

基于强化学习的相关反馈图像检索算法

孙惠萍 龚声蓉 王朝晖 刘 全

SUN Hui-ping GONG Sheng-rong WANG Zhao-hui LIU Quan

苏州大学 计算机科学与技术学院 江苏 苏州 215006

School of Computer Science & Technology, Soochow University, Suzhou, Jiangsu 215006, China

E-mail: shrgong@suda.edu.cn

SUN Hui-ping, GONG Sheng-rong, WANG Zhao-hui et al. Relevance feedback in image retrieval algorithm based on reinforcement learning. Computer Engineering and Applications, 2008, 44(34): 175-178.

Abstract: Relevance feedback algorithm has been an important compose in image retrieval and it's a hotspot in the research of image retrieval recently. In this paper, the Reinforcement Learning (RL) based relevance feedback algorithm is proposed. According to Q -learning function of RL, a matrix named Q is established and for every image counterpart one item $Q_i (i=1, 2, \dots, n)$, which is used to record the image's total value of feedback in current retrieval. Each feedback calculates new features depending on weighted features method, then calculates every image's current total value of feedback base on Q -learning function. The bigger of the value of Q the more relevant with example image. RL acquires the best path by keeping feed back to the environment, which is consistent the idea of relevance feedback acquires the best answer by grouping the retrieval intention of users. The experiment proves the algorithm is more superior.

Key words: reinforcement learning Q -learning relevance feedback image retrieval

摘 要 相关反馈算法是图像检索不可或缺的重要组成部分,是近来图像检索中研究的一个热点。提出了基于强化学习的相关反馈算法。根据强化学习中的 Q -学习函数,建立矩阵 Q ,对每幅图像建立对应的一项 $Q_i (i=1, 2, \dots, n)$,记录每幅图像的本次检索中的累计反馈值,并根据加权特征法计算新的特征,对于每幅反馈的图像根据 Q -学习函数计算其当前的累计反馈值。 Q 值越大即越与例子图像相关。由于强化学习是通过不断对环境的反馈来获得最佳的路径,这与相关反馈通过对用户检索意图的摸索来获得最优答案的思想一致。实验表明,提出的相关反馈算法具有更大的优越性。

关键词 强化学习 Q -学习 相关反馈 图像检索

DOI: 10.3778/j.issn.1002-8331.2008.34.054 **文章编号:** 1002-8331(2008)34-0175-04 **文献标识码:** A **中图分类号:** TP391

1 引言

20 世纪 70 年代,由于政府、医院、学校等机构出现大量的图像、图形,查找非常麻烦,因而出现了基于文本的多媒体信息检索技术 (MIR)^[1-2]。该方法采用人工给每个图像标上文字信息,通过检索文字信息来搜索图像。该方法存在着人工标注麻烦、不同的人标注的不一致等缺点。为了解决这些问题,出现了基于内容的图像检索 (CBIR)^[3-4]。它借助于图像理解、模式识别、计算机视觉技术和数据库管理及信息检索技术,首先对图像采用特征提取算法提取相应的特征,如图像的颜色、纹理、形状等,并对这些特征按有效的方式进行索引和存储,形成对图像内容的有效描述;在检索阶段,根据用户要求进行查询,并返回满足条件的一组图像。然而,由于机器对图像的理解与人的理解存在差别,检索出的图像难以满足用户的需求。因此,20 世纪 90 年代,提出了相关反馈机制,成为了图像检索不可或缺的重要组成部分^[5-6]。

相关反馈的基本步骤为:

- (1) 用户提交一个例子,系统返回初始的查询结果;
- (2) 用户对查询结果标上相关或不相关,提交给系统;
- (3) 系统通过反馈进行学习,给出新的查询结果;
- (4) 返回到第(2)步,直到用户得到满意的结果,结束查询。

典型的相关反馈算法有,早期的 Yong Rui^[2]等人实现的 MARS 系统采用的基于特征向量权重调整和查询点优化的算法,它通过反馈的值计算特征权重值并忽略相关性很小的特征,同时修改查询点的位置,使之更接近相关的图像。该方法计算量大,查询点偏移容易错过相关的图像。近年来,很多研究者将机器学习方法应用到图像检索上。目前较流行的支持向量机 (SVM)^[7],它在有限样本的情况下具有良好的推广性,将首次反馈的 N 幅图像中,相关的 n 幅图像作为正样本,余下的 $N-n$ 为负样本,对正负样本进行学习训练,建立 SVM 模型分类器,再根据分类器对库中的图像进行分类。缺点在于将图像绝对的分

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.60673092) 教育部科研重点项目(No.205059)。

作者简介:孙惠萍,女,硕士研究生,主要研究方向为数字图像处理、图像检索;龚声蓉,男,教授,主要研究方向为模式识别与智能信息处理;王朝晖,女,副教授,主要研究方向为图像处理;刘全,男,教授,主要研究方向为智能信息处理。

收稿日期:2007-12-17 修回日期:2008-02-29

成两类,不能准确地描述非相关图像的分布。

本文通过对机器学习中强化学习算法的研究,提出基于强化学习的相关反馈算法。这是个合作型的相关反馈,对于每次的反馈,不给出已经标记为极相关的图像,而是给出相似度比较大且还未标记过的图像,直到用户提出结束查询才给出最终的答案。这样可以给用户更多的选择,同时,结合了特征加权算法来调整特征权重。实验证明,提出的相关反馈算法可取得满意的结果。

本文先给出纯特征加权算法、强化学习及 Q -学习算法的描述,再给出本文算法的思想,最后是本文的实验结果和与加权算法的比较。

2 纯特征加权反馈算法

相关反馈模型可以定义为 $RF=(R, Q, I, Fe, A)$ 五元组合。其中 R 为用户反馈的结果。 Q 为给定的查询对象。 $I=[I_1, I_2, \dots, I_N]$ 为图像对象的集合, N 是图像的个数。 Fe 为提取的图像各个特征的集合, $Fe=[Fe_{ij}]$ 是图像 I_i 的特征集合。即 $Fe_{ij}=[Fe_{ij1}, Fe_{ij2}, \dots, Fe_{ijk}]$, 其中 k 表示特征的长度。 A 为相似性度量方法,用于度量查询 Q 和图像 I_i 的特征的相似度。不同的相似性度量可用于不同的特征表示中。

在特征加权算法中^[2,8],图像的特征集合 Fe 定义为 $Fe_{ij}'=[Fe_{ij1} \times W_{ij1}, Fe_{ij2} \times W_{ij2}, \dots, Fe_{ijk} \times W_{ijk}]$, 即由权重 W 和原图像的特征集的乘积决定。

在相关反馈中,如果所有正反馈在某个维度上的值相差很大,说明这个维度和用户查询的关系不大,可以降低该维度的权重;反之,如果所有正反馈在某个维度上的值很接近,说明该维度很好地反映查询中的某个特点,应该提高该维度的权重。

一个新的权向量由 $W=[W_{ij1}, W_{ij2}, W_{ijk}]$ 组成,它们分别对应于图像间、不同特征间和同一特征不同分量间的权重。其中 $w_{ijk}=1/\sigma_{ijk}$, σ_{ijk} 是 Fe_{ijk} 序列的标准偏差, $W_i=[w_{i1}, w_{i2}, \dots, w_{ik}]$ 。

反馈算法的具体步骤为:

步骤1 用户提交一个例子,检索系统根据相似度算法 A 给出初始的查询结果。

步骤2 用户对查询结果给予标注相关或不相关,将奖赏值 R_i 反馈给系统。

步骤3 根据用户的反馈,获得当前权重 W ,库中图像特征间、特征内分别进行权重调整,得到新的特征 Fe_{ij}' ,根据调整后的特征给出新的查询结果。

步骤4 返回到步骤2,重复,直到用户结束查询。

3 提出的相关反馈算法

在一般的相关反馈算法中,用户每次的反馈值,仅是本次反馈给出新特征的依据。事实上,累计的反馈值,也可以提供有用的信息。强化学习法正是通过累积的反馈值来获得最终的最优策略。

3.1 强化学习

强化学习^[9-11](Reinforcement Learning)是机器学习中的一种方法,在没有先验知识的情况下,也可以从有延迟的汇报中学习并获得最优策略。它是指一个能感知环境的智能体,通过与环境的不断学习能获得达到目标的最佳途径。

如图1,当智能体(Agent)在某个状态(State)做出每个动作(Action)时,环境(Environment)都会提供一个反馈信号,即奖赏值(Reward)。智能体通过与环境不断地进行试探交互,从而

获得一个最优策略 $\pi: S \rightarrow A$ (S 为状态集, A 为动作集),即累计奖赏值最大的策略。Watkins 于 1989 年提出了 Q -学习算法,通过 Q -学习算法来获得最大的累计奖赏值。

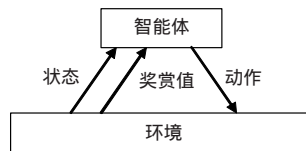


图1 强化学习原理图

Watkins 的 Q -学习算法,定义了一个评估函数 Q :从状态 s 开始,动作 a 为第一个有回报值的动作,所产生的奖赏值 r ,再加上将来的最大回报值 $\max_{a'} Q(s', a')$,以此评估函数的形式

实现最优策略 π 。 Q 评估函数可用公式表示为:

$$Q(s, a) = r(s, a) + \gamma \max_{a'} Q(s', a') \quad (1)$$

其中 $\gamma \max_{a'} Q(s', a')$ 表示将来的动作的 a' 的回报值, $\gamma \in [0, 1]$

是折算因子,当 $\gamma=0$ 表示 Q 函数只考虑当前的反馈值;反之 γ 取值越接近 1 表示将来的反馈值越重要。

实际算法中,用 \hat{Q} 函数来估计 Q 函数,即使用函数 $\hat{Q}(s, a) \leftarrow r + \gamma \max_{a'} \hat{Q}(s', a')$ 。

具体算法中,学习器通过一个大表 $Q(s, a)$ 表示,其中对每个状态-动作对 $\langle s, a \rangle$ 有一个表项来存储 \hat{Q} 的值。具体步骤为:

(1) 初始化表格 $Q(s, a)$;

(2) Agent 重复地观察当前状态 s , 选择某动作 a , 执行该动作;

(3) 观察结果回报 $r=r(s, a)$ 以及新状态 s' ;

(4) 根据 Q 函数来修改表 $Q(s, a)$ 的表项, 返回第(2)步。

在强化学习中,任一个状态对应的动作 a 的个数非常有

限,若 a 较多,还可根据概率公式 $P(a|s) = \frac{k^{Q(a)}}{\sum_j k^{Q(a_j)}}$ 来选择。当

$0 < k < 1$ 时,可使智能体探索目前 Q 值不是很高的动作。当 $k > 1$ 时, Q 值越大的动作 P 越大,本文利用 $k > 1$,在检索的最后给出 Q 值最大的相关图像。

3.2 算法思想

本文在利用加权反馈算法的基础上结合强化学习中的 Q -学习算法。根据强化学习的概念,给出图像检索的五元素 $IR(S, A, R, Q, E)$ 。其中,状态 S :当前用于匹配的图像的特征;动作 A :库中所有的图像;奖赏值 R :用户的反馈值;矩阵 Q :记录图像的累计奖赏值;环境 E 表示用户。

相关反馈与强化学习的区别是,在强化学习中对于动作的选择没有一定的特征参考,而相关反馈通过提取图像的可视化特征,来选择相关的图像,因此,相关的图像数是有限的。

(1) 图像 a

把相关反馈算法比做智能体,每一次反馈相当于智能体行动的一步,则相关反馈的动作就是寻找相关的图像。有一点区别的是,智能体每一步的动作是任意选择,而相关反馈需要根据匹配算法来给出概率较大的前几个图像。

(2) 反馈值 r

让用户选择单纯的相关或不相关有些绝对,可能有些图像界于相关和不相关之间,用户难以判断。Yong Rui^[12]等人将用

户的反馈分成 5 级的相关尺度, 分别为: 极相关、相关、无意见、无关、极无关, 对应的数值为 3、1、0、-1、-3。本文将负例看作是孤立的、无联系的, 因此分成 4 级。分别是: 极相关、相关、无关、极不相关, 即 r 有 4 个值 r^3, r^1, r^0, r^{-1} 。

(3) 特征 s

图像检索中的特征首先由用户提交的例子给出。新的特征值根据用户的反馈来决定。将矩阵中 $Q_i \geq 3$ 的图像进行特征加权, 作为新的特征 S' 。

3.3 算法具体步骤

由强化学习的性质可知, 这是个合作型的用户模式。即认为用户是有耐心的, 每次反馈给用户的是中间结果, 直到用户提出要结束查询时, 才给出最终的结果。具体步骤如下:

步骤 1 初始化矩阵 $Q, Q(a_i)=0 (i=1, 2, \dots, n)$;

步骤 2 用户提交一个例子, 提取特征 S , 根据匹配算法给出最匹配的 12 幅图像 a_1, a_2, \dots, a_{12} ;

步骤 3 用户对系统给出的 12 个图像给出反馈值 r_1, r_2, \dots, r_{12} 并提交, 系统根据评估函数 $\hat{Q}(s, \mu) \leftarrow r + \gamma \max_{a'} \hat{Q}(s', \mu')$ 计算这 12 幅图像的累计奖赏值并修改矩阵 Q ;

步骤 4 计算新的特征值 S' , 并提交匹配的图像, 对于已经获得奖赏值 r^3 和 r^{-1} 的图像不再显示给用户。这样便于给用户更多机会选择还未标记过的图像;

步骤 5 返回步骤 3;

步骤 6 当用户提出结束查询时, 根据 Q 矩阵, 由概略函数

$$P(a_i) = \frac{k^{Q(a_i)}}{\sum_j k^{Q(a_j)}} \text{ 给出 } P \text{ 值较大的前 12 幅图像即为最终的结果。}$$

其中, \sum_j 是所有的图像 $k > 1$ 。

为了 Q 矩阵中有标记的图像最少为 12 个, 最好检索 2 次以上。

4 实验分析

实验采用了基于内容的图像检索中颜色特征^[13-15]和形状特征^[15-17]的提取, 图库包括自然风景、豹、鹰、花朵等 300 幅图像, 实验在 matlab7.0 环境实现。实验设定反馈次数 n , 再给出最终结果。 Q 评估函数中折算因子 γ 取值 0.5, 结束查询时的概率函数中 k 值取 2。

图 2 所示是对豹的检索结果。由图可知检索结果比较准确, 因此不需要再加大 n 的值。



图 2 提交的例子图像(豹)

由于检索结果是根据 Q 矩阵给出, 因此获得正反馈值的图像个数至少为 12 个, 图 3、图 4 由于 $Q_i > 0$ 的个数少于 12, 因此出现了其他图像的填补。图 5、图 6 的反馈结果相同, 则当 $n=6$ 时就可以结束本次检索。

表 1 给出了使用基于强化学习的反馈算法检索结果, 表 2 给出了使用纯特征加权算法检索结果。 n 是指反馈的次数, 本



图 3 $n=2$ 的豹图像反馈结果

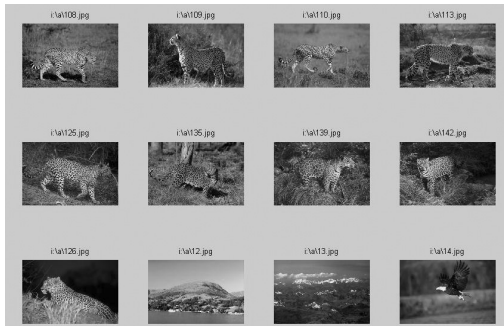


图 4 $n=4$ 时豹图像的反馈结果

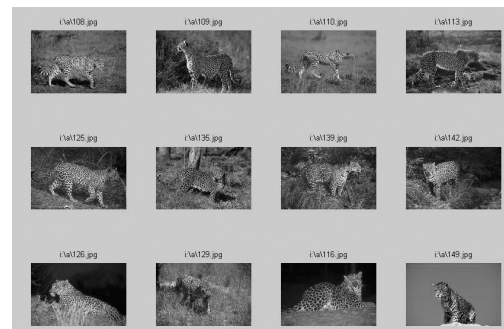


图 5 $n=6$ 时豹图像的反馈结果

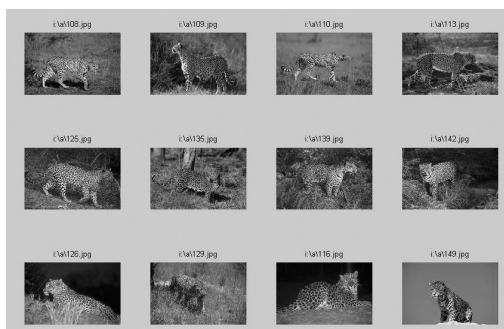


图 6 $n=8$ 时豹图像的反馈结果

文共进行了 8 次反馈; 两种方法分别进行了 3 次图像检索, 分别是花朵、鹰和豹, 每项数值是对应次数的查准率。

表 1 基于强化学习的反馈算法检索结果

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$
花朵	45.8	47.9	50	52.1	58.3	60.4	72.9	83.3
鹰	75.0	75.0	77.1	81.3	83.3	87.5	91.7	100.0
豹	33.3	33.3	50.0	70.8	79.2	87.5	87.5	87.5

由这两张表的数据比较可知, 基于强化学习的反馈算法明显优于纯特征加权算法。因为强化学习用 Q 函数记录累计反馈值, 同时合作型的方式可以更好地了解用户的需求。

表2 纯特征加权反馈算法检索结果

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$
花朵	37.5	39.5	41.7	41.7	41.7	47.9	47.5	50.0
鹰	66.7	68.9	68.9	70.8	70.8	72.9	72.9	75.0
豹	25.0	41.7	41.7	50.0	58.3	66.7	70.8	70.8

实验数据还表明,检索次数增多可获得更准确的反馈图像,但不是 n 的值越大越好。当用户认为已得到所需要的反馈结果,或反馈的中间结果图像中没有满意的图像,就可以结束这次的检索。可看出,本文算法明显优于特征加权算法。

5 结论

每次中间结果由 Q 函数记录,每次反馈舍去上一轮反馈值 $r=3$ 和 $r=-1$ 的图像,给出匹配概率比较大、 Q 值为空的图像,这样用户可以得到更多的信息,有更多的选择。

将强化学习方法用于相关反馈,通过 Q 数组记录每次的反馈结果,用户检索完成后,如果想保留当前的检索结果,则可以保留当前的 Q 数组。在本实验中,将 Q 设置成对应于一个正例图像的一维的数组,如果将 Q 设成 $n \times n$ 的二维矩阵,例子图像从库中选择,则可通过用户的使用记录,记录整个图像库中图像间的相互关联程度。这样,该方法就有了记忆性。

本文方法只使用了图像的颜色和形状特征,如果加上纹理特征或区域特征,将获得更好的检索结果。

参考文献:

- [1] Rui Yong, Huang T S, Mehrotra S et al. A relevance feedback architecture for content-based multimedia information retrieval systems[C]//CBAIVL, Jun 1997: 82-90.
- [2] Rui Yong, Huang T S. Content-based image retrieval with relevance feedback in MARs[C]//ICIP, Oct 1997: 815-818.
- [3] Smeulders A W M, Worring M, Santini S et al. Content-based image retrieval at the end of the early years[C]//TPAMI, Dec 2000, 22(12): 1349-1380.
- [4] Veltkamp R C, Tanase M. Content-based image retrieval systems a

survey[C]//The Knowledge Engineering Review Archive, June 2003, 18(2): 95-145.

- [5] 黄德才, 胡嘉, 郑月锋. 交互式图像检索中相关反馈进展研究[J]. 计算机应用研究, 2005, 9: 15-18.
- [6] 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.
- [7] Chen Yun-qiang, Zhou X S, Huang T S. One-class SVM for learning in image retrieval[C]//Proceedings of the 8th IEEE International Conference on Image Processing, Thessaloniki, Greece, Oct 7-10 2001, 1: 34-37.
- [8] 龚声蓉, 熊璋, 韩军, 等. 基于内容的自适应图像检索系统的设计[J]. 计算机工程, 2001, 27(10): 45-47.
- [9] Mitchell T M. 机器学习[M]. 北京: 机械工业出版社, 2005: 263-276.
- [10] 王珏, 周志华, 周傲英. 机器学习及其应用[M]. 北京: 清华大学出版社, 2006: 116-133.
- [11] Sudhakara R P, Bapi R S, Bhagvati C et al. Concept pre-digestion method for image relevance reinforcement learning[C]//Proceedings of the International Conference on Computing Theory and Application, Mar 2007: 605-610.
- [12] Rui Yong, Huang T S, Ortega M et al. Relevance feedback: a powerful tool in interactive content-based image retrieval[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5): 644-655.
- [13] 刘忠伟, 章毓晋. 十种基于颜色特征图像检索算法的比较和分析[J]. 信号处理, 2000, 16(1): 79-84.
- [14] 章毓晋. 基于颜色特征的图像检索[M]. 北京: 科学出版社, 2003: 57-130.
- [15] 章毓晋. 图像工程(下册)图像理解[M]. 2版. 北京: 清华大学出版社, 2007: 372-377.
- [16] 姚玉荣, 章毓晋. 利用小波和矩进行基于形状的图像检索[J]. 中国图象图形学报, 2000, 5(3): 206-210.
- [17] Liu Reng-yu, Jia Ke-bin, Wang Zhuo-zheng et al. A new and effective image retrieval method based on combined features[C]//Fourth International Conference on Image and Graphics, Aug 2007(22/24): 786-790.

(上接174页)

分析表2可以看出,小规模图像融合(256×256),多处理器并行实现效率仍然是很低的。但由于串行处理的时间已经很短了,该部分可以不用并行。当融合结果图像为512×512时,多处理器并行处理的优势显示出来,可以预测随着图像精度增高、图像规模的变大并行效率会更高。实验中,四处理器并行实现所用时间最短,计算量和多处理进程间通信量比例合适,使得加速比和并行效率最高。该算法的并行效率表明,该并行实现通信量较大,可扩展性不高。

表3是多处理器上对序列图像进行并行融合的时间统计,并行处理显著地减少了处理时间。进一步分析该算法的并行效率,维持在0.97附近,算法具有较高的可扩展性,可以预测在更多处理器上该并行算法处理时间会更短。

5 结论

本文首先给出了基于小波变换的多模医学图像融合方法,分析了其潜在的并行性,并给出并行实现方法。通过在不同机群数量上的测试,结果表明,改进单幅图像融合并行算法在结果图像较大时取得较好的并行效率,序列图像并行融合算法有很好的效果,具有较高的可扩展性,并行性能理想。

参考文献:

- [1] 刘贵喜, 赵曙光, 陈文锦. 双正交小波变换多分辨率图像融合方法[J]. 光电工程, 2004, 31(4): 50-53.
- [2] 孙夫雄, 孙东, 余智欣, 等. 基于小波提升算法的快速图像融合技术[C]//第五届全球智能控制与自动化大会, 2004: 4112-4115.
- [3] Wang P W, Wu X Q, Tang J H. Watershed segmentation based on multiscale morphological fusion[C]//The 8th International Conference on Signal Processing, 2006, 2: 16-20.
- [4] 康原原, 李彬, 田联房, 等. 基于小波变换和纹理测度的多模医学图像融合[C]//第26届中国控制会议, 2007: 697-699.
- [5] 陈文光, 武永卫. MPI与OpenMP并行程序设计[M]. 北京: 清华大学出版社, 2004.
- [6] Foster I. Designing and building parallel Programs[M]//Concepts and Tools for Parallel Software Engineering. Reading, MA: Addison-Wesley, 1995.
- [7] Cheng Y L, Li Y, Zhao R C. A parallel image fusion algorithm based on wavelet packet[C]//The 8th International Conference on Signal Processing, 2006, 2.
- [8] 程英蕾, 胡伏原, 赵荣椿. 一种多传感器图像并行融合新方法[J]. 空军工程大学学报, 2006, 7(6): 40-43.