

# 图像检索中 IRRL 模型研究

王朝晖, 孙惠萍

(苏州大学 计算机科学与技术学院, 江苏 苏州 215006)

**摘要:** 相关反馈实现了人机交互, 是图像检索中的不可缺少的部分, 一般图像检索中都使用一种反馈算法。IRRL 模型将机器学习中的强化学习原理应用到图像检索的相关反馈中来。它将现有的查询点优化、特征加权、贝叶斯分类器等算法作为系统学习的动作, 通过不同的状态选择不同的动作, 最终为不同类的图像寻找到合适的反馈算法策略, 最后根据策略进行具体的图像检索。文中对 IRRL 模型具体算法进行了研究, 并在此基础上提出了一些改进意见。

**关键词:** 强化学习; Q-学习; 相关反馈; 图像检索; IRRL 模型

中图分类号: TP391.3

文献标识码: A

文章编号: 1673-629X(2008)12-0035-03

## Research of IRRL Model in Image Retrieval

WANG Zhao-hui, SUN Hui-ping

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

**Abstract:** Relevance feedback realized a good man-machine interaction in image retrieval, therefore it became an indispensable part of image retrieval. Generally, there is only one feedback algorithm in an image retrieval system. IRRL model integrate query vector modification, feature relevance estimation and Bayesian inference as actions for learning of the system, through various state choosing different action, at last looking for different feedback algorithm strategies for different classes of images, then use the strategies to detailed retrieval. In this paper, based the research of IRRL model, advanced some advices.

**Key words:** reinforcement learning; Q-learning; relevance feedback; image retrieval; IRRL model

## 0 引言

相关反馈机制的出现解决了由于机器对图像的理解与人的理解存在差别, 基于内容的图像检索系统检索出的图像难以满足用户的需求的问题, 目前已成为图像检索不可缺少的重要组成部分<sup>[1,2]</sup>。相关反馈的基本步骤为: a. 用户提交一个例子, 系统返回初始的查询结果; b. 用户对查询结果标上相关或不相关, 提交给系统; c. 系统通过反馈进行学习, 给出新的查询结果; d. 返回到第 b 步, 直到用户得到满意的结果, 结束查询。

一个图像检索中只有一个相关反馈算法, 不能适用于不同类的图像的检索, 而且每种反馈算法都有其局限和缺点。IRRL (Image Relevance Reinforcement Learning) 模型, 由 Yin 等人提出, 他们将现有的三种反馈算法: 查询点优化、特征加权、贝叶斯分类器作为检

索系统要研究的“动作”, 选择一副图像来进行检索, 最后得到由这 3 个算法组成的一个序列, 其中以能够得到最优结果的序列作为最优策略。保存该策略, 其他的图像检索的时候都使用这个策略。

文中先给出三种反馈算法的概述, 然后是强化学习及 Q-学习算法的描述, 再给出 IRRL 模型的思想, 最后是 IRRL 模型的缺点及提出的改进方法。

## 1 相关反馈

将相关反馈模型定义为:  $RF = (R, X, I, Fe, Al)$  五元组合。其中:

\*  $R$ : 用户反馈的结果, 如果只有相关和不相关两个尺度, 则  $R^+$  表示相关,  $R^-$  表示不相关。

\*  $X$ : 给定的查询对象。

\*  $I = [I_1, I_2, \dots, I_N]$  是图像对象的集合,  $N$  是图像的个数。

\*  $Fe$ : 是提取的图像各个特征的集合,  $Fe = \{Fe_i\}$  是图像  $I_i$  的特征集合。

\*  $Al$ : 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

收稿日期: 2008-03-05

基金项目: 国家自然科学基金 (60673092); 教育部科研重点项目 (205059); 江苏省高技术研究计划项目 (BG2005019)

作者简介: 王朝晖 (1967-), 女, 硕士, 副教授, 研究方向为多媒体技术。

特征表示中。下面给出常见的欧氏距离法。

设查询对象的特征表示为:  $X = \{X_{ij}\}$ , 库中某个图像  $Y$ , 则两者的相似度表示为:

$$\text{Dis}_i(X_i, Y_i) = \sqrt{\sum_{j=1}^K (x_{ij} - y_{ij})^2 / K} \quad (1)$$

下面介绍 IRRL 中用到三种常见的相关反馈算法。

### 1.1 查询点优化算法

设图像检索<sup>[3,4]</sup>到  $t$  步, 当前用于检索的特征已经被优化成  $X_t^i$ , 当前反馈标注为相关  $R^+$  的图像是  $Y^+$ , 标注为不相关  $R^-$  的图像是  $Y^-$ 。则新的查询点:

$$X_t^{i+1} = \alpha X_t^i + \beta \sum \frac{Y^+}{R^+} - \gamma \sum \frac{Y^-}{R^-} \quad (2)$$

由于不相关的图像有可能是无关联的, 因此这个方法并不能在任何时候都有效地接近用户所需要的答案。

### 1.2 特征加权算法

在特征加权算法中<sup>[4,5]</sup>, 图像的特征集合  $Fe$  定义为:

$$Fe'_i = [Fe_i \times W_i \mid Fe_{i1} \times w_{i1}, Fe_{i2} \times w_{i2}, \dots, Fe_{iK} \times w_{iK}] \quad (3)$$

即由权重  $W$  和原图像的特征集的乘积决定。一个新的权重向量由  $W = [W_1, W_2, \dots, W_d]$  表示, 同一个特征间的权重表示为  $W_i = [w_{i1}, w_{i2}, \dots, w_{ij}, \dots, w_{iK}]$ 。其中  $w_{ij} = 1/\sigma_{ij}$ ,  $\sigma_{ij}$  是  $Fe_{ij}$  序列的标准偏差。该方法额外要计算标准偏差, 计算量比查询优化大的多。

### 1.3 贝叶斯分类器算法

使用通用的高斯分布模型作为概率分布模型<sup>[6]</sup>, 设库中图像  $Y$  满足高斯分布, 则其概率密度函数为:

$$P(Y) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(Y-\epsilon)^T \Sigma^{-1}(Y-\epsilon)} \quad (4)$$

其中  $\epsilon = [\epsilon(I_1), \epsilon(I_2), \dots, \epsilon(I_N)]$ ,  $\Sigma = E\{(Y - \epsilon)(Y - \epsilon)^T\}$ 。基于这种分布假设建立贝叶斯分类器, 图像  $Y$  属于正例的概率函数:

$$g(Y) = \ln P(Y \mid R^+) p(R^+) = \ln P(Y) \quad (5)$$

将式(4)代入即可。引入对数  $\ln$  是为了计算方便。该方法需要大量样本来学习。

## 2 强化学习

强化学习是机器学习中的一种, 主要用于智能机器人的研究。它是通过计算累积的反馈值的大小来获得最终的最优策略。

### 2.1 强化学习(Reinforcement Learning)

强化学习<sup>[7,8]</sup>是机器学习中的一种方法, 一个能

感知环境的智能体, 通过与环境的不断学习从而获得达到目标的最佳途径。

如图 1, 当智能体(Agent)在某个状态(Station)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。

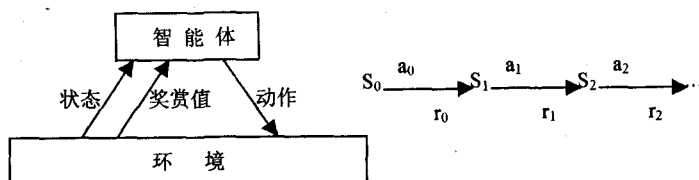


图 1 强化学习原理图

由此可得出最优策略的表达式:

$$\pi^* = \arg \max_a [r(s, a) + \gamma V^*(\delta(s, a))] \quad (6)$$

其中,  $\gamma \in [0, 1]$  是折算因子,  $\delta(s, a)$  代表应用动作  $a$  到状态  $s$  的结果状态。如果回报函数  $r$  和状态转换函数  $\delta$  都未知, 则智能体不能通过上式进行最优策略的评估。

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, 解决了式(6)的问题。

### 2.2 Q-学习算法

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  $\pi$ 。实际算法中, 用  $Q$  函数来估计  $Q$  函数, 其中  $\gamma \max_{a'} Q(s', a')$  表示将来的动作的  $a'$  回报值:

$$Q(s, a) \leftarrow r + \gamma \max_{a'} Q(s', a') \quad (7)$$

算法实现中, 学习器用一个大表  $Q(s, a)$  表示, 其中对每个  $\langle s, a \rangle$  有一个表项来存储  $Q$  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。

动作  $a$  的选择可根据概率公式来选择, 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:

$$P(a_i \mid s) = \frac{k^{Q(a_i)}}{\sum_j k^{Q(a_j)}} \quad (8)$$

### 3 IRRL 模型

IRRL 模型<sup>[9]</sup>结合强化学习中的 Q-学习算法。根据强化学习的概念,本模型先将图像分成  $N$  类,每一类选择一幅图像来进行检索,保存最终得到的最优策略,该策略适用于这个类的所有其他图像。

如图 2 所示是将三种反馈算法(Q 查询优化,F 特征权重,B 贝叶斯分类器)作为强化学习中的动作,产生的所有可能的反馈算法。下面先给出图像检索的五要素 IR( $S, A, R, Q, E$ )。

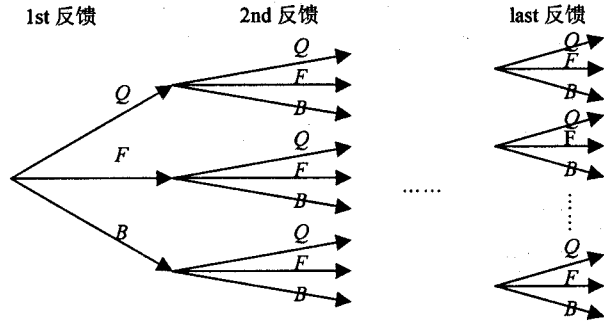


图 2 所有可能的相关反馈算法图

\*  $S = \{s_{i,j,l} | 1 \leq i \leq N, 0 \leq j \leq M, 0 \leq l \leq L\}$ : 状态  $S$  包含了三个信息,  $i$  表示图像属于的类别;  $j$  表示反馈的次数;  $l$  表示所使用的反馈算法。  $j = 0$  的时候表示反馈还没开始, 则  $l = 0$ 。

\*  $A = \{a_h | 1 \leq h \leq L\}$ : 动作  $a$  表示图像检索系统中的反馈算法, 有查询优化、特征加权、贝叶斯分类器 3 种。

\* 奖赏值  $r$ : 需要记录的不是用户反馈的每幅图像奖赏值, 而是本次反馈算法的奖赏值, 因此,  $r = \text{反馈为 } R^+ \text{ 的图像数} / \text{总反馈的图像数}$ 。

\* 环境  $E$ : 用户。

\* 状态转换函数  $\delta: \delta(s_{i,j,l}, a_h) = s_{i,j+l,h}$

算法:

初始化: 对于所有的  $i, j, l, h, \hat{Q}(s_{i,j,l}, a_h) = 1$ ;

开始: 1) 选择类  $i$  的一幅图像,  $j = 0$ , 令  $I_i$  表示当前检索的图像类。

2) 如果  $I_i$  还没有最优策略, 跳到第 4 步,  $l = 0$ 。

3) 将策略中反馈算法  $a_h$  用于  $I_i$ , 开始检索。

4) 将当前状态赋给  $s$ 。

5) 如果  $l = 0$ , 根据欧氏距离计算与例子图像最相似的  $n$  幅图像输出给用户; 否则根据策略中的  $a_h$  计算最相似的  $n$  幅图像。

6) 如果用户不满意该检索结果:

(1) 给这  $t$  幅图像标记相关或不相关;

(2) 反馈算法  $a_h$  根据式(8)来选择, 这是具有随机性的;

(3) 根据选择出的  $a_h$ , 计算并给出  $n$  幅图像;

(4) 计算反馈值  $r: r = r(s_{i,j,l}, a_h)$  和新状态

$s_{i,j+l,h} = \delta(s_{i,j,l}, a_h)$ ;

(5) 根据式(7)更新  $\hat{Q}(s_{i,j,l}, a_h)$ ;

(6)  $j = j + 1, k = h$ ;

End

End

### 4 提出的改进方案

不难看出, 该方法主要是产生一个策略, 以方便同类中其他图像的检索, 但是有关类中图像的选择存在一些问题。如图 2 所示, 随着反馈次数的增加, 可供选择的方案就越多, 也就是说矩阵  $Q$  将非常庞大。

#### 4.1 图像的选择

首先, 同一个类中的图像不易太多, 类的划分应该跟图像提取的特征有关。如果图像提取的是颜色和纹理特征, 则在手动分类的时候, 应该将纹理、颜色相近的图像放在一类中。如果是将形状相似的图像放在同一类, 那么该类的策略不一定适合类中所有的图像。

有关类中图像的选择, 比较简单的方法是计算中值图像。设某一类有  $N$  幅图像, 系统按照颜色和纹理来进行检索。则先提取该类中图像的颜色和纹理特征。  $F^{\text{col}} = (f_1^{\text{col}}, f_2^{\text{col}}, \dots, f_N^{\text{col}})$  和  $F^{\text{tex}} = (f_1^{\text{tex}}, f_2^{\text{tex}}, \dots, f_N^{\text{tex}})$ , 然后求均值:

$$(f_i^{\text{avgcol}}, f_i^{\text{avgtex}}) = \frac{1}{N} \left( \sum_{j=1}^m f_{ij}^{\text{col}}, \sum_{j=1}^k f_{ij}^{\text{avgtex}} \right) \quad (9)$$

其中  $m$  表示颜色特征的个数,  $k$  表示纹理特征的个数, 最后计算该类中与均值最相近的图像, 以此作为中值图像来进行策略的选择。

#### 4.2 反馈算法的选择

如果不做限制, 那么随着反馈次数的增多, 矩阵越来越庞大, 而且每一轮所花费的时间也逐渐增多。

提出的方法是, 在每轮反馈之前, 并不是将所有可能的动作都执行, 而是通过公式(9), 令  $k > 1$ , 选择概率最大的一个或者两个动作, 如果某个动作的概率明显的很低, 则放弃该方法。

假设反馈 4 次, 如果全部动作都选择, 则有  $3^4 = 81$  种方案, 如果按照改进的方法, 假设从第一轮反馈后, 每次只选择两种动作, 则有  $3 + 2^1 * 3 + 2^2 * 3 + 2^3 * 3 = 45$  种。这样就可以有效地缩减  $Q$ , 从时间和空间上都得到了优化。

### 5 结束语

有关相关反馈, 一直是图像检索研究的热点, 并产生了很多种反馈算法, 每种方法都有各自的长处和不

(下转第 40 页)

区域  $R_i$  和  $R_j$  之间的边为  $l$ , 则边界的强度定义如下:

$$\text{Str}(l) = \frac{1}{N} \sum_{p \in l} |\text{Grad}(p)|$$
 其中,  $p$  为边  $l$  上的点,  $\text{Grad}(p)$  为像素  $p$  的梯度, 这样, 当  $\text{Str}(l) < T_2$  时, 则将两区域进行融合, 其中  $T_2$  为门限, 对于某区域来说, 区域灰度相似性是全局性的, 而边界的强度准则是局部性的。

#### 4 实验结果及分析

实验结果如图 1~图 4 所示。图 1 为原始图像; 图 2 为直接分水岭分割后的图像, 由该图可以看到过度分割的情况, 此时的区域个数最多, 且区域最零散; 图 3 为经平滑后, 再进行分水岭变换的结果图像, 由图可见, 平滑过程使得区域个数减少很多; 图 4 为经平滑、分水岭变换、再经区域融合后而得到的结果图像, 效果非常理想。

实验结果表明, 文中的方法获得了较好的分割结果<sup>[4]</sup>。

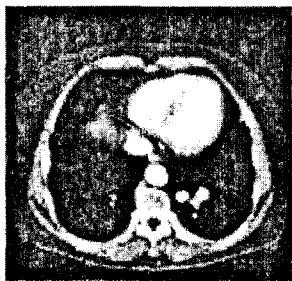


图 1 原图

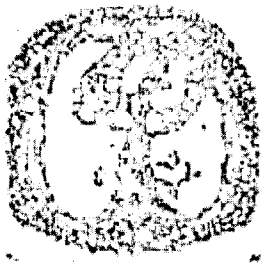


图 2 直接分水岭变换后的图

#### 5 结束语

首先使用均值滤波对原图像进行平滑, 然后对平滑后的图像进行分水岭分割, 得到初始的分割结果, 最后进行区域的融合, 从而较好地解决了图像的过度分割问题, 使图像完全从背景中分离出来。

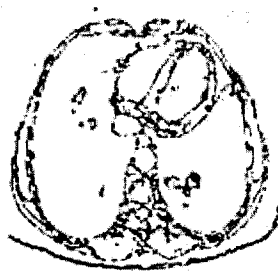


图 3 经平滑后分水岭变换后的图

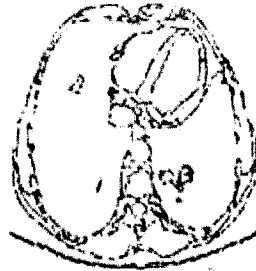


图 4 区域合并后的图像

#### 参考文献:

- [1] Gonzalez R C, Woods R E. 数字图像处理[M]. 北京: 电子工业出版社, 2006: 500-506.
- [2] 杜啸晓, 扬新, 施鹏飞. 一种新的基于区域和边界的图像分割方法[J]. 中国图像图形学报, 2001, 6(8): 1-4.
- [3] 高秀娟, 张树功. 图像分割的理论、方法及应用[D]. 长春: 吉林大学, 2006: 20-30.
- [4] 章毓晋. 图像处理与分析[M]. 北京: 清华大学出版社, 2000: 200-203.
- [5] Makovoz D. Image Segmentation[M]. Chicago: [s. n.], 2007.
- [6] 刘喜英, 吴淑泉, 徐向民. 基于改进分水岭算法的医学图像分割的研究[J]. 微电子技术, 2003, 31(4): 3-4.

(上接第 37 页)

足, IIRL 模型没有依照一般的追求新颖的反馈算法的思路, 而是将已有的几种反馈算法结合起来, 这样可以在一定程度上很好地扬长避短。

当然, 可用于 IIRL 模型的不止这三个算法, 其他的算法相结合也许有更好的效果。

#### 参考文献:

- [1] 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.
- [2] 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.
- [3] Yong Rui, Huang T S. Content-Based Image Retrieval with Relevance Feedback in MARs[J]. ICIP, 1997, 2: 815-818.
- [4] Yong Rui, Huang T S, Mehrotra S, et al. A Relevance Feedback Architecture for Content-based Multimedia Information Retrieval Systems[C]// In: Proceeding of IEEE Workshop on

Content Based Access of Image and Video Libraries. Porto Rico: [s. n.], 1997: 82-90.

- [5] 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.
- [6] Su Z, Zhang H, Ma S. Using Bayesian classifier in relevant feedback of image retrieval[C]// In: Titsworth M. Proceedings of the 12th IEEE International Conference on Tools with Artificial Intelligence (IEEE ICTAI 2000). Vancouver: IEEE CS Press, 2000: 258-261.
- [7] Mitchell T M. 机器学习[M]. 北京: 机械工业出版社, 2005: 263-276.
- [8] 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.
- [9] Yin P Y, Bhanu B, Chang K C, et al. Integrating Relevance Feedback Techniques for image retrieval Using Reinforcement Learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1536-1551.