

基于 Borda 计数的多特征融合图像检索技术研究

白明昌¹, 李文萃²

(1. 甘肃农业大学 信息科学技术学院, 甘肃 兰州 730070;

2. 西北师范大学 第二附属中学, 甘肃 兰州 730070)

摘要: 多个特征融合查询被认为是提高检索效率的有效方法. 在以往图像检索技术的基础上, 提出了新的加入多个分类器相关信息的 Borda 计数的特征融合图像检索方法. 实验证明, 该方法能够提高图像的检索效率, 方法有效.

关键词: 图像检索; 分类器; Borda 计数方法; 算法

中图分类号: TP 391.41

文献标识码: A

文章编号: 1001-988X(2008)01-0031-03

Image retrieval based on Borda count method of integrates correlative information

BAI Ming-chang¹, LI Wen-cui²

(1. College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730070, Gansu, China;

2. No. 2 Middle School Attached to Northwest Normal University, Lanzhou 730070, Gansu, China)

Abstract: The method of multi-features cooperation retrieval has been recognized as an effective method in improving image retrieval performance. The thesis presents a novel Borda count method that integrates correlative information between multi-classifiers. The experimental result shows that this approach can improved the effect of the image retrieval. And it is very efficient.

Key words: image retrieval; classifiers; Borda count method; algorithm

基于内容的图像检索, 是根据媒体和媒体对象的内容及上下文联系, 利用图像处理、模式识别、计算机视觉、图像理解等学科的相关理论, 从认知科学、用户模型、知识库系统、计算机图形学、数据库管理系统, 以及信息检索等领域中获得启发, 引入新的媒体数据表示和数据模型, 产生出有效、可靠的查询处理算法和可视化查询接口, 在大型数据库中进行检索. 笔者在以往图像检索技术的基础上, 提出了新的加入多个分类器相关信息的 Borda 计数的特征融合图像检索方法. 该方法考虑到各个图像检索中分类器之间强烈的相关性, 对各个分类器检索出来的图像按照一定的策略进行奖惩, 重新对图像进行排序. 该算法能够保持分类能力强的分类器结果, 同时能够抑制分类能力弱的分类器结

果, 实现了特征的有效融合.

1 Borda 计数原理

Borda 计数方法是基于投票选取策略, 最早用于推测在竞选中可能获胜的候选人, 曾经成功地应用于手写数字识别^[1]. 其原理描述如下:

$$\begin{aligned} e^1(x) &= C_0^1, \dots, C_{M-1}^1, \\ e^2(x) &= C_0^2, \dots, C_{M-1}^2, \\ &\dots \\ e^k(x) &= C_0^k, \dots, C_{M-1}^k, \end{aligned} \quad (1)$$

其中, $e^1(x), \dots, e^k(x)$ 是 k 个不同的分类器; x 是输入特征; $C_0^k, \dots, C_{M-1}^k, 1 \leq k \leq K$ 是第 k 个分类器对 M 个物体进行排序的结果.

收稿日期: 2007-11-12; 修改稿收到日期: 2007-12-17

作者简介: 白明昌 (1975—), 男, 甘肃靖远人, 讲师, 硕士. 主要研究方向为计算机图形学.

E-mail: baimc@gsau.edu.cn

$$B_i = \sum_{k=1}^K w^k \cdot G^k(i), \quad (2)$$

$$G^k(i) = M - r, \quad (3)$$

其中, B_i 为物体 $i(1 \leq i \leq M)$ 的 Borda 数; w^k 为权重, 对于不同的分类器 $e^k(x)$, 权重 w^k 可以为不同的值; r 为物体 i 在分类器 $e^k(x)$ 中位于排序 $C_0^k, \dots, C_{M-1}^k, 1 \leq k \leq K$ 中的位置.

$$\text{if } B_i = \max_{1 \leq i \leq M} \{B_i\} \implies \text{Borda}(x) = j.$$

综上, Borda 计数是通过对于 $i(1 \leq i \leq M)$ 个物体中最大的 Borda 数来确定输入模式的最接近的类别.

2 适用于图像检索的 Borda 计数方法

Borda 计数虽然能很好地运用于手写字符识别, 但不能直接运用于图像检索, 所以需对 Borda 计数方法进行必要的修改^[21].

1) Borda 计数方法用于模式识别, 输出结果是与输入最接近的模式. 在图像检索中输出的是一系列和示例图像相似的图像而不是单一类别, 故应修改输出值为一系列的 Borda 值, 并且对输出的系列值进行排序.

$$\text{Borda}(x) = C_0^{\text{Borda}}, C_1^{\text{Borda}}, \dots, C_{M-1}^{\text{Borda}}, \quad (4)$$

$$\text{Borda}(x) = \text{sort} \{C_0^{\text{Borda}}, C_1^{\text{Borda}}, \dots, C_{M-1}^{\text{Borda}}\}, \quad (5)$$

其中, C_i^{Borda} 为图像 i 的 Borda 数.

2) Borda 计数方法的基本要求是各个分类器对于相同的一组图像进行排序, 而通过形状、颜色和纹理等分类器得到的图像不全相同. 故图像检索中 M 的值应为各分类器输出图像集合并集的势.

$$M = |\{N(e_1(x)) \cup N(e_2(x)) \cup \dots \cup N(e_k(x))\}|,$$

其中 $N(e_k(x))$ 为第 i 个分类器输出图像的集合.

3) Borda 计数方法采用代数和的方法得到 Borda 数, 这种方法容易受到分类器中最坏分类器的影响. 可采用几何均值代替算术均值:

$$C_i^{\text{Borda}} = \prod_{k=1}^K w^k \cdot G^k(i). \quad (6)$$

这样改进过后适应于图像检索的 Borda 计数表示为

$$\begin{aligned} e^1(x) &= C_0^1, \dots, C_{A-1}^1, \\ e^2(x) &= C_0^2, \dots, C_{B-1}^2, \\ &\dots \end{aligned} \quad (7)$$

$$\begin{aligned} e^k(x) &= C_0^k, \dots, C_{k-1}^k; \\ M &= |\{N(e_1(x)) \cup N(e_2(x)) \cup \dots \cup N(e_k(x))\}|; \end{aligned} \quad (8)$$

$$C_i^{\text{Borda}} = \prod_{k=1}^K w^k \cdot G^k(i); \quad (9)$$

$$\text{Borda}(x) = C_0^{\text{Borda}}, C_1^{\text{Borda}}, \dots, C_{M-1}^{\text{Borda}}. \quad (10)$$

3 增加各类分类器相关信息的 Borda 计数

在 Borda 计数方法中待融合的数据属性是已知的、完备的. 图像检索中分别用颜色、纹理和形状分类器对图像进行检索, 发现各个分类器检索出来图像差异很大. 图像在各个分类器中重新排序过程中, 会出现这样的情况: 在 i 分类器中排在前面的图像, 在 $j(i \neq j)$ 分类器的排序号很低, 造成在融合过程中丢掉该图像, 其中没有考虑到各个分类器的相关信息, 因此必须在重新排序过程中引入距离变化和惩罚因子增加各个分类器的相关信息.

设置一个惩罚因子 C 对出现在各个分类器中的图像进行奖励, 对于不相关的进行处罚.

$$C_i^{\text{Borda}} = \prod_{k=1}^K (w^k \cdot G^k(i)) + f \cdot C, \quad (11)$$

其中 C 为惩罚因子; f 为图像 m 在多个分类器中出现的频率.

若每次查询 n 个相似图像, 假设在 i 分类器中前 $[n/k]$ 图像和示例图像在语义上较后面的 $n - [n/k]$ 更相关, 在 $j(i \neq j)$ 分类器中重排序过程中须对 i 分类器中出现在 $[n/k]$ 之前的图像 m 进行奖励, 其在 $j(i \neq j)$ 分类器中的序号重新排序, 序号向前移动, 保证不会因为最坏分类器造成检索时丢失该图像. 其在 $j \neq i$ 分类器中的排序号调整为

$$\begin{aligned} \text{order}_m^j &= C_m^i + C_m^j - [n/k], \\ \text{if } C_m^j &< [n/k], \end{aligned} \quad (12)$$

其中 C_m^i 为图像 m 在 i 分类器中的排序值; C_m^j 为图像 m 在 $j(i \neq j)$ 分类器中的排序值.

从中可以看到图像 m 在 i 分类器排序号越靠前, 那么在 $j(i \neq j)$ 分类器中向前移动的位置越多.

4 基于 Borda 计数的多特征融合图像检索

对于图像检索用增加相关信息的 Borda 计数方法分别对颜色、纹理和形状特征检索得到的图像做结果融合, 算法如下:

1) 按照基于“特征+距离”的方法对用户提供的例子图像. 分别用颜色、纹理和形状分类器对图像检索前 M 个图像得到 3 类检索结果, 并对结果进行排序.

$$\text{Color}(q) = C_0^1, \dots, C_{M-1}^1, \quad (13)$$

$$\text{Texture}(q) = C_0^2, \dots, C_{M-1}^2, \quad (14)$$

$$\text{Shape}(q) = C_0^3, \dots, C_{M-1}^3. \quad (15)$$

2) 求出 3 个分类器得到的图像集合的并集

M_{set} :

$$M_{\text{set}} = \{\text{Color}(q) \cup \text{Texture}(q) \cup \text{Shape}(q)\}. \quad (16)$$

3) 各个分类器对图像集合 M_{set} 重新排序:

$$\text{Color}(q) = C_0^1, \dots, C_{|M_{\text{set}}|-1}^1, \quad (17)$$

$$\text{Texture}(q) = C_0^1, \dots, C_{|M_{\text{set}}|-1}^1, \quad (18)$$

$$\text{Shape}(q) = C_0^1, \dots, C_{|M_{\text{set}}|-1}^1. \quad (19)$$

4) 对各个分类器对集合 M_{set} 排序, 根据(11)式和(12)式分别对各个分类器的图像变更序号和奖励.

5) 设置各个分类器的权重为 w^k , 对集合 M_{set} 中图像计算 Borda 数:

$$C_i^{\text{Broda}} = \prod_{k=1}^3 (w^k \cdot G^k(i)) + f \cdot C, \quad (20)$$

$$1 \leq i \leq |M_{\text{set}}|.$$

6) 对集合 M_{set} 的计算结果进行排序, 并返回前 M 个图像给用户.

为了讨论方便, 记上述算法为 A^+ , 记不考虑分类器相关信息的 Borda 计数算法为 A .

5 实验结果及分析

5.1 实验环境

为验证文中提出的算法的有效性, 实现了一个网络环境下基于的内容图像检索系统^[3]. 该系统用 Visual C++6.0 编程, 用 SQL Server 2000 建立图像特征库和图像库, 在 Windows 2000 Professional 操作系统下实现.

5.2 测试图像库

测试图像库是从大约有 1 000 000 幅图的图像

库中, 选出 13 个类别 2 500 幅图像: bear (285), elephant (165), plane (216), horse (231), butterfly (110), tower (241), rose (110), sunset (317), mushroom(200), eagle (133), fruit (238), rainbow (48), train(277).

5.3 评价方法

在传统的信息检索中, 对试验结果的评价常常采用查全率 r (recall)和查准率 p (precision)2 个参数:

$$r = a/(a+c), \quad p = a/(a+b),$$

其中 a 为计算机自动检索输出的图像中从主观上满足条件的图像数目; b 为计算机自动检索输出的图像中从主观上不满足条件的图像数目; c 为计算机未检索输出的图像中从主观上满足条件的图像数目.

5.4 实验结果及分析

表 1 是对 eagle, rose, sunset, butterfly, mushroom 5 类图像利用未考虑分类器相关信息的 (A 算法)和考虑了相关信息的 Borda 计数算法(A^+ 算法)选择 $k=2, C=40$ 情况下查准率、平均查准率的比较.

从表 1 可以看出, 当有分类器分类能力下降的时候, 对 A 算法影响较大. 考虑到各分类器相关信息后的 A^+ 算法, 分类能力较好的分类器性能得到保持, 较差的分类器结果得到抑制, 从而有效地对各个分类器得到的图像进行融合.

表 2 是对 5 类图像查全率和平均查全率的比较结果, 实验中每次查询图像数目为该类图像在图像库中数目的 1.5 倍. 从中可以看到查全率有明显的提高.

表 1 A 算法和 A^+ 算法对 5 类图像的查准率和平均查准率

	颜色	纹理	形状	平均距离	A	A^+
eagle	0.575	0.350	0.400	0.625	0.650	0.675
rose	0.525	0.125	0.475	0.500	0.475	0.550
sunset	0.625	0.200	0.550	0.650	0.600	0.750
butterfly	0.225	0.075	0.050	0.250	0.275	0.375
mushroom	0.525	0.175	0.000	0.475	0.450	0.525
平均查准率	0.495	0.185	0.295	0.500	0.490	0.575

表 2 A 算法和 A^+ 算法对 5 类图像的查全率和平均查全率

算法	平均查全率	eagle	rose	sunset	butterfly	train
A	0.326	0.564	0.373	0.360	0.145	0.191
A^+	0.348	0.571	0.418	0.388	0.163	0.202

(下转第 51 页)

加接近 1.5%。可见, 算法对训练数据的规模不敏感, 避免了统计模型易受数据稀疏影响的问题。

表 6 显示了本文中提出的算法与文献[7]中基于 TBL 算法及文献[8]中基于最大熵模型算法的比较结果。从表中可以看出, 本文中的算法预测结果高出文献[7]的结果近 15%~20%, 也好于文献[8]的结果, 说明本文中的方法优于其他方法。

表 6 各学习算法预测结果的对比

Table 6 comparison result for 3 predicting approaches of prosodic words

	Prosodic work/%		
	precision	recall	F-score
TBL	78.2	84.8	81.4
ME model	93.6	96.3	94.9
our approach	97.5	97.5	97.5

4 结论

在 TTS 系统中, 为了能够合成出高度自然和清晰度的合成语音, 需要预测出语句的韵律结构。本文在 1 万句人工标注过的语料的基础上, 分析了语法词与韵律词之间的关系, 发现语法词的词长、语法词是否为韵律词、语法词是否和前后的语法词组合以及语法词的组合方式这 4 个特征对韵律词边界的预测具有重要作用。基于以上分析, 实现了一种基于 TBL 算法的韵律词预测方法, 充分考虑了以上 4 个特征。实验结果表明, 本文提出的方法在测试集上能够达到 97.5% 的预测精度, 优于其它几种韵律词边界预测方法。

参考文献:

[1] 曹剑芬. 普通话节奏的声学语音学特性[C]//吕士

楠. 现代语音学论文集. 1999: 155-159.

[2] GEE J P, GROSJEAN F. Performance structures: A psycholinguistic and linguistic appraisal [J]. *Cognitive Psychology*, 1983, **15**: 411-458.

[3] 曹建芬. 基于语法信息的汉语韵律结构预测[J]. *中文信息学报*, 2003, **17**(3): 41-46.

[4] 应 宏, 蔡莲红. 基于结构助词驱动的韵律短语界定的研究[J]. *中文信息处理学报*, 1999, **13**(6): 8-13.

[5] WANG M, HIRSCHBERG J. Predicting intonational boundaries automatically from text: the ATIS domain[C]//*Proceedings of the Workshop on Speech and Natural Language*. California, 1991: 378-383.

[6] 赵 晟, 陶建华, 蔡莲红. 基于规则学习的韵律结构预测[J]. *中文信息学报*, 2002, **16**(5): 30-37.

[7] 李剑锋, 胡国平, 王仁华. 基于最大熵模型的韵律短语边界预测[J]. *中文信息学报*, 2004, **18**(5): 56-63.

[8] ZHANG Xiao-nan, XU Jun, CAI Lian-hong. Prosodic structure prediction based on maximum entropy model with Error-driven modification[C]//*Proceedings of Internation Symposium of Chinese Spoken Language Processing*. Singapore. 2006: 149-160.

[9] BRILL ERIC. Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging[J]. *Computational Linguistics*, 1995, **21**(4): 543-565.

[10] C J van RIJSBERGEN. *Information Retrieval*[M]. London: Butterworths, 1979.

(责任编辑 孙晓玲)

(上接第 33 页)

从表 2 可以看出加入分类器相关信息的 A⁺ 算法比 A 算法更具有鲁棒性, 能有效的减少差分类器性能的影响。

参考文献:

[1] 季 艳, 高大启. 基于 Borda 规则的分类器组合方

法及其在手写字符识别中的应用[J]. *计算机工程*, 2005, **5**(33): 92-94.

[2] 苏为华, 陈 骥. 模糊 Borda 法的缺陷分析及其改进思路[J]. *统计研究*, 2007, **24**(7): 58-64.

[3] 宋麦玲, 李 欢, 杨 捷. 基于内容的图像检索系统[J]. *自动化博览*, 2005(4): 91-95.

(责任编辑 惠松骐)