

# 基于集成学习的交互式图像分割

刘金平<sup>1</sup>, 陈 青<sup>2</sup>, 张 进<sup>2</sup>, 唐朝晖<sup>2</sup>

(1. 湖南师范大学数学与计算机科学学院, 湖南长沙 410081; 2. 中南大学信息科学与工程学院, 湖南长沙 410083)

**摘 要:** 针对交互式图像分割人工标记示例匮乏、不同目标区域难以均衡标记, 单一分类器难以获得有效分割结果的问题, 提出一种多分类器集成学习的交互式图像分割方法. 采用多元自适应回归样条(MARS)方法构造第一个分类器; 同时引入光滑薄板样条回归函数(TPSR)构造与之互补的第二个分类器, 综合组成 bagging 集成学习器, 以降低单一分类器对噪声的敏感度并进一步提高人工标记样本特征空间的利用率. 随后, 基于半监督学习中的聚类假设, 结合 bagging 多学习器并联特点, 提出一种 REG-Boosting 半监督学习算法, 实现半监督图像分割. 在不同数据集上的验证性和对比性实验表明所提方法的有效性和优越性.

**关键词:** 交互式图像分割; 多元自适应回归样条; 集成学习; 薄板样条回归; 半监督学习

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 0372-2112 (2016)07-1649-07

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2016.07.019

## Interactive Image Segmentation Based on Ensemble Learning

LIU Jin-ping<sup>1</sup>, CHEN Qing<sup>2</sup>, ZHANG Jin<sup>2</sup>, TANG Zhao-hui<sup>2</sup>

(1. College of Mathematics and Computer Science, Hunan Normal University, Changsha, Hunan 410081, China;

2. School of Information Science and Engineering, Central South University, Changsha, Hunan 410083, China)

**Abstract:** A kind of interactive image segmentation method based on ensemble multi-classifiers is put forward to solve the problem of unsatisfactory segmentation results based on scarce or unbalanced labelling labels on different object areas by single learner. The first classifier is established based on multivariate adaptive regression splines (MARS) method. A complementary thin plate spline regression (TPSR) classifier is simultaneously established. By combination of these two classifiers, a bagging ensemble learner is achieved to reduce the noise sensitivity and make further efforts of improving the use of the feature space information of the labeling samples. Ultimately, a kind of REG-Boosting algorithm for semi-supervised image segmentation is put forward based on the clustering hypothesis in the ensemble learning combining with the parallel characteristic of the bagging multi-learners. Abundant validation experiments and comparative experiments on different test sets confirm the effectiveness and out-performance of the proposed method.

**Key words:** interactive image segmentation; multivariate adaptive regression splines (MARS); ensemble learning; thin-plate spline regression (TPSR); semi-supervised learning

## 1 引言

图像分割是计算机视觉处理中的一项基本任务, 为完成图像中复杂目标的精细分割往往需要借助于人工交互<sup>[1]</sup>. 从本质上来说交互式图像分割是典型的模式分类(识别)问题, 研究者已经提出了许多相关的交互式图像分割方法, 比如支持向量机(SVM)<sup>[2]</sup>、割图(Graph-Cut, GC)<sup>[3]</sup>、随机游走(Random Walk, RW)<sup>[4]</sup>等. 在交互式图像分割中, 对于那些目标形态复杂、表面

细节丰富的图像, 往往需要大量增加手工标签才能获得有效的分割结果. 然而, 详尽的手动标记会给使用者带来极大不便, 也给后续的分类器学习带来沉重的计算负担, 从而严重制约交互式图像分割的实际应用.

在交互式图像分割中, 图像中除了少量的人工标记, 绝大多数是廉价的未标记像素点信息, 因此, 要想获得理想的分割结果, 有效的交互式图像分割方法显然需紧扣以下两点: (1) 最大限度利用人工标记样本点的区域分类信息; (2) 充分利用大量未标记像素点所蕴含

的分类信息,也就是希望在充分考虑有限人工标记像素点信息的同时,通过引入半监督学习的思想,充分考虑大量廉价的未标记像素点的区域分割信息,以获得更为精细的分割结果。

当前的基于单分类器学习的交互式图像分割方法还存在改进的空间.本文细致分析了文献[5]中所描述的基于薄板样条回归(Thin Plate Spline Regression, TPSR)的交互式图像分割算法的性能,发现 TPSR 方法仅从标记像素点特征向量的整体描述上挖掘各目标区域的分类信息,忽略了各特征向量因子间的局部相互作用(本文将样本标记点的特征向量的任意一维称为一个参量因子),这和最大限度利用手动标记样本点信息这一原则还存在一定差距。

与 TPSR 方法不同的是,多元自适应回归样条(Multivariate Adaptive Regression Splines, MARS)<sup>[6]</sup>方法在利用了特征参量单因子的贡献度的同时,综合利用了多个因子之间交叉作用.因此,从理论上来说,在相同手工标记下 MARS 方法可挖掘更多的分类信息,也将获得更好的分割效果.但是,如果仅采用 MARS 方法进行半监督图像分割,却性能难以稳定.其根本原因是, MARS 分类器在重复利用特征参量各因子间的交互作用同时,单个或数个特征参量因子间的干扰信息被累积放大了,因此 MARS 方法容易受到噪声干扰.当然,不管是 TPSR 方法、MARS 方法还是其他经典的交互式图像分割方法(比如 GC),如果人工标记点过少或者分布不均衡的话,这些常用的交互式图像分割方法往往难以获得有效的分割结果。

为改善基于单分类器的交互式图像分割方法的有效性和鲁棒性,本文引入集成学习<sup>[7]</sup>思想,将具有较好抗噪性能且具备准确分类能力的 TPSR 分类器与对噪声敏感却综合利用了特征向量因子之间交叉作用的 MARS 分类器并联组合,将集成学习中的 Boosting<sup>[8]</sup>策略引入到像素点集成分类器的训练中.同时,基于半监督学习中的聚类假设原则,进一步选择置信度足够高的未标记像素点映射到已标记的示例集合并供 TPSR 和 MARS 分类器进行模型更新,以进一步提高交互式分割模型的区域分割准确性.论文在几个经典的图像数据集上进行了算法性能验证与参数调整测试,并与其经典的交互式分割方法进行了性能对比。

## 2 集成学习交互式图像分割

### 2.1 交互式图像分割问题描述

图像分割是要通过一定的图像像素点分类模型 $h(\cdot)$ 给图像中的任意像素点 $p_i$ 分配合适的区域分割标记 $O_{p_i}$ .交互式图像分割通过借助人工标记的具有明确区域标签的图像样本像素点集 $U = \{x_i, y_i | 1 \leq i \leq n, y_i = 1, 2, 3, \dots, K\}$ 进

行分类器学习( $K$ 代表图像中目标区域的个数, $n$ 表示人工标记点的个数)获得相应的分类模型 $h$ ,进而可以获得任意像素点 $p_i$ 对应的目标标签响应 $y_{p_i}$ ,即

$$y_{p_i} = h(x_{p_i}) + \varepsilon \quad (1)$$

其中 $\varepsilon$ 表示模型残差, $x_{p_i}$ 表示像素点 $p_i$ 的特征向量.公式(1)表明,像素点 $p_i$ 与其对应的像素点的区域类型响应标记 $y_i$ 存在某种映射关系.这种映射关系可以用一组分段光滑的低次多项式去逼近。

### 2.2 多分类器构造

本文从独立且互补的视角构建两个学习器 $h_1(x)$ 、 $h_2(x)$ (为便于描述,这里统一将基于回归函数的分类器称为学习器),并将其并联组合成集成学习模型 $H(x)$ .由于这两个学习器因是从互补的视角构建的,那么这两个学习器的分类偏差都偏大的概率会比较小,这是集成分类模型与单一分类器模型相比所具有的最大优势。

#### 2.2.1 多元自适应回归样条(MARS)方法

MARS 方法是一种非线性、非参数的局部回归方法,通过样条函数来模拟复杂非线性关系,具备准确分类的优点,在模式识别、系统辨识、过程控制等领域得到广泛应用<sup>[6]</sup>,其一般形式为

$$h_1(x) = \lambda_0 + \sum_{m=1}^M \lambda_m \cdot \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)} - t_{km})]_+ \quad (2)$$

其中 $M$ 是样条基的个数, $\lambda = \{\lambda_0, \lambda_1, \dots, \lambda_M\}$ 是输出权重, $K_m$ 是第 $m$ 个样条的分段次数, $t_{km}$ 是样条节点或者说分割阈值, $v(k,m) \in \{1, 2, \dots, p\}$ 代表 $p$ 个预测子中哪一个进入了第 $m$ 个样条中的第 $k$ 个分段中, $s_{km} \in \{+1, -1\}$ 决定了样条分割的方向, $[s_{km}(x_{v(k,m)} - t_{km})]_+$ 为半截多项式,即

$$[x_{v(k,m)} - t_{km}]_+ = \begin{cases} x_{v(k,m)} - t_{km}, & x_{v(k,m)} > t_{km} \\ 0, & \text{其它} \end{cases} \quad (3)$$

MARS 模型构建算法包括前向逐步选择基函数、剪枝和最优模型确定这三个过程.每次选取一对最优的样条函数来提高模型的准确度,建模过程不断重复,直到达到用户给定基函数个数的最大值 $M$ 和交互的基函数数目最大值 $N$ .剪枝过程基于广义交互验证标准(GCV)进行。

MARS 的基函数是单一样条函数或者两个(多个)样条函数的交互结果,因此, $h_1(x)$ 可分解成如下便于分析的直观形式:

$$h_1(x) = \lambda_0 + \sum_{k_i=1} h_{1_i}(x_i) + \sum_{k_{i,j}=2} h_{1_{i,j}}(x_i, x_j) + \sum_{k_{i,j,k}=3} h_{1_{i,j,k}}(x_i, x_j, x_k) + \dots \quad (4)$$

其中,第1个累加项是所有只包含单因子作用的基函数的累加,第2个累加项是所有包含且仅含有两个因子相互作用的基函数的和,依此类推,第 $m$ 个累加和是所有

包含且仅含  $m$  个因子相互作用的基函数的累加和。

式(4)表明, MARS 方法从标记样本点特征参量的多因子角度建立回归模型, 综合利用了多个因子之间的协同作用, 深入挖掘了隐藏在多维特征参量中的复杂信息结构。

### 2.2.2 薄板样条回归 (TPSR) 方法

基于 TPSR 方法的学习器  $h_2(x)$  是在兼顾插值准确性与样条平滑性的同时, 从 Sobolev 空间构造的一种回归函数分类器<sup>[5]</sup>, 该分类器通过通过最小化  $J(h_2)$  获得:

$$J(h_2) = \sum_{i=1}^{N_o} [y_i - h_2(x_i^{O_r})]^2 + \sum_{i=1}^{N_o} [y_i - h_2(x_i^{O_s})]^2 + \lambda S(h_2) \quad (5)$$

其中  $x_i^{O_r}$  和  $x_i^{O_s}$  分别为图像中前景和背景样本点所对应的图像特征,  $S(h_2)$  为  $h_2$  的光滑度惩罚函数:

$$S(h_2) = \int_{\mathbb{R}^d} |D^2(f)|^2 dx = \int_{\mathbb{R}^d} \sum_{i=1}^d \sum_{j \neq i}^d \left( h_{2,x_i x_j}^2 + 2h_{2,x_i x_j}^2 + h_{2,x_i x_j}^2 \right) dx_1 dx_2 \cdots dx_d \quad (6)$$

研究表明, 满足式(5)在约束条件(6)下的解为格林公式

$$\varphi(r) = \begin{cases} r^{4-d} \ln r, & d=2 \text{ 或 } d=4 \\ r^{4-d}, & \text{others} \end{cases}, r = \|x - x_i\| \quad (7)$$

上式中  $d$  表示函数输入参数  $x_i$  的维数,  $D(y)$  代表对函数求导. 通过最小化  $J(h_2)$  得到理想的目标函数的最优估计  $\hat{h}_2$ .  $\hat{h}_2$  的某种唯一样条函数形式可使  $J(\hat{h}_2)$  达到最小值, 即

$$\hat{h}_2(x) = \sum_{i=1}^l \beta_i p_i(x) + \sum_{j=1}^m \alpha_j \varphi_j(x) \quad (8)$$

从而得到 TPSR 分类器  $h_2(x)$

$$h_2(x) = \beta_0 + \sum_{i=1}^d \beta_i x_i + \sum_{j=1}^{N_o} \alpha_j^{O_r} \varphi_j^{O_r}(x) + \sum_{j=1}^{N_o} \alpha_j^{O_s} \varphi_j^{O_s}(x) \quad (9)$$

其中  $\varphi_j^{O_r} = (\|x - x_j^{O_r}\|)^2 \log(\|x - x_j^{O_r}\|)$ ,  $\varphi_j^{O_s} = (\|x - x_j^{O_s}\|)^2 \log(\|x - x_j^{O_s}\|)$ .

TPSR 方法中格林解  $\varphi(r)$  与有标记训练样本关联, 因为  $r = \|x - x_i\|$  是  $x$  与  $x_i$  之间的距离, 它无法利用样本特征向量的单个因子乃至至少于  $d$  个因子的协同作用, 但是 TPSR 方法是从特征向量的整体上(也即基于样本示例的空间分布)建立相应的回归模型, 与综合利用多个因子间协调作用的 MARS 方法具有互补性。

### 2.2.3 分类器集成

最简单的集成模型为两个分类器的加权形式, 比

如,  $H(x) = [h_1(x) + h_2(x)]/2$ . 基于 TPSR 和 MARS 方法的性质, 从理论上来说, 这种简单的加权模型不仅利用了特征矢量各因子的单独作用和交互作用, 还利用了特征矢量在空间上的分布特征, 这在使用样本特征的深度和广度上都是单个 MARS 模型或 TPSR 模型所无法比拟的. 因此, 在抗噪的方面, 集成模型  $H(x)$  对同一噪声同时达到突变点的概率远低于单一模型. 图 1 显示了在不同噪声水平和不同的手工标记下采用 MARS 分类器和 TPSR 分类器进行交互式图像分割的结果。

图 1 中的  $\sigma$  代表噪声水平. 从图中可以很明显看出, 在不带噪声情况下综合利用了特征向量多因子间交互作用的 MARS 方法与 TPSR 方法相比, 能获得更好的分割结果, 但是 MARS 方法对噪声比较敏感, 而 TPSR 方法则相对稳定. 因此, 采用简单的集成模型  $H$  就可以有效提高图像分割的性能. 但是, 如果人工标签信息过少, 少量噪声污染, 将使图像分割性能明显下降。

因此, 有必要进一步考虑大量的未标记像素点的分割信息对分割模型的贡献程度, 以便能在较少的标签下获得更好的分割结果。

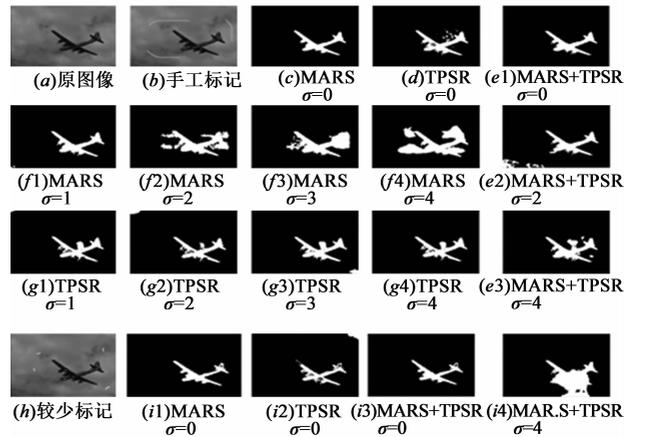


图1 飞机图像交互式分割结果

## 2.3 半监督学习策略

### 2.3.1 算法思想

半监督分类学习中的协同训练把图像中大量廉价的未标记像素点通过试分配合理的目标区域标签当作手工标记样本点以重新训练分类器. 本文受 Zhou<sup>[9]</sup> 提出的基于协同训练的半监督曲线回归方法 COREG 的启发, 结合 bagging 并联学习器的特点, 提出一种既不需要充分冗余视图, 也不需要通过十倍交叉验证来评价像素点区域标签分配信度的 REG-Boosting 算法, 以实现半监督的交互式图像分割. 算法的总体思想是有放回地多次从未标记集合中随机选择  $M'$  个未标记的待标记样例, 采用预先构造的 MARS 分类器和 TPSR 分类器分别进行标签预标记, 选择其中置信度高的样本点送

入 MARS 分类器和 TPSR 分类器进行分类器模型更新, 最终获得性能良好的半监督分类器。

### 2.3.2 REG-boosting 算法流程

REG-boosting 算法流程如下:

(1) 根据用户手工标记指定的样本像素构建前景特征集  $\{x_i^{0^r}, 1\}_{i=1}^{N_{o^r}}$  和背景特征集  $\{x_j^{0^s}, -1\}_{j=1}^{N_{o^s}}$ ; 其中  $+1$  和  $-1$  分别代表前景和背景区域的分割标记响应值,  $x_i^{0^r}$  和  $x_j^{0^s}$  分别代表前景中第  $i$  个标记点和背景区域中第  $j$  个标记点的特征向量,  $N_{o^r}$  和  $N_{o^s}$  分别代表背景和前景区域中人工手动标记的样本点数目;

(2) 分别对  $\{x_i^{0^r}, 1\}_{i=1}^{N_{o^r}}$ 、 $\{x_j^{0^s}, -1\}_{j=1}^{N_{o^s}}$  采用 kNN 方法聚成  $k$  个类得前景和背景区域的像素特征集  $\{c_i^{0^r}, 1\}_{i=1}^k$  和  $\{c_j^{0^s}, -1\}_{j=1}^k$ ;

(3) 基于  $\{c_i^{0^r}, 1\}_{i=1}^k$  和  $\{c_j^{0^s}, -1\}_{j=1}^k$  构建 MARS 分类学习器  $h_1$  和 TPSR 分类学习器  $h_2$ ;

(4) 在未标记像素点集合  $U = \{x_1, x_2, \dots, x_M\}$  中随机选择  $M'$  个未标记样例构成待测样例  $U'$ ;

(5) 评价将待测样例  $x_i \in U'$  分配目标区域标签后可以当作手工标记样本像素点的置信度:

(a) 分别计算  $h_1(x_i)$  和  $h_2(x_i)$ , 以 0 为阈值评价  $h_1(x_i)$  和  $h_2(x_i)$  是否属于同一区域分割标签, 若属于同一区域标签则转(b); 若不是, 则继续对  $x_{i+1}$  执行该操作过程, 直至取完  $U'$  中全部元素;

(b) 若  $h_1$  和  $h_2$  对  $x_i$  均标记为前景, 即  $L_{x_i} = 1$ , 则将  $x_i$  聚类到  $\{c_i^{0^r}, 1\}_{i=1}^k$ ; 同理, 若  $h_1$  和  $h_2$  对  $x_i$  均标记为背景 ( $L_{x_i} = -1$ ), 则将  $x_i$  聚类到  $\{c_j^{0^s}, -1\}_{j=1}^k$  中, 并记录下  $x_i$  所属类别的所有元素. 聚类过程中, 若  $\{c_i^{L_{x_i}}, 1\}_{i=1}^k$  依 kNN 法则不能接受  $x_i$ , 则将  $L_{x_i}$  中所有元素重新聚成  $k$  个类;

(c) 采用均方误差规则  $MSE(h_i)$  评价  $x_i \in U'$  分配目标区域标签后可以当作手工标记样本点的置信度. 若满足条件, 则采用  $h_1^*(x)$ 、 $h_2^*(x)$  替代  $h_1(x)$ 、 $h_2(x_i)$ ; 否则拒绝  $h_1^*(x)$ 、 $h_2^*(x)$  替代  $h_1(x)$ 、 $h_2(x_i)$ , 并从  $U'$  选择  $x_{i+1}$  继续执行步骤(a), 直至取完  $U'$  中全部元素.  $MSE(h)$  计算规则如下:

$$MSE(h_i) = \sum_{i=1}^{2k+N(c_i)} \left\{ \begin{array}{l} [y_i - h_i(x_i)]^2 \\ [y_i - h_i^*(x_i)]^2 \end{array} \right\} \geq 0 \quad (10)$$

(6) 算法训练结束, 对余下未标记像素使用最终的集成学习器  $H(x) = [h_1(x) + h_2(x)]/2$  分配目标区域标签.

由于 REG-Boosting 从未标记集合  $U$  中随机选择了  $M'$  个未标记的待测样例  $U'$ , 这就保证了未标记示例相较于用户手工标记的训练示例有较大的变异成分, 丰富了目标与背景的差异性信息. 同时, 只有在式  $MSE(h_1)$  和  $MSE(h_2)$  同时满足, 即均方误差朝着减小的方向变化, 才可判定对未标记像素点  $x_i$  分配目标区域标签

后当作手工标记样本像素点来更新学习器, 否则不利用  $x_i$  的预标记结果进行学习器更新.

REG-Boosting 算法不同于 COREG<sup>[7]</sup> 的交叉验证的学习模式, 主要是为了保证分类模型在再学习的过程中尽可能少地受噪点干扰. COREG 对置信点的交叉验证是通过一个学习器做均方误差验证, 而实际情况是: 用于做均方误差的学习器是基于少量样本训练出来的, 其鲁棒性较差, 难免把噪点错误地投影到有标记示例集合. 相比之下, REG-Boosting 通过两个互补的学习器共同验证, 这在滤除噪声点的规则上更为严格, 并且在做均方误差验证时是从多个角度同时进行的, 这就保证了后续的参数更新不会对伪目标形成误判. 这个类似于集成学习中弱分类器经过不断地 Boosting 训练成强分类器.

## 3 实验验证

实验验证主要包含三大部分: (1) 验证 MARS 分类器与 TPSR 分类器在交互式图像分割中的互补特性, 证明所提算法的有效性; (2) 分析不同参数设计对算法性能的影响; (3) 与该领域中的经典算法如 SVM、TPSR、GraphCut (GC) 的性能比较, 验证本文方法的有效性和性能优越性.

实验中所选择的自然场景图像数据库来自于 Berkeley Segmentation Data Set (BSDS500)<sup>[10]</sup> 和 Pascal Visual Object Classes (Pascal VOC) 数据库<sup>[11]</sup>. 至于图像像素点特征向量描述, 目前已经有相当多的特征提取方法, 比如最简单的基于图像 RGB 颜色特征, 另外还有 Gabor 小波特征、基于稠密采样的一些局部特征描述符等. 根据 Xiang<sup>[5]</sup> 的研究结果, 对图像进行一些频率滤波处理后提取的特征并不一定能有效改善交互式图像分割的性能. 因此, 在对比实验中, 本文只选用最简单的 RGB 颜色信息和位置信息共同组成像素点特征向量, 即图中任意像素点的特征向量采用一个 5 维的特征向量来描述.

### 3.1 MARS 分类器与 TPSR 分类器的互补性分析

TPSR 分类器可以充分利用手工标签样本点特征向量的整体空间结构信息, 因此, 从理论上说, TPSR 方法在绝大多数自然场景图像的交互式分割中均能取得较好的分割效果, 但该方法的准确性严重依赖手工标记位置的选择及人工样本标记的个数, 图 2 中实验结果也正好证明了这一点.

如图 2 中的蝴蝶图像, 当蝴蝶标记位置发生改变时, TPSR 的分割结果也随之发生了大的变化. 当目标多、形状复杂多变时, 要想获得更好的分割结果, TPSR 方法必须通过大量增加标签数目, 如图 2 中的多人物目标图像的分割结果.

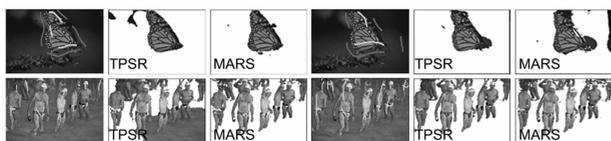


图2 不同位置手工标记以及标记样本数目对分割性能影响

MARS方法与TPSR方法具有较好互补性.在相同的手工标记条件下,MARS方法难以分割的图像,TPSR方法往往能得到不错的分割结果;同样,对于TPSR方法难以分割的图像,MARS方法也能获得很好的结果,如对图3中的鸟、岩石、蝴蝶图像.

另外,从图3中对汽车、房子、女人的图像分割结果还可以明显看出,即便是MARS方法与TPSR方法均不能完全准确分割的图像,但MARS方法与TPSR方法各自分错的图像块几乎不重叠.这时只需对MARS和TPSR所分割出来的图像加以集成即可获取跟手工分割图像相媲美的结果.

图2和图3中的分割结果表明MARS方法和TPSR方法具有明显的互补特性.MARS方法善于利用特征参量因子间的关系,在同样贫乏的标记点条件下,MARS方法能获得相对较好的分割结果.而TPSR从特征参量的整体上进行分类器训练,因此对噪声敏感度比MARS方法低,图1的实验结果已经验证了这一点.

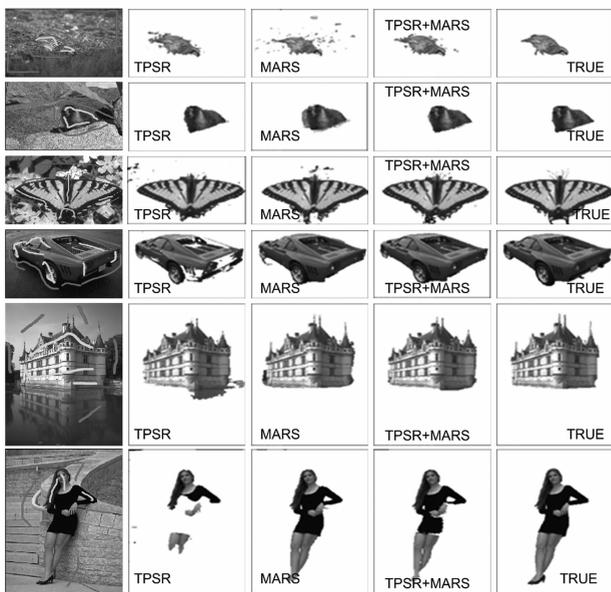


图3 MARS和TPSR方法在交互式图像分割中的互补性表现

### 3.2 参数设置对图像分割性能的影响

#### 3.2.1 特征聚类数目 $k$ 对图像分割性能的影响

目标与背景区域的特征聚类数目  $k$  是本文图像分割算法中第一需设定的参数.图4给出了在不同聚类数目  $k$  下,采用本文方法进行交互式图像分割的结果.为了更直观地展示不同聚类数目  $k$  对分割准确性的影响,

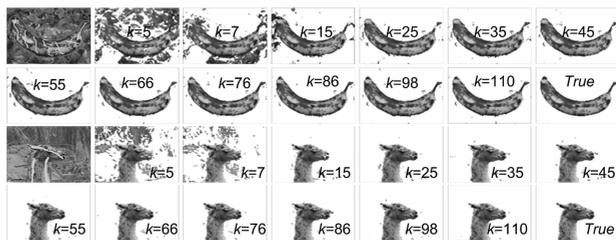


图4 不同聚类数目下的图像分割结果

图5中的分割准确性测量曲线显示了不同  $k$  值下的TPSR方法、MARS方法和本文提出的集成学习的交互式图像分割方法的准确度测量结果.从中可以看出,香蕉图像在  $k=66$  时逼近最优分割结果,而羊驼图像在  $k=45$  时就已逼近最优分割结果.

不难发现,不同的自然场景图像达到最优分割结果的  $k$  值是不同的,这主要与自然场景图像的复杂程度有关.图像中目标形态越复杂其特征就越丰富,需要更多聚类来刻画这些差异较大的图像特征.但并不是  $k$  值设置的越大就越好,如图5所示,在  $k$  值达到某一临界时,再增大  $k$ ,对图像分割的准确性几乎不再提高;并且随着  $k$  值的增大,分类器学习和判断的时间开销也随之增大.大量的分割实验表明,当  $k$  处于 45 到 86 的区间时,能可靠分割出自然场景图像中的绝大多数目标.基于这一实验结果,在后续的与其它方法进行性能比较时,本文的交互式图像分割算法中的特征聚类数目  $k$  取为 45 ~ 86 间的值.

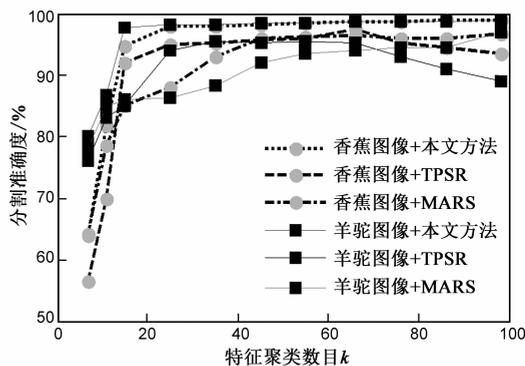


图5 聚类数目  $k$  对图像分割性能的影响

#### 3.2.2 随机选择未标记点数目 $M'$ 对半监督图像分割性能影响

本文所提出的半监督图像分割算法的另一参数是从未标记集合中随机选择  $M'$  个未标记待测样例进行 REG-Boosting 学习.图6展示了在不同取值下,本文提出的基于集成学习的图像分割算法的图像分割结果.从图中可以发现,在同等条件下,  $M'$  越大,被错分的细小区域就越少.因此,为了获得更好的分割性能,应倾向于让更多的未标记像素点有机会加入到标记示例中进

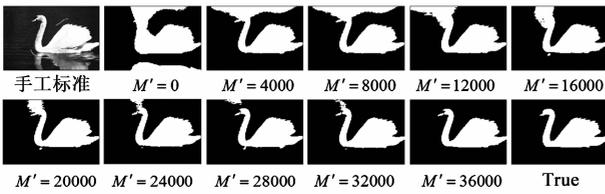


图6 用于REG-Boosting的未标记像素数目对集成学习交互式图像分割准确性的影响

行分类模型更新。

### 3.3 与其它经典算法的性能比较

将本文所提出的图像分割方法与目前经典的交互式图像分割算法 SVM、TPSR、GraphCut (GC)<sup>[12]</sup> 方法进行了性能对比。本文提到的 MARS 方法、TPSR 方法也分别作为独立的分割方法纳入到对比分析中。实验中 MARS 选用 ARESLab 包<sup>[13]</sup>, TPSR<sup>[5]</sup> 和 GraphCut (GC) 是直接对应作者的主页上下载的程序包, SVM 库选用 Chang 开发的 libSVM<sup>[14]</sup> 包, 对自然图像的目标与背景区域的特征聚类数目  $k$  值均设定为 86, 在进行 REG-boosting 训练时,  $M'$  设为 10000。

图 7 和图 8 分别是 BSDS500 数据库和 Pascal VOC 图像库中自然场景图像的实验结果, 同时为了定量比较不同算法针对每幅图像的分割准确性, 还计算了不同分割方法的分割准确度, 如表 1 所示。

对比结果表明, 本文提出的基于集成学习的图像分割算法的性能明显优于单纯只使 TPSR 分类器和 MARS 分类器的单分类器分割方法。在分割过程中, MARS 方法和 TPSR 方法对于大部分错误分割的区域回归值的绝对值在 0 附近, 而能准确分割的区域其回归值的绝对值则逼近于 1 甚至大于 1。因此, 当将 MARS 和

TPSR 所得到的结果进行集成时, 任一学习器的细小回归值将被另一学习器的较大回归值修正, 最终的分割结果也就被修正。基于 SVM 的交互式图像分割方法对前景或背景比较简单的环境下可以获得不错的分割结果, 对于复杂图像则往往难以满足实际要求。究其原因, 主要是 SVM 仅仅考虑前景与背景样本像素特征空间的最大间距。与经典的图像分割算法 GC 相比, 对于一般的自然图像, 如果 GC 能获得较好分割结果的图像, 本文方法同样能获得好的分割结果。从图像分割的视觉结果可以看出, GC 的分割结果一般比较平滑且几乎没有孤立的错割细小区域, 究其原因, 主要是该方法既利用了像素特征的空间信息又结合了待割像素与种子点的相似度, 但是对于前景和背景区域比较相似的图像, GC 容易产生较大的分割误差。

表 1 图 7 和图 8 中 8 幅图像的分割准确性测量

| 测试图像 | 方法选择   |        |        |        |        |
|------|--------|--------|--------|--------|--------|
|      | SVM    | TPSR   | MARS   | GC     | 本文方法   |
| 1    | 88.59% | 79.89% | 99.29% | 98.79% | 99.50% |
| 2    | 98.08% | 98.99% | 90.87% | 98.52% | 99.14% |
| 3    | 97.03% | 96.57% | 98.68% | 92.29% | 99.35% |
| 4    | 93.76% | 96.15% | 95.03% | 94.56% | 96.83% |
| 5    | 96.00% | 98.26% | 97.00% | 97.77% | 98.35% |
| 6    | 97.73% | 98.68% | 97.89% | 95.81% | 98.86% |
| 7    | 98.85% | 92.89% | 97.78% | 35.25% | 99.19% |
| 8    | 97.38% | 99.04% | 97.21% | 97.31% | 99.05% |

本文提出的图像分割算法集成了两个互补的分类器, 在最大限度利用了人工标记像素点分类信息的同时通过 REG-boosting 算法充分挖掘了未标记像素点的分类信息, 因此, 在同样的人工标记条件下, 本文提出的基于集成学习的交互式半监督图像分割方法对那些形状复杂、表面细节丰富的图像能获得较 GC、SVM、TPSR 和 MARS 更细致的分割结果。

## 4 结论

介绍了一种基于集成学习的交互式半监督图像分割方法, 将协同训练的思想引入到半监督的交互式图像分割中, 采用 Bagging 方法集成对有限标签样本数据信息利用率高的 MARS 分类器和对噪声敏感度相对较低的 TPSR 分类器, 并基于半监督学习中的聚类假设, 提出一种 REG-Boosting 半监督分类器学习以实现交互式图像分割。通过大量的验证性和对比性实验, 验证了所提方法的有效性和性能优越性, 并分析了不同参数设定对本文算法性能的影响。

### 参考文献

[1] 罗希平, 田捷. 一种改进的交互式医学图像序列分割方法

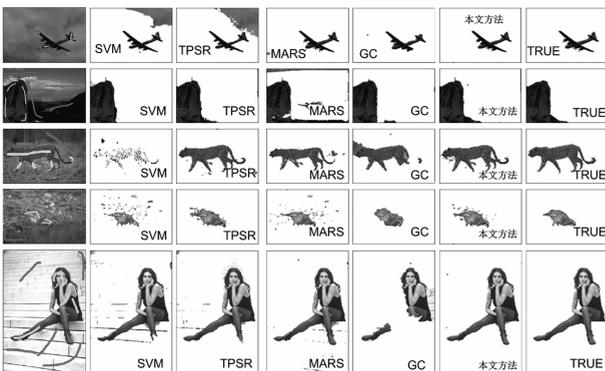


图7 不同分割算法在BSDS500数据库上的分割结果对比

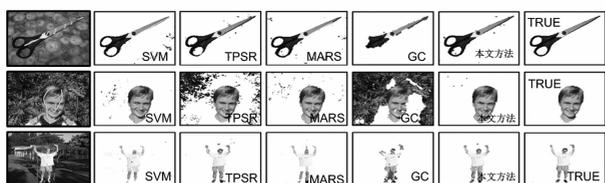


图8 不同分割算法在Pascal VOC数据库上的分割结果对比

- [J]. 电子学报,2003,31(1):29-32.
- Luo X-P, Tian J. A modified interactive segmentation of medical image series[J]. Acta Electronica Sinica,2003,31(1):29-32. (in Chinese)
- [2] Wang X-Y, Wang T, Bu J. Color image segmentation using pixel wise support vector machine classification[J]. Pattern Recognition,2011,44(4):777-787.
- [3] Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2004,26(2):147-159.
- [4] Grady L. Random walks for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006,28(11):1768-1783.
- [5] Xiang S, Nie F, Zhang C, et al. Interactive natural image segmentation via spline regression[J]. IEEE Transactions on Image Processing,2009,18(7):1623-1632.
- [6] Menon R, Bhat G, Saade G R, et al. Multivariate adaptive regression splines analysis to predict biomarkers of spontaneous preterm birth[J]. Acta Obstetrica et Gynecologica Scandinavica,2014,93(4):382-391.
- [7] Fu Z, Wang L, Zhang D. An improved multi-label classification ensemble learning algorithm[A]. Pattern Recognition[C]. Berlin:Springer,2014:243-252.
- [8] Galar M, Fernandez A, Barrenechea E, et al. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews,2012,42(4):463-484.
- [9] Zhou Z-H, Li M. Semisupervised regression with cotraining-style algorithms[J]. IEEE Transactions on Knowledge and Data Engineering,2007,19(11):1479-1493.
- [10] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[A]. Proceedings of Eighth IEEE International Conference on Computer Vision[C]. IEEE,2001. 416-423.
- [11] Everingham M, Van Gool L, Williams C K, et al. The pascal visual object classes (voc) challenge[J]. International Journal of Computer Vision,2010,88(2):303-338.
- [12] Jekabsons G. ARESLab: adaptive regression splines toolbox for Matlab[J]. Institute of Applied Computer Systems Riga Technical University, Meza,2010,1(3):1-19
- [13] Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(2): 147-159.
- [14] Chang C-C, Lin C-J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology (TIST),2011,2(3):1-27.

### 作者简介



刘金平 男,1983 年生于湖南邵阳. 湖南师范大学数学与计算机科学学院讲师. 研究方向为计算机视觉与模式识别.  
E-mail: ljp202518@163.com



陈青 女,1967 年生于湖南长沙,中南大学信息科学与工程学院博士研究生,研究方向为计算机信息技术、智能化信息检测与处理等领域的研究工作.