

基于 G-GMM 的视频序列运动目标检测算法研究

盛家川 杨 巍

(天津财经大学理工学院 天津 300200)

摘 要 为了能够从视频序列中快速准确地检测运动目标,在混合高斯背景差分法的基础上引入 Grabcut 算法,提出了一种新的运动目标检测 G-GMM(Grabcut-Gaussian Mixture Model)算法。首先通过混合高斯模型背景差分法提取运动目标初始二值轮廓,构建其最小的外接矩形;然后初始化矩形内图像信息,寻找潜在前景区域;最后采用迭代算法实现最优化分割,得到准确的运动目标轮廓。实验结果表明,在静止摄像机户外视频监控系统中,提出算法具有较高的准确性和鲁棒性,对刚性和非刚性两类目标都具有较好的检测结果。

关键词 目标检测, G-GMM, 运动前景, 图像分割

中图分类号 TP391.41 **文献标识码** A

Research on Moving Objects Detection in Video Sequences Based on Grabcut-gaussian Mixture Model

SHENG Jia-chuan YANG Wei

(School of Polytechnic, Tianjin University of Finance and Economics, Tianjin 300200, China)

Abstract To detect moving objects accurately and rapidly from the videos sequences, this paper proposed a novel G-GMM method for automatic detection via combination of GMM and Grabcut techniques in image processing. Firstly, this algorithm uses GMM(Gaussian Mixture Model) based background subtraction to produce binary images for every moving object and then constructs their minimum marking rectangles. And then it follows the image information initialization of each marking rectangle via Grabcut. Finally, an iterative algorithm with foreground parameters is adopted to optimize the object segmentation and thus the moving object contour is obtained. Experimental results indicate that the proposed method achieves good accuracy and robustness in the still camera outdoor video surveillance system, providing promising detection results for both rigid and non-rigid objects.

Keywords Object detection, Grabcut-gaussian mixture model(G-GMM), Foreground motion, Image segmentation

1 引言

随着视觉监控、人机交互、人脸识别、MPEG 压缩等技术的发展与普及,视频序列中的运动目标检测逐渐成为一个热门的研究领域,在计算机视觉、智能监控系统和智能人机交互等诸多领域得到了广泛的应用。

针对运动目标的检测问题,现有的研究主要包括以下 3 种:帧间差分法^[1,2]实时性较强,实现较简单,但很难分辨出变化不明显的像素点,容易出现物体内部空洞;光流法^[3,4]无需场景的任何信息,根据光流的不连续性特点进行运动目标检测,但是庞大的计算量降低了其实时性;而背景差分法^[5,6]可以检测出比较完整的运动区域,其核心是建立准确的背景模型,在当前帧图像与背景图像的差分结果中检测运动目标,但是对于变化幅度较大的动态场景,其需要及时更新背景。Stauffer 和 Grimson^[7,8]利用混合高斯模型来建立背景模型,该方法在运动目标检测领域得到了广泛的应用^[9-11],常用于静止摄像机户外视频监控系统。然而,基于混合高斯模型背景差分法存在以下两方面的问题:(1)在研究多运动目标检

测时,该检测算法抗噪性能低;(2)对大型、慢速的运动目标的检测效果差,只能获得其部分轮廓。

本文针对静止摄像机户外视频监控系统中混合高斯模型背景差分法存在的不足,提出了新的解决方案。为解决问题(1),本文提出了一种新的 G-GMM 算法,该方法适用于室外和正常光照条件,在给定阈值的前提下,视频序列图像中小面积的噪音点都被设置为背景,抗噪性能强。为解决问题(2),本文采用迭代算法实现最优化分割,可以简单有效地提取自然图像中的目标,在不同场景下都能够准确地检测前景运动目标,因此具有较高的准确性和鲁棒性。

2 Grabcut 算法

Grabcut 算法是一种基于 Graphcut 算法的有效图像分割方法,利用图像中的颜色信息和边界信息,将 Graphcut 算法的应用领域拓展到彩色图像和不完整的掩码图像,使分割精度更高,执行速度更快,操作交互更少。

Grabcut 分割算法主要分 3 步:初始化、迭代最小化和用户编辑。Grabcut 算法的初始化交互只需简单地在目标对象

本文受国家自然科学基金(61502331),天津市应用基础与前沿技术研究计划(15JCQNJC00800),天津财经大学优秀青年学者计划项目,天津财经大学“大学生创新创业训练计划”项目(201410070014)资助。

盛家川(1982-),女,博士,讲师,主要研究方向为多媒体处理、机器学习, E-mail: jc_sheng@163.com(通信作者);杨 巍(1991-),女,硕士,主要研究方向为多媒体处理。

周围画出矩形框,框内的所有像素点初始化为前景,框外的像素点初始化为背景。Grabcut 算法最关键的迭代最小化包含以下步骤:

- (1) 给像素赋 GMM 的初值: 求出未知区域 T_U 中每个像素 i 对应的高斯参数 $k_i, k_i = \arg \min_{k_j} D_i(\alpha_i, k_j, \theta, z_i)$;
- (2) 从图像数据集 z 中获得高斯参数 $\theta, \theta = \arg \min_{\theta} U(\alpha, k, \theta, z)$;
- (3) 用最大流 maxflow 算法求解最小化 Gibbs 能量函数: $\min_{\{\alpha_i, i \in T_U\}} \min_k E(\alpha, k, \theta, z)$;
- (4) 重复步骤(1)–(3), 直至收敛。

用户可以根据图像分割结果添加少量标记以纠正分割, 最终得到精确的分割结果。

在现有的研究^[12,13]中, Grabcut 的初始化需要手动设定掩码图像的初始值, 操作较为复杂且获得的结果具有一定主观性, 并且其迭代方式直接影响分割结果, 因此该算法不能在视频序列分析系统中执行自动分割。针对 Grabcut 算法在视频应用中的不足, 本文采用自动初始化视频图像掩码图像的方式作为 Grabcut 的初始值, 简化了用户交互步骤, 并且能够自动分割出运动目标。

3 G-GMM 算法

在研究过程中发现混合高斯模型背景差分法存在运动目标检测不完整和有噪音点的问题^[14], 为获取更精确的检测结果, 本文提出了一种新的运动目标检测方法: G-GMM 算法。首先采用 GMM 背景差分算法过滤掉无效背景, 然后锁定包含运动目标前景的有效前景区域, 最后再分别对每个前景区域执行 Grabcut 分割操作, 获得完整的运动目标检测结果。

G-GMM 算法主要分为 3 个处理模块: G-GMM 背景建模、初始化潜在前景区域和执行 Grabcut 分割。

3.1 G-GMM 背景建模

在 G-GMM 算法的背景建模过程中, 每个像素点的强度自适应地采用 K 个混合高斯分布来建模(K 的数值根据计算机的处理性能决定, 通过实验, 本文取 $K=5$), 通过一个时间序列的像素点值估计任意像素点 x 的混合高斯分布函数。在任意时刻 t , 对于一个特定像素点 $\{x_0, y_0\}$, 它的历史值是:

$$\{X_1, \dots, X_t\} = \{I(x_0, y_0, i); 1 \leq i \leq t\} \quad (1)$$

当前像素值的概率是当前多维彩色像素分布的高斯概率密度函数的权值总和, 其概率 $P(X_t)$ 为:

$$P(X_t) = \sum_{k=1}^K \omega_{k,t} \times \eta(X_t, \mu_{k,t}, \Sigma_{k,t}) \quad (2)$$

$$\eta(X_t, \mu_{k,t}, \Sigma_{k,t}) = \frac{1}{(2\pi)^{n/2} |\Sigma_{k,t}|^{1/2}} e^{-\frac{1}{2}(X_t - \mu_{k,t})^T \Sigma_{k,t}^{-1} (X_t - \mu_{k,t})} \quad (3)$$

$$\Sigma_{k,t} = \sigma_k^2 I \quad (4)$$

其中, K 是高斯分布个数, $\omega_{k,t}$ 是第 k 个高斯分布在 t 时刻的权值估计值, $\mu_{k,t}$ 是第 k 个高斯分布在 t 时刻的均值, $\Sigma_{k,t}$ 表示第 k 个高斯分布在 t 时刻的协方差矩阵, η 是高斯分布概率密度函数。

3.2 初始化潜在前景区域

在为视频图像中每一个像素点建立 G-GMM 背景模型之后, 用当前帧减去背景帧, 并进行阈值化处理, 获得该视频序列的前景, 然后对检测结果采取平滑处理。对于检测得到的

前景区域存在的小空隙等情况, 对其进行形态学的膨胀处理, 扩充前景, 填补空洞。对于当前检测结果存在的大面积空洞、断层等情况, 本文提出采取寻找目标前景最小外接矩形的方式改善检测结果: 首先, 根据前景区域边缘轮廓线获得最小外接矩形; 然后判断是否为相邻矩形框: 任意两个矩形框之间的垂直距离小于阈值 d , 且它们垂直方向与中心连线的夹角小于阈值 θ (d 和 θ 一般为两个常数, 其值取决于实际场景), 如果满足条件则合并; 最后, 在 RGB 颜色空间中, 获得每个矩形框内子图像的 Grabcut 初始化信息。图 1 为本文算法的流程图。

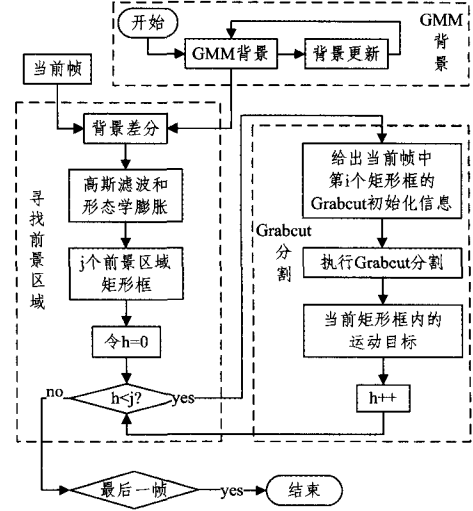


图 1 G-GMM 算法流程图

与传统的 Grabcut 算法相比, 本文算法的初始化不需要用户绘制矩形框进行手工交互, 而是自动标定掩码图像的前景、背景和未知区域, 简化了操作步骤。假设第 s 张视频帧子图像的掩码图像为 $mask = \{M_B, M_U, M_F\}$, 其中 M_B 表示背景区域, M_U 表示未知区域, M_F 表示前景区域。背景区域 M_B 由图像在垂直方向的特征获得。

假设第 s 个矩形框中连通的前景区域集合 $D = \{d_1, \dots, d_n\}$ (n 表示矩形框中连通的前景区域总个数), 则设定前景区域 M_F 为:

$$M_F = D_1 \cup D_2 \cup \dots \cup D_n \quad (5)$$

统计横向的垂直投影直方图, 已知图的垂直像素点个数为 h , 水平像素点个数为 w , 并假设矩形框 r_i 内的二值图像为 $g(x, y)$, 当 $g(x, y) = 1$ 时, 则以 (x, y) 为坐标的像素点是背景像素点; 当 $g(x, y) = 0$ 时, 则是前景像素点。

W_j 表示第 j 列的所有像素值的累加和, 当 W_j 小于给定的阈值时, 则判定第 j 列没有出现前景像素点, 将整个 j 列设置为背景, 如式(6)所示:

$$W_j = \sum_{i=1}^h g(i, j), 1 \leq j \leq w \quad (6)$$

前景和背景并集的补集是未知区域, 其数学表达式如下:

$$M_U = \overline{M_B \cup M_F} \quad (7)$$

设置掩码图像之后, 对其进行初始化, 然后对其子图像执行 Grabcut 分割。

3.3 执行 Grabcut 分割

假设在 RGB 颜色空间, 视频图像由像素点 z_n 组成, $n \in \{1, \dots, N\}$, N 为像素点个数。图像的分割可以由数组 $\alpha = \{\alpha_1, \dots, \alpha_n\}$ 表示, 对于硬分割 $\alpha_n \in \{0, 1\}$, 0 表示背景, 1 表示前景。本文算法对视频序列采用两个 GMM, 分别对应前景

和背景,每个 GMM 由 K 个高斯模型混合组成(本文取 $K=5$)。为了更好地处理高斯混合模型,本文在优化框架中对每一个像素点赋予参数 k_n ,其中 $k_n \in \{1, \dots, K\}$,表示该像素属于特定的 GMM 分组,通过设定 $\alpha_n = 0$ 或 $\alpha_n = 1$ 以界定该分组属于背景或前景。

通过引入 k_n ,Gibbs 能量分割^[15]公式改写为:

$$E(\alpha, k, \theta, z) = U(\alpha, k, \theta, z) + V(\alpha, z) \quad (8)$$

其中, E 表示 Gibbs 能量, U 表示数据, V 表示平滑度, θ 表示背景和前景的灰度直方图: $\theta = \{h(z, \alpha), \alpha = 0, 1\}$ 。

根据颜色高斯混合模型,数据 U 被定义为:

$$U(\alpha, k, \theta, z) = \sum_n D(\alpha_n, k_n, \theta, z_n) \quad (9)$$

在 RGB 空间中,平滑度 V 采用欧几里得距离计算,分割可估计为全局最优:

$$\alpha = \arg \min_{\alpha} E(\alpha, k, \theta, z) \quad (10)$$

本文算法通过多次迭代实现最小化分割,可以有效地提取出完整的运动目标区域,解决了运动目标检测区域存在的“c”字形的边缘残缺、空洞与断层等问题。

4 实验结果与分析

为了验证算法的有效性,本文在实验中采用实际的静止摄像机监控视频序列进行运动目标检测。图 2 是本文对运动目标的检测结果,处理的视频图像序列依次为 Single 序列、Double 序列、Noise 序列和 Traffic 序列(视频序列源于 Visor 和校园监控视频),Single 是对单人的提取,Double 是对多人的提取,Noise 是实验视频中出现噪音点(骑自行车的人)的情况下对运动目标的提取,Traffic 是对汽车的提取。实验设计了以下 3 组对比:单个目标和多个目标,刚性物体和非刚性物体(汽车和人),有噪音和无噪音。由实验结果可以得出,无论是单个目标还是多个目标,运动目标自动检测算法都能很好地分割运动物体与背景图像,结果精度较高。同时,算法对于刚性物体和非刚性物体都有较好的分割结果。当实验视频中出现噪音点时,如图 2 场景 Noise 中有摆动的树叶和远处的路人,在动态背景运动范围不大,即没有大块背景物体持续运动的情况下,本文算法都对其忽略不计,只对前景人物进行目标检测,可见算法具有鲁棒性。

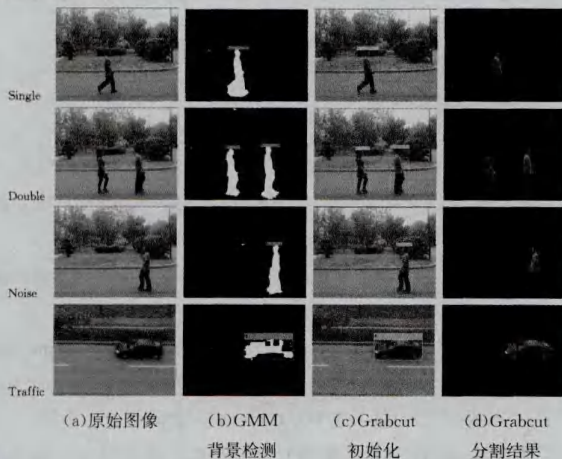


图 2 运动目标检测结果

针对这 4 个不同场景,分别用帧间差分^[16]、Vibe^[17]、GMM^[10]和本文提出的 G-GMM 算法在准确性方面做了分析对比,实验结果如图 3 所示。

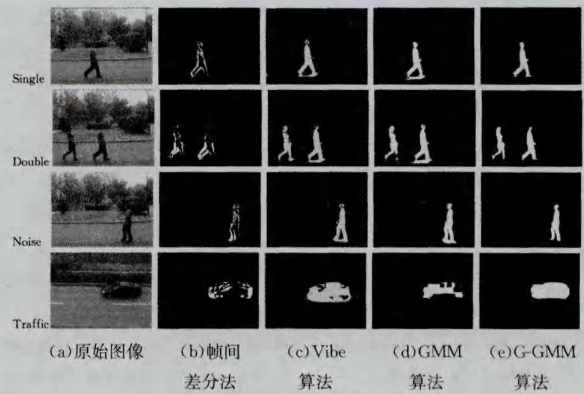


图 3 4 种运动检测算法运行结果对比

受到文献^[18]的启发,本文采用正确率来评价算法的执行结果的准确率。假设算法识别出的像素总数为 R ,真实前景(groundtruth)的像素总数为 GT ,真实前景是人工标定的正确前景,则正确率 ρ 定义如下:

$$\rho = \frac{R \cap GT}{R \cup GT} \quad (11)$$

对比表 1 中数据可以得出,不同的运动检测算法在不同的场景中检测结果的精度各不相同,G-GMM 算法基本适用于各种场景,正确率达到 80%,检测精度相对较高。其中帧差法的正确率整体比较低,约为 30%,虽然该方法容易实现、算法比较简单,但是其整体正确率基本上无法满足生产生活的需要。Vibe 算法正确率约为 50%,GMM 算法约为 70%,这两种算法对噪声敏感,无法有效抑制场景中的干扰,因此检测结果中有噪声点。本文提出的 G-GMM 算法先过滤掉无效背景,再对运动目标区域进行检测,能在过滤掉无效背景的同时过滤掉噪声,并采用迭代算法实现了最优化分割,因此正确率能够达到 80%。实验结果表明,G-GMM 算法能满足准确性和鲁棒性的要求,并取得较好的实验结果。

表 1 各种算法正确率对比(%)

视频	帧差法	Vibe 算法	GMM 算法	G-GMM 算法
Single	35.2	47.34	74.87	75.88
Double	23.13	73.32	79.86	86.22
Noise	36.57	44.21	81.46	87.59
Traffic	24.71	43.61	53.36	85.3
平均值	29.9	52.12	72.39	83.75

结束语 本文针对静止摄像机户外视频监控,提出一种新的 G-GMM 运动目标检测方法。首先用高斯混合模型对背景建模,采用背景差分法获得运动目标的初始轮廓,完成运动目标的粗检测;然后以初始轮廓二值图像作为 trimap 图为 Grabcut 算法提供初始化信息;最后采用迭代算法实现最优化分割,检测得到准确的运动目标轮廓。实验结果表明,在室外和正常光照条件下,该方法具有良好的鲁棒性和准确性。该算法的优点是抗噪性能高,能够完整地检测运动前景目标。今后的研究工作主要是对算法进行优化,引入背景运动补偿技术和多级假设检验等方法以有效地解决复杂地面背景下运动小目标的检测问题。

参考文献

- [1] Khare M, Srivastava R K, Khare A. Moving object segmentation in Daubechies complex wavelet domain[J]. Signal Image and Video Processing, 2015, 9(3): 635-650

- [2] 文嘉俊,徐勇,战荫伟. 基于 AdaBoost 和帧间特征的人数统计[J]. 中国图象图形学报, 2014, 16(9): 1729-1735
- [3] 刘赏,董林芳. 人群运动方向异常检测算法[J]. 计算机科学, 2013, 40(11A): 337-340
- [4] Trulls E, Tsogkas S, Kokkinos I, et al. Segmentation-aware deformable part models[C]// IEEE International Conference on Computer Vision and Pattern Recognition (CVPR). 2014: 168-175
- [5] 张欢,安国成,张凤军,等. 多颜色空间融合的人体检测算法研究[J]. 中国图象图形学报, 2011, 16(10): 1944-1950
- [6] Ramirez-Quintana J A, Chacon-Murguia M I. Self-adaptive SUM-CNN neural system for dynamic object detection in normal and complex scenarios[J]. Pattern Recognition, 2015, 48(4): 1137-1149
- [7] Stauffer C, Grimson W E L. Adaptive background mixture models for real time tracking[C]// IEEE International Conference on Computer Vision and Pattern Recognition (CVPR). 1999: 246-252
- [8] Stauffer C, Grimson W E L. Learning patterns of activity using real-time tracking[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(8): 747-757
- [9] 程全,马军勇. 基于改进高斯混合模型的运动目标检测方法[J]. 计算机科学, 2014, 41(7): 318-321
- [10] Zhao Qin-pei, Ville H, Ismo K. Random swap EM algorithm for Gaussian Mixture Models [J]. Pattern Recognition Letters, 2012, 33(16): 2120-2126
- [11] Nazre B, Rama R. Detection and inpainting of facial wrinkles using texture orientation fields and markov random field modeling[J]. IEEE Transactions on Image Processing, 2014, 23(9): 3773-3788
- [12] 周良芬,何建农. 基于 GrabCut 改进的图像分割算法[J]. 计算机应用, 2013, 33(1): 49-52
- [13] Tao Wen-bing, Li Kun-qian, Sun Kun. SaCoseg: object cosegmentation by shape conformability[J]. IEEE Transactions on Image Processing, 2015, 24(3): 943-955
- [14] 周爱民,张青富,张桂成. 一种基于混合高斯模型的多目标进化算法[J]. 软件学报, 2014, 25(5): 913-928
- [15] Zhu H, Ding M, Li Y. Gibbs phenomenon for fractional Fourier series[J]. IET Signal Processing, 2011, 5(8): 728-738
- [16] Chen Ning, Ding Fei. Flame object segmentation by an improved frame difference method[C]// Third International Conference on Digital Manufacturing and Automation (ICDMA). 2012: 422-425
- [17] Barnich O, Droogenbroeck M V. Vibe: a powerful random technique to estimate the background in video sequences[C]// IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2009: 945-948
- [18] Mofaddel M A, Abd-Elhafiez W M. Fast and accurate approaches for image and moving object segmentation[C]// International Conference on Computer Engineering & Systems (ICES). 2011: 252-259

(上接第 198 页)

- [10] Wang R, Wan B, Ma Z, et al. Computer-aided detection of microcalcifications in digital mammograms using a synthetic technique [C]// International Society for Optics and Photonics Second International Conference on Image and Graphics. 2002: 639-644
- [11] 万柏坤,王瑞平,朱欣,等. SVM 算法及其在乳腺 X 片微钙化点自动检测中的应用[J]. 电子学报, 2004, 32(4): 587-590
- [12] 许向阳. 乳腺钼靶图像中肿块检测方法研究[D]. 武汉: 华中科技大学, 2010
- [13] 朱景升. 基于乳腺 X 射线片的肿块检测方法研究 [D]. 武汉: 华中科技大学, 2012
- [14] 方玲玲. 图像分割的活动轮廓模型研究[D]. 苏州: 苏州大学, 2013
- [15] 李妍. 活动轮廓模型影像分割方法综述[J]. 遥感信息, 2014, 29(1): 102-107
- [16] 王沛,周鑫,彭荣鲲,等. 结合边缘和区域的活动轮廓模型 SAR 图像目标轮廓提取[J]. 中国图像图形学报, 2014, 19(7): 1095-1103
- [17] 吕泽华,赵盛荣,梁虎,等. 基于 Gmac 模型的乳腺肿块分割算法[J]. 电子学报, 2014, 42(2): 398-404
- [18] Timp S, Karssemeijer N. A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography[J]. Medical Physics, 2004, 3(5): 957-971
- [19] 王小芳. 基于活动轮廓模型的图像分割算法研究[D]. 长沙: 中南大学, 2011
- [20] 严学强,叶秀清,刘济林,等. 基于量化图像直方图的最大熵阈值处理算法[J]. 模式识别与人工智能, 1998, 11(3): 352-358
- [21] 俞勇,施鹏飞,赵立初. 基于最小能量的图像分割方法[J]. 红外与激光工程, 1998, 28(4): 20-27
- [22] 程杰. 一种基于直方图的分割方法[J]. 华中理工大学学报, 1999, 27(1): 20-23
- [23] 张建,汪天富,李德玉,等. 基于对称区域生长算法的超声医学图像分割方法[J]. 生物医学工程杂志, 2007, 24(3): 500-503
- [24] 王广君,田金文,柳健. 基于二叉树结构的图像分割技术[J]. 华中科技大学学报, 2001, 30(1): 12-14
- [25] 屈彬,王景熙. 一种基于区域生长规则的快速边缘跟踪算法[J]. 四川大学学报, 2002, 34(2): 100-103
- [26] 杨斌,宋立新. 基于自适应区域生长的乳腺肿块分割方法[J]. 计算机工程与应用, 2014(20): 171-175, 210
- [27] 张深毅. 基于参考图像的乳腺肿块诊断方法研究[D]. 武汉: 华中科技大学, 2011
- [28] 姜雯. 基于乳腺 X 线摄片的计算机辅助检测肿块方法研究[D]. 武汉: 华中科技大学, 2009
- [29] 兰义华. 基于图像内容检索的乳腺肿块诊断方法研究 [D]. 武汉: 华中科技大学, 2011
- [30] 陈桂林,汪家旺. 改良区域生长算法自动分割乳腺肿块图像的断价值[J]. 江苏医药, 2011, 37(13): 1551-1553
- [31] Dominguez A R, Nandi A K. Improved Dynamic Programming-based Algorithms for Segmentation of Masses in Mammograms [J]. Medical Physics, 2007, 34(11): 4256-4269
- [32] 乔剑敏. 基于 GAC 模型和 C-V 模型的图像分割方法的改进 [D]. 哈尔滨: 哈尔滨工业大学, 2011