

# 基于轮廓特征的人体行为识别

赵海勇 贾保先

(聊城大学计算机学院 聊城 252000)

**摘要** 提出了一种以人的动作序列图像的轮廓为特征、基于隐条件随机场的行为识别方法。首先,利用背景差分法和阴影消除技术提取运动人体轮廓。星型骨架方法只采用单一质心-边界距离来描述人体轮廓,因此对轮廓的局部特征描述能力较弱。定义一种新的基于距离组的轮廓描述方法,它将时变的 2D 轮廓形状转换为对应的 1D 距离向量。最后利用判别隐条件随机场对行为进行训练和识别。实验结果表明,本方法的正确识别率达到 91.4% 以上,识别结果较为理想。

**关键词** 轮廓,判别式模型,隐条件随机场,行为识别

中图分类号 TP391.4 文献标识码 A

## Human Action Recognition Using Image Contour

ZHAO Hai-yong JIA Bao-xian

(School of Computer Science and Technology, Liaocheng University, Liaocheng 252000, China)

**Abstract** A hidden conditional random fields-based method for action recognition was presented by using the contours of image sequence as representative descriptors of human posture. First, human silhouette contour extraction is obtained by background subtraction and shadow elimination. Star skeleton which uses the centroid-border distance function describing the contour is not good at describing the local features. A new contour descriptor based on distance array was defined, which is employed to transform the contour into a 1D distance signal. At last, the discriminative hidden conditional random fields are used to recognize human actions. The experimental results show that this method can achieve the correct recognition rate above 91.4%.

**Keywords** Contour, Discriminative model, Hidden conditional random fields, Action recognition

## 1 引言

人体行为识别由于在智能视频监控、基于内容的视频检索、体育赛事分析、人机交互和虚拟现实等领域具有广泛的应用价值,已经成为计算机视觉中的热点研究领域<sup>[1,2]</sup>。然而,现实环境的复杂性、动作序列中各种姿态的非刚性特点及行为分类边界的模糊性,使得人体行为识别成为一个非常具有挑战性的研究课题。

人体行为识别最基本的两个问题是行为的描述和识别。目前,常用的行为描述方法可分为两类:一类是基于低层图像信息的描述方法;另一类是基于高层人体结构的描述方法。基于低层图像信息的方法可以快速鲁棒地获取行为特征,一般只用于描述简单的行为。基于高层人体结构的方法可以更精确地描述行为特征,但是行为特征提取比较困难,往往要依赖于人体姿势估计的准确性<sup>[3]</sup>。在行为描述中可利用的低层图像信息包括:人的运动速度、光流、运动轨迹信息、人的侧影和轮廓等等。其中轮廓特征是人体行为的内在重要线索之一,对人体运动状态的描述具有一定的代表性。常用的轮廓特征提取方法有星型骨架<sup>[4]</sup>、傅立叶描述子等。星型骨架方

法的本质是利用质心-边界距离函数来描述轮廓。质心-边界距离描述方法简单,但不能精确描述轮廓,对轮廓曲线的局部特征描述较弱。傅立叶描述算子用系数的低频部分描述轮廓的大体特征,用高频部分来反映轮廓的微观特征即详细特征,但是在噪声环境下,其转换性能不是很高。针对星型骨架方法对人体轮廓局部特征描述能力较弱的缺点,本文采用一种称为距离组<sup>[5]</sup>的轮廓描述方法,即通过等弧长的分割人体轮廓线来构建轮廓线的多边形近似,利用多边形的顶点到形状的几何中心的距离和其两个后继顶点的距离构成的 3 个距离串来描述形状。这种描述方法计算非常简单,而且对人体轮廓的全局特征和局部特征都能进行很好的描述。最后使用具有隐变量结构的判别式模型 HCRF (Hidden Conditional Random Fields, 简称 HCRF)<sup>[6]</sup> 对行为进行识别。

## 2 行为识别相关研究

由于行为是一个时变信号,因此行为识别问题可以看成是时变信号的分类问题。用于解决这类问题的常用方法有两种,一是模板匹配法,一是状态空间法。基于模板匹配的算法可以分为帧对帧匹配方法和融合匹配方法。帧对帧匹配方法

到稿日期:2012-04-11 返修日期:2012-07-04 本文受国家自然科学基金(61104179)资助。

赵海勇(1981-),男,博士,讲师,主要研究方向为计算机视觉、模式识别, E-mail: zhaohaiyong@lcu.edu.cn; 贾保先(1982-),男,硕士,讲师,主要研究方向为电子商务、数据挖掘。

是指直接用测试集的行为特征序列与参考集的行为特征序列进行逐帧匹配;而融合匹配方法是指先将整个行为过程融合为一个整体模板或者若干个固定数目的模板,然后利用这有限个模板进行匹配。基于模板匹配的行为识别算法计算简单,但是对行为的执行时间比较敏感,造成相同的行为由于执行时间的长短不同而识别结果不同。状态空间的行为识别算法充分考虑了行为的动态过程,并且将时间尺度和空间尺度上的运动的微小变化采用概率的方法进行建模,因此逐渐成了主要的行为识别方法。HMM是目前应用最为广泛的行为识别方法,但作为以满足模型解释要求为目的的产生式模型,HMM通过观察当前时刻的运动状态往往很难判断该运动所属的类别。此外,产生式模型为了降低计算复杂度,往往假设观察序列是基于条件独立的,这种不确切的假设导致其不能表示时间序列中的依赖关系以及序列间的特征多重重叠。

为了改进产生式模型对时间序列建模所带来的问题,人们开始研究基于判别式模型(Discriminative Models, DM)的方法。判别模型最早是根据最大熵马尔可夫模型(Maximum Entropy Markov Models, MEMM)的思想建立起来的。针对MEMM的标记偏置问题,Lafferty等人首次提出了条件随机场(Conditional Random Fields, 简称 CRF)<sup>[7]</sup>。近年来,条件随机场模型受到了越来越多研究人员的关注。Sminchisescu等人<sup>[8]</sup>采用CRF进行动作识别,实验表明,他们的模型能够比MEMM取得更好的识别效果。HCRF通过引入隐状态对CRF进行了扩展。通过引入隐状态,HCRF增强了对随机变量为结构型数据时的处理能力。当图像序列相邻帧之间存在时空邻域约束关系时,利用HCRF对结构型数据的处理能力可以更加自然地对这种时空邻域关系进行建模。Wang等人<sup>[9]</sup>采用HCRF对人体行为进行识别,实验结果表明HCRF的识别效果要好于CRF和HMM。图1给出了HMM模型、CRF模型和HCRF模型示意图。

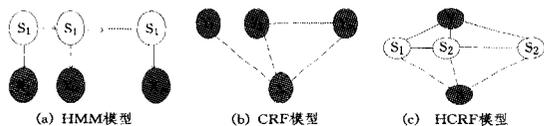


图1 HMM模型、CRF模型和HCRF模型示意图

### 3 特征提取与表示

#### 3.1 人体轮廓提取

假定包含人体行为的数据采集环境为:1)摄像机静止不动;2)摄像机方向与人体运动方向近似垂直;3)环境中只有一个人在运动。人体轮廓提取步骤如下:

(1)图像去噪:视频序列在获取和传输时常常会受到各种各样的噪声干扰,例如高斯白噪声、脉冲噪声和乘性噪声等。为了改善图像质量并便于各种后续处理,本算法采用高斯滤波去噪和邻域去噪相结合的方法。

(2)目标检测与阴影去除:采用背景差分法检测视频序列中的运动人体。首先建立一个随时间和空间变化而自适应变化的背景模型<sup>[10]</sup>,然后将当前帧与背景模型进行差分,设定阈值,二值化分割图像得到前景和背景。目标检测的目的是从背景图像中分割提取出运动变化区域,因此,在检测运动目标的同时,运动目标投射的阴影也会被检测为运动前景的一部分。本文采用文献<sup>[11]</sup>提出的算法对前景/阴影进行分割,

该方法适用于室内、户外等多种环境。

(3)轮廓提取:采用基于连通性的边界跟踪算法提取整个图像序列中行人的轮廓。图2给出了其中一帧的人体侧影和轮廓提取结果。



图2 运动人体轮廓提取

#### 3.2 人体轮廓表示

一个二维的人体轮廓线可以通过边界解卷绕转换为对应的一维距离信号<sup>[12]</sup>。将轮廓形式化地表示为沿逆时针方向的点的有序集  $C = \{p_i = (x_i, y_i), i = 1, 2, \dots, N\}$ , 其中  $(x_i, y_i)$  是点  $p_i$  在图像中的坐标,  $N$  是轮廓线上点的个数。定义一种基于距离组的轮廓描述函数。

(1)计算轮廓的质心  $p_c$ :

$$x_c = \frac{1}{N} \sum_{i=1}^N x_i, y_c = \frac{1}{N} \sum_{i=1}^N y_i \quad (1)$$

(2)定义轮廓的直径:

$$D = \max_{i,j \in \{1, \dots, N\}} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (2)$$

(3)对轮廓线做多边形近似:

选定轮廓线头顶边缘点作为起点,沿逆时针方向将轮廓线分割成  $M$  段相同长度的弧:  $\overline{s_0 s_1}, \overline{s_1 s_2}, \dots, \overline{s_{m-1} s_m}$ , 其中  $s_i$  为割点,且  $s_0 = s_m$ 。弧段  $s_i s_{i+1}$  表示从  $s_i$  出发沿轮廓线按逆时针方向到达  $s_{i+1}$  所经过的轮廓线上的点的有序集,其长度为集合中相邻点  $s_{i+1}$  的欧式距离的和。顺序连接这些割点所形成的线段的集合  $V = \{\overline{s_0 s_1}, \overline{s_1 s_2}, \dots, \overline{s_{m-1} s_m}\}$  构成了一个与起点有关的基于等弧长分割的近似轮廓线的多边形。

(4)描述轮廓线的3个归一化距离组表示如下:

$$\lambda_1 = \{\alpha_i = d(s_i, p_c) / D, i = 0, \dots, M-1\} \quad (3)$$

$$\lambda_2 = \{\beta_i = d(s_i, s_{i+1}) / D, i = 0, \dots, M-1\} \quad (4)$$

$$\lambda_3 = \{\gamma_i = d(s_i, s_{i+2}) / D, i = 0, \dots, M-1\} \quad (5)$$

其中,  $d(\cdot)$  表示两点之间的欧式距离。图3给出了3个距离  $\alpha_i, \beta_i, \gamma_i$  的示意图。

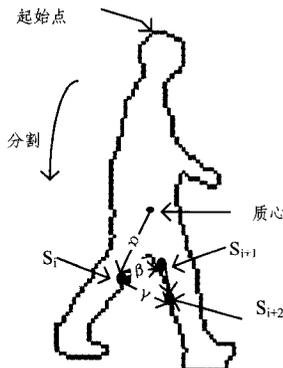


图3  $\alpha_i, \beta_i, \gamma_i$  的示意图

距离组与中心距离函数一样利用了轮廓线上的点到质心的距离来刻画轮廓,而质心是轮廓的全局特征点,因此对全局特征具有一定的描述能力。轮廓线上一段弧的弯曲程度是局部特征,本文利用弦长来刻画,即多边形的顶点到其两个后继顶点的距离( $\beta_i, \gamma_i$ )用于描述轮廓线的局部弧段的弯曲程度,因此距离串对局部特征也进行了描述。

### 3.3 人体轮廓之间的匹配

为了计算任意两个人体轮廓  $p$  和  $q$  之间的距离,采用距离串描述轮廓。设轮廓  $p$  的描述为:  $\lambda_1^p = [\alpha_i^p, i=0, \dots, M-1]$ ,  $\lambda_2^p = [\beta_i^p, i=0, \dots, M-1]$ ,  $\lambda_3^p = [\gamma_i^p, i=0, \dots, M-1]$ 。  $q$  的描述为:  $\lambda_1^q = [\alpha_i^q, i=0, \dots, M-1]$ ,  $\lambda_2^q = [\beta_i^q, i=0, \dots, M-1]$ ,  $\lambda_3^q = [\gamma_i^q, i=0, \dots, M-1]$ 。由于轮廓线是闭合的,因此得到的距离串是环形的。计算两个环形串的距离,首先要确定串的距离串起始位置,用偏移量  $t$  代表匹配的起始位置,  $t$  的取值为:  $0, 1, \dots, M-1$ 。对于  $t$  的每一个取值,完成一次串匹配即计算两个串的距离,在  $M$  次串匹配中,取最小距离作为形状差异的度量。轮廓  $p$  和  $q$  的距离定义如下:

$$diff(p, q) = \min_{t=0}^{M-1} \sum_{i=0}^{M-1} ((\alpha_i^p - \alpha_{i+t}^q)^2 + (\beta_i^p - \beta_{i+t}^q)^2 + (\gamma_i^p - \gamma_{i+t}^q)^2) \quad (6)$$

## 4 行为建模与识别

### 4.1 HCRF 模型

HCRF 可以根据输入序列  $X$  推断动作类别  $Y$ 。  $X$  由  $m$  个局部观测向量组成:  $X = \{x_1, x_2, \dots, x_m\}$ , 每个局部观测  $x_i$  对应一个特征向量。对任意观察序列  $X$ , 存在对应向量  $s = \{s_1, s_2, \dots, s_t\}$ , 其中  $s_i \in S$ ,  $S$  表示所有可能出现的隐状态集合。对于给定的观察数据序列, HCRF 按照如下方式对行为的条件概率进行建模:

$$P(y|X, \theta) = \sum_s P(y, s|X, \theta) = \frac{\sum_s e^{\Psi(y, s, X; \theta, \omega)}}{\sum_{y' \in Y, s' \in S^m} e^{\Psi(y', s', X; \theta, \omega)}} \quad (7)$$

式中, 势能函数  $\Psi(y, s, X; \theta, \omega)$  表示隐状态结构、观测数据和行为类别之间的相互影响程度,  $\omega$  是窗口大小,  $\theta$  是参数。

$$\Psi(y, s, X; \theta, \omega) = \sum_{j=1}^n \varphi(X, j, \omega) \cdot \theta_s[s_j] + \sum_{j=1}^n \theta_y[y, s_j] + \sum_{(j, k) \in E} \theta_e[y, s_j, s_k] \quad (8)$$

图 E 是链状结构, 每个节点都对应  $t$  时刻的一个隐状态,  $\varphi(X, j, \omega)$  可以表示观测窗口中的任意特征。

### 4.2 训练与识别

参数  $\theta$  由 3 部分组成:  $\theta = [\theta_e, \theta_y, \theta_s]$ ,  $\theta_e$  度量的是连续状态  $j$  和  $k$  与动作类别  $s_j \in S$  之间的兼容度;  $\theta_y$  度量的是隐藏状态  $j$  与动作类别  $Y$  之间的兼容度;  $\theta_s$  表示对应状态  $s_j \in S$  的参数  $\theta_s$ 。训练中, 采用目标函数:

$$L(\theta) = \sum_{i=1}^n \log P(y_i | x_i, \theta) - \frac{1}{2\sigma^2} \|\theta\|^2 \quad (9)$$

式中,  $n$  表示训练样本序列的总个数。

给定新的测试序列  $X$  和学习得到的参数  $\theta^*$ , 测试序列  $X$  的动作类别由如下方式确定:

$$y = \arg \max_{y \in Y} P(y|X, \omega, \theta^*) \quad (10)$$

## 5 实验结果与分析

采用 Blank 等人<sup>[13]</sup> 的数据库对算法进行验证。该数据库共计包含 90 个低分辨率的视频序列 (180 × 144 像素, 50 帧/秒), 所有动作分别由 9 名实验人员完成, 每人完成 10 种行为, 包括“run”、“walk”、“skip”、“jumping-jack”、“jump”、“jump-in-place-on-two-legs”、“gallop side ways”、“wave-two-hands”、“wave one-hand”、“bend”等。

### (1) 人体轮廓提取

利用轮廓提取方法提取行为序列的所有人体轮廓, 图 4 给出了示例图像及其相应的人体轮廓。

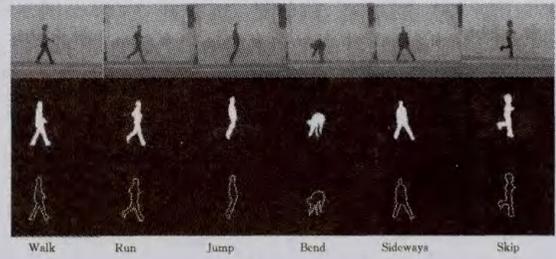


图 4 数据库的示例图像及对应的人体轮廓

### (2) 行为识别

轮廓描述方法的参数是轮廓线的分割段数, 分割段数越多, 获得的近似多边形对轮廓的近似效果越好, 但段数过多一方面会增加计算量, 另一方面会降低抗干扰的能力。本文根据星型骨架方法对人体轮廓的描述能力, 将分割段数设为 8。

将所有 9 人所完成的动作序列分成 3 组, 每组序列由 3 人完成。其中一组序列用于训练一个“多类分类 (multi-class)”的 HCRF 模型, 将训练得到的 HCRF 模型在其余两组动作序列上测试。参数设置如下: 窗口参数  $\omega = 2$ 、隐藏状态数等于 8, 最后选择概率最高的行为类别作为识别结果。表 1 显示了识别结果混淆矩阵。

表 1 测试数据的识别结果混淆矩阵

	run	walk	Skip	Jack	Jump	Pjump	Side	Wave1	Wave2	bend
run	0.90	0.05	0.05							
walk	0.05	0.95								
skip	0.05		0.95							
jump					1.00					
Sideways							1.00			
bend										1.00

### (3) 算法对比

为了测试不同模型结构对识别结果的影响, 在相同数据库上对 HMM, CRF, HCRF 3 种模型的识别效果进行了对比。

HMM 模型: 为 10 类行为分别训练一个 HMM 模型, 每个 HMM 模型的隐藏状态根据经验设为 6 个。识别测试时, 分别在每类行为模型上进行行为识别, 选择概率最大的那个模型类别作为识别结果。

CRF 模型: 训练一个线性链状 CRF 模型 (见图 1(b)), 每类行为对应一种状态。在训练过程中, 状态序列  $\{s_1, s_2, \dots, s_m\}$  定义为行为类别  $s$  的  $m$  次常量。对于所有的测试序列, CRF 模型为序列中每帧图像预测状态, 最后根据状态序列中出现次数最多的状态判定行为类别。

HCRF 模型: 分别采用“一对多 (one-vs.-all)”和“多类分类 (multi-class)”两种模式训练 HCRF 模型。在“一对多”实验中, 为每类行为训练一个 HCRF 模型, 每个模型包含 6 个隐藏状态; 识别时, 分别向每类模型输入测试序列, 选择概率

最大的那个模型类别为识别结果。在“多类分类”实验中,训练一个具有 10 个隐藏状态的 HCRF 模型;识别测试时,选择概率最大的模型类别为识别结果。

表 2 对比了在给定数据库上,利用不同算法模型进行行为识别的实验结果。由表 2 可以看出:(1) CRF 模型和 HCRF 模型的识别率要高于 HMM 模型。这在一定程度上也验证了对有限动作集合进行行为识别时,判别式模型要优于产生式模型。(2) 加入长相关约束(即窗口大小  $\omega$  取不同的值)有助于提高算法的识别率,随着窗口大小  $\omega$  的增加,识别率也在增加。(3) “多类分类”模式训练的 HCRF 模型的识别效果要优于“一对多”模式训练的 HCRF 模型。

表 2 不同模型的行为识别率对比

模型	识别率
HMM	83.5
CRF( $\omega=0$ )	86.7
CRF( $\omega=1$ )	89.9
HCRF(multiclass, $\omega=0$ )	93.5
HCRF(multiclass, $\omega=1$ )	95.4
HCRF(multiclass, $\omega=2$ )	96.8
HCRF(one-vs.-all, $\omega=0$ )	91.4

**结束语** 本文将人体行为的侧影轮廓特征作为出发点,提出了一种新的行为识别方法;使用了一种基于距离组的轮廓描述方法,提高了星型骨架方法对轮廓局部特征的描述能力;使用隐条件随机场对行为进行建模和识别,结果表明,对于有限动作集合进行动作识别时,判别式模型要优于产生式模型。本算法的不足之处在于要求能比较精确地提取轮廓;此外,还需要在如何合理选择和建立模型方面展开更加深入的研究。

## 参考文献

[1] 王亮,胡卫明,谭铁牛. 人运动的视觉分析综述[J]. 计算机学报, 2002, 25(3): 225-237

[2] 杜友田,陈峰,徐文立,等. 基于视觉的人的运动识别述[J]. 电子学报, 2007, 35(1): 84-90

[3] 谷军霞,丁晓青,王生进. 行为分析算法综述[J]. 中国图像图形学报, 2009, 14(3): 377-387

[4] Chen H-S, Chen H-T, Chen Y-W, et al. Human action recognition using star skeleton[C]//Proc of the 4th ACM international workshop on Video surveillance and sensor networks. New York: ACM Press, 2006: 171-178

[5] 王斌,舒华忠,施朝建,等. 一种基于轮廓线的形状描述与匹配方法[J]. 电子与信息学报, 2008, 30(4): 949-952

[6] Quattoni A, Wang S, Morency LP, et al. Hidden conditional random fields [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(10): 1848-1852

[7] Lafferty J, McCallum A, Pereira F. A conditional random fields, Probabilistic models for segmenting and labeling sequence data [C]//Proc of the Annual International Conference on Machine Learning. San Francisco, USA, 2001: 282-289

[8] Sminchisescu C, Kanaujia A, Li Z, et al. Conditional models for contextual human motion recognition [J]. Journal of Computer Vision and Image Understanding, 2006, 104(2/3): 1808-1815

[9] Wang S B, Quattoni A, Morency L P, et al. Hidden conditional random fields for gesture recognition [C]//Proc of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2006: 1521-1527

[10] 侯叶,郭宝龙. 基于图切割的人体运动检测[J]. 光电子·激光, 2007, 18(6): 725-728

[11] 褚一平,陈勤,黄叶珏,等. 基于判别模型的视频前景/阴影自动分割算法[J]. 模式识别与人工智能, 2008, 21(6): 849-855

[12] 王亮,胡卫明,谭铁牛. 基于步态的身份识别[J]. 计算机学报, 2003, 26(3): 353-360

[13] Gorelick L, Blank M, Shechtman E, et al. Action as space-time shapes [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(12): 2247-2253

(上接第 303 页)

$\delta=1.0$ ,应用本文算法计算得到限定 Voronoi 图,如图 6 所示,网格单元数量为 1694,生成图示限定 Voronoi 图所用的时间为 3946ms。实验的结果证明了本文给出的限定 Voronoi 图生成算法对于内部边界约束、线束约束条件以及不规则区域均可以得到质量较好的满足约束条件的限定 Voronoi 图。

**结束语** 本文针对已有限定 Voronoi 图算法的不足,给出了改进的梯形检测带细分算法生成二维 Voronoi 网格。主要是引入了几个控制因子,简化了限定条件附近的处理,保证了算法对于复杂约束条件下的收敛,同时提高了整个算法的柔性,使生成的限定 Voronoi 图网格单元具有较好的质量,能满足煤层气数值模拟等计算的需要。下一步工作考虑在本文算法的基础上,研究三维限定条件下限定 Voronoi 图的生成算法,并对生成的三维限定 Voronoi 图进行质量与尺寸控制。

## 参考文献

[1] 马立玲,张有会. 分区加权 Voronoi 图的性质及其面积计算[J]. 计算机科学, 2011, 38(2): 195-198

[2] Okabe A, Boots B, Sugihara K, et al. Spatial tessellations: concepts and applications of Voronoi diagram (2nd edition) [M]. UK: John Wiley & Sons, Ltd. 2000: 242-245

[3] KAPPA Engineering. E06-Voronoi numerical module [EB/OL]. Technicalreference, <http://www.egr.uh.edu/courses/Engi 6314/TechnicalReferences/TR6Advanced Interpretation. PDF>

[4] 蔡强,杨钦,孟宪海,等. 二维 PEBI 网格生成的研究[J]. 工程图学学报, 2005, 26(2): 69-72

[5] 杨权一,蔡强,李吉刚. 二维限定 PEBI 网格生成技术的研究[J]. 计算机工程与应用, 2004, 22(7): 80-83

[6] 李吉刚,孟宪海,杨钦,等. 二维约束 Voronoi 网格构造及其尺寸、质量控制[J]. 计算机辅助设计与图形学学报, 2005, 17(9): 1950-1956

[7] Verma S. A flexible gridding scheme for reservoir simulation [Z]. Presented at SPE Intl, Student Paper Contest, 1995: 657-672

[8] 蔡强,王长飞,李海生,等. 二维复杂域 PEBI 网格细化生成算法[J]. 东南大学学报, 2009, 10(4): 343-347

[9] 蔡强. 限定 Voronoi 网格剖分的理论及应用研究[M]. 北京: 北京邮电大学出版社, 2010

[10] 汪嘉业,王文平,屠长河,等. 计算几何及应用[M]. 北京: 科学出版社, 2011

[11] 杨承磊,汪嘉业,孟祥旭. 多边形外部 Voronoi 图顶点和边数的上界[J]. 计算机辅助设计与图形学学报, 2005, 17(4): 689-693

[12] A Delaunay Refinement Algorithm for Quality 2-Dimensional Mesh Generation [J]. Journal of Algorithms, 1995, 18(3): 548-585