow estimation based on a theory for warping. In: Pajdla T, Matas J, eds. *Computer Vision-ECCV 2004. Lecture Notes in Computer Science*, vol 3024. Springer. Berlin, Heidelberg. 2004. 25-36. [doi: 10.1007/978-3-540-24673-2\_3]
- Sun DQ, Roth S, Black MJ. Secrets of optical flow estimation and their principles. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, CA, USA. IEEE. 2010. 2432-2439. [doi: 10.1109/CVPR.2010.5539939]
- Xu L, Jia JY, Matsushita Y. Motion detail preserving optical flow estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(9): 1744-1757. [doi: 10.1109/TPAMI.2011.236]
- Thomas Brox, Jitendra Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(3): 500-513. [doi: 10.1109/TPAMI.2010.143]
- Brox T, Bregler C, Malik J. Large displacement optical flow. 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA. IEEE. 2009. 41-48. [doi: 10.1109/CVPR.2009.5206697]
- Tu Z, Poppe R, Veltkamp RC. Weighted local intensity fusion method for variational optical flow estimation. *Pattern Recognition*, 2016, 50(C): 223-232.
- Weinzaepfel P, Revaud J, Harchaoui Z, *et al.* Deepflow: Large displacement optical flow with deep matching. 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia. IEEE. 2013. 1385-1392. [doi: 10.1109/ICCV.2013.175]
- Yang JL, Li HD. Dense, accurate optical flow estimation with piecewise parametric model. 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. IEEE. 2015. 1019-1027. [doi: 10.1109/CVPR.2015.7298704]
- Bao LC, Yang QX, Jin HL. Fast edge-preserving patchmatch for large displacement optical flow. *IEEE Transactions on Image Processing*, 2014, 23(12): 4996-5006. [doi: 10.1109/TIP.2014.2359374]
- Bailer C, Taetz B, Stricker D. Flow fields: Dense correspondence fields for highly accurate large displacement optical flow estimation. 2015 IEEE International Conference on Computer Vision. Santiago, Chile. IEEE. 2015. 4015-4023. [doi: 10.1109/ICCV.2015.457]
- Hu YL, Song R, Li YS. Efficient coarse-to-fine patch match for large displacement optical flow. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. IEEE. 2016. 5704-5712. [doi: 10.1109/CVPR.2016.615]
- Barnes C, Shechtman E, Goldman DB, *et al.* The generalized patchmatch correspondence algorithm. In: Daniilidis K, Maragos P, Paragios N, eds. *Computer Vision-ECCV 2010. Lecture Notes in Computer Science*, vol 6313. Springer. Berlin, Heidelberg. 2010. 29-43. [doi: 10.1007/978-3-642-15558-1\_3]
- Felzenszwalb PF, Huttenlocher DP. Efficient belief propagation for early vision. *International Journal of*

- Computer Vision, 2006, 70(1): 41–54. [doi: [10.1007/s11263-006-7899-4](https://doi.org/10.1007/s11263-006-7899-4)]
- 16 Li Y, Min D, Brown MS, *et al.* Spm-bp: Sped-up patchmatch belief propagation for continuous mrfs. 2015 IEEE International Conference on Computer Vision. Santiago, Chile. IEEE. 2015. 4006–4014. [doi: [10.1109/ICCV.2015.456](https://doi.org/10.1109/ICCV.2015.456)]
- 17 Chen QF, Koltun V. Full flow: Optical flow estimation by global optimization over regular grids. arXiv: 1604.03513.
- 18 Achanta R, Shaji A, Smith K, *et al.* Slic superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274–2282. [doi: [10.1109/TPAMI.2012.120](https://doi.org/10.1109/TPAMI.2012.120)]
- 19 Liu C, Yuen J, Torralba A. Sift flow: Dense correspondence across scenes and its applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 978–994. [doi: [10.1109/TPAMI.2010.147](https://doi.org/10.1109/TPAMI.2010.147)]
- 20 Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence. In: Eklundh JO, ed. Computer Vision-ECCV'94. Lecture Notes in Computer Science, vol 801. Springer. Berlin, Heidelberg. 1994. 151–158. [doi: [10.1007/BFb0028345](https://doi.org/10.1007/BFb0028345)]
- 21 Kennedy R, Taylor CJ. Hierarchically constrained optical flow. 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. IEEE. 2015. 3340–3348. [doi: [10.1109/CVPR.2015.7298955](https://doi.org/10.1109/CVPR.2015.7298955)]
- 22 Revaud J, Weinzaepfel P, Harchaoui Z, *et al.* Epicflow: Edge-preserving interpolation of correspondences for optical flow. 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. IEEE. 2015. 1164–1172. [doi: [10.1109/CVPR.2015.7298720](https://doi.org/10.1109/CVPR.2015.7298720)]
- 23 Butler DJ, Wulff J, Stanley GB, *et al.* A naturalistic open source movie for optical flow evaluation. In: Fitzgibbon A, Lazebnik S, Perona P, *et al.*, eds. Computer Vision-ECCV 2012. Lecture Notes in Computer Science, vol 7577. Springer. Berlin, Heidelberg. 2012. 611–625. [doi: [10.1007/978-3-642-33783-3\\_44](https://doi.org/10.1007/978-3-642-33783-3_44)]
- 24 Li Y, Min D, Do MN, *et al.* Fast guided global interpolation for depth and motion. In: Leibe B, Matas J, Sebe N, *et al.*, eds. Computer Vision-ECCV 2016. Lecture Notes in Computer Science, vol 9907. Springer. Cham. 2016. 717–733. [doi: [10.1007/978-3-319-46487-9\\_44](https://doi.org/10.1007/978-3-319-46487-9_44)]
- 25 Sun DQ, Roth S, Black MJ. A quantitative analysis of current practices in optical flow estimation and the principles behind them. International Journal of Computer Vision, 2014, 106(2): 115–137. [doi: [10.1007/s11263-013-0644-x](https://doi.org/10.1007/s11263-013-0644-x)]