

基于结构化预测的细胞跟踪方法

陈旭¹ 万九卿¹

摘要 提出一种新的多细胞联合检测与跟踪方法, 通过椭圆拟合构建细胞观测假说的完备集合, 定义了多种局部事件来描述细胞的行为以及检测阶段可能出现的错误. 通过引入相应的标签变量, 将细胞跟踪建模为结构化预测问题, 通过求解一个带约束的整数规划问题得到细胞轨迹的全局最优解. 针对结构化预测模型中的参数学习问题, 本文采用 Block-coordinate Frank-Wolfe 优化算法根据给定的训练样本求解模型的最优参数, 同时给出了该算法的非线性核化版本. 本文在多个公开数据集上对提出的算法进行了验证, 结果表明, 本文的实验表现相比于传统算法有着显著的提升.

关键词 细胞跟踪, 结构化预测, 结构化学习, Block-coordinate Frank-Wolfe 算法

引用格式 陈旭, 万九卿. 基于结构化预测的细胞跟踪方法. 自动化学报, 2017, 43(3): 376–389

DOI 10.16383/j.aas.2017.c160039

Cell Tracking Using Structured Prediction

CHEN Xu¹ WAN Jiu-Qing¹

Abstract In this work we propose a new joint detection and tracking method for cell tracking. First we develop a new procedure for generating an over complete set of detection hypothesis via ellipse fitting methods. Then we define several local events and corresponding labeling variables to account for the biological behavior of cells and the imperfection in segmentation, and formulate the task of cell tracing as an integer programming problem with constraints. In addition, instead of learning local classifiers, we exploit a recently proposed block-coordinate Frank-Wolfe algorithm to automatically learn optimal parameters of our model. We also present the kernelized version of the learning algorithm which can boost the tracking performance even further. We conduct extensive experiments on public datasets, showing that our method consistently outperforms traditional counterparts.

Key words Cell tracking, structured prediction, structured learning, Block-coordinate Frank-Wolfe algorithm

Citation Chen Xu, Wan Jiu-Qing. Cell tracking using structured prediction. *Acta Automatica Sinica*, 2017, 43(3): 376–389

了解细胞活动及其生存环境是解开许多未知的生命活动机理与疾病机理的关键. 其核心在于描述和量化多种不同的细胞行为, 包括细胞迁移、有丝分裂和细胞凋亡. 现代的时延显微镜技术^[1–2] 可以产生数量巨大的广空间、长时间、多维度细胞图像序列, 这对精确并且鲁棒的自动细胞跟踪算法提出了巨大的需求^[3–7]. 作为多目标跟踪问题的应用之一, 细胞跟踪具有以下几个额外难点: 1) 与传统的行人或车辆^[8–10] 不同, 细胞可以分裂、凋亡或是在跟踪过程中忽然消失, 这导致了跟踪目标的数量变化, 增加了跟踪的复杂性; 2) 显微图像序列中的细胞通常数量巨大, 并且存在密集的重叠, 难以从外观上进行

区分; 3) 由于图像的低信噪比和复杂的细胞轮廓, 细胞分割过程容易出错, 并且这种错误常常被传递到跟踪阶段.

本文结构如下: 第 1 节介绍目前主流的细胞跟踪方法和本文贡献; 第 2 节介绍如何生成细胞观测假说和矛盾集合; 第 3 节和第 4 节分别介绍结构化预测与结构化学习在细胞跟踪和模型训练方面的具体实现; 第 5 节给出实验结果; 第 6 节进行总结.

1 背景介绍

1.1 相关工作

目前主流的细胞跟踪方法分为以下三类:

1) 模型演化方法 (Model evolution)

模型演化方法通常将细胞检测、分割与跟踪联合进行, 在图像序列的第一帧初始化细胞模型, 在后续的帧中更新这些模型^[11–17]. 其通常先建立细胞的外观或几何模型, 例如主动轮廓模型 (Active contours)^[13–16] 和高斯混合模型^[17]. 其中高斯混合模型方法在跟踪时将每个细胞的质心和近似的高斯形状由上一帧传递到当前帧, 同时检测细胞分裂

收稿日期 2016-01-19 录用日期 2016-04-18
Manuscript received January 19, 2016; accepted April 18, 2016
国家自然科学基金 (61174020) 资助
Supported by National Natural Science Foundation of China (61174020)

本文责任编辑 贺威

Recommended by Associate Editor HE Wei

1. 北京航空航天大学自动化科学与电气工程学院 北京 100191
1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191

事件,取得了一定的效果.虽然这类方法有着较为坚实的数学基础,但由于其只考虑相邻帧之间的信息,无法保证全局最优性,在实验表现上也有所欠缺.

2) 先检测-再跟踪方法 (Tracking by detection)

先检测-再跟踪指的是先在图像序列的每一帧中独立地检测并分割出细胞,然后使用一定的寻优算法,例如:多假说跟踪^[18]、整数规划^[19-21]、动态规划^[22]、最小费用流跟踪^[23]等,求解出最有可能的细胞连接轨迹.将细胞的检测与跟踪分开处理使得这类方法在算法替换方面比较灵活,有利于算法的通用性.然而,先检测-再跟踪方法不可避免地会将细胞分割阶段产生的错误传递到跟踪阶段中,跟踪阶段得到的有用信息却无法反过来改善分割的效果,使得跟踪结果严重地依赖于检测和分割的精度.

3) 联合检测与跟踪方法 (Joint detection and tracking)

这类方法联合利用检测和跟踪阶段的信息,使得在跟踪阶段中能够修正分割过程中的错误.通常其先构建出一个全局目标函数,同时建立将不同细胞事件间的约束,通过最小化这个全局目标函数来求解最优轨迹,因此跟踪效果较前两种方法好,也更为鲁棒.其中文献 [24] 使用概率图模型来构建每个检测区域内细胞数目的守恒规则,以保持解的全局一致性.文献 [25] 使用超像素/超体素 (Superpixels/Voxels) 分割方法来产生潜在的观测假说,随后利用全局的时空信息选择概率最大的假说连接轨迹.文献 [26] 使用维特比 (Viterbi) 算法,在全部图像序列上,以一种贪心的方式来增添或修改轨迹,在相关数据集上取得了不错的效果.文献 [27] 将检测与跟踪视作一个网络流规划问题,使用椭圆拟合方法在初始分割图像上生成完备的假说集合,最终通过求解一个整数规划问题得到全局最优的跟踪方案,虽然其假说集的过完备性导致生成的图规模较大,但该方法不需要引入启发式规则就可以应对漏检的观测.以上这些方法几乎都包含能量术语,其参数通常较多且难以进行人工调整,其中大部分方法选择训练局部分类器 (Local classifiers) 来获得局部近似最优参数,因而忽略了各种分配方案之间的联系,无法保证参数以及最终跟踪结果的全局最优性.

1.2 本文贡献

本文在文献 [24-27] 的基础上,提出了一种新的细胞联合检测与跟踪方法.本文主要贡献在于: 1) 使用一种新的流程来产生细胞观测假说和矛盾集合,与文献 [27] 相比,大大减少了生成的假说数量,从而降低了问题的复杂程度; 2) 设计了多个细胞事件以及相应的约束条件,用以解释细胞的各种生物学行

为以及分割中的错误,并且将细胞跟踪问题建模为结构化问题,使用整数规划求解器可以对其进行高效的求解; 3) 不同于文献 [24-27] 训练局部分类器以及文献 [21] 只在成对的相邻帧上进行训练,本文应用了一种高效的结构化支持向量机 (Structured support vector machine, Structured SVM) 学习算法^[28],该算法可以接受图像序列中任意长度的子序列作为训练样本,并且支持在线学习.同时本文也给出了该算法的核化版本 (Kernelized version),以进一步提升跟踪性能; 4) 本文在几个公开的数据集上进行了实验,结果表明本文方法在几个数据集上的表现均超过传统方法.

2 预处理

预处理步骤的主要目的是将原始输入图片转化成基本观测的集合,基本观测是处理的最小单元,每个基本观测都代表一个可能存在的细胞.部分基本观测之间互相矛盾,即它们无法同时存在,我们用矛盾集合的形式表达这种矛盾关系.

2.1 基本观测

预处理的第一步是产生基本的候选细胞观测.首先在每个数据集上利用预先标注好的部分图像,训练一个二值随机森林分类器^[29].此过程首先要对输入图像进行高斯平滑,然后使用 4 种不同类型的低阶特征,分别为像素灰度值、梯度、黑塞值 (Hessian values) 和高斯差分特征进行训练.采用随机森林分类器是一种有监督的分类方法,可以有效地利用训练集中自带的标准分割答案,原理简单,参数较少,对于不同数据集的适应程度较高.采用的特征也是图像分割领域常用的基础特征,不需要针对数据集进行分析选取,进一步加强了分割算法的通用性.缺点是对于复杂的细胞图像,往往存在一定程度的欠分割或过分割.随后将得到的分类器应用于整个图像序列,得到初始分割图像,其中每个前景区域中包含数目未知的细胞,见图 1.

接着利用 Canny 边缘算子从初始分割图像中提取出图像的边缘,再利用文献 [30] 中的方法获得每个前景区域的轮廓数据.然后使用 Ramer-Douglas-Peucker 算法的一个变体^[31]来对轮廓进行多边形拟合,将轮廓分段线性化从而获得轮廓上各关键点的近似曲率.随后通过计算多边形的每个夹角找出轮廓上的凹点,并以这些凹点为分割点将原始轮廓分割为多个单元轮廓.与文献 [27] 将轮廓上每个曲率极值点都作为分割点相比,本方法产生的单元轮廓数量大大减少.接着使用非迭代的最小二乘椭圆拟合算法^[32]将每个单元轮廓以及它们的组合拟合为椭圆.为了进一步减少拟合出的椭圆数目,我们定义

了一个距离来度量原始轮廓和椭圆之间的拟合程度

$$d_{\text{fitness}}(O, E) = \sup_{x \in X} \inf_{y \in Y} (x, y) \quad (1)$$

式 (1) 中 X 是组成原始轮廓 O 的点的集合, Y 是组成椭圆 E 的点的集合, $d(x, y)$ 是点 x 和 y 之间的欧氏距离. 当一个前景区域中的椭圆数目较多时, 拟合值较低的椭圆将被删去以降低后续运算的复杂度, 由剩下的椭圆组成基本观测集合.

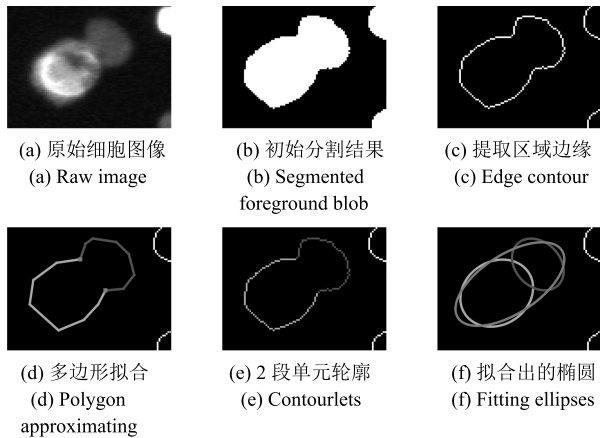


图 1 生成基本观测的流程

Fig. 1 Generating basic detections

图 1 展示本节方法的优势. 经过以上一系列操作, 本算法从原本严重欠分割的初始分割方案图 1(b) 中将 3 个潜在的候选细胞 (见图 1(f)) 以椭圆的方式近似恢复了出来, 其中存在 2 个正确的观测和 1 个冗余的观测. 有效地防止了将分割时产生的误差传递到跟踪阶段, 同时将细胞信息抽象化为椭圆的信息, 将细胞图像表现为椭圆集合的形式, 摒弃了大量的冗余信息, 有利于后续计算机的快速处理. 这也是本算法与传统的先检测-再跟踪方法的主要区别之一.

2.2 矛盾集合

利用第 2.1 节中的方法产生的基本观测可能存在互相矛盾的情况. 假设给定一个前景区域, 其轮廓上存在 k 个凹点, 那么该轮廓将被分割为 k 个单元轮廓, 其组合方式有 $2^k - 1$ 种, 其中每种组合轮廓都可以拟合出一个相应的椭圆 $e_j, j = 1, \dots, 2^k - 1$. 规定每个单元轮廓至多只能被使用 1 次, 即若 2 段不同的组合轮廓中包含了同一个单元轮廓, 则由这 2 段组合轮廓拟合而来的椭圆是相互矛盾的. 这样一来, 对于一个前景区域可以构造出 k 个矛盾椭圆集合 $C_r, r = 1, \dots, k$, 其中 C_r 由包含了第 r 段单元轮廓的所有组合轮廓对应的椭圆构成, 图 2 中给出了一个具体的例子. 图 2(a) 中有 3 段单元轮廓, 图

2(b) 中每行代表一个单元轮廓, 每一列代表一个椭圆, 如果单元轮廓被某个椭圆使用了, 则矩阵的相应位置为 1, 否则为 0. 图 2(c) 中第 r 个矛盾集合由所有使用了第 r 段单元轮廓的椭圆构成. 对于一个给定的前景区域, 首先构造出一个单元轮廓-椭圆的逻辑矩阵用以表明某个椭圆中是否用到了某个单元轮廓. 以图 2 为例, 3 个单元轮廓一共可以产生 7 种可能的椭圆, 其中椭圆 e_3 在矩阵中的列向量为 $[1, 1, 0]^T$, 表示其是由单元轮廓 x_1 和 x_2 的组合拟合而来. 定义矛盾集合的目的主要是为了符合细胞图像的真实情况, 同时限制有效观测的数目. 例如图 2 中若选定 e_7 作为有效基本观测, 则 $e_1 \sim e_6$ 由于和 e_7 矛盾而均无法被选取, 从图像上直观地看, 是因为 e_7 已经占据了该前景区域的全部面积, 不可能再允许其他基本观测的出现. 而 e_1, e_2, e_4 三者之间彼此都不矛盾, 若选定其作为 3 个有效观测, 从图像上看, 三者面积有部分重叠, 与细胞的实际情况相符. 因此, 即使选择了错误的观测, 由于其矛盾集合的存在, 可以避免集合内其他错误观测的产生, 从而达到限制错误观测数目的目的.

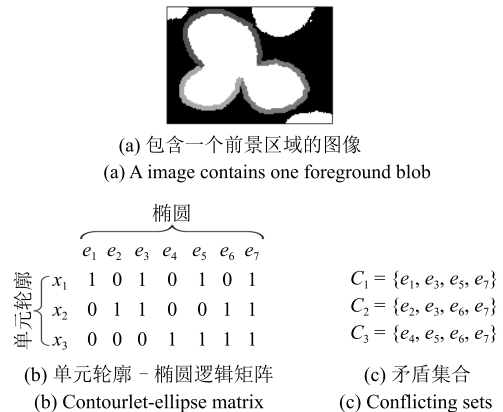


图 2 矛盾集合的图示

Fig. 2 Example of conflicting sets

3 结构化预测

本节通过求解基本观测的最优连接来解决细胞跟踪问题. 首先, 定义 6 种细胞局部事件 (局部指的是事件只发生在某一细胞及其邻近细胞之间) 来描述基本观测和其邻居观测之间的连接模式, 并为每个事件定义一个二值的标签变量用以指示该事件是否发生. 接着, 将细胞局部事件和基本观测间的依赖关系表述为标签变量间的线性约束, 通过求解带约束的整数规划问题来找出与原始图像序列最为匹配的连接轨迹.

3.1 细胞局部事件与标签变量

首先定义一些细胞局部事件以描述相邻 2 帧间

的基本观测和/或观测对(即一对基本观测)的连接模式(见图3), 以此解释细胞的各种生物学行为, 包括细胞迁移、有丝分裂、凋亡; 以及检测和分割中的错误, 例如过分割(Over-segmentation)与欠分割(Under-segmentation). 本文定义了6种不同类型的局部事件: 迁移、分裂、出现、消失、分离、合并. 首先定义第 t 帧中, 基本观测的集合为 $\{v_j\}$, 为了便于描述分裂、分离和合并这3个局部事件, 引入观测对的概念, 由 $\{v_j\}$ 中的相邻观测两两组合得到观测对的集合 $\{q_\beta\}$. 以下对各个局部事件进行阐述:

1) 迁移: 若第 $t-1$ 帧中的细胞 v_i 迁移到第 t 帧中的 v_j , 则认为 v_i 和 v_j 之间的迁移事件为真.

2) 分裂: 若第 $t-1$ 帧中的母细胞 v_i 有丝分裂为到第 t 帧中的观测对 q_β (包含2个子细胞), 则认为 v_i 和 q_β 之间的分裂事件为真.

3) 出现: 出现事件是从源节点 S 到细胞 v_j 的连接, 用于解释一个新的细胞(往往是从图像边界)迁移进入了图像区域.

4) 消失: 与出现类似, 消失事件是从细胞 v_j 到凋亡节点 D 的连接, 表明细胞的自然凋亡或是迁移出了图像区域.

5) 分离: 分离事件表示第 $t-1$ 帧中的细胞 v_i 分离为第 t 帧中的观测对 q_β , 与分裂事件不同的是, 此处 v_i 实际上由欠分割产生(即 v_i 内实际包含2个细胞)或者 q_β 由过分割产生(即 q_β 实际是单个细胞).

6) 合并: 合并事件表示第 t 帧中的观测对 q_β 合并为第 $t+1$ 帧的细胞 v_k , 此处 v_k 实际上由欠分割产生或者 q_β 由过分割产生.

注意在以上6种局部事件中, 迁移、分裂、消失和出现是对真实细胞行为的描述, 传统方法如文献[26-27]大多只定义了这4种细胞事件. 而本文在其基础上又加入了两个虚拟事件: 合并与分离, 因此不仅能够描述真实的细胞行为, 而且还能对细胞分割阶段产生的欠分割与过分割错误进行解释并纠正, 从而防止分割阶段的错误传递到跟踪阶段中. 这也是本算法相比于传统的先检测-再跟踪方法的创新点之一.

我们为每个细胞局部事件分配一个0/1标签变量 z 来表示该事件的发生与否. 所有标签变量的任一分配方案都代表了一种可能的细胞连接轨迹, 且必须满足下一节中将要提到的约束条件, 图4给出了一个分配方案的示例.

图4(a)为所有可能的连接方式中的一种, 实线箭头表示发生的事件, 虚线箭头表示未发生的事件. 图4(b)为图4(a)边的一种等价表示, 使用标签变量来标识事件的发生与否. 由于篇幅的限制, 图中只列出了部分局部事件和对应的标签变量.

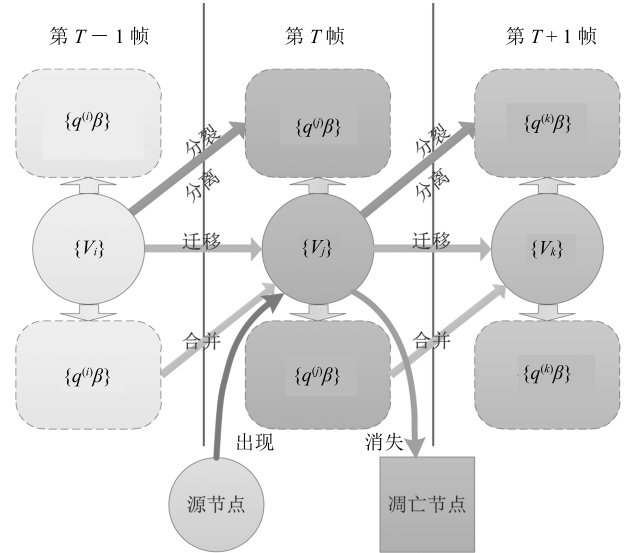
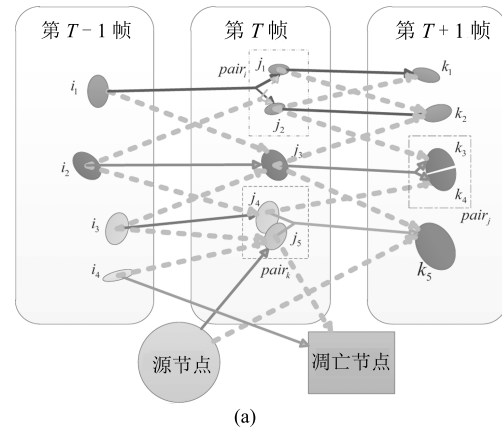


图3 基本观测、观测对和细胞局部事件
Fig. 3 Basic detections, detection pairs and local events



t	局部事件	$t+1$	z
i_1	分裂	$\{j_1, j_2\}$	$z_{i_1\alpha}^{divide} = 1, \alpha = pair_i$
i_2	分裂	$\{j_1, j_2\}$	$z_{i_2\alpha}^{divide} = 0, \alpha = pair_i$
i_2	迁移	j_3	$z_{i_2j_3}^{move} = 1$
i_3	迁移	j_3	$z_{i_3j_3}^{move} = 0$
i_4	消失		$z_{i_4}^{disappear} = 1$
j_5	消失		$z_{j_5}^{disappear} = 0$
	出现	j_5	$z_{j_5}^{appear} = 1$
	出现	k_5	$z_{k_5}^{appear} = 0$
j_3	分离	$\{k_3, k_4\}$	$z_{j_3\alpha}^{split} = 1, \alpha = pair_j$
j_4	迁移	k_4	$z_{j_4k_4}^{move} = 1$
$\{j_4, j_5\}$	合并	k_5	$z_{\alpha k_5}^{merge} = 1, \alpha = pair_k$

图4 细胞连接轨迹、局部事件和标签变量的示例
Fig. 4 Illustration of links, local events and labeling variables

理论上, 所有可能的标签变量的数量与所有基本观测的数量呈现三次方关系, 这使得变量数目过多. 因此本文只考虑发生在局部近邻之间的细胞事件, 具体来讲, 对于迁移事件, 只考虑每个基本观测

与其相邻帧中 6 个最近的基本观测之间的连接; 对于分裂、分离和合并事件, 只考虑每个基本观测与其相邻帧中 15 个最近的观测对之间的连接. 通过这种方式, 标签变量的个数降低为与基本观测的个数呈现线性关系.

3.2 约束关系

本文将标签变量级联成一个 N 维的向量 \mathbf{z} , 此处 N 是所有标签变量的数目. 注意组成 \mathbf{z} 的标签变量并不是互相独立的, 而是满足如下几个约束条件.

约束 1. 任何一个基本观测的来路和去路的个数应该相等, 即

$$\begin{aligned} & \sum_{i \in \mathcal{N}^1(j)} \mathbf{z}_{ij}^{\text{move}} + \mathbf{z}_j^{\text{appear}} + \sum_{\alpha \in \mathcal{N}^2(j)} \mathbf{z}_{\alpha j}^{\text{merge}} + \\ & \sum_{\beta: j \in q_\beta} \sum_{i: \beta \in \mathcal{N}^2(i)} \mathbf{z}_{i\beta}^{\text{divide}} + \sum_{\beta: j \in q_\beta} \sum_{i: \beta \in \mathcal{N}^2(i)} \mathbf{z}_{i\beta}^{\text{split}} = \\ & \sum_{k \in \mathcal{N}^1(j)} \mathbf{z}_{jk}^{\text{move}} + \mathbf{z}_j^{\text{disappear}} + \sum_{\gamma \in \mathcal{N}^2(j)} \mathbf{z}_{j\gamma}^{\text{divide}} + \\ & \sum_{\gamma \in \mathcal{N}^2(j)} \mathbf{z}_{j\gamma}^{\text{split}} + \sum_{\beta: j \in q_\beta} \sum_{k: \beta \in \mathcal{N}^2(k)} \mathbf{z}_{\beta k}^{\text{merge}}, \quad \forall r \end{aligned} \quad (2)$$

式中, $\mathcal{N}^1(j)$ 和 $\mathcal{N}^2(j)$ 分别代表基本观测 v_j 的邻近基本观测集合和邻近观测对集合. 等号左边描述 v_j 的来路: 即 v_j 可能由前一帧中的 v_i 迁移而来, 或是从图像外部进入图像区域, 或是由前一帧中的观测对 q_α 合并得到, 或是属于观测对 q_β 中的一个, 而 q_β 是由前一帧中的 v_i 分裂或分离得到. 等号右边描述 v_j 的去路: 即 v_j 可能迁移为下一帧中的 v_k , 或是凋亡/迁移出图像区域, 或是分裂/分离为下一帧中的观测对 q_γ , 或是属于某一个观测对 q_β , 而 q_β 将会合并为下一帧中的基本观测 v_k . 这里将式 (2) 中等号左边或右边定义为观测 v_j 的值, 记作 V_j .

约束 2. 一个基本观测的来路或去路至多为 1 条, 即

$$V_j \leq 1, \quad \forall j \quad (3)$$

式 (3) 规定任意观测 v_j 的值只能为 0 或 1. $V_j = 0$ 说明 v_j 是一个虚观测.

约束 3. 在每个矛盾集 C_r 中 (参见第 2 节), 至多只能有一个观测的值为 1, 即

$$\sum_{j \in C_r} V_j \leq 1, \quad \forall r \quad (4)$$

注意一个矛盾集中的所有观测值均为 0 的情况也是可能发生的.

定义 \mathcal{Z} 为向量 \mathbf{z} 的所有可能的取值, 即

$$\mathcal{Z} := \left\{ \mathbf{z} \in (\mathbf{0}, \mathbf{1})^N \mid \mathbf{z} \text{ s.t. (2), (3), (4)} \right\} \quad (5)$$

式 (2)~(4) 阐明了标签变量之间的相关关系, 构造出了 \mathbf{z} 的取值空间的结构. 注意这 3 个约束的构建并没有涉及人为定义的启发式信息, 而是细胞在迁移、分裂等各种行为中必须遵守的基本逻辑约束. 跟踪过程中如果违背了这 3 个约束, 例如一个细胞没有来路, 却有去路, 是不符合现实情况的. 本文创新性地将分离与合并事件融入到 3 条约束中, 从而无需另外为这 2 个虚拟事件定义额外约束, 简化了约束形式.

3.3 整数规划

\mathbf{z} 的一种取值对应于图像序列中的一种细胞跟踪轨迹, 因此可以通过找出与观测最为匹配的标签 $\mathbf{z}^* \in \mathcal{Z}$ 来解决细胞跟踪问题. 本文将局部事件 e 的匹配度定义为内积 $\langle \mathbf{x}_{(\cdot)}^e, \mathbf{w}^e \rangle$, 其中 $\mathbf{z}_{(\cdot)}^e$ 是从图像中提取出的与该事件相关的特征向量, \mathbf{w}^e 是参数向量, 其编码了从训练样本中学习得到的先验知识. 因此 \mathbf{z} 的匹配度就可以用所有局部事件的匹配度之和表示, 即

$$\begin{aligned} \langle \mathbf{w}, \Phi(\mathbf{x}, \mathbf{z}) \rangle &= \sum_{(i,j)} \langle \mathbf{x}_{ij}^{\text{move}}, \mathbf{w}^{\text{move}} \rangle \mathbf{z}_{ij}^{\text{move}} + \\ & \sum_j \langle \mathbf{x}_j^{\text{appear}}, \mathbf{w}^{\text{appear}} \rangle \mathbf{z}_j^{\text{appear}} + \\ & \sum_j \langle \mathbf{x}_j^{\text{disappear}}, \mathbf{w}^{\text{disappear}} \rangle \mathbf{z}_j^{\text{disappear}} + \\ & \sum_{(j,\gamma)} \langle \mathbf{x}_{j\gamma}^{\text{divide}}, \mathbf{w}^{\text{divide}} \rangle \mathbf{z}_{j\gamma}^{\text{divide}} + \\ & \sum_{(j,\gamma)} \langle \mathbf{x}_{j\gamma}^{\text{split}}, \mathbf{w}^{\text{split}} \rangle \mathbf{z}_{j\alpha}^{\text{split}} + \\ & \sum_{(\alpha,j)} \langle \mathbf{x}_{\alpha j}^{\text{merge}}, \mathbf{w}^{\text{merge}} \rangle \mathbf{z}_{\alpha j}^{\text{merge}} \end{aligned} \quad (6)$$

式中, \mathbf{w} 为各个事件参数向量的级联, Φ 为所有发生了的事件的特征向量之和, 则最优标签变量可以从下式中求出

$$\mathbf{z}^* = \arg \max_{\mathbf{z} \in \mathcal{Z}} \langle \mathbf{w}, \Phi(\mathbf{x}, \mathbf{z}) \rangle \quad (7)$$

当给定参数 \mathbf{w} 的情况下, 通过求解上述带有线性约束和线性目标函数的整数规划问题, 就可以得到 \mathbf{z}^* . 本文使用商业数学规划求解器 CPLEX 中的分支剪枝算法 (Branch-and-cut) 来高效地求解式 (7), 通常求解一个含有 15 万个变量左右的问题只需要花费 1~2s. 由于整数规划的特点, 使用其求解式 (7) 可以得到满足相关约束下的全局最优的标签

分配方案, 文献 [24–25, 27] 中也应用了该技术, 并取得了不错的效果, 因此本文也选取整数规划作为最终求解的方法.

式 (7) 定义了一个从特征空间到标签空间的预测函数 $f_w: \mathcal{X} \rightarrow \mathcal{Z}$, 其中含有参数 $w \in \mathbf{R}^D$ (D 是参数的数目). 第 4 节中将探究如何从给定的训练集中学习出最优参数.

4 参数学习

设置合适的参数 w 对于预测函数 f_w 的精度是至关重要的. 然而, 这个任务并不容易, 主要因为: 1) 每个参数或参数组合的意义和影响通常比较模糊. 2) 对于高维特征来说, 参数空间会变得非常大. 因此, 本文诉诸于基于学习的方法, 它能够从带有真实跟踪轨迹的训练图像序列中, 以先验知识的形式, 自动地给出参数.

4.1 结构风险最小化

给定 K 个带标签的训练样本 $\{(\mathbf{x}^k, \mathbf{z}^k)\}$, 参数向量 w^* 的最优解可以通过求解以下的结构风险最小化问题^[33] 得到:

$$w^* = \arg \max_{w \in \mathbf{R}^D} \frac{\lambda}{2} \|w\|^2 + \frac{1}{K} \sum_{k=1}^K L(\mathbf{z}^k, f_w(\mathbf{x}^k)) \quad (8)$$

式中, 加号左侧是 L2 正则化项, 用来防止过拟合, 其中 λ 是该项的权重. 加号右侧是期望风险的经验估计, 损失函数 $L: \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbf{R}_+$ 定义为真实标签 z 和预测标签 $\hat{z} = f_w(\mathbf{x})$ 的汉明距离

$$L(z, \hat{z}) := \frac{1}{|z|} \sum_e \sum_{(\cdot, \cdot)} z_{(\cdot, \cdot)}^e (1 - \hat{z}_{(\cdot, \cdot)}^e) + \hat{z}_{(\cdot, \cdot)}^e (1 - z_{(\cdot, \cdot)}^e) \quad (9)$$

注意损失函数相对于 z 来说是线性的, 这也为式 (8) 的高效求解提供了可能.

直接求解式 (8) 是不可行的. 然而, 统计学习理论中的结论指出: 如果能够最小化式 (8) 中目标函数的凸上界, 一样可以在无限数据的情况下获得最优的预测精度^[33]. 这导出了以下的支持向量训练过程:

$$z^* = \arg \max_{w \in \mathbf{R}^D} \frac{\lambda}{2} \|w\|^2 + \frac{1}{K} \sum_{k=1}^K H_k(w) \quad (10)$$

其中

$$H_k(w) := \max_{z \in \mathcal{Z}_k} L(z^k, z) - \langle w, \Phi(\mathbf{x}^k, z^k) \rangle + \langle w, \Phi(\mathbf{x}^k, z) \rangle \quad (11)$$

式中, $\mathcal{Z}_k = \mathcal{Z}(\mathbf{x}^k)$. 注意 $H_k(w)$ 是 $L(z^k, f_w(\mathbf{x}^k))$ 的一个凸上界, 并且相对于 w 来说是凸的, 得到 $H_k(w)$ 需要求解一个整数规划问题, 也称作损失增强型解码 (Loss-augmented decoding) 子问题, 与式 (7) 相似, 这个问题也可以利用 CPLEX 进行高效的求解. 为便于后面讨论, 此处定义

$$L_k(z) = L(z^k, z), \quad \Psi_k(z) = \Phi(z^k, z^k) - \Phi(\mathbf{x}^k, z) \\ H_k(z; w) = L_k(z) - \langle w, \Psi_k(z) \rangle$$

4.2 Block-coordinate Frank-Wolfe 优化算法

式 (10) 的一种带有 K 个松弛变量的等价形式如下:

$$(w^*, \xi^*) = \arg \max_{w, \xi \geq 0} P(w, \xi) = \\ \frac{\lambda}{2} \|w\|^2 + \frac{1}{K} \sum_{k=1}^K \xi_k \\ \text{s.t.} \quad \langle w, \Psi_k(z) \rangle \geq L_k(z) - \xi_k, \quad \forall k, \forall z \in \mathcal{Z}_k \quad (12)$$

注意目标函数 $P(w, \xi)$ 相对于 w 来说是二次的, 对于 ξ 是线性的, 并且所有的约束都是线性的. 但约束的数量过多, 对于每个样本均有 $|\mathcal{Z}_k|$ 个约束条件.

式 (12) 的拉格朗日对偶形式如下:

$$\alpha^* = \arg \max_{\alpha \geq 0} D(\alpha) = \frac{\lambda}{2} \|A\alpha\|^2 + b^T \alpha \\ \text{s.t.} \quad \sum_{z \in \mathcal{Z}_k} \alpha_k(z) = 1, \quad \forall k \quad (13)$$

其中, $\alpha_k(z)$ 代表原问题的第 (k, z) 个约束条件. 式 (13) 中 A 的每一列定义为 $(1/\lambda K)\Psi_k(z)$, 向量中的每个元素定义为 $(1/K)L_k(z)$. 注意对偶问题的变量数目是标签变量的指数量级, 与原问题中的约束数目相同. 给定一个对偶向量 α , 可以使用 Karush-Kuhn-Tucker 最优条件, 选择出对应的原向量 $w = A\alpha$. 而对偶目标函数的梯度具有如下简单形式:

$$\nabla D(\alpha) = \lambda A^T A \alpha - b = \lambda A^T w - b \quad (14)$$

最终, 对偶问题的可行域是 K 个概率单纯形的乘积.

经典的 Frank-Wolfe 算法 (1956) 是一种迭代的优化算法, 可以用它来求解式 (13). 在每一轮迭代中, 该算法使用当前循环的 α 在可行域 \mathcal{M} 上通过最小化 D 的线性化超平面得到搜索角点 (Search corner) s

$$s = \arg \max_{s' \in \mathcal{M}} \langle s', \nabla D(\alpha) \rangle \quad (15)$$

注意式 (15) 可以被解耦为 K 个线性规划子问题, 其中的每个子问题都对应于一个训练样本, 并且可以被独立求解. 而下一轮迭代则可视作 \mathbf{s} 和上一轮迭代的凸组合, 并且使用最优步长 γ . 经典 Frank-Wolfe 算法的主要缺点在于, 当其应用于式 (15) 时, 每一轮迭代中都需要求解 K 个线性规划问题, 当数据集含有大量的训练样本时, 学习过程会变得相当慢.

因此本文使用了近期提出的 Block-coordinate Frank-Wolfe 算法^[28] 来解决式 (15), 此算法继承了 Frank-Wolfe 算法的优异特性, 同时在每一轮迭代中只需要求解一次线性规划子问题. 不同于 Frank-Wolfe 算法需要在整个可行域 \mathcal{M} 上搜索 \mathbf{s} , 该算法每次迭代中只在单纯形 Δ_k 上搜索 \mathbf{s} 的一个区块. 这里定义 $\mathbf{s}_{(k)}$ 为向量 \mathbf{s} 的第 k 个坐标区块, $\nabla_{(k)} D(\boldsymbol{\alpha})$ 为第 k 个坐标区块的梯度向量. 注意, 在概率单纯形 Δ_k 上最小化线性函数 $\langle \mathbf{s}_{(k)}, \nabla_{(k)} D(\boldsymbol{\alpha}) \rangle$ 等价于求解损失增强型解码子问题 (11), 这也是此算法的一个关键点. 这个等价关系来自于: 在 Δ_k 上最小化 $\langle \mathbf{s}_{(k)}, \nabla_{(k)} D(\boldsymbol{\alpha}) \rangle$ 相当于搜索 Δ_k 的角点, 或是找出向量 $\nabla_{(k)} D(\boldsymbol{\alpha})$ 的最小组成部分, 而 $\nabla_{(k)} D(\boldsymbol{\alpha})$ 的第 z 个组成部分正是 $-(1/K)H_k(\mathbf{z}; \mathbf{w})$.

完整的 Block-coordinate Frank-Wolfe 算法在算法 1^[28] 中给出.

算法 1. Block-coordinate Frank-Wolfe 优化算法

1. Let $\mathbf{w}^{(0)} = \mathbf{w}_k^{(0)} = 0$, $\ell^{(0)} = \ell_k^{(0)} = 0$
2. For $t = 0, \dots, T$ do
3. Pick k at random in $\{1, \dots, K\}$
4. Solve $\mathbf{z}^* = \arg \max_{\mathbf{z} \in \mathcal{Z}_k} H_k(\mathbf{z}, \mathbf{w}^{(t)})$
5. Let $\mathbf{w}_s = (1/\lambda K)\Psi_k(\mathbf{z}^*)$ and $\ell_s = (1/K)L_k(\mathbf{z}^*)$
6. Let $\gamma = 2K/(t + 2K)$
7. Update $\mathbf{w}_k^{(t+1)} = (1 - \gamma)\mathbf{w}_k^{(t)} + \gamma\mathbf{w}_s$
8. and $\ell_k^{(t+1)} = (1 - \gamma)\ell_k^{(t)} + \gamma\ell_s$
9. Update $\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} + \mathbf{w}_k^{(t+1)} - \mathbf{w}_k^{(t)}$
10. and $\ell^{(t+1)} = \ell^{(t)} + \ell_k^{(t+1)} - \ell_k^{(t)}$
11. End for

注意在算法 1 中只得到了原变量. 第 5 行中, $\mathbf{w}_s = A\mathbf{s}_{[k]}$ 和 $\ell_s = \mathbf{b}^T \mathbf{s}_{[k]}$, 其中的 $\mathbf{s}_{[k]}$ 为 $\mathbf{s}_{(k)} = \mathbf{e}^{\mathbf{z}^*}$ 的补 0 扩展结果. 第 6 行中, 步长 γ 可以通过最小化 $D(\boldsymbol{\alpha}^{(t)} + \gamma(\mathbf{s} - \boldsymbol{\alpha}^{(t)}))$ 得到, 然后映射到 $[0, 1]$ 中. 本文创新性地将 Block-coordinate Frank-Wolfe 算法应用到了细胞跟踪领域, 主要是基于速度与运算能力的考虑. 相比于传统的割平面法^[21], Block-coordinate Frank-Wolfe 算法每轮循环只选取一个样本进行问题 (11) 的求解, 大大缩短了每轮循环的时间, 同时能够得到相同的收敛速率保证^[28]. 另一方面, 某些规模较大的训练样本可能会超出计算机的处理能力, 此时可将其拆分为数个小样本, 并使用

Block-coordinate Frank-Wolfe 算法进行训练, 在保证数据总量不变的情况下往往可以得到相当的效果, 下文中的实验部分将对这一点进行展示.

4.3 核化学习算法

线性函数 $\langle \mathbf{x}, \mathbf{w} \rangle$ 对于某些局部事件来说可能缺少区别能力. 因此, 本文使用非线性映射函数 φ 将特征向量 \mathbf{x} 映射到一个更高维或是无限维空间中, 并在这个新的空间中使用线性函数进行事件的分类, 即核化的训练算法. 为了实现这一目的, 需要对算法 1 进行以下两方面的调整: 首先, 需要在每轮迭代中显式地得到对偶向量 $\boldsymbol{\alpha}^{(t)}$. 注意 $\boldsymbol{\alpha}^{(t)}$ 是起始点 $\boldsymbol{\alpha}^{(0)}$ 和先前每轮迭代中找出的角点 \mathbf{s} 的凸组合, 因此非 0 对偶变量的个数至多只能与迭代轮数相同. 算法只需跟踪 $\boldsymbol{\alpha}^{(t)}$ 的非 0 元素即可, 避免了指数的产生. 其次, 需要重写 $\langle \mathbf{w}, \Phi \rangle$, 将原特征向量 \mathbf{x} 的内积形式替代为适当的核函数. 使用等式 $\mathbf{w} = A\boldsymbol{\alpha}$, 得到:

$$\begin{aligned} \langle A\boldsymbol{\alpha}, \Phi(\mathbf{x}, \mathbf{z}) \rangle &= \\ &= \frac{1}{\lambda K} \sum_{k', \mathbf{z}' \in \mathcal{Z}_{k'}} \boldsymbol{\alpha}_{(k')}(\mathbf{z}') \langle \Psi_{k'}(\mathbf{z}'), \Phi(\mathbf{x}, \mathbf{z}) \rangle = \\ &= \frac{1}{\lambda K} \sum_{e \in \mathcal{E}} \sum_{k', \mathbf{z}' \in \mathcal{Z}_{k'}} \boldsymbol{\alpha}_{(k')}(\mathbf{z}') \langle \Psi_{k'}^{(e)}(\mathbf{z}'), \Phi^{(e)}(\mathbf{x}, \mathbf{z}) \rangle = \\ &= \frac{1}{\lambda K} \sum_{e \in \mathcal{E}} \sum_{k', \mathbf{z}' \in \mathcal{Z}_{k'}} \boldsymbol{\alpha}_{(k')}(\mathbf{z}') \langle \Phi^{(e)}(\mathbf{x}^{k'}, \mathbf{z}^{k'}) - \\ & \quad \Phi^{(e)}(\mathbf{x}^{k'}, \mathbf{z}') \rangle, \Phi^{(e)}(\mathbf{x}, \mathbf{z}) \rangle \end{aligned} \quad (16)$$

式中, $\Psi^{(e)}$ 和 $\Phi^{(e)}$ 表示事件 e 对应的子向量. 不妨以出现事件为例 (即 $e = \text{appear}$), $\Phi^{(e)}$ 之间的内积可以被展开如下:

$$\begin{aligned} \langle \Phi^{(\text{appear})}(\mathbf{x}^{k'}, \mathbf{z}^{k'}), \Phi^{(\text{appear})}(\mathbf{x}, \mathbf{z}) \rangle &= \\ \left\langle \sum_i \mathbf{x}_i^{k'(\text{appear})} \mathbf{z}_i^{k'(\text{appear})}, \sum_j \mathbf{x}_j^{(\text{appear})} \mathbf{z}_j^{(\text{appear})} \right\rangle &= \\ \sum_i \sum_j \mathbf{z}_i^{k'(\text{appear})} \mathbf{z}_j^{(\text{appear})} \langle \mathbf{x}_i^{k'(\text{appear})}, \mathbf{x}_j^{(\text{appear})} \rangle &= \\ \sum_i \sum_j \mathbf{z}_i^{k'(\text{appear})} \mathbf{z}_j^{(\text{appear})} \mathcal{K}(\mathbf{x}_i^{k'(\text{appear})}, \mathbf{x}_j^{(\text{appear})}) & \end{aligned} \quad (17)$$

式 (16) 中的其他项也可以按相似方法展开. 注意, 核化的 $\langle A\boldsymbol{\alpha}, \Phi \rangle$ 仍然是 \mathbf{z} 的线性函数, 使得问题 (7) 和 (11) 仍然可以被高效求解. 将 $\langle A\boldsymbol{\alpha}, \Phi \rangle$ 基于不同事件 $e \in \mathcal{E}$ 进行解耦的一大好处是对不同的事件使用不同的核函数. 式 (16) 和式 (17) 是

通过细胞事件中基于核函数的原理推导而出, 文献 [28] 中虽然给出了核化的 Block-coordinate Frank-Wolfe 算法, 但其只是一个原理性的框架, 对于实际应用必须以具体的公式去替代其中的内容, 这也是本文的创新点之一.

完整的核化 Block-coordinate Frank-Wolfe 算法在算法 2^[28] 中给出.

算法 2. 核化 Block-coordinate Frank-Wolfe 算法

1. Let $\alpha^{(0)} = (\mathbf{e}^{z^1}, \dots, \mathbf{e}^{z^K}) \in \mathcal{M}$
2. and $\alpha_{(k)}^{(0)} = \mathbf{e}^{z^k} \in \Delta_k$
3. For $t = 0, \dots, T$ do
4. Pick k at random in $\{1, \dots, K\}$
5. Solve $\mathbf{z}^* = \arg \max_{\mathbf{z} \in \mathcal{Z}_k} H_k(\mathbf{z}, A\alpha^{(t)})$
6. Let $\mathbf{s}_{(k)} = \mathbf{e}^{z^*} \in \Delta_k$
7. Let $\gamma = 2K/(t + 2K)$
8. Update $\alpha_{(k)}^{(t+1)} = (1 - \gamma)\alpha_{(k)}^{(t)} + \gamma\mathbf{s}_{(k)}$
9. Update $\alpha^{(t+1)} = \alpha^{(t)} + \alpha_{[k]}^{(t+1)} - \alpha_{[k]}^{(t)}$
10. End for

第 5 行需要求解一个核化的损失增强型解码子问题, 其中 $\Phi^{(e)}$ 间的内积使用式 (17) 得到. 第 9 行中, 为了保持与 α 的长度相同, $\alpha_{[k]}$ 表示 $\alpha_{(k)}$ 的补 0 扩展结果.

5 实验结果

本文使用由细胞跟踪挑战文献 [34] 提供的四个公开数据集, 包括两个仿真数据集: Fluo-N2DH-SIM+-01、SIM+-02, 两个真实数据集: Fluo-N2DL-HeLa-1、HeLa-2. 四者均包含大量的细胞, 同时伴随着多种细胞行为, 细胞与背景的对比较低、细胞形态复杂以及细胞间的相互重叠进一步加强了分割与跟踪的难度. 表 1 中列出了与数据集相关的一些统计数据, 其中也包含了使用第 2.1 节中的方法得到的分割结果的 F-值 (F-measures).

5.1 算法实现平台

为了便于文献 [29] 的集成调用, 本文算法的检测分割部分在 Java 平台上实现. 其他部分在 Matlab 20011b 上完成. 本文使用了文献 [35] 以调用

CPLEX 来求解整数规划问题. 为了解决数据集缺少真实标签的问题, 本文开发了一个 GUI 以进行高效的半自动标注. 跟踪完成后, 使用不同的符号和灰度值将跟踪结果可视化以便于分析结果, 如图 5 和图 6 所示.

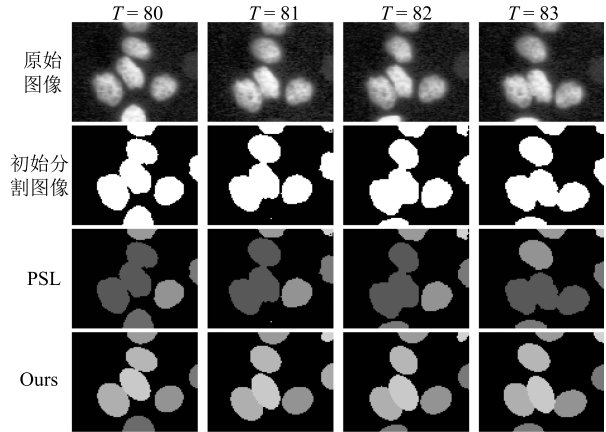


图 5 PSL 和 Ours 在 HeLa-1 数据集上的跟踪结果比较
Fig. 5 Tracking results by PSL and Ours on HeLa-1 datasets

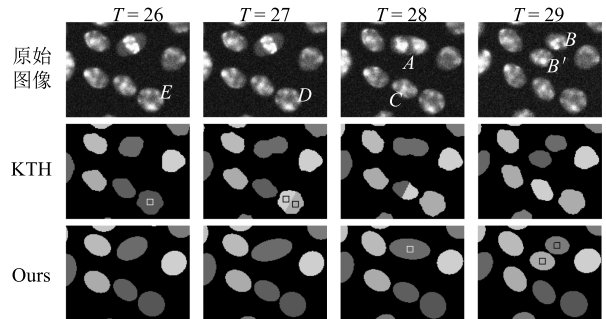


图 6 KTH 和 Ours 在 SIM+-01 数据集上的跟踪结果比较
Fig. 6 Tracking results by KTH and Ours on SIM+-01 datasets

图 5 使用相同灰度值串联一条迁移轨迹. 可以看见, 当原始图像背景噪声较大, 细胞存在密集重叠时, 初始分割结果往往较差, 通常表现为一定程度的欠分割. 此时 PSL 由于将一个前景区域视作一个细胞, 导致出现了较多的跟踪错误. 而 Ours 能够以椭

表 1 数据集的相关统计信息

Table 1 Statistics of dataset used in our study

数据集名称	训练图像序列长度	测试图像序列长度	图像大小	初始分割 F-值 (%)
Fluo-N2DL-HeLa-01	92	92	1 100 × 700	94.2
Fluo-N2DL-HeLa-02	92	92	1 100 × 700	92.7
Fluo-N2DH-SIM+-01	65	130	660 × 718	96.4
Fluo-N2DH-SIM+-02	150	138	664 × 790	95.7

圆假说的形式恢复出隐藏在前景区域中的多个细胞, 因此有效地修正了分割阶段产生的错误, 避免其传递到跟踪过程中.

整个处理过程在台式机上完成, 其中 CPU 为 Intel Core i3-2100 双核, 主频 3.1 GHz, 内存 4 GB, 操作系统为 Windows 7.

图 6 使用相同灰度值代表迁移轨迹, 正方形代表正在进行有丝分裂的母细胞和随后产生的子细胞.

5.2 结构化预测性能

将提出的算法 (Ours) 与当前几种最好的算法进行比较, 使用精度 (Precision)、召回率 (Recall) 和 F-值 (F-measure) 这三个指标对算法检测局部事件的能力进行度量. 表 2 展示了实验结果, 由于对迁移和分裂事件的检测是重建细胞族系的关键, 表 2 中将这两项事件的检测结果详细列出, 同时全部 6 种事件的综合结果在全部事件一栏中给出.

5.3 与网络流规划方法 (NFP) 的比较

将本文算法与网络流规划方法^[27] 进行比较. 在

参数学习阶段, 文献 [27] 在迁移和分裂事件上分别 (局部) 训练梯度提升树 (Gradient boosted tree, GBT) 分类器. 由于文献 [27] 没有公开源码及实现细节, 如参数设置、特征选择等, 本文采用与本文算法相同的特征以及线性支持向量机 (Linear SVM) 分类器来复现文献 [27] 的结果, 命名为 NFP. 另一方面, 文献 [27] 的假说生成方法会产生数量巨大的标签变量 (在 HeLa-2 数据集上其标签变量的数目在 800 万左右), 这个量级超出了本实验所用计算机的处理能力. 因此将文献 [27] 中的假说生成方法替换为第 2.1 节阐述的方法, 这在 HeLa-2 数据集上只产生 45 万个标签变量. 为了更具说服力, 本文将文献 [27] 的实验结果也列在表 2, 命名为 EPFL.

从表 2 可以看出, 本文算法在三个数据集上的表现均好于 NFP 方法, 特别是在对分裂事件的检测上有显著提升. 而 EPFL 尽管使用了相当复杂的局部 GBT 分类器, 在实验表现上还是逊于本文算法. 这种提升得益于本文方法中使用的结构化学习策略, 该策略充分利用了输出空间的结构化信息, 将各种局部事件分类器联合起来进行参数学习, 从而可以

表 2 本文算法与当前几种最好的算法之间的比较 (%)

Table 2 Comparison of our algorithm against state-of-the-art methods (%)

		迁移事件			分裂事件			全部事件		
		精度	召回率	F-值	精度	召回率	F-值	精度	召回率	F-值
HeLa-1	KTH	98.1	96.4	97.2	75.3	74.8	75.1	97.4	95.3	96.3
	NFP	98.9	99.1	99.0	85.1	76.2	80.4	98.1	98.1	98.1
	EPFL	97	99	98	92	79	85	N/A	N/A	N/A
	PSL	97.7	91.8	94.7	82.7	73.4	77.8	94.8	91.6	93.1
	Ours-P	98.2	98.7	98.5	92.1	74.1	82.1	97.2	98.2	97.7
	Ours	99.1	99.3	99.2	88.5	86.0	87.2	97.9	98.2	98.1
HeLa-2	KTH	95.1	98.3	96.7	76.3	77.5	76.9	94.7	97.5	96.1
	NFP	98.0	98.3	98.2	84.5	84.5	84.5	96.1	96.4	96.3
	EPFL	96	99	97	86	83	84	N/A	N/A	N/A
	PSL	96.8	90.7	93.7	73.6	73.6	73.6	94.3	90.1	92.2
	Ours-P	97.3	97.9	97.6	82.8	86.0	84.4	96.1	97.3	96.7
	Ours	97.9	98.6	98.3	92.8	89.9	91.3	97.2	98.0	97.6
SIM+-01	KTH	97.7	98.9	98.3	73.8	51.7	60.8	96.8	97.9	97.3
	NFP	99.1	96.8	98.0	76.1	85.0	80.3	97.3	96.6	97.0
	EPFL	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	PSL	98.9	94.3	96.7	33.3	73.3	45.8	91.7	94.1	92.9
	Ours-P	99.7	96.5	98.1	60.9	93.3	73.7	93.6	96.5	95.0
	Ours	99.8	99.9	99.9	94.6	89.8	92.2	99.7	99.8	99.7
SIM+-02	KTH	94.4	92.6	93.5	80.2	72.7	76.3	93.2	92.1	92.6
	NFP	97.1	90.0	93.4	79.6	81.8	80.7	94.4	91.1	92.7
	EPFL	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	PSL	96.4	94.1	95.2	73.8	84.1	78.6	94.7	93.6	94.1
	Ours-P	96.4	94.0	95.2	82.0	86.4	84.1	95.0	93.6	94.3
	Ours	97.6	93.9	95.7	86.4	90.1	88.2	96.6	93.6	95.1

在理论上确保学习到的参数为全局最优参数; 而局部学习方法为每个事件单独训练一个局部分类器, 最终应用于全局的跟踪时, 不同分类器之间的分类结果可能出现矛盾, 如某细胞即被出现分类器判定为新出现的, 又被迁移分类器判定为迁移而来的, 此时只能选取其判别值大的事件作为最终事件, 而两者的判别值是基于不同的特征空间计算而来, 不具有可比性, 因此往往造成错误. 因此 NFP 和 EPFL 的实验表现与理论分析是一致的.

5.4 与邻帧结构化学习方法 (PSL) 的比较

文献 [21] 提出的 PSL 方法同样基于结构化预测和结构化学习, 然而, 该方法与本文方法具有两点本质性区别: 1) PSL 是一种先检测-再跟踪的方法, 直接使用初始分割图像中的每一个前景区域作为细胞观测, 无法应对图像中潜在的欠分割和过分割, 而二者实际上是无法避免的, 因而该方法无法阻止检测分割阶段产生的错误传递到跟踪过程中. 而本文方法通过构造观测假说的完备集合的方式, 能够处理分割过程中潜在的错误, 并在随后的整数规划寻优过程中找出最佳观测假说轨迹. 2) PSL 中的预测和学习模型都只建立在相邻帧之间. 在学习阶段, PSL 将相邻 2 帧图像作为一个训练样本; 而在预测阶段, PSL 在每 2 帧之间求解一个整数规划问题, 最后将所有结果串联起来得到完整跟踪轨迹. 文献 [21] 还提出了一种主动学习 (Active learning) 的策略以减小对训练样本的标记花费, 但作为代价, 会轻微降低在学习和跟踪上的表现. 为了公平比较, 本文不采用主动学习策略, 而是使用完整的训练集. 另外, PSL 与本文方法使用相同的初始分割结果.

表 2 的结果显示, 本文方法的实验结果要好于 PSL. 原因可能在于, PSL 的表现严重的依赖于检测分割的质量. 笔者注意到在文献 [21] 中, PSL 使用了一个相当精确的分割结果, 其 F-值高达 99.3%. 而本文方法有能力修复检测阶段的错误, 尽管使用了一个极其朴素的分割算法, 也能取得较好的跟踪效果. 图 5 给出了一个典型例子, 其中由于分割的误差导致 PSL 产生了大量的跟踪错误. 为了体现本文方法使用的全局模型的能力, 加入本文方法 Ours 的一个变体 Ours-P, 其中约束条件 (2) 被删去了, 而约束条件 (3) 和 (4) 保留, 并且仿照 PSL, 在相邻 2 帧之间进行学习和预测. 从表 2 中可见, Ours 的表现优于 Ours-P, 因为 Ours 的模型使用了全部图像序列的信息.

5.5 与 KTH 方法的比较

KTH^[36] 中提出的方法在第一届细胞跟踪挑战^[34] 中获得了冠军, 其使用维特比 (Viterbi) 方法来构建细胞轨迹, 而模型的参数则通过训练局部分

类器得到. KTH 公开了可执行文件 (exe), 并附带了为每个数据集分别设置好的最佳参数; 相反, 本文方法不需要为特定数据集单独设置参数.

从表 2 得知, 本文方法表现胜过 KTH. 图 6 中给出了 SIM+-01 数据集上的部分跟踪结果. 可以看到, 本文方法正确地检测到了分裂事件 $A \rightarrow BB'$, 而 KTH 错误地将 B 和 B' 解释成了 2 个迁移事件 $A \rightarrow B$ 和 $C \rightarrow B'$. 此外, KTH 错误地将前景区域 D 和 C 中的细胞个数判断为 2 个; 对于细胞 E , 则将其误检为分裂. 推断这可能与 KTH 中采用的贪心求解策略有关, 其通过最大化一个打分函数的方式每次添加一条轨迹, 虽然每条轨迹是通过动态规划计算得来, 可以保证其单条轨迹在当前评判尺度下的最优性, 但无法保证最后所有的轨迹分配是全局最优的分配方案; 而本文方法则是将轨迹分配转换为整数规划问题, 通过求解整数规划问题同时获得所有的跟踪轨迹, 因而保证了理论上的全局最优性. 另外, 在运行时间上, 以细胞数目最多的 HeLa-2 数据集为例, 由于 KTH 没有给出其训练时间, 因此只比较测试时间: 本文算法大约花费 360 秒来构建矛盾集合及约束条件, 花费 30 秒进行 ILP 的求解, 总共时间大约为 390 秒; KTH 运行测试花费大约 393 秒.

5.6 结构化学习性能

学习过程以一个预先设计好的特征集合和带标注的图像序列作为输入, 最后输出每个特征的权值. 与文献 [21] 中使用割平面法 (Bundle method) 不同, 本文采用了更加高效的 BCFW (Block-coordinate Frank-Wolfe) 学习算法, 此方法可以显著地缩减训练时间和内存需求. BCFW 算法基于单个样本来更新权值, 因此可以以一种在线的方式进行, 甚至有处理能力处理无限数目的训练样本.

5.6.1 特征和对应的权值

结构化学习使大量特征的使用变得可行, 这些复杂的特征有利于从不同层面描述事件, 从而提升跟踪的表现. 本文从先前的文献中借鉴了一些特征, 同时选取了几个新的特征. 由于本文以椭圆作为基本观测对象, 因此使用了椭圆的一些几何特征, 例如大小、偏心率、半长轴的方向等. 其他特征则是从椭圆区域对应的原图上提取得到, 如区域的灰度和/或平均灰度、灰度直方图等.

图 7 展示了从 HeLa-2 数据集上学习得到的特征权值. 权值参数的绝对值代表相应特征的重要性. 各个特征的意义及部分说明如下: 以 diff 开头的代表两个椭圆的特征差异, 其中 position 表示两个椭圆的中心距离, size 表示两个椭圆的面积差, eccentricity 代表偏心率, orientation 代表半长轴方向差

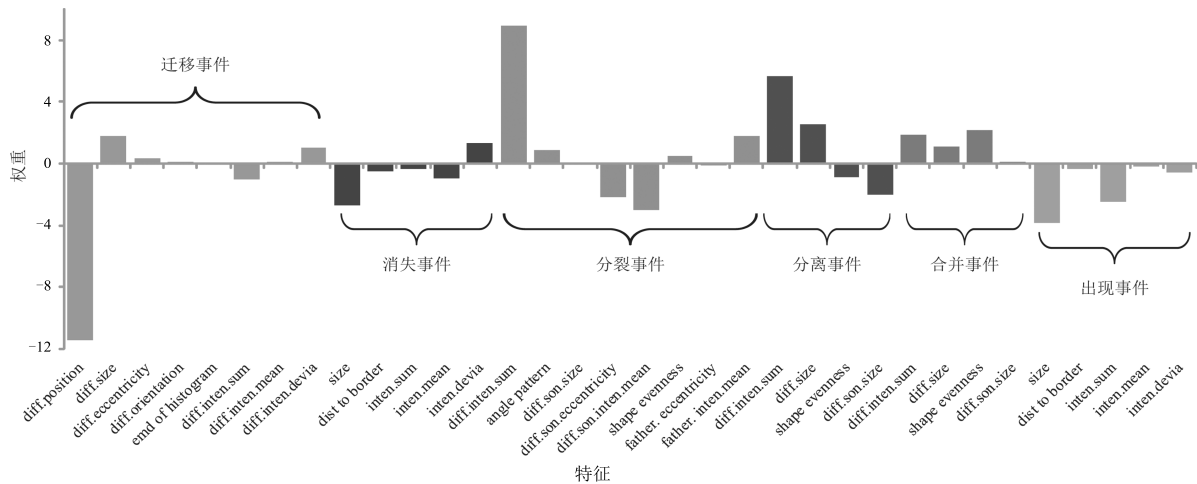


图 7 HeLa-2 数据集上学习得到的特征权值结果

Fig. 7 Learning results on HeLa-2

等; `inten.sum` 代表灰度和, `inten.mean` 代表灰度均值, `inten.devia` 代表灰度标准差; `diff.son.size` 代表两个子细胞的面积差, `diff.son.inten.mean` 代表两个子细胞的灰度均值差; `angle pattern` 则是两个子细胞与母细胞构成的角度大小 (在一次真实的分裂事件中, 这个角度往往趋向于 180°), `dist to border` 代表细胞距图像边缘的距离 (理论上此特征对判定细胞的出现和消失有一定的作用); `emd of histogram` 则是两个灰度直方图之间的地面移动距离 (Earth mover's distance), 用来比较两个直方图的相似度. 学习过程选取了 60 帧训练图像, 并分割为 6 个样本, 每个样本包含 10 帧连续图像. 在训练集上运行 BCFW 算法, 设置最大迭代次数为 200, 从图 7 可以发现, 学习得到的各特征权值与直观感觉是一致的. 例如, 特征 `diff.postion` 代表一个椭圆和其邻近迁移目标之间的欧氏距离, 其在迁移事件的所有 8 个特征中占据了极大的权重, 说明距离是影响迁移事件的主导因素. 为验证此点, 统计了真实轨迹中的所有迁移事件, 结果表明 99% 的迁移事件发生在一个细胞和它下一帧最近邻细胞之间.

表 3 展示了当使用训练集分割时, 拆分方式的不同 (即样本长度和样本数不同) 对训练误差、训练时间、迁移和分裂事件的 F- 值的影响. 表 3 中加粗的部分是表 2 中使用的配置, 在 HeLa-1 数据集中选取 80 帧图像作为训练集, HeLa-2 数据集中选取 60 帧图像作为训练集. 从表 3 可以看出, 使用数目较多但长度较短的训练样本可以显著地缩短训练事件, 但这是以训练误差的增加作为代价, 同时会轻微地降低迁移事件的预测精度. 此外, 减短单个训练样本的长度能够有效地避免求解大规模整数规划问题时遇到的内存不够的问题, 使得算法在较小内存的

个人计算机也能顺利运行. 但当样本长度过小时 (例如 HeLa-1 中的 4 帧长度样本和 HeLa-2 中的 3 帧长度样本), 分裂事件的预测精度会有较为明显的下降. 这与本文观点一致, 即较长样本中包含了更多的全局信息, 其对于预测类似分裂的复杂事件相当关键.

5.7 核化学习的效果

以上实验均是基于线性 S-SVM 学习和预测算法. 第 4.3 节提到, 通过应用核技巧 (Kernel tricks) 可以进一步提升算法表现. 然而, 对全部事件进行核化训练是不切实际的, 因为这需要两两计算所有特征向量间的核函数, 代价过大. 但根据式 (16), 核化算法可以只对那些难以被线性函数判别的特定事件使用. 因此, 在实验中, 本文只对分裂事件使用核化的分类器, 其他事件依然使用线性函数. 表 4 展示了仿真数据集 SIM+-01 上的核化学习算法的表现, 训练过程将 65 帧训练图像序列拆分为 5 个样本, 每个样本包含 13 帧图像.

从表 4 可以看出, 采用适当的核函数和参数值, 相比于线性函数可以提升大约 4% 的分裂精度. 但是注意, 尽管所有的核函数的值已经在训练之前提前计算完成, 非线性核的训练时间仍然为线性的 3 倍左右.

5.8 可选约束

本文方法中可以灵活地添加标签变量的可选约束. 将来自细胞生物学领域的先验知识编码为可选约束的形式, 再加入到学习和预测过程中, 可能会进一步改善跟踪的结果. 在实验中发现, 分裂事件有时会和分离事件混淆, 特别是当这两者特征不具有足够的判别力时. 然而, 可以利用一些先验假设来减少

表 3 HeLa 数据集上训练样本长度对于训练和预测的影响 (%)

Table 3 Effects of training sample length on training and prediction HeLa (%)

数据集	样本数量	单个样本长度	训练时间 (min)	训练误差	迁移事件 F-M	分裂事件 F-M	全部事件 F-M
HeLa-1	1	80	101.5	3.5	99.3	86.8	98.9
	2	40	41.0	2.1	99.2	82.5	98.1
	4	20	25.3	2.0	99.3	87.5	98.9
	8	10	14.6	1.7	99.1	87.7	98.7
	10	8	15.4	10.3	98.6	72.5	98.0
	16	5	12.3	10.6	98.9	77.8	98.3
	20	4	4.4	10.7	98.5	69.7	97.7
HeLa-2	1	80	96.3	4.3	98.1	86.2	97.1
	2	30	28.8	4.8	97.8	82.7	96.7
	4	15	8.9	5.2	97.9	90.4	97.1
	6	10	4.8	4.8	98.3	91.3	97.6
	10	6	4.0	6.0	98.3	89.4	97.5
	15	4	1.8	6.4	98.1	89.7	97.0
	20	3	1.1	7.3	95.9	55.3	93.6

表 4 SIM+-01 上的核化学习效果 (%)

Table 4 Effects of kernelization SIM+-01 (%)

核函数类型	参数	训练时间 (min)	迁移事件			分裂事件		
			精度	召回率	F- 值	精度	召回率	F- 值
RBF	100	5.8	99.8	100	99.9	94.6	88.1	91.2
	10	4.5	99.8	100	99.9	92.9	88.1	90.4
	1	4.5	99.8	100	99.9	98.2	91.5	94.7
	0.1	4.8	99.8	100	99.9	96.4	91.5	93.9
	0.01	4.3	99.8	100	99.9	96.4	91.5	93.9
	0.001	5.0	99.9	100	99.9	98.3	94.9	96.6
Linear		2.0	99.8	99.9	99.9	94.6	89.8	92.2

这种混淆的发生. 例如: 如果 2 个细胞是由前一帧的一个大细胞 (实际上是一个欠分割细胞) 分离得到, 则它们应该在同一个前景区域中, 这种假设可以被编码为如下约束:

当 e_j, e_k 不在同一个前景中时

$$z_{i\alpha}^{\text{split}} = 0, \quad q_\alpha = \{v_j, v_k\} \quad (18)$$

实验中发现, 通过有丝分裂产生的 2 个子细胞通常会将彼此推开, 因此往往位于不同的前景区域中. 而相反, 分离事件的设计目的只是为了解释由严重的欠分割产生的大细胞, 因此当该事件发生时, 生成的 2 个小细胞很有可能位于同个前景中. 从表 5 可以看出, 加入此项可选约束后虚检测 (False positive) 的个数明显降低了, 在一定程度上避免了分裂事件和分离事件的相互混淆.

6 结论

本文提出了一种细胞跟踪的新方法, 在预测和

学习过程中充分利用了轨迹连接的结构信息. 实验结果证明了该方法确实是有效的. 本文方法的主要缺陷在于只能处理规则形状的近似椭圆细胞, 这一点是由预处理步骤中采用椭圆拟合细胞造成的. 但只要能够提供合适的细胞检测假说, 该方法可以应用于任何场合. 未来工作中, 希望能将本文方法扩展到其他形状的细胞跟踪^[36] 问题上, 这一点需要借助于使用其他分割算法如超像素^[37] 代替当前算法实现.

表 5 在 HeLa-1 数据集上加入可选约束的效果

Table 5 Effects of optional constraints on

HeLa-1 dataset

细胞局部事件 (真实发生数目)	迁移		分裂		分离	
	TP	FP	TP	FP	TP	FP
不加可选约束	12 575	115	119	20	1	12
加入可选约束	12 590	119	125	17	1	0

References

- 1 Krzic U, Gunther S, Saunders T E, Streichan S J, Hufnagel L. Multiview light-sheet microscope for rapid in toto imaging. *Nature Methods*, 2012, **9**(7): 730–733
- 2 Tomer R, Khairy K, Amat F, Keller P J. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods*, 2012, **9**(7): 755–763
- 3 Meijering E, Dzyubachyk O, Smal I, van Cappellen W A. Tracking in cell and developmental biology. *Seminars in Cell and Developmental Biology*, 2009, **20**(8): 894–902
- 4 Kanade T, Yin Z Z, Bise R, Huh S, Eom S, Sandbothe M F, Chen M. Cell image analysis: algorithms, system and applications. In: Proceedings of the 2011 IEEE Workshop on Applications of Computer Vision (WACV). Kona, HI: IEEE, 2011. 374–381
- 5 Meijering E, Dzyubachyk O, Smal I. Methods for cell and particle tracking. *Methods in Enzymology*, 2012, **504**: 183–200
- 6 González G, Fusco L, Benmansour F, Fua P, Pertz O, Smith K. Automated quantification of morphodynamics for high-throughput live cell time-lapse datasets. In: Proceedings of the 10th IEEE International Symposium on Biomedical Imaging. San Francisco, CA: IEEE, 2013. 664–667
- 7 Maška M, Ulman V, Svoboda D, Matula P, Matula P, Ederra C, Urbiola A, España T, Venkatesan S, Balak D M W, Karas P, Bolcková T, Štreitová M, Carthel C, Coraluppi S, Harder N, Rohr K, Magnusson K E G, Jaldén J, Blau H M, Dzyubachyk O, Krížek P, Hagen G M, Pastor-Escuredo D, Jimenez-Carretero D, Ledesma-Carbayo M J, Muñoz-Barrutia A, Meijering E, Kozubek M, Ortiz-de-Solorzano C. A benchmark for comparison of cell tracking algorithms. *Bioinformatics*, 2014, **30**(11): 1609–1617
- 8 Li Zhen-Xing, Liu Jin-Mang, Li Song, Bai Dong-Ying, Ni Peng. Group targets tracking algorithm based on box particle filter. *Acta Automatica Sinica*, 2015, **41**(4): 785–798 (李振兴, 刘进忙, 李松, 白东颖, 倪鹏. 基于箱式粒子滤波的群目标跟踪算法. *自动化学报*, 2015, **41**(4): 785–798)
- 9 Yang Feng, Wang Yong-Qi, Liang Yan, Pan Quan. A survey of PHD filter based multi-target tracking. *Acta Automatica Sinica*, 2013, **39**(11): 1944–1956 (杨峰, 王永齐, 梁彦, 潘泉. 基于概率假设密度滤波方法的多目标跟踪技术综述. *自动化学报*, 2013, **39**(11): 1944–1956)
- 10 He Tao, Zhang Jia-Shu, Hua Jian, Gong Xiao-Biao. Multi-target tracking framework for fixed camera. *Application Research of Computers*, 2015, **32**(6): 1892–1896 (何涛, 张家树, 华见, 龚小彪. 基于单固定摄像头的多目标跟踪框架. *计算机应用研究*, 2015, **32**(6): 1892–1896)
- 11 Alexandre D, Vasily S, Shahragim T, Dufour A, Shinin V, Tajbakhsh S, Guillen-Aghion N, Olivo-Marin J C, Zimmer C. Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. *IEEE Transactions on Image Processing*, 2005, **14**(9): 1396–1410
- 12 Li K, Chen M, Kanade T, Miller E D, Weiss L E, Campbell P G. Cell population tracking and lineage construction with spatiotemporal context. *Medical Image Analysis*, 2008, **12**(5): 546–566
- 13 Dzyubachyk O, van Cappellen W A, Essers J, Niessen W J, Meijering E. Advanced level-set-based cell tracking in time-lapse fluorescence microscopy. *IEEE Transactions on Medical Imaging*, 2010, **29**(3): 852–867
- 14 Dufour A, Thibeaux R, Labruyere E, Guillen N, Olivo-Marin J C. 3-D active meshes: fast discrete deformable models for cell tracking in 3-D time-lapse microscopy. *IEEE Transactions on Image Processing*, 2011, **20**(7): 1925–1937
- 15 Delgado-Gonzalo R, Chenouard N, Unser M. Fast parametric snakes for 3D microscopy. In: Proceedings of the 9th IEEE International Symposium on Biomedical Imaging (ISBI). Barcelona, Spain: IEEE, 2012. 852–855
- 16 Maška M, Daněk O, Garasa S, Rouzaut A, Muñoz-Barrutia A, Ortiz-de-Solorzano C. Segmentation and shape tracking of whole fluorescent cells based on the Chan-Vese model. *IEEE Transactions on Medical Imaging*, 2013, **32**(6): 995–1006
- 17 Amat F, Lemon W, Mossing D P, McDole K, Wan Y, Branson K, Myers E W, Keller P J. Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. *Nature Methods*, 2014, **11**(9): 951–958
- 18 Chenouard N, Bloch I, Olivo-Marin J C. Multiple hypothesis tracking for cluttered biological image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(11): 2736–2750
- 19 Li F H, Zhou X B, Ma J W, Wong S T C. Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. *IEEE Transactions on Medical Imaging*, 2010, **29**(1): 96–105
- 20 Kausler B X, Schiegg M, Andres B, Lindner M, Koethe U, Lette H, Wittbrodt J, Hufnagel L, Hamprecht F A. A discrete chain graph model for 3d+t cell tracking with high misdetection robustness. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer, 2012. 144–157
- 21 Luo X H, Schiegg M, Hamprecht F A. Active structured learning for cell tracking: algorithm, framework, and usability. *IEEE Transactions on Medical Imaging*, 2014, **33**(4): 849–860
- 22 Magnusson K E G, Jaldén J. A batch algorithm using iterative application of the Viterbi algorithm to track cells and construct cell lineages. In: Proceedings of the 9th IEEE International Symposium on Biomedical Imaging (ISBI). Barcelona, Spain: IEEE, 2012. 382–385
- 23 Padfield D, Rittscher J, Roysam B. Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis. *Medical Image Analysis*, 2011, **15**(4): 650–668
- 24 Schiegg M, Hanslovsky P, Kausler B X, Hufnagel L, Hamprecht F A. Conservation tracking. In: Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013. 2928–2935
- 25 Schiegg M, Hanslovsky P, Haubold C, Koethe U, Hufnagel L, Hamprecht F A. Graphical model for joint segmentation and tracking of multiple dividing cells. *Bioinformatics*, 2015, **31**(6): 948–956
- 26 Magnusson K E G, Jaldén J, Gilbert P M, Blau H M. Global linking of cell tracks using the viterbi algorithm. *IEEE Transactions on Medical Imaging*, 2015, **34**(4): 911–929

- 27 Türetken E, Wang X C, Becker C J, Haubold C, Fua P. Network flow integer programming to track elliptical cells in time-lapse sequences. *IEEE Transactions on Medical Imaging*, 2017, **PP**(99): 1–1, DOI: 10.1109/TMI.2016.2640859
- 28 Lacoste-Julien S, Jaggi M, Schmidt M, Pletscher P. Block-coordinate Frank-Wolfe optimization for structural SVMs. *Eprint Arxiv*, 2012, 53–61
- 29 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J Y, White D J, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 2012, **9**(7): 676–682
- 30 Kovese P. MATLAB and Octave Functions for Computer Vision and Image Processing [Online], available: <http://www.peterkovese.com/matlabfns/>, November 3, 2015
- 31 Prasad D K, Quek C, Leung M K H, Cho S Y. A parameter independent line fitting method. In: Proceedings of the 1st Asian Conference on Pattern Recognition. Beijing, China: IEEE, 2011. 441–445
- 32 Prasad D K, Leung M K H, Quek C. ElliFit: an unconstrained, non-iterative, least squares based geometric ellipse fitting method. *Pattern Recognition*, 2013, **46**(5): 1449–1465
- 33 Nowozin S, Lampert C H. Structured learning and prediction in computer vision. *Foundations and Trends® in Computer Graphics and Vision*, 2010, **6**(3–4): 185–365
- 34 de Solórzano C O, Kozubek M, Meijering E, Barrutia A M. ISBI Cell Tracking Challenge [Online], available: <http://www.codesolorzano.com/celltrackingchallenge/CellTrackingChallenge/Welcome.html>, November 15, 2015
- 35 Lófberg J. YALMIP: a toolbox for modeling and optimization in MATLAB. In: Proceedings of the 2004 IEEE International Symposium on Computer Aided Control Systems Design. Taipei, China: IEEE, 2004. 284–289
- 36 Bensch R, Ronneberger O. Cell segmentation and tracking in phase contrast images using graph cut with asymmetric boundary costs. In: Proceedings of the 12th IEEE International Symposium on Biomedical Imaging. New York, United States: IEEE, 2015. 1220–1223
- 37 Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(11): 2274–2282



陈旭 北京航空航天大学自动化科学与电气工程学院硕士研究生。主要研究方向为数字图像处理和目标跟踪。

E-mail: xchen0530@163.com

(CHEN Xu Master student at the School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics. His research interest covers digital signal processing and target tracking.)



万九卿 北京航空航天大学自动化科学与电气工程学院副教授。主要研究方向为信号/图像/视频处理, 统计推理与机器学习, 目标检测、跟踪与识别, 复杂系统故障诊断与健康管理。本文通信作者。

E-mail: wanjiuqing@buaa.edu.cn

(WAN Jiu-Qing Associate professor at the School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics. His research interest covers signal/image/video processing, statistical inference, machine learning, target detection, tracking and recognition, and fault diagnosis and prognosis of complex system. Corresponding author of this paper.)