

基于视频时间段检索的多媒体数据库系统研究^①

Research on Multimedia Database System Based on Time Slice Video Retrieval

郑力明 易平 (暨南大学 信息科学技术学院 广东 广州 510632)

摘要: 本文在对当前视频检索的数据库系统进行分析的基础上,设计了基于时间段检索的多媒体数据库系统模型。该系统在关键帧提取时增加了对时间偏移量 **offset** 的提取,进一步改进了目前视频建立索引的方法,提出了基于时间段的分层检索的概念,能为基于移动终端的短视频检索系统的研究与开发提供参考。

关键词: 多媒体数据库 关键帧 时间段索引 分层检索

1 引言

随着流媒体技术的发展与普及,视频系统的研究越来越受到人们的关注,其中多媒体数据库系统就是该领域的研究热点之一。目前,视频检索主要有两种检索方式,一是目前很成熟的基于标题和注释的文本检索方式,二是基于内容的视频检索方式。随着视频系统的发展及用户需求的提高,第一种方式已无法满足用户的需求,而第二种方式实现难度太大,技术上很多方面无法在短时间内突破。对于目前充斥在网络上的众多长视频中,大多数情况下,人们想要观看的只是某一长视频中的一段短视频。因此,如何快速地从长视频中检索出我们所需要的短视频就显得很重要。

本文结合目前大多数用户的需求及手机用户对短视频的需求,设计了一种全新的基于视频时间段检索的多媒体数据库系统模型。在继承了一般对视频镜头分割和关键帧的提取方法的同时,加入了对视频关键帧的时间偏移量 **offset** 的提取,并将该图像关键帧与 **offset** 建立对应关系,从而提出了按时间段建立索引的方法。最后提出了分层检索的概念,使得用户可以先按关键字,再按直观的视频图像检索不同视频内容的视频片段。

2 基于时间段检索的多媒体数据库系统模型

如图 1 所示,基于时间段检索的多媒体数据库系

统模型由以下几部分组成:

(1)用户接口:用于用户查询检索、上传、观看视频,并提供个性化的人机交互界面。

(2)数据库管理系统:与文本信息数据库通信,按用户检索的需求,触发“文本信息数据库”中的数据,再根据“文本信息数据库”中的特定信息从多媒体数据库中提取出满足用户需求的视频数据。

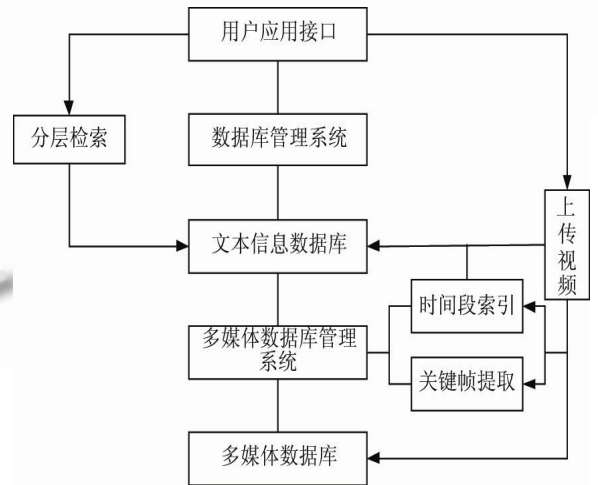


图 1 基于时间段检索的多媒体数据库系统模型

(3)文本信息数据库:“数据库管理系统”和“多媒体数据库”的桥梁,是文本信息(标题,注释等信息)和多媒体数据的交汇点。除了用户上传的视频标题,

^① 基金项目:广东省自然科学基金项目(8151063201000051);广东省科技计划项目(2006B11601001)

收稿时间:2008-10-18

注释等一般的文本信息字段外,还包括经“多媒体数据库管理系统”对视频处理后加入了特定时间视频段的“offset(时间偏移)”和“Msrc(视频文件源地址)”。

(4)多媒体数据库管理系统:整个数据库系统的关键,其中最重要的是:一是建立基于时间段方式的索引,二是对关键帧的抓取,并将关键帧的offset(时间偏移)和Msrc(视频文件源地址)插入到“文本信息数据库”。

(5)多媒体数据库:主要是存储视频文件及按特定时间段抓取的关键帧图像文件。

(6)上传视频:用户将自己的视频上传时,将视频标题及注释插入“文本信息数据库”并交由“多媒体数据库管理系统”对其进行关键帧的提取和建立时间段方式的索引。

(7)分层检索:用户可以根据不同的需求先按文本,再通过点击感兴趣的图像帧按时间段搜索到自己想要的短视频。

3 基于时间段检索的多媒体数据库系统的关键技术

3.1 镜头分割及关键帧的提取

本文提出的基于视频时间段检索方式和基于视频内容检索方式一样,其视频中的关键帧的提取,是整个系统的关键技术之一。而在提取关键帧时,对于不同镜头的分割又是实现该技术的关键。镜头分割的方法有很多种:如基于灰度差的方法,基于边缘的镜头分割方法,基于彩色直方图的分割方法,MPEG 视频的镜头分割方法等^[1],其中最为普遍的是基于边缘的镜头分割方法,而基于边缘的镜头分割方法中最为典型的三种类型分别是:帧间差检测法,模型检测法,压缩域检测法^[2]。本文结合了文献[1,3,4]的方法,提出了一种改进的提取关键帧的方法。

一般情况下,不同镜头之间的转换会引起视频内容的底层特征发生显著的变化。因此,在这种情况下,我们可以按照镜头转换处视频底层特征的变化来分割镜头^[4]。从而可以根据一定的数学模型,将不同镜头分割出来。

(1)设定一个阈值概率 t ,及设定各相邻帧间差的阈值为 T ,再计算出该视频中各连续帧之间底层视觉特征的帧间差值。

(2)记录每一帧的时间偏移量 $offset$,并默认第一

个镜头的起始帧的 $offset(0)=0$ 。

(3)帧间差值由边缘变化的百分比表示,即边缘从一帧到另一帧移进和移出的比例。设 $P(k)$ 为第 k 帧 $F(k)$ 与第 $k+1$ 帧 $F(k+1)$ 最近边缘的距离大于给定的阈值 T 的百分比,同时 $P(k+1)$ 为 $F(k+1)$ 中与 $F(k)$ 的最近边缘的距离大于给定的阈值 T 的百分比。则当帧间差为 $t(k)=\text{Max}(P(k), P(k+1))$ ^[1],当 $t(k)$ 大于给定的帧间差阈值 t 时,则将 $F(k)$ 分为第 K 个镜头 $S(k)$ 的结尾帧, $F(k+1)$ 为第 $K+1$ 镜头 $S(k+1)$ 的起始帧。

(4)记录 $S(k)$ 中的首尾图像帧 ($F(k-1), F(k)$) 的时间偏移量 $offset(offset(k-1), offset(k))$,并存入文本数据库的相应字段,如图 2 所示。

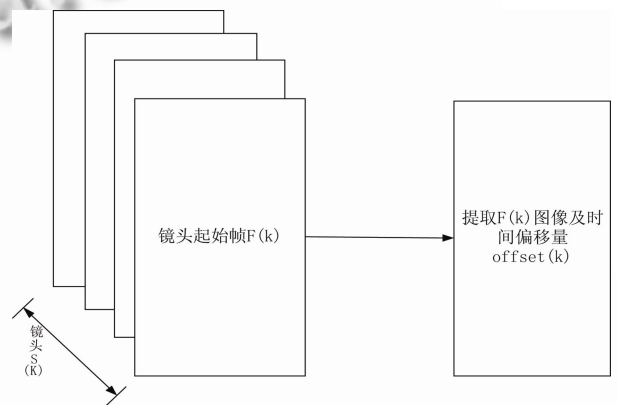


图 2 镜头分割及关键帧的提取

依此将视频分割成不同的镜头,同时提取出每个镜头的首尾帧存入多媒体数据库,并将各图像帧和 $offset$ 建立时间对应关系。当用户点击图像帧检索相应时间段的视频时,就可以直接发现视频跳到对应的时间段,而不至于要观看整段视频。这给现在快节奏生活中想观看特定时间段的短视频的人们,特别是手机用户提供了非常便利的功能。

3.2 索引的建立

如何建立索引是实现海量视频数据库进行按时间段快速检索的前提。目前,基于现有视频检索技术,建立索引主要有如下两种基本方法^[2,5,6]:

(1)基于注释的索引:按视频内容的镜头分割方法,将视频分割为不同的镜头后,将不同特征的镜头分层地对各个镜头进行注释,该方法的优点是可以详细、丰富地描述视频数据的语义特性,因此,基于这种索引方式的检索是比较精准的。但由于目前的图像处理技术水平还无法由机器自动完成这种索引的建

立,而不得不由人工手动完成。这对于海量的不同视频的镜头片段来讲,是非常繁琐,低级的方法。

(2)基于特征的索引:该种索引的建立,主要是按照视频图像的底层特征的提取而建立索引。如:颜色特征,纹理特征,轮廓特征等。该方法的特性和基于注释的索引方法的特性恰好相反,能取代人工手动由机器自动生成,却往往无法形成符合人类语言逻辑的自然语义信息,这也就不便于用户使用适当的文本语言对其进行搜索查询。

结合前面提出的多媒体数据库系统模型,并针对方法1的繁琐和方法2的低效率等缺点,本文对视频索引的建立提出以下的改进方法:

(1)对用户上传时添加的标题,注释建立索引,这种基于自然语言的方法对用户来讲是最人性化,最自然的。

(2)结合提取出的各镜头的首尾帧的时间偏移量 **offset** 建立索引。并在用户应用接口层,直接以图像的形式显示出各镜头的首尾图像帧。这一方式避免了手动加入各视频段注释的困难。另外,该方式比基于注释的索引方法更形象、直观地列出了各镜头的内容。在本方案中,虽然也要手动加入注释,但一段视频只要加一次,并且不同视频可以由不同上传用户输入,从而不用对一段视频的不同镜头在后台分割后再手动加入注释。

3.3 分层检索的建立

在后台完成关键帧的提取及索引的建立,是为了为用户提供快速,精准的查询检索功能。由于同时搜索视频及时间段对用户来说是非常困难的,因为用户并不清楚具体的视频段的具体时间段。因此,基于前述对关键帧的提取及索引的建立,本文提出分层检索的方法,其流程如图3所示。

当用户输入关键词时,系统根据索引与注释及标题匹配开始查询,并列出与之匹配的视频,同时附上提取出的各关键帧图像。用户根据浏览到的图像,按不同的喜好,作出选择。若直接点击播放,则播放整个视频;若点击图像,这时系统就按 **Msrc**(视频文件源地址)路径直接启动与该图像相对应的 **offset** 时间点的视频,即可直接观看该段短视频,而不用从头至尾地全部浏览整段视频。这对于目前快节奏生活中的手机用户而言,想观看自己喜好的短视频是非常实用的功能。

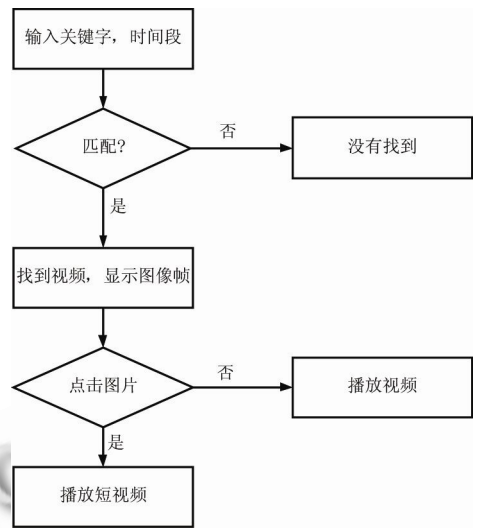


图3 分层检索流程图

另外,对于某些特定用户,如:手机等移动设备用户只想观看短时间的视频(如3分钟),则可以输入时间段,直接查询与之匹配的各种短视频。

4 结束语

随着数字娱乐信息产业的发展,人们对各种形式的多媒体数据的检索要求越来越高。针对人们对短视频检索的需求,通过对现有检索技术的改进,本文介绍了一种基于视频时间段检索的多媒体数据库系统,该系统可以直接按直观的图像检索到个人喜好的短视频。该系统能为基于移动终端的短视频检索系统的研究与开发提供参考。

参考文献

- 1 刘政凯,汤晓鸥.视频检索中镜头分割方法综述.计算机工程与应用,2002,38(23):84-87.
- 2 柯和平.基于内容检索的视频数据库系统研究.计算机与数字工程,2007,35(5):69-72.
- 3 季春.基于内容的视频检索中的关键帧提取技术.情报杂志,2006,11:116-119.
- 4 钱刚,曾贵华.典型视频镜头分割方法的比较.计算机工程与应用,2004,40(32):51-55.
- 5 郑鹏,周晓嵘,周洞汝.按内容访问视频数据库.计算机工程与应用,1999,35(8):73-75.
- 6 陈志圣,陈福生,刘庆.视频数据库索引技术研究及应用.微型电脑应用,2004,20(4):11-12.