

# 基于教与学优化算法的相关反馈图像检索

毕晓君, 潘铁文

(哈尔滨工程大学信息与通信工程学院, 黑龙江哈尔滨 150001)

**摘 要:** 为提高基于内容的图像检索的检索性能和检索速度, 克服低层视觉特征与高层语义概念间的“语义鸿沟”, 提出一种基于教与学优化的图像检索相关反馈算法 (TLBO-RF). 结合图像检索问题的特殊性和粒子群优化算法的优点, 对 TLBO 算法中个体的更新机制进行了改进, 通过将相关图像集的中心作为教师以及引入学员最好学习状态 Pbest, 使之朝用户感兴趣的相关图像区域快速收敛. 将该算法与目前效果最好的两种基于进化算法的相关反馈技术在两套标准图像测试集上进行对比, 结果表明本文算法相较于另外两种算法具有明显的优势, 不仅提高了图像检索性能, 同时也加快了图像检索速度, 更好地满足了用户的检索要求.

**关键词:** 基于内容的图像检索; 相关反馈; 教与学优化算法; 粒子群优化算法

**中图分类号:** TP391      **文献标识码:** A      **文章编号:** 0372-2112 (2017)07-1668-09

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2017.07.017

## Relevance Feedback Image Retrieval Based on Teaching-Learning-Based Optimization Algorithm

BI Xiao-jun, PAN Tie-wen

(College of Information and Communication Engineering, Harbin Engineering University, Harbin, Heilongjiang 150001, China)

**Abstract:** To improve the performance of image retrieval, and accelerate the speed of image retrieval in content-based image retrieval and reduce the “semantic gap” between visual low-level features and high-level semantic, relevance feedback image retrieval based on teaching-learning-based optimization algorithm is proposed (TLBO-RF). Considering the specificity of image retrieval and the advantage of the PSO, the update strategy of individual is modified in TLBO, the center of the relevant images is regarded as the teacher and the personal best is introduced, which makes the algorithm converge fast to the region of relevant images that the user is interested in. TLBO-RF is compared to two state-of-the-art RFs based on evolutionary algorithm on two benchmark images. The results show that TLBO-RF has obvious advantage in comparison with other two algorithms, not only increases the performance of image retrieval, but also improves the image retrieval speed, and can better meet the user needs of image retrieval.

**Key words:** content-based image retrieval; relevance feedback; TLBO; PSO

### 1 引言

近几年, 随着互联网技术和大数据技术的迅速发展, 产生了越来越多的图像数据, 如何快速从海量的图像库中检索出用户需要的信息成为一个非常亟待解决的问题<sup>[1]</sup>. 由于基于文本的图像检索方式不仅耗时长, 而且带有很强主观性, 因此使得基于内容的图像检索 (Content-Based Image Retrieval, CBIR) 方式逐渐成为主流<sup>[2-4]</sup>. 但是, 低层视觉特征与高层语义概念间存在“语义鸿沟”, CBIR 系统给出的检索结果往往不能很好

地满足用户的查询需求<sup>[5]</sup>. 为此, 引入相关反馈 (Relevance Feedback, RF) 技术到当前 CBIR 系统中, 通过在交互式进程中引入用户的偏好信息, 不断学习更新参数, 提高检索性能<sup>[6]</sup>. 但如今大多数基于 RF 技术的 CBIR 系统只注重于对用户标记的相关图像区域的开发, 而缺乏对特征空间未知区域的探索. 实际上, 最能提供信息的图像往往并不是相关区域附近的图像, 而是未知区域不能确定其相关性类别的图像<sup>[7]</sup>. 因此, 如何在对相关图像区域的开发与对未知区域的探索之间保持平衡成为图像检索相关反馈技术的关键<sup>[8]</sup>.

近期,基于进化算法(Evolutionary Algorithms)的相关反馈技术为两者之间的平衡提供了新的思路.由于进化算法本身就是在寻求全局探索与局部开发之间的平衡,与图像检索相关反馈技术中的平衡问题相似,两者结合的有效性已在文献[9]中得到证实. Mattia Broilo 等人<sup>[10]</sup>提出基于粒子群优化算法(Particle Swarm Optimization, PSO)的相关反馈技术,通过粒子在特征空间的运动来探索相关图像区域,并根据用户的反馈信息动态调整特征向量的权值,但是由于引入的分组策略,使得算法过分注重对特征空间未知区域的探索而相关图像区域开发不足. Lai 等人<sup>[11]</sup>提出基于交互式遗传算法的相关反馈算法,并采用新型的图像特征描述算子更好地体现用户的查询意图,尽管如此,由于遗传算法的交叉和变异操作,使得对相关区域过分开而未知区域的探索具有盲目性. Arevalillo-Herráez 等人<sup>[9]</sup>结合遗传算法和最近邻域法,并配以特征权重自适应和局部搜索策略,显著提高了 CBIR 系统的检索精度,但是却过多的引入了调节参数. Kanimozhi T 等人<sup>[12]</sup>提出将萤火虫优化算法与支持向量机相结合的相关反馈算法,利用萤火虫探索特征空间,用相关图像集与不相关图像集训练支持向量机的参数,取得了很好的检索结果但检索时间过长.

针对上述问题,本文提出了一种基于教与学优化的图像检索相关反馈算法(Relevance Feedback based on Teaching-Learning-Based Optimization algorithms, TLBO-RF). TLBO 算法<sup>[13]</sup>是 Rao 等人于 2011 年提出的一种新型群智能优化算法,该算法具有参数少、求解速度快和收敛能力强等优点,在数值优化和工程优化上都取得了很好的效果. 本文结合图像检索相关反馈问题的特殊性,对教与学优化算法进行专门的改进,旨在利用 TLBO 算法求解速度快和收敛能力强的特点,加快图像检索速度以及提高检索性能. 其创新点主要包括:①结合相关反馈技术的特点,选取相关图像区域的中心位置作为教阶段的教师以提高收敛性;②结合 PSO 算法的位置更新公式,在 TLBO 算法中引入学员最好学习状态 Pbest,以加强对未知区域的探索能力;③在教阶段学员的更新公式中增加自我学习因子一项,提高局部搜索能力.

## 2 RF 技术和 TLBO 算法

### 2.1 RF 技术

相关反馈技术首先是在文本检索领域提出的,于 1998 年由 Rui 等人引入到 CBIR 系统中. 在图像检索中,由于低层视觉特征与高层语义概念间存在一定间隔,初始的检索结果通常不能让用户满意. 而相关反馈技术可以根据用户对相关和不相关图像的标记,来构

造更好的查询表达或修改检索策略中的参数,更好地把握用户的查询需求<sup>[7,14]</sup>.

相关反馈算法根据检索模型的不同大致可以分为三种方法:基于概率框架的方法、基于机器学习的方法和基于距离度量的方法<sup>[7]</sup>. 其中,基于距离度量的方法由于模型简单、检索速度快而受到广泛应用. 它是把图像检索看作是一种  $k$  近邻搜索,图像的特征向量看作是特征空间中的一个点,通过距离大小来寻找靠近相关图像而远离不相关图像的点<sup>[7]</sup>. 本文方法在此基础上,将图像检索进一步建模成一个优化问题,从而可以通过优化算法的不断迭代寻找距离相关图像最近的图像<sup>[9]</sup>.

### 2.2 TLBO 算法

TLBO 算法模拟了班级学员向教师和其他学员学习的过程,分为教阶段和学阶段,具体步骤如下:

**步骤 1** 算法在解空间随机初始化种群  $\mathbf{P}^0 = [\mathbf{p}_1^0, \mathbf{p}_2^0, \dots, \mathbf{p}_N^0]$ , 其中  $N$  为种群规模,优化问题的解为  $\mathbf{P}_i^0 = [p_{i1}^0, p_{i2}^0, \dots, p_{id}^0]$ ,  $D$  为决策变量的维数.

**步骤 2** 教阶段:模拟教师通过“教学”使学生获取知识,提高班级学员的平均水平. 班级中的每个学员根据教师  $\mathbf{p}_{\text{teacher}}^t$  和学员的平均值  $\mathbf{p}_m^t$  之间的差异进行学习,学员更新公式如式(1)所示.

$$\begin{aligned} \mathbf{p}_i^{t+1} &= \mathbf{p}_i^t + r_i \times (\mathbf{p}_{\text{teacher}}^t - TF_i \times \mathbf{p}_m^t) \\ \mathbf{p}_m^t &= \frac{1}{N} \sum_{i=1}^N \mathbf{p}_i^t \end{aligned} \quad (1)$$

其中学习步长  $r_i = \text{rand}(0, 1)$ , 教学因子  $TF_i = \text{round}[1 + \text{rand}(0, 1)]$ ,  $\mathbf{p}_{\text{teacher}}^t$  为第  $t$  代种群中适应度值最高的个体,  $\mathbf{p}_m^t$  代表第  $t$  代种群中的平均水平个体.

学员更新之后进行学员选择操作,选择公式如式(2)所示.

$$\mathbf{p}_i^{t+1} = \begin{cases} \mathbf{p}_i^{t+1}, & f(\mathbf{p}_i^{t+1}) < f(\mathbf{p}_i^t) \\ \mathbf{p}_i^t, & \text{else} \end{cases} \quad (2)$$

**步骤 3** 学阶段:该阶段是模拟班级中学员互相交流的学习过程. 具体而言,第  $t$  代种群中,在经过教阶段后的种群中随机选取一个个体  $\mathbf{p}_j^t (i \neq j)$ , 通过分析两个学员之间成绩的差异,进行学习调整. 学员更新公式如式(3)所示.

$$\mathbf{p}_i^{t+1} = \begin{cases} \mathbf{p}_i^t + r_i \times (\mathbf{p}_i^t - \mathbf{p}_j^t), & f(\mathbf{p}_i^t) < f(\mathbf{p}_j^t) \\ \mathbf{p}_i^t + r_i \times (\mathbf{p}_j^t - \mathbf{p}_i^t), & f(\mathbf{p}_j^t) < f(\mathbf{p}_i^t) \end{cases} \quad (3)$$

其中  $r_i = \text{rand}(0, 1)$  表示第  $i$  个学员的学习因子.

学阶段的学员选择操作与教阶段一样.

**步骤 4** 如果满足结束条件,则优化结束,否则转至步骤 2 继续.

### 3 基于 TLBO 的图像检索 RF 算法 (TLBO-RF)

TLBO-RF 算法主要包括用户相关反馈和种群进化两个迭代过程,其整体流程图如图 1 所示。

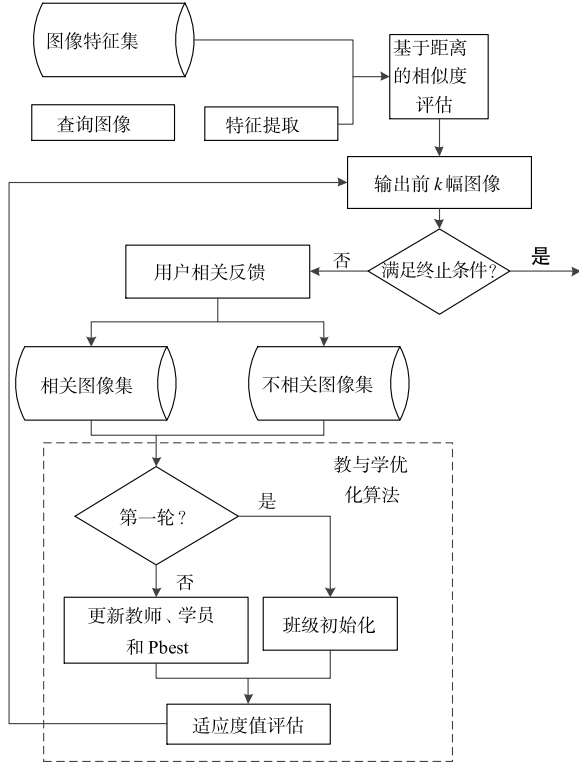


图1 基于TLBO的图像检索RF算法流程图

#### 3.1 特征描述和距离模型

为了显示 TLBO-RF 算法在克服“语义鸿沟”上的作用,本文采用颜色、形状和纹理三种最基本的视觉特征对图像进行描述,向量表示形式为  $\mathbf{x}_i = [\mathbf{x}_i^{cm}, \mathbf{x}_i^{ch}, \mathbf{x}_i^{tu}]$ , 其中  $\mathbf{x}_i^{cm}$  表示  $N_{cm}$  维的颜色矩特征向量,  $\mathbf{x}_i^{ch}$  表示  $N_{ch}$  维的颜色直方图特征向量,  $\mathbf{x}_i^{tu}$  表示  $N_{tu}$  维的 Tamura 纹理特征向量, 则一幅图像是由  $D = N_{cm} + N_{ch} + N_{tu}$  维向量来表征。当用户输入查询图像,其与图像集中图像的距离可表示为:

$$\text{Dist}(\mathbf{x}_q, \mathbf{x}_j) = E(\mathbf{x}_q^{cm}, \mathbf{x}_j^{cm}) + H(\mathbf{x}_q^{ch}, \mathbf{x}_j^{ch}) + H(\mathbf{x}_q^{tu}, \mathbf{x}_j^{tu}) + E(\mathbf{x}_q^{tu}, \mathbf{x}_j^{tu}) \quad (4)$$

其中  $E$  表示的是欧式距离,  $H$  表示直方图相交距离,具体公式如式(5)、式(6)所示。

$$E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \quad (5)$$

$$H(\mathbf{x}, \mathbf{y}) = \frac{\sum_{i=1}^N \min(x_i, y_i)}{\min(\sum_{i=1}^N x_i, \sum_{i=1}^N y_i)} \quad (6)$$

其中  $\mathbf{x}, \mathbf{y}$  表示图像特征向量,  $N$  表示对应特征向量的维数。

#### 3.2 种群初始化和适应度函数确定

如前所述,图像检索可以看作一个优化问题,本文采用 TLBO 算法求解。首先,我们需要在  $D$  维特征空间中定义种群个体  $\mathbf{p}_n$ , 这里选取图像集中距离查询图像最近的  $N_p$  个图像构成初始种群  $\mathbf{P}$ , 其种群大小  $N_p$  满足  $k \leq N_p < N_{DB}$ , 其中  $k$  表示每轮检索反馈给用户的图像数,  $N_{DB}$  表示图像集的大小。在进化算法中最重要的是定义适应度函数,综合考虑相关图像集 ( $X_{REL}$ ) 和不相关图像集 ( $X_{IRR}$ ), 期望适应度函数起到靠近相关图像而远离不相关图像的作用,这里采用文献 [10] 中的适应度函数:

$$F(\mathbf{p}_n) = \frac{1}{N_{rel}} \sum_{r=1}^{N_{rel}} \text{Dist}(\mathbf{p}_n; \mathbf{x}_r) + \frac{1}{\frac{1}{N_{irr}} \sum_{i=1}^{N_{irr}} \text{Dist}(\mathbf{p}_n; \mathbf{x}_i)} \quad (7)$$

其中  $\mathbf{x}_r \in X_{REL}, r = 1, 2, \dots, N_{rel}, \mathbf{x}_i \in X_{IRR}, i = 1, 2, \dots, N_{irr}$ 。

当个体越靠近相关图像集区域而远离不相关图像集区域时,适应度值越小,因此这是一个最小化问题。此外,由多维特征向量确定的特征空间是复杂不可预测的,其相关图像的位置分布也是不规律的,要求相应的 RF 技术在对已知相关区域具有一定开发能力的同时,对未知的相关区域也具有较强的探索能力,也即要求所采用的进化算法具有很好的全局收敛能力和局部搜索能力。而教与学优化算法具有快速收敛到全局最优的优点,但局部搜索能力不佳,为此本文对其进行改进,使之适合图像检索相关反馈问题,具体细节详见第 3.3 节。

#### 3.3 进化

为了加强 TLBO 算法的全局收敛能力和局部搜索能力,本文对学员的更新策略进行了改进。首先,为了更快地朝相关图像区域收敛,教师不再是每代种群中适应度值最高的个体,而是相关图像集区域的中心,由相关图像集的平均值得出:

$$\mathbf{p}_{teacher}^k = \frac{1}{N_{rel}} \sum_{r=1}^{N_{rel}} \mathbf{x}_r, \mathbf{x}_r \in X_{REL} \quad (8)$$

如果相关图像集为空集,教师则取查询图像位置。

其次,为了提高局部搜索能力,结合 PSO 算法的优点,在 TLBO 中引入每个学员的最好学习状态 Pbest, 添加到教阶段学员更新公式中,通过向自己的最好状态学习,提高学习成绩,改进后的更新公式如式(9)所示。

$$\mathbf{p}_i^{k+1} = \mathbf{p}_i^k + r_{i1} \times (\mathbf{p}_{teacher}^k - TF_i \times \mathbf{p}_i^k) + r_{i2} \times (\mathbf{l}_i^k - \mathbf{p}_i^k) \quad (9)$$

其中  $\mathbf{l}_i^k$  表示第  $k$  轮第  $i$  个学员的最好学习状态,  $r_{i1}, r_{i2} = \text{rand}(0, 1)$  表示学习步长。

教完成后,更新后的学员需要进行如下边界条件判定:

$$P_{ij}^{k+1} = \begin{cases} \text{Feature}_j^u, & P_{ij}^{k+1} \geq \text{Feature}_j^u \\ \text{Feature}_j^l, & P_{ij}^{k+1} \leq \text{Feature}_j^l \end{cases} \quad (10)$$

其中,  $P_{ij}^{k+1}$  表示第  $k+1$  轮第  $i$  个学员的第  $j$  维,  $\text{Feature}_j^u$ 、 $\text{Feature}_j^l$  分别表示特征空间第  $j$  维的上界和下界. 修正之后的学员不再进行学员选择操作,而是通过与最好学习状态进行对比,更新 Pbest,更新公式如式(11)所示.

$$l_i^{k+1} = \begin{cases} p_i^{k+1}, & f(p_i^{k+1}) < f(l_i^k) \\ l_i^k, & \text{else} \end{cases} \quad (11)$$

学阶段学员的更新公式与原始 TLBO 一样,但删除学员选择操作,变为学员最好学习状态 Pbest 的更新,更新公式与教阶段相同.

通过以上三方面的改进,TLBO 算法在 RF 中的作用将不仅仅是一种寻优工具,更是一种优良的空间探索工具,因为教、学阶段完成后,只是更新学员 Pbest 来引导下一轮学员的搜索,而更新后的学员会直接进入下一轮搜索,不需要进行原始 TLBO 算法中学员的选择操作,以期望学员通过教师和 Pbest 的引导,在特征空间中探索更多的相关图像.

### 3.4 匹配

由于图像的特征表示是在离散空间,而学员的移动是在连续空间,因此还需根据式(4)在学员的最小邻域内确定不重复的离散图像点(排除已被用户标记为不相关的图像点),即在对学员进行适应度值评估之后,需选取适应度值排名前个学员与图像库进行匹配,将幅图像输出给用户,经过相关和不相关判定后,进行种群的再进化.直至达到设定的反馈次数后,将所有的相关图像展示给用户.

### 3.5 TLBO-RF 图像检索原理分析

TLBO-RF 算法不同于以往基于单个查询点移动的方式,而是以多个查询点并行移动的方式探索更多的相关图像区域,使得对特征空间的探索能力更强,基于 TLBO-RF 的图像检索示意图如图 2 所示.

由图 2 可以看出,教阶段的  $p_i^{k+1}$  是通过把教师与班级平均水平个体的加权差向量和学员最好学习状态与当前学员的加权差向量加至基向量  $p_i^k$  上产生的,相当于在基向量上加了一个随机偏差扰动;学阶段的  $p_i^{k+1}$  则是将当前学员与其他学员的加权差向量加至基向量  $p_i^k$  上产生的,也相当于在基向量上加了一个随机偏差扰动.由于不存在教师与学员最好学习状态的牵引,使得算法的探索能力得到增强.在相关反馈前期,由于种群个体之间的差异较大,使得差向量对基向量的扰动较大,保证了算法前期对整个特征空间的探索能力,尽

可能探索到分散的相关图像区域,避免算法后期在一个相关图像区域开发.再加上学员最好学习状态 Pbest 和自我学习因子的引入,使得对未知区域的探索有了明确的方向,算法的局部搜索能力也得到一定增强.而随着反馈轮数的增加,种群个体之间的差异度减小,使得算法后期的开发能力和全局搜索能力加强,有助于加快收敛速度,且此时由于算法前期已将大部分相关图像区域探明,后期的开发将不再只针对一个相关区域.由此可见,TLBO-RF 算法前期注重对未知区域的探索,后期注重对相关图像区域的开发,有效地解决了图像检索相关反馈技术中开发相关图像区域和探索未知区域之间的平衡问题.

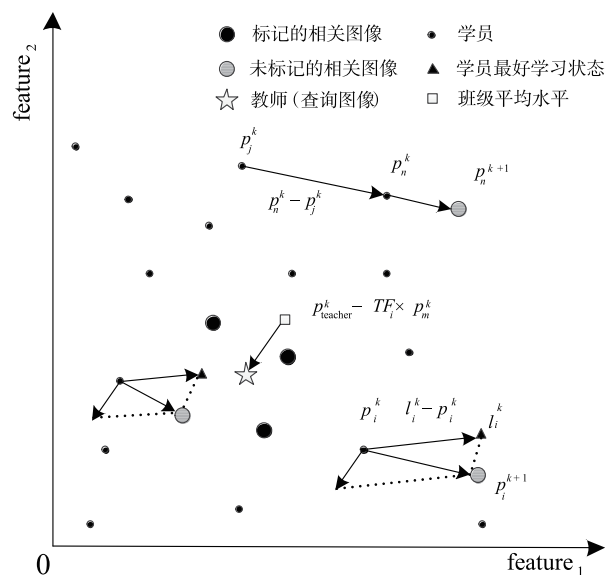


图2 基于TLBO-RF的图像检索示意图(特征空间为二维空间)

## 4 实验及结果分析

### 4.1 图片库及参数设置

#### 4.1.1 图片库

为了验证本文算法的有效性,本文选取在 SIMPLicity 标准图像集<sup>[15]</sup>和 Caltech-256 标准图像集<sup>[16]</sup>上进行测试.其中 SIMPLicity 图像集是 Corel 图像库的子集,共包含 10 个类别 1000 幅图像,每个类别均包含 100 幅图像;Caltech-256 图像集共包含 257 类 30607 幅图像.本文选取整个图像集的前 50 类,共包含 5617 幅图像(不包含背景图像和灰度图像),每类包含从 70 到 278 幅图像不等.

#### 4.1.2 视觉特征

每幅图像均是由 4 种特征构成的 55 维综合特征向量来表征:第一种特征是从 HSV 颜色空间中提取的 9 位颜色距;第二种特征是在 HSV 颜色空间下的 32 位颜色直方图(H 量化成 8 级,S 量化成 4 级);第三种特征

是 8 位边界方向直方图;第四种特征是 6 位 Tamura 纹理特征. 所有的特征向量都根据文献[17]归一化在  $[0, 1]$  之间.

#### 4.1.3 对比算法

为了验证本文算法的先进性,选取目前效果最好的两种基于进化算法的相关反馈技术作为对比算法,具体参数设置如下.

基于遗传算法的图像检索相关反馈 GA-RF<sup>[9]</sup>:交叉算子采用 SBX 算子,交叉概率  $p_c = 1$ ,  $\eta_c = 20$ ,无变异算子.

基于粒子群优化算法的图像检索相关反馈 PSO-RF<sup>[10]</sup>:  $c_1 = c_2 = 2$ ,惯性权重采用线性下降法,其中  $w_{\max} = 0.7$ ,  $w_{\min} = 0.2$ .

为了实验的公平性,三种算法的种群大小均为  $N = 100$ ,反馈轮数  $\text{Iter} = 10$ ,每轮反馈图像数均为  $k = 20$ .检索系统在 Intel Pentium、4G 内存、2.6GHz 主频,win7 64 位操作系统的计算机上运行,程序采用 MATLAB R2011b 编写.

#### 4.1.4 评价指标

为了评估本文算法的有效性,选取信息检索系统中常用的三种指标:查准率、查全率和综合评价指标  $F$  值,其公式分别如式(12)、式(13)和式(14)所示.

$$\text{Precision} = \frac{N_{R(q)}}{N(q)} \quad (12)$$

$$\text{Recall} = \frac{N_{TR(q)}}{N_{T(q)}} \quad (13)$$

$$F = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

式中:  $N_{R(q)}$  表示每轮检索出的相关图像个数;  $N(q)$  表示每轮反馈给用户的图像个数;  $N_{TR(q)}$  表示总共检索出的相关图像个数;  $N_{T(q)}$  表示图像库中所有相关图像个数.

## 4.2 实验结果及分析

### 4.2.1 检索性能分析

为了获得可靠的数据,从 SIMPLicity 图像集的每一类图像中按间隔 5 均匀选取 20 幅图像,组成 200 幅的图像查询集;从 Caltech-256 图像集中按间隔 11 均匀选取 500 幅图像作为查询集. 由于进化算法是随机算法,每次查询重复 5 次,取平均值作为每次查询的统计输出结果,最后 200(500)次查询的结果再取平均值作为最终的输出结果. SIMPLicity 图像集和 Caltech-256 图像集上 10 轮反馈查准率、查全率和  $F$  值的平均结果分别如图 3 和图 4 所示.

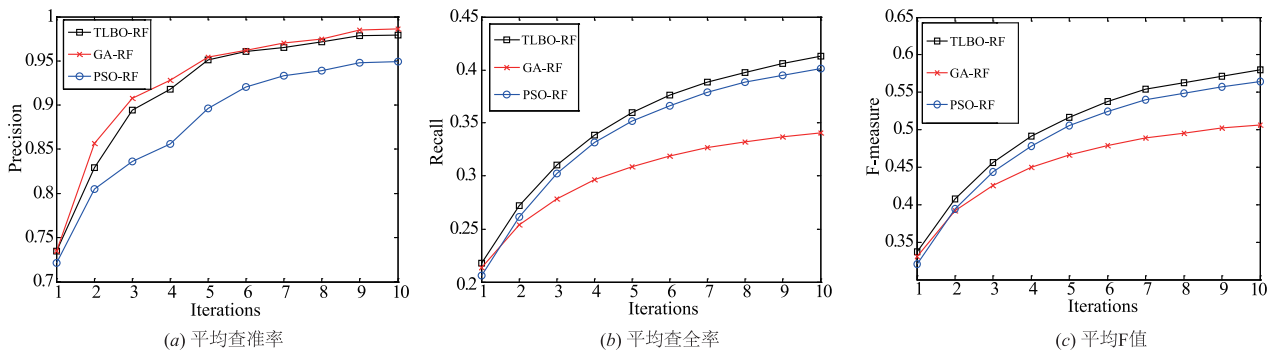


图3 SIMPLicity图像集上10轮反馈的平均查准率、平均查全率和平均F值

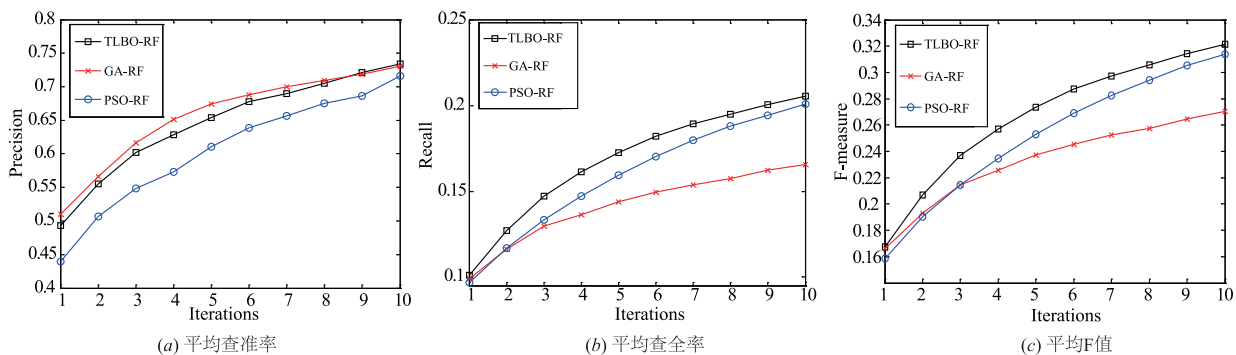


图4 Caltech-256图像集上10轮反馈的平均查准率、平均查全率和平均F值

从图 3(a) 中可以看出, TLBO-RF 算法得到的每轮查准率稍差于 GA-RF 算法, 但明显好于 PSO-RF 算法,

且结合图 3(b) 的查全率结果可以看出, TLBO-RF 算法在保证查准率的同时, 对特征空间的探索能力更强, 能够找到更多的相关图像, 提高查全率, 而 GA-RF 算法虽然能得到很好的查准率, 但查全率却不是理想, 到后期查全率曲线几乎水平, 说明该算法陷入局部最优, 无法找到更多的相关图像; PSO-RF 算法虽得到的查全率与 TLBO-RF 算法相近, 但查准率却是最差的. 最后从图 3(c) 的综合指标 F 值来看, TLBO-RF 算法在 SIMPLicity 标准图像集上的效果也是优于 PSO-RF 算法和 GA-RF 算法.

从图 4(a) 中可以看出, 随着图像集规模的增大, TLBO-RF 算法的总体查准率还是稍差于 GA-RF 算法, 但依旧保持着最优查全率, PSO-RF 算法的查全率与 TLBO-RF 算法相近, 但在查准率上还是劣于另外两种算法. 最后从图 4(c) 的综合指标 F 值也能看出, TLBO-RF 算法在 Caltech-256 标准图像集上的效果还是优于 PSO-RF 算法和 GA-RF 算法. 究其原因, 本文算法不旨在特征空间中找到最优解, 而是通过学员的位置更新在特征空间中探索更广泛的区域, 寻找更多相关图像点, GA 算法本身易陷入局部最优, 这会导致每轮搜索到的相关图像会重复出现在下一代, 阻碍算法搜索新的相关图像, 从图 3(a) 和图 4(a) 中 GA-RF 算法查全率偏低就能说明这个问题. 相比之下, TLBO-RF 算法会在寻优同时具备更好的空间探索能力, 会搜索出更多新的相关图像. 文献[9]的 PSO-RF 算法虽然也具备一定的空间探索能力, 但很显然是低于 TLBO-RF 算法的.

表 1 给出的是三种算法在 SIMPLicity 图像集的不同类别图片集中 10 轮反馈平均查全率的统计结果.

由于查准率反映的是每轮检索出的相关图像数与反馈给用户的图像数之比, 在很多情况下, 上一轮检索到的相关图像会重复出现在下一轮中, 导致查准率很高, 但相关图像的数量却没有明显增加, 因此查准率往往无法真实反映算法搜索相关图像的能力. 所以本文采用查全率来比较 3 种算法的性能优劣. 从表 1 中可以看出, TLBO-RF 算法在海滩、恐龙、马和食物这 4 类图像上的查全率明显优于 PSO-RF 算法和 GA-RF 算法, 虽然在建筑和山脉两类图像上前二轮查全率差于 PSO-RF 算法以及在公共汽车、花和山脉这三类图像前 3 轮差于 GA-RF 算法. 但是随着反馈次数的增加, TLBO-RF 算法搜索到的相关图像越来越多, 而 GA-RF 算法和 PSO-RF 算法却很少再能找到新的相关图像, 陷入局部最优, 特别是在恐龙这类图像上, 从第一轮开始就停滞, 陷入局部最优. 这再次说明本文提出的 TLBO-RF 算法具备更好的空间探索能力, 能找到更多的相关图像.

表 1 3 种算法在 10 个不同类别的图片集中 10 轮反馈平均查全率的统计结果

类别	算法	反馈轮数									
		1	2	3	4	5	6	7	8	9	10
非洲	TLBO	0.21	<b>0.31</b>	<b>0.36</b>	<b>0.38</b>	<b>0.39</b>	<b>0.39</b>	<b>0.40</b>	<b>0.40</b>	<b>0.41</b>	<b>0.42</b>
	PSO	<b>0.23</b>	0.30	0.34	0.36	0.37	0.38	0.39	0.39	0.40	0.41
	GA	0.23	0.29	0.32	0.34	0.36	0.36	0.37	0.37	0.37	0.38
海滩	TLBO	<b>0.16</b>	<b>0.23</b>	<b>0.29</b>	<b>0.32</b>	<b>0.35</b>	<b>0.37</b>	<b>0.38</b>	<b>0.39</b>	<b>0.41</b>	<b>0.42</b>
	PSO	0.15	0.20	0.23	0.26	0.28	0.29	0.31	0.32	0.33	0.34
	GA	0.14	0.18	0.21	0.23	0.23	0.24	0.25	0.25	0.25	0.26
建筑	TLBO	0.19	0.25	<b>0.30</b>	<b>0.33</b>	<b>0.34</b>	<b>0.36</b>	<b>0.37</b>	<b>0.39</b>	<b>0.40</b>	<b>0.41</b>
	PSO	<b>0.19</b>	<b>0.26</b>	0.29	0.32	0.34	0.36	0.37	0.38	0.39	0.40
	GA	0.18	0.23	0.26	0.28	0.29	0.30	0.30	0.31	0.31	0.31
公共汽车	TLBO	0.25	0.32	0.36	<b>0.40</b>	<b>0.43</b>	<b>0.46</b>	<b>0.48</b>	<b>0.50</b>	<b>0.51</b>	<b>0.53</b>
	PSO	0.26	0.32	0.36	0.39	0.41	0.41	0.43	0.43	0.44	0.44
	GA	<b>0.28</b>	<b>0.35</b>	<b>0.38</b>	0.39	0.40	0.40	0.40	0.40	0.40	0.40
恐龙	TLBO	<b>0.23</b>	<b>0.34</b>	<b>0.39</b>	<b>0.41</b>	<b>0.43</b>	<b>0.44</b>	<b>0.45</b>	<b>0.45</b>	<b>0.45</b>	<b>0.45</b>
	PSO	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
	GA	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
大象	TLBO	<b>0.19</b>	<b>0.25</b>	<b>0.29</b>	<b>0.31</b>	<b>0.33</b>	<b>0.35</b>	<b>0.36</b>	<b>0.38</b>	<b>0.39</b>	<b>0.39</b>
	PSO	0.19	0.25	0.28	0.30	0.31	0.32	0.33	0.33	0.34	0.35
	GA	0.19	0.25	0.28	0.29	0.30	0.30	0.30	0.31	0.31	0.31
花	TLBO	0.21	0.26	0.29	0.31	<b>0.34</b>	<b>0.35</b>	<b>0.37</b>	<b>0.39</b>	<b>0.40</b>	<b>0.41</b>
	PSO	0.22	0.26	0.29	0.32	0.33	0.35	0.36	0.36	0.37	0.38
	GA	<b>0.22</b>	<b>0.27</b>	<b>0.29</b>	0.31	0.31	0.32	0.32	0.33	0.33	0.33
马	TLBO	<b>0.28</b>	<b>0.37</b>	<b>0.40</b>	<b>0.42</b>	<b>0.44</b>	<b>0.45</b>	<b>0.46</b>	<b>0.47</b>	<b>0.47</b>	<b>0.47</b>
	PSO	0.26	0.31	0.33	0.34	0.35	0.35	0.36	0.36	0.37	0.37
	GA	0.26	0.30	0.32	0.33	0.34	0.35	0.35	0.35	0.36	0.36
山脉	TLBO	0.20	0.27	0.32	<b>0.34</b>	<b>0.37</b>	<b>0.38</b>	<b>0.39</b>	<b>0.40</b>	<b>0.41</b>	<b>0.42</b>
	PSO	<b>0.21</b>	<b>0.28</b>	<b>0.32</b>	0.33	0.36	0.37	0.39	0.39	0.40	0.40
	GA	0.18	0.23	0.28	0.29	0.30	0.31	0.31	0.32	0.32	0.32
食物	TLBO	0.26	0.32	<b>0.37</b>	<b>0.40</b>	<b>0.42</b>	<b>0.45</b>	<b>0.46</b>	<b>0.47</b>	<b>0.49</b>	<b>0.50</b>
	PSO	0.25	0.31	0.34	0.36	0.37	0.37	0.38	0.39	0.39	0.40
	GA	<b>0.27</b>	<b>0.33</b>	0.35	0.37	0.37	0.37	0.37	0.37	0.37	0.37

#### 4.2.2 检索结果分析

为了更深入地理解 TLBO-RF 的检索机制, 本文给出 TLBO-RF 在 SIMPLicity 图像集的 10 个类别上 10 轮反馈的平均查准率, 按照 4.2.1 节设置方法, 其统计结果如表 2 所示.

从表 2 中可以看出: 在非洲、海滩、建筑和山脉类别上, 后 4 轮平均查准率不是逐步提高, 而是会出现回落的情况, 正如 4.2.1 节所述, TLBO-RF 中的学员

是通过教师和自身最好学习状态  $P_{best}$  的引导来不断更新自身位置,以期在特征空间中探索到更多新的相关图像点,因此并不能保证平均查准率稳步上升,即便如此,TLBO-RF 在 10 类图像集上最后一轮的平均查准率都能在 90% 以上,尤其是在恐龙、花和马的类别上达到了 100%. 另外,通过对比不同类别图像集间的平均查准率可知,海滩相对于其他类别的图像集来说,其查准率稍微偏低,可能是因为海滩图像集在语义上比较模糊,有的图片不只是有海滩,还含有山

脉或是建筑,而本文为了突显 TLBO-RF 在克服“语义鸿沟”上的作用,只采用了颜色、形状和纹理三种最基本的视觉特征,其在语义明确的图像集上都能取得很好的结果,但对语义模糊的图像集(比如海滩)检索效果略差,需要进一步结合其他特征算子加以区分. 由于篇幅有限,只选取含有山脉的海滩图像(NO. 157)作为检索示例说明上述情况,检索结果如图 5 所示. 其中,绿色边框代表相关图像,红色边框代表不相关图像.



(a) 基于底层视觉特征的反馈结果 (查准率50%)



(b) 基于TLBO-RF的第10轮反馈结果 (查准率90%)

图5 海滩 (NO.157) 的检索结果

表 2 TLBO-RF 在 10 个不同类别的图片集中  
10 轮反馈平均查准率的统计结果

类别	反馈轮数									
	1	2	3	4	5	6	7	8	9	10
非洲	0.68	0.80	0.83	0.86	0.90	0.92	0.95	0.94	0.95	0.97
海滩	0.50	0.62	0.71	0.77	0.84	0.88	0.92	0.93	0.92	0.92
建筑	0.66	0.71	0.81	0.87	0.88	0.93	0.98	0.97	0.95	0.96
汽车	0.80	0.81	0.86	0.92	0.92	0.94	0.98	0.99	0.99	0.99
恐龙	0.95	1	1	1	1	1	1	1	1	1
大象	0.65	0.82	0.84	0.85	0.91	0.91	0.95	0.95	0.95	0.98
花	0.80	0.83	0.84	0.89	0.89	0.93	0.95	0.96	0.97	1
马	0.97	0.93	0.98	0.99	0.97	0.99	0.99	1	1	1
山脉	0.58	0.7	0.82	0.89	0.88	0.91	0.94	0.92	0.95	0.95
食物	0.84	0.92	0.93	0.97	0.98	0.99	0.99	0.99	0.99	0.99

从图 5 中可以看出:基于低层视觉特征相似度计算的检索结果只有 50%,不相关图像全是山脉,说明低层视觉特征与高层语义概念间存在很大鸿沟;经过 TLBO-RF 的 10 轮相关反馈,查准率提高到 90%,但由于查询图像中含有山脉的语义,而三种最基本视觉特征的语义区分能力还不是很强,导致检索结果中还是存在两幅山脉图像。

#### 4.2.3 检索速度分析

将本文算法与另外两种算法的检索速度作对比,分析三种算法在两套图像测试集上的检索效率.检索时间为各自图像集上前 100 幅图像作为查询图像时所得的平均时间,代表一次查询中 10 轮相关反馈所需的时间,如表 3 所示.从表 3 中可以看出,TLBO-RF 算法在检索速度上高于另外两种算法,其原因可能在于 GA-RF 算法需要在每轮反馈中进行最近邻样本的判断,而 PSO-RF 算法除了分组策略需要花费时间,其根据用户的反馈信息动态调整每一维特征向量的权值也需要花费不少时间.TLBO-RF 算法虽然没有分组策略和特征向量权值自适应调整,但由于 TLBO 自身优势以及引入的自我学习因子却能达到同样的效果,且带来的时间开销也很少。

表 3 三种算法的检索速度比较(单位:s)

算法	GA-RF 算法	PSO-RF 算法	TLBO-RF 算法
SIMPLicity 图像集	16.3	20.9	<b>10.5</b>
Caltech-256 图像集	89.4	116.5	<b>55.2</b>

## 5 结论

针对现有基于内容的图像检索相关反馈算法存在检索性能低、速度慢等问题,提出一种基于教与学优化的图像检索相关反馈算法 TLBO-RF.首先为了使 TLBO

算法快速收敛到相关图像区域,结合用户相关反馈信息,对教阶段教师的选择进行了改进;其次为了提高 TLBO 算法的空间探索能力,结合 PSO 算法的优点,引入学员最好学习状态 Pbest,通过不断更新 Pbest 来指导学员在特征空间中的探索,加强对未知区域的探索能力;最后为了提高算法局部搜索能力,对教阶段学员的更新公式进行了改进,增加了自我学习因子.在 SIMPLicity 和 Caltech-256 标准图像测试集上的实验结果表明,相对于同样基于进化算法的 GA-RF 和 PSO-RF 算法,TLBO-RF 算法在保证查准率的前提下明显提升查全率,使得图像检索的性能得到提升,同时加快了图像检索速度,更好地满足了用户的检索需求.接下来的研究工作将研究在 TLBO-RF 算法中结合更具表达能力的特征算子,比如 SIFT 特征,以提升其检索性能,并在大规模图像集上进行检索验证。

#### 参考文献

- [1] KUNDU Malay Kumar, CHOWDHURY Manish, BULò Samuel Rota. A graph-based relevance feedback mechanism in content-based image retrieval [J]. Knowledge-Based Systems, 2015, 73: 254 - 264.
- [2] 冯松鹤,郎丛妍,须德.一种融合图学习与区域显著性分析的图像检索算法[J].电子学报,2011,39(10):2288 - 2294.  
FENG Song-he, LANG Cong-yan, XU De. Combining graph learning and region saliency analysis for content-based image retrieval [J]. Acta Electronica Sinica, 2011, 39(10): 2288 - 2294. (in Chinese)
- [3] THOMEE Bart, LEW Michael S. Interactive search in image retrieval: a survey [J]. International Journal of Multimedia Information Retrieval, 2012, 1(2): 71 - 86.
- [4] 周燕,曾凡智,赵慧民,卢炎生,周月霞.一种基于精细化稀疏自适应匹配追踪算法的图像检索方法研究[J].电子学报,2014,42(12):2457 - 2466.  
ZHOU Yan, ZENG Fan-zhi, ZHAO Hui-min, et al. An image retrieval method based on meticulous sparsity adaptive matching pursuit algorithm [J]. Acta Electronica Sinica, 2014, 42(12): 2457 - 2466. (in Chinese)
- [5] LIU Ying, ZHANG Deng-sheng, LU Guo-jun, MA Wei-ying. A survey of content-based image retrieval with high-level semantics [J]. Pattern Recognition, 2007, 40(1): 262 - 282.
- [6] 许相莉,张利彪,刘向东,等.基于粒子群的图像检索相关反馈算法[J].电子学报,2010,38(8):1935 - 1940.  
XU Xiang-li, ZHANG Li-biao, LIU Xiang-dong, et al. Image retrieval relevance feedback algorithm based on particle swarm optimization [J]. Acta Electronica Sinica, 2010, 38(8): 1935 - 1940. (in Chinese)

- [7] 吴洪, 卢汉清, 马颂德. 基于内容图像检索中相关反馈技术的回顾[J]. 计算机学报, 2005, 28(12): 1969 – 1979.  
WU Hong, LU Han-qing, MA Song-de. A survey of relevance feedback techniques in content-based image retrieval [J]. Chinese Journal of Computers, 2005, 28(12): 1969 – 1979. (in Chinese)
- [8] AREVALILLO-HERRÁEZ Miguel, FERRI Francesc J, MORENO-PICOT Salvador. A hybrid multi-objective optimization algorithm for content based image retrieval [J]. Applied Soft Computing, 2013, 13(11): 4358 – 4369.
- [9] AREVALILLO-HERRÁEZ Miguel, FERRI Francesc J, MORENO-PICOT Salvador. Distance-based relevance feedback using a hybrid interactive genetic algorithm for image retrieval [J]. Applied Soft Computing, 2011, 11(2): 1782 – 1791.
- [10] BROILO Mattia, DE NATALE Francesco Gb. A stochastic approach to image retrieval using relevance feedback and particle swarm optimization [J]. IEEE Transactions on Multimedia, 2010, 12(4): 267 – 277.
- [11] LAI Chi-chi, CHEN Ying-chuan. A user-oriented image retrieval system based on interactive genetic algorithm [J]. IEEE Transactions on Instrumentation and Measurement, 2011, 60(10): 3318 – 3325.
- [12] KANIMOZHI T, LATHA K. An integrated approach to region based image retrieval using firefly algorithm and support vector machine [J]. Neurocomputing, 2015, 151: 1099 – 1111.
- [13] RAO Ravipudi V, SAVSANI Vimal J, VAKHARIA Dp. Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems [J]. Computer-Aided Design, 2011, 43(3): 303 – 315.
- [14] ZHANG Yong-dong, YANG Xiao-peng, MEI Tao. Image search reranking with query-dependent click-based relevance feedback [J]. IEEE Transactions on Image Processing, 2014, 23(10): 4448 – 4459.
- [15] WANG James Z, LI Jia, WIEDERHOLD Gio. SIMPLiCity: Semantics-sensitive integrated matching for picture libraries [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(9): 947 – 963.
- [16] GRIFFIN G, HOLUB A, Perona P. Caltech-256 Object Category Dataset (Technology Report 7694) [R]. California Institute of Technology, 2007.
- [17] AKSOY Selim, HARALICK Robert M. Feature normalization and likelihood-based similarity measures for image retrieval [J]. Pattern recognition letters, 2001, 22(5): 563 – 582.

#### 作者简介



毕晓君 女, 1964 年 11 月生于黑龙江省哈尔滨. 哈尔滨工程大学信息与通信工程学院教授、博士生导师. 主要研究方向为智能信息处理、图像处理.

E-mail: bixiaojun@hrbeu.edu.cn



潘铁文 男, 1984 年生于辽宁省鞍山市. 哈尔滨工程大学信息与通信工程学院博士研究生, 研究方向为进化算法、图像处理.

E-mail: pantiewen@hrbeu.edu.cn