

# 基于图割的图像分割方法及其新进展

刘松涛<sup>1,2</sup> 殷福亮<sup>1</sup>

**摘要** 鉴于图割的理论意义和实际应用价值,系统综述了基于图割的图像分割方法.首先,深入分析了基于图割的图像分割方法的基本原理,主要从定性和定量角度剖析了图割与能量函数最小化之间的关系,然后,概括了基于图割的图像分割方法的基本步骤,包括能量函数的设计、图的构造和最小割/最大流方法,其次,系统梳理和评述了基于图割的图像分割方法的国内外研究现状,最后,指出了基于图割的图像分割方法的发展方向.

**关键词** 图像分割,图割,能量最小化,图论

**引用格式** 刘松涛,殷福亮.基于图割的图像分割方法及其新进展.自动化学报,2012,38(6):911-922

**DOI** 10.3724/SP.J.1004.2012.00911

## The Basic Principle and Its New Advances of Image Segmentation Methods Based on Graph Cuts

LIU Song-Tao<sup>1,2</sup> YIN Fu-Liang<sup>1</sup>

**Abstract** In view of the theoretical significance and practical value of graph cuts, the image segmentation methods based on graph cuts are reviewed in this paper. Firstly, the basic principle of image segmentation method based on graph cuts is analyzed in detail, which mainly focuses on the relation between graph cuts and energy minimization involving both qualitative and quantitative analysis. Secondly, the steps of image segmentation methods based on graph cuts are generalized as designing energy function, constructing graph, and minimum cut/maximum flow approaches. Thirdly, the current status of image segmentation methods based on graph cuts is combed and commented. Finally, the future for these segmentation methods is pointed out.

**Key words** Image segmentation, graph cuts, energy minimization, graph theory

**Citation** Liu Song-Tao, Yin Fu-Liang. The basic principle and its new advances of image segmentation methods based on graph cuts. *Acta Automatica Sinica*, 2012, 38(6): 911-922

图像分割是图像处理与计算机视觉领域的基础性工作.图像分割方法的数量非常多<sup>[1]</sup>,其中,能量最小化方法在过去 30 年引起学者们的广泛关注,形成一大流派.能量最小化方法的基本步骤为<sup>[2]</sup>: 1) 设计一个目标函数(能量函数),其最小值对应最优解.常用的两个约束是数据和先验知识.数据约束限制了理想解应该和真实数据尽量接近;先验约束要求理想解的形式应该和先验知识保持一致; 2) 最小化目标函数.大多数感兴趣的能量函数是非凸

的,有多个极小值,导致多数方法只能找到逼近解,因此,最小化过程通常比较困难.应该说,合适的编码问题的约束形成能量函数和找到一个好的最小化方法是同等重要的,因为任何一步出错都会造成分割的失效.能量最小化方法的显著优点有: 1) 给出了分割的统一框架,用标准优化方法可以求解; 2) 提供了明确的思路来编码要解决问题的约束; 3) 由于最小化能量函数和 MRF (Markov random field) 的最大后验估计是等效的,从贝叶斯统计角度也可以求证能量最小化方法的正确性.能量最小化方法有多种分类方法,包括全局和局部方法、确定性和随机性方法、连续性和离散性方法等.本文采用全局和局部的分类方法,常见的全局能量最小化方法包括模拟退火方法、动态规划方法和图论方法等;局部能量最小化方法包括变分 (Variational) 方法和 ICM (Iterated conditional modes) 方法等.

模拟退火方法代表了一类随机优化方法<sup>[3]</sup>.马尔科夫链蒙特卡罗 (Markov chain Monte Carlo, MCMC) 是这类方法的核心.计算机视觉领域 MCMC 的发展历程为: 1) Metropolis 采样.逐渐降低温度的 Metropolis 采样就是模拟退火;

收稿日期 2011-03-18 录用日期 2011-07-25  
Manuscript received March 18, 2011; accepted July 25, 2011  
中国博士后科学基金 (20100471451), 水下测控技术国家级重点实验室基金 (9140C2603051003) 资助  
Supported by Postdoctoral Science Foundation of China (20100471451), Science and Technology Foundation of State Key Laboratory of Underwater Measurement and Control Technology (9140C2603051003)  
本文责任编辑 章毓晋  
Recommended by Associate Editor ZHANG Yu-Jin  
1. 大连理工大学电子信息与电气工程学部 大连 116024 2. 海军大连舰艇学院信息与通信工程系 大连 116018  
1. Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024 2. Department of Information and Communication Engineering, Dalian Naval Academy, Dalian 116018

2) Metropolis-Hastings 采样. 当状态转移概率  $q$  对称时, 简化为 Metropolis 采样; 3) Gibbs 采样. Metropolis-Hastings 采样的特例, 接受概率永远是 1; 4) S-W 方法<sup>[4]</sup>. 将相邻的磁旋子聚类, 状态更新聚类而非单个磁旋子, 加快了收敛速度. 不足是仅适用于 Ising 和 Potts 模型, 每一个磁旋子的状态数固定且已知, 以及存在外场时收敛慢; 5) DDMCMC (Data driven MCMC) 方法<sup>[5]</sup>. 该方法结合了 Top-down 和 Bottom-up 方法来搜索最优解, 比 MCMC 方法更有效, 但是由于执行分裂和合并时, 要维持区域的拓扑, 难以有大的移动; 6) S-W 割<sup>[6]</sup>. 主要优点为: 从 Metropolis-Hastings 的角度设计, 可以应用于任意概率, 而且状态数可变, 存在外场时, 收敛速度加快, 将分裂和合并结合为单一的过程, 大大增加了 DDMCMC 方法的速度.

动态规划是一个多步决策过程, 通过把一个  $N$  步过程转化为  $N$  个单步过程实现将全局最优转化为局部最优之和<sup>[7]</sup>. 在局部最优的搜索过程中, 标准的动态规划采用 Viterbi 算法进行求解, 耗时较长<sup>[8]</sup>. 若运用启发性信息来指导搜索, 会加快收敛速度, 但不能保证找到最优解. 标准的动态规划方法仅适用于一维能量函数, 不能有效地求解二维能量函数. 实际应用中, 可将二维问题转化为一维问题, 转化思路有两种: 1) 给定初始的包围目标的封闭曲线, 然后离散化曲线, 并对离散点应用动态规划来求解最优路径<sup>[9]</sup>; 2) 在目标轮廓内部找一个点作为极坐标变换的中心点, 然后实施极坐标变换将图像变成一个矩阵, 则目标轮廓检测转化为求解矩阵最左列到最右列的最优路径问题. 该思路中极坐标变换中心点的选择会影响轮廓检测结果, 而且只适合分割星形的目标轮廓<sup>[10]</sup>.

图论方法的主要思想是将图像映射成加权图, 把图像像素看作图的顶点, 邻接像素之间的关系看作图的边, 邻接像素之间的相似性看作边的权值, 根据边的权值设计能量函数, 通过最小化能量函数完成对图的分割, 从而实现图像分割. 基于图论的分割方法对图像进行分割时优点明显: 1) 图论是一门研究比较早而且已经发展成熟的学科, 具有较好的数学基础. 针对某个问题, 图论中有多个方法可以解决; 2) 图像和图之间非常相似. 在图像映射为图之后, 便可以利用图论中的各种理论和数学工具进行图像分割. 目前常用的基于图论的分割方法包括最小割 (通常称为图割)、归一化割等. 和图割相比, 归一化割存在以下几点不足<sup>[11]</sup>: 1) 没有嵌入一元 (Unary) 项, 如各自图结点的先验, 相当于对所有结点都是零先验; 2) 需要计算大矩阵的广义特征向量, 尽管采取了复杂度抑制措施, 计算量仍很大;

3) 只能逼近全局最优解; 4) 倾向于分割出的类具有相同的类内相似度. 因此, 尽管图割容易分离出小目标, 但图割方法的优点更加明显<sup>[12]</sup>: 1) 在全局最优的框架下进行分割, 保证了能量函数的全局最优解; 2) 同时利用了图像的像素灰度信息和区域边界信息, 分割效果好; 3) 用户交互简单且方便, 只需在目标内部和背景区域标记少量的种子点, 对种子点的具体位置也没有严格要求, 而且通过预处理方法自动确定种子点, 也可让图割方法自动化. 因此, 近 10 年来, 基于图割的图像分割方法的发展和应用尤为广泛. 经典的图割只能精确求解特定的能量函数, 而对其他能量函数需要用线性规划松弛方法来求解, 有两个思路可实现: 1) 基于移动制造的方法, 包括  $\alpha$  扩展和  $\alpha-\beta$  交换等, 不试图精确地求解松弛问题; 2) 基于消息传递的方法, 包括信念传播 (Belief propagation) 和 TRW (Tree-reweighted message-passing) 等, 试图精确地求解松弛问题.

变分方法是几何活动轮廓模型 (Geometry active contour model) 的能量最小化方法. 几何活动轮廓模型将演化曲线隐含地表示为更高维曲面 (称为水平集函数), 然后通过变分方法得到水平集函数的演化方程, 当水平集函数停止演化时, 其零水平集对应点集合即为要求的轮廓线或面<sup>[13]</sup>. 该模型有许多优点: 1) 曲线的演化过程始终处于固定网格, 数值实现较为简单; 2) 能够自动处理边界的拓扑变化; 3) 得到光滑、闭合的高精度分割曲线; 4) 自然地将边界信息和区域信息集成在一起. 其主要不足是曲线演化收敛较慢、对初始值敏感. 根据能量函数中的外力项利用图像特性的不同, 几何活动轮廓模型分为基于边界的模型、基于区域的模型和二者的结合. 基于边界的模型包括几何活动轮廓模型<sup>[14]</sup> 和测地线活动轮廓 (Geodesic active contour, GAC) 模型<sup>[15]</sup> 等. 该类模型利用图像梯度信息将初始曲线牵引到目标边界, 依赖于初始化, 且无法提取以非梯度形式定义的边界. 基于区域的模型包括 M-S 模型<sup>[16]</sup> 与 C-V 模型<sup>[17]</sup> 等. 该类模型利用区域信息控制曲线的移动, 通过分片平滑的区域逼近原始图像, 能够解决非梯度定义的边界提取问题, 其中 C-V 模型还具有求解速度较快、与初始曲线位置无关、能够提取内部边界等特点, 但由于 C-V 模型没有考虑边界附近的具体情况, 对渐进型边界的图像, 分割误差较大. 考虑到基于区域的模型可以较好地体现灰度信息, 而基于边界的模型能更好地体现边缘信息. 因此, 两者的结合才是最好的<sup>[18]</sup>.

ICM 方法是基于确定性贪婪策略来寻优的能量最小化方法. 从初始估计的分割结果开始, 对每一个像素, 选择一个可以使能量函数值下降最多的标记

(目标或背景), 并重复该过程直至能量函数收敛到某一稳定值. 该方法的特点是保证收敛、速度快, 但是对初始估计非常敏感, 且易收敛到局部极值点<sup>[19]</sup>.

综上所述, 在用于图像分割的能量最小化方法中, 模拟退火方法的优点是适用能量函数类型广, 不足是计算效率偏低. 动态规划方法仅适用于一维能量函数, 不能有效地求解二维能量函数, 而且耗时较长. 图割只能应用到某类能量函数, 需要扩展应用到更多的能量函数, 但可以得到全局最优解或逼近最优解, 计算效率高. 对任意能量函数, 可使用基于消息传递的方法, 但是这些方法计算速度很慢、不一定收敛、在图割能应用时比图割找到的能量高. 不管是变分方法, 还是 ICM 方法, 都对初始化敏感、易得到局部最小解. 显然, 基于图割的图像分割方法优势更加明显, 从而使其成为近年来的研究热点.

鉴于图割的理论意义和实际应用价值, 本文系统综述了基于图割的图像分割方法. 目的在于让感兴趣的学者能迅速理解和掌握该方法的基本原理, 并能了解它的最新进展, 为进一步深入研究图割理论和拓展其应用范围奠定良好的基础. 具体内容安排为: 第 1 节从定性和定量角度剖析了图割与能量函数最小化之间的关系, 并概括了基于图割的图像分割方法的基本步骤; 第 2 节和第 3 节系统梳理和评述了基于图割的图像分割方法的国内外研究现状; 第 4 节总结全文, 并指出了基于图割的图像分割方法的发展方向.

## 1 基本原理

图像分割可以看作不同的划分问题, 而图可以进行不同的划分, 将图像映射为图后, 分割问题就可以用图论方法(如: 最小割)求解. 这只是直观地解释, 具体如何将图像分割转化为求图的最小割呢? 本文的分析思路是首先将图像分割问题看作能量函数最小化问题, 当能量函数最小时, 实现最优图像分割. 然后, 对给定的能量函数定义一个图, 让图的割集的代价正好等于给定的能量函数. 最后, 通过求图的最小割(即图割), 实现给定能量函数的最小化, 也就实现了最优图像分割. 关于图像分割如何表示为能量函数最小化, 可参考文献[20–21]. 本文重点分析图割与能量函数最小化之间的关系.

### 1.1 图割与能量函数最小化

考虑一个具有端点  $s$  和  $t$  的图  $G = (V, E)$ , 其中  $G$  的每个割集将  $V$  分成两个互不相交的子集  $S$  和  $T$ , 其中  $s \in S, t \in T$ . 对于一个  $n$  维的二值向量  $\mathbf{X} = (x_1, x_2, \dots, x_n)$ , 定义其取值规则为: 若  $v_i \in S$ , 则  $x_i = 0$ ; 若  $v_i \in T$ , 则  $x_i = 1$ . 这样每一个二值向量都可以唯一地对应图  $G$  的一个割集(见

图 1), 而每一个二值向量也可对应一个图像分割结果. 如何求最优图像分割呢? 对于图  $G$ , 可以定义能量函数  $E(x)$ , 它的取值是向量  $(x_1, x_2, \dots, x_n)$  所对应的  $G$  的割集的代价, 而  $G$  的最小割集对应能量函数的最小值. 再令分割的能量函数等价于  $E(x)$ , 由于能量函数最小时, 分割最优. 所以, 最小割集对应的二值向量所确定的分割结果就是最优图像分割结果.

这里有两个问题值得思考: 1) 如何将给定的能量函数转化为图? 2) 为什么图的割集的代价最小值正好等于给定的能量函数最小值?

对第 1 个问题, 假定给定的能量函数为

$$E(x_1, \dots, x_n) = \sum_i E^i(x_i) + \sum_{i < j} E^{i,j}(x_i, x_j) \quad (1)$$

式中,  $E^i(x_i)$  是似然能, 表示当顶点  $v_i$  的标记为  $x_i$  时的代价, 强调数据的一致性;  $E^{i,j}(x_i, x_j)$  是先验能, 表示当相邻顶点  $v_i$  和  $v_j$  的标记为  $x_i$  和  $x_j$  的代价, 强调解的平滑性.

首先, 构造一个图, 包括  $n + 2$  个顶点:  $V = \{s, t, v_1, \dots, v_n\}$ , 其中每一个非末端顶点  $v_i$  对应二值变量  $x_i$ . 然后, 对能量函数中的每一项, 将增加一个或多个边缘, 具体方法如下<sup>[22]</sup>:

1)  $E^i$  只依赖于变量  $x_i$ . 如果  $E^i(0) < E^i(1)$ , 说明顶点  $v_i$  和源点  $s$  更像, 故增加边缘  $(s, v_i)$ , 权重为  $E^i(1) - E^i(0)$ , 见图 2(a); 否则, 说明顶点  $v_i$  和汇点  $t$  更像, 增加边缘  $(v_i, t)$ , 权重为  $E^i(0) - E^i(1)$ , 见图 2(b).

2)  $E^{i,j}$  依赖于两个变量  $x_i$  和  $x_j$ , 见式(2). 式(2)中第 1 项是常数, 不需要增加边缘. 第 2 项和第 3 项均依赖于一个变量, 可用第 1) 步的方法增加边缘. 第 4 项增加一个边缘  $(v_i, v_j)$ , 权重为  $B + C - A - D$ , 见图 2(c). 注意到, 由于能量函数的正则性, 这个权重是非负的. 对  $E^{i,j}$  来说, 完整的图包括三个边缘  $(C - A > 0$  和  $C - D > 0)$ , 见图 2(d).

$$E^{i,j} = \begin{bmatrix} E^{i,j}(0,0) & E^{i,j}(0,1) \\ E^{i,j}(1,0) & E^{i,j}(1,1) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = A +$$

$$\begin{bmatrix} 0 & 0 \\ C-A & C-A \end{bmatrix} + \begin{bmatrix} 0 & D-C \\ 0 & D-C \end{bmatrix} + \begin{bmatrix} 0 & B+C-A-D \\ 0 & 0 \end{bmatrix} \quad (2)$$

上述图的构造过程中, 没有给出边缘权重的确定方法. 本文通过代数变换方法来进一步解释边缘权重如何加和加多少. 首先介绍如何用图表示似然项, 代数变换过程为

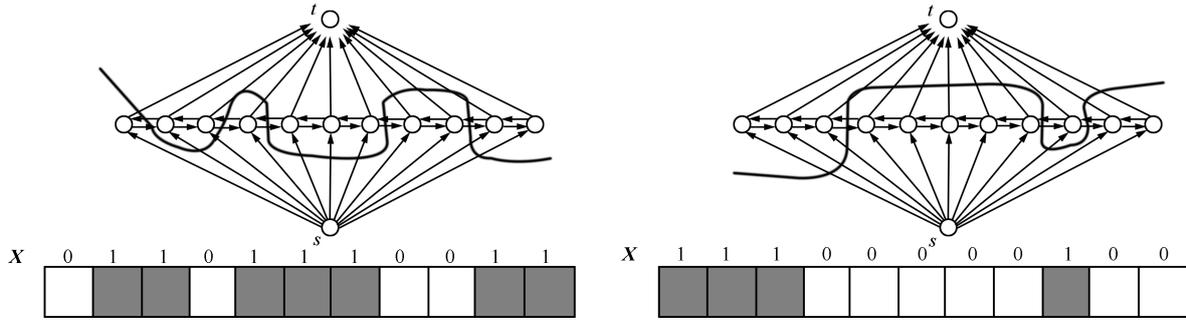


图 1 割集和二值向量  $\mathbf{X}$  之间的对应关系

Fig. 1 The corresponding relation between cut sets and binary vector  $\mathbf{X}$

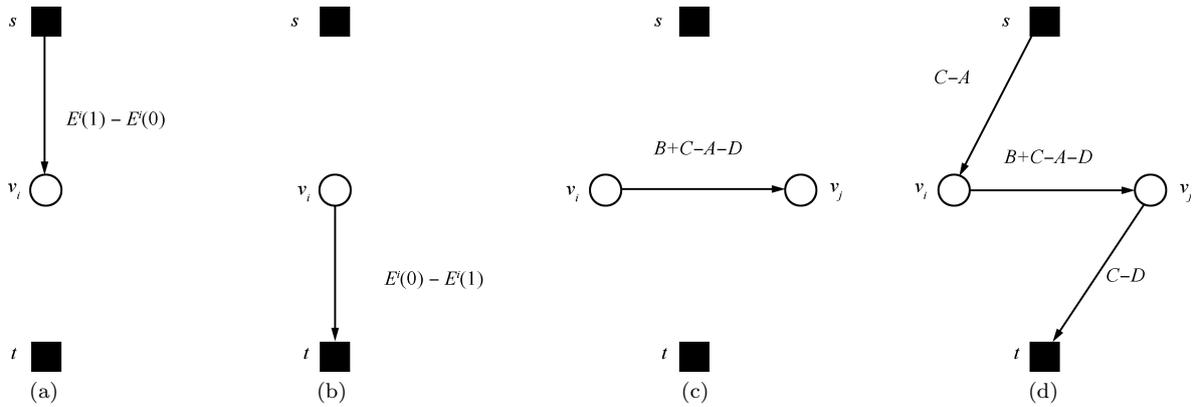


图 2 能量函数与图的对应关系 ((a)  $E^i(E^i(0) < E^i(1))$  的图; (b)  $E^i(E^i(0) \geq E^i(1))$  的图; (c)  $E^{i,j}$  第 3 条边缘的图; (d)  $E^{i,j}$  ( $C > A$  且  $C > D$ ) 完整的图)

Fig. 2 The corresponding relation between energy function and graph ((a) Graph for  $E^i$ , where  $E^i(0) < E^i(1)$ ; (b) Graph for  $E^i$ , where  $E^i(0) \geq E^i(1)$ ; (c) The third edge for  $E^{i,j}$ ; (d) Complete graph for  $E^{i,j}$  if  $C > A$  and  $C > D$ )

$$\begin{aligned} \sum_i E^i(x_i) &= \sum_i [x_i E^i(1) + (1 - x_i) E^i(0)] = \\ &= \sum_i x_i [E^i(1) - E^i(0)] + \sum_i E^i(0) = \\ &= \sum_i x_i \max(0, E^i(1) - E^i(0)) + \\ &= \sum_i (1 - x_i) \max(0, E^i(0) - E^i(1)) + \\ &= \sum_i [E^i(0) - \max(0, E^i(0) - E^i(1))] \end{aligned} \quad (3)$$

上式中最后一项与  $x_i$  无关, 不需考虑. 如果  $E^i(1) - E^i(0) > 0$ , 对顶点  $i$ , 增加一个边缘  $(s, i)$ , 权重为  $E^i(1) - E^i(0)$ ; 否则, 增加一个边缘  $(i, t)$ , 权重为  $E^i(0) - E^i(1)$ .

对平滑项, 代数变换过程如下:

$$\begin{aligned} \sum_{i < j} E^{i,j}(x_i, x_j) &= \\ &= \sum_{i < j} [x_i x_j E^{i,j}(1, 1) + x_i (1 - x_j) E^{i,j}(1, 0) + \\ &= (1 - x_i) x_j E^{i,j}(0, 1) + (1 - x_i) (1 - x_j) E^{i,j}(0, 0)] = \\ &= \frac{1}{2} \sum_{i < j} (x_i - x_j)^2 [E^{i,j}(1, 0) + E^{i,j}(0, 1) - \\ &= E^{i,j}(1, 1) - E^{i,j}(0, 0)] + \frac{1}{2} \sum_{i < j} x_i [E^{i,j}(1, 1) + \\ &= E^{i,j}(1, 0) - E^{i,j}(0, 1) - E^{i,j}(0, 0)] + \\ &= \frac{1}{2} \sum_{i < j} x_j [E^{i,j}(1, 1) + E^{i,j}(0, 1) - E^{i,j}(1, 0) - \\ &= E^{i,j}(0, 0)] + \sum_{i < j} E^{i,j}(0, 0) \end{aligned} \quad (4)$$

显然, 需要在图的边缘中表现前 3 项. 对第 1 项, 在相邻顶点之间加一个边缘, 权重是  $[E^{i,j}(1, 0) + E^{i,j}(0, 1) - E^{i,j}(1, 1) - E^{i,j}(0, 0)]/2$ ; 对第 2 项, 如

果  $E^{i,j}(1,1) + E^{i,j}(1,0) - E^{i,j}(0,1) - E^{i,j}(0,0) > 0$ , 增加一个边缘  $(s, i)$ , 权重为:  $[E^{i,j}(1,1) + E^{i,j}(1,0) - E^{i,j}(0,1) - E^{i,j}(0,0)]/2$ , 否则, 增加一个边缘  $(i, t)$ , 权重为:  $[E^{i,j}(0,1) + E^{i,j}(0,0) - E^{i,j}(1,1) - E^{i,j}(1,0)]/2$ . 对第 3 项, 处理方法和第 2 项相同. 如果某个边缘已经存在, 就更新它的权重. 注意, 这里的权重和文献 [22] 中的权重不同, 所有权重都差了一个数:  $[E^{i,j}(1,0) + E^{i,j}(0,1) - E^{i,j}(1,1) - E^{i,j}(0,0)]/2$ , 这只会影响能量最小化的值, 而不会影响能量最小化的分割结果.

对第 2 个问题, 可表示出构成图的割集的代价, 看是否等价于给定的能量函数. 对给定图, 背景的顶点为:  $B = \{s\} \cup \{v_i : x_i = 0\}$ , 目标的顶点为:  $T = \{t\} \cup \{v_i : x_i = 1\}$ , 则  $B$  中的顶点和  $T$  中的顶点构成的边缘的集合称为割集, 而这些割集的代价的和称为割集的容量, 即:

$$C(\mathbf{x}) = \sum_{k \in B} \sum_{l \in T} c_{kl} \quad (5)$$

从分解的角度,  $C(\mathbf{x})$  可以重写为

$$C(\mathbf{x}) = \sum_{i=1}^n x_i \max(0, -\lambda_i) + \sum_{i=1}^n (1 - x_i) \max(0, \lambda_i) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} (x_i - x_j)^2 \quad (6)$$

其中,  $\lambda_i = E^i(1) - E^i(0)$ ,  $\beta_{ij} = B + C - A - D$ . 注意, 因为  $C(\mathbf{x})$  和  $E(\mathbf{x})$  的差值是常数, 故最小割等价于能量函数最小化.

## 1.2 基于图割的图像分割步骤

从图割可以精确求解的能量函数入手, 可概括出基于图割的图像分割步骤, 主要包括三步: 1) 能量函数的设计; 2) 图的构造; 3) 最小割/最大流方法.

### 1.2.1 能量函数的设计

在满足一定前提条件下, 能量函数最小化可用最小割来精确求解<sup>[22]</sup>. 这个前提条件是: 1) 二值标记; 2) 所有权重是非负的. 这反映到能量函数中, 对数据项函数, 可以是任意的, 因为如果权重是负的, 加一个常数即可. 对平滑项函数要满足子模函数的条件, 比如: 二值 Potts 能. 如果是其他能量函数, 通常只能得到逼近最优解, 比如: 多标记 Potts 能, 可用  $\alpha$  扩展算法或  $\alpha - \beta$  交换算法来逼近求解.

图割模型的核心问题在于如何设计适当的数据项和光滑项, 来得到更加准确的分割结果. 让  $\mathbf{x} = (x_1, \dots, x_p, \dots, x_{|P|})$  为一个二值向量, 其中,

$x_p$  表示分配给图像中像素  $p$  的值 (0 表示背景, 1 表示目标), 向量  $\mathbf{x}$  对应图像分割结果. 常见的能量函数设计如下<sup>[23]</sup>:

$$E(\mathbf{x}) = \sum_{\{p,q\} \in N} V_{\{p,q\}}(x_p, x_q) + \lambda \sum_{p \in P} D_p(x_p) \quad (7)$$

其中, 似然项:  $D_p(1) = -\ln(P(I(p); x_p = 1))$ ,  $D_p(0) = -\ln(P(I(p); x_p = 0))$ , 平滑项:  $V_{\{p,q\}}(x_p, x_q) = B_{\{p,q\}} \cdot T(x_p \neq x_q) = \frac{T(x_p \neq x_q)}{\text{dist}(p,q)} \times \exp(-\frac{(I_p - I_q)^2}{2\sigma^2})$ . 直观地解释为:  $P$  是基于背景或前景种子点估计得到的直方图,  $p$  是顶点,  $x_p$  是标记值,  $x_p = 0$  和  $x_p = 1$  分别对应背景和前景直方图. 将  $p$  的像素值  $I(p)$  代入直方图, 如果代入前景直方图的值大于代入背景直方图的值, 也就是代入前景直方图的能量小于代入背景直方图的能量, 则  $p$  属于前景, 否则, 属于背景. 如果  $x_p = x_q$ ,  $V_{pq} = 0$ , 否则, 当  $|I_p - I_q| < \sigma$  时,  $V_{pq}$  对相似灰度像素之间的不连续性惩罚非常大, 当  $|I_p - I_q| > \sigma$  时, 惩罚比较小.  $\sigma$  可以通过传感器噪声来估计.  $\text{dist}(p, q)$  可以惩罚  $p$  和  $q$  之间距离的大小. 如果在邻域系统内, 该值通常为 1. 系数  $\lambda \geq 0$  表示区域特性和边界特性的相对重要性.

### 1.2.2 图的构造

1) 构建图. 图的顶点和图像的像素或区域对应. 每个顶点有两个边, 连接源 ( $s$ ) 和汇 ( $t$ ), 称为  $t$ -links, 反映了每个标记的偏好程度. 邻域连接  $n$ -links 反映了平滑项, 指示顶点之间的不连续性. 图定义后, 由所有顶点和所有边组成, 即:

$$G = \langle V, E \rangle \quad (8)$$

其中,  $V = P \cup \{s, t\}$ ,  $E = N \cup_{p \in P} \{\{p, s\}, \{p, t\}\}$ .

2) 对图的各个边缘权重进行赋值, 具体见文献 [23].

### 1.2.3 最小割/最大流方法

最小割/最大流方法主要包括两大类<sup>[24]</sup>: 推进重标记 (Push relabel) 方法和增广路径 (Augmenting paths) 方法.

推进重标记方法<sup>[25-26]</sup> 沿着非饱和边缘给一个到汇的距离的低界估计, 然后, 面向具有到汇最小估计距离的顶点来推进剩余的流. 随着推进操作, 边缘逐渐饱和, 距离逐渐增加. 该方法易于并行实现, 通常采用 GPU 加速实现来提高效率.

Ford 和 Fulkerson 的标号方法 (简称 FF 方法) 是基于增广路径的方法<sup>[27]</sup>, 通过标号不断生长一棵树, 直到找不到关于可行流的增广路径为止. FF 方法的计算复杂性和网络的节点数或者边数无关, 而与边上的权值有关. 为了避免求最大流时计算复杂

度依赖于边的权值大小的缺点, Dinic 设计了一种分层算法<sup>[28]</sup>. 为了进一步提高最小割/最大流方法的效率, Boykov 等提出了基于增广路径的新方法<sup>[12]</sup>, 该方法在当前计算机视觉领域的应用最为广泛. 其核心是建立两棵搜索树  $S$  和  $T$ ,  $S$  以源点  $s$  为根,  $T$  以汇点  $t$  为根. 树  $S$  中所有父结点到孩子结点的边都是不饱和的. 树  $S$  中的结点分为“主动 (Active) 结点”和“被动 (Passive) 结点”, 主动结点可以通过从树  $T$  获得新的后代来使得搜索树“生长 (Grow)”, 被动结点不能生长. 算法重复以下三个阶段:

- 1) 生长阶段 (Growth stage): 搜索树  $S$ 、 $T$  生长, 直到找到汇点;
- 2) 扩展阶段 (Augmentation stage): 扩展路径, 搜索树变成森林;
- 3) 收养阶段 (Adoption stage): 收养孤立结点, 恢复搜索树.

## 2 国外研究现状

Greig 等于 20 世纪 80 年代末首次将图割理论应用于图像处理领域<sup>[29]</sup>, 同时他最早提出使用组合优化理论中的最小割/最大流算法来最小化计算机视觉中的能量函数. 20 世纪 90 年代末 Boykov 等<sup>[30]</sup> 将该理论应用到图像分割领域, 提出了基于图割的交互式图像分割方法, 正式引发了国际上对图割理论及其应用的研究热潮.

### 2.1 基于图割的交互式和自动分割方法

与自动分割方法相比, 交互式分割方法有很多优点: 1) 更准确地理解用户的需求; 2) 更准确地提取目标区域.

2001 年, Boykov 等<sup>[23]</sup> 将图割应用到灰度图像前景提取领域. 用户分别指定一部分区域作为前景和背景, 算法自动实现图像分割. 该方法以其简洁的交互方式, 较快的处理速度, 以及能够综合颜色、纹理等各种信息引起了人们的广泛关注.

Blake 等<sup>[31]</sup> 使用高斯混合模型对前景和背景的颜色空间进行建模, 从而能够利用图像的颜色信息, 对彩色图像的分割效果较好. Rother 等<sup>[32]</sup> 提出的 Grab cut 方法采用迭代的图割进行图像分割. Grab cut 总体上仍采用高斯混合模型和图割的框架, 但是在用户的交互方式上更为简洁, 用户只需提供两个点, 形成一个将目标套住的包围盒. 该方法的不足是当包围盒中的背景不能充分表示背景区域时, 分割效果较差. 用户提供包围盒是一种简单和流行的交互方式, 然而, Rother 等仅利用提供包围盒来排除外部区域和初始化能量函数, 文献 [33] 进一步讨论将包围盒作为一个拓扑形状先验, 来阻止过度收缩, 确保用户提供的包围盒以有效的、紧的方式来约束

分割结果.

2004 年, 出现了懒人抠图 (Lazy snapping)<sup>[34]</sup> 技术, 这是一种新颖的能够从粗到细地提取出前景对象的方法. 懒人抠图技术包括两步: 快速的对象标记和简单的边界编辑. 对象标记是在较粗糙的尺度上进行, 用户绘制几条标记线来指定感兴趣的对象; 边界编辑是在较细致的尺度上进行, 用户通过简单的点击或者拖动附着在前景边界上的多边形顶点来编辑对象边界.

图割方法易于分割出小目标, 而测地 (Geodesic) 分割方法可以避免这个不足<sup>[35]</sup>, 但是对种子点的位置非常敏感, 而且由于缺乏边缘建模, 很难精确定位目标的边界. 因此, 文献 [36] 在图割优化框架下结合了测地距离信息和边缘信息, 不仅克服了传统图割方法的缺点, 而且提高了分割精度, 减少了用户交互. 文献 [37] 提出的方法仅需提供一个目标种子点, 就可以实现自动分割, 将交互程度降到了最低.

当然, 图割方法也可完全自动化, 思路是将图割方法和其他方法结合, 比如 PUMA (Phase unwrapping max-flow)<sup>[38]</sup> 方法, 利用相位展开方法获取种子点, 然后用图割进行自动分割; 文献 [39] 利用 SIFT 算子来获取种子点, 从而实现自动分割; 文献 [40] 利用显著性检测算法来自动粗定位目标, 然后用加权核密度估计和图割算法来优化分割.

考虑到分割图像的复杂程度和图像是否有明显的感兴趣特征是影响自动分割的关键因素, 也限制着自动分割的适用范围. 因此, 当前实用的方法仍然是基于图割的交互式分割方法. 对交互方式有两点值得注意, 其一是交互方便; 其二是交互要少, 如参与的次数要少, 每次参与的时间也要少. 总之, 在交互式分割方法中, 既要发挥用户的判断能力, 指导分割的进行, 又要充分利用计算机的运算性能, 使分割方法具有实用性.

### 2.2 能量函数中融入各种先验的图割方法

1) 在能量函数中体现纹理信息. 考虑到马尔科夫随机场模型和条件随机场模型, 只含有单个像素信息 (一阶势) 和成对像素信息 (二阶势), 而一阶势和二阶势并不适合表示纹理信息, Russell 等<sup>[41]</sup> 在能量函数中引入了高阶 Potts 模型. 鉴于结构张量进行纹理分析非常有效, 可有效呈现边缘及其方向, 且运算速度快, Malcolm 等<sup>[42]</sup> 在张量空间中引入了图割理论.

2) 在能量函数中引入形状先验信息. Malcolm 等<sup>[43]</sup> 采用一个训练形状集, 并使用 KPCA (Kernel principal component analysis) 来构建一个统计形状空间, 取得了喜人的效果, 但它的形状能量不是基

于形状测度的,不能处理形状的仿射变换,也不能同时分割多个目标. Veksler<sup>[44]</sup> 采用星形先验信息,没有局限于具体的形状,可部分解决多目标问题,但不能区分目标. Wang 等<sup>[45]</sup> 提出了自适应形状先验方法,主要思想是基于图像灰度信息推导出概率图,让低概率值的像素比高概率值的像素具有更强的形状先验.

3) 在能量函数中加入序约束. 比如目标分割中,标记车轮的像素不可能在标记车顶的像素上面. 加入序约束后,普通  $\alpha$  扩展移动图割会陷入局部最小,文献 [46] 提出同时作用在所有标记上的序保护的移动方法.

4) 在能量函数中体现全局属性. 普通 MRF 仅建模了局部连接,文献 [47] 推导出一个势函数,可以让输出的标记具有连接性. 连接性可以让非均匀目标形成一个整体,也可以滤除干扰杂波.

### 2.3 能量函数中正则参数的自适应确定方法

对整幅图像采用固定的正则参数来进行数据项和平滑项的平衡是不合适的,即使采用最适宜的正则参数,也会导致过分割或欠分割.

文献 [48] 利用 Canny 边缘检测算子来计算每个像素是边缘的概率,并依赖概率值来调整每个像素的正则参数. 为了增强边缘检测方法的鲁棒性,文献 [49] 通过结合图像曲率的局部测度和边缘置信度来自动计算局部正则权重. 上述方法受限于边缘检测方法的性能,文献 [50] 设计了新的能量函数,将区域项并入边界项,没有正则参数,分割结果非常稳定.

### 2.4 非标准图割能量函数的最小化方法

标准图割的能量函数最小化方法仅适用于可度量势的 MRF<sup>[22]</sup>. 对非标准图割的能量函数归纳为以下几类: 无度量势能量函数、多标记能量函数、非子模能量函数和高阶子模能量函数等,下面分别探讨其最小化方法.

1) 无度量势能量函数. Komodakis 等<sup>[51]</sup> 结合线性规划的对偶理论使图割适合无度量势的 MRF 的优化,这样在 MRF 中可以利用更多的先验信息,而且新的 MRF 工作在原始对偶域,每个最大流有较少的增强路径,也可提高 MRF 的计算速度. 基于对偶理论的能量最小化方法依赖于原始和对偶变量的初始化以及原始变量的数目,动态  $\alpha$  扩展方法可以通过缩减未知变量的数目来简化能量函数和重用对偶变量来生成初始化,分割效果优于  $\alpha$  扩展方法<sup>[52]</sup>.

2) 多标记能量函数.  $\alpha$  扩展移动和  $\alpha-\beta$  交换移动是多标记能量函数最小化的基本方法,计算

速度较快,同时能逼近全局能量函数的最小值. 基本  $\alpha$  扩展方法有两个问题: a) 仅适合于正则势的成对 MRF; b) 当每个变量的基数比较大时,迭代次数多,获得局部最小值的风险变大. 问题 a) 的解决办法是扩展图割到非正则势,可以用 QPBO (Quadratic pseudo-binary optimization) 来逼近解决成对 MRF<sup>[53]</sup>, 或者研究特殊的能量函数,比如剪平的凸先验,利用  $\alpha-\beta$  距离移动来最小化截平的凸势能<sup>[54]</sup>. 这两种方法的不足是局限于成对势,不能优化一般的能量函数. 问题 b) 的解决办法是通过缩减变量来减小数值空间,基于融合移动和逼近推断来改进  $\alpha-\beta$  交换和  $\alpha$  扩展算法. 多标记能量函数中到底有多少标记合适? 文献 [55] 提出了未知标记数量的 MRF 的优化方法,具体是采用划分合并策略将复杂问题分解为一系列子问题,基于聚类运算得到子最优标记. 新方法可以自动确定标记数目,平衡标记精度和空间一致性,同时计算效率较高.

3) 非子模能量函数. 传统图割能以多项式时间精确求解子模能量函数,而最小化非子模能量函数是 NP 困难的<sup>[56]</sup>. 通常采用的方法有: a) 线性规划松弛,比如 Roof-dual 松弛; b) 最小化一个相关的子模函数.

4) 高阶子模能量函数. 高阶模型有很多优点: a) 允许编码的先验多; b) 能捕获多个目标间的交互,更好地反映目标特征. 但是高阶 MRF 的优化非常困难,计算量大、缺乏有效的优化算法. 随着最近新的能量最小化方法的成功应用,掀起了新的高潮来寻找有效的优化高阶 MRF 的方法<sup>[57]</sup>. Ishikawa<sup>[58]</sup> 引入一个变换,将二值标记的高阶 MRF 变换为一阶 MRF,两者具有相同的最小值,同时结合缩减、融合移动和 QPBO 算法,建立了一个高阶多值 MRF 能量函数最小化的逼近框架<sup>[59]</sup>.

### 2.5 图割方法实时性的提高措施

常用的措施是超像素方法、多分辨率方法、并行实现方法和重用方法. 直接对像素进行计算,效率不高,且有很多冗余. 超像素表示通过基于灰度相似性合并邻域,减少了顶点数,因此计算时间大大减少,而且易于保留边界,获得精确的分割结果. 常见的超像素方法有: Mean shift<sup>[60]</sup>, Watershed<sup>[34]</sup> 和 Quick shift<sup>[61]</sup> 等.

多分辨率方法将计算时间几乎降低到线性复杂度<sup>[62]</sup>,但存在三个主要的难点: 1) 低分辨率导致分割效果差; 2) 低分辨率时,细小目标丢失; 3) 需要防止邻近的种子点在低分辨率时并入到相同的区域. 另外,多分辨率方法还有一些固有的缺点,比如: 底层金字塔上的图割处理必须等高层金字塔图割处理完成,因而串行执行并没能有效地提高计算速度. 超

像素方法和多分辨率方法的核心都是通过减少图的顶点数来提高速度,图割算法本身的优缺点并没有变化,而且粗化算子引入了不确定性,不再保证最终分割为最小割。

并行实现方法通过划分图为多个子图,然后以并行方式迭代求解子问题来得到全局最优解。具体实现时能以多核 GPU 方式实现,也能以多计算机方式并行实现,解决图太大时导致计算机内存无法存储的难题。Strandmark 等<sup>[63]</sup>通过对偶分解将图划分为多个子图,用并行方式解决最大流/最小割问题,既保证了解的最优性,又提高了图割的计算效率。

重用方法适用于提高序列图像分割的实时性,具体措施包括重用树、重用流<sup>[64]</sup>和重用割<sup>[65]</sup>等。

### 2.6 基于图割的序列图像分割方法

文献 [66] 将序列图像中提取目标的任务转化为超图割 (Hypergraph cut) 问题。对序列图像中每一帧都进行过分割,然后将分割区域作为图的顶点,建立超图来表示这些顶点之间复杂的时空邻域关系。

传统的交互式分割方法因为频繁的交互,不能用于视频分割。文献 [67] 提出了一个基于 MRF 的轮廓跟踪和基于图割的图像分割的视频目标分割方法。轮廓跟踪在序列图像间传播目标的形状,而图割可以优化形状和改善视频分割精度。该方法需要较少的关键帧和用户交互。为了进一步减少交互,文献 [68] 将多标记图割应用于连续帧图像,数据项和平滑项在时间轴上增量更新和传播,优点在于: 1) 面向区域内的运动流而不是跟踪区域边界的二维窗口来传播数据和标记先验; 2) 不需要用户交互。

## 3 国内研究情况

和国际上的研究热潮相比,国内对图割的深入研究较少,而对图谱聚类研究较多,比如:在降低谱聚类算法的计算量方面,陶文兵等<sup>[69]</sup>采用均值漂移方法,马秀丽等<sup>[70]</sup>基于分水岭算法都使图的顶点数大大减少,取得了较好的实时性。闫成新等<sup>[71]</sup>首次综述了基于图论的图像分割方法,重点介绍了图谱聚类,但在最优剪切准则设计部分提到了最小割准则,并明确指出基于图论的图像分割方法涉及的理论知识较多,应用尚处在初级阶段。当前对基于图割的图像分割方法的主要研究有:

1) 在基于图割的交互式分割方面,刘嘉等<sup>[72]</sup>以文献 [23] 的工作为基础,提出了一种新的交互式图像分割算法,改进之处有: a) 在贝叶斯框架下重新建立模型,加入纹理和色彩特征; b) 提出用 Fisher 判别准则对 Texton 直方图进行降维,使其适用于交互式应用; c) 在再编辑阶段用新加种子点重新估计分割模型,以弥补初始模型的不准确性。韩守东

等<sup>[73]</sup>使用高斯超像素来构建 Graph cuts 模型以实现加速,其主要特色是使用分量形式的期望最大化混合高斯算法对用户交互进行聚类,并用高斯统计对超像素进行特征描述。该课题组又将多级带状闭合形状和 Grab cut 算法结合,在分割精度、计算效率和内存使用方面优于原始的 Grab cut 算法<sup>[74]</sup>。

2) 在能量函数设计方面,田铮课题组设计了一个基于加权割的图像分割算法<sup>[75]</sup>,其特点是整合了图像的局部和整体分割信息,能同时实现类间最大相异性和类内最大一致性。该课题组又从研究图割模型与加权核 K 均值之间的等价性入手,设计了一个基于图的加权核 K 均值的图像多尺度分割方法,既避免了利用图割模型进行图像分割时图谱的求解问题,又避免了利用加权核 K 均值方法进行数据聚类时核矩阵的确定问题<sup>[76]</sup>。考虑到连通性是反映图像相关性的一种自然的假设,文献 [77] 在能量函数中引入了具有自适应权重调节功能的连通性约束项,有效地改善了 Shrinking bias 现象,提高了分割结果的精确性。鉴于目标形状信息对分割的指导作用,唐鹏等<sup>[78]</sup>将动态形状模型引入到能量函数中,再用图割获取全局最优分割。与一般的形状先验不同,文献 [79] 提出了图割优化框架下的泛化形状先验。泛化形状可以从抽象形状到具体形状的任意过渡与组合,用户不需要考虑输入的具体形式,从而放宽了对交互的约束,降低了交互负担。

3) 在实时设计实现方面,考虑到 GPU 的不足在于动态随机访问存储器严格的存储模式和固定的流水管线,而统一计算设备架构 (Compute unified device architecture, CUDA) 技术提供了更便捷的 GPU 编程及存储模式,文献 [80] 基于 CUDA 的设计思想和编程方式,对图割算法进行了并行改造,在 CUDA 架构下实现了并行图割运算,在保证分割效果的同时,提高了图像处理速度。

## 4 结束语

MRF 进行图像分割时最大后验估计就得到了图像分割结果,而 MRF 最大后验估计等效于离散能量最小化。离散能量最小化方法包括随机性方法和确定性方法,其中,图割是非常重要的一种确定性离散能量最小化方法,被广泛应用于图像分割领域。本文重点阐述了基于图割的图像分割方法的基本原理,主要特色是从代数变换角度给出了边缘权重的确定方法,而且论证了图的割集的代价最小值等于给定能量函数的最小值,并将基于图割的图像分割方法概括为三步:能量函数的设计、图的构造和最小割/最大流方法,然后系统梳理和评述了基于图割的图像分割方法的国内外现状,最后,除了前面总结的

研究热点, 作者认为, 以下几个研究方向值得继续关注:

1) 克服现有图割方法的不足. 如: a) 现有图割在融入形状先验信息时, 要么是一般的形状, 要么是人工提供的形状模板, 不能以无监督方式挖掘目标. 可考虑采用平均形状方法<sup>[81]</sup>, 让模型适应真实图像, 不需要人工定义目标形状就可以分割目标; b) 将重计算这种加快计算速度的方法应用到非子模函数的最小化中; c) 对非子模函数, 目前没有方法能以多项式时间优化, 而且获得局部极小值, 没有逼近保证, 可研究更紧逼近界限的有效算法; d) 当前高阶势的图割优化限制在子模高阶势函数, 需要研究精确的一般高阶势的有效最小化逼近算法.

2) 扩展图割方法的分割对象. 比如: 多视图像分割. 当前大多数分割方法是单视图像分割, 然而, 随着采集设备的性能提升, 可以很容易建立多视采集系统, 这促进了多视技术及其应用的发展. 感兴趣区域的有效分割在多视技术应用中非常重要<sup>[82-83]</sup>, 比如基于图像的渲染、三维目标模型重建等. 在基于图像的渲染中, 用户可能更希望只渲染感兴趣区域, 而不是整个场景, 这就要求精确地分割目标. 在三维模型重建中, 首要任务是如何有效分割感兴趣目标, 然后, 才能集成从不同角度采集的二维图像来重构三维目标模型.

3) 研究图割和动态规划、模拟退火、贝叶斯网络、信念传播、水平集等优化方法之间的关系. 比如: 水平集和图割是完全不同的图像分割方法, 文献<sup>[84]</sup>通过核技术在这两类方法之间建立了等价关系, 特别是 C-V 函数的核化, 正好为平均互联割的能量函数. 活动轮廓模型通常在连续水平集框架下构成, 利用梯度下降方法进行优化, 文献<sup>[85]</sup>在离散域重新构建了 C-V 模型, 并分析了离散能量函数的图可表示性和子模性, 然后利用最大流/最小割算法进行优化. 传统离散图割方法存在度量误差, 而连续图割方法不仅度量误差小, 而且可以并行执行, 文献<sup>[86]</sup>使连续图割方法可以优化的函数包括了各向同性 TV-norms, 并将  $\alpha$  扩展算法推广到连续域.

## References

- 1 Pal N R, Pal S K. A review on image segmentation techniques. *Pattern Recognition*, 1993, **26**(9): 1277-1294
- 2 Veksler O. Efficient Graph-based Energy Minimization Methods in Computer Vision [Ph.D. dissertation], Cornell University, USA, 1999
- 3 Bhandarkar S M, Zhang H. A comparison of stochastic optimization techniques for image segmentation. *International Journal of Intelligent Systems*, 2000, **15**(5): 441-476
- 4 Wang J S, Swendsen R H. Cluster Monte Carlo algorithms. *Physica A: Statistical Mechanics and Its Applications*, 1990, **167**(3): 565-578
- 5 Tu Z W, Zhu S C. Image segmentation by data-driven Markov chain Monte Carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 657-673
- 6 Barbu A, Zhu S C. Generalizing Swendsen-Wang to sampling arbitrary posterior probabilities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(8): 1239-1253
- 7 Martelli A. An application of heuristic search methods to edge and contour detection. *Communications of the ACM*, 1976, **19**(2): 73-83
- 8 Asano T, Chen D, Katoh N, Tokuyama T. Efficient algorithms for optimization-based image segmentation. *International Journal of Computational Geometry and Applications*, 2001, **11**(2): 145-166
- 9 Mortensen E, Morse B, Barrett W, Udupa J. Adaptive boundary detection using "live-wire" two-dimensional dynamic programming. In: *Proceedings of the Computers in Cardiology*. Durham, USA: IEEE, 1992. 635-638
- 10 Cheng D C, Jiang X. Detections of arterial wall in sonographic artery images using dual dynamic programming. *IEEE Transactions on Information Technology in Biomedicine*, 2008, **12**(6): 792-799
- 11 Ghanem B, Ahuja N. Dinkelbach NCUT: an efficient framework for solving normalized cuts problems with priors and convex constraints. *International Journal of Computer Vision*, 2010, **89**(1): 40-55
- 12 Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(11): 1222-1239
- 13 Zhao H K, Chan T, Merriman B, Osher S. A variational level sets approach to multiphase motion. *Journal of Computational Physics*, 1996, **127**(1): 179-195
- 14 Caselles V, Catte T, Coll T, Dibos F. A geometric model for active contours in image processing. *Numerische Mathematik*, 1993, **66**(1): 1-31
- 15 Caselles V, Kimmel R, Sapiro G. Geodesic active contours. *International Journal of Computer Vision*, 1997, **22**(1): 61-79
- 16 Mumford D, Shah J. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 1989, **42**(5): 577-685
- 17 Chan T, Vese L. Active contours without edges. *IEEE Transactions on Image Processing*, 2001, **10**(2): 266-277
- 18 Paragios N, Deriche R. Geodesic active regions and level set methods for motion estimation and tracking. *Computer Vision and Image Understanding*, 2005, **97**(3): 259-282
- 19 Besag J. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, Series B*, 1986, **48**(3): 259-302

- 20 Stuart G, Donald G. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, **6**(6): 721–741
- 21 Besag J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 1974, **36**(2): 192–236
- 22 Kolmogorov V, Zabini R. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(2): 147–159
- 23 Boykov Y, Jolly M P. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001. 105–112
- 24 Boykov Y, Veksler O. Graph cuts in vision and graphics: theories and applications. *Handbook of Mathematical Models in Computer Vision*. New York: Springer, 2006. 79–96
- 25 Cherkassky B V, Goldberg A V. On implementing the push-relabel method for the maximum flow problem. *Algorithmica*, 1997, **19**(4): 390–410
- 26 Goldberg A V, Tarjan R E. A new approach to the maximum-flow problem. *Journal of the ACM*, 1988, **35**(4): 921–940
- 27 Ford L, Fulkerson D. *Flows in Networks*. Princeton: Princeton University Press, 1962
- 28 Dinic E A. Algorithm for solution of a problem of maximum flow in networks with power estimation. *Soviet Mathematics-Doklady*, 1970, **11**(5): 1277–1280
- 29 Greig D, Porteous B, Seheult A. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society, Series B*, 1989, **51**(2): 271–279
- 30 Boykov Y, Jolly M P. Interactive organ segmentation using graph cuts. In: Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention. Pittsburgh, USA: Springer, 2000. 276–286
- 31 Blake A, Rother C, Brown M, Perez P, Torr P. Interactive image segmentation using an adaptive GMMRF model. In: Proceedings of the 8th European Conference on Computer Vision. Prague, Czech Republic: Springer, 2004. 428–441
- 32 Rother C, Kolmogorov V, Blake A. “GrabCut”: interactive foreground extraction using iterated graph cuts. In: Proceedings of the ACM SIGGRAPH Conference. Los Angeles, USA: ACM, 2004. 309–314
- 33 Lempitsky V, Kohli P, Rother C, Sharp T. Image segmentation with a bounding box prior. In: Proceedings of the 12th IEEE International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 277–284
- 34 Li Y, Sun J, Tang C K, Shum H Y. Lazy snapping. In: Proceedings of the ACM SIGGRAPH Conference. Los Angeles, USA: ACM, 2004. 303–308
- 35 Bai X, Sapiro G. A geodesic framework for fast interactive image and video segmentation and matting. In: Proceedings of the 11th IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil: IEEE, 2007. 1–8
- 36 Price B L, Morse B, Cohen S. Geodesic graph cut for interactive image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 3161–3168
- 37 Das P, Veksler O, Zavadsky V, Boykov Y. Semiautomatic segmentation with compact shape prior. *Image and Vision Computing*, 2009, **27**(1–2): 206–219
- 38 Bioucas-Dias J, Valadao G. Phase unwrapping via graph cuts. *IEEE Transactions on Image Processing*, 2007, **16**(3): 698–709
- 39 Suga A, Fukuda K, Takiguchi T, Ariki Y. Object recognition and segmentation using SIFT and graph cuts. In: Proceedings of the 19th International Conference on Pattern Recognition. Tampa, USA: IEEE, 2008. 1–4
- 40 Tang Z, Miao Z J, Wan Y L, Li J. Automatic foreground extraction for images and videos. In: Proceedings of the 17th IEEE International Conference on Image Processing. Hong Kong, China: IEEE, 2010. 2993–2996
- 41 Russell C, Restif C, Metaxas D, Torr P. Using the  $P^n$  Potts model with learning methods to segment live cell images. In: Proceedings of the 11th IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil: IEEE, 2007. 1–8
- 42 Malcolm J, Rathi Y, Tannenbaum A. A graph cut approach to image segmentation in tensor space. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Minnesota, USA: IEEE, 2007. 1–8
- 43 Malcolm J, Rathi Y, Tannenbaum A. Graph cut segmentation with nonlinear shape priors. In: Proceedings of the IEEE International Conference on Image Processing. San Antonio, USA: IEEE, 2007. 365–368
- 44 Veksler O. Star shape prior for graph-cut image segmentation. In: Proceedings of the 10th European Conference on Computer Vision. Marseille, France: Springer, 2008. 454–467
- 45 Wang H, Zhang H. Adaptive shape prior in graph cut segmentation. In: Proceedings of the 17th IEEE International Conference on Image Processing. Hong Kong, China: IEEE, 2010. 3029–3032
- 46 Liu X Q, Veksler O, Samarabandu J. Order preserving moves for graph-cut-based optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(7): 1182–1196
- 47 Nowozin S, Lampert C H. Global connectivity potentials for random field models. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 818–825
- 48 Candemir S, Akgul Y S. Adaptive regularization parameter for graph cut segmentation. In: Proceedings of the 7th International Conference on Image Analysis and Recognition. Povo de Varzim, Portugal: Springer, 2010. 117–126

- 49 Rao J, Abugharbieh R, Hamarneh G. Adaptive regularization for image segmentation using local image curvature cues. In: Proceedings of the 11th European Conference on Computer Vision. Crete, Greece: Springer, 2010. 651–665
- 50 Le T H, Jung S W, Choi K S, Ko S J. Image segmentation based on modified graph-cut algorithm. *Electronics Letters*, 2010, **46**(16): 1121–1122
- 51 Komodakis N, Tziritas G. Approximate labeling via graph cuts based on linear programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(8): 1436–1453
- 52 Alahari K, Kohli P, Torr P H S. Dynamic hybrid algorithms for MAP inference in discrete MRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(10): 1846–1857
- 53 Lempitsky V, Rother C, Roth S, Blake A. Fusion moves for Markov random field optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(8): 1392–1405
- 54 Veksler O. Multi-label moves for MRFs with truncated convex priors. In: Proceedings of the 7th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition. Bonn, German: Springer, 2009. 1–13
- 55 Feng W, Jia J Y, Liu Z Q. Self-validated labeling of Markov random fields for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(10): 1871–1887
- 56 Kolmogorov V, Rother C. Minimizing nonsubmodular functions with graph cuts — a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(7): 1274–1279
- 57 Ishikawa H. Higher-order clique reduction in binary graph cut. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 2993–3000
- 58 Ishikawa H. Transformation of general binary MRF minimization to the first-order case. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(6): 1234–1249
- 59 Ishikawa H. Higher-order gradient descent by fusion-move graph cut. In: Proceedings of the 20th IEEE International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 568–574
- 60 Ning J F, Zhang L, Zhang D, Wu C K. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 2010, **43**(2): 445–456
- 61 Alper A, Stefano S. Motion segmentation with occlusions on the superpixel graph. In: Proceedings of the 12th IEEE International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 727–734
- 62 Lombaert H, Sun Y, Grady L, Xu C Y. A multilevel banded graph cuts method for fast image segmentation. In: Proceedings of the 10th IEEE International Conference on Computer Vision. Beijing, China: IEEE, 2005. 259–265
- 63 Strandmark P, Kahl F. Parallel and distributed graph cuts by dual decomposition. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 2085–2092
- 64 Kohli P, Torr P H S. Dynamic graph cuts for efficient inference in markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(12): 2079–2088
- 65 Juan O, Boykov Y. Active graph cuts. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2006. 1023–1029
- 66 Huang Y C, Liu Q S, Metaxas D. Video object segmentation by hypergraph cut. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 1738–1745
- 67 Chung C Y, Chen H H. Video object extraction via MRF-based contour tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, **20**(1): 149–155
- 68 Wang T H, Guillemaut J Y, Collomosse J. Multi-label propagation for coherent video segmentation and artistic stylization. In: Proceedings of the 17th IEEE International Conference on Image Processing. Hong Kong, China: IEEE, 2010. 3005–3008
- 69 Tao Wen-Bing, Jin Hai. A new image thresholding method based on graph spectral theory. *Chinese Journal of Computers*, 2007, **30**(1): 110–119  
(陶文兵, 金海. 一种新的基于图谱理论的图像阈值分割方法. 计算机学报, 2007, **30**(1): 110–119)
- 70 Ma Xiu-Li, Jiao Li-Cheng. SAR image segmentation based on watershed and spectral clustering. *Journal of Infrared and Millimeter Waves*, 2008, **27**(6): 452–456  
(马秀丽, 焦李成. 基于分水岭 — 谱聚类的 SAR 图像分割. 红外与毫米波学报, 2008, **27**(6): 452–456)
- 71 Yan Cheng-Xin, Sang Nong, Zhang Tian-Xu. Survey on graph theory based image segmentation technique. *Computer Engineering and Applications*, 2006, **42**(5): 11–14  
(闫成新, 桑农, 张天序. 基于图论的图像分割研究进展. 计算机工程与应用, 2006, **42**(5): 11–14)
- 72 Liu Jia, Wang Hong-Qi. A graph cuts based interactive image segmentation method. *Journal of Electronics and Information Technology*, 2008, **30**(8): 1973–1976  
(刘嘉, 王宏琦. 一种基于图割的交互式图像分割方法. 电子与信息学报, 2008, **30**(8): 1973–1976)
- 73 Han Shou-Dong, Zhao Yong, Tao Wen-Bing, Sang Nong. Gaussian super-pixel based fast image segmentation using graph cuts. *Acta Automatica Sinica*, 2011, **37**(1): 11–20  
(韩守东, 赵勇, 陶文兵, 桑农. 基于高斯超像素的快速 Graph Cuts 图像分割方法. 自动化学报, 2011, **37**(1): 11–20)
- 74 Han S D, Tao W B, Wu X L, Tai X C, Wang T J. Fast image segmentation based on multilevel banded closed-form method. *Pattern Recognition Letters*, 2010, **31**(3): 216–225

- 75 Li Xiao-Bin, Tian Zheng, Liu Mi-Ge, Xu Hai-Xia. Weighted cut based image segmentation. *Acta Electronica Sinica*, 2008, **36**(1): 76–80  
(李小斌, 田铮, 刘密歌, 徐海霞. 基于加权割的图像分割. 电子学报, 2008, **36**(1): 76–80)
- 76 Li Yu-Chuan, Tian Zheng. Multiscale image segmentation based on graph weighted kernel K-means. *Acta Optica Sinica*, 2009, **29**(10): 2762–2767  
(李昱川, 田铮. 基于图的加权核 K 均值的图像多尺度分割. 光学学报, 2009, **29**(10): 2762–2767)
- 77 Zheng Jia-Ming, Chen Zhao-Jiong. Connectivity constrained graph-cut for fast interactive image segmentation. *Journal of Computer-Aided Design and Computer Graphics*, 2011, **23**(3): 399–405  
(郑加明, 陈昭炯. 带连通性约束的快速交互式 Graph-Cut 算法. 计算机辅助设计与图形学学报, 2011, **23**(3): 399–405)
- 78 Tang Peng, Gao Lin, Sheng Peng. A novel dynamic shape based infrared object subtraction method. *Journal of Optoelectronics · Laser*, 2009, **20**(8): 1049–1052  
(唐鹏, 高琳, 盛鹏. 基于动态形状的红外目标提取算法. 光电子·激光, 2009, **20**(8): 1049–1052)
- 79 Liu Chen, Li Feng-Xia, Zhang Yan. An interactive object cutout algorithm based on graph-cut and generalized shape prior. *Journal of Computer-Aided Design and Computer Graphics*, 2009, **21**(12): 1753–1760  
(刘陈, 李凤霞, 张艳. 基于图割与泛形信息的对象分割方法. 计算机辅助设计与图形学学报, 2009, **21**(12): 1753–1760)
- 80 Li Shuan-Qiang, Feng Qian-Jin. Liver tumor segmentation using graph cuts based on CUDA. *Chinese Journal of Biomedical Engineering*, 2010, **29**(5): 641–647  
(李拴强, 冯前进. 统一计算设备架构并行图割算法用于肝脏肿瘤图像分割. 中国生物医学工程学报, 2010, **29**(5): 641–647)
- 81 Jiang T T, Jurie F, Schmid C. Learning shape prior models for object matching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 848–855
- 82 Campbell N, Vogiatzis G, Hernandez C, Cipolla R. Automatic 3D object segmentation in multiple views using volumetric graph-cuts. In: Proceedings of the British Machine Vision Conference. Warwick, UK: BMVA, 2007. 530–539
- 83 Zhang Q, Ngan K N. Multi-view video based multiple objects segmentation using graph cut and spatiotemporal projections. *Journal of Visual Communication and Image Representation*, 2010, **21**(5–6): 453–461
- 84 Tan S, Kakadiaris I A. Kernel active contour. In: Proceedings of the 12th IEEE International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 521–528

- 85 El-Zehiry N, Sahoo P, Elmaghraby A. Combinatorial optimization of the piecewise constant Mumford-Shah functional with application to scalar/vector valued and volumetric image segmentation. *Image and Vision Computing*, 2011, **29**(6): 365–381
- 86 Olsson C, Byrod M, Overgaard N C, Kahl F. Extending continuous cuts: anisotropic metrics and expansion moves. In: Proceedings of the 12th IEEE International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 405–412



**刘松涛** 海军大连舰艇学院信息与通信工程系副教授. 大连理工大学信息与通信工程博士后流动站出站博士后. 2000 年获得海军航空工程学院航空雷达专业学士学位, 并分别于 2003 年和 2006 年获得该校信号与信息处理专业硕士和博士学位. 主要研究方向为图像处理, 成像制导, 光电对抗. 本文通信作者.

E-mail: navylst@163.com

**(LIU Song-Tao** Associate professor in the Department of Information and Communication Engineering, Dalian Naval Academy, and also postdoctor at the Postdoctoral Station of Information and Communication Engineering, Dalian University of Technology. He received his bachelor degree in aviation radar, master degree and Ph.D. degree in signal and information processing from Naval Air Engineering Institute, in 2000, 2003 and 2006, respectively. His research interest covers image processing, imaging guidance, and optoelectronic countermeasures. Corresponding author of this paper.)



**殷福亮** 大连理工大学信息与通信工程学院教授. 1984 年获得大连理工大学无线电技术专业学士学位, 1987 年获得大连理工大学通信与电子系统专业硕士学位. 主要研究方向为数字信号处理, 语音处理, 图像处理, 宽带无线通信技术.

E-mail: flyin@dlut.edu.cn

**(YIN Fu-Liang** Professor at the School of Information and Communication Engineering, Dalian University of Technology. He received his bachelor degree in electronic engineering and master degree in communications and electronic systems from Dalian University of Technology, in 1984 and 1987, respectively. His research interest covers digital signal processing, speech processing, image processing, and broadband wireless communication.)