

基于FSRM的相关反馈图像检索算法

郭士会 杨明 王晓芳 赵东方 邓世涛

(西南大学计算机与信息科学学院 重庆 400715)

摘要 在基于内容的图像检索中,低层视觉特征和高层语义之间的“语义鸿沟”一直是基于内容图像检索技术前进的一大障碍。相关反馈机制在一定程度上缩小了图像检索中的“语义鸿沟”。提出了一种基于模糊语义相关矩阵(FSRM)的相关反馈算法。该算法根据用户对检索结果的反馈调整模糊语义相关矩阵中的权值,从而捕捉用户的检索意图,通过对模糊语义相关矩阵中数据的学习不断修正语义矩阵,达到低层视觉特征到高层语义特征的过渡,最终提高了查询的准确度。实验结果证明了该算法的有效性。

关键词 基于内容的图像检索,语义鸿沟,相关反馈,模糊语义相关矩阵

中图法分类号 TP391

文献标识码 A

Relevance Feedback Algorithm Based on Fuzzy Semantic Relevance Matrix in Image Retrieval

GUO Shi-hui YANG Ming WANG Xiao-fang ZHAO Dong-fang DENG Shi-tao

(School of Computer and Information Science, Southwest University, Chongqing 400715, China)

Abstract The semantic gap between low level visual features and high level semantic concepts is an obstacle to the development of image retrieval. Relevance feedback techniques narrow the semantic gap to some extent. A relevance feedback algorithm was presented based on fuzzy semantic relevance matrix (FSRM). During the retrieval process, the weights in the FSRM are adjusted according to user's feedback and the FSRM are modified by learning more time. Experimental results show the effectiveness of the algorithm.

Keywords CBIR, Semantic gap, Relevance feedback, Fuzzy semantic relevance matrix

1 引言

自基于内容的图像检索(CBIR)提出以来,它对图像检索领域做出了很大的贡献,使我们摆脱了先前基于文本的图像检索中人工标注的繁琐性。但是随着多媒体数据的海量增加以及用户对检索要求的不断提高,基于低层视觉特征的图像检索和高层语义之间的“语义鸿沟”成为图像检索技术前进的一大障碍。

为了缩小“语义鸿沟”间的差距,人们提出了将机器学习、分类和聚类技术应用到图像检索的预处理和相关反馈中。其中相关反馈机制通过人机交互的方式让用户参与到图像检索中来,以更好地捕捉用户的检索意图,提高了检索的效率,从而得到了广泛关注。近年来很多相关反馈算法不断涌现^[1-6],它们都在一定程度上提高了图像检索的效率。目前人们把CBIR中的相关反馈技术大致分为两类^[7]:一类是利用用户的反馈信息调整相似度量中的一些参数;另一类则是从概率的观点出发,利用用户的反馈信息计算出图像库中每一幅图像符合用户要求的概率,并将概率高的图像作为检索结果返回给用户。

本文通过对图像之间存在的模糊语义建立相关矩阵^[8],提出了一种基于模糊语义相关矩阵(FSRM)的相关反馈算

法。在每次反馈中根据用户标记的相关和不相关样本来调整FSRM中的权值,在经过有限次的反馈之后,通过对FSRM中数据的学习来修正我们的语义矩阵。该算法是在FSRM的基础上进行反馈的,不需要特定问题的先验知识,在有限的反馈次数下,可以达到很好的检索效果。

2 基于模糊语义相关矩阵的相关反馈算法

2.1 模糊语义相关矩阵

1965年,Zadeh提出模糊数学的概念以研究模糊不确定性的问题。模糊数学把特征值的范围从集合 $\{0,1\}$ 扩充到区间 $[0,1]$ 的连续取值。因此一个对象符合某概念的程度,就可以用区间 $[0,1]$ 内的数来表示,对象所对应的数值愈接近1,表示该对象符合概念的程度愈大;反之,符合程度愈小。下面是模糊矩阵的定义^[9]:

设 $A=(a_{ij})_{m \times n}$,且 $a_{ij} \in [0,1] (i=1,2,\dots,m; j=1,2,\dots,n)$,则称 A 为Fuzzy矩阵。

由于人类语义和语言的模糊性,我们很难完整地対图像进行精确的语义或者文字标注,因此,用模糊语义相关矩阵(FSRM)来搭建语义桥梁^[8]。该矩阵满足:

$$R(i,j) \in [0,1], R(i,i)=1, \text{ 且 } R(i,j)=R(j,i), i,j=0, 1, \dots, N-1 \quad (1)$$

本文受重庆自然科学基金计划项目(2009BB2006),西南大学基本科研业务费专项资金项目(2010C028)资助。

郭士会(1986-),女,硕士生,主要研究方向为图像处理,E-mail:gshui@swu.edu.cn;杨明(1970-),女,博士,副教授,主要研究方向为人工智能、图像处理;王晓芳(1971-),女,硕士,讲师,主要研究方向为电子商务、语义Web。

式中, N 是图像库中图像的数目。称 $R(i, j)$ 为图像 i 和 j 的“直接相似度”, 它反映了图像 i 和 j 的相似关系。根据模糊矩阵的定义可知, FSRM 是 Fuzzy 相似矩阵。

2.2 在图像检索中的相关反馈算法

根据文献[2]中提到的基于模糊语义相关矩阵(FSRM)的图像检索方法可知, FSRM 的大小由图像库中图像的数目决定, 因此 FSRM 是一个非常大的矩阵, 相应的运算量就会很大。为了减少 FSRM 的大小, 本文使用将图像库中的图像先进行分类, 然后为每一类建立一个模糊语义相关矩阵的方法, 这样就大大减小了 FSRM 的大小。本文的图像检索方法主要分为 3 部分: 首先选用图像库中的图像作为用户待查询的样图(便于训练学习), 在每次反馈中根据用户标记的相关和不相关样本来调整对应类的 FSRM 中的权值; 其次在进行有限次的反馈之后, 根据 2.3 节中的算法来调整每一个 FSRM; 最后在前两步的基础上, 对非图像库中的图片进行样图检索, 同时根据用户的反馈结果来不断地重复前两步的操作。详细的检索算法如下:

(1) FSRM 的初始化, 图像 i 和 j 的相似度 $R(i, j)$ 应该满足:

$$0 \leq R(i, j) \leq 1 (i, j=1, 2, \dots, N) \quad (2)$$

如果 $i=j$, 那么

$$R(i, j)=1 (i, j=1, 2, \dots, N) \quad (3)$$

$$R(i, j)=R(j, i) (i, j=1, 2, \dots, N) \quad (4)$$

式中, N 为每个图像类中图像的个数。由于每个图像类中的各图像之间都有一定的相似性, 因此 FSRM 的初始值为 0 到 1 之间最模糊的值 0.5, 即 $R(i, j)=0.5 (i \neq j \text{ 且 } i, j=1, 2, \dots, N)$ 。

(2) 采用传统方法(底层视觉特征)对查询样图进行检索, 每次查询结果返回前 $n (n < N)$ 张最相似的图像。

(3) 用户对查询结果中的 n 张图像进行标记得到: 相关图像集为 I^+ , 不相关图像集为 I^- , 一般图像集为 I^0 。

(4) 首先考察 I^+ 中的图像, 在同一个图像类中的相关图像就根据式(5)调整它们所在 FSRM 中对应的权值, 否则不修改; 然后考察 I^- 中的图像, 在同一个图像类中的不相关图像就根据式(6)调整它们所在 FSRM 中对应的权值; 最后考察 I^0 中的图像, 在同一个图像类中的一般图像根据式(7)调整相关的权值。

$$R(i, j)_{\text{new}} = R(i, j)_{\text{old}} + \alpha(1 - R(i, j)_{\text{old}}) \quad (5)$$

$$R(i, j)_{\text{new}} = R(i, j)_{\text{old}} - \beta(1 - R(i, j)_{\text{old}}) \quad (6)$$

$$R(i, j)_{\text{new}} = R(i, j)_{\text{old}} + \chi(1 - R(i, j)_{\text{old}}) \quad (7)$$

式中, α, β 分别表示权值增大或减少的程度参数^[10]。此处取 $\alpha=0.3, \beta=0.6, \chi=0.1$ 。

(5) 通过步骤(2)~(4)的有限次的训练之后, 每个图像类的 FSRM 中的部分权值都已经由初始值 0.5 变为了更为贴近现实语义的值, 即每个图像类的各图像之间建立了一定的相似关联度。此时使用 2.3 节中的算法对每个 FSRM 中的数据进行学习来修正语义矩阵, 以提高检索效果。

(6) 根据用户提交的查询样图, 转入步骤(2)开始下一次循环。经过有限次的反馈之后, 执行步骤(5)来调整各个 FSRM 中的权值。

(7) 根据用户的反馈, 对标记为相关的图像所对应的 FSRM 中的权值从大到小排序, 返回最相似的图像。

由上述基于 FSRM 的相关反馈算法可以发现, 该算法具有一定的学习能力, 只要经过有效的训练之后就可以在后期的反馈中取得很好的结果。

2.3 基于 FSRM 的学习算法

在检索中, 由于低级视觉特征的相似度很难反映语义相似度, 语义信息的记忆和学习功能对系统检索性能的提高具有更重要的作用。

在经过有限次的反馈之后, 各个模糊语义相关矩阵中的部分权值已经由初始值(0.5)变成更为贴近现实语义的权值。例如, 图 1 所示为某个 FSRM 中前 10 张图像之间的模糊语义相关度。

	1	2	3	4	5	6	7	8	9	10
1	1	0.5	0.72	0.5	0.84	0.5	0.5	0.88	0.5	0.5
2	0.5	1	0.5	0.5	0.5	0.5	0.9	0.5	0.5	0.5
3	0.72	0.5	1	0.7	0.5	0.9	0.5	0.5	0.5	0.84
4	0.5	0.5	0.7	1	0.5	0.5	0.95	0.5	0.5	0.5
5	0.84	0.5	0.5	0.5	1	0.5	0.2	0.5	0.1	0.5
6	0.5	0.5	0.9	0.5	0.5	1	0.5	0.34	0.5	0.5
7	0.5	0.9	0.5	0.95	0.2	0.5	1	0.5	0.85	0.5
8	0.88	0.5	0.5	0.5	0.5	0.34	0.5	1	0.5	0.5
9	0.5	0.5	0.5	0.5	0.1	0.5	0.5	0.5	1	0.5
10	0.5	0.5	0.84	0.5	0.5	0.5	0.5	0.5	0.5	1

图 1 FSRM 学习之前

图 1 中的权值表示的是各个图像之间的直接相似度。例如通过几次反馈之后 $R(10, 3)$ 由初值 0.5 更新为 0.84, $R(3, 6)$ 由初值 0.5 更新为 0.9, 但是由于反馈次数有限, $R(5, 8)$ 仍然为初值 0.5 或者变化很小, 此时图像 5 和 8 的“直接相似度” $R(5, 8)$ 就很难反映它们隐含的语义相似度。相较而言, $R(5, 1) * R(1, 8) = 0.7392$ 更能反映图像 5 和 8 的隐含语义信息。

为了考虑隐含语义信息, 让语义信息能快速地在 FSRM 中传播, 定义 FSRM 中图像 i 和 j 的相似度 $R(i, j)$ 如下:

如果 $R(i, k) \geq T, R(k, j) \geq T, R(i, k) * R(k, j) \geq T$ 且 $R(i, k) * R(k, j) \geq R(i, j) \geq 0.5$, 那么 $R(i, j) = R(i, k) * R(k, j)$ 。

其中, T 为阈值, 我们取 $T=0.7$ 。

经过以上方法的学习之后, 图 1 中的 FSRM 被调整为图 2 中的结果。

	1	2	3	4	5	6	7	8	9	10
1	1	0.5	0.72	0.5	0.84	0.5	0.5	0.88	0.5	0.5
2	0.5	1	0.5	0.705	0.5	0.6	0.9	0.5	0.765	0.5
3	0.72	0.5	1	0.7	0.5	0.9	0.5	0.5	0.8	0.84
4	0.5	0.705	0.7	1	0.5	0.5	0.85	0.5	0.725	0.5
5	0.84	0.5	0.5	0.5	1	0.5	0.2	0.7392	0.1	0.5
6	0.5	0.6	0.9	0.5	0.5	1	0.5	0.34	0.5	0.765
7	0.5	0.9	0.5	0.85	0.2	0.5	1	0.5	0.85	0.5
8	0.88	0.5	0.5	0.5	0.7392	0.34	0.5	1	0.5	0.5
9	0.5	0.765	0.8	0.725	0.1	0.5	0.5	0.5	1	0.5
10	0.5	0.5	0.84	0.5	0.5	0.765	0.5	0.5	0.5	1

图 2 FSRM 学习之后

我们发现在经过有限次的反馈之后, 通过以上方法的学习可以使得 FSRM 中的数据更新得更快, 更接近现实的语义, 从而达到更好的检索效果。

3 仿真实验及分析

3.1 实验环境的建立

仿真实验基于 Matlab 平台, 所使用的是 Corel 图像库, 图像库包含 1000 幅大小为 $384 * 256$ 或者 $256 * 384$ 的 JPG 图像, 分为 10 类, 每一类包含 100 幅图像。

我们让 10 个现实中不知道本文算法的用户来使用该系统, 每个用户对于每个类(共 10 个类)检索 2 次(共 20 次), 每次检索他们可以根据实际情况反馈 1~5 次。在每次反馈中,

他们根据自己的要求对检索结果标注“相关”或者“不相关”。

3.2 实验分析

实验分析分为两个部分,第一部分从基于低层视觉特征的检索、基于无学习相关反馈的检索以及基于有学习相关反馈检索 3 个方面来比较该算法的有效性;第二部分从反馈次数的角度来考察该算法的有效性。

图 3—图 5 给出了检索一个“花”图像的检索结果,检索结果中返回了 15 幅最相似的图像。图 3 是基于低层视觉特征(纹理)的检索结果,其中包含了一些“错误图像”;图 4 是用户通过反馈之后,系统根据反馈信息调整相应的 FSRM 中的权值,然后返回的新的检索结果。检索结果表明该反馈算法有效地掌握了用户的语义要求。图 5 是在经过有限次的反馈之后,使用 2.3 节中的学习算法调整相应的 FSRM 中的权值后返回的新的检索结果。从图 4 和图 5 的对比可以看到,本文中的学习算法有效地提高了系统的性能。

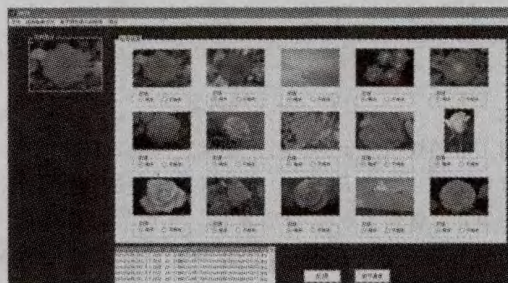


图 3 基于低层视觉特征(纹理)的检索结果

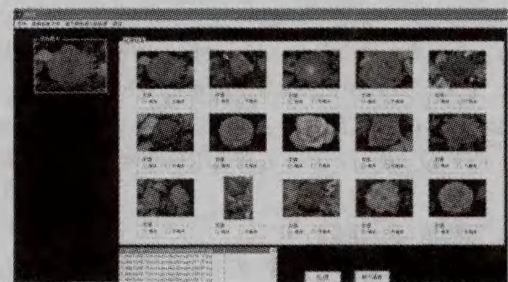


图 4 未学习之前第一次反馈后的结果

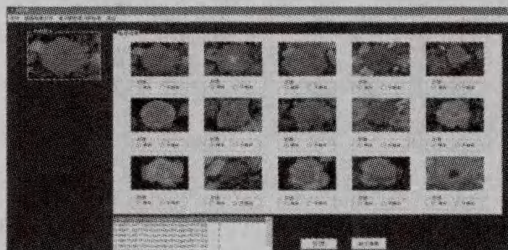


图 5 经过学习之后第一次反馈后的结果

图 6 是在不同的反馈次数下的有学习和无学习的查准率曲线。每条曲线包含 11 个点,它们分别对应于不同反馈次数的检索结果的综合评价指标。取定一个反馈次数值,对图像库中的 100 幅测试图像(随机抽取)进行检索,我们取检索结果的平均值作为该反馈次数的一个综合评价指标。查准率就是用户反馈为相关图像的个数与检索个数(本文为 15)的比

值。通过图 6 可以得到如下结论:(1)反馈次数一定的情况下,有学习比无学习的检索效果好;(2)随着反馈次数的增加,有学习和无学习的查准率都会增加;(3)在经过一定次数的反馈之后,有学习和无学习的查准率曲线趋于平稳状态。也就是说,我们的反馈算法在通过有限次的用户反馈之后能取得良好的检索效果。

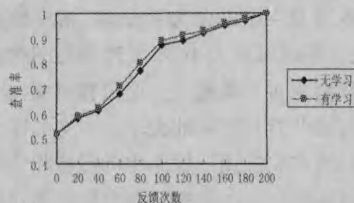


图 6 基于 FSRM 的相关反馈算法的检索实验

结束语 将模糊语义相关矩阵运用到图像检索的相关反馈过程中,根据用户的反馈标记相关和不相关图像来调整 FSRM 中的权值,不断捕捉用户的检索意图;同时经过有限次的反馈之后不断地使用本文提出的学习算法来更新 FSRM 中的值,从而在更少的反馈次数中将尽量好的结果反馈给用户。本文中的算法类似于人脑的经验积累机制,并且具有初步的学习机制。通过对 Corel 图像库的实验表明了该算法的有效性。

参考文献

- [1] Carneiro G, Chan A B, Moreno P J, et al. Supervised Learning of Semantic Classes for Image Annotation and Retrieval[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2007, 29: 545-554
- [2] Fechner S, Chang R, Qi Xiao-jun. Inter-query Semantic Learning Approach to Image Retrieval[C]// IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'10). 2010: 1246-1249
- [3] Han J, Ngan K N, Li M, et al. A memory learning framework for effective image retrieval[J]. IEEE Transactions on Image Processing, USA, 2005, 14(4): 511-524
- [4] Müller H, Müller W, Squire D M, et al. Performance evaluation in content-based image retrieval: overview and proposals[J]. Pattern Recognition Letters, 2001, 22: 593-601
- [5] Smeulders A W M, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years[J]. IEEE Transactions on pattern Analysis and Machine Intelligence, 2000, 22: 624-639
- [6] 张磊, 林福宗, 张钱. 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: 80-83
- [7] 曹奎, 冯玉才, 王元珍. 图像检索中一种新的相关反馈机制[J]. 计算机科学, 2002, 29: 65-68
- [8] 金海军. 基于语义特征的图像检索算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2006
- [9] 彭祖赠, 孙毓玉. 模糊数学及其应用[M]. 武汉: 武汉大学出版社, 2003: 75-80
- [10] 刘恒. 支持语义的图像检索系统研究与实现[D]. 上海: 上海海事大学, 2006