

# 一种基于概率距离尺度学习的图像标注方法\*

田璟 郭智 黄宇 黄廷磊 付琨

(中国科学院电子学研究所地理空间信息处理与应用系统技术重点实验室 北京 100190)

**摘要:** 大部分传统的基于距离尺度学习的图像标注方法假设训练样本之间存在确定的约束关系。这样的假设在小规模、精准的数据集上可以取得一定的效果。但是当数据量较大,每幅图像有多个标注且部分标注含有噪声时,这种过于理想的假设则不成立。提出了一种基于概率主题模型的距离尺度学习方法,通过概率主题模型挖掘这种模糊的、潜在的不确定辅助信息,并应用于后续的距离尺度学习中得到语义上距离尺度,改进基于搜索的图像标注性能。在 Flickr 数据集上的实验证明所提出方法比当前方法具有更好的图像自动标注性能。

**关键词:** 图像标注;距离尺度学习;概率主题模型;Gibbs 采样

**中图分类号:** TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.2040

## Automatic image annotation method based on probabilistic distance metric learning

Tian Jing Guo Zhi Huang Yu Huang Tinglei Fu Kun

(Key Laboratory of Technology in Geo-spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing100190, China)

**Abstract:** Most of the traditional approaches for image annotation based on distance metric learning generally suppose that the constraints on training data is explicit, which only works on small datasets with exact labels. As the scale of the dataset becomes larger and most images are accompanied by noisy labels, this ideal assumption will be not efficient. In this paper, we propose a novel distance metric learning method based on probabilistic topic model. The uncertain and latent side information can be mined by a probabilistic topic model and afterwards used in distance metric learning. The learned semantic distance metric then can be used in the search-based image annotation. Experiments on Flickr dataset demonstrate that the proposed model outperforms the state-of-the-art annotation methods.

**Keywords:** image annotation; distance metric learning; probabilistic topic model; Gibbs sampling

### 1 引言

大多数机器学习算法依赖于输入样本之间的距离度量,学习器的性能除了由自身的结构决定以外,很大程度上依赖于一个能够真实反映样本之间关系的距离度量。对基于搜索的图像标注<sup>[1]</sup>方法而言,最重要的环节是选取科学、有效的距离尺度来衡量待标注图像与搜索结果图像之间的相似度,因此,距离尺度学习算法是基于搜索的图像标注方法的基石。传统意义上的距离尺度学习方法,如相关成分分析(relevant con analysis, RCA)<sup>[2]</sup>,通过一个全局的线性变换,使变换后的数据形成同类模式相距较近,不同模式相距较远的集合。其输入是确定的约束信

息,例如类别标签、成对约束等,其中成对约束信息定义了样本两两之间的相似(“must-link”)与不相似(“cannot-link”)关系,相似的样本组成一个“chunklet”。以上形式的辅助信息一般来源于人工标注,如CBIR的相关反馈系统<sup>[3]</sup>。这样的假设在小规模、精准的数据集上可以取得一定的效果。当数据量较大,每幅图像有多个标注且部分标注含有噪声时,则不能应用这种过于理想的约束信息。例如,当两幅图像的标注之间只存在一部分重叠时便无法区分二者是否为“must-link”约束。

针对以上问题,本文提出了一种基于多模态相关主题模型的概率相关成分分析方法(multi-modal correlated topic model based relevant component analysis,

收稿日期:2015-03

\* 基金项目:国家 863 计划(2014AA7013033)资助项目

MMCTM-RCA)。首先,采用多模态 LDA (multi-modal LDA, MMLDA) [4] 的框架来建模图像、图像标签文本 2 种模态的数据,同时引入相关主题模型 (correlated topic model, CTM) [5] 中的逻辑正态分布作为主题分布参数的先验分布,利用相关系数矩阵来显式建模同一文档中各维度之间的相关性先验。然后,采用基于 Gibbs 采样 [6] 算法求解概率模型,得到训练样本与隐含主题之间的概率从属关系,并以此作为概率化的辅助信息。进而,将得到的概率辅助信息作为改进的相关成分分析算法的输入,并提出相应的优化算法,学习得到一个最优的距离尺度  $M$ 。最后,所学习得到的距离尺度  $M$  可以作为相似性图像搜索阶段的度量标准,通过对相似图像的标签进行整合、排序、筛选可得到最终的图像标注结果。在 NUS-WIDE 数据集 [7] 上的实验表明,所提出的 MMCTM-RCA 方法在含有噪声的大规模图像集合上可以显著提高标注的性能。

## 2 基于多模态相关主题模型的概率相关成分分析方法

本文用基于多模态相关主题模型的概率相关成分分析方法,实现图像的自动标注。共分 2 个阶段:语义距离尺度学习阶段和图像自动标注阶段。

在语义距离尺度学习阶段:

1) 采用多模态主题模型生成概率辅助信息;

2) 将得到的概率辅助信息作为改进的相关成分分析算法的输入,学习得到一个语义距离尺度  $M$ 。

在图像自动标注阶段:将所学习得到的距离尺度  $M$  作为相似性图像搜索阶段的度量标准,通过对相似图像的标签进行整合、排序、筛选得到最终的图像标注结果。

### 2.1 基于多模态相关主题模型的概率约束关系生成

本文提出的多模态相关主题模型,采用 MMLDA 的框架来建模 2 种模态的数据,并在建模主题分布的先验时借鉴 CTM 的主题分布先验设定,因此称之为多模态相关主题模型 (multi-modal correlated topic models, MMCTM)。

其概率图模型如图 1 所示,与原始的 CTM 相比,MMCTM 增加了对图像模态的建模,有 2 条路径分别代表文本模态与视觉模态。从图模型中可以看出,每个文档都由两种模态的数据组成,假设整个文档集中文档的数目为  $D$ ,对其中的任意一篇文档  $d$ ,均有  $M_d$  个文本词,  $N_d$  个视觉词组成,两种词的词典维度分别为  $W$  和  $V$ ,主题的维度为  $K$ 。为了对其进行完整的后验估计,给模型的参数  $\mu$ 、 $\Sigma$  设置了超参数先验分布,为正态-逆-威沙特分布 [8]。

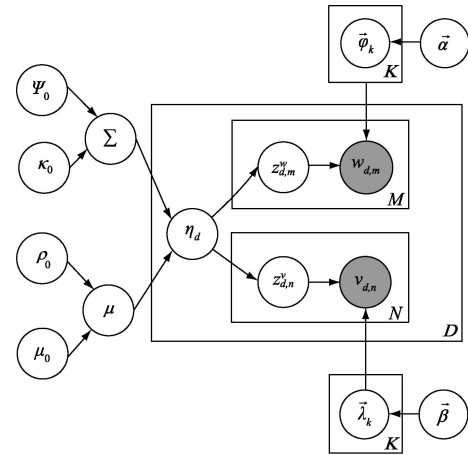


图 1 MMCTM 概率模型

对每一个文档  $d$  中的每一个文本词  $w_{d,m}$  和视觉词  $v_{d,n}$ ,该模型所描述的产生式过程如下:

1) 生成主题分配情况。依据正态分布选择文档的主题分配情况  $\eta_d, \eta_d | \{\mu, \Sigma\} \sim N(\mu, \Sigma)$ 。其中  $(\mu, \Sigma) \sim NIW(\mu_0, \rho_0, \Psi_0, \kappa_0)$ 。

2) 生成主题分配的分布。依据逻辑斯谛回归将上一步中得到的  $\eta_d$  转化成为主题分布  $\theta_d = f(\eta_d) = \exp \frac{\eta_d}{\sum_i \exp \eta_i}$ 。

3) 生成文本词  $w_{d,m}$

①按照多项式分布选择选择一个主题,  $z_{d,m}^w | \theta_d \sim Mult(\theta_d)$ 。

②按照选定主题  $z_{d,m}^w$  条件下的多项式分布选择一个词  $w_{d,m} | z_{d,m}^w, \varphi_{1:K} \sim Mult(\varphi_{z_{d,m}^w})$ 。

4) 生成视觉词  $v_{d,n}$

按照多项式分布选择选择一个主题,  $z_{d,n}^v | \theta_d \sim Mult(\theta_d)$ 。

5) 按照选定主题  $z_{d,n}^v$  条件下的多项式分布选择一个词  $v_{d,n} | z_{d,n}^v, \lambda_{1:K} \sim Mult(\lambda_{z_{d,n}^v})$ 。

模型的任务是计算上述的后验概率。虽然基于变分-EM 的方法可以实现其求解,但是一般情况下,变分 EM 容易受初始化状态的影响。因此,提出基于 Gibbs 采样的方法,根据共轭先验,简化模型的计算。

#### 2.1.1 采样主题变量 $z_{d,m}, z_{d,n}$

对主题变量  $z_{d,m}$ ,根据贝叶斯准则,其更新公式可直接得到为:

$$p(z_{d,m} = k | \cdot) \propto p(z_{d,m} = k | \eta_d) \cdot p(w_{d,m} | z_{d,m} = k, \alpha) \quad (1)$$

式中:第 1 项展开可得:

$$p(z_{d,m} = k | \eta_d) \propto e^{\eta_k} \quad (2)$$

式(2)的第 2 项展开为:

$$p(w_{d,m} | z_{d,m} = k, \alpha) = \int_{\varphi_k} p(\varphi_k | \beta) p(w_{d,m} | \varphi_k) d\varphi_k \quad (3)$$

且因为  $p(\varphi_k | \beta)$  与  $p(w_{d,m} | \varphi_k)$  组成一对狄利克雷-多项共轭分布,因此上式中的  $\varphi_k$  可约去,可得:

$$p(w_{d,m} | z_{d,m} = k, \alpha) = \frac{n_{z_{d,m}}^{w_{d,m}} + \alpha}{\sum_w (n_{z_{d,m}}^{w_{d,m}} + \alpha)} \quad (4)$$

式中:  $n_{z_{d,m}}^{w_{d,m}}$  表示文本词  $w_{d,m}$  被分配到当前主题  $z_{d,m}$  的次数。

进而,得到  $z_{d,m}$  的更新公式:

$$p(z_{d,m} = k | \cdot) \propto e^{\eta} \cdot \frac{n_{z_{d,m}}^{w_{d,m}} + \alpha}{\sum_w (n_{z_{d,m}}^{w_{d,m}} + \alpha)} \quad (5)$$

同理,  $z_{d,n}$  的更新公式如下所示:

$$p(z_{d,n} = k | \cdot) = p(v_{d,n} = k | \eta_d) \cdot p(v_{d,n} | z_{d,n} = k, \beta) \propto e^{\eta} \cdot \frac{n_{z_{d,n}}^{v_{d,n}} + \beta}{\sum_v (n_{z_{d,n}}^{v_{d,n}} + \beta)} \quad (6)$$

### 2.1.2 采样主题分配情况 $\eta$

给定文档  $d$  中文本部分以及视觉部分的主题分配  $z_d^w$  和  $z_d^v$ 、参数  $\mu$  和  $\sum$ , 以及当前文档在除  $k$  以外的其他主题上的权重  $\eta_d^k$ , 需要对当前文档在主题  $k$  上的权重  $\eta_d^k$  进行采样。根据贝叶斯规则,  $\eta_d^k$  的条件概率为:

$$p(\eta_d^k | \cdot) \propto p(\eta_d^k, \eta_d^{-k} | \mu, \sum) \cdot p(z_d^w | \eta_d^k, \eta_d^{-k}) \cdot p(z_d^v | \eta_d^k, \eta_d^{-k}) \quad (7)$$

式(7)的第一项为正态分布:

$$p(\eta_d^k, \eta_d^{-k} | \mu, \sum) = N(\eta_d^k, \eta_d^{-k} | \mu, \sum) \quad (8)$$

式(7)的第二、三项为:

$$p(z_d^w | \eta_d^k, \eta_d^{-k}) = \prod_{i=1}^M \frac{\exp(\eta_d^i)}{\sum_{j=1}^K \exp(\eta_d^j)} \quad (9)$$

$$p(z_d^v | \eta_d^k, \eta_d^{-k}) = \prod_{i=1}^N \frac{\exp(\eta_d^i)}{\sum_{j=1}^K \exp(\eta_d^j)} \quad (10)$$

因此,先验与似然之间属于非共轭分布,一般不能从后验概率中直接采样。而基于数据推广(aata argumentation)<sup>[9]</sup>的方法可以从此概率分布采样。首先,假设存在一个  $\lambda$  满足:

$$p(\eta_d^k, \lambda_d^k | \eta_d^{-k}, z_d) \propto \frac{1}{2^N} \exp\left(\kappa_d^k \rho_d^k - \frac{\lambda_d^k (\rho_d^k)^2}{2^N}\right) PG(\lambda_d^k, N, 0) N(\eta_d^k | \mu_d^k, (\sigma^2)^k_d) \quad (11)$$

通过求该联合概率的边缘分布可得到  $\eta_d^k$  的条件概率,具体的操作方法是进行若干轮 Gibbs 采样来获得一个  $(\eta_d^k, \lambda_d^k)$ , 然后丢弃  $\lambda_d^k$  来获得  $\eta_d^k$ 。 $\eta_d^k$  和  $\lambda_d^k$  相对于彼此的条件概率分别为:

$$p(\eta_d^k | \lambda_d^k) \propto \exp\left(\kappa_d^k \rho_d^k - \frac{\lambda_d^k (\rho_d^k)^2}{2^N}\right) N(\eta_d^k | \mu_d^k, (\sigma^2)^k_d) \quad (12)$$

$$p(\lambda_d^k | \eta_d^k) \propto \exp\left(-\frac{\lambda_d^k (\rho_d^k)^2}{2^N}\right) PG(\lambda | N, 0) \quad (13)$$

按照式(12)和(13)进行循环采样,即可得到符合式

(11)分布的  $(\eta_d^k, \lambda_d^k)$ 。

### 2.1.3 采样超参数先验

对模型的参数设置了超参数先验分布  $(\mu, \sum)$ , 采样时利用正态分布  $N(\eta_d^k, \eta_d^{-k} | \mu, \sum)$  和正态-逆-威沙特分布  $NIW(\mu_0, \rho_0, \Psi_0, \kappa_0)$  之间的共轭性,计算得到后验概率:

$$p(\mu, \sigma^2 | \cdot) \sim NIW(\mu_0', \rho_0', \Psi_0', \kappa_0') \quad (14)$$

$$\mu_0' = \frac{\rho_0}{\rho_0 + D} \mu_0 + \frac{\rho_0 + D}{\rho_0} \bar{\eta} \quad (15)$$

$$\rho_0' = \rho_0 + D \quad (16)$$

$$\Psi_0' = \Psi_0 + Q + \frac{\rho_0 D}{\rho_0 + D} (\bar{\eta} - \mu_0) (\bar{\eta} - \mu_0)^T \quad (17)$$

$$\kappa_0' = \kappa_0 + D \quad (18)$$

式中:  $\bar{\eta} = \frac{1}{D} \sum_d \eta_d$  为  $\eta$  的均值,  $Q = \sum_d (\eta_d - \bar{\eta}) (\eta_d - \bar{\eta})^T$  为  $\eta$  的散布矩阵。

最终,已知模型的参数,对于任一图像  $x$  可以计算条件概率  $p(x | z^w, z^v)$  并将其作为隐“chunklet”与图像  $x$  之间隐含从属关系的表征,并作为距离尺度学习算法的输入。

## 2.2 概率距离尺度度量

假设  $x_i$  为图像  $i$  所代表的特征向量,  $z_k$  为  $K$  个隐“chunklet”中的一个,且  $\mu_k$  为隐“chunklet” $z_k$  的均值,这样所有的隐“chunklet”的均值组成一个均值矩阵  $\mu = (\mu_1, \dots, \mu_K)$ 。而矩阵  $P = (p_1, \dots, p_I)$  代表  $I$  个实例中的每一个与隐“chunklet”的从属关系,其中  $p_i = (p_i^{(1)}, \dots, p_i^{(K)})$  为第  $i$  个实例的概率分布,且  $p_i^{(k)}$  代表给定隐“chunklet” $z_k$  其分配给第  $i$  个实例的概率,即  $p_i^{(k)} = p(x_i | z_k)$ 。

在算法中,使用前一阶段得到的估计量对从属关系矩阵  $P$  进行了初始化。即令  $P_0 = [p(x_i | z_k)]_{I \times K}$ , 其中  $p(x_i | z_k)$  为前一阶段得到的结果。

距离尺度学习的目标为学习一个最优的距离尺度  $M \in \mathbf{R}^{d \times d}$ , 而结合隐含从属关系估计的概率相关成分分析其所对应的优化问题如下:

$$\begin{aligned} & \min_{M > 0, \mu, P} \sum_{i=1}^I \sum_{k=1}^K p_i^{(k)} \|x_i - \mu_k\|_M^2 - \lambda \log \|M\| \\ \text{s.t. } & \|P - P_0\|_F^2 \leq \gamma, \\ & \sum_k p_i^{(k)} = 1, p_i^{(k)} \geq 0, i = 1, \dots, I \end{aligned} \quad (19)$$

式中:参数  $\gamma \geq 0$  代表先验从属关系矩阵  $P_0$  与待求矩阵  $P$  之间的差值,  $\lambda$  为正则常量,  $\|\cdot\|_F$  代表矩阵的  $F$  范数。目标函数的第 1 项是最小化实例与“chunklet”均值的平方和,第 2 项是避免求解  $M$  时对整个求解空间的压缩;约束条件中,第一项旨在约束求得的矩阵  $P$  与先验矩阵  $P_0$  的偏差尽量小,第 2 项为概率约束。

所以该优化问题为包含 3 个变量  $M, P$ , 以及  $\mu$  的非线性优化问题,且  $\mu$  可以在求得  $P$  之后求解。而全局最优解

往往难于得到。为此,借鉴文献[10]中基于不动点迭代的迭代算法,提出本文的迭代优化算法,共包含3部分:1)固定  $P$  和  $\mu$ , 优化  $M$ ; 2) 固定  $M$  和  $\mu$ , 优化  $P$ ; 3) 固定  $P$  和  $M$ , 求解  $\mu$ 。

第1步,等价于求解标准的RCA,即  $M = \frac{1}{\lambda} C^{-1}$ , 式中:  $C$  为给定“chunklet”在给出矩阵  $P$  的协方差。而第3步可通过公式  $\mu = P^T X$  求解得到。

第2步,当固定  $M$  和  $\mu$  时,该优化问题可以重新写作:

$$\begin{aligned} \min_P \sum_{i=1}^n \sum_{k=1}^m p_i^{(k)} \|x_i - \mu_k\|_M^2 + \frac{\gamma}{2} \|P - P_0\|_F^2 \\ \text{s. t. } \sum_k p_i^{(k)} = 1, p_i^{(k)} \geq 0, i = 1, \dots, n \end{aligned} \quad (20)$$

这样,原始的问题转化为一个二次规划优化问题。假设所有的  $p_i$  在式(9)中是相互独立的。则式(9)可以转化为  $n$  个相互独立的优化任务,如下:

$$\begin{aligned} \min_{p \in \mathbf{R}^m} \sum_{k=1}^m p_k \|x_i - \mu_k\|_M^2 + \frac{\gamma}{2} \|p - p_0\|_2^2 \\ \text{s. t. } \sum_{k=1}^m p_k = 1, p_k \geq 0, k = 1, \dots, m \end{aligned} \quad (21)$$

使用拉格朗日乘法对上述目标函数进行求解,得到目标函数如下:

$$L = f^T p + \frac{\gamma}{2} \|p - p_0\|_2^2 + \rho \cdot \left( \sum_k p_k - 1 \right) - \eta \cdot \rho \quad (22)$$

式中:  $f^T = (\|x_i - \mu_1\|_M^2, \dots, \|x_i - \mu_m\|_M^2)$ ,  $\rho$  为拉格朗日乘子。对  $p_k$  求导,可得:

$$\frac{\partial L}{\partial p_k} = f_k + \gamma(p_k - p_{0k}) + \rho - \eta_k = 0 \quad (23)$$

通过运用KKT条件,即只要  $p_k > 0$ , 则  $\eta_k = 0$ 。这样,若  $p_k > 0$ , 可得:

$$p_k = p_{0k} - \frac{1}{\gamma}(\rho + f_k) \quad (24)$$

因为已知  $p_k \geq 0$ , 可得:

$$p_k = \max(0, p_{0k} - \frac{1}{\gamma}(\rho + f_k)) \quad (25)$$

然后,通过排序的方法即可求得  $\rho$  的最优解。

第3步,通过  $\mu = P^T X$  求解得到。

### 3 实验

#### 3.1 数据集及参数设置

将所提出模型与多个基于距离尺度学习的方法进行了对比,包括:相关成分分析(RCA)[2],区分性成分分析(discriminative component analysis, DCA)[11]以及UDML(unified distance metric learning)[12]。

实验在NUS-WIDE数据集[7]的一个子集上进行。共选取了200 000幅图像,并分成3部分:

- 1) 16 000幅做概率约束估计;
- 2) 2 000幅做检索测试;

3) 剩余做数据库。在特征提取时,第1组提取了SIFT特征并生成视觉词。以便于在主题模型中进行建模;第2,3组按照提取颜色、边缘、纹理等特征,方便进行特征的相似性度量。

对于传统的距离尺度学习方法,包括RCA, DCA以及UDML,对训练数据按照视觉特征和文本标注进行聚类,然后将聚类结果作为距离尺度学习的输入,其中“chunklet”的取值统一设为500个。对MMCTM-RCA,在概率约束评估阶段,设置隐“chunklet”的个数为500个,设置视觉词的维度为1 000维,文本词维度为2 000维;在概率距离尺度学习阶段,惩罚因子  $\gamma$  取值为0.5;在相似性图像搜索阶段,设定搜索得到的相似图像的数量为30幅,并将得到的候选标注中的前10个按照顺序赋给待标注图像。

#### 3.2 正确率、召回率评估

图2和图3给出了不同模型在前1~10个标注词上的平均正确率和平均召回率比较情况。图中横坐标分别代表前1~10个标注词。

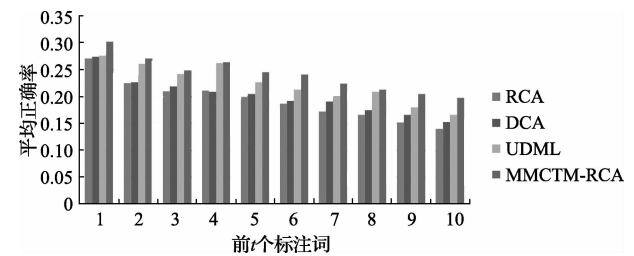


图2 各方法在标注词个数设定下的平均正确率比较

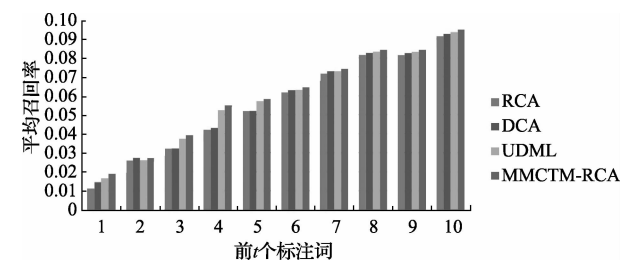


图3 各方法在标注词个数设定下的平均召回率比较

观察图2所示的平均正确率可以发现,该文所提出的MMCTM-RCA模型1~10个标注词的情形下性能均优于其他对比模型。例如,在前6个标注词的情况下,MMCTM-RCA相对于RCA, DCA以及UDML分别有2.8%, 4.9%和5.4%的提升。同样,MMCTM-RCA模型的平均召回率在全部标注词的情况下,同样优于对比方法。

#### 3.3 不同近邻数目下的标注性能

相似性搜索阶段近邻图像数目同样影响标注的性能。图4给出了当近邻数目分别取10、20、30、40和50

时 MMCTM-RCA 的标注平均正确率。从图中可见,总体而言,当相似性图像为 40 时标注性能最佳。这主要是因为,当近邻数目过大时,噪声的影响同样随之增大;而当数目过小时,又有可能丢失部分有用的标注词。

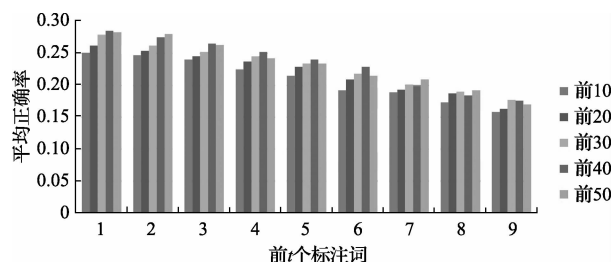


图 4 在不同近邻数目设定下的平均正确率比较

#### 4 结 论

提出了一种基于多模态相关主题模型的概率相关成分分析方法。首先采用多模态 LDA 与相关主题模型相结合的框架来建模图像、图像标签文本 2 种模态的数据,学习得到概率化的辅助信息。然后,将得到的概率辅助信息作为改进的相关成分分析算法的输入,学习得到一个最优的语义距离尺度。最后,所学习得到的距离尺度可以作为相似性图像搜索阶段的度量标准,通过对相似图像的标签进行整合、排序、筛选可得到最终的图像标注结果。在 Flickr 数据集上的实验表明,所提出的 MMCTM-RCA 方法在含有噪声的大规模图像集合上可以显著提高标注的性能。

#### 参 考 文 献

- [1] 张滢,齐美彬,周云,等. 基于特征提取和多示例学习的图像区域标注[J]. 电子测量与仪器学报, 2014, 28(8):909-914.
- [2] BAR-HILLEL A, HERTZT, SHENTALN, et al. Learning a mahalanobis metric from equivalence constraints [J]. Journal of Machine Learning Research, 2005, 6(6):937-965.
- [3] HOI C H, LYU M R. A novel log-based relevance feedback technique in content-based image retrieval [C]. Proceedings of ACM Multimedia Conference, 2004: 24-31.
- [4] RAMAGE D, HEYMANN P, MANNING C D, et al. Clustering the tagged web[C]. Proceedings of the ACM International Conference on Web Search and Data Mining, 2009: 54-63.
- [5] BLEI D, LAFFERTY J. Correlated topic models[J]. Advances in neural information processing systems, 2006(18): 147.
- [6] NEAL R M. Markov chain sampling methods for Dirichlet process mixture models [J]. Journal of computational and graphical statistics, 2000, 9(2): 249-265.
- [7] CHUA T S, TANG J, HONG R, et al. NUS-WIDE: a real world web image database from National University of Singapore[C]. Proceedings of ACM International Conference on Image and Video Retrieval, Greece, 2009:48.
- [8] MURPHY K P. Conjugate bayesian analysis of the Gaussian distribution[J]. def, 2007:16.
- [9] TANNER M A, WONG W H. The calculation of posterior distributions by data augmentation [J]. Journal of the American Statistical Association, 1987, 82(398):528 - 540.
- [10] BEZDEK J C, HATHAWAY R J. Convergence of alternating optimization[J]. Neural, Parallel & Scientific Computations, 2003, 11(4):351-368.
- [11] HOI S C, LIU W, LYU MR, et al. Learning distance metrics with contextual constraints for image retrieval[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2006: 2072-2078.
- [12] WU P, HOI S C, ZHAO P, et al. Mining social images with distance metric learning for automated image tagging [C]. Proceedings of the ACM International Conference on Web Search and Data Mining, 2011: 197-206.

#### 作 者 简 介

田璟, 1987 年出生, 博士研究生。主要研究方向为数据挖掘、图像自动标注。  
E-mail: tianjing0303@163.com