

# 一种高效的图像检索方法

陈旭文, 朱红丽

(浙江大学城市学院信息与电气工程分院 杭州 310015)

**摘要:** 本文在分析了基于内容的图像检索中用户反馈信息的空间特性和时间特性的基础上提出了一种综合用户反馈信息时空特性的用户反馈学习新方法——时空特性支持向量机(SKSVM)方法。实验证明SKSVM方法和普通支持向量机方法相比,特别是在图像数据库复杂的情况下,能够明显的提高用户反馈的学习能力,提高检索结果的准确率;同时还能在检索过程中准确实时的跟踪用户的查询需求,是一种高效的图像检索方法。

**关键词:** SKSVM; 图像检索; 用户反馈; 学习

## Efficient image retrieval method

Chen Xuwen, Zhu Hongli

(College of Information & Electrical Engineering, ZheJiang University City College, Hangzhou 310015, China)

**Abstract:** The spatial and temporal characteristics of user feedback information in content-based image retrieval are analyzed, based on which a novel user feedback learning method, namely, SKSVM is proposed. The method treats the positive and negative examples returned at different rounds with different weights in the learning process according to the spatial and temporal characteristics of the relevance feedback information. With these features, the proposed SKSVM method improves the performance of the learning significantly as compared with the conventional SVM method.

**Key words:** SKSVM; image retrieval; user feedback; learning

### 1 引言

由于图像的底层视觉特征很难直接反映用户的高层语义,系统通常采用用户反馈等人机交互技术来减小用户的高层感知与图像的底层特征之间的差距,以改善检索性能。在用户反馈过程中,系统首先要求用户对一些待评价图像的相关性做出评价,然后通过学习获得的用户用户反馈信息来重新检索图像,改进检索结果,同时产生出新的一些待评价图像供用户做进一步的评价。若用户对检索结果不满意,则对待评价图像再次进行评价,系统再次学习用户反馈信息,改进检索结果,如此循环反馈,直到检索出用户满意的结果为止。

目前,基于内容的图像检索中采用了很多不同的用户反馈技术来提高检索性能,其中,一些用户反馈技

术试图通过寻找代表用户查询需要的对象以及对每一底层特征的权重来改善查询结果<sup>[1-2]</sup>;一些技术采用概率模型来估计用户查询需要<sup>3-4]</sup>。无论检索系统采用何种用户反馈技术,如何合理的利用用户反馈信息,快速准确的学习到用户的查询需求,从而行之有效的提高检索精度是系统的关键性问题之一。在对用户反馈信息的用户反馈学习中,一些学习方法<sup>[5]</sup>应用启发式知识,同时利用了用户反馈信息中的正例和负例,但通常对图像底层特征做出了独立性的强制假设;一些方法<sup>[6]</sup>将正例集合和负例集合作为两个平等的类来看待,却忽视了两个集合具有不同特性的事实。

为了在用户反馈学习中充分考虑到用户反馈信息的空间特性和时间特性,本文提出了一种基于内容的图像检索中的综合了用户反馈信息时空特性的新方法——时空特性支持向量机(SKSVM)方法。SKSVM将用户反馈学习视为一种时空有偏的两类分类学习问

题,其学习目的是根据用户反馈信息训练出一个用于分类的分界面,继而从图像数据库中检索出满足用户需求的图像。

2 SKSVM 用户反馈学习方法

用户反馈信息中的反馈图像  $O$  用四元组  $(x, y, t, s)$  来表示,其中  $x \in R^N$ , 是反馈图像  $O$  的图像底层特征向量,  $y \in \{-1, 1\}$  是反馈图像  $O$  的用户评价,  $t$  是  $O$  对应的用户反馈时间,  $s$  是  $O$  对应的时空权重 ( $0 < s \leq 1$ )。于是,用户反馈学习中的用户反馈信息构成了训练集  $S = \{(x_i, y_i, t_i, S_i) | 1 \leq i \leq L\}$ , 其中,  $L$  是训练集中反馈图像的总数目。

对于给定的训练集  $S$ , 求最优分界面  $(w, b)$ , 使得

$$\Phi(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^L S_i \xi_i \tag{1}$$

最小, 并且满足约束条件

$$y_i(x_i \cdot w + b) \geq 1 - \xi_i, (1 \leq i \leq L), \xi_i \geq 0, (1 \leq i \leq L)$$

式中:  $w \in R^N; b \in R; C$  为平衡常数;  $\xi_i$  是松弛变量。

一般情况下, 用户反馈学习中的分类学习问题是线性不可分的问题, 式 (1) 中的松弛变量  $\xi_i$  实际上表示了第  $i$  幅反馈图像在分类中被分界面错分的程度,  $\sum_{i=1}^L S_i \xi_i$  是分界面对所有反馈图像的加权错分程度的总和。在本方法中,  $C$  往往取得足够大, 使加权错分程度的总和在分界面的确定中起到足够大的作用。

反映用户反馈信息时空特性的时空权重  $S_i$  定义如下:

$$S_i = \begin{cases} s_+ - g(t - t_i)(s_+ - s_-), & y_i = +1 \\ s_-, & y_i = -1 \end{cases}$$

式中:  $s_+$  和  $s_-$  是常数, 满足  $s_+ \geq s_- > 0$ ;  $t$  是当前用户反馈的循环次数;  $g(t - t_i)$  为时间因子。

$$g(t - t_i) = \begin{cases} \frac{t - t_i}{w}, & 0 < t - t_i < w \\ 1, & t - t_i \geq w \end{cases}$$

当  $s_+ > s_-$  时, 在用户反馈学习中任何正例对应的时空权重都大于或等于任何负例对应的时空权重。这样的设定, 强化了用户感兴趣的正例集合, 弱化用户不感兴趣的负例集合, 达到合理对待正负两个集合的效果。

当  $s_+ = s_-$  时, 所有正例和负例对应的时空权重相同, 此时用户反馈学习就不再考虑用户反馈信息的空间特性。

对于正例 ( $y_i = 1$ ) 的时空权重  $s_i$  来说, 当  $t = t_i$  时,

$s_i = s_+ - g(t_i - t_i)(s_+ - s_-) = s_+$ ,  $s_i$  取到最大值  $s_+$ ; 当  $t \geq t_i + w$  时,  $s_i = s_+ - (s_+ - s_-) = s_-$ ,  $s_i$  取到最小值  $s_-$ 。由此可见, 一个正例对应的时空权重  $s_i$  随着反馈循环的次数增大而减小。

而对于负例, 用户一旦将其评价为不相关图像, 在以后的反馈过程中一般仍保持不相关性。

当查询需求确实发生转移时, 正例的时空权重  $s_i$  随着反馈循环次数增加而逐渐减小, 经过  $w$  次循环后, 查询转移前被评价为相关的图像对分类面的确定的影响将达到最小, 等同于负例的影响程度; 而对于查询转移后的新查询需求, 代表该查询需求的相关图像成为用户反馈信息中的新近正例, 具有较大的影响程度, 能够迅速的调整分界面使之符合新的查询需求。

3 用 SKSVM 方法的图像检索

由于 SKSVM 将检索作为分类问题处理, 初始用户反馈信息中需要一些相关图像和一些非相关图像。因此, 用户在开始查询时首先要向系统提交一幅或者多幅初始正例图像和一幅或者多幅初始负例图像。

接着, SKSVM 对反馈信息进行学习, 训练出最优分界面  $w \cdot x + b = 0$  后。根据得到的分界面, 系统按以下方法生成检索结果和待评价图像, 返回给用户:

1) 系统通过计算图像数据库中的每一幅图像 (底层特征为  $Z$ ) 到最优分界面的距离来判断该图像属于哪一类。系统从相关类中按照距离绝对值从大到小顺序选取  $k$  幅图像作为检索结果返回给用户。

2) 系统将图像数据库中到分界面的距离绝对值最小的若干幅图像作为待评价图像返回给用户。

3) 若用户对检索结果不满意, 则对待评价图像做出相关性评价, 系统将这些评价过的图像加入到反馈信息中, 开始新的一次反馈循环; 若用户对检索结果满意, 则检索结束。

4 实验结果与分析

为了验证综合时空特性的用户反馈学习方法的有效性, 本文分别在四个不同大小的图像数据集上针对 SKSVM 对用户反馈信息的空间特性和时间特性的体现做了实验。如表 1 所示, 实验从图库中选取多个主题的图像构造了四个大小不同但是依次包含的图像集, 随着图像集中主题数目的增加, 图像在底层特征空间上的分布也不断趋于复杂。

表 1 中每一图像主题都是由相关领域的专家根据经验选择的在视觉及语义上相关的 100 幅图像组成。因此,实验将此作为衡量检索正确与否的标准,即视属于同一主题的两幅图像为相关,而视不属于同一主题的两幅图像为不相关。

表 1 四个图像集的组成

图像集	构成图像集的主题	说明
1	风景、河流等四个主题	每主题包含 120 幅图像
2	图像集 1 及鸟、云等 10 个主题	每主题包含 100 幅图像
3	图像集 2 及建筑物等 15 个主题	每主题包含 100 幅图像
4	图像集 3 及飞机等 42 个主题	每主题包含 100 幅图像

实验对图像提取了两类底层视觉特征:颜色特征和纹理特征。其中,颜色特征提取  $HSV$  色彩空间中  $H$  维 16 bin 和  $S$  维 4 bin,组成  $16 \times 4 = 64$  bin 的直方图;纹理特征由共生矩阵提取  $0^\circ, 45^\circ, 90^\circ, 135^\circ$  四个方向上的能量、熵、对比度以及同一性共 16 维特征。图像采用由这两类特征组成的 80 维向量来表示。

实验在所有图像集均包含的图像数据集 1 中的 4 个主题分别进行检索,对每一主题分别做 20 次随机检索。表 2 和表 3 是分别采用 SKSVM 方法和传统 SVM 方法在四个图像集上经 3 次用户反馈后得到的查准率。

5 结 论

基于内容的图像检索中,用户反馈学习是减小图像底层视觉特征和用户高层语义之间差距的主要方法之一。如何有效利用用户的用户反馈信息,是提高系统性能的一个重要问题。我们分析了用户反馈信息的空间特性和时间特性,针对这些特性提出了一种综合用户反馈信息时空特性的用户反馈学习方法——时空特性支持向量机(SKSVM)方法。在学习,SKSVM 根据用户反馈信息的空间特性,强化用户感兴趣的正例集合,弱化用户不感兴趣的负例集合,合理利用正负例子蕴涵的信息;同时在时间上考虑反馈信息的时间特性,强化最近的反馈信息,实现快速准确的跟踪用户

的查询需求。实验结果表明,和传统的 SVM 方法相比,SKSVM 方法能够显著的提高检索的精度,同时能够快速的跟踪用户的查询需求,尤其在查询需求发生转移时尤为有效。

表 2 平均前 20 幅图的查准率

图像集	1	2	3	4
SKSVM 方法	98%	89%	82%	66%
传统 SVM 方法	97%	86%	79%	58%

表 3 平均前 100 幅图的查准率

图像集	1	2	3	4
SKSVM 方法	88%	71%	60%	45%
传统 SVM 方法	86%	68%	57%	38%

参考文献

[ 1 ] ZHOU X S, HUNAG T S. Exploring the Nature and Variants of Relevance Feedback[ J] . in Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries, Hawaii, Dec. 2001

[ 2 ] PENG J, BHANU B, QING S. Probabilistic feature relevance learning for content-based image retrieval, Computer Vision and Image Understanding, 1999, 75: 150-164.

[ 3 ] VASCONCELOS N, LIPPMAN A. Bayesian relevance feedback for content-based image retrieval[ J] . in Proc. IEEE Workshop on Content-based Access of Image and Video Libraries, CVPR 00, Hilton Head Island, SC, 2000.

[ 4 ] SU Z, ZHANG H, MA S. Relevant Feedback using a Bayesian Classifier in Content-Based Image Retrieval [ J] . SPIE Electronic Imaging 2001, January

[ 6 ] TONG S, CHANG E. Support Vector Machine Active Learning for Image Retrieval[ C] , ACM International Conference on Multimedia, pp. 107-118, Ottawa, October 2001.

[ 5 ] RUI Y, HUANG T S, ORTEGA M, et al. Relevance feedback: a power tool in interactive content based image retrieval[ J] . IEEE Trans. Circuits and Systems for Video Technology, 1998, 8(5): 644-655.