

中图法分类号: TP391.3 文献标志码: A 文章编号: 1006-8961(2010)11-1658-06

检索信息: 邓昌葛, 朱俊株, 尤庆成, 高如如. 基于非对称打包和 FSVM 的图像检索[J]. 中国图象图形学报, 2010, 15(11): 1658-1663

## 基于非对称打包和 FSVM 的图像检索

邓昌葛, 朱俊株, 尤庆成, 高如如

(中国科学技术大学电子科学与技术系, 合肥 230027)

**摘要:** 在图像检索的相关反馈中, 引入支持向量机分类方法虽可以提升图像的检索性能, 但是传统的支持向量机存在正样本数少、样本非对称、过学习和弱实时性的局限。针对上述问题, 提出了一种基于非对称打包的 FSVM 算法。该算法首先对负样本进行非对称打包处理, 最后结合模糊理论与 SVM 实现图像检索。Corel 图片集上的实验表明, 当正样本数较小时, 该新算法的平均查准率-查全率要优于已有算法。  
**关键词:** 基于内容的图像检索; 非对称打包; 模糊支持向量机

### Image retrieval using asymmetric bagging and FSVM

DENG Changge, ZHU Junzhu, YOU Qingcheng, GAO Ruru

(Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230027)

**Abstract:** Recently, SVMs (support vector machines) have been widely used in image retrieval as a method to improve the retrieval performance. However, conventional SVMs encounter four problems: small size of positive samples, asymmetry problem of training samples, over-fitting and weakly real-time. To solve these problems, an asymmetric bagging based fuzzy support vector machine (AB-FSVM) is proposed. An asymmetric bagging is made to negative samples, and then based on fuzzy theory and SVM, the retrieval images are gotten. Experimental results based on a set of Corel images show that the proposed system performs much better than the previous methods, especially when the size of positive samples is small.

**Keywords:** content-based image retrieval (CBIR); asymmetric bagging (AB); fuzzy support vector machine (FSVM)

## 0 引言

随着基于内容的图像检索 (CBIR) 技术的发展, 包含‘交互系统’和‘用户参与’的相关反馈构架, 在缩减高层语义概念与低层视觉特征之间的‘语义鸿沟’方面体现了较好的性能。特别是基于支持向量机 (SVM) 的相关反馈, 由于具有良好的泛化能力, 因而进一步提高了检索性能。

但是, 由于以下 4 个原因<sup>[1]</sup> 基于 SVM 的相关反馈性能会变差: 1) 小样本导致 SVM 分类器性能不

稳定, 若训练样本数较少, 则 SVM 的最优分类面将对训练样本较敏感, SVM 分类器性能会很差; 2) 不对称问题, 大多数情况下, 由于用户反馈的正样本数远小于负样本数, 因此在训练过程中, 如果将正负样本同等对待的话, 则决策面会发生倾斜, 会导致将不相关的图像误判为相关的; 3) 过度拟合问题, 由于特征向量的维数远大于用户每次提供的反馈样本数, 多数分类器会发生抖动问题, 导致 SVM 分类器的分类性能变差; 4) 由于用户参与到训练的过程中, 因此算法的实时性难以满足。

为此, Hoi 等人提出了基于日志的相关反馈方

收稿日期: 2009-05-18; 改回日期: 2009-09-08

第一作者简介: 邓昌葛 (1985—), 男, 中国科学技术大学电子科学与技术系电路系统专业硕士研究生。主要研究方向为数字信号处理及基于内容的图像检索。E-mail: adamdcg@gmail.com。

法<sup>[2]</sup>;Kim 等人使用一种自适应的方法<sup>[3]</sup>;Tong 和 Chang 研究了 SVM 的主动学习<sup>[4]</sup>;Tao 等人提出基于非对称打包和随机子空间的支持向量机 (ABRS-SVM)<sup>[5]</sup>,但该方法训练及判断速度缓慢,实时性不够好;Rui 等人在支持向量机中应用了模糊技术,提出了模糊支持向量机 (FSVM),使检索准确度和实时性得到很大提高,但是该算法在正样本数较小时仍存在局限性。针对 Rui 等人提出算法的不足,本文借鉴 Tao 等人的负样本非对称打包 (AB) 的思想<sup>[5]</sup>,并结合模糊理论,提出了一种基于非对称打包的模糊支持向量机算法 (AB-FSVM),从而提高了基于相关反馈的 SVM 图像检索性能。

## 1 本文算法

在基于支持向量机的相关反馈中,首先 CBIR 系统提供给用户一个初步的查询结果,其中往往只有很少量可以被用户按语义概念视为相关(即正样本)的图像。此时,正样本数远小于负样本数,面临正负样本不对称及正样本数小的问题,这也正是 Rui 等人提出的算法<sup>[6]</sup>局限所在。本文采用负样本打包的思想,提出了一种负样本非对称打包的 SVM 算法。该算法首先设定非对称打包个数  $N_{AB}$ ,同时通过对负样本进行随机抽样来得到若干负样本的  $N_{AB}$  个子集,称为负样本包;然后分别对每个负样本包与正样本进行 SVM 训练;最后对正样本求出归一化距离和,与最大归一化距离和相对应的负样本包将视为最优负样本包。

本文将负样本非对称打包 SVM 算法与模糊理论结合,形成了基于负样本非对称打包的 FSVM 算法。当反馈的正样本数小于等于一定值  $T$  时,就采用负样本非对称打包 SVM 算法来处理;当反馈的正样本数大于阈值  $T$  时,则应用 FSVM 算法来处理。

### 1.1 负样本非对称打包 SVM 算法

输入: 正样本序列  $S^+$ 、正样本个数  $N_p$ 、负样本序列  $S^-$ 、负样本个数  $N_n$ 、分类器  $I(\text{SVM})$ 、非对称打包个数  $N_{AB}$  (整数) 和阈值  $T$  (整数)。

For  $i = 1$  to  $N_{AB}$  {

$$S_i^- = \text{subset}(S^-), |S_i^-| = |S^+| \quad (1)$$

$$C_i = I(S_i^-, S^+) \quad (2)$$

$$f_i^{\max} = |\max(f(x_p), f(x_q))| \quad (3)$$

$$x_p \in S^+, x_q \in S_i^-; p, q = 1, 2, \dots, N_p$$

$$Sum_i = \sum_{p=1}^{N_p} \frac{f(x_p)}{f_i^{\max}}, x_p \in S^+ \quad (4)$$

$$C^*(x) = C_j(x, S_i^-, S^+) \quad (5)$$

且  $j$  满足  $Sum_j = \max(Sum_i | i=1, 2, \dots, M)$

输出: 分类器  $C^*$

上述算法中,式(1)用于对负样本进行非对称打包处理,式(4)用于求取正样本的归一化距离和。

### 1.2 FSVM 算法

所有使用 SVM 算法的图像检索系统都是基于以下假设:图像的特征可以代表图像;相似的图像在特征空间具有聚类性。对于一个基于 SVM 的图像检索系统,图像在原始的特征空间可能不具有可分性,但将原始的特征空间变换到高维的特征空间后,则可以进行区分。

#### 1.2.1 训练及判别模型

如今模糊理论已经成功地引入到 SVM 中<sup>[7]</sup>。训练时,给训练序列中的每个样本分配一个因子  $s_i \in [0, 1]$ 。因子  $s_i$  表明样本的重要性; $s_i$  为 1,意味着样本非常重要; $s_i$  为 0,意味着样本非常重要。一个标记过的训练样本序列为

$$S = \{(x_1, y_1, s_1), (x_2, y_2, s_2), \dots, (x_n, y_n, s_n)\} \quad (6)$$

式中  $x_i \in \mathbf{R}^D$ ,  $D$  是训练样本的维数,  $y_i$  是对应的标记,  $s_i$  是模糊因子,且  $0 \leq s_i \leq 1, i = 1, 2, \dots, n$ 。给定的标记  $y_i$  只有以下两个值:  $-1$  和  $+1$ , 则最优分类面问题等价于使下式最小化,

$$\Phi(w, \xi, s) = \frac{1}{2} (w^T w) + C \left( \sum_{i=1}^n s_i \xi_i \right) \quad (7)$$

且满足约束条件

$$y_i (w^T x_i + b) \geq 1 - \xi_i \quad (8)$$

$$\xi_i \geq 0; \quad i = 1, 2, \dots, n$$

其中  $C$  是一个常量,  $\xi_i$  是样本  $i$  产生的误差,  $s_i$  对  $\xi_i$  起限制作用。由于较小的  $s_i$  可减弱  $\xi_i$  的影响,因此对应的样本  $x_i$  重要性减弱,其产生的误差也较小。经过推导<sup>[7]</sup>,使用核函数的 FSVM 算法等价于使下式拉格朗日函数最大化

$$W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (9)$$

并满足以下约束条件:

$$\sum_{i=1}^n y_i \alpha_i = 0 \quad (10)$$

$$0 \leq \alpha_i \leq s_i C; \quad i = 1, 2, \dots, n$$

判别函数为

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \quad (11)$$

SVM 算法和 FSVM 算法的区别在于样本的重要性能否由模糊因子  $s_i$  来衡量。通常来讲, 对于一个给定的样本, 若  $|f(\mathbf{x})|$  值越大, 则对应的预测函数的可信度越高; 反之, 若  $|f(\mathbf{x})|$  值较小, 意味着样本靠近分类面, 则其对应的预测函数可信度就低。

### 1.2.2 聚类松散度

在式(6)中, 每个训练样本都被分配一个权值, 及模糊因子  $s_i$ 。通过边界聚类, Barbar 等人提出了一种计算样本聚类松散度  $\beta_i$  的方法<sup>[8]</sup>。本文通过  $\beta_i$  来求取  $s_i$ 。对样本  $i$ , 有

$$\beta_i = \frac{\sum_{j=1}^K d_{i,j}^1}{\sum_{j=1}^K d_{i,j}^{-1}} \quad (12)$$

其中  $K$  是一个常数。  $d_{i,j}^1$  是属于同一类的样本  $i$  和样本  $j$  之间的欧氏距离;  $d_{i,j}^{-1}$  是不属于同一类的样本  $i$  和样本  $j$  之间的欧氏距离。聚类松散度越大, 说明该样本的聚类性越差。考虑下述情形: 第  $i$  幅图像由  $K$  个非常相似的样本围绕, 同时远离  $K$  个不相似的样本, 此时  $\sum_{j=1}^K d_{i,j}^1$  很小,  $\sum_{j=1}^K d_{i,j}^{-1}$  很大,  $\beta_i$  的值趋于很小。理想情况是,  $\sum_{j=1}^K d_{i,j}^1$  趋于 0,  $\sum_{j=1}^K d_{i,j}^{-1}$  趋于  $\infty$ , 则  $\beta_i$  趋于 0, 此时聚类效果最好。  $K$  的值计算如下:

$$K = \min(N_p - 1, N_n - 1) \quad (13)$$

式中  $N_p, N_n$  分别是正、负样本数。

试验中, 对于  $d_{i,j}^1$  和  $d_{i,j}^{-1}$  的求取, 若 SVM 的核函数采用 RBF 函数, 则  $d_{i,j}^1$  和  $d_{i,j}^{-1}$  在 SVM 高维空间的欧氏距离如下<sup>[6]</sup>:

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{2 - 2\exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{2\sigma^2}\right)} \quad (14)$$

### 1.2.3 模糊因子

聚类松散度  $\beta_i$  属于 0 到无穷, 但由于在该区间并不是均匀分布, 因此引入映射因子<sup>[6]</sup>

$$\mu_i(\beta) = \begin{cases} 2^{-3\beta} \times 0.5^{-2} & \beta \leq 1 \\ (1 - 2^{-3/\beta}) \times 0.5^{-2} & \beta > 1 \end{cases} \quad (15)$$

则模糊因子为

$$s_i = \begin{cases} \mu_i \times \frac{N_n}{N_p + N_n} & i \in S^+ \\ \mu_i \times \frac{N_p}{N_p + N_n} & i \in S^- \end{cases} \quad (16)$$

### 1.3 构建的检索系统

为评估本文算法的性能, 笔者开发了一个 CBIR 系统(如图 1 所示), 由图 1 可见, 该系统流程为: 首先对查询图像进行低层特征提取; 然后按照欧氏距离, 对图像库中的所有图像进行排序, 同时用户将排列靠前的图像标记为正反馈或负反馈; 接着利用这些反馈样本, 经过 SVM 训练, 得到 SVM 分类模型; 最后将该分类模型作为新的相似性判断准则, 再对图像库中的所有图像进行重新排序。这样的反馈过程迭代进行, 直到用户对查询结果满意为止。

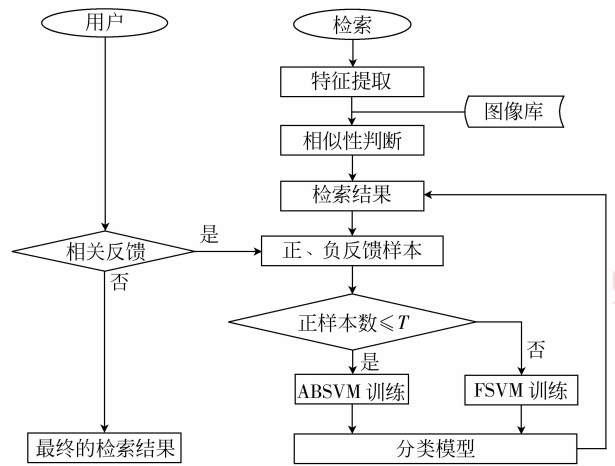


图 1 系统流程图

Fig. 1 Flow chart of the system

## 2 实验结果与分析

实验中为了将本文算法与现有的算法进行比较。借鉴文献[1]的实验方法, 通过编程由计算机自动进行模拟。实验在 CPU Pentium4 2.66 G, 内存 512 M, VC++ 6.0 环境下进行。

### 2.1 图像数据库和特征

实验采用 Corel 图像的一个子集, 共 50 个类别, 每个类别包含 100 幅语义相似的图像。实验时先对每幅图像进行特征提取, 并保存到特征库。本文采用下述 3 种特征:

1) 颜色特征, 即 80 维的主颜色描述子<sup>[6]</sup>;

2) 纹理特征。先通过金字塔小波变换<sup>[5]</sup>将图像变换到  $YCbCr$  颜色空间, 并对 Y 分量进行多分辨率 2 维 Haar 小波变换, 层数为 3, 共得到 10 个子带, 然后分别计算每个子带的均值和方差, 构成 20 维的向量;

3) 形状特征。采用 Manjunath 等人提出的局部边缘方向直方图<sup>[9]</sup>来构成形状特征向量,向量的维数为 80。

为了与相关算法进行比较,本文还实现了其他两种 SVM 反馈系统,即 ABRS-SVM 算法<sup>[4]</sup>,FSVM 算法<sup>[6]</sup>。前者使用的特征是 256 维的 HSV 颜色直方图,20 维金字塔小波变换<sup>[5]</sup>,80 维的局部边缘方向直方图<sup>[9]</sup>;后者与本文算法使用相同的特征。

### 2.2 算法性能

对算法性能进行评价时,先从图像数据库随机选取一幅图像,然后在图像库中按欧氏距离判断相似性,并返回查询结果。设前 100 幅图像中相关图像数为  $N_R$ 。按  $N_R \leq 5, N_R \leq 10, N_R \leq 15$  将图像库中的图像分为 A, B, C 组。

#### 2.2.1 平均查准率比较

每轮反馈中用户标记的样本数为  $N = 20$  时,取阈值  $T = 5$ ,从 A 组图像随机选择 100 幅图像,分别在图像数据库中进行检索。取样本数  $N = 20$ ,本文算法 AB-FSVM 与 ABRS-SVM,FSVM 的平均查准率的比较如图 2 所示。同理,当取阈值  $T = 10, T = 15$ ,再分别从 B 组、C 组图像中随机选择 100 幅图像,分别在图像数据库中进行检索,其平均查准率( $N = 20$ )比较如图 3、图 4 所示。

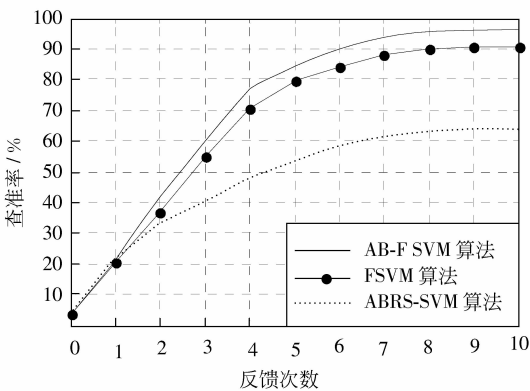


图 2 T = 5 平均查准率比较

Fig.2 Average retrieval precision comparison (T=5)

由图 3、图 4 可以看出,随着反馈次数的增加,本文算法优于其余两种算法,当取阈值  $T = 5$  时,优势更明显。前文的算法描述中,在正样本数有限的情况下,本文对负样本进行了非对称的打包处理,首先选择出最优的负样本包,与正样本一起参与训练,得出了最优的支持向量机模型。随着正样本数的增加,负样本对分类器的决定作用相对减弱,因此,当

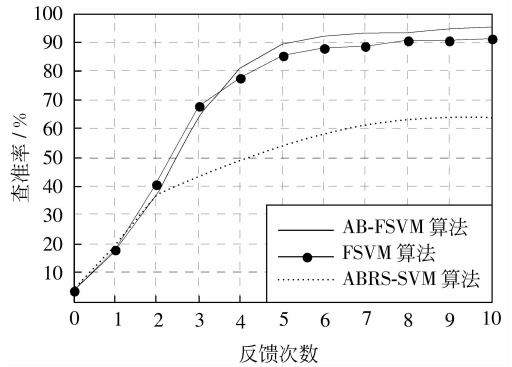


图 3 T = 10 平均查准率比较

Fig.3 Average retrieval precision comparison (T=10)

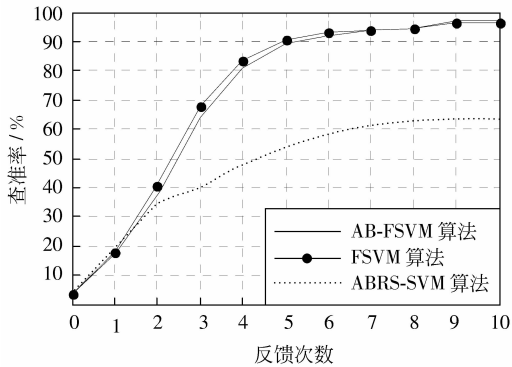


图 4 T = 15 平均查准率比较

Fig.4 Average retrieval precision comparison (T=15)

取  $T = 15$  时,本文算法 AB-FSVM 的查准率稍微优于 FSVM 算法;而当取阈值  $T = 5$  时,由于正样本数很小,负样本数对分类器的决定作用相对增强,所以本文算法 AB-FSVM 的查准率优于 FSVM 算法。由于 ABRS-SVM 算法没有采用模糊技术,没有利用到样本间的聚类信息,所以该算法的检索性能明显低于本文的算法及 FSVM 算法。

#### 2.2.2 查准率-查全率

随着反馈次数的增加,当  $T$  越小时,AB-FSVM 算法的检索性能越好。作为折中,本文取阈值  $T = 10$ ,设计了以下两组实验。第 1 组实验,用户标记  $N = 20$  个正样本。第 2 组实验,用户标记的样本数  $N = 40$  个正样本。另外,取反馈次数为 5 次,AB-FSVM 算法的查准率即可超过 85%。5 次迭代反馈后,算法的平均查准率-查全率曲线的比较如图 5、图 6 所示。

由图 5、图 6 查准率-查全率曲线可以看出以下几点: 1) 样本数  $N$  从 20 增加到 40,ABRS-SVM 算

法的查准率-查全率提升较少; 2) FSVM 算法的查准率-查全率提升较明显, 在查全率为 50% 处, 查准率可提高 22%; 3) AB-FSVM 算法的查准率-查全率始终优于 ABRS-SVM 算法和 FSVM 算法。由于 FSVM 算法及本文的算法都采用模糊技术, 且充分利用了样本间的聚类信息, 因此二者的查准率-查全率明显优于 ABRS-SVM 算法; 本文算法在查全率较小阶段, 查准率-查全率均优于 FSVM 算法, 随着查全率增加, 即随着正样本数的增加, 二者的查准率-查全率曲线逐渐接近。

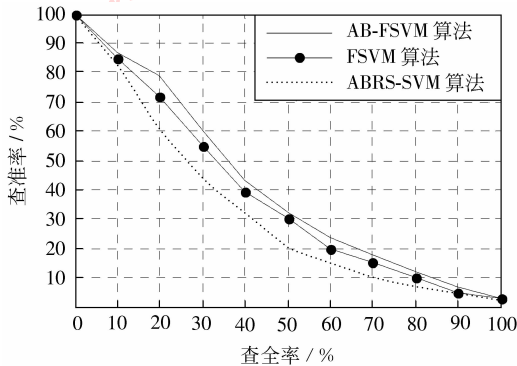


图 5 查准率-查全率曲线比较 ( $N = 20$ )

Fig. 5 Average precision vs. recall comparison ( $N = 20$ )

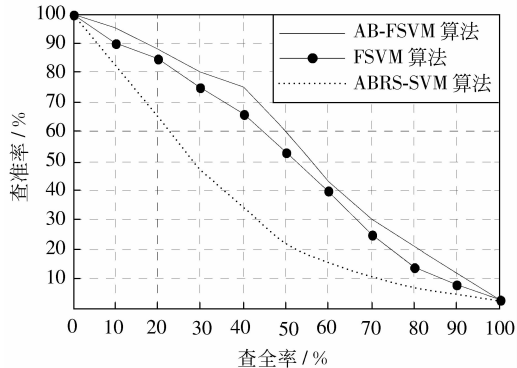


图 6 查准率-查全率曲线比较 ( $N = 40$ )

Fig. 6 Average precision vs. recall comparison ( $N = 40$ )

### 2.2.3 时间复杂度分析

若  $S$  为训练样本集, SVM 的训练时间  $t_T$  只跟训练样本数  $|S|$  有关。分类时间  $t_C$  由图像数、支持向量个数和特征向量维数共同决定。 $t_{fuzzy}$  为模糊聚类时间。非对称打包个数  $N_{AB}$  是常量, 即负样本被打包的个数。基于 SVM 算法的计算时间至少应包括训练时间和分类时间。对 ABRS-SVM 还要考虑提取空间特征子向量的耗费时间  $t_S$ ; 而对于 FSVM 算

法和本文的算法, 则都要考虑计算模糊因子的时间。理论上 3 种算法的时间复杂度如表 1 所示。

表 1 时间复杂度比较

Tab. 1 Comparison of time complexities

算法	ABRS-SVM 算法	FSVM 算法	AB-FSVM 算法
训练	$N_{AB} \cdot t_T$	$t_T$	$N_{AB} \cdot t_T, N_p \leq T$ $t_T, N_p > T$
分类	$N_{AB} \cdot t_C$	$t_C$	$t_C$
其他	$t_S$	$t_{fuzzy}$	$t_{fuzzy}$

在表 1 中,  $t_T$  由训练样本数  $|S|$  决定,  $t_C = O(|I| \times N_{SVM} \times D_{feature})$ , 其中  $|I|$  为图像数,  $N_{SVM}$  为支持向量个数,  $D_{feature}$  为特征向量维数。 $t_{fuzzy} = O((N_p + N_n)^2)$ , 其中  $N_p, N_n$  分别为正、负样本数。

从 B 组图片随机选择 100 幅图像, 在 5 000 幅的图像库中进行检索。当取阈值  $T = 10$  时, 每次反馈用户标记  $N = 20$ , 5 次迭代反馈后的 3 种算法的平均运行时间如表 2 所示, ABRS-SVM 算法, FSVM 算法及本文算法的平均运行时间比约为 15.5:1:1.51。

表 2 平均运行时间/s

Tab. 2 Average running time /s

ABRS-SVM 算法	FSVM 算法	AB-FSVM 算法
8.943	0.577	0.871

由表 2 可以看出, ABRS-SVM 算法训练、提取特征子向量和分类的时间都相当可观。本文算法由于将非对称打包与模糊理论相结合, 从而减少了分类检索的时间。但是与 FSVM 算法相比, 本文算法对负样本的非对称打包处理, 使训练时间有所增加。

## 3 结 论

基于 SVM 的相关反馈 CBIR 系统中, 可能面临以下 4 个问题: 1) 小样本问题; 2) 样本不对称问题; 3) 过度拟合问题; 4) 实时性问题。本文算法采用改进的非对称打包算法, 与 FSVM 算法有机地结合, 解决了前两个问题。另外, 采用 80 维的主颜色描述子, 有效地降低了特征向量的维数, 即缓解了第 3 个问题。本文算法与 ABRS-SVM 算法相比, 运算时间大幅减少; 与传统的 SVM 算法相比, 则略微增加了训练时间; 当正反馈样本数较少时, 虽然训练时间比 FSVM 算法有所增加, 但是由于此时正负样本数量较少, 因此检索性能得到明显提升。

## 参考文献 (References)

- [ 1 ] Tao Dacheng, Tang Xiaoou. Random sampling based SVM for relevance feedback image retrieval [ C ] // Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: IEEE Press, 2004:647-652.
- [ 2 ] Hoi S C H, Lyu M R, Jin R. A unified log-based relevance feedback scheme for image retrieval [ J ]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(4): 509-524.
- [ 3 ] Kim D H, Chung C W. Qcluster: relevance feedback using adaptive clustering for content-based image retrieval [ C ] // Proceedings of the ACM SIGMOD International Conference on Management of Data. San Diego, California, USA: ACM Press, 2003:599-610.
- [ 4 ] Tong S, Chang E. Support vector machine active learning for image retrieval [ C ] // Proceedings of the Ninth ACM International Conference on Multimedia. Ottawa, Canada: ACM Press, 2001: 107-118.
- [ 5 ] Tao Dacheng, Tang Xiaoou, Li X L, et al. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(7): 1088-1099.
- [ 6 ] Rui M, Cheng H D. Effective image retrieval using dominant color descriptor and fuzzy support vector machine [ J ]. Pattern Recognition, 2009, 42(1):147-157.
- [ 7 ] Lin Chunfu, Wang Shengde. Fuzzy support vector machines [ J ]. IEEE Transactions on Neural Networks, 2002, 13(2): 464-471.
- [ 8 ] Barbar D, Domeniconi C, Rogers J P. Detecting outliers using transduction and statistical testing [ C ] // Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Philadelphia, PA, USA: ACM Press, 2006: 55-64.
- [ 9 ] Manjunath B S, Ohm J R, Vasudevan V V, et al. Color and texture descriptors [ J ]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(6):703-715.