

分布预取式 VOD 系统的设计与实现^{*}

Design and Implementation of Distributed Pre-fetch VOD System

高攀 周皓峰 朱扬勇

(复旦大学计算机系 上海 200433)

Abstract VOD is one of the hottest topics in multimedia technology. To implement a good VOD system, an intelligent multimedia search tool as well as a proper system architecture and a efficient scheduling are needed. This paper presents a distributed pre-fetch architecture, an efficient scheduling and a flexible query/search method based on this architecture.

Keywords VOD, Scheduling, Content-based query, MPEG-7

VOD (Video on Demand, 视频点播) 是伴随着音频、视频处理技术及网络技术的发展而迅速兴起的一门综合性技术。视频服务器通常采用较高档硬件设备,并在操作系统级采取相应优化策略,这类集中式解决方案价格昂贵,专用性强。在给定的硬件条件下,构建一个合理的体系结构,采用高效的调度策略,始终是任何多媒体应用所必要考虑的问题,文[4,7]均提出了分布式的解决方案,以期获得良好的可扩展性和较好的性价比。另一方面,多媒体数据缺乏结构性和直观的语义信息,如何更好地完成“点播”过程,从大量的多媒体信息中找到自己感兴趣的内容也是一个难题。文[5,6]等所描述的多媒体检索系统,制定中的 MPEG-7 标准^[2,3],都对这个问题作了一定的探索。

本文提出了一个分布预取式 VOD 系统的体系结构,基于这个结构提出了合适的调度算法,其优势在于可扩充性强、响应快、吞吐量大。同时提出了一个灵活的基于内容的检索办法,它允许用户精确地描述自己感兴趣的多媒体资料,将基于非语义信息和基于语义信息的查询纳入一个统一的框架,并智能地利用领域知识返回相关的查询结果。

1 VOD 系统结构

VOD 系统的特征是视频音频数据量特别巨大,而依照 Zipf 分布规律,大部分用户请求会局限在少数的节目上。所以数据的存储应当是层次型的,维护一个数据源中心(即内容服务器),让直接处理用户请求的服务器动态地向数据源中心请求数据,并发送回放数据,形成一个三层分布式结构。

内容服务器:负责视频资料的输入和存取控制,实现多媒体资料的基于内容的查询检索,并为缓冲服务器提供原始的视频音频数据(位于光盘库等慢速设备中)。

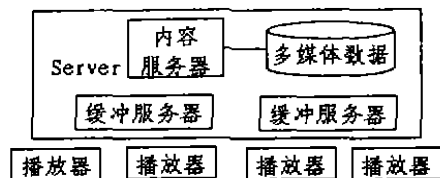


图 1

缓冲服务器:维护视频数据的换进换出,接受用户请求,执行允许控制,分割请求,提交请求,并发送流式数据,在系统要求较高的情况下,它可以用高端单机型服务器产品充当。

播放器:可以是浏览器、独立的应用程序或是机顶盒等,负责读取视频流并播放。它从缓冲服务器获取流式数据的方式可以利用现有的技术和有关标准,如基于因特网的多点组播, H. 323 标准^[10]以及基于电缆电视网 HFC 的 DAVIC1.1^[9], IEEE802.14^[6]等。

内容服务器、缓冲服务器和播放器位于不同子网内,构成一个分布式的构架,同一子网可以有多个服务器,它们之间动态地交换数据,使处理的瓶颈分布到整个网络上。在网络连接上,缓冲服务器应当离播放器尽可能近一些,为播放器缓存将要播放的数据。例如,一个较好的结构是:内容服务器和缓冲服务器用主干网连接,以便进行快速数据交换,缓冲服务器再通过非

^{*} 本文研究得到上海教委曙光项目基金资助,高攀 硕士生,主要研究领域为数据库和知识库,周皓峰 博士生,主要研究领域为数据库和知识库,朱扬勇 教授,博士生导师,主要研究领域为数据库,数据仓库和数据采掘、多媒体。

主干网与用户桌面上的播放器相连,形成一个层次结构。由于视频音频数据必须先从慢速设备中读取并写入缓冲服务器之后,播放器才有可能获得数据开始回放,故属于预取式 VOD 系统。

2 服务器调度策略

2.1 缓冲服务器的调度策略

一部片子持续时间可能很长,处于片头和片尾的数据的紧迫程度是不同的;另一方面,本系统信息点播的最小单位是一个镜头,远小于一部片子,要让最受欢迎的多媒体资料始终离用户最近,必须降低调度的粒度。缓冲服务器根据现存的片段切分用户请求,或重复利用已有片段,或生成新的待请求片段,这些分片后的请求是五元组 (MID, F_b, F_e, T_R, S) , MID 唯一确定一部片子, F_b, F_e 分别表示需要播放的开始和结束帧数, T_R 是所有需要这个片段的用户请求的播放时间的最小值。S 是缓冲片段的状态,包括:等待;请求;写入;就绪待用;正在使用;就绪无用。后三个状态统称就绪状态。

文[4]提出片段的状态随着播放不停地改变,同时用户的请求又不断增加着新的片段,缓冲服务器的多线程的任务是使服务器不断地向“平衡态”转变。

定义 1 平衡态:以服务器的缓冲容量为限,如果最早需要时间最早的那些片段都已经处于就绪状态,则称服务器处于平衡态。平衡态时,下式成立:

$$\forall C_i, C_i(S) = \text{就绪}, \forall C_i, C_i(S) < \text{就绪} \rightarrow C_i(T_R) < C_i(T_R) + T_b, T_b \text{ 是给定阈值。}$$

在有能力存贮的情况下,缓冲服务器会尽早发出请求,以便内容服务器进一步调度。当缓冲服务器接收到视频数据时,可能要换出系统中不重要的片段。判断缓冲服务器是否有能力存贮和寻找系统中不重要的片段是同一问题。若系统中不重要的片段的大小超过当前请求处理的片段大小,服务器就有能力存贮该片段。我们将系统中片段的重要程度做这样规定:就绪无用的片段最不重要,就绪待用的片段的重要程度根据其 T_R 值不同而不同。对于就绪无用的片段,我们可以根据其过去的使用情况,估算它的下次使用时间 T_n ,从而区分它们不同的重要程度。 T_n 可以定义为: $T_n = T_i + (\sum \alpha^i \Delta t_i) / (\sum \alpha^i)$ ($0 < \alpha < 1$), 其中 Δt_i 上第 i 次与上第 $i+1$ 次访问该片段的时间间隔, T_i 为当前时间。

2.2 内容服务器的调度策略

内容服务器处理来自缓冲服务器的请求也需要通过调度来优化系统性能。要使响应时间小,这些请求应该按照回放时间的先后顺序依次处理;但考虑到光盘库中盘片的调度和读取头的移动在整个响应时间中占的比例较大,故处理请求时应当使盘片的换入和读取

头的移动次数最少,才能获得最大的吞吐量,因此本系统采取的是一种折衷的处理办法。

内容服务器维护的处理队列是个双重队列,一方面按照请求回放的时间从小到大排序,即表示这个请求的紧迫程度,另一方面按照请求数据所处的文件进行归类,片段属于同一文件的请求属于一类,同一类中的各请求仍按照请求回放的时间顺序排列。设有请求中回放时间最小的请求是 R_1 , 其回放时间为 T_1 , 设上次内容服务器工作线程处理的请求要求的数据在文件 F 中,而下一个要求的数据在文件 F 中的请求为 R_2 , 其回放时间为 T_2 。

1. 如果 R_1, R_2 都不存在,则工作线程挂起等待;

2. 如果 R_1 存在, R_2 不存在,则处理请求 R_1 , 反之处理请求 R_2 ;

3. 如果 $T_1 - T_2 < h$ (h 为事先确定的阈值,与所选光盘库的参数有关,如换入光盘的所需时间和平均寻道时间),则这次处理请求 R_1 , 否则处理 R_2 。

2.3 视频数据的动态抽取与合成

以目前最常见的视频数据格式即 VCD 的 dat 格式为例,一个片子即一个 .dat 文件。其结构为:

| | | | | |
|------------|---------------|---------------|-------|---------------|
| DAT 文件头 | MPEG 数据块 1 | MPEG 数据块 2 | | MPEG 数据块 n |
|------------|---------------|---------------|-------|---------------|

图 1

DAT 文件头大小为 70628 字节,每个 MPEG 数据块大小为 2352 字节,其中最后 28 个字节仅作填充用。MPEG 数据块中的有效数据形成 MPEG 数据流,其中前两个 MPEG 数据块中的数据是 MPEG 系统流头数据,其余则按照一定比例交错存放 MPEG 视频流和音频流的数据。DAT 文件头较大,为了更好地利用网络资源,让播放器能尽早开始播放,本系统从数据源中动态抽取该片子的 MPEG 系统流头、所需片段的 MPEG 视频和音频流数据并存储在缓冲服务器中,在播放时,将 MPEG 系统流头、相应多个片段的 MPEG 视频音频流数据依次写入播放流,供播放器使用。其他常见数据格式一般都可以用类似方式动态抽取合成处理。

3 基于内容的多媒体资料检索

上述系统结构和调度策略构成了 VOD 系统运作的基石,另一方面,如何更好地完成点播过程,从大量的多媒体信息中迅速找到自己感兴趣的内容也是一个完整的 VOD 系统所必须解决的问题。

3.1 多媒体数据的分段

原始的视频音频数据是非结构化的数据流,其非

语义特征和语义信息随时间的改变而改变, 要让用户准确地找到自己感兴趣的视频音频资料, 必须将原始数据流分段以减小检索粒度。我们将其分解成为一个三层结构: 镜头、场景和片段。镜头指一个摄像机拍摄的一系列连续帧, 帧之间是渐变而不是突变的。一个场景由一个或多个反映同一事件的在时间或空间上相关联的镜头组成, 而一个片段则由一个或多个语义上相关的场景组成, 例如新闻报道中的一则新闻或一个故事。镜头是非语义相关的, 已有不少对镜头边界检测的研究, 并已取得一定成果; 而场景和片段都是语义相关的, 只能通过人工的办法确定其界限。

我们采用镜头作为音频视频数据分段的最基本的单位, 因为一般一个镜头中各帧的非语义特性比较相近, 而具有独立的语义含义的场景由一个或多个镜头构成, 将一个包含单个或多个镜头的场景作为检索的结果其粒度是合适的。从原始多媒体数据中动态抽取与合成视频数据, 这个问题已经在上文中解决了。

3.2 查询表达方式的分析与索引的建立

一个查询由一条或多条用“与”, “或”, “非”运算连接的原子查询组成。原子查询可以有两种, 一种是有关非语义信息的限定条件, 另一种则是涉及语义信息的限定。前者的形式通常可以归结为要求某参数如整体色调的 RGB 取值在某范围之内; 而后者可以看作是一个可能残缺不全的对自己感兴趣的资料的描述, 也就是说, 一个可能缺少一个或多个语法成分的陈述句。一句陈述句的语法成分有主、谓、宾、定、状、补。主语和宾语通常是实体, 而谓语一般是动作, 可以看成是关联实体的谓词, 定语, 描述的是实体的属性。状语, 则有几种情况: 1) 直接修饰动作, 例如“慢慢(地走)”可以看作是谓词的属性; 2) 地点, 可以将它看作是另一个原子查询中出现的实体; 3) 时间, 对整个一段视频应当是全局的, 我们可以考虑将时间作为一类独立的属性。而补语, 在纯粹以描述事实的陈述中用得不多, 或可以转化成其他不用补语的句式。例如, “激动得流泪”可以转化为“激动地流泪”。

当两个以上的原子查询用“与”连接时, 这些原子查询所描述的场景还可以有一定的相互关系, 如场景 A 出现在场景 B 之前。

本文认为, 上述对基于内容的查询条件表达的分析所列出的情况已能够表达绝大部分检索要求, 因此在建立索引时将以上述分析作为基准。

视频音频资料的索引信息分为两类, 一类是与内容无关的, 如作者, 制作时间等, 这类信息的组织方式是简单的, 故本文不做详细分析; 另一类则是反映视频音频本身的内容的信息, 其基本组织形式由实体和谓词组成, 它们可以抽象为内容对象。内容对象在应用领

域中指有意义的有明确界定范围的事物, 可以是某个人物, 动作(简单对象), 或是所有人物的集合(复合对象)。内容对象通过其属性来描述, 属性根据其抽象程度的不同, 大致分为两大类: 非语义属性和语义属性。前者是比较底层的, 无须领域知识介入且一般可以用机器提取的特征; 语义属性则是比较高层的, 需领域知识介入且一般无法用机器提取而需手工进行标注的特征, 如电影镜头中的人物名称, 风景的地点等。为了更确切地表达信息, 我们将语义属性进行分类, 使之形成层次结构, 层次越高, 反映的语义信息(内涵)就越丰富, 所指定的对象(外延)就越明确, 粗略地分成三个层次: 一般、特定及抽象。层次越高, 涉及的语义信息越丰富, 实现了内容对象属性的层次管理, 也就实现了实体和 E-R 关系谓词属性层次的管理。

3.3 智能查询的实现

将属性分为三层在一定程度上对多媒体信息的描述作了限定, 有助于查询更明确地表达和检索结果的准确性, 但事实上对象的属性是一个多层次的复杂的层次结构, 图 2 所示的是一个简单的例子。当进行语义标注时, 如果已知这一片段有关汤姆·克鲁斯时, 通常不会再另外指出这也是有关演员的片段。事实上, 这种从属关系是与领域知识相关的, 并不是多媒体素材自身所反映的, 不应当作为元数据的一部分。另一方面, 对内容进行标注时, 可能发生不同文字描述反映相同或相似语义的情况, 这类语义关系, 同样也是与领域知识相关而非多媒体数据本身所能体现的。

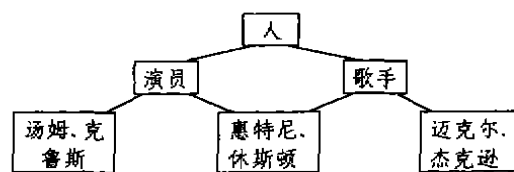


图 2

为了更好地维护属性间的层次或语义关系, 我们另外引入一个知识库, 专门用来存储表达属性间关系的知识。其中每一条这样的知识是一个形如 $A \xrightarrow{(a)} B$ 的蕴涵式, A 是前提, B 是结论, a 是这条知识的肯定系数, 取值范围在 0 和 1 之间, 表示从前提能够推出结论的肯定程度。例如, 草坪 \Rightarrow 户外的肯定系数为 90%, 表示草坪的场景有 90% 的可能是户外场景, 记做草坪 $\xrightarrow{(0.9)}$ 户外。知识具有以下性质:

自反性: 即, $A \xrightarrow{(1)} A$
 传递性: 即, $A \xrightarrow{(a)} B \wedge B \xrightarrow{(b)} C \Rightarrow A \xrightarrow{(ab)} C$

(下转封四)

(上接第 113 页)

在处理查询时,假设我们有一个原子查询要求查询有关 A 和 B 的内容,我们考虑所有从属于 A, B 的对象,假设有一片段的描述具有属性 C, D, 且有 $C \xrightarrow{(a)}$ A, $D \xrightarrow{(b)}$ B, 则我们定义这个片段符合该查询条件的相似度为 ab。

知识可以由人工传授,即手工输入,也可以通过访问历史记录进行知识发现来获得。

由于非语义信息的匹配结果通常也是一个相似度的数值,此时我们就把基于非语义信息的查询和基于语义信息的查询统一起来了,如果原子查询间用“与”连接,则相似度取较小的一个,如果原子查询间用“或”连接则取较大的一个。最后将相似度大于给定阈值的片段按照相似度排序后作为查询结果返回。

3.4 与其他查询系统的比较

现有的多媒体资料查询系统可以分为两类,一类是基于如颜色、纹理、形状等特征的查询,通过特定算法与库存资料的相应特征进行比较,以相似度为衡量标准判断是否将其作为查询结果。另一种做法沿袭了针对文本或类文本(例如超文本)资料的查询检索方法,使用关键字,与预先对多媒体资料标注时所用的关键字进行匹配,如有匹配,则认为该段资料符合查询要求。

本系统中对象属性包含非语义属性,也就是说能够将第一种做法纳入这个查询体系中去。与第二种做法相比,它将元数据先横向分为内容元素和谓词,再纵向地分为多个层次,使语义的表达更为精确,使元数据能够比关键字更有效地得到利用,同时让用户也以这种方式精确地描述自己感兴趣的多媒体资料,使查询的结果更准确。这种查询方式还进一步考虑了领域知识的引入,通过一定的推理返回相关的查询结果,与现有的查询系统相比更加智能。通过引入推理的肯定系

数,让语义信息的匹配也有了一个相似度的概念,从而较好地非语义信息和语义关系纳入了一个统一的查询框架之中。

结束语 本文提出了一个分布预取式 VOD 系统的体系结构,并基于这个结构提出了合适的调度算法,其优势在于硬件的基本要求不高,可扩充性强,响应请求快,吞吐量较大。另外,针对多媒体信息检索这个问题,本文提出了一个灵活的基于内容的检索办法,综合考虑非语义信息和语义信息,将其纳入统一的框架之中,对元数据进行了分类、分层,使查询的表达及其结果更准确,它还进一步考虑了领域知识的引入,通过一定的推理返回相关的查询结果,与现有的查询系统相比更加智能。

参考文献

- 1 Gemmell D J, et al. Multimedia Storage Servers: A Tutorial. IEEE Computer, 1995, 28(5): 40~49
- 2 Requirements Group. MPEG-7: Context and Objectives (version-10 Atlantic City), MPEG98 October 1998/Atlantic City, USA
- 3 James A, et al. Multiple Level Classification of Visual Descriptors in the Generic AV DS. MPEG99 Melbourne, Australia, Oct 1999
- 4 Brubeck D W, Rowe L A. Hierarchical Storage Management in a Distributed Video-On-Demand System. IEEE Multimedia, 1996, 3(3): 37~47
- 5 Chang S -F, et al. Visual Information Retrieval from Large Distributed On-line Repositories. CACM, 1997, 40(12): 63~71
- 6 Benitez A B, et al. Using Relevance Feedback in Content-based Image Metasearch. IEEE Internet Computing, 1998, 2(4): 59~69
- 7 向哲,钟玉琢,龙白滔. 视频服务器的体系结构. 计算机世界, 1999, 13
- 8 Available at: <http://www.walkingdog.com>
- 9 Available at: <http://www.davic.org>
- 10 Available at: <http://www.openh323.org/standards.html>

计算机科学

(1974年1月创刊)

第28卷第1期(月刊)

2001年1月25日出版

中国标准刊号: ISSN 1002-137X
CN50-1075/TP

定价: 12.00元 国外定价: 5美元

邮发代号: 78-68

发行范围: 国内外公开

主管单位: 国家科学技术部

主办单位: 国家科技部西南信息中心

编辑出版: 《计算机科学》杂志社

重庆市渝中区胜利路132号 邮政编码: 400013

电话: (023) 63500828 E-mail: jsjcx@swic.ac.cn

主 编: 朱宗元

印刷者: 国家科技部西南信息中心印刷厂

总发行处: 重 庆 市 邮 政 局

订购处: 全 国 各 地 邮 政 局

国外总发行: 中国国际图书贸易总公司(北京399信箱)

国外代号: 6210-MO