

基于网络搜索量的扩展属性图像检索

张 芬¹, 孔祥维¹, 宁 斐², 贾 则³

(1. 大连理工大学 信息与通信工程学院, 辽宁 大连 116024; 2. 中国舰船研究设计中心, 上海 221108;
3. 中国人民解放军 91439 部队, 辽宁 大连 116041)

摘 要: 现有基于属性的图像检索主要依赖于预标签属性, 使用户只能通过预定义的属性来搜索目标。基于扩展属性的方法则可使用户输入与预标签属性相关的查询词, 而非仅选择预定义属性。为此, 设计基于网络检索量的扩展属性学习方法。利用 Wiktionary 挖掘扩展属性, 将其与 WordNet 所得结果相结合, 使用由百度指数和谷歌趋势获得的预定义属性及其相应扩展属性的相对平均检索量度量用户偏好, 并通过一致性度量方法验证扩展属性的可靠性。实验结果表明, 该方法可有效提高图像检索性能。

关键词: 图像检索; 扩展属性; 语义关系度量; 相对平均检索量; 用户偏好; 一致性度量

中文引用格式: 张 芬, 孔祥维, 宁 斐, 等. 基于网络搜索量的扩展属性图像检索[J]. 计算机工程, 2017, 43(9): 276-280, 287.

英文引用格式: ZHANG Fen, KONG Xiangwei, NING Fei, et al. Image Retrieval by Extended Attributes Based on Web Search Amount[J]. Computer Engineering, 2017, 43(9): 276-280, 287.

Image Retrieval by Extended Attribute Based on Web Search Amount

ZHANG Fen¹, KONG Xiangwei¹, NING Fei², JIA Ze³

(1. School of Information and Communication Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China;
2. China Ship Research and Design Center, Shanghai 221108, China; 3. Unit 91439 of PLA, Dalian, Liaoning 116041, China)

[Abstract] Existing attribute-based image retrieval principally relies on pre-labeled attributes, which restricts users to use only the pre-defined attribute to search the intended targets. Extended attribute-based methods turn users from passively choosing pre-defined attributes to actively inputting query words which are pertinent to the pre-labeled attributes. This paper proposes an extended attribute learning method based on Web search amount. It uses Wiktionary to mine extended attributes and combines the results with that of WordNet. After that, it exploits relative average retrieval amount of attributes obtained from Baidu Index and Google Trends to measure user preference, then adopts a consistency measure method to validate the reliability of the extended attributes. Experimental results demonstrate the significant image retrieval performance improvements of the proposed method.

[Key words] image retrieval; extended attribute; Semantic Relation (SR) measure; relative average retrieval amount; user preference; consistency measure

DOI: 10.3969/j.issn.1000-3428.2017.09.048

0 概述

近年来, 属性的概念已在计算机视觉领域引起广泛关注。属性被视为一个富有表现力的中间层, 在链接低层特征和高层语义之间具有重要作用。迄今为止, 已完成大量属性相关的研究工作, 内容涉及属性学习^[1-3]、目标识别^[4-6]、人脸验证^[7-8]、图像分类^[9-10]和图像检索^[11-14]等。

目前在图像检索方面的研究主要有基于文本的图像检索^[15]、基于内容的图像检索^[16]和基于属性的

图像检索, 本文主要研究基于属性的图像检索。属性表示具有直观性和跨类别泛化性, 已在大量图像检索工作中显示出巨大的潜力。当前, 基于属性的图像检索系统主要依赖于预定义属性集, 如独立属性^[4]、相关属性^[2, 11]和相对属性^[1]图像检索方面的研究。在这些图像检索方法中, 用户只能被动选择预定义属性集中的属性, 因而当属性发生变化或属性集过大时会给用户选择带来不便。文献[12]提出一种基于弱属性的大规模图像检索方法。尽管弱属性的维数远高于预定义查询属性, 使得弱属性空间

基金项目: 中央大学基础研究基金 (DUT14QY03, DUT14RC(3)103)。

作者简介: 张 芬 (1982—), 女, 博士研究生, 主研方向为图像检索; 孔祥维, 教授、博士、博士生导师; 宁 斐, 贾 则, 工程师、硕士。

收稿日期: 2016-06-28 **修回日期:** 2016-09-13 **E-mail:** dlut_zhangfen@mail.dlut.edu.cn

更具表达性,但仍然没有扩大实际的查询属性集范围。而扩展属性检索方法扩大了查询属性集范围,并让用户能够主动输入查询属性进行图像检索。文献[13]通过 WordNet^[17]学习扩展属性,提高了图像检索的匹配率。但由于该方法只选择 WordNet 作为唯一的外部语言知识库,且采用候选属性的一词多义性数量表示用户熟悉度,其结果有失偏颇。以学习到的扩展属性“shoes”为例,其一词多义性数量很小,用该数据表示的用户熟悉度很低,但实际上,该属性词却是大众颇为熟悉且常用的。

针对上述问题,本文提出一种基于网络检索量的语义关系度量方法用以学习扩展属性。首先利用 Wiktionary^[18]挖掘扩展属性,并与 WordNet 所得结果相结合;然后根据由百度指数^[19]和谷歌趋势^[20]获得的数据信息计算相对平均检索量,对用户偏好进行度量;最后采用一致性度量方法验证所学扩展属性的可靠性。

1 基于扩展属性的检索模型

基于扩展属性的图像检索使用户不仅可以不使用原始预定义属性,还可使用学习的扩展属性来检索与多属性查询最匹配的所有相关图像。如图 1 所示,采用预定义属性“jet engine”和所学的扩展属性“fly”来检索飞机图像。当用户输入查询属性“fly”时,因为“fly”为扩展属性,所以首先要通过扩展属性集与预定义属性集之间的语义相似度关系找到相应的预定义属性“wing”(图中虚箭头所示),再通过“wing”和“jet engine”来检索飞机图像。

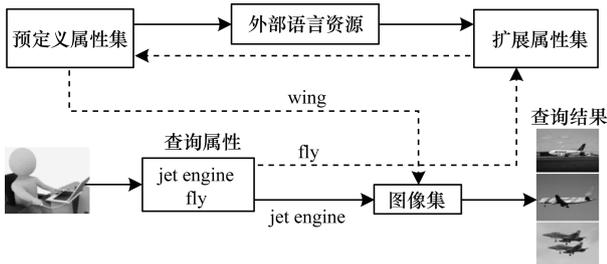


图 1 基于扩展属性的图像检索过程

事实上,预定义查询属性之间也或多或少存在一定的语义关系,而通过扩展属性挖掘所获得的语义信息正可弥补该缺失的连接关系。本文构建一个模拟预定义属性及其扩展属性间相关性,并基于结构化支持向量机(Support Vector Machine, SVM)^[11-13]的检索模型。

1.1 检索过程

给定 $X_q \in \mathbb{R}^m$ 为预定义查询属性集, $X_e \in \mathbb{R}^n$ 为扩展属性集,则完整的属性集 $X \in \mathbb{R}^{m+n}$ 可表示为 $X = \{X_q, X_e\}$ 。给定多属性查询 $Q, Q \subset X$ 以及图像集 Y , 目标为选择与 Q 最相关的图像子集 $y^* \subset Y$ 。因此,预测函数 $f_\omega: Q \rightarrow y$ 返回图像子集 y^* 作为给定的多属性查询 Q 的结构化响应,如式(1)所示。

$$y^* = \arg \max_{y \in Y} \omega^T \psi(Q, y) \quad (1)$$

$\omega^T \psi(Q, y)$ 定义为:

$$\omega^T \psi(Q, y) = \sum_{x_i \in Q} \sum_{x_j \in X} \omega_{ij} \varphi(x_j, y) \quad (2)$$

其中:

$$\varphi(x_j, y) = \begin{cases} \sum_{y_k \in y} \phi(x_j, y_k), & x_j \in X_q \\ \sigma_{jq} \sum_{y_k \in y} \phi(x_q, y_k), & x_j \in X_e \text{ and } x_q \Rightarrow x_j \end{cases} \quad (3)$$

在式(3)中, $\phi(x_j, y_k)$ 为特征向量,表明图像 y_k 具有属性 x_j 。将 $\phi(x_j, y_k)$ 设为独立训练的属性检测器的输出,其有效性在文献[11]中已作阐述。本文致力于在整个属性集 X 上学习相关模型 ω ,而非仅利用预定义查询属性集本身。对于预定义属性 $x_j \in X_q$,式(2)表示各查询属性 $x_i \in Q$ 的权重;对于扩展属性 $x_j \in X_e$,将输出定义为相应的预定义属性的加权分数,如式(3)第二部分所示。 $x_q \Rightarrow x_j$ 表示与预定义属性相一致的扩展属性 x_j , σ_{jq} 为衡量 x_q 与 x_j 之间语义关系的参数。

1.2 训练过程

给定多标签训练图像集 Y_{tr} 及其真值预定义查询属性集 X_q 和相应的扩展属性集 X_e ,本文致力于学习相关模型 ω ,使得对于各个查询 $Q \subset X, X = \{X_q, X_e\}$,检索得到的图像集 y^* 在所有 $y \in Y_{tr}$ 中具有最高分值。上述问题可通过标准的最大间隔训练来得到,由于存在训练误差,需要引入松弛变量 ξ ,并在松弛变量中加入线性惩罚因子 C 。直观地讲,违反存在高损失 $\Delta(y^*, y)$ 的边缘约束要比违反带有更小损失输出值的边缘约束受到更严厉的惩罚,一种特殊情况就是加入汉明(Hamming)损失^[21]。上述方法可表示如下:

$$\begin{aligned} & \arg \min_{\omega, \xi} \omega^T \omega + C \sum_t \xi_t \\ & \forall t \omega^T \psi(Q_t, y_t^*) - \omega^T \psi(Q_t, y_t) \geq \Delta(y_t^*, y_t) - \xi_t \end{aligned} \quad (4)$$

其中, $C > 0$ 为常量,是平衡训练误差最小化和间隔最大化的惩罚因子; ξ_t 为第 t 个训练样本的松弛变量; $\Delta(y_t^*, y_t)$ 为损失函数,本文采用 Hamming 损失,如式(5)所示。

$$\Delta(y_t^*, y_t) = 1 - \frac{|y_t \cap y_t^*| + |\bar{y}_t \cap \bar{y}_t^*|}{|Y_{tr}|} \quad (5)$$

式(4)所示的优化问题可用 Cutting Plane 算法^[21]来解决。初始时,该算法不包含任何约束条件,之后在每一次迭代中都加入当前优化问题所违反的最严重的约束条件。这一约束条件的形式为:

$$\xi_t \geq \max_{y_i \in Y_{tr}} [\Delta(y_t^*, y_i) - (\omega^T \psi(Q_t, y_t^*) - \omega^T \psi(Q_t, y_i))] \quad (6)$$

在 Hamming 损失函数下,时间复杂度为 $O(|Y_{tr}|)$ 。该模型通过采用预定义属性和扩展属性语义关系的联合分布信息,提高了泛化能力。

2 基于网络检索量的扩展属性学习

给定预定义属性集 $X_q \in \mathbb{R}^m$, 意在语义空间中学习最相关的扩展属性集 $X_e \in \mathbb{R}^n$ 。此过程可采用多种语言知识库, 如 WWW, Wikipedia, WordNet 等。本文选用具有显式结构的 Wiktionary 和 WordNet 作为扩展属性学习的外部言语来源。

尽管 Wiktionary 和 WordNet 具有高度组织化的结构, 但如何选取合适的关系类型来构建查询词的候选集仍是困难的问题。文献[13]利用 WordNet 中概念的同义词集 S_x , 同位词集 C_x , 解释词集 E_x 以及衍生词集 D_x 来构建候选集, 并用一词多义性数目表示用户熟悉度来进行扩展属性挖掘。除上述关系类型外, 本文还利用 Wiktionary 中的相关词集 R_x 构建候选集, 并将 Wiktionary 和 WordNet 所得结果相结合。另外, 为改善用户体验, 本文采用基于网络检索量的预定义属性及其相应扩展属性间的相对平均检索量来度量用户偏好。

给定预定义属性 $x \in X_q$ 及其候选集 $F_x = F_x^1 \cup F_x^2$, 其中 F_x^1 和 F_x^2 分别为由 Wiktionary 和 WordNet 所得的候选集。对于 Wiktionary, F_x^1 由五部分组成:

$$F_x^1 = \{S_x^1, C_x^1, E_x^1, D_x^1, R_x^1\} \quad (7)$$

本文意在寻找与 x 具有最近关系的子集 $W^1 \subset F_x^1$ 。

给定词 $\omega_i \in W^1$, 相应的释义集 $Sen_{\omega_i}^1$; 预定义属性 x 的释义集 Sen_x^1 , 则 ω_i 和 x 的相似性得分可计算如下:

$$Sim^1(x, \omega_i) = MS^1(x, \omega_i) + MS^1(\omega_i, x) \quad (8)$$

其中:

$$MS^1(x, \omega_i) = \frac{\sum_{s_m^1 \in Sen_x^1, s_n^1 \in Sen_{\omega_i}^1} \max S(s_m^1, s_n^1)}{|Sen_x^1| + |Sen_{\omega_i}^1|} \quad (9)$$

$$MS^1(\omega_i, x) = \frac{\sum_{s_n^1 \in Sen_{\omega_i}^1, s_m^1 \in Sen_x^1} \max S(s_n^1, s_m^1)}{|Sen_{\omega_i}^1| + |Sen_x^1|} \quad (10)$$

$|Sen_x^1|$ 和 $|Sen_{\omega_i}^1|$ 分别为 x 和 ω_i 的释义的数量; $S(s_m^1, s_n^1)$ 表示释义 s_m^1 和 s_n^1 之间的相似性:

$$S(s_m^1, s_n^1) = S(s_n^1, s_m^1) = \frac{1}{R_{s_m^1} R_{s_n^1}} \cdot \frac{Q(S^1) + Q(C^1) + Q(E^1) + Q(D^1) + Q(R^1)}{Q(F_{s_m^1}^1) + Q(F_{s_n^1}^1)} \quad (11)$$

其中:

$$Q(S^1) = \sum_{q_i \in S_{s_m^1}^1 \cap S_{s_n^1}^1} \lambda_{s_i} idf(q_i)^2 \quad (12)$$

$$Q(F_{s_m^1}^1) = \sum_{q_i \in F_{s_m^1}^1} \lambda_1 idf(q_i)^2 \quad (13)$$

$R_{s_m^1}$ 是对 x 的释义 s_m^1 的排名, $R_{s_n^1}$ 是对 ω_i 的释义 s_n^1 的排名。 $Q(S^1)$ 表示在同义词集上 s_m^1 和 s_n^1 的相似性度量, $q_i \in S_{s_m^1}^1 \cap S_{s_n^1}^1$ 表示同时存在于 s_m^1 和 s_n^1 的同义词集。 λ_s^1 为同义词集 S_x^1 的加权参数, $idf(q_i)$ 为 q_i

的逆向文档频率。同理可得 $Q(C^1), Q(E^1), Q(D^1)$ 和 $Q(R^1)$ 。 $Q(F_{s_m^1}^1)$ 表示 s_m^1 的整个候选集的总和, 其中:

$$F_{s_m^1}^1 = \{S_{s_m^1}^1, C_{s_m^1}^1, E_{s_m^1}^1, D_{s_m^1}^1, R_{s_m^1}^1\} \\ \lambda_1 \in \{\lambda_{S^1}, \lambda_{C^1}, \lambda_{E^1}, \lambda_{D^1}, \lambda_{R^1}\}$$

对于 WordNet, 只需考虑前 4 种词集, 同理可得相似性得分 $Sim^2(x, \omega_i)^{[13]}$ 。因此, x 和 ω_i 之间相似性得分总和为:

$$Sim(x, \omega_i) = Sim^1(x, \omega_i) + Sim^2(x, \omega_i) \quad (14)$$

若 ω_i 只存在于 Wiktionary, 则 $Sim^2(x, \omega_i) = 0$, 反之亦然。

事实上, 通过上述相似性度量方法获得的候选集中还存在许多不常用的词, 这在实际系统中可能导致较差的用户体验。简言之, 用户通常使用“head”和“brain”来检索他们所需的图片, 而并非不常用的词“caput”。因此, 在学习扩展属性时需考虑用户偏好。本文根据 2015 年 6 月—2016 年 5 月的百度指数^[19]和谷歌趋势^[20]的相关数据信息, 采用预定义属性及其相应扩展属性间的相对平均检索量来度量用户偏好。

对于百度指数, 定义 N_x 为预定义属性 x 的平均检索量, N_{ω_i} 为相应候选属性 ω_i 的平均检索量, 则相对平均检索量可表示为:

$$RAR_1(x, \omega_i) = \frac{N_{\omega_i}}{N_x} \quad (15)$$

对于谷歌趋势, 相对平均检索量可表示为:

$$RAR_2(x, \omega_i) = \frac{M_{\omega_i}}{M_x} \quad (16)$$

其中, M_x 和 M_{ω_i} 分别表示 x 和 ω_i 的平均检索量。

候选属性 ω_i 的最终得分为:

$$Score(\omega_i) = Sim(x, \omega_i) + \beta(RAR_1(x, \omega_i) + RAR_2(x, \omega_i)) \quad (17)$$

其中, β 为权衡语义相似度和用户偏好的控制参数。

3 实验与结果分析

3.1 实验设置

本文采用数据集 a-Pascal 和 a-Yahoo 进行实验^[4]。a-Pascal 数据集包含 12 695 张图像(6 340 张图像用于训练, 6 355 张图像用于测试), 这些图像都采自 PASCAL VOC 2008 挑战赛。a-Pascal 数据集中的图像被分为 20 类, 如飞机、自行车、沙发等。每张图像都由 64 个预定义属性标记, 如“jet engine”, “window”, “metal”等。a-Yahoo 数据集被视作 a-Pascal 的补充数据集, 其图像均采自雅虎图片搜索, 共包含 2 644 幅图像, 并被分为 12 类。a-Yahoo 中的图像与 a-Pascal 中一样, 都由同样的 64 个预定义属性标记, 只图像类别与 a-Pascal 的不同。

本文采用与文献[4]相同的设置, 即利用 a-Pascal

中预定义的训练图像进行训练,其他图像(包括 a-Pascal 中预定义的测试图像和 a-Yahoo 中的所有图像)则用于测试。

实验所用的每张图像都由一个 9 751 维特征向量表示,内容涉及色彩、纹理、视觉词和图像边缘。实验采用参考文献[22]中的结构化 SVM 的相关设置。在训练阶段,为每个预定义属性学习扩展属性。对于 Wikitionary,式(11)中语义相似度度量的权重值分别为: $\lambda_s^1 = 1.5, \lambda_c^1 = 1, \lambda_e^1 = 0.5, \lambda_d^1 = 1, \lambda_r^1 = 1$;对于 WordNet,权重值分别为: $\lambda_s^2 = 1.5, \lambda_c^2 = 1, \lambda_e^2 = 0.5, \lambda_d^2 = 1$ 。另外,通过重复实验及个人经验,式(17)中的 β 值设为 0.6 时获得最佳实验结果。本文采用平均 AUC 标准来评估实验性能。

3.2 扩展属性

在 Wikitionary 和 WordNet 语言库中,每个预标签属性都具有多个释义,但只有少数具有视觉表达含义。例如预标签属性“wood”,在 Wikitionary 和 WordNet 中都具有 8 个释义,而实际上其在 Wikitionary 中的释义应为“the substance making up the central part of the trunk and branches of a tree that used as a material for construction, to manufacture various items,etc. or as fuel”;在 WordNet 中的释义应为“the hard fibrous lignified substance under the bark of trees”,这两个释义的意思相同。因此,首先必须为每个预标签属性确定一个正确的释义,对全部 64 个预标签属性共可获得 324 个候选扩展属性。

一些候选扩展属性和预标签属性完全相同,例如,“glass”为学习的预定义属性“window”的候选扩展属性,但“glass”本身也是预定义属性。另外,一些预定义属性的候选扩展属性集中含有相同项,例如“leaf”,“flower”,“stem”这 3 个预定义属性的候选扩展属性中都有“plant”。因此,必须剔除这些重复项,最终可获得 279 个扩展属性。这 279 个扩展属性中,有 27 个扩展属性对应多于一个预定义属性,因此,扩展属性检索算法除了使用原属性对应的 64 个分类器以外,还需要学习 27 个新的分类器,共 91 个分类器。

表 1 随机显示 5 个预定义属性及其相应的扩展属性。

表 1 扩展属性

预定义属性	扩展属性
foot	feet, ankle, shoes, toe, heel, walk
wood	wooden, woody, timber, tree
label	tag, mark, sign, identification
handlebars	handlebar, handle, bicycle
window	windowed, glass, windowpane, windowglass

3.3 相对平均检索量

由于用户习惯使用常用词来检索他们想要的图片,因此在学习扩展属性时需考虑用户偏好。本文采用相对平均检索量来度量用户偏好,其数据信息来源于 2015 年 6 月—2016 年 5 月百度指数和谷歌趋势的统计数据。由于属性平均检索量的总体趋势基本一致,因此个别数据的浮动对统计特性的影响可忽略不计。

表 2 显示上述 5 个预定义属性及其相应的扩展属性之间的相对平均检索量的计算结果,括号内数据为百度指数和谷歌趋势的平均值。

表 2 相对平均检索量

预定义属性	扩展属性
foot	feet(0.385), ankle(0.098), shoes(0.183), toe(0.162), heel(0.119), walk(0.251)
wood	wooden(0.406), woody(0.227), timber(1.256), tree(1.023)
label	tag(1.161), mark(1.693), sign(1.525), identification(0.432)
handlebars	handlebar(0.506), handle(4.331), bicycle(2.726)
window	windowed(0.502), glass(0.781), windowpane(0.392), windowglass(0.682)

3.4 扩展属性的验证

本文提出一种一致性度量方法对学习的扩展属性可靠性进行验证。考虑:1)若 B 为 A 的扩展属性($B \Rightarrow A$), C 为 B 的扩展属性($C \Rightarrow B$),则 C 为 A 的扩展属性($C \Rightarrow A$);2)令 F_A 和 F_B 分别为 A 和 B 的扩展属性集,且 $B \in F_A$,若 $C \in (F_A \cap F_B)$,则可通过 C 的数量来确定 B 的可靠性。

给定预定义属性 x 及其相应的扩展属性 ω_i 。令 F_x 和 F_{ω_i} 分别表示 x 和 ω_i 的扩展属性集,且 $\omega_i \in F_x$,则 $\tilde{F}_x = \{x, F_x\}, \tilde{F}_{\omega_i} = \{\omega_i, F_{\omega_i}\}$ 。一致性度量系数可表示如下:

$$CM(x, \omega_i) = \frac{|\tilde{F}_x \cap \tilde{F}_{\omega_i}|}{|\tilde{F}_x|} \quad (18)$$

其中, $|\tilde{F}_x \cap \tilde{F}_{\omega_i}|$ 表示同时存在于 \tilde{F}_x 和 \tilde{F}_{ω_i} 的属性的数量。

表 3 显示了上述 5 个预定义属性的扩展属性验证结果,其中括号内数据为一致性度量系数。

表 3 扩展属性验证结果

预定义属性	扩展属性
foot	feet(1.000), ankle(0.714), shoes(0.571), toe(0.714), heel(0.714), walk(0.429)
wood	wooden(0.8), woody(0.8), timber(1.0), tree(0.4)
label	tag(1.0), mark(0.8), sign(0.6), identification(0.4)
handlebars	handlebar(1.00), handle(0.75), bicycle(0.50)
window	windowed(0.4), glass(0.8), windowpane(0.6), windowglass(0.8)

3.5 实验性能比较

为了进行实验性能比较,首先将64个属性扩展到91个代表性扩展属性。扩展属性与其相应的预定义属性符合一对一关系,则扩展属性与原属性的标注标签相同,因此,只需重新学习27个一对多的扩展属性对应的分类器。比较扩展属性中91个代表性扩展属性的检索结果与其他算法64个属性的检索结果即可。

图2显示了本文方法与现有方法在数据库 a-Pascal 上的检索性能比较。比较方法为 TagProp^[4], Weak Attributes^[12], Extended Attributes^[13] 和 Hypergraph-based Attribute Predictor(HAP)^[2]。前3种比较方法的结果来自文献[13],第4种方法是在第3种方法的基础上用 HAP 方法学习预定义属性,提高了属性学习精度,其性能较前3种方法有一定的提高。本文基于网络检索量对预定义属性及其相应的扩展属性间的语义关系进行度量,提高了语义相似度精度;同时由于引入另一外部语言库 Wiktionary,增加并完善了扩展属性集中的属性。实验结果表明,本文方法在全部查询类别下均优于其他4种方法,尤其对于双属性和三属性的查询,检索性能有明显改善。

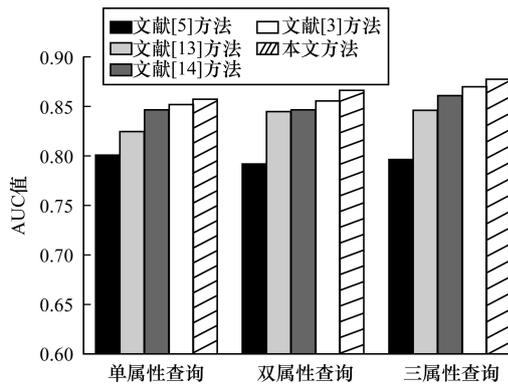


图2 在 a-Pascal 数据集上的检索性能比较

图3显示了本文方法与上述4种方法在数据库 a-Yahoo 上的检索性能比较。a-Yahoo 数据库只参与测试,在训练图像集中学习的属性间相互关系不能很好地适用于测试集中的图像,因此,上述所有方法在 a-Pascal 数据库上的检索性能都优于在 a-Yahoo 数据库上的实验结果。

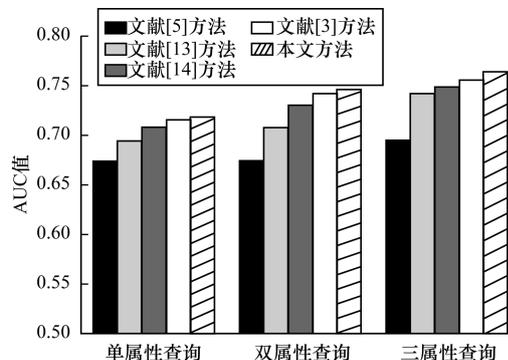


图3 在 a-Yahoo 数据集上的检索性能比较

从图3可看出,本文方法能够实现跨数据集图像检索,并仍具有最佳性能。

4 结束语

现有基于属性的图像检索方法主要依赖于预标签属性,使用户只能通过该预定义的属性来搜索预定目标。本文提出一种基于网络检索量的扩展属性学习方法。首先利用外部语言库 Wiktionary 挖掘扩展属性,并与 WordNet 所得扩展属性相结合,其结果较文献[13]只采用 WordNet 的方法所得结果更为全面;然后采用由百度指数和谷歌趋势获得的2015年6月—2016年5月的预定义属性及其相应扩展属性的相对平均检索量度量用户偏好。从网络所获得的数据信息直接反映了所有用户对某个属性词的使用量和熟悉度,因而更具实际意义;最后提出一种一致性度量方法用于验证所学扩展属性的可靠性。实验结果表明,本文方法可有效提高检索性能。如何更准确地表示属性及更好地改进属性学习方法,将是下一步的研究方向。

参考文献

- [1] LIANG L, GRAUMAN K. Beyond Comparing Image Pairs: Setwise Active Learning for Relative Attributes[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA; IEEE Press, 2014; 228-215.
- [2] HUANG Sheng, ELGAMMAL M, ELGAMMAL A, et al. Learning Hypergraph-regularized Attribute Predictors[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA; IEEE Press, 2015; 409-417.
- [3] SCHWARTZ G, NISHINO K. Automatically Discovering Local Visual Material Attributes[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA; IEEE Press, 2015; 3565-3573.
- [4] FARHADI A, ENDRES I, HOIEM D, et al. Describing Objects by Their Attribute[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA; IEEE Press, 2009; 1778-1785.
- [5] ELANGO VAN V, ALSAIDI B, SHIRKHODAIE A. A Multi-attribute Based Methodology for Vehicle Detection and Identification [J]. Proceedings of the SPIE, 2013, 8745.
- [6] FEDOROVSKAYA E A, LAWRENCE D R. Identifying Image Preferences Based on Demographic Attributes [J]. Proceedings of the SPIE, 2014, 9014.
- [7] KUMAR N, BERG A C, BELHUMEUR P N, et al. Attribute and Simile Classifiers for Face Verification[C]// Proceedings of the 12th IEEE International Conference on Computer Vision. Washington D. C., USA; IEEE Press, 2009; 365-372.
- [8] KUMAR N, BERG A C, BELHUMEUR P N, et al. Describable Visual Attributes for Face Verification and Image Search [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(10): 1962-1977.

(下转第287页)