

基于贝叶斯分类器的图像检索相关反馈算法*

苏中^{1,2}, 张宏江³, 马少平^{1,2}

¹(清华大学 计算机科学与技术系,北京 100084);

²(清华大学 智能技术与系统国家重点实验室,北京 100084);

³(微软亚洲研究院,北京 100080)

E-mail: suzhong_bj@hotmail.com

http://www.tsinghua.edu.cn

摘要: 由于图像底层特征及其本身所包含的上层语义信息的巨大差距,使得基于内容的图像检索很难取得令人满意的效果.作为一种有效的解决方案,在过去的几年中,相关反馈在该研究领域取得了一定的成功.提出了一种新的具有学习能力的反馈算法.该算法基于贝叶斯分类原理,运用不同的反馈策略分别处理正、负反馈,同时它具有学习能力,可以运用用户的反馈信息不断地修正检索参数,使系统的检索能力得到不断的提高.通过在大图片库上的检索实验,该算法产生的效果大大优于当前其他的反馈方法.

关键词: 基于内容的图像检索;相关反馈;贝叶斯分类器

中图法分类号: TP311 文献标识码: A

伴随着多媒体数据的大规模运用,对基于内容的多媒体数据检索的研究已成为当前多媒体领域的热点之一,特别是基于内容的图像检索(content based image retrieval),已经进入实用阶段.CBIR 主要是利用提取图像中的颜色、纹理或者形状特征,并将这些特征用于检索的一种信息检索手段.很多人在该领域中取得了出色的研究成果,但是由于图像检索的复杂性,特别是由于在图像底层特征和高层语义信息之间存在很大的差别,我们还不能利用已有的计算机视觉和人工智能的技术获得二者之间满意的映射关系,因此只能通过特征提取进行检索,在效果上往往不能满足用户需求.为了解决这个问题,最近人们试图运用相关反馈(relevance feedback)的技术来帮助 CBIR,以获得更好的检索效果.

相关反馈在信息检索中是一种指导性学习的技术,用以提高系统的检索能力.在基于内容的图像检索中,相关反馈的过程是这样的:对于用户提供的样本图片,系统根据相似矩阵和样本图片及图像数据库中图片的图像特征,给出一个基于相似度的排序列表,数据库中图片和样本图片如果在特征空间中相似程度越高,则排序越靠前.然后,用户在该列表中选择一组符合该次检索内容的正反馈图片和一组不符合检索内容的负反馈图片,并提交给检索系统.系统根据用户的提交信息来优化相似矩阵并提交给用户新的检索结果.因此,在相关反馈中最重要的问题就是如何有效地利用用户提供的交互信息,并通过优化检索向量和相似矩阵来提高检索的准确度.

本文提出了一种基于贝叶斯分类的相关反馈的算法.对于每个检索样本图片,假定该图片和与之对应的所有正反馈图片构成了一个拥有共同语义含义的图像集合,则通过对于该图片集的高斯拟合构造出该检索样本图片在特征空间的贝叶斯分类器.该分类器参数反映了语义类在特征空间中的分布信息.同时,利用用户的正反馈信息不断地更新该分类器的参数,使检索效果不断得到提高,方便以后用户的检索.对于负反馈信息,通过对

* 收稿日期: 2000-12-21; 修改日期: 2001-04-03

基金项目: 国家自然科学基金资助项目(69823001);国家重点基础研究发展规划 973 资助项目(G1998030509)

作者简介: 苏中(1976 -),男,上海人,博士生,主要研究领域为基于内容图像检索,模式识别,网络数据挖掘;张宏江(1960 -),男,黑龙江哈尔滨人,博士,主要研究领域为视频和图像内容分析与检索,计算机视觉,信息系统;马少平(1961 -),男,河北唐山人,博士,教授,博士生导师,主要研究领域为模式识别,信息检索,网络数据挖掘.

大量试验的观察,我们有理由假定负反馈样本之间是彼此独立的,这样,通过构造一个简单的混合高斯模型来拟合出负反馈样本在特征空间的分布,并利用一种类似于“挖洞”的过程将负反馈信息用于检索过程.通过对大图像库的检索实验,结果证明,我们的方法比其他算法可取得更好的效果.

本文第 1 节介绍一些研究背景.第 2 节介绍贝叶斯分类.第 3、4 节分别介绍正负反馈算法.第 5 节介绍试验结果.第 6 节是结论.

1 相关工作

作为一种有效的检索手段,相关反馈在文本检索中取得了巨大的成功^[1,2],因此,由于 CBIR 检索效果的局限性,人们开始试图将相关反馈用于图像甚至视频检索^[3-5],并取得了不错的效果.我们也在这方面的研究中取得了许多成果^[6,7].按照反馈的方法,我们可以将诸多方法大致分为两种:检索向量转移^[3,5]和权值的调整^[4,8].

转移检索向量的思想是希望利用用户的反馈信息,改变用户的检索向量,将它向最理想的方向转移,通过这种方式来优化检索结果.对于用户提交的正反馈集合 D'_R 和负反馈集合 D'_N ,一种方法是利用 Rocchio 公式(式 (1))改变用户的检索向量.

$$Q' = \alpha Q + \beta \left(\frac{1}{N_{R'}} \sum_{i \in D'_R} D_i \right) - \gamma \left(\frac{1}{N_{N'}} \sum_{i \in D'_N} D_i \right), \quad (1)$$

其中 α, β 和 γ 分别为常量, $N_{R'}$ 和 $N_{N'}$ 分别为 D'_R 和 D'_N 中的反馈图像个数.这种技术在 Mars 系统^[4]中得到了实现,试验结果表明,通过用户的反馈,系统检索精度得到一定的提高.另外一种检索向量转移的方法是贝叶斯推理. Cox^[3]和 Nuno^[5]结合用户的反馈过程,假定对于给定的样本图,正、负反馈图像是一系列彼此独立的检索单元,利用贝叶斯规则来优化检索结果,试图将检索错误降到最低.

权值调整的方法试图通过调整各个特征、特征各维的权重来优化检索结果. Rui 和 Huang 提出了一种利用特征标准差的算法^[4,8],其基本思路是对在特征空间中对特征分布标准差较大的特征维降低权重,反之则增加权重以提高检索的效果.同时他们提出了一种多层的图像模型,并以此模型为基础,使用全局优化的方法来处理权值调整的问题.实验证明,这种方法大大提高了系统的检索能力.

2 贝叶斯分类器(Bayesian classifier)^[9]

高斯分布模型是一种通用的概率分布模型.这种模型运算简单,而且现实世界的很多事件都和高斯分布有着极大的相似性.假定 n 维空间 R^n 中的向量 x 满足高斯分布,那么 x 的概率密度函数就是

$$p(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\varepsilon)^T \Sigma^{-1}(x-\varepsilon)}, \quad (2)$$

其中 $x = [x_1, \dots, x_n]$, $\varepsilon = [\varepsilon(x_1), \dots, \varepsilon(x_n)]$, $\Sigma = E\{(x-\varepsilon)(x-\varepsilon)^T\}$. 基于这种分布的假设并建立贝叶斯分类器,我们可以得到 x 属于类 w_i 的概率判别函数为

$$\begin{aligned} g_i(x) &= \ln P_i(x) = \ln p(x | w_i) + \ln P(w_i) \\ &= -\frac{1}{2}(x-\varepsilon_i)^T \Sigma_i^{-1}(x-\varepsilon_i) - \frac{d}{2} \ln 2\pi - \frac{1}{2} \ln |\Sigma_i| + \ln P(w_i) \\ &= -\frac{1}{2}(x-\varepsilon_i)^T \Sigma_i^{-1}(x-\varepsilon_i) + \ln P(w_i) + c_i. \end{aligned} \quad (3)$$

3 正反馈

我们使用贝叶斯分类器来解决正反馈的问题.在检索中,每一幅图像都属于某个特定的类,除非手工标定,这个类我们事先并不能确定.正如第 5 节中表述的那样,我们的系统是一个样例检索系统,即用户通过提交给系统样本图像来检索系统数据库中的图像.在检索中,用户提供的样本图在语义上可以反映用户的检索值,例如用户提交一张狗的相片,想让系统找到其他关于狗的图片.可是,在底层特征空间中,它往往不能描述用户的检索

信息,这正是由于图像检索中高层的语义信息和底层特征的分布之间存在严重的差异所造成的.我们的思路是利用用户给系统提供的正反馈图片的分布特征,构造出一个贝叶斯分类器,以此来决定数据库中所有图像属于该类的概率或一种距离度量.同时,分类器的参数可以不断地由新加入的正反馈信息进行更新.

我们假定在图像检索中,图像的特征满足高斯分布.根据对实际图像库进行大量的统计分析可以发现,这种假定是有根据的.下面给定一些数据定义.每一张图片 P_k 可以表示为特征空间中的一个特征向量组 $\vec{x}_k = [\vec{x}_{k1}, \dots, \vec{x}_{kn_k}]$,其中向量的每一维分别代表某种特征,例如颜色直方图.当然它们仍然是向量,我们将其表示为 $\vec{x}_{ki} = [x_{ki1}, \dots, x_{kin_i}]$.对于每个特征向量 \vec{x}_{ki} 分别有一个 $n_i \times n_i$ 维的协方差矩阵 Σ_{ki} 和一个 n_i 维的均值向量 ε_{ki} 来存储它的检索分类器参数. n_k 是对图片 P_k 作检索时正反馈图片的个数.由于在现实的检索中,很难获得足够的反馈信息,一方面由于用户在检索中并不希望而且不喜欢多次的交互过程,另一方面也是因为图像库中同类别的图像数目与图像特征维数的不可比性,这样,很难估计准确的 $n_i \times n_i$ 的协方差矩阵 Σ_{ki} ,因此将其简化为一个对角阵 $diag\{\sigma_{ki}^2\}$,其中 $\sigma_{ki}(m) = \Sigma_{ki}(m, m)$.

从式(2)中,我们清楚地看到类 i 的分类器拥有 3 个参数: ε_i, Σ_i 和 $P(w_i)$.如果保存这些参数并根据用户的交互过程不断地更新这些参数,就可能不断地提高系统的检索精度,试验结果证实了这个假设.假定对于检索样本图 P_k ,用户新提交的正反馈图片集合为 C_p ,其中 $C_p = \{P_{p1}, \dots, P_{pq}\}$,根据参数的定义,将新的反馈集合类和已有的 P_k 所对应的类进行合并,这样很容易得到以下的参数更新方程:

$$(n_k + q)\sigma_{ki}^2 = n_k\sigma_{ki}^2 + \frac{n_k q \varepsilon_{ki}^2 - 2n_k \varepsilon_{ki} \sum P_{pi}}{n_k + q} + \sum P_{pi}^2 - \frac{(\sum P_{pi})^2}{n_k + q}, \tag{4}$$

$$\varepsilon_{ki} = \frac{n_k \times \varepsilon_{ki} + \text{sum}(C_p)}{n_k + q}, \tag{5}$$

$$n_k = n_k + q. \tag{6}$$

下面是我们处理正反馈的算法流程:

(1) 特征的规格化:规格化的目的是将所有特征规格在同样范围中取值.对于向量 \vec{x}_{ki} ,其规格化后的值 $\vec{x}'_{ki} = [x'_{ki1}, \dots, x'_{kin_i}]$,其中

$$x'_{ki_m} = \frac{x''_{ki_m} - \varepsilon(x''_{ki_m})}{3\sigma(x''_{ki_m})},$$

$$x''_{ki_m} = \frac{x_{ki_m} - \min(X_{ki_m})}{\max(X_{ki_m}) - \min(X_{ki_m})}.$$

(2) 初始化:初始化 σ_{ki} 为单位对角阵, $\varepsilon_{ki} = \vec{x}_{ki}, n_k = 1$.

(3) 提交反馈并根据式(4)~式(6)分别更新参数.

(4) 距离的计算:对于所有数据库中的图片 P_i ,我们运用式(2)计算该图片属于样本检索图片所在语义类的概率,当然这种概率彼此间具有可比性,所以可把它看成是一种距离的度量,并记作 $d_{i,k}$,注意,式(3)中的

$\frac{d}{2} \ln 2\pi, -\frac{1}{2} \ln |\Sigma_i|, \ln P(w_i)$ 对于相同类都是常数.可以看到,这样的距离定义其实就是一种 Mahalanobis 距离.

(5) 如果没有负反馈,那么就利用得到的距离进行排序,并提交给用户最新的结果.

4 负反馈

许多算法,例如 Rui 和 Huang^[5]采用了同样的方式来处理正、负反馈的结果,在他们的方法中,简单地将所有图片和正、负反馈对应的中心距离值相减,并将得到的结果作为最终的距离.这种方法有着严重的不合理性.在没有反馈或只有正反馈的情况下,假定特征分布在满足高斯分布的前提下,距离类中心的等高线在二维空间为

椭圆形的.当引入负反馈后,采用上述方法,其距离的等高线就成为双曲线了,这与我们对数据分布的假定严重矛盾.同时,通过大量试验我们发现,用户提交的正反馈的图片在语义上有着极强的相似性,通常可以看做是属于同一个语义类,但是提交的负反馈的图片却恰恰相反,语义上没有必然的联系.因此,我们认为负反馈的图片在语义上彼此是孤立的、独立的.

基于上述结论,我们采用如下方法处理负反馈信息.我们只惩罚那些距离负反馈样本很近的图片,而基本上不影响其他图片,方法是对于那些距离负反馈样本很近的图片,增加它们和检索类间的距离.因此,这样一个惩罚函数的图像看起来就像这样:函数值在绝大多数定义域中都接近 0,只是在每个负反馈样本的地方有一个尖峰.作为一种模拟的手段,我们采用了混合高斯模型来拟合这条曲线.假定对于样本图片 P_k 存在一个负反馈集合 $C_n = \{P_{n_1}, \dots, P_{n_l}\}$.惩罚函数定义为

$$P(d_{i,k}) = \sum_{j=1}^l (p_{P_{n_j}}(P_i) \times d_{k,n_j}), \quad (7)$$

其中 $p_{P_{n_j}}(x)$ 的定义见式(2), $\varepsilon = P_{n_i}$, $\Sigma = I$.因此最终图片 P_i 到检索 P_k 间的距离为

$$d_{i,k} = d_{i,k} + P(d_{i,k}). \quad (8)$$

5 实验结果

表 1 显示的是我们的系统使用的 6 种图像特征.

Table 1 Low level features used in our system

表 1 我们的系统中使用的底层特征

Low level features	
1	Color histogram in HSV space with quantization 256
2	First and second color moments in Lab space
3	Color coherence vector in LUV space with quantization 64
4	MRSAR texture feature
5	Tamura coarseness histogram
6	Tamura directionality

底层特征, 在 HSV 颜色空间中量化到 256 维的颜色直方图, 在 Lab 颜色空间加一阶和二阶矩, 在 LUV 颜色空间中量化到 64 维的颜色一致性特征, 多尺度的自回归纹理特征, Tamura 稀疏度特征, Tamura 方向性特征.

我们使用的图像来自于“Corel Image Gallery”,从中随机选取了 1 600 幅图,这些图被手工标注在 32 个语义类中,每个类包含 50 幅图片.在检索过程中,我们认为和样本图片同属于同一个语义类的结果是正确的,反之则是错误的.我们的正反馈过程是这样的:排在前 50 幅的正确图片被自动标注为正反馈,对于负反馈,实验是将前两个不正确的结果自动标注为负反馈样本并提交给系统.我们同文献[5,6]分别做了对比性实验.结果显示在表 2~表 4 中.我们对于正确率的定义如下:

$$R = \frac{\text{relevant ones retrieved in window}}{\text{window size}}. \quad (9)$$

Table 2 Retrieval accuracy without relevance feedback

表 2 没有相关反馈的检索正确率

	Top10	Top20	Top30	Top40	Top50	Top100
R&H ^[5]	0.031	0.031	0.121	0.151	0.160	0.142
R&H ^[6]	0.432	0.362	0.317	0.284	0.258	0.183
Our algorithm	0.430	0.362	0.316	0.283	0.258	0.182

Table 3 Retrieval accuracy after 1st relevance feedback

表 3 第 1 轮相关反馈后的检索正确率

	Top10	Top20	Top30	Top40	Top50	Top100
R&H ^[5]	0.231	0.201	0.187	0.172	0.181	0.153
R&H ^[6]	0.650	0.529	0.463	0.415	0.375	0.246
Our algorithm	0.749	0.611	0.529	0.467	0.416	0.268

Table 4 Retrieval accuracy after 10th relevance feedback

表 4 第 10 轮相关反馈后的检索正确率

	Top10	Top20	Top30	Top40	Top50	Top100
R&H ^[5]	0.251	0.203	0.192	0.176	0.190	0.152
R&H ^[6]	0.776	0.597	0.512	0.455	0.408	0.264
Our algorithm	0.831	0.737	0.652	0.564	0.483	0.292

通过对上述几个数据表格的分析不难发现,在没有反馈的情况下,我们的算法给出的结果并不是最好的,但是只要存在反馈信息,我们的算法就会比其他两种方法好得多.表 5 显示了经过一轮的负反馈系统的检索正确率.可以清楚地看到,我们对于负反馈采用的策略大大提高了检索的精度.

Table 5 Retrieval accuracy after negative feedback

表 5 负反馈后的检索正确率

	<i>iFind</i>	R&H ^[5]
Before feedback	0.749	0.231
After feedback	0.827	0.232
Increase	0.078	0.001

反馈前, 反馈后, 正确率增加.

我们还在一个更大的图片库中进行了对比实验,图片数量有 12 000 个,分属于 100 个不同的语义类.结果显示,我们的算法取得的结果同样优于对比的几种算法.

6 结 论

本文提出了一种基于贝叶斯分类的新的相关反馈的算法.我们分别采用贝叶斯分类器来处理反馈信息,同时采用不同的策略分别处理正、负反馈.与其他算法相比,我们的算法具有学习能力,分类器的参数可以随着用户的反馈过程得到不断的更新,提高了系统的检索能力.而且,通过在大图像库中得到的实验结果我们发现,无论是处理正反馈还是负反馈,我们的算法都比其他算法取得了更高的准确率.

致谢 所有这些工作均在微软亚洲研究院完成,感谢 Chunhui Hu, Xingquan Zhu 和 Jacky Shen,他们给予了本文许多有价值的建议.

References:

- [1] Aalbersberg, I.J. Incremental relevance feedback. In: Belkin, N.J., ed. Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information retrieval. Copenhagen: ACM Press, 1992. 11~22.
- [2] Harman, D. Relevance feedback revisited. In: Belkin, N.J., ed. Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Copenhagen: ACM Press, 1992. 1~10.
- [3] Cox, I.J., Minka, T.P., Papatomas, T.V., et al. The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments. IEEE Transactions on Image Processing, 2000,9(1):20~37.
- [4] Rui, Y., Huang, T.S. Relevance feedback: a power tool for interactive content-based image retrieval. IEEE Circuits and Systems for Video Technology, 1999,8(5):644~655.
- [5] Vasconcelos, N., Lippman, A. Learning from user feedback in image retrieval systems. In: Proceedings of the NIPS'99. 1999. <http://www.media.mit.edu/people/nuno/publications.html>.
- [6] Su, Z., Zhang, H., Ma, S. Relevant feedback using a Bayesian classifier in content-based image retrieval. In: Yeung, M.M., et al, eds. Proceedings of the SPIE Storage and Retrieval for Media Databases, Vol 4315. San Jose: SPIE Press, 2001. 97~106.
- [7] Su, Z., Zhang, H., Ma, S. Using Bayesian classifier in relevant feedback of image retrieval. In: Tittsworth, M., ed. Proceedings of the 12th IEEE International Conference on Tools with Artificial Intelligence (IEEE ICTAI 2000). Vancouver: IEEE CS Press, 2000. 258~261.
- [8] Rui, Y., Huang, T.S. A novel relevance feedback technique in image retrieval. In: Buford, J., ed. Proceedings of the 7th ACM International Conference (part 2) on Multimedia (Part 2). New York, NY: ACM Press, 1999. 67~70.
- [9] Duda, R.O., Hart, P.E. Pattern Classification and Scene Analysis. New York: John Wiley & Sons, 1973.

An Image Retrieval Relevance Feedback Algorithm Based on the Bayesian Classifier*

SU Zhong^{1,2}, ZHANG Hong-jiang³, MA Shao-ping^{1,2}

¹(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China);

²(The State Key Laboratory of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China);

³(Microsoft Research Asia, Beijing 100080, China)

E-mail: suzhong_bj@hotmail.com

http://www.tsinghua.edu.cn

Abstract: The biggest problem in content-based image retrieval (CBIR) is a big gap between high-level semantic contents and low-level features. As an effective solution, relevance feedback has been put on many efforts for the past few years. In this paper, a new relevance feedback approach with progressive learning capability is proposed. It is based on a Bayesian classifier and treats positive and negative feedback examples with different strategies. It can utilize previous users' feedback information to improve its retrieval ability. The experimental results show that this algorithm achieves high accuracy and effectiveness on real-world image collections.

Key words: content-based image retrieval; relevance feedback; Bayesian classifier

* Received December 21, 2000; accepted April 3, 2001

Supported by the National Natural Science Foundation of China under Grant No.69823001; the National Grand Fundamental Research 973 Program of China under Grant No.G1998030509

=====

第3届全国虚拟现实与可视化学术会议(CCVRV 2003)

征文通知

由中国计算机学会虚拟现实与可视化技术专业委员会和中国图像图形学会虚拟现实专业委员会联合主办、国防科学技术大学承办的第3届全国虚拟现实与可视化技术及应用学术会议将于2003年10月18日~20日在长沙/张家界举行。

一、征文范围(包括但不限于)

建模技术,动画技术,可视化技术,遥操作技术,多媒体技术,人机交互技术,图形图像,仿真技术,分布式系统,VRML 技术,空间化声音,模式识别技术,虚拟制造,人机工效,网络技术,VR 传感器技术,可视化地理信息系统,基于图像的视景生成技术,虚拟现实与可视化应用系统。

二、征文要求

- 1、论文未被其他会议、期刊录用或发表;
- 2、来稿一式3份,并提交电子文档(word 格式)软盘,同时接受电子投稿;
- 3、论文包含:题目、中英文摘要、正文、参考文献等,正式格式见论文录用通知;
- 4、投稿者请务必写清姓名、单位、通信地址、电话、及 E-mail 地址。

三、重要日期

征文截止日期:2003年5月15日(收到) 录用通知日期:2003年6月30日(发出)

四、会议网址 <http://vrlab.buaa.edu.cn/ccvrv2003>

五、来稿联系方式(请注明 CCV RV2003 会议论文)

通信地址:100083 北京航空航天大学 6863 信箱

联系人:吴威 陈小武 电话:(010)82317109 传真:(010)82317644 E-mail: ccvrv2003@vrlab.buaa.edu.cn