

稀疏表示下的非监督显著对象提取

李仲生^{1,2}, 李仁发², 蔡则苏³, 赵乘麟¹

(1. 邵阳学院信息工程系, 湖南邵阳 422000;

2. 湖南大学网络与信息安全湖南省重点实验室, 湖南长沙 410082;

3. 哈尔滨工业大学计算机学院, 黑龙江哈尔滨 150001)

摘要: 针对现有显著对象提取算法时间复杂度高和未考虑显著对象的完整性等问题, 提出了一种能适应资源有限环境的显著对象提取算法. 首先建立了稀疏表示的数学模型, 归纳出了显著对象与稀疏表示的对应关系、区域间的边能近似模式和邻接区域间的渐变模式. 然后依据对应关系确定候选区域, 依据渐变模式和边能近似模式实现显著对象的局部提取. 对比实验证实: 本文算法高速、精确地捕捉到了显著对象, 并能在一定条件下保持显著对象的完整性.

关键词: 时间复杂度; 对象局部提取; 稀疏表示; 渐变模式; 边能近似模式

中图分类号: TP311 **文献标识码:** A **文章编号:** 0372-2112 (C110076) 06-1097-06

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2012.06.005

Unsupervised Salient Object Extraction Based on Sparse Representation

LI Zhong-sheng^{1,2}, LI Ren-fa², CAI Ze-su³, ZHAO Cheng-lin¹

(1. Department of Information Engineering, Shaoyang University, Shaoyang, Hunan 422000, China;

2. Hunan Provincial Key Laboratory of Network and Information Security, Hunan University, Changsha, Hunan 410082, China;

3. Computer Science and Technology School of Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: The existing salient object extraction algorithms had high time complexity and didn't take the integrity of the objects into account, and a salient object extraction algorithm is proposed, which is adaptive to the resource-constrained environment. The mathematic model of sparse representation is built. The corresponding relation between the salient objects and sparse representation, edge energy approximation pattern among regions, and gradual change pattern between adjacent regions are induced. The candidate regions are determined based on the relation, and the salient objects are locally extracted based on gradual change pattern and edge energy approximation pattern. The contrast experiments indicate that the salient objects are captured accurately and the integrity of the salient objects are well-kept under given conditions with the proposed algorithm.

Key words: time complexity; object local-abstraction; sparse representation; gradual change pattern; approximation pattern of edge energy

1 引言

随着大量的嵌入式设备和传感器纳入信息系统, 每个服务器的客户端设备数量可达到几万个之多, 对这些嵌入式设备和传感器发送的海量信息进行存储、搜索、校对、汇兑和分析, 将是 21 世纪信息领域新的挑战任务^[1]. 如何提取显著对象, 高效处理海量信息中的图像、视频内容, 成为信息处理领域的焦点之一.

近期, 非监督提取显著对象的研究活跃^[2~10]. 这些文献有以下共性: 需要先对图像作分割或作边缘分析,

时间复杂度高; 在显著对象被部分遮挡时未考虑其完整性等. 为此, 本文提出了一种稀疏表示下的非监督显著对象提取算法, 有以下特点: 处理基本单元可任意形状, 任意分布(即可稀疏); 引入边能近似模式, 以保显著对象的完整性; 引入渐变模式, 实现显著对象的局部提取.

2 稀疏表示的数学描述及显著对象提取的基本原理

2.1 数学描述

稀疏表示是对图像中潜存对象内在的聚集性的一

种描述,以下给出对它作定量分析的一系列数学定义.

定义 1 拓扑信息系统:定义为 $I = (U, A, P, V, f, V_p, f_p)$, 其中 U 称为论域, A, P 都是有限属性集, $V = \bigcup_{a \in A} V_a$, V_a 表示属性 a 的值域; $V_p = \bigcup_{p \in P} V_p$, V_p 是拓扑属性 p 的值域; $f: U \times A \rightarrow V$ 为信息映射函数, 完成论域、属性到值域的唯一映射; $f: U \times P \rightarrow V_p$ 是论域对象的拓扑位置映射函数.

定义 2 不可分辨关系及等价类:不可分辨关系是一种由任意属性子集 $B \subseteq A$ 决定的二值关系, 定义如下.

$$T(B) = \{(x, y) \mid f_a(x) = f_a(y) \text{ or } f_a(y) = * \text{ or } f_a(x) = *, \forall a \in B \text{ and } x, y \in U\} \quad (1)$$

依据一不可分辨关系划分论域成类, 同类对象不可分辨, 这些类即等价类, 其合集记为 $U/T(B)$.

定义 3 邻接关系及拓扑等价类:对任意拓扑属性子集 $Q \subseteq P$, 相应的邻接关系定义为:

$$T(Q) = \{(x, y) \mid D(f_p(x), f_p(y)) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \leq r, \forall p \in Q \text{ and } x, y \in U\} \quad (2)$$

n 为维数, r 为阈值. 例如, 在二维图像中且要求八邻域时, $r = \sqrt{2}$. 对等价类依据邻接关系作进一步划分, 把符合邻接关系的对象归为一类, 形成拓扑等价类. 合集记为 $U/T(BQ)$. $T(BQ)$ 是 $T(B)$ 和 $T(Q)$ 的交.

定义 4 维表:将拓扑等价类按维排列形成的一种数据结构. 其本质是将拓扑等价类作分段投影. 例如在二维图中, 基本维表有横维表(记为 \mathcal{R}_r)和纵维表(记为 \mathcal{R}_c).

定义 5 维表纯容量:维表中包含的论域对象总和.

定义 6 维扫:依据第 τ 维排列成的维表, m 切分了拓扑等价类, 得 m 个分量, 记为 $\tau = (\tau_1, \tau_2, \dots, \tau_m)$, 沿 $\tau_i (i = 1, \dots, m)$ 依次扫描, 记录非该拓扑等价类的对象或统计相关参数, 这个过程称维扫. 维扫所记录的对象合集记为 \mathcal{C} .

定义 7 维扫梯度:对于二个相邻维表, 在各维作维扫, 计算所有维上相邻点的梯度, 其方差的均值称维扫梯度. 设维表为 D 维, 第 i 维有邻点 d_i 对, C_{ij}, N_{ij} 为一对邻点的概念向量, 则维扫梯度

$$G_r = \sum_{i=1}^D \sum_{j=1}^{d_i} |C_{ij} - N_{ij}|^2 / \sum_{i=1}^D d_i \quad (3)$$

定义 8 相对边能:二个拓扑等价类边能差异的度量, 计算如下.

$$E_e = \left(\frac{1}{d_i} \sum_{i=1}^{d_i} |L_{ni}|^2 - \frac{1}{d_j} \sum_{j=1}^{d_j} |M_{kj}|^2 \right) / \sigma^2 \quad (4)$$

其中 L_{ni} 是第 n 个拓扑等价类第 i 个点的概念向量, M_{kj} 为第 k 个拓扑等价类第 j 个边点的概念向量, d_n, d_k 为对应边点总数, σ 为相应概念下拓扑划分标准差.

定义 9 维表视在容量:维表自身对象个数与集合 \mathcal{C} 中的对象个数的总和.

定义 10 拓扑分布密度:维表纯容量在维表视在容量中占的权重.

定义 11 拓扑连通强度:在拓扑等价类中, 非边点与维表纯容量的比值称拓扑连通强度.

定义 12 显著度:对于大于一定尺度的拓扑等价类, 记其拓扑分布密度为 $d_{\mathfrak{M}}$, 拓扑连通强度 $h_{\mathfrak{M}}$, 则其显著度 $S_{\mathfrak{M}} = \alpha \times d_{\mathfrak{M}} + \beta \times h_{\mathfrak{M}}$ (α, β 为权系数).

以上为在同一概念模式下的拓扑等价类分析, 在不同概念模式下, 拓扑空间中的拓扑等价类之间有下近似、上近似关系, 并可类推出粗糙区, 相关形式化定义可参见文[12], 与文[12]不同的是, 本文的上、下近似关系等在拓扑位置空间内确定.

2.2 基本原理

根据大量实验, 发现在稀疏表示下有以下规律:显著对象必有显著的拓扑等价类与其对应; 存在渐变模式; 同类或同一对象的不同部分间存在边能近似模式.

根据定义 10、11、12 可计算出拓扑等价类的拓扑显著度.

模式 1 渐变模式:如果某对相邻的拓扑等价类满足某种渐变约束, 则它们之间构成渐变模式.

以定义 4、6、7 等定量分析渐变模式. 实验显示, 渐变约束条件要以具体概念模式、具体位置定, 例如, 在采用 HSI 颜色空间的 I 下, 对应于 H 确定的显著对象候选区的下近似集内, 两个拓扑等价类的渐变约束为等价类连续或维扫梯度小于 $0.3\sigma^2$ (在做了预合并后所用的参数, 后同); 但在粗糙区, 则必须满足维扫梯度小于 $0.3\sigma^2$. σ 为标准差, 在实现数据稀疏表示过程中自适应确定.

图 1 是一个实例(注:本文实验用的全是彩图, 出版时转为灰度图). 各子图的等价类号[拓扑等价类号]依次为: 5[2262], 5[2283], 5[2288], 6[2933], 7[3443], 7[3444], 8[3847], 4[128]. 全位于图的上半部, 用白色标注, 其中的 4[128]属于概念模式 1, 前面七个之间符合渐变模式, 属于概念模式 2. 二者交叉验证, 能准确定位天空边缘.

模式 2 边能近似模式:对于等价类分布近似相同的显著对象候选区, 或等价类连续、拓扑等价类相邻的候选区, 若其相对边能(按定义 8 计算)低于 0.1, 则这些候选区构成边能近似模式. 图 2 是一个实例.

边能近似模式用于聚集分开的同类对象或同一对象因为遮挡而分离的多个部分. 此外, 为判定对象的居



图2 边能近似模式示例(依次为7[474],7[477],原图)

中程度,引入定义 13.

定义 13 位权:记为 P_{w_i} ,是一个对象在图像中所处位置的居中性度量.

位权可以以四面到四边距离的标准差均值来度量.设任意候选对象 i , d_o 是对象四面离四边的距离,其中 o 为边的编号,依次从左 1 起,顺时针编. l_o 表示相应边垂直方向的维高. ρ_o 表示归一距离: $\rho_o = \frac{d_o}{l_o}$,均值为 $\bar{\omega}$, $\bar{\omega} = \frac{1}{n} \sum_{i=1}^n \rho_o$, n 为边数.位权为:

$$P_{w_i} = 1 - \frac{1}{n} \sum_{i=1}^n |\rho_o - \bar{\omega}| \quad (5)$$

当显著对象包含多个区域时,以其中位权最大的区域确定位权.

3 显著对象提取的实现

以前期双概念拓扑标注为基础,按序以下面步骤完成:概念模式 1 下尺度过滤;依据显著度筛选;候选区合并;双概念交叉验证,局部语义分割;候选对象的删除、合并;位权分析;输出显著对象.

第一步,过滤尺度太小、不可能显著的拓扑等价类.第二步,计算选中的拓扑等价类的显著度.

$$\Gamma_{d_i} = S_{\mathcal{N}_i} = \alpha \times d_{\mathcal{N}_i} + \beta \times h_{\mathcal{N}_i} \quad (6)$$

其中的 $i = 1, \dots, m$, m 是尺度过滤后的拓扑等价类总数.按显著度筛选,找出候选区域.

算法 1 候选区域初选算法



图1 渐变模式示例

输入: Γ_{d_i} 序列

输出: 候选区域序列.

(1) 如果 Γ_{d_i} 序列长度为 1, 对应的拓扑等价类直接保留为候选区域. 转 4.

(2) 如果一个候选区域与图的上下对边的归一距离 ρ_o (见定义 13) 均近于 0, 视为背景.

(3) 将 Γ_{d_i} 以升序排序, 采用文 [13] 改进的 Fisher 线性判别算法作计算, 确定序列中的最佳切分位. 如果切分位两边的拓扑显著度差值超过 0.1, 则只输出切分位后的拓扑等价类为显著对象候选区域, 否则保留全部拓扑等价类.

(4) 输出, 算法结束.

第三步, 合并候选区.

算法 2 相近候选区启发式合并算法

输入: 当前候选区序列

输出: 经合并后的候选区域序列.

(1) 计算相对边能, 判定是否有符合边能近似模式的候选区存在, 若无, 则转第 4 步, 若有, 舍去超过 1 的, 执行以下各步.

(2) 调用改进的 Fisher 线性判别算法找到切分点, 分成二类, 合并低值对应的各组候选区.

(3) 在合并二个并不相邻的候选区时, 以中心连接线为启发, 洪泛式地合并等价类相等或连续的小区域. 若不可合并区已隔断二个区域, 则停止. 如隔断区大于二个候选区的尺度均值, 则二候选区不并. 否则合并并做标记.

(4) 全部处理完后, 退出.

算法 3 候选区域合并

输入: 候选区域序列

输出: 经合并后的候选区域序列.

(1) 在概念模式 1 下, 一旦最低类包围了最高类, 合并.

(2) 如果一个区包含另一个相邻区达百分之六十以上, 并二个区等价类连续且外围等价类号低、内部的高, 合并二区.

(3) 调用算法 2, 处理上面二步未处理的可合并情况.

(4) 在各区内去孔, 输出, 完毕.

算法 3 的第 1、2 步是经验规则, 在实验的数百幅图和视频监控的原型系统应用中, 未发现反例, 能使处理更精简.

接下来, 根据所确定的候选拓扑等价类, 在概念模式 2 下的拓扑标注上依据渐变模式作局部语义分割, 提取出候选显著对象, 具体见算法 4.

算法 4 局部粗糙对象提取

输入: 概念模式 1、2 下的拓扑标注, 概念模式 1 下

的候选拓扑等价类维表 R_{1i} .

输出: 候选对象.

(1) 如果 R_{1i} 当前待处理, 在概念 2 的拓扑标注的对应范围内求 R_{1i} 的上近似类集、下近似类集.

(2) 如果在 R_{1i} 的上下近似类集只有一个拓扑等价类, 即存在一个拓扑等价类 R_{2i} , 成立 $R_{1i} \subseteq R_{2i}$, 在概念模式 2 下计算 R_{1i} 与邻接区域的维扫梯度, 如果小于阈值 $0.3\sigma^2$, 则 R_{1i} 被忽略, 查找 R_{2i} 逆向上近似中存在的大拓扑等价类, 对它们递归调用本算法, 随后按渐变模式合并, 去孔, 输出; 否则, R_{1i} 对应区域去孔后单独作为一候选对象输出. 转 5.

(3) 否则, 处理下近似类集. 合并 R_{1i} 的下近似类中等价类号连续且相邻、或维扫梯度低于 $0.3\sigma^2$ 的拓扑等价类.

(4) 处理粗糙区内的拓扑等价类. 如果粗糙区内的某一拓扑等价类向内 (即 R_{1i} 区内) 与内部其他拓扑等价类连续渐变、向外与 R_{1i} 区外的相邻拓扑等价类构成渐变模式, 则切开, 分别并入各邻接拓扑等价类; 如果仅与内部连续渐变, 就归于内部; 如果只与外部连续渐变, 则归于外部.

(5) 也可继续引入其他已积累的知识模式以形成更多模态, 完成更逼近的语义分割.

(6) 对新形成区域作邻接分析, 如果维扫梯度低于阈值 $0.3\sigma^2$, 合并之. 去孔, 并记录各候选显著对象在概念模式 2 下的拓扑分布. 本次处理完毕, 取下一个 R_{1i} , 返回 1; 如果已处理完, 算法结束.

在算法 4 完成局部粗糙分割后, 为进一步强化所提取的对象的语义, 调用算法 5 合并满足相应约束的候选对象.

算法 5 候选对象合并

输入: 候选显著对象集

输出: 处理后的候选显著对象集.

(1) 如果一个候选对象完全包围一个或几个与外部邻接对象不存在权重近于一致的概念 2 下拓扑等价类分布的候选对象, 合并.

(2) 合并有着权重近于一致的概念 2 下拓扑等价类分布的候选显著对象.

(3) 处理完毕后, 算法返回.

接下来, 按式 (7) 更新候选显著对象的拓扑显著度.

$$\Gamma_{di} = \sum_{i=1}^t \frac{a_i}{a_{\mathcal{R}}} \Gamma_{di} \quad (7)$$

式中, $a_{\mathcal{R}}$ 是合并后的某个候选显著对象总面积, t 为相应对象所覆盖的尺度过滤后、未作 Fisher 切分前的拓扑等价类总数, i 代表第 i 个拓扑等价类, a_i 为第 i 个扑等

价类被覆盖的面积, Γ_{di} 为对应拓扑显著度.

对 Γ_{di} 序列调用算法 1, 得到最终候选对象. 如果只剩一个候选对象, 直接输出. 否则, 合并相邻且 Γ_{di} 相差不超过 0.086 (微调参数) 的候选对象, 接着作位权分析, 取位权最大的作显著对象.

$$O_{\text{label}} = \arg \max_{i \in L_c} P_{w_i} \quad (8)$$

其中 L_c 是候选显著对象的标号集.

4 时间复杂性分析、对比

加上前期工作, 本文算法有三步: 单概念分析、拓扑标注和显著对象提取. 设论域对象总数为 N , 则单概念分析器的时间复杂度为: $\leq C_1 \times N \times \log_2^{\text{Max}}$, 其中 C_1 表示一次减法运算的耗时, Max 代表最终聚类数, 小于等于指在第二轮迭代以后有类别变动的数据已极少, 远小于 N . 标注算法需要扫描的次数为 $N + 2 \times M$, 其中 M 为行连通段数, 一般远小于 N . 由于大的拓扑连通区数不多, 显著对象局部提取耗时极少, 因此, 本算法的时间复杂度近似为 $C_1 \times (N \times \log_2^{\text{Max}} + N + 2M)$.

将它与文 [8] 等在提取显著对象前用来作前期分割的 JSEG 算法作比较. 仅量化和计算 J 值, JSEG 算法的时间复杂度将不低于 $C_2 \times [N \times \text{Max} + (2 \times \text{offset} + 1)^2 \times N]$, 其中 C_2 为乘法用时, Max 为量化类数, offset 为计算 J 值时的偏移量, 高于本文算法.

5 实验结果与分析

实验在一台笔记本电脑上完成 (1.60 GHz, 1GB 内存), 软件配置为 GhostXP, SP2, VC60, OpenCV 2, Matlab 7. 采用的图像有非标准的自拍图像和标准的 Berkley 图像库等, 共三百多幅.

计算显著度的参数如下: $\alpha = \beta = 0.5$, 其他参数由实验系统运行时自动产生. 概念模式选取的是 HSI 颜色空间中之二概念 H 和 I.



图 3 pho1.jpg 的显著对象提取

图 3 是对自然照片 pho1.jpg (750 * 499) 提取显著对象的示例. 图 3(a) 为原图, 图 3(b) 为算法的中间结果, 本文称之为主控候选对象, 图 3(c) 是所提取的显著对象, 显然, 捕捉到了聚焦中心, 直观上有语义贴近性 (即与客观显著对象的一致性).

图 4 中的第一、二行与文 [2] 作比较, 比较标准采用文 [2] 定义的“命中率”, 命中率越高效果越好. 第二行中,

文[2]的命中率为 0.575,本文为 0.921232.第一行中,由于文[2]提供的“真实”图像只是鱼的一部分,不适合作量化对比,本文的方法,找到了被隔断的尾巴,保持了显著对象的完整性.第三行是与文[7]的比较结果,比较标准采用文[7]定义的“误分率”,误分率越低效果越好.其中文[7]的误分率为 0.380,若由客观对象(熊)直接计算,本文分割的误分率 0.071425,低于文[7].第四行是与文[8]的比较结果,直观可见本文结果更准确地捕捉住了原显著对象(例如其中的箭头).所有实验结果显示:本文算法所得结果边缘的直观语义贴近性较优,并在一定情况

下保持了显著对象的完整性.

上文所比较的三文在处理图像时各有侧重,文[2]采用多尺度分析,文[7]注重焦点区域,文 9 依赖 canny 边缘和 JSEG 分割,本文算法与它们比较都取得了较优的效果,在一定程度上显示出了本文算法的通用性.此外,由于本文以任意区域为处理基元并以局部方式提取显著对象,且极少涉及浮点运算,在资源相对有限的笔记本上速度都达到了毫秒级,远低于其他基于图像分割的算法.例如在同样的环境下,文[8]所用的 JSEG 分割仅计算 J 值就接近分级.

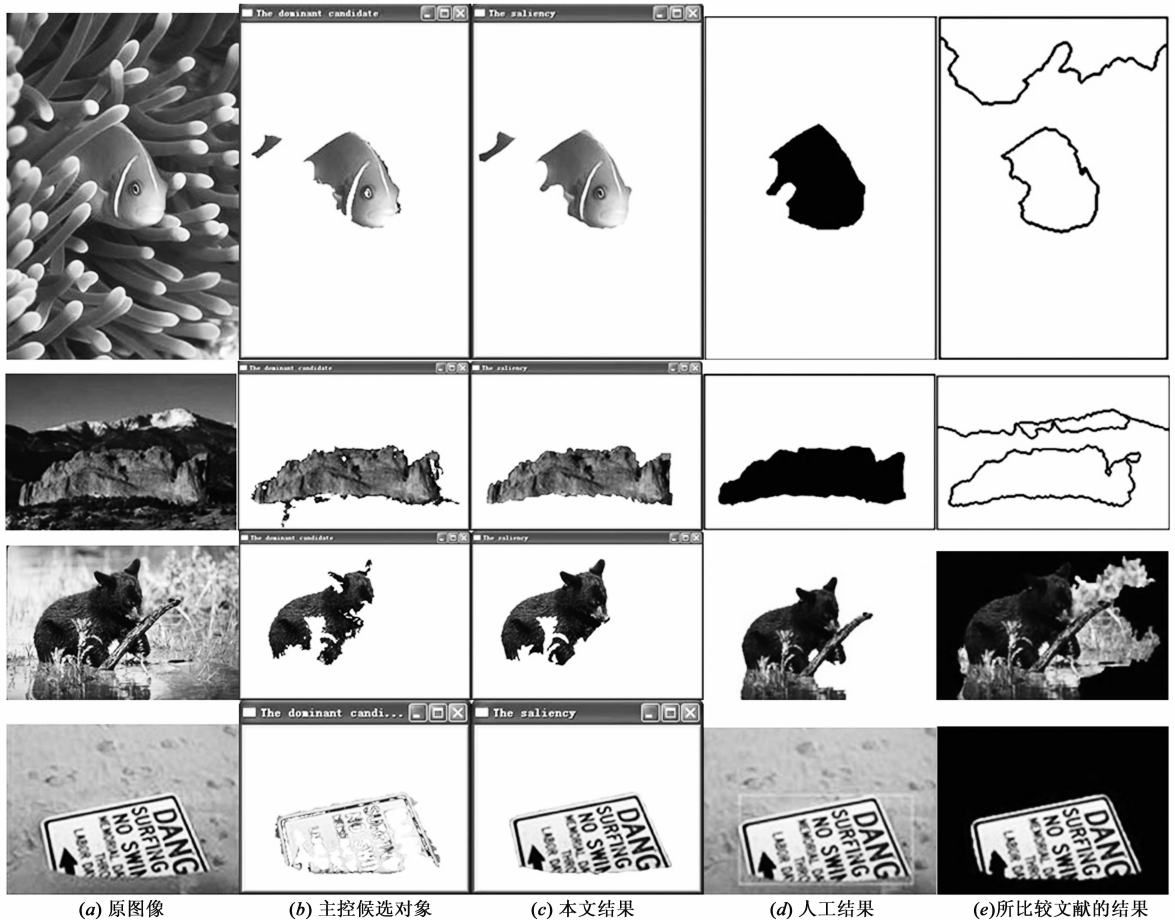


图4 显著对象提取及对比分析

6 结语

本文给出了图像稀疏表示的数学模型,并在此基础上实现了显著对象的非监督局部提取,所得结果与客观对象有着较好的直观语义逼近性,并在一定条件下保持了显著对象的完整性.未来的工作,将继续深化本文工作,寻求通用的物类识别方法.

参考文献

[1] 李国杰. 信息科学技术的长期发展趋势和我国的战略取

向[J]. 中国科学信息科学, 2010, 40(1): 128 - 138.

- [2] Liu Zhi, Shen Liquan, Zhang Zhaoyang. Unsupervised image segmentation based on analysis of binary partition tree for salient object extraction[J]. Signal Processing, 2011, 91(2): 290 - 299.
- [3] Nock R, Nielsen F. Statistical region merging[J]. IEEE Trans Pattern Anal Machine Intel, 2004, 26(11): 1452 - 1458.
- [4] Yang X, Zhao W, Chen Y, et al. Image segmentation with a fuzzy clustering algorithm based on ant-tree[J]. Signal Process, 2008, 88(10): 2453 - 2462.

- [5] Vanhamel I, Pratikakis I, Sahli H. Multi-scale gradient watersheds of color images[J]. IEEE Trans Image Process, 2003, 12(6): 617 – 626.
- [6] 苗启广, 等. 遥感图像中无水桥梁识别新算法[J]. 电子学报, 2011, 39(7): 1698 – 1701.
MIAO Qi-guang, et al. Novel algorithm for recognition of bridge in remote sensing image[J]. Acta Electronica Sinica, 2011, 39(7): 1698 – 1701. (in Chinese).
- [7] Liu Zhi, Li Weiwei, Shen Liqian, et al. Automatic segmentation of focused objects from images with low depth of field[J]. Pattern Recognition Letters, 2010, 31(7): 572 – 581.
- [8] Feng Songhe, Xu De, Yang Xu. Attention-driven salient edge (s) and region(s) extraction with application to CBIR[J]. Signal Processing, 2010, 90(1): 1 – 15.
- [9] WANG W, et al. Measuring visual saliency by site entropy rate [A]. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2010 [C]. San Francisco, USA: IEEE Computer Society Press, 2010. 2368 – 2375.
- [10] Goferman S, Zelnik-Manor L, Tal A. Context-Aware saliency detection [A]. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2010 [C]. San Francisco, USA: IEEE Computer Society Press, 2010. 2376 – 2383.
- [11] Deng Y, Manjunath B S. Unsupervised segmentation of color-texture regions in images and video[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(8): 800 – 810.

- [12] 张腾飞, 等. 粗糙集理论中属性相对约简算法[J]. 电子学报, 2005, 33(11): 2080 – 2083.
ZHANG Teng-fei, et al. Algorithms of attribute relative reduction in rough set theory[J]. Acta Electronica Sinica, 2005, 33(11): 2080 – 2083. (in Chinese).
- [13] Deng Y, Kenney C, Moore M S, et al. Peer group filtering and perceptual color image quantization [A]. Proceedings of the 1999 IEEE International Symposium on Circuits and Systems [C]. Orlando, FL, USA: IEEE Computer Society Press, 1999. 4: 21 – 24.

作者简介



李仲生 男, 1967 年 8 月生于湖南邵阳市, 副教授, 博士, 主要研究方向为多媒体信息处理、粒计算.

E-mail: zsl1666 @ 163 . com



李仁发 男, 1956 年生, 中国计算机学会高级会员, 博士, 教授, 博士生导师, 主要研究方向为嵌入式计算、数字媒体与网络、无线网络.

E-mail: lirenfa@vip.sina.com