

一种 MPEG-2 流的索引模型及其应用*

王伟强¹, 高文^{1,2}

¹(中国科学院 计算技术研究所, 北京 100080):

²(哈尔滨工业大学 计算机科学与工程系, 黑龙江 哈尔滨 150001)

E-mail: wqwang@ict.ac.cn; wgao@ict.ac.cn

http://www.ict.ac.cn

摘要: MPEG(moving picture expert group)标准目前以及未来都将在多媒体应用中起到越来越重要的作用. 提出了一种有效的针对 MPEG-1、2 码流的索引模型, 并详细给出了 MPEG 流的索引文件生成算法以及基于该模型对码流中的任意帧随机访问显示算法. 实验结果表明, 该模型可提供对一般的 MPEG-1、2 码流异常快速的定位访问能力, 并可有效地应用于码流中的任意帧随机访问.

关键词: MPEG; 索引; 节目流; 传输流; 图片组

中图法分类号: TP316 文献标识码: A

桌面计算机计算能力的提高、存储介质费用的不断降低、宽带网络的扩展使得更多的数字视频应用逐渐成为现实. 每一个数字视频素材都由一系列时间上连续的帧构成, 诸多的多媒体视频应用都希望系统具有快速定位到视频流某一特定帧的能力. 目前, 微机环境下的 VCD 软件播放器虽然提供了一定程度的随机访问帧的能力, 如类似 VCR(video cassette recorder)的快进、快退功能, 但其在快速性、准确性上均有一定的局限性, 无法满足许多高级多媒体应用的需要. 比如, 一个支持基于内容浏览的大型视频数据库系统, 当用户对视频流文件进行快速浏览时, 经常会提出跳转到另一个场景、故事单元或镜头的要求, 这就需要系统能以最快的响应速度准确地跳转到目标场景、故事单元或镜头的起始第 1 帧, 而且由于系统一般都以客户/服务模式并发地响应多个用户的请求, 从而对系统的任意帧随机定位访问能力的快速性、准确性提出了更高的要求. 这种应用背景即便采用文献[1]中所叙述的快速定位算法也无法给予用户满意的服务质量, 因为一方面, 文献[1]在确定一帧的起始位置时仍存在字节内容比较运算, 帧定位的性能还不能达到最佳效率. 另一方面, 一个场景、故事单元或镜头的起始帧为 P 帧或 B 帧的情况很普遍, 而文献[1]的帧定位仅局限于对 I 帧. 这时, 通过对视频流建立合适的索引来实现对特定帧视频内容访问的快速性、准确性是一种有效的解决途径.

文献[1]提出了一种在 MPEG-2 系统流上采用粗跳步与细跳步快速定位 I 帧图像的方法, 该方法较适用于 MPEG-2 演播器这样的应用背景. 与之相比, 本文讨论的索引模型具有如下特点: (1) 对节目流仅假设其 GOP(group of picture)结构在整个节目流是规则的, 该假设对一般的商用流来说普遍成立; (2) 可定位访问任意一帧, 无论其帧编码类型为 I, P, 还是 B; (3) 对于特定帧的定位访问具有近乎完美的响应时间, 且帧定位的计算费用近似为一个常数. 因此, 该模型可更好地

* 收稿日期: 1999-12-29; 修改日期: 2000-04-18

基金项目: 国家自然科学基金资助项目(69789301); 国家 863 高科技发展计划资助项目(863-306-ZT02-01 2; 863 306 ZD11-03); 中国科学院百人计划资助项目

作者简介: 王伟强(1972-), 男, 浙江绍兴人, 博士, 助理研究员, 主要研究领域为多媒体技术、人工智能; 高文(1956-), 男, 山东牟平人, 教授, 博士生导师, 主要研究领域为多媒体数据压缩, 图像处理, 计算机视觉, 多模式接口, 人工智能, 虚拟现实.

适用于许多高级的多媒体应用背景.

1 MPEG-2 流索引模型的描述

一般地,在一个 MPEG(本文后面若不作特殊说明,MPEG 均指 MPEG-1 和 MPEG-2)节目流中存在 3 种编码类型的图:I 图、P 图与 B 图.在 MPEG-1 中,每个图都由一帧来构成;而在 MPEG-2 中,它既可为一帧,也可能是一场.I 图和 P 图被称作参考图.从一个 I 图起始到下一个 I 图之前称为一个图片组 GOP.

本文讨论的索引模型假设被索引节目流的 GOP 结构在整个节目流是规则的,该假设对于一般的商用流是普遍成立的.所谓规则的是指 GOP 结构在整个节目流中保持不变,且按显示顺序排列的参考帧之间所包含的 B 帧的数目相同.

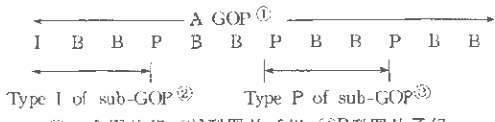
下面给出该索引模型的形式化描述.

定义 1. MPEG 流的索引 X 是一个二元组 $\langle G, D \rangle$, G 代表对压缩码流结构特征及索引 X 的特征描述项集, D 是按显示顺序排列的索引对象集.

定义 2. 图片子组是按播放顺序从一个参考帧到下一个参考帧之前所有帧的集合.若一个图片子组中的参考帧含有 I 图,则该图片子组称为 I 型图片子组,否则称为 P 型图片子组.

如图 1 中所示的图片组便由一个 I 型图片子组和 3 个 P 型图片子组构成.一般地,每一个图片组 GOP 均由一个 I 型图片子组和若干个 P 型图片子组构成.

定义 3. I 和 P 索引子分别是对 I 型图片子组和 P 型图片子组的索引表示形式,是一个四元组 $\langle tag, numB, ipPos, dif_Bs \rangle$,其中 tag 用于标志索引子的类型, $numB$ 为该索引子所对应的图片子组中包含的 B 帧数目, $ipPos$ 为该索引子对应的图片子组中 I 帧或 P 帧在流中位置的相关描述信息, dif_Bs 为一个由 $numB$ 个元素组成的有序正数集,其中每个元素是图片子组中各个 B 帧位置相对于 $ipPos$ 的位置差值的描述信息.



①-一个图片组,②I型图片子组,③P型图片子组.

Fig.1 Conception of sub-group of picture
图1 图片子组概念的图示

每个索引子构成了 D 中的一个索引对象.

定义 4. MPEG 流的索引 X 中的元素 $G = \langle FrameRate, Frames, tagRegular, Gop_NP, Gop_NB, BL_GopIndex, BaseFrmNum, TypeStream \rangle$,其中 $FrameRate$ 为 MPEG 流的帧率, $Frames$ 为该视频流的总帧数, $tagRegular$ 标志索引的 MPEG 流的 GOP 结构是否是规则的,本文假设它总为真(该元素的存在可提供未来对该模型进行扩展), Gop_NP 和 Gop_NB 为被索引的 MPEG 流的每个 GOP 中 P 帧和 B 帧的数目, $BL_GopIndex$ 为在索引 X 中索引一个 GOP 中所有帧花费的字节数, $BaseFrmNum$ 为 MPEG 视频流中按播放顺序第 1 个 I 帧的帧序号, $TypeStream$ 为被索引视频流的类型.

正常情况下, $BaseFrmNum$ 应为 0,但本模型考虑到有些情况下会使生成的 MPEG 码流第 1 个 I 帧的帧序号不为 0.例如,我们通过特定的设备录制的通过卫星传送的数字电视节目就经常遇到这种情况,这时,播放帧序号为 $0 \sim BaseFrmNum - 1$ 的帧不能被正确解码输出.所以,该索引模型仅对播放帧序号从 $BaseFrmNum$ 到 $Frames$ 的帧进行索引. $TypeStream$ 是一个重要的元素,它决定着定义 3 中所定义的索引子的 $ipPos, dif_Bs$ 的语法结构及构成元素的语义.

从上面的定义不难得到如下的结论.

结论 1. 本模型不仅可以按照播放帧的序号来定位访问特定帧,也可以利用时间信息来定位访

间特定帧.可用式(1)来进行时间与播放帧序号之间的换算.

$$Dest_FrmNum = Current_FrmNum + TimeLen * FrameRate. \quad (1)$$

$Dest_FrmNum, Current_FrmNum$ 分别为目标与当前的帧序号, $TimeLen$ 为快进或快退的以秒为单位的时间间隔.

结论 2. 给定一个待定位帧的帧序号 $DestFrmNum$, 利用 MPEG 流的索引 X 中的信息, 可以确定该帧所在 GOP 的 I 型图片子组对应的索引子在 I 中 D 元素中的位置. 可用式(2)和式(3)来获得.

$$GopSize = 1 + Gop_NB + Gop_NP, \quad (2)$$

$$PosIdxFile = BL_GopIndex * INT((DestFrmNum - BaseFrmNum) / GopSize). \quad (3)$$

其中 $GopSize$ 为该视频流的 GOP 长度, $PosIdxFile$ 为帧序号是 $DestFrmNum$ 的帧所属 GOP 的 I 型图片子组对应的索引子在 I 中 D 元素中的字节位置, $INT()$ 为下取整函数.

结论 3. 若在 MPEG 视频流的索引 X 中所有 I, P 索引子占用的字节长度是等长的, 设其为 $LItem$, 则对任意一帧利用 X 中的信息其对应的索引子在 D 中的位置是计算可定的.

设该帧的播放序号为 $DestFrmNum$, 利用式(2)和式(3)可求得 $PosIdxFile$, 进一步利用下面的式(4)~式(6)可求得该帧对应的索引子的位置 $PosIdxItem$. 其中 Mod 为取模操作符.

$$nFrmsItem = Gop_NB / (1 + Gop_NP) + 1, \quad (4)$$

$$RmFrms = (DestFrmNum - BaseFrmNum) Mod GopSize, \quad (5)$$

$$PosIdxItem = LItem * INT(RmFrms / nFrmsItem) + PosIdxFile. \quad (6)$$

结论 4. 若一个索引子中的各元素是定长已知的, 则对任意一帧利用 X 中的信息, 则其在 MPEG 流中的位置是可求的.

对于序号为 $DestFrmNum$ 的帧, 利用结论 2 和结论 3 易求出在索引 X 中对应的索引子的位置, 若一个索引子中的各元素是定长已知的, 则各元素的位置可求, 便可得到该帧在 MPEG 流中的位置.

通过上面的讨论可知, 若按上面的模型建立了一个 MPEG-2 流的索引 X, 当对流中任意一帧访问时, 可通过非常小的计算复杂度获得任何一帧在流文件中的位置, 在整个定位的过程中不涉及任何字节内容比较搜索计算, 因此, 它提供了非常高效的在视频流中定位任何编码类型帧的能力.

2 实现与应用

2.1 建造 MPEG 流索引文件的算法

参照第 1 节中的索引模型, 我们设计实现了一个对 MPEG 流的索引工具. 它可对 MPEG 节目流、视频基本流建立索引. 算法通过对 MPEG 流中的视频数据的内容进行扫描来发现两种重要的位置信息: 视频 PES 包头起始位置 and 对应于一帧开始的图片头位置. 对于视频基本流, 则仅提取图片头位置信息. 同时, 算法完成了 MPEG 流编码次序向播放次序的转换, 使得索引中的帧排列顺序与播放次序相一致. 这种有序规则性, 不但使通过索引检索目标帧在流文件中的位置信息高效, 而且无须在索引文件中存储每个索引项对应的帧号信息, 从而节约了索引文件的空间费用(索引文件的空间耗费情况见表 1). 算法的完整描述如下:

MPEG 流的索引文件生成算法.

输入: MPEG-2 或 MPEG-1 的节目流, 或视频流文件 mpeg-fp.

输出:相应的索引文件 `mpgIdx.fp`.

中间变量:

`FwdFrmInfo`, `BwdFrmInfo`, `AuxFrmInfo` 分别用于存储前向参考帧、后向参考帧、B 帧的帧编码类型、图片头信息在 `mpeg.fp` 的位置信息.

`CurFrmInfo`, `DisFrmInfo` 分别用于指向引用解码当前帧、播放当前帧的索引信息.

- (1) 打开文件 `mpeg.fp`, 打开一个临时文件 `Tmp.fp`.
- (2) `Frames=0`, `Snd.Field=FALSE`.
- (3) 若视频缓冲区空, 则搜索、解析 `mpeg.fp`, 将视频基本流读入视频缓冲区.
- (4) 搜索、解析视频缓冲区中的各种 MPEG 头信息数据(从序列头信息中提取出帧率), 若遇到序列结束码或文件 `mpeg.fp` 结束符, 则转(15); 若遇到一个视频 PES 包头, 记下该包头的起始位置于 `LastVPESHdrPos`; 若遇到一个图片头, 记下该图片头的起始位置于 `PicHdrPos`, 并转(5).
- (5) 解析该图片头.
- (6) 若图片编码类型(`PictCodingType`)为 B 类型, 则 `CurFrmInfo` 指向 `AuxFrmInfo`, 转(9); 否则转(7).
- (7) 若 `Snd.Field=FALSE`, 则 `FwdFrmInfo`, `BwdFrmInfo` 的内容互换.
- (8) 令 `CurFrmInfo` 指向 `BwdFrmInfo`.
- (9) 若 `Snd.Field=FALSE`, 则将 `PictCodingType` 和 `PicHdrPos` 的信息存入到 `CurFrmInfo` 指向的对象. 若流类型为节目流, 同时也将 `LastVPESHdrPos` 信息存入到 `CurFrmInfo` 指向的对象, 验证 MPEG-2 码流的 GOP 结构是否出现非规则情况, 若是, 则失败, 退出.
- (10) 若 `Frames≠0`, 且满足图片结构(`PictStruct`)为帧类型或者满足 `Snd.Field=FALSE`, 则按照如下的规则选择 `DisFrmInfo` 所要指向的对象; 若 `PictCodingType` 为 B 类型, 则令 `DisFrmInfo` 指向对象 `AuxFrmInfo`, 否则 `DisFrmInfo` 指向对象 `FwdFrmInfo`.
- (11) 若当前图片结构 `PictStruct` 不为帧类型, 则 `Snd.Field` 取反.
- (12) 若 `Snd.Field=FALSE`, 则将 `DisFrmInfo` 指向对象的信息放入到队列缓冲区 `OutputBuffer` 中, `Frames=Frames+1`.
- (13) 若缓冲区 `OutputBuffer` 超过一定的满度, 则取出一定量的数据构造出若干按定义3来定义的索引子, 并按预定的文件格式写入到一个临时文件 `Tmp.fp` 中.
- (14) 转(3).
- (15) 若缓冲区 `OutputBuffer` 不空, 则取出其中所有的数据, 构造出若干按定义3所定义的索引子, 并按预定的文件格式写入到一个临时文件 `Tmp.fp` 中.
- (16) 打开索引文件 `mpgIdx.fp`.
- (17) 根据定义4所定义的内容按一定的文件格式将描述数据写入到文件 `mpgIdx.fp` 中. 将文件 `Tmp.fp` 的内容追加到 `mpgIdx.fp` 中, 并关闭文件 `mpgIdx.fp` 和 `mpeg.fp`, 关闭并删除临时文件 `Tmp.fp`.
- (18) 结束.

算法的步骤(1)和(2)完成一些初始化的工作; 步骤(3)~(14)用于构造索引 X 中按显示顺序排列的索引对象集 D . 其中变量 `Snd.Field` 用于标记当前的图是否为一帧中的第2个场图; 步骤(4)通过对码流的读取与解析抽取用于构造索引子的帧位置描述信息 `LastVPESHdrPos` 与 `PicHdrPos`; 步骤(6)~(8)选择用于存储当前解码帧位置描述信息的对象; 步骤(9)在当前图为帧图或构成一帧的第1个场图时, 将用于构造索引子的帧位置描述信息 `LastVPESHdrPos` 与 `PicHdrPos` 存入到步骤(6)~(8)所选择的对象中; 步骤(10)选择对应解码顺序的帧索引信息的描述对象; 步骤(12)和(13)将步骤(10)所选择的对象缓冲并输出到临时文件 `Tmp.fp` 中; 步骤(16)和(17)完成索引文件的最后生成.

2.2 利用 MPEG 流的索引随机访问显示节目流中的任意一帧

对 MPEG 码流中的帧进行解压缩并加以显示是许多涉及到视频的多媒体应用的一项基本操作功能. 由于 MPEG 码流编码的特点, 只有编码类型为 I 的图其内容才可独立解码显示, 其他类型的图都需要与其相关的某些帧正确解码后该图才能够正确解码, 这也正是一般的 MPEG 节目流演播器的快进、快退的分辨粒度一般为 GOP 的原因, 实际进行的是与目标位置最为接近的 I 图的搜索.

我们可以利用索引文件对一帧在码流中的位置的快速定位能力及 MPEG 码流编码的特点, 使演播工具具有快速、精确地访问显示节目流中的任意一帧的能力. 图2给出了一个按播放顺序排列的 GOP 中的各图, 这里不妨令每副图的图结构为帧. 在一个 MPEG 的压缩码流中, 帧数据的排列顺序与其播放次序是不一致的, 但与其解码次序却完全一致. 当我们仅想访问显示某个 GOP 中的 P 帧和 B 帧时, 必须先解码与之相关的一些帧, 但我们可以利用 MPEG 码流编码的特点和索引, 仅挑选对于目标帧的正确解码必不可少的那些帧的内容进行解码. 比如, 若要访问显示 B12(这里13及后面的数字序号均指播放次序), 仅需按如下顺序定位解码帧序列: I1→P4→P7→P10→P13→B12.

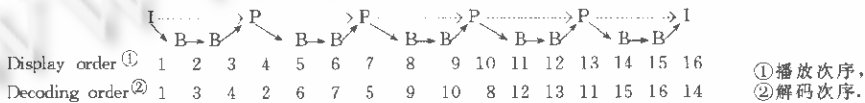


Fig. 2 Random access and display of a frame in a GOP

图2 对一个GOP中的帧的随机访问和显示

下面, 我们给出按照本文描述的 MPEG 节目流的索引模型随机访问显示码流中的任意帧的一般算法.

利用 MPEG 节目流的索引对码流中的任意帧随机访问显示算法的描述如下:

输入: ① MPEG-2或 MPEG-1的节目流, 或视频基本流文件 mpeg_fp;

② 对相应码流的索引文件 mpgIdx_fp;

③ 显示顺序的帧序号 DisFrmNum(第1帧的序号 DisFrmNum=0).

输出: 解码显示与帧序号 DisFrmNum 相对应的帧.

描述:

(1) 打开文件 mpeg_fp, mpgIdx_fp.

(2) 从 mpgIdx_fp 中提取相关的参数, 利用步骤(4)计算 nFrmsItem.

(3) 从 mpgIdx_fp 中提取相关的参数, 利用前文中的式(2)和式(3), 求出帧序号为 DisFrmNum 的帧所在 GOP 的 I 型图片子组对应的索引子在 I 中 D 元素中的位置 PosIdxFile.

(4) 移动 mpgIdx_fp 的文件指针到与 PosIdxfile 对应的地方, 并读取一个索引子大小的内容到 CurrentIdx-er.

(5) 利用前文中的式(5), 求得 RmFrms.

(6) 移动 mpeg_fp 的文件指针到 CurrenIdx-er.ipPos 处, 利用 MPEG 解码引擎解码一帧的数据.

(7) 若 RmFrms=0, 则转(16).

(8) LastIdx-er=CurrenIdx-er.

(9) 若 RmFrms<nFrmsItem, 则转(13).

(10) 从 mpgIdx_fp 中读取一个索引子大小的内容到 CurrentIdx-er.

(11) 移动 mpeg_fp 的文件指针到 CurrenIdx-er.ipPos 处, 利用 MPEG 解码引擎解码一帧的数据.

(12) 令 RmFrms=RmFrms-nFrmsItem, 转(8).

(13) 若 RmFrms=0, 则转(16).

(14) 从 mpgIdx_fp 中读取一个索引子大小的内容到 CurrentIdx-er, 移动 mpeg_fp 的文件指针到 CurrenIdx-er.ipPos 处, 利用 MPEG 解码引擎解码一帧的数据.

- (15) 移动 mpeg_ip 的文件指针到 LastIdxer.ipPos + LastIdxer.cif_Bs[RmFrms - 1]处,利用 MPEG 解码引擎解码一帧的数据.
- (16) 显示出最后解码的帧,结束.

3 性能实验与分析

我们实现了前文所述的算法,并设计了两组实验来测试通过 MPEG 索引以支持对 MPEG 流的帧内容的快速访问的性能.第1组实验测试在 MPEG 节目流中快速定位帧的速度性能,第2组实验测试在 MPEG 节目流中快速定位访问任意帧的速度性能,主要是测试随机显示任意帧的速度性能.我们选择了不同 GOP 结构的 MPEG-1和 MPEG-2的多个流来作测试.第1组实验的测试在配置分别为 PIII-450,64M 内存与奔腾133,内存为32M 的微机上进行;第2组实验考虑到奔腾133,内存为32M 的机器对 MPEG-2流解码的运算能力,因此仅在配置为 PIII-450,64M 内存上进行.表1给出了各个测试码流的结构特征信息以及为它们生成的索引文件的大小信息.

Table 1 Some feature parameters about the tested streams and their index file
表1 测试码流及其索引文件的一些特征参数

Stream file name ^①	Stream type ^②	Size of stream (KB) ^③	Frame dimension ^④	Structure of a GOP ^⑤	Size of index file (KB) ^⑥	Size ratio between index files and program streams ^⑦
Avseq01.dat	MPEG-1	47 762	352 * 288	IBBPBBPBB	25	0.000 52
Flight.mpg	MPEG-1	15 109	352 * 240	IBBPBBPBBPBBPBB	12	0.000 79
Dance.m2v	MPEG-2	31 237	752 * 576	IBBPBBPBBPBB	4	0.000 13
Dvdemo.m2v	MPEG-2	23 552	720 * 480	IBBPBBPBBPBB	4	0.000 17

①码流文件名,②码流类型,③码流大小(KB),④帧尺寸,⑤GOP 帧结构,⑥索引文件大小(KB),⑦索引文件与节目流大小比.

· 测试在 MPEG 节目流中快速定位帧的速度性能实验

在本组实验中,对任意一个待测码流,我们将随机生成待定位的帧的帧序号,并针对该帧序号进行如下的操作:打开索引文件,利用索引文件返回该帧内容在 MPEG 码流中的位置,关闭索引文件,统计若干次进行这样的操作后的时间.表2给出了实验的结果数据.

Table 2 Experimental result about speed of positioning frames in different MPEG streams
表2 对不同 MPEG 流帧定位速度的测试实验结果

Stream file name ^①	Operation times of random positioning frame ^②	PC with PIII 450 CPU and 64M memory ^③		PC with Pentium 133 and 32M memory ^④	
		Total response time (s) ^⑤	Frames positioned per second ^⑥ (Frames/s)	Total response time (s)	Frames positioned per second (Frames/s)
Avseq01.dat	40 000	7	5 714	99	404
Flight.mpg	80 000	15	5 333	169	473
Dance.m2v	40 000	7	5 714	84	476
Dvdemo.m2v	80 000	14	5 714	189	423

①码流名,②随机定位操作次数,③PIII-450,64M 内存的机器,④总响应时间(秒),⑤每秒定位帧数,⑥奔腾133,内存为32M 的机器.

实验结果表明,利用该索引模型具有无比快速的定位帧的速度,它要比文献[1]中的最快速度还快近4倍,并且它对帧的定位具有接近于一个常数的速度,与要定位的帧在 MPEG 流中的位置、流的码率等因素无关,仅与机器的硬件速度性能有关.这里,我们可以对上述特点进行分析.

算法通过索引文件确定视频帧在流文件中的位置主要涉及两部分的费用:一部分是待定位帧所在 GOP 的对应索引数据信息从硬盘(或其他介质)读入内存的时间费用 w ,另一部分是利用式(2)~(6)进行简单计算以及从索引数据块中抽取相关位置信息的时间费用 λ .由于目前 CPU 的处理速度与硬盘的读取速度属于不同的数量级,即 $w \gg \lambda$,从而总的费用 $\pi = w + \lambda \approx w$.另外,设操作系统一次读取磁盘块的大小为 μ (一般至少为几 K),所用的时间费用为 η ,在我们的实现中,一个

GOP 所对应的索引数据信息的长度为 ρ (一般至多几十字节), 因此访问任意一帧的时间费用最坏为 $\pi \approx w = 2\eta$, 这时一个 GOP 的对应索引数据信息将跨越两个磁盘块. 对于其他情况的时间费用为 $\pi = \eta$, 假设对各帧的访问是等概率的, 则访问跨越两个磁盘块的索引数据块的概率约为 $1 / \lceil \frac{\mu}{\rho} \rceil$, 从而算法的平均时间复杂度为

$$2\eta / \lceil \frac{\mu}{\rho} \rceil + \eta * \left(1 - 1 / \lceil \frac{\mu}{\rho} \rceil \right), \text{ 即 } \eta / \lceil \frac{\mu}{\rho} \rceil + \eta. \quad (7)$$

若 $\mu \gg \rho$, 则 $\eta / \lceil \frac{\mu}{\rho} \rceil + \eta$ 会趋近于 η . 通过上述分析得到的结果与表2中的实验结果是一致的.

· 测试在 MPEG 节目流中快速定位显示任意帧的速度性能

在本组实验中, 对任意一个待测码流, 我们将随机生成待显示帧的帧序号, 并依据节 2.2 中的算法利用索引文件中的信息, 进行快速地定位、解码与待显示帧相关的帧, 最后定位、解码、显示目标帧. 表3给出了一些实验的结果数据.

Table 3 Experimental result about fast positioning and displaying a frame in different MPEG streams on PC with PIII-450 CPU and 64M memory

表3 在 PIII-450、64M 内存的机器上对不同 MPEG 流快速定位显示任意帧的测试实验结果

Stream file name ^①	Capability of decoding engine used in experiment ^② (Frames/s)	Operation times of random access of frames ^③	Constraints of random access ^④	Total response time (s) ^⑤	Frames positioned and displayed per second ^⑥ (Frames/s)	Ratio between average positioning time and average decoding time ^⑦
Avseq01.dat	31.2	90	1 frame only ^⑧	3	30	0.005 25
			Any frame ^⑨	7	12.9	0.002 26
Flight.mpg	41.2	150	1 frame only	4	37.5	0.00703
			Any frame	13	11.5	0.002 01
Dance.m2v	6.5	120	1 frame only	17	7.1	0.00124
			Any frame	59	2	0.000 35
Dvdemo.m2v	7.8	60	1 frame only	7	8.6	0.001 51
			Any frame	25	2.4	0.000 42

①码流名, ②实验用的播放器的解码引擎性能 (Frames/s), ③随机访问操作次数, ④随机访问限制条件, ⑤总响应时间 (秒), ⑥每秒定位显示帧数, ⑦平均定位与解码的时间费用比, ⑧仅 I 帧, ⑨任意帧.

实验结果表明, 利用本索引模型来支持对任意帧的随机访问显示具有令人满意的效果. 当随机访问的对象仅限制于 I 帧时, 对 MPEG-1 流播放器具有逼近于解码器最大播放能力的性能, 对 MPEG-2 流播放器甚至具有更高的性能. 当随机访问的对象为任意帧时, 与随机访问的对象仅为 I 帧相比, 性能有明显的下降, 这与码流的 GOP 结构、访问帧的位置有关. 例如, 若要访问显示 Flight.mpg 的任意一个 GOP 中的最后一帧, 就要解码 7 帧的数据, 但对 I 帧仅需解码 1 帧的数据. 在整个过程中, 解码的费用要远远高于其他费用 (定位、显示) 的. 最后需要指出的是, 本实验采用的解码引擎的效率是较低的, 在相同的环境下, 一些著名商业软件能够提供相当于我们所用的解码引擎平均 4 倍多的性能, 可以想像在高性能解码引擎的支持下, 本索引模型对任意帧的随机访问能力的支持将会有进一步的质的飞跃.

4 结束语

本文提出并讨论了一种 MPEG-2 节目流的索引模型, 该模型可推广至 MPEG-2 传输流. 本文给出了基于该模型的 MPEG-2 索引文件生成算法及利用该索引对码流中的任意帧随机访问显示算法. 实验表明, 利用索引来访问 MPEG 码流中的任意帧具有相当快速的系统响应时间, 可以更好地

适用于一些并发服务于多个客户的采用 MPEG 格式作为视频源的多媒体服务器端应用系统,如 VOD、大型视频数据库等,也同样适用于对内容浏览定位具有较高精度要求的多媒体应用,如面向 MPEG 流的非线性编辑系统。索引文件虽然具有一定的存储空间的耗费,但其在空间上的耗用量与 MPEG 流文件的空间耗用量相比是微乎其微的,仅为其万分之几,而且 MPEG 流的码率越高,索引文件所占空间的相对比例越小。目前,单位容量存储介质的价格在不断地降低。所以,我们认为建立并利用 MPEG 流的索引是一种提高某些高级多媒体应用系统性能和服务质量的有效途径。虽然建立索引的过程要有一定的时间开销,开销的大小为系统从硬盘上读取扫描一遍流文件花费的时间,但由于一旦建立了索引,便可以一劳永逸地多次使用,因此为视频流建立索引适用于需频繁访问特定帧的系统。

References:

- [1] Chen, Wei-qiang, Gao, Wen. A method of fast searching I pictures in MPEG-2 system stream. *Journal of Software*, 1998, 9(7):520~524 (in Chinese).
- [2] Zhong, Yu-zhuo, Qiao, Bing-xin, Qi, Wei. *International Standard ISO/IEC 13818-Information Technology Generic Coding of Moving Pictures and Associated Audic.* Beijing: Tsinghua University Press, 1997 (in Chinese).
- [3] Yu, Chong-zhi, Li, Cun-zhu, Zhang, Zhi-li, *et al.* *Compression and Coding of Moving Pictures and Associated Audio-MPEG1.* Nanjing: Nanjing University Press, 1995 (in Chinese).

附中文参考文献:

- [1] 陈维强,高文.一种在 MPEG-2 系统流上快速搜索 I 图像的方法. *软件学报*, 1998, 9(7):520~524.
- [2] 钟玉琢,乔秉新,祁卫. *运动图像及其伴音通用编码国际标准-MPEG-2(ISO/IEC 13818)*. 北京:清华大学出版社,1997.
- [3] 余崇志,李存珠,张志力,等. *运动图像及其伴音的压缩与编码-MPEG1*. 南京:南京大学出版社,1995.

An Index Model for MPEG-2 Streams and Its Application*

WANG Wei-qiang¹, GAO Wen^{1,2}

¹*Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100080, China;*

²*Department of Computer Science and Engineering, Harbin Institute of Technology, Harbin 150001, China)*

E-mail: wqwang@ict.ac.cn; wgao@ict.ac.cn

http://www.ict.ac.cn

Abstract: MPEG standards are playing a more and more important role in many multimedia applications now, and will continue in the future. An effective index model for MPEG-1,2 streams is presented in this paper. Moreover, an algorithm to construct the index for MPEG-1,2 streams is designed, as well as a solution to access any frame in the streams rapidly based on the model. The experimental results have shown the index model can support the ability of positioning a frame in the stream quickly, and it can be applied to access any frame randomly effectively.

Key words: MPEG (moving picture expert group); index; program stream; transport stream, GOP (group of picture)

* Received Decmber 29, 1999; accepted April 18, 2000

Supported by the National Natural Science Foundation of China under Grant No. 69789301; the National High Technology Development Program of China under Grant Nos. 863-306-ZT03-01-2, 863-306-ZD11-03; the 100 Outstanding Scientists Foundation of the Chinese Academy of Sciences