

# 基于内容图像检索的特征子空间抽取\*

苏中<sup>1,2+</sup>, 马少平<sup>1,2</sup>, 张宏江<sup>3</sup>

<sup>1</sup>(清华大学 计算机科学与技术系,北京 100084)

<sup>2</sup>(清华大学 智能技术与系统国家重点实验室,北京 100084)

<sup>3</sup>(微软亚洲研究院,北京 100080)

## Feature Subspaces Extraction for Content-Based Image Retrieval

SU Zhong<sup>1,2+</sup>, MA Shao-Ping<sup>1,2</sup>, ZHANG Hong-Jiang<sup>3</sup>

<sup>1</sup>(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

<sup>2</sup>(State Key Laboratory of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China)

<sup>3</sup>(Microsoft Research Asia, Beijing 100080, China)

+ Corresponding author: Phn: 86-10-62986677 ext 305, Fax: 86-10-82899633, E-mail: suzhong\_bj@hotmail.com

<http://www.tsinghua.edu.cn>

Received 2001-07-23; Accepted 2001-10-16

**Su Z, Ma SP, Zhang HJ. Feature subspaces extraction for content-based image retrieval. *Journal of Software*, 2003,14(2):190~193.**

**Abstract:** Relevance feedback (RF) is used as an effective solution for content-based image retrieval (CBIR). Although it is effective, the RF-CBIR framework does not address the issue of feature extraction for dimension reduction and noise reduction. In this paper, a novel method is proposed for extracting features for the class of images represented by the positive images provided by subjective RF. Principal component analysis (PCA) is used to reduce both noise contained in the original image features and dimensionality of feature spaces. The method increases the retrieval speed and reduces the memory significantly without sacrificing the retrieval accuracy.

**Key words:** content-based image retrieval (CBIR); principal component analysis (PCA); relevance feedback

**摘要:** 作为一种有效的解决手段,相关反馈(relevance feedback)技术在基于内容图像检索(content based image retrieval)的研究中得到了深入的发展.尽管有效,已有的反馈算法却始终没有解决特征空间的有指导降维和特征中的噪声去除这两个问题.提出了一种新的方法,通过对用户在检索过程中提供的正反馈样本在各特征空间中的分布特性,利用主成分分析(principal component analysis)来消除特征中的噪声,实现了对特征空间进行有效的降维.试验结果显示,该方法在不牺牲检索精度的前提下提高了检索速度,降低了存储复杂度.

**关键词:** 基于内容的图像检索;主成分分析;相关反馈

**中图法分类号:** TP391 **文献标识码:** A

\* Supported by the National Natural Science Foundation of China under Grant No.69823001 (国家自然科学基金); the National Grand Fundamental Research 973 Program of China under Grant No.G1998030509 (国家重点基础研究发展规划(973))

**第一作者简介:** 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.

CBIR 就是在图像数据库中,检索与用户所提交样本图像在内容上一致或相似的图像集合的过程,通过对图像底层特征的比较来实现检索.一般使用的底层特征是从图像中利用算法自动抽取的有关颜色、纹理和形状特征.CBIR 中最大的问题在于图像所包含内容的语义信息和图像底层特征间的巨大差异,即语义上相似的图像在特征空间中距离往往很远.相反,语义上完全不同的图像有时在特征空间中却非常相近.为了解决这个矛盾,在 CBIR 中引入了 RF 的办法<sup>[1,3-5]</sup>.反馈过程是一个机器学习过程,特征维度越高,在学习时所需样本也越多,但实际系统不可能期望用户在检索过程中提交非常多的反馈信息.同时,在特征中存在大量和检索无关的噪声信息,这将大大影响检索的精度.高维度还会加大检索时的运算负担和更高的数据存储需求.因此,运用相关反馈,有指导地进行特征降维是解决上述问题的一种有效途径.本文要研究的问题就是利用反馈进行特征空间有指导降维和噪声去除.

本文第 1 节介绍我们提出的 PCA 子空间中的反馈算法.第 2 节介绍试验结果.第 3 节给出结论.

### 1 PCA 子空间中的反馈算法

下面给定一些符号定义.图像数据库表示为  $D$ .假定对于图像我们抽取了  $m$  种不同的特征,每种特征构成了一个维度为  $n_i$  的特征空间  $S_i$ ,这样每幅图像可以表示为一组特征向量  $\bar{x} = [\bar{x}_1, \dots, \bar{x}_m]$ ,其中  $\bar{x}_i$  是该图像在  $S_i$  中的投影.假定同类图像中的每种特征满足高斯分布,则  $\bar{x}_i \sim N(\varepsilon_i, \Sigma_i)$ , $\Sigma_i$  是  $n_i \times n_i$  维的协方差矩阵, $\varepsilon_i$  是  $n_i$  维的均值向量.实际运用中用简化的对角阵  $diag\{\sigma_i^2\}$  来替代  $\Sigma_i$ ,其中  $\sigma_i(m) = \Sigma_i(m, m)$ .简化的原因是因为在反馈样本数目不多的情况下,太多的参数估计会引入大量的误差.

反馈算法的基本思路是:假定在用户任何一轮检索和反馈中,所有正反馈包括用户提交样本图像构成的某一图像语义类在特征空间中满足高斯分布,这种假设经过试验证明是基本成立的,基于这种假定,就可以根据正反馈图像的特征不断地修正该高斯类的参数并优化检索结果.对于特征向量  $x$  属于该语义类  $c_i$  的对数后验概率,可以由下面的贝叶斯公式估计出来:

$$g_i(x) = \ln P(c_i | x) \propto \ln p(x | c_i) | \ln P(c_i) \\ = -\frac{1}{2}(x - \varepsilon_i)^T \Sigma_i^{-1} (x - \varepsilon_i) + \text{const}, \tag{1}$$

其中  $\text{const}$  在一轮反馈过程中是固定常数.该公式只针对一种特征.假定不同种特征在分布上是彼此独立的,对于多种特征的对数后验概率就可以表示为各个  $g_i(x)$  的和.然后利用该概率值作为检索排序的依据.根据式(1),当有新的正反馈样本提交时,有两个参数需要更新:  $\varepsilon_i, \Sigma_i$ .实际上,这样一个过程可以看做是对于两个高斯类的合并,因此,假定新提交的正反馈集合为  $U$ ,可以很容易地得到下列的参数更新方法:

$$(n+|U|)\sigma_i^2 = n\sigma_i^2 + \frac{n|\varepsilon_i - \sum_{u \in U} u|}{n+|U|} + \sum_{u \in U} u^2 - \frac{\left(\sum_{u \in U} u\right)^2}{n+|U|}, \tag{2}$$

$$\varepsilon_i = \frac{n \times \varepsilon_i + \sum_{u \in U} u}{n+|U|}, \tag{3}$$

$$n = n+|U|. \tag{4}$$

为了从原始特征空间抽出更能有效地体现用户需求的特征子空间,我们利用用户的交互信息,将主成分分析和相关反馈进行了有效的融合.在对原始特征空间进行 PCA 过程后,特征值的大小决定了对应维度的重要性,因此对于投影到该空间的图像特征在各维度间也有了不同的重要性.子空间反馈的基本思路就是只用一定比率( $\psi$ )重要维度进行检索而抛弃不重要的维度,这样初始的检索利用维度为  $m_i = \psi \times n_i$ .这里可以将比率  $\psi$  看成是特征降维的比率,例如  $\psi = 10\%$ ,即意味着只用大约十分之一维度的特征进行检索.对于以后的反馈过程,根据下面的策略对各特征子空间的维度进行修正.在检索中,检索使用的特征维度总和保持基本不变,根据每一轮反馈的结果,对各种特征进行评价,好的特征可以引入更多的维度进入下一轮检索,反之,将在已有的基础上再次降低检索中使用的维度.因此,反馈中维度的变化也就是对  $m_i$  的不断调整.假定  $Q$  为检索样本图,当前的正反馈

集合为  $U$ ,对每种特征的评价函数定义如下:

$$M_i = \frac{\sum_{p_i \in D} \frac{d(p_i, Q)}{|D|}}{\sum_{u \in U} \frac{d(u, Q)}{|U|} \times (1 - \alpha_i)}, \quad (5)$$

其中  $d(x, y)$  表示在 PCA 子空间向量  $x$  和  $y$  之间的距离,  $\alpha_i$  为该特征当前轮次中使用维度的对应特征值之和. 从上述评价函数可以看到,好的特征应该具有这样的性质,即在该空间中,正反馈图像与其他图像相比,到样本图的距离要小得多.这种评价方法比较直观,当然,其他评价函数在这里也可以运用.在计算  $M_i$  后,将  $M_i$  降序排序,这样可以得到  $M_i$  对应的排序位置  $r_i$ ,该排序位置给出了不同种特征间的比较,特征越好,排序越靠前.因此,可以根据排序值对  $m_i$  进行修正:

$$m_i = \left( m_i + \tau \times \left( \left\lfloor \frac{m}{2} \right\rfloor - r_i - 1 \right) \right). \quad (6)$$

因此,对于好的特征,其对应的  $m_i$  会增加,从而使更多的维度引入检索.  $\tau$  是比率常数,我们的试验中将  $\tau$  置为 1.下面描述该反馈算法:

离线(offline)初始化.

对于每种特征  $i=1, \dots, m$ ,进行下列操作:

- (1) 对原始特征空间  $S_i$  作 PCA 变换,得到特征值和特征向量.按特征值大小对相应的特征向量降序排列.
- (2) 将所有原始特征投影到 PCA 空间中,并白化(whitened)特征.注意,在这一步对特征并未进行降维.
- (3) 对每个图像的检索参数  $\sigma_i, \varepsilon_i$  和  $n$  进行初始化  $\sigma_i = I$  (单位对角阵),  $\varepsilon_i = \bar{x}_i, n = 1$ .

反馈.

- (1) 在反馈过程中按照式(2)、式(3)和式(4)对高斯类的参数进行修正.使用的特征空间是 PCA 空间.
- (2) 根据式(5)、式(6)修正  $m_i$ .
- (3) 在  $m_i$  维的特征子空间中计算图像库中所有图像到检索样本间的距离.
- (4) 根据所有计算出的距离排序,并将新的排序列表呈现给用户.

我们的方法有以下优点:(1) 白化特征分布,从而使得距离的度量在新的子空间中更为有效;(2)降低了原始特征中的噪声,对有利于检索的特征进行了加强,对那些不利于检索的特征进行了削弱,从而提高了检索精度;(3) 对特征子空间进行了有效的降维,存储空间和计算速度与特征维度呈线性正比关系,从而减少了存储空间并提高了检索速度;(4) 反馈中的参数估计数目的减少降低了对用户提交反馈样本数目的要求.

## 2 试验结果

试验图像来源于 Corel 图像库,共选择了 79 个类共 10 009 张彩色照片.试验中,用 80%的图像学习并构造 PCA 特征子空间,其余 20%作为样本检索图在整个图库上进行检索测试.检索的正确答案以 Corel 图库本身的分类信息作为标准.检索系统中使用了 6 种特征:3 种颜色、3 种纹理特征,总维度为 435 维.反馈试验是这样进行的,从测试集中选取一个图像进行检索,在每一次反馈中,随机选取一个正确答案作为正反馈,这样反复进行.评价检索的准确率定义如下:

$$Accuracy = \frac{\text{relevant images retrieved in top } N \text{ returns}}{N}.$$

首先,通过试验比较了在原始特征空间我们提出的基于贝叶斯估计的相关反馈算法和 Vasconcelo<sup>[1]</sup>以及 Rui<sup>[3]</sup>提出的反馈算法在试验图像库上的检索准确率.图 1 显示了该反馈算法在检索精度上的优越性.下面,我们将对特征进行降维,看看在降维后检索精度会有什么变化.我们对 PCA 子空间中的图像反馈准确率和原始特征空间中反馈准确率进行了对比.从图 2 可以看到,在特征只用了 30%维度的 PCA 反馈就达到了和原始空间几乎相同的检索准确率.也就是说,检索速度提高了 3 倍,内存消耗只有原始的 1/3,而检索准确率并没有下降,在有些

时候还略有上升.

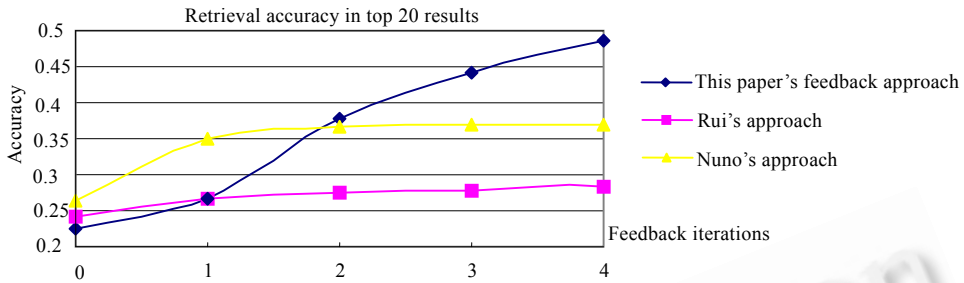


Fig.1 Retrieval accuracy for top 20 results in original feature space

图1 原始特征空间前 20 个结果的检索正确率

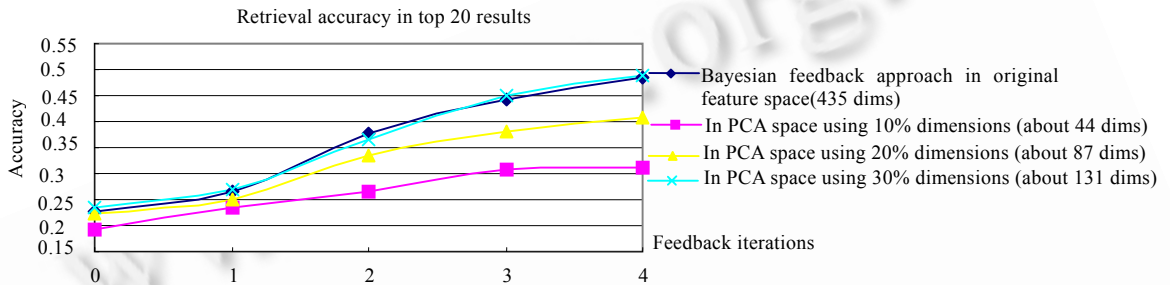


Fig.2 Comparing the retrieval accuracy in top 20 results between the original feature space and PCA feature space

图2 对原始特征空间和 PCA 特征空间中前 20 个结果的检索正确率比较

### 3 结论

本文提出了一种新的图像反馈算法,即将主成分分析与贝叶斯反馈进行了有机的结合,从而由原始特征空间抽取有效的特征子空间,并根据反馈动态调整各个特征子空间.反馈算法在这里扮演了两个角色:提供贝叶斯反馈中高斯参数的调整依据以及对特征子空间维度变化进行指导.后者体现了与已有的相关反馈算法的本质区别.试验结果显示,本文提出的贝叶斯反馈算法与其他算法相比,在检索准确率上有一定的优势,而特征子空间的抽取则提高了检索的时间和空间效率.

#### References:

- [1] Cox IJ, Minka TP, Papathomas TV, Yianilos PN. The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments. *IEEE Transactions on Image Processing*, 2000,9(1):20~37.
- [2] Diamantaras I, Kung SY. *Principal Component Neural Networks, Theory and Applications*. New York: John Wiley & Sons, Inc., 1996.
- [3] Rui Y, Huang TS. A novel relevance feedback technique in image retrieval. In: *Proceedings of the 7th ACM International Conference (Part 2) on Multimedia (Part 2)*. Orlando: ACM Press, 1999. 67~70.
- [4] Su Z, Li S, Zhang H. Extraction of feature subspaces for content-based retrieval using relevance feedback. In: *Proceedings of the 9th ACM Multimedia Conference*. Orlando: ACM Press, 2001. 98~106.
- [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>.