



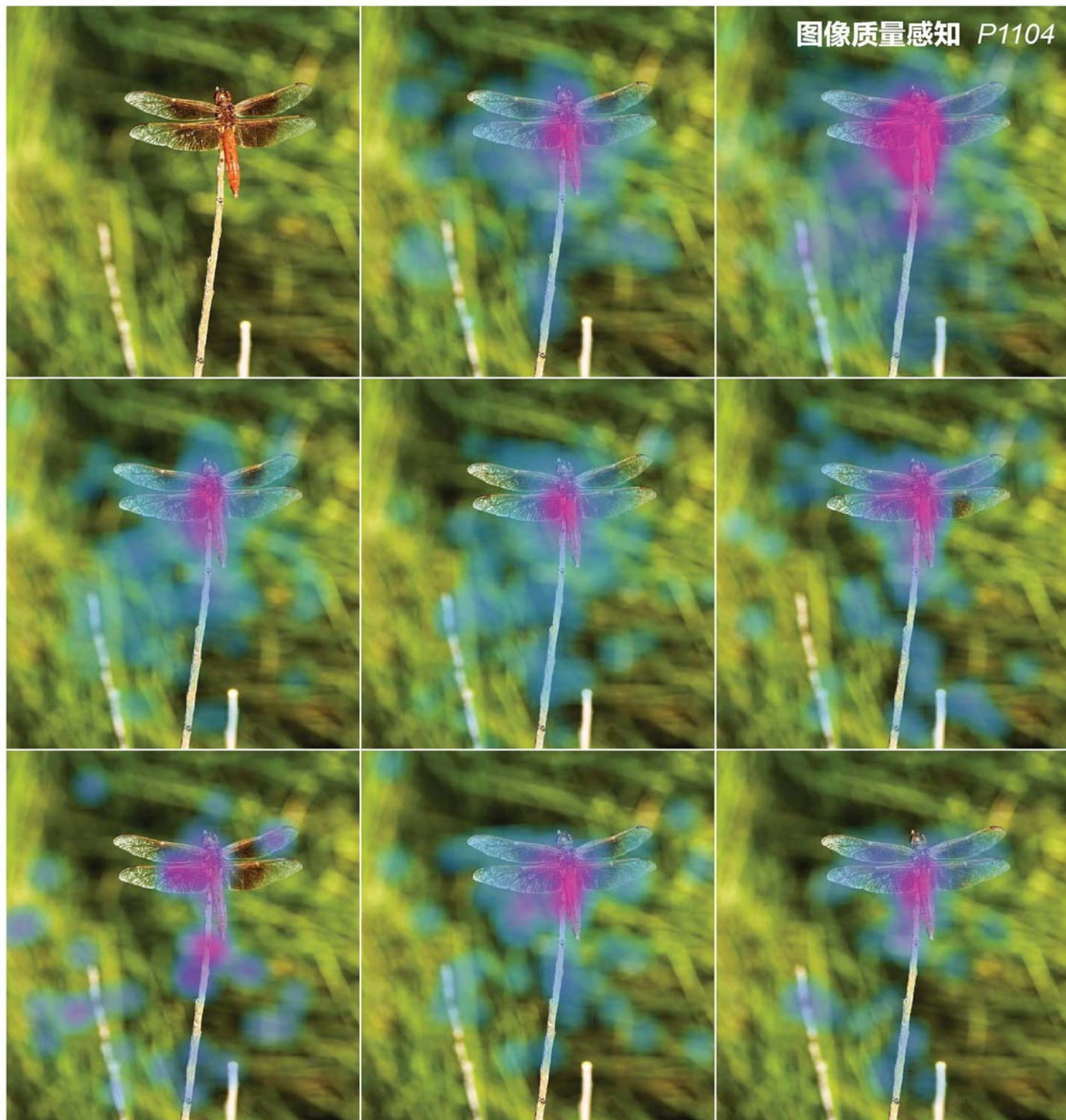
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

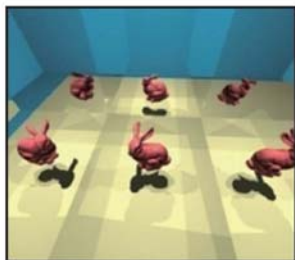
2014
07
VOL.19

ISSN1006-8961
CN11-3758/TB



多特征融合的图像自动变形
(第1012页)

多通道Haar-like特征多示例学习目标跟踪(第1038页)

快速虚顶点光子映射的焦散
绘制(第1062页)

综述

打印机光谱特性化方法研究综述

万晓霞, 刘强 0985

图像处理和编码

图像场方向导数的局部区域重建

叶学义, 齐珍珍, 何志伟, 李汪兵 0998

Lytro相机的光场图像校正与重对焦方法

周文晖, 林丽莉 1006

多特征融合的图像自动变形

肖红光, 肖树根, 刘京, 李淳芃, 陈立福 1012

结合频谱扩散的指纹图像滤波增强

卞维新, 封举富, 罗永龙, 徐德琴 1021

图像分析和识别

中心投影变换的抗噪性分析

黄永东, 李三三, 杨建伟 1031

多通道Haar-like特征多示例学习目标跟踪

宁纪锋, 赵耀博, 石武祯 1038

自主移动机器人自适应室外道路检测

杜明芳, 王军政, 李静, 崔广涛, 方建军, 曹海青 1046

分层树结构字典编码的行为识别

周同驰, 程旭, 吴镇扬 1054

计算机图形学

快速虚顶点光子映射的焦散绘制

李瑞瑞, 秦开怀 1062

双重视图下的3维网格模型可见水印

陆尘, 朱长青, 王玉海 1068

虚拟现实与增强现实

个性化编辑的轻量化3维树木模型构建

董天阳, 纪磊, 刘思远, 范菁, 熊丽荣 1074

遥感图像处理

带距离徙动校正的压缩感知PFA成像方法

段化军, 朱岱寅 1085

结合边缘和区域的活动轮廓模型SAR图像目标轮廓提取

王沛, 周鑫, 彭荣赜, 符鹏 1095

第2届全国图像图形联合学术会议专栏

视线跟踪技术的图像质量主观数据库构建及其可视化

张昀, 牟轩沁 1104

丰富图像标签的正则化非负矩阵分解方法

夏召强, 冯晓毅, 彭进业 1112

CONTENTS

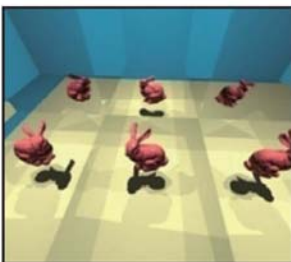
JOURNAL OF IMAGE AND GRAPHICS



Multi-feature fusion-based image morphing(P1012)



Multiple instance learning based object tracking with multi-channel Haar-like feature(P1038)



Fast caustics rendering by photon beam tracing based on virtual vertex transformations(P1062)

Review

- Review of spectral printer characterization
Wan Xiaoxia, Liu Qiang 0985

Image Processing and Coding

- Reconstructing image local regions based on directional derivative of a field
Ye Xueyi, Qi Zhenzhen, He Zhiwei, Li Wangbing 0998
- Light field image rectification and refocus method for Lytro camera
Zhou Wenhui, Lin Lili 1006
- Multi-feature fusion-based image morphing
Xiao Hongguang, Xiao Shugen, Liu Jing, Li Chunpeng, Chen Lifu 1012
- The fingerprint enhancement using filtering combined with spectrum diffusion
Bian Weixin, Feng Jufu, Luo Yonglong, Xu Deqin 1021

Image Analysis and Recognition

- Robustness to noise of central projection transformation
Huang Yongdong, Li Sansan, Yang Jianwei 1031
- Multiple instance learning based object tracking with multi-channel Haar-like feature
Ning Jifeng, Zhao Yaobo, Shi Wuzhen 1038
- Adaptive outdoor road detection for autonomous mobile robot
Du Mingfang, Wang Junzheng, Li Jing, Cui Guangtao, Fang Jianjun, Cao Haiqing 1046
- Action recognition using hierarchically tree-structured dictionary encoding
Zhou Tongchi, Cheng Xu, Wu Zhenyang 1054

Computer Graphics

- Fast caustics rendering by photon beam tracing based on virtual vertex transformations
Li Ruirui, Qin Kaihuai 1062
- Visible watermarking for three-dimensional mesh model in two views
Lu Chen, Zhu Changqing, Wang Yuhai 1068

Virtual Reality and Augmented Reality

- Construction of lightweight three-dimensional tree models supporting personalization editing
Dong Tianyang, Ji Lei, Liu Siyuan, Fan Jing, Xiong Lirong 1074

Remote Sensing Image Processing

- PFA imaging with range migration correction by compressive sensing
Duan Huajun, Zhu Daiyin 1085
- Active contour model based on edge and region attributes for target contour extraction in SAR image
Wang Pei, Zhou Xin, Peng Rongkun, Fu Peng 1095

Column of UCIG'2013

- Eye tracking based image quality assessment database and visualization method
Zhang Yun, Mou Xuanqin 1104
- Regularized non-negative matrix factorization method for image tag enrichment
Xia Zhaoqiang, Feng Xiaoyi, Peng Jinye 1112

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2014)07-1112-06

论文引用格式: Xia Z Q, Feng X Y, Peng J Y. Regularized non-negative matrix factorization method for image tag enrichment[J]. Journal of Image and Graphics, 2014, 19(7): 1112-1117. [夏召强, 冯晓毅, 彭进业. 丰富图像标签的正则化非负矩阵分解方法[J]. 中国图象图形学报, 2014, 19(7): 1112-1117.] [DOI:10.11834/jig.20140716]

丰富图像标签的正则化非负矩阵分解方法

夏召强, 冯晓毅, 彭进业

西北工业大学电子信息学院, 西安 710129

摘要: 目的 随着 Web2.0 技术的进步, 以用户生成内容为中心的社交网站蓬勃发展, 也使得基于图像标签的图像检索技术越来越重要。但是, 由于用户标注时的随意性和个性化, 导致用户提交的图像标签不够完备, 降低了图像检索的准确性。方法 针对这一问题, 提出一种正则化的非负矩阵分解方法来丰富图像欠完备的标签, 提高图像标签的完备性。利用非负矩阵分解的方法将原始的标签—图像矩阵投影到潜在的低秩空间里消除噪声, 同时利用图像的类内视觉离散度作为正则化项提高消除噪声、丰富标签的效果。结果 利用从社交网站 Flickr 上下载的大量社交图像进行对比实验, 验证了本文方法对丰富图像标签的有效性。通过对比目前流行的优化算法, 本文算法获得较高的性能提升, 算法平均准确度提高了 12.3%。结论 将图像类内视觉离散度作为正则化项的非负矩阵分解算法, 能较好地丰富社交图像的标签, 解决网络图像标签的欠完备问题。

关键词: 图像标签丰富; 正则化; 非负矩阵分解; 投影梯度法

Regularized non-negative matrix factorization method for image tag enrichment

Xia Zhaoqiang, Feng Xiaoyi, Peng Jinye

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China

Abstract: Objective With the development of Web 2.0, social websites centered on user-generated content are arising. Therefore, tag-based image retrieval becomes more and more important. However, the image tags that users upload are incomplete because users label images freely and arbitrarily and thus decrease the performance of image retrieval. **Method** To solve the problem of image tag incompleteness, this paper proposes an algorithm based on regularized non-negative matrix factorization to enrich the tags of social images and make these tags complete. This proposed algorithm casts the original tag-image matrix to a latent low-rank space and discovers the correlations between tags with the matrix factorization technique. The relationships among tags are utilized to enrich tags for social images. Meanwhile, the overall visual diversity as a regularization term is utilized to restrict the impact of content-irrelevant tags and enrich image tags. **Result** This paper constructs comparison experiments on images downloaded from sharing website Flickr. Accuracy is used to evaluate these comparison experiments. These experiments demonstrate the effectiveness of our proposed algorithm for enriching image tags. Compared with state-of-the-art approaches, our approach could improve average accuracy by 12.3%. **Conclusion** This paper proposes a regularized non-negative matrix factorization framework with overall visual diversity as the regularization term and enriches the tags of images effectively. Our proposed algorithm can solve the problem of incomplete tags.

Key words: image tag enrichment; regularization; non-negative matrix factorization; projected gradient method

收稿日期: 2013-11-29; 修回日期: 2014-02-24

基金项目: 国家自然科学基金项目 (61075014); 西北工业大学博士创新基金项目 (CX201113)

第一作者简介: 夏召强 (1986 年—), 男, 西北工业大学电路与系统专业在读博士研究生, 主要研究方向为多媒体内容分析与理解, 图像检索。E-mail: xiayan91@126.com

0 引言

随着社会化媒体的快速发展,用户上传和共享了数以亿计的数字图像,这给图像的管理和检索带来了巨大的挑战。如何在如此大规模的图像中,快速准确地查找所需要的图像,成为当前国内外研究的热点^[1]。

传统的基于内容的图像检索(CBIR)技术通常使用图像的视觉特征来匹配查询图像与数据库中的图像。传统的方法经常使用的视觉特征包括全局视觉特征^[2](如颜色直方图、纹理特征和形状特征)和局部视觉特征^[3](如SIFT关键点)。但是,基于视觉特征的图像查询方法受制于语义鸿沟的存在,无法突破底层视觉特征与高层语义之间的障碍。而为了克服语义鸿沟的问题,基于标签的图像检索方法(TBIR)使用用户为社会化图像(social image)提供的标签进行图像检索。由于用户给定的图像标签能够很好地描述图像的视觉内容,因此能够减少图像语义与视觉内容之间的差异性。并且用户只需要输入关键字即可搜索相关的图像,而不需要上传图像或者提供图像的网络地址。这种检索方式减少了用户的检索难度,使图像检索更加有效率。

但是,基于标签的图像检索方法要依赖于标签的数量和质量。通常,标签由图像的上传者提供,但其他用户在得到上传者的授权后也可以对上传图像进行标注。由于用户为了使用最少的代价管理和共享图像,往往使用相对较少^[4]和抽象的标签^[5]对图像进行标注。因此,图像的标签往往欠完备,大部分社会化图像只有较少的标签。图像标签的丰富(Tag Enrichment 或 Tag Completion),就是增加能够描述图像视觉内容的标签,使图像标签逐步完备的过程。因此,图像标签的丰富可以加强用户对社会化图像的管理,提高图像检索的效率。

为丰富图像的标签,可以使用传统的自动图像标注的方法^[6],如支持向量机、人工神经网络等分类器方法。但是,针对社会化图像,自动图像标注的方法存在一些难以克服的缺点:1)需要标注精确、图像数量较大的训练数据库;2)需要定义相对明确的概念,而不是用户自定义的标签。对于大量的网络图像,无法获取数量足够和标注完备的训练图像,也无法指定用户使用定义明确的词语,因此自动图像标注方法无法很好地应用到网络图像标签丰富这

一问题上。

于是,开始有学者使用其他的方法来解决这一问题。Jäschke 等人^[7]提出使用网络图像标签间的相关性来丰富网络图像的标签。该方法通过统计方法来计算标签间的共现关系,然后利用计算得出的标签共现关系来推断新的标签,借以丰富图像的标签。但是该方法只考虑标签间的共现关系,并没有考虑图像的视觉内容,因此容易增加与图像视觉内容无关的标签。Lee 等人^[8]提出了一种基于视觉分类法(visual folksonomy)的图像标签丰富方法,该方法利用最近邻方法为每幅图像构造一个视觉分类(即在特征空间内最相近的 N 幅图像),然后利用视觉分类中图像的相似度和标签共生关系改善标签的质量。但是如何确定一个合理的视觉分类,是该方法的关键点和难点。Shen 等人^[9]提出利用多示例学习算法建立图像标签与视觉内容上的关联,利用这种关联来进行图像标签的完备,但是该方法仍需要标注完备的训练数据库。Wu 等人^[10]提出了一种新的最优化方法来丰富图像的标签,通过最小化每幅图像的标签与图像视觉内容之间的差异来丰富标签,但很难保证标签间相似性与图像视觉内容相似性度量空间的一致性。

针对图像标签的丰富问题,提出了一种基于正则化的非负矩阵分解方法。该方法不仅利用了标签间语义上的相关性,而且还结合了图像视觉内容上的相似性。首先,该方法利用非负矩阵分解将原始的标签空间投影到潜在的低秩空间中,在潜在的低秩空间中利用标签间的相关性去除噪声,寻找部分丢失的图像标签关系。然后通过增加衡量图像类内视觉离散度的正则化项,降低与图像视觉内容不相关的标签的相关关系。通过两者的共同作用,提高标签丰富的效果。

1 问题描述

使用标签—图像矩阵(tag-image matrix)来表示标签与图像之间的相关关系。如果使用 T_0 表示用户标定的标签—图像矩阵,则 $T_0(i,j)$ 表示用户是否将第 i 个标签赋给第 j 幅图像。而图像标签的丰富问题,即转化为通过正则化的非负矩阵分解方法,估计出满足最优原则的标签—图像矩阵 T , $T(i,j)$ 表示第 i 个标签与第 j 幅图像的相关性大小,数值越大越相关。关于图像标签丰富的过程如图1所示,通

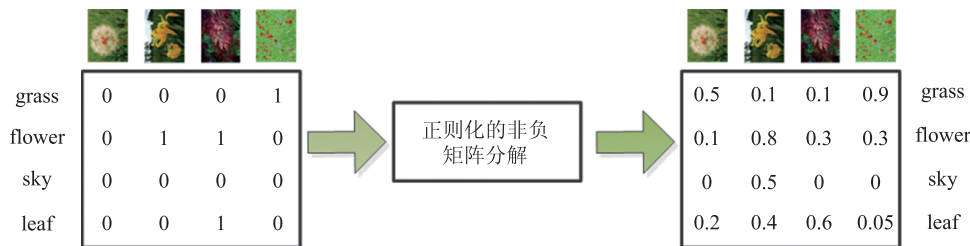


图1 图像标签丰富示意图

Fig. 1 Demonstration of image tag enrichment

过正则化非负矩阵分解估计得到的 $T(i,j)$ 为 $[0,1]$ 范围的概率值。

1.1 非负矩阵分解

非负矩阵分解^[11] (non-negative matrix factorization) 是一种带有约束条件的矩阵分解方法, 约束条件为分解的两个矩阵的各个元素必须为非负。非负矩阵分解已经应用到多个领域^[12], 如图像复原、文本分析和数据挖掘。其非负性的约束条件保证了分解的矩阵的元素可以解释, 有其特定的意义, 如在标签—图像矩阵中, 负的矩阵元素没有明确的物理含义, 标签与图像之间不能负相关。非负矩阵分解的基本思想是将一个较大的非负矩阵分解为两个相对较小的非负矩阵。

非负矩阵分解的问题可以描述为: 将标签—图像矩阵 T 分解为两个低秩的矩阵 M 和 H , H 表示潜在的在低秩空间, M 表示 T 投影到低秩空间 H 上的系数。非负矩阵分解的目标是使得 T 与 $M \cdot H$ 的差值最小。假设 n 和 m 为标签和图像的数目, 则 $T \in \mathbf{R}^{n \times m}$, $M \in \mathbf{R}^{n \times r}$, $H \in \mathbf{R}^{r \times m}$, 其中 r 为潜在低秩空间的维度。非负矩阵分解问题可表示为

$$\begin{aligned} \min_{M, H} & \|T - MH\|_F^2 \\ \text{s. t. } & M \geq 0, H \geq 0 \end{aligned} \quad (1)$$

式中, 矩阵 H 中各个列向量构成了矩阵分解的潜在在低秩空间的基, M 的各个行向量可以表示矩阵 T 的各个列向量在新的空间中投影所得的系数。约束条件保证了分解得到的两个子矩阵的所有元素都是非负的。

在通过式(1)求解最优解的过程中, 非负矩阵分解方法通过在潜在的在低秩空间里寻找标签间的相关性来消除噪声, 得到丰富的标签—图像矩阵。在求解得到最优解 M 和 H 后, 可以利用矩阵相乘运算 $M \cdot H$, 重新得到估计后的最优标签—图像矩阵 T 。但是由于标签之间的相关性, 部分图像也会被赋予与视觉内容不相关的标签。因此, 通过引入类

内视觉离散度来抑制与视觉内容不相关标签的丰富。

1.2 正则化因子

非负矩阵分解的过程, 即是在以 H 各列向量为基所构建的潜在低秩空间里对 T 进行消除噪声的过程。为保证在原有空间里标签与图像对应关系的准确性, 加入了一个衡量类内视觉离散度的正则化项, 然后利用该类内视觉离散度来约束潜在在低秩空间里噪声的消除, 调整标签间相关关系的影响。

如果将 $T(i,j)$ 看做是第 j 幅图像的权重系数, 则第 i 个标签的类内视觉特征的加权值可以表示为 $T(i,j) V_j$, 其中, V_j 表示第 j 幅图像对应的视觉特征向量(提取方法在实验部分有详细介绍)。根据一般类内离散度的定义, 图像类内视觉离散度可以表示为

$$\sum_{i=1}^n \left\| \sum_{j=1}^m T(i,j) V_j - U_i \right\|_2^2 \quad (2)$$

式中, n 和 m 为标签与图像的数量。 U_i 表示第 i 个标签初始的视觉特征均值向量, 即

$$U_i = \frac{\sum_{j=1}^m T_0(i,j) V_j}{\sum_{j=1}^m T_0(i,j)} \quad (3)$$

当图像类内视觉离散度作为正则化项时, 如果与图像视觉内容不相关的标签给予的权重太高, 则会引起此正则化的增大。因此, 通过正则化项可以约束 $T(i,j)$ 的变化。

将以上定义的正则化项和非负矩阵分解方法结合在正则化框架中, 新的目标函数可以使用矩阵的方式表示为

$$\begin{aligned} \min_{T, M, H} & \frac{1}{2} \|T - MH\|_F^2 + \frac{\lambda}{2} \|VT^T - U\|_F^2 \\ \text{s. t. } & M \geq 0, H \geq 0, 1 \geq T \geq 0 \end{aligned} \quad (4)$$

式中, λ 为正则化因子, V 为 V_j 级联而成的特征矩

阵, U 为 U_i 级联而成的均值矩阵。矩阵的差值度量方式使用的是 F-范数。图像标签的丰富问题即可转为求解式(4)定义的最优化问题。

1.3 问题求解

针对式(4)定义的约束最优化问题, 使用投影梯度法^[13]进行求解。式(4)中的目标函数使用 $L(T, M, H)$ 来表示, 即

$$L(T, M, H) = \frac{1}{2} \|T - MH\|_F^2 + \frac{\lambda}{2} \|VT^T - U\|_F^2$$

该目标函数是连续可微的。则针对优化问题中的变量 M 、 H 和 T , 其梯度分别表示为

$$\begin{aligned}\nabla_M L_t &= (MH_t - T_t)H_t \\ \nabla_H L_t &= (MH_t - T_t)M_t \\ \nabla_T L_t &= (MH_t - T_t) + \lambda(VT_t^T - U)V^T\end{aligned}\quad (5)$$

下标 t 为当前迭代的状态, 表示第 t 次迭代。

在使用投影梯度方法之前, 需要定义投影函数 $P(x)$ 来限制问题求解的范围

$$P(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & 1 \leq x \end{cases} \quad (6)$$

结合式(5)(6), 变量 M 、 H 和 T 的迭代求解公式为

$$\begin{aligned}M_{t+1} &= \max(0, M_t - \eta \nabla_M L_t) \\ H_{t+1} &= \max(0, H_t - \eta \nabla_H L_t) \\ T_{t+1} &= P(T_t - \eta \nabla_T L_t)\end{aligned}\quad (7)$$

式中, η 为变量学习效率因子, t 和 $t+1$ 表示相邻的两次迭代状态。算符 $\max$ 表示取两个因子间的最大值。

求解出式(4)最优解如下:

1) 输入初始值: 标签—图像矩阵 T_0 , $\lambda = 10$, $\eta = 0.5$, $\varepsilon = 0.000\,000\,1$, $r = 100$, 随机初始化 M 、 H , $T = T_0$ 。

2) 开始迭代, 设定最大迭代次数为 1 000:

(1) 固定变量 H 、 T 的值, 根据式(7)定义的 M 迭代更新公式求解最优的 M ;

(2) 使用 M 的最优解替代其上一次迭代的值, 固定变量 M 、 T 的值, 根据式(7)定义的 H 迭代更新公式求解 H 的最优解;

(3) 使用 M 、 H 的最优解替代其上一次迭代的值, 同时固定变量 H 、 M , 根据式(7)给出的迭代公式求解最优的 T ;

(4) 进行条件判断, 如果 $\|L_{t+1}(T, M, H) - L_t(T, M, H)\| \leq \varepsilon$ 或已达最大运行次, 终止迭代,

否则继续重复以上步骤。

3) 输出迭代结果, 估计出的最优的标签—图像矩阵 T 。

在求解步骤中, 各种初始参数由网格搜索的方法获取而得。其中, 正则化因子 λ 由 0.5 至 50, 以 0.5 的步长搜索最佳参数值; 学习因子 η 由 0.1 至 1, 以 0.05 的步长搜索; 低秩空间维数 r 由 50 至 1 000, 以 50 的步长搜索。参数 ε 决定算法迭代的终止条件。

最后, 需要根据式(4)的最优解 T 将标签重新赋给对应的图像。由于矩阵 T 中元素表示为标签与图像相关的程度, $T(i, j)$ 的取值范围为 $[0, 1]$ 。为了获得精确的标注结果, 需要将 $T(i, j)$ 二值化为 0 和 1。本文采用阈值法对取 $T(i, j)$ 二值化。设定阈值 θ , 如果 $T(i, j) > \theta$, 则 $T(i, j)$ 重新取值为 1; 否则取值为 0。

2 实验结果及分析

为了验证本文图像标签丰富算法的有效性, 从国外图像分享网站 Flickr (www.flickr.com) 上下载了 21 744 幅图像, 共有 3 539 个独立标签, 每幅图像平均拥有 2.171 个标签。由于网络图像的标签中会出现错误的拼写, 因此在进行标签丰富之前对所有图像的标签进行了预处理。使用语义词典 WordNet (<http://wordnet.princ-eton.edu/>) 对所有单词进行拼写检查, 删除错误的拼写。同时, 删除了标签数目出现少于 10 次的标签, 剩余 1 529 个独立标签。经过预处理后, 构建标签—图像矩阵 T_0 , 矩阵大小为 $1\,529 \times 28\,744$ 。由于该矩阵需要的内存较大, 使用了 12 GB 的内存来进行实验。

为了准确描述图像的视觉内容, 采用 CPAM (colored pattern appearance model) 视觉特征^[14]。CPAM 从多幅图像里学习出一个视觉轮廓词典, 然后将一幅图像划分成不重复的 4×4 方格。每个方格的像素用词典里单词表示, 统计整幅图像里各种视觉轮廓单词出现的频率。CPAM 中包含两类视觉轮廓词典, 一类是灰度轮廓, 用 32 维特征向量表示; 另一类是色彩轮廓, 也用 32 维特征向量表示。将两者级联成一个 64 维的特征向量来表示图像的视觉内容, 即上文所提到的 V_j 。视觉词典中的视觉单词如图 2 所示, 每个视觉单词可以反映出一种纹理变化。

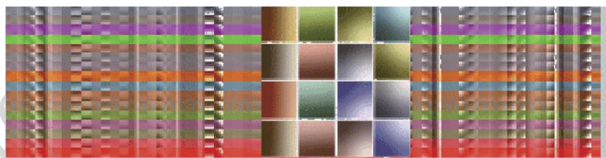


图2 视觉词典单词示例

Fig. 2 Samples of appearance prototypes

为了进一步对比验证本文算法,使用图像标签丰富后的平均标注准确度作为算法的测试指标。平均标注准确度定义为

$$p = \frac{N_p}{N} \quad (8)$$

式中, N_p 表示所有图像增加的标签中能准确描述图像视觉内容的标签数目; N 表示丰富算法后所有增加的标签数目。由于本文采用的图像较多,无法一一进行验证。因此,将图像标签丰富后标签增加最多的1 000幅图像进行人工标注,以此作为算法对比的基准。

同时,采用文献[7]和文献[10]的算法进行对比实验,来验证本文算法的有效性。文献[7]只利用标签间的共现关系来丰富图像的标签。文献[10]通过最大化图像之间标签语义与视觉内容的一致性来丰富图像的标签。对比结果如表1所示。从表1中可以看出,本文算法采用正则化的最优方法,取得了更好的标签丰富效果。与两种对比方法相比,本文算法的优势是不仅利用了标签的语义相关性,而且从总体上利用视觉离散度最小化来寻找图像与标签的关系。

表1 算法比较结果统计

Table 1 Comparisons of algorithms

算法	增加的标签总数数目	平均标注准确度/%
文献[7]	5 462	32.5
文献[10]	3 627	67.5
本文	5 104	79.8

图3是利用本文算法对图像标签进行丰富的结果。本文算法中二值化阈值 θ 取值为0.1,所有小于0.1的 $T(i,j)$ 都被设置为0。从图3可以看出,本文算法可以丰富图像不完备的标签。如图3(b)所示,通过本文算法,可以弥补与初始的标签‘flower’语义相关的标签‘grass’、‘leaf’和图像内容相关的标签‘sky’。



源标签: grass
丰富后: grass, flower
(a)



源标签: flower
丰富后: grass, flower, sky, leaf
(b)



源标签: flower
丰富后: flower, sky, people
(c)



源标签: people
丰富后: sky, people, water, sand
(d)



源标签: car
丰富后: car, road, building
(e)



源标签: dog
丰富后: dog, sky
(f)

图3 图像标签丰富示例

Fig. 3 Exemplar images of tag enrichment

3 结 论

利用正则化的矩阵分解算法,将图像标签的丰富问题,转化为矩阵分解问题。利用非负矩阵的分解构建了潜在的低秩空间,并在低秩空间内消除噪声,寻找标签间的相关性,以此提高标签与图像之间的相关性。同时,利用类内视觉离散度作为正则化项,降低与图像视觉内容不相关的标签的作用,提高算法的有效性效果。最后,在构建的欠完备数据集上验证本文算法的有效性。本文算法能运行在较大规模的图像集上,有效地丰富图像的标签。

参考文献 (References)

[1] Liu Y, Fan J L. A survey on state-of-the-art techniques in content-based image retrieval [J]. Journal of Xi'an Institute of Posts and Telecommunications, 2012, 17(2):1-8. [刘颖, 范九伦. 基于内容的图像检索技术综述[J]. 西安邮电学院学报, 2012, 17(2):1-8.] [DOI:10.3969/j.issn.1007-3264.2012.02.002]
[2] Jia J X, Wang X P, Ren E E. Image retrieval technique based

- on global features [J]. Automation & Instrumentation, 2011, (1):101-103. [贾君霞, 王小鹏, 任恩恩. 基于全局特征的图像检索技术[J]. 自动化与仪器仪表, 2011, (1): 101-103.] [DOI:10.3969/j.issn.1001-9227.2011.01.039]
- [3] Sun H, Wang C, Wang R S. A review of local invariant features [J]. Journal of Image and Graphics, 2011, 16(2):141-151. [孙浩, 王程, 王润生. 局部不变特征综述[J]. 中国图象图形学报, 2011, 16(2): 141-151.] [DOI: 10.11834/jig.20110207]
- [4] Halpin H, Robu V, Shepherd H. The complex dynamics of collaborative tagging [C]//Proceedings of the 16th International Conference on World Wide Web. Canada: ACM, 2007, 211-220. [DOI: 10.1145/1242572.1242602]
- [5] Xia Z Q, Peng J Y, Feng X Y, et al. Social tag enrichment via automatic abstract tag refinement [C]//Proceedings of the 13th Pacific-rim Conference on Multimedia. Singapore: Springer Berlin Heidelberg, 2012, 198-209. [DOI: 10.1007/978-3-642-34778-8_18]
- [6] Carneiro G, Chan A B, Moreno P J, et al. Supervised learning of semantic classes for image annotation and retrieval [J]. IEEE transactions on Pattern Analysis and Machine Intelligence, 2007, 29(3): 394-410. [DOI: 10.1109/TPAMI.2007.61]
- [7] Jäschke R, Marinho L, Hotho A, et al. Tag recommendations in folksonomies [C]// Knowledge Discovery in Databases: PKDD 2007. Poland: Springer Berlin Heidelberg, 2007, 506-514. [DOI: 10.1007/978-3-540-74976-9_52]
- [8] Lee S, DeNeve W, Ro Y M. Tag refinement in an image folksonomy using visual similarity and tag co-occurrence statistics [J]. Signal Processing: Image Communication, 2010, 25(10):761-773. [DOI:10.1016/j.image.2010.10.002]
- [9] Shen Y, Peng J Y, Feng X Y, et al. Multiple instance learning with missing object tags [C]//Proceedings of the 3rd International Conference on Internet Multimedia Computing and Service. Chengdu: ACM, 2011: 9-12. [DOI: 10.1145/2043674.2043677]
- [10] Wu L, Jin R, Jain A K. Tag completion for image retrieval [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(3):716-727. [DOI:10.1109/TPAMI.2012.124]
- [11] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization [J]. Nature, 1999, 401(6755), 788-791. [DOI: 10.1038/44565]
- [12] Li L, Zhang Y J. A survey on algorithms of non-negative matrix factorization [J]. Acta Electronica Sinica, 2008, 36(4):737-743. [李乐, 章毓晋. 非负矩阵分解算法综述[J]. 电子学报, 2008, 36(4): 737-743.] [DOI: 10.3321/j.issn:0372-2112.2008.04.023]
- [13] Lin C J. Projected gradient methods for nonnegative matrix factorization [J]. Neural Computation, 2007, 19(10): 2756-2779. [DOI: 10.1162/neco.2007.19.10.2756]
- [14] Qiu G. Indexing chromatic and achromatic patterns for content-based colour image retrieval [J]. Pattern Recognition, 2002, 35(8): 1675-1686. [DOI:10.1016/S0031-3203(01)00162-5]