

基于空间关系的手绘草图检索^{*})

李彬 梁爽 孙正兴

(南京大学计算机软件新技术国家重点实验室 南京 210093)

摘要 手绘草图是人类最自然的人机交互方式,在普适计算模式下手绘草图将得到越来越多的应用,手绘草图的检索也成为一个新的课题。本文提出一种高效的草图检索方法,该方法以独立于用户绘制习惯的手绘草图统一表示为基础,以手绘草图空间关系为检索相似度匹配特征,并采用特征点调整的相关反馈机制优化检索结果。实验表明本文提出的方法具有很好的检索效果。

关键词 草图检索,空间关系,特征向量,相关反馈

Sketch Retrieval Based on Topological Relations

LI Bin LIANG Shuang SUN Zheng-Xing

(State Key Lab for Novel Software Technology, Nanjing University, Nanjing 210093)

Abstract Sketching is a natural way for recording and conveying ideas. Along with more and more sketches created in ubiquitous computing, sketch retrieval has become a new issues. This paper presented an effective approach for sketch retrieval to improve the retrieval performance. The topological relationship of sketch constitutes is used as the feature of similarity matching of the sketches, based on a user-independent sketch representation. Relevance feedback was also introduced to adjustment to refine the retrieval results by adjustment of feature point. Experiments prove the effectiveness and efficiency of the method.

Keywords Sketch retrieval, Spatial relation, Feature vector, Relevance feedback

1 引言

手绘草图作为自然而直接的人机交互方式,越来越受到人们的关注。它是辅助人类思维创新的主要方式之一,有效地提高了人类协作和交流的效率。20世纪90年代,小型化、便携式和无线设备的快速发展及平板电脑(Tablet PC)的出现,尤其是W3C推出的InkML标准草案实际上将“数字墨水”作为一种新型多媒体数据类型^[1],极大地推动了手绘草图领域的发展,大量的蕴涵丰富信息的草图不断涌现,对这些信息的检索和利用也就成为一个新的研究课题^[2]。

草图检索属于基于内容的图像检索的范畴。对此已存在一定的相关研究,但是大多数基于内容的图像检索的方法并不能够满足草图检索的要求。QBIC^[3]利用全局特征进行形状匹配;Sciascio^[4]利用基于傅立叶算子描述的形状信息进行形状匹配;Lopresti等^[5]提出将草图按照顺序笔划串进行匹配以及将草图看作图像分块匹配,但该方法对草图绘制笔划顺序过于敏感;Kato等^[6]提出的QVE系统利用草图对应块之间的相互关系进行相似度的计算,但该策略不适合粗略绘制的草图。近来,Leung等^[7]利用草图的多形式表示来消除用户绘制方式对草图表示的影响,同时利用相关反馈技术优化查询结果。然而,作者过多考虑笔划的细节信息,导致检索结果对用户绘制的过程非常敏感;Fonseca等^[8]利用笔划的层次结构、形状信息和高维度的索引来检索复杂的图形,但该方法不适合层次关系不明确的草图。

本文提出一种基于空间拓扑关系的草图检索方法。对一幅草图,我们首先通过预处理来获得不同用户对同一草图的相对一致的表示。本文抽取空间拓扑结构信息以获取草图的特征向量,并据此进行相似度计算。最后将相似度较高的草图返回给用户,并根据用户对返回结果集的评价进行反馈来优化查询的结果。试验表明,本文所提出的检索方法能够高效准确地进行草图检索。

2 系统框架

图1给出了草图检索的系统结构图。草图主要由两大部分组成:相似度计算和相关反馈。

首先,本文对草图进行预处理^[9],获得草图相对统一的表示。我们根据笔划的笔速和曲率信息,将笔划分割成为由基本图元(直线、弧线、椭圆)构成的图元集合。由于草图的笔划构成是不唯一的,不同的用户可能绘制出不同笔划构成的同一幅草图,而草图的图元构成是唯一的,所以利用图元构成可以获得草图的统一的图元表示。本文所介绍的草图检索的技术是在草图进行预处理的基础之上建立的。

在相似度计算模块中,本文考虑图元之间的8种空间关系,首先根据这8种关系抽取草图的空间关系特征。我们引入图谱来表示草图的空间关系信息,有效降低后期由图匹配带来的复杂性。我们再通过降维将获得的空间关系信息转换成草图的特征向量。最后,我们计算输入草图的特征向量和草图库中草图的特征向量之间的欧拉距离,从而通过计算进

^{*})本文得到国家自然科学基金(编号:69903006,60373065)和教育部“新世纪优秀人才支持计划”资助。李彬 硕士生,研究方向为视觉信息检索、智能用户接口;梁爽 博士研究生,研究方向为视觉信息检索、智能用户接口;孙正兴 博士,教授、博导,研究方向为智能用户接口、多媒体挖掘。

行草图匹配,并将较高相似度的草图作为候选集返回给用户。

在相关反馈模块中,用户首先对返回的候选集进行判定。如果返回的结果令用户满意,则检索成功。否则,用户标识出候选集中相关的结果并进行反馈,系统通过向量点位移的反

馈机制对结果集进行调整,从而不断提高检索效果,最终得到用户满意的检索效果。

系统的两个主要模块将在本文的第3部分和第4部分详细介绍。

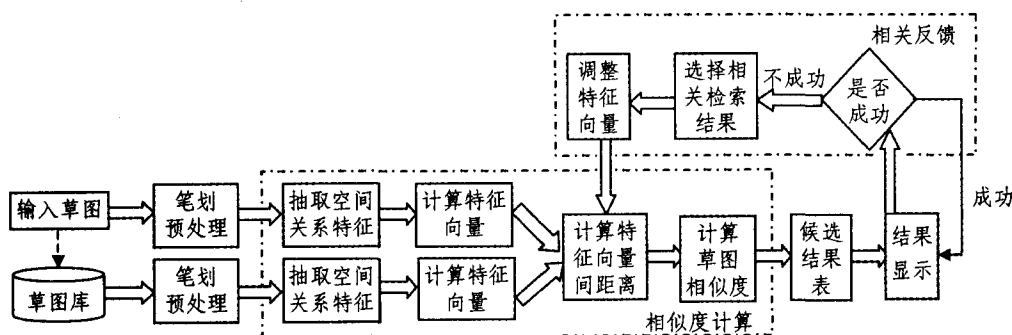


图1 基于空间关系的草图检索框架

3 相似度计算

预处理后,对于采用基本图元表示的草图,首先根据空间关系特征抽取特征,并根据抽取的特征生成基于图谱描述的特征描述子,再通过降维生成草图的特征向量,最后通过计算草图特征向量的相似度得到草图之间的相似度。

3.1 特征选取和特征抽取

为了使选取的特征较为准确地表达草图整体结构信息并能够应用到草图检索中,本文考虑到以下几个因素的影响。首先,选择的特征准确地反映出草图的整体信息。其次,选取的特征能够应不限制用户的绘制方式。最后,选取的特征在不同的用户绘制风格中应保持一致性,从而在不同用户的草图间进行检索。

考虑到以上因素,我们选取空间关系作为草图的特征描述。虽然不同的用户有不同的绘制习惯,但是整个草图的空间拓扑结构是相对统一的。我们根据图元之间的交点数和总体形状选取了8种图元间的空间关系,如图2所示。



图2 空间关系

我们利用空间关系拓扑图来描述抽取的图元之间的空间

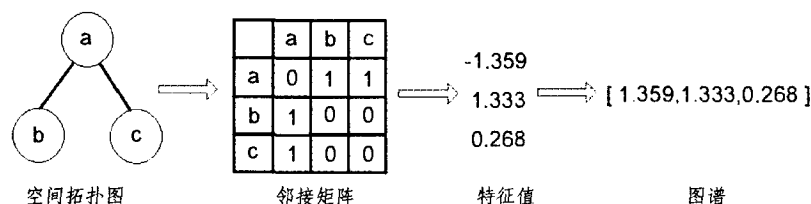


图4 图谱计算过程示例

关系, $G_i=(V, E_i)$ 表示第*i*种空间关系信息。其中*V*是图的顶点集, $E_i \subseteq V \times V$ 是图的边集。顶点表示草图中经预处理获得的基本图元,边 $(j, k) \in E_i$ 存在,当且仅当顶点图元*j*和顶点图元*k*之间存在第*i*种空间关系。以相邻关系为例,如图3所示,原始草图被分割为3个基本图元*a*、*b*和*c*,其中*a*与*b*、*c*相邻,而*b*与*c*不相邻(*b*、*c*为半相交)。所以草图对应的相邻关系拓扑图中,*a*与*b*相连,*a*与*c*相连,而顶点*b*与*c*之间没有边。

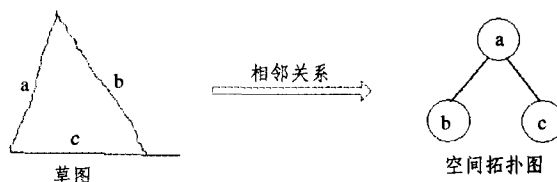


图3 相邻关系的拓扑图表示

因为我们选取了8种空间关系,所以共有8个对应的空间关系拓扑图。在得到草图的空间关系图后,计算草图相似度的问题就转化为计算草图对应的空间关系拓扑图的相似度这样一个计算问题。

3.2 基于图谱的特征描述子

众所周知,图的同构问题是NP完全问题,所以本文引入图谱^[10]来生成基于图谱的特征描述子,以便有效地降低相似度计算的复杂度。

图谱是图所对应的邻接矩阵的特征值的集合。利用图谱可以将空间关系拓扑图转化为其所对应的图谱,即本文中所述的基于图谱的特征描述子。不同的图可能有相同的图谱,这样就在转换中带来冲突。然而,在文[11]中作者论证了这种冲突发生的可能性很小,而且图谱相似是图相似的必要条件。这与检索的目的并不冲突,所以这种简化计算复杂度的方法在检索中仍然是有效的。

图4给出了计算图谱的方法。首先,我们获取拓扑图的

邻接矩阵表示。然后,计算邻接矩阵的特征值,由于邻接矩阵

是实对称矩阵,所以特征值都是实数。最后,把获得的特征值按照绝对值的大小降序,获得拓扑图对应的图谱。

经过拓扑图到图谱的转化之后,我们获得8种空间关系对应的图谱描述子。这些描述子能够有效地表示草图的结构特征,同时我们成功地把拓扑图的匹配问题简化为对应图谱描述子的匹配问题。

3.3 特征降维

如果一个草图由 n 个基本图元构成,其对应的空间关系拓扑图则有 n 个顶点,拓扑图对应的图谱是一个 n 维向量。构成不同草图的基本图元数目可能不等,所以草图对应的图谱描述子就具有不同的维度。不同维度的向量之间计算相似度比较困难,我们利用降维的思想将不同维度的向量转化到一维空间。

一个 n 维的向量可以看作 n 维空间的一个点,我们利用该点到原点的欧拉距离将 n 维的向量降维到一维的空间中。二维空间向量点降维如图5所示。如果两个草图的结构比较相似,则两幅草图的图谱描述子降维到一维空间的两个点就比较接近,所以我们降维后仍然可以有效地计算草图之间的相似度。

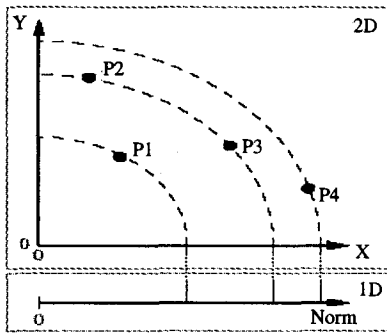


图5 二维空间向量点降维

图谱描述子降维后,每个草图可以得到8种空间关系所对应的一维空间的点,即8个数值。我们把这8个数值组成一个8维的向量 $D=[d_1, d_2, d_3, \dots, d_8]$,作为我们描述草图的特征向量。

3.4 相似度计算

我们首先计算查询草图对应的特征向量与草图库中的特征向量之间的欧拉距离,然后根据该距离计算两者之间的相似度。

在本文中,我们采用经典的二次欧拉距离来衡量两个向量之间的相似度距离。设两幅草图对应的特征向量分别为 $D=[d_1, d_2, d_3, \dots, d_8]$ 和 $D'=[d'_1, d'_2, d'_3, \dots, d'_8]$,则它们之间的欧拉距离为:

$$Dis = \sqrt{\sum_{i=1}^8 [FeatureVec(i)' - FeatureVec(i)]^2}$$

$$= \sqrt{\sum_{i=1}^8 (d'_i - d_i)^2}$$

欧拉距离 Dis 越小,则草图之间的相似度越高。实际上, Dis 是两幅草图之间的相异度,我们设定阈值(threshold)将相异度归一化为相似度,相似度的计算方法如下所示:

$$Similarity = \begin{cases} 0 & \text{if } Dis \geq threshold \\ 1 - \frac{Dis}{threshold} & \text{if } Dis < threshold \end{cases}$$

当欧拉距离 Dis 超过这个阈值时,则相似度取0;否则进行归一化。这种归一化比较简单,保证相似度顺序保持不变,并且得到的相似度就能保证在 $[0, 1]$ 之间。最后我们将相似度较高的草图作为结果集返回给用户。

与传统的基于图表示的草图匹配相比较,本文将图匹配的问题转化为特征向量的相似度计算,避免了图匹配的复杂性,有效提高了匹配的效率。

4 相关反馈

我们引入相关反馈来更好地捕捉用户的意图。相似度最高的草图作为候选集返回给用户后,相关反馈根据用户对结果集的评价和原始草图信息来优化检索结果。

4.1 用户评价

对于系统返回的候选结果集,用户根据自己的意图进行评价。本文主要从草图的空间拓扑结构上来捕获用户输入草图的意图,把握用户对草图的概念理解。如图6所示,相关的和相似的结果被标注出来,系统通过反馈算法对原始的检索请求进行调整,从而优化检索的结果。

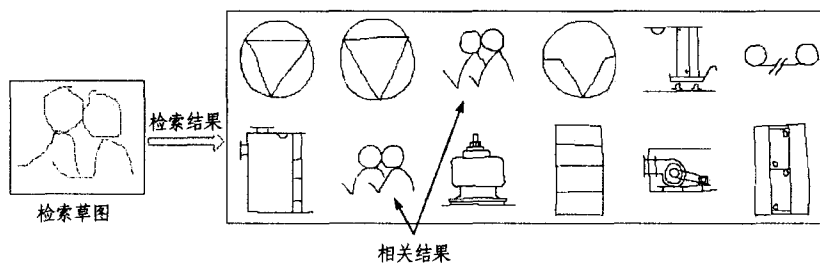


图6 用户评价

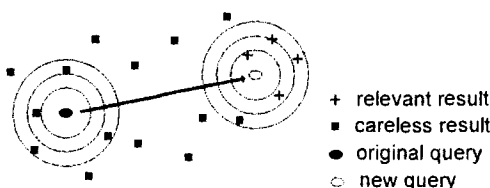


图7 向量点位移

4.2 特征向量调整

本文采用查询点位移的反馈算法。草图检索是以特征向量相似度计算为基础,所以草图对应的特征向量可以看作特征空间的向量点。查询点位移方法也使用于基于内容的图像检索的相关反馈^[12],其思想非常直观:根据原始的查询点和相关的向量点构造新的查询点,使该查询更接近相关的向量点,如图7所示。

本文采用文本信息检索系统 Rocchio^[13]中提出的特征向

量调整的方法。 $S_{rel} = \{O_j | O_j \text{ 是用户标注的相关结果}\}$ 是用户反馈中相关结果集。 $FeatureVec_j$ 是 O_j 对应的特征向量, 即第 j 个相关草图所对应的特征向量。 $FeatureVec_{new}$ 是调整之后生成的查询向量, $FeatureVec_{old}$ 是原始查询向量。调整生成的查询向量由相关结果集对应的特征向量和原始查询向量计算得到:

$$FeatureVec_{new} = \alpha * FeatureVec_{old} + \frac{\beta}{|S_{rel}|} * \sum_{O_j \in S_{rel}} FeatureVec_j$$

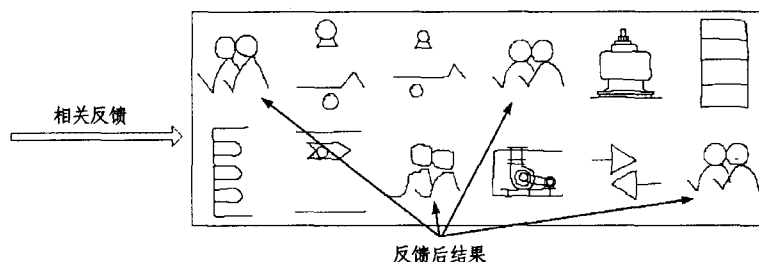


图 8 相关反馈后的检索结果

5 实验结果及分析

我们用实验来验证本文所提出的草图检索方法的有效性。为了验证在单用户和多用户情况下该检索策略的检索效率和相关反馈的效果, 我们设计了单用户实验和多用户试验。

实验在 Intel P4 PC (1.6G Hz CPU, 256MB 内存), Microsoft Windows XP 环境下进行。实验采用了如图 9 所示的 55 个草图类, 其中包含 30 类工程零件草图、25 类电器元件草图。我们共收集 10 个用户绘制的草图, 每个用户对每类草图绘制两幅, 数据库中共计 1100 幅草图。

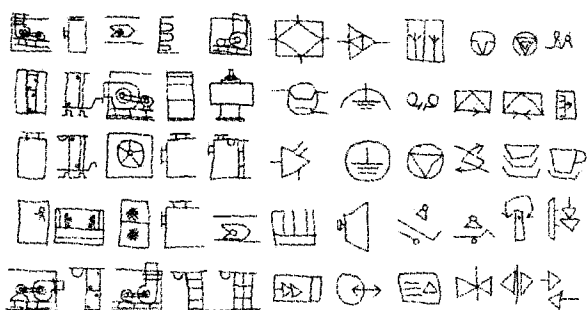


图 9 试验使用的 55 个草图类

本文采用查全率-查准率图^[14]来衡量检索的效率。查全率是根据候选集中相关结果的数目与整个草图库中相关结果的数目的比值, 其计算如下:

$$recall = \frac{|relevant \cap candidates|}{|relevant|}$$

查准率是候选集中相关结果的数目与整个候选集中结果数目的比值, 其计算如下:

$$precision = \frac{|relevant \cap candidates|}{|candidates|}$$

显然, 在查全率-查准率图中, 曲线越高, 检索效果越好。因为相同的查全率下, 曲线越高代表查准率越高, 反之亦然。

在单用户试验中, 我们采用的是 10 个使用者各自单独的草图库, 用户输入检索草图后在自己的草图库上进行检索。我们取 10 个不同使用者的试验结果的平均值作为最后的试验结果。在多用户试验中, 我们把 10 个独立的草图库合并到

新的查询点向相关结果集移动的速度由参数 α 和 β 控制, 参数满足条件: $\alpha + \beta = 1$ 。为了防止过分适应 (Over-fitting) 相关结果集, 所以要保留原始查询向量, 即 α 不直接取 0。在我们的系统中, α 和 β 的取值分别为 $\alpha = 0.2$ 和 $\beta = 0.8$ 。

经过查询向量调整之后, 我们利用调整之后的查询向量重新进行检索。经过若干次反馈和查询向量调整后, 查询向量逐渐逼近理想值, 检索效果也更好地满足用户需求。图 8 给出了由图 6 的检索候选集经过一次反馈后检索结果。

一起, 用户输入检索请求后, 在合并后的草图库上进行检索。两个试验分别进行 3 次反馈来优化检索结果。

图 10 和图 11 是根据试验结果绘制的查全率-查准率图。其中, “No feedback” 表示的是初始查询的查全率-查准率曲线, “1st feedback”、“2nd feedback” 和 “3rd feedback” 分别表示经过一次、两次和三次反馈后的查全率-查准率曲线。从图中可以看出, 经过逐次反馈, 检索的效果越来越好, 最终得到用户满意的检索结果。

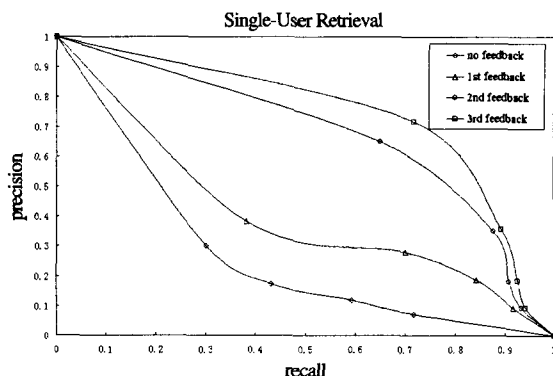


图 10 单用户草图检索试验结果

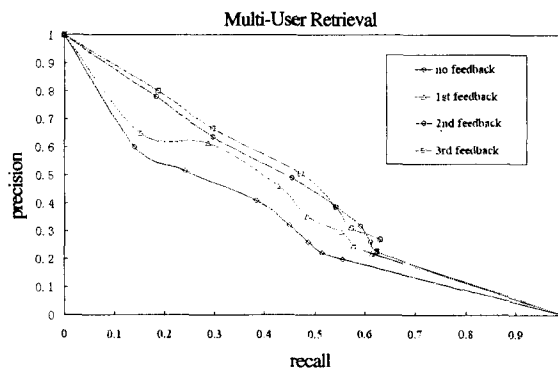


图 11 多用户草图检索试验结果

此外, 时间因素在检索系统中也是检验系统效率的关键因素之一。特别是在实时交互的检索系统中, 检索时间越短, 也就是时延越小, 越能满足实时应用需求。试验中, 单用户数

数据库的检索时间为 2.4ms,多用户数据库的检索时间为 8.7ms。总的来说,本文提出的方法能够满足实时检索应用的需求。

总结 本文提出了一种高效快速的草图检索方法。该检索方法基于草图的空间拓扑信息,对草图绘制的笔划顺序、草图的旋转、草图的大小等用户习惯并不做严格的限制。该方法抽取草图的向量化特征,大大降低草图匹配的时间和存储空间。并且,利用相关反馈技术有效地提高草图检索的效率,不断改进检索效果,最终准确捕获用户意图。然而,草图的模糊性和用户适应性是手绘草图领域中的一个难点问题。如何更好地表示草图的特征信息以及如何准确地捕捉用户的意图,还需要更加深入的研究。我们将在特征的选取和反馈的机制上做进一步的研究。

参考文献

- 1 Aref W G, Barbará D, Lopresti D P. Ink as a First-Class Datatype in Multimedia Databases. *Multimedia Database System, Issues and Research Direction*, 1996, 113~163
- 2 Kamel I, Barbará D. Retrieving Electronic Ink by Content. In: *Proc. of the International Workshop on Multimedia Database Management Systems*, 1996, 54~61
- 3 Faloutsos C, Barber R, Flickner M, et al. Efficient and Effective Querying by Image Content. *Journal of Intelligent Information Systems*, 1994, 3:231~262
- 4 Sciascio E D, Mongiello M. Query by Sketch and Relevance

- Feedback for Content-Based Image Retrieval over the Web. *Journal of Visual Languages and Computing*, 1999, 10(6):565~584
- 5 Lopresti D, Tomkins A, Zhou J. Algorithms for matching hand-drawn sketches. In: *Proc. of the Fifth Intl. Workshop on Frontiers in Handwriting Recognition*, 1996, 233~238
- 6 Kato T, Kurita T, Otsu N, et al. A Sketch Retrieval Method for Full Color Image Database. In: *Proc. of the 11th Intl. Conf. On Pattern Recognition*, 1992, 530~533
- 7 Ho Leung H W. Representations, Feature Extraction, Matching and Relevance Feedback for Sketch Retrieval; [Ph. D dissertation]. 2003
- 8 Fonseca M J, Jorge J A. Towards Content-Based Retrieval of Technical Drawings through High-Dimensional Indexing. In: *Proc. of the 1st Ibero-American Symposium in Computer Graphics (SIACG'02)*, 2002, 263~270
- 9 Feng Y J, Xing S Z. Temporal-based Multi-Stroke Sketchy Graphics Recognition. In: *Proc. of Casia Annual Conf. on Automation and Information Technology*, Sept, 2004, 538~545
- 10 Cvertkovi c D, Rowlinson P, Simi c S. *Eigenspaces of Graphs*. Cambridge; Cambridge University Press, 1997
- 11 Shokoufandeh A, Dickson S, Siddiqi K, et al. Indexing Using a Spectral Encoding of Topological Structure. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 1999, 491~497
- 12 Aigrain P, Zhang H, Petkovic D. Content-Based Representation and Retrieval of Visual Media; A State of the Art Review. *Multimedia Tools and Applications*, 1996, 3:179~202
- 13 Rocchio J J. Relevance feedback in formation retrieval. In: Gerard S ed. *The Smart Retrieval System; Experiments in Automatic Document Processing*, 1971, 313~323
- 14 Van Rijsbergen C J. *Information Retrieval*. Butterworths, London, 1979

(上接第 215 页)

表 3 SMA 方法在实验 C 中的检索结果

K	r	查全率%				AVRR			
		0rf	1rf	2rf	3rf	0rf	1rf	2rf	3rf
1	100%	13.3	14.0	14.5	14.7	27.5	25.6	24.7	24.4
	80%	13.3	18.8	20.5	21.3	27.5	21.6	21.2	21.3
	50%	13.3	19.7	21.7	22.5	27.5	22.0	20.4	19.7
	20%	13.3	17.7	19.9	20.9	27.5	20.6	19.3	18.9
2	100%	10.7	12.5	13.3	13.7	40.3	35.5	34.6	34.2
	80%	14.1	20.0	21.8	22.5	24.5	20.7	19.9	19.9
	50%	11.9	19.2	21.9	23.0	25.4	20.8	19.0	19.1
	20%	5.6	9.6	12.5	14.1	21.0	18.0	16.0	15.4
5	100%	11.3	13.0	13.5	13.8	44.1	37.3	35.9	35.3
	80%	20.5	24.1	25.0	25.3	23.0	21.1	20.8	20.9
	50%	19.1	24.4	25.9	26.5	24.1	20.2	19.6	19.0
	20%	13.9	19.3	22.2	23.3	23.1	18.4	17.5	16.8

表 4 SMA 方法在实验 C+T 中的检索结构

K	r	查全率%				AVRR			
		0rf	1rf	2rf	3rf	0rf	1rf	2rf	3rf
1	100%	18.2	21.2	22.0	22.4	27.9	25.8	25.4	25.3
	80%	18.2	27.0	30.5	32.1	27.9	23.3	22.9	22.3
	50%	18.2	27.8	31.4	33.7	27.9	23.9	22.5	22.8
	20%	18.2	25.6	29.2	31.4	27.9	23.8	22.5	22.4
2	100%	19.1	21.4	22.4	23.0	41.0	33.7	33.1	32.4
	80%	22.7	31.6	34.3	35.5	28.2	23.8	23.3	23.3
	50%	20.1	30.9	34.7	36.5	27.9	23.3	22.6	22.4
	20%	11.4	20.4	25.1	27.7	27.5	21.3	20.1	20.1
5	100%	19.8	22.2	23.1	23.7	43.3	34.7	32.2	31.5
	80%	28.8	36.5	38.6	39.6	26.5	24.8	24.1	24.0
	50%	27.2	36.7	40.0	41.5	26.2	24.4	23.8	23.7
	20%	20.0	30.1	34.3	36.3	25.5	22.4	21.8	21.6

结论 本文提出了一种基于相关反馈和样本散布矩阵的图像检索方法,实验结果说明对多例图像查询有较好的检索效果。但是如何评价不同投影方向在检索过程中的作用,还值得进一步研究。另外,还可以结合反例图像的分布来寻找更好的投影空间。

参考文献

- 1 徐杰,施鹏飞. 基于内容的图像检索技术. *中国图像图形学报*, 2003, 8(9):977~983
- 2 许月华. CBIR 系统中的 Bayesian 相关反馈技术研究; [学位论文]. 合肥:中国科技大学, 2004
- 3 Zhang L, Lin F Z, Zhang B. Support Vector Machine Learning for Image Retrieval. In: *Intl. Conf. on Information Processing*, 2001, 721~724
- 4 Cox I J, Minka T P, Papathomas T V, Yianilos P N. The Bayesian Image Retrieval System, PicHunter; Theory, Implementation, and Psychophysical Experiments. *IEEE Trans on Image*

- Processing, 2000, 9(1):20~37
- 5 Vasconcelos N, Lippman A. Learning From User Feedback in Image Retrieval Systems. *Advances in Neural Information Processing Systems*, 2000, 12
- 6 Ishikawa Y, Subramanya R, Faloutsos C. Mindreader; Query databases through multiple examples. In: *Proc. of the 24th VLDB Conf.*, 1998, 651~675
- 7 Rui Y, Huang T S. Optimizing Learning in Image Retrieval. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2000, 6:236~243
- 8 Zhou X S, Huang T S. Relevance feedback for image retrieval; a comprehensive review. *Multimedia Systems*, 2003, 8(6):536~560
- 9 Duda R O, Hart P E, Stork D G. *模式分类. 第二版*. 北京:机械工业出版社, 2003, 94~96
- 10 Muller H, Muller W, Squire D, et al. Performance Evaluation in Content-based Image Retrieval; Overview and Proposals. *Pattern Recognition Letters*, 2001, 22:593~604