

基于内容的视频检索关键技术研究综述

蒲筱哥

(徐州师范大学 图书馆, 江苏 徐州 221116)

摘要:对视频数据模型、视频结构化、镜头聚类 and 场景提取、视频索引和浏览等基于内容的视频检索关键技术的当前研究现状做了深入地分析,并总结了基于内容的视频检索关键技术当前研究中存在的问题以及其研究的未来发展趋势。

关键词:视频检索;镜头边界检测;关键帧提取

中图分类号:TP391;G354 **文献标识码:**A **文章编号:**1007-7634(2010)03-0464-06

A Survey on Key Techniques of Content-based Video Retrieval

PU Xiao-ge

(Library of Xuzhou Normal University, Xuzhou 221116, China)

Abstract: This article has analyzed the Insufficiency and trend of current research progress of key techniques of content-based video retrieval, which beside t Video data model, video structurization, shot cluster and the scene extraction and the video frequency index and the browsing and so on.

Keyword: video retrieval; shot boundary detection; key frame extraction

基于内容的视频检索技术是针对音视频这类非结构化数据,使用了自动数字化、语音识别、镜头检测、关键帧抽取和内容自动关联等技术,真正做到了从内容上对视频进行搜索。而其中视频数据模型、视频结构化、场景构造、视频索引和浏览等是其中最具有代表性的关键技术。要实现基于内容的视频检索,首先必须进行视频镜头分割、关键帧提取、镜头聚类,经过这些处理,然后才能通过对视频段之间特征空间的比较来进行视频段内容的比较。然而由于视频内容繁多且复杂,对视频的检索十分困难。从20世纪90年代初,国际上就开始了对于基于内容的视频检索方面的研究,但到目前基于内容的视频检索关键技术方面的研究仍然存在着不少问题。

1 视频数据模型

数据模型是数据库系统中用于提供信息表示和操作手段的形式构架。在视频数据库管理系统中,视

频数据模型起着十分重要的作用,是基于内容的视频分析与检索的基础。由于视频数据内容的丰富性、多样性、结构的复杂性,在视频模型建立中必须考虑一些特殊的问题,所以有很多视频数据模型被提出,当前有基于分段的模型、基于层次注释的模型、面向对象的模型、代数视频数据模型和其它视频数据模型,而顺序单元模型^[1]和层次化模型^[2]是目前最主要的两种视频数据模型。

2 视频结构化

视频结构化是指通过镜头边界的检测,把视频分割成基本的组成单元—镜头,镜头边界检测是建立视频结构的第一步,是视频结构分析的主要内容,直接影响到视频检索的成败。

2.1 镜头检测

镜头检测是将视频自动地分割为镜头,以作为

收稿日期:2009-05-26

作者简介:蒲筱哥(1972-),男,陕西汉中人,馆员,硕士,从事网络信息资源开发与利用研究。

基本的索引单元,因此镜头的自动分割是视频结构化的基础。现有的关于镜头边界检测已经提出了很多算法,大体上可将这些算法分为两大类:基于解压的全图像序列的算法和直接基于压缩视频的算法。

2.1.1 基于解压的全图像序列的算法

该类算法都是基于阈值的。当前帧间差别主要通过直方图间、像素间或块间的差值而获得,也可以通过双值门限比较、对运动矢量的分析以及基于聚类/模糊聚类的检测等方法而获得,如果这个差值超过预先设定的阈值则就说明镜头发生了转换。

(1)基于直方图的算法。直方图描述了一幅图像的灰度和颜色的分布情况,通过直方图的相似度可以判断图像之间的相似情况,并且以此来判断是否有场景切换。灰度/彩色帧 f 的亮度/彩色直方图是一个 n 维向量, n 是灰度/色彩层次数, $H(f, k)$ 是帧 f 中属于某一灰度/色彩层次 k 的像素数。

Tonomura^[3]提出了基于灰度直方图的算法。Zhang等^[4]提出了一种所谓“二值比较”的方法来检测镜头转换。由于直方图法使用像素亮度和色彩的统计值,不考虑像素的位置信息,抗噪声能力较强,但有时会漏掉场景切换。

(2)基于像素的算法(模板匹配法)。此方法将两帧对应像素差的绝对值之和作为帧间差,若大于某阈值便认为有镜头切换。这种方法对检测镜头突变很有效,但对镜头中的运动物体,也可能计算出大于阈值的帧间差,造成错误判断,所以阈值较难确定,该方法也容易受噪声影响。

(3)基于运动矢量的技术。从视频序列中估计出来的运动矢量表现了单个镜头中相对连续的改变,这种连续性在不同镜头之间会被破坏,利用这一点可以检测镜头边界。运动矢量的估计有两种方法^[5]:一是传统的基于块的运动估计,这种方法在基于块的编码方法中已经被证明是有效的;二是基于点的运动估计,这种方法主要在基于网格的编码技术中使用,即估计关键点的运动矢量,这一技术也是目前发展的方向之一。

(4)双值门限比较法。双值门限比较的方法适用于像分解(dissolve)、淡入(fade in)、淡出(fade out)这类缓变的镜头的检测。该方法设置两个门限值 T_b 和 T_s ,当帧差大于 T_b 时,存在镜头突变。当帧差大于 T_s 而小于 T_b 时,存在镜头缓变,当相邻帧帧差超过 T_s 时,该段连续帧的首帧称为镜头缓变的起始帧。计算两种帧差:一种是相邻帧帧差 F_{dl} ,另一种是相隔 L 帧的帧差 F_{dL} ,从缓变起始帧开始逐渐增加 L ,

显然相隔帧的帧差随 L 的增加而增加,因为相隔帧帧差是一个累计帧差,当相隔帧的帧差累计超过 T_b 而相邻帧帧差 F_{dl} 低于 T_s 时,这一帧便为镜头缓变的终止帧^[6]。

(5)基于聚类/模糊聚类的检测方法。该方法不但可用于检测镜头突变也可用于检测镜头缓变。基于模糊聚类的镜头分割方法先将一段视频进行模糊聚类后便得到各帧属于明显变化(SC)和非明显变化(NSC)两类场景的隶属度。如果某帧属于 SC 的隶属度大于它属于 NSC 的隶属度,则该帧属于明显变化类,并用 1 表示,反之用 0 表示。这样便把视频表示成二进制序列。视频中镜头的突变和渐变具有一定模式,因此可对二进制序列进行模式判别,便可以以此来检测镜头的突变与缓变。

2.1.2.直接基于压缩视频的算法

由于越来越多的视频序列是以压缩的形式保存的,如何对这些压缩形式的视频序列进行镜头边界检测已经成为一个当前急待解决的问题。这方面的研究才刚刚开始,相应的方法和手段还比较少,而且效果也并不理想。

(1)基于学习的算法。文献[7]认为应将复杂的镜头变换特别是叠化视为模式识别和机器学习问题。这样做势必会需要许多学习的例子,而搜集大量的复杂镜头变换的例子并不容易(因为其在镜头变换中占的比例很小),因此可以建立复杂镜头变换的合成系统(transition synthesizer system):第一,在真实的视频中其转换也是有差不多的方法生成;第二,如新的方法对于合成的转换都不能很好地识别,就更别说是实际的视频了,识别合成转换是必要条件而非充分条件;第三,最终算法都要在实际的视频中进行测试以确认其效果。可见,建立复杂镜头变换的合成系统不失为一个新的研究方向。

(2)子带分解算法。Lee和Dickinson^[8]提出了一种在子带域中基于直方图的边界检测技术。首先对要比较的两帧图像作子带分解,然后对各级子带的直方图从低分辨率到高分辨率作分层比较。这种技术的优点是计算复杂度比较低,如作2级子带分解,子带级2的大小只有原始图像的1/16。

(3)基于确定变换模型的算法。前面所介绍的算法都是完全基于图象处理技术的。实际上,可以通过建立镜头变换的精确数学模型来帮助提高 SCD 识别率和识别速度^[9],这一类方法属于自上而下(Top-down)^[10]的方法。它的优点是可以建立一套基于数学模型的系统的方法,而且对于特定领域的视频图象

可以通过在数学模型上加上一一定的限制条件来提高方法的有效性。这一类方法的表现主要取决于它们所基于的数学模型。

2.2 关键帧提取

关键帧是反映一组镜头中主要信息内容的一帧或若干帧图像,关键帧的作用类似于文本检索中的关键词。用关键帧来代表镜头,使得对视频镜头可用图像的技术进行检索。关键帧的提取方法可以分为在非压缩域和压缩域中的方法。

2.2.1 非压缩域关键帧提取算法

(1)基于镜头的方法。一段视频流被分割成若干镜头后,一种最直接、最简单的关键帧提取方法就是将每个镜头的首帧、中间帧以及末帧作为镜头的关键帧^[11]一般说来,从镜头中选取一帧或固定数目的关键帧的方法并不是很好,因为当处理变化很少的镜头时,这样选取的关键帧过多,而对于运动较多的镜头,用一两个关键帧又无法充分描述其内容。一个比较有效的方法是 Zhang^[12]提出的基于颜色直方图来提取关键帧。此外, Yeung^[13]采用了通过计算特征空间的距离来提取关键帧的方法。

此种方法简单,而且无需计算,但效果不稳定,因为首帧、中间帧及末帧并不一定能反映镜头的内容,对于比较复杂的镜头更是如此。

(2)基于聚类的方法。聚类是一种被广泛应用在模式识别、语音分析和信息检索中的非常有效的技术。使用聚类分析提取关键帧主要的思路是首先初始化一个聚类中心,然后根据当前帧与中心的距离来判断是归为该类还是作为新的聚类中心,最后将各类中离聚类中心最近的帧作为关键帧。这种基于聚类的关键帧提取算法不仅计算效率高,它还有有效的获取视频镜头的显著视觉内容。该方法仅依赖于当前帧与前面的帧,易于在线实现。

文献【14】提出先从每个镜头中找出一个代表帧,再把所有这些代表帧放在一起做聚类,最后从每个类里面选出一帧作为关键帧。最近又有人提出了一种基于时间约束的关键帧选取算法, Chi 等将直方图与聚类结合提出了基于直方图的模糊 C-均值聚类算法(HBFCM),这种方法的优点是不需预先设定阈值。而 Hanjalic 和 Zhang 提出了一种带有聚类正确性分析的划分聚类算法^[15],它为镜头选择最优的聚类数,最靠近聚类质心的帧选为关键帧。基于聚类的算法力图通过聚类的方法消除镜头间的相关性,但是却没有很好的保留原来镜头内图像帧的时

间顺序和动态信息,因此从这个意义上讲,这些算法在对视频序列进行关键帧提取的时候也是有缺陷的。

(3)基于运动分析的方法。Wolf^[16]通过光流分析来计算镜头中的运动量,在运动量取局部最小值处选取关键帧,它反映了视频数据中的一个“静止”特点,视频中通过摄像机在一个新的位置上停留或通过人物的某一动作的短暂停留来强调其重要性。但这个算法也存在问题:一是由于算法依赖于局部信息,所以鲁棒性不强;二是算法没有足够地重视由动态累加带来的内容变化。Divakaran 在 Wolf 方法的基础上进行了改进,该方法是用 MPEG-7 标准的运动行为描述符的强度作为视频序列摘要的测量^[17]。而 Divakaran 和 Sun 在着手研究有关运动行为的空间分布的描述符^[18]。

(4)帧平均法和直方图平均法。帧平均法和直方图平均法是关键帧提取的经典方法^[19]。帧平均值法取一个镜头中所有帧的某个位置上的像素值的平均值,将镜头中该点位置的像素值等于平均值的帧作为代表帧;直方图平均法即将镜头中所有帧的统计直方图取平均,选择与该平均直方图最接近的帧作为代表帧。这些方法的优点是计算比较简单,所选取的帧具有平均代表意义。缺点是从一个镜头中选取一个关键帧,无法描述有多个物体运动的镜头。

(5)基于图论的方法。这是关键帧抽取理论上的最新进展,该方法将视频看成高维特征空间上的点。这样,抽取关键帧就是在这些点中选取一个子集,这个子集中的点或者能在指定特征距离内覆盖其它点,或者反映了镜头内容上的显著变化。Chang^[20]将镜头看成邻接图,每一帧都是图上的一个顶点,因此关键帧的抽取问题就等价于顶点覆盖问题,也即希望找到一个最小的顶点覆盖以使顶点和它们的邻接点之间的特征距离最小,因此使用了一个次优化方法。与 Chang 观点不同, Dementhon 和 Zhao^[21]将关键帧抽取转化为曲线分割问题。他们把镜头看成特征轨迹或高维空间点组成的曲线,所以关键帧选取转化为检测曲线结合或分裂的点。基于这种想法, Dementhon 进一步提出了利用递归的标识曲线连接点得到分层次的关键帧的观点。

2.2.2 压缩视频关键帧的提取算法

基于非压缩域关键帧提取算法都是直接分析镜头内的帧,对一些非压缩的视频帧很方便,然而目前很多视频都是以 MPEG 等压缩形式存取的,如果将上述算法应用到 MPEG^[22]压缩流,需要先还原每一

帧图像,这样需要很大的计算量。文献[23]提出的一种思路是通过检测 MPEG 压缩视频流中已有的离散余弦变换(DCT)的 DC 系数和运动向量(MV)来提取关键帧,而无需全部解压缩。最近随着 MPEG7 标准的提出,有人提出了基于镜头活动性提取关键帧的方法。首先计算内部帧和参考帧的直方图,然后计算出一个活动指示量,根据活动性的曲线,选取曲线中局部最小值的帧作为关键帧。

3 镜头聚类 and 场景提取

在镜头聚类及场景生成过程中,镜头不仅在时间上是连续的,更重要的是它们在内容含义上是一致的,这是镜头聚类的关键。视频聚类的过程也就是镜头匹配的过程,即在一组特征参数度量下将相似的镜头合并为镜头组,进而聚类生成对应的场景。很多种方法都可以用于这里的特征聚类,例如 K-均值法、ISODATA 法、松弛迭代法、基于关联规则的算法、基于模糊图论聚类法等。最著名的工作来自 Princeton 大学, B. L. Yeo 和 M. M. Young 提出用时间约束的聚类方法^[24]对镜头代表帧聚类,并根据聚类结果的时间特性探测对话、动作和一般故事单元。这种方法有两个问题:一个是固定的时间约束给故事单元探测带来人为的约束;另一个是镜头聚类过程中采取了难以控制的阈值约束。文献[25]为了解决自动电影摘要的问题,设计了从电影中提取对话、大运动、高对比度片段的方法,但是并没有研究对整段视频分段的问题。

场景一般由多个镜头组成,实际中也可以先确定场景中的关键镜头,再用这些镜头的关键帧来代表场景。要想有效地表达场景,一种直观的方法就是选择场景中最主要的镜头,用这个镜头作为场景的代表。但仅用一个镜头代表场景容易丢失其他镜头的信息,不够全面;另一方面,一个完整镜头的数量也是相当庞大的,用来表达一个情节不够紧凑。由于对每个镜头可以提取关键帧,所以可从场景所包含的镜头关键帧集合中选择代表该场景的代表帧。

4 视频数据索引与浏览

4.1 视频索引

视频索引是将视频内容分析,视频结构分解和摘要的过程中提取的结构和内容属性归入同类别或

进行索引结构的聚类建立的属性表。视频索引可分为三类:基于注释的索引,基于特征的索引以及基于特定领域的索引。基于注释的索引,也称为高级索引。基于特征的索引也可称之为低级索引。基于特定领域的索引是指专门针对某个应用领域建立的索引,它们一般有固有的模式。

当前高维索引方法很多,也都有各自的优点。文献[26]提出了一种特殊的二叉树索引法,在 KD-树中,结点的值用一个多维向量而不是一个单一的数值来表示,每个特征向量代表了数据库中的一个对象。文献[27]是对固定网格结构的扩展。在固定网格结构中,一个 n 维的空间被划分成为大小相同的网格矩阵,每个网格矩阵中包含了零个和多个特征向量。它们通过和网格矩阵相联系的指针进行访问。R-树及其变种^[28]是最广泛使用的一种多维索引结构。R-树存在的问题是每个叶节点只指向一个区域,而这个区域有可能被许多不同的矩阵所覆盖,这样就会导致仅跟踪结构的一条通路进行搜索时会遇到失败的情况。SS-树^[29]也是通过改进 R-树得到的,对于每个节点,SS-树利用最小边界球来代替最小边界矩阵,降低了存储空间,提高了最近邻检索的性能。但是,在高维数据集合中,SS-会增加覆盖,从而异致最近邻检索性能的降低,提高了检索性能,但是树的更新更加复杂。

4.2 视频浏览

视频浏览是视频检索系统中交互查询的一个组成部分。为了有效地浏览,视频文档的内容应表示成用户易于理解的静态画面的形式,并且必须提供非线性的访问。通常每个镜头的关键帧被用作为“浓缩”了的视频序列;然而,在许多视频中,常常有几百个镜头。另外仅用静态的画面常常不足以表示动态的信息;因此仅将代表帧排列起来的方法无法满足用户有效的浏览要求。己有不少文章就结构化的浏览进行了研究,其中典型的方法有简单层次浏览、视频内容的目录结构、场景转换图、视频摘要等。

5 当前研究存在的问题及未来的发展趋势

5.1 当前研究存在的问题

基于内容的视频检索研究已有十多年的历史,已取得了不少成果,对它的应用也已全面展开,同时基于内容的视频检索又是一个具有相当挑战性的研究课题,当前的研究还存在某些需要进一步解决的

课题:

5.1.1 视频的结构化问题

视频具有非结构化的特点,这就要求在基于内容的检索系统的设计过程中首先解决视频的结构化问题。合理的结构化表示将有助于后续的特征和内容分析及用户检索,但是怎样划分具体的结构仍然是值得探讨的问题^[30-31]。

(1) 特征向量的选择。与图像分割问题一样,正确的特征选择对视频分析具有十分重要的意义。像素差和边缘变化率对运动和噪声比较敏感,会造成误检测;而颜色直方图又丢失了位置信息,两幅完全不同的图像可能具有相似的颜色分布,会造成漏检测。找到更好的特征或特征组合将更有利于镜头边界的检测。

(2) 阈值的选择。阈值选择是利用帧间差的镜头分割算法的一个重要问题。阈值过大会漏掉镜头转换,阈值太小会引起误检测,则把镜头内摄像机或物体的运动(此时帧间差值增大)误检测为镜头转换。不同类型的视频应选择不同的阈值,阈值应根据视频的内容自适应的选定,当前研究中大多数算法都采用依靠经验人工选择阈值的方法,这不利于镜头分割的实现。

(3) 渐变与镜头运动的区别。渐变与镜头运动都会造成帧间差连续的增大,从本质上说利用帧间差的方法无法从根本上区别渐变和镜头运动。基于模型的方法是一种可能的途径,但是为各种渐变建立起模型也非常困难。

(4) 闪光灯及光照条件的变化引起的误检测。闪光灯及光照条件的变化都会造成视频帧亮度的变化,引起各种视频特征的变化,从而容易导致误检测为镜头边界。

5.1.2 有效的特征提取问题

传统的文本数据库的检索可以用关键字,是因为其形式单一,信息量小。而对于结构复杂、含有大量信息的视频数据,需要从多方面提取其客观低级特征,并从低层次的视觉听觉特征中提取高层次的语义信息^[32]。

5.1.3 视频的底层特征和高层语义的结合问题

从视频数据中获得的低层视觉特征与用户自身对数据理解的不一致而出现的“语义鸿沟”(semantic gap)是目前基于内容的视频检索系统难以被普通用户所接受的根本原因,如何建立这些底层的特征与高层语义概念的关联,从而使计算机自动抽取视频语义是当前研究中的难点所在。

5.2 未来发展趋势

(1) 检索技术的多特征融合化。随着信息化进程的深入,需要对数字音频、语音和音乐等进行基于内容的检索,对合成媒体如动画、VRML 数据进行检索等。在研究单一媒体的检索同时,注意研究多种媒体的互相关联和互补关系,以提高检索算法的效率。而如何有机地组织这些多种视频特征,使应用能够调用合适的特征和特征表示来支持查询,从而达到较高的检索率,是今后研究的一个必然趋势。

(2) 视频高层语义和低层特征的结合化。CBVR 要走上实用化之路就必须解决将高层语义(人的理解)和视频的低层特征有机的结合,从而使得检索能直接基于人的理解来进行,然而这这将是一个需要长期解决的问题。

(3) 多维化索引技术。多维索引技术是从如何改进索引算法以适应大规模数据库的角度来考虑问题的,是 CBVR 在大型图像和视频数据库中发挥优势的保证。在 CBVR 中,用于表述图像内容的视觉特征有其自身的表示方法(如直方图等),这无疑大大增加索引的难度,如果利用多维化索引技术,从语义层次上对图像进行有效的分类,对于高效的系统索引处理和精确的用户查询都将起到重要的作用。

(4) 视频检索反馈的交互化。信息获取过程的交互性是 CBVR 的一个重要特征。反馈是一种调整技术,用以适应用户需求和提高检索精度的常用手段。在借助语义概念进行检索时,反馈是实现人机交互,进而将用户知识结合进查询中的重要方法。因此,如何更好地实现人机交互、实现用户查询接口的智能化是今后发展的一大趋势。

(5) CBVR 的网络化。从未来应用的需求来看,Web 上视频信息的检索是当前对视频基于内容的组织和处理所面临的又一个问题。面向 Web 的 CBVR 大致可分为 WWW 服务器、查询处理器、视频数据库和特征提取器四大部分。WWW 服务器用于互联网上的通信,查询处理器用于处理客户的查询需求,特征提取器则是负责视频特征提取的模块,查询处理器涉及到对视频语义分析技术的研究,视频数据库则涉及到具有视频特征含义的视频数据库的组织,然而这些都还有待进一步的探索。

综上所述,目前在基于内容的视频检索领域尚存在许多问题,不可能一蹴而就。从当前研究来看,在理论领域,目前视频数据的结构化、视频数据模型、场景构造、视频索引和浏览等关键技术已经取得

了很大进展,但各项技术的实用化还有待提高,同时应在理论方面深入探讨视频语义特征的构造和有效的检索机制,从而为促进镜头检测和特定领域的视频检索走向实用化和网络化打下坚实的基础。

参考文献

- 1 B.Rubin and G Davenport.Structured Content Modeling for Cinematic Information [J].SIGCIHI Bulletin,1989,21 (2):51-57.
- 2 G.Davenport,T.A.Smith,N.Pincever.Cinematic Primitives for Multimedia [J].IEEE Computer Graphics&Applications,1991,(7):67-74.
- 3 Y.Tonomura,Video handling based on structured information for hypermedia systems[C].ACM Proceedings'91,International Conference on Multimedia Information Systems, McGraw - Hill, Inc. USA:New York, NY, 1991:333-344.
- 4 H.J.Zhang,A.Kankanhalli,S.W.Smoliar,Automatic partitioning of full-motion video [C].ACM Multimedia Systems,1993:10-28.
- 5 Gerald Kühne,Stephan Richter,Markus Beier.Motion -based Segmentation and Contour based Classification of Video Objects.Proc [C].ACM Multimedia, New York, NY:ACM Press, 2001:56-59.
- 6 刘政凯,汤晓鸥.视频检索中镜头分割方法综述[J].计算机工程与应用,2002,(23):84-87.
- 7 Rainer Lienhart.Reliable Dissolve Detection In Storage and Retrieval for Media Databases 2001 [J].Proc.SPIE,2001,(1): 219-230.
- 8 J.Lee and B.W.Dickinson.Multiresolution video indexing for subband coded video databases.In IS&T/SPIE[C].Conference on Storage and Retrieval for Image and Video Databases, San Jose, CA, USA ,1994:111-115.
- 9 Arun Hampapur,Ramesh Jain and Terry Weymouth,Digital Video egmentation [C].Proceedions second Annual Acm Multimedia Conference and Exposition, SanFrancisco,CA, USA ,1994:65-68.
- 10 P.Aigrain and P.Joly,Automatic Real-time Analysis of Film Editing and Transition Effects and its Applications[J].Computer and Graphics 18, 1994,(1):93-103.
- 11 A. Nagasaka and Y Tanaka.Automatic video indexing and full-video search for object appearances[C]. Budapest: Proceedings of the second Working Conference on Visual Database Systems. 1991: 119-133.
- 12 H. Zhang, J. Y A. Wang, and Y Altunbasak.An integrated system for content-based video retrieval and browsing [J] . Pattern Recognition, 1997,30(3): 643-648.
- 13 M. M. Yeung and B. Liu. "Efficient matching and clustering of video shots," [C].Proceedings of IEEE ICIP, NY,USA: IEEE ,1995: 338-341.
- 14 Chi C L,Shuenn J W.Video Segmentation Using a Histogram-based Fuzzy C-means Clustering Algorithm[J].Computer Standards&Interface,2001,23,(5):429-438.
- 15 Hanjalic A,Zhang H J.An Integrated Scheme for Automated Video Abstraction Based on Unsupervised Cluster-validity Analysis [J].IEEE Trans on Circuits and Systems for Video Technology,1999,9(8):1280-1289.
- 16 W.Wolf.Key frame selection by motion analysis[J].Acoustics, Speech, and Signal Processing, 1996.ICASSP -96.Conference Proceedings,1996 IEEE International Conference on, Volume: 2,7-10,1996,(4):1228-1231.
- 17 A. Divakaran, R. Regunathan, and K. A. Peker.Video Summarization Using Descriptors of Motion Activity: A Motion Activity Based Approach to Key -Frame Extraction from Video Shots [J].Journal of Electronic Imaging, 2001,10(10): 909-916.
- 18 A. Divakaran and H. Sun.Descriptor for spatial distribution of motion activity for compressed video [J]. Proceedings of SPTE on Storage and Retrieval for Media Databases 2000, 2000,3972(1): 24-28.
- 19 许 源,薛向阳.一种视频局部高层语义特征提取算法[J].计算机科学,2006, 33(11): 134-138.
- 20 H S Chang, S S Sull,S U Lee. Efficient Video Indexing Scheme for Content-based Retrieval [J].IEEE Trans on Circuits and Systems for Video Technology,1999,9(8):135-139.
- 21 L Zhao,W Qi,S Z Li et al.Key-frame Extraction and Shot Retrieval Using Nearest Feature Line (NFL)[C].In:International Workshop on Multimedia Information Retrieval, Geneva, Switzerland ,2000:57-61.
- 22 杨 胜,钟玉琢.一种从 MPEG 压缩视频流中提取关键帧的方法[J].中国图像图形学报,2001,6(3):254-258.
- 23 朱映映,周洞汝.一种从压缩视频流中提取关键帧的方法 [J].计算机工程与应用,2003,(18):13-14,48.
- 24 P. Aigrain,P. Joly,V. Longueville, Medium Knowledge Based Macro Segmentation of Video into Sequences[M]. Chapter 8,in Intelligent Multimedia Information Retrieval, MIT Press Cambridge, MA, USA ,2000:89-93.
- 25 D. Zhong, H. J. Zhang and S-F. Chang. Clustering Methods for Video Browsing and Annotation [C], Proc. of SPIE Conf. on Storage and Retrieval for Image and Video Databases IV, San Jose,CA. Calif.USA, 1996:46-51.
- 26 H. B. Lu,Y.,Zhang,Y. R. Yao. Robust Gradual Scene Change Detection [C]. In Int.. Conf. On Image Processing, volume 3,Yokoama, Japan,pages 1999: 304-308.
- 27 J.Nievergelt, H. Hinterger, R. Midstraumm. Web STAR-Video Database on WWW[C], Proc. of Multimedia Computing and Networking 2000, San Jose, California, USA, 2000: 24-26.

(下转第 476 页)

保证搜索正确性和真实性,每个搜索引擎都有它的长处,山寨搜索引擎的理念是不错的,可以满足一部分用户的需求,但从商业模式上讲是不成熟的。山寨搜索引擎必须走由模仿到创新,最后到可持续发展的道路,发扬山寨搜索引擎结果丰富等优点,同时改进其检索时间长等不足,特别是不支持高级搜索的缺陷。而在其发展趋向方面,专业化搜索、个性化搜索将成为山寨搜索引擎主要的发展方向。

参考文献

- 1 2008 搜索引擎市场调查报告 [EB/OL]. <http://tech.sina.com.cn/focus/2008ssdc/>, 2008-05-07.
- 2 唐 琴,等.搜索引擎发展阶段研究及热点发现[J].情报学报,2008,(10):664-669.
- 3 谷萌萌,高茂庭.搜索引擎研究与发展[J].计算机与数字工程,2008,(7):74-77.
- 4 张卫峰,等.元搜索引擎研究[J].计算机科学,2001,(8):36-41.
- 5 赵立刚.搜索引擎的研究与设计[D].长春:吉林大学,2005:5-9.
- 6 王 芳,张晓林.元搜索引擎原理与利用[J].现代图书情报技术,1998,(6):18-19.
- 7 陈伟雄.基于元搜索的中文搜索引擎研究与实现[D].北京:清华大学,2004:77-83.
- 8 李广建,等.元搜索引擎及其主要技术[J].情报科学,2002,(2):173-179.
- 9 郭少友.元搜索引擎的原理与设计[J].情报科学,2005,(2):245-248.
- 10 门风超,等.元搜索引擎的实现技术与发展趋势[J].现代情报,2008,(7):61-62.
- 11 丁成杰.搜索引擎技术的研究与实现——元搜索引擎和文本聚类[D].上海:上海交通大学,2006:20-36.
- 12 王 亮.搜索引擎及其相关性排序研究[D].武汉:武汉大学,2004:27-42.
- 13 王 敏,杨炳儒.基于主题的个性化元搜索引擎的设计与实现[J].情报杂志,2005,(7):43-81.
- 14 曾伟忠,徐 昕.搜索引擎及元搜索引擎工作原理及存在的不足[J].图书馆学刊,2004,(5):58-59.
- 15 李志文.搜索引擎发展中的问题与对策[J].情报科学,2002,(5):556-558.
- 16 武助宇,刘文青.中文搜索引擎发展趋势 [J]. 情报科学,2002,(9):990-992.
- 17 赵 静,王玉平.目前我国搜索引擎研究的现状与发展[J].情报科学,2003,(8):877-879.
- 18 王 敏,杨炳儒.基于主题的个性化元搜索引擎的设计与实现[J].情报杂志,2005,(7):43-81.

(责任编辑:刘凤勤)

(上接第 469 页)

- 28 A. Gunman. R-Trees: A Dynamic Index Structure for Spatial Searching, Proc [C]. Boston:ACM SICMOD Conference,, USA, 1984:47-57.
- 29 D. White, R. Jain. Similarity Indexing with the SS-tree[C]. New Orleans:in Proc. 12th IEEE Int.Conf. On Data Engineering New Orleans, Louisiana,1996: 516-523.
- 30 Ren Wei, Weal Paul, Singh Maneesha, Singh Sameer. Visual information retrieval: Minerva video benchmark [C]. Proceedings of the Third IASTED International Conference on Signal Processing, Pattern Recognition and Applications, Academic Press, Inc. Orlando, FL, USA, 2006: 256-261.
- 31 Yang Xiaoyan, Xiang Hui, Xu Hui. MPEG-7 based video retrieval technology and its application in digital home system[J]. Journal of Computational Information System, 2006, 2(2):707-712.
- 32 Liu J, Bhanu R. learning semantic visual concepts from video. In[C]:IEEE Proceedings of 16th International Conference on Recognition Pattern. Quebec: IEEE Press, Calgary, Canada, 2002:1061-1064.

(责任编辑:刘凤勤)