

多媒体系统中的媒体同步问题研究^{*})

A Study of Media Synchronization

陈杰军 鲁东明 潘云鹤

(浙江大学人工智能研究所 杭州310027)

摘 要 Media synchronization is a very difficult problem in multimedia system, it is a kernel technology of multimedia technologies. This article begins with the basic notion of media synchronization, and probes into the key problem of media synchronization such as synchronization model and synchronization specification

关键词 Media synchronization, Synchronization model, Synchronization specification

一、引言

随着当今网络计算与存储领域的发展,多媒体系统将为人们提供广阔的应用前景,如多媒体娱乐、多媒体会议、多媒体演示等。

区分一个系统是否为多媒体系统有三条准则:(1)媒体的数量。(2)所支持的媒体类型。(3)媒体间集成的程度。最简单的一条准则是媒体的数量,如果按照这条准则,一个能处理文本与图形的普通编辑器也可看作是一个多媒体系统。媒体的类型可按与时间的相关性来划分为时间相关媒体与时间无关媒体。时间相关媒体是指一种媒体流,由一系列连续的媒体单元组成。如视频媒体对象是由一连串视频帧组成。时间相关媒体所包含的语义需要一段时间来进行展示。时间无关媒体是指一些传统的媒体如文本、图象等。这些媒体对象所包含的语义的展示与时间域无关。对时间相关媒体的处理是一个多媒体系统必不可少的条件。媒体的集成程度是第三条准则。集成意味着不同类型的媒体保持独立,但是能一起处理与展示。综合这三条准则,一个多媒体系统可看作是:能集成处理多种类型的媒体类型,并且至少包含一种时间相关媒体^[1,2]。

正是由于多媒体系统引入了时间相关媒体类型,以及由于多种媒体集成所带来的媒体间的时间关系的存在,导致了多媒体系统中的媒体同步这一核心技术的存在。媒体对象的时间相关性可能隐藏在建立过程中,如视频媒体,其内部视频帧之间的时

间关系在视频采集时就已定义好。媒体对象的时间相关性也可能由用户定义,如多媒体演示中指定各种媒体的播放时序。不管是哪种情况,这些时间相关性必须在多媒体系统中得到体现与支持。另外,一个多媒体系统必须能够克服在数据演示过程中的存储、通信和计算所引起的系统延迟,所有这些都需要媒体同步技术加以解决。本文将从媒体同步技术的基本概念出发,较为详细地论述媒体同步技术中的同步模型、同步规范等关键问题。

二、媒体同步的概念

2.1 媒体同步关系

多种媒体集成处理是多媒体系统的一个重要特征。多种媒体之所以要集成处理,是由于多种媒体组成一个信息对象造成媒体之间具有一种相互依赖关系。这种依赖关系必须在媒体集成处理当中表现出来。媒体集成处理包括存储、操作、通信特别是展示。一般来讲,媒体间的相互依赖关系包括以下三种:

(1)内容关系 定义了媒体对象一些数据间的依赖关系。例如,一张反映人口密度的表与人口密度分布图。这两个媒体对象都依赖于同一数据:人口统计数据。在多媒体系统中,内容关系的实现一般都是基于共同的数据结构。

(2)空间关系 定义了各种媒体对象某一时刻在输出设备上的一种位置关系。一个典型的例子是幻灯片。幻灯片中一系列文本、图象媒体对象按一定的位置排列起来表现了一定的语义。

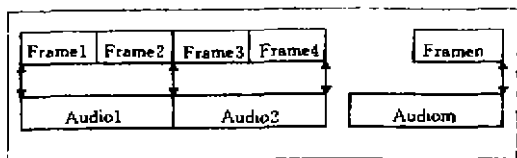
^{*})本项目得到浙江省自然科学基金和九五国家科技项目的资助。

(3)时序关系 定义了媒体对象一种时间上的依赖关系。一个典型的例子是记录一个人说话时所获得的音频对象与视频对象。当展示这些对象时,这两种媒体对象之间的时序关系必须与记录时相一致。

所有这三种同步关系,在多媒体系统都需要加以处理,但其中媒体的时序关系是关键因素,它所带来的同步需求是媒体同步问题中主要要处理的。

2.2 逻辑数据单元(LDU)

逻辑数据单元是时间相关媒体流中的一个信息单元。在视频流媒体中,一个视频帧可看作是一个逻辑数据单元。音频也可按照采样频率分成一个个逻辑数据单元。把时间相关媒体细分成逻辑数据单元有利于简化时间相关媒体的同步处理。如下图就是用 LDU 来表示一段音频与一段视频之间的同步关系,一个音频 LDU 的展示时间与两个视频 LDU 的展示时间相同。

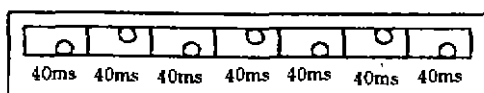


三、媒体同步的分类

3.1 媒体内同步与媒体间同步

媒体的时序关系有两种。一种是时间相关媒体内部的 LDU 之间的一种时序关系,如视频媒体内部的视频帧序列。一种是不问媒体之间的时序关系。

□媒体内部同步 指的就是时间相关媒体内部各 LDU 之间的一种同步。如一段每秒 25 帧的视频媒体,展示时每帧的显示时间应为 40ms,然后才能显示下一帧。如下图表示一个球跳动的一段视频。



□媒体间同步 处理不同媒体之间的一种时序上的关系。如图1展示了一个多媒体演示时的媒体间同步关系。

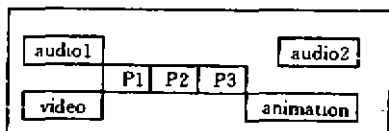


图1

该图表示首先展示一段音/视频流,紧接着顺序展示三张图片,最后展示了一段动画,展示动画的同时有一段伴音进行讲解。

3.2 实况同步与综合同步

1)实况同步 目标是要在演示时再现媒体获取时的时序关系。一个典型的例子是:两个在不同地点的人使用一个协同讨论系统讨论一个新产品,该协同讨论环境由一个视频会议模块与一个共享的白板组成。两个人一边用共享光标指点着产品,一边发表意见。这时就有两个同步问题。一个是口型同步,一个人看到另一个人的视频一边听他说,口型与声音应该同步在用户感觉不出的范围内。另一个同步问题是共享光标与用户的语言之间的同步。当一个用户看到另一个用户用光标指着产品发表意见时,他看到的共享光标的移动应与听到的声音同步起来。这两种同步都属于实况同步。实况同步就是要在媒体展示时真实再现媒体获取时的状态^[6]。

2)综合同步 媒体间时序关系是由人指定的。综合同步的重点是支持媒体间灵活的时序关系。综合同步有两个阶段,一是同步规范制定阶段,该阶段明确指定媒体间的时序关系。二是展示阶段,系统按规定好的同步规范展示媒体数据。综合同步经常出现在多媒体演示系统中。

四、媒体同步模型

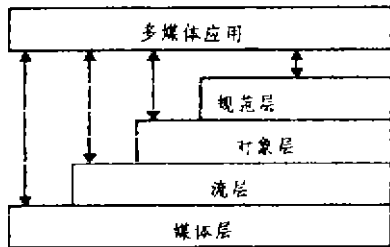
媒体同步模型用来从整体上描述多媒体同步,以利于媒体同步处理。其主要目的是用于表征多种媒体之间的时序关系。目前较为流行的多媒体同步模型有三种:参考点同步模型、时间轴同步模型、层次同步模型^[3,4]。

参考点同步模型将连续媒体看作是一系列离散子单元构成的序列,一个子单元所在位置称为一个参考点。不同媒体之间的同步是通过将其在某一时刻的子单元相连接来定义的。参考点同步模型的缺点是将连续媒体离散化,从而破坏了连续媒体之间的相互依赖关系,破坏了连续媒体的整体特性。

时间轴同步模型中,所有媒体对象都是彼此独立地依赖于同一时间轴。这样,一个对象的丢失或删除不会影响其他对象的同步。这种模型对于那些起始时间点固定的多媒体对象而言,无疑非常合适。时间轴同步模型应用比较广泛,象现在一些非线性视频编辑系统、多媒体演示编辑器大都是基于时间轴的。

层次同步模型的特点是对多媒体同步分层考

虑,一个典型的层次媒体同步模型如下所示:



媒体层:在媒体层,多媒体应用操作单个的连续媒体流。该媒体流被当成一个 LDU 序列,该层提供的接口是一些设备无关的操作如:read(devicehandle, LDU) write(devicehandle, LDU)等。应用程序如果使用媒体层,则应用程序自己将使用流量控制机制来进行媒体内部同步。

流层:流层操作一组连续媒体流,在媒体流组中,所有流都并行处理,以解决媒体间的同步。在流层提供的典型接口有 start(stream), stop(stream), create_group(list-of-streams), start(group), stop(group)等。应用程序使用流层来启动媒体流,终止媒体流,对媒体流进行分组等。

对象层:对象层操作所有类型的媒体,掩盖了时间相关媒体与时间无关媒体之间的差异。这一层主要完成媒体对象的同步展示。对象层的输入为同步规范,负责整个展示的调度。对象层并不处理媒体流间的同步与媒体流内部的同步,这些功能由对象层调用流层的功能来实现。

规范层:规范层为一开放层,它没有提供明确的接口。这一层包括创建同步规范的应用程序与工具。例如同步编辑器、多媒体文档编辑器等。制定同步规范的方法也很多,它也是多媒体同步研究的一个重点。关于同步规范的制定本文将在下一节加以论述。

层次同步模型中对多媒体同步进行了分层处理,这样有利于得到一个比较完善、较为通用的多媒体同步模型。

五、媒体同步规范

多媒体对象的同步规范描述了所有媒体对象的时间上的依赖关系。在层次同步模型中,同步规范的产生是利用规范层中的工具来得到的。同步规范决定了多媒体对象的展示,它是多媒体系统中的一个中心问题。一个多媒体的同步规范应该包括以下内

容:

- ★媒体对象的对象内部同步规范。
- ★对象内部同步的服务质量描述。
- ★媒体对象的媒体间同步规范。
- ★媒体对象间同步的服务质量描述。

同步规范是多媒体对象的一部分属性,同步规范制定的方法应满足以下要求:

★方法应支持对象的一致性与同步规范的维护。

★方法应能抽象出媒体对象的内容。

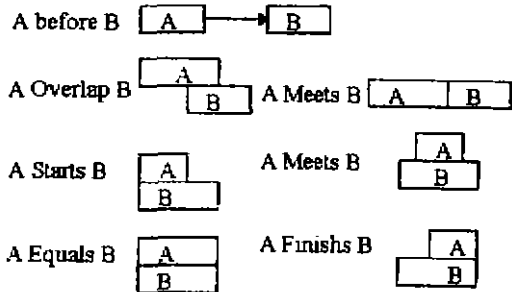
★所有类型的同步关系应能简单地描述出来。

★必须支持时间相关媒体对象与时间无关媒体对象的集成。

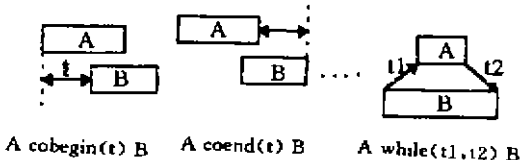
★必须能够定义同步服务质量。

目前,同步规范的描述方法已提出许多,主要分为以下几类:

1) 基于时间片的同步规范^[4] 在此规范中,时间片指的是一个对象的展示时间。两个时间片对象之间的时序关系一般有以下七种:



针对这些时间片之间的关系,我们可以定义一系列动作,如:



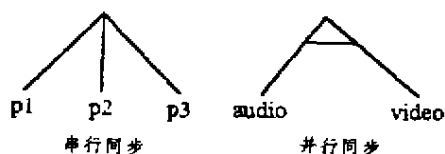
应用这些动作,前面图1中的同步关系可用如下规范描述出来:

```
audiol while(0,0) video
audiol before(0) p1
p1 before(0) p2
p2 before(0) p3
p3 before(0) animation
audio2 coend(t) animation
```

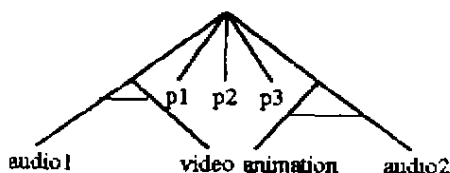
基于时间片的同步规范的优点是容易集成时间无关媒体对象,能很好地抽象出媒体内容。但基于时间片的同步规范比较复杂,它直接定义媒体间的时间关系,对媒体对象的 LDU 之间的时间关系则不定义。

2) 层次同步规范^[5] 层次同步规范中有一个动

作(action)的概念。同步规范就是描述这些动作之间的时间关系。动作表示发生的事情,事件(event)是动作的最小单位。一个动作就是一系列事件的复合体。如果一个动作的某些事件与别的动作的某些事件有对应关系,则这些事件称为同步点。一个动作有启动点(startpoint)与结束点(endpoint)。一个动作分以下两种:(1)原子动作,在启动点与结束点之间不再有同步点的动作。(2)复合动作,由多个原子动作组成。为了同步,一个动作需分成原子动作,通过适当的同步机制使它们在启动点与结束点之间同步。动作间的同步可分为两种:(1)串行同步;(2)并行同步,用图可表示如下:

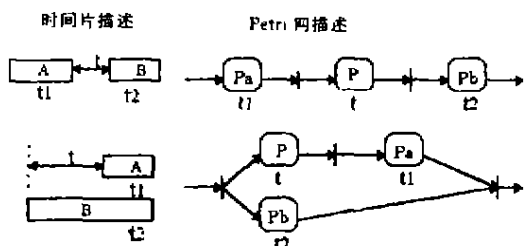


应用层次同步规范,图1中的同步关系可表示如下:



层次同步规范的优点是简单明了,缺点则是必须额外指定时间无关媒体的展示时间,为同步需要必须对媒体对象进行切分。

3)基于 Petri 网的同步规范^[7] Petri 网是一带权有向图,在描述同步规范时,Petri 网中的权可用来表示媒体对象的持续时间。下面是时间片描述与 Petri 网描述的两个对应关系:



在 Petri 网同步规范中,两个媒体对象的起始点或结束点之间的时间间隔都用增加一个时间持续节点的方式来表示,如上图中的 P 节点。Petri 网是一种比较好的同步描述方法,它适合于所有的同步情况,它的主要缺点是同步规范比较复杂。

描述同步规范的方法还有许多,如基于参考点的同步规范,基于事件的同步规范等。这些同步规范都各有优缺点。象同步模型一样,同步规范描述方法也是媒体同步中的一个主要研究内容。

参考文献

- [1] Gerold Blakowski and Ralf Steinmetz, A Media Synchronization Survey: Reference Model, Specifications, and Case Studies, IEEE J. on Select Areas in Commun., 14(1)1996
- [2] R. Steinmetz, Synchronization properties in multimedia systems, IEEE J. Select Areas Commun. Vol8, Apr. 1990
- [3] Chun-Chuan Yang and Jau-Hsiung Huang, A Multimedia Synchronization Model and Its Implementation in Transport Protocols, Same to [1]
- [4] T. D. C. Little and A. Ghafoor, Synchronization and storage models for multimedia objects, Same to [2]
- [5] D. Shepherd and M. Salmony, Extending OSI to support synchronization required by multimedia applications, Computer Commun., Vol. 13, Sept. 1990
- [6] Human perception of jitter and media synchronization, IEEE J. Select. Areas Commu. pp61-72
- [7] Spatio-temporal composition of distributed multimedia objects for value added networks, IEEE Computer, Vol. 24, Oct. 1991
- [8] 陈俊华、杨品、巴林凤,多媒体二级同步模型,计算机研究与发展,33(5)1996