

Journal of Image  
and Graphics

# 中国图象图形学报



ISSN1006-8961  
CN11-3758/TB

2012 **10**  
Vol.17 No.

中国科学院遥感应用研究所  
中国图象图形学学会主办  
北京应用物理与计算数学研究所

# 中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年10月 第17卷 第10期(总第198期)

## 目次

### 综述

张量值图像插值方法综述 ..... 邵宇, 刘莹, 孙富春(1197)

### 图像处理和编码

利用模运算及其周期性特点的安全隐写算法 ..... 廖琪男(1206)

结合快速融合与颜色传递的双波段夜视图像染色 ..... 李郁峰, 冯晓云, 徐铭蔚, 黄文丽(1213)

Weber 定律下尺度空间的自适应构建 ..... 刘立, 张瑞军, 万亚平, 黄欣阳, 彭复员(1222)

改进权值函数的非局部均值去噪算法 ..... 单建华(1227)

### 图像分析和识别

非相相似度保持投影 ..... 陈才扣, 侯钰(1232)

非正交二值子空间模板表示的并行生成 ..... 杨培, 武港山, 杨扬, 任桐炜(1237)

Adaboost 和随机图划分的无监督图像分类 ..... 李巍, 杨素锦, 段晓华(1245)

aiNet 背景抑制的单帧红外弱小目标检测 ..... 陈炳文, 王文伟, 秦前清(1252)

韦伯-中心环绕结构的图像显著性检测模型 ..... 林丽莉, 周文晖(1261)

气象传真图信息提取 ..... 李存东, 肖传毅, 潘海朗, 陈日清, 杨劲松(1268)

### 图像理解和计算机视觉

多要素空间场景相似性匹配模型及应用 ..... 宋腾义, 汪闽(1274)

### 计算机图形学

Delaunay 三角网通用合并算子及分治算法的简化 ..... 刘永和, 冯锦明, 郭维栋, 田根, 金毅(1283)

基于 Laplace 谱嵌入和 Mean Shift 的三角网格一致性分割 ..... 马亚奇, 李忠科, 赵静(1292)  
断裂面匹配的破碎刚体复原 ..... 李群辉, 周明全, 耿国华(1298)

### 虚拟现实与增强现实

核爆炸外观景象实时模拟 ..... 郑涛, 徐晓刚, 邵承永(1305)

### 医学图像处理

全局脑白质纤维群智能跟踪算法 ..... 冯远静, 王哲进, 张贵军, 俞立(1312)

### 遥感图像处理

对偶四元数线阵遥感影像几何定位 ..... 盛庆红, 姬亭, 刘微微, 王惠南(1319)

修正安装矩阵提高 FY-3B/MERSI 的地理定位精度 ..... 吴荣华, 杨忠东, 关敏, 李翔翔(1327)

## 中国图象图形学报

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 10 期

2012 年 10 月 16 日出版

主管单位 中国科学院

主 办 中国科学院遥感应用研究所  
中国图象图形学学会  
北京应用物理与计算数学研究所

主 编 李小文

编辑出版 《中国图象图形学报》编辑出版委员会

北京 9718 信箱 邮编 100101  
电子信箱:jig@irsa.ac.cn  
电话:010-64807995 010-82614429  
网 址:www.cjig.cn

印刷装订 北京北林印刷厂

广告经营许可证 京朝工商广字第 0346 号

总 发 行 北京报刊发行局

订 购 全国各地邮局

国外发行 中国国际图书贸易总公司  
(中国国际书店)  
(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing Application,  
CAS China Society of Image and Graphics  
Institute of Applied Physics and Computational  
Mathematics

Chief editor LI Xiaowen

Editor, Publisher Editorial and Publishing Board  
of Journal of Image and Graphics  
(P. O. Box 9718, Beijing 100101, China)  
E-mail:jig@irsa.ac.cn

Distributed by Beijing Bureau for Distribution of Newspapers  
and Journals

Domestic All Local Post Offices in China

Foreign China International Book Trading Corporation  
(P. O. Box 399, Beijing 100044, China)

Printed by Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

# Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 10 October 2012

## Contents

### Review

Overview of tensor valued images interpolation technology ..... Shao Yu, Liu Ying, Sun Fuchun (1197)

### Image Processing and Coding

Secure steganography based on modulo and its cyclical characteristic ..... Liao Qinan (1206)

Night vision dual-band images coloration using fast fusion and color transfer  
..... Li Yufeng, Feng Xiaoyun, Xu Mingwei, Huang Wenli (1213)

Adaptive algorithm of scale-space construction method based on Weber's law  
..... Liu Li, Zhang Ruijun, Wan Yaping, Huang Xinyang, Peng Fuyuan (1222)

Non-local means denoising algorithm with enhanced weight function ..... Shan Jianhua (1227)

### Image Analysis and Recognition

Dissimilarity preserving projection ..... Chen Caikou, Hou Yu (1232)

Parallel algorithm for generating template representation based on non-orthogonal binary subspace  
..... Yang Pei, Wu Gangshan, Yang Yang, Ren Tongwei (1237)

Unsupervised image categorization based on Adaboost and stochastic graph partition ..... Li Wei, Yang Sujin, Duan Xiaohua (1245)

Infrared dim target detection in single image based on background suppression by aiNet  
..... Chen Bingwen, Wang Wenwei, Qin Qianqing (1252)

Image saliency detection model with Weber's law based on center-surround structure ..... Lin Lili, Zhou Wenhui (1261)

Information extraction from meteorological facsimile maps  
..... Li Cundong, Xiao Chuanyi, Pan Hailang, Chen Riqing, Yang Jinsong (1268)

### Image Understanding and Computer Vision

Multi-feature based spatial scene matching model and its application ..... Song Tengyi, Wang Min (1274)

### Computer Graphics

Merging planar Delaunay triangulations based on universal operators and the implementation of a divide-conquer algorithm  
..... Liu Yonghe, Feng Jinming, Guo Weidong, Tian Gen, Jin Yi (1283)

Consistence segmentation of triangle mesh using Laplace spectral embedding and Mean Shift  
..... Ma Yaqi, Li Zhongke, Zhao Jing (1292)

Reassembly of broken 3D solids based on fractured surfaces matching ..... Li Qunhui, Zhou Mingquan, Geng Guohua (1298)

### Virtual Reality and Augmented Reality

Real-time simulation of nuclear explosion scene ..... Zheng Tao, Xu Xiaogang, Shao Chengyong (1305)

### Medical Image Processing

Global white matter tractography using swarm optimization ..... Feng Yuanjing, Wang Zhejin, Zhang Guijun, Yu Li (1312)

### Remote Sensing Image Processing

Geo-positioning line-array CCD images with dual quaternion ..... Sheng Qinghong, Ji Ting, Liu Weiwei, Wang Huinan (1319)

Improved FY-3B/MERSI geolocation accuracy using installation matrix  
..... Wu Ronghua, Yang Zhongdong, Guan Min, Li Xiangxiang (1327)

中图法分类号: TN911 文献标识码: A 文章编号: 1006-8961(2012)10-1245-07

论文引用格式: 李巍, 杨素锦, 段晓华. Adaboost 和随机图划分的无监督图像分类[J]. 中国图象图形学报, 2012, 17(10): 1245-1251.

## Adaboost 和随机图划分的无监督图像分类

李巍<sup>1</sup>, 杨素锦<sup>1</sup>, 段晓华<sup>2</sup>

1. 周口师范学院计算机科学与技术学院, 周口 466001; 2. 中山大学信息科学与技术学院, 广州 510006

**摘要:** 针对当前大多数无监督图像分类方法不能对每个图像类进行特征选择和自动确定图像类别的数量问题, 提出一种基于 Adaboost 和随机图划分的无监督图像分类方法。该方法包括两个部分: 1) 将图像分类问题看做是一个自动的随机图划分问题, 其中图的每一个顶点代表一幅图像, 通过划分形成的子图代表了图像类。再采用 Adaboost 算法对每一个形成的图像类进行特征选择, 从而得到每类图像的表达模型。2) 采用一种基于蒙特卡洛马尔可夫链(MCMC)的随机采样算法(SWC)来对图进行划分。相比传统的随机采样算法, SWC 具有更快的收敛速度。在两个图像数据集上的实验结果表明, 本文方法的分类性能明显优于其他现有的无监督分类法。

**关键词:** 无监督图像分类; Adaboost; 图划分; Swendsen-Wang cuts

## Unsupervised image categorization based on Adaboost and stochastic graph partition

Li Wei<sup>1</sup>, Yang Sujin<sup>1</sup>, Duan Xiaohua<sup>2</sup>

1. School of Computer Science and Technology, Zhoukou Normal University, Zhoukou 466001, China;  
2. School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, China

**Abstract:** In this paper, we present a general framework to discover image categories automatically. The algorithm includes two parts: 1) we pose the problem of category discovery as an automated graph partition task. Each graph vertex indicates an image, and a partitioned sub-graph consisting of connected graph vertices representing a clustered category. The model of each image category can be learned by stepwise feature selection using the Adaboost algorithm. 2) A MCMC-based stochastic algorithm, the Swendsen-Wang Cuts (SWC), is adopted to solve the graph partition fast. Compared to traditional random cluster sampling techniques, SWC converges faster. We apply our method on two image datasets, and the experimental results demonstrate superior performance of our method over other popular state-of-the-arts methods, including Kmeans, pLSA, and Affinity Propagation.

**Key words:** unsupervised image categorization; Adaboost; graph partition; Swendsen-Wang cuts (SWC)

## 0 引言

目前, 对图像或物体无监督分类的研究已经取得了一定的成果, 并在图像检索、目标检测和识别等应用中扮演着非常重要的角色<sup>[1-4]</sup>。但是当前的无监督图像分类方法仍然存在如下两个问题:

1) 对不同的图像类别自动选择特征。不同类别的图像具有不同的特征构成, 例如夕阳场景的图像具有大量的红色, 城市街道场景的图像则具有大量的结构和线条特征。因此为不同类别的图像选择不同的特征, 将对图像分类起着重要作用。在有监督的图像分类方法中, 可以通过最小化某个损耗函数(loss function)来进行特征的选择。而大多数现

收稿日期: 2011-12-05; 修回日期: 2012-04-13

第一作者简介: 李巍(1982—), 男, 讲师, 2009年于电子科技大学获软件工程专业硕士学位, 主要从事计算机应用及开发、图形图像处理技术等方面的研究。E-mail: zhkliwei@126.com

有的无监督图像分类方法都没有考虑特征选择的问题,而对所有类别使用统一的预定义的特征向量来表示<sup>[5-7]</sup>。

2) 自动地选择类别数量。以往的无监督聚类方法<sup>[4-5]</sup>,通常需要事先指定类别的数量。但是在无任何先验知识的情况下,很难判断待分类图像数据到底有多少类别的。而类别数量在无监督图像分类中起着至关重要的作用,也是影响分类效果的重要因素。所以无监督分类方法应当具有自动确定类别数量的能力。因此一个无监督图像分类方法应当具备在聚类的过程中能够产生新类别、删除已有类别以及让数据在不同类别间的跳转的能力。一般来说,有两类方法可以完成以上任务:一种是基于层次 Dirichlet 过程的方法<sup>[8-9]</sup>,但是这类方法不支持在分类的同时进行特征选择;另一种是基于马尔可夫链蒙特卡罗 (MCMC) 的方法<sup>[10-11]</sup>,这类方法非常适合在分类的同时进行特征选择,但是还没有很好地应用于无监督图像分类。

基于上述两点,提出一种基于 Adaboost<sup>[12]</sup>和随机图划分的无监督图像分类方法。将图像的分类问题看做是贝叶斯理论下后验概率最大化的图划分问题。图中的每个顶点代表一幅图像,每条边赋予一个连接概率。根据边的概率随机地断开图中的边,形成一个图划分,每个子图可以看做一个图像类别。因为边的连接概率是根据图像之间的相似性计算所得,相似图像之间的边断开的可能性会很小,而不相似图像之间的边则更容易断开,亦即相似的图像容易被分到同一类,而不相似的图像则很难被分到同一类别。并且每次对图进行划分的同时,对每个形成的图像类别利用 Adaboost 算法进行自动选择特征,从而学习出每个类别的表达模型。将图的一个划分以及在当前划分上每个图像类别的概率表达模型称为一个解。那么无监督图像分类的目标则是寻求一个在贝叶斯理论下的最优解。基本过程概括如下:

1) 初始化,给定一个初始邻接图,并根据图像间的相似性计算每条边的连接概率。同时给定一个初始解状态  $S_0$ 。然后重复后面的两个步骤。

2) 图划分,对当前的解状态  $S$ ,断开所有不同类别图像之间的边。然后对剩下的边根据对应的边概率随机地断开或者连接。这样将得到一系列连通分量。

3) 随机选择一个连通分量,将该连通分量随机

地合并到当前划分状态中的一个图像类别当中,或者独立产生一个新的图像类。与此同时利用 Adaboost 算法对每一个类别进行特征选择,形成每一个类别的概率表达模型,形成一个新的解状态  $S'$ ,并以概率  $\alpha(S \rightarrow S')$  接受新的解状态。如果新的解状态被接受,那么令  $S = S'$ ,否则当前解状态不变。重复步骤 2) 和步骤 3),直到超过一定的迭代次数或者达到一个满意的后验概率。

本文方法主要有以下两个方面的贡献:首先,提出了一种新的通用的无监督图像分类框架;其次,在进行图划分的同时采用 Adaboost 算法为每个图像类别进行特征选择,从而学习出该类图像的表达模型;最后,将无监督图像聚类看做随机图划分问题,无需事先指定分类的类别数量,而由方法自动形成类别数量。

## 1 基于贝叶斯理论的无监督图像分类

给定一个没有任何标注信息的图像数据集  $D = \{I_1, I_2, \dots, I_N\}$  以及特征集合  $F = \{f_1, f_2, \dots, f_m\}$ 。则每幅图像可表示为  $R(I) = (f_1(I), f_2(I), \dots, f_m(I))$ 。无监督图像分类的目标则是  $D$  划分成未知的  $K$  个不相交的类别,即

$$\pi = \{D_1, D_2, \dots, D_K\} \quad (1)$$

式中,  $\cup_{i=1}^K D_i = D, D_i \cap D_j = \emptyset, \forall i \neq j$ 。在本文中,将分类问题看做是一个邻接图  $G_0 = \langle V, E_0 \rangle$  的划分问题。图中的顶点集合  $V = D$ ,即每个顶点代表一幅图像,  $E_0$  为边的集合。并且为每条边赋予一个连接概率,根据边的概率随机地断开或者连接邻接图的每条边,从而形成一系列子图。如此形成一个图的划分,即一个分类状态。同时为每个图像类  $D_k$  学习其概率表达模型

$$\psi = \{P_k(F_k, \Theta_k), F_k \subset F, k = 1, \dots, K\} \quad (2)$$

式中,  $P_k$  为第  $k$  类图像的概率表达模型,由对该类所选择的特征集合  $F_k$  和对应的参数  $\Theta_k$  决定。

因此无监督图像分类的解可以表示为

$$S = (K, \pi, \psi) \quad (3)$$

式中,  $K$  表示类别数量,  $\pi$  和  $\psi$  为类别划分和对应的表达模型。那么无监督图像分类的目标则是通过最大化贝叶斯后验概率来寻找最优的解

$$S^* = \arg \max_{S \in \Omega} p(S | D) \quad (4)$$

式中,  $\Omega$  为整个解空间。基于贝叶斯理论,有

$$p(S | D) \propto p(S)p(D | S) \quad (5)$$

式中,  $p(S)$  为先验概率,  $p(\mathbf{D} | S)$  为似然概率。先验概率  $p(S)$  由  $K$ 、 $\boldsymbol{\pi}$  和  $\boldsymbol{\psi}$  决定。为了便于计算, 假定  $\boldsymbol{\pi}$  和  $\boldsymbol{\psi}$  服从均匀分布, 而  $K$  的分布通常假定为指数模型, 即  $p(K) = \exp(-\beta K)$ ,  $\beta$  为一经验常数用于限制图像类别的数量。假设每个图像类  $\mathbf{D}_k$  包含了  $n_k$  幅图像, 则

$$p(\mathbf{D} | S) = p(\mathbf{D} | \boldsymbol{\pi}, \boldsymbol{\psi}) = \prod_{k=1}^K P_k(\mathbf{F}_k, \boldsymbol{\Theta}_k) = \prod_{k=1}^K \prod_{i=1}^{n_k} P_k(I_{k,i}, \mathbf{F}_k, \boldsymbol{\Theta}_k) \quad (6)$$

因此, 式(5)的后验概率可以写为

$$p(S | \mathbf{D}) \propto \exp(-\beta K) \prod_{k=1}^K \prod_{i=1}^{n_k} P_k(I_{k,i}, \mathbf{F}_k, \boldsymbol{\Theta}_k) \quad (7)$$

式中, 表达模型  $P_k$  是利用 Adaboost 算法进行特征选择和参数学习而得到的。下面将首先介绍特征选择和模型学习的算法。

## 2 特征选择和模型学习

对每一个图像类  $\mathbf{D}_k = \{I_1, I_2, \dots, I_{n_k}\}$ , 特征选择的目的是选择一个具有判别力的特征子集  $\mathbf{F}_k \subset \mathbf{F}$ , 同时给  $\mathbf{D}_k$  建立一个概率模型  $P_k$ 。通常很难直接得到  $\mathbf{D}_k$  的表达模型。因此将图像类的表达模型问题看做是一个基于 Adaboost 算法的分类问题。将  $\mathbf{D}_k$  当做一个假设类别(假设其中的每一幅图像属于同一类), 并且随机选择其中的一半图像作为训练集的正样本, 从自然图像中随机选择一个包含  $n_{\text{ref}}$  幅图像的图像数据集  $\mathbf{D}_{\text{ref}}$  作为负样本来进行特征选择和模型学习。则假设训练样本为  $\{(I_1, y_1), (I_2, y_2), \dots, (I_m, y_m)\}$ , 共包含  $m = n_k/2 + n_{\text{ref}}$  幅图像, 其中  $y_i = \{+1 | -1\}$ ,  $+1$  表示图像  $I_i$  为正样本,  $-1$  表示  $I_i$  为负样本。则特征选择和分类器学习过程如下:

1) 初始化样本分布  $d_1(I_i) = 1/m$ 。

2) For  $t = 1, \dots, T$

(1) 在当前的样本分布  $d_t$  下, 选择一个错误率  $\varepsilon_t$  最小的弱分类器  $h_t$  (即一个最佳的特征  $f_t$  和对应的阈值);

(2) 给  $h_t$  赋予一个权值  $\alpha_t = \frac{1}{2} \ln\left(\frac{1 - \varepsilon_t}{\varepsilon_t}\right)$ ;

(3) 更新样本分布

$$d_{t+1}(I_i) = \frac{d_t(I_i) \exp(-\alpha_t y_i h_t(I_i))}{Z_t}$$

式中,  $Z_t$  为归一化常量。

3) 输出最终分类器

$$H(I) = \text{sgn}\left(\sum_{t=1}^T \alpha_t h_t(I)\right) \quad (8)$$

一旦通过 Adaboost 算法学习得到了一个如式(8)所示的强分类器, 根据  $H(I)$  的符号可以预测测试样本属于哪一类。

将  $\mathbf{D}_k$  的剩下一半图像当作测试样本, 假设为  $\{I_1, \dots, I_{n_k/2}\}$ , 则每个样本的裕量 (margin) 为

$$M_i = \sum_{t=1}^T \alpha_t h_t(I_i) \quad (9)$$

而每个测试样本的似然概率定义为<sup>[12]</sup>

$$p(I_i) = \frac{1}{Z} \exp\{y_i M_i\} = \frac{1}{Z} \exp\left\{y_i \sum_{t=1}^T \alpha_t h_t(I_i)\right\} \quad (10)$$

式中,  $Z = \sum_{i=1}^T \alpha_t$  为归一化常量。

无监督图像分类的目标是寻求后验概率(式(11))的最大化, 而当每个聚类内部具有较强的一致性时, 则具有较大的似然概率, 从而使得分类结果有较大的后验概率; 反之, 当每个聚类内部不一致时, 则会有较小的似然概率, 从而使得分类结果具有较小的后验概率。因为本文方法为一种随机聚类采样方法, 因此不可避免在采样过程中形成的图像聚类内部很不一致, 例如某假设的图像类(或图划分中的子图)实际包括了建筑、森林、人脸等不同类型的图像, 以其一半为正样本, 事实上是选择了不好(不属于一类)的样本, 基于这种样本学习出来的模型势必影响分类的结果。但是这种情况往往会得到较小的似然概率, 从而导致当前划分状态的后验概率较小而被算法拒绝。即使在迭代的过程中被接受, 也会最终会被新的更好的划分替代。

## 3 基于随机图划分的聚类采样

根据式(7)和式(10), 可将式(4)改写为

$$S^* \propto \arg \min_{S \in \Omega} \exp(-\beta K) \times$$

$$\prod_{k=1}^K \prod_{i=1}^{n_k/2} \frac{1}{Z^{(k)}} \exp\left\{y_i^{(k)} \sum_{t=1}^T \alpha_t^{(k)} h_t^{(k)}(I_i^{(k)})\right\} \quad (11)$$

为了得到最优的解  $S^*$ , 引入基于随机聚类采样

的 SWC 算法来对图进行分割。该算法由 Swendsen 和 Wang 在 1980 年提出,一开始是用于仿真 Ising/Potts 图模型。后来由 Bardu 等人<sup>[11]</sup>扩展到图划分领域,称为 Swendsen-Wang cuts (SWC) 算法。这里将其进一步应用于图像分类领域,在介绍 SWC 算法之前先介绍邻接图的初始化。

### 3.1 邻接图和边概率

图像的无监督分类可看做是一个图的分割问题,要采用 SWC 算法进行图划分,需要构造一个邻接图和给出每条边的连接概率(edge probability)。

给定一个初始邻接图  $G_0 = \langle V, E_0 \rangle$ , 其中的每一个顶点表示一个图像,并且所有顶点的集合即为待分类的图像数据集合。对每个图像表示为  $R(I) = (f_1(I), f_2(I), \dots, f_M(I))$ , 而给每条边  $e_{s,t} \in E_0 (s, t \in [1, n], s < t)$  赋予一个伯努利随机变量  $\mu_{e_{s,t}} \in \{\text{on}, \text{off}\}$  用于指示边  $e_{s,t}$  连接或者断开。而边连接的概率  $q_{e_{s,t}} = p(\mu_{e_{s,t}} = \text{on} | I_s, I_t)$ , 反映了这条边所连接的两个图像之间的相似性,其中边概率定义为

$$q_{e_{s,t}} = \exp\left\{-\frac{\tau[K(R(I_s) \| R(I_t)) + K(R(I_t) \| R(I_s))]}{2}\right\} \quad (12)$$

式中,  $K(\cdot)$  为 Kullback-Leibler 散度,  $\tau$  为一个常数。通常如果两幅图像分别属于两个不同类别(即不相似)  $q_{e_{s,t}}$  将接近 0, 即边  $e_{s,t}$  连接的概率非常小; 反之两个图像属于同一类别,  $q_{e_{s,t}}$  将接近 1。

为了使邻接图  $G_0 = \langle V, E_0 \rangle$  尽量稀疏,不使用完全图,限制每个顶点的度数不超过一定的阈值  $d$  (例如  $d = 5$ ), 这样可以更快地搜索到最优的划分状态。根据边的连接概率独立地断开每一条边,可以得到一个新的邻接图  $G = \langle V, E \rangle$ 。该邻接图由若干个连通分量组成,即得到了邻接图  $G_0 = \langle V, E_0 \rangle$  的一个划分。但这样的划分依赖于全部的特征集合  $F$ , 而没有对每个不同的图像类别自动选择不同的特征。因此接下来将基于 AdaBoost 的模型学习方法与图划分进行结合,在每次对图进行划分的同时对每个形成的图像类自动选择特征构造不同的概率表达模型。

### 3.2 SWC 算法

引入 SWC 算法来实现无监督图像分类。其好处是无需事先指定类别的数量,而由算法根据最大

后验概率自动确定类别的数量。如图 1(b) 所示,该算法主要包括两个步骤:

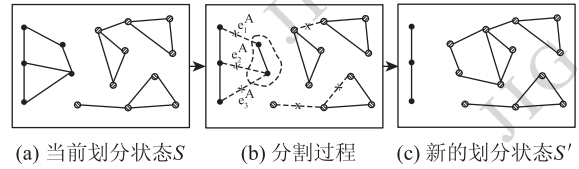


图 1 SWC 图划分算法示意图

Fig. 1 Illustration of graph partition by SWC algorithm

1) 对当前邻接图中的每一条边根据其边概率随机地断开,产生一系列连通分量;

2) 随机地选择一个连通分量,将该连通分量随机地合并到当前的任意图像类中,或者独立出来形成一个新的图像类别。

详细的 SWC 算法如下:

输入:  $G_0 = \langle V, E_0 \rangle$ ,  $q_e, \forall e \in E_0$ , 式中  $V = D$ 。

输出: 无监督分类结果  $S^* = (K, \{D_k, P_k\}_{k=1}^K)$ 。

初始化: 给定一个初始的划分状态  $\Pi_0 = (D_1, D_2, \dots, D_{K_0})$ , 对每个类别利用第 3 节所述方法学习其表达模型,从而得到初始的解状态  $S_0 = (K_0, \{D_k, P_k\}_{k=1}^{K_0})$ 。重复以下步骤:

1) 对于当前解状态  $S = (K, \{D_k, P_k\}, k = 1, \dots, K)$ , 以概率  $1 - q_e$  断开当前邻接图的每一条边,每一个图像类  $D_k = (D_{k,1}, D_{k,2}, \dots, D_{k,n_k})$  被分割成  $n_k$  个连通分量,  $k = 1, \dots, K$ 。  $CP = \{D_{ki} : k = 1, \dots, K, i = 1, \dots, n_k\}$  表示形成的所有连通分量集合。

2) 随机选择一个连通分量  $D_0 \in CP$ , 假设  $D_0 \subset D_k$ 。将  $D_0$  的所有图像随机地标定为第  $k'$  类,  $k' \in [1, K+1]$  并且以式(13)所示概率接受这个标定。如果该标定被接受,使用第 3 节所述的模型学习算法学习每一类的表达模型,得到一个新的解状态  $S'$ , 并且有

(1) 如果  $k' = k$ , 或者  $k' = K+1$  并且  $D_0 = D_k$ , 那么  $S' = S$ , 即解状态未发生变化;

(2) 如果  $k' = K+1$  并且  $D_0 \neq D_k$ , 那么产生一个新类别  $D_{k'} = D_0$ , 然后使用第 3 节所述算法选择其特征  $F_{k'}$ , 形成表达模型  $P_{k'}$ ;

(3) 如果  $k' \in [1, K]$  并且  $k' \neq k$ , 那么将进行类别重组, 即  $D_k = D_k - D_0$  (如果  $D_0 = D_k$ , 那么第  $k$  类将消亡), 而  $D_{k'} = D_{k'} \cup D_0$ , 同样对新的类别选择特征  $F_k$  和  $F_{k'}$ , 分别得到表达模型  $P_k$  和  $P_{k'}$ 。

接受新的解状态的概率为

$$\alpha(S \rightarrow S') = \min\left(1, \frac{Q(S' \rightarrow S) p(S' | \mathbf{D})}{Q(S \rightarrow S') p(S | \mathbf{D})}\right) \quad (13)$$

式中,  $Q(S' \rightarrow S)$  和  $Q(S \rightarrow S')$  为状态跳转概率

$$\frac{Q(S' \rightarrow S)}{Q(S \rightarrow S')} = \frac{\prod_{e \in C(\mathbf{D}_0, \mathbf{D}_{k'} - \mathbf{D}_0)} (1 - q_e)}{\prod_{e \in C(\mathbf{D}_0, \mathbf{D}_k - \mathbf{D}_0)} (1 - q_e)} \cdot \frac{q(k | \mathbf{D}_0)}{q(k' | \mathbf{D}_0)} \quad (14)$$

式中,  $C(\mathbf{D}_i, \mathbf{D}_j)$  为连通分量  $\mathbf{D}_i$  和  $\mathbf{D}_j$  之间的边的集合。

## 4 实验结果及评价

### 4.1 评价方法

对于给定的数据集  $\mathbf{D}$ , 包含  $L$  个类别, 共  $N$  幅图像, 并且对应的真实类别标签 (ground truth) 为

$$\mathbf{C} = \{c_1, c_2, \dots, c_N\} \quad (15)$$

式中,  $c_i \in \{1, \dots, L\}$ 。

对于算法输出的分类结果  $S = (K, \{\mathbf{D}_k, P_k\}_{k=1}^K)$ , 分类标签为

$$\hat{\mathbf{C}} = \{\hat{c}_1, \hat{c}_2, \dots, \hat{c}_N\} \quad (16)$$

式中,  $\hat{c}_i \in \{1, \dots, K\}$ , 而  $K$  可能大于  $L$  也可能小于  $L$ 。

那么对于每一个算法输出类别  $\mathbf{D}_k \subset \mathbf{D}$  (其分类标签为  $k$ ), 其中的每一个图像有其对应的真实类别标签

$$\mathbf{C}_k = \{c_1^{(k)}, \dots, c_{n_k}^{(k)}\} \subset \mathbf{C} \quad (17)$$

为评价方法的性能, 采用文献 [4] 中所用的纯度 (Purity) (越大越好) 和条件熵 (conditional entropy) (越小越好) 来评价无监督图像分类结果的性能。其中纯度和条件熵分别定义为

$$Purity(\hat{\mathbf{C}} | \mathbf{C}) = \sum_{k \in \hat{\mathbf{C}}} p(k) \max_{l \in \mathbf{C}} p(l | k) \quad (18)$$

$$H(\hat{\mathbf{C}} | \mathbf{C}) = \sum_{k \in \hat{\mathbf{C}}} p(k) \sum_{l \in \mathbf{C}} p(l | k) \log \frac{1}{p(l | k)} \quad (19)$$

式中,  $p(k) = \frac{|\mathbf{D}_k|}{N}$ , 而  $p(l | k)$  可以从  $\mathbf{C}_k$  中估计得到。

### 4.2 实验结果

为了测试本文方法的性能, 在两个数据集上进行了测试。其中一个数据集是从互联网上搜集的 10 类图片, 分别是建筑、汽车、山、高速公路、森林、海滩、人脸、瀑布、杯子和飞机, 每类 300 幅图像, 共 3 000 幅。另一个数据集是 MIT 的图像分类数据集 (<http://people.csail.mit.edu/torralba/code/spatiale->

[nvelope/](http://people.csail.mit.edu/torralba/code/spatiale-)), 包含 8 类场景, 共 2 688 幅图像。

实验中, 使用 3 种不同类型的特征, 分别是 Gist 特征<sup>[7]</sup>、SIFT 特征<sup>[13]</sup> 和颜色特征。其中 Gist 特征 512 维; SIFT 特征采用 Bag-of-words 模型来表示图像, 通过聚类共产生 500 个词 (visual words), 这样图像的 SIFT 特征为 500 维; 颜色特征采用 256 维的 HSV 颜色直方图, 其中色调 (H) 128 维, 饱和度 (S) 和亮度 (V) 各 64 维。因此每幅图像可以表示为一个 1 268 维的特征向量。而本文方法完全可以不局限于这 3 种特征, 可以使用更多的特征。

在 Core Duo 3.0 GHz CPU、4 G 内存的 PC 上进行实验。实验中设置式 (7) 中的参数  $\beta = 6$ , 式 (12) 中的常数  $\tau = 0.25$ 。因为本文方法是一种随机算法, 每次的结果并不相同, 所以对每个数据集运行本文方法 10 次, 每次迭代 10 000 次, 然后采用平均值来与其他方法进行比较。对于本文数据集运行本文方法 10 次的分类平均类别数量为 10.5 类; 而对于 MIT 数据集平均为 9 类。

本文方法的时间效率主要取决于两个因素: 数据规模 (图像数量) 和采样迭代次数。在 MIT 的数据集上, 约 0.12 s 每步; 而在本文数据集上约为 0.14 s 每步。为了比较 SWC 算法在聚类采样中的优势, 与传统的随机采样方法 Gibbs 采样<sup>[14]</sup> 方法进行了对比。图 2 显示了 SWC 和 Gibbs 采样方法分别在两个数据集上的能量 (Energy =  $-\log p(S | \mathbf{D})$ ) 收敛曲线。能量越小表示后验概率越大。从图 2 中可以看出, SWC 算法的收敛速度明显快于 Gibbs 采样方法。

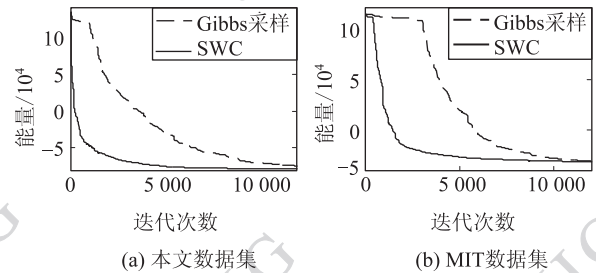


图 2 传统 Gibbs 采样方法与 SWC 算法的收敛速度比较。

Fig. 2 Convergence comparison between Gibbs sampler and SWC algorithm

为了对比分类结果的性能, 还与另外 3 个无监督分类方法 Kmeans、pLSA<sup>[15]</sup> 和 AP<sup>[16]</sup> 进行了比较。这些方法不包含特征选择步骤, 所以采用全部的 1 268 维特征来进行聚类, 聚类的数量预先指定为真实类别数量, 即对于本文数据集为 10, 而对于 MIT 数据集为 8。实验的对比结果如表 1 和表 2 所示。

表 1 纯度实验结果

Table 1 Experimental results on purity

|         | 方法     |      |      |      |
|---------|--------|------|------|------|
|         | Kmeans | pLSA | AP   | 本文   |
| 本文数据集   | 0.42   | 0.54 | 0.47 | 0.65 |
| MIT 数据集 | 0.45   | 0.59 | 0.53 | 0.68 |

表 2 条件熵实验结果

Table 2 Experimental results on conditional entropy

|         | 方法     |      |      |      |
|---------|--------|------|------|------|
|         | Kmeans | pLSA | AP   | 本文   |
| 本文数据集   | 2.12   | 1.56 | 1.77 | 1.24 |
| MIT 数据集 | 1.94   | 1.43 | 1.58 | 1.15 |

从表 1、表 2 可以看出本文方法在两个数据集上均取得了最高的纯度、最低的熵。总之,本文方法在分类性能上均比其他 3 个方法取得了更好的结果。

图 3 显示了一些图像类别的特征选择结果。这里只显示了判别性最好的前 40 个特征。纵轴表示了特征的权重,其越大说明判别性越强,横轴表示选择的某个特征,其颜色代表了该特征的类型。

从图 3 可以看出不同图像类别,具有不同的显著特征。例如对于建筑类的图像具有较多的 Gist 特征(因为 Gist 特征善于捕捉结构信息),而森林的场景则包含较多的颜色特征。因此实验说明对于不同类别的图像选择不同的特征来构造表达模型是非常有必要的。

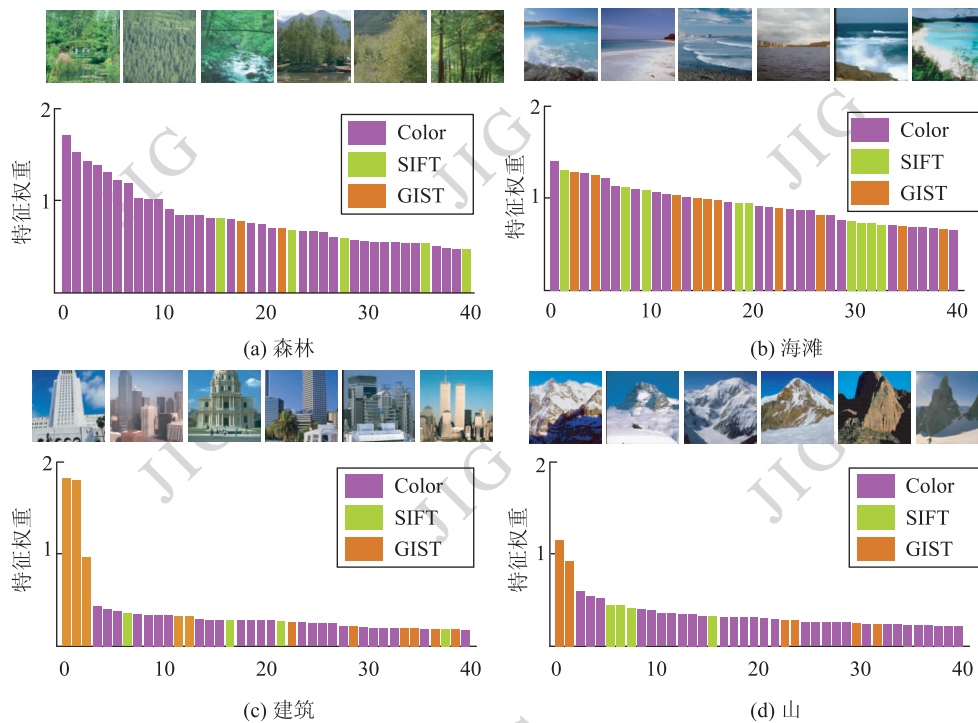


图 3 Adaboost 的特征选择的结果

Fig. 3 The selected features for 4 categories by Adaboost

## 5 结论

提出了一种新的无监督图像分类方法,将图像的分类问题看做是一个基于贝叶斯理论下的图划分问题。图的每个顶点代表待分类的图像集合中的一幅图像,每一条边表示了对应两幅图像之间的相似

性。在实现过程中,采用 SWC 算法对图不断进行随机分割、合并和重组,并且在分割的同时对当前划分状态中形成的每个图像类采用 Adaboost 方法进行特征的自动选择,从而学习出每一类的概率模型。本文方法是一种无监督分类方法,并且无需指定分类的数量。实验结果表明本文方法在图像分类中取得了令人满意的结果。

## 参考文献(References)

- [ 1 ] Fergus R, Perona P, Zisserman A. Object class recognition by unsupervised scale-invariant learning [ C ] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Madison, USA; IEEE Press, 2003: 264-271.
- [ 2 ] Simon I, Snavely N, Seitz S. Scene summarization for online image collections [ C ] // Proceedings of IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil; IEEE Press, 2007: 1-8.
- [ 3 ] Sivic J, Russell C, Efros A, et al. Discovering object categories in image collections [ C ] // Proceedings of IEEE International Conference on Computer Vision. Beijing, China; IEEE Press, 2005: 2-12.
- [ 4 ] Tuytelaars T, Lampert C, Blaschko M, et al. Unsupervised object discovery: A comparison [ J ]. International Journal of Computer Vision, 2009, 88(2): 284-302.
- [ 5 ] Liu D, Chen T. Unsupervised image categorization and object localization using topic models and correspondences between images [ C ] // Proceedings of IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil; IEEE Press, 2007: 1-7.
- [ 6 ] Waydo S, Koch C. Unsupervised learning of individuals and categories from images [ J ]. Neural Computing, 2008, 20(5): 1165-1178.
- [ 7 ] Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope [ J ]. International Journal of Computer Vision, 2001, 42(3): 145-175.
- [ 8 ] Sudderth B, Torralba B, Freeman T, et al. Describing visual scenes using transformed objects and parts [ C ] // Proceedings of IEEE International Conference on Computer Vision. Anchorage, USA; IEEE Press, 2008: 291-330.
- [ 9 ] Fox E, Sudderth B, Jordan I, et al. An HDP-HMM for systems with state persistence [ C ] // Proceedings of International Conference on Machine Learning. Helsinki, Finland; International Machine Learning Society, 2008: 312-319.
- [ 10 ] Tu Z W, Chen X, Yuille A, et al. Image parsing: Unifying segmentation, detection and recognition [ J ]. International Journal of Computer Vision, 2005, 63(2): 113-140.
- [ 11 ] Barbu A, Zhu S C. Generalizing Swendsen-Wang to sampling arbitrary posterior probabilities [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(8): 1239-1253.
- [ 12 ] Freund Y, Schapire R. A short introduction to boosting [ J ]. Journal of Japanese Society for Artificial Intelligence, 1999, 14(5): 771-780.
- [ 13 ] Lowe D. Distinctive image features from scale-invariant keypoints [ J ]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [ 14 ] Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, 6(6): 721-741.
- [ 15 ] Bosh A, Zisserman A, Muñoz X. Scene classification via pLSA [ C ] // Proceedings of European Conference on Computer Vision. Graz, Austria; University of Ljubljana, 2006: 517-530.
- [ 16 ] Frey B, Duech D. Clustering by passing messages between data points [ J ]. Science, 2007, 315(2): 972-976.