

# 图像检索中基于长期学习的动态用户模型\*

张亮<sup>1</sup>, 周向东<sup>1+</sup>, 张琪<sup>2</sup>, 施伯乐<sup>1</sup>

<sup>1</sup>(复旦大学 计算机与信息技术系, 上海 200433)

<sup>2</sup>(Department of Computer Science, University of North Carolina at Chapel Hill, USA)

## A Long-Term Learning-Based Dynamic User Model in Image Retrieval

ZHANG Liang<sup>1</sup>, ZHOU Xiang-Dong<sup>1+</sup>, ZHANG Qi<sup>2</sup>, SHI Bai-Le<sup>1</sup>

<sup>1</sup>(Department of Computing and Information Technology, Fudan University, Shanghai 200433, China)

<sup>2</sup>(Department of Computer Science, University of North Carolina at Chapel Hill, USA)

+ Corresponding author: Phn: +86-21-65642219, E-mail: xdzhou@fudan.edu.cn, <http://www.fudan.edu.cn>

Received 2003-07-25; Accepted 2004-09-08

Zhang L, Zhou XD, Zhang Q, Shi BL. A long-term learning-based dynamic user model in image retrieval. *Journal of Software*, 2005,16(2):233-238. <http://www.jos.org.cn/1000-9825/16/233.htm>

**Abstract:** This paper proposes a probabilistic model incorporating long-term learning to estimate a dynamic user model. By using RF sequence as the user pattern, the approach can gradually update the prediction of current user based on matching the current user pattern with the user patterns in log. Compared with the invariant user model in PicHunter, the model is capable of dynamically adjusting when more user actions are observed, thus provides more accurate prediction for probability distribution. Experimental results on 11 000 images show that this approach can improve the retrieval accuracy apparently.

**Key words:** content-based image retrieval; user model; long-term learning; relevance feedback; Bayesian formulation

**摘要:** 提出一个结合长期学习的概率检索模型,即通过估算一个动态的用户模型来预测目标图像的概率分布.系统以反馈序列为用户模式,通过与反馈日志中用户模式的匹配,动态地调整对当前用户的预测.与 PicHunter 传统用户模型相比,此系统基于长期学习的方法,通过对用户反馈的观察,动态地调整当前的用户模型,从而更精确地推断当前用户的反馈行为.对 1.1 万幅图像数据库的实验表明,与已知方法相比,此方法对检索精度有明显的改善.

**关键词:** 基于内容的图像检索;用户模型;长期学习;用户相关反馈;Bayesian 公式

中图法分类号: TP311 文献标识码: A

图像数据库检索是一个具有挑战性的问题.由于图像的内在复杂性,用户衡量图像相似的语义空间与机器

---

\* Supported by the National Natural Science Foundation of China under Grant Nos.69933010, 60403018 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2002AA4Z3430 (国家高技术研究发展计划(863)); the Natural Science Foundation of Shanghai of China under Grant No.04ZR14011 (上海市自然科学基金)

**作者简介:** 张亮(1963-),男,湖北武汉人,博士,教授,主要研究领域为支持多媒体应用的数据库技术,信息集成;周向东(1969-),男,博士,讲师,主要研究领域为多媒体数据库,信息检索;张琪(1979-),女,博士生,主要研究领域为数据挖掘,多媒体信息检索;施伯乐(1935-),男,教授,博士生导师,主要研究领域为数据库理论与应用.

度量图像相似的特征空间之间难以建立起稳定的对应关系,使得检索的有效性受到影响.为了缩小用户对系统在图像理解上的偏差,用户相关反馈机制被引入到图像检索领域.基于短期学习的用户相关反馈假定目标图像集在特征向量空间中为单一簇连续分布,利用反馈信息修改查询向量使其向相关检索对象的分布中心移动<sup>[1]</sup>,或根据反馈信息调整距离度量公式中各分量的权值<sup>[2]</sup>.近来,机器学习方法,如支持向量机(SVM)<sup>[3]</sup>等也与相关反馈方法相结合来学习更复杂的目标集分布.然而,只利用当前检索的反馈信息,使得基于短期学习的检索效果的改进十分有限.系统积累的反馈历史信息包含了大量的用户对图像库中图像的语义相关性评价.采用适当的形式记录并合理地利用这些信息,将明显提高系统检索的性能<sup>[4-11]</sup>.

用户和相关反馈中起着关键作用,因此,对用户模型的研究具有特别重要的意义.然而,已知的长期学习方法通常是从图像间的相关性出发,试图通过某种机制建立起数据库图像之间的语义相关关系,忽略了不同检索用户对图像的语义判断多样性和灵活性的特点.事实上,对于每一次特定检索,通常都有一个特定的用户模式(用户类别),它决定着一类用户在特定时刻对图像语义的特殊判断.而在短期学习方法中,已知的用户模型往往是一种简化的、静态的模型,难以满足用户对检索性能的要求.

本文提出一种基于动态学习用户模型的贝叶斯概率检索方法.通过利用历史反馈信息的长期学习,动态估算用户模型,从而预测目标图像的概率分布.与 PicHunter 中的贝叶斯检索框架不同,本文方法首先将短期学习与利用用户反馈日志的长期学习结合,得到扩展的概率检索模型;然后利用反馈序列作为用户模式对用户行为进行描述;最后,通过当前用户模式与日志中的用户模式的相似性匹配,动态地估算用户模型.与 PicHunter 中的用户模型相比,本文给出的用户模型能够通过长期学习动态地调整,以适应不同用户的检索特点.

本文第 1 节对相关工作进行介绍.第 2 节给出结合长期学习的概率模型框架.第 3 节介绍动态用户模型及其学习的过程.第 4 节介绍检索流程.第 5 节给出实验结果与分析.最后是本文的研究结论.

## 1 相关工作

已有不少图像检索系统结合长期学习来提高检索性能.Lu 等人<sup>[5]</sup>利用已有的相关反馈信息建立、调整语义网络,并且结合语义网络与图像的视觉特征,改进图像的相似性度量.Bartolin 等人<sup>[6]</sup>提出了 FeedbackBypass 系统,在用户提交的检索样本与检索中通过相关反馈得到的“最优”参数之间建立映射关系,并记录到数据库中.当新的检索进行时,系统首先在数据库中进行相似匹配得到“最优”检索参数,然后使用这些参数进行检索.该方法可以减少反馈轮次,提高检索效率.Muller 等人<sup>[7]</sup>给出了一个使用大图像特征集的基于文本检索方法的图像检索系统,通过分析用户的反馈日志来调整图像特征相关度公式中的相关度因子(即该特征的权重).该系统对用户的反馈日志进行离线分析.Yang 等人<sup>[8]</sup>采用一种两级对等索引(peer index)来记录历史的反馈信息.该方法通过用户注册机制来识别用户/用户类别.

图像检索中的用户模型概念由 Cox 等人<sup>[9]</sup>在 PicHunter 中提出.PicHunter 采用基于贝叶斯(Bayes)公式的概率检索框架.根据贝叶斯公式,系统对检索目标的概率分布的计算转化为对用户模型的估计.但是,系统在检索过程中忽略用户状态的变化,没有考虑历史反馈信息的利用.

基于 PicHunter 用户模型概念,我们在概率检索模型中建立了一个动态的、考虑不同用户的用户模型.它能够通过用户模式匹配,动态地更新对当前用户的估计,从而更加精确地预测用户行为,估算概率分布.本文工作包括:在 PicHunter 的贝叶斯概率模型基础上,利用已有反馈历史的长期学习;利用相关反馈序列作为用户模式来描述执行检索的用户,从而使用户模型能够区分检索中的特定用户;通过编辑距离匹配查找相似用户模式,从而不断更新对当前用户的估计,动态更新用户模型.

## 2 结合长期学习的概率检索模型

本文考虑常见的目标类别检索问题,即我们要估算的是图库中每一幅图像  $T_i$  属于检索目标类别的概率  $P(T_i \in \mathcal{T}H_i)$ ,其中  $\mathcal{T}$  为图库中属于当前检索目标类别的所有图像的集合.在此基础上,结合对反馈历史数据  $L$  的长期学习,可以得到扩展后的检索概率  $P(T_i \in \mathcal{T}H_i, L)$ ,记为  $P_i^{t+1}$ .我们给出下面的定义.

**定义 1.** 短期历史  $H_t$ .

短期历史  $H_t$  定义为  $H_t = \langle D_1, A_1, D_2, A_2, \dots, D_t, A_t \rangle$ , 其中  $D_i$  是第  $i$  轮系统的检索结果显示,  $A_i$  是用户在第  $i$  轮做出的反馈(同文献[9]中  $H_t$  的定义).

**定义 2.** 长期历史  $L$ .

长期历史  $L$  定义为  $L = \{H_1^1, \dots, H_{i_N}^N\}$ , 其中  $H_i^i$  为记录了第  $i$  次检索过程的历史, 反馈轮次为  $i$ .  $L$  代表了当前检索可利用的已有检索的历史反馈信息.

按照 PicHunter 增量式计算概率  $P_i^{t+1}$  的思想<sup>[9]</sup>, 有:

$$\left. \begin{aligned} P_i^{t+1} &= P(T_i \in T | H_t, L) = P(T_i \in T | D_t, A_t, H_{t-1}, L) \\ &= \frac{P(D_t, A_t | T_i \in T, H_{t-1}, L) P(D_t, T_i \in T | H_{t-1}, L)}{\sum_{j=1}^n P(D_t, A_t | T_j \in T, H_{t-1}, L) P(T_j \in T | H_{t-1}, L)} \\ &= \frac{P(A_t | T_i \in T, D_t, H_{t-1}, L) P(T_i \in T | H_{t-1}, L)}{\sum_{j=1}^n P(A_t | T_j \in T, D_t, H_{t-1}, L) P(T_j \in T | H_{t-1}, L)} \\ &= \frac{P(A_t | T_i \in T, D_t, H_{t-1}, L) P_i^t}{\sum_{j=1}^n P(A_t | T_j \in T, D_t, H_{t-1}, L) P_j^t} \end{aligned} \right\} \quad (1)$$

则概率  $P_i^{t+1}$  的计算问题转化为  $P(A_t | T_i \in T, H_{t-1}, L)$  的计算, 也称其为用户模型, 记做  $UM_i^t$ .

### 3 动态用户模型 UM

用户模型的估计是整个概率模型检索框架的核心. 用户模型  $UM_i^t = P(A_t | T_i \in T, D_t, H_{t-1}, L)$  代表用户在本轮检索时已有前  $t-1$  轮短期历史  $H_{t-1}$ , 以及系统积累的长期历史  $L$  的情况下, 假设图像  $T_i \in T$  时用户做出反馈行为  $A_t$  的概率. 同样可以将  $UM_i^t$  写做  $UM_i^t = P(A_t | T_i \in T, D_t, U)$ , 其中  $U = (H_{t-1}, L)$ , 表示通过匹配当前部分用户模式  $H_{t-1}$  与已有用户模式  $L$  来推测当前用户  $U$ .

#### 3.1 一般用户模型和特殊用户模型

我们从两个层面来描述用户模型  $UM_i^t$ : 一般用户模型和特殊用户模型. 一般用户模型旨在通过所有用户对图像理解的共性来描述用户模型, 特殊用户模型则通过提取当前用户理解的特性来描述用户模型. 根据一般用户模型, 用户模型  $UM_i^t = P(A_t | T_i \in T, D_t, H_{t-1}, L)$  可写为  $P(A_t | T \in T, D_t)$ , 用  $GUM_i^t$  表示, 因为它不考虑用户变量  $U$ , 认为所有的用户都将采取一致的行为. 根据特殊用户模型, 用户模型可以写为  $P(A_t | T \in T, U)$ , 它通过区分不同的用户来描述用户模型, 即动态地识别当前用户  $U$ , 用  $PUM_i^t$  表示. 将两个模型结合起来, 得到下面估算用户模型的公式:

$$\left. \begin{aligned} UM_i^t &\approx \alpha GUM_i^t + (1 - \alpha) PUM_i^t \\ &= \alpha P(A_t | T_i \in T, D_t) + (1 - \alpha) P(A_t | T_i \in T, H_{t-1}, L) \end{aligned} \right\} \quad (2)$$

其中,  $\alpha (0 < \alpha < 1)$  为调节两个模型影响大小的因子. 本文方法同时结合了短期与长期的学习策略, 实现了用户模型的动态调整.

#### 3.2 用户模式

一般情形下, 用户在检索中的反馈行为在一定程度上反映了该用户的检索特征. 因此, 本文采用用户在检索过程中的行为序列(这里是指反馈序列)作为用户模式来描述特定的用户. 例如, 一个包含  $l$  轮反馈的检索的用户模式可以记为  $\langle A_1, \dots, A_l \rangle$ . 注意, 在特殊用户模型  $PUM_i^t = P(A_t | T_i \in T, H_{t-1}, L)$  中, 从当前用户的  $H_{t-1} = \langle D_1, A_1, D_2, A_2, \dots, D_t, A_t \rangle$  抽取用户的反馈序列是非常容易的. 因此, 在具体的实现中, 我们采用用户反馈的正例序列作为用户模式.

**定义 3.** 用户模式  $UP$ .

用户模式  $UP$  定义为一次特定检索中用户的反馈正例序列, 即  $UP = \langle I_1, \dots, I_s \rangle$ , 并且序列满足: 如果  $i < j$ , 则逻辑表达式  $(\text{Round}(I_i) < \text{Round}(I_j)) \vee (\text{Round}(I_i) = \text{Round}(I_j) = t \wedge (P_i^t > P_j^t))$  为真. 这里,  $\text{Round}(I_i)$  表示  $I_i$  在某一轮被标注为正例. 实际上,  $UP$  由每一轮的正例序列按轮数号由小到大拼接而成, 而每一轮中的正例则是按照概率大小

排序.

在具体实现中,系统累积的反馈历史信息也以用户模式的形式存储,即  $L=\{UP_1, \dots, UP_N\}$ .而在一次检索的过程中,当前的部分用户模式可以记为  $CUP=\langle I_1, \dots, I_c \rangle$ ,其中  $I_1, \dots, I_c$  为已有的  $t$  轮反馈中所标注的正例.

### 3.3 动态用户模型的估计

一般用户模型的估算类似 PicHunter.为了计算特殊用户模型  $PUM=P(A_t|H_{t-1}, L)$ ,需要有效的机制来匹配当前用户模式  $CUP$  和长期历史  $L$  中记录的用户模式.

在匹配用户模式之前,首先根据假定  $T_i \in T$ ,生成相关模式集  $RelPatSet$ .其定义为:  $RelPatSet = \{UP_i | T_i \in UP_i, i=1, \dots, N\}$ .

#### 3.3.1 利用编辑距离匹配用户模式

通过匹配当前用户模式与相关模式集  $RelPatSet$  中的模式,我们可以计算相似模式集  $SimPatSet$ .由于当前用户模式  $CUP$  与  $L$  中模式  $UP$  的匹配以序列形式表示,因此参照字符串相似匹配中编辑距离的概念,可以用两个序列的编辑距离来表示它们之间的相似程度.相关的概念与定义见文献[10].

**定义 4.** 当前检索用户模式  $CUP$  的相似模式集.

$$SimPatSet = \{UP_i | sim(CUP, UP_i) > \alpha, i=1, \dots, N\} \quad (3)$$

其中,  $\alpha$  为相似度阈值,  $sim(CUP, UP_i)$  为当前用户模式  $CUP$  与日志中的模式  $UP_i$  的相似度.

#### 3.3.2 估算特殊用户模型

根据相似模式集  $SimPatSet$  中的模式推断用户做出  $A_t$  行为的概率.对于  $SimPatSet$  中的每一个相似模式  $UP=\langle I_1, \dots, I_s \rangle, s > |CUP|=c$ [10].我们认为,紧跟在模式  $UP$  与当前模式  $CUP$  匹配的  $c$ -前缀  $[UP]^c = \langle I_1, \dots, I_c \rangle$  之后的  $|A_t|$  个图像  $I_{c+1}, \dots, I_{c+|A_t|}$  最有可能出现在  $A_t$  中,而在序列中位于  $I_{c+|A_t|}$  后面或  $I_{c+1}$  前面的图像出现在  $A_t$  中的概率则分别随它们的位置与  $c+|A_t|$  或  $c+1$  的距离的增大而减小.因此有下面的定义.

**定义 5.** 相似模式  $UP$  中的图像  $I_j(j=1, \dots, s)$  出现在  $A_t$  中的概率的位置因子  $pos(j)$ :

$$pos(I_j) = \begin{cases} 1 & \text{if } c < j \leq c + |A_t| \\ e^{-|j-(c+A_t)|} & \text{if } j > c + |A_t| \\ e^{-|j-(c+1)|} & \text{if } j \leq c \end{cases} \quad (4)$$

考虑到相似模式的相似度,有如下定义.

**定义 6.** 图库中图像  $T_k(k=1, \dots, n)$  出现在  $A_t$  中的概率  $pInAt(T_k)$ .

$$pInAt(T_k) = \sum_{i=1}^N \sum_{j=1}^{c_i} sim(UP_i, CUP) pos(I_j^i) \delta(T_k - I_j^i) \quad (5)$$

其中  $\text{if } x=0, \delta(x)=0$ .  $CUP$  表示当前用户模式,  $UP_i \in SimPatSet$  为  $CUP$  的相似用户模式,  $pos(I_j^i)$  代表用户模式  $UP_i$  中的第  $j$  个图像  $I_j^i$  的位置因子.  $sim(UP_i, CUP)$  代表  $UP_i$  对于  $CUP$  的相似度.其中,  $\delta(\cdot)$  用于滤除图库中没有出现在  $A_t$  中的图像.

设  $At = \langle T_1, \dots, T_i, \dots, T_d \rangle$ , 假设  $A_t$  中的图像被选为正例的概率是相互独立的,则可根据下式计算特殊用户概率

$$PUM_t^i = P(A_t | H_{t-1}, L) = \prod_{i=1}^d pInAt(T_i), T_i \in A_t \quad (6)$$

即对于  $A_t$  中出现的每一幅图像  $T_i$ ,求得相应的  $pInAt(T_i)$ ,并计算得到特殊用户概率.

由式(6)可知,  $A_t$  中的任一图像  $T_i$  的  $pInAt(T_i)$  为 0 将导致  $PUM_t^i$  为 0,为了避免这种情况的出现,对  $pInAt(T_i)$  的计算稍加一些变化:假设图库中的所有图像已计算  $pInAt(T_i)$ ,并组成相应的向量  $\vec{a} = [pInAt(T_1), \dots, pInAt(T_n)]$ .对于  $\vec{a}$  中值为 0 的元素,赋予它一个小于  $\vec{a}$  中最小元素的值,即  $\text{if } \vec{a}(i) = 0, \vec{a}(i) = \min(\vec{a}(j))\omega, i, j = 1, \dots, n, 0 < \omega < 1$ .最后,将向量  $\vec{a}$  归一化从而得到变更过后的每一幅图像的  $pInAt(T_i)$ .

## 4 检索系统及流程

整个检索模型以贝叶斯框架为基础,增量式地计算每一轮的检索目标概率  $P_i^{t+1} = P(T_i \in T | H_t, L)$ , 概率越

大,  $T_i$  属于检索类别的可能性越大, 根据  $P_i^{t+1}$  的大小对  $T_i$  排序, 概率最大的 100 幅图片作为该轮的结果集返回给用户。

根据式(1),  $P_i^{t+1}$  的计算由用户概率  $P(A_i | T_i \in T, H_{t-1}, L)$  及上一轮的检索概率  $\{P_i^t\}, i=1, \dots, N$  根据贝叶斯公式得到。初始(首轮检索)时, 即  $t=1$  时,  $P_i^1$  由一定的初始函数决定, 满足  $P_i^1 \propto \frac{1}{D(Q, T_i)}$ 。这里, 图像之间的距离度量

$D(Q, T_i)$  采用特征向量之间的带权欧氏距。当  $t > 1$  时, 检索概率  $P_i^{t+1}$  的计算转化为用户模型  $UM_i^t$  的计算。根据公式(2), 需要估算一般用户概率  $GUM_i^t$  及特殊用户概率  $PUM_i^t$ 。一般用户概率的计算公式为(见文献[9])

$$P_{soft\ min}(A = a | X_1, \dots, X_{N_D}, T) = \frac{\exp(-d(X_a, T)/\sigma)}{\sum_{i=1}^{N_D} \exp(-d(X_i, T)/\sigma)}$$

其中  $\{X_1, \dots, X_{N_D}\}$  为上一轮显示的结果图片,  $X_a \in X_A$  为本轮所选中的正例图片,  $X_A$  为正例图片的集合。假设所有图片  $X_a$  被选中为正例是两两独立的, 则我们将  $\prod_{X_a \in X_A} P_{soft\ min}(X_a)$  作为最终的一般用户概率。对于特殊用户模型

$PUM_i^t$ , 首先通过当前用户模式  $CUP$  与  $L$  中模式  $UP$  的匹配, 计算当前用户模式的相似模式集  $SimPattSet$ (式(3)), 然后考虑图片在  $SimPattSet$  模式中的每次出现(在所在的相似模式中的位置, 所在的相似模式与当前模式的相似度)来计算  $PUM_i^t$ (式(4)~式(6))。根据式(2), 用  $GUM_i^t$  和  $PUM_i^t$  计算用户模式  $UM_i^t$ , 根据式(1), 利用  $\{UM_i^t\}$  和  $\{P_i^t\}$  计算  $\{P_i^{t+1}\}$ , 最后, 将  $\{P_i^{t+1}\}$  从大到小排序并输出前 100 个检索结果。用户反馈后可以进行新一轮检索。

## 5 实验与分析

### 5.1 实验系统

实验图像检索系统中图像的视觉特征用颜色直方图与小波子段标准差来描述。系统在图像颜色特征上使用  $u=r/(r+g+b), v=g/(r+g+b)$  变换颜色空间, 并在  $u, v$  上各取 16 个 bin, 构成一个 32 维的向量, 再对图像进行 3 次小波变换得到 10 个子段, 取每个子段的标准差合成后得到一个 10 维的向量, 即数据库中的每一个图像由两个向量来描述。每一轮按照 PicHunter<sup>[3]</sup> 的概率模型进行隶属概率的估算, 在实现 PicHunter 方法的基础上, 实现了本文基于动态用户模型的长期学习方法。

实验中的图像库包含 11 000 幅图像。检索的主题固定为 4 个: 飞机、赛马、火车、海浪(经人工分类, 每个主题含 100 幅相关图像)。实验中用于对比的是 PicHunter 的方法。反馈日志文件包含对 10 个主题进行的共 700 次不同检索的反馈记录。由于数据库中每个主题有 100 幅相关图像, 我们统计每次检索后系统输出的前 100 幅图像中实际相关的图像个数, 按照上述查全率和查准率的公式, 实验中二者数值上是相等的, 统称为检索精度。

### 5.2 实验结果与分析

#### 5.2.1 检索性能的改进

首先考察本系统的检索性能, 使用本文方法分别和 PicHunter 方法对每一个子类各进行 10 次检索, 每一次检索进行 5 次反馈, 实验得到的每次检索的最大检索精度的对比结果在图 1 中给出。这里, 我们用 default 表示 PicHunter 方法, Long-Term 表示本文方法。两种方法的所有 40 次检索的平均检索精度对比在图 2 中给出。总共 40 次检索中有 39 次检索超过传统方法, 使检索精度比传统反馈平均提高了 11.175%(初始无反馈为 23.75%, PicHunter 为 42.275%, 本文方法为 53.45%)。所以实验显示本方法具有更好的检索精度。

#### 5.2.2 检索精度的提高

检索精度的提高可以用公式  $prc_{gain} = (prc(UM)/prc(cox) - 1) \times 100$  度量<sup>[6]</sup>, 其中  $prc_{gain}$  表示精度提高的倍数,  $prc(Long-term)$  表示本文方法的检索精度,  $prc(default)$  表示 PicHunter 方法的检索精度。每类主题的检索对检索精度的改进随反馈轮次的变化在图 3 中给出。注意到, 我们没有显示第 1 轮反馈的检索精度的提高, 因为在第 1 轮反馈前没有短期历史  $H_t$ , 也就是没有特殊用户模式的部分。所以我们的方法在第 1 轮反馈时与 PicHunter 方法是相同的。从图 3 可以看出, 本文方法通过前两三次反馈得到检索精度的改进最多。

由于需要对反馈日志进行匹配,因此与 PicHunter 相比,本系统在提高检索精度的同时需要付出更多的检索时间代价.进一步的工作将考虑如何提高动态用户模型学习的效率.

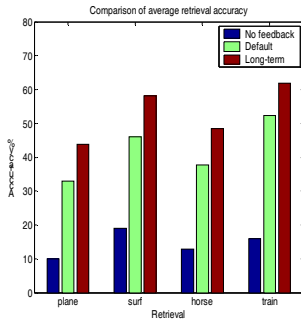


Fig.1 Comparison of maximum retrieval accuracy

图1 最大检索精度对比

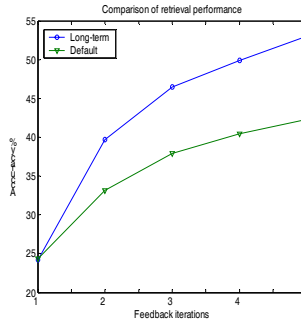


Fig.2 Comparison of average retrieval performance

图2 平均检索性能对比

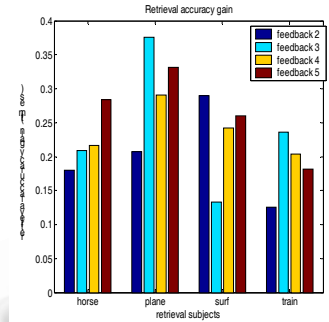


Fig.3 Retrieval accuracy gain

图3 检索精度的改进

## 6 结论

利用相关反馈的累积历史信息进行长期学习,是提高基于内容图像检索的效率和效果的有效途径.由于图像的特定检索语义可能随用户、时间的变化而变化,使得无论是反馈信息的记录还是利用都必须考虑到特定的用户模式.本文在 PicHunter 的概率模型基础上,完善了原有的用户模型,通过基于用户模式匹配的长期学习,识别当前用户的特定检索语义,从而真正实现了包含不同用户  $U$  的动态的用户模型,合理地利用了历史反馈信息.实验显示,与 PicHunter 相比,本文方法能够明显地提高相关反馈的效果.

## References:

- [1] Yu CT, Luk WS, Cheung TY. A statistical model for relevance feedback in information retrieval. *Journal of the ACM*, 1976,23(2): 273-286.
- [2] Rui Y, Huang TS, Mehrotra S. Content-Based image retrieval with relevance feedback in MARS. In: *Proc. of the IEEE Int'l Conf. on Image Processing*. Vol.2, New York: IEEE Press, 1997. 815-818.
- [3] Tong S, Chang E. Support vector machine active learning for image retrieval. In: *Proc. of the 9th ACM Int'l Multimedia Conf*. Ottawa: ACM Press, 2001. 107-119.
- [4] Fournier J, Cord M. Long-Term similarity learning in content-based image retrieval. In: *Proc. of the IEEE Int'l Conf. on Image Processing*. New York: IEEE Press, 2002. 441-444.
- [5] Lu Y, Hu C, Zhu X, Zhang H, Yang Q. A unified semantics and feature based image retrieval technique using relevance feedback. In: *Proc. of the 8th ACM Int'l Multimedia Conf*. Los Angeles: ACM Press, 2000. 31-37.
- [6] Bartolini I, Ciaccia P, Waas F. FeedbackBypass: A new approach to interactive similarity query processing. In: *Proc. of the 27th Int'l Conf. on Very Large Data Bases*. Roma: Morgan Kaufmann Publishers, 2001. 201-210.
- [7] Muller H, Muller W, Squire D. Learning feature weights from user behavior in content-based image retrieval. In: *Proc. of the Int'l Workshop on Multimedia Data Mining (MDM/KDD 2000)*. 2000. 67-72. [http://www.cs.ualberta.ca/~zaiane/mdm\\_kdd2000/proceedings.html](http://www.cs.ualberta.ca/~zaiane/mdm_kdd2000/proceedings.html)
- [8] Yang J, Li Q, Zhuang Y. Octopus: Aggressive search of multi-modality data using multifaceted knowledge base. In: *Proc. of the 11th Int'l World Wide Web Conf*. Honolulu: ACM Press, 2002. 54-64.
- [9] Cox I, Miller M, Minka T, Papathomas T, Yianilos P. The bayesian image retrieval system, picHunter: Theory, implementation and psychophysical experiments. *IEEE Trans. on Image Processing*, 2000,9(1):20-37.
- [10] Zhou XD, Shi BL, Zhang Q, Zhang L, Liu L. A long-term learning based similarity retrieval of multimedia database. *Journal of Software*, 2004,15(1):86-93 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/15/86.htm>

## 附中中文参考文献:

- [10] 周向东,施伯乐,张琪,张亮,刘莉.基于长期学习的多媒体数据相似性检索. *软件学报*,2004,15(1):86-93. <http://www.jos.org.cn/1000-9825/15/86.htm>