

基于区域图的智能目标检测

张艳娇,魏志成\*,王艳玲

(河北师范大学信息技术学院, 河北省石家庄市 050024)

**摘要:**提出了一种基于区域图的智能目标检测方法。将颜色稀有和局部边界两个属性通过一个模糊函数整合，在*YIQ*颜色空间下提取图像特征，利用特征构造像素8-邻域关系图，并加入用户动作标示种子像素，最终实现交互式目标与背景的分割。在用户动作硬性约束条件下，利用遗传算法优化区域图中作为软约束的评价函数，寻找图像的最优二值分割，得到全局优化的分割目标和背景。得到的结果保证了颜色和边界的最佳平衡。实验结果表明，颜色稀有和局部边界属性特征与遗传算法的结合优化了检测性能，对多目标的识别效果优于最大流方法和广义霍夫变换算法。

**关键词:**目标检测； 颜色稀有；局部边界；图；评价函数；遗传算法

**中图分类号：TP391 文献标志码：A**

**An Intelligent Object Detection Method Based On Local Graphs**

Zhang Yanjiao, Wei Zhicheng\*, Wang Yanling

(\*corresponding author College of Information Technology, Hebei Normal University, Shi Jiazhuang 050024)

**Abstract**: We propose an intelligent object detection method based on local graphs. Color rarity and local edge property, the soft constraints, which are combined by a fuzzy function are used as image features on YIQ color space. These features are chosen to construct 8 neighborhood graphs. Then a user marks certain pixels as ‘seed’ to provide hard constraints for segmentation. Genetic Algorithm is used to find the globally optimal local graphs cost function of an image. The obtained solution gives the best balance of color and boundary properties. Some experiments with several images indicate that the proposed approach is able to detect the objects better than max-flow method and generalized hough transform algorithm.

Keywords: object detection; color rarity; local edge; graph; evaluation function; genetic algorithm

1. 引言

目标检测是计算机视觉的重要组成部分，已广泛应用于工业、体育、商业、医学等领域中，所以目标检测技术的研究也越来越受到重视。

到目前为止已经有大量关于目标检测研究的技术方法，总结来看可包括目标检测器、背景去除、图像分割等等[1]。目标检测器通常是一个分类器，通过一个滑动窗口扫描图像并定义子图像，把子图像标记为目标或者背景。一般来说，分类器可通过不同数据集的离线学习[2-3]或在视频开始时对一个手动标记帧的在线学习获得[4-5]。背景去除技术则是建立一个背景模型和图像对比，把差异部分检测为目标，该方法常常假设建立背景模型时图像中无目标出现[6]，。典型的方法包括单一高斯分布[7]、混合高斯[8]、k-ernel密度估计[9-10]、块相关性[11]、隐马尔可夫模型[12-13]和线性自回归模型[6] [14]等。但这些方法对训练样本的各种要求则限制了其实用性。图像分割技术中研究比较多的是把一帧图像分割成两部分：目标和背景。图像分割中一些简单的方法，如区域增长[15]，没有定义一个明确的代价函数，在实际应用中鲁棒性较差。于是出现了基于评价函数的分割方法，这种方法可分成两大类：一类是对定义在一个连续轮廓或表面函数的最优化；另一类是对定义在离散变量集的一个代价函数的最优化。第一类方法包括snakes[16]、区域竞争[17]、测地线轮廓法[18]和基于水平集的一些方法[19-20]；第二类方法大体可分为基于路径的和基于区域的，其中一些基于路径的方法采用动态程序设计(DP)计算最优路径，例如Mortensen and Barrett[21]使用Dijkstra算法，Amir et al. [22]采用Viterbi算法，但我们注意到，所有基于路径的方法能够编码基于边界的分割但分割中区域属性却不明显。基于区域的方法中最经典的是图切割方法，它根据图论中最大流最小割定理，利用最大流[23-24]、push-relabel[25]等算法对图进行全局最优分割，得到两个互不交迭顶点集。在[26]中，采用福特-福克森标号法找最大流，它记录每次扩展过程中每个节点的前驱与该节点的增广最大流量，从源点开始直到扩展到汇节点结束，该算法是对最大流算法的简单实现，其缺陷是该方法对多目标图像识别效果并不理想。

近些年来，目标检测技术发展迅速，许多学者提出了针对性的创新方法与改进意见。鉴于灵活性好和学习过程简单的特点，基于霍夫变换的方法[27-28]广泛应用于基于分离的目标检测。Okada[29]利用局部图像块训练目标分离模型，通过投票过程预测目标位置及大小，该方法是对基于霍夫变换的一种概率描述，应用灵活但模型缺乏一致性，存在理论和实践性的问题。考虑在基于霍夫变换的框架中检测多目标距离，Leibe[27]对霍夫变换重塑并融合于评价函数，用目标分割和MDL（Minimum Description Length）删除错误假设。但该算法也涉及对单独目标及其重叠区域的分割假设，使得对框架的概率和评价变得不明确。[30]提出了对目标检测关于霍夫变换概率描述的新框架，绕过霍夫图像中多峰值识别问题，不调用非最大值抑制启发法直接进行多目标检测，但付出了计算量增加的代价。

针对上述方法中存在的问题，本文提出了一种基于区域图的目标检测算法，该算法是对基于区域的分割方法的改进，引入遗传算法对构造图的评价函数进行最优化，并结合颜色稀有和局部边界两个属性提取图像特征，对大量图像进行了仿真实验，该算法显示了良好的性能。用遗传算法优化时还加入了用户动作硬性约束条件，由用户来选择自己感兴趣的作为目标，实验表明该算法的目标、背景检测可包含多个孤立的部分，尤其是在处理多目标图像时效果更为明显。图1是本文提出方法的结构框架图。

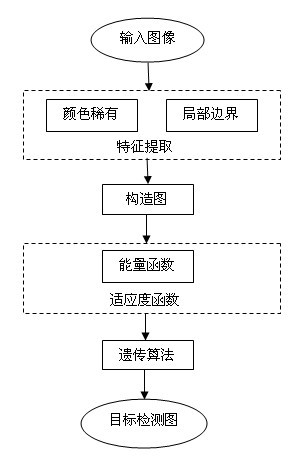


图1 基于区域图的遗传算法实现目标

检测框架图

Fig.1 Graph-based genetic algorithm (Graph-based GA) for object detection frame diagram

本文首先介绍图像特征提取的颜色稀有和局部边界两个属性，然后利用该特征和用户动作标记的目标和背景种子来构造区域图，建立评价函数，利用遗传算法优化评价函数，找到图像帧的最优二进制分割，最后是本文的方法与最大流算法和广义霍夫变换算法进行目标检测结果的比较。

1. 基于区域图的智能目标检测
2. 特征定义

目前常用目标检测方法都是基于强度的，但由于计算强度值时颜色信息被忽略，导致部分目标区域溶解于背景区域中。该问题的存在使得应用其他任何方法之前都降低了目标检测能力。另外在全局图像中，强边界区域会影响弱边界区域的检测。

鉴于上述两个问题，本文在空间下采用颜色稀有和局部边界属性提取特征。因为相比其他颜色空间像RGB、HSV、YcbCr和Ohta[31]，颜色空间会产生更好的分割效果[32]，进而实现高性能的目标检测。颜色稀有是表征那些仅用强度信息不能检测到的区域，像素的颜色稀有属性定义为：

 (1)

分别表示像素的颜色值，整个图像颜色直方图和直方图分布的平均值。

局部边界属性用来检测那些在全局上弱而局部强的边界，该属性用8-邻域窗口评价像素，定义为：

 (2)

其中和的定义如下所示：

 (3)

 (4)

分别是像素依次在三个颜色部分的颜色值、与8-邻域像素的差值以及差值的平均值。

本文用模糊操作将颜色稀有和局部边界属性结合来评价像素，像素的代价函数定义如下：

 (5)

1. 基于区域图的遗传算法实现目标检测
2. 图的构建

基于区域图的目标检测是通过构造一个图，建立图像像素8-邻域的关系，得到图像分割的评价函数，最后用遗传算法获得图像目标和背景的最优分割。

算法要求用户点击待分割图像中感兴趣的部分目标区域和背景区域，作为种子像素。将种子像素和图像普通像素构建成一个图，其中

 (6)

是图的节点，包含图像普通像素构成图的中间节点以及两个辅助节点，构成图的源点和汇点。

是图中节点-节点边集，分为两种类型：和。像素集中像素之间边为构成边集合,每个像素有两个，分别为.因此有

 (7)

记录和边权值，权值分配过程如下表1所示:

表1 图边集的权重分配表

Table 1 Weights of edges for graph

|  |  |  |
| --- | --- | --- |
| 边 | 权值 | 条件 |
|  |  |  |
|  |  |  |
|  |  |
|  |  |
|  |  |  |
|  |  |
|  |  |

其中

 (8)

 (9)

指定图中间节点与边的权重，分别是图像像素的特征值，用来计算像素与的空间距离（本文采用领域）。作为摄像机噪点的参数，当像素和比较相似即时，值较大；反之则越小。因此，参数越大会使差异大的像素被看做属于一个区域，更适宜划分大片的连续区域。

1. 评价函数

本文定义一个二元向量,每个为或，因此向量定义了一个分割，该分割消耗的评价函数定义为：

 (10)

其中

 (11)

在公式(10)中，是区域属性，计算像素与种子像素和的相似程度，进而指定一个，由式(12)(13)给出

 (12)

 (13)

称为边界属性，是像素差异的惩罚因子。评价函数通过参数实现对区域属性和边界属性的权衡。

在本文的目标检测方法中，图像的最优分割就是使得区域图顶点集分割消耗能量最小的分割，考虑到遗传算法具有很强的全局寻优能力，采用遗传算法寻找评价函数的全局最小值。遗传算法中，我们定义适应度越大个体越好，要根据适应度大小对个体进行选择来保证适应度大的个体有更多的机会繁殖后代，使优良特性得以遗传。将公式(10)取倒数

 (14)

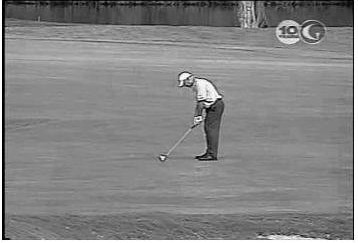
用公式(14)作为遗传算法适应度函数。

1. 实验结果与分析
2. 实验数据集及实验说明

本文仿真实验采用两个数据集和，其中包含99张图像，涉及运动员、动物、车辆、普通人等为目标的内容，目标类型不同；是一段摩托车比赛现场的视频图像，基本上以摩托车为目标，目标类型相似，共589张。为了降低运算成本，算法将输入图像统一处理成大小像素块。基于区域图的遗传算法中设定交叉概率，变异概率，种群染色体条数为50条，迭代循环代数设为3000代。为了验证提出算法的有效性，将本文方法与最大流算法[24]、广义霍夫变换算法进行比较实验。其中提出的方法和最大流都是基于图的目标检测方法，另外为了进一步验证颜色稀有和局部边界属性与遗传算法的结合优化了检测性能，广义的霍夫变换算法中也利用我们提出的颜色稀有和局部边界提取图像特征。

1. 实验结果与分析

考虑到颜色空间相比其他颜色空间具有更好的图像分割效果，本文在空间下结合颜色稀有和局部边界两个属性进行特征提取，图2是利用颜色稀有和局部边界属性进行特征提取的结果图与面向强度方法的灰度图比较,该图像的背景是草地部分,相对于背景而言颜色和边界差异较大，可作为目标的是球员和右上角的台标。由图2可见本文的特征提取方法将图像中的目标都进行了标记，克服了像面向强度提取特征方法最后目标检测结果中部分目标区域溶解于背景区域的问题。而且在本文基于区域图的遗传算法中具有弱边界的台标区域并没有受到强边界目标的影响。



(a)



(b)

图2 高尔夫比赛图像的特征提取结果图: (a)灰度图像;（b）利用颜色稀有和局部边界属性进行特征提取的结果图

Fig.2 Features extraction results of golf image. (a)Gray image; (b) Features extraction image referring to color rarity and local edge property

鉴于本文提出的方法与最大流都是基于图的目标检测方法，先用适应度验证提出方法的优越性。图3给出了基于区域图的遗传算法适应度函数与最大流算法适应度的比较结果。本文算法迭代循环3000代，在第929代时收敛于最大适应度值0.53508，而最大流算法得到适应度值为0.52738。对大量图像进行两种方法的比较结果如表2所示。实验数据证明用本文的算法得到的适应度比最大流算法的适应度值大，本文的算法找到了更优的结果进而保证了优良特性的更好遗传。



图3. 基于区域图的遗传算法适应度函数图

Fig.3 Fitness chart of graph-based GA

表2 基于区域图的遗传算法与最大流算法适应度值比较表

Table 2 The fitness comparison between graph-based GA and max-flow

|  |  |  |  |
| --- | --- | --- | --- |
| 图像 | 大小 | 适应度值 | |
| Graph-based GA | Max-flow |
| Golf | 240352 | 0.35285 | 0.19875 |
| Penguin | 431283 | 0.36497 | 0.33591 |
| Football | 533800 | 0.28342 | 0.18019 |
| Basketball 1 | 500300 | 0.53508 | 0.52738 |
| Basketball 2 | 10241024 | 0.21879 | 0.13994 |

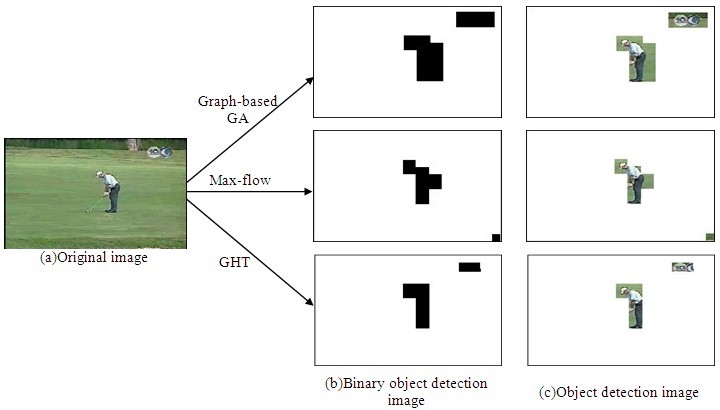
图4是分别针对单目标和多目标情况下，三组不同图像(一组单目标，两组多目标)的目标检测结果，将最大流、广义霍夫变换算法和本文算法进行了实验对比。其中(a)是原始输入图像，(b)为三种不同算法得到的二值目标检测图，最后目标检测结果如(c)所示。由图4(1)可见对单一目标图像进行目标检测时，因为采用颜色稀有和局部边界属性提取图像特征，在基于区域图的遗传算法和广义的霍夫变换算法中图像右上角具有弱边界的台标也被检测出来而最大流算法结果中台标部分却被溶解于背景区域中。由于三种算法中图像是大小像素块处理，目标检测结果会以矩阵块的形式呈现，例如图4(1)中基于区域图的遗传算法选择了相对多的几个矩阵块来检测更完整的高尔夫球员，而广义霍夫变换算法检测的球员区域则不完整；在多目标 (见图4(2)和图4(3))情况下，基于区域图的遗传算法显示了更为优越的目标检测性能。相比最大流算法，基于区域图的遗传算法不仅能够检测出存在弱边界的目标(图4(1))，而且改善了部分目标溶解于背景区域的问题，提升了目标检测效果。由于广义霍夫变换算法对图像中的直线比较敏感，一定程度上影响了不是直线边界的目标区域的检测，例如在图4(3)对企鹅进行目标检测时，作为背景的海浪(直线边界)也被检测为目标区域；而且在同时采用颜色稀有和局部边界提取特征的前提下，加入了用户动作硬性约束来标签种子目标的基于区域图的遗传算法，不仅能够检测出广义霍夫变换漏检的目标，如图4(2)所示，而且能够实现目标与背景的交互式分割，检测性能要优于广义霍夫变换的目标检测。

三种算法目标检测性能的对比参数评价如表3所示，表最后一行是每个算法得到的最优评价结果(黑色字体标注)数量，括号内是次优结果数量。由表可看出基于区域图的遗传算法在数据集和都得到了半数以上(分别是4和3个)的最优评价结果数，优于最大流和广义霍夫变换算法。图5给出了多张图像以及利用本文算法分别得到的的目标检测结果，由图看出基于区域图的遗传算法检测的目标区域既具有很好的全局连通性又具有良好的局部连通性，对于目标关联性很强的图像，本文方法能够检测出连通的多个目标，对于多个目标分离的图像，本文算法也能够识别出局部上孤立的目标区域，兼顾了全局和局部性。

1. 结论

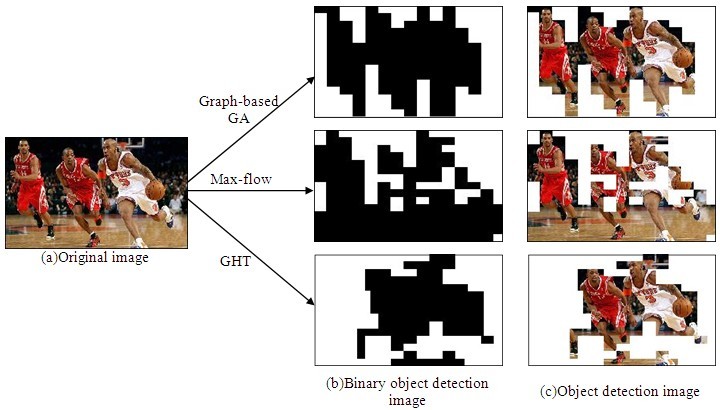
考虑到常见目标检测技术中存在部分目标溶解于背景区域以及强边界区域会影响弱边界检测的问题，本文提出了一种基于区域图的智能目标检测方法。实验证明该方法在多目标检测任务中优于经典的最大流算法和广义霍夫变换算法。硬性条件约束允许用户选择自己感兴趣目标，这样就可以满足不同用户的需求，实现目标与背景的交互式分割。采用遗传算法可以实现对评价函数的全局寻优，使目标检测变得智能化。本文算法是一种普适的算法，在特定应用中，还可以引入一些关于目标的先验知识，如形状、颜色等，这些特征必将提高目标检测的性能。

改善遗传算法的时间复杂度将作为我们接下来的工作，时间复杂度的降低会进一步提高基于区域图智能目标检测算法的性能。针对现在目标检测技术存在的问题(如排除光照干扰、摄像机运动、多场景等)，目前仍没有一种方法可以广泛的应用于各类目标检测中。所以提高检测的鲁棒性、有效性，加入摄像机动作以及多摄像机、多场景中的目标检测等内容将作为本文进一步研究的对象。



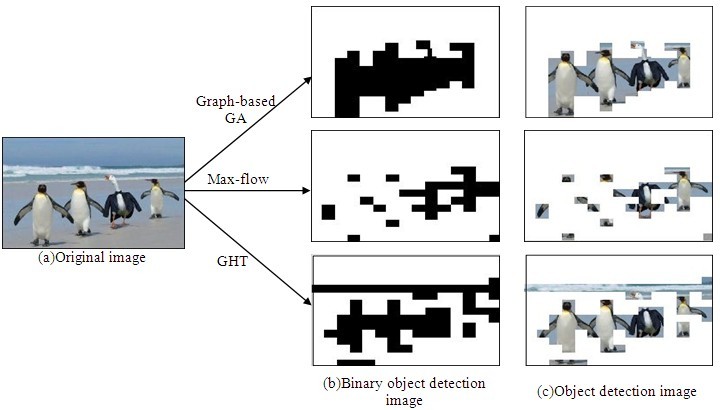
(1)单一目标的高尔夫比赛图像的目标检测

(1)Object detection of golf image with single object



(2)多目标篮球比赛图像的目标检测

(2) Object detection of basketball image with muti-object



(3)多目标企鹅图像的目标检测

(3) Object detection of penguin image with muti-object

图4 单目标与多目标图像下，基于区域图的遗传算法与最大流算法、广义霍夫变换算法的目标检测结果比较图

Fig.4 The comparison between graph-based GA and max-flow、generalized hough transform(GHT) for object detection under single and muti-object respectively

表3 基于区域图的遗传算法、最大流算法和广义霍夫变换三种算法的评价结果表

Table 3 Evaluation results of Graph-based GA, Max-flow and GHT (generalized hough transform)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 评测方法 |  | | |  | | |
| Graph-based GA | Max-flow | GHT | Graph-based GA | Max-flow | GHT |
| Hamming loss | **0.0758** | 0.0821 | 0.1066 | **0.0325** | 0.0597 | 0.0405 |
| accuracy | **0.9242** | 0.8744 | 0.8934 | 0.8516 | **0.8943** | 0.8274 |
| F-measure | **0.9255** | 0.8726 | 0.8912 | **0.8741** | 0.8452 | 0.8735 |
| Balanced accuracy | **0.9243** | 0.8745 | 0.8939 | **0.8726** | 0.8691 | 0.8644 |
| #wins(#better) | 4(0) | 0(1) | 0(3) | 3(1) | 1(1) | 0(2) |



图5 原始图像(包含单目标和多目标)及本文算法得到的目标检测结果

Fig.5 Object detection performance of the proposed method for some images (including single object and muti-object)

**参考文献：**

[1] Yilmaz A, Javed O, Shah M. Object racking: A survey [C]. ACM computing surveys, 2006:1-45.

[2] Papageorgiou C, Oren M., Poggio T. A general framework for object detectionin [C] //Proc. of IEEE International Conference on Computer Vision. 1998,555.

[3] Viola P, Jones M, Snow D. Detecting pedestrians using patterns of motion and appearance [J]. International Journal of Computer Vision, 2005, 63(2):153-161.

[4] Grabner H, Bischof H. On-line boosting and vision [C]//Proc. of IEEE International Conference on Computer Vision and Pattern Recognition. 2006, 260-267.

[5] Babenko B, Yang M.-H., Belongie S. Robust object tracking with online multiple instance learning [J].IEEE Transactions on Pattern Analysis and Machine Intelligence , 2011,33(8):1619 -1632.

[6] Toyama K,Krumm J, Brumitt B, Meyers B. Wallflower: Principles and practice of background maintenance [C] //Proc. of IEEE International Conference on Computer Vision . 1999.

[7] Wren C, Azarbayejani A, Darrell T, Pentland A. Pfind-er: Real-time tracking of the human body [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 780–785.

[8] Stauffer C, Grimson W. Adaptive Background Mixture Models for Real-Time Tracking [C] //Proc. of IEEE International Conference on Computer Vision and Pattern Recognition. 1999.

[9] Elgammal A, Harwood D,Davis L. Non-parametric model for background subtraction [C] //Proc. of IEEE International Conference on Computer Vision . 2000.

[10] Mittal A, Paragios N. Motion-based background subtraction using adaptive kernel density estimation [C] //Proc. of IEEE International Conference on Computer Vision and Pattern Recognition. 2004.

[11] Matsuyama T, Ohya T, Habe H. Background subtraction for non-stationary scenesin [C] //Proc. of Asian Conference on Computer Vision .2000.

[12] Friedman N, Russell S. Image segmentation in video sequences: A probabilistic approach [M]. Uncertainty in Artificial intelligence, 1997.

[13] Rittscher J, Kato J, Joga S, Blake A. A probabilistic back-ground model for tracking [C] //Proc. of IEEE Eur. Conference on Computer Vision. 2000.

[14] Monnet A , Mittal A , Paragios N , Ramesh V .Background modeling and subtraction of dynamic scenes [C] //Proc. of IEEE International Conference on Computer Vision . 2003.

[15] Pallavi V, Mukherjee J, Majumdar Arun K, Sural S. Graph-Based Multiplayer Detection and Tracking in Broadcast Soccer Videos [J]. IEEE Transactions on Multimedia, 2008, 10(5): 794-805.

[16] Kass M., Witkin A, Terzolpoulos D. Snakes: Active con-tour models [J]. International Journal of Computer Vision, 2008, 1(4):321-331.

[17] Zhu S, Yuille A. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(9):884-900.

[18] Caselles V, Kimmel R., Sapiro G. Geodesic active contours [J]. International Journal of Computer Vision, 1997, 22(1):61–79.

[19] Osher S, Fedkiw R. Level Set Methods and Dynamic Implicit Surfaces [M]. Springer Verlag, 2000.

[20] Osher S, Paragios N. Geometric Level Set Methods in Imaging, Vision, and Graphics [M]. Springer Verlag, 2003.

[21] Mortensen E, Barrett W. Interactive segmentation with intelligent scissors [J].Graphical Models and Image Processing, 1998, 60:349-384.

[22] Amini A, Weymouth T, Jain R. Using dynamic programming for solvin variational problems in vision [J] . IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(9):855-867.

[23] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 2004, 26(9): 1124-1137.

[24] Boykov Y, Funka-Lea G. Graph cuts and efficient N-D image segmentation [J]. International Journal of Computer Vision, 2006, 70(2): 109-131.

[25] Goldberg A, Tarjan R. A New Approach to the Maximum-Flow Problem [J]. ACM, 1988, 35(4): 921-940.

[26] Ford L, Fulkerson D. Flows in Networks [M]. Princeton Univ. Press, 1962.

[27] Leibe B, Leonardis A, Schiele B. Robust Object Detection with Interleaved Categorization and Segmentation [J].International Journal of Computer Vision , 2008,77(1-3):258-289 .

[28] Gall J, Lempitsky V. Class-Specific Hough Forests for Object Detection [C]. //Proc. of IEEE International Conference on Computer Vision and Pattern Recognition .2009.

[29] Okada R. Discriminative Generalized Hough Transform for Object Detection[C] //Proc. 12th IEEE Int’l Conf. Computer Vision. 2009.

[30] Barinova O, Lempitsky V, Kholi P. On Detection of Multiple Object Instances Using Hough Transforms [C].//IEEE Conference on Computer Vision and Pattern Recognition, vol. 2008: 259-289.

[31] Ohta Y, Kanade T, Sakai T. Color information for region segmentation [J]. Computer Graphical Models and Image Processing, 1980, 3(13): 22-241.

[32] Mukherjee J. MRF clustering for segmentation of colorimages [J]. Pattern Recognit. Lett., 2002, 23(8): 917-929 .