

DOI: 10.7652/xjtuxb201608016

一种判别极端学习的相关反馈图像检索方法

黄晓冬¹, 孙亮¹, 刘胜蓝²

(1. 解放军信息工程大学信息工程学院, 450001, 郑州;

2. 大连理工大学电子信息与电气工程学部, 116024, 大连)

摘要: 针对基于支持向量机(SVM)的相关反馈图像检索方法计算复杂度高、缺乏判别能力以及图像特征提取不充分的问题,提出一种基于判别极端学习的相关反馈图像检索(DELM)方法。在图像特征提取阶段,通过连接图像的颜色、纹理及边缘直方图实现图像的特征提取,解决了以往多数检索方法仅使用单一图像特征造成的图像描述不充分的问题;在检索的反馈阶段,将最大边际准则(MMC)引入到极端学习机中,通过分析极端学习机隐层空间的类内离散度和类间离散度得到包含判别信息的分类模型,并给出降维和不降维两种形式,以提高相关反馈图像检索系统的检索能力。DELM方法能有效应用于基于内容的图像检索中,并显著提高图像检索的性能。实验结果表明,DELM方法和采用SVM、ELM和最小类别方差ELM的方法相比,在Corel-1K数据集下检索平均准确率分别提高了11.06%、5.28%和6.40%。

关键词: 图像检索;相关反馈;支持向量机;极端学习机;判别信息

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2016)08-0096-07

A Retrieval Method of Relevance Feedback Images Based on Discriminative Extreme Learning

HUANG Xiaodong¹, SUN Liang¹, LIU Shenglan²

(1. Institute of Information System Engineering, PLA Information Engineering University, Zhengzhou 450001, China;

2. Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: A novel retrieval method of relevance feedback images based on discriminative extreme learning, named DELM, is proposed to concern the high computational complexity, low discriminant ability and insufficient image feature extraction of content-based image retrieval (CBIR) with the relevance feedback (RF) methods based on support vector machine (SVM). The proposed method extracts image features in the phase of image feature extraction through color, texture and edge histogram of the image to solve the problem that image feature extraction of the existing methods that based on single feature is insufficient. A maximum margin criterion (MMC) is introduced to extreme learning machine (ELM) in the phase of retrieval feedback. A classification model including discriminative information is obtained through analyzing discrete degrees within and between scatters of feature space in ELM hidden layer, and two versions of DELM are proposed to improve the retrieval performance of RF based image retrieval system, that is, a dimension reduction free based version and a dimension reduction based version. The

收稿日期: 2016-03-20。 作者简介: 黄晓冬(1991—),男,硕士生;孙亮(通信作者),男,教授。 基金项目: 国家自然科学基金资助项目(61372172)。

网络出版时间: 2016-06-28

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20160628.2026.002.html>

DELM method is effectively applied to CBIR, and significantly improves the quality of retrieval performance. Experimental results on Corel-1K dataset and comparisons with the methods using SVM, ELM, and minimum class variance ELM (MCVELM) show that the average retrieval precision of the DELM method increases by about 11.06%, 5.28% and 6.40%, respectively.

Keywords: image retrieval; relevance feedback; support vector machine; extreme learning machine; discriminative information

近年来,随着电子信息技术及互联网图像资源的迅速发展,人们对图像检索的需求越来越强烈。基于内容的图像检索方法(content-based image retrieval, CBIR)主要包含特征提取、高维特征向量降维、相似度匹配以及检索结果分类等几个环节,往往存在维度灾难和语义鸿沟等问题。目前主要通过维数约简方法来解决维度灾难问题,而针对语义鸿沟问题,目前主流的方法是采用相关反馈技术(relevance feedback, RF),利用人的主观认知能力提高 CBIR 系统检索的准确性。

区别于传统的 CBIR 只进行一次分类, RF-CBIR 系统需要根据反馈回来的信息进行再学习。目前已有大量的相关反馈方法用于改善图像检索的性能。He 等人提出一种半监督学习方法^[1],在相关反馈中较好地保持了查询图像的局部空间信息。为解决排序问题, Kundu 等人提出了一种基于图的相关反馈图像检索机制^[2],提高了检索准确率。此外,基于支持向量机(support vector machine, SVM)的主动学习算法也被用于 RF-CBIR 中,如半监督的 SVM 批处理主动学习算法^[3]、集成的 SVM 学习机^[4]等。然而,由于 RF-CBIR 中有标签的样本数量很少,因此基于 SVM 的算法很难获得令人满意的分类结果。

近年来,基于极端学习机(extreme learning machine, ELM)的方法在高维数据的处理分类中得到了广泛应用。Huang 等人证明了通过随机增加激活函数作为节点, SLFNs 可以近似于任意的连续目标函数^[5]。基于此,他们又提出了增量 ELM 算法^[6]和凸增量 ELM 算法^[7]。文献[8]提出的鲁棒的 ELM 算法改善了离群点问题对 ELM 算法精度的影响,但是没有充分利用未标记样本。文献[9]提出了一种基于图的半监督 ELM 算法,提高了 ELM 的泛化能力。除了上述关于 ELM 的理论改进之外,一些基于 ELM 方法的应用也相继提出。文献[10]提出一种鲁棒的激活函数(robust activation function, RAF),该方法将 Gaussian 核函数中的欧氏距离转化为余弦度量,避免了离群点对核函数的

影响,提高了 ELM 算法的性能,并应用于人脸识别中。文献[11]提出一种基于 ELM 的 RF-CBIR 系统框架,检索性能明显优于基于 SVM 的算法,但是该方法没有利用数据的类判别信息。为处理人体运动识别中的分类问题,最小类别方差 ELM 算法(minimum class variance ELM, MCVELM)^[12]将线性判决分析(linear discriminant analysis, LDA)引入到 ELM 隐层的特征值降维中,保证类内离散度最小来实现降维。基于此方法的启发,并考虑到 LDA 容易出现小样本问题,本文将最大边准则(maximum margin criterion, MMC)^[13]引入到 ELM 中,提出一种包含判别信息的 ELM 算法(discriminative ELM, DELM)。该算法通过分析 ELM 隐层特征空间中的类内离散度和类间离散度得到了一种判别分类模型,增强了分类器的判别能力。实验结果表明,本文算法能够有效应用于图像检索,取得了较好的检索结果,同时避免了小样本问题。

1 ELM 理论

对于一个含有 N 个输入节点、 $\tilde{N}$ 个隐层节点、 m 个输出节点的 SLFN 网络,给定 N 个不同的训练样本 $(\mathbf{x}_i, \mathbf{c}_i) \in \mathbf{R}^D \times \mathbf{R}^m$, 其中 $\mathbf{c}_i = [c_{i1}, c_{i2}, \dots, c_{im}]^T$ 。在分类问题中,隐层节点的个数 $\tilde{N}$ 通常都远大于类别数 m , 即 $\tilde{N} \gg m$ ^[14], 如图 1 所示。网络的每一个目标输出值 $\mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T$ 对应于一个样本 $\mathbf{x}_i$, 当样本属于一个类别 k 时, 即当 $c_i = k$ 时, $t_{ik} = 1$; 否则 $t_{ik} = -1$ 。对于一个给定的激活函数 $g(\cdot)$, SLFNs 网络的输出 $\mathbf{o}_i = [o_{i1}, o_{i2}, \dots, o_{im}]^T$ 可以通过下式计算

$$\mathbf{o}_i = \sum_{j=1}^{\tilde{N}} \boldsymbol{\beta}_j g(\mathbf{v}_j \mathbf{x}_i + b_j), \quad i = 1, 2, \dots, N \quad (1)$$

式中: $\mathbf{v}_j = [v_{j1}, v_{j2}, \dots, v_{jN}]^T$ 为连接网络第 j 个隐层节点和输入节点的输入权值; $\boldsymbol{\beta}_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jm}]^T$ 为连接网络第 j 个隐层节点和输出节点的输出权值; b_j 为第 j 个隐层节点的偏置值。

式(1)可以写成矩阵形式 $\mathbf{O} = \mathbf{G}\boldsymbol{\beta}$, 则网络隐层输出矩阵为

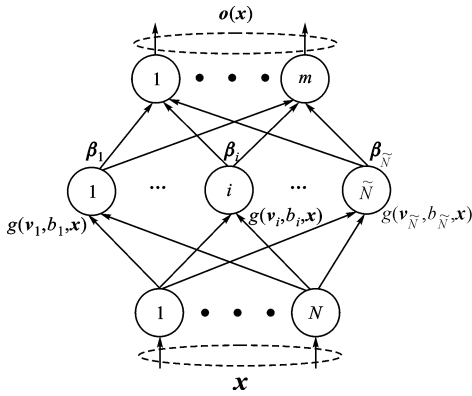


图1 单隐层前馈神经网络

$$G = \begin{bmatrix} g(v_1, b_1, x_1) & \cdots & g(v_N, b_N, x_1) \\ \vdots & & \vdots \\ g(v_1, b_1, x_N) & \cdots & g(v_N, b_N, x_N) \end{bmatrix} = \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_N) \end{bmatrix}_{N \times \tilde{N}} \quad (2)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_N^T \end{bmatrix}_{\tilde{N} \times m}; O = \begin{bmatrix} o_1^T \\ \vdots \\ o_N^T \end{bmatrix}_{N \times m} \quad (3)$$

假设网络预测输出 O 和网络目标输出 T 相等,

即 $\sum_{i=1}^{\tilde{N}} \|o_i - t_i\| = 0$, 其中 $i=1, 2, \dots, N$, 则式(1)可写成矩阵形式 $T = G\beta$, 其中 $T = [t_1^T, t_2^T, \dots, t_N^T]^T$ 是网络目标输出矩阵。

为了训练这样一个 SLFN, 希望能够找到一些合适的 $\hat{v}_j, \hat{b}_j$ 和 $\hat{\beta} (j=1, 2, \dots, \tilde{N})$, 使得

$$\|G(\hat{v}_1, \dots, \hat{v}_{\tilde{N}}, \hat{b}_1, \dots, \hat{b}_{\tilde{N}})\hat{\beta} - T\| = \min_{v_j, b_j, \beta} \|G(v_1, \dots, v_N, b_1, \dots, b_N)\beta - T\| \quad (4)$$

式(4)可以利用梯度下降算法求解。由式(4)可知, SLFNs 网络的解可通过线性系统 $T = G\beta$ 的最小二乘解得到 β 。

当网络输出节点选择一个线性激活函数时, β 可以通过 $\beta = G^+T$ 计算, 其中 $G^+ = (G^T G)^{-1} G^T$ 是 G^T 的 MP 广义逆矩阵。 G^+ 可以利用奇异值分解法和最小二乘法计算。同时, 为防止过拟合现象, 希望求解的权值尽可能小, 因此 ELM 的优化模型可以用下式表示

$$\min_{\beta} \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^N \|\xi_i\|^2$$

s. t. $g(x_i)\beta = t_i^T - \xi_i^T, i = 1, 2, \dots, N$ (5)

式中: $\xi_i = [\xi_{i,1}, \xi_{i,2}, \dots, \xi_{i,m}]^T$ 是 m 个输出节点关于训练样本 x_i 的训练误差向量。式(5)可以通过拉格

朗日方法转化为无约束最优化问题求解 β 。

ELM 算法可以概述如下:

- (1) 随机指定输入权值 v_j 和隐层节点的偏置值 $b_j, j=1, 2, \dots, \tilde{N}$;
- (2) 根据激活函数计算隐层输出矩阵 G ;
- (3) 计算网络输出权值 $\beta = G^+T$ 。

2 基于判别极端学习机的相关反馈图像检索算法

2.1 LDA 及 MMC 的优化模型

假定 ELM 隐层输出矩阵 $G = [g_1, g_2, \dots, g_N]^T$ 被分成如下形式 $G = [G_1, G_2, \dots, G_m]^T$, 其中 $G_i \in \mathbf{R}^{\tilde{N} \times N_i}$ 表示包含 N_i 个样本的第 i 类的实数矩阵, 满足 $N = \sum_{i=1}^m N_i, \mu_g^i = \frac{1}{N_i} \sum_{j=1}^{N_i} g_j^i$ 为第 i 类的均值向量, 矩阵 G 的全局均值向量为 $\mu_g = \frac{1}{N} \sum_{i=1}^N g_i$ 。线性投影矩阵 $W = [w_1, w_2, \dots, w_d] \in \mathbf{R}^{\tilde{N} \times d}$, 其中 d 为维数约简中的目标维数。LDA 的最优化过程可由如下的优化函数实现

$$W_{\text{opt}}^{\text{LDA}} = \arg \max_W \frac{\text{tr}(W^T S_b W)}{\text{tr}(W^T S_w W)} \quad \text{s. t. } W^T W = I \quad (6)$$

式中: I 为单位矩阵; S_b 为类间离散度矩阵, $S_b = \frac{1}{N} \sum_{i=1}^m N_i (\mu_g^i - \mu_g)(\mu_g^i - \mu_g)^T = H_b H_b^T$, 其中 $H_b = \frac{1}{N^{1/2}} [N_1^{1/2}(\mu_g^1 - \mu_g), \dots, N_m^{1/2}(\mu_g^m - \mu_g)]^T$; S_w 为类内离散度矩阵, $S_w = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{N_i} (g_j^i - \mu_g^i)(g_j^i - \mu_g^i)^T = H_w H_w^T$, 其中 $H_w = \frac{1}{N^{1/2}} [(g_1^1 - \mu_g^1), \dots, (g_{N_1}^1 - \mu_g^1), \dots, (g_{N_m}^m - \mu_g^m)]^T$ 。优化函数 $W_{\text{opt}}^{\text{LDA}}$ 可通过 $S_w^{-1} S_b$ 特征分解前 d 个最大特征值对应的特征向量得到。虽然 LDA 考虑同时满足类内离散度最小, 类间离散度最大来实现降维, 然而在实际应用中, LDA 经常会面临小样本问题, 即 $\tilde{N} \gg N$, 导致 S_w 不可逆。为解决该问题, MMC 在 LDA 的基础上构建新的优化函数来避免小样本问题。

MMC 中 2 个不同类的距离 $d(m^i, m^j)$ 定义为

$$d(m^i, m^j) = d(\mu_g^i, \mu_g^j) - (S(m^i) + S(m^j)) \quad (7)$$

式中: $d(\mu_g^i, \mu_g^j)$ 为 μ_g^i 和 μ_g^j 的距离; $S(m^i)$ 和 $S(m^j)$ 为每类样本的离散度矩阵的迹。根据式(6), MMC 的低维子空间 $W_{\text{opt}}^{\text{MMC}}$ 可通过如下的优化函数得到

$$\begin{aligned} \mathbf{W}_{\text{opt}}^{\text{LDA}} = \arg \max_{\mathbf{W}} \text{tr}(\mathbf{W}^T(\mathbf{S}_b - \mathbf{S}_w)\mathbf{W}) \\ \text{s. t. } \mathbf{W}^T\mathbf{W} = \mathbf{I} \end{aligned} \quad (8)$$

式中: $\mathbf{W}_{\text{opt}}^{\text{MMC}}$ 可通过 $\mathbf{S}_b - \mathbf{S}_w$ 特征分解前 d 个最大特征值对应的特征向量得到。

2.2 判别极端学习机(DELM)

传统的ELM直接利用隐层输出矩阵 $\mathbf{G}$ 计算输出权值 $\boldsymbol{\beta}$ 。实际上,在ELM的隐层特征空间中有许多非常重要的结构。MCVELM算法^[12]在ELM空间中只对训练数据的类内离散度矩阵求最小化,与此对应,本文同时考虑最小化类内离散度,最大化类间离散度来构造一个基于ELM的分类器。

当 $\tilde{N} \geq N$ 时,考虑到最小化类内离散度,最大化类间离散度,DELM方法可以转化为解决如下最优化问题

$$\begin{aligned} \min_{\boldsymbol{\beta}} \|\mathbf{H}_w\boldsymbol{\beta}\|_2^2 - \|\mathbf{H}_b\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_2^2 + C \sum_{i=1}^N \|\boldsymbol{\xi}_i\|_2^2 \\ \text{s. t. } \mathbf{g}_p = \mathbf{t}_i^T - \boldsymbol{\xi}_i^T, i=1,2,\dots,N \end{aligned} \quad (9)$$

式中: $\|\mathbf{H}_w\boldsymbol{\beta}\|_2^2$ 和 $\|\mathbf{H}_b\boldsymbol{\beta}\|_2^2$ 的目标是平衡ELM特征空间的分类结构; $\lambda \|\boldsymbol{\beta}\|_2^2$ 称为结构风险; $C \sum_{i=1}^N \|\boldsymbol{\xi}_i\|_2^2$ 称为经验风险,可使约束条件的错误最小,可消除过拟合问题。式(9)可通过拉格朗日乘子法求解,得到 $\boldsymbol{\beta}$ 的表达式如下

$$\boldsymbol{\beta} = \left(\frac{1}{C}(\mathbf{S}_w - \mathbf{S}_b) + \frac{1}{C\lambda}\mathbf{I} + \mathbf{G}^T\mathbf{G} \right)^{-1} \mathbf{G}^T\mathbf{T} \quad (10)$$

根据式(10)得到DELM分类器的输出函数为

$$\begin{aligned} f(\mathbf{x}) = \mathbf{g}(\mathbf{x})\boldsymbol{\beta} = \\ \mathbf{g}(\mathbf{x}) \left(\frac{1}{C}(\mathbf{S}_w - \mathbf{S}_b) + \frac{1}{C\lambda}\mathbf{I} + \mathbf{G}^T\mathbf{G} \right)^{-1} \mathbf{G}^T\mathbf{T} \end{aligned} \quad (11)$$

当 $\tilde{N} < N$ 时, $\mathbf{P} = \frac{1}{C}(\mathbf{S}_w - \mathbf{S}_b) + \frac{1}{C\lambda}\mathbf{I} + \mathbf{G}^T\mathbf{G}$ 可能会不可逆,为了保持判别信息并使 $\mathbf{P}$ 可逆,式(8)的降维方法可用于DELM分类器,因此类内离散度和类间离散度可分别改写为 $\hat{\mathbf{S}}_w = \mathbf{W}^T\mathbf{S}_w\mathbf{W}$ 和 $\hat{\mathbf{S}}_b = \mathbf{W}^T\mathbf{S}_b\mathbf{W}$ 。所以, $\boldsymbol{\beta}$ 可以通过下式得到

$$\boldsymbol{\beta} = \left[\mathbf{W}^T \left(\frac{1}{C}(\mathbf{S}_w - \mathbf{S}_b) + \frac{1}{C\lambda}\mathbf{I} + \mathbf{G}^T\mathbf{G} \right) \mathbf{W} \right]^{-1} \mathbf{W}^T\mathbf{G}^T\mathbf{T} \quad (12)$$

根据式(12)得到带降维的DELM分类器的输出函数为

$$\begin{aligned} f(\mathbf{x}) = \mathbf{g}(\mathbf{x})\boldsymbol{\beta} = \\ \mathbf{g}(\mathbf{x}) \left[\mathbf{W}^T \left(\frac{1}{C}(\mathbf{S}_w - \mathbf{S}_b) + \frac{1}{C\lambda}\mathbf{I} + \mathbf{G}^T\mathbf{G} \right) \mathbf{W} \right]^{-1} \mathbf{W}^T\mathbf{G}^T\mathbf{T} \end{aligned} \quad (13)$$

式中: $\mathbf{W}$ 可由式(8)计算得到。需要说明的是, $\tilde{N}$ 由

使用者决定,然后就可以通过 $\tilde{N} \geq N$ 或 $\tilde{N} < N$ 来选择式(11)的方法或者式(13)的方法。

2.3 基于DELM的反馈算法

在CBIR系统中基于SVM的相关反馈方法已经得到广泛应用,然而,当SVM中特征维数非常大时引起的过拟合问题却使系统性能不令人满意。此外,SVM需要产生大量的支持向量,导致分类的时间复杂度增加。为了解决上述问题,本文针对CBIR系统提出了基于DELM的相关反馈方法。

在RF-CBIR系统中,用户可以对上一步的检索图像进行标记。通过将检索图像作为训练样本再学习,可以得到一个基于DELM分类器的二分类学习算法。给定一个反复变换的样本集 $(\mathbf{x}_i, t_i)$, $\mathbf{x}_i \in \mathbf{R}^D$ 表示包含一幅图像的颜色、纹理、形状特征的特征向量, t_i 是 $\mathbf{x}_i$ 的标签, $t_i \in [1, -1]$,DELM分类器的输出函数为 $f(\mathbf{x}_i) = (f^1(\mathbf{x}_i), \dots, f^m(\mathbf{x}_i))$ 。然后,构造一个基于DELM算法的分类器,把剩余图像分成相关类和不相关类。基于此原理,经过几次反馈后,依据与查询图像的关联度将图像进行排序可以得到更好的结果。具体步骤如下:

(1)利用传统的图像检索算法对查询样例进行检索,得到前 N 幅图像;

(2)通过用户标注,将这 N 幅图像分为两类,相关的标记为正例集合 U^+ ,不相关的标记为负例集合 U^- ;

(3)准备训练样本集 $(\mathbf{x}_i, t_i)$,其中 $\mathbf{x}_i \in U^+ \cup U^-$, $t_i \in \begin{cases} 1, \mathbf{x}_i \in U^+ \\ -1, \mathbf{x}_i \in U^- \end{cases}$;

(4)利用式(11)或(13)构造基于DELM算法的分类器;

(5)计算数据集中每幅图像的输出向量,通过选择分类器 $f(\mathbf{x}_i)$ 中输出向量的第一个值得到与查询图像的相似距离,因此,每幅图像的得分 I_i 可以表示为 $I_i = f^1(\mathbf{x}_i)$;

(6)依据图像的得分进行排序,并返回最新的检索结果。

选择Sigmoid函数作为ELM的激活函数。上述算法过程与基于SVM的相关反馈框架^[15]十分相似,但是由于利用了ELM分类器,该算法比基于SVM的相关反馈要快几千倍,并且在完成相同的任务时性能表现更好。实验结果阐明了算法的有效性。图2给出了基于DELM的RF-CBIR系统工作流程图。

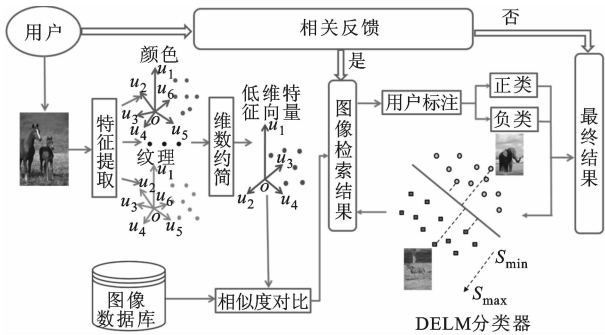


图 2 基于 DELM 的 RF-CBIR 系统工作流程图

3 实验结果及分析

3.1 实验设置

本实验采用的图像数据集为 Corel-1K、Corel-5K 和 Corel-10K。Corel-1K 包含 10 类总计 1 000 张图片,有汽车、恐龙、食物等等。Corel-5K 和 Corel-10K 数据集比 Corel-1K 更大。这 3 个数据集的基本信息如表 1 所示。

表 1 实验数据集分布

数据集	图片大小/像素	类别数	每类图片数	总数
Corel-1K	256×384	10	100	1 000
Corel-5K	128×192	50	100	5 000
Corel-10K	126×187	100	100	10 000

参数 C 用于平衡式(9)中的结构风险和经验风险,由于检索中有标签样本数量较少,因此结构风险比经验风险更加重要,这里设置 $C=0.05$;参数 λ 用于避免式(10)中最优矩阵不可逆,这里设置 $\lambda=1$,实际 λ 可以通过 $\lambda=\text{tr}(\mathbf{S}_w-\mathbf{S}_b)$ 设置。在 ELM 中,隐层节点数 $\tilde{N}$ 的设置也十分重要,具体细节可见文献[6-7],这里设置 $\tilde{N}=200$ 。训练样本数等于检索出的排在最前面的 N 幅图像,这里取 $N=20,30,40,50$ 。反馈(RF)的次数为 4 次,并分别表示为 RF1、RF2、RF3、RF4。FD1 表示一种特征提取方法,包含全局颜色直方图(GCH)、LBP^[16]纹理特征

表 3 $N=20$ 时采用 2 种特征对 4 种方法在 Corel-1K 数据集上的检索准确率比较

方法	$P/\%$		$P_1/\%$		$P_2/\%$		$P_3/\%$		$P_4/\%$	
	FD1	MSD	FD1	MSD	FD1	MSD	FD1	MSD	FD1	MSD
SVM	52.46	67.12	62.17	82.37	65.81	85.73	66.71	86.59	67.47	87.02
ELM	52.46	67.12	65.59	81.97	74.27	85.46	79.78	86.65	83.72	87.38
MCVELM	52.46	67.12	64.89	81.05	72.47	82.10	78.06	83.77	81.71	85.61
DELM	52.46	67.12	69.05	82.28	75.34	85.75	79.76	87.05	82.56	87.84

注: P 、 P_1 、 P_2 、 P_3 、 P_4 分别表示原始图像检索准确率和经过 1、2、3、4 次相关反馈后的检索准确率;黑体数据为比较最优值。

和边缘朝向直方图(EDH)。微观结构描述子(micro-structure descriptor, MSD)^[17] 利用边缘定位来提取图像的微观结构。实验采用的不同图像特征描述子的维数如表 2 所示。

表 2 图像特征描述子的维数

方法	GCH	LBP	EDH	MSD
维数	256	256	5	72

3.2 实验结果及分析

本实验的数据集中的所有图片都轮流作为查询图片。用于评价数据集的查准率和查全率的定义与文献[18]一致。

表 3 为 $N=20$ 时采用 4 种方法对 2 种特征在 Corel-1K 数据集上的检索准确率比较。通过表 3 可以看到,在 Corel-1K 数据集中,多数情况下 DELM 要比 SVM、ELM 和 MCVELM 的性能更好,这主要是因为 DELM 空间中图像的类别能保持较好的结构。本文发现一个比较有意思的结果是 MSD 的检索结果要优于 FD1,这说明一个好的特征描述子对提高图像检索性能十分重要,达到较高的查准率所需要的反馈次数也更少。由于基于 MSD 的检索结果非常相似,因此很难对基于 MSD 特征的相关反馈方法给进行评价,所以表 4 和表 5 只列出 FD1 特征的检索结果来评价不同的相关反馈算法。

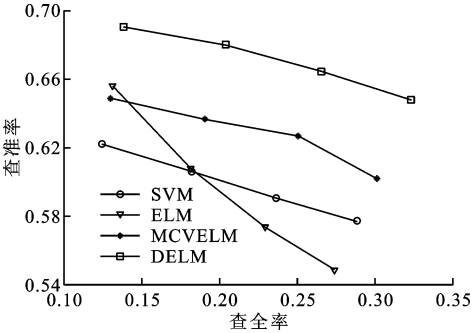
通过表 3~5 可以看到,第一次相关反馈的结果更利于评价相关反馈方法的性能。因此,在 N 为 20、30、40、50,反馈次数为 1 时,给出不同数据集上基于 FD1 特征的不同方法的查准-查全率曲线,如图 3 所示。从图 3 可以看出:由于查找范围与查全率是正相关的,因此随着查全率的增加,参与反馈的样本数增多,相关反馈效果越来越好,但是检索平均准确率降低,这是因为检索准确率和查全率成反比;DELM 的效率和 MCVELM 相似,主要原因是这 2 种方法在实现降维时都有着相似的处理过程。DELM 中降维的时间复杂度为 $O(D^3)$,这是比 ELM 多出的额外时间,因此 DELM 需要花费更多

表 4 $N=20$ 时采用 4 种方法对 FDI 特征在 Corel-5K 数据集上的检索准确率比较

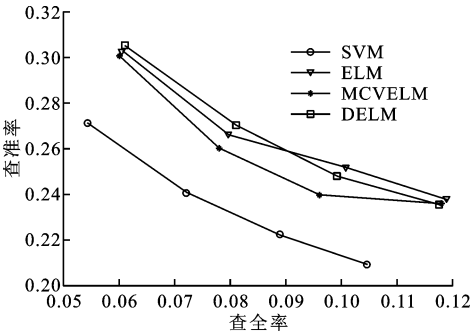
方法	$P/\%$	$P_1/\%$	$P_2/\%$	$P_3/\%$	$P_4/\%$
SVM	24.78	27.12	29.10	29.93	30.35
ELM	24.78	30.26	34.17	36.81	39.41
MCVELM	24.78	30.04	33.75	36.66	38.83
DELM	24.78	30.51	34.17	37.14	39.73

表 5 $N=20$ 时采用 4 种方法对 FDI 特征在 Corel-10K 数据集上的检索准确率比较

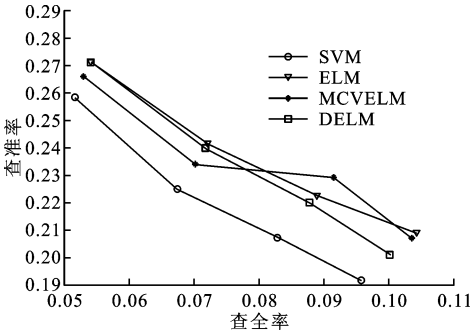
方法	$P/\%$	$P_1/\%$	$P_2/\%$	$P_3/\%$	$P_4/\%$
SVM	23.90	25.83	26.84	27.27	27.56
ELM	23.90	27.12	30.92	33.29	35.68
MCVELM	23.90	26.58	30.11	33.30	35.77
DELM	23.90	27.12	30.86	33.26	35.68



(a)Corel-1K



(b)Corel-5K



(c)Corel-10K

图 3 反馈次数为 1 时在不同数据集上基于 FDI 特征的不同方法的查准-查全率曲线

的时间来得到反馈函数,但是这一点在图像特征的维数不大(小于 500)时的影响非常小。

4 结 论

本文深入分析了 ELM 特征空间的分类结构,并针对 CBIR 系统的相关反馈提出了一种判别极端学习的相关反馈图像检索方法—DELM。该方法的主要优点如下:①DELM 包含了隐层特征空间的类内离散度和类间离散度,增强了基于 ELM 的相关反馈算法的判别能力;②DELM 可以根据用户需求选择是否包含降维方法,避免了小样本问题。基于标准数据集下的实验结果表明,ELM 的相关方法比基于 SVM 的方法更有效。此外,第一次的反馈结果也说明了图像特征的提取在图像检索过程中的重要性。

参考文献:

[1] HE X. Incremental semi-supervised subspace learning for image retrieval [C]// Proceedings of the 12th Annual ACM International Conference on Multimedia. New York, USA: ACM, 2004: 2-8.

[2] KUNDU M K, CHOWDHURY M, BULÒ S R. A graph-based relevance feedback mechanism in content-based image retrieval [J]. Knowledge-Based Systems, 2015, 73: 254-264.

[3] HOI S C H, JIN R, ZHU J, et al. Semi-supervised SVM batch mode active learning for image retrieval [C]// IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2008: 1-7.

[4] WANG X Y, CHEN J W, YANG H Y. A new integrated SVM classifiers for relevance feedback content-based image retrieval using EM parameter estimation [J]. Applied Soft Computing, 2011, 11(2): 2787-2804.

[5] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications [J]. Neurocomputing, 2006, 70(1): 489-501.

[6] HUANG G B, CHEN L, SIEW C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes [J]. IEEE Transactions on Neural Networks, 2006, 17(4): 879-892.

[7] HUANG G B, CHEN L. Convex incremental extreme learning machine [J]. Neurocomputing, 2007, 70(16): 3056-3062.

[8] HORATA P, CHIEWCHANWATTAN S, SUNAT

- K. Robust extreme learning machine [J]. Neurocomputing, 2013, 102: 31-44.
- [9] LIU S, FENG L, XIAO Y, et al. Robust activation function and its application: semisupervised kernel extreme learning method [J]. Neurocomputing, 2014, 144: 318-328.
- [10] 冯林, 刘胜蓝, 张晶, 等. 高维数据中鲁棒激活函数的极端学习机及线性降维 [J]. 计算机研究与发展, 2014, 51(6): 1331-1340.
- FENG Lin, LIU Shenglan, ZHANG Jing, et al. Robust activation function of extreme learning machine and linear dimensionality reduction in high-dimensional data [J]. Journal of Computer Research and Development, 2014, 51(6): 1331-1340.
- [11] LIU S, WANG H, WU J, et al. Incorporate extreme learning machine to content-based image retrieval with relevance feedback [C]//Proceeding of the 11th World Congress on Intelligent Control and Automation. Piscataway, NJ, USA: IEEE, 2014: 1010-1013.
- [12] IOSIFIDIS A, TEFAS A, PITAS I. Minimum class variance extreme learning machine for human action recognition [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(11): 1968-1979.
- [13] LI H, JIANG T, ZHANG K. Efficient and robust feature extraction by maximum margin criterion [J]. IEEE Transactions on Neural Networks, 2006, 17(1): 157-165.
- [14] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multiclass classification [J]. IEEE Transactions on Systems, Man, and Cybernetics: Part B Cybernetics, 2012, 42(2): 513-529.
- [15] 张磊, 林福宗. 基于支持向量机的相关反馈图像检索算法 [J]. 清华大学学报(自然科学版), 2002, 42(1): 80-83.
- ZHANG Lei, LIN Fuzong. Support vector machine based relevance feedback algorithm in image retrieval [J]. Journal of Tsinghua University (Science and Technology), 2002, 42(1): 80-83.
- [16] OJALA T, PIETIKÄINEN M, MÄENPÄÄ T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.
- [17] LIU G H, LI Z Y, ZHANG L, et al. Image retrieval based on micro-structure descriptor [J]. Pattern Recognition, 2011, 44(9): 2123-2133.
- [18] LIU G H, YANG J Y, LI Z Y. Content-based image retrieval using computational visual attention model [J]. Pattern Recognition, 2015, 48(8): 2554-2566.

[本刊相关文献链接]

- 周远, 周玉生, 刘权, 等. 一种适用于图像拼接的 DSIFT 算法研究. 2015, 49(9): 84-90. [doi:10.7652/xjtuxb201509015]
- 毛彦斌, 张选平, 杨晓刚. 伪 DNA 密码图像加密算法研究. 2015, 49(9): 91-98. [doi:10.7652/xjtuxb201509016]
- 刘凯, 张立民, 孙永威, 等. 利用深度玻尔兹曼机与典型相关分析的自动图像标注算法. 2015, 49(6): 33-38. [doi:10.7652/xjtuxb201506006]
- 符均, 牟沁沁, 季文博. 亮色分离的饱和图像校正方法. 2014, 48(10): 101-107. [doi:10.7652/xjtuxb201410016]
- 岳桂华, 滕奇志, 何小海, 等. 岩心三维图像修复算法. 2014, 48(9): 37-42. [doi:10.7652/xjtuxb201409007]
- 任茂栋, 梁晋, 唐正宗, 等. 数字图像相关法中的优化插值滤波器. 2014, 48(7): 65-70. [doi:10.7652/xjtuxb201407012]
- 靳峰, 冯大政. 利用空间序列描述子的快速准确的图像配准算法. 2014, 48(6): 19-24. [doi:10.7652/xjtuxb201406004]
- 张琳彦, 朱利. 以多幅图像非几何约束线段匹配重建建筑物外立面三维线段模型. 2014, 48(4): 15-19 + 25. [doi:10.7652/xjtuxb201404003]
- 袁飞, 朱利, 张磊. 利用超图图割的图像共分割算法. 2014, 48(2): 20-24. [doi:10.7652/xjtuxb201402004]
- 王芳梅, 范虹, YiWANG. 利用改进 CV 模型连续水平集算法的核磁共振乳腺图像分割. 2014, 48(2): 38-43. [doi:10.7652/xjtuxb201402007]
- 马元魁, 张树生, 白晓亮. 用球面混合曲率图像比较三角网格模型的相似性. 2012, 46(5): 97-101. [doi:10.7652/xjtuxb201205017]
- 史金钢, 齐春. 一种双约束稀疏模型图像修复算法. 2012, 46(2): 6-10. [doi:10.7652/xjtuxb201202002]
- 吴攀超, 王宗义, 林欣堂. 采用局部差分模型描述的彩色图像配准技术. 2011, 45(10): 30-37. [doi:10.7652/xjtuxb201110006]
- 徐胜军, 韩九强, 赵亮, 等. 用于图像分割的局部区域能量最小化算法. 2011, 45(8): 7-12. [doi:10.7652/xjtuxb201108002]
- 罗涛, 牟沁沁. 一种胸部 X 射线摄影图像中结节检测的多尺度匹配滤波器. 2011, 45(4): 30-35. [doi:10.7652/xjtuxb201104006]
- 赵海峰, 姚丽莎, 罗斌. 改进的人工鱼群算法和 Powell 法结合的医学图像配准. 2011, 45(4): 46-52. [doi:10.7652/xjtuxb201104009]
- 陈蓉, 孙剑, 徐宗本. 彩色图像分割中基于图上半监督学习算法研究. 2011, 45(2): 11-14. [doi:10.7652/xjtuxb201102003]

(编辑 刘杨)