

文章编号: 1000-5862(2013)03-0263-05

基于连续肯德尔相关系数 学习相似度函数的图像检索方法

黄 伟

(南昌大学信息工程学院 江西 南昌 330031)

摘要: 提出了一种基于连续肯德尔相关系数学习图像间相似度函数和运用学习的相似度函数进行图像检索的方法. 通过对 500 幅图像所组成的图像数据库以及和其他传统相似度函数学习方法在图像检索中检索效果的比较实验可以得出: 该方法的图像检索效果要优于其他相比较的传统方法.

关键词: 图像检索; 相似度学习; 连续肯德尔相关系数

中图分类号: TP 242.6⁺2

文献标志码: A

0 引言

随着数字图像在人们日常工作、生活各领域中无所不在的运用,如新闻报道、互联网、医学影像等,根据用户需要从图像数据库找出用户所感兴趣的图像(即图像检索)已成为近 20 年来计算机视觉领域研究的热门问题之一^[1]. 根据用户在检索过程中所给出的图像信息类别来划分,图像检索方法可以分为基于文字的图像检索方法和基于内容的图像检索方法. 基于文字的图像检索一般指通过比较用户所给出感兴趣图像(即目标图像)的文字信息和数据库中其他图像的文字信息,将具有和目标图像相同或相似文字信息的数据库图像作为检索结果返回给用户的方法;而基于内容的图像检索则通过比较目标图像与数据库图像之间的视觉特征,将具有和目标图像相同或相似视觉特征的数据库图像作为检索结果返回给用户的方法. 一般来说,基于内容的图像检索比基于文字的图像检索更受研究者和用户青睐,这是因为基于内容的图像检索方法不需要像基于文字的图像检索方法一样要求用户事先对图像加上额外的文字信息. 基于内容的图像检索既能节省人为的文字标记劳动,又能避免由于用户个人不同的主观判断所带来的文字标示误差,影响检索效

果. 因此,本文提出了一种基于连续肯德尔系数来进行相似度学习,并运用学习的相似度函数进行基于内容的图像检索方法,其检索效果得到了实验的验证.

1 相关技术

1.1 肯德尔相关系数

肯德尔相关系数主要基于排序理论. 排序是根据某种标准来反应一系列研究对象之间相互关系的序列形式. 以图像检索为例,数据库中每幅图像与用户指定的目标图像之间的相似度大小不同. 以相似度作为标准,将数据库中所有图像按照相似度从高到低排列并作为检索结果返回给用户就能构成一个图像检索结果的排序. 为了衡量不同图像检索结果的排序效果好坏,通常采用肯德尔相关系数(KT),它的定义为^[2]:

$$KT = N/N_n = (P - Q) / N_n = (P - Q) / (n(n-1)/2), \quad (1)$$

其中 P 和 Q 分别为一个图像检索结果的排序序列中一致排序对和非一致排序对的数量; N_n 为归一化因子,它代表了由 n 幅图像所组成的排序序列中所包含个数为 2 的组合个数(即 $C_n^2 = n(n-1)/2$). 一致排序对和非一致排序对的思想如图 1 所示. 假定 2 副

收稿日期: 2013-03-26

基金项目: 国家“863”计划(2013AA013804)和江西省科技计划(20123BBG70208 20123BBE50103)资助项目.

作者简介: 黄 伟(1983-),男,江西南昌人,讲师,博士,主要从事计算机视觉、模式识别和医学图像计算等方面的研究.

图像: X 与 Y . 图像 X 相对于目标图像的相似度比图像 Y 相对于目标图像的相似度更高(即 $X > Y$). 如果在一个相似度从高到低排序的序列中图像 X 比图像 Y 的位置更靠前(即满足相似度高的图像应作为检索结果更早返回给用户的图像检索宗旨), 那么图像 X 与 Y 即构成一个一致排序对; 反之, 它们称为一个非一致排序对. 一般来说, 方程(1) 所表示的归一化的肯德尔相关系数的值域为 $[-1, 1]$, 且肯德尔相关系数越高, 该图像序列按照相关度排序的程度也就越好(即更准确的图像检索结果). 在这里也可以注意到, 方程(1) 所示的肯德尔相关系数实际上是一个离散函数(因为 P 和 Q 为离散项). 因此, 基于方程(1) 的肯德尔相关系数、通过最优化求解方法来学习相似度函数就需要对方程(1) 进行适当的连续化和近似化.



图1 图像序列中一致/非一致排序对示意图

1.2 连续肯德尔相关系数

造成方程(1) 离散的原因在于离散项 P 和 Q , 因此需要对离散项 P 和 Q 进行改写. 在这里假设用户感兴趣的目标图像为图像 q , 图像检索序列结果中包含图像 x 与图像 y . 方程(1) 中的离散项 P 和 Q 可以被分别改写为

一致排序对(P):

$$\text{sgn}(s_{(q,x)} - s_{(q,y)}) \text{sgn}(l_{(q,y)} - l_{(q,x)}) = 1, \quad (2)$$

非一致排序对(Q):

$$\text{sgn}(s_{(q,x)} - s_{(q,y)}) \text{sgn}(l_{(q,y)} - l_{(q,x)}) = -1, \quad (3)$$

其中 sgn 代表一个符号函数: 当符号函数的自变量为非负时, 符号函数的函数值为 1; 反之, 符号函数的函数值为 -1. $s_{(q,x)}$ 表示目标图像 q 与图像检索序列中图像 x 的相似度函数. 在本文中, $s_{(q,x)}$ 通过定义在加权欧几里德范数空间的指数形式函数来表示: $s_{(q,x)} = \exp(-\langle \mathbf{a}, \|\mathbf{q} - \mathbf{x}\| \rangle)$, 其中 \langle, \rangle 代表两向量间的内积; \mathbf{x} 和 \mathbf{q} 分别表示从图像 x 和 q 中抽取的视觉特征信息向量; 向量 \mathbf{a} 表示该相似度函数中需要确定的未知数向量, 进行相似度函数学习的目的. 方程(2) 和方程(3) 中的 $l_{(q,x)}$ 为目标图像 q 与图像检索序列中的图像 x 之间的数值化标识差异. 由方程(2) 和方程(3) 可知, 它们能较好地代表方程(1) 中的离散项 P 和 Q . 假定图像检索序列中图像 x 相对

于目标图像 q 的相关度比图像 y 相对于目标图像 q 的相关度更高(即 $s_{(q,x)} > s_{(q,y)}$), 那么图像 x 相对于目标图像 q 的数值化标识差异应该比图像 y 相对于目标图像 q 的数值化标识差异要更小(即 $l_{(q,x)} < l_{(q,y)}$), 方程(2) 左式中 2 个符号函数的自变量都为非负, 即方程(2) 成立; 图像 x 与图像 y 构成了相对于目标图像 q 的一组一致排序对; 反之, 方程(3) 成立, 图像 x 与图像 y 构成了相对于目标图像 q 的一组非一致排序对. 将方程(2) 和方程(3) 代入方程(1) 中可得

$$KT = \frac{N}{N_n} = \frac{1}{N_n} \sum_{x, y \in D, x \neq y} (\text{sgn}(s_{(q,x)} - s_{(q,y)}) \cdot \text{sgn}(l_{(q,y)} - l_{(q,x)})), \quad (4)$$

其中 D 代表由目标图像 q 所产生的图像检索序列. 由于方程(4) 中的符号函数具有阶跃变化特性(在阶跃点符号函数不连续); 在这里可以采用双曲正切函数(\tanh) 来近似符号函数

$$\begin{aligned} \text{sgn}(\xi) &\approx \tanh(\xi) = \frac{\sinh(\xi)}{\cosh(\xi)} = \\ &= \frac{(e^\xi - e^{-\xi})/2}{(e^\xi + e^{-\xi})/2} = \frac{e^{2\xi} - 1}{e^{2\xi} + 1}. \end{aligned} \quad (5)$$

将方程(4) 和方程(5) 合并, 即可得到连续化肯德尔相关系数(CKT)

$$CKT = \frac{1}{N_n} \sum_{x, y \in D, x \neq y} \left(\frac{\exp(2(s_{(q,x)} - s_{(q,y)})) - 1}{\exp(2(s_{(q,x)} - s_{(q,y)})) + 1} \cdot \frac{\exp(2(l_{(q,y)} - l_{(q,x)})) - 1}{\exp(2(l_{(q,y)} - l_{(q,x)})) + 1} \right). \quad (6)$$

1.3 基于连续肯德尔相关系数学习相似度函数

基于连续肯德尔相关系数(方程(6)) 学习相似度函数的任务可以具体化为求解相似度函数 $s_{(q,x)}$ 中的未知变量 \mathbf{a} . 在这里, 采用最优化中常见的梯度下降法来求解^[3]. 对方程(6) 求解关于变量 \mathbf{a} 的梯度结果为

$$\begin{aligned} \nabla CKT(\mathbf{a}) &= \frac{1}{N_n} \cdot \\ &\sum_{x, y \in D, x \neq y} \frac{4 \left(\frac{\partial s_{(q,x)}}{\partial \mathbf{a}} - \frac{\partial s_{(q,y)}}{\partial \mathbf{a}} \right) \frac{\exp(2(l_{(q,y)} - l_{(q,x)})) - 1}{\exp(2(l_{(q,y)} - l_{(q,x)})) + 1}}{\exp(2(s_{(q,x)} - s_{(q,y)})) + \exp(2(s_{(q,y)} - s_{(q,x)})) + 2} \quad (7) \end{aligned}$$

基于连续肯德尔相关系数学习相似度函数 $s_{(q,x)}$ 中未知变量 \mathbf{a} 的主要步骤如图 2 所示.

因为在图像检索过程中每幅图像中被抽取的视觉特征信息是以向量的形式存在, 所以相似度函数 $s_{(q,x)}$ 中的未知变量 \mathbf{a} 实际上是与被抽取的视觉特征信息具有相同维度的未知向量. 求解 \mathbf{a} 的过程实际

上是对一个求解多元函数进行最优化的过程. 一般来说, 求解多元函数的全局最优解因具体实际研究问题的不同, 求解思路也会不同; 其中有些问题还会比较难以求解. 因此, 在实际操作中, 研究者通常转而求解一个足够好的局部最优解, 使得该局部最优

解能保证在此类实际问题中带来比较好的实验结果.

在本文的工作中, 求解未知变量 a 的主要步骤基于梯度上升法思想被提出; 求解结果实际上也是一个能给该图像检索问题带来满意检索结果的局

输入量	(i) 训练集图像: $(q, x \in D)$, (ii) 确认集图像: $(q_v, x_v \in D_v)$, (iii) 迭代次数 T , (iv) 步长 η .
训练步骤	T1 将相似度函数 $s_{(q, x)}$ 中的未知变量 a 初始化: a_0 T2 从 $t = 1$ 到 $t = T$ T3 令 $a = a_{t-1}$ T4 将训练集图像 $(q, x \in D)$ 代入到方程 (7) 中计算 $\nabla CKT(a)$ T5 运用梯度上升法更新 $a : a = a + \eta \cdot \nabla CKT(a)$ T6 令 $a_t = a$ T7 结束 T2 中的一次循环计算
确认步骤	V1 从 $j = 1$ 到 $j = T$ V2 将 a_j 所对应的相似度函数代入确认集图像: $(q_v, x_v \in D_v)$ V3 利用方程 (1) 计算此时检索序列的肯德尔相关系数 V4 结束 V1 中的一次循环计算
输出量	具有最大肯德尔相关系数的 a_{opt} 所对应的相似度函数

图 2 利用连续肯德尔系数学习相似度函数的主要步骤

部最优解. 在图 2 的步骤中, 输入量中的迭代次数 $T = 100$, 步长 $\eta = 0.01$. 训练集 D 和确认集 D_v 的组成遵循交叉验证 (cross-validation) 的原则, 即每一次实验中训练集和确认集彼此没有交集, 但它们的合集组成了图 2 中所用的所有图像. 在本工作的实验中, 500 幅图像被平均分为 5 份: 每一次实验时 4 份图像被用作训练集 D , 1 份图像被用作确认集 D_v . 图 2 的训练步骤所包含了 T 次循环. T1 步骤的初始化 a_0 采取归一化无偏执初始原则; 即 a_0 中所包含的每个元素都相等, 而所有元素初始化值等于 1. 通过 T4 和 T5 步骤的计算可以用递归方法得出 T 个不同的未知变量 a 结果. 在接下来在确认步骤中 (V1 ~ V4), 将该 T 个未知变量 a 的递归结果代入到确认集图像中并计算对应于每个不同 a 的肯德尔相关系数

(方程 (1)). 具有最大肯德尔相关系数的未知变量结果 a_{opt} 就可以作为本问题中最满意的局部最优解; a_{opt} 所对应的相似度函数 $s_{(q, x)} = \exp(-\langle a_{opt}, \|q - x\| \rangle)$ 即为通过连续肯德尔相关系数学习的图像检索中相似度函数.

2 实验验证与结果分析

在本文实验验证中, 500 幅拍摄了患者眼球的斜射灯图像作为需要进行检索的图像组成了一个医学图像数据库. 图 3 列举了该数据库中的一些图像. 斜射灯图像在医学诊断中的重要运用之一就是用来诊断患者白内障的病情. 白内障的病理具体表现在患者眼睛晶体内蛋白质凝结所形成的“雾状”; 白内

障是全世界造成患者失明的首要病因^[4].

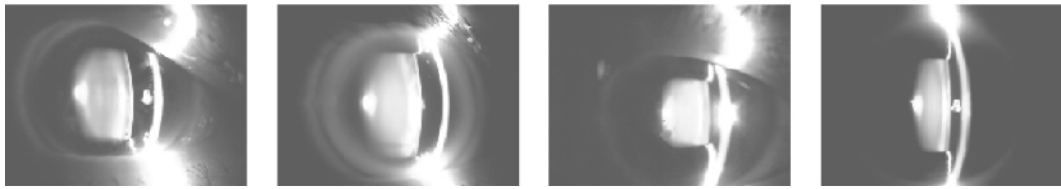


图3 本工作所使用的图像数据库中斜射灯图像实例

本工作实验验证首先需要从描述患者白内障病情的斜射灯图像中抽取图像的视觉特征信息. 在这里, 一个包含了晶体内部 HSI(即色泽/饱和度/灰度)颜色空间分量、晶体边际信息在内的 6 维视觉特征信息向量从每幅图像中被抽取出来. 因此, 本工作中所需要学习的相似度函数 $s_{(q, x)}$ 中的未知变量 a 是一个 6 维特征向量, 包括每幅图像中眼球区域内的平均灰度、灰度方差、灰度熵, 以及眼球区域内的另外 3 维 HSI 颜色信息. 在具体实验验证过程中, 一种 5 次交叉验证机制被采用. 由于该 500 幅图像包含 4 种不同程度的患者白内障病情(即可由 4 种不同的数值化标识来表示), 并且每种病情的患者数量均相等(即该数据库是个完美的平衡数据库, 每种数值化标识所对应的患者个数都为 125 个); 可以由不同病情患者的图像检索结果的平均值来得到最终该数据库的检索结果. 为了评价图像检索效果的好坏, 在图像检索工作中被广泛运用的 precision 和 recall 衡量标准在本工作中也被采用, 它们的定义为

$$\text{precision} =$$

$$\frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}, \quad (8)$$

$$\text{recall} =$$

$$\frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{relevant documents}\}|}. \quad (9)$$

为了突出本工作介绍方法的优点, 其他图像检索工作中常用的相似度函数学习方法, 例如大间距最近领域学习法(记为 LMNN)^[5]、支持向量机中的留一法(记为 SVM-LOO)^[6]、支持向量机中的半径边缘法(记为 SVM-RM)^[7]、支持向量机中的依据最大法(记为 SVM-EV)^[8]、半监督学习中的 Side-information 法(记为 Side + Semi)^[9] 也被一一复现并运用在本工作的图像数据库中进行图像检索实验验证. 图 4 表示了本工作方法与其他比较方法的 precision-recall 曲线^[10] 和 precision 箱状统计图. 由

图 4(a) 可知, 本工作方法的曲线比其他比较方法的曲线位置高, 这说明本工作方法检索图像的平均精度要高于其他方法. 在图 4(b) 中, 每个统计箱中水平线表示了不同方法检索精度的中位数、中位数上下的 2 条水平线(即统计箱的水平边缘)为不同方法精度的上四分位数和下四分位数, 从上下四分位数出发的垂直虚线一直延续到 1.5 倍的四分位数间距区间处. 由图 4(b) 可以看出, 本工作方法所产生的统计箱位置要高于其他比较方法所产生的统计箱位置. 因此, 本工作方法在由 500 幅图像所组成的图像数据库中的图像检索效果要好于其他比较方法在同一图像数据库中的图像检索效果.

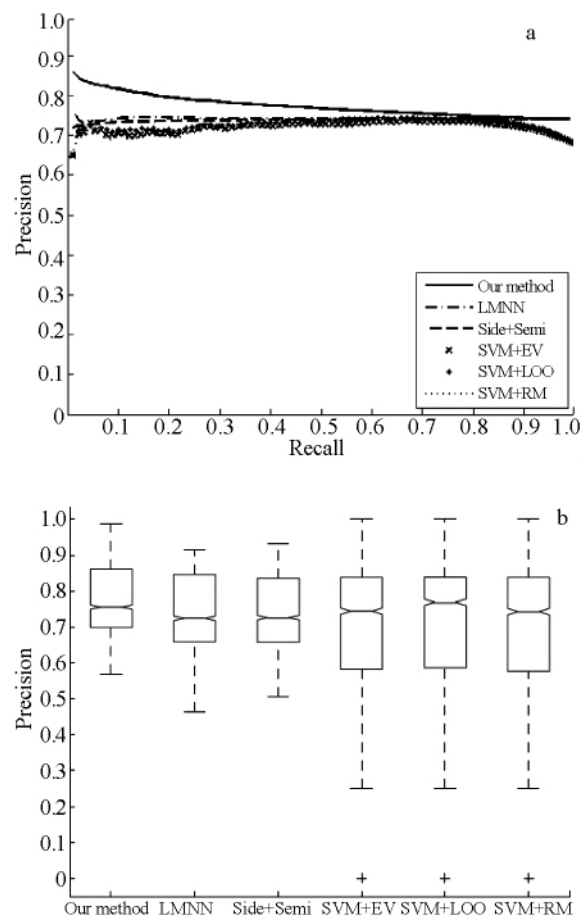


图4 本工作方法与比较方法的 precision-recall 曲线(上) 和 precision 箱状图(下)

3 结论

在本工作中,一种基于连续肯德尔相关系数进行相似度函数学习的图像检索方法被提出并研究.在500幅医学图像所组成的图像数据库上的图像检索实验表明,该方法比其他传统的具有相似度函数学习功能的图像检索方法具有更好的检索效果.从本工作的主要思路出发,未来基于排序思想进行图像检索的工作也将延续;其他排序衡量标准函数以及它们所对应的相似度函数学习方法也将得到研究.

4 参考文献

- [1] Smeulder A, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2000 22(12): 1349-1380.
- [2] Kendall M. A new measure of rank correlation [J]. Biometrika 1938 30(1/2): 81-93.
- [3] 傅湧. 用广义梯度刻画集值优化的强有效解 [J]. 江西师范大学学报: 自然科学版 2009 33(1): 47-51.
- [4] World Health Organization. The world health report: life in the 21st century: a vision for all [M]. Netherland: Geneva, 1998.
- [5] Weinberger K, Blitzer J, Saul L. Distance metric learning for large margin nearest neighborhood classification [C]. Vancouver: Neural Information Processing Systems 2005: 265-272.
- [6] Weston J. Leave-one-out support vector machines [C]. Stockholm: International Joint Conferences on Artificial Intelligence 1999: 727-733.
- [7] Chapelle O, Vapnik V, Bousquet O, et al. Choosing multiple parameters for support vector machines [J]. Machine Learning 2002 46(1/3): 131-159.
- [8] Gold C, Sollich P. Model selection for support vector machine classification [J]. Neurocomputing 2003 55(1/2): 221-249.
- [9] Xing E, Ng A, Jordan M, et al. Distance metric learning with application to clustering with side-information [C]. Vancouver: Neural Information Processing Systems 2002: 505-512.
- [10] 马永成, 肖诗斌, 林春雨, 等. 基于内容的文档图像检索的特征抽取研究 [J]. 江西师范大学学报: 自然科学版 2008 32(2): 138-141.

A Similarity Learning Method in Image Retrieval via Continuous Kendall-Tau Rank Correlation Coefficient

HUANG Wei

(School of Information Engineering, Nanchang University, Nanchang Jiangxi 330031, China)

Abstract: An image retrieval method with similarity learning via continuous kendall-tau rank correlation coefficient has been introduced. Experimental evaluation based on a database composed of 500 images reveals that the introduced method outperforms several other conventional similarity learning methods in this image retrieval application.

Key words: image retrieval; similarity learning; kendall-tau rank correlation coefficient

(责任编辑: 冉小晓)