

# 参数自适应的条件随机场视频分割方法

郑河荣<sup>1</sup> 褚一平<sup>2</sup> 潘翔<sup>1</sup>

(浙江工业大学计算机学院 杭州 310014)<sup>1</sup> (杭州电子科技大学 杭州 310032)<sup>2</sup>

**摘要** 针对已有算法需要采用一系列参数经验值得到最优视频分割结果的问题,提出根据视频特征自适应地计算视频邻域关系的特征强度函数,构造参数自适应的条件随机场视频分割模型。算法的核心思想是采用视频中像素之间的邻域关系自适应计算各个模型的特征函数,通过条件随机场模型对这些特征能量函数进行约束并利用 Gibbs 采样算法对该模型进行求解,获得全局优化的分割结果。针对不同环境下的视频分割实验表明,该算法能够很好地逼近最优经验参数所得到的视频分割结果,从而避免定义经验值所导致的算法局限性问题。

**关键词** 视频分割,背景建模,参数自适应,条件随机场

**中图分类号** TP391 **文献标识码** A

## Adaptive Parameters Conditional Random Field Video Segmentation Algorithm

ZHENG He-rong<sup>1</sup> CHU Yi-ping<sup>2</sup> PAN Xiang<sup>1</sup>

(School of Computer Science, Zhejiang University of Technology, Hangzhou 310014, China)<sup>1</sup>

(School of Computer Science, Hangzhou Dianzi University, Hangzhou 310032, China)<sup>2</sup>

**Abstract** Aimed at the problem that the video segmentation method based on conditional random fields needs set empirical values to obtain optimal segmentation results, the video segmentation model was proposed to compute adaptively video neighboring relationship feature function by video features and construct adaptive parameters conditional random fields. The core idea of algorithm is to calculate different kinds of model feature function by pixel neighboring relationship in video, these feature energy functions are constrained by conditional random field model, which is solved via Gibbs sampling algorithm to obtain globally optimal segmentation results. The experiment shows that the results of adaptive parameter algorithm are the same as one of optimal empirical parameters.

**Keywords** Video segmentation, Background modeling, Adaptive parameters, Conditional random fields

## 1 引言

计算机视觉领域一个基础而关键的问题是如何从视频序列中准确地提取出运动目标。随着视频技术在智能交通、楼宇监控等领域的广泛应用,视频分割更是成为一个迫切需要解决的问题。

近年来,国内外学术界对视频分割进行了大量的研究,提出了大量的算法。其中,背景建模方法<sup>[9,10]</sup>由于具有较好的实时性,被广泛应用于智能交通等领域。而学术界也针对背景建模方法提出了一系列改进算法。其改进主要集中在如下两方面:1)采用概率统计假设提高像素级分类的准确率。赵明<sup>[25]</sup>等人则在高斯分布模型基础上,提出采用统计推断对后续帧进行假设检验,进一步提高视频分割质量。邹正平<sup>[21]</sup>等人则把背景建模问题转化为包含“状态、决策、状态转移规律、代价函数和目标函数”等5个部分的动态规划问题。Elgammal等人<sup>[3]</sup>提出使用非参数估计方法对背景进行建模,这种方法采用高斯核函数估计视频序列上像素的概率密度。Stenger等人<sup>[12]</sup>提出一个基于拓扑结构可变的隐马尔可夫模

型的背景建模方法,它可以适应光线突然变化等情况。陈睿等人<sup>[20]</sup>在文献[3]的基础上结合了边界信息,用于提高视频序列中运动目标分割和检测的健壮性。文献[25]提出了一种利用贝叶斯学习方法来进行视频分割的算法,通过对每个像素点不断变化的背景建模,用贝叶斯学习方法判断在光线缓慢或者突然变化的时候,每个像素点是属于前景还是属于背景。高丽等人<sup>[22]</sup>针对复杂背景的情况,设计了一种视频对象自动分割新算法,它利用随机信号的统计特性累计获得背景,并利用背景累积过程中分类讨论的观点,有效地将运动对象从视频序列中自动地分割出来。2)在像素级分类基础上,采用上下文关系消除前景目标中产生的“虫洞”等误分割。因此,需要通过视频序列的相关性信息进行纠正,提高视频分割精确率。Zhou等人<sup>[18]</sup>则使用多分辨的马尔可夫随机场模型进一步对像素之间的时空信息进行约束,使得算法可以较好地抵抗户外环境中图像的噪声以及摄像机的抖动。Yaser等人<sup>[16]</sup>对Elgammal的算法<sup>[3]</sup>进行了扩展,增加了对前景目标的建模能力,同时利用马尔可夫随机场模型对同一视频帧像素之间的邻域关系进行建模,通过求解马尔可夫随机场模型

到稿日期:2009-12-10 返修日期:2010-03-02 本文受国家自然科学基金(60703001)资助。

郑河荣(1971-),男,副教授,主要研究方向为图像处理和识别、视频分析等,E-mail:hailiang@zjut.edu.cn;褚一平(1979-),男,博士,主要研究方向为视频处理等;潘翔(1977-),男,副教授,主要研究方向为计算机图形图像等,E-mail:panx@zjut.edu.cn(通信作者)。

实现前景背景的分割。而二维马尔可夫随机场模型可以较好地处理像素之间的邻域约束关系,文献[7]利用二维马尔可夫随机场模型对像素之间的时空信息进行约束,通过求解最大后验概率实现对视频前景目标的分割。Wang 等人<sup>[14,15]</sup>把条件随机场模型应用于视频目标分割中,该算法通过条件随机场模型对视频序列的时空邻域信息进行建模,构造了一个可以根据前一帧视频图像的条件随机场模型数据对当前帧视频图像的条件随机场模型进行迭代更新的滤波器。文献[23]提出了基于隐条件随机场模型的视频目标分割算法,通过使用隐条件随机场可以更直观地对视频序列的多种上下文特征进行建模。

从已有研究可以发现,采用随机场进行上下文建模,能够有效地提高视频分割准确率。但是,随机场建模面临的一个主要问题是需要用户自定义参数来得到最优分割结果,从而导致算法不具有普遍性。针对该问题,本文提出参数自适应的条件随机场建模视频分割方法,该方法根据视频中像素之间的邻域关系来构造模型,用以自适应计算条件随机场中的特征强度,从而避免已有研究需要通过经验值限定视频邻域关系无法自适应处理各种视频的弊端。实验结果表明,参数自适应算法能够很好地逼近采用最优经验参数得到的分割效果。

## 2 参数自适应的条件随机场视频分割模型

### 2.1 基于后验概率的视频分割模型

视频序列中相邻的像素之间存在相互影响的关系,这些相互影响的像素被定义成邻域关系。合理地利用视频序列的邻域信息有助于提高视频分割的质量。因此,可以利用条件随机场对视频的颜色、运动特征以及邻域关系进行建模,形成条件随机场视频分割模型。通过概率图模型推断算法求解这个模型,可获得视频分割结果。把视频中的像素集合定义为随机变量集  $X$ ,其分类集概率被定义成一个二维的条件随机场:

$$P(L|X) = \frac{1}{Z(X)} \exp(-E(L;X)) \quad (1)$$

$$E(L;X) = U(L;X) + V(L;X)$$

式中,能量项  $U(\cdot)$  表示单个像素根据背景模型、阴影模型和前景模型分类计算得到的能量, $V(\cdot)$  表示邻域关系能量。 $Z(X)$  可以保证计算得到的结果是一个归一化的概率,它的计算公式如下:

$$Z(X) = \sum_L \exp(-E(L;X)) \quad (2)$$

条件随机场对视频的分割结果可通过式(3)来计算。

$$\hat{L} = \operatorname{argmax}_L P(L|X) \quad (3)$$

### 2.2 参数自适应的特征强度函数

显然,对于上述条件随机场分割建模,其分割结果的好坏主要取决于能量函数的定义。因此各个相关的特征强度函数需要根据实际的情况来设计。在此,能量项  $U$  表示局部的能量之和,它由局部特征函数加权所得,其定义如下:

$$U(L;X) = \sum_{n=1}^N \delta(L_i, L_{i,m}) \log P(X_i | L_i) \quad (4)$$

式中, $\delta$  为 Kronecker delta 函数, $L_{i,m}$  为结点  $i$  的初始分类标签,它根据像素级别上的背景模型、阴影模型以及前景模型获得;式中  $\log P(X_i | L_i = 0)$ ,  $\log P(X_i | L_i = 1)$  和  $\log P(X_i | L_i =$

2) 是分别根据混合高斯分布模型计算得到的背景、阴影以及前景的概率的对数<sup>[23]</sup>。而邻域关系能量项  $V^S$  反映了视频帧中像素之间的邻域强度关系,空域邻域关系能量项  $V^S$  定义如下:

$$V^S(L;X) = \sum_{j \in N(i)} v^S(L_i, L_j) \quad (5)$$

其中,邻域关系对应的特征函数定义为:

$$v^S(L_i, L_j) = \begin{cases} \beta_1, & \text{if } L_i = L_j = 0 \\ \beta_2, & \text{if } L_i = L_j = 1 \\ \beta_3, & \text{if } L_i = L_j = 2 \\ \beta_4, & \text{if } L_i \neq L_j \end{cases} \quad (6)$$

式中, $L_j$  为  $L_i$  的空域邻域上的结点, $\beta_1, \beta_2, \beta_3$  和  $\beta_4$  为常量,表示空域邻域中各结点之间的相互影响程度。显然,这些参数直接影响到最终的分割结果。已有算法针对不同分割视频预定义不同的经验值,用以得到最优视频分割结果。

显然,对于这种经验参数方法,没有考虑邻域关系中不同像素之间的相关性。这种相关性体现在相邻两个像素如果同属于前景或者背景,通常会具有类似的颜色特征。颜色越相似,表明相邻像素相关性越强。因此,可以通过像素之间的颜色相似性定义特征强度函数,从而避免上述方法设计视频邻域能量函数时需要设置经验值进行分割。为此,我们根据视觉特征构造模型,用于计算阈值化过程中的邻域关系能量项,从而针对不同的视频自适应地计算视频邻域强度,避免需要用户设定上述经验参数  $\beta_1, \beta_2, \beta_3$  和  $\beta_4$ , 而无法适用于不同视频。对于邻域关系,其能量项  $V^S(\cdot)$  定义如下:

$$V^S(L;X) = \sum_{j \in N(i)} \delta(L_i, L_j) \left( \frac{1 + T(x_i, x_j)}{2} \right) \quad (7)$$

式中, $N(i)$  表示结点  $i$  的邻域, $T(x_i, x_j)$  是用于自适应计算邻域能量强度的函数,定义如下:

$$T(x_i, x_j) = \exp\left(-\frac{\langle ||x_i - x_j||^2 \rangle}{2 \langle ||x_i - x_j||^2 \rangle}\right) \quad (8)$$

式中, $\langle ||x_i - x_j||^2 \rangle$  表示当前帧所有相邻像素强度差的期望值。从上面等式的右边指数内部可以看到,当  $||x_i - x_j|| < \sqrt{\langle ||x_i - x_j||^2 \rangle}$  时,式(8)的值将会变得非常大;而当  $||x_i - x_j|| > \sqrt{\langle ||x_i - x_j||^2 \rangle}$  时,式(8)的值将会变得很小。由于  $x_i$  和  $x_j$  为视频帧中相邻的两个像素点的值,它们的强度差异决定了  $T(x_i, x_j)$  的值,通过式(8)可以根据邻域像素之间的强度较好地反映出它们之间的能量关系。通过该式可以针对不同的视频自适应地计算视频邻域关系的能量,从而避免已有研究采用经验参数进行能量值计算。

### 2.3 基于 Gibbs 采样算法的约束求解

显然,针对条件随机场建模,视频分割转化为后验概率的约束求解问题。在这里,可以通过 Gibbs 采样算法进行约束求解。由于像素级别上的点分布模型可以获得对视频背景、阴影以及前景目标的初步分割,这个初步的分割结果可以作为 Gibbs 采样算法的初始输入,这相对 Gibbs 采样算法原本以一个随机值作为初始输入而言可以大大减少迭代次数。Gibbs 采样算法是从一个初始值开始,通过不断地迭代来达到总能量的平衡。使用背景模型和阴影模型可以获得视频帧的分类集,以这个分类集作为 Gibbs 采样算法的初始输入。Gibbs 采样算法的每一次迭代都会对式(1)进行采样,根据温度常量进行模拟退火,最终达到一个总能量的平衡。Gibbs 采样算法可以较好地利用视频序列之间的连贯信息,除了使

用点分布模型的分割结果作为初始输入外,也可以把视频前一帧的分割结果作为后一帧中 Gibbs 采样算法的初始输入。利用视频序列的这种连贯信息,也可以大大减少 Gibbs 采样算法的迭代次数。

### 3 实验结果和分析

为了验证参数自适应算法的有效性和稳定性,本节针对不同类型的视频进行分割实验,并和类似算法的分割结果进行比较。实验视频一段来源于室外,阴影相对比较弱,背景也相对比较简单。另外一段来源于室内,由于光照等问题,具有明显的阴影,用于比较的算法包括加权模型、二维条件随机场模型。其中加权模型属于像素级别分类器,二维条件随机场模型是在像素级别分类器基础上,采用随机场进行邻域关系建模。在分割结果中,背景用白色表示,黑色为前景,阴影用粉红色表示。

首先,对背景相对比较简单视频进行分割实验。这段视频为室外一人在行走,背景基本保持不变。对于这种简单背景的视频,采用合理的像素级分类器进行建模,能够得到比较好的分割效果。但是由于该视频前景目标像素颜色比较接近背景颜色,造成部分前景像素误分类。二维条件随机场比较稳定,能够较好地去除分割噪声。而从图 1 中的分割效果比较可以看出,本文提出的参数自适应算法能够较好地逼近采用经验参数得到的分割效果。

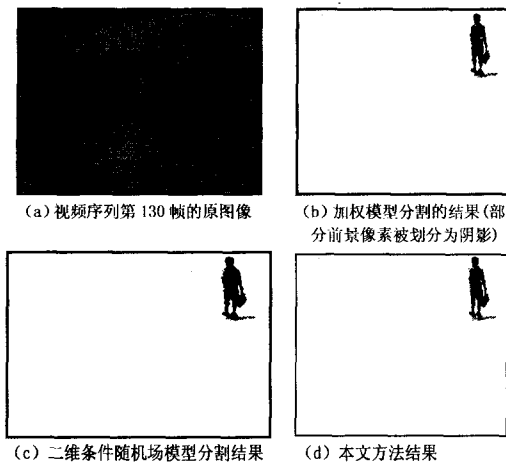


图 1 室外行人不同算法分割结果

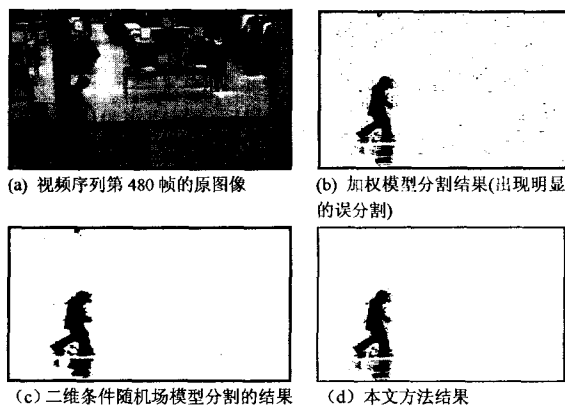


图 2 室内行人不同算法分割结果

为进一步验证本文算法的有效性,对背景相对比较复杂的视频进行测试。该段视频是一行人在商场室内行走的序

列,图 2 是视频序列中第 480 帧的原始图像。这段视频中背景比较复杂,而且商场灯光等效果所造成的阴影也比较明显。由于像素级分类器加权模型没有进行相邻关系建模,使得背景区域出现大量的前景和阴影像素点,形成误分割。和这些算法相比较,通过条件随机场对视频序列空域邻域进行约束,参数自适应模型和二维条件随机场的实验结果非常类似,但是二维条件随机场模型根据具体的场景需要手工来设置经验值,而本文提出的算法无须根据场景设置经验值。

**结束语** 在背景建模基础上,采用随机场进行视频分割建模成为提高分割准确率的有效方法。但是,已有随机场建模方法,需要用户定义各种参数来获得最优分割结果。本文提出参数自适应分割模型,以根据当前视频特征自动计算出所需参数。实验结果表明,改进的算法能够很好地逼近最优参数获得的视频分割结果,大大提高算法的通用性。

近年来,在视频分割应用领域,如智能交通监控等,一个主要特点是更多采用例如球形摄像机等运动式监控方法。这种方法所形成的视频具有背景不停变化的特点,因此,难以直接采用现有的建模方法。但是无论背景如何变化,人类视觉都能够准确有效地从视频中分离出运动目标。因此,如果能够建立起反映人类视觉的感知模型,显然能显著提高视频分割效果。因此,如何在已有视频分割研究成果基础上,融入视觉心理学内容,将是后续研究的重点。

### 参考文献

- [1] Boykov Y, Veksler O. Fast approximate energy minimization via graph cuts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(11): 1222-1239
- [2] Chum O, Zisserman A. An Exemplar Model for Learning Object Classes[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2007
- [3] Elgammal A, Harwood D, Davis L S. Non-parametric background model for background subtraction, 2000
- [4] Leibe B, Schiele B. Interleaved Object Categorization and Segmentation[C]//British Machine Vision Conference, 2003
- [5] Li H, King N N. Face Segmentation in Head-and-Shoulder Video Sequence Based on Facial Saliency Map[C]//IEEE Int'l. Symp. Circuits and Syst. Kos, Greece, May 2006
- [6] Mikolajczyk K, Leibe B, Schiele B. Multiple Object Class Detection with a Generative Model[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2006
- [7] Migdal J, Grimson E. Background Subtraction Using Markov Thresholds[C]//IEEE Workshop on Motion and Video Computing, 2005
- [8] Seemann E, Fritz M, Schiele B. Towards Robust Pedestrian Detection in Crowded Image Sequences[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2007
- [9] Stauffer C, Grimson W E L. Adaptive Background Mixture Models for Real-time Tracking[C]//IEEE Conference on Computer Vision and Pattern Recognition, 1999
- [10] Stauffer C, Grimson W E L. Learning Patterns of Activity Using Real-time Tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 747-757

为观察半方差约束对决策的影响,在不同的半方差约束下,通过运行智能算法,得到不同的最大本原机会测度和决策方案,如表 2 所列。

表 2 不同半方差约束下的最大机会测度和决策方案

半方差 约束	第 i 种贷款比例					最大本原 机会测度
	i=1	i=2	i=3	i=4	i=5	
0.004	0.2690	0.2122	0.1897	0.2234	0.1056	0.446751
0.005	0.2310	0.3008	0.1236	0.3401	0.0045	0.501930
0.006	0.8195	0.0484	0.0316	0.1004	0.0001	0.586291
0.007	0.9308	0.0048	0.0124	0.0518	0.0002	0.616019
0.008	0.9267	0.0126	0.0116	0.0474	0.0017	0.618533
0.009	0.9306	0.0092	0.0110	0.0472	0.0020	0.620623
0.01	0.9341	0.0085	0.0105	0.0452	0.0016	0.621724

**结束语** 本文使用半方差作为贷款组合中的风险度量方式,讨论在模糊随机环境下的贷款组合优化模型,将贷款收益率刻画为模糊随机变量,建立半方差约束下的模糊随机收益率贷款组合优化模型,并应用混合智能算法进行求解,通过实例验证了模型和算法的有效性。

### 参 考 文 献

- [1] Markowitz H. Portfolio selection[J]. Journal of Finance, 1952, 7:71-93
- [2] Altman E I. Corporate Bond and Commercial Loan Portfolio Analysis[R]. New York: New York University Salomon Center, 1997
- [3] 马志卫,刘应宗. 基于蒙特卡洛模拟的贷款组合优化决策方法[J]. 管理科学, 2006, 19(3): 66-70
- [4] Li D, Ng Wan-Lung. Optimal dynamic portfolio selection: Multi-period mean-variance formulation[J]. Mathematical Finance, 2000, 10(3): 387-406
- [5] Trevor S E, Wood R J. Asset-allocation decisions when risk is changing [J]. Journal of Financial Research, 1999, 22(3): 301-315
- [6] 迟国泰,姜大治,奚扬,等. 基于 VaR 收益率约束的贷款组合优化决策模型[J]. 中国管理科学, 2002, 10(6): 1-7
- [7] 宁玉富,唐万生,严维真. 商业银行贷款组合优化决策的机会准则模型[J]. 计算机工程与应用, 2008, 44(21): 235-237, 244
- [8] Huang Xiaoxia. Mean-semivariance models for fuzzy portfolio selection[J]. Journal of Computational and Applied Mathematics, 2008(217): 1-8
- [9] 马慧民,叶春明. 粒子群算法在贷款组合优化决策中的应用[J]. 计算机工程与应用, 2006(14): 219-221, 224
- [10] 陆源,朱邦毅. 基于半方差的投资项目风险度量模型研究[J]. 数量经济技术经济研究, 2005(7): 90-95
- [11] Liu Y K, Liu B. Fuzzy random variables: A scalar expected value operator[J]. Fuzzy Optimization and Decision making, 2003, 2(2): 143-160
- [12] Kwakernaak H. Fuzzy random variables- I. Definitions and theorems[J]. Information Sciences, 1978, 15: 1-29
- [13] Kwakernaak H. Fuzzy random variables- II. Algorithm and examples for the discrete case[J]. Information Sciences, 1979, 17: 253-278
- [14] Liu B, Liu Y-K. Expected value of fuzzy variable and fuzzy expected value models[J]. IEEE Transactions on Fuzzy Systems, 2002(10): 445-450
- [15] Liu B. Fuzzy random chance-constrained programming[J]. IEEE Transactions on Fuzzy System, 2001, 9(5): 713-720
- [16] Liu B. Fuzzy random chance-constrained programming[J]. IEEE Transactions on Fuzzy System, 2001, 9(5): 721-726
- [17] Liu B, Zhao R Q, Wang G. Uncertain programming with Application[M]. Tsinghua University Press, 2003
- [18] Gao J, Liu B. New primitive chance measures of fuzzy random event[J]. International Journal of Fuzzy System, 2001, 3(4): 527-531
- [19] 占德胜,唐加山. 关于扰动广义 cox 保险风险模型的破产概率[J]. 重庆邮电大学学报: 自然科学版, 2007, 21(6): 782-784
- [10] (上接第 256 页)
- [11] Stenger B, Ramesh V, Paragios N, et al. Topology free Hidden Markov models: Application to background modeling[C]//IEEE International Conference on Computer Vision. 2001
- [12] Sun Y R. Hierarchical Object-Based Visual Attention for Machine Vision[D]. 2003
- [13] Wang Y, Ji Q. A Dynamic Conditional Random Field Model for Object Segmentation in Image Sequences[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005
- [14] Wang Y, Loe K F, Wu J K. A Dynamic Conditional Random Field Model for Foreground and Shadow Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(2): 279-289
- [15] Yaser S, Mubarak S. Bayesian Modeling of Dynamic Scenes for Object Detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(11): 1778-1792
- [16] Yin P, Criminisi A, Winn J, et al. Tree-based Classifiers for Bilinear Video Segmentation[C]// Proc. IEEE Computer Vision and Pattern Recognition (CVPR). 2007
- [17] Zhou Y, Xu W, Tao H, et al. Background Segmentation Using Spatial-Temporal Multi-Resolution MRF[C]// IEEE Workshop on Motion and Video Computing. 2005
- [18] Zhai Y, Shah M. Visual attention detection in video sequences using spatiotemporal cues[C]// ACM International Conference on Multimedia. 2006: 815-824
- [19] 陈睿,邓宇,向世明,等. 结合强度和边界信息的非参数前景/背景分割方法[J]. 计算机辅助设计与图形学学报, 2005, 17(6): 1278-1284
- [20] 邹正平,卜佳俊,陈纯. 一种基于动态规划的视频分割方法[J]. 计算机辅助设计与图形学学报, 2002, 14(8): 743-746
- [21] 高丽,杨树元,李海强. 一种有效的基于时空联合的视频对象自动分割新算法[J]. 中国图象图形学报, 2005, 9(10): 1096-1104
- [22] 褚一平. 基于条件随机场模型的视频目标分割算法研究[D]. 杭州: 浙江大学, 2008
- [23] 吴思,林守勋,张勇东. 基于动态背景构造的视频运动对象自动分割[J]. 计算机学报, 2005, 8(28): 1386-1392
- [24] 王林波,赵杰煜. 基于贝叶斯学习的视频图像分割[J]. 中国图象图形学报, 2005, 9: 1073-1078
- [25] 赵明,李娜,陈纯. 采用统计推断的自动视频对象分割[J]. 计算机辅助设计与图形学学报, 2003, 15(3): 318-323